Bayesian Cosmological inference beyond statistical isotropy

Tarun Souradeep[1], Santanu Das[1] and Benjamin Wandelt[2]

1. IUCAA, Post Bag 4 Ganeshkhind, Pune 411007 India
2. IAP, Lagrange Institute, Sorbonne University, Paris

Abstract.   With advent of rich data sets, computationally challenge of inference in cosmology has relied on stochastic sampling method. First, I review the widely used MCMC approach used to infer cosmological parameters and present a adaptive improved implementation SCoPE developed by our group. Next, I present a general method for Bayesian inference of the underlying covariance structure of random fields on a sphere. We employ the Bipolar Spherical Harmonic (BipoSH) representation of general covariance structure on the sphere. We illustrate the efficacy of the method with a principled approach to assess violation of statistical isotropy (SI) in the sky maps of Cosmic Microwave Background (CMB) fluctuations. The general, principled, approach to a Bayesian inference of the covariance structure in a random field on a sphere presented here has huge potential for application to other many aspects of cosmology and astronomy, as well as, more distant areas of research like geosciences and climate modelling.

1. Introduction

With the different experimental data the cosmology has become a play room for the data analysts. In recent past we get data from WMAP, Planck, BICEP, BOSS etc and there are many new experiments coming up. We need to analyse the high volume of data and draw inferences from the data. For many data analysis process we require advanced numerical techniques. In this proceeding we discuss two cosmological parameter estimation methods in which we used some advanced numerical methods.

First, we demonstrate a cosmological parameter estimation for different theoretical model by comparing the C_l with the observed C_l. We develop a MCMC code, named as SCoPE [1, 2, 3, 4, 5], based on global Metropolis Hastings algorithm [6, 7, 8, 9]. In general the Metropolis Hastings algorithm is pretty slow. Therefore, we have developed few novel technique that makes the MCMC method faster. In the first section we briefly discuss about these techniques that make SCoPE faster then standard Monte Carlo methods.

In the second part of the proceeding we have discussed about a Monte Carlo method that we use for Bayesian estimation of the isotropy violation in the CMB sky. For the first time, we develop an algorithm for fast basyan estimation of the isotropy violation on the CMB sky. We use Hamiltonion Monte Carlo for the estimation process. The stability of the algorithm was a big issue there. Therefore, we have applied different integration method for testing the stability of the method. We have discussed that in the second part of the proceeding.

2. Metropolis Hastings

For cosmological parameter estimation we use Metropolis Hastings (MH) algorithm. The Metropolis-Hastings (MH) is one of the most widely used MCMC sampler, in which the posterior i.e. $\pi(\theta)$ is sampled using a random walk. A standard Markov Chain at each step i, randomly chooses a candidate value θ_{i+1} from the proposal distribution $q(.|\theta_i)$. The candidate value only depends on the current data point θ_i. The new data point is then accepted with probability $\alpha = \min(1, \pi(\theta_{i+1})/\pi(\theta_i))$. If the new data point rejected, the previous point is replicated by increasing its weight by $+1$. The chain of data-points thus generated, approximate the target posterior distribution $\pi(\theta)$. The proposal distribution is generally taken to be a Gaussian distribution i.e. $q(\theta_{i+1}|\theta_i) = N\exp(-\frac{s}{2}u_i[C_{ij}]^{-1}u_j)$, where $u_i = \theta_{i+1} - \theta_i$ and C_{ij} is the covariance matrix. s is the step size. Theoretical optimum step size for an ideal distribution that provide the best acceptance rate is $s = 2.4/\sqrt{n}$ for a n dimensional MCMC sampler. The covariance matrix is provided as an input to the program. As the exact covariance matrix is unknown before the analysis, in practice an approximate covariance matrix, often based on some previous analysis, is provided. If no prior information is available about the covariance between parameters then some approximate diagonal matrix is also often used. However, in such cases the acceptance rate of the sampler may reduce drastically and can be ensured to remain reasonable only by trial and error. Therefore a prior knowledge of the covariance is required. Parallelization of MH sampler is generally done by running multiple chains. Whether it is better to run a longer chain than running multiple short chains has been addressed and argued by many authors [10, 11]. But, for running multiple parallel chains proper mixing between the chains has to be ensured. Therefore, each of the multiple chains has to be long enough so that the it can represent an unbiased sample of the population. Gelman-Rubin "R" statistics is generally used for testing the mixing of chains. For convergence, the chains have to be long enough such that R is very close to unity[12]. For practical purposes it is taken as $R < 1.2$. However, this criterion though necessary is often not sufficient for ensuring proper sampling. It is desirable to devise an implementation of MH algorithm that allows the individual chains to run in parallel and increase the acceptance rate of the models of a chain. Apart from that the mixing of the chains are also necessary. Therefore in SCoPE we have modified standard MH algorithm to accomplish effective parallelization through prefetching together with all other above mentioned features, namely, enhanced acceptance, regular covariance update from samples.

