Cosmic Microwave Background Anisotropies: Simulations with Pixel-based Likelihood

Jeong Hwa Kim*
Korea University, Department of Physics
145, Anam-ro, Seongbuk-gu, Seoul, Republic of Korea

In this paper, we study the pixel-based likelihood method for cosmic microwave background data analysis through simulations. while there are many different likelihood methods for CMB, the pixel-based likelihood is the only method that calculates the exact likelihood function directly. Therefore, it can illustrate the principles of CMB statistics much clearly, providing a nice introduction to the fundamentals of the field. We start by briefly outlining the basics of CMB and the history of its measurements. We will then discuss how to analyze these measurements in the scope of pixel-based likelihood. We then address the limitations and uncertainties arising from this method based on our simulation results.

I. INTRODUCTION

The cosmic microwave background(CMB) is one of the most important probes in cosmology. It is remarkably uniform in all directions—resembling a black body radiation spectrum at a temperature of 2.73 K. However, most of the information is hidden under the anisotropies of $O(10^{-5})$, which was discovered decades later after the first observations by Penzias and Wilson in 1964[1]. The matter perturbations caused by the quantum fluctuations in the early universe are imprinted in the CMB temperature(T) spectrum at the time of matter-radiation decoupling. Scattering between photons and electrons up to that period leaves a 'curl-free' polarization(E-mode) , whereas gravitaional lensing–and possibly early tensor perturbations–is a source to 'divergence-free' (B-mode) polarization spectrum in the CMB[2].

There is an abundance of observational efforts to measure the CMB temperature and polarization fields, including space missions[3][4], balloon experiments[5], and ground based telescopes[6][7][8]. Diverse observational method allows precise measurements over a wide range of angular scales. The data is also in good agreement with the standard model in cosmology; a homogeneous and isotropic universe. In the context of the ΛCDM framework, current data[3] suggests a spatially flat($\Omega_k h^2 = 0$) universe composed of baryons($\Omega_b h^2 = 0.022$), dark matter($\Omega_c h^2 = 0.12$), and dark energy($\Omega_\Lambda h^2 = 0.68$) up to sub-percent accuracy. Moreover, the level of precision cosmology achieved in the recent years can provide valuable data for particle physics, e.g. constraining neutrino properties from the thermal relics of the early universe[9].

The likelihood function to constrain the model parameters from observed CMB data is far from unique. Making a suitable choice for a certain type of experiment can often be difficult, and there have been cases that a sensitive model providing accurate results in some experiments turned out to be misleading for others. With that said, the pixel-likelihood method that will be discussed in this paper captures the core properties of CMB data analysis in a relatively simple manner.

Statistics of CMB

The CMB temperature field $T(\hat{\mathbf{n}}) = T(1 + \Theta(\hat{\mathbf{n}}))$ observed in a given direction $\hat{\mathbf{n}}$ has a fluctuation $\Theta(\hat{\mathbf{n}})$. The direct values of $\Theta(\hat{\mathbf{n}})$ cannot be predicted by our model parameters, because our own universe is only one particular realization from a set of universe sharing the same parameter set. So we can only infer statistical properties of the observed perturbation field. We start by decomposing $\Theta(\hat{\mathbf{n}})$ in terms of spherical harmonics [2][10]:

$$\Theta(\hat{\mathbf{n}}) = \sum_{l,m} a_{lm} Y_{lm}(\hat{\mathbf{n}}) \tag{1}$$

where the harmonic Y_{lm} corresponds to an angular scale $\theta \sim \pi/l$, with $(2l+1)$ m-modes for each multipole l. Assuming the a_{lm}'s follow a Gaussian distribution, the power spectrum

$$\langle a_{lm} a^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l \tag{2}$$

fully characterizes a given field. The subscript m from C_l can be dropped using rotational invariance of CMB.

Now the two-point correlation function $C(\theta) = \langle \Theta(\hat{\mathbf{n}}_1)\Theta(\hat{\mathbf{n}}_2) \rangle$ can be written in terms of C_l

$$C(\theta) = \sum_l \frac{2l+1}{4\pi} C_l P_l(\cos\theta) \tag{3}$$

where $\hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 = \cos\theta$.

