A SPECTACULAR TITLE

Daniel Heinesen

Draft version November 22, 2017

ABSTRACT

State problem. Briefly describe method and data. Summarize main results.

Subject headings: cosmic microwave background — cosmology: observations — methods: statistical

1. INTRODUCTION

Since its discovery in 1964 *REF* the Cosmic Microwave Background (CMB) radiation has been an important field of research in cosmology. The CMB is the red shifted remains of the photons made about 380 000 years after the Big Bang, when the plasma of the primordial universe had cold down enough to be transparent and light could escape. The CMB can now be observed as a $\sim$ 2.7 K blanket of microwave radiation covering the whole sky. This radiation was at first shown to behave more or like a black body. In the 60's and 70's a theoretical basis for permutations in the primordial background radiation was found*REF*, which would lead to a small fluctuation in the observable CMB(anisotropy) on the order

$$\Delta T/T \approx 10\mu K. \tag{1}$$

These fluctuations predicted to be both isotropic and Gaussian. This is a clear testable prediction of inflationary theory.

An other prediction of the inflationary model is that these fluctuations would be nearly the same across all scales, but not entirely the same.

The Cosmic Background Explorer (COBE) was mission launched in 1989 by NASA, and collected data to 1993. The satellite was launched with three onboard instruments: the "Far-InfraRed Absolute Spectrophotometer(FIRAS)", the "Diffuse InfraRed Background Experiment(DIRBE)" and the "Differential Microwave Radiometer(DMR)". The latter being of interest in this article.

We are going to use the 53 GHz and 90 GHz maps from the COBE-DMR. We are going to use these maps to estimate the values of the fluctuations of the temperature of the CMB, given by the parameter Q; and these fluctuations dependence on scale, given by the parameter n.

We are going to use a power-law model for the power spectrum 2 to estimate a theoretical covariance matrix for data for the parameter Q and n(14). We are then going apply Bayesian analysis to this covariance matrix calculate the posterior probability that this model is correct given the data obtained from COBE-DMR. We are then going to use a Markov Chain Monte Carlo simulation to optimize these parameters to find the best possible parameters Q and n.

We are only going to look at fluctuations with modes

daniel.heinesen@sf-nett.no
[1] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, N-0315 Oslo, Norway

with an effective wavelength2 of $\ell \leq 47$, since for these mode the fluctuations are a direct product of the inflation, and thus gives clear evidence for the model. Since COBE-DMR only measures the difference between points in the sky, there is a difficulty in measuring monopoles. It also faces problems with dipoles, since motions trough space causes red-shifts in monopoles, manifesting as dipoles; thus making observations of dipoles more likely to originate from this effect than from actual dipoles. Thus we are only looking at multipoles of order $2 \leq \ell \leq 47$.

Should we be able to values for Q and n is line with the prediction of the inflationary model, this would be an important confirmation of the model, and would go a long way of cementing its role as the leading model of the expansion of the universe.

2. BACKGROUND THEORY

The fluctuations in the CMB can be described with the multipole expansion

$$\Delta T(\hat{n}) = \sum_{\ell=0}^{\ell_{max}} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}, \tag{2}$$

where $a_{\ell m}$ is the amplitude of the different modes of the spherical harmonics $Y_{\ell m}$. ℓ is the effective wavelength $\ell = 180/\lambda$, and m is the phase. From this we can find the power spectrum defined though the variance of $a_{\ell m}$

$$\langle a_{\ell m} a_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell. \tag{3}$$

Where $C_\ell = |a_\ell|^2$ is the power spectrum. This is where our investigation starts. Inflationary theory predicts that in the areas mapped out by COBE-DMR, this power spectrum should follow power law of the form

$$P(k) \propto k^n, \tag{4}$$

where n should be a bit smaller than 1. This makes it possible to parametrize the power spectrum with the parameters Q and n, where Q is the amplitudes of the fluctuations, and n is the spectral index and describes the fluctuations dependence on the scale.

3. METHOD

The method described here is a taken from *REF OPP-GAVETEKST*. For a more detailed description we refer you to this paper *EVT SAME REF*.

3.1. *Data Model*

The data from the COBE-DRM is on the form

$$d(\hat{n}) = s(\hat{n}) + f(\hat{n}) + n(\hat{n}), \qquad (5)$$

where f is the foreground radiation, n is the instrumental noise and s is the actual signal from the CMB.

