A SPECTACULAR TITLE

DANIEL HEINESEN

Draft version November 20, 2017

ABSTRACT

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

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

1. INTRODUCTION

Discuss background, physical importance and possibly some history of the problem that is being studied in this paper.

2. METHOD

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

2.1. *Data Model*

CMB can be well approximated as a Gaussian distribution *REF*

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

Here $\mathbf{d}$ is the observed data, and $\mathbf{C}$ the covariance matrix for the pixels in the CMB field.

We are taking the observed CMB data $d(\hat{n})$ to be consisting of three parts

$$d(\hat{n}) = s(\hat{n}) + f(\hat{n}) + n(\hat{n}), \tag{2}$$

where f is the foreground radiation, n is the noise and s is the actual signal from the CMB. We assume that non of these are correlated, and we can therefore write the covariance matrix for the data d as

$$\mathbf{C} = \langle \mathbf{d}\mathbf{d}^T \rangle = \mathbf{S} + \mathbf{N} + \mathbf{F}, \tag{3}$$

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 start with the foreground. We want to marginalize over the the dipole and monopole so to not 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:

$$\mathbf{F} = \lambda \mathbf{f}\mathbf{f}^T. \tag{4}$$

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

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

$$\mathbf{N} = \langle n_i n_j \rangle = \sigma_i^2 \delta_{ij}. \tag{5}$$

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

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

$$S_i j = \frac{1}{4\pi} \sum_{\ell=0} (2\ell+1)(b_\ell p_\ell)^2 C_\ell P_\ell(\cos\theta_{ij}), \tag{6}$$

where b_ℓ is the instrumental beam, p_ℓ is the pixel window *REF EVT THEORY*, P_ℓ are Legendre polynomials and θ_{ij} is the angle between two pixels i and j. C_ℓ refers to the power spectrum

$$C_\ell = |a_\ell|^2. \tag{7}$$

We use the power law model of the power spectrum given in *REF*, with $P(k) \propto k^n$. In this model it can be shown that C_ℓ can be given as

$$C_\ell = \frac{4\pi}{5} Q^2 \frac{\Gamma\left(\ell + \frac{n-1}{2}\right)\Gamma\left(\frac{9-n}{2}\right)}{\Gamma\left(\ell + \frac{5-n}{2}\right)\Gamma\left(\frac{3+n}{2}\right)}. \tag{8}$$

To ease the numerical calculation we write this recursively

$$C_{\ell+1} = C_\ell \frac{2\ell + n - 1}{2\ell + 5 - n}. \tag{9}$$

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.

2.2. *Likelihood and Markov Chain Mote Carlo Methods*

We are interested in finding the parameters Q and n which gives a power spectrum that represent the observed data. We are looking for the posterior probability

$$P(Q, n|\mathbf{d}) = \mathcal{L}(Q, n) P(Q, n), \tag{10}$$

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

$$\mathcal{L}(Q, n) = p(\mathbf{d}|Q, n), \tag{11}$$

where $p(\mathbf{d}|Q, m)$ is the conditional probability that we are going to we will observe our data $\mathbf{d}$ given the parameters Q and n. Due to numerical limits of our numerical methods, we are not interested in $\mathcal{L}$ directly, but rather its logarithm, given as *REF HANS-KRISTIAN*

$$-2\log\mathcal{L}(Q, n) = \mathbf{d}^T\mathbf{C}^{-1}\mathbf{d} + \log|\mathbf{C}|. \tag{12}$$

To make everything more computationally faster, we do a Cholesky decomposition of the covariance matrix:

$$\mathbf{C} = \mathbf{L}\mathbf{L}^T. \tag{13}$$

This makes the inversion easier

$$\mathbf{d}^T\mathbf{C}^{-1}\mathbf{d} = \mathbf{x}^T\mathbf{x} \Leftrightarrow \mathbf{x} = \mathbf{L}^{-1}\mathbf{d}, \tag{14}$$

