



CHORUS

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

Quadratic estimators for CMB weak lensing

Abhishek S. Maniyar, Yacine Ali-Haïmoud, Julien Carron, Antony Lewis, and Mathew S. Madhavacheril

Phys. Rev. D **103**, 083524 — Published 23 April 2021

DOI: [10.1103/PhysRevD.103.083524](https://doi.org/10.1103/PhysRevD.103.083524)

Quadratic estimators for CMB weak lensing

Abhishek S. Maniyar and Yacine Ali-Haïmoud
*Center for Cosmology and Particle Physics, Department of Physics,
New York University, New York, NY 10003, USA*

Julien Carron
*Université de Genève, Département de Physique Théorique et CAP,
24 Quai Ansermet, CH-1211 Genève 4, Switzerland*

Antony Lewis
Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

Mathew S. Madhavacheril
Perimeter Institute for Theoretical Physics, Waterloo, ON, Canada N2L 2Y5

In recent years, weak lensing of the cosmic microwave background (CMB) has emerged as a powerful tool to probe fundamental physics, such as neutrino masses, primordial non-Gaussianity, dark energy, and modified gravity. The prime target of CMB lensing surveys is the lensing potential, which is reconstructed from the observed CMB temperature T and polarization E and B fields. Until very recently, this reconstruction has been performed with quadratic estimators (QEs), which, although known to be suboptimal for high-sensitivity experiments, are numerically efficient, and useful to make forecasts and cross-check the results of more sophisticated likelihood-based methods. It is expected that ongoing and near-future CMB experiments such as AdvACT, SPT-3G and the Simons Observatory (SO), will also rely on QEs. In this work, we review different QEs, and clarify and quantify their differences. In particular, we show that the Hu-Okamoto (HO02) estimator is not the absolute optimal lensing estimator that can be constructed out of quadratic combinations of T , E and B fields. Instead, we derive the global-minimum-variance (GMV) lensing quadratic estimator. Although this estimator can be found elsewhere in the literature, it was erroneously described as equivalent to the HO02 estimator, and has never been used in real data analyses. Here, we show explicitly that the HO02 estimator is suboptimal to the GMV estimator, with a reconstruction noise larger by up to $\sim 9\%$ for a SO-like experiment. We further show that the QE used in the Planck, and recent SPT lensing analysis is suboptimal to both the HO02 and GMV estimator, and would have a reconstruction noise up to $\sim 11\%$ larger than that of the GMV estimator for a SO-like experiment. In addition to clarifying differences between different QEs, this work should thus provide motivation to implement the GMV estimator in future lensing analyses relying on QEs.

I. INTRODUCTION

Weak gravitational lensing of the cosmic microwave background (CMB) arises from the deflection of CMB photons as they travel to us from the last scattering surface, through the inhomogeneous Universe [1]; see e.g. Ref. [2] for a review. The deflection angle is proportional to the gradient of the lensing potential ϕ , which is determined by the projected mass distribution along the line of sight. Reconstructing ϕ is therefore a powerful cosmological tool, as it gives direct access to the projected distribution of the *total* matter – baryonic and dark – without relying on biased tracers [3]. Among other applications, the power spectrum of the lensing potential and its cross-correlation with other tracers of large-scale structure are a sensitive probe of the growth of matter fluctuations, primordial non-Gaussianity, neutrino masses, dark energy, and modified gravity [2, 4, 5]. CMB lensing has been successfully measured by ACT, SPT, *Planck*, BICEP and POLARBEAR [6–17]. Current and upcoming wide-field CMB experiments such as AdvACT [18], SPT-3G [19] and the Simons Observa-

tory (SO) [20] will measure the lensing potential with even higher signal-to-noise ratio. Looking ahead, next-generation “Stage-4” instrumental concepts with unprecedented depth and angular resolution are currently under development [21], with CMB lensing as one of their main science goals [22].