As mentioned in the introduction, the fluctuations can be taken as Gaussian, so we can expect that the distribution of the real signal *DOBBELSJEKK* in the data goes like

$$p(\mathbf{d}) = \frac{1}{\sqrt{|\mathbf{C}|}} e^{-\frac{1}{2}\mathbf{d}^T \mathbf{C}^{-1}\mathbf{d}}. \qquad (6)$$

Here $\mathbf{d}$ is the observed data, and $\mathbf{C}$ the covariance matrix for the pixels in the data. The covariance matrix is a quantity that describes how much the signal for each pixel is correlated to the others. The covariance matrix contains all the relevant information of the data we need. We expect that s,f and n is non-correlated, and we can therefore write the covariance matrix for the data d as

$$\mathbf{C} = \langle \mathbf{d}\mathbf{d}^T \rangle = \mathbf{S} + \mathbf{N} + \mathbf{F}, \qquad (7)$$

where $\mathbf{S}$, $\mathbf{N}$ and $\mathbf{F}$ are the covariance matrices from s, n and f respectively. We so need to find expression for these matrices.

We'll start with the foreground. We want to marginalize over the the dipole and monopole so not to get any contributions from these. This is done with a map of the structures we want to ignore $\mathbf{f}$ and a large number λ:

$$\mathbf{F} = \lambda \mathbf{f}\mathbf{f}^T. \qquad (8)$$

This gives these regions near infinite variance, and is thus ignored.

We do not expect the instrumental noise in the different pixels to be correlated to one another, so N will be a diagonal matrix defined as

$$\mathbf{N} = \langle n_i n_j \rangle = \sigma_i^2 \delta_{ij}. \qquad (9)$$

The last part of the covariance matrix is the actual CMB signal: $\mathbf{S}$. *REF* shows that this can be written as

$$S_i j = \frac{1}{4\pi}\sum_{\ell=0}(2\ell+1)(b_\ell p_\ell)^2 C_\ell P_\ell(\cos\theta_{ij}), \qquad (10)$$

where b_ℓ is the instrumental beam, p_ℓ is the pixel window *REF DATA*, P_ℓ are Legendre polynomials and θ_{ij} is the angle between two pixels i and j. The $(b_\ell p_\ell)^2$ is there to correct for the beam and the pixel window, and will be discussed more in the *data* section 4. C_ℓ refers to the power spectrum2

$$C_\ell = |a_\ell|^2. \qquad (11)$$

As mentioned in the theory section 2 we use the power law model of the power spectrum predicted for the inflation model, with $P(k) \propto k^n$. In this model it can be shown*REF* that C_ℓ can be given as

$$C_\ell = \frac{4\pi}{5}Q^2 \frac{\Gamma\left(\ell+\frac{n-1}{2}\right)\Gamma\left(\frac{9-n}{2}\right)}{\Gamma\left(\ell+\frac{5-n}{2}\right)\Gamma\left(\frac{3+n}{2}\right)}. \qquad (12)$$

To ease the numerical calculation we write this recursively

$$C_{\ell+1} = C_\ell \frac{2\ell+n-1}{2\ell+5-n}. \qquad (13)$$

Since we are not interested in the mono- and dipole we let $C_0 = C_1 = 0$, and for simplicity we let $C_2 = 4\pi/5Q^2$.

We see that this model for the power spectrum is parametrized the amplitude Q and the spectral index n. This also make the covariance matrix*REF COV MAT* and the probability *REF PROB* dependent on these parameters:

$$\mathbf{C}(Q, n) = \mathbf{S}(Q, n) + \mathbf{N} + \mathbf{F}. \qquad (14)$$

3.2. *Likelihood and Markov Chain Mote Carlo Methods*

We now have a modelled covariance matrix dependent on Q and n. We are interested in finding the values Q and n which gives a power spectrum that represent the observed data. For this we look at the posterior probability

$$P(Q, n|\mathbf{d}) = \mathcal{L}(Q, n)P(Q, n), \qquad (15)$$

where $P(Q, n)$ is an unknown prior probability *DOUBLE CHECK* and $\mathcal{L}$ is the likelihood

$$\mathcal{L}(Q, n) = p(\mathbf{d}|Q, n), \qquad (16)$$

where $p(\mathbf{d}|Q, m)$ is the conditional probability that we will observe our data $\mathbf{d}$ given the parameters Q and n. We want to maximize this quantity with respect to Q and n to get the best posterior probability.

