Lawrence Berkeley National Laboratory
LBL Publications

Title
Foreground removal with ILC methods for AliCPT-1

Permalink

Journal

ISSN

Authors
Dou, Jiazheng
Ghosh, Shamik
Santos, Larissa
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Foreground removal with ILC methods for AliCPT-1

Jiazheng Dou,[a,b] **Shamik Ghosh,**[c] **Larissa Santos,**[d] **Wen Zhao**[a,b]

[a]CAS Key Laboratory for Researches in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui 230026, P.R.China

[b]School of Astronomy and Space Sciences, University of Science and Technology of China, Hefei 230026, P.R.China

[c]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, U.S.A.

[d]Center for Gravitation and Cosmology, Yangzhou University, Yangzhou 224009, P.R.China

E-mail: doujzh@mail.ustc.edu.cn, shamik-ghosh@outlook.com, larissa@yzu.edu.cn, wzhao7@ustc.edu.cn

Abstract. One of the main goals of most future CMB experiments is the precise measurement of CMB B-mode polarization, whose major obstacle is the Galactic foregrounds. In this paper, we evaluate the foreground cleaning performance of the variants of the ILC method on partial sky B-modes and analyze the main sources of biases on the BB power spectrum. Specially, we compare the NILC, the cILC (in three domains) and the cMILC methods for AliCPT-1 simulations. We find that the cILC methods implemented in harmonic space and needlet space are both competent to clean different models of foregrounds, which bias the tensor-to-scalar ratio about 0.008 at maximum, and constrain the tensor-to-scalar ratio to $r < 0.043\,(95\%\mathrm{CL})$ for the AliCPT-1 configuration. We also note that the deviation of the estimated noise bias from the actual one for ILC, dubbed the noise bias error (NBE) in this paper, might make significant effects on the power spectrum for a small footprint and low signal-to-noise ratio CMB experiment. We finally obtain its relation with respect to the noise residual which fits well with the simulated results.

Contents

1 Introduction

The Cosmic Microwave Background (CMB) anisotropies provide direct information about the origin and history of our universe. Observations of the CMB radiation in the last two decades (e.g., COBE [1], WMAP [2] and Planck [3]) have been in excellent agreement with

the ΛCDM model with very tight constraints on its parameters. However, despite ΛCDM success in explaining the observations, it has some well-known problems: horizon, flatness and magnetic monopole. To alleviate them, the universe is supposed to undergo an inflationary epoch with a nearly exponential expansion. One of the most important predictions of cosmic inflation is the existence of primordial gravitational waves (PGW) generated from vacuum fluctuations in the early Universe (about 10^{-35}s after the Big Bang). These PGWs would leave a faint imprint on the CMB polarization, known as primordial B-modes [4]. The amplitude of the primordial B-mode power spectrum at degree angular scales is related to the amount of the PGWs, which is parameterized by the tensor-to-scalar ratio r. Thus, measuring r would enable us to distinguish different classes of early Universe models (e.g., [5, 6]).

Attracted by the 'smoking-gun' evidence of inflation, the measurement of the CMB primordial B-modes has been the main scientific goal of current CMB experiments, such as POLARBEAR [7], ACTpol [8], SPTpol [9] and BICEP/Keck [10]. However, measuring the primordial CMB B-mode can be challenging due to the tiny signal and the various contaminants, such as the Galactic emissions, the instrumental noise and the CMB gravitational lensing, which converts E-modes into B-modes [11]. The current tightest upper limit of the tensor-to-scalar ratio, $r < 0.032\,(95\%\mathrm{CL})$, is given by a combination of Planck PR4 and BICEP2/Keck Array 2018 (BK18), BAO and CMB lensing data [12]. Many on-going and future ground-based telescopes, as well as, space-based CMB experiments aim to reach higher sensitivity in order to be able to measure the primordial B-modes, such as the Simons Observatory [13], QUBIC [14], AliCPT [15], LiteBIRD [16], CMB-S4 [17].

Contamination from polarized foreground signals, especially the Galactic thermal dust and the synchrotron emission, must then be disentangled from the B-mode signal by the so-called 'component separation' methods. These methods are usually divided into blind, parametric and template removal techniques. Blind methods, such as ILC (Internal Linear Combination) [18, 19], FastICA (Fast Independent Component Analysis) [20] and SMICA (Spectral Matching Independent Component Analysis) [21, 22], usually rely on the statistical independence of components of various physical origins for their separation in multi-frequency observations, without assumptions of foregrounds. Parametric methods, such as Commander [23] and FG-Buster [24, 25], model the foregrounds with spectral laws where parameters are fitted to the observed data. The parametric methods build an end-to-end Bayesian framework to estimate the parameters, but their effectiveness relies on the model. Template removal methods like SEVEM (Spectral Estimation Via Expectation Maximisation) [26] construct a set of foreground templates from the data.

The ILC method and its variants, such as the Needlet ILC (NILC) [27], play an important role in the CMB analysis of future experiments given our poor knowledge of the polarized foregrounds' properties. One of the caveats on the partial-sky analysis of B-modes is what is known as EB leakage [28, 29]: the decomposition from partial-sky QU maps into E- and B-modes will give rise to a considerable leakage from E-modes to B-modes. Directly applying ILC to Q and U maps separately is conceptually problematic since subtracting a constant contribution from Q or U is equivalent of introducing a pattern on the E and B modes. One should either linearly combine the QU maps to obtain the cleaned QU signals, e.g. PILC (Polarization ILC) [30], or try to reconstruct the pure B-modes from QU maps before applying the ILC methodology. The PILC (constructing a complex-weight linear combination of $Q \pm iU$ such that the quantity $\langle P^2 = Q^2 + U^2 \rangle$ of the final map is minimized) makes a tradeoff between minimizing the variance of E-modes and B-modes. Thereby, it may not be effective as a cleaning method to be applied in the CMB B-modes map recovery. Therefore,

we adopt a novel template cleaning method [31] to extract the pure B-modes maps, and then apply several component separation methods on the multi-frequency B maps to clean the foregrounds.

Some authors already forecast on the performance of the standard ILC and NILC on B-modes mock data for some future experiments [32–39]. However, considering smaller sky fractions and lower sensitivities, such blind methods may give rise to a foreground-induced bias of the order $r \sim 0.01$ [33, 37]. Semi-blind extensions of ILC, such as cILC (constrained ILC) [40] and cMILC (constrained Moment ILC) [33], are proposed to project out some foreground components by constraining the multi-frequency weights using the assumed-known average of the spectral energy distribution (SED). This methodology is able to effectively reduce the foreground bias with the disadvantage of rising the uncertainty. In this paper, we consider both NILC and cILC algorithms implemented in the pixel, harmonic and needlet domains, as well as, the cMILC implemented in the needlet domain.

In our analysis, we use the mock data of seven frequency bands combining the WMAP K-band, the 4 Planck HFI bands and the 2 AliCPT-1 bands, which in total spans from 23GHz to 353GHz. The AliCPT (Ali CMB Polarization Telescope), a ground-based experiment located in Tibet, will measure CMB polarization in the northern hemisphere at 90GHz and 150GHz [15]. The focus is on the trapezoid-shaped sky patch at the northern hemisphere with $f_{\mathrm{sky}} \sim 7\%$, centering at RA = $170°$ and DEC = $40°$ for the Celestial coordinate.

Our paper is organized as follows. In Section 2, we introduce the standard ILC method and its variants. In Section 3, we describe the sky simulations used in our analysis. In Section 4, we present the implementation details of the cILC method. In Section 5, we show the results of B-mode reconstruction and analyze the sources of biases. Finally, we conclude in Section 6.

2 Methods

2.1 Standard ILC

The ILC method is widely used in recovering the CMB map from the observed sky using several frequency channels without any assumption of foreground emissions and noise. The cleaned CMB map is constructed as a linear combination of the multi-frequency data where the variance is minimized, leaving the CMB component unchanged.

First we model the observed sky as a linear mixture of different emissions. For the cth emission component, let us assume it can be decomposed into a spatial template $s_c(p)$, which varies with the position in the sky, and a SED $a_c(\nu)$, which only depends on the frequency. The model for the observed data in the frequency band ν, where $\nu \in \{1, 2, \ldots, n_\nu\}$, at pixel p can be written as:

$$\boldsymbol{d}(p) = \boldsymbol{A}\boldsymbol{s}(p) + \boldsymbol{n}(p), \tag{2.1}$$

where $\boldsymbol{n}(p) = [n(\nu, p)]^T$ is a $n_\nu \times 1$ column vector of multi-frequency instrumental noise, $\boldsymbol{s}(p) = [s_c(p)]^T$ is a $n_c \times 1$ column vector of multi-component templates, and $\boldsymbol{d}(p) = [d(\nu, p)]^T$ is a $n_\nu \times 1$ column vector of the observed data. $\boldsymbol{A}$ denotes the $n_\nu \times n_c$ mixing matrix of the emission components with elements $a_c(\nu)$. Note that hereafter, the observed data is already deconvolved to a common beam and resolution: $d_{\ell m}^\nu = \frac{b_\ell^0}{b_\ell^\nu} d_{\ell m}^{\nu, obs}$, where b_ℓ^0 and b_ℓ^ν are the common beam window function and the beam for each frequency band, respectively. $d_{\ell m}^\nu$ is the harmonic coefficients of $d(\nu, p)$, and $d_{\ell m}^{\nu, obs}$ is the harmonic coefficients of the raw data.

The ILC method defines the estimated CMB map as a linear combination of all observed frequency maps with a constraint of weights $\sum_\nu w_\nu(p) = 1$:

$$
\begin{aligned}
\hat{s}_{\rm CMB}(p) &= \boldsymbol{w}^T(p) \cdot \boldsymbol{d}(p) \\
&= s_{\rm CMB}(p) + \sum_\nu w_\nu(p) \sum_c a_c(\nu) s_c(p) + \sum_\nu w_\nu(p) n(\nu, p) \,,
\end{aligned} \tag{2.2}
$$

where $\hat{s}_{\rm CMB}(p)$ is the estimate of the true CMB anisotropies $s_{\rm CMB}(p)$, $\boldsymbol{w}(p) = [w_\nu(p)]^T$ is a $n_\nu \times 1$ column vector, and c sums over foregrounds in the second term. Here, we assume that the CMB signal is frequency independent after calibration with respect to CMB, i.e. $a_{\rm CMB}(\nu) = 1$. The weights $w_\nu(p)$ are obtained by minimizing the ensemble variance of the ILC-cleaned map $\langle \hat{s}_{\rm CMB}^2(p) \rangle$, using the Lagrange multiplier method, when the variance of the error $(\hat{s}_{\rm CMB} - s_{\rm CMB})$ is therefore minimized. In practice, the variance localized at pixel p is estimated empirically by the averaged variance of the observed data over the vicinity of the p-th pixel. Defining the $n_\nu \times 1$ CMB mixing vector as $\boldsymbol{a} = [1, \ldots, 1]^T$, the ILC weights are given as:

$$
\boldsymbol{w}^{\rm ILC}(p) = \frac{\hat{\boldsymbol{C}}^{-1} \boldsymbol{a}}{\boldsymbol{a}^T \hat{\boldsymbol{C}}^{-1} \boldsymbol{a}} \,, \quad \text{where } \hat{\boldsymbol{C}}(p) = \frac{1}{N_{p'}} \sum_{p' \in V(p)} q(p', p) \boldsymbol{d}(p')^\dagger \boldsymbol{d}(p') \,. \tag{2.3}
$$

Here $\hat{\boldsymbol{C}}$ is the $n_\nu \times n_\nu$ covariance matrix of the data as an estimate of the real covariance matrix, $V(p)$ is the vicinity of the pixel p, and $q(p', p)$ is some type of weight of p' around the p-th pixel (for instance a Gaussian smooth kernel centered at p).

The method above implemented in pixel space is denoted as pILC. Likewise, the ILC implemented in harmonic space is denoted hILC, just replacing the pixel p to the harmonic mode (ℓ, m), and the vicinity $V(p)$ to the binning of (ℓ, m)s around ℓ. The hILC weights are given by:

$$
\boldsymbol{w}_\ell = \frac{\hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}}{\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} \,, \quad \text{where } \hat{\boldsymbol{C}}_\ell = \sum_{\ell' \in B(\ell)} \sum_{m=-\ell'}^{\ell'} \boldsymbol{d}_{\ell'm}^\dagger \boldsymbol{d}_{\ell'm} \,. \tag{2.4}
$$

The covariance matrix $\hat{\boldsymbol{C}}_\ell$ is summed over the bin $B(\ell)$ centered at ℓ and the foreground-cleaned CMB map is $\hat{s}_{\ell m} = \boldsymbol{w}_\ell^T \cdot \boldsymbol{d}_{\ell m}$. The ILC in needlet domain is introduced in next section.

2.2 NILC

The needlet ILC, as a refinement of ILC, has been applied to extract CMB T-modes and E-modes for various experiments [41–43]. The sky map of each channel is decomposed into a set of harmonic-filtered maps, which are localized in harmonic domain, and for each filtered map a perfect localization in pixel domain can be done by ILC independently. The observed data is filtered by the needlet bands:

$$
d_{\ell m}^{\nu, j} = h_\ell^j d_{\ell m}^\nu \,, \tag{2.5}
$$

where the needlet bands h_ℓ^j satisfying $\sum_j (h_\ell^j)^2 = 1$ are responsible for the localization in harmonic space. The cosine and Gaussian needlet bands are commonly used in NILC.

Assuming `HEALPix` pixelization [44], the spherical needlet function is defined as:

$$
\psi_{jk}(p) = \sqrt{\frac{4\pi}{N_j}} \sum_{\ell m} h_\ell^j Y_{\ell m}(p) Y_{\ell m}^*(n_{jk}) \,, \tag{2.6}
$$

where N_j is the NPIX parameter of the j-th needlet map, and n_{jk} refers to the k-th pixel of the j-th needlet map. The j-th needlet map transformed from the observed data $d(\nu, p)$ is then given by:

$$
\begin{aligned}
b_j^\nu(n_{jk}) &= \frac{4\pi}{N_p} \sum_p d(\nu, p)\psi_{jk}^*(p) \\
&= \sqrt{\frac{4\pi}{N_j}} \sum_{\ell m} h_\ell^j d_{\ell m}^\nu Y_{\ell m}(n_{jk}) ,
\end{aligned}
\tag{2.7}
$$

while its inverse transformation is given by:

$$
d_{\ell m}^\nu = \sum_{jk} b_j^\nu(n_{jk}) \sqrt{\frac{4\pi}{N_j}} h_\ell^j Y_{\ell m}^*(n_{jk}) .
\tag{2.8}
$$

The needlet transformation is a linear operation. The final NILC map is transformed from the individually-cleaned needlet maps, which are linearly combined by the input multi-frequency needlet maps for each needlet band j:

$$
\hat{s}_{\ell m}^{\mathrm{NILC}} = \sum_{jk} b_j^{\mathrm{NILC}}(n_{jk}) \sqrt{\frac{4\pi}{N_j}} h_\ell^j Y_{\ell m}^*(n_{jk}) ,
\tag{2.9a}
$$

$$
b_j^{\mathrm{NILC}}(n_{jk}) = \sum_\nu w_{\nu,j}^{\mathrm{NILC}}(n_{jk}) b_j^\nu(n_{jk}) .
\tag{2.9b}
$$

Just as the pixel domain ILC method, the NILC weights are given by:

$$
\boldsymbol{w}_j^{\mathrm{NILC}}(n_{jk}) = \frac{\hat{\boldsymbol{C}}_{jk}^{-1} \boldsymbol{a}}{\boldsymbol{a}^T \hat{\boldsymbol{C}}_{jk}^{-1} \boldsymbol{a}} .
\tag{2.10}
$$

To estimate the covariance matrices for the k-th pixel of the j-th needlet scale $\boldsymbol{C}_{jk} = C_{jk}^{\nu_1 \times \nu_2} = \langle b_j^{\nu_1}(n_{jk}) b_j^{\nu_2}(n_{jk}) \rangle$, we compute the empirical covariance $\hat{\boldsymbol{C}}_{jk}$ by averaging the needlet coefficient products $b_j^\nu(n_{jk}) b_j^\nu(n_{jk})$ over some domain of pixels around pixel k, which can be written as:

$$
\hat{C}_{jk}^{\nu_1 \times \nu_2} = \frac{1}{N_k} \sum_{k'} q_j(k, k') b_j^{\nu_1}(n_{jk'}) b_j^{\nu_2}(n_{jk'}) ,
\tag{2.11}
$$

where $q_j(k, k')$ are weights dependent on the needlet scale. More details are introduced in Section 4.3.

