Binning of the Primordial Power Spectrum and Optimal Galaxy Survey Design

Paniez Paykari

Astrophysics Group
Department of Physics
Imperial College London

Abstract

The power spectrum describes the fluctuations in the Universe and encodes much of the cosmological information. Hence measuring different types of power spectra is of great importance in cosmology today. In this thesis we try to constrain two of these power spectra, the primordial power spectrum and the galaxy power spectrum. First we analyse the binning of the primordial power spectrum. The primordial power spectrum describes the initial perturbations in the Universe which eventually grew into the large-scale structure we observe today, and thereby provides an indirect probe of inflation or other structure-formation mechanisms. We will investigate which scales the primordial power spectrum can best be probed, in accordance with the knowledge about other cosmological parameters. The aim is to find the most informative way of measuring the primordial power spectrum at different length scales, using different types of surveys, such as Planck and SDSS (Bright Red Galaxy). For this we make use of the Fisher matrix formalism, principal component analysis and Hermitian square root of the Fisher matrix. This method of binning of the primordial power spectrum is then applied to the reconstruction of this power spectrum from WMAP and simulated Planck data. Here a new method for the reconstructing, directly from observations of the Cosmic Microwave Background (CMB), is introduced. Finally, we investigate the optimal observing strategy for measuring galaxy/matter power spectrum. These power spectra are of great importance in cosmology. Measuring this spectrum will enable us to measure other cosmological parameters. If we are only interested in the large scale power spectrum then we might gain more by sparsely observing a larger patch of sky, for the same observing time, rather than observing a smaller contiguous area. We will investigate the advantages and disadvantages of this strategy using Bayesian Experimental Design.

Contents

List of Tables

List of Figures

Declaration and Copyright

This thesis is my own work, except where explicitly indicated in the text.

Paniez Paykari
October 2009

Acknowledgements

No PhD is possible without the help and support of others. Many people have contributed to the work within this thesis and helped me get where I am now.

First I would like to thank my family, for their unwavering support in whatever I have achieved so far. Without their encouragement and love none of this would have been possible. I cannot thank them enough.

My greatest thanks goes to my supervisor, Andrew Jaffe – he has been an invaluable source of information and inspiration and has given me a strong foundation as a scientist.

My thanks also go to my collaborators and friends at Imperial, who gave me guidance and many interesting scientific projects. The research presented in Chapter 3 was carried out in collaboration with Carlo Contaldi and Gavin Nicholson, who I gratefully acknowledge here for their great help.

Foreword

This thesis is based on work carried out at Imperial College London between October 2006 and September 2009. It is presented in five chapters.

Chapter 1 is a brief introduction to cosmology and a literature review of previous work that supports the thesis.

Chapter 2 is taken from Paykari & Jaffe 2009 — arXiv:0902.4399.

Chapter 3 is taken from Nicholson et. al. 2009 — arXiv:0909.5092 - **JCAP**01(2010)016.

Chapter 4 reports an analysis on the ongoing work that investigates the optimal design for galaxy surveys.

Chapter 5 concludes and summarizes the thesis.

Chapter 1

Introduction to Cosmology

1.1 A Brief Overview

About $300,000$ years after the big bang the temperature of the Universe fell to about a few thousand degrees and radiation was released which we now observe as the cosmic microwave background (CMB). This radiation was discovered by radio astronomers Arno Penzias and Robert Wilson in 1965 [1] and earned them the 1978 Nobel Prize. This discovery was in support of the Big Bang theory and ruled out the only other available theory at that time — the steady state theory. The crucial observations were made by the Far-Infrared Absolute Spectrophotometer (FIRAS) instrument on the Cosmic Background Explorer (COBE) satellite [2], — orbited in $1989 - 1996$. COBE made the most accurate measurements of the CMB frequency spectrum and confirmed it as being a black body to within experimental limits. This made the CMB spectrum the most precisely measured black body spectrum in nature. The CMB has a thermal black body spectrum at a temperature of 2.725 K: the spectrum peaks in the microwave range frequency of 160.2 GHz, corresponding to a 1.9 mm wavelength — see Figure 1.1.

The results of COBE inspired a series of ground- and balloon-based experiments, which measured CMB anisotropies on smaller scales over the next decade. These experiments were aiming to measure the angular scale of the first acoustic peak (explained in Section 1.4), for which COBE did not have sufficient resolution. These measurements ruled out cosmic strings as the theory of cosmic structure formation, and suggested cosmic inflation as the right theory. During the 1990s, the first acoustic peak was measured with increasing sensitivity and by 2000 the BOOMERanG
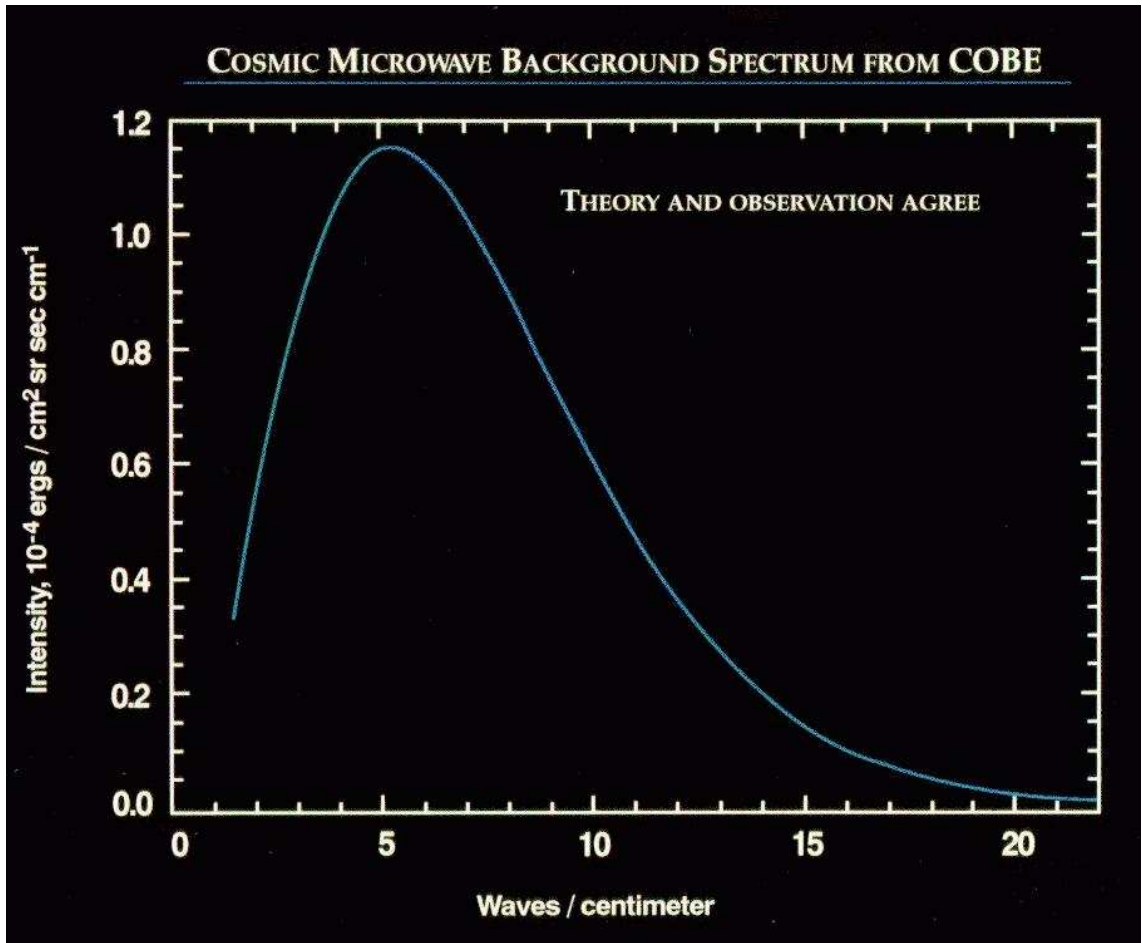
Figure 1.1: Cosmic Microwave Background (CMB) spectrum plotted in waves per centimetre vs. intensity. The solid curve shows the expected intensity from a single temperature black body spectrum, as predicted by the hot Big Bang theory [3].

experiment [4] reported that the highest power fluctuations occur at scales of about one degree. These results, together with other cosmological data, implied that the geometry of the Universe is flat. A number of ground-based interferometers provided measurements of the fluctuations with higher accuracy over the next three years, including the Very Small Array [5], Degree Angular Scale Interferometer (DASI) [6] and the Cosmic Background Imager (CBI) [7]. DASI was the first to detect the polarisation of the CMB and the CBI provided the first E-mode polarisation spectrum with compelling evidence that it is out of phase with the T-mode spectrum.

In June 2001, NASA launched another CMB mission (COBE was its first mission), Wilkinson Microwave Anisotropy Explorer (WMAP) [8], to make much more precise measurements of the CMB sky. Although WMAP provided very accurate measurements of the large angular-scale fluctuations in the CMB (0.2 degrees), it

did not have the angular resolution to cover the smaller scale fluctuations which had been observed by previous ground-based interferometers. A third space mission, the Planck Surveyor [9], was launched in May 2009. Planck will measure the CMB on smaller scales than WMAP as well as precise measurements of the polarisation of CMB. Additional ground-based instruments such as the South Pole Telescope [10], the Clover Project [11] and the QUIET telescope [12] provide additional data not available from satellite observations, possibly including the B-mode polarisation.

The CMB measurements have made the inflationary Big Bang theory the standard model of the early Universe. This theory predicts a roughly Gaussian distribution for the initial conditions of the Universe. The power spectrum of these fluctuations agrees well with the observations, although certain observables, such as the overall amplitude of the fluctuations, remain as free parameters of the cosmic inflation model. The model uses a Gaussian random field with a nearly scale invariant spectrum to represent the primeval inhomogeneities. Measuring the primordial power spectrum is one of the tasks of the modern cosmology and the main goal of this thesis.

Not only can we observe the structure of the Universe at recombination but also we can observe the current state of the Universe. Technological developments over the past two decades have accelerated the progress in observational cosmology. Observations now involve construction of huge maps of the distribution of galaxies in space, showing the remarkable large scale structure of filaments and sheets in the Universe. Early work involved making catalogs of galaxy positions on the sky, which showed some pattern of clustering. An example is the APM galaxy survey [13] which played a crucial roll in downfall of the Standard Cold Dark Matter (SCDM) model. The next generation of observations was the Two-degree-Field Galaxy Redshift Survey (2dF or 2dFGRS) [14]. The 2dF is a redshift survey which used the 3.9m Anglo-Australian Telescope and conducted between 1997 and 2002. It obtained a total of $250,000$ redshifts and covered an area of about 1500 square degrees and determined the large-scale structure in one section of the local Universe. Then came the Sloan Digital Sky Survey (SDSS) [15] which uses a dedicated telescope aiming to redshift a million galaxies, covering π steradians to a depth of several hundred mega parsecs. The survey started in 2000 and aims to obtain photometric observations on around 100 million objects and spectra for 1 million objects. The SDSS is a major multi-filter imaging and spectroscopic redshift survey using a dedicated 2.5m wide-angle optical telescope at the Apache Point Observatory in New Mexico. There are redshifts for luminous red galaxies (LRG) as far as $z = 0.4$, for quasars as far as $z = 5$

and beyond. In 2006 the survey extended to the SDSS-II by making observations to explore the structure and stellar makeup of the Milky Way, the SEGUE and the Sloan Supernova Survey, which watches after supernova Ia events to measure the distances to far objects.

The outcome of these surveys is some type of a power spectrum which is a convolution of the primordial power spectrum and (the square of) a transfer function which, in turn, depends on the cosmological parameters. This will be explained in more depth in the rest of this thesis. Observations in other areas such as weak lensing, peculiar velocity, etc. have helped a great deal in what we understand about our Universe today. The interplay between the new theoretical ideas and new observational data has taken cosmology from a purely theoretical domain into a field of rigorous experimental science and we are now in what is called the precision cosmology era.

1.2 Introduction

On the largest scales, the Universe is well described by the Friedmann equation

$$H^2 = (\frac{\dot{a}}{a})^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{K}{a^2} \tag{1.1}$$

where H = Hubble parameter, a = scale factor, $\dot{a}$ = rate of change a with respect to time, G = Gravitational Constant, ρ = total energy density of Universe, Λ = Cosmological Constant and K = Curvature constant of Universe. The Friedmann equation describes the expansion of space in homogeneous and isotropic models of the Universe within the context of general relativity. While this works perfectly well on large scales, it cannot be applied on small scales. On small scales, the Universe is neither homogeneous nor isotropic. The typical over-density on galaxy scales is about $\delta\rho/\rho = 10^5$ and on cluster scales is about 10^{2-3} and also, we know from CMB experiments that, at the time of decoupling, the fluctuations were of the order of 10^{-5}. How the homogeneous and isotropic early Universe turned into the clumpy Universe, made of galaxies and clusters, is believed to be due to *gravitational instability*. This is responsible for the structure in the Universe today.

The density contrast is usually described by

$$\delta(\boldsymbol{x}) = \frac{\delta\rho(\boldsymbol{x})}{\bar{\rho}} = \frac{\rho(\boldsymbol{x}) - \bar{\rho}}{\bar{\rho}} \ , \tag{1.2}$$

where $\bar{\rho}$ is the average density of the Universe. We will assume linearity, $\delta \lesssim 1$, otherwise modes become coupled and cannot be treated analytically; this assumption is valid while the perturbations are small. In the case of small perturbations around a smooth Universe we would benefit from Fourier transforming; different Fourier modes decouple and therefore can be treated independently. The fundamental equations governing fluid motion are

$$\text{Euler} \qquad \frac{D\mathbf{v}}{Dt} = -\frac{\nabla p}{\rho} - \nabla\Phi \ , \qquad (1.3)$$

$$\text{energy} \qquad \frac{D\rho}{Dt} = -\rho\nabla \cdot \mathbf{v} \ , \qquad (1.4)$$

$$\text{Poisson} \qquad \nabla^2\Phi = 4\pi G\rho \ , \qquad (1.5)$$

where $D/Dt = \partial/\partial t + \mathbf{v} \cdot \nabla$ is the convective derivative, p =pressure, $\mathbf{v}$ =velocity. For sufficiently small perturbation amplitude, terms containing a product of perturbations are negligible in comparison with the first order terms and hence are ignored. In the linear regime these partial differential equations turn into ordinary differential equations with $\nabla = -\imath k$; this means on Fourier transform we get $A(\boldsymbol{x}) \longrightarrow \tilde{A}(\boldsymbol{k})$, meaning every Fourier mode evolves independently.

Consider the Universe as a box of side L with periodic boundary conditions; taking L to infinity, it is possible to expand in plane waves

$$\delta(\boldsymbol{x}) = \frac{1}{(2\pi)^3}\int \delta(\boldsymbol{k})\exp(-\imath\boldsymbol{k}\cdot\boldsymbol{x})d^3k$$

$$\delta(\boldsymbol{k}) = \int \delta(\boldsymbol{x})\exp(\imath\boldsymbol{k}\cdot\boldsymbol{x})d^3x \ . \qquad (1.6)$$

This prescription is only valid in flat models in which the exponential above form an orthonormal harmonic basis with k being the eigenvalue of the expansion. Therefore, the plane waves satisfy the flat space Laplacian, $\nabla^2 Y = -k^2 Y$, where Y is the plane wave.

The perturbations can be described statistically using the correlation function ξ between two points. This is expected to depend only on the separation of two points, as isotropy has been assumed. The correlation function is defined as

$$\xi(\boldsymbol{r}) \equiv \langle\delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r})\rangle = \left\langle \frac{1}{(2\pi)^6}\int \delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')e^{-\imath(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{x}}e^{\imath\boldsymbol{k}\cdot\boldsymbol{r}}d^3k d^3k' \right\rangle , \qquad (1.7)$$

where $\delta^*(\boldsymbol{k}) = \delta(-\boldsymbol{k})$ (perturbations are real in coordinate space) and $\langle\cdots\rangle$ indicates

15

ensemble average. Having only one Universe to observe, the ergodic theorem is used, which says that on a large enough part of sky, the ensemble average is equivalent to volume average.

Now, in Fourier space, the perturbations can be described as uncorrelated, Gaussian random fields. Therefore, the second moment of $\delta(\boldsymbol{k})$ is enough to define the perturbations completely. The second moment is called the Power Spectrum of the field and is defined as

$$\langle \delta(\boldsymbol{k}) \delta^*(\boldsymbol{k}') \rangle = (2\pi)^3 \delta_D(\boldsymbol{k} - \boldsymbol{k}') P_\delta(k) \ , \tag{1.8}$$

where $k = |\boldsymbol{k}|$ (i.e., isotropy is assumed) and δ_D is the Dirac delta function. From the above equations, it can be seen that the correlation function is the Fourier transform of the power spectrum

$$\xi(\boldsymbol{r}) \equiv \langle \delta(\boldsymbol{x}) \delta(\boldsymbol{x} + \boldsymbol{r}) \rangle = \frac{1}{(2\pi)^3} \int P_\delta(\boldsymbol{k}) e^{-\imath \boldsymbol{k} \cdot \boldsymbol{r}} d^3 k \ . \tag{1.9}$$

Expanding $d^3 k = k^2 dk \sin(\theta) d\theta d\phi$ and using $e^{-\imath \boldsymbol{k} \cdot \boldsymbol{r}} = \cos(kr \cos \theta) + \imath \sin(kr \cos \theta)$, where θ is the angle between r and k, and k is aligned along the polar coordinate axis, we get

$$\xi(\boldsymbol{r}) = \frac{1}{2\pi^2} \int k^3 P_\delta(\boldsymbol{k}) \frac{\sin(kr)}{kr} d\ln k = \int \Delta_\delta^2 \frac{\sin(kr)}{kr} d\ln k \ , \tag{1.10}$$

where a dimensionless power spectrum is defined as

$$\Delta_\delta^2 = \frac{1}{2\pi^2} k^3 P_\delta(k) \ . \tag{1.11}$$

This is the variance per logarithmic k interval (per decade) of the field. $\Delta^2 \simeq 1$ is the boundary between linear and non-linear regime.

The scalar perturbations discussed above are sourced by the density fluctuations. However, many theories of the structure formation produce an additional *tensor* perturbations. These perturbations produce observable distortions in the CMB, especially on large scales. In Section 1.4.2 we will explain how these perturbations are measured.

1.3 Matter Power Spectrum

We need to find a way of relating the structure we see today to the initial pertur-
bations. If we know how perturbations were initially distributed, then it is possible
to predict what they should look like in the future if a time-dependent equation,
which describes the evolution of perturbations, is available. However, other than
the time-dependence, what mode we are considering is very important as well, as
the evolution of perturbations depends on the ratio of the scale of the perturbations
to that of the horizon. The power spectrum today needs factors accounting for

- time-dependence,

- mode-dependence,

- initial perturbation.

Therefore

$$\Phi(\boldsymbol{k}, a) \propto \Phi_0(\boldsymbol{k}) T(k) D(a) \ , \tag{1.12}$$

where

- $\Phi(\boldsymbol{k}, a)$ = the potential.

- $\Phi_0(\boldsymbol{k})$ = the primordial potential set by inflation.

- $T(\boldsymbol{k})$ = Transfer Function; it accounts for the mode horizon-crossing, and es-
  pecially takes care of the effects on the modes at the matter-radiation equality.
  It is defined as

$$T(k) \equiv \frac{\Phi(k, a_{Late})}{\Phi_{LS}(k, a_{Late})} \ , \tag{1.13}$$

  where, a_{Late} denotes an epoch well after the transfer function regime and Φ_{LS}
  is the potential on large scales (this is the primordial Φ decreased by a small
  amount — explained on the next page).

- $D(a)$ = Growth Function; it takes care of the time evolution of modes af-
  ter matter-radiation equality and it is mode-independent. In a flat, matter-
  dominated Universe the potential is constant, which means the over-density
  grows with scale factor, $\delta \propto a$.

It is worth mentioning that the separation of variables in Equation 1.12 is only
possible in the linear theory; on Fourier transforming the partial differential linear

17

equations, we get $\partial/\partial x \rightarrow k$. This produces a set of ordinary differential equations for the Fourier modes, and this set of equations is uncoupled.

[Aside; in reality we have

$$\Phi(\boldsymbol{k}, a) = \frac{9}{10}\Phi_0(\boldsymbol{k})T(k)D(a) . \tag{1.14}$$

The 9/10 factor accounts for the transition from radiation- to matter-dominated era. The transition is not instantaneous and happens over some time. This makes the modes, even the largest ones, still decay some time after the matter-radiation equality. Therefore, the final potential after the equality is a factor of 0.9 of what it was before (Remember, that the potential on the largest scales are meant to present the primordial potential, as they should not have been effected by the evolution throughout the history of the Universe, being super-horizon. But the transition through the matter-radiation equality makes the potential to drop by a factor of 10). Sometimes this factor is taken into the transfer function [16].]

Now, we can use the Poisson equation to relate the dark matter density perturbations to the gravitational potential

$$\Phi(\boldsymbol{k}, a) = \frac{4\pi G \rho_m a^2 \delta}{k^2} . \tag{1.15}$$

Using

$$\begin{aligned} \rho_m &= \rho_0 (\frac{a_0}{a})^3 \\ &= \frac{\rho_0}{\rho_{cr}}\rho_{cr}(\frac{a_0}{a})^3 \\ &= \Omega_m \rho_{cr} a^{-3} , \end{aligned} \tag{1.16}$$

having $a_0 = 1$ and the critical density definition $4\pi G \rho_{cr} = 3H_0^2/2$, for today, we obtain

$$\delta(\boldsymbol{k}, a) = \frac{3k^2}{5\Omega_m H_0^2}\Phi_0(\boldsymbol{k})T(k)D(a) . \tag{1.17}$$

The power spectrum is defined as

$$\langle\delta(\boldsymbol{k}, a)\delta^*(\boldsymbol{k}', a)\rangle = (2\pi)^3\delta_D(\boldsymbol{k} - \boldsymbol{k}')P_\delta(k, a) . \tag{1.18}$$

18

The simplest form to assume for the primordial power spectrum is a dimensionless power that is constant on all scales and is isotropic

$$\langle \Phi(\boldsymbol{k}, a)\Phi^*(\boldsymbol{k}', a) \rangle = (2\pi)^3 \delta_D(\boldsymbol{k} - \boldsymbol{k}') P_\Phi(k, a) , \qquad (1.19)$$

with $\Delta_\Phi^2 = (1/2\pi^2)k^3 P_\Phi(k) = Ak^{n_s-1}$. The $n_s = 1$ case refers to the **Harrison-Zeldovich (HZ)** scale-invariant spectrum.

[Aside; In the inflationary picture, scalar, vector and tensor perturbations of the metric are produced during inflation [17, 18]. The scalar perturbations grow by gravitational instability to structures in the Universe today. Vector perturbations decay during the fast expansion and have no importance. The tensor perturbations are due to the background of gravity waves (and they also induce some anisotropy in the CMB photons). The vacuum energy of the scalar field is the dominant energy density at the epoch of inflation. The kinetic energy is small and so the scalar field is changing very slowly. This slow change in the scalar field, while the Universe is going through rapid expansion, means the power spectrum after inflation is expected to be scale-invariant, i.e. the HZ spectrum.]

From above, a matter power spectrum can be obtained

$$P_\delta(k) \simeq Ak^{n_s} T^2(k) D^2(a) . \qquad (1.20)$$

The observable in the present-day Universe is the galaxy distribution and so a galaxy power spectrum is what we obtain from observations. This is related to the matter power spectrum via *bias* , which is defined as $P_g(k) = b^2 P_\delta(k)$. This linear relationship suggests where there is an over-density in dark matter, there is also a collection of galaxies, but with somewhat more power. This means galaxies trace the underlying dark matter distribution. This is really a simplified version, as bias can be scale-dependent or morphology-dependent. So far, some galaxy power spectra for different types of galaxies have been obtained and they all seem to have the same shape as the dark matter power spectrum, with a different amplitude; an example is [19]. The measurements of the galaxy power spectrum will be under investigation in Chapter 4, where we intend to optimise galaxy surveys for power spectrum measurements.

We can have a general idea of what the overall power spectrum should look like; the dominant component of the Universe is of great importance here. This is what makes the modes evolve differently. The driving force for the modes to grow is gravitation.

In the radiation-dominated era, the energy in radiation is so great that it provides the dominant gravitational force. The density grows in radiation and the matter distribution tends to follow the radiation distribution. However, there is a limit to how much radiation can clump, as radiation exerts pressure. The combination of the gravitation and the radiation pressure forces introduces an oscillation in the perturbations. This is exactly the photon-baryon acoustic oscillation (baryons and photons are coupled because of Thompson scattering of photons and electrons and the electrostatic attraction between the electrons and nuclei). The oscillation means that over some time the over-densities do not grow on small scales. On larger scales though, over-densities can grow as the radiation does not have time to cross the whole scale. The boundary between the two scales is defined by the Jeans length, which is when the sound crossing time (speed of radiation pressure) is equal to the gravitational in-fall time [1] . As time goes on, Jeans length increases as radiation pressure waves can travel further. This means that perturbations on larger and larger scales stop growing as time passes. This is true until the epoch of matter-radiation equality. After this, the main gravitational component is matter, which has a small pressure. This means that all scales can now grow. This induces a break in the matter power spectrum, frozen at the scale of the horizon at the epoch of matter-radiation equality.

How much growth there is on each scale is translated to the amplitude of the relevant scale in the power spectrum (through the Fourier transform); more growth means larger amplitudes. This is depicted in Figure 1.2. The first plot is the scale-invariant HZ power spectrum. The series of plots represent the matter power spectrum as the Universe grows and goes through different phases of its evolution. The peak moving across is the scale of the Jeans length at that time in the history of the Universe. The last plot is the matter power spectrum today, with the peak at the scale of matter-radiation equality. Scales smaller than the scale of equality became sub-horizon at some time in the past. The smaller the scale, the earlier it becomes sub-horizon and the less power it has in the power spectrum. Scales larger than the scale of matter-radiation equality have been super-horizon all along.

Knowing the physics, it is easy to formulate a general form for the transfer function and hence a general shape for the power spectrum. Perturbations with $kd_{eq} \ll 1$ (d_{eq} size of the Universe at the matter-radiation equality) have always been super-horizon and will always grow. Perturbations with $kd_{eq} \gg 1$ will stop growing as

---

[1] Jeans length matches with the horizon scale in the radiation-dominated era. This is due to the fact that the furthest sound can travel in this epoch is determined by the sound speed of radiation.
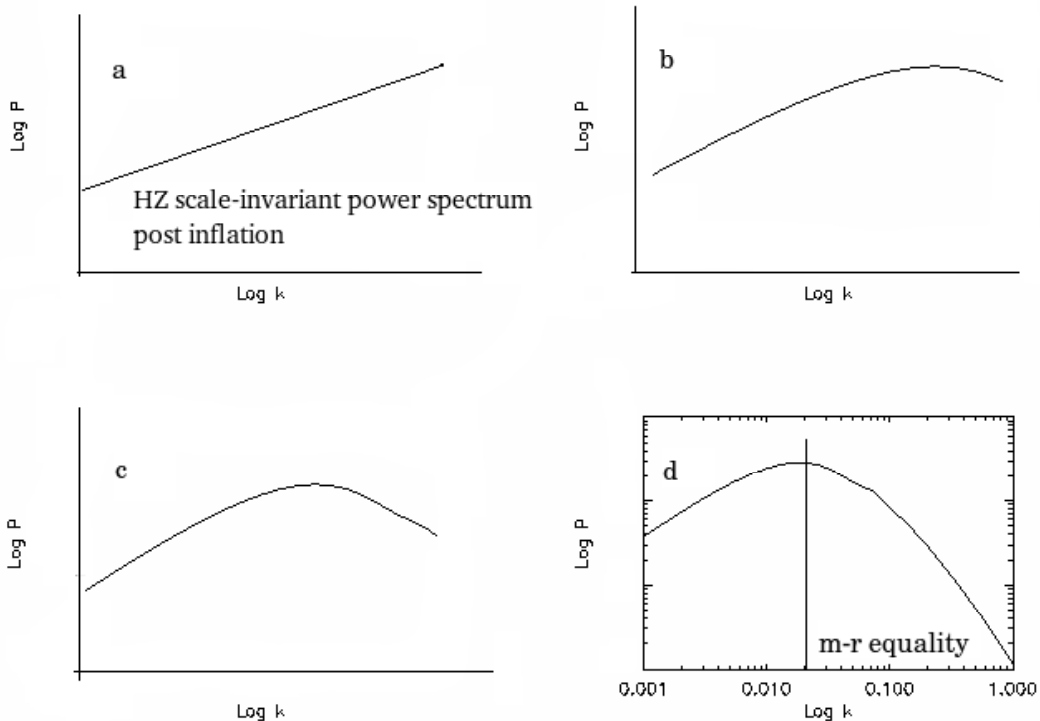
Figure 1.2: Power Spectrum from just after inflation (a), until today (d), with arbitrary normalisation.

soon as they become sub-horizon— which depends on their size k. Remembering the growth of perturbations, $\delta \propto d^2$, we get

$$
T = \begin{cases} 1 & kd_{eq} \ll 1 \\ (kd_{eq})^{-2} & kd_{eq} \gg 1 \end{cases} ,
$$
(1.21)

where $(kd_{eq})^{-2}$ is the ratio of the amplitude of the sub-horizon k-mode to its super-horizon amplitude — had it not entered the horizon. This is shown in Figure 1.3. Therefore, for $P_\delta(k) \propto T^2 k$ we have

$$
P_\delta(k) = \begin{cases} k & kd_{eq} \ll 1 \\ k^{-3} & kd_{eq} \gg 1 \end{cases} .
$$
(1.22)

This is Figure 1.2-d for a ΛCDM model.

$kd_{eq}<1$ $kd_{eq}>1$

Scale of m-r equality

0.001   0.010   0.100

Wavenumber k/hMpc^{-1}
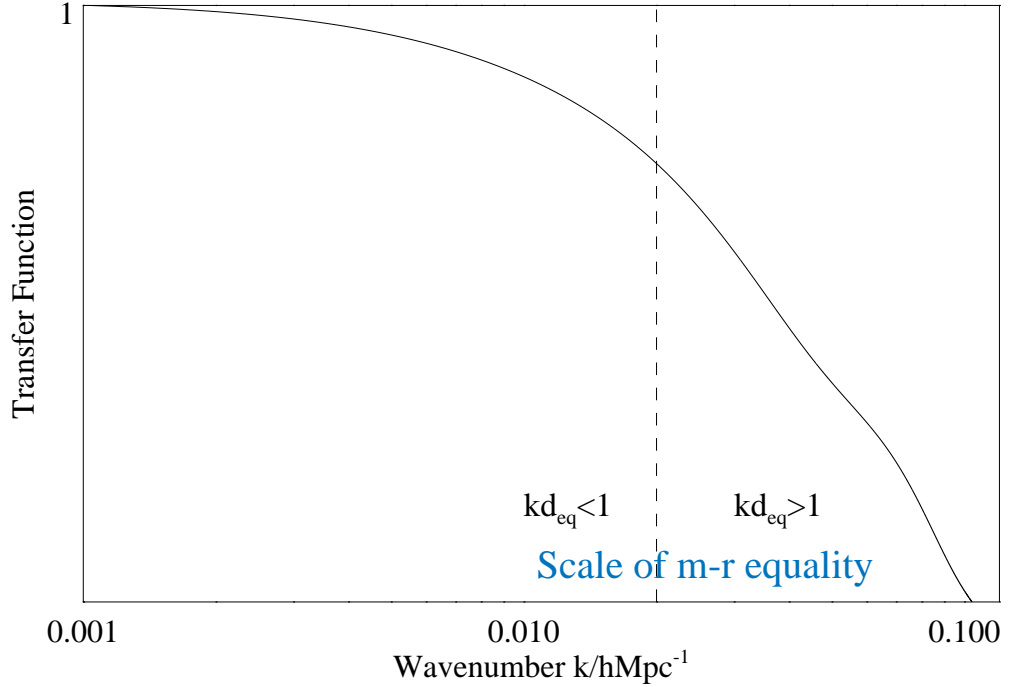
Figure 1.3: Matter Transfer Function.

1.4   CMB Power Spectrum

1.4.1   Temperature Power Spectrum

In the case of the CMB, the observable quantity is the temperature of the photons coming from different directions, at our position and time. The primordial photons can be separated out from other nuisance photons (coming, for instance, from our galaxy) by analysing the incoming flux and frequency from different directions. The primordial photons seem to have the same flux in all directions. However, when analysed in detail, fluctuations can be seen, which are only perturbed to one part in $100,000$.

As our position and time is fixed, the only variation we can measure is the direction the photon is coming from, $\hat{p}$. This temperature can be described as

$$T(\boldsymbol{x}, \hat{p}, \eta) = T_{CMB}(\eta)[1 + \Theta(\boldsymbol{x}, \hat{p}, \eta)] \,, \tag{1.23}$$

where $\Theta(\boldsymbol{x}, \hat{p}, \eta)$ is the temperature anisotropy. Just like the matter power spectrum, we need to expand the field using an orthonormal basis which, in this case, is that

of spherical harmonic functions, with coefficients $a_{\ell m}$ described like

$$
\begin{aligned}
a_{\ell m}(\boldsymbol{x}, \eta) &= \int d\Omega \Theta(\boldsymbol{x}, \hat{p}, \eta) Y_{\ell m}^*(\hat{p}) \\
&= \int \frac{d^3 k}{(2\pi)^3} e^{-\imath \boldsymbol{k} \cdot \boldsymbol{x}} \int d\Omega \Theta(\boldsymbol{k}, \hat{p}, \eta) Y_{\ell m}^*(\hat{p}) ,
\end{aligned}
\tag{1.24}
$$

where in the second line of the equation $\Theta(\boldsymbol{x})$ is replaced by its Fourier transform $\Theta(\boldsymbol{k})$. The multipole moment, ℓ, is related to the angular size on the sky like $\ell \sim 180^0/\theta$, m is the phase ranging from $-\ell$ to ℓ, $d\Omega$ refers to integration over the whole sky and $Y_{\ell m}^*(\hat{p})$ is the spherical harmonic function. Assuming a Gaussian random field again, all we need is the average and variance of all quantities. In case of $a_{\ell m}$, the average vanishes and the variance is given by

$$
\langle a_{\ell m} a_{\ell m}^* \rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell .
\tag{1.25}
$$

C_ℓ is the angular power spectrum of the temperature field, and is m-independent, as isotropy has been assumed. C_ℓ is the variance of the Gaussian distribution and for each ℓ, $a_{\ell m}$ is drawn from the same width Gaussian. The angular power spectrum is related to the angular correlation function as follows

$$
C(\theta) \equiv \langle \Theta(\hat{n}) \Theta^*(\hat{m}) \rangle = \frac{1}{4\pi} \sum_\ell (2\ell + 1) C_\ell P_\ell(\hat{n} \cdot \hat{m}) ,
\tag{1.26}
$$

where $\hat{n} \cdot \hat{m} = \cos(\theta)$.

Knowing that only the initial perturbations (initial amplitude and phase of the perturbations) are stochastic and the evolution of the initial perturbations to the anisotropy today is not a stochastic process, we can have

$$
\Theta(\boldsymbol{k}, \hat{p}, \eta) = \Phi_0(\boldsymbol{k}) \frac{\Theta(\boldsymbol{k}, \hat{p}, \eta)}{\Phi_0(\boldsymbol{k})} ,
\tag{1.27}
$$

where it has been assumed that initial condition is isotropic, so that there is no $\hat{p}$ dependence. The ratio on the right hand side of the above equation represents the evolution of the perturbations. This means that each mode has the same phase, independent of the direction it is coming from. This is known as *coherence* and this is what allows us to see a coherent pattern of acoustic oscillations in the CMB at the last scattering surface.

Taking the ensemble average of the above

$$
\begin{aligned}
\langle \Theta(\boldsymbol{k}, \hat{p}, \eta) \Theta^*(\boldsymbol{k}', \hat{p}', \eta) \rangle &= \frac{\Theta(\boldsymbol{k}, \hat{p}, \eta)}{\Phi_0(\boldsymbol{k})} \frac{\Theta^*(\boldsymbol{k}', \hat{p}', \eta)}{\Phi_0^*(\boldsymbol{k}')} \langle \Phi_0(\boldsymbol{k}) \Phi_0^*(\boldsymbol{k}') \rangle \\
&= \frac{\Theta(k, \hat{p}, \eta)}{\Phi_0(k)} \frac{\Theta^*(k', \hat{p}', \eta)}{\Phi_0^*(k')} (2\pi)^3 \delta_D(\boldsymbol{k} - \boldsymbol{k}') P_\Phi(k) \quad (1.28)
\end{aligned}
$$

The ensemble average is only taken around the initial values as this is the stochastic process, whereas, the ratio of perturbations to their initial values describes the evolution of the mode and is the same in any realisation. Therefore,

$$
\begin{aligned}
C_\ell &= \int \frac{d^3k}{(2\pi)^3} P_\Phi(k) \\
&\times \int d\Omega d\Omega' Y_{\ell m}^*(\hat{p}) Y_{\ell m}^*(\hat{p}') \frac{\Theta(k, \hat{p}, \eta)}{\Phi_0(k)} \frac{\Theta^*(k', \hat{p}', \eta)}{\Phi_0^*(k')} \\
&= \int \frac{d^3k}{(2\pi)^3} P_\Phi(k) \int d\Omega d\Omega' Y_{\ell m}^*(\hat{p}) Y_{\ell m}^*(\hat{p}') \frac{\Theta(k, \hat{p}, \eta)}{\Phi_0(k)} \frac{\Theta^*(k, \hat{p}', \eta)}{\Phi_0^*(k)} . \quad (1.29)
\end{aligned}
$$

Using $\hat{p} \to \hat{p} \cdot \hat{k} = \mu$ and $\Theta(k, \mu, \eta) = \sum_\ell (-\imath)^\ell (2\ell + 1) P_\ell(\mu) \Theta_\ell(k, \eta)$, we get

$$
\begin{aligned}
C_\ell &= \int \frac{d^3k}{(2\pi)^3} P_\Phi(k) \sum_{\ell\ell'} (-\imath)^\ell (-\imath)^{\ell'} (2\ell + 1)(2\ell' + 1) \frac{\Theta_\ell(k, \eta) \Theta_{\ell'}^*(k, \eta)}{|\Phi_0(k)|^2} \\
&\times \int d\Omega P_\ell(\mu) Y_{\ell m}^*(\hat{p}) \int d\Omega' P_{\ell'}(\mu') Y_{\ell' m}^*(\hat{p}') , \quad (1.30)
\end{aligned}
$$

where $\mu' = \hat{k} \cdot \hat{p}'$. Using $P_\ell(\hat{k} \cdot \hat{p}) = 4\pi/(2\ell+1) \sum_m Y_{\ell m}(\hat{k}) Y_{\ell' m}^*(\hat{p})$, the two integrals in the last line of above equation $\to (4\pi)^2 / ((2\ell + 1)(2\ell' + 1)) \delta_{\ell\ell'} Y_{\ell m}(\hat{k}) Y_{\ell' m}^*(\hat{k})$, making use of the orthonormality of the Legendre Polynomials. Then integrating over the angular part of the wavenumber, using $d^3k = k^2 dk \sin(\theta) d\theta d\phi$, we get

$$
C_\ell = \frac{2}{\pi} \int dk \, k^2 P_\Phi(k) \left| \frac{\Theta_\ell(k, \eta)}{\Phi(k)} \right|^2 , \quad (1.31)
$$

where the expression in the modulus is $\Delta_\ell(k, \eta)$, the angular transfer function of the radiation anisotropies. It is a two-dimensional transfer function, since it converts the real 3D perturbations into the angular perturbations on sky. Substituting the initial power spectrum of the metric introduced earlier, $k^3 P_\Phi(k) = A k^{n_s - 1}$, we have

$$
C_\ell = \frac{2A}{\pi} \int d\ln k \, k^{n_s - 1} |\Delta_\ell(k, \eta)|^2 . \quad (1.32)
$$

This tells us the relation between what we observe today and the primordial power

spectrum. The angular transfer function is calculated using the *Boltzmann* equation. The Boltzmann equations describe the statistical distribution in an out of equilibrium fluid. Therefore, it can perfectly describe the distribution and evolution of particles throughout the history of Universe. It is defined, in its most general form, as

$$\frac{df(t, x, p)}{dt} = C[f] \,, \tag{1.33}$$

where f describes the distribution of particles as a function of time, position and momentum and d/dt contains gravitational effects. $C[f]$ accounts for all the collision terms between the particles. This is zero in absence of collisions, which is the case for photons after decoupling. The collision-less case states that the number density of particles in a phase space element $(dxdp)$ is constant with time.

Writing this equation for all the components in the Universe, we get a set of equations for photons, dark matter, baryons (and mass-less neutrinos). For the case of the CMB, the Boltzmann equations for photons are of interest [20];

$$\dot{\Theta} + \imath k \mu \Theta = -\dot{\Phi} - \imath k \mu \Psi - \dot{\tau} \left[\Theta_0 - \Theta + \mu v_b - \frac{1}{2} P_2(\mu) \Pi \right] \,, \tag{1.34}$$

where Φ is the perturbations to spatial curvature, Ψ is the Newtonian potential and Π is the anisotropic shear. This equation is expanded in multipole moments using

$$\Theta(\boldsymbol{k}, \boldsymbol{n}) = \sum_{\ell} (2\ell + 1)(-\imath)^{\ell} \Theta_{\ell} P_{\ell}(\mu) \,, \tag{1.35}$$

which gives a hierarchy of coupled equations. Up to 1996, the angular transfer function was calculated by solving these coupled equations simultaneously. This meant some thousands of equations had to be solved at the same time and so it was a very time-consuming procedure. In 1996 Seljak and Zaldarriaga [21] solved this problem by separating the geometrical and dynamical contributions. This meant that the dynamical contribution, which is the *source* term expressed in terms of photons, baryons and metric perturbations, as in Equation 1.34, had to be solved for only a couple of equations at low multipole and integrated through the photon's past light cone. This improved the speed of the calculations greatly and this is exactly what the CMBFAST code is based on, [21].

The shape of the angular power spectrum (Figure 1.4) can be easily predicted by understanding the physics; in the case of the CMB, the oscillations of photons in the photon-baryon fluid are of interest. After matter-radiation equality dark matter perturbations can grow easily. However, the photon-baryon fluid is still
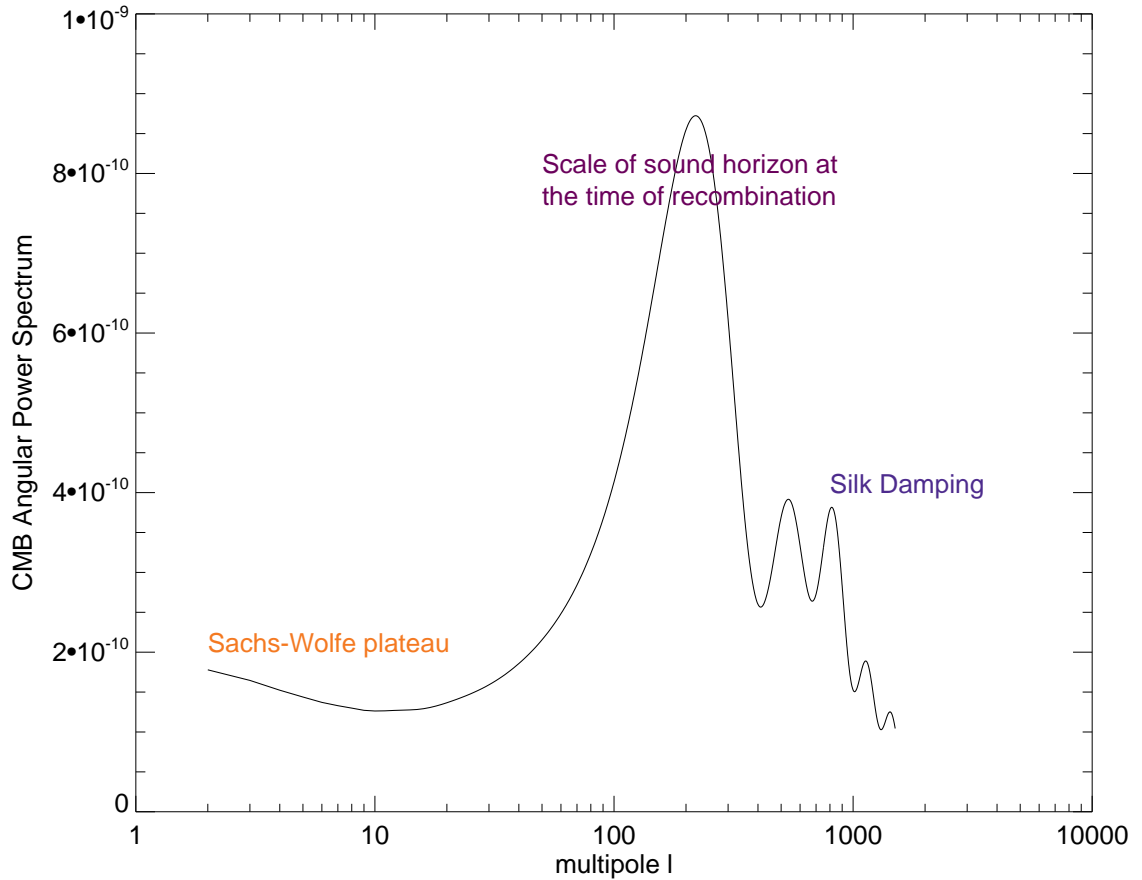
Figure 1.4: CMB Angular Power Spectrum, showing the important scales discussed in the text.

undergoing oscillations, as discussed before. They fall in the gravitational potential wells of dark matter, but get pushed apart by the radiation pressure of photons. This oscillation carries on until the time of recombination. This is when photons decouple from baryons and travel unimpeded through the Universe, reaching us today. What is frozen in the distribution of photons that we receive today is the signature of their last oscillation at the time of recombination. Large enough scales have not had time to undergo any oscillations by the time of recombination. There are some scales which have had just enough time to compress maximally under gravity. There are smaller scales which have had time for many oscillations. The scales with maximal compression represent the first peak in the power spectrum. The second peak represents the scales which have had time for a compression and are caught when maximally rarefied at the time of recombination. The third peak is again a compression. If we look at the CMB map, we expect to see lots of patterns on these scales — Figure 1.5.
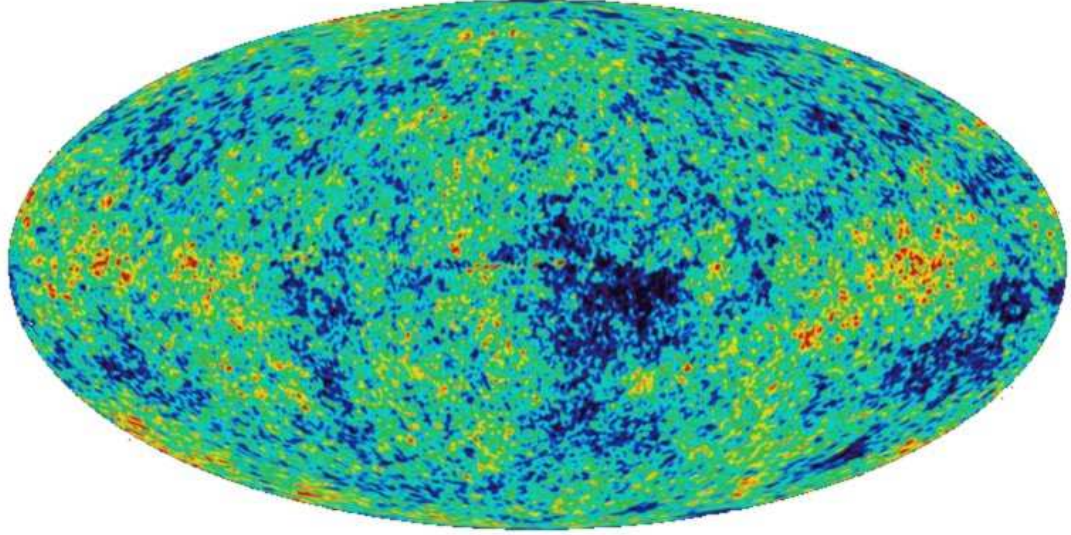
Figure 1.5: CMB map seen by WMAP — `http://map.gsfc.nasa.gov/`.

The position of the peaks are of great importance in cosmology. The first peak corresponds to the horizon scale at the time of recombination. It shows how far radiation has travelled since the Big Bang until the time of recombination. However, what we see on sky is the *apparent* angular size of the physical scale, defined by $\theta = d/D_A$, where θ is the angular size of the scale on sky, d is the physical size of the object and D_A is the angular diameter distance. Having θ from CMB observations and d from theory, it is possible to work out D_A. This can tell us about the geometry of the Universe. D_A for a flat Universe is just the physical distance, $a\chi$. However, for a closed or open Universe, it is

$$D_A = \frac{a}{H_0\sqrt{|\Omega_K|}} \begin{cases} \sinh[\sqrt{\Omega_K}H_0\chi] & \Omega_K > 0 \\ \sin[\sqrt{-\Omega_K}H_0\chi] & \Omega_K < 0 \end{cases} \tag{1.36}$$

where χ is the comoving distance. In a closed Universe, patterns in the CMB map is expected to look bigger and the opposite in an open Universe. What we see from the CMB seems to match with a flat Universe — Figure 1.6.

From roughly the third peak onwards, there is less power in the spectrum. This is due to *Silk Damping*. Silk damping occurs because the photons have a non-zero mean free path. This means they can stream out of the over-densities on small scales and make them decay. This scale turns out to approximately match the position of
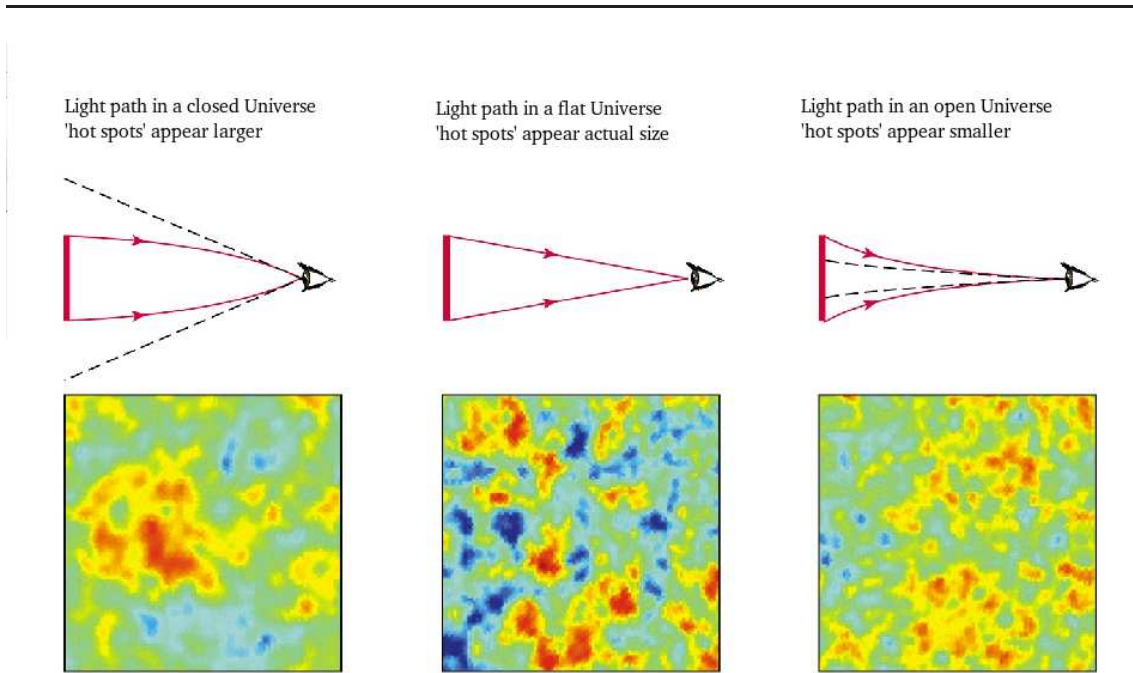
Figure 1.6: CMB map for closed/flat/open Universes — `http://find.uchicago.edu/~pryke/teaching/natsci102/spring-2005/cmbmaplab/`.

the third peak. The slope of the power spectrum here can also provide information about the neutrinos; their streaming length can put a limit on their mass [22, 23].

