Partially constrained internal linear combination: A method for low-noise CMB foreground mitigation

Y. Sultan Abylkairov,[1,2,3,*] Omar Darwish,[3] J. Colin Hill,[4,5] and Blake D. Sherwin[3,6]

[1]*Physics Department, Nazarbayev University, Nur-Sultan 010000, Kazakhstan*
[2]*Energetic Cosmos Laboratory, Nazarbayev University, Nur-Sultan 010000, Kazakhstan*
[3]*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom*
[4]*Department of Physics, Columbia University, New York, New York 10027, USA*
[5]*Center for Computational Astrophysics, Flatiron Institute, New York, New York 10010, USA*
[6]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 OHA, United Kingdom*

(Received 25 January 2021; accepted 12 April 2021; published 11 May 2021)

Internal linear combination (ILC) methods are some of the most widely used multifrequency cleaning techniques employed in cosmic microwave background (CMB) data analysis. These methods reduce foregrounds by minimizing the total variance in the coadded map (subject to a signal-preservation constraint), although often significant foreground residuals or biases remain. A modification to the ILC method is the constrained ILC, which explicitly nulls certain foreground components; however, this foreground nulling often comes at a high price for ground-based CMB datasets, with the map noise increasing significantly on small scales. In this paper we explore a new method, the partially constrained ILC, which allows us to optimize the tradeoff between foreground bias and variance in ILC methods. In particular, this method allows us to minimize the variance subject to an inequality constraint requiring that the constrained foregrounds are reduced by at least a fixed factor, which can be chosen based on the foreground sensitivity of the intended application. We test our method on simulated sky maps for a Simons Observatory–like experiment; we find that for cleaning thermal Sunyaev-Zel'dovich contamination at $\ell \in [3000, 4800]$, if a small thermal Sunyaev-Zel'dovich residual of 20% of the standard ILC residual can be tolerated, the variance of the CMB temperature map is reduced by at least 50% over the constrained ILC value. We also demonstrate an application of this method to reduce noise in CMB lensing reconstruction.

DOI: 10.1103/PhysRevD.103.103510

I. INTRODUCTION

The cosmic microwave background radiation (CMB) is one of our most important sources of information about cosmology and fundamental physics. Over the past decades, much of its constraining power has arisen from the primary CMB anisotropies. However, increasingly, the CMB is also being used as a backlight to understand the distribution of matter, gas and tracers lying between us and the last scattering surface, using the secondary anisotropies these imprint into the microwave background.

To robustly analyze either the primary CMB or the individual astrophysical contributions, multifrequency *component separation methods*, which use different frequency dependences to disentangle the different components, are becoming increasingly important.

There are several methods that have been proposed to separate the CMB signal, or another astrophysical signal of interest, from the other components that are present in an observed CMB map. Perhaps the most widely used method is the internal linear combination (ILC [1–9]) method, which combines in a linear fashion multifrequency observations in order to extract an unbiased estimate of the desired component (e.g., CMB).[1] This method employs a linear combination of frequency channels that minimizes the total map variance, subject to the constraint of an unbiased recovery of the desired component, with weights calculated from an empirically determined covariance matrix. A frequently used extension of ILC is the constrained ILC (CILC [10]), where the linear combination is constructed in such a way to minimize the variance subject to the additional constraint that a particular component, with known spectral dependence, is nulled in the extracted map. Without such nulling (also known as

*sultan.abylkairov@nu.edu.kz

[1]Note that, because of frequency dependence, these frequency methods in general will extract a combination of CMB + kSZ + other frequency independent elements.

"deprojection"), ILC maps can possess significant foreground residuals (e.g., [11,12]). The resulting deprojected maps have a wide range of applications, including primordial non-Gaussianity [13,14], cross-correlations (e.g., studying kinetic Sunyaev-Zel'dovich (kSZ) [15,16] or ISW [17]), CMB lensing reconstruction [11,18], and primordial B modes [19]. However, a major downside of using constrained ILC methods is that the additional constraints often lead to a substantial noise increase in the resulting map, particularly for ground-based CMB experiments with a moderate number of frequency channels.

However, this large noise penalty is not generally necessary if we only wish to obtain a map with a reduced level of foregrounds. Depending on the application, the complete nulling of foreground contamination may not be required, and it may instead be sufficient to merely reduce the contamination by a large factor in amplitude. This is the goal of this paper, which presents the partially constrained ILC (PCILC) method. This method guarantees an overall foreground bias that is reduced by at least a fixed factor, but, by not requiring foregrounds to be completely nulled, can result in a significantly lower noise than the CILC. The method is easily applicable; as an example, we will show an application to foreground cleaning for CMB gravitational lensing reconstruction.

In Sec. II, following a brief review of existing multifrequency component separation methods, we will introduce the PCILC. In Sec. III we will show and discuss the results of our method when applied to simulations, and to an example of CMB lensing reconstruction. We conclude in Sec. IV. Further technical results are collected in the Appendixes.

II. METHOD

Obtaining a clean and accurate CMB map from observational data is difficult due to various foreground signals such as the thermal Sunyaev-Zeldovich (TSZ) effect and the cosmic infrared background (CIB). In this section, we will quickly review the standard ILC, the constrained ILC, and finally our new proposed method, the partially constrained ILC.

A. The ILC method

The ILC is a commonly used method because it requires minimal modeling assumptions about the data and has considerable flexibility in the choice of domain in which to extract the signal of interest. If we have N_ν frequency channels in our observational data, then for each pixel p we can write an $N_\nu \times 1$ vector where each row represents the observed map at the corresponding frequency channel:

$$\mathbf{y}(p) = \mathbf{a}s(p) + \mathbf{A}_f\mathbf{s}_f(p) + \mathbf{n}(p) \tag{1}$$

where $\mathbf{a}$ is the spectral energy distribution (SED) response vector of the desired signal $s(p)$, $\mathbf{A}_f$ is the mixing matrix for the foreground components $\mathbf{s}_f$ (to know how much a specific foreground i contributes to the map at the observed frequency j), and $\mathbf{n}(p)$ is the noise. We will focus solely on the case of CMB ILC reconstruction, and thus $\mathbf{a}$ is the CMB SED, which is unity when working in thermodynamic temperature units. Note that p can be a point in any desired space, e.g., in harmonic space, real space, or a needlet frame. The ILC solution provides a linear combination of maps $\hat{s} = \mathbf{w}^T\mathbf{y}$ that recover the component of interest, in this case the CMB, and it has a minimum variance

$$\min\left(\langle\hat{s}^2\rangle - \langle\hat{s}\rangle^2\right) = \min(\mathbf{w}^T\mathbf{R}\mathbf{w}), \tag{2}$$

where $\mathbf{R} = \langle\mathbf{y}\mathbf{y}^T\rangle - \langle\mathbf{y}\rangle\langle\mathbf{y}^T\rangle$ is the covariance matrix of the data. Solving Eq. (2) under the constraint $\mathbf{w}^T\mathbf{a} = 1$, to ensure an unbiased recovery of the component of interest, gives the ILC weights (e.g., [3]):

$$\mathbf{w}_{\mathrm{ILC}}^T = (\mathbf{a}^T\mathbf{R}^{-1}\mathbf{a})^{-1}\mathbf{a}^T\mathbf{R}^{-1}. \tag{3}$$

B. The constrained ILC method

The CILC similarly involves building a linear combination of observed maps, at different frequencies, $\hat{s} = \mathbf{w}^T\mathbf{y}$ that recovers the component of interest with minimum possible variance; however, the CILC involves the additional constraint of nulling some unwanted foreground or other components.

To recover the CMB signal, while deprojecting some foregrounds, i.e., nulling some components of the $\mathbf{s}_f$, we use the corresponding SED vectors from the mixing matrix $\mathbf{A}_f$. We define these SED vectors as $(\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_m})$. Then we can write a condition under which we completely deproject these components:

$$\begin{cases} \mathbf{w}^T\mathbf{b_1} = 0 \\ \mathbf{w}^T\mathbf{b_2} = 0 \\ \vdots \\ \mathbf{w}^T\mathbf{b_m} = 0 \end{cases}. \tag{4}$$

In this way we guarantee that the contribution of the selected foregrounds to the final linear combination map will be zero, which is not necessarily true for the ILC case. However, this deprojection comes at a price: since we have used one or more degrees of freedom for the deprojection, the noise in the final CILC map is guaranteed to be higher than that in the standard ILC case.

Formally, the constrained ILC solution provides a linear combination of maps $\hat{s} = \mathbf{w}^T\mathbf{y}$ such that it has minimal variance, subject to the constraints $\mathbf{w}^T\mathbf{a} = 1$, $\mathbf{w}^T\mathbf{b_1} = 0$, $\mathbf{w}^T\mathbf{b_2} = 0, \ldots, \mathbf{w}^T\mathbf{b_m} = 0$. In this case, the weights are (e.g., [10]):

TABLE I. Table showing which variables are zero, out of all the combinations described in step three, and whether for this combination of zero-valued variables it is possible (Yes or No) to find a solution for the system of equations in Eq. (9).

λ_1			s_1			λ_1 & s_1		
λ_2	s_2	λ_2 & s_2	λ_2	s_2	λ_2 & s_2	λ_2	s_2	λ_2 & s_2
Yes	Yes	No	Yes	No	No	No	No	No

$$\mathbf{w}_{\mathrm{CILC}}^T = \mathbf{e}^T(\mathbf{A}^T\mathbf{R}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{R}^{-1}, \tag{5}$$

where $\mathbf{A} = [\mathbf{a}\,\mathbf{b_1}\cdots\mathbf{b_m}]$ is a matrix of size $N_\nu \times (m+1)$, and $\mathbf{e}^T = [1\,0\cdots 0]$ is a vector of $1\times(m+1)$, so that we can recover the CMB (which, in our formalism, is always the first component) if the unwanted components are foregrounds.

C. The partially constrained ILC method

In the constrained ILC, by deprojection we lose one degree of freedom for each deprojected component; this inevitably leads to an increase of variance in the combined map. Here we propose a new method where we partially deproject foregrounds to get an intermediate solution between the ILC and the constrained ILC, i.e., to achieve a balance between foreground bias and variance reduction.

