



This is the accepted manuscript made available via CHORUS. The article has been published as:

CMB internal delensing with general optimal estimator for higher-order correlations

Toshiya Namikawa

Phys. Rev. D **95**, 103514 — Published 24 May 2017

DOI: [10.1103/PhysRevD.95.103514](https://doi.org/10.1103/PhysRevD.95.103514)

CMB internal delensing with general optimal estimator for higher-order correlations

Toshiya Namikawa^{1,2}

¹*Department of Physics, Stanford University, Stanford, California 94305, USA*

²*Kavli Institute for Particle Astrophysics and Cosmology, SLAC National Accelerator Laboratory,
2575 Sand Hill Road, Menlo Park, California 94025, USA*

(Dated: May 4, 2017)

We present a new method for delensing B modes of the cosmic microwave background (CMB) using a lensing potential reconstructed from the same realization of the CMB polarization (CMB internal delensing). The B -mode delensing is required to improve sensitivity to primary B modes generated by, e.g., the inflationary gravitational waves, axion-like particles, modified gravity, primordial magnetic fields, and topological defects such as cosmic strings. However, the CMB internal delensing suffers from substantial biases due to correlations between observed CMB maps to be delensed and that used for reconstructing a lensing potential. **Since the bias depends on realizations, we construct a realization-dependent (RD) estimator for correcting these biases by deriving a general optimal estimator for higher-order correlations. The RD method is less sensitive to simulation uncertainties. Compared to the previous ℓ -splitting method, we find that the RD method corrects the biases without substantial degradation of the delensing efficiency.**

I. INTRODUCTION

The cosmic microwave background (CMB) anisotropies are one of the best probe in cosmology. B modes of the CMB polarization could be generated by the inflationary gravitational waves, axion-like particles, primordial magnetic fields, modified gravity, and topological defects such as cosmic strings (e.g., Refs. [1–8]). In addition, the gravitational lensing converts some of the E modes into B modes [9], and this effect has been measured by multiple CMB experiments (e.g., Refs. [10–14]). The recent precise measurement of the CMB polarization by the BICEP2/Keck Array experiment indicates that the observed B modes are dominated by the gravitational lensing B modes [14]. Thus, removal of the lensing-induced B modes, usually referred to as *delensing*, will be soon required to improve sensitivity to the primary B modes [15, 16]. The temperature and E -mode delensing are also important to constrain, e.g., the effective number of the relativistic species [17].

Several works have discussed methods to remove the lensing contributions in the CMB anisotropies. Refs. [18, 19] provide a method to construct a template of the lensed CMB fluctuations and to remove it from the observed CMB fluctuations. Ref. [17] proposes a delensing technique to remove the lensing-induced CMB anisotropies including the higher-order terms of the lensing potential. Refs. [20, 21] show the delensing method using the cosmic infrared background (CIB) (and this technique is demonstrated by Refs. [22, 23]).

The delensing methods, however, suffer from biases if the CMB maps are delensed with a lensing potential derived from the CMB maps (CMB internal delensing). In the CMB internal delensing, we use the same realization of the observed CMB anisotropies for the delensing and reconstruction of the lensing potential. Therefore, the CMB anisotropies to be delensed correlates with that used in the lensing potential reconstruction. As we dis-

cuss latter in Sec. II, this correlation substantially biases the power and scatter of the delensed CMB spectrum (delensing bias).

Methods to correct the delensing bias have been explored in several works. The delensing bias is pointed out by Ref. [8] and is characterized by Ref. [24] (see also Ref. [25] for the CMB temperature and Ref. [26] for the Planck cases). Refs. [8, 25] proposed a method to mitigate the delensing bias by splitting CMB anisotropies into annuli in multipole space so that CMB multipoles to be delensed and that used for the lensing reconstruction are uncorrelated. In this method, however, the efficiency of the delensing is degraded because we lose the CMB multipoles in the lensing reconstruction and precision of the reconstructed lensing potential decreases.

In this paper, we study an alternative way of correcting the delensing bias to avoid degradation of the delensing efficiency and to estimate the delensed spectrum more accurately than using the simulation alone. To construct an estimator for the delensing bias, we use the similar analogy of the “realization-dependent” (RD) method in the lensing reconstruction [27–29]. In the reconstruction of the lensing potential, the leading-order bias in estimating the lensing-potential power spectrum comes from the disconnected part of the observed CMB trispectrum. In the RD method, the bias is estimated by combining both simulated and observed CMB fluctuations so that the estimator is optimal and less sensitive to the error in the simulated CMB covariance. The RD method in the lensing reconstruction is motivated by the optimal trispectrum estimator [30].

Compared to the lensing reconstruction, the delensing bias contains higher order correlations such as six point correlations. Thus, we first generalize the optimal trispectrum estimator to the case with high-order correlations (polyspectra). Then we employ the optimal polyspectra estimator to construct RD estimators for the delensing bias. We demonstrate that the RD method removes the delensing bias without degradation of the de-

lensing efficiency. In this paper, for simplicity, we derive the RD method in the flat-sky approximation but the extension to the full sky approach is straightforward.

This paper is organized as follows. Sec. II reviews the B -mode delensing and the delensing bias. Sec. IV derives an optimal estimator for polyspectra based on the Edgeworth expansion of the CMB likelihood. Sec. V provides the RD estimators for the delensing bias and shows results of numerical simulation. Sec. VI is devoted to summary and discussion.

II. CMB INTERNAL DELENSING

Here we briefly summarize notations in this paper and methods for the lensing reconstruction and delensing.

A. CMB polarization

The lensing effect on the CMB anisotropies is expressed by a remapping of the primary CMB anisotropies. Denoting the primary CMB fluctuations at position \hat{n} on the last scattering surface as $X(\hat{n})$, where X is either the CMB temperature Θ , or polarization anisotropies $Q \pm iU$, the lensed CMB anisotropies in a direction \hat{n} , are given by (e.g., Ref. [31])

$$\tilde{X}(\hat{n}) = X(\hat{n} + \nabla\phi(\hat{n})). \quad (1)$$

Here the two-dimensional vector, $\nabla\phi(\hat{n})$, is the deflection angle. The Fourier modes of the temperature anisotropies, and the E and B modes are defined as

$$\Theta_\ell = \int d^2\hat{n} e^{-i\hat{n}\cdot\ell} \Theta(\hat{n}), \quad (2)$$

$$E_\ell \pm iB_\ell = - \int d^2\hat{n} e^{-i\hat{n}\cdot\ell} [Q \pm iU](\hat{n}) e^{\mp 2i\varphi_\ell}, \quad (3)$$

where φ_ℓ is the angle of ℓ measured from the x -axis. Substituting Eq. (1) into the above equation gives the following expression of the lensed CMB anisotropies up to first order of the lensing potential [32]:

$$\tilde{\Theta}_\ell = \Theta_\ell - \int \frac{d^2\mathbf{L}}{(2\pi)^2} \mathbf{L} \cdot (\ell - \mathbf{L}) \phi_{\mathbf{L}} \Theta_{\ell-\mathbf{L}}, \quad (4)$$

$$\tilde{E}_\ell = E_\ell - \int \frac{d^2\mathbf{L}}{(2\pi)^2} \phi_{\mathbf{L}} (E_{\ell-\mathbf{L}} z_{\ell\mathbf{L}} - B_{\ell-\mathbf{L}} w_{\ell\mathbf{L}}), \quad (5)$$

$$\tilde{B}_\ell = B_\ell - \int \frac{d^2\mathbf{L}}{(2\pi)^2} \phi_{\mathbf{L}} (B_{\ell-\mathbf{L}} z_{\ell\mathbf{L}} + E_{\ell-\mathbf{L}} w_{\ell\mathbf{L}}), \quad (6)$$

where the weight functions are given by

$$z_{\ell\mathbf{L}} = \mathbf{L} \cdot (\ell - \mathbf{L}) \cos 2(\varphi_{\ell-\mathbf{L}} - \varphi_\ell), \quad (7)$$

$$w_{\ell\mathbf{L}} = \mathbf{L} \cdot (\ell - \mathbf{L}) \sin 2(\varphi_{\ell-\mathbf{L}} - \varphi_\ell). \quad (8)$$

If the primordial B modes are ignored, the B modes induced by gravitational lensing are expressed as a convolution of the lensing potential and primary E modes.

B. Lensing reconstruction

Precision of the lensing reconstruction in the future will be almost determined by the EB quadratic estimator, and we focus on the case using the EB estimator. For a fixed realization of the lensing potential in the universe, the lensing effect on CMB maps produces off-diagonal elements of the CMB covariance. From Eq. (6), the off-diagonal elements become [32]

$$\langle \tilde{E}_\ell \tilde{B}_{\mathbf{L}-\ell} \rangle_{\text{CMB}} = -\tilde{C}_\ell^{\text{EE}} w_{\mathbf{L}-\ell, \mathbf{L}} \phi_{\mathbf{L}}, \quad (9)$$

where $\langle \cdots \rangle_{\text{CMB}}$ denotes the ensemble average over unlensed CMB anisotropies with a fixed realization of the lensing potential ϕ . We ignore the higher-order terms of the lensing potential and contributions from primary B modes¹. $\tilde{C}_\ell^{\text{EE}}$ is the lensed E -mode spectrum to mitigate the higher-order terms of the lensing-potential power spectrum [33, 34]. The lensing estimators are then described as [32]

$$\hat{\phi}_{\mathbf{L}}^{EB} = A_L \int \frac{d^2\ell}{(2\pi)^2} \frac{-\tilde{C}_\ell^{\text{EE}} w_{\mathbf{L}-\ell, \mathbf{L}}}{\hat{C}_\ell^{\text{EE}} \hat{C}_{|\mathbf{L}-\ell|}^{\text{BB}}} \hat{E}_\ell \hat{B}_{\mathbf{L}-\ell}. \quad (10)$$

Here \hat{E} (\hat{B}) is the observed E (B) modes, \hat{C}^{EE} (\hat{C}^{BB}) is the observed E (B) mode spectrum, and A_L is the quadratic estimator normalization given by Ref. [32]. In practical analysis, the filtering functions, $1/\hat{C}_\ell^{\text{EE}}$ and $1/\hat{C}_\ell^{\text{BB}}$, are generalized to the inverse of the covariance matrix of E and B modes, and are determined by simulation (e.g., see Ref. [35]). Unless otherwise stated, $\hat{\phi}$ is equal to $\hat{\phi}^{EB}$ with the correction of the mean-field bias, $\langle \hat{\phi} \rangle$.

C. Delensing

A method to remove lensing B modes from observed B modes is to make a template of the lensing B modes with a measured lensing potential and E modes. Hereafter, we refer to this technique as the template method. The template of the lensing B modes up to first order of the lensing potential is defined as [19]

$$(\hat{E} \star \hat{\phi})_\ell = - \int \frac{d^2\ell'}{(2\pi)^2} \frac{C_{\ell'}^{\text{EE}}}{\hat{C}_{\ell'}^{\text{EE}}} \frac{C_{|\ell-\ell'|}^{\phi\phi}}{\hat{C}_{|\ell-\ell'|}^{\phi\phi}} w_{\ell, \ell-\ell'} \hat{E}_{\ell'} \hat{\phi}_{\ell-\ell'}. \quad (11)$$

Here the optimal weight is given as a product of the Wiener filters of the E modes and lensing potential. In

¹ Note that the presence of the primordial gravitational-wave (GW) B modes does not affect the precision of the lensing reconstruction and delensing efficiency. This is because the GW B modes are significant only at the large scale ($\ell < 100$) and the lensing potential is reconstructed from small scale B modes.

practical analysis, the denominator of the Wiener filter is determined by simulations. This expression is derived by assuming that the reconstructed lensing potential is uncorrelated with the CMB anisotropies [19]. While it would be possible to make a template including this correlation, Eq. (11) is numerically tractable and useful for delensing analysis in practical cases. A lensing B -mode template to remove the higher-order terms of ϕ is also possible to construct but Eq. (11) is enough to reproduce most of the lensing B modes [19]. The delensed B -mode power spectrum is then given by [19]

$$C_{\ell}^{\hat{B}^d \hat{B}^d} = \int \frac{d^2 \ell'}{(2\pi)^2} w_{\ell, \ell'}^2 \left(1 - \frac{(C_{\ell'}^{\text{EE}})^2 (C_{|\ell-\ell'|}^{\phi\phi})^2}{\hat{C}_{\ell'}^{\text{EE}} \hat{C}_{|\ell-\ell'|}^{\phi\phi}} \right). \quad (12)$$

Alternatively, in high sensitivity polarization experiments, the delensed B modes are obtained by the inverse remapping of the observed CMB map using the Wiener-filtered lensing potential [17]. Hereafter, we call this approach the remapping method. While the remapping method is a simple way to remove higher order terms of the lensing potential, the inverse-remapping procedure simultaneously remaps noise fluctuations which should be taken into account for a noisy experiments [26].

III. DELENSING BIAS

To delens B modes, the observed B modes to be delensed and that used in the lensing reconstruction are correlated and the delensed B modes have significant biases. Here we discuss the bias in the delensed B -mode power spectrum.

A. CMB internal delensing with simulation

To study the CMB internal delensing, we simulate the lensed CMB, noise and reconstructed lensing potential maps as follows. We first generate 1200 realizations of unlensed CMB and lensing potential maps in $10\text{deg} \times 10\text{deg}$ region as random Gaussian fields using CAMB output spectra [36]². We assume the cosmological parameters consistent with Ref. [37] without the inflationary B modes. The lensed CMB maps are obtained by remapping the unlensed CMB maps with the lensing potential [38]. The instrumental noise spectrum is computed with the formula of Ref. [39] assuming $\sqrt{2} \times 3 \mu\text{K-arcmin}$ white noise in polarization map and 1 arcmin Gaussian beam, and the noise maps are generated with this noise spectrum. The lensing potential is reconstructed from the

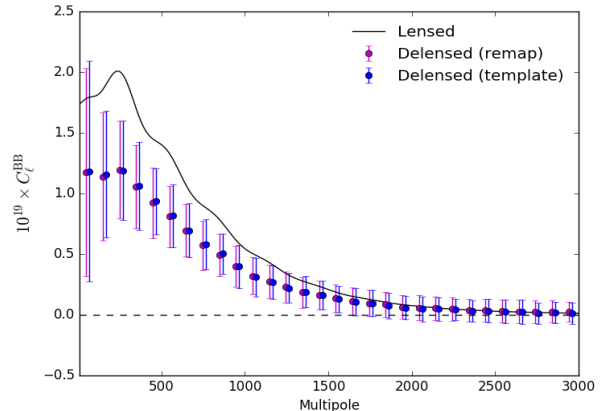


FIG. 1: Delensed B -mode power spectrum with the template and remapping methods. The solid line shows the lensing B -mode spectrum. The error bar denotes 1σ statistical error of one realization of observation in a 100deg^2 patch.

lensed+noise (observed) maps with the EB quadratic estimator using CMB multipoles in the range between 200 and 3000. We do not use the multipoles at $\ell < 200$ since the large-scale B modes are dominated by the Galactic foregrounds and are also not important in the lensing potential reconstruction. To characterize the delensing bias, we also use the input lensing potential with a random Gaussian noise whose power corresponds to the reconstruction noise³.

