

**Data Analysis**  
of  
**the Cosmic Microwave Background.**

Dissertation der Fakultät für Physik

der

Ludwig-Maximilians-Universität München

vorgelegt von: **Frode Kristian Hansen**

aus Tønsberg

München, den 11.01.2002



**Data Analysis**  
of  
**the Cosmic Microwave Background.**

Dissertation der Fakultät für Physik

der

Ludwig-Maximilians-Universität München

vorgelegt von: **Frode Kristian Hansen**

aus Tønsberg

München, den 11.01.2002

1.Gutachter: Prof. S. White

2.Gutachter: Prof. A. Schenzle

Tag der Mündlichen Prüfung: 12.06.02

# Contents

Einleitung und Zusammenfassung . . . . .	1
Introduction and Summary . . . . .	5
<b>1 The Physics of the CMB fluctuation</b>	<b>9</b>
1.1 Basic Concepts . . . . .	10
1.1.1 The Friedmann Equations . . . . .	10
1.1.2 Problems in Standard Cosmology . . . . .	14
1.2 Theories of the Early Universe . . . . .	16
1.2.1 Inflation . . . . .	16
1.2.2 Cosmic Strings . . . . .	19
1.3 The Growth of Perturbations in the Early Universe . . . . .	20
1.3.1 Linear Evolution . . . . .	20
1.3.2 Nonlinear Evolution . . . . .	25
1.3.3 The Matter Power Spectrum . . . . .	26
1.4 The Recombination Era and the CMB . . . . .	28
1.4.1 The Origin of the CMB Anisotropies . . . . .	29
1.4.2 The CMB Power Spectrum . . . . .	32
1.4.3 Reionization and Secondary Anisotropies . . . . .	38
1.4.4 Polarisation of the CMB and Tensor Perturbations . . . . .	39
<b>2 The Analysis of CMB Data</b>	<b>45</b>
2.1 CMB Experiments, the Past, Present and Future . . . . .	45
2.1.1 COBE . . . . .	46
2.1.2 BOOMERANG . . . . .	46
2.1.3 MAP . . . . .	47
2.1.4 Planck . . . . .	49
2.1.5 Other Experiments . . . . .	52
2.2 The Analysis of CMB Data Sets . . . . .	53
2.2.1 Map Making . . . . .	53
2.2.2 Foregrounds . . . . .	58
2.3 Power Spectrum Estimation . . . . .	62
2.3.1 Likelihood Estimation . . . . .	63
2.3.2 Quadratic Estimators . . . . .	68
2.3.3 Two New Methods for Power Spectrum Estimation . . . . .	70

<b>3</b>	<b>Fast Exact Power Spectrum Analysis for a Special Type of Scanning Strategies</b>	<b>71</b>
3.1	Fast Fourier Space Convolution . . . . .	72
3.2	Power Spectrum Estimation Using Scanning Rings . . . . .	76
3.2.1	Theory . . . . .	80
3.2.2	An Example . . . . .	86
3.2.3	Discussion . . . . .	91
<b>4</b>	<b>Gabor Transforms on the Sphere and Application to CMB Analysis</b>	<b>95</b>
4.1	The Gabor Transformation and the Temperature Power Spectrum	97
4.1.1	The One Dimensional Gabor Transform . . . . .	97
4.1.2	Gabor Transform on the Sphere . . . . .	98
4.1.3	Rotational Invariance . . . . .	108
4.2	Likelihood Analysis . . . . .	110
4.2.1	The Form of the Likelihood Function . . . . .	110
4.2.2	The Correlation Matrix . . . . .	114
4.2.3	Including White Noise . . . . .	116
4.2.4	Likelihood Estimation and Results . . . . .	119
4.3	Extensions of the Method . . . . .	134
4.3.1	Multiple Patches . . . . .	135
4.3.2	Monte Carlo Simulations of the Noise Correlations and Extension to Correlated Noise . . . . .	144
4.4	Discussion . . . . .	147
<b>5</b>	<b>Gabor Transform on the Polarised CMB Sky</b>	<b>149</b>
5.1	The Gabor Transformation . . . . .	150
5.1.1	Polarisation Powerspectra . . . . .	150
5.1.2	Rotational Invariance . . . . .	170
5.2	Likelihood Analysis . . . . .	171
5.2.1	The Form of the Likelihood Function for Polarisation . . . . .	171
5.2.2	The Polarisation Correlation Matrix . . . . .	178
5.2.3	Polarisation with Noise . . . . .	182
5.3	Results of Likelihood Estimations . . . . .	186
5.4	Discussion . . . . .	187
<b>A</b>	<b>Rotation Matrices</b>	<b>191</b>
<b>B</b>	<b>Spin-s Harmonics</b>	<b>192</b>
<b>C</b>	<b>Some Wigner Symbol Relations</b>	<b>193</b>
<b>D</b>	<b>Recurrence Relation</b>	<b>194</b>

<b>E</b>	<b>Extention of the Recurrence Relation to Polarisation</b>	<b>197</b>
	Acknowledgements . . . . .	205
	Lebenslauf . . . . .	206





## Einleitung und Zusammenfassung

Das frühe Universum bestand aus einem heißen, dichten und ionisierten Gas aus Elektronen, Protonen, Neutronen, einigen leichten Atomkernen und Photonen. Das Universum war in dieser Zeit optisch dicht da aufgrund der häufigen Zusammenstöße mit Elektronen in dem dichten Gas, die Photonen sich nicht sehr weit bewegen konnten. Die Temperatur war so hoch, dass die Elektronen, Protonen und Neutronen nicht zu Atomen rekombinieren konnten. Als das Universum expandierte, kühlte es sich ab. Ungefähr 300 000 Jahre nach dem Urknall betrug die Temperatur des Gases etwa 3000 Grad Kelvin. Unterhalb dieser Temperatur ist die Bildung der ersten Atome des Universums möglich. Die Elektronen und Protonen rekombinierten zu Atomen. Im Allgemeinen ist die Wahrscheinlichkeit dafür, dass ein Photon mit einem Atom zusammenstößt, viel geringer als die Wahrscheinlichkeit dafür, dass es mit freien Elektronen oder Protonen kollidiert. Aus diesem Grund sagt man, dass das Universum durchsichtig wurde, als die Elektronen mit den Protonen rekombinierten. Die Photonen bewegten sich entlang einer Geraden ohne dass sie abgelenkt wurden. Ungefähr 12 Milliarden Jahre später trafen die Photonen einen Detektor auf einem Planeten namens 'die Erde' und gaben den Wissenschaftlern wertvolle Informationen über den Anfang des Universums. Diese Strahlung, die auf ihrer Reise von den frühesten Zeiten bis heute kaum eine Veränderung erfahren hat, wird 'der kosmische Mikrowellenhintergrund' genannt, und ist das Thema dieser Doktorarbeit.

Das dichte Gas des frühen Universums war nicht ganz homogen. Prozesse der ersten  $10^{-34}$  Sekunden nach dem Urknall verursachten kleine Schwankungen in der Dichte. Die Entwicklung dieser Schwankungen unter dem Einfluss von Gravitation und Druck kann man mit Hilfe der Hydrodynamik einfach berechnen. Nachdem das Universum durchsichtig wurde (Rekombination), entwickelten sich die Schwankungen und sind heute als Sterne, Galaxien und Galaxienhaufen sichtbar. Die Bildung dieser Objekte ist eine sehr komplizierte Prozess, die noch nicht vollständig verstanden ist. Die Entwicklung der Strukturen des Universums bevor Rekombination kann man mit linearer Physik beschreiben. Hätte man die Möglichkeit, das Universum in diesem früheren Stadium zu beobachten, hätte man Informationen über die Eigenschaften und Anfang des Universums auf eine einfache Weise erhalten können. Der kosmische Mikrowellenhintergrund (CMB) gibt eine solche Möglichkeit. Diese Strahlung ist vom Zeitpunkt der Rekombination bis heute nahezu ohne Störung gereist und ist somit ein Bild vom Universum und der Verteilung der Dichteschwankungen zu diesem Zeitpunkt.

Der CMB wurde zuerst von Penzias und Wilson 1965 als isotrope Schwarzkörper-

strahlung mit einer Temperatur von ca. 2.7 Grad Kelvin entdeckt. Wie oben erwähnt war die Temperatur des Gases zum Zeitpunkt der Rekombination ca. 3000 Grad Kelvin, da das Universum aber seitdem ungefähr 1000 mal größer geworden ist, ist die Temperatur dieser Schwarzkörperstrahlung auch genau um diesen Faktor kleiner geworden. Die Schwankungen der Dichte des Universums zur Zeit der Rekombination verursachten kleine Unregelmäßigkeiten in der Temperatur des CMB, die zuerst 1992 mit dem Cosmic Background Explorer (COBE) Satellit entdeckt wurden. In den letzten Jahren sind mehrere Experimente mit immer höherer räumlicher Auflösung von der Erde und von Ballons aus durchgeführt worden. Diese Experimente haben schon wertvolle Informationen über das frühe Universum gegeben. Ein anderer Satellit namens MAP führt jetzt Beobachtungen des CMB durch und der Planck Satellit, der in einigen Jahren ins All geschickt wird, wird Beobachtungen mit noch höherer räumlicher und spektraler Auflösung durchführen.

Da die Experimente immer höhere Auflösung erreichen und immer größere Teile des Himmels beobachten, wird es auch immer schwieriger, die Daten zu analysieren. Die Analyse der Daten jetziger Experimente ist bereits ein Problem und die Menge der Daten von MAP und Planck werden um ein Vielfaches größer sein. Die Standardmethode, um Informationen (kosmologische Parameter) aus CMB Daten zu extrahieren ist die Maximum-Likelihood Methode. Diese Methode fordert die Invertierung einer  $N \times N$  Matrix, wobei  $N$  die Anzahl der Pixel eines Experimentes ist. Diese Invertierung erfordert  $N^3$  Berechnungen. Für das Planck Experiment wird  $N$  einen Wert von mehr als zehn Millionen haben. Die Invertierung einer Matrix dieser Größe dauert mehrere Tausend Jahre mit den schnellsten Supercomputern. Daher muss man andere Methoden finden um die Daten analysieren zu können. In dieser Doktorarbeit werde ich zwei neue Methoden präsentieren, die es ermöglichen, CMB Daten schnell zu analysieren. Mit diesen Methoden wird man das CMB Leistungsspektrum -die sphärisch harmonischen Koeffizienten der CMB Temperatur Daten- schnell extrahieren können. Mit Hilfe des Leistungsspektrums kann man die kosmologischen Parameter einfach berechnen.

In Kapitel (1) wird die Physik des frühen Universums beschrieben. Eine Zusammenfassung der Entstehung und der Entwicklung der Strukturen des Universums vom Urknall bis heute wird präsentiert. Die physikalischen Prozesse, die die Hintergrundstrahlung verursacht haben, werden zusammengefasst, und es wird gezeigt, wie diese Prozesse eine Abhängigkeit des Leistungsspektrums von den kosmologischen Parametern verursachen. In Kapitel (2) wird einen Überblick über vergangenen, heutigen und zukünftigen CMB Experimente gegeben, und die Standardmethoden für Analyse der CMB Daten werden erläutert. In diesem Kapitel wird auch beschrieben, wie man eine CMB Karte von den 'Zeitgeordneten Daten' (TOD) eines CMB Experimentes macht und wie man Störungen durch

die Mikrowellenstrahlung anderer Körper entfernen kann. Zum Schluss des Kapitels werden die neusten Methoden für die Extraktion des Leistungsspektrums aus CMB Daten beschrieben.

In Kapitel (3) wird beschrieben, wie man mit einer Maximum-Likelihood-Methode das Leistungsspektrum direkt von den TOD extrahieren kann, ohne zuerst eine CMB Karte zu erstellen. Diese Methode ist auf Experimente wie MAP und Planck anwendbar die den Himmel ‘in Ringen von Ringen’ vermessen. Symmetrien dieser Abtastmethode machen die Korrelationsmatrizen für Signal und Störung blockdiagonal im Fourierraum. Aus diesem Grund kann man die Korrelationsmatrix mit  $N^2$  Berechnungen statt  $N^3$  invertieren. Das ermöglicht die Verwendung der Maximum-Likelihood Methode für solche Experimente. Weil wirkliche Experimente, Schwankungen von der hier angenommenen idealen Abtastmethode haben werden, werden Erweiterungen dieser Methode für solche Experimente diskutiert.

In Kapitel (4) wird eine andere Methode mit der man das Leistungsspektrum von CMB Daten schnell extrahieren kann, vorgestellt. In diesem Kapitel wird auch erläutert wie man ‘Windowed Fourier Transforms’ (auch als ‘Gabor Transforms’ bekannt) vom Eindimensionalen auf die 2-dimensionale Kugeloberfläche erweitert. Es wird gezeigt, wie man das Leistungsspektrum von einem Ausschnitt des Himmels (Pseudo-Leistungsspektrum) als Eingangsdaten für eine Maximum-Likelihood Analyse verwenden kann. Die Korrelationsmatrix der Maximum-Likelihood Methode ist in diesem Fall viel kleiner als wenn alle Pixeln eines Experimentes benutzt werden. Aus diesem Grund ist die Invertierung der Matrix schnell. Es wird gezeigt, dass die Annahme einer Gaußschen Verteilung der Pseudo-Leistungsspektrumskoeffizienten eine gute Annäherung ist. Es wird auch gezeigt, wie man CMB Daten vom ganzen Himmel analysieren kann, wenn man den Himmel in viele kleinere Ausschnitte aufteilt und alle Ausschnitte gleichzeitig analysiert. In Kapitel (5) wird die Methode für Analyse der CMB Polarisation Daten erweitert.



## Introduction and Summary

The early universe consisted of a dense, hot and ionised gas of electrons, protons, neutrons, some light atomic nuclei and photons. The universe at that time was optically thick as the photons could not travel very far before being scattered on electrons in the dense gas. The temperature was too hot for the electrons to combine with the nuclei and form atoms. But as the universe expanded it cooled. About 300 000 years after the Big Bang, the temperature of the dense gas filling the universe was about 3000 degree Kelvin. This temperature allowed the formation of the first atoms in the universe. The electrons combined with the protons to form atoms. The probability for a photon to scatter on a neutral atom is much less than the probability to scatter on free electrons and protons. For this reason it is said that the universe got transparent when the electrons were bound to the protons. The photons continued travelling in a straight line without being scattered. About 12 billion years later some of these photons hit a detector on the planet called ‘the earth’. And it provided scientists with valuable information about the origin of the universe. This radiation which has travelled more or less unchanged from the earliest times until today is called *the cosmic microwave background radiation*, and is the topic of this Ph.D. thesis.

The dense gas in the early universe was not completely homogeneous. Processes during the first  $10^{-34}$  seconds after the Big Bang created small density fluctuations. The evolution of these fluctuations under the influence of gravitation and pressure in the dense gas can be easily calculated using hydrodynamics. After the universe became transparent (known as ‘the recombination era’), these small density fluctuations evolved to be the structure that can be seen in the present universe, stars, galaxies and clusters of galaxies. The formation of these objects is physically very complicated and not completely understood. If one had the possibility to observe the universe at an earlier stage when structure formation was still governed by known linear physics, it would be easier to gain information about the properties of the universe and its origin. The cosmic microwave background (CMB) radiation provides such a possibility. The radiation has travelled from the recombination epoch until today and gives a picture of the universe and the distribution of density fluctuations at the recombination epoch.

The CMB was first detected by Penzias and Wilson in 1965 as a black body radiation with a temperature of 2.7 Kelvin coming from all directions. As explained above the temperature of the gas at recombination was about 3000 Kelvin, but since that time the universe has expanded by a factor of roughly 1000 which has decreased the temperature of the CMB photons by a similar factor. The density

fluctuations in the universe at the recombination epoch gave rise to small fluctuations in the CMB temperature. These fluctuations were first detected by the Cosmic Background Explorer (COBE) satellite in 1992. In the recent years several ground and balloon borne experiments have made observations of the CMB at increasing angular resolution. These observations have already provided valuable information about the early universe. Another satellite called Microwave Anisotropy Probe (MAP) is currently observing the CMB and a satellite called Planck will make observations with an even higher angular and spectral resolution in a few years. As the experiment are getting higher resolution and larger sky coverage the task of analysing the data is gradually getting harder. Already the analysis of the present datasets have presented a challenge and the datasets from MAP and Planck will be several times larger. The standard method of extracting information in terms of cosmological parameters from the CMB data has been the method of maximum-likelihood. Unfortunately this method requires the inversion of a  $N \times N$  correlation matrix where  $N$  is the number of pixels in the experiment. This takes of the order  $N^3$  operations. For Planck  $N$  will be of the order of tens of millions and the inversion of the corresponding matrix takes thousands of years on the fastest supercomputers. Other methods for analysing the data have to be found. The aim of this thesis is to present two new methods to analyse CMB temperature and polarisation data. The methods described in the thesis aim at extracting the angular power spectrum of the CMB which is the spherical harmonic transform of the CMB sky. From this power spectrum the cosmological parameters can be extracted.

In chapter (1) a description of the physics of the early universe is given. A summary of the formation and evolution of structure from the Big Bang until the present era is presented. The processes giving rise to the CMB is also outlined and it is shown how these processes give a dependency on cosmological parameters in the angular power spectrum of the CMB. Then in chapter (2) the past, present and future CMB experiments are reviewed and the standard methods of data analysis are presented. The chapter also contains a description of the process of making a CMB sky map from the time ordered data (TOD) of a CMB experiment and the removing of foreground contamination. Finally the last methods of extracting the angular power spectrum are reviewed.

In chapter (3) a new method is presented for extracting the CMB power spectrum using the maximum-likelihood method directly on the TOD instead of making a CMB map first. The method is applicable to experiments scanning on rings of rings on the sky like the MAP and Planck experiments. Symmetries of this scanning strategy make the signal and noise correlation matrix for the likelihood calculation block diagonal in Fourier space. For this reason the matrix can be inverted in  $N^2$  operations instead of  $N^3$ . This makes the maximum-likelihood method feasible for this kind of experiments. The method is exact and naturally

takes into account arbitrary beam shapes and side lobes which other existing methods can not deal with. As realistic experiments can have deviation from the ideal scanning strategy assumed, extensions of the method to these cases is discussed.

In chapter (4) a different approach of estimating the angular power spectrum of the CMB is shown. In the chapter it will be explained how the windowed Fourier transforms (known as Gabor transforms) can be extended from the line to the sphere. It will be shown how one can take a patch on the CMB sky and use the power spectrum of this patch (the so called pseudo power spectrum) as input to a maximum-likelihood procedure. The correlation matrix for this likelihood is much smaller than when using all the pixels of the experiment and matrix inversion is very fast. The assumption is that the coefficients of the pseudo power spectrum is Gaussian distributed which is shown to be a good approximation. It is also shown how many such patches can be combined and in this way it is possible to analyse datasets with observations of the full CMB sky by breaking them up into several patches which are all analysed simultaneously. In chapter (5) the method is extended to CMB polarisation.





# Chapter 1

## The Physics of the CMB fluctuation

In this chapter I will discuss the physical processes that gave rise to the fluctuations in the Cosmic Microwave Background (CMB). I will start by introducing some basic concepts in cosmology (sect. 1.1), and discuss some main problems that standard cosmology (sect. 1.1.2) is facing. In sect. 1.2.1 I will briefly discuss the theory of inflation in the early universe and show how this theory may solve some of the problems in cosmology as well as explain the origin of structure in the universe. In sect. 1.2.2 an alternative theory of structure formation will briefly be reviewed.

I will proceed by a description of how the density perturbation produced in the early universe were growing. In the early stages, the evolution of primordial density perturbation were linear (sect. 1.3.1). In the last stages also nonlinear effects became important (sect. 1.3.2). I will discuss the physics of the recombination era where light decoupled from matter (sect. 1.4.1) and gave rise to the CMB. It will be shown that by studying the fluctuations in the CMB, one actually studies the density fluctuations at the recombination epoch when structure formation was still governed by linear evolution. The density distribution in the universe today has undergone nonlinear evolution by physical processes that are not completely known. For this reason it is easier to reconstruct the primordial field of density fluctuations and to estimate the values of the cosmological parameters from the spectrum of density fluctuations at the last scattering surface. How this can be done will be reviewed in sect. 1.4.2. In sect. 1.4.3 I will discuss how processes that took place after recombination also contributed to small scale power in the CMB.

The discussion in sect. 1.1 and 1.2 are when no other references are given, built on the book (Kolb and Turner 1990) and the reviews (Watson 2000; Narlikar and Padmanabhan 1991). Section 1.3 mostly follows the books (Kolb and Turner

1990; Coles and Lucchin 1995) and the review (Padmanabhan 1999). The paper (Hu and Sugiyama 1995) and (Zaldariagga and Harari 1995) are the main sources for section 1.4. All the power spectra used in this thesis were obtained using the publicly available code CMBFAST (Seljak and Zeldariagga 1996). The code can be found at the site <http://physics.nyu.edu/matiasz/CMBFAST/cmbfast.html>

## 1.1 Basic Concepts

In this section I will review the basic principles of cosmology (sect.1.1.1). Then I will outline the main problems that the standard Big Bang scenario is facing (sect. 1.1.2).

### 1.1.1 The Friedmann Equations

The Cosmological Principle (CP), is the foundation of almost all modern cosmology. The CP consists of two parts,

1. Homogeneity: The universe is homogeneous on large scales
2. Isotropy: The universe is isotropic on large scales

The homogeneity principle says that the universe is supposed to look the same at every point. The isotropy principle says that the universe is supposed to look similar in all directions. On small scales this is obviously not the case, but observations indicate that this seems to be true on larger scales. One of the best probes is the CMB which has been observed to be isotropic to one part in  $10^5$ .

The metric in a homogeneous and isotropic universe is the Friedmann-Robertson-Walker (FRW) metric,

$$ds^2 = c^2 dt^2 - a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right]. \quad (1.1)$$

Here,  $(r, \theta, \phi, t)$  are the comoving coordinates (coordinates moving with the expansion of the universe). The expansion factor of the universe is  $a(t)$ , and  $k$  is the curvature. By a proper rescaling of the coordinates,  $k$  can take the values  $-1$  (hyperbolic space),  $0$  (flat space) or  $1$  (spherical space).

The goal now is to find the Friedmann equations, which are the 'equations of motion' of the universe. This can be achieved by inserting the FRW metric into the Einstein equations. The Einstein equations read

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \frac{\Lambda g_{\mu\nu}}{c^2} = -\frac{8\pi G}{c^4}T_{\mu\nu}, \quad (1.2)$$

where  $R_{\mu\nu}$  is the Ricci tensor,  $R$  is the Ricci scalar,  $\Lambda$  is the cosmological constant,  $g_{\mu\nu}$  is the metric,  $G$  is the gravitational constant and  $T_{\mu\nu}$  is the stress-energy tensor of the universe. These quantities are described in more detail elsewhere (e.g. (Kolb and Turner 1990)). The left side of the equation contains terms which are purely geometry dependent and the right side has all the matter and energy in it. The energy/matter content of the universe was very different at different epochs, but a quite general form of the stress-energy tensor is that of a perfect fluid. In a perfect fluid  $T_{\mu\nu}$  is diagonal and takes the form (I have set  $c = 1$  for simplicity),

$$T_{\nu}^{\mu} = \text{diag}(\rho, -p, -p, -p). \quad (1.3)$$

Here  $\rho$  is the energy density and  $p$  is the pressure of the fluid.

So inserting the FRW metric (1.1) and the stress-energy tensor for a perfect fluid (1.3) into the Einstein equations (1.2), one arrives at the Friedmann equations:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho_T + 3p_T), \quad (1.4)$$

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G\rho_T}{3} - \frac{k}{a^2}. \quad (1.5)$$

Here the dots denote derivatives with respect to time. The cosmological constant is included here, by letting

$$\rho_T = \rho + \rho_{\Lambda}, \quad (1.6)$$

$$p_T = p - \rho_{\Lambda}, \quad (1.7)$$

where  $\rho_{\Lambda} = \frac{\Lambda}{8\pi G}$ .

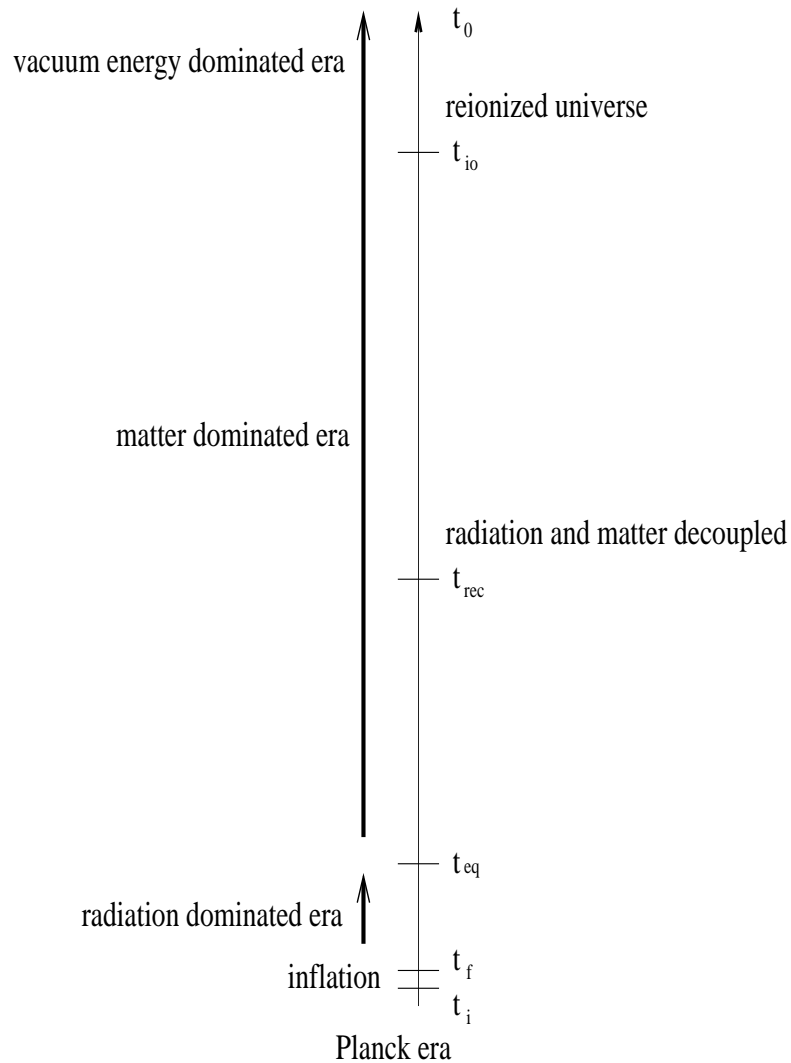


Figure 1.1: A sketch of the history of the universe (the time axis is **not** linear).

To show some applications of the Friedmann equations I will now concentrate on two stages in the history of the universe. The *radiation dominated era* and the *matter dominated era*. At cosmic time  $t_{eq} \approx 10^4$  years the energy densities of matter and radiation in the universe were equal. Well before  $t_{eq}$  there was an era when the energy density in the universe was dominated by radiation. In that epoch, the energy density due to radiation was much higher than the en-

ergy density due to matter, and one can use the equation of state for radiation  $p = \alpha\rho$  with  $\alpha = 1/3$ . After  $t_{eq}$ , matter was dominating the energy density of the universe. In the matter dominated era one can neglect the pressure ( $\alpha = 0$ ). The matter dominated era includes the recombination epoch ( $t_{rec} \approx 10^5$  years) at which matter and radiation decoupled. Recent observation indicate that there is a cosmological constant (Jaffe et al. 2001) and that the universe at present is dominated by vacuum energy ( $\rho_{\text{matter}}/\rho_{\Lambda} \approx 0.4$ ). The equation of state of vacuum energy is  $\alpha = -1$ . The history of the universe is sketched in figure (1.1).

First I will show how the energy density  $\rho$  in the universe is changing with the scale factor  $a$ . To do this, one can combine equation (1.4) and (1.5) to get

$$\frac{d}{da}(\rho a^3) + 3a^2 p - \frac{a^2 \Lambda}{4\pi G} = 0. \quad (1.8)$$

Using the equation of state, one can rewrite this as

$$\frac{d}{da}(\rho a^{3(1+\alpha)}) = \frac{a^{2+3\alpha} \Lambda}{4\pi G}. \quad (1.9)$$

One easily sees that for the matter dominated era,  $\rho \propto a^{-3}$  (actually  $\rho = k_1 a^{-3} + k_2 \Lambda$  where  $k_1$  and  $k_2$  are constants), for the radiation dominated era  $\rho \propto a^{-4}$  and for the vacuum energy dominated epoch  $\rho \propto \ln a$ .

Using the Friedmann equations one can find how the expansion parameter varies with time. As a simple example I will just demonstrate the flat universe case with  $\Lambda = 0$ . In this case one can set  $\rho = k_1 a^{-3}$  for a matter dominated universe into equation (1.5) and integrate the equation. One sees that  $a \propto t^{2/3}$ . Similarly for a radiation dominated universe one has  $a \propto \sqrt{t}$ . In a radiation dominated universe, temperature  $T$  and density are related by  $\rho \propto T^4$  giving  $T \propto t^{-1/2}$ . With this relation one can study how the temperature in the universe varies with time in the early stages of the history of the universe. A similar argument yields for the matter dominated universe  $T \propto t^{-2/3}$ .

I now define some new parameters, namely

- the *Hubble constant*

$$H = \frac{\dot{a}(t)}{a(t)}. \quad (1.10)$$

- The redshift that one observes today from a given epoch  $t$

$$z(t) = \frac{a_0}{a(t)} - 1, \quad (1.11)$$

where  $a_0 = a(t_0)$ , where  $t_0$  is the present time.

- The *density parameter*  $\Omega$  is defined as

$$\Omega = \frac{\rho}{\rho_c}, \quad (1.12)$$

where

- $\rho_c$  is the *critical density*. This is the density at which the universe is exactly flat ( $k = 0$ ). This can be directly found from equation (1.5), setting  $k = 0$  and solving for the density  $\rho_T$ . This gives

$$\rho_c(t) = \frac{3H^2(t)}{8\pi G}. \quad (1.13)$$

The critical density today is  $\rho_c \approx 2 \times 10^{-29} h^2 \text{gcm}^{-3}$  (where I used  $H_0 = H(t_0) = h \times 100 \text{kms}^{-1} \text{Mpc}^{-1}$ ).

Finally I will explain some expressions which will be useful. These are *the particle horizon*, *the event horizon* and *the Hubble radius*.

- The particle horizon is the distance which limits the range of casual communications from the past. It is simply the most distant point from which one can receive a light signal which was emitted at the beginning of the universe.
- Similarly the event horizon limits casual communications to the future. I light signal emitted from us today will never reach beyond the event horizon.
- Finally, the Hubble radius which is given by  $r_H = cH^{-1}$  is in a normal universe a good approximation to the particle horizon. The Hubble radius is often used as an approximation to the size of a causally connected region in the universe, and I will do this also in the following.

### 1.1.2 Problems in Standard Cosmology

The standard Big Bang scenario suffers from several problems. I will now describe some of the most important ones, known as the horizon problem, the flatness problem and the monopole problem. At the end I will also explain the problem of formation of structure in the universe.

- **The horizon problem**

The horizon problem is usually illustrated by comparing the size of casually connected regions at the recombination epoch  $t = t_{rec}$  and today  $t = t_0$ . One has that

$$\int_0^{t_{rec}} \frac{dt}{a(t)} \ll \int_{t_{rec}}^{t_0} \frac{dt}{a(t)}, \quad (1.14)$$

where the left hand side is the distance that a ray of light could have traveled from the beginning of the universe  $t = 0$  to recombination  $t = t_{rec}$  and the right hand side is the distance to the last scattering surface (the recombination epoch is often called the *last scattering surface* as this is the most remote era that can be observed and is therefore where the photos from this era was last scattered on electrons before they free streamed to the present era). This means that any casually connected region at recombination was much smaller than the size of the observable universe today. Still the CMB which was produced in the recombination epoch has a temperature which is uniform to 1 part in  $10^5$  in all direction. The horizon problem is the problem of explaining how the CMB can have the same temperature in two opposite direction, when these parts of the universe have never been casually connected.

- **The flatness problem**

By using equation (1.4) and (1.5), one finds,

$$\Omega(t) - 1 = \frac{k}{a^2(t)H^2(t)}. \quad (1.15)$$

Using this equation at three epochs,  $t_0$ ,  $t_{eq}$  and  $t$ , where  $t$  is now some time before  $t_{eq}$ , one finds

$$\Omega(t) - 1 \approx \left(\frac{a}{a_{eq}}\right)^2 \left(\frac{a_{eq}}{a_0}\right) (\Omega_0 - 1) \approx 10^4(1+z)^{-2}(\Omega_0 - 1), \quad (1.16)$$

where  $H^2 \propto a^{-3}$  and  $H^2 \propto a^{-4}$  was used for the matter- and radiation-dominated eras respectively. One sees that if the universe is not flat ( $k \neq 0$ ), then it was very close to flat at early times because  $\Omega_0$  is of the order 1 today. To explain this one would need a process which makes the universe very close to flat in the beginning. And if the universe is exactly flat, one should be able to explain why. These questions are known as the flatness problem.

- **The monopole problem**

In the earliest stages of the evolution of the universe, the universe was so dense and hot that one needs to take quantum field theory into account to describe the physics of that era. As the universe went from a very hot to a colder phase, several symmetries in field theory should have been broken, for instance as the electroweak interaction split up into the electromagnetic and the weak interaction. Under such symmetry breakings, field theory predict the production of magnetic monopoles. Most theories however, predict a very high density of such monopoles, so high that they should

have dominated the universe. The question of why one doesn't see any of these monopoles today is known as the monopole problem.

Finally I will discuss some problems connected with the formation of large scale structure in the universe. There are two main problems.

- **Primordial seeds**

In the Big Bang model one assumes a completely homogeneous and isotropic universe. There is no process which can produce structure on the scales at which one sees structure today. The problem is twofold, first one needs some process which is able to create inhomogeneities. Second, as will be discussed in sect. 1.3, a density perturbation with length  $\lambda_0 = \lambda(t_0)$  today grows as  $\lambda(t) = \lambda_0(a(t)/a_0)$ . If one assumes  $a(t) \propto t^n$  where  $n < 1$  then  $\lambda \propto t^n$ . The Hubble radius  $H^{-1} \propto t$ . This means that at early times (small  $t$ ), the density perturbation was larger than a casually connected region. There is no known physical process which could create such a perturbation.

- **Density perturbations in the CMB**

The density perturbations at the recombination epoch which were discovered by observations of the CMB, have a density of the order  $10^{-5}$  relative to the background density. Assuming linear evolution, these could not have grown more than a factor  $\Omega_{z_{rec}}$  from decoupling till today. This means that they should not have a density higher than a factor  $10^{-2}$  relative to the background density today. Observations show that they are of order unity for scales larger than  $\lambda \approx 8h^{-1}Mpc$ . There must have been other effects contributing to the evolution.

## 1.2 Theories of the Early Universe

In this section I will briefly review the theory of inflation, as this theory seems to be able to solve the problems mentioned in the last section in a very natural manner. It also provides a mechanism for production of structure in the very early universe. This is important as one can together with linear perturbation theory, predict some properties of the density fluctuations at the recombination epoch which is what one observes when studying the temperature fluctuations in the CMB. I will also briefly discuss cosmic strings which is an alternative theory of structure formation.

### 1.2.1 Inflation

The idea of inflation comes from fundamental physics. Most theories of fundamental physics predict the existence of one or more scalar fields in the universe (for instance the Higgs field giving rise to the Higgs particle which existence is



predicted by field theory but not yet confirmed). This scalar field could have undergone a phase transition in the early universe as the temperature of the universe was dropping. Such a phase transition is expected in several Grand Unified Theories (GUTs), where a higher symmetry group is broken down to  $SU(3) \times SU(2) \times U(1)$ . In this phase transition the minimum of the potential energy for the scalar field is shifted and the scalar field rolls or tunnels from the previous minimum to the new one.

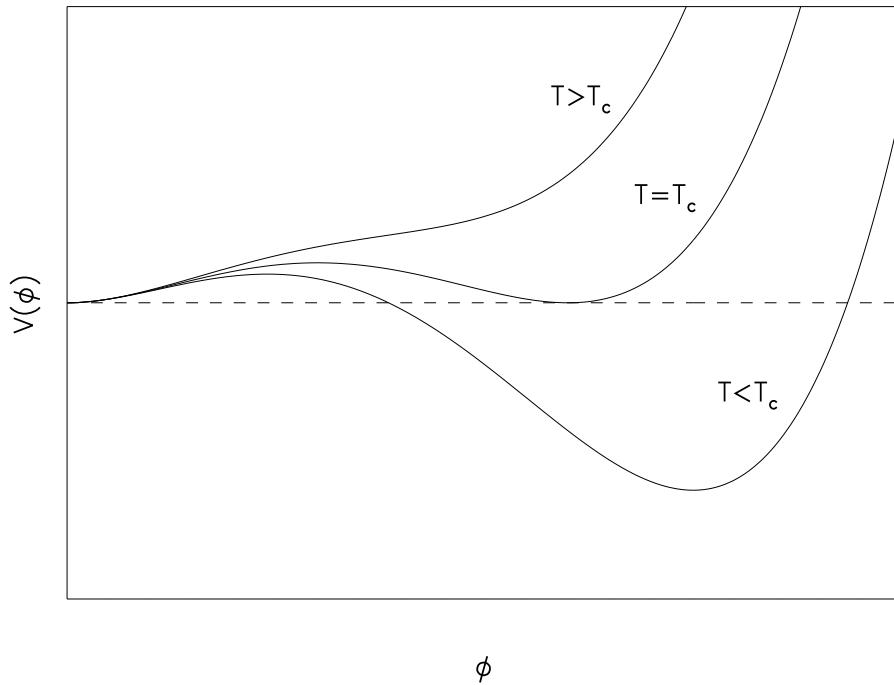


Figure 1.2: The potential of a scalar field

In Figure (1.2), I have sketched one possible potential for a scalar field  $\Phi$ . One sees that for a temperature  $T$  higher than the critical temperature  $T_c$ , the potential had a minimum at  $\Phi = 0$ . After the temperature dropped below  $T_c$ , a deeper minimum arose at  $\Phi = \sigma$ . From this point, the scalar field was not at its real minimum and the universe is said to have a false vacuum. Then the scalar field rolled or tunneled down to the real minimum (or real vacuum). At the end, the scalar field oscillated around the new minimum, interacting with matter fields and creating particles. This is called the reheating period since the universe again got hot after it cooled significantly during the expansion period.

It can be shown that in such a phase transition, the equation of state of the universe is  $p = -\rho$ . Looking at equation 1.4, one sees that the universe in this case has positive acceleration. Moreover, if the phase transition occur in such a way so that the energy density of the universe is roughly constant for a short period of time, then equation 1.4 says that  $\ddot{a}/a$  is constant. This means that  $a(t) \propto e^{At}$ . The expansion of the universe becomes *exponential*. A short period with exponential expansion in the very early universe can solve most of the problems in the standard Big Bang scenario.

If one assumes that the scale factor during a short period of inflation increased by a factor of  $10^{30}$ , one can immediately see how the flatness problem is solved. By looking at equation (1.15), one can see that  $\Omega$  is driven to 1 by such a large growth of the scale factor. *So inflation actually predicts the universe to be very nearly flat.* For this reason, one can set  $k = 0$  in the Friedmann equations (Although  $k \neq 0$  the term is so small compared to the other terms, that  $k$  is effectively 0).

The horizon problem is also naturally solved in this scenario. By evaluating the integrals in equation (1.14) using  $a(t) \propto e^{Ht}$  in the inflationary epoch, one sees that a light ray could have been traveling much longer (in physical coordinates) from the beginning of the universe till recombination, than from recombination and till today. In other words, the whole observable universe today was inside a casually connected region before inflation. This could explain the uniformity of the CMB temperature. The monopole problem is similarly solved. During the inflation, the number of monopoles per comoving volume element was constant, but the physical size of this volume element increased, so the physical density of monopoles decreased with a large factor. Estimates with inflation predicts at the order of one monopole inside the current observable universe.

Finally inflation can solve the problem of the origin of structure in the universe. Before and under inflation, quantum fluctuations in the scalar field  $\Phi$  made some regions of space to have a higher field value  $\Phi + \delta\Phi$  than the background. These regions had a size of the order of a Planck length before inflation. During inflation, the size of these regions (in physical coordinates) expanded with the scale factor. At the end of inflation, the size of these perturbation could have been big enough as to act as seeds for the large scale structure that the universe exhibits today. One important thing to notice here is that most theories predict these quantum fluctuations to be Gaussian distributed. This means that *inflation predicts Gaussian density perturbations in the early universe.* The evolution of the perturbations from the end of inflation to the recombination era is linear so the density fluctuations in that epoch should also be Gaussian. For this reason checking that the temperature fluctuations in the CMB are indeed Gaussian will be an important test of inflation.

Inflation also has a prediction for the density fluctuations. The prediction says that *the density fluctuations should be almost scale invariant*, meaning that there should be a roughly equal number of fluctuations on all physical scales. To understand this one needs to study how the fluctuations in the energy density can arise during inflation. Because of the presence of a horizon, there is confinement in space which due to the uncertainty principle leads to fluctuations in the energy/momentum of the fields present. The size of these fluctuations will be dependent on the horizon size which is almost constant during inflation. This means that the size of the fluctuations arising during inflation will be the same during the whole inflationary period. Inflation then expands these fluctuations out of the causal horizon and they are frozen in. For this reason, the fluctuations on all scales will have a constant amplitude when they reenter the horizon at a later stage. This will be discussed further in section (1.3).

### 1.2.2 Cosmic Strings

As mentioned before, during phase transformations in the early universe, magnetic monopoles should have been created. Magnetic monopoles are just one type of the so called *topological defects*. In this section I will briefly mention another type, called cosmic strings. The theory of cosmic strings was competing with inflation about being the main source of structure in the universe. During the last years, this theory has been more or less ruled out, leaving inflation as the main explanation for the origin of large scale structure. However the cosmic strings could still have played a role, even if they didn't give the main contribution to structure formation. After all, theories of fundamental physics predict these phase transformations which give rise to cosmic strings to have taken place, and for this reason cosmic strings should also have been formed.

To explain what a cosmic string is, I will use as an example a complex scalar field  $\Phi$  which for a temperature  $T$  higher than a critical temperature  $T_c$ , has its minimum potential energy at  $\Phi = 0$ . At this temperature the Lagrangian and the vacuum expectation value  $\langle \Phi \rangle = 0$  are invariant under local  $U(1)$  transformations  $\Phi \rightarrow \Phi e^{i\alpha(\mathbf{x})}$ . At a temperature  $T < T_c$ , the potential energy has a new minimum at  $|\Phi| = \sigma$ . In this case, the  $U(1)$  symmetry is broken, as the vacuum expectation value is now *not* invariant under the transformation  $\Phi \rightarrow \Phi e^{i\alpha(\mathbf{x})}$ . The new vacuum expectation value  $\langle \Phi \rangle = \sigma e^{i\theta(\mathbf{x})}$  can have a different phase at different location (and will also necessarily have different phases, as different points in space which are not casually connected cannot 'know' what phase the other points have.) But if one starts at one point and go through a loop in a two dimensional plane and end up at the same point, the phase must change through this loop with an integer multiple of  $2\pi$ . Say that one has such a loop with  $\Delta\theta = 2\pi$ . If one shrinks this loop to a point,  $\Delta\theta$  can not change continuously

from  $\Delta\theta = 2\pi$  to  $\Delta\theta = 0$ . For this reason, there must be a point in the middle which doesn't have a phase, namely the old minimum  $\langle \Phi \rangle = 0$ . So one ends up with a string in space having a false vacuum, that is the potential energy of the scalar field along this string is higher than that for the true vacuum (where the vacuum expectation value is at the minimum of the potential).

The cosmic string can act as gravitational attractors and seed structure formation. *Cosmic strings lead to non-Gaussian density fluctuations* and can in this way be distinguished by the fluctuations created by inflation. Strings that formed before inflation probably got so diluted during inflation that there are hardly any left (this is what solved the monopole problem). If they formed after inflation, they could still be around and to some extent have contributed to structure formation in the early universe. If they are there, they should be visible in the CMB as non-Gaussian features.

## 1.3 The Growth of Perturbations in the Early Universe

In this section I will review how density fluctuations which arose in the very early universe evolve. I will first discuss the linear regime (sect.1.3.1) where the density fluctuations can be represented as  $\rho = \rho_0 + \rho_1$ , where  $\rho_0$  is the background density and  $\rho_1$  is a perturbation  $\rho_1 \ll \rho_0$ . I will discuss how the fluctuations evolved during the radiation dominated era, in the matter dominated era and beyond the recombination epoch. In section 1.3.2 I will briefly discuss the evolution when the fluctuations became so large that the linear perturbation theory becomes invalid. Finally, the matter power spectrum which is used to quantify the fluctuations is defined in section 1.3.3.

### 1.3.1 Linear Evolution

In the early universe, the density fluctuations were so small compared to the background density that the use of linear perturbation theory is justified. Two cases need to be distinguished. The perturbations which are significantly smaller than the Hubble radius can be treated with Newtonian dynamics. The super-horizon size perturbations however, need a fully relativistic treatment.

I will start by the superhorizon sized perturbations. When the superhorizon perturbations are treated with the relativistic formalism, there is an additional complication due to a *gauge freedom*. When doing perturbation theory, one works with a uniform 'background spacetime' and a physical spacetime which is similar to the background spacetime but with small perturbations added to it. The

gauge freedom which arises, is the freedom to choose the correspondence between coordinated of points in the background spacetime and in the perturbed physical spacetime. A *gauge transformation* is a change in this correspondence, keeping the coordinates of the background spacetime fixed (Bardeen 1980). The treatment of superhorizon sized perturbations is traditionally done in the *synchronous* gauge, but as this gauge has problems with coordinate singularities, the *Newtonian conformal gauge* has been used in newer treatments. Some studies of perturbations growth have been using gauge invariant quantities. For a review of these different approaches, see (Ma and Bertschinger 1995). To avoid these problems, I will use some simple arguments instead of the complete relativistic formalism to estimate the growth of superhorizon sized perturbations. This gives the same result as the relativistic approach.

Consider a flat ( $k = 0$ ) universe with a uniform density  $\rho_0$  and Hubble constant  $H$ . The Friedmann equations give for this universe,

$$H^2 = \frac{8\pi G}{3}\rho_0. \quad (1.17)$$

In this universe, one has a super horizon sized area with a slightly different density  $\rho_0 + \rho_1$ . This part of the universe can be treated as a part of the universe which is no longer flat (because of the difference in density) and therefore with  $k \neq 0$ . Inside this area one has,

$$H^2 + \frac{k}{a^2} = \frac{8\pi G}{3}(\rho_0 + \rho_1). \quad (1.18)$$

These two equations together give for the size of the density fluctuations described by  $\delta = \rho_1/\rho_0$

$$\delta \equiv \frac{\rho_1}{\rho_0} \propto (\rho_0 a^2)^{-1}. \quad (1.19)$$

This gives for matter and radiation dominated universes  $\delta \propto a$  and  $\delta \propto a^2$  respectively.

As the fluctuations grow, their sizes will at some point become smaller than the horizon at a given epoch. They *enter* the horizon at a time  $t_{enter}$ . When the fluctuations have entered the horizon, one can use Newtonian physics to describe their further evolution by using the equations of motions for an expanding perfect fluid. The Eulerian equations for a perfect fluid are given as,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (1.20)$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} + \frac{1}{\rho} \nabla p + \nabla \phi = 0, \quad (1.21)$$

$$\nabla^2 \phi = 4\pi G \rho, \quad (1.22)$$

where the expansion of the universe is not taken into account. This will be included later. Here  $\rho = \rho(\mathbf{r}, t)$ ,  $p = p(\mathbf{r}, t)$  and  $\phi = \phi(\mathbf{r})$  are the density, pressure and gravitational potential respectively at position  $\mathbf{r}$  at time  $t$ . The fluid velocity is  $\mathbf{v} = \mathbf{v}(\mathbf{r}, t)$ . Now consider small perturbations of these quantities around the background value. One can denote the background by subscript 0 and the perturbation by subscript 1. One can then use

$$\epsilon = \epsilon_0 + \epsilon_1 \quad (1.23)$$

in the Eulerian equations. Here  $\epsilon$  can be  $\rho$ ,  $p$ ,  $\mathbf{v}$  and  $\phi$ . The resulting equations for the perturbations are

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \nabla \cdot \mathbf{v}_1 = 0, \quad (1.24)$$

$$\frac{\partial \mathbf{v}_1}{\partial t} + \frac{v_s^2}{\rho_0} \nabla \rho_1 + \nabla \phi_1 = 0, \quad (1.25)$$

$$\nabla^2 \phi_1 = 4\pi G \rho_1. \quad (1.26)$$

Here  $v_s$  is the speed of sound

$$v_s^2 \equiv \left( \frac{\partial p}{\partial \rho} \right)_{\text{adiabatic}} = \frac{p_1}{\rho_1}, \quad (1.27)$$

where the last equality comes about because of the assumption that there is now spatial variations in the equations of state. Combining these equations one gets the differential equation

$$\frac{\partial^2 \rho_1}{\partial t^2} - v_s^2 \nabla^2 \rho_1 = 4\pi G \rho_0 \rho_1. \quad (1.28)$$

Solving for the density perturbation  $\rho_1$ , one gets

$$\rho_1(\mathbf{r}, t) = A \rho_0 e^{-i\mathbf{k} \cdot \mathbf{r} + i\omega t}. \quad (1.29)$$

Inserting this solution into the differential equation (equation 1.28) one finds the dispersion relation

$$\omega^2 = v_s^2 k^2 - 4\pi G \rho_0. \quad (1.30)$$

Here  $k = |\mathbf{k}|$ . From the dispersion relation (equation 1.30), one easily sees that for

$$k < k_J, \quad k_J = \sqrt{\frac{4\pi G \rho_0}{v_s^2}}, \quad (1.31)$$

one gets imaginary values for  $\omega$ . The wavenumber  $k_J$  is called the *Jeans wavenumber* and determines whether  $\omega$  is real or imaginary. In the case of an imaginary  $\omega$ , one sees that the solution (equation 1.29) is exponentially growing with time. In the other case, the perturbation is oscillating. So the *Jeans wavelength*

$\lambda_J = 2\pi/k_J$  is the limiting wavelength. Perturbations bigger than this will be exponentially growing and perturbations which are smaller will be oscillating. Physically one can understand this as the battle between gravitation and pressure. Gravitation is trying to make the perturbation bigger whereas pressure is trying to make it smaller by resisting contraction. When the size of the perturbation is bigger than the Jeans wavelength, the force of Gravity is big enough to win over pressure.

To make this analysis cosmologically relevant, the expansion of the universe has to be taken into account. For an expanding universe, the zeroth order solution to the fluid equations is

$$\rho_0 = \rho_0(t_0)a^{-3}, \quad (1.32)$$

$$\mathbf{v}_0 = H\mathbf{r}, \quad (1.33)$$

$$\nabla\phi_0 = \frac{4\pi G\rho_0}{3}\mathbf{r}. \quad (1.34)$$

Perturbing around this solution,

$$\rho = \rho_0(t_0)a^{-3} + \rho_1, \quad (1.35)$$

$$\mathbf{v}_0 = H\mathbf{r} + \mathbf{v}_1, \quad (1.36)$$

$$\nabla\phi_0 = \frac{4\pi G\rho_0}{3}\mathbf{r} + \nabla\phi_1, \quad (1.37)$$

and solving the Eulerian equations after Fourier transforming one finds the differential equation

$$\ddot{\delta}_k + 2H\dot{\delta}_k + \left( \frac{v_s^2 k^2}{a^2} - 4\pi G\rho_0 \right) \delta_k = 0. \quad (1.38)$$

here  $\delta_k$  is the Fourier transform of the perturbation  $\delta = \rho_1/\rho_0$ ,

$$\delta(\mathbf{r}, t) = (2\pi)^{-3}V \int \delta_k(t) e^{-i\mathbf{k}\cdot\mathbf{r}/a(t)} d^3r, \quad (1.39)$$

where  $V$  is the volume of integration. As before, the sign of the factor before the last term in equation (1.38) is determining whether one gets growing or oscillating solutions. The Jeans wavelength in the expanding universe is given by

$$k_J = \sqrt{\frac{4\pi G\rho_0 a^2}{v_s^2}}. \quad (1.40)$$

If  $k \ll k_J$  the perturbation is growing. In this case one can write equation (1.38) as

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\rho_0\delta = 0, \quad (1.41)$$

where the  $k$  subscript is dropped, as this is valid for both  $\delta$  and  $\delta_k$ . The exact solution is dependent on the parameters of the universe. In a flat  $\Omega = 1$  universe in the matter dominated era,  $H = \frac{2}{3}t^{-1}$  and  $\rho_0 = \frac{1}{6\pi G}t^{-2}$ . Trying with the solution  $\delta = t^n$  inserted into equation (1.41), one finds that  $\delta \propto t^{2/3}$  is a growing mode solution. Apparently the expansion of the universe prevents the fast exponential growth of perturbations in the non-expanding model described above.

To study the evolution in the radiation dominated era, one has to take into consideration that there are different types of particles present. In this case, the total mass density  $\rho_0$  is a sum over densities  $\rho_j$  of each species  $j$  present. For a species  $i$  which is non-relativistic (this is what was assumed in the above derivation), equation (1.41) can then be written

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\rho_0 \sum_j \epsilon_j \delta_j = 0, \quad (1.42)$$

where the sum over  $j$  is a sum over the non-relativistic species and  $\epsilon_j = \rho_j/\rho_0$ . Assuming that there are only two species, photons and the non-relativistic species  $i$ , the equation can be written as ( $\epsilon_j \ll 1$  as photons are dominating in the radiation dominated era),

$$\ddot{\delta} + \frac{1}{t}\dot{\delta} = 0, \quad (1.43)$$

with the solution  $\delta \propto \ln t$ . This shows that in the radiation dominated era, the perturbations can hardly grow.

From the above analysis, it seems that a significant growth of perturbations will be postponed to after  $t = t_{eq}$ . After matter-radiation equality, the perturbations could in principle have started growing, but only if there were perturbations present which were big enough to start growing. If one assumes a spherical density perturbation of size  $\lambda$ , the mass of this perturbation is (assuming uniform density)

$$M = \frac{4\pi}{3}\rho \left(\frac{\lambda}{2}\right)^3. \quad (1.44)$$

For this perturbation to start growing, its size  $\lambda$  must be bigger than the Jeans length  $\lambda_J$ , or stated differently, the mass of the perturbation  $M$ , must be bigger than the Jeans mass  $M_J$ , corresponding to the mass of a spherical perturbation with length  $\lambda_J$ . Assuming a model with two species, baryons and photons, it can be shown that before recombination when the photons are coupled to the baryons, the Jeans mass is bigger than the total mass inside the horizon. Hence, the perturbation cannot start growing as there are now perturbation within the horizon being big enough to have a stable growing mode. After recombination when the pressure is dominated by the hydrogen atoms, the Jeans mass becomes considerably smaller and perturbations can start growing.



From observations of the CMB, the density fluctuations  $\delta$  in the recombination epoch were of the order  $10^{-5}$  and as mention in section (1.1.2), this is not big enough for the perturbations to be able to grow to the current size today. From the analysis in this section, one can see a solution to this problem. If there was a particle species present which does not couple to the photons, density perturbations in this species could have started growing just after  $t = t_{eq}$  and before recombination. The non-baryonic dark matter which existence has been indicated by several observations (e.g. of rotation curves of galaxies), could be such a species. As the species is not coupled to the photons, the evolution of density perturbations was independent of the photon pressure and could therefor have started growing before recombination. The perturbations in this species would have grown between  $t = t_{eq}$  and  $t = t_{rec}$ . After recombination when the baryonic matter perturbations started growing, the gravitational force from the density perturbations in the non-baryonic component would increase the speed of the growth of baryonic density perturbations.

### 1.3.2 Nonlinear Evolution

The linear perturbations theory for density fluctuations in the early universe can be used when the density perturbations are small  $\delta \ll 1$ . At some point, as the perturbations were growing, it had reached a density which was considerably higher than the background density of the universe and linear perturbations theory breaks down. In the nonlinear regime, the evolution of the density perturbations was very complex. Here I will just briefly consider the properties of this nonlinear evolution.

The model used in the beginning of sect.(1.3.1), describing a flat universe ( $k = 0$ ) containing a region with slightly different density can be used here. In that model, the flat universe has a scale factor  $a(t)$  and the part of the universe containing an overdense region has a scale factor  $a'(t)$ . As I now describe a region which has a higher density than the surrounding flat universe, this region will expand and eventually collapse as the density is higher than the critical one. The evolution of the scale factor  $a'$  in this closed 'universe' with Hubble constant  $H'$  and density parameter  $\Omega'$  can be represented by a parameterized set of equations,

$$\frac{a'}{a'_i} = (1 - \cos \theta) \frac{\Omega'_i}{\Omega'_i - 1}, \quad (1.45)$$

$$H'_i t = (\theta - \sin \theta) \frac{\Omega'_i}{2(\Omega'_i - 1)^{3/2}}. \quad (1.46)$$

Here the index  $i$  describes the parameters at some fixed time  $t_i$ . Knowing the parameters  $\Omega'_i$ ,  $a'_i$  and  $H'_i$  one can use the equations to find the expansion parameter  $a'$  at any given time. The density fluctuation will have reached its maximum

expansion when  $\theta = \pi$ , giving

$$\frac{a'_{max}}{a'_i} = \frac{\Omega_i}{\Omega_i - 1}, \quad (1.47)$$

$$H'_i t_{max} = \frac{\pi}{2} \frac{\Omega_i}{(\Omega_i - 1)^{3/2}} \approx \frac{\pi}{2} \left( \frac{a'_{max}}{a'_i} \right)^{3/2}, \quad (1.48)$$

with a density

$$\rho'_{max} = \rho'_i \left( \frac{a'_i}{a'_{max}} \right)^3 = \frac{3\pi}{32Gt_{max}^2 (\rho'_C)_i}. \quad (1.49)$$

Taking the time  $t_i$ , so early that the density was the same as the background density  $\rho'_i = \rho_i = (\rho_C)_i = (\rho'_C)_i$ . Then the density at the maximum expansion of the perturbation is

$$\rho'_{max} = \frac{3\pi}{32Gt_{max}^2}, \quad (1.50)$$

giving

$$\delta_{max} = \frac{\rho'_{max}}{\rho_{max}} - 1 = \frac{9\pi^2}{16} - 1 \approx 4.55. \quad (1.51)$$

At the maximum expansion, the kinetic energy of the overdense region is  $E_k = 0$ , whereas the gravitational energy  $E_g = C/r_{max}$  where  $C$  is a constant. After the maximum expansion, the region starts collapsing and collapses until  $E_k = -E_g/2$ . This is called virialisation and when the overdense region has virialised it has a gravitational energy of  $E_g = C/r_{vir}$ . Equalizing the total energy  $E_t = E_k + E_g$  at maximum and after virialisation one finds that  $r_{vir} = r_{max}/2$ , meaning that the density has increased by a factor of 8 from maximum expansion to virialisation.

### 1.3.3 The Matter Power Spectrum

In the previous sections I have described how the structures in the universe might have arisen in the very early universe and how they should have evolved to the present era. The next step is to find some way of characterising this structure so that the theoretical predictions can be compared to observations. The density fluctuations  $\delta(\mathbf{x})$  at a position  $\mathbf{x}$  in the universe at some early time  $t$ , is a random variable. I will in the following be assuming that this is a Gaussian distributed random variable with ensemble averaged mean value  $\langle \delta \rangle = 0$ . This is the prediction of inflation. As will be discussed, parts of the formalism is also valid if these assumptions do not hold.

The first thing I will define is the *autocorrelation function* describing the ensemble average correlation between  $\delta(\mathbf{x})$  and the fluctuation at some other point  $\delta(\mathbf{x} + \mathbf{r})$  displaced by a vector  $\mathbf{r}$ . This is simply defined as,

$$\xi(\mathbf{r}) = \langle \delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r}) \rangle. \quad (1.52)$$

This can easily be Fourier transformed to give,

$$\xi(\mathbf{r}) = (2\pi)^{-3}V^{-1} \int |\delta_k|^2 e^{-i\mathbf{k}\cdot\mathbf{r}} d^3k, \quad (1.53)$$

$$|\delta_k|^2 = V \int \xi(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} d^3r, \quad (1.54)$$

where  $\delta_k$  is the Fourier transformed density fluctuation given by equation (1.39) and  $V$  is the integration volume. The *power spectrum* of matter density fluctuations is defined in this way as  $P(k) = |\delta_k|^2$ . Since the density fluctuations evolve in time due to gravitational interactions, so does the power spectrum. The *initial* power spectrum is often assumed to be a power law of the form  $P(k) = AVk^n$ , where  $n$  is the *spectral index* and  $A$  is the amplitude.

Since  $\langle \delta \rangle$  is zero, one could calculate what the rms fluctuation  $\sigma^2 = \langle \delta^2 \rangle$  is. Some algebra gives,

$$\sigma^2 = \frac{1}{2\pi^2} \int_0^\infty P(k) k^2 dk. \quad (1.55)$$

This is the integral of the power spectrum over  $k$ -space. Unfortunately the quantity  $\sigma^2$  does not say anything about the relative contribution from different scales  $k$ . For this reason, it is convenient to define the *mass variance* for a given length scale  $R$ . Defining a sphere of radius  $R$ , the mean mass  $\langle M \rangle$  found inside such a sphere is

$$\langle M \rangle = \langle \rho \rangle V = \frac{4\pi}{3} \langle \rho \rangle R^3. \quad (1.56)$$

This leads to the definition of the mass variance for a certain mass scale  $M$ ,

$$\sigma_M^2 = \frac{(\delta M)^2}{\langle M \rangle^2} = \frac{\langle (M - \langle M \rangle)^2 \rangle}{\langle M \rangle^2}. \quad (1.57)$$

For the power law initial power spectrum, the initial mass variance for a mass scale  $M$  can be shown to scale as

$$\sigma_M \propto k^{3/2} \delta_k \propto k^{3\alpha} \propto M^{-\alpha}, \quad (1.58)$$

where  $\alpha = (3 + n)/6$  for the power law power spectrum.

As discussed in the section on inflation (sect.1.2.1), the perturbations produced in inflation are supposed to have the same amplitude at all scales. They are supposed to be *scale invariant*. In this case, this refers to the perturbations

produced in the gravitational potential  $\phi$ . These perturbations can be estimated using the shape of the Newtonian gravitational potential

$$\delta\phi \propto G \frac{\delta M}{M} \propto \frac{\delta M}{M} \frac{M}{R} \propto \sigma_M M^{2/3}. \quad (1.59)$$

Looking at equation (1.58), one sees that for the initial power spectrum of perturbations to be scale invariant,  $\alpha$  must be  $2/3$  and therefore the spectral index  $n$  must be one.

The evolution of the power spectrum with time can be found by studying the evolution of  $\delta$  with time. In linear perturbation theory, one can describe the evolution of  $\delta_k(t)$  as

$$\delta_k(t) = T_k(t, t_i) \delta_k(t_i). \quad (1.60)$$

Here  $T_k(t, t_i)$  is the transfer function which describes the evolution of the  $k$  mode of density perturbations from some initial time  $t_i$  to time  $t$ . This gives for the power spectrum

$$P(k, t) = |T_k(t, t_i)|^2 P(k, t_i). \quad (1.61)$$

In the linear regime, one can find the transfer functions, using the linear theory of the evolution of density perturbations outlined above. In the nonlinear regime, finding the transfer function is non trivial as it is not quite known which physical processes were dominant during the epoch of nonlinear growth of perturbations.

Finally, when using the linear theory one finds that the power spectrum of fluctuations at the time when the fluctuation with wavenumber  $k$  entered the horizon is given by

$$P(k, t_{enter}) \propto k^{n-4}, \quad (1.62)$$

so that

$$\sigma_M(t_{enter}) \propto k^{n-1}, \quad (1.63)$$

which for the Harrison-Zeldovich spectrum is independent of  $k$ . This means that when  $n = 1$ , all fluctuations entered the horizon with the same amplitude.

## 1.4 The Recombination Era and the CMB

In this section I will discuss how the primary anisotropies in the CMB arose (sect.1.4.1). I will describe how the linearized Einstein equations for the density perturbations can be used together with the Boltzmann equation to predict the anisotropies at the last scattering surface. Then in section 1.4.2 I will define the CMB power spectrum and describe how it varies with some of the cosmological parameters. The small scale secondary anisotropies which arose when the CMB photons passed through the reionized part of the universe will be briefly described in section 1.4.3.

### 1.4.1 The Origin of the CMB Anisotropies

To predict the shape of the temperature fluctuations  $\Delta T/T$  observed in the CMB, one has to apply the Boltzmann equation describing the evolution of the photon distribution function in a photon-baryon gas with Compton scattering. The linear density perturbations will now be described using the gauge invariant potentials  $\Psi$  and  $\Phi$  described in (Bardeen 1980). These are related to the total density perturbations  $\delta$  and the anisotropic stress  $\Pi$  by

$$k^2\Phi = 4\pi G\rho_0 \left(\frac{a}{a_0}\right)^2 \delta_k, \quad (1.64)$$

$$\Phi + \Psi = -\frac{8\pi Gp}{k^2} \left(\frac{a}{a_0}\right)^2 \Pi. \quad (1.65)$$

Here (and in the rest of the section),  $\delta_k$ ,  $\Phi$ ,  $\Psi$  and  $\Pi$  denote the Fourier transforms of the quantities  $\delta$ ,  $\Phi$ ,  $\Psi$  and  $\Pi$  respectively. The scale factor is as before  $a$ , the pressure is  $p$ . In a completely matter dominated universe where the pressure  $p$  is unimportant, one sees that the two potentials are related by  $\Psi = -\Phi$ .

The temperature anisotropy at position  $\mathbf{x}$  in space at conformal time  $\eta = \int dt(a_0/a)$  in the direction  $\mathbf{n}$  will now be denoted  $\Delta(\mathbf{x}, \mathbf{n}, \eta)$ , and its Fourier transform  $\Delta_k(\mu, \eta)$  where  $k = |\mathbf{k}|$  and  $\mu = \mathbf{k} \cdot \mathbf{n}/k$ . The Legendre expansion is given by,

$$\Delta_k(\mu, \eta) = \sum_{\ell} (-i)^{\ell} \Delta_{k\ell}(\eta) P_{\ell}(\mu), \quad (1.66)$$

where  $P_{\ell}(\mu)$  is the Legendre Polynomial of order  $\ell$ . Using these definitions, the Boltzmann equation for the temperature anisotropies becomes

$$\dot{\Delta}_k(\mu, \eta) + ik\mu(\Delta_k(\mu, \eta) + \Psi) = -\dot{\Phi} + \dot{\tau}(\eta)[\Delta_{k0}(\eta) - \Delta_k(\mu, \eta) - \frac{1}{10}\Delta_{k2}(\eta)P_2(\mu) - i\mu V_k(\eta)]. \quad (1.67)$$

Here a dot represents the derivative with respect to conformal time,  $V_k$  is the Fourier transform of the baryon velocity and  $\dot{\tau} = x_e n_e \sigma_T a/a_0$  is the differential optical depth to Thompson scattering. The ionization fraction is given by  $x_e$ , the electron density by  $n_e$  and the Thompson cross section by  $\sigma_T$ .

The evolution of the linear density perturbations in baryons  $\delta_b$  and photons  $\delta_{\gamma}$  is given again by the continuity and Euler equations

$$(\dot{\delta}_b)_k = -k(V_k - \Delta_{k1}) + \frac{3}{4}(\dot{\delta}_{\gamma})_k \quad (1.68)$$

$$\dot{V}_k = -\frac{\dot{a}}{a}V_k + k\Psi + \frac{\dot{\tau}(\Delta_{k1} - V_k)}{R}. \quad (1.69)$$

Here  $R = 3\rho_b/(4\rho_{\gamma})$ . The universe is now supposed to consist of 3 species, the photons, the baryons and the dark matter. The density fluctuations in the coupled baryon-photon fluid is described by the Boltzmann equation (1.67) and the

Euler equations (1.68, 1.69). The dark matter is not coupled to any of these species and only shows up through its influence on the gravitational potentials  $\Psi$  and  $\Phi$ .

Before recombination the photons were continuously scattered on the electrons. In such a case the photon distribution is isotropic in the electron rest frame, so the dipole  $\Delta_{k1}$  is only caused by the movement of the baryons,  $\Delta_{k1} = V_k$ . Inserting this into equation (1.68) one gets  $(\dot{\delta}_b)_k = 3/4(\dot{\delta}_\gamma)_k$ , which is the equation for for adiabatic evolution of the density fluctuations. This also gives that  $\Delta_{k\ell} = 0$  for  $\ell \geq 2$ . This is called the *tight coupling limit*.

Close to the tight coupling limit one can use the *the tight coupling approximation*. In this approximation equations (1.67), (1.68) and (1.69) can be expanded to first order in the Compton scattering time  $\dot{\tau}^{-1}$ . This gives the resulting differential equation

$$\ddot{\Delta}_{k0} + \frac{\dot{a}}{a} \frac{R}{1+R} \dot{\Delta}_{k0} + k^2 c_s^2 \Delta_{k0} = F(\eta), \quad (1.70)$$

where  $F(\eta)$  is a term containing the gravitational potentials and  $c_s$  is the sound speed

$$c_s = \frac{1}{3} \frac{1}{1+R}. \quad (1.71)$$

This is the equation of a forced damped oscillator. The gravitational attraction of the over density regions pull the matter into this region, but the photon pressure resists the contraction, giving acoustic oscillations in the photon-baryon fluid.