1.4.2 Polarisation Power Spectrum

Apart from the temperature anisotropy, the CMB radiation is polarised. This polarisation is due to the Compton scattering at the time of decoupling which thermalizes the CMB radiation. The polarisation of the CMB is not a scalar field; light travelling in the z direction has electric and magnetic fields oscillating in the $x - y$ plane (perpendicular to the propagation direction). If the intensities along the two transverse directions are not equal, then the light is linearly polarised — this is a spin-2 field. This sort of polarisation is best represented by a tensor field.

The Stokes parameters, Q and U, are used to measure the polarisation of the CMB radiation[2]. These parameters represent modes oscillating at 90° to each other; at $(0, \pi/2)$ and $(\pi/4, 3\pi/4)$ with respect to the reference direction. Q is positive if the temperature anisotropy is larger along the x axis (i.e. along 0) relative to the y axis (i.e. along $\pi/2$) and U is positive if the perturbation is large along the $\pi/4$ direction

---

[2]Compton scattering cannot produce any circular polarisation. Hence the stokes parameter $V = 0$.

relative to the $3\pi/4$ direction. Maps of the Q and U polarisation are not physical quantities (i.e. are basis-dependent) and are components of the polarisation tensor:

$$P_{ab}(\hat{n}) = \frac{1}{2} \begin{pmatrix} Q(\hat{n}) & -U(\hat{n})\sin\theta \\ -U(\hat{n})\sin\theta & -Q(\hat{n})\sin^2\theta \end{pmatrix} \, , \tag{1.37}$$

in spherical polar coordinates. We then decompose this field onto a complete set of orthonormal basis functions, using the spin-weighted spherical harmonics:

$$\frac{P_{ab}}{T_0} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{+\ell} (a^E_{\ell m} Y^E_{\ell m-ab} + a^B_{\ell m} Y^B_{\ell m-ab}) \, , \tag{1.38}$$

where T_0 is the mean CMB temperature and E and B represent the curl and divergence free modes respectively. The basis functions $Y^{E/B}$ can be given in terms covariant derivative (on a sphere) of the spherical harmonics:

$$Y^E_{ab} = \sqrt{2}(\nabla_a \nabla_b Y_{\ell m} - \frac{1}{2} g_{ab} \nabla_c \nabla^c Y_{\ell m}) \, , \tag{1.39}$$

$$Y^B_{ab} = \frac{1}{\sqrt{2}} (\epsilon^c_b \nabla_a \nabla_c Y_{\ell m} - \frac{1}{2} \epsilon^a_c \nabla_b \nabla_c Y_{\ell m}) \, , \tag{1.40}$$

where $g_{ab} = diag(1, \sin^2\theta)$ and ϵ_{ab} is a completely antisymmetric tensor. The decomposition into E and B modes are particularly helpful because scalar/density fluctuations cannot produce B modes (B modes are only produced by directed quantities such as gravitational waves or lensing). Hence a B type detection is a direct signature of the presence of a stochastic background of gravitational waves. This would provide an invaluable information about inflation. Figure 1.7 shows a polarisation pattern on sky for E and B modes.

The three types of a^E, a^B and a^T can form six types of power spectra;

$$\left\langle a^X_{\ell m} a^Y_{\ell m} \right\rangle = \delta_{\ell\ell'} \delta_{mm'} C^{XY}_\ell \, , \tag{1.41}$$

where $X, Y \in \{T, B, E\}$. The C^{TT}_ℓ power spectrum is the temperature power spectrum obtained in the previous section. We expect $C^{BE}_\ell = C^{BT}_\ell = 0$ as any correlation between B and T or E would correspond to parity violation at decoupling. Figure 1.8, 1.9, 1.10 and 1.11 show some of these power spectra from WMAP5 [24].
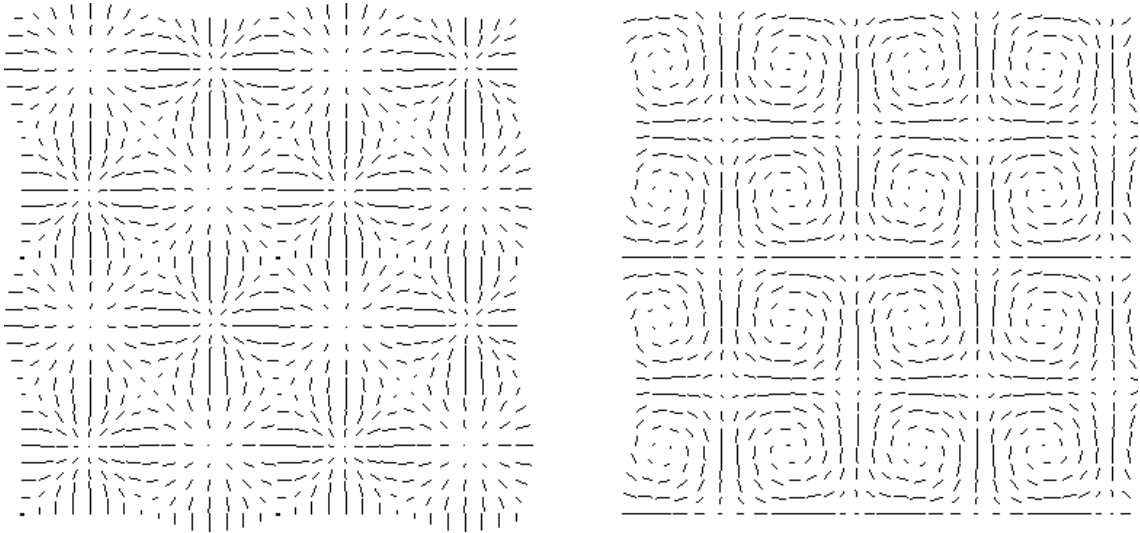
Figure 1.7: Examples of E (left) and B (right) modes of polarisation. Vectors show the direction and magnitude of the polarisation — Courtesy Joe Zuntz.

1.5 Weak Lensing Power Spectrum

A powerful probe of the *total* matter (dark and baryonic) in the Universe is gravitational lensing. Most other surveys, such as galaxy surveys, probe *galaxies* in the Universe. However, the early Universe models predict the inhomogeneities in the total matter. This deficiency in the galaxy surveys introduces an extra parameter in our models, the *bias*, which can be avoided in lensing surveys. A gravitational lens distorts light from a background galaxy by its total mass. Therefore, gravitational lensing surveys are of great importance and would be great help if included in the analysis here. The first thing to do is to obtain an observable related to the matter power spectrum. To do this we define a power spectrum of convergence, κ. The convergence is the weighted integral of mass density along the line of sight:

$$\kappa \simeq \frac{3}{2}H_0^2\Omega_m \int d\chi \, g(\chi)\frac{\delta}{a} \qquad (1.42)$$

where δ is the over-density, χ is the comoving distance and $g(\chi)$ is the lensing weight defined as

$$g(\chi) \equiv \int_\chi^{\chi_\infty} d\chi_s \, \frac{\chi(\chi_s - \chi)}{\chi_s}p(\chi_s) \,, \qquad (1.43)$$

for sources with distributions $p(\chi_s)$, normalised to $\int dp = 1$. Using Limber and Born approximations [30] we write the dimensionless convergence power spectrum
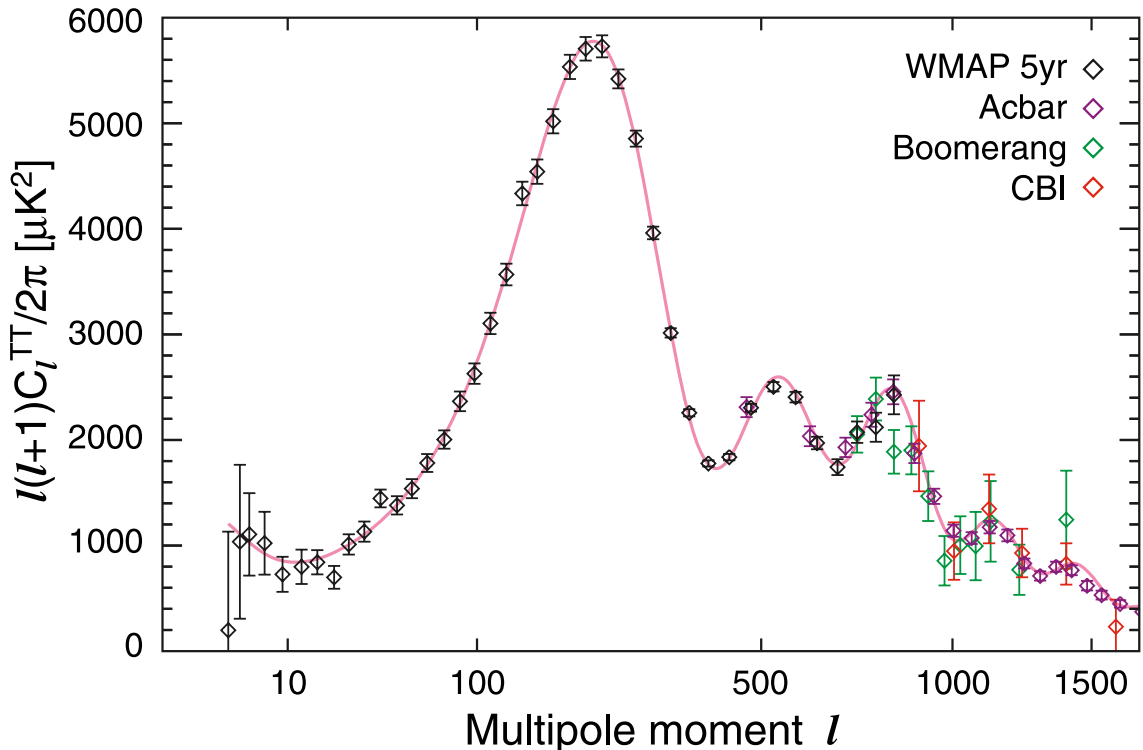
Figure 1.8: The WMAP5 TT power spectrum along with recent results from the ACBAR ([25], purple), Boomerang ([4], green) and CBI ([7], red) experiments. The red line is the best-fit ΛCDM model. Figure from `http://map.gsfc.nasa.gov/`.

as an integral over the mass power spectrum

$$\Delta_\kappa^2(\ell) = \frac{9\pi}{4\ell}\Omega_\delta^2 H_0^4 \int \chi'd\chi' \left[\frac{g(\chi')}{a(\chi')}\right]^2 \Delta_m^2(k = \frac{\ell}{\chi}, a)\,, \qquad (1.44)$$

where $\Delta_\kappa^2(\ell)$ is the contribution to κ_{rms} per logarithmic interval in angular wavenumber ℓ. Just like the other power spectra, this power spectrum depends in the primordial power spectrum through the dimensionless matter power spectrum [31].

1.6 Peculiar Velocity Field and Matter Power Spectrum

A galaxy's velocity measured away from us is composed of two effects. One is the general Hubble expansion, and the other is the *peculiar velocity* component,

31
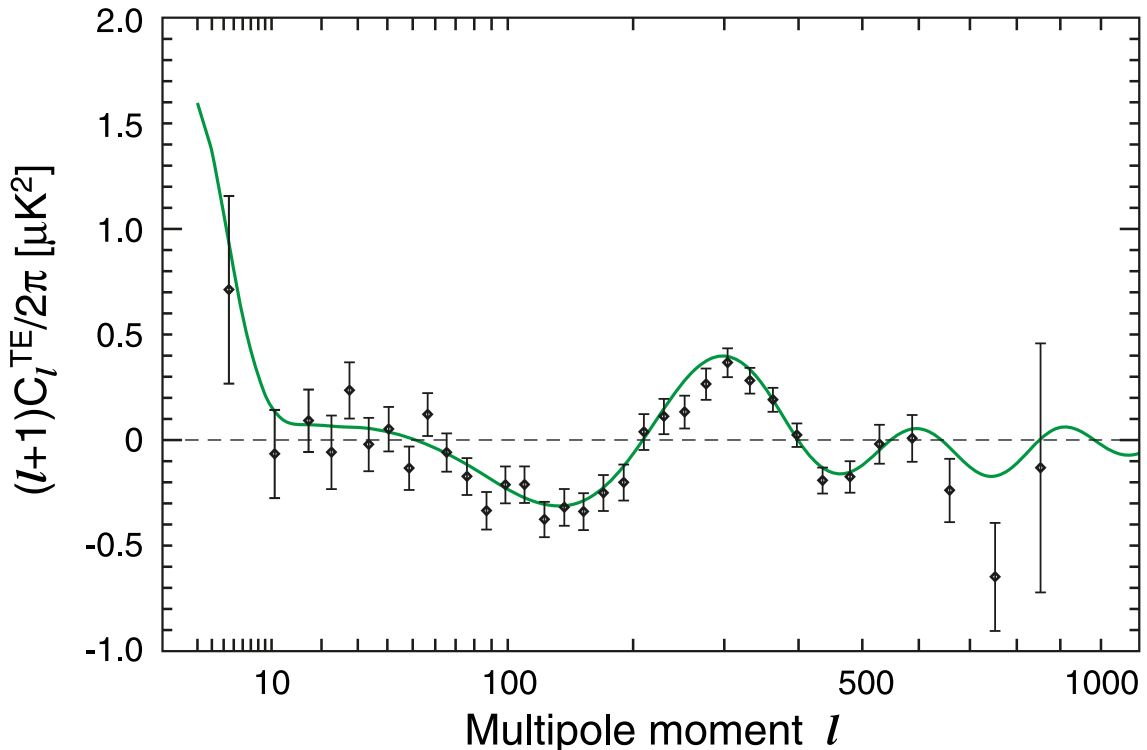
Figure 1.9: The WMAP5 TE power spectrum. The green curve is the best-fit theory spectrum from the ΛCDM/WMAP Markov chain ([24]). The clear anticorrelation between the primordial density (T) and velocity (E) in causally disconnected regions of the sky indicates that the primordial perturbations must have been on a super-horizon scale. Figure from `http://map.gsfc.nasa.gov/`.

superimposed on the Hubble expansion. Therefore the total velocity of a galaxy is

$$\boldsymbol{v} = \boldsymbol{v}_{pec} + \hat{\boldsymbol{x}} v_H \ , \tag{1.45}$$

where $\boldsymbol{v}_H = H\boldsymbol{x}$, with $H =$ Hubble Constant, and $\boldsymbol{x}$ is the radial distance of the galaxy from us. In redshift surveys, the total v is measured and this can be used to correct for the distortions induced. This distortion caused by the peculiar velocity component is of great importance for probing the total (dark and baryonic) matter distribution. This is because, in linear theory, the peculiar velocities are determined by the surrounding density field. Therefore, it is worth investigating how a matter power spectrum can be obtained from redshift surveys.

The *Continuity Equation* can be used to relate the velocity and the density fields. On scales well within the horizon, the continuity equation has the form

$$\dot{\rho} + \nabla \cdot (\rho v) = 0 \ , \tag{1.46}$$
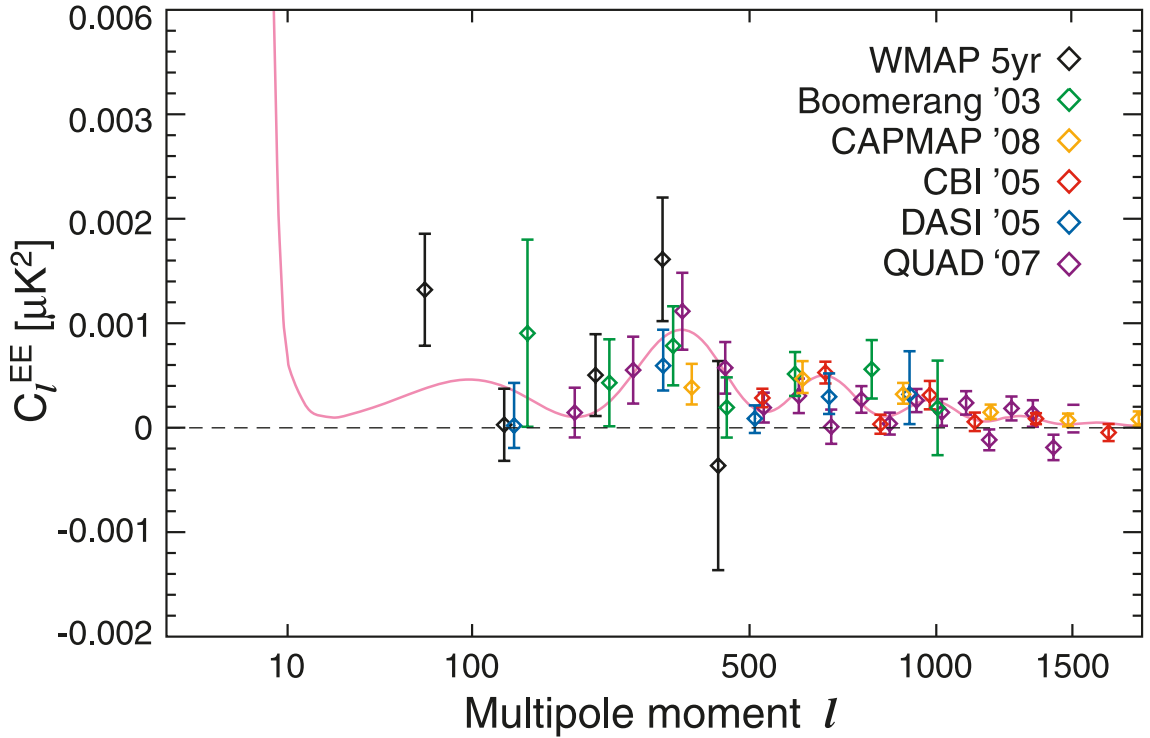
Figure 1.10: The WMAP5 EE power spectrum along with results from the Boomerang ([26], green), CBI ([27], red), CAPMAP ([28], orange), QUAD ([29], purple), and DASI ([6], blue) experiments. The pink curve is the best-fit theory spectrum from the ΛCDM/WMAP Markov chain ([24]). Figure from `http://map.gsfc.nasa.gov/`.

where $v = v_{pec}$. In Fourier space $\nabla = -\imath k$ and again defining perturbations like $\rho = \bar{\rho}(1 + \delta)$ and $v = 0 + \delta v$ (where the mean velocity perturbation, $\bar{v} = 0$) and ignoring quadratic terms, we get

$$\dot{\delta} + \imath k v = 0 \ . \tag{1.47}$$

Therefore,

$$v(k, \eta) = \frac{\imath}{k} \frac{d\delta}{d\eta} \ , \tag{1.48}$$

where η is the conformal time. To relate velocity to density, the *dimensionless linear growth rate* function is used

$$f = \frac{a}{\delta} \frac{d\delta}{da} \ , \tag{1.49}$$

and using $d/d\eta = (a^2 H)d/da$, the velocity is related to density like

$$v(k, \eta) = \frac{\imath a H f \delta(k, \eta)}{k} \ . \tag{1.50}$$

33
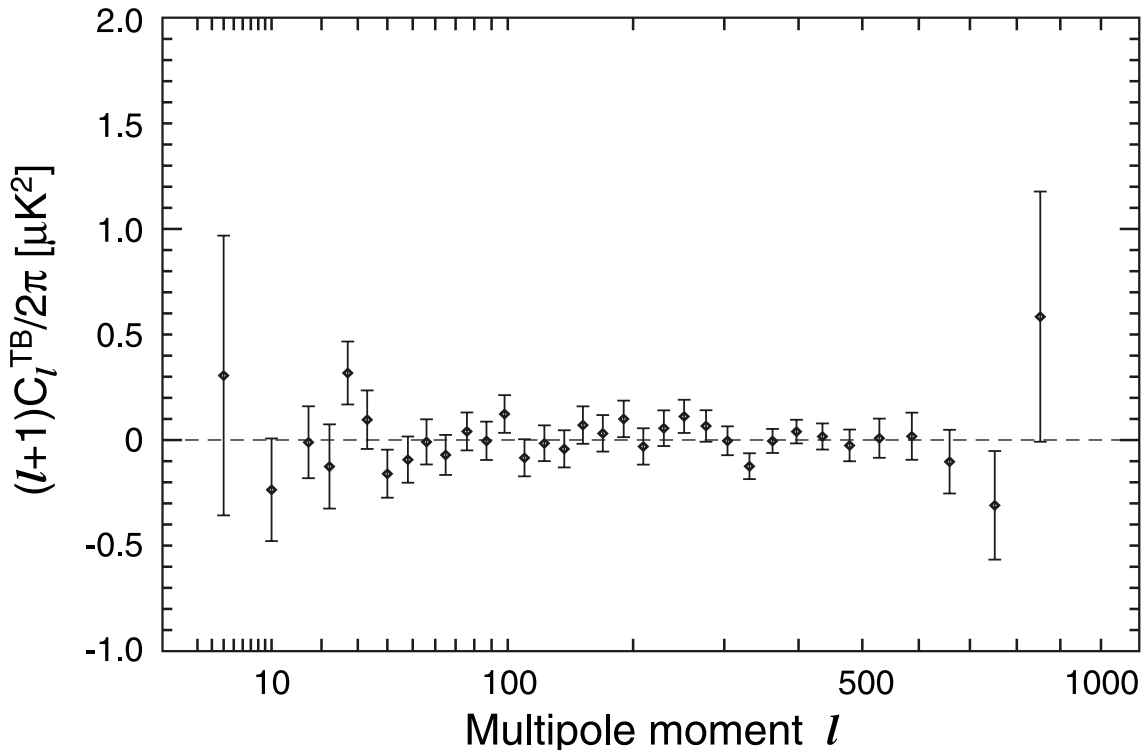
Figure 1.11: The WMAP5 TB power spectrum, showing no evidence of cosmo-logical signal. Figure from `http://map.gsfc.nasa.gov/`.

This is just a proportionality relation and defining $P_v(k) = \langle v(k)v^*(k) \rangle$, we can relate the matter power spectrum to the velocity power spectrum as

$$k^2 P_v(k) \propto P_\delta(k) \ . \tag{1.51}$$

As can be seen all the above power spectra are a convolution of the primordial power spectrum with some type of transfer function. This just shows the importance of the primordial power spectrum. In Chapters 2 and 3 the primordial power spectrum will be under investigation from two different aspects.

1.7 Data Analysis — Bayesian Statistics

To understand our Universe, we have to observe it in great detail. This has been possible only recently through different experiments such as COBE [2], WMAP [32], SDSS [33], 2dF [34], etc.. As a result, data analysis has been one of the important subjects of the recent years with huge improvements, alongside the improvement of

our experiments. The aim of data analysis is to get the most information out of data. The new surveys, such as Planck [35], SDSS, etc., are expected to produce large amounts of data. Tackling these large data sets is a big challenge in cosmology. Below, I will explain how some data analysis is done for cosmology.

The approach here is based on Bayesian Statistics, which suggests that the concept of probability can take into account any prior knowledge about the hypothesis or its parameters. The probability law $P[A \cap B] = P[A|B]P[B] = P[B|A]P[A]$, where $P[x|y]$ is the probability of x given y, can be applied here and so we obtain

$$P[H|DI] = \frac{P[D|HI]P[H|I]}{P[D|I]} \; , \tag{1.52}$$

where D = data, I = other information we have about our experiment (eg. background, etc.) and H = hypothesis. The terms in the above equation are

- $P[H|I]$ is the *prior probability*, which represents our state of knowledge about the hypothesis. This is known without having any data. We often assume a uniform prior if we do not know anything about the hypothesis, so that $P[H|I]$ = constant.

- $P[D|HI]$ is the *likelihood function*, which is defined as the probability of getting the data in an experiment given a theory.

- $P[H|DI]$ is the *posterior distribution* which, according to the Bayesian theorem, has all the information about our hypothesis and its parameters.

- $P[D|I]$ is the *evidence*. This can be thought of as the normalisation factor, realising that the posterior distribution should have a probability of 1 integrated over the parameters of the hypothesis. Therefore, it is the integration of the numerator over the parameter space. This is important as we can then ignore this term, as it shouldn't have any effect on where the likelihood function peaks in the parameter space.

This makes the equation above a proportionality relation

$$P[H|DI] \propto P[D|HI] = \mathcal{L} \; . \tag{1.53}$$

However, care needs to be taken, as ignoring the prior is not always appropriate — it can only be ignored if the data is very informative about the parameters. This means

that the likelihood function peaks at the same place as the posterior distribution in the parameter space, which indicates the best values of the parameters. Below, the likelihood function for CMB and galaxy surveys will be derived.[3]

1.7.1 CMB

In the case of the CMB, the noise is assumed Gaussian. Also, the CMB temperature at a given spot on the sky has a Gaussian distribution. This is predicted by most early Universe models like *inflation theory* [18, 17]. To get the likelihood function, we need to convolve the equations for the noise and the signal. The probability that a given spot on the sky has a temperature between s and $s + ds$ is

$$P(s)ds = \frac{1}{\sqrt{2\pi C_S}} \exp\{\frac{-s^2}{2C_S}\}ds \ , \tag{1.54}$$

where C_S is the signal variance only. The probability distribution for the noise is

$$P(\Delta) = \frac{1}{\sqrt{2\pi C_N}} \exp\{\frac{-(\Delta - s)^2}{2C_N}\} \ , \tag{1.55}$$

where C_N is the noise variance only and Δ is the estimated value for the temperature on the sky. To combine the two equations, we use $P[\Delta|C_S] = \sum_s P[\Delta|s]P[s|C_S]$ to get

$$\mathcal{L} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi C_N}} \exp\{\frac{-(\Delta - s)^2}{2C_N}\} \frac{1}{\sqrt{2\pi C_S}} \exp\{\frac{-s^2}{2C_S}\}ds \ . \tag{1.56}$$

Going through the calculation and having $C = C_N + C_S$ leads to

$$\mathcal{L} = \sqrt{\frac{1}{2\pi C}} \exp \frac{-\Delta^2}{2C} \ . \tag{1.57}$$

This is the likelihood function for 1 pixel and it looks as expected; a Gaussian in Δ, with a variance that is the sum of the two variances. We need to generalise the above equation to account for all the pixels. This leads to

$$\mathcal{L} = \frac{1}{(2\pi)^{N_p/2}|C|^{1/2}} \exp\{-\frac{\Delta^T C^{-1} \Delta}{2}\} \ , \tag{1.58}$$

where Δ is now a data *vector* which contains all the estimated values of the temperature for all the N_p pixels, C is the full covariance matrix; C_N is often diagonal

---

[3]Discussion here is from Modern Cosmology book of Scott Dodelson [16].

(white noise), but C_S is not. C_S is related to theory, C_ℓ;

$$(C_S)_{ij} = \sum_\ell \frac{2\ell + 1}{4\pi} C_\ell B_\ell P_\ell(\hat{n}_i \cdot \hat{n}_j)\,, \qquad (1.59)$$

where B_ℓ is the beam pattern and $P_\ell(\hat{n}_i \cdot \hat{n}_j)$ is the Legendre polynomial.

1.7.2 Galaxy Surveys

There are some differences between galaxy survey analysis and the CMB;

- In the galaxy surveys, we are dealing with a 3D map, while in the CMB case, we deal with a 2D projection of a 3D map.

- There are discrete objects in the galaxy surveys, while the temperature field in the CMB case is continuous.

- The CMB temperatures are assumed Gaussian, while in the galaxy survey case, we are looking at non-linearities which induce non-Gaussianity.

However, it is still possible to apply the same methods. The first step is to define a *pixel* for these surveys. From [36], a generalised data pixel i in galaxy surveys can be defined as

$$\Delta_i \equiv \int d^3x\,\psi_i(\boldsymbol{x}) \left(\frac{n(\boldsymbol{x}) - \bar{n}(\boldsymbol{x})}{\bar{n}(\boldsymbol{x})} \right)\,, \qquad (1.60)$$

where $\bar{n}(\boldsymbol{x})$ is the expected number of galaxies at position $\boldsymbol{x}$; it is the selection function of the survey and it is a constant for volume-limited surveys ($= N_{tot}/V$ — N_{tot} = total number of galaxies in the survey, V = total volume of the survey). $n(\boldsymbol{x})$ is the actual galaxy number density at that position. There are two choices for the weighting function, $\psi_i(\boldsymbol{x})$;

- CIC, *counts-in-cells*; $\psi_i^{CIC}(\boldsymbol{x}) = \begin{cases} \bar{n}(\boldsymbol{x}) & \text{if within a subvolume} \\ 0 & \text{otherwise} \end{cases}$

- *Fourier*; $\psi_i^{Fourier}(\boldsymbol{x}) = \frac{e^{i\boldsymbol{k}_i \cdot \boldsymbol{x}}}{V} \begin{cases} 1 & \text{inside survey volume} \\ 0 & \text{outside survey volume} \end{cases}$

where V is the total volume of the survey.

In both cases, Δ_i is the fractional over-density in the pixel. A pixel in this case refers to a sub-volume in the surveys that contains some number of galaxies — should have enough galaxies so that a Gaussian approach is reliable[45]. A complete expression for the likelihood function is too complicated in the case of galaxy surveys. The only progress in this case is to realise that, exactly like the CMB case, the covariance matrix is

$$\langle \Delta_i \Delta_j \rangle = (C_S)_{ij} + (C_N)_{ij} . \tag{1.61}$$

The signal covariance matrix can be found by setting the term in brackets in Equation 1.60 to $\delta(\boldsymbol{x})$. This leads to

$$(C_S)_{ij} = \langle \Delta_i \Delta_j^* \rangle \Big|_{no\,noise} = \int d^3x d^3x' \; \psi_i(\boldsymbol{x}) \psi_j(\boldsymbol{x}') \xi(\boldsymbol{x} - \boldsymbol{x}') . \tag{1.62}$$

Remembering the correlation function is just the Fourier transform of the power spectrum, it is possible to obtain

$$(C_S)_{ij} = \int \frac{d^3k}{(2\pi)^3} P_g(k) \tilde{\psi}_i(\boldsymbol{k}) \tilde{\psi}_j^*(\boldsymbol{k}) , \tag{1.63}$$

where $P_g(k)$ is the galaxy power spectrum. We will also assume that the galaxy power spectrum is *linearly* related to the underlying matter power spectrum via bias, $P_g(k) = b^2 P_\delta(k)$, where bias is scale/colour-independent. A *window function* can be defined as the angular part of the above integral

$$W_{ij}(k) \equiv \int \frac{d\Omega_k}{4\pi} \tilde{\psi}_i(\boldsymbol{k}) \tilde{\psi}_j^*(\boldsymbol{k}) , \tag{1.64}$$

so that

$$(C_S)_{ij} = \int_0^\infty \frac{dk}{k} \left[\frac{k^3 P_g(k)}{2\pi^2} \right] W_{ij} , \tag{1.65}$$

where the term in the brackets is the dimensionless power spectrum, $\Delta_g^2(k)$. For a volume-limited survey, which observes galaxies out to the radius of R, the Fourier transform of the weighting function is defined as

$$\tilde{\psi}_i(\boldsymbol{k}) = \int_{|\boldsymbol{x}|<R} \frac{d^3x}{V} e^{-\imath \boldsymbol{k}.\boldsymbol{x}} e^{\imath \boldsymbol{k}_i.\boldsymbol{x}} . \tag{1.66}$$

---

[4]By the central limit theorem, any variable that is the sum of a large number of independent factors is likely to obey the Gaussian distribution.

[5]A Poisson distribution (which is the case here as we have discrete galaxies) with parameters n approaches a Gaussian distribution for large n.

In the case where the wavelengths are much smaller than the survey volume, it is possible to obtain

$$(C_S)_{ii} \simeq \frac{P_g(k_i)}{V} \ , \tag{1.67}$$

for the signal covariance matrix and for the noise covariance matrix we obtain

$$(C_N)_{ii} = \frac{1}{\bar{n}V} \ , \tag{1.68}$$

with $1/\bar{n}$ being the Poissonian shot noise. Note that in the above calculation, any non-linear clustering and redshift distortion has been ignored. It is assumed that all the information about the parameters is given by the galaxy power spectrum. In Chapter 4 we will present an extensive data analysis for galaxy surveys and how a Fisher matrix is obtained for them. At this point, this simplified version suffices.

1.7.3 Optimal Quadratic Estimator

This is a root finding algorithm and is a fast way of performing likelihood calculations. Approximating the likelihood function as a multivariate Gaussian leads to the quadratic estimator; if we think of the likelihood function as being continuous, having a peak at the maximum and also make sure that it is structure-less (to avoid getting trapped in a local extremum), then we can think of it as a Gaussian near the peak. This is a good approximation only for a well-constrained parameter. The likelihood function is a maximum in the parameter space where the parameters have the best estimated values. Therefore,

$$\frac{\partial \ln \mathcal{L}}{\partial \lambda}\big|_{\lambda=\bar{\lambda}} = 0 \ , \tag{1.69}$$

where $\ln \mathcal{L}$ is maximised at the same position in the parameter space as $\mathcal{L}$. We can use a Taylor expansion

$$(\ln \mathcal{L})_{,\lambda}(\bar{\lambda}) = (\ln \mathcal{L})_{,\lambda}(\lambda^{(0)}) + (\ln \mathcal{L})_{,\lambda\lambda}(\lambda^{(0)})(\bar{\lambda} - \lambda^{(0)}) + \cdots \ , \tag{1.70}$$

where $\lambda^{(0)}$ is some trial point, $\bar{\lambda}$ is the value of the parameter at the maximum of the likelihood function and $_{,\lambda}$ means derivative with respect to λ. The estimated value is $\hat{\lambda} = \bar{\lambda}$. Truncating the expansion above is justified if the error bars are small; $\ln \mathcal{L}$'s drop is quite fast and therefore the third-order terms can be ignored. This also says that Gaussianity is a good approximation for $\mathcal{L}$. Calculating the derivatives

above and going through the calculation, we get

$$\hat{\lambda} = \lambda^{(0)} + F_{\lambda\lambda}^{-1} \frac{\Delta C^{-1} C_{,\lambda} C^{-1} \Delta - Tr[C^{-1} C_{,\lambda}]}{2} \ , \tag{1.71}$$

where $\Delta\Delta \rightarrow \langle \Delta\Delta \rangle = C$. F is called the *Fisher Matrix* and is the ensemble average of the *curvature* of a function (i.e. it is the average of the curvature over many realisations of signal and noise)

$$F_{\lambda\lambda} = \langle \mathcal{F} \rangle = \frac{1}{2} Tr[C_{,\lambda} C^{-1} C_{,\lambda} C^{-1}] \ , \tag{1.72}$$

where $\mathcal{F}$ is the curvature of the likelihood function defined as

$$\mathcal{F} = -\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda^2} \ . \tag{1.73}$$

It is worth emphasising that this is only possible if the likelihood function approximates to a Gaussian near the peak. $\mathcal{F}$ is inversely proportional to the *radius of curvature*. The curvature of the likelihood function is very important; it tells us how rapidly the function is moving away from the maximum point (since the first derivative vanishes, as it measures the change of likelihood function at the peak). For instance, if the curvature is small it means the function is changing slowly at the peak which, means data is not very constraining (this *is* where the prior information is important and cannot be ignored), and vice versa. Generalising the Fisher matrix to a many parameter case we have

$$\begin{aligned} F_{\alpha\beta} &= \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial \lambda_\alpha \partial \lambda_\beta} \right\rangle \\ &= \frac{1}{2} Tr[C_{,\alpha} C^{-1} C_{,\beta} C^{-1}] \ . \end{aligned} \tag{1.74}$$

The Fisher matrix allows us to estimate the errors on parameters without having to cover the whole parameter space. Authors of [37] have compared the Fisher matrix analysis with the full likelihood function analysis and found there was perfect agreement between the two methods if the likelihood function is Gaussian.

1.7.4 Fisher Matrix/Covariance matrix

The aim here is to find out how well the future experiments can measure the cosmological parameters, especially the primordial power spectrum. This can be achieved by a Fisher matrix analysis. Fisher matrices can be computed analytically and

they are suitable for *forecasting* how well the future experiments can constrain the cosmological parameters as they do not depend on data. All we need to feed into a Fisher matrix is the theory and the experimental specifications, such as beam, survey volume/sky coverage etc..

There is a large motivation to include galaxy surveys in estimating errors on cosmological parameters. There are huge degeneracies between parameters from the CMB alone and it is hoped to break this degeneracy by including galaxy surveys — the degeneracies between parameters will be more discussed in Section 1.8 and also in Chapter 2. Galaxy surveys are as important as CMB experiments and a way of analysing these large amounts of data sets that they produce is certainly needed. Also, more accuracy can be obtained if other surveys such as weak lensing, peculiar velocity, Lyman α forest surveys are included too (although, Lyman α forest may be probing too small a scale to be suitable for Fisher matrix analysis). They each probe different scales and all together can give great accuracies on parameters.

The Gaussianity assumption of the likelihood function made above is still a very good approximation for galaxy surveys as well. On large scales they are Gaussian, but even on small scales where things have gone non-linear (and therefore non-Gaussian), this method is quite competitive to other traditional estimators of error calculations.

The inverse of the Fisher matrix is an approximation of the covariance matrix of the parameters, by analogy with a Gaussian distribution in the θ_α, for which this would be exact. Some powerful theorems have been proved[6];

- The maximum likelihood estimator is asymptotically the best unbiased estimator ($\langle \Theta \rangle = \Theta^0$). This gives the errors on the parameters which are equal to $1/\sqrt{F}$.

- The Cramer-Rao inequality[7] states that the smallest error measured, for θ_α, by any unbiased estimator (such as the maximum likelihood) is $1/\sqrt{F_{\alpha\alpha}}$ and $\sqrt{(F^{-1})_{\alpha\alpha}}$, for non-marginalised and marginalised[8] one-sigma errors respectively. This does make sense as the errors obtained from a full likelihood analysis could not be smaller than the width of the likelihood function.

---

[6]For instance, Kenney, J. F. & Keeping, E. S. 1951 Mathematics of Statistics, Part II.

[7]It should be noted that the Cramer-Rao inequality is a statement about the so-called "Frequentist" confidence intervals and is not strictly applicable to "Bayesian" errors.

[8]Integration of the joint probability over other parameters.

All the Fisher matrices from different surveys can be added to account for all the information

$$F = F_{galaxy} + F_{CMB} + F_{lensing} + \dots \ . \tag{1.75}$$

This can be then inverted to give the total covariance matrix, where data from different experiments are included. Below, I will explain how the general formalism of the Fisher matrix above can be turned into the useful form for the CMB and Galaxy surveys.

1.7.4.1   CMB

To calculate the Fisher matrix explicitly for the CMB we need to know the total covariance matrix $C = C_N + C_S$ and what parameters we want the Fisher matrix for. The easiest choice for the parameters would be the C_ℓs themselves, as they have dependence on other parameters of the theory and therefore it is easy to get the Fisher matrix for other parameters using the Jacobian.

C_S is the signal covariance matrix and is related to the C_ℓs, as explained above. Therefore, we need the coefficients of the spherical harmonic functions defined as $a_{\ell m} = \int d\Omega Y_{\ell m}^*(\hat{n})\Theta(\hat{n})$, remembering the definition of C_ℓs from Equation 1.25. To take the effect of window function in, we need to multiply these C_ℓs by $e^{-\ell^2\sigma^2}$, where σ is the beam size in the experiment. This is the signal covariance matrix.

Next, we need the noise covariance matrix. One way to obtain this is to specify an error for the pixels in the map, σ_{pix}. Assuming the errors are uncorrelated and have uniform variance, $\sigma_{pix}^2\delta_{ij}$ is the variance in the noise with i and j referring to the pixels. The observing time per pixel is inversely proportional to the pixel size. The pixel solid angle can be defined as $\Omega_{pix} = \theta_{fwhm}^2$, where θ_{fwhm} is the full-width at half-maximum of the Gaussian beam in the experiment. The best way to express noise is in *weight* per solid angle [38], which gives

$$w = (\sigma_{pix}^2\Omega_{pix})^{-1} \ . \tag{1.76}$$

So, now we have as the total covariance matrix

$$C_{\ell m \ell' m'} = \delta_{\ell\ell'}\delta_{mm'}[C_\ell e^{-\ell^2\sigma^2} + w^{-1}] \ , \tag{1.77}$$

where C on the left hand side is the total covariance matrix and the C on the right hand side indicates the signal covariance matrix ($= C_\ell$). Insert this in Equation 1.74

and what we get is

$$F_{\ell\ell'} = \frac{2\ell+1}{2}\delta_{\ell\ell'}e^{-2\ell^2\sigma^2}[C_\ell e^{-\ell^2\sigma^2} + w^{-1}]^{-2} \ . \tag{1.78}$$

In an all sky survey, F above is diagonal and there are no correlations between C_ℓ. The error on C_ℓ is then

$$\delta C_\ell = \sqrt{\frac{2}{2\ell+1}}[C_\ell + w^{-1}e^{-\ell^2\sigma^2}] \ . \tag{1.79}$$

This equation looks as expected:

- $2\ell+1$ in the number independent measurement for a specific scale ℓ. This is related to the *sampling variance*, which means no matter how precisely each realisation is measured, i.e. $w = \infty$, δC_ℓ never goes to zero due to the finite number of samples.

- 2 in the numerator of the square root comes due to taking difference of the variances. The sampling variance for a Gaussian distribution is twice the square of the variance divided by the number of independent measurements.

- In general, there are two sources of error; one is from the sampling variance, $C_\ell/2\ell+1$, and the other is the noise, rising from atmospheric/instrumental limitations.

Figure 1.12 shows an angular power spectrum, its Fisher matrix and errors for Planck type survey. The large error on large scales is due to cosmic variance and on small scales is due to the resolution limit of the experiment. Planck is almost cosmic-variance limited up to a very large ℓ.

It is important to note that in reality, an all-sky survey is not really possible, at least in the near future. This induces a factor of f_{sky} in the equation above, multiplying the sampling variance term. Another effect is that the different phases, ms (on a particular scale ℓ), will not be completely independent, and there is a leakage of power from one mode to another. It also causes a lack of power as the whole of the sky is not taken into account. How much these effect the equation above need to be investigated and corrected for any specific experiment. However, they will be ignored in this report as they are not expected to have large effects.

As can be seen, experiments themselves introduce three characteristic effects into the errors; the sky coverage, the weight and beam width. These are all shortfalls

of our experiments, which affect all the scales. We are also limited by the cosmic variance on large scales. Therefore, there is only a finite range where cosmology can be probed at the moment, and we need to make the most of this.

The Fisher matrix for other parameters, λ_α, is obtained by using the Jacobian as follows

$$F_{\alpha\beta} = \sum_\ell \frac{1}{(\delta C_\ell)^2} \frac{\partial C_\ell}{\partial \lambda_\alpha} \frac{\partial C_\ell}{\partial \lambda_\beta} \; . \tag{1.80}$$

Note that the above Fisher matrix formula does not account for the systematic errors arising from Galactic/extra-Galactic foregrounds, scanning strategy etc. (although some work has been done to account for these, for example in [39, 40, 41]).

This analytical Fisher matrix will be used in Chapters 2 and 3 for our analysis.

1.7.4.2   Galaxy Surveys

A similar approach is taken for galaxy surveys. A volume-limited galaxy survey is an analogue of the all-sky CMB survey. The signal and noise covariance matrices are given in Equations 1.67 and 1.68. The inverses and derivatives are obtained and inserted into Equation 1.74. This leads to

$$F_{nn'} = \delta_{nn'} \frac{k_n^2 \Delta k V}{(2\pi)^2 (P_n + \frac{1}{\bar{n}})^2} \; , \tag{1.81}$$

where n and n' refer to the bins and P_n is the average power in the nth bin. This has a similar form to Equation 1.78, where the numerator is the number of independent modes measured in the survey. To get the Fisher matrix for other parameters, the Jacobian is used again, as

$$F_{\alpha\beta} = \sum_n \frac{k_n^2 \Delta k V}{(2\pi)^2 (P_n + \frac{1}{\bar{n}})^2} \frac{\partial P_n}{\partial \lambda_\alpha} \frac{\partial P_n}{\partial \lambda_\beta} \; . \tag{1.82}$$

Note that the sum is over n, which is associated with the bins. Using $d \ln P = dP/P$, and turning sum to an integral it is possible to obtain

$$F_{\alpha\beta} = 2\pi \int_{k_{min}}^{k_{max}} d \ln k \, \frac{V}{\lambda^3} \left(\frac{P_n}{P_n + \frac{1}{\bar{n}}} \right)^2 \frac{\partial \ln P_n}{\partial \lambda_\alpha} \frac{\partial \ln P_n}{\partial \lambda_\beta} \; , \tag{1.83}$$

where $V_{eff} = V \left(P_n/(P_n + \frac{1}{\bar{n}}) \right)^2$ is the *effective* volume for a volume-limited survey, and $w(k) = V_{eff}/\lambda^3$ is the weight function. This is shown in Figure 1.13, taken

from [36]. This shows the weight function for the main northern part and BRG sample of SDSS. Note how small wavenumbers are not weighted much. The approach above has assumed a volume-limited survey, for which $\bar{n}$ is constant. This means the volume of the survey is the only free parameter. If the survey is flux limited, the selection function is much more complicated. Also, it has been assumed that $P(k) \gtrsim 1/\bar{n}$ (the Poissonian shot noise) in the region of the survey so that the *effective* volume available for the survey is the survey volume itself. At large scales this assumption breaks down, but these scales get the least weighting anyways — Figure 1.13. The weight function, $w(k) = V_{eff}/\lambda^3$, gives the number of independent Fourier modes available. This is just like the $f_{\mathrm{sky}}(2\ell+1)$ factor in the CMB case.

The above formula assumes that there are no features in the power spectrum that are smaller than k_{min} (inverse scale of survey). Also, the choice of k_{max} is of importance as the above formula assumes Gaussianity. This will be explained in detail later.

Note also that edge effects have been ignored here. The author of [36] has tested the edge effects in the equation and have concluded that it is accurate to within a factor of two, if the survey size, $\lambda \gg 200h^{-1}\mathrm{Mpc}$. In case of SDSS, as it is deep and wide, the features in the power spectrum (such as baryonic oscillations) should be accessible and also, the edge effects can be resolved [42, 43].

The above calculations are taken from *Modern Cosmology* of Scott Dodelson [16]. However [36] is the first author to attempt to get the equation above.

This analytical version of the galaxy Fisher matrix will be used for the analysis in Chapter 2. In Chapter 4 we will present an extension of the Bayesian statistics; Bayesian Experimental Design, and also a full numerical version of the galaxy Fisher matrix, used to optimise galaxy surveys.

1.8   Measuring Cosmological Parameters

Different sets of experiments such as the CMB (Planck), galaxy (SDSS) and gravitational lensing (DUNE) surveys are underway to probe the Universe. These surveys probe different scales and constrain different parameters and all together they would provide a great handle on cosmology. On their own they may not be able to constrain the parameters due to the degeneracies amongst them. However, as different surveys constrain parameters in different directions, they can constrain more parameters when put together. For example, it has been pointed out by authors, such

as [44, 45], that CMB power spectra on their own cannot provide enough constrains on parameters due to degeneracies. This means altering the model parameters in certain combinations produces power spectra that are indistinguishable. One consequence of these degeneracies is that they break the Gaussianity assumption about the likelihood function. This means, for example, that the Fisher matrix analysis is not reliable in these cases and we might get biased results using this technique blindly! One way of breaking the degeneracies is to include other experiments as explained above.

Efstathiou and Bond in [46] refer to the degeneracy between the parameters in the CMB anisotropies as "cosmic confusion". They analyse the degeneracies between parameters for WMAP and Planck. They perform an exact likelihood analysis for a ΛCDM Universe with an adiabatic power spectrum and show that the usual Fisher matrix analysis overestimates the errors on the parameters. Their work shows that 'significant biases' are induced on the parameters due to correlated errors that exist in estimation of CMB power spectra. There are generally two types of correlations that exist in parameter estimations from CMB; One is the 'geometrical' degeneracy which means that CMB anisotropies are exactly the same for Universes with the same baryonic and cold dark matter densities, identical primordial power spectrum and identical values of the parameter [47, 48]

$$\Re = \frac{\omega_m^{1/2}}{\omega_k^{1/2}} \sinh(\omega_k^{1/2} y) \,,$$

where $\omega_i = \Omega_i h^2 = \rho_i 8\pi G/(3 \times (100 \text{kms}^{-1}\text{Mpc}^{-1})^2)$ (physical densities of different contents) and

$$y = \int_{a_{rad}}^{1} \frac{da}{(\omega_m a + \omega_k a^2 + \omega_\Lambda a^4)^{1/2}} \,.$$

Figure 1.14 shows five models with identical C_ℓs. In linear theory, no matter how precisely experiments are done, observations of primary anisotropy cannot break this degeneracy. This introduces a limit on measuring the curvature of the Universe and the Hubble constant, just from CMB anisotropies. The other parameter degeneracies, to some extent, depend on how well experiments can be done. For instance, for a cut sky experiment, there are correlations between C_ℓ of neighbouring ℓs. The authors conclude that including external information, such as measurements of the Hubble constant, measurements of the Ω_Λ from supernovae experiments, etc. are indeed needed to break the degeneracies.

On this note, some have started combining CMB data with large scale structure data [49, 50, 51, 52, 53]. An important example is [53], where the authors performed an intensive Fisher matrix analysis on thirteen cosmological parameters of the adiabatic CDM model. A consequence of having a large parameter space is to get weaker constraints on the parameters. Indeed, the errors obtained form their work is larger than any previous work that has been done with smaller parameter spaces, such as [45, 44]. They explain this difference by numerical differences, such as derivative methodology or fiducial models, that is different in previous works. In this work, they have assumed a typical HZ primordial power spectrum having the usual power law form $P_p = Ak^{n_s}$ and they have included these parameters as part of their parameter space. They start by measuring the errors on the parameters from CMB data alone and show the decrease in the errors when the redshift survey data is added. It is shown that good constraints can be achieved on some parameters, such as $\Omega_b h^2$, $\Omega_m h^2$, n_s, α ($= d \ln n_s / dk$ — running of spectral index) and a *combination* of Ω_Λ and Ω_K from CMB data alone. But other parameters such as h and Ω_m are not constrained very well due to the degeneracies, such as the angular diameter degeneracy. A combination of Ω_Λ and Ω_K at fixed $\Omega_b h^2$ and $\Omega_m h^2$ keep the height and the position of the acoustic oscillations same; having $\Omega_m = 1 - \Omega_K - \Omega_\Lambda$, we can multiply the equation by h^2, to get the physical densities of the parameters, which is really what defines the CMB power spectrum. We then get $\omega_m = h^2 - \omega_K - \omega_\Lambda$. Now, ω_m and ω_b are enough to fix the shape and the position of the acoustic peaks as ω_Λ and ω_K do not have much effect at the time of recombination. This means that CMB data alone cannot constrain h, Ω_m and $\Omega_m h$ very well. By adding redshift surveys, the angular diameter degeneracy is lifted. This is because in the case of CMB, what we see is the *apparent* size of the scales on the sky and, as explained previously, this depends on the angular diameter distance to last scattering surface. A combination of Ω_Λ and Ω_K can have the same angular diameter distance [45, 46, 44], which means we can only know about certain combination of the two parameters — Figure 1.14 shows this degeneracy [46]. Once ω_m and ω_b are known from the CMB data, the matter power spectrum is known as well. Comparing the features (eg. the baryonic oscillations) in the real matter power spectrum with those in the redshift space gives an estimation of h and therefore, the degeneracy is lifted.