We then perform delensing with the above lensing potentials and observed CMB maps. In the template method, the lensing B modes are estimated from Eq. (11) and are subtracted from the observed B modes. In the remapping method, the observed Q/U maps are remapped by the Wiener-filtered lensing potential and we obtain the delensed B modes from the delensed Q/U maps. We do not apply Wiener filter for the CMB polarizations since its impact is negligible in our case.

B. Comparison between template and remapping methods

Fig. 1 shows delensed B modes with the template and remapping methods. For simplicity, we use the input lensing potential with the random Gaussian reconstruction noise since it does not create the delensing bias. The delensed B -mode power spectra using the template and remapping methods are similar since the lensing B modes

² The first 300 realizations are used as “observed” CMB fluctuations while the other 3×300 realizations are used for estimating the (RD) delensing biases.

³ The leading order reconstruction noise is expressed as the disconnected bias, but the higher-order biases are not negligible in the following discussion. Thus we denote the reconstruction noise spectrum as $C_L^{\hat{\phi}\hat{\phi}} - C_L^{\phi\phi}$.

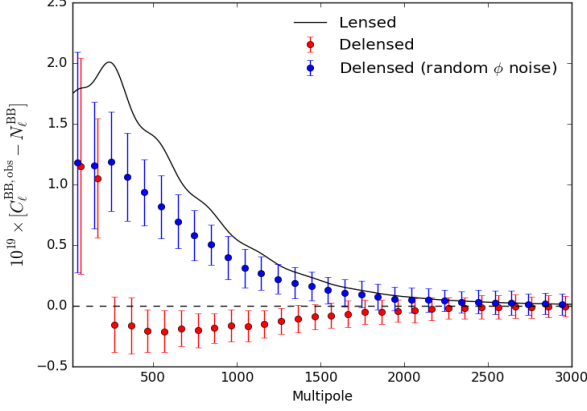


FIG. 2: Delensed B -mode power spectra using the lensing potential reconstructed from the CMB maps or the input lensing potential with the random Gaussian reconstruction noise. The instrumental noise spectrum is subtracted from the delensed spectra. The blue points are equivalent to the B -mode spectrum with the template method shown in Fig. 1. The error bars denote the 1σ statistical error for one realization of observation at a 100deg^2 patch.

are approximately expressed as Eq. (6) [40]. In the following sections, we focus on the method of the (linear order) template method.

C. Delensing bias in B mode power spectrum

Next we use the reconstructed lensing potential for delensing. Fig. 2 shows the delensed B -mode spectrum using the input lensing potential with the random Gaussian reconstruction noise or the reconstructed lensing potential.⁴ Since the delensed B -mode spectrum with the random Gaussian reconstruction noise does not contain the delensing bias, the difference of the B -mode spectra between the two cases describes the delensing bias. The discrepancy appears at $L \gtrsim 200$ and the scatter of the power spectrum is also reduced.

To understand the delensing bias, we expand the delensed B -mode spectrum into the four and six point correlations. If the lensing potential is estimated from an external data or the reconstruction noise is uncorrelated with the CMB anisotropies, the delensed B -mode power spectrum is simply equivalent to the power spectrum given in Eq. (12):

$$\langle |\hat{B}^d|^2 \rangle = \langle |\hat{B} - (\hat{E} \star \hat{\phi})|^2 \rangle = (2\pi)^2 \delta_0^D C^{\hat{B}^d \hat{B}^d}. \quad (13)$$

⁴ The instrumental noise power spectrum is subtracted from the delensed B -mode spectrum and thus the debiased spectra can be negative.

Here $\hat{\phi}'$ is given as a sum of the true lensing potential and the random Gaussian reconstruction noise, δ_0^D is the Delta function, and we omit the multipole ℓ . In the CMB internal delensing, the delensed spectrum has additional contributions. The additional terms are given as

$$\begin{aligned} N^{\text{bias}} &= \langle |\hat{B} - (\hat{E} \star \hat{\phi})|^2 \rangle - \langle |\hat{B} - (\hat{E} \star \hat{\phi}')|^2 \rangle \\ &= -2\langle \hat{B}(\hat{E} \star (\hat{\phi} - \hat{\phi}'))^* \rangle \\ &\quad + \langle |(\hat{E} \star \hat{\phi})|^2 - |(\hat{E} \star \hat{\phi}')|^2 \rangle \\ &= N^{(4)} + N^{(6)}, \end{aligned} \quad (14)$$

where we define

$$N^{(4)} \equiv -2\langle \hat{B}(\hat{E} \star (\hat{\phi} - \hat{\phi}'))^* \rangle, \quad (15)$$

$$N^{(6)} \equiv \langle |(\hat{E} \star \hat{\phi})|^2 - |(\hat{E} \star \hat{\phi}')|^2 \rangle, \quad (16)$$

The first term, $N^{(4)}$, is a four point correlation and leads to the following bias terms,

$$N^{(4)} = -2\langle \hat{B}^*(\hat{E} \star \hat{\phi}) \rangle_d - 2\langle \hat{B}^*(\hat{E} \star (\hat{\phi} - \hat{\phi}')) \rangle_c, \quad (17)$$

where $\langle \cdots \rangle_d$ and $\langle \cdots \rangle_c$ are the disconnected and connected parts, respectively.

The second term, $N^{(6)}$, is a six point correlation and is decomposed into a connected term, products of the power spectrum and trispectrum, and products of three power spectra. To see this, from Eqs. (10) and (11), we explicitly rewrite the lensing B -mode template as

$$(\hat{E} \star \hat{\phi})_\ell = \int \frac{d^2 \ell'}{(2\pi)^2} \int \frac{d^2 \ell''}{(2\pi)^2} W_{\ell\ell'\ell''} \hat{E}_{\ell'} \hat{E}_{\ell''} \hat{B}_{\ell-\ell'-\ell''}, \quad (18)$$

where we define

$$W_{\ell\ell'\ell''} \equiv \frac{C_{\ell'}^{\text{EE}} C_{|\ell-\ell'|}^{\phi\phi} w_{\ell,\ell-\ell'}}{\hat{C}_{\ell'}^{\text{EE}} \hat{C}_{|\ell-\ell'|}^{\phi\phi}} \frac{A_{|\ell-\ell'|} \tilde{C}_{\ell''}^{\text{EE}} w_{\ell-\ell'-\ell'',\ell-\ell'}}{\hat{C}_{\ell''}^{\text{EE}} \hat{C}_{|\ell-\ell'-\ell''|}^{\text{BB}}}. \quad (19)$$

The power spectrum of the above quantity is the six point correlation, and we obtain

$$\begin{aligned} \langle |(\hat{E} \star \hat{\phi})|_\ell^2 \rangle &= \int \frac{d^2 \ell'_1}{(2\pi)^2} \int \frac{d^2 \ell''_1}{(2\pi)^2} \int \frac{d^2 \ell'_2}{(2\pi)^2} \int \frac{d^2 \ell''_2}{(2\pi)^2} \\ &\quad \times W_{\ell\ell'_1\ell''_1} W_{\ell\ell'_2\ell''_2} \langle \hat{E}_{\ell'_1} \hat{E}_{\ell'_2} \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{E}_{\ell''_1}^* \hat{E}_{\ell''_2}^* \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle. \end{aligned} \quad (20)$$

The disconnected part of the six point correlation is expanded as

$$\begin{aligned} &\langle \hat{E}_{\ell'_1} \hat{E}_{\ell'_2} \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{E}_{\ell''_1}^* \hat{E}_{\ell''_2}^* \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle_d \\ &= 2\langle \hat{E}_{\ell'_1} \hat{E}_{\ell'_2} \rangle \langle \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{E}_{\ell''_1}^* \hat{E}_{\ell''_2}^* \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle_c \\ &\quad + 2\langle \hat{E}_{\ell'_1} \hat{E}_{\ell''_1}^* \rangle \langle \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{E}_{\ell'_2} \hat{E}_{\ell''_2}^* \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle_c \\ &\quad + \langle \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle \langle \hat{E}_{\ell'_1} \hat{E}_{\ell'_2} \hat{E}_{\ell''_1}^* \hat{E}_{\ell''_2}^* \rangle_c \\ &\quad + \langle \hat{E}_{\ell'_1} \hat{E}_{\ell''_1}^* \rangle \langle \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{E}_{\ell'_2} \hat{E}_{\ell''_2}^* \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle_c \\ &\quad + \langle \hat{E}_{\ell'_1} \hat{E}_{\ell'_2} \rangle \langle \hat{E}_{\ell''_1}^* \hat{E}_{\ell''_2}^* \rangle \langle \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle \\ &\quad + \langle \hat{E}_{\ell'_1} \hat{E}_{\ell''_1}^* \rangle \langle \hat{E}_{\ell'_2} \hat{E}_{\ell''_2}^* \rangle \langle \hat{B}_{\ell-\ell'_1-\ell'_2} \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle, \end{aligned} \quad (21)$$

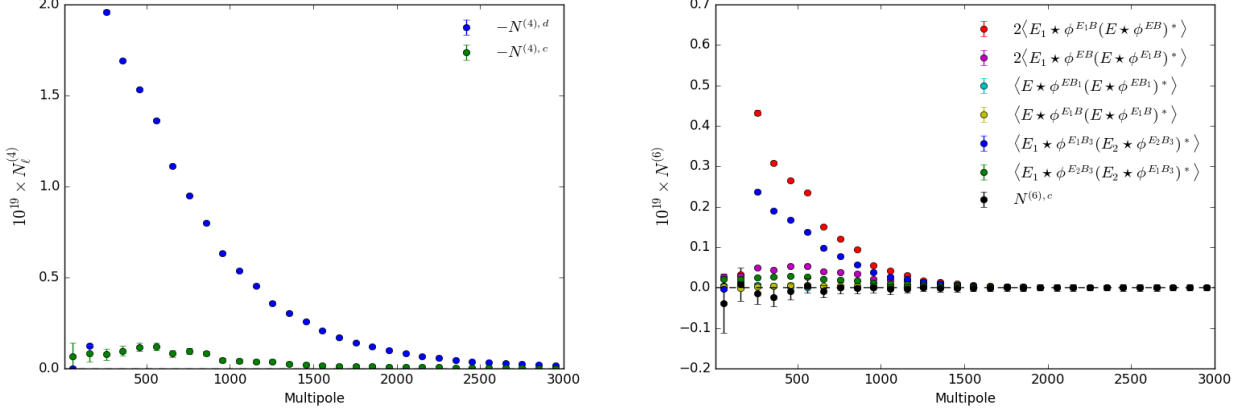


FIG. 3: Delensing biases from the four point (Left) and six point correlations (Right). The four point correlation is decomposed into the disconnected and connected parts. The six point correlation is divided into the seven terms described in Eq. (24). The error bars denote the 1σ variance of the Monte Carlo simulation.

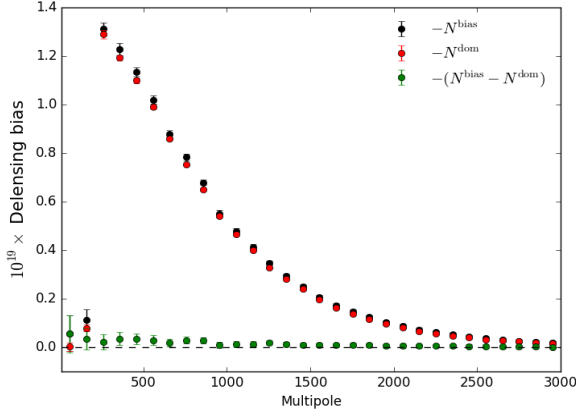


FIG. 4: Dominant contributions to the delensing bias defined as Eq. (25).

where we omit the terms irrelevant to the delensing bias (i.e., the terms contained in $|\langle \hat{E} \star \hat{\phi}' \rangle|^2$) and use the fact that the correlation between E and B modes vanishes.