As a starting point, suppose we have just the CMB and one foreground component, with the SED vector $\mathbf{b_1}$, that we wish to reduce in the final combination. Partial deprojection can be expressed as

$$|\mathbf{w}^T\mathbf{b_1}| \le \epsilon, \tag{6}$$

where ϵ is some arbitrary positive number which controls the level of residual foregrounds in the final map. By defining new "slack variables" s_1 and s_2 to turn inequality constraints to equality constraints, we write the modulus in Eq. (6) as two equations with different signs [20]; the inequality constraint can then be expressed as follows:

$$\begin{aligned}\epsilon - \mathbf{w}^T\mathbf{b_1} - s_1^2 = 0,\\ \epsilon + \mathbf{w}^T\mathbf{b_1} - s_2^2 = 0.\end{aligned} \tag{7}$$

To find weights $\mathbf{w}$ such that the combined map has minimal variance under constraints, we use the method of Lagrange multipliers:

$$\begin{aligned}\mathcal{L} = \mathbf{w}^T\mathbf{R}\mathbf{w} + \lambda(1-\mathbf{w}^T\mathbf{a}) + \lambda_1(\epsilon - \mathbf{w}^T\mathbf{b_1} - s_1^2)\\ + \lambda_2(\epsilon + \mathbf{w}^T\mathbf{b_1} - s_2^2)\end{aligned} \tag{8}$$

Minimizing this, we obtain a linear system of equations:

$$\begin{cases}\frac{\partial\mathcal{L}}{\partial\mathbf{w}^T} = 2\mathbf{R}\mathbf{w} - \lambda\mathbf{a} - \lambda_1\mathbf{b_1} + \lambda_2\mathbf{b_1} = \mathbf{0}\\ \frac{\partial\mathcal{L}}{\partial s_1} = -2\lambda_1 s_1 = 0\\ \frac{\partial\mathcal{L}}{\partial s_2} = -2\lambda_2 s_2 = 0\\ \frac{\partial\mathcal{L}}{\partial\lambda} = 1 - \mathbf{a}^T\mathbf{w} = 0\\ \frac{\partial\mathcal{L}}{\partial\lambda_1} = \epsilon - \mathbf{b_1}^T\mathbf{w} - s_1^2 = 0\\ \frac{\partial\mathcal{L}}{\partial\lambda_2} = \epsilon + \mathbf{b_1}^T\mathbf{w} - s_2^2 = 0\end{cases}. \tag{9}$$

We will now outline a step-by-step solution of the system of Eq. (9):

Step one: the first equation of the system gives

$$\mathbf{w} = \frac{1}{2}\mathbf{R}^{-1}(\lambda\mathbf{a} + \lambda_1\mathbf{b_1} - \lambda_2\mathbf{b_1}). \tag{10}$$

Step two: substituting Eq. (10) into other equations that contain $\mathbf{w}$ in the system of equations

$$\mathbf{a}^T\frac{1}{2}\mathbf{R}^{-1}(\lambda\mathbf{a} + \lambda_1\mathbf{b_1} - \lambda_2\mathbf{b_1}) = 1, \tag{11}$$

$$\mathbf{b_1}^T\frac{1}{2}\mathbf{R}^{-1}(\lambda\mathbf{a} + \lambda_1\mathbf{b_1} - \lambda_2\mathbf{b_1}) = \epsilon - s_1^2, \tag{12}$$

$$\mathbf{b_1}^T\frac{1}{2}\mathbf{R}^{-1}(\lambda\mathbf{a} + \lambda_1\mathbf{b_1} - \lambda_2\mathbf{b_1}) = s_2^2 - \epsilon. \tag{13}$$

Step three: from the second and third equations of the system of equations it follows: $\lambda_1 = 0$ or $s_1 = 0$ or $\lambda_1 = 0$& $s_1 = 0$, and $\lambda_2 = 0$ or $s_2 = 0$ or $\lambda_2 = 0$ & $s_2 = 0$. By simply substituting the possible combinations into Eqs. (11)–(13), we can build the following table:

Table I shows all possible combinations of different variables being zero: $\lambda_1 = 0$ or $s_1 = 0$ or both, and $\lambda_2 = 0$ or $s_2 = 0$ or both. Specifically, the first row labels which of the λ_1, s_1 variables are zero (or whether both are zero); the second row similarly labels which of the λ_2, s_2, λ_2 & s_2 variables are zero, and the third row shows whether this combination of zero-valued variables allows the system of equations to have an answer (Yes/No).

Step four: finally, using simple algebra for $\lambda_1 = 0$ and $\lambda_2 = 0$ we have the standard ILC solution, for $\lambda_1 = 0$ and $s_2 = 0$ we have $\mathbf{w}_{\mathrm{PCILC-}}$ [see Eq. (14)], and for $\lambda_2 = 0$ and $s_1 = 0$ we have $\mathbf{w}_{\mathrm{PCILC+}}$ [see Eq. (14)].

Therefore, solving the linear system of Eq. (9) gives us three solutions:

(i) The first solution is the ILC solution.

(ii) The other two take values at the boundary, i.e., $\mathbf{w}_{\mathrm{PCILC}_\pm}^T\mathbf{b_1} = \pm\epsilon$, and have the following form:

$$\begin{aligned}&\mathbf{w}_{\mathrm{PCILC}_\pm}^T\\ &\quad= \frac{\mathbf{a}^T\mathbf{R}^{-1}(K_b \mp K_{ab}\epsilon) + \mathbf{b_1}^T\mathbf{R}^{-1}(\pm K_a\epsilon - K_{ab})}{K_aK_b - K_{ab}^2},\end{aligned} \tag{14}$$

where $K_a = \mathbf{a}^T\mathbf{R}^{-1}\mathbf{a}$, $K_b = \mathbf{b_1}^T\mathbf{R}^{-1}\mathbf{b_1}$ and $K_{ab} = \mathbf{a}^T\mathbf{R}^{-1}\mathbf{b_1}$.

The answers above make sense if we consider this problem as follows: the variance function $\mathbf{w}^T\mathbf{R}\mathbf{w}$: $\mathbb{R}^n \to \mathbb{R}$, where $n = N_\nu$ is a number of frequency channels. Geometrically this is an elliptic paraboloid with a minimum at $\mathbf{w} = \mathbf{0}$. With our first constraint $\mathbf{w}^T\mathbf{a} = 1$, our working domain is reduced to $\mathbb{R}^{n-1}$ and geometrically it is still an elliptic paraboloid with minimum at $\mathbf{w}_{\rm ILC}$. Adding the inequality constraint $|\mathbf{w}^T\mathbf{b_1}| \le \epsilon$ is equivalent to considering an allowed interval I in $\mathbb{R}^{n-1}$. Therefore, the minimum will be the standard ILC solution if the interval includes $\mathbf{w}_{\rm ILC}$, otherwise it will lie at the boundaries (i.e., if $\mathbf{w}_{\rm ILC} \notin I$).

To find the correct overall solution, we first have to check if the inequality condition is satisfied for the standard ILC solution, i.e., $|\mathbf{w}_{\rm ILC}^T\mathbf{b_1}| \le \epsilon$. If so, then the answer ($\mathbf{w}_{\rm PCILC}$) is equal to the standard ILC solution. Otherwise, we calculate the variance of the combined map for both weights $\mathbf{w}_{\rm PCILC_\pm}$, i.e., $\mathbf{w}_{\rm PCILC_\pm}^T\mathbf{R}\mathbf{w}_{\rm PCILC_\pm}$, and compare them. The answer in this case is the weight vector for which we get the smallest variance.

In the Appendix we discuss how to generalize this derivation to multiple partially deprojected components.

III. RESULTS AND DISCUSSION

A. Simulations

We test our proposed method on the high-resolution simulations of the microwave sky[2] generated by The Simons Observatory Collaboration [22,23]. For simplicity, we perform the ILC in harmonic space, but the novel aspects of our formalism can be straightforwardly applied in pixel space or on a needlet frame as well. The simulation maps are constructed for six frequency channels at which the Simons Observatory (SO) will operate: 27, 39, 93, 145, 225, and 280 GHz. For simplicity, delta-function passbands are assumed.

In this work, we use a simple sky model that includes the lensed CMB signal, the TSZ effect, and the CIB. The TSZ effect is the inverse-Compton scattering of CMB photons off hot, free electrons, which generates a unique spectral distortion in the mm-wave bands [24]. The CIB is the cumulative thermal emission from dust grains heated by starlight in galaxies over cosmic history. The lensing, TSZ, and CIB fields were constructed in these simulations by post processing a large N-body simulation with prescriptions for each observable [23]. All components are thus realistically correlated. Further details on each individual component can be found in Refs. [22,23], including adjustments that were made to more closely match recent measurements of these fields. The noise model in the simulations is generated from the properties of the planned SO surveys, i.e., the "baseline" level for the SO Large Aperture Telescope with observed sky fraction $f_{\rm sky} = 40\%$ (see [22]). Note that the noise maps are correlated at 27 and 39, 93 and 145, and at 225 and 280 GHz due to the atmospheric correlations for frequency channels in the same optics tube (see [22]). We combine the lensed CMB, TSZ, CIB, and noise components for each frequency channel.

[2]The simulations can be found at [21].

B. Frequency dependence of components

To apply the CILC and PCILC methods, we need to know the frequency response models of the components that we wish to deproject or partially deproject. In this work, the components we will focus on are the TSZ and CIB.

For the TSZ effect [24–26] the frequency dependence in thermodynamic CMB temperature units is given by

$$f_{\rm TSZ}(\nu) = x\frac{e^x+1}{e^x-1} - 4, \tag{15}$$

where $x = h\nu/(k_B T_{\rm CMB})$. In contrast to the TSZ effect, the CIB is not a single field that is rigidly rescaled across frequency channels according to a fixed SED. However, as an approximation, we adopt the following modified blackbody SED for the CIB [8]:

$$f_{\rm CIB}(\nu) \propto \frac{\nu^{3+\beta}}{e^{h\nu/(k_B T_{\rm CIB})}-1}\left(\left.\frac{dB(\nu,T)}{dT}\right|_{T=T_{\rm CMB}}\right)^{-1} \tag{16}$$

where $\beta = 1.2$, $T_{\rm CMB} = 24$ K, and $B(\nu, T)$ is the Planck function, needed here to convert from specific intensity to thermodynamic CMB temperature units. We emphasize that the CIB component in the simulated sky maps is not generated assuming this SED, but rather from detailed post processing of a light cone from an N-body simulation, using semianalytic star formation prescriptions. Thus, the simulated CIB maps do not follow a single, rigid SED, and they exhibit realistic decorrelation across frequency channels [22].