A similar analysis can be done for polarization fields E and B by introducing spin-2 spherical harmonics ${}_{\pm 2}Y_{lm}$ along with the Stokes Parameters Q and U:

$$(Q + iU)(\hat{\mathbf{n}}) = \sum_{l,m} {}_2a_{lm}\, {}_2Y_{lm}(\hat{\mathbf{n}}) \tag{4}$$

$$(Q - iU)(\hat{\mathbf{n}}) = \sum_{l,m} {}_{-2}a_{lm}\, {}_{-2}Y_{lm}(\hat{\mathbf{n}}) \tag{5}$$

* jailbraker@korea.ac.kr

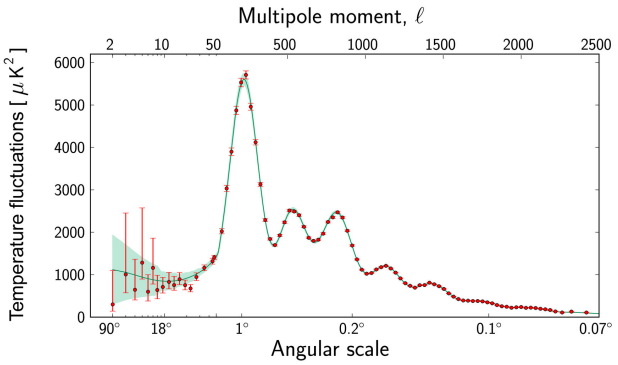

FIG. 1. The TT power spectrum constructed from the Planck measurements along with best fit results[3].

where the coefficients for E and B are related by

$$a^E_{lm} = -\frac{1}{2}({}_2a_{lm} + {}_{-2}a_{lm}) \tag{6}$$

$$a^B_{lm} = \frac{i}{2}({}_2a_{lm} - {}_{-2}a_{lm}). \tag{7}$$

The actual calculation can be performed in terms of either E and B or Q and U since they are physically equivalent. We use Q and U for our simulation.

The correlation functions derieved from the above relation is then later used to construct the covariance matrix. See reference [10] for more details.

Cosmic Variance As mentioned earlier, our analysis is fundamentally limited by the fact that there is only one observable universe. We cannot detect a monopole-shift of our vicinity, nor can we distinguish a cosmological dipole from our peculiar motion with respect to the CMB rest frame[2]. The exact relation between C_l and a_{lm}

$$C_l = \frac{1}{2l+1}\sum_m \langle |a_{lm}| \rangle \tag{8}$$

requires an *ensemble* average operation, meaning that we need to average over all possible realizations. It is apparent that this is not possible. So we replace $\langle |a_{lm}| \rangle$ with $|a_{lm}|$ to obtain the estimator for C_l:

$$\hat{C}_l = \frac{1}{2l+1}\sum_m |a_{lm}| \tag{9}$$

$\hat{C}_l$ has a mean of C_l and a variance[11]

$$\left\langle \left(\frac{\hat{C}_l - C_l}{C_l} \right)^2 \right\rangle = \frac{2}{2l+1}. \tag{10}$$

This is known as the cosmic variance and reflects our limited$(2l+1)$ sample size for each l. The effect is apparant in Fig. 1, where the disagreement between the data and the fit curve is significant in the low-l region.

II. PIXEL-BASED LIKELIHOOD

The basic object of our interest is the likelihood function $\mathcal{L}$, i.e., the probability of the observed data d given a model, regarded as a function of the model itself. If the model is defined in terms of a vector of parameters Θ_p, we have:

$$\mathcal{L}(\Theta_p) = p(d|\Theta_p). \tag{11}$$

We will simulate our data using the six independent ΛCDM parameters[12]

Hubble parameter	$H_0 = 67.5$ km s^{-1} Mpc^{-1}
Baryon density	$\Omega_b h^2 = 0.022$
Dark matter density	$\Omega_c h^2 = 0.122$
Neutrino mass	$m_\nu = 0.06$ eV
Curvature	$\Omega_k = 0$
Optical depth	$\tau = 0.06$

and three parameters to generate the primordial power spectrum.

Comoving curvature power	$A_s = 2 \times 10^{-9}$
Scalar spectral index	$n_s = 0.96$
Tensor to scalar ratio at pivot	$r = 0$

In the following analysis, we will fit the optical depth to reionization(τ). It is a unitless quantity which provides a measure of the line-of-sight free-electron opacity to CMB radiation. Assuming a fixed dependence of electron density on redshift, a larger value of τ implies eariler periods of reionization and thus an earlier onset of star and galaxy formation[13].