The evaluation of the above equation is simplified by introducing an operator, $\langle \cdots \rangle_i$, which is the ensemble average over quantities in i th set of realizations. For example, denoting E_i and B_i as the E and B modes in i th set of simulation, we can rewrite the last term of Eq. (21) as

$$\begin{aligned}
 & \langle \hat{E}_{\ell'_1} \hat{E}_{\ell'_2}^* \rangle \langle \hat{E}_{\ell''_1} \hat{E}_{\ell''_2}^* \rangle \langle \hat{B}_{\ell-\ell'_1-\ell''_1} \hat{B}_{\ell-\ell'_2-\ell''_2}^* \rangle \\
 &= \langle \hat{E}_{1,\ell'_1} \hat{E}_{1,\ell'_2}^* \rangle_1 \langle \hat{E}_{2,\ell''_1} \hat{E}_{2,\ell''_2}^* \rangle_2 \langle \hat{B}_{3,\ell-\ell'_1-\ell''_1} \hat{B}_{3,\ell-\ell'_2-\ell''_2}^* \rangle_3 \\
 &= \langle \hat{E}_{1,\ell'_1} \hat{E}_{1,\ell'_2}^* \hat{E}_{2,\ell''_1} \hat{E}_{2,\ell''_2}^* \hat{B}_{3,\ell-\ell'_1-\ell''_1} \hat{B}_{3,\ell-\ell'_2-\ell''_2}^* \rangle_{1,2,3}, \quad (22)
 \end{aligned}$$

where the index ($i = 1, 2, 3$) denotes independent set of

realizations. Substituting the above term into Eq. (20), we obtain

$$\langle |\langle \hat{E} \star \hat{\phi} \rangle|^2 \rangle \supset \langle (E_1 \star \hat{\phi}^{E_2 B_3})(E_2 \star \hat{\phi}^{E_1 B_3})^* \rangle_{1,2,3}. \quad (23)$$

Here we again omit the multipole. Substituting Eq. (21) into Eq. (20), and including the connected part of the six point correlation, we obtain

$$\begin{aligned}
 N^{(6)} = & 2 \langle \langle E_1 \star \hat{\phi}^{E_1 B_2} (E_2 \star \hat{\phi}^{E_2 B_2})^* \rangle_1 \rangle_{2,c} \\
 & + 2 \langle \langle E_1 \star \hat{\phi}^{E_2 B_2} (E_2 \star \hat{\phi}^{E_1 B_2})^* \rangle_1 \rangle_{2,c} \\
 & + \langle \langle E_2 \star \hat{\phi}^{E_2 B_1} (E_2 \star \hat{\phi}^{E_2 B_1})^* \rangle_1 \rangle_{2,c} \\
 & + \langle \langle E_2 \star \hat{\phi}^{E_1 B_2} (E_2 \star \hat{\phi}^{E_1 B_2})^* \rangle_1 \rangle_{2,c} \\
 & + \langle \langle (E_1 \star \hat{\phi}^{E_1 B_3})(E_2 \star \hat{\phi}^{E_2 B_3})^* \rangle_1 \rangle_{2,3} \\
 & + \langle \langle (E_1 \star \hat{\phi}^{E_2 B_3})(E_2 \star \hat{\phi}^{E_1 B_3})^* \rangle_1 \rangle_{2,3} \\
 & + \langle |\hat{E} \star \hat{\phi}|^2 \rangle_c. \quad (24)
 \end{aligned}$$

Here $\langle \cdots \rangle_{i,c}$ extracts the connected part of the ensemble average. This expression is convenient to evaluate the bias using simulation. In the above equation, the first four terms are the product of the trispectrum and power spectrum, and the next two terms are the product of the three power spectra. The last term contains the pentaspectrum.

Fig. 3 shows the power spectra of each contribution in the four point and six point delensing bias. The most significant contribution comes from the disconnected part of the four point correlation $N^{(4),d}$. Some terms involved in $N^{(6)}$ are also significant. The last term in Eq. (24) is negligible.

Fig. 4 plots the dominant contribution of the delensing

bias which is given by

$$\begin{aligned} N^{\text{dom}} = & -2\langle \hat{B}^* \hat{E} \star \hat{\phi} \rangle_{\text{d}} \\ & + 2\langle (\hat{E}_1 \star \hat{\phi}^{E_1 B_2})(\hat{E}_2 \star \hat{\phi}^{E_2 B_2})^* \rangle_{1,2,c} \\ & + \langle (\hat{E}_1 \star \hat{\phi}^{E_1 B_3})(\hat{E}_2 \star \hat{\phi}^{E_2 B_3})^* \rangle_{1,2,3}. \end{aligned} \quad (25)$$

$N^{(4),c}$ cancels with a term involved in $2\langle \langle E_1 \star \hat{\phi}^{E_2 B_2} (E_2 \star \hat{\phi}^{E_1 B_2})^* \rangle_{1,2,c} \rangle$. The above expression can be derived by rewriting the delensed B modes as [24]

$$\hat{B}^{\text{d}} \simeq \hat{B} - (\hat{E} \star \hat{\phi}') - \langle (\hat{E}_1 \star \hat{\phi}^{E_1 B}) \rangle_1. \quad (26)$$

Because the third term is proportional to \hat{B} [24], the bias depends on the realization of the observed B mode. The primary B modes which we want to detect are also simultaneously subtracted.

D. Mitigating delensing bias

Let us discuss how to mitigate the delensing bias.

Several works have discussed a method to avoid the delensing bias by splitting multipole range of CMB into annuli in ℓ space so that CMB maps to be delensed and that used for the lensing reconstruction are uncorrelated [8, 25]. Fig. 5 shows the delensed B -mode spectra using the ℓ -splitting method of Ref. [8] where CMB multipole is split into four annuli with equal spacing between 200 and 3000. We also show the delensing efficiency using the standard and the ℓ -splitting methods where we use the random Gaussian reconstruction noise to avoid the delensing bias. The efficiency of delensing is defined as the ratio of the residual lensing B -mode spectrum to the original lensing- B mode spectrum, indicating that how significantly the lensing B mode is removed. In the ℓ -splitting method, the delensing bias (difference between cyan and magenta points) is not significant. However, comparing the standard method (black) with the ℓ -splitting technique (blue) in the random ϕ noise cases, the use of the ℓ -splitting method decreases the delensing efficiency. This is because the CMB multipoles used for the lensing reconstruction decrease and the precision of the lensing potential reconstruction is degraded.

As an alternative method, in the following sections, we discuss the RD method which utilizes observed CMB maps to estimate the delensed B -mode power spectrum without degradation of the delensing efficiency and is more reliable than using simulation alone. **This method emerges naturally as the delensing bias depends on observed B mode.** To construct an estimator, we employ the similar analogy of the RD method for reconstructing the lensing potential [27] which is motivated by the optimal trispectrum estimator. We extend this technique to the delensing bias by deriving the optimal estimator for the polyspectra in Sec. IV, and applying the optimal polyspectra estimator to the disconnected part of the delensing bias which has most significant contribution in the bias in Sec. V.

IV. OPTIMAL ESTIMATOR FOR POLYSPECTRA

A. Derivation

We consider a set of variables $\mathbf{x} = (x_1, x_2, \dots, x_N)$. For example, the index of x_i denotes the multipoles (ℓ) in the flat sky analysis and the harmonic coefficients (ℓm) in the full sky analysis, and specifies the temperature, E or B modes. Expanding the likelihood of these variables around the Gaussian likelihood, a term which contains n th order of the correlation is given by (e.g., Refs. [30, 41])

$$\mathcal{L} \propto \sum_{i_j} k_{i_1 \dots i_n} \frac{\partial}{\partial x_{i_1}} \dots \frac{\partial}{\partial x_{i_n}} \mathcal{L}_g, \quad (27)$$

where $k_{i_1 \dots i_n}$ is the coefficients of the n th order derivative, $i_1, \dots, i_n \in [1, N]$, and the operator \sum_{i_j} sums over all possible combinations of i_1, \dots, i_n . For example, if $n = 6$, k is the sum of the six point correlation and products of the three point correlations. The Gaussian likelihood is given as

$$\mathcal{L}_g \propto \exp \left(-\frac{1}{2} \sum_{ij} c_{ij} x_i x_j \right), \quad (28)$$

where $c_{ij} = (C^{-1})^{x_i x_j}$ with C denoting the covariance between \mathbf{x} .

To derive an explicit expression of the derivative of the Gaussian likelihood with respect to $x_{i_1}, x_{i_2}, \dots, x_{i_n}$, we first define the following quantities:

$$p_m^n = \langle \bar{x}_{i_1}^1 \dots \bar{x}_{i_{2m}}^m \bar{x}_{i_{2m+1}} \dots \bar{x}_{i_n} + (\text{perms.}) \rangle_{1, \dots, m}. \quad (29)$$

Here m is a non-negative integer and $p_m^n = 0$ if $m > [n/2]$. Again, to simplify the numerical calculation, we introduce the operator, $\langle \dots \rangle_i$, which takes the ensemble average over i th set of realizations, \bar{x}^i , and \bar{x} is an inverse-variance filtered variable fields as

$$\bar{x}_j = \sum_i c_{ij} x_i. \quad (30)$$

For example, if $n = 6$, we obtain

$$p_0^6 = \bar{x}_{i_1} \dots \bar{x}_{i_6}, \quad (31)$$

$$p_1^6 = \langle \bar{x}_{i_1}^1 \bar{x}_{i_2}^1 \bar{x}_{i_3} \dots \bar{x}_{i_6} + (14 \text{ perms.}) \rangle_1, \quad (32)$$

$$p_2^6 = \langle \bar{x}_{i_1}^1 \dots \bar{x}_{i_4}^2 \bar{x}_{i_5} \bar{x}_{i_6} + (44 \text{ perms.}) \rangle_{1,2}, \quad (33)$$

$$p_3^6 = \langle \bar{x}_{i_1}^1 \dots \bar{x}_{i_6}^3 + (14 \text{ perms.}) \rangle_{1,2,3}, \quad (34)$$

and $p_m^6 = 0$ for $m > 3$. Using

$$\frac{\partial \bar{x}_j}{\partial x_i} = c_{ij} = \langle \bar{x}_i \bar{x}_j \rangle, \quad (35)$$

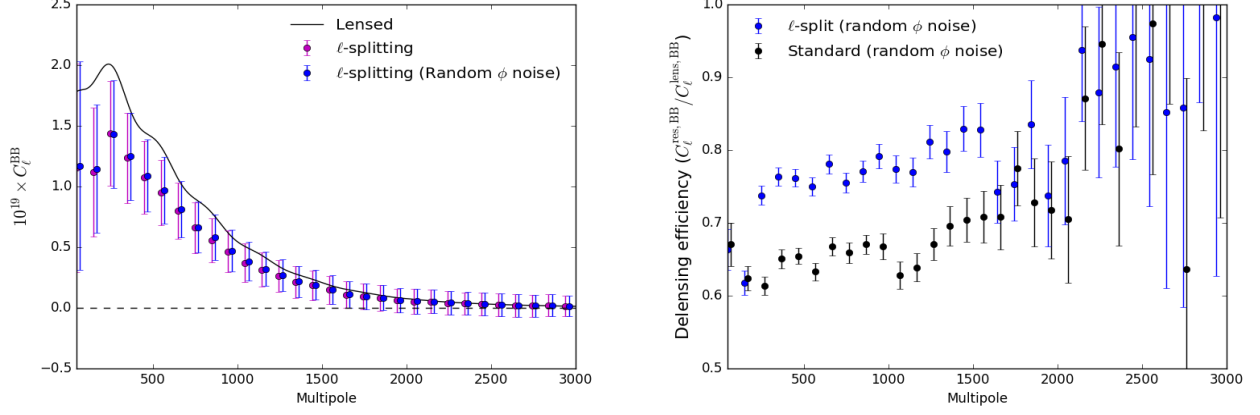


FIG. 5: *Left:* Same as Fig. 2 but for the ℓ -splitting method. *Right:* Delensing efficiency of the B -mode polarization. We show the standard and ℓ -splitting methods with the random Gaussian reconstruction noise. The error bar denotes the 1σ simulation error.

the quantities p_m^n satisfy

$$\begin{aligned} p_0^{n+1} &= \bar{x}_{i_{n+1}} p_0^n, \\ p_{m+1}^{n+1} &= \frac{\partial p_m^n}{\partial x_{i_{n+1}}} + \bar{x}_{i_{n+1}} p_{m+1}^n. \end{aligned} \quad (36)$$

With p_m^n , we find that the n th order derivative of the Gaussian likelihood is simply expressed as

$$(-1)^n \frac{\partial^n \mathcal{L}_g}{\partial x_{i_1} \cdots \partial x_{i_n}} = \mathcal{L}_g \sum_{m=0}^n (-1)^m p_m^n \quad (n \geq 1). \quad (37)$$

To show this, we first consider the case with $n = 1$. The left-hand side of the above equation becomes

$$-\frac{\partial \mathcal{L}_g}{\partial x_{i_1}} = \bar{x}_{i_1} \mathcal{L}_g, \quad (38)$$

On the other hand, the quantity defined in Eq. (29) is given by $p_0^1 = \bar{x}_{i_1}$ and the right-hand side of Eq. (37) coincides with Eq. (38). To show the case with n greater than 1, we differentiate Eq. (37) with respect to $x_{i_{n+1}}$ and obtain

$$\begin{aligned} \frac{(-1)^n \partial^n \mathcal{L}_g}{\partial x_{i_1} \cdots \partial x_{i_{n+1}}} &= \sum_{m=0}^n (-1)^m \left[\mathcal{L}_g \frac{\partial p_m^n}{\partial x_{i_{n+1}}} + p_m^n \frac{\partial \mathcal{L}_g}{\partial x_{i_{n+1}}} \right] \\ &= \sum_{m=0}^n (-1)^m \mathcal{L}_g \frac{\partial p_m^n}{\partial x_{i_{n+1}}} - \sum_{m=0}^n (-1)^m \mathcal{L}_g p_m^n \bar{x}_{i_{n+1}}. \end{aligned} \quad (39)$$

Replacing m with $m + 1$ in the second term, we obtain

$$\begin{aligned} \frac{(-1)^n \partial^n \mathcal{L}_g}{\partial x_{i_1} \cdots \partial x_{i_{n+1}}} &= \sum_{m=0}^n (-1)^m \mathcal{L}_g \left[\frac{\partial p_m^n}{\partial x_{i_{n+1}}} + p_{m+1}^n \bar{x}_{i_{n+1}} \right] \\ &\quad - \mathcal{L}_g \bar{x}_{i_{n+1}} p_0^n. \end{aligned} \quad (40)$$

Using Eq. (36), the first term becomes

$$\sum_{m=0}^n (-1)^m \mathcal{L}_g p_{m+1}^{n+1} = \sum_{m=0}^{n+1} (-1)^{m+1} \mathcal{L}_g p_m^{n+1} + \mathcal{L}_g p_0^{n+1}. \quad (41)$$

Substituting Eq. (41) into the first term of Eq. (40) leads to Eq. (37) and Eq. (37) is satisfied for any n .

