A Foreground Model Independent Estimation of Joint Posterior of CMB E mode Polarization over Large Angular Scales

Ujjal Purkayastha¹, Vipin Sudevan¹, Rajib Saha¹

¹Physics Department, Indian Institute of Science Education and Research Bhopal, Bhopal, M.P, 462066, India.

Ever since Cosmic Microwave Background (CMB) signal is being measured by various satellites based observations with increasing experimental accuracies there has been a parallel increase in the demand for a CMB reconstruction technique which can provide accurate estimates of CMB signal and the theoretical angular power spectrum along with reliable statistical error estimates associated with them. In this work, we estimate the joint posterior of CMB E mode signal (S) and corresponding theoretical angular power spectrum (C^E_ℓ) over large angular scales given the simulated polarization observations of future generation CorE satellite mission. To generate samples from the joint distribution we employ the internal-linear-combination (ILC) technique with prior information of CMB E mode covariance matrix augmented by a Gibbs sampling technique of Sudevan and Saha 2020. We estimate the marginalized densities of S and C^E_ℓ using the samples from full-posterior. The best fit cleaned E mode map and the corresponding angular power spectrum agree well with the input E mode map and the sky power spectrum implying accurate reconstruction using CorE like observations. Using the samples C^E_ℓ of all Gibbs chains we estimate the likelihood function P(C^E_ℓ|D) of any arbitrary C^E_ℓ given simulated observed maps (data, D) of CorE mission following Blackwell-Rao estimator. The likelihood function can be seamlessly integrated to the cosmological parameter estimation method. Apart from producing an accurate estimate of E mode signal over large angular scales our method also builds a connection between the component reconstruction and reliable cosmological parameter estimation using CMB E mode observations over large angular scales. The entire method does not assume any explicit models for E mode foreground components in order to remove them, which is an attractive property since foreground modelling uncertainty does not pose as a challenge in this case.

I. INTRODUCTION

The weak E mode polarization signal of Cosmic Microwave Background (CMB) radiation [28] anisotropies were generated in early Universe by Thomson scattering at the surface of last scattering and by the scattering of the CMB photons by the ionized Hydrogen and Helium atoms in the intergalactic medium during reionization epoch [7] in relatively recent past. CMB E mode signal provides valuable insights into the era of reionization by breaking degeneracies between scalar field fluctuation amplitude and optical depth of the reionization era. The reionization bump over large angular scales of CMB E mode signal encodes unique information about the reionization optical depth [47]. The shape and location of this peak is sensitive to the epoch of reionization [20, 21]. For accurate estimation of cosmological parameters from the CMB E mode signal over the large angular scales a foreground removal technique that can provide accurate estimates of the signal along with the statistical error estimates is necessary to be employed.

An important derivable in context of CMB analysis is the joint posterior P(S, C^E_ℓ|D) of the CMB E mode signal S and theoretical angular power spectrum C^E_ℓ given the observed (E mode) maps, D since following Bayesian concept all the information that can be derived about S and C^E_ℓ from D is encoded in this function. The posterior density can be used to best-fit values of the random variables S and C^E_ℓ along with the associated error-intervals. The posterior can be marginalized over the E mode signal to estimate the likelihood function P(C^E_ℓ|D) which plays a central role in estimating the cosmological parameters. The posterior and hence the likelihood function derived from the cleaned CMB E mode signal are however expected to be different for different CMB component reconstruction method since each is characterized effectively by a different foreground removal filter. Hence these functions need to be accurately determined for a CMB reconstruction technique for accurate estimation and correct interpretation of cosmological parameters. In this article, we estimate both the joint posterior density and likelihood function of the cleaned E mode maps over large angular scales using internal linear combination (ILC) method along with the Gibbs sampling technique [45]. Since the likelihood function estimated by us is calculable for any choice of C^E_ℓ it can be readily used to calculate cosmological parameters integrating with a Markov Chain Monte Carlo (MCMC) method [55].

A major step towards estimating the joint conditional distribution and the likelihood function discussed above is process of reconstruction of foreground minimized CMB E maps. Two major polarization foreground components are synchrotron at low frequency and thermal dust at high frequency. These contaminations are very strong at large angular scales and can almost cover the reionization bump [1]. Reconstruction of weak E mode background CMB signal by removing the strong foregrounds therefore becomes a challenging task. The
method may be further limited by the presence of substantial amount of detector noise in the observed maps. Fortunately, many future generation sensitive CMB experiments are being designed to accurately measure the weak CMB polarization anisotropy with sufficiently large signal to noise ratio.

Two distinct routes have been followed in the literature in order to accurately reconstruct the CMB component by removing the foregrounds. The first type of them minimizes contributions from all non-CMB components thereby leaving only a CMB signal, without, however, using any explicit foreground models. The other type of the methods requires prior knowledge about the frequency dependence and or morphological pattern of the different foreground components present in the observed CMB maps. Such methods are known as foreground model-dependent methods like Wiener Filtering [5, 6], Gibbs sampling approach [48–51], template fitting method [53], the maximum entropy method [56], a MCMC [13] method. This method works based on principle of inclusiveness in which one exploits freedom of large scale modelling of each and every physical component present in the sky. The first type of method, on the contrary, requires very little assumption about the foreground components which are to be removed and as far as CMB reconstruction alone is concerned as is the case for cosmological analysis, becomes a much simpler problem to focus. Various model-independent methods have been developed such as Independent Component Analysis (ICA) [4, 25, 26], ILC [43], Correlated Component Analysis (CCA) [57].

