

Reconstruction of primordial power spectrum(PPS) from CMB angular power spectrum

An Ill-posed inverse problem

Student

Harsh Prajapati

*(Institute of Engineering and Technology, Ahmedabad University, Ahmedabad
(India))*

Supervisor

Dr. Jayanti Prasad

(CMS, Savitribai Phule Pune University, Pune (India))

Acknowledgement

First and foremost I would like to thank **Dr.Jayanti Prasad** under whose guidance I carried out this project. He gave me an opportunity to work on one of the important problems in Cosmology. His constant supervision,tremendous help and support througout this project is invaluable. I was new to this field but he provided me with all the background and information as and when needed. I got to learn a lot of valuable things from him. I would also like to thank **Prof. Tarun Souradeep** who arranged for my visit at IUCAA and provided valuable insights into my project. I wish to extend my gratitude to the **IUCAA staff** who provided an excellent hospitality and took care of all my needs.

I would also take this opportunity to thank **Dr.Gaurav Goswami** who provided me with this opportunity of being at IUCAA and providing all the academic support needed.

Last but not the least, I would like to thank my parents,sister and friends who provided me with the much needed emotional and moral support,who believed in me and gave me strength.

Harsh Prajapati

Contents

1	Cosmic Microwave Background Radiation	5
1.1	What is Cosmic Microwave Background	5
1.2	History	7
1.3	Observing the CMB	8
1.4	CMB Anisotropies	10
1.5	Angular Power Spectrum	10
1.6	Primordial Power Spectrum	12
2	Ill posed inverse problems	14
2.1	Inverse problems	14
2.2	Ill-posed problems	16
2.3	Linear inverse problems	17
2.4	Solving ill-posed inverse equations	18
2.4.1	Regularization techniques	18
2.4.2	Bayesian Analysis	19
2.5	SVD Regularization	20
2.5.1	SVD theory	20
2.5.2	Regularization using SVD Decomposition	22
3	Reconstructing primordial power spectrum of CMB data	23
3.1	Introduction to the problem	23
3.2	Approach	25

3.2.1	SVD Reconstruction	25
3.2.2	SVD basis	26
3.2.3	Maximum Likelihood Estimation	28
4	Results	30
4.1	Numerical codes and Data	30
4.2	Reconstruction from Simulated Data	32
4.3	Reconstruction from WMAP-9 year and Planck Data	33
4.3.1	WMAP-9 year data	34
4.3.2	Planck	37
4.4	Reconstruction of features	38
5	Discussion & Conclusions	46

Abstract

CMB anisotropies as we observe in the sky are sourced by the primordial density fluctuations. These density fluctuations are responsible for the evolution of structure such as galaxies and clusters of galaxies in our universe. Thus studying these primordial density fluctuations is an important aspect of modern cosmology. CMB anisotropies at different angular scales are represented by the angular power spectrum C_l , similarly density fluctuations at different length scales are represented by primordial power spectrum(PPS) $P(k)$. Thus reconstructing the PPS from CMB angular power spectrum is an important problem in cosmology. These spectra are related by the radiation transportation kernel $G(l, k)$. It is well known that this reconstruction problem is an ill-posed inverse problem. In this work we have employed Singular Value Decomposition(SVD) to find the pseudo-inverse of the kernel G_{lk} . SVD method does not incorporate noise and hence no mechanism to estimate the error band on reconstructed PPS. Thus we have reduced the problem into SVD basis and augmented it with Maximum Likelihood Estimation(MLE) to get the least square solution along with error bands. We use the Planck and WMAP-9 year data to obtain the reconstructed PPS. From our study we find that the power law PPS cannot be ruled out, still PPS which has cut-off at large scales is more favored.

Chapter 1

Cosmic Microwave Background Radiation

1.1 What is Cosmic Microwave Background

According to the Big Bang model of the universe, just after the big bang, the early universe was a hot and dense plasma of photons, electrons and protons. Matter and Radiation existed in thermal equilibrium. Temperature was well above the binding energies of proton and electron and hence neutral atoms were disintegrated by the high energy photons as soon as they were formed. The photons were scattered by the free moving electrons.

Due to the rapid expansion of the universe, eventually the temperature started falling. At a point in time, temperature went below the binding energies of the typical nuclei. This allowed the formation of neutral atoms and hence light elements began to form. This process is known as recombination. As a result of recombination the photons were no longer scattered off the electrons and hence radiation decoupled from matter. This process is known as decoupling. These process did not happen instantaneously, they happened over a period of time.

The surface where the photons were last scattered by the electron is

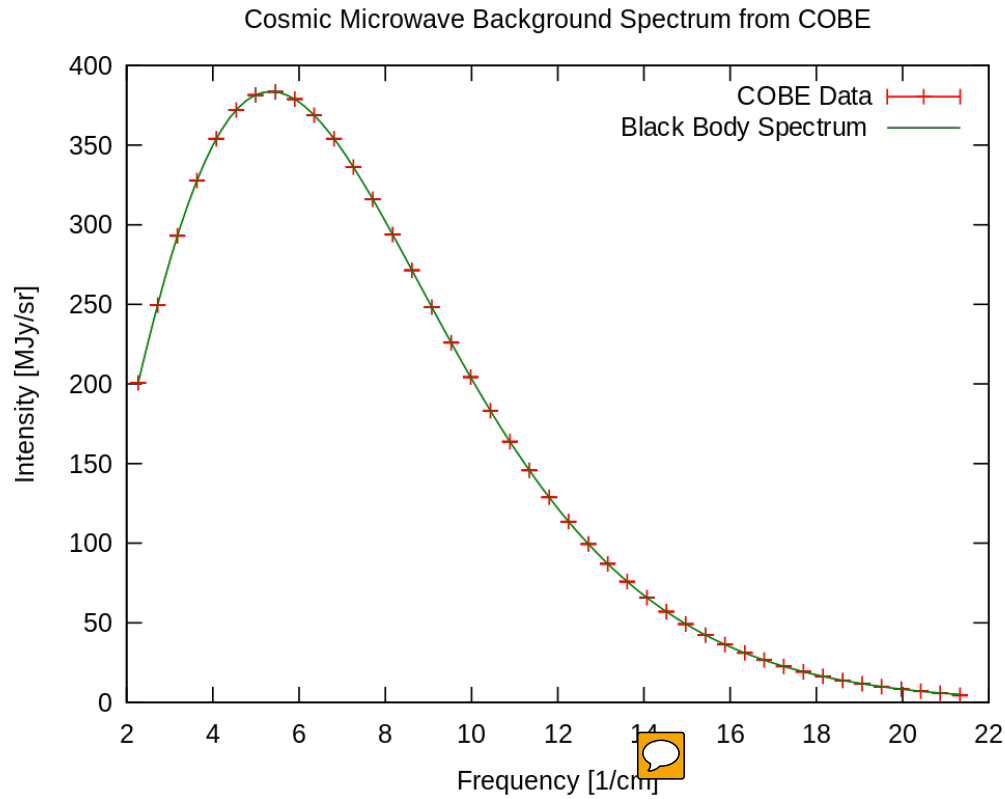






Figure 1.1: CMB spectrum mimicking the spectrum of black-body at 2.732K
 (Source: <http://i.stack.imgur.com/EAUGG.png>)

known as the last scattering surface. These photons  since then moving freely in the universe. The structures formed thereafter were transparent to these photons. These photons from the last scattering surface constitute the cosmic microwave background that we see today. It is an isotropic radiation equivalent to the black body radiation emitted by a black body having  temperature around 2.732K.

1.2 History

The discovery of Cosmic microwave Background was rather accidental.  In 1965 Arno Penzias and Robert Wilson, two radio astronomers at the Bell laboratories were carrying out radio astronomy experiments with a Dicke radiometer. While carrying out their experiments their Antennae picked up an excess of noise apart from the thermal noise. This noise was isotropic and caused the antennae temperature to read 4.7K more than the usual. They could not account for this excess noise.

During that period, the speculations for existence of a radiation background were being made simultaneously in different parts of the world. In 1940's Big bang theory of nucleosynthesis was developed by George Gamow, Ralph Alpher and Robert Herman. In 1948 as an extension of this theory Ralph Alpher and Robert Herman speculated a radiation background with a present temperature of around 5K. In early 1960's Yakov Zel'dovich re-discovered this background radiation. During the same time Robert Dicke independently speculated this radiation background. Robert Dicke and his colleagues Peebles, Rolls and Wilkinson were building a Dicke radiometer to  detect this radiation when they heard about the discovery of excess noise by Penzias and Wilson. These people met and confirmed that the excess noise was indeed due to the background radiation and published a pair of companion letters in The Astrophysical Journal. In 1978, Penzias and Wilson received a Nobel Prize in Physics for their discovery.

There were many reasons provided for this background radiation. But in 1970s a consensus was reached that it was the remnant of big bang. This was due to a number of measurements of the background radiation made at different frequencies which showed that spectrum followed the thermal black body spectrum. Thus the discovery of CMB gave the big bang theory an upper hand among various other theories explaining the evolution of universe.

1.3 Observing the CMB

CMB is an open window to the early universe and hence researchers wanted it to study it at wide range of frequencies with better resolution and sensitivity.

Penzias and Wilson discovered it at $2.33GHz$ and found it to be completely isotropic. After this many studies were made.