The optimal estimator of the n th order polyspectra, k , is proportional to the derivative of the likelihood with respect to k as similar to the trispectrum [27]. From Eq. (37), the estimator of k is given as

$$\hat{k}_{i_1 \cdots i_n} = A_{i_1 \cdots i_n} \left[p_0^n + \sum_{m=1}^n (-1)^{m+n} p_m^n \right] \quad (42)$$

$$\equiv A_{i_1 \cdots i_n} \bar{k}_{i_1 \cdots i_n}, \quad (43)$$

where $A_{i_1 \cdots i_n}$ is a normalization of the estimator so that the estimator is unbiased and $\bar{k}_{i_1 \cdots i_n}$ is the unnormalized estimator. The second term in the parenthesis is the estimator of the disconnected part of k :

$$\hat{k}_{i_1 \cdots i_n}^d = -A_{i_1 \cdots i_n} \sum_{m=1}^n (-1)^{m+n} p_m^n. \quad (44)$$

If the products of the polyspectra in k are negligible (e.g., the product of the three-point correlations in $k^{i_1 \cdots i_6}$), Eq. (42) becomes the estimator for the connected part of the n -point correlation.

B. Suboptimal estimator

Here we discuss the accuracy of the estimator given in Eq. (42). The estimator requires a set of simulated data, \bar{x}^i . In the trispectrum case ($n = 4$), when we use an incorrect covariance $\langle \bar{x}_i \bar{x}_j \rangle = \langle \bar{x}_i \bar{x}_j \rangle^{\text{true}} + \delta \langle \bar{x}_i \bar{x}_j \rangle$ in the

simulation, the estimator of Eq. (42) has no contribution at linear order of $\delta\langle\bar{x}_i\bar{x}_j\rangle$ [27]. Thus, the estimator is less sensitive to a mismatch between the simulated and observed covariance and more accurate than correcting the disconnected bias with the simulation alone.

In general cases, however, the estimator has contributions from linear order of $\delta\langle\bar{x}_i\bar{x}_j\rangle$. For example, in $n = 6$, the following terms involved in Eq. (32) produce biases:

$$\langle p_1^6 \rangle \supset \langle \bar{x}_{i_1}\bar{x}_{i_2} \rangle \langle \bar{x}_{i_3} \cdots \bar{x}_{i_6} \rangle_c + (14 \text{ perms.}). \quad (45)$$

A possible way to mitigate these terms is to redefine p_1^6 as

$$(p_1^6)' \equiv p_1^6 - \langle \bar{x}_{i_1}\bar{x}_{i_2} \rangle \langle \bar{x}_{i_3} \cdots \bar{x}_{i_6} \rangle_c + (14 \text{ perms.}) \\ + \bar{x}_{i_1}\bar{x}_{i_2} \langle \bar{x}_{i_3} \cdots \bar{x}_{i_6} \rangle_c + (14 \text{ perms.}). \quad (46)$$

The second term is added so that its ensemble average cancels with Eq. (45), while the third term is introduced such that the additional terms have zero mean. Combining p_m^6 in Eqs. (32) and (33) and (34), the estimator for the disconnected part of the polyspectra is given by

$$\hat{k}_{i_1 \cdots i_n}^{\text{d,sub}} \propto (p_1^6)' - p_2^6 + p_3^6 \\ = [\langle \bar{x}_{i_1}\bar{x}_{i_2} - \langle \bar{x}_{i_1}\bar{x}_{i_2} \rangle \rangle \langle \bar{x}_{i_3} \cdots \bar{x}_{i_6} \rangle + (14 \text{ perms.})] \\ + p_1^6 - 2p_2^6 + 4p_3^6. \quad (47)$$

Note that we discuss an alternative way to derive the above estimator in Appendix B by generalizing the discussion in Ref. [26] to the polyspectra. The above estimator has no contributions from linear order of $\delta\langle\bar{x}_i\bar{x}_j\rangle$ and also from error of the trispectrum at linear order, $\delta\langle\bar{x}_i\bar{x}_j\bar{x}_k\bar{x}_l\rangle_c$. Thus, the suboptimal estimator of Eq. (47) is more accurate than the estimator defined in Eq. (44).

V. REALIZATION-DEPENDENT DELENSING

We here discuss the RD estimators for the removal of the delensing bias. Since the dominant contribution to the delensing bias comes from the disconnected part, we apply the above optimal polyspectra estimators given in Eqs. (44) and (47) to the disconnected part of the delensing bias. Since the derivation requires a lengthy calculation, here we only show the results and the details of the calculation are given in Appendix C.

From Eqs. (14) and (16), the bias is described as

$$N^{\text{bias}} = N^{(4)} + N^{(6)}. \quad (48)$$

$N^{(4)}$ is a four point correlation and its disconnected part is the most dominant contribution in the delensing bias. The disconnected part of $N^{(4)}$ can be replaced with the following RD estimator (see Appendix C 1);

$$\hat{N}^{(4),d} = -2[\langle \hat{E}_1 \star \hat{\phi} + \hat{E}_{012} \star (\hat{\phi}^{E_{0,-2}B_1} + \hat{\phi}^{E_1B_{0,-2}}) \rangle \hat{B}_{01}^*]_{1,2}. \quad (49)$$

Here we define $X_{i,\pm j \dots} = X_i \pm X_j + \dots$ and $i = 0$ means the real data. $N^{(6)}$ is a six point correlation and its disconnected part is the dominant source of the bias. According to the optimal polyspectra estimator, the RD estimator for the disconnected part of the six point correlation is given by $\hat{N}^{(6)} \propto p_1^6 - p_2^6 + p_3^6$ where p_i^6 is defined in Eq. (29) with four E and two B modes. We find that the RD estimator for the disconnected part of the six point correlation becomes (see Appendix C 2)

$$\hat{N}^{(6),d} = \langle \text{Re}[\hat{E}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13})\mathcal{B}_{0123}^*] \rangle_{1,2,3} \\ + \langle \hat{E}_{0,-3} \star \hat{\phi}_{11} (\hat{E}_{03} \star \hat{\phi}_{11})^* \rangle_{1,3}, \quad (50)$$

where we define $2\hat{\phi}_{ij} = \hat{\phi}^{E_iB_j} + \hat{\phi}^{E_jB_i}$ and

$$\mathcal{B}_{0123} \equiv \hat{E}_{012} \star (2\hat{\phi} - 2\hat{\phi}_{02} - 4\hat{\phi}_{03} + 2\hat{\phi}_{23}) \\ - 2\hat{E}_3 \star \hat{\phi} + \hat{E}_{0,-2} \star (2\hat{\phi}_{01} - \hat{\phi}_{13}). \quad (51)$$

As discussed in the previous section, the suboptimal estimator of Eq. (47) is more accurate than the estimator of Eq. (50). The suboptimal RD estimator for the six point correlation is given by (see Appendix C 3)

$$\hat{N}^{(6),\text{d,sub}} = \langle \hat{E}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13})(\mathcal{B}_{0123}^{\text{sub}})^* \rangle_{1,2,3}, \quad (52)$$

where we define

$$\mathcal{B}_{0123}^{\text{sub}} \equiv \hat{E}_{012} \star (2\hat{\phi} - 4\hat{\phi}_{02} - 8\hat{\phi}_{03} + 8\hat{\phi}_{23} + 2\hat{\phi}_{11}) \\ - 2\hat{E}_3 \star (2\hat{\phi} + \hat{\phi}_{11} + \hat{\phi}_{33}) \\ + 2(\hat{E}_{01} - 2\hat{E}_2) \star (\hat{\phi}_{01} - \hat{\phi}_{13}). \quad (53)$$

As shown in Fig. 4, the dominant terms of the delensing bias are given by Eq. (25). The RD estimator for Eq. (25) can be constructed by extracting the corresponding terms in the suboptimal estimator. The details are given in Appendix C and the result is

$$\hat{N}^{(6),\text{d,dom}} = 2\hat{E}_{0,-1} \star \hat{\phi}^{E_{01}B_{02}} [\hat{E}_{23} \star (\hat{\phi}_{22} - 2\hat{\phi}_{23})]^* \\ + \hat{E}_1 \star \hat{\phi}^{E_{01}B_{02}} [2\hat{E} \star (\hat{\phi} - \hat{\phi}_{22} - \hat{\phi}^{E_2B}) \\ - \hat{E}_2 \star (\hat{\phi}^{(2E+E_2)B_{02}} - \hat{\phi}_{22}) + 2\hat{E}_3 \star \hat{\phi}^{E_2B_3}]^*. \quad (54)$$

The above RD estimators are more reliable than the use of the delensing bias estimated from simulation alone. In the followings, we check how the use of $\hat{N}^{(6),\text{d,sub}}$ or $\hat{N}^{(6),\text{d,dom}}$ degrades the performance of the delensing.

Fig. 6 shows the difference between the B -mode power spectra using the RD estimators and with the random Gaussian reconstruction noise, divided by the 1σ simulation error. After the correction of the delensing bias with the RD estimators, the delensed spectrum becomes consistent with that in the case with the random Gaussian reconstruction noise. Fig. 7 shows the 1σ statistical error of the delensed B -mode spectrum after the bias

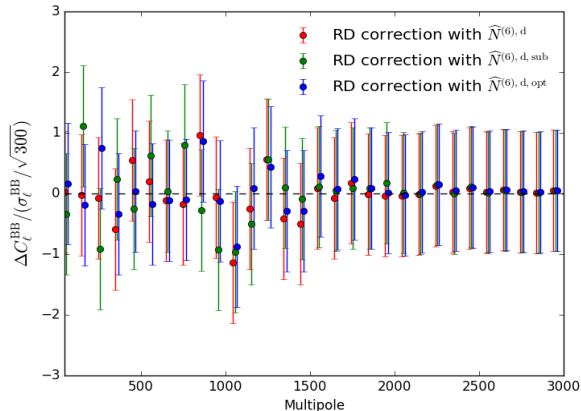


FIG. 6: Difference between the B -mode spectra using the RD estimators and with the random Gaussian reconstruction noise, divided by the 1σ simulation error. We show the cases where the delensing bias is estimated using $\hat{N}^{(6),d}$, $\hat{N}^{(6),d,sub}$ or $\hat{N}^{(6),d,dom}$.

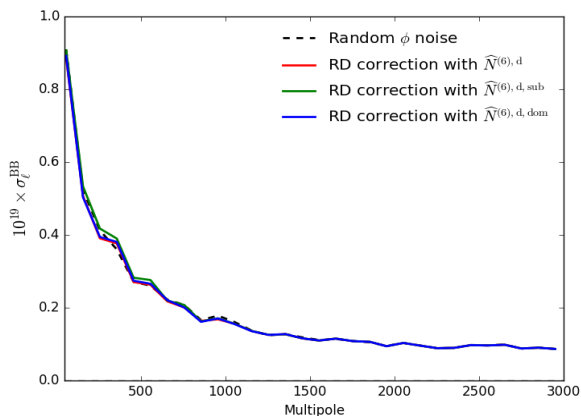


FIG. 7: 1σ statistical error of the delensed B -mode power spectrum with corrections of the delensing bias using the RD estimators. We also show the case with the random Gaussian reconstruction noise.

correction. The statistical errors are almost unchanged between the three RD estimators. We conclude that the use of the suboptimal RD estimators (\hat{N}^{sub} or \hat{N}^{dom}) does not degrade the performance of the delensing.

VI. SUMMARY AND DISCUSSION

We have derived the RD estimators to correct the delensing bias. We first showed that the delensing bias is

significant in the delensed B -mode power spectrum. To correct the bias, we derived the optimal polyspectra estimator in general cases. We then formulated the RD estimators for the delensing bias with the polyspectra estimator. Unlike the ℓ splitting method proposed in previous works, the RD method corrects the delensing bias without substantial loss of the delensing efficiency.

A similar bias on B modes discussed in this paper would be appeared when we remove, for example, the spatially varying components of foregrounds estimated from the same realization of CMB maps with a quadratic estimator. The RD method presented in this paper would be also useful to mitigate such biases.

We have focused on the B -mode delensing but the delensing bias is also significant in the CMB temperature and E -mode delensing [25, 26] (see also Appendix A). Unlike the B -mode delensing focused in this paper, the higher-order terms of the lensing potential is not negligible in the temperature and E -mode delensing, and the remapping method or the template method including the higher-order terms of the lensing potential is more efficient than the linear-order template method. The delensing bias contains non negligible contributions from the higher-order correlations (e.g. eight-point correlation of the observed CMB maps). To apply the RD method described in this paper to the temperature and E -mode delensing, we need to include the optimal polyspectra up to more than six point correlations.

In this paper, we have focused on the application of the optimal polyspectra estimator to the delensing. The estimator is also useful to estimate, e.g., the bispectrum of the reconstructed lensing potential which is a six point correlation of the observed CMB anisotropies, and higher-order statistics of the kinetic SZ effects. The lensing bispectrum is generated by the non-linear growth of the large-scale structure and the post-Born corrections. Since measurements of the lensing bispectrum will be useful to extract additional cosmological information in future CMB experiments [42–45], we will investigate the feasibility of the RD method for a lensing bispectrum measurement in our future work.