The solution to equation (1.70), can be written as

$$\hat{\Delta}_{k0}(\eta) = a \cos kr_s(\eta) + b \sin kr_s(\eta) + c(F(\eta)), \quad (1.72)$$

where the hat denotes that this is the solution in the tight coupling approximation. The dipole goes as  $k\hat{\Delta}_{k1} = -3(\dot{\Delta}_{k0} + \dot{\Phi})$  which means that it oscillates with an angle  $\pi/2$  out of phase with the monopole. Here  $r_s$  is the sound horizon, given as

$$r_s(\eta) = \int_0^\eta c_s d\eta'. \quad (1.73)$$

Clearly, the amplitude of the dipole will be damped by a factor  $\dot{r}_s \propto (1+R)^{-1/2}$  with respect to the amplitude of the monopole.

In equation (1.72), the cos-term comes from *adiabatic perturbations* and the sin-term from *isocurvature perturbations*. Adiabatic fluctuations are perturbations for which the entropy is constant causing the relation between the baryon and photon perturbation to  $\delta_b = 3/4\delta_\gamma$ . These kind of fluctuations are predicted by inflation. The isocurvature perturbations are perturbations in the entropy. For these perturbations which are expected by cosmic string models,  $\delta_b = -\delta_\gamma$

so that the total energy density is constant. Recent observations indicate that adiabatic perturbations are dominant so that  $b \approx 0$  in the equation. For this reason I will mainly discuss adiabatic perturbations in this thesis.

One sees that different  $k$  modes oscillate with a different frequency. At recombination  $\eta = \eta_{rec}$ , the monopole for some scales  $k$  is at its maximum or minimum of the oscillations and will show up as peaks in the CMB temperature anisotropy power spectrum. For some other scales the monopole will be zero, and the CMB temperature power spectrum has a trough, which is not zero due to the dipole (which has its maximum/minimum here due to the  $\pi/2$  phase shift). The peaks and troughs in the temperature power spectrum resulting from these oscillations will be described in more detail in section (1.4.2).

When recombination started, the mean free path of the photons became longer as there were less and less electrons present on which they could scatter. The photons were able to diffuse through the baryons and in this way the temperature anisotropies at small scales were getting smaller. This is called Silk damping (Silk 1968). Another result of the increasing diffusion length, is that the CMB photons that one can observe did not last scatter at the same time. The last scattering surface has a finite width and the photons come from different depths in the surface where the oscillations in the monopole  $\Delta_{k0}$  had different phases. This has the similar effect of smearing out the temperature anisotropies on small scales.

When scattering became less frequent the tight coupling limit breaks down and the equations (1.67), (1.68) and (1.69) have to be expanded to second order in the Compton scattering time  $\dot{\tau}^{-1}$ . The solution to this set of equations can be written as

$$(\Delta_{k0} + \Psi) = (\hat{\Delta}_{k0} + \Psi)e^{-(k/k_D(\eta))^2}, \quad (1.74)$$

where  $k_D(\eta)$  is dependent on the diffusion length of the photons at time  $\eta$ . This exponential damping of small scales is caused by the above mentioned effects of photon diffusion and finite width of the last scattering surface.

To find the CMB temperature anisotropies today, the multipoles  $\ell \geq 2$  of  $\Delta_{k\ell}$  have to be found. One can get these by solving the boltzmann equation (1.67), ignoring the quadrupole which disappears in the tight coupling limit and make a multipole expansion of the solution. The result is

$$\begin{aligned} \Delta_{k\ell}(\eta_0) \approx & (\Delta_{k0} + \Psi)(\eta_{rec})(2\ell + 1)j_\ell(k\Delta\eta_{rec}) \\ & + \Delta_{k1}(\eta_{rec})[\ell j_{\ell-1}(k\Delta\eta_{rec}) - (\ell + 1)j_{\ell+1}(k\Delta\eta_{rec})] \\ & + (2\ell + 1) \int_{\eta_{rec}}^{\eta_0} (\dot{\Psi} - \dot{\Phi})j_\ell(k\Delta\eta)d\eta, \end{aligned} \quad (1.75)$$

where  $j_\ell(x)$  is the Bessel function,  $\Delta\eta = \eta_0 - \eta$  and  $\Delta\eta_{rec} = \eta_0 - \eta_{rec}$ . The Silk

damping and the damping due to finite width of the last scattering surface are included by

$$(\Delta_{k0} + \Psi)(\eta_{rec}) = (\hat{\Delta}_{k0} + \Psi)(\eta_{rec})D(k) \quad \Delta_{k1}(\eta_{rec}) = \hat{\Delta}_{k1}(\eta_{rec})D(k), \quad (1.76)$$

where

$$D(k) = \int_0^{\eta_0} \dot{\tau}(\eta) e^{-\tau(\eta, \eta_0)} e^{-(k/k_D(\eta))^2} d\eta. \quad (1.77)$$

One can see from equation (1.75) that the anisotropies in the CMB have three main contributions. The first term in the equation is called *the ordinary Sachs-Wolf effect* (Sachs and Wolfe 1967). This is the sum of the monopole temperature differences on the last scattering surface and the gravitational potential  $\Psi$ . The gravitational potential accounts for the redshift of the photons as they climbed out of the potential wells of the density inhomogeneities at the last scattering surface. The second term in equation (1.75) is due to a Doppler shift. Because of the acoustic oscillations before and at recombination, the baryons were moving. When the photons were last scattered on these moving electrons, they experienced a Doppler shift. Finally, the last main contributor to the CMB anisotropy is the *integrated Sachs-Wolf effect*. This effect arose *after* the photons last scattered. This is due to a possible time change in the gravitational potential after the recombination epoch. This term becomes important if matter-radiation equality occurred close to recombination, or if the universe has got  $\Lambda$ -dominated after recombination.

Clearly, these three terms are dependent on cosmological parameters. If the amount of dark matter and baryons are changed, the size of the density perturbations and the gravitational potentials are also changed. This clearly results in changes of the CMB anisotropies. The speed of expansion of the universe measured by the Hubble constant, is of course also important in these calculations. As are the initial perturbations power spectrum produced by (possibly) inflation. In the next section, I will define the CMB temperature power spectrum of fluctuations and outline how it is dependent on some of these parameters. Equation (1.75) shows how important the CMB power spectrum can be for measuring the cosmological parameters.

## 1.4.2 The CMB Power Spectrum

I start by defining the  $a_{\ell m}$  which are the coefficients of a spherical harmonic expansion of the temperature fluctuations  $\Delta T/T(\theta, \phi)$  on the sky,

$$a_{\ell m} = \int_S d\hat{\mathbf{n}} T(\hat{\mathbf{n}}) Y_{\ell m}(\hat{\mathbf{n}}), \quad (1.78)$$



where  $T(\hat{\mathbf{n}})$  is the fluctuation  $\Delta T/T$  in the position  $\hat{\mathbf{n}}$  and  $Y_{\ell m}(\hat{\mathbf{n}})$  are the spherical harmonic functions. The integration is performed over the full sky denoted by  $S$ . If the initial density perturbations in the universe are Gaussian as predicted by most inflationary theories, so are the  $a_{\ell m}$ . The angular power spectrum of the CMB is defined as,

$$C_\ell = \sum_{m=-\ell}^{\ell} \frac{a_{\ell m} a_{\ell m}}{2\ell + 1}, \quad (1.79)$$

where  $\ell$  is the multipole. Small  $\ell$ -values represent large angular scales and high  $\ell$ -values represent small angular scales. The relation between an  $\ell$ -value and an angular scale  $\Delta\theta$  on the sky can be written as  $\Delta\theta \approx 180^\circ/\ell$ .

If the  $a_{\ell m}$  are Gaussian distributed, the power spectrum  $C_\ell$  contain all the information contained in the CMB. Assuming isotropy of the universe, one has that  $\langle T(\hat{\mathbf{n}})T(\hat{\mathbf{n}}') \rangle = C(\Delta)$  where  $\Delta = \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'$ . This gives that the  $a_{\ell m}$  are uncorrelated  $\langle a_{\ell m} a_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} \langle C_\ell \rangle$ . One sees from equation (1.79) that for a given coefficient  $C_\ell$ , there are  $2\ell + 1$  independent statistical samples of the average value for  $\langle C_\ell \rangle$ . This means that for the higher multipoles, the statistics is better and  $C_\ell$  for a given sky will be closer to the ensemble average  $\langle C_\ell \rangle$ . The smaller multipoles has lower statistics and therefore suffer from the problem of *cosmic variance*.

The angular power spectrum can be determined from the power spectrum in  $k$ -space derived in the last section, by

$$C_\ell = \frac{2V}{\pi(2\ell + 1)^2} \int k^2 |\Delta_{k\ell}(\eta_0)|^2, \quad (1.80)$$

where  $V$  is an arbitrary normalisation constant.

The Sachs-Wolfe effect is the dominant effect on large scales. I will now estimate how the power spectrum  $C_\ell$  depends on  $\ell$  from the Sachs-Wolfe effect. At recombination, one can approximately set

$$\Delta_{k0}(\eta_{rec}) \approx -\frac{2}{3}\Psi(\eta_{rec}) \approx -\frac{2}{3}\Psi(\eta_0). \quad (1.81)$$

Here it is assumed that the pressure is zero at recombination and that the potential has not changed from the last scattering surface and until today. So, only looking at the Sachs-Wolfe effect one gets from equation (1.75)

$$\Delta_{k\ell}(\eta_0) \propto \Psi(\eta_{rec})(2\ell + 1)j_\ell(kx), \quad (1.82)$$

where  $x = \Delta\eta_{rec}$ . Using equation (1.80), one finds

$$C_\ell \propto \int k^2 |\Psi(\eta_0)|^2 j_\ell^2(kx) dk. \quad (1.83)$$

Now using equation (1.64), one sees that  $\Psi(\eta_0) \propto \delta_k k^{-2}$ , so that

$$C_\ell \propto \int k^{n-2} j_\ell^2(kx) dk \propto \frac{\Gamma((2\ell + n - 1)/2)}{\Gamma((2\ell + 3 - n)/2)}. \quad (1.84)$$

In the limit of a Harrison Zeldovich spectrum  $n = 1$ , one has,

$$C_\ell \propto \frac{1}{\ell(\ell + 1)} \quad (1.85)$$

Because of the strong  $\ell$ -dependence from the Sachs-Wolfe effect, it is normal to plot  $C_\ell \ell(\ell + 1)$  in stead of  $C_\ell$ . In figure (1.3) I have plotted the power spectrum of a model with total density  $\Omega_t = 1$ , cosmological constant  $\Omega_\Lambda = 0.7$ , Hubble constant  $H_0 = 100 h$  km/s/Mpc with  $h = 0.82$ , baryon density  $\Omega_b h^2 = 0.03$  and power law index of the scalar power spectrum  $n = 0.975$ . These were the parameters found from the joint Maxima-Boomerang-COBE analysis (Jaffe et al. 2001).

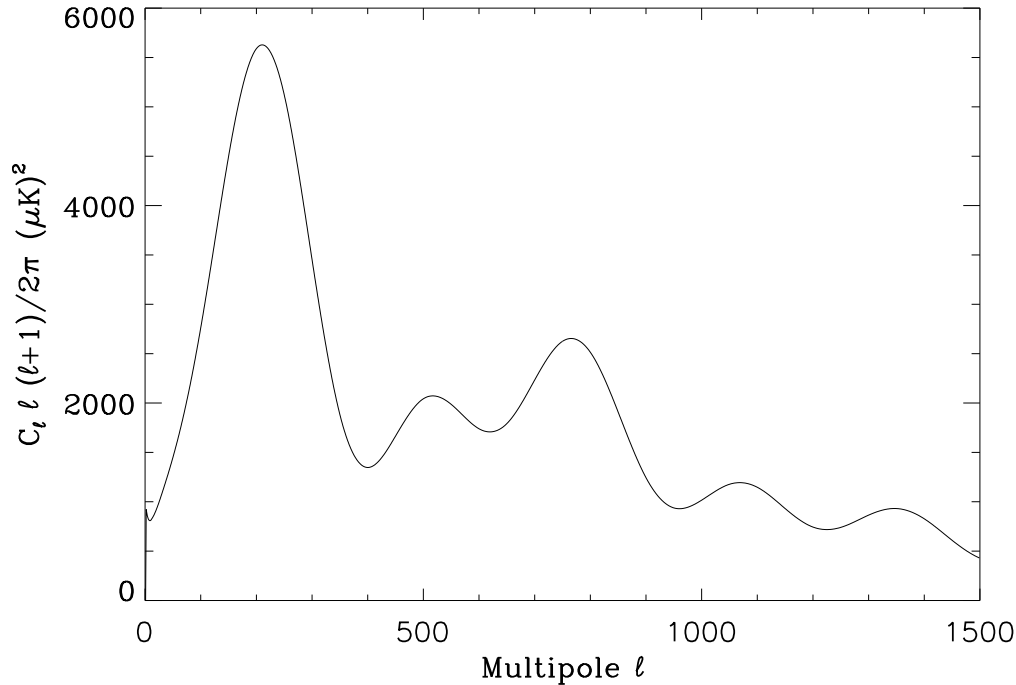


Figure 1.3: The angular power spectrum with  $\Omega_t = 1$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$ ,  $h = 0.82$ ,  $n = 0.975$

Studying figure (1.3), the first apparent features are the peaks and troughs. As mentioned in the previous section, these are the result of the acoustic oscillations which were taking place in the photon-baryon plasma before recombination.

The peaks are resulting from the  $k$  modes of the monopole  $\Delta_{k0}(\eta_{rec})$  which were at their minimum/maximum at recombination. The troughs occur where the monopole was 0. These are filled in by the dipole which was oscillating  $\pi/2$  out of phase and for this reason had its minima/maxima in the troughs. The dipole however had a lower amplitude.

Assuming adiabatic fluctuations, equation (1.72) tells us that one can expect the first peak at  $k = \pi/r_s$ , which is the scale of the sound horizon at last scattering. In a flat universe, the sound horizon at the last scattering surface is expected to have an angular size of about one degree on the sky. This means that for flat universes with adiabatic fluctuations one should expect to see the first peak at a position of  $\ell \approx 200$ . This fits very well with the current data. When the universe has less matter, the universe is open and the geometry of the universe is hyperbolic. In this case the same physical scales at recombination corresponds to a smaller angle on the CMB sky than in the flat universe. So for an open universe, the first peak was expected to be found at higher  $\ell$  values. Also in the case of isocurvature perturbations, the peak should have been at higher  $\ell$  since in this case the wave number for the first peak should be  $k = \pi/(2r_s)$  corresponding to larger scales. One already sees here how the CMB power spectrum can be used to measure values and properties of the universe.

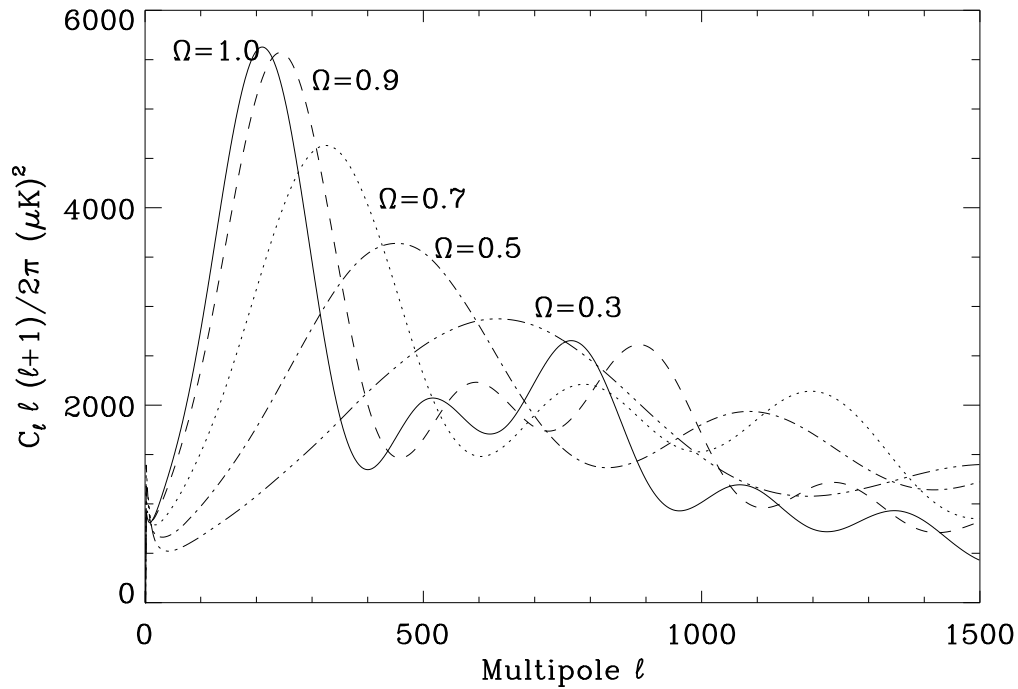


Figure 1.4: The solid line shows the angular power spectrum with  $\Omega_t = 1.0$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$ ,  $h = 0.82$ ,  $n = 0.975$ . The other curves are the power spectra for lower total density  $\Omega = \Omega_t$ . The ratios between  $\Omega_b$ ,  $\Omega_c$  and  $\Omega_\Lambda$  are kept constant.

In figure (1.4) one can see how the peaks are moving to larger scales as the total density is decreasing. I have chosen the same model as in figure (1.3), but varied the total density  $\Omega$  from 0.3 to 1.0 by keeping the ratios of baryon density, dark matter density and vacuum energy density constant.

Another effect is the change of heights of the peaks as the baryon fraction is changed. As described above, the peaks of the power spectrum come from the maximum and minimum of the acoustic oscillations. The oscillations are driven by the gravitational force contracting the density perturbations and the photon pressure resisting this contraction. Every second peak, starting with the first one are coming from the contraction phase. The stronger the contraction, the higher these peaks should be. The other peaks come from the rarefaction phase.

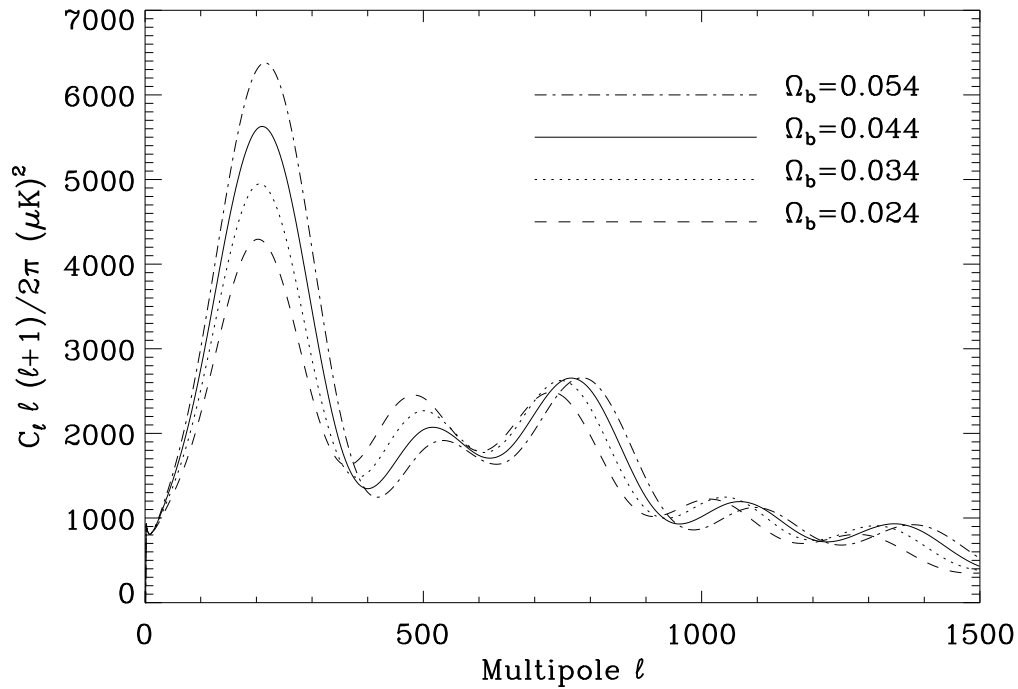


Figure 1.5: The solid line shows the angular power spectrum with  $\Omega_t = 1.0$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$ ,  $h = 0.82$ ,  $n = 0.975$ . The other curves are the power spectra for different baryon densities  $\Omega_b$ . The sum of baryons and dark matter is kept constant,  $\Omega_b + \Omega_c = 0.3$ .

When the baryon content of the universe is higher, the mean free path of the photon and the photon pressure decrease. This means that the contractions will have a higher amplitude and the rarefactions a smaller amplitude. In other words, raising the baryon fraction should increase the height of the odd numbered peaks in the angular power spectrum. Similarly the even numbered peaks should get lower. This is clearly seen in figure (1.5). In the figure I have again used the model in figure (1.3), but this time varied the baryon density to dark matter density ratio, keeping the total density  $\Omega_b + \Omega_c = 0.3$  and  $\Omega_\Lambda = 0.7$ .

From these discussion it should be apparent that the angular power spectrum of the CMB is an important source of the cosmological parameters. By estimating the power spectrum from measurements of the CMB one should be able to measure the cosmological parameters with a very high precision. Some other cosmological parameters that one can measure from the power spectrum include the Hubble parameter  $h$ , the dark matter content  $\Omega_c$ , the power law index of the

power spectrum of scalar and tensor perturbations (tensor perturbations will be discussed in the next section)  $n_s$  and  $n_t$ , the redshift of reionization  $z_{reion}$ , the number of neutrino degrees of freedom at decoupling  $N_\nu$ , the normalisation of the matter power spectrum  $Q$  and the ration of tensor to scalar perturbations  $Q_T^2/Q_S^2$  (Jungman, Kamionkowski, Kosowsky, and Spergel 1996).

### 1.4.3 Reionization and Secondary Anisotropies

After recombination, the universe was *reionized* at some redshift  $z_{reion}$ , which is still unknown. For this reason CMB photons might have been scattered on electrons after recombination. This scattering might have influenced the CMB power spectrum. If reionization happened early ( $z_{reion} > 30$ ) this rescattering might have erased the peaks in the CMB power spectrum and thereby also important information about the universe. As discussed above, the present observations have revealed a peak structure in the power spectrum indicating that reionization happened in an epoch which was so late that the scattering after recombination did not have a significant influence on the CMB on the scales where most of the cosmological information is stored. They could however be the dominant contribution to the power spectrum on smaller scales ( $\ell > 1000$ ).

Here I list some effects that reionization might have on the power spectrum:

1. *Vishniac effect*: During the epoch of galaxy formation, there were huge moving gas masses on which the CMB photons were scattered. The Dopple shift from the scattering on these moving electrons should have influenced the angular power spectrum on smaller scales (Ostriker and Vishniac 1986; Vishniac 1987; Persi, Spergel, Cen, and Ostriker 1995). The strength of this effect and the scales at which it is important is dependent on the reionization history of the universe.
2. *Thermal Sunyaev Zeldovich effect*: When CMB photons passed through hot electron gas in clusters of galaxies and superclusters, their frequencies were changed due to Compton scattering (Zeldovich and Sunyaev 1969). The hot electrons in the clusters transferred energy to the photons resulting in an upscattering of the photon frequency. For this reason the clusters resulted in a temperature decrement at lower frequencies and increment at higher frequencies. This effect has been observed and has been used to measure the Hubble constant (see e.g. (Myers, Baker, Readhead, and Leitch 1997)). The SZ effect is only expected to give significant contributions to the CMB power spectrum on smaller scales (Persi, Spergel, Cen, and Ostriker 1995).
3. *Kinetic Sunyaev Zeldovich effect*: Because of the peculiar velocities of the clusters of galaxies, there will be an additional Doppler shift of the CMB

photons scattered on the electron gas in clusters. This is called the kinematic SZ effect and is about an order of magnitude smaller than the thermal effect. The two effects can be distinguished because of their different spectral signatures. The kinetic effect can in principle be used to measure the peculiar velocities of clusters (Sunyaev and Zeldovich 1980).

There are also other small scale effects caused by gravitation. One of them is associated with the integrated Sachs-Wolfe effect for the time varying gravitational potentials from nonlinear evolution of density perturbation. The movements of the gas in the density fluctuations in the epoch of nonlinear growth was making the gravitational potential time varying and thereby causing a gravitational frequency shift of the CMB photons passing through the perturbation (see equation (1.75)). This is called *the Rees-Sciama effect* (Rees and Sciama 1968) and is contributing (being of the order of magnitude of the primordial perturbations) to the power spectrum on very small scales (typically  $\ell$ -values of a few thousand).

Finally *gravitational lensing* of the CMB by foreground galaxies and galactic clusters might also have influenced the power spectrum. This effect tends to smooth out the peak structure in the power spectrum (Seljak 1996).

#### 1.4.4 Polarisation of the CMB and Tensor Perturbations

Some of the cosmological parameters have similar effects on the power spectrum. For this reason several combinations of cosmological parameters may give the same power spectrum. This *degeneracy* of parameters may be broken by looking at the *polarisation* of the CMB. Since Thompson scattering produce linear photon polarisation, the CMB photons are expected to be linearly polarised.

Polarisation of photons are usually characterized by the Stokes parameters  $I$ ,  $Q$ ,  $U$  and  $V$ . Here  $I$  is the total intensity,  $Q$  and  $U$  describe linear polarisation and  $V$  circular polarisation. Since Thompson scattering is unable to produce circular polarisation,  $V$  is expected to be zero in the CMB and will for this reason not be treated further here.

I will now define the Stokes parameters following (Chandrasekhar 1960). The  $Q$  and  $U$  Stokes parameters can be defined by the components of the electric vector of electromagnetic radiation decomposed in two directions  $x$  and  $y$  which are mutually perpendicular and also perpendicular to the direction of propagation of the radiation. For a plane electromagnetic wave, the components can be written as

$$E_x = E_{x0} \cos(\omega t - \delta_x), \quad (1.86)$$

$$E_y = E_{y0} \cos(\omega t - \delta_y), \quad (1.87)$$

where  $E_{x0}$  and  $E_{y0}$  are the wave amplitudes and  $\delta_x$  and  $\delta_y$  are phases factors. One can then define intensity and the Stokes parameters as

$$I = E_{x0}^2 + E_{y0}^2, \quad (1.88)$$

$$Q = E_{x0}^2 - E_{y0}^2, \quad (1.89)$$

$$U = E_{x0}E_{y0} \cos(\delta_x - \delta_y). \quad (1.90)$$

As with the CMB temperature, the CMB polarisation will also have fluctuations which arose from the acoustic oscillation at the last scattering surface. It is practical to make a spherical harmonic expansion of the CMB polarisation in the same way as for the CMB temperature and describe the fluctuations in terms of polarisation power spectra. There are two approaches to this (Zaldarriaga and Seljak 1997; Kamionkowski, Kosowsky, and Stebbins 1997). In this thesis I will follow (Zaldarriaga and Seljak 1997).

As for the temperature anisotropies, the fluctuations in  $Q$  and  $U$  can be expanded in spherical harmonic coefficients  $a_{\ell m}$ . For polarisation one has to use *tensor harmonics*  ${}_2Y_{\ell m}(\hat{\mathbf{n}})$  and  ${}_{-2}Y_{\ell m}(\hat{\mathbf{n}})$  instead of normal scalar harmonics  $Y_{\ell m}(\hat{\mathbf{n}})$ . These tensor harmonics are described in appendix B. The coefficients can be defined as (from now on  $Q(\hat{\mathbf{n}})$  and  $U(\hat{\mathbf{n}})$  are taken to be  $\Delta Q/Q$  and  $\Delta U/U$  to simplify notation.)

$$(Q + iU)(\hat{\mathbf{n}}) = \sum_{\ell m} a_{2,\ell m} {}_2Y_{\ell m}(\hat{\mathbf{n}}), \quad (1.91)$$

$$(Q - iU)(\hat{\mathbf{n}}) = \sum_{\ell m} a_{-2,\ell m} {}_{-2}Y_{\ell m}(\hat{\mathbf{n}}). \quad (1.92)$$

The inverse relations are

$$a_{2,\ell m} = \int d\hat{\mathbf{n}} {}_2Y_{\ell m}^*(\hat{\mathbf{n}})(Q + iU)(\hat{\mathbf{n}}), \quad (1.93)$$

$$a_{-2,\ell m} = \int d\hat{\mathbf{n}} {}_{-2}Y_{\ell m}^*(\hat{\mathbf{n}})(Q - iU)(\hat{\mathbf{n}}). \quad (1.94)$$

These spherical harmonic coefficients can be redefined in terms of their linear combinations

$$a_{E,\ell m} = -(a_{2,\ell m} + a_{-2,\ell m})/2, \quad (1.95)$$

$$a_{B,\ell m} = i(a_{2,\ell m} - a_{-2,\ell m})/2. \quad (1.96)$$

The indices  $E$  and  $B$  only reflects the fact that these two coefficients have some properties similar to the electric and magnetic fields. Using these coefficients, one can finally define the three power spectra of polarisation

$$C_{E\ell} = \frac{1}{2\ell + 1} \sum_m \langle a_{E,\ell m}^* a_{E,\ell m} \rangle, \quad (1.97)$$



$$C_{B\ell} = \frac{1}{2\ell+1} \sum_m \langle a_{B,\ell m}^* a_{B,\ell m} \rangle, \quad (1.98)$$

$$C_{C\ell} = \frac{1}{2\ell+1} \sum_m \langle a_{T,\ell m}^* a_{E,\ell m} \rangle, \quad (1.99)$$

where the last definition is the *temperature-polarisation cross power spectrum*. From statistical isotropy, one has in analogy to the temperature case

$$\langle a_{E,\ell' m'}^* a_{E,\ell m} \rangle = C_{E\ell} \delta_{\ell\ell'} \delta_{mm'}, \quad (1.100)$$

$$\langle a_{B,\ell' m'}^* a_{B,\ell m} \rangle = C_{B\ell} \delta_{\ell\ell'} \delta_{mm'}, \quad (1.101)$$

$$\langle a_{T,\ell' m'}^* a_{E,\ell m} \rangle = C_{C\ell} \delta_{\ell\ell'} \delta_{mm'}, \quad (1.102)$$

$$\langle a_{B,\ell' m'}^* a_{E,\ell m} \rangle = \langle a_{B,\ell' m'}^* a_{T,\ell m} \rangle = 0. \quad (1.103)$$

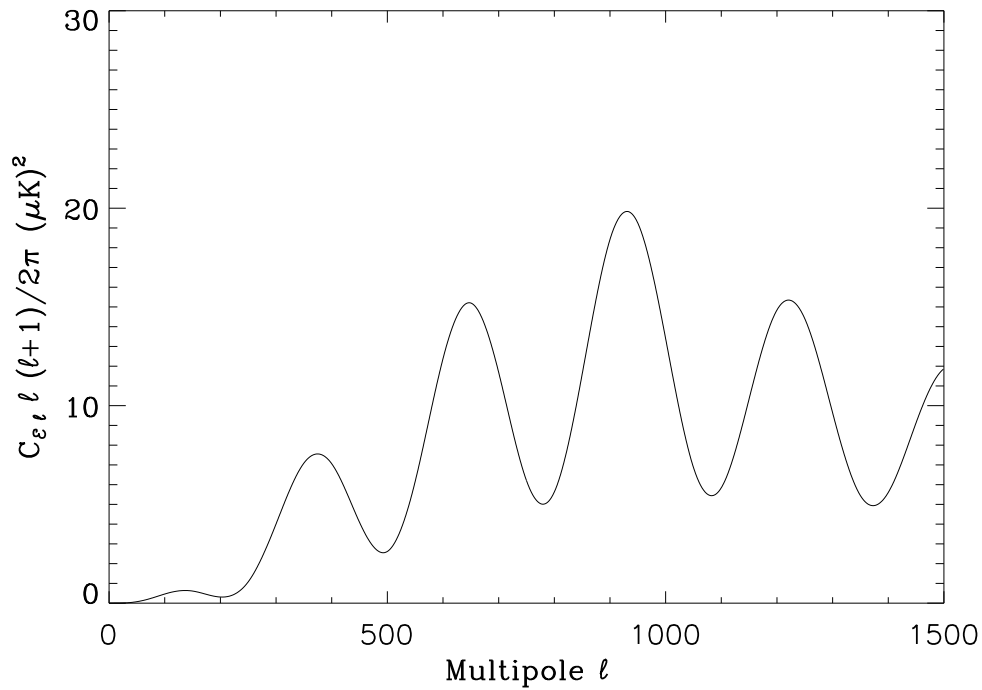


Figure 1.6: The  $E$  power spectrum with  $\Omega_t = 1$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$ ,  $h = 0.82$ ,  $n = 0.975$ .

As showed in (Zaldarriaga and Seljak 1997), the polarisation power spectra for some given model of the universe can be found along the same lines as described in section (1.4.1) and (1.4.2). In addition to the Boltzmann equation for the photon intensity (1.67) one also has to use the Boltzmann equation describing the

evolution of photon polarisation in the presence of Thompson scattering. These power spectra are again dependent on cosmological parameters and will give an independent measure of these. The polarisation power spectra depend differently than the temperature power spectrum on the cosmological parameters. In this way, the polarisation of the CMB can be used to break degeneracies between parameters. In figures (1.6, 1.7) I have plotted the  $E$  and  $C$  power spectra for the same model as in figure (1.3).

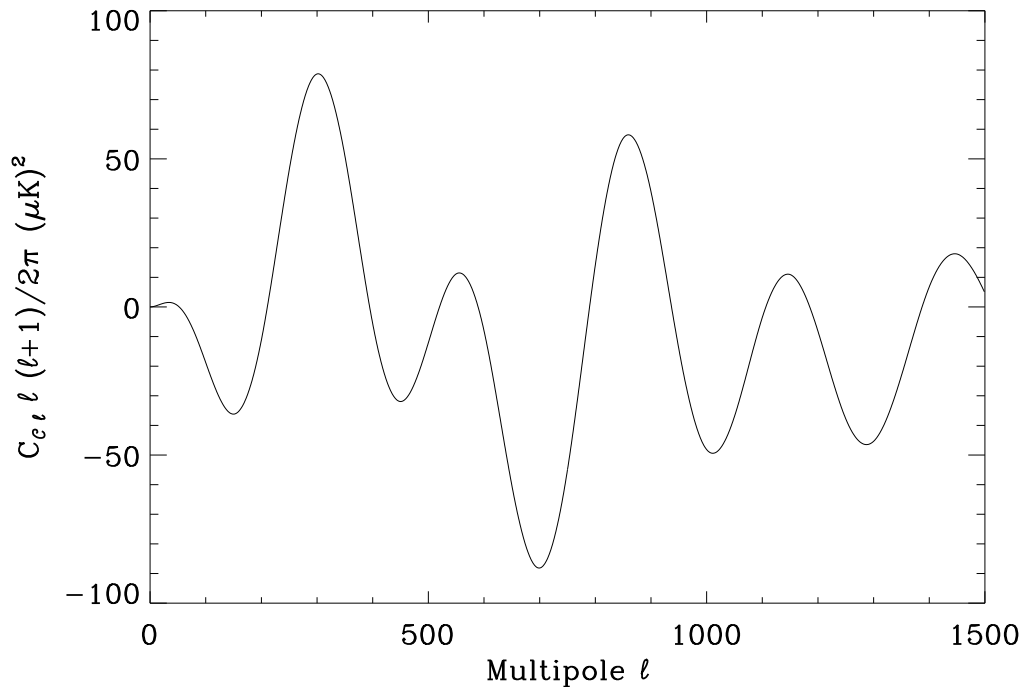


Figure 1.7: The angular power spectrum with  $\Omega_t = 1$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$ ,  $h = 0.82$ ,  $n = 0.975$ .

The discussion so far has been concerned with *scalar perturbations* of the metric. There can also be contributions from *tensor perturbations*. I will now briefly sketch the difference between the two and what the effect of the tensor perturbations can be on the CMB power spectra. In this I will follow (Ma and Bertschinger 1995). For a metric tensor  $g_{\mu\nu}$  where  $\mu, \nu = 0, 1, 2, 3$  the components  $g_{00}$  and  $g_{0i}$  where  $i = 1, 2, 3$  are by definition unperturbed in the synchronous gauge. In this case, the perturbed line element can be written

$$ds^2 = a^2(\tau)(-d\tau^2 + (\delta_{ij} + h_{ij})dx^i dx^j), \quad (1.104)$$

where  $\tau$  is conformal time. There is an implicit sum over all equal indices. The metric perturbation  $h_{ij}$  can be written as

$$h_{ij} = h\delta_{ij}/3 + h_{ij}^{\parallel} + h_{ij}^{\perp} + h_{ij}^T. \quad (1.105)$$

Here the first term is the trace part. The three last terms which together make the traceless part, have the following properties

$$\epsilon_{ijk}\partial_j\partial_i h_{ik} = 0, \quad (1.106)$$

$$\partial_i\partial_j h_{ij} = 0, \quad (1.107)$$

$$\partial_i h_{ij}^T = 0, \quad (1.108)$$

where  $\epsilon_{ijk}$  is the totally antisymmetric Levi-Civita symbol. Obviously,  $h_{ij}^{\parallel}$  can be written in terms of a scalar field  $\mu$ , and  $h_{ij}^{\perp}$  can be written in terms of a divergenceless vector field  $\mathbf{A}$ ,

$$h_{ij}^{\parallel} = (\partial_i\partial_j - 1/3\delta_{ij}\nabla^2)\mu, \quad (1.109)$$

$$h_{ij}^{\perp} = \partial_i A_j + \partial_j A_i \quad \partial_i A_i = 0. \quad (1.110)$$

The  $h$  and  $\mu$  parts make up the *scalar metric perturbation*. The  $h_{ij}^{\perp}$  is the *vector perturbation* and  $h_{ij}^T$  is the *tensor perturbation*. The scalar perturbation have already been treated above and the vector perturbations are not expected to make any significant contribution to the CMB. The tensor perturbations (gravity waves) however, are predicted by most inflationary theories.

Following (Turner, White, and Lidsey 1993),  $h_{ij}^T$  can be Fourier expanded as

$$h_{ij}^T(\mathbf{x}, \tau) = (2\pi)^{-3} \int d^3k h^i(\mathbf{k}, \tau) p_{jk}^s e^{-i\mathbf{k}\cdot\mathbf{x}}, \quad (1.111)$$

where  $p_{ij}^s$  are polarisation tensors ( $s = 1, 2$ ) which satisfy  $p_{ij}^s k^j = 0$  and  $p_{ij}^s \delta^{ij} = 0$ . The evolution of these gravity wave perturbations can be described by the massless Klein-Gordon equation

$$\ddot{h}^i(\mathbf{k}, \tau) + 2\frac{\dot{a}(\tau)}{a(\tau)}\dot{h}^i(\mathbf{k}, \tau) + k^2 h^i(\mathbf{k}, \tau) = 0. \quad (1.112)$$

The solutions of this equation for modes that are well outside of the horizon says that  $h^i(\mathbf{k}, \tau)$  is constant. For modes which are well inside the horizon, one gets  $h^i(\mathbf{k}, \tau) \propto \cos k\tau/a$ .

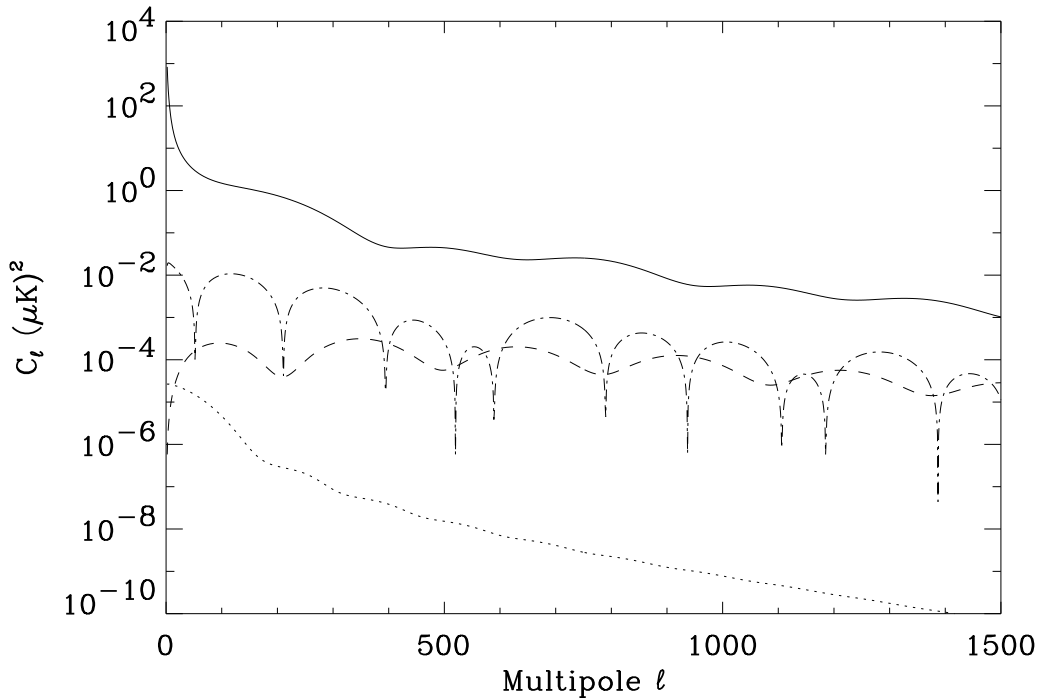


Figure 1.8: The angular power spectrum with  $\Omega_t = 1$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$ ,  $h = 0.82$ ,  $n = 0.975$ ,  $n_T = 1 - n$ ,  $Q_T/Q_S = -7n_T$ . The solid line is  $C_\ell$ , the dashed line is  $C_{E\ell}$ , the dotted line is  $C_{B\ell}$  and the dashed-dotted line is  $|C_{C\ell}|$ .

If the initial power spectrum of scalar fluctuations goes as  $P(k) = Q_S k^n$  and of tensor fluctuations as  $P(k) = Q_T k^{n_T}$ , then most inflationary theories predict  $n_T \approx 1 - n$  and  $Q_T/Q_S \approx -7n_T$ . Using this, I have plotted the different CMB power spectra including  $B$  polarisation in figure (1.8). The model is the same as in the previous figures, but now contributions from tensor perturbations are also included. Clearly for a completely scale invariant  $n = 1$  model, there will be now contributions from tensor modes. Most theories of inflation however, do predict small deviation from a scale invariant power spectrum. This suggests that  $n_T$  and thereby  $Q_T$  are also very small. For this reason the  $C_{B\ell}$  power spectrum which only has contributions from the tensor perturbations is a magnitude of order smaller than the other spectra.

# Chapter 2

## The Analysis of CMB Data

The previous chapter was discussing the physical theories behind the fluctuations in the CMB and how the observation of the CMB power spectrum can test these theories as well as estimate their parameters. This chapter is a review of the previous, current and future experiments intended to measure the CMB fluctuations and how information can be extracted from these observations.

Section (2.1) is an overview of some of the biggest experiments. In section (2.1.1), (2.1.2), (2.1.3) and (2.1.4) I will discuss the COBE satellite, the Boomerang balloon borne experiment, the MAP satellite and the Planck satellite respectively. I will briefly mention the scan strategy and some of the parameters of the instruments. Other experiments will briefly be mentioned in section (2.1.5).

The problem of making CMB sky maps from a time stream of data will be discussed in section (2.2.1). The CMB maps from observations contain in addition to CMB, other sources of radiation which radiate at the same wavelengths as the CMB. These *foregrounds* and how they can be removed from CMB maps is discussed in section (2.2.2). Different methods of extraction of the CMB power spectra from CMB maps with foregrounds removed will be reviewed in section (2.3).

### 2.1 CMB Experiments, the Past, Present and Future

This section is a short review of some of the CMB experiments. I will start with the experiment which first measured the CMB temperature fluctuations. Then I will discuss a balloon borne experiments which is already conducted but will fly again in an improved version. Then two satellite experiments are described, one is already collecting data as this is written and the other will be launched in a few years.

### 2.1.1 COBE

COBE was the first experiment to measure the fluctuations in the CMB. The COBE satellite was launched in November 1989 into a circular orbit around the earth at an altitude of 900km with a  $99^\circ$  inclination. The spacecraft was always pointing away from the earth and was pointing in a direction being almost perpendicular to the direction of the sun. The orbital period was 103 minutes, giving about 14 full revolutions around the earth a day. COBE was scanning the entire sky once every 6 months, and it was operating for a total of 4 years.

COBE was equipped with three instruments, FIRAS, DMR and DIRBE. FIRAS was measuring the absolute temperature of the CMB, whereas DMR was measuring the fluctuations. The DIRBE instrument was measuring the cosmic infrared background radiation. The FIRAS instrument was pointing along the spin axis of the space craft. The two other instruments were both offset with  $30^\circ$  from the spin axis around which the space craft was spinning at a rate of 0.8 revolutions per minute.

The DMR beam was observing at the frequencies 31.5, 53 and 90GHz with a  $7^\circ$  FWHM beam. After 4 years of observation, maps with 6144 pixels were made for all 3 channels (Bennet et al. 1996). By cutting out the contaminated galactic plane (a  $40^\circ$  belt around the equator of the map in galactic coordinates), the power spectrum of the CMB up to  $\ell = 30$  could be estimated. This was done for the 4 year data in (Górski et al. 1996). The power spectrum was shown to be consistent with a  $n_s = 1$  Harrison-Zeldovich spectrum.

### 2.1.2 BOOMERANG

The BOOMERANG experiment is a cooperation between the universities in Rome and institutes in America and Great Britain. This is a balloon borne experiment which had its test flight in Texas in August 1997, lasting for 6 hours. They collected data from a 200 square degree patch of the sky at two frequencies 90GHz and 150GHz. From the data, the power spectrum was estimated, the first peak was found and  $\Omega_t$  was estimated (Mauskopf et al. 2000; Melchiorri et al. 2000).

The main flight of BOOMERANG was conducted in Antarctica at the end of 1998 (Bernardis et al. 2000). The balloon was launched from the McMurdo Station on December 29th, and returned after 259 hours at an altitude of about 38 km. The sky temperature was measured by comparing the temperature of the incoming photons with an onboard thermal reference load. The gondola was equipped with a 1.2m mirror and a cryogenic mm-wave bolometric receiver. Data was collected from an area with minimal contamination from galactic dust(see

section (2.2.2)). This time, also the frequencies 240GHz and 400GHz were observed. A rectangular area of about  $100^\circ$  by  $30^\circ$  was scanned and maps with  $14'$  pixels have been made.

The different frequency channels were used to estimate the foregrounds. The main foreground contamination in the BOOMERANG maps comes from thermal emission from interstellar dust grains. The 410GHz channel which is dust dominated, was used to measure the amount of dust contaminating the data, but was found to be negligible (Netterfield et al. 2000). Three quasars were however cut away from the map before the map was analysed to get the power spectrum. In (Netterfield et al. 2000), the results of the power spectrum and cosmological parameter estimations are showed. The two first peaks of the angular power spectrum are clearly visible and there are indications of a third peak. Using this power spectrum, they estimated several cosmological parameters. Their estimate of the total  $\Omega$  is close to one, confirming the prediction of inflation. Also the index of scalar fluctuations  $n_s$  is found to be very close to one, which fits very well with the prediction of inflation. The baryon fraction  $\Omega_b$  is in very good agreement with what has previously been derived from nucleosynthesis. BOOMERANG is expected to fly again in the end of 2001. This time it will also attempt to measure the polarisation of the CMB.

### 2.1.3 MAP

MAP (Microwave Anisotropy Probe) is a NASA project. The satellite was launched on June 30th 2001 and has at the time when this is written, reached its destination which is the Lagrange L2 point. MAP is orbiting the L2 point which is situated 1.5 mill kilometers away from the earth on the opposite side of the sun, on the line between the sun and earth (see figure (2.1)). The advantage of observing from this point is that the telescope can always point away from the sun, earth and moon. The information in this section is taken from the MAP homepage <http://map.gsfc.nasa.gov/>.

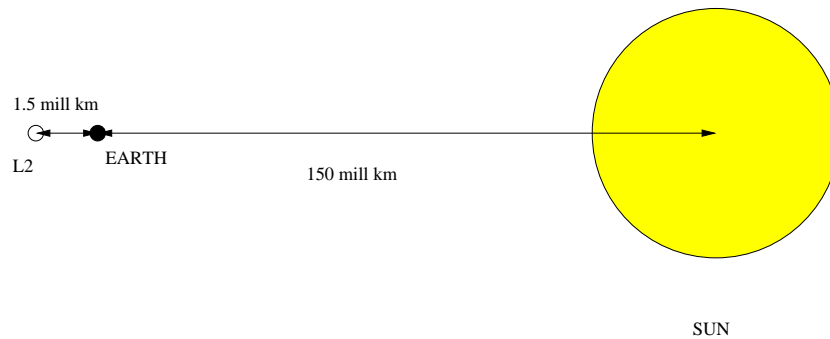


Figure 2.1: The position of the Lagrange L2 point. NB! The figure does not have a linear length scale!!!

MAP will be spinning about its axis with a period of 2.2 minutes in its orbit around the sun. The axis will also be precessing with a  $22.5^\circ$  angle about the Sun-MAP line with a period of 1 hour. In this way MAP is scanning the sky on circles with diameter  $141^\circ$  and observing 30% of the sky every day. It takes 6 months for the spacecraft to observe the full sky. MAP is designed to operate for two years and produce 4 full sky observations.

MAP is a differential experiment meaning that it measures the temperature at one position in the sky and compares with the temperature at another. In



this way it measures the temperature difference between points in the sky. This is different from Planck and BOOMERANG which both compare the sky temperature to a source with known temperature. The two points on the sky are for MAP about  $141^\circ$  away (depending on the channel). This large distance is necessary to maintain sensitivity to signals at large angular separation.

MAP is observing in 5 different frequency bands in order to separate galactic foregrounds signals from the CMB. The first channel is at 22GHz which is just at the atmospheric water line. Frequencies below this can be observed from the ground. This channel is dominated by galactic synchrotron and free-free radiation (see section (2.2.2)). The other channels are at the frequencies 30, 40, 60 and 90GHz where the CMB is dominating. Above 100GHz, thermal dust radiation is starting to dominate.

MAP is equipped with two Gregorian telescopes with  $1.4\text{m} \times 1.6\text{m}$  (elliptical) primary mirror. The angular resolution is varying with the different frequency channels. The lowest frequency channel has a beam with a FWHM of about 56 arcminutes. The highest frequency has 14 arcminutes FWHM. The sensitivity per  $0.3^\circ \times 0.3^\circ$  is expected to be about  $35\mu\text{K}$  for each channel which means that the power spectrum can be measured with reasonable accuracy to a multipole of about  $\ell = 1000$ .

#### 2.1.4 Planck

Planck is a project of the European Space Agency ESA. The Planck satellite is scheduled for launch in year 2007 together with the FIRST satellite. As with MAP, Planck is also going to an orbit about the L2 Lagrange point. Planck will also be spinning about its own spin axis about once every minute and thereby scan the sky with rings. The telescope will be pointing  $85^\circ$  away from the spin axis, giving circles with a diameter of  $170^\circ$ . All the information about Planck in this section is from the Planck Homepage <http://astro.estec.esa.nl/SA-general/Projects/Planck/>.

As MAP, Planck will be observing the full sky once every 6 months. It will be using a  $1.5\text{m} \times 1.3\text{m}$  (off axis paraboloid) telescope. It will be observing in 10 frequency channels located between 30 and 1000GHz. Two different instruments will be observing at different frequencies. The *high frequency instrument* (HFI) is a bolometer array operating at  $0.1\text{K}$  and covering the frequencies above 100GHz. The *low frequency instrument* (LFI) is an array of tuned radio receivers based on HEMT amplifiers and observes the frequency range below 100GHz.

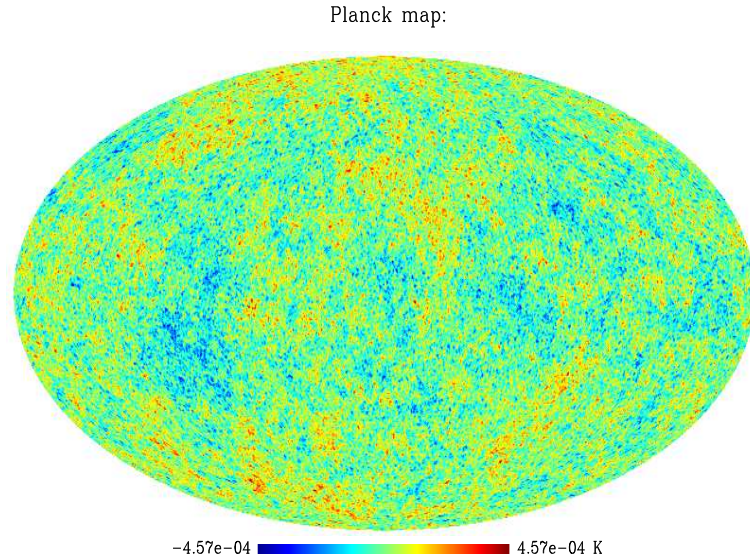


Figure 2.2: The sky map of a realisation of the power spectrum in figure (1.3). Here a 10 arc minute FWHM beam has been used. This is typically what one could expect from one of the Planck channels

The HFI will consist of an array of 48 bolometers split into 6 frequency channels at 100, 143, 217, 353, 545 and 857GHz. The choice of frequencies have been made in order to remove foregrounds (thermal radiation from dust is the most important sources at these frequencies), observe the CMB and to study the Sunyaev-Zeldovich effect (see section (1.4.3)). For instance the channel at 217GHz is at the zero point of the thermal SZ effect, allowing for studies of the kinematic effect. The angular resolution in the different channels varies from 5 arcminutes for the highest frequencies to 11 arcminutes (FWHM) for 100GHz. The expected sensitivity goes from  $6670\mu\text{K}$  per pixel for 857GHz to  $1.7\mu\text{K}$  per pixel for the 100GHz channel.

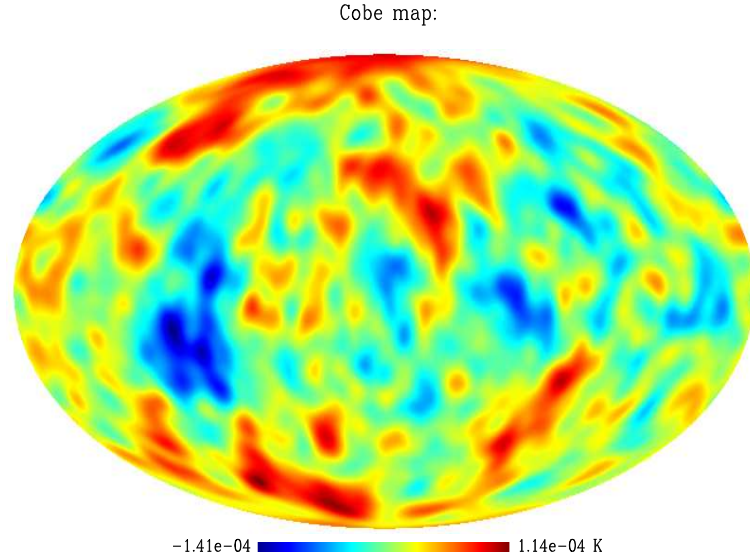


Figure 2.3: This is the same map as in figure (2.2). The only difference is that the beam has  $7^\circ$  FWHM beam, the same as COBE.

The LFI will have 56 detectors covering the frequencies 30, 44, 70 and 100GHz. These frequencies are mostly dominated by CMB. The angular resolution for these channels will be 33, 23, 14 and 10 arcminutes respectively. The sensitivity will go from  $1.6\mu\text{K}$  to  $4.3\mu\text{K}$  for the lowest and highest frequency respectively.

In figures (2.2) and (2.3), the difference in angular resolution by Planck and COBE is illustrated. Planck is expected to get a reasonable estimate of the power spectrum up to a multipole of  $\ell = 2000$ . It will also attempt to measure the polarisation power spectra.

In figure (2.4), I have plotted the power spectrum in figure (1.3) together with the expected noise power spectrum for 2 Planck channels (143GHz HFI and 100GHz LFI) and for a combination of 3 MAP channels. The colored areas show the optimal error bars from estimations of single multipoles from the Planck 143GHz HFI channel and the three combined MAP channels. These error bars are of course smaller by a factor  $\sqrt{n}$  if  $n$   $C_\ell$  are binned together.

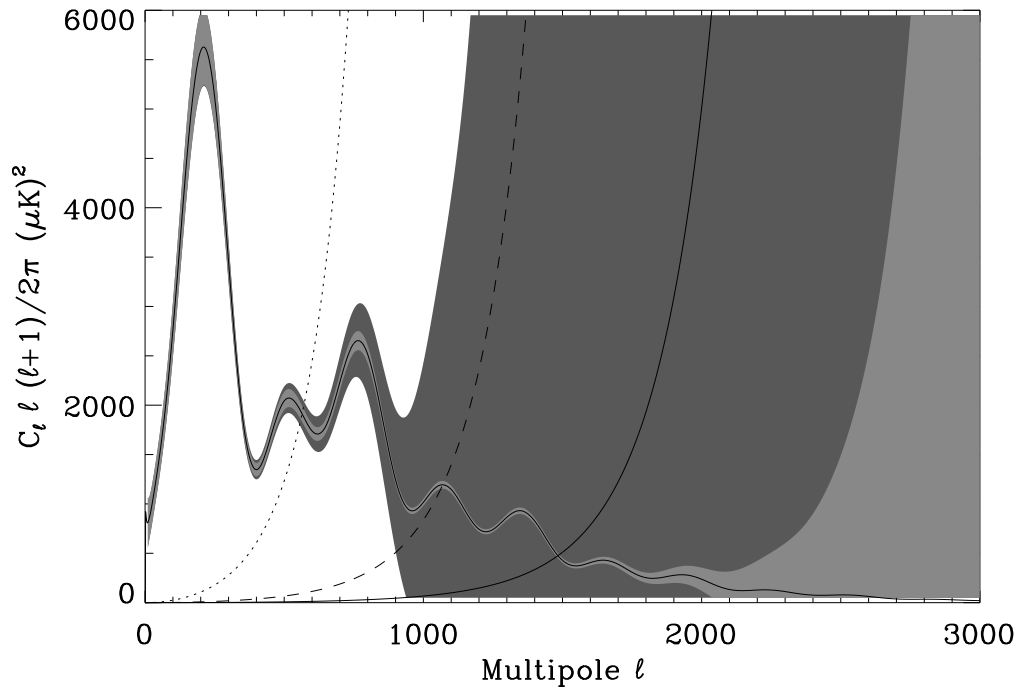


Figure 2.4: The CMB power spectrum in figure (1.3) and the noise power spectrum of the Planck 143GHz HFI channel (solid line), the Planck 100GHz LFI channel (dashed) and of the combination of the 3 highest frequency channels on MAP (dotted line). The light red colored area shows the expected error bars from the 143GHz channel on Planck and the dark red area shows the error bars from the three combined MAP channels. These are the optimal error bars if each single multipole is estimated.

### 2.1.5 Other Experiments

The MAXIMA-1 balloon borne experiment was conducted on August 2, 1998. Data was collected for a few hours by a 1.3 meter telescope. The detectors which were bolometers similar to those planned for Planck HFI, were centered at the frequencies 150, 240 and 410GHz. The sky was mapped with a 10 arc minute FWHM beam in all frequencies. The map and power spectrum estimated from the MAXIMA-1 data is presented in (Hanany et al. 2000). The data shows the two first peaks in the power spectrum. In (Jaffe et al. 2001), the results of a joint analysis of the MAXIMA-1 and BOOMERANG data sets are presented.

The Tophat experiment is another balloon borne experiment. It was launched

on the south pole January 4th 2001 and observed till January 31th, mapping 6% of the sky. Tophat differs from other experiments in that it has the payload on the top of the balloon. It was observing in the frequency range 150 to 630GHz and is expected to give an estimate of the power spectrum up to multipole  $\ell = 700$ . More information can be found at <http://topweb.gsfc.nasa.gov/>.

Archeops is a balloon borne experiment which had its first scientific flight flight in Kiruna in Sweden, January 29th 2001. Archeops aims on large sky coverage to reduce the sample variance in the power spectrum. The first flight which lasted 7 and a half hours, covered more than 20% of the sky. Archeops uses a 1.5 meter telescope similar to that which will be used for Planck. The telescope is spinning during flight, producing a data set consisting of rings on the sky. The detectors used are bolometers similar to those used for the Planck HFI. The frequencies observed are 143, 217, 353 and 545GHz and the angular resolution will allow power spectrum estimation up to  $\ell = 1000$ . More information can be found on (Benoit et al. 2001) or at [http://archeops01.free.fr/main\\_archeops/index\\_english.html](http://archeops01.free.fr/main_archeops/index_english.html)

Finally, I will mentioned a ground based experiment called the *Very Small Array* (VSA). The VSA is a fourteen-element interferometer situated at Tenerife. It observes since September 2000 at the wavelengths 26 to 36GHz. It will observe fluctuations in the CMB in the multipole range  $\ell = 150$  to 1800. More information can be found in (Taylor 2001)

## 2.2 The Analysis of CMB Data Sets

In this section I will discuss how one makes sky maps of a time stream of data from a CMB experiment. In this part I used (Tegmark 1997a) as the main source but other sources have also been use where this is indicated. Then the process of removing galactic foreground contamination from the maps is discussed. The resulting map is supposed to contain only the CMB signal and noise from the detectors. The map is then ready for CMB power spectrum analysis which will be the topic of section (2.3).

### 2.2.1 Map Making

The data stream from a CMB experiment, called the *time ordered data* (TOD), is a vector  $\mathbf{d}$  of temperature measurements from different points on the sky. The telescope beam has a certain beam profile, so that the measurement in element  $d_i$  of the TOD is a convolution of the real underlying sky with this beam profile. The beam profile is often assumed to be symmetric and usually a Gaussian since this simplifies further calculations. This is not always a good assumption as there

are also small contributions from sidelobes.

A sky map  $\mathbf{m}$  can be written in terms of a pointing matrix  $\mathbf{P}$  and a noise vector  $\mathbf{n}$  as

$$\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{n}. \quad (2.1)$$

Here the TOD  $\mathbf{d}$  is supposed to have  $N_d$  elements and the map  $N_p$  pixels. The pointing matrix is a  $N_d \times N_p$  matrix containing the information about the sky position (the pointing of the telescope) for each element of the TOD. In the general case,  $\mathbf{m}$  is the real underlying map and  $\mathbf{P}$  is a complicated matrix containing a complicated beamprofile at each pointing. However if one assumes the beam profile to be simple and symmetric, then  $\mathbf{m}$  can represent the beam smeared map and  $\mathbf{P}$  is a matrix with one element per row, giving the pixel number of each element  $d_i$  in the TOD.

The problem of map making, is the problem of finding the best estimate of the map  $\tilde{\mathbf{m}}$ . The linear solutions to the map can be written

$$\tilde{\mathbf{m}} = \mathbf{W}\mathbf{d}. \quad (2.2)$$

One now needs to find a matrix  $\mathbf{W}$  which constructs a map  $\mathbf{m}$  which has as much as possible (and preferably all) of the information contained in  $\mathbf{d}$ . The map making method used for COBE had

$$\mathbf{W} = (\mathbf{P}^T\mathbf{N}^{-1}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{N}^{-1}, \quad (2.3)$$

where  $\mathbf{N} = \langle \mathbf{nn}^T \rangle$  is the noise correlation matrix.

This choice of  $\mathbf{W}$  has the advantage that the map making is lossless. That is, the map has all the information that is contained in the TOD when assuming that the data is Gaussian. It is shown in (Tegmark, Taylor, and Heavens 1997) that the inverse of the *fisher information matrix*  $\mathbf{F}$  is a good approximation to the covariance matrix for the best possible estimates of parameters from a data set. For Gaussian data with zero mean and covariance matrix given by  $\mathbf{C}$ , the fisher matrix is

$$F_{ij} = \frac{1}{2}Tr(\mathbf{C}^{-1}\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{C}_{,j}), \quad (2.4)$$

where the  $,i$  subscript means derivative with respect to parameter  $i$ . It is shown in (Tegmark 1997a) that the fisher information matrix for the map and for the TOD is the same when the map making procedure is given by equation (2.3).

In principle making the lossless map should be easy. Given equations (2.2) and (2.3) the estimate of the map is

$$\tilde{\mathbf{m}} = (\mathbf{P}^T\mathbf{N}^{-1}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{N}^{-1}\mathbf{d}. \quad (2.5)$$

Knowing the TOD  $\mathbf{d}$ , the noise properties of the detectors  $\mathbf{N}$  and the pointing matrix  $\mathbf{P}$  the map could easily be computed by this formula. There are however problems. The first one being the matrix inverses in equation (2.5). The matrices to inverse scale as  $N_p \times N_p$ . For the forthcoming CMB experiments  $N_p$  will be of the order of millions. Matrix inversion of a  $N \times N$  matrix takes  $N^3$  operations and for  $N \propto 10^6$  this becomes unfeasible with existing computers. The second problem is that the noise properties are normally not known beforehand. They must be estimated from the TOD together with the map.

Assuming that the noise correlation matrix  $\mathbf{N}$  is known, (Wright 1996) presented a solution to the map making problem. His approach was to solve equation (2.5) iteratively using the fact that given a vector  $\mathbf{b}$  and a matrix  $\mathbf{M}$  one can find the vector  $\mathbf{a} = \mathbf{M}^{-1}\mathbf{b}$  iteratively knowing  $\mathbf{M}$  without having to compute  $\mathbf{M}^{-1}$ . This makes map-making using equation (2.5) feasible as the scaling of this is of the order  $N_d \times N_{iter}$  where  $N_{iter}$  is the number of iterations needed to be performed (of the order 10-20).

The next problem to be dealt with is the problem of estimating the noise  $\mathbf{N}$  from the TOD together with the map. I will first briefly review the typical properties of noise in detectors used for CMB experiments. Detector noise is usually a sum of *white noise* which is noise having the same amplitude on all scales along the time stream and *1/f noise* which is having increasing amplitude for lower frequencies. If one represents the noise at position  $j$  in the time stream as  $n_j$ , then the Fourier transform of the noise along a time stream can be written as,

$$n_k = \sum_j n_j e^{2\pi i k j / N_d}. \quad (2.6)$$

The *power spectrum* of the noise is then defined as

$$P(k) = \langle n_k n_k^* \rangle = \sum_{jj'} \langle n_j n_{j'} \rangle e^{2\pi i k / N_d (j-j')}. \quad (2.7)$$

The typical power spectrum represented as a sum of white noise and  $1/f$  noise can be written as (in terms of frequency  $f$ )

$$P(f) = A \left( 1 + \left( \frac{f_{knee}}{f} \right)^\alpha \right), \quad (2.8)$$

where  $A$ ,  $f_{knee}$  and  $\alpha$  are parameters different in each experiment. The frequency  $f_{knee}$  is the frequency where the white noise begins to dominate over  $1/f$  noise.

As is described in (Natoli et al. 2001), the noise correlation matrix  $\mathbf{N}$  can easily be found when knowing the power spectrum and vice versa. If one assumes

that the noise is stationary, i.e. that it has the same properties along the whole time stream, one can write

$$N_{ij} = \langle n_i n_j \rangle = \chi(|i - j|). \quad (2.9)$$

Clearly  $\mathbf{N}$  is a Toeplitz matrix. As discussed in (Natoli, de Gasperis, Gheller, and Vittorio 2001),  $\mathbf{N}$  can be well approximated by a circulant matrix if  $\chi$  vanished for some length  $|i - j| > N_\chi < N_d$ . The Fourier transform of a circulant matrix is diagonal, so one can write

$$\mathbf{N}^{-1} \simeq Q^\dagger \Lambda^{-1} Q, \quad (2.10)$$

where  $\Lambda = \text{Diag}\{P(f_1), \dots, P(f_{N_d/2})\}$ . The dagger denotes complex conjugation and  $Q$  is given by

$$Q_{jk} = \frac{e^{ijf_k}}{\sqrt{2\pi N_d}}. \quad (2.11)$$

(See (Natoli et al. 2001)). This establishes the connection between the power spectrum of the noise and the noise correlation matrix  $\mathbf{N}$ .

I will now discuss the available methods for estimating the map and the noise power spectrum (and thereby  $\mathbf{N}$ ) from the TOD of an experiment. In (Ferreira and Jaffe 2000) an iterative approach was found which for each step in an iterative scheme one finds a solution to equation (2.5) and to the noise power spectrum. Then for each step these two solutions are improved until convergence is reached. As discussed in the paper, the method is not feasible for experiments with a very large number of pixels. In (Prunet, Netterfield, Hivon, and Crill 2000; Dore et al. ) a similar iterative method is presented which is faster and therefore feasible also for big experiments.

Finally another simple and fast method was presented in (Natoli, de Gasperis, Gheller, and Vittorio 2001). I will briefly outline this method. Their approach is basically the approach of (Wright 1996). They first estimate  $\mathbf{N}$  from the time stream and then once  $\mathbf{N}$  is known, equation (2.5) can be solved iteratively using a conjugate gradient method. The noise correlation matrix  $\mathbf{N}$  is found by the use of a *Maximum-Likelihood (ML) technique* which will be described in more detail in section (2.3.1). The method says that given the pure noise data vector  $\mathbf{n}$ , the best estimate of the noise power spectrum  $P(f)$  is given by the  $P(f)$  which minimizes the likelihood

$$L = \frac{e^{-\frac{1}{2}\mathbf{n}^T \mathbf{N}^{-1} (P(f)) \mathbf{n}}}{\sqrt{2\pi \det \mathbf{N}}}. \quad (2.12)$$

Finding a good parameterization of  $P(f)$ , one can easily minimize this function, using the approximation equation (2.10). The only problem now is that one doesn't have the pure noise data stream  $\mathbf{n}$ , only the signal+noise data stream  $\mathbf{d}$ .



It is shown in (Natoli et al. 2001), that using an expression for the noise given by

$$\tilde{n} = \mathbf{d} - \mathbf{P}(\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{d}, \quad (2.13)$$

one gets results in good agreement with the iterative methods. The last term in this equation can be understood as the signal:  $\mathbf{P}^T\mathbf{d}$  makes a map from the TOD,  $(\mathbf{P}^T\mathbf{P})^{-1}$  divides each pixel with the number of values added in the pixel and  $\mathbf{P}$  makes a time stream from the map. Using this as the noise vector one can solve the maximum-likelihood problem to find the noise power spectrum and thereby  $\mathbf{N}$ . When  $\mathbf{N}$  is found the map can be found iteratively as described above.