The likelihood in real space(in terms of observed values, rather than a_{lm}'s) is defined as[11]

$$\mathcal{L} = p(\mathbf{m}|C_l) = \frac{1}{2\pi|M|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{m}^T M^{-1}\mathbf{m}\right) \tag{12}$$

where $\mathbf{m}$ is the observed map with $3N$ elements containing the data of $T(\Theta)$, Q and U. The covariance matrix M is calculated from the model parameters by the correlation function introduced earlier.

$$M(\hat{\mathbf{n}}_i \cdot \hat{\mathbf{n}}_j) = \begin{pmatrix} \langle T_iT_j\rangle & \langle T_iQ_j\rangle & \langle T_iU_j\rangle \\ \langle T_iQ_j\rangle & \langle Q_iQ_j\rangle & \langle Q_iU_j\rangle \\ \langle T_iU_j\rangle & \langle Q_iU_j\rangle & \langle U_iU_j\rangle \end{pmatrix} \tag{13}$$

The above likelihood is exact, but highly expensive from a computational standpoint. The computational cost to evaluate M as well as its inverse scales roughly with l_{max}, limiting the use of pixel likelihood method to only large angular scales. Small scale analysis should resort to approximation methods[14][15] such as pseudo-C_l formalism for efficient calculation.

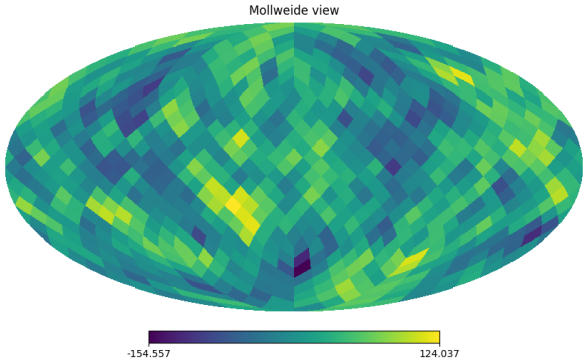

FIG. 2. An example of a TT map simulated with $N_{side} = 8$.

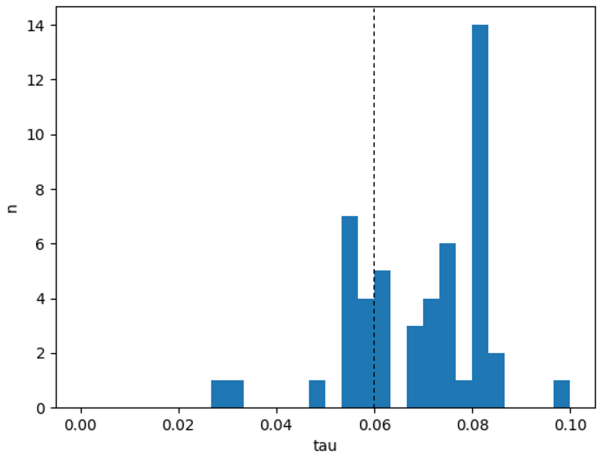

FIG. 3. The result of fitting 50 maps with $N_{side} = 8$ generated from optical depth $\tau = 0.06$, identity matrix regularization.

A. Simulation Methods

We used `CAMB`[12] to generate C_l power spectrums from the cosmological parameters and `healpy`[16][17] to generate maps from C_l. `CAMB` is a library for calculating CMB, lensing, galaxy count, and other CMB-anisotropy related calculations. `healpy` is a `python` implementation of the `HEALPix` scheme, designed to efficiently process CMB data in spherical sky coordinates.

The `HEALPix` algorithm starts by dividing the sphere into twelve identical square-shaped facets. Each square is further divided into N_{side}^2 segments where N_{side} is a user defined parameter. A `HEALPix` map $\mathbf{m}$ is a 1-dimentional object containing the values of each pixel. Fig. 2 shows an example map generated from `healpy`.

