Bayesian Analysis of Hemispherical Power Asymmetry in the Cosmic Microwave Background

Candidatus Scientiarum Thesis by
Jostein Hoftuft
Institute of Theoretical Astrophysics
University of Oslo

June, 2009

Preface

The current cosmological standard model is based on inflation. This theory predicts that the initial perturbation in the universe had a Gaussian distribution on any given scale, and we would expect that the universe is statistically isotropic and homogeneous, in accordance with the cosmological principle.

The cosmic microwave background (CMB) radiation is basically a snapshot of the universe when it was around 400 000 years old. This radiation has small fluctuations in its temperature as imprinted by the density perturbations at the time it was sent out. After that time, the change in the radiation has been linear, so the characteristics of the radiation still remain the same today. By measuring this radiation today, we are basically probing the earliest times of the universe. From this radiation we can extract information about the universe, such as its age, contents and geometry. Several experiments have set out to detect this ancient radiation. COBE was the first experiment to find the small perturbations in the CMB radiation. Using a beam with FWHM at 7°, the resolution of the observations was low. Later, experiments like BOOMERanG observed the CMB anisotropies at a considerably higher resolution, but only measuring a small patch of the sky. This gave the rise to the standard cosmological model. WMAP performed a high resolution measurement of the entire sky and confirmed the cosmological model and the cosmological parameters could now be determined with greater accuracy. The predicted theory agrees well with the observations of the CMB, and we should therefore expect the universe to be isotropic and homogeneous. But the full-sky maps from WMAP also showed some unexpected features. Several studies of the maps have revealed strong hints of non-Gaussianity and violation of statistical isotropy.

One of these properties was an observed asymmetry in the map. On large scales there were large fluctuations in temperature in one hemisphere, while the other did not have as large fluctuations on the same scales. This asymmetric distribution of power on opposing hemispheres seems to be a violation of isotropy. Several independent works have pointed out this feature. The fluctuation power for large scales seems to have a dipolar distribution.

To try and characterize this distribution, I adopt a simple model where the isotropic CMB field is modulated by a dipole field. The analysis is done using a Bayesian framework. The purpose is to find the strength (amplitude) and direction of this field, and for how small scales this asymmetry is present. The method used here to find these parameters is a time consuming process, as it demands an inversion of a large

covariance matrix. This process scales as $\mathcal{O}(N_{\rm pix}^3)$, where $N_{\rm pix}$ is the number of data points used. The current computational resources therefore sets the limit for how small scales we could compute using this method.

The analysis was done using the five-year WMAP data. The results shows evidence for this model at the 3.5σ level, and an article on the results is to be published in Astrophysical Journal.

Acknowledgments

I would like to give a big thanks to Hans Kristian Eriksen, who has been my supervisor. He has been a source of inspiration through out the whole period, and has always been helpful, conveyed his knowledge in a understandable way and given me much positive feedback. I would also like to thank all the people in Astrokjelleren who have made my time here more fun and exiting, and everybody on Uglelaget.

The computation of this thesis were done using Titan, a computer cluster owned and maintained by the University in Oslo and NOTUR. Throughout the thesis I have used HEALPix.

Jostein Hoftuft

Contents

List of Figures

Part I

Introducing the CMB and its properties

Chapter 1

Introduction

1.1 Introduction

For centuries mankind has been searching for how the universe was created. Almost every culture has a story of how the world was created. Most of these myths are tremendous tales of battles of Gods and other supernatural happenings. The first to explain how the the universe originated using physical laws and scientific description were the Greeks. The word cosmology also came from Greek, *kósmos* meaning world; emphlógos meaning knowledge or science. But it is only in the present century that we have developed good theories that describe the universe as a whole, and have observations that support these theories. The modern picture of the universe is that it started with a hot Big Bang, and one observation that strengthen this view of the universe is the presence of a low energy radiation which fills the whole universe. This radiation stems from the Big Bang and is called the cosmic microwave background. In this chapter I will look at how it was discovered and how it changed our view of the universe.

1.2 The discovery of the CMB

In 1915 Albert Einstein submitted his paper on the general theory of relativity, where the field equations describe gravity as a curved spacetime caused by matter and energy [1]. In 1922 Alexander Friedmann found a solution to the equations in which the universe may expand or contract. At the time, Einstein believed that the universe was static, and he introduced a new term with a cosmological constant to the field equations for them to allow a static universe. But a static universe was unstable in this theory, and the universe would start to expand or contract if it was slightly brought out of equilibrium. Later, in 1927, George Lemaître also published a report where he presented the idea of an expanding universe [2]. Here he also derived the connection between distance and redshift, later to be known as the Hubble law.

By studying the redshift of distant galaxies, Edwin Hubble found a linear relation between the distance and velocity of galaxies. He released this in a paper in 1929,

which strongly supports an expanding universe [3]. It is then easy to imagine that the universe was smaller before, and that it also was much denser and more hot.

Lemaître later proposed that the universe expanded from an initial point, called the "Primeval Atom". This idea was published in Nature in 1931 [4]. He described his theory as "the Cosmic Egg exploding at the moment of the creation".

But not everyone agreed with this interpretation of the expanding universe. This creation seemed philosophically troubling, as this imply a cause or a creator. Therefore, many scientist, e.g., Fred Hoyle, Thomas Gold and Hermann Bondi, argued for a steady-state universe. To explain that the universe is expanding while still obeying the cosmological principle, matter must be continuously created as it expands in order for the average density to remain at the same level at any given time. That way the universe looks the same at all times. This universe model has no beginning and no end, and therefore, the universe is infinitely old.

Fred Hoyle, who was against this expanding universe model, would later give this model the now famous name Big Bang. In a BBC radio broadcast in 1949 he said, "This big bang idea seemed to me to be unsatisfactory even before examination show that it leads to serious difficulties. For when we look at our own galaxy there is not the smallest sign that such an explosion ever occurred." However, there were some who predicted that there would be a remnant signal filling the universe after the Big Bang.

In 1946 George Gamow was pondering the cosmic abundances of the elements. He realized that a newborn, dense universe must be hot enough for the nuclear reactions that create the elements to occur. Two years later Gamow and Ralph Alpher published a paper called "The Origin of Chemical Elements". Later, detailed calculations by Alpher and Ralph Herman showed that Gamow's idea did not produce elements heavier than helium. However, in 1957 Fred Hoyle released a paper about stellar nucleosynthesis which explained the formation of heavier elements. Gamow's theory of the Big Bang nucleosynthesis explained that the early universe must have been very dense and hot, and where radiation and matter was in thermal equilibrium. Under these conditions the photons should have a blackbody spectrum. In 1948 Alpher and Herman published a paper on how this blackbody radiation would cool off as the universe expanded, and they predicted that the universe today would be filled with a blackbody radiation with a temperature of 5 K. But astronomers and scientist did not make any effort to detect this leftover blackbody radiation from the early universe, much due to the lack of interest and the immaturity of microwave observations.

In the sixties Bell Labs built a giant antenna in Holmdel, New Jersey, for long distance radio communication. Two employees of Bell Labs, Arno Penzias and Robert Wilson, saw the potential to use this as a radio telescope. When a satellite took over the job of the antenna, it could now be used for research. When Penzias and Wilson started to use it, they registered a faint signal, or background noise, in the microwave range that remained no matter where they turned their antenna on the sky. Numerous attempts were done to find the source of the signal. They ruled out urban interference and radiation from our own galaxy. The signal did not change with the seasons either. They even removed the pigeons living in the antenna and scrubbed it clean, but the signal remained. Penzias and Wilson were aware that a 3 K blackbody would produce

their interference, but unaware of what might be the source of this.

At the same time and just miles away, Robert Dicke at Princeton University and his postdoctoral student, P.J.E. Peebles, had been calculating the blackbody radiation left over from the Big Bang. Peebles calculated that the radiation would have a temperature of 10 K. They were interested in searching for this remnant of the early universe. When a friend told Penzias about the calculation of the radiation left over from the Big Bang by Peebles, he started to realize the significance of their discovery. Penzias called Dicke and in 1965 the pieces of the puzzle fell together. Penzias and Wilson had detected the 2.73 K blackbody radiation that fills the Universe. This relic of the Big Bang is called the Cosmic Microwave Background(CMB), and is uniform over the full sky, i.e., the temperature is the same in any direction. Dicke, Peebles, and their co-workers at Princeton wrote an article for the Astrophysical Journal Letters where they explain that the CMB strongly supports the Big Bang theory [5], while Penzias and Wilson wrote an accompanying letter where they confirmed the existence of the background noise [6]. Penzias and Wilson received the Nobel Prize for Physics for their discovery in 1978.

1.3 The search for CMB anisotropies

The measurements from Penzias and Wilson showed an approximate isotropy in the CMB temperature, just as one would expect if the radiation is produced by the Big Bang. But there are many physical effects that can cause inhomogeneities in the CMB. Inhomogeneities in the density or velocity of matter in the Universe would cause fluctuations or anisotropies in the CMB. At last-scattering, gravitational instability theory predicts that fractional density perturbations must have been $\delta \gtrsim 10^{-3}$ in order for galaxies and clusters to develop to what we observe today. A challenge in cosmology has been to detect a corresponding fluctuation in the temperature of the CMB radiation. During the sixties and seventies, experimental and theoretical work put harder constraints on the observable fluctuations in the CMB. Sunyaev predicted in 1977 that fluctuations in the CMB must exist a level $\Delta T/T_0 \sim 10^{-5}$. New experiments with higher accuracy had to be conducted. The earth's atmosphere consists of water vapor and it absorbs much of the microwave radiation. To reduce the effects of atmospheric disturbance in order to get a better signal from the CMB, new experiments were performed from balloons or in space.

1.3.1 COBE - the first fluctuations

The Cosmic Background Explorer (COBE) satellite was launched in 1989. The FIRAS instrument on board showed a perfect fit of the CMB and the theoretical curve for a black body at a temperature of 2.7K. See Figure 1.1. Despite the measurement of the CMB by Penzias and Wilson, not everybody embraced the Big Bang theory as the explanation of this. They claimed other sources could explain the radiation. But this new measurement that shows a remarkable agreement between the prediction of Big Bang theory and observations should be enough to convince most skeptics. On

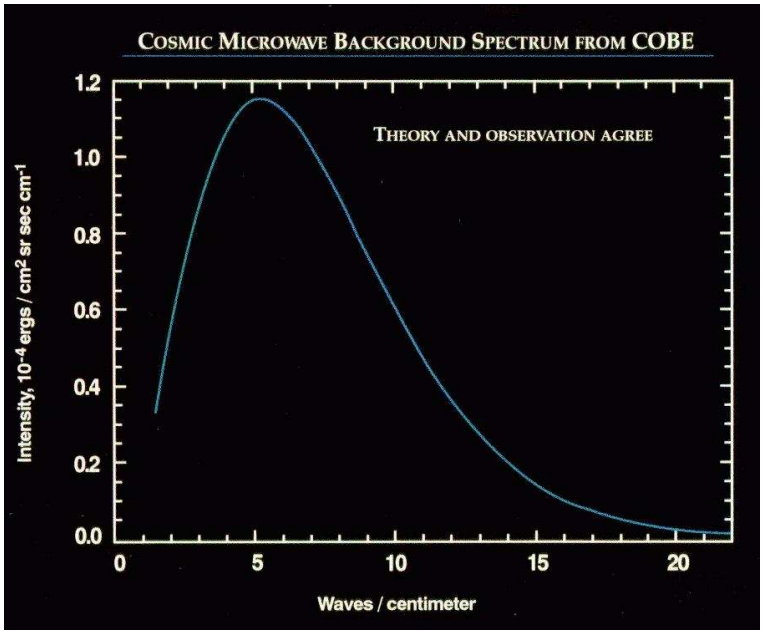


Figure 1.1: Intensity of the cosmic microwave background as a function of wavelengthas measured by the Far InfraRed Absolute Spectrophotometer (FIRAS). The full line shows the theoretical blackbody curve. Inside of this curve are dozens of measured points, and the error bars are so small that they are obscured by the thickness of the curve. Courtesy of the COBE science team.

board the satellite was also the Differential Microwave Radiometer (DMR) whose job was to measure the anisotropy in the CMB. This is very difficult to measure since it is only deviates from the average temperature as one part in 100 000. The instrument measured the temperature with a beam covering 7° of the sky. The anisotropies are showed in Figure 1.2. The two principal investigators on the DMR and FIRAS, George Smoot and John Mather, respectively, won the Nobel Price in Physics in 2006 for their work. The Nobel Price committee stated that "These measurements also marked the inception of cosmology as a precise science."

The next step was now to measure the anisotropies on smaller angels. The Toco experiment was the first to localize the first acoustic peak in the power spectrum [7].

1.3.2 BOOMERanG - a flat universe

The BOOMERanG experiment (**B**alloon **O**bservations **o**f **M**illimetric **E**xtragalactic **R**adiation **an**d **G**eophysics) was a balloon based experiment to measure the the CMB. By sending a probe 42,000 meters above mean sea level, it was possible to reduce the atmospheric absorption of microwaves and also save a lot of money compared to a satellite mission. The probe was only able to scan a small part of the sky, but it

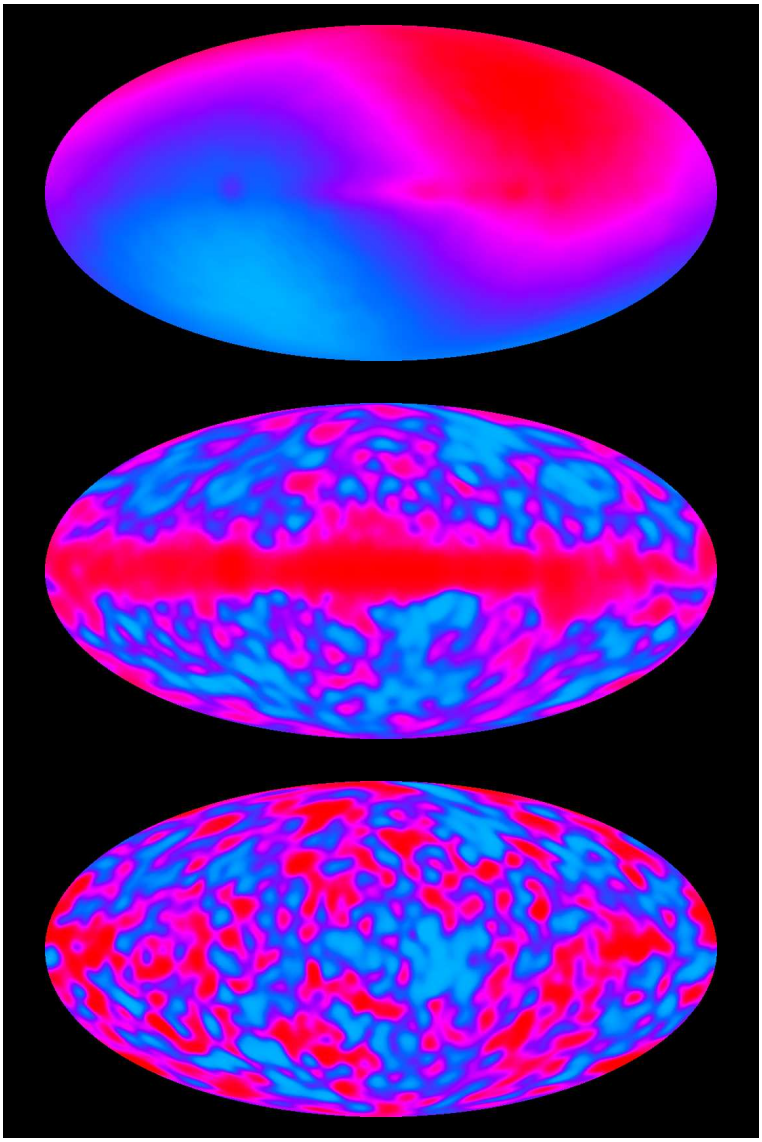


Figure 1.2: These pictures shows the anisotropy in the CMB as measured by DMR. Top: A map including the dipole and galaxy. Due to the earths movement through space relative to the CMB, the frequency of the radiation will be shifted. Moving away relative to the radiation causes a drop in temperature (redshift), while moving towards it causes a rise (blueshift). Middle: The dipole is removed. The band that stretches across the picture is radiation caused by our own Galaxy. Bottom: Galactic foreground emission is reduced and shows us the anisotropies in the CMB. Courtesy of the COBE science team.

captured large, high fidelity images of the CMB temperature anisotropies. Together with other experiments it could determine the angular diameter distance of the surface of last scattering with high precision, and it determined the geometry of the Universe to be flat. For the 2003 flight it gathered extremely high signal-to-noise ratio maps of the CMB temperature anisotropy, and a measurement of the polarization of the CMB.

1.3.3 WMAP - concordance model with slight flaws

In 2001 NASA launched the **W**ilkinson **M**icrowave **A**nisotropy **P**robe (**WMAP**). Its mission was to measure the temperature difference in the CMB radiation. These anisotropies can then be used to measure the universes geometry, contents and evolution. To do so, the instrument creates a full-sky map with a 13 arcminutes resolution of the temperature anisotropy. Compared to COBE with a resolution of 7° full-sky map, this had 33 times better resolution. WMAP was the first to localize the third peak.

The data from WMAP gave rise to a standard cosmological model: A flat universe composed of matter, baryons and dark energy with a nearly scale-invariant spectrum of primordial fluctuations. Together with other cosmological data, values of the different cosmological parameters could now be calculated. The age of the Universe is 13.72 ± 0.12 Gyr. It consists of 4.56 ± 0.15 % baryonic matter, 22.8 ± 1.3 % dark matter and 72.6 ± 1.5 % dark energy [8].

But even if the data described the universe very well, there were some strange features in the map from WMAP. (1): alignments and symmetry features among low low-l multipoles (Tegmark et al. 2003 [9]; de Oliveira-Costa et al. 2004 [10]; Eriksen et al. 2004 [11]) (2): an apparent asymmetry in the distribution of fluctuation power in two opposing hemispheres (Eriksen et al. 2004 [12]; Hansen et al. 2004 [13]) (3): a peculiar cold spot in the southern hemisphere (Vielva et al. 2004 [14]; Cruz et al. 2005 [15])

The reason for this asymmetry is not known, but three possible candidates that can cause this are systematics, foregrounds or some new, exotic physics.

1.3.4 Future missions

New experiments are planned to get better data about the cosmic microwave background radiation. The Planck satellite was launched on 14th of May 2009. This will be the third space CMB mission after COBE and WMAP. The goals of Planck are to make maps for the temperature and polarization anisotropies of the CMB with high resolution (down to 5') over the entire sky. It will take measurements in a wide range of frequencies (nine frequencies between 30 - 857 GHz) to discriminate between Galactic emission and cosmic signal and to study galaxy clusters and extragalactic point sources. By measuring the CMB on smaller scales than WMAP, it will be able to extract more information about the cosmological parameters.

It will give us a higher resolution of the CMB over the whole sky. It will detect both the intensity and polarization of photons at nine different frequencies.

Chapter 2

The physics of the CMB

In this section I will be looking at the physics in the early universe and explain how the CMB radiation originated and later evolved with the universe. I will mostly just give a short introduction to the various physics and not introduce a lot of mathematics. The main target in my thesis is not to explain how the universe started, but how to compute cosmological parameters. But it is nice to have a feeling of what we are studying. For more complete texts on this matter I refer the interested reader to Dodelson [16] and Tegmark [17]. Some of the material in this section is taken from the lecture notes for the course AST4220 by Øystein Elgarøy [18].

2.1 Inflation

The Big Bang explains many features of our universe such as the origin of light elements (Big Bang nucleosynthesis), the formation of the cosmic microwave background, and the relation between distance and redshift of cosmological objects (Hubble's law).

However there are some features that may seem very unlikely. The first is the isotropy of the CMB, which we today observe around 2.7 K on the whole sky. The natural thing to assume is that there is some physical process that smooth out any temperature variation that existed in the early universe. For this to happen, these regions must be connected by causal physics. The distance that causal physics can act is given by the particle horizon. Assume that we live in a spatially flat universe with dust and a cosmological constant. Then, when calculating the angular size of the particle horizon at last scattering as seen on the sky today, this only covers a few degrees. How is it then possible that regions on the sky today that are separated by as much as 180 degrees have almost the same temperature? We may assume that that the uniform temperature was a part of the initial conditions of the Big Bang model, but many may not be satisfied with this assumption. To explain this, cosmologists have postulated the existence of a process called inflation. Inflation is a short period of exponential growth in the early universe, i.e., the universe expands rapidly. Before inflation the regions visible to us were inside the particle horizon, thus there is no problem understanding the isotropy of the CMB.

Inflation also solves other problems in the universe. Observations tell us that the curvature of the universe is close to flat, or equivalently that the energy density in the universe is close to the critical density. This is a problem because, as time goes by, the deviation of the density from the critical one increases. Since the density is close to the critical today, it would be much closer to the critical in the past. Again, we could assume that the universe did start in this fine tuned state, but we do not get an explanation why it is so. Inflation solves this problem because during the rapid expansion of the universe, any deviation from the critical will be wiped out.

Another puzzle with the observed universe is the magnetic monopole problem. In a very hot, early universe it is predicted that a number of heavy, stable particles, such as the magnetic monopole, would be produced. The problem is that we have not observed these particles. Inflation may solve this if the rapid expansion occurred after the production of monopoles. During expansion of space the density of these particles would decrease, hence becoming very rare in the present observable universe.

The idea of inflation was proposed by Guth (1981) [19], Linde (1982) [20] and Albrecht et al. (1982) [21] to solve the horizon, flatness and monopole problems. It was soon discovered that the same mechanism that explains the uniformity of the temperature in the CMB can also explain the origin of perturbations in the universe. In its simplest form, the inflation is driven by a single scalar field. Quantum mechanics limits how homogeneous this inflationary field can be. The Heisenberg uncertainty principle for energy and time, $\Delta E \Delta t \approx \hbar$, sets a limit for how precise we can know the energy of the field in a given time interval. As a consequence of this, inflation will begin and end at different times in different regions of space and this leads to perturbations in the energy density. These perturbations will manifest themselves in the radiation as well as in the matter distribution we observe today. The simplest form of inflation theory is a single scalar field with with adiabatic fluctuations. This CMB anisotropies should then have an approximately scale-invariant spectral index of primordial fluctuations, $n \simeq 1$, and have produce density perturbations that are random-phase and have amplitudes with a Gaussian distribution.

2.2 Recombination/decoupling

In the early, hot universe matter and radiation were tightly coupled. The high energy radiation did not give protons and electrons the possibility to form a stable atom. Due to the high density of free electrons the photons are often scattered off baryons. This cause the mean free path of the photons to be short and they can not move far in the universe. This interaction between light and baryons coupled them in thermodynamic equilibrium, i.e., they shared the same temperature and they should have a blackbody spectrum.

The temperature drops as the universe expands. The binding energy of hydrogen is $B_{\mathrm{H}} = 13.6eV$, and as the temperature drops such that $k_{\mathrm{B}}T = B_{\mathrm{H}}$ one should expect the production of neutral hydrogen. But since there are many more photons than matter at this point, there is still enough high energy photons to ionize any hydrogen as the

temperature in the universe goes below 13.6 eV. As the temperature drops further free electrons can finally combine with hydrogen and helium nuclei. The formation of neutral atoms is referred to as *recombination*, despite the fact that electrons and nuclei never have been combined into atoms before. The loss of free electrons due the recombination and dilution from expansion decoupled the photons from the matter, and they could now move freely through the universe. This happened about 400 000 years after the Big Bang. We can see this radiation in all directions, and it forms a spherical shell around us with a radius of 13.7 billion light years. But this shell also has a thickness since the photons where released from the matter at different times. The thickness is about 50 million light years. This radiation is the our best source to the early universe and we call it the cosmic microwave background. At the time of decoupling the temperature of the photons was about 3000 K. Since then the photons have lost their energy due to the expansion of the universe, and today we observe it with a temperature of 2.73 K.

But the CMB is not completely isotropic. There are temperature fluctuations of the order of 10^{-5}. These have their origin from the quantum fluctuations in the scalar field responsible for inflation. These manifest themselves in the matter and photon densities, and then evolve with time. Because of the density differences at decoupling, the photons have different temperatures at different locations. The next section describe how the density varies with time after inflation has set up the original density perturbation and up to decoupling.

2.3 Acoustic oscillations in the baryon-photon fluid

Before recombination, the photons were coupled to protons and electrons, and they could be described as a single fluid. This fluid basically have two competing forces. The gravity of the baryons makes the fluid clump together, resulting in over- and under-densities. The pressure from the photons, on the other hand, washes out these differences. As the over-density in the fluid grows, the pressure also grows. When the pressure from the photons gets large enough, it can withstand the contraction and cause it to stream out of the over-density. This makes the fluid being more homogeneous. If the pressure is large enough it may cause an under-density during this dissipation. But as the pressure is reduced, gravity will again cause the fluid to clump together.

These two forces will cause the fluid to oscillate. It oscillates at different scales at different times. The small scales starts first as they are first affected by causal physics. The oscillations in the fluid are sound waves, and the sound speed in fluid, c_s, depends on the baryon density in the fluid. The sound speed determines how fast different parts in space can have causal contact with each other. The baryons makes the fluid heavier and reduces the sound speed. The maximum comoving distance traveled by a sound wave distance one can travel in the fluid by time η is given by the sound horizon,

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

where η is the conformal time. The sound horizon is the distance a sound wave could

propagate from the Big Bang to recombination, and is the largest scale on which causal physics can act in this fluid. On scales larger than this, we do not expect to see any fluctuations caused by acoustic oscillations. Instead, we only expect to see the effects that were caused by inflation. When decoupling occurs, the different scales are at different states in this oscillation. After decoupling the photons and baryons are no longer a single fluid, and the oscillation stops. This leaves us with various densities in space on different scales.

These various densities are the origin of the anisotropies in the CMB radiation. In the remainder of this chapter I will look at the physical effects that caused the fluctuations in the Cosmic Microwave Backgound radiation. The first part looks at the changes to the CMB radiation that was sent out at decoupling. We may divide them into two groups, the primary and secondary fluctuations. The primary fluctuations were generated before the universe became transparent. The secondary fluctuations are the changes in frequency, apart from the change caused by universal expansion, while the radiation moved from the surface of last scattering and to us. The CMB radiation is basically a snapshot of the universe when it is around 400 000 years old. This gives us a unique chance to observe the early universe. The anisotropies in the CMB remain small because the photons do not clump, so the distribution looks quite the same today as when the photons were sent out. There has only been linear changes to the photons.

There are also other sources causing microwave radiation other than the CMB. These will contaminate the "pure" signal sent out. All of these effects are summarized in Table 2.1

2.4 Primary fluctuations

In this section I will look at the effects that cause anisotropies in the CMB during recombination.

2.4.1 Sachs-Wolfe effect

First is the Sachs-Wolfe effect. If the radiation has to climb out of the gravitational potential it is thereby redshifted. This has two effects: i) Photons loose energy as they move out of the gravitational field.

$$\left(\frac{\Delta T}{T}\right)_I = -\frac{\Delta\phi}{c^2},$$

where $\Delta\phi$ is the difference in gravitational potential at the emitter position and at the observer position; ii) Because of the gravitational time dilatation time proceeds at a slower rate down in the gravitational field. We then seem to be looking at a younger, and hence hotter, universe where there is an overdensity.

$$\left(\frac{\Delta T}{T}\right)_{II} = \frac{2}{3(1+\omega)}\frac{\Delta\phi}{c^2},$$

Primary	Sachs-Wolfe (Gravity)	
	Doppler	
	Density fluctuations	
	Damping	
Secondary	Gravity	Early ISW
		Late ISW
		Rees-Sciama
		Lensing
	Local reionization	Thermal SZ
		Kinematic SZ
	Global reionization	Suppression
		New Doppler
		Vishniac
Other sources	Extragalaxtic	Radio point sources
		IR point sources
	Galactic	Dust
		Free-Free
		Synchrotron
	Local	Solar system
		Atmosphere
		Noise, etc.

Table 2.1: Sources of temperature fluctuations in the CMB. The two topmost categories shows what effects that change the CMB directly. The last category shows physical processes that produce microwave radiation that contaminate the CMB signal.

Here ω is the equation of state factor that determine the relation between pressure p and matter/energy density ρ. For non-relativistic matter, also just called dust, this constant is equal to 0. In this case the effect becomes

$$\left(\frac{\Delta T}{T}\right)_{II} = \frac{2}{3}\frac{\Delta\phi}{c^2},$$

The total of this then gives

$$\left(\frac{\Delta T}{T}\right)_{SW} = -\frac{1}{3}\frac{\Delta\phi}{c^2}.$$

2.4.2 Doppler effect

At last scattering, the fluid moves in a random direction with a velocity, $\mathbf{v}$, relative to us. If the fluid in a direction $\hat{n}$ moves towards us then the photons sent out will be blueshifted, while if the fluid moves away from us, the photons will be redshifted.

$$\left(\frac{\Delta T}{T}\right)_{D} = \frac{\mathbf{v}\cdot\mathbf{n}}{c}.$$

2.4.3 Adiabatic fluctuations

In a region with higher matter density the photon density will also be higher. For an adiabatic density fluctuation the photon density fluctuation is related to the matter density fluctuation as

$$\left(\frac{\Delta\rho}{\rho}\right)_\gamma = \frac{4}{3}\left(\frac{\Delta\rho}{\rho}\right)_m$$

The density is proportional to temperature, $n_\gamma \propto T^3$, hence a region with increased photon density will have a higher temperature. Last scattering happens at roughly the same temperature. The temperature decrease as the universe expands, so we might expect that the fluctuations in photon temperature can not be observed as they all were sent out with the same temperature. But there will be an observable difference due to the fact that regions that originally had different temperature will send out the photons at different times. Overdense regions with higher temperature will scatter off photons at later times than the regions with a lower temperature. This will cause radiation from overdense regions to have a lower redshift from universal expansion. The temperature difference caused by this effect is

$$\left(\frac{\Delta T}{T}\right)_A = -\frac{\Delta z}{1+z} = \frac{\Delta\rho}{\rho},$$

where z is the redshift and the last equality assumes linear growth, $\Delta\rho \sim (1+z)^{-1}$.