Finally, for our frequency channels, the CMB SED is a constant and equal to unity, since we work in thermodynamic CMB temperature units.

C. Choosing a value for ϵ

To understand how to choose a reasonable threshold value for partial deprojection [see Eq. (6)], we will first explain the calculation of the foreground bias values. We define a foreground bias fraction as

$$B := \frac{\mathbf{w}^T\mathbf{F}\mathbf{w}}{F_{145\times145}}, \tag{17}$$

where $\mathbf{F}$ is the empirically determined frequency-frequency covariance matrix for this foreground and $F_{145\times145}$ is the

power spectrum of this foreground at 145 GHz obtained from the simulations. This bias variable B represents the size of the residual foreground power after applying the PCILC method (equal to $\mathbf{w}^T\mathbf{F}\mathbf{w}$), relative to the original foreground power at 145 GHz. We can then define a positive number B^{th} such that $B \leq B^{\text{th}}$, which defines a threshold value of the foreground bias. Note that from Eq. (17) we obtain the foreground bias for the standard ILC, by inserting the standard ILC weights in the numerator. Based on this result, we know a reasonable upper bound for the threshold bias value B^{th}, since any threshold bias value above the standard ILC bias will just reproduce the standard ILC weights, as discussed in Sec. II C. Depending on how much we want to reduce the variance in the final ILC map, we can thus choose any value between the standard ILC foreground bias and zero. Finally, we can calculate the threshold value to be used for the PCILC weights determination using the following equation:

$$\epsilon = \sqrt{B^{\text{th}} \cdot F_{145\times145}}. \tag{18}$$

In this paper, for simplicity, we use the same threshold value for all ℓ, even if for some ℓ it exceeds the standard ILC bias. Specifically, we constrain the TSZ bias to be below a threshold $B_{\text{TSZ}} \leq 0.2$ or 0.3 and CIB bias to be below a threshold $B_{\text{CIB}} \leq 0.2$ or 0.3.

D. ILC: Results and discussion

Figure 1 shows the results obtained using ILC, CILC, and PCILC for CMB map reconstruction on the SO-like simulations, where the constraints are applied only to the TSZ foreground component. As mentioned previously, for the PCILC method, we show results for both $B_{\text{TSZ}} \leq 0.2$ and $B_{\text{TSZ}} \leq 0.3$. All calculations are performed in linearly spaced multipole bins of width $\Delta\ell = 21$. The upper left panel shows the total power spectrum of the lensed CMB signal reconstructed by various methods, and the upper right panel shows the ratio of these power spectra to the power spectrum of the lensed CMB obtained using the standard ILC method. The lower left and right (lr) panels show the TSZ bias and the CIB bias, as defined in Eq. (17). The same configuration is shown in Figs. 3 and 5, but in Fig. 3 the PCILC method is implemented for the CIB foreground component with $B_{\text{CIB}} \leq 0.2$ and $B_{\text{CIB}} \leq 0.3$, and in Fig. 5 the PCILC method is implemented for both foreground components simultaneously, i.e., TSZ and CIB

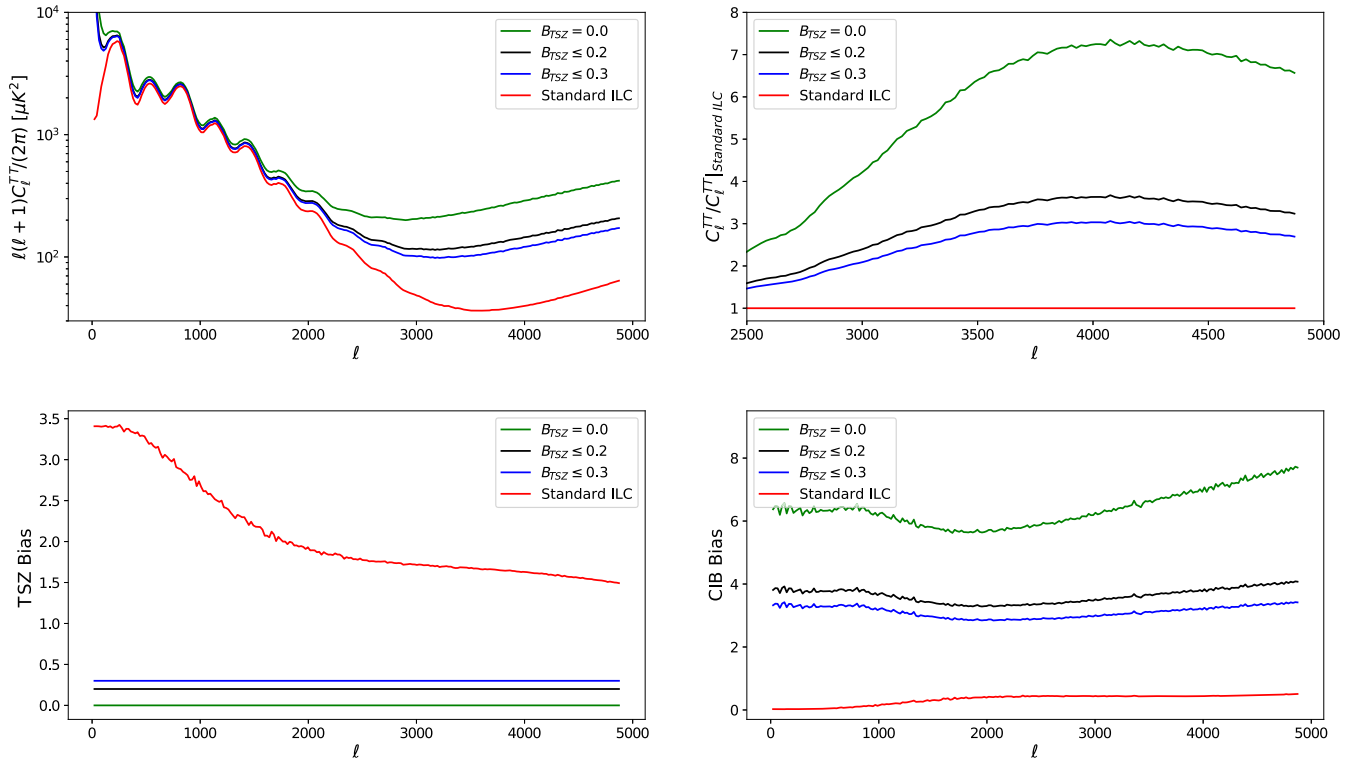


FIG. 1. Partial deprojection of the TSZ component. Results from our new PCILC method applied to simulated sky maps with various foreground bias threshold values as defined in Eq. (17) (blue and black curves) compared to the standard ILC (red) and the constrained ILC, or CILC, (green) results. Upper left: total power spectra of the reconstructed (PC)ILC CMB maps. Upper right: the ratio of the power spectra to the total CMB power spectrum obtained with the standard ILC. Lower left and lower right: the residual TSZ power (left) and CIB power (right) in the coadded maps, measured relative to the power of the TSZ or CIB at 145 GHz [see Eq. (17)]. It can be seen that, if a small bias can be tolerated, the PCILC method provides a significant variance reduction when compared with the CILC.

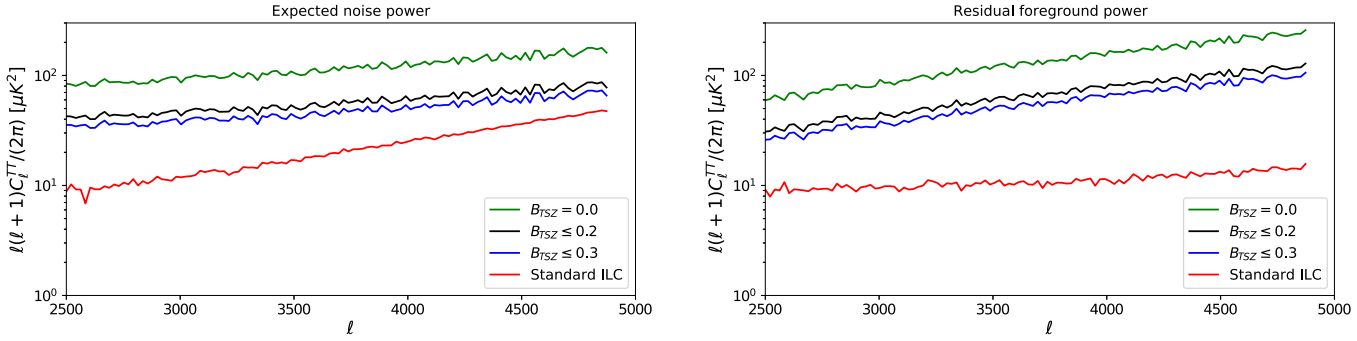


FIG. 2. Complementary results to Fig. 1. Expected noise power (left) and residual foreground power (right).

with $B_{\rm TSZ} \leq 0.2$ and $B_{\rm CIB} \leq 0.2$, and $B_{\rm TSZ} \leq 0.3$ and $B_{\rm CIB} \leq 0.3$.

Figures 2, 4, and 6 show complementary results to Figs. 1, 3, and 5, respectively. In each figure, the right panel shows the residual foreground power, and the left panel shows the expected noise power (instrumental and atmospheric) obtained by subtracting from the total result power the input theory CMB power and the residual foreground power.

The results in Fig. 1 show that when the TSZ component is fully deprojected with the CILC, the variance of the reconstructed CMB map increases by more than a factor of six compared to the standard ILC for some multipoles, and the residual CIB power becomes much higher than its value for the standard ILC (as expected, since the constrained weights have less freedom to adjust to minimize CIB contamination). However, with partial deprojection such that $B_{\rm TSZ} \leq 0.2$, 0.3 using the PCILC, the variance and the CIB bias only increase by a moderate amount over the standard ILC results. In particular, for $B_{\rm TSZ} \leq 0.3$ the resulting total power spectrum is less than half the power spectrum of the CILC map for $\ell \in [3000, 4800]$.