ILC is a foreground minimization method in which in order to obtain a cleaned CMB signal one makes a simple assumption that the foreground and noise spectra are different from the CMB power spectrum, which follows a black-body spectrum [54]. The cleaned CMB map obtained using the ILC method is not susceptible to inaccuracies in foreground modeling. In the ILC method, a cleaned CMB map is obtained by linearly combining multi-frequency observed foreground contaminated CMB maps using some amplitude terms known as weight factors. These weights are subjected to the constrain that the sum of all weights should be unity. In the ILC method these weights are estimated analytically by performing a constrained minimization of the variance of the cleaned CMB map. The analytical nature of the weights is an added advantage of the method since it avoids issues related to numerical convergences that may be present in numerical approximate methods.

Since we use ILC method for E mode CMB reconstruction in the current article the posterior density and the likelihood function estimated by us become independent on the explicit modelling of polarized foregrounds. Modelling uncertainty for the polarized foregrounds does not become any issue for these estimates. Moreover, as we demonstrate in this work and since we use a large number of frequency maps in our analysis the effects of residual foregrounds in the cleaned E modes map and angular power spectrum are negligible. The posterior density and likelihood function derived in this work therefore can safely be assumed to be free from effects of such foreground residuals. For estimation of the joint conditional posterior using explicit models of foreground components we refer to [11].

The ILC method has been studied extensively in [2, 8, 11, 23, 24, 33–35, 42, 43]. A global ILC method was proposed [37] where the weights are estimated by minimizing a CMB covariance weighted variance instead of the usual variance in the cleaned maps. The authors [58] extended this method to obtain a cleaned CMB E mode map at large angular scales. Recently [62] a detailed analysis of the impact of random residual calibration errors present in the observed CMB temperature maps have been studied for the case of [41] and was shown that the effects of these random calibration errors are small.

We have organized our paper as follows. In Section II we review the basic formalism of our method. In Section III we discuss the input maps that we have used in the present analysis and the methodology in Section IV. We present the results of our analysis in Section V. In Section VI we discuss the Blackwell-Rao estimator for estimating continuous likelihood distribution of the CMB theoretical E mode angular power spectrum. Finally we check for the convergence of the Gibbs chains in Section VII and present the discussion and conclusion of our work in Section VIII.

II. BASIC FORMALISM

In this section we briefly review the basic formalism of our method of estimating joint conditional density of CMB E mode signal. The method consists of two interleaved steps of estimating the cleaned E mode CMB map by employing a modified version of ILC algorithm [37, 58] and sampling the E mode theoretical angular power spectrum from the respective conditional density [41]. Following Gibbs theorem after burn in rejection the collection of samples of cleaned E mode map and angular power spectrum forms samples from the desired joint posterior density.

The observables for the linear polarization of CMB are described by the Stokes Q and U parameters. Linear combinations of these variables under a local coordinate transformation transform as spin $\pm 2$ quantities [21, 46]. These spin functions can be expanded in terms of the spin $\pm 2$ basis functions $\pm 2Y_{\ell m}(p)$ following,

$$Q(p) \pm iU(p) = \sum_{\ell m} \pm 2a_{\ell m} Y_{\ell m}(p),$$

(1)

where $p$ denotes the pixel index and $\pm 2a_{\ell m}$ represents the $\pm 2$ spherical harmonic coefficients. One can form linear combinations of the spin harmonic coefficients to obtain

$$a_{\ell m}^E = \frac{1}{2} (\pm 2a_{\ell m} + 2a_{\ell m}),$$

(2)
using which spin-0 CMB E mode map can be defined as follows,

\[ E(p) = \sum_{\ell m} d_{\ell m}^E Y_{\ell m}(p), \tag{3} \]

where \( Y_{\ell m}(p) \) represents the spin-0 spherical harmonics.

In this article we work directly in the E mode basis instead of Q, U basis from which E mode angular power spectrum \( C_E^E \) can be obtained by a simple spin-0 spherical harmonic transformation. Using E mode maps in the analysis helps since it avoids additional spin 2 spherical harmonic transformations necessary to obtain the E mode angular power spectrum at each Gibbs iteration while reducing the total disk storage requirement by half. The E mode maps can of course be converted to equivalent Q, U maps in a lossless fashion if desired. Since we use full-sky Q, U maps for converting to E mode maps we do not arise.

Let us assume that we have observations of CMB E mode signal \( S \) at \( n \) different frequencies. The observed map \( d_i \) at frequency \( \nu_i \) is given by,

\[ d_i = S + F_i + n_i, \tag{4} \]

where \( F_i \) denotes net E mode polarization from all the foreground components at the \( i^{th} \) frequency and \( n_i \) represents the detector noise contamination. We assume each of the input frequency maps are already smoothed by a common beam and pixel window functions so that we can safely remove any explicit reference to these smoothing effects in Eqn. 4. Each of the bold-faced variables in Eqn. 4 represents a column vector of size \( N_{\text{pix}} \), total number of pixels in the observed maps. The observed data \( D = \{ d_1, d_2, ..., d_n \} \) is represented by a matrix of size \( N_{\text{pix}} \times n \).

To map out the posterior density \( P(S, C_E^E|D) \) one approach is to draw large number samples from it without direct evaluation of the function. An alternative method to achieve this is to employ Gibbs sampling in which samples are drawn from the two conditional densities which are easier to sample. In particular the \( (i+1)^{th} \) Gibbs CMB E mode signal is obtained by drawing sample from

\[ S_i^{i+1} \sim P_1(S|D, C_{E,i}^E), \tag{5} \]

where \( C_{E,i}^E \) represents the corresponding theoretical angular power spectrum sampled at the previous step. \( C_{E,i}^{E,i+1} \) is obtained by sampling the conditional density

\[ C_{E,i}^{E,i+1} \sim P_2(C_{E,i}^E|D, S_i^{i+1}). \tag{6} \]

Using the pair of samples at the \( (i+1)^{th} \) iteration we repeat the two sampling steps (Eqns. 5 and 6), for a large number of iterations until convergence is achieved. Ignoring the initial burn-in phase the samples drawn from the two conditional densities are then samples from the joint posterior density.