The final NILC cleaned map is inverse spherical harmonic transformed (SHT) to pixel space as:

$$
\hat{s}_{\mathrm{NILC}}(p) = \sum_{\ell m} \hat{s}_{\ell m}^{\mathrm{NILC}} Y_{\ell m}(p) .
\tag{2.12}
$$

2.3 Constrained ILC

The constrained ILC (cILC) method [40] introduces an assumption about the SED of the main foreground components. By construction, it adds constraints to cancel those unwanted components. However, while this process reduces the residual foregrounds, it increases the noise level as an expense.

Given the mixing matrix, the constraints are given by:

$$\sum_\nu w_\nu^{\mathrm{cILC}} a_c(\nu) = e_c \,, \tag{2.13}$$

where e_c is 1 for CMB and 0 for all foreground components modeled in the mixing matrix. The cILC weights are given by:

$$\boldsymbol{w}^T = \boldsymbol{e}^T \left(\boldsymbol{A}^T \hat{\boldsymbol{C}}^{-1} \boldsymbol{A} \right)^{-1} \boldsymbol{A}^T \hat{\boldsymbol{C}}^{-1} \,, \tag{2.14}$$

where $\hat{\boldsymbol{C}}$ is the covariance matrix of the data. The cILC cleaned CMB map is given by $\hat{s}_{\mathrm{CMB}} = \boldsymbol{w}^T \cdot \boldsymbol{d}$.

As the ILC, the cILC can also be implemented in pixel, harmonic and needlet domain, hereafter denoted by cpILC, chILC and cNILC, respectively.

2.3.1 The mixing matrix

The mixing matrix used in cILC is determined by the assumed SEDs of the foreground components of interest, which, in our analysis, are synchrotron and thermal dust. We model the SED of CMB anisotropies as a differential black body, of synchrotron emission as a power law with a fixed spectral index of $\beta_{\mathrm{sync}} = -3$, and of the thermal dust as a modified black body with the dust temperature of $T_{\mathrm{dust}} = 19.6\mathrm{K}$ and the dust spectral index of $\beta_{\mathrm{dust}} = 1.59$, for data in Rayleigh-Jeans brightness temperature units. These SED parameters are from the Planck constraints on polarized foregrounds [45, 46]. The SEDs after calibration to the CMB thermodynamic temperature are written as:

$$\begin{cases} a_{\mathrm{CMB}}(\nu) = 1 \,, \\ a_{\mathrm{sync}}(\nu) = \dfrac{g_\nu}{g_{\nu_s}} (\dfrac{\nu}{\nu_s})^{\beta_s} \,, \\ a_{\mathrm{dust}}(\nu) = \dfrac{g_\nu}{g_{\nu_d}} (\dfrac{\nu}{\nu_d})^{\beta_d+1} \dfrac{\exp(x_d(\nu_d)) - 1}{\exp(x_d(\nu)) - 1} \,, \end{cases} \tag{2.15}$$

where $x_d(\nu) \equiv \frac{h\nu}{k_B T_d}$, $\nu_s = 23\mathrm{GHz}$ and $\nu_d = 353\mathrm{GHz}$ are the reference frequencies of synchrotron and dust emissions, respectively. g_ν is the unit conversion factor from μK_{RJ} to μK_{CMB}.

In principle, the adopted SED parameters should fit the real sky foregrounds to optimize the foreground cleaning, and should be position-dependent in the best case for the actual data with parameters varying over the sky. On the other hand, given lack of prior knowledge of the polarized foregrounds, the mixing matrix should be robust enough to be applied on different sky models. Therefore, in Section 5.4 we test the cILC method on four different foreground models (introduced in Section 3.4).

2.4 Extensions: constrained Moment-ILC

The constrained Moment-ILC (cMILC) [33] is an extension of cILC that adds several nulling constraints on the SEDs of the main moments of the foreground emissions. Assuming that the SED of the cth component varies over the sky, we can expand the mixing amplitude around some fixed pivot SED parameters $\bar{\boldsymbol{\beta}} = (\bar{\beta}_1, \ldots \bar{\beta}_n)^T$:

$$a_c(\nu, \boldsymbol{\beta}(\boldsymbol{p})) = \sum_k \sum_{\alpha_1 + \cdots + \alpha_n = k} \frac{(\beta_1(p) - \bar{\beta}_1)^{\alpha_1} \cdots (\beta_n(p) - \bar{\beta}_n)^{\alpha_n}}{\alpha_1! \cdots \alpha_n!} \frac{\partial^k a_c(\nu, \bar{\boldsymbol{\beta}})}{\partial \bar{\beta}_1^{\alpha_1} \cdots \partial \bar{\beta}_n^{\alpha_n}} \,. \tag{2.16}$$

Table 1: Summary Table of seven channels used in our simulations, including WMAP K-band, AliCPT 95GHz and 150GHz bands, Planck HFI bands. σ_n^{P} is the map noise level for polarization. Note that the noise levels are computed by fitting the noise power spectra in the sky patch with the white noise spectra, where the Planck levels are slightly lower than the values from Table 4 of [48]. The ideal channel-combined noise level is about 10μK-arcmin.

Instruments	WMAP	AliCPT		Planck HFI			
Frequency(GHz)	23	95	150	100	143	217	353
Beam size(arcmin)	52.8	19	11	9.7	7.3	5.0	4.9
$\sigma_n^{\mathrm{P}}(\mu$K-arcmin)	496	13	18	78	65	91	404

Under the approximation of the finite-order expansion, the derivative terms can all be analytically computed in terms of the channel ν and the fixed $\bar{\boldsymbol{\beta}}$. Similar with the cILC method, we add constraints to null these high-order moments of the foreground SED:

$$\sum_\nu w_\nu^{\mathrm{cMILC}} \frac{\partial^k a_c(\nu, \bar{\boldsymbol{\beta}})}{\partial \bar{\beta}_1^{\alpha_1} \cdots \partial \bar{\beta}_n^{\alpha_n}} = e_c \,. \tag{2.17}$$

It can be seen that the constraint turns to that of cILC when $k = 0$. The additional constraints (of $k = 1, 2, \ldots$) further reduce the residual foreground with an expense of increasing the residual noise power. Besides, the number of observation channels cannot be less than the number of constraints to ensure the solvability of the weights. In this work, with seven bands involved, cMILC is considered up to the first order ($k = 1$) and implemented in needlet space with an additional constraint corresponding to the first order derivative of the thermal dust with respect to the dust temperature T_d. More specifically, the new constraint is given by:

$$\sum_\nu w_\nu^{\mathrm{cMILC}} \frac{\partial a_{\mathrm{dust}}(\nu)}{\partial T_d}|_{\bar{T}_d} = 0 \,, \tag{2.18}$$

where the pivot SED parameters are the same as the parameters of the mixing matrix.

3 Simulations

The simulated data sets are based on the first season observation of the future CMB experiment AliCPT, with Planck HFI four bands and WMAP K band combined as ancillary simulations to improve the efficiency of foreground removal. Unlike simulations in [47], our maps are not generated from time-ordered data (TOD), and instrumental systematics like filtering effects are not involved in our analysis, which can be generally modeled as a set of transfer functions independent of the foreground removal process. Each realization includes observations of seven frequency channels: the 95GHz and 150GHz bands of AliCPT mission, four bands of Planck HFI instruments(100, 143, 217 and 353GHz), and the WMAP K band (23GHz). The input maps are the sum of CMB, noise and foregrounds smoothed with the instrumental Guassian beams information listed in Table 1. We adopt a `HEALPix` $N_{side} = 1024$ pixelization scheme.

3.1 Masks

The AliCPT scanning region in the first observation season centers at RA $= 180°$, DEC $= 30°$ in the northern sky, covering about 17% of the sky area, as shown in Figure 1. In the

(a) binary masks   (b) inverse variance
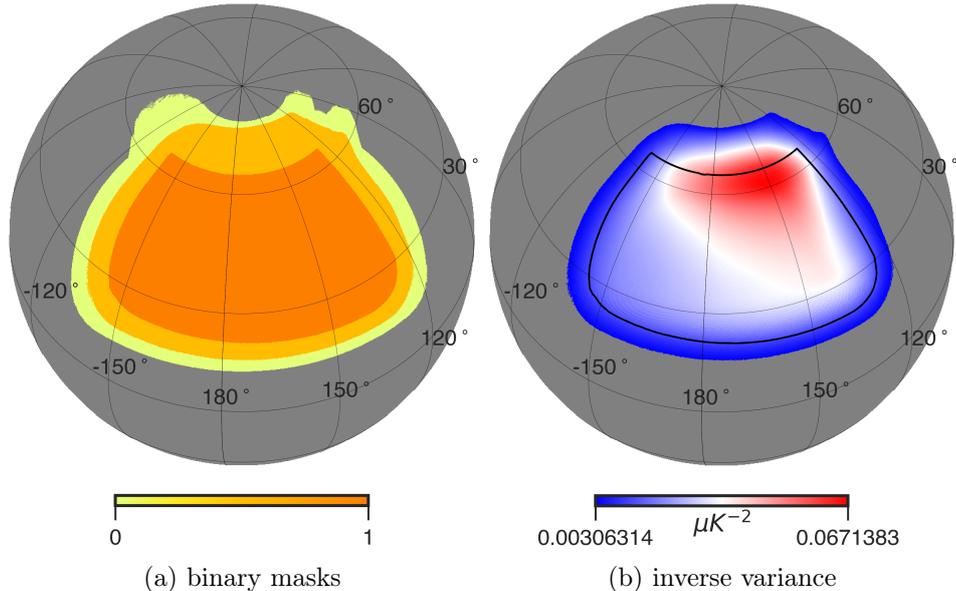
Figure 1: Masks in the AliCPT-1 data analysis. (a) From the outer to inner region, the figure shows the AliCPT-1 observation patch (yellow), the $20\mu K$ mask (light orange), and the UNP mask(dark orange), with sky fractions of 13%, 10%, 6.7%, respectively. (b)The inverse variance of the polarization noise at 150GHz. The black curve shows the boundary of the UNP mask, inside which is the UNP-inv mask.

foreground cleaning pipeline, we adopt two binary masks: the '$20\mu K$' mask and the 'UNP' mask. They are shown in Figure 1. The $20\mu K$ mask is defined by masking the pixels with the noise standard deviation above $20\mu K$ at `NSIDE`=1024 for the 150GHz channel. We then further remove the sky above declination of $65°$, obtaining a mask with $f_{sky} = 6.7\%$ named the UNP mask. We also divide the UNP mask by the variance of the polarization noise at 150GHz, hereafter named the UNP-inv mask.

3.2 CMB

We produce two sets of lensed CMB data ($A_L = 1$) with a fiducial tensor-scalar ratio of $r = 0.03$ and the null value $r = 0.0$, with each of the sets containing 300 seven-channel simulations. The CMB maps are Gaussian realizations generated from the power spectra generated by the `CAMB` [49] package using the best-fit Planck 2018 parameters[1] [3]. The lensing effects are added by the `LensPyx` [50] package. Unless specified, the CMB realizations with $r = 0.03$ are used in the following analysis.

3.3 Noise

The instrumental noises for AliCPT-1 and WMAP-K channels are Gaussian white realizations generated from the noise variance maps of AliCPT-1 bands and WMAP K band (upgraded to $N_{side} = 1024$), respectively. For Planck HFI four bands, we use Planck FFP10 noise simulations from the Planck Legacy Archive[2].

---

[1]Specifically, the parameters include dark matter dencity $\Omega_c h^2 = 0.120$, baryon density $\Omega_b h^2 = 0.02237$, scalar spectral index $n_s = 0.9649$, optical depth $\tau = 0.0544$, Hubble constant $H_0 = 69.36$ km s^{-1} Mpc^{-1} and the primordial comoving curvature power spectrum amplitude $A_s = 2.10 \times 10^{-9}$.

[2]http://pla.esac.esa.int/pla
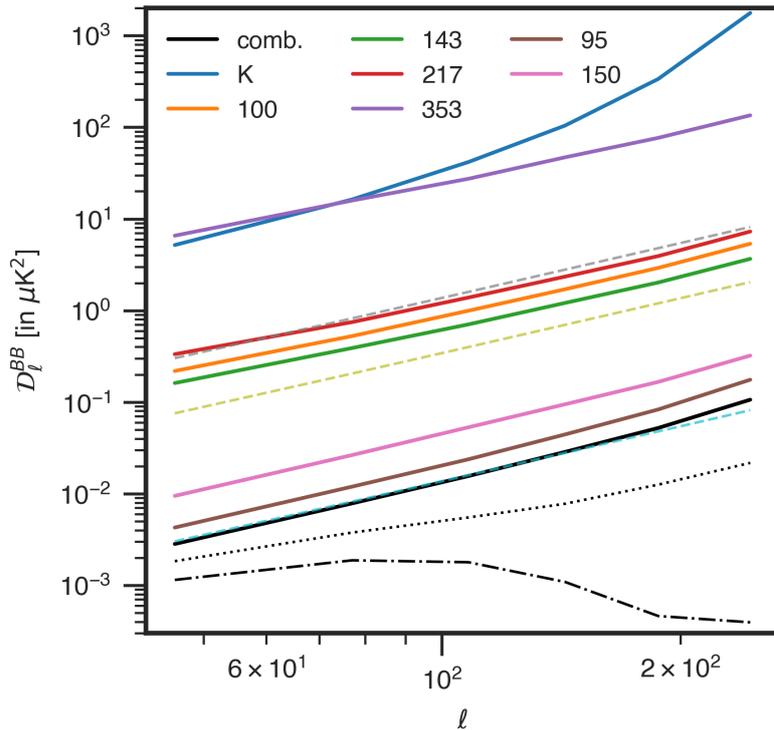
Figure 2: The debeamed polarization noise power spectra $\mathcal{N}_\ell/b_\ell^2$ of our seven frequency channels. For reference, the white noise spectra with noise levels of 10, 50 and 100 μK-arcmin are plotted as three dashed curves from bottom to top. The black solid curve is the channel-combined noise spectrum, which is about 10 μK-arcmin. The lensed CMB BB spectrum (black dotted curve) and the tensor BB spectrum with $r = 0.03$ (black dot-dashed curve) are also plotted for comparison.

The debeamed polarization noise power spectra $\mathcal{N}_\ell/b_\ell^2 = \frac{\ell(\ell+1)}{2\pi} N_\ell/b_\ell^2$ for seven channels are shown in Figure 2. We hereafter omit the b_ℓ^2 factor since the noises are all debeamed in the final cleaned map. The ideal noise level combining the data of all channels is about 10 μK-arcmin, given by $\mathcal{N}_\ell = [\sum_\nu \mathcal{N}_{\ell,\nu}^{-1}]^{-1}$.

3.4 Foregrounds

The diffuse foregrounds include various components, such as synchrotron, thermal dust, anomalous microwave emission (AME), free-free and CO emissions. However, the synchrotron emission and the thermal dust are known to be two main polarization contaminants among them. We do not consider the polarized point sources in our main analysis since they are not modeled in the cILC method, but we show in Appendix A that their effects are negligible to the extent of current sensitivity. In order to properly assess the performance of cILC methods on various sky models, we use four Galactic models based on the Planck Sky Model (PSM) [51] to generate our foreground simulations using the Python Sky Model (pysm [52]) package. The models are specified as follows:

- *AliCPT foregrounds:* In this model, the thermal dust polarization maps are generated from the GNILC template of the Planck 2018 release [43], scaled to different frequencies by a

modified blackbody SED with the dust temperature and spectral indices from the GNILC Planck 2015 dust maps best-fit [53]. The synchrotron polarization template is based on the Planck 2018 SMICA map, and it is scaled by a power law SED with a fixed β_s of -3.08. In this work, this is our baseline model.

- *d0s0:* The d0s0 pysm model uses the 353GHz map from Planck 2015 release as the thermal dust polarization template, and the WMAP 9-year 23 GHz Q/U map [54] as the synchrotron template, with fixed spectral indicies of $\beta_s = -3$, $\beta_d = 1.54$ and $T_d = 20K$. The polarization templates are smoothed with a Gaussian kernel of FWHM 2.6° for thermal dust and 5° for synchrotron, and have small scales added via the procedure described in [55].