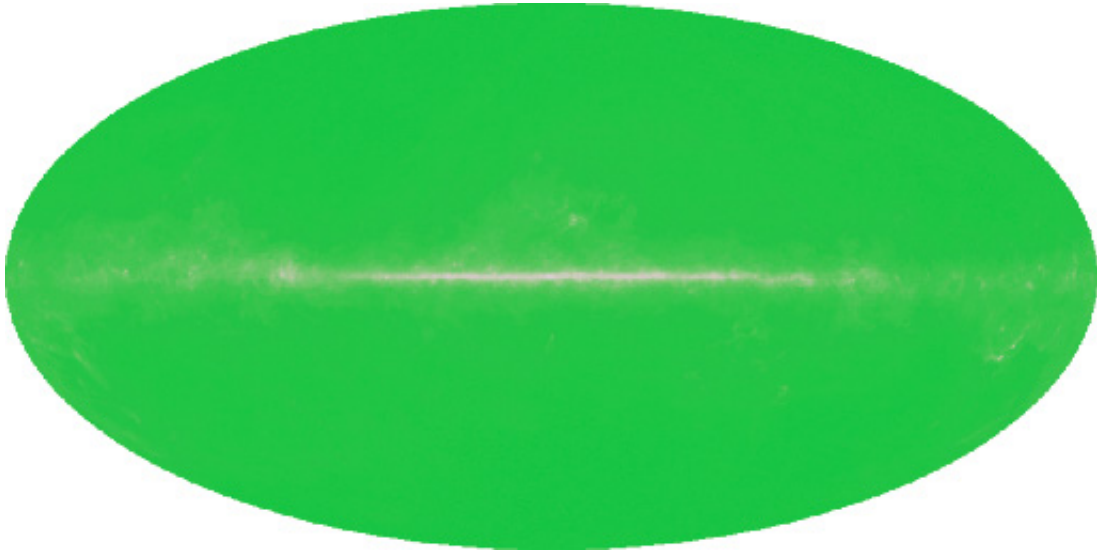


Figure 1.2: CMB sky as observed by Penzias and Wilson
Source:<http://planck.cf.ac.uk>

The next remarkable observation was made by NASA's COBE mission. COBE had an angular resolution of 7° and made observations in the frequency range of $2.24GHz-2.99GHz$.^[15] It was the first to observe anisotropies in the CMB sky. This inspired future missions as a better resolution and wider frequency range would help researchers study the anisotropies more closely.

In June, 2001 NASA launched WMAP, which had a much higher angular resolution than COBE. It could measure anisotropies at a resolution of 0.2° . It made observations in the frequency range $22GHz-90GHz$. The sensitivity

with which it measured temperature was $35\mu K$ per 0.3 square pixel.[15]

The latest is the ESA's Planck mission launched in May,2009. It works in two frequency ranges.The low frequency instrument(30-100GHz) has an angular resolution around 10 arc-minute and sensitivity of $12\mu K$. The high frequency instrument(100-850GHz) has an angular resolution of 5 arc-minute and sensitivity of $5\mu K$.[15]

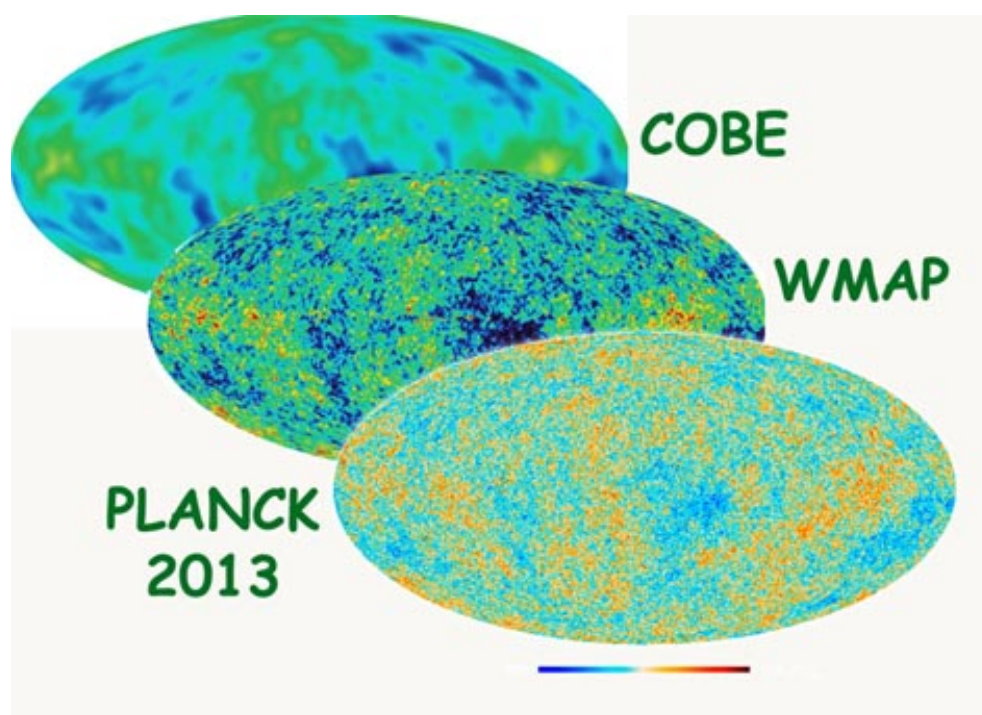


Figure 1.3: CMB sky as observed by different missions
Source:<http://www.planetastronomy.com/astronews/astrn-2013/04/astron3.jpg>

1.4 CMB Anisotropies

The CMB radiation is believed to be uniform with a mean temperature of 2.732K, but on precise measurements you see slight deviations from the mean temperature. These deviations are known as CMB anisotropies. These temperature anisotropies give us an insight that photons from the last scattering surface came from different density regions. This tells us that there were density perturbations in the early universe which is responsible for the evolution of structure in the universe. Thus studying these anisotropies would provide more insights of the fluctuations in density and the geometry of space-time at the time of last scattering surface (Primary Anisotropies) and that caused later (Secondary Anisotropies).

In the coming sections we will introduce the angular power spectrum which represents the temperature fluctuations in the CMB sky and the primordial power spectrum which represents the density fluctuations in the early universe.

1.5 Angular Power Spectrum

The temperature anisotropies are represented by a scalar function ΔT . It is a dimensionless quantity and is defined as:

$$\Delta T(\hat{n}) = \frac{T(\hat{n}) - T_{mean}}{T_{mean}} \quad (1.1)$$

where T represents the temperature.

Any function over a sphere can be expressed as the sum of spherical harmonics (3D Fourier transform in spherical co-ordinates). Thus we can write ΔT as:

$$\Delta T(\hat{n}) = \sum_2^{lmax} \sum_{-l}^l a_{lm} Y_{lm}(\hat{n}) \quad (1.2)$$

Note that the l value starts from two instead of zero. This is because:

- The quantity represents deviation from mean and hence for $l = 0$, $a_{00} = 0$.
- The $l = 1$ term also known as the dipole is dominated by the Doppler shift due to the motion of solar system with respect to the last scattering surface.

Hence we are only interested in values $l \geq 2$ for which the anisotropies are related to the perturbations in the early universe.

The coefficients a_{lm} are given by:

$$a_{lm} = \int Y_{lm}^*(\hat{n}) \Delta T(\hat{n}) d\Omega \quad (1.3)$$

These a_{lm} 's are Gaussian random variables with mean 0 and variance C_l . Also temperature is isotropic and hence the a_{lm} 's depend only upon l (angular scale) and not m (orientation). Thus we define the angular power spectrum C_l as:

$$C_l = \frac{1}{2l+1} \sum_m | \langle a_{lm} \rangle |^2 \quad (1.4)$$

Note that we have used angular brackets here $\langle \rangle$, which means an ensemble variance, that is variance with respect to the average taken over all possible universe. But we can make observations only in our universe and hence have only one realization of anisotropies. Thus the observed angular spectrum can be written as:

$$C_l^{obs} = \frac{1}{2l+1} \sum_m |a_{lm}|^2 \quad (1.5)$$

As a result of this, there is always an inherent error in the observation of C_l^{obs} , which limits the accuracy in C_l^o no matter how precise instrument we use. This error is known as the cosmic variance and is defined as:

$$\frac{C_l^{obs} - C_l}{C_l} = \sqrt{\frac{2}{2l+1}} \quad (1.6)$$

As we can see from the definition, the cosmic variance is high at large angular scales and reduces at lower scales.

Summarizing the above discussion we can say that the angular spectrum C_l gives us information about the temperature anisotropies at different angular scales.

1.6 Primordial Power Spectrum

Primordial power spectrum is the representation of density fluctuations in the early universe. Let $\delta(\vec{x})$ be a scalar function representing the density fluctuations. It is defined as:

$$\delta(\vec{x}) = \frac{\rho(\vec{x}) - \rho_{mean}}{\rho_{mean}} \quad (1.7) \img alt="comment icon" data-bbox="810 421 846 445"/>$$

where ρ represents the density.

Taking the 3-D Fourier transform of $\delta(\vec{x})$ we get:

$$\delta(\vec{x}) = \int \delta_k e^{-i\vec{k} \cdot \vec{x}} d\vec{k} \quad (1.8) \img alt="comment icon" data-bbox="800 523 836 547"/>$$

where δ_k are the Fourier components.

Taking the ensemble average of these Fourier components gives the primordial power spectrum $P(k)$:

$$\langle \delta_k, \delta_{k'} \rangle = \frac{2\pi^2}{k^3} \delta(k - k') P(k) \quad (1.9) \img alt="comment icon" data-bbox="795 637 826 661"/>$$

The primordial power spectrum is believed to be a featureless power law:

$$P(k) \propto k^{n_s-1} \quad (1.10)$$

where n_s is the scalar spectral index.

Having defined the angular power spectrum and primordial power spectrum, it is now important to know the relation between them and the methods

we have used to reconstruct the primordial power spectrum from the angular power spectrum. This reconstruction is an ill-posed inverse problem and hence first we will have some information about ill-posed inverse problems in the next chapter.



Chapter 2

Ill posed inverse problems

2.1 Inverse problems

Given a mathematical model and a set of inputs to the model we can always predict the outcome for deterministic systems. This problem of predicting the outcome is known as the direct or the forward problem. The inverse problem as the name suggests, involves solving for the parameter or the input using the measured output. Inverse problems are one of the most important problems in mathematics, science and engineering as they allow us to calculate parameters and quantities that are not directly measurable.

Inverse problems come paired with the direct problems. A very simple example would be, given two numbers we can multiply them to find their product. This is the direct problem of multiplication. The corresponding inverse problem is, given a number find its factors. More practical problems include, calculating electric field induced by the known distribution of electric charges is a direct problem. The corresponding inverse problem is deducing the distribution of the electric charges from the measurements of the field.


While the direct problems typically have a unique solution and are simpler to deal with, inverse problems do not necessarily have a unique solution thus making it more complex. Let us go back to our simple example of mul-

multiplication of two numbers. The product obtained is a unique number but the factorization does not always yield a unique solution.