Hence the total effect of these three become

$$\left(\frac{\Delta T}{T}\right)_{tot} = -\frac{1}{3}\frac{\Delta\phi}{c^2} + \frac{\mathbf{v}\cdot\mathbf{n}}{c} - \frac{\Delta\rho}{\rho}$$

2.4.4 Damping by photon diffusion

We treat the baryons and photons as a single fluid, but this is just an approximation as the photons will travel a finite length between scatters. This way photons can travel between the hot and cold regions, and any perturbation on scales smaller than this length can be expected to be washed out. This effect is named Silk damping after Joseph Silk who first discussed the effect in 1968 [22]. It is also known as diffusion damping.

2.5 Secondary fluctuations

Next we got the effects that altered the CMB radiation after decoupling when traveling to us today.

2.5.1 Change in gravitational potential

The first effect is caused by gravity. Even if the photons now are decoupled from the baryons, they are still affected by the gravitational potential of matter.

This effect can be split into three parts, all due to time-variations in the gravitational potential. I.e., $\dot{\phi} \neq 0$ where $\dot{\phi}$ denotes the conformal time derivative of the gravitational potential. When a photon flies through a potential well it is blueshifted when falling in to it, and it is then redshifted when climbing out again. As long as the potential is the same on entrance and exit the net shift in frequency of the photon is zero. But if the potential changes while the photon travels through it, the frequency of the photon will change. During the matter-dominated era the potentials remains the same. But when other components dominate the universe these potentials may change.

At the time of last scattering the photon-contribution to the density of the universe is significant. This makes the potentials decay, causing the early ISW effect.

At recent times the universe seems to be dominated by vacuum energy which causes the universe's expansion to accelerate. This vacuum energy also causes the potentials to vary with time, and the radiation's energy will change. This is called the late ISW effect.

When non-linear structures (such as galaxy clusters) form, linear perturbation theory breaks down and the perturbation theory result $\dot{\phi} = 0$ is no longer valid. This is called the Rees-Sciama effect.

The other effect caused by gravity change the direction of the photons and not the energy. If a pair of photons would have arrived separated by an angle θ in the absence of fluctuations in ϕ, they will in reality arrive separated by some angle $\theta + \Delta\theta$. This change in the photons trajectory is called weak gravitational lensing. The total power of the in the fluctuations are conserved, but the power is redistributed from the peaks to the troughs.

2.5.2 The Sunyaev-Zel'dovich effect

The photons can be directly influenced by baryonic matter as they free-stream towards us. Baryons may be reionized after recombination locally e.g., in hot clusters of galaxies or globally throughout all of space.

Local reionization is called the Sunyaev-Zel'dovich (SZ) effect and is caused by two effects: i) If a cluster of galaxies is moving towards us, Thomson scattering of the CMB photons off free electrons in the hot gas in the galaxies will cause a Doppler blueshift in the direction of the cluster. This is known as the kinematic SZ effect. ii) The high energy of the free electrons will change the Planck spectrum by reducing the Rayleigh-Jeans (low ν) tail and increase the Wien (high ν) tail. This happens independently of the clusters velocity and is known as the thermal SZ-effect.

2.5.3 Global reionization

If global reionization happens throughout space it will suppress fluctuations on small scales in the CMB power spectrum. This reionization causes the photons to Thomson scatter off free electrons, thus changing the original direction of the photon. Then we don not know exactly which direction the original photon came from. It could originally

have come from anywhere within a certain region on the sky, and the signal is therefore smeared out at these scales.

2.6 Other sources to microwave radiation

There are of course other sources for radiation at microwave wavelength other than the radiation from the last scattering. This will contaminate the signal from the cosmic microwave background. All sources of microwave radiation that might come between us and the radiation left over from the Big Bang are called foregrounds. But the problems with foregrounds are manageable. The various effects causing the radiation is frequency dependent, and this is something we may take advantage of if we observe the sky in many different frequencies (see section 7.1).

2.6.1 Extragalactic radiation

Extragalactic point sources comes from outside our own galaxy. We distinguish between radio and infra-red sources. To get rid of this radiation we may subtract it from known point source catalogues or we may throw away the pixels containing a bright point source.

2.6.2 Galactic radiation

There are at least three sources for radiation from our own Galaxy: dust, free-free emission and synchrotron radiation. The amplitude of each component varies across the sky, but the relative amplitudes are quite typical. The effects from dust is caused by two processes, namely thermal radiation.and radiation from charged spinning dust grains.

Free-free emission (Bremsstrahlung) is produced when a charged particle, e.g., an electron, is passes near another charged particle, such as an atomic nucleus. The electron loses energy as it passes near the ion and the energy difference due to this acceleration is sent out as electromagnetic radiation. Synchrotron radiation is generated when ultrarelativistic, charged particles are accelerated through magnetic fields.

The amplitudes of the of these foregrounds have proven to be smaller than the cosmic signal in a wide part of the frequency range. At high frequencies, dust dominates the signal, and at low frequencies synchrotron and free-free radiation become dominant.But in the range from 30 to 200 GHz, the CMB signal often has the largest intensity.

2.6.3 Local radiation

Contamination may also come from our own solar system. Radiation can come from the sun, moon, earth and other planets. The instrument we are using to register the radiation may have some electronic receiver noise. Ground- and balloon-experiments

will be affected by radiation in the atmosphere. There are of course many potential sources of systematic errors.

COBE is a real life example where local interference can false data. For about two months each year the satellite went behind the shadow of the Earth. This caused the signal to be non-Gaussian. It took quite some time to discover, but after removing the data from the two months when the Earth obscured the satellite, the non-Gaussian behaviour was gone [23].

Chapter 3

Mathematical description of the CMB

The physics described in Chapter 2 comes from a thorough mathematical description of the evolution of the perturbations in the universe from the time of inflation to today. These can be described using the Boltzmann equations and the theory of general relativity. The deduction of the equations describing be perturbations as the universe evolve is a demanding task, and is outside the scope of this thesis. For great texts on this subject I recommend Dodelson [16] and Tegmark [17]. Instead, I will in this chapter go through how the we describe the CMB signal we are receiving and how we handle it in order to do calculations with them.

3.1 Data model

First we need a model of the observed signal. The signal $\mathbf{d}(\hat{n})$ we receive from a direction $\hat{n}$ on the sky can be written as

$$\mathbf{d}(\hat{\mathbf{n}}) = \mathbf{s}(\hat{\mathbf{n}}) + \mathbf{n}(\hat{\mathbf{n}}) + \mathbf{f}(\hat{\mathbf{n}}) \tag{3.1}$$

Here $\mathbf{s}$ is the clean signal from the CMB with power spectrum C_l, $\mathbf{n}$ is the instrumental noise and $\mathbf{f}$ is the sum of foreground contributions. The direction may also be written in polar coordinates θ and ϕ. Many instruments, like COBE and WMAP, do not measure the absolute temperature of the CMB radiation, but the temperature difference in two different directions. Hence the signal is given as $\mathbf{s}(\hat{n}) = \Delta T(\hat{n}) = T(\hat{n}) - T_0$, where T_0 is the mean temperature of the CMB radiation, 2.73 K.

However the instruments that register the CMB signal do not just observe a single point on the sky, but a finite solid angle. This smears out the signal and the small scales get filtered away;

$$\mathbf{d}(\hat{\mathbf{n}}) = \mathbf{A}\left(\mathbf{s}(\hat{\mathbf{n}}) + \mathbf{f}(\hat{\mathbf{n}})\right) + \mathbf{n}(\hat{\mathbf{n}}), \tag{3.2}$$

where $\mathbf{A}$ denotes convolution by an instrumental beam. This process can be described mathematically as a convolution in pixel space or a multiplication in harmonic space.

There are several advantages working in harmonic space, so we will introduce spherical harmonics.

In normal flat space we can expand a function into wave-functions through a Fourier transform. This makes it possible to decompose the function into different scales, or modes. The function we are interested in is the CMB temperature field we observe on the sky. The desired wave-functions are given by Laplace's equation $\Delta^2\psi = 0$. As we observe the temperature field on a sphere, we must solve this equation in spherical coordinates. The solution to this equation is

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

for $\ell \geq 0$ and $-\ell \leq m \leq \ell$. The $Y_{\ell m}(\theta,\phi)$'s are called spherical harmonics and are eigenfunctions of the angular part of the Laplacian. and the P_ℓ^m's are the Legendre polynomials.

The temperature field can be expanded in terms of the spherical harmonics

$$\Delta T(\hat{n}) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell m}Y_{\ell m}(\hat{n}) \tag{3.3}$$

The ℓ and m are conjugate to the real direction we are observing, $\hat{n}$. This decomposition is completely analogous to a Fourier decomposition in flat space. The $a_{\ell m}$s contain all the information about the temperature field. These functions are orthogonal, with normalization

$$\int d\Omega Y_{\ell m}\left(\hat{p}\right)Y^*_{\ell' m'}\left(\hat{p}\right) = \delta_{\ell\ell'}\delta_{mm'}$$

Multiplying both sides of Equation 3.3 by $Y^*_{\ell' m'}\left(\hat{p}\right)$ and integrating gives

$$a_{\ell m} = \int d\Omega Y^*_{\ell m}\left(\hat{p}\right)\Delta T$$

The different modes are referred to as multipoles. For a value ℓ, the typical scale for a spot for is $\theta \sim 180°/\ell$. For $\ell = 0, 1, 2, 3$ and 4 the modes are referred to as the monopole, dipole, quadrupole, octopole and hexadecapole. $\ell = 0$ picks up the constant part, $\ell = 1$ picks up the linear part, $\ell = 2$ picks up the quadratic part $\ell = 3$ picks up the cubic part, etc. This way we decompose a temperature map into the different ℓ-modes that corresponds to different scales. This decomposition is shown in Figure 3.1, where all the different "ℓ-maps" sums up to produce the temperature map.

With the convolution by the instrumental beam, the CMB signal now becomes

$$A\Delta T(\hat{n}) = \sum_{\ell=0}^{\ell_{\max}}\sum_{m=-\ell}^{\ell} b_\ell a_{\ell m}Y_{\ell m}(\hat{n}),\ \text{beam convolved map} \tag{3.4}$$

Here, b_ℓ is the Legendre transformation of the experimental beam. Since small scales are being smudged out, we can only observe the scales up to some $\ell_{\max}$. Smaller scales are simply not resolved by the instrument. How to choose this value will be determined later.

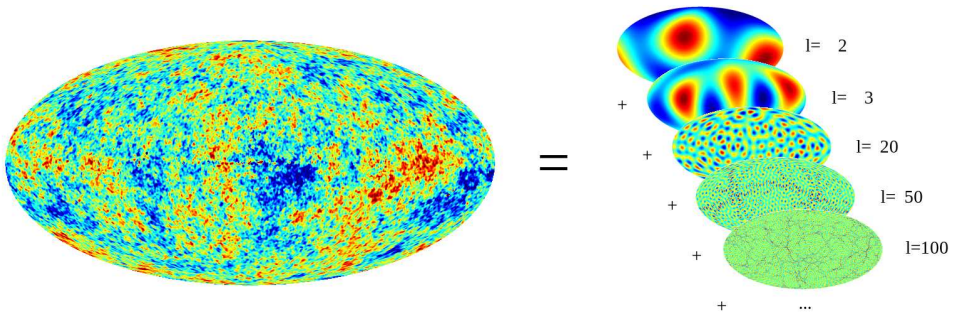


Figure 3.1: Decomposing the temperature map using spherical harmonics. Each set of ℓ corresponds to a certain scale on the sky.

3.2 The power spectrum

We need a way to quantify and describe the observed data. If we got a set of data, what could it tell us? We need to be able to characterize the distribution so that theory can be compared to experiments. It is one thing is to look at a map of the distribution of matter or CMB, and another thing to see what these maps quantitatively can tell us about cosmological models. We are only interested in the amplitude of the fluctuations for the different scales. Under the assumption of statistical isotropy and homogeneity, the specific positions of the maximum or minimum are irrelevant.

Maybe the most important statistics for the CMB and the large-scale distribution of matter is the *two-point correlation function*, called the *power spectrum* in Fourier space. In the case with distribution of mass, the mean density of the galaxies is $\bar{n}$, and the inhomogeneities in the distribution at a position $\vec{x}$ is given by $\delta(\vec{x}) = (n(\vec{x}) - \bar{n})/\bar{n}$. If we work in Fourier space it is easier to separate large scales from small scales. The Fourier transform of the inhomogeneity is $\tilde{\delta}(\vec{k})$, where k denotes the scales. The power spectrum $P(k)$ is defined via the two-point correlation function

$$\langle \tilde{\delta}(\vec{k})\tilde{\delta}(\vec{k}')\rangle = (2\pi)^3 P(k)\delta^3(\vec{k} - \vec{k}') \tag{3.5}$$

The angular brackets denote an average over the whole distribution and $\delta^3()$ is the Dirac delta function which constrains $\vec{k} = \vec{k}'$. The power spectrum tells us about the spread, or variance, in the distribution. If there is a lot of over- and under-dense regions, then the power spectrum will be large. If the amplitude is small, then the distribution of matter is smooth.

The two-point correlation function is also the best measure of the anisotropies in the CMB. Instead of Fourier transforming the CMB temperature, it is better to expand it in spherical harmonics, since the temperature is a two-dimensional field defined on the sky. The temperature field expanded in spherical harmonics is given in Equation 3.3.

The $a_{\ell m}$s are drawn from a distribution which traces its origin to the quantum fluctuations first laid down during inflation. This distribution is said to be Gaussian.

The mean value for the $a_{\ell m}$s are zero, $\langle a_{\ell m} \rangle = 0$, but their variance is not zero. The $a_{\ell m}$s are Gaussian random variables, and if there is statistical isotropy, then they are statistically independent for different ℓ and m. We can then write $\langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{\ell\ell'}\delta_{mm'}C_\ell$ where C_ℓ is the variance of the $a_{\ell m}$s. It is important to note that for each ℓ the $a_{\ell m}$s have the same variance, hence we only write the variance as C_ℓ.

For each ℓ we get $(2\ell + 1)$ different $a_{\ell m}$s. For low values of ℓ we are only sampling a few modes. For $\ell = 2$ we only measure five components, and we do not get much information about the variance for C_2. Therefore is a fundamental uncertainty in the knowledge we may get about the C_ℓs. This uncertainty is called the cosmic variance.

$$\left(\frac{\Delta C_\ell}{C_\ell}\right)_{\text{cosmic variance}} = \sqrt{\frac{2}{2\ell + 1}} \tag{3.6}$$

This is most pronounced at low ℓ, while for higher values of ℓ we can sample more modes and the variance decreases. This effect is indicated by the blue band present in Figure 3.2.

We are not interested in the details in the observed CMB, but in its statistical properties. The quantum fluctuations from inflation were random, so our universe is just one realization of a stochastic ensemble of universes. Just as we know the statistical properties of throwing two dice, we will get different outcomes in different series of throwing them. So although the outcomes looks different, they have the same statistical properties.

The angular power spectrum (or anisotropy spectrum) is defined by,

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} a_{\ell m} a^*_{\ell m} = \left\langle |a_{\ell m}|^2 \right\rangle \tag{3.7}$$

Here l is the multipole number and is related to the angular extension on the sky. It is given by $\theta \simeq \pi/\ell$ radians $= 180^\circ/\ell$. This tells us how strong variations there are on various scales. This is a very important function for cosmologists because the cosmological parameters alters the shape of the power spectrum. The parameters change the location, height and width of the peaks in different ways. So by measuring the power spectrum accurately, we can estimate the values of the different parameters. Figure 3.2 shows the power spectrum from the WMAP five year data.

3.3 Likelihood

Inflation theory predicts that the signal we observe on the sky is drawn from a Gaussian distribution. The probability distribution for a multivariate normal distribution is

$$p(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \frac{1}{\sqrt{|\mathbf{C}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{x}-\boldsymbol{\mu})} \tag{3.8}$$

Here, $\mathbf{x}$ is the data from the observed sky consisting of n elements and $\mathbf{C}$ is the covariance matrix. We call this distribution the likelihood and from now write it as $\mathcal{L}(\mathbf{x})$. We

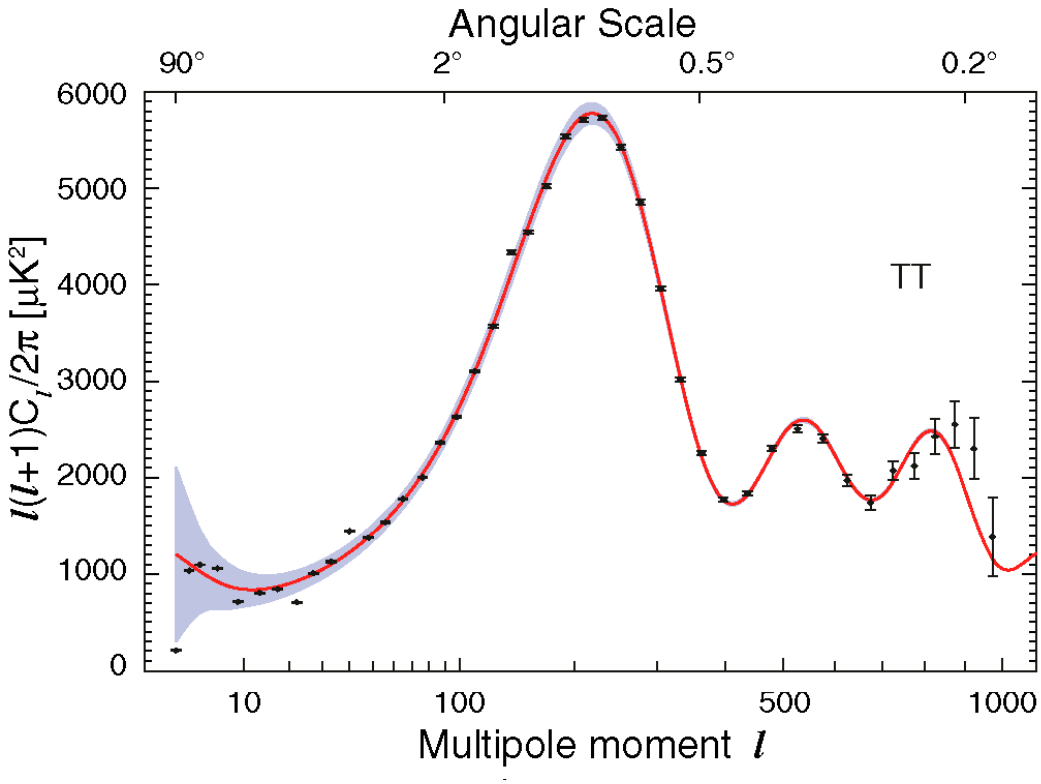


Figure 3.2: The temperature power spectrum with points from WMAP five year data. The full line is the best fit for the ΛCDM model and the shaded region around the curve shows the uncertainty due to cosmic variance. The horizontal axes show how the angles on the sky (upper axis) corresponds to the multipoles ℓ (lower). Courtesy of the WMAP science team.

will study the centered distribution, i.e., $\mu = 0$. Also, it is standard practice to work with the logarithm of the likelihood, as this gives more manageable numbers. Using thousand pixels gives us a log likelihood of the order of 1000, while the likelihood is e^{1000}. The large number in natural units may cause some numerical errors, hence it is better to work with the logarithm of the likelihood. Taking the logarithm of the above expression gives

$$-2\ln\mathcal{L}(\mathbf{x}) = n\ln 2\pi + \ln|\mathbf{C}| + \mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} \tag{3.9}$$

In this analysis I will be dropping the first term, as it is just a normalizing constant, and it does not change where the parameters gives the highest likelihood.

3.4 Signal covariance matrix

To calculate the likelihood we must first find the covariance matrix for the data. For a vector

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

the covariance matrix is defined as

$$C_{ij} = cov(X_i, X_j) \tag{3.10}$$
$$= E[(X_i - \mu_i)(X_j - \mu_j)] \tag{3.11}$$
$$= \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle \tag{3.12}$$

The last equality is just a different type of nomenclature. Both $E[]$ and $\langle\rangle$ denotes the expectation value of the given element.

Adapting the data model in Equation 3.1 the covariance matrix of the observed signal $\mathbf{d}$ is given by $\mathbf{C} = \langle (d_i - \langle d_i \rangle)(d_j - \langle d_j \rangle) \rangle = \langle (\mathbf{d} - \langle \mathbf{d} \rangle) \left(\mathbf{d}^t - \langle \mathbf{d} \rangle^t \right) \rangle = \langle (\mathbf{s} + \mathbf{n} + \mathbf{f} - \langle \mathbf{s} + \mathbf{n} + \mathbf{f} \rangle) \left((\mathbf{s} + \mathbf{n} + \mathbf{f})^t - \langle \mathbf{s} + \mathbf{n} + \mathbf{f} \rangle^t \right) \rangle$. The expectation value of the three components are zero, $\langle s \rangle = \langle n \rangle = \langle f \rangle = 0$. We assume that none of the three components are internally correlated, so the cross terms all are zero, $\langle sn^t \rangle = \langle sf^t \rangle = \langle nf^t \rangle = 0$. The covariance matrix becomes

$$\mathbf{C} = \langle \mathbf{s}\mathbf{s}^t \rangle + \langle \mathbf{n}\mathbf{n}^t \rangle + \langle \mathbf{f}\mathbf{f}^t \rangle \equiv \mathbf{S} + \mathbf{N} + \mathbf{F} \tag{3.13}$$

Here $\mathbf{S}$ is the covariance matrix of the clean CMB radiation, $\mathbf{N}$ is the covariance matrix of the noise from the instrument, and $\mathbf{F}$ is the covariance matrix of the foregrounds.

We utilize the fact that we are observing the signal on a sphere and expand the signal in terms of spherical harmonics, as described in Equation 3.3. Inserting the expression for the spherical harmonics into the CMB covariance matrix gives us

$$S_{ij} = \langle s(\hat{n_1}) s^*(\hat{n_2}) \rangle = \left\langle \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n_1}) \sum_{\ell'=0}^{\infty} \sum_{m'=-\ell'}^{\ell'} a^*_{\ell' m'} Y^*_{\ell' m'}(\hat{n_2}) \right\rangle$$

The spherical harmonics are just constants and we may pull them out of the averaging brackets,

$$S_{ij} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \sum_{\ell'=0}^{\infty} \sum_{m'=-\ell'}^{\ell'} Y_{\ell m}(\hat{n_1}) Y^*_{\ell' m'}(\hat{n_2}) \langle a_{\ell m} a^*_{\ell' m'} \rangle$$

We assume that the CMB field is Gaussian and isotropic, hence we do not depend on the actual directions. This means that $\ell = \ell'$ and $m = m'$, or $\langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell$. The Legendre polynomial P_ℓ may be written in terms of a sum of products of the spherical harmonics $P_\ell(\hat{x} \cdot \hat{x}') = \frac{4\pi}{2\ell+1} \sum_{m=-l}^{\ell} Y_{\ell m}(\hat{x}) Y_{\ell m}(\hat{x}')$, and the signal covariance matrix therefore reads,

$$S_{ij} = \sum_{\ell=0}^{\infty} \sum_{m=-l}^{\ell} Y_{\ell m} Y^*_{\ell m} C_\ell = \frac{1}{4\pi} \sum_{l=0}^{\infty} (2\ell + 1) C_\ell P_\ell(\cos\theta_{ij}), \tag{3.14}$$

where $\hat{n_1} \cdot \hat{n_1} = \cos\theta_{ij}$. Here we choose to parameterize the power spectrum with a free amplitude q and tilt n like

$$C_\ell = q \left(\frac{\ell}{\ell_0} \right)^n C_\ell^{\rm fid}, \tag{3.15}$$

where $C_\ell^{\rm fid}$ is a fiducial model and ℓ_0 is a pivot multipole. For a beam convolved map, the signal covariance matrix becomes

$$S_{ij} = \frac{1}{4\pi} \sum_{l=0}^{\ell_{\rm max}} (2\ell + 1) \left(b_\ell p_\ell\right)^2 C_\ell P_\ell(\cos\theta_{ij}), \tag{3.16}$$

where p_ℓ is the effect of finite pixelization and principally acts the same way as the beam function b_ℓ. Due to the beams finite resolution, the instrument can only resolve scales up to a multipole $\ell_{\rm max}$. The value of this will be discussed in Section 3.5.1.

We assume that the noise is Gaussian and uncorrelated from pixel to pixel, with a standard-deviation of σ_p. The covariance matrix then becomes

$$N_{ij} = \langle n_i n_j \rangle = \delta_{ij} \sigma_i^2, \tag{3.17}$$

where i and j are two pixel indices. Thus the noise covariance matrix is diagonal, with variances on the diagonal.

We also need a term that could remove unwanted effects. In practice this is done by adding an extra term in the covariance matrix. If we want to be insensitive to a signal $f(\hat{n})$ then we want it to have zero statistical weight. To do this we say that the uncertainty of this component is "infinitely" large. In practice it is done by adding the following term to the covariance matrix,

$$\mathbf{F} = \lambda \mathbf{f} \mathbf{f}^t \tag{3.18}$$

Here, λ is a large constant and f is a signal on the sky, also called template, we wish to be insensitive to. $\mathbf{ff^t}$ is the outer product of the template. More detailed information on the various signals will given in the WMAP-data section 7.1.

3.5 Technical considerations

3.5.1 Determining resolution parameters

There are two things we must think about when choosing $\ell_{\rm max}$ and the pixelization to represent the observed data. First of all we know that the resolution of the instrument beam determines what scales are measurable. $\ell_{\rm max}$ must therefore be chosen large enough so that there is only negligible power beyond this multipole.

A Gaussian beam is a good approximation to the instruments that carry out the CMB experiments. The total width of the beam profile extends forever, so as a measure of the beam size, the Full Width Half Maximum (FWHM), $\theta_{\rm FWHM}$, of the beam is used. As the name suggests, FWHM is the width of the beam pattern where the beam drops

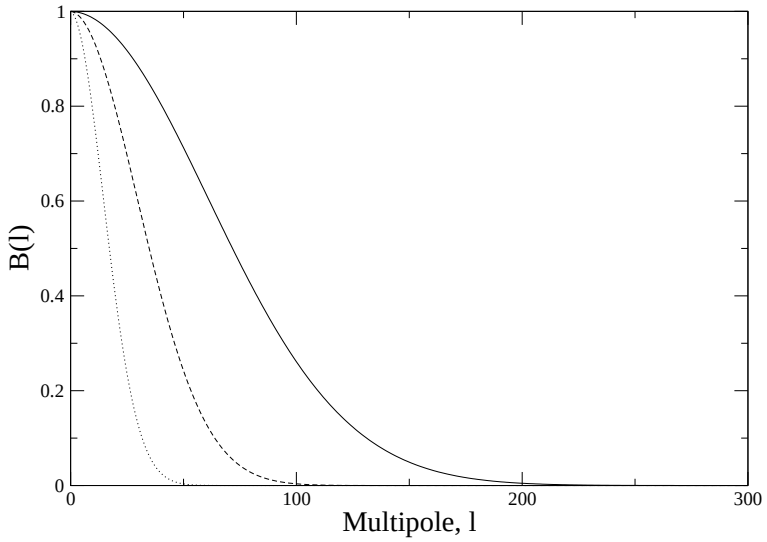


Figure 3.3: Show three different Gaussian beams. The full line shows a beam with $\theta_{\rm FWHM} = 2.20$ which corresponds to a resolution supported by $N_{\rm side} = 64$. Using this beam, the typical scales observed is $\ell_{\rm FWHM} \sim 80$, but it registers signal for scales much smaller than this. We can see that we can choose $\ell_{\rm max}$ considerably larger, around 130 for $\ell_{\rm max} = 2N_{\rm side}$ or 190 for $\ell_{\rm max} = 3N_{\rm side}$. The dashed line is a beam with $\theta_{\rm FWHM} = 4.50$ and is supported by resolution at $N_{\rm side} = 32$. The dotted line shows a beam with $\theta_{\rm FWHM} = 9.00$ which corresponds to a resolution supported by $N_{\rm side} = 16$.

to half of its maximum. This gives us a number on the scales that are resolved by the instrument.

The angular extension on the sky is related to the multipole number as $\theta \simeq \pi/\ell$. So for the FWHM, the typical scales the beam can resolve is $\ell_{\rm FWHM} \simeq \pi/\theta_{\rm FWHM}$. But the FWHM is just a measure of the typical scales for the beam, and the beam does resolve scales that are smaller than this. The Fourier transform of the beam profile is also a Gaussian, given as,

$$B(\ell) = e^{-\ell(\ell+1)\sigma^2/2} \tag{3.19}$$

where $\sigma = \theta_{\rm FWHM}/(\sqrt{8\ln(2)})$. Using this formula, I have plotted five different beam functions in Figure 3.3. Here we can see how the beam falls off with ℓ. For high enough ℓ the beam suppresses the signal, and can not register the fluctuations on these small scales. Or put another way, the beam can not resolve these small fluctuations, and they are just smeared out. This gives us the largest multipole $\ell_{\rm max}$. We must choose $\ell_{\rm max}$ such that there is negligible power beyond this multipole. For the case of the smallest beam, indicated in red, the typical scales are $\ell_{\rm FWHM} \sim 80$. We can see that the beam does have much power above this multipole, and has almost gone down to zero around $\ell \sim 200$. This may seem like a good value for choosing $\ell_{\rm max}$.