- *d1s1:* The templates for the d1s1 pysm model are the same as d0s0. The spatially varying spectral indices for thermal dust are obtained from the Planck 2015 Commander dust map [56]; for synchrotron β_s is derived using a combination of the Haslam 408 MHz data and WMAP 23 GHz 7-year data [57].

- *d10s5:* The thermal dust maps are based on the Planck 2018 GNILC maps [43], which present a larger variation of β_d and T_d parameters at low resolution than the d1 model. The synchrotron templates are the same as s1, with the spectral index map from s1 rescaled based on the S-PASS data [58]. Small-scale structures are added as Gaussian realizations of power-law power spectra.

Four sets of multi-frequency foregrounds are generated using the above models. A set of B-mode maps (CMB, noise and foreground) at 150GHz for AliCPT-1 is shown in Figure 3. The B-mode maps are obtained by the template cleaning method (see Section 4.1).

4 Pipelines: CMB B-mode reconstruction

We now implement the NILC, chILC, cpILC, cNILC and cMILC methods on the multi-frequency data sets, where the cMILC is implemented in needlet space. First we have to extract the pure B-mode maps from the input partial-sky QU maps considering the partial-sky effects. Then we apply the ILC cleaning process on the B-mode maps in pixel, harmonic or needlet domain. Finally, we compute the noise debiased power spectra and extract the tensor-to-scalar ratio value from it.

4.1 EB leakage and Power spectrum estimation

Two effects must be considered when reconstructing the full-sky power spectrum from a partial-sky map: the EB leakage, and the mode coupling of the pseudo power spectrum. The template cleaning method introduced by Liu et al. [31] is an efficient way to remove the EB leakage in pixel domain, which can suppress the leakage to a level of $r \sim 10^{-4} - 10^{-5}$ over $\ell \sim 60$.

The template cleaning steps are as follows.

1. The masked polarization map $\boldsymbol{P} = (Q, U)$ is decomposed into the so-called E and B family maps: $\boldsymbol{P_E} = (Q_E, U_E)$ and $\boldsymbol{P_B} = (Q_B, U_B)$, which is done by $(Q, U) \to E; B \to (Q_E, U_E); (Q_B, U_B)$.

2. Obtain $\boldsymbol{P'_B}$ in the same way from the masked E family $\boldsymbol{P_E}$, i.e. $\boldsymbol{P_E} \to \boldsymbol{P'_E}; \boldsymbol{P'_B}$.

(a) CMB   (b) Noise   (c) AliCPT FG

(d) d0s0 FG   (e) d1s1 FG   (f) d10s5 FG
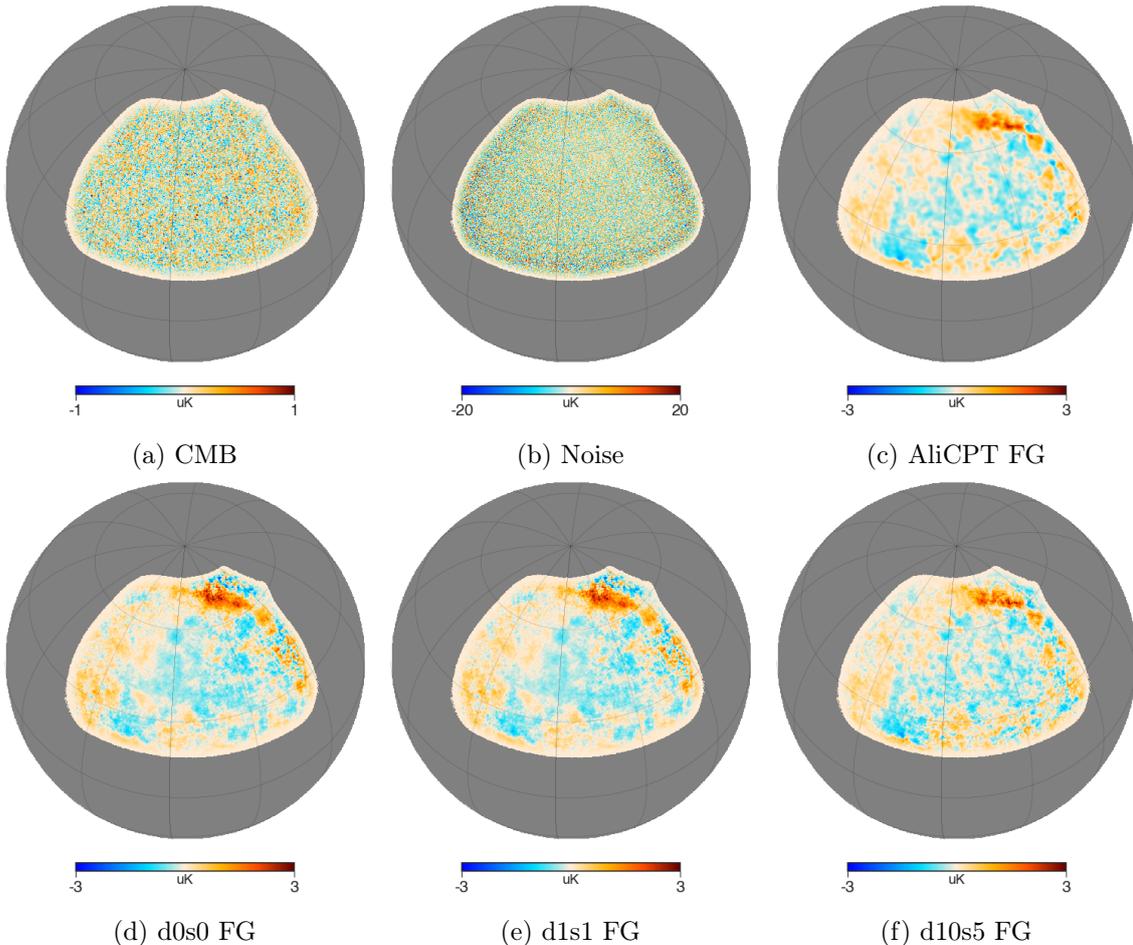
Figure 3: One sample of the simulated CMB, noise and foreground B-mode maps at 150GHz band after template cleaning with 6° C2 apodization of the 20μK mask.

3. Use the masked $\boldsymbol{P'_B}$ as a E-B leakage template to remove the leakage from the masked $\boldsymbol{P_B}$ by linear fitting.

After correcting the EB leakage and foreground cleaning, we compute the B-mode power spectrum with the pseduo-C_ℓ (PCL) estimator using the `NaMASTER` [59][3] python package, which takes the masking, binning, pixel-window and beaming effects into account.

In order to assess the accuracy of power spectrum estimation on foreground-free maps, we apply the template cleaning and the PCL process on 100 CMB-only simulations smoothed with Gaussian beams of either FWHM=52.8 arcmin or 11 arcmin as the possible common beams, corresponding to the largest beam among our frequency channels and the beam used in AliCPT for NILC [60], respectively. First, we obtain the real B-mode maps from the full-sky CMB simulations as reference. The simulations using the 20μK mask go through the template cleaning pipeline to correct for the leakage. The corrected B-mode maps are then masked again by the 20μK mask with 6 degree C2 type apodization in order to reduce the high leakage residual on the boundary of the mask [31]. Finally, we apply PCL on the leakage-cleaned maps to reconstruct the full-sky power spectra. The estimated BB power

---

[3]https://github.com/LSSTDESC/NaMaster
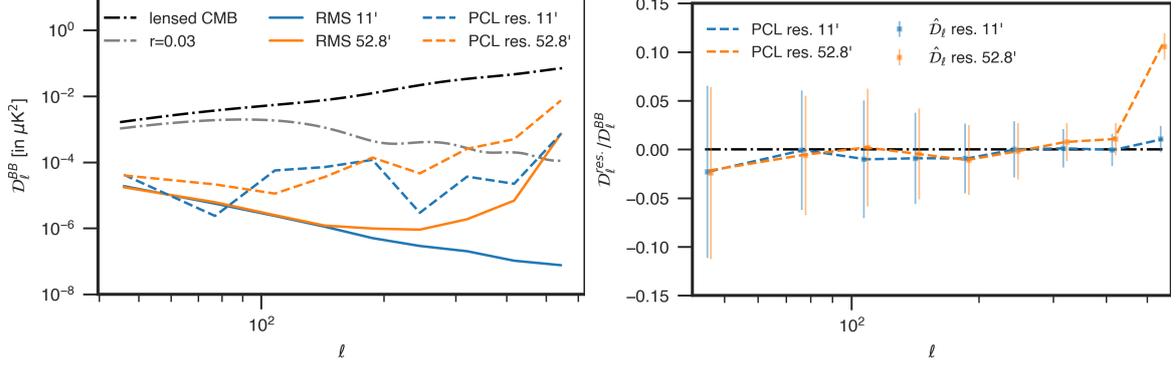
Figure 4: Left: the RMS defined as Eq.(4.1) and the PCL residual averaged over 100 CMB-only simulations with the beam size of either 11 arcmin or 52.8 arcmin. Right: the ratios between the residual of the estimated power spectrum and the input BB power spectrum. The relative PCL residuals are also shown by dashed curves.

spectra are binned with $\Delta\ell = \max\{30, 0.3\ell_{\min}\}$ (where $\ell_{\min}$ is the starting multipole of a bin) from $\ell = 40$ to 600.

To evaluate the residual leakage, we compute the RMS error which depends only on the error of template cleaning as:

$$\Delta = \sqrt{\frac{1}{N_{\text{sim}}} \sum_{i=1}^{N_{\text{sim}}} (\hat{\mathcal{D}}_\ell - \mathcal{D}_\ell^{\text{real}})^2}, \qquad (4.1)$$

where $\hat{\mathcal{D}}_\ell$ is the estimated power spectrum for each simulation and $\mathcal{D}_\ell^{\text{real}}$ is obtained by running `NaMASTER` on the real full-sky B maps with the $6°$ C2 apodized $20\mu K$ mask. Besides, the difference between $\mathcal{D}_\ell^{\text{real}}$ and the input CMB spectrum $\mathcal{D}_\ell^{\text{BB}}$(which are bandpowers in practice) reflects the residual power spectrum only from the PCL estimation. The RMS error (i.e. the error from the residual EB leakage) as well as the mean PCL residual $\langle\mathcal{D}_\ell^{\text{real}} - \mathcal{D}_\ell^{\text{BB}}\rangle$ are shown in the left panel of Figure 4. The right panel shows the ratio between the residual of the estimated power spectrum and the input BB power spectrum, $\langle\hat{\mathcal{D}}_\ell - \mathcal{D}_\ell^{\text{BB}}\rangle/\mathcal{D}_\ell^{\text{BB}}$, averaged over 100 CMB-only simulations. It can be seen that the RMS or the residual leakage is negligible. Thus nearly all the residual comes from the PCL estimator (as shown in the right panel), while the PCL residual surpasses the power spectrum for $r = 0.03$ over $\ell \sim 300$ in the case of FWHM=52.8 arcmin. The PCL residual might come from the large computational uncertainty of the deconvolved power spectrum after dividing by the beam function squared (see Figure 5 for the beam functions). Since the bias of the power spectrum estimation at the beam size of 52.8 arcmin exceeds the 1σ threshold over $\ell \sim 400$, the multipole range is restricted from 40 to 400 when adopting the common beam of 52.8 arcmin. Hereafter, unless specially mentioned, we adopt FWHM=11 arcmin as the common beam while discarding the K-band data at scales of $\ell > 300$ to avoid the extreme noise amplification of this specific frequency channel.

4.2 ILC implemention in harmonic domain

Here, we clarify the detailed procedure of foreground cleaning using cILC in harmonic space. After template cleaning, the multi-band B-mode maps masked by the UNP-inv mask are
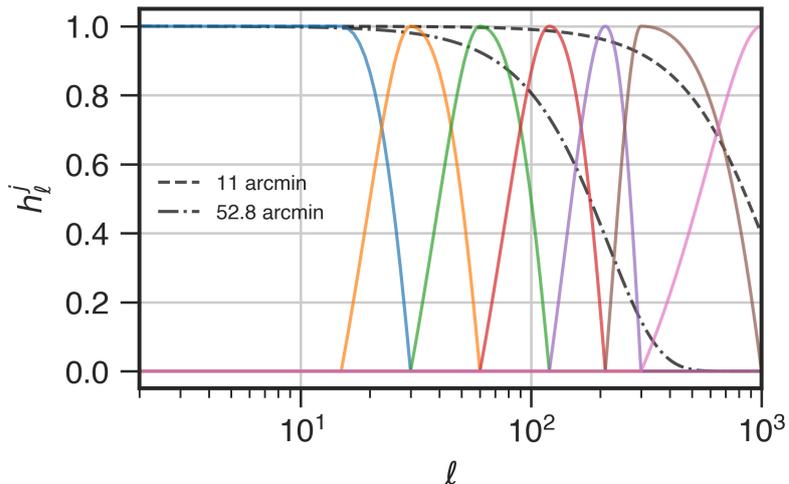
Figure 5: The window fuctions of seven needlet bands. The beam functions of 11 arcmin and 52.8 arcmin are added for ancillary view.

transformed to harmonic space and reconvolved to a common beam. The ℓ-dependent empirical covariance is averaged by the $a_{\ell m}$s over the bin range $[0.6\ell, 1.4\ell]$ with weights of $(2\ell + 1)$. We then multiply the cILC weights computed from the covariance matrix by the channelwise B-mode $a_{\ell m}$s, resulting in the cleaned B-mode $a_{\ell m}$. The final full-sky power spectrum is obtained by PCL after debiasing the noise power spectrum.

4.3 ILC implemention in needlet (pixel) domain

The ILC implemention in needlet space is somewhat more complicated than in harmonic space due to the discriminatory handling of multi-scale needlet bands. The maps after template cleaning are reconvolved to a common beam and decomposed into needlet coefficients, where the cosine bands are adopted as the needlet bands in this work, as shown in Figure 5. The cosine window functions can be written as:

$$
h_\ell^j = \begin{cases}
\cos(\dfrac{\pi}{2} \dfrac{\ell_{mid}^j - \ell}{\ell_{mid}^j - \ell_{mid}^{j-1}}) , \ell_{mid}^{j-1} \le \ell < \ell_{mid}^j , \\[3mm]
\cos(\dfrac{\pi}{2} \dfrac{\ell - \ell_{mid}^j}{\ell_{mid}^{j+1} - \ell_{mid}^j}) , \ell_{mid}^j \le \ell < \ell_{mid}^{j+1} , \\[3mm]
0 , \text{otherwise} ,
\end{cases}
\tag{4.2}
$$

where $j \in \{1, 2, \ldots, n_j\}$, $h_\ell^1 = 1$ when $0 \le \ell < \ell_{mid}^1$, and $h_\ell^{n_j} = 1$ when $\ell \ge \ell_{mid}^{n_j}$. The seven needlet bands used in our tests are listed in Table 2. Note that the needlet maps of larger scales containing less modes are therefore downgraded to lower N_{side}s. Just like for chILC, we discard K-band data for those needlet bands covering the range of $\ell > 350$, i.e. the 6th and 7th bands. The needlet coefficients are also reconvolved with a common beam of 11 arcmin FWHM.

The covariance matrices $\hat{C}_{jk}$ for the k-th pixel of the j-th needlet scale are estimated as follows: first we downgrade the j-th needlet maps of every channel to a super N_{side} of $N_{side}(j)/4$. Then, the product of downgraded needlets of two channels ν_1 and ν_2,

– 13 –

Table 2: The needlet bands used for needlet decomposition and the corresponding smoothing FWHMs when computing the covaraince matrix.

Band(j)	1	2	3	4	5	6	7
ℓ_{mid}^{j}	15	30	60	120	210	300	1000
$N_{side}(j)$	32	64	128	256	512	1024	1024
FWHM($^{\circ}$)	124	72	36	20	15	4.2	3.5

$b_j^{\nu_1}(n_{jk})b_j^{\nu_2}(n_{jk})$, is smoothed with a Gaussian beam whose FWHM varies with the scale j. Each channel pair of needlets returns a covariance map, of which the k-th pixel is taken as the (ν_1, ν_2) element of the covariance matrix $\hat{\boldsymbol{C}}_{jk}$. The estimated covariance contains the information of pixels in a disk centered at the k-th pixel, averaged by a Gaussian beam. The smoothing FWHMs of seven needlet bands are listed in Table 2. The FWHM values are chosen such that the ILC bias [27] is under 3% of the CMB.