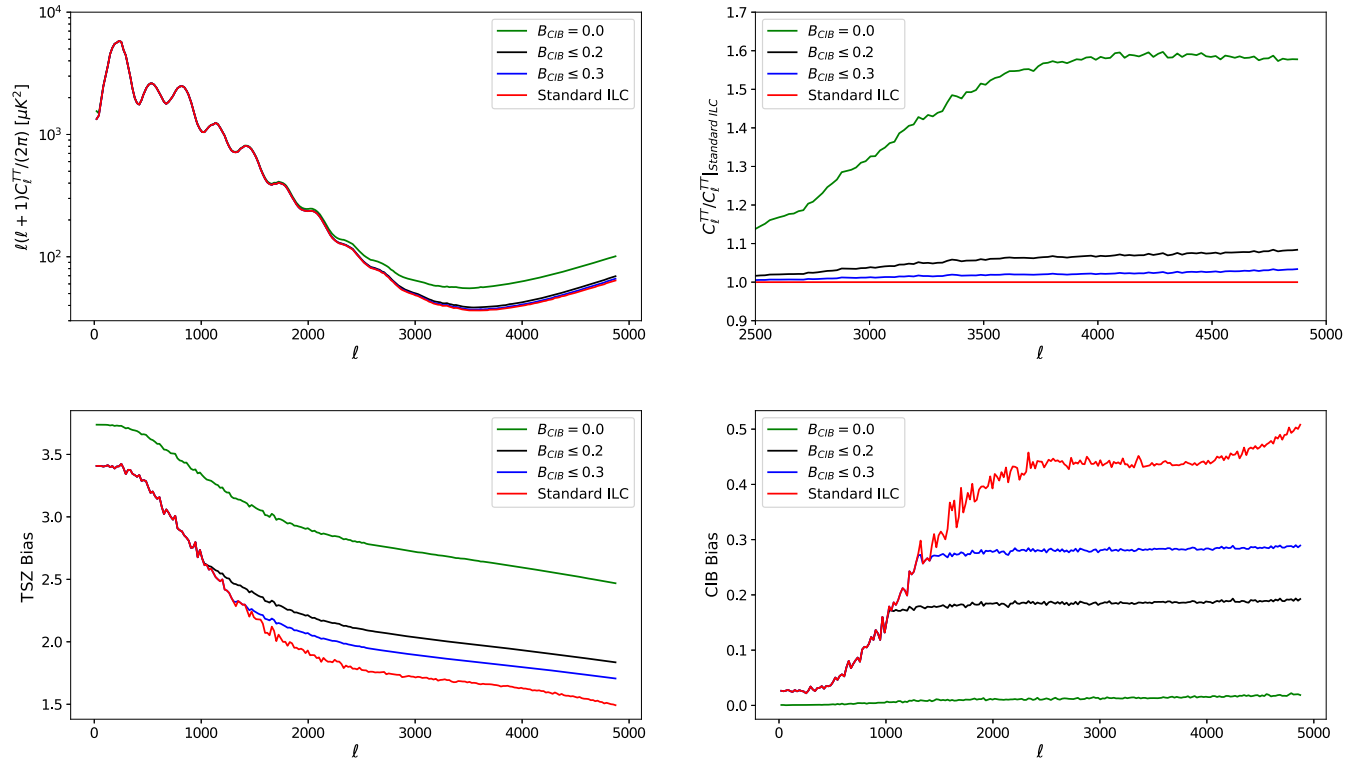


FIG. 3. Partial deprojection of the CIB component. The same configuration of plots as in Fig. 1, but now with the CIB foreground component deprojected or partially deprojected, as labeled in the plot legends. The CIB SED is taken to be a modified blackbody [see Eq. (16)] in the ILC constraints, but the sky simulations are constructed with a realistic model that produces decorrelation and a nonrigid SED that varies with frequency and sky position. This is why the residual CIB bias is slightly nonzero even when we set $B_{\rm CIB} = 0$ (see the green curve in the lower right panel).

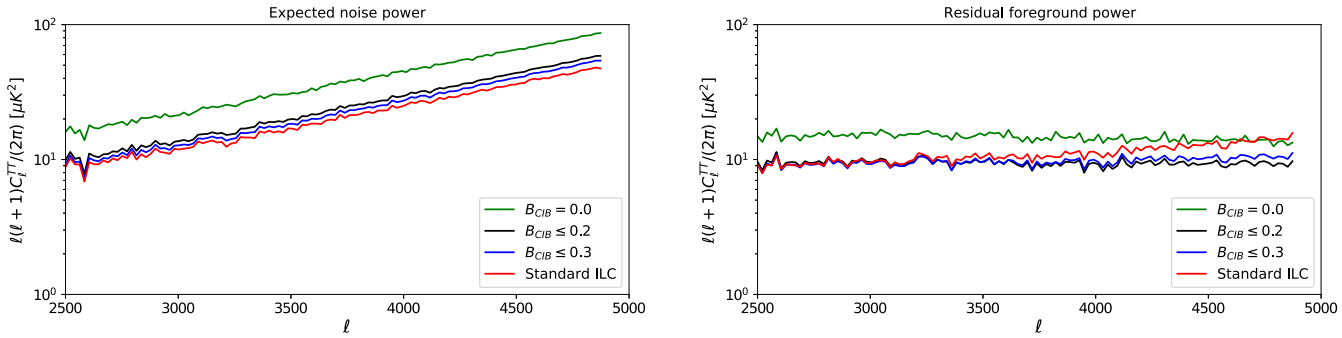


FIG. 4. Complementary results to Fig. 3. Expected noise power (left) and residual foreground power (right).

For the CIB deprojection in Fig. 3, the improvements when using the PCILC method are not as striking as for TSZ deprojection. Nevertheless, if we can tolerate a small residual CIB bias, we still can lower the effective power spectrum by tens of percent and reduce the residual TSZ bias substantially by using the PCILC method instead of the CILC, as shown in Fig. 3.

The main disadvantage of single-component CILC and PCILC is that we do not control the bias of other foregrounds, and therefore these biases can become larger, as shown previously. By deprojection and partial deprojection for two or more components, we control the level of bias of multiple foregrounds, or two for the case of our sky simulations here. In this case, where we deproject both TSZ and CIB, CILC increases the variance more than 40 times compared to the ILC for high ℓ, as shown in Fig. 5. Using the PCILC method, we can significantly reduce the variance.

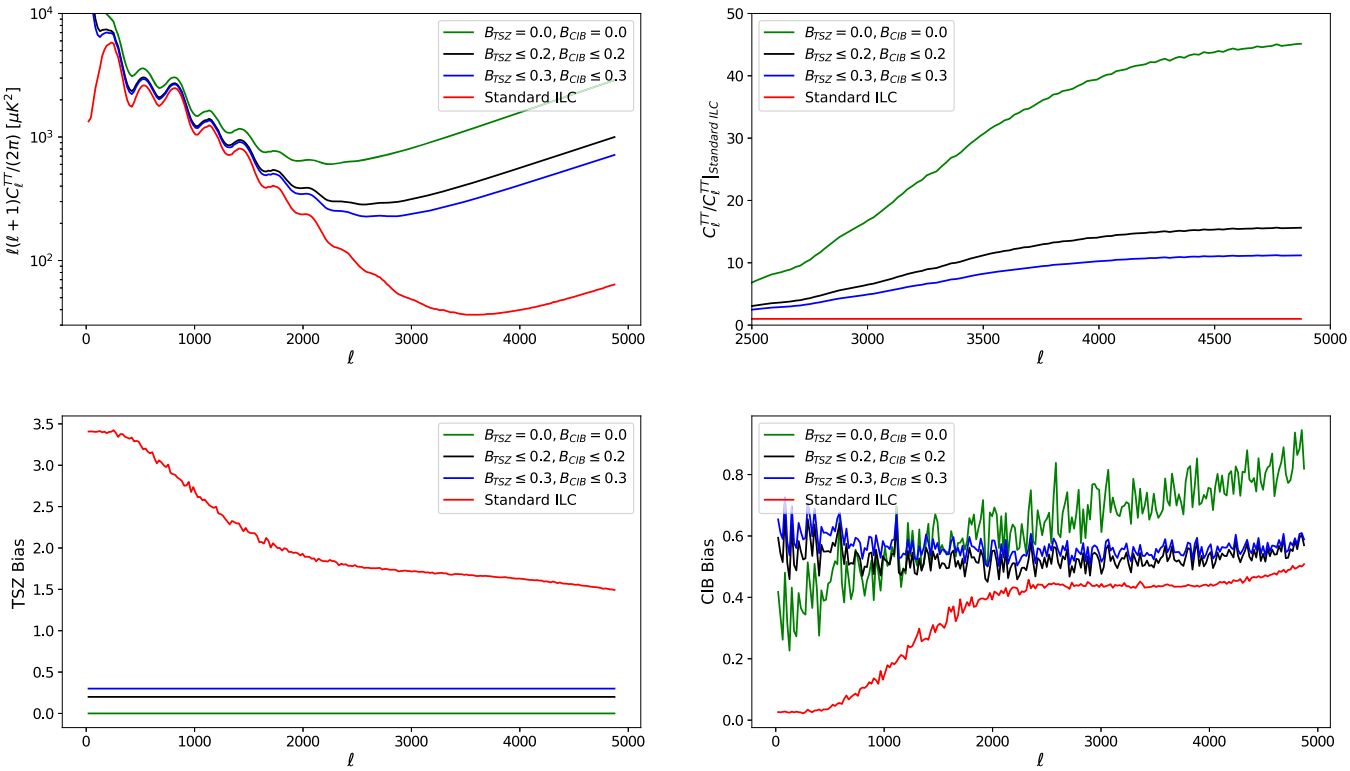


FIG. 5. Partial deprojection of both TSZ and CIB components. The same configuration of plots as in Fig. 1, but with simultaneous deprojection or partial deprojection of both the TSZ and CIB foreground components, as labeled in the plot legends. Again, it can be seen in the top right panel that, if a small bias can be tolerated, the PCILC method provides a significant variance reduction when compared with the CILC. However, the bottom right panel shows that the CIB bias remains significantly nonzero even when we attempt to fully deproject this component. This situation can arise when (partially) deprojecting multiple foregrounds with a small number of frequency channels, such that even small inaccuracies in the modeling lead to non-negligible residual foreground biases (see text for further discussion).

