

Extended delta-map: A map-based foreground removal method for CMB polarization observations

Yuto Minami ^{1,*} and Kiyotomo Ichiki ^{2,3,4}

¹*Research Center for Nuclear Physics, Osaka University, Ibaraki, Osaka 567-0047, Japan*

²*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University, Furouchou, Chikusa-ku, Nagoya, Aichi 464-8602, Japan*

³*Graduate School of Science, Nagoya University, Furouchou, Chikusa-ku, Nagoya, Aichi 464-8602, Japan*

⁴*Institute for Advanced Research, Nagoya University, Furouchou, Chikusa-ku, Nagoya Aichi 464-8602, Japan*

*E-mail: yminami@rcnp.osaka-u.ac.jp

Received December 16, 2022; Revised January 20, 2023; Accepted January 21, 2023; Published January 28, 2023

.....
 In order to extract information about inflationary gravitational waves using B -mode patterns of cosmic microwave polarization anisotropy, we need to remove the foreground radiation from the Milky Way. In our previous delta-map method for foreground removal, the number of observation bands was limited to the number of parameters of the assumed foreground model, and therefore it was difficult to improve the sensitivity by increasing the number of observation bands. Here, we extend the previous method so that it can be adapted to an arbitrary number of observation bands. Using parametric likelihood and realistic foreground and CMB simulations, we show that our method can increase the sensitivity to the tensor-to-scalar ratio r without inducing any significant bias.

Subject Index E63, F11, F14

1. Introduction

Precise measurements of linear polarization patterns of the cosmic microwave background (CMB) provide us with plentiful information of our Universe [1–9]. In particular, a measurement of the B -mode pattern in CMB polarization anisotropies is a key to detecting primordial gravitational waves (PGWs) expected in the inflation scenario of the early Universe [10–14]. However, celestial sources also emit microwave photons with linear polarization that mimic the cosmological B -mode, the so-called foreground emission, which disturbs the precise measurement of the CMB B -mode [15]. The two representative foreground emissions in CMB polarization measurements are synchrotron and thermal dust emissions, which have different spectral energy distributions from the black body and dominate at lower and higher frequencies, respectively. Therefore, we can remove the foreground emissions using multiple-band observations owing to their frequency dependences.

Many foreground removal and component separation methods exist to extract the cosmological B -mode signal, e.g., Commander [16], SEVEM [17], SMICA [18], NILC [19], and GNILC [20]. Among these, we proposed a method called “delta-map” based on a linear combination of the observed maps to remove the foreground components [21]. This method allows the directional dependence of the frequency spectrum of foreground emissions up to the

first order to be taken into account by considering the differences in the observation maps¹. Because the method uses one additional frequency map to eliminate the directional dependence of one foreground parameter, the usable number of maps is limited by the number of foreground parameters in the assumed model. Specifically, suppose that we assume one foreground model with N parameters and CMB signal in all the multi-frequency maps. In this case, the delta-map method requires one map for the CMB, one map for the zeroth (or spatially uniform) foreground parameter, and N maps for the first-order expansion of the N parameters; we need exactly $(2 + N)$ maps. This prevents us from obtaining improvement of the sensitivity in observation by increasing the number of observing bands. More concretely, if we consider a power-law synchrotron foreground model with one parameter β_s and a one-component dust model with two parameters β_d and T_d , the number of observation bands should be exactly six. This is a waste of resources since some future CMB missions plan to have plenty of bands; e.g., LiteBIRD will have 15 bands [26,27]. In this paper, we improve the method by constructing a parametric likelihood in a Bayesian way so that more observation bands can be used.

The rest of this paper is organized as follows. In Sect. 2, we first review the previous delta-map method that is based on the linear combination, and then derive a new delta-map method based on a parametric likelihood. In Sect. 3, we explain the foreground models and simulations used in this work. We show the results of the measurements of CMB parameters with some sky simulation setups in Sect. 4. In Sect. 5, we discuss the results and conclude this work.

2. Methodology

We first introduce the delta-map method described in Ref. [21]. We decompose a linear polarization signal of a microwave component at a frequency, ν , from a line of sight, \hat{n} , to two orthogonal Stokes parameters, $Q(\hat{n})$ and $U(\hat{n})$, which we observe in thermodynamic temperature units. Let us vectorize all $Q(\hat{n})$ and $U(\hat{n})$ from all \hat{n} in use and concatenate them into one column vector as

$$\vec{s} = (Q(\hat{n}_1), \dots, Q(\hat{n}_{N_{\text{pix}}}), U(\hat{n}_1), \dots, U(\hat{n}_{N_{\text{pix}}}))^T, \quad (1)$$

where N_{pix} is the number of pixels in use and the subscript T represents a transpose.

The CMB signal is common in any frequency band except for beam effects of instruments, and can be written as $\vec{s}_{\text{CMB}}(\nu) = \vec{s}_{\text{CMB}}$. We assume that each frequency has each independent Gaussian noise, \vec{s}_N . A foreground signal varies in frequency space according to N parameters that vary over the sky, \vec{p}^I ($I = 1, 2, \dots, N$), as

$$\vec{s}_f(\nu) = g_\nu \mathbf{D}_\nu(\vec{p}^I) \vec{s}_b, \quad (2)$$

where \vec{s}_b is a signal vector at a pivot frequency in brightness temperature units, $\mathbf{D}_\nu(\vec{p}^I)$ is a diagonal matrix with the same dimension as \vec{s}_f , and g_ν is the conversion factor from the brightness temperature to the CMB thermodynamic temperature given by

$$g_\nu \equiv \frac{(e^x - 1)^2}{e^x x^2} \text{ with } x \equiv \frac{h\nu}{k_B T_{\text{CMB}}}. \quad (3)$$

In this paper, we use the power-law synchrotron and one-component modified black body (1MBB) models to fit the synchrotron and dust foreground emissions, respectively. The func-

¹A similar perturbative approach in multipole space has also been investigated in Refs. [22–25].

tional form of one pixel of $\mathbf{D}_v(\vec{p}^I)$ for the power-law synchrotron is

$$D_v^s(\beta_s(\hat{n})) = \left(\frac{\nu}{\nu_{s*}} \right)^{\beta_s(\hat{n})}, \quad (4)$$

where ν_* is the reference frequency, which we set as $\nu_{s*} = 23$ GHz, and $\beta_s(\hat{n})$ is the synchrotron spectral index. The one for 1MBB is

$$D_v^d(T_d(\hat{n}), \beta_d(\hat{n})) = \left(\frac{\nu}{\nu_{d*}} \right)^{\beta_d(\hat{n})+1} \frac{e^{x_{d*}(\hat{n})} - 1}{e^{x_d(\hat{n})} - 1}, \quad (5)$$

where $x_{d*}(\hat{n}) \equiv \frac{h\nu_{d*}}{k_B T_d(\hat{n})}$, $x_d(\hat{n}) \equiv \frac{h\nu}{k_B T_d(\hat{n})}$, $T_d(\hat{n})$ is the dust temperature, $\beta_d(\hat{n})$ is the dust spectral index, and we set $\nu_{d*} = 353$ GHz.

The basic idea of the delta-map method is to consider a spatial variation of foreground signal parameters up to the first-order expansion² as

$$g_v \mathbf{D}_v(\vec{p}^I) \vec{s}_b = g_v D_v(\vec{p}^I) \mathbf{I} \vec{s}_b + g_v \sum_{I=1}^N D_{v,p^I}(\vec{p}^I) \mathbf{I} (\delta \vec{p}^I \circ \vec{s}_b) + \mathcal{O}(\delta \vec{p}^{I2}), \quad (6)$$

where \vec{p}^I represents the mean value of the parameter p^I over the sky, \mathbf{I} is an identity matrix, and “ \circ ” means a Hadamard product. Hereafter, we consider two representative foreground components: synchrotron radiation and thermal dust emission. With these two foreground components, the foreground signal is written as

$$\begin{aligned} g_v \mathbf{D}_v(\vec{p}^I) \vec{s}_f &= g_v D_v^d(\vec{p}^I) \mathbf{I} \vec{s}_d + g_v \sum_{I=1}^{N_d} D_{v,p_d^I}(\vec{p}^I) \mathbf{I} (\delta \vec{p}_d^I \circ \vec{s}_d) + \mathcal{O}(\delta \vec{p}_d^{I2}) \\ &\quad + g_v D_v^s(\vec{p}^I) \mathbf{I} \vec{s}_s + g_v \sum_{I=1}^{N_s} D_{v,p_s^I}(\vec{p}^I) \mathbf{I} (\delta \vec{p}_s^I \circ \vec{s}_s) + \mathcal{O}(\delta \vec{p}_s^{I2}) \\ &= \tilde{\mathbf{D}}_v \vec{s}_f, \end{aligned} \quad (7)$$

where the superscripts and subscripts “ s ” and “ d ” denote synchrotron and dust emission foreground components, respectively, and N_s and N_d denote the numbers of parameters for synchrotron and dust foreground emissions, respectively. Here $\tilde{\mathbf{D}}_v$ represents the frequency dependence of the signals up to the first order \vec{s}_f and is given by

$$\tilde{\mathbf{D}}_v = \begin{pmatrix} g_v D_v^d(\vec{p}_d^1) \mathbf{I} & g_v D_{v,p_d^1}(\vec{p}_d^1) \mathbf{I} & \cdots & g_v D_{v,p_d^{N_d}}(\vec{p}_d^{N_d}) \mathbf{I} & g_v D_v^s(\vec{p}_s^1) \mathbf{I} & g_v D_{v,p_s^1}(\vec{p}_s^1) \mathbf{I} & \cdots & g_v D_{v,p_s^{N_s}}(\vec{p}_s^{N_s}) \mathbf{I} \end{pmatrix} \quad (8)$$

and, in the last line, we define the foreground signal vector as

$$\vec{s}_f \equiv \begin{pmatrix} \vec{s}_d \\ \delta \vec{p}_d^1 \circ \vec{s}_d \\ \vdots \\ \delta \vec{p}_d^{N_d} \circ \vec{s}_d \\ \vec{s}_s \\ \delta \vec{p}_s^1 \circ \vec{s}_s \\ \vdots \\ \delta \vec{p}_s^{N_s} \circ \vec{s}_s \end{pmatrix}. \quad (9)$$

In summary, we assume that the observed data can be decomposed as

$$\vec{m}(\nu) = \vec{s}_{\text{CMB}} + \tilde{\mathbf{D}}_v \vec{s}_f + \vec{s}_N(\nu). \quad (10)$$

²The standard deviation of directional variation of foreground parameters considered in this paper is less than 10% of the spatial uniform parameter. We can assume the second-order variation to be less than 1%; hence we neglect higher-order terms.