We compute the weights for each n_{jk} from the covariance matrix independently and project them on the channel-wise needlet maps to obtain the foreground-cleaned needlet maps. Finally we synthesize them to the final cNILC cleaned map, as the inversion of needlet decomposition.

The cILC in pixel space (cpILC) is basically a simple version of cNILC without needlet filtering. Note that in this case ℓ is limited up to 300 since the debeamed K band data over the limit would be too large to recover any useful information. As mentioned in Section 2.4, the cMILC is implemented in needlet space adding one more constraint. Thereby, the only processing difference between cMILC and cNILC is that the mixing matrix of cMILC contains an additional column vector describing the SED of the derivative of the thermal dust.

4.4 Noise debiasing

It is essential to debias the noise power spectrum from the foreground-cleaned power spectrum where the residual noise power is the predominant component for B-modes. The noise bias can be properly corrected either by the cross spectrum of two data splits between which the noises are uncorrelated, or by estimating the residual noise power spectrum from noise simulations. In this work, we estimate the noise bias from 100 noise simulations projected by the cILC(or NILC) weights obtained by the mock data.

4.5 Fitting the tensor-to-scalar ratio

We fit the tensor-to-scalar ratio r with the estimated power spectra by the Markov Chain Monte Carlo (MCMC) simulations. Ignoring the slight non-Gaussianity from the foregrounds given the large number of multipoles in a bin, the Gaussian likelihood of r given the binned cILC cleaned power spectrum $\hat{C}_{\ell_b}$ is:

$$-2\ln\mathcal{L}(r) = \sum_{\ell_b \ell_{b'}} \left[\hat{C}_{\ell_b} - rC_{\ell_b}^{r=1} - C_{\ell_b}^{\mathrm{lens}}\right] \left[M_{\mathrm{fid}}^{-1}\right]_{\ell_b \ell_{b'}} \left[\hat{C}_{\ell_{b'}} - rC_{\ell_{b'}}^{r=1} - C_{\ell_b}^{\mathrm{lens}}\right], \qquad (4.3)$$

where $\ell_b, \ell_{b'}$ are indices for the multipole bins, $rC_{\ell_b}^{r=1} + C_{\ell_b}^{\mathrm{lens}}$ is the binned theoretical B-mode power spectrum with multipole range $[40, 200]$ including 5 ℓ_bs, and the fiducial covariance matrix, $[\boldsymbol{M}_{\mathrm{fid}}]_{\ell_b \ell_{b'}} = \langle(\hat{C}_{\mathrm{fid},\ell_b} - \langle\hat{C}_{\mathrm{fid},\ell_b}\rangle)(\hat{C}_{\mathrm{fid},\ell_{b'}} - \langle\hat{C}_{\mathrm{fid},\ell_{b'}}\rangle)\rangle$ is computed from 300 cILC-cleaned simulations with $r = 0$ and the foregrounds same with the fitted data. The theoretical

spectrum is computed from `CAMB` with the best-fit parameters of Planck 2018 results [3]. We do no delensing to the B-mode power spectrum. The above likelihood is multiplied by a uniform prior of $r \in [0, 1]$ to obtain the unnormalized posterior $P(r|\boldsymbol{d})$. We set a MCMC chain including 10000 samples satisfying the posterior distribution using the `emcee` [61] python package. Therefore, the mean r-value extracted from the distribution of 10000 samples can be taken as the best-fit r for the estimated power spectrum, with the uncertainty $\sigma(r)$ (also extracted from the distribution) due to the cosmic variance of noise and residual foregrounds embedded in the fiducial covariance matrix.

5   Results

In this section, we present the results of the foreground-cleaned maps and the estimated B-mode power spectra using different ILC methods. The input realizations are the sum of CMB, noise and AliCPT foregrounds unless specified. We analyze the possible sources of bias and quantitatively analyze the bias in our results in Section 5.3 and 5.3.1. We present the results considering various foregrounds while using the assumed mixing matrix in Section 5.4. We compute the best-fit r and its uncertainty for different cases in Section 5.5.

5.1   Maps

First we illustrate the effectiveness of ILC methods on recovering the CMB map. The cleaned CMB maps and the residual foreground and noise contaminations of NILC, cNILC and cMILC, three ILC methods with increasing number of constraints, are shown in Figure 6. The residual contaminations are produced by applying the same ILC weights on the input foreground and noise maps with the input simulations. It can be seen that the additional constraints on foregrounds improve the foreground removal at the expense of increasing the noise residual in the recovered maps. Which method is the best depends on the object to study. On the map level, adding the constraints introduces much more noise contaminations than the foregrounds further removed. On the power spectrum level, as the noise residual can be properly estimated by noise realizations, it becomes significant to reduce the foreground bias.

5.2   Power spectra

The comparison of the estimated power spectra using NILC, chILC, cpILC, cNILC and cMILC and their respective uncertainties are shown in Figure 7. The effectiveness and the accuracy of foreground cleaning can be quantified in a straightforward way by analyzing both noise and the foreground residuals. They are obtained by linearly combining the input noise and foregrounds simulations with the ILC weights obtained from the input realization.

It can be seen that compared to NILC, adding constraints reduces the foreground residual significantly while increasing the noise residual since the variance $\boldsymbol{w}^T \hat{\boldsymbol{C}} \boldsymbol{w}$ with constrained weights cannot remain as small as the unconstrained variance (see Eq.(39) of [33]). The cMILC adds more constraints than cILC, resulting in the lowest residual foreground at the expense of the largest noise level. Moreover, the performance of the cpILC is poor, meaning that its foreground residual and noise residual are both higher than chILC and cNILC. The chILC result is fairly similar to the cNILC result, except its overall lower noise level and slightly lower foreground residual at scales of $\ell \sim 100$.

The ILC method makes a tradeoff between reducing the residual foregrounds and reducing the residual noise. The ILC in harmonic space does not take into account the spatial
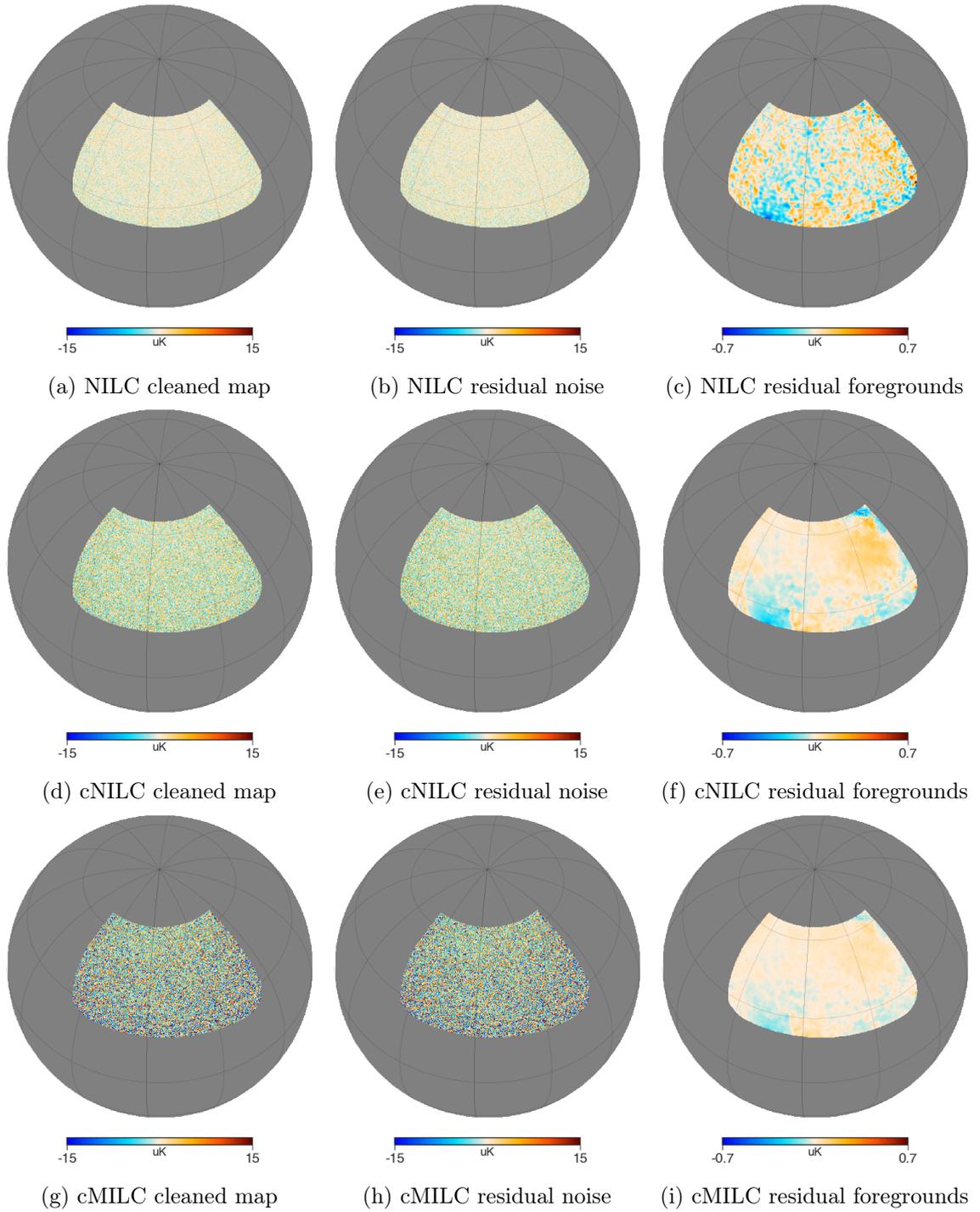
Figure 6: The cleaned CMB maps, the residual foreground maps and the residual noise maps. From top to bottom are results from NILC, cNILC and cMILC.

variation of the noise and the foregrounds, where the noise dominates at high latitude and the foregrounds dominate near the Galactic plane, thus being unfit for use over large sky patches. On the other hand, the ILC in pixel space does not take into account the harmonic
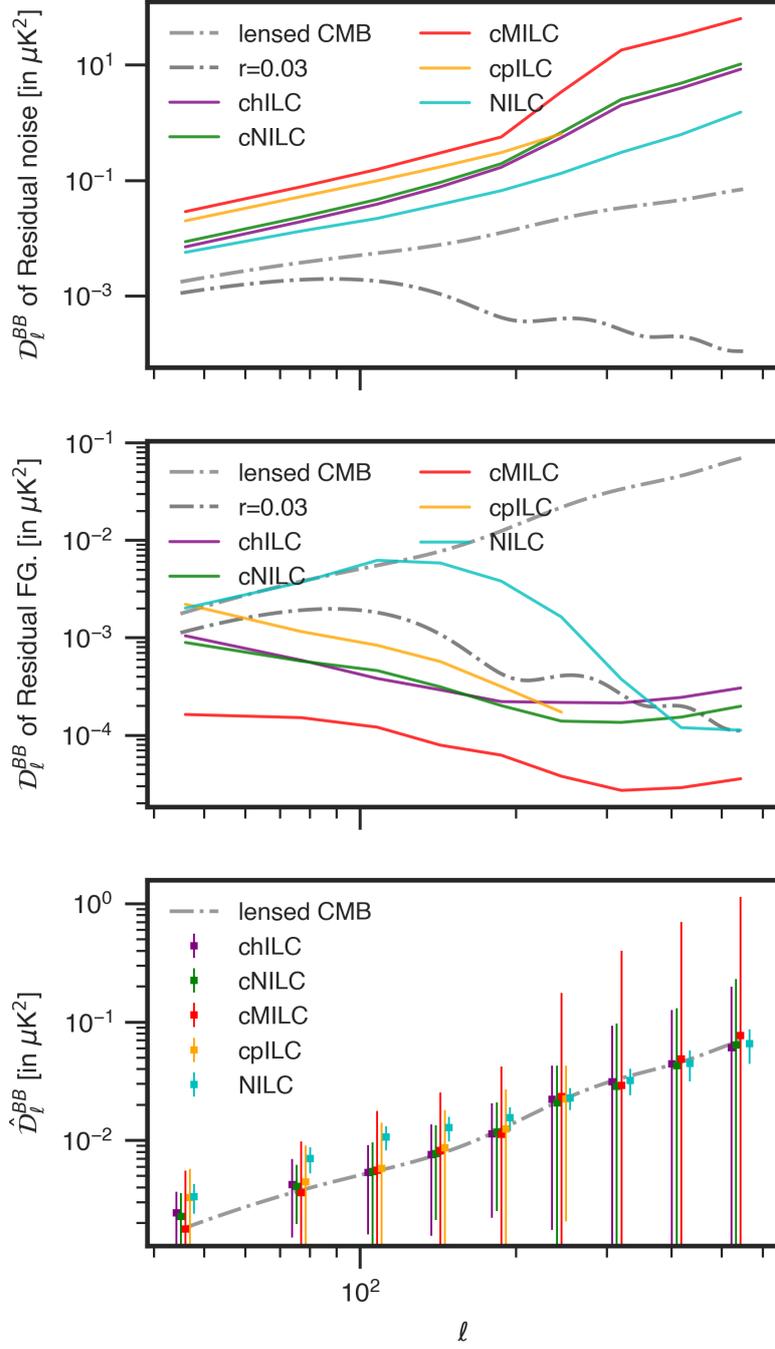
Figure 7: The comparison between the BB residual noise power spectra (upper), the BB residual foreground power spectra (middle) and the estimated CMB BB power spectra $\hat{\mathcal{D}}_\ell^{BB}$ (lower) with different methods. Different colors represent different foreground cleaning methods. The power spectrum of cpILC has a limit of $\ell_{\max} = 300$ as mentioned in Section 4.3.

variation where the noise dominates at small scales and the foregrounds dominate at large scales. Typically, the ILC in needlet space computes the optimal weights both across the sky

and the scales given the localization both in pixel space and in harmonic space.

In our sky patch, the foreground maps are nearly uniform, which narrows the advantage of localization in both pixel space and favours a fine division in harmonic space. Therefore, the chILC, which assigns weights at each multipole, performs nearly equally well with the cNILC. The cNILC would be expected more advantaged than chILC in cleaning a patch of spatially-varying foregrounds. As for different foreground models in Figure 10, we find that the cNILC does render lower residual foregrounds but a slightly higher noise level in exchange.

Both chILC and cNILC reduces the foregrounds to a low level, which is already low enough. Hence adding yet more constraints with the disadvantage of increasing the noise level might not be a wise choice. For this reason, we will not adopt cMILC in the following tests in view of its large uncertainty of power spectrum estimation. Furthermore, neither NILC nor cpILC will be discussed in the following analysis since they are not able to clean the foregrounds substantially.

5.3   Bias analysis

In this section, we compute analytical estimates of the potential biases of cILC other than the foreground and noise residuals and compute its level in practice. The cILC cleaned map can be decomposed into three components, the CMB signal, the residual foreground and the residual noise:

$$\hat{s}_{\ell m}^{\mathrm{cILC}} = s_{\ell m} + \boldsymbol{w}_\ell^T \boldsymbol{a}_{\ell m}^{\mathrm{fg}} + \boldsymbol{w}_\ell^T \boldsymbol{n}_{\ell m} \,. \tag{5.1}$$

where the residual noise is the dominant component, while the residual foreground level is of the order $r \sim 0.01$ for the tensor spectrum at the large scales. Therefore, it is safe to omit the correlations between the residual foreground and other emission components. Then the power spectrum of the cILC cleaned map is given by:

$$\hat{C}_\ell^{\mathrm{cILC}} = C_\ell^{\mathrm{lens}} + r C_\ell^{r=1} + C_\ell^{\mathrm{res.fg}} + N_\ell^{\mathrm{res.}} + 2 C_\ell^{\mathrm{s \times res.n}} \,, \tag{5.2}$$

where the chance correlation between CMB and the residual noise $C_\ell^{\mathrm{s \times res.n}}$, also denoted as the 'ILC bias' [27], can be restricted to a negligible level given enough modes to sample the covariance. The residual noise $N_\ell^{\mathrm{res.}}$ is debiased by the average of 300 noise realizations weighted by the same cILC weights, denoted by $\bar{N}_\ell^{\mathrm{res.}}$. The lensing term C_ℓ^{lens} is assumed to be known in this work, with the lensing noise to be studied in the future.