But there is a second aspect we must think about. We can not choose $\ell_{\rm max}$ to be too

large as aliasing can become a problem when we represent the data on the computer. It may seem reasonable to make the pixel size the same size as the smallest scales that the instrument can resolve. as this should seem sufficient to represent the smallest scales. However this may cause some trouble. This has to do with the sampling rate of the signal, just as with digitalized music, where we must sample the music signal often enough to get good quality. From signal processing, the Nyquist theorem tell us that we must sample the signal well enough to represent it without losing much information. More specifically, the Nyquist rate , $\nu_{\rm N}$, at which we must sample a signal must be at least twice the largest frequency, $\nu_{\rm max}$, in the signal in order to avoid aliasing effects. The largest frequency is also called the bandwidth. Thus the sample rate is $\nu_{\rm N} = 2\nu_{\rm max}$. This analysis is true for flat space, while we on the other hand are working with functions on a sphere. So a good rule-of-thumb for our case is that the beams FWHM must be at least 2.5 times the pixel size,

$$\theta_{FWHM} = 2.5\theta_{pix} \tag{3.20}$$

HEALPix has $N_{pix} = 12N_{\rm side}^2$ pixels of equal area. The solid angle over the sphere is 4π sr. The area for each pixel then becomes $\Omega_{pix} = \frac{4\pi}{N_{pix}} = \frac{\pi}{3N_{side}^2}$. The angular resolution of each pixel is $\theta_{pix} \equiv \sqrt{\Omega_{pix}} = \sqrt{\pi/3}/N_{\rm side}$. By inserting this into Equation 3.20, we can now calculate what resolution value $N_{\rm side}$ that supports our beam. With the above equations we can also express out choice of $\ell_{\rm max}$ by $N_{\rm side}$ by choosing a relation between $\ell_{\rm FWHM}$ and $\ell_{\rm max}$. The recommended upper limit is $\ell_{\rm max} = 2N_{\rm side}$, but through various numerical techniques we can be pushed to $\ell_{\rm max} = 3N_{\rm side}$. If we set it lager than this we are likely to get nonsensical results, as aliasing will become a problem.

HEALPix has a lot of advantages in its choice of pixelization to make it fast and efficient. Further information on HEALPix is found in Appendix A.

3.5.2 Downgrading

Sometimes we can not work all the data that an experiment gives us. The data may demand too much computer resources and the calculations may take too long time. E.g., computing the likelihood, Equation 3.9, requires us to invert and calculate the determinant of the covariance matrix. This operation scales as $\mathcal{O}(N_{pix}^3)$.

HEALPix consists of twelve base-pixels. These pixels can be divided into smaller pixels. $N_{\rm side}$ is the resolution parameter and denotes the number of pixels along the edge of a base-pixel. The total number of pixels in the map is $N_{\rm pix} = 12N_{\rm side}^2$. This way we can go between different resolutions for a given data set. When we downscale our map to a lower resolution, the pixel sizes will also become larger, and we will not be able to represent as small scales as we did before. We therefore need to make sure that the experimental beam suppresses the signal for the small scales that will not be supported by this resolution.

First we choose a pixelization that makes it possible for us to calculate the desired value, here the likelihood, within a reasonable amount of time. To go from the original

resolution to another we must first deconvolve the data as described in Equation 3.4. This is simply done by division of the beam function in harmonic space, so it is best to work with spherical harmonics. The original beam describing the instrument is usually given by the experiment team. Then we must convolve with a new beam function that corresponds to the new resolution. At a lower resolution we can not resolve scales smaller than $\ell_{\max}$, thus we need an experimental beam suppresses all signal beyond this value. Most experimental beams act like a Gaussian beam, so therefore we approximate our new beam to a Gaussian beam.

The steps to downscale a temperature map are as follows:

1. First we transform the temperature field to spherical harmonics.
2. Then we deconvolve the beam and pixel window for the original resolution. In harmonic space this is just a division.
3. Then we must convolve with the beam and pixel window corresponding to the desired lower resolution.
4. Then we transform back to the temperature.

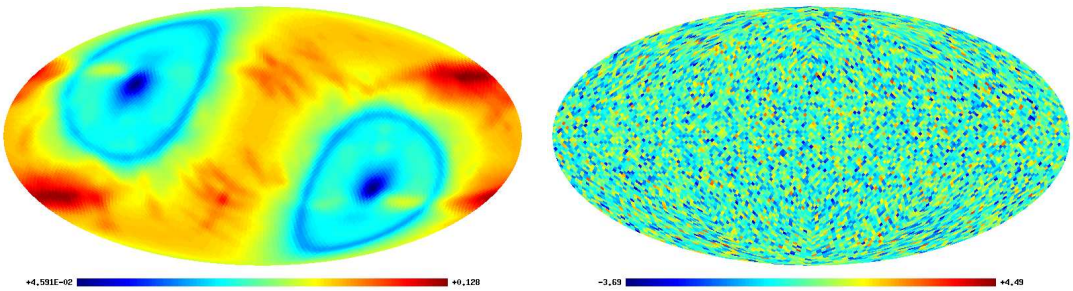
This gives us a temperature map that is connected to the physical scales we could observe using an instrument with a larger beam.

3.5.3 Covariance matrix regularization by noise

The pixel noise rms can be evaluated from the expression $\sigma_i = \sigma_0/\sqrt{N_{\text{obs,i}}}$, where $N_{\text{obs,i}}$ is the number of times the pixel i is observed and σ_0 is the overall noise level per observation. For current CMB experiments the signal is strongly sub-dominant to the signal term on large angular scales. The signal-to-noise ratio of the WMAP is high at large angular scales, about 150 for the V-band at $\ell = 100$, and we can therefore neglect it when working on large scales. However, the noise can be of use for us as it may be used to regularize the total covariance matrix.

We can approximate the noise covariance matrix. Lets assume that the noise is Gaussian and uncorrelated from pixel to pixel, with a standard deviation σ_n The covariance matrix then become $N_{ij} = \sigma_n^2 \delta_{ij}$. In Eriksen et al. 2007 [24],they also computed the noise covariance from the smoothed instrumental noise variance for the V-band data, but they found out that it had no effect on the results, since its amplitude is much lower than the CMB signal. I will therefore only use the approximated noise. The noise for the two cases can be seen in Figure 3.4. The shape of Figure 3.4(a) comes from the scanning pattern the instrument had when registering the CMB signal.

In Figure 3.5 we can see how the CMB signal compares to the instrumental noise. The signal is the power spectrum C_ℓ^{fid}, and the noise is given as $N_\ell = \sigma_N^2 \frac{4\pi}{N_{\text{pix}}} \frac{l(l+1)}{2\pi}$. The top line shows the original signal, the middle shows the signal after it has been beam convolved, and the lower line shows the noise. The signal is convolved with a $\theta_{\text{FWHM}} = 4.5^\circ$ beam, which corresponds to a $N_{\text{side}} = 32$ resolution. The signal to noise is unity at $\ell \sim 80$.



(a) Noise estimates from experimental data

(b) Gaussian noise estimates

Figure 3.4: Figure 3.4(a) shows the noise rms per pixel. The noise in each pixel goes down as a function of how many times the area is scanned over. Thus, it is the scanning pattern of the experiment that gives it the peculiar shape. Figure 3.4(b) shows the noise for each pixel when we approximate it to be Gaussian an uncorrelated from pixel to pixel. Both images is given in $N_{\text{side}} = 32$.

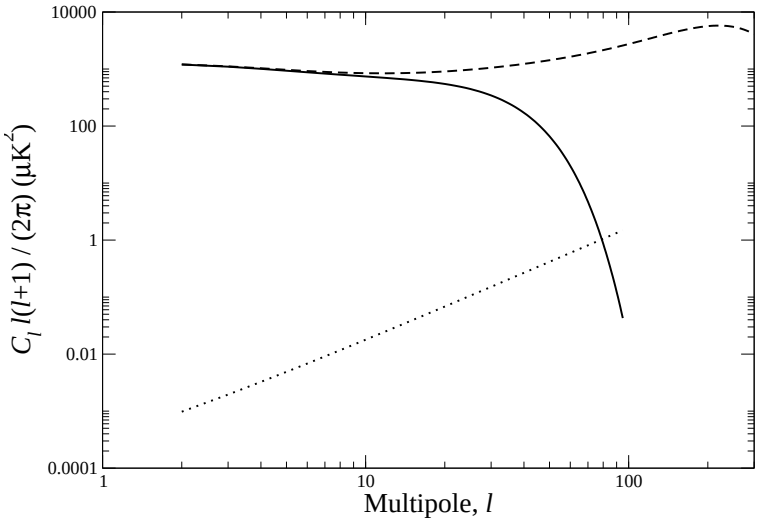


Figure 3.5: The dashed line shows the theoretical best fit power spectrum to the observed data. The full line in the middle shows the power spectrum after it has been convolved with a beam and pixel window. Due to the beam, the amplitude decreases with growing ℓ. The beam is $\theta_{\text{FWHM}} = 4.5°$, which is supported by a pixelization at $N_{\text{side}} = 32$. The dotted line at the bottom shows the instrumental noise for a resolution at $N_{\text{side}} = 32$. Compared to the signal, the noise is very small at low ℓ. This results in a signal-to-noise ratio of unity at $\ell = 79$. The noise signal is actually constant, but appears to grow as it is plotted as $\ell(\ell + 1)/(2\pi)$.

For a pure CMB signal , Equation (3.16) with an upper bandwidth limit $\ell_{\max}$ there are $\sum_{\ell=0}^{\ell_{\max}} = (\ell_{\max} + 1)^2$ real modes. Using the limit of $\ell_{\max} = 3N_{\text{side}}$ we can only get $\sim 9N_{\text{side}}^2$ independent modes. This is less than the number of pixels $N_{\text{pix}} = 12N_{\text{side}}^2$. By simply counting the number of modes, it is clear that this signal-only covariance matrix must be singular, and thus cannot be inverted. Since the matrix cannot be inverted we cannot calculate the likelihood in Equation 3.9 either. This is a quantity we want to be able to calculate.

To regularize the matrix it could be tempting to increase the value of $\ell_{\max}$ to $4N_{\text{side}}$. This will make the matrix invertible since the number of (projected) modes are greater than the number of pixels. But this also makes it mathematically arbitrary, highly pixelization dependent and not connected to the observed data since we cannot observe scales for so large ℓs.

A better way to make the matrix non-singular is by adding noise to it. We therefore add Gaussian, uncorrelated noise to each pixel. This has N_{pix} independent nodes and makes the total covariance matrix (Equation 3.13) invertible.

Chapter 4

Statistical methods

4.1 Bayesian parameter estimation

We now have a model of the CMB radiation. The power spectrum, C_ℓ, depends on cosmological parameters and by varying these we get different correlation properties in the CMB field.

There are various statistical methods to estimate the cosmological parameters from the CMB data. It is common to use maximum likelihood analysis. The best estimate of the parameters is the place in parameter space where the likelihood is the largest, and we may also calculate the errors which is the width of the likelihood function.

We could calculate the likelihood at uniform locations in parameter space, e.g., in a uniform grid or with a Monte Carlo method using random,uniformly distributed points. But likelihood evaluation is a time consuming process, and when using many parameters, this would take too long. Besides we could also risk to calculate many points where the likelihood is small and does not contribute much in estimating the parameters. Instead we want a method that focuses on the peak of the likelihood function. Sophisticated frequentist techniques are developed to maximize the likelihood, such as the Levenberg-Marquardt method and downhill simplex method. The problem with these maximum likelihood methods is that they are not always good for parameter estimation, because for a given parameter, the maximization depends on whether other parameters have been integrated out. Getting the marginal best-fit value for a parameter by integrating over the others is a procedure that gets more difficult as the number of parameters increase.

A Monte Carlo method uses random numbers to sample the distribution $\pi(\boldsymbol{\theta})$. Two ways of drawing samples from a probability distribution are importance sampling and Markov Chain Monte Carlo (MCMC) sampling. For importance sampling we can use our knowledge about the shape of the likelihood-function in order to sample wisely. Random samples are generated from a trial distribution different from (but close to) the target one and then weighted according to the importance ratio. When using MCMC sampling, a Markov chain is created where a Monte Carlo method is used to propose new points. The chain will eventually converge to the target distribution we

are looking for. I will be using the MCMC method in order to sample the likelihood function. This way of sampling also solves the problem of finding the marginal best-fit values of the parameters, as we can make one-dimensional histograms directly from the samples.

4.1.1 MCMC - Metropolis-Hastings sampling

One of the most famous MCMC methods is the Metropolis-Hastings algorithm. This algorithm was proposed by Metropolis et al. in 1953 [25] and later generalized by W.K. Hastings in 1970 [26]. It has been adopted by researchers in many different scientific fields such as biology, chemistry, computer sciences, economics, engineering, material sciences, physics and statistics. The Metropolis-Hastings algorithm can be used for generating random samples for any target distribution $\pi(\boldsymbol{\theta})$, regardless of analytical complexity and dimensionality. The MCMC method uses the current point to calculate the new one, and due to this correlation the estimates we get from these samples often have much greater variances than those resulting from independent samples. In MCMC a Markov chain is constructed, going through the various states in parameter space. After a number of samples, the chain will reach its equilibrium or stationary distribution, which is just the target distribution $\pi(\boldsymbol{\theta})$. The period before the chain reaches its equilibrium is called the "burn-in" period.

From Bayes theorem we get that the probability for a set of parameters, θ, given some observed data d is

$$P(\theta|d) = \frac{P(d|\theta)P(\theta)}{P(d)}$$

This is the target distribution we are want. $P(\theta|d)$ is called the posterior and $P(\theta)$ is the prior. $P(d|\theta)$ is the likelihood, and quantifies the probability for a data set given the theory with parameters θ. We change the notation of the likelihood to the one in Equation 3.9, $P(d|\theta) \equiv \mathcal{L}(\theta)$. If we want to normalize this we know that when we integrate the probability $P(\theta|d)$ over all values of the parameters θ, we must get 1. Hence the denominator of the expression (called the marginal likelihood) may be viewed as a normalizing constant and is equal to the integral of the numerator over θ, i.e.,

$$P(d) = \int \mathcal{L}(d|\theta)P(\theta)d\theta$$

This is independent of the parameters and does not affect the position in parameter space where the likelihood function peaks or the width of it. We may therefore ignore it when we are only interested in the maximum likelihood,

$$P(\theta|d) \propto \mathcal{L}(\theta)P(\theta)$$

If we got some prior information about the parameters, θ, we can use them here. But if we do not want to assume anything, we put in a uniform prior for the parameters.

$$P(\theta|d) \propto \mathcal{L}(\theta)$$

The steps in the Metropolis-Hastings algorithm are as follows:

1. Start with an initial point θ_0 with likelihood $\ln \mathcal{L}(\theta_0)$.
2. Take a random step to a new point θ and calculate the new likelihood $\ln \mathcal{L}(\theta)_{prop}$. Then calculate the difference between the new and old likelihood $\Delta \ln \mathcal{L}(\theta) = \ln \mathcal{L}(\theta)_{prop} - \ln \mathcal{L}(\theta_0)$
3. If $\Delta\mathcal{L}(\theta) \geq 0$, then accept the new point.
4. If $\Delta\mathcal{L}(\theta) < 0$, compute $\omega = e^{\Delta\mathcal{L}(\theta)}$.
5. Compare ω with a random number r. If $r \leq \omega$ we accept the new point, else we keep the old one.
6. Repeat steps 2-5 above until some criterion is satisfied. This could be a predetermined number of samples or some convergence criterion.

The MCMC method makes it possible to reach all states in the system, given enough time. This way we can find the most likely state of the system. We can use the samples from the chain when it has reached equilibrium to estimate the values of the parameters.

A smart choice for choosing new points is crucial for the chains to converge. If the random steps are too large we may not get many points that are accepted. On the other side, if we make the random steps too small then we may risk to never reach the most likely state. A rule of thumb is that the step lengths should be be chosen so that "roughly 50%" of all new points are accepted. To find the optimal step length we often use the try and fail method.

I will be using a Gaussian proposal density for Q, n and A, and a Euler-based, uniform proposal density for $\hat{p}$.

4.2 Model Selection

In Section 4.1 I gave an outline of how to determine the parameters in a certain model given a set of data. But how can we decide what model is the best? It is smart to keep in mind Occam's razor: the explanation of any phenomenon should make as few assumptions as possible, eliminating those that make no difference in the observable predictions of the explanatory hypothesis or theory. Model selection statistics is the task of selecting a statistical model from a set of potential models, given data. Using this we can choose between models and determine the need for new parameters. There is of course several methods for model selection. Two methods are Bayesian evidence (model comparison) and Bayesian information criterion. The Bayesian evidence is unbiased as opposed to approximations such as the information criteria. I will be looking at the Bayesian evidence in this analysis.

In parameter fitting we could ignore the marginal likelihood as it did not change the position in parameter space where the likelihood function peaks or the width of it. But in Bayesian evidence this is the quantity we are interested in. Hence, the marginal likelihood is often called the *evidence*.

For a specific model M the posterior probability of the parameters, θ, of that model given the data is

$$P(\theta|d, M) = \frac{\mathcal{L}(\theta, M)P(\theta|M)}{P(d|M)}. \tag{4.1}$$

Here, $\mathcal{L}(\theta, M)$ is the likelihood of the data given the model and its parameters and $P(\theta|M)$ is the prior on the parameters. The denominator $P(d|M)$ is the evidence for the model M. Again, this is not dependent on the parameters and is a normalizing constant setting the posterior probability to unity. This is done by marginalizing the posterior $P(\theta|d, M)$ over θ. The evidence is thus

$$E = P(d|M) = \int d\theta \mathcal{L}(\theta, M)P(\theta|M), \tag{4.2}$$

with the prior also normalized to unity.

The evidence is the average likelihood of a model over its prior parameter space. More complex models will result in better fits to the data, i.e., higher likelihood, but it will also occupy a larger volume. The evidence is proportional to the volume occupied by the prior relative to the volume occupied by the prior. This automatically implements Occam's razor. It favors simpler models with greater predictive power provided they give a good fit to the data, quantifying the tension between model simplicity and the ability to fit the data in the Bayesian sense. A good model must balance goodness of fit with simplicity. Jeffreys (1961) [27] gives us a guideline to what is a significant difference between two models. The evidences of the two models is E_1 and E_2. The difference between the logarithm of the evidences is $\Delta \ln E = \ln E_1 - \ln E_2$, and a positive value favours model M_1. $1 < \Delta lnE < 2.5$ is substantial, $2.5 < \Delta lnE < 5$ is strong and $\Delta lnE > 5$ is decisive. $\Delta lnE = 2.5$ corresponds to odds of about 1 in 13.

4.2.1 Nested Sampling - calculating the evidence

So how do we implement an algorithm for calculating the evidence? Using a Markov Chain Monte Carlo (MCMC) method to map out the posterior is a good approach to find estimates for the parameters that gives the maximum likelihood. But when using this to calculate the evidence, it would generally not be correct since the technique sample the peaks of the probability distribution well, but under-sample the tails that might occupy a large volume of the prior.

In order to sample from the whole prior volume, John Skilling introduced a method called nested sampling [28]. This scheme traces the variations of the likelihood with prior mass, with the effects of topology, dimensionality and everything else implicitly built into it. We are interested in the shape of the "nested" contours of likelihood. We can not sort all the points by their likelihood values since there are so may of them, so nested sampling simulates the operation statistically.

The following procedure for implementing this method is described by Mukherjee et al. [29]. We start of by breaking up the prior volume into a large number of "equal mass" points and order them by likelihood. To rewrite Equation 4.2 in the notation of

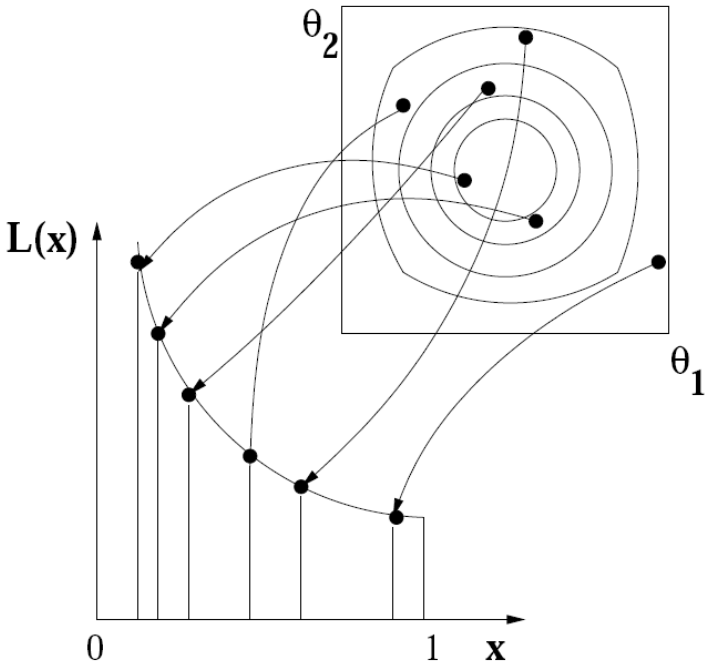


Figure 4.1: Using the nested sampling, we find the evidence by integrating over the prior volume X by moving up the nested contours of equal likelihood. Reprinted from [29].

Skilling with $dX = P(\theta, M)d\theta$ being the element of prior mass, we get

$$E = \int_0^1 \mathcal{L}(X)dX \tag{4.3}$$

The calculation of the evidence is now only a one-dimensional integral. Suppose we can evaluate the likelihood as $\mathcal{L}_j = \mathcal{L}(X_j)$, where X_j is a decreasing sequence of points

$$0 < X_m < ... < X_2 < X_1 < 1 \tag{4.4}$$

This ordering of likelihoods are shown in Figure 4.1. We then sum up these values, e.g., with the trapezoid rule, in order to get the evidence.

$$E = \sum_{j=1}^{m} E_j, E_j = \frac{\mathcal{L}_j}{2}(X_{j-1} - X_{j+1}) \tag{4.5}$$

The evidence can be calculated by the nested sampling scheme in the following way:

1. Sample N points randomly from the full prior volume and evaluate their likelihoods. It is important that the whole prior volume is sampled over, i.e., (0, X_0=1), in order to get the tails of the distribution.

2. Select the point with the lowest likelihood, L_j. The prior volume corresponding to this point, X_j, can be estimated statistically. The average volume decrease is given as $X_j/X_{j-1} = t$, where t is the expectation value of the largest of N random numbers drawn from a uniform distribution (0,1), which is N/(N+1)

3. Increment the evidence with $E_j = \frac{L_j}{2}(X_{j-1} - X_{j+1})$

4. Replace the point with the lowest likelihood with a new point, which is within the remaining prior volume (0,X_j). The new point must have a likelihood greater than the point X_j have.

5. Repeat steps 2-4 until the evidence has been estimated to some desired accuracy.

This method works its way up the surface of the likelihood function, through the nested contours of equal likelihood. Each time we replace a point the volume (or mass) decreases, and this decrease is estimated statistically. The decrease factor t is the expectation value of the largest of N random numbers drawn from a uniform distribution (0,1), which is $N/(N+1)$. After j steps the volume has shrunk to $X_j \sim t^j = [N/(N+1)]^j$.

We need a way to decide when to terminate the process. This can be done by a pre-set number of points. Or it could be when even the largest current likelihood multiplied with the remaining full mass only increase the current evidence with a small fraction f: $(L_{max})_i X_i < fE_i \Rightarrow$ terminate process After termination the N points with the highest likelihood may be added to the evidence.

We can solve a multidimensional integral by Monte Carlo sampling. Unlike the MCMC method we set a hard constraint on the likelihood to work our way up to higher values. The most difficult task in implementing the algorithm is to sample uniformly from the remaining prior volume, and not only sample from the peak of the likelihood distribution. The new point can not be correlated to the existing points, but we may use them as a guide. We find the covariance of the existing points, rotate our coordinates to the principal axes, and create an ellipsoid that just touches the maximum coordinate of the already existing points. If this ellipsoid is made too small, the new points may be chosen too near the center of the likelihood distribution and thereby overestimating the evidence. So to be sure that the new points are not biased towards the center, we multiply our limit with a constant enlargement factor, thus making it possible to sample from the full remaining prior volume with likelihood larger than the current limit. This factor also allows the contour of equal likelihood not to be exactly elliptical. New points can now be proposed uniformly within the ellipsoid until one with greater likelihood replaces the old minimum. It is important to choose an adequate enlargement factor. In [29], they made tests on the computed evidence as a function of number of initial points for different enlargement factors. Based on their tests, they chose to work with $N = 300$ and an enlargement factor of 1.5 for a five-dimensional model.

4.3 Assessment of MC convergence

To check whether an MCMC algorithm has converged, i.e., if the output can be regarded as samples from the target distribution and when a sufficient number of samples has been produced, we must run some statistics. There has been developed several tests concerning convergence, and many of them are referred in Cowles and Carlin [30]. One of them is the Gelman-Rubin-statistics, and I will here use this to determine convergence. The set up for this test is the following. When running m independent MCMC chains in parallel, each producing n samples, we compute the following quantities

$$W = \frac{1}{m(n-1)} \sum_{j=1}^{m} \sum_{i=1}^{n} \left(\theta_i^j - \hat{\theta}^j\right)^2 \tag{4.6}$$

$$B = \frac{n}{m-1} \sum_{j=1}^{m} \left(\theta^j - \hat{\theta}\right)^2 \tag{4.7}$$

$$V = \left(1 - \frac{1}{n}\right) W + \frac{1}{n} B \tag{4.8}$$

$$R = \frac{V}{W} \tag{4.9}$$

$\hat{\theta}$ is the average over all $m \cdot n$ samples and $\hat{\theta}^j$ is the average of the samples in chain number j. W estimates the variance within each chain, B estimates the variance between all chains. When the chains have converged properly, V and W should be identical and R should be close to unity. Gelman and Rubin [31] recommends that the chains should run until R < 1.2. Figure 4.2 show how different MCMC chains evolve. Four of the chains end up in the right equilibrium state after just some hundred samples. The last chain is stuck in a local maximum, until it finally reach the same state as the other chains after about 2 500 samples. The chains then fluctuate around this equilibrium, and map out the posterior distribution.

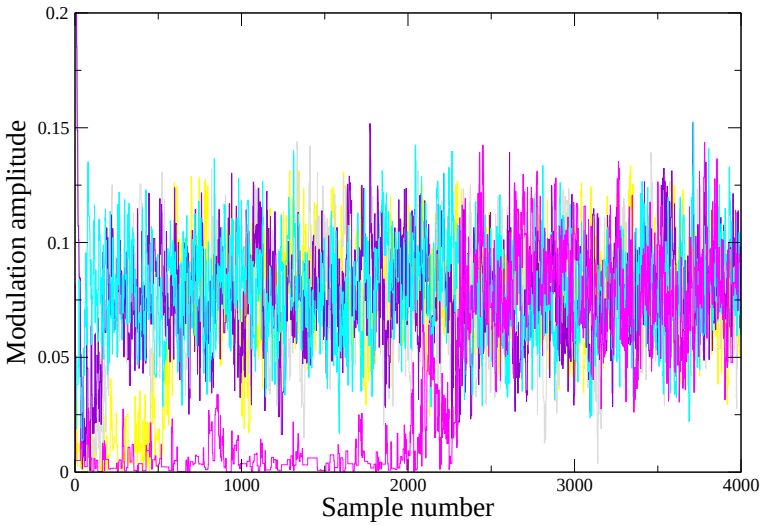


Figure 4.2: The evolution of the modulation amplitude for five MCMC chains. Four of the chains converge quite fast, while the last (pink) chain is stuck in a local maximum. Eventually it does end up with the other chains, and it is when a chain has reached this equilibrium that the desired posterior distribution can be mapped out.

Part II

Analysis of models with hemispherical asymmetry

Chapter 5

Anisotropic universe models

5.1 Introduction

After the release of the first-year WMAP data in 2003, the question on statistical isotropy has received much attention. The reason is two-fold.

The current cosmological concordance model is based on inflation theory, which predicts a statistically and homogeneous universe. This theory is very successful in describing cosmological observations, such as the CMB and large-scale power spectra, which makes us favour the isotropic and homogenic model.

On the other hand, several studies of the WMAP sky maps have showed strong hints of violation of statistical and non-Gaussianity.

For example, on large scales there is a deficit in temperature fluctuations compared to the isotropic model. This was first seen by COBE, and WMAP now confirmed and sharpened the observation. In an isotropic model this feature is unlikely and would only happen in 0.7% to 10% of the realizations, depending on assumptions, choice of statistic and analysis ([32] and references therein).

Also, statistical isotropy requires that all multipole moments of the CMB temperature field in spherical harmonics to be uncorrelated at the two point level, i.e., the signal is $S \propto N(0, I)$, where I is the identity matrix. But the data from WMAP show correlations or alignments in the temperature field. The quadrupole and especially the octopole of the WMAP temperature field are planar and they are aligned with each other. The chance for such a close alignment to happen in an isotropic universe is 1/62 [9]. This can be seen in Figure 5.1. This alignment of the CMB multipoles on very large scales has been dubbed the "axis of evil".

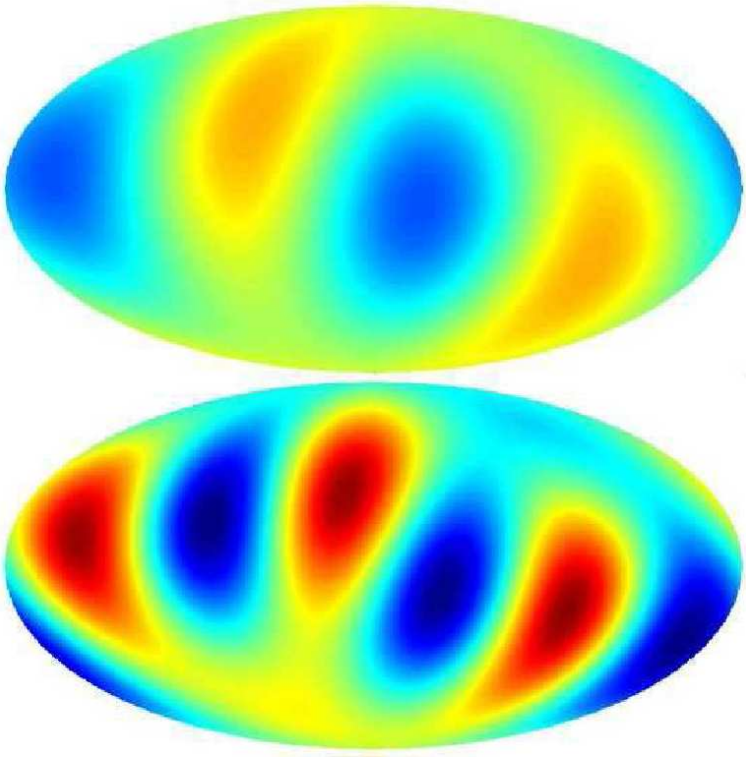


Figure 5.1: The quadrupole (top) and the octopole (bottom) components for the cleaned all-sky CMB map made by Tegmark et al. (2003) [9]. Both of them have all of their power perpendicular to a common axis in space.