Let us model a practical problem [1] to understand the complexities involved in solving an inverse problem:

Suppose we want to study a spectrum, this would require us to know the function $\hat{x}(\nu)$, $\nu \in \mathbb{R}$, but due to experimental constraints we only have access to the time series data $x(t)$ of which $\hat{x}(\nu)$ is the Fourier transform (FT). Further, due to instrumentation constraints $x(t)$ can only be observed within a window $h(t)$ which gives us the observed function $y(t)$:

$$y(t) = x(t)h(t) \quad (2.1)$$

Considering the fact that the experimental data would have finite regularly spaced samples of the function $y(t)$ incorporated with the measurement errors we can write eq 2.1 

$$y_n = h_n x_n + b_n \quad (2.2)$$

where x_n and h_n are the samples drawn from the functions $x(t)$ and $y(t)$ respectively. Thus to obtain the spectrum we want to study, first we need to calculate x_n and then take the FT of x_n to get the frequency spectrum. The process involves many difficulties:

Firstly, even if we do not consider the weighted window $h(t)$, the finite samples of the data suggest that we would have information only on $\hat{x}_1(\nu)$. This $\hat{x}_1(\nu)$ is the result of the periodization that takes place due to sampling and hence is a periodic function deduced from $\hat{x}(\nu)$:

$$x_n = \int \hat{x}_1(\nu) e^{(2j\pi\nu n)} d\nu \quad (2.3)$$

Thus $\hat{x}(\nu)$ must have finite number of non-zero elements and the sampling must be such that there is no aliasing, to be able to obtain the correct spectrum.



Now eq2.2 can be written as:

$$y_n = \int h(\nu) \star \hat{x}_1(\nu) e^{(2j\pi\nu n)} d\nu \quad (2.4)$$

where $\hat{h}(\nu)$ is the FT of $h(t)$. The convolution indicates the loss in resolution power(differentiate between two distinct peaks) of the spectrum, due to the window $h(t)$.


Secondly, it is impossible to get an exact mathematical inverse of the eq2.4 even in the absence of noise. Meaning, since there are finite number of sample points, there is infinite number of solutions $\hat{x}(\nu)$ that satisfies the constraint:

$$x_n = \frac{y_n}{h_n} \quad \text{if } n \in [1..., N] \quad \text{and} \quad h_n \neq 0 \quad (2.5)$$


Thirdly, noise in the measurement of y_n adds to the misery. For small values of h_n the noise in the corresponding measurement will get amplified.


The above example illustrates the posedness and the difficulties that arise in most of the practical inverse problems. Typically, inverse problems are ill-posed.


2.2 Ill-posed problems

To define ill-posed problems it is necessary to know what are well-posed problems. Well-posed problems defined by Hadamard are those problems that satisfy these conditions: 

- A solution to the problem exist. (*existence*)
- The solution is unique. (*uniqueness*)
- The solution depends continuously on data and parameter. (*Stability*)

Any  problem which fails to satisfy one of these condition is said to be ill-posed. The problem discussed in the previous section is thus ill-posed as it

violates the uniqueness condition. One may argue that the example problem in the previous section is ill-posed due to its discretization. But, the corresponding continuous case faces the problem of stability.[for details refer 
[\[1\]](#)]

Here it is also important to mention about ill-conditioned problems.
A well-posed problem may be ill-conditioned in the sense that small error in the data may result into large deviations in the solution. The continuity condition is necessary to ensure stability and robustness of the solution but its not sufficient [\[1\]](#) and hence well-posed problems may become ill-conditioned which makes the solution non-robust. Inverse problems may be ill-conditioned.

2.3 Linear inverse problems


Consider a linear system characterized by the equation:


$$y_{m \times 1} = A_{m \times n} x_{n \times 1} \quad (2.6)$$

where $y \in Y$ and $x \in X$ are vectors, $A : X \rightarrow Y$, is the coefficient matrix. The inverse problem here is solving for x , given y and A . The posedness of the problem here depends upon the relation between m and n and the properties of the coefficient matrix A . Let us look at the different cases:

- For $m = n$. If A is non-singular ($\det A \neq 0$), then $A^{-1}y$ is a unique solution for x . But if A is singular then either no solution exist or there is no unique solution
- For $m > n$. The system is over-determined. No exact solution exist¹. However, if the rank of A is n then a least square solution can be found, which may be treated as unique.

¹exact solution would exist only when $\text{rank}(A) = n$ and the rows of A are linearly dependent which is never true in practical examples

- For $m < n$. The system is under-determined that is there are more unknowns than equations. No unique solutions exists, there are infinitely many solutions. To find a solution one needs to incorporate prior information or constrain the unknowns. 

Most of the practical linear inverse problems are either under-determined or over-determined and hence are ill-posed. In a very crude sense one can say that linear inverse problems are mostly about calculating A^{-1} .  Here it is important to note that matrix A is ill-conditioned and hence its inversion would amplify the noise.



2.4 Solving ill-posed inverse equations

We now know that inverse problems are difficult to solve due to three main reasons:

- No unique solution exists i.e. infinitely many solutions are possible.
- The solution does not depend continuously on the data.
- The solution is not robust i.e. the problem is ill-conditioned.

These conditions can be dealt with by incorporating the prior knowledge about the solution into the formation of the problem or by constraining the required unknown. This can be achieved using various methods:

2.4.1 Regularization techniques

Regularization techniques are one of the most used techniques to solve inverse problems.  These techniques allow us to incorporate the prior knowledge about the solution in the form of regularization function and regularization parameter. It helps us rule out the other possible solutions by converging onto one solution.  In some cases they also help us establish the continuity condition[[2]]. Various regularization techniques include:

- Dimensionality control for under determined systems.
- Minimising a composite criteria

A very common regularization technique is use of a penalty function with the least square method:

$$x = \operatorname{argmin}(\|y - Ax\|^2 + \lambda\|x\|^2) \quad (2.7)$$

It is primarily used to compensate for over-fitting of data. Here λ is the regularization parameter and $\|x\|^2$ is the regularization function. The choice of regularization technique depends upon the problem at hand.

2.4.2 Bayesian Analysis

Bayesian Analysis is a natural way to incorporate prior knowledge into the problem. It is based on the Bayes' theorem which relates posterior probability to the likelihood and prior probability:


$$P(\theta|d) = \frac{P(d|\theta)P(\theta)}{P(d)} \quad (2.8)$$

where $P(\theta|d)$ is the posterior probability, $P(d|\theta)$ is the likelihood and $P(\theta)$ is the prior probability and $P(d)$ is the evidence.

Maximizing this probability would give us the most probable solution among the infinite solutions and hence ensure uniqueness. This is known as Maximum likelihood estimate(ML). Bayesian analysis also allows us to estimate the error bars on the solution making it a more informative tool for solving inverse problems.

In the present work we use dimensionality control as the zeroth order regularization and ML to estimate error in the solution.

2.5 SVD Regularization

Singular Value Decomposition(SVD) acts as a zeroth order regularization scheme. It helps us to get rid of those parts of the ill-conditioned matrix that can corrupt the solution. 

2.5.1 SVD theory

We know that a $m \times n$ matrix A maps the vectors from $\mathbb{R}^n \rightarrow \mathbb{R}^m$, more specifically all the vectors in its row space ($R(A)$) to the corresponding vector in the column space($C(A)$). We are now interested in finding those orthonormal basis of $R(A)$ that get transformed to orthonormal basis in $C(A)$. Note that not all orthonormal basis of $R(A)$ will get transformed to orthonormal basis of $C(A)$, thus these would be unique set of vectors.

Suppose $v_1, v_2, v_3, \dots, v_r$ are the unit vectors in $R(A)$ and $u_1, u_2, u_3, \dots, u_r$ are the unit vectors in $C(A)$, then:

$$A[v_1 \ v_2 \ v_3 \ \dots v_r \ v_{r+1} \ \dots v_n] = [u_1 \ u_2 \ u_3 \ \dots u_r \ u_{r+1} \ \dots u_m]\Sigma$$
$$AV = U\Sigma$$

$$A = U\Sigma V^T \tag{2.9}$$

where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, 0, 0)$, r is the rank of matrix A and $v_{r+1} \dots v_n$ are vectors in the null space($N(A)$) of A and $u_{r+1} \dots u_m$ are vectors in the null space($N(A^T)$) of A^T . σ 's are the scaling factors.

Eq 2.9 is the Singular value decomposition equation. Now, we need to find U, V and Σ . We can write:

$$A^T = V\Sigma^T U^T \tag{2.10}$$

Multiplying eq 2.9 by A^T we get:

$$\begin{aligned} A^T A &= V \Sigma^T \Sigma V^T \\ A^T A &= V \Sigma^2 V^T \end{aligned} \quad (2.11)$$

where $\Sigma^2 = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, 0, 0, 0)$ Eq 2.11 is similar to the diagonalization equation of the positive definite matrix[3], given by:

$$B = Q \lambda Q^T \quad (2.12)$$

where B is a positive definite matrix, λ is a diagonal matrix containing eigen values of B and Q is an orthogonal matrix containing orthonormal eigen vectors of B .

Since, $A^T A$ is positive definite, we can say that V is an orthogonal matrix containing orthonormal eigen vectors of $A^T A$ and diagonals of Σ^2 the eigen values of $A^T A$. Similarly we can say that U is an orthogonal matrix containing orthonormal eigen vectors of AA^T and diagonals of Σ^2 the eigen values of AA^T .

We can see that the eigen values of AA^T and $A^T A$ are same. Thus, eigen values of A are square root of eigen values of AA^T or $A^T A$.

Summarizing the above discussion, a $m \times n$ matrix A can be decomposed as follows:

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T \quad \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, 0, \dots, 0) \quad (2.13)$$

where U is an orthogonal matrix whose column contain eigen vectors of AA^T , V^T is also an orthogonal matrix whose rows contain eigen vectors of AA^T and Σ is a diagonal matrix containing singular values of A .

Some important points to take under notice are:

- Given any matrix A , the matrices $A^T A$ and AA^T are always positive semi-definite.
- The eigen values of a positive semi definite matrices are always greater than equal to zero.

- If v_1, v_2, \dots, v_r are the eigen vectors of $A^T A$ then Av_1, Av_2, \dots, Av_r are the eigen vectors of AA^T .
- For a rank-deficient matrix where $r < m, n$ we have r non-zero eigen values and the rest $\min(m - r, n - r)$ eigen values are zero.