All the authors above have used Fisher matrix analysis for their analysis and there are some questions about the reliability of this technique; as mentioned previously the Fisher matrix analysis is only reliable when $\mathcal{L}$, the likelihood function, has approximately a Gaussian distribution in the parameters. Authors of [54] have tested the reliability of the Fisher matrix technique against a Markov Chain Monte

Carlo (MCMC) simulation. They use synthetic data that can be obtained from the experiment under consideration. The errors on the parameters are expected to be more reliable as they do not assume any Gaussianity about the parameters. They also investigate the whole range of the likelihood function instead of using information at only the best fit point, which makes the technique even more reliable. They perform this investigation on eleven parameters for Planck. Their results show that indeed for some parameters (for instance neutrinos) the likelihood function is not Gaussian, and therefore the MCMC results can be as different as a factor of two from the Fisher matrix approach. This definitely indicates that other external information should indeed be included to break the degeneracies (and make the likelihood function Gaussian for the Fisher matrix analysis to be reliable).

A very important set of cosmological parameters is the one parametrising primordial power spectrum. The inflationary model of the Big Bang [17, 18] and other models of the early Universe (such as ekpyrotic models [55]) predict an adiabatic scalar primordial perturbations, with an almost Gaussian distribution. The shape of the primordial power spectrum depends on the model and its parameters. However, there are quite a lot of complications involved in this. For example, there is a degeneracy between the shape of the primordial power spectrum and the cosmological parameters, such as the densities of the different components of the Universe, and constraints on the geometry of the Universe. It has been shown by [56, 57] that changing the cosmological parameters can have the same effect as changing the shape of the primordial power spectrum. This alone has a great impact on determining the shape of the primordial power spectrum; it means that it is possible to completely misinterpret the cosmological parameters and be totally biased. This leads us to rethink the *assumptions* made about the early Universe and investigate more about the physics of the CMB rather than making assumptions that just fit the data. Understanding the shape of the primordial power spectrum helps to narrow down the set of broad models of the early Universe. It helps to understand the physics of the early Universe and the initial conditions of the structure formation, which reflect on the shape of the primordial power spectrum. This will in turn help in determining other cosmological parameters [58, 59]. All this makes the primordial power spectrum one of the hot topics in cosmology.

For example a low power of the quadruple of CMB data has been detected by WMAP [32], which is not predicted by the standard model. While other possibilities of this observation is under investigation, such as [60, 61, 62], it might really be an indication of deviation from the assumed scale-invariant (HZ) initial power spectrum.

The quadrupole corresponds to the largest observable scales in the Universe, and it is not expected to have changed much from the early Universe, reflecting the primordial power spectrum. These issues make the extraction of the primordial power spectrum one of the leading problems in cosmology. Observing the Universe *today* is the only handle we have on determining the primordial power spectrum.

The authors of [59] have used the data from WMAP and 2dfGRF to construct the shape of the primordial power spectrum. They have applied a model-independent parametrisation with 16 bands in wavenumber. They have investigated four different approaches to construct the shape; power law spectra with or without a running spectral index, power spectrum having a cut-off scale, above which there is no power, power spectrum in wavenumber bands and broken power spectrum, in which, the effect of a double inflation has been investigated. They conclude that in each of above cases, a scale-invariant spectrum still fits the data perfectly well. A red tilt is slightly preferred and there is some marginal signature for a drop in power on the largest scales. Although, generally, over their 16 wavebands, the primordial power spectrum seems featureless and scale-invariant. They have also shown that allowing the shape of the primordial power spectrum to change significantly amplifies the errors on other parameters.

Another approach is taken by [63] to reconstruct the shape of the primordial power spectrum. They use the wavelet band power method of [64] and the top-hat binning method of [65] to reconstruct $P_p(k)$ as a free function. They use both CMB and LSS data. In this approach, both a scale-invariant and a power law form of $P_p(k)$ is consistent with data and they conclude that there is indeed a need for a more stringent and independent way of constraining parameters so that $P_p(k)$ can be constrained without making any assumptions about the inflationary models. The thing to note here is that binning the data have some advantages (reduces the errors in each bin) and disadvantages (there might be features smaller that the bin size and they get completely lost). Therefore, we need to find a compromise between the two to arrange the bins optimally.

Knox, in [38], has used the Fisher matrix technique to analyse parameters of initial power spectrum (as opposed to a non-parametric approach by the previously mentioned authors). These parameters include the scalar and tensor spectral indices, n_s and n_t, and the amplitude of the spectra, A_s and A_t. Therefore, he has assumed a

HZ initial power spectra defined as

$$
\begin{aligned}
P_p^s(k) &= A_s k^{n_s} \\
P_p^t(k) &= A_t k^{n_t - 3} .
\end{aligned}
\tag{1.84}
$$

Knox tries to constrain these parameters for different sets of θ_{fwhm} and w-specifications (noise) of CMB experiments. He concludes that in getting acceptable errors on the above parameters, the angular resolution is the most important factor (especially in an open Universe where, apparent sizes of objects on the sky are smaller, a very narrow beam is of great interest). For example, to be able to measure n_s with an accuracy of 0.016, an experiment with $\theta_{fwhm} = 20'$ and $w^{-1} = 6.85 \times 10^{-14}$ is necessary (cf. Planck: $\theta_{fwhm} = 10.7'$ and $w = 5.4 \times 10^{-15}$ for $\nu = 100\mathrm{GHz}$ channel). However, trying to constrain parameters other than n_s might mean that other specifications of the experiments might become more important. For instance, if we want to measure Ω_b as well as n_s, sky coverage might be more important than the angular resolution [66]. This is because of the degeneracy between n_s and Ω_b; n_s affects the tilt of the *whole* spectrum and so its slope, even at very large ℓs, is of great importance. Therefore a narrow beam is most desired for the determinations of n_s only.

On the other hand, Hu & Okamoto in [58], show that there are *physical* limitations on how best it is possible to probe the primordial power spectrum from CMB experiments alone. This is the case even if the experiments give perfect results. These limitations are imposed due to

- Geometric projection (projection of a 3D reality onto the 2D surface of a sphere around us),

- Acoustic physics (trajectory of photon from the last scattering surface to us and how they are effected on the way, eg. by ISW effect),

- Gravitational lensing,

- Degeneracy of cosmological parameters.

They suggest the best way to tackle the degeneracy of the cosmological parameters, and especially measuring the deviations from the scale-invariant initial power spectrum is to parametrise the deviations from a simple power law and to expand them in an orthonormal basis. This is based on the principal components of the covariance

matrix of the deviation parameters. These modes are like the Fourier expansion of the initial spectrum, which is localised to the acoustic regime. This provides a reliable platform to find out about the deviations from the assumed primordial spectrum.

The rest of the thesis is organised in this format:

- In Chapter 2 we will investigate the best scales upon which the primordial power spectrum can be probed, in accordance with the knowledge about other cosmological parameters such as Ω_b, Ω_c, Ω_Λ, h and τ. We will find the optimal binning of the primordial power spectrum, by making use of the Fisher matrix formalism. To investigate the correlations between the cosmological parameters, mentioned above, and a set of primordial power spectrum bins, we make use of principal component analysis (PCA) and Hermitian square root of the Fisher matrix. The surveys used in this project are Planck and SDSS (BRG), but the formalism can easily be extended to any windowed measurements of the perturbation spectrum.

- In Chapter 3 we introduce a new method for reconstructing the primordial power spectrum directly from the observations of the CMB. We employ Singular Value Decomposition (SVD) to invert the radiation perturbation transfer function. We present best–fit $P(k)$ obtained with this method along with other cosmological parameters. Only the vectors associated with the strongest signal in the decomposition are used, this assumption allows the 'pseudo-inversion' of a singular transfer function.

- In Chapter 4 we will investigate the advantages and disadvantages of a sparse sampling strategy for galaxy surveys using Bayesian Experimental Design. Measuring the galaxy/matter power spectrum is of great importance in cosmology and measuring this spectrum will enable us to measure other cosmological parameters. If we are only interested in the large scale of the power spectrum then it might be a good idea to sparsely observe a larger patch of sky, for the same observing time, rather than observing a smaller contiguous area.

- In Chapter 5 we summarise and conclude.
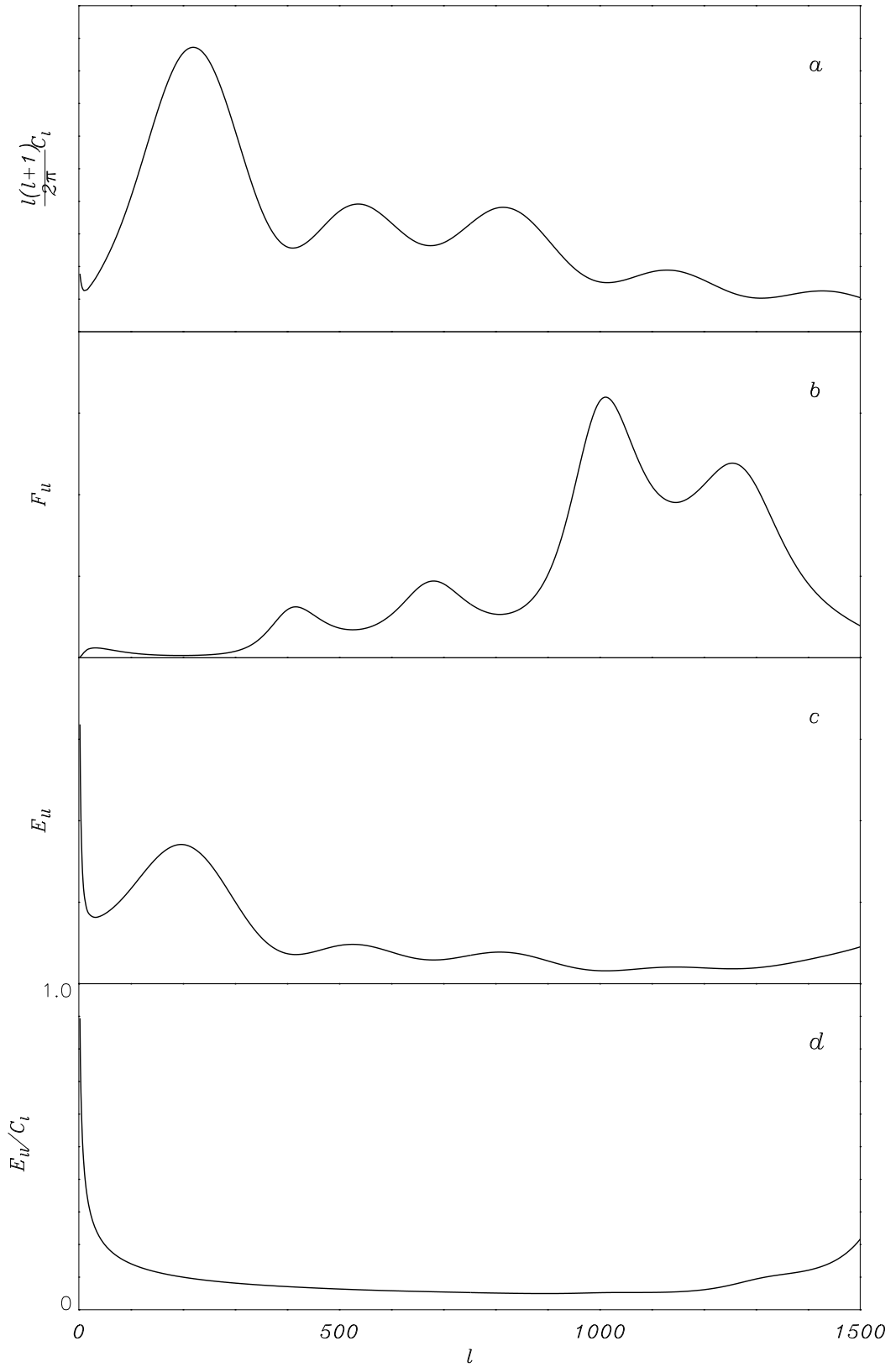
Figure 1.12: a)Angular power spectrum of the fiducial model. b)Fisher matrix of $\mathcal{C}_\ell$. c)Absolute error on $\mathcal{C}_\ell$. d)Relative error on $\mathcal{C}_\ell$.
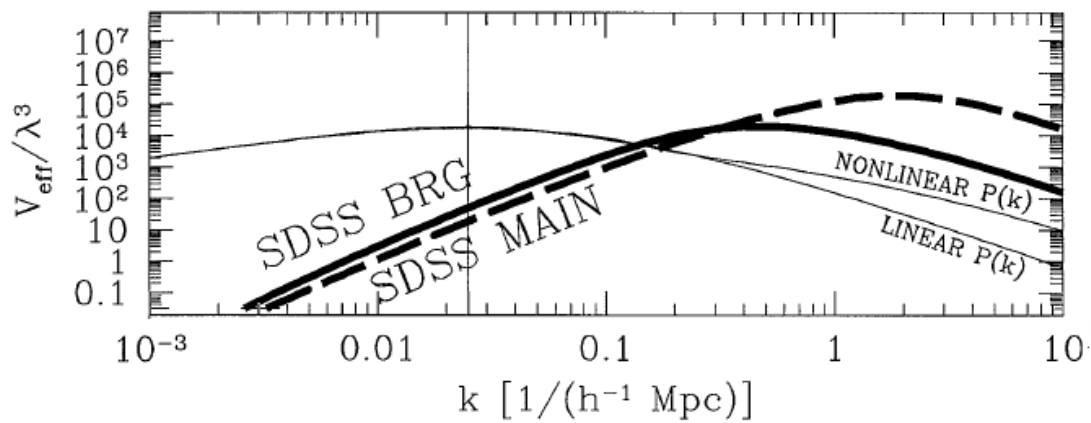
Figure 1.13: Weight function for the main part and the BRG sample of SDSS
— Figure taken from [36].
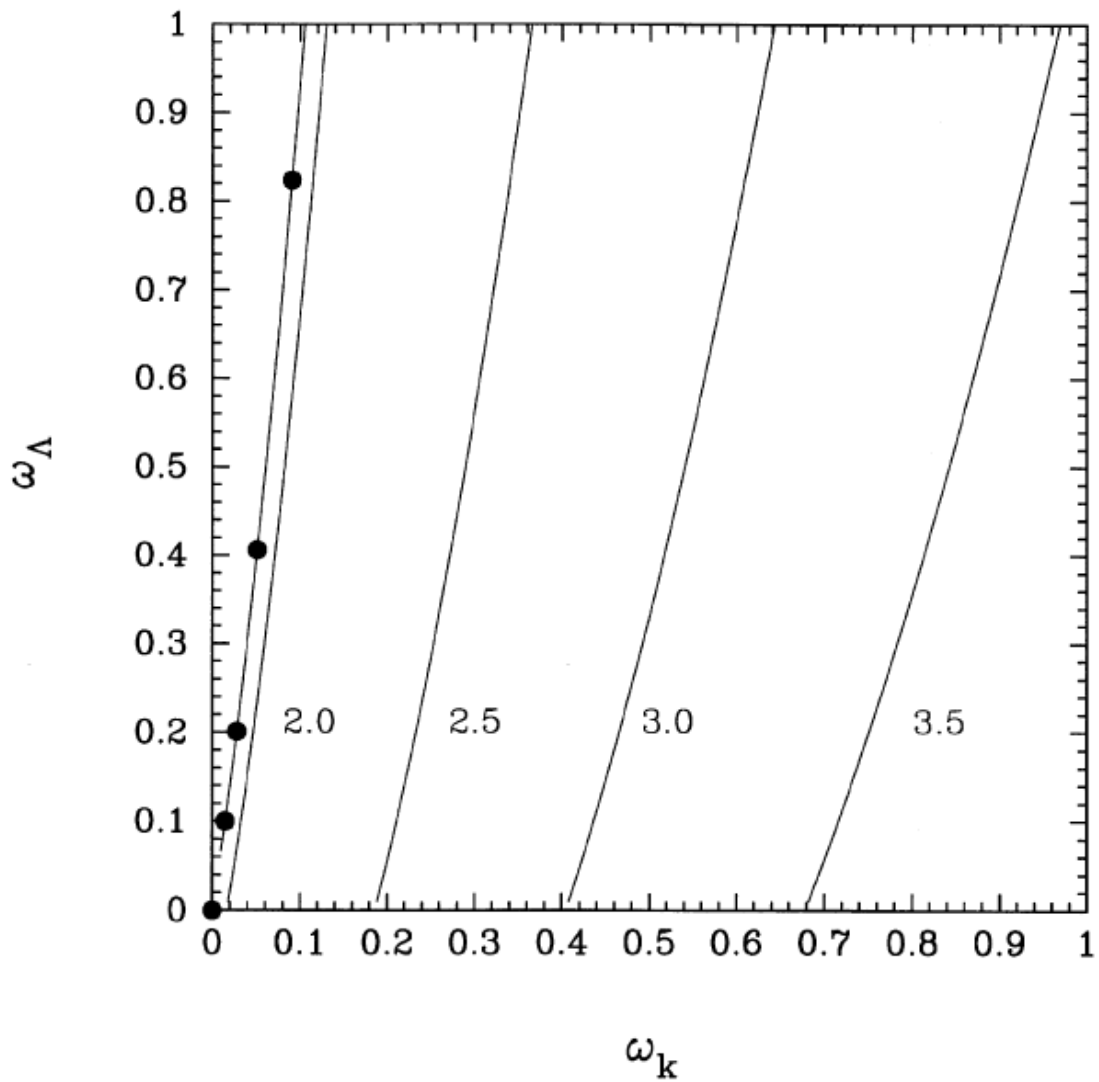
Figure 1.14: Solid lines show constant values of $\Re$ in the $w_\Lambda - w_k$ plane for $w_m = 0.250$. The five dots show the location of five models with nearly identical C_ℓs for one particular $\Re$ — Figure taken fron [46].

Chapter 2

Optimal Binning of the Primordial Power Spectrum

2.1 Introduction

The primordial power spectrum encodes the physics of the early Universe and its measurement is one of the key research areas in modern cosmology. There are many proposed models that try to describe the early Universe, out of which the theory of inflation [18, 67] is currently the most favoured. The simplest models of inflation predict almost purely adiabatic primordial perturbations with a nearly scale-invariant power spectrum (i.e., $P(k) \propto k$) — the so-called Harrison-Zeldovich spectrum. Indeed this form of power spectrum fits the current observations very well. However, there are various models for the generation of the perturbations with deviations from the perfectly scale-invariant power spectrum. The simplest are the slow-roll inflationary models which describe the deviations through a minimal scale dependence ('running') of the power law index of the power spectrum. Other models generating deviations from scale-invariance include, for example, multiple scalar fields during inflation [59, 62], multiple inflation, and various phenomenological models resulting in features such as an exponential cut-off on large scales or a power law with superimposed harmonic wiggles (due to features in the inflation potential for example) — refer to [68] and references therein. Determining the primordial power spectrum will therefore give us a better intuition into the conditions of the early Universe and help us choose from the many proposed models of the early Universe.

The drawback, however, is that we cannot measure the primordial power spectrum

directly, and our path to its measurement is through different experimental techniques such as cosmic microwave background (CMB) measurements and large scale structure (LSS) surveys. The outcome of such surveys is a convolution of the primordial power spectrum and (the square of) a transfer function which, in turn, depends on the cosmological parameters (which we will collectively call θ_i). Here we list some examples;

- For galaxy surveys, the observable power spectrum is related to the primordial power spectrum through the matter power spectrum, $P_\delta(k)$ as

$$P_g(k) = b^2(k)P_\delta(k) \simeq b^2(k)2\pi^2 k T^2(k)\Delta_\Phi^2(k) \,, \tag{2.1}$$

where Δ_Φ^2 is the primordial power spectrum and $T(k)$ is the matter transfer function and $b(k)$ is the *bias*.

- For CMB surveys, the angular power spectrum is

$$C_\ell = 4\pi \int_0^\infty d\ln k \, \Delta_\ell^2(k)\Delta_\Phi^2(k) \,, \tag{2.2}$$

where ℓ is the angular wavenumber on the sky and roughly corresponds to an angular scale on sky via $\ell \sim 180^0/\theta$ and $\Delta_\ell(k)$ is the angular transfer function of the radiation anisotropies.

As explained in the previous chapter, the primordial curvature power spectrum is defined as $\Delta_\Phi^2(k) = (k^3/2\pi)P_\Phi(k) = A(k/0.05)^{n_s-1}$, where A is the amplitude and n_s is the spectral index. Other types of power spectra, such as the weak lensing and peculiar velocity power spectra, have similar forms; they depend on the cosmological parameters, through a transfer function, as well as the primordial power spectrum. As these power spectra are *jointly* sensitive to the primordial power spectrum and θ_i, there is an induced statistical degeneracy between them.

To recover the continuous primordial power spectrum, we need to deconvolve it from discrete data such as the CMB power spectrum C_ℓ or the band powers of the LSS power spectrum. This deconvolution, along with our lack of knowledge of the cosmological parameters that determine the transfer function, induces a correlation between those parameters and those that determine the primordial power spectrum, which limits our ability to recover the primordial power spectrum completely. In some cases, such as measurements of the CMB, even a perfect survey cannot recover the primordial power spectrum completely [58]. However, the transfer function

of different techniques vary in the scales and parameters that they measure and this means the induced degeneracy between the parameters is different for different surveys. Therefore, one survey can help fill the gaps of others and together they can make significant improvements. Hence, combining surveys will enable us to improve, for example, the resolution of our measurements of the primordial power spectrum.

One method to recover the primordial power spectrum is to measure its amplitude in a series of bins. The more bins with high signal-to-noise (S/N), the more accurately the power spectrum can be reconstructed. The aim of this work is to find an optimal binning for the primordial power spectrum based on the knowledge (or, better, lack of knowledge) of other cosmological parameters from some specific surveys, such as SDSS [33] and Planck [40]. We then want to quantify the correlations between these carefully chosen bins. Therefore, the aim is to explore, as the data improve, what new information can be learnt about the primordial power spectrum and what exactly needs to be improved to better constrain the primordial power spectrum. The motivation is to test the assumptions about the initial conditions and get better constraints on parameters based on the same set of assumptions. Acknowledging the degeneracy between the cosmological parameters and the primordial power spectrum, we want to investigate the scales the primordial power spectrum can be probed best with future experiments.

One common method for error estimation is to use a Fisher matrix analysis. The Fisher matrix is generally used to determine the sensitivity of a particular survey to a set of parameters and has been largely used for forecasting and optimisation. It is defined as the ensemble average of the *curvature* of a function $\mathcal{F}$ (i.e., it is the average of the curvature over many realisations of signal and noise);

$$F_{\alpha\beta} = \langle \mathcal{F} \rangle = \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_\alpha \partial \theta_\beta} \right\rangle , \qquad (2.3)$$

where $\mathcal{L}$ is the likelihood function. Its inverse is an approximation of the covariance matrix of the parameters, by analogy with a Gaussian distribution in the θ_α, for which this would be exact. Therefore, it allows us to estimate the errors on parameters without having to cover the whole parameter space. The authors of [37] have compared the Fisher matrix analysis with the full likelihood function analysis and found there was great agreement between the two methods if the likelihood function is approximately Gaussian near the peak.

The Cramer-Rao inequality[1] states that the smallest error measured, for θ_α, by any unbiased estimator (such as the maximum likelihood) is $1/\sqrt{F_{\alpha\alpha}}$ and $\sqrt{(F^{-1})_{\alpha\alpha}}$, for non-marginalised and marginalised[2] one-sigma errors respectively.

We further note, as in all uses of the Fisher matrix, that any results thus obtained must be taken with the caveat that these relations only map onto realistic error bars in the case of a Gaussian distribution, usually most appropriate in the limit of high signal-to-noise ratio and/or relatively small scales, so that the conditions of the central limit theorem apply. As long as we do not find extremely degenerate parameter directions, we expect that our results will certainly be indicative of a full analysis, using simulations and techniques such as Bayesian Experimental Design [69, 70] (and P. Paykari & A. H. Jaffe (2009, in preparation)).

The Fisher matrix for CMB surveys is given by

$$F_{\ell\ell'} = f_{\text{sky}} \frac{2\ell+1}{2} \delta_{\ell\ell'} [C_\ell + w^{-1}e^{\ell^2\sigma^2}]^{-2} \; , \tag{2.4}$$

where C_ℓ is the angular power spectrum, w is the weight defined as $(\Delta\Omega\sigma_n^2)^{-1}$ with $\Delta\Omega$ being the real space pixel size and σ_n^2 being the noise per pixel, $e^{-\ell^2\sigma^2}$ is the window function[3] for a Gaussian beam (where $\sigma = \theta_{fwhm}/\sqrt{8\ln 2}$) and f_{sky} is the fraction of the sky observed. The factor $f_{\text{sky}}(2\ell+1)$ gives the number of independent modes at a given wavenumber; the term proportional to C_ℓ is the sample (or cosmic) variance contribution, and the $w^{-1}e^{\ell^2\sigma^2}$ term is the noise contribution. Note that the diagonal form for the matrix implies diagonal (uncorrelated) errors on the C_ℓs.

The $F_{\ell\ell'}$ gives the errors on the C_ℓs. Therefore, to find the errors on other parameters, we use the Jacobian

$$F_{\alpha\beta} = \sum_{\ell\ell'} F_{\ell\ell'} \frac{\partial C_\ell}{\partial\theta_\alpha} \frac{\partial C_{\ell'}}{\partial\theta_\beta} \; , \tag{2.5}$$

where θ_α and θ_β are different parameters.

---

[1] It should be noted that the Cramer-Rao inequality is a statement about the so-called "Frequentist" confidence intervals and is not strictly applicable to "Bayesian" errors.

[2] Integration of the joint probability over other parameters.

[3] This damps power on larger ℓs; as we get closer to the resolution limit of the survey C_ℓs start to correlate.

For a volume-limited galaxy survey the Fisher matrix [36] is [4]

$$F_{nn'} = \delta_{nn'} \frac{k_n^2 \Delta k V}{(2\pi)^2 (P_n + 1/\bar{n})^2} \, , \qquad (2.6)$$

where V is the total volume of the survey, $\bar{n}$ is the number density of the survey (N_{tot}/V), P_n is the galaxy power spectrum in each k_n bin and Δk is the binwidth. Similar to the CMB power spectrum case, $k_n^2 \Delta k V$ counts the number of independent modes, P_n gives the sample variance, and $1/\bar{n}$ the noise variance due to Poisson counting errors. This, again, gives us the errors on the galaxy power spectrum and we use the Jacobian to get the errors on other parameters

$$F_{\alpha\beta} = \sum_{nn'} F_{nn'} \frac{\partial P(k_n)}{\partial \theta_\alpha} \frac{\partial P(k_{n'})}{\partial \theta_\beta} \, . \qquad (2.7)$$

Fisher matrices for different surveys can easily be combined by simple summation: $\mathbf{F} = \mathbf{F}_{galaxy} + \mathbf{F}_{CMB}$; they are proportional to the log of the likelihood function and this is equivalent to the multiplication of independent likelihoods to combine them. Equivalently, we can think of them as the *weights* (inverse noise variance) of the experiments, which add for a Gaussian distribution. The nonzero correlation between the parameters in the covariance matrix makes interpreting the errors somewhat more difficult than the uncorrelated case. We will discuss various methods for decorrelating the power spectra and cosmological parameters.

2.2 Method

The aim is to investigate the primordial power spectrum in a "non-parametric" way (we use quotations remarks to remind the reader that "non-parametric" merely means that we use a very general model, potentially with a very large number of parameters). For this purpose, we define the primordial power spectrum as a series of top-hat bins:

$$\Delta_\Phi^2(k) = \sum_B w_B(k) Q_B \, , \qquad (2.8)$$

where Q_B is the power in each bin B and $w_B = 1$ if $k \in B$ and 0 otherwise. The cosmological parameters under investigation are (and of the form) Ω_c, Ω_b, Ω_Λ,

---

[4]Note that this equation only applies to linear regime, as non-linearities impose non-Gaussianities.

h and τ. We will choose a geometrically flat ΛCDM model with adiabatic perturbations and the WMAP5 [24] values for the parameters: $\Omega_m = 0.214 \pm 0.027$, $\Omega_b = 0.044 \pm 0.003$, $\Omega_\Lambda = 0.742 \pm 0.03$, $\tau = 0.087 \pm 0.017$ and $h = 0.719 \pm 0.0265$, where $H_0 = 100h\mathrm{km}^{-1}\mathrm{Mpc}^{-1}$. $\Omega_\nu = 0.0$ was chosen, as massive neutrinos introduce some difficulties in the Fisher matrix analysis [53] and therefore were ignored for now. A CMB temperature of 2.726 and He fraction of 0.24 were chosen. The CMBfast software [21] was used to generate the matter and CMB power spectra. The surveys chosen for this initial investigation are the projected results from the SDSS Bright Red Galaxies (BRG)[5] sample and the Planck Surveyor CMB Power Spectrum[6].

2.2.1 Application to Surveys

2.2.1.1 Galaxy Surveys — SDSS(BRG)

A galaxy power spectrum is related to the matter power spectrum via a parameter called *bias* — Equation 2.1. For the BRG sample of SDSS, this is assumed linear and scale-independent with the form $P_g(k) = b^2 P_\delta(k)$, where the bias is $b \simeq 2.0$ [71, 72, 73, 74]. The survey specifications for BRG sample are $\bar{n} = 10^5/V$ and $V = (1h^{-1}\mathrm{Gpc})^3$[33].

For the θ_i the derivatives in the Jacobian were obtained numerically using the Taylor expansion

$$P(\theta_i) = P(\theta_0) + (\frac{\partial P}{\partial \theta_i})\Delta(\theta_i) . \tag{2.9}$$

The width and direction of the step are quite important here. A two-sided derivative was chosen, so that the derivative is centred on the default value θ_0, with a step size of $\Delta(\theta_i)/2$ on each side. This *is* accurate to 2nd order in $\Delta(\theta_i)$ (a one-sided derivative would be at a slightly shifted place of $\theta_i + \Delta(\theta_i)/2$, and is only accurate to 1st order [53]). The width of the step should be small enough to give accurate results and yet big enough to avoid numerical difficulties. This was taken to be a 5% variation, therefore a 2.5% width on each side. Other studies have shown that this turns out to be the best step size, giving the most accurate results [53].

---

[5]These are bright galaxies, which means the survey will be quite deep, with $z \sim 0.25 - 0.5$. Also, these trace the elliptical galaxies, which are thought to be better tracers of mass at this redshift range.

[6]`http://www.rssd.esa.int/SA/PLANCK/docs/Bluebook-ESA-SCI\textbackslash{}\`
`%282005\textbackslash{}\%291\textbackslash{}\_V2.pdf`

For the primordial power spectrum bins, the derivative is proportional to the matter transfer function

$$\frac{\partial P_g(k_n)}{\partial \Delta_\Phi^2(k_{n'})} = 4 \times 2\pi^2 k T^2(k) \delta_{nn'} \; , \tag{2.10}$$

where n and n' refer to the bins. The k-range for SDSS is $0.006 \lesssim k/(h\mathrm{Mpc}^{-1}) \lesssim 0.1$. The minimum value for the wavenumber, k_{min}, is obtained from the largest scale of the survey — $(2\pi/V^{1/3})$. Its maximum value, k_{max}, is chosen to avoid non-linearities — scale at which $\Delta^2 \sim 1$. Simulations of a very similar flat model [75] suggested a k_{max} of $0.1h\mathrm{Mpc}^{-1}$. This is also very close to the scale at which departures from linear theory was seen by [76].

The derivatives in the Jacobian need to be averaged into bins. In Section 2.2.2 we will explain the criteria for choosing the widths and locations of the bins.

2.2.1.2 CMB Surveys — Planck

One thing to note in the case of CMB power spectra is that the output of CMBfast is of the form $\mathcal{C}_\ell = [\ell(\ell+1)/2\pi]C_\ell$, so the CMB Fisher matrix, Equation 2.4, is multiplied by this traditional factor $\ell(\ell+1)/2\pi$. The specifications for Planck HFI (we use the $\nu = 100\mathrm{GHz}$ channel and conservatively assume that other frequencies are used for foreground cleaning) are $\theta_{fwhm} = 10.7' = 0.003115$ radians, $\sigma_{pix} = 1.7 \times 10^{-6}$, $w^{-1} = 0.028 \times 10^{-15}$ [40]. The derivatives in the Jacobian were again obtained numerically by the Taylor expansion

$$\mathcal{C}_\ell(\theta_i) = \mathcal{C}_\ell(\theta_0) + (\frac{\partial \mathcal{C}_\ell}{\partial \theta_i})\Delta(\theta_i) \; . \tag{2.11}$$

The same arguments as in the SDSS case applies for the width and direction of the step here. In the case of the primordial power spectrum bins, the derivative becomes [7].

$$\frac{\partial \mathcal{C}_\ell}{\partial \Delta_\Phi^2(k)} = 2\ell(\ell+1) \int_{k_{min}^B}^{k_{max}^B} dk \; |\Delta_\ell(k)|^2 \; . \tag{2.12}$$

This needs to be averaged into k bins, as discussed in Section 2.2.2. The chosen k-range for Planck is $0.0001 \lesssim k/(h\mathrm{Mpc}^{-1}) \lesssim 0.1$, where k_{min} was obtained from $k_{min} = \ell_{min}/d_A = 2/d_A$, where d_A is the angular diameter distance to the surface of last scattering obtained to be $\sim 14\mathrm{Gpc}$ [24].

---

[7]To obtain $\Delta_\ell(k)$, CMBfast needed to be altered to give the radiation transfer functions at all ℓs. Then, for each ℓ, this was interpolated in k.

As explained above, to combine data from different surveys, we can add the Fisher matrices obtained for each of them. We expect to see an improvement on the errors of both the bins and cosmological parameters. Equivalently, this will enable us to have narrower bins without sacrificing Signal-to-Noise per bin.

2.2.2 Optimal Binning

As explained before, a set of primordial power spectrum bins form part of our parameter space. In this section we will explain how these bins are chosen. We construct the bins to have the same contribution to the Fisher matrix; that is, they each have the same S/N. We take the signal in each bin to be the amplitude of the primordial power spectrum in that bin and the noise to be given by the inverse of the square root of the diagonal elements of the Fisher matrix. For this, we construct a diagonal signal matrix, $\mathbf{S}$, which contains the amplitude of the primordial power spectrum for all the bins and the values of the cosmological parameters as its diagonal elements. We weight our Fisher matrix by this matrix

$$F'_{\alpha\beta} = S_\alpha F_{\alpha\beta} S_\beta \,. \tag{2.13}$$

This now gives us a $(S/N)^2$ matrix, where the square root of its diagonal elements are the S/N for the bins, and the weighted errors for θ_i. It is worth emphasising that it is this $(S/N)^2$ Fisher matrix that will be diagonalised later on.

Our algorithm will result in more bins where the signal-to-noise ratio is greater, sampling more finely where the signal is strongest (we will see this explicitly in our discussion of the CMB power spectrum in Section 2.3.1.2, which has considerable structure and therefore varying S/N). There are circumstances in which we might instead want to place bins by hand if we are looking for specific features (but of course we will always be limited by the S/N of our measurements).

For SDSS, we start with the maximum number of bins possible in our k-range, set by the usual properties of the Fourier transform. These imply that the scale of the survey not only determines k_{min}, but also limits our resolution: $k_{min} = (\Delta k)_{min} = (2\pi/V^{1/3})$; narrower bins would become highly correlated. Therefore, we set up a series of bins with this minimum width in our k-range. We then construct a Fisher matrix for this set of bins (and θ_i) and weight it by the signal matrix, $\mathbf{S}$, for this set. With this binning adopted, the S/N values range from 3.7 in the first bin to 35.1 in the last bin. Knowing that the binwidths chosen are the minimum possible and that

increasing binwidths will increase the S/N value, we conclude that the bin with the maximum S/N cannot be changed and hence we make other bins wide enough to reach the S/N in this bin. To obtain this "optimal" binning we add small bins until their S/N is within 15% of the maximum S/N:

$$\frac{\text{Max}(S/N) - (S/N)_i}{\text{Max}(S/N)} = 0.15 \,, \tag{2.14}$$

where i refers to the bins. This finally gives us 8 bins with their S/N ranging $30-35$.

For Planck, the bins are obtained so that their S/N matches that of SDSS. The reason for applying this criteria to Planck is to allow for a fair comparison between the results from SDSS and Planck. This criteria gives us a total of 23 bins for Planck.

In the case of the combined Planck and SDSS we require only that the S/N of the bins are equal within 50%. This now gives us the optimal resolution of the primordial power spectrum we can achieve from SDSS and Planck. We have a total of 48 bins with S/N being in the vicinity of ~ 20.0 and, therefore, still comparable to the S/N values in the other cases.

An alternative, and perhaps more appropriate, way to determine the binning would be to take the marginalised errors as the noise. This would be obtained by inverting the Fisher matrix in each iteration to get the covariance matrix, which gives the marginalised covariances of the bins and θ_i. We would then take the sub-block of this covariance matrix that refers to the bins only, and invert it to get a *marginalised* Fisher matrix for the bins. We would then feed this Fisher matrix into Equation 2.13. However, this method could not be implemented because the SDSS Fisher matrix is not invertible; the SDSS Fisher matrix is not a positive definite matrix because it is asked to estimate too many parameters. There are a total of n data points (n galaxy power spectrum bins) and we are asking these to predict $n + m$ parameters (n primordial power spectrum bins and m θ_i). Also, note that whichever of the methods presented uses the *correlated* errors as the noise. We now discuss the decorrelation of the parameters.

2.2.3 Decorrelating the Parameters

2.2.3.1 Principal Component Analysis

Having obtained a set of bins, and therefore parameters we wish to estimate (or, in this case, forecast), we of course find that for these realistic experiments the parameter errors are often highly correlated (i.e., the Fisher matrix has significant off-diagonal components). To overcome this problem, we make use of principal component analysis (PCA) to obtain an orthogonal basis (onto which the original parameters will be projected); the covariance(/Fisher) matrix is a symmetric $n \times n$ matrix and therefore, can be diagonalised using its eigenvectors. This has the form $\mathbf{C} = \mathbf{E^T \Lambda E}$, where $\mathbf{C}$ is the covariance matrix, $\mathbf{E}$ is an orthogonal matrix with the eigenvectors of $\mathbf{C}$ as its rows and $\mathbf{\Lambda}$ is the diagonal matrix with the eigenvalues of $\mathbf{C}$ as its diagonal elements[8]. This constructs a new set of variables $\mathbf{X}$ that are orthogonal to each other and are a linear combination of the original parameters $\mathbf{O}$, through the eigenvectors

$$\mathbf{X} = \mathbf{EO} \,. \tag{2.15}$$

The X_i are called the principal components of the experiment and are ordered so that so that X_1 has the smallest eigenvalue and X_n the largest. In this construction, the eigenvalues are the variances of the new parameters, so X_1 and X_n are the best- and worst-measured components respectively. The eigenvectors have been normalised so that $\sum_j e_j^2 = 1$, where e_j are the elements of E_i. We list some properties of PCA below:

- The main point of PCA is to assess the degeneracies (correlations) amongst the parameters that are not resolved by the experiments, be they fundamental as from cosmic variance or due to the noise and coverage of the experiment. In our case, it will especially help us to see the correlation amongst the bins of the primordial power spectrum, and between the bins and the cosmological parameters.

- The eigenvalues obtained measure the performance of the experiment — a larger number of small eigenvalues means a better experiment. Another measure of the performance of the experiments is to see how they mix physically independent parameters such as, say, n_s, the spectral index, and Ω_b. This sort

---

[8]It is common to construct the covariance matrix for PCA. However, the Fisher matrix can be used instead; the eigenvectors stay the same, but eigenvalues are reciprocals.

of mixture may be improved by improving the experiment's noise properties or increasing its area or volume[9].

All the above points may be summed up to conclude that in a perfect setting (if we could resolve cosmic variance and the geometrical degeneracy) we would expect a one-to-one relation between the old and the new parameters — the Fisher matrix would be diagonal. Each of the original parameters would contribute to one and only one of the new parameters, with zero contribution from the others.

Note that the principal components obtained are not strictly unique and depend on the form of the variables (e.g., whether we use Ω_b or $\log \Omega_b$), as well as where they are evaluated.

2.2.3.2 Hermitian Square Root

Another approach to remove the correlations between the uncertainties is to use the Hermitian square root of the Fisher matrix as a linear transformation on the parameter space [37, 77, 78]. This transformation is obtained by

$$\mathbf{F}^{1/2} = \mathbf{E^T} \mathbf{\Lambda}^{1/2} \mathbf{E} \,, \tag{2.16}$$

where, like before, $\mathbf{E}$ is the eigenvector matrix and $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues. It has the property $\mathbf{F} = \mathbf{F}^{1/2}\mathbf{F}^{1/2} = (\mathbf{F}^{1/2})^{\mathbf{T}}\mathbf{F}^{1/2}$ and therefore the condition $(\mathbf{F}^{-1/2})\mathbf{F}(\mathbf{F}^{-1/2}) = (\mathbf{F}^{-1/2})^{\mathbf{T}}\mathbf{F}(\mathbf{F}^{-1/2}) = diag$ is satisfied. Unlike PCA, $\mathbf{F}^{1/2}$ does not give us an orthogonal basis and instead can be thought of as giving "window functions" for the primordial power spectrum resulting in uncorrelated parameters (in the Gaussian limit). We define a window matrix by

$$H_{nm} = \frac{(F^{1/2})_{nm}}{\sum_n (F^{1/2})_{nm}} \,, \tag{2.17}$$

which satisfies the normalisation condition $\sum_n H_{nm} = 1$. Hence the windowed power spectrum is defined as

$$\tilde{P}_m = \sum_n H_{nm} P(k_n) \,, \tag{2.18}$$

---

[9]However, the so-called "geometrical degeneracy" [47, 48] cannot be improved by improving the experiments; two models with same primordial power spectrum, the same matter content, and the same comoving distance to the surface of last scattering produce identical CMB power spectrum.

where $P(k_n)$ is the original primordial power spectrum. Note that this windowed power spectrum is constructed for a visual presentation and understanding of the underlying correlations (indeed, it can be manifestly unphysical if, as we will see, the window function is negative). Again, in a perfect setting — a diagonal Fisher matrix — we would expect this windowed power spectrum to be equal to the primordial power spectrum (i.e., with each window function comprising a single bin).

We obtain this window matrix for the marginalised Fisher matrix of the bins and hence it can only be applied to the Fisher matrices of Planck and the combination of Planck and SDSS, which are invertible.

2.3 Results

2.3.1 PCA

The principal components,X_i, obtained for SDSS, Planck and their combination are shown as colour-coded matrix plots; X_i are shown from left to right with increasing errors (which is equal to $1/\Lambda_i^{1/2}$, as the eigenvalues are constructed for the Fisher matrix). The original parameters, O_i, are shown vertically starting with the bins on the bottom to θ_i on the top. For the bins, the vertical width of the box is an indication of the binwidth. The colour coding is shown at the bottom of Figure 2.1; the plots are showing $X_i = \sum_j E_{ij}O_j$, where j^{th} element in E_{ij} is shown as a colour according to the colour plot.

2.3.1.1 SDSS

The result is shown in Figure 2.1. There are a total of 8 bins that could be obtained to meet the S/N criteria as explained above. Together with the 5 θ_i, we have a total of 13 original parameters, O_i, and 13 principal components, X_i. The last 5 principal components are not measured well (they have large/negative eigenvalues — Table 2.1). This is because, as explained before, the SDSS Fisher matrix is not a positive definite matrix; we have a total of 8 data points and this means only 8 parameters (or 8 different combinations of the parameters, i.e., X_i) can be measured.

The best measured principal component, X_1, has contributions from the cosmological parameters (θ_i) only, with h dominating. The fact that there is more than one cosmological parameter contributing to this principal component means that SDSS

can only measure a [linear] combination of them — evincing a degeneracy between these parameters. X_2 measures a combination of the bins and θ_i, which suggests a degeneracy between the highlighted bins, h and Ω_c. Other principal components, $X_3 - X_8$, measure the bins only, with no contribution from θ_i at all, and the correlation amongst the bins is between neighbouring ones only. Intuitively, one expects more correlation between the bins and the cosmological parameters. The errors for the bins are related to the matter transfer function — Equation 2.10. Therefore, one might expect a change in θ_i would induce a change in the matter transfer function and hence a correlation between bins and θ_i. However, consider Figure 2.4 showing the individual derivatives that contribute to the Jacobian. The derivatives with respect to θ_i are of similar magnitudes (apart from τ which was multiplied by 200 for presentation). However, the derivative with respect to the primordial power spectrum bins is rescaled by 10^{-8} to fit in the same range as the rest of the derivatives. This suggests that perhaps the changes in θ_i are not large enough in this setting to have a significant effect on the matter transfer function and therefore the correlation is not that significant to show effects in the PCA.

Note that the correlation between the bins shows the limits to what we can learn about the primordial power spectrum. This correlation arises due to our lack of knowledge of the cosmological parameters. If we knew the parameters perfectly, we would have what is shown in Figure 2.3 (which is in fact the Fisher matrix itself). Generally it seems that SDSS measures cosmological parameters better than the primordial power spectrum and considering the primordial power spectrum, measures small scales better than large.

We also investigated what improvements we would see given better — realistic — knowledge of the cosmological parameters. Hence, WMAP5 priors [24] were added to constrain the θ_i in the Fisher matrix, by adding the inverse variance of each parameter to the Fisher matrix, i.e., ignoring the WMAP5 correlations. The result is shown in Figure 2.2. Note that the errors on the principal components have reduced and now all X_i can be measured well — Table 2.1. Some of the degeneracies between the cosmological parameters have been broken. For example, Ω_b and τ dominate completely in X_{10} and X_{12} respectively, with no contribution from any other parameter. This is expected as WMAP5 does a good job measuring these cosmological parameters. With respect to the primordial spectrum, these improved constraints on cosmological parameters have only helped to measure linear combinations of the bins better and have not been able to break the degeneracy between them.

2.3.1.2 Planck

For Planck there are a total of 23 bins and, with the 5 θ_i, we have 28 principal components, shown in Figure 2.5. They are all measured well and better than SDSS — Table 2.1. The reflection of the acoustic peaks of C_ℓs on the bin sizes can clearly be seen; those corresponding to the peaks are measured with better resolution (see Figure 2.9 to see a pictorial version of the contributions to the Jacobian. The summation over ℓ in the Fisher matrix gives the oscillatory feature seen in k space.)

Just like SDSS, Planck measures the cosmological parameters better than the primordial power spectrum and overall, gives smaller errors and smaller correlations between them than SDSS. This is not surprising as we already know Planck does a good job measuring the cosmological parameters; it measures Ω_Λ and h very well, with only slight correlation with other cosmological parameters.

The rest of principal components contain the highly-correlated bins only, with no single large contribution from any of them. Intuitively one might expect the correlation to be between neighbouring bins only. The reason for the longer-range correlation lies in the form of the radiation transfer function; for each ℓ, this transfer function spans a k-range around $k = \ell/d_A$, where d_A is the angular-diameter distance to the last-scattering surface. This is due to the projection of a 3D Universe onto a 2D sphere around us. Equation 2.12 shows the contribution to the Jacobian for the Fisher matrix analysis. For each ℓ, this derivative integrates the radiation transfer function over the k-range of the bins. This would be reflected as correlation between neighbouring bins. However, remember that in the Fisher matrix analysis the ℓs are summed over (equation 2.5) and this induces correlation between all bins; Figure 2.9 shows a pictorial version of Equation 2.12, weighted by the primordial power spectrum. Note how each ℓ spans a range of k. The summation over all ℓs means that, for example, the bin with $\ell = 400$ dominating has contributions from all ℓs from 100 to 500, with each spanning a different range of k. This induces correlation between bins of all scales.

This correlation between small and large scales might even be worse when there is a degeneracy between the measured cosmological parameters. For example, consider an experiment with which we attempt to measure two parameters, such as the spectral index n_s and Ω_b, where n_s is dominant on large scales and Ω_b is dominant on small scales — Figure 2.8. The degeneracy between these parameters could induce a degeneracy between large and small scale bins.

Figure 2.6 shows the principal components for the bins with no θ_i — i.e., assuming cosmological parameters are known perfectly. Since Planck measures the cosmological parameters very well, we might expect little change. Indeed, not much is changed. Note, however, that the smallest error for this set is still larger than the smallest error for the set including θ_i. This is because the θ_i are individually measured better than the primordial power spectrum bins and hence they reduce the errors; instead, comparing the largest errors of both sets shows the improvements. Despite the smaller errors for this set, the correlation between the bins is not significantly improved.

We now consider the correlation between the bins for the *marginalised* Fisher matrix (that is, marginalised over the other cosmological parameters, θ_i). This is obtained by inverting the parent Fisher matrix to get a covariance matrix, giving the marginalised errors for all the parameters (in the Gaussian limit). We can further take the sub-block of this matrix which holds the errors for the bins and invert it to obtain a *marginalised* Fisher matrix for the bins alone. The principal components for this Fisher matrix are shown in Figure 2.7. The first thing to note is that bins contribute more significantly to some of the principal components. In particular there are some mid-scale bins which seem to be measured well. For example, consider X_{19} - X_{22}; the marginalisation has uncorrelated some mid-scale bins from the rest of the bins.

Another interesting result is that very large and very small scales never really dominate in the principal components with large errors. They only contribute to them at levels of $\lesssim 0.01$. Recall that X_i with large errors are the most highly-correlated and therefore the fact that mid-scale bins do not contribute to these principal components means that they are measured quite well.

In conclusion, Planck will largely decorrelate the primordial power spectrum from the θ_i (and therefore the transfer function) but cannot completely uncorrelate the bins themselves.

2.3.1.3 Planck & SDSS

The results are shown in Figure 2.10. Combining surveys has clearly helped to improve the resolution of the primordial power spectrum. The data now support a total of 48 bins in the same k-range. Again the cosmological parameters are measured better than the primordial power spectrum and there is also smaller correlation between the cosmological parameters compared to the previous cases. There is also

smaller correlation between the bins themselves. Features of both SDSS and Planck can clearly be seen here. For example, acoustic oscillations in the C_ℓs still influence the bin sizes and resolution of the primordial power spectrum. It also seems like small scales are measured better than large scales, which is a feature seen in the SDSS case.

Figure 2.11 shows the results for the marginalised Fisher matrix of the bins for SDSS and Planck combined. Compare to Figure 2.10; not much change can be seen.