We need to observe the sky at multiple frequencies to remove foreground components using its frequency dependence. When we observe the sky at N_ν frequencies, the total data can be expressed as

$$\vec{m} = \mathbf{D}^{\text{CMB}} \vec{s}_{\text{CMB}} + \tilde{\mathbf{D}} \vec{s}_f + \begin{pmatrix} \vec{s}_N(\nu_1) \\ \vdots \\ \vec{s}_N(\nu_{N_\nu}) \end{pmatrix}, \quad (11)$$

where

$$\vec{m} \equiv \begin{pmatrix} \vec{m}(\nu_1) \\ \vdots \\ \vec{m}(\nu_{N_\nu}) \end{pmatrix}, \quad (12)$$

$$\mathbf{D}^{\text{CMB}} \equiv \begin{pmatrix} \mathbf{I} \\ \vdots \\ \mathbf{I} \end{pmatrix} \quad \text{and} \quad \tilde{\mathbf{D}} \equiv \begin{pmatrix} \tilde{\mathbf{D}}_{\nu_1} \\ \vdots \\ \tilde{\mathbf{D}}_{\nu_{N_\nu}} \end{pmatrix}. \quad (13)$$

This is the baseline expression of the ‘‘delta-map’’ method.

In Ref. [21], one frequency, ν_1 , was chosen as the CMB channel and other channels with weights $\vec{\alpha}^\top = (\alpha_{\nu_2}, \dots, \alpha_{\nu_{N_\nu}})$ were added to remove foreground contributions and to have a cleaned CMB map as

$$\vec{m}_{\text{CMB}} = \frac{\vec{m}_{\nu_{\text{CMB}}} + \sum_{i=\nu_2}^{\nu_{N_\nu}} \alpha_i \vec{m}_{\nu_i}}{1 + \sum_{i=\nu_2}^{\nu_{N_\nu}} \alpha_i} \quad (14)$$

with

$$[1, \vec{\alpha}^\top] \tilde{\mathbf{D}} = \mathbf{O}. \quad (15)$$

The above equation can be solved only when the number of frequencies, N_ν , is equal to the number of degrees of freedom, $(N_d + 1) + (N_s + 1) + 1$, where N_s and N_d are the number of parameters of the synchrotron radiation and thermal dust emission models, respectively. This is the caveat and weak point of the previous delta-map method because we cannot increase the sensitivity by increasing N_ν . We shall mitigate this point in the following sections.

Since the cleaned CMB map should contain only the CMB and the combined noise, we can construct our likelihood as

$$-2 \ln \mathcal{L} = \vec{m}_{\text{CMB}}^\top \mathbf{C}^{-1} \vec{m}_{\text{CMB}} + \ln |2\pi \mathbf{C}|, \quad (16)$$

where the covariance matrix \mathbf{C} is given by [21,28]

$$\mathbf{C} = r \times \mathbf{C}^{\text{tens}} + \mathbf{C}^{\text{scal}} + \frac{\mathbf{N}_{\nu_{\text{CMB}}} + \sum_{i=\nu_2}^{\nu_{N_\nu}} \alpha_i^2 \mathbf{N}_{\nu_i}}{(1 + \sum_{i=\nu_2}^{\nu_{N_\nu}} \alpha_i)^2}. \quad (17)$$

By minimizing this likelihood, we can determine a CMB parameter, r , and foreground parameters, $p_d^{I_d}$ and $p_s^{I_s}$, where $I_d \in (1, \dots, N_d)$ and $I_s \in (1, \dots, N_s)$. It is known that the determined foreground parameters are biased when we use the full likelihood function [29]. We can avoid the bias by using the χ^2 term, namely, the first term of Eq. (16). Technically, the following iteration scheme was adapted in Ref. [21] to determine both the CMB and foreground parameters.

2.1 Extended delta-map

In this section we introduce our method using a parametric likelihood. We describe the details of the derivation in Appendix A.

Algorithm 1 Iteration algorithm of the minimization

Set initial values of $p_d^{I^d}$ and $p_s^{I^s}$
 $r_{\text{pre}} = \infty$
 $r_{\text{out}} = 1e2$
 $-2\mathcal{L}_{\text{pre}} = \infty$
 $-2\mathcal{L}_{\text{out}} = 1e10$
Set initial values of foreground parameters
while $(-2\mathcal{L}_{\text{pre}} + 2\mathcal{L}_{\text{out}}) > 1e-2$ & $(r_{\text{pre}} - r_{\text{out}}) > 1e-5$ **do**
 Fix r_{out} and minimize χ^2 term against $p_d^{I^d}$ and $p_s^{I^s}$
 Fix foreground and minimize all the $-2\mathcal{L}$ terms against r_{out}
 Set r_{pre} with r_{out}
 Set $-2\mathcal{L}_{\text{pre}}$ with $-2\mathcal{L}_{\text{out}}$
 Set $-2\mathcal{L}_{\text{out}}$ with the minimized $-2\mathcal{L}$
end while
Return r_{out} , $p_d^{I^d}$, and $p_s^{I^s}$

We start from Eq. (11). By subtracting the CMB and foreground terms from the observed maps \vec{m} , we can construct the likelihood of data as

$$-2 \ln \mathcal{L}(\vec{m} | \vec{s}_{\text{CMB}}, \vec{s}_f, \tilde{\mathbf{D}}) = (\vec{m} - \mathbf{D}^{\text{CMB}} \vec{s}_{\text{CMB}} - \tilde{\mathbf{D}} \vec{s}_f)^\top \mathbf{N}^{-1} (\vec{m} - \mathbf{D}^{\text{CMB}} \vec{s}_{\text{CMB}} - \tilde{\mathbf{D}} \vec{s}_f) + \ln |2\pi \mathbf{N}|, \quad (18)$$

where $\mathbf{N} = \text{diag}(\mathbf{N}_{v_1}, \dots, \mathbf{N}_{v_{N_v}})$ is a noise covariance matrix.

We use Bayes' theorem to relate the posterior distribution to the likelihood as

$$P(\vec{s}_{\text{CMB}}, \vec{s}_f, \vec{p}^f, \mathbf{S}_0^{\text{CMB}}, \mathbf{S}^f | \vec{m}) = \frac{\mathcal{L}(\vec{m} | \vec{s}_{\text{CMB}}, \vec{s}_f, \tilde{\mathbf{D}}(\vec{p}^f)) \cdot P(\vec{s}_{\text{CMB}}, \vec{s}_f, \tilde{\mathbf{D}}(\vec{p}^f), \mathbf{S}_0^{\text{CMB}}, \mathbf{S}^f)}{P(\vec{m})}, \quad (19)$$

where $\mathbf{S}_0^{\text{CMB}}$ and \mathbf{S}^f are covariance matrices of the CMB, \vec{s}_{CMB} , and foreground signal, \vec{s}_f , respectively; \vec{p}^f represents the foreground parameters of both synchrotron and thermal dust emissions; and $P(\vec{m})$ is a normalization factor. Here we use a flat prior on \vec{p}^f and have $P(\vec{s}_{\text{CMB}}, \vec{s}_f, \tilde{\mathbf{D}}(\vec{p}^f), \mathbf{S}_0^{\text{CMB}}, \mathbf{S}^f) \propto P(\vec{s}_{\text{CMB}}, \vec{s}_f, \mathbf{S}_0^{\text{CMB}}, \mathbf{S}^f)$. We marginalize it over the CMB signal assuming a Gaussian distribution,

$$-2P(\vec{s}_{\text{CMB}}, \mathbf{S}_0^{\text{CMB}}) = \vec{s}_{\text{CMB}}^\top \mathbf{S}_0^{\text{CMB}^{-1}} \vec{s}_{\text{CMB}} + \ln |2\pi \mathbf{S}_0^{\text{CMB}}|, \quad (20)$$

to have

$$-2 \ln P(\vec{p}^f, \vec{s}_f, \mathbf{S}^f, \mathbf{S}_0^{\text{CMB}} | \vec{m}) = (\vec{m} - \tilde{\mathbf{D}} \vec{s}_f)^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} (\vec{m} - \tilde{\mathbf{D}} \vec{s}_f) + \ln |2\pi (\mathbf{S}^{\text{CMB}} + \mathbf{N})| - 2 \ln P(\vec{s}_f, \mathbf{S}^f) + \text{const.}, \quad (21)$$

where $\mathbf{S}^{\text{CMB}} = \mathbf{D}^{\text{CMB}} \mathbf{S}_0^{\text{CMB}} \mathbf{D}^{\text{CMB}^\top}$. This is Eq. (58) in Ref. [21].