2.5.2 Regularization using SVD Decomposition

In eq 2.9 we see that singular values σ_i for $i > r$ are zero, where r is the rank of the matrix. We can discard this singular values and the corresponding eigen vectors in U and V^T . Thus we can write the SVD decomposition as:

$$A_{m \times r} = U_{m \times r} \Sigma_{r \times r} V_{r \times n}^T \quad (2.14)$$

In practice we also discard the singular values which are very small i.e. $\sigma_i < \epsilon$, where ϵ is a threshold value.

Thus this tunable parameter r allows us to get rid of those parts of the matrix that corrupt the solution. Also, by discarding the singular values that are zero, we are getting rid of the null-space vectors which are the reason for infinite solutions for an under-determined system. Thus ensuring a unique solution.

Now we can use the SVD to calculate the pseudo-inverse or the Moore-Penrose inverse, given by:

$$A^\dagger = V \Sigma^{-1} U^T \quad \Sigma^{-1} = \text{diag}(1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_r) \quad (2.15)$$

The solution to our linear inverse problem using SVD would be:

$$\begin{aligned} x &= A^\dagger y \\ x &= V \Sigma^{-1} U^T y \end{aligned} \quad (2.16)$$

This is the minimum norm solution².

²for a system $Ax = b$ a minimum norm solution is the solution which has the least value of $\|x\|_2$

Chapter 3

Reconstructing primordial power spectrum of CMB data

3.1 Introduction to the problem

As C_l ¹ represents temperature anisotropies of the CMB sky in the present time, P_k represents the density perturbations in the early universe during the time of recombination. These perturbations help us explain the evolution of structure in the Universe. Thus reconstruction of the primordial power spectrum from the angular power spectrum is an important problem in Cosmology.

The angular power spectrum is related to the primordial power spectrum by the following convolution integral:

$$C_l = \int \frac{dk}{k} |\Delta_l(k, \tau_0)|^2 P(k) \quad (3.1)$$

where $\Delta_l(k, \tau_0)$ is the radiation transport kernel and can be obtained by solving Einstein-Boltzmann set of equations for given cosmological parameters. [[4]]

¹ C_l^{TT}

Discretizing in k-space we can write eq 3.1 as:

$$C_l = \sum_{i=1}^N G_{li} P_i \quad (3.2)$$

where G_{li} is defined as:

$$G_{li} = \frac{\Delta k_i}{k_i} |\Delta_l(k, \tau_0)|^2 \quad (3.3)$$

Eq 3.2 in linear algebra notation can be written as:

$$C_l = G_{lk} P_k \quad (3.4)$$

where $C_l \in \mathbb{R}^m$, $G_{lk} \in \mathbb{R}^{m \times n}$ and $P_k \in \mathbb{R}^n$. G_{lk} is known as the transfer kernel or simply kernel.

For WMAP and PLANCK data l_{max} is 1200 and 2500 respectively. The transformation kernel G_{lk} is a rapidly oscillating function hence dense sampling of k-space is needed for appropriate representation of the continuous function. Thus k_{max} is taken as 6080.

Thus, as seen, here $n > m$ and hence the system is under-determined i.e. there are less equations and more unknown. This also means the kernel G_{lk} is rectangular and hence no direct inverse exist. Thus infinitely many solutions are possible. Also the condition number of kernel G_{lk} is of the order of 10^{10} making it ill-conditioned. These issues makes the problem an ill-posed inverse problem.

To solve this problem, first we use Singular Value decomposition(SVD) as zeroth order regularization scheme. We then augment the SVD with Maximum Likelihood(ML) scheme to get an error band on the reconstructed primordial power spectrum.

In the coming sections, we will discuss our approach in detail.

3.2 Approach

We know that G_{lk} is rectangular, and hence to invert the matrix we need to find its pseudo-inverse. Also G_{lk} is ill-conditioned thus it would be necessary to get rid of those parts of the matrix that will corrupt the solution. Singular value decomposition of the matrix G_{lk} serves as a useful tool for achieving these goals.

3.2.1 SVD Reconstruction

As discussed in the previous section we can diagonalize the transfer kernel G_{lk} using the SVD so that $G_{lk} = U\Sigma V^T$. So we can write eq3.4 as:

$$C_{m \times 1}^l = U_{m \times r} \Sigma_{r \times r} V_{r \times n}^T P_{n \times 1}^k \quad (3.5)$$

where U and V^T are orthogonal matrices having eigen vectors of $G_{lk}^T G_{lk}$ and $G_{lk} G_{lk}^T$ and Σ a diagonal matrix containing singular values of G_{lk} .

Here r is a tunable parameter which allows us to discard the singular values less than some threshold, i.e. for $i > r, \sigma_i < \epsilon$. After keeping r singular values and discarding the remaining singular values we can invert the kernel G_{lk} . Thus the pseudo-inverse or Moore-Penrose inverse is given by:

$$G_{lk}^\dagger = V \Sigma^{-1} U^T \quad (3.6)$$

Using this pseudo-inverse we can reconstruct the Pk and find the minimum norm solution. The reconstruction is given by:

$$Pk = G_{lk}^\dagger Cl \quad (3.7)$$

In fig(3.1) we show the singular values of kernel G_{lk} for some typical values of cosmological parameters.

To know the number of SVD modes, that would faithfully represent the kernel G_{lk} , we try to recover C_l for different values of r . From fig(3.2) we

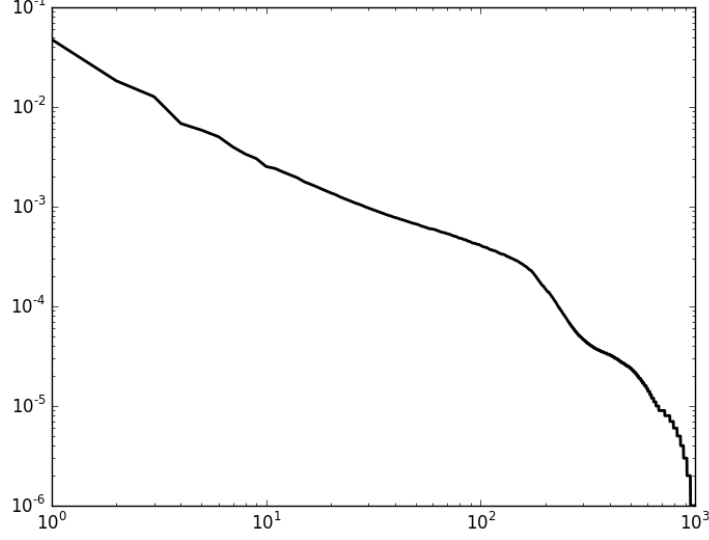



Figure 3.1: Singular Value Decomposition of the kernel G_{lk} for  LAP-9 best fit parameters.

can see that 300 singular values are enough to represent the transformation kernel.

We now use this SVD decomposition to reduce the problem into SVD basis as shown in the next subsection.

3.2.2 SVD basis

The tunable parameter r in eq 3.5 gives us the freedom to replace $\Sigma \in \mathbb{R}^{m \times n}$ by $\Sigma' \in \mathbb{R}^{r \times r}$ of lower dimension, where we choose r to be specifically less than n . Thus we can rewrite the eq 3.4 as:

$$C_l = \tilde{G}_{lk} \tilde{P}_k \quad (3.8)$$

where

$$\tilde{G}_{lk} = U \Sigma' \quad (3.9)$$

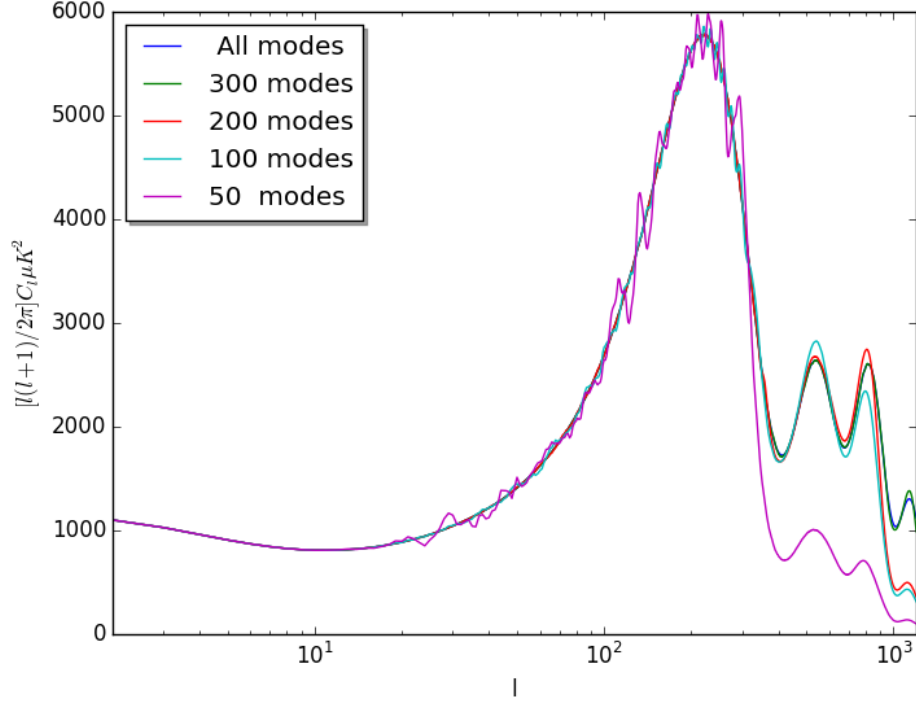


Figure 3.2: Angular power spectrum C_l with different number of SVD modes of G_{lk}

and

$$\tilde{P}_k = V^T P_k \quad (3.10)$$

Thus $\tilde{G}_{lk} \in \mathbb{R}^{m \times r}$ and $\tilde{P}_k \in \mathbb{R}^{r \times 1}$. Here $r < m$ and hence the problem now has become over-determined with more equations and less unknowns. Now we can use Maximum Likelihood Estimation to get the solution.

The motivation behind reducing the problem into SVD basis and the method of Maximum Likelihood is discussed in the next subsection.

3.2.3 Maximum Likelihood Estimation

In the SVD reconstruction you can notice that we ignored the noise in C_l . But noise has a significant contribution in the reconstruction of P_k and hence it is important to take it under consideration. The Maximum Likelihood Estimation would allow us to use noise as an inverse weighting to data meaning the data points with greater noise should be given less weight. Also MLE would allow us to obtain the error band on the reconstructed PPS.