The main steps of the simulations are as follows:

- Generate a power spectrum $C_{l,0}$ with a set of cosmological parameters $\Theta_{p,0}$.
- Generate a map $\mathbf{m}_0$ from the above power spectrum.
- Generate n sets of $C_{l,n}$ from parameters $\Theta_{p,n}$ for sufficiently large n. Can be optimized by using a minimization algorithm rather than brute-force calculations.
- Minimize the negative log-likelihood by computing $-\log p(\mathbf{m}_0|C_{l,n})$ for each n.

It is important to note that the likelihood in (12) is not well-defined. Since the mean values of T, Q, and U maps are all zero by definition, the degree of freedom of each maps are reduced to $N - 1$. This implies that the covariance matrix M is singular and therefore does not have a proper inverse matrix. So in real analysis, the covariance matrix is always corrected with a noise matrix N involving the systematic uncertainties of the setup. We used an identity matrix of $O(10^{-5})$ to regularize our $\langle TT \rangle$ matrix and $O(10^{-10})$ for the other matrices.

III. RESULTS

To test the performance of our method, we fitted the optical depth τ for 50 maps generated from $\tau = 0.06$. The resolution is $N_{side} = 8$, corresponding to an angular scale of roughly $\sim 7°$ and an l_{max} of 23. The N_{side} parameter of our map cannot exactly determine the value of l_{max} since different m components of the same l spherical harmonics can represent a variety of angular scales.

The fit result is shown in Fig. 3. While we cannot make accurate statistical statements due to our limited sample size, there seems to be a significant bias towards $\tau \simeq 0.08$. The reason behind this bias is unclear, although it was reproduced in most of our parameter fits. Recovered values of curvature(Ω_k) from maps generated with $\Omega_k = 0$ were also biased towards $\Omega_k \simeq 0.1$. Changing the regularization matrix to a random matrix of a similar scale did not remove this bias(Fig. 4).

The resolution of 50 simulations should correspond to an error of approximately $\sigma \sim 1/\sqrt{50} \sim 14\%$. Considering the effects of the cosmic variance, rest of the fit results (excluding the peak around $\tau = 0.08$) seems to be in good agreement with the initial value of $\tau = 0.06$.

IV. CONCLUSION

CMB data analysis has provided the most accurate measurements of our universe for decades. Many ongning experiments (BICEP/Keck, SPIDER, GroundBird, etc.) aim to uncover more information from CMB polarization fields, in hopes of better understanding the early phase of our universe and its subsequent evolution.

There are many likelihood methods designed for CMB analysis. Pixel-based likelihood, the one we've seen in this paper, is an exact likelihood function that does not involve any approximations. However, it is only appropriate for certain low-l analysis because of its large computa-

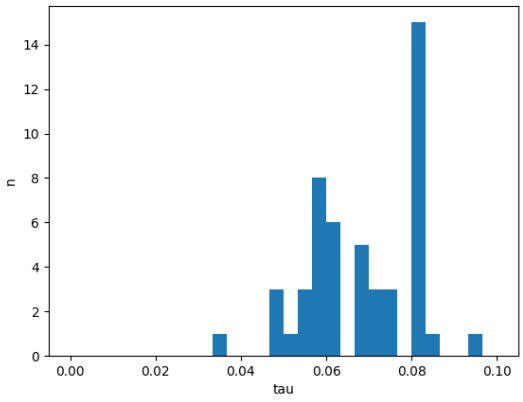

FIG. 4. The result of fitting 50 maps with $N_{side} = 8$ generated from optical depth $\tau = 0.06$, random matrix regularization.

tional costs. Other methods—such as the Blackewll-Rao estimator or the Hamimeche-Lewis method—also have their weaknesses which makes deciding a suitable likelihood method for a certain experiment very difficult.

Due to the limited timeframe available fo write this paper, we could not address all of the subtleties that took place during our study. Most notably, we could not fully understand the cause of the bias mentioned in section III.

A. Research limitations

- Regularizing the covariance matrix through either an identity matrix or a random matrix seems to cause a bias in the fit results. We need a physically sound regularization scheme for the covariance matrix.

- The likelihood depends on our choice of l_{max} near $3N_{side} - 1$. In particular, the likelihood diverges to unphysical value when the maximum value of l is reduced by 1, to $3N_{side} - 2$. The relation between l_{max} and $\mathcal{L}$ should be studied.