2.3.2 Hermitian Square Root of Fisher Matrix

Figure 2.12 shows the window functions for Planck derived from the Hermitian square root decorrelation. Note that only the magnitude of the components of H_ms are important and not their sign. However, it is worth mentioning that for the non-marginalised Fisher matrix (both for Planck and its combination with SDSS), the window functions have only positive values. Therefore, the lack of knowledge of the cosmological parameters (and the induced correlation between the bins) induces non-physical negative values into the window functions. The window functions, H_ms, are plotted in the order of increasing errors, so that H_1 is the best-measured and H_{23} the worst-measured vector, respectively. Here, small scales are measured best and large scales worst, contributing to H_ms with the lowest and highest errors respectively. Planck cannot decorrelate the bins completely and some correlations between *neighbouring* bins can be seen. In addition, large scales (i.e., bins in the range of $k \sim 0.02 - 0.04h\mathrm{Mpc}^{-1}$) have a large contribution to the H_m, compared to the bins on smaller scales. These window functions clearly show the influence of cosmic variance. Compare this to Figure 2.13, where we diagonalised the marginalised Fisher matrix through its eigenvectors (This is exactly Figure 2.7 plotted in this form for easier comparison). In the PCA case, the correlations seem not to be only between neighbouring bins, but between bins of all scales, which is not seen in this case. Also, the compactness seen here (i.e. more of a traditional window-function feature) is not seen in the PCA case; there is no particular scale that contributes significantly to the principal components.

Figure 2.16 shows the windowed power spectrum for Planck. It is plotted so that each $\tilde{P}_m$ is placed at the k_n from which it receives the largest contribution. The vertical errors bars shown are $\Delta_\Phi^2(k_n)(\mathbf{H}\mathbf{F}^{-1}\mathbf{H}^T)$, where $\Delta_\Phi^2(k_n)$ is the amplitude of the primordial power spectrum in the bins and $(\mathbf{H}\mathbf{F}^{-1}\mathbf{H}^T)$ is the errors propagated through the H_m distribution. The horizontal error bars are the half-width at half-

SDSS	X_1	X_2	X_3	X_5	X_7	X_9	X_{11}	X_{13}	—
No priors	0.0038	0.0196	0.0288	0.0331	0.0349	3E5	NaN	NaN	—
No θ_i	0.0282	0.0299	0.0317	0.0340	0.0359	—	—	—	—
WMAP5 priors	0.0037	0.0189	0.0287	0.0330	0.0344	0.0423	0.1233	0.5871	—

Planck	X_1	X_2	X_3	X_5	X_{10}	X_{15}	X_{20}	X_{25}	X_{28}
PCA-No priors	0.0004	0.0022	0.0094	0.0209	0.0300	0.0402	0.0516	0.0716	0.2576
PCA-No θ_i	0.0110	0.0149	0.0181	0.0224	0.0333	0.0422	0.0572	—	—
PCA-Margin.	0.0181	0.0210	0.0250	0.0297	0.0384	0.0514	0.0679	—	—
Hermitian Sqrt	0.0165	0.0184	0.0188	0.0190	0.0218	0.0496	0.1164	—	—

Planck & SDSS	X_1	X_2	X_3	X_{10}	X_{20}	X_{30}	X_{40}	X_{50}	X_{54}
PCA-No priors	0.0004	0.0020	0.0073	0.0353	0.0469	0.0552	0.0643	0.1135	0.2068
PCA-No θ_i	0.0561	0.0568	0.0575	0.0624	0.0729	0.0891	0.1003	—	—
PCA-Margin.	0.0227	0.0288	0.0304	0.0408	0.0514	0.0578	0.0734	—	—
Hermitian Sqrt	0.0233	0.0234	0.0235	0.0243	0.0384	0.0736	0.5616	—	—

Table 2.1: Errors for different sets for SDSS, Planck and combination of Planck and SDSS.

maximum in each direction of the main peak of each H_m. The original primordial power spectrum is plotted for comparison. Remember that $\tilde{P}_m$ is not a physical power spectrum *per se*. However, the observed differences from the original power spectrum arise due to the induced correlations between the bins. In a perfect setting, where there are no correlations between bins, we expect $\tilde{P}_m = \Delta^2_\Phi(k)$. Note that the main feature of this plot is that vertical errors, unlike those for the original primordial power spectrum, are *not* correlated. The correlation between the errors has been transferred to overlaps between the window functions — as shown in Figure 2.12.

Figures 2.14 and 2.17 show the same set of results for combination of Planck and SDSS. Again large scales contribute to H_m with the largest errors. There is less correlation between neighbouring bins compared to the Planck case. Also, note that bins in this case are narrower and therefore correlation between neighbouring bins still means correlation between a narrower range of k. Compare Figure 2.14 to Figure 2.15 (same as Figure 2.11). Again, the PCA case has wider effective bins, more than observed for Planck on its own. Figure 2.14 indicates that bins in the vicinity of $k \sim 0.02 - 0.025 h\mathrm{Mpc}^{-1}$ contribute very strongly to the H_m compared to other bins, in particular the last window function, H_{48}. This effect is shown in $\tilde{P}_m$, with $\tilde{P}_{11}$ having a very large amplitude — Figure 2.17.

2.4 Conclusions

The primordial power spectrum holds precious information about the physics of the early Universe and constraining it is one of the key goals of modern cosmology. However, the induced degeneracy between the cosmological parameters determining the matter/radiation transfer functions and the primordial power spectrum limit our ability to recover the primordial power spectrum, even from a perfect survey, especially in the case of CMB measurements [58]. Different surveys probe different scales with different accuracies and might not be able to constrain the primordial power spectrum to a desired resolution on their own. In combination, however, they make significant improvements. In this paper we have investigated these limits and improvements for CMB and large-scale structure surveys, exemplified by Planck and the SDSS BRG sample. For this purpose, we have assumed a non-parametric form for the primordial power spectrum and have constructed a parameter space containing a set of *carefully chosen* bins of the primordial power spectrum along with a set of cosmological parameters. We constructed a Fisher matrix for this parameter space for the two different surveys separately and combined. By diagonalising these Fisher matrices, via two different methods of eigenvector decomposition (PCA) and the Hermitian square root, we have investigated the induced correlation between the primordial power spectrum bins and the cosmological parameters.

In the PCA case, we conclude that SDSS and Planck together measure the cosmological parameters to a better extent, and even break the degeneracy between them. More importantly for our purposes, they can increase the obtainable resolution of the primordial power spectrum by a factor of two and can also condense the correlation between bins to be only amongst neighbouring ones. On the whole these experiments combined will constrain small scales better than large scales.

By the use of Hermitian square root of the Fisher matrix we managed to divert the correlation amongst the marginalised errors of the bins to the correlation between the bins themselves. In this case, the combination of SDSS and Planck helped to decrease the level of correlation between neighbouring bins, but also, because it has helped to increase the resolution of the bins, correlation between neighbouring bins means correlation between a smaller range of k.

Clearly combining the two surveys will constrain the primordial power spectrum better than current measurements, and better than each experiment on its own. Obviously, further surveys of other phenomena related to the evolution of fluctuations, such as Ly-α (e.g., SDSS LyαF power spectrum), weak lensing (e.g., Euclid),

peculiar velocities (e.g., Cluster Imaging Experiment (CIX)), etc. will help further measure the primordial power spectrum.
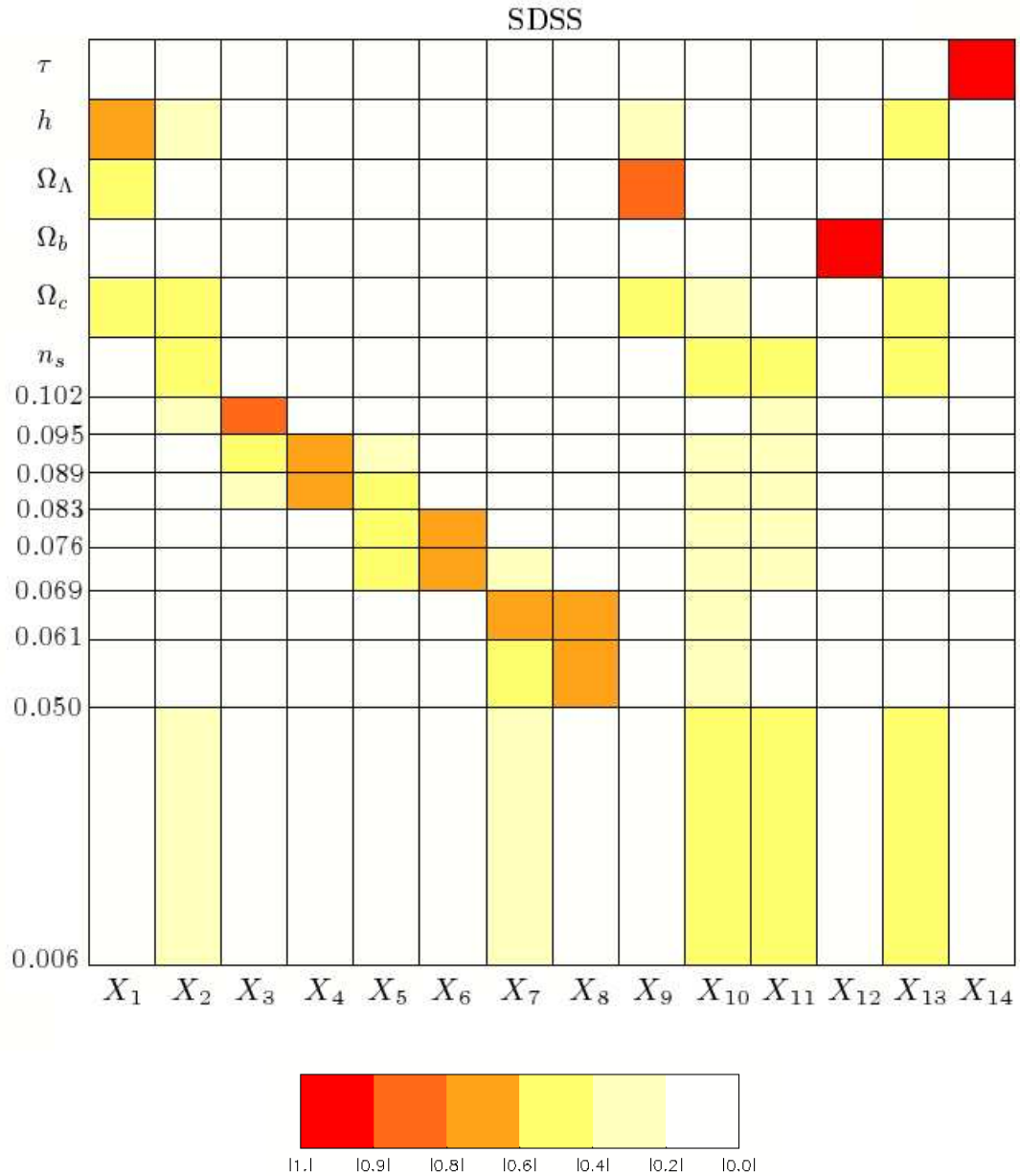
Figure 2.1: The principal components of SDSS with no priors on θ_i. X_i are shown from left to right with increasing errors $(= 1/\sqrt{\Lambda_i})$. Original parameters, O_i, are shown vertically starting with the bins on the bottom to θ_i on the top. For the bins, the vertical width of the box is an indication of the binwidth. The last 5 principal components can be ignored as they are not measured — refer to Table 2.1 and text for more details. At the bottom we show the colour plot indicating different levels of contribution to the principal components.
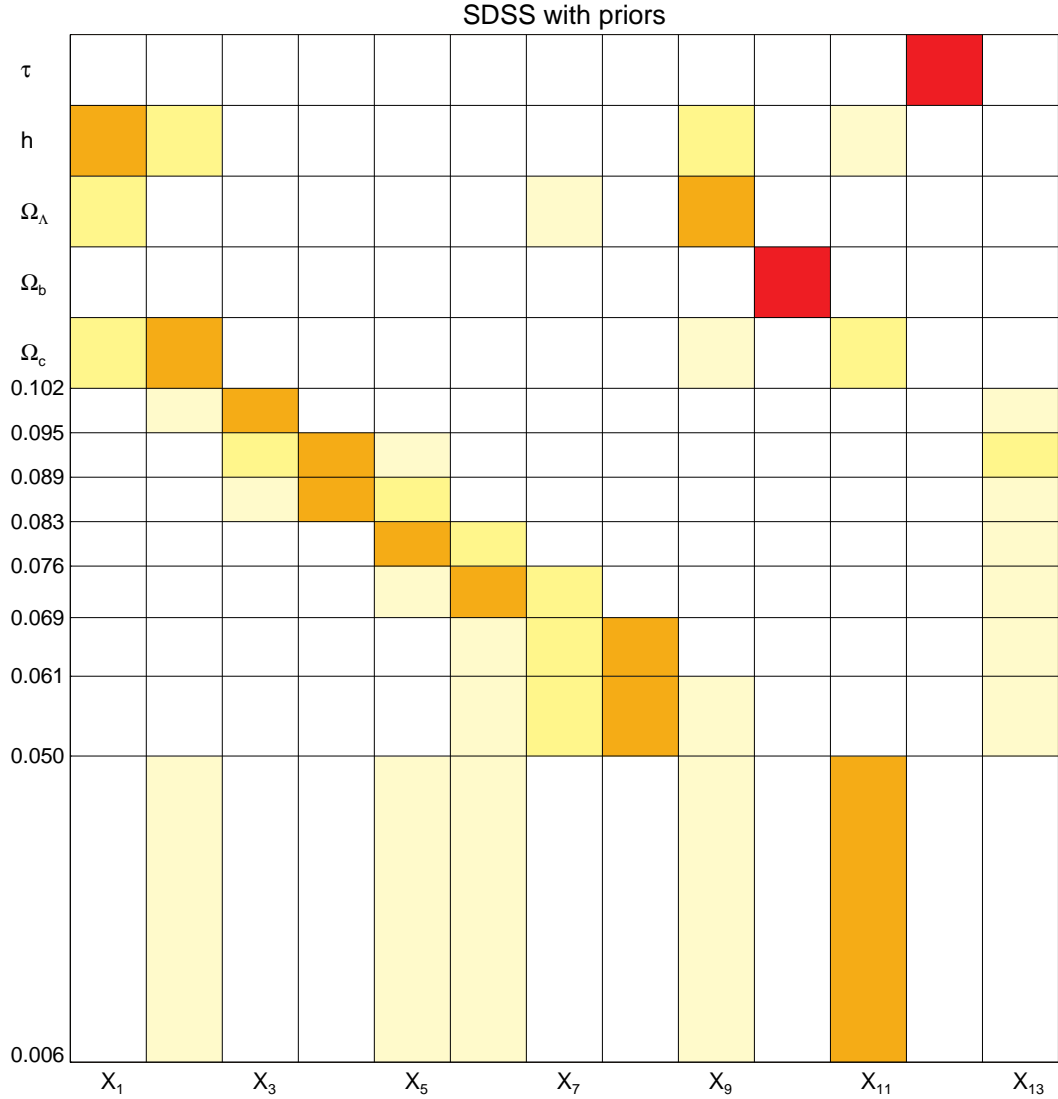
Figure 2.2: The principal components of SDSS with WMAP5 priors. The priors have helped to reduce the errors on X_i and now all of them can be measured well — Table 2.1. Some of the degeneracies between θ_i have also been broken; for example, Ω_b and τ dominate completely in X_{10} and X_{12} respectively, with no contribution from any other parameter. With respect to the primordial spectrum, these improved constraints on cosmological parameters have only helped to measure linear combinations of the bins better and have not been able to break the degeneracy between them.
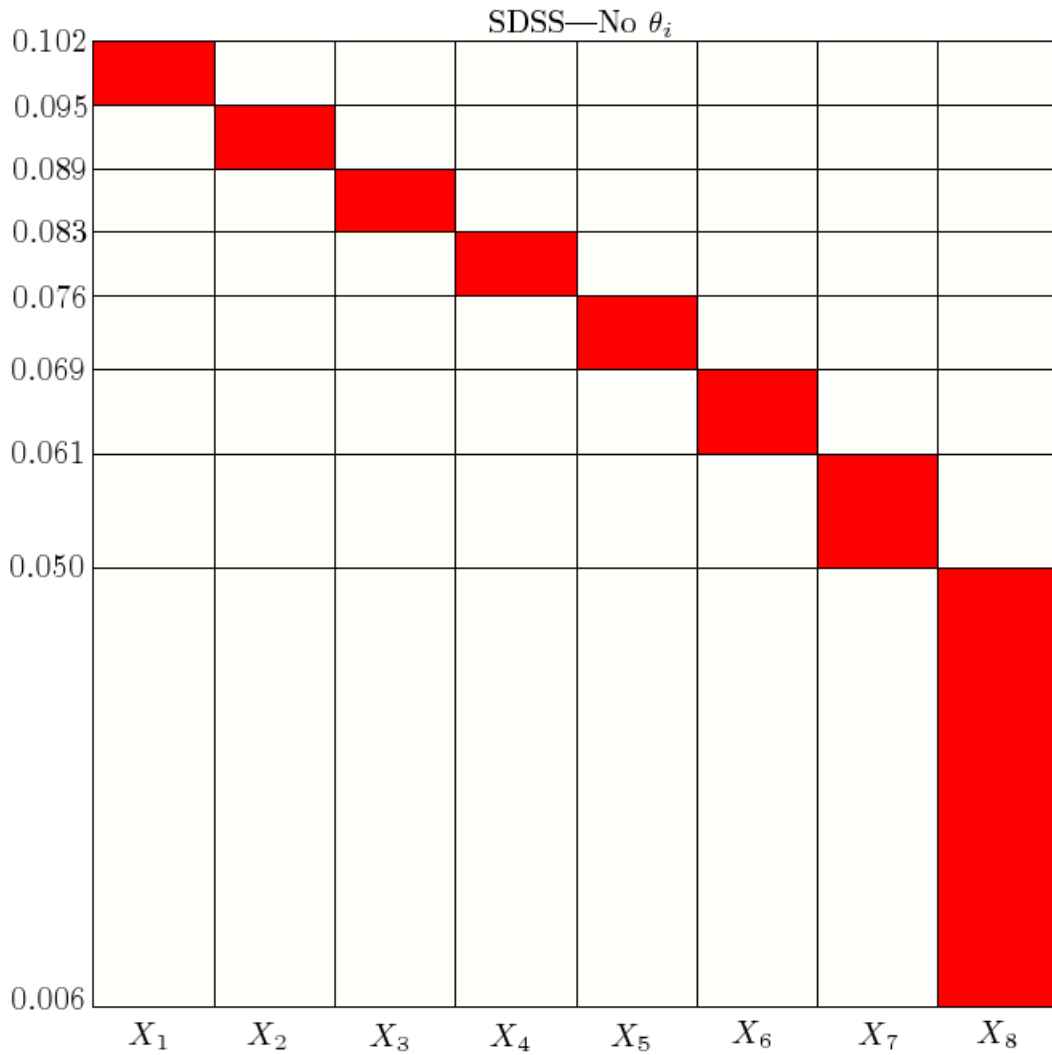
Figure 2.3: The principal components of SDSS for the primordial power spectrum bins only, assuming θ_i are known perfectly. No correlation exists between the bins at all. Compare to Figures 2.1 and 2.2, and see how the lack of knowledge of the cosmological parameters induce correlation between (neighboring) bins. This shows our limits in recovering the primordial power spectrum due to the uncertainties in the cosmological parameters.
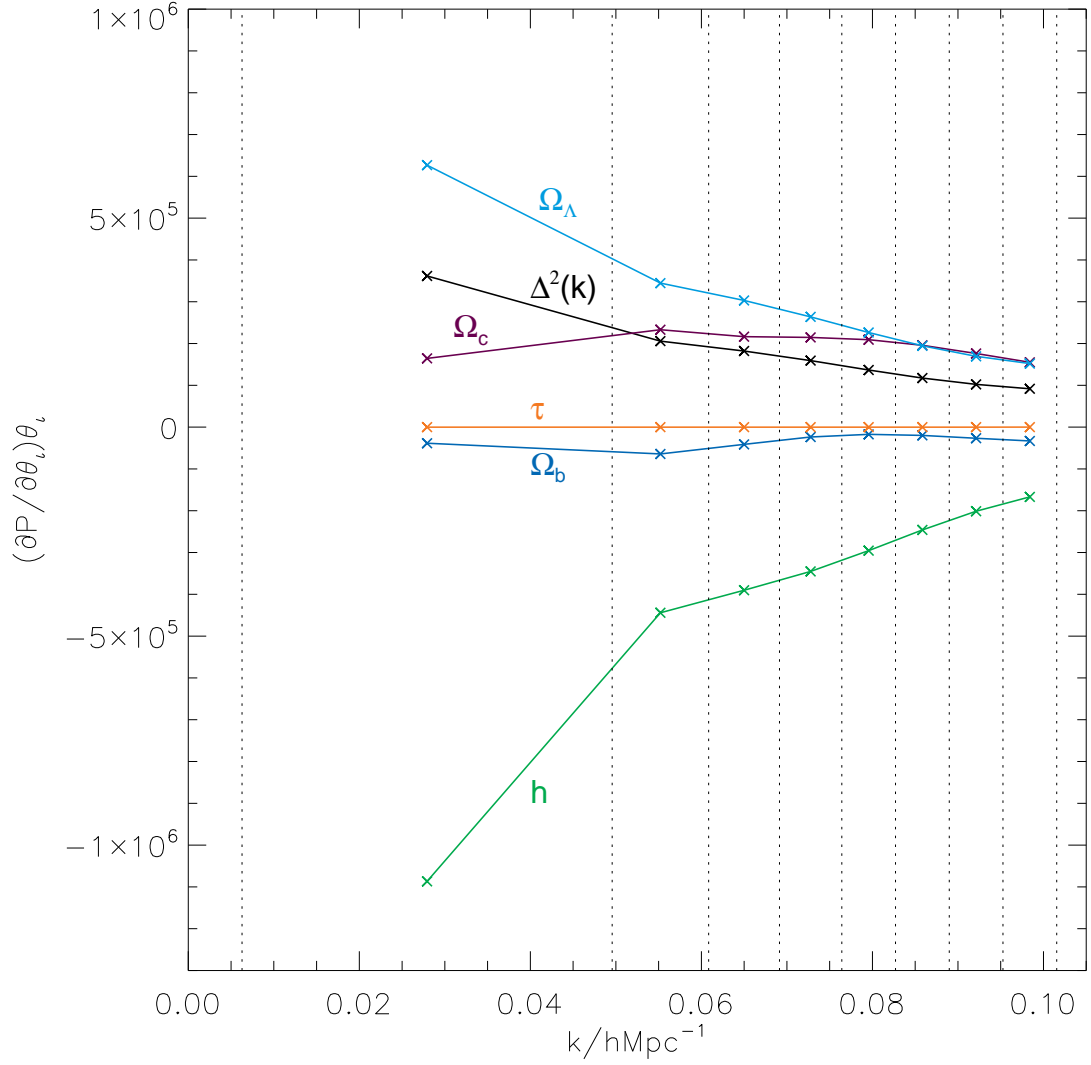
Figure 2.4: The derivative of galaxy power spectrum with respect to the primordial power spectrum bins and θ_i, weighted by the parameters values — this is exactly what goes in the Jacobian. It shows that, in the k-range we have chosen, much of the variation is on large scales and Hence there is no surprise that SDSS measures small scales better.
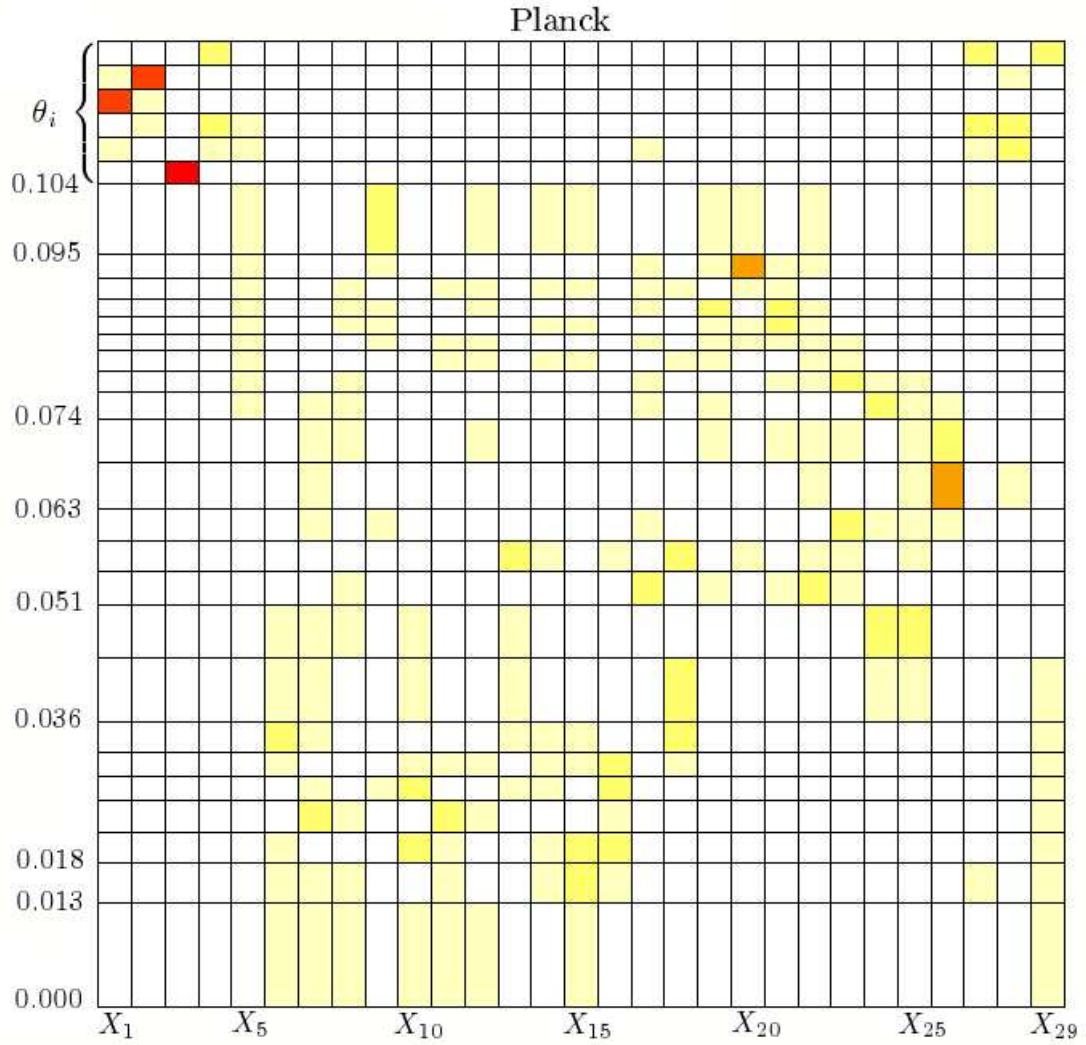
Figure 2.5: The principal components of Planck. Again bins are shown on the bottom and θ_i on the top, ordered in the same way shown in previous figures. Just like SDSS, Planck measures θ_i better than the primordial power spectrum and overall, gives smaller errors and smaller correlations between them. The rest of principal components contain the highly-correlated bins only, with no single large contribution from any of them.
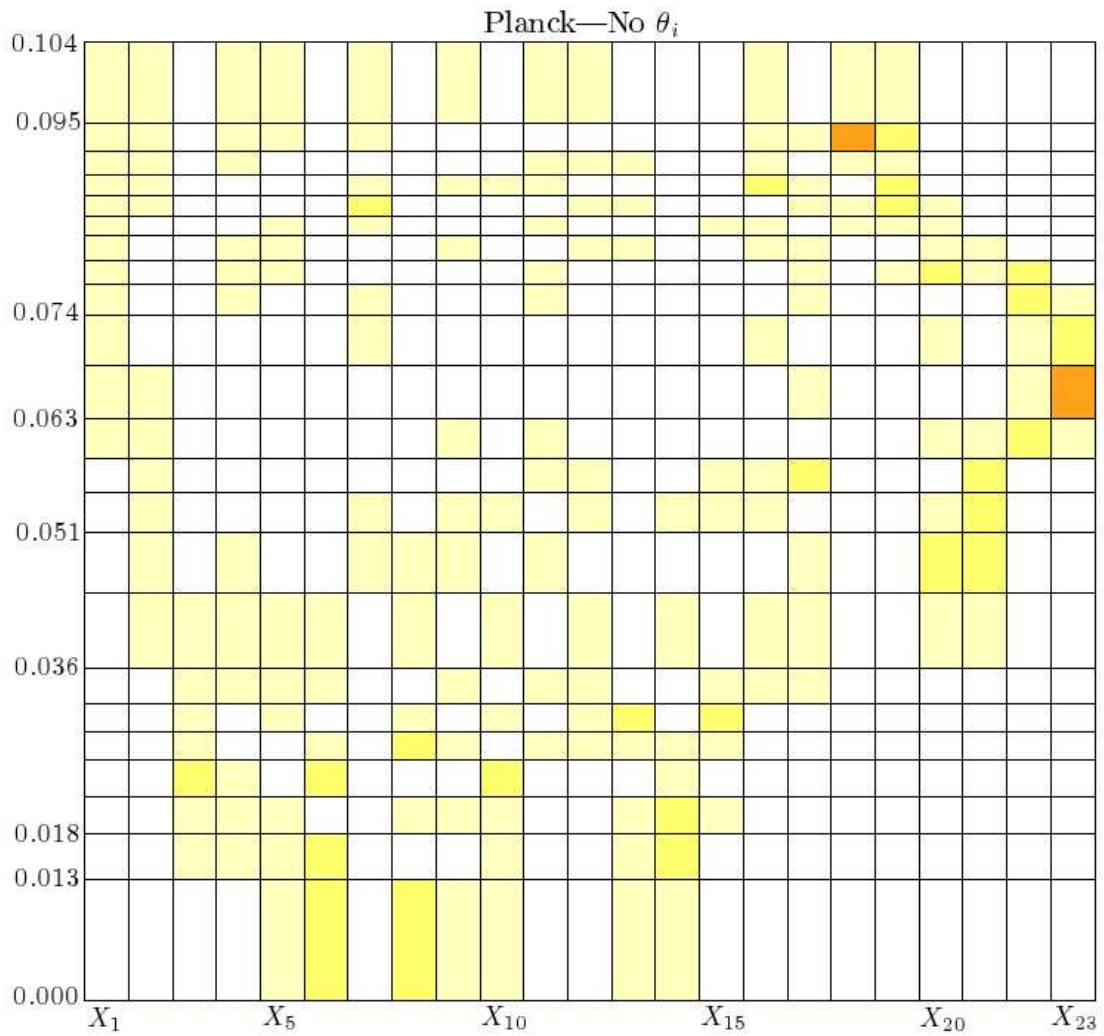
Figure 2.6: The principal components of Planck for the primordial power spectrum bins only, assuming θ_i are known perfectly. Compare to Figure 2.5. It seems like lack of knowledge of cosmological parameters does not have much of an effect in measuring the primordial power spectrum bins in this case. This is because Planck measures θ_i very well.

79
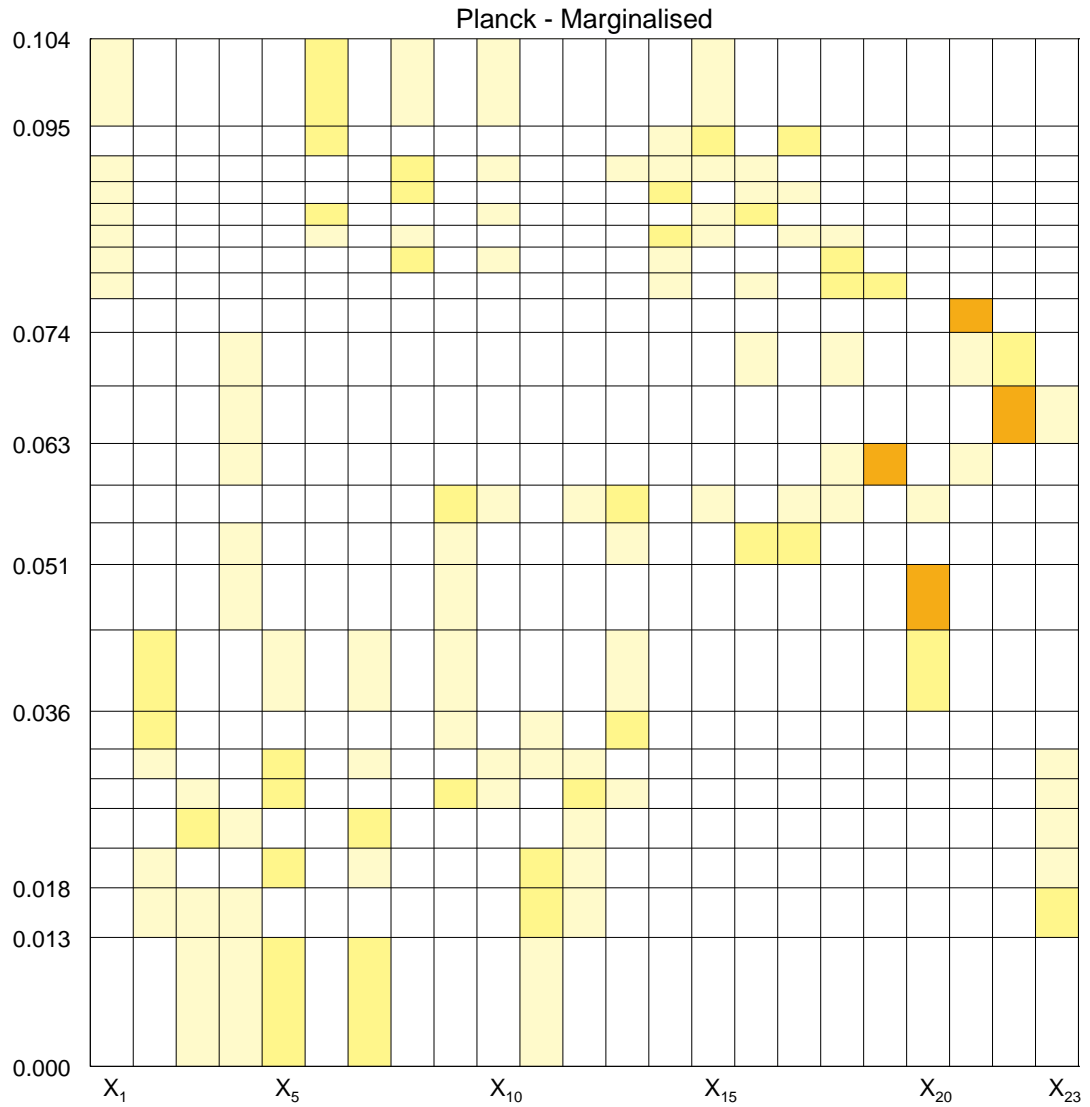
Figure 2.7: The principal components of the 'marginalised' Fisher matrix of Planck. Compare to Figure 2.5 where the non-marginalised case is shown.
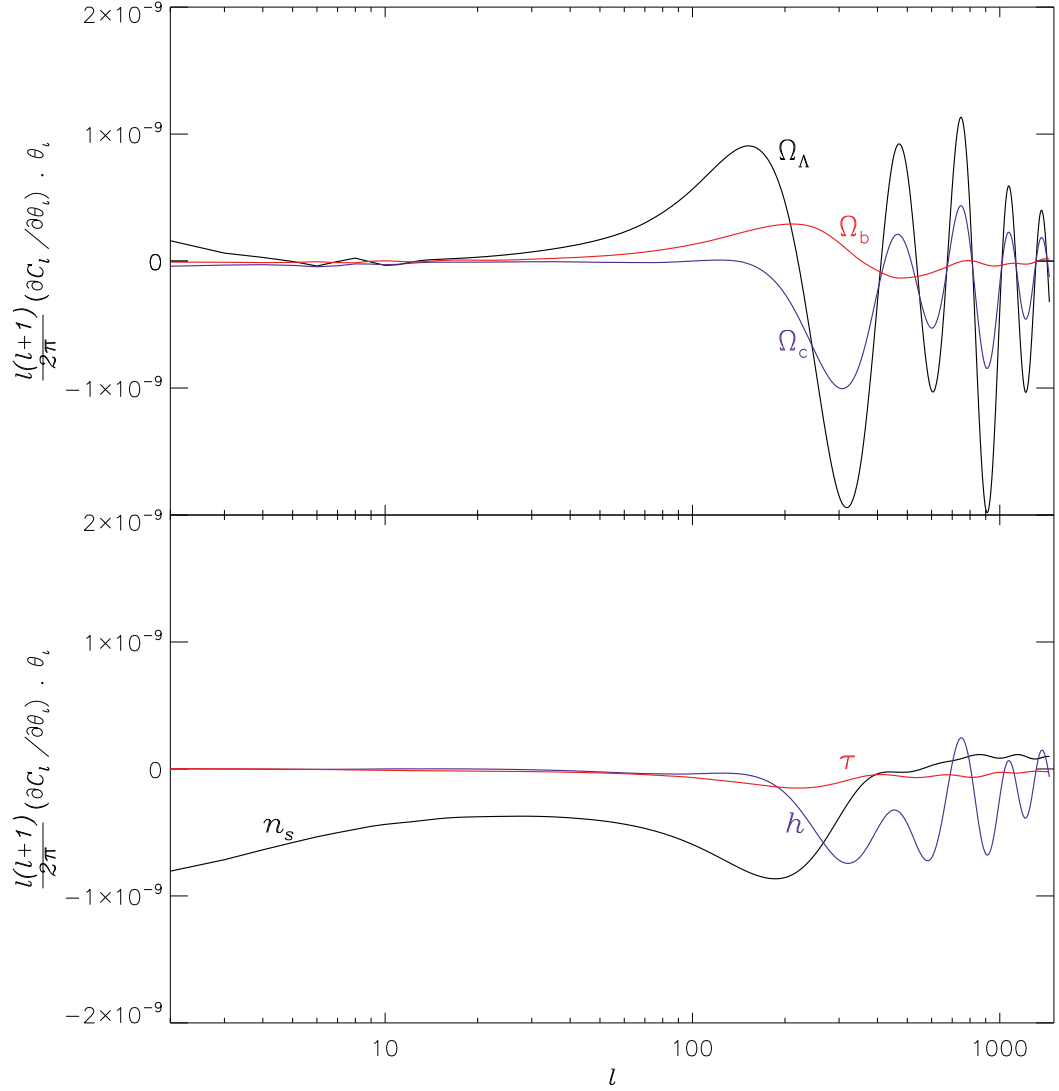
Figure 2.8: The derivative of radiation power spectrum with respect to θ_i. Note how different parameters dominate on different scales. For example, n_s dominates on large scales and Ω_b dominates on small scales. If an experiment can only measure a linear combination of n_s and Ω_b, the degeneracy between these parameters could induce a degeneracy between large and small scale bins!
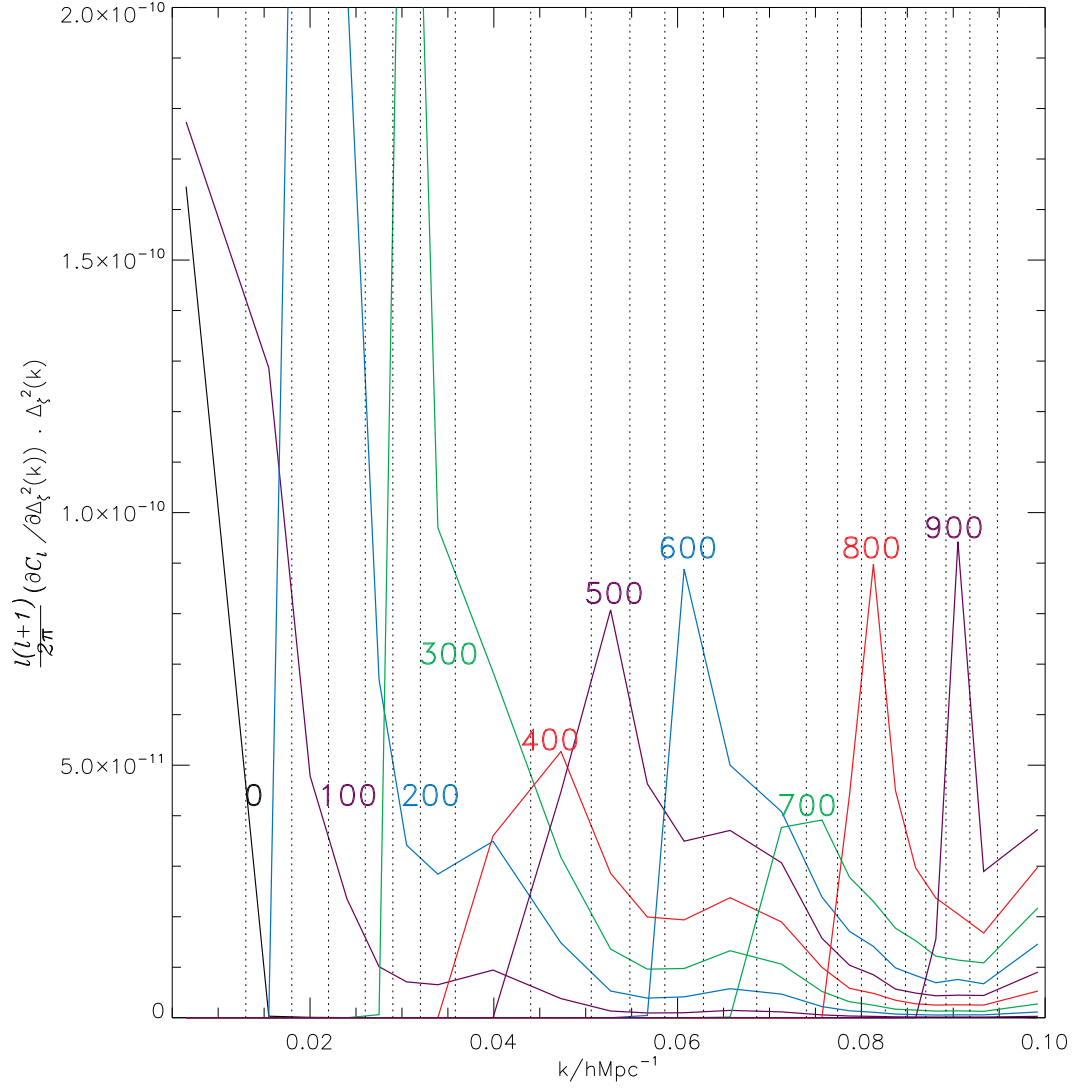
Figure 2.9: The derivative of radiation power spectrum with respect to the primordial power spectrum bins, Equation 2.12, weighted by the primordial power spectrum. Note that the bin with $\ell = 400$ dominating, gets contributions from all ℓs from 100 to 500. This makes the correlation between the bins on all scales possible.

Figure 2.10: The principal components of Planck & SDSS. Clearly, resolution of the primordial power spectrum has improved, now having 48 bins in the same k range. An almost diagonal trend can be seen now, showing small scales are measured better than the large scales. There is also less correlation between θ_i.