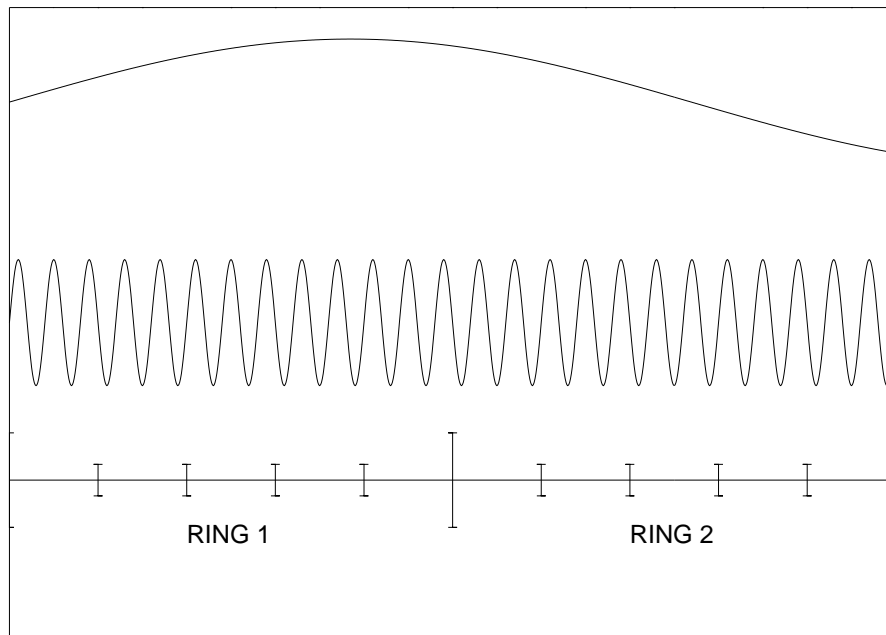


Figure 2.5: The figure illustrates high and low frequency noise (lower and upper sine-curve respectively) and 2 scan rings each scanned 5 times. The purpose of the figure is to show the origin of striping. This is explained in the text.

Another issue strongly related to map-making is the problem of *striping*. Because of  $1/f$  noise the scan path will be visible as stripes in the map. Why this is so is easy to see from figure (2.5). I take the Planck scanning strategy as an example. Planck will be scanning on circles and each circle will be scanned several time before the next circle is scanned. On the figure, the upper cosine

curve shows low frequency noise and the lower shows high frequency noise. The lower line shows 2 rings with 5 scans in each ring. Obviously by averaging all the scans in a ring, the high frequency noise will be averaged out. However the low frequency noise will clearly not. In the figure, the first ring will clearly have a higher noise level than the second. Each ring get a different average noise level because of the low frequency noise. This is visible in the map as stripes along the scanning rings. These offsets can be partially removed (called *destriping*) by comparing the observed temperatures at the points where the rings are crossing. Obviously the more crossing points there are between the rings, the more efficient can the stripes be removed. In (Bersanelli et al. 1996; Delabrouille 1998; Maino et al. 1999) destriping algorithms for the Planck surveyor are discussed.

### 2.2.2 Foregrounds

There are several other sources in the universe apart from the CMB which are emitting radiation on the same wavelengths as the CMB. In order to estimate the CMB power spectrum, one must either observe in areas of the sky without any of these *foregrounds*, or one must somehow remove them from the data. There are 5 main types of foregrounds which must be dealt with in order to study the CMB in contaminated areas of the sky. Galactic dust emission, free-free and synchrotron radiation are contaminants from our own galaxy. In addition there is the radiation from extra galactic point sources and the SZ effect from clusters of galaxies (the thermal AND kinetic) (see section (1.4.3)). Thanks to different spectral behavior of these foregrounds and the CMB, one can in principle separate the CMB from the foreground contaminants. I will now discuss in detail some of these foregrounds. When no other reference is given, the information comes from (Bersanelli et al. 1996; Kogut et al. 1996; Bouchet and Gispert 1999; Hobson, Jones, Lasenby, and Bouchet 1998).

In interstellar space in our own galaxy there are dust grains (consisting of graphite and silicates) heated by the surrounding stars and thereby emitting electromagnetic radiation. The spectrum of the *dust emission* from the galaxy was measured by the FIRAS instrument on the COBE satellite (see section (2.1.1)) with a  $7^\circ$  FWHM beam. Other balloon borne experiments have measured the galactic dust spectrum at higher resolution up to 30 arcminutes. At high galactic latitudes (away from the galactic plane) it has been found that the dust can be described well with a single dust component at 18K with an emissivity which goes as  $\nu^2$  where  $\nu$  is the frequency of radiation. In the direction of the galactic plane, there seems to be another component with a temperature of 21K and spectral dependency  $\nu^{1.4}$ . Maps of galactic dust have been made by DIRBE (another COBE instrument, see section (2.1.1)) at a resolution of 42 arc minutes and by the *Infrared Astronomical Satellite* IRAS at 4 arcminutes. The high resolution images from IRAS has revealed an angular power spectrum of the dust of

$C_\ell \propto \ell^{-3}$ . At high galactic latitudes (outside of the galactic plane  $\theta > 30^\circ$ , dust is the dominating foregrounds component at frequencies above 100GHz. Recently there have been detection of emission from what appears to be spinning dust grains. This emission seems to be dominant at frequencies below 30GHz.

Another galactic foregrounds component operating at lower frequencies is the free-free emission from ionized hydrogen  $H2$ . The spectrum of the free-free emission is well known to be  $\nu^{-0.16}$ , but good maps of interstellar  $H2$  is lacking. Attempts to map the free-free emission directly was made using data from the COBE DMR instrument (Bennett et al. 1992; Bennett et al. 1994). Unfortunately the data was so noisy that only the quadrupole could be measured. Other attempts to map  $H2$  has been observation of  $H_\alpha$  emission (Reynolds 1984; Reynolds 1992) but these experiments suffered from undersampling and selection biases. For this reason, a correlation which is detected between free-free emission and dust emission is used. Dust is correlated with neutral hydrogen  $H1$  which is again correlated with ionized hydrogen  $H2$ . The dust-free-free correlation has been detected in correlations of data between the DMR and DIRBE instruments and between data from the Saskatoon (Tegmark et al. 1997) experiment and DIRBE at smaller angular scales (Oliveira-Costa et al. 1997). For this reason, the same maps used to map dust (IRAS/DIRBE) are also used for free-free emission. The angular power spectrum of free-free emission is assumed to be the same as for dust. Outside of the galactic plane, free-free emission is the dominant galactic foreground contaminant in the frequency range 5GHz to 100GHz.

The third galactic source of microwave radiation is synchrotron emission resulting from the acceleration of cosmic ray electrons in the galactic magnetic field. Due to the varying magnetic field strength in the galaxy, the spectral index of synchrotron emission is varying. Observations at 408GHz show that the spectrum of synchrotron emission goes as  $\nu^\beta$  where  $\beta$  is between  $-2.7$  and  $-3.1$ . Observations at higher frequencies indicate an index of  $\beta = -0.9$  for the frequency range of interest for CMB experiments. For synchrotron emission, the maps used are the 408MHz map by (Haslam et al. 1981) and the 1420MHz map by (Reich and Reich 1988). For CMB experiments, the data of synchrotron emission in these maps are extrapolated to the higher frequencies. These maps have an angular resolution of  $0.85^\circ$  and  $0.6^\circ$  FWHM respectively. There is no data available at higher resolution. As these maps seem to indicate that the power spectrum falls of as  $C_\ell \propto \ell^{-3}$ , this is the assumption usually adopted. This is the same power spectrum observed for dust at small scales. Apparently synchrotron emission is the galactic contaminant for which one has the least information. Synchrotron radiation is together with free-free important at frequencies below 100GHz and is dominant below 5GHz.

The CMB maps will also be contaminated by extra galactic point sources.

These can be AGNs (Active galactic nuclear), radio galaxies, quasars or BL Lac objects in the radio domain of the spectrum. In the Far-IR (Far-Infrared) domain, the dust emission from dust dominated infrared galaxies is present. The spectrum of these objects is dependent on redshift. Unfortunately the spectrum and the evolution of these objects are not well known. In (Toffolatti et al. 1998), a last update on the observations and evolutionary models is found. They conclude that on Planck resolution, the radio sources is the dominating point source contaminant on frequencies  $\nu < 100\text{GHz}$ , the Far-IR sources are dominating at  $\nu > 200\text{GHz}$  and in the intermediate range both types are comparable. They do however point out, that even in the most pessimistic models, the amplitude of fluctuations due to point sources is well below the amplitude of the CMB at the frequencies  $100 - 200\text{GHz}$ . In (Hobson et al. 1999), it is shown how these models can be adopted to remove point sources in CMB data.

Another source of extra galactic point sources is the SZ-effect from clusters of galaxies. The change of the CMB spectrum due to the thermal motion of electrons in the clusters (thermal SZ effect) and the bulk motion of the cluster (kinematic SZ effect) is also contaminating the underlying CMB. Fortunately, the SZ effect has a certain spectral signature that makes it easy to identify. According to (Zeldovich and Sunyaev 1969), the relative temperature change of the CMB due to the SZ effect is given by

$$\frac{\Delta T}{T} = y \left( x \coth \frac{x}{2} - 4 \right), \quad (2.14)$$

where  $x = h\nu/kT$ ,  $h$  is Planck's constant,  $k$  is Boltzmann's constant and  $y$  is the Compton  $y$  parameter measuring the line of sight density of electrons given as

$$y = \int_0^s \frac{kT_e}{mc^2} ds, \quad (2.15)$$

where  $T_e$  is the electron temperature,  $m$  is the electron mass and  $s$  is the Thompson optical depth  $s = \int \sigma_T \rho_e dl$ . Here  $\sigma_T$  is the Thompson cross section and  $\rho_e$  is the electron density. Figure (2.6) shows the spectral signature.

For the kinetic SZ effect, the spectrum is constant (Sunyaev and Zeldovich 1980)

$$\frac{\Delta T}{T} = \frac{v_r}{c} s, \quad (2.16)$$

where  $v_r$  is the radial velocity of the cluster. The kinematic effect is typically an order of magnitude lower than the thermal effect.

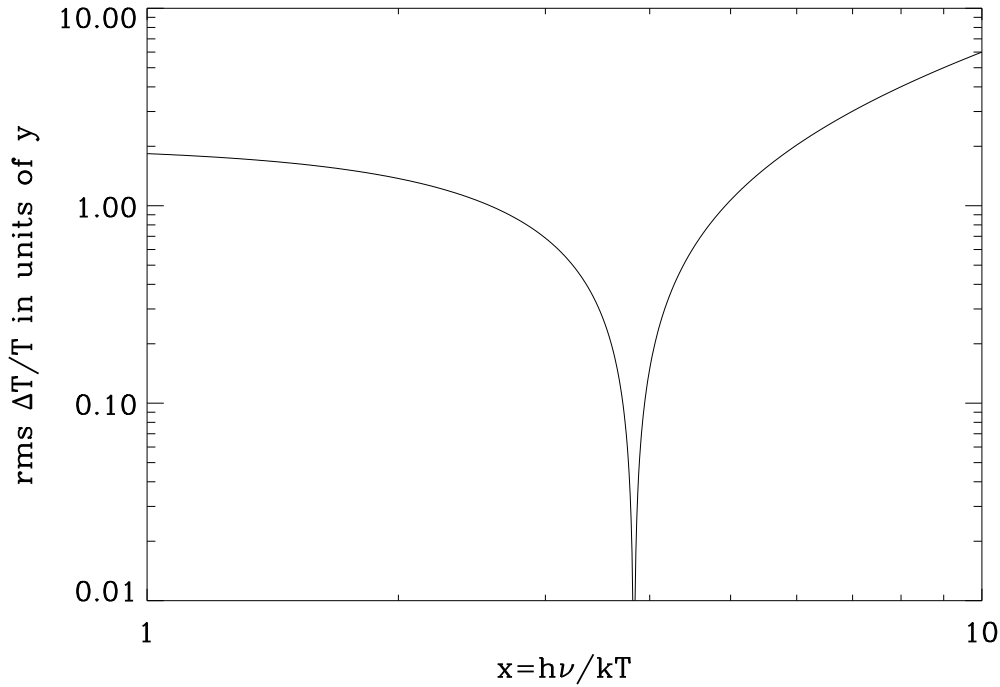


Figure 2.6: The spectral change in the CMB due to the thermal SZ effect.

Finally I will review some of the standard method of separating the different components of the CMB. One can define a function  $f(\hat{\mathbf{n}}, \nu)$  which is the intensity in the direction  $\hat{\mathbf{n}}$  at frequency  $\nu$ . This intensity consists of contributions from different components like the CMB, galactic emission or the SZ effect. Assuming that the different components can be factorised into a spatial part  $x(\hat{\mathbf{n}})$  and a spectral part  $s(\nu)$  one can write the intensity as

$$f(\hat{\mathbf{n}}, \nu) = \sum_{j=1}^{n_p} s_j(\nu) x_j(\hat{\mathbf{n}}), \quad (2.17)$$

where the sum goes over the  $n_p$  different components. For an experiment with  $n_c$  frequency channels one can assign the observation in the  $i$ th channel to the  $i$ th element of a vector  $\mathbf{y}$  given as

$$\mathbf{y}(\hat{\mathbf{n}}) = \mathbf{P}\mathbf{x}(\hat{\mathbf{n}}) + \mathbf{n}(\hat{\mathbf{n}}). \quad (2.18)$$

Here  $\mathbf{x}(\hat{\mathbf{n}})$  is an  $n_p$  element vector, the elements being the  $x_j(\hat{\mathbf{n}})$  above for each channel,  $\mathbf{n}(\hat{\mathbf{n}})$  is an  $n_p$  element vector containing the noise in each channel and  $\mathbf{P}$  is a  $n_c \times n_p$  matrix assigning contributions from different components to different

channels. The elements of  $\mathbf{P}$  are given by

$$P_{ij} = \int_0^\infty t_i(\nu) s_j(\nu) d\nu, \quad (2.19)$$

here  $t_i(\nu)$  is the spectral transmission of channel  $i$ .

The problem of component separation now becomes the problem of finding the vector  $\mathbf{x}$  given the observations  $\mathbf{y}$  and general information about the different components. In other words, one is looking for a  $n_p \times n_c$  matrix  $\mathbf{W}$  satisfying

$$\mathbf{x}(\hat{\mathbf{n}}) = \mathbf{W}\mathbf{y}(\hat{\mathbf{n}}). \quad (2.20)$$

There have been different approaches to find the matrix  $\mathbf{W}$ . One of the standard ones has been Wiener filtering. In this method one tries to minimize the quantity

$$\epsilon^2 = \langle |\mathbf{W}\mathbf{y} - \mathbf{x}|^2 \rangle. \quad (2.21)$$

As this technique is time consuming in pixel space, one normally transforms all components to spherical harmonic space for the Wiener filtering. The method is described in more detail in e.g. (Bersanelli et al. 1996; Tegmark and Efstathiou 1996; Prunet et al. 2001). Another method is the Maximum Entropy Method (MEM) (Hobson, Jones, Lasenby, and Bouchet 1998; Stolyarov, Hobson, Ashdown, and Lasenby 2001). These methods use Bayes theorem (this is described in more detail in section (2.3.1)) to maximize the posterior probability (the probability of the theory given the data) expressed in terms of the likelihood (the probability of the data given a theory) and an entropic prior. In this case the Gaussian likelihood can be written using equation (2.18)

$$L \propto e^{-\frac{1}{2}(\mathbf{y}-\mathbf{P}\mathbf{x})^\dagger \mathbf{N}^{-1}(\mathbf{y}-\mathbf{P}\mathbf{x})}, \quad (2.22)$$

where  $\mathbf{N} = \langle \mathbf{n}^T \mathbf{n} \rangle$  is the noise correlation matrix. Other methods include wavelet techniques (Tenorio 1999), neural networks (Baccigalupi et al. 2000) and Fast Independent Component Analysis (FastICA) (Maino et al. 2001). All these techniques rely on good spectral and spatial information about the foregrounds which as described above, is not always available.

## 2.3 Power Spectrum Estimation

In this section I will describe how the angular power spectrum of the CMB can be extracted from the maps where the foregrounds are removed. I will discuss the standard approach which is the maximum-likelihood method. I will describe why this method is not feasible for large data sets and why new and faster methods are needed. I will describe some attempts to find fast power spectrum estimation methods. The rest of this thesis will discuss new methods which should make the power spectrum analysis of huge data sets feasible.

### 2.3.1 Likelihood Estimation

The standard method for analysing CMB data is the *maximum-likelihood method*. This is the method which gives the smallest error bars. I will now outline the method in general and for CMB analysis. For more information, see e.g. (Parratt 1961).

The maximum-likelihood method (MLM) is based on *Bayes Theorem*. If one considers two events  $A$  and  $B$  which have the probabilities  $P(A)$  and  $P(B)$  respectively for occurring, then Bayes theorem connect the probability for  $A$  to occur given that  $B$  has occurred  $P(A|B)$  and the probability for  $B$  to occur given that  $A$  has occurred  $P(B|A)$ . Bayes theorem can in this case be written as

$$P(A|B) = P(A) \frac{P(B|A)}{P(B)}. \quad (2.23)$$

Bayes theorem is simply another way of writing the product theorem for probabilities,

$$P(A, B) = P(A)P(B|A), \quad (2.24)$$

which says that the probability for both  $A$  and  $B$  to occur, is just the product of the probability for the occurrence of  $A$  times the probability for the occurrence of  $B$  given that  $A$  has already happened. In data analysis one is interested in the probability for a hypothesis  $H$  to be true given the events (data)  $d$  from an experiment,  $P(H|d)$ . With these parameters Bayes theorem can be written

$$P(H|d) = P(d|H) \frac{P(H)}{P(d)}. \quad (2.25)$$

The function  $P(H|d)$  is called the *posterior probability* since it expresses the belief in a hypothesis *after* the data from an experiment has been considered. The function  $P(H)$  is called the *prior* since it expresses the prior belief in the hypothesis before the experiment was done. The function  $P(d|H)$  is called the *likelihood* and is the probability of getting the data  $d$  given that the hypothesis  $H$  is true. Finally, the factor  $P(d)$  is independent of the hypothesis and is just a normalisation factor.

When an experiment has been done which has given the data set  $d$ , one would like to find which hypothesis  $H$  of several hypothesis has the best fit to the data. Expressed differently, one would like to find the hypothesis  $H$  for which the probability of the hypothesis given the data  $P(H|d)$  is highest. Looking at equation (2.25) this means that one has to calculate the product of the likelihood and the prior for the different hypothesis available and find the hypothesis for which this product is highest. This is the method of Maximum likelihood (ML).

As an example I will use an experiment where one has measured a set of numbers given as a vector  $\mathbf{x}$ . The hypothesis is that this set of numbers has a multivariate Gaussian distribution. The mean  $\langle \mathbf{x} \rangle$  and correlation matrix  $\mathbf{C} = \langle \mathbf{x} \mathbf{x}^T \rangle$  are given by the hypothesis as a function of some parameters  $\alpha$  which one wants to estimate from the data. I will use a flat (constant) prior (as in the rest of this thesis) meaning that there is no a priori values of  $\alpha$  which are known to be better or more probable than others. In this case one has to maximize the likelihood given as a multivariate Gaussian

$$L(\alpha) = \frac{e^{-\frac{1}{2}(\mathbf{x} - \langle \mathbf{x} \rangle(\alpha))^T \mathbf{C}^{-1}(\alpha)(\mathbf{x} - \langle \mathbf{x} \rangle(\alpha))}}{\sqrt{2\pi \det \mathbf{C}}}. \quad (2.26)$$

It often simplifies to take the logarithm of the likelihood and *minimize* the quantity

$$\mathcal{L}(\alpha) = -2 \log L(\alpha) = (\mathbf{x} - \langle \mathbf{x} \rangle(\alpha))^T \mathbf{C}^{-1}(\alpha)(\mathbf{x} - \langle \mathbf{x} \rangle(\alpha)) + \log \det \mathbf{C}. \quad (2.27)$$

This quantity can be minimized using different available computer algorithms. Which algorithm to use depends on whether one has knowledge about the derivatives of the likelihood. If one knows the first (and possibly also the second) derivative, the minimization procedure will be quicker.

Some words have to be said about the estimate  $\hat{\alpha}$  of  $\alpha$  that one gets from the MLM. As discussed in (Tegmark, Taylor, and Heavens 1997),  $\hat{\alpha}$  is the best unbiased estimator of  $\alpha$ . It is unbiased in the sense that  $\hat{\alpha}$  goes asymptotically to the true value of  $\alpha$  when many likelihood estimations of  $\alpha$  from different experiments are made. It is the best in the sense that no other method gives smaller error bars. The *Cramer-Rao inequality* says that

$$\Delta\alpha_i \geq \sqrt{(\mathbf{F}^{-1})_{ii}}. \quad (2.28)$$

Here  $\Delta\alpha_i = \sqrt{(\hat{\alpha}_i - \langle \hat{\alpha}_i \rangle)^2}$  and  $\mathbf{F}$  is given by the *Fisher information matrix*

$$F_{ij} = \frac{1}{2} \left\langle \frac{\partial \mathcal{L}}{\partial \alpha_i \partial \alpha_j} \right\rangle. \quad (2.29)$$

For the ML estimator, the Cramer-Rao inequality is an equality and for this reason no other method can have smaller error bars.

For CMB analysis, likelihood analysis is very attractive for the reasons given above. It gives an unbiased estimate of the parameters one is estimating for and it gives the smallest error bars. Also, the objects that one gets from the experiments, the sky temperature in a given pixel or the Spherical Harmonic transform  $a_{\ell m}$  are (with Gaussian initial conditions assumed) Gaussian distributed so a simple likelihood of the form given in equation (2.26) can be used. The parameters  $\alpha$



that one can estimate for could be the power spectrum coefficients  $C_\ell$  or directly the cosmological parameters  $\Omega_0, \Omega_b, h$  etc.

I will now assume that I have a sky map from a full sky experiment with foregrounds removed. The aim is to find the best estimate of the power spectrum  $C_\ell$ . The datavector for the likelihood will now be the pixels of the CMB map arranged into a datavector  $\mathbf{d} = \mathbf{s} + \mathbf{n}$  where  $\mathbf{s}$  is the CMB signal and  $\mathbf{n}$  is the noise. The monopole (average) has been subtracted from the map so that  $\langle \mathbf{d} \rangle = 0$ . Before one can formulate the likelihood (equation 2.26), one also needs the correlation matrix  $\mathbf{C} = \langle \mathbf{d}^T \mathbf{d} \rangle = \mathbf{S} + \mathbf{N}$  (where  $\mathbf{S} = \langle \mathbf{s}^T \mathbf{s} \rangle$  and  $\mathbf{N} = \langle \mathbf{n}^T \mathbf{n} \rangle$ ) expressed through the parameters  $C_\ell$  for which one wants to find an estimate. This can be done using the relation

$$T(\hat{\mathbf{n}}_i) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}_i), \quad (2.30)$$

where  $\hat{\mathbf{n}}_i$  is the unit vector in the direction of pixel  $i$ . One then gets for  $S_{ij}$

$$\langle d_i d_j \rangle = \langle T(\hat{\mathbf{n}}_i) T(\hat{\mathbf{n}}_j) \rangle, \quad (2.31)$$

$$= \sum_{\ell m} \sum_{\ell' m'} \langle a_{\ell m} a_{\ell' m'} \rangle Y_{\ell m}(\hat{\mathbf{n}}_i) Y_{\ell' m'}(\hat{\mathbf{n}}_j), \quad (2.32)$$

$$= \sum_{\ell} C_\ell \sum_m Y_{\ell m}(\hat{\mathbf{n}}_i) Y_{\ell m}(\hat{\mathbf{n}}_j), \quad (2.33)$$

$$= \sum_{\ell} \frac{2\ell + 1}{4\pi} C_\ell P_\ell(\cos(\Delta\theta_{ij})), \quad (2.34)$$

where  $\Delta\theta_{ij}$  is the angular distance between pixel  $i$  and  $j$ . The following two relations were used

$$\langle a_{\ell m} a_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'}, \quad (2.35)$$

as discussed in section (1.4.2) and

$$\sum_m Y_{\ell m}(\hat{\mathbf{n}}_i) Y_{\ell m}(\hat{\mathbf{n}}_j) = \frac{2\ell + 1}{4\pi} P_\ell(\cos(\Delta\theta_{ij})). \quad (2.36)$$

How the noise correlation matrix  $\mathbf{N}$  can be found was discussed in section (2.2.1). For a given power spectrum  $C_\ell$  one can now calculate the log-likelihood

$$\mathcal{L}(C_\ell) = \mathbf{d}^T \mathbf{C}^{-1} \mathbf{d} + \log \det \mathbf{C}. \quad (2.37)$$

The MLM method now says that one shall minimize this likelihood with respect to  $C_\ell$ . The set of parameters  $\hat{C}_\ell$  which minimizes equation (2.37) is the ML estimate.

This was the ideal case. The first problem with the MLM is the size of the data set. The MAP and Planck satellites will produce sky maps with  $10^6 - 10^7$  pixels. This means that the correlation matrix  $\mathbf{C}$  has say  $10^7 \times 10^7$  elements.

Finding the inverse and determinant of a  $N \times N$  matrix takes of the order  $N^3$  operations. On current computers, these operations take of the order a few seconds for  $N = 1000$ . To make a very optimistic estimate, lets say the the operations take 1sec for  $N = 1000$ . then for  $N = 10^7$  this will take  $10^{12}$ sec  $\approx 32000$ years.

The problem of finding the inverse and determinant of the correlation matrix could be solved if the matrix had some symmetries making it faster to invert. One possibility to explore would be to go to spherical harmonic space and use the coefficients  $a_{\ell m}$  in the datavector instead of the pixels in the map. In theories with Gaussian initial conditions, these  $a_{\ell m}$  are also Gaussian distributed with a zero mean. With such a datavector, the correlation matrix would have the form  $\mathbf{S}_{\ell m, \ell' m'} = \langle a_{\ell m} a_{\ell' m'} \rangle = C_{\ell} \delta_{\ell \ell'} \delta_{m m'}$ . In this case the correlation matrix would be diagonal and the inversion would take  $N$  operations. Unfortunately the noise correlation matrix does not simplify in spherical harmonic space and so the problem of doing a complete maximum likelihood analysis remains unsolved.

Another problem with the likelihood formalism arises when only parts of the sky is covered. From the discussion about foregrounds it should be clear that it can be difficult to remove all the foreground contaminants from the map. In this case, parts of the sky has to be cut out. When the integration is no longer done over the whole sky, the orthogonality of spherical harmonics is destroyed and the  $a_{\ell m}$  on the cut sky are no longer orthogonal, i.e. the correlation matrix in spherical harmonic space is no longer diagonal. In (Górski 1994) a method to orthogonalize the spherical harmonics on the cut sky is presented. I will now briefly outline the method.

Following (Górski 1994), one can define a scalar product of two functions  $f_i$  and  $g_i$  on a pixelised sphere as

$$\langle fg \rangle_{(fullsky)} = \Omega_{pix} \sum_i f_i g_i. \quad (2.38)$$

Defining a vector  $\mathbf{y}$  which elements are the spherical harmonic functions  $Y_{\ell m}$  for each pair  $\ell m$ , one gets

$$\langle \mathbf{y} \mathbf{y}^T \rangle_{(fullsky)} \approx I \quad (2.39)$$

by the orthogonality of spherical harmonics (this is not an identity here since a pixelised sphere is considered). On the cut sky however, one gets

$$\langle \mathbf{y} \mathbf{y}^T \rangle_{(cutsky)} = W \quad (2.40)$$

By defining a new set of harmonics  $\psi$  defined as

$$\psi = \mathbf{\Gamma} \mathbf{y}, \quad (2.41)$$

where  $\mathbf{\Gamma}$  is the inverse  $\mathbf{L}^{-1}$  of the factors of the Cholesky decomposition of the coupling matrix  $\mathbf{W} = \mathbf{L}\mathbf{L}^T$ . With these new harmonics one can now define a new set of harmonic coefficients. Letting  $\mathbf{a}$  be a vector of  $a_{\ell m}$  components, one can define the vector  $\mathbf{c}$  of new coefficients as

$$\mathbf{c} = \mathbf{L}^T \mathbf{a}, \quad (2.42)$$

which correlation matrix  $\langle \mathbf{c}\mathbf{c}^T \rangle$  is again diagonal.

Estimating the power spectrum using the full MLM is clearly not feasible so other approximate methods have to be sought. One method was developed for the MAP experiment by (Oh, Spergel, and Hinshaw 1999). By assuming uncorrelated axis symmetric noise, they estimate the maximum of the likelihood using conjugate gradient techniques. This takes of the order  $N_{pix}^2$  operations instead of the  $N_{pix}^3$  required for the exact likelihood solution. Their starting point is the second order Taylor expansion of the log-likelihood around the minimum at the estimate  $\hat{C}_\ell$

$$\mathcal{L} \approx \mathcal{L}[\hat{C}_\ell] + \sum_\ell \frac{\partial \mathcal{L}}{\partial C_\ell} [\hat{C}_\ell] (C_\ell - \hat{C}_\ell) + \sum_{\ell\ell'} \frac{\partial^2 \mathcal{L}}{\partial C_\ell \partial C_{\ell'}} [\hat{C}_\ell] (C_\ell - \hat{C}_\ell) (C_{\ell'} - \hat{C}_{\ell'}), \quad (2.43)$$

where  $[\hat{C}_\ell]$  means that the function is supposed to be taken at the minimum where  $C_\ell = \hat{C}_\ell$  for all  $\ell$ . The derivatives of the log-likelihood are

$$\frac{\partial \mathcal{L}}{\partial C_\ell} = -\mathbf{d}^T \mathbf{C}^{-1} \mathbf{P}^\ell \mathbf{C}^{-1} \mathbf{d} + Tr(\mathbf{C}^{-1} \mathbf{P}^\ell), \quad (2.44)$$

$$\frac{\partial^2 \mathcal{L}}{\partial C_\ell \partial C_{\ell'}} = 2\mathbf{d}^T \mathbf{C}^{-1} \mathbf{P}^\ell \mathbf{C}^{-1} \mathbf{P}^{\ell'} \mathbf{C}^{-1} - Tr(\mathbf{C}^{-1} \mathbf{P}^\ell \mathbf{C}^{-1} \mathbf{P}^{\ell'}), \quad (2.45)$$

$$\mathbf{P}^\ell = \frac{\partial \mathbf{C}}{\partial C_\ell}. \quad (2.46)$$

The Fisher matrix (equation (2.29)) is the half of the expectation value of this second derivative

$$F_{\ell\ell'} = \frac{1}{2} Tr(\mathbf{C}^{-1} \mathbf{P}^\ell \mathbf{C}^{-1} \mathbf{P}^{\ell'}). \quad (2.47)$$

The approach in (Oh, Spergel, and Hinshaw 1999) is to minimize the log-likelihood by iteratively finding the zero point of the first derivative, using the second derivative approximated by its expectation value given by the fisher matrix. Calculating the first and second derivatives of the log-likelihood consist of inverting the correlation matrix  $\mathbf{C}$  which needs to be avoided as discussed above. They solve this problem by means of the following techniques

- They find the term  $\mathbf{C}^{-1} \mathbf{d}$  iteratively using conjugate gradient techniques which do not require the inversion of  $\mathbf{C}$ . However a good guess for  $\mathbf{C}^{-1}$  is needed for the iteration to converge. They find this by inverting a simplified form of the correlation matrix.

- The trace  $Tr(\mathbf{C}^{-1}\mathbf{P}^\ell)$  is found using the fact that

$$\langle \mathbf{d}^T \mathbf{C}^{-1} \mathbf{P}^\ell \mathbf{C}^{-1} \mathbf{d} \rangle = Tr(\mathbf{C}^{-1} \mathbf{P}^\ell). \quad (2.48)$$

In this way, the trace term can be calculated using Monte Carlo simulations.

- The fisher matrix (which is not needed exactly in order to converge to the minimum) is approximated by using the preconditioner for  $\mathbf{C}^{-1}$  in the first step.

Another way to increase the speed of the likelihood estimation is by means of approximate likelihood expressions (Bond, Jaffe, and Knox 2000; Bartlett, Douspis, Blanchard, and Dour 2000). In (Bond, Jaffe, and Knox 2000) this is motivated by the fact that the likelihood is not a Gaussian function of the power spectrum  $C_\ell$ . For the very simplified scenario with full sky coverage and uniform white noise the likelihood can be written in terms of the coefficients  $C_\ell = \ell(\ell + 1)C_\ell/(2\pi)$  as

$$\mathcal{L} = \sum_\ell (2\ell + 1) \left( \ln(C_\ell B_\ell^2 + \mathcal{N}_\ell) + \frac{\tilde{\mathcal{C}}_\ell}{C_\ell B_\ell^2 + \mathcal{N}_\ell} \right), \quad (2.49)$$

where  $B_\ell$  accounts for the beam,  $\mathcal{N}_\ell$  is the noise power spectrum defined in the same way as  $C_\ell$  and  $\tilde{\mathcal{C}}_\ell$  is the measured power spectrum. As expected this is not Gaussian in  $C_\ell$  which is clearly seen by the fact that the second derivative is not a constant in this quantity. In (Bond, Jaffe, and Knox 2000) they construct a new quantity  $Z_\ell = \ln(C_\ell + x_\ell)$  with  $x_\ell = \mathcal{N}_\ell/B_\ell^2$  which is approximated to be normally distributed. The approximate likelihood then becomes

$$\mathcal{L} = \sum_{\ell\ell'} Z_\ell M_{\ell\ell'}^{(Z)} Z_{\ell'}, \quad (2.50)$$

where  $M_{\ell\ell'}^{(Z)} = (C_\ell + x_\ell) M_{\ell\ell'}^{(C)} (C_{\ell'} + x_{\ell'})$  and  $M^C$  is the inverse covariance matrix. The idea is now to keep this simplified form of the likelihood for a more complicated case with partial sky coverage and non-uniform noise. In this case the parameter  $x_\ell$  will have to be changed. Unfortunately the calculation of  $x_\ell$  will in the general case scale as  $N_{pix}^3$ .

A similar approximation was used by (Wandelt, Hivon, and Górski 2000). In that approach, the so called pseudo- $C_\ell$  ( $\tilde{C}_\ell$ ) were approximated to be Gaussian and uncorrelated and the corresponding likelihood-ansatz was used. This method will be described in more detail and extended in chapter (4).

### 2.3.2 Quadratic Estimators

Some authors have avoided the ML estimation and instead attempted to use *quadratic estimators*. Quadratic estimators mean estimators of the power spectrum  $C_\ell$  which are quadratic in the data. In (Bond, Jaffe, and Knox 1998),

the Taylor expansion of the likelihood was used to make a quadratic estimator which is calculated using iterations. Starting with an initial guess for the power spectrum, the correction at each step can be calculated using

$$\delta C_\ell = \sum_{\ell'} \frac{1}{2} F_{\ell\ell'}^{-1} \text{Tr} \left[ (\mathbf{d}\mathbf{d}^T - \mathbf{C})(\mathbf{C}^{-1} \mathbf{P}^\ell \mathbf{C}^{-1}) \right], \quad (2.51)$$

where  $F_{\ell\ell'}$  is the Fisher matrix. As this is a Taylor expansion about the maximum-likelihood solution, the expression becomes more and more correct, the closer one is to the solution. For this reason the iteration converges to the maximum-likelihood solution. But clearly  $N_{pix}^3$  operations are needed to calculate this quantity. A similar quadratic estimator with optimal error bars was found by (Tegmark 1996; Tegmark 1997b).

In (Doré, Knox, and Peel 2001) the quadratic method (equation (2.51)) was extended to yield a faster power spectrum estimation method. Their method involves splitting the map into several non-overlapping pieces having a size for which power spectrum estimation using the quadratic estimator (equation 2.51) is feasible. To obtain the power spectrum for large angular scales, the map is coarsened and the quadratic estimator is used on the coarsened map. In the end, all the different estimates are averaged in an optimal way.

Another quadratic estimator using correlation functions was suggested by (Szapudi et al. 2001) and further developed by (Szapudi, Prunet, and Colombi 2001). The idea is to calculate an unbiased quadratic estimator of the pixel-pixel correlation function in bins

$$\tilde{\xi}(\cos \theta) = \sum_{ij} f_{ij} (d_i d_j - N_{ij}), \quad (2.52)$$

where  $f_{ij}$  determines which pixel pairs belong to the bin (can also be used for noise weighting). Then the power spectrum is integrated from the correlation function. The error bars in the method are found using Monte Carlo simulations. The error bars are found to be up to 10% larger than the optimal ML error bars. The method has been tested on uncorrelated noise, but it should in principle also work with noise correlations present.

Finally an unbiased quadratic estimator based on pseudo  $C_\ell$  was presented in (Hivon et al. 2001). In this paper, the relation between  $\tilde{C}_\ell$  and the full sky  $C_\ell$  was derived and inverted giving an equation for  $C_\ell$  expressed by the  $\tilde{C}_\ell$ . These relations will be discussed further in chapter (4). In this method, the noise correlation matrix and the variance of the estimates was computed using Monte Carlo simulations. The method was tested on an asymmetric patch on the sky with correlated noise and the error bars turned out to be almost optimal.

### 2.3.3 Two New Methods for Power Spectrum Estimation

This thesis describes two new methods for fast power spectrum estimation. The main goal of a CMB experiment is usually the estimation of cosmological parameters, but the estimation of parameters from a CMB map is usually done with an intermediate step where the angular power spectrum is estimated. As discussed in (Bond, Jaffe, and Knox 1998; Bond, Efstathiou, and Tegmark 1997) it has several advantages to estimate the cosmological parameters from the power spectrum rather than directly from the map. A near-degeneracy of cosmological parameters makes it harder for search algorithms to find the minimum of the likelihood in cosmological parameter space. The power spectrum space is much simpler as the different  $C_\ell$ s are only weakly dependent on each other. This makes it significantly simpler to search for the likelihood minimum in  $C_\ell$  space. Once the power spectrum is found, it is now easier to find the log-likelihood minimum of the cosmological parameters with the power spectrum as input data. Also, the joint analysis of data from several CMB experiments is easier using the power spectra from each experiment instead of the maps.

In chapter (3) a new method for extracting the power spectrum directly from the time stream of an experiment without doing the intermediate map-making step will be discussed. This method is developed for an ideal experiment (stationary noise, no foregrounds) which is scanning on rings. It is an exact likelihood maximization algorithm which takes advantage of the signal and noise symmetries in experiments like MAP and Planck which are scanning on rings. The scaling is  $N^2$  where  $N$  is the number of elements in the time stream (after averaging all the scans in each scan ring reducing the length of the time stream significantly). One of the advantages with the method is that it takes a general non symmetric beam pattern and correlated noise into account at no extra cost.

In chapter (4) another method which is an extension of the method in (Wandelt, Hivon, and Górski 2000) will be discussed. The pseudo- $C_\ell$ s from a map which is multiplied with a general axis-symmetric window is used as data in a ML estimation. The chapter describes the extension of *Gabor Transforms* to the sphere. Gabor Transforms or Windowed Fourier Transforms are just Fourier Transforms of data which has been multiplied with a window function to smooth the data at the edges or to increase  $S/N$ . The extension of the method to analyse several patches on the sky with different windows is discussed. The method is fast and will in principle allow analysis of huge data sets as those coming from MAP and Planck. In chapter (5) the method is extended to polarisation.

# Chapter 3

## Fast Exact Power Spectrum Analysis for a Special Type of Scanning Strategies

In this chapter I will present work done under supervision of Ben Wandelt. A new method to estimate the power spectrum from a CMB experiment using the maximum-likelihood method directly on the data time stream instead of the sky map will be presented. The method works for experiments scanning on rings (like MAP, Planck and Archeops) where the data time stream has symmetries which makes exact likelihood analysis possible in  $\mathcal{O}(N^2)$  operations,  $N$  being the number of elements in the time stream.

The method uses a technique presented in (Wandelt and Górski 2000) for fast convolution of two functions on the sphere. For this reason I will make a summary of this paper in section (3.1). Some of the techniques and notation in this paper is used in section (3.2) for the power spectrum estimation. In the paper it is described how one can simulate the time stream from an experiment with an asymmetric beam profile and far side lobes. The task of convolving a general beam profile with a given sky is speeded up by doing the convolution in spherical harmonic space.

For experiments which are scanning on rings, the correlation between rings will only depend on the distance between the rings. This is true both for noise and signal, making the total correlation matrix for the time stream block circulant (Wandelt 2000). In Fourier space, such a matrix is block diagonal allowing fast inversion of the matrix. In section (3.2) I will describe the method of extracting the angular power spectrum with the MLM exploiting this symmetry of the correlation matrix. This method naturally deals with general beam patterns and side lobes which so far no other power spectrum estimation method can handle. An example will be presented. The method and its results are published in

(Wandelt and Hansen 2000)

### 3.1 Fast Fourier Space Convolution

This section is a summary of the paper (Wandelt and Górski 2000). The results and definitions here will be used in the next section. In (Wandelt and Górski 2000) a fast method to make simulated time streams of a general CMB experiment is presented. To test the different steps of the data analysis pipeline of a CMB experiment one needs to simulate the experiment. Generally when the CMB telescope is looking in a certain direction, it does not only observe the CMB temperature in that given direction. There are also contributions from other directions. Usually one approximates this beam pattern to be a 2D Gaussian centered in the direction in which the telescope is pointing. In general however, this beam pattern will be different and often not axisymmetric causing distortions of the CMB patterns. In addition, if the instrument allows diffraction of stray light into the detectors, the observed CMB temperature in one direction might have contributions from completely other directions. Because of high intensity radiation from the galactic plane (see section (2.2.2)) or reflected radiation from planets in the solar system this can contaminate the CMB data and should be taken into account in the data analysis.

In the following, rotations and directions on the sky will be given by the Euler angles  $(\Phi_2, \Theta, \Phi_1)$ . These angles describe a right handed rotation in a fixed coordinate system by the angles  $\Phi_1$ ,  $\Theta$  and  $\Phi_2$  about the axis  $z$ ,  $y$  and  $z$  respectively in that order. As an example one can consider a point at the north pole. After a rotation  $(0^\circ, -90^\circ, 90^\circ)$  the point will be in the direction of the  $y$  axis.

Following (Wandelt and Górski 2000), one can consider two functions on the sphere  $b(\hat{\mathbf{n}})$  and  $s(\hat{\mathbf{n}})$  which can be the beam profile (with all contributions including sidelobes) and the sky of a CMB experiment. The observed temperature  $T(\Phi_2, \Theta, \Phi_1)$  in the direction that a beam centered at the north pole will have after the rotation  $(\Phi_2, \Theta, \Phi_1)$  can be written

$$T(\Phi_2, \Theta, \Phi_1) = \int d\hat{\mathbf{n}} \left[ \hat{D}(\Phi_2, \Theta, \Phi_1) b \right] (\hat{\mathbf{n}}) s(\hat{\mathbf{n}}). \quad (3.1)$$

This is a convolution of the sky with the beam rotated into the correct position. The rotation operator  $\hat{D}$  is described in appendix (A) and rotates the function with the given Euler angles. So given a sky and a beam profile centered at the north pole, this integral gives the observed temperature in a given direction. In a CMB experiment the beam will take a certain path on the sky as a function of time  $t$ . So the direction in which the beam is pointing is given as a function of time as  $(\Phi_2(t), \Theta(t), \Phi_1(t))$ . This is called a *scan path*.



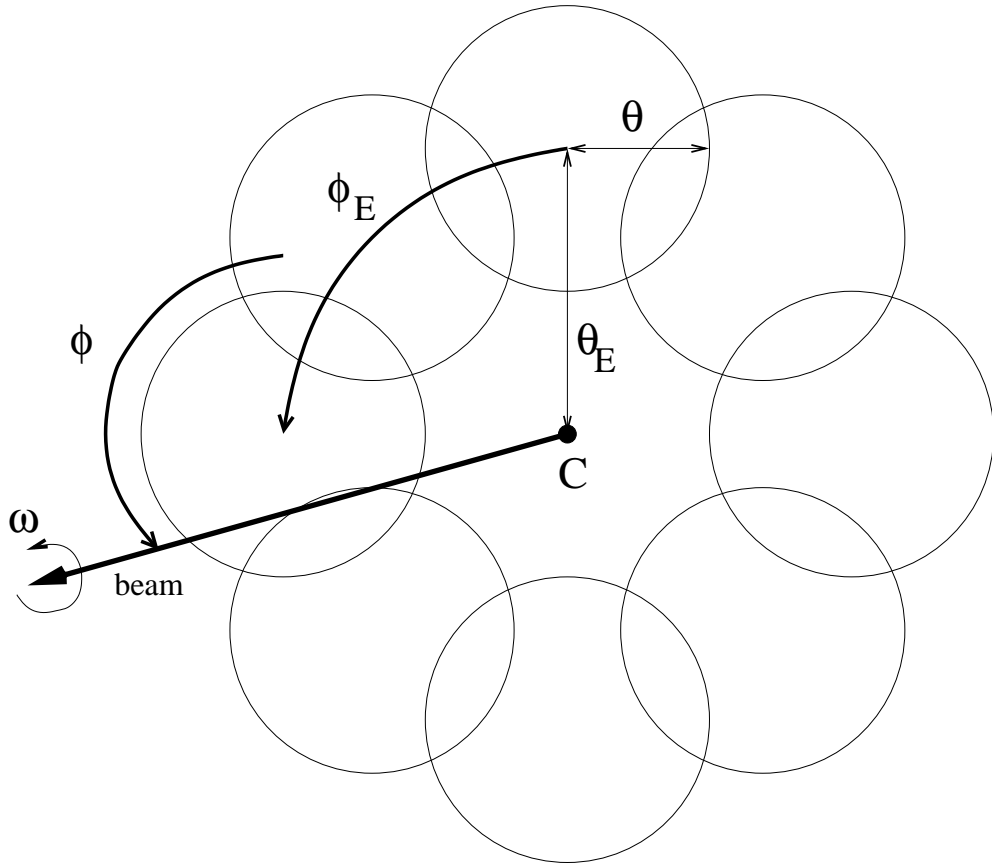


Figure 3.1: The structure of a basic scan path described in the text. The scan rings are all at an angular distance  $\theta_E$  from the center  $C$  and the angular radius of the rings is  $\theta$ . The angle  $\phi_E$  is the azimuthal angle of the center of a given ring and  $\phi$  gives the position in a specified ring.

To do the integral (equation (3.1)) is very time consuming. Say that the expansion of the sky and beam in spherical harmonics can be truncated at the multipoles  $L_s$  and  $L_b$  respectively. Then defining  $L = \min(L_s, L_b)$ , the scaling of the

integral (equation (3.1)) is  $\mathcal{O}(L^2)$  operations for each position  $(\Phi_2(t), \Theta(t), \Phi_1(t))$  of the beam. The temperature has to be evaluated at at least  $\mathcal{O}(L^3)$  such positions giving a total scaling of the convolution of  $\mathcal{O}(L^5)$  operations.

The aim of (Wandelt and Górski 2000) is to increase the speed of this convolution by going to spherical harmonic space. Before doing this step the *basic scan path* should be introduced. This is the scan path of an experiment scanning on rings such as the MAP and Planck satellite. This basic scan path is the basis of the fast convolution method and also of the power spectrum estimation technique presented in the next section.

The basic scan path consists of a ring of rings on the sphere. This is illustrated in figure (3.1). Given a certain centre axis (the centre axis intersects the sphere in the point C on the figure), the experiment scans on rings of angular radius  $\theta$  with the centres situated at an angular distance  $\theta_E$  from the point C. A given ring can be characterized by the azimuthal angle (about C)  $\phi_E$  of its centre and a certain point on the ring can be given by the angle  $\phi$  which goes from 0 to  $2\pi$  along the scanning ring. The angle  $\phi$  is zero at the point on the ring which is furthest away from C and increases in the right handed direction about the outward normal from the centre of the sphere. Finally as a generalization of the basic scan path, an angle  $\omega$  can describe a right handed rotation of the beam itself about the normal from the centre of the sphere in the direction of the beam. For a certain experiment,  $\theta_E$  and  $\theta$  can be fixed so that the timestream is written as a function of the remaining angles as  $T(\phi_E, \phi, \omega)$ . As discussed in the next section this will not always be the case, but will be assumed in the rest of the thesis.

With these definition, the rotation in equation (3.1) can be factorised into two rotations

$$\hat{D}(\Phi_2, \Theta, \Phi_1) = \hat{D}(\phi_E, \theta_E, 0)\hat{D}(\phi, \theta, \omega). \quad (3.2)$$

This can be thought of in the following way. With the beam pointing in the direction of the north pole, the beam is first rotated into the position in the scanning ring, such that the north pole is now the centre of the scanning ring. Then the centre of the scanning ring is moved into its correct position on the sphere.

A general function  $s(\hat{\mathbf{n}})$  on the sphere can be expanded in spherical harmonics as

$$s(\hat{\mathbf{n}}) = \sum_{\ell=0}^{L_s} \sum_{m=-\ell}^{\ell} s_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}). \quad (3.3)$$

With this expansion, equation (3.1) can be written

$$T(\phi_E, \phi, \omega) = \sum_{\ell m} \sum_{\ell' m'} s_{\ell m} \left[ \hat{D}(\phi_E, \theta_E, 0) \hat{D}(\phi, \theta, \omega) b \right]_{\ell' m'}^* \int d\hat{\mathbf{n}} Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell' m'}(\hat{\mathbf{n}}), \quad (3.4)$$

where the limits on the  $\ell$  and  $m$  sums are taken to be the ones in equation (3.3). By the orthogonality of spherical harmonics this can be written as

$$T(\phi_E, \phi, \omega) = \sum_{\ell m} \sum_{MM'} s_{\ell m} D_{mM}^{\ell*}(\phi_E, \theta_E, 0) D_{MM'}^{\ell*}(\phi, \theta, \omega) b_{\ell M'}^*, \quad (3.5)$$

where the rotation operators are written in spherical harmonic space as described in appendix (A). As given in the appendix the spherical harmonic rotation coefficients can be written as

$$D_{mm'}^{\ell}(\phi_2, \theta, \phi_1) = e^{-im\phi_2} d_{mm'}^{\ell}(\theta) e^{-im'\phi_1}, \quad (3.6)$$

where the  $d_{mm'}^{\ell}(\theta)$  factors can be calculated by recursion. This gives

$$T(\phi_E, \phi, \omega) = \sum_{\ell m} \sum_{MM'} s_{\ell m} b_{\ell M'}^* d_{mM}^{\ell}(\theta_E) d_{MM'}^{\ell}(\theta) e^{i(m\phi_E + M\phi + M'\omega)}. \quad (3.7)$$

Now defining the 3D Fourier transform of the time stream as

$$T_{mm'm''} = \frac{1}{(2\pi)^3} \int_0^{2\pi} d\phi_E d\phi d\omega T(\phi_E, \phi, \omega) e^{-i(m\phi_E + m'\phi + m''\omega)}, \quad (3.8)$$

one gets after doing the summations (which give delta functions)

$$T_{mm'm''} = \sum_{\ell} s_{\ell m} d_{mm'}^{\ell}(\theta_E) d_{m'm''}^{\ell}(\theta) b_{\ell m''}^*. \quad (3.9)$$

The timestream is easily found from the inverse Fourier transformation giving

$$T(\phi_E, \phi, \omega) = \sum_{mm'm''=-L}^L T_{mm'm''} e^{i(m\phi_E + m'\phi + m''\omega)}. \quad (3.10)$$

Using these formulae, the convolution of a sky with a general beam can be done in  $\mathcal{O}(L^4)$  operations (actually even less as described in (Wandelt and Górski 2000)).

In the work in the next section, the basic scan path will be assumed. The basic scan path is the path with  $\omega = 0$ . In this case, the Fourier transform can be written in terms of  $m$  and  $m'$ . The variable  $m''$  has to be summed over

$$T_{mm'} = \sum_{m''} T_{mm'm''} \quad (3.11)$$

$$= \sum_{\ell} s_{\ell m} d_{mm'}^{\ell}(\theta_E) X_{\ell m'}(\theta), \quad (3.12)$$

where

$$X_{\ell m} = \sum_M d_{mM}^{\ell}(\theta) b_{\ell M}^*. \quad (3.13)$$

The reason for the sum over  $m''$  is clear from equation (3.10),

$$T(\phi_E, \phi, 0) = \sum_{mm'm''=-L}^L T_{mm'm''} e^{i(m\phi_E + m'\phi + m'')}. \quad (3.14)$$

This Fourier transformed time stream for a basic scan path will be used in the next section for fast power spectrum estimation from the time stream.

## 3.2 Power Spectrum Estimation Using Scanning Rings

As discussed in section (2.3.1), the problem of calculating the full likelihood of a CMB sky map is the inversion of the pixel correlation matrix. Because the CMB pixel correlation matrix has the nice property that it is only dependent on the angular distance between two pixels (assuming statistical isotropy), the correlation matrix in spherical harmonic space is diagonal. But the noise has no symmetries on the sphere and for this reason the correlation matrix does not have any properties that makes it easy to invert.

On the time stream however, the noise has a very nice symmetry. As discussed in section (2.2.1), the correlation of the noise between two elements  $d_i$  and  $d_j$  in the timestream is only dependent on the distance  $|i - j|$  (along the timestream) between the two elements. In Fourier space this makes the correlation matrix diagonal. But the signal does not in general have any symmetries on the timestream.

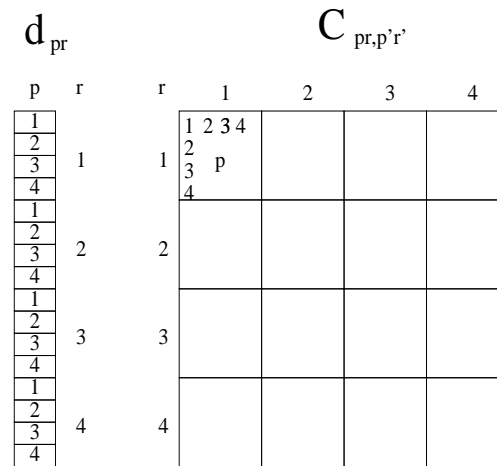


Figure 3.2: The figure shows the structure of a vector  $d_{pr}$  as a column vector and a correlation matrix  $C_{pr,p'r'}$  for a data set with 4 rings and 4 pixels per ring. The purpose of the figure is to show the position of the elements in a vector and a matrix with a ring-pixel index. The ring number is shown on the right side of the vector and the pixel number in each ring is shown inside the vector. For the matrix, the ring numbers are showed on the outside (above and left) and the pixel numbers are only shown in the upper left block.

This fact changes for an experiment scanning on rings using a basic scan path as described in the previous section. When one assumes that the angular distance between the centre of two adjacent rings is the same and one uses the coadded rings (when an experiment scans the same ring several times, the coadded ring is just the average of these scans) as input datavector in the likelihood, it turns out that some symmetries in both the noise and signal correlation matrices are kept. When using the coadded rings as datavector  $\mathbf{d}$ , the noise correlation between two elements in the vector,  $d_i$  and  $d_j$  is no longer only dependent on the distance  $|i-j|$ . But when writing this datavector indexed with pixel number  $p$  (or  $\phi$  in the continuous case discussed in the previous section) in ring number  $r$  ( $\phi_E$ ) as  $d_{pr}$  (see figure (3.2)) one can see that there still is some symmetry present. The correlation between two rings  $r$  and  $r'$  will of course only dependent on the distance between the rings  $|r-r'|$ . The signal correlation between two elements in the

timestream is only dependent on the distance on the sphere, but because of the geometrical positioning of the rings on the sphere, also the correlation between  $d_{pr}$  and  $d_{p'r'}$  for the signal will depend on  $|r - r'|$ . This can also easily be seen by studying figure (3.1). Consider two positions  $p$  and  $p'$  on the rings  $r$  and  $r'$  and the same positions  $p$  and  $p'$  on the rings  $R$  and  $R'$ . If the distance between the two pairs of rings is the same  $|r - r'| = |R - R'|$  then the angular distance on the sphere between the points  $p$  and  $p'$  on these two different pairs of rings is the same.

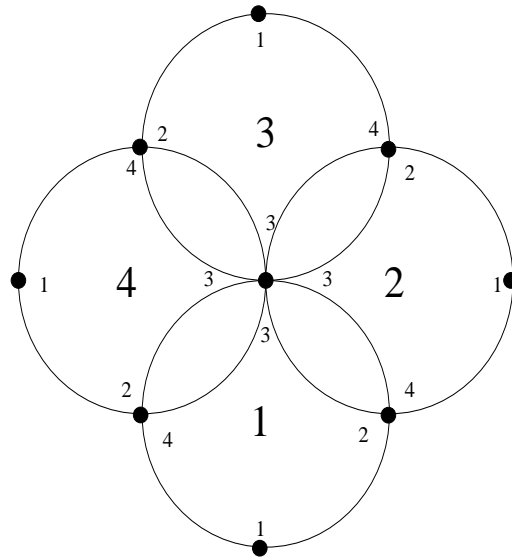


Figure 3.3: An example of an experiment scanning on 4 rings with 4 sample points (shown as black dots) on each ring. The big numbers indicate the rings numbers  $r$  and the smaller numbers indicate the number of the sample points  $p$  in each ring.

The fact that the correlation  $\langle d_{pr}d_{p'r'} \rangle$  between two elements in the ringset is only dependent on the distance between rings  $|r - r'|$  both for the signal and noise simplifies the structure of the correlation matrix. The correlation matrix will be *block circulant*, meaning that each row of blocks will be equal to the previous row of blocks shifted one block to the right. In addition a block circulant matrix is block symmetric meaning that the matrix is symmetric when the blocks are considered as the elements of the matrix. This is illustrated in figure (3.3)

and (3.4). In this example, one has 4 rings with 4 sample points on each (figure (3.3)). The  $4 \times 4$  correlation matrix between points  $p$  and  $p'$  for ring 1 is called  $A$ . One has  $A_{pp'} = \langle d_{1p}d_{1p'} \rangle$ . The correlation matrix between a point  $p$  on ring 1 and a point  $p'$  on ring 2 is  $B_{pp'} = \langle d_{1p}d_{2p'} \rangle$ . In the same way the correlations between rings 1 and 3 is given by the matrix  $C$ . Obviously the correlations between rings 1 and 4 will be the same as for rings 1 and 2 as the absolute value of the distance between the rings  $|r - r'|$  is the same. The full  $16 \times 16$  correlation matrix  $\langle d_{pr}d_{p'r'} \rangle$  for this case is shown in figure (3.4) (see also figure (3.2)). One clearly sees the block circulant structure. It has to be noted that this example has very few rings and is therefore unrealistic. With this small number of rings, the approximation that the noise correlations are dependent on the absolute value of the distance between rings breaks down. This is because the correlations between the first and the last ring cannot be the same as between the first and the second. The noise is not circulant. But as mentioned in section (2.2.1) this is a good approximation when the number of rings is large as in a real CMB experiment.

The block circulant structure of the correlation matrix is what makes the likelihood evaluation fast. A block circulant matrix is block diagonal in Fourier space and one can therefore find the determinant and inverse faster. This will be formalized in the next section.

**A B C B**  
**B A B C**  
**C B A B**  
**B C B A**

Figure 3.4: The correlation matrix  $\langle d_{pr}d_{p'r'} \rangle$  for the ringset shown in figure (3.3). Here  $A$  is the  $4 \times 4$  correlation matrix for points  $p$  and  $p'$  within the same ring  $A_{pp'} = \langle d_{pr}d_{p'r} \rangle$ ,  $B$  is the  $4 \times 4$  correlation matrix between a point  $p$  on ring  $r$  and point  $p'$  on ring  $r \pm 1$ . Finally  $C$  is the correlation matrix between a point  $p$  on ring  $r$  and point  $p'$  on ring  $r \pm 2$ .

### 3.2.1 Theory

As a generalization of the previous example, one can consider an experiment which has  $N_r$  rings,  $N_R$  pixels per ring and scans each ring  $N_c$  times. To go from the time stream  $d_n$  with  $N_{TOD} = N_r N_R N_c$  elements to the coadded ringset  $d_{pr}$  with  $N_r N_R$  elements one has to average the  $N_c$  scans of each ring. This can be written as

$$d_{pr} = \sum_{n=0}^{N_{TOD}-1} A_{pr\ n} d_n. \quad (3.15)$$

Here the *coadding matrix*  $A_{pr\ n}$  selects which elements of the whole timestream that belongs to ring  $r$  and pixel  $p$ . The index  $n$  can be written in terms of three indices  $x$ ,  $y$  and  $z$  as  $n = N_R N_c x + N_R y + z$ . Here  $x$  is the ring number going from 0 to  $N_r - 1$ ,  $y$  is the scan number for a given ring going from 0 to  $N_c - 1$  and  $z$  is the pixel number in the given ring being in the interval  $[0, N_R - 1]$ . With this indexing equation (3.15) can be written as

$$d_{pr} = \sum_{x=0}^{N_r-1} \sum_{y=0}^{N_c-1} \sum_{z=0}^{N_R-1} A_{pr\ xyz} d_{xyz}. \quad (3.16)$$

It is now obvious that the co-adding matrix can be written as

$$A_{pr\ xyz} = \frac{1}{N_c} \delta_{rx} \delta_{zp}, \quad (3.17)$$

where the prefactor makes the division with the number of scans per ring after the measurements from all of the scans have been added together. Using this expression equation (3.16) can be written

$$d_{pr} = \frac{1}{N_c} \sum_{y=0}^{N_c-1} d_{r y p}. \quad (3.18)$$

In the Likelihood estimation, the signal and noise correlation matrices will be needed in Fourier space. Because of this and since the signal correlation matrix is simpler to calculate directly in Fourier space, the full expressions will be derived directly in this space. However in pixel space it is easier to see how the symmetries work to give the simple form of the correlation matrices. For this reason I will first outline the structure of the correlation matrices in pixel space.

The CMB temperature on pixel  $p$  in ring  $r$  can be written as  $d_{pr} = T(\theta, \phi)$  where  $\theta$  and  $\phi$  are polar coordinates in a coordinate system in which C (see figure (3.1)) is the north pole. For a given pixel and ring  $pr$  these angles can be written as  $\theta_{pr} = \theta_E + \Delta\theta_p$  and  $\phi_{pr} = \phi_r + \Delta\phi_p$ . Here  $\theta_E$  is as before the polar coordinate of the center of the rings and  $\Delta\theta_p$  is the offset from the center depending on



the pixel  $p$ . The  $\theta$  coordinate is clearly independent of the ring number. The azimuthal coordinate depends on the angle  $\phi_r$  which is the azimuthal position of the centre of ring  $r$  and  $\Delta\phi_p$  which is the offset depending only on the pixel  $p$  in the ring. One can now write

$$C_{pr,p'r'}^S = \langle d_{pr}^S d_{p'r'}^S \rangle, \quad (3.19)$$

$$= \langle T(\theta_p, \phi_{pr}) T(\theta_{p'}, \phi_{p'r'}) \rangle, \quad (3.20)$$

$$= \sum_{\ell m} \sum_{\ell' m'} \langle a_{\ell m} a_{\ell' m'} \rangle Y_{\ell m}(\theta_p, \phi_{pr}) Y_{\ell' m'}(\theta_{p'}, \phi_{p'r'}), \quad (3.21)$$

$$= \sum_{\ell} C_{\ell} \frac{2\ell + 1}{4\pi} P_{\ell}(\cos \delta\theta), \quad (3.22)$$

where  $\delta\theta = \cos \theta_p \cos \theta_{p'} + \sin \theta_p \sin \theta_{p'} \cos(\phi_r - \phi_{r'} + \Delta\phi_p - \Delta\phi_{p'})$  is the angle between the two points  $pr$  and  $p'r'$ . The superscript S means 'signal'. Clearly  $\delta\theta$  and thereby  $C_{pr,p'r'}^S$  is a function  $f(|r - r'|, p, p')$  of the distance between the rings. The matrix  $C_{pr,p'r'}^S$  is block circulant with blocks of size  $N_R \times N_R$ . To show the same for the noise correlation matrix one can use that  $\langle d_n d_{n'} \rangle = \chi(|n - n'|)$ . One then gets

$$\begin{aligned} C_{pr,p'r'}^N &= \langle d_{pr}^N d_{p'r'}^N \rangle \\ &= \frac{1}{N_c^2} \sum_{yy'} \langle d_{ryp}^N d_{r'y'p'}^N \rangle \\ &= \frac{1}{N_c^2} \sum_{yy'} \chi(|N_R N_c (r - r') + N_R (y - y') + p - p'|) \\ &= \frac{1}{N_c^2} \sum_{yy'} \chi(|r - r'|, p, p'), \end{aligned} \quad (3.23)$$

where the superscript N means 'noise'. So the correlation matrix  $C_{pr,p'r'}$  is block circulant both for signal and noise.

It is well known that the Fourier transform of a circulant matrix is diagonal. I will now show this generally. By the Fourier transform of a  $N \times N$  matrix  $\mathbf{A}$  I mean

$$\mathbf{A}' = \mathbf{F}^\dagger \mathbf{A} \mathbf{F}, \quad (3.24)$$

where  $\mathbf{F}$  is an  $N \times N$  matrix with elements

$$F_{jk} = \frac{1}{\sqrt{N}} e^{-2\pi i j k / N}, \quad (3.25)$$

and the dagger means Hermitian conjugate. The elements of the Fourier transformed matrix can be written

$$A'_{kk'} = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{j'=0}^{N-1} e^{-2\pi i j k / N} e^{2\pi i j' k' / N} A_{jj'}, \quad (3.26)$$

$$= \frac{1}{N} \sum_{jj'} e^{-2\pi i(kj-k'j')/N} A_{jj'},$$

where the last sum has the same limits as the upper sum. For a circulant matrix  $A_{jj'} = A_{|j-j'|}$  and one has

$$\begin{aligned} A'_{kk'} &= \frac{1}{N} \sum_{j=0}^{N-1} \left[ \sum_{j'=0}^j e^{-2\pi i(kj-k'j')/N} A_{|j-j'|} \right. \\ &\quad \left. + \sum_{j'=j+1}^{N-1} e^{-2\pi i(kj-k'j')/N} A_{|j-j'|} \right], \\ &= \frac{1}{N} \sum_{j=0}^{N-1} \left[ \sum_{\Delta=0}^j e^{-2\pi ij(k-k')/N} A_{\Delta} e^{-2\pi ik'\Delta} \right. \\ &\quad \left. + \sum_{\Delta=-1}^{j-(N-1)} e^{-2\pi ij(k-k')/N} A_{-\Delta} e^{-2\pi ik'\Delta/N} \right], \end{aligned} \quad (3.27)$$

where the substitution  $\Delta = j - j'$  has been made. Now setting  $\Delta' = \Delta + N$  in the last sum this can be written as

$$\begin{aligned} A'_{kk'} &= \frac{1}{N} \sum_{j=0}^{N-1} \left[ \sum_{\Delta=0}^j e^{-2\pi ij(k-k')/N} A_{\Delta} e^{-2\pi ik'\Delta} \right. \\ &\quad \left. + \sum_{\Delta=j+1}^{N-1} e^{-2\pi ij(k-k')/N} A_{-(\Delta+N)} e^{-2\pi ik'\Delta} \right]. \end{aligned} \quad (3.28)$$

Finally defining  $A_{-(\Delta+N)} = A_{\Delta}$  one can write

$$\begin{aligned} A'_{kk'} &= \frac{1}{N} \sum_{\Delta=0}^{N-1} A_{\Delta} e^{-2\pi ik'\Delta} \underbrace{\sum_{j=0}^{N-1} e^{-2\pi ij(k-k')/N}}_{N\delta_{kk'}} \\ &= \delta_{kk'} \sum_{\Delta=0}^{N-1} A_{\Delta} e^{-2\pi ik\Delta/N}, \end{aligned} \quad (3.29)$$

which shows that the Fourier transform of a circulant matrix is diagonal. In the same way the Fourier transform of a block circulant matrix is block diagonal. Hence the Fourier transformed of the correlation matrix  $C_{pr,p'r'}$  is block diagonal and the determinant and inverse can be found fast. This was shown in pixel space where the reason can be easily understood physically. Now the derivation of the full expressions which will be needed for likelihood estimation will be shown.

Equation (3.8) defines the 2D Fourier transform of  $d_{pr}$  ( $\omega = 0$ ). In discrete form the equation can be written,

$$d_{mm'} = \frac{1}{N_r N_R} \sum_{r=0}^{N_r} \sum_{p=0}^{N_R} e^{-2\pi ipm/N_R} e^{-2\pi ir m'/N_r} d_{pr}. \quad (3.30)$$

For the signal correlation matrix, expression (3.12) can be used by replacing  $s_{\ell m} = a_{\ell m}$ . One then gets for the correlation matrix in Fourier space  $C_{mm',MM'}^S$

$$C_{mm',MM'}^S = \langle d_{mm'}^S d_{MM'}^S \rangle \quad (3.31)$$

$$= \sum_{\ell\ell'} \langle a_{\ell m} a_{\ell' M} \rangle d_{mm'}^\ell(\theta_E) d_{MM'}^{\ell'}(\theta_E) X_{\ell m'}(\theta) X_{\ell' M'}(\theta) \quad (3.32)$$

$$= \delta_{mM} \sum_{\ell} C_{\ell} d_{mm'}^\ell(\theta_E) d_{MM'}^\ell(\theta_E) X_{\ell m'}(\theta) X_{\ell M'}(\theta). \quad (3.33)$$

The first delta function shows that the signal correlation matrix is block diagonal in Fourier space. The diagonal blocks can be calculated as a sum over the power spectrum, rotation matrices which can be calculated by recursion and pre-computed  $X_{\ell m}$  factors.