From here, we deal with the $P(\vec{s}_f, \mathbf{S}^f)$ term. Our strategy is to marginalize the foreground probability function, assuming that its mean is zero, $\vec{s}_f = \vec{0}$, and its covariance is ‘‘vague’’, $\mathbf{S}^f \rightarrow \mathbf{O}$. We follow the methodology described in Sect. 2 of Ref. [30] to deal with the ‘‘vague’’

foreground covariance matrix. Following the method, we obtain

$$\begin{aligned}
-2 \ln P(\vec{p}^f, \mathbf{S}_0^{\text{CMB}} | \vec{m}) &= -\vec{m}^T \mathbf{N}^{-1} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}T} \mathbf{N}^{-1} \vec{m} \\
&\quad - \vec{M} \mathbf{H} \vec{M} - \vec{M} \mathbf{H} \mathbf{B}^{-1} \mathbf{H} \vec{M} \\
&\quad + \ln |\mathbf{S}_0^{\text{CMB}}| + \ln |\tilde{\mathbf{D}}^T \mathbf{N}^{-1} \tilde{\mathbf{D}}| + \ln |B| + \text{const.}, \tag{22}
\end{aligned}$$

where

$$\mathbf{A} = \left(\mathbf{S}_0^{\text{CMB}^{-1}} + \sum_{j=1}^{N+1} \mathbf{N}_{v_j}^{-1} \right), \tag{23}$$

$$\mathbf{H} = \mathbf{N}^{-1} \tilde{\mathbf{D}} [\tilde{\mathbf{D}}^T \mathbf{N}^{-1} \tilde{\mathbf{D}}]^{-1} \tilde{\mathbf{D}}^T \mathbf{N}^{-1}, \tag{24}$$

$$\mathbf{B} = \mathbf{A} - \mathbf{D}^{\text{CMB}T} \mathbf{H} \mathbf{D}^{\text{CMB}}, \tag{25}$$

$$\vec{M} = \vec{m} - \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}T} \mathbf{N}^{-1} \vec{m}, \tag{26}$$

and put the noise terms, $\vec{m}^T \mathbf{N}^{-1} \vec{m} + \ln |\mathbf{N}|$, into the const. term. For details of the derivation of Eq. (22), see Appendix A. To determine the CMB parameter (r) and foreground parameters (\vec{p}^f), we follow the same procedure, namely Algorithm 1, as in Ref. [21]. We set the initial values of all parameters as $r = 0.5$, $\beta_s = -3.0$, $\beta_d = 1.5$, and $T_d = 20.1$ K. We set a lower bound on $r \geq 0.0$ so that the CMB covariance matrix is not singular and set boundaries on the foreground parameters as $\beta_s \in (-10.0, -0.01)$, $\beta_d \in (0.1, 10.0)$, and $T_d \in (5.0, 40.0)$.

We obtain the ‘‘extended delta-map’’ likelihood as Eq. (22), which allows us to determine the CMB and foreground parameters without the band number constraint. This likelihood has another benefit. Because we use the ‘‘matrix inversion lemma’’ (Appendix B) in the derivation, we can reduce the computational cost in using the Cholesky solver for $\mathbf{S}^{\text{CMB}} + \mathbf{N}$ in Eq. (21), which is a positive definite symmetric matrix with a dimension of $2N_{\text{pix}}N_v$. In Eq. (22), in contrast, we only need to use the Cholesky solver for some positive definite symmetric matrices, e.g., \mathbf{B} , with a smaller size of $2N_{\text{pix}}(N_d + N_s + 3)$.

Because the form of Eq. (22) looks so different from that of Eq. (16), one may suspect that they are totally different methods. We find that estimated values of the parameters with this likelihood are equivalent to those estimated from the previous delta-map method in the case where $N_v = (N_d (= 1) + 1) + 1$. We describe the comparison in Appendix C.

All of the numerical codes for the calculations above have been implemented in an extended-deltamap GitHub repository³.

3. Models and simulations

We use simulations to validate our methodology. We use the ‘‘PySM’’ package [31] to produce polarized synchrotron and thermal dust emission maps with direction-dependent spectral parameters. For the synchrotron map, we use the power-law synchrotron model, ‘‘s1’’, which is based on the Q and U maps from WMAP-9 [32] and the spectral index map from Ref. [33]. For the thermal dust emission map, we use the one-component modified black body (MBB) model, ‘‘d1’’ [34], and the two-component MBB model, ‘‘d4’’ [35], both of which are based on the Planck HFI results. Note that the results used the intensity map in addition to the Q and U polarization maps, and the model adopted common foreground parameters β_d and T_d for the

³<https://github.com/YutoMinami/extended-deltamap>.

Table 1. Polarization sensitivity and beam size of the telescopes similar to the LiteBIRD mission [26]. The parameters in brackets are for the modified high-frequency model used in Sect. 4.4.

Frequency (GHz)	Polarization sensitivity ($\mu\text{K}'$)	Beam size in FWHM (arcmin)
40	37.5	69
50	24.0	56
60	19.9	48
68	16.2	43
78	13.5	39
89	11.7	35
100	9.2	29
119	7.6	25
140	5.9	23
166	6.5	21
195	5.8	20
235	7.7	19
280	13.2	24
337 (500)	19.5 (67.8)	20 (13.7)
402 (600)	37.5 (113.5)	17 (11.5)

Q and U Stokes parameters at each sky pixel. Cosmological CMB maps are generated using the `synfast` function of the HEALPix package [36] from the power spectra calculated using CAMB [37] with the Planck 2018 cosmological parameters for “TT,TE,EE+lowE+lensing” [2]: $\Omega_b h^2 = 0.02237$, $\Omega_c h^2 = 0.1200$, $h = 0.6736$, $\tau = 0.0544$, $A_s = 2.100 \times 10^{-9}$, and $n_s = 0.9649$. We generate CMB maps with some values of the tensor-to-scalar ratio, r .

We use experimental parameters such as frequency bands and angular resolutions similar to the LiteBIRD mission [26] (Table 1). For the input noise, we assume white noise with a standard deviation of $\sigma_N = (\pi/10800)(w_p^{-1/2}/\mu\text{K}') \mu\text{K str}^{-1/2}$ [28], where we use the “Polarization sensitivity” column of Table 1 for the values of $w_p^{-1/2}$.

We incorporate the beam smearing effect from the finite angular resolution, whose full-width-half-maximum (FWHM) values are given by the “Beam size in FWHM” column of Table 1, by multiplying the spherical harmonics coefficients of the CMB and foreground maps by the corresponding Gaussian beam transfer function for individual frequencies. In this paper, we use the map resolution parameter of $N_{\text{side}} = 4$ and set $\ell_{\text{max}} = 2N_{\text{side}}$. To ensure that the maps are limited to low resolution, we follow the method described in Ref. [38]; specifically, we deconvolve each frequency map and reconvolve all the maps with a Gaussian 2200 arcmin beam, which is 2.5 times the pixel size of $N_{\text{side}} = 4$. Even though our method can be applied to higher-resolution maps, we choose $N_{\text{side}} = 4$ (or $\ell_{\text{max}} = 8$), which covers the so-called reionization bump. Since we do not apply any delensing scheme, we cannot effectively increase the sensitivity to r because of the cosmic variance of the lensing B -mode. If we can increase ℓ_{max} further to cover the so-called recombination bump, we could increase the sensitivity. However, such high-resolution analysis is limited by our computational resources.

In Sect. 4.4, we will show the case in which the frequencies of the two highest-frequency bands are increased. For that, we show the replaced parameters in brackets. The replaced polarization sensitivity and beam size in FWHM are extrapolated using a power-law function.

Since our method approximates the directional dependence of foreground parameters up to the first order, we need to mask the very bright Galactic plane. We follow Ref. [21] and use the

Table 2. Summary of our analysis. Estimated r_{out} values are from the 50th percentile. For $r_{\text{in}} = 1.0 \times 10^{-2}$, we use the (16th, 84th) percentiles for the uncertainties. For null r_{in} , we use the 95% C.L. as the upper bounds. “(High)” indicates the result with the “modified high-frequency model”, where we increase the frequencies of the two highest-frequency bands.

Number of bands	Input synchrotron	Input dust	T_d prior	$r_{\text{in}} \times 10^{-2}$	$r_{\text{out}} \times 10^{-2}$
3	s1	–	–	1.0	$1.00^{+0.18}_{-0.33}$
9	s1	–	–	1.0	$0.99^{+0.15}_{-0.28}$
3	s1	–	–	0.0	<0.12 (95% C.L.)
9	s1	–	–	0.0	<0.07 (95% C.L.)
4	–	d1	flat	1.0	$0.95^{+1.43}_{-0.95}$
9	–	d1	flat	1.0	$1.13^{+0.61}_{-0.49}$
15	s1	d1	flat	1.0	$0.96^{+0.49}_{-0.96}$
15	s1	d1	1σ	1.0	$1.08^{+0.52}_{-0.90}$
15	s1	d1	$10^{-5}\sigma$	1.0	$1.18^{+0.50}_{-0.83}$
15	s1	d1	flat	0.0	<1.15 (95% C.L.)
15	s1	d1	1σ	0.0	<1.26 (95% C.L.)
15 (High)	s1	d1	flat	1.0	$1.40^{+0.63}_{-0.48}$
15 (High)	s1	d1	1σ	1.0	$1.41^{+0.56}_{-0.42}$
15	s1	d4	flat	1.0	$2.00^{+1.09}_{-0.85}$
15	s1	d4	1σ	0.0	$2.19^{+1.26}_{-0.93}$

“P06 mask” by WMAP polarization analysis, whose $f_{\text{sky}} = 0.56$. Since our method uses pixel space maps, we do not apply any apodization to the mask.

4. Results

We apply our new method to simulated CMB maps with foregrounds. In the following sections, we show the results of estimating the tensor-to-scalar ratio r with CMB + synchrotron (Sect. 4.1), CMB + 1MBB dust (Sect. 4.2), and CMB + 1MBB dust + synchrotron (Sect. 4.3) maps. Recently, it has been shown that shifting the observation bands toward higher frequencies can improve the determination of foreground parameters [27]. Thus, we show the results of estimating the tensor-to-scalar ratio r , with the two highest observation frequencies given in Table 1 shifted even higher, in Sect. 4.4. Finally, we show the results with the two-component dust model as a case of mismodeling of the foreground emissions.

We summarize all the estimated r_{out} values for all the setups in Table 2.

4.1 Estimation with synchrotron radiation foreground only

In this section we consider the case where only synchrotron radiation is the foreground source. Because we use a power-law synchrotron radiation model, we need at least three observation bands to fit our parametric model. We first estimate the tensor-to-scalar ratio, r , and synchrotron spectrum index, β_s , with three exact bands $\nu \in (40, 60, 140)$ GHz against input $r_{\text{in}} = 0.01$. We show histograms of the estimated r and β_s as blue boxes in the left and right panels of Fig. 1 from 1000 realizations, respectively. The black vertical line shows the naive average of the β_s of the PySM input in the unmasked sky area.