How to draw a samples of CMB E mode signal given \( D \) and \( C_E^E \) ? This is achieved by estimating the cleaned E mode CMB signal given the data and a theoretical CMB E mode angular power spectrum following the foreground removal method described by [37, 58]. The cleaned CMB E mode map \( \hat{S} \) which is an estimator of the true CMB E mode signal \( S \) is obtained following the ILC method,

\[ \hat{S} = \sum_{i=1}^{n} w_i d_i, \tag{7} \]

where, \( w_i \) represents the weight corresponding to the \( i^{th} \) frequency and \( \nu_i \) is independent of frequency due to the black-body nature of its frequency spectrum we perform a constrained minimization of the CMB E mode covariance weighted variance, \( \sigma^2 \) defined as

\[ \sigma^2 = S^T C_i^{E} S, \tag{8} \]

to estimate the weights. \( C \) represents the CMB E mode theoretical covariance matrix and \( e \) denotes the Moore-Penrose generalized inverse [52]. Using a Lagrange’s multiplier approach, we obtain the choice of weights which minimizes \( \sigma^2 \) as

\[ W = e^T A^T e, \tag{9} \]

where \( W \) is a \( (1 \times n) \) weight vector with elements \( W = \{ w_1, w_2, ..., w_n \} \) and the \( (i, j)^{th} \) element of the \( A \) matrix is in pixel space is given by,

\[ A_{ij} = d_i^T C_i^{E} d_j. \tag{10} \]

Since \( C \) is a dense \( N_{\text{pix}} \times N_{\text{pix}} \) matrix computing Eqn. 10 at every Gibbs iteration becomes computationally very expensive. It is greatly advantageous to work in the harmonic space where Eqn 10 is simpler to compute,

\[ A_{ij} = \sum_{\ell=2}^{\ell_{\text{max}}} (2\ell + 1) \sigma_{ij}^{\ell} C_{E,\ell}^E, \tag{11} \]

where \( \ell_{\text{max}} \) represents the maximum multipole used in the analysis, \( \sigma_{ij}^{\ell} \) denotes the cross angular power spectrum between \( d_i \) and \( d_j \). \( B_{\ell} \) and \( P_{\ell} \) are respectively the beam and pixel smoothed CMB EE theoretical power spectrum i.e.,

\[ C_{E,\ell}^E = C_{E,\ell}^E B_{\ell}^2 P_{\ell}^2, \tag{12} \]

where \( C_{E,\ell}^E \) does not contain any smoothing effects and \( B_{\ell} \) and \( P_{\ell} \) are respectively the beam and pixel window functions.

To draw samples of \( C_{E,\ell}^E \) given \( S \) and \( D \) we first obtain the conditional density on \( P_2(C_{E,\ell}^E|S, D) \) in terms of the variable \( y = \hat{C}_{E,\ell}^E (2\ell + 1)/C_{E,\ell}^E \) as [41]

\[ P_2(C_{E,\ell}^E|\hat{C}_{E,\ell}^E) \propto y^{-(2\ell-1)/2-1} \exp \left[ -\frac{y}{2} \right], \tag{13} \]
where, $\hat{C}^E_\ell$ is estimated from the cleaned CMB E mode map and a uniform prior on $C^E_\ell$ is assumed. Eqn. 13 shows that the variable $y$ follows a $\chi^2$ distribution with $2\ell - 1$ degrees of freedom. Therefore, in order to sample a CMB theoretical power spectrum, we draw $y$ from the $\chi^2$ distribution of $2\ell - 1$ degrees of freedom and then find $C^E_\ell$ as follows

$$C^E_\ell = \hat{C}^E_\ell (2\ell + 1)/y.$$  \hspace{1cm} (14)

III. INPUT MAPS

In our analysis, we simulate foreground and noise contaminated CMB E mode maps at the CORe frequency channels ranging from 60 GHz to 340 GHz. The higher frequencies are avoided as they have comparatively large level of detector noise contamination. The frequency maps used in this work along with the beam widths and noise levels are listed in Table I.

A. CMB Maps

We generate random realizations of the CMB Stokes Q and U maps using the anafast compatible to Planck theoretical CMB angular power spectrum [59] at HEALPix (Hierarchical Equal Area IsoLatitude Pixellation of sky) [15] resolution $N_{\text{side}} = 16$. We then smooth the resulting map using a Gaussian beam window of FWHM $9^\circ$ and convert it to the CMB E mode map using the synfast facility.

B. Foreground model

Synchrotron and thermal dust are two major foreground components that contaminate CMB polarization signal. We first generate Stokes Q and U maps for synchrotron and thermal dust components at the 15 CORe frequencies between 60 to 340 GHz following a procedure similar to [38]. To generate the synchrotron Stokes maps we use WMAP 23 GHz (K band) Stokes maps smoothed by a effective Gaussian beam window function of FWHM $1^\circ$ at $N_{\text{side}} = 512$. In thermodynamic $\mu K$ temperature unit the synchrotron Stokes maps $Q_s(\nu, p), U_s(\nu, p)$ at a frequency $\nu$ (in GHz) are given by,

$$Q_s(\nu, p) = 1000 \left( \frac{a(\nu)}{a(\nu_s)} \right) Q_s(p) \left( \frac{\nu}{\nu_s} \right)^{\beta_s},$$  \hspace{1cm} (15)

$$U_s(\nu, p) = 1000 \left( \frac{a(\nu)}{a(\nu_s)} \right) U_s(p) \left( \frac{\nu}{\nu_s} \right)^{\beta_s},$$

where $p$ denotes the pixel index, $Q_s(p)$ or $U_s(p)$ represents the synchrotron Q or U map at the reference frequency $\nu_s = 23$ GHz, as observed by WMAP in mk thermodynamic temperature unit, $a(\nu)$ represents the antenna to thermodynamic conversion factor and $\beta_s = -3.0$ which represents the pixel independent synchrotron spectral index. As mentioned in [38] the choice of such a spectral index corresponds to the typical mean synchrotron spectral index measured at the CMB observation window.