One of the main signatures of weak lensing is the induced correlations between unequal Fourier modes of the CMB temperature and polarization fields. It is therefore natural to seek to estimate ϕ out of linear combinations of terms quadratic in different modes of the observed fields [23, 24]; and indeed, almost all CMB lensing analyses thus far have relied on such quadratic estimators. For the next-generation Stage-4-like CMB experiments (CMBS4), quadratic estimators are known to be suboptimal, especially for polarization [25]. More elaborate algorithms are being developed, such as the gradient-inversion method [26], or likelihood-based methods [25, 27–29]. Meanwhile, quadratic estimators remain the workhorse tool for current and near-future CMB experiments like AdvACT, SPT-3G, and SO. They have the advantages of being very simple to implement and computationally efficient, and will serve as useful cross-checks

even when more accurate and computationally demanding methods are employed with future data.

The main goal of this paper is to clarify and quantify the differences between several quadratic estimators commonly used for CMB lensing reconstruction. Our most important point is that the well-known Hu and Okamoto [30] (hereafter, HO02) estimator is *not* the optimal quadratic estimator that can be constructed from temperature and polarization maps, even if generalized to the full sky, and even when using non-perturbative response functions [31, 32]. Instead, we derive the global-minimum-variance (hereafter GMV) quadratic estimator built out of all possible quadratic combination of T, E and B . The GMV estimator was in fact first derived in Hirata and Seljak [25], as the weak-signal limit of their likelihood-based method. Nevertheless, it was stated there and in subsequent works that this estimator is equivalent to that of HO02. We explicitly show that this is not the case, and that the reconstruction noise of the GMV estimator can be up to $\sim 9\%$ lower than that of the HO02 estimator on large angular scales. We also generalize it to be accurately unbiased accounting for higher-order lensing effects. Furthermore, we show that the quadratic estimator used in the Planck collaboration [9, 10] and SPT collaboration [13] lensing analyses, obtained by neglecting C_l^{TE} in the inverse filter matrix, is suboptimal to *both* the GMV and HO02 estimators. For a SO-like experiment, this suboptimal estimator is up to $\sim 11\%$ noisier than the GMV estimator. This may motivate implementing the GMV estimator in future analyses, despite the possible added complexity of jointly filtering temperature and polarization maps.

The remainder of this paper is organized as follows. After introducing our notation and convention in Sec. II, we review the HO02 estimator and its close cousin, the Okamoto-Hu [33] (hereafter OH03) estimator in Sec. III. We then derive the GMV estimator in Sec. IV and explicitly show how it differs from the HO02 estimator. We describe the suboptimal lensing estimator of Ref. [10] in Sec. V. Finally, we compare estimators in Sec. VI for different instrumental setups, and conclude in Sec. VII.

II. NOTATION AND CONVENTIONS

We denote by capital letters $X, Y = T, E, B$ the observed (lensed and noisy) CMB temperature and polarization fields, and by ϕ the projected lensing potential. Throughout we work in the flat-sky approximation; we denote two-dimensional Fourier wavenumbers by \mathbf{l} for CMB fields and \mathbf{L} for the lensing potential.

The power spectra of the observed temperature and polarization fields are defined as

$$\langle X(\mathbf{l})Y(\mathbf{l}') \rangle = (2\pi)^2 \delta(\mathbf{l} + \mathbf{l}') C_l^{XY}, \quad (1)$$

where C_l^{XY} is the total cross-power spectrum of the lensed fields, including detector noise added in quadrature (for $X = Y$). It can also include contributions

α	$f_\alpha(\mathbf{l}_1, \mathbf{l}_2)$
TT	$\tilde{C}_{l_1}^{T\nabla T}(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{T\nabla T}(\mathbf{L} \cdot \mathbf{l}_2)$
TE	$\tilde{C}_{l_1}^{T\nabla E} \cos 2\varphi_{l_1 l_2}(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{T\nabla E}(\mathbf{L} \cdot \mathbf{l}_2)$
EE	$[\tilde{C}_{l_1}^{E\nabla E}(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{E\nabla E}(\mathbf{L} \cdot \mathbf{l}_2)] \cos 2\varphi_{l_1 l_2}$
TB	$\tilde{C}_{l_1}^{T\nabla E} \sin 2\varphi_{l_1 l_2}(\mathbf{L} \cdot \mathbf{l}_1)$
EB	$[\tilde{C}_{l_1}^{E\nabla E}(\mathbf{L} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{E\nabla B}(\mathbf{L} \cdot \mathbf{l}_2)] \sin 2\varphi_{l_1 l_2}$