Thus we now consider the equation:

$$C_l = \tilde{G}_{lk} \tilde{P}_k + n \quad (3.11)$$

C_l can be considered as a Gaussian random variable with data values as mean and corresponding noise value as variance. Thus we can write:

$$P(C_l|\tilde{P}_k) = \left(\frac{1}{2\pi}\right)^{M/2} \frac{1}{\sqrt{|C_N|}} \exp\left(-\frac{\chi^2(\tilde{P}_k)}{2}\right) \quad (3.12)$$

where

$$\chi^2(\tilde{P}_k) = (C_l - \tilde{G}\tilde{P}_k)C_N^{-1}(C_l - \tilde{G}\tilde{P}_k) \quad (3.13)$$

Thus we need to find \tilde{P}_k which maximizes the probability of the occurrence of the observed C_l i.e. $P(C_l|\tilde{P}_k)$. This is equivalent to minimizing the $\chi^2(\tilde{P}_k)$. Thus the Maximum Likelihood solution would be :

$$\tilde{P}_{kML} = (G^T C_N^{-1} G)^{-1} G^T C_N^{-1} C_l \quad (3.14)$$

where $(G^T C_N^{-1} G)^{-1}$ is the hessian of $\chi^2()$ whose diagonal elements would give variance on \tilde{P}_{kML} .

Now the original PPS can be recovered from \tilde{P}_{kML} by inverting the eq 3.11 :

$$P_k = V \tilde{P}_{kML} \quad (3.15)$$

Also the covariance matrix for P_k can be obtained by:

$$\text{cov}(P_k) = V(G^T C_N^{-1} G)^{-1} V^T \quad (3.16)$$

The diagonal elements of the covariance matrix would give the variance of P_k . We will use this variance to estimate error bands.

Chapter 4

Results

4.1 Numerical codes and Data

CAMB and CMBfast are two publicly available codes that can generate the radiation transportation kernel and best-fit C_l 's for a set of cosmological parameters [15]. gtFast is the modified version of CMBfast which also serves the same purpose. We have used kernel and best-fits obtained from both CAMB and gTfast to check the consistency in our results.

We have used the Wmap-9 best fit parameters as the values of the cosmological parameters. The parameters and values are:

Cosmological Parameter	Description	Value
$\Omega_b h^2$	Physical Density of baryonic matter	0.02264
$\Omega_c h^2$	Physical Density of Dark matter	0.1138
H_0	Hubble Parameter	$70 \frac{km/sec}{MPc}$
A_s	Amplitude of PPS	2.41×10^{-9}
n_s	Spectral index of PPS	0.972
τ	Reionization optical depth	0.089

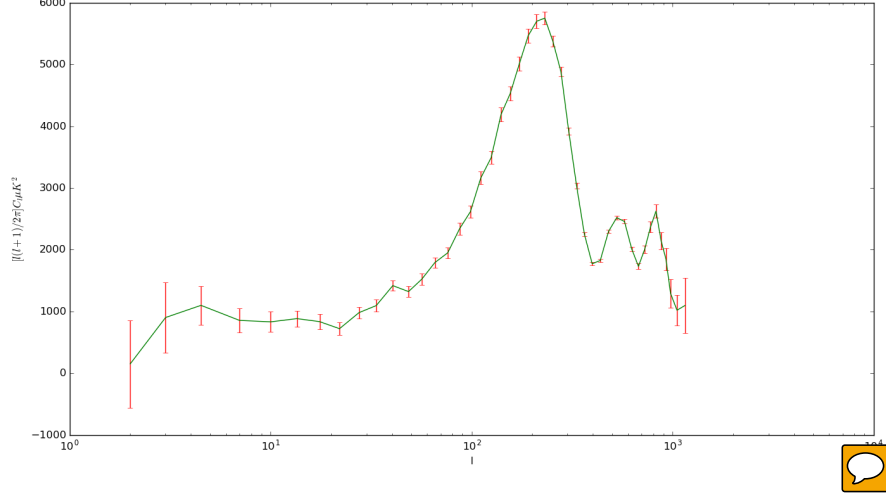


Figure 4.1: WMAP-9 year binned data with error bars.

We reconstruct the PPS using simulated C_l 's and observed C_l 's. For the simulated case we use the best-fit C_l 's obtained from CAMB and gTfast. For observed C_l 's we use the Planck and WMAP-9 year data. Both of them provide binned as well as un-binned data.

For reconstruction of PPS from WMAP-9 year data we use binned data interpolated using cubic spline. For Planck we use un-binned data upto $l=49$ and interpolated binned data from $l=50$.

Note that we have used only temperature data to carry out our study. The same can be done using polarization data.

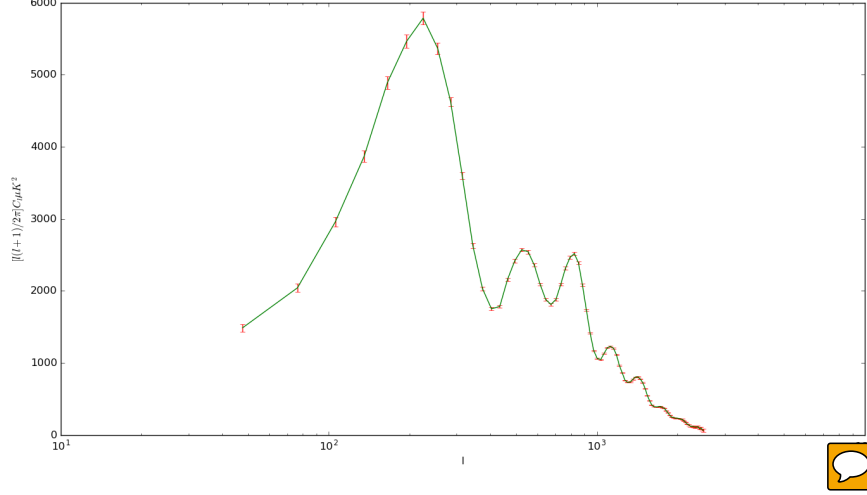


Figure 4.2: Planck binned data with error bars.

4.2 Reconstruction from Simulated Data

In order to validate the SVD method we use Simulated Data that is best-fit C_l and G_{lk} obtained from CAMB and gTfast to reconstruct the PPS. Fig(4.3) shows the reconstructed PPS for different SVD modes(no of singular values chosen) without any smoothing. As seen earlier 300 modes are enough for accepted representation of C_l .

We can see that broad features are recovered even for lower SVD modes. One of the advantages of using SVD is it gives positive reconstruction which is not the case in every reconstruction method.

Fig(4.4) shows the same reconstruction after logarithmic binning(40 bins) and fig(4.5) shows reconstruction using only high-l data.

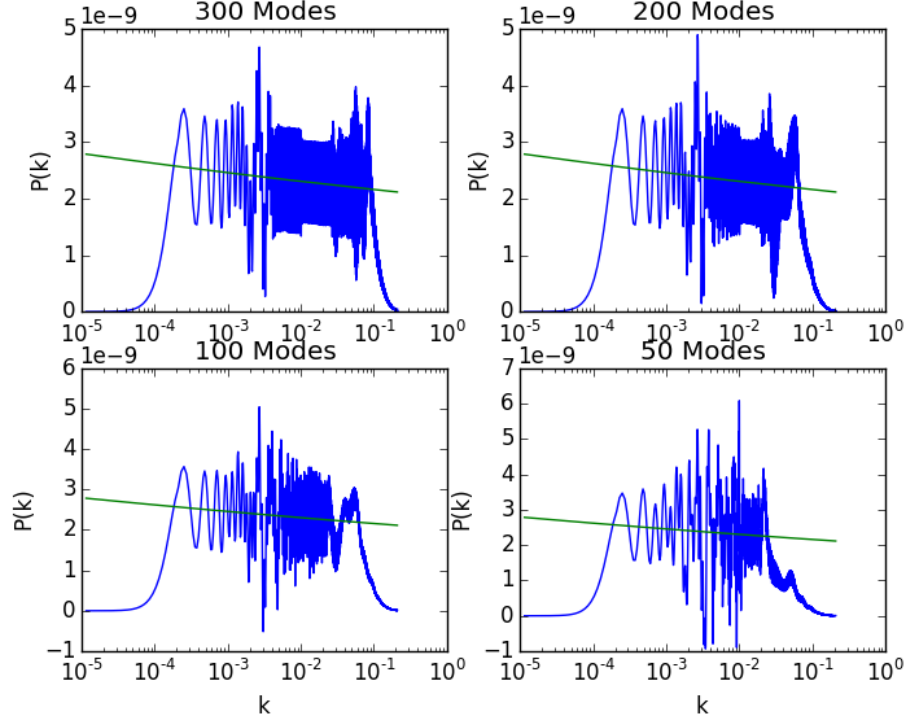



Figure 4.3: SVD reconstruction of power law PPS from simulated data with different number of SVD modes of G_{lk} 

4.3 Reconstruction from WMAP-9 year and Planck Data

After validating the method we use it on the observed data to reconstruct the PPS. Below we show reconstruction obtained from WMAP-9 year data and Planck data using the SVD and MLE methods.