We next estimate r and β_s with nine bands $\nu \in (40, 50, 60, 68, 78, 89, 100, 119, 140)$ GHz, and show the histograms of the estimated r and β_s in the left and right panels of Fig. 1 in orange,

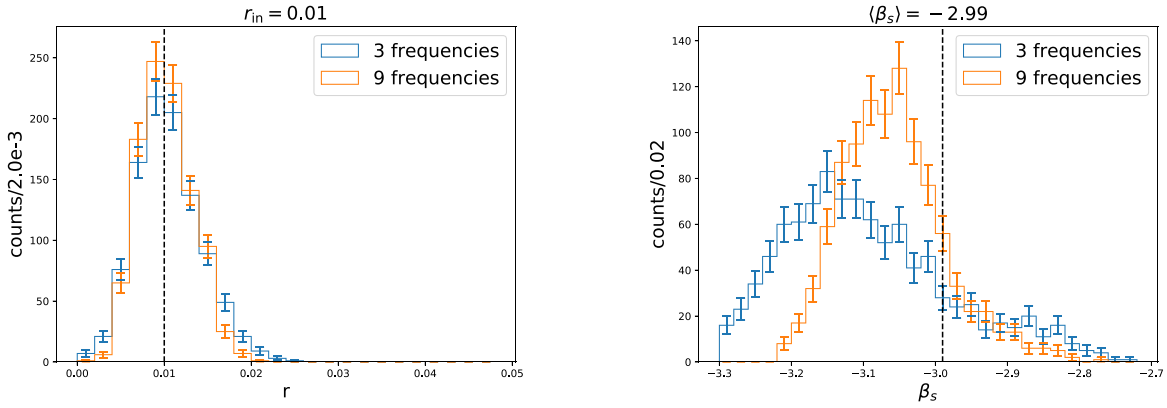


Fig. 1. Estimated r_{out} (left) and β_s (right) from 1000 realizations. Blue histograms show the estimates with three bands and orange histograms show the estimates with nine bands.

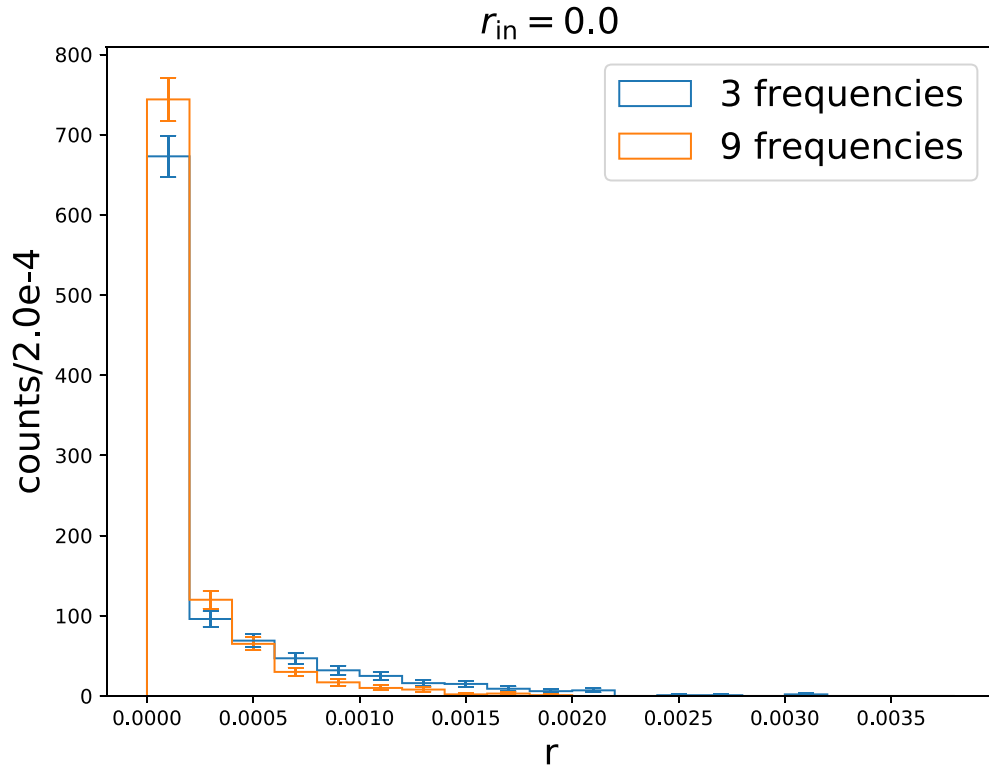


Fig. 2. Estimated r_{out} from 1000 realizations for null r_{in} . Blue histograms show the estimated results with three bands and orange histograms show those with nine bands.

respectively. This is one of the main results of this paper showing that we are free from the constraint of the number of bands that existed in the previous delta-map method. We find that r and β_s are determined better by increasing the number of observation bands.

We show the results of the estimation of r_{out} for the case in which r_{in} is null in Fig. 2. We find that r is constrained to $r_{\text{out}} < 0.12 \times 10^{-2}$ and $r_{\text{out}} < 0.07 \times 10^{-2}$ using three and nine bands, respectively.

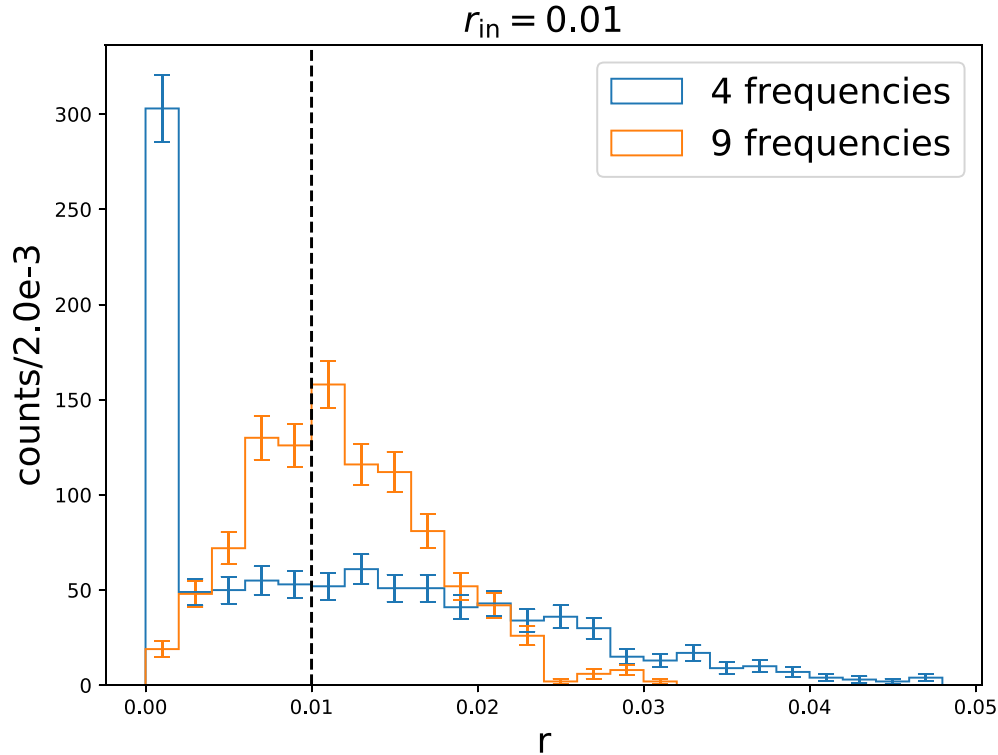


Fig. 3. Histograms of estimated r from 1000 realization samples. Blue box shows the estimated r with four bands. Orange box shows the estimated r with nine bands.

4.2 Estimation with thermal dust emission foreground only

Next, we consider the case where only the thermal dust emission is the foreground source. We choose the one-component modified black body (1MBB) model as the foreground dust model, for which we need at least four bands to fit our parametric model, and we set r_{in} to be 0.01.

We first estimate the tensor-to-scalar ratio parameter r and foreground parameters, T_d and β_d , with the four exact bands, $\nu \in (140, 235, 280, 402)$ GHz. We show the histograms of the estimated r_{out} from 1000 realizations in Fig. 3. We find that the uncertainty on r_{out} is large and the lowest bin is dominant, which reflects the fact that we impose an $r \geq 0$ prior on r so that $\mathbf{S}_0^{\text{CMB}}$ is a positive definite matrix.

We show 2D histograms of the estimates of T_d and β_d in the left panel of Fig. 4. We find that foreground parameters T_d and β_d are not precisely determined with the four bands. The undetermined foreground parameters cause the large error in estimating r .

This can be mitigated by increasing the number of frequency bands. We show the estimated r_{out} with nine bands, $\nu \in (100, 119, 140, 166, 195, 235, 280, 337, 402)$ GHz, in the orange histogram of Fig. 3, and the estimated T_d and β_d with nine bands using the 2D histogram in the right panel of Fig. 4. We find that we can determine foreground parameters well using nine bands, and thereby we can determine r_{out} more precisely.

4.3 Estimation with thermal dust and synchrotron foreground emissions

Finally, we consider the one-component modified black body dust and power-law synchrotron as our foreground models, and estimate the CMB parameter, r , and the foreground parameters, T_d , β_d , and β_s . We use all 15 bands in Table 1. Because we found that the number of bands is

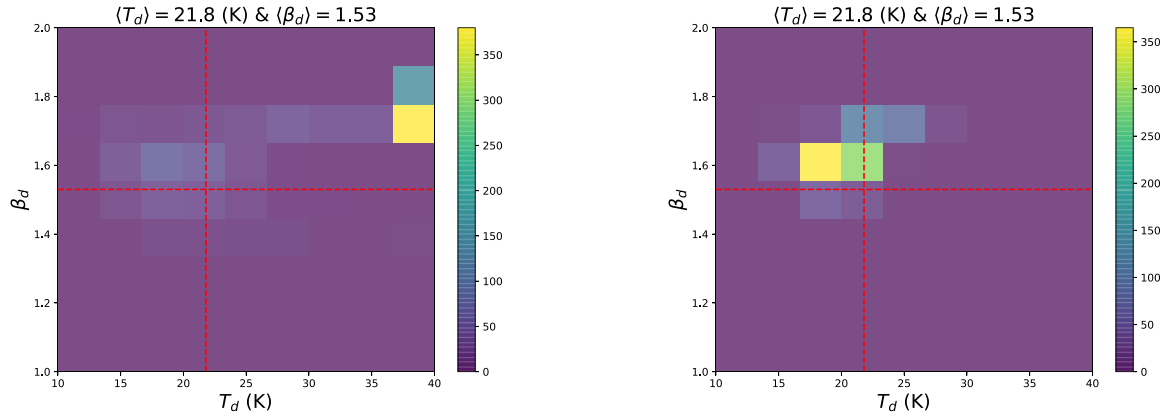


Fig. 4. 2D histograms of estimated dust parameters, β_d vs T_d , with four bands (left) and nine bands (right).