2.1. Embellishing the standard Metropolis-Hastings algorithm
2.1.1. Prefetching   In MCMC it is extremely useful to speed up the generation of a single chain, through parallelization rather than using multiple chains. In our work we make the individual chains parallel by precomputing several draws from the posterior distribution ahead of time via multiple evolution of models simultaneously in parallel and then use only the values that are needed. Prefetching is a draw level parallelization in a single chain. The method can be explained by taking the binary tree of a Metropolis algorithm as shown in Fig. 1. In a kth level binary tree there are total $2^k - 1$ nodes, each of which represents a possible future state of a metropolis algorithm. The branches at the left child of any node represent the accepted steps and the right child represents the rejected states. If we have enough computational resources then all $2^k - 1$ nodes can be evaluated simultaneously and k steps of a MCMC chain can be carried out in parallel simultaneously. Though the method of prefetching allows to parallelize a single chain, it only uses k steps out of $2^k - 1$ computations. The rest of the computations are not utilized. Therefore, many argue against computing all the nodes of the binary tree. Rather if we know the acceptance rate at a point of time from the previously accepted data points, we can statistically identify and precompute only the most probable chain and hence avoid any unnecessary wastage of computation power. It is easy to see that if the acceptance probability at any point of time is less than 0.5 then the extreme right chain (1-3-7-15-..) of Fig. 1 will
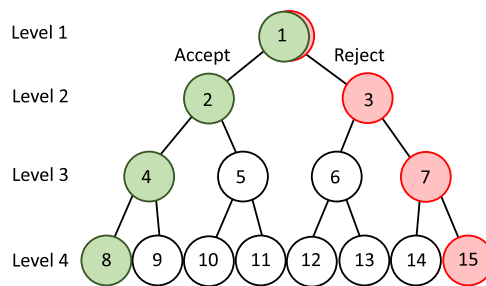
Figure 1. Prefetching scheme is explained in the text with the help of above figure

be most probable chain. In a similar manner if the acceptance rate is more than 0.5 then the extreme left chain (1-2-4-8-..) will be the most probable chain [13]. Therefore, by pre-evoluting only the most probable chain, we parallelize the code and at the same time we can manage the computational resources in a better way.

2.1.2. Delayed rejection   One of the major problems with the MCMC method is the choice of the step size for the proposal distribution. It will be better idea if the rejected sample from one step can be used to determine the proposal distribution for the next sample by varying stepsize accordingly. This increases acceptance rate but at a cost of violation of the Markovian property. But if we can find some method that can change the acceptance probability of the sample point to compensate the step size variation then that will be useful. The concept of delayed rejection can briefly be explained as follows. Suppose at some step i, the position of a chain is $\theta_i = x$. Suppose at this time a candidate y_1 is accepted from $q_1(x, y_1)$ and accepted with probability $\alpha_1(x, y_1) = \min(1, \pi(y_1)q_1(y_1, x)/\pi(x)q_1(x, y_1))$ as in the standard MH algorithm. For a Markov chain, $q_1(y_1, x)$ is time symmetric, i.e. $q_1(y_1, x) = q_1(x, y_1)$. Therefore, the acceptance ratio only depends on the posterior. A rejection at any step suggests that there is a local bad fit of the correct proposal and a better one, $q_2(x, y_1, y_2)$, can be constructed in light of this. But, in order to maintain the same stationary distribution the acceptance probability of the new candidate, y_2, has to be properly computed. A possible way to reach this goal is to impose detailed balance separately at each stage and derive the acceptance probability that preserves it. It can be done by changing the acceptance probability as [14, 15, 16, 17, 18, 19]