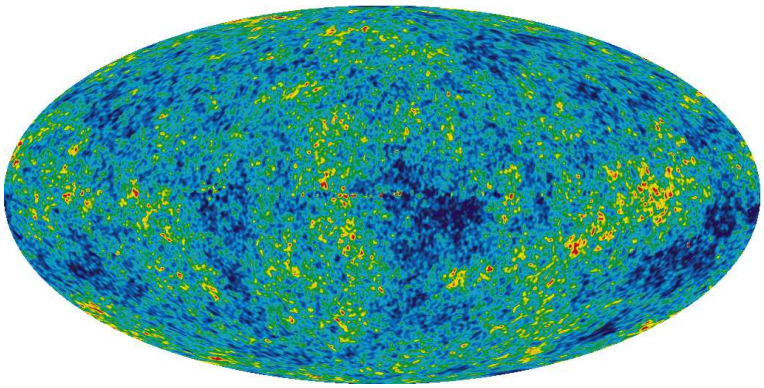


Figure 5.2: WMAPs ILC sky map. Here you can see the asymmetry in fluctuation power in two opposing hemispheres. The south-eastern hemisphere exhibit large differences in temperature on scales larger than $3 - 5°$, while the north-western hemisphere does not have as large fluctuation on the same scales. Courtesy of the WMAP science team.

A peculiar large cold spot in the southern Galactic hemisphere was reported by Vielva et al. 2004 [14]. The probability for this non-Gaussian detection is $\sim 0.4\%$. Eriksen et al. 2004 [12] found a dipolar distribution of large-scale power. The literature on the violation of statistical isotropy and non-Gaussianity have today grown very large (e.g., Bernui et al. 2006 [33]; Bielewicz et al. 2005 [34]; Copi et al. 2006 [35]; Cruz et al. 2005 [15]; Eriksen et al. 2004b,c [11][36], 2005 [37]; Jaffe et al. 2005 [38], 2006 [39]; Martínez-González et al. 2006 [40]; McEwen et al. 2008 [41]; Räth et al. 2007 [42]; Yadav & Wandelt 2008) [43]), and it would be unwise not to take these aspects seriously.

In this thesis I will be looking at the dipolar distribution of large-scale power first reported by Eriksen et al. 2004 [12]. There is an asymmetry in the distribution of fluctuation power in two opposing hemispheres. This can be seen in Figure 5.2. The southern hemisphere contains large hot and cold spots, while the northern hemisphere does not have similar large spots. The effect can be seen on angular scales down to $3-5^\circ$ or $l \lesssim 40$. The asymmetry first reported by Eriksen et al. 2004 [12] has later been confirmed by Hansen et al. 2004 [13], Eriksen et al. 2007 [24], and Hansen et al. 2008 [44].

In the paper by Eriksen et al. 2004 [12] they measure the amplitude of the large-scale fluctuations on opposing hemispheres and study the ratio of the two amplitudes. This is done with the WMAP data and with simulated maps with WMAP like properties. The amplitude is computed using both the angular power spectrum and the N-point correlation function. The results using the power spectrum show that the amplitudes on the northern Galactic hemisphere are generally lower in the WMAP data than in the simulated maps. On the southern hemisphere on the other hand, the amplitudes from WMAP are generally higher than the simulated. When comparing the ratio between the northern and southern hemisphere, only 0.5% of the simulated maps have a ratio as low as the WMAP data. The axes are rotated to find what coordinate system that gives the lowest ratio between the amplitudes of the opposing hemispheres. The analysis is done for different multipole ranges between 2-63. Depending on the multipole range included, the ratio for WMAP is high at the 95%-99% level, with the axes rotated to maximize the asymmetry between northern and southern hemisphere. The coordinates of the north pole in this rotated system is $(\theta, \phi) = (80, 57)$ in Galactic coordinates. For the N-point correlation function analysis the ratio of the large-scale fluctuation amplitudes is high at the 98%-98% level.

Hansen et al. 2004 [13] consider the variations in the angular power spectrum by sampling different patches on the sky. For an isotropic universe there should be no preferred direction and we would expect the same values for each patch on the whole sky. They also calculate the power spectrum on large scales by calculating it for the northern and southern hemisphere, and it clearly indicate an asymmetry between the two. The amplitude in the power spectrum in the range $l = 5 - 40$ is generally low for the northern hemisphere, and high in the southern hemisphere. The coordinate axes are rotated to give the maximal difference between the amplitudes of the northern and southern hemisphere. The north pole of the coordinate system that maximizes the asymmetric distribution is $(80^\circ, 57^\circ)$ in Galactic co-latitude and longitude over the

whole multipole range $\ell = 5 - 40$. This asymmetric distribution of power on the sky provides a serious test for the cosmological principle of isotropy.

What is the reason for these features is unknown, but three potential candidates are systematics, astrophysical and cosmological. The WMAP data seems to be remarkably free for systematic errors. The instrument has passed stringent tests that would make a error at this level difficult [45]. An astrophysical explanation to this could be contamination from residual foregrounds. But the foregrounds are accounted for very successfully [46] [47]. The asymmetry of fluctuation power is also very stable with respect to frequency and sky coverage.

Even if these two candidates are not the reason for the anomalies we observe in the CMB, they can weaken the significance of the statistical anomalies.

The asymmetry in the CMB may of course be a realization of an isotropic universe, even if it is unlikely. But it would be more satisfactory if there is some physical phenomenon that causes these effects. After the first year release of the WMAP data showed hints of violation of statistical isotropy and/or non-Gaussianity, there has been much interest in anisotropic universe models. Physicists and theorists have come up with new theories to explain these different features. I will here present some theories.

5.2 Anisotropic inflation models

The inflation theory proposed by Guth solved the horizon, flatness and monopole problems and in addition provides an mechanism for generating primordial density perturbations. These perturbations give rise to galaxy formation and temperature anisotropies in the CMB we see today. However inflation predicts statistical homogeneity and isotropy. The observed asymmetry in the CMB has caused theorists to come up with new models with anisotropic inflation, and inflation models can be constructed to violate statistical isotropy (e.g, Gümrükçüoglu et al. 2007 [48]; Pullen & Kamionkowski 2007 [49]; Kanno et al. 2008 [69]; Yokoyama & Soda 2008 [74], Ackerman et al. 2007 [50], Erickcek 2008 [65]).

5.2.1 CMB statistics for a direction-dependent primordial power spectrum

The assumption of statistical isotropy of primordial perturbations is an assumption that should be tested. Pullen & Kamionkowski [49] develop cosmic microwave background statistics for a primordial power spectrum that depends on the direction and the magnitude of the Fourier wavevector. The density perturbation to a point in space $\mathbf{x}$ is given by $\delta(\mathbf{x})$. If we Fourier expand δ we get the Fourier transform $\delta(\mathbf{k})$, where $\mathbf{k}$ is the conjugate to the real space vector $\mathbf{x}$. As in Equation 3.5 the power spectrum is defined by

$$\langle \delta(\mathbf{k}) \delta^*(\mathbf{k}') \rangle = \delta_D(\mathbf{k} - \mathbf{k}') P(\mathbf{k}) \tag{5.1}$$

The angular brackets denote an average over all realizations of the random field, and δ_D is Dirac's delta function. In a statistically isotropic universe the power spectrum,

$P(\mathbf{k})$, only depend on the magnitude, k, of the wavevector $\mathbf{k}$, as in Equation 3.5. But if it is not isotropic, the power spectrum also depend on the direction $\hat{\mathbf{k}}$. The most general power spectrum can be written as

$$P(\mathbf{k}) = A(k)\left[1 + \sum_{LM} g_{LM}(k) Y_{LM}(\hat{\mathbf{k}})\right] \tag{5.2}$$

Here $Y_{LM}(\hat{\mathbf{k}})$ are the spherical harmonics, and $g_{LM}(k)$ quantifies the departure from statistical isotropy as a function of k. For $g_{LM}(k) = 0$, we get the usual statistical isotropic power spectrum $A(k)$, which only depend on the magnitude k.

This will change our predictions of the CMB. The spherical harmonic expansion of the temperature field was given in Equation 3.3. For statistical isotropy, the $a_{\ell m}$s are statistically independent for different ℓ and m, and we get $\langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'}$. C_ℓ is the angular power spectrum and is the spherical version of the power spectrum $P(\mathbf{k})$. When statistical isotropy is broken there are correlations between $a_{\ell m}$s for different ℓ and m. The two-point correlation for statistical isotropy is given by Equation 3.14, i.e., it only depends on the angular separation between two points. When breaking statistical isotropy this is no longer necessarily true.

If we consider a primordial power spectrum as given in Equation 5.2, we can write the covariance matrix as

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'} + \sum LM \xi^{LM}_{\ell m \ell' m'} D^{LM}_{\ell\ell'} \tag{5.3}$$

The first term on the right side denotes the usual power spectrum for statistical isotropy, while the second term describes departures from statistical isotropy. The upper-case indices LM are used for power anisotropies, and the lower-case indices ℓm are used for temperature/polarization anisotropies.

So if the primordial perturbations are statistically isotropic and Gaussian, then the statistics of the CMB can be fully described by the set of C_ℓs. But if the perturbations deviates from statistical isotropy and this can be described in terms of spherical harmonics $Y_{LM}(\hat{\mathbf{k}})$, then the two-point statistics can be described by an additional term consisting of the set of multipole moments $D^{LM}_{\ell\ell'}$.

They propose two tests for departures from statistical isotropy. The first is called "The power multipole moments" and can be used to search in a model independent way for departures from statistical isotropy. But if a specific model is introduced with a particular parameterization of the functions $g_{LM}(k)$ then the "minimum variance estimator" test can improve the precision with which these parameters can be constrained.

5.2.2 Imprints of a primordial preferred direction on the CMB

Ackerman et al. [50] proposed that a violation of rotational invariance during inflation could cause certain effects that might be observable today, such as an imprint on the CMB. This violation may have been caused by an effect that has disappeared in the

later universe. The effects of this violation are expected to be scale-invariant, just as the amplitudes of the perturbations set up during inflation are scale-invariant.

The breaking of rotational invariance gives rise to correlations between multipole moments that would normally vanish and also alters the predictions for the usual multipole moment correlations.

Here they change the power spectrum for the primordial density perturbations as

$$P'(\mathbf{k}) = P(k)\left(1 + g(k)(\hat{\mathbf{k}}\cdot\mathbf{n})^2\right) = P(k)\left(1 + g_*(\hat{\mathbf{k}}\cdot\mathbf{n})^2\right) \tag{5.4}$$

The last equality comes from the assumption that the same arguments that lead to the scale-invariant Harrison-Zel'dovich spectrum, imply that $g(k)$ is independent of k, which makes it a constant g_*. Note that this is a special case of Equation 5.2. The change of power spectrum also change the predictions for the expectation values, $\langle a_{\ell m} a_{\ell' m'*}\rangle$, of the microwave background anisotropies.

$$\langle a_{\ell m} a_{\ell' m'*}\rangle = \langle a_{\ell m} a_{\ell' m'*}\rangle_0 + \Delta(\ell m; \ell' m') \tag{5.5}$$

The subscript 0 denotes the usual invariant piece, while the last term is the perturbation of the isotropic part. The expectation values only depend on three parameters, the amplitude g_* and a direction on the sky defined by a unit vector $\mathbf{n}$. This theory can be directly compared with observations to probe the existence of small Lorentz-violating effects in the very early universe.

This was done by Groeneboom & Eriksen 2008 [51] using the WMAP data. For $\ell \leq 400$ using the W-band data, they found that the amplitude $g_* = 0.15 \pm 0.039$ and the preferred direction is $(l, b) = (110°, 10°)$. Using the V-band data they found a similar result with the amplitude $g_* = 0.10 \pm 0.04$ and the preferred direction is $(l, b) = (130°, 20°)$. Figure 5.3 shows the anisotropic contribution for a simulated map with parameters corresponding to the W map posterior.

5.3 Dark energy

Dark-energy models may also explain the the departure from statistical isotropy. (Battye and Moss [52], Koivisto and Mota [53])

5.3.1 Late, anisotropic acceleration of the universe

Koivisto and Mota [53] proposed that anisotropies in the dark energy may be the reason for the observed anomalies in the CMB. One motivation for this idea is that the large scales that we are observing the anomalies in the CMB are the same that enter the horizon at the time when the dark energy started to dominate. Dark energy has negative pressure, and causes the expansion rate of the universe to accelerate. If the pressure varies with direction, then the universal expansion becomes anisotropic and this could affect the low multipoles of the CMB. A similar CMB pattern may occur in an universe which is ellipsoidal at the time of last scattering. To distinguish these

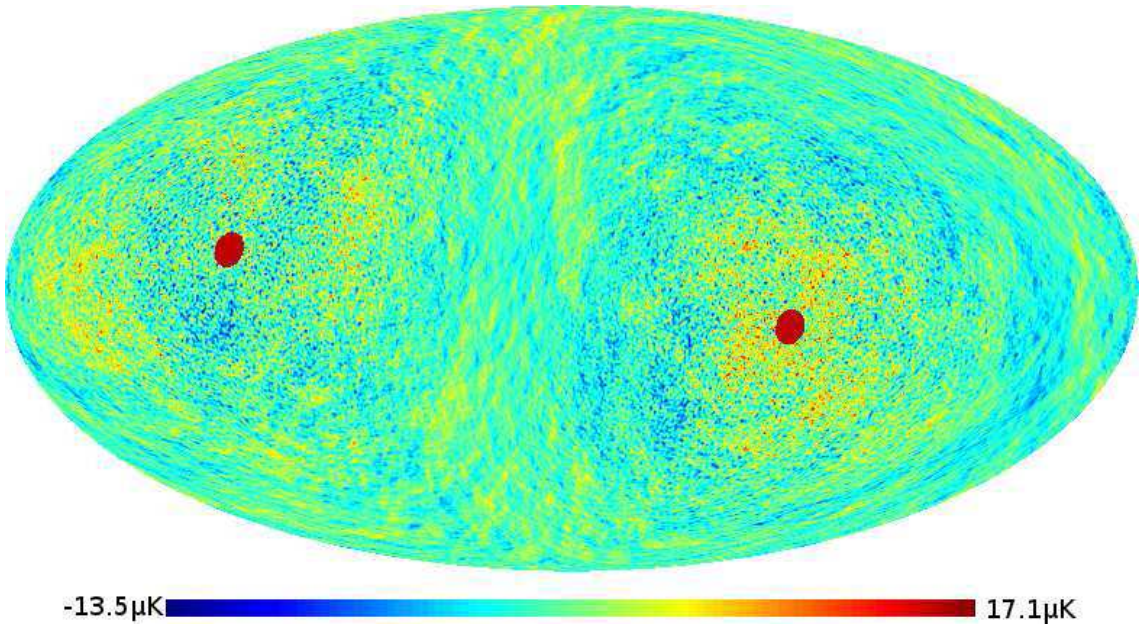


Figure 5.3: A simulated realization drawn from a Gaussian distribution with zero mean and a covariance matrix given by the anisotropic term in the ACW model, computed for an asymmetry amplitude of $g_* = 0.14$ and a preferred direction $(l, b) = (110°, 10°)$, marked by red dots. Notice the rotational structure about the preferred direction. The amplitude of the anisotropic component is $\pm 15 \mu K$, or ~ 3 % of the isotropic component. Reprinted from [51].

two scenarios one can use data from supernovas. The Robertson-Walker (RW) metric cannot describe an anisotropic expansion, so the authors introduce the Bianchi type 1, B(1), metric. This way the universe can expand at different rates in different directions. Their results are very interesting. Even if the CMB formed isotropically at early times, it could be distorted by the anisotropic acceleration caused by dark matter at a later time so that it would appear anomalous at large scales. The anisotropic distribution of dark energy may explain the observed expansion of the universe and the large scale anomalies in the CMB.

Gordon et al. 2005 [32] examine the possibility that statistical isotropy is spontaneously broken in the CMB. They introduce two models for isotropy breaking: additive contributions and multiplicative modulation of the intrinsic anisotropy. The additive contribution is less likely to produce the observed alignments in the CMB than the isotropic fluctuations. Multiplicative models overcome these difficulties. They generate stronger alignments than WMAP in 30-45% of the realizations.

5.4 Phenomenological description by a dipolar modulation

Instead of the standard isotropic CMB temperature we may apply a modulated field to the isotropic signal as introduced by Gordon et al. [32]. Here we suggest a model with a multiplicative modulation field to the isotropic CMB signal. The physical process that created this dipole field is unknown. I am mostly interested in how well the parameters of the model fit the data. The signal that we are receiving can be written as

$$\mathbf{d}(\hat{n}) = \mathbf{s}(\hat{n})(1 + \mathbf{a}(\hat{n})) + \mathbf{n}(\hat{n}) \tag{5.6}$$

where $\mathbf{a}(\hat{n}) = A(\hat{p} \cdot \hat{n})$ is the modulation field, with $\hat{p}$ as the direction of the modulation axis and A is the amplitude of the modulation field. The amplitude A of the modulation field is not to be confused with the beam convolution A of the signal. $\mathbf{s}(\hat{n})$ is a statistically isotropic and Gaussian random field with power spectrum C_ℓ, just as in Section 3.1. It is the modulated signal that is anisotropic, but it is still a Gaussian, random field.

Doing a similar calculation as for the isotropic model, we end up with the covariance matrix

$$\mathbf{C}(\hat{n}, \hat{m}) = \tilde{\mathbf{S}} + \mathbf{N} + \mathbf{F}, \tag{5.7}$$

where

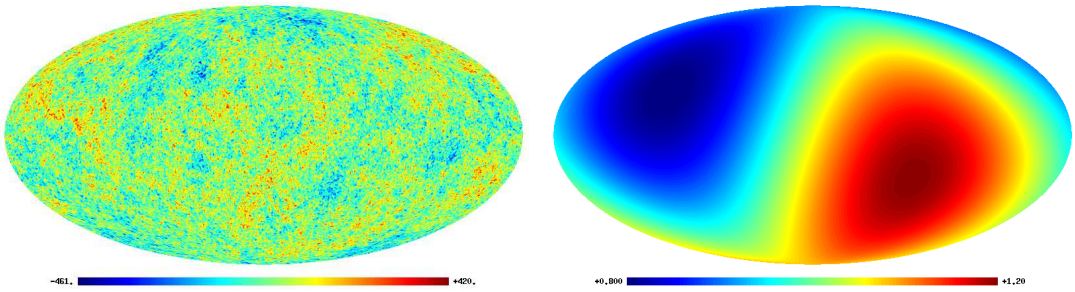
$$\tilde{\mathbf{S}}(\hat{n}, \hat{m}) = [1 + \mathbf{a}(\hat{n})]\mathbf{S}(\hat{n}, \hat{m})[1 + \mathbf{a}(\hat{m})], \tag{5.8}$$

$\mathbf{S}(\hat{n}, \hat{m})$ is the same as in Equation 3.14, the noise, $\mathbf{N}$, as in Equation 3.17 and the foreground template,$\mathbf{F}$, as in Equation 3.18.

To get a better feeling of what this field do, I have created a simulated, isotropic CMB map. This map is then multiplied with a modulation field, where the amplitude $A = 0.1$ and the direction of the axis is $(\theta,\ \phi) = (2,\ 5)$ in radians. This gives us a modulated map, where the temperature in one hemisphere become more uniform, while the other get larger temperature differences. This is illustrated in Figure 5.4.

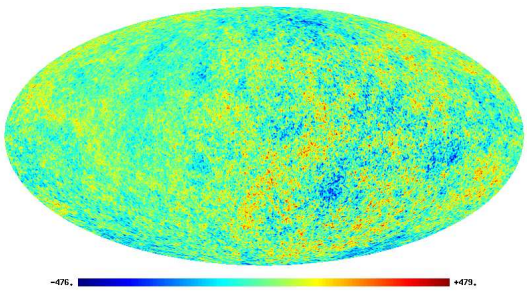
The first reports on hemispherical asymmetry showed the effect on scales up to $\ell = 63$ by Eriksen et al. [12] and up to $\ell = 40$ by Hansen et al. 2004 [13]. This asymmetry between the northern and southern hemisphere was also maximized when rotating to a special coordinate system. An analysis based on the modulated model in Equation 5.6 was done by Eriksen et al. 2007a [24] by using an optimal Bayesian framework on the three-year WMAP data for scales up to $\ell \lesssim 40$. This way we can quantify the asymmetry using the amplitude A along with the direction of the field, $\hat{p}$. The model with the isotropic CMb sky modulated by a dipole field gave a substantially better fit to the WMAP data than the purely isotropic model. The Bayesian log-evidence difference is ~ 1.8 in favour of the modulated model, which ranks as substantial. The amplitude is found to be 0.114 and the direction of the dipole axis points towards (l, b) = (225°, −27°). The frequentist probability of obtaining such a high amplitude in an isotropic universe is $\sim 1\%$.

A paper by Hansen et al. 2008 [44] shows that the hemispherical power asymmetry also is evident at higher ℓs, up to 600. This is the motivation for doing an analysis based



(a) Isotropic map

(b) Modulation field



(c) Modulated map

Figure 5.4: 5.4(a) shows a simulated, isotropic map. The fluctuations are statistically the same on all scales over the entire sky. 5.4(b) is the modulation field $1+A(\hat{p}\cdot\hat{n})$, with amplitude A = 0.4 and the direction of the modulation field $\hat{p}$ at coordinates (θ, ϕ)=(2, 5) in radians. In 5.4(c) the isotropic temperature field has been multiplied with the modulation field. We can here see how the modualtion field change the temperature in two opposing hemispheres. The temperature in the northern hemisphere now looks more uniform, while the temperature fluctuations in the southern hemisphere are larger than in the isotropic map. The amplitude has been set large to make the difference in the two maps more distinct.

on this model again for higher values of ℓ. They also found that the dipole amplitude tends to decrease with ℓ.

The problem now is to try and estimate the three parameters introduced here, the amplitude A and direction $\hat{p}$ of the modulation field, for larger values of ℓ. We have also chosen to parameterize the angular power spectrum with a free amplitude q and tilt n as defined in Equation 3.15. There is a total of five parameters we must determine. This can be done by using the method of parameter estimation given in Chapter 4. In the next chapter I will go through how this method is implemented in a program. The code for calculating the Bayesian evidence is also presented here.

Even if it is easy come up with new theories, it might be hard to test them. Power spectrum estimators scales as $\mathcal{O}\left({N_{pix}}^{\frac{3}{2}}\right)$, while the calculation of the likelihood scales as $\mathcal{O}\left({N_{pix}}^{3}\right)$. We must therefore do some tests to see for what resolution it is possible to do an analysis using this brute force way of inverting and calculating the determinant of the covariance matrix in the likelihood.

Chapter 6

Implementational Details

In this chapter I will go through the codes that carry out the calculations of the methods introduced in Chapter 4. The codes are only pseudo-coded, as a complete code would take up too much space, and it a pseudo-code can be easier to read.

6.1 The Program

I have used two programs that basically do the same thing. They both calculate the values of the desired parameters and the evidence. The difference is the way they handle matrices. The first program use LAPACK for the matrix operations and MPI to run multiple chains of the same program on different processors. Every chain has to go through a burn-in period before it converges.

The second program uses ScaLAPACK. This program library basically does the same as LAPACK. The difference is that it can split up a matrix into smaller parts and then given to different processors. All processors can then work on their parts of the matrix at the same time, and this way the total matrix operation will be completed faster.

When starting the program the first thing to do is to initialize all the necessary parts. One must allocate arrays to contain the data and read it in, exclude the pixels that are masked out, initialize the random number generators, start MPI, distribute the data to all the processors, etc. Only the pixels that are not excluded by the mask are included in the data vector, so the computational time goes down when the mask exclude more pixels.

6.1.1 MCMC code

I here present the pseudo-code for estimating the parameters using Metropolis-Hastings algorithm. A Gaussian proposal density is used for Q, n and A and an Euler-matrix-based, uniform proposal density is used for $\hat{p}$. When using a Gaussian proposal density, one simply make a small perturbation to the existing point, q_i. This perturbation is determined by a small amplitude q_{rms} and a random Gaussian number. The random

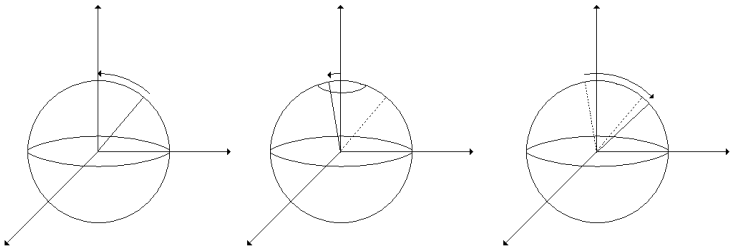


Figure 6.1: Here is a sketch of how the direction of the modulation axis, $\hat{p}$ is proposed. Left:Rotate vector so it is parallel to the z-axis. Middle: The new point is randomly chosen within the disc indicated around the z-axis, and the vector is rotated to the new point. Right: The vector is now rotated back relative to the old point.

number is picked from a Gaussian distribution with $\mu = 0$ and $\sigma = 1$. So the new point is $q_{i+1} = q_i$ + G(0,1)q_{rms}.

The proposal for new points of the axis on the sky, $\vec{p}$, is done in a different way. The new point, $\vec{p}_{i+1}$, must lie within a disc with the original point $\vec{p}_i$ as the center. The angle of this disc is θ_p. The original point $\vec{p}_i$ has coordinates (θ_i, ϕ_i). We know the rotation matrices

$$M_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix}, M_y = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix}, M_z = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

First we rotate our original point so its parallel to the z-axis, i.e. $\hat{p}_i \parallel \hat{z}$. To do this we must first rotate around the z-axis with angle $-\phi_i$, then rotate around the y-axis with $-\theta_i$. From this position we can now propose a new point. The new point must be chosen uniformly within $\theta \sim U[0, \theta_p]$ and $\phi \sim U[0, 2\pi]$. We then rotate to this new point: around the y-axis with θ and then around the z-axis with angle ϕ. This new point is then rotated back relative to the old point. We rotate around the y-axis with θ_i, then rotate around the z-axis with angle ϕ_i. So the complete set of transformations for proposing a new point is given as $p^{i+1} = M_z(\phi_0)M_y(\theta_0)M_z(\phi)M_y(\theta)M_y(\text{-}\theta_0)M_z(\text{-}\phi_0)p^i$. This process is shown in Figure 6.1.

The following is the code for parameter estimation using MCMC chains.

```
! Propose initial point for all parameters within prior volume
q, n, a, theta, phi

! Compute likelihood for this point
lnL_0

! Run N samples
```

```
do i = 1, N

   ! Propose new point
   q_prop     = q + Gaussian random number * q_rms
   n_prop     = n + Gaussian random number * n_rms
   a_prop     = a + Gaussian random number * a_rms
   ! theta_prop and phi_prop are proposed using an Euler-matrix-based,
   ! uniform proposal density

   ! check if parameters are within borders

   ! Compute new likelihood
   lnL_prop

   ! Apply Metropolis-Hastings rule
   if (lnL_i > lnL_0) then
      accept point
   else
      if (uniform random number < exp(lnL_prop - lnL_0)) then
         accept
      else
         reject
      end if
   end if

   if (accept) then
      numaccept = numaccept + 1

      ! replace the parameters with the new
      q     = q_prop
      n     = n_prop
      a     = a_prop
      theta = theta_prop
      phi   = phi_prop
      lnL   = lnL_prop
   end if

   !write to file

end do
```

The calculation of the likelihood is done as in Equation 3.9. First, the covariance matrix is constructed by calculating all the parts in 5.7 with the proposed parameters. This matrix is symmetric and positive definite. Then a library (LAPACK or

ScaLAPACK) is used to Cholesky factorize this matrix, and the factorization is used to find the inverse and determinant of the covariance matrix.