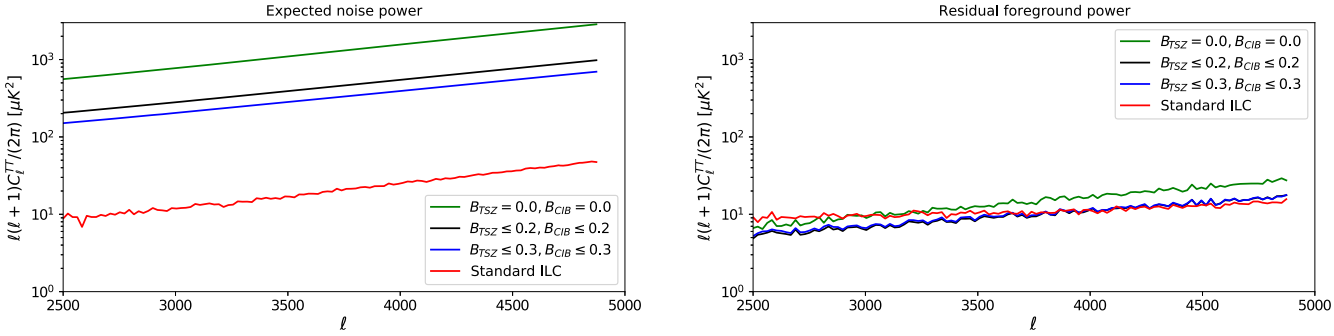


FIG. 6. Complementary results to Fig. 5. Expected noise power (left) and residual foreground power (right).

Although PCILC performs well at reducing variance, the CIB bias becomes more sensitive to decorrelation and suboptimal SED modeling as additional constraints are added. The small decorrelation and suboptimal SED model of the CIB signal affect the accuracy of the CIB bias removal, as shown in Fig. 5(lr). However, we would expect the same variance reduction performance for more accurate models (see next subsection); experiments with more frequency channels should, in addition, suffer less from bias residuals when (partially) deprojecting multiple components (see Ref. [19]). In Appendix A, we analyze in detail the problem of accurately reducing the CIB bias. We also note that including additional CIB components via a moment expansion [27] could reduce the bias seen in Fig. 5, although this may be challenging for experiments with a relatively small number of frequency channels.

Aside from the CIB bias calculation, we have verified that the simulated results match our forecast performance well. This gives further confidence in our method and our pipeline performance.

E. Validation of CIB results using simplified sky simulations

In this subsection, we use a simplified version of the simulated sky described in Sec. III A, with the only difference that instead of the standard CIB signals, we use one CIB signal at 145 GHz and scale it to the other frequency channels using Eq. (16). Thus, in these simplified simulations, the CIB field is comprised of a single component, which is rescaled across frequencies with a rigid SED. We combine this component with the CMB, TSZ, and noise as done for the original simulations. With these simplified sky maps, we check how well our method works for partially deprojecting both TSZ and CIB when we have a perfectly understood, one-component CIB signal. The results are shown in Fig. 7, and they demonstrate that the CIB bias is now successfully removed when $B_{\rm TSZ} = 0 = B_{\rm CIB}$, and behaves as expected in the other cases shown. This validates our claim that the residual CIB biases seen in Fig. 5 are indeed due to SED variations and decorrelation in the CIB signal in the original simulations.

Also note that for two-component partial deprojection, in order to obtain the standard ILC solution, the corresponding intervals from the inequality constraints for TSZ and CIB, that is $I_{\rm TSZ}$ and $I_{\rm CIB}$, must include $\mathbf{w}_{\rm ILC}$ (i.e., $\mathbf{w}_{\rm ILC} \in I_{\rm TSZ}$ and $\mathbf{w}_{\rm ILC} \in I_{\rm CIB}$). This is why, in Fig. 7, the CIB bias may be higher than its corresponding bias from the standard ILC if at the same time the TSZ bias is lower than its corresponding bias from the standard ILC.

Figure 8 shows how the ratio of the power spectra obtained with PCILC for two components (TSZ and CIB) to the spectrum of the total lensed CMB obtained with standard ILC at $\ell = 3500$ varies with $B^{\rm th}_{\rm TSZ}$ and $B^{\rm th}_{\rm CIB}$. Note that when $B^{\rm th}_{\rm TSZ}$ and $B^{\rm th}_{\rm CIB}$ are equal to or greater than the

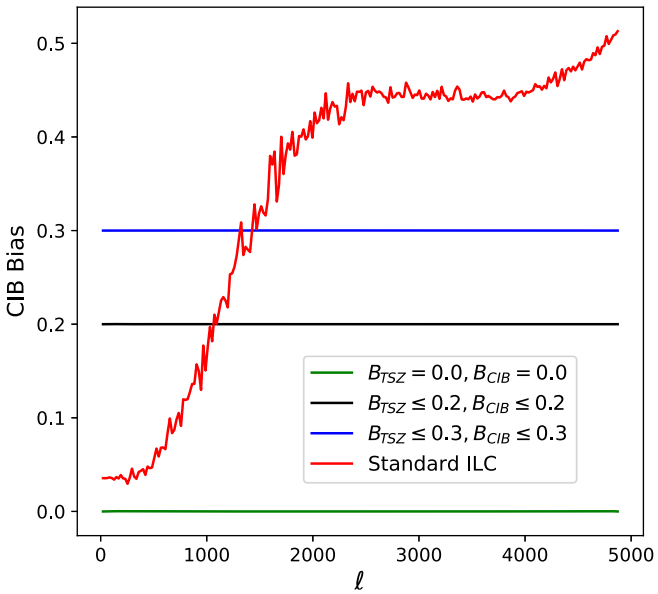


FIG. 7. Partial deprojection of both TSZ and CIB components. The residual CIB power in the coadded maps, measured relative to the power of the CIB at 145 GHz. The results shown here are obtained for the simplified sky simulations described in Sec. III E, in which the CIB field is comprised of a single component that is simply rescaled in frequency using Eq. (16). In contrast to Fig. 5, the CIB bias now behaves as expected, which demonstrates that the behavior seen previously was due to CIB decorrelation and SED variations.

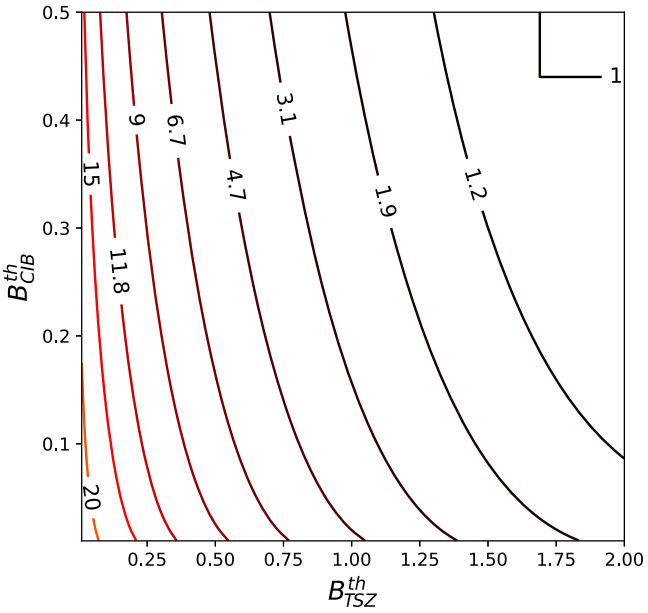


FIG. 8. Contour plot of the ratio of the total CMB map power spectra obtained with PCILC for two components (TSZ and CIB) to the spectrum of the total CMB obtained with standard ILC, at $\ell = 3500$ as a function of $B^{\rm th}_{\rm TSZ}$ and $B^{\rm th}_{\rm CIB}$.

corresponding standard ILC bias values, the ratio becomes equal to one, as expected, since the standard ILC weights are recovered in this case.

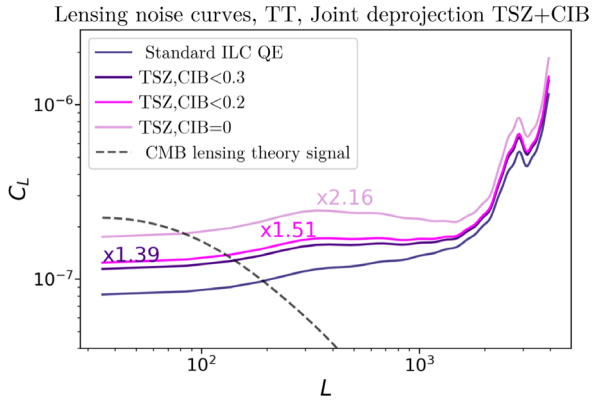
F. Application to CMB lensing reconstruction

In this subsection, we will present, as an example, an application of our new PCILC method for foreground mitigation: foreground reduction for CMB lensing analysis.

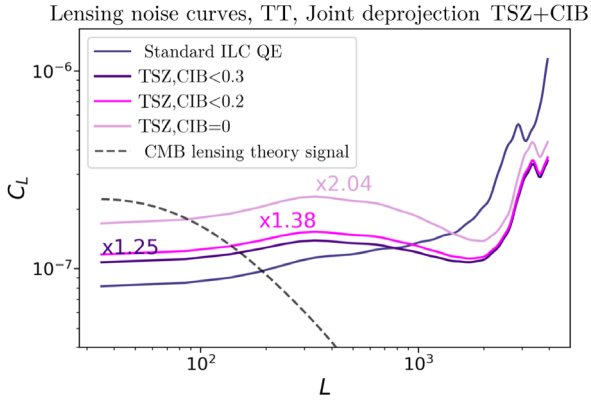
Along their paths to our telescopes, CMB photons are deflected, or lensed, by the gravitational influence of matter in our Universe. CMB lensing measurements allow us to constrain key cosmological parameters, such as the equation of state of dark energy, the sum of neutrino masses, or the amplitude of density fluctuations (e.g., [28–30]).