$$\alpha_2(x, y_1, y_2) = \min\left(1, \frac{\pi(y_2)q_1(y_2, y_1)q_2(y_2, y_1, x)\left[1 - \alpha_1(y_2, y_1)\right]}{\pi(x)q_1(x, y_1)q_2(x, y_1, y_2)\left[1 - \alpha_1(x, y_1)\right]}\right). \tag{1}$$

The Markovian property of the chain will not get destroyed, but still the sample choice can be made dependent on the previously accepted data point. We implement this in our algorithm.

2.1.3. Inter-chain covariance adaptation   The practical problem in implementing MH is the tuning problem of the proposal distribution q so that the sampling is efficient. One of the recent improvements in the MCMC efficiency is to introduce adaptive samplers. The adaptive MCMC uses the sample history and automatically tune the proposal distribution in the sampling process [20, 21, 22, 23]. In adaptive metropolis algorithm, the covariance matrix from the samples obtained so far is used as the covariance of a Gaussian proposal. The most common parallelization scheme of MCMC method is to run parallel chains instead of running a single one. If in each chains proposal distribution is adapted using the local covariance matrix then the acceptance probability a chain may improve, however, the inter-chain mixing will not improve. If some chain stuck at some local minima then the local covariance matrix corresponding to that chain will be erroneous. So, in case of a local peak the local covariance update will give covariance

corresponding to the local peak. In that case, the mixing of chains will slow down and sampling may not be proper. Therefore, we have adapted the concept of the inter-chain covariance update in the adaptation technique. We run several parallel chains, and randomly update the covariance matrix taking the data points accepted till then from all the chains. This means we have used the covariance after ith accepted step as $\text{Cov}(\theta_{11}.., \theta_{i1}, \theta_{12}, .., \theta_{i2}....\theta_{1n}, ..., \theta_{in})$, where, n is the number of chains. This speeds up the mixing of the chains and covers the sample space faster. The value of the covariance matrix will freeze after few adaptations and hence we will be using same Gaussian proposal after few steps, which is important to guarantee proper sampling. However, if the adaptive covariance is not frozen, it may give rise to unfair sampling as the Gaussian proposal will vary between steps. Therefore, the process of the adaptive covariance calculation is only used for the initial burn in process and after that the adaptation is stopped, during which the covariance calculation attains partial convergence.

3. Estimation of Isotrypy violation with Hamiltonion Monte Carlo

The previous section we briefly demonstrate how the standard MH algorithm can be modified to make a faster parameter estimation algorithm. However, MH algorithm is mainly applicable when the number of parameters is small, of order of 10's. However, if the parameter space is high of the order of 10^6-10^7 then MH is not at all efficient. To explore such a big parameter space we need to use Gibbs sampling or Hamiltonian Monte Carlo (HMC) etc. In the problem of estimating isotropy violation in CMB we use HMC [24, 25, 26, 27].