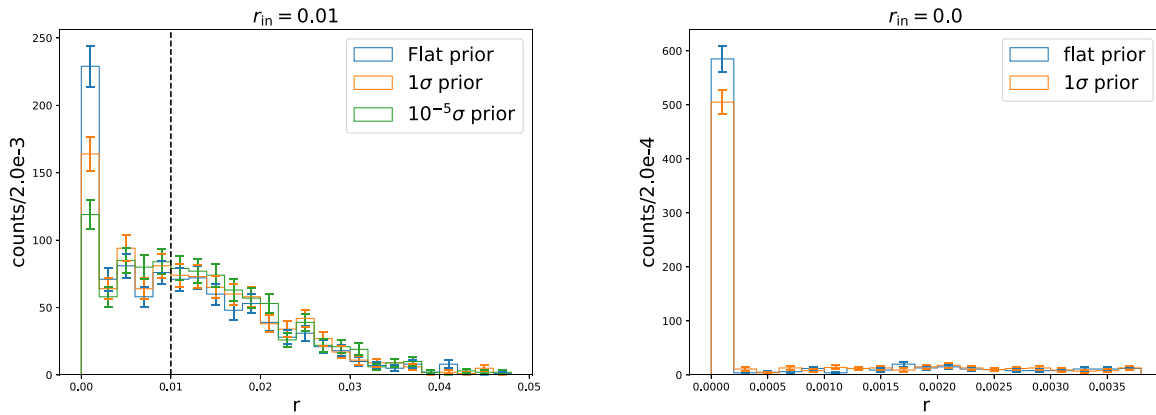


Fig. 5. Histograms of estimated r_{out} from 1000 samples. (Left) r_{out} histograms for input $r_{\text{in}} = 0.01$. Blue, orange, and green histograms show the estimated r_{out} with flat, 1σ , and $10^{-5}\sigma$ priors on T_d , respectively. (Right) r_{out} histograms for input $r_{\text{in}} = 0.0$. Blue and orange histograms show the estimated r_{out} with flat and 1σ priors on T_d , respectively.

not sufficient to determine T_d , we impose some priors, $T_d = 21.8 \pm 1\sigma$ K and $\pm 1.0 \times 10^{-5}\sigma$ K, on T_d , where σ is the standard deviation of the dust temperature measured by Planck [34].

The estimated r_{out} values are shown in the left panel of Fig. 5. We find that stronger constraints on T_d lead to a more precise estimate of r_{out} . However, the 50th percentile is slightly biased to positive, as shown in Table 2. We can see this bias in the estimation of r for null r_{in} , as shown in the right panel of Fig. 5, which shows that r_{out} is slightly biased to positive and the upper bound with the 95% C.L. becomes larger (Table 2).

4.4 Modified high-frequency case

We next see the results with the modified high-frequency model, which increases the frequencies of the high-frequency bands of telescopes.

To see the determination of foreground parameters, we show 2D histograms of β_d and T_d for 15 normal bands and 15 modified bands in the left and right panels of Fig. 6, respectively. We find that we can determine the foreground parameters precisely in the case of the modified high-frequency model.

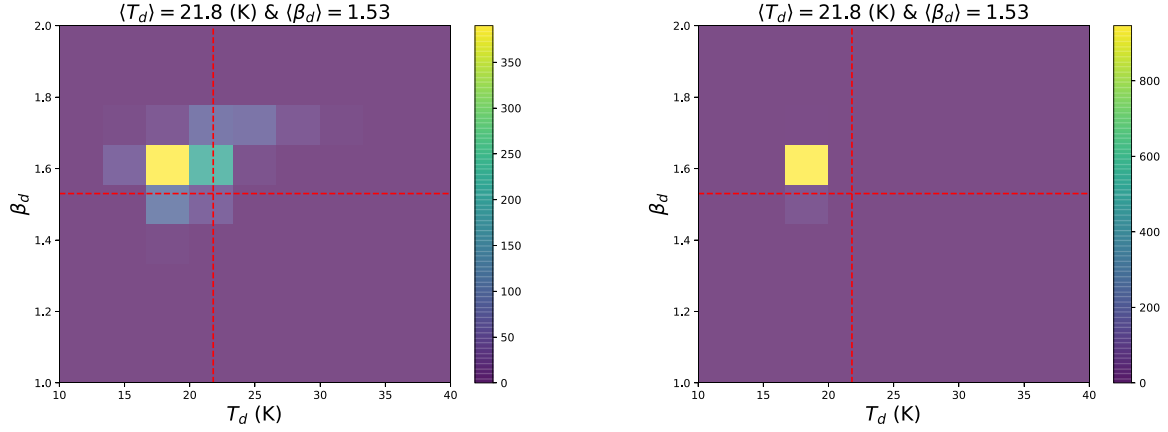


Fig. 6. 2D histograms of T_d vs β_d with 15 normal bands (left) and 15 modified high-frequency bands (right).

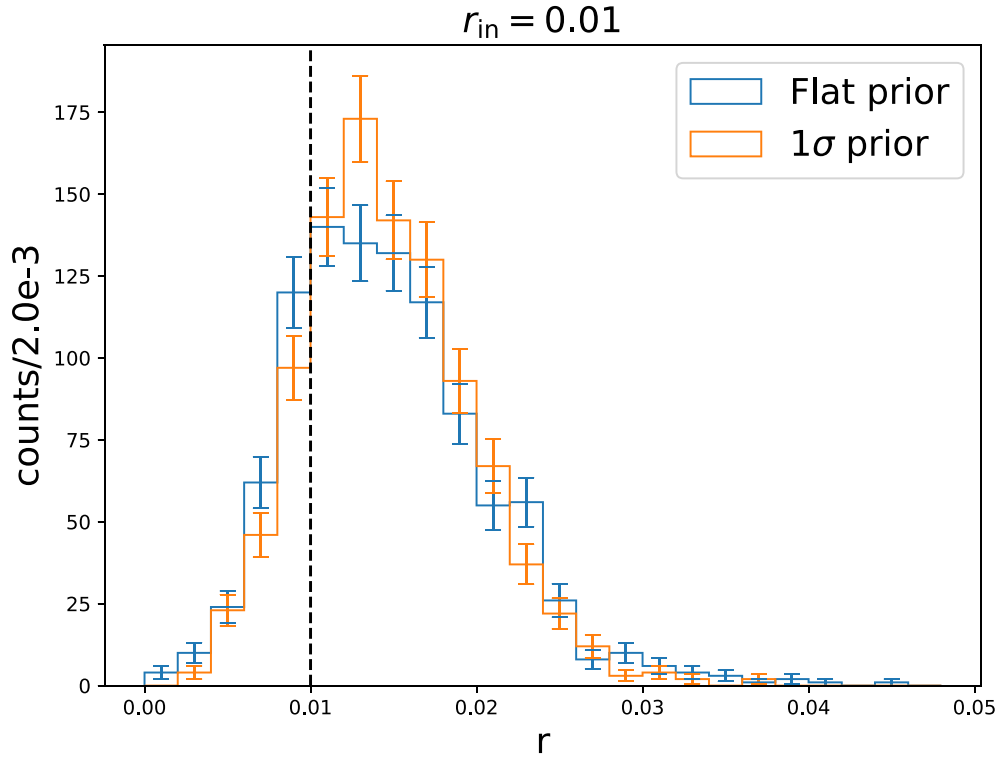


Fig. 7. Histograms of r_{out} from 1000 realizations for a flat prior and a 1σ prior on T_d . The input $r_{\text{in}} = 0.01$ and the dust model is 1MBB. The modified high-frequency model is used.

We show histograms of r_{out} against $r_{\text{in}} = 0.01$ with a flat prior and a 1σ prior in Fig. 7 for the modified high-frequency model. Compared to the result with the fiducial frequency band setting shown in the left panel of Fig. 5, the uncertainties on r become much smaller while the 50th percentile values of r are positively biased as shown in Table 2 if the modified high-frequency model is considered. This tendency was also found in Ref. [21]; the smaller uncertainty comes with a larger systematic bias if one sets the foreground frequency bands far away from the CMB bands.

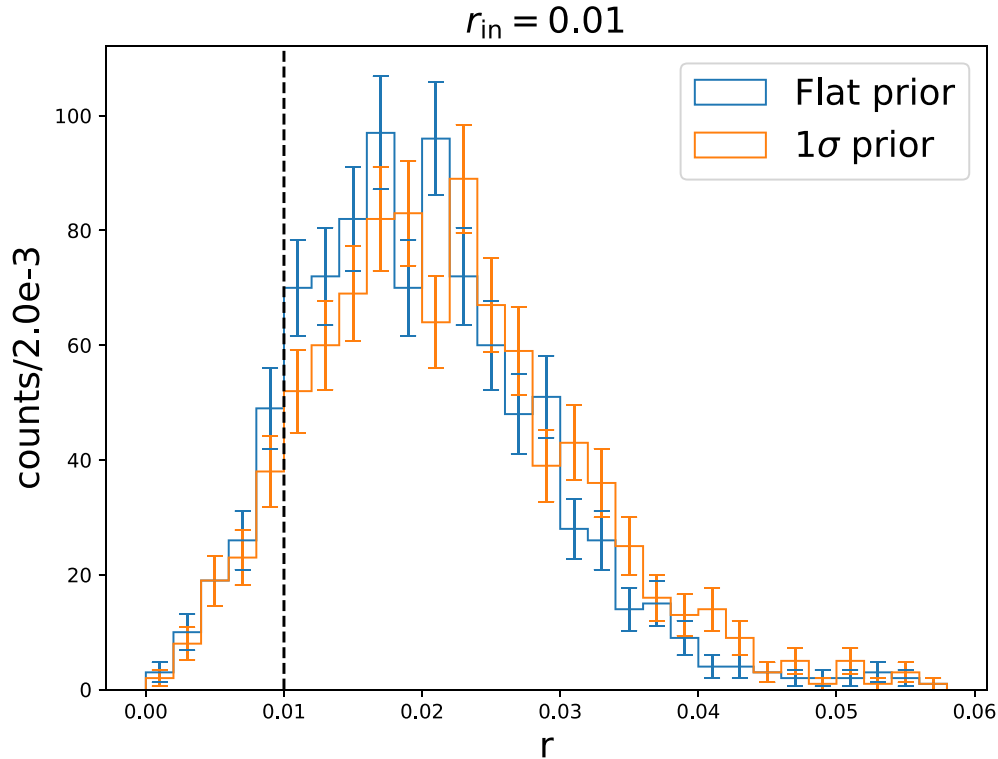


Fig. 8. Histograms of r_{out} from 1000 realizations. The input $r_{\text{in}} = 0.01$ and the dust model is 2MBB.