We generate thermal dust polarization maps at the five CORe frequencies by extrapolating Planck 353 GHz thermal dust optical depth map as produced by generalized needlet ILC (GNILC) algorithm [39]. Using the GNILC map is advantageous since it provides a significantly improved picture of galactic thermal dust emission by separating the cosmic infrared background (CIB) anisotropies. We downgrade the 353 GHz GNILC optical depth map at $N_{\text{side}} = 512$ and smooth it by a Gaussian beam function of FWHM = $\sqrt{(60')^2 - (5')^2}$ to bring the resulting map to the effective beam resolution of $1^\circ$. We represent this map by $\tau_{353}(p)$. The thermal dust intensity map $I(\nu, p)$ is then obtained following,

$$I(\nu, p) = 10^{20} B(\nu, T_d) \tau_{353}(p) \left( \frac{\nu}{\nu_d} \right)^{\beta_d},$$  \hspace{1cm} (16)

where, $\beta_d = 1.6$ and $T_d = 19.4 K$ (e.g., [38, 39]), $\nu_d = 353$ GHz and $B(\nu, T_d)$ represents the Planck function at a
thermal dust temperature $T_d$,
\[
B(\nu, T_d) = \frac{2h^3}{c^2} \exp\left(\frac{h\nu}{kT_d}\right) - 1.
\] (17)

Since $B(\nu, T_d)$ physically represents intensity, in metric system its unit is Joule sr$^{-1}$m$^{-2}$s$^{-1}$Hz$^{-1}$. Therefore, the numerical factor of $10^{20}$ in Eqn. 16 estimates dust intensity in MJy sr$^{-1}$ unit. We estimate thermal dust Stokes Q, U polarization maps using the intensity map of Eqn. 16 following [38, 40],
\[
Q(\nu, p) = f_d g_d(\nu) \cos(2\gamma_d(\nu))I(\nu, p),
\]
\[
U(\nu, p) = f_d g_d(\nu) \sin(2\gamma_d(\nu))I(\nu, p),
\]
(18)

where $f_d, g_d(\nu)$ and $\gamma_d(\nu)$ represent respectively pixel independent intrinsic dust polarization fraction and pixel dependent dust geometric depolarization and polarization angle. As in [38, 40] we take $f_d = 0.15$. Following these literatures the polarization angles are estimated using WMAP K band Stokes Q and U maps (after smoothing to effective resolution of $3^\circ$) using,
\[
\gamma_d(\nu) = \frac{1}{2} \tan^{-1} \left( -U_{23}(\nu), Q_{23}(\nu) \right).\]
(19)

The depolarization factor is computed from the properly scaled version of the degree of linear polarization at 23 GHz as obtained by extrapolating the 408 MHz Haslam synchrotron template ($I_{0.408}(p)$) to 23 GHz assuming a constant spectral index $-3.0$. Prior to using the Haslam template we subtract a residual monopole offset of 8.33K in antenna temperature unit and smooth the corrected map to an effective resolution of $3^\circ$ at $N_{side} = 512$. The depolarization factor is obtained following,
\[
g_d(\nu) = \frac{\sqrt{Q_{23}(\nu)^2 + U_{23}(\nu)^2}}{1000fI_{0.408}(p)(23.0/0.408)^{-3}},
\]
(20)

where the factor $f$ in the denominator represents the product of synchrotron polarization fraction (0.75) and antenna to thermodynamic conversion factor at 23GHz. Once we have obtained the polarization angles and depolarization factors respectively following Eqs. 19 and 20 we use them in Eqn. 18 to estimate thermal dust Stokes parameters at CoRE frequencies.

Using the full sky Stokes Q and U maps for synchrotron and thermal dust components as discussed above we obtain full sky E mode maps for these components at all the frequencies of this work at $N_{side} = 16$ as follows. First we
generate $a^E_{\ell m}$ at each frequency up to the maximum multipole $\ell_{\text{max}} = 32$ following Eqn 2 and smooth them by the COrE science team and simulate foreground and noise contaminated CMB E mode input maps at all 15 COrE frequencies. The noise levels for Stokes Q and U for the 15 COrE frequencies in the unit of $\mu$K.arcmin are given in the third column of Table I (e.g., see [60]). Noise levels for Stokes Q and U map at frequency $\nu$ are given in the third column of Table I. We assume that the noise is Gaussian, isotropic and is uncorrelated from pixel to pixel. We also make a further assumption that the Q and U noise maps are pixel uncorrelated. Mathematically this means that Stokes Q and U noise maps at pixel index $p$ can be represented as

$$\langle Q_i(p)Q_i(p') \rangle = \sigma^2_{Q_i}\delta_{pp'}, \quad (21)$$

$$\langle U_i(p)U_i(p') \rangle = \sigma^2_{U_i}\delta_{pp'}, \quad (22)$$

$$\langle Q_i(p)U_i(p') \rangle = 0. \quad (23)$$

where $\sigma^2_{Q_i}$ and $\sigma^2_{U_i}$ are the pixel noise variances for Q and U maps at frequency $\nu_i$ respectively and are assumed to be identical. The noise variances are given by

$$\sigma^2_{Q_i} = \sigma^2_{U_i} = \left(\Delta Q^2_{\text{e}} a_r \right)^2 / (\Delta \Omega), \quad (24)$$

where $a_r$ represents the conversion factor from arcmin to radian and $\Delta \Omega = 4\pi/N_{\text{pix}}$ represents area for each $N_{\text{side}} = 16$ pixel in unit of steradian ($N_{\text{pix}} = 12N_{\text{side}}^2$).