In CMB the observed sky map is a convolution of the real sky temperature with the instrumental beam with an addition of instrumental noise. Therefore, $\tilde{T}(\gamma)$ the actual temperature signal of the CMB sky along the direction γ is linearly related to the observed sky temperature, $\tilde{d}(\gamma)$, as

$$\tilde{d}(\gamma) = \hat{B}(\gamma, \gamma_0)\tilde{T}(\gamma_0)d\Omega\gamma_0 + \tilde{N}(\gamma) \tag{2}$$

where $B(\gamma, \gamma_0)$ is the instrumental beam profile and $\tilde{N}(\gamma)$ is the instrumental noise. For a perfectly circular beam profile, $B(\gamma, \gamma_0) \equiv B(\gamma \cdot \gamma_0)$, assumed in this work, it is easy to deconvolve the effect of the beam after inferring the power spectra. However, if the beam is not circular symmetric then the effect of the beam depends on the full scan pattern of the experiment and its deconvolution may be non-trivial [28, 29, 30, 31, 32]. For data defined on a sphere, it is convenient to work in the spherical harmonic space. The CMB signal, $\tilde{T}(\gamma)$, then can be expanded in terms of spherical harmonics as

$$\tilde{T}(\gamma) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\gamma), \tag{3}$$

where $Y_{lm}(\gamma)$ are the spherical harmonic functions and a_{lm} are the coefficients in the spherical harmonic basis. Similarly, the observed data, $\tilde{d}(\gamma)$, can also be expanded in spherical harmonics with coefficients, d_{lm}. For a perfectly statistically isotropic sky, the two point correlation function on sky can be expressed in terms of the angular power spectrum, C_l, alone as

$$\langle a_{lm} a_{l'm'}^* \rangle = S_{lml'm'} = C_l \delta_{ll'} \delta_{mm'} \tag{4}$$

Here $\langle \ldots \rangle$ denotes the ensemble average. However, in presence of SI violation, C_l does not provide a full description of the covariance matrix. A general covariance matrix can be expanded in the BipoSH representation as

$$\langle a_{lm} a_{l'm'}^* \rangle = \sum_{m,m'} (-1)^{m'} A_{ll'}^{LM} \mathcal{C}_{lml'-m'}^{LM}, \tag{5}$$

where $\mathcal{C}^{LM}_{lml'-m'}$ are the Clebsch Gordon coefficients and $A^{LM}_{ll'}$ are the BipoSH spectra that provide a natural generalisation of the angular power spectrum.

The observed sky map contains the noise with the CMB signal and our goal is to calculate the posterior distribution of $A^{LM}_{ll'}$ from the observed sky map, i.e. $P(A^{LM}_{ll'}|d_{lm})$ or $P(S_{lml'm'}|d_{lm})$. However, its difficult to calculate this Probability directly as there are no direct method to find out the a_{lm}s. Therefore, in practice instead of finding this distribution, people find a joint probability distribution of the $S_{lml'm'}$ and a_{lm}, i.e. $P(S_{lml'm'}, a_{lm}|d_{lm})$ and then marginalize over $a_{lm.}$. This can be obtained directly by using the Bayesian Technique

$$
\begin{aligned}
P(S_{lml'm'}, a_{lm}|d_{lm}) &= P(d_{lm}|a_{lm})P(a_{lm}|S_{lml'm'})P(S_{lml'm'}) \\
&= \frac{1}{\sqrt{|N_l|}} \exp\left[-\frac{1}{2}(d^*_{lm} - a^*_{lm})^T N^{-1}_l (d_{lm} - a_{lm})\right] \\
&\quad \times \frac{1}{\sqrt{|S_{lml'm'}|}} \exp\left(-\frac{1}{2}a^{*T}_{lm} S^{-1}_{lml'm'} a_{lm}\right) P(S_{lml'm'})
\end{aligned}
\tag{6}
$$

$P(S_{lml'm'})$ is the prior on $S_{lml'm'}$. Here we use a flat prior on $S_{lml'm'}$ i.e. $P(S_{lml'm'}) = 1$. We also consider a isotropic noise and N_l is the noise power spectrum.

We tried to sample a_{lm} amd $A^{LM}_{ll'}$ simultaneously from a given map d_{lm} and noise covariance matrix N_l. Therefore we have of the order of 10^7 variables.