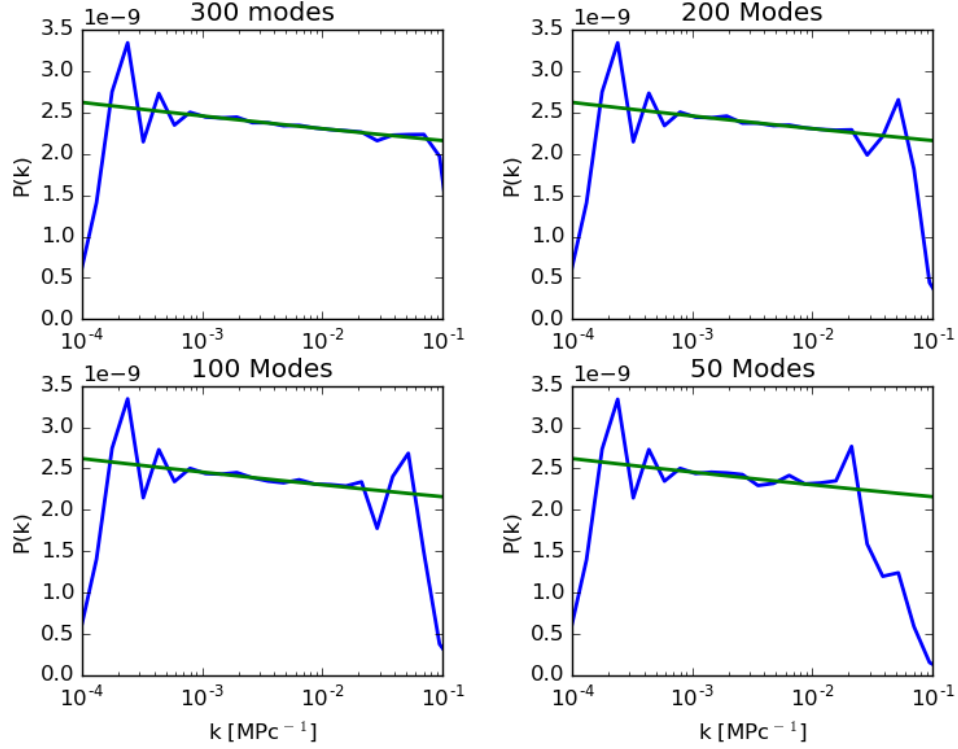


Figure 4.4: SVD reconstruction of power law PPS with different svd modes for simulated data(all l 's) after logarithmic binning of k -space(40 bins).As we can see reconstructed PPS nearly obeys the power law for mid-range k modes even for the least number of SVD modes.As we reduce the number of SVD modes the reconstructed PPS approaches zero for high k 's.

4.3.1 WMAP-9 year data

Fig(4.6) shows the SVD reconstruction from the WMAP-9 year interpolated binned data with all- l values. Fig(4.7) shows the same but only using high- l values.

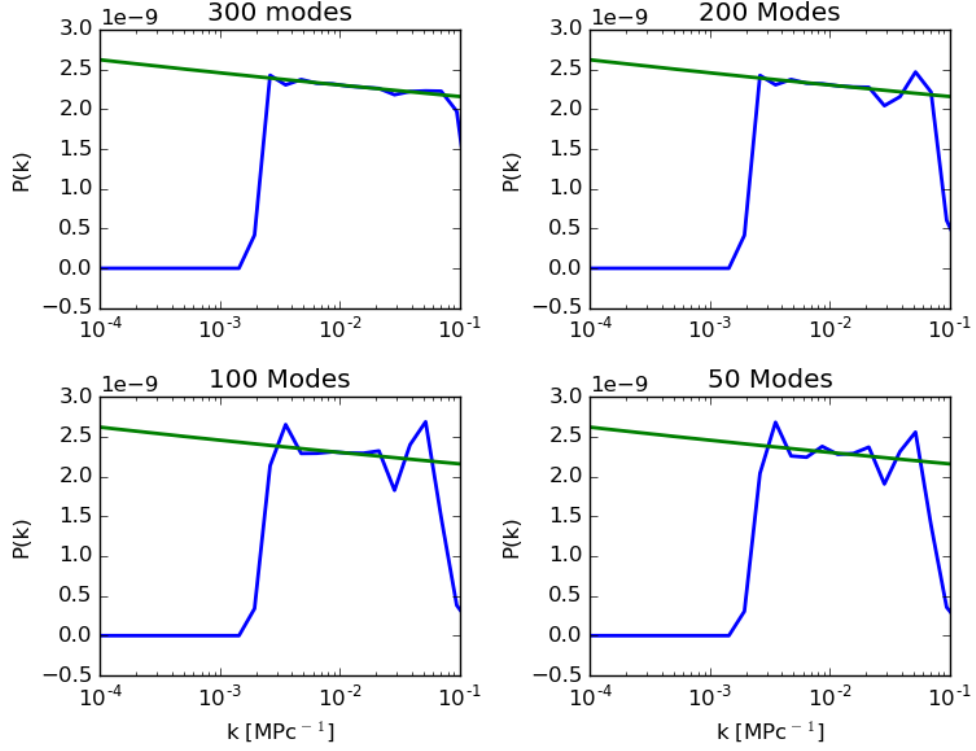


Figure 4.5: SVD reconstruction of power law PPS with different svd modes for simulated data (high l 's from $l=30$) after logarithmic binning of k -space (40 bins). As we can see there is a sharp cut off at low k modes due to absence of large angular scales.

Fig(4.8) shows the C_l obtained from the reconstructed PPS plotted along with the simulated C_l and the WMAP-9 year data. At large angular scales we can see that the C_l obtained from the reconstructed PPS is trying to fit the WMAP-9 year data.

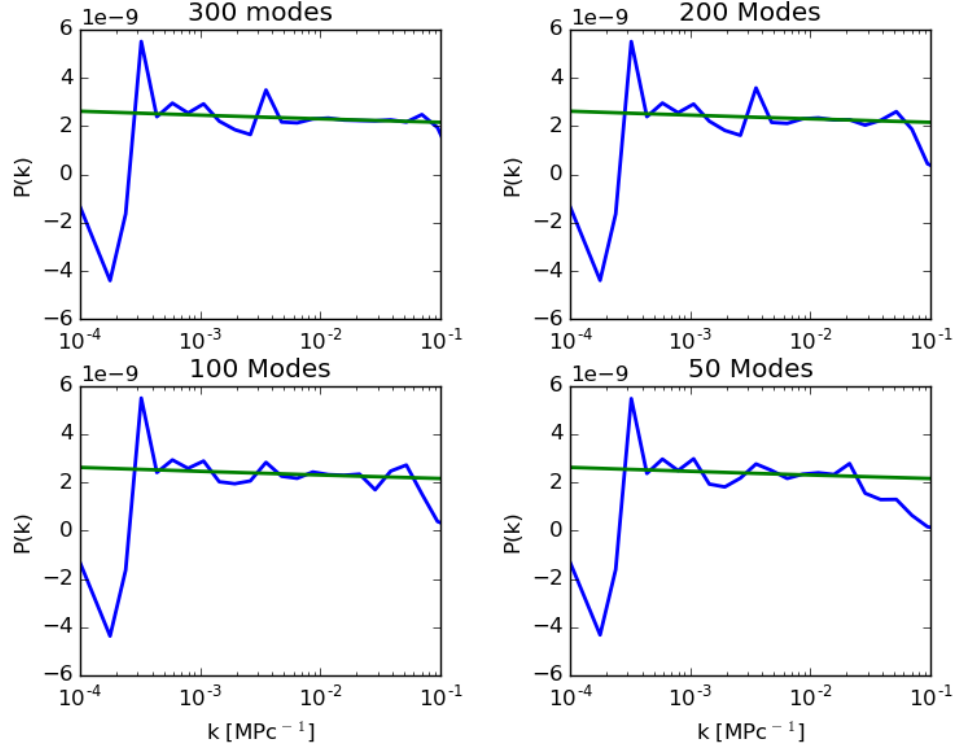


Figure 4.6: SVD reconstruction of power law PPS with different svd modes for interpolated WMAP-9 year data(all l's) after logarithmic binning of k-space(40 bins). As we can see broadly the reconstructed PPS obeys the power law. The negative values at very low k modes is due to glitches in the interpolated WMAP-9 years data.

Fig(4.9) shows the reconstruction using the MLE methods using all-l and fig(4.10) using high-l. The error band estimates still need some improvement at high-k.

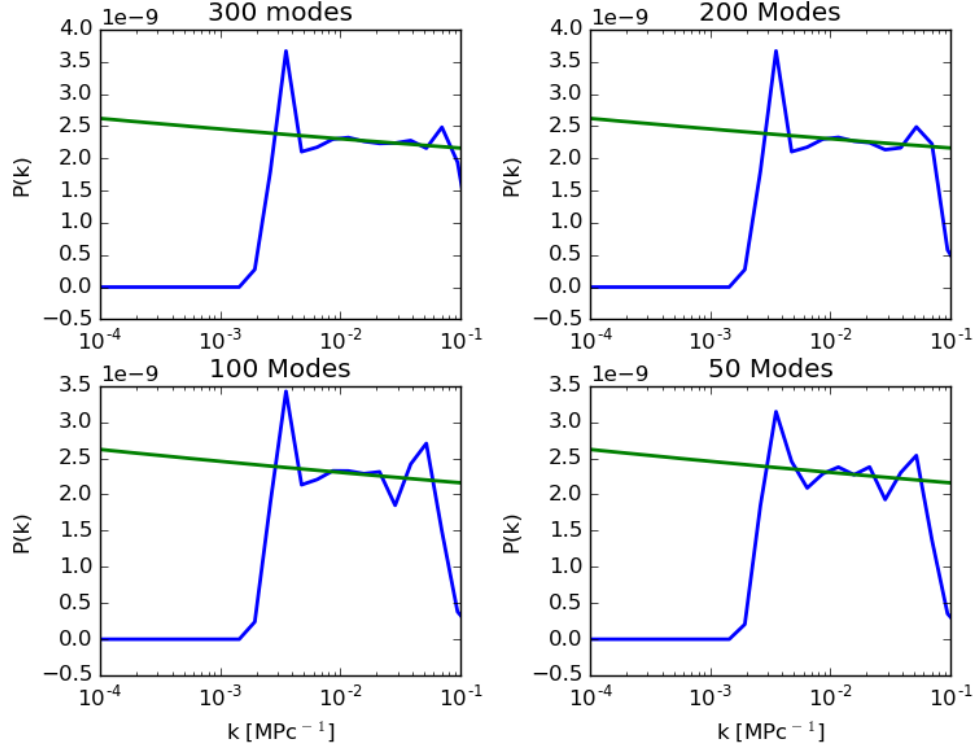


Figure 4.7: Same as in fig(4.6) but with high- l (from $l=30$) data.

4.3.2 Planck

For Planck data we have results obtained using the SVD reconstruction method.