The noise matrix is a bit more complicated. As discussed in section (2.2.1), the noise can be described by a power spectrum  $P(k)$  given as  $P(k) = \langle d_k d_k^* \rangle$  where  $d_k$  are the 1D Fourier coefficients of the noise,

$$d_k^N = \sum_n d_n^N e^{-2\pi i k n / N}, \quad (3.34)$$

where  $N = N_{TOD}$ . The reverse transformation is

$$d_n^N = \frac{1}{N} \sum_k d_k^N e^{2\pi i k n / N}, \quad (3.35)$$

which gives

$$\langle d_n^N d_{n'}^N \rangle = \frac{1}{N^2} \sum_{kk'} \langle d_k^N d_{k'}^N \rangle e^{2\pi i (k n - k' n') / N}, \quad (3.36)$$

$$= \frac{1}{N} \sum_k P(k) e^{2\pi i k (n - n') / N}, \quad (3.37)$$

which follows from the fact that the noise is (assumed to be) circulant. The correlation  $C_{pr,p'r'}^N = \langle d_{pr}^N d_{p'r'}^N \rangle$  is easiest to calculate in pixel space. For this reason a Fourier transform will be applied to  $C_{pr,p'r'}$  to get the correlation matrix in Fourier space  $C_{mm',MM'}^N$ . Using equation (3.18) one has that

$$\langle d_{pr} d_{p'r'} \rangle = \frac{1}{N_c^2} \sum_{yy'} \langle d_{r_y p} d_{r'_y p'} \rangle. \quad (3.38)$$

Using equation (3.37) one can write this as

$$C_{pr,p'r'}^N = \frac{1}{N_c^2} \sum_{k=0}^{N-1} P(k) \sum_{y,y'=0}^{N_c-1} e^{-2\pi i k / N (N_R N_c (r-r') + N_R (y-y') + (p-p'))} \quad (3.39)$$

$$\equiv C^N(r - r', p - p'). \quad (3.40)$$

The Fourier transform is found by the use of equation (3.30),

$$C_{mm',MM'}^N = \frac{1}{N_r^2 N_c^2} \sum_{p,p'=0}^{N_R-1} \sum_{rr'=0}^{N_r-1} C^N(r-r', p-p') \quad (3.41)$$

$$\times e^{-2\pi i(rm-r'M)/N_r} e^{-2\pi i(pm'-p'M')/N_R}$$

$$= \frac{1}{N_r^2 N_c^2} \sum_{p,p'=0}^{N_R-1} \sum_{r=0}^{N_r-1} \sum_{\Delta=r}^{r-(N_r-1)} C(\Delta, p-p') e^{-2\pi ir/N_r(m-M)} \quad (3.42)$$

$$\times e^{-2\pi iM\Delta/N_r} e^{-2\pi i(pm'-p'M')/N_R}$$

$$= \frac{1}{N_c^2 N_r} \delta_{mM} \sum_{p,p'=0}^{N_R-1} e^{-2\pi i(m'p-M'p')/N_R} \quad (3.43)$$

$$\times \sum_{\Delta=0}^{N_r-1} e^{-2\pi iM\Delta/N_r} C(\Delta, p-p'),$$

where the last transition follows from the fact that  $C(r-r'+mN_r, p-p') = C(r-r', p-p')$  where  $m$  is an integer.

Before embarking on the likelihood estimation, one practical issue has to be mentioned. When only parts of the sky are observed one does not have enough data to estimate the whole power spectrum multipole by multipole. Instead the power spectrum is estimated in  $N_b$  bins of size  $\Delta\ell_b$  such that  $C_\ell = C_b^B$  when  $\ell_b \leq \ell < \ell_{b+1}$ , where  $\ell_{b+1} = \ell_b + \Delta\ell_b$ .

The likelihood can be written in the form given in equation (2.26)

$$L(C_b^B) = \frac{e^{-\frac{1}{2}\mathbf{d}^\dagger \mathbf{C}^{-1}(C_b^B)\mathbf{d}}}{\sqrt{2\pi \det \mathbf{C}(C_b^B)}}. \quad (3.44)$$

The data vector  $\mathbf{d}$  here is the Fourier transformed timestream with elements  $d_{mm'}$  and the correlation matrix is the sum of the signal covariance matrix  $C_{mm',MM'}^S$  in equation (3.33) and the noise correlation matrix  $C_{mm',MM'}^N$  in equation (3.43). If  $N$  is the number of elements in  $\mathbf{d}$  (NOT the same  $N$  as above!!!) which is  $N = N_r N_R$  then because of the block diagonal structure of  $C_{mm',MM'}^S$  the correlation matrix can be inverted in  $N_R^3 N_r$  operations instead of  $N^3$ . If the number of rings is equal to the number of pixel per ring  $N_r = N_R$  then this means that the inversion of the correlation matrix takes  $N^2$  instead of  $N^3$ . This is also the scaling of the determinant evaluation. This is easily seen: The inverse of a block diagonal matrix is just a block diagonal matrix containing the inverted blocks. To invert each  $N_R \times N_R$  block takes  $N_R^3$  operations and there are  $N_r$  blocks to invert. The determinant of a block diagonal matrix is the product of the determinant of the individual blocks. Finding the determinant of a  $N_R \times N_R$  matrix takes  $N_R^3$  operations and there are  $N_r$  blocks giving the same scaling to inversion

and determinant evaluation.

The inversion and determinant evaluation are fast but one has to check the operation count for the calculation of the correlation matrices themselves. From expression (3.33) the calculation of the signal correlation matrix is just a sum over  $N_b$  bins for each of the  $N_R^2 N_r$  elements of the block diagonal matrix. Also the calculation of the rotation matrices  $d_{mm'}^\ell$  are fast using recursion or precalculation. The noise matrix can be calculated using 3 FFTs (see equation (3.43)) scaling as  $(N_R \log N_R)^2 N_r \log N_r$ . The precalculation of the function  $C^N(r-r', p-p')$  takes  $N \log N$  operations for the FFT in equation (3.39) times  $N_r N_R$  for each element.

To quickly find the maximum of the likelihood equation (3.44) with respect to the binned power spectrum  $C_b^B$  one needs to find the derivative. In practice the log-likelihood

$$\mathcal{L}(C_b^B) = -2 \log L(C_b^B) = \mathbf{d}^\dagger \mathbf{C}^{-1}(C_b^B) \mathbf{d} + \log \det \mathbf{C}(C_b^B). \quad (3.45)$$

is minimized and its derivative is given by equation (2.44). The operations involved in the derivative calculations are fast. The first term only needs multiplications of blocks diagonal matrices with vectors which takes  $N_R^2 N_r$  operations. The last term involves finding the trace of a matrix product of two block diagonal matrices also taking  $N_R^2 N_r$  operations.

A final issue involves finding the error bars on the estimate. According to section (2.3.1), the error bars can be found by taking the inverse of the Fisher matrix. Unfortunately the evaluation of the Fisher matrix scales almost as  $N^3$ . This is seen from equation (2.47). Multiplication of the block diagonal matrices takes  $\mathcal{O}(N_R^3 N_r)$  operations. These matrix multiplications have to be done for each of the  $N_b^2$  elements in the Fisher matrix. If  $N_R = N_r$  then the operation count is  $\mathcal{O}(N^2 N_b^2)$  which could be almost  $N^3$  for many bins. Another approach would be to explore the likelihood surface around the maximum. By finding where the likelihood has fallen by a certain ratio compared to its value at the maximum, approximate confidence regions can be found. However each likelihood evaluation takes  $N^2$  operations and this has to be done for each bin giving  $N^2 N_b$  operations or  $N^{5/2}$  for many bins. To find the full covariance of the estimate this way would scale as  $N^2 N_b^2$ . A faster way would be to approximate the fisher matrix (which is the average of the second derivative of the log-likelihood at the maximum) by the numerical second derivative at the likelihood maximum. Evaluating the gradient of the likelihood takes only  $\mathcal{O}(N^2)$  operations. The  $i$ -th row of the covariance matrix of the estimates can be found by taking the numerical derivative of the gradient with respect to the  $i$ -th bin which then also takes  $\mathcal{O}(N^2)$  operations. Finding the whole covariance matrix of the estimates then takes  $\mathcal{O}(N^2 N_b)$  operations.

### 3.2.2 An Example

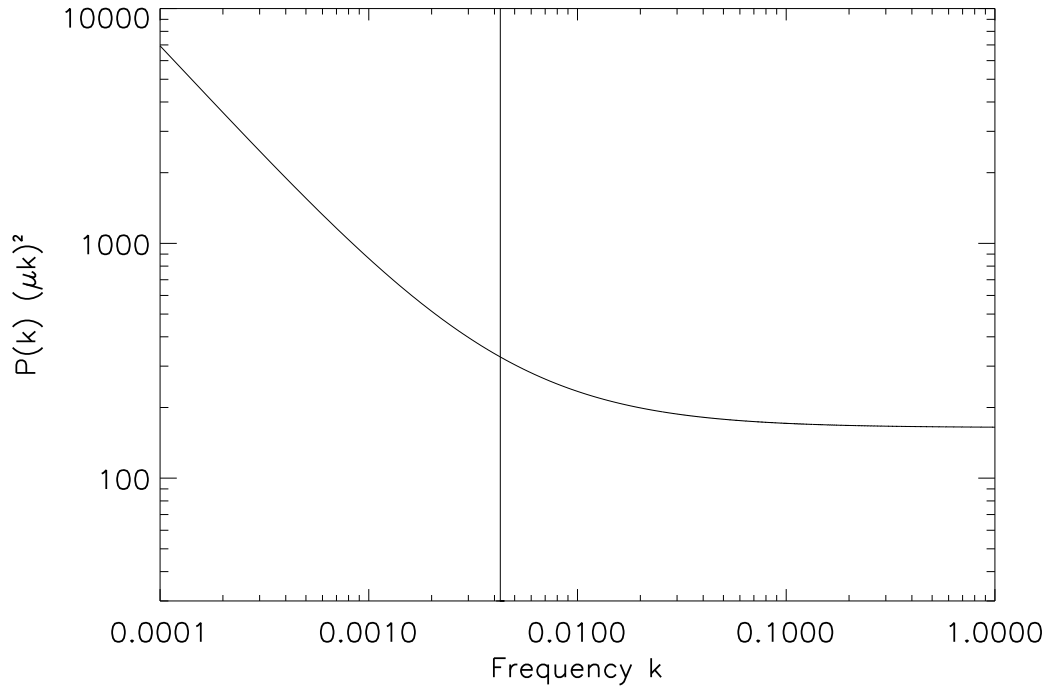


Figure 3.5: The noise power spectrum used in the simulation. The vertical line shows the knee frequency. The frequency  $k$  is shown in units of the Nyquist frequency  $N_{TOD}/2$ .

To test the above method for power spectrum estimation I simulated an example where  $N_r = N_R = 243$  meaning that there are 243 scanning rings with 243 pixels in each. Each ring is only scanned once  $N_c = 1$  so that the total number of pixels  $N_{TOD} = 59049$ , a similar number of pixels to the BOOMERANG experiment. The opening angle  $\theta_E = 5^\circ$  and the radius of each ring  $\theta = 5^\circ$ . All the rings intersect at C (figure (3.1)). For simplicity a symmetric Gaussian beam with  $18'$  FWHM was used. This is however not necessary as the method can deal with any beam pattern. The noise power spectrum was of the form (2.8) with  $\alpha = 1$ . Writing  $P(k)$  in terms of wavenumber  $k$  one has

$$P(k) = \sigma^2 \left( 1 + \frac{k_{knee}}{k} \right), \quad (3.46)$$

with  $\sigma = 9\mu\text{K}$  and  $k_{knee} = 4.3 \times 10^{-3}$  in units of the Nyquist frequency  $N_{TOD}/2$ . The power spectrum is showed in figure (3.5). With this noise model, the signal

to noise ratio was one at a multipole of about  $\ell = 600$ .

The simulation was carried out with the following steps:

1. First a standard CDM power spectrum was selected and a timestream was simulated using the Wandelt-Górski method described in section (3.1). A set of  $a_{\ell m}$ s were created using a random number generator and the CDM power spectrum up to a multipole  $L = 1024$ . Then with equation (3.12) the Fourier transformed time stream was created and finally equation (3.14) (FFT) was used to create the time stream in pixel space. The ringset is shown in figure (3.6).

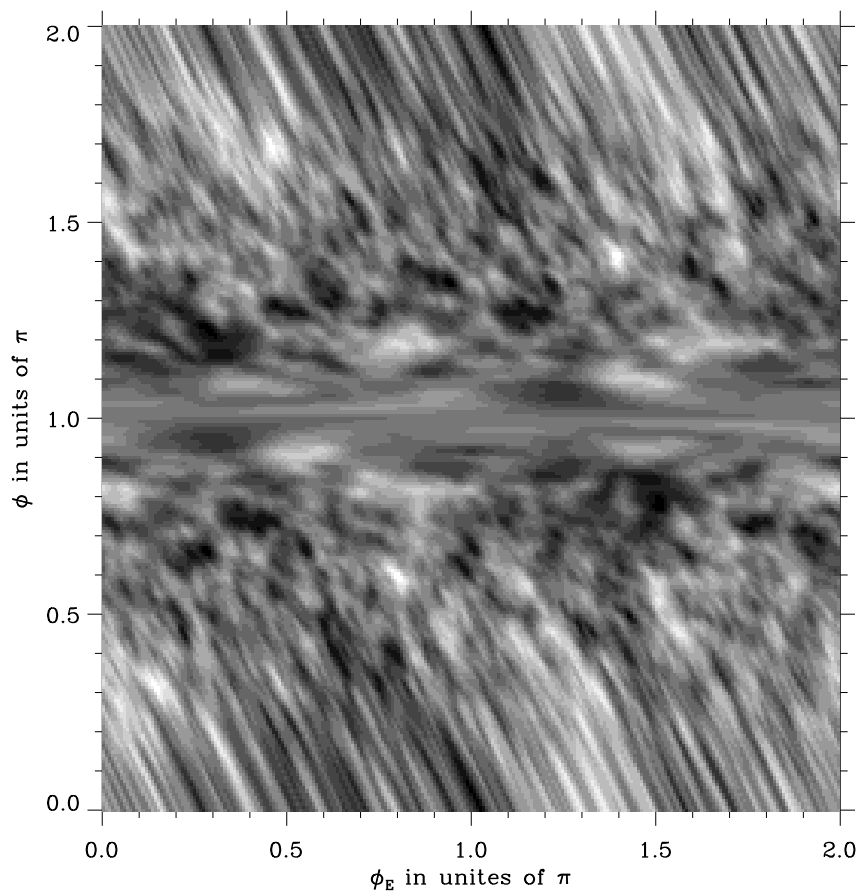


Figure 3.6: The ringset from a scan on a simulated sky without noise. On the plot a line from the bottom to the top describes one ring. In this way the rings are put next to each other from the left to the right. The central point  $C$  where all rings intersect corresponds to the horizontal line of constant value in the middle of the plot.

In the figure, each ring goes from the bottom to the top along the  $\phi$  axis. The rings are put next to each other from the left to the right along the  $\phi_E$  axis. At  $\phi = \pi$  each ring intersects the central point  $C$  and therefore this value is the same for all the rings giving the horizontal line of constant value in the middle. The lines of constant value going from upper left to lower right are the other points of intersection between rings which are close to each other.

2. Noise was generated using the power spectrum above. In Fourier space Gaussian random numbers were created with the power spectrum  $P(k)$ . The noisestream was then Fourier transformed to pixel space (with FFT). The noise ringset is shown in figure (3.7).

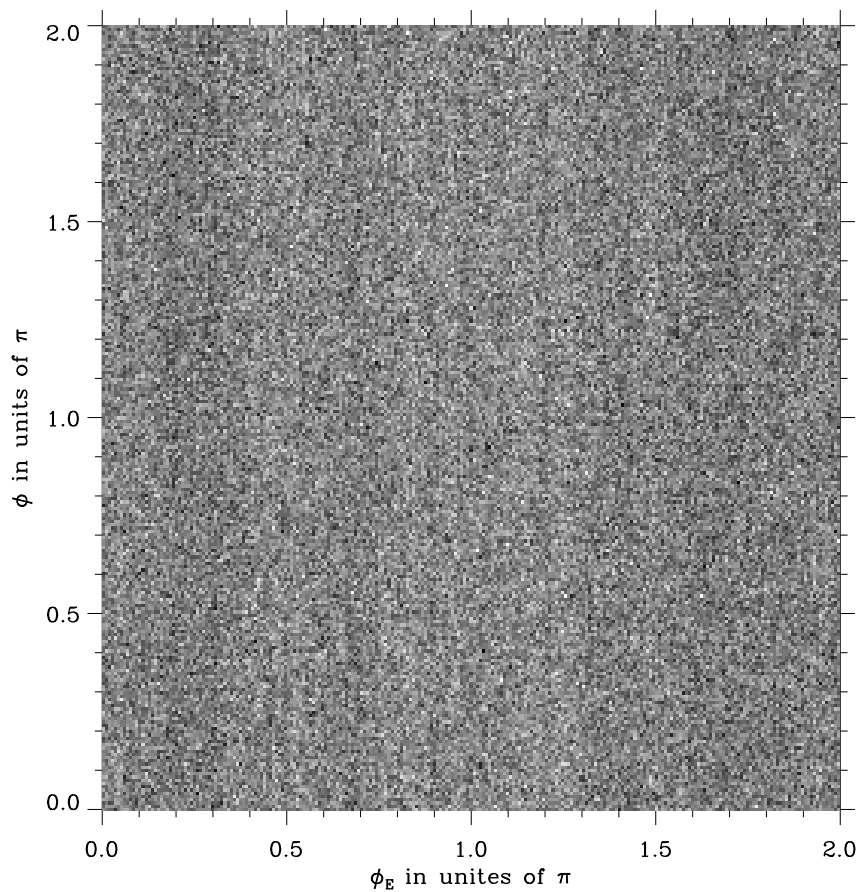


Figure 3.7: Same as figure (3.6) for the noise.

3. The signal and noise ringsets were added together to give a data set from an



experiment scanning on rings. The ringset is shown in figure (3.8). This is the input data from which the power spectrum was to be estimated. From the figure the vertical striping caused by the noise which was discussed in section (2.2.1) is clearly visible.

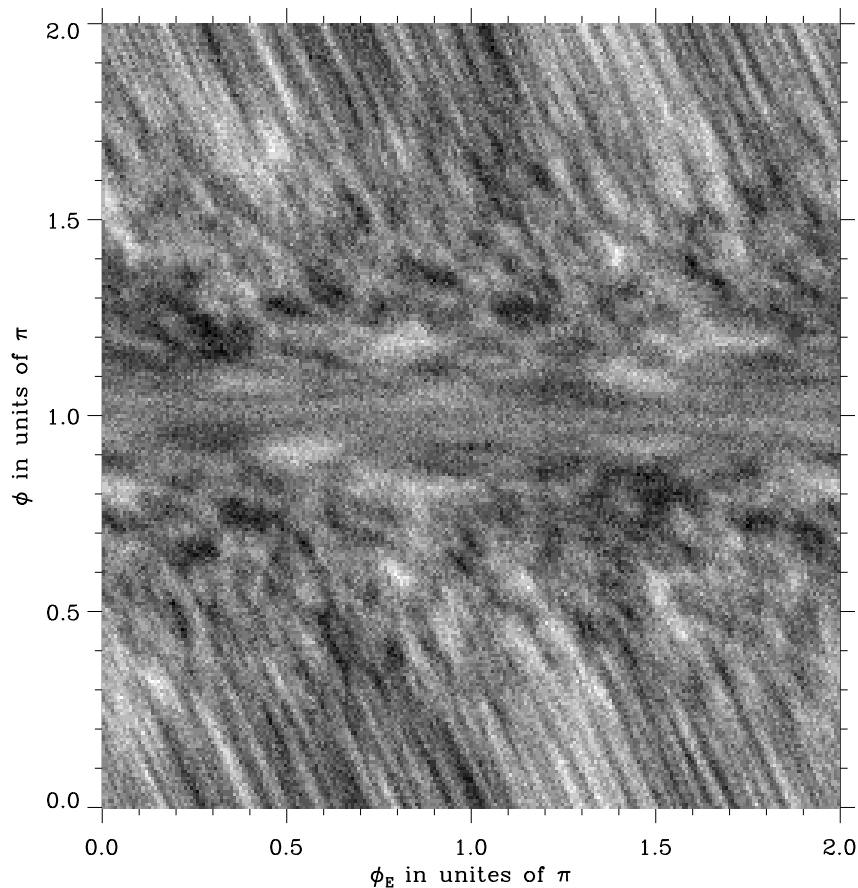


Figure 3.8: Same as figure (3.6) for signal plus noise. This is the result of adding the ringsets in figure (3.6) and (3.7)

4. Now the simulated data set was used for likelihood estimation. A conjugate gradient solver using only first derivative to find the minimum of the log-likelihood was used (Press, Teukolsky, Vetterling, and Flannery 1992). Since the solver can easier handle a solution in which all the parameters have roughly the same order of magnitude, the following change of variables was made:

$$C_\ell = \frac{D_b e^{-\sigma_b^2 \ell(\ell+1)}}{\ell(\ell+1)}, \quad (3.47)$$

where  $D_b$  is the binned power spectrum ( $b$  is bin number) for which the log-likelihood was minimized. The Gaussian beam has  $\sigma = \sigma_b$ . For this minimization 20 bins with 50 multipoles in each were used. The starting guess for the minimizer was  $D_b = 10^4(\mu K)^2$  for all  $b = [0, 19]$ . For this test the correlation matrix was calculated in pixel space (equation 3.22) and then Fourier transformed using FFT. With the binning, equation (3.22) takes the form

$$C_{pr,p'r'}^S = \sum_b D_b \sum_{\ell \in b} \frac{e^{-\sigma_b^2 \ell(\ell+1)}}{\ell(\ell+1)} \frac{2\ell+1}{4\pi} P_\ell(\cos \delta\theta), \quad (3.48)$$

and its derivative is

$$\frac{\partial C_{pr,p'r'}^S}{\partial D_b} = \sum_{\ell \in b} \frac{e^{-\sigma_b^2 \ell(\ell+1)}}{\ell(\ell+1)} \frac{2\ell+1}{4\pi} P_\ell(\cos \delta\theta), \quad (3.49)$$

where the sum over  $\ell \in b$  is the sum over all multipoles  $\ell$  in the bin  $b$ . The angles  $\delta\theta$  between all possible pairs of points  $pr$  and  $p'r'$  were precalculated.

5. After about 30 likelihood and derivative evaluations taking about 2 hours each on a single processor on a 500MHz DEC Alpha Work Station the minimum was found. If the initial guess was better (it could have been taken from a faster approximate power spectrum estimation algorithm as discussed in section (2.3)), the number of likelihood evaluations could have been significantly reduced. The result is shown in figure (3.9). The solid line is the input average power spectrum and the crosses are the estimated bins. The crosses are plotted in the middle of each bin.

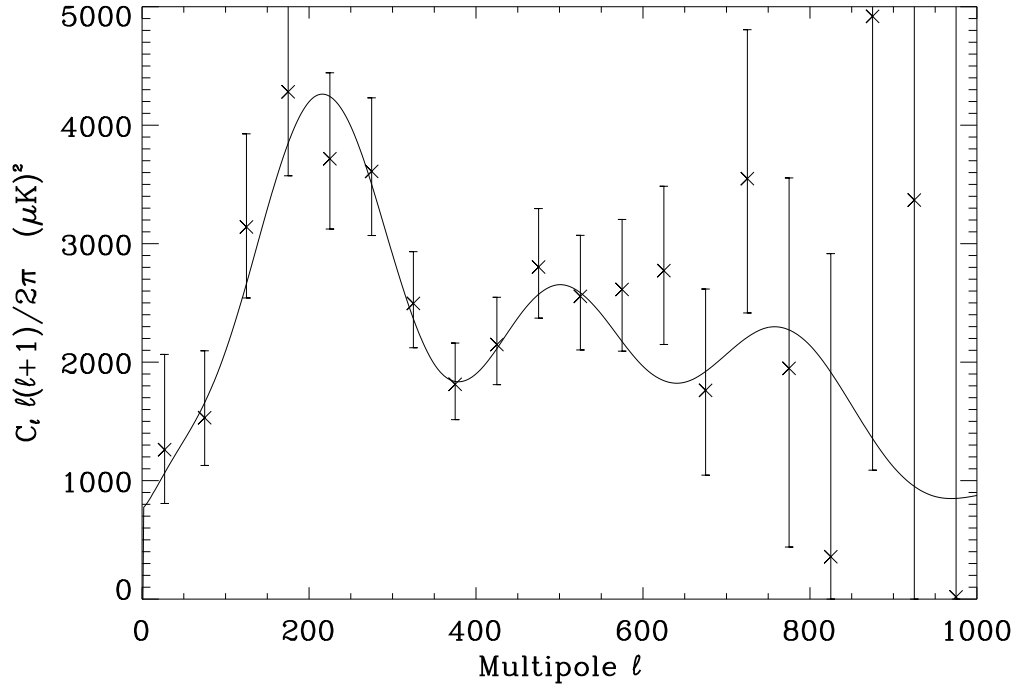


Figure 3.9: The figure shows the result of the likelihood maximisation using the Fourier transformed ringset as input data. The solid line shows the input average power spectrum and the crosses show the estimated bins with error bars. The error bars are approximate asymmetric  $2\sigma$  confidence regions showing where the log-likelihood drops by 2 compared to the maximum.

6. Finally the error bars were found. First the  $2\sigma$  (95%) confidence regions were found by finding where the log-likelihood had been reduced by 2 compared to the maximum. These asymmetric error bars are plotted on the estimates in figure (3.9). Then these error bars were compared to the  $2\sigma$  error bars found from approximating the Fisher matrix by taking the numerical second derivative of the log-likelihood. These error bars were consistent with the ones found from the likelihood contours.

### 3.2.3 Discussion

The power spectrum estimation method described here is the first exact method which takes into account all the following effects:

- **Correlated noise** As discussed in previous sections, most CMB experiments have correlated noise. This can usually be seen in the CMB maps

as stripes. Taking noise correlations into account when estimating the power spectrum is usually hard and involves time consuming calculations or Monte Carlo simulations. When operating on the timestream instead of the sky map, the noise correlations have a much simpler structure and are for this reason easily incorporated into the power spectrum estimation by the method described here.

- **Non-uniform distribution of integration time** Most CMB sky experiments do not observe each point on the sky the same amount of time. For this reason the noise for each point on a sky map will be different. On the timeline integration time per pixel is by definition constant.
- **Beam distortions and far side lobes** Normally the beam of the telescope is approximated to be azimuthally symmetric and Gaussian. This is however not the case for all CMB experiments. Also diffraction of the microwave radiation around the telescope and other apparatus used for the CMB experiment usually leads to far side lobes.
- **Partial sky coverage** Normally only parts of the sky are observed. This leads to coupling between the power spectrum coefficients  $C_\ell$  on the cut sky. These couplings have to be taken into account when attempting to estimate the full sky power spectrum from partial sky data.
- **Pixelization effects** When the CMB data are mapped before the power spectrum estimation is performed, there is an additional smoothing of the CMB fluctuations on small scales due to the finite sized pixels in the map. On the method presented here, the CMB data is not mapped and hence this effect is not present.

The MAP mission is design in such a way that the noise correlations are kept at a minimum. For this case there already exist fast methods for power spectrum estimation (Oh, Spergel, and Hinshaw 1999; Wandelt, Hivon, and Górski 2000). These methods take into account most of the above mentions effect. They do however not take into account a non-symmetric beam shape and far side lobes. If these issues appear to be important then the method presented here is important for exact estimation of the power spectrum.

For the MAP mission and possibly for the Planck mission, the spin axis will be precessing. In this case the centre of the scanning rings will not have the same  $\Theta_E$ . One more rotation has to be added to equation (3.5) describing this precession. For a precessing experiment where the amplitude of the precessions (variations in  $\theta_E$ ) is given by the angle  $\theta_P$ , the formula for the Fourier transform of the signal takes the following form

$$d_{mm'm''}^S = \sum_{\ell} a_{\ell m} d_{mm'}^{\ell}(\theta_E) d_{m'm''}^{\ell}(\theta_P) X_{\ell m''}^* \quad (3.50)$$

Obviously the correlation matrix  $\langle d_{mm'm''} d_{MM'M''} \rangle$  is still block diagonal but the blocks are now larger.

Another problem with the full sky experiment is that parts of the sky (around the galactic plane) has to be cut out due to foreground contamination. This is equivalent to cutting out parts of the rings which destroys the symmetries making the fast likelihood evaluation possible.

The power spectrum estimation method on rings, the way it is described here is clearly for an ideal experiment where there are no foregrounds and no irregularities in the scanning strategy. In reality however this will often not be the case. To deal with these problems the method could be extended using techniques similar to those described in (Oh, Spergel, and Hinshaw 1999) and outlined in section (2.3.1). The idea is that if one knows the exact solution in the ideal situation then the case with cut rings and imperfect scanning strategy can be considered as perturbations around the ideal case. In this way one can find the product  $\mathbf{C}^{-1}\mathbf{d}$  without having to invert  $\mathbf{C}$ . As described in section (2.3.1) this can be done by preconditioned conjugate gradient techniques. By using the approximate  $\mathbf{C}^{-1}$  by the methods described here one can iterate to find the correct  $\mathbf{C}^{-1}\mathbf{d}$  without having to invert the full matrix  $\mathbf{C}$ . This is delayed to future work.



# Chapter 4

## Gabor Transforms on the Sphere and Application to CMB Analysis

In this and the next chapter I will present work done under supervision of Kris Górski. In this chapter the formalism of windowed Fourier transforms known as *Gabor transforms* is extended from the one dimensional line to the sphere. The Gabor coefficients on the CMB sky will be studied and it is shown how these represent a compression of CMB data which allow for fast likelihood estimation.

Gabor transforms, or windowed Fourier transforms are just Fourier transforms where the function  $f(x)$  to be Fourier transformed is multiplied with a *Gabor Window*  $W(x)$  (Gabor 1946). In the discrete case  $f(x_i)$  can be a data set. If parts of the data set are of poor quality or are missing, this can be represented as  $W(x_i)f(x_i)$  where the window  $W$  is zero where there are missing parts. The window can also be formed so that it smooths the edges close to the missing parts and in this way avoid ringing in the Fourier spectrum.

In the first part of this chapter I will study the effect of Gabor transforms on the sphere in particular for the CMB sky. The Gabor transform in this context is just the multiplication of the CMB sky with a window function before using the spherical harmonic transform to get the pseudo power spectrum. The window can be a tophat to take out certain parts of the sky for limited sky coverage. Another window can be a Gaussian Gabor window for smoothing the transition between the observed and unobserved area of the sky. The Gabor window can also be designed in such a way as to reduce signal-to-noise by giving pixels with high signal-to-noise higher significance in the analysis. In the first part I will focus on two types of Gabor windows, the Gaussian window and the tophat window. The effect of these two windows will be compared. For a given patch on the sphere, the effect of multiplying this patch with a Gaussian to smooth the edges instead

of keeping it as a tophat with sharp edges will be discussed. I will show how the Gaussian window cuts of long range correlations between the power spectrum coefficients on the cut sphere (pseudo power spectrum) at the cost of increased short range correlations. The use of the windowed Fourier transform was already studied in (Hobson and Magueijo 1996) in the flat sky approximation. I show that some of their qualitative results are also valid on the sphere.

In (Hivon et al. 2001) the pseudo power spectrum was used as a quadratic estimator for CMB power spectrum analysis. In the second part of the chapter I will discuss the use of the pseudo power spectrum as a compression of a CMB data set to be used in likelihood estimation. This was already done for a tophat window in (Wandelt, Hivon, and Górski 2000) under the assumption that the pseudo spectrum coefficients are not coupled so that the correlation matrix for the likelihood is diagonal. This is only a good assumption when a large part of the sky is covered. As I will show here, the couplings between the pseudo power spectrum coefficients can be quite large depending on the size of the observed area. For this reason the full covariance matrix will be used. Analytical and recursive techniques will be developed which allows one to do fast calculation of the full covariance matrix.

The distribution of the pseudo power spectrum coefficients will be assumed Gaussian. Using a Gaussian likelihood ansatz the full sky power spectrum will be estimated from the noisy pseudo power spectrum. The likelihood function can in this case easily be calculated as the datavector and thereby the correlation matrix are small (only having a few hundred elements even if the original whole data set had millions of pixels). The likelihood evaluation takes of the order  $(N^{bin})^2(N^{in})^2$  where  $N^{bin}$  is the number of  $C_\ell$  bins estimated and  $N^{in}$  is the number of pseudo spectrum coefficients used as input to the likelihood. The most time consuming part is the precomputation of some factors which scales as about  $l_{max}^3$  for the signal part. For the noise correlation matrix this precalculation takes of the order  $\sqrt{N_{pix}l_{max}}(N^{in})^2$  for axissymmetric noise or up to about  $\sqrt{N_{pix}l_{max}^2}(N^{in})^2$  for general noise. Typically  $N^{in} \propto l_{max}$  with a prefactor which is dependent on the size of the window. For a  $18^\circ$  degree radius with Gaussian apodisation which will be used in some examples here  $N^{in} \approx 0.1l_{max}$ . Calculations of noise prefactors can also be done by Monte-Carlo. It then takes of the order  $N_{pix}^{3/2}N_{sim}$ , where  $N_{sim}$  is the number of Monte Carlo simulations needed, which is typically a few thousand.

The method is developed for uncorrelated noise on an axissymmetric patch on the sky. In the last part of the chapter extensions to non-symmetric patches and correlated noise will be worked out and further developments will be discussed.



## 4.1 The Gabor Transformation and the Temperature Power Spectrum

In this section I will first describe the Gabor transform for functions on a one dimensional line. Then I extend the formalism to functions on the sphere and the properties of the Gabor transform coefficients on the sphere are discussed.

### 4.1.1 The One Dimensional Gabor Transform

For a data set  $d_j$  with  $N$  elements, the normal Fourier transform is defined as,

$$\tilde{d}_k = \sum_j d_j e^{i2\pi jk/N}. \quad (4.1)$$

A tilde on  $\tilde{d}$  shows that these are the Fourier coefficients. The inverse transform is then,

$$d_j = \frac{1}{N} \sum_k \tilde{d}_k e^{-i2\pi jk/N}. \quad (4.2)$$

Sometimes it is useful to study the spectrum of just a part of the data set. This could be if parts of it are of poor quality or the spectrum is changing along the data set. In this case, one can multiply the data set with a function, removing the unwanted parts and taking out a segment to be studied. The function can be a step function cutting out the segment to study with sharp edges or a function which smooths the edges of the segment to avoid ringing (typically a Gaussian).

The Fourier transform with such a multiplication was studied by Gabor (Gabor 1946) and is called the *Gabor Transform*. It is defined for a segment centered at  $j = M$  and with wavenumber  $k$  as,

$$\tilde{d}_{kM} = \sum_j d_j G_{j-M} e^{i2\pi jk/N}. \quad (4.3)$$

Here  $G_{j-M}$  is the *Gabor Window*, the function to multiply the data set with. The transform is similar to the Wavelet transform. The difference is that the window function in the Wavelet transform is frequency dependent so that the size of the segment is changing with frequency.

Analogously to the Fourier transform, there is also an inverse Gabor transform. To recover the whole data set from a Gabor transform, one needs the Fourier coefficients taken with different windows  $G_{j-M}$  being centered at several different points  $M$ . This means that the data set has to be split up into several segments. The center of each segment is set to  $M = mK$  where  $K$  determines

the density of segments and  $m$  is an integer specifying the segment number. One then has for the inverse transform

$$d_j = \sum_m \sum_k \tilde{d}_{km} g_{km}. \quad (4.4)$$

Due to the non-orthogonality of the Gabor transform, the *dual Gabor window*  $g_{km}$  is not trivial to find, but several techniques have been developed for calculating this dual window (e.g. (Strohmer 1997) and references therein).

In this chapter I will study the Gabor transform on the sphere and apply it to CMB analysis. I will take out a disc on the CMB sky, using either tophat or Gaussian apodisation and then create the pseudo power spectrum  $\tilde{C}_\ell$  on this apodised sky. These  $\tilde{C}_\ell$  will be used for likelihood estimation of the underlying full sky power spectrum. I also outline how several discs (segments) centered at different points can be combined to yield the full sky power spectrum.

### 4.1.2 Gabor Transform on the Sphere

I start by defining the  $\tilde{C}_\ell$  for a Gabor Window  $G(\hat{\mathbf{n}})$  as,

$$\tilde{C}_\ell = \sum_m \frac{\tilde{a}_{\ell m}^* \tilde{a}_{\ell m}}{2\ell + 1}, \quad (4.5)$$

where

$$\tilde{a}_{\ell m} = \int d\hat{\mathbf{n}} T(\hat{\mathbf{n}}) G(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}). \quad (4.6)$$

Here  $T(\hat{\mathbf{n}})$  is the observed temperature in the direction of the unit vector  $\hat{\mathbf{n}}$ ,  $Y_{\ell m}(\hat{\mathbf{n}})$  is the spherical Harmonic function and  $G(\hat{\mathbf{n}})$  is the Gabor window.

I will here use a Gabor Window which is azimuthally symmetric about a point  $\hat{\mathbf{n}}_0$  on the sphere, so that the window is only a function of the angular distance from this point on the sphere  $\cos \theta = \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}_0$ . Then one can write the Legendre expansion of the window as,

$$G(\theta) = \sum_\ell \frac{2\ell + 1}{4\pi} g_\ell P_\ell(\cos \theta) = \sum_{\ell m} g_\ell Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}_0). \quad (4.7)$$

One can also write,

$$T(\hat{\mathbf{n}}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}). \quad (4.8)$$

Inserting these two expressions in equation (4.6) one gets

$$\tilde{a}_{\ell m} = \sum_{\ell' m'} a_{\ell' m'} \sum_{\ell'' m''} g_{\ell''} Y_{\ell'' m''}^*(\mathbf{n}_0) \int Y_{\ell m}^*(\hat{\mathbf{n}}) Y_{\ell' m'}(\hat{\mathbf{n}}) Y_{\ell'' m''}(\hat{\mathbf{n}}) d\hat{\mathbf{n}} \quad (4.9)$$

$$= \sum_{\ell' m'} a_{\ell' m'} \sum_{\ell'' m''} g_{\ell''} Y_{\ell'' m''}^*(\hat{\mathbf{n}}_0) \sqrt{\frac{(2\ell + 1)(2\ell' + 1)(2\ell'' + 1)}{4\pi}} \quad (4.10)$$

$$\times \begin{pmatrix} \ell & \ell' & \ell'' \\ -m & m' & m'' \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} (-1)^m, \quad (4.11)$$

where relation (C.3) for Wigner 3j Symbols were used. Using this expression, the relation  $\langle a_{\ell m}^* a_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'}$  and the orthogonality of Wigner symbols (equation (C.1)), one can write  $\langle \tilde{C}_\ell \rangle$  as,

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} C_{\ell'} K(\ell, \ell'). \quad (4.12)$$

With  $C_\ell$  I will always mean  $\langle C_\ell \rangle$  when I am referring to the full sky  $C_\ell$ . In this expression,  $K(\ell, \ell')$  is the Gabor Kernel,

$$\begin{aligned} K(\ell, \ell') &= (2\ell' + 1) \sum_{\ell'' m''} g_{\ell''}^2 \frac{1}{4\pi} |Y_{\ell'' m''}^2(\hat{\mathbf{n}}_0)|^2 \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 \\ &= (2\ell' + 1) \sum_{\ell''} g_{\ell''}^2 \frac{(2\ell'' + 1)}{(4\pi)^2} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 \end{aligned} \quad (4.13)$$

The Legendre coefficients  $g_\ell$ , are found by the inverse Legendre transformation,

$$g_\ell = 2\pi \int_{\theta=0}^{\theta=\theta_C} G(\theta) P_\ell(\cos \theta) d \cos \theta, \quad (4.14)$$

where  $\theta_C$  is the cut-off angle where the window goes to zero. One sees from the expression for the kernel, that there is no dependency on  $\hat{\mathbf{n}}_0$ . This means that  $\langle \tilde{C}_\ell \rangle$  is the same, independent on where the Gabor window is centered.

In Figure (4.1) I have plotted the kernel for a Gaussian Gabor window,

$$G(\theta) = e^{-\theta^2/(2\sigma^2)} \quad \theta \leq \theta_C, \quad (4.15)$$

$$G(\theta) = 0 \quad \theta > \theta_C, \quad (4.16)$$

with 5 and 15 degrees FWHM (corresponding to  $\sigma = 2.12^\circ$  and  $\sigma = 6.38^\circ$ ) and  $\theta_C = 3\sigma$ . One sees that the kernel is centred about  $\ell = \ell'$ , and falls off rapidly. Figure (4.2) shows the same for the corresponding tophat Gabor windows,

$$G(\theta) = 1 \quad \theta \leq \theta_C, \quad (4.17)$$

$$G(\theta) = 0 \quad \theta > \theta_C. \quad (4.18)$$

The tophat windows are covering the same area on the sky as the corresponding Gaussian windows in figure (4.1) ( $\theta_C$  is the same). One sees clearly that the diagonal is broader for the smaller windows indicating stronger couplings. Another thing to notice is that whereas the kernel for the tophat Gabor window only falls by about 4 orders of magnitude from the diagonal to the far off-diagonal elements, the Gaussian Gabor kernel falls by about 8 orders of magnitude (the vertical axis on the four plots are the same). The smooth cut-off of the Gaussian Gabor Window cuts off long range correlations in spherical harmonic space. One of the aims of the first part of this paper is to see how the pseudo power spectrum of a given disc on the sky (tophat window) is affected by the multiplication with

a Gaussian Gabor window. For this reason the pseudo spectrum will be studied for a tophat and a Gaussian always covering the same area on the sky.

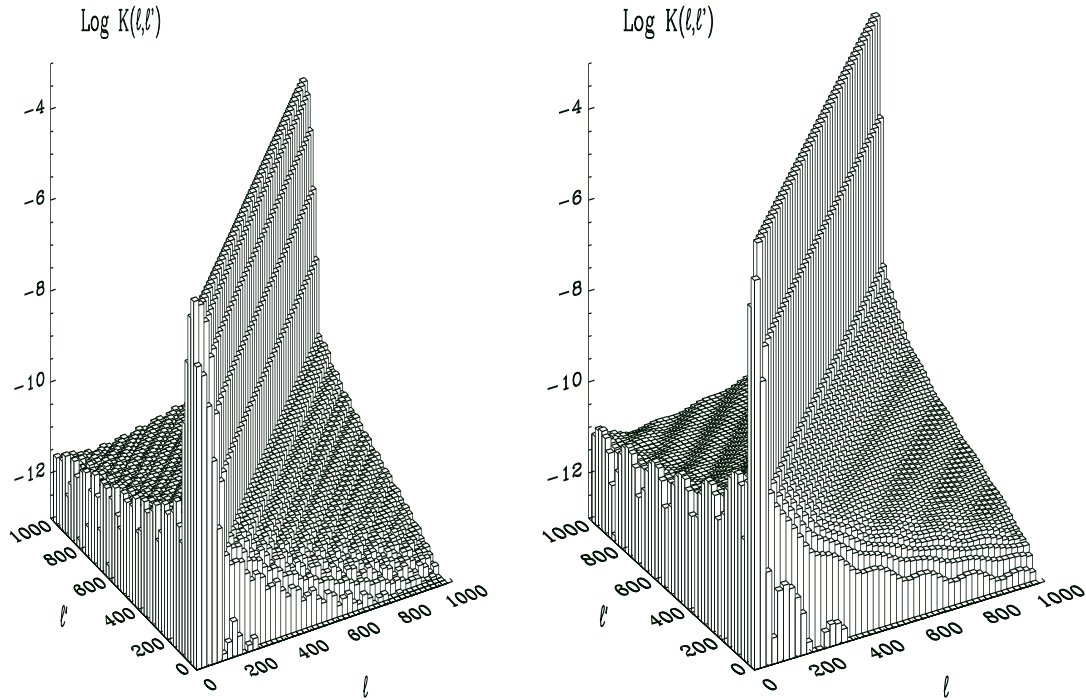


Figure 4.1: The logarithm of the kernel  $K(\ell, \ell')$  describing the connection between the spherical harmonic coefficients  $C_\ell$  on the full sky and the corresponding coefficients  $\tilde{C}_\ell$  on the apodised sky via the relation  $\tilde{C}_\ell = \sum_{\ell'} K(\ell, \ell') C_{\ell'}$ . The figure shows the kernel for a  $5^\circ$  and  $15^\circ$  FWHM Gaussian Gabor window with  $\theta_C = 3\sigma$  (left and right respectively).

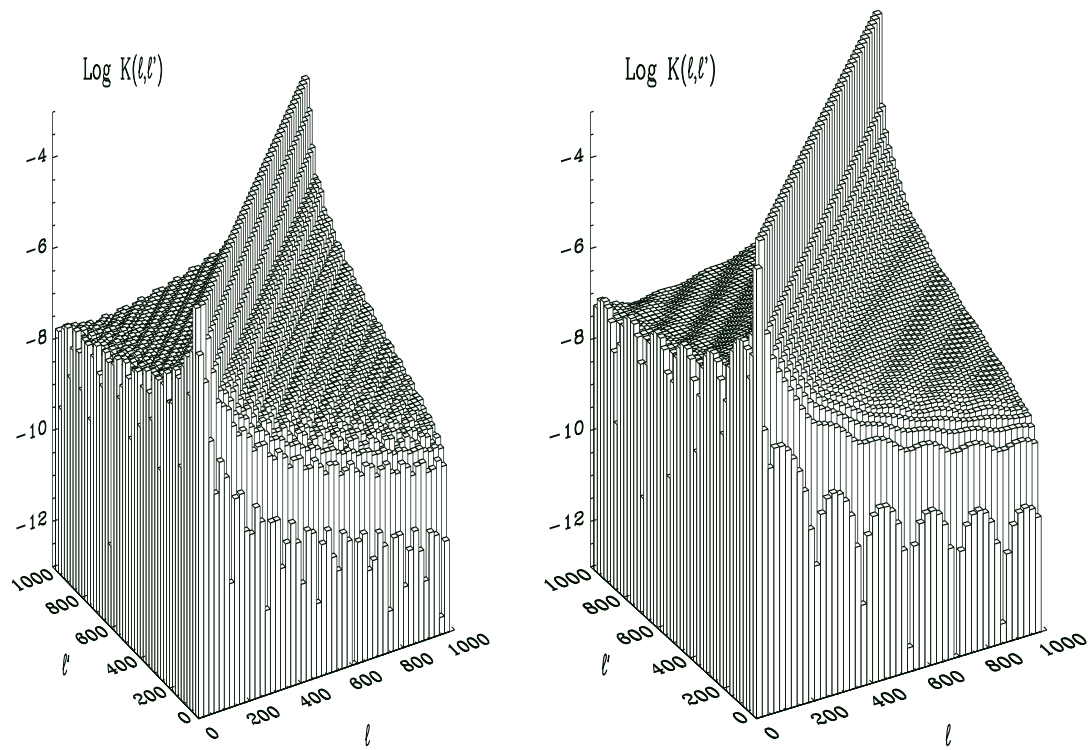


Figure 4.2: Same as figure (4.1) for tophat Gabor windows covering the same area on the sky as the Gaussian windows.

In figure (4.3), I have plotted cuts through the kernel at  $\ell = 200$  and  $\ell = 500$  for the 5 and 15 degree FWHM Gaussian Gabor windows (dashed line). The solid line is the corresponding kernel (same area on the sky) when using the tophat Gabor window. One sees that the Gaussian window effectively cuts off long range correlations whereas the tophat window is narrower close to the diagonal. The Gaussian window has larger short range correlations.

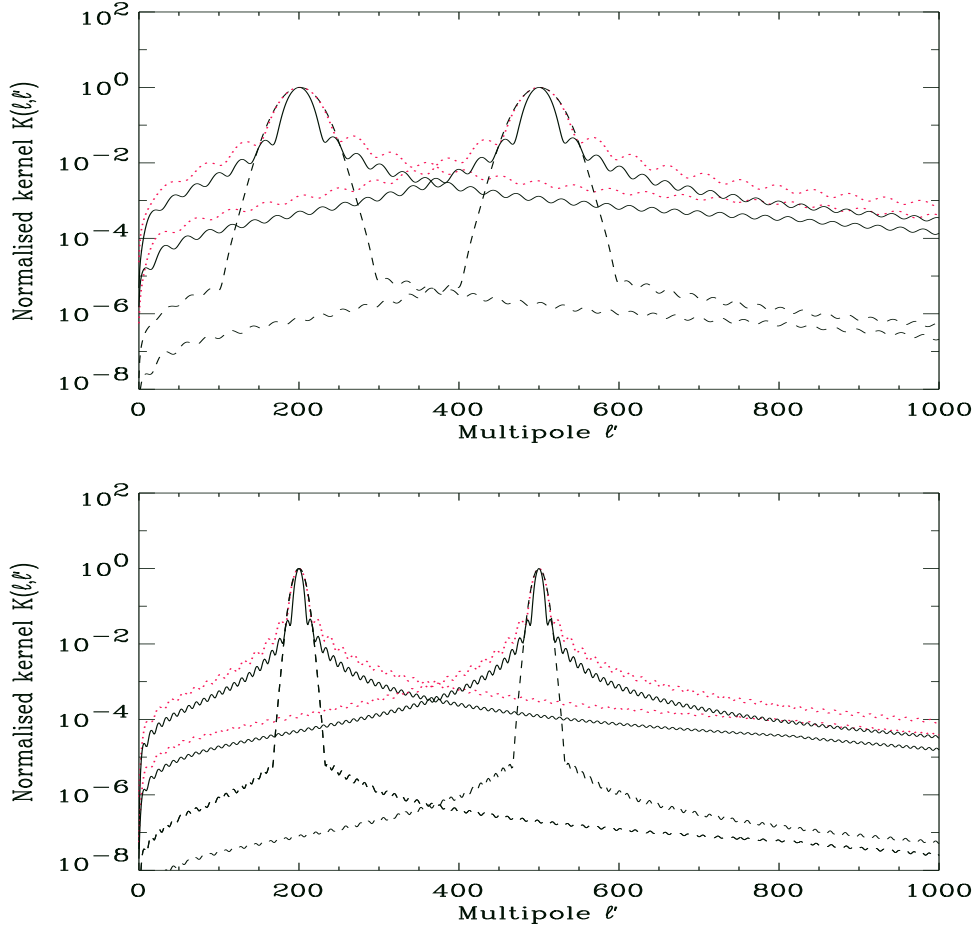


Figure 4.3: The figures show cuts through the kernel  $K(\ell, \ell')$  connecting the full sky and cut sky spherical harmonic coefficients. The full kernels are shown in figures (4.1) and (4.2). The cuts are taken at  $\ell = 200$  and  $\ell = 500$  for the 5 (upper plot) and 15 (lower plot) degree Gaussian Gabor window (dashed line). The solid line is for the corresponding (same area on the sky) tophat window. The kernels are here normalised so that the peak at the given cut has its maximum at 1. In this way one can easier compare the kernels.

Figure (4.4) shows how the width of the kernel gets narrower and the correlations smaller as the Gabor window opens up. The four kernels are shown for  $\ell = 500$  and the Gaussian windows have 5, 10, 15 and 30 degree FWHM with  $\theta_C = 3\sigma$ . The same relation for the corresponding (in the sense described above) tophat window is also shown. Gaussian functions are plotted on top of the kernels and show that the kernels are very close to Gaussian functions near the diagonal.

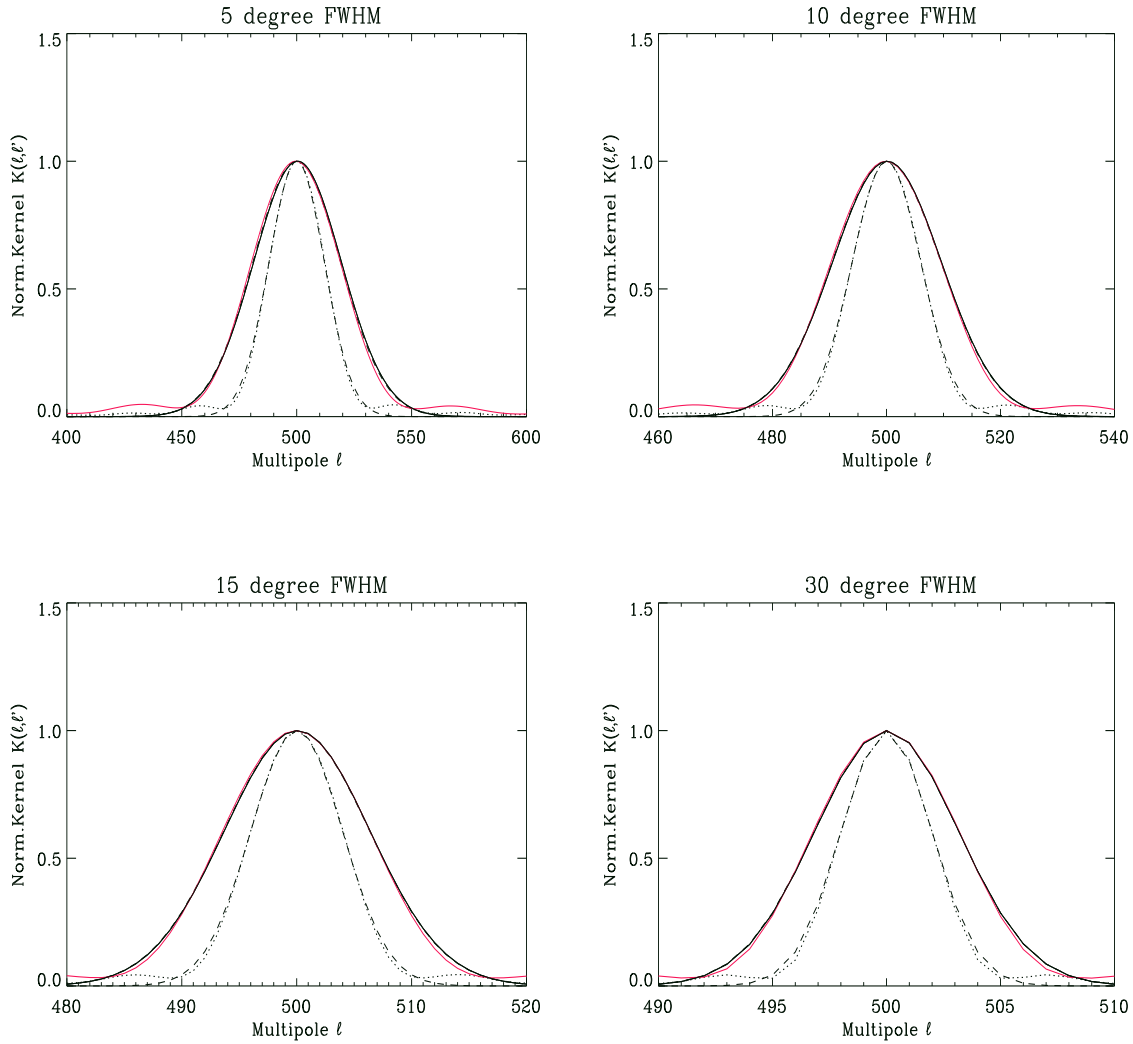


Figure 4.4: The figures show a cut through the kernel  $K(\ell, \ell')$  connecting the full sky and cut sky spherical harmonic coefficients. The cuts are taken at  $\ell = 500$  for a 5, 10, 15 and 30 degree FWHM Gaussian Gabor window (solid line) with a  $\theta_C = 3\sigma$  cut-off. The dotted line shows the kernel for a tophat window covering the same area on the sky. Dashed lines are Gaussian fits to the curves.

In figure (4.5) I have plotted the relation between the FWHM width  $\Delta\ell$  of the kernel and the size  $\Delta\theta$  of the window for Gaussian and tophat windows. The two curves are very well described by  $\Delta\ell = 220/\theta_{FWHM}$  for the Gaussian window ( $\theta_{FWHM}$  in degrees) and  $\Delta\ell = 140/\theta_{FWHM}$  for the corresponding tophat window. Clearly for a given observed area of the sky, multiplying with a Gaussian will increase the FWHM of the kernel. This is also what was seen in figure (4.3)

and (4.4).

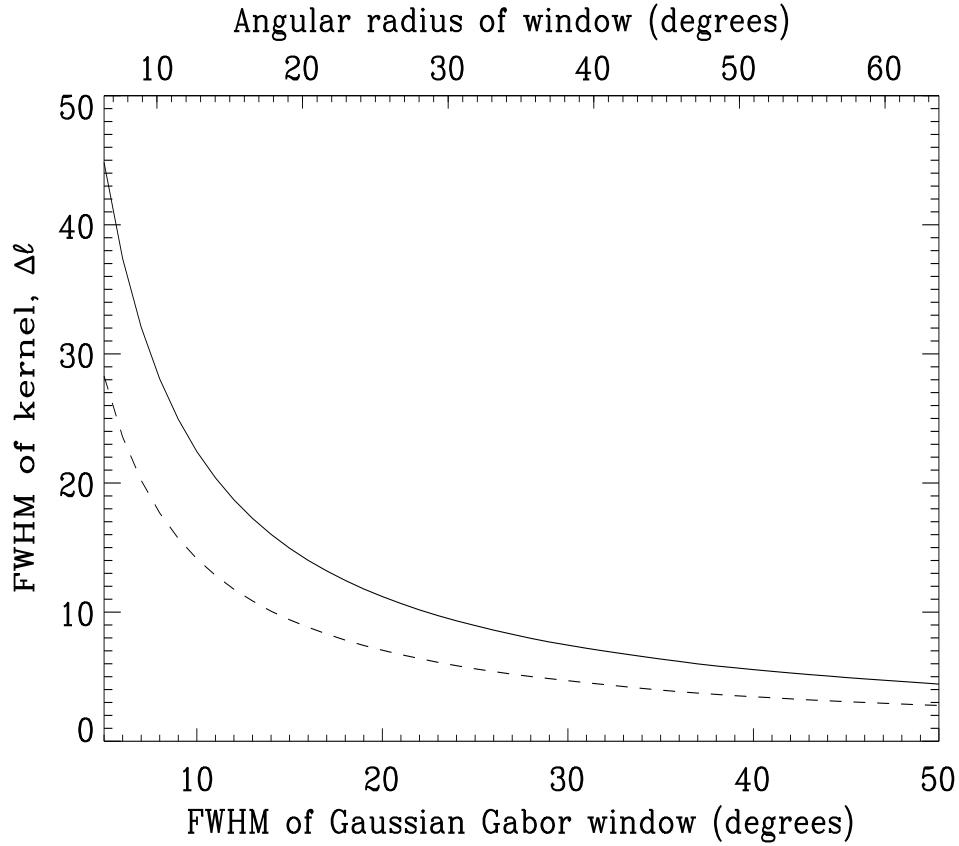


Figure 4.5: The figure shows the uncertainty relation  $\Delta\ell\Delta\theta = \text{const}$  for a Gabor transform on the sphere. The solid line shows the width  $\Delta\ell$  of the Gabor kernel  $K(\ell, \ell')$  connecting the full sky and the cut sky power spectra when applying a Gaussian Gabor window with a cut  $\theta_C = 3\sigma$ . The FWHM is showed on the lower abcissa. The dashed line shows the width of the kernel for a tophat window covering the same area of the sky as the Gaussian. The full radius of the tophat window is shown on the upper abcissa. The curves are well described by  $\Delta\ell = 220/\theta_{FWHM}$  and  $\Delta\ell = 140/\theta_{FWHM}$  for the Gaussian and tophat windows respectively.

In Figure (4.6), I show the shapes of the  $\tilde{C}_\ell$  for Gaussian and tophat windows compared to the full sky spectrum. The plots which were made using the analytical formula (4.12) show  $\tilde{C}_\ell$  for a 5 and 15 degree FWHM Gaussian Gabor window (solid line) cut at  $3\sigma$ . The corresponding spectra for the tophat Gabor windows are shown as dotted lines. The spectra are normalised in such a way that



they can be compared to the full sky power spectrum (dashed line). For the  $5^\circ$  FWHM window one can still distinguish the three lines. At this window size the pseudo spectra are very similar to the full sky spectra but with small deviations depending on the shapes of the kernel and the shape of the power spectrum. In this case the spectrum for the Gaussian window seems to be smaller at the peaks and larger at the troughs whereas the spectrum for the tophat window is always larger. For the  $15^\circ$  FWHM windows the pseudo spectrum using the Gaussian Gabor window are on top of the full sky power spectrum. For the tophat window it is still possible to distinguish the pseudo spectrum from the full sky power spectrum although the lines are still very close. This means that the  $\tilde{C}_\ell$  could be good estimators of the underlying full sky  $C_\ell$  provided that the window is big enough. This is consistent with the conclusion in (Hobson and Magueijo 1996), where this was shown in the flat sky approximation.

One feature which is very prominent is the additional peak at low  $\ell$  for the Gaussian window. The reason for this peak comes from the fact that the diagonal in the Gaussian kernel is broader than in the tophat kernel. For the low multipoles the power spectrum is dropping rapidly because of the Sachs-Wolfe effect and the lowest multipole  $C_\ell$  are much bigger than the  $C_\ell$  for higher multipoles. Since the Gaussian kernel is broad, the  $\tilde{C}_\ell$  at low multipoles will pick up more from the  $C_\ell$  at lower multipoles than the narrower tophat kernel (see figure (4.3)). These low multipole  $C_\ell$  have very high values compared to the higher multipole  $C_\ell$  and for that reason the  $\tilde{C}_\ell$  for the Gaussian window will get a higher value. This is illustrated in figure (4.7) where a cut through the kernel at  $\ell = 50$  is shown for the  $5^\circ$  FWHM Gaussian Gabor window (solid line) and the corresponding tophat (dashed line) normalised to one at the peak. The dotted line shows a typical power spectrum. Clearly the Gaussian kernel will pick up more of the high value  $C_\ell$  at low multipoles.

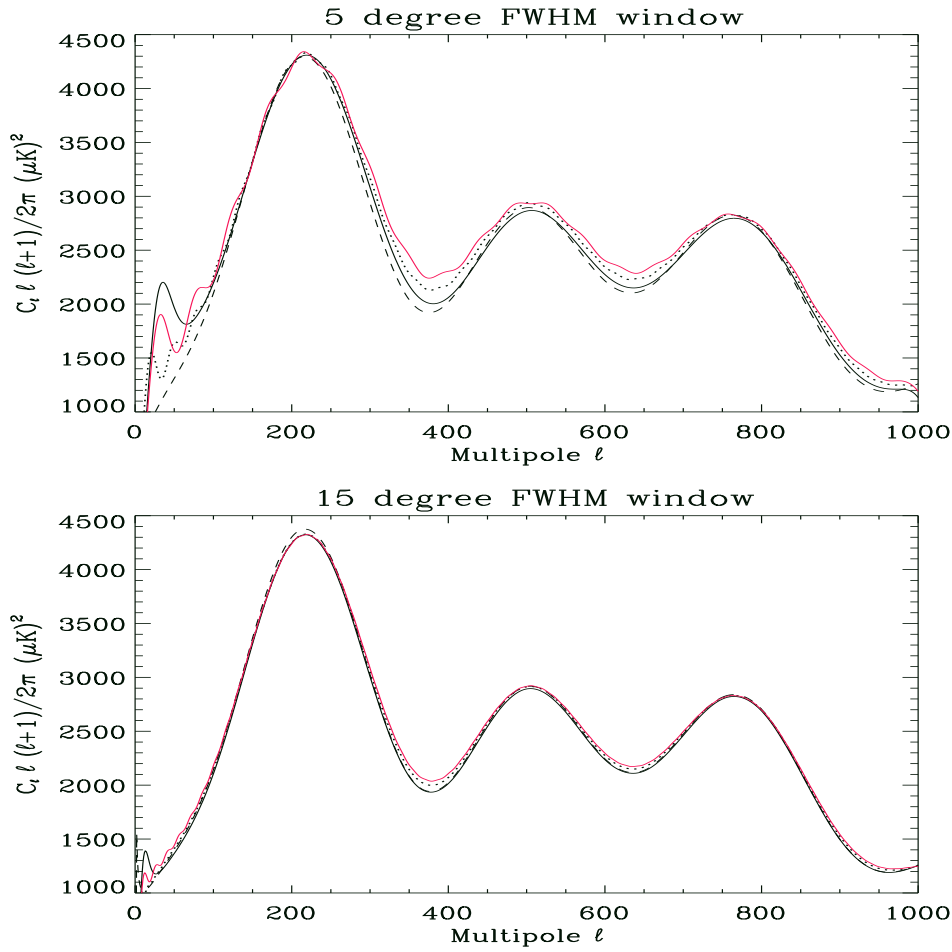


Figure 4.6: The windowed power spectra  $\tilde{C}_l$  for a 5 and 15 degree FWHM Gaussian Gabor window cut at  $\theta_C = 3\sigma$  (solid line) and for a tophat window covering the same area on the sky (dotted line). All spectra are normalised in such a way that they can be compared directly with the full sky spectrum which is shown on each plot as a dashed line. Only in the first plot are all three lines visible. In the three last plots, the full sky spectrum and the Gaussian pseudo spectrum (dashed and solid line) are only distinguishable in the first few multipoles.

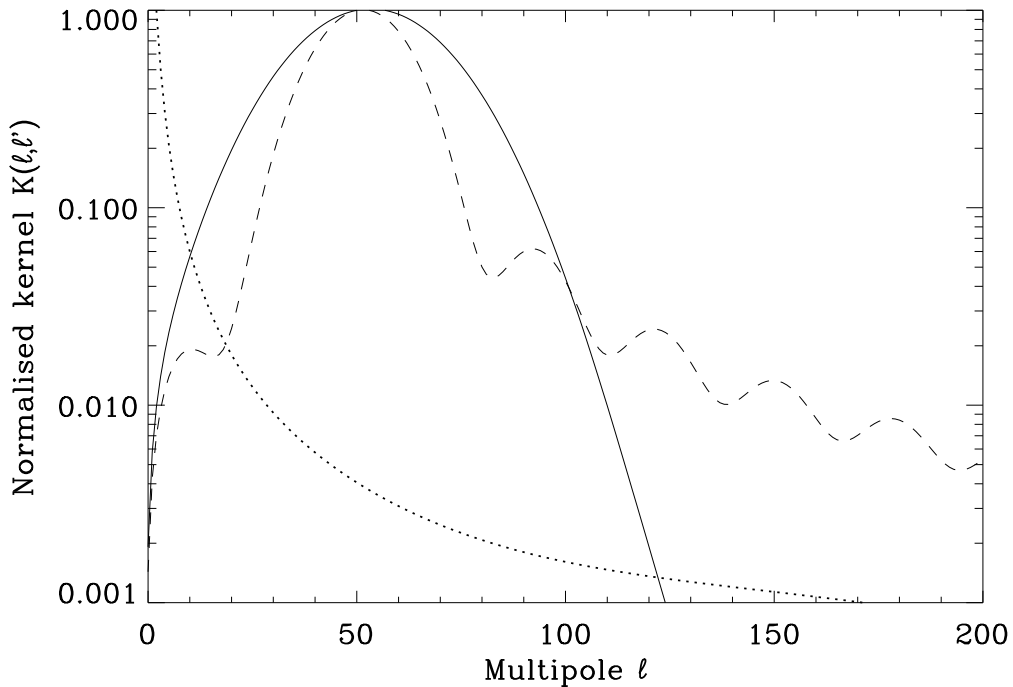


Figure 4.7: The figure show a cut through the kernel  $K(\ell, \ell')$  connecting the full sky and cut sky spherical harmonic coefficients. The cut is taken at  $\ell = 50$  for a 5 degree FWHM Gaussian Gabor window (solid line) and a corresponding tophat window (dashed line). The kernels are normalised to one at the peak. A typical power spectrum normalised to one at the quadrupole is plotted as a dotted line. The figure aims at explaining the extra peak in the pseudo power spectrum at low multipoles for the Gaussian Gabor window shown in figure (4.6).

In figure (4.8) I show the pseudo power spectra for a particular realisation using a 15 degree FWHM Gaussian window (upper plot) and a tophat window (lower plot). The pseudo spectra are compared to the average full sky spectra shown as a dashed line. The dark shaded area shows the expected  $1\sigma$  cosmic and sample variance on the pseudo spectra taken from the formulae to be developed in the next sections. The lighter shaded area shows only cosmic variance. Note that the pseudo spectrum for the Gaussian window is smoother than the pseudo spectrum for the tophat window. This is again a result of the broader Gaussian kernel.