Figure 2.11: The principal components of the marginalised Fisher matrix of Planck & SDSS. Compare to Figure 2.10 where the non-marginalised case is shown.
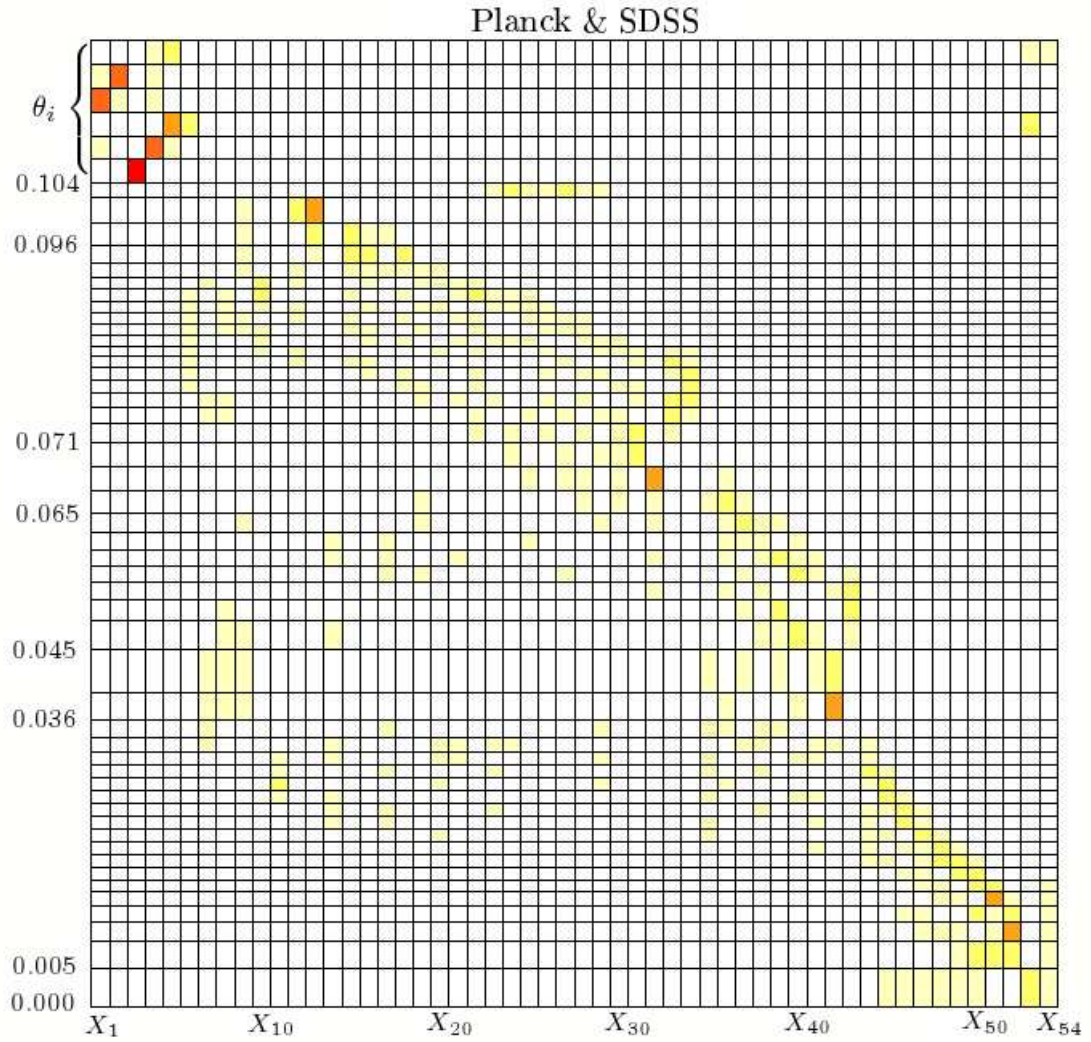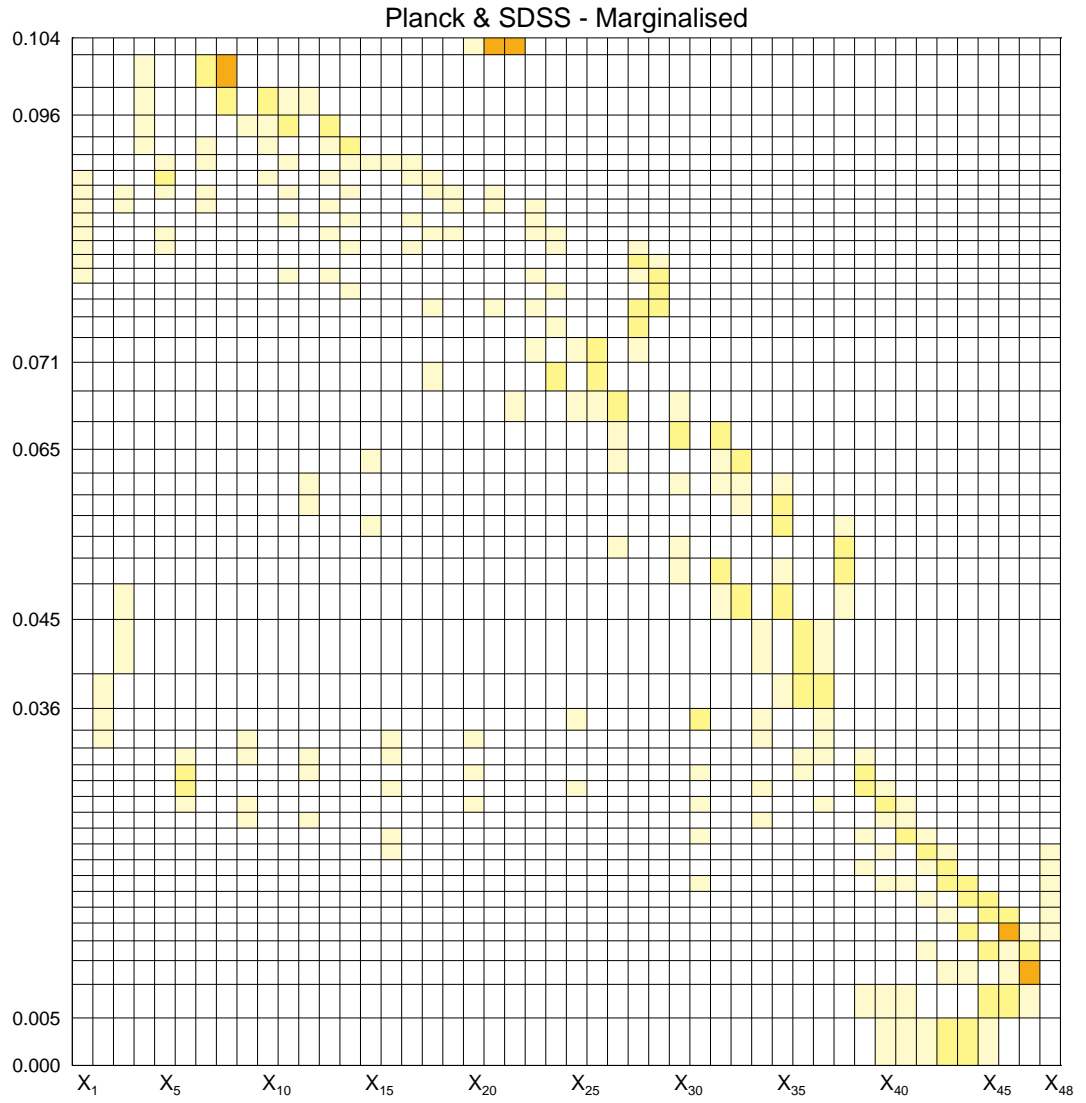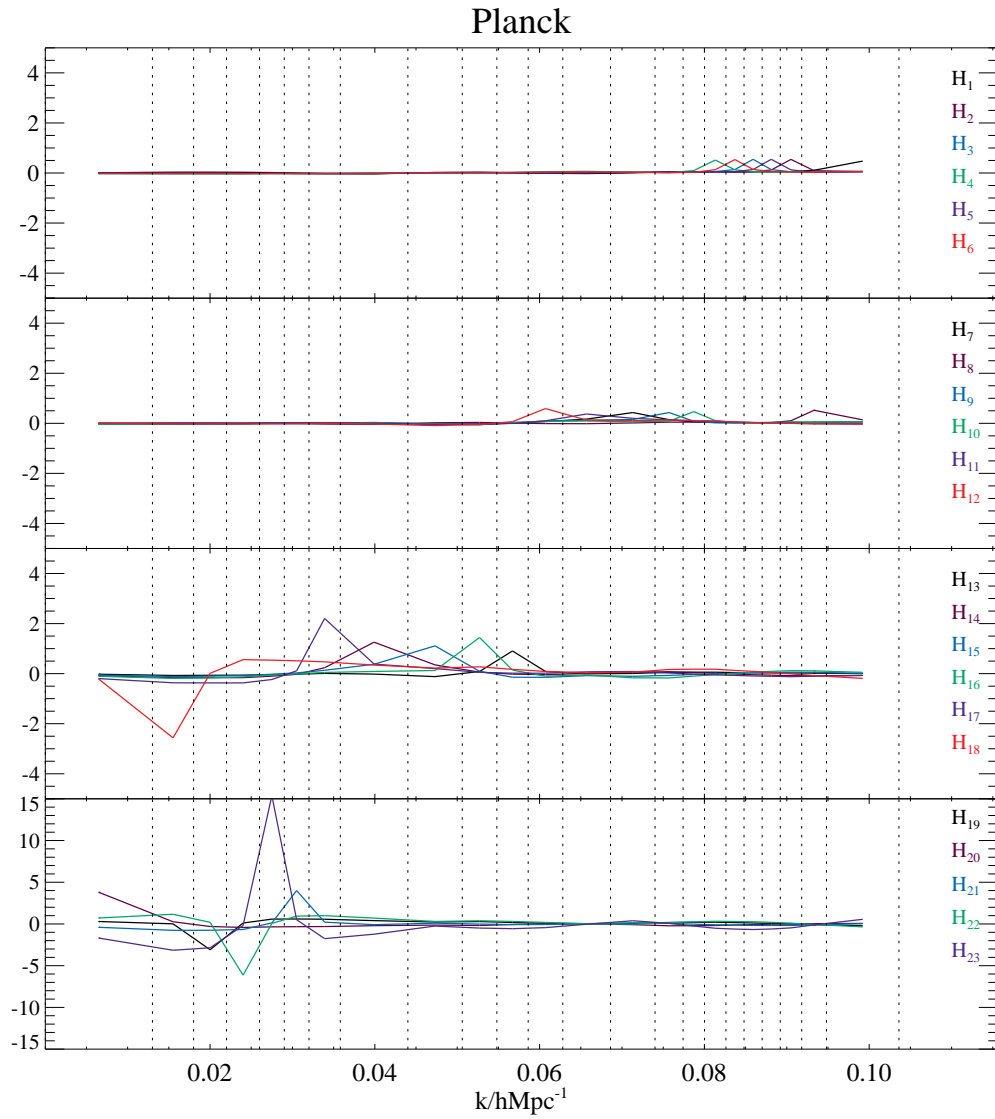
Figure 2.12: The row vectors of **H** for the marginalised Fisher matrix of Planck. These "window functions" are ordered with increasing errors, so that H_1 is the best and H_{23} is the worst measured vector. This shows that correlation is only between *neighboring* bins and, that bins on small scales are measured better than the ones on large scales (just like the PCA case).
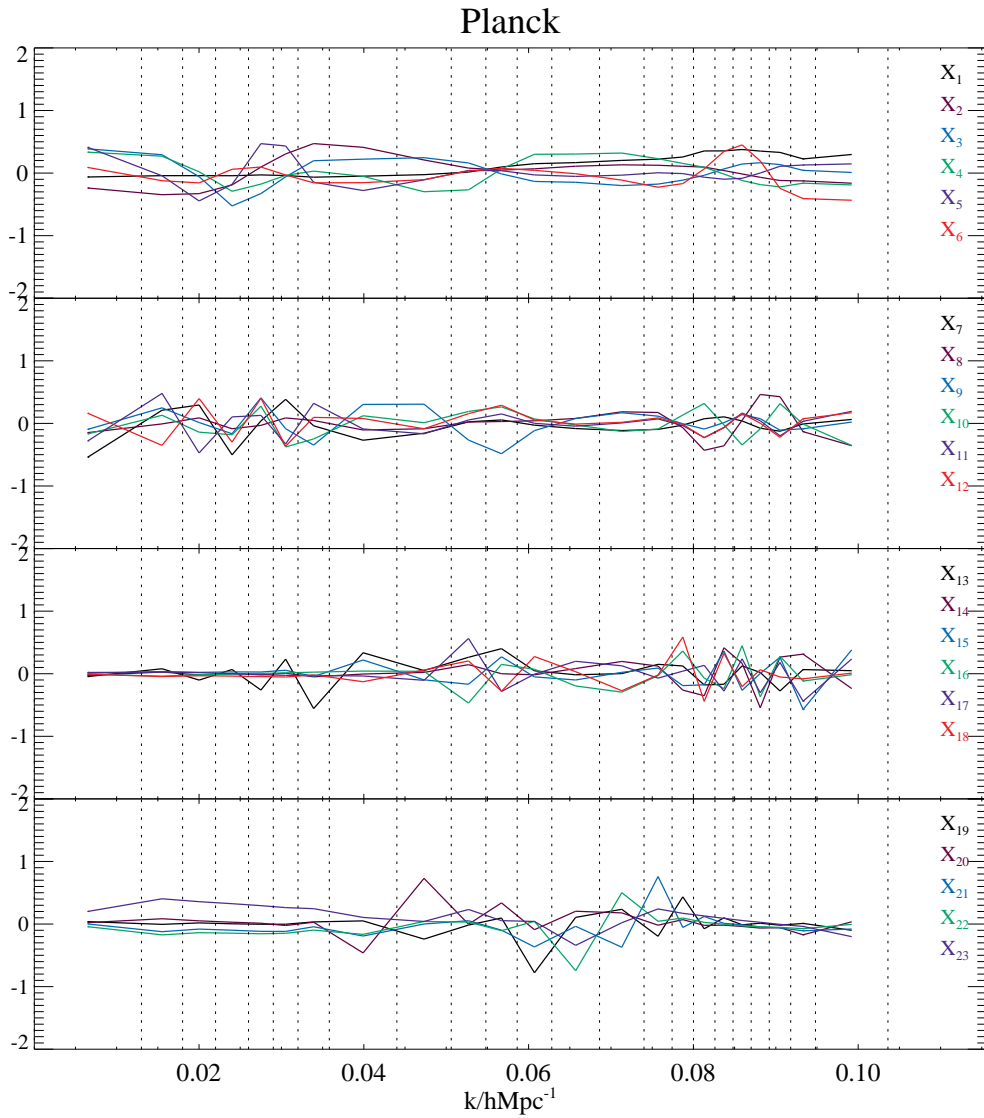
Figure 2.13: Showing Figure 2.7 — plotted in this way for easier comparison with Figure 2.12.
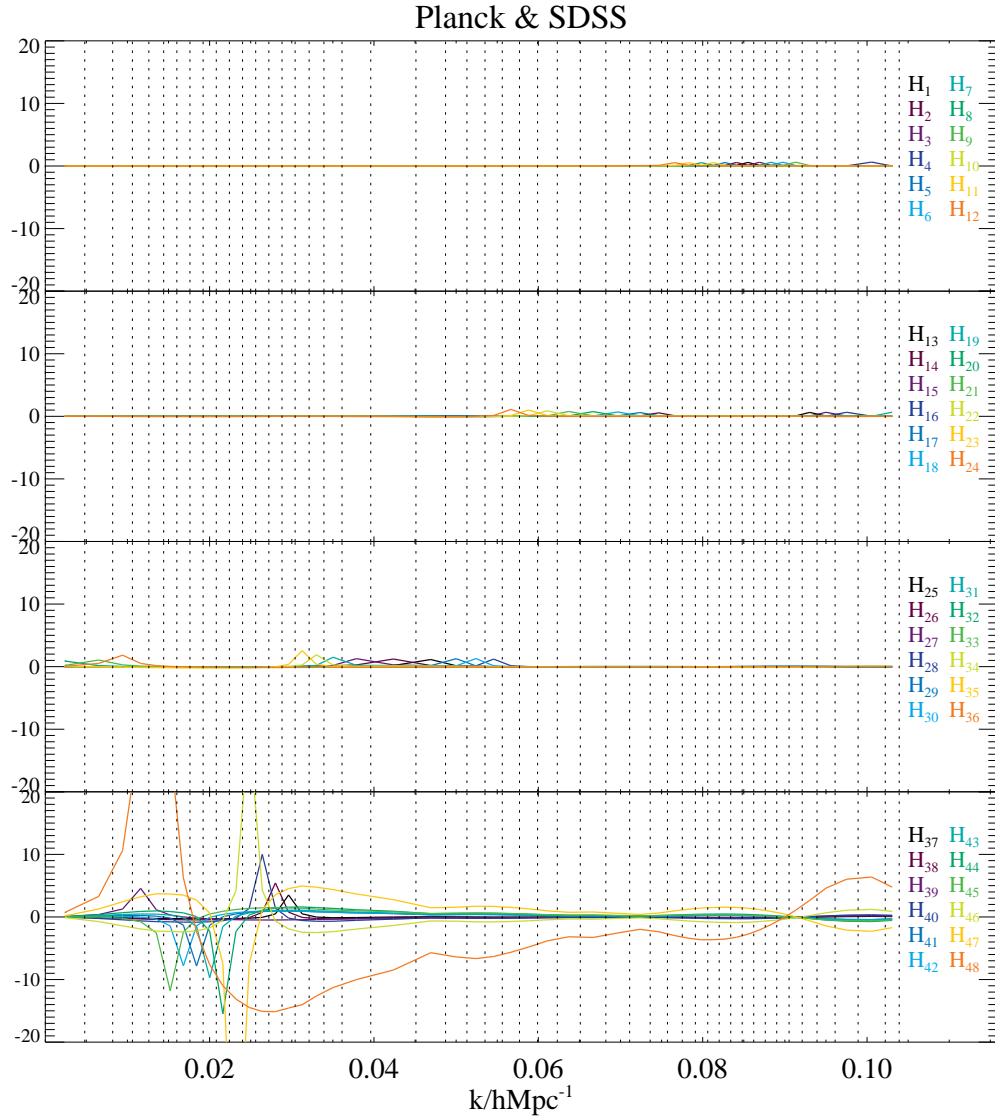
Figure 2.14: The row vectors of **H** ("window functions") for the marginalised Fisher matrix of Planck & SDSS. Like before, they are ordered with increasing errors. The correlation between neighboring bins still exists but to a lesser extent. Also, note that the bins are narrower here so that correlation between neighboring bins still means a correlation within a narrower k-range.
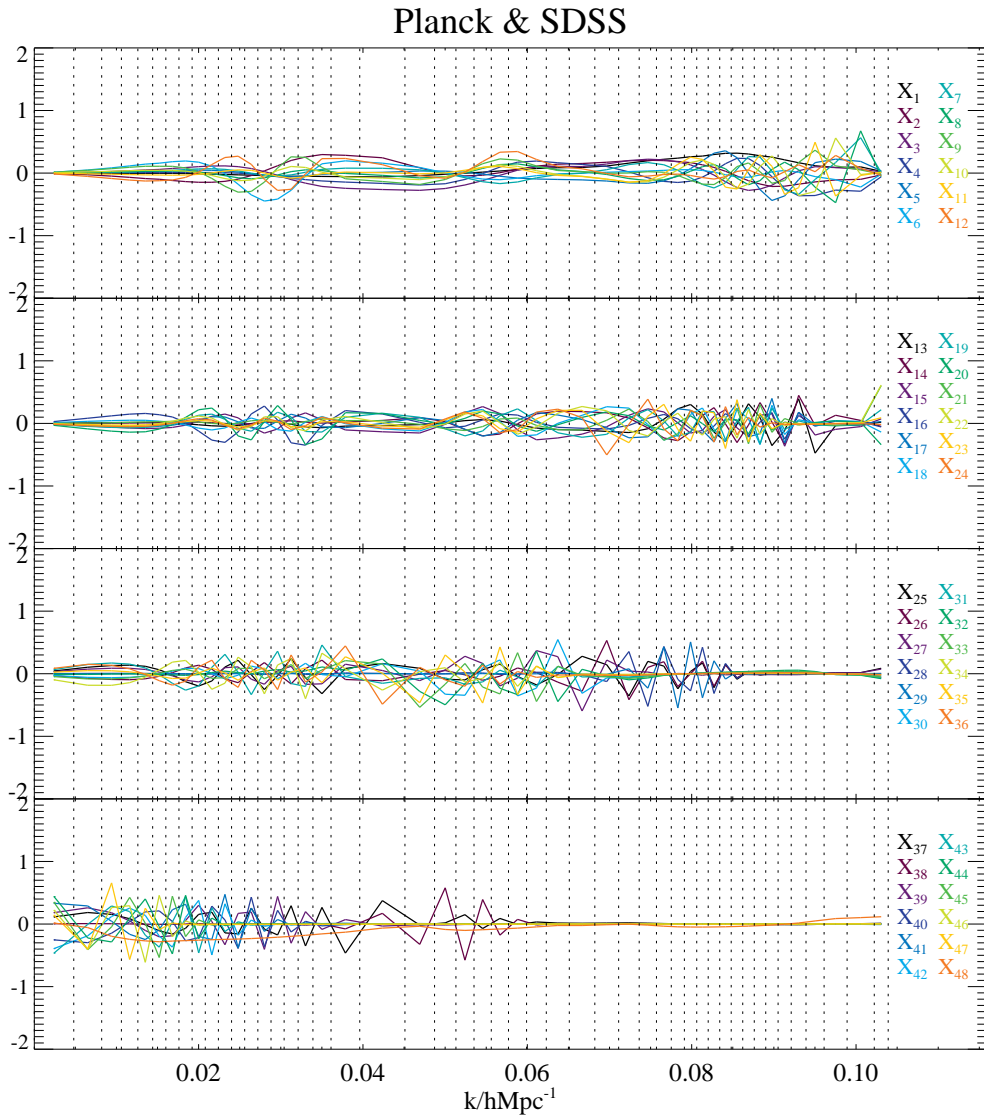
Figure 2.15: Showing Figure 2.11 — plotted in this way for easier comparison with Figure 2.14.
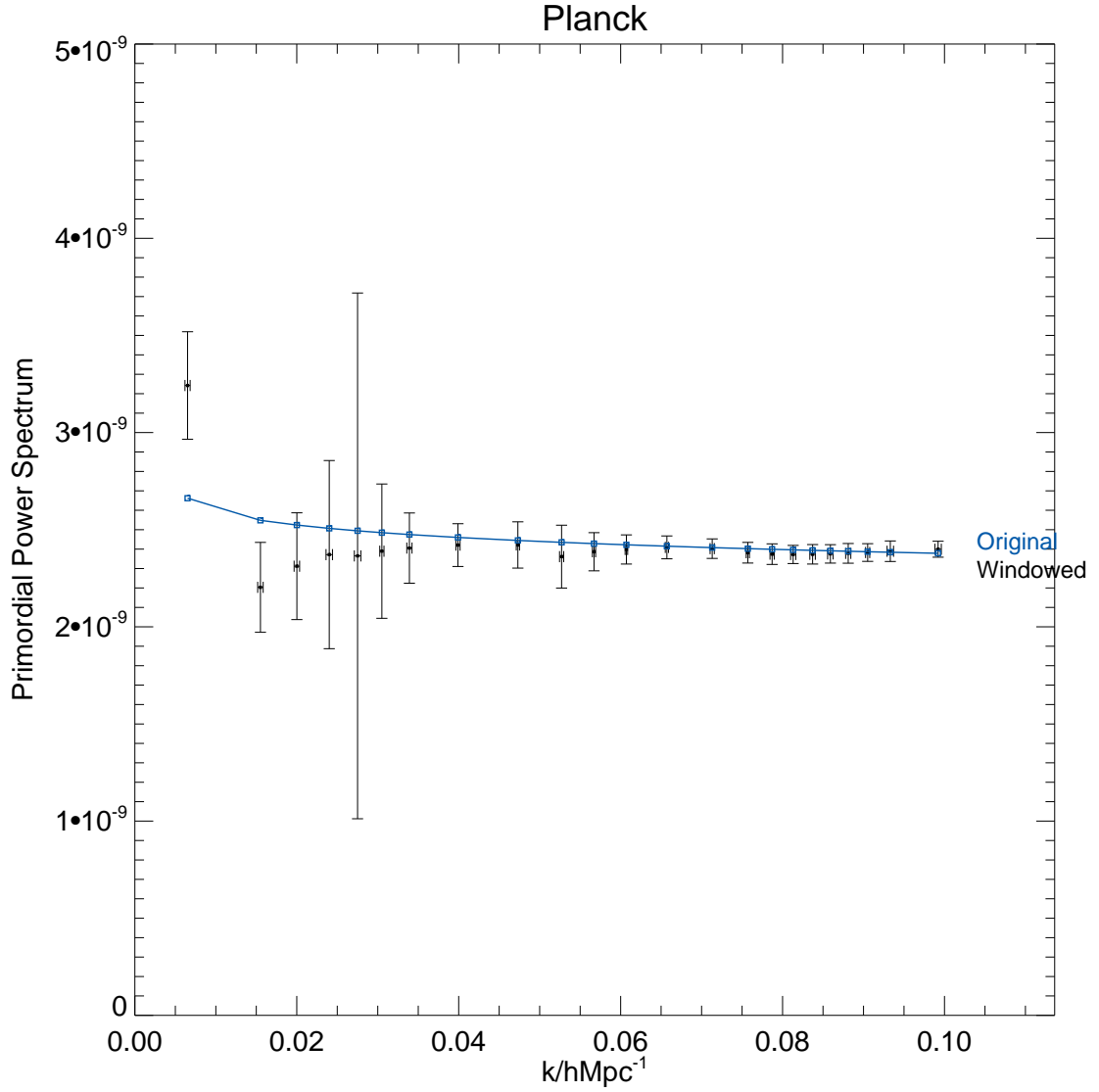
Figure 2.16: The windowed power spectrum obtained from Planck. The original primordial power spectrum is plotted for comparison. Remember that $\tilde{P}_m$ is not a physical power spectrum *per se*. However, the observed differences from the original power spectrum arise due to the induced correlations between the bins. In a perfect setting, where there are no correlations between bins, we expect $\tilde{P}_m = \Delta_\zeta^2(k)$. Note that the main feature of this plot is that vertical errors, unlike those for the original primordial power spectrum, are *not* correlated.
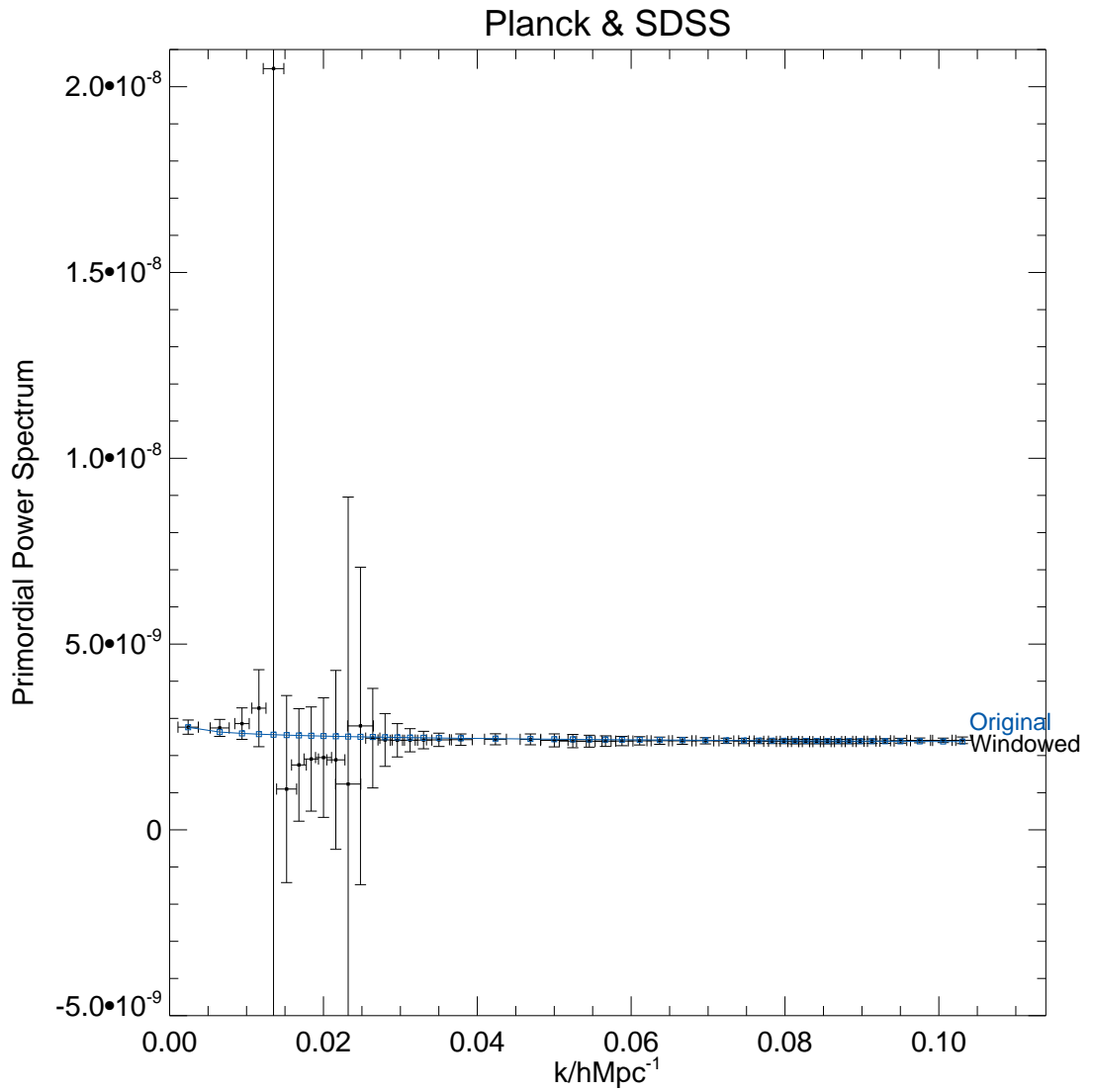
Figure 2.17: The windowed power spectrum obtained from Planck & SDSS. There is a better match between the two power spectrum on small scales compared to Planck on its own. However, the differences on large scales still remain.

Chapter 3

Reconstruction of the Primordial Power Spectrum by Direct Inversion

The work presented in this chapter was in collboration with Carlo Contaldi and Gavin Nicholson. My contribution to this work is presented in Section 3.3.

3.1   Introduction

As explained in the previous chapters, the primordial power spectrum of curvature perturbations $\Phi(\boldsymbol{k})$ is defined by,

$$P(k) = \Delta_\Phi^2 \equiv \frac{k^3}{2\pi^2} \delta^3(\boldsymbol{k} - \boldsymbol{k}') \langle \Phi(\boldsymbol{k})\Phi^*(\boldsymbol{k}') \rangle, \qquad (3.1)$$

where k is the wavenumber. We will call the primordial power spectrum $P(k)$ for simplicity. This spectrum represents the initial conditions set by some high-energy physical processes in the early universe. Most commonly it is thought that this process is inflation.

There are very few probes of the physics of the early universe. Other probes include non-Gaussianity, the primordial tensor power spectrum, a cosmic gravitational wave background and a cosmic neutrino background, none of which have been observed. Although there have been hints of a detection of non-Gaussianity [79], further studies disagree with this analysis [80].

On the other hand we can observe $P(k)$ through the windows of the Cosmic Microwave Background (CMB) and Large Scale Structure (LSS). It is therefore an incredibly important and powerful insight into the early universe. The simplest inflationary models predict that $P(k)$ takes the form of a power law $\propto k^{n_s-1}$ over the observable range in k, as opposed to the Harrison-Zeldovich spectrum parametrised by a single amplitude over all k. The spectral index n_s is predicted to be close to 1, current limits agree with this predication placing $n_s = 0.963 \pm 0.015$ [8].

Now that we have strong evidence suggesting inflation did indeed occur we would like to distinguish between different inflationary models. Many in fact propose alterations to the standard power law parametrisation. This includes those with features on the potential [81, 82, 83, 84, 85, 86, 87, 88], a small number of e-folds [89, 90, 91], or other exotic inflationary models [92, 93, 94, 95, 96].

There are two approaches to reconstructing $P(k)$, parametrisation and direct inversion. None of the various methods have shown conclusive evidence for a departure from near scale-invariance of $P(k)$. Despite this there have been tantalising hints of anomalous features in the CMB. One example of this is that the first year WMAP results gave an indication of a cut-off in $P(k)$ on large scales. With subsequent data releases the significance of this feature has been reduced, although future observations of the polarisation of the CMB may provide more conclusive evidence [91, 97]. In [98] we also showed evidence for a dip in power at $k \approx 0.002 \mathrm{Mpc}^{-1}$.

Numerous parametric searches for features with a similar form to those in complex inflationary models have been performed along with simple binning of $P(k)$ [99, 89, 100, 101, 102, 103, 104, 105, 106, 107, 108]. Methods of direct inversion which make no assumptions about the early universe model being probed [109, 110, 111, 112, 99, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 98] are hampered by the singular nature of the transfer function that takes $P(k)$ and transfers it onto the CMB or LSS. In general this causes the process of estimation to be prohibitively slow so as not to allow joint estimation of a free $P(k)$ with cosmological parameters. This hides any degeneracies between cosmological parameters and the form of $P(k)$. Thus it is not clear what the significance of any features found in the reconstructed $P(k)$ should be. It has been pointed out [124, 98, 125] that adding polarisation information or LSS data can help break any degeneracies.

We propose a new method for direct inversion of $P(k)$ using Singular Value Decomposition (SVD), which is fast enough to allow us to carry out a joint estimation of $P(k)$ and the cosmological parameters. The form of $P(k)$ is derived from the SVD

inversion. We also show that there are regions of k for which polarisation data has the potential to more accurately constrain $P(k)$.

3.2 Direct Inversion of P_k by Singular Value Decomposition

The theorem of SVD [126] states that for any m-by-n matrix, $\mathbf{\Sigma}$, with real or complex entries there exists a factorisation of the form

$$\mathbf{\Sigma} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*, \tag{3.2}$$

where $\mathbf{U}$ is an m-by-m unitary matrix, $\mathbf{\Lambda}$ is a m-by-n non-negative diagonal real matrix and $\mathbf{V}^*$ is the conjugate transpose of $\mathbf{V}$, an n-by-n unitary matrix. The diagonal entries of $\mathbf{\Lambda}$ are the singular values which can be loosely thought of as eigenvalues. The columns of $\mathbf{U}$ form a set of orthonormal 'output' basis vectors for $\mathbf{\Sigma}$, whilst the columns of $\mathbf{V}$ form a set of orthonormal 'input' basis vector directions for $\mathbf{\Sigma}$.

The factorisation is 'almost' unique in that making the same permutation of columns of $\mathbf{U}$, elements of $\mathbf{\Lambda}$ and columns of $\mathbf{V}$ will not change the result of the multiplication. Neither will forming linear combinations of any columns of $\mathbf{U}$ and $\mathbf{V}$ whose corresponding elements of $\mathbf{\Lambda}$ happen to be exactly equal.

In the case where there are fewer equations than unknowns, $m < n$, there is not a unique solution. It then becomes important to ensure that $n-m$ zero λ_is are found, where the λ_i are the diagonal entries of $\mathbf{\Lambda}$. If degeneracies exist in the equations to be solved then the number of λ_i that should be zero or negligible will correspond to the number of degeneracies. In practice this is normally achieved by setting the smallest λ_i to zero. However in the case where it is not clear how many degeneracies exist it is possible that information may be lost by doing this.

The properties of SVD allow an 'inverse' to be calculated for any $\mathbf{\Sigma}$. In the case of singular matrices the inverse is not well defined, there is in fact a solution space of 'pseudo-inverse' matrices given by all the permutations that maintain the validity of the factorisation. Where $m \neq n$ there is both a 'right-inverse' and 'left-inverse' of $\mathbf{\Sigma}$, although these are simply the conjugate transpose of each other. The 'inverses'

are given by

$$
\begin{aligned}
\mathbf{\Sigma}^{-1}_{\text{left}} &= \mathbf{U}^*\mathbf{\Lambda}^{-1}\mathbf{V} \text{ and} \\
\mathbf{\Sigma}^{-1}_{\text{right}} &= \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}^*,
\end{aligned}
\tag{3.3}
$$

$\mathbf{\Lambda}^{-1} = [\text{diag}(1/\lambda_i)]$, where any of the λ_i are 0 we replace $\text{diag}(1/\lambda_i)$ by 0.

Direct primordial power spectrum reconstruction requires the inversion of the following relations

$$
C_\ell^{XY} = \int\limits_0^\infty \frac{\mathrm{dk}}{k} \Delta_\ell^X(k)\Delta_\ell^Y(k)P(k),
\tag{3.4}
$$

where X and Y represent T, E, or B-type anisotropies, C_ℓ^{XY} are the angular power spectra for the XY combination and the $\Delta_\ell^X(k)$ are the photon perturbation transfer functions. The transfer functions are obtained by integrating the full Einstein-Boltzmann system of differential equations. These describe the evolution of perturbations in the photon distribution functions in the presence of gravity and other sources of stress-energy. The functions determine all of the structure in the anisotropy spectra which arises after the initial conditions are set. Most notably the C_ℓ^{XY} contain distinct peaks due to the acoustic oscillation of the tightly coupled photon-baryon fluid in gravitational potential wells at the time of last scattering. The aim of any inversion method is to distinguish such features from any structure in the initial perturbation spectrum.

For a finite sampling of the wavenumber space k Equation 3.4 can be recast as an operator acting on the primordial spectrum P_k

$$
C_\ell = \sum_k M_{\ell k}P_k,
\tag{3.5}
$$

with operator

$$
M_{\ell k}^{XY} = \Delta \ln k \, \Delta_{\ell k}^X \Delta_{\ell k}^Y,
\tag{3.6}
$$

where $\Delta \ln k$ are the logarithmic k intervals for the discrete sampling chosen in the integration of the system of equations.

A solution for P_k cannot be obtained from a traditional direct inversion of the $M_{\ell k}^{XY}$ as it is numerically singular. This is due to the high level of degeneracy in the transfer functions relating the power at any wavenumber k to angular multipoles ℓ. We instead approximate the inversion by using the SVD method, first reducing the degeneracy of the system and then inverting using the remaining orthogonal modes.

The transfer functions can be factorised using SVD,

$$M_{\ell k} = \sum_{\ell' k'} U_{\ell \ell'} \Lambda_{\ell' k'} V^{*}_{k' k} \,, \tag{3.7}$$

where the matrices $\mathbf{U}$ and $\mathbf{V}$ are unitary and of dimensions n_ℓ and n_k respectively, and $\mathbf{\Lambda}$ is a non-negative, diagonal matrix with elements λ_k. For this application, the dimensions of the matrices are that $n_\ell < n_k$ i.e. there are more equations than modes of interest. This results in some of the diagonal elements of $\mathbf{\Lambda}$ being singular (numerically zero) which prevent the inversion of the transfer matrix. The SVD method allows one to invert such a system by nulling the singular modes. This is achieved by creating an inverse $\mathbf{\Lambda^{-1}}$ where the diagonal elements are $1/\lambda_k$ except where the value of λ_k is singular in which case it is replaced by 0.

In practice we rank order the factorised modes in descending order of λ_k and all modes with condition number less than a threshold $\epsilon_{max}(\lambda_k)$ are nulled, $\epsilon \approx 0.038$ for this work. Thus the method is a 'k-to-ℓ' compression of the system where we keep the least degenerate modes connecting the 3d Fourier space to the 2d angular multipole space. This is not to be confused with a signal-to-noise compression of the data, which aims to select with respect to orthogonal modes of the covariance of the observations [37].

It is instructive to look at the first few orthogonal modes given by the columns of the $\mathbf{U}$ matrix. We plot the first six in Figure 3.1. These ℓ-space modes are the least degenerate (or best determined) in the mapping provided by the CMB physics. In other words, in the absence of sample and noise variance, these modes pick out the ℓ-range where observing the CMB will have the highest impact upon the reconstructed $P(k)$. Not surprisingly, the first few modes are peaked around the angular scales where the acoustic signal from each polarisation combination is maximised. As these are the best constrained vectors in ℓ space in the absence of all errors, they are not necessarily the basis vectors of P_k which are most accurately constrained — [124] show what they are for WMAP. We assume that this ordering of singular values is the optimal method for sorting the columns in $\mathbf{U}$ and $\mathbf{V}$.

Once $\mathbf{\Lambda^{-1}}$ has been computed the primordial power spectrum can be reconstructed by inverting a set of observed C_ℓ using

$$
\begin{aligned}
P_k &= \sum_{\ell' k' \ell} V_{kk'} \Lambda^{-1}_{k' \ell'} U^{\dagger}_{\ell' \ell} C_\ell \,, \\
&\approx \sum_{\ell} M^{-1}_{k\ell} C_\ell \,.
\end{aligned} \tag{3.8}
$$

Our choice of $\epsilon \approx 0.038$ is a conservative one with approximately 200 non-singular modes for a typical transfer matrix. Reconstructing the primordial power spectrum in this way means that we are not using any degenerate modes which carry no information in k space but can increase the scatter in the reconstructed spectrum. However we are still susceptible to the scatter in the observed C_ℓ since we have not used any noise weighting in this scheme.

In practice we start our inversion process with a guess input spectrum, parametrised by the usual form $A_s k^{n_s-1}$. From this and our fiducial cosmological model we obtain a C_ℓ^{model} spectrum. We use this to calculate the residual spectrum,

$$C_\ell^{\text{res}} = C_\ell^{\text{obs}} - C_\ell^{\text{model}}, \tag{3.9}$$

so as to minimise the error induced in k-space by the cut-offs in ℓ, both on large and small scales.

To remove high frequency oscillations in the data we apply a low-pass filter to the resultant P_k. The following algorithm was used,

$$P_k^{\text{low-pass}} = \alpha P_k + (1 - \alpha) P_{k-1}, \tag{3.10}$$

where α was taken to be 0.05. This method of smoothing leaves the first few points in the series strongly influenced by P_1. Therefore one should take any effects seen at low k with a pinch of salt as these points are highly correlated. The covariance matrix was altered appropriately by a lower-triangular matrix representing this filter.

3.3   Optimal Binning

We then proceed to bin the reconstructed power spectrum using the optimal binning method of [127]. This binning method estimates a series of ranges in k over which the signal-to-noise in the measured primordial power spectrum is constant. Many of the data points are highly correlated with their nearest neighbours and optimal binning gives a clear indication of the scales on which we have independent information. The centre of the k bins chosen by the cosmological tool CAMB [128] change when the input cosmological parameters are altered. This is a problem when it is run over many realizations as in the case of a Markov Chain. We choose the optimal binning method to find the standardised output of centres and sizes of k bins for each call of the CAMB routine.

To find the optimal binning of the reconstructed power spectrum we investigate how the uncertainty in the C_ℓs transfers into uncertainty in the primordial power spectrum. For this purpose, we need to define the primordial power spectrum as a series of top-hat bins:

$$P(k) = \sum_B w_B(k)Q_B \,, \qquad (3.11)$$

where Q_B is the power in each bin B and $w_B = 1$ if $k \in B$ and 0 otherwise. To obtain the errors for these bins we define the Fisher matrix for the C_ℓs by

$$F_{\ell\ell'} = (\delta C_{\ell\ell'})^{-1} \,, \qquad (3.12)$$

where $\delta C_{\ell\ell'}$ is the diagonal matrix of the squares of the variances in each measurement of C_ℓ. We assume symmetrical errors and ignore the correlation between multipoles. To transfer the given errors from the C_ℓs to other parameters we need to use the Jacobian,

$$F_{\alpha\beta} = \sum_{\ell\ell'} F_{\ell\ell'} \frac{\partial C_\ell}{\partial \theta_\alpha} \frac{\partial C_{\ell'}}{\partial \theta_\beta} \,, \qquad (3.13)$$

where here θ_α and θ_β represent the bins of the primordial power spectrum. The derivative of the C_ℓs with respect to the primordial power spectrum is the average radiation transfer function in each bin:

$$\frac{\partial C_\ell}{\partial P(k)} = \int_{k_{min}^B}^{k_{max}^B} \frac{dk}{k} \Delta_\ell^X(k) \Delta_\ell^Y(k) \,. \qquad (3.14)$$

To calculate the signal-to-noise ratio in each bin we take the inverse square root of the diagonal elements of $F_{\alpha\beta}$ to be the noise and the amplitude of the primordial power spectrum to be the signal. We will then arrange the bins to have the same S/N over our k range: as the reconstructed spectrum seems to be featureless, we do not favour a particular k range and hence, will treat the whole range equally. There are cases in which we might want to place the bins by hand if we are looking for specific features in a particular k range (but of course we will always be limited by the S/N of our measurements). Our algorithm will result in more bins where the signal-to-noise ratio is greater, sampling more finely where the signal is strongest.

We construct a diagonal signal matrix, $\mathbf{S}$, which contains the amplitude of the primordial power spectrum for all the bins as its diagonal elements and with which

we weight our Fisher matrix as

$$\left(\frac{S}{N}\right)^2_{\alpha\beta} = S_\alpha F_{\alpha\beta} S_\beta \,. \qquad (3.15)$$

The square root of the diagonal elements of this matrix are the S/N of the bins. We start our algorithm with the maximum number of bins possible in our k range. This is set by the usual properties of the Fourier transform. These imply that the scale of the survey not only determines k_{min}, but also gives a lower bound upon the resolution, Δk_{min}: narrower bins would become highly correlated. Therefore, we set up a series of bins with the properties

$$k_{min} = \frac{\ell_{min}}{d_A} = \frac{2}{d_A} \;\; \text{and} \;\; (\Delta k)_{min} = \frac{\Delta\ell}{d_A} = \frac{1}{d_A} \,, \qquad (3.16)$$

where $d_A = 14.12\text{Gpc}$ (value given by WMAP5) is the angular diameter distance to the surface of last scattering. We set $k_{max} = 0.08\text{Mpc}^{-1}$ as the reconstruction process is limited past this regime due to the cut-off in ℓ at 1000. A Fisher matrix is then constructed for this set of bins and weighted by the signal matrix. We choose the maximum S/N in this set to be the target S/N. We iterate until the bins with smaller S/N are increased until they reach within 5% of the target value at each bin.

In the case where the intrinsic error in estimating the inverse operator, $M_{\ell k}$, is accounted for, the above procedure could be repeated while a set of cosmological parameters (the ones for which their change on the inverse operator was investigated for) form part of out parameter space, along with the primordial power spectrum bins. A Fisher matrix is constructed for this parameter space, where the same arguments as above are applied to get the initial binning. We then invert this Fisher matrix to get a covariance matrix, take the sub-block of this matrix that refers to the bins only and invert it back to get a *marginalised* Fisher matrix (the errors for the bins are marginalised over the cosmological parameters — assuming Gaussian errors). We then start an iteration to obtain a set of bins with equal S/N. Therefore, the cosmological parameters participate in choosing the bins. However, the reconstruction method adopted in this work induces errors and correlations that are so much larger than the ones induced due to the inclusion of parameters in the parameters space. Hence this method was ignored.

Ideally we would need to perform PCA (as in Chapter 2) on this set of bins to analyse the correlation between bins and combine the ones with strong correlations.

However, the correlations induced between the bins when the primordial power spectrum is reconstructed using our method is so much stronger than what is picked up by our binning method and hence this procedure was not pursued.

3.4 Reconstruction with cosmological parameter fitting

The reconstruction method described above is fast and can be carried out at each random sample of a MCMC exploration of the cosmological parameter space. Inserting the reconstruction as part of a MCMC exploration we can account for the variance induced in the primordial power spectrum due to the dependence of the radiation transfer function on the cosmology.

We have modified the `cosmomc` [129] package by introducing the reconstruction at each chain evaluation using the inverse of the transfer function computed for each combination of parameters acting on the 'observed' CMB angular power spectrum. The reconstructed spectrum is then itself used to compute the final C_ℓ which are used to calculate the likelihood at the chain step.

In principle the chains would probe the reduced set of parameters; the physical densities of baryons $\Omega_b h^2$, and of cold dark matter $\Omega_c h^2$, the angular diameter distance parameter θ, and optical depth parameter τ. The primordial power spectrum parameters n_s and A_s become irrelevant and need not be probed since the power spectrum is being reconstructed directly. However, in practice, we *do* include power law spectral parameters which determined the shape of the template model and then marginalise over the spectral parameters in order to account for any sensitivity of the reconstruction to the assumed C_ℓ^{model}.

The immediate advantage of combining the reconstruction with an MCMC method is that we can then calculate the variance in the resulting spectrum due to the random nature of the transfer function. We do this by including the binned amplitudes for the reconstructed spectrum as 'derived' parameters when analysing the chains. The covariance of the chains is then mapped into a covariance for the binned power spectrum.

We also need to account for the variance due to the errors in the observed CMB data. This is not accounted for in the MCMC chains since we always use the same observed C_ℓ data to reconstruct the spectrum. In principle this contribution to the

variance and that from the transfer function are correlated, however this is difficult to quantify without including MCMC steps over realisations of the observations. We therefore make a conservative estimate of the final error in the reconstructed spectrum by adding the covariance matrix obtained from the MCMC chain and that obtained by rotating the error matrix of the observed C_ℓ as

$$\delta P_{BB'} = \sum_{\ell\ell'} M_{B\ell}^{-1} \delta C_{\ell\ell'} M_{B'\ell'}^{-1} \qquad (3.17)$$

where $M_{B\ell}^{-1}$ is the bin-averaged contribution from $M_{k\ell}^{-1}$.

3.5 Application of the SVD Inversion

3.5.1 Tests on simulated CMB data

We start by testing the reconstruction algorithm on a set of input spectra with known features. We used a fiducial cosmological model of $\Omega_b h^2 = 0.0226$, $\Omega_c h^2 = 0.108$, $\theta = 1.041$ and $\tau = 0.076$. The starting guess spectrum used to obtain C_ℓ model is shown by the black dot-dashed line in Figure 3.2. We assume there are no errors on the input C_ℓ^{obs}s (between $\ell = 2$ and $\ell = 1000$) and observe that the reconstructed P_k generally picks out the input features. We limit our method to fit between $k = 0.0013$ and $k = 0.08$ as this is the range with the highest signal-to-noise in the WMAP data. The features chosen are the same as in [98] to allow comparisons with this method. They are a standard power law with $n_s = 0.96$ but with an amplitude 90% of that of the best-fit WMAP model, a running n_s model with $dn_s/d\ln k = -0.037$, a power law with a sharp, compensated feature at $k = 0.02\,\mathrm{Mpc}^{-1}$ [81, 92, 82, 93, 94, 85, 86, 95, 88] and a power law with superimposed sinusoidal oscillations [83, 130, 131, 87, 96]. All four features are clearly recovered to varying degrees when using TT, TE or EE, however we find a phase offset between the reconstructed and the input spectrum for the case where the input spectrum includes oscillations as in the lower panels of Figure 3.2. The offset is stable with respect to the presence of the smoothing kernel, the number of singular values cut from the inversion, and with respect to the number of k bins and range. The reconstructed P_k given by both TE and EE contains 'glitches' not present in the TT reconstruction. These regions correspond to regions where there is little information in the TE and EE spectra and the reconstruction is still degenerate.

The cosmological parameters obtained from a traditional MCMC search are used to give us our fiducial operator, $M_{\ell k}$. The parameters however have errors present upon them, which are not usually incorporated into errors on the final reconstructed P_k. If the wrong parameters have been used to calculate the fiducial operator very specific signatures would be expected to show up in the reconstructed form of P_k. We show in Figure 3.3 how these signatures appear in TT for four cosmological parameters, $\Omega_b h^2$, $\Omega_c h^2$, h and τ. If any features with this form are observed one should attribute this to an incorrect estimation of the parameters and not some fundamental physics. A check for this would be if the corresponding features also show up in the reconstructed TE and EE spectra.

We have also tested the inversion on simulated CMB data with similar experimental properties as the recently launched Planck satellite mission [9]. We assume a total of 12 detectors with NETs (Noise-Equivalent Temperature) of $64\mu K/\sqrt{s}$ observing 80% of the sky over 12 months with a resolution of 7 arcminutes FWHM. We calculate errors around our fiducial CMB best-fit models in both total intensity and polarisation spectra for this experimental setup and use these together with C_ℓ samples on the fiducial model to test the inversion method's properties. We consider multipoles of $\ell < 1000$ for both total intensity spectra and polarisation. We have not taken into account any residual error from foreground subtraction in our forecasts. Thus our forecast are on the optimistic side of the accuracy achievable in the case of polarisation where foreground removal will certainly have a significant impact on errors at $\ell < 1000$. In the case of total intensity spectra we are significantly underestimating the accuracy achievable by Planck as we expect to obtain well measured C_ℓs well above ℓ of 1000. We do not consider these modes as it significantly increases the time required to perform the SVD. Consideration of total-intensity modes past ℓ of 1000 will increase the accuracy obtainable for Planck on P_k, it also allows one to probe k past 0.08. It would be desirable to perform this process with a greater ℓ range when the Planck data becomes available.

To compare the accuracy obtainable with each of the anisotropy types we show the degree to which they each reconstruct a simple power law P_k in Figure 3.4. We use the same fiducial cosmological model as in Figure 3.2 with a standard power law input P_k equal to the spectra being reconstructed. All the C_ℓ^{obs} were placed on the fiducial model. Over the whole range of k total-intensity modes best constrain P_k. But it is also true that there are regions in k-space, for example between $k = 0.022$ and $k = 0.035$, where considering only the TE measurements can give us a more accurate estimation of P_k. EE measurements approach the accuracy of TT at a

number of points, there are however regions where the errors become so large that any reconstructed P_k is meaningless (these correspond to the troughs of the EE spectrum). The regions of ℓ-space corresponding to the most accurately measured k regions are the peaks of the spectra. Both TE and EE spectra are not significantly affected by changes to P_k at very low k, so errors in this region are artificially small.

We also tested the combined $P(k)$ and parameter estimation MCMC search as described in the previous section. The optimal binning method found 128 bins for the Planck experiment, where our target signal-to-noise value in each bin is 10. We choose this value so as to have approximately the same number of bins across the range $0.01 < k < 0.03$ where WMAP best probes P_k.

In Figure 3.5 we show the results of a cosmological parameter estimation from this process for a simulated Planck experiment. We find that the input parameters were accurately recovered after this process. We show the final reconstructed P_k in Figure 3.6. The red line is the reconstructed P_k at its best-fit point. The error bars we show are obtained from combining the error from the marginalised distributions with the reconstruction errors given the observed CMB data. The errors are centred around the mean of the marginalised distribution for each bin. It is important to note that the errors are highly correlated. This explains the reduced scatter in the mean values compared to the size of the plotted errors. As was seen in Figure 3.3 changing a single parameter by a small amount (in the manner a MCMC search does) creates a very specific signature on P_k for each parameter, where changing it slightly has a correlated effect upon the whole range of k. This explains the very high correlations observed across the whole k range. It is at odds with the errors on $P(k)$ associated with those of C_ℓ at any best-fit point which are not correlated across large ranges of k.

3.5.2   Constraints from current CMB observations

We used two sets of currently available data to estimate the cosmological parameters in conjunction with a free unparametrised P_k. The first is the WMAP 5-year data alone [8], in the second we combine this with that of SNIa, HST and BBN [132, 133, 134, 135, 136, 137]. The second set of datasets were chosen because the non-CMB sets do not depend upon the form of P_k and can therefore give us an independent and tighter constraint on most of the cosmological parameters. We run this data through the MCMC tool `COSMOMC` in the same manner as we did for the test Planck data.

In Figure 3.7 we show the current constraints on $\Omega_b h^2$, $\Omega_c h^2$, θ and τ. The solid red vertical lines indicates the WMAP only best fit values when $P(k)$ is parametrised in the usual fashion, the black solid curves represent the marginalised probability distribution of the WMAP only data. There is no disagreement between this and the WMAP bestfit model. The blue dotted line shows the marginalised probability distribution of WMAP including the other datasets. Here we observe some tension at around the 1σ level in $\Omega_c h^2$. The inclusion of this data does not move the position of the likelihood peak significantly for each parameter.

The optimal binning method gave us 86 bins in our k range for WMAP. We used the minimum Δk as described previously, this gave a target signal-to-noise ratio of 6. The binned reconstructed P_k are shown in Figs. 3.8 and 3.9 for each of the two data sets. As in the Planck case, the results of the MCMC are highly correlated across the whole k range, in contrast to a single reconstruction which is not. We have similarly added the errors from the MCMC to a single reconstruction in each bin in quadrature. The red lines are the reconstructed P_k at their best-fit points. The error bars are centred around the best-fit marginalised value of each bin. For the WMAP only run we find that the limiting errors on P_k in the range $0.0075 < k < 0.05$ come from uncertainty in the cosmological parameters, whereas this limiting range is around $0.0075 < k < 0.04$ when we include the other datasets. No significant deviation from the standard power-law case is observed in either run.

3.6   Discussion

The SVD process provides a uniquely fast method of reconstructing the primordial power spectrum and so has been incorporated into full MCMC parameter fitting runs. We have tested the method and shown that it recovers the overall features of input spectra. We have applied the method to forecasted Planck data and current WMAP 5-year results. These results allow for the consistent combination of a reconstruction method with a full exploration of the parameter likelihoods for the first time.

We have seen that the limiting factor in constraining the primordial spectrum over a large range of wavenumbers k comes from the uncertainty in cosmological parameters. Any claims of a detection of a feature in P_k must necessarily confront the degeneracy with the cosmological parameter space. This effect will be less important when the Planck data is released, however it must still be considered as the

unprecedented accuracy offered by future data may lead to premature claims of a detection of an interesting feature.

We observed some tension between the WMAP only best-fit model, when $P(k)$ is parametrised with the normal amplitude and tilt, and our method in the marginalised probability distribution of $\Omega_c h^2$. The SVD method disagrees very mildly (at 1σ) with the WMAP only best-fit model in the value of $\Omega_c h^2$. This is not overly significant, however it may be an indication of some departure from a standard power law in $P(k)$.

There are other currently available CMB datasets, which could expand the range of k probed. There however exist difficulties in incorporating these into the current method. By expanding the range of k you stand to increase the amount of time required to perform the inversion. Also, when considering binned C_ℓ data and non-binned data (WMAP) it is not clear how the terms in $M_{\ell k}$ should be weighted.

In the future as CMB polarisation data becomes increasingly accurate it will be desirable to perform a joint inversion of total intensity data along with polarisation data. It is not clear how to extend the SVD based method to include all polarisation modes simultaneously since a HOSVD (Higher-Order SVD) step would probably be required. On the other hand this would give the best estimate of P_k given any dataset and would help to reduce the correlations found in the reconstructed P_k by increasing the degrees of freedom that can be effectively constrained.