4.5 Mismodeling

To demonstrate the case in which we assume a wrong foreground model, we use the two-component modified black body (2MBB) model as the input but the 1MBB model to estimate r and the foreground parameters. We show the histograms of the estimated r from 1000 realizations with flat and 1σ priors on T_d in Fig. 8. We find that the estimates of r are biased by ~ 0.01 .

This bias was not found in our previous paper [21]. This could be because the determinative power of the foreground parameters depends on the number of frequency bands. In Ref. [21], r and 1MBB foreground parameters were estimated using only six bands. Because of the small number of bands used in the analysis, the foreground parameters were not well determined and therefore the bias on r was small. To confirm this argument, we estimated r with seven and 10 bands using the extended delta-map method. The results showed that the bias was small for the case with seven bands, and large with 10 bands as well as with 15 bands.

We show 2D histograms of the foreground parameters, T_d and β_d , in the left and right panels of Fig. 9 with flat and 1σ priors on T_d , respectively. Without a prior, the estimated T_d reaches the bound.

5. Summary and discussion

In this paper, we have improved the “delta-map method” [21] by constructing a parametric likelihood in a Bayesian way so that more observation bands can be used. By incorporating the covariance of the foreground emission as “vague”, we have extended the method. Sample codes are available at the extended-deltamap GitHub repository <https://github.com/Yutominami/extended-deltamap>.

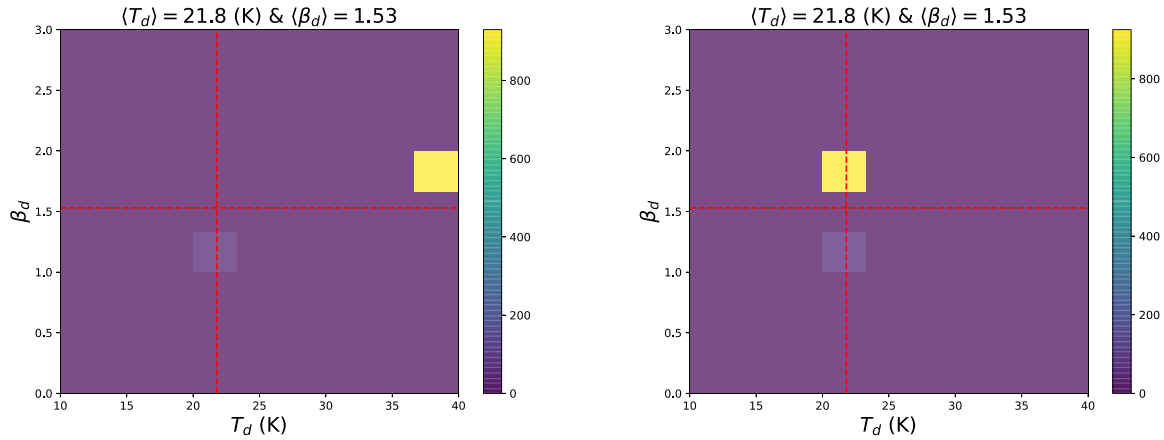


Fig. 9. 2D histograms of estimated T_d and β_d in the case of mismodeling with flat (left) and 1σ (right) priors on T_d .

We have tested the “extended delta-map method” with realistic simulations assuming LiteBIRD-like telescopes. In the case with one foreground component, we find that the extended delta-map method can estimate both the CMB parameter (tensor-to-scalar ratio r) and the foreground parameters, even if we use more frequency bands than the minimum required bands (Sects. 4.1 and 4.2). Moreover, we also found that the parameters of the foreground model are better determined when more observation frequency bands are used, and the tensor-to-scalar ratio r is better estimated accordingly. This is an improvement on and benefit compared to the previous delta-map method.

Next, we apply our method to the model with two foreground components, synchrotron and dust. It turns out that determining dust foreground parameters β_d and T_d becomes difficult in this case. The reason for this probably lies in the interplay between the synchrotron and dust foreground emissions. To determine β_d and T_d simultaneously, one needs to observe the dust spectrum over a wide frequency range. However, in the two-component foreground model, synchrotron radiation dominates the low-frequency side and masks the dust component, effectively reducing the number of observed frequency bands that can be used to estimate the dust foreground parameters. To aid the determination of the foreground parameter, therefore, we have imposed a prior on dust temperature based on the Planck results. In this case, the error in the estimate of r becomes small, while we have a small positive bias in the estimate. This small bias may indicate that the mean value of T_d estimated in the delta-map method is not necessarily the same as T_d derived by maps with higher angular resolutions, e.g., Planck.

A positive bias is also found in the case using a modified band configuration with higher frequencies. If we increase the frequencies of the two highest observation bands, we can determine the foreground parameters better, as we discussed above, while the estimated CMB parameter is biased. This tendency was already found in Ref. [21]. The reason for this is probably the breakdown of the perturbative treatment of the foreground parameters in our method. Although it is easier to remove foreground radiation when the frequency bands are closer to each other, internal template methods, including ours, have the disadvantage of removing the CMB at the same time, resulting in relatively larger noise. Thus, we are faced with the familiar dilemma of systematic and statistical errors that are inherent in statistical analysis.

Lastly, we test the case with mismodeling by fitting the two-component modified black body model with the one-component modified black body model. We find that the estimated CMB parameter is biased.

In this paper, we have not considered detailed characteristics of telescopes or detectors, e.g., the bandpass average discussed in Ref. [21]. We can include the detailed characteristics in the transfer matrices. The basic idea of the delta-map method is to consider the spatial variation of foreground signal parameters perturbatively. Though we have only considered the first-order expansions of spatial variations of foreground parameters, we can consider higher-order expansions, which may improve the estimates of the foreground and CMB parameters. We leave this study for future work.

Acknowledgements

We thank E. Komatsu, A. Nishizawa, Y. Chinone, and S. Takakura for useful discussions. This work is supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI, Grants Nos. JP18K03616 (K.I.) and JP20K14497 (Y.M.), JSPS core-to-core program number JPJSCCA20200003 (K.I.), JST AIP Acceleration Research Grant JP20317829 (K.I.), and JST FOREST Program JPMJFR20352935 (K.I.).

Appendix A. Derivation of likelihood

We use the methodology described in Ref. [30] to deal with the “vague” foreground covariance matrix.

When we marginalize $P(\vec{s}_f, \mathbf{S}^f)$ in Eq. (21), we have

$$-2 \ln P(\vec{p}^I, \vec{s}_f, \mathbf{S}|\vec{m}) = (\vec{m} - \tilde{\mathbf{D}}\vec{s}_f)^\top ((\mathbf{S}^{\text{CMB}} + \mathbf{N}) + \tilde{\mathbf{D}}\mathbf{S}^f\tilde{\mathbf{D}}^\top)^{-1} (\vec{m} - \tilde{\mathbf{D}}\vec{s}_f) + n \ln 2\pi + \ln |(\mathbf{S}^{\text{CMB}} + \mathbf{N}) + \tilde{\mathbf{D}}\mathbf{S}^f\tilde{\mathbf{D}}^\top|. \quad (\text{A1})$$

When we assume that the mean of the foreground signal is zero, $\vec{s}_f = \vec{0}$, we have

$$-2 \ln P(\vec{p}^I, \mathbf{S}|\vec{m}) = \vec{m}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \vec{m} - \vec{m}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \tilde{\mathbf{D}} \left[\mathbf{S}^{f^{-1}} + \tilde{\mathbf{D}}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \tilde{\mathbf{D}} \right]^{-1} \tilde{\mathbf{D}}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \vec{m} + \ln |\mathbf{S}^{\text{CMB}} + \mathbf{N}| + \ln |\mathbf{S}^f| + \ln |\tilde{\mathbf{D}}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \tilde{\mathbf{D}}| + \text{const.}, \quad (\text{A2})$$

where we have used the *matrix inversion lemma* (B1), (B2).

When we are ignorant about the covariance of the foreground signals, we take the limit where $\mathbf{S}^{f^{-1}} \rightarrow \mathbf{O}$ and have

$$-2 \ln P(\vec{p}^I, \mathbf{S}_0^{\text{CMB}}|\vec{m}) = \vec{m}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \vec{m} - \vec{m}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \tilde{\mathbf{D}} \left[\tilde{\mathbf{D}}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \tilde{\mathbf{D}} \right]^{-1} \tilde{\mathbf{D}}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \vec{m} + \ln |\mathbf{S}^{\text{CMB}} + \mathbf{N}| + \ln |\tilde{\mathbf{D}}^\top (\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \tilde{\mathbf{D}}| + \text{const.}, \quad (\text{A3})$$

where we discard the terms $\ln |\mathbf{S}^f|$ following Ref. [30].

This is similar to the likelihood function (Eq. (61) of Ref. [21]) except for the additional term, $\ln |\tilde{\mathbf{D}}(\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} \tilde{\mathbf{D}}|$. We will revisit this difference in Appendix C and show that this term is necessary to reproduce the results of Ref. [21].

Table A1. Summaries of matrices and the vector defined in this section.