Acknowledgments

We would like to thank Chao-Lin Kuo for support of this work and Antony Lewis for helpful comments. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

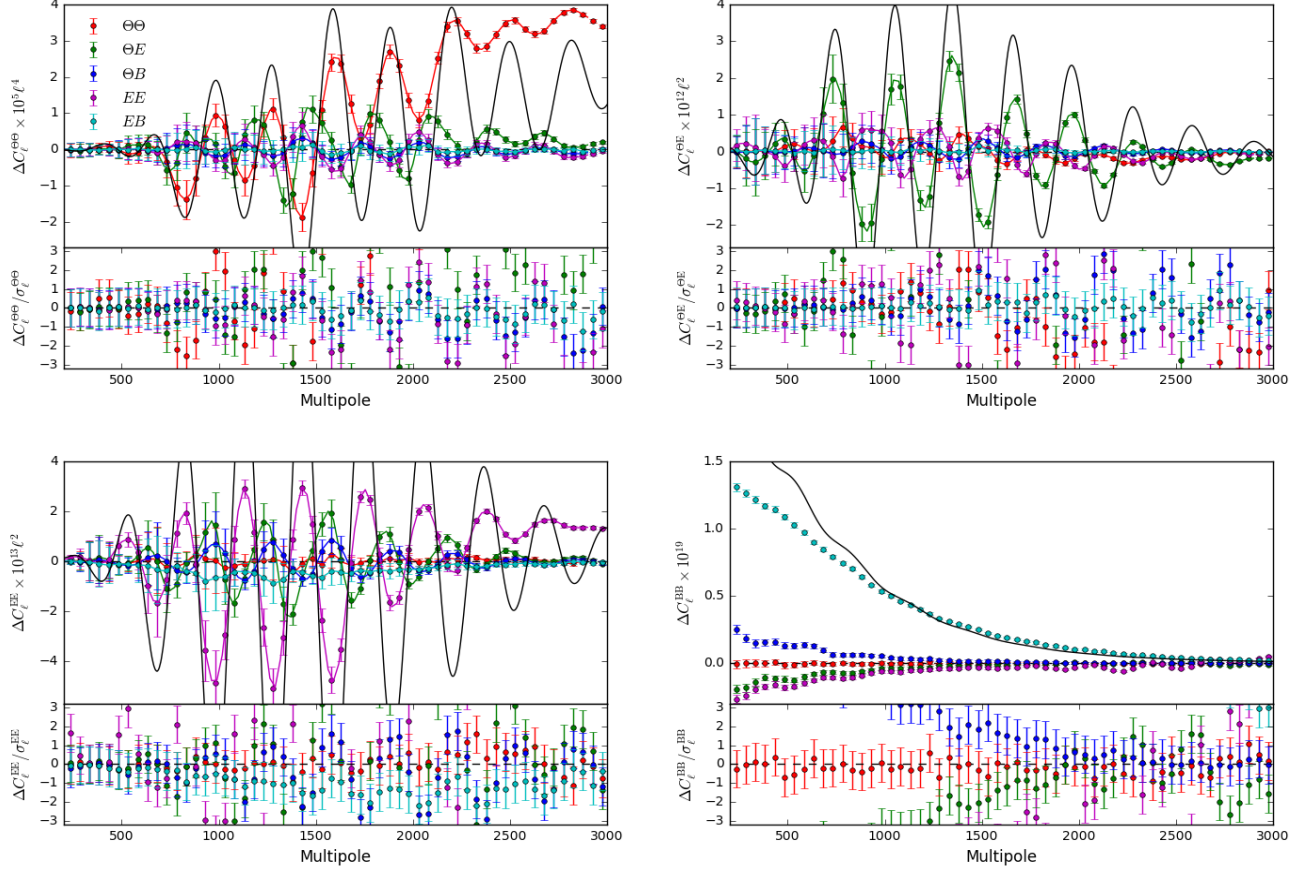


FIG. 8: Delensing bias in $C_\ell^{\Theta\Theta}$, $C_\ell^{\Theta E}$, C_ℓ^{EE} , and C_ℓ^{BB} with reconstructed lensing potentials from $\Theta\Theta$ (red), ΘE (green), ΘB (blue), EE (magenta), and EB (cyan) quadratic estimators of Ref. [32]. The solid lines show the difference between the lensed and unlensed CMB power spectra. We assume $3\mu\text{K-arcmin}$ and $\sqrt{2} \times 3\mu\text{K-arcmin}$ white noises in the temperature and polarization maps, respectively, and the observed CMB maps are convolved with 1 arcmin Gaussian beam. The error denotes the 1σ statistical error for one realization of a full sky observation.

Appendix A: Delensing bias in other cases

In this section, we show the delensing biases in the delensed temperature, E , and B mode auto power spectra and temperature- E -mode cross power spectrum. We reconstruct lensing potentials from $\Theta\Theta$, ΘE , ΘB , EE , and EB quadratic estimators of Ref. [32]. The temperature, E , and B modes are delensed by the lensing potentials with the remapping method.

Fig. 8 shows the delensing biases using a lensing potential reconstructed from these five estimators. The delensing biases are defined as $\Delta C_\ell = C_\ell^{\text{res}, \phi^{\text{inp}}} - C_\ell^{\text{res}}$ where C_ℓ^{res} and $C_\ell^{\text{res}, \phi^{\text{inp}}}$ are the delensed spectra using a reconstructed lensing potential and the input lensing potential with the random Gaussian reconstruction noise, respectively. We also show the difference between the lensed and unlensed power spectra. In the delensed $C_\ell^{\Theta\Theta}$, $C_\ell^{\Theta E}$, C_ℓ^{EE} , and C_ℓ^{BB} , the most significant delensing biases come from $\Theta\Theta$, ΘE , EE , and EB quadratic estimators, respectively.

Appendix B: Derivation of the suboptimal polyspectra estimator

Here we derive Eq. (47) by generalizing the discussion of Ref. [26] to the case with the six point correlation. In the six point correlation, errors in the covariance and trispectrum could lead to errors in the estimate of the delensing

bias. To remove the errors, we consider the following estimator:

$$\hat{k}'_{i_1 \dots i_6} = K_{i_1 \dots i_6}^{(2,4)} + K_{i_1 \dots i_6}^{(2,2,2)} + \sum_{jk} \frac{\partial [K_{i_1 \dots i_6}^{(2,4)} + K_{i_1 \dots i_6}^{(2,2,2)}]}{\partial c_{jk}} \delta c_{jk} + \sum_{jklm} \frac{\partial K_{i_1 \dots i_6}^{(2,4)}}{\partial t_{jklm}} \delta t_{jklm}. \quad (\text{B1})$$

where we define $c_{ij} = \langle \bar{x}_i \bar{x}_j \rangle$, $t_{jklm} = \langle \bar{x}_j \bar{x}_k \bar{x}_l \bar{x}_m \rangle_c$, and

$$K_{i_1 \dots i_6}^{(2,4)} \equiv c_{i_1 i_2} t_{i_3 i_4 i_5 i_6} + (14 \text{ perms.}), \quad (\text{B2})$$

$$K_{i_1 \dots i_6}^{(2,2,2)} \equiv c_{i_1 i_2} c_{i_3 i_4} c_{i_5 i_6} + (14 \text{ perms.}) = p_3^6, \quad (\text{B3})$$

$$\delta c_{jk} \equiv \bar{x}_j \bar{x}_k - c_{jk}, \quad (\text{B4})$$

$$\delta t_{jklm} \equiv \bar{x}_j \dots \bar{x}_m - [\bar{x}_j \bar{x}_k c_{lm} + (5 \text{ perms.})] + [c_{jk} c_{lm} + (2 \text{ perms.})] - t_{jklm}. \quad (\text{B5})$$

The sum of $K^{(2,4)}$ and $K^{(2,2,2)}$ is the disconnected part of the six point correlation. The expectation of the above estimator is equal to the first term, $K^{(2,4)} + K^{(2,2,2)}$. If there are errors in the covariance and trispectrum, δc and δt , the above estimator does not contain biases from the linear order of δc and δt .

To simplify the above estimator, we use

$$\sum_{jk} \frac{\partial K_{i_1 \dots i_6}^{(2,4)}}{\partial c_{jk}} \delta c_{jk} = (\bar{x}_{i_1} \bar{x}_{i_2} - c_{i_1 i_2}) t_{i_3 i_4 i_5 i_6} + (14 \text{ perms.}), \quad (\text{B6})$$

$$\sum_{jk} \frac{\partial K_{i_1 \dots i_6}^{(2,2,2)}}{\partial c_{jk}} \delta c_{jk} = [(\bar{x}_{i_1} \bar{x}_{i_2} - c_{i_1 i_2}) c_{i_3 i_4} c_{i_5 i_6} + (\bar{x}_{i_3} \bar{x}_{i_4} - c_{i_3 i_4}) c_{i_1 i_2} c_{i_5 i_6} + (\bar{x}_{i_5} \bar{x}_{i_6} - c_{i_5 i_6}) c_{i_1 i_2} c_{i_3 i_4}] + (14 \text{ perms.}), \quad (\text{B7})$$

$$\begin{aligned} \sum_{jklm} \frac{\partial K_{i_1 \dots i_6}^{(2,4)}}{\partial t_{jklm}} \delta t_{jklm} &= c_{i_1 i_2} \{ \bar{x}_{i_3} \dots \bar{x}_{i_6} - [\bar{x}_{i_3} \bar{x}_{i_4} c_{i_5 i_6} + (5 \text{ perms.})] + [c_{i_3 i_4} c_{i_5 i_6} + (2 \text{ perms.})] - t_{i_3 \dots i_6} \} + (14 \text{ perms.}) \\ &= p_1^6 - 2p_2^6 + 3p_3^6 - K_{i_1 \dots i_6}^{(2,4)}. \end{aligned} \quad (\text{B8})$$

Substituting the above equations into Eq. (B1), we find

$$\hat{k}'_{i_1 \dots i_6} = [(\bar{x}_{i_1} \bar{x}_{i_2} - c_{i_1 i_2}) \langle \bar{x}_{i_3} \bar{x}_{i_4} \bar{x}_{i_5} \bar{x}_{i_6} \rangle + (14 \text{ perms.})] + p_1^6 - 2p_2^6 + 4p_3^6. \quad (\text{B9})$$

This estimator coincides with Eq. (47).

Appendix C: Derivation of the RD estimator for delensing bias

In this section, we derive the RD estimator in general cases of the CMB internal delensing. We consider delensed CMB anisotropies \hat{X} using a reconstructed lensing potential $\hat{\phi}^{YZ}$ obtained from observed CMB anisotropies, \hat{Y} and \hat{Z} .

1. Four point correlation

We first discuss the RD estimator for the disconnected part of the four point correlation $N^{(4),d}$. In the CMB internal delensing, the delensing bias from the disconnected part of the four point correlation is given by

$$N^{(4),d} = -2 \langle (\hat{X} \star \hat{\phi}^{YZ}) \widehat{W}^* \rangle. \quad (\text{C1})$$

From Eq. (44), the optimal trispectrum estimator for the disconnected part of $-2(\hat{X} \star \hat{\phi}^{YZ}) \widehat{W}^*$ is given as $p_1^4 - p_2^4$ where p_m^4 is defined in Eq. (29). The quantity, p_1^4 , becomes

$$p_1^4 = -2 \langle 2(\hat{X}_1 \star \hat{\phi}_{01}) \widehat{W}^* + 2(\hat{X} \star \hat{\phi}_{01}) \widehat{W}_1^* + (\hat{X}_1 \star \hat{\phi}) \widehat{W}_1^* \rangle_1. \quad (\text{C2})$$

Here we define a symmetric estimator,

$$2\widehat{\phi}_{ij} = \widehat{\phi}^{Y_i Z_j} + \widehat{\phi}^{Y_j Z_i}, \quad (\text{C3})$$

with $i = 0$ meaning the real data, and ignore the mean-field bias because its contribution is absorbed into the mean-field corrected estimator $\widehat{\phi} = \widehat{\phi}^{YZ} - \langle \widehat{\phi} \rangle$ after combining p_m . We then replace $\widehat{\phi}^{YZ}$ with $\widehat{\phi}$. The last term is the correction for the mean-field bias in $(\widehat{X} \star \widehat{\phi})\widehat{W}^*$. The term, p_2^4 , becomes

$$p_2^4 = -4\langle (\widehat{X}_1 \star \widehat{\phi}_{12})\widehat{W}_2^* \rangle_{1,2}. \quad (\text{C4})$$

Using Eqs. (C2) and (C4), the optimal estimator for $N^{(4),d}$ is written as

$$\begin{aligned} \widehat{N}^{(4),d} &= p_1^4 - p_2^4 \\ &= -2\langle [\widehat{X}_1 \star \widehat{\phi} + \widehat{X}_{012} \star (\widehat{\phi}^{Y_0, -2Z_1} + \widehat{\phi}^{Y_1 Z_0, -2})]\widehat{W}_{01}^* \rangle_{1,2}, \end{aligned} \quad (\text{C5})$$

where we define $X_{i,\pm j\dots} = X_i \pm X_j + \dots$.

The RD estimator of Eq. (49) is derived by substituting $X = Y = E$, $Z = W = B$ into Eq. (C5). For the temperature delensing with the temperature quadratic estimator, the RD estimator is obtained by substituting $X = Y = Z = W = \Theta$ into Eq. (C5) and the result is

$$\widehat{N}^{(4),d} = -2\langle [\widehat{\Theta}_1 \star \widehat{\phi} + 2\widehat{\Theta}_{012} \star \widehat{\phi}^{\Theta_0, -2\Theta_1}]\widehat{\Theta}_{01}^* \rangle_{1,2}, \quad (\text{C6})$$

where $\widehat{\phi}^{\Theta\Theta}$ is the temperature quadratic estimator [46].

2. Six point correlation

Next we consider the estimator for the six point correlation. For clarity, we omit the operation, $\langle \dots \rangle_i$. In the CMB internal delensing using the template method, the six point correlation is written as $|\widehat{X} \star \widehat{\phi}^{YZ}|^2$. We first derive the estimator based on Eq. (42) and the suboptimal estimator of Eq. (47) is derived in the next section.