Due to numerical limits of our numerical methods, we are not interested in $\mathcal{L}$ directly, but rather its logarithm. From *REF HANS-KRISTIAN* we get that this is given as

$$-2\log\mathcal{L}(Q, n) = \mathbf{d}^T \mathbf{C}^{-1}\mathbf{d} + \log|\mathbf{C}|. \qquad (17)$$

So for the different values of Q and n we need to calculate $\log\mathcal{L}$, and since $\mathbf{C}$ is large matrix, this can be computationally heavy. To make everything more faster, we do a Cholesky decomposition of the covariance matrix:

$$\mathbf{C} = \mathbf{L}\mathbf{L}^T. \qquad (18)$$

This makes the inversion of $\mathbf{C}$ easier, since

$$\mathbf{d}^T \mathbf{C}^{-1}\mathbf{d} = \mathbf{x}^T\mathbf{x} \Leftrightarrow \mathbf{x} = \mathbf{L}^{-1}\mathbf{d}. \qquad (19)$$

This is because solving the triangular matrix system $\mathbf{L}\mathbf{d} = \mathbf{x}$ is much faster than inverting $\mathbf{C}$ directly. The triangular matrix $\mathbf{L}$ also gives us the advantage at

$$\log|\mathbf{C}| = \log|\mathbf{L}\mathbf{L}^T| = 2\sum_i \log L_{ii}. \qquad (20)$$

With a effective method of finding $\log\mathcal{L}$ we need to maximize it, thus finding the best posterior probability. To do this we are using a Markov Chain Monte Carlo method, namely the Metropolis-Hastings algorithm *REF BOOK*[2], with the acceptance probability

$$P(Q_i, n_i) = \min\left(1, e^{\mathcal{L}(Q_i,n_i)-\mathcal{L}(Q_{i-1},n_{i-1})}\right). \qquad (21)$$

But due to $\mathcal{L}(Q_i, n_i) - \mathcal{L}(Q_{i-1}, n_{i-1})$ being able to take from very small to very large values, the exponent of

---

[2] To be more precise: Since the probability distribution is symmetric, the Markov chain is reversible. Our proposal distribution is thus the same backwards and forward, and the algorithm reduces to a normal Metropolis algorithm.

this is numerically dangerous. We instead calculate this probability as

$$P(Q_i, n_i) = e^p \tag{22}$$

where

$$p = \min\left(0, \mathcal{L}(Q_i, n_i) - \mathcal{L}(Q_{i-1}, n_{i-1})\right). \tag{23}$$

This is equivalent, but way safer numerically. With this acceptance probability we implement the algorithm

- Choose initial guess for Q and n

- loop

  - move Q and n a step length times a uniformly random number
  - calculate the acceptance probability: (23) and (22)
  - draw a uniformly distributed number $\mathbf{x} \in [0, 1]$
  - if $\mathbf{x} < P(Q_i, n_i)$
    * save Q_i and n_i
  - else
    * save the previous Q_{i-1} and n_{i-1} again

This algorithm will randomly walk through the posterior probability, landing more often on the Q's and n's with a higher probability, and will therefore quickly converge towards the most likely Q and n.

4. DATA

In this article the 53 and 90 GHZ maps from the COBE-DMR were used. We used these maps since they gives the cleanest data, with the lowest noise levels. The original CMB data consisted of 3072 data points, each having a corresponding direction vector. 1131 data points was masked out to remove foreground contamination, leaving us with 1941 usable data points.

ℓ was used in an interval of 0 to 47. To remove mono- and dipole only $C_\ell \geq 2$ was given a given a non-zero value.

COBE-DMR used a instrumental beam width of $7°$, meaning that the signal get smeared out/convoluted by a Gaussian function with FWHM of $7°$. Signals with wavelength smaller than this will get convoluted away, thus the data loses accuracy at this scale.*DOBBELSJEKK*

COBE-DMR also have a pixel window of $220'$ or $3.7°$. Details smaller than this will not registered by the pixels and therefore lost instrument. This therefore have some of the same effects as the beam.

The instrumental noise used in the calculation of the covariance matrix(14) was supplied by the COBE mission, and thus not calculated by us.

5. RESULTS

The Monte Carlo simulations run with $2 \cdot 10^4$ Metropolis iterations and take around about 80 minutes. The result of simulation for 53 and 90 GHz are summarized in table 1.