3.1. Calculating the equations of motion

Hamiltonian Monte Carlo (HMC) technique based on the Classical Hamiltonian Mechanics relies on the fact that the density of a group of particles with random momenta placed in a potential will trace the potential given that all of them start from random momentum drawn from a normal distribution with mean 0 and co-variance M, where M is a positive definite matrix called the mass matrix and can be chosen independently. It is known that HMC method can sample the distribution more effectively even in very high dimensional space in comparison to other conventional MCMC methods. In a Hamiltonian Monte Carlo algorithm we need to define a conjugate momentum and a mass corresponding to each of its parameters. We consider the conjugate momentum to a_{lm} and $A^{LM}_{ll'}$ are $p_{a_{lm}}$ and $p_{A^{LM}_{ll'}}$ and a corresponding mass $m_{a_{lm}}$, $m_{A^{LM}_{l_1 l_2}}$ respectively. The mass matrices are the positive definite quantity by their definition. Thus the Hamiltonian for the motion of this ensemble of particles is

$$
H = \sum_{lm} \frac{p^2_{a_{lm}}}{2m_{a_{lm}}} + \sum_{LMll'} \frac{p^2_{A^{LM}_{ll'}}}{2m_{A^{LM}_{ll'}}} - \ln(P(S_{lml'm'}, a_{lm}|d_{lm})).
\tag{7}
$$

Using Hamiltonian mechanics

$$
\dot{a}_{lm} = p_{a_{lm}}/m_{a_{lm}}
\tag{8}
$$

and

$$
\dot{p}_{a_{lm}} = -\frac{\partial H}{\partial a_{lm}} = N^{-1}_l (d^*_{lm} - a^*_{lm}) - S^{-1}_{lml'm'} a^*_{lm},
\tag{9}
$$

Similarly, the equation of motion for $A^{LM}_{ll'}$ will be

$$
\dot{A}^{LM}_{ll'} = p_{A^{LM}_{ll'}}/m_{A^{LM}_{ll'}}
\tag{10}
$$

and

$$\dot{p}_{A_{ll'}^{LM}} = -\frac{\partial H}{\partial A_{ll'}^{LM}} = -\frac{1}{2}\partial_{A_{ll'}^{LM}} \ln |S| + \partial_{A_{ll'}^{LM}} \left(a_{lm}^* \left(S^{-1} \right)_{lml'm'} a_{l'm'} \right) \tag{11}$$

HMC is performed in two steps. In the first step, values of the momentum variables are chosen from the Gaussian distribution of mean 0 and variance m_x, where $x \in (a_{lm}, A_{ll'}^{LM})$. In the next step a Matropolis update is performed from the state $(p_{a_{lm}}, p_{A_{ll'}^{LM}}, a_{lm}, A_{ll'}^{LM})$ to a new state $(p_{a_{lm}}^1, p_{A_{ll'}^{LM}}^1, a_{lm}^1, A_{ll'}^{LM1})$ by integrating the equation equation of motions through a trajectory of length Δt. The Hamiltonian is computed in this new state and the state is accepted with probability $\min(1, \exp(-\Delta H))$, where ΔH is the change in the Hamiltonian between these two states. If the new state is accepted then similar operation is performed considering $(a_{lm}^1, A_{ll'}^{LM1})$ as the new position variable, otherwise the position is not updated from $(a_{lm}, A_{ll'}^{LM})$. Choice of $m_{a_{lm}}$ and $m_{A_{ll'}^{LM}}$ decides the stability of the integration process. HMC algorithm in general uses the Leapfrog integration algorithm due to its time reversal symmetry and almost symplectic nature. However, in our calculation we have used a fourth order Forest and Ruth integrator , which is a symplectic integrator that involves three Leapfrog steps, works better then a simple Leapfrog.