Thanks to high-resolution, low-noise CMB surveys, it is possible to extract the CMB lensing signal with quadratic estimators, exploiting the lensing-induced couplings between different modes of the CMB (e.g., [31]). However, these estimators are susceptible to the presence of foreground contamination in mm-wave maps, leading to potential biases in the extracted cosmological parameters [32–35]. Foregrounds are a more significant limitation for CMB temperature-derived lensing reconstruction than for polarization-derived reconstruction, as small-scale foregrounds in polarization are smaller compared to the CMB signal [36]. If CMB temperature foregrounds are left untreated, the resulting CMB lensing autospectrum and cross-spectrum analyses may be biased at the 3–20% levels, much larger than the statistical error bars (e.g., [11,32,34,37–39]). As many current- and next-generation lensing maps will still depend to a large extent on temperature data, rather than on polarization, it is important to mitigate these foregrounds for lensing analyses.

To mitigate foreground contamination in CMB lensing maps, different methods exist; these broadly divide into geometrical (e.g., [33,35,40]) and multifrequency methods (e.g., [11,18,36]). Here we will focus on the latter, although it is worth keeping in mind that this approach cannot mitigate foreground biases with the same SED as the primary CMB [34]. In particular, we will use the multifrequency symmetric cleaned estimator presented in



(a) Case with $\ell_{\rm max} = 3000$ for reconstruction with the standard quadratic and multi-frequency symmetric cleaned lensing estimators.



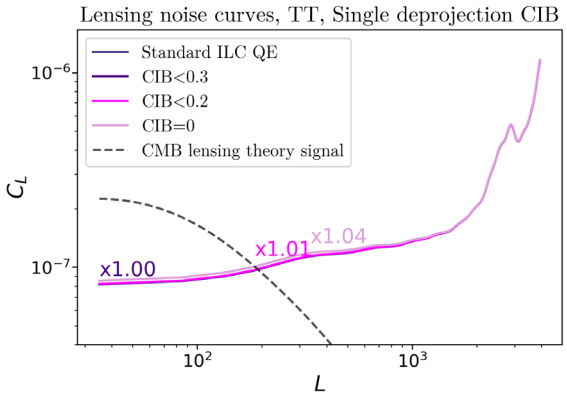
(b) Case with $\ell_{\rm max} = 3500$ for reconstruction with the standard quadratic and multi-frequency symmetric cleaned lensing estimators.

FIG. 9. (Partial) deprojection of both TSZ and CIB in CMB maps used for lensing reconstruction, using the temperature CMB noise curves in Fig. 5. In both panels (a) and (b) the lensing noise curves are shown for four cases: standard ILC CMB maps used in a quadratic lensing estimator; a CMB map cleaned with CILC, used in a quadratic lensing estimator; and PCILC-cleaned CMB maps, again used in a quadratic lensing estimator. The multifrequency-cleaned maps are labeled with a number that represents the average increase in noise with respect to lensing reconstruction using the standard ILC CMB map. It can be seen that the PCILC gives a significant noise reduction on large scales of around 30%, when compared with CILC foreground mitigation methods.

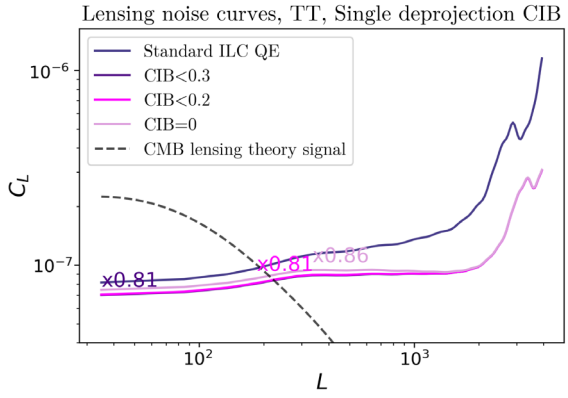
Ref. [18] (building on the estimator of Ref. [11]), which has the advantage of degrading the S/N of a CMB lensing measurement less than many other methods, while still significantly reducing the foreground-induced contamination.

In Ref. [18] this method was applied to temperature TSZ-deprojected CILC maps, constructed using data from the Atacama Cosmology Telescope and the *Planck* satellite [8]. This lead to a reduction of the TSZ contamination in the reconstructed CMB lensing map, albeit at the price of somewhat decreased S/N in the measurement, when compared to the case without any deprojection. For current CMB lensing goals, a complete nulling of the foreground contamination is likely unnecessary; given that foreground biases are already quite small (a few percent of the signal), only a substantial reduction of the CMB lensing foreground contamination is likely required in order for it to be negligible. The PCILC method allows for the required reduction in the foreground contamination while improving the S/N relative to complete deprojection with a CILC.

We illustrate this S/N improvement from the use of the PCILC in Figs. 9–11, where temperature CMB lensing reconstruction noise curves arising from the ILC, CILC, and PCILC temperature CMB maps are shown. We apply the standard quadratic estimator to the ILC map, and apply the method of Ref. [18] to the CILC and PCILC maps. The CMB modes we use for reconstruction have $\ell_{\rm min} = 30$; in addition, we use $\ell_{\rm max} = 3000$ for all the CMB lensing estimators; we also show results with $\ell_{\rm max} = 3500$ for the multifrequency cleaned ones. Figure 9 shows temperature estimator CMB lensing noise curves when constraining both TSZ and CIB (from Fig. 5), Fig. 10 shows constraints

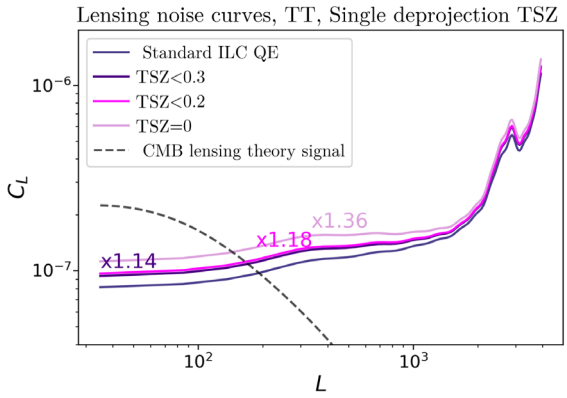


(a) Case with $\ell_{\rm max} = 3000$ for reconstruction with the standard quadratic and multi-frequency symmetric cleaned lensing estimators.

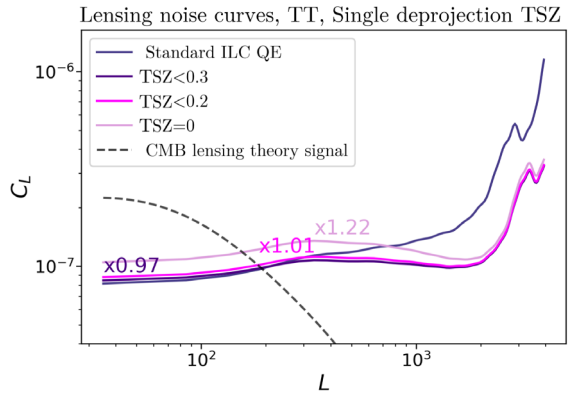


(b) Case with $\ell_{\rm max} = 3500$ for reconstruction with the standard quadratic and multi-frequency symmetric cleaned lensing estimators.

FIG. 10. As for Fig. 9, but (partially) deprojecting only CIB, using the temperature CMB noise curves in Fig. 3. In this case there are not relevant CMB lensing noise improvements between CILC and PCILC, as CIB deprojection does not lead to a huge blowing up in CMB temperature noise for the CMB scales of lensing reconstruction.



(a) Case with $\ell_{\rm max} = 3000$ for reconstruction with the standard quadratic and multi-frequency symmetric cleaned lensing estimators.



(b) Case with $\ell_{\rm max} = 3500$ for reconstruction with the standard quadratic and multi-frequency symmetric cleaned lensing estimators.

FIG. 11. As for Fig. 9, but (partially) deprojecting only TSZ, using the temperature CMB noise curves in Fig. 1. On large scales, the PCILC derived CMB lensing noise performs better than CILC one at around 10%.

on only CIB (from Fig. 3), and Fig. 11 only on TSZ (from Fig. 1). It is clear that in terms of CMB lensing noise, the PCILC outperforms the CILC by a factor of around 10–30%, even though the foreground bias is expected to be reduced to a level that is negligible for current and upcoming experiments. Multifrequency cleaning with the PCILC method is therefore a powerful technique for mitigation of foreground biases in CMB lensing, while minimizing degradation in signal-to-noise.[3]

IV. CONCLUSIONS

In this work, we have developed a new tool for CMB foreground cleaning—the PCILC method. This method finds the minimum-variance linear combination of different frequency channels in cases where residual foreground biases must be controlled to be below a threshold value but do not need to strictly be nulled. By allowing for, in many cases, negligibly small but nonzero foreground residuals, this method provides significant reductions in variance—often by factors of 2–3—when compared with a constrained ILC in which foregrounds are strictly nulled. We test and validate our method using realistic SO-like simulations, finding that we can reproduce the expected, forecast performance. Our method can be easily applied to current and upcoming CMB surveys, and has several possible applications; as an example, we show that it is capable of mitigating foreground biases in CMB lensing at lower noise than previous multifrequency methods.

ACKNOWLEDGMENTS

We thank Mathieu Remazeilles for useful discussions. S. A. acknowledges support from the Energetic Cosmos Laboratory. B. D. S. acknowledges support from a European Research Council (ERC) Starting Grant under the European Unions Horizon 2020 research and innovation programme (Grant Agreement No. 851274) and from an STFC Ernest Rutherford Fellowship. J. C. H. thanks the Simons Foundation for support. O. D. thanks the STFC.