We first generate random Gaussian Stokes Q and U noise maps at 15 COrE frequencies at $N_{\text{side}} = 16$ and smooth them by the ratio of a polarized Gaussian beam of FWHM $9^\circ$ and the original polarized beam given in Table I. This operation is performed in order to bring all the frequency maps to same beam resolution. Finally, we convert the noise Stokes maps to full sky E noise maps for all frequencies.

We show the mean noise power spectrum obtained from 100 Monte Carlo simulations at some of the COrE frequencies in the Fig. 2 along with the CMB theoretical EE power spectrum $C^{EE}_E$. We see that the mean noise power is lower than the theoretical CMB EE power spectrum for all multipoles.

### IV. METHODOLOGY

We implement our method to estimate the joint posterior for CMB E mode given the foreground contaminated CMB E mode maps in a (foreground) model independent manner. We use the foreground and noise model provided by the COrE science team and simulate foreground and noise contaminated CMB E mode input maps at all 15 frequencies.

| Frequency $\nu$ (GHz) | Beam FWHM $\Delta Q = \Delta U$ (arcmin) | $\mu$K.arcmin |
|-----------------------|------------------------------------------|----------------|
| 60                    | 17.87                                    | 7.49           |
| 70                    | 15.39                                    | 7.07           |
| 80                    | 13.52                                    | 6.78           |
| 90                    | 12.08                                    | 5.16           |
| 100                   | 10.92                                    | 5.02           |
| 115                   | 9.56                                     | 4.95           |
| 130                   | 8.51                                     | 3.89           |
| 145                   | 7.68                                     | 3.61           |
| 160                   | 7.01                                     | 3.68           |
| 175                   | 6.45                                     | 3.61           |
| 195                   | 5.84                                     | 3.46           |
| 220                   | 5.23                                     | 3.81           |
| 255                   | 4.57                                     | 5.58           |
| 295                   | 3.99                                     | 7.42           |
| 340                   | 3.49                                     | 11.10          |

TABLE I. COrE frequency maps used in this work.
COrE frequency channels at a pixel resolution defined by the HEALPix pixel resolution parameter $N_{\text{side}} = 16$ and smoothed by a Gaussian beam of FWHM $9^\circ$. Our foreground model consists of synchrotron and thermal dust. The noise model used is the one provided by 4 year COrE mission. The random CMB E mode realization used in the simulated input frequency maps is generated using the CMB theoretical angular power spectrum consistent with Planck 2015 results [59]. The details of the simulated input maps are given in the Table I. After simulating the foreground and noise contaminated input CMB E mode maps we remove the monopole and dipole components from all the maps. In our analysis, in order to sample the joint posterior density of CMB $P(\mathbf{S}, C_E^{\ell} | \mathbf{D})$ our algorithm has 10 independent chains and each chain consist of 10000 Gibbs steps. The initial point of each chain is obtained by randomly generating an initial value of $C_E^{\ell}$ from a uniform distribution within $\pm 3\Delta C_E^{\ell}$ around the Planck best fit theoretical power spectrum, where $\Delta C_E^{\ell}$ is the cosmic variance error.

We sample a CMB theoretical EE angular power spectrum and a cleaned CMB E mode map given the input CMB maps following the procedure as outlined in the Section II at each Gibbs step for every chain. At any given instance, we use Eqn. 7 to sample $\mathbf{S}$. The weights to be used in this equation are obtained by using Eqn. 9 in terms of matrix $\mathbf{A}$. The elements of $\mathbf{A}$ are computed following the Eqn. 11 using the last sampled $C_E^{\ell}$ and the corresponding sky $\hat{C}_E^{\ell}$. After an initial burn-in phase each of the Gibbs chains stabilizes. We remove a conservative choice of 50 samples for the burn-in period. Hence for analysis of the results we are left with 99500 samples of $C_E^{\ell}$ and $\mathbf{S}$. We note that the joint posterior density CMB E mode signal over large angular scales derived in this work is independent on the polarized foreground models since we use ILC method to remove the foregrounds. This causes the posterior density to be insensitive to any uncertainty that may exists in any one of the polarized foregrounds.

V. RESULTS

In this section we discuss the results obtained by using the sampled E mode cleaned maps and theoretical angular power spectrum after burn-in rejection. we also show that the noise levels of COrE 15 frequencies used in this work are sufficiently small and can be ignored for large angular analysis of the CMB E mode signal using our method.

A. Cleaned Maps

We estimate the marginalized probability density of CMB E mode polarization signal at each pixel by using a total of 99500 samples of the cleaned CMB E mode maps obtained from all the chains. Each of these marginalized probability densities is converted to a normalized density corresponding to each pixel by dividing the respective marginalized probability densities with its mode value. We show these normalized probability densities corresponding to some randomly chosen pixels in Fig. 3. We see from this figure that the normalized densities are very nearly symmetric with minor asymmetric tails. Finally, we estimate the best-fit cleaned CMB E mode map by assigning to each pixel the modal value of the corresponding histogram. We show our estimated best-fit cleaned CMB E mode map in the top panel of the Fig. 4. We see that the estimated posterior cleaned CMB E mode map do not contain any visible foreground residual contamination. The input CMB E mode map used in the simulation is shown in the middle panel of this figure.
FIG. 5. In the top panel we show the difference between the best fit cleaned CMB E mode map and the mean of all cleaned CMB E mode maps. We see that our best fit cleaned map matches well with the mean cleaned CMB E mode map. In the bottom panel we show the standard deviation map obtained from all the difference maps obtained by subtracting the input CMB map from the cleaned E mode maps. We see that the error while reconstructing the pure CMB E mode signal using the input foreground and noise contaminated E mode maps using our method is very small.