ACKNOWLEDGMENTS

This paper was written as part of the Advanced Physics experiment class in Korea University. I thank PhD. Kyungmin Lee and Prof. Eunil Won for providing a much needed guidance to CMB physics.

[1] A. A. Penzias and R. W. Wilson, A measurement of excess antenna temperature at 4080 mc/s., The Astrophysical Journal **142**, 419 (1965).
[2] W. Hu and S. Dodelson, Cosmic microwave background anisotropies, Annual Review of Astronomy and Astrophysics **40**, 171 (2002).
[3] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, N. Bartolo, S. Basak, *et al.*, Planck 2018 results. vi. cosmological parameters, arXiv preprint arXiv:1807.06209 (2018).
[4] G. Hinshaw, D. Larson, E. Komatsu, D. Spergel, C. Bennett, J. Dunkley, M. Nolta, M. Halpern, R. Hill, N. Odegard, *et al.*, Apjs 208, 19 (2013), arXiv preprint arXiv:1212.5226 **3**.
[5] A. A. Fraisse, P. A. Ade, M. Amiri, S. Benton, J. Bock, J. Bond, J. Bonetti, S. Bryan, B. Burger, H. Chiang, *et al.*, Spider: probing the early universe with a suborbital polarimeter, Journal of Cosmology and Astroparticle Physics **2013** (04), 047.
[6] T. Louis, E. Grace, M. Hasselfield, M. Lungu, L. Maurin, G. E. Addison, P. A. Ade, S. Aiola, R. Allison, M. Amiri, *et al.*, The atacama cosmology telescope: two-season actpol spectra and parameters, Journal of Cosmology and Astroparticle Physics **2017** (06), 031.
[7] P. A. Ade, R. Aikin, J. Bock, J. Brevik, J. Filippini, H. Hui, S. Kefeli, M. Lueker, R. O'Brient, A. Orlando, *et al.*, Improved constraints on cosmology and foregrounds from bicep2 and keck array cosmic microwave background data with inclusion of 95 ghz band, Physical Review Letters **116**, Art (2016).
[8] S. Oguri, J. Choi, T. Damayanthi, M. Hattori, M. Hazumi, H. Ishitsuka, K. Karatsu, S. Mima, M. Minowa, T. Nagasaki, *et al.*, Groundbird: observing cosmic microwave polarization at large angular scale with kinetic inductance detectors and high-speed rotating telescope, Journal of Low Temperature Physics **184**, 786 (2016).
[9] J. Lesgourgues and S. Pastor, Neutrino cosmology and planck, New Journal of Physics **16**, 065002 (2014).
[10] M. Zaldarriaga, Cosmic microwave background polarization experiments, The Astrophysical Journal **503**, 1 (1998).
[11] M. Gerbino, M. Lattanzi, M. Migliaccio, L. Pagano, L. Salvati, L. Colombo, A. Gruppuso, P. Natoli, and G. Polenta, Likelihood methods for cmb experiments, Frontiers in Physics **8**, 15 (2020).
[12] A. Lewis, A. Challinor, and A. Lasenby, Efficient computation of CMB anisotropies in closed FRW models, Astrophys. J. **538**, 473 (2000), arXiv:astro-ph/9911177 [astro-ph].
[13] L. M. Griffiths, D. Barbosa, and A. R. Liddle, Cosmic microwave background constraints on the epoch of reionization, Monthly Notices of the Royal Astronomical Society **308**, 854 (1999).
[14] M. Tegmark and A. de Oliveira-Costa, How to measure cmb polarization power spectra without losing information, Physical Review D **64**, 063001 (2001).
[15] R. E. Upham, L. Whittaker, and M. L. Brown, Exact joint likelihood of pseudo-c_l estimates from correlated gaussian cosmological fields, Monthly Notices of the Royal Astronomical Society **491**, 3165 (2020).
[16] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset,

E. Hivon, and K. Gorski, healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python, Journal of Open Source Software **4**, 1298 (2019).

[17] K. Gorski, E. Hivon, A. Banday, B. Wandelt, F. Hansen, M. Reinecke, and M. Bartelman, Healpix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere, apj **622** (2004).

[18] M. Zaldarriaga and U. Seljak, All-sky analysis of polarization in the microwave background, Physical Review D **55**, 1830 (1997).