According to the optimal estimator of Eq. (42), the RD estimator for the disconnected part of the six point correlation is given as $p_1^6 - p_2^6 + p_3^6$. The quantity, p_1^6 , has 15 terms and becomes

$$\begin{aligned} p_1^6 &= 2\text{Re}[\widehat{X}_1 \star \widehat{\phi}^{Y_1 Z}(\widehat{X} \star \widehat{\phi})^*] + 2\text{Re}[\widehat{X}_1 \star \widehat{\phi}(\widehat{X} \star \widehat{\phi}^{Y_1 Z})^*] + |\widehat{X}_1 \star \widehat{\phi}|^2 + |\widehat{X} \star \widehat{\phi}^{Y_1 Z}|^2 + |\widehat{X} \star \widehat{\phi}^{Y Z_1}|^2 \\ &\quad + 2\text{Re}[\widehat{X}_1 \star \widehat{\phi}^{Y Z_1}(\widehat{X} \star \widehat{\phi})^*] + 2\text{Re}[\widehat{X}_1 \star \widehat{\phi}(\widehat{X} \star \widehat{\phi}^{Y Z_1})^*] + 2\text{Re}[\widehat{X} \star \widehat{\phi}^{Y_1 Z}(\widehat{X} \star \widehat{\phi}^{Y Z_1})^*]. \end{aligned} \quad (\text{C7})$$

Using the symmetric estimator of Eq. (C3), the above quantity is reduced to

$$\begin{aligned} p_1^6 &= 2\text{Re}[\widehat{X}_1 \star (\widehat{\phi}^{Y_1 Z} + \widehat{\phi}^{Y Z_1})(\widehat{X} \star \widehat{\phi})^*] + 2\text{Re}[\widehat{X}_1 \star \widehat{\phi}(\widehat{X} \star (\widehat{\phi}^{Y_1 Z} + \widehat{\phi}^{Y Z_1}))^*] + |\widehat{X}_1 \star \widehat{\phi}|^2 + |\widehat{X} \star (\widehat{\phi}^{Y_1 Z} + \widehat{\phi}^{Y Z_1})|^2 \\ &= 4\text{Re}[\widehat{X}_1 \star \widehat{\phi}_{01}(\widehat{X} \star \widehat{\phi})^*] + 4\text{Re}[\widehat{X}_1 \star \widehat{\phi}(\widehat{X} \star \widehat{\phi}_{01})^*] + |\widehat{X}_1 \star \widehat{\phi}|^2 + 4|\widehat{X} \star \widehat{\phi}_{01}|^2. \end{aligned} \quad (\text{C8})$$

The quantity, p_2^6 , consists of 45 terms. The terms which do not have $\langle \hat{\phi} \rangle$ are given by

$$\begin{aligned}
p_2^6 = & \hat{X} \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z_2})^* + \hat{X} \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z_1})^* \\
& + \hat{X} \star \hat{\phi}^{Y_1 Z} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z_2})^* + \hat{X} \star \hat{\phi}^{Y_1 Z} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z_1})^* \\
& + \hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X} \star \hat{\phi}^{Y_2 Z_1})^* + |\hat{X} \star \hat{\phi}^{Y_1 Z_2}|^2 \\
& + \hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_2 \star \hat{\phi}^{Y Z_1})^* + \hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_1 \star \hat{\phi}^{Y Z_2})^* \\
& + \hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_1 \star \hat{\phi}^{Y_2 Z})^* + \hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z})^* \\
& + \hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z_1})^* + \hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z_2})^* \\
& + \hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X} \star \hat{\phi}^{Y_2 Z_1})^* + \hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X} \star \hat{\phi}^{Y_1 Z_2})^* \\
& + \hat{X}_1 \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y Z_2})^* + \hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X}_2 \star \hat{\phi}^{Y Z_1})^* + |\hat{X}_1 \star \hat{\phi}^{Y Z_2}|^2 \\
& + \hat{X}_1 \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z})^* + \hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z})^* + \hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X}_1 \star \hat{\phi}^{Y_2 Z})^* \\
& + \hat{X}_1 \star \hat{\phi}^{Y_2 Z} (\hat{X} \star \hat{\phi}^{Y_1 Z_2})^* + \hat{X}_1 \star \hat{\phi}^{Y_2 Z} (\hat{X} \star \hat{\phi}^{Y_2 Z_1})^* \\
& + \hat{X}_1 \star \hat{\phi}^{Y_2 Z} (\hat{X}_1 \star \hat{\phi}^{Y Z_2})^* + \hat{X}_1 \star \hat{\phi}^{Y_2 Z} (\hat{X}_2 \star \hat{\phi}^{Y Z_1})^* + \hat{X}_1 \star \hat{\phi}^{Y_1 Z} (\hat{X}_2 \star \hat{\phi}^{Y Z_2})^* \\
& + \hat{X}_1 \star \hat{\phi}^{Y_1 Z} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z})^* + \hat{X}_1 \star \hat{\phi}^{Y_2 Z} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z})^* + |\hat{X}_1 \star \hat{\phi}^{Y_2 Z}|^2 \\
& + \hat{X}_1 \star \hat{\phi}^{Y_2 Z_1} (\hat{X} \star \hat{\phi}^{Y Z_2})^* + \hat{X}_1 \star \hat{\phi}^{Y_2 Z_1} (\hat{X} \star \hat{\phi}^{Y Z_1})^* \\
& + \hat{X}_1 \star \hat{\phi}^{Y_2 Z_1} (\hat{X} \star \hat{\phi}^{Y_2 Z})^* + \hat{X}_1 \star \hat{\phi}^{Y_1 Z_2} (\hat{X} \star \hat{\phi}^{Y_2 Z})^* \\
& + \hat{X}_1 \star \hat{\phi}^{Y_2 Z_1} (\hat{X}_2 \star \hat{\phi})^* + \hat{X}_1 \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_2 \star \hat{\phi})^* .
\end{aligned} \tag{C9}$$

To reduce the above terms, we first rewrite the above equation as

$$\begin{aligned}
p_2^6 = & 2\text{Re}[\hat{X} \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z_2})^*] + 2\text{Re}[\hat{X} \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z_1})^*] \\
& + 2\text{Re}[\hat{X} \star \hat{\phi}^{Y_1 Z} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z_2})^*] + 2\text{Re}[\hat{X} \star \hat{\phi}^{Y_1 Z} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z_1})^*] \\
& + \hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X} \star \hat{\phi}^{Y_2 Z_1})^* + |\hat{X} \star \hat{\phi}^{Y_1 Z_2}|^2 + 2\text{Re}[\hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X} \star \hat{\phi}^{Y_1 Z_2})^*] \\
& + 2\text{Re}[\hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_2 \star \hat{\phi}^{Y Z_1})^*] \\
& + 2\text{Re}[\hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_1 \star \hat{\phi}^{Y_2 Z})^*] + 2\text{Re}[\hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z})^*] \\
& + 2\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z_1})^*] + 2\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z_2})^*] \\
& + \hat{X}_1 \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y Z_2})^* + \hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X}_2 \star \hat{\phi}^{Y Z_1})^* + |\hat{X}_1 \star \hat{\phi}^{Y Z_2}|^2 \\
& + 2\text{Re}[\hat{X}_1 \star \hat{\phi}^{Y Z_1} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z})^*] + 2\text{Re}[\hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z})^*] + 2\text{Re}[\hat{X}_1 \star \hat{\phi}^{Y Z_2} (\hat{X}_1 \star \hat{\phi}^{Y_2 Z})^*] \\
& + \hat{X}_1 \star \hat{\phi}^{Y_1 Z} (\hat{X}_2 \star \hat{\phi}^{Y_2 Z})^* + \hat{X}_1 \star \hat{\phi}^{Y_2 Z} (\hat{X}_2 \star \hat{\phi}^{Y_1 Z})^* + |\hat{X}_1 \star \hat{\phi}^{Y_2 Z}|^2 .
\end{aligned} \tag{C10}$$

Using the symmetric estimator of Eq. (C3), we obtain

$$\begin{aligned}
p_2^6 = & 2\text{Re}[\hat{X} \star (\hat{\phi}^{Y Z_1} + \hat{\phi}^{Y_1 Z}) (\hat{X}_2 \star (\hat{\phi}^{Y_1 Z_2} + \hat{\phi}^{Y_2 Z_1}))^*] + 2\text{Re}[\hat{X}_1 \star (\hat{\phi}^{Y Z_2} + \hat{\phi}^{Y_2 Z}) (\hat{X} \star (\hat{\phi}^{Y_1 Z_2} + \hat{\phi}^{Y_2 Z_1}))^*] \\
& + 2\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star (\hat{\phi}^{Y_2 Z_1} + \hat{\phi}^{Y_1 Z_2}))^*] \\
& + \text{Re}[\hat{X}_1 \star (\hat{\phi}^{Y_1 Z} + \hat{\phi}^{Y Z_1}) (\hat{X}_2 \star (\hat{\phi}^{Y_2 Z} + \hat{\phi}^{Y Z_2}))^*] + \text{Re}[\hat{X}_1 \star (\hat{\phi}^{Y Z_2} + \hat{\phi}^{Y_2 Z}) (\hat{X}_2 \star (\hat{\phi}^{Y_1 Z} + \hat{\phi}^{Y Z_1}))^*] \\
& + \hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X} \star (\hat{\phi}^{Y_2 Z_1} + \hat{\phi}^{Y_1 Z_2}))^* + |\hat{X}_1 \star (\hat{\phi}^{Y_2 Z} + \hat{\phi}^{Y Z_2})|^2 \\
= & 8\text{Re}[\hat{X} \star \hat{\phi}_{01} (\hat{X}_2 \star \hat{\phi}_{12})^*] + 8\text{Re}[\hat{X}_1 \star \hat{\phi}_{02} (\hat{X} \star \hat{\phi}_{12})^*] \\
& + 4\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star \hat{\phi}_{12})^*] \\
& + 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{01} \star (\hat{X}_2 \star \hat{\phi}_{02})^*] + 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{02} (\hat{X}_2 \star \hat{\phi}_{01})^*] \\
& + 2|\hat{X} \star \hat{\phi}_{12}|^2 + 4|\hat{X}_1 \star \hat{\phi}_{02}|^2 ,
\end{aligned} \tag{C11}$$

where we use $\hat{X} \star \hat{\phi}^{Y_1 Z_2} (\hat{X} \star (\hat{\phi}^{Y_2 Z_1} + \hat{\phi}^{Y_1 Z_2}))^* = 2|\hat{X} \star \hat{\phi}_{12}|^2$. The quantity, p_3^6 , is given by

$$\begin{aligned} p_3^6 &= 2\hat{X}_1 \star \hat{\phi}^{Y_1 Z_2} (\hat{X}_3 \star \hat{\phi}_{23})^* + 2\hat{X}_1 \star \hat{\phi}^{Y_2 Z_1} (\hat{X}_3 \star \hat{\phi}_{23})^* \\ &\quad + 2\hat{X}_1 \star \hat{\phi}^{Y_2 Z_3} (\hat{X}_1 \star \hat{\phi}_{23})^* + 2\hat{X}_1 \star \hat{\phi}^{Y_2 Z_3} (\hat{X}_3 \star \hat{\phi}_{12})^* \\ &\quad + 2\hat{X}_1 \star \hat{\phi}^{Y_2 Z_3} (\hat{X}_2 \star \hat{\phi}_{13})^* \\ &= 4\hat{X}_1 \star \hat{\phi}_{13} (\hat{X}_2 \star \hat{\phi}_{23})^* + 4\hat{X}_1 \star \hat{\phi}_{23} (\hat{X}_2 \star \hat{\phi}_{13})^* + 2|\hat{X}_1 \star \hat{\phi}_{23}|^2. \end{aligned} \quad (C12)$$

Combining the above quantities, we obtain

$$\begin{aligned} \hat{N} &\equiv 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{01} (\hat{X} \star \hat{\phi})^*] + 4\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X} \star \hat{\phi}_{01})^*] + |\hat{X}_1 \star \hat{\phi}|^2 + 4|\hat{X} \star \hat{\phi}_{01}|^2 \\ &\quad - 8\text{Re}[\hat{X} \star \hat{\phi}_{01} (\hat{X}_2 \star \hat{\phi}_{12})^*] - 8\text{Re}[\hat{X}_1 \star \hat{\phi}_{02} (\hat{X} \star \hat{\phi}_{12})^*] - 4\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star \hat{\phi}_{12})^*] \\ &\quad - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{01} (\hat{X}_2 \star \hat{\phi}_{02})^*] - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{02} (\hat{X}_2 \star \hat{\phi}_{01})^*] - 2|\hat{X} \star \hat{\phi}_{12}|^2 \\ &\quad - 4|\hat{X}_1 \star \hat{\phi}_{02}|^2 + 4\hat{X}_1 \star \hat{\phi}_{13} (\hat{X}_2 \star \hat{\phi}_{23})^* + 4\hat{X}_1 \star \hat{\phi}_{23} (\hat{X}_2 \star \hat{\phi}_{13})^* + 2|\hat{X}_1 \star \hat{\phi}_{23}|^2. \end{aligned} \quad (C13)$$

Using $\hat{X}_{i,\pm j,\dots} = \hat{X}_i \pm \hat{X}_j + \dots$ with $i = 0$ denoting the real data, we find

$$\begin{aligned} 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{01} (\hat{X} \star \hat{\phi})^*] + 4\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X} \star \hat{\phi}_{01})^*] &= 4\text{Re}[\hat{X}_{012} \star \hat{\phi}_{01} (\hat{X}_{01} \star \hat{\phi})^*] \\ - 8\text{Re}[\hat{X} \star \hat{\phi}_{01} (\hat{X}_2 \star \hat{\phi}_{12})^*] - 8\text{Re}[\hat{X}_1 \star \hat{\phi}_{02} (\hat{X} \star \hat{\phi}_{12})^*] &= -8\text{Re}[\hat{X}_{012} \star \hat{\phi}_{13} (\hat{X}_{01} \star \hat{\phi}_{03})^*] \\ - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{01} (\hat{X}_2 \star \hat{\phi}_{02})^*] - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{02} (\hat{X}_2 \star \hat{\phi}_{01})^*] &= -4\text{Re}[\hat{X}_{012} \star \hat{\phi}_{01} (\hat{X}_{12} \star \hat{\phi}_{02})^*] \\ 4\hat{X}_1 \star \hat{\phi}_{13} (\hat{X}_2 \star \hat{\phi}_{23})^* + 4\hat{X}_1 \star \hat{\phi}_{23} (\hat{X}_2 \star \hat{\phi}_{13})^* &= 4\text{Re}[\hat{X}_{012} \star \hat{\phi}_{13} (\hat{X}_{12} \star \hat{\phi}_{23})^*] \\ - 4\text{Re}[\hat{X}_1 \star \hat{\phi} (\hat{X}_2 \star \hat{\phi}_{12})^*] &= -4\text{Re}[\hat{X}_{012} \star \hat{\phi}_{13} (\hat{X}_3 \star \hat{\phi})^*]. \end{aligned} \quad (C14)$$

Here we use the fact that the left-hand side of the above fourth equation is real, and we also exchange the index ($1 \leftrightarrow 3$) in the right-hand side of the last equation. The sum of the above terms becomes

$$\text{Re}[\hat{X}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13}) [\hat{X}_{012} \star (2\hat{\phi} - 2\hat{\phi}_{02} - 4\hat{\phi}_{03} + 2\hat{\phi}_{23}) - 2\hat{X}_3 \star \hat{\phi}]^*]. \quad (C15)$$