There are further extensions of the both methods that could increase their effectiveness. The addition of other observables with different transfer functions such as galaxy redshift surveys or cosmic shear surveys will provide complementary information in the reconstruction.
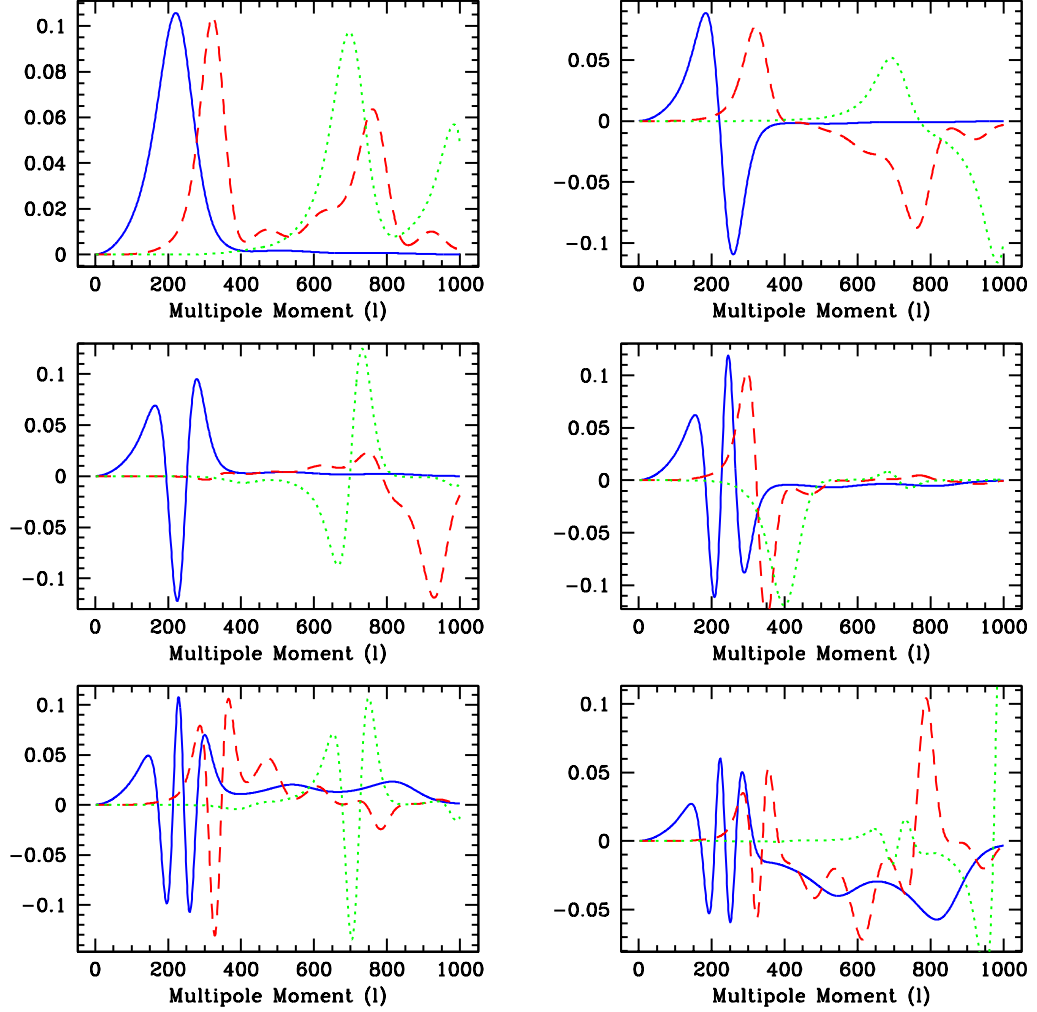
Figure 3.1: We plot the vectors in the matrix **U** associated with the 6 highest singular values. The vector corresponding to the highest singular value is shown in the top left panel, the second highest is shown in the top right panel, etc.. These vectors are the modes best constrained in C_ℓ in the absence of all sources of error. We decomposed $M_{k\ell}$ for a cosmology of $\Omega_b h^2 = 0.0226$, $\Omega_c h^2 = 0.108$, $\theta = 1.041$ and $\tau = 0.076$. The blue/solid line is the TT mode, red/dashed is the TE mode and green/dotted is the EE mode.
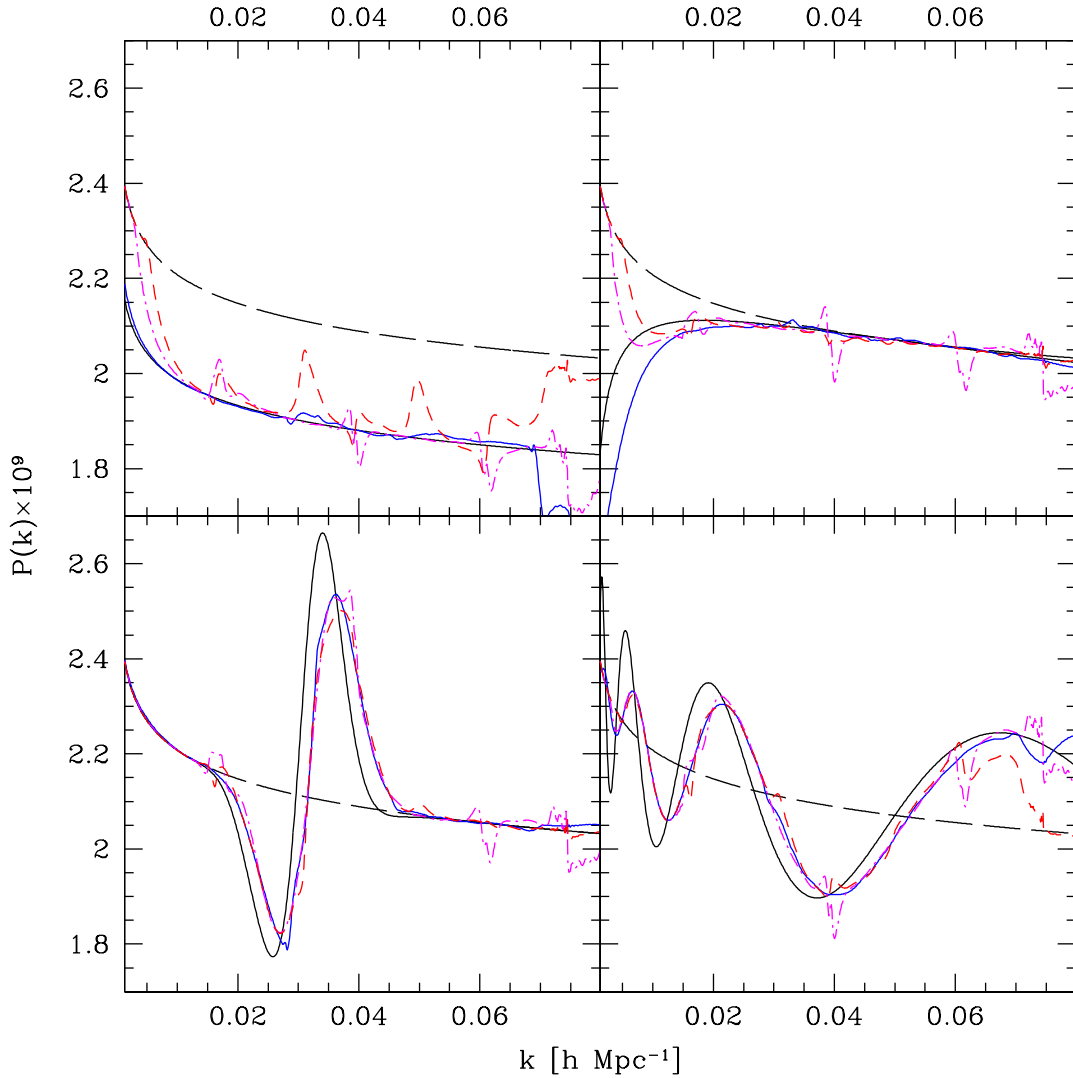
Figure 3.2: The reconstruction of several test spectra. The test models used to generate the simulated C_ℓ, shown in black (thick/solid) curves, are (a) A 10% decrease in power from the WMAP5 best fit amplitude, (b) the WMAP5 best fit model including running $dn_s/d\ln k = -0.037$, (c) a localised feature at around $k = 0.02$ Mpc^{-1}, and (d) a model with sinusoidal oscillations superimposed on the best fit power law spectrum. The black (long-dashed) curves show the best fit spectrum used to minimise any cut-off effects at the ends of the ℓ regions. The blue (solid) curves are the reconstructions using total intensity data whereas the red (dashed) curves and magenta (dot-dashed) curves use TE and EE data respectively. The C_ℓ forecasts assumed an experiment with no noise and an $\ell_{\mathbf{max}} = 1000$.
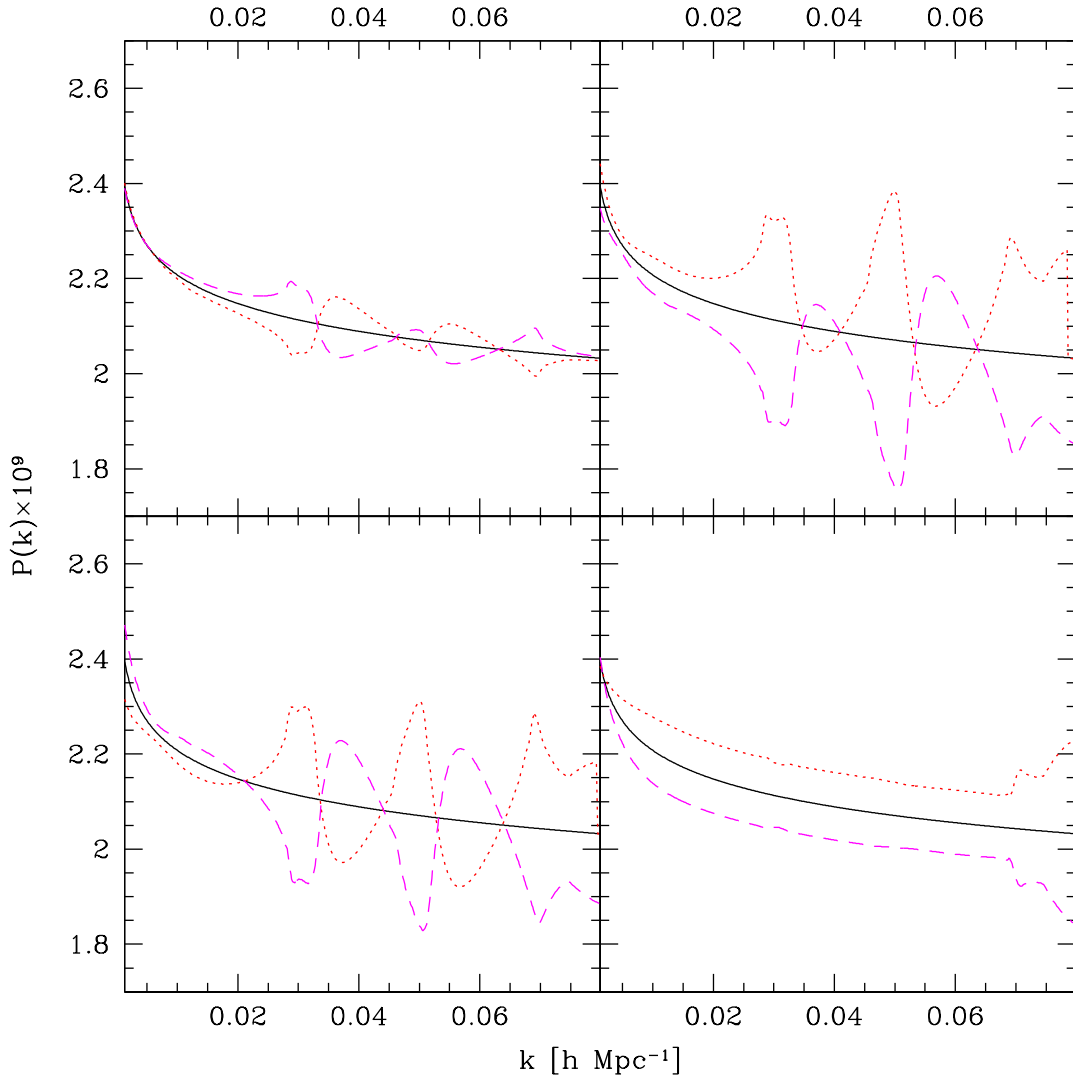
Figure 3.3: The reconstruction of several spectra when the incorrect parameters are used for the fiducial cosmological model. The black (solid) line is the correct input spectrum. In each panel we change a single parameter in the the fiducial model ($\Omega_b h^2 = 0.0226$, $\Omega_c h^2 = 0.108$, $\theta = 1.041$ and $\tau = 0.076$). The red (dot-dashed) line shows the effect of changing a parameter by 1σ in a positive direction from the WMAP best fit model, the magenta (dashed) line shows the effect of a 1σ shift in a negative direction. The top left panel shows the effect of varying $\Omega_b h^2$, top right of $\Omega_c h^2$, bottom left of θ and the bottom right of τ.
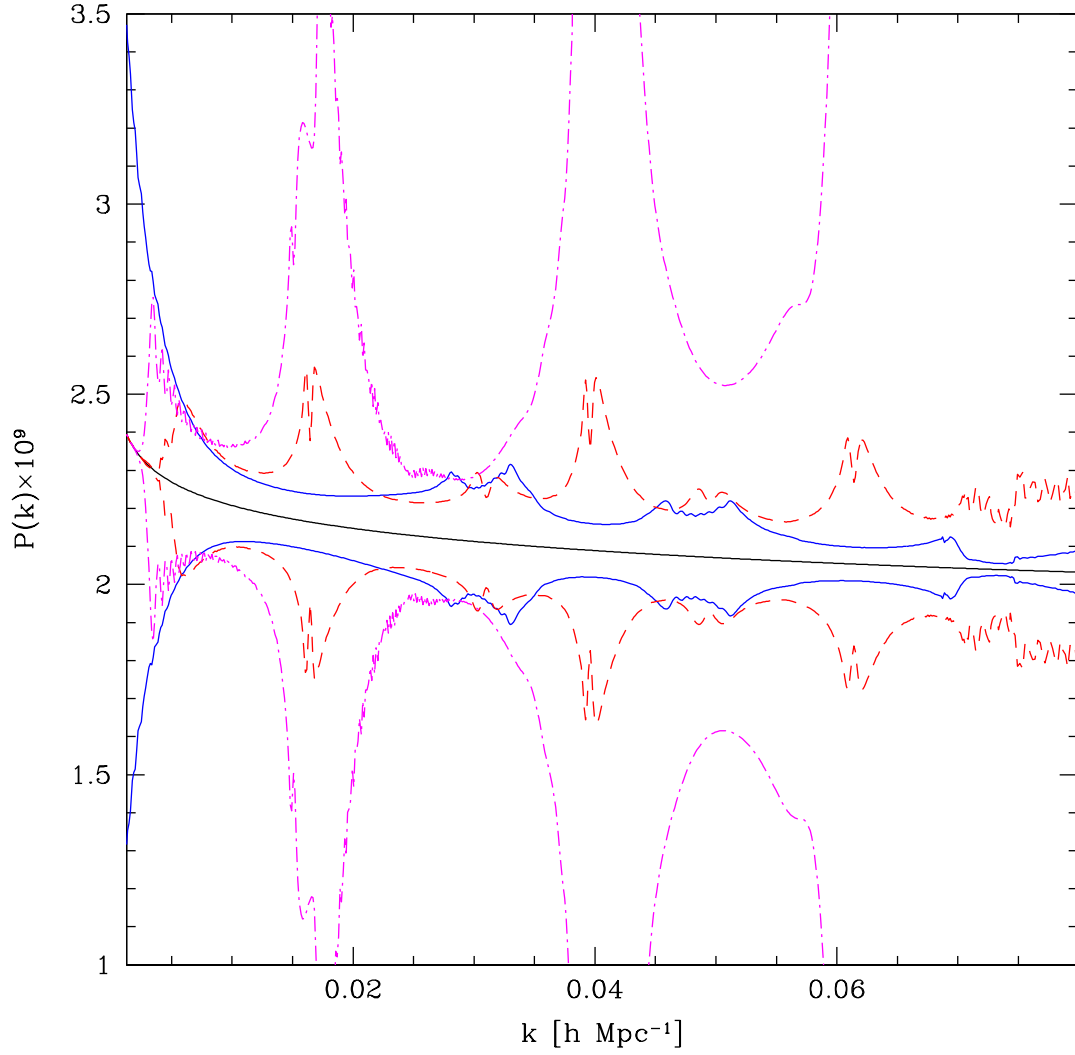
Figure 3.4: We show the unbinned errors on a reconstructed $P(k)$ for a forecasted Planck dataset. The cosmological parameters have been fixed. The blue (solid) lines show the 1σ confidence regions obtained from TT measurements, with red (dashed) and magenta (dot-dashed) showing the same 1σ bounds for TE and EE respectively.
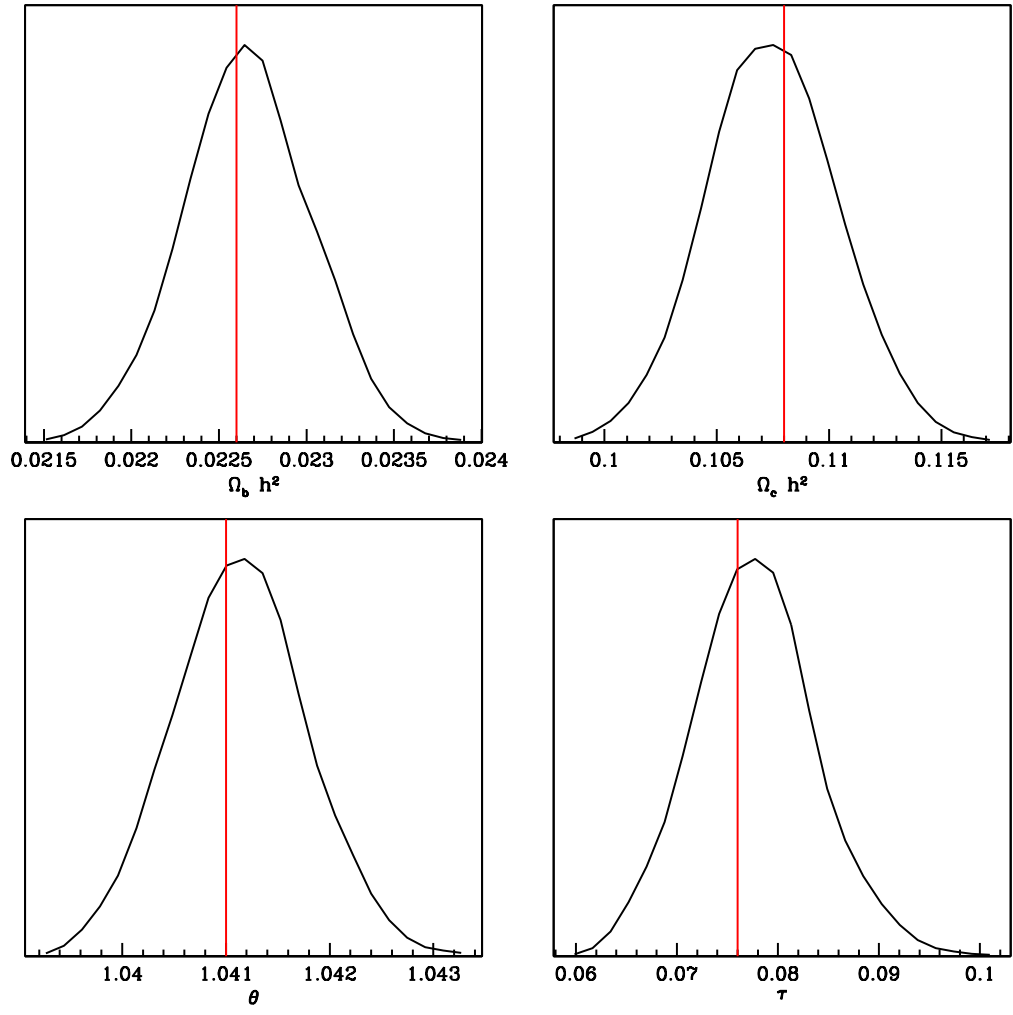
Figure 3.5: Predicted constraints on $\Omega_b h^2$, $\Omega_c h^2$, θ and τ from Planck, when P_k is given total freedom. The solid red vertical lines indicates the input values for each of the parameters. The black solid curves show the marginalised probability distribution.
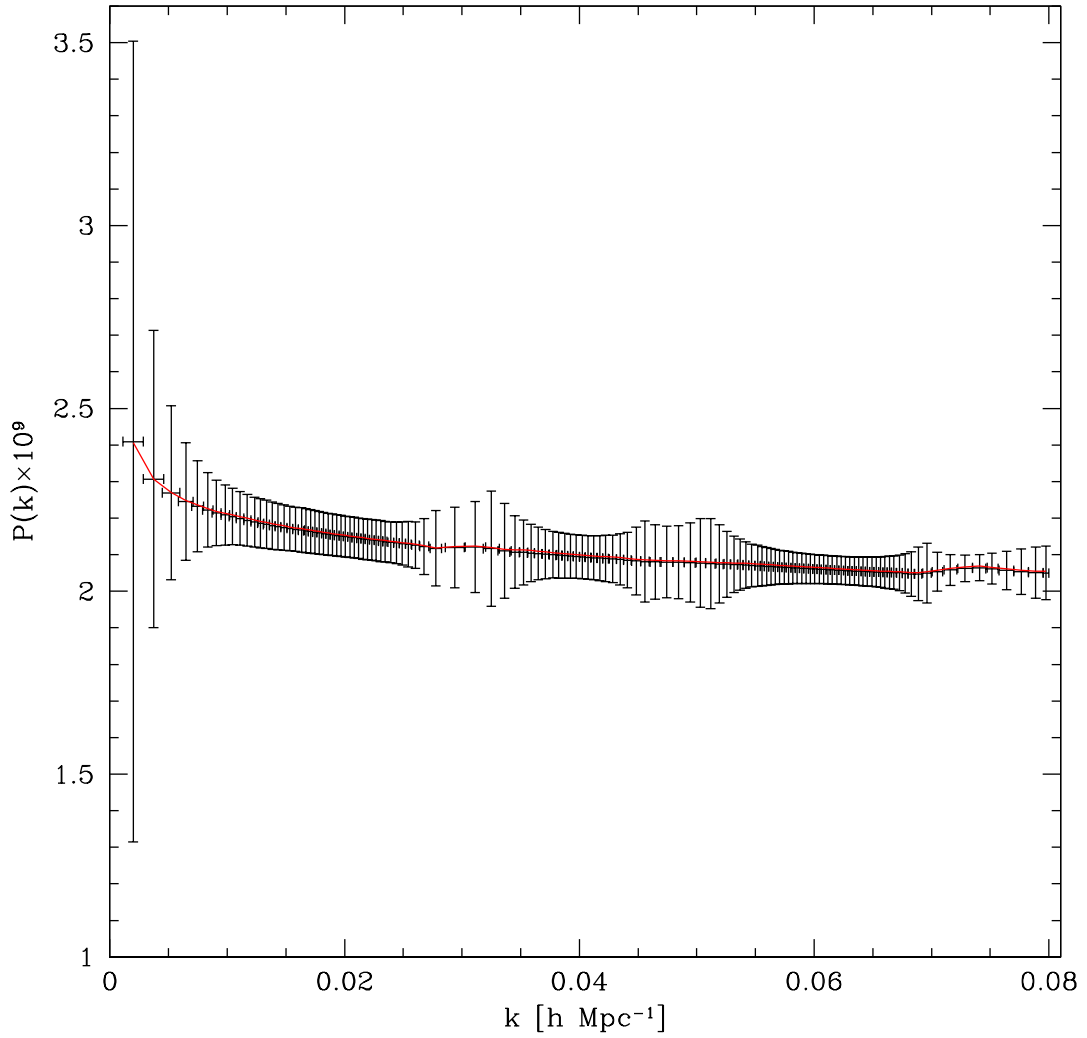
Figure 3.6: Predicted constraints on the form of P_k from Planck. The red line is the reconstructed P_k at the best-fit point. We show in black the marginalised values of each bin, the error bars represent the 1σ error.
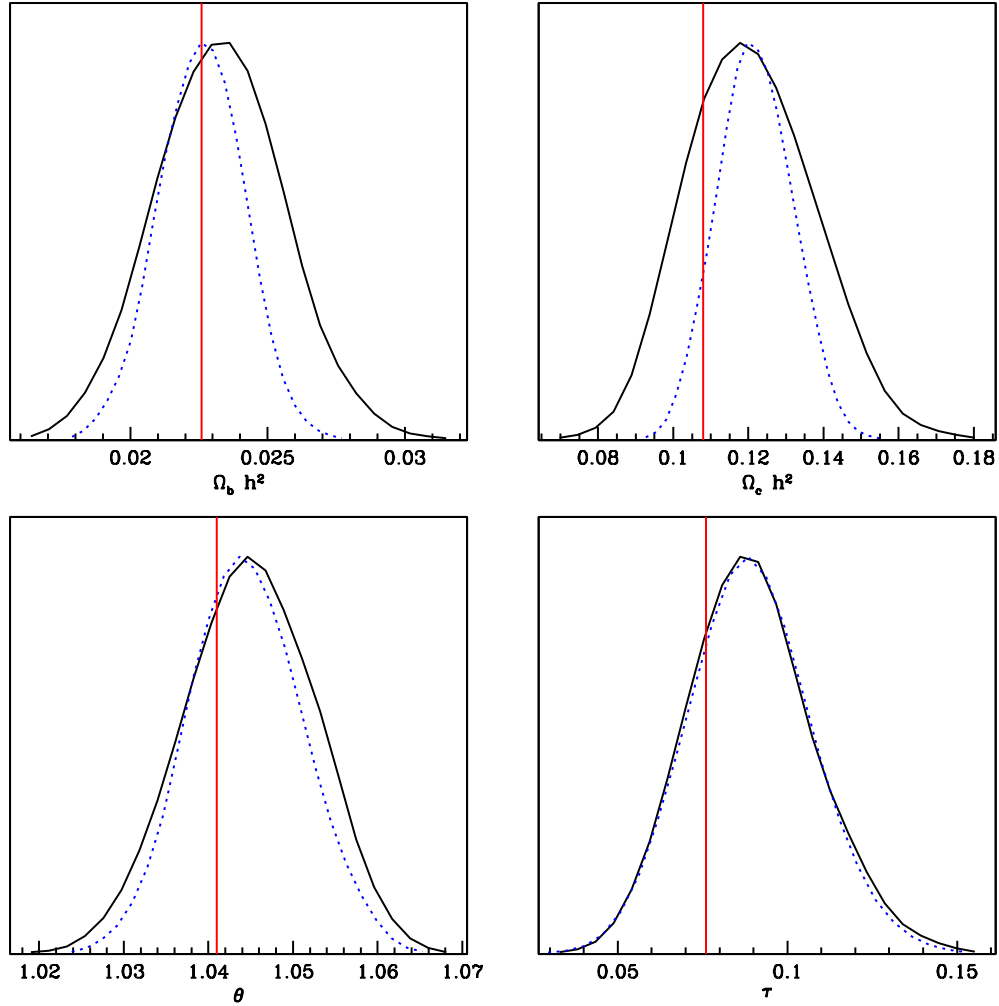
Figure 3.7: The current constraints on $\Omega_b h^2$, $\Omega_c h^2$, θ and τ, when P_k is given total freedom. The solid red vertical lines indicates the WMAP only best fit values when $P(k)$ is parametrised in the usual fashion, the black solid curves represent the marginalised probability distribution of the WMAP only data. The blue dotted line shows the marginalised probability distribution of WMAP including the other datasets.
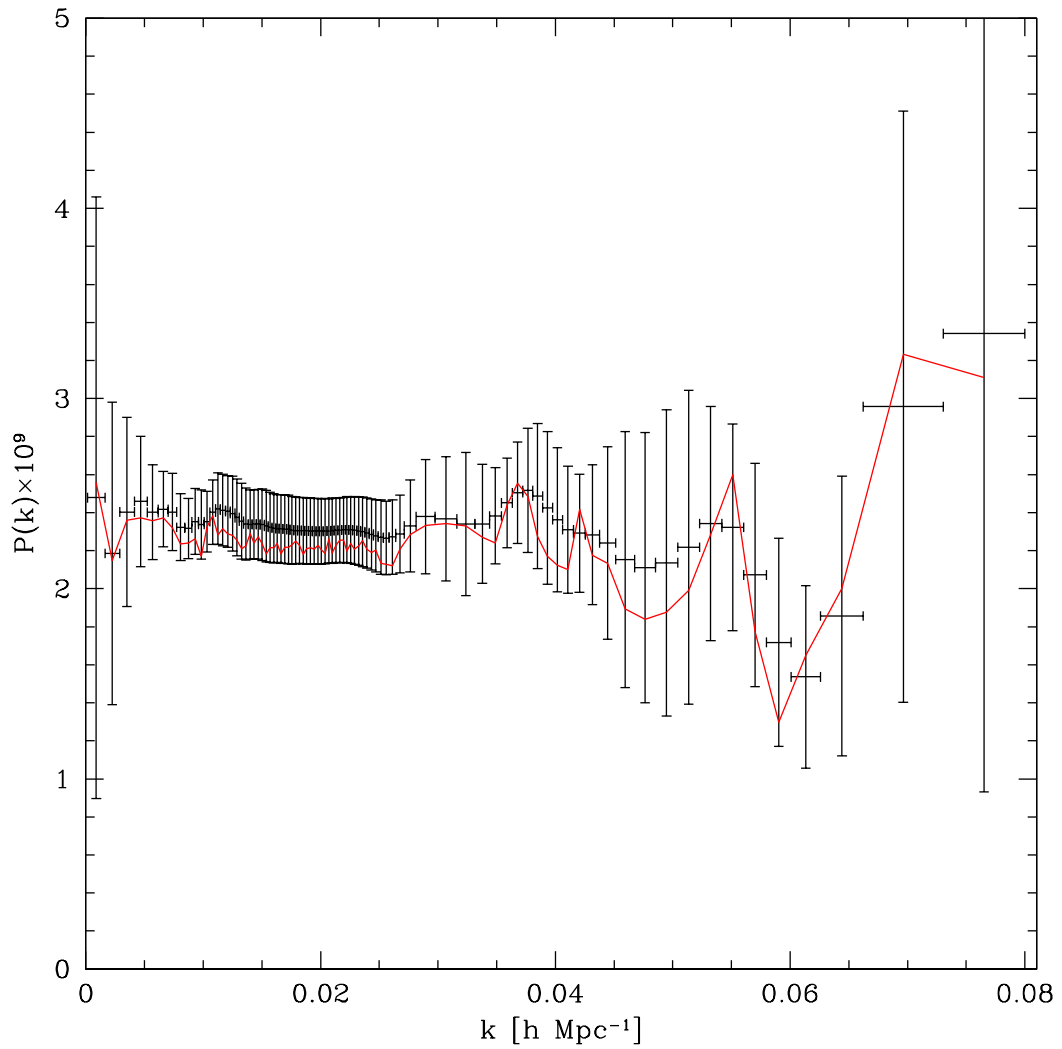
Figure 3.8: The current constraints on the form of P_k from WMAP data only. The red line is the reconstructed P_k at the best-fit point. We show in black the marginalised values of each bin, the error bars represent the 1σ error. Neighbouring points have a covariance value of > 0.95.

Figure 3.9: The current constraints on the form of P_k from WMAP and the other datasets. The red line is the reconstructed P_k at the best-fit point. We show in black the marginalised values of each bin, the error bars represent the 1σ error.

Chapter 4

Bayesian Experimental Design for Sparse Sampling

4.1 Introduction

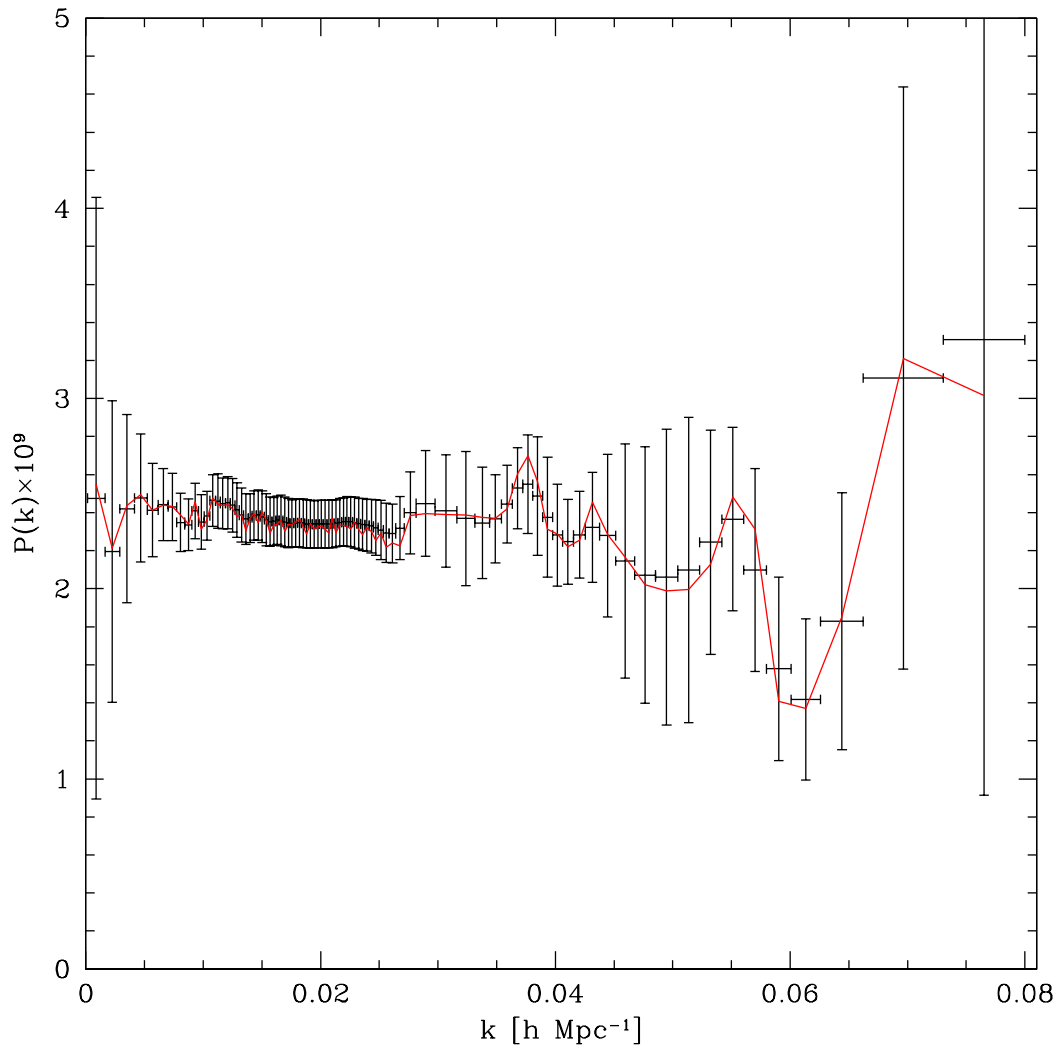As seen in the previous chapters measuring the primordial power spectrum (and indeed some of the cosmological parameters) relies heavily on measurements of the galaxy/CMB/etc. power spectra. Therefore, an accurate measurements of these power spectra would mean more accurate measurements of the primordial power spectrum. In this chapter we investigate the optimisation of the measurements of one of these power spectra: the galaxy power spectrum.

The form of the galaxy power spectrum on large scales provides us with unique information about galaxy clustering. On very large scales one is probing structure which is less affected by clustering and hence it is still in the linear regime and therefore, has a memory of the initial state. The information from these regimes are therefore the cleanest since the Big Bang and any knowledge on these large scales would shed light on the questions about the primordial power spectrum.

To investigate larger scales, it may be more efficient to observe a larger, but sparsely sampled, area of sky instead of a smaller contiguous area. In this case we would gather a larger density of states in Fourier space, but at the expense of an increased correlation between different scales — aliasing. This would smooth out features on these scales and decrease its significance if any observed. However, a common feature amongst many models is that the power spectrum should be smooth over these scales anyways: approaching a Harrison-Zeldovich form, proportional to the wavenumber,

k. In this era of cosmology where the statistical errors have reduced greatly and are now comparable with systematics, observing, for example, a greater number of galaxies may not necessarily improve our results. We need to devise more strategic ways to make our observations and take control of our systematics. Some previous work on sparse sampling include [138, 139]. They try to constrain a parameter, such as the dark energy parameters, from sparse sampling observation. Here, by making use of Bayesian Experimental Design we will investigate the advantages and disadvantages of the sparse sampling and see if a complete contiguous survey is indeed the most efficient way of observing the sky for our purposes. The parameter of interest here is the galaxy power spectrum itself.

4.2   Bayesian Experimental Design and Figure-of-Merit

In Section 1.7 we presented an introduction to the Bayesian statistics. As mentioned in that section, Bayesian methods have recently been used in cosmology for model comparison and for deriving posterior probability distributions for parameters of different models. However, Bayesian statistics can do even more by handling questions about the performance of future experiments, based on our current knowledge [140, 141, 142]. Authors of [70] use a Bayesian approach to constrain the Baryon Acoustic Oscillations (BAO) from a survey parameter space. Here we will use this strength of the Bayesian statistics for optimising the strategy to observe the sky for galaxy surveys. For this, we need to define a quantity of interest, generally called the figure of merit (FoM), associated with the proposed experiment. The choice of FoM depends on the questions being asked. We then want to maximise/minimise (depending on the definition of the FoM) the FoM subject to constraints imposed by the experiment or by our knowledge about the nature of the universe. Below, we will explain the procedure.

Assume e denotes the different experimental designs that we can build and M^i are the different models under consideration with their parameters θ^i. Assume that experiment o has been performed, so that this experiment's posterior $P(\theta|o)$ forms our prior probability function for the new experiment. The FoM — let's now call it U (for Utility function) — will depend on the set of parameters under investigation, the performed experiment (data) and the characteristics of the future experiment;

$U(\theta, e, o)$. From the utility function we can build the expected utility

$$E[U|e,o] = \sum_i P(M^i|o) \int d\hat{\theta}^i \, U(\hat{\theta}^i, e, o) P(\hat{\theta}^i|o, M^i) \,, \qquad (4.1)$$

where $\hat{\theta}^i$ represent the fiducial parameters for model M^i. This says: If a set of fiducial parameters, $\hat{\theta}$, correctly describe the universe and we build an experiment e, then we can compute the utility function for that experiment, $U(\hat{\theta}, e, o)$. However, our knowledge of the universe is described by the current posterior distribution $P(\hat{\theta}|o)$. Averaging the utility over the posterior accounts for the present uncertainty in the parameters and summing over all the available models would account for the uncertainty in the underlying true model.

The aim is to select an experiment that maximises/minimises the expected utility E. The utility function takes into account the current models and the uncertainties in their parameters and therefore maximising it takes into account the lack of knowledge of the true model of the universe.

One of the common choices for a utility is some form of function of the Fisher matrix (We will explain in the next section how a Fisher matrix is obtained). The three most common utility functions are

- A-optimality: trace of the Fisher matrix (or its log) and is proportional to sum of the variances. This prefers a spherical error region, but may not necessarily select the smallest volume.

- D-optimality: determinant of the Fisher matrix (or its log), which is inversely proportional to the square of the parameters volume enclosed by the posterior. This is a good indicator of the overall size of the error over all parameter space, but is not sensitive to any degeneracies amongst the parameters.

- Entropy (also called the Kullback-Leibler divergence):

$$\begin{aligned} E &= \int d\theta \, P(\theta|\hat{\theta}, e, o) \log \frac{P(\theta|\hat{\theta}, e, o)}{P(\theta|o)} \\ &= -\frac{1}{2} \left[\log|\mathbf{F}| - \log|\mathbf{\Pi}| - \text{trace}(1 - \mathbf{\Pi}\mathbf{F}^{-1}) \right] \,, \qquad (4.2) \end{aligned}$$

where $P(\theta|\hat{\theta}, e, o)$ is the posterior distribution with Fisher matrix $\mathbf{F}$ and $P(\theta|o)$ is the prior distribution with the Fisher matrix $\mathbf{\Pi}$. The entropy forms a nice compromise between the A-optimality and D-optimality.

In the next section we will explain how a Fisher matrix is formulated.

4.3   Fisher Matrix Analysis

The Fisher matrix is generally used to determine the sensitivity of a particular survey to a set of parameters and has been largely used for optimisation (and forecasting). Consider the likelihood function for a future experiment with experimental parameters e, $\mathcal{L}(\theta|e) \equiv P(D_{\hat{\theta}}|\theta, e)$, where $D_{\hat{\theta}}$ are simulated data from the future experiment assuming that $\hat{\theta}$ are the true parameters in the given model. We Taylor expand the log-likelihood around its maximum value:

$$\ln \mathcal{L}(\theta|e) = \ln \mathcal{L}(\theta^{ML}) + \frac{1}{2}\sum_{ij}(\theta_i - \theta_i^{ML})\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j}(\theta_j - \theta_j^{ML})\,, \qquad (4.3)$$

where the first term is a constant and only affects the height of the function, the second term describes how fast the likelihood function falls around the maximum. Fisher matrix is defined as the ensemble average of the *curvature* of a function $\mathcal{F}$ (i.e., it is the average of the curvature over many realizations of signal and noise);

$$F_{ij} = \langle \mathcal{F} \rangle = \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial \theta_i \partial \theta_j} \right\rangle \qquad (4.4)$$

$$= \frac{1}{2}\text{trace}[C_{,i}C^{-1}C_{,j}C^{-1}] \qquad (4.5)$$

where C is the total covariance matrix, θ_i/θ_j are the different parameters and $\mathcal{L}$ is the likelihood function. Its inverse is an approximation of the covariance matrix of the parameters, by analogy with a Gaussian distribution in the θ_i/θ_j, for which this would be exact. The Cramer-Rao inequality[1] states that the smallest error measured, for θ_i, by any unbiased estimator (such as the maximum likelihood) is $1/\sqrt{F_{ii}}$ and $\sqrt{(F^{-1})_{ii}}$, for non-marginalised and marginalised[2] one-sigma errors respectively. The derivatives in Equation 4.4 generally depend on where in the parameter space they are calculated and hence it is clear that the Fisher matrix is function of the fiducial parameters.

The Fisher matrix allows us to estimate the errors on parameters without having

---

[1]It should be noted that the Cramer-Rao inequality is a statement about the so-called "Frequentist" confidence intervals and is not strictly applicable to "Bayesian" errors.

[2]Integration of the joint probability over other parameters.

to cover the whole parameter space. The authors of [37] have compared the Fisher matrix analysis with the full likelihood function analysis and found there was great agreement between the two methods if the likelihood function is approximately Gaussian near the peak. It also makes the calculations easier. For example, if we are only interested in a subset of parameters, then marginalising over unwanted parameters is just the same as inverting the Fisher matrix, taking only the rows and columns of the wanted parameters and inverting the smaller matrix back. It is also very straightforward to combine constraints from different independent parameters: we just sum over the Fisher matrices of the experiments (remember Fisher matrix is the log of the likelihood function).

We further note, as in all uses of the Fisher matrix, that any results thus obtained must be taken with the caveat that these relations only map onto realistic error bars in the case of a Gaussian distribution, usually most appropriate in the limit of high signal-to-noise ratio and/or relatively small scales, so that the conditions of the central limit theorem obtain. As long as we do not find extremely degenerate parameter directions, we expect that our results will certainly be indicative of a full analysis, using simulations and techniques such as Bayesian Experimental Design [69].

4.3.1   Fisher Matrix for Galaxy Surveys

We follow the approach of [36] to define the pixelisation for galaxy surveys. First we define the data in pixel i as

$$\Delta_i \equiv \int d^3x \, \psi_i(\boldsymbol{x}) \left[\frac{n(\boldsymbol{x}) - \bar{n}}{\bar{n}} \right] , \qquad (4.6)$$

where $n(\boldsymbol{x})$ is the galaxy density at $\boldsymbol{x}$ and $\bar{n}$ is the expected number of galaxies at $\boldsymbol{x}$. The weighting function, $\psi_i(\boldsymbol{x})$, which determines the pixelisation, is defined as a set of Fourier pixels

$$\psi_i(\boldsymbol{x}) \;=\; \frac{e^{i\boldsymbol{k}_i \cdot \boldsymbol{x}}}{V} \begin{cases} 1 & \boldsymbol{x} \text{ inside survey volume} \\ 0 & \text{otherwise} \end{cases} , \qquad (4.7)$$

where V is the volume of the survey. Here we have divided the volume into sub-volumes, each being much smaller than the total volume of the survey, but being large enough to contain many galaxies. This means Δ_i is the fractional over-density

in pixel i. Using this pixelisation we can define a covariance matrix as

$$\langle \Delta_i \Delta_j^* \rangle = C = (C_S)_{ij} + (C_N)_{ij} , \qquad (4.8)$$

where C_S and C_N are the signal and noise covariance matrices respectively. The signal covariance matrix can be defined as (assuming noise and signal are independent)

$$
\begin{aligned}
(C_S)_{ij} &= \langle \Delta_i \Delta_j^* \rangle \\
&= \int d^3x \, d^3x' \, \psi_i(\boldsymbol{x}) \psi_j^*(\boldsymbol{x}') \left\langle \frac{n(\boldsymbol{x}) - \bar{n}}{\bar{n}} \cdot \frac{n(\boldsymbol{x}') - \bar{n}}{\bar{n}} \right\rangle . \qquad (4.9)
\end{aligned}
$$

By setting the term in the brackets to $\delta(\boldsymbol{x})$ — the continuous over-density described in Equation 1.6 — we obtain

$$
\begin{aligned}
(C_S)_{ij} &= \int \frac{d^3k}{(2\pi)^3} P(k) \tilde{\psi}_i(\boldsymbol{k}) \tilde{\psi}_j^*(\boldsymbol{k}) \\
&= \int \frac{dk}{(2\pi)^3} k^2 P(k) \int d\Omega_k \, \tilde{\psi}_i(\boldsymbol{k}) \tilde{\psi}_j^*(\boldsymbol{k}) \\
&= \int \frac{dk}{(2\pi)^3} k^2 P(k) W_{ij}(k) , \qquad (4.10)
\end{aligned}
$$

where we have defined the window function, $W_{ij}(k)$, as the angular average of the square of the Fourier transform of the weighting function.

The design of the survey will shape the form of the weighting function in Equation 4.7. For simplicity, we will approximate the design of the sparsely sampled area of the sky as a regular grid of square patches of size $M \times M$. This design is shown in Figure 4.1. We therefore define the structure on the sky as a square wave in both x and y directions

$$
\sum_i \Pi(x - x_i) = \begin{cases} 1 & 0 < |x - x_i| < M/2 \\ 0 & \text{otherwise} \end{cases} , \qquad (4.11)
$$

$$
\sum_j \Pi(y - y_j) = \begin{cases} 1 & 0 < |y - y_j| < M/2 \\ 0 & \text{otherwise} \end{cases} , \qquad (4.12)
$$

where x_i and y_j mark the centres of the steps in our coordinate system. In the z
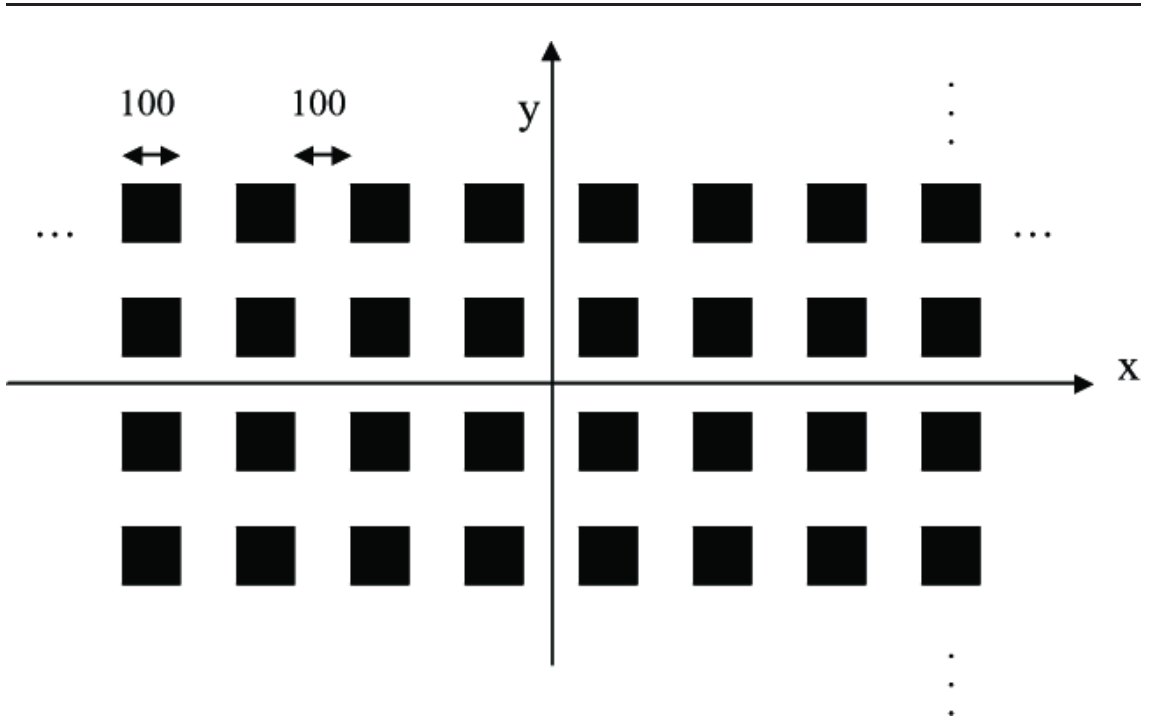
Figure 4.1: Design of the sparsely sampled area of the sky. We assume a series of top-hat functions (of size 100Mpc) in x and y directions on the surface of the sky — flat-sky approximation has been assumed.

direction we use the step function, which is defined as:

$$\Theta(z) \;=\; \begin{cases} 1 & z > 0 \\ 0 & \text{otherwise} \end{cases}. \qquad (4.13)$$

Using these the weight function takes the form:

$$
\begin{aligned}
\tilde{\psi}_i(\boldsymbol{k}) \;&=\; \frac{1}{V}\int d^3x\; e^{i(\boldsymbol{k}_i-\boldsymbol{k}).\boldsymbol{x}} \cdot \Theta(z+\tfrac{L}{2})\Theta(\tfrac{L}{2}-z) \cdot \sum_n \Pi(x-x_n) \cdot \sum_m \Pi(y-y_m) \\
&=\; \frac{1}{V}\int dx\; e^{iq_x x}\sum_n \Pi(x-x_n)\int dy\; e^{iq_y y}\sum_m \Pi(y-y_m)\int dz\; e^{iq_z z}\Theta(z+\tfrac{L}{2})\Theta(\tfrac{L}{2}-z) \\
&=\; \frac{M^2 L}{V} sinc(q_x\tfrac{M}{2})\sum_n 2\cos(q_x x_n)\cdot sinc(q_y\tfrac{M}{2})\sum_m 2\cos(q_y y_m)\cdot sinc(q_z\tfrac{L}{2})\,, \qquad (4.14)
\end{aligned}
$$

where $\boldsymbol{q} = \boldsymbol{k}_i - \boldsymbol{k}$ and V is the total sparsely sampled volume, M is the size of the observed patch on the surface of the sky and L is the observed depth. The last equality is obtained by using the symmetry in the design and using the *Dirichlet*

120

Kernel defined as:

$$D_n(x) = \sum_{k=-n}^{n} e^{ikx} = 1 + 2\sum_{k=1}^{n} \cos(kx) \,. \tag{4.15}$$

The first term on the right hand side of this equation can be removed if there is no step at the origin of the coordinate. Therefore the window function, defined in Equation 4.10, takes the form

$$
\begin{aligned}
W_{ij}(k) &= \int_{-1}^{1} d\mu \int_{0}^{2\pi} d\phi \; \tilde{\psi}(\boldsymbol{k}_i - \boldsymbol{k})\tilde{\psi}^*(\boldsymbol{k}_j - \boldsymbol{k}) \\
&= \int_{-1}^{1} d\mu \int_{0}^{2\pi} d\phi \; \frac{M^4 L^2}{V^6} \\
&\quad \left[sinc(q_x \frac{M}{2}) \sum_n 2\cos(q_x x_n) \cdot sinc(q'_x \frac{M}{2}) \sum_{n'} 2\cos(q'_x x_{n'}) \right] \\
&\quad \left[sinc(q_y \frac{M}{2}) \sum_m 2\cos(q y_m) \cdot sinc(q'_y \frac{M}{2}) \sum_{m'} 2\cos(q'_y y_{m'}) \right] \\
&\quad \left[sinc(q_z \frac{L}{2}) \cdot sinc(q'_z \frac{L}{2}) \right] \,, \tag{4.16}
\end{aligned}
$$

where $q_x = q\sin\theta\cos\phi$, $q_y = q\sin\theta\sin\phi$, $q_z = q\cos\phi$ and $d\mu = d\cos\theta$.