6.1.2 Code for nested sampling

In this section I will present the code for calculating the evidence using the nested scheme. Here, a Gaussian proposal density is used for all the parameters.

```
! Setting up the prior volume
priors(1,1)   =  -1.d0                  ! low prior for cos(theta)
priors(1,2)   =  1.d0                   ! high prior for cos(theta)
priors(2,1)   =  0.d0                   ! low prior for phi
priors(2,2)   =  2.d0*pi                ! high prior for phi
priors(3,1)   =  0.d0                   ! low prior for A
priors(3,2)   =  prior_f                ! high prior for A
priors(4,1)   = 1.d0 - prior_q          ! low prior for q
priors(4,2)   = 1.d0 + prior_q          ! high prior for q
priors(5,1)   = -prior_n                ! low prior for n
priors(5,2)   =  prior_n                ! high prior for n


! Initialise N points randomly from the prior
do i = 1, N

   ! Propose a random point uniformly within the prior volume
   do j = 1, numpar
      samples(i,j) = priors(j,1) + rand_uni(rng_handle) * (priors(j,2)-priors(j,1))
   end do

   ! Compute likelihood
   lnL_x(i)
end do

! Find the point with the lowest likelihood, L_j
do i = 1, N
   if (L_x(i) < L_j) then
      L_j = L_x(i)
      j   = i
   end if
end do

!Initialize mass fraction
X = 1
t = N / N + 1
```

```
! Compute new points within the remaining volume
do i = 2, num_samp

          ! Compute the mean (center) of the N points
          do k = 1, numpar
                    centroid(k) = sum(samples(:,k)) / real(n,dp)
          end do

          ! Compute ellipsoid that just touch the outer points
          ! Expand the ellipsoid with an enlargement factor
          L_param

          ! propose new points until one has a higher likelihood
          reject_sample = .true.
          do while (reject_sample)

                    ! Propose a point near the edge of the ellipsoid,
                    ! making a small perturbation eta
                    params_prop = centroid + matmul(L_param, eta)

                    ! Check if the new point is within the prior volume
                    if (outside_prior) cycle
                    ! i.e. go to the start of the while-loop and propose a new point

                    ! Compute likelihood for this new point

                    ! Compare likelihood with the lowest one
                    if (lnL_prop > L_x(j)) then

                              ! accept new point
                              reject_sample = .false.
                    else
                              ! reject sample
                    end if
          end do

          ! write likelihood and fraction of the prior mass to file
          write(unit,*) lnL_j, X

          ! Replace lowest point with the proposed one
          lnL_j = lnL_prop
          X     = X * t
```

```
            ! Find a new lowest point within the N samples
            L_j = 1.d30
            do k = 1, n
                      if (L_x(k) < L_j) then
                                L_j = L_x(k)
                                j   = k
                      end if
            end do

end do
```

After the program has computed the points and written them to file, I can calculate the evidence by using the midpoint rule as in 4.5. The file writes out the logarithm of the likelihood, so we need to take the exponent of this to use it. Since these exponents become so small the computer has problems computing the evidence. So we subtract the largest value of the $\mathcal{L}_j$ to make the exponents manageable.

$$E_j = \frac{\mathcal{L}_j}{2}(X_{j-1}-X_{j+1}) = \frac{e^{\ln \mathcal{L}_j - \mathcal{L}_{max} + \mathcal{L}_{max}}}{2}(X_{j-1}-X_{j+1}) = \frac{e^{\mathcal{L}_{max}} e^{\ln \mathcal{L}_j - \mathcal{L}_{max}}}{2}(X_{j-1}-X_{j+1})$$

So the evidence becomes

$$E = \frac{e^{\mathcal{L}_{max}}}{2} \sum_{j=1}^{m} e^{\ln \mathcal{L}_j - \mathcal{L}_{max}}(X_{j-1} - X_{j+1})$$

6.2 Computational expense

We now want to find out how much time it takes to produce one sample. There are two things that take a lot of time when doing the computation. First, we must construct the covariance matrix, and then we must evaluate the likelihood. When using LAPACK, only one processor

Using a Monte Carlo method to map out the posterior or calculating the evidence, we must calculate the likelihood several thousand times. To reduce the time of each evaluation of the likelihood, we could parallelize the matrix operations. When using LAPACK, only one processor alone can perform the needed operations on the covariance matrix. But by using ScaLAPACK, several processors can get a part of the covariance matrix, so all processors can work simultaneously, hence reducing the time of calculating the likelihood. An introduction to ScaLAPACK can be found in Appendix B.

For ScaLAPACK we have to find out what block size and number of processors that does the most time efficient calculation. An initial test using 25 processors showed that the computational time was minimized when using a block-sizes of MB=NB=32. This number was then used for calculating the time per sample as a function of the number of pixels used, N_{pix}. It was tested for two maps at $N_{\text{side}} = 16$ and $N_{\text{side}} = 32$.

The used block-sizes may not be the most ideal when using a number of processors different from 25. As the initial test was done using 25, this may be the best (cubic)

N_{pix}	1 proc	4 procs	9 procs	16 procs	25 procs
1408	2.71	0.33	0.21	0.17	0.17
1664	3.87	0.36	0.22	0.18	0.17
1920	5.20	0.66	0.45	0.30	0.28
2176	7.14	0.99	0.69	0.43	0.40
2432	9.20	0.79	0.53	0.33	0.30
2944	13.84	2.11	1.33	0.84	0.73
5888	92.08	39.31	18.63	24.05	9.39
6912	131.44	74.51	34.78	104.24	18.21
7936	157.53	78.80	37.64	75.05	19.59
8960	209.30	122.49	58.03	71.60	29.98
9984	266.11	122.59	58.17	131.25	30.47
11520	423.86	212.60	101.05	263.49	53.27

Table 6.1: Time per sample when using different number of processors. The first column show how many pixels that are used. The second column show the time per sample when only using one processor. The following columns show the time per sample when using 4, 9, 16 and 25 processors. ScaLAPACK is used to calculate the likelihood by distributing the covariance matrix to the processors. The block sizes are MB=NB=32 and were choosen based on a test using 25 processors. Therefore the may not be optimal for the other processors. This is evident for the case when using 16 processors, as the processing time is much larger than when using less processors.

block-size for other choices of processors. Using a cubic block-size (MB = NB), may also not be the fastest choice. Further tests should be run in order to find the best partition factors MB and NB when using a certain number of pixels and number of processors. The time per sample using 1, 4, 9, 16 and 25 processors are presented in Table 6.1 and illustrated in Figure 6.2. The times are calculated when using different Here we can clearly see that something is not working out right for the computation when using 16 processors. The choice of block-size used here is not very optimal when using a 4×4 processor grid.

6.3 Testing the program

As a simple test I created a synthetic, isotropic map at $N_{\text{side}} = 16$, multiplied it with a modulation field as in Equation 5.6 and then added random, Gaussian noise to each pixel. The amplitude was set to $A = 0.1$ and the direction was set to (2,5). This map was then run through the program. For $\ell = 47$ the results gave an amplitude at A $= 0.118 \pm 0.036 \sim 3.3\sigma$. So the initial sat amplitude is within the uncertainty of the results.

The program was also ran with the same isotropic map, without the modulation field, and added noise. This gave an amplitude of A $= 0.035 \pm 0.026 \sim 1.3\sigma$. So even if I sat A = 0, it still finds a non-zero amplitude, even if it is very small. So a non-zero amplitude can originate from an isotropic universe as well. But the uncertainty of this

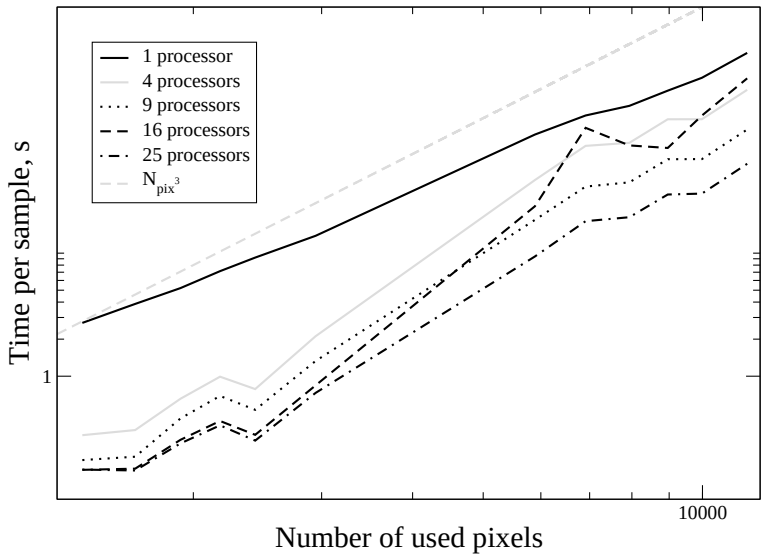


Figure 6.2: The average time used for each sample as a function of the number of pixels used. The benchmarks have been done for 1, 4, 9, 16 and 25 processors. The time per sample goes down as more processors are used. But the scale up is not optimal. When using 25 processors, the speed-up is only about 8 times faster than with one processor. The case when using 16 processors does not compute well with the chosen block size. For reference, N_{pix}^3, is also plotted. For a high number of pixels, the time per sample scales a bit better than N_{pix}^3.

calculation is very large. The distribution of the amplitudes and axes is given in Figure 6.3. The direction is clearly preferred in the case of the modulated temperature map, while the isotropic map does not seem to favor any direction.

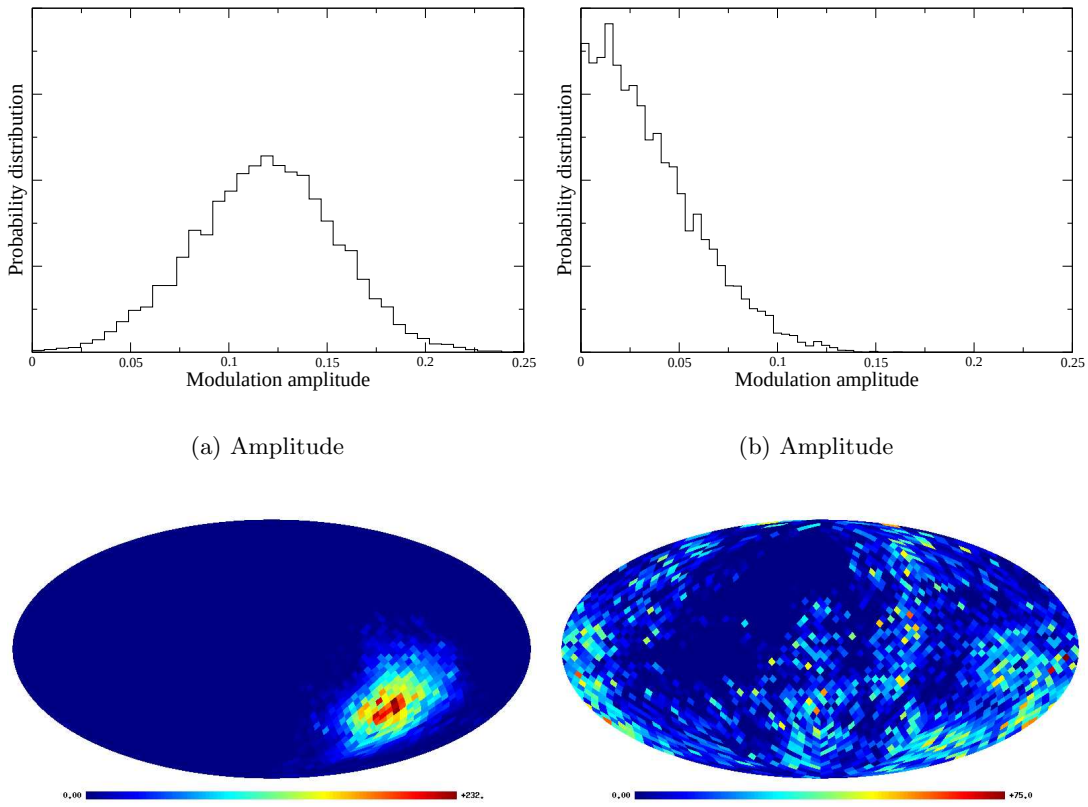
So for this simple test, the program finds the amplitude and direction of the modulation field.

When running maps with different seeds for noise based on the same data set, I got different values for the amplitude. The reason for this was that the noise dominates at small scales, just as showed in Figure 3.5. To fix this problem I split the signal covariance matrix of Equation 5.8 into two parts, one modulated low-ℓ part and one isotropic high-ℓ part, $\mathbf{S} = \mathbf{S}_{\mathrm{mod}} + \mathbf{S}_{\mathrm{iso}}$. The first part is given as

$$S_{\mathrm{ij,mod}} = \frac{1}{4\pi} \sum_{\ell=2}^{\ell_{\mathrm{mod}}} (2\ell+1) C_\ell (p_\ell b_\ell)^2 P_\ell(\cos\theta_{\mathrm{ij}}) \tag{6.1}$$

where we let all of the parameters to vary. The second part is

$$S_{\mathrm{ij,iso}} = \frac{1}{4\pi} \sum_{\ell_{\mathrm{mod}}+1}^{\ell_{\mathrm{max}}} (2\ell+1) C_{\ell,\mathrm{iso}} (p_\ell b_\ell)^2 P_\ell(\cos\theta_{\mathrm{ij}}) \tag{6.2}$$



(a) Amplitude

(b) Amplitude

(c) Position

(d) Position

Figure 6.3: The upper row show the posterior distribution of the amplitude in the modulation field, while the two lower pictures shows the posterior distribution of the dipole axis of the field. Posterior distribution for the amplitude and axis of the dipole modulation field for two simulated maps. The left column shows the distribution for an isotropic map where the temperature field has been modulated. Here, an axis is clearly preferred. The two pictures on the right shows the distribution for a purely isotropic, simulated map. This does not show any preferred axis. The isotropic map is used to make the modulated map. Both maps have Gaussian noise added to them.

For the isotropic part we set $A = 0$, $q = 1$ and $n = 0$.

The isotropic part acts as a buffer for the calculations. There are two things that cause the signal not to be dominant at high values of ℓ. First of all we have signal-to-noise ratio at some ℓ_{unity}, so for ℓ values above this, the noise will dominate over the signal. Secondly, cutting off the multipoles at ℓ_{max} may give some aliasing effects, causing the signal to be distorted around this multipole. We can now choose ℓ_{mod} lower than ℓ_{max} and investigate the properties for the modulation field up to this value. This splitting may not seem like physically well-motivated, but the purpose here is to characterize its properties.

After changing the signal covariance matrix, the amplitude of the modulation field was much more stable when using different noise seeds, as the noise did not dominate at ℓ_{mod}.

Part III

Application to the WMAP data

Chapter 7

The WMAP experiment

7.1 Data

The data I will be using in this analysis are the five year WMAP data. These are available from LAMBDA[1]. The satellite observes the Cosmic Microwave Background radiation in five frequencies. Specifications for each frequency is given in Table 7.1. The purpose of producing multi-frequency maps is to separate the CMB anisotropy from the foreground emission. The four main physical mechanisms that contribute to the galactic foreground emission are synchrotron radiation, free-free radiation (Bremsstrahlung), thermal radiation from dust and radiation from charged spinning dust and grains. Figure 7.1 shows the intensity for the different foregrounds and the CMB. At high frequencies, dust dominates, and at very low frequencies synchrotron and free-free emission are important. For the frequency range in between, from 30 to 200 GHz, the CMB signal often have the largest intensities. The spectral shape of the foregrounds are all different from one another and different from the blackbody spectrum of the CMB. Detections at different frequencies can be used to to extract the CMB signal from the foregrounds.

We wish to remove as much foreground emission as possible, in order to get a better view of the clean CMB signal. Bennett et al.(2003) [54] proposes three conceptual approaches that can be used, both individually and in combination, to remove the dominating emission from the galactic foregrounds.

The first is to use existing measurements of the emission from galactic foregrounds

[1]Legacy Archive for Microwave Background Data Analysis (LAMBDA), http://lambda.gsfc.nasa.gov/

Properties	K-band	Ka-band	Q-band	V-band	W-band
Frequency(GHz)	23	33	41	61	94
Number of Differencing Assemblies	1	1	2	2	4
Beam Size (deg)	0.88	0.66	0.51	0.35	0.22

Table 7.1: Specifications for the five differencing assemblies onboard WMAP

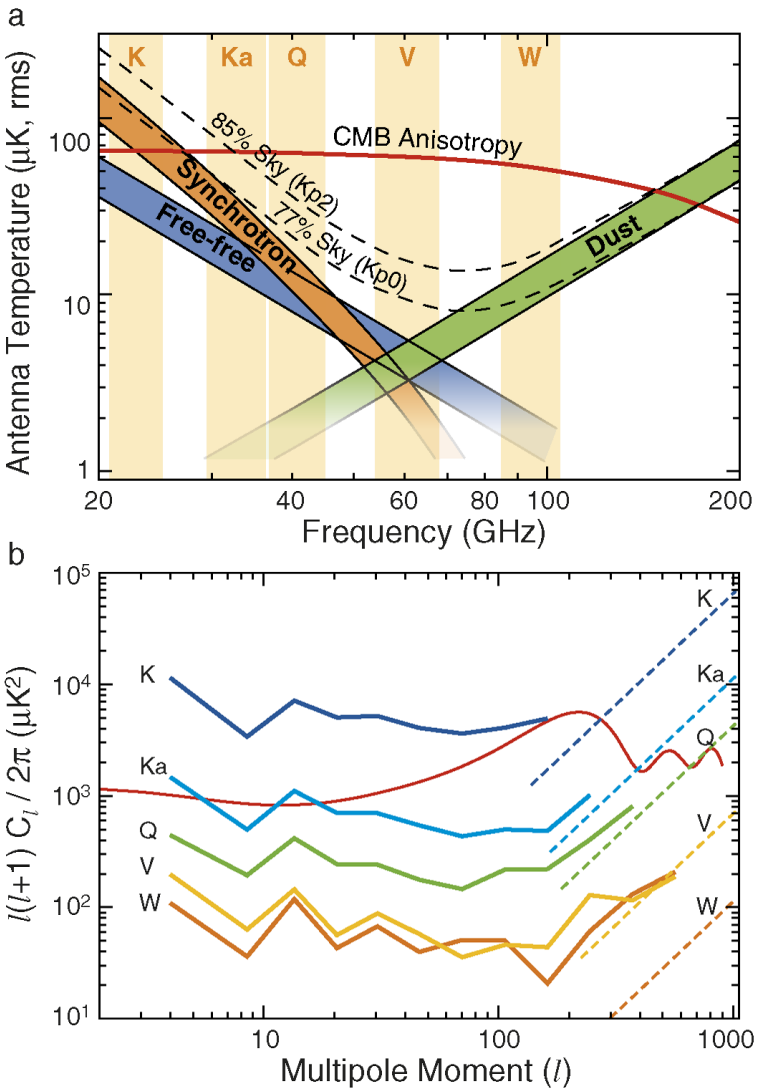


Figure 7.1: The frequency bands were chosen to be in the frequency region where the CMB anisotropy is dominant over the foreground emissions. a) Spectral shapes of the CMB blackbody temperature anisotropy and the galactic foregrounds. b) Foreground power spectra for each of the WMAP band using a Kp2 mask. Reprinted from Bennett et al. 2003 [55].

	σ_0 for each differencing Assembly									
	K1	Ka1	Q1	Q2	V1	V2	W1	W2	W3	W4
Stokes I	1.4391	1.4638	2.2449	2.1347	3.3040	2.9458	5.8833	6.5324	6.8849	6.7441

Table 7.2: Pixel noise in units of mK can be evaluated from N_{obs} with the expression $\sigma = \sigma_0/\sqrt{N_{obs}}$

at lower (radio) and higher (far-infrared) frequencies. The maps of the foreground emission are called foreground emission templates. These can be scaled to the WMAP frequencies and subtracted. The weakness of this method is the uncertainties of the external data and the scaling errors due to position-dependent spectral index variations.

The second approach is to use the different WMAP maps so that the galactic signals with specified spectra are canceled, giving us a map consisting only of CMB. This is less problematic than the first approach since the frequency-difference between the maps is smaller. Here we do not make any assumptions about the foreground signal strength, but we require knowledge of the spectra of the foregrounds. Assuming a constant spectral index for the different frequencies at which WMAP observes is less problematic than the first technique, since the frequency range is smaller. Since the technique only depends on WMAP data, it does not get systematic errors from other external experiments. But this technique does add a significant amount of extra noise the the resulting CMB map.

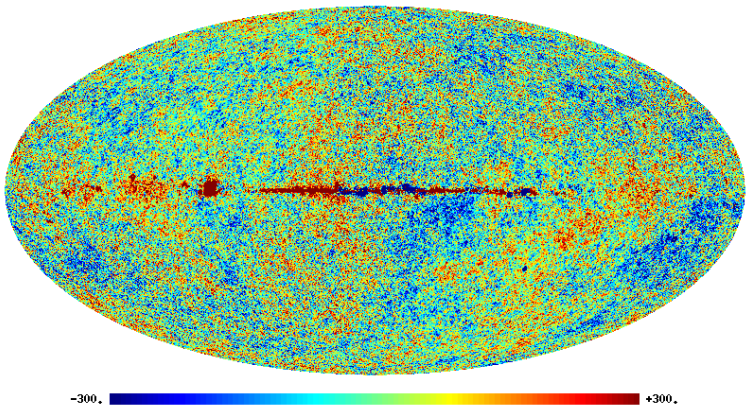
The third approach is to determine the spatial and/or spectral properties of each of the galactic emission mechanisms by performing a fit to either the WMAP data alone or in combination with external data sets.

For the analysis I will be using the ILC map and the Q-, V- and W-band maps that are foreground corrected. In the foreground corrected maps much of the foreground emission has been removed to get a better view of the clean CMB-signal. The K and Ka are not foreground reduced since they were used to produce the foreground models. Therefore I only use the maps from the Q-, V- and W-bands. A thorough explanation of the technique on how these maps were produced can be found in Hinshaw et al. (2007) [56]. The maps are given for each of the differencing assembly, and I use the average of these for each band. Figure 7.2(a) show the anisotropy map from the V-band.

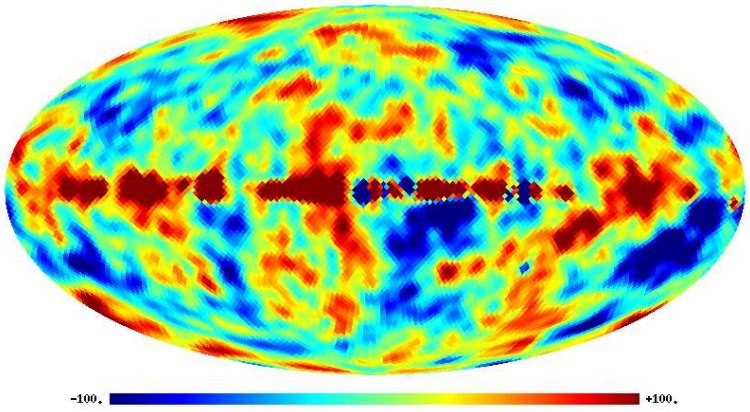
The Internal Linear Combination (ILC) map is made from 5 smoothed temperature maps to maintain the CMB signal while minimizing the Galactic foreground contribution. For angular scales larger than $\sim 10^\circ$ the ILC map provides a reliable estimate of the CMB signal with negligible instrument noise over the full sky. But for small scales there is a significant structure in the bias correction map that is still uncertain. All maps are converted into units of micro Kelvin (μK).

The satellite observes a part of the sky, i, $N_{obs,i}$ times. From this we can calculate the noise per pixel, i, by the formula $\sigma_i = \sigma_0/\sqrt{N_{obs,i}}$. Here σ_0 is the noise per observation for each channel and is given in Table 7.2.

Even though the temperature maps are foreground reduced, they still contain foreground-contamination. Most of this contamination is from our own galaxy, but



(a) Full resolution map



(b) Downgraded map at $N_{\text{side}} = 32$

Figure 7.2: The foreground reduced V-band sky map from WMAP. But even if it is foreground reduced, there are still some strong foregrounds in the galactic plane left in this map. The resolution of the map at $N_{\text{side}} = 512$ is given in 7.2(a) and shows great details in the CMB. 7.2(b) shows the same sky map downgraded to a resolution at $N_{\text{side}} = 32$. All the small structures in the map are now smudged out because of the beam, but the large scales from the high resolution map are still present. Also notice how the galactic plane gets wider when the beam smudge out the signal.

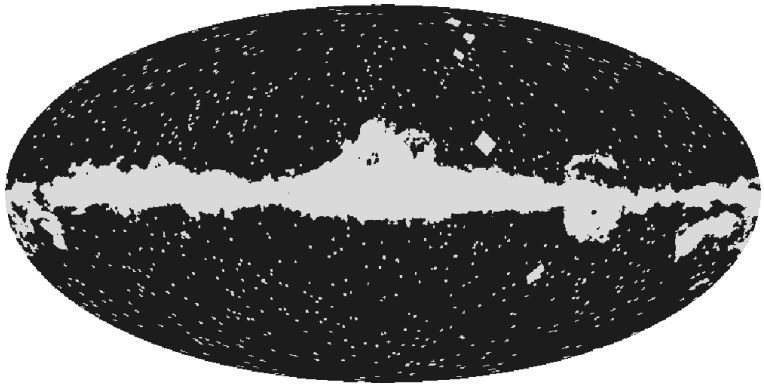


Figure 7.3: The KQ85 mask in $N_{\text{side}} = 512$ resolution. This masks out the emission from the galactic plane and 390 point sources.

some are point-sources beyond our galaxy as well. To get rid of this we apply a mask to the temperature maps to exclude the signals that are not CMB-radiation. LAMBDA provides several masks and I will be using a mask for temperature analysis called KQ85 [57]. The mask-file simply contains an array of size $12N_{side}^2$ with values of 0 and 1. If the value of a pixel is zero the point is rejected, while pixels with a value of one are accepted. This mask is given in Figure 7.3 and remove 18% of the sky.

Now what happens when we cut out certain parts of the sky with a mask? When we got an ideal full-sky coverage of the anisotropies we can at best measure $2\ell + 1$ $a_{\ell m}$s for each ℓ. As mentioned in Section 3.2, this gave rise to the cosmic variance, a fundamental uncertainty due to the fact that we can only sample $2\ell + 1$ independent modes and that we only got one sky we can observe. This is a fundamental limit on the accuracy with which we can measure the C_ℓs. When we cut away a part of the sky because of the foreground emissions, we only sample a part of the sky. We are then probing fewer independent modes and this increases the uncertainty. This is called the sampling variance. The sample variance σ^2_{sam} relates to the cosmic variance σ^2_{cos} as $\sigma^2_{\text{sam}} \simeq (4\pi/\Omega)\sigma^2_{\text{sam}}$ where Ω is the solid angle covered by the experiment. When using a Bayesian technique or a Monte-Carlo simulation to analyze the data, the sample variance is automatically accounted for.[58]

The beam functions for each differencing assembly is provided by LAMBDA. These are the Legendre transform of the beam profile.

The fiducial power spectrum is chosen to be the best-fit ΛCDM power law spectrum of Komatsu et al (2009) [59].

7.2 Downgrading data

To do this analysis in the given resolution is very computational demanding. When computing the likelihood as described above, we must invert the covariance matrix C. This operation scales as $\mathcal{O}(N_{pix}^3)$. The WMAP maps given from LAMBDA are

pixelized in the HEALPix scheme and I use the one with resolution $N_{side} = 512$. If we should do the analysis with this resolution we would get a covariance matrix that is of size $N_{pix} \times N_{pix} = \left(12 \cdot 512^2\right)^2 = 9\,895\,604\,649\,984$. This alone would take 72 TB of memory, and to calculate the determinant and the inverse of this is a tremendous task. So we need to downsize the resolution to get matrices we may be able to calculate. The basic outline for downgrading the resolution of a temperature map was given in 3.5.2.

First the map is expanded into spherical harmonics. Then it is deconvolved with the original beam function and pixel window. I choose to downgrade the original maps down to $N_{\text{side}} = 32$. The recommended beam size for this resolution is $\theta_{\text{FWHM}} = 4.5^\circ$. The map is then convolved with this beam and the corresponding pixel window. If the band consists of many differencing assemblies, then the average of the maps are taken. At last the map is added regularization noise. The downgraded version of the V-band map to a resolution at $N_{\text{side}} = 32$ is given in Figure 7.2(b).

The pure signal covariance matrix may now be written as

$$S_{ij} = \frac{1}{4\pi} \sum_{l=2}^{l_{max}} (2l+1) C_\ell (p_l b_l)^2 P_l(\cos\theta_{ij})$$

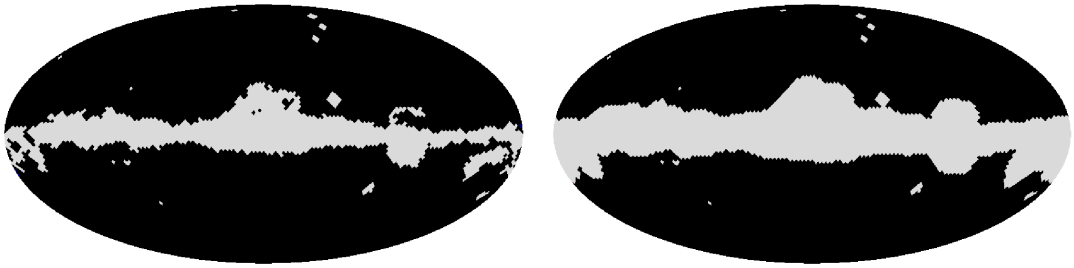
where p_ℓ is the effect of finite pixelization and it principally acts the same way as the beam. Note that we set lower boundary $\ell = 2$ since we remove the mono- and di-pole moments. The monopole cannot be found from the WMAP data since the instrument measuring the anisotropy is a differential instrument. This means that it measures the differences in temperature in two different directions, and not the absolute value of the temperature. Due to our own velocity through space we cannot observe the real dipole. The Doppler-effect caused by our movement makes it unobservable to us.

When we downgrade the temperature map, we must also downgrade our mask file to the same resolution. We know that HEALPix divides the each pixel up into four new pixels for each N_{side}. When downgrading the masks we simply need to look at the sub-pixels of a main pixel at a given resolution. Using the NESTED scheme makes it easy to find the sub-pixels that belong to a main pixel, as the ascending indices belong to the same main pixel. The link between pixels at different N_{side} is given in Figure A.1. The number of sub-pixels inside of a main pixel is given by

$$N_{\text{sub-pixels}} = \frac{12 N_{\text{side,original}}^2}{12 N_{\text{side,downgrade}}^2} = \left(\frac{N_{\text{side,original}}}{N_{\text{side,downgrade}}} \right)^2$$

I then simply add the $N_{sub-pixel}$ succeeding pixels that make up a lower resolution pixel. The pixels that now consists of more than 50 % ones is set to one.