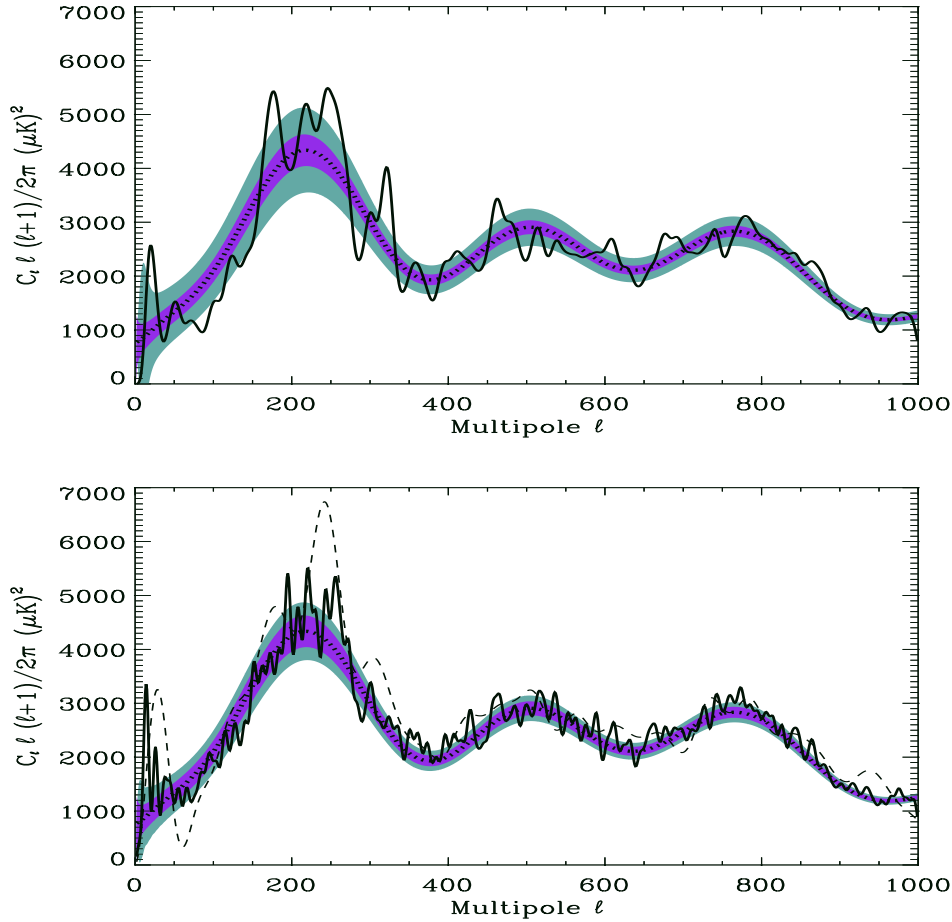


Figure 4.8: One realisation of the windowed power spectra. The upper plot shows a realisation of a pseudo power spectrum using a 15 degree FWHM Gaussian Gabor window. The pseudo spectrum is normalised in such a way that it can be compared directly to the full sky spectrum which average is shown as a dotted line. The lower plot shows the same realisation using a corresponding tophat window. The light shaded area shows  $1\sigma$  cosmic variance around the full sky average spectrum. The darker area shows  $1\sigma$  cosmic and sampling variance taken from the theoretical formula.

### 4.1.3 Rotational Invariance

It was shown that the average  $\langle \tilde{C}_\ell \rangle$  is invariant under rotations of the Gabor window. I will now show that the non averaged  $\tilde{C}_\ell$  are rotationally invariant under any rotation of the sky AND Gabor window by the same angle. This fact justifies that I in the following always put the window on the north pole since this simplifies the calculations. In the following I will use the rotation matrices

$D_{mm'}^\ell$  described in Appendix (A). Consider a rotation of the sky and window by the angles  $(-\gamma - \beta - \alpha)$ . Then the  $\tilde{a}_{\ell m}$  becomes,

$$\tilde{a}_{\ell m}^{\text{rot}} = \int d\hat{\mathbf{n}}[\hat{D}(-\gamma - \beta - \alpha)T(\hat{\mathbf{n}})G(\hat{\mathbf{n}})]Y_{\ell m}^*(\hat{\mathbf{n}}). \quad (4.19)$$

If one makes the inverse rotation of the integration angle  $\hat{\mathbf{n}}$ , one can write this as

$$\tilde{a}_{\ell m}^{\text{rot}} = \int d\hat{\mathbf{n}}T(\hat{\mathbf{n}})G(\hat{\mathbf{n}})[\hat{D}^*(\alpha\beta\gamma)Y_{\ell m}^*(\hat{\mathbf{n}})], \quad (4.20)$$

which is just

$$\tilde{a}_{\ell m}^{\text{rot}} = \sum_{m'} D_{m'm}^{\ell*}(\alpha\beta\gamma) \int d\hat{\mathbf{n}}T(\hat{\mathbf{n}})G(\hat{\mathbf{n}})Y_{\ell m'}^*(\hat{\mathbf{n}}). \quad (4.21)$$

One can identify the last integral as the normal  $\tilde{a}_{\ell m}$ .

$$\tilde{a}_{\ell m}^{\text{rot}} = \sum_{m'} D_{m'm}^{\ell*}(\alpha\beta\gamma)\tilde{a}_{\ell m}. \quad (4.22)$$

So the  $\tilde{a}_{\ell m}$  are NOT rotationally invariant. Rotation mixes  $m$ -modes for a given  $\ell$ -value.

Now to the  $\tilde{C}_\ell$ . One has that

$$\tilde{C}_\ell^{\text{rot}} = \frac{1}{2\ell + 1} \sum_m a_{\ell m}^{\text{rot}} a_{\ell m}^{\text{rot}*} \quad (4.23)$$

$$= \frac{1}{2\ell + 1} \sum_m \sum_{m'} \sum_{m''} D_{m'm}^\ell(\alpha\beta\gamma) D_{m''m}^{\ell*}(\alpha\beta\gamma) \tilde{a}_{\ell m'} \tilde{a}_{\ell m''}^* \quad (4.24)$$

$$= \frac{1}{2\ell + 1} \sum_{m'm''} \tilde{a}_{\ell m'} \tilde{a}_{\ell m''}^* \sum_m D_{m'm}^\ell(\alpha\beta\gamma) D_{m''m}^{\ell*}(\alpha\beta\gamma). \quad (4.25)$$

Using the properties given in Appendix (A), one can write the last D-function on the last line as,

$$\hat{D}_{m''m}^{\ell*}(\alpha\beta\gamma) = D_{m''m}^\ell(-\alpha\beta - \gamma) \quad (4.26)$$

$$= e^{-im''\alpha} d_{m''m}^\ell(\beta) e^{-im\gamma} \quad (4.27)$$

$$= e^{-im''\alpha} d_{mm''}^\ell(-\beta) e^{-im\gamma} \quad (4.28)$$

$$= D_{mm''}^\ell(-\gamma - \beta - \alpha). \quad (4.29)$$

Knowing that  $(-\gamma - \beta - \alpha)$  is the inverse rotation of  $(\alpha\beta\gamma)$  one can write,

$$\sum_m D_{m'm}^\ell(\alpha\beta\gamma) D_{m''m}^{\ell*}(\alpha\beta\gamma) = \sum_m D_{m'm}^\ell(\alpha\beta\gamma) D_{mm''}^\ell(-\gamma - \beta - \alpha) \quad (4.30)$$

$$= D_{m'm''}^\ell(000) = \delta_{m'm''} \quad (4.31)$$

So one gets,

$$\tilde{C}_\ell^{\text{rot}} = \tilde{C}_\ell. \quad (4.32)$$

## 4.2 Likelihood Analysis

In this section I will show how the pseudo power spectrum can be used as input to a likelihood analysis for estimating the full sky power spectrum from an observed disc on the sky multiplied with a Gabor window. I will in this section concentrate on a Gaussian Gabor window, but the formalism is valid for any azimuthally symmetric Gabor window.

### 4.2.1 The Form of the Likelihood Function

To know the form of the likelihood function, one needs to know the probability distribution of  $\tilde{C}_\ell$ . In Figure (4.9) and (4.10) I show the probability distribution from 10000 simulations with a  $5^\circ$  and  $15^\circ$  FWHM Gaussian Gabor window respectively. The dashed line shows a Gaussian with mean value and standard deviation found from the formulae given in the previous and next section. One can see that the probability distribution is slightly skewed for low  $\ell$ , but for high  $\ell$  it seems to be very well approximated by a Gaussian. Also the small window shows more deviations from a Gaussian than the bigger window. In figure (4.11) I show the probability distribution from a simulation with a tophat Gabor window covering the same area on the sky as the  $15^\circ$  FWHM Gaussian window. Also for this window the probability distribution is close to Gaussian.

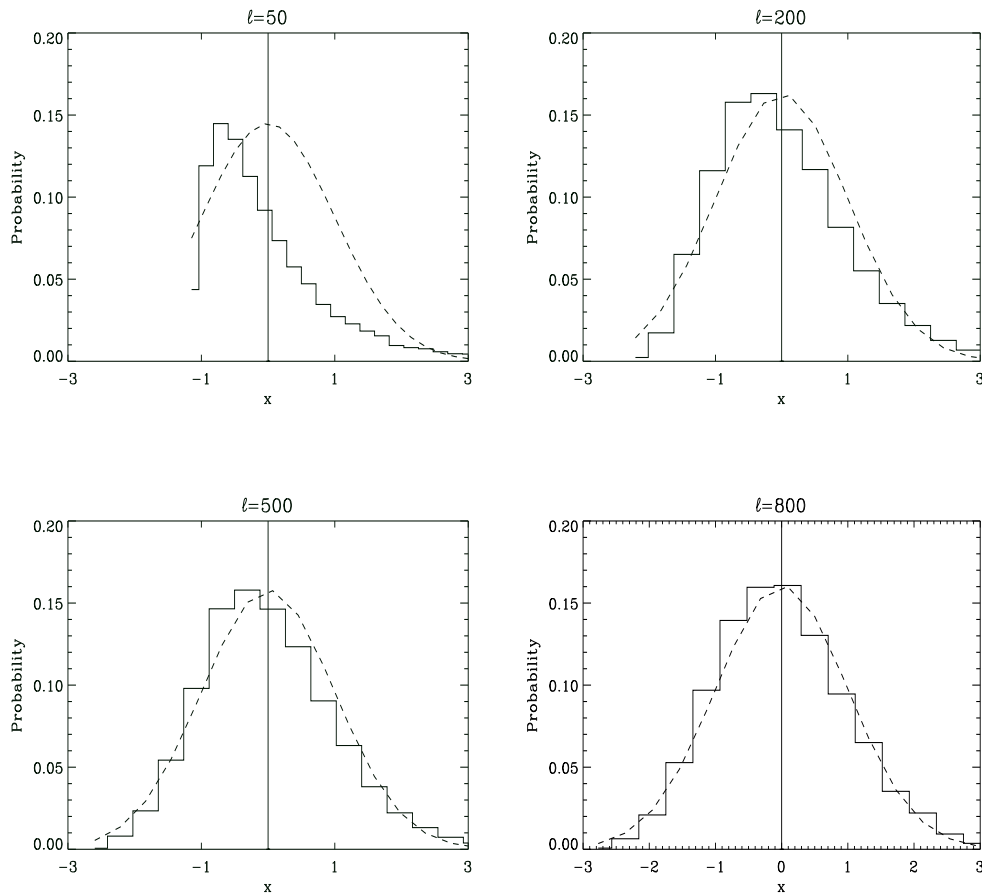


Figure 4.9: The probability distribution of  $\tilde{C}_\ell$  taken from 10000 simulations with a  $5^\circ$  FWHM Gaussian Gabor window truncated at  $\theta_C = 3\sigma$ . The variable  $x$  is given as  $x = (\tilde{C}_\ell - \langle \tilde{C}_\ell \rangle) / \sqrt{\langle (\tilde{C}_\ell - \langle \tilde{C}_\ell \rangle)^2 \rangle}$ . The dashed line is a Gaussian with the theoretical mean and standard deviation of the  $\tilde{C}_\ell$ . The plot shows the  $\tilde{C}_\ell$  distribution for  $\ell = 50$ ,  $\ell = 200$ ,  $\ell = 500$ , and  $\ell = 800$ . The probabilities are normalised such that the integral over  $x$  is 1.

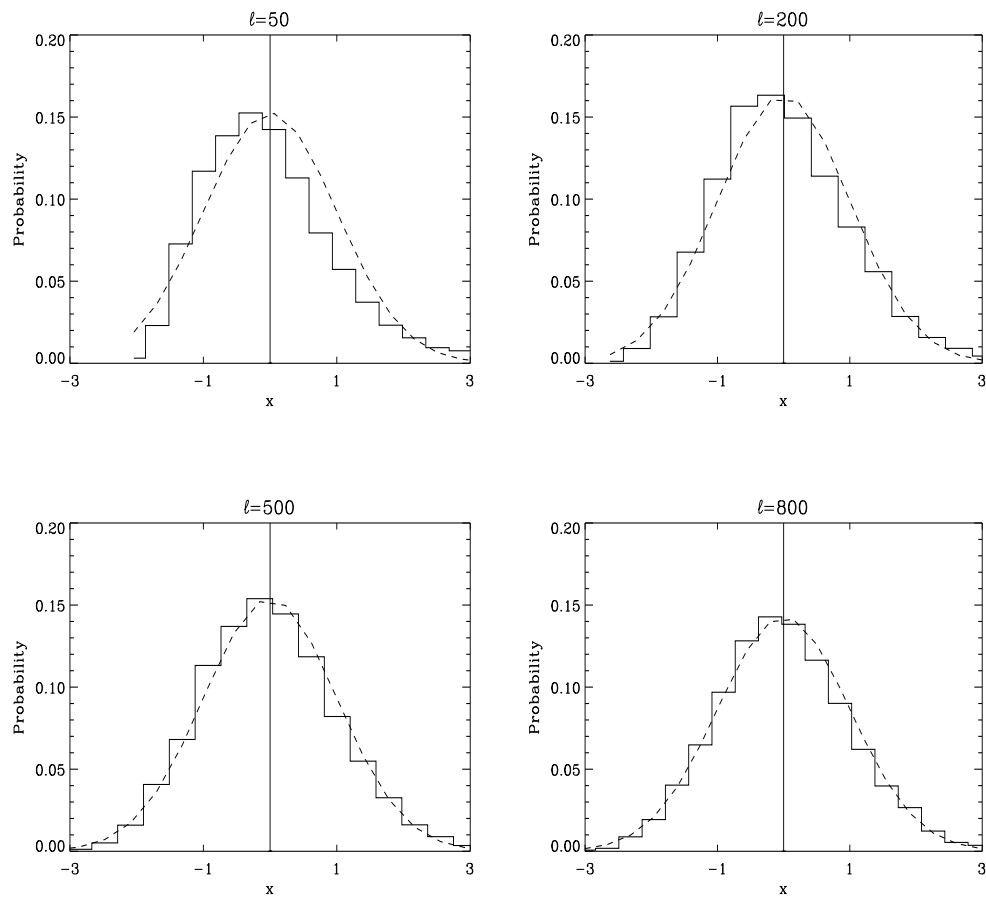


Figure 4.10: Same as figure (4.9) but for a  $15^\circ$  FWHM Gaussian Gabor window.



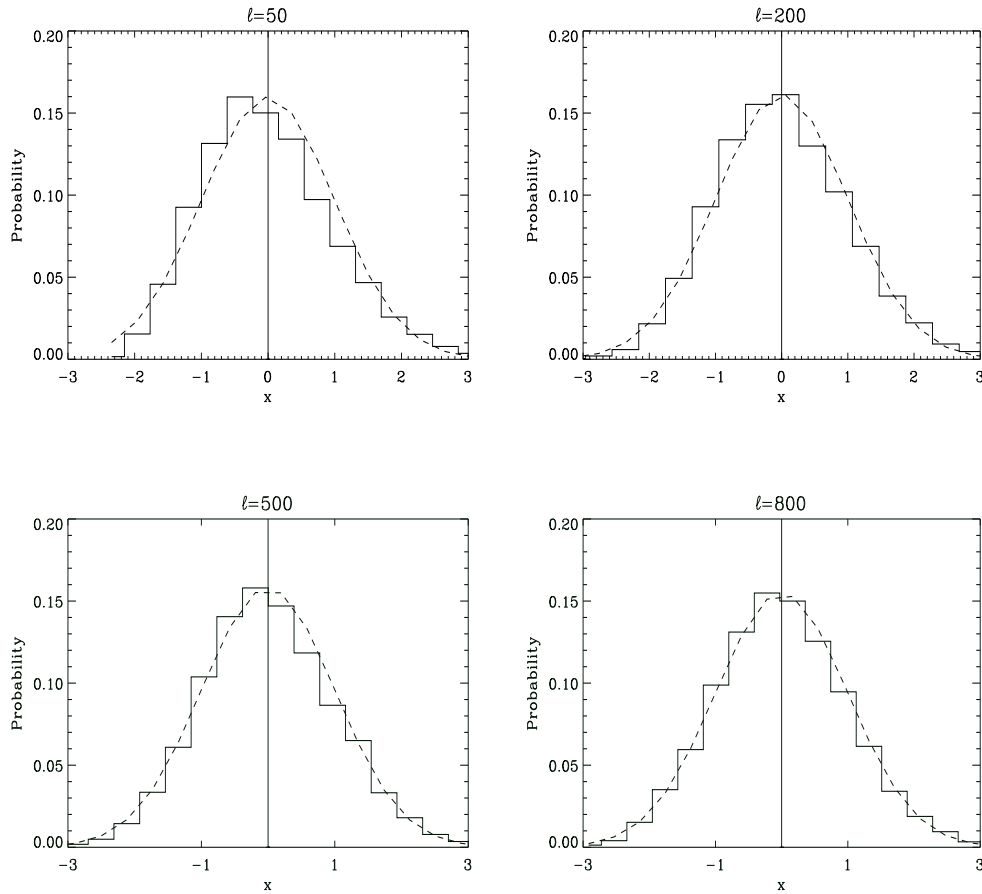


Figure 4.11: Same as figure (4.10) but for a tophat Gabor window covering the same area on the sky.

From the above plots it seems to be reasonable to approximate the likelihood function with a Gaussian provided the window is big enough and multipoles at high enough  $\ell$  values are used,

$$\mathcal{L} = \frac{e^{-\frac{1}{2}\mathbf{d}^T\mathbf{M}^{-1}\mathbf{d}}}{\sqrt{2\pi \det \mathbf{M}}}. \quad (4.33)$$

Omitting all constant terms and factors, the log-likelihood can then be written:

$$L = \mathbf{d}^T\mathbf{M}^{-1}\mathbf{d} + \ln \det \mathbf{M}. \quad (4.34)$$

Here  $\mathbf{d}$  is the datavector which contains the observed  $\tilde{C}_\ell$  for the set of sample  $\ell$ -values  $\ell_i$ . The datavector is taken from the observed windowed sky in the

following way:

$$d_i = \tilde{C}_{\ell_i} - \langle \tilde{C}_{\ell_i} \rangle. \quad (4.35)$$

The matrix  $\mathbf{M}$  is the covariance between pseudo- $C_l$  which elements are given by:

$$M_{ij} = \langle \tilde{C}_{\ell_i} \tilde{C}_{\ell_j} \rangle - \langle \tilde{C}_{\ell_i} \rangle \langle \tilde{C}_{\ell_j} \rangle. \quad (4.36)$$

The next step is to find an expression for  $M_{ij}$  first assuming that no noise is present.

### 4.2.2 The Correlation Matrix

To do fast likelihood analysis with  $\tilde{C}_\ell$  one needs to be able to calculate  $\langle \tilde{C}_\ell \rangle$  and the correlations  $\langle \tilde{C}_\ell \tilde{C}_{\ell'} \rangle$  fast. Calculating the average  $\langle \tilde{C}_\ell \rangle$  by formula (4.12) using the analytic expression (4.13) for the kernel is not very fast. It turns out that a faster way of evaluating the kernel is by using direct integration (summation on the pixelised sphere) and then, as shown in Appendix (D), recurrence. By means of an integral, one can then write the  $\tilde{a}_{\ell m}$  as (now assuming that  $\hat{\mathbf{n}}_0$  is on the north pole),

$$\begin{aligned} \tilde{a}_{\ell m} &= \sum_{\ell' m'} a_{\ell' m'} \int G(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}) Y_{\ell' m'}(\hat{\mathbf{n}}) d\hat{\mathbf{n}} \\ &= \sum_{\ell' m'} a_{\ell' m'} \int G(\theta) \lambda_{\ell m}(\theta) \lambda_{\ell' m'}(\theta) d \cos \theta \underbrace{\int e^{-i\phi(m-m')} d\phi}_{2\pi\delta_{mm'}} \\ &= \sum_{\ell'} a_{\ell' m} 2\pi \int G(\theta) \lambda_{\ell m}(\theta) \lambda_{\ell' m}(\theta) d \cos \theta \\ &\equiv \sum_{\ell'} a_{\ell' m} h(\ell, \ell', m), \end{aligned} \quad (4.37)$$

where the last line defines  $h(\ell, \ell', m)$  and  $\lambda_{\ell m}(\theta)$  is given by,

$$Y_{\ell m}(\theta, \phi) = \lambda_{\ell m}(\theta) e^{-i\phi m}. \quad (4.38)$$

Using this form, one gets,

$$\langle \tilde{C}_\ell \rangle = \frac{1}{2\ell + 1} \sum_{\ell'} C_{\ell'} \sum_m h^2(\ell, \ell', m). \quad (4.39)$$

To obtain this expression,  $\hat{\mathbf{n}}_0$  was on the north pole, but as was shown, the  $\langle \tilde{C}_\ell \rangle$ s are rotationally invariant, that is  $\langle \tilde{C}_\ell \rangle$  remains the same if one rotates the Gabor window so that it is centered on the north pole.

When using real CMB data, the observed temperature map is always pixelised. So an integral over the sphere has to be replaced by a sum over pixels. In this case, the formula for  $h(\ell, \ell', m)$  has to be replaced by

$$h(\ell, \ell', m) = \sum_j G_j \lambda_{\ell m}^j \lambda_{\ell' m}^j \Delta_j, \quad (4.40)$$

where the index  $j$  is the pixel number replacing the angle  $\theta$  and  $\Delta_j$  is the area of pixel  $j$ . Using a pixelisation scheme like HEALPix (Górski et al. ) which has a structure of azimuthal rings going from the north to the south pole with  $N_r$  pixels in ring  $r$  and equal area for each pixel  $\Delta_j = \Delta$  this can be written as

$$\begin{aligned} h(\ell, \ell', m) &= \Delta \sum_r \sum_{p=0}^{N_r-1} G_r \lambda_{\ell m}^r \lambda_{\ell' m}^r, \\ &= \Delta \sum_r N_r G_r \lambda_{\ell m}^r \lambda_{\ell' m}^r. \end{aligned} \quad (4.41)$$

Here the sum over pixels is split into a sum over rings  $r$  and a sum over the pixels in each ring  $p$ . The first sum goes over all rings which have  $\theta < \theta_C$ .

Using this expression for the  $\tilde{a}_{\ell m}$  one can now find the correlation matrix

$$\langle \tilde{C}_\ell \tilde{C}_{\ell'} \rangle = \sum_{mm'} \frac{\langle \tilde{a}_{\ell m}^* \tilde{a}_{\ell m} \tilde{a}_{\ell' m'}^* \tilde{a}_{\ell' m'} \rangle}{(2\ell + 1)(2\ell' + 1)}. \quad (4.42)$$

In this expression one can use relation (4.37) to get,

$$\langle \tilde{C}_\ell \tilde{C}_{\ell'} \rangle = \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} \sum_{LL'KK'} \langle a_{Lm}^* a_{L'm} a_{K'm'}^* a_{K'm'} \rangle \quad (4.43)$$

$$\times h(\ell, L, m) h(\ell, L', m) h(\ell', K, m') h(\ell', K', m') \quad (4.44)$$

$$= \frac{1}{(2\ell + 1)(2\ell' + 1)} \quad (4.45)$$

$$\times \sum_{mm'} \sum_{LL'KK'} [\langle a_{Lm}^* a_{L'm} \rangle \langle a_{K'm'}^* a_{K'm'} \rangle \quad (4.46)$$

$$+ \langle a_{Lm}^* a_{K'm'}^* \rangle \langle a_{L'm} a_{K'm'} \rangle \quad (4.47)$$

$$+ \langle a_{Lm}^* a_{K'm'} \rangle \langle a_{L'm} a_{K'm'}^* \rangle ] \quad (4.48)$$

$$\times h(\ell, L, m) h(\ell, L', m) h(\ell', K, m') h(\ell', K', m'). \quad (4.49)$$

Clearly the first term is just the product  $\langle \tilde{C}_\ell \rangle \langle \tilde{C}_{\ell'} \rangle$ , and the two last terms are equal (using  $a_{K'm'}^* = a_{K'm'} (-1)^{m'}$  and  $a_{K'm'} = (-1)^{m'} a_{K'm'}^*$ ) so one gets,

$$M_{ij} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m \left( \sum_L C_L h(\ell_i, L, m) h(\ell_j, L, m) \right)^2. \quad (4.50)$$

This is one of the main results of this thesis since the formula allows one to analytically calculate the correlation matrix needed for likelihood analysis. Another

main result is the recurrence deduced in appendix (D) which allows fast evaluation of the  $h(\ell, \ell', m)$  functions and thereby this correlation matrix.

In Figure (4.12) one can see the correlation matrix for a typical power spectrum with a 15° FWHM Gaussian Gabor window (note that in the figure, the correlation matrix is normalised with the pseudo power spectrum). The correlation between different  $\tilde{C}_\ell$  is clearly small as the diagonal is quite sharp. There is only the small 'wall' at low multipoles which again comes from the coupling to the smallest multipoles which have very high values.

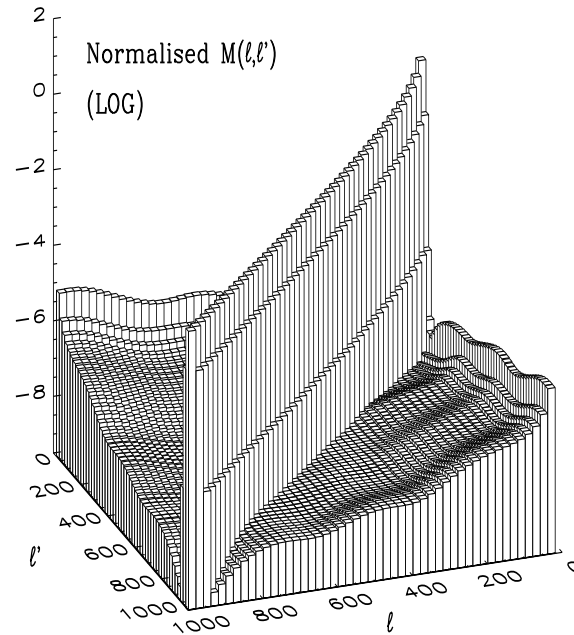


Figure 4.12: The figure shows the correlation matrix  $M_{\ell\ell'}$  between pseudo spectrum coefficients normalised with the pseudo spectrum  $(\langle \tilde{C}_\ell^T \tilde{C}_{\ell'}^T \rangle - \langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^T \rangle) / (\langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^T \rangle)$  for a 15 degree FWHM Gaussian Gabor window. A standard CDM power spectrum was used to produce this matrix.

### 4.2.3 Including White Noise

I now assume that each pixel  $j$  has a noise temperature denoted by  $n_j$ , with the following properties,

$$\langle n_j \rangle = 0, \quad \langle n_j n_{j'} \rangle = \delta_{jj'} \sigma_j^2, \quad (4.51)$$

where  $\sigma_j$  is the noise variance in pixel  $j$ . Then one has the following expressions for the  $a_{\ell m}$  and  $C_\ell$  (I use superscript  $N$  for noise quantities),

$$a_{\ell m}^N = \sum_j n_j Y_{\ell m}^{j*} \Delta_j \quad (4.52)$$

$$\langle a_{\ell m}^N a_{\ell' m'}^{N*} \rangle = \sum_{jj'} \langle n_j n_{j'} \rangle Y_{\ell m}^j Y_{\ell' m'}^{j'*} \Delta_j \Delta_{j'} = \sum_j \sigma_j^2 Y_{\ell m}^j Y_{\ell' m'}^{j*} \Delta_j^2 \quad (4.53)$$

$$\langle C_\ell^N \rangle = \frac{1}{2\ell + 1} \sum_m \langle a_{\ell m}^N a_{\ell m}^{N*} \rangle = \frac{1}{4\pi} \sum_j \sigma_j^2 \Delta_j^2. \quad (4.54)$$

Here  $Y_{\ell m}^j$  is the Spherical Harmonic of the pixel centre of pixel  $j$ . For the windowed coefficients, one gets similarly,

$$\tilde{a}_{\ell m}^N = \sum_j G_j n_j \Delta_j Y_{\ell m}^{j*} \quad (4.55)$$

$$\tilde{C}_\ell^N = \frac{1}{4\pi} \sum_j \sigma_j^2 G_j^2 \Delta_j^2 \quad (4.56)$$

The next step is to find the noise correlation matrix,

$$\langle \tilde{C}_\ell^N \tilde{C}_{\ell'}^N \rangle = \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} \sum_{jj'kk'} \Delta_j \Delta_{j'} \Delta_k \Delta_{k'} \quad (4.57)$$

$$\times \langle n_j n_{j'} n_k n_{k'} \rangle G_j G_{j'} G_k G_{k'} Y_{\ell m}^j Y_{\ell' m'}^{j'*} Y_{\ell m}^k Y_{\ell' m'}^{k'*} \quad (4.58)$$

$$= \langle C_\ell^N \rangle \langle C_{\ell'}^N \rangle + M_{\ell \ell'}^N, \quad (4.59)$$

where  $M_{\ell \ell'}^N$  can be written as,

$$M_{\ell \ell'}^N = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} \left( \sum_j \Delta_j^2 G_j^2 \sigma_j^2 Y_{\ell m}^j Y_{\ell' m'}^{j*} \right)^2 \quad (4.60)$$

For pixelisation schemes like HEALPix, the expression can be evaluated fast using FFT. This is apparent when one writes the sum over pixels as a double sum over rings and pixels per ring.

$$\sum_j \Delta_j^2 G_j^2 \sigma_j^2 Y_{\ell m}^j Y_{\ell' m'}^{j*} = \sum_r \Lambda_{\ell m}^r \Lambda_{\ell' m'}^r \sum_{p=0}^{N_r-1} e^{-i\phi_p(m-m')} \Delta^2 G_r^2 \sigma_{r,p}^2. \quad (4.61)$$

In the case of an axissymmetric noise model, this expression becomes even easier which is apparent writing this as

$$\sum_r \Lambda_{\ell m}^r \Lambda_{\ell' m'}^r \Delta^2 G_r^2 \sigma_r^2 \underbrace{\sum_{p=0}^{N_r-1} e^{-i\phi_p(m-m')}}_{N_r \delta_{mm'}} = \Delta \sum_r \Lambda_{\ell m}^r \Lambda_{\ell' m'}^r \underbrace{\Delta G_r^2 \sigma_r^2}_{G_r'} \equiv h'(\ell, \ell', m). \quad (4.62)$$

The sum is equivalent to the previous expression for  $h(\ell, \ell', m)$  (equation (4.41)) with a new window  $G'_r$ . This motivates the definition of  $h(\ell, \ell', m, m')$  such that

$$M_{\ell\ell'}^N = \frac{2}{(2\ell+1)(2\ell'+1)} \sum_{mm'} h^2(\ell, \ell', m, m'), \quad (4.63)$$

where

$$h'(\ell, \ell', m, m') \equiv \Delta \sum_j G'_j Y_{\ell m}^j Y_{\ell' m'}^j, \quad (4.64)$$

where  $G'_j = \Delta G_r^2 \sigma_{r,p}^2$  for the discrete case. These function can also be calculated using the recursion which I deduce in appendix (D).

One can then find the total correlation matrix, splitting it up into one part due to signal, one part due to noise and a cross term,

$$\tilde{a}_{\ell m} = \tilde{a}_{\ell m}^S + \tilde{a}_{\ell m}^N \quad (4.65)$$

$$\tilde{C}_\ell = \frac{1}{2\ell+1} \sum_m \langle \tilde{a}_{\ell m} \tilde{a}_{\ell' m} \rangle = \tilde{C}_\ell^S + \tilde{C}_\ell^N + \tilde{C}_\ell^X \quad (4.66)$$

$$\langle \tilde{C}_\ell \rangle = \langle \tilde{C}_\ell^S \rangle + \langle \tilde{C}_\ell^N \rangle \quad (4.67)$$

$$\tilde{C}_\ell^X = \frac{1}{2\ell+1} \sum_m \left( a_{\ell m}^S a_{\ell m}^{N*} + a_{\ell m}^N a_{\ell m}^{S*} \right), \quad (4.68)$$

where the assumption that there is no correlation between signal and noise was used. One can then see that the correlation matrix can be written in a similar manner,

$$\langle \tilde{C}_{\ell_i} \tilde{C}_{\ell_j} \rangle - \langle \tilde{C}_{\ell_i} \rangle \langle \tilde{C}_{\ell_j} \rangle = M_{ij}^S + M_{ij}^N + \langle \tilde{C}_{\ell_i}^X \tilde{C}_{\ell_j}^X \rangle. \quad (4.69)$$

This is another major result of this thesis showing the full correlation matrix of  $\tilde{C}_\ell$  including noise. One can write the cross term as,

$$\langle \tilde{C}_\ell^X \tilde{C}_{\ell'}^X \rangle = 4 \sum_{mm'} \frac{\langle \tilde{a}_{\ell m}^S \tilde{a}_{\ell' m'}^{S*} \rangle \langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell' m'}^{N*} \rangle}{(2\ell+1)(2\ell'+1)} \quad (4.70)$$

$$= 4 \sum_m \frac{\langle \tilde{a}_{\ell m}^S \tilde{a}_{\ell' m}^{S*} \rangle \langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell' m}^{N*} \rangle}{(2\ell+1)(2\ell'+1)}, \quad (4.71)$$

where the relation  $\langle \tilde{a}_{\ell m}^S \tilde{a}_{\ell' m'}^{S*} \rangle = \delta_{mm'} \langle \tilde{a}_{\ell m}^S \tilde{a}_{\ell' m}^{S*} \rangle$  was used. From the above, one can see that these two factors can be written as,

$$\langle \tilde{a}_{\ell m}^S \tilde{a}_{\ell' m}^{S*} \rangle = \sum_{\ell''} C_{\ell''} h(\ell, \ell'', m) h(\ell', \ell'', m), \quad (4.72)$$

$$\langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell' m}^{N*} \rangle = \sum_i G_i^2 Y_{\ell m}^i Y_{\ell' m}^i \Delta_i^2 \sigma_i^2, \quad (4.73)$$

$$= h'(\ell, \ell', m). \quad (4.74)$$

### 4.2.4 Likelihood Estimation and Results

The expressions for the covariance matrix and mean psuedo- $C_\ell$  have now been found. Because of the limited information content in one patch of the sky one can not estimate the full sky  $C_\ell$  for all multipoles  $\ell$ . For this reason the full sky power spectrum has to be estimated in  $N^{bin}$  bins. Also the algorithm to minimize the log-likelihood needs the numbers to be estimated to be of roughly the same order of magnitude. For this reason I estimate for some parameters  $D_b$  which for bin  $b$  is defined as

$$C_\ell = \frac{D_b}{\ell(\ell+1)}, \quad \ell_b \leq \ell < \ell_{b+1}, \quad (4.75)$$

where  $\ell_b$  is the first multipole in bin  $b$ . Using this binning, one can calculate the likelihood significantly faster by writing the signal correlation matrix equation (4.50) as

$$M_{ij}^S = \sum_b \sum_{b'} D_b D_{b'} \chi(b, b', i, j), \quad (4.76)$$

where  $\chi(b, b', i, j)$  is given as,

$$\begin{aligned} \chi(b, b', i, j) &\equiv \frac{2}{(2\ell_i + 1)(2\ell_j + 1)} \\ &\times \sum_m \left( \sum_{\ell \in b} B_\ell^2 \ell(\ell+1) h(\ell, \ell_i, m) h(\ell, \ell_j, m) \right) \\ &\times \left( \sum_{\ell \in b'} B_\ell^2 \ell(\ell+1) h(\ell, \ell_i, m) h(\ell, \ell_j, m) \right), \end{aligned} \quad (4.77)$$

which is precomputed. The sums over  $\ell$  here go over the  $\ell$  values in each specific bin  $b$ . One sees that computing the likelihood takes of the order  $(N^{bin})^2 (N^{in})^2$  operations whereas the precomputation of the factor  $\chi(k, k', i, j)$  goes as  $(N^{bin})^2 (N^{in})^2 N_m$  where  $N_m$  is the number of  $m$  values used. Note that the multipole coefficients of the beam  $B_\ell$  are also included. The reason is that the input data is always affected by the beam and this is corrected for by using the beam convolved full sky power spectrum  $C_\ell B_\ell^2$ .

The sum over  $m$  in the expressions for the covariance matrix and  $\langle \tilde{C}_\ell \rangle$  can be limited. The  $h$ -functions are rapidly decreasing for increasing  $m$  for Gaussian and tophat windows. For Gaussian Gabor windows it seems that one can cut the sums over  $m$  at 200 to high accuracy. For top-hat windows, the sum should be extended to  $m = 400$ .

The noise cross term can be written in a similar way as

$$M_{ij}^X \equiv \langle \tilde{C}_{\ell_i}^X \tilde{C}_{\ell_j}^X \rangle = \sum_k D_b \chi'(b, i, j), \quad (4.78)$$

where

$$\chi'(b, i, j) \equiv \frac{2}{(2\ell_i + 1)(2\ell_j + 1)} \sum_m \left( \sum_{\ell b} B_\ell^2 \ell(\ell + 1) h(\ell, \ell_i, m) h(\ell, \ell_j, m) \right) h'(\ell_i, \ell_j, m). \quad (4.79)$$

In the minimization of the likelihood, one also needs the first and second derivative of the log-likelihood with respect to the bin values  $D_b$ . These can be found to be,

$$\frac{\partial L}{\partial D_b} = Tr(\mathbf{A}_b) + \mathbf{f}^T \mathbf{h}_b + 2 \frac{\partial \mathbf{d}^T}{\partial D_b} \mathbf{f} \quad (4.80)$$

$$\frac{\partial^2 L}{\partial D_b \partial D_{b'}} = -Tr(\mathbf{A}_b \mathbf{A}_{b'}) + Tr(\mathbf{C}^{-1} \frac{\partial^2 \mathbf{C}}{\partial D_b \partial D_{b'}}) + 2 \mathbf{h}_b^T \mathbf{C}^{-1} \mathbf{h}_{b'} \quad (4.81)$$

$$-2 \mathbf{h}_b^T \mathbf{g}_{b'} - 2 \mathbf{h}_{b'}^T \mathbf{g}_b - \mathbf{f}^T \frac{\partial^2 \mathbf{C}}{\partial D_b \partial D_{b'}} \mathbf{f} + 2 \mathbf{f}^T \mathbf{k}_{bb'} + 2 \frac{\partial \mathbf{d}^T}{\partial D_b} \mathbf{g}_{b'} \quad (4.82)$$

I have used the following definitions,

$$\mathbf{A}_b = \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial D_b} \quad (4.83)$$

$$\mathbf{h}_b = \frac{\partial \mathbf{C}}{\partial D_b} \mathbf{C}^{-1} \mathbf{d} \quad (4.84)$$

$$\mathbf{g}_b = \mathbf{C}^{-1} \frac{\partial \mathbf{d}}{\partial D_b} \quad (4.85)$$

$$\mathbf{f} = \mathbf{C}^{-1} \mathbf{d} \quad (4.86)$$

$$\mathbf{k}_{bb'} = \frac{\partial^2 \mathbf{d}}{\partial D_b \partial D_{b'}} \quad (4.87)$$

Here the derivatives of  $\mathbf{d}$  are,

$$\frac{\partial d_i}{\partial D_b} = -\frac{\partial \langle \tilde{C}_{\ell_i} \rangle}{\partial D_b} = \frac{-1}{2\ell_i + 1} \sum_m \left( \sum_{L'} h(\ell_i, L', m)^2 \frac{\partial C_{L'}}{D_b} B_\ell^2 \right), \quad (4.88)$$

$$\frac{\partial^2 d_i}{\partial D_b \partial D_{b'}} = -\frac{\partial^2 \langle \tilde{C}_{\ell_i} \rangle}{\partial D_b \partial D_{b'}} = \frac{-1}{2\ell_i + 1} \sum_m \left( \sum_{L'} h(\ell_i, L', m)^2 \frac{\partial^2 C_{L'}}{D_b D_{b'}} B_\ell^2 \right). \quad (4.89)$$

Obviously for this binning, the double derivative of  $\mathbf{d}$  disappears.

Since the  $\tilde{C}_\ell$  are coupled, one can not use all multipoles in the datavector, the covariance matrix would in this case become singular. One has to choose a number  $N^{in}$  of multipoles  $\ell_i$  for which one finds  $d_i$ . How many multipoles to use depends on how tight the  $\tilde{C}_\ell$  are coupled which depends on the size  $\Delta\ell$  of the kernel (figure (4.5)). I found that  $N^{in} \approx 2/3 \ell_{max}/\Delta\ell$  seems to be optimal. To use a lower  $N^{in}$  increases the error bars on the estimates and a higher  $N^{in}$  will



not improve the estimates. One can at most fit for as many  $C_\ell$ s as the number of  $\tilde{C}_\ell$  ( $N^{in}$ ) one has used in the analysis. So one needs to find a number  $N^{bin} \leq N^{in}$  of bin values  $D_b$  from which one can construct the full sky  $C_\ell$ .

I will now describe some test simulations to see how the method works. As a first test, I used the same model as was used in (Hivon et al. 2001), with  $\Omega_{total} = 1$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$  and  $n_s = 0.975$ . These are the parameters from the combined Maxima-Boomerang analysis (Jaffe et al. 2001). I used a circular patch with  $15.5^\circ$  radius covering the same fraction of the sky as in (Hivon et al. 2001). Using HEALPix I simulated a sky using a standard CDM power spectrum with  $l_{max} = 1024$  and a  $7'$  pixel size ( $N_{side} = 512$  in HEALPix language). I smoothed the map with a  $10'$  beam and added non-correlated non-uniform noise to it. Here a Gaussian Gabor window with  $FWHM = 12^\circ$  was used with a cut-off  $\theta_C = 3\sigma$ . For the likelihood estimation, I had  $N^{bin} = 20$  full sky  $C_\ell$  bins and  $N^{in} = 100$   $\tilde{C}_\ell$  values between  $\ell = 2$  and  $\ell = 960$ . In figure (4.13) one can see the result. The shaded areas are the expected  $1\sigma$  variance with and without noise. These were found from the approximate theoretical formula for a uniform noise model (using the formula in (Hivon et al. 2001)). The formula is similar to the one used in most publications and is in this case a very good approximation even with non-uniform noise. In the next example however I will show that the formula has to be used with care. In the figure, the error bars on the estimates are taken from the Fisher matrix and the signal-to-noise ratio  $S/N = 1$  at  $\ell = 575$ .

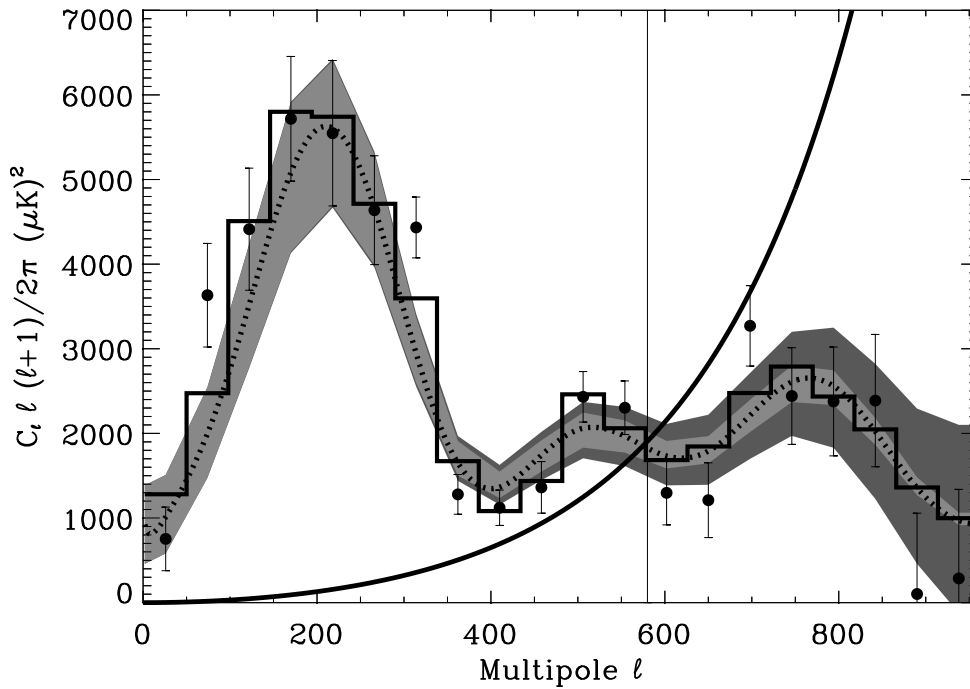


Figure 4.13: The analysis of an input model with  $\Omega_{total} = 1$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b h^2 = 0.03$  and  $n_s = 0.975$ . I used a non-uniform white noise model with  $S/N = 1$  at  $\ell = 575$ . The dotted line is the input average full sky power spectrum and the histogram shows the binned pseudo power spectrum for this realisation (without noise). I used  $N^{bin} = 20$  bins and  $N^{in} = 100$  input sample points to the likelihood. The shaded areas about the binned average full sky spectrum (which is not plotted) are the theoretical variance with and without noise. The bright shaded area shows cosmic and sample variance and the dark shaded area also has variance due to noise included. The variance due to noise was found from the approximate formula for uniform noise. The  $1\sigma$  error bars on the estimates are taken from the inverse Fisher matrix. The solid line increasing from the left to the right is the noise power spectrum.

In figure (4.14), I have plotted the average of 1000 simulations, with different noise and sky realisations. From the plot, the method seems to give an unbiased estimate of the power spectrum bins  $D_b$ . For the lowest multipoles the estimates are slightly lower than the binned input spectrum. This is a result of the slightly skewed probability distribution of  $\tilde{C}_\ell$  for small windows at these low multipoles (see figure (4.9) and (4.10)). The probability that the  $\tilde{C}_\ell$  at lower multipoles have a value lower than the average  $\langle \tilde{C}_\ell \rangle$  is high and the assumption about a

Gaussian distribution about this average leads the estimates to be lower. When a bigger area of the sky is available such that several patches can be analysed jointly to give the full sky power spectrum, this bias disappears. This will be shown in section (4.3.1).

In this example one can see that the  $1\sigma$  error bars from Monte Carlo coincide very well with the theoretical error shown as shaded areas from the formula in (Hivon et al. 2001). Note that the error bars on the higher  $\ell$  are smaller than in (Hivon et al. 2001) because the noise model used in that paper was not white. Also they took into account errors due to map making which is not considered here.

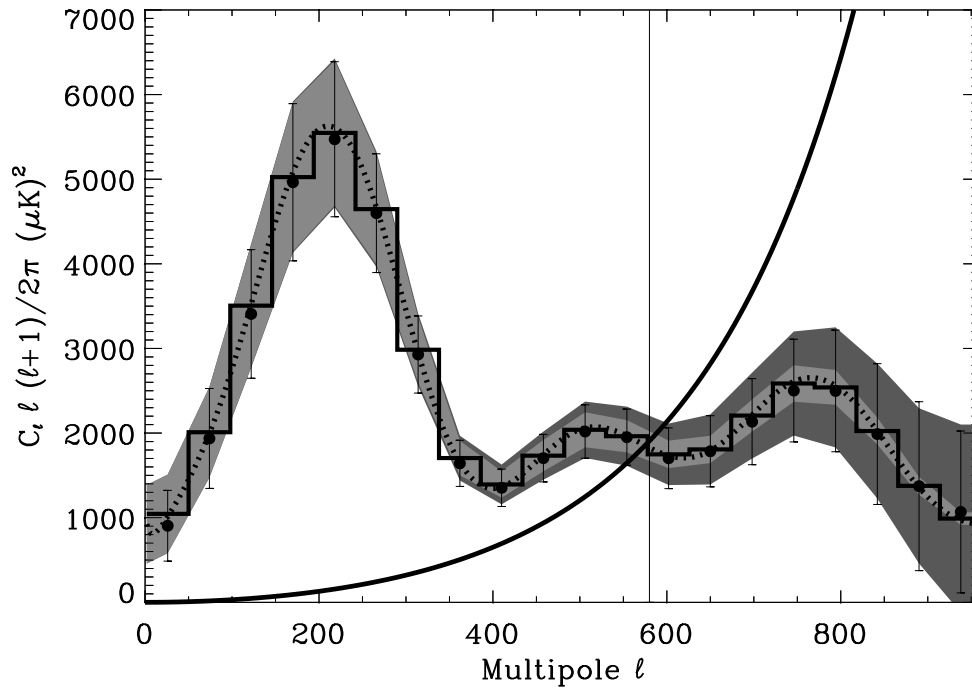


Figure 4.14: Same as in figure (4.13) but this is an average over 1000 simulations and estimations. The histogram is now the binned average full sky power spectrum. The error bars on this plot are the  $1\sigma$  variances taken from Monte-Carlo.

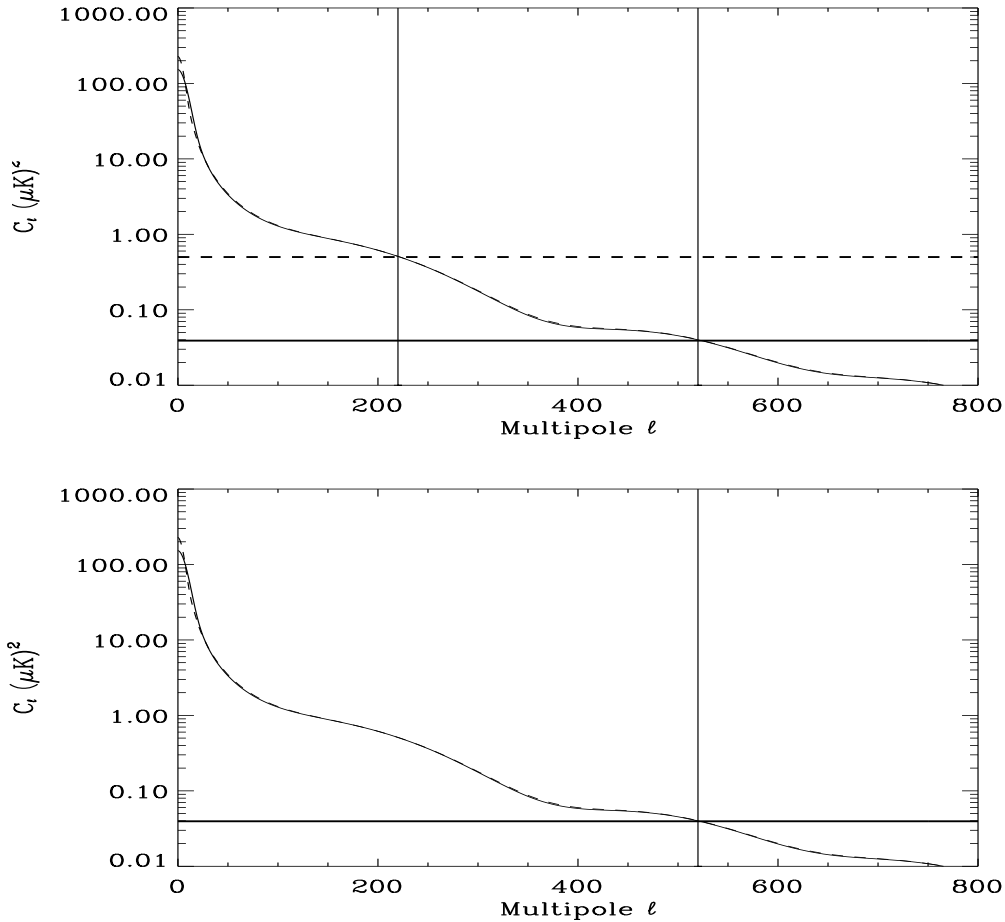


Figure 4.15: The plots show average signal and noise pseudo power spectra plotted separately. The spectra are normalised so that they can be compared to the full sky power spectrum. The solid and dashed curves which almost fall together are the signal pseudo power spectra for a 15 degree FWHM Gaussian Gabor window and a corresponding tophat window respectively. In the upper plot the noise model showed in figure (4.16) was used. This noise model is increasing from the north pole and down to the edges of the patch. This is opposite of the Gaussian window and for this reason the Gaussian window gets higher signal to noise ratio. The solid horizontal line in the upper plot shows the noise pseudo power spectrum for the Gaussian window and the dashed horizontal line shows the noise pseudo power spectrum for the tophat window. In the lower plot a uniform noise model was used so that the noise pseudo power spectra fall together and are showed as a solid vertical line. The figure shows how a Gaussian window can be used to increase signal to noise.

As a next test, I used a simulation with the same resolution and beam size.

The power spectrum was this time a standard CDM power spectrum. I used an axissymmetric noise model with noise increasing from the centre and outwards to the edges (see figure (4.16)). This is the kind of noise model which could be expected from an experiment scanning on rings, with the rings crossing in the center. I now use a circular patch with  $18.5^\circ$  radius and a  $FWHM = 15^\circ$  Gaussian Gabor window cut at  $\theta_C = 3\sigma$  (radius  $\approx 18^\circ$ ). An interesting point now is that the Gabor window is decreasing from the center and outwards, which is opposite of the noise pattern. This gives the pixels with low noise high significant in the analysis and the pixels with high noise low significance. One sees from the expressions for the signal and noise pseudo power spectra that the Gabor window will work differently on both. This means that  $S/N$  is different depending on the Gabor window. For this case, I have plotted the average pseudo power spectrum for signal and noise separately in figure (4.15). This shows the described effect. The  $S/N$  ratio is much higher for the Gaussian Gabor window in this case, favoring the use of this window for the analysis.

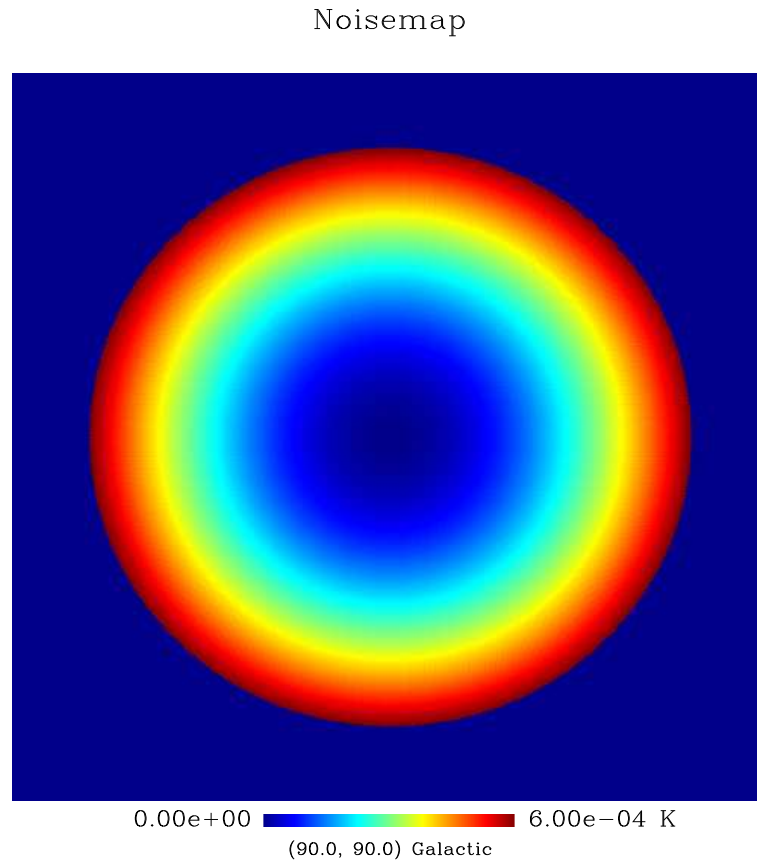


Figure 4.16: The noise map with noise increasing from the north pole and downwards. The figure shows a gnomonic projection with the north pole in the centre.

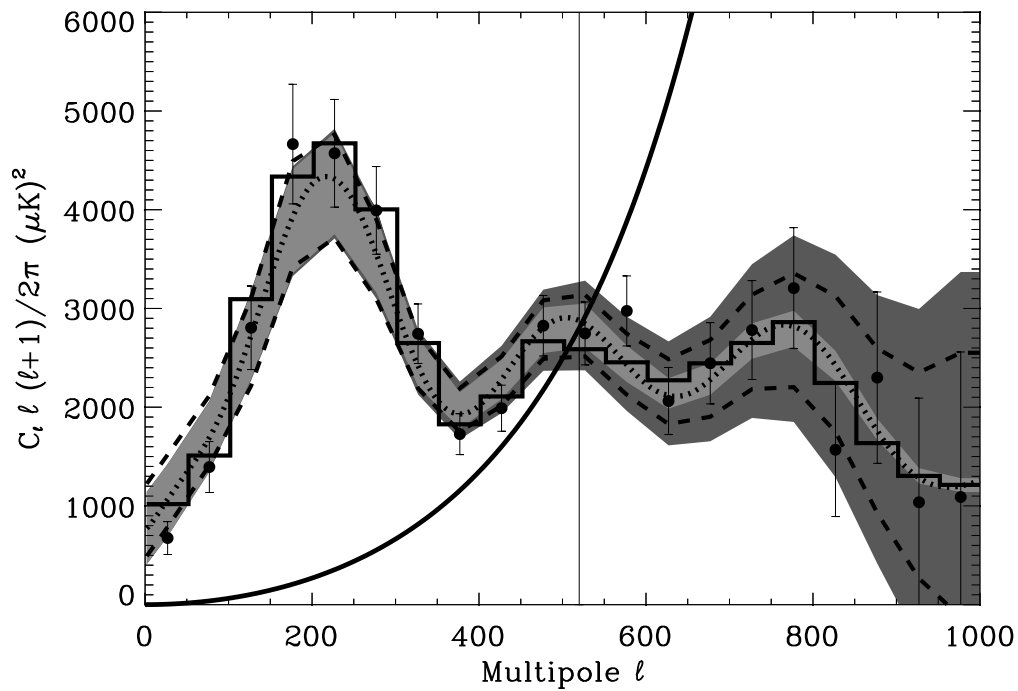


Figure 4.17: Same as figure (4.13) but for a standard CDM model. The noise is increasing from the center and out to the edges while the Gaussian Gabor window has the opposite effect, giving an increased significance to pixels with less noise. As in figure (4.13) the shaded areas show the analytically calculated variance using the ‘naive’ formula for the uniform noise case. The dashed lines show the expected variance using the inverse of the Fisher matrix.

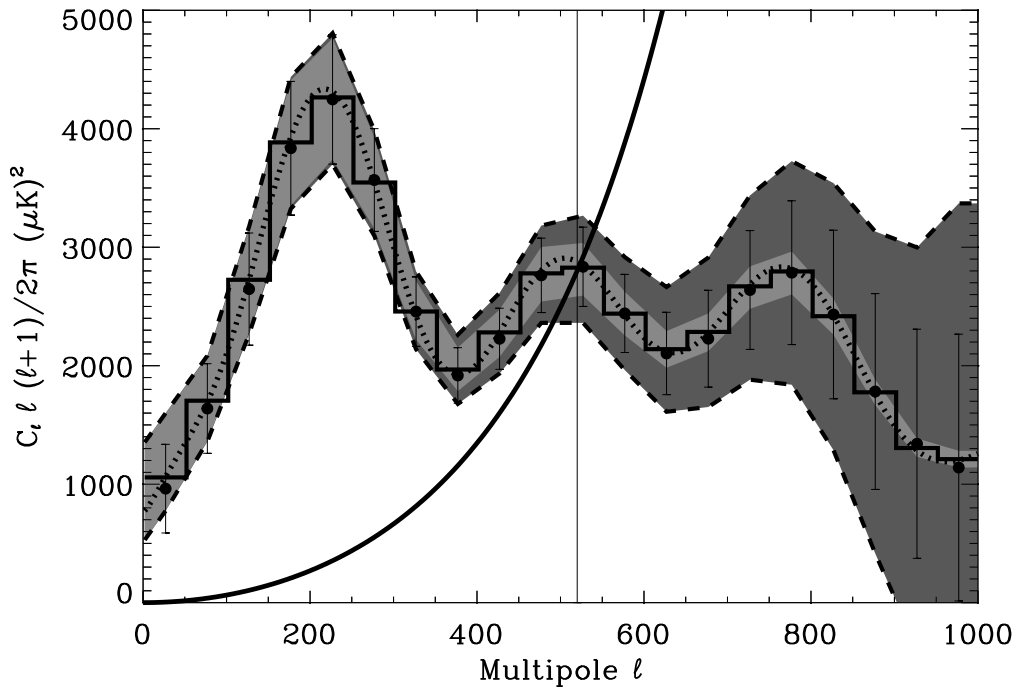


Figure 4.18: Same as figure (4.17) but with 5000 simulations and estimations averaged. The histogram now shows the binned average full sky power spectrum. The error bars on the estimates are here the  $1\sigma$  averages from Monte Carlo.

Again I used  $N^{bin} = 20$  and  $N^{in} = 100$ . The result is shown in Figure (4.17). In figure (4.18) the average over 5000 simulations and estimations is shown. One can see that the estimate also does well beyond  $\ell = 520$  which is where the effective  $S/N = 1$ . The method is clearly still unbiased. The error bars in the part where noise dominates are here lower than the theoretical approximation for uniform noise showed as the dark shaded area. The dashed lines show the theoretical  $1\sigma$  variance taken from the inverse Fisher matrix which here gives a very good agreement with Monte Carlo.



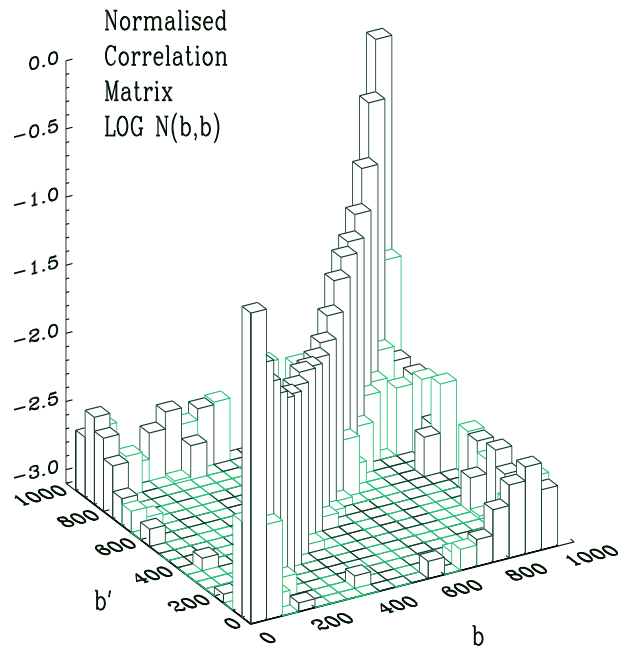


Figure 4.19: The average correlation matrix  $N(b, b') \equiv \langle D_b D_{b'} \rangle / (\langle D_b \rangle \langle D_{b'} \rangle) - 1$  of the estimates in figure (4.18). The negative elements are bright coloured.

In figure (4.19) I show the average (over 5000 estimations) of the correlation between the estimates  $D_b$  between different bins. The figure shows that the correlations between estimates are low and in fact in each line all off-diagonal elements are more than an order of magnitude lower than the diagonal element of that line. In Figure (4.20) I show that the probability distribution of the estimates in Figure (4.18) are almost Gaussian distributed.

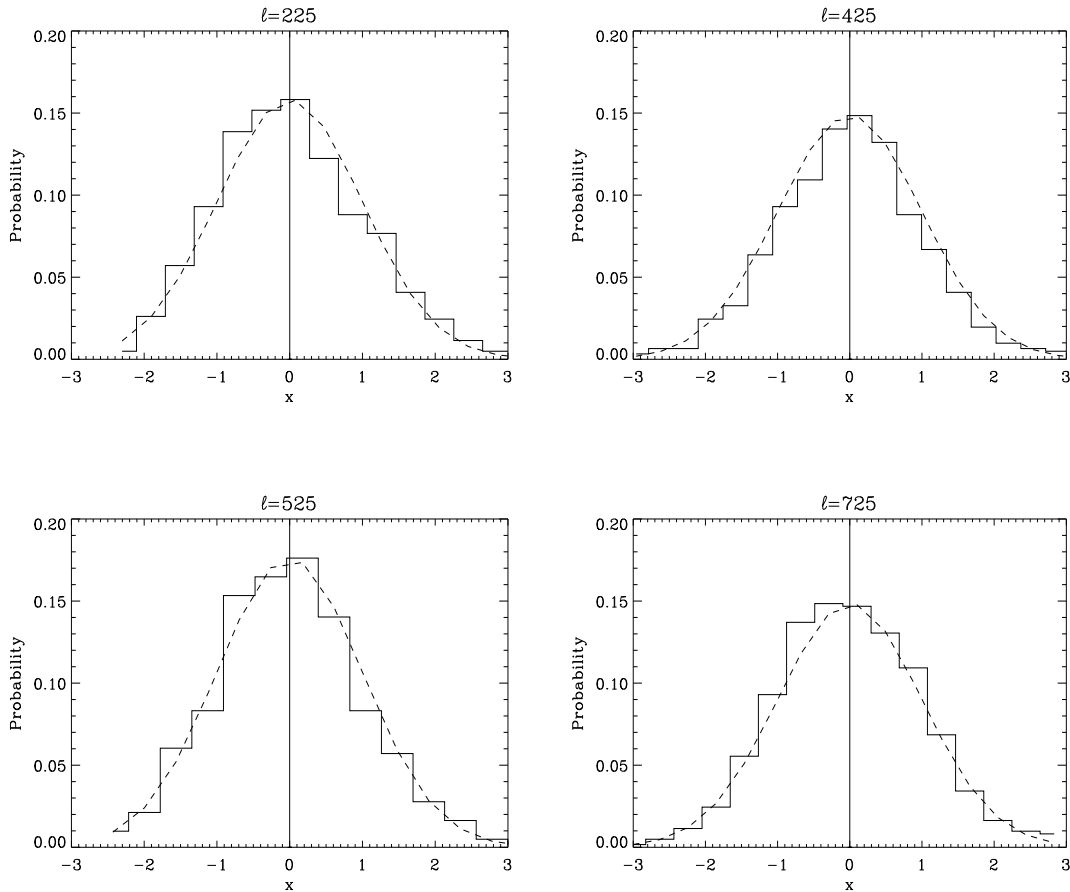


Figure 4.20: The probability distribution of the estimates  $D_b$  in figure 4.18. The variable  $x$  is given as  $x = (D_b - \langle D_b \rangle) / \sqrt{\langle (D_b - \langle D_b \rangle)^2 \rangle}$ . The dashed line is a Gaussian with mean and standard deviation taken from Monte-Carlo. The plot shows bin estimates centered at  $\ell = 225$ ,  $\ell = 425$ ,  $\ell = 525$  and  $\ell = 725$ .

To test the method at higher multipoles I also did one estimation up to multipole  $\ell = 2048$ . I used HEALPix resolution  $N_{side} = 1024$  and simulated a sky with a  $8'$  Gaussian beam and added non-uniform noise. Both the beam and noise level were adjusted according to the specifications for the Planck HFI 143GHz detector (Bersanelli et al. 1996). I used again a  $15^\circ$  FWHM Gaussian Gabor window cut at  $3\sigma$  away from the centre. In the estimation I used  $N^{bin} = 40$  bins and  $N^{in} = 200$  input  $\tilde{C}_\ell$  between  $\ell = 7$  and  $\ell = 2048$ . The average of 100 such simulations is shown in figure (4.21). Each complete likelihood estimation (which includes a total of about 25 likelihood evaluations) took about 8 minutes on a single processor on a 500MHz DEC Alpha work station.

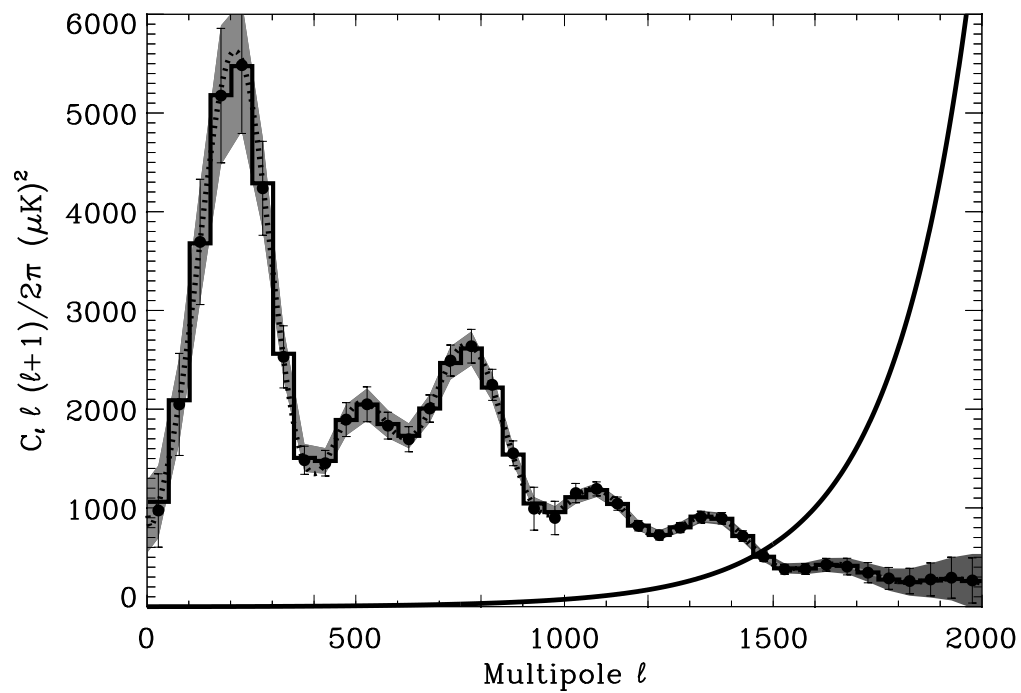


Figure 4.21: Same as figure (4.14) for 100 simulations where the beam and noise level was set according to the specifications of the Planck 143GHz channel. Again a  $15^\circ$  FWHM Gaussian Gabor window was used. The power spectrum was estimated in 40 bins between  $\ell = 2$  and  $\ell = 2048$

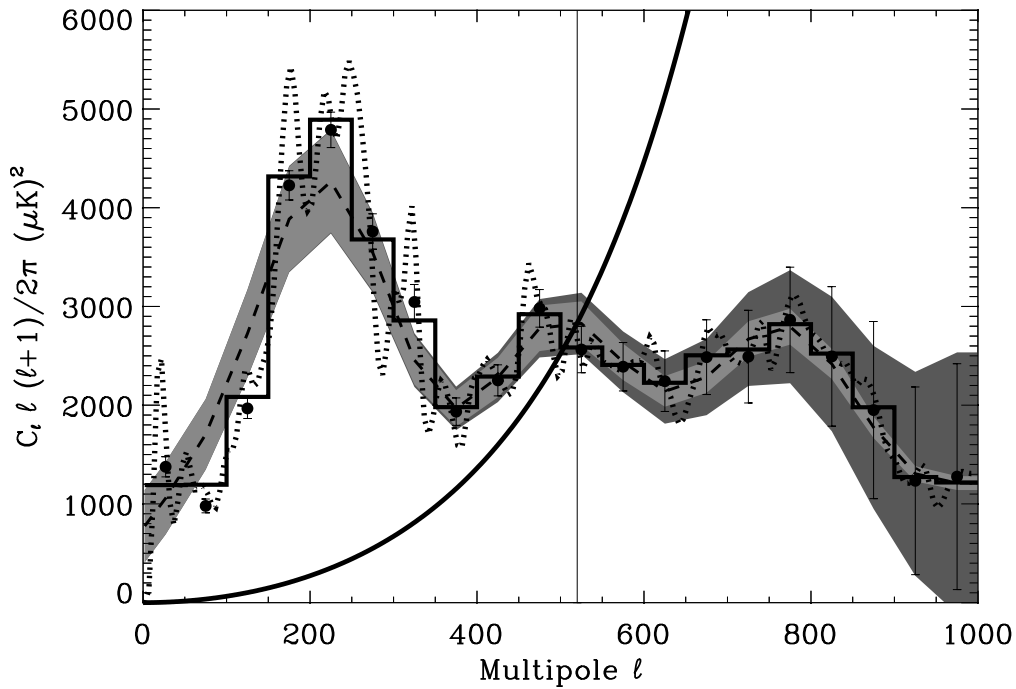


Figure 4.22: The average of 300 estimations where the input  $\tilde{C}_\ell$  were taken from simulations with a fixed CMB realisation but varying noise realisations. The dotted line is the  $N^{in}$  input  $\tilde{C}_\ell$  from the CMB realisation without noise. The histogram is the same spectrum binned in  $N^{bin}$  bins. The dashed line is the binned average full sky power spectrum from which this realisation was made. The shaded areas around the binned full sky spectrum show the variance with (dark) and without (bright) noise. The solid line rising from the left to the right is the noise power spectrum.