Averaged over realizations, the residual bias of the noise-debiased cILC power spectrum is given by:

$$R_\ell \equiv \langle \hat{C}_\ell^{\mathrm{cILC}} - \bar{N}_\ell^{\mathrm{res.}} - C_\ell^{\mathrm{s}} \rangle = \langle C_\ell^{\mathrm{res.fg}} \rangle + \langle N_\ell^{\mathrm{res.}} - \bar{N}_\ell^{\mathrm{res.}} \rangle + 2 \langle C_\ell^{\mathrm{s \times res.n}} \rangle \,, \tag{5.3}$$

where $C_\ell^{\mathrm{s}} = C_\ell^{\mathrm{lens}} + r C_\ell^{r=1}$ is the input CMB power spectrum and the brakets denote the ensemble averaging. More specifically,

$$R_\ell = \langle C_\ell^{\mathrm{res.fg}} \rangle + \langle \boldsymbol{w}_\ell^T (\boldsymbol{N}_\ell - \bar{\boldsymbol{N}}_\ell) \boldsymbol{w}_\ell \rangle + \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle s_{\ell m}^* \boldsymbol{w}_\ell^T \boldsymbol{n}_{\ell m} + s_{\ell m} \boldsymbol{w}_\ell^T \boldsymbol{n}_{\ell m}^* \rangle \,, \tag{5.4}$$

where $\boldsymbol{N}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \boldsymbol{n}_{\ell m} \boldsymbol{n}_{\ell m}^\dagger$. The foreground residual can be easily obtained by projecting the input foregrounds on the weights. The second term, named the noise bias error
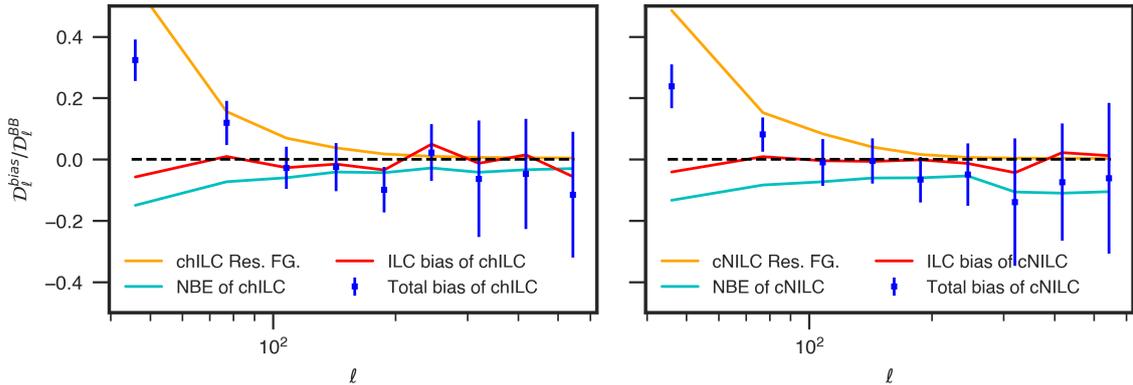
Figure 8: Ratios of biases from different sources, with the left panel for chILC and the right panel for cNILC. Each panel shows the total bias (blue) with the error bar of $\sigma(\hat{C}_\ell)/\sqrt{N_{\rm sim}}$ and its components: the residual foregrounds (orange), the ILC bias (red) and the noise bias error (cyan).

(NBE), is fully discussed in Section 5.3.1. The third term, i.e. the ILC bias, is described theoretically by the following formula:

$$2\langle C_\ell^{\rm s\times res.n}\rangle = \frac{2(n_c - n_\nu)}{N_k f_{\rm sky}}\langle C_\ell^{\rm s}\rangle\,, \tag{5.5}$$

where N_k is the number of coefficients in the needlet (harmonic) domain as computing the covariance matrix, $f_{\rm sky} \approx 5.9\%$[4] is the sky fraction, $n_c = 3$ represents the number of constraints, and $n_\nu = 7$ is the number of involved channels. In practice, we exclude the ℓ (n_{jk}) mode from the empirical covariance matrix used to compute the weights of ℓ (n_{jk}), i.e. $\ell' \neq \ell$ in Eq.(2.4)($k' \neq k$ in Eq.(2.11)), which is commonly used to reduce the ILC bias by avoiding the correlation between the ILC weights and the data sets which the weights are applied on. The actual ratios between the mean ILC bias and the input CMB power spectrum $2\langle C_\ell^{\rm s\times res.n}\rangle/\langle C_\ell^{\rm s}\rangle$ are plotted in Figure 8. As seen in the plot, the actual ILC bias, mostly negative, is of order 1% of the CMB power spectrum. The deviation of the actual ILC bias from its negative expectation (Eq.(5.5)) could be attributed to its large uncertainty compared to its small amplitude.

The mean biases of the residual foreground and the noise bias error are also obtained from simulations. We compute their ratios to the input CMB power spectrum, as shown in the right panel of Figure 8. We also plot the total biases $\langle \hat{C}_\ell - C_\ell^s\rangle/C_\ell^s$, whose error bars are given by the uncertainty of the mean residual bias $\sigma(\hat{C}_\ell)/\sqrt{N_{\rm sim}} = \sigma(\hat{C}_\ell)/10$. As seen in the plot, the biases for chILC and cNILC are mainly from NBE at small scales of $\ell>100$, and from residual foregrounds at large scales. Since the NBE and the ILC bias are both within the uncertainty of the estimated noise bias at all scales, we do not consider it necessary to debias them.

---

[4]The effective sky fraction of the inverse noise variance weighted mask is given by $f_{\rm sky}^{\rm eff} = \langle M\rangle^2/\langle M^2\rangle$ where M is the weight of the mask.

5.3.1 Noise bias error

In our analysis of the bias, we find that the estimated noise bias is systematically lower than the actual noise bias (obtained from the actual noise projected by the cILC weights) due to the fact that the cILC weights are conceptually designed to minimize the power spectrum of the actual data which is dominated by the actual noise. The error, however, is rarely mentioned by the previous works of ILC either because they use the cross spectra of data splits to cancel the noise bias or their simulations are not dominated by the noise. To make it clear, the third term of Eq.(5.1) is the residual noise, whose power spectrum is dubbed the noise bias $N_\ell^{\text{res.}} = \boldsymbol{w}_\ell^T \hat{\boldsymbol{N}}_\ell \boldsymbol{w}_\ell$, while the estimated noise bias averaged over noise simulations is $\bar{N}_\ell^{\text{res.}} = \boldsymbol{w}_\ell^T \bar{\boldsymbol{N}}_\ell \boldsymbol{w}_\ell$. Since the noise covariance dominates the total covariance matrix, as the weights (dependent on the realization) varying, the minimum point of $\boldsymbol{w}_\ell^T \hat{\boldsymbol{C}}_\ell \boldsymbol{w}_\ell$ is approximately the minimum point of $\boldsymbol{w}_\ell^T \hat{\boldsymbol{N}}_\ell \boldsymbol{w}_\ell$; thus the cILC weights also nearly minimizes the noise bias of the particular noise realization, i.e. $\boldsymbol{w}_\ell^T \hat{\boldsymbol{N}}_\ell \boldsymbol{w}_\ell \leq \boldsymbol{w}_\ell^T \bar{\boldsymbol{N}}_\ell \boldsymbol{w}_\ell$ for almost any noise realization. The difference between the actual and estimated noise bias, named the noise bias error, is approximated (See Appendix B for the derivation) as:

$$\langle N_\ell^{\text{res.}} - \bar{N}_\ell^{\text{res.}} \rangle = \langle \boldsymbol{w}_\ell^T (\hat{\boldsymbol{N}}_\ell - \bar{\boldsymbol{N}}_\ell) \boldsymbol{w}_\ell \rangle \approx \frac{2(n_c - n_\nu)}{N_k f_{\text{sky}}} \bar{N}_\ell^{\text{res.}} . \tag{5.6}$$

In our application, $n_c = 3$, $n_\nu = 7$ at $\ell < 300$ or 6 otherwise, the effective sky fraction of the UNP-inv mask $f_{\text{sky}} \approx 5.9\%$ and N_k is the number of coefficients involved when computing the covariance matrix in an cILC implementation, e.g., $N_k \sim 14000$ at $\ell \sim 100$ for the chILC. Typically we use larger N_k to estimate the covariance matrix at smaller scales.

We compute the noise bias estimated from 100 noise simulations and the actual noise power spectrum for the chILC using both different binning sizes of the covariance matrix and different noise levels. In chILC, the covariance matrix $\hat{\boldsymbol{C}}_\ell$ is computed by averaging the covariance over harmonic modes in $\ell \pm \Delta\ell/2$. The noise bias error and its theoretical expectations for chILC are plotted in Figure 9. Being close to the theoretical results, the noise bias error reduces with the binning size (or N_k) increasing.

As seen in the lower plot, the ratio between the noise bias error and the mean residual noise does not vary with the noise level, as expected from Equation 5.6. The noise bias error is about 0.8% of the mean residual noise at $\ell \sim 100$, which is negligible compared to the uncertainty of the noise bias $\sigma(N_\ell^{\text{res.}})/\sqrt{N_{\text{sim}}} \approx \sqrt{\frac{1}{(2\ell+1)\Delta\ell_b f_{\text{sky}} N_{\text{sim}}}} \bar{N}_\ell^{\text{res.}}$ where $\Delta\ell_b$ is the binning size in the NaMASTER power spectrum estimation and N_{sim} is the number of noise simulations used to debias the noise.

We investigate possible ways to mitigate the ILC noise bias error. Theoretically, either fixing the weights independent of the realizations, or crossing two different data splits to null the average noise bias should eliminate this kind of error. Here we use the Jackknife method, where the input maps are two data splits with the same CMB and foregrounds but different Gaussian instrumental noises. We first obtain the cILC weights from the average of two data splits, then apply the weights onto two splits individually and compute their cross power spectrum to get rid of the noise bias. The results shown in Appendix C have verified that the Jackknife would not produce the noise bias error.

5.4 Other foreground models

We apply chILC and cNILC on four sets of simulations with four types of foregrounds introduced in Section 3.4 to test the robustness of the mixing matrix. As seen in Figure 10, both
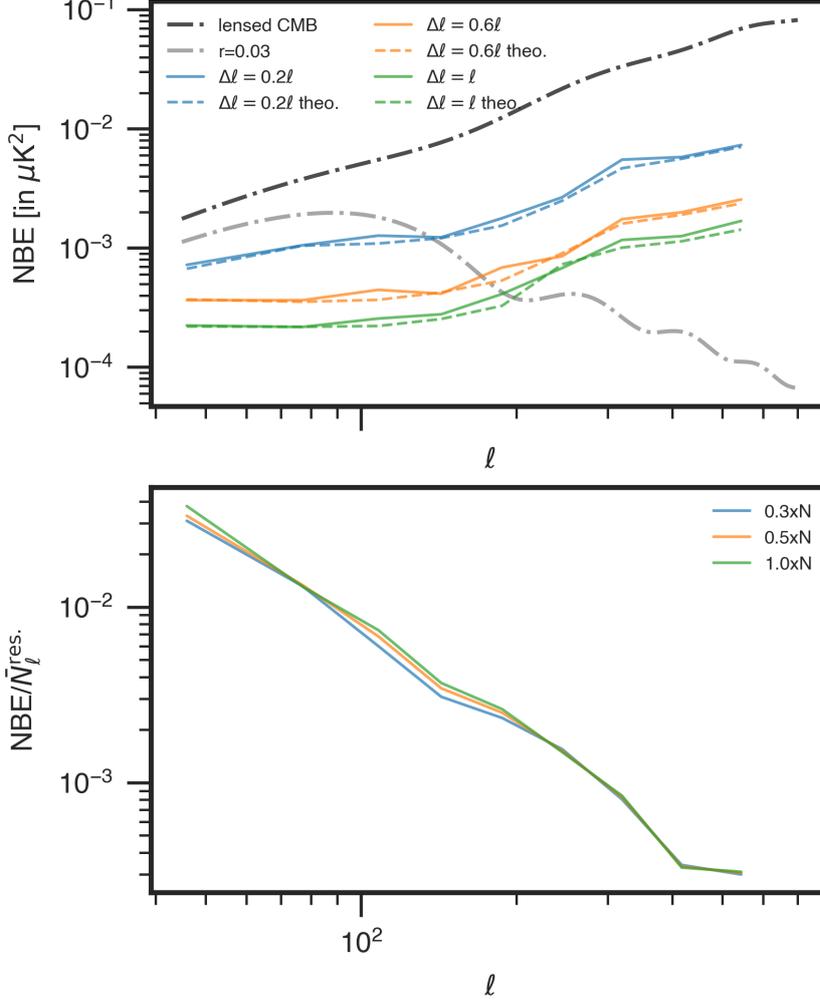
Figure 9: Upper: The noise bias errors from simulations (solid) compared to their theoretical expectations (dashed) in chILC. Lower: The ratios between the noise bias error and the mean residual noise under different noise levels (0.3, 0.5 and 1.0 times the real noise level) in chILC.

methods are able to reduce all types of foregrounds effectively. The noise residual hardly varies among models since the data share the same noise realizations, while the foreground residual depends on the particular foreground model. The new PySM synchrotron model `s5` has more power than the older models `s0` and `s1`. The PySM dust model `d10` has more decorrelation than either `d0` or `d1` dust models. The larger residual foregrounds for `d10s5` compared to `d0s0` and `d1s1` likely originates in the complexity of the `d10s5` model.

5.5   Estimation of the tensor-to-scalar ratio

We estimate the tensor-to-scalar ratio r from the mean measured power spectrum after foreground cleaning using chILC and cNILC on our simulated data with four different foreground models. The results are summerized in Figure 11. The best-fit r values and the corresponding 1σ error bars are shown in red, with the input value $r = 0.03$. We can see that the true value $r = 0.03$ is within 1σ error bars for all cases. For AliCPT foregrounds we obtain $r = 0.037 \pm 0.018$ ($r = 0.036 \pm 0.018$) for chILC (cNILC). For the null tests with the input
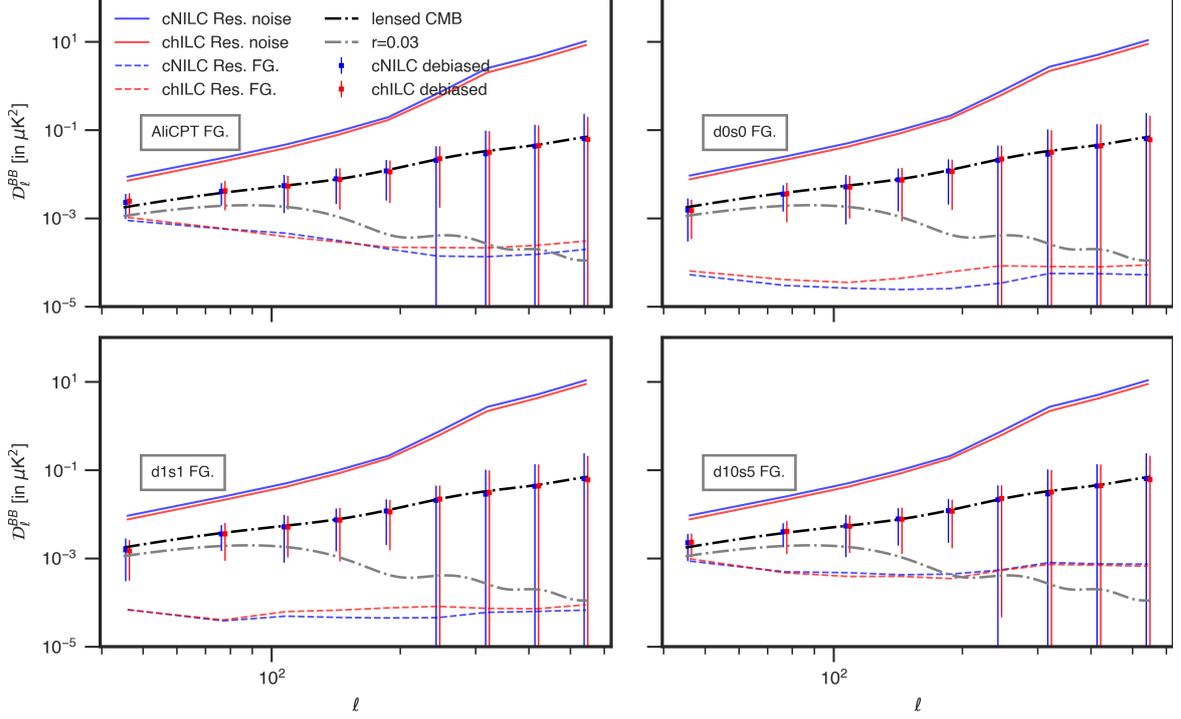
Figure 10: Results of chILC (red) and cNILC (blue) debiased power spectra and the residual noise and foreground spectra for four different foreground models. The noise residuals are plotted as solid curves and the foreground residuals are plotted as dashed curves.