When downgrading a temperature map, we use a beam to smooth out the signal. This causes the small scales to be washed out and smudge out the signal. Due to this smudging, the original mask used to block out any foreground contamination may be too small. We therefore need a new mask that can account for this. In a similar way as with the temperature map, we can use a beam to expand the original mask. To make the expanded mask I smooth out the map with a beam that corresponds to the wanted



(a) KQ85 mask at $N_{\text{side}} = 32$ (b) Expanded mask KQ85e

Figure 7.4: The KQ85 sky cut downgraded to $N_{\text{side}} = 32$ resolution and its extended improvement against foreground contamination, the KQ85e mask. Only the galactic plane has been extended by the Gaussian beam at $\theta_{\text{FWHM}} = 10°$, so the point sources remain the same.

resolution. If the value of the pixel is greater than 0.99 then I keep it, i.e., I set it to one. Else I set it to zero. Only the galactic plane is expanded by a Gaussian beam, while point sources are not expanded.

The KQ85 mask at $N_{\text{side}} = 32$ resolution cuts away 16.3 % of the sky, while KQ85e cuts away 26.9 %. Figure 7.4 shows the two masks.

To get an additional assurance against foreground contamination we marginalize over a fixed set of "foreground templates", $\mathbf{t}_i$, as described in Equation 3.18. One monopole and three dipole templates are used in addition to foreground templates made from the downgraded temperature maps. The ILC map is basically the clean CMB signal. This is not entirely true, but it is close. To get a map only consisting of the foregrounds I take the maps that are not foreground reduced for each frequency and subtract the signal from the ILC map. No regularization noise is added to any of these maps. The V-ILC difference template is used for both the V-band and ILC maps, the Q-ILC difference for the Q-band, and W-ILC difference for the W-band.

These templates do however not affect the results noticeably, since the mask cuts out most of the pixels that are important in this map. The template made from the V-band is given in Figure 7.5.

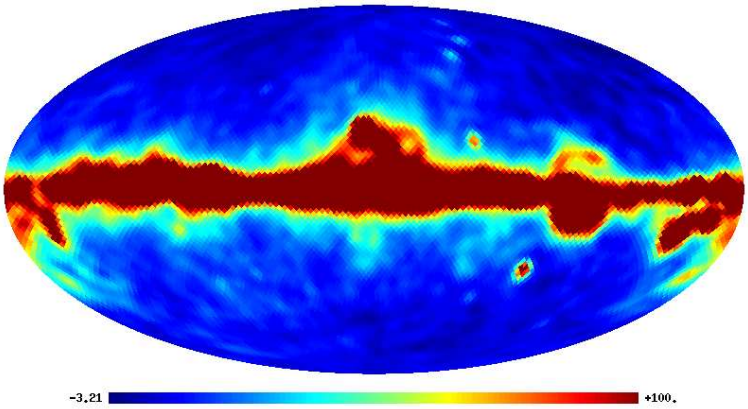


Figure 7.5: The foreground template made from the ILC map and the original V-band map that is not foreground reduced. The two maps are first downgraded to $N_{\text{side}} = 32$ with and smoothed out with a 4.5° beam. None of the maps are added the regularization noise. Then the ILC map is subtracted from the V-band map, just leaving us with the foregrounds.

Chapter 8

Results

8.1 Results

The results for the $N_{\text{side}} = 32$ analysis is given in Table 8.1. The analysis is done for four different maps (ILC, Q, V and W), using two different masks and different multipole ranges, as described in Chapter 7. The Table shows the best-fit modulation axis and amplitude, both with 68% confidence regions, and the statistical significance for the amplitude (i.e. A/σ_A) for all maps. The Bayesian log-evidence for the modulated model over the isotropic model is only calculated for the ILC map with the KQ85 mask. The reason for this is because of the computational time. It took 25 processors 679 hours each to compute the evidence for this one map.

However, there is room for improvement in this algorithm. The acceptance ratio is very low as the volume increases, about 10% is accepted. In [29] the typical acceptance ratio is found to be roughly constant at $\sim 20\% - 25\%$, after an initial period of almost 100% acceptance. What is the reason for my low acceptance ratio is unknown, but it might be the enlargement factor or the choice of the perturbation to the ellipsoid.

The evidence is closely related to the significance level A/σ_{A}, and we can therefore estimate the evidence for the other cases given the evidence and significance for ILC KQ85. The computational resources were instead focused on computing the amplitudes for each case.

The best-fit axis is the coordinate of the pixel that has been sampled over most times. The uncertainty of the best-fit axis is calculated by summing up the number of times the pixels within an angular distance radius θ_{sigma} from the best-fit axis has been sampled over. The radius is extended until the disc contain a certain percent of the samples. The results are presented with a 68% confidence region.

For the ILC map with the KQ85 mask and $\ell_{\text{mod}} = 64$, the best-fit amplitude is 0.072 ± 0.022, non-zero at the 3.3σ confidence level. The best-fit axis for the modulation field points towards the Galactic coordinates $(l, b) = (224^\circ, -22^\circ)$ with a 68% uncertainty of 24°. These results are consistent with the results presented by Eriksen et al. 2007 [24], who found an amplitude at 0.11 ± 0.04 and a best-fit axis at $(225^\circ, -27^\circ)$ for $\ell \lesssim 40$.

The results are not highly dependent on frequency. Both the V- and W-band with same mask and ℓ-range have an amplitude close to the corresponding ILC map. For the KQ85-mask and $\ell_{\rm mod} = 64$ the amplitudes of the V- and W-band are within 0.5σ of the ILC map. These are plotted in Figure 8.1(a). The distribution of the amplitude is almost equal for the three maps. The Q-band analysis using the KQ85-mask is not included here as the foregrounds clearly are visible outside the mask, and will bias the result. For the KQ85e-mask the amplitudes of the V- and W-band are within 0.1σ of the ILC map. These are plotted in Figure 8.1(b) together with the Q-band. Just as for the KQ85 case, ILC, V and W shows a remarkable consistency in the amplitude distribution. The amplitude for Q-band however is larger than these. For the Q-band case using the KQ85e-mask there are still some foregrounds visible outside the mask, and this may be the cause of the high amplitude and high σ-detection. As we get this high amplitude when using the Q-band, we may raise the question if foregrounds could be the reason for the amplitudes in the other maps as well. This can not be the case as the amplitude distribution is quite stable for the ILC, V and W maps. The intensity of the foregrounds is different for the different frequencies, and we would therefore expect the amplitude of the detected modulation field to be different as well for the different maps.

The results are not strongly dependent on the sky cut. Even if the expanded KQ85e mask cut away 10.6% more than the KQ85 mask, the amplitudes for both masks are quite consistent. The shape and peak for the amplitude distribution of ILC, V and W are very close. From Table 8.1 we can see how the uncertainty in the amplitude increases as the number of pixels are reduced. When using the liberal mask, using 83.7% of the sky, we see that the significance for all the maps are above 2σ. But when using the conservative mask, only using 73.1% of the sky, the significance of the detections are below 2σ (exept for the Q-band). This is due to the sample variance as discussed in Section 7.1. Both the amplitude and direction of the modulation field shows a consistency between the two sky cuts.

The quantities are also computed for different values of $\ell_{\rm mod}$ using the V-band data and KQ85 mask. The best-fit modulation amplitude are $A = 0.12$ for $\ell_{\rm mod} = 40$, $A = 0.08$ for $\ell_{\rm mod} = 64$ and $A = 0.07$ for $\ell_{\rm mod} = 80$. The results for Eriksen et al. 2007 [24] also had larger mean amplitude, 0.11, for $\ell \sim 40$. So it seems as the amplitude decrease with increasing values of $\ell_{\rm mod}$. This could indicate a non-scale invariant behaviour of A, as also noted by Hansen et al. 2008 [44]. But all the cases have a single common value for the amplitude that is within their uncertainties, A ~ 0.07. So even if the amplitude shows a tendency to decay with growing $\ell_{\rm mod}$, the statistical significance is so far quite low. To determine this further we need better measurements at higher $\ell_{\rm mod}$s. There may also be a tendency for increasing statistical significance with ℓ. It increases from 3.5σ at $\ell_{\rm mod} = 40$ to 3.8σ at $\ell_{\rm mod} = 64$ and 3.7σ at $\ell_{\rm mod} = 80$. The signal-to-noise ratio is unity at $\ell = 79$, and this may alter the result for $\ell_{\rm mod} = 80$.

The best-fit axes for the maps using the KQ85 mask are plotted in Figure 8.1(c) and show a close alignment of the axes. For reference, the axis that maximize the asymmetry from Eriksen et al. 2004 [12] and the best-fit axis found in Eriksen et al.

Data	Mask	ℓ_{mod}	(l_{bf}, b_{bf}) (deg)	A_{bf}	significance (σ)	$\Delta \ln \mathcal{L}$	$\Delta \ln E$
ILC	KQ85	64	$(224°, -22°) \pm 24°$	0.072 ± 0.022	3.3	7.3	2.6
V-band	KQ85	64	$(232°, -22°) \pm 23°$	0.080 ± 0.021	3.8	···	···
V-band	KQ85	40	$(224°, -22°) \pm 24°$	0.119 ± 0.034	3.5	···	···
V-band	KQ85	80	$(235°, -17°) \pm 22°$	0.070 ± 0.019	3.7	···	···
W-band	KQ85	64	$(232°, -22°) \pm 24°$	0.074 ± 0.021	3.5	···	···
ILC	KQ85e	64	$(215°, -19°) \pm 28°$	0.066 ± 0.025	2.6	···	···
Q-band	KQ85e	64	$(245°, -21°) \pm 23°$	0.088 ± 0.022	3.9	···	···
V-band	KQ85e	64	$(228°, -18°) \pm 28°$	0.067 ± 0.025	2.7	···	···
W-band	KQ85e	64	$(226°, -19°) \pm 31°$	0.061 ± 0.025	2.5	···	···

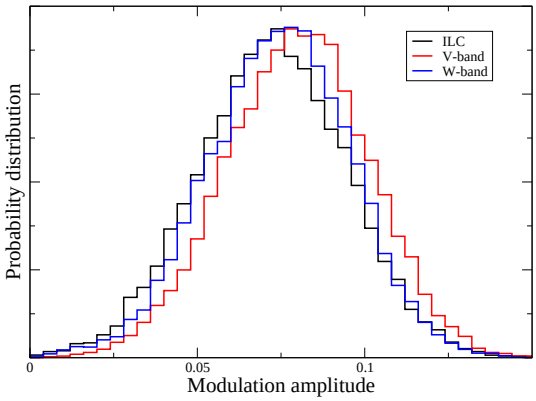
Table 8.1: Results for $N_{\rm side} = 32$. The listed quantities are data set (*first column*), mask (*second column*), maximum multipole used for modulation covariance matrix $l_{\rm mod}$ (*third column*), marginal best-fit dipole axis (*fourth column*) and amplitude (*fifth column*) with corresponding 68% confidence regions indicated, statistical significance of non-zero detection of A (*sixth column*), change in maximum likelihood between the modulated and isotropic models, $\Delta \ln \mathcal{L} = \ln \mathcal{L}_{\rm mod} - \ln \mathcal{L}_{\rm iso}$ (*seventh column*), logarithm of the Bayes factor, i.e. logarithm of the Bayesian evidence difference, $\Delta \ln E = \ln E_{\rm mod} - \ln E_{\rm iso}$ (*eighth column*). The two latter quantities are only calculated for one data set, due to the computational cost. But these values can be estimated for the other cases by comparing to the significance in the sixth column.

2007 [24] are also plotted. The coordinates θ and ϕ given in [12] are measured in the HEALPix convention, corresponding to co-latitude and longitude. It also gives the "north pole" of the dipole field. Converting it into the south pole we get that it is (l, b) = $(237°, -10°)$. In Figure 8.1(c), the axes for the cases using the KQ85e mask are plotted. The best fit axes for all the cases are fairly close to each other. Both plots of the best-fit axes use the same background of the posterior distribution of the V-band, $\ell_{\rm mod} = 80$ and KQ85 sky cut. The best-fit axis of the modulation field all point in the same direction for all the cases.

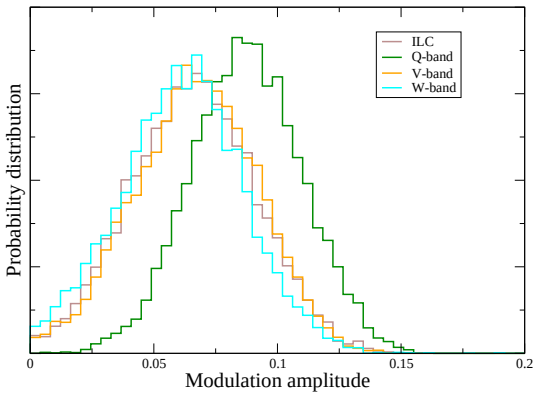
The Bayesian log-evidence difference for the ILC map is 2.6 in favour for the modulated model. This is an improvement from the analysis by Eriksen et al. 2004, where the log-evidence was ~ 1.8. The increase in evidence takes it up from "substantial" to the "strong to very strong" category on Jeffrey's scale. This corresponds to odds of 13:1.

The choice of prior volume is important when computing the evidence. Based on the distribution of the amplitude in Figure 8.1(a) and 8.1(b), the prior volume for the amplitude for the evidence computation was chosen to be 0.2. All the amplitudes of the ILC, V and W drops to zero at 0.15. However the amplitude of the Q-band stretches above 0.15, hence the prior amplitude is chosen to be 0.2. The prior volume for the direction $\hat{p}$ is the whole area of the sphere, i.e. $\theta = (0, \pi)$ and $\phi = (0, 2\pi)$. The range for the two power spectrum parameters are $0.5 \leq Q \leq 1.5$ and $-0.5 \leq n \leq 0.5$.

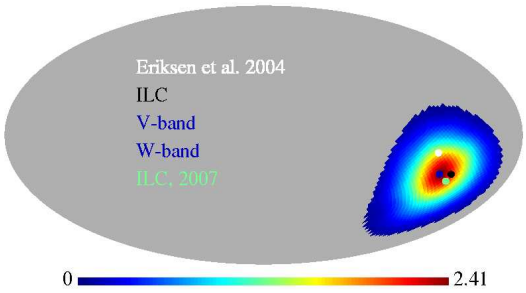
There are also other ways to test what model is preferred besides the Bayesian evid-



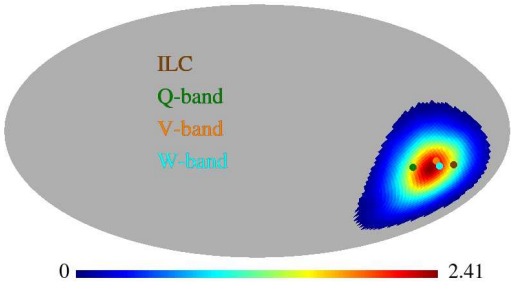
(a) Amplitude distribution using the KQ85 sky cut



(b) Amplitude distribution using the KQ85e sky cut



(c) Best-fit axes using the KQ85 sky cut



(d) Best-fit axes using the KQ85e sky cut

Figure 8.1: The two upper figures show the posterior distribution of the modulation amplitude marginalized over all other parameters. The dots in the two lower figures show the best-fit axes for the dipole modulation field. The two figures on the left show the distribution when using the KQ85 mask, while the two on the right are for the KQ85e mask. The background in the two lower figures show the posterior distribution for the V-band map using the KQ85-mask and $\ell_{\rm mod} = 64$. It has been smoothed out with a beam for plotting purposes to reduce the pattern caused by Monte Carlo simulations. The grey area indicates pixels outside the 2σ confidence region. All of my results in the figures are for $\ell_{\rm mod} = 64$.

ence. We can calculate the frequentist probability of obtaining such a large amplitude as observed in the WMAP data for simulated, isotropic universes. As I mentioned before, the asymmetric distribution may be a realization of an isotropic universe. To check this, we may run the same analysis for thousands of simulated, isotropic maps and see how many of these that gives an modulation amplitude as large as the one given by the WMAP data. In the article by Eriksen et al. 2007 [24] for $N_{\text{side}} = 16$, and he found that the frequentist probability for getting such a high amplitude is 0.9-1.5% depending on data-set and sky-cut. In this thesis I unfortunately didn't have time to do a similar test for the $N_{\text{side}} = 32$ analysis since it demands thousands of simulated maps to get reliable results. As I pointed out before, it takes a lot of time to calculate the likelihood at this resolution, and a similar test is therefore omitted here.

Part IV

Conclusion and summary

Chapter 9

Conclusions

9.1 A short review of the thesis

After the first-year release of the WMAP data in 2003, there were reports on violation of statistical isotropy. This seems troublesome as it conflicts with current theories about the universe. The simplest inflation theories predicts that the universe should be statistically homogeneous and isotropic. This theory is highly successful in describing the observations, such as the CMB and large-scale power spectra, and we would therefore expect isotropy and homogeneity in the universe as well. But many analysis of the the CMB maps from WMAP show hints of violation of statistical isotropy and non-Gaussianity. Several theories have been developed in order to describe these features, but none of them are yet completely satisfactory.

One of the early reports on the violation of isotropy was an asymmetry in the distribution of fluctuation power in the CMB on large scales. Eriksen et al. (2004) [12] and Hansen et al. 2004 [13] both presented an early evidence for a dipolar distribution of power in the CMB temperature anisotropy sky, on scales up to $\ell \sim 63$ and 40, respectively.

In a paper by Eriksen et al. 2007 [24], they presented further evidence for a dipolar distribution of fluctuation power in the CMB. By introducing a model with an isotropic CMB sky modulated by a dipolar field using a Bayesian framework, they could quantify the large-scale power asymmetry. This model gave a better fit to the data than the isotropic model for $\ell \lesssim 40$.

A possibly physically related work were presented in Groeneboom & Eriksen (2009) [51], in which they present evidence for a quadrupolar distribution power, that extended over the whole range $\ell = 2 - 400$. Hansen et al. 2008 [44] showed that the dipolar power distribution is present on scales up to $\ell = 600$. There appears to be increasing evidence for both a dipolar and quadrupolar structure in the CMB power distribution on all scales. Based on this claim, I adopt the modulated CMB model presented in Eriksen et al. 2007 [24] and apply it to the five-year WMAP data and go to higher multipoles, even though far lower than those investigated by Hansen et al. 2008 [44].

Compared to the analysis done by Eriksen et al. 2007 [24] the statistical significance

of a non-zero amplitude increases from 2.8σ to 3.3σ when increasing ℓ from 40 to 64. The statistical significance increases from 3.5σ at $\ell_{\rm mod} = 40$ to 3.8σ at $\ell_{\rm mod} = 64$ and 3.7σ at $\ell_{\rm mod} = 80$.

So there is substantial evidence for both a dipolar and quadrupolar power distribution in the WMAP data. The magnitude of the dipolar mode is stronger than the quadrupolar mode, as a $\sim$3.5 significance level is already found at $\ell \sim 64$ for the dipole, while the same significance level is found for the quadrupole at $\ell \sim 400$.

The Bayesian evidence also increase with ℓ. In [24] the Bayesian evidence difference between the modulated model and the isotropic is $\Delta \ln E = 1.8 \pm 0.2$ for $\ell \sim 40$ using the ILC map. Here, using the ILC map for $\ell = 64$, the evidence has increased to $\Delta \ln E = 2.6$.

The Bayesian evidence found here at 2.6 ranks as "strong to very strong" on Jeffrey's scale. But the Bayesian evidence is strongly dependent on the choice of priors. Choosing a prior twice as large as the only actually used, will cause the log-evidence to drop, ranking only as "substantial". Because of this dependence it is in may respects easier to get a firm statistical interpretation to the posterior distribution than the Bayesian evidence.

As the amplitude of the Q-band is higher than the other maps, it can seem like the foregrounds are the reason for the increase. The choice of prior volume can therefor be chosen to be smaller, A = [0, 0.15]. This can cause an increase in the evidence.

The results also show hints that the amplitude A decrease with ℓ. Whether this is due to a statistical fluctuation or that a physical effect cause this non-scale invariant behaviour is unknown. The data allow both possibilities. Looking at the V-band data using the KQ85 mask, a common amplitude A ~ 0.07 is within the error bars of the amplitude for the all values of ℓ. If the case is that the amplitude is decreasing with ℓ we should replace the amplitude A with the function $A(\ell)$.

9.2 Problems for future work

The main problem for future work is to be able to do compute the amplitude and direction for higher values of ℓ. Then we can find out for how small scales this dipole behavior is seen, and if the amplitude decrease with ℓ or if it is scale-invariant.

The approach in this thesis relies on a brute force inversion of a $N_{\rm pix} \times N_{\rm pix}$ covariance matrix which scales as $\mathcal{O}(N_{\rm pix}^3)$ or $\mathcal{O}(N_{\rm side}^6)$. The total use of CPU hours for the $N_{\rm side} = 32$ analysis was $\sim 50\,000$. The next supported resolution by HEALPix is with $N_{\rm side} = 64$. By increasing $N_{\rm side}$ with a factor of two would require a total of ~ 3 million CPU hours. Clearly, new and more efficient algorithms are needed.

The computation of the parameters amplitude and direction of the dipole modulation field would be possible to compute with ScaLAPACK in a reasonable amount of time. But the evidence calculation would take a lot of time. However, to compute the Bayesian evidence may not be necessary as it is in many respects easier to attach a firm statistical confirmation to the posterior distribution. If to use the brute force method described in this thesis, there are some things that could further reduce the computa-

tional time. One can use more parallel processors that work together on one matrix and optimize the block-size for the number of processors. The use of multiple-try Metropolis algorithm could also reduce the computational time. Maybe the best method for this problem is by using another MCMC method called Gibbs sampling. This technique of mapping out the posterior does not depend on the brute force calculation of the likelihood. It scales as $\mathcal{O}(N_{\rm pix}^{3/2})$ or $\mathcal{O}(N_{\rm pix}^{2})$ for data with white or correlated noise, respectively.

These results may be useful for theorists trying to construct models for the observed asymmetric distribution of power across the sky, either phenomenological or fundamental. Many attempts of creating such theories have already been done (e.g., Ackerman et al. 2007 [50], Böhmer & Mota 2008 [60]; Carroll et al. 2008 [61], 2009 [62]; Chang et al 2009 [63]; Erickcek et al 2008 [64][65]; Gordon et al. 2005 [32]; Gümrükcüoglu et al. 2007 [48]; Himmetoglu et al. 2009 [66][67]; Kahniashvili et al. 2008 [68]; Kanno et al. 2008 [69]; Koivisto & Mota 2008 [53][70]; Pereira et al. 2007 [71]; Pitrou et al. 2008 [72]; Pullen & Kamionkowski 2007 [49]; Watanabe et al 2009 [73]; Yokoyama & Soda 2008 [74]), but none have so far been fully convincing. More work is needed on both the theoretical and observational side.

The successful launch of Planck on 27th of May 2009 also means that we can get new data within a few years. This will measure the temperature anisotropies down to arcminute scales. Using the additional frequency coverage, we can get better foreground treatment and a larger sky coverage can be achieved. Full-sky and high-sensitivity polarisation maps can also provide useful information on the origin of the effects. This can make it possible to measure the properties of the dipole, quadrupole and, possibly higher-order modes of the modulation field to a much better accuracy than with today's data.

Bibliography

[1] A. Einstein. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.*, pages 844–847, 1915.

[2] G. Lematre. Un Univers homogne de masse constante et de rayon croissant rendant compte de la vitesse radiale des nbuleuses extra-galactiques. *Annales de la Socit Scientifique de Bruxelles*, 47:49, april 1927.

[3] E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929.

[4] G. Lemaître. The Beginning of the World from the Point of View of Quantum Theory. *Nature*, 127:706–+, May 1931.

[5] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. *ApJ*, 142:414–419, July 1965.

[6] A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *ApJ*, 142:419–421, July 1965.

[7] E. Torbet, M. J. Devlin, W. B. Dorwart, T. Herbig, A. D. Miller, M. R. Nolta, L. Page, J. Puchalla, , and H. T. Tran. A measurement of the angular power spectrum of the microwave background made from the high chilean andes. *The Astrophysical Journal Letters*, 521(2):L79–L82, 1999.

[8] G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, N. Jarosik, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, M. Halpern, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Data Processing, Sky Maps, and Basic Results. *The Astrophysical Journal Supplement*, 180:225–245, February 2009.

[9] M. Tegmark, A. de Oliveira-Costa, and A. J. Hamilton. High resolution foreground cleaned CMB map from WMAP. *Phys. Rev. D*, 68(12):123523–+, December 2003.

[10] A. de Oliveira-Costa, M. Tegmark, M. Zaldarriaga, and A. Hamilton. Significance of the largest scale CMB fluctuations in WMAP. *Phys. Rev. D*, 69(6):063516–+, March 2004.

[11] H. K. Eriksen, A. J. Banday, K. M. Górski, and P. B. Lilje. On Foreground Removal from the Wilkinson Microwave Anisotropy Probe Data by an Internal Linear Combination Method: Limitations and Implications. *ApJ*, 612:633–646, September 2004.

[12] H. K. Eriksen, F. K. Hansen, A. J. Banday, K. M. Górski, and P. B. Lilje. Asymmetries in the Cosmic Microwave Background Anisotropy Field. *ApJ*, 605:14–20, April 2004.

[13] F. K. Hansen, A. J. Banday, and K. M. Górski. Testing the cosmological principle of isotropy: local power-spectrum estimates of the WMAP data. *MNRAS*, 354:641–665, November 2004.

[14] P. Vielva, E. Martínez-González, R. B. Barreiro, J. L. Sanz, and L. Cayón. Detection of Non-Gaussianity in the Wilkinson Microwave Anisotropy Probe First-Year Data Using Spherical Wavelets. *ApJ*, 609:22–34, July 2004.

[15] M. Cruz, E. Martínez-González, P. Vielva, and L. Cayón. Detection of a non-Gaussian spot in WMAP. *MNRAS*, 356:29–40, January 2005.

[16] S. Dodelson. *Modern Cosmology*. Academic Press, 2003.

[17] M. Tegmark. Doppler peaks and all that. In *Enrico Fermi, Course CXXXII, Varenna*, 1995. astro-ph/9511148.

[18] Ø. Elgarøy. Ast4220: Cosmology i. Lecture notes for the course AST4220: Cosmology I, fall 2008, 2008.

[19] A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23:347–356, January 1981.

[20] A. D. Linde. A new inflationary Universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108:389–393, 1982.

[21] A. Albrecht and P. J. Steinhardt. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Physical Review Letters*, 48:1220–1223, April 1982.

[22] J. Silk. Cosmic Black-Body Radiation and Galaxy Formation. *ApJ*, 151:459–+, February 1968.

[23] A. J. Banday, S. Zaroubi, and K. M. Górski. On the Non-Gaussianity Observed in the COBE Differential Microwave Radiometer Sky Maps. *ApJ*, 533:575–587, April 2000.

[24] H. K. Eriksen, A. J. Banday, K. M. Gorski, F. K. Hansen, and P. B. Lilje. Hemispherical power asymmetry in the three-year Wilkinson Microwave Anisotropy Probe sky maps. *ApJ*, 660:L81–L84, 2007. arXiv:astro-ph/0701089.

[25] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21:1087–1092, June 1953.

[26] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[27] Harold Jeffreys. *Theory of probability*. Oxford University Press, 3rd edition edition, 1998, c1961.

[28] J. Skilling. Nested Sampling. In R. Fischer, R. Preuss, and U. V. Toussaint, editors, *American Institute of Physics Conference Series*, volume 735 of *American Institute of Physics Conference Series*, pages 395–405, November 2004.

[29] P. Mukherjee et al. A Nested Sampling Algorithm For Cosmological Model Selection. *ApJ*, 638:51–54, 2006.

[30] Mary Kathryn Cowles and Bradley P. Carlin. Markov chain monte carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, 91:883–904, 1996.

[31] A. Gelman and D. Rubin. Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.*, 7(4):457–472, 1992.

[32] C. Gordon, W. Hu, D. Huterer, and T. Crawford. Spontaneous isotropy breaking: A mechanism for CMB multipole alignments. *Physical Review D*, 72(10):103002–+, November 2005.

[33] A. Bernui, T. Villela, C. A. Wuensche, R. Leonardi, and I. Ferreira. On the cosmic microwave background large-scale angular correlations. *A&A*, 454:409–414, August 2006.

[34] P. Bielewicz, H. K. Eriksen, A. J. Banday, K. M. Górski, and P. B. Lilje. Multipole Vector Anomalies in the First-Year WMAP Data: A Cut-Sky Analysis. *ApJ*, 635:750–760, December 2005.

[35] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman. On the large-angle anomalies of the microwave sky. *MNRAS*, 367:79–102, March 2006.

[36] H. K. Eriksen, D. I. Novikov, P. B. Lilje, A. J. Banday, and K. M. Górski. Testing for Non-Gaussianity in the Wilkinson Microwave Anisotropy Probe Data: Minkowski Functionals and the Length of the Skeleton. *ApJ*, 612:64–80, September 2004.

[37] H. K. Eriksen, A. J. Banday, K. M. Górski, and P. B. Lilje. The N-Point Correlation Functions of the First-Year Wilkinson Microwave Anisotropy Probe Sky Maps. *ApJ*, 622:58–71, March 2005.

[38] T. R. Jaffe, A. J. Banday, H. K. Eriksen, K. M. Górski, and F. K. Hansen. Evidence of Vorticity and Shear at Large Angular Scales in the WMAP Data: A Violation of Cosmological Isotropy? *ApJ*, 629:L1–L4, August 2005.

[39] T. R. Jaffe, A. J. Banday, H. K. Eriksen, K. M. Górski, and F. K. Hansen. Bianchi type VIIh models and the WMAP 3-year data. *A & A*, 460:393–396, December 2006.

[40] E. Martínez-González, M. Cruz, L. Cayón, and P. Vielva. The non-Gaussian cold spot in WMAP. *New Astronomy Review*, 50:875–879, December 2006.

[41] J. D. McEwen, M. P. Hobson, A. N. Lasenby, and D. J. Mortlock. A high-significance detection of non-Gaussianity in the WMAP 5-yr data using directional spherical wavelets. *MNRAS*, 388:659–662, August 2008.

[42] C. Räth, P. Schuecker, and A. J. Banday. A scaling index analysis of the Wilkinson Microwave Anisotropy Probe three-year data: signatures of non-Gaussianities and asymmetries in the cosmic microwave background. *MNRAS*, 380:466–478, September 2007.