[Aside: In case of the contiguous sampling of the sky where we are observing through a contiguous square, the window function takes the form:

$$
\begin{aligned}
W_{ij}(k) &= \int_{-1}^{1} d\mu \int_{0}^{2\pi} d\phi \; \tilde{\psi}(\boldsymbol{k}_i - \boldsymbol{k})\tilde{\psi}^*(\boldsymbol{k}_j - \boldsymbol{k}) \\
&= \int_{-1}^{1} d\mu \int_{0}^{2\pi} d\phi \; \frac{M^4 L^2}{V^2} \\
&\quad \left[sinc(q_x \frac{M}{2}) \cdot sinc(q'_x \frac{M}{2}) \right] \\
&\quad \left[sinc(q_y \frac{M}{2}) \cdot sinc(q'_y \frac{M}{2}) \right] \\
&\quad \left[sinc(q_z \frac{L}{2}) \cdot sinc(q'_z \frac{L}{2}) \right] \,, \tag{4.17}
\end{aligned}
$$

which is a square tube.]

The noise covariance matrix — which is due to Poisson shot noise — is given by

$$
\begin{aligned}
(C_N)_{ij} &= \int d^3x \frac{\psi_i(\boldsymbol{x})\psi_j^*(\boldsymbol{x})}{\bar{n}} \\
&= \frac{1}{\bar{n}V^2} \int d^3x \; e^{i(\boldsymbol{k}_i - \boldsymbol{k}_j)\cdot\boldsymbol{x}} \Theta(z + \frac{L}{2})\Theta(\frac{L}{2} - z) \sum_n \Pi(x - x_n) \sum_m \Pi(y - y_m) \\
&= \frac{4M^2 L}{V N_g} \cdot sinc(l_z \frac{L}{2}) \\
&\quad sinc(l_x \frac{M}{2}) \sum_i \cos(l_x x_i) \cdot sinc(l_y \frac{M}{2}) \sum_j \cos(l_y y_j) \,,
\end{aligned}
\tag{4.18}
$$

where $\boldsymbol{l} = \boldsymbol{k}_i - \boldsymbol{k}_j$, N_g is the total number of galaxies in the survey, M is the width of the steps in x and y directions and L is the width of the step in the z direction just like above.

This prescription gives us a data covariance matrix for a galaxy survey. What we actually need is a Fisher matrix for the parameters we are interested in. For this we will use Equation 4.4 above, which defines the Fisher matrix of parameters in terms of the inverse of the data covariance matrix and its differentiation with respect to the parameters of interest. We are interested in the galaxy power spectrum and hence the differentiation of the covariance matrix in Equation 4.4 is taken with respect to the bins of this power spectrum. As the noise covariance matrix does not depend on the power spectrum, we differentiate the signal covariance matrix of Equation 4.10. Taking the galaxy power spectrum as a series of top-hat bins

$$
P(k) = \sum_B w_B(k) P_B \,,
\tag{4.19}
$$

where P_B is the power in each bin B and $w_B = 1$ if $k \in B$ and 0 otherwise, the differentiation takes the form

$$
\frac{\partial (C_S)_{ij}}{\partial P(k)} = \int_{k_B^{min}}^{k_B^{max}} \frac{dk}{(2\pi)^3} k^2 W_{ij}(k) \,.
\tag{4.20}
$$

The bins are arranged using the total volume of the survey: $k_{min} = (2\pi/V)^{1/3} = dk$. Therefore, a larger volume gives more number of bins. We take this and the inverse of the data covariance matrix and insert into Equation 4.4 to get a Fisher matrix for the galaxy power spectrum bins.

4.4 Results

We will choose a geometrically flat ΛCDM model with adiabatic perturbations and the WMAP5 [24] values for the parameters: $\Omega_m = 0.214 \pm 0.027$, $\Omega_b = 0.044 \pm 0.003$, $\Omega_\Lambda = 0.742 \pm 0.03$, $\tau = 0.087 \pm 0.017$ and $h = 0.719 \pm 0.0265$, where $H_0 = 100h\mathrm{km}^{-1}\mathrm{Mpc}^{-1}$. The utility functions used are

$$\mathrm{Entropy} = -\left[-\frac{1}{2}\left[\log|\mathbf{F}| - n_B \right] \right], \quad (4.21)$$

$$\mathrm{A\text{-}optimality} = \log(\mathrm{trace}(\mathbf{F})), \quad (4.22)$$

$$\mathrm{D\text{-}optimality} = \log(|\mathbf{F}|), \quad (4.23)$$

where n_B is the number of bins of the galaxy power spectrum. These functions are defined so that they need to be maximised for an optimal design.

4.4.1 Sparse Vs Contiguous: Constant Observing Time

We first compare the two observing designs for the same investment of observing time. In case of the sparse sampling we can observe a larger area of the sky for the same observing time, which means larger Fourier density. However, due to the sparse nature of the observation the correlation between the Fourier states is expected to increase.

Galaxy surveys contain a list of galaxies with their positions, where the positions are obtained through measuring the redshift of these galaxies. The further away or the fainter a galaxy is, the more time will be needed to observe the galaxy and obtain its redshift. A galaxy with an intrinsic luminosity L_* at redshift z has a flux

$$Flux = L_* 4\pi d_L(z), \quad (4.24)$$

where $d_L(z)$ is the luminosity distance to the galaxy. If this galaxy is observed in time t, with energy E, with a telescope of an area A_{tel}, we have a flux given by

$$Flux = \frac{E}{A_{tel} \cdot t}. \quad (4.25)$$

From Equations 4.24 and 4.25, we can see that we need an observing time of

$$time = \frac{E}{L_* 4\pi d_L(z) \cdot A_{tel}}, \quad (4.26)$$

to observe a galaxy with luminosity L_* at redshift z. We will assume that galaxies with this luminosity are randomly distributed in the universe with number density $\bar{n}$ (i.e. no clustering assumed — we will look far away enough so that this assumption is valid). This means within each redshift bin there are $N = \bar{n}V_{bin}$ (V_{bin} = volume of each redshift bin) galaxies that need to be observed and therefore we need a total time of $t \times N$ to observe all the galaxies in that redshift bin. The further away the redshift bin is the larger t and hence the larger $t \times N$ is. Summing over all the redshift bins gives the total observing time T of the survey. It is then our choice how to spend this observing time T in depth and area. We can observe closer galaxies but on a larger surface area on the sky or we can observe far away galaxies but through a smaller area. This depends on what we are trying to measure. Here we investigate the two choices for both sparse and contiguous sampling and find the optimal design for measuring the galaxy power spectrum.

We start by having 4 redshift bins in the range $z = 0.3 - 1.1$, each having a binwidth of $\delta z = 0.2$. We therefore have; $z = 0.3 - 0.5,\ 0.5 - 0.7,\ 0.7 - 0.9,\ 0.9 - 1.1$. For this initial redshift range we have 400 patches on the surface of sky (20 in each direction), each having a size 100Mpc $\times$ 100Mpc. The grid, shown in Figure 4.1, is designed so that the patches are placed at 100Mpc from each other. We then start to remove the last redshift bins (one at a time) and observe a larger area of the sky in the remaining redshift bins for the same observing time. The different cases are summarised in Table 4.1. It would be useful to compare the values given in this Table with some of the real surveys — given in Figure 4.2 .

	z range	Patches	A_{obs}/Mpc^2	A_{sps}/Mpc^2	$A_{obs/sps}$	V_{obs}/Mpc^3	V_{sps}/Mpc^3	N_g
1	$0.3 - 1.1$	20×20	4×10^6	1.5×10^7	0.263	2.4×10^{10}	8.9×10^{10}	1.74×10^8
2	$0.3 - 0.9$	30×30	9×10^6	3.5×10^7	0.259	3.8×10^{10}	1.5×10^{11}	3.24×10^8
3	$0.3 - 0.7$	46×46	2.1×10^7	8.3×10^7	0.256	5.7×10^{10}	2.2×10^{11}	5.09×10^8

Table 4.1: Summarising the three different designs explained in Section 4.4.1 for a constant observing time. The total observed area is shown in the 4th column and the 5th column shows the total sparsely sampled area. In the 6th column we show the ratio of different areas. In the 7th and 8th columns we show the volumes observed. Note that in the case of the contiguous sampling $A_{obs} = A_{sps}$ and $V_{obs} = V_{sps}$. In the 9th column we show the total number of galaxies observed in each case.

Figures 4.3-4.8 show the diagonal elements of the window functions, i.e. $W_{ii}(k)$, for the sparse and contiguous sampling for the three different cases of Table 4.1. To compare the aliasing of scales between the sparse and contiguous sampling, we zoom in in Figures 4.4, 4.6 and 4.8. The sparse sampling seems to have a more
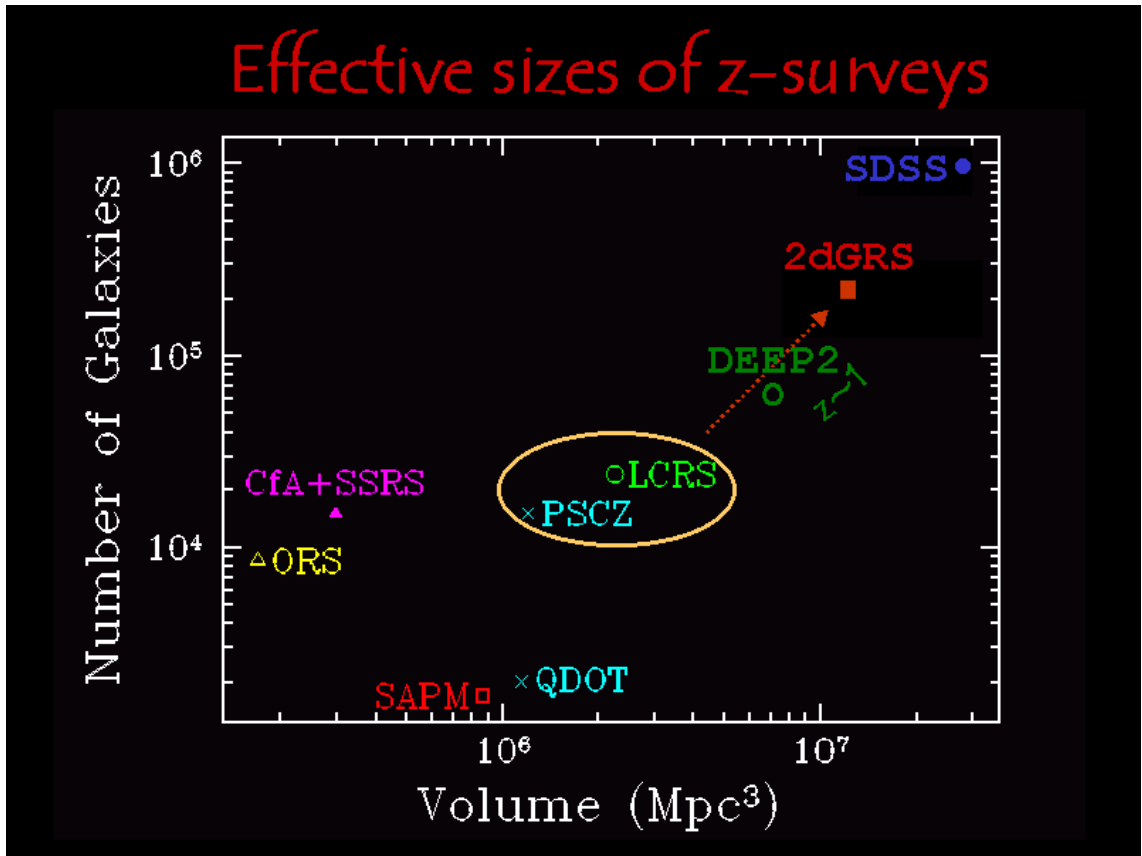
Figure 4.2: Geometries of some of the redshift surveys [14].

complex behaviour in all three cases, and hence is expected to have a more complex aliasing between the scales. This will cause correlation between k scales in the power spectrum. The utility functions will show the extent of the importance of this aliasing and will decide if this can be forgiven for what we are trying to measure. The variations seen in where the maximum of the window functions lie, especially apparent in the sparse cases, are a result of the choice of the Fourier pixels, k_is. Remember that the window functions are a function of $q = k_i - k$ — Equation 4.16. These variations are a reflection of $\text{sinc}(q) \cdot \cos(q)$; for example in case 1 of the sparse sampling q varies in a sinusoidal way. Also note that the window functions suffer from edge effect, where they have cut off at both small and large k in the range. Therefore, any features at the far ends of the k-range should be ignored.

The utility functions are shown in Figure 4.9. In all three cases sparse sampling does a better job than the contiguous sampling. Also, all the utility functions seem to rise from case 1 to case 3. This means to measure the galaxy power spectrum we would gain more by observing near galaxies but observing a larger surface area. This might be completely different if we were to measure other parameters. For example,

to constrain the running of dark energy parameters, we would require a large z range instead of large surface area. The entropy, which is a nice summary of the other two utility functions, shows that overall the sparse sampling is the preferred observing design for galaxy power spectrum measurements. The power spectrum is indeed resolved better due to the larger volume in the case of the sparse sampling, and is also measured more accurately. And for the sparse sampling, case 3 is preferred over the other cases.

The marginalised errors, i.e. $(F^{-1})_{ii}$, for all cases are shown in Figure 4.10. Note that the features at the very far ends should be ignored due to edge effects. As explained above sparse sampling measures the bins of power spectrum with a better accuracy for all cases. Also, note how case 3 resolves the power spectrum better while decreasing the errors of the bins.

4.4.2 Sparse: varying step sizes

Here we investigate the influence of step sizes and the extent they could improve the performance of the survey. For the sparse sampling of case 1 above, we will vary the step sizes in the range $M = 10 - 150\text{Mpc}$ and monitor the variation of the utility functions as M is increased. Here we do not try to keep the observing time constant, knowing the observing time will be higher for larger steps anyways. Instead we want to see the gain for spending more time observing larger steps. The results are shown in Figures 4.11 and 4.12. These Figures are the same but plotted with different x axes: Figure 4.11 shows the utility functions against A_{obs}/A_{sparse} (which might be thought of as the information gain) and Figure 4.12 shows the same results against the step sizes. An overall but slow rise is seen in both D-optimality and entropy. In the mid ranges, it seems like a saturation limit is reached where larger patches don't gain much more. The A-optimality shows a more interesting behaviour where it seems that having larger patches may even increase the errors of the bins of the power spectrum. Therefore, depending on what we are more interested in, larger patches might be improving or worsening the performance of the survey.

4.4.3 Sparse: a power law spectrum instead of galaxy power spectrum

Since the parameters of interest are bins of the galaxy power spectrum we wanted to see how the shape of this power spectrum influences the performance. Hence we have

	-Entropy	A-optimality	D-optimality
Contig: galaxy PS	15.78	3.50	74.66
Contig: primordial PS	9.37	3.08	61.73
Sparse: galaxy PS	25.98	3.73	119.95
Sparse: primordial PS	15.62	3.30	99.23

Table 4.2: Utility functions for a power law spectrum compared to the galaxy power spectrum.

repeated case 1 for both sparse and contiguous sampling for a power law spectrum with the form $P(k) = Bk^{0.96}$, which is the primordial power spectrum—normalized so that it has the same as the galaxy power spectrum at $k = 0.05h\text{Mpc}^{-1}$ as this is our pivot scale.. The marginalised errors are shown in Figure 4.13. The influence of the shape of the power spectrum is clearly seen in the errors. Depending solely on the shape of the spectrum the same survey measures some scales better for one spectrum and other scales better for the other spectrum. The utility functions are shown in Table 4.2. It seems the survey measures the galaxy power spectrum much better than the primordial power spectrum for both sparse and contiguous sampling.

4.4.4 Sampling Pixels

One set of parameters that have great influence on the final results in all the above cases is the choice of the sampling pixels; k_i in Equation 4.7. In theory we would prefer an infinite number of sampling pixels, however in reality, we are limited to these Fourier pixels — Equation 4.7. Here we investigate how the number and range of the sampling pixels influence our results. So far, we have let the largest scales of the survey choose the pixeling, i.e. from the usual properties of the Fourier transform, the largest scales of the survey determines the minimum k, k_{min}, and the minimum binwidth, $(dk)_{min}$. This was then applied to both the binning of the power spectrum and the sampling pixels, i.e. $k_i = k_B$ for all the cases. Here we investigate other choices of the sampling pixels and compare the utility functions to previous cases.

We will focus on case 1 above for both contiguous and sparse sampling and monitor the variation in the utility functions as k_i is changed. Previously, we had 43 and 68 bins for the contiguous and sparse sampling respectively. We will now try

$k_i = 0.00001, 0.00101, 0.00201, \cdots, 0.10001\, h\text{Mpc}^{-1}$, having 101 pixels, and

$k_i = 0.00001, 0.00068, 0.00134, \cdots, 0.10001\, h\text{Mpc}^{-1}$, having 151 pixels.

Another interesting choice would be to extend the k-range to beyond the current limits. So far we have limited ourselves to $k_{max} < 0.1h\mathrm{Mpc}^{-1}$ as we did not want to enter the non-linear regime. Here we want to the extend the range of the sampling pixels to beyond this limit and investigate how much information we can actually recover from the non-linear regime. For this, we use k_{min}, $(dk)_{min}$ as obtained by the largest scales of the surveys (just like the previous cases) but extend k_{max} to $0.15h\mathrm{Mpc}^{-1}$. This means the number of pixels would now be 66 and 103 for the contiguous and sparse sampling cases respectively (compared to 43 and 68 pixels in the previous cases).

The utility functions are shown in Figures 4.14. As the sampling pixels increase for the same k-range, more information is fed into the measurements of the bins (Equation 4.4) and hence the utility functions also increase. Generally, the improvement is milder for the sparse sampling. This is because a change from 68 pixels to 101 and 151 pixels is less effective than from 43 pixels in the contiguous case. Note that in the last case, where $k_{max} = 0.15h\mathrm{Mpc}^{-1}$, the utility functions have only a slight rise from the fiducial case even though there is more number of sampling pixels. This suggests that not much information lies in the non-linear regime that can help us measure the bins in the linear regime.

The errors are shown in Figure 4.15 for different number of pixels in the same k-range. The errors decrease for the larger number of pixels as expected.

4.5   Discussion

The importance of an accurate measurements of the galaxy power spectrum has been pointed out many times in this thesis. In this chapter we have investigated the optimisation of the galaxy surveys to improve the measurements of the galaxy power spectrum. The galaxy power spectrum on large scales probes structures that are less affected by clustering and therefore retain more memory about the initial state. Hence the information from these regimes are the cleanest since the Big Bang and any knowledge on these large scales would give us more information about the initial state of the Universe. In the case where we are mostly interested in the large scales, we might benefit more by investing our money to observe a larger, but sparsely sampled, area of the sky. In this chapter, by making use of Bayesian Experimental Design, we have investigated the advantages and disadvantages of the sparse sampling of sky as opposed to the contiguous sampling and ways of enhancing

the performance of the sparse sampling. Overall it seems like sparse sampling could be the way forward for future observations and measurements of the galaxy power spectrum. We have seen that sparse sampling improves the resolution of the power spectrum while measuring these bins more accurately. However, note that these results could be different if we were to measure a parameter other than the galaxy power spectrum itself.

We would still like to extend our analysis further; a project that we have begun to undertake is the investigation of the performance of the sparse sampling strategy in the BAO measurements and hence the measurements of the dark energy parameters.
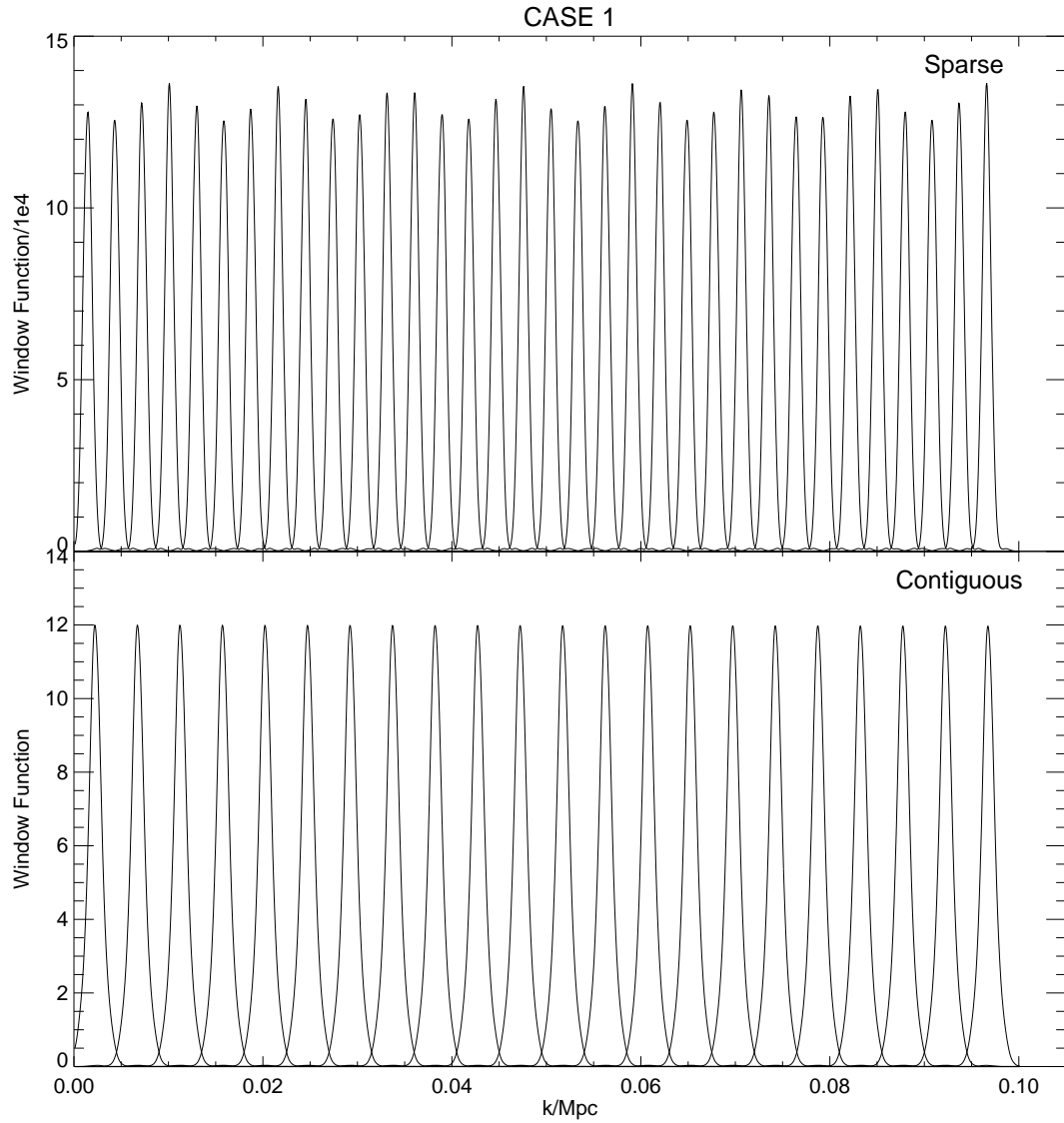
Figure 4.3: Diagonal elements of the window Function for the sparse and contiguous design for case 1. The sinusoidal feature in the sparse case is due to the cos term — see Equation 4.16.
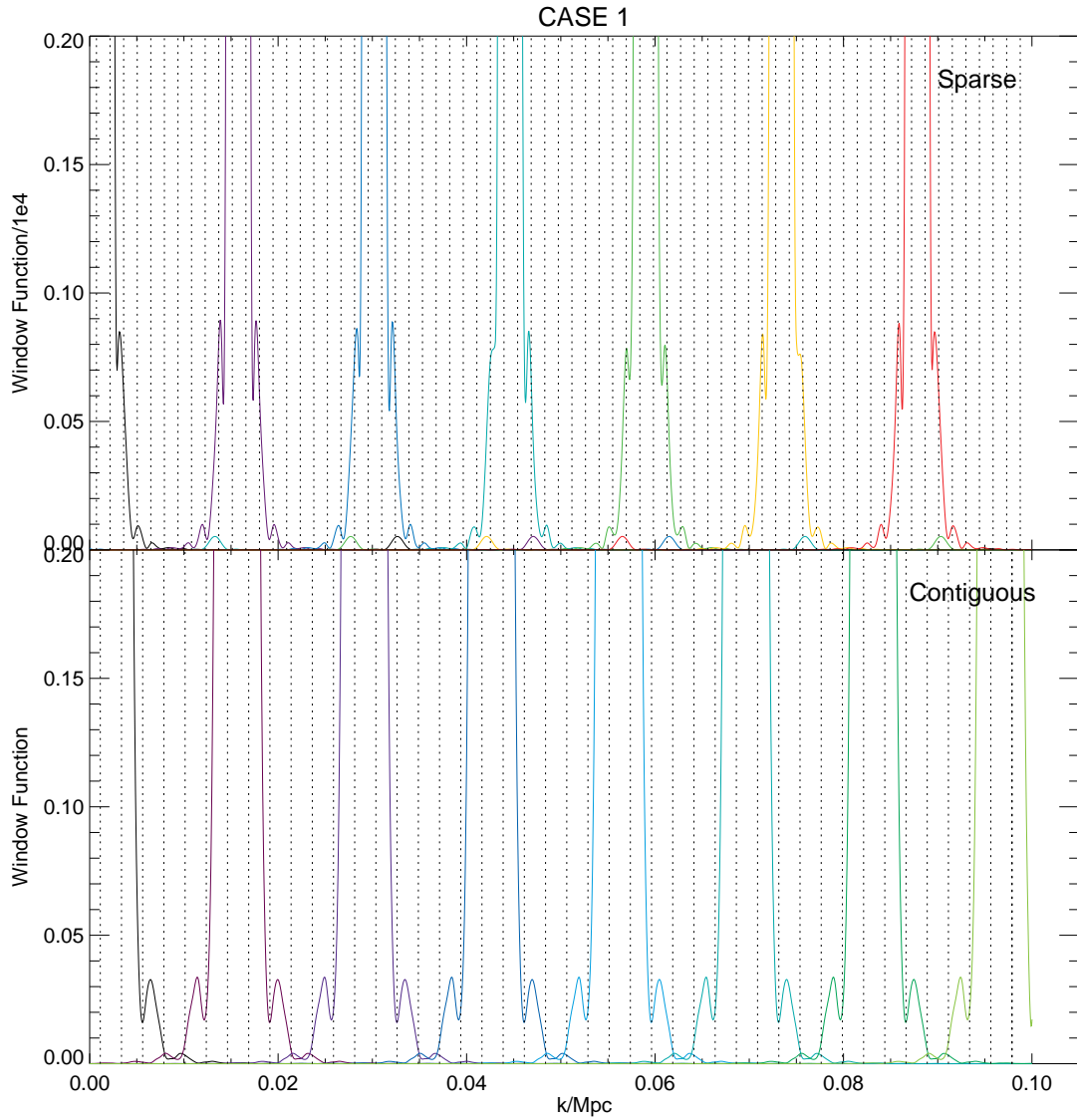
Figure 4.4: The same as Figure 4.3, where we have zoomed to compare the aliasing of scales between sparse and contiguous sampling. The dotted lines show the power spectrum bins.
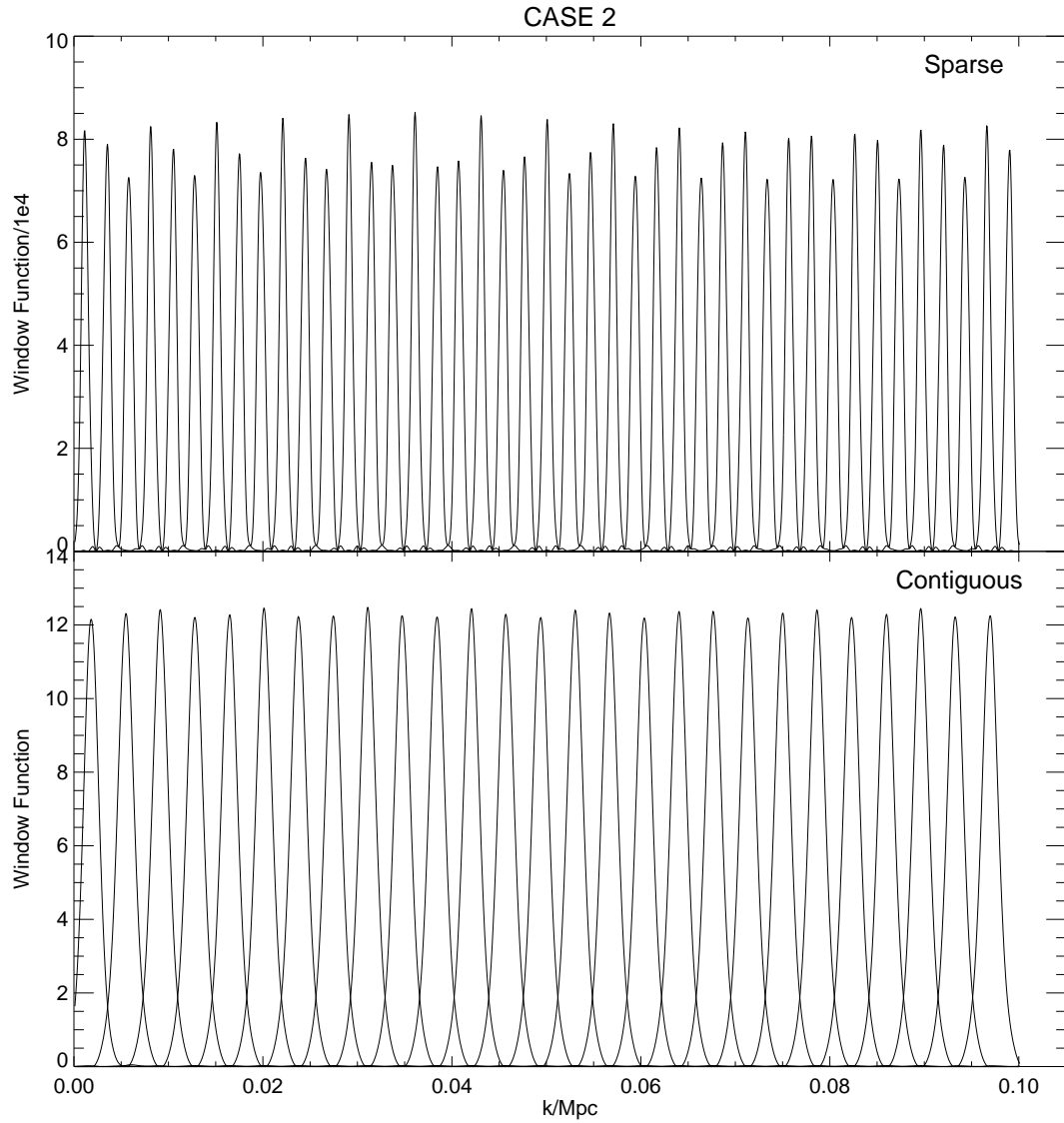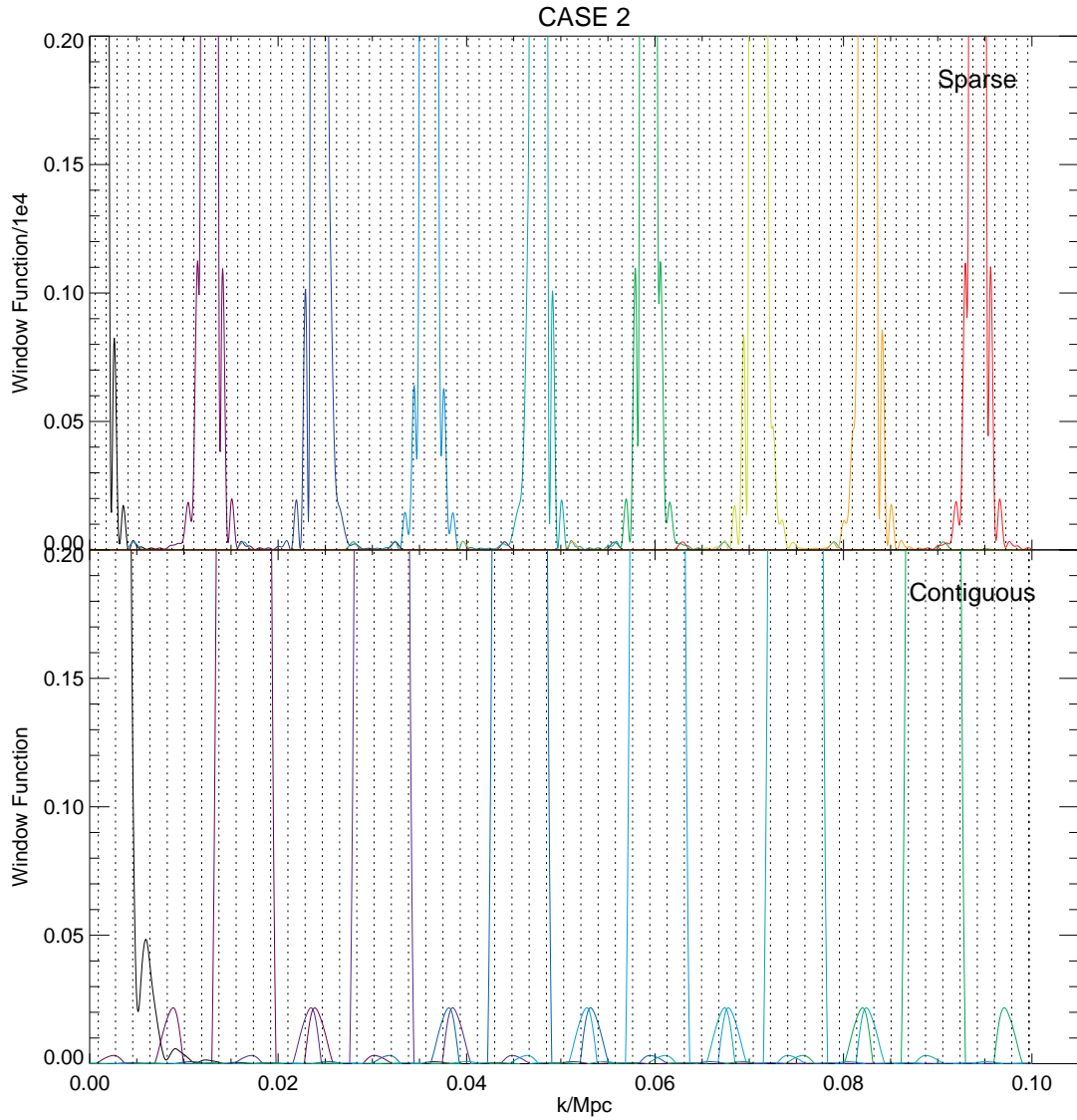
Figure 4.5: Diagonal elements of the window Function for the sparse and contiguous design for case 2. The sinusoidal feature in the sparse case is due to the cos term — see Equation 4.16.

Figure 4.6: The same as Figure 4.5, where we have zoomed to compare the aliasing of scales between sparse and contiguous sampling. The dotted lines show the power spectrum bins.

133

Figure 4.7: Diagonal elements of the window Function for the sparse and contiguous design for case 3. The sinusoidal feature in the sparse case is due to the cos term — see Equation 4.16.
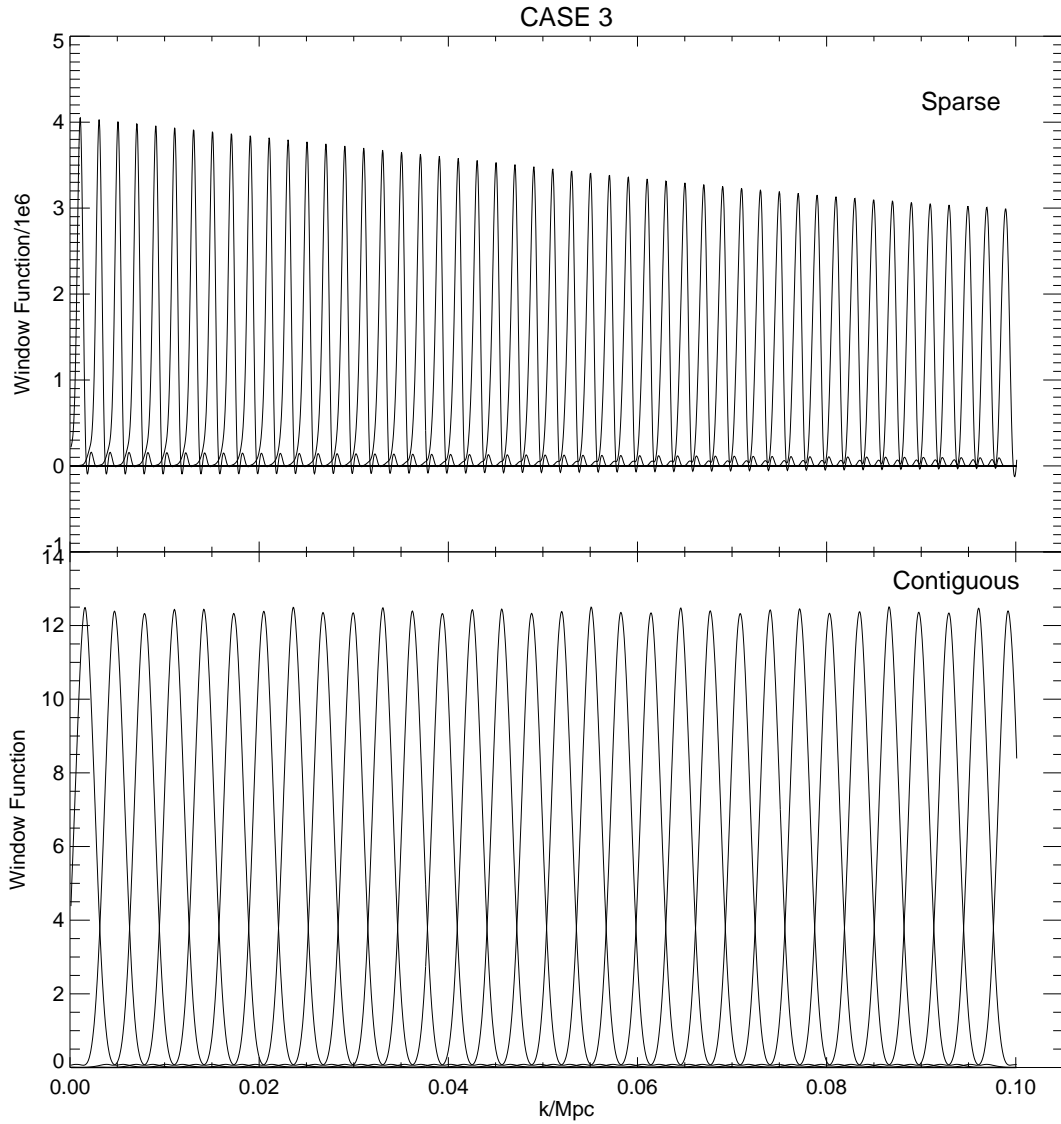
Figure 4.8: The same as Figure 4.7, where we have zoomed to compare the aliasing of scales between sparse and contiguous sampling. The dotted lines show the power spectrum bins.
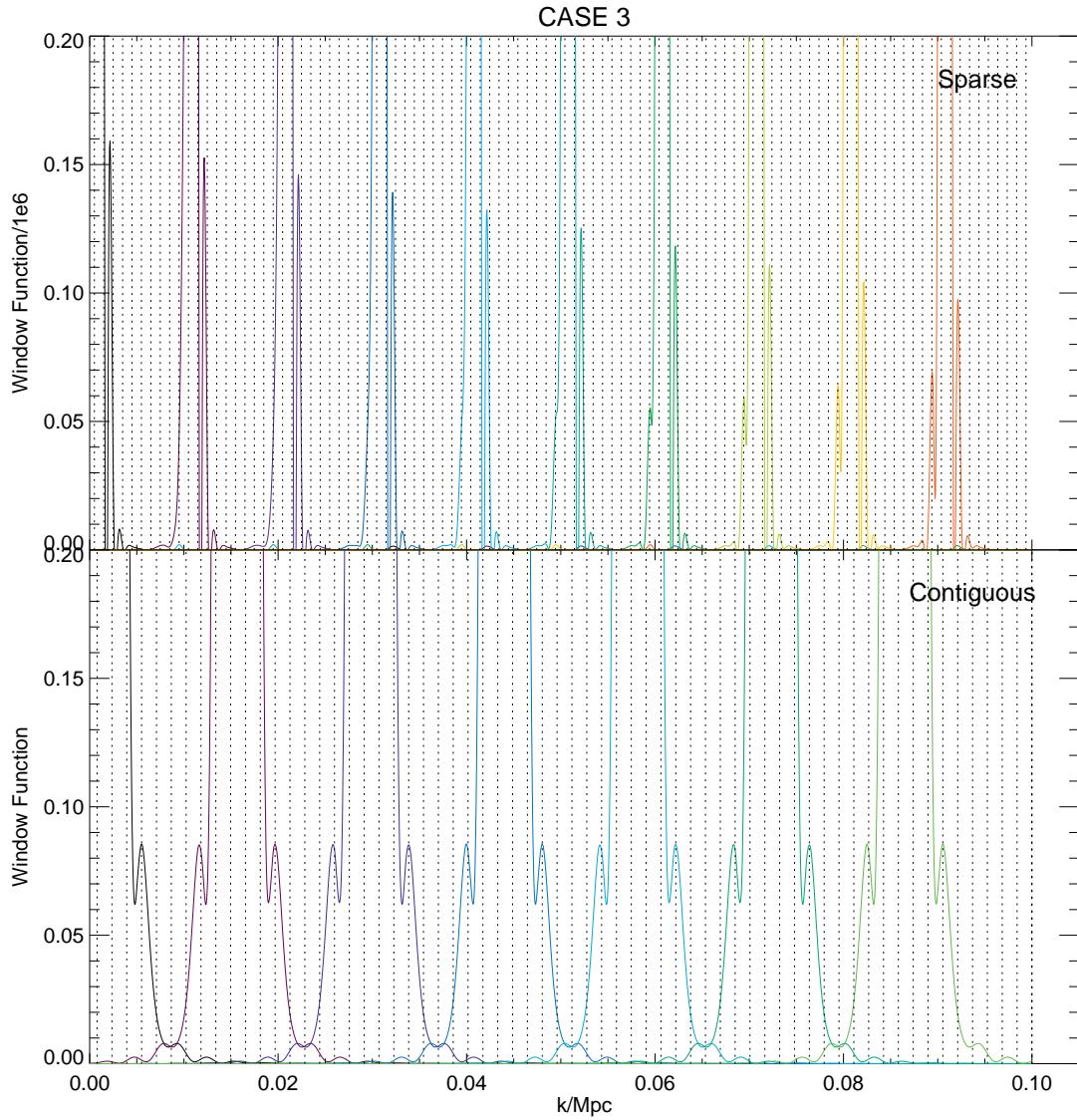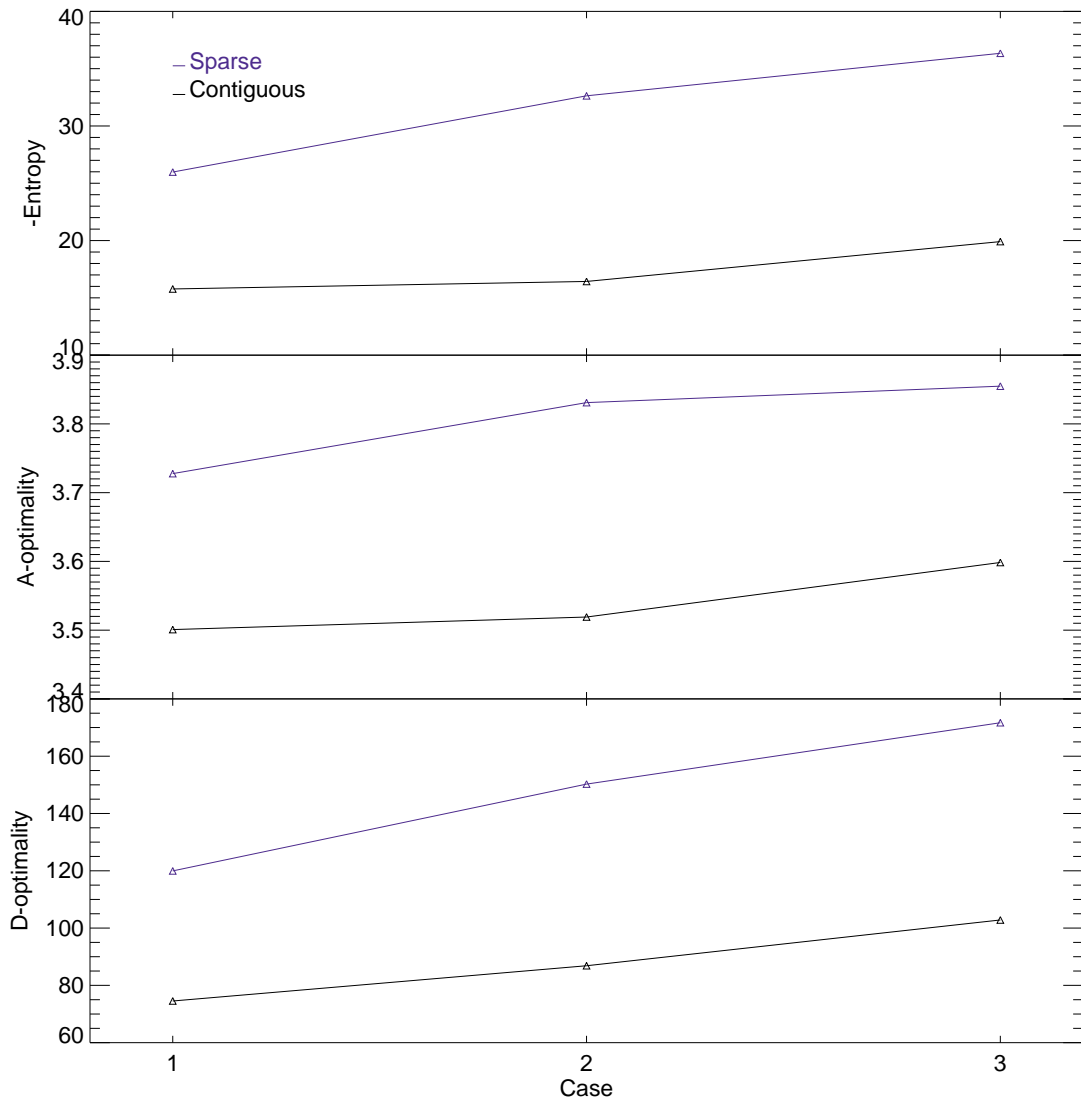
Figure 4.9: Utility functions for the three cases explained.
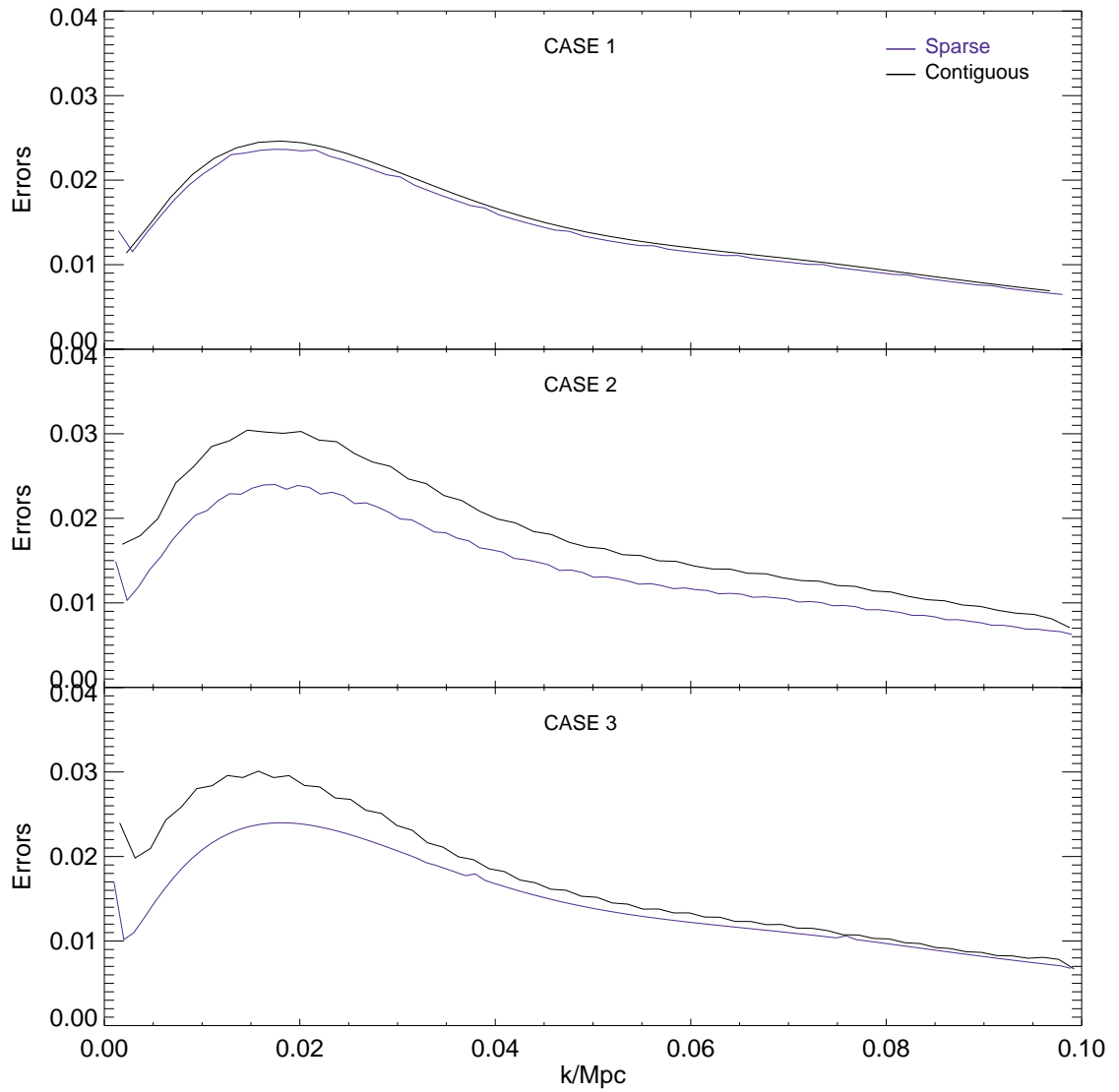
Figure 4.10: Marginalised errors for the three different cases, comparing sparse and contiguous designs in each case. Any features on $k < 0.005$ and $k > 0.095$ should be ignored due to edge effects.