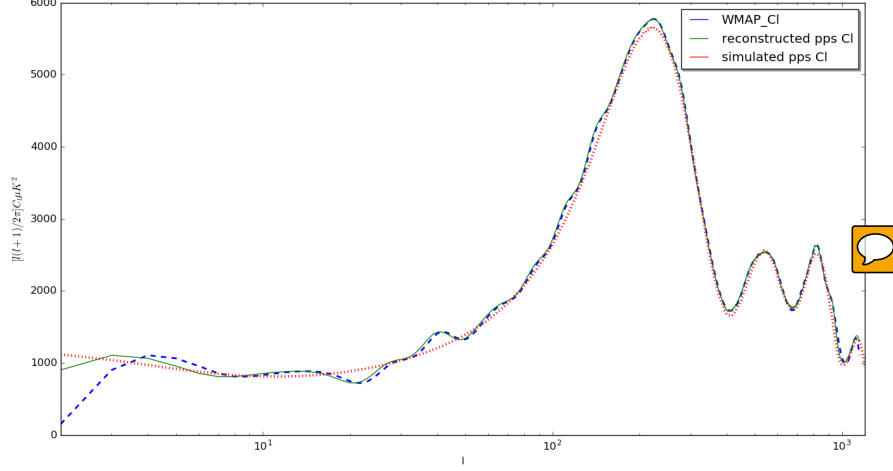


Figure 4.8: C_l obtained from reconstructed PPS along with WMAP-9 year interpolated binned data and simulated C_l

4.4 Reconstruction of features

To check whether the SVD formalism is able to recover features in the PPS if present, we introduce some features in the power law PPS. We obtain the corresponding C_l by multiplying this featured P_k with the transportation kernel G_{lk} . Using this C_l we try to reconstruct the PPS and check whether the feature added is recovered.

Fig(4.14) shows the reconstruction of PPS with Gaussian bump and dip introduced as features at low-k and high-k. Fig(4.15) shows the reconstruction of PPS with sharp cut-off introduced at low-k and high-k.

We see that SVD reconstruction is capable of capturing strong features in PPS. Further if needed we can carry out a detailed quantitative analysis as to how broad and strong and at what scales the features can be recovered.

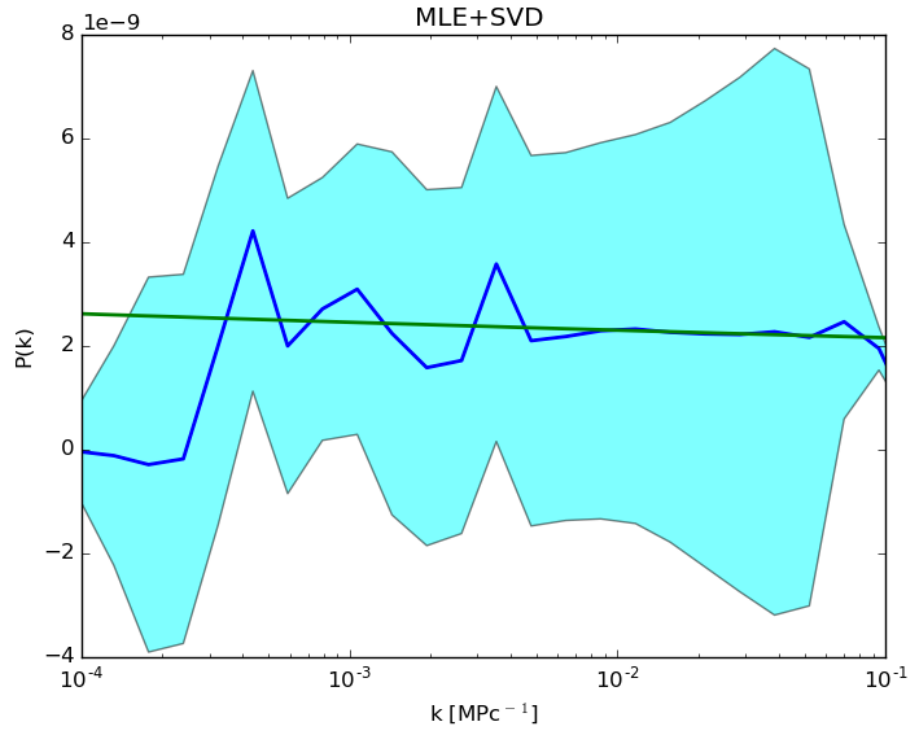


Figure 4.9: Reconstruction of power law PPS from interpolated WMAP-9 years data(all l 's) using SVD + ML after logarithmic binning of k -space(40 bins). The light blue band is the error band. Error band is estimated from ML.

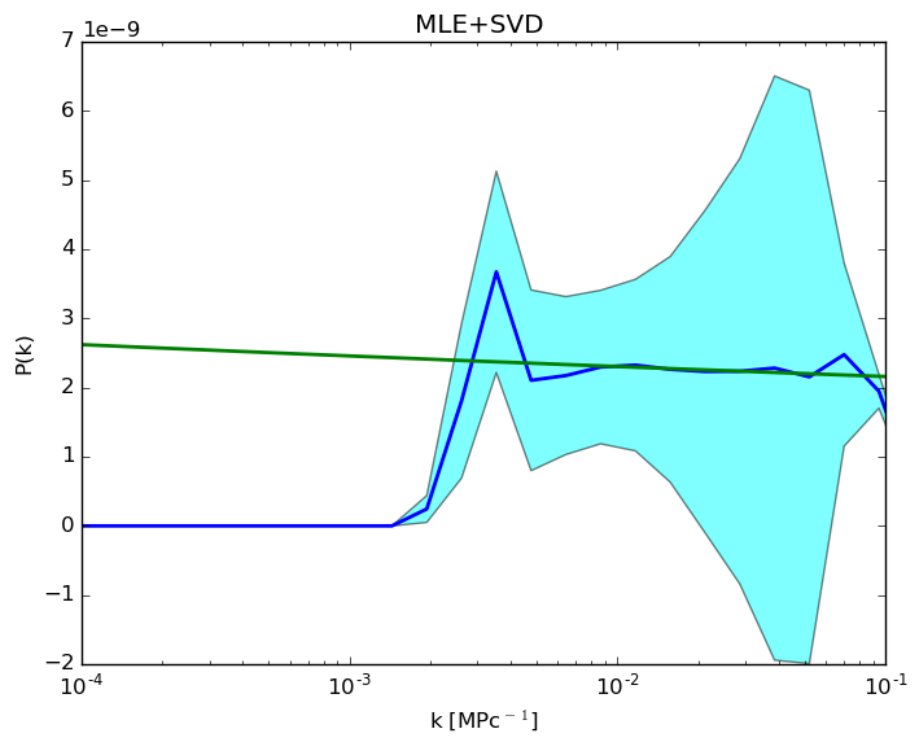


Figure 4.10: Same as in fig(4.9) but with high- l (from $l=30$) data.

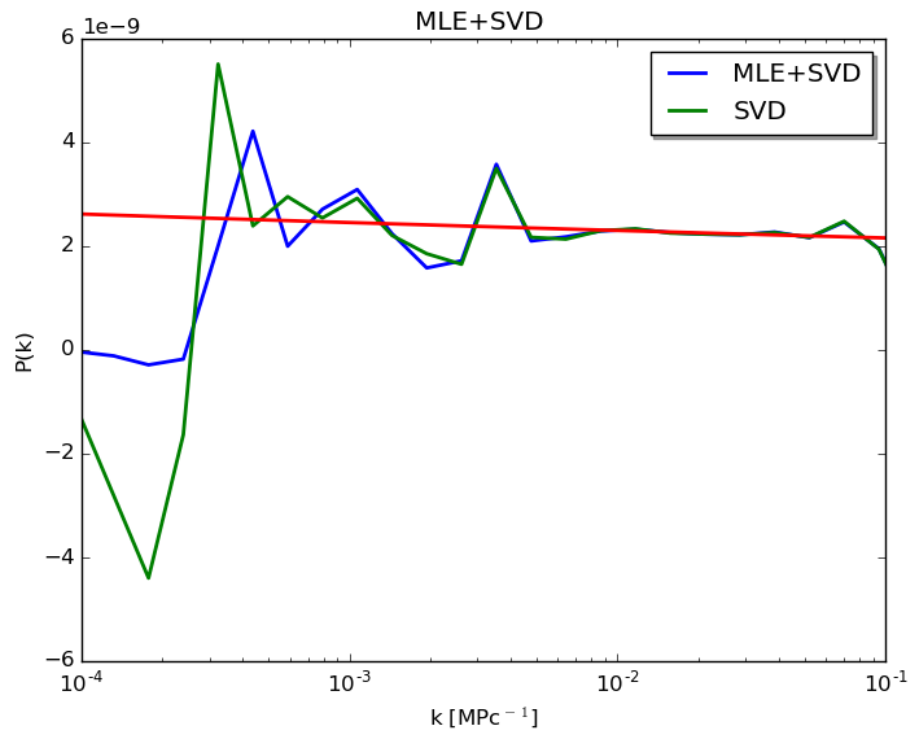


Figure 4.11: Reconstruction for WMAP 9 binned data for all l range (2-1200) with SVD and SVD+ML.

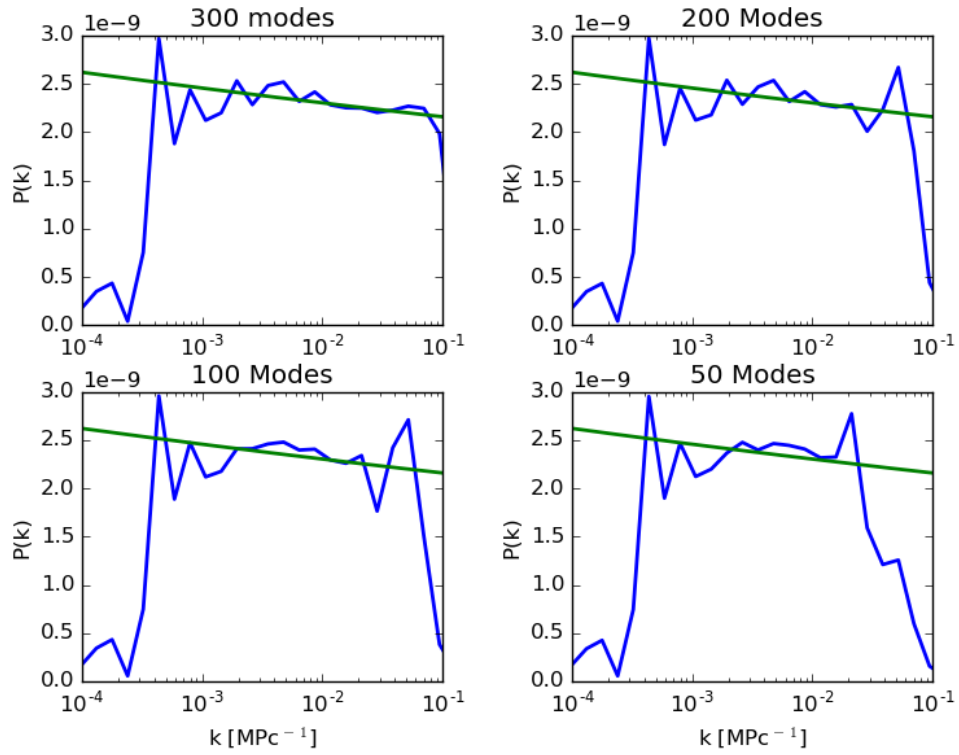


Figure 4.12: SVD reconstruction of power law PPS with different svd modes for interpolated Planck data(all l 's) after logarithmic binning of k -space(40 bins).