In Figure (4.22), I have plotted the average of 300 estimations. The data input in these 300 estimations were the  $\tilde{C}_\ell$  from simulations with a fixed CMB realisation and varying noise realisation. The dotted line shows the  $N^{in}$   $\tilde{C}_\ell$ s used as input to the likelihood but without the noise. The histogram is as before, the input pseudo spectrum without noise binned in  $N^{bin}$  bins. This means that each histogram line shows the average of the dotted line over the bin. One can see that the result is partly following the  $N^{in}$  input  $\tilde{C}_\ell$  and partly the binned power spectrum.

Finally, I made a comparison between a tophat window and a Gaussian Gabor window. In this case I used uniform noise, so that the Gaussian and tophat Gabor windows have the same  $S/N$  ratio which I set to 1 at  $\ell = 520$ . I used a

disc with  $18^\circ$  radius,  $N^{in} = 200$  and  $N^{bin} = 20$ . In Figure (4.23) one can see the result. The lower plot shows the estimates with the Gaussian Gabor window ( $15^\circ$  FWHM) and the upper with the tophat window. The Gaussian window is suppressing parts of the data and for this reason gets a higher sample variance than the tophat. This effect is seen in the plot. Clearly when no noise weighting is required the tophat window seems to be the preferred window (which was also discussed in (Hivon et al. 2001)). This chapter has been concentrating on the Gaussian window to study power spectrum estimation in the presence of a window different from a tophat. It has been shown that a different window can be advantageous when the noise is not uniformly distributed as one can then give data with different quality different significance.

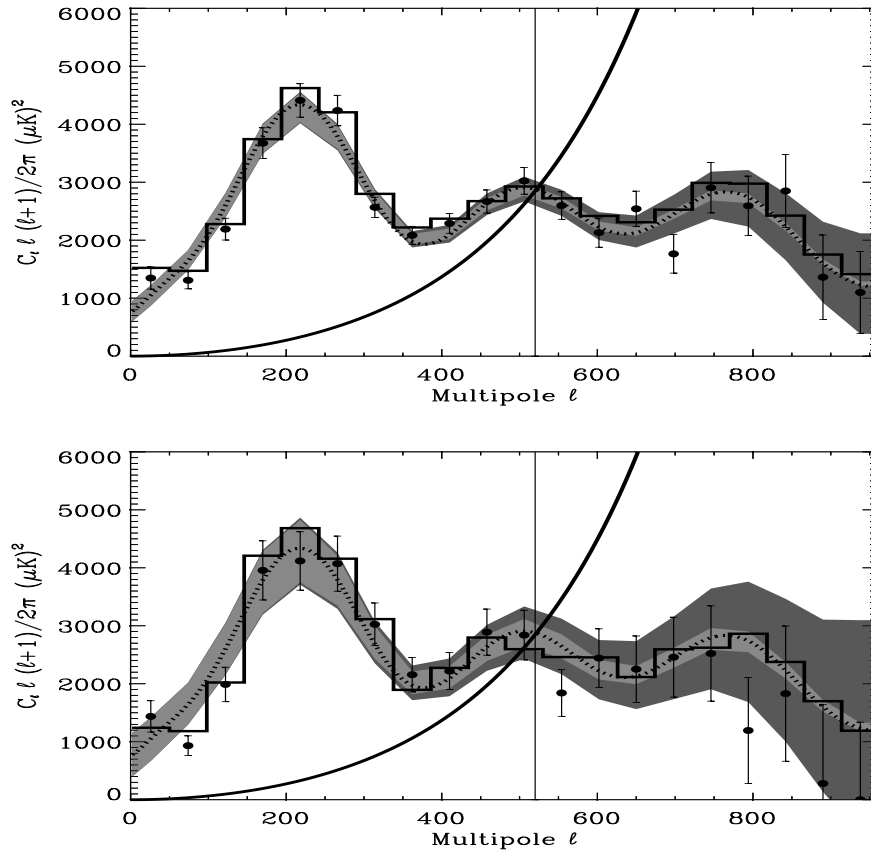


Figure 4.23: Estimates of  $\tilde{C}_\ell$  using a Gaussian Gabor window (lower plot) and a tophat window (upper plot). Here I used a uniform noise model with  $S/N = 1$  at  $\ell = 520$ . The dotted line shows the average full sky power spectrum and the histogram shows the input pseudo power spectrum without noise for this realisation, binned in the same way as the estimates. The bright shaded area shows the cosmic and sample variance around the binned average spectrum (not plotted). The dark shaded area has the variance due to noise included. The  $1\sigma$  error bars on the estimates are taken from the inverse Fisher matrix. The solid line increasing from left to right is the noise power spectrum.

### 4.3 Extensions of the Method

In this section I will discuss two possible extensions of the method. The formalism for the extensions are worked out and some simple examples are shown. Further investigations of these extensions are left for future work.

### 4.3.1 Multiple Patches

It has been shown how one can do power spectrum estimation on one axissymmetric patch on the sky. The next question that arises is what to do when the observed area on the sky is not axissymmetric. In this case one can split the area into several axissymmetric pieces and use the  $\tilde{C}_\ell$  from each piece. Then these  $\tilde{C}_\ell$  from all the patches are used together in the likelihood maximization. The first thing to check before embarking on this idea is the correlation between  $\tilde{C}_\ell$ s in different patches.

Suppose one has two axissymmetric Gabor windows,  $G^A(\Omega)$  and  $G^B(\Omega)$ , centered at two different positions  $A$  and  $B$  on the sky. Suppose also that the rotation operators  $\hat{D}^A$  and  $\hat{D}^B$  will rotate these patches so that the centers are on the north pole. Considering patch  $A$ , one can define,

$$\tilde{a}_{\ell m}^A = \int G_0^A(\Omega) [\hat{D}^A T(\Omega)] Y_{\ell m}(\Omega), \quad (4.90)$$

where  $G_0^A$  is the window  $G^A$  rotated to the north pole. Since  $T(\Omega) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\Omega)$ , one gets that

$$\hat{D}^A T(\Omega) = \sum_{\ell m} a_{\ell m} \sum_{m'} D_{m'm}^{lA} Y_{\ell m'}(\Omega). \quad (4.91)$$

Here the  $D_{m'm}^{\ell A}$  coefficients are described in appendix (A). One now gets,

$$\tilde{a}_{\ell m}^A = \sum_{\ell' m'} a_{\ell' m'} \sum_{m''} D_{m'' m'}^{\ell' A} h^A(\ell, \ell', m) \delta_{m m''} \quad (4.92)$$

$$= \sum_{\ell' m'} a_{\ell' m'} D_{m m'}^{\ell' A} h^A(\ell, \ell', m), \quad (4.93)$$

where  $h^A(\ell, \ell', m)$  is just the  $h(\ell, \ell', m)$  function for the Gabor window  $G^A(\Omega)$ .

The next step is to find the correlations between  $\tilde{C}_\ell^A$  and  $\tilde{C}_\ell^B$ , defined for patch  $A$  as,

$$\tilde{C}_\ell^A = \sum_m \frac{\tilde{a}_{\ell m}^A \tilde{a}_{\ell m}^{A*}}{2\ell + 1}. \quad (4.94)$$

Following the procedure I used for a single patch one gets,

$$\begin{aligned} \langle \tilde{C}_\ell^A \tilde{C}_{\ell'}^B \rangle &= \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{m m'} \langle \tilde{a}_{\ell m}^{A*} \tilde{a}_{\ell m}^A \tilde{a}_{\ell' m'}^{B*} \tilde{a}_{\ell' m'}^B \rangle \\ &= \langle \tilde{C}_\ell^A \rangle \langle \tilde{C}_{\ell'}^B \rangle + \frac{2 \sum_{m m'} |\langle \tilde{a}_{\ell m}^{A*} \tilde{a}_{\ell' m'}^B \rangle|^2}{(2\ell + 1)(2\ell' + 1)}. \end{aligned} \quad (4.95)$$

One can use the expression for  $\tilde{a}_{\ell m}^A$  to find,

$$\langle \tilde{a}_{\ell m}^A \tilde{a}_{\ell m}^{B*} \rangle = \sum_{\ell'' m''} \sum_{L'' M''} \langle a_{\ell'' m''} a_{L'' M''}^* \rangle D_{m m''}^{\ell'' A} D_{m' M''}^{L'' B*} \quad (4.96)$$

$$\times h^A(\ell, \ell'', m)h^B(\ell', L'', m') \quad (4.97)$$

$$= \sum_{\ell''} C_{\ell''} \underbrace{\left( \sum_{m''} D_{mm''}^{\ell''A} D_{m'm''}^{\ell''B*} \right)}_{D_{mm'}^{\ell''}(\Delta)} h^A(\ell, \ell'', m)h^B(\ell', \ell'', m') \quad (4.98)$$

$$= \sum_{\ell''} C_{\ell''} d_{mm'}^{\ell''}(\Delta) h^A(\ell, \ell', m)h^B(\ell', \ell'', m'), \quad (4.99)$$

where  $\Delta$  is the angel between the centers of the patches. Relations from appendix (A) were used here.

The next step is to see what happens when noise is introduced. I assume that the noise is uncorrelated. The noise in pixel  $j$  is  $n_j$  and  $\langle n_j n_{j'} \rangle = \delta_{jj'} \sigma_j^2$ . From above one has,

$$\tilde{C}_\ell^A = \tilde{C}_\ell^{AS} + \tilde{C}_\ell^{AN} + \tilde{C}_\ell^{AX}, \quad (4.100)$$

where

$$\tilde{C}_\ell^{AN} = \sum_m \frac{\tilde{a}_{\ell m}^{NA} \tilde{a}_{\ell m}^{NA*}}{2\ell + 1} \quad (4.101)$$

$$\tilde{C}_\ell^{AX} = \sum_m \frac{\tilde{a}_{\ell m}^{NA} \tilde{a}_{\ell m}^{SA*}}{2\ell + 1} \quad (4.102)$$

$$\tilde{a}_{\ell m}^{AN} = \sum_j G_j^A n_j^A Y_{\ell m}^j, \quad (4.103)$$

where the last sum is over pixels,  $G_j^A$  and  $n_j^A$  being the window and noise for pixel  $j$  respectively.

The correlation between the two patches then becomes,

$$\langle \tilde{C}_\ell^A \tilde{C}_{\ell'}^B \rangle - \langle \tilde{C}_\ell^A \rangle \langle \tilde{C}_{\ell'}^B \rangle = M_{\ell\ell'}^S + M_{\ell\ell'}^N + \langle \tilde{C}_\ell^{AX} \tilde{C}_{\ell'}^{BX} \rangle, \quad (4.104)$$

where,

$$M_{\ell\ell'}^S = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} | \langle \tilde{a}_{\ell m}^{AS*} \tilde{a}_{\ell' m'}^{BS} \rangle |^2 \quad (4.105)$$

$$M_{\ell\ell'}^N = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} | \langle \tilde{a}_{\ell m}^{AN*} \tilde{a}_{\ell' m'}^{BN} \rangle |^2. \quad (4.106)$$

Finally,

$$\langle \tilde{C}_\ell^{AX} \tilde{C}_{\ell'}^{BX} \rangle = \frac{4}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} \langle \tilde{a}_{\ell m}^{AS*} \tilde{a}_{\ell' m'}^{BS} \rangle \langle \tilde{a}_{\ell m}^{AN*} \tilde{a}_{\ell' m'}^{BN} \rangle. \quad (4.107)$$

Now one needs an expression for  $\langle \tilde{a}_{\ell m}^{AN} \tilde{a}_{\ell' m'}^{BN*} \rangle$ . One gets,

$$\langle \tilde{a}_{\ell m}^{AN} \tilde{a}_{\ell' m'}^{BN*} \rangle = \sum_{jj'} G_j^A G_{j'}^B \langle n_j^A n_{j'}^B \rangle Y_{\ell m}^j Y_{\ell' m'}^{j'}. \quad (4.108)$$



Here there are only correlations between overlapping pixels. If there are no overlapping pixels between the patches, this term is zero. Otherwise this can be written as a sum over the overlapping pixels

$$\langle \tilde{a}_{\ell m}^{AN} \tilde{a}_{\ell' m'}^{BN*} \rangle = \sum_j G_j^A G_j^B \sigma_j^2 Y_{\ell m}^j Y_{\ell' m'}^{j'}. \quad (4.109)$$

After the expression (4.95) was tested with Monte Carlo simulations, I computed the correlations between  $\tilde{C}_\ell$  for two patches  $A$  and  $B$  where I varied the distance  $\theta$  between the centers of  $A$  and  $B$ . I used a standard CDM power spectrum and both patches  $A$  and  $B$  had a radius of  $18^\circ$  apodised with a  $15^\circ$  FWHM Gaussian Gabor window. In figure (4.24) I have plotted the diagonal of the normalised correlation matrix  $(\langle \tilde{C}_\ell^A \tilde{C}_\ell^B \rangle - \langle \tilde{C}_\ell \rangle^2) / \langle \tilde{C}_\ell \rangle^2$ . The angles  $\theta$  I used were  $6^\circ$ ,  $12^\circ$ ,  $24^\circ$ ,  $30^\circ$ ,  $36^\circ$  and  $180^\circ$ . One sees clearly how the correlations drop with the distance. In the two last cases, there were no common pixels in the patches. As one could expect, the correlations for the largest angles (the first few multipoles) do not drop that fast.

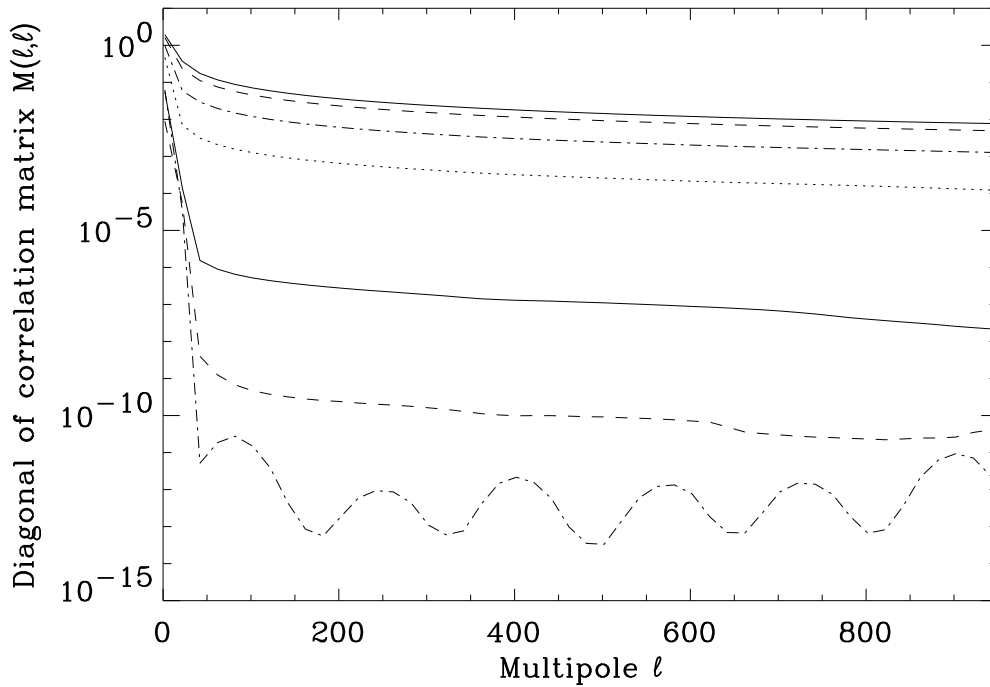


Figure 4.24: The correlation between  $\tilde{C}_\ell$  between two patches  $A$  and  $B$  with an angular distance  $\theta$  between the centers. A normal CDM power spectrum was used and the patches had an  $18^\circ$  radius apodised with a  $15^\circ$  FWHM Gaussian Gabor window. The figure shows the diagonal of the normalised correlation matrix  $(\langle \tilde{C}_\ell^A \tilde{C}_\ell^B \rangle - \langle \tilde{C}_\ell \rangle^2) / \langle \tilde{C}_\ell \rangle^2$  where of course  $\langle C_\ell \rangle = \langle C_\ell^A \rangle = \langle C_\ell^B \rangle$ . The angles used are (from top to bottom on the figure)  $0^\circ$ ,  $6^\circ$ ,  $12^\circ$ ,  $18^\circ$ ,  $30^\circ$ ,  $36^\circ$  and  $180^\circ$ .

In figure (4.25) I have plotted two slices through the correlation matrix of  $\tilde{C}_\ell$  for a single patch at  $\ell = 400$  and  $\ell = 800$ . On the top I plotted the diagonals of the correlation matrices for separation angle  $\theta = 30^\circ$ ,  $\theta = 36^\circ$  and  $\theta = 180^\circ$ . One sees that for the case where the patches do not have overlapping pixels, the whole diagonals have the same level as the far-off-diagonal elements in the  $\theta = 0^\circ$  matrix. When doing power spectrum estimation on one patch, the result did not change significantly when these far-off-diagonal elements were set to zero. For this reason one expects that when analysing several patches which do not overlap simultaneously, the correlations between non-overlapping patches do not need to be taken into account. Note however that for the  $\theta = 30^\circ$  which means that there are only a few overlapping pixels, the approximation will not be that good as the level is orders of magnitude above the far-off-diagonals of the  $\theta = 0^\circ$  matrix. Another thing to note is that for the lowest multipoles, the correlation between

patches is still high but I will also assume this part to be zero and attempt a joint analysis of non-overlapping patches.

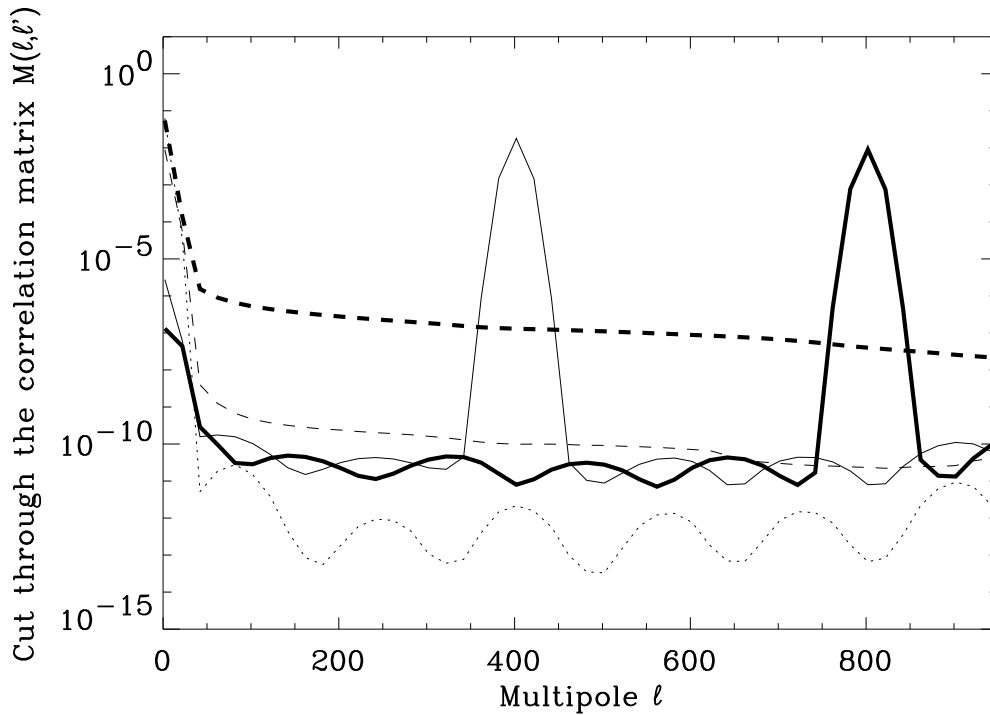


Figure 4.25: Cuts through the correlation matrices which diagonals are shown in figure (4.24). The solid lines (thin and thick) show a cut through the  $\theta = 0^\circ$  correlation matrix at  $\ell = 400$  and  $\ell = 800$  respectively. The dashed lines (thin and thick) show the diagonal of the correlation matrices for  $\theta = 36^\circ$  and  $\theta = 30^\circ$  respectively. The dotted line is the diagonal of the  $\theta = 180^\circ$  matrix.

The full correlation matrices for 0 and 30 degree separation are shown in figure (4.26). The figures show how the diagonal is dropping relative to the far off-diagonal elements.

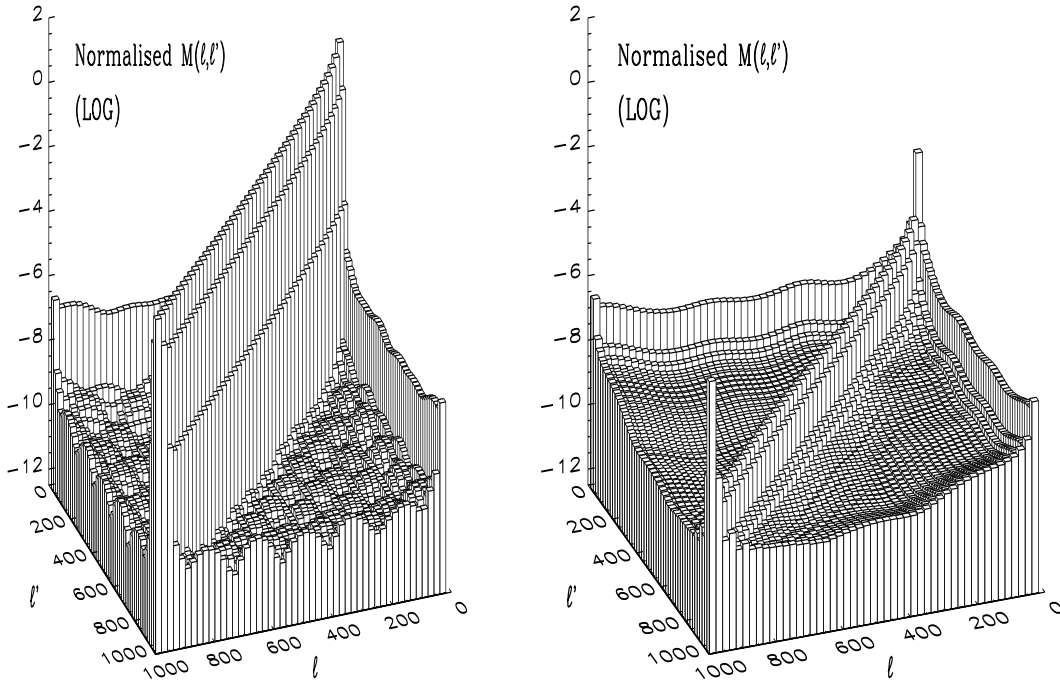


Figure 4.26: The figure shows the correlation matrices  $M(\ell, \ell')$  normalised ( $\langle C_\ell^A C_{\ell'}^B \rangle - \langle C_\ell^A \rangle \langle C_{\ell'}^B \rangle$ ) / ( $\langle C_\ell^A \rangle \langle C_{\ell'}^N \rangle$ ) between pseudo spectrum coefficients for two patches  $A$  and  $B$  of  $18^\circ$  radius and with  $0$  degree (left plot) and  $30$  degree separation. A Gaussian Gabor window with  $15$  degree FWHM was used. The aim of the plot is to show how correlations between  $C_\ell$  from different patches drop when the distance between the two patches is about the FWHM of the Gaussian kernel.

In figure (4.27) the full correlation matrices for  $36$  and  $180$  degree separation is shown. For  $36$  degree separation one can see that the diagonal has almost disappeared with respect to the rest of the matrix whereas for  $180$  degree the diagonal has vanished completely. But the 'wall' at low multipoles remains.

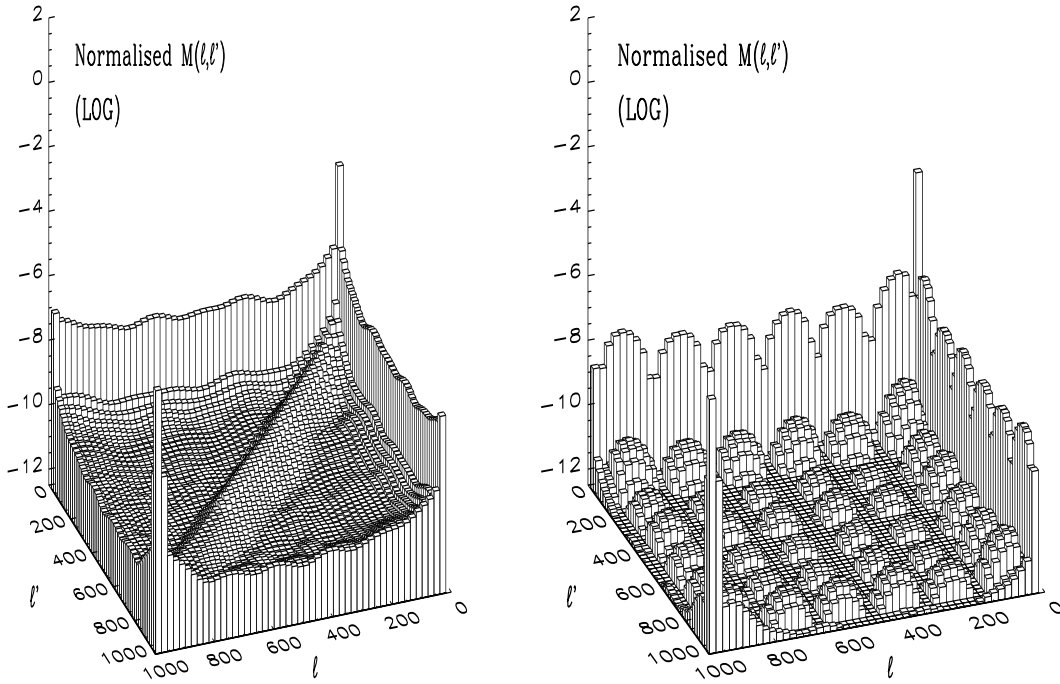


Figure 4.27: This figure shows the same as figure (4.26) but for 36 and 180 degree separation of the patches.

In figure (4.28), I did a separate  $C_\ell$  estimation on 146 non-overlapping patches with radius  $18^\circ$  apodised with a  $15^\circ$  FWHM Gaussian Gabor window. The patches were uniformly distributed over the sphere and uniform noise was added to the whole map. The figure shows the average of the 146  $C_\ell$  estimates. One can see that the estimate seems to be approaching the full sky power spectrum even at small multipoles.

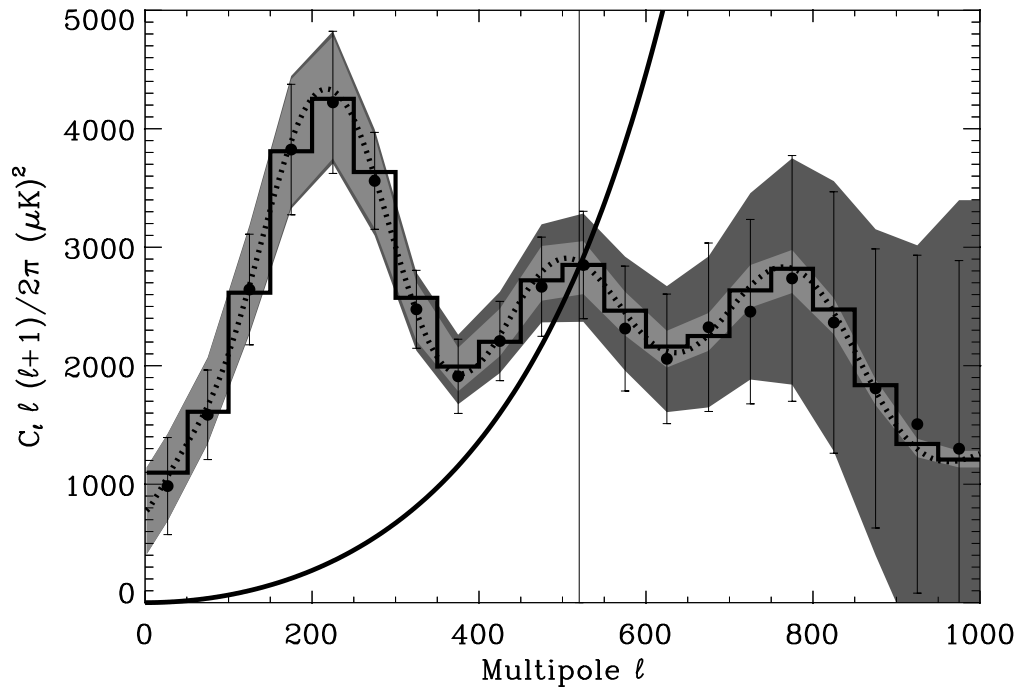


Figure 4.28: The average of 146 individual power spectrum estimations of 146 non-overlapping patches on the same CMB map with uniform noise added to it. The patches all had radius  $18^\circ$  degrees apodised with a  $15^\circ$  FWHM Gaussian Gabor window. The histogram shows the binned average of the 146  $\tilde{C}_\ell$  from the different patches without noise. The dotted line is the average full sky power spectrum and the shaded areas around the binned full sky power spectrum (not plotted) show the theoretical  $1\sigma$  variance with (dark) and without (bright) noise. The solid line rising from the left to the right is the noise power spectrum.

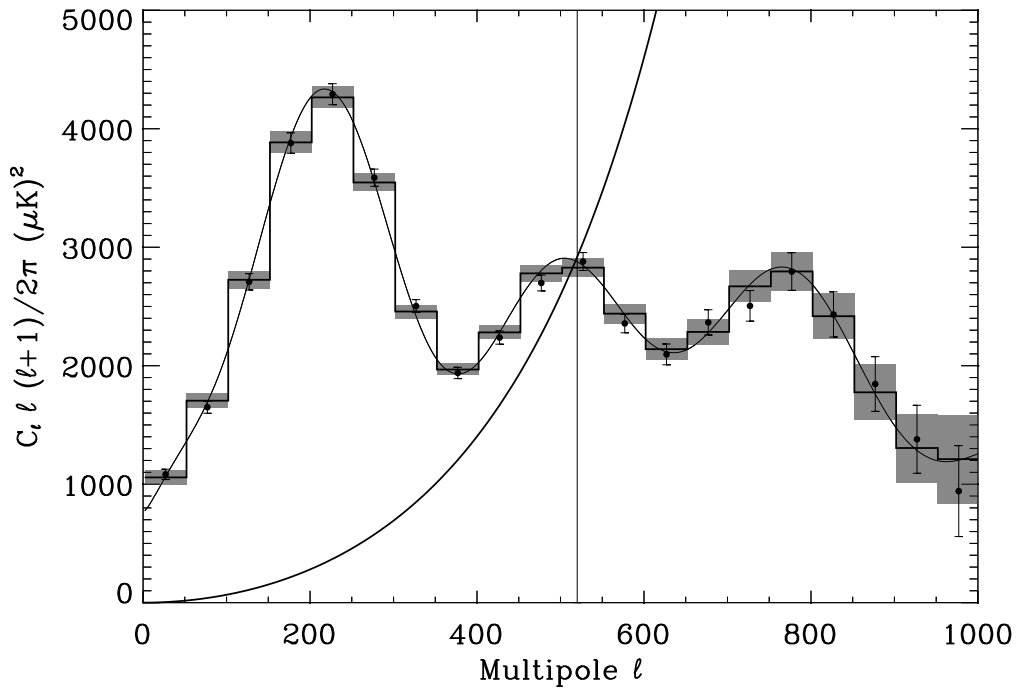


Figure 4.29: The result of a joint analysis of 146 patches on the same CMB sky. The solid line shows the average full sky power spectrum, the histogram shows the binned full sky power spectrum and the shaded boxes show the expected  $2\sigma$  deviations due to noise, cosmic and sample variance. The sizes of the shaded boxes were calculated from the approximate formula for uniform noise. The dots show the estimates with  $2\sigma$  error bars taken from the Fisher matrix. As before the rising solid line is the noise power spectrum and the vertical line shows where  $S/N = 1$ .

Finally I made a joint analysis of all the 146 patches. The idea was to extend the datavector in the likelihood so that it contained all the  $\tilde{C}_\ell$  from all the 146 patches. The datavector can then be written as  $\mathbf{d} = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{146}\}$  where  $\mathbf{d}_i$  now denotes the whole datavector for patch number  $i$ . From the results above it seems to be a good approximation to assume that the correlation between  $\tilde{C}_\ell$  from different patches is zero so that the correlation matrix will be block diagonal. Each block is then the correlation matrix for each individual patch. The log-likelihood can then simply be written as

$$L = \sum_{i=1}^{146} \mathbf{d}_i^T \mathbf{M}_i^{-1} \mathbf{d}_i + \sum_{i=1}^{146} \ln \det \mathbf{M}_i, \quad (4.110)$$

where  $\mathbf{M}_i$  is the correlation matrix for patch number  $i$ . In figure (4.29) the result of this joint analysis is shown. One can see that the full sky power spectrum is well within the two sigma error bars of the estimates.

The method of combining patches on the sky for power spectrum analysis will be developed further in a forthcoming paper (Górski and Hansen 2002).

### 4.3.2 Monte Carlo Simulations of the Noise Correlations and Extention to Correlated Noise

The computation of the noise correlation matrix in the general case takes  $\sqrt{N_{pix}}l_{max}^2(N^{in})^2$  which is approximately  $N_{pix}^{3/2}(N^{in})^2$ . When  $N^{in}$  is getting large this can be calculated quicker using Monte Carlo simulations (as was done in (Hivon et al. 2001)). Finding the  $\tilde{C}_\ell^N$  from one noise map takes  $\mathcal{O}(N_{pix}^{3/2})$  operations so using Monte Carlo simulations to find the whole noise matrix takes  $\mathcal{O}(N_{pix}^{3/2}N_{sim})$  operation where  $N_{sim}$  is the number of Monte Carlo simulations needed. So when  $N_{sim} \ll (N_{in}^2)$  it will be advantageous using MC if this gives the same result.

Also when the noise gets correlated, the analytic calculation of  $\langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell' m}^N \rangle$  will be very expensive. In this case another method for computing  $\langle \tilde{a}_{\ell m} \tilde{a}_{\ell' m} \rangle$  will be necessary and Monte Carlo simulations could also prove useful. For a given noise model several noise realisations can be made and averaged to yield the noise correlation matrix and the  $\langle \tilde{a}_{\ell m} \tilde{a}_{\ell' m} \rangle$  term needed in the estimation process. Using Monte Carlo the operations count for precalculations of noise properties will still be  $\mathcal{O}(N_{pix}^{3/2})$  also with correlated noise assuming that each noise realisation for a given noise model can be calculated quickly.



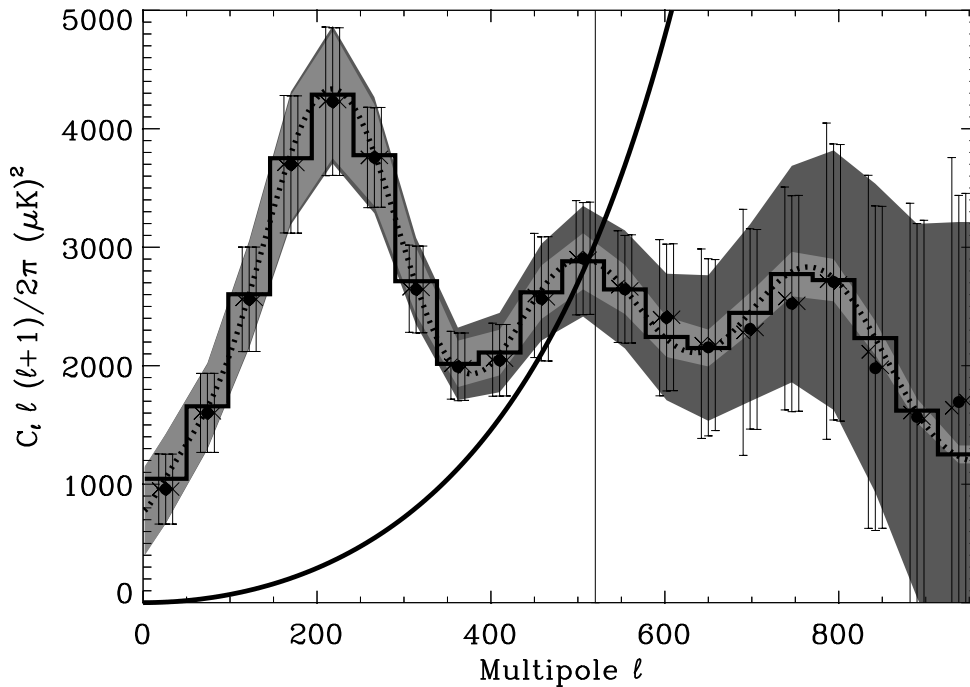


Figure 4.30: Same as figure (4.18) but with a different noise model. Here the average of 100 estimations is shown. The big dots in the middle of the bins show the result using analytical expressions for the noise matrices. The crosses on the left side of each big dot show the results of using noise matrices from  $N_{sim} = 1000$  MC simulations. The crosses on the right side are for  $N_{sim} = 20000$  noise simulations.

In figure (4.30), the result of  $C_\ell$  estimation with noise matrix and  $\langle a_{\ell m}^N a_{\ell' m}^N \rangle$  computed with Monte Carlo is shown. Again a standard CDM power spectrum was used with a non-uniform white noise model and a  $15^\circ$  FWHM Gaussian Gabor window. In the  $C_\ell$  estimation  $N^{in} = 200$   $\tilde{C}_\ell$  were used and  $N^{bin} = 20$  power spectrum bins were estimated. The noise matrices were calculated using (1) the analytical expression, (2) MC with  $N_{sim} = 20000$  and (3) MC with  $N_{sim} = 1000$ . In figure (4.31), a cut through the correlation matrices for the different cases is shown for  $\ell = 500$ . The dashed line (case (3)) follows the solid line (case (1)) to a level of about  $10^{-2}$  of the diagonal. The dotted line (case (2)) is roughly correct to about  $10^{-1}$  times the value at the diagonal.

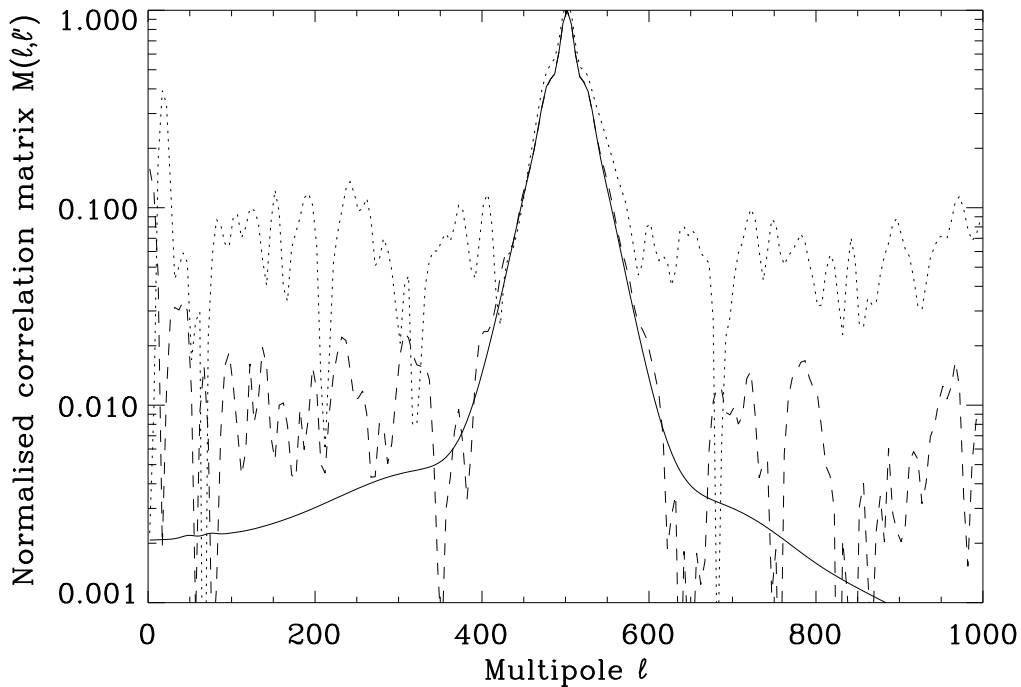


Figure 4.31: A cut through the noise correlation matrix at multipole  $\ell = 500$ . The correlation matrix was evaluated using the analytical formulae (solid line), 20000 MC simulations (dashed line) and 1000 MC simulations (dotted line). The matrix is here normalised to be 1 at the diagonal.

I did 100 estimations for each case and the average result is plotted in figure (4.30). The big dots are the results from case (1), the crosses on the right hand side are the results from (2) and the crosses on the left hand side the results from (3). The average estimates seem to be consistent, only in the highly noise dominated regime they start to deviate. For case (2), the error bars are for some multipoles higher and for some lower than the analytic case. The differences are at most 3%. I conclude that using this many simulations, the error bars do not increase significantly over the analytic case. For case (3) the error bars are up to 17% higher (and only higher) than the analytic case. It seems that 1000 simulations was not sufficient to keep the same accuracy of the estimates as when using analytic noise matrices. To keep the error bars as low as possible it seems that  $N_{sim} = 20000$  is a reasonable number of simulations.

## 4.4 Discussion

In this chapter I have discussed the Gabor transform on the sphere and how this can be used for CMB power spectrum estimation. A powerful tool has been presented for estimation of the full sky CMB power spectrum using small patches on the sky. The method is very fast and gives an unbiased estimate for the multipoles for which the Gaussian likelihood ansatz is valid. For the lowest multipoles a small downwards biasing is observed for small windows due to the skewed distribution of  $\tilde{C}_\ell$ . This can easily be solved by using information about low multipole  $C_\ell$  from larger patches. When estimating the power spectrum for larger parts of the sky this small bias disappears.

The method has been demonstrated to work very well on a single azimuthally symmetric patch on the CMB sky. It was also shown how the method can be extended to non-symmetric patches or even full sky by combining several patches in a joint likelihood analysis. This was only tested for the uniform noise case but will be explored further in future work. Also the case with correlated noise should in principle work using Monte Carlo methods to precompute noise matrices. This was tested for uncorrelated noise and compared to the analytic precalculation scheme. The results were getting more and more similar as the number of Monte Carlo simulations was increased. This is another extension which needs more work in the future.



# Chapter 5

## Gabor Transform on the Polarised CMB Sky

As discussed in chapter (1) and (2) the observation of the CMB polarisation power spectra will help breaking some parameter degeneracies. Several different sets of values for the cosmological parameters can give the same CMB temperature power spectrum. The polarisation power spectra will however be different and can distinguish the different models. Also the error bars on the cosmological parameters can be reduced by also exploiting the information present in the CMB polarisation power spectra.

Much effort has been made recently in order to find methods to analyse the CMB temperature power spectrum. For the even harder task of estimating the polarisation power spectra there has been very few publications. The framework for analysing the polarisation power spectra has been set in (Zaldarriaga and Seljak 1997; Kamionkowski, Kosowsky, and Stebbins 1997) but these papers only describe the full likelihood method which is far too time consuming also when only considering the temperature power spectrum.

In this chapter I will extend the method of using the pseudo power spectrum as input to a likelihood estimation of the power spectrum. I will include the  $E$  and  $C$  mode polarisation pseudo power spectra in the datavector and use techniques similar to those described in the previous chapter to estimate the power spectra. This can be done because the kernels that connect the full sky polarisation power spectra with the polarisation pseudo power spectra on an apodised sky are similar to the kernel for the temperature power spectrum. In the first part of this chapter I will derive the formulae for these kernels and for the polarisation pseudo power spectra and discuss their shapes. Then in the second part this will be used for likelihood estimation.

In this chapter the  $B$  component polarisation will mostly be neglected. The

$B$  polarisation power spectrum is expected to be very small and will hardly be detectable by the *MAP* or *Planck* experiments. Also the  $E$  and  $B$  components of polarisation mixes on the cut sky as will be discussed in this chapter, making the  $B$  polarisation pseudo spectrum to be dominated by the  $E$  component. This was also found independently by (Lewis, Challinor, and Turok 2001; Chiueh and Ma 2001).

## 5.1 The Gabor Transformation

### 5.1.1 Polarisation Powerspectra

As described in the first chapter the polarisation spherical harmonic coefficients are defined by means of the tensor spherical harmonics  ${}_2Y_{\ell m}(\hat{\mathbf{n}})$  as

$$a_{2,\ell m} = \int d\hat{\mathbf{n}} {}_2Y_{\ell m}^*(\hat{\mathbf{n}})(Q + iU)(\hat{\mathbf{n}}), \quad (5.1)$$

$$a_{-2,\ell m} = \int d\hat{\mathbf{n}} {}_{-2}Y_{\ell m}^*(\hat{\mathbf{n}})(Q - iU)(\hat{\mathbf{n}}), \quad (5.2)$$

and the inverse transforms are given as

$$(Q + iU)(\hat{\mathbf{n}}) = \sum_{\ell' m'} a_{2,\ell' m'} {}_2Y_{\ell' m'}(\hat{\mathbf{n}}) \quad (5.3)$$

$$(Q - iU)(\hat{\mathbf{n}}) = \sum_{\ell' m'} a_{-2,\ell' m'} {}_{-2}Y_{\ell' m'}(\hat{\mathbf{n}}). \quad (5.4)$$

It will be advantageous to write these spherical harmonics in terms of the rotation matrices  $D_{mm'}^\ell$  defined in Appendix (A). Using the formulae in Appendix (B) one can write

$${}_2Y_{\ell m}(\hat{\mathbf{n}}) = \sqrt{\frac{2\ell + 1}{4\pi}} D_{-2m}^\ell(\phi, \theta, 0), \quad (5.5)$$

$${}_{-2}Y_{\ell m}(\hat{\mathbf{n}}) = \sqrt{\frac{2\ell + 1}{4\pi}} D_{2m}^\ell(\phi, \theta, 0). \quad (5.6)$$

$$(5.7)$$

The corresponding complex conjugates can be written as (using the relations in Appendix (A))

$${}_2Y_{\ell m}^*(\hat{\mathbf{n}}) = \sqrt{\frac{2\ell + 1}{4\pi}} D_{-2m}^{\ell*}(\phi, \theta, 0) \quad (5.8)$$

$$= \sqrt{\frac{2\ell + 1}{4\pi}} (-1)^m D_{2-m}^\ell(\phi, \theta, 0), \quad (5.9)$$

$${}_{-2}Y_{\ell m}^*(\hat{\mathbf{n}}) = \sqrt{\frac{2\ell + 1}{4\pi}} D_{2m}^{\ell*}(\phi, \theta, 0) \quad (5.10)$$

$$= \sqrt{\frac{2\ell + 1}{4\pi}} (-1)^m D_{-2-m}^\ell(\phi, \theta, 0). \quad (5.11)$$

Finally as explained in chapter (1) the power spectrum can be written in terms of a ‘divergence free’  $E$  component and a ‘curl free’  $B$  component

$$a_{E,\ell m} = -\frac{1}{2}(a_{2,\ell m} + a_{-2,\ell m}), \quad (5.12)$$

$$a_{B,\ell m} = \frac{1}{2}i(a_{2,\ell m} - a_{-2,\ell m}) \quad (5.13)$$

Now I will define the windowed coefficients  $\tilde{a}_{\ell m}$  and  $\tilde{C}_\ell$  for polarisation in an analogous way as for temperature. As in chapter (4) I Legendre expand the Gabor Window  $G(\theta)$  which is an axisymmetric function centered at  $\hat{\mathbf{n}}_0$ ,

$$G(\theta) = \sum_{\ell''} \frac{2\ell'' + 1}{4\pi} g_{\ell''} P_{\ell''}(\cos \theta) = \sum_{\ell'' m''} g_{\ell''} Y_{\ell'' m''}(\hat{\mathbf{n}}) Y_{\ell'' m''}^*(\hat{\mathbf{n}}_0). \quad (5.14)$$

$$= \sum_{\ell'' m''} g_{\ell''} \sqrt{\frac{2\ell'' + 1}{4\pi}} D_{0m''}^{\ell''}(\phi, \theta, 0) Y_{\ell'' m''}(\hat{\mathbf{n}}_0). \quad (5.15)$$

I define the windowed coefficients  $\tilde{a}_{\ell m}$  as

$$\tilde{a}_{2,\ell m} = \int d\hat{\mathbf{n}} {}_2Y_{\ell m}^*(\hat{\mathbf{n}}) (Q + iU)(\hat{\mathbf{n}}) G(\hat{\mathbf{n}}, \hat{\mathbf{n}}_0) \quad (5.16)$$

Using the expression for (5.3)  $(Q + iU)(\hat{\mathbf{n}})$  and writing all  ${}_2Y_{\ell m}$  as  $D$ -matrices using expressions (5.5), (5.6), (5.9) and (5.11) one gets,

$$\tilde{a}_{2,\ell m} = \sum_{\ell' m'} a_{2,\ell' m'} \sum_{\ell'' m''} g_{\ell''} \frac{\sqrt{(2\ell + 1)(2\ell' + 1)(2\ell'' + 1)}}{(4\pi)^{3/2}} (-1)^m \quad (5.17)$$

$$\times Y_{\ell'' m''}(\hat{\mathbf{n}}_0) \int d\hat{\mathbf{n}} D_{-2m'}^{\ell'}(\phi, \theta, 0) D_{2-m}^{\ell}(\phi, \theta, 0) D_{0m''}^{\ell''}(\phi, \theta, 0) \quad (5.18)$$

$$= \sum_{\ell' m'} a_{2,\ell' m'} \sum_{\ell'' m''} g_{\ell''} \frac{\sqrt{(2\ell + 1)(2\ell' + 1)(2\ell'' + 1)}}{(4\pi)^{3/2}} (-1)^m \frac{1}{2\pi} \quad (5.19)$$

$$\times Y_{\ell'' m''}(\hat{\mathbf{n}}_0) \int d\hat{\mathbf{n}} d\gamma D_{-2m'}^{\ell'}(\phi, \theta, \gamma) D_{2-m}^{\ell}(\phi, \theta, \gamma) D_{0m''}^{\ell''}(\phi, \theta, \gamma) \quad (5.20)$$

$$= \sum_{\ell' m'} a_{2,\ell' m'} h_2(\ell, \ell', m, m', \hat{\mathbf{n}}_0). \quad (5.21)$$

By using equation (5.16), one can also write this as,

$$\tilde{a}_{2,\ell n} = \sum_{\ell' m'} a_{2,\ell' m'} \int d\hat{\mathbf{n}} G(\hat{\mathbf{n}}, \hat{\mathbf{n}}_0) {}_2Y_{\ell m}^*(\hat{\mathbf{n}}) {}_2Y_{\ell' m'}(\hat{\mathbf{n}}). \quad (5.22)$$

Using the two last expressions, the  $h_2$  function can be written in two ways (using relation (C.2) for the last expression),

$$h_2(\ell, \ell', m, m', \hat{\mathbf{n}}_0) \equiv \int d\hat{\mathbf{n}} G(\hat{\mathbf{n}}, \hat{\mathbf{n}}_0) {}_2Y_{\ell m}^*(\hat{\mathbf{n}}) {}_2Y_{\ell' m'}(\hat{\mathbf{n}}) \quad (5.23)$$

$$= \sum_{\ell'' m''} g_{\ell''} \sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}} (-1)^m \quad (5.24)$$

$$\times Y_{\ell'' m''}(\hat{\mathbf{n}}_0) \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m' & -m & m'' \end{pmatrix}. \quad (5.25)$$

As I soon will show, the polarisation pseudo power spectra are rotationally invariant under rotation of the Gabor window. For that reason, one can put the centre of the Gabor window on the north pole giving,

$$\tilde{a}_{2,\ell m} = \sum_{\ell'} a_{2,\ell' m} h_2(\ell, \ell', m), \quad (5.26)$$

where,

$$h_2(\ell, \ell', m) = h_2(\ell, \ell', m, m, 0) \quad (5.27)$$

$$= \sum_{\ell''} g_{\ell''} \frac{\sqrt{(2\ell+1)(2\ell'+1)(2\ell''+1)}}{4\pi} \quad (5.28)$$

$$\times (-1)^m \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & -m & 0 \end{pmatrix}. \quad (5.29)$$

Similarly one gets,

$$\tilde{a}_{-2,\ell m} = \sum_{\ell'} a_{-2,\ell' m} h_2(\ell, \ell', -m) \quad (5.30)$$

and

$$\tilde{a}_{E,\ell m} = -\frac{1}{2}(\tilde{a}_{2,\ell m} + \tilde{a}_{-2,\ell m}) \quad (5.31)$$

$$= \sum_{\ell'} a_{E,\ell' m} H_2(\ell, \ell', m) + i \sum_{\ell'} a_{B,\ell' m} H_{-2}(\ell, \ell', m) \quad (5.32)$$

$$\tilde{a}_{B,\ell m} = i\frac{1}{2}(\tilde{a}_{2,\ell m} - \tilde{a}_{-2,\ell m}) \quad (5.33)$$

$$= \sum_{\ell'} a_{B,\ell' m} H_2(\ell, \ell', m) - i \sum_{\ell'} a_{E,\ell' m} H_{-2}(\ell, \ell', m) \quad (5.34)$$

Please note that whereas  $h(\ell, \ell', -m) = h(\ell, \ell', m)$  a similar relation does not exist for  $h_2(\ell, \ell', -m)$ . Using the expression above, one has that,

$$h_2(\ell, \ell', -m) = \sum_{\ell''} g_{\ell''} \frac{\sqrt{(2\ell+1)(2\ell'+1)(2\ell''+1)}}{4\pi} (-1)^m \quad (5.35)$$

$$\times \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & -m & 0 \end{pmatrix} (-1)^{\ell+\ell'+\ell''} \quad (5.36)$$



The reason why this is not equal to  $h_2(\ell, \ell', m)$  is that the first Wigner symbol is not zero when  $\ell + \ell' + \ell''$  is even, which is the case when the whole lower row in the Wigner symbol is 0, as in the case with  $h(\ell, \ell', m)$ . It is also obvious from the expression (5.22). For the scalar case, the relation  $Y_{\ell(-m)}(\hat{\mathbf{n}}) = (-1)^m Y_{\ell m}(\hat{\mathbf{n}})$  ensures that there is no dependency on  $m$  in  $h(\ell, \ell', m)$  whereas a similar relation does not exist for the tensor harmonics (but see relation (B.5)).

I have further defined,

$$H_2(\ell, \ell', m) = \frac{1}{2} (h_2(\ell, \ell', m) + h_2(\ell, \ell', -m)) \quad (5.37)$$

$$H_{-2}(\ell, \ell', m) = \frac{1}{2} (h_2(\ell, \ell', m) - h_2(\ell, \ell', -m)) \quad (5.38)$$

which contrary to  $h_2(\ell, \ell', m)$  have an  $m$ -symmetry

$$H_2(\ell, \ell', -m) = H_2(\ell, \ell', m) \quad (5.39)$$

$$H_{-2}(\ell, \ell', -m) = -H_{-2}(\ell, \ell', m) \quad (5.40)$$

To find the  $\tilde{C}_\ell$  and (later) the correlation matrices, the following quantities will be needed

$$\langle \tilde{a}_{E,\ell m} \tilde{a}_{E,\ell',m'}^* \rangle = \delta_{mm'} \left[ \sum_{\ell''} C_{\ell''}^E H_2(\ell, \ell'', m) H_2(\ell', \ell'', m) \right] \quad (5.41)$$

$$+ \sum_{\ell''} C_{\ell''}^B H_{-2}(\ell, \ell'', m) H_{-2}(\ell', \ell'', m) \quad (5.42)$$

$$\langle \tilde{a}_{B,\ell m} \tilde{a}_{B,\ell',m'}^* \rangle = \delta_{mm'} \left[ \sum_{\ell''} C_{\ell''}^B H_2(\ell, \ell'', m) H_2(\ell', \ell'', m) \right] \quad (5.43)$$

$$+ \sum_{\ell''} C_{\ell''}^E H_{-2}(\ell, \ell'', m) H_{-2}(\ell', \ell'', m) \quad (5.44)$$

$$\langle \tilde{a}_{E,\ell m} \tilde{a}_{B,\ell',m'}^* \rangle = \delta_{mm'} i \left[ \sum_{\ell''} C_{\ell''}^E H_2(\ell, \ell'', m) H_{-2}(\ell', \ell'', m) \right] \quad (5.45)$$

$$+ \sum_{\ell''} C_{\ell''}^B H_{-2}(\ell, \ell'', m) H_2(\ell', \ell'', m) \quad (5.46)$$

$$\langle \tilde{a}_{E,\ell m} \tilde{a}_{\ell',m'}^* \rangle = \delta_{mm'} \sum_{\ell''} C_{\ell''}^C H_2(\ell, \ell'', m) h(\ell', \ell'', m) \quad (5.47)$$

$$\langle \tilde{a}_{B,\ell m} \tilde{a}_{\ell',m'}^* \rangle = -i \delta_{mm'} \sum_{\ell''} C_{\ell''}^C H_{-2}(\ell, \ell'', m) h(\ell', \ell'', m) \quad (5.48)$$

For  $\langle \tilde{C}_\ell^E \rangle$  one now has,

$$\langle \tilde{C}_\ell^E \rangle = \sum_m \frac{\langle \tilde{a}_{E,\ell m}^* \tilde{a}_{E,\ell m} \rangle}{2\ell + 1} = \sum_{\ell'} C_{\ell'}^E K_2(\ell, \ell') + \sum_{\ell'} C_{\ell'}^B K_{-2}(\ell, \ell'). \quad (5.49)$$

Where,

$$K_{\pm 2}(\ell, \ell') = \frac{1}{2\ell + 1} \sum_m H_{\pm 2}^2(\ell, \ell', m) \quad (5.50)$$

Using the expression for  $h_2(\ell, \ell', m)$ , one gets

$$K_{\pm 2}(\ell, \ell') = \frac{1}{2} \frac{1}{2\ell + 1} \sum_m \left( h_2^2(\ell, \ell', m) \pm h_2(\ell, \ell', m)h_2(\ell, \ell', -m) \right) \quad (5.51)$$

$$= \sum_{\ell'' L''} g_{\ell''} g_{L''} \frac{(2\ell' + 1)(2\ell'' + 1)(2L'' + 1)}{32\pi^2} \quad (5.52)$$

$$\times \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & L'' \\ -2 & 2 & 0 \end{pmatrix} (1 \pm (-1)^{\ell + \ell' + \ell''}) \quad (5.53)$$

$$\times \underbrace{\sum_m \begin{pmatrix} \ell & \ell' & \ell'' \\ m & -m & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & L'' \\ m & -m & 0 \end{pmatrix}}_{(2\ell'' + 1)^{-1} \delta_{\ell'' L''}} \quad (5.54)$$

$$= \sum_{\ell''} g_{\ell''}^2 \frac{(2\ell' + 1)(2\ell'' + 1)}{32\pi^2} \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix}^2 \quad (5.55)$$

$$\times (1 \pm (-1)^{\ell + \ell' + \ell''}). \quad (5.56)$$

Similarly,

$$\langle \tilde{C}_\ell^B \rangle = \sum_{\ell'} C_{\ell'}^B K_2(\ell, \ell') + \sum_{\ell'} C_{\ell'}^E K_{-2}(\ell, \ell') \quad (5.57)$$

$$\langle \tilde{C}_\ell^C \rangle = \sum_{\ell'} C_{\ell'}^C K_{20}(\ell, \ell') \quad (5.58)$$

$$K_{20}(\ell, \ell') = \frac{1}{2\ell + 1} \sum_m H_2(\ell, \ell', m) h(\ell, \ell', m) \quad (5.59)$$

$$= \sum_{\ell''} g_{\ell''}^2 \frac{(2\ell' + 1)(2\ell'' + 1)}{16\pi^2} \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}. \quad (5.60)$$

Before studying these kernels I will first show the rotational invariance. To show that  $\langle \tilde{C}_\ell^E \rangle$ ,  $\langle \tilde{C}_\ell^B \rangle$  and  $\langle \tilde{C}_\ell^C \rangle$  are rotationally invariant under rotations of the window, one can keep the dependency on the angle and the expression for the kernel turns out to be,

$$K_{\pm 2}(\ell, \ell', \hat{\mathbf{n}}_0) = \frac{1}{2} \frac{1}{2\ell + 1} \sum_{mm'} \left( h_2^2(\ell, \ell', m, m', \hat{\mathbf{n}}_0) \right) \quad (5.61)$$

$$\pm h_2(\ell, \ell', m, m', \hat{\mathbf{n}}_0) h_2(\ell, \ell', -m, -m', \hat{\mathbf{n}}_0) \quad (5.62)$$

$$= \sum_{\ell'' L''} \sum_{m'' M''} g_{\ell''} g_{L''} \frac{(2\ell' + 1) \sqrt{(2\ell'' + 1)(2L'' + 1)}}{8\pi} \quad (5.63)$$

$$\times Y_{\ell''m''}(\hat{\mathbf{n}}_0)Y_{L''M''}^*(\hat{\mathbf{n}}_0)\begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix}\begin{pmatrix} \ell & \ell' & L'' \\ -2 & 2 & 0 \end{pmatrix} \quad (5.64)$$

$$\times \underbrace{\sum_{mm'} \begin{pmatrix} \ell & \ell' & \ell'' \\ m' & -m & m'' \end{pmatrix} \begin{pmatrix} \ell & \ell' & L'' \\ m' & -m & M'' \end{pmatrix}}_{(2\ell''+1)^{-1}\delta_{\ell''L''}\delta_{m''M''}} \quad (5.65)$$

$$\times (1 \pm (-1)^{\ell+\ell'+\ell''}) \quad (5.66)$$

$$= \sum_{\ell''} g_{\ell''}^2 \frac{2\ell'+1}{8\pi} \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix}^2 \underbrace{\sum_{m''} |Y_{\ell''m''}(\hat{\mathbf{n}}_0)|^2}_{\frac{2\ell''+1}{4\pi}} \quad (5.67)$$

$$\times (1 \pm (-1)^{\ell+\ell'+\ell''}) \quad (5.68)$$

$$= K_{\pm 2}(\ell, \ell'). \quad (5.69)$$

This is independent on the angle  $\hat{\mathbf{n}}_0$  which shows the rotational invariance of the pseudo polarisation power spectra.

As with the temperature kernels, the polarisation kernels can be evaluated either using the analytical Wigner symbol expressions (5.55) and (5.58) or faster by direct integration and recursion of the  $H_{\pm 2}$  functions. The recursion for the  $h_2(\ell, \ell', m, m')$  functions in appendix (E) is one of the major results in the thesis. This is an extension of the recursion for  $h(\ell, \ell', m, m')$  in appendix (D).

Studying equation (5.49) and (5.57) one sees that the  $E$  and  $B$  modes are mixing when only a portion of the sky is observed. The kernel  $K_2(\ell, \ell')$  is the kernel which takes full sky  $C_\ell^E$  modes to the pseudo coefficients  $\tilde{C}_\ell^E$  and similarly for  $C_\ell^B$ . The kernel  $K_{-2}(\ell, \ell')$  is the one which causes the mixing. In figure (5.1) and (5.2) I have plotted the kernels  $K_2(\ell, \ell')$  together with  $K_{-2}(\ell, \ell')$  for a 5 and 15 degree FWHM Gaussian Gabor window with  $\theta_C = 3\sigma$ . One can see that the diagonal of the  $K_2(\ell, \ell')$  kernel is about an order of magnitude larger than the diagonal of the mixing kernel  $K_{-2}(\ell, \ell')$ . This means that  $C_\ell^E$  would dominate  $\tilde{C}_\ell^E$  and  $C_\ell^B$  would dominate  $\tilde{C}_\ell^B$  provided that the two spectra  $C_\ell^E$  and  $C_\ell^B$  were of the same order of magnitude. However as discussed before, the  $C_\ell^B$  are expected to be considerably smaller than  $C_\ell^E$  in most cosmological models. For the  $E$  mode this is not a problem as  $C_\ell^E$  will then hardly be affected. The problem is the  $B$  mode which in this case will be dominated by the  $E$  mode.

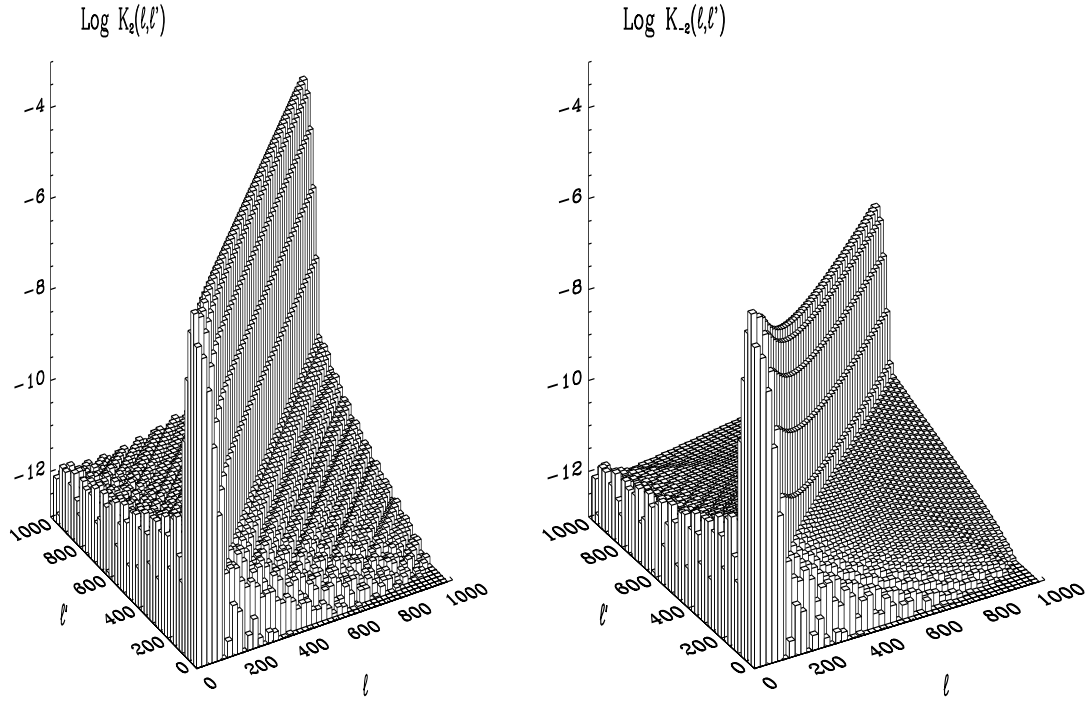


Figure 5.1: The kernels  $K_2(\ell, \ell')$  (left plot) and  $K_{-2}(\ell, \ell')$  (right plot) connecting full and cut sky polarisation power spectra  $C_\ell^E$  and  $C_\ell^B$ . The left kernel is the one which takes full sky  $C_\ell^E$  into cut sky  $\tilde{C}_\ell^E$  and full sky  $C_\ell^B$  into cut sky  $\tilde{C}_\ell^B$ . The right kernel is the one which mixes the two giving contributions from full sky  $C_\ell^E$  in cut sky  $\tilde{C}_\ell^B$  and vice versa.