137
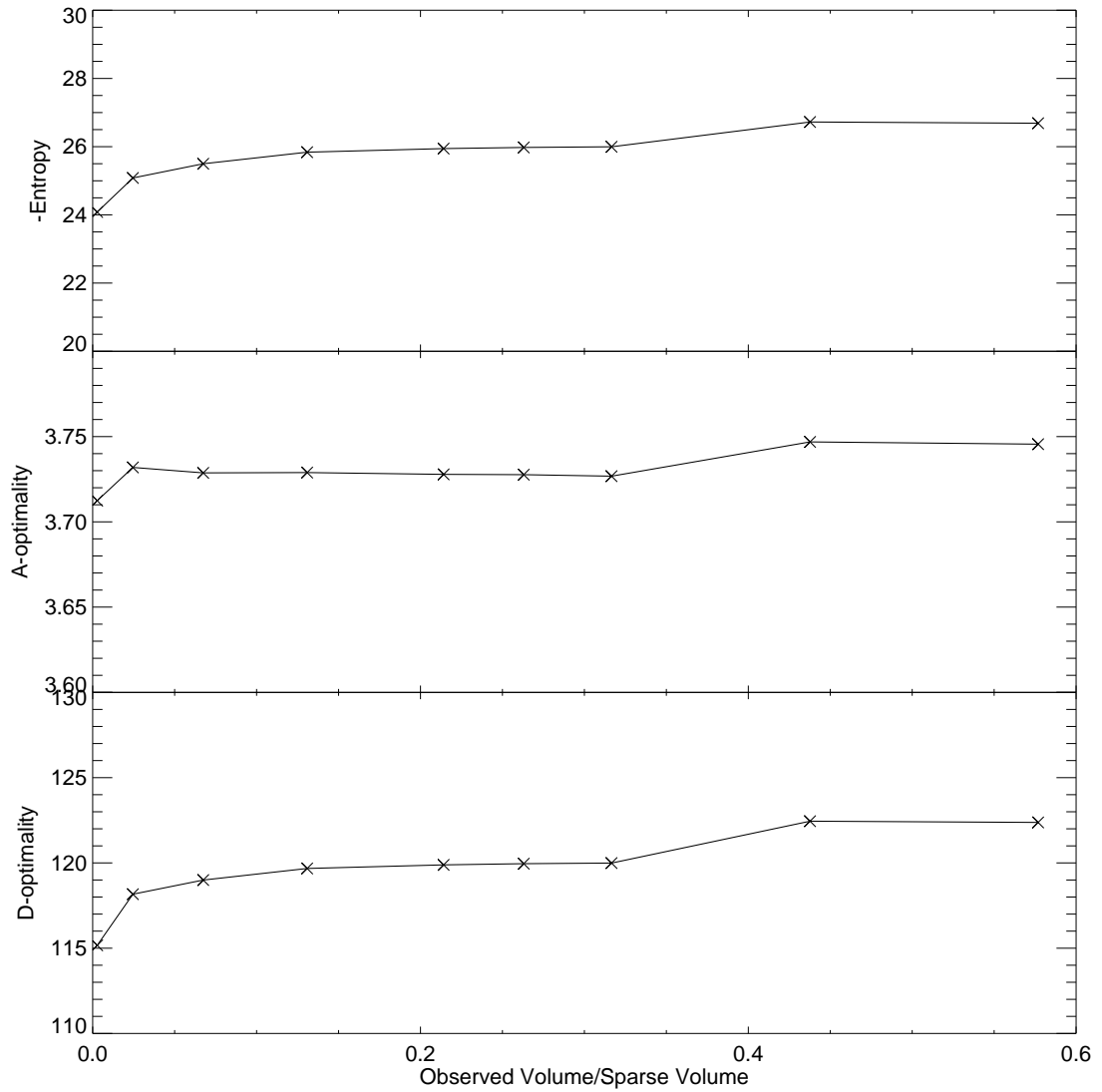
Figure 4.11: The variation of the utility functions as the step sizes vary in case 1 of the sparse sampling. This is shown against A_{obs}/A_{sparse}.
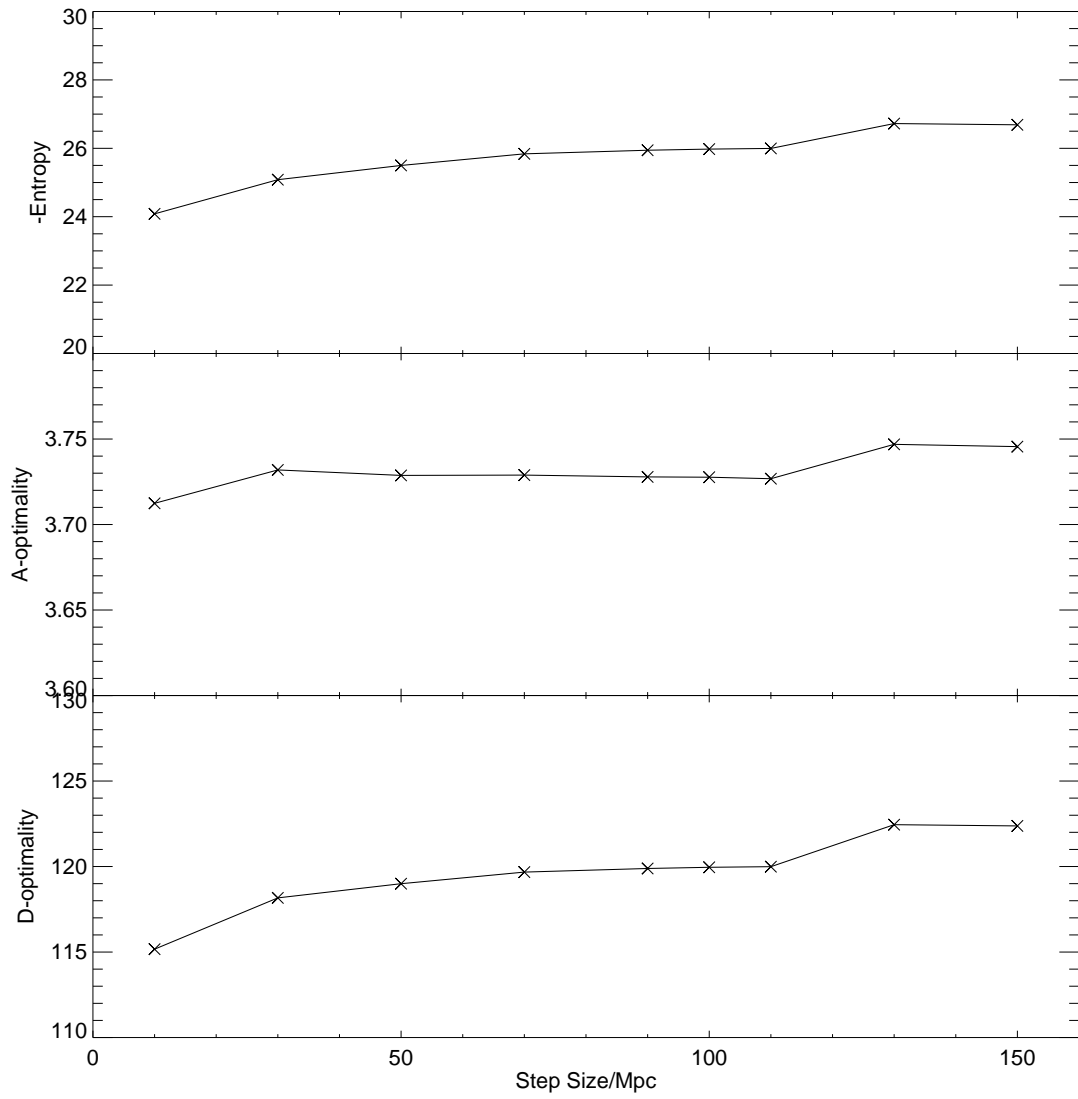
Figure 4.12: The same as Figure 4.11, but shown against the step sizes instead.
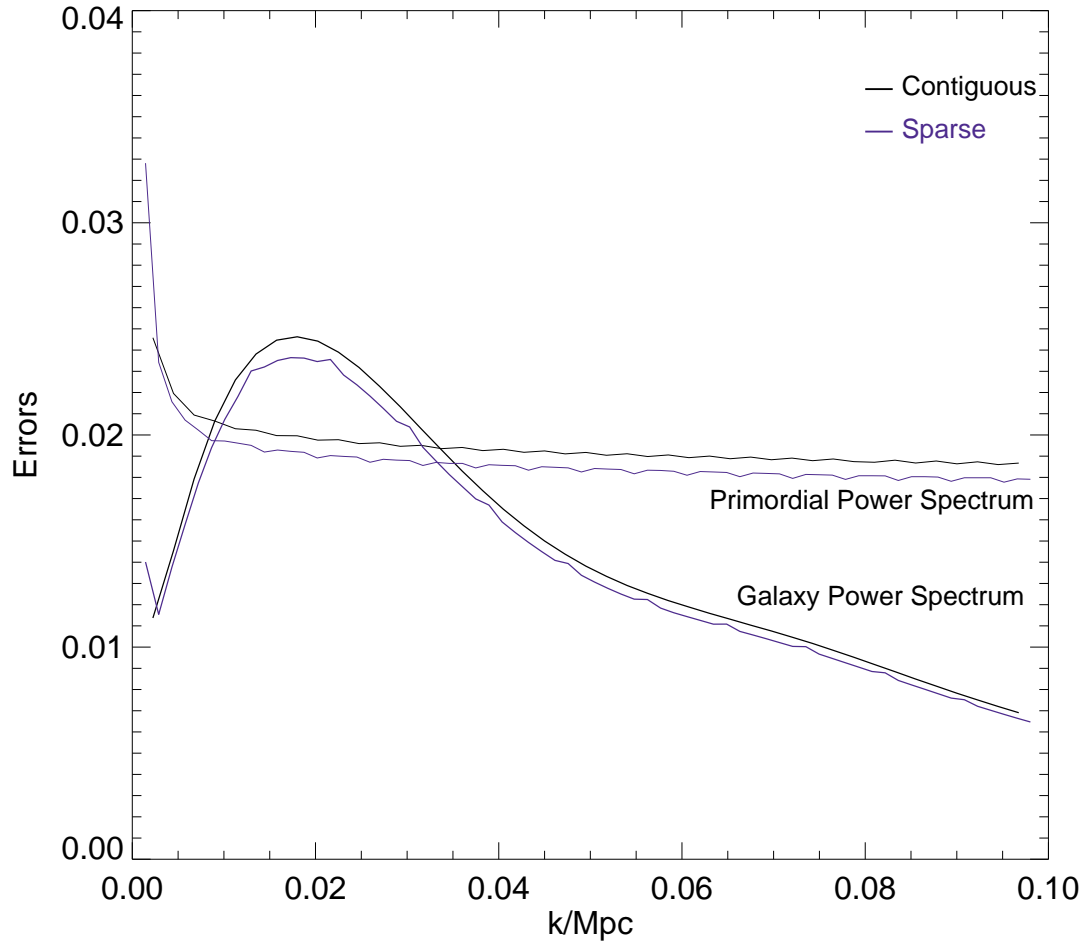
Figure 4.13: Marginalised errors shown for a galaxy power spectrum and a power law spectrum. Note that it is only that overall shape of this function is under investigation; the wiggles superimposed are a result of the numerical integration in IDL as this is a computationally intensive process.
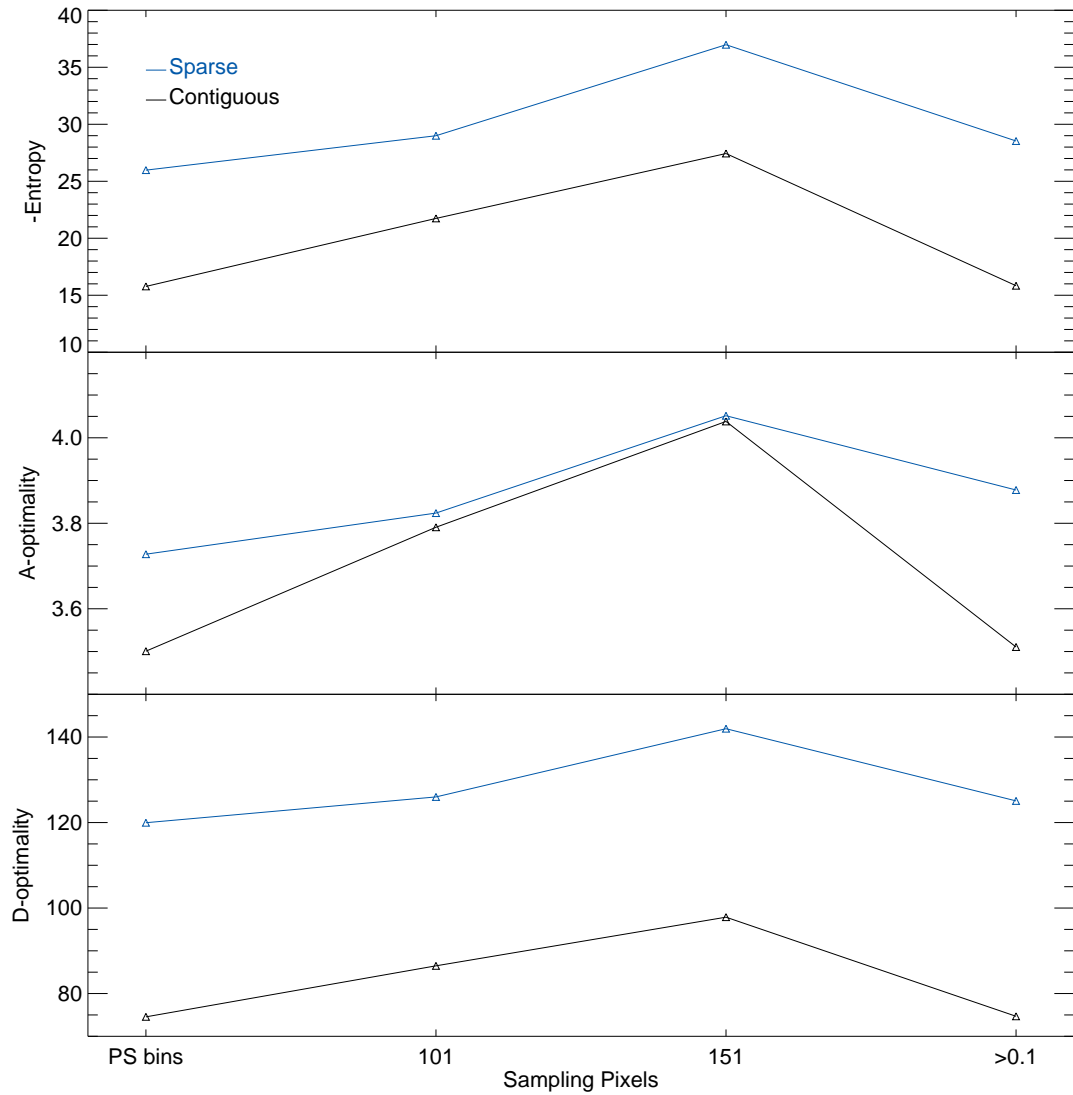
Figure 4.14: Utility functions for different sampling Fourier pixels. As the sampling pixels increases, surveys do a better job.
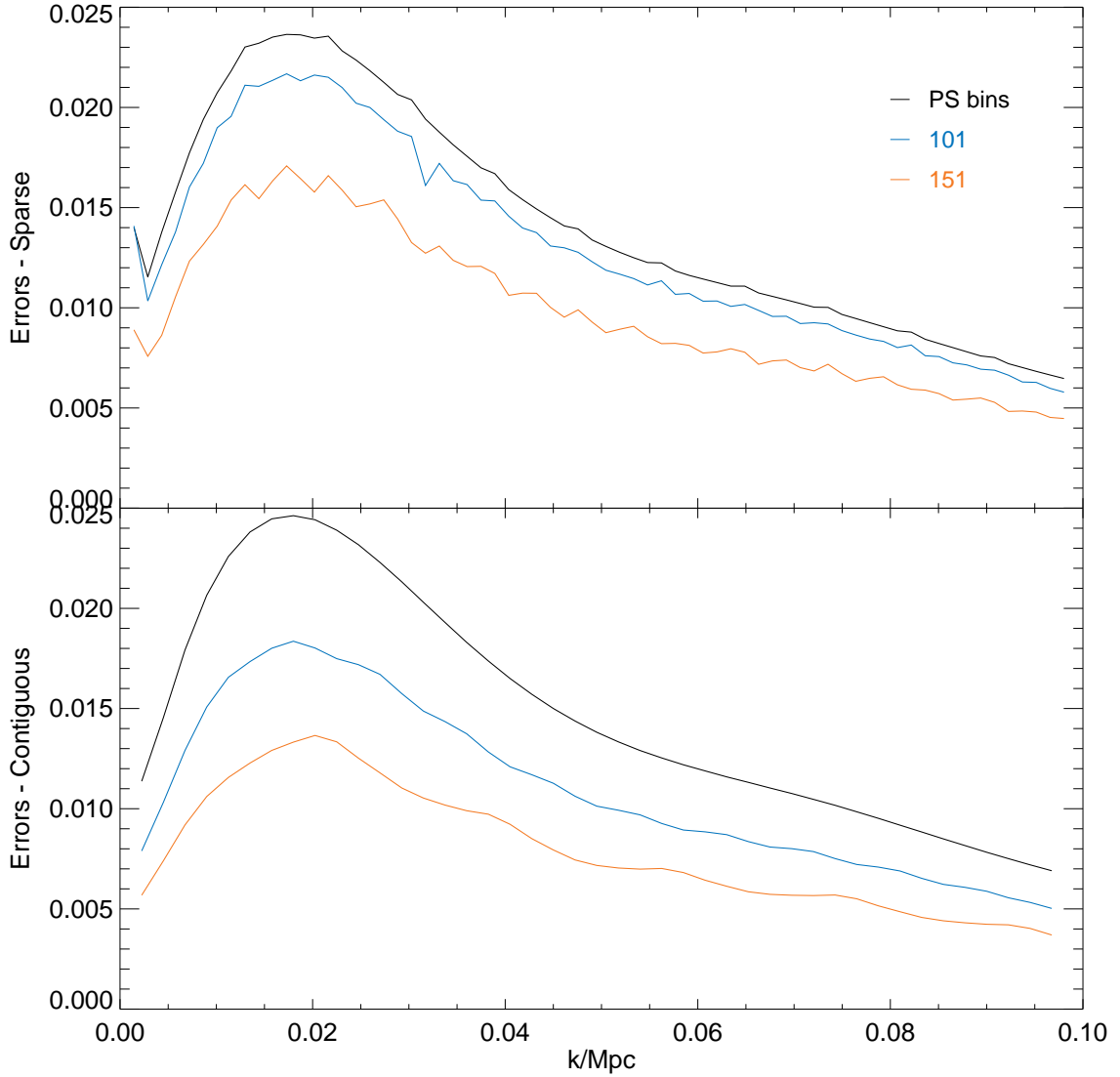
Figure 4.15: Marginalised errors shown for the different sampling Fourier pixels. As the sampling pixels increase, the errors measured on the parameters of interest decrease. Just like Figure 4.13, we only concentrate on the overall behaviour of the errors as function of k.

Chapter 5

Conclusions

We have presented work on different aspects of the measurement of the primordial and galaxy power spectra. The primordial power spectrum encodes precious information about the physics of the early Universe and constraining it is one of the key goals of modern cosmology. However, we cannot measure the primordial power spectrum directly and our path to its measurement is through different surveys (CMB, galaxy surveys, etc.). The outcome of such surveys is a convolution of the primordial power spectrum and a transfer function (which holds the cosmological parameters) — discussed in Chapter 1; galaxy power spectrum, CMB angular power spectrum, etc.. Hence the primordial power spectrum is reconstructed by deconvolving it from these power spectra: this shows the importance of the accurate measurements of the different types of power spectra. Also, the deconvolution induces degeneracy between the cosmological parameters determining the transfer functions and the primordial power spectrum.

In this thesis we first investigated the induced degeneracy between the cosmological parameters and the primordial power spectrum in Chapter 2. Then in Chapter 3 we presented a new method for deconvolving the primordial power spectrum from the WMAP5 CMB power spectrum. And finally in Chapter 4 we investigated the optimisation of galaxy surveys for the measurements of the galaxy power spectrum in the hope of measuring the primordial power spectrum more accurately.

In Chapter 2 we have investigated these degeneracies using the Planck and SDSS surveys. The induced degeneracy limits our ability to recover the primordial power spectrum, even from a perfect survey, especially in the case of CMB measurements [58]. We have constructed a parameter space containing a set of *carefully chosen* bins of the primordial power spectrum along with a set of cosmological parameters.

We then constructed a Fisher matrix for this parameter space for the two different surveys separately and combined. By diagonalising these Fisher matrices, via two different methods of eigenvector decomposition (PCA) and the Hermitian square root, we have investigated the induced correlation between the primordial power spectrum bins and the cosmological parameters. In the PCA case, we conclude that SDSS and Planck together measure the cosmological parameters to a better extent than each separately, and even break the degeneracy between them. More importantly for our purposes, they can increase the obtainable resolution of the primordial power spectrum by a factor of two and can also condense the correlation between bins to be only amongst neighbouring ones. By the use of the Hermitian square root of the Fisher matrix we managed to divert the correlation amongst the marginalised errors of the bins to the correlation between the bins themselves. In this case, the combination of SDSS and Planck helped to decrease the level of correlation between neighbouring bins, but also, because it has helped to increase the resolution of the bins, decreasing the correlation between neighbouring bins results in a correlation over a smaller range of k. We have shown that combining the two surveys will constrain the primordial power spectrum better than current measurements, and better than each experiment on its own.

In Chapter 3 we have employed the SVD process to reconstruct the primordial power spectrum directly from WMAP data. We have applied the method to forecasted Planck data and current WMAP5 results. These results allow for the consistent combination of a reconstruction method with a full exploration of the parameter likelihoods for the first time. We have seen that the limiting factor in constraining the primordial spectrum over a large range of wavenumbers k comes from the uncertainty in cosmological parameters. In the future as CMB polarisation data becomes increasingly accurate it will be desirable to perform a joint inversion of total intensity data along with polarisation data. It is not clear how to extend the SVD based method to include all polarisation modes simultaneously since a HOSVD (Higher-Order SVD) step would probably be required. On the other hand this would give the best estimate of P_k (the reconstructed primordial power spectrum) given any dataset and would help to reduce the correlations found in the reconstructed P_k by increasing the degrees of freedom that can be effectively constrained.

As thoroughly presented in the previous chapters, measurements of the primordial power spectrum heavily relies on the measurements of the galaxy/CMB power spectra. This is due to lack of a direct access to the primordial power spectrum measurements. In Chapter 4 we focus on galaxy surveys, and try to improve their

performance. In this Chapter we investigate the performance of the sparse sampling of sky to measure the galaxy power spectrum. This power spectrum on large scales probes structure that is less affected by clustering and therefore, has a memory of the initial state. In the case where we are mostly interested in the large scales, a sparse sampling design seems more advantageous. By making use of Bayesian Experimental Design, we have investigated the advantages and disadvantages of the sparse sampling of sky as opposed to contiguous sampling. Overall it seems like sparse sampling could be the way forward for future observations and measurements of the galaxy power spectrum. We have seen that sparse sampling enhances the resolution of the power spectrum while measuring the bins more accurately. However, these results may be totally different if we were to measure a parameter other than the galaxy power spectrum. We would still like to extend our analysis further; a project that we have begun to undertake is the investigation of the performance of the sparse sampling strategy in the BAO measurements and hence the measurements of the dark energy parameters.

References

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

[2] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. *APJ*, 396:L1–L5, September 1992.

[3] J. C. Mather, E. S. Cheng, D. A. Cottingham, R. E. Eplee, Jr., D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jensen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, S. H. Moseley, Jr., R. F. Silverberg, G. F. Smoot, R. Weiss, and D. T. Wilkinson. Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument. *APJ*, 420:439–444, January 1994.

[4] P. de Bernardis, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, K. Coble, B. P. Crill, G. De Gasperis, P. C. Farese, P. G. Ferreira, K. Ganga, M. Giacometti, E. Hivon, V. V. Hristov, A. Iacoangeli, A. H. Jaffe, A. E. Lange, L. Martinis, S. Masi, P. V. Mason, P. D. Mauskopf, A. Melchiorri, L. Miglio, T. Montroy, C. B. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, S. Prunet, S. Rao, G. Romeo, J. E. Ruhl, F. Scaramuzzi, D. Sforna, and N. Vittorio. A flat Universe from high-resolution maps of the cosmic microwave background radiation. *NAT*, 404:955–959, April 2000.

[5] R. A. Watson, P. Carreira, K. Cleary, R. D. Davies, R. J. Davis, C. Dickinson, K. Grainge, C. M. Gutiérrez, M. P. Hobson, M. E. Jones, R. Kneissl, A. Lasenby, K. Maisinger, G. G. Pooley, R. Rebolo, J. A. Rubiño-Martin,

B. Rusholme, R. D. E. Saunders, R. Savage, P. F. Scott, A. Slosar, P. J. Sosa Molina, A. C. Taylor, D. Titterington, E. Waldram, and A. Wilkinson. First results from the Very Small Array - I. Observational methods. *MNRAS*, 341:1057–1065, June 2003.

[6] E. M. Leitch, J. M. Kovac, N. W. Halverson, J. E. Carlstrom, C. Pryke, and M. W. E. Smith. Degree Angular Scale Interferometer 3 Year Cosmic Microwave Background Polarization Results. *APJ*, 624:10–20, May 2005.

[7] A. C. S. Readhead, B. S. Mason, C. R. Contaldi, T. J. Pearson, J. R. Bond, S. T. Myers, S. Padin, J. L. Sievers, J. K. Cartwright, M. C. Shepherd, D. Pogosyan, S. Prunet, P. Altamirano, R. Bustos, L. Bronfman, S. Casassus, W. L. Holzapfel, J. May, U.-L. Pen, S. Torres, and P. S. Udomprasert. Extended Mosaic Observations with the Cosmic Background Imager. *APJ*, 609:498–512, July 2004.

[8] G. Hinshaw et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations:Data Processing, Sky Maps, & Basic Results. 2008.

[9] *Planck: The scientific programme.* 2006.

[10] SPT Collaboration: J. E. Ruhl, P. A. R. Ade, J. E. Carlstrom, H. M. Cho, T. Crawford, M. Dobbs, C. H. Greer, N. W. Halverson, W. L. Holzapfel, T. M. Lantin, A. T. Lee, J. Leong, E. M. Leitch, W. Lu, M. Lueker, J. Mehl, S. S. Meyer, J. J. Mohr, S. Padin, T. Plagge, C. Pryke, D. Schwan, M. K. Sharp, M. C. Runyan, H. Spieler, Z. Staniszewski, and A. A. Stark. The south pole telescope. *PROC.SPIEINT.SOC.OPT.ENG*, 11, 2004.

[11] C. E. North, B. R. Johnson, P. A. R. Ade, M. D. Audley, C. Baines, R. A. Battye, M. L. Brown, P. Cabella, P. G. Calisse, A. D. Challinor, W. D. Duncan, P. G. Ferreira, W. K. Gear, D. Glowacka, D. J. Goldie, P. K. Grimes, M. Halpern, V. Haynes, G. C. Hilton, K. D. Irwin, M. E. Jones, A. N. Lasenby, P. J. Leahy, J. Leech, B. Maffei, P. Mauskopf, S. J. Melhuish, D. O'Dea, S. M. Parsley, L. Piccirillo, G. Pisano, C. D. Reintsema, G. Savini, R. Sudiwala, D. Sutton, A. C. Taylor, G. Teleberg, D. Titterington, V. Tsaneva, C. Tucker, R. Watson, S. Withington, G. Yassin, and J. Zhang. Detecting the B-mode Polarisation of the CMB with Clover. *ArXiv e-prints*, May 2008.

[12] D. Samtleben and for the QUIET collaboration. QUIET - Measuring the CMB polarization with coherent detector arrays. *ArXiv e-prints*, June 2008.

[13] S. J. Maddox, G. Efstathiou, W. J. Sutherland, and J. Loveday. The APM galaxy survey. I - APM measurements and star-galaxy separation. *MNRAS*, 243:692–712, April 1990.

[14] Scott M. Croom et al. The 2dF QSO Redshift Survey - XII. The spectroscopic catalogue and luminosity function. *Mon. Not. Roy. Astron. Soc.*, 349:1397, 2004.

[15] Jennifer K. Adelman-McCarthy et al. The Sixth Data Release of the Sloan Digital Sky Survey. *Astrophys. J. Suppl.*, 175:297–313, 2008.

[16] S. Dodelson. *Modern cosmology*. 2003.

[17] 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.

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

[19] H. Tadros, G. Efstathiou, and G. Dalton. The power spectrum of rich clusters of galaxies on large spatial scales. *MNRAS*, 296:995–1003, June 1998.

[20] J. R. Bond and G. Efstathiou. The statistics of cosmic background radiation fluctuations. *MNRAS*, 226:655–687, June 1987.

[21] U. Seljak and M. Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *APJ*, 469:437–+, October 1996.

[22] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. *APJS*, 170:377–408, June 2007.

[23] W. Hu and N. Sugiyama. Small-Scale Cosmological Perturbations: an Analytic Approach. *APJ*, 471:542–+, November 1996.

[24] J. Dunkley, E. Komatsu, M. R. Nolta, D. N. Spergel, D. Larson, G. Hinshaw, L. Page, C. L. Bennett, B. Gold, N. Jarosik, J. L. Weiland, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and

E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data. *ArXiv e-prints*, 803, March 2008.

[25] C. L. Reichardt, P. A. R. Ade, J. J. Bock, J. R. Bond, J. A. Brevik, C. R. Contaldi, M. D. Daub, J. T. Dempsey, J. H. Goldstein, W. L. Holzapfel, C. L. Kuo, A. E. Lange, M. Lueker, M. Newcomb, J. B. Peterson, J. Ruhl, M. C. Runyan, and Z. Staniszewski. High-Resolution CMB Power Spectrum from the Complete ACBAR Data Set. *APJ*, 694:1200–1219, April 2009.

[26] T. E. Montroy, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, P. Cabella, C. R. Contaldi, B. P. Crill, P. de Bernardis, G. De Gasperis, A. de Oliveira-Costa, G. De Troia, G. di Stefano, E. Hivon, A. H. Jaffe, T. S. Kisner, W. C. Jones, A. E. Lange, S. Masi, P. D. Mauskopf, C. J. Mac-Tavish, A. Melchiorri, P. Natoli, C. B. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, G. Polenta, S. Prunet, S. Ricciardi, G. Romeo, J. E. Ruhl, P. Santini, M. Tegmark, M. Veneziani, and N. Vittorio. A Measurement of the CMB EE Spectrum from the 2003 Flight of BOOMERANG. *APJ*, 647:813–822, August 2006.

[27] J. L. Sievers, C. Achermann, J. R. Bond, L. Bronfman, R. Bustos, C. R. Contaldi, C. Dickinson, P. G. Ferreira, M. E. Jones, A. M. Lewis, B. S. Mason, J. May, S. T. Myers, N. Oyarce, S. Padin, T. J. Pearson, M. Pospieszalski, A. C. S. Readhead, R. Reeves, A. C. Taylor, and S. Torres. Implications of the Cosmic Background Imager Polarization Data. *APJ*, 660:976–987, May 2007.

[28] C. Bischoff, L. Hyatt, J. J. McMahon, G. W. Nixon, D. Samtleben, K. M. Smith, K. Vanderlinde, D. Barkats, P. Farese, T. Gaier, J. O. Gundersen, M. M. Hedman, S. T. Staggs, and B. Winstein. New Measurements of Fine-Scale CMB Polarization Power Spectra from CAPMAP at Both 40 and 90 GHz. *APJ*, 684:771–789, September 2008.

[29] P. G. Castro, P. Ade, J. Bock, M. Bowden, M. L. Brown, G. Cahill, S. Church, T. Culverhouse, R. B. Friedman, K. Ganga, W. K. Gear, S. Gupta, J. Hinderks, J. Kovac, A. E. Lange, E. Leitch, S. J. Melhuish, Y. Memari, J. A. Murphy, A. Orlando, C. Pryke, R. Schwarz, C. O'Sullivan, L. Piccirillo, N. Rajguru, B. Rusholme, A. N. Taylor, K. L. Thompson, A. H. Turner, E. Y. S. Wu, M. Zemcov, and The QUa D collaboration. Cosmological Parameters

from the QUAD CMB Polarization Experiment. *APJ*, 701:857–864, August 2009.

[30] M. Bartelmann and P. Schneider. Weak gravitational lensing. *PHYSREP*, 340:291–472, January 2001.

[31] M. White. Baryons and weak lensing power spectra. *Astroparticle Physics*, 22:211–217, November 2004.

[32] 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*, 148:1, 2003.

[33] J. Gunn and D. Weinberg. The Sloan Digital Sky Survey. In S. J. Maddox and A. Aragon-Salamanca, editors, *Wide Field Spectroscopy and the Distant Universe*, pages 3–+, 1995.

[34] K. Taylor. The Anglo-Australian Telescope's 2dF Facility: Progress and Predicted Performance. In S. J. Maddox and A. Aragon-Salamanca, editors, *Wide Field Spectroscopy and the Distant Universe*, pages 15–+, 1995.

[35] M. Bersanelli, G. F. Smoot, M. Bensadoun, G. de Amici, M. Limon, and S. Levin. Measurements of the CMB Spectrum at Centimeter Wavelengths. *Astrophysical Letters Communications*, 32:7–+, 1995.

[36] M. Tegmark. Measuring Cosmological Parameters with Galaxy Surveys. *Physical Review Letters*, 79:3806–3809, November 1997.

[37] J. R. Bond, A. H. Jaffe, and L. Knox. Estimating the power spectrum of the cosmic microwave background. *PRD*, 57:2117–2137, February 1998.

[38] L. Knox. Determination of inflationary observables by cosmic microwave background anisotropy experiments. *PRD*, 52:4307–4318, October 1995.

[39] F. Bouchet, R. Gispert, and J. L. Puget. The COBRAS/SAMBA CMB project. In S. Maurogordato, C. Balkowski, C. Tao, and J. Tran Thanh van, editors, *Clustering in the Universe*, pages 537–+, 1995.

[40] J. Delabrouille, J. . Puget, R. Gispert, and J. . Lamarre. Scanning strategies for the Planck mission. *ArXiv Astrophysics e-prints*, October 1998.

[41] M. P. Hobson, A. W. Jones, A. N. Lasenby, and F. R. Bouchet. Foreground separation methods for satellite observations of the cosmic microwave background. *MNRAS*, 300:1–29, October 1998.

[42] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay. Measuring the Galaxy Power Spectrum with Future Redshift Surveys. *APJ*, 499:555–+, May 1998.

[43] A. F. Heavens and A. N. Taylor. Design and analysis of redshift surveys. *MNRAS*, 290:456–464, September 1997.

[44] M. Zaldarriaga, D. N. Spergel, and U. Seljak. Microwave Background Constraints on Cosmological Parameters. *APJ*, 488:1–+, October 1997.

[45] J. R. Bond, G. Efstathiou, and M. Tegmark. Forecasting cosmic parameter errors from microwave background anisotropy experiments. *MNRAS*, 291:L33–L41, November 1997.

[46] G. Efstathiou and J. R. Bond. Cosmic confusion: degeneracies among cosmological parameters derived from measurements of microwave background anisotropies. *MNRAS*, 304:75–97, March 1999.

[47] M. Zaldarriaga, D. N. Spergel, and U. Seljak. Microwave Background Constraints on Cosmological Parameters. *APJ*, 488:1–+, October 1997.

[48] G. Efstathiou and J. R. Bond. Cosmic confusion: degeneracies among cosmological parameters derived from measurements of microwave background anisotropies. *MNRAS*, 304:75–97, March 1999.

[49] D. Scott, J. Silk, and M. White. From Microwave Anisotropies to Cosmology. *Science*, 268:829–835, May 1995.

[50] J. R. Bond and A. H. Jaffe. Cosmic Parameter Estimation Combining Sub-Degree CMB Experiment with COBE. In F. R. Bouchet, R. Gispert, B. Guiderdoni, and J. Tran Thanh van, editors, *Microwave Background Anistropies*, pages 197–+, 1997.

[51] C. H. Lineweaver. The Cosmic Microwave Background and Observational Convergence in the Omega_m-Omega_Lambda Plane. *APJL*, 505:L69–L73, October 1998.

[52] E. Gawiser and J. Silk. Extracting Primordial Density Fluctuations. *Science*, 280:1405–+, May 1998.

[53] D. J. Eisenstein, W. Hu, and M. Tegmark. Cosmic Complementarity: Joint Parameter Estimation from Cosmic Microwave Background Experiments and Redshift Surveys. *APJ*, 518:2–23, June 1999.

[54] L. Perotto, J. Lesgourgues, S. Hannestad, H. Tu, and Y. Y Y Wong. Probing cosmological parameters with the CMB: forecasts from Monte Carlo simulations. *Journal of Cosmology and Astro-Particle Physics*, 10:13–+, October 2006.

[55] P. J. Steinhardt and N. Turok. A Cyclic Model of the Universe. *Science*, 296:1436–1439, May 2002.

[56] W. H. Kinney. How to fool cosmic microwave background parameter estimation. *PRD*, 63(4):043001–+, February 2001.

[57] T. Souradeep, J. R. Bond, L. Knox, G. Efstathiou, and M. S. Turner. Prospects for Measuring Inflation Parameters with the CMB. In L. Roszkowski, editor, *COSMO-97, First International Workshop on Particle Physics and the Early Universe*, pages 389–+, 1998.

[58] W. Hu and T. Okamoto. Principal power of the CMB. *PRD*, 69(4):043004–+, February 2004.

[59] S. L. Bridle, A. M. Lewis, J. Weller, and G. Efstathiou. Reconstructing the primordial power spectrum. *MNRAS*, 342:L72–L78, July 2003.

[60] A. Niarchou and A. Jaffe. Imprints of spherical non-trivial topologies on the CMB. *ArXiv Astrophysics e-prints*, February 2007.

[61] G. Efstathiou. Is the low cosmic microwave background quadrupole a signature of spatial curvature? *MNRAS*, 343:L95–L98, August 2003.

[62] C. R. Contaldi, M. Peloso, L. Kofman, and A. Linde. Suppressing the lower multipoles in the CMB anisotropies. *Journal of Cosmology and Astro-Particle Physics*, 7:2–+, July 2003.

[63] P. Mukherjee and Y. Wang. Model-independent Reconstruction of the Primordial Power Spectrum from Wilkinson Microwave Anistropy Probe Data. *APJ*, 599:1–6, December 2003.

[64] P. Mukherjee and Y. Wang. Wavelet Band Powers of the Primordial Power Spectrum from Cosmic Microwave Background Data. *APJ*, 593:38–47, August 2003.

[65] Y. Wang, D. N. Spergel, and M. A. Strauss. Cosmology in the Next Millennium: Combining Microwave Anisotropy Probe and Sloan Digital Sky Survey Data to Constrain Inflationary Models. *APJ*, 510:20–31, January 1999.

[66] G. Hinshaw, C. L. Bennett, and A. Kogut. Simulated cosmic microwave background maps at 0.5 deg resolution: Basic results. *APJL*, 441:L1–L4, March 1995.

[67] 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.

[68] L. Verde and H. Peiris. On minimally parametric primordial power spectrum reconstruction and the evidence for a red tilt. *Journal of Cosmology and Astro-Particle Physics*, 7:9–+, July 2008.

[69] R. Trotta. Forecasting the Bayes factor of a future observation. *MNRAS*, 378:819–824, July 2007.

[70] D. Parkinson, C. Blake, M. Kunz, B. A. Bassett, R. C. Nichol, and K. Glazebrook. Optimizing baryon acoustic oscillation surveys - I. Testing the concordance ΛCDM cosmology. *MNRAS*, 377:185–197, May 2007.

[71] R. J. Scherrer and D. H. Weinberg. Constraints on the Effects of Locally Biased Galaxy Formation. *APJ*, 504:607–+, September 1998.

[72] G. Hütsi. Acoustic oscillations in the SDSS DR4 luminous red galaxy sample power spectrum. *AAP*, 449:891–902, April 2006.

[73] R. G. Mann, J. A. Peacock, and A. F. Heavens. Eulerian bias and the galaxy density field. *MNRAS*, 293:209–+, January 1998.

[74] U. Seljak and M. S. Warren. Large-scale bias and stochasticity of haloes and dark matter. *MNRAS*, 355:129–136, November 2004.

[75] A. Meiksin, M. White, and J. A. Peacock. Baryonic Signatures in Large-Scale Structure. In J. Paul, T. Montmerle, and E. Aubourg, editors, *Abstracts of the 19th Texas Symposium on Relativistic Astrophysics and Cosmology, held in Paris, France, Dec. 14-18, 1998. Eds.: J. Paul, T. Montmerle, and E. Aubourg (CEA Saclay), meeting abstract.*, December 1998.

[76] W. J Percival and M. White. Testing cosmological structure formation using redshift-space distortions. *ArXiv e-prints*, July 2008.

[77] A. J. S. Hamilton. Towards optimal measurement of power spectra - I. Minimum variance pair weighting and the Fisher matrix. *MNRAS*, 289:285–294, August 1997.

[78] A. J. S. Hamilton. Towards optimal measurement of power spectra - II. A basis of positive, compact, statistically orthogonal kernels. *MNRAS*, 289:295–304, August 1997.

[79] Amit P. S. Yadav and Benjamin D. Wandelt. Evidence of Primordial Non-Gaussianity (f_{NL}) in the Wilkinson Microwave Anisotropy Probe 3-Year Data at 2.8σ. *Phys. Rev. Lett.*, 100:181301, 2008.

[80] Kendrick M. Smith, Leonardo Senatore, and Matias Zaldarriaga. Optimal limits on $f_N L^l ocal$ from WMAP 5-year data. 2009.

[81] Alexei A. Starobinsky. Spectrum of adiabatic perturbations in the universe when there are singularities in the inflation potential. *JETP Lett.*, 55:489–494, 1992.

[82] Jennifer A. Adams, Bevan Cresswell, and Richard Easther. Inflationary perturbations from a potential with a step. *Phys. Rev.*, D64:123514, 2001.

[83] Xiulian Wang, Bo Feng, Mingzhe Li, Xue-Lei Chen, and Xinmin Zhang. Natural inflation, Planck scale physics and oscillating primordial spectrum. *Int. J. Mod. Phys.*, D14:1347, 2005.

[84] Paul Hunt and Subir Sarkar. Multiple inflation and the WMAP 'glitches'. *Phys. Rev.*, D70:103518, 2004.

[85] Minu Joy, Varun Sahni, and Alexei A. Starobinsky. A New Universal Local Feature in the Inflationary Perturbation Spectrum. *Phys. Rev.*, D77:023514, 2008.

[86] Paul Hunt and Subir Sarkar. Multiple inflation and the WMAP 'glitches' II. Data analysis and cosmological parameter extraction. *Phys. Rev.*, D76:123504, 2007.

[87] Cedric Pahud, Marc Kamionkowski, and Andrew R Liddle. Oscillations in the inflaton potential? 2008.

[88] Rose Lerner and John McDonald. Space-Dependent Step Features: Transient Breakdown of Slow-roll, Homogeneity and Isotropy During Inflation. *Phys. Rev.*, D79:023511, 2009.

[89] Carlo R. Contaldi, Marco Peloso, Lev Kofman, and Andrei Linde. Suppressing the lower Multipoles in the CMB Anisotropies. *JCAP*, 0307:002, 2003.

[90] Brian A. Powell and William H. Kinney. The pre-inflationary vacuum in the cosmic microwave background. *Phys. Rev.*, D76:063512, 2007.

[91] Gavin Nicholson and Carlo R. Contaldi. The large scale CMB cut-off and the tensor-to-scalar ratio. *JCAP*, 0801:002, 2008.

[92] Julien Lesgourgues. Features in the primordial power spectrum of double D-term inflation. *Nucl. Phys.*, B582:593–626, 2000.

[93] Bo Feng and Xinmin Zhang. Double inflation and the low CMB quadrupole. *Phys. Lett.*, B570:145–150, 2003.

[94] Grant J. Mathews, D. J. H. Chung, K. Ichiki, T. Kajino, and M. Orito. Constraints on resonant particle production during inflation from the matter and CMB power spectra. *Phys. Rev.*, D70:083505, 2004.

[95] Rajeev Kumar Jain, Pravabati Chingangbam, Jinn-Ouk Gong, L. Sriramkumar, and Tarun Souradeep. Double inflation and the low CMB multipoles. *JCAP*, 0901:009, 2009.

[96] Antonio Enea Romano and Misao Sasaki. Effects of particle production during inflation. *Phys. Rev.*, D78:103522, 2008.

[97] Rajeev Kumar Jain, Pravabati Chingangbam, L. Sriramkumar, and Tarun Souradeep. The tensor-to-scalar ratio in punctuated inflation. 2009.

[98] Gavin Nicholson and Carlo R. Contaldi. Reconstruction of the Primordial Power Spectrum using Temperature and Polarisation Data from Multiple Experiments. 2009.

[99] S. L. Bridle, A. M. Lewis, J. Weller, and G. Efstathiou. Reconstructing the primordial power spectrum. *Mon. Not. Roy. Astron. Soc.*, 342:L72, 2003.

[100] David Parkinson, Shinji Tsujikawa, Bruce A. Bassett, and Luca Amendola. Testing for double inflation with WMAP. *Phys. Rev.*, D71:063524, 2005.

[101] Rita Sinha and Tarun Souradeep. Post-WMAP assessment of infrared cutoff in the primordial spectrum from inflation. *Phys. Rev.*, D74:043518, 2006.

[102] Carolyn Sealfon, Licia Verde, and Raul Jimenez. Smoothing spline primordial power spectrum reconstruction. *Phys. Rev.*, D72:103520, 2005.

[103] Pia Mukherjee and Yun Wang. Primordial Power Spectrum Reconstruction. *JCAP*, 0512:007, 2005.

[104] Michael Bridges, A. N. Lasenby, and M. P. Hobson. A Bayesian analysis of the primordial power spectrum. *Mon. Not. Roy. Astron. Soc.*, 369:1123–1130, 2006.

[105] Michael Bridges, A. N. Lasenby, and M. P. Hobson. WMAP 3-year primordial power spectrum. 2006.

[106] Laura Covi, Jan Hamann, Alessandro Melchiorri, Anze Slosar, and Irene Sorbera. Inflation and WMAP three year data: Features have a future! *Phys. Rev.*, D74:083509, 2006.

[107] Minu Joy, Arman Shafieloo, Varun Sahni, and Alexei A. Starobinsky. Is a step in the primordial spectral index favored by CMB data ? 2008.

[108] Licia Verde and Hiranya V. Peiris. On Minimally-Parametric Primordial Power Spectrum Reconstruction and the Evidence for a Red Tilt. *JCAP*, 0807:009, 2008.

[109] Steen Hannestad. Reconstructing the inflationary power spectrum from CMBR data. *Phys. Rev.*, D63:043009, 2001.

[110] Yun Wang and Grant Mathews. A measurement of the primordial power spectrum from Maxima and Boomerang data. *Astrophys. J.*, 573:1, 2002.

[111] Makoto Matsumiya, Misao Sasaki, and Jun'ichi Yokoyama. Cosmic inversion: Reconstructing primordial spectrum from CMB anisotropy. *Phys. Rev.*, D65:083007, 2002.

[112] Arman Shafieloo and Tarun Souradeep. Primordial power spectrum from WMAP. *Phys. Rev.*, D70:043523, 2004.

[113] Noriyuki Kogo, Makoto Matsumiya, Misao Sasaki, and Jun'ichi Yokoyama. Reconstructing the primordial spectrum from WMAP data by the cosmic inversion method. *Astrophys. J.*, 607:32–39, 2004.

[114] Pia Mukherjee and Yun Wang. Wavelet Band Powers of the Primordial Power Spectrum from CMB Data. *Astrophys. J.*, 593:38, 2003.

[115] Pia Mukherjee and Yun Wang. Model-Independent Reconstruction of the Primordial Power Spectrum from WMAP Data. *Astrophys. J.*, 599:1–6, 2003.

[116] Steen Hannestad. Reconstructing the primordial power spectrum - a new algorithm. *JCAP*, 0404:002, 2004.

[117] Noriyuki Kogo, Misao Sasaki, and Jun'ichi Yokoyama. Reconstructing the Primordial Spectrum with CMB Temperature and Polarization. *Phys. Rev.*, D70:103001, 2004.

[118] Domenico Tocchini-Valentini, Marian Douspis, and Joseph Silk. Are there features in the primordial power spectrum? *Mon. Not. Roy. Astron. Soc.*, 359:31–35, 2005.

[119] Samuel M. Leach. Measuring the primordial power spectrum: Principal component analysis of the cosmic microwave background. *Mon. Not. Roy. Astron. Soc.*, 372:646–654, 2006.

[120] Arman Shafieloo, Tarun Souradeep, P. Manimaran, Prasanta K. Panigrahi, and Raghavan Rangarajan. Features in the Primordial Spectrum from WMAP: A Wavelet Analysis. *Phys. Rev.*, D75:123502, 2007.

[121] Arman Shafieloo and Tarun Souradeep. Estimation of Primordial Spectrum with post-WMAP 3 year data. *Phys. Rev.*, D78:023511, 2008.

[122] Ryo Nagata and Jun'ichi Yokoyama. Reconstruction of the primordial fluctuation spectrum from the five-year WMAP data by the cosmic inversion method with band-power decorrelation analysis. *Phys. Rev.*, D78:123002, 2008.

[123] Ryo Nagata and Jun'ichi Yokoyama. Band-power reconstruction of the primordial fluctuation spectrum by the maximum likelihood reconstruction method. *Phys. Rev.*, D79:043010, 2009.

[124] Wayne Hu and Takemi Okamoto. Principal Power of the CMB. *Phys. Rev.*, D69:043004, 2004.

[125] Michael J. Mortonson, Cora Dvorkin, Hiranya V. Peiris, and Wayne Hu. Things that go bump in the CMB polarization: features from inflation versus reionization. 2009.

[126] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*, pages 59–70. Cambridge University Press, 1992.

[127] Paniez Paykari and Andrew H. Jaffe. Optimal Binning of the Primordial Power Spectrum. 2009.

[128] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, 538:473–476, 2000.

[129] Antony Lewis and Sarah Bridle. Cosmological parameters from cmb and other data: a monte-carlo approach. *Physical Review D*, 66:103511, 2002.

[130] Jerome Martin and Christophe Ringeval. Superimposed Oscillations in the WMAP Data? *Phys. Rev.*, D69:083515, 2004.

[131] Jerome Martin and Christophe Ringeval. Exploring the superimposed oscillations parameter space. *JCAP*, 0501:007, 2005.

[132] S. Perlmutter et al. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *Astrophys. J.*, 517:565–586, 1999.

[133] Adam G. Riess et al. Type ia supernova discoveries at z¿1 from the hubble space telescope: Evidence for past deceleration and constraints on dark energy evolution. *Astrophys. J.*, 607:665–687, 2004.

[134] Pierre Astier et al. The supernova legacy survey: Measurement of $omega_m, omega_lambda$ and w from the first year data set. *Astron. Astrophys.*, 447:31–48, 2006.

[135] M. Kowalski et al. Improved Cosmological Constraints from New, Old and Combined Supernova Datasets. 2008.

[136] W. L. Freedman et al. Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant. *Astrophys. J.*, 553:47–72, 2001.

[137] R. A. Alpher, H. Bethe, and G. Gamow. The origin of chemical elements. *Phys. Rev.*, 73(7):803–804, Apr 1948.

[138] C. Blake, D. Parkinson, B. Bassett, K. Glazebrook, M. Kunz, and R. C. Nichol. Universal fitting formulae for baryon oscillation surveys. *MNRAS*, 365:255–264, January 2006.

[139] N. Kaiser. A sparse-sampling strategy for the estimation of large-scale clustering from redshift surveys. *MNRAS*, 219:785–790, April 1986.

[140] A. Liddle, P. Mukherjee, and D. Parkinson. Model selection in cosmology. *Astronomy and Geophysics*, 47(4):040000–4, August 2006.

[141] R. Trotta. Forecasting the Bayes factor of a future observation. *MNRAS*, 378:819–824, July 2007.

[142] R. Trotta. Applications of Bayesian model selection to cosmological parameters. *MNRAS*, 378:72–82, June 2007.