3.2. Computational implementation

3.2.1. Overcoming the time complexity
There are practical limitations arising from carrying out the computationally challenging analysis in reasonable time. The first practical issue that we face is related to inverting the covariance matrix. A brute force inversion of the matrix is computationally prohibitive. The inversion of $\left(S^{-1} \right)_{lml'm'} a_{l'm'}$ can be done by using Gauss Seidel method. Brute force inversion of $S_{lml'm'}$ is difficult and time consuming except for small l_{max} . However, using the fact that, in case of CMB signal the off-diagonal components of the matrix $S_{lml'm'}$ are expected to be much smaller than the diagonal components dominated by A_{ll}^{00} , we can use Taylor series expansion to invert the matrix. Taylor series up to the first order gives us considerably good results.

3.2.2. Stability of numerical integration
Another computational issue is the choice of the numerical integration method and the mass matrix. In Hamiltonian integrators, though Leapfrog integrator is common because the integrator preserves the Hamiltonian in phase space (symplectic), the propagation error being huge we have to use a fourth order symplectic integrator, namely Forest and Ruth integrator [25, 33], which performs better and the propagational errors are contained at a manageable level. Forest-Ruth algorithm is a combination of three Leapfrog steps. It can be shown that for the absolute stability of the integration process we need to take the mass matrix of the particles to the inverse of the covariance matric of the corresponding quantities [24].

4. Conclusion
In this proceeding we discuss two MCMC techniques that we use for cosmological parameter estimation and discuss the techniques which we use to make the algorithm faster. In our modified global Metropolis Hastings algorithm, SCoPE, the individual chains can run in parallel and a rejected sample can be used to locally modify the proposal distribution without violating the Markovian property. The latter increases the acceptance probability of the samples in chains. The prefetching algorithm allows us to increase the acceptance probability as much as required, provided requisite number of multiple cores are available in the computer. Apart from these, due to the introduction inter-chain covariance update the code can start without specifying any input covariance matrix. The mixing of the chains is also faster in SCoPE.

In the second part of the proceeding we discuss about a general method to infer the underlying covariance structure of a CMB sky using a completely Bayesian technique. We employ the Bipolar Spherical Harmonic representation of the general covariance matrix underlying random fields on a sphere and outline the method for a fully Bayesian inference of the angular power spectrum and the BipoSH coefficients simultaneously from a single observed map. We use Hamiltonian Monte Carlo for sampling the posterior distributions of the BipoSH parametrization of the covariance. We also discuss the time complexity and choice of the mass matrix and integration method for the HMC algorithm that can provide stable numerical integration.