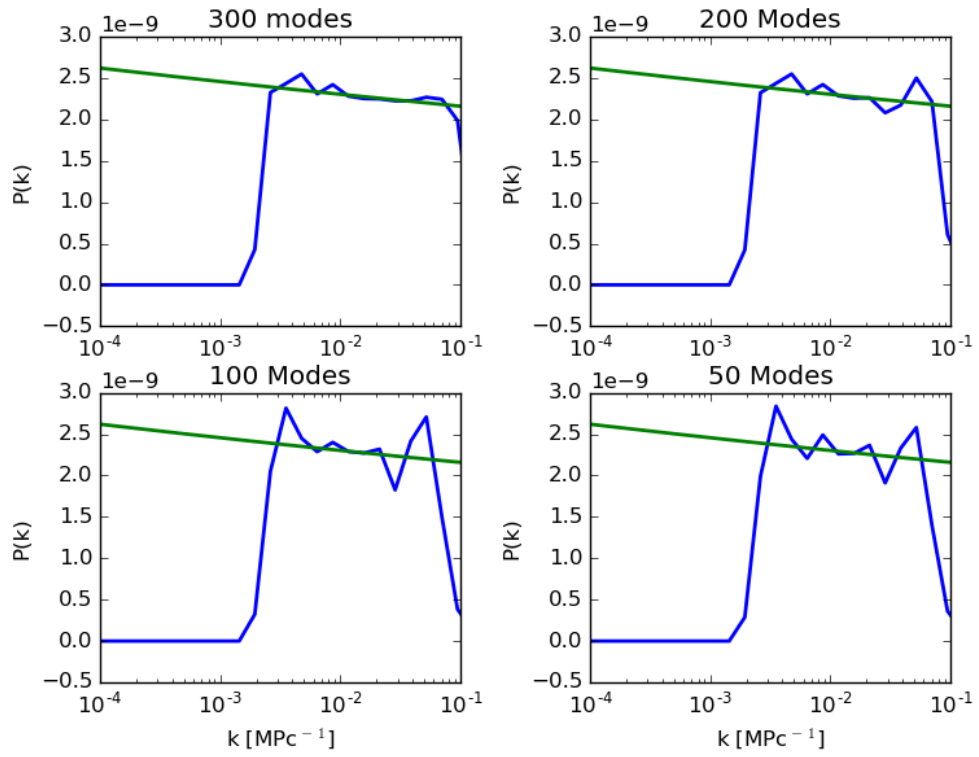


Figure 4.13: Same as in fig(4.12) but with high- l (from $l=30$) data.

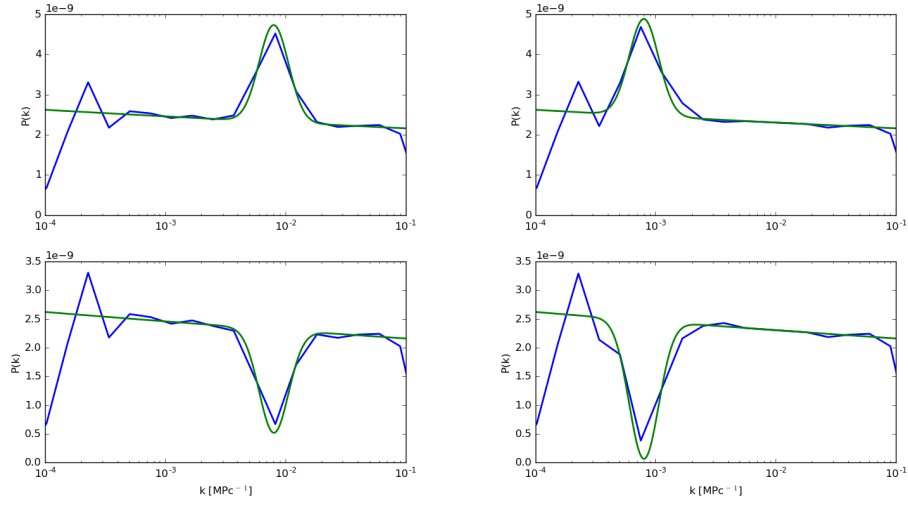


Figure 4.14: The left and right panels in the top row show the reconstruction of a primordial power spectrum having a Gaussian bump at small (high- k) and large scales (low- k) respectively, using singular value decomposition. The input PPS is shown by the green smooth line and the reconstructed by the fluctuating blue line. Width and height for both the cases are the same. The same exercise is done for the case of a Gaussian dip also in the lower panel. From this figure we can see the such features can be easily reconstructed using SVD given that they are strong.

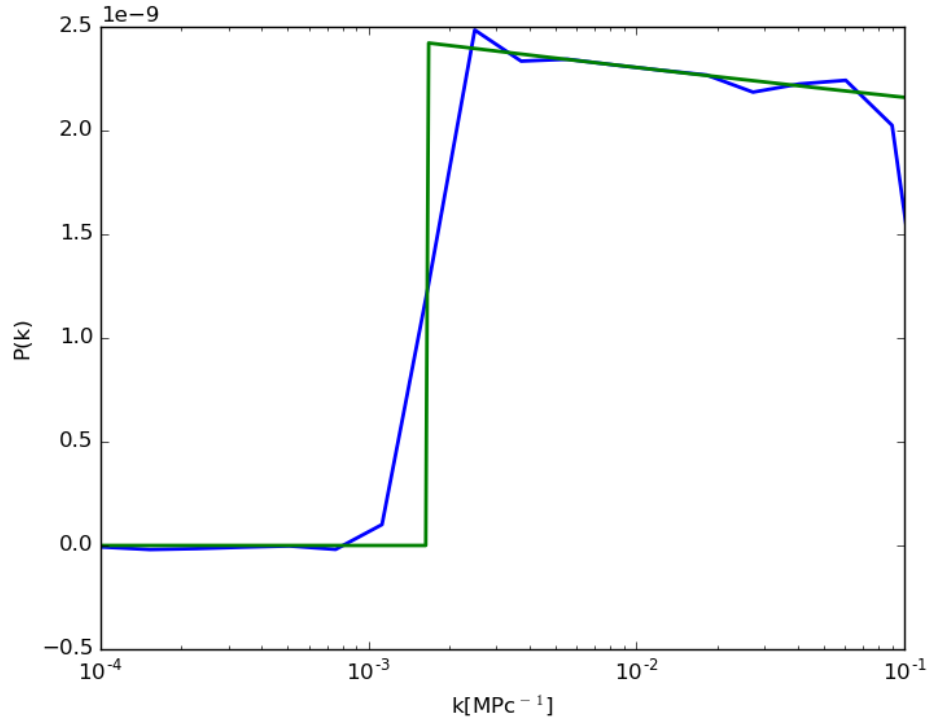


Figure 4.15: SVD reconstruction of power spectrum with a sharp cut of at large scales (low- k). The input PPS is shown by the green smooth line and the reconstructed by the fluctuating blue line. In this case also we can see the method is successful to reconstruct the features.

Chapter 5

Discussion & Conclusions

Looking at the results in the above section we can say that the PPS broadly obeys the power law. There is a sharp cut-off followed by a spike at low- k which has been observed previously in studies carried out using other methods. We still have some features at high- k and low- k and we need to carry out a study as to determine whether these features are due to the method employed or they are actually present.

Further we are working on getting better estimation of error bands on the reconstructed PPS. Now that the qualitative structure of the method is ready we can carry out more detailed quantitative studies of this method.

Bibliography

- [1] Edited by Jerome Idier, *Bayesian Approach to Inverse Problems*
- [2] A.N. Tikhnov, A.V. Goncharsky, V.V. Stepanov, A.G. Yagola, *Numerical Methods for the solution of Ill-posed Problems*
- [3] Gilbert Strang, *Linear Algebra and Its Applications, 4th Edition*
- [4] S. Dodelson, *Modern cosmology*
- [5] Wayne Hu & Martin White , Scientific American, Feb. 2004.
- [6] Charles L. Bennett, Nature 440, 1126-1131 (27 April 2006), *Cosmology from start to finish*.
- [7] Bevington, Philip R., *Data reduction and error analysis for the physical sciences*, New York: McGraw-Hill (1969).
- [8] Press, William H.; Flannery, Brian P.; Teukolsky, Saul A., *Numerical recipes. The art of scientific computing*, Cambridge: University Press (1986).
- [9] Gene H. Golub, Matrix Computations 3e (Johns Hopkins Studies in the Mathematical Sciences).
- [10] A line of sight integration approach to Cosmic Microwave Background Anisotropies Uros Seljak and Matias Zaldarriaga, astro-ph/9603033 Ap.J. 469:2 437-444, 1996

- [11] Gavin Nicholson, Carlo R. Contaldi, Paniez Paykari, JCAP (2010) **2010** 1001:016. *Reconstruction of the primordial power spectrum by direct inversion.*
- [12] Paul Hunt and Subir Sarkar, JCAP (2014) **2014** 01:025, *Reconstruction of the primordial power spectrum of curvature perturbations using multiple data sets.*
- [13] <http://background.uchicago.edu/~whu/beginners/introduction.html>
- [14] http://www.damtp.cam.ac.uk/research/gr/public/bb_home.html
- [15] <http://lambda.gsfc.nasa.gov/>
- [16] <http://pla.esac.esa.int/pla/>
- [17] https://en.wikipedia.org/wiki/Cosmic_microwave_background
- [18] <http://camb.info/>