The other terms in Eq. (C13) except $|\hat{X}_1 \star \hat{\phi}|^2$ are summarized as

$$- 2|\hat{X} \star \hat{\phi}_{23}|^2 + 2|\hat{X}_1 \star \hat{\phi}_{23}|^2 + 4|\hat{X} \star \hat{\phi}_{02}|^2 - 4|\hat{X}_1 \star \hat{\phi}_{02}|^2 = 2\hat{X}_{012} \star (\hat{\phi}_{02} + \hat{\phi}_{23}) [\hat{X}_{0,-1} \star (2\hat{\phi}_{02} - \hat{\phi}_{23})]^*. \quad (C16)$$

The sum of Eqs. (C15) and (C16) becomes

$$\text{Re}[\hat{X}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13}) \mathcal{B}_{0123}^*], \quad (C17)$$

where we define a realization dependent B mode as

$$\mathcal{B}_{0123} \equiv \hat{X}_{012} \star (2\hat{\phi} - 2\hat{\phi}_{02} - 4\hat{\phi}_{03} + 2\hat{\phi}_{23}) - 2\hat{X}_3 \star \hat{\phi} + \hat{X}_{0,-2} \star (2\hat{\phi}_{01} - \hat{\phi}_{13}) \quad (C18)$$

The estimator for the disconnected part of the six point correlation is recast as

$$\hat{N} = \text{Re}[\hat{X}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13}) \mathcal{B}^*] + |\hat{X}_1 \star \hat{\phi}|^2. \quad (C19)$$

To obtain the delensed power spectrum of Eq. (12), we need to add the term $|\hat{X}_2 \star \hat{\phi}_{11}|^2$ to the above estimator because \hat{N} is the estimator for the all disconnected parts of the six point correlation and $|\hat{X}_2 \star \hat{\phi}_{11}|^2$ is involved in \hat{N} . To recover this term, we add

$$|\hat{X}_2 \star \hat{\phi}|^2 + |\hat{X} \star \hat{\phi}_{11}|^2 - |\hat{X}_2 \star \hat{\phi}_{11}|^2. \quad (C20)$$

This correction is motivated by the following estimator [26]

$$\hat{f} = \langle f \rangle + \frac{\partial \langle f \rangle}{\partial x} (x - \langle f \rangle). \quad (C21)$$

where x is any quantity and f is a function of x . Subtracting Eq. (C20) from \hat{N} , we obtain

$$\hat{N}^{(6),d} \equiv \text{Re}[\hat{X}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13}) \mathcal{B}_{0123}^*] + \hat{X}_{0,-3} \star \hat{\phi}_{11} (\hat{X}_{03} \star \hat{\phi}_{11})^*. \quad (C22)$$

The RD estimator of Eq. (50) is derived by substituting $X = Y = E$, $Z = B$ into Eq. (C22).

3. Six point correlation with suboptimal estimator

Next we derive the suboptimal estimator for the six point correlation. From Eq. (47), the suboptimal estimator is given by

$$\hat{N}^{(6),d,\text{sub}} = p_1^6 - 2p_2^6 + 4p_3^6 + [(\hat{X}\hat{\phi}^{XY_2} - \hat{X}_1\hat{\phi}^{Y_1Z_2})\hat{X}_2\hat{\phi}^{Y_2Z_2} + (14 \text{ perms.})]. \quad (\text{C23})$$

Using Eqs. (C8), (C11) and (C12), we obtain

$$\begin{aligned} p_1^6 - 2p_2^6 + 4p_3^6 = & 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{01}(\hat{X} \star \hat{\phi})^*] + 4\text{Re}[\hat{X}_1 \star \hat{\phi}(\hat{X} \star \hat{\phi}_{01})^*] + 4|\hat{X} \star \hat{\phi}_{01}|^2 + |\hat{X}_1 \star \hat{\phi}|^2 \\ & - 16\text{Re}[\hat{X} \star \hat{\phi}_{01}(\hat{X}_2 \star \hat{\phi}_{12})^*] - 16\text{Re}[\hat{X}_1 \star \hat{\phi}_{02}(\hat{X} \star \hat{\phi}_{12})^*] \\ & - 8\text{Re}[\hat{X}_1 \star \hat{\phi}(\hat{X}_2 \star \hat{\phi}_{12})^*] - 8\text{Re}[\hat{X}_1 \star \hat{\phi}_{01}(\hat{X}_2 \star \hat{\phi}_{02})^*] - 8\text{Re}[\hat{X}_1 \star \hat{\phi}_{02}(\hat{X}_2 \star \hat{\phi}_{01})^*] \\ & - 4|\hat{X} \star \hat{\phi}_{12}|^2 - 8|\hat{X}_1 \star \hat{\phi}_{02}|^2 \\ & + 4[4\hat{X}_{12} \star \hat{\phi}_{13}(\hat{X}_{12} \star \hat{\phi}_{23})^* + 2|\hat{X}_1 \star \hat{\phi}_{23}|^2]. \end{aligned} \quad (\text{C24})$$

The last term of Eq. (C23) is given by

$$\begin{aligned} & 4\text{Re}[\hat{X} \star \hat{\phi}_{01}(\hat{X}_1 \star \hat{\phi}_{11})^*] + 4\text{Re}[\hat{X} \star \hat{\phi}_{11}(\hat{X}_1 \star \hat{\phi}_{01})^*] + 4|\hat{X}_1 \star \hat{\phi}_{01}|^2 \\ & - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{12}(\hat{X}_2 \star \hat{\phi}_{22})^*] - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{22}(\hat{X}_2 \star \hat{\phi}_{12})^*] - 4|\hat{X}_1 \star \hat{\phi}_{12}|^2 + |\hat{X} \star \hat{\phi}_{11}|^2 - |\hat{X}_1 \star \hat{\phi}_{22}|^2. \end{aligned} \quad (\text{C25})$$

Combining Eqs. (C24) and (C25), we obtain

$$\begin{aligned} \hat{N}^{(6),d,\text{sub}} = & 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{01}(\hat{X} \star \hat{\phi})^*] + 4\text{Re}[\hat{X}_1 \star \hat{\phi}(\hat{X} \star \hat{\phi}_{01})^*] + 4|\hat{X} \star \hat{\phi}_{01}|^2 \\ & - 16\text{Re}[\hat{X} \star \hat{\phi}_{01}(\hat{X}_2 \star \hat{\phi}_{12})^*] - 16\text{Re}[\hat{X}_1 \star \hat{\phi}_{02}(\hat{X} \star \hat{\phi}_{12})^*] \\ & - 8\text{Re}[\hat{X}_1 \star \hat{\phi}(\hat{X}_2 \star \hat{\phi}_{12})^*] - 8\text{Re}[\hat{X}_1 \star \hat{\phi}_{01}(\hat{X}_2 \star \hat{\phi}_{02})^*] - 8\text{Re}[\hat{X}_1 \star \hat{\phi}_{02}(\hat{X}_2 \star \hat{\phi}_{01})^*] \\ & - 4|\hat{X} \star \hat{\phi}_{12}|^2 - 8|\hat{X}_1 \star \hat{\phi}_{02}|^2 \\ & + 4[4\hat{X}_{12} \star \hat{\phi}_{13}(\hat{X}_{12} \star \hat{\phi}_{23})^* + 2|\hat{X}_1 \star \hat{\phi}_{23}|^2] \\ & + 4\text{Re}[\hat{X} \star \hat{\phi}_{01}(\hat{X}_1 \star \hat{\phi}_{11})^*] + 4\text{Re}[\hat{X} \star \hat{\phi}_{11}(\hat{X}_1 \star \hat{\phi}_{01})^*] + 4|\hat{X}_1 \star \hat{\phi}_{01}|^2 \\ & - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{12}(\hat{X}_2 \star \hat{\phi}_{22})^*] - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{22}(\hat{X}_2 \star \hat{\phi}_{12})^*] - 4|\hat{X}_1 \star \hat{\phi}_{12}|^2, \end{aligned} \quad (\text{C26})$$

where $|\hat{X}_1 \star \hat{\phi}|^2 + |\hat{X} \star \hat{\phi}_{11}|^2 - |\hat{X}_1 \star \hat{\phi}_{22}|^2$ is subtracted from the above equation.

To simplify the above equation, we use

$$4|\hat{X} \star \hat{\phi}_{01}|^2 - 4|\hat{X} \star \hat{\phi}_{12}|^2 - 8|\hat{X}_1 \star \hat{\phi}_{02}|^2 + 8|\hat{X}_1 \star \hat{\phi}_{23}|^2 = \text{Re}[\hat{X}_{012} \star (2\hat{\phi}_{02} + 2\hat{\phi}_{23})(2(\hat{X}_0 - 2\hat{X}_1) \star (\hat{\phi}_{02} - \hat{\phi}_{23}))^*]. \quad (\text{C27})$$

Using Eq. (C14) and the above equation (with exchanging the index, $1 \leftrightarrow 2$), we obtain

$$\begin{aligned} \hat{N}^{(6),d,\text{sub}} = & \text{Re}[\hat{X}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13})[\hat{X}_{012} \star (2\hat{\phi} - 4\hat{\phi}_{02} - 8\hat{\phi}_{03} + 8\hat{\phi}_{23}) - 4\hat{X}_3 \star \hat{\phi} + 2(\hat{X}_0 - 2\hat{X}_2) \star (\hat{\phi}_{01} - \hat{\phi}_{13})]^*] \\ & + 4\text{Re}[\hat{X} \star \hat{\phi}_{01}(\hat{X}_1 \star \hat{\phi}_{11})^*] + 4\text{Re}[\hat{X} \star \hat{\phi}_{11}(\hat{X}_1 \star \hat{\phi}_{01})^*] + 4|\hat{X}_1 \star \hat{\phi}_{01}|^2 \\ & - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{12}(\hat{X}_2 \star \hat{\phi}_{22})^*] - 4\text{Re}[\hat{X}_1 \star \hat{\phi}_{22}(\hat{X}_2 \star \hat{\phi}_{12})^*] - 4|\hat{X}_1 \star \hat{\phi}_{12}|^2, \end{aligned} \quad (\text{C28})$$

The sum of the second and third lines of the above equation is given by

$$4\text{Re}[\hat{X}_{013} \star (\hat{\phi}_{01} + \hat{\phi}_{12})[\hat{X}_{013} \star \hat{\phi}_{11} - \hat{X}_2 \star (\hat{\phi}_{11} + \hat{\phi}_{22}) + \hat{X}_1 \star (\hat{\phi}_{01} - \hat{\phi}_{12})]^*]. \quad (\text{C29})$$

Substituting the above equation into Eq. (C28), we find

$$\hat{N}^{(6),d,\text{sub}} = \text{Re}[\hat{X}_{012} \star (2\hat{\phi}_{01} + 2\hat{\phi}_{13})(\mathcal{B}_{0123}^{\text{sub}})^*], \quad (\text{C30})$$

where we define a realization-dependent B mode as

$$\mathcal{B}_{0123}^{\text{sub}} = \hat{X}_{012} \star (2\hat{\phi} - 4\hat{\phi}_{02} - 8\hat{\phi}_{03} + 8\hat{\phi}_{23} + 2\hat{\phi}_{11}) - 2\hat{X}_3 \star (2\hat{\phi} + \hat{\phi}_{11} + \hat{\phi}_{33}) + 2(\hat{X}_{01} - 2\hat{X}_2) \star (\hat{\phi}_{01} - \hat{\phi}_{13}). \quad (\text{C31})$$

Eq. (52) is derived with $X = Y = E$ and $Z = B$ in Eq. (C30).

4. Six point correlation relevant to the dominant term of the delensing bias

Here we consider the RD estimator for Eq. (25). To extract the corresponding terms of Eq. (25) in Eq. (47), we first rewrite Eq. (47) as

$$[(\bar{x}_{i_1}\bar{x}_{i_2} - \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle)(\langle\bar{x}_{i_3}\cdots\bar{x}_{i_6}\rangle_c + \langle\bar{x}_{i_3}\cdots\bar{x}_{i_6}\rangle_d) + \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle\bar{k}_{i_3i_4i_5i_6}] + (14 \text{ perms.}) + p_3^6. \quad (\text{C32})$$

Here $\bar{k}_{i_3i_4i_5i_6}$ is the unnormalized estimator of the trispectrum defined in Eq. (43).