[3]Assessing the precise impact of our method on CMB lensing power spectrum foreground biases requires simulation-based calculations and is beyond the scope of our paper. Nevertheless, we can make an approximate estimate for the reduction in foreground biases in the standard quadratic power spectrum estimator, assuming the CMB lensing filters do not change and neglecting any scale dependence in ϵ. Depending on the lensing scale, the lensing foreground biases generally are dominated by either the primary terms, which have the form $\langle \kappa T_f T_f \rangle \propto \epsilon$, or the trispectrum terms, which have the form $\langle T_f T_f T_f T_f \rangle \propto \epsilon^2$; here T_f is the foreground contribution to the temperature map. These scalings with powers of T_f imply that lensing biases should therefore be suppressed by at least a factor of ϵ (although on small scales, where the trispectrum term dominates, a more significant reduction by a factor ϵ^2 can be expected.) The original, unmitigated biases are typically less than ∼5% [32] of the signal power for conventional analysis choices; therefore, a reduction by even a moderate factor $\epsilon \sim \leq 0.2$ may be sufficient to reduce the biases to a sub-percent level, which is negligible for current-generation experiments.

APPENDIX A: CIB BIAS

In this section, we discuss why we find problems with the CIB bias obtained from the PCILC and CILC methods implemented for the simultaneous deprojection of two components.

As mentioned in Sec. III, these problems arise due to slight decorrelation of the CIB across frequencies and the imperfect SED model of the CIB signal. To understand why this is so, let us first express the CIB cross-frequency power spectra as follows:

$$F^{\mathrm{CIB}}_{\nu\times\nu'} = r^{\mathrm{CIB}}_{\nu\times\nu'}\sqrt{F^{\mathrm{CIB}}_{\nu\times\nu}F^{\mathrm{CIB}}_{\nu'\times\nu'}} \tag{A1}$$

where $r^{\mathrm{CIB}}_{\nu\times\nu'}$ is the correlation coefficient between ν and ν' frequency channels, and all auto- and cross-frequency spectra obtained are from the simulations. Next, we split this equation into two parts,

$$F^{\mathrm{CIB}}_{\nu\times\nu'} = \sqrt{F^{\mathrm{CIB}}_{\nu\times\nu}F^{\mathrm{CIB}}_{\nu'\times\nu'}} + (r^{\mathrm{CIB}}_{\nu\times\nu'} - 1)\sqrt{F^{\mathrm{CIB}}_{\nu\times\nu}F^{\mathrm{CIB}}_{\nu'\times\nu'}}, \tag{A2}$$

where the first part is decorrelation free and the second is the decorrelation estimate. Since the off-diagonal values of the covariance matrix $\mathbf{F}^{\mathrm{CIB}}$ consist of cross-frequency spectra, by analogy we can express the covariance matrix as follows:

$$\mathbf{F}^{\mathrm{CIB}} = \mathbf{q}\mathbf{q}^T + \mathbf{q}\mathbf{q}^T \circ (\mathbf{r} - \mathbf{1}) = \mathbf{Q} + \mathbf{D}, \tag{A3}$$

where $\mathbf{q}^T = \left[\sqrt{F^{\mathrm{CIB}}_{\nu_1\times\nu_1}}\sqrt{F^{\mathrm{CIB}}_{\nu_2\times\nu_2}}\cdots\sqrt{F^{\mathrm{CIB}}_{\nu_{N_\nu}\times\nu_{N_\nu}}}\right]$, $\mathbf{r}$ is the $N_\nu \times N_\nu$ matrix where each element corresponds to the correlation coefficient between frequency channels, and $\mathbf{1}$ is the $N_\nu \times N_\nu$ matrix where each element is equal to one. Then, using Eq. (17), the CIB bias can be decomposed as follows:

$$B_{\mathrm{CIB}} = \frac{\mathbf{w}^T\mathbf{Q}\mathbf{w} + \mathbf{w}^T\mathbf{D}\mathbf{w}}{F^{\mathrm{CIB}}_{145\times145}}. \tag{A4}$$

The first part of Eq. (A4) corresponds to the case where the CIB maps are perfectly correlated, so we can use it as a SED test:

$$\Delta B_{\mathrm{SED}} = \frac{\mathbf{w}^T\mathbf{Q}\mathbf{w}}{F^{\mathrm{CIB}}_{145\times145}} - \frac{\mathbf{w}^T\mathbf{b}\mathbf{b}^T\mathbf{w}}{b^2_{145}} \tag{A5}$$

where $\mathbf{b}$ is the model CIB SED [see Eq. (16)], an $N_\nu \times 1$ vector, and b_{145} is the model CIB SED evaluated at 145 GHz. In Fig. 12 (left), we show how the suboptimal CIB SED affects the CIB bias error.

The second part of the Eq. (A4) shows how decorrelation affects the CIB bias:

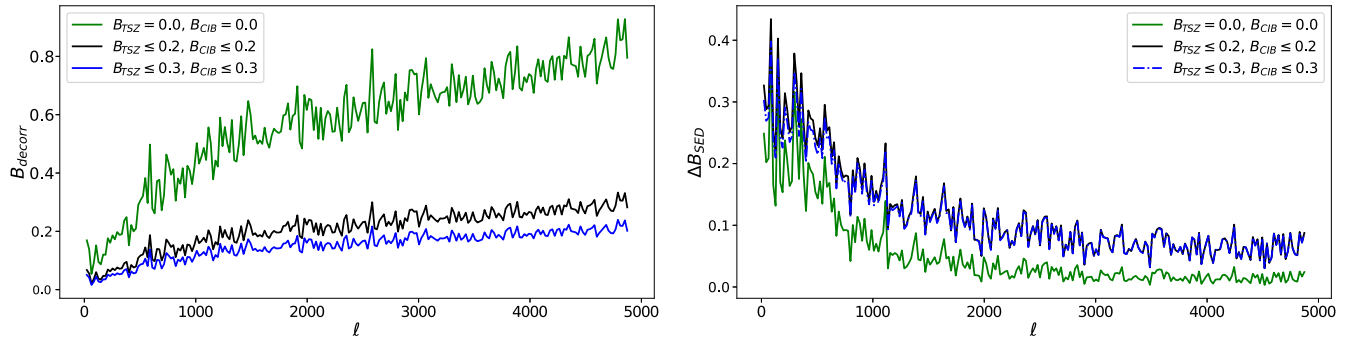


FIG. 12. Effect of decorrelation between frequency maps (right) and suboptimal CIB spectral response model (left) on CIB bias.

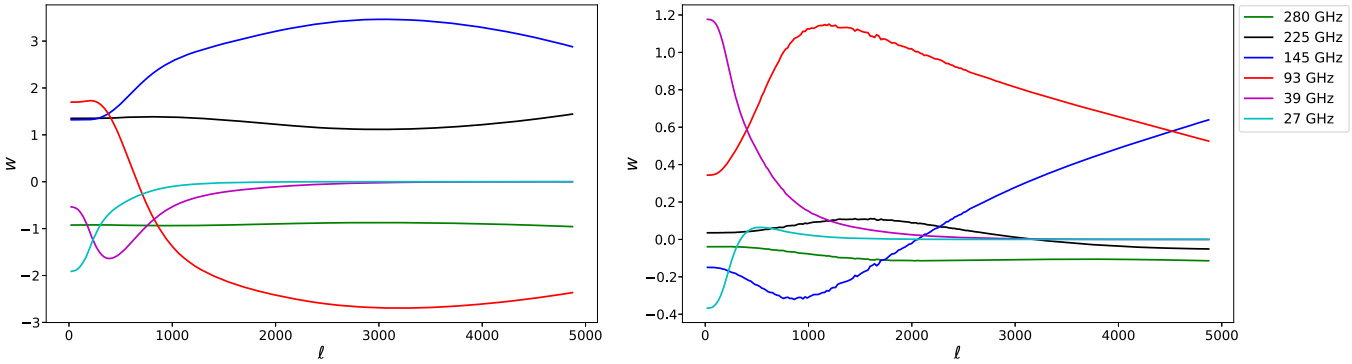


FIG. 13. Weight values for each frequency channel. Left: weights obtained from CILC for two-component (CIB and TSZ) deprojection. Right: weights obtained for CIB deprojection.

$$B_{\rm decorr} = \frac{\mathbf{w}^T \mathbf{D} \mathbf{w}}{F^{\rm CIB}_{145\times145}}. \tag{A6}$$

In Fig. 12 (right) we can see how decorrelation affects the CIB bias error. As expected, the decorrelation errors grow with ℓ, as the CIB decorrelation itself does (e.g., [41]).

In Fig. 3 (lr), we can see that the CIB bias residual when reducing or removing only one component is much lower compared with the CIB bias when reducing or removing two components. Since Eq. (A4) also applies for single-inequality ILC, the only difference is in the weights, which are shown in Fig. 13. An explanation for the imperfect CIB removal is therefore that the high weight values when deprojecting two components amplify small errors in the CIB model and small amounts of decorrelation, leading to significant biases.

APPENDIX B: MULTIPLE PARTIALLY CONSTRAINED ILC

Suppose we have N_f frequencies, N_c components, with $N_c - 1$ foregrounds. Suppose we would like to constrain P foregrounds.

Let us define a few quantities: consider a "projection" operator M described by a $2P \times P$ matrix,

$$M \equiv \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & \ldots & 1 \\ 0 & 0 & 0 & 0 & \ldots & \ldots & 1 \end{pmatrix} \tag{B1}$$

and another $2P \times P$ operator N,

$$N \equiv \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ -1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 0 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & \ldots & 1 \\ 0 & 0 & 0 & 0 & \ldots & \ldots & -1 \end{pmatrix}. \tag{B2}$$

Also define a $2P$-dimensional vector containing the "slack" variables,

$$\vec{s} \equiv (\, s_{11} \quad s_{12} \quad \ldots \quad \ldots \quad s_{j1} \quad s_{j2} \quad \ldots \quad \ldots \quad s_{P1} \quad s_{P2} \,)^T \tag{B3}$$

and another one for their squares

$$\vec{s}_2 \equiv (\, s_{11}^2 \quad s_{12}^2 \quad \ldots \quad \ldots \quad s_{j1}^2 \quad s_{j2}^2 \quad \ldots \quad \ldots \quad s_{P1}^2 \quad s_{P2}^2 \,)^T, \tag{B4}$$

where j refers to the jth foreground to be deprojected.

Note that we can write this vector as

$$\vec{s}_2^T = \sum_k (\mathbf{P}_k \vec{s})^T (\mathbf{P}_k \vec{s}) \vec{e}_k^T, \tag{B5}$$

where $\vec{e}_k$ is an orthonormal basis vector and $\mathbf{P}_k$ is a projection matrix.