Table 3: Upper: the best-fit r and its uncertainties for the mean cILC cleaned power spectrum with the input $r = 0.03$. Lower: the 2σ constraints (95%CL) on r for the null test without the tensor B modes.

$r_{\mathrm{in}} = 0.03$	chILC	cNILC
AliCPT	$r = 0.037 \pm 0.018$	$r = 0.036 \pm 0.018$
d0s0	$r = 0.027 \pm 0.016$	$r = 0.028 \pm 0.017$
d1s1	$r = 0.027 \pm 0.016$	$r = 0.029 \pm 0.017$
d10s5	$r = 0.038 \pm 0.018$	$r = 0.038 \pm 0.018$

Null Test($r_{\mathrm{in}} = 0$)	chILC	cNILC
AliCPT	$r < 0.043$	$r < 0.043$
d0s0	$r < 0.033$	$r < 0.034$
d1s1	$r < 0.034$	$r < 0.034$
d10s5	$r < 0.044$	$r < 0.044$

of $r = 0$, the chILC results of best-fit r and 2σ upper limits are shown in green arrows and bars respectively, which are similar to the cNILC results. The results are also summarized in Table 3. In the null tests of the AliCPT foregrounds, the 95% CL upper limit of r is 0.043 for both chILC and cNILC. With the simpler foreground models like d0s0 and d1s1, the constraint could reach $r < 0.033$ due to the lower foreground bias.
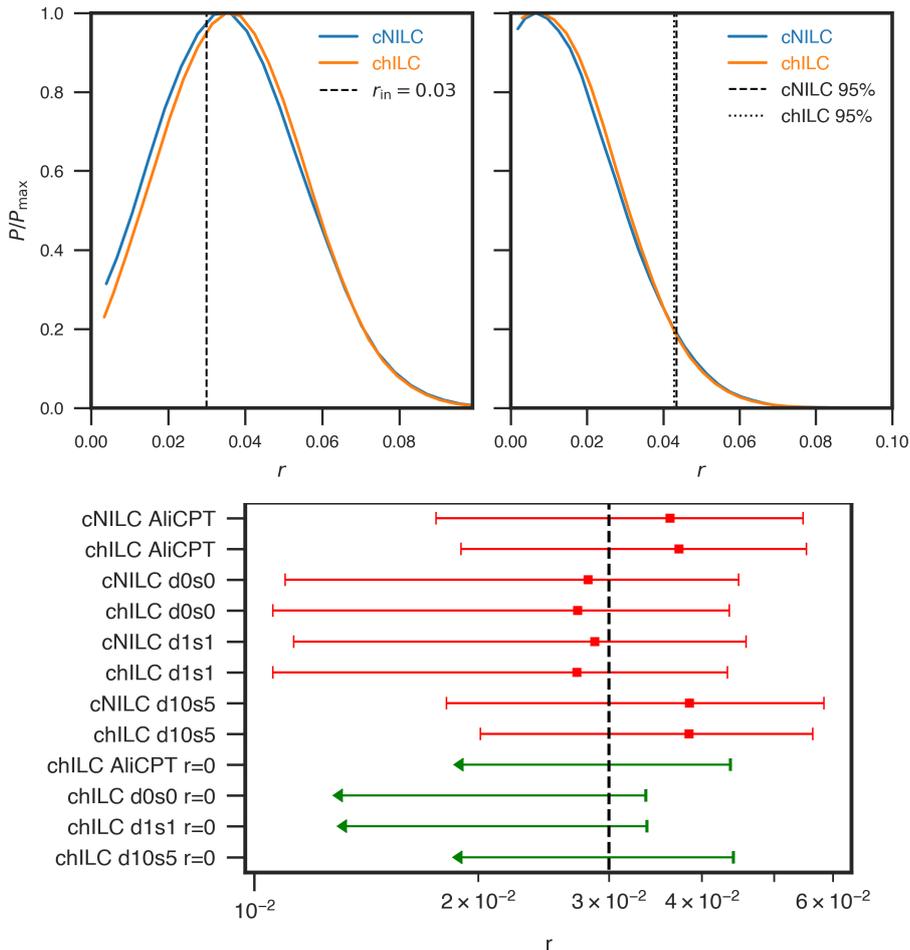
Figure 11: Left upper: The posterior distribution of r for both chILC (orange) and cNILC (blue) tests with the input value $r = 0.03$ and AliCPT foregrounds. Right upper: The posterior distribution of r for both chILC and cNILC null tests with the input $r = 0$ and AliCPT foregrounds. Lower: The best-fit tensor-to-scalar ratio r of chILC and cNILC results with the four models considered here. For the input $r = 0.03$, the best-fit values for r are shown in red squares with 1σ error-bars. For the null tests with input $r = 0$, the best-fit r values are shown in green arrows with 2σ upper limits in gree bars, and only the chILC results are shown since they are close to the cNILC results.

6 Conclusions

The precise measurement of the CMB B-mode power spectrum as well as the tensor-to-scalar ratio demands an elaborate component separation method in order to balance the minimization of the statistical uncertainty and the systematic bias. Without many priors on the Galactic foregrounds, the ILC method and its variants are competent candidates for the future ground-based CMB experiments.

We evaluate the foreground cleaning performance of some variants of the ILC method on partial sky B-modes. We carry out NILC, cMILC, and the cILC methods in different domains (cpILC, chILC, cNILC) on a northern sky patch using the AliCPT, WMAP-K and Planck

HFI combined mock data. We mainly focus on the cILC methods in harmonic (chILC) and needlet (cNILC) domain since the NILC results show larger foreground bias. According to our results, there are advantages in different aspects when we use both chILC and cNILC.

In terms of the power spectrum, the cILC in harmonic space renders slightly less noise residuals and smaller uncertainties on estimation of the power spectrum and the tensor-to-scalar ratio. The cNILC method renders less foreground residuals, thus reducing the systematic errors. The results of both methods are similar due to the relatively low and uniform foreground contamination in our sky patch. On the other hand, the foreground cleaning performance of the cILC in pixel space is poor due to the properties of the data. Moreover, using cMILC in the AliCPT patch results in high noise uncertainty, despite lowest foreground residuals.

Additionally, doing ILC in harmonic space is about two times computationally faster than that in needlet space, since the size of the covariance matrix for harmonic ILC is smaller on average. Applying the NILC costs plenty of computational hours on needlet transformation. Our results using different foregrounds also demonstrate that our mixing matrix is compatible to various foreground models.

Analysis of main biases on power spectra shows that other than the foreground residual and the ILC bias, the estimated noise bias cannot accurately counteract the actual noise bias given the large noise level and the finite size of the domain where we compute the covariance. Fortunately, this error (named the noise bias error) could be limited to a negligible order compared to the uncertainty of the estimated power spectrum. We firstly derive its relation with respect to the noise residual which fits well with the results from simulations. We also illustrate that the Jackknife approach crossing two data splits would not produce this error.

We estimate the tensor-to-scalar ratio r for our estimated power spectra, concluding that the methods bias r in about 0.008 at maximum for all cases with an uncertainty no larger than 0.018. For the null tests without primordial B-modes, we obtain a constraint of $r < 0.043$ (95%CL) applying cNILC/chILC on AliCPT-1 mock data, and $r < 0.033$ (95%CL) using the simpler d0s0 sky model. Note that since currently we have not yet involved all possible systematics, the constraints might be underestimated and could not be directly compared to the best limit $r < 0.034$ (95%CL) of pure CMB experiments from combined Planck PR4 + BK18 data, or $r < 0.032$ (95%CL) with BAO data involved further [12]. Nevertheless, the future observations with lower noise would be promising to further constrain the results.

A   Effects of point sources

Here, we add a simulated polarized point source map to the data sets. The input point sources are modeled based on the radio source catalogue of the Planck Sky Model [51]. These point sources are masked by a point source mask produced by the matched filter and inpainted before the cILC implementation. Since the point sources are independent of other components, we compute the residual point source power spectra from cNILC and chILC methods by projecting the input point source maps on the cILC weights, as shown in Figure 12. It can be seen that the contribution of residual point sources is negligible compared to the bias of residual foregrounds.

We preprocess the input maps to diminish the effects of point sources. First, we identify the point sources on the HFI 100GHz temperature map using the outlier method [62] and obtain the point source mask using the matched filter with a threshold of 3σ. Then we mask the QU maps with the point source mask and use the template cleaning method to obtain
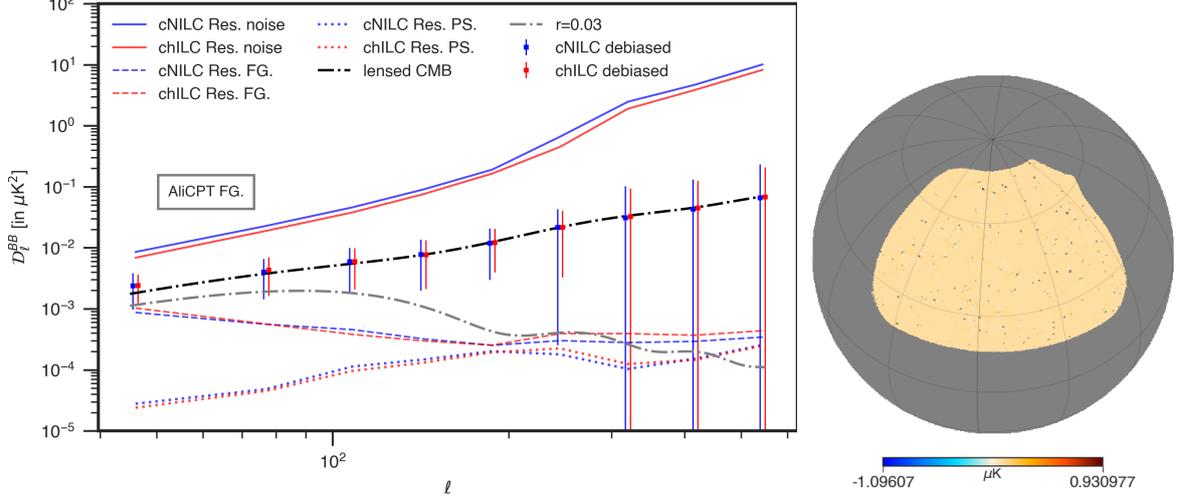
Figure 12: Testing chILC and cNILC methods considering the point sources. Left: the debiased power spectra and the noise, foreground and point source residuals. Right: the masked point source B-mode map at 150GHz after template cleaning.

the B-mode maps with the point sources masked. Finally, we inpaint the B maps using the `iSAP` [63] package. The procedure after point source preprocessing is the same with the basic B-mode reconstruction pipeline.

B   Derivation of the noise bias error

In this appendix we compute the analytical form of the noise bias error, which must be carefully accounted for under cases of low signal-to-noise ratio. The derivation is in harmonic space while the result is general in all domains.

B.1   The noise bias error in standard ILC

The input data is modelled as $\boldsymbol{d}_{\ell m} = \boldsymbol{a} s_{\ell m} + \boldsymbol{f}_{\ell m} + \boldsymbol{n}_{\ell m}$ where $s_{\ell m}$ is the CMB signal, $\boldsymbol{f}_{\ell m}$ the foregrounds and $\boldsymbol{n}_{\ell m}$ the noise term. Ignoring the binning scheme till the final form of the noise bias error, the estimated covariance matrix (in harmonic ILC) at ℓ is given by:

$$\hat{\boldsymbol{C}}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \boldsymbol{d}_{\ell m} \boldsymbol{d}_{\ell m}^\dagger = \hat{C}_\ell^s \boldsymbol{a}\boldsymbol{a}^T + \hat{\boldsymbol{N}}_\ell + \hat{\boldsymbol{F}}_\ell + \boldsymbol{\Lambda}_\ell \,, \qquad (\text{B.1})$$

where

$$\hat{C}_\ell^s = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} s_{\ell m} s_{\ell m}^* \,, \qquad (\text{B.2})$$

$$\hat{\boldsymbol{N}}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \boldsymbol{n}_{\ell m} \boldsymbol{n}_{\ell m}^\dagger \,, \qquad (\text{B.3})$$

$$\hat{\boldsymbol{F}}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \boldsymbol{f}_{\ell m} \boldsymbol{f}_{\ell m}^\dagger \,, \qquad (\text{B.4})$$

$$\mathbf{\Lambda}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \left(s_{\ell m} \boldsymbol{a} \boldsymbol{n}_{\ell m}^\dagger + s_{\ell m}^* \boldsymbol{n}_{\ell m} \boldsymbol{a}^T \right). \tag{B.5}$$

We ignore the cross terms between the foregrounds and the CMB or noise. Hereafter, the 'hat' over some quantity denotes the empirical estimation from one realization while the 'bar' denotes the ensemble average. Then the average terms

$$\bar{\boldsymbol{C}}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle \boldsymbol{d}_{\ell m} \boldsymbol{d}_{\ell m}^\dagger \rangle = \bar{C}_\ell^s \boldsymbol{a} \boldsymbol{a}^T + \bar{\boldsymbol{N}}_\ell + \bar{\boldsymbol{F}}_\ell\,, \tag{B.6}$$

$$\bar{C}_\ell^s = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle s_{\ell m} s_{\ell m}^* \rangle\,, \tag{B.7}$$

$$\bar{\boldsymbol{N}}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle \boldsymbol{n}_{\ell m} \boldsymbol{n}_{\ell m}^\dagger \rangle\,, \tag{B.8}$$

$$\bar{\boldsymbol{F}}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle \boldsymbol{f}_{\ell m} \boldsymbol{f}_{\ell m}^\dagger \rangle\,, \tag{B.9}$$

where $\bar{\mathbf{\Lambda}}_\ell = 0$ assuming the CMB is not correlated to the noise.

The (harmonic) ILC weights

$$\hat{\boldsymbol{w}}^T = \frac{\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1}}{\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} \tag{B.10}$$

are applied on the ℓ-mode input data sets.

Given the residual noise power spectrum $N_\ell^{\text{res.}} = \hat{\boldsymbol{w}}^T \hat{\boldsymbol{N}}_\ell \hat{\boldsymbol{w}}$ and the estimated noise bias $\bar{N}_\ell^{\text{res.}} = \hat{\boldsymbol{w}}^T \bar{\boldsymbol{N}}_\ell \hat{\boldsymbol{w}}$, the noise bias error is defined as the average of their difference:

$$\langle N_\ell^{\text{res.}} - \bar{N}_\ell^{\text{res.}} \rangle = \langle \hat{\boldsymbol{w}}^T (\hat{\boldsymbol{N}}_\ell - \bar{\boldsymbol{N}}_\ell) \hat{\boldsymbol{w}} \rangle\,. \tag{B.11}$$

Assuming the noise realizations are independent among channels:

$$\bar{N}_{\ell,ij} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle n_{\ell m,i} n_{\ell m,j}^* \rangle = \delta_{ij} \sigma_{\ell,i}^2\,, \tag{B.12}$$

where $\sigma_{\ell,i}^2$ is the mean noise power spectrum of the channel i.