[43] A. P. S. Yadav and B. D. Wandelt. Evidence of Primordial Non-Gaussianity (f_{NL}) in the Wilkinson Microwave Anisotropy Probe 3-Year Data at 2.8σ. *Physical Review Letters*, 100(18):181301, May 2008.

[44] F. K. Hansen, A. J. Banday, K. M. Gorski, H. K. Eriksen, and P. B. Lilje. Power Asymmetry in Cosmic Microwave Background Fluctuations from Full Sky to Sub-degree Scales: Is the Universe Isotropic? *ArXiv e-prints*, December 2008. submitted to ApJ.

[45] G. Hinshaw, C. Barnes, C. L. Bennett, M. R. Greason, M. Halpern, R. S. Hill, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, L. Page, D. N. Spergel, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing Methods and Systematic Error Limits. *The Astrophysical Journal Supplement*, 148:63–95, September 2003.

[46] C.L. Bennett et al. The Microwave Anisotropy Probe (MAP) Mission. *ApJ*, 583:1–23, 2003.

[47] D. P. Finkbeiner. Microwave Interstellar Medium Emission Observed by the Wilkinson Microwave Anisotropy Probe. *ApJ*, 614:186–193, October 2004.

[48] A. Emir Gümrükçüoglu, C. R. Contaldi, and M. Peloso. Inflationary perturbations in anisotropic backgrounds and their imprint on the cosmic microwave background. *Journal of Cosmology and Astro-Particle Physics*, 11:5–+, November 2007.

[49] A. R. Pullen and M. Kamionkowski. Cosmic microwave background statistics for a direction-dependent primordial power spectrum. *Physical Review D*, 76(10):103529–+, November 2007.

[50] Lotty Ackerman, Sean M. Carroll, and Mark B. Wise. Imprints of a Primordial Preferred Direction on the Microwave Background. *Physical Review*, D75, 2007. 083502.

[51] Nicolaas E. Groeneboom and Hans Kristian Eriksen. Bayesian analysis of sparse anisotropic universe models and application to the 5-yr WMAP data. *ApJ*, submitted, 2008. arXiv:0807.2242.

[52] R. A. Battye and A. Moss. Anisotropic perturbations due to dark energy. *Physical Review D*, 74(4):041301, August 2006.

[53] T. Koivisto and D. F. Mota. Accelerating Cosmologies with an Anisotropic Equation of State. *ApJ*, 679:1–5, May 2008.

[54] C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright, C. Barnes, M. R. Greason, R. S. Hill, E. Komatsu, M. R. Nolta, N. Odegard, H. V. Peiris, L. Verde, and J. L. Weiland. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. *The Astrophysical Journal Supplement Series*, 148:1–27, September 2003.

[55] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission. *The Astrophysical Journal Supplement*, 148:97–117, September 2003.

[56] G. Hinshaw, M. R. Nolta, C. L. Bennett, R. Bean, O. Doré, M. R. Greason, M. Halpern, R. S. Hill, N. Jarosik, A. Kogut, E. Komatsu, M. Limon, N. Odegard, S. S. Meyer, L. Page, H. V. Peiris, D. N. Spergel, G. S. Tucker, L. Verde, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature Analysis. *ApJS*, 170:288–334, June 2007.

[57] B. Gold, C. L. Bennett, R. S. Hill, G. Hinshaw, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, J. Dunkley, M. Halpern, N. Jarosik, A. Kogut, E. Komatsu, D. Larson, S. S. Meyer, M. R. Nolta, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Galactic Foreground Emission. *ApJS*, 180:265–282, February 2009.

[58] Douglas Scott, Mark Srednicki, and Martin White. "Sample Variance" in Small-Scale CMB Anisotropy Experiments. *ApJ*, 421:L5, 1994.

[59] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation. *ApJS*, 180:330–376, February 2009.

[60] C. G. Böhmer and D. F. Mota. CMB anisotropies and inflation from non-standard spinors. *Physics Letters B*, 663:168–171, May 2008.

[61] S. M. Carroll, C.-Y. Tseng, and M. B. Wise. Translational Invariance and the Anisotropy of the Cosmic Microwave Background. *ArXiv e-prints*, November 2008.

[62] S. M. Carroll, T. R. Dulaney, M. I. Gresham, and H. Tam. Instabilities in the aether. *Physical Review D*, 79(6):065011–+, March 2009.

[63] S. Chang, M. Kleban, and T. S. Levi. Watching worlds collide: effects on the CMB from cosmological bubble collisions. *Journal of Cosmology and Astro-Particle Physics*, 4:25–+, April 2009.

[64] A. L. Erickcek, S. M. Carroll, and M. Kamionkowski. Superhorizon perturbations and the cosmic microwave background. *Physical Review D*, 78(8):083012–+, October 2008.

[65] A. L. Erickcek, M. Kamionkowski, and S. M. Carroll. A hemispherical power asymmetry from inflation. *Physical Review D*, 78(12):123520–+, December 2008.

[66] B. Himmetoglu, C. R. Contaldi, and M. Peloso. Instability of the Ackerman-Carroll-Wise model, and problems with massive vectors during inflation. *Physical Review D*, 79(6):063517–+, March 2009.

[67] B. Himmetoglu, C. R. Contaldi, and M. Peloso. Instability of Anisotropic Cosmological Solutions Supported by Vector Fields. *Physical Review Letters*, 102(11):111301–+, March 2009.

[68] T. Kahniashvili, G. Lavrelashvili, and B. Ratra. CMB temperature anisotropy from broken spatial isotropy due to a homogeneous cosmological magnetic field. *Physical Review D*, 78(6):063012–+, September 2008.

[69] S. Kanno, M. Kimura, J. Soda, and S. Yokoyama. Anisotropic inflation from vector impurity. *Journal of Cosmology and Astro-Particle Physics*, 8:34–+, August 2008.

[70] T. Koivisto and D. F. Mota. Vector field models of inflation and dark energy. *Journal of Cosmology and Astro-Particle Physics*, 8:21–+, August 2008.

[71] T. S. Pereira, C. Pitrou, and J.-P. Uzan. Theory of cosmological perturbations in an anisotropic universe. *Journal of Cosmology and Astro-Particle Physics*, 9:6–+, September 2007.

[72] C. Pitrou, T. S. Pereira, and J.-P. Uzan. Predictions from an anisotropic inflationary era. *Journal of Cosmology and Astro-Particle Physics*, 4:4–+, April 2008.

[73] M.-A. Watanabe, S. Kanno, and J. Soda. Inflationary Universe with Anisotropic Hair. *Physical Review Letters*, 102(19):191302–+, May 2009.

[74] S. Yokoyama and J. Soda. Primordial statistical anisotropy generated at the end of inflation. *Journal of Cosmology and Astro-Particle Physics*, 8:5–+, August 2008.

[75] K. M. Gorski, B. D. Wandelt, F. K. Hansen, E. Hivon, and A. J. Banday. The HEALPix Primer. *ArXiv Astrophysics e-prints*, May 1999.

Part V

Appendices

Appendix A

An introduction to the HEALPix software

In this thesis I have been using the HEALPix pixelization. Here is a short introduction to the software package. It is mostly based on 'A HEALPix Primer" by Gorski et al. [75].

A.1 Introduction

HEALPix is a data structure with an associated library of computational algorithms and visualization for discretized spherical maps. It has an efficient discretization and offers a fast analysis of functions defines on the sphere. To be able to analyse the information on a computer, it must discretized in some way. We must partition the sphere into finite area elements. There are several ways one can do this, but a smart choice of partitioning can give us many advantages. First of all the scheme should give small systematic errors.

HEALPix is an acronym for Hierarchical, Equal Area, iso-Latitude Pixelisation. These properties makes it well fitted for fast and efficient calculations of functions on the sphere.

1. Hierarchical structure of the database. Elements that are nearby on the sphere are also nearby in the tree structure of the database. Hence, the elements are optimally placed in the storage medium for fast access between neighboring pixels Figure A.1 shows a hierarchical partition with quadrilateral structure.

2. All pixels have the same area at a given resolution. This is advantageous because white noise generated by the instrument gets integrated exactly into white noise in the pixel space. Sky signals are sampled without regional dependence.

3. The center of the pixel lie on rings of equal latitude. The advantage of this property is the computational speed in all operations of spherical harmonics. Associated Legendre polynomials are evaluated via slow recursions, and if the

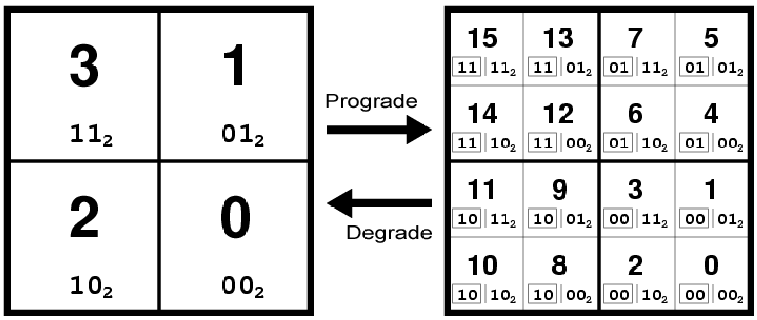


Figure A.1: This hierarchical partition has a quadrilateral tree pixel numbering. To increase the resolution, each pixel is split into four sub-pixels. The structure allows for a binary vector indexing of the database. The sub-pixels inherit the index of the original (in the box).

sampling grid deviates from an iso-latitude distribution will cause a loss of computational performance with the growing number of sampling points. This makes it impractical for high resolution data.

Figure A.1 shows a hierarchical partition with quadrilateral structure.

A.2 The HEALPix pixelization

HEALPix consists of twelve base-pixels in three rings around the poles and equator. Each of these pixels can be divided into smaller ones in order to achieve a higher resolution. Pixels are divided into four new pixels as showed in Figure A.1. The resolution is parameterized by N_{side} which gives the number of divisions along the edge of the base-pixel. , $N_{side} = i^2$ where i is an integer. The total number of pixels on the sphere is given by $N_{pix} = 12N_{side}^2$. The pixels are placed on $4N_{side} - 1$ rings of constant latitude.

To identify each pixel, HEALPix have two different numbering schemes.The first is the RING scheme, where the pixels are numbered from the north to the south, going along each iso-latitude ring. The second one is the NESTED scheme, where the pixel indices are arranged in twelve tree structures corresponding to the base-resolution pixels.

A.3 The software

The HEALPix software package features Fortran90 and IDL source codes. It uses the standardised FITS I/O interface for easy reading and writing. These contains many routines for various calculations, that can easily be linked to your own applications.

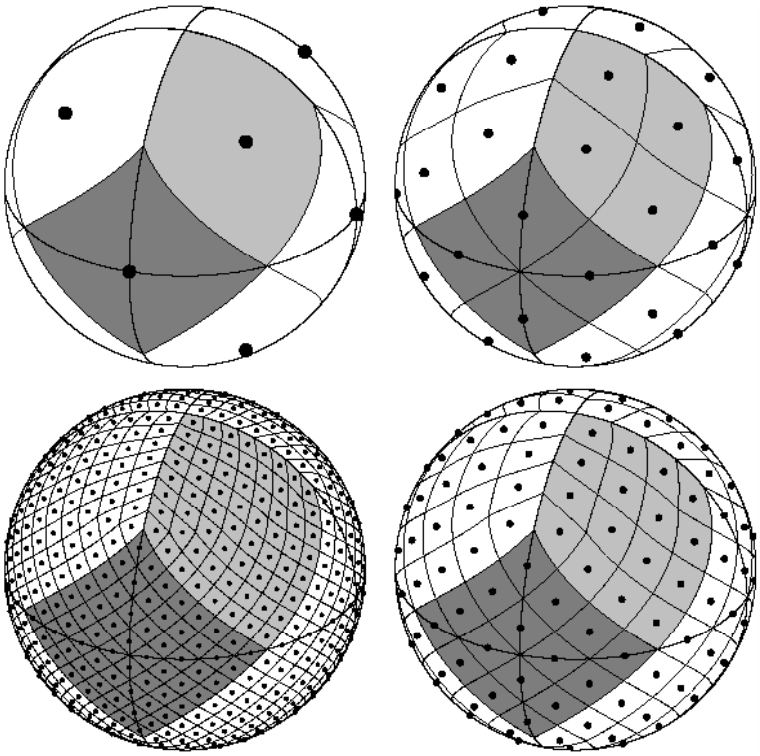


Figure A.2: Shows four different resolutions in the HEALPix scheme. The upper left panel shows the base-resolution pixels, i.e. $N_{side} = 1$. Moving in a clockwise fashion then shows how the base-pixels are hierarchically subdivided into the resolution of N_{side} = 2, 4 and 8. The grey areas shows two base-pixels.

The package also features applications for different purposes. These are:alteralm (alters the a_{lm}'s), anafast (creates temperature and/or polarisation maps as realisations of random Gaussian fields on a sphere characterized by a theoretical power spectra), hotspot (finds the local maximum and minimum in a map), map2gif (is a very good program for turning a HEALPix map into a gif image), median_filter, plmgen, smoothing (convolves a map with a Gaussian beam), synfast (makes a simulated HEALPix map, based on a theoretical power spectrum) and ud_grade (translate a map from one $N_{\rm side}$ to another).

Appendix B

An introduction to ScaLAPACK

B.1 Introduction

ScaLAPACK is a library for solving linear equations, least squares problems, eigenvalue problems and singular value problems. The name ScaLAPACK stands for Scalable Linear Algebra PACKage, or Scalable LAPACK, Some of the goals for the library is efficiency (to run as fast as possible) and scalability (to be able to run well as the problem's size and the number of processors increase). The interface is very similar to LAPACK to make it easy to use. The advantage with ScaLAPACK is that it could be parallelized in a manner where many processors can collaborate in solving a matrix. This will reduce the time for solving a linear algebra set. Each processor is given parts of a matrix and only solves it's part. To make this run efficiently, the matrix is distributed in a block-cyclic manner.

There are four basic steps required to call a ScaLAPACK routine.

1. Initialize the process grid
2. Distribute the matrix on the process grid
3. Call the ScaLAPACK routine
4. Release the process grid

B.2 Initialization of the process grid

B.2.1 BLACS

First we need to initialize the process grid. We want to use NPROW $\times$ NPCOL processors. ScaLAPACK uses BLACS (Basic Linear Algebra Communication Subprograms) for communication between the processors. Initiating this grid is done by calling the routine BLACS_GRIDINIT(ICONTXT, ORDER, NPROW, NPCOL). Here ICONTXT is a number that identifies the process grid. This way the grid can safely communicate even if some other process grid consists of the same processors. ORDER

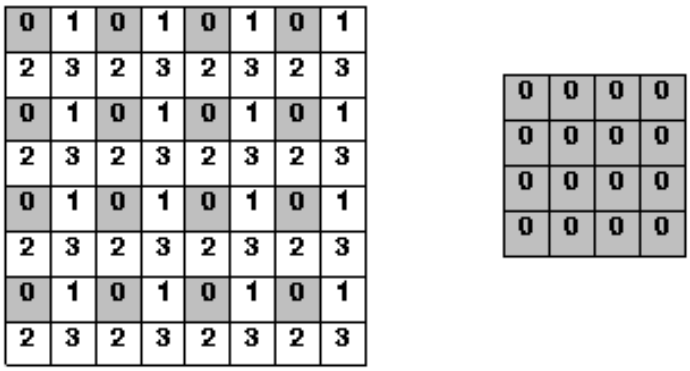


Figure B.1: How the matrix is distributed in a block-cyclic manner. The process grid is 2×2 and 0-3 represents the process number. Left: The global matrix is partitioned into blocks and these are distributed to the different processors. Right: The local matrix of processor 0.

tell if the process grid should be arranged in a row- or column-major way. The user can get the coordinates of the processes by the call BLACS_GRIDINFO(ICTXT, NPROW, NPCOL, MYROW, MYCOL).

B.3 Distributing the matrix

The global matrix must be distributed on the process grid. The ScaLAPACK routines assume that the global matrix has been distributed on the processors. The ScaLAPACK software library provides routines that operate on three types of matrices: in-core dense matrices, in-core narrow band matrices and out-of-core dense matrices. I will only look at in-core dense matrices here, since these are the ones I must solve.

B.3.1 Block cyclic distribution

Like LAPACK, the ScaLAPCK routines are based on block-partitioned algorithms in order to minimize the frequency of data between different levels of the memory hierarchy. A wide variety of distribution schemes have potential for achieving high performance for dense matrix computations. ScaLAPACK has chosen the block-cyclic data layout due to its scalability, load balance and communication properties. The global matrix is first divided into blocks of size $MB \times NB$. Each block is then distributed on the process grid. Figure B.1 shows how the blocks are distributed to the different processors.

B.3.2 Local storage scheme

The distribution of the matrix on the process grid is the user's responsibility. The following scheme gives the coupling between the coordinates in the global matrix and the local matrix of the different processes. A two-dimensional global matrix A with dimension $M_A \times N_A$ is to be distributed to different processes. This matrix is partitioned in blocks of size MB × NB, and the first block is given to the process with coordinates (RSRC, CSRC). The processor grid is of dimension $P_r \times P_c$, and the coordinates in this grid is (p_r, p_c). The matrix entry (I,J) of matrix A is distributed to the process with coordinates (p_r, p_c) local matrix within the (l,m) block at the position (x,y) given by:

$$
\begin{aligned}
(l,m) &= \left(\frac{I-1}{P_r MB}, \frac{J-1}{P_c nB}\right) \\
(p_r,p_c) &= \left(mod\left(RSRC + \frac{I-1}{MB}, P_r\right), mod\left(CSRC + \frac{J-1}{NB}, P_c\right)\right) \\
(x,y) &= (mod(I-1,MB)+1, mod(J-1,NB)+1)
\end{aligned}
\tag{B.1}
$$

After distributing the matrix it now need an array descriptor. This contains information on the dimension of the global matrix, the block size, what processor to receive the first block of the global matrix and the leading dimension of the local matrix. The array descriptor is called via DESCINIT(DESCA, M, N, MB, NB, RSRC, CSRC, ICTXT, MXLLDA, INFO).

- DESCA - the array containing information on the matrix and processors
- M - Number of rows in the global matrix
- N - Number of columns in the global matrix
- MB - Block size used to partition the rows
- NB - Block size used to partition the columns
- RSRC - Process row on which the first row in the global matrix is distributed
- CSRC - Process column on which the first row in the global matrix is distributed
- ICTXT - Context number identifying the grid on which the matrix is distributed
- MXLLDA - Leading dimension of the local matrix, $MXLLDA \geq MAX(1, LOC_r(M))$
- INFO - Information about the distribution. Equals zero upon successful exit.

The dimensions of the local matrix can be calculated with the ScaLAPACK tool function NUMROC. The row dimension is $LOC_r(M) = NUMROC(M, MB_A, MYROW, RSRC_A, NPRO$ and the column dimension is $LOC_c(N) = NUMROC(N, NB_A, MYCOL, CSRC_A, NPCOL)$

If the matrix dimension does not divide the block factor evenly, then last block will contain fewer elements. To make it more clear how it works, I'll give an example.

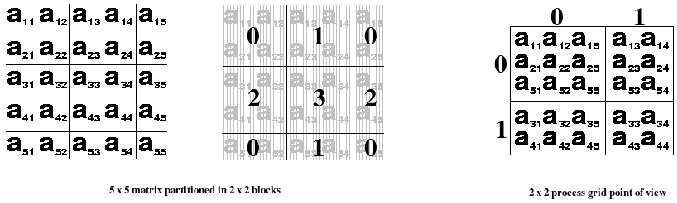


Figure B.2: This shows how a 5×5 matrix is partitioned into blocks of size 2×2 and distributed on to a 2×2 process grid. Left: The 5×5 matrix is partitioned into blocks of 2×2. Middle: The figure shows which block is assigned to which processor, here indicated by the processor number in black. Right: Shows the local matrices for each processor with the corresponding grid coordinates listed outside the frame. The processor grid coordinates and number is (0,0) = 0, (0,1) = 1, (1,0) = 2 and (1,1) = 3.

The global matrix has dimensions $M_A \times N_A = 5 \times 5$, and it is partitioned into blocs of size $MB \times NB = 2 \times 2$. These blocks are now distributed to the processors in a block-cyclic manner. How they are distributed is given in figure B.2. The processor with coordinates (0,0) in the grid is given the first block consisting of elements a_{11}, a_{12}, a_{21} and a_{22}. The next block given to this processor consists only of the two elements a_{15} and a_{25} since the matrix dimension and block size does not divide evenly.

B.4 Computational Routines

We may now call the desired routine we wish to perform. The routines of ScaLAPACK are very similar to the ones used in LAPACK. This was done to make it easy to use. ScaLAPACK features many different routines and they are classified into three categories. These are driver routines, computational routines and auxiliary routines. The driver routines simply calls a sequence of computational routines. So the computational routines can perform a wider range of tasks than the driver routines. Auxiliary routines perform a certain sub-task or common low-level computation.

To compute the inverse and determinant of the covariance matrix, I need to use the computational routines, so I will mainly look at these here. The computational routines gives many options for what type of matrix you start with. You may have a general matrix, a symmetric and Hermitian positive definitie, general band, etc. The available routines are given in the list below. The last three letters denotes the type of the computational routine.

- PxyyTRF - Factorize
- PxyyTRS - use the factorization to solve equation B.2.
- PxyyCON - Estimate the reciprocal of the condition number $\kappa(A) = |A||A|^{-1}$

Matrix Type and Storage	Operation	Single Precision		Double Precision	
		real	complex	real	complex
General	factorize	PSGETRF	PCGETRF	PDGETRF	PZGETRF
partial pivoting	solve using fact.	PSGETRS	PCGETRS	PDGETRS	PZGETRS
	est. cond. number	PSGECON	PCGECON	PDGECON	PZGECON
	error bnds. for solu.	PSGERFS	PCGERFS	PDGERFS	PZGERFS
	invert using fact.	PSGETRI	PCGETRI	PDGETRI	PZGETRI
	equilibrate	PSGEEQU	PCGEEQU	PDGEEQU	PZGEEQU
General band	factorize	PSGBTRF	PCGBTRF	PDGBTRF	PZGBTRF
partial pivoting	solve using fact.	PSGBTRS	PCGBTRS	PDGBTRS	PZGBTRS
General band	factorize	PSDBTRF	PCDBTRF	PDDBTRF	PZDBTRF
no pivoting	solve using fact.	PSDBTRS	PCDBTRS	PDDBTRS	PZDBTRS
General tridiag.	factorize	PSDTTRF	PCDTTRF	PDDTTRF	PZDTTRF
no pivoting	solve using fact.	PSDTTRS	PCDTTRS	PDDTTRS	PZDTTRS
Symmetric/	factorize	PSPOTRF	PCPOTRF	PDPOTRF	PZPOTRF
Hermitian	solve using fact.	PSPOTRS	PCPOTRS	PDPOTRS	PZPOTRS
positive definite	est. cond. number	PSPOCON	PCPOCON	PDPOCON	PZPOCON
	error bnds. for solu.	PSPORFS	PCPORFS	PDPORFS	PZPORFS
	invert using fact.	PSPOTRI	PCPOTRI	PDPOTRI	PZPOTRI
	equilibrate	PSPOEQU	PCPOEQU	PDPOEQU	PZPOEQU
Symmetric/	factorize	PSPBTRF	PCPBTRF	PDPBTRF	PZPBTRF
Hermitian	solve using fact.	PSPBTRS	PCPBTRS	PDPBTRS	PZPBTRS
positive definite band					
Symmetric/	factorize	PSPTTRF	PCPTTRF	PDPTTRF	PZPTTRF
Hermitian	solve using fact.	PSPTTRS	PCPTTRS	PDPTTRS	PZPTTRS
positive definite tridiagonal					
triangluar	solve	PSTRTRS	PCTRTRS	PDTRTRS	PZTRTRS
	est. cond. number	PSTRCON	PCTRCON	PDTRCON	PZTRCON
	error bnds. for solu.	PSTRRFS	PCTRRFS	PDTRRFS	PZTRRFS
	invert	PSTRTRI	PCTRTRI	PDTRTRI	PZTRTRI

Table B.1: Computational routines for linear equations

- PxyyRFS - Compute bounds on the error in the computed solution
- PxyyTRI - Use the factorization to compute A^{-1}
- PxyyEQU - Compute scaling factors to equilibrate A.

The letters replacing xyy gives what kind of matrix is used. The computational routines for different types of matrices are given in Table B.1.

$$AX = B \tag{B.2}$$

B.5 Releasing the Process Grid

When the routines have been executed, we must release the process grid. This is done by calling of BLACS_GRIDEXIT(ICTXT). At last we exit by CALL BLACS_EXIT(0).

Appendix C

Article featured in Astrophysical Journal

The Astrophysical Journal, 697:1–5, 2009 ??? doi:10.1088/0004-637X/697/1/1

INCREASING EVIDENCE FOR HEMISPHERICAL POWER ASYMMETRY IN THE FIVE-YEAR *WMAP* DATA

J. Hoftuft[1], H. K. Eriksen[1,2], A. J. Banday[3,4], K. M. Górski[5,6,7], F. K. Hansen[1], and P. B. Lilje[1,2]
[1] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, N-0315 Oslo, Norway; h.k.k.eriksen@astro.uio.no
[2] Centre of Mathematics for Applications, University of Oslo, P.O. Box 1053 Blindern, N-0316 Oslo, Norway
[3] Centre d'Etude Spatiale des Rayonnements, 9, av du Colonel Roche, BP 44346, 31028 Toulouse Cedex 4, France
[4] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, Postfach 1317, D-85741 Garching bei München, Germany
[5] Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[6] California Institute of Technology, Pasadena, CA 91125, USA
[7] Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland
Received 2009 March 6; accepted 2009 April 30; published 2009 ???

ABSTRACT

Motivated by the recent results of Hansen et al. concerning a noticeable hemispherical power asymmetry in the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) data on small angular scales, we revisit the dipole-modulated signal model introduced by Gordon et al.. This model assumes that the true cosmic microwave background signal consists of a Gaussian isotropic random field modulated by a dipole, and is characterized by an overall modulation amplitude, A, and a preferred direction, $\hat{p}$. Previous analyses of this model have been restricted to very low resolution (i.e., $3\overset{\circ}{.}6$ pixels, a smoothing scale of 9° FWHM, and $\ell \lesssim 40$) due to computational cost. In this paper, we double the angular resolution (i.e., $1\overset{\circ}{.}8$ pixels and $4\overset{\circ}{.}5$ FWHM smoothing scale), and compute the full corresponding posterior distribution for the five-year *WMAP* data. The results from our analysis are the following: the best-fit modulation amplitude for $\ell \leqslant 64$ and the ILC data with the *WMAP* KQ85 sky cut is $A = 0.072 \pm 0.022$, non-zero at 3.3σ, and the preferred direction points toward Galactic coordinates $(l, b) = (224^\circ, -22^\circ) \pm 24^\circ$. The corresponding results for $\ell \lesssim 40$ from earlier analyses were $A = 0.11 \pm 0.04$ and $(l, b) = (225^\circ, -27^\circ)$. The statistical significance of a non-zero amplitude thus increases from 2.8σ to 3.3σ when increasing $\ell_{\rm max}$ from 40 to 64, and all results are consistent to within 1σ. Similarly, the Bayesian log-evidence difference with respect to the isotropic model increases from $\Delta \ln E = 1.8$ to $\Delta \ln E = 2.6$, ranking as "strong evidence" on the Jeffreys' scale. The raw best-fit log-likelihood difference increases from $\Delta \ln \mathcal{L} = 6.1$ to $\Delta \ln \mathcal{L} = 7.3$. Similar, and often slightly stronger, results are found for other data combinations. Thus, we find that the evidence for a dipole power distribution in the *WMAP* data increases with ℓ in the five-year *WMAP* data set, in agreement with the reports of Hansen et al.

Key words: cosmic microwave background – cosmology: observations – methods: statistical

1. INTRODUCTION

The question of statistical isotropy in the cosmic microwave background (CMB) has received much attention within the cosmological community ever since the release of the first-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*; Bennett et al. 2003a) in 2003. The reasons for this are twofold. On the one hand, the current cosmological concordance model is based on the concept of inflation (Starobinsky 1980; Guth 1981; Linde et al. 1982; Mukhanov et al. 1981; Starobinsky 1982; Linde et al. 1983, 1994; Smoot 1992; Ruhl 2003; Rynyan 2003; Scott 2003), which predicts a statistically homogeneous and isotropic universe. Since inflation has proved highly successful in describing a host of cosmological probes, most importantly the CMB and large-scale power spectra, this undeniably imposes a strong theoretical prior toward isotropy and homogeneity.

On the other hand, many detailed studies of the *WMAP* sky maps, employing higher-order statistics, have revealed strong hints of both violation of statistical isotropy and non-Gaussianity. Some early notable examples include unexpected low-ℓ correlations (de Oliveira-Costa et al. 2004), a peculiar large cold spot in the southern Galactic hemisphere (Vielva et al. 2004), and a dipolar distribution of large-scale power (Eriksen et al. 2004b). Today, the literature on non-Gaussianity and violation of statistical isotropy in the *WMAP* data has grown very large, indeed (e.g., Bernui et al. 2006; Bielewicz et al. 2005; Copi et al. 2006; Cruz et al. 2005, 2006; Eriksen et al. 2004a, 2004c, 2005; Jaffe et al. 2005, 2006; Martínez-González et al. 2006; McEwen et al. 2008; Räth et al. 2007; Yadav & Wandelt 2008), and it would be unwise not to consider these issues very seriously.

Of particular interest to us is the question of hemispherical distribution of power in the *WMAP* data, first reported by Eriksen et al. (2004b) and later confirmed by, e.g., Hansen et al. (2004) and Eriksen et al. (2005). The most recent works on this topic include those presented by Hansen et al. (2008), who found that the power asymmetry extends to much smaller scales than previously thought, and by Eriksen et al. (2007a), who quantified the large-scale power asymmetry in the three-year *WMAP* data using an optimal Bayesian framework.