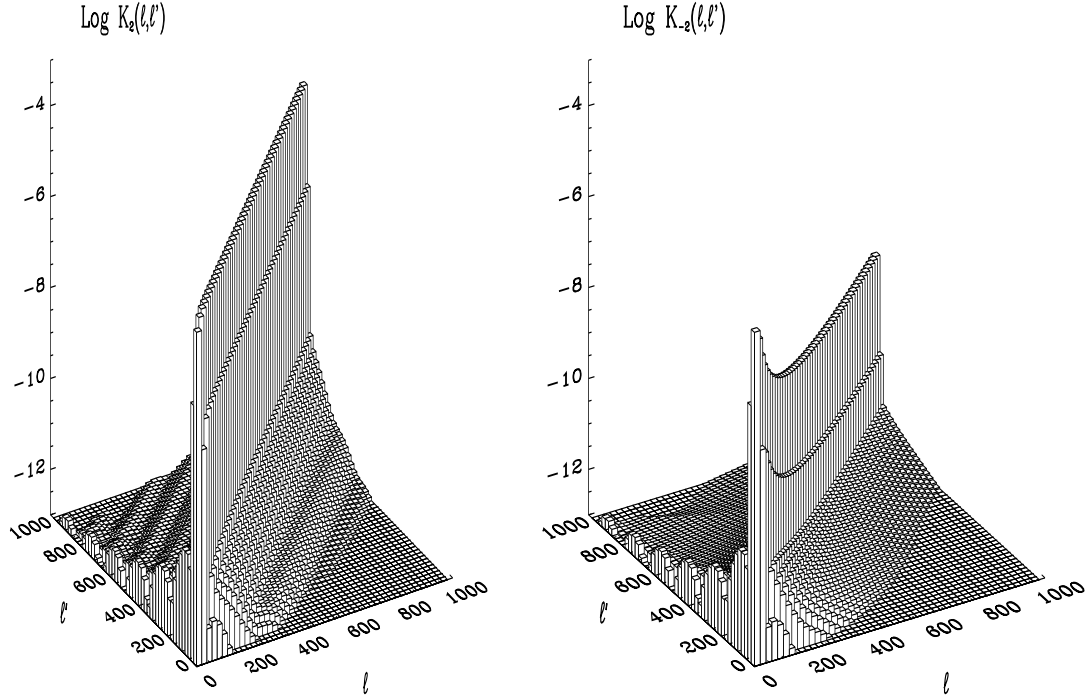


Figure 5.2: The same as figure (5.1) for a 15 degree FWHM Gaussian Gabor window.

The separation of  $E$  and  $B$  modes of polarisation on the cut sky was already discussed in (Lewis, Challinor, and Turok 2001; Chiueh and Ma 2001). Here I will only note that an easy way of separating the two would be to construct two new coefficients

$$a_{+, \ell m} = a_{E, \ell m} + i a_{B, \ell m} \quad (5.70)$$

$$a_{-, \ell m} = a_{E, \ell m} - i a_{B, \ell m}, \quad (5.71)$$

with the corresponding pseudo quantities

$$\tilde{a}_{+, \ell m} = \sum_{\ell'} a_{+, \ell' m} (H_2(\ell, \ell', m) + H_{-2}(\ell, \ell', m)) \quad (5.72)$$

$$\tilde{a}_{-, \ell m} = \sum_{\ell'} a_{-, \ell' m} (H_2(\ell, \ell', m) - H_{-2}(\ell, \ell', m)) \quad (5.73)$$

$$(5.74)$$

For the power spectra one gets,

$$\langle C_\ell^+ \rangle \equiv \sum_m \frac{\langle a_{+, \ell m} a_{+, \ell m}^* \rangle}{2\ell + 1} = \langle C_{\ell m}^E \rangle + \langle C_{\ell m}^B \rangle, \quad (5.75)$$

$$\langle C_\ell^- \rangle \equiv \sum_m \frac{\langle a_{-\ell m} a_{-\ell m}^* \rangle}{2\ell + 1} = \langle C_{\ell m}^E \rangle - \langle C_{\ell m}^B \rangle. \quad (5.76)$$

$$(5.77)$$

To get the pseudo power spectra one can use equations (5.72) and (5.73) to get

$$\langle \tilde{C}_\ell^+ \rangle = \sum_{\ell'} C_\ell^+ K_+(\ell, \ell'), \quad (5.78)$$

$$\langle \tilde{C}_\ell^- \rangle = \sum_{\ell'} C_\ell^- K_-(\ell, \ell'), \quad (5.79)$$

where the kernels can be written

$$K_+(\ell, \ell') = \frac{1}{2\ell + 1} \sum_m (H_2(\ell, \ell, m) + H_{-2}(\ell, \ell', m))^2, \quad (5.80)$$

$$K_-(\ell, \ell') = \frac{1}{2\ell + 1} \sum_m (H_2(\ell, \ell, m) - H_{-2}(\ell, \ell', m))^2. \quad (5.81)$$

In figure (5.3) and (5.4) I have plotted the  $K_2(\ell, \ell')$  and  $K_{-2}(\ell, \ell')$  kernels for a tophat window covering the same area on the sky as the Gaussian windows in figure (5.1) and (5.2).

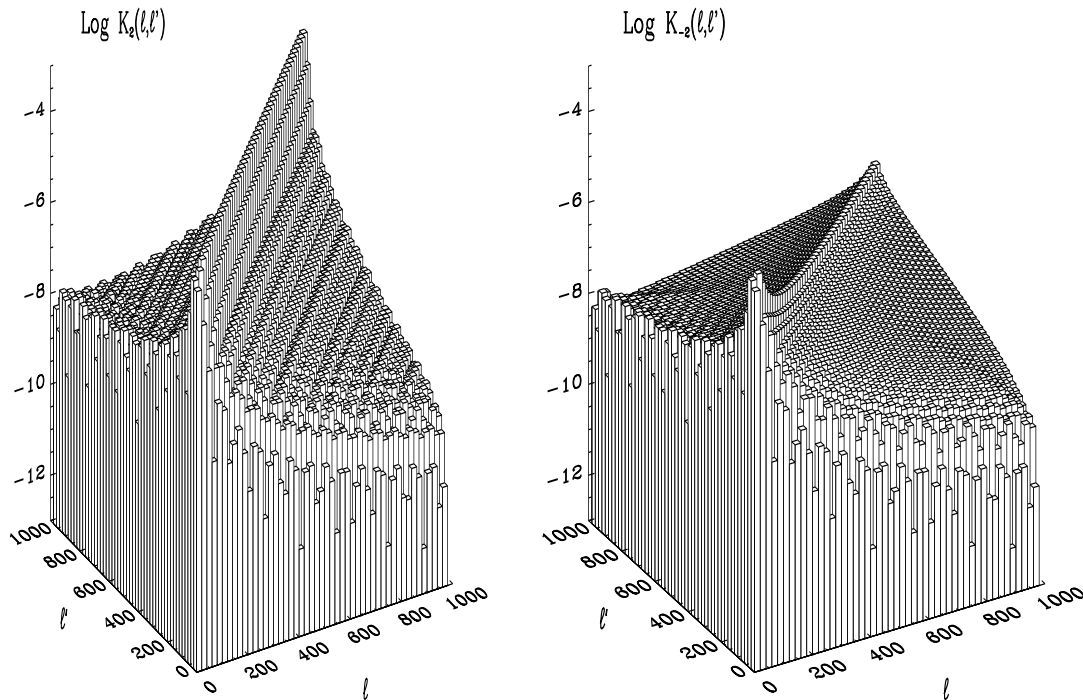


Figure 5.3: The same as figure (5.1) for a tophat Gabor window covering the same area on the sky as the Gaussian window in figure (5.1)

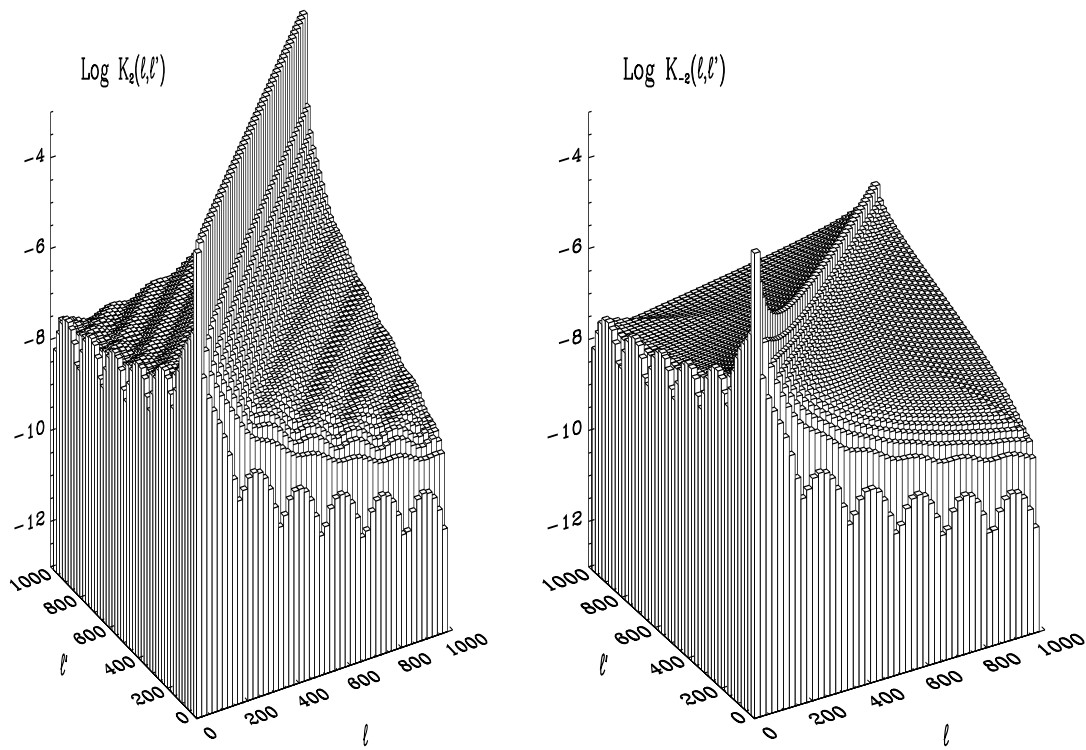


Figure 5.4: The same as figure (5.2) for a tophat Gabor window covering the same area on the sky as the Gaussian window in figure (5.2)

The kernel  $K_{20}(\ell, \ell')$  for the cross polarisation power spectrum  $C_\ell^C$  is shown in figure (5.5) for a 5 and 15 degree Gaussian window and in figure (5.6) for the corresponding tophat windows.

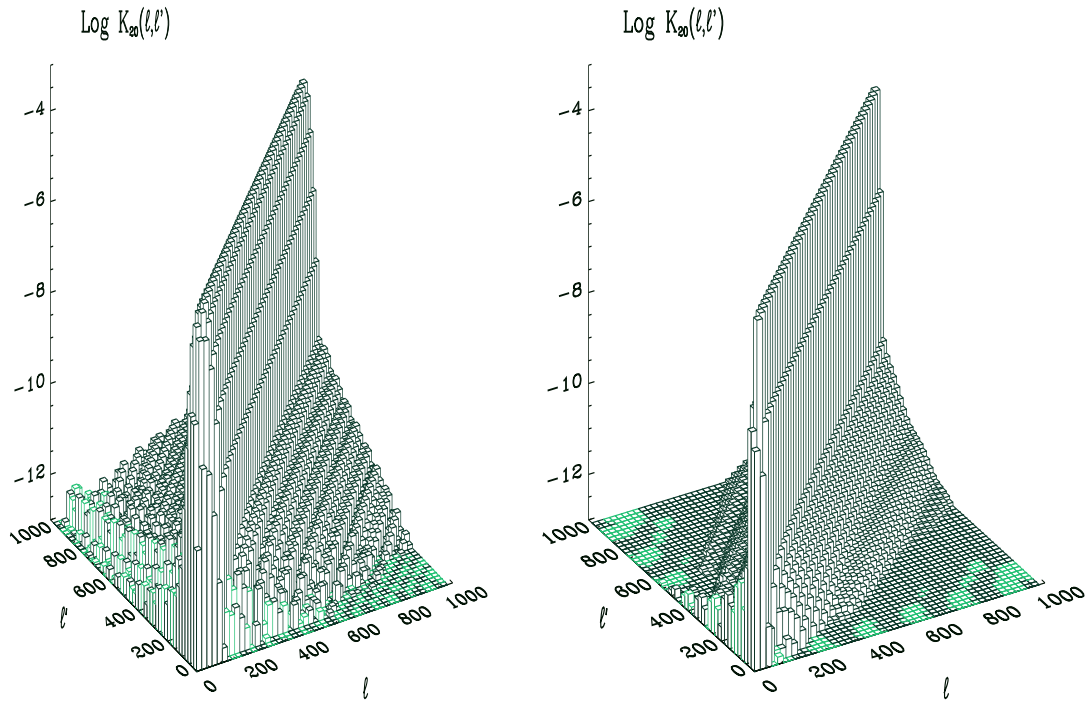


Figure 5.5: The kernel  $K_{20}(l, l')$  connecting the full sky cross polarisation spectrum  $C_\ell^C$  and the cut sky spectrum  $\tilde{C}_\ell^C$  for a 5 (left plot) and 15 (right plot) degree FWHM Gaussian Gabor window. The negative elements have a brighter colour.



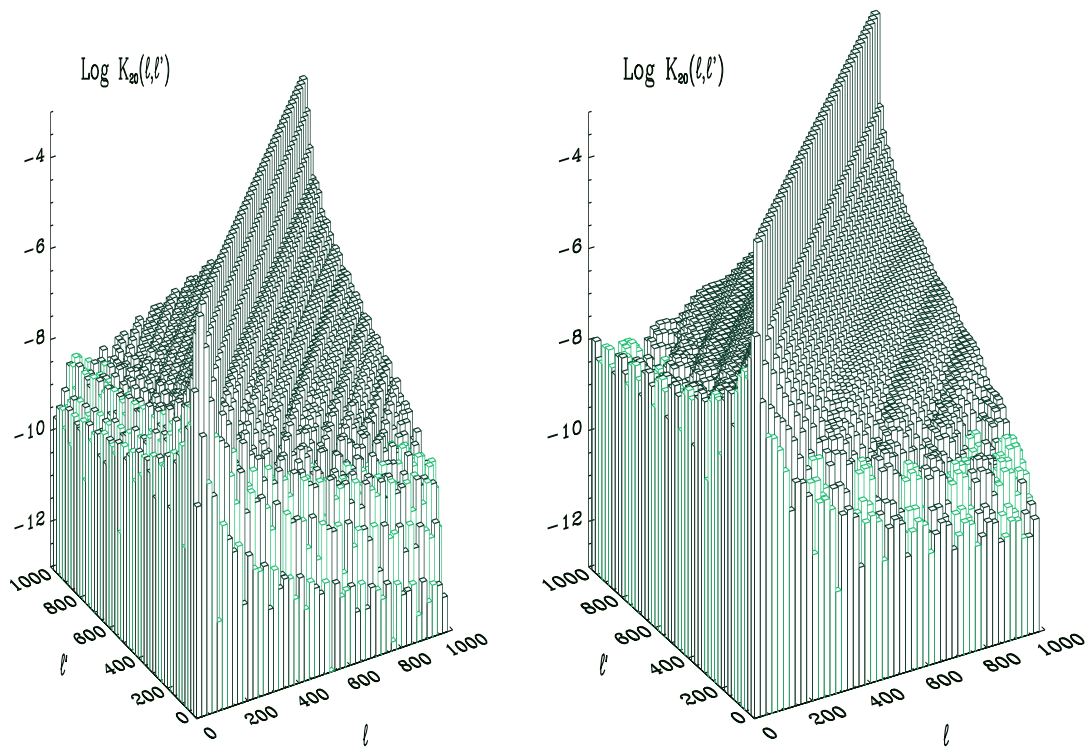


Figure 5.6: Same as figure (5.5) for the corresponding tophat windows.

As for the temperature kernels, all the polarisation kernels show the same behaviour when changing type and size of the window. When going from smaller to larger windows, the diagonals get sharper. Also the tophat kernels have more long range correlations than the Gaussian kernels (note that all the plots have the same vertical scale and can be compared directly).

In figure (5.7) I have plotted a cut through the different kernels at  $\ell = 200$  for comparison. The cuts are made through the kernels for 5 and 15 degree FWHM Gaussian Gabor windows. The first thing to note is that the temperature kernel  $K(\ell, \ell')$ , the  $E$  and  $B$  kernel  $K_2(\ell, \ell')$  and the temperature-polarisation cross spectrum kernel  $K_{20}(\ell, \ell')$  only differ for the far off-diagonal elements. At the diagonal their shape and size are the same. For this reason the relation shown in figure (4.5) between the width of the kernel and the width of the size of the window is also valid for polarisation. This is an important result to be used for the likelihood estimation of the polarisation power spectra in the next section. It shows that the number of polarisation pseudo spectrum coefficients to be used in the likelihood analysis should be the same as for the likelihood estimation of the temperature power spectrum.

In figure (5.8) a similar plot is shown for the corresponding tophat windows. The plot shows that the conclusions made for the Gaussian windows are also valid in this case. The shape and size of the three kernels are the same around the diagonal. For this reason the results shown for the temperature kernel that the tophat window has larger long range correlations whereas the Gaussian window has large short range correlations and therefore a wider kernel is also valid for polarisation.

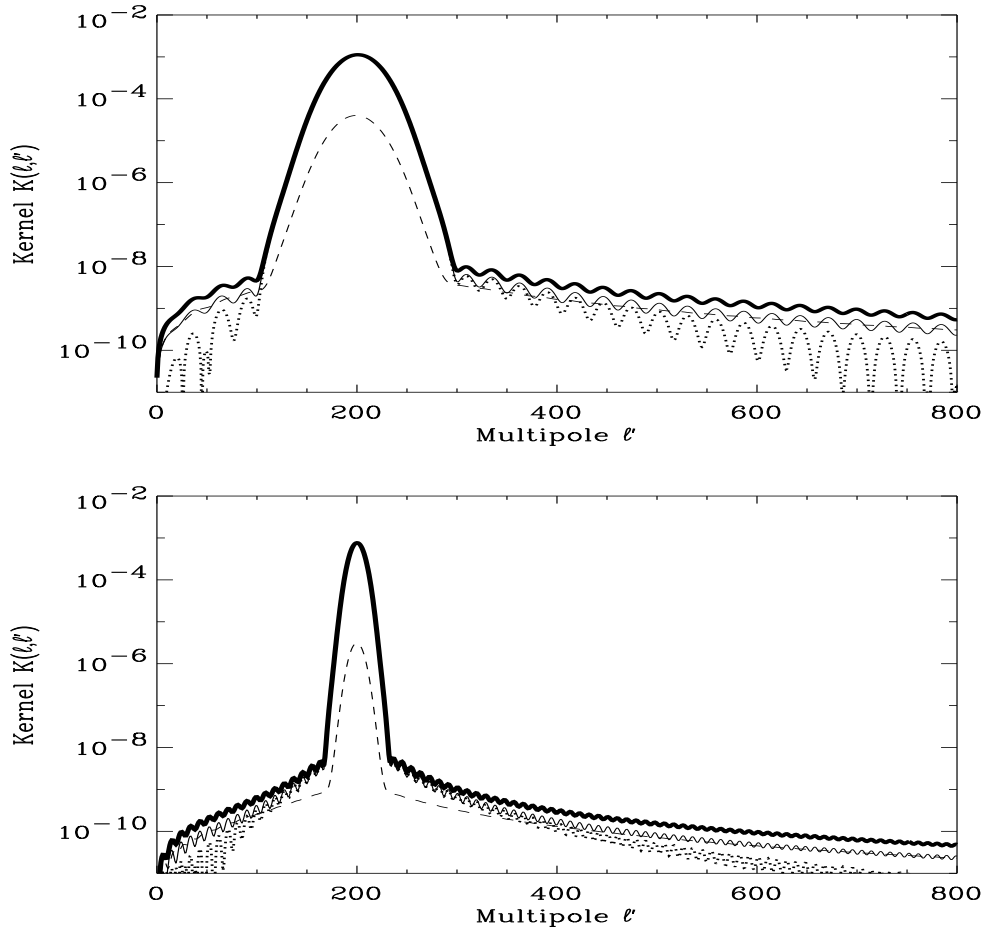


Figure 5.7: A cut at  $\ell = 200$  through the kernels combining the full sky and cut sky power spectra. The thick solid line is the kernel  $K(\ell, \ell')$  for the temperature power spectrum, the thin solid line is the kernel  $K_2(\ell, \ell')$  for  $E$  and  $B$  mode polarisation and the dotted line is the kernel for the temperature-polarisation cross power spectrum  $K_{20}(\ell, \ell')$ . All these kernels go together around the diagonal. They only differ for the far off-diagonal elements. The dashed line is the mixing kernel  $K_{-2}(\ell, \ell')$  which mixes the  $E$  and  $B$  mode polarisation power spectra on the cut sky. This kernel is lower than the other kernels. The upper plot is for a 5 degree Gaussian Gabor window and the lower plot for a 15 degree FWHM Gaussian window.

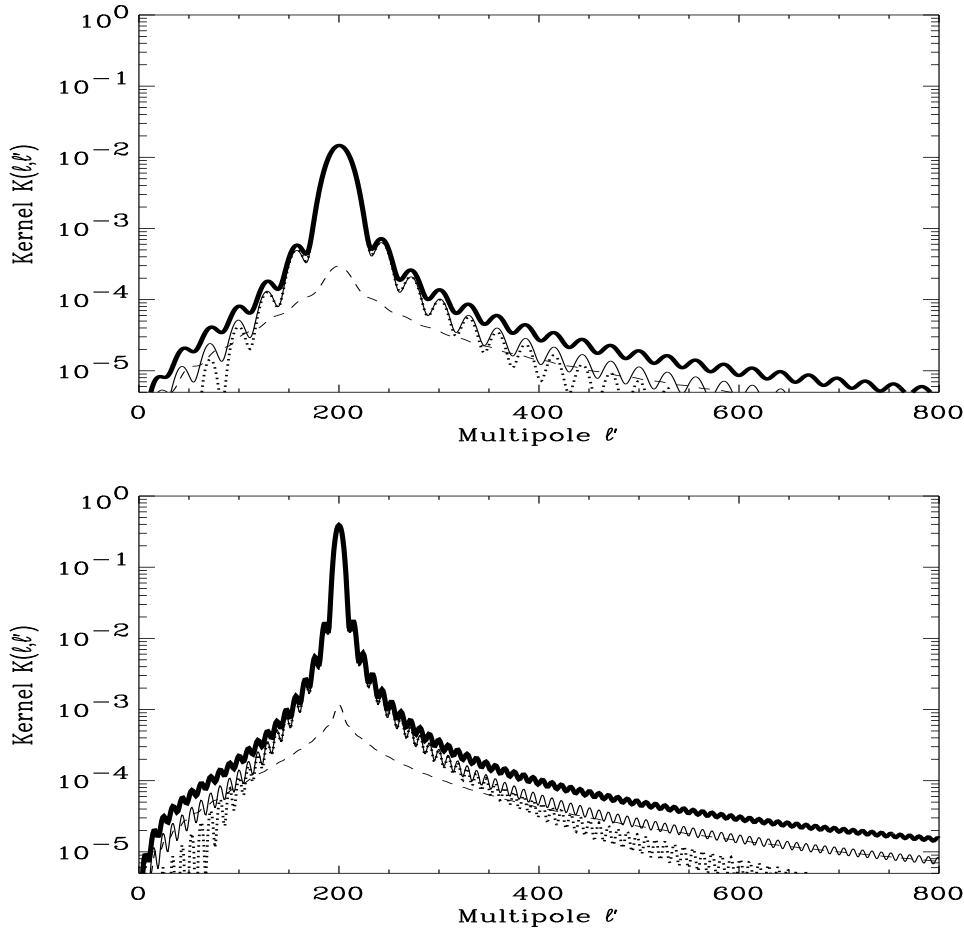


Figure 5.8: Same as figure (5.7) but for tophat windows covering the same area on the sky.

The kernel  $K_{-2}(\ell, \ell')$  which mixes the  $E$  and  $B$  modes on the cut sky is plotted as a dashed line in figures (5.7) and (5.8). It is much smaller than the three other kernels and the shape seems to differ as well. Note that the height of the mixing kernel relative to the other kernels is lower for the 15 degree window than for the 5 degree window. That the size of the mixing kernel relative to the other kernels is dropping when the size of the window is increasing was to be expected since in the limit of full sky coverage the  $E$ - $B$  mixing disappears and the mixing kernel must go to zero.

In figure (5.9) a cut at  $\ell = 200$  through the temperature kernel and the mixing kernel is shown for the 5 and 15 degree FWHM Gaussian Gabor window. The kernels are normalised to one at the peak so that the shapes can be compared. For the Gaussian window, the shapes of the kernels still seem to be the same.

But the kernels for the corresponding tophat windows shown in figure (5.10) does not have a Gaussian shape and differs significantly from the other kernels.

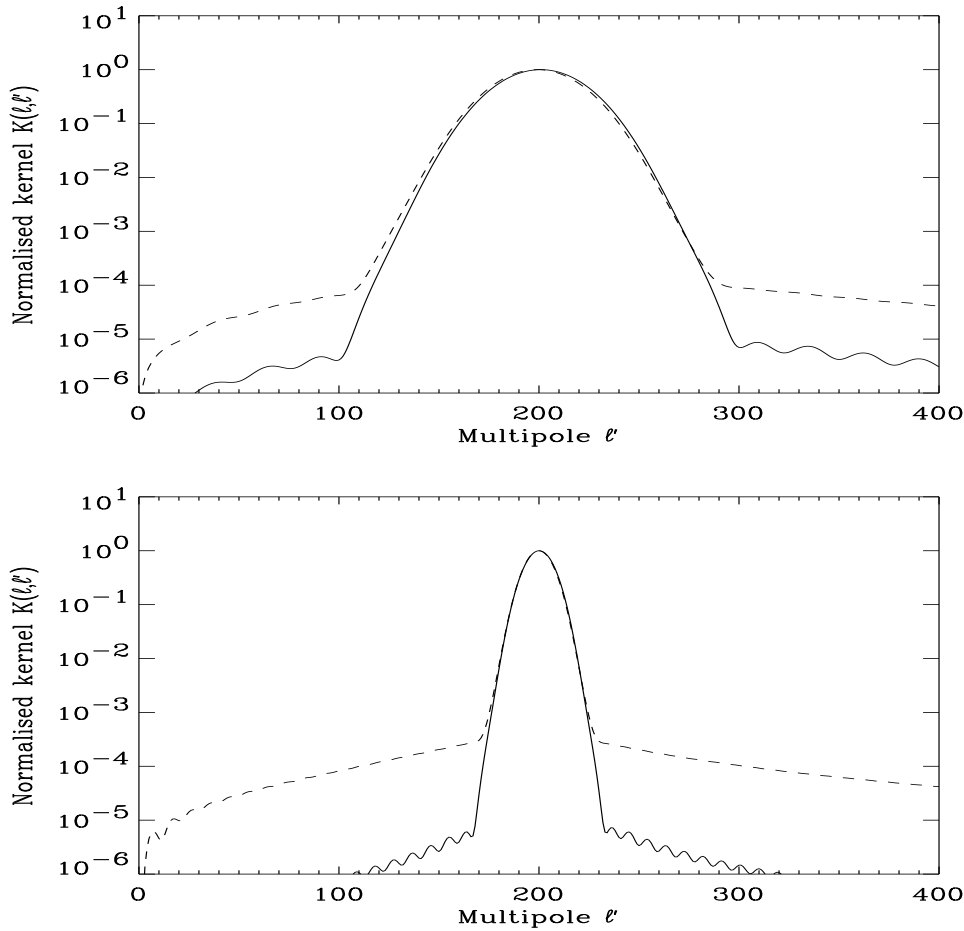


Figure 5.9: A cut at  $\ell = 200$  through the kernel  $K(\ell, \ell')$  connecting the full sky temperature power spectrum with the cut sky temperature power spectrum (solid line) and the kernel  $K_{-2}(\ell, \ell')$  which is mixing the  $E$  and  $B$  mode polarisation power spectra on the cut sky. The upper plot is for a 5 degree FWHM Gaussian Gabor window and the lower plot for a 15 degree Gaussian window. The kernels are here normalised to 1 at the peak at  $\ell = 200$  in order to compare the shapes of the kernels.

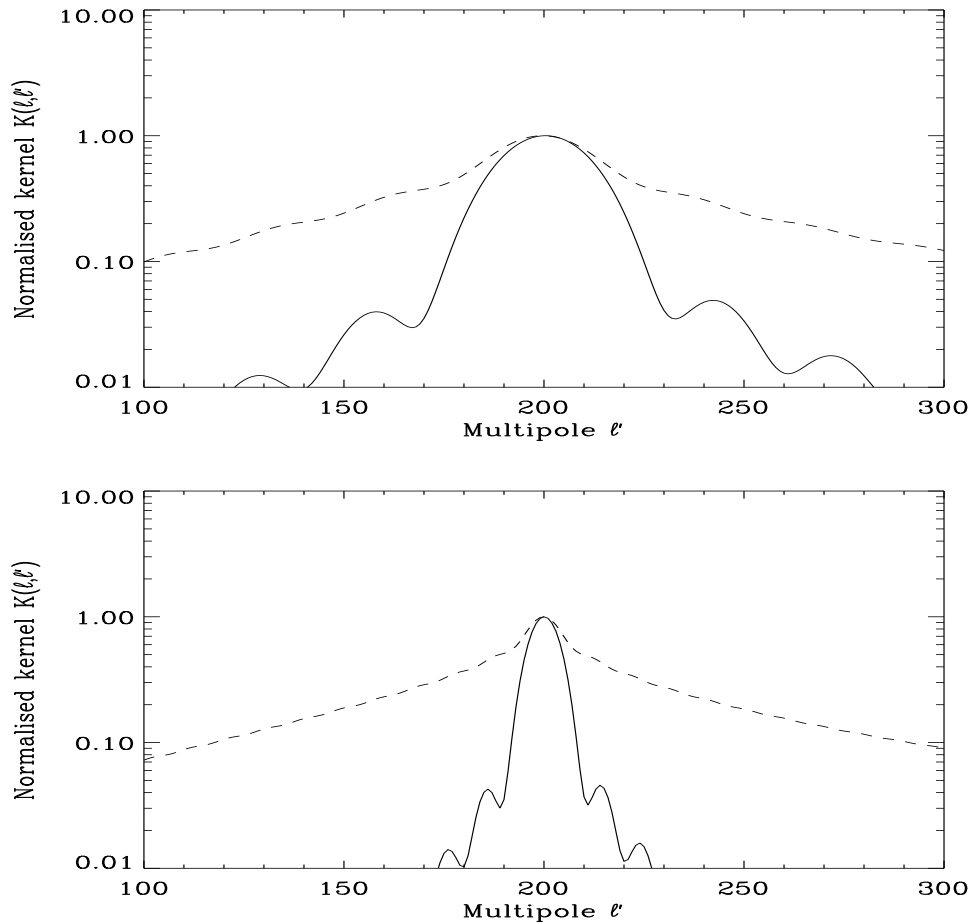


Figure 5.10: This figure is the same as figure (5.9) for tophat windows covering the same area on the sky.

Since the kernels for the polarisation power spectra have a shape similar to that of the temperature power spectrum the effect of a Gabor window on the shape of the power spectrum should also be similar. This can be seen in figure (5.11) and (5.12). The figures show the full sky polarisation power spectra (dashed line)  $C_\ell^E$  (figure (5.11)) and  $C_\ell^C$  (figure (5.12)) for a standard CDM model. In this model the  $B$  component of polarisation is zero. On top of the full sky power spectra I have plotted the polarisation pseudo power spectra for a 5 and 15 degree Gaussian Gabor window (upper and lower plots respectively) normalised so that it can be compared to the full sky spectrum. The pseudo spectra for the corresponding tophat windows are plotted as dotted lines. As expected the shape of the polarisation pseudo spectra relative to the full sky spectra is similar to that for the temperature spectrum (figure (4.6)). One difference is that the polarisation pseudo spectra for the Gaussian window do not have

the characteristic extra peak at low multipole which is seen in the temperature pseudo spectrum. This peak in the temperature spectrum arose due to the step  $1/\ell(\ell+1)$  fall-off of the temperature spectrum at low multipole. The polarisation spectra do not have this steep fall-off and for this reason there is no extra peak.

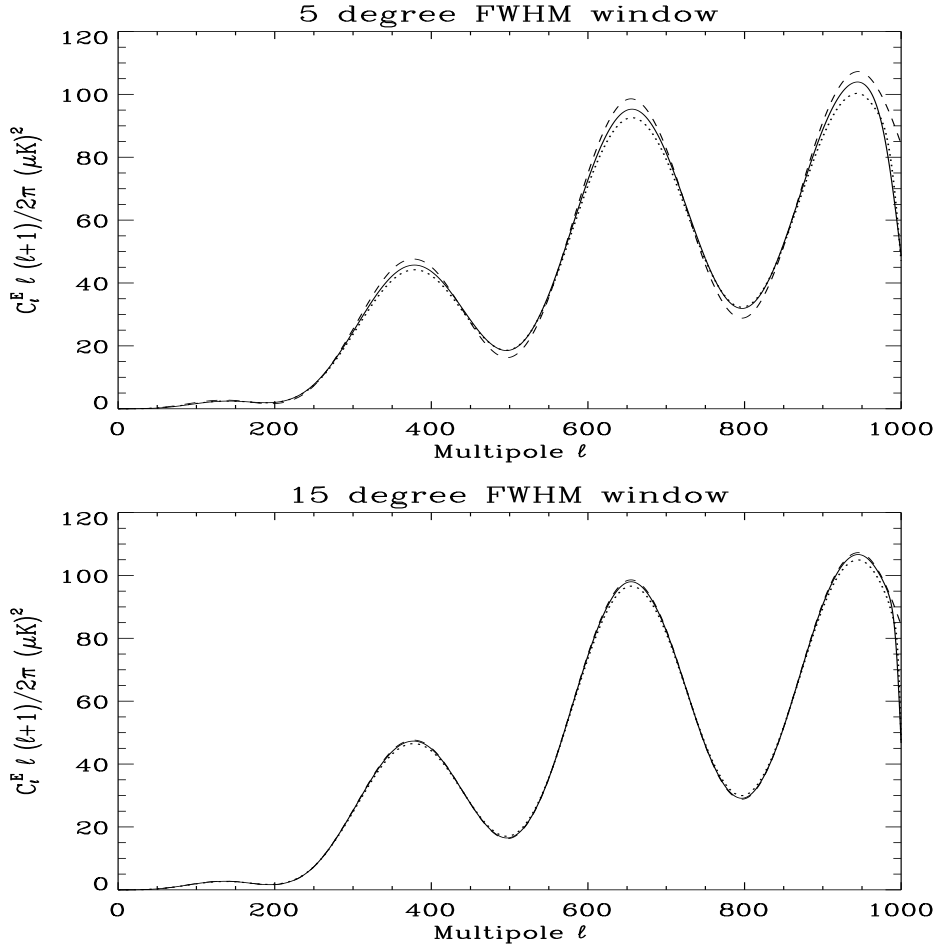


Figure 5.11: The windowed polarisation power spectra  $\tilde{C}_\ell^E$  for a 5 and 15 degree FWHM Gaussian Gabor window cut at  $\theta_C = 3\sigma$  (solid line) and for a tophat window covering the same area on the sky (dotted line). All spectra are normalised in such a way that they can be compared directly with the full sky spectrum which is shown on each plot as a dashed line. Only in the first plot are all three lines visible. In the three last plots, the full sky spectrum and the Gaussian pseudo spectrum (dashed and solid line) are hardly distinguishable.

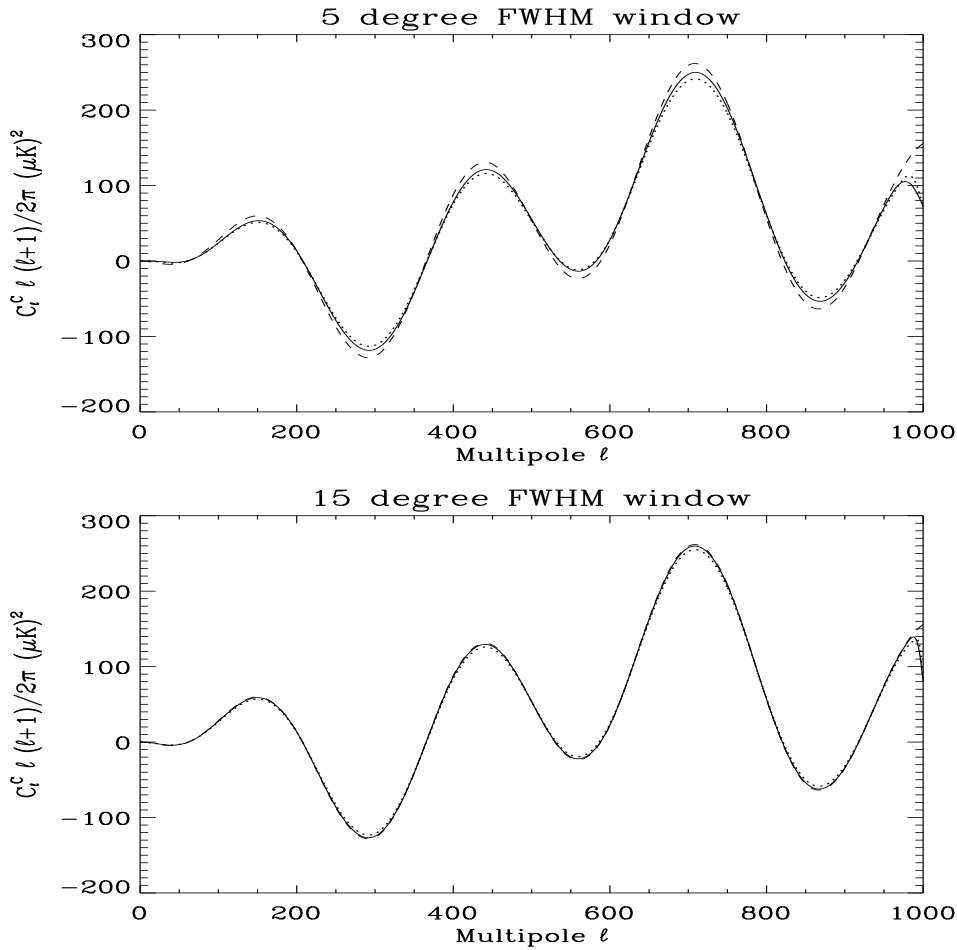


Figure 5.12: Same as figure (5.11) for the temperature-polarisation cross power spectrum  $C_\ell^C$ .

Because of the mixing of  $E$  and  $B$  modes there is also a  $B$  polarisation component  $\tilde{C}_\ell^B$  for the pseudo spectrum even when the input full sky  $C_\ell^B$  were zero. This is shown in figure (5.13) where I have plotted the full sky spectrum  $C_\ell^E$  and the pseudo spectra  $\tilde{C}_\ell^B$  for the 5 and 15 degree FWHM Gaussian gabor windows and corresponding tophat windows. The pseudo spectra are normalised so that they can be compared directly to the full sky spectrum. The dashed lines show the pseudo spectra for the Gaussian window. The upper line is for the 5 degree window and the lower line for the 15 degree window. As expected the size of the  $B$  component is dropping with increasing window size. The  $\tilde{C}_\ell^B$  for the tophat windows are plotted as dotted lines. The shape of the pseudo spectra  $\tilde{C}_\ell^B$  for the Gaussian windows are roughly following the shape of the full sky  $C_\ell^E$ . This could be expected because the mixing kernel  $K_{-2}(\ell, \ell')$  for the Gaussian window has a Gaussian shape close to the diagonal, similar to the other kernels (see fig-



ure (5.9)). The pseudo spectra  $C_\ell^B$  for the tophat windows however are much smoother due to the much broader kernel (figure (5.10)).

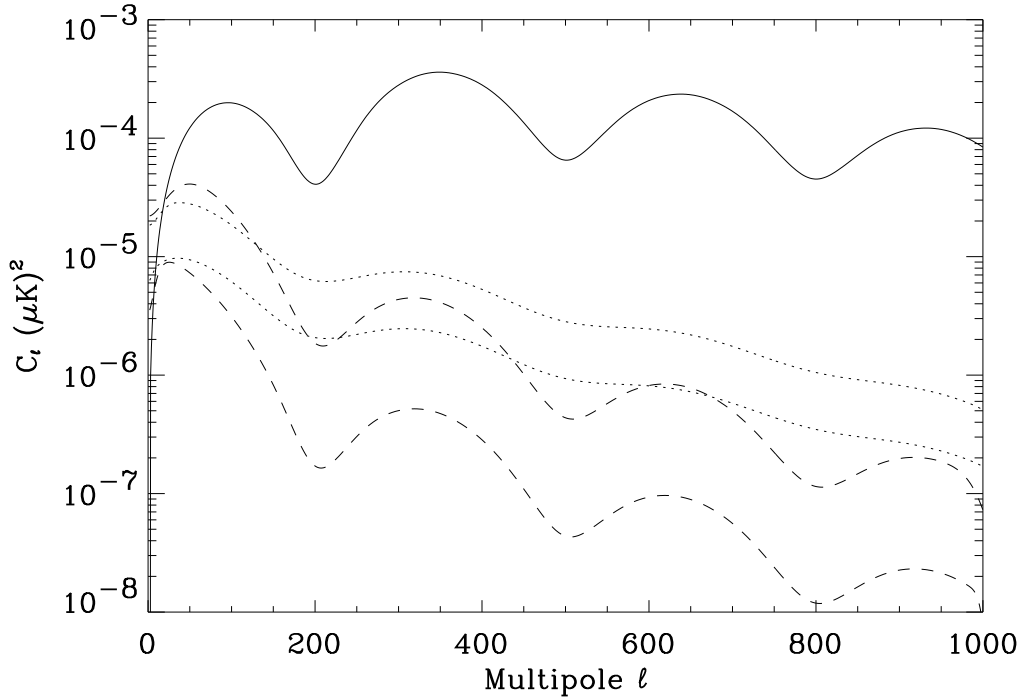


Figure 5.13: The full sky  $C_\ell^E$  power spectrum plotted together with the spectra  $\tilde{C}_\ell^B$  on the windowed sky. The dashed lines show the  $B$  spectra for a 5 and 15 degree FWHM Gaussian Gabor window (upper and lower line respectively). The dotted lines are for the corresponding tophat windows. The pseudo spectra are normalised so that they can be compared directly with the full sky spectrum. In the model used, there was no  $B$  polarisation spectrum for the full sky. The  $\tilde{C}_\ell^B$  shown arise due to the mixing of  $E$  and  $B$  modes on the cut sky only.

In the same way as for the temperature power spectrum, it seems that the polarisation pseudo spectra resemble the full sky polarisation spectra when the patches on the sky are large enough. This motivates the use of the polarisation pseudo power spectra as input to a likelihood estimation of the polarisation power spectra in the same way as for the temperature power spectrum showed in the previous chapter.

### 5.1.2 Rotational Invariance

I now want to show that the pseudo power spectra for polarisation are (as the temperature power spectra) rotationally invariant under a common rotation of the sky and window. First, note that the rotation matrices  $D_{mm'}^\ell$  are rotating both the normal spherical harmonics and the spin-s harmonics. This is easy to show. Assume that one wants to rotate  ${}_s Y_{\ell m}(\hat{\mathbf{n}})$  with the Euler angles  $(\alpha, \beta, \gamma)$ . Using the formula for the normal spherical harmonics one gets,

$${}_s Y_{\ell m}^{\text{rot}}(\hat{\mathbf{n}}) = \sum_{m'} D_{m'm}^\ell(\alpha, \beta, \gamma) {}_s Y_{\ell m'}(\hat{\mathbf{n}}) \quad (5.82)$$

$$= \sqrt{\frac{2\ell+1}{4\pi}} \sum_{m'} D_{m'm}^\ell(\alpha, \beta, \gamma) D_{-sm'}^\ell(\phi, \theta, 0) \quad (5.83)$$

$$= D_{-sm}^\ell(\phi_{\text{rot}}, \theta_{\text{rot}}, 0) \sqrt{\frac{2\ell+1}{4\pi}} \quad (5.84)$$

$$= {}_s Y_{\ell m}(\hat{\mathbf{n}}_{\text{rot}}), \quad (5.85)$$

where  $\hat{\mathbf{n}}_{\text{rot}}$  is the rotation of the angle  $\hat{\mathbf{n}}$  by  $(\alpha, \beta, \gamma)$ . This is clearly general for all spin-s harmonics. Therefore I use the method from chapter (4) to show that polarisation pseudo power spectra are rotationally invariant,

Consider a rotation of the sky and window by the angles  $(-\gamma - \beta - \alpha)$ . Then the  $\tilde{a}_{s,\ell m}$  becomes,

$$\tilde{a}_{s,\ell m}^{\text{rot}} = \int d\hat{\mathbf{n}} [\hat{D}(-\gamma - \beta - \alpha) T(\hat{\mathbf{n}}) G(\hat{\mathbf{n}})] {}_s Y_{\ell m}^*(\hat{\mathbf{n}}). \quad (5.86)$$

If one makes the inverse rotation of the integration angle  $\hat{\mathbf{n}}$ , one can write this as;

$$\tilde{a}_{s,\ell m}^{\text{rot}} = \int d\hat{\mathbf{n}} T(\hat{\mathbf{n}}) G(\hat{\mathbf{n}}) [\hat{D}^*(\alpha\beta\gamma) {}_s Y_{\ell m}^*(\hat{\mathbf{n}})], \quad (5.87)$$

which is just

$$\tilde{a}_{s,\ell m}^{\text{rot}} = \sum_{m'} D_{m'm}^{\ell*}(\alpha\beta\gamma) \int T(\hat{\mathbf{n}}) G(\hat{\mathbf{n}}) {}_s Y_{\ell m'}^*(\hat{\mathbf{n}}). \quad (5.88)$$

The last integral can be identified as the normal  $\tilde{a}_{s,\ell m}$ .

$$\tilde{a}_{s,\ell m}^{\text{rot}} = \sum_{m'} D_{m'm}^{\ell*}(\alpha\beta\gamma) \tilde{a}_{s,\ell m}. \quad (5.89)$$

Thus,

$$\tilde{a}_{E,\ell m}^{\text{rot}} = \sum_{m'} D_{m'm}^{\ell*}(\alpha\beta\gamma) \tilde{a}_{E,\ell m} \quad \tilde{a}_{B,\ell m}^{\text{rot}} = \sum_{m'} D_{m'm}^{\ell*}(\alpha\beta\gamma) \tilde{a}_{B,\ell m}. \quad (5.90)$$

For  $\tilde{C}_{E,\ell m}$  (and analogously for  $\tilde{C}_{B,\ell m}$  and  $\tilde{C}_{C,\ell m}$ ) one gets

$$\tilde{C}_{E,\ell}^{\text{rot}} = \frac{1}{2\ell+1} \sum_m a_{E,\ell m}^{\text{rot}} a_{E,\ell m}^{\text{rot}*} \quad (5.91)$$

$$= \frac{1}{2\ell+1} \sum_m \sum_{m'} \sum_{m''} D_{m'm}^\ell(\alpha\beta\gamma) D_{m''m}^{\ell*}(\alpha\beta\gamma) \tilde{a}_{E,\ell m'} \tilde{a}_{E,\ell m''}^* \quad (5.92)$$

$$= \frac{1}{2\ell+1} \sum_{m'm''} \tilde{a}_{E,\ell m'} \tilde{a}_{E,\ell m''}^* \underbrace{\sum_m D_{m'm}^\ell(\alpha\beta\gamma) D_{m''m}^{\ell*}(\alpha\beta\gamma)}_{\delta_{mm'}} \quad (5.93)$$

$$= \tilde{C}_{E,\ell}. \quad (5.94)$$

Since the polarisation power spectra are rotationally invariant, I will in the rest of the chapter put the centre of the window on the north pole. This makes the calculations easier while keeping the generality of the results. Now, to do the likelihood analysis one needs to find theoretical expressions for the correlations between different  $\tilde{C}_\ell^Z$  ( $Z=\{T,E,C\}$ ).

## 5.2 Likelihood Analysis

For the temperature power spectrum a Gaussian likelihood ansatz was used (chapter (4)). Because of the similarities between the kernels of the polarisation power spectra and the temperature power spectrum this is also to be expected for the polarisation spectra. I will now show the results of some Monte Carlo simulations confirming this assumption. In this chapter I will assume that the  $B$  component of polarisation is so small that it can be neglected. I will only concentrate on the  $T$ ,  $E$  and  $C$  components of polarisation.

### 5.2.1 The Form of the Likelihood Function for Polarisation

In figure (5.14) and (5.15) I have plotted the probability distribution of the  $\tilde{C}_\ell^E$  and  $\tilde{C}_\ell^C$  from 10000 simulations. The probability distribution (histogram) is plotted on top of a Gaussian (dashed line) with mean and FWHM taken from the theoretical expressions to be derived in the next section. In these simulations I was using a 15 degree FWHM Gaussian Gabor window with  $\theta_C = 3\sigma$ . In figure (5.16) and (5.17) I show the results of similar simulations with a 15 degree FWHM Gaussian Gabor window. As expected the trend is that the distributions get more and more Gaussian for higher multipoles and for bigger windows. For the 15° FWHM window, the distribution is very close to a Gaussian for the multipoles above  $\ell = 50$ .

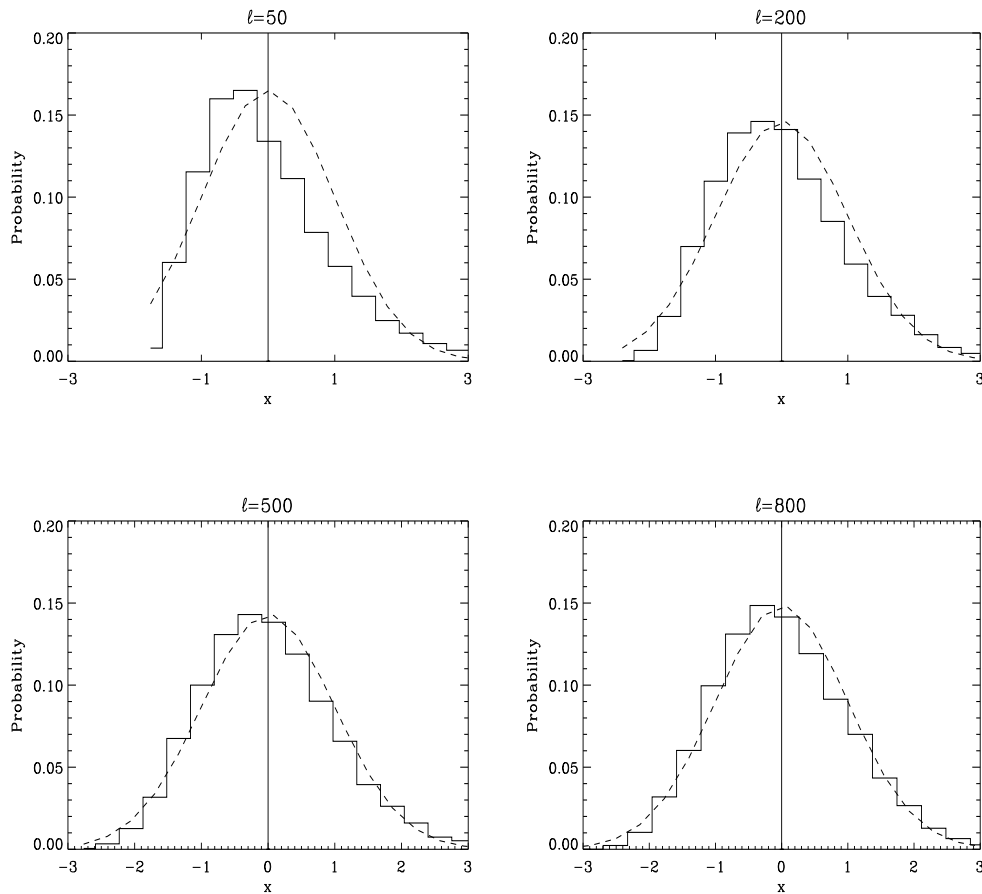


Figure 5.14: The probability distribution of  $\tilde{C}_\ell^E$  taken from 10000 simulations with a  $5^\circ$  FWHM Gaussian Gabor window truncated at  $\theta_C = 3\sigma$ . The variable  $x$  is given as  $x = (\tilde{C}_\ell^E - \langle \tilde{C}_\ell^E \rangle) / \sqrt{\langle (\tilde{C}_\ell^E - \langle \tilde{C}_\ell^E \rangle)^2 \rangle}$ . The dashed line is a Gaussian with the theoretical mean and standard deviation of the  $\tilde{C}_\ell^E$ . The plot shows the  $\tilde{C}_\ell^E$  distribution for  $\ell = 50$ ,  $\ell = 200$ ,  $\ell = 500$ , and  $\ell = 800$ . The probabilities are normalised such that the integral over  $x$  is 1.

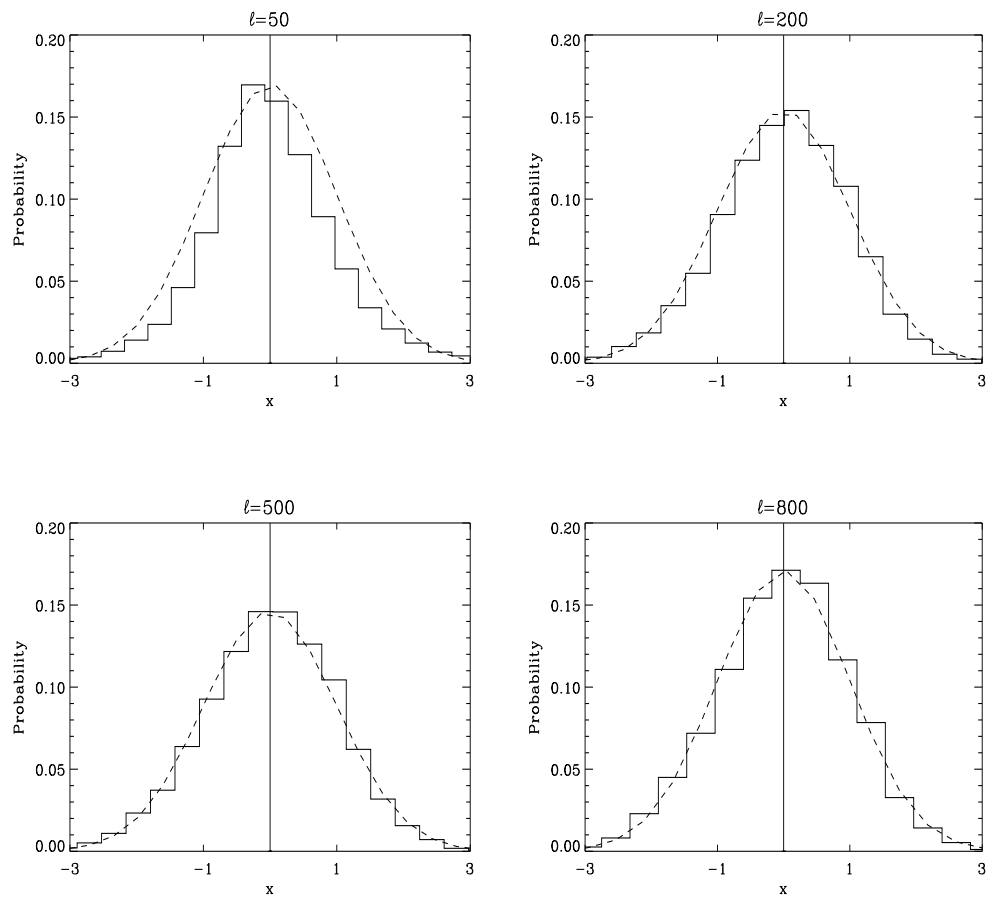


Figure 5.15: Same as figure (5.14) for the temperature-polarisation cross spectra  $\tilde{C}_\ell^C$ .

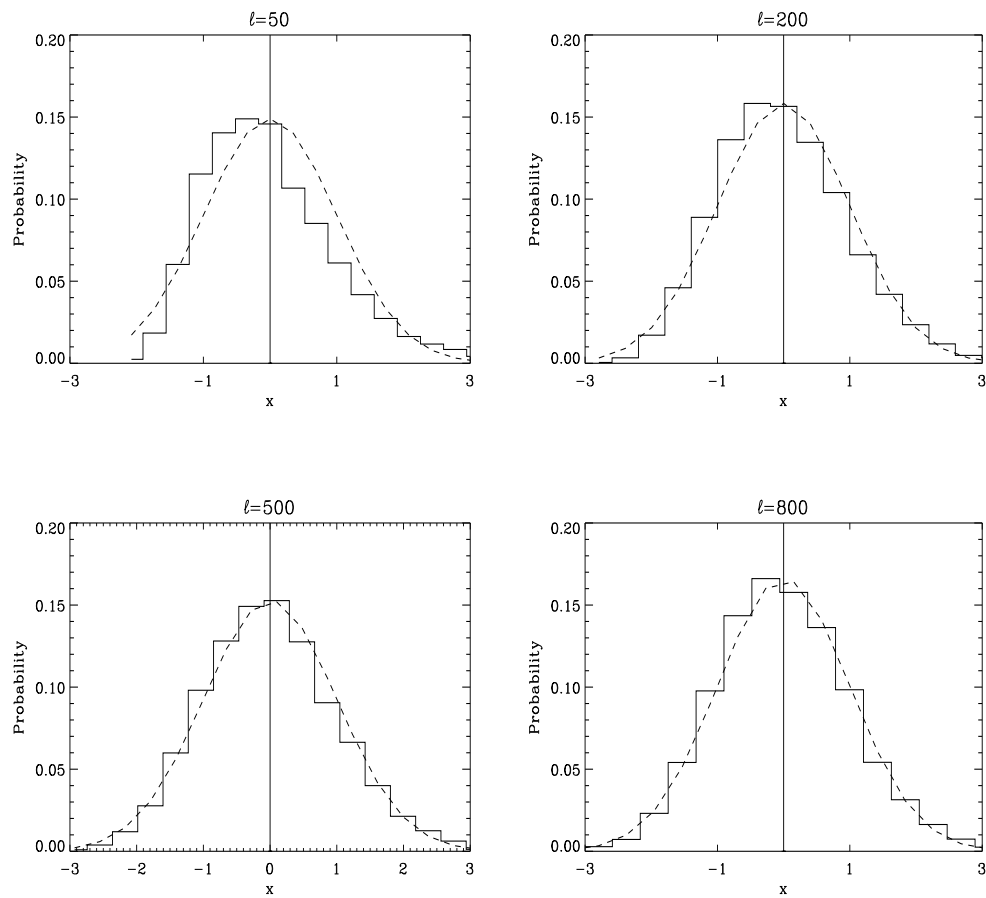


Figure 5.16: Same as figure (5.14) for a  $15^\circ$  FWHM Gaussian Gabor window.

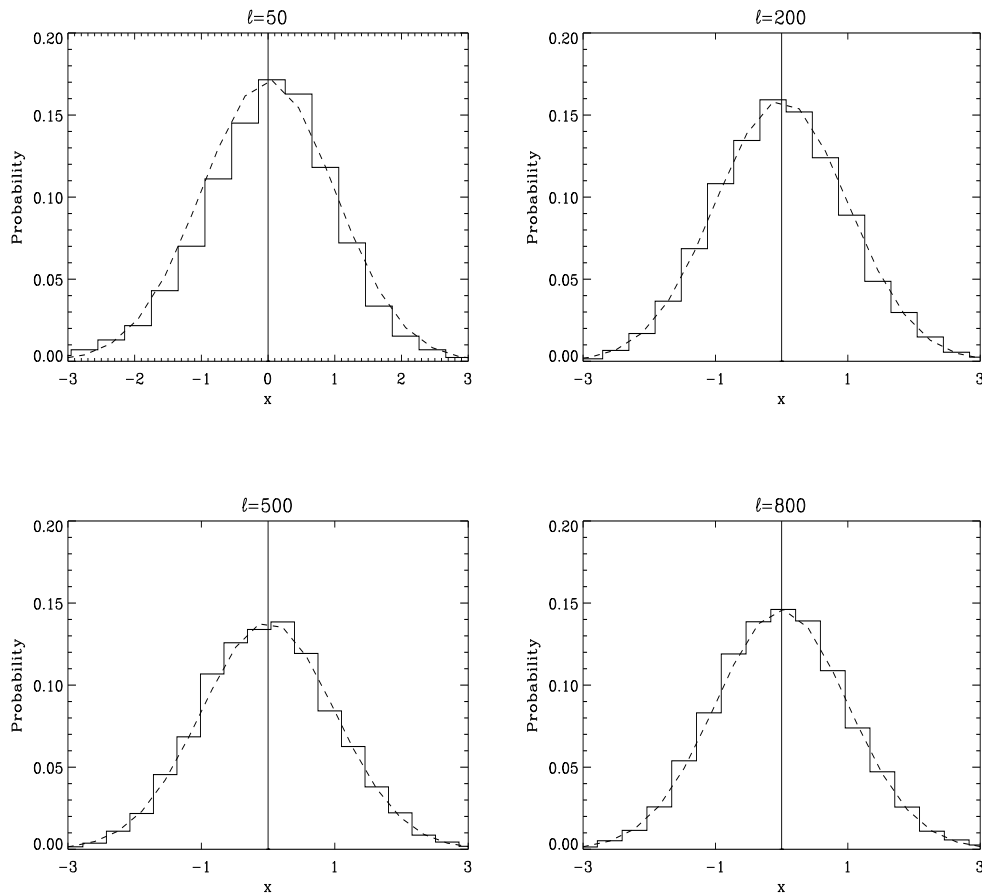


Figure 5.17: Same as figure (5.15) for a  $15^\circ$  FWHM Gaussian Gabor window.

Figure (5.18) and (5.19) show the probability distribution for a tophat window covering the same area on the sky as the Gaussian window used in figure (5.16) and (5.17). Also this distribution is very close to a Gaussian.

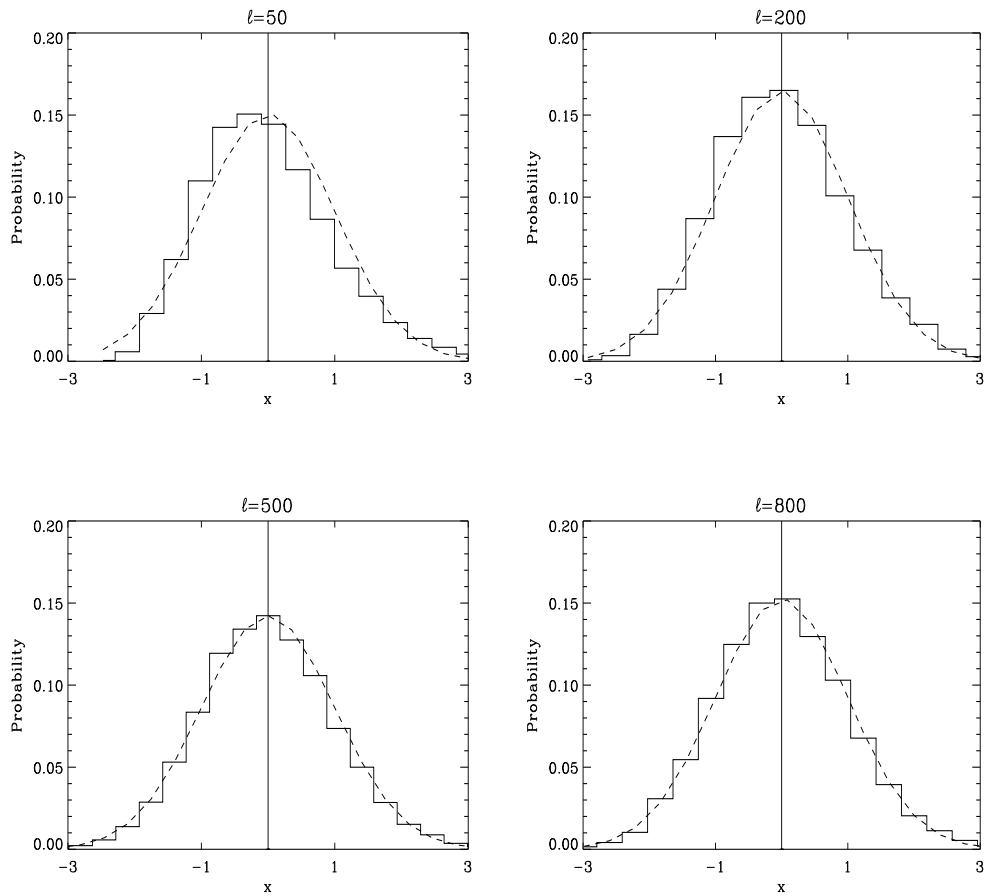


Figure 5.18: Same as figure (5.16) for a tophat window covering the same area on the sky



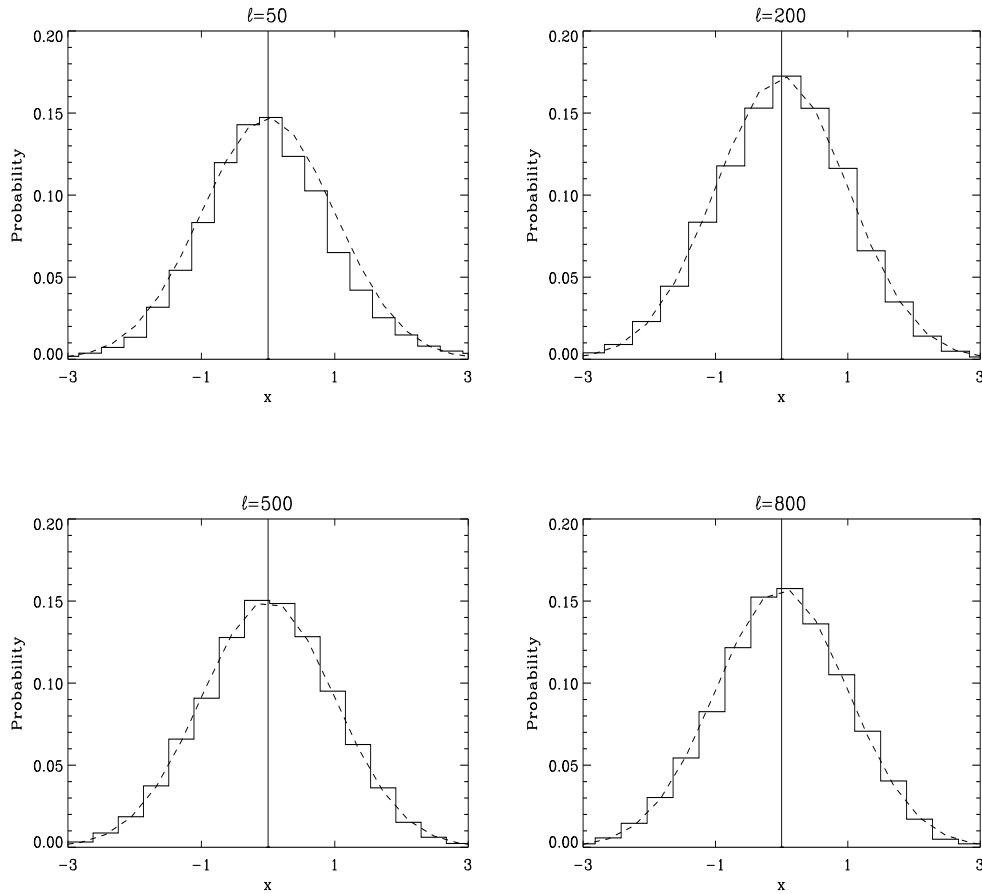


Figure 5.19: Same as figure (5.17) for a tophat window covering the same area on the sky

The previous plots have shown that a Gaussian likelihood ansatz for the polarisation pseudo spectra seems to be a very good approximation provided that the window is big enough. As for the temperature spectrum, the approximation is no longer valid for the lowest multipoles, but as was shown for the temperature power spectrum, this might only give rise to a very small downward bias for the estimates of the lowest multipoles.

The form of the log-likelihood to minimize is therefore still

$$L = \mathbf{d}^T \mathbf{M}^{-1} \mathbf{d} + \ln \det \mathbf{M}, \quad (5.95)$$

where the datavector now consists of the temperature and polarisation power spectra  $\mathbf{d} = \{\mathbf{d}^T, \mathbf{d}^E, \mathbf{d}^C\}$ . Here the  $\mathbf{d}^Z$  vectors are given as

$$d_i^Z = \tilde{C}_{\ell_i}^Z - \langle \tilde{C}_{\ell_i}^Z \rangle, \quad (5.96)$$

where  $Z = \{T, E, B\}$ . Similarly the correlation matrix  $\mathbf{M}$  will consist of blocks  $\mathbf{M}_{ZZ'}$  defined as

$$M_{ZZ',ij} = \langle \tilde{C}_{\ell_i}^Z \tilde{C}_{\ell_j}^{Z'} \rangle - \langle \tilde{C}_{\ell_i}^Z \rangle \langle \tilde{C}_{\ell_j}^{Z'} \rangle. \quad (5.97)$$

This structure of the datavector and correlation matrix is shown in figure (5.20).

T	$\langle \text{TT} \rangle$	$\langle \text{TE} \rangle$	$\langle \text{TC} \rangle$
E	$\langle \text{TE} \rangle$	$\langle \text{EE} \rangle$	$\langle \text{EC} \rangle$
C	$\langle \text{TC} \rangle$	$\langle \text{EC} \rangle$	$\langle \text{CC} \rangle$

Figure 5.20: The figure shows the structure of the datavector  $\mathbf{d}$  on the left hand side and the correlation matrix  $\mathbf{M}$  on the right hand side used for joint likelihood estimation of temperature and polarisation power spectra.

For fast likelihood estimation, it is crucial that one can calculate the average pseudo spectra  $\langle \tilde{C}_{\ell}^Z \rangle$  and correlation matrix  $\mathbf{M}$  fast. The formalism in the previous chapter which enabled fast calculations of these quantities for the temperature power spectrum will now be extended to polarisation.

### 5.2.2 The Polarisation Correlation Matrix

To find the correlation matrix  $\mathbf{M}$  for likelihood estimation of the polarisation power spectra one needs the formulae given in equations (5.41) to (5.48). As shown there the correlations of the pseudo  $\tilde{a}_{\ell m}$  coefficients can be written in terms

of the  $h(\ell, \ell', m)$  function from chapter (4) and the  $H_2(\ell, \ell', m)$  and  $H_{-2}(\ell, \ell', m)$  functions. These function can be quickly calculated using the important recursion formulae deduced in appendix (D) and (E). The starting points of these recursions can also be quickly provided using summations and FFT as explained in chapter (4). I will now show that the correlation function  $\mathbf{M}$  can be expressed in terms of these functions and for this reason can be calculated quickly.