Symbols	Definition
\mathbf{A}	$(\mathbf{S}_0^{\text{CMB}^{-1}} + \sum_{j=1}^{N+1} \mathbf{N}_{\nu_j}^{-1})$
\mathbf{H}	$\mathbf{N}^{-1} \tilde{\mathbf{D}} [\tilde{\mathbf{D}}^T \mathbf{N}^{-1} \tilde{\mathbf{D}}]^{-1} \tilde{\mathbf{D}}^T \mathbf{N}^{-1}$
\mathbf{B}	$\mathbf{A} - \mathbf{D}^{\text{CMB}^T} \mathbf{H} \mathbf{D}^{\text{CMB}}$
\vec{M}	$\vec{m} - \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m}$

Next, we apply the *matrix inversion lemma* (B1), (B2) to $(\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1}$ and $\ln |(\mathbf{S}^{\text{CMB}} + \mathbf{N})|$, to reduce the calculation cost of the large covariance matrix. Using $(\mathbf{S}^{\text{CMB}} + \mathbf{N})$ expressed as $(\mathbf{D}^{\text{CMB}} \mathbf{S}_0^{\text{CMB}} \mathbf{D}^{\text{CMB}^T} + \mathbf{N})$, we have

$$(\mathbf{S}^{\text{CMB}} + \mathbf{N})^{-1} = \mathbf{N}^{-1} - \mathbf{N}^{-1} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \quad (\text{A4})$$

$$\ln |\mathbf{S}^{\text{CMB}} + \mathbf{N}| = \ln |\mathbf{S}_0^{\text{CMB}}| + \ln |\mathbf{N}| + \ln |\mathbf{A}|, \quad (\text{A5})$$

where we define $\mathbf{A} = (\mathbf{S}_0^{\text{CMB}^{-1}} + \sum_{j=1}^{N+1} \mathbf{N}_{\nu_j}^{-1})$. When we substitute Eqs. (A4) and (A5) into Eq. (A3), we have

$$\begin{aligned} -2 \ln P(\vec{p}^I, \mathbf{S}_0^{\text{CMB}} | \vec{m}) &= \vec{m}^T \mathbf{N}^{-1} \vec{m} - \vec{m}^T \mathbf{N}^{-1} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} - \vec{M}^T \mathbf{H} \vec{M} - \vec{M}^T \mathbf{H} \mathbf{B}^{-1} \mathbf{H} \vec{M} \\ &+ \ln |\mathbf{S}_0^{\text{CMB}}| + \ln |\mathbf{N}| + \ln |\tilde{\mathbf{D}}^T \mathbf{N}^{-1} \tilde{\mathbf{D}}| + \ln |\mathbf{B}|, \end{aligned} \quad (\text{A6})$$

where we define

$$\vec{M} = \vec{m} - \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} \quad (\text{A7})$$

$$\mathbf{H} = \mathbf{N}^{-1} \tilde{\mathbf{D}} [\tilde{\mathbf{D}}^T \mathbf{N}^{-1} \tilde{\mathbf{D}}]^{-1} \tilde{\mathbf{D}}^T \mathbf{N}^{-1} \quad (\text{A8})$$

$$\mathbf{B} = \mathbf{A} - \mathbf{D}^{\text{CMB}^T} \mathbf{H} \mathbf{D}^{\text{CMB}}. \quad (\text{A9})$$

Here we try to keep the symmetry to write down each term because we find that the estimated parameters are biased when we calculate the log-likelihood value using asymmetric terms. Finally, we summarize the newly defined matrices and the vector in Table A1.

Appendix B. Matrix inversion lemma

In the derivation of the likelihood function, we use the Woodbury, Sherman, and Morrison formula, the so-called “matrix inversion lemma”,

$$(\mathbf{Z} + \mathbf{U} \mathbf{W} \mathbf{V}^T)^{-1} = \mathbf{Z}^{-1} - \mathbf{Z}^{-1} \mathbf{U} (\mathbf{W}^{-1} + \mathbf{V}^T \mathbf{Z}^{-1} \mathbf{U})^{-1} \mathbf{V} \mathbf{Z}^{-1}, \quad (\text{B1})$$

where \mathbf{Z} and \mathbf{W} are invertible matrices and \mathbf{U} and \mathbf{V} are matrices with corresponding dimensions.

For log-determinants, a similar equation exists:

$$\ln |\mathbf{Z} + \mathbf{U} \mathbf{W} \mathbf{V}^T| = \ln |\mathbf{Z}| + \ln |\mathbf{W}| + \ln |\mathbf{W}^{-1} + \mathbf{V}^T \mathbf{Z}^{-1} \mathbf{U}|. \quad (\text{B2})$$

Appendix C. Comparison of likelihood

In this section, we show that the estimated parameters with Eq. (22) are equivalent to those estimated from Eq. (C4) in the case in which we assume one foreground component and one parameter, β , where the number of frequencies is $N_\nu = (N_d (= 1) + 1) + 1 = 3$.

The cleaned CMB map (14) can also be expressed as $[\mathbf{D}^{-1}\vec{m}]_{\text{CMB}}$, where the subscript ‘‘CMB’’ means that we take the elements related to the CMB, as described above Eq. (50) in Ref. [21].

First, let us calculate this cleaned CMB map. The elements of the \mathbf{D} matrix are

$$\begin{aligned}\mathbf{D} &= [\mathbf{D}^{\text{CMB}} \tilde{\mathbf{D}}] \\ &= \begin{bmatrix} \mathbf{I} & g_{v_1} D_{v_1} \mathbf{I} & g_{v_1} D_{v_1, \beta} \mathbf{I} \\ \mathbf{I} & g_{v_2} D_{v_2} \mathbf{I} & g_{v_2} D_{v_2, \beta} \mathbf{I} \\ \mathbf{I} & g_{v_3} D_{v_3} \mathbf{I} & g_{v_3} D_{v_3, \beta} \mathbf{I} \end{bmatrix}.\end{aligned}\quad (\text{C1})$$

The inverse matrix of \mathbf{D} is

$$\mathbf{D}^{-1} = \frac{-1}{|\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}}|} \begin{bmatrix} -\xi_{23} \mathbf{I} & -\xi_{31} \mathbf{I} & -\xi_{12} \mathbf{I} \\ (g_{v_2} D_{v_2, \beta} - g_{v_3} D_{v_3}) \mathbf{I} & (-g_{v_1} D_{v_1, \beta} + g_{v_3} D_{v_3, \beta}) \mathbf{I} & (g_{v_1} D_{v_1, \beta} - g_{v_2} D_{v_2, \beta}) \mathbf{I} \\ (-g_{v_2} D_{v_2} + g_{v_3} D_{v_3}) \mathbf{I} & (g_{v_1} D_{v_1} - g_{v_3} D_{v_3}) \mathbf{I} & (-g_{v_1} D_{v_1} + g_{v_2} D_{v_2}) \mathbf{I} \end{bmatrix}, \quad (\text{C2})$$

where $\boldsymbol{\Xi} = (\xi_{23} \mathbf{I} \ \xi_{31} \mathbf{I} \ \xi_{12} \mathbf{I})$. Then the cleaned CMB map can be expressed as

$$\begin{aligned}\text{CMB}^{\text{ML}}(\hat{n}) &= [\mathbf{D}^{-1}\vec{m}]_{\text{CMB}} \\ &= (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \boldsymbol{\Xi} \vec{m}.\end{aligned}\quad (\text{C3})$$

The corresponding covariance matrix (17) is calculated as

$$\mathbf{C} = \mathbf{S}_0^{\text{CMB}} + |\mathbf{K}| (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \quad (\text{C4})$$

where $\mathbf{K} = \boldsymbol{\Xi} \mathbf{N} \boldsymbol{\Xi}^{\text{T}}$.

Therefore, we can write the likelihood function used in Ref. [21] as

$$-2 \ln \mathcal{L} = ((\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \boldsymbol{\Xi} \vec{m})^{\text{T}} (\mathbf{S}_0^{\text{CMB}} + |\mathbf{K}| (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1})^{-1} (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \boldsymbol{\Xi} \vec{m}. \quad (\text{C5})$$

We will show that Eq. (22) is equal to this equation.

First, we calculate $\tilde{\mathbf{D}}^{\text{T}} \mathbf{N}^{-1} \tilde{\mathbf{D}}$ matrix and its determinant as

$$\tilde{\mathbf{D}}^{\text{T}} \mathbf{N}^{-1} \tilde{\mathbf{D}} = \begin{bmatrix} \sum_i N_i^{-1} g_{v_i}^2 D_{v_i}^2 & \sum_i N_i^{-1} g_{v_i} D_{v_i} D_{v_i, \beta} \\ \sum_i N_i^{-1} g_{v_i} D_{v_i} D_{v_i, \beta} & \sum_i N_i^{-1} g_{v_i}^2 D_{v_i, \beta}^2 \end{bmatrix} \quad (\text{C6})$$

$$|\tilde{\mathbf{D}}^{\text{T}} \mathbf{N}^{-1} \tilde{\mathbf{D}}| = |\mathbf{K}| \prod_i |N_i^{-1}|. \quad (\text{C7})$$

Using them, we have

$$\begin{aligned}\mathbf{H} &= \frac{1}{|\mathbf{K}|} \begin{bmatrix} N_1^{-1} (N_3 \xi_{12}^2 + N_2 \xi_{31}^2) \mathbf{I} & -\xi_{31} \xi_{23} \mathbf{I} & -\xi_{12} \xi_{23} \mathbf{I} \\ -\xi_{31} \xi_{23} \mathbf{I} & N_1^{-2} (N_1 \xi_{23}^2 + N_3 \xi_{12}^2) \mathbf{I} & -\xi_{12} \xi_{31} \mathbf{I} \\ -\xi_{12} \xi_{23} \mathbf{I} & -\xi_{12} \xi_{31} \mathbf{I} & N_1^{-3} (N_2 \xi_{31}^2 + N_1 \xi_{23}^2) \mathbf{I} \end{bmatrix} \\ &= \mathbf{N}^{-1} - \frac{\boldsymbol{\Xi}^{\text{T}} \boldsymbol{\Xi}}{|\mathbf{K}|}.\end{aligned}\quad (\text{C8})$$