The distribution plots are made with a bin size of 250. This is done to lump close values of Q and n together. So the number of possible values they are allowed to have

(a) Posterior distribution for 53 GHz

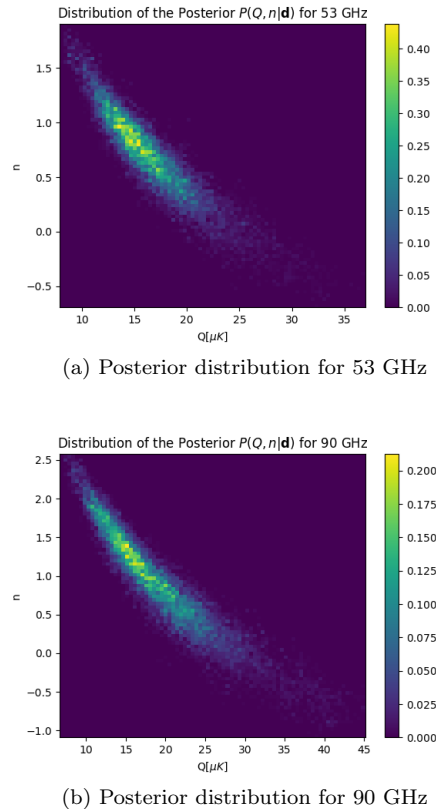(b) Posterior distribution for 90 GHz

FIG. 1.— Posterior distribution for the amplitude Q and the tilt n, found with $2 \cdot 10^4$ Metropolis iteration. The heat maps were made with a bin size of 250. Due to the relatively low number of iteration, the data is quite sparse and the contour plot fragmented.

	53 GHz	90 GHz
n	0.696 ± 0.381	0.915 ± 0.623
$Q[\mu K]$	16.3 ± 3.81	18.9 ± 6.12

TABLE 1
RESULTS FOR THE MONTE CARLO SIMULATION WITH $2 \cdot 10^4$ METROPOLIS ITERATION. THE RESULTS ARE ON THE FORM $\mu \pm \sigma$ AND ARE TAKEN DIRECTLY FROM THE VALUES WE GET FROM THE RANDOM WALK, AND IS NOT CALCULATED FROM THE DISTRIBUTION IN FIG. 2

is #Metropolis iterations/Bin Size. For higher bin sizes the trends are easier to see, but the fine details disappears, and for low bin sizes the details become prominent but trends are difficult to see. A bin size of 250 seemed like the best compromise.

The values of Q and n in table 1 are taken from the raw data, and are not affected by the bin size.

6. CONCLUSIONS

Who do you want to thank for helping out with this project?

REFERENCES

Górski, K. M., Hinshaw, G., Banday, A. J., Bennett, C. L., Wright, E. L., Kogut, A., Smoot, G. F., and Lubin, P. 1994, ApJL, 430, 89
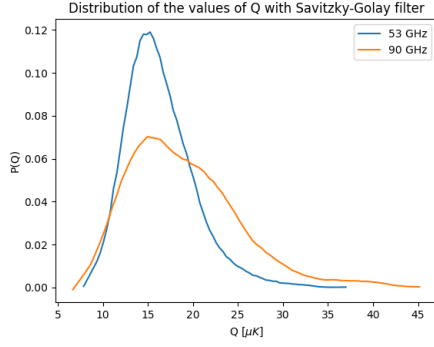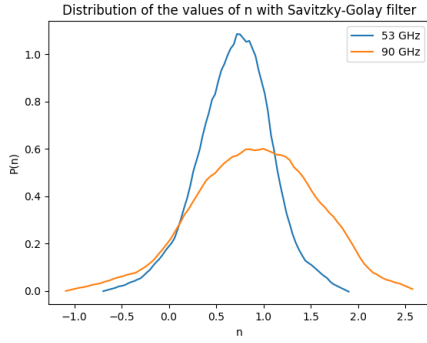
(a) Posterior distribution $P(Q|\mathbf{d})$

(b) Posterior distribution $P(n|\mathbf{d})$

FIG. 2.— Posterior distribution $P(n|\mathbf{d})$ and $P(Q|\mathbf{d})$ with $2 \cdot 10^4$ Metropolis iteration and a bin size of 250. Due to the few Metropolis iterations, the raw results are quite messy, so a Savitzky-Golay filter is used to smooth the data with out loss of generality *SKJEKK DETTE!!!!*