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

$$\log|\mathbf{C}| = \log|\mathbf{L}\mathbf{L}^T| = 2\sum_i \log L_{ii}. \tag{15}$$

To find this distribution *REF DIST* we are using a Markov Chain Monte Carlo method, namely the Metropolis-Hastings algorithm *REF BOOK*[2], with the acceptance probability

$$P(Q_i, n_i) = \min\left(1, e^{\mathcal{L}(Q_i,n_i)-\mathcal{L}(Q_{i-1},n_{i-1})}\right). \tag{16}$$

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 this is numerically dangerous. We instead calculate this probability as

$$P(Q_i, n_i) = e^p \tag{17}$$

where

$$p = \min\left(0, \mathcal{L}(Q_i, n_i) - \mathcal{L}(Q_{i-1}, n_{i-1})\right) \tag{18}$$

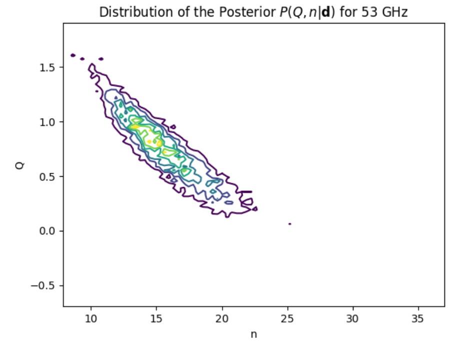
This is equivalent, but way safer numerically. With this acceptance probability we implement the algorithm

- Choose initial guess for Q and n
- loop
  - move Q and n a step length times a uniformly random number
  - calculate the acceptance probability (17)
  - draw a uniformly distributed number $\mathbf{x}$
  - 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

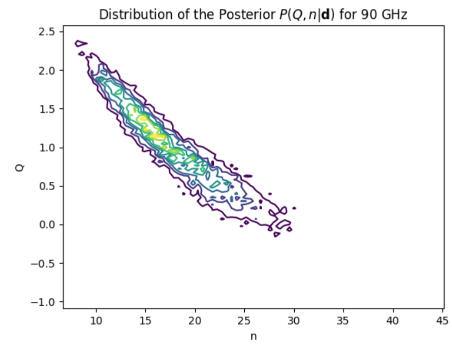
This algorithm will quickly converge towards the most likely Q and n.

3. DATA

In this article data from the COBE satellite is used. We look at 53 and 90 GHZ map, since they gives the cleanest data. For the collection of the data a 7° beam was used. The foreground was masked out, so not to contaminate the data with noise from stars and galactic dust. ℓ was used in an interval of 0 to 47.



(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, and the contour plots made with a resolution 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

4. 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.

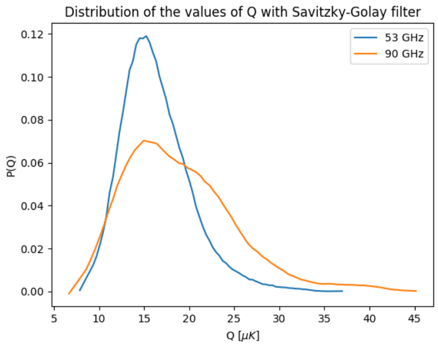
5. CONCLUSIONS

Summarize results. Discuss their importance, referring to the discovery to the initial seeds for structure formation. Mention that these results are in good agreement with expectations from inflationary theory.

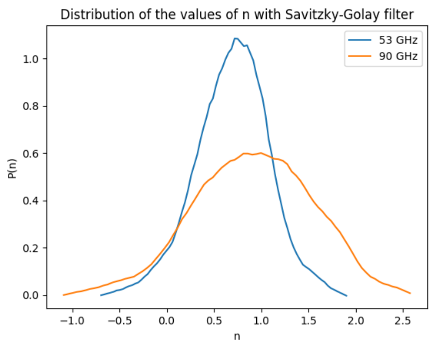
Who do you want to thank for helping out with this project?

REFERENCES

[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.



(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 resolution 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!!!!*

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