Also consider an $N_f \times P$ matrix defining the foregrounds to be constrained,

$$F \equiv (\, \vec{f}^1 \quad \vec{f}^2 \quad \ldots \quad \vec{f}^P \,), \tag{B6}$$

with $\vec{f}^j$ the $N_{\rm f}$-dimensional vector where $(\vec{f}^j)_i$ is the foreground component j SED at frequency i.

And finally define the constraints vector

$$\vec{\epsilon} \equiv (\, \epsilon_1 \quad \epsilon_2 \quad \ldots \quad \epsilon_P \,)^T. \tag{B7}$$

Then we can write a Lagrangian

$$\mathcal{L} \equiv \vec{w}^T R \vec{w} + \lambda (1 - \vec{w}^T \vec{e}) + \vec{\lambda}_{\rm ineq}^T M \vec{\epsilon} + \vec{\lambda}_{\rm ineq}^T N F^T \vec{w} - \vec{\lambda}_{\rm ineq}^T \vec{s}_2. \tag{B8}$$

To solve for the weights, we then have to look at the surfaces of minimum functional

$$\begin{cases} \vec{\nabla}_{\vec{w}^T} \mathcal{L} = 2R\vec{w} - \lambda \vec{e} + F N^T \vec{\lambda}_{\rm ineq} = \vec{0} \\ \frac{\partial \mathcal{L}}{\partial \lambda} = 1 - \vec{e}^T \vec{w} = 0 \\ \vec{\nabla}_{\vec{\lambda}_{\rm ineq}^T} \mathcal{L} = M\vec{\epsilon} + N F^T \vec{w} - \vec{s}_2 = \vec{0} \\ \vec{\nabla}_{\vec{s}^T} \mathcal{L} = -2 \sum_k (\mathbf{P}_k^T \mathbf{P}_k \vec{s}) \vec{e}_k^T \vec{\lambda}_{\rm ineq} = \vec{0} \end{cases}. \tag{B9}$$

If the constraints are active, for the $\lambda_{{\rm ineq},jl}$ multiplier, then we have equality constraint, otherwise if the constraints are nonactive, then we reduce to the standard ILC, as this is allowed.

[1] C. L. Bennett *et al.*, First-year Wilkinson microwave anisotropy probe (WMAP) observations: Foreground emission, Astrophys. J. Suppl. Ser. **148**, 97 (2003).

[2] M. Tegmark, A. de Oliveira-Costa, and A. J. S. Hamilton, High resolution foreground cleaned CMB map from WMAP, Phys. Rev. D **68**, 123523 (2003).

[3] 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, Astrophys. J. **612**, 633 (2004).

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

[5] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown *et al.* (Planck Collaboration), Planck 2013 results. XXI. Power spectrum and high-order statistics of the Planck all-sky Compton parameter map, Astron. Astrophys. **571**, A21 (2014).

[6] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown *et al.* (Planck Collaboration), Planck 2013 results. XII. Diffuse component separation, Astron. Astrophys. **571**, A12 (2014).

[7] N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi *et al.* (Planck Collaboration), Planck 2015 results. XXII. A map of the thermal Sunyaev-Zeldovich effect, Astron. Astrophys. **594**, A22 (2016).

[8] M. S. Madhavacheril, J. C. Hill, S. Næss, G. E. Addison, S. Aiola, T. Baildon *et al.*, Atacama Cosmology Telescope: Component-separated maps of CMB temperature and the thermal Sunyaev-Zel'dovich effect, Phys. Rev. D **102**, 023534 (2020).

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

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

[11] M. S. Madhavacheril and J. C. Hill, Mitigating foreground biases in CMB lensing reconstruction using cleaned gradients, Phys. Rev. D **98**, 023534 (2018).

[12] T. Chen, M. Remazeilles, and C. Dickinson, Impact of SZ cluster residuals in CMB maps and CMB-LSS cross-correlations, Mon. Not. R. Astron. Soc. **479**, 4239 (2018).

[13] J. C. Hill, Foreground biases on primordial non-Gaussianity measurements from the CMB temperature bispectrum: Implications for Planck and beyond, Phys. Rev. D **98**, 083542 (2018).

[14] Y. Akrami *et al.*, Planck 2018 results. IX. Constraints on primordial non-Gaussianity, Astron. Astrophys. **641**, A9 (2020).
[15] J. C. Hill, S. Ferraro, N. Battaglia, J. Liu, and D. N. Spergel, Kinematic Sunyaev-Zel'dovich Effect with Projected Fields: A Novel Probe of the Baryon Distribution with Planck, WMAP, and WISE Data, Phys. Rev. Lett. **117**, 051301 (2016).
[16] N. Aghanim *et al.*, Planck intermediate results. LIII. Detection of velocity dispersion from the kinetic Sunyaev-Zeldovich effect, Astron. Astrophys. **617**, A48 (2018).
[17] Y. Akrami *et al.*, Planck 2018 results. VII. Isotropy and statistics of the CMB, Astron. Astrophys. **641**, A7 (2020).
[18] O. Darwish *et al.*, The Atacama Cosmology Telescope: A CMB lensing mass map over 2100 square degrees of sky and its cross-correlation with BOSS-CMASS galaxies, Mon. Not. R. Astron. Soc. **500**, 2250 (2020).
[19] M. Remazeilles, A. Rotti, and J. Chluba, Peeling off foregrounds with the constrained moment ILC method to unveil primordial CMB *B*-modes, arXiv:2006.08628.
[20] MA 1024 Lagrange multipliers for inequality constraints, https://users.wpi.edu/~pwdavis/Courses/MA1024B10/1024_Lagrange_multipliers.pdf.
[21] https://lambda.gsfc.nasa.gov/simulation/tb_sim_ov.cfm
[22] P. Ade *et al.*, The Simons Observatory: Science goals and forecasts, J. Cosmol. Astropart. Phys. 02 (2019) 056.
[23] N. Sehgal, P. Bode, S. Das, C. Hernandez-Monteagudo, K. Huffenberger, Y.-T. Lin, J. P. Ostriker, and H. Trac, Simulations of the microwave sky, Astrophys. J. **709**, 920 (2010).
[24] Y. B. Zeldovich and R. A. Sunyaev, The interaction of matter and radiation in a hot-model universe, Astrophys. Space Sci. **4**, 301 (1969).
[25] R. A. Sunyaev and Y. B. Zeldovich, Small-Scale Fluctuations of Relic Radiation, Astrophys. Space Sci. **7**, 3 (1970), https://link.springer.com/article/10.1007/BF00653471.
[26] S. Nozawa, N. Itoh, Y. Suda, and Y. Ohhata, An improved formula for the relativistic corrections to the kinematical Sunyaev-Zeldovich effect for clusters of galaxies, Nuovo Cimento B Ser. **121**, 487 (2006).
[27] J. Chluba, J. C. Hill, and M. H. Abitbol, Rethinking CMB foregrounds: Systematic extension of foreground parametrizations, Mon. Not. R. Astron. Soc. **472**, 1195 (2017).
[28] B. D. Sherwin, A. van Engelen, N. Sehgal, M. Madhavacheril, G. E. Addison, S. Aiola *et al.*, Two-season Atacama Cosmology Telescope polarimeter lensing power spectrum, Phys. Rev. D **95**, 123529 (2017).
[29] W. L. K. Wu, L. M. Mocanu, P. A. R. Ade, A. J. Anderson, J. E. Austermann, J. S. Avva *et al.*, A measurement of the cosmic microwave background lensing potential and power spectrum from 500 deg^2 of SPTpol temperature and polarization data, Astrophys. J. **884**, 70 (2019).
[30] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi *et al.* (Planck Collaboration), Planck 2018 results. VIII. Gravitational lensing, Astron. Astrophys. **641**, A8 (2020).
[31] W. Hu and T. Okamoto, Mass reconstruction with cosmic microwave background polarization, Astrophys. J. **574**, 566 (2002).
[32] A. van Engelen, S. Bhattacharya, N. Sehgal, G. P. Holder, O. Zahn, and D. Nagai, CMB lensing power spectrum biases from galaxies and clusters using high-angular resolution temperature maps, Astrophys. J. **786**, 13 (2014).
[33] S. J. Osborne, D. Hanson, and O. Doré, Extragalactic foreground contamination in temperature-based CMB lens reconstruction, J. Cosmol. Astropart. Phys. 03 (2014) 024.
[34] S. Ferraro and J. C. Hill, Bias to CMB lensing reconstruction from temperature anisotropies due to large-scale galaxy motions, Phys. Rev. D **97**, 023512 (2018).
[35] E. Schaan and S. Ferraro, Foreground-Immune Cosmic Microwave Background Lensing with Shear-Only Reconstruction, Phys. Rev. Lett. **122**, 181301 (2019).
[36] D. Beck, J. Errard, and R. Stompor, Impact of polarized galactic foreground emission on CMB lensing reconstruction and delensing of B-modes, J. Cosmol. Astropart. Phys. 06 (2020) 030.
[37] Y. Omori *et al.* (DES and SPT Collaborations), Dark Energy Survey Year 1 Results: Tomographic cross-correlations between Dark Energy Survey galaxies and CMB lensing from South Pole Telescope + Planck, Phys. Rev. D **100**, 043501 (2019).
[38] E. J. Baxter *et al.* (DES and SPT Collaborations), Dark Energy Survey Year 1 results: Methodology and projections for joint analysis of galaxy clustering, galaxy lensing, and CMB lensing two-point functions, Phys. Rev. D **99**, 023508 (2019).
[39] N. Sailer, E. Schaan, and S. Ferraro, Lower bias, lower noise CMB lensing with foreground-hardened estimators, Phys. Rev. D **102**, 063517 (2020).
[40] T. Namikawa, D. Hanson, and R. Takahashi, Bias-hardened CMB lensing, Mon. Not. R. Astron. Soc. **431**, 609 (2013).
[41] D. S. Y. Mak, A. Challinor, G. Efstathiou, and G. Lagache, Measurement of CIB power spectra over large sky areas from Planck HFI maps, Mon. Not. R. Astron. Soc. **466**, 286 (2017).