The difference of the best-fit cleaned E mode map and the input map is shown in the bottom panel of Fig. 4. The best-fit map agree very well with the pure CMB E mode map. We also estimate a mean cleaned CMB E mode map by taking the mean of all 99500 samples of cleaned maps. We show the difference of the best-fit cleaned CMB E mode map and the mean of all cleaned CMB E modes maps in the top panel of the Fig. 5. We see that both the best-fit and the mean cleaned CMB E mode signal using the input foreground and noise contaminated E mode maps using our method is very small.

B. Cleaned Power Spectrum

We estimate the marginalized densities corresponding to multipoles of the CMB E mode theoretical angular power spectrum. In Fig. 6, we show the normalized densities for certain multipoles which are obtained in the same way as discussed in the Section V A. In Fig. 6, the horizontal axis of each sub plot represents \( \ell(\ell + 1)C_\ell^E/2\pi \) in units of \( \mu K^2 \). The blue vertical lines in each sub figure indicate the boundaries for the asymmetric error bars with 1\( \sigma \) (68.27\%) confidence limits for the corresponding sampled \( C_\ell^E \). From the histogram plots it is evident that the lower multipoles (e.g., \( \ell = 2, 3, ..., 6 \)) \( C_\ell^E \) has asymmetric distribution. The asymmetry gradually reduces as we go to higher multipoles (\( \ell > 20 \)). In the top panel of Fig. 5, we show the best-fit CMB E mode theoretical angular power spectrum at all multipoles in blue obtained by using the locations of the peaks of the marginalized angular power spectrum density function shown in Fig. 6. Also plotted on the best-fit spectrum are the asymmetrical error bars corresponding to 1\( \sigma \) confidence interval for each multipole. The error bars are large at low multipoles and gradually decreases as we go to higher multipoles. The red line indicates the the sky CMB E mode angular power spectrum of this work. Both the spectra indicated by the blue and green match with each other on the scales of this plot. In the bottom panel of Fig. 7, we show the difference between best-fit CMB E mode theoretical angular power spectrum and input CMB E mode \( C_\ell^E \). We see that both the angular power spectra matches well with each other at all multipoles.

C. Residual Noise in the Cleaned E-mode CMB Maps

In the current analysis our simulated input CMB maps have realistic level of detector noise generated using the noise models corresponding to COrE polarization maps after 4 years of observations. Since we are working at pixel resolution \( N_{\text{side}} = 16 \) and smoothed by a Gaussian beam of FWHM = 9\(^\circ\), at this large smoothing, the detector noise present in the smoothed CMB E mode maps is low. The comparison of detector noise power for the multipole range \( \ell = 2 - 32 \) for all COrE frequencies is shown in Fig. 2 along with the CMB EE power spectrum, obtained after removing the beam and pixel effects, as brown line. Even though, we see that the noise power at all multipoles for all COrE frequencies is lower than the CMB E mode power, it is important to study whether the presence of detector noise can lead to any residual noise bias in the best fit E mode cleaned map.

In order to understand the level of residual noise present in the cleaned CMB E-mode maps obtained at each iteration in our method, we performed Wiener filtering of the corresponding cleaned map. We first estimated the 15 \( \times \) 15 empirical noise covariance matrix \( C_N \) in pixel space by cross-correlating the simulated noise.
FIG. 6. We show the normalized histogram plots of theoretical cleaned angular power spectrum of some of multipoles whose horizontal axis represents the $\ell(\ell+1)C_{\ell}^E/2\pi$ in units of $\mu K^2$. The region between the sky blue vertical lines represents the $1-\sigma$ error interval around the best fit. The lower multipoles prominently show asymmetry for chi square distribution with low number of degrees of freedom which becomes gradually symmetrical as we move towards higher multipoles.

FIG. 7. In the top panel we show the cleaned CMB EE power spectrum $C_{\ell}^E$ with asymmetric error bars in blue. The vertical axis is in log scale $\ell(\ell+1)C_{\ell}^E/2\pi$ in units of $\mu K^2$. Logarithmic axis is chosen to highlight the variation of the angular power spectrum over a few orders of magnitude. The power spectrum for the input CMB E mode map is shown in red. We show the difference between the green and in the bottom panel. Both these spectra match well with each other.
Finally, we estimate the Wiener filter, $F_i$, values estimated from the cleaned map obtained at each iteration for a randomly chosen Gibbs chain and for chosen 500 samples. We see clearly that the $\theta$ values of $F_i$ are very close to unity which implies that negligible noise levels are present in the cleaned E mode maps. In the bottom panel we show the difference between the standard deviation maps obtained using the Wiener filtered cleaned E mode map and the set of cleaned E mode maps obtained without any Wiener filtering. The difference between both these standard deviation maps is very small indicating negligible detector noise level is present in the cleaned E maps.

![Figure 8](image.png)

**Figure 8.** In the top panel we show the trace plot of Wiener filter, $F_i$, values estimated from the cleaned map obtained at each iteration for a randomly chosen Gibbs chain and for chosen 500 samples. We see clearly that the $\theta$ values of $F_i$ are very close to unity which implies that negligible noise levels are present in the cleaned E mode maps. In the bottom panel we show the difference between the standard deviation maps obtained using the Wiener filtered cleaned E mode map and the set of cleaned E mode maps obtained without any Wiener filtering. The difference between both these standard deviation maps is very small indicating negligible detector noise level is present in the cleaned E maps.