We first consider the term containing the connected four point correlation in Eq. (25). The corresponding terms in Eq. (C32) are given by

$$(\bar{x}_{i_1}\bar{x}_{i_2} - \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle)(\langle\bar{x}_{i_3}\cdots\bar{x}_{i_6}\rangle_c + \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle\bar{k}_{i_3i_4i_5i_6}). \quad (\text{C33})$$

The permutations of the above quantity are not relevant to Eq. (25). The above equation is recast as

$$\begin{aligned} & (\bar{x}_{i_1}\bar{x}_{i_2} - \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle)(\langle\bar{x}_{i_3}\cdots\bar{x}_{i_6}\rangle - \langle\bar{x}_{i_3}\cdots\bar{x}_{i_6}\rangle_d) + \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle(\bar{x}_{i_3}\cdots\bar{x}_{i_6} - \langle\bar{x}_{i_3}\bar{x}_{i_4}\rangle\bar{x}_{i_5}\bar{x}_{i_6} - (5 \text{ perms.}) + \langle\bar{x}_{i_3}\cdots\bar{x}_{i_6}\rangle_d) \\ &= (\bar{x}_{i_1}\bar{x}_{i_2} - \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle)(\bar{x}_{i_3}\cdots\bar{x}_{i_6}) + \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle(\bar{x}_{i_3}\cdots\bar{x}_{i_6} - \langle\bar{x}_{i_3}\bar{x}_{i_4}\rangle\bar{x}_{i_5}\bar{x}_{i_6} - (5 \text{ perms.})) \\ &+ (2\langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle - \bar{x}_{i_1}\bar{x}_{i_2})(\langle\bar{x}_{i_3}\cdots\bar{x}_{i_6}\rangle_d). \end{aligned} \quad (\text{C34})$$

From the above equation, the RD estimator for the second term of Eq. (25) is obtained by replacing $\bar{x}_{i_1} \rightarrow \hat{E}$, $\bar{x}_{i_2}\bar{x}_{i_3} \rightarrow \hat{\phi}^{EB}$, $\bar{x}_{i_4} \rightarrow \hat{E}$ and $\bar{x}_{i_5}\bar{x}_{i_6} \rightarrow \hat{\phi}^{EB}$. The result is

$$\begin{aligned} \hat{N}^{\text{dom},1} &\equiv (\hat{E} \star \hat{\phi}^{EB_2} - \hat{E}_1 \star \hat{\phi}^{E_1 B_2})(\hat{E}_2 \star \hat{\phi}_{22})^* \\ &+ \hat{E}_1 \star \hat{\phi}^{E_1 B}(\hat{E} \star \hat{\phi})^* - \hat{E}_1 \star \hat{\phi}^{E_1 B}(\hat{E} \star \hat{\phi}_{22} + 2\hat{E}_2 \star \hat{\phi}_{02})^* - \hat{E}_1 \star \hat{\phi}^{E_1 B_2}(2\hat{E} \star \hat{\phi}_{02} + \hat{E}_2 \star \hat{\phi})^* \\ &+ 2(2\hat{E}_1 \star \hat{\phi}^{E_1 B_2} - \hat{E} \star \hat{\phi}^{EB_2})(\hat{E}_3 \star \hat{\phi}_{23})^* \end{aligned} \quad (\text{C35})$$

$$\begin{aligned} &= (\hat{E} \star \hat{\phi}^{EB_2} - \hat{E}_1 \star \hat{\phi}^{E_1 B_2})(\hat{E}_2 \star \hat{\phi}_{22} - 2\hat{E}_3 \star \hat{\phi}_{23})^* \\ &+ \hat{E}_1 \star \hat{\phi}^{E_1 B}[\hat{E} \star (\hat{\phi} - \hat{\phi}_{22}) - 2\hat{E}_2 \star \hat{\phi}_{02}]^* \\ &+ \hat{E}_1 \star \hat{\phi}^{E_1 B_2}(2\hat{E}_3 \star \hat{\phi}_{23} - 2\hat{E} \star \hat{\phi}_{02} - \hat{E}_2 \star \hat{\phi})^*. \end{aligned} \quad (\text{C36})$$

To simplify the above equation, we use

$$\begin{aligned} &(\text{the first line of Eq. (C36)}) = \hat{E}_{0,-1} \star \hat{\phi}^{E_{01} B_{02}}[\hat{E}_{23} \star (\hat{\phi}_{22} - 2\hat{\phi}_{23})]^* \\ &(\text{the second line of Eq. (C36)}) = \hat{E}_1 \star \hat{\phi}^{E_{01} B_{02}}[\hat{E} \star (\hat{\phi} - \hat{\phi}_{22}) - 2\hat{E}_2 \star \hat{\phi}_{02}]^* \\ &(\text{the third line of Eq. (C36)}) = \hat{E}_1 \star \hat{\phi}^{E_{01} B_{02}}(2\hat{E}_3 \star \hat{\phi}_{23} - 2\hat{E} \star \hat{\phi}_{02} - \hat{E}_2 \star \hat{\phi})^*. \end{aligned} \quad (\text{C37})$$

Note that the ensemble average of the above first and third equations is zero while that of the second equation gives the connected part in Eq. (25). With the above equations, Eq. (C36) becomes

$$\hat{N}^{\text{dom},1} = \hat{E}_{0,-1} \star \hat{\phi}^{E_{01} B_{02}}[\hat{E}_{23} \star (\hat{\phi}_{22} - 2\hat{\phi}_{23})]^* + \hat{E}_1 \star \hat{\phi}^{E_{01} B_{02}}[\hat{E} \star (\hat{\phi} - \hat{\phi}_{22} - 2\hat{\phi}_{02}) - \hat{E}_2 \star (\hat{\phi} + 2\hat{\phi}_{02}) + 2\hat{E}_3 \star \hat{\phi}_{23}]^*. \quad (\text{C38})$$

Next we consider the full disconnected part of the six point correlation (i.e., the third term) involved in Eq. (25). The corresponding terms in Eq. (C32) are described by

$$[(\bar{x}_{i_1}\bar{x}_{i_2} - \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle)(\langle\bar{x}_{i_3}^1\bar{x}_{i_4}^2\bar{x}_{i_5}^2\bar{x}_{i_6}^1\rangle_{1,2} + (i_1,i_2) \leftrightarrow (i_3,i_4) + (i_1,i_2) \leftrightarrow (i_5,i_6))] + \langle\bar{x}_{i_1}\bar{x}_{i_2}\rangle\langle\bar{x}_{i_3}^1\bar{x}_{i_4}^2\bar{x}_{i_5}^2\bar{x}_{i_6}^1\rangle_{1,2}. \quad (\text{C39})$$

Replacing $\bar{x}_{i_1} \rightarrow \hat{E}$, $\bar{x}_{i_2}\bar{x}_{i_3} \rightarrow \hat{\phi}^{EB}$, $\bar{x}_{i_4} \rightarrow \hat{E}$ and $\bar{x}_{i_5}\bar{x}_{i_6} \rightarrow \hat{\phi}^{EB}$, we obtain the RD estimator for the third term of Eq. (25) as

$$\begin{aligned} \hat{N}^{\text{dom},2} &\equiv 2\hat{E} \star \hat{\phi}^{EB_1}(\hat{E}_2 \star \hat{\phi}^{E_2 B_1})^* + \hat{E}_1 \star \hat{\phi}^{E_1 B}(\hat{E}_2 \star \hat{\phi}^{E_2 B})^* - 2\hat{E}_1 \star \hat{\phi}^{E_1 B_2}(\hat{E}_3 \star \hat{\phi}^{E_3 B_2})^* \\ &= \hat{E}_1 \star \hat{\phi}^{E_{01} B_{02}}(2\hat{E} \star \hat{\phi}^{EB_2} + \hat{E}_2 \star \hat{\phi}^{E_2 B} - 2\hat{E}_3 \star \hat{\phi}^{E_3 B_2})^*. \end{aligned} \quad (\text{C40})$$

Combining $\hat{N}^{\text{dom},1}$ and $\hat{N}^{\text{dom},2}$, the RD estimator of Eq. (25) is given as

$$\begin{aligned}
\hat{N}^{(6),\text{dom}} &= 2\hat{N}^{\text{dom},1} + \hat{N}^{\text{dom},2} \\
&= 2\hat{E}_{0,-1} \star \hat{\phi}^{E_{01}B_{02}} [\hat{E}_{23} \star (\hat{\phi}_{22} - 2\hat{\phi}_{23})]^* \\
&\quad + \hat{E}_1 \star \hat{\phi}^{E_{01}B_{02}} [2\hat{E} \star (\hat{\phi} - \hat{\phi}_{22} - 2\hat{\phi}_{02} + \hat{\phi}^{E_{B_2}}) - \hat{E}_2 \star (2\hat{\phi} + 4\hat{\phi}_{02} - \hat{\phi}^{E_{B_2}}) + 2\hat{E}_3 \star (2\hat{\phi}_{23} - \hat{\phi}^{E_{B_3}})]^* \\
&= 2\hat{E}_{0,-1} \star \hat{\phi}^{E_{01}B_{02}} [\hat{E}_{23} \star (\hat{\phi}_{22} - 2\hat{\phi}_{23})]^* \\
&\quad + \hat{E}_1 \star \hat{\phi}^{E_{01}B_{02}} [2\hat{E} \star (\hat{\phi} - \hat{\phi}_{22} - \hat{\phi}^{E_{B_2}}) - \hat{E}_2 \star (2\hat{\phi} + 2\hat{\phi}^{E_{B_2}} + \hat{\phi}^{E_{B_3}}) + 2\hat{E}_3 \star \hat{\phi}^{E_{B_3}}]^* \\
&= 2\hat{E}_{0,-1} \star \hat{\phi}^{E_{01}B_{02}} [\hat{E}_{23} \star (\hat{\phi}_{22} - 2\hat{\phi}_{23})]^* \\
&\quad + \hat{E}_1 \star \hat{\phi}^{E_{01}B_{02}} [2\hat{E} \star (\hat{\phi} - \hat{\phi}_{22} - \hat{\phi}^{E_{B_2}}) - \hat{E}_2 \star (\hat{\phi}^{(2E+E_2)B_{02}} - \hat{\phi}_{22}) + 2\hat{E}_3 \star \hat{\phi}^{E_{B_3}}]^*. \tag{C41}
\end{aligned}$$

-
- [1] A. G. Polnarev, *Sov. Astron.* **29**, 607 (1985).
[2] A. Lue, L. Wang, and M. Kamionkowski, *Phys. Rev. Lett.* **83**, 1506 (1999), [astro-ph/9812088](#).
[3] M. Kamionkowski, A. Kosowsky, and A. Stebbins, *Phys. Rev. Lett.* **78**, 2058 (1997), [astro-ph/9609132](#).
[4] U. Seljak and M. Zaldarriaga, *Phys. Rev. Lett.* **78**, 2054 (1997), [astro-ph/9609169](#).
[5] U. Seljak and A. Slosar, *Phys. Rev. D* **74**, 063523 (2006).
[6] J. R. Shaw and A. Lewis, *Phys. Rev. D* **81**, 043517 (2010), [arXiv:0911.2714](#).
[7] L. Pogosian, P. S. A. Yadav, Y.-F. Ng, and T. Vachaspati, *Phys. Rev. D* **84**, 043530 (2011), [arXiv:1106.1438](#).
[8] W.-H. Teng, C.-L. Kuo, and J.-H. P. Wu (2011), [arXiv:1102.5729](#).
[9] M. Zaldarriaga and U. Seljak, *Phys. Rev. D* **58**, 023003 (1998), [astro-ph/9803150](#).
[10] D. Hanson et al. (SPTpol), *Phys. Rev. Lett.* **111**, 141301 (2013), [arXiv:1307.5830](#).
[11] POLARBEAR Collaboration, *Astrophys. J.* **794**, 171 (2014).
[12] ACT Collaboration (Sherwin, Blake D. et. al.) (2016), [arXiv:1611.09753](#).
[13] *Planck* Collaboration (2016), [arXiv:1512.02882](#).
[14] BICEP2 / *Keck Array* Collaboration VIII, *Astrophys. J.* **833**, 228 (2016), [arXiv:1606.01968](#).
[15] M. Kesden, A. Cooray, and M. Kamionkowski, *Phys. Rev. Lett.* **89**, 011304 (2002), [astro-ph/0202434](#).
[16] L. Knox and Y.-S. Song, *Phys. Rev. Lett.* **89**, 011303 (2002), [astro-ph/0202286](#).
[17] D. Green, J. Meyers, and A. van Engelen (2016), [arXiv:1609.08143](#).
[18] U. Seljak and C. M. Hirata, *Phys. Rev. D* **69**, 043005 (2004), [astro-ph/0310163](#).
[19] K. M. Smith et al., *J. Cosmol. Astropart. Phys.* **1206**, 014 (2012), [arXiv:1010.0048](#).
[20] G. Simard, D. Hanson, and G. Holder, *Astrophys. J.* **807**, 166 (2015).
[21] B. D. Sherwin and M. Schmittfull, *Phys. Rev. D* **92**, 043005 (2015).
[22] P. Larsen, A. Challinor, B. D. Sherwin, and D. Mak, *Phys. Rev. Lett.* **117**, 151102 (2016), [arXiv:1607.05733](#).
[23] A. Manzotti et al. (2017), [arXiv:1701.04396](#).
[24] T. Namikawa and R. Nagata, *J. Cosmol. Astropart. Phys.* **1409**, 009 (2014), [arXiv:1405.6568](#).
[25] N. Sehgal, M. S. Madhavacheril, B. Sherwin, and A. van Engelen (2016), [arXiv:1612.03898](#).
[26] J. Carron, A. Lewis, and A. Challinor (2017), [arXiv:1701.01712](#).
[27] T. Namikawa, D. Hanson, and R. Takahashi, *Mon. Not. R. Astron. Soc.* **431**, 609 (2013), [arXiv:1209.0091](#).
[28] *Planck* Collaboration, *Astron. Astrophys.* **571**, A17 (2014).
[29] *Planck* Collaboration (2015), [arXiv:1502.01591](#).
[30] D. M. Regan, E. P. S. Shellard, and J. R. Fergusson, *Phys. Rev. D* **82**, 023520 (2010), [arXiv:1004.2915](#).
[31] A. Lewis and A. Challinor, *Phys. Rep.* **429**, 1 (2006), [astro-ph/0601594](#).
[32] W. Hu and T. Okamoto, *Astrophys. J.* **574**, 566 (2002), [astro-ph/0111606](#).
[33] D. Hanson, A. Challinor, G. Efstathiou, and P. Bielewicz, *Phys. Rev. D* **83**, 043005 (2011), [arXiv:1008.4403](#).
[34] A. Lewis, A. Challinor, and D. Hanson, *J. Cosmol. Astropart. Phys.* **03**, 018 (2011), [arXiv:1101.2234](#).
[35] K. T. Story et al. (SPTpol), *Astrophys. J.* **810**, 50 (2015).
[36] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000), [astro-ph/9911177](#).
[37] *Planck* Collaboration, *Astron. Astrophys.* **594**, A13 (2016), [arXiv:1502.01589](#).
[38] T. Louis, S. Naess, S. Das, J. Dunkeley, and B. Sherwin, *Mon. Not. R. Astron. Soc.* **435**, 2040 (2013), [arXiv:1306.6692](#).
[39] L. Knox, *Phys. Rev. D* **60**, 103516 (1999), [astro-ph/9902046](#).
[40] A. Challinor and A. Lewis, *Phys. Rev. D* **71**, 103010 (2005), [astro-ph/0502425](#).
[41] L. Amendola, *Mon. Not. R. Astron. Soc.* **283**, 983 (1996).
[42] T. Namikawa, *Phys. Rev. D* **93**, 121301 (2016), [arXiv:1604.08578](#).
[43] V. Boehm, M. Schmittfull, and B. Sherwin (2016), [arXiv:1605.01392](#).
[44] G. Pratten and A. Lewis (2016), [arXiv:1605.05662](#).
[45] J. Liu, J. C. Hill, B. Sherwin, A. Petri, V. Boehm, and H. Zoltan, *Phys. Rev. D* **94**, 103501 (2016), [arXiv:1608.03169](#).
[46] W. Hu, *Astrophys. J.* **557**, L79 (2001), [astro-ph/0105424](#).