Then we have

$$\begin{aligned}\mathbf{B} &= \mathbf{A} - \mathbf{D}^{\text{CMB}} \mathbf{T} \left(\mathbf{N}^{-1} - \frac{\boldsymbol{\Xi}^{\text{T}} \boldsymbol{\Xi}}{|\mathbf{K}|} \right) \mathbf{D}^{\text{CMB}} \\ &= \mathbf{S}_0^{\text{CMB}^{-1}} + \frac{\mathbf{D}^{\text{CMB}} \mathbf{T} \boldsymbol{\Xi}^{\text{T}} \boldsymbol{\Xi} \mathbf{D}^{\text{CMB}}}{|\mathbf{K}|},\end{aligned}\quad (\text{C9})$$

and its inverse as

$$\begin{aligned}
\mathbf{B}^{-1} &= \left(\mathbf{S}_0^{\text{CMB}^{-1}} + \frac{\mathbf{D}^{\text{CMB}^T} \boldsymbol{\Xi}^T \boldsymbol{\Xi} \mathbf{D}^{\text{CMB}}}{|\mathbf{K}|} \right)^{-1} \\
&= |\mathbf{K}| (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \mathbf{S}_0^{\text{CMB}} (\mathbf{S}_0^{\text{CMB}} + |\mathbf{K}| (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1})^{-1} \\
&= |\mathbf{K}| (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \mathbf{S}_0^{\text{CMB}} \mathbf{C}^{-1} \\
&= (\mathbf{A}^{-1} + \mathbf{B}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1}) \\
&= (\mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{B}^{-1}). \tag{C10}
\end{aligned}$$

Therefore, we can rewrite Eq. (22) as

$$\begin{aligned}
-2 \ln P(\vec{p}^I, \mathbf{S}_0^{\text{CMB}} | \vec{m}) &= \vec{m}^T \mathbf{N}^{-1} \vec{m} - \vec{m}^T \mathbf{N}^{-1} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} - \vec{M}^T \mathbf{H} \vec{M} \\
&\quad - \vec{M} \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{B}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{H} \vec{M} \\
&\quad + \ln |\mathbf{S}_0^{\text{CMB}}| + \ln |\mathbf{N}| + \ln |\tilde{\mathbf{D}}^T \mathbf{N}^{-1} \tilde{\mathbf{D}}| + \ln |\mathbf{B}| \\
&= \vec{m}^T \mathbf{N}^{-1} \vec{m} - \vec{m}^T \mathbf{N}^{-1} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} - \vec{m}^T \mathbf{H} \vec{m} \\
&\quad + 2 \vec{m}^T \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} \\
&\quad - \vec{m}^T \mathbf{N}^{-1} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} \\
&\quad - \vec{m}^T \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{B}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{H} \vec{m} + 2 \vec{m} \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{B}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{H} \mathbf{D}^{\text{CMB}} \\
&\quad \times \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} - \vec{m} \mathbf{N}^{-1} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{B}^{-1} \mathbf{D}^{\text{CMB}^T} \\
&\quad \times \mathbf{H} \mathbf{D}^{\text{CMB}} \mathbf{A}^{-1} \mathbf{D}^{\text{CMB}^T} \mathbf{N}^{-1} \vec{m} + 2 \ln |\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}}| + \ln |\mathbf{C}| \\
&= \vec{m}^T (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \frac{\mathbf{I} - \mathbf{S}_0^{\text{CMB}} \mathbf{C}^{-1}}{|\mathbf{K}|} (\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}})^{-1} \vec{m} + 2 \ln |\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}}| + \ln |\mathbf{C}| \\
&= \text{CMB}^{\text{ML}}(\hat{n})^T \mathbf{C}^{-1} \text{CMB}^{\text{ML}}(\hat{n}) + 2 \ln |\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}}| + \ln |\mathbf{C}|. \tag{C11}
\end{aligned}$$

This agrees with Eq. (C4) except for the term $\ln |\boldsymbol{\Xi} \mathbf{D}^{\text{CMB}}|$. This does not change the estimated parameters, since, in Ref. [21], the foreground parameters are determined only with the chi-squared term and the CMB parameters are determined with the total likelihood function by fixing the foreground parameter.

References

- [1] E. Komatsu et al., Prog. Theor. Exp. Phys. **2014**, 06B102 (2014) [arXiv:1404.5415 [astro-ph.CO]] [Search inSPIRE].
- [2] Planck Collaboration VI, Astron. Astrophys. **641**, A6 (2020) [arXiv:1807.06209 [astro-ph.CO]] [Search inSPIRE].
- [3] The POLARBEAR Collaboration, Astrophys. J. **931**, 101 (2022) [arXiv:2203.02495 [astro-ph.CO]] [Search inSPIRE].
- [4] The POLARBEAR Collaboration, Astrophys. J. **904**, 65 (2020) [arXiv:2005.06168 [astro-ph.CO]] [Search inSPIRE].
- [5] S. Aiola et al., J. Cosmol. Astropart. Phys. **2012**, 047 (2020) [arXiv:2007.07288 [astro-ph.CO]] [Search inSPIRE].
- [6] SPTpol Collaboration, Phys. Rev. D **101**, 122003 (2020) [arXiv:1910.05748 [astro-ph.CO]] [Search inSPIRE].

- [7] SPT-3G Collaboration, Phys. Rev. D **104**, 022003 (2021) [[arXiv:2101.01684 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [8] BICEP/Keck Collaboration, Phys. Rev. Lett. **127**, 151301 (2021) [[arXiv:2110.00483 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [9] Spider Collaboration, Astrophys. J. **927** 174 2022 [arXiv:2103.13334 \[astro-ph.CO\]](#) [[Search inSPIRE](#)].
- [10] L. P. Grishchuk, Zh. Eksp. Teor. Fiz. **67**, 825 (1974).
- [11] A. A. Starobinsky, JETP Lett. **30**, 682 (1979).
- [12] U. Seljak and M. Zaldarriaga, Phys. Rev. Lett. **78**, 2054 (1997) [[arXiv:astro-ph/9609169](#)] [[Search inSPIRE](#)].
- [13] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Phys. Rev. Lett. **78**, 2058 (1997) [[arXiv:astro-ph/9609132](#)] [[Search inSPIRE](#)].
- [14] M. Kamionkowski and E. D. Kovetz, Ann. Rev. Astron. Astrophys. **54**, 227 (2016) [[arXiv:1510.06042 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [15] K. Ichiki, Prog. Theor. Exp. Phys. **2014**, 06B109 (2014).
- [16] H. K. Eriksen, J. B. Jewell, C. Dickinson, A. J. Banday, K. M. Gorski, and C. R. Lawrence, Astrophys. J. **676**, 10 (2008) [[arXiv:0709.1058 \[astro-ph\]](#)] [[Search inSPIRE](#)].
- [17] R. Fernández-Cobos, P. Vielva, R. B. Barreiro, and E. Martínez-González, Mon. Not. R. Astron. Soc. **420**, 2162 (2012) [[arXiv:1106.2016 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [18] J. Delabrouille, J. F. Cardoso, and G. Patanchon, Mon. Not. R. Astron. Soc. **346**, 1089 (2003) [[arXiv:astro-ph/0211504](#)] [[Search inSPIRE](#)].
- [19] J. Delabrouille, J. F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, and F. Guilloux, Astron. Astrophys. **493**, 835 (2009) [[arXiv:0807.0773 \[astro-ph\]](#)] [[Search inSPIRE](#)].
- [20] M. Remazeilles, J. Delabrouille, and J.-F. Cardoso, Mon. Not. R. Astron. Soc. **418**, 467 (2011) [[arXiv:1103.1166 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [21] K. Ichiki, H. Kanai, N. Katayama, and E. Komatsu, Prog. Theor. Exp. Phys. **2019**, 033E01 (2019) [[arXiv:1811.03886 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [22] J. Chluba, J. C. Hill, and M. H. Abitbol, Mon. Not. R. Astron. Soc. **472**, 1195 (2017) [[arXiv:1701.00274 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [23] A. Mangilli, J. Aumont, A. Rotti, F. Boulanger, J. Chluba, T. Ghosh, and L. Montier, Astron. Astrophys. **647**, A52 (2021) [[arXiv:1912.09567 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [24] L. Vacher, J. Aumont, L. Montier, S. Azzoni, F. Boulanger, and M. Remazeilles, Astron. Astrophys. **660**, A111 (2022) [[arXiv:2111.07742 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [25] L. Vacher, J. Chluba, J. Aumont, A. Rotti, and L. Montier, Astron. Astrophys. **669**, A5 (2023) [[arXiv:2205.01049 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [26] M. Hazumi et al., J. Low Temp. Phys. **194**, 443 (2019).
- [27] LiteBIRD Collaboration, Prog. Theor. Exp. Phys. **2022** ptac150 2022 [arXiv:2202.02773 \[astro-ph.IM\]](#) [[Search inSPIRE](#)].
- [28] N. Katayama and E. Komatsu, Astrophys. J. **737**, 78 (2011) [[arXiv:1101.5210 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [29] R. Stompor, S. M. Leach, F. Stivoli, and C. Baccigalupi, Mon. Not. R. Astron. Soc. **392**, 216 (2009) [[arXiv:0804.2645 \[astro-ph\]](#)] [[Search inSPIRE](#)].
- [30] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning (MIT Press, Cambridge, MA, 2006).
- [31] B. Thorne, J. Dunkley, D. Alonso, and S. Naess, Mon. Not. R. Astron. Soc. **469**, 2821 (2017) [[arXiv:1608.02841 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [32] C. L. Bennett et al., Astrophys. J. Suppl. Ser. **208**, 20 (2013).
- [33] M. A. Miville-Deschenes, N. Ysard, A. Lavabre, N. Ponthieu, J. F. Macias-Perez, J. Aumont, and J. P. Bernard, Astron. Astrophys. **490**, 1093 (2008) [[arXiv:0802.3345 \[astro-ph\]](#)] [[Search inSPIRE](#)].
- [34] Planck Collaboration, Astron. Astrophys. **594**, A10 (2016) [[arXiv:1502.01588 \[astro-ph.CO\]](#)] [[Search inSPIRE](#)].
- [35] A. M. Meisner and D. P. Finkbeiner, Astrophys. J. **798**, 88 (2015) [[arXiv:1410.7523 \[astro-ph.GA\]](#)] [[Search inSPIRE](#)].

- [36] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, *Astrophys. J.* **622**, 759 (2005) [[arXiv:astro-ph/0409513](https://arxiv.org/abs/astro-ph/0409513)] [[Search inSPIRE](#)].
- [37] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000) [[arXiv: astro-ph/9911177](https://arxiv.org/abs/astro-ph/9911177)] [[Search inSPIRE](#)].
- [38] H. K. Eriksen et al., *Astrophys. J.* **656**, 641 (2007) [[arXiv: astro-ph/0606088](https://arxiv.org/abs/astro-ph/0606088)] [[Search inSPIRE](#)].