maps of all 15 COre frequencies. The noise covariance matrix is then used to estimate the noise variance ($\sigma^2_N$) in the cleaned E mode map at each Gibbs step, $i$, following,

$$
\sigma^2_N = \mathbf{W}_i \mathbf{C}_N \mathbf{W}_i^T,
$$

(25)

where, $\mathbf{W}_i$ denotes the $n \times 1$ weight vector obtained at $i^{th}$ Gibbs step. In order to obtain the CMB variance, we randomly generate 50000 different CMB realizations at $N_{\text{side}} = 16$ and at Gaussian beam resolution of FWHM = $9^\circ$. The CMB variance ($\sigma^2_{\text{CMB}}$) is calculated by taking the mean of variance of each individual CMB realizations. Finally, we estimate the Wiener filter, $F_i$ at each Gibbs iteration as follows

$$
F_i = \frac{\sigma^2_{\text{CMB}}}{\sigma^2_{\text{CMB}} + \sigma^2_N}.
$$

(26)

We calculate the $F_i$ values of all the 99500 cleaned maps obtained after burn-in rejection from all the 10 chains. In the top panel of Fig. 8 we show the filter values corresponding to one chain for 500 sample values. We see that the filter values are very close to unity indicating negligible residual noise in the cleaned maps. Filter from each iteration is then used to suppress noise from the corresponding cleaned CMB E mode map. In the bottom panel of Fig. 8 we show the difference between the standard deviation map obtained from all such noise weighted cleaned maps and the standard deviation map from our method which does not use any Wiener filtering. Since the difference of the two standard deviation maps are very small we conclude that both the standard deviation maps matches closely indicating no significant residual noise bias is present in our cleaned E mode maps. Therefore noise bias in the best fit cleaned CMB map is also negligible.

VI. ESTIMATING LIKELIHOOD FUNCTION USING BLACKWELL RAO ESTIMATOR

For cosmological parameter estimation, it is of utmost importance to be able to compute the likelihood of any arbitrary theoretical CMB EE angular power spectrum given the data. Although our method computes posterior density of theoretical CMB E mode angular power spectrum, it is in effect a discrete estimation of the underlying power spectrum. Interestingly, Blackwell-Rao [61] theorem can be used for an improved estimation of the likelihood of the CMB EE power spectrum consistent with the foreground removal algorithm used in this article. The theorem says that we can always find an estimator with an equal or better efficiency than the initial estimator by taking its conditional expectation with respect to a sufficient statistic The transformed estimator using Blackwell-Rao theorem is called Blackwell-Rao estimator. In this section we calculate the Blackwell-Rao estimates of likelihood function of theoretical CMB E mode angular power spectrum by using 99500 samples of theoretical angular power spectrum from our method. The Blackwell-Rao expression for likelihood function of $C^E_\ell$ given $\hat{C}^E_\ell$ is given by

$$
P(C^E_\ell | \hat{C}^E_\ell) \propto \frac{1}{\hat{C}^E_\ell} \left( \frac{\hat{C}^E_\ell}{C^E_\ell} \right)^{(2\ell + 1)/2} \exp \left[ - \frac{2\ell + 1}{2} \frac{\hat{C}^E_\ell}{C^E_\ell} \right],
$$

(27)

where $\hat{C}^E_\ell$ is the $i^{th}$ realization of the power spectrum obtained after burn-in rejection from all Gibbs chains. The likelihood function for $C^E_\ell$ given data $\mathbf{D}$ now can be approximated as [61]

$$
P(C^E_\ell | \mathbf{D}) \approx \frac{1}{N} \sum_{i=1}^{N} P(C^E_\ell | \hat{C}^E_\ell),
$$

(28)

where $N$ denotes total the number of Gibbs samples from all chains after burn-in rejection.

We show the results in Fig. 9. At low multipoles the likelihood distribution is highly asymmetric with a long tail and as we go to higher multipoles the distribution becomes more symmetric. The positions corresponding to
the theoretical power spectrum that was used to generate random realization of CMB E mode map used in this analysis is also shown by blue vertical lines for comparison. The peak positions of likelihood functions of this plot match very well with the best-fit spectrum which was shown in Fig. 7. We show the difference between the two in Fig. 10. An interesting property of the estimated likelihood functions is that they are entirely independent on the explicit foreground models. Hence the likelihood functions are not affected by any modeling uncertainty of the polarized foregrounds. Moreover, these likelihood functions are not affected by the residual E mode foregrounds since we use a large number of input frequency maps which results in a negligible foreground residual effects in the cleaned E mode map and its theoretical angular power spectrum. In a future work these estimators will be directly integrated in cosmological parameter estimation using CMB EE angular power spectrum over large angular scales.

VII. TESTING CONVERGENCE OF THE METHOD

Our implementation of the algorithm consists of 10 chains with different starting points and each chain with
10000 Gibbs steps. Overall after rejecting 50 initial samples of CMB cleaned maps and CMB theoretical angular power spectrum in each chain, we have a total of 99500 samples. It is important to check whether in each chain, a chain length of 10000 Gibbs steps the samples have converged in order to ascertain the sampled distribution converged to the targeted CMB joint posterior distribution. Convergence of the chains also ensures that the final samples no longer depends on the initial points. A powerful diagnostic to check for the convergence is the Gelman Rubin statistic [12] which proposes that lack of convergence in a chain is better diagnosed in the presence of parallel chains with different initial points. We discuss this method.