References

[1] S. Das and T.Souradeep, SCoPE: An efficient method of Cosmological Parameter Estimation JCAP, 2014, 1407, 018
[2] S. Das, S. Mukherjee and T. Souradeep, Revised cosmological parameters after BICEP 2 and BOSS JCAP, 2015, 1502, 016
[3] S. Mukherjee, P. K. Aluri, S. Das, S. Shaikh and T. Souradeep, Direction dependence of cosmological parameters due to cosmic hemispherical asymmetry 2015
[4] S. Das, S. Mitra and S. T. Paulson, Effect of noncircularity of experimental beam on CMB parameter estimation JCAP, 2015, 1503, 048
[5] S. Mukherjee, S. Das, M. Joy and T. Souradeep, Estimation of Inflation parameters for Perturbed Power Law model using recent CMB measurements JCAP, 2015, 1501, 043
[6] W. K. Hastings, Monte Carlo Sampling Methods Using Markov Chains and their Applications, Biometrika, 1970, 57, 97-109
[7] A. Lewis, S. Bridle, Cosmological parameters from CMB and other data: a Monte-Carlo approach, Phys. Rev. D66:103511, 2002
[8] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller, Equations of State Calculations by Fast Computing Machines, Journal of Chemical Physics, 1953, 21, 1087-1092
[9] A. Lewis, Efficient sampling of fast and slow cosmological parameters, Phys. Rev., 2013, D87, 103529
[10] C. J. Geyer, Practical Markov Chain Monte Carlo, Statistical Science, 1992, 7(4): 473-483
[11] A. Gelfand and A. Smith, Sampling-Based Approaches to Calculating Marginal Densities, Journal of the American Statistical Association, 1990, 85: 398-409
[12] A. Gelman and D. B. Rubin, Inference from iterative simulation using multiple sequences, Statist. Sci. 7 (1992) 457-511
[13] J. Rosenthal, Parallel computing and Monte Carlo algorithms, Far East J. Theor. Stat., 4: 207-236, 2000
[14] A. Mira, On Metropolis-Hastings algorithms with delayed rejection, Metron, Vol. LIX, No. 3-4, pp. 231-241 (2001)
[15] L. Tierney and A. Mira Some, Adaptive Monte Carlo methods for Bayesian inference, Statist. Med. 18 2507, 1999
[16] P. J. Green and A. Mira, Delayed rejection in reversible jump Metropolis-Hastings, Biometrika, 2001 88 1035-1053
[17] http://helios.fmi.fi/~lainema/dram/
[18] H. Haario, M. Laine, A. Mira and E. Saksman, DRAM: Ecient adaptive MCMC, Statistics and Computing, 2006, 16(3): 339-354
[19] M. Trias, A. Vecchio, J. Veitch, Delayed rejection schemes for efficient Markov-chain Monte-Carlo sampling of multimodal distributions, 2009, arXiv:0904.2207
[20] W. Gilks, G. Roberts and S. Sahu, Adaptive Markov chain Monte Carlo through regeneration, Journal of the American Statistical Association, 1998, 93(443): 1045-1054
[21] H. Haario, E. Saksman and J. Tamminen, An adaptive Metropolis algorithm, Bernoulli, 2001, 7(2): 223-242
[22] G. O. Roberts and J. S. Rosenthal, Coupling and Ergodicity of Adaptive MCMC, Journal of Applied Probability, 2007, 44(2): 458-475
[23] C. Andrieu and E. Moulines, On the ergodicity properties of some adaptive MCMC algorithms, Annals of Applied Probability, 16 (2006), no. 3, 1462-1505
[24] S .Das, B. D.Wandelt and T. Souradeep, Bayesian inference on the sphere beyond statistical isotropy JCAP, 2015, 1510, 050
[25] J. Taylor, M. Ashdown, and M. Hobson, *Fast optimal CMB power spectrum estimation with Hamiltonian sampling, Mon.Not.Roy.Astron.Soc.* **389** (2008) 1284, [`arXiv:0708.2989`].

[26] A. Hajian, *Efficient Cosmological Parameter Estimation with Hamiltonian Monte Carlo*, *Phys.Rev.* **D75** (2007) 083525, [`astro-ph/0608679`].

[27] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth, *Hybrid monte carlo*, *Physics Letters B* **195** (1987), no. 2 216 – 222.

[28] S. Das, S. Mitra, A. Rotti, N. Pant, and T. Souradeep, *Statistical isotropy violation in WMAP CMB maps due to non-circular beams*, `arXiv:1401.7757`.

[29] N. Joshi, S. Das, A. Rotti, S. Mitra, and T. Souradeep, *Revealing Non-circular beam effect in WMAP-7 CMB maps with BipoSH measures of Statistical Isotropy*, `arXiv:1210.7318`.

[30] C. Armitage and B. D. Wandelt, *Deconvolution map-making for cosmic microwave background observations*, *Phys. Rev.* **D70** (2004) 123007, [`astro-ph/0410092`].

[31] E. Keihänen and M. Reinecke, *ArtDeco: a beam-deconvolution code for absolute cosmic microwave background measurements*, *Astronomy & Astrophysics* **548** (Dec., 2012) A110, [`arXiv:1208.1399`].

[32] E. Keihänen, K. Kiiveri, V. Lindholm, M. Reinecke, and A.-S. Suur-Uski, *The impact of beam deconvolution on noise properties in CMB measurements: Application to Planck LFI*, *ArXiv e-prints* (June, 2015) [`arXiv:1506.0855`].

[33] R. K. J. M. K. Jain, R. K. Iyengar, *Numerical Methods for Scientific and Engineering Computation*. New age international (P) limited, publishers, 2005.