Defining $\mathbf{\Omega}_\ell = \hat{\boldsymbol{N}}_\ell - \bar{\boldsymbol{N}}_\ell$, the cosmic variance of the Gaussian white noise is given by:

$$\langle \Omega_{\ell,ij} \Omega_{\ell,i'j'} \rangle = \frac{1}{(2\ell+1)^2} \sum_{mm'} \left(\langle n_{\ell m,i} n_{\ell m,j}^* n_{\ell m',i'} n_{\ell m',j'}^* \rangle - \langle n_{\ell m,i} n_{\ell m,j}^* \rangle \langle n_{\ell m',i'} n_{\ell m',j'}^* \rangle \right)$$
$$= \frac{\sigma_{\ell,i}^2 \sigma_{\ell,j}^2}{2\ell+1} \left(\delta_{ii'} \delta_{jj'} + \delta_{ij'} \delta_{ji'} \right). \tag{B.13}$$

Defining $\mathbf{\Delta}_\ell = \hat{\boldsymbol{C}}_\ell - \bar{\boldsymbol{C}}_\ell$, from Eq.(B.1) and Eq.(B.6), we can write

$$\mathbf{\Delta}_\ell = (\hat{C}_\ell^s - \bar{C}_\ell^s) \boldsymbol{a} \boldsymbol{a}^T + \mathbf{\Omega}_\ell + (\hat{\boldsymbol{F}}_\ell - \bar{\boldsymbol{F}}_\ell) + \mathbf{\Lambda}_\ell\,. \tag{B.14}$$

When we consider the correlation between $\boldsymbol{\Delta}_\ell$ and $\boldsymbol{\Omega}_\ell$, only the second term of Eq.(B.14) gives rise to the correlation. The first and the third correlation term vanish since the noise and the CMB or foregrounds are uncorrelated, while the last correlation term vanishes since it contains the third power of $\boldsymbol{n}_{\ell m}$ and the first power of $s_{\ell m}$. Therefore, we have

$$\langle \Delta_{\ell,ij} \Omega_{\ell,i'j'} \rangle = \langle \Omega_{\ell,ij} \Omega_{\ell,i'j'} \rangle = \frac{\sigma_{\ell,i}^2 \sigma_{\ell,j}^2}{2\ell+1} (\delta_{ii'} \delta_{jj'} + \delta_{ij'} \delta_{ji'}) \,. \tag{B.15}$$

Inserting the ILC weights to the noise bias error:

$$\langle N_\ell^{\text{res.}} - \bar{N}_\ell^{\text{res.}} \rangle = \langle \frac{\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1}}{\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} (\hat{\boldsymbol{N}}_\ell - \bar{\boldsymbol{N}}_\ell) \frac{\hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}}{\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} \rangle$$

$$= \langle \frac{\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Omega}_\ell \hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}}{(\boldsymbol{a}^T \hat{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^2} \rangle \,. \tag{B.16}$$

Since $\boldsymbol{\Delta}_\ell$ is a small perturbation to $\hat{\boldsymbol{C}}_\ell$, we use the first order expansion:

$$\hat{\boldsymbol{C}}_\ell^{-1} = [\bar{\boldsymbol{C}}_\ell + \boldsymbol{\Delta}_\ell]^{-1} \approx \bar{\boldsymbol{C}}_\ell^{-1} - \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \,. \tag{B.17}$$

Writing (Eq.(A.11) of [27]):

$$\frac{1}{\boldsymbol{a}^T [\bar{\boldsymbol{C}}_\ell^{-1} - \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1}] \boldsymbol{a}} = \frac{1}{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} \frac{1}{1-\epsilon} \approx \frac{1}{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} (1+\epsilon) \,, \tag{B.18}$$

where

$$\epsilon = \frac{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}}{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} \,. \tag{B.19}$$

The terms without $\boldsymbol{\Delta}_\ell$ vanish since $\langle \boldsymbol{\Omega}_\ell \rangle = 0$. Keeping up to the first order terms of $\boldsymbol{\Delta}_\ell$:

$$\langle N_\ell^{\text{res.}} - \bar{N}_\ell^{\text{res.}} \rangle \approx \frac{2}{(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^3} \langle (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Omega}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}) \rangle$$

$$- \frac{2}{(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^2} \langle (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Omega}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}) \rangle \,. \tag{B.20}$$

For the first term of Eq.(B.20)

$$\langle (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Omega}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}) \rangle$$

$$= \langle [\sum_{ijkp} (\bar{\boldsymbol{C}}_\ell^{-1})_{ij} \Omega_{\ell,jk} (\bar{\boldsymbol{C}}_\ell^{-1})_{kp}][\sum_{i'j'k'p'} (\bar{\boldsymbol{C}}_\ell^{-1})_{i'j'} \Delta_{\ell,j'k'} (\bar{\boldsymbol{C}}_\ell^{-1})_{k'p'}] \rangle$$

$$= \sum_{ijkp} (\bar{\boldsymbol{C}}_\ell^{-1})_{ij} (\bar{\boldsymbol{C}}_\ell^{-1})_{kp} \sum_{i'j'k'p'} (\bar{\boldsymbol{C}}_\ell^{-1})_{i'j'} (\bar{\boldsymbol{C}}_\ell^{-1})_{k'p'} \langle \Omega_{\ell,jk} \Delta_{\ell,j'k'} \rangle$$

$$= \frac{2}{2\ell+1} [\sum_{iji'} (\bar{\boldsymbol{C}}_\ell^{-1})_{ij} \sigma_{\ell,j}^2 (\bar{\boldsymbol{C}}_\ell^{-1})_{ji'}][\sum_{pkp'} (\bar{\boldsymbol{C}}_\ell^{-1})_{pk} \sigma_{\ell,k}^2 (\bar{\boldsymbol{C}}_\ell^{-1})_{kp'}]$$

$$= \frac{2}{2\ell+1} (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^2 \,. \tag{B.21}$$

Here we use the symmetry of $\bar{\boldsymbol{C}}_\ell$. For the second term of Eq.(B.20)

$$
\begin{aligned}
&\langle \boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Omega}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a} \rangle \\
&= \sum_{ijkpqr} (\bar{\boldsymbol{C}}_\ell^{-1})_{ij} (\bar{\boldsymbol{C}}_\ell^{-1})_{kp} (\bar{\boldsymbol{C}}_\ell^{-1})_{qr} \langle \Delta_{\ell,jk} \Omega_{\ell,qr} \rangle \\
&= \frac{1}{2\ell+1} \sum_{ijkr} (\bar{\boldsymbol{C}}_\ell^{-1})_{ij} \sigma_{\ell,j}^2 \sigma_{\ell,k}^2 [(\bar{\boldsymbol{C}}_\ell^{-1})_{jk}(\bar{\boldsymbol{C}}_\ell^{-1})_{kr} + (\bar{\boldsymbol{C}}_\ell^{-1})_{kk}(\bar{\boldsymbol{C}}_\ell^{-1})_{jr}] \\
&= \frac{1}{2\ell+1}[\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a} + \boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a} \, \mathrm{Tr}(\bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1})] \, .
\end{aligned}
\tag{B.22}
$$

Ignoring the foreground term we have $\bar{\boldsymbol{N}}_\ell = \bar{\boldsymbol{C}}_\ell - \bar{C}_\ell^s \boldsymbol{a} \boldsymbol{a}^T$. Using

$$
\begin{aligned}
\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a} &\approx \boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} (\bar{\boldsymbol{C}}_\ell - \bar{C}_\ell^s \boldsymbol{a} \boldsymbol{a}^T) \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a} \\
&= (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})[1 - \bar{C}_\ell^s (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})] \, ,
\end{aligned}
\tag{B.23}
$$

and

$$
\begin{aligned}
\mathrm{Tr}(\bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1}) &\approx \mathrm{Tr}[(\bar{\boldsymbol{C}}_\ell - \bar{C}_\ell^s \boldsymbol{a} \boldsymbol{a}^T) \bar{\boldsymbol{C}}_\ell^{-1}] \\
&= \mathrm{Tr}(\boldsymbol{I}) - \bar{C}_\ell^s \mathrm{Tr}(\boldsymbol{a} \boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1}) \\
&= n_\nu - \bar{C}_\ell^s (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}) \, ,
\end{aligned}
\tag{B.24}
$$

Eq.(B.22) turns out to be

$$
\begin{aligned}
&\langle \boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Delta}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{\Omega}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a} \rangle \\
&= \frac{1}{2\ell+1} (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})[n_\nu + 1 - 2\bar{C}_\ell^s (\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})] \, .
\end{aligned}
\tag{B.25}
$$

Bring Eq.(B.21) and Eq.(B.25) to Eq.(B.20):

$$
\begin{aligned}
\langle N_\ell^{\mathrm{res.}} - \bar{N}_\ell^{\mathrm{res.}} \rangle &\approx \frac{4}{2\ell+1} \frac{(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^2}{(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^3} - \frac{2(n_\nu+1)}{2\ell+1} \frac{(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})}{(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^2} \\
&+ \frac{4}{2\ell+1} \bar{C}_\ell^s \frac{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}}{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}} \, .
\end{aligned}
\tag{B.26}
$$

Given the mean residual noise $\bar{N}_\ell^{\mathrm{res.}} \approx \frac{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \bar{\boldsymbol{N}}_\ell \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}}{(\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a})^2}$ and the mean power spectrum of the ILC map $\bar{C}_\ell^{\mathrm{ILC}} = \langle \hat{\boldsymbol{w}}^T \hat{\boldsymbol{C}}_\ell \hat{\boldsymbol{w}} \rangle \approx \frac{1}{\boldsymbol{a}^T \bar{\boldsymbol{C}}_\ell^{-1} \boldsymbol{a}}$, replacing $\bar{C}_\ell^s$ with $(\bar{C}_\ell^{\mathrm{ILC}} - \bar{N}_\ell^{\mathrm{res.}})$, we get the final form of the noise bias error

$$
\langle N_\ell^{\mathrm{res.}} - \bar{N}_\ell^{\mathrm{res.}} \rangle \approx -\frac{2(n_\nu - 1)}{2\ell+1} \bar{N}_\ell^{\mathrm{res.}} \, .
\tag{B.27}
$$

Considering the binning and partial-sky effects, we get

$$
\langle N_\ell^{\mathrm{res.}} - \bar{N}_\ell^{\mathrm{res.}} \rangle \approx -\frac{2(n_\nu - 1)}{(2\ell+1)\Delta\ell_b f_{\mathrm{sky}}} \bar{N}_\ell^{\mathrm{res.}} \, ,
\tag{B.28}
$$

where $\Delta\ell_b$ is the binning size of the empirical covariance matrix and f_{sky} the sky fraction. The general form in all domains (harmonic, pixel and needlet) is

$$
\langle N_\ell^{\mathrm{res.}} - \bar{N}_\ell^{\mathrm{res.}} \rangle \approx -\frac{2(n_\nu - 1)}{N_k f_{\mathrm{sky}}} \bar{N}_\ell^{\mathrm{res.}} \, .
\tag{B.29}
$$

B.2 The noise bias error in cILC

The data is modelled as $\boldsymbol{d}_{\ell m} = \boldsymbol{A}\boldsymbol{s}_{\ell m} + \boldsymbol{n}_{\ell m}$ where $\boldsymbol{A}$ is the $n_\nu \times n_c$ mixing matrix and $\boldsymbol{s}_{\ell m}$ is the 3-dim vector of three sky components, the CMB, the thermal dust and the synchrotron. The cILC weights are given by:

$$\hat{\boldsymbol{w}}^T = \boldsymbol{e}^T(\boldsymbol{A}^T\hat{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\hat{\boldsymbol{C}}_\ell^{-1}\,, \tag{B.30}$$

where $\boldsymbol{e} = [1,0,0]^T$. The noise bias error is

$$\langle N_\ell^{\text{res.}} - \bar{N}_\ell^{\text{res.}}\rangle = \langle \boldsymbol{e}^T(\boldsymbol{A}^T\hat{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\hat{\boldsymbol{C}}_\ell^{-1}\boldsymbol{\Omega}_\ell\hat{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\hat{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e}\rangle\,. \tag{B.31}$$

Using Eq.(B.17),

$$\begin{aligned}
(\boldsymbol{A}^T\hat{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1} &\approx (\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A} - \boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{\Delta}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1} \\
&= (\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}(\boldsymbol{I} - \boldsymbol{\epsilon})^{-1} \\
&\approx (\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}(\boldsymbol{I} + \boldsymbol{\epsilon})\,,
\end{aligned} \tag{B.32}$$

where

$$\boldsymbol{\epsilon} = \boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{\Delta}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\,. \tag{B.33}$$

Keeping up to the first order terms of $\boldsymbol{\Delta}_\ell$, the noise bias error includes two terms, of which the first term:

$$\begin{aligned}
&2\langle \boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{\epsilon}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{\Omega}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e}\rangle \\
=&2\sum_{ijkp}(\boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1})_i\langle\Delta_{\ell,ij}\Omega_{\ell,kp}\rangle \\
&(\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1})_{jk}(\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e})_p \\
=&\frac{2}{2\ell+1}(\boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e} \\
&+ \boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e}\text{Tr}[\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}])\,.
\end{aligned} \tag{B.34}$$

Inserting $\bar{\boldsymbol{N}}_\ell = \bar{\boldsymbol{C}}_\ell - \boldsymbol{A}\bar{\boldsymbol{S}}_\ell\boldsymbol{A}^T$ where $\bar{\boldsymbol{S}}_\ell = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\langle\boldsymbol{s}_{\ell m}\boldsymbol{s}_{\ell m}^\dagger\rangle$, the first term of Eq.(B.34) becomes

$$\frac{2}{2\ell+1}[\boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e} - 2\boldsymbol{e}^T\bar{\boldsymbol{S}}_\ell\boldsymbol{e} + \boldsymbol{e}^T\bar{\boldsymbol{S}}_\ell(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})\bar{\boldsymbol{S}}_\ell\boldsymbol{e}]\,, \tag{B.35}$$

and the trace term of Eq.(B.34)

$$\text{Tr}[\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1} - \boldsymbol{A}\bar{\boldsymbol{S}}_\ell\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}] = n_c - \text{Tr}[\boldsymbol{A}\bar{\boldsymbol{S}}_\ell\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}]\,. \tag{B.36}$$

The second term of the noise bias error:

$$\begin{aligned}
&-2\langle \boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{\Delta}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{\Omega}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e}\rangle \\
=&-2\sum_{ijkp}(\boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1})_i\langle\Delta_{\ell,ij}\Omega_{\ell,kp}\rangle(\bar{\boldsymbol{C}}_\ell^{-1})_{jk}(\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e})_p \\
=&-\frac{2}{2\ell+1}(\boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e} \\
&+ \boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e}\text{Tr}[\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}])\,.
\end{aligned} \tag{B.37}$$

Inserting $\bar{\boldsymbol{N}}_\ell = \bar{\boldsymbol{C}}_\ell - \boldsymbol{A}\bar{\boldsymbol{S}}_\ell\boldsymbol{A}^T$, the first term of Eq.(B.37) becomes

$$-\frac{2}{2\ell+1}[\boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e} - 2\boldsymbol{e}^T\bar{\boldsymbol{S}}_\ell\boldsymbol{e} + \boldsymbol{e}^T\bar{\boldsymbol{S}}_\ell(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})\bar{\boldsymbol{S}}_\ell\boldsymbol{e}]\,, \tag{B.38}$$

and the trace term of Eq.(B.37)

$$\text{Tr}[\boldsymbol{I} - \boldsymbol{A}\bar{\boldsymbol{S}}_\ell\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}] = n_\nu - \text{Tr}[\boldsymbol{A}\bar{\boldsymbol{S}}_\ell\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}]\,. \tag{B.39}$$

Summing all the terms above, the noise bias error

$$\langle N_\ell^{\text{res.}} - \bar{N}_\ell^{\text{res.}}\rangle \approx \frac{2(n_c - n_\nu)}{2\ell+1}\boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e}\,. \tag{B.40}$$

Given the mean residual noise $\bar{N}_\ell^{\text{res.}} \approx \boldsymbol{e}^T(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\bar{\boldsymbol{N}}_\ell\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A}(\boldsymbol{A}^T\bar{\boldsymbol{C}}_\ell^{-1}\boldsymbol{A})^{-1}\boldsymbol{e}$, the general form of the noise bias error for cILC is

$$\langle N_\ell^{\text{res.}} - \bar{N}_\ell^{\text{res.}}\rangle \approx \frac{2(n_c - n_\nu)}{N_k f_{\text{sky}}}\bar{N}_\ell^{\text{res.}}\,. \tag{B.41}$$

A comparison between the theoretical amplitude and the actual error from simulations has been plotted in the lower panel of Figure 9, indicating that the above approximation is sensible.