Let us consider a simulation for estimating the posterior of a model with parameter \( \phi \). Let there are \( N \) number of chains with \( L \) being the number of steps after rejecting the samples during the burn-in phase. Let us assume that the sample posterior mean and variance are given by \( \bar{\phi}_n \) and \( (\bar{\sigma}_n)^2 \) respectively for \( n^{th} \) chain using all the \( L \) samples in the chain. Then the between-chain \( (B/L) \) and within-chain variances \( (W) \) are respectively given by

\[
B = \frac{L}{N - 1} \sum_{n=1}^{N} (\bar{\phi}_n - \bar{\phi})^2 \tag{29}
\]

\[
W = \frac{1}{N} \sum_{n=1}^{N} (\bar{\sigma}_n)^2 \tag{30}
\]

where \( \bar{\phi} \) is the mean of the samples from all the sample chains, \( \bar{\phi} = \frac{1}{N} \left( \sum_{n=1}^{N} \bar{\phi}_n \right) \)

The pooled posterior variance is defined as follows

\[
V = \frac{L - 1}{L} W + \frac{N + 1}{NL} B \tag{31}
\]

The Gelman-Rubin statistic \( R \) is then given by

\[
R = \sqrt{\frac{V}{W}} \tag{32}
\]

According to [12], the chain converges when the value of \( R \) approaches unity \( (L \to \infty) \) which in turn implies the sampled distribution is same as the targeted distribution.

In the top panel of Fig. 11, we shown the Gelman-Rubin statistic \( R \) value for each multipole. The bottom panel shows the same for the pixel values. From the figure its clear that the \( R \) values are close to 1 implying the chains have achieved convergence.
VIII. DISCUSSION AND CONCLUSION

Accurate measurement of CMB E mode signal provide valuable physical information about the ionization history of the universe. In our work, we estimate CMB E mode signal and its theoretical angular power spectrum joint posterior density over large angular scales of the sky using simulated observations of 15 frequency maps between 60 to 340 GHz of future generation COrE satellite observations. We employ the ILC method augmented by Gibbs sampling technique to draw samples from the joint density. The input frequency maps contain E mode foregrounds and detector noise contamination along with CMB E mode signal compatible to Planck 2015 theoretical angular power spectrum [59]. Foregrounds in input maps consist of two major polarized components, namely, the synchrotron and thermal dust emissions. The detector noise level used in this work is compatible to 4 years of observations of COrE. We run a total of 10 Gibbs chains each comprising 10000 numbers. We reject 50 samples from each chain due to burn-in removal which results in a total of 99500 usable samples. We test convergence of these chains by evaluating Gelman-Rubin statistic and conclude reliable convergence has been achieved in these chains.

Using the complete joint posterior density we evaluate the marginal densities of the CMB E mode map and theoretical angular power spectrum. We see that the best-fit cleaned CMB E map agree excellently with the input CMB E mode map. The best-fit E mode angular power spectrum agree with the sky angular power spectrum very well. The error in reconstruction is CMB E mode map is small for almost all over the sky. The reconstruction error is seen to be large only at a couple of pixels at the galactic center with a standard deviation of \(~0.0061 \mu K\). The samples of theoretical E mode angular power spectra are used to estimate the asymmetric error levels on the corresponding best-fit estimates. The best-fit E mode angular power spectrum does not show any indication of presence of a bias in the multipole range \(2 \leq \ell \leq 32\) studied in this work.

Since polarization signal is weak, presence of significant detector noise in the input frequency maps may lead to foreground residuals in the cleaned maps as well as foreground residuals and detector noise bias in the cleaned E mode angular power spectrum. In order to assess the level of detector noise in the cleaned maps obtained in the Gibbs chains we perform Wiener filtering of all the sampled clean CMB E mode maps obtained. We find that the Wiener filtered CMB E maps agree very well with the unfiltered maps implying detector noise can safely be ignored for the large angular scales analysis for CMB E mode using 4 year COrE noise levels.

Using all Gibbs samples of CMB theoretical E mode angular power spectrum after burn in rejection we estimate the likelihood function of any arbitrary angular power spectrum given the simulated data of 15 COrE frequency maps using Blackwell-Rao estimator. The likelihood function can be directly integrated to a Markov Chain Monte Carlo algorithm to estimate the cosmological parameters relevant for large angular scales CMB E mode observations. The CMB E mode component reconstruction method of this article thus possesses an attractive property of accurately and reliably used in cosmological parameter estimation methodology by predicting exact likelihood function of the E mode theoretical angular power spectrum given the observed data after performing E mode foreground removal.

The method outlined in this work has another excellent advantage that the estimated CMB joint posterior is not affected by the inaccuracies in foreground modelling since the weight factors for linear combination of input maps are estimated without assuming any foreground model. This is particularly very advantageous since the CMB E mode signal is very weak compared to the E mode foregrounds. A small modelling uncertainty in the foregrounds if exists may lead to large error (or even a bias) in the reconstructed CMB E mode signal, if explicit foreground modelling is necessary in order to remove the foregrounds. Moreover, the effects of foreground residuals in the reconstructed joint posterior density following our method is negligible since we have a large number of input frequency maps which is larger than typical number of independent spectral index parameters required for modelling the polarized foregrounds. In a future article, we will extend the work for reconstruction of CMB B mode over large angular scales.

This work is based on the CMB observations of Planck an ESA mission. For our work we have used the freely publicly available [15] HEALPix Software package \url{http://healpix.sourceforge.net} for some of the analysis of this work. We acknowledge the use of Planck Legacy Archive (PLA) and the Legacy Archive for Microwave Background Data Analysis (LAMBDA). LAMBDA is a part of the High Energy Astrophysics Science Archive Center (HEASARC). HEASARC/LAMBDA is supported by the Astrophysics Science Division at the NASA Goddard Space Flight Center.

[1] Baccigalupi, C. 2003, New Astron. Rev., 47, 1127
[2] Basak, S., & Delabrouille, J. 2011, Mon. Not. Roy. Astron. Soc. 419, 1163
[3] Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003, The Astrophysical Journal Supplement, 148, 1
[4] Bottino, M., Banday, A. J., & Maino, D. 2010, Mon. Not. Roy. Astron. Soc. 402, 207
[5] Bouchet, F. R., Prunet, S., & Sethi, S. K. 1999, Mon. Not. Roy. Astron. Soc. 302, 663
[6] Bunn, E. F., Fisher, K. B., Hoffman, Y., et al. 1994, The