The pseudo power spectra can be written as

$$\langle \tilde{C}_\ell^T \rangle = \sum_m \frac{\langle \tilde{a}_{\ell m}^T \tilde{a}_{\ell m}^{T*} \rangle}{2\ell + 1}, \quad (5.98)$$

$$\langle \tilde{C}_\ell^E \rangle = \sum_m \frac{\langle \tilde{a}_{\ell m}^E \tilde{a}_{\ell m}^{E*} \rangle}{2\ell + 1}, \quad (5.99)$$

$$\langle \tilde{C}_\ell^C \rangle = \sum_m \frac{\langle \tilde{a}_{\ell m}^T \tilde{a}_{\ell m}^{E*} \rangle}{2\ell + 1}. \quad (5.100)$$

To find the correlation function between  $\tilde{C}_\ell$  for polarisation one can follow the same steps as for the temperature correlation functions, and get,

$$M_{EE, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m \langle \tilde{a}_{E, \ell m} \tilde{a}_{E, \ell' m}^* \rangle^2, \quad (5.101)$$

$$M_{BB, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m \langle \tilde{a}_{B, \ell m} \tilde{a}_{B, \ell' m}^* \rangle^2, \quad (5.102)$$

$$M_{CC, \ell \ell'} = \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_m \left[ \langle \tilde{a}_{E, \ell m} \tilde{a}_{E, \ell' m}^* \rangle \langle \tilde{a}_{\ell m} \tilde{a}_{\ell' m}^* \rangle \right. \quad (5.103)$$

$$\left. + \langle \tilde{a}_{E, \ell m} \tilde{a}_{\ell' m}^* \rangle \langle \tilde{a}_{E, \ell' m} \tilde{a}_{\ell m}^* \rangle \right], \quad (5.104)$$

$$M_{ET, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m \langle \tilde{a}_{E, \ell m} \tilde{a}_{\ell' m}^* \rangle^2, \quad (5.105)$$

$$M_{BT, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m | \langle \tilde{a}_{B, \ell m} \tilde{a}_{\ell' m}^* \rangle |^2, \quad (5.106)$$

$$M_{CT, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m \langle \tilde{a}_{E, \ell m} \tilde{a}_{\ell' m}^* \rangle \langle \tilde{a}_{\ell m} \tilde{a}_{\ell' m}^* \rangle, \quad (5.107)$$

$$M_{EB, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m | \langle \tilde{a}_{E, \ell m} \tilde{a}_{B, \ell' m}^* \rangle |^2, \quad (5.108)$$

$$M_{EC, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m \langle \tilde{a}_{E, \ell m} \tilde{a}_{E, \ell' m}^* \rangle \langle \tilde{a}_{E, \ell m} \tilde{a}_{\ell' m}^* \rangle, \quad (5.109)$$

$$M_{BC, \ell \ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_m \langle \tilde{a}_{E, \ell' m} \tilde{a}_{B, \ell m}^* \rangle \langle \tilde{a}_{\ell' m}^* \tilde{a}_{B, \ell m} \rangle, \quad (5.110)$$

$$(5.111)$$

where the correlation between  $\tilde{a}_{\ell m}$  are given in equations (5.41) to (5.48) as sums of  $h(\ell, \ell', m)$  and  $H_{\pm 2}(\ell, \ell', m)$ .

In figure (5.21) I have plotted the correlation matrix  $M_{\ell\ell'}^{TT}$  next to the matrix  $M_{\ell\ell'}^{EE}$ . A standard CDM power spectrum without  $B$  mode polarisation was used. The two matrices are very similar. One big difference is that the matrix for  $E$  mode polarisation is missing the 'wall' at low multipoles present in the temperature matrix. As discussed before this is because of the different shapes for the  $T$  and  $E$  power spectra at low multipoles. The temperature power spectrum drops steeply at low  $\ell$  while this is not the case for the  $E$  mode polarisation spectrum.

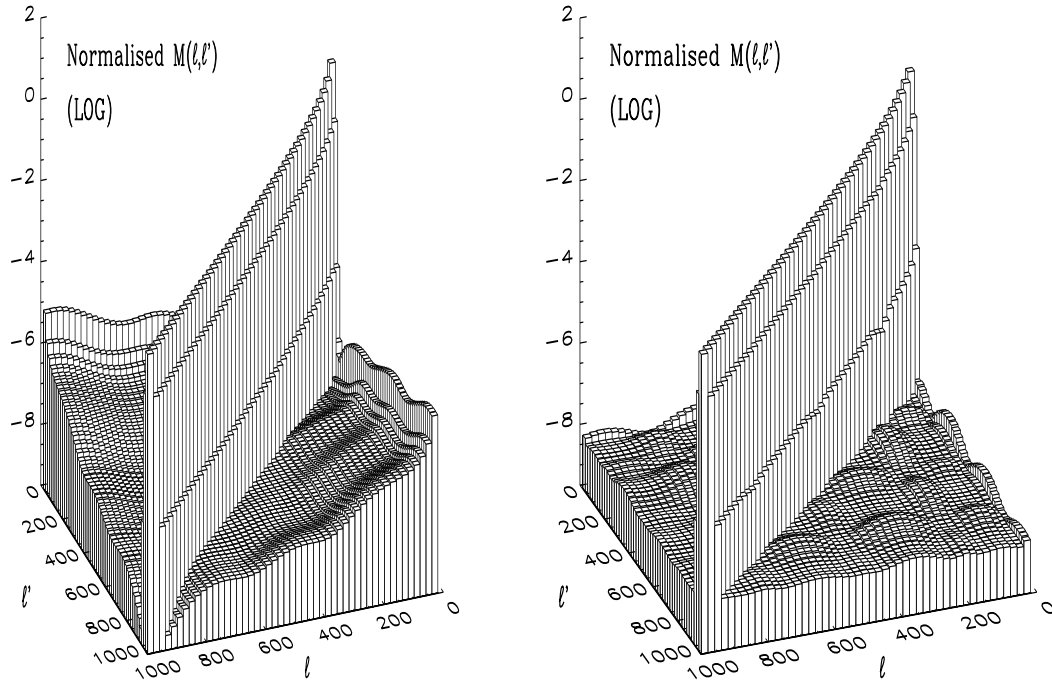


Figure 5.21: The correlation matrices  $M(\ell, \ell')$  in the figure show the correlations between the temperature pseudo power spectrum coefficients and between the  $E$  mode polarisation pseudo spectrum coefficients for a 15 degree FWHM Gaussian Gabor window. The left plot shows  $(\langle \tilde{C}_\ell^T \tilde{C}_{\ell'}^T \rangle - \langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^T \rangle) / (\langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^T \rangle)$  and the right plot shows  $(\langle \tilde{C}_\ell^E \tilde{C}_{\ell'}^E \rangle - \langle \tilde{C}_\ell^E \rangle \langle \tilde{C}_{\ell'}^E \rangle) / (\langle \tilde{C}_\ell^E \rangle \langle \tilde{C}_{\ell'}^E \rangle)$ . A standard CDM power spectrum was used to produce the plots.

In figures (5.22) and (5.23) the  $M_{\ell\ell'}^{CC}$ ,  $M_{\ell\ell'}^{TC}$ ,  $M_{\ell\ell'}^{TE}$  and  $M_{\ell\ell'}^{EC}$  matrices are shown. All matrices are diagonally dominant and the values on the diagonals are not differing significantly between the matrices. For this reason all the matrices have to be computed and used in likelihood analysis.

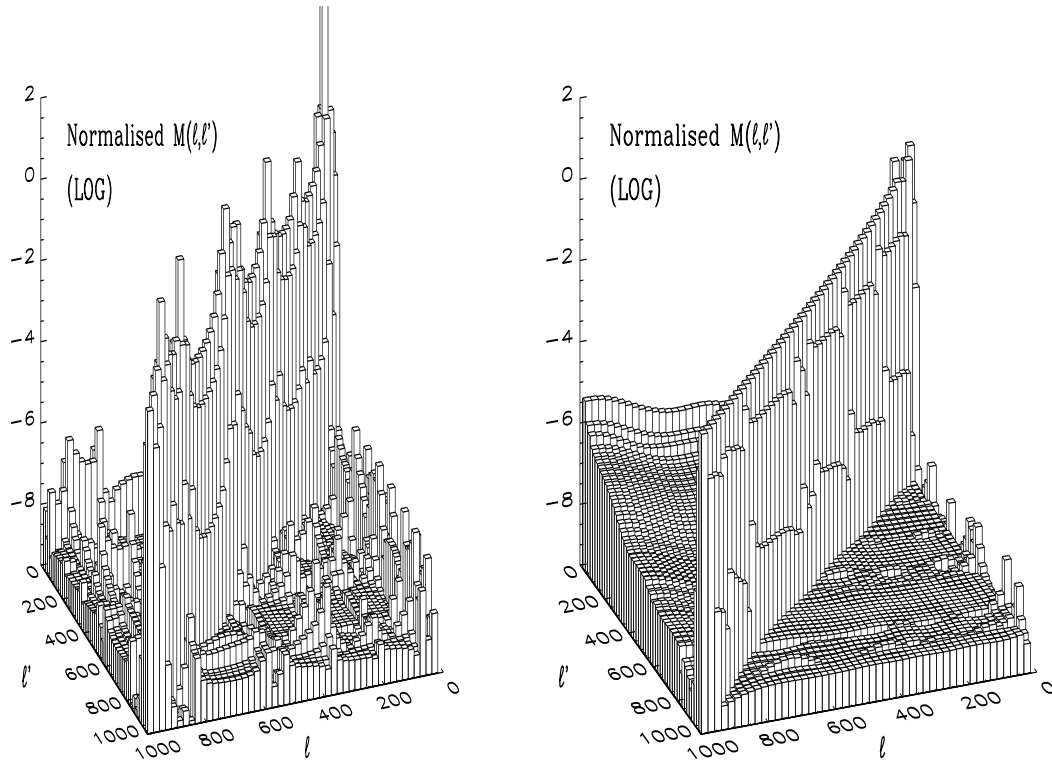


Figure 5.22: The correlation matrices  $M(\ell, \ell')$  in the figure show the correlations between the cross-correlation pseudo power spectrum  $C$  coefficients and between the temperature and cross correlation pseudo spectrum coefficients  $C$  for a 15 degree FWHM Gaussian Gabor window. The left plot shows  $(\langle \tilde{C}_\ell^C \tilde{C}_{\ell'}^C \rangle - \langle \tilde{C}_\ell^C \rangle \langle \tilde{C}_{\ell'}^C \rangle) / (\langle \tilde{C}_\ell^C \rangle \langle \tilde{C}_{\ell'}^C \rangle)$  and the right plot shows  $(\langle \tilde{C}_\ell^T \tilde{C}_{\ell'}^C \rangle - \langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^C \rangle) / (\langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^C \rangle)$ . A standard CDM power spectrum was used to produce the plots.

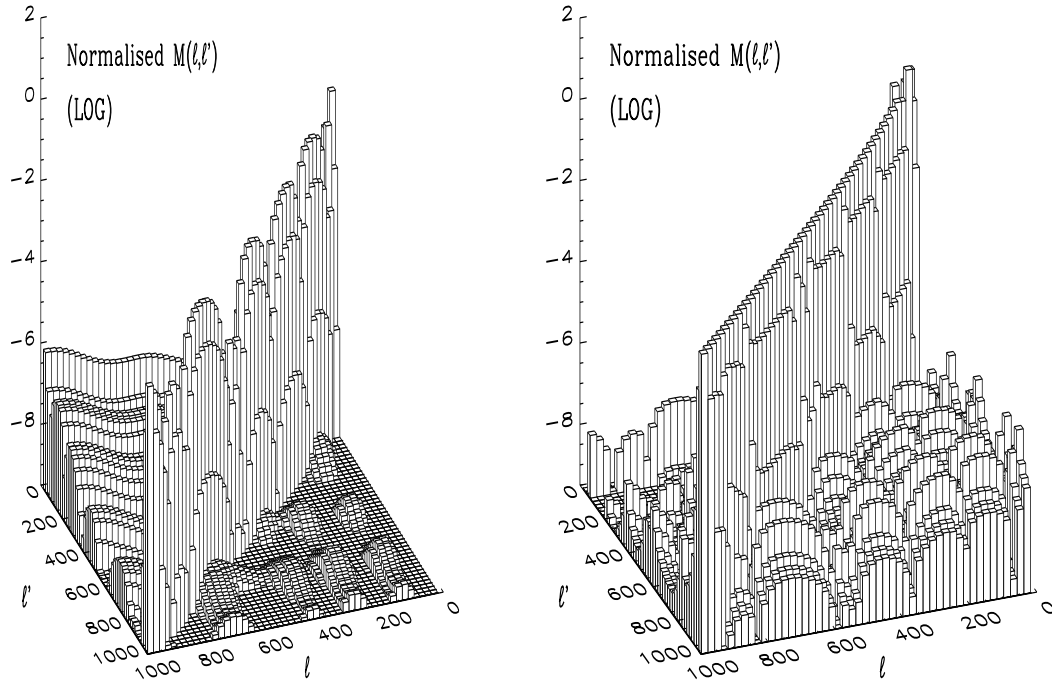


Figure 5.23: The correlation matrices  $M(\ell, \ell')$  in the figure show the correlations between the temperature and  $E$  mode polarisation pseudo spectrum coefficients and between the  $E$  mode polarisation and cross correlation pseudo spectrum coefficients  $C$  for a 15 degree FWHM Gaussian Gabor window. The left plot shows  $(\langle \tilde{C}_\ell^T \tilde{C}_{\ell'}^E \rangle - \langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^E \rangle) / (\langle \tilde{C}_\ell^T \rangle \langle \tilde{C}_{\ell'}^E \rangle)$  and the right plot shows  $(\langle \tilde{C}_\ell^E \tilde{C}_{\ell'}^C \rangle - \langle \tilde{C}_\ell^E \rangle \langle \tilde{C}_{\ell'}^C \rangle) / (\langle \tilde{C}_\ell^E \rangle \langle \tilde{C}_{\ell'}^C \rangle)$ . A standard CDM power spectrum was used to produce the plots.

### 5.2.3 Polarisation with Noise

Analogously to chapter (4) I will now discuss the noise pseudo power spectra and the noise correlation matrix for polarisation. Each pixel in the temperature map has a noise temperature  $n_j$  and for the polarisation maps I assume  $n_j^Q$  and  $n_j^U$  to have the following properties,

$$\langle n_j \rangle = 0, \quad \langle n_j n_{j'} \rangle = \delta_{jj'} (\sigma_j^T)^2, \quad (5.112)$$

$$\langle n_j^Q \rangle = 0, \quad \langle n_j^Q n_{j'}^Q \rangle = \delta_{jj'} (\sigma_j^P)^2, \quad (5.113)$$

$$\langle n_i^U \rangle = 0, \quad \langle n_i^U n_j^U \rangle = \delta_{ij} (\sigma_j^P)^2, \quad (5.114)$$

I also assume that there is no correlation between noise in the different maps T,Q and U. For the full sky one has,

$$\langle a_{\ell m}^N a_{\ell m}^{N*} \rangle = \sum_j (\sigma_j^T)^2 |Y_{\ell m}^j|^2 \quad (5.115)$$

$$\langle a_{2,\ell m}^N a_{2,\ell m}^{N*} \rangle = 2 \sum_j (\sigma_j^P)^2 |{}_2Y_{\ell m}^j|^2, \quad (5.116)$$

$$\langle a_{-2,\ell m}^N a_{-2,\ell m}^{N*} \rangle = 2 \sum_j (\sigma_j^P)^2 |{}_{-2}Y_{\ell m}^j|^2, \quad (5.117)$$

$$\langle a_{2,\ell m}^N a_{-2,\ell m}^{N*} \rangle = 0, \quad (5.118)$$

$$\langle a_{E,\ell m}^N a_{E,\ell m}^{N*} \rangle = 2 \sum_j (\sigma_j^P)^2 (|{}_2Y_{\ell m}^j|^2 + |{}_{-2}Y_{\ell m}^j|^2), \quad (5.119)$$

$$\langle a_{B,\ell m}^N a_{B,\ell m}^{N*} \rangle = \langle a_{E,\ell m}^N a_{E,\ell m}^{N*} \rangle, \quad (5.120)$$

$$\langle a_{E,\ell m}^N a_{B,\ell m}^{N*} \rangle = 0, \quad (5.121)$$

which for this type of noise gives  $C_\ell^{EN} = C_\ell^{BN}$ .

The pseudo  $a_{2,\ell m}$  coefficients can now be found using equations (5.1) and (5.2) I define

$$\tilde{a}_{\pm 2,\ell m}^N = \sum_j (n_j^Q \pm i n_j^U) G_j \pm 2Y_{\ell m}^j, \quad (5.122)$$

for an axissymmetric Gabor window  $G$  having the value  $G_j$  in pixel  $j$ . The  $E$  and  $B$  components are then similarly

$$\tilde{a}_{E,\ell m}^N = -\frac{1}{2} \sum_j \left[ n_j^Q ({}_2Y_{\ell m}^j + {}_{-2}Y_{\ell m}^j) + i n_j^U ({}_2Y_{\ell m}^j - {}_{-2}Y_{\ell m}^j) \right] G_j, \quad (5.123)$$

$$\tilde{a}_{B,\ell m}^N = \frac{1}{2} i \sum_j \left[ n_j^Q ({}_2Y_{\ell m}^j - {}_{-2}Y_{\ell m}^j) + i n_j^U ({}_2Y_{\ell m}^j + {}_{-2}Y_{\ell m}^j) \right] G_j. \quad (5.124)$$

$$(5.125)$$

The correlations between these coefficients are

$$\langle \tilde{a}_{E,\ell m}^N \tilde{a}_{E,\ell' m'}^{N*} \rangle = \frac{1}{4} \sum_j \left[ (\sigma_j^P)^2 ({}_2Y_{\ell m}^j + {}_{-2}Y_{\ell m}^j) \right] \quad (5.126)$$

$$\times ({}_2Y_{\ell' m'}^j + {}_{-2}Y_{\ell' m'}^j) \quad (5.127)$$

$$+ (\sigma_j^P)^2 ({}_2Y_{\ell m}^j - {}_{-2}Y_{\ell m}^j) \quad (5.128)$$

$$\times ({}_2Y_{\ell' m'}^j - {}_{-2}Y_{\ell' m'}^j) \Big] G_j^2 \quad (5.129)$$

$$= \frac{1}{2} [h'_2(\ell, \ell', m, m')] \quad (5.130)$$

$$+ (-1)^{m+m'} h'_2(\ell, \ell', -m, -m')] \quad (5.131)$$

$$\equiv H'_2(\ell, \ell', m, m'),$$

where the last line defines  $H'_2(\ell, \ell', m, m')$ . The  $h'_2(\ell, \ell', m, m')$  is defined similar to the  $h'(\ell, \ell', m, m')$  function in chapter (4)

$$h'_2(\ell, \ell', m, m') = \sum_j G_j^2(\sigma_j^P)^2 {}_2Y_{\ell m}^j {}_2Y_{\ell' m'}^j, \quad (5.132)$$

Note the following relation which was used to obtain equation (5.130)

$$\sum_j G_j^2(\sigma_j^P)^2 {}_2Y_{\ell m}^j {}_2Y_{\ell' m'}^j = (-1)^{m+m'} h_2(\ell, \ell', -m, -m'). \quad (5.133)$$

Again, one can see that when the Gabor window AND noise have azimuthal symmetry this reduces simply to,

$$h'_2(\ell, \ell', m, m') = h'_2(\ell, \ell', m). \quad (5.134)$$

In a similar manner the other  $a_{\ell m}$  relations can be found

$$\langle \tilde{a}_{B, \ell m}^N \tilde{a}_{B, \ell' m'}^{N*} \rangle = \langle \tilde{a}_{E, \ell m}^N \tilde{a}_{E, \ell' m'}^{N*} \rangle \quad (5.135)$$

$$\langle \tilde{a}_{E, \ell m}^N \tilde{a}_{B, \ell' m'}^{N*} \rangle = \frac{1}{2} i [h'_2(\ell, \ell', m, m') - \quad (5.136)$$

$$\begin{aligned} & (-1)^{m+m'} h'_2(\ell, \ell', -m, -m')] \\ & \equiv i H'_{-2}(\ell, \ell', m, m'), \end{aligned} \quad (5.137)$$

where the last line again defines  $H'_{-2}(\ell, \ell', m, m')$ . The  $H_{\pm 2}(\ell, \ell', m, m')$  functions which are needed to find the noise correlation matrices can be quickly calculated using the recursion in appendix (E).

Using these relations one can now find the polarisation pseudo spectra

$$\langle \tilde{C}_\ell^{EN} \rangle = \langle \tilde{C}_\ell^{BN} \rangle = \frac{1}{2\ell + 1} \sum_m \langle \tilde{a}_{E, \ell m}^N \tilde{a}_{E, \ell m}^{N*} \rangle, \quad (5.138)$$

$$\langle \tilde{C}_\ell^{CN} \rangle = 0 \quad (5.139)$$

One can further use this to find the noise correlation matrices  $M_{ZZ', \ell \ell'}^N$ , defined as,

$$M_{ZZ', \ell \ell'}^N = \langle C_\ell^{ZN} C_{\ell'}^{Z'N} \rangle - \langle C_\ell^{ZN} \rangle \langle C_{\ell'}^{Z'N} \rangle, \quad (5.140)$$

where  $Z = \{T, E, B, C\}$ . I find,

$$M_{TT, \ell \ell'}^N = \frac{2}{2\ell + 1} \sum_{mm'} |\langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell' m'}^{N*} \rangle|^2, \quad (5.141)$$

$$M_{EE, \ell \ell'}^N = \frac{2}{2\ell + 1} \sum_{mm'} |\langle \tilde{a}_{E, \ell m}^N \tilde{a}_{E, \ell' m'}^{N*} \rangle|^2, \quad (5.142)$$

$$M_{BB, \ell \ell'}^N = \frac{2}{2\ell + 1} \sum_{mm'} |\langle \tilde{a}_{B, \ell m}^N \tilde{a}_{B, \ell' m'}^{N*} \rangle|^2, \quad (5.143)$$



$$M_{CC,\ell\ell'}^N = \frac{1}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{E,\ell m}^N \tilde{a}_{E,\ell'm'}^{N*} \rangle \langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell'm'}^{N*} \rangle, \quad (5.144)$$

$$M_{EB,\ell\ell'}^N = \frac{2}{2\ell+1} \sum_{mm'} | \langle \tilde{a}_{E,\ell m}^N \tilde{a}_{B,\ell'm'}^{N*} \rangle |^2, \quad (5.145)$$

$$(5.146)$$

all others combinations are zero. I then find the total correlation matrix  $M_{ZZ',\ell\ell'}$  consisting of both signal and noise. As for temperature, this is not simply the sum of the correlation matrix for signal and noise, one also gets cross terms. The final result is

$$M_{TT,\ell\ell'} = M_{TT,\ell\ell'}^S + M_{TT,\ell\ell'}^N + \quad (5.147)$$

$$\frac{4}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{\ell m}^S \tilde{a}_{\ell'm'}^{S*} \rangle \langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell'm'}^{N*} \rangle, \quad (5.148)$$

$$M_{EE,\ell\ell'} = M_{EE,\ell\ell'}^S + M_{EE,\ell\ell'}^N + \quad (5.149)$$

$$\frac{4}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{E,\ell m}^S \tilde{a}_{E,\ell'm'}^{S*} \rangle \langle \tilde{a}_{E,\ell m}^N \tilde{a}_{E,\ell'm'}^{N*} \rangle, \quad (5.150)$$

$$M_{BB,\ell\ell'} = M_{BB,\ell\ell'}^S + M_{BB,\ell\ell'}^N + \quad (5.151)$$

$$\frac{4}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{B,\ell m}^S \tilde{a}_{B,\ell'm'}^{S*} \rangle \langle \tilde{a}_{B,\ell m}^N \tilde{a}_{B,\ell'm'}^{N*} \rangle, \quad (5.152)$$

$$M_{CC,\ell\ell'} = M_{CC,\ell\ell'}^S + M_{CC,\ell\ell'}^N + \quad (5.153)$$

$$\frac{1}{2\ell+1} \sum_{mm'} \left( \langle \tilde{a}_{E,\ell m}^S \tilde{a}_{E,\ell'm'}^{S*} \rangle \langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell'm'}^{N*} \rangle \right. \quad (5.154)$$

$$\left. + \langle \tilde{a}_{\ell m}^S \tilde{a}_{\ell'm'}^{S*} \rangle \langle \tilde{a}_{E,\ell m}^N \tilde{a}_{E,\ell'm'}^{N*} \rangle \right), \quad (5.155)$$

$$M_{TE,\ell\ell'} = M_{TE,\ell\ell'}^S \quad (5.156)$$

$$M_{TB,\ell\ell'} = M_{TB,\ell\ell'}^S \quad (5.157)$$

$$M_{TC,\ell\ell'} = M_{TC,\ell\ell'}^S + \frac{2}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{\ell m}^S \tilde{a}_{E,\ell'm'}^{S*} \rangle \langle \tilde{a}_{\ell m}^N \tilde{a}_{\ell'm'}^{N*} \rangle, \quad (5.158)$$

$$M_{EB,\ell\ell'} = M_{EB,\ell\ell'}^S + M_{EB,\ell\ell'}^N + \quad (5.159)$$

$$\frac{4}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{E,\ell m}^S \tilde{a}_{B,\ell'm'}^{S*} \rangle \langle \tilde{a}_{E,\ell m}^N \tilde{a}_{B,\ell'm'}^{N*} \rangle, \quad (5.160)$$

$$M_{EC,\ell\ell'} = M_{EC,\ell\ell'}^S + \frac{2}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{E,\ell m}^S \tilde{a}_{\ell'm'}^{S*} \rangle \langle \tilde{a}_{E,\ell m}^N \tilde{a}_{\ell'm'}^{N*} \rangle, \quad (5.161)$$

$$M_{BC,\ell\ell'} = M_{BC,\ell\ell'}^S + \frac{2}{2\ell+1} \sum_{mm'} \langle \tilde{a}_{B,\ell m}^S \tilde{a}_{\ell'm'}^{S*} \rangle \langle \tilde{a}_{B,\ell m}^N \tilde{a}_{\ell'm'}^{N*} \rangle, \quad (5.162)$$

$$(5.163)$$

## 5.3 Results of Likelihood Estimations

The likelihood estimation was carried out in the same way as for the temperature power spectrum. The power spectra were estimated in bins defined as

$$C_\ell^T = \frac{D_b^T}{\ell(\ell+1)}, \quad \ell_b \leq \ell < \ell_{b+1}, \quad (5.164)$$

$$C_\ell^E = \frac{D_b^E}{\ell(\ell+1)}, \quad \ell_b \leq \ell < \ell_{b+1}. \quad (5.165)$$

$$(5.166)$$

A similar binning does not work for the temperature-polarisation cross correlation power spectrum. The reason for this is the Schwarz inequality  $C_\ell^C \leq \sqrt{C_\ell^T C_\ell^E}$ . During likelihood maximization one must make sure that the estimated value of  $C_\ell^C$  never exceeds  $\sqrt{C_\ell^T C_\ell^E}$ . The way I solved this problem was to estimate for  $C_\ell^C / \sqrt{C_\ell^T C_\ell^E}$  under the constraint that this value never exceeds 1. So the binning is then

$$C_\ell^C = D_b^C \sqrt{C_\ell^T C_\ell^E} \quad (5.167)$$

$$= \frac{D_b^C \sqrt{D_b^T D_b^E}}{\ell(\ell+1)}, \quad (5.168)$$

where as before  $\ell_b \leq \ell < \ell_{b+1}$ .

As an example I simulated a sky using  $N_{side} = 512$  resolution in Healpix and a  $10'$  beam. I added non-uniform noise to the map. A reasonable assumption about the size of the noise deviations for polarisation is to take  $\sigma_j^P = \sqrt{2}\sigma_j^T$  (Zaldarriaga and Seljak 1997). This is what I used. The noise level was set so that the signal to noise ration for the temperature power spectrum was always well above 1 below the maximum multipole  $\ell = 1024$  whereas for the  $E$  mode polarisation power spectrum it was mostly below 1 (see figure (5.25)). This is close to the values expected for the Planck  $143GHz$  channel. For the analysis I used a 15 degree Gaussian Gabor window. The result of one single estimation is shown in figure (5.25).

To test whether the method is bias or not, I did 60 Monte Carlo simulations. The result of the average of these simulations is shown in figure (5.24). The method seems to be unbiased also for the estimates of the polarisation power spectra. Note that the expected noise variance taken from the approximate analytic formula for uniform noise (shaded areas on the plot) here fails to predict the size of the error bars on the estimates. The expected variance taken from the inverse Fisher matrix (dashed lines) fits better with the error bars from Monte

Carlo. The reason is that I as in chapter (4) used a noise profile with increasing noise from the centre of the disc and down to the edges, opposite of the Gaussian window.

## 5.4 Discussion

An extension of the power spectrum estimation method developed in chapter (4) has been made in order to estimate for polarisation power spectra. The method has been tested here under the assumption that the  $B$  mode polarisation is negligible. In this case the method appears to give unbiased estimates of the polarisation power spectra also in the presence of non-uniform noise and a Gabor window.

The kernels connecting the full sky polarisation power spectra with the cut sky polarisation pseudo power spectra were studied and found to be very similar to the kernel for the temperature power spectrum. For this reason the effect of a cut sky and a Gabor window on the polarisation power spectra is similar to the effect on the temperature power spectrum. This explains that the method of estimating the power spectrum from the pseudo power spectrum also worked for polarisation.

One issue which has not been studied here is the inclusion of the  $B$  mode polarisation. I demonstrated that the  $E$  and  $B$  mode polarisation power spectra are mixing on the cut sky making detections of the much weaker  $B$  component difficult. Further work needs to be done in order to include the  $B$  component in the likelihood analysis.

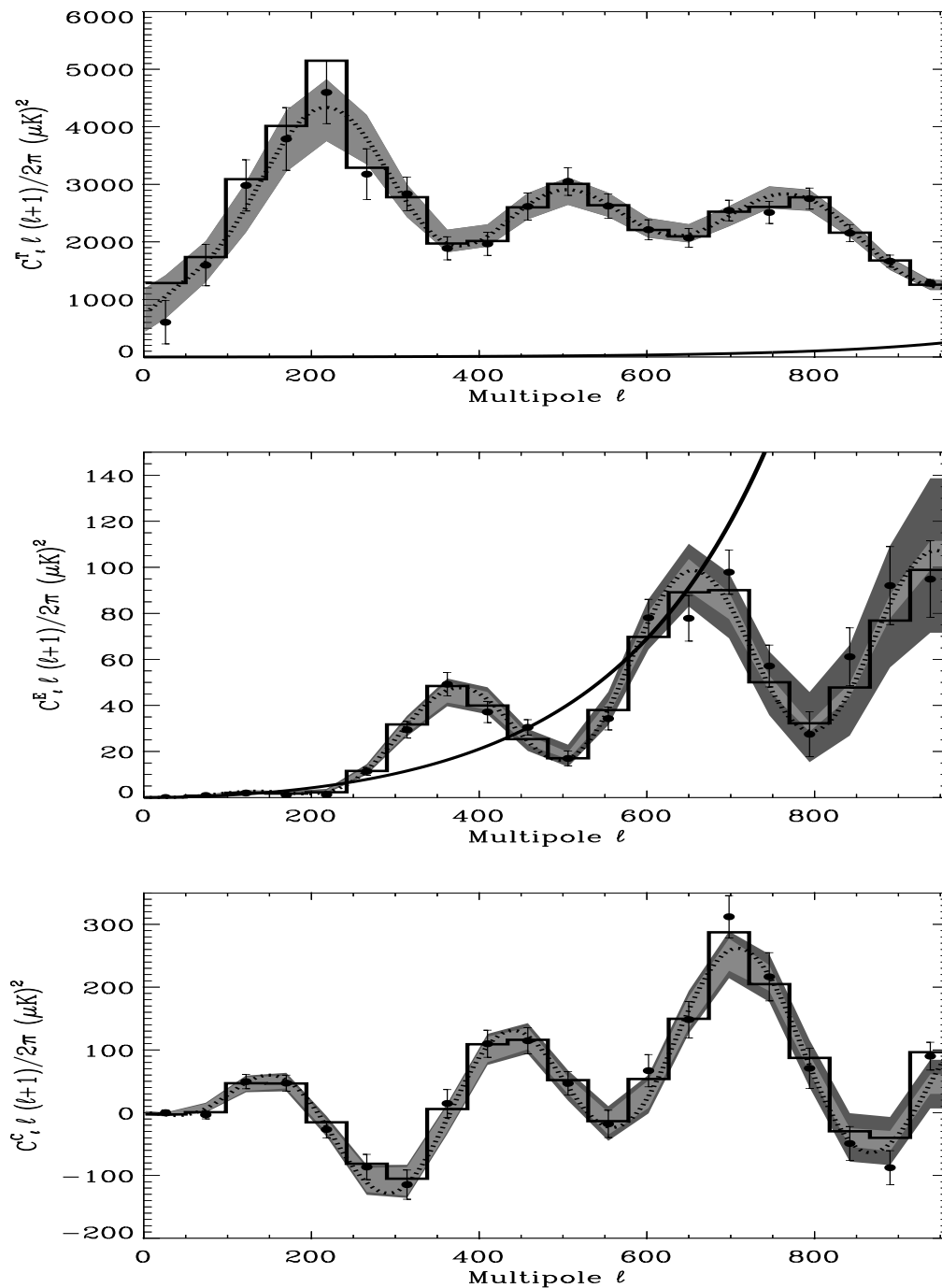


Figure 5.24: The result of a joint likelihood estimation of the temperature power spectrum (upper plot) and the  $E$  (middle plot) and  $C$  (lower plot) polarisation power spectra. The dotted line shows the full sky average spectrum. The histogram shows the binned input pseudo spectrum without noise. The shaded areas around the binned average full sky power spectrum (not shown) show the expected deviations from the average using the approximate formula for uniform noise. The bright shaded area shows the cosmic and sample variance only whereas the dark shaded area also shows expected variance due to noise. The dots show the estimate with  $1\sigma$  error bars taken from the inverse Fisher matrix. In the analysis a 15 degree FWHM Gaussian Gabor window with a  $\theta_C = 3\sigma$  cutoff was used.

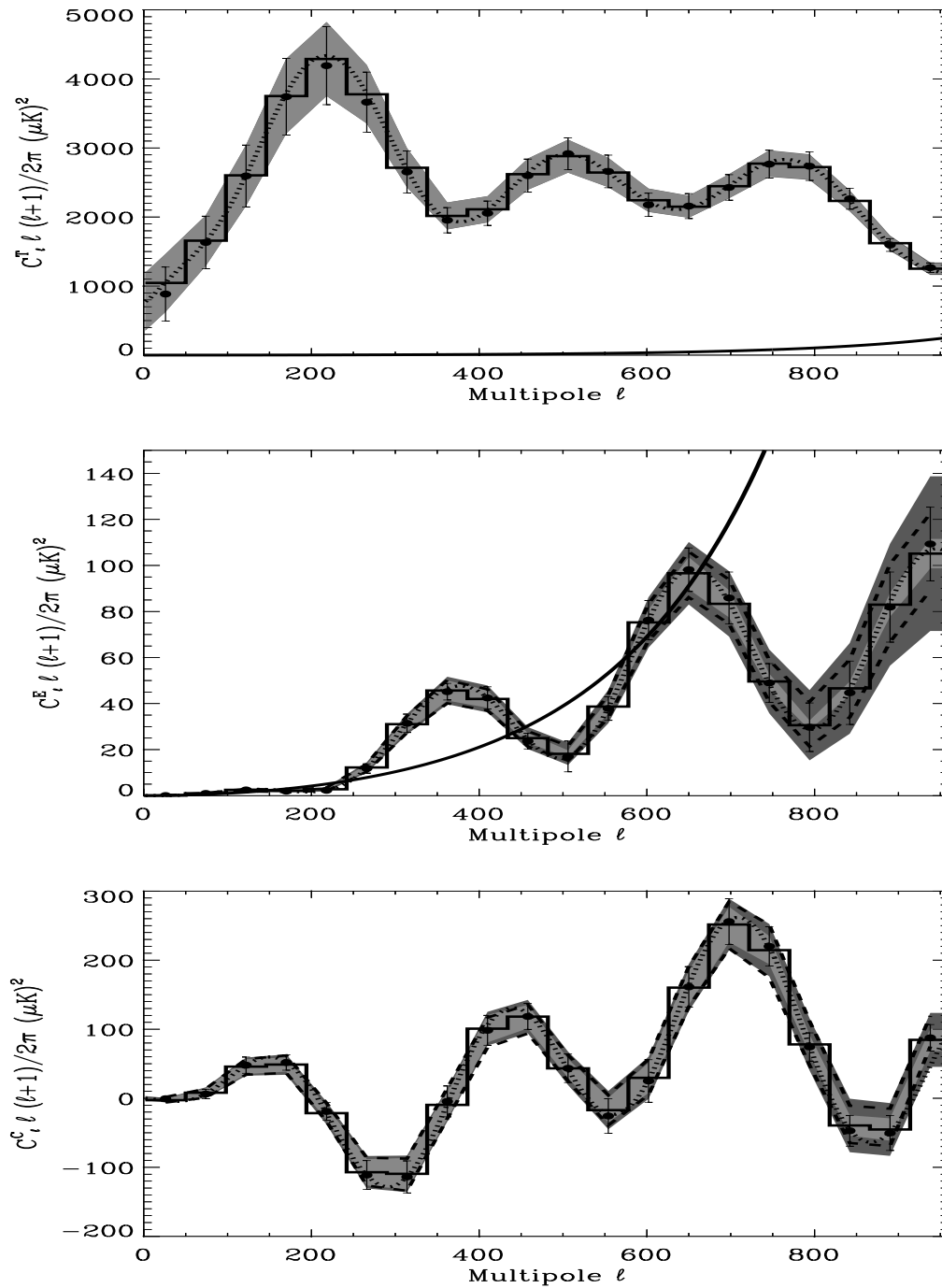


Figure 5.25: Same as figure (5.24) but the dots here are the average of 60 estimates from Monte Carlo simulations. The error bars are the average deviations taken from the simulations. The dotted line shows the average full sky spectrum. The shaded areas which are plotted around the binned full sky power spectrum (not shown) show the variance taken from the approximate variance formula for uniform noise. The dashed lines show the expected variance taken from the inverse Fisher matrix.



# Appendix A

## Rotation Matrices

A spherical function  $T(\hat{\mathbf{n}})$  is rotated by the operator  $\hat{D}(\alpha\beta\gamma)$  where  $\alpha\beta\gamma$  are the three Euler angles for rotations (See T.Risbo 1996) and the inverse rotation is  $\hat{D}(-\gamma - \beta - \alpha)$ . For the spherical harmonic functions, this operator takes the form,

$$Y_{\ell m}(\hat{\mathbf{n}}') = \sum_{m'=-\ell}^{\ell} D_{m'm}^{\ell}(\alpha\beta\gamma) Y_{\ell m'}(\hat{\mathbf{n}}), \quad (\text{A.1})$$

where  $D_{m'm}^{\ell}$  has the form

$$D_{m'm}^{\ell}(\alpha\beta\gamma) = e^{im'\alpha} d_{m'm}^{\ell}(\beta) e^{im\gamma}. \quad (\text{A.2})$$

Here  $d_{m'm}^{\ell}(\beta)$  is a real coefficient with the following property:

$$d_{m'm}^{\ell}(\beta) = d_{mm'}^{\ell}(-\beta). \quad (\text{A.3})$$

The D-functions also have the following property:

$$D_{m'm}^{\ell}(\alpha\beta\gamma) = \sum_{m''} D_{m'm''}^{\ell}(\alpha_2\beta_2\gamma_2) D_{m''m}^{\ell}(\alpha_1\beta_1\gamma_1), \quad (\text{A.4})$$

where  $(\alpha\beta\gamma)$  is the result of the two consecutive rotations  $(\alpha_1\beta_1\gamma_1)$  and  $(\alpha_2\beta_2\gamma_2)$ .

The complex conjugate of the rotation matrices can be written as

$$D_{mm'}^{\ell*} = (-1)^{m+m'} D_{(-m)(-m')}^{\ell}. \quad (\text{A.5})$$

See also Appendix (B).

# Appendix B

## Spin-s Harmonics

The spherical harmonic functions  $Y_{\ell m}(\hat{\mathbf{n}})$  can be generalized to *spin-s harmonics* using the rotation matrices in Appendix (A). The general definition is

$$D_{-sm}^{\ell}(\phi_2, \theta, \phi_1) = \sqrt{\frac{4\pi}{2\ell+1}} {}_s Y_{\ell m}(\theta, \phi_2) e^{-is\phi_1}, \quad (\text{B.1})$$

or in the form which will be mostly used in this thesis

$${}_s Y_{\ell m}(\theta, \phi) = \sqrt{\frac{2\ell+1}{4\pi}} D_{-sm}^{\ell}(\phi, \theta, 0). \quad (\text{B.2})$$

The spin-s harmonics have the orthogonality and completeness relations given by

$$\int d\hat{\mathbf{n}} {}_s Y_{\ell m}(\hat{\mathbf{n}}) {}_s Y_{\ell m'}(\hat{\mathbf{n}}) = \delta_{\ell\ell'} \delta_{mm'} \quad (\text{B.3})$$

$$\sum_{\ell m} {}_s Y_{\ell m}(\hat{\mathbf{n}}) {}_s Y_{\ell m}(\hat{\mathbf{n}}_0) = \delta(\hat{\mathbf{n}} - \hat{\mathbf{n}}_0). \quad (\text{B.4})$$

The complex conjugate of the spin harmonics can be written

$${}_s Y_{\ell m}^*(\hat{\mathbf{n}}) = (-1)^{s+m} {}_{-s} Y_{\ell(-m)}(\hat{\mathbf{n}}). \quad (\text{B.5})$$



# Appendix C

## Some Wigner Symbol Relations

Throughout the paper, the Wigner 3j Symbols will be used frequently. Here are some relations for these symbols, which are used. The orthogonality relation is,

$$\sum_{mm'} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix} \begin{pmatrix} \ell & \ell' & L'' \\ m & m' & M'' \end{pmatrix} = (2\ell'' + 1)^{-1} \delta_{\ell'' L''} \delta_{m'' M''}. \quad (\text{C.1})$$

The Wigner 3j Symbols can be represented as an integral of rotation matrices (see Appendix(A)),

$$\frac{1}{8\pi^2} \int d\cos\theta d\phi d\gamma D_{m_1 m'_1}^\ell D_{m_2 m'_2}^{\ell'} D_{m_3 m'_3}^{\ell''} = \begin{pmatrix} \ell & \ell' & \ell'' \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m'_1 & m'_2 & m'_3 \end{pmatrix}. \quad (\text{C.2})$$

This expression can be reduced to,

$$\int d\hat{\mathbf{n}} Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell' m'}(\hat{\mathbf{n}}) Y_{\ell'' m''}(\hat{\mathbf{n}}) = \sqrt{\frac{(2\ell + 1)(2\ell' + 1)(2\ell'' + 1)}{4\pi}} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{C.3})$$

# Appendix D

## Recurrence Relation

It is important for the precalculations to the likelihood analysis that the calculation of  $h(\ell, \ell', m)$  is fast. For this reason a recurrence relation for  $h(\ell, \ell', m)$  would be helpful. To speed up the calculation of the noise correlation matrix for non axisymmetric noise, it would also help if one had a more general recurrence relation for  $h(\ell, \ell', m, m')$ . I will now show how to find such a recurrence for these objects which I now call  $A_{\ell'\ell}^{m'm}$  to simplify notation (and for the notation to comply with (Wandelt, Hivon, and Górski 2000)). The definition is again,

$$A_{\ell'\ell}^{m'm} = \int_a^b G(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}) Y_{\ell' m'}(\hat{\mathbf{n}}) d\hat{\mathbf{n}}, \quad (\text{D.1})$$

where  $G(\hat{\mathbf{n}}) = G(\theta, \phi)$  is a general function and  $Y_{\ell m}$  are the spherical harmonics which can be factorised into one part dependent on  $\theta$  and one dependent on  $\phi$  in the following way,

$$Y_{\ell m}(\theta, \phi) = \lambda_{\ell m}(\cos \theta) e^{-i\phi m}. \quad (\text{D.2})$$

Now writing,

$$A_{\ell'\ell}^{m'm} = \int d \cos \theta \lambda_{\ell m}(\cos \theta) \lambda_{\ell' m'}(\cos \theta) \int d\phi e^{-i\phi(m-m')} G(\theta, \phi) \quad (\text{D.3})$$

$$\equiv \int d \cos \theta \lambda_{\ell m}(\cos \theta) \lambda_{\ell' m'}(\cos \theta) F_{m'm}(\theta), \quad (\text{D.4})$$

where  $F_{m'm}(\theta)$  is simply the Fourier transform of the window at each  $\theta$ . The quantities  $A_{\ell'\ell}^{m'm}$  and  $F_{m'm}(\theta)$  are in general complex quantities obeying,

$$A_{\ell'\ell}^{m'm} = (A_{\ell\ell'}^{mm'})^* \quad F_{m'm}(\theta) = (F_{mm'}(\theta))^* \quad (\text{D.5})$$

The  $A_{\ell'\ell}^{m'm}$  can be expressed as

$$A_{\ell'\ell}^{m'm} = \frac{\sqrt{(2\ell' + 1)(2\ell + 1)}}{2} \sqrt{\frac{(\ell' - m')!(\ell - m)!}{(\ell' + m')!(\ell + m)!}} I_{\ell'\ell}^{m'm}, \quad (\text{D.6})$$

where  $I_{\ell'\ell}^{m'm}$  is defined as:

$$I_{\ell'\ell}^{m'm} = \int_a^b F_{m'm}(x) P_\ell^m(x) P_{\ell'}^{m'}(x) dx. \quad (\text{D.7})$$

The following relation for the Legendre Polynomials will be used:

$$x P_\ell^m = \frac{\ell - m + 1}{2\ell + 1} P_{\ell+1}^m + \frac{\ell + m}{2\ell + 1} P_{\ell-1}^m \quad (\text{D.8})$$

I now define the object  $X_{\ell'\ell}^{m'm}$  as

$$X_{\ell'\ell}^{m'm} = \int_a^b F_{m'm}(x) x P_\ell^m P_{\ell'}^{m'} dx \quad (\text{D.9})$$

Using relation (D.8) in this definition, one gets,

$$X_{\ell'\ell}^{m'm} = \frac{\ell - m + 1}{2\ell + 1} I_{\ell'(\ell+1)}^{m'm} + \frac{\ell + m}{2\ell + 1} I_{\ell'(\ell-1)}^{m'm} \quad (\text{D.10})$$

One can also exchange  $(\ell, \ell')$  and  $(m, m')$  to get

$$X_{\ell'\ell}^{mm'} = \frac{\ell' - m' + 1}{2\ell' + 1} I_{\ell(\ell'+1)}^{mm'} + \frac{\ell' + m'}{2\ell' + 1} I_{\ell(\ell'-1)}^{mm'} \quad (\text{D.11})$$

Taking the complex conjugate of the first expression and subtracting the last, one has

$$(X_{\ell'\ell}^{m'm})^* - X_{\ell'\ell}^{mm'} = 0 = \frac{\ell - m + 1}{2\ell + 1} (I_{\ell'(\ell+1)}^{m'm})^* + \frac{\ell + m}{2\ell + 1} (I_{\ell'(\ell-1)}^{m'm})^* \quad (\text{D.12})$$

$$- \frac{\ell' - m' + 1}{2\ell' + 1} I_{\ell(\ell'+1)}^{mm'} - \frac{\ell' + m'}{2\ell' + 1} I_{\ell(\ell'-1)}^{mm'} \quad (\text{D.13})$$

Then setting  $\ell' = \ell - 1$  one gets:

$$I_{\ell'\ell}^{m'm} = \frac{2\ell' - 1}{\ell' - m'} \left( \frac{\ell - m + 1}{2\ell + 1} I_{(\ell'-1)(\ell+1)}^{m'm} + \frac{\ell + m}{2\ell + 1} I_{(\ell'-1)(\ell-1)}^{m'm} - \frac{\ell' + m - 1}{2\ell' - 1} I_{(\ell'-2)\ell}^{m'm} \right) \quad (\text{D.14})$$

Using equation (D.6), one can express this as

$$A_{\ell'\ell}^{m'm} = \frac{1}{\sqrt{\ell'^2 - m'^2}} \left( \sqrt{\frac{(4\ell'^2 - 1)((\ell + 1)^2 - m^2)}{(2\ell + 1)(2\ell + 3)}} A_{(\ell'-1)(\ell+1)}^{m'm} \right. \\ \left. + \sqrt{\frac{(4\ell'^2 - 1)(\ell^2 - m^2)}{4\ell^2 - 1}} A_{(\ell'-1)(\ell-1)}^{m'm} - \sqrt{\frac{(2\ell' + 1)((\ell' - 1)^2 - m'^2)}{2\ell' - 3}} A_{(\ell'-2)\ell}^{m'm} \right), \quad (\text{D.15})$$

which is the final recurrence relation. The  $A_{m'\ell}^{m'm}$  elements must be provided before the recurrence is started. Then for each  $(m, m')$ , set  $\ell' = m' + 1$  and let  $\ell$  go from

$\ell'$  and upwards, then set  $\ell' = m' + 2$  and again let  $\ell$  go from  $\ell'$  and upwards. Continue to the desired size of  $\ell'$ . Note that, in order to get all objects up to  $A_{\ell_{max}\ell_{max}}^{m'm}$  one need to go up to  $\ell = 2\ell_{max}$  all the time during recursion. This is because of the  $A_{(\ell'-1)(\ell+1)}^{m'm}$  term which all the time demands an object indexed  $(\ell + 1)$  in the previous  $\ell'$  row.

To start the recurrence, one can precompute the  $A_{m'\ell}^{m'm}$  factors fast and easily using FFT and a sum over rings on the grid. F.ex. for the HEALPix grid, I did it the following way,

$$A_{m'\ell}^{m'm} = \sum_r \lambda_{m'm'}^r \lambda_{\ell m}^r \sum_{j=0}^{N_r-1} e^{-2\pi i j / N_r (m-m')} G_{rj}, \quad (\text{D.16})$$

where the last part is the Fourier transform of the Gabor window, calculated by FFT,  $r$  is ring number on the grid and  $j$  is azimuthal position on each ring. Ring  $r$  has  $N_r$  pixels.

It turns out that the recurrence can be numerically unstable dependent on the window and multipole, and in order to avoid problems I (using double precision numbers) restart the recurrence with a new set of precomputed  $A_{\ell'\ell}^{m'm}$  for every 50th  $\ell'$  row. However for some windows and multipoles the recurrence can run for hundreds of  $\ell$ -rows without problems.

# Appendix E

## Extention of the Recurrence Relation to Polarisation

For fast calculations of correlation matrices for polarisation, it would be pleasant to have a recurrence relation for

$$h_2(\ell, \ell', m, m') \equiv \int d\hat{\mathbf{n}} G(\hat{\mathbf{n}}) {}_2Y_{\ell m}(\hat{\mathbf{n}}) {}_2Y_{\ell' m'}(\hat{\mathbf{n}}), \quad (\text{E.1})$$

similar to the one for  $h(\ell, \ell', m, m')$  in Appendix (D). Again I simplify the notation by calling the function  $A_{\ell m}^{\ell' m'}$ . Separating the spin-2 harmonic one can write

$${}_2Y_{\ell m}(\theta, \phi) = {}_2\lambda_{\ell m}(\cos \theta) e^{-im\phi}. \quad (\text{E.2})$$

In this way one can write  $A_{\ell m}^{\ell' m'}$  as

$$A_{\ell m}^{\ell' m'} = \int dx {}_2\lambda_{\ell m}(x) {}_2\lambda_{\ell' m'}(x) \underbrace{\int_0^{2\pi} d\phi G(\theta, \phi) e^{-i\phi(m-m')}}_{\equiv F_{mm'}(x)}, \quad (\text{E.3})$$

where  $x = \cos \theta$ . As before I define

$$X_{\ell m}^{\ell' m'} = \int dx x {}_2\lambda_{\ell m}(x) {}_2\lambda_{\ell' m'}(x) F_{mm'}(x), \quad (\text{E.4})$$

where obviously  $F_{mm'}(x)^* = F_{m'm}(x)$  and therefore  $X_{\ell m}^{\ell' m'} = (X_{\ell' m'}^{\ell m})^*$ . The next step is to use a recurrence relation for spin-2 harmonics

$$x {}_2\lambda_{\ell m}(x) = p_{\ell m} {}_2\lambda_{(\ell+1)m}(x) - q_{\ell m} {}_2\lambda_{\ell m}(x) + p_{(\ell-1)m} {}_2\lambda_{(\ell-1)m}(x), \quad (\text{E.5})$$

where

$$p_{\ell m} = \frac{1}{\ell+1} \sqrt{\frac{((\ell+1)^2 - m^2)((\ell+1)^2 - 4)}{4(\ell+1)^2 - 1}}, \quad (\text{E.6})$$

$$q_{\ell m} = \frac{2m}{\ell(\ell+1)}. \quad (\text{E.7})$$

In this way one has

$$X_{\ell m}^{\ell' m'} = p_{\ell m} A_{(\ell+1)m}^{\ell' m'} - q_{\ell m} A_{\ell m}^{\ell' m'} + p_{(\ell-1)m} A_{(\ell-1)m}^{\ell' m'}, \quad (\text{E.8})$$

$$X_{\ell' m'}^{\ell m} = p_{\ell' m'} A_{(\ell'+1)m'}^{\ell m} - q_{\ell' m'} A_{\ell' m'}^{\ell m} + p_{(\ell'-1)m'} A_{(\ell'-1)m'}^{\ell m}, \quad (\text{E.9})$$

Subtracting the complex conjugate of equation (E.9) from equation (E.8) the left side is zero and one is left with

$$A_{\ell m}^{\ell' m'} = \frac{1}{p_{(\ell'-1)m'}} \left[ p_{\ell m} A_{(\ell+1)m}^{(\ell'-1)m'} + (q_{(\ell'-1)m'} - q_{\ell m}) A_{\ell m}^{(\ell'-1)m'} \right. \quad (\text{E.10})$$

$$\left. + p_{(\ell-1)m} A_{(\ell-1)m}^{(\ell'-1)m'} - p_{(\ell'-2)m'} A_{\ell m}^{(\ell'-2)m'} \right]. \quad (\text{E.11})$$

This is the recursion formula. All the  $A_{\ell m}^{\ell' m'}$  can be found using the same scheme as for the scalar harmonics described in appendix (D). This equation together with equation (D.15) are two of the main results in the thesis. These formulae allow fast calculations of the couplings between scalar and tensor harmonics on a cut sphere.

# Bibliography

- Baccigalupi, C. et al. (2000). Neural networks and the separation of cosmic microwave background and astrophysical signals in sky maps. *MNRAS* *318*, 769.
- Bardeen, J. M. (1980). Gauge-invariant cosmological perturbations. *Phys. Rev. D* *22*, 1882.
- Bartlett, J. G., M. Douspis, A. Blanchard, and M. L. Dour (2000). An approximation to the likelihood function for band-power estimates of cmb anisotropies. *A&ASS* *146*, 507.
- Bennet, C. L. et al. (1996). Four-year coBE dmr cosmic microwave background observations: Maps and basic results. *ApJ* *464*, L1.
- Bennett, C. L. et al. (1992). Preliminary separation of galactic and cosmic microwave emission for the coBE differential microwave radiometer. *ApJ* *396*, L7.
- Bennett, C. L. et al. (1994). Cosmic temperature fluctuations from two years of coBE differential microwave radiometers observations. *ApJ* *436*, 423.
- Benoit, A. et al. (2001). Archeops: A high resolution, large sky coverage balloon experiment for mapping cmb anisotropies. *astro-ph/0106152*.
- Bernardis, P. D. et al. (2000). A flat universe from high-resolution maps of the cosmic microwave background radiation. *Nature* *404*, 955.
- Bersanelli, M. et al. (1996). Cobras/samba: report on the phase a study.
- Bond, J. R., G. Efstathiou, and M. Tegmark (1997). Forecasting cosmic parameter errors from microwave background anisotropy experiments. *MNRAS* *291*, L33.
- Bond, J. R., A. H. Jaffe, and L. Knox (1998). Estimating the power spectrum of the cosmic microwave background. *Phys.Rev.D* *57*, 2117.
- Bond, J. R., A. H. Jaffe, and L. Knox (2000). Radical compression of cosmic microwave background data. *ApJ* *533*, 19.
- Bouchet, F. R. and R. Gispert (1999). Foregrounds and cmb experiments i. semi-analytical estimates of contamination. *NewA* *4*, 443.

- Chandrasekhar, S. (1960). *Radiative Transfer*. Dover, New York.
- Chiueh, T. and C. Ma (2001). The annulus-filtered e and b modes in the cmb polarisation. *astro-ph/0101205*.
- Coles, P. and F. Lucchin (1995). *Cosmology*. John Wiley & Sons.
- Delabrouille, J. (1998). Analysis of the accuracy of a destriping method for future cosmic microwave background mapping with the planck surveyor satellite. *A&ASS 127*, 555.
- Dore, O. et al.
- Doré, O., L. Knox, and A. Peel (2001). Cmb power spectrum estimation via hierarchical decomposition. *astro-ph/0104443*.
- Ferreira, P. G. and A. H. Jaffe (2000). Simultaneous estimation of noise and signal in the cosmic microwave background experiments. *MNRAS 312*, 89.
- Gabor, D. (1946). *J. Inst. Elect. Eng. 93*, 429.
- Górski et al. Homepage: <http://www.eso.org/science/healpix/>.
- Górski, K. M. (1994). On determining the spectrum of primordial inhomogeneity from the coBE dmr sky maps: method. *ApJ 430*, L85.
- Górski, K. M. et al. (1996). Power spectrum of primordial inhomogeneity determined from the four-year coBE dmr sky maps. *ApJ 464*, L11.
- Górski, K. M. and F. K. Hansen (2002). Cmb power spectrum estimation by combining multiple patches on the sky. *in prep.*
- Hanany, S. et al. (2000). Maxima-1: A measurement of the cosmic microwave background anisotropy on angular scales of  $10'$ - $5^\circ$ . *ApJ 545*, L5.
- Haslam, C. G. T. et al. (1981). A 408 mhz all-sky continuum survey. i - observations at southern declinations and for the north polar region. *A&A 100*, 209.
- Hivon, E. et al. (2001). Master of the cmb anisotropy power spectrum: A fast method for statistical analysis of large and complex cmb data sets. *astro-ph/0105302*.
- Hobson, M. P. et al. (1999). The effect of point sources on satellite observations of the cosmic microwave background. *MNRAS 306*, 232.
- Hobson, M. P., A. W. Jones, A. N. Lasenby, and F. R. Bouchet (1998). Foreground separation methods for satellite observations of the cosmic microwave background. *MNRAS 300*, 1.
- Hobson, M. P. and J. Magueijo (1996). *MNRAS 283*, 1133.
- Hu, W. and N. Sugiyama (1995). Anisotropies in the cosmic microwave background. *ApJ 444*, 489.



- Jaffe, A. H. et al. (2001). Cosmology from maxima-1, boomerang, and cobe dmr cosmic microwave background observations. *Phys. Rev. Lett* 86, 3475.
- Jungman, G., M. Kamionkowski, A. Kosowsky, and D. N. Spergel (1996). Cosmological-parameter determination with microwave background maps. *Phys. Rev. D* 54, 1332.
- Kamionkowski, M., A. Kosowsky, and A. Stebbins (1997). Statistics of cosmic microwave background polarization. *Phys. Rev. D* 55, 7368.
- Kogut, A. et al. (1996). High-latitude galactic emission in the cobe differential microwave radiometer 2 year sky maps. *ApJ* 460, 1.
- Kolb, E. W. and M. S. Turner (1990). *The Early Universe*. Addison-Wesley Publishing Company.
- Lewis, A., A. Challinor, and N. Turok (2001). Analysis of cmb polarisation on an incomplete sky. *astro-ph/0106536*.
- Ma, C. P. and E. Bertschinger (1995). Cosmological perturbation theory in the synchronous and conformal newtonian gauges. *ApJ* 455, 7.
- Maino, D. et al. (1999). The planck-lfi instrument: Analysis of the  $1/f$  noise and implications for the scanning strategy. *A&ASS* 140, 383.
- Maino, D. et al. (2001). All-sky astrophysical component separation with fast independent component analysis (fastica). *astro-ph/0108362*.
- Mauskopf, P. D. et al. (2000). Measurement of a peak in the cosmic microwave background power spectrum from the north american test flight of boomerang. *ApJ* 536, L59.
- Melchiorri, A. et al. (2000). Measurement of a peak in the cosmic microwave background power spectrum from the north american test flight of boomerang. *ApJ* 536, L63.
- Myers, S. T., J. E. Baker, A. C. S. Readhead, and E. M. Leitch (1997). Measurements of the sunyaev-zeldovich effect in the nearby clusters a478, a2142, and a2256. *ApJ* 485, 1.
- Narlikar, J. V. and T. Padmanabhan (1991). Inflation for astronomers. *ARAA* 29, 325.
- Natoli, P. et al. (2001). Non-iterative methods to estimate the in-flight noise properties of cmb detectors. *astro-ph/0110508*.
- Natoli, P., G. de Gasperis, C. Gheller, and N. Vittorio (2001). A map-making algorithm for the planck surveyor. *A&A* 372, 346.
- Netterfield, C. B. et al. (2000). A measurement by boomerang of multiple peaks in the angular power spectrum of the cosmic microwave background. *astro-ph/0104460*.

- Oh, S. P., D. N. Spergel, and G. Hinshaw (1999). An efficient technique to determine the power spectrum from cosmic microwave background sky maps. *ApJ* 510, 551.
- Oliveira-Costa, A. D. et al. (1997). Galactic microwave emission at degree angular scales. *ApJ* 482, 17.
- Ostriker, J. P. and E. T. Vishniac (1986). Generation of microwave background fluctuations from nonlinear perturbations at the era of galaxy formation. *ApJ* 306, L51.
- Padmanabhan, T. (1999). Aspects of gravitational clustering. *astro-ph/9911374*.
- Parratt, L. G. (1961). *Probability and experimental errors in science*. John Wiley and Sons, Inc.
- Persi, F., D. N. Spergel, R. Cen, and J. P. Ostriker (1995). Hot gas in superclusters and microwave background distortions. *ApJ* 442, 1.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes*. Cambridge University Press.
- Prunet, S. et al. (2001). Error estimation for the map experiment. *A&A* 373, L13.
- Prunet, S., C. B. Netterfield, E. Hivon, and B. P. Crill (2000). Iterative map-making for scanning experiments. *astro-ph/0006052*.
- Rees, M. J. and D. W. Sciama (1968). *Nature* 217, 511.
- Reich, P. and W. Reich (1988). A map of spectral indices of the galactic radio continuum emission between 408 mhz and 1420 mhz for the entire northern sky. *A&AS* 74, 7.
- Reynolds, R. J. (1984). A measurement of the hydrogen recombination rate in the diffuse interstellar medium. *ApJ* 282, 191.
- Reynolds, R. J. (1992). The optical emission-line background and accompanying emissions at ultraviolet, infrared, and millimeter wavelengths. *ApJ* 392, L35.
- Sachs, R. K. and A. M. Wolfe (1967). Perturbations of a cosmological mode and angular variations of the microwave background. *ApJ* 147, 73.
- Seljak, U. (1996). Gravitational lensing effect on cosmic microwave background anisotropies: A power spectrum approach. *ApJ* 463, 1.
- Seljak, U. and M. Zeldariagga (1996). A line of sight approach to cosmic microwave background anisotropies. *ApJ* 469, 437.
- Silk, J. (1968). Cosmic black-body radiation and galaxy formation. *ApJ* 151, 459.

- Stolyarov, V., M. P. Hobson, M. A. J. Ashdown, and A. N. Lasenby (2001). All-sky component separation for the planck mission. *astro-ph/0105432*.
- Strohmer, T. (1997). *Proc. SampTA - Sampling Theory and Applications, Aveiro/Portugal*, 297.
- Sunyaev, R. A. and Y. B. Zeldovich (1980). The velocity of clusters of galaxies relative to the microwave background. the possibility of its measurement. *MNRAS 190*, 413.
- Szapudi, I. et al. (2001). Fast cosmic microwave background analyses via correlation functions. *ApJ 548*, L115.
- Szapudi, I., S. Prunet, and S. Colombi (2001). Fast analysis of inhomogenous megapixel cosmic microwave background maps. *ApJ 561*, L11.
- Taylor, A. C. (2001). The very small array. *astro-ph/0109343*.
- Tegmark, M. (1996). A method for extracting maximum resolution power spectra from microwave sky maps. *MNRAS 280*, 299.
- Tegmark, M. (1997a). How to make maps from the cosmic microwave background data without losing information. *ApJ 480*, L87.
- Tegmark, M. (1997b). How to measure cmb power spectra without losing information. *Phys.Rev.D 55*, 5895.
- Tegmark, M. et al. (1997). A high-resolution map of the cosmic microwave background around the north celestial pole. *ApJ 474*, 77.
- Tegmark, M. and G. Efstathiou (1996). A method for subtracting foregrounds from multifrequency cmb sky maps. *MNRAS 281*, 1297.
- Tegmark, M., A. N. Taylor, and A. F. Heavens (1997). Karhunen-loeve eigenvalue problems in cosmology: How should we tackle large data sets? *ApJ 480*, 22.
- Tenorio, L. (1999). Applications of wavelets to the analysis of cosmic microwave background maps. *MNRAS 310*, 823.
- Toffolatti, L. et al. (1998). Extragalactic source counts and contributions to the anisotropies of the cosmic microwave background: predictions for the planck surveyor mission. *MNRAS 297*, 117.
- Turner, M. S., M. White, and J. E. Lidsey (1993). Tensor perturbations in inflationary models as a probe of cosmology. *Phys. Rev. D 48*, 4613.
- Vishniac, E. T. (1987). Reionization and small-scale fluctuations in the microwave background. *ApJ 322*, 597.
- Wandelt, B. D. (2000). Advanced methods for cosmic microwave background data analysis: the big  $n^3$  and how to beat it. *astro-ph/0012416*.

- Wandelt, B. D. and K. M. Górski (2000). Fast convolution on the sphere. *astro-ph/0008227*.
- Wandelt, B. D. and F. K. Hansen (2000). Fast, exact cmb power spectrum estimation for a certain class of observational strategies. *astro-ph/0106515*.
- Wandelt, B. D., E. Hivon, and K. M. Górski (2000). The pseudo- $c_\ell$  method: cosmic microwave background anisotropy power spectrum statistics for high precision cosmology. *astro-ph/0008111*.
- Watson, G. S. (2000). An exposition on inflationary cosmology. *astro-ph/0005003*.
- Wright, E. L. (1996). Scanning and mapping strategies for cmb experiments. *proceeding of the IAS CMB Workshop astro-ph/9612006*.
- Zaldariagga, M. and D. D. Harari (1995). Analytic approach to the polarization of the cosmic microwave background in flat and open universes. *Phys. Rev. D* 52, 3276.
- Zaldarriaga, M. and U. Seljak (1997). An all-sky analysis of the polarisation in the microwave background. *Phys. Rev. D* 55, 1830.
- Zeldovich, Y. B. and R. A. Sunyaev (1969). *Ap&SS* 4, 301.

# Acknowledgements

A big thanks to Kris Górski for excellent supervision of my Ph.D. thesis, for always motivating me and giving suggestions to overcome problems, and for always being available for questions and discussions during these 3 years.

A big thanks also goes to Ben Wandelt for supervising me in a part of the thesis, for being a great teacher in CMB analysis and all other parts of physics, for giving me a lot of inspiration and for all the long nights of fruitful discussions.

Thanks to Simon White for making it possible for me to make my Ph.D. work at the Max Planck Institute for Astrophysics.

Thanks to Anthony Banday and Eric Hivon for discussions and helpful suggestions.

Cristina: Thanks a lot for your patience during the last weeks of writing my Ph.D. I had far too little time for you.... And thanks for all support, energy and motivation.

To my parents: Thanks for always supporting me and for always accepting my decisions.

To all my friends in Munich: Thanks a lot for making Munich such a nice place to live.

To all my colleagues at MPA: Thanks for making MPA a great place to work.

And finally I want to thank the Norwegian Research Council (Norges Forskningsråd) for awarding me a Ph.D. grant without which this work would have been impossible.

# Lebenslauf

## Persönliche Daten

Name: Frode Kristian Hansen  
Geboren am: 18.04.1974 in Tønsberg, Norwegen.  
Staatsangehörigkeit: Norwegisch.  
Familienstand: ledig  
Jetzige Adresse:  
Christoph-Probst-Strasse 16-727,  
80805 München.  
Email: frodekh@astro.uio.no

## Hochschulausbildung

seit Okt.1999	Doktorand am Max-Planck-Institut für Astrophysik, Garching bei München.
Jan.1999-Sep.1999	Doktorand am Theoretical Astrophysics Center (TAC), Kopenhagen.
Jan.1999-Dec.2001	Promotionsstipendium vom Norwegischen Forschungsrat. Doktorarbeit betreut von Prof. Kris Górski.
Jan.1997-Nov.1998	Diplomarbeit in Astrophysik betreut von Prof. Per Lilje.
Aug.1993-Dec.1996	Mathematik-, Physik- und Astrophysikkurse mit Examen jedes Semester.
Aug.1993-Nov.1998	Astrophysikstudium an der Universität in Oslo.

## Schulausbildung

Aug.1990-Jun.1993	Tønsberg Gymnasium (Norwegen), Linie für Naturwissenschaft.
Aug.1981-Jun.1990	Grundschule in Tønsberg (Norwegen)