C Verifying NBE of Jackknife

As discussed in section 5.3.1, we are about to verify that the Jackknife method would produce no noise bias error. First we add the same CMB and baseline foregrounds to different noise realizations ($\sqrt{2}$ of the original level) to generate the mock data splits. In the Jackknife pipeline, we apply the same cILC weights obtained from the combined data averaging two data splits onto two splits respectively, and finally compute the cross power spectrum between two cILC-cleaned split maps. The averaged residual noise power spectrum and its error bar over 100 simulations are plotted in Figure 13. Indeed, the mean noise power spectrum for Jackknife is consistent with zero at all scales (except for the largest one, where the multipole modes are not enough to manifest the expectation).

Figure 13: The mean residual noise cross spectrum for the Jackknife test. The plotted error bars are $\sigma(\bar{N}_\ell^{\mathrm{res.}}) = \sigma(N_\ell^{\mathrm{res.}})/10$, denoting the uncertainty of the noise bias averaged over 100 noise simulations.

References

[1] C.L. Bennett, A.J. Banday, K.M. Górski, G. Hinshaw, P. Jackson, P. Keegstra et al., *Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results*, *Astrophys. J. Lett.* **464** (1996) L1 [astro-ph/9601067].

[2] WMAP collaboration, *First-Year Wilkinson Microwave Anisotropy Probe WMAP Observations: Preliminary Maps and Basic Results*, *Astrophys. J. Suppl.* **148** (2003) 1 [astro-ph/0302207].

[3] Planck collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [1807.06209].

[4] A.A. Starobinsky, *The Perturbation Spectrum Evolving from a Nonsingular Initially De-Sitter Cosmology and the Microwave Background Anisotropy*, *Sov. Astron. Lett.* **9** (1983) 302.

[5] D. Baumann, M.G. Jackson, P. Adshead, A. Amblard, A. Ashoorioon, N. Bartolo et al., *CMBPol Mission Concept Study: Probing Inflation with CMB Polarization*, in *AIP Conf. Proc.*, vol. 1141, pp. 10–120, June, 2009, DOI [0811.3919].

[6] J. Martin, C. Ringeval and V. Vennin, *Encyclopaedia Inflationaris*, *Phys. Dark Univ.* **5–6** (2014) 75 [1303.3787].

[7] POLARBEAR collaboration, *A Measurement of the Degree Scale CMB B-mode Angular Power Spectrum with POLARBEAR*, *Astrophys. J.* **897** (2020) 55 [1910.02608].

[8] ACTPol collaboration, *The Atacama Cosmology Telescope: Two-Season ACTPol Spectra and Parameters*, *JCAP* **06** (2017) 031 [1610.02360].

[9] J.T. Sayre, C.L. Reichardt, J.W. Henning, P.A.R. Ade, A.J. Anderson, J.E. Austermann et al., *Measurements of B-mode Polarization of the Cosmic Microwave Background from 500 Square Degrees of SPTpol Data*, *Phys. Rev. D* **101** (2020) 122003 [1910.05748].

[10] B. Collaboration, P.A.R. Ade, Z. Ahmed, M. Amiri, D. Barkats, R.B. Thakur et al., *BICEP / Keck XIII: Improved Constraints on Primordial Gravitational Waves using Planck, WMAP,*

and BICEP/Keck Observations through the 2018 Observing Season, *Phys. Rev. Lett.* **127** (2021) 151301 [2110.00483].

[11] M. Zaldarriaga and U. Seljak, *Gravitational Lensing Effect on Cosmic Microwave Background Polarization*, *Phys. Rev. D* **58** (1998) 023003 [astro-ph/9803150].

[12] M. Tristram, A.J. Banday, K.M. Górski, R. Keskitalo, C.R. Lawrence, K.J. Andersen et al., *Improved limits on the tensor-to-scalar ratio using BICEP and Planck*, *Phys. Rev. D* **105** (2022) 083524 [2112.07961].

[13] T.S.O. Collaboration, P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali et al., *The Simons Observatory: Science goals and forecasts*, *J. Cosmol. Astropart. Phys.* **2019** (2019) 056 [1808.07445].

[14] A. Mennella, P. Ade, G. Amico, D. Auguste, J. Aumont, S. Banfi et al., *QUBIC: Exploring the Primordial Universe with the Q&U Bolometric Interferometer*, *Universe* **5** (2019) 42.

[15] H. Li, S.-Y. Li, Y. Liu, Y.-P. Li, Y. Cai, M. Li et al., *Probing Primordial Gravitational Waves: Ali CMB Polarization Telescope*, *Natl. Sci. Rev.* **6** (2019) 145 [1710.03047].

[16] M. Hazumi, P.A.R. Ade, Y. Akiba, D. Alonso, K. Arnold, J. Aumont et al., *LiteBIRD: A Satellite for the Studies of B-Mode Polarization and Inflation from Cosmic Background Radiation Detection*, *J Low Temp Phys* **194** (2019) 443.

[17] T.C.-S. Collaboration, K. Abazajian, G.E. Addison, P. Adshead, Z. Ahmed, D. Akerib et al., *CMB-S4: Forecasting Constraints on Primordial Gravitational Waves*, *ApJ* **926** (2022) 54 [2008.12619].

[18] H.K. Eriksen, A.J. Banday, K.M. Gorski and P.B. Lilje, *On Foreground Removal from the Wilkinson Microwave Anisotropy Probe Data by an Internal Linear Combination Method: Limitations and Implications*, *ApJ* **612** (2004) 633 [astro-ph/0403098].

[19] M. Tegmark, A. de Oliveira-Costa and A. Hamilton, *A high resolution foreground cleaned CMB map from WMAP*, *Phys. Rev. D* **68** (2003) 123523 [astro-ph/0302496].

[20] D. Maino, A. Farusi, C. Baccigalupi, F. Perrotta, A.J. Banday, L. Bedini et al., *All-sky astrophysical component separation with Fast Independent Component Analysis (FastICA)*, *Mon. Not. R. Astron. Soc.* **334** (2002) 53 [astro-ph/0108362].

[21] J. Delabrouille, J.-F. Cardoso and G. Patanchon, *Multi-Detector Multi-Component spectral matching and applications for CMB data analysis*, *Mon. Not. R. Astron. Soc.* **346** (2003) 1089 [astro-ph/0211504].

[22] J.-F. Cardoso, M. Martin, J. Delabrouille, M. Betoule and G. Patanchon, *Component separation with flexible models. Application to the separation of astrophysical emissions*, *arXiv:0803.1814 [astro-ph]* (2008) [0803.1814].

[23] H.K. Eriksen, J.B. Jewell, C. Dickinson, A.J. Banday, K.M. Gorski and C.R. Lawrence, *Joint Bayesian component separation and CMB power spectrum estimation*, *ApJ* **676** (2008) 10 [0709.1058].

[24] R. Stompor, S. Leach, F. Stivoli and C. Baccigalupi, *Maximum Likelihood algorithm for parametric component separation in CMB experiments*, *Mon. Not. Roy. Astron. Soc.* **392** (2009) 216 [0804.2645].

[25] J. Errard, F. Stivoli and R. Stompor, *Framework for performance forecasting and optimization of CMB B-mode observations in presence of astrophysical foregrounds*, *Phys. Rev. D* **84** (2011) 063005 [1105.3859].

[26] R. Fernández-Cobos, P. Vielva, R.B. Barreiro and E. Martínez-González, *Multi-resolution internal template cleaning: An application to the Wilkinson Microwave Anisotropy Probe 7-yr polarization data*, *Mon. Not. Roy. Astron. Soc.* **420** (2012) 2162 [1106.2016].

[27] J. Delabrouille, J.-F. Cardoso, M.L. Jeune, M. Betoule, G. Fay and F. Guilloux, *A full sky, low foreground, high resolution CMB map from WMAP*, *Astron. Astrophys.* **493** (2009) 835 [0807.0773].

[28] W. Zhao and D. Baskaran, *Separating E and B types of polarization on an incomplete sky*, *Phys. Rev. D* **82** (2010) 023001 [1005.1201].

[29] S. Ghosh, J. Delabrouille, W. Zhao and L. Santos, *Towards ending the partial sky E-B ambiguity in CMB observations*, *J. Cosmol. Astropart. Phys.* **2021** (2021) 036 [2007.09928].

[30] R. Fernández-Cobos, A. Marcos-Caballero, P. Vielva, E. Martínez-González and R.B. Barreiro, *Exploring two-spin internal linear combinations for the recovery of the CMB polarization*, *Mon. Not. R. Astron. Soc.* **459** (2016) 441 [1601.01515].

[31] H. Liu, J. Creswell, S. von Hausegger and P. Naselsky, *Methods for pixel domain correction of EB leakage*, *Phys. Rev. D* **100** (2019) 023538 [1811.04691].

[32] M. Remazeilles, A.J. Banday, C. Baccigalupi, S. Basak, A. Bonaldi, G. De Zotti et al., *Exploring Cosmic Origins with CORE: B-mode Component Separation*, 1704.04501.

[33] M. Remazeilles, A. Rotti and J. Chluba, *Peeling off foregrounds with the constrained moment ILC method to unveil primordial CMB B-modes*, 2006.08628.

[34] D. Adak, A. Sen, S. Basak, J. Delabrouille, T. Ghosh, A. Rotti et al., *B-mode forecast of CMB-Bhārat*, 2110.12362.

[35] N. Krachmalnicoff, T. Matsumura, E. de la Hoz, S. Basak, A. Gruppuso, Y. Minami et al., *In-flight polarization angle calibration for LiteBIRD: Blind challenge and cosmological implications*, *J. Cosmol. Astropart. Phys.* **2022** (2022) 039 [2111.09140].

[36] Z. Zhang, Y. Liu, S.-Y. Li, D.-L. Wu, H. Li and H. Li, *Efficient ILC analysis on polarization maps after EB leakage correction*, *J. Cosmol. Astropart. Phys.* **2022** (2022) 044 [2109.12619].

[37] A. Carones, M. Migliaccio, D. Marinucci and N. Vittorio, *Analysis of NILC performance on B-modes data of sub-orbital experiments*, *Astron. Astrophys.* **677** (2023) A147 [2208.12059].

[38] R. Aurlien, M. Remazeilles, S. Belkner, J. Carron, J. Delabrouille, H.K. Eriksen et al., *Foreground Separation and Constraints on Primordial Gravitational Waves with the PICO Space Mission*, *JCAP* **06** (2023) 034 [2211.14342].

[39] U. Fuskeland, J. Aumont, R. Aurlien, C. Baccigalupi, A.J. Banday, H.K. Eriksen et al., *Tensor-to-scalar ratio forecasts for extended LiteBIRD frequency configurations*, 2302.05228.

[40] M. Remazeilles, J. Delabrouille and J.-F. Cardoso, *CMB and SZ effect separation with Constrained Internal Linear Combinations*, *Mon. Not. R. Astron. Soc.* **410** (2011) 2481 [1006.5599].

[41] S. Basak and J. Delabrouille, *A needlet ILC analysis of WMAP 7-year data: Estimation of CMB temperature map and power spectrum*, *Mon. Not. R. Astron. Soc.* **419** (2012) 1163 [1106.5383].

[42] S. Basak and J. Delabrouille, *A needlet ILC analysis of WMAP 9-year polarisation data: CMB polarisation power spectra*, *Monthly Notices of the Royal Astronomical Society* **435** (2013) 18 [1204.0292].

[43] P. Collaboration, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., *Planck 2018 results. IV. Diffuse component separation*, *Astron. Astrophys.* **641** (2020) A4 [1807.06208].

[44] K.M. Gorski, E. Hivon, A.J. Banday, B.D. Wandelt, F.K. Hansen, M. Reinecke et al., *HEALPix – a Framework for High Resolution Discretization, and Fast Analysis of Data Distributed on the Sphere*, *ApJ* **622** (2005) 759 [astro-ph/0409513].

[45] P. Collaboration, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., *Planck 2018 results. XI. Polarized dust foregrounds*, *Astron. Astrophys.* **641** (2020) A11 [1801.04945].

[46] P. Collaboration, R. Adam, P.A.R. Ade, N. Aghanim, M. Arnaud, J. Aumont et al., *Planck intermediate results. XXX. The angular power spectrum of polarized dust emission at intermediate and high Galactic latitudes*, *Astron. Astrophys.* **586** (2016) A133 [1409.5738].

[47] S. Ghosh, Y. Liu, L. Zhang, S. Li, J. Zhang, J. Wang et al., *Performance forecasts for the primordial gravitational wave detection pipelines for AliCPT-1*, 2205.14804.

[48] P. Collaboration, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, *Astron. Astrophys.* **641** (2020) A1 [1807.06205].

[49] A. Lewis, A. Challinor and A. Lasenby, *Efficient Computation of CMB anisotropies in closed FRW models*, *ApJ* **538** (2000) 473 [astro-ph/9911177].

[50] M. Reinecke, S. Belkner and J. Carron, *Improved CMB (de-)lensing using general spherical harmonic transforms*, *Astron. Astrophys.* **678** (2023) A165 [2304.10431].

[51] J. Delabrouille, M. Betoule, J.-B. Melin, M.-A. Miville-Deschênes, J. Gonzalez-Nuevo, M.L. Jeune et al., *The pre-launch Planck Sky Model: A model of sky emission at submillimetre to centimetre wavelengths*, *Astron. Astrophys.* **553** (2013) A96 [1207.3675].

[52] A. Zonca, B. Thorne, N. Krachmalnicoff and J. Borrill, *The Python Sky Model 3 software*, *JOSS* **6** (2021) 3783 [2108.01444].

[53] P. Collaboration, N. Aghanim, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., *Planck intermediate results. XLVIII. Disentangling Galactic dust emission and cosmic infrared background anisotropies*, *Astron. Astrophys.* **596** (2016) A109 [1605.09387].

[54] C.L. Bennett, D. Larson, J.L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard et al., *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, *ApJS* **208** (2013) 20 [1212.5225].

[55] B. Thorne, J. Dunkley, D. Alonso and S. Naess, *The Python Sky Model: Software for simulating the Galactic microwave sky*, *Mon. Not. R. Astron. Soc.* **469** (2017) 2821 [1608.02841].

[56] P. Collaboration, R. Adam, P.A.R. Ade, N. Aghanim, M.I.R. Alves, M. Arnaud et al., *Planck 2015 results. X. Diffuse component separation: Foreground maps*, *Astron. Astrophys.* **594** (2016) A10 [1502.01588].

[57] M.-A. Miville-Deschenes, N. Ysard, A. Lavabre, N. Ponthieu, J.F. Macias-Perez, J. Aumont et al., *Separation of anomalous and synchrotron emissions using WMAP polarization data*, *Astron. Astrophys.* **490** (2008) 1093 [0802.3345].

[58] N. Krachmalnicoff, E. Carretti, C. Baccigalupi, G. Bernardi, S. Brown, B.M. Gaensler et al., *The S-PASS view of polarized Galactic Synchrotron at 2.3 GHz as a contaminant to CMB observations*, *Astron. Astrophys.* **618** (2018) A166 [1802.01145].

[59] D. Alonso, J. Sanchez and A. Slosar, *A unified pseudo-C_ℓ framework*, *Mon. Not. R. Astron. Soc.* **484** (2019) 4127 [1809.09603].

[60] J. Han, B. Hu, S. Ghosh, S. Li, J. Dou, J. Delabrouille et al., *Forecasts of CMB lensing reconstruction of AliCPT-1 from the foreground cleaned polarization data*, 2303.05705.

[61] D. Foreman-Mackey, D.W. Hogg, D. Lang and J. Goodman, *Emcee: The MCMC Hammer*, *Publications of the Astronomical Society of the Pacific* **125** (2013) 306 [1202.3665].

[62] M. Tegmark and A. de Oliveira-Costa, *Removing point sources from CMB maps*, *Astrophys. J.* **500** (1998) L83 [astro-ph/9802123].

[63] J.L. Starck, J. Bobin, F. Sureau, J. Schmitt, O. Fourt, Y. Moudden et al., *iSAP: Interactive Sparse Astronomical Data Analysis Packages*, Astrophys. Source Code Libr. (2013) ascl:1303.029.