A separate, but possibly physically related, line of work was recently presented by Groeneboom & Eriksen (2009), who considered the specific model for violation of Lorenz invariance in the early universe, proposed by Ackerman et al. (2007). This model involves CMB correlations with a quadrupolar distribution on the sky, and is thus orthogonal to the current dipolar model. Surprisingly, when analyzing the five-year *WMAP* data, Groeneboom & Eriksen (2009) found supportive evidence for this model at the 3.8σ significance level, when considering angular scales up to $\ell \leqslant 400$. Thus, assuming that the *WMAP* observations are free of unknown systematics, there appears to be increasing evidence for both dipolar and quadrupolar structure in the CMB power distribution, at all angular scales.

In this paper, we repeat the Bayesian analysis of Eriksen et al. (2007a), but double the angular resolution of the data. Nevertheless, we are still limited to relatively low angular

resolutions, since the method inherently relies on brute-force evaluation of a pixel-based likelihood, and therefore scales as $\mathcal{O}(N_{\rm pix}^3)$. Yet, simply by spending more computer resources we are able to increase the pixel resolution from $N_{\rm side} = 16$ to 32 and decrease the degradation smoothing scale from 9° to 4.°5 FWHM. This provides additional support for multipoles between $\ell \approx 40$ and 80. While not sufficient to provide a full and direct comparison with the results of Hansen et al. (2008), it is a significant improvement over the results presented by Eriksen et al. (2007a).

2. OVERVIEW OF MODEL AND ALGORITHMS

The Bayesian analysis framework used in this paper is very similar to that employed by Eriksen et al. (2007a). We therefore only give a brief overview of its main features here, and refer the reader interested in the full details to the original paper and references therein.

2.1. Data Model and Likelihood

The starting point for our analysis is the phenomenological CMB signal model first proposed by Gordon et al. (2005),

$$\mathbf{d}(\hat{n}) = [1 + f(\hat{n})]\mathbf{s}(\hat{n}) + \mathbf{n}(\hat{n}). \quad (1)$$

Here $\mathbf{d}(\hat{n})$ denotes the observed data in direction $\hat{n}$, $\mathbf{s}(\hat{n})$ is an intrinsically isotropic and Gaussian random field with power spectrum C_ℓ, $f(\hat{n})$ is an auxiliary modulating field, and $\mathbf{n}(\hat{n})$ denotes instrumental noise.

Obviously, if $f = 0$, one recovers the standard isotropic model. However, we are interested in a possible hemispherical asymmetry, and we therefore parameterize the modulation field in terms of a dipole with a free amplitude A and a preferred direction $\hat{p}$,

$$f(\hat{n}) = A\,(\hat{n} \cdot \hat{p}). \quad (2)$$

The modulated signal component is thus an anisotropic, but still Gaussian, random field, with covariance matrix

$$\mathbf{S}_{\rm mod}(\hat{n}, \hat{m}) = [1 + A\,(\hat{n} \cdot \hat{p})]\mathbf{S}_{\rm iso}(\hat{n}, \hat{m})[1 + A\,(\hat{m} \cdot \hat{p})], \quad (3)$$

where

$$\mathbf{S}_{\rm iso}(\hat{n}, \hat{m}) = \frac{1}{4\pi}\sum_\ell (2\ell + 1) C_\ell P_\ell(\hat{n} \cdot \hat{m}). \quad (4)$$

We now introduce one new feature compared to the analysis of Eriksen et al. (2007a), for two reasons. First, we are interested in studying the behavior of the modulation field as a function of ℓ-range, and therefore want a mechanism to restrict the impact of the modulation parameters in harmonic space. Second, we also want to minimize the impact of the arbitrary regularization noise (see Section 3) on the modulation parameters at high ℓ's. Therefore, we split the signal covariance matrix into two parts, one modulated low-ℓ part and other isotropic high-ℓ part,

$$\mathbf{S}_{\rm total} = \mathbf{S}_{\rm mod} + \mathbf{S}_{\rm iso}, \quad (5)$$

where only multipoles between $2 \leqslant \ell < \ell_{\rm mod}$ are included in $\mathbf{S}_{\rm mod}$, and only multipoles at $\ell \geqslant \ell_{\rm mod}$ are included in $\mathbf{S}_{\rm iso}$. (Note that we are not proposing a physical mechanism for generating the modulation field in this paper, but only attempt to characterize its properties. This split may or may not be physically well motivated, but it does serve a useful purpose in the present paper as it allows us to study the scale dependence of the modulation field in a controlled manner.)

Including instrumental noise and possible foreground contamination, the full data covariance matrix reads

$$\mathbf{C} = \mathbf{S}_{\rm mod}(A, \hat{p}) + \mathbf{S}_{\rm iso} + \mathbf{N} + \mathbf{F}. \quad (6)$$

The noise and foreground covariance matrices depend on the data processing, and will be described in greater detail in Section 3.

We also have to parameterize the power spectrum for the underlying isotropic component, C_ℓ. Following Eriksen et al. (2007a), we choose a simple two-parameter model with a free amplitude q and tilt n for this purpose,

$$C_\ell = q\left(\frac{\ell}{\ell_0}\right)^n C_\ell^{\rm fid}. \quad (7)$$

Here ℓ_0 is a pivot multipole and $C_\ell^{\rm fid}$ is a fiducial model, in the following chosen to be the best-fit ΛCDM power-law spectrum of Komatsu et al. (2009).

Since both the signal and noise are assumed to be Gaussian, the log-likelihood now reads

$$-2\log\mathcal{L}(A, \hat{p}, q, n) = \mathbf{d}^T\mathbf{C}^{-1}\mathbf{d} + \log|\mathbf{C}|, \quad (8)$$

up to an irrelevant constant, with $\mathbf{C} = \mathbf{C}(A, \hat{p}, q, n)$.

2.2. The Posterior Distribution and Bayesian Evidence

The posterior distribution for our model is given by Bayes' theorem,

$$P(q, n, A, \hat{p}|\mathbf{d}, H) = \frac{\mathcal{L}(q, n, A, \hat{p})P(q, n, A, \hat{p}|H)}{P(\mathbf{d}|H)}. \quad (9)$$

Here $P(q, n, A, \hat{p}|H)$ is a prior, and $P(\mathbf{d}|\mathbf{H})$ is a normalization factor often called the "Bayesian evidence." Note that we now have included an explicit reference to the hypothesis (or model), H, in all factors, as we will in the following compare two different hypotheses, namely "H1: The universe is isotropic ($A = 0$)" versus "H2: The universe is anisotropic ($A \neq 0$)."

We adopt uniform priors for all priors in the following. Specifically, we adopt $P(q) = \text{Uniform}[0.5, 1.5]$ and $P(n) = \text{Uniform}[-0.5, 0.5]$ for the power spectrum, and a uniform prior over the sphere for the preferred axis, $\hat{p}$. The modulation amplitude prior is chosen uniformly over $[0, A_{\rm max}]$, where $A_{\rm max} = 0.15$ is sufficiently large to fully encompass the non-zero parts of the likelihood. If more liberal priors are desired, the interested reader can easily calculate the corresponding evidence by subtracting the logarithm of the volume expansion factor from the results quoted in this paper.

With these definitions and priors, the posterior distribution, $P(q, n, A, \hat{\mathbf{p}}|\mathbf{d}, H)$, is mapped out with a standard MCMC sampler. The Bayesian evidence, $E = P(\mathbf{d}|H)$, is computed with the "nested sampling" algorithm (Skilling 2004; Mukherjee et al. 2006). For further details on both procedures, we refer the interested reader to Eriksen et al. (2007a).

For easy reference, we recall Jeffreys' interpretational scale for the Bayesian evidence (Jeffreys 1961): a value of $\Delta\ln E < 1$ indicates a result "not worth more than a bare mentioning;" a value of $1 < \Delta\ln E < 2.5$ is considered as "significant" evidence; a value of $2.5 < \Delta\ln E < 5$ is considered "strong to very strong;" and $\Delta\ln E > 5$ ranks as "decisive."

Table 1
Summary Statistics for Modulated CMB Model Posteriors

Data	Mask	$\ell_{\rm mod}$	$(l_{\rm bf}, b_{\rm bf})$	$A_{\rm bf}$	Significance (σ)	$\Delta\log\mathcal{L}$	$\Delta\log E$
ILC	KQ85	64	$(224^\circ, -22^\circ) \pm 24^\circ$	0.072 ± 0.022	3.3	7.3	2.6
V band	KQ85	64	$(232^\circ, -22^\circ) \pm 23^\circ$	0.080 ± 0.021	3.8	...	...
V band	KQ85	40	$(224^\circ, -22^\circ) \pm 24^\circ$	0.119 ± 0.034	3.5	...	...
V band	KQ85	80	$(235^\circ, -17^\circ) \pm 22^\circ$	0.070 ± 0.019	3.7	...	...
W band	KQ85	64	$(232^\circ, -22^\circ) \pm 24^\circ$	0.074 ± 0.021	3.5	...	...
ILC	KQ85e	64	$(215^\circ, -19^\circ) \pm 28^\circ$	0.066 ± 0.025	2.6	...	...
Q band	KQ85e	64	$(245^\circ, -21^\circ) \pm 23^\circ$	0.088 ± 0.022	3.9	...	...
V band	KQ85e	64	$(228^\circ, -18^\circ) \pm 28^\circ$	0.067 ± 0.025	2.7	...	...
W band	KQ85e	64	$(226^\circ, -19^\circ) \pm 31^\circ$	0.061 ± 0.025	2.5	...	...
ILC[a]	Kp2	~ 40	$(225^\circ, -27^\circ)$	0.11 ± 0.04	2.8	6.1	1.8

Notes. Listed quantities are data set (first column); mask (second column); maximum multipole used for modulation covariance matrix, $\ell_{\rm mod}$ (third column); marginal best-fit dipole axis (fourth column) and amplitude (fifth column) with 68% confidence regions indicated; statistical significance of non-zero detection of A (sixth column); the change in maximum likelihood between modulated and isotropic models, $\Delta\log\mathcal{L} = \log\mathcal{L}_{\rm mod} - \log\mathcal{L}_{\rm iso}$ (seventh column); and the Bayesian evidence difference, $\Delta\log E = \log E_{\rm mod} - \log E_{\rm iso}$ (eighth column). The latter two were only computed for one data set, due to a high computational cost. However, other values can be estimated by comparing the significances indicated in the sixth column.
[a] Results computed from $N_{\rm side} = 16$ and 9° FWHM data, as presented by Eriksen et al. (2007a).

3. DATA

In this paper, we analyze several downgraded versions of the five-year *WMAP* temperature sky maps, namely the template-corrected *Q*-, *V*-, and *W*-band maps, as well as the "foreground cleaned" Internal Linear Combination (ILC) map (Gold et al. 2009). Each map is downgraded to low resolution as follows (Eriksen et al. 2007b): first, each map is downgraded to a HEALPix[8] resolution of $N_{\rm side} = 32$, by smoothing to an effective resolution of 4.°5 FWHM and properly taking into account the respective pixel windows. We then add uniform Gaussian noise of $\sigma_{\rm n} = 1\,\mu$K rms to each pixel, in order to regularize the pixel–pixel covariance matrix at small angular scales. The resulting maps have a signal-to-noise ratio of unity at $\ell = 80$, and are strongly noise dominated at $\ell_{\rm max} = 95$.

Two different sky cuts are used in the analyses, both of which are based on the *WMAP* KQ85 mask (Gold et al. 2009). In the first case, we directly downgrade the KQ85 cut to the appropriate $N_{\rm side}$, by excluding any HEALPix pixel for which more than half of the corresponding sub-pixels are missing. This mask is simply denoted by KQ85. In the second case, we smooth the mask image (consisting of 0's and 1's) with a beam of 4.°5 FWHM, and reject all pixels with a value less than 0.99. We call this expanded mask KQ85e. The two masks remove 16.3% and 26.9% of the pixels, respectively.

The instrumental signal-to-noise ratio of the *WMAP* data is very high at large angular scales, at about 150 for the *V* band at $\ell = 100$. The only important noise contribution in the downgraded sky maps is therefore the uniform regularization noise, which is not subject to the additional beam smoothing. We therefore approximate the noise covariance matrix by $N_{ij} = \sigma_n^2\delta_{ij}$. Note that this approximation was explicitly validated by Eriksen et al. (2007a) for the three-year *WMAP* data, which have higher instrumental noise than the five-year data.

We also marginalize over a fixed set of "foreground templates," $\mathbf{t}_i$, by adding an additional term to the data covariance matrix of the form $\mathbf{F}_i = \alpha_i\mathbf{t}_i\mathbf{t}_i^T$, with $\alpha_i \gtrsim 10^3$, for each template. In addition to one monopole and three dipole templates,[9] we use the *V*–ILC difference map as a template for both the *V* band and ILC maps, the *Q*–ILC difference for the *Q* band, and the *W*–ILC difference for the *W* band. However, these foreground templates do not affect the results noticeably in either case, due to the sky cuts used.

4. RESULTS

The main results from the analysis outlined above are summarized in Table 1. We consider nine different data combinations (i.e., frequency bands, masks, and multipole range), and show (1) the best-fit modulation axis and amplitude, both with 68% confidence regions; (2) the statistical significance of the corresponding amplitude (i.e., A/σ_A); and (3) the raw improvement in χ^2 and Bayesian log-evidence for the modulated model over the isotropic model. The last items are shown for the ILC with the KQ85 sky cut only. For reference, we also quote the ILC result for the Kp2 mask (Bennett et al. 2003b) reported by Eriksen et al. (2007a) when analyzing the $N_{\rm side} = 16$ and 9° FWHM data.

The reason for providing the full evidence for only one data set is solely computational. The total CPU cost for the full set of computations presented here was ~50,000 CPU hr, and the evidence calculation constitutes a significant fraction of this. On the other hand, the evidence is closely related to the significance level A/σ_A, and one can therefore estimate the evidence level for other cases in Table 1 given the two explicit evidence values and significances. We have therefore chosen to spend our available CPU time on more MCMC posterior analyses, rather than on more evidence computations.

We first consider the results for the ILC map with the KQ85 mask and $\ell_{\rm mod} = 64$. In this case, the best-fit amplitude is $A = 0.073 \pm 0.022$, non-zero at the 3.3σ confidence level. The best-fit axis points toward Galactic coordinates $(l, b) = (224^\circ, -22^\circ)$, with a 68% uncertainty of 24°. These results are consistent with the results presented by Eriksen et al. (2007a), who found an amplitude of $A = 0.11 \pm 0.04$ and a best-fit axis of $(l, b) = (225^\circ, -27^\circ)$ for $\ell \lesssim 40$.

Second, we see that these results are only weakly dependent on frequency, as both the *V* band and *W* band for the same mask and ℓ-range have amplitudes within 0.5σ of the ILC map, with $A = 0.080$ and $A = 0.074$, and non-zero at 3.8σ and 3.5σ,

[8] http://healpix.jpl.nasa.gov
[9] For an explicit demonstration of the importance of monopole and dipole marginalization on this specific problem, see Gordon (2007a).

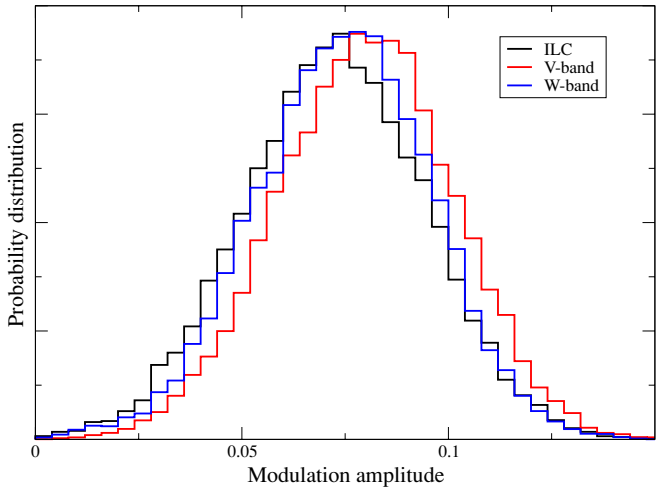


Figure 1. Posterior distributions for the dipole modulation amplitude, marginalized over direction and CMB power spectrum, computed for the KQ85 sky cut and $\ell_{\rm mod} = 64$.

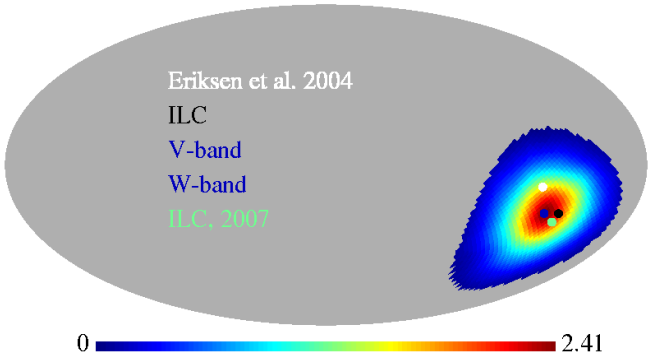


Figure 2. Posterior distribution for the dipole modulation axis, shown for the *V*-band map and KQ85 sky cut, marginalized over power spectrum and amplitude parameters. Gray sky pixels indicate pixels outside the 2σ confidence region. The dots indicate the axis (1) reported by Eriksen et al. (2004b) in white; (2) for both the ILC and *V*-band maps (these have the same best-fit axis) with the KQ85 sky cut in black; (3) for the *W* bands in blue, and the axis reported by Eriksen et al. (2007a) in green. Note that the background distribution has been smoothed for plotting purposes to reduce visual Monte Carlo noise.

respectively. (We have not included the *Q*-band analysis for the KQ85 mask, as there were clearly visible foreground residuals outside the mask for this case.) The corresponding marginal posteriors are shown in Figure 1, clearly demonstrating the consistency between data sets. Figure 2 compares the best-fit axes of the three data sets, and also indicates the axes reported by Eriksen et al. (2004b) and Eriksen et al. (2007a).

Next, we also see that the results are not strongly dependent on the choice of mask, as the amplitudes for the extended KQ85e mask are consistent with the KQ85 results, even though it removes an additional 10% of the sky. However, we do see, as expected, that the error bars increase somewhat by removing the additional part of the sky, and this reduces the absolute significances somewhat.

Finally, the best-fit modulation amplitudes for the *V*-band data and KQ85 mask are $A = 0.11$ for $\ell_{\rm mod} = 40$, $A = 0.075$ for $\ell_{\rm mod} = 64$, and $A = 0.066$ for $\ell_{\rm mod} = 80$ at 3.5σ, 3.8σ, and 3.7σ, respectively. This is an interesting observation for theoreticians who are interested in constructing a fundamental model for the effect: taken at face value, these amplitudes could indicate a non-scale invariant behavior of A, as also noted by Hansen et al. (2008). On the other hand, the statistical significance of this statement is so far quite low, as a single common value $A \sim 0.07$ is also consistent with all measurements. Better measurements at higher ℓ's are required to unambiguously settle this question.

5. CONCLUSIONS

Shortly, following the release of the first-year *WMAP* data in 2003, Eriksen et al. (2004a) presented the early evidence for a dipolar distribution of power in the CMB temperature anisotropy sky, considering only the large angular scales of the *WMAP* data. Next, Groeneboom & Eriksen (2009) presented the evidence for a quadrupolar distribution of CMB power, and found that this feature extended over all ℓ's under consideration. Finally, Hansen et al. (2008) found that the dipolar CMB power distribution is also present at high ℓ's. The evidence for violation of statistical isotropy in the CMB field is currently increasing rapidly, and the significance of these detections are approaching 4σ.

In this paper, we revisit the high-ℓ claims of Hansen et al. (2008), by applying an optimal Bayesian framework based on a parametric modulated CMB model to the *WMAP* data at higher multipoles than previously considered with this method, albeit lower than those considered by Hansen et al. (2008). In doing so, we find results very consistent with those presented by Hansen et al. (2008): the evidence for a dipolar distribution of power in the *WMAP* data increases with ℓ. For example, when considering the *V*-band data and KQ85 sky cut, the statistical significance of the modulated model increases from 3.2σ at $\ell_{\rm mod} = 40$ to 3.8σ at $\ell_{\rm mod} = 64$, and 3.7σ at $\ell_{\rm mod} = 80$.

The Bayesian evidence now also ranking within the "strong to very strong" category on Jeffreys' scale. However, it should be noted that the Bayesian evidence is by nature strongly prior dependent, and if we had chosen a prior twice as large as the one actually used, the corresponding log-evidence for the ILC map would have fallen from $\Delta \ln E = 2.6$ to 1.7, ranking only as "substantial" evidence. For this reason, it is in many respects easier to attach a firm statistical interpretation to the posterior distribution than the Bayesian evidence.

It is interesting to note that the absolute amplitude A may show hints of decreasing with ℓ. It is premature to say whether this is due simply to a statistical fluctuation, or whether it might point toward a non-scale invariant underlying physical effect, in which case the amplitude A should be replaced with a function $A(\ell)$. Either case is currently allowed by the data.

To answer this question, and further constrain the overall model, better algorithms are required. The current approach relies on brute-force inversion of an $N_{\rm pix} \times N_{\rm pix}$ covariance matrix, and therefore scales as $\mathcal{O}(N_{\rm pix}^3)$ or $\mathcal{O}(N_{\rm side}^6)$. However, already the present analysis, performed at $N_{\rm side} = 32$, required $\sim$50,000 CPU hr, and increasing $N_{\rm side}$ by an additional factor of 2 would require $\sim$3 million CPU hr. More efficient algorithms are clearly needed.

To summarize, there is currently substantial evidence for both dipolar (Hansen et al. 2008 and this work) and quadrupolar power distribution (Groeneboom & Eriksen 2009) in the *WMAP* data, and this is seen at all probed scales. The magnitude of the dipolar mode is considerably stronger than the quadrupolar mode, as a $\sim 3.5\sigma$ significance level is reached already at $\ell \sim 64$ for the dipole, while the same significance was obtained at $\ell \sim 400$ for the quadrupole.

These observations may prove useful for theorists attempting to construct alternative models for these features, either

phenomenological or fundamental. Considerable efforts have gone toward this goal already (e.g., Ackerman et al. 2007; Böhmer & Mota 2008; Carroll et al. 2008a, 2008b; Chang et al. 2008; Erickcek et al. 2008a, 2008b; Gordon et al. 2005; Emir Gümrükçüoglu et al. 2007; Himmetoglu et al. 2008a, 2008b; Kahniashvili et al. 2008; Kanno et al. 2008; Koivisto & Mota 2008a, 2008b; Pereira et al. 2007; Pitrou et al. 2008; Pullen & Kamionkowski 2007; Watanabe et al. 2009; Yokoyama & Soda 2008), but so far no fully convincing model has been established. Clearly, more work is needed on both the theoretical and observational side of this issue. Fortunately, it is now only a few years until Planck will open up a whole new window on these issues by producing high-sensitivity maps of the CMB polarization, as well as measuring the temperature fluctuations to arcminute scales. We will then be able to measure the properties of the dipole, quadrupole, and, possibly, higher-order modes of the modulation field to unprecedented accuracy.

H.K.E. acknowledges financial support from the Research Council of Norway. The computations presented in this paper were carried out on Titan, a cluster owned and maintained by the University of Oslo and NOTUR. Some of the results in this paper have been derived using the HEALPix (Górski et al. 2005) software and analysis package. We acknowledge use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science.

REFERENCES

Ackerman, L., Carroll, S. M., & Wise, M. B. 2007, Phys. Rev. D, 75, 083502
Bennett, C. L., et al. 2003a, ApJS, 148, 1
Bennett, C. L., et al. 2003b, ApJS, 148, 97
Bernui, A., Villela, T., Wuensche, C. A., Leonardi, R., & Ferreira, I. 2006, A&A, 454, 409
Bielewicz, P., Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2005, ApJ, 635, 750
Böhmer, C. G., & Mota, D. F. 2008, Phys. Lett. B, 663, 168
Bridges, M., Mc Ewen, J. D., Lasenby, A. N., & Hobson, M. P. 2006, arXiv:astro-ph/0605325
Carroll, S. M., Dulaney, T. R., Gresham, M. I., & Tam, H. 2008a, arXiv:0812.1049
Carroll, S. M., Tseng, C.-Y., & Wise, M. B. 2008b, arXiv:0811.1086
Chang, S., Kleban, M., & Levi, T. S. 2008, arXiv:0810.5128
Copi, C. J., Huterer, D., Schwarz, D. J., & Starkman, G. D. 2006, MNRAS, 367, 79
Cruz, M., Martínez-González, E., Vielva, P., & Cayón, L. 2005, MNRAS, 356, 29
Cruz, M., Tucci, M., Martínez-González, E., & Vielva, P. 2006, MNRAS, 369, 57
de Oliveira-Costa, A., Tegmark, M., Zaldarriaga, M., & Hamilton, A. 2004, Phys. Rev. D, 69, 063516
Emir Gümrükçüoglu, E. A., Contaldi, C. R., & Peloso, M. 2007, J. Cosmol. Astropart. Phys. JCAP11(2007)5
Eriksen, H. K., Banday, A. J., Górski, K. M., Hansen, F. K., & Lilje, P. B. 2007a, ApJ, 660, L81
Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004a, ApJ, 612, 633
Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2005, ApJ, 622, 58
Erickcek, A. L., Carroll, S. M., & Kamionkowski, M. 2008a, Phys. Rev. D, 78, 083012
Eriksen, H. K., Hansen, F. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004b, ApJ, 605, 14
Erickcek, A. L., Kamionkowski, M., & Carroll, S. M. 2008b, Phys. Rev. D, 78, 123520
Eriksen, H. K., Novikov, D. I., Lilje, P. B., Banday, A. J., & Górski, K. M. 2004c, ApJ, 612, 64
Eriksen, H. K., et al. 2007b, ApJ, 656, 641
Finkbeiner, D. P. 2003, ApJS, 146, 407
Finkbeiner, D. P., Davis, M., & Schlegel, D. J. 1999, ApJ, 524, 867
Gold, B., et al. 2009, ApJS, 180, 265
Gordon, C. 2007, ApJ, 656, 636
Gordon, C., Hu, W., Huterer, D., & Crawford, T. 2005, Phys. Rev. D, 72, 103002
Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelman, M. 2005, ApJ, 622, 759
Groeneboom, N. E., & Eriksen, H. K. 2009, ApJ, 690, 1807
Guth, A. H. 1981, Phys. Rev. D, 347
Hansen, F. K., Banday, A. J., & Górski, K. M. 2004, MNRAS, 354, 641
Hansen, F. K., Banday, A. J., Gorski, K. M., Eriksen, H. K., & Lilje, P. B. 2008, ApJ, submitted (arXiv:0812.3795)
Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. 1982, A&AS, 47, 1
Himmetoglu, B., Contaldi, C. R., & Peloso, M. 2008a, arXiv:0809.2779
Himmetoglu, B., Contaldi, C. R., & Peloso, M. 2008b, arXiv:0812.1231
Hinshaw, G., et al. 2006, ApJ, submitted (arXiv:astro-ph/0603451)
Jaffe, T. R., Banday, A. J., Eriksen, H. K., Górski, K. M., & Hansen, F. K. 2005, ApJ, 629, L1
Jaffe, T. R., Banday, A. J., Eriksen, H. K., Górski, K. M., & Hansen, F. K. 2006, A&A, 460, 393
Jeffreys, H. 1961, Theory of Probability (3rd ed.; Oxford: Oxford Univ. Press)
Kahniashvili, T., Lavrelashvili, G., & Ratra, B. 2008, Phys. Rev. D, 78, 063012
Kanno, S., Kimura, M., Soda, J., & Yokoyama, S. 2008, J. Cosmol. Astropart. Phys. JCAP8(2007)34
Koivisto, T., & Mota, D. F. 2008a, J. Cosmol. Astropart. Phys. JCAP8(2007)21
Koivisto, T., & Mota, D. F. 2008b, ApJ, 679, 1
Komatsu, E., et al. 2009, ApJS, 180, 330
Linde, A. D. 1982, Phys. Lett. B, 108, 389
Linde, A. D. 1983, Phys. Lett. B, 155, 295
Linde, A. D. 1994, Phys. Rev. D, 49, 748
Martínez-González, E., Cruz, M., Cayón, L., & Vielva, P. 2006, New. Astron. Rev., 50, 875
McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2008, MNRAS, 388, 659
Muhkanov, V. F., Chibishov, G. V., & Pis'mah, Zh. 1981, Eskp. Teor. Fiz., 33, 549
Mukherjee, P., Parkinson, D., & Liddle, A. R. 2006, ApJ, 638, L51
Pereira, T. S., Pitrou, C., & Uzan, J.-P. 2007, J. Cosmol. Astropart. Phys. JCAP9(2007)6
Pitrou, C., Pereira, T. S., & Uzan, J.-P. 2008, J. Cosmol. Astropart. Phys. JCAP4(2007)4
Pullen, A. R., & Kamionkowski, M. 2007, Phys. Rev. D, 76, 103529
Räth, C., Schuecker, P., & Banday, A. J. 2007, MNRAS, 380, 466
Ruhl, J. E. 2003, ApJ, 599, 786
Runyan, M. C. 2003, ApJS, 149, 265
Scott, P. F. 2003, MNRAS, 341
Skilling, J. 2004, in AIP Conf. Proc. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. von Toussaint (Melville, NY: AIP), 395
Smoot, G. F. 1992, ApJ, 396, L1
Starobinsky, A. A. 1980, Phys. Lett. B, 91, 99
Starobinsky, A. A. 1982, Phys. Lett. B, 117, 175
Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, ApJ, 609, 22
Watanabe, M.-a., Kanno, S., & Soda, J. 2009, arXiv:0902.2833
Yadav, A. P. S., & Wandelt, B. D. 2008, Phys. Rev. D, 100, 181301
Yokoyama, S., & Soda, J. 2008, J. Cosmol. Astropart. Phys. JCAP8(2007)5