TABLE I: CMB lensing correlation coefficients. $\varphi_{l_1 l_2}$ is the angle between \mathbf{l}_1 and \mathbf{l}_2 . The quantity $\tilde{C}_l^{X\nabla Y}$ is the lensed gradient spectrum, defined in Refs. [31, 32]. Note that we do not include curl-like terms $\tilde{C}_l^{TP_1}, \tilde{C}_l^{PP_1}$, which are always subdominant [32].

from other sources of variance such as residual foreground contamination. In this expression the angular brackets denote taking ensemble averages over the primordial CMB, detector noise, along with the underlying large scale structure.

Gravitational lensing affects the auto- and cross-power spectra of CMB fields, and moreover produces correlations between non-opposite \mathbf{l} modes, proportional to the projected lensing potential. The response of the non-opposite correlations to lensing can be quantified by non-perturbative response functions f_{XY} defined by

$$\left\langle \frac{\delta}{\delta\phi(\mathbf{L})} (X(\mathbf{l})Y(\mathbf{l}')) \right\rangle = \delta(\mathbf{l} + \mathbf{l}' - \mathbf{L}) f_{XY}(\mathbf{l}, \mathbf{l}'). \quad (2)$$

The coupling coefficients f_{XY} appearing in Eq. (2) are given explicitly in Table I. They depend on the lensed gradient spectra $\tilde{C}_l^{X\nabla Y}$ defined in Refs. [31, 32], which generalize the unlensed spectra used in the original work of HO02 so that the response function for each lensing mode includes the important higher-order effect of other lensing modes. The BB term has negligible contribution to the signal-to-noise ratio of the reconstructed ϕ field and thus we neglect it in our analysis. Note that different foregrounds can also contribute to off-diagonal correlations [e.g. 34, 35], but we do not include them in this work.

In the remainder of this work we will describe different estimators $\hat{\phi}_\alpha$ for the lensing potential. All these estimators are required to be unbiased, i.e. such that $\langle \hat{\phi}_\alpha \rangle = \phi$. They are however noisy, and we define their variance (or reconstruction noise) $N_\alpha(\mathbf{L})$ through

$$\langle (\hat{\phi}_\alpha - \phi)(\mathbf{L})(\hat{\phi}_\alpha - \phi)(\mathbf{L}') \rangle = (2\pi)^2 \delta_{\mathbf{D}}(\mathbf{L} + \mathbf{L}') N_\alpha(\mathbf{L}). \quad (3)$$

Here, for optimizing the signal to noise, we only consider the primary Gaussian disconnected contractions of the lensed fields, $N_\alpha^{(0)}(\mathbf{L})$; in Appendix. A we give an explicit form for the $N_\alpha^{(1)}(\mathbf{L})$ contractions [36] that should also be included in any full data likelihood analysis. The superscript values 0 and 1 represent the order to which the variance $N_\alpha(\mathbf{L})$ explicitly depends on $C_L^{\phi\phi}$.

We will often deal with convolutions in Fourier space, and for brevity, introduce the compact notation

$$\int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \dots \equiv \iint \frac{d^2l_1 d^2l_2}{(2\pi)^2} \delta_{\mathbf{D}}(\mathbf{l}_1 + \mathbf{l}_2 - \mathbf{L}) \dots \quad (4)$$

We define the Fourier transform of a configuration-space function $A(\hat{\mathbf{n}})$ as

$$\mathcal{F}[A(\hat{\mathbf{n}})](\mathbf{l}) \equiv \int d^2\hat{\mathbf{n}} A(\hat{\mathbf{n}}) e^{-i\mathbf{l}\cdot\hat{\mathbf{n}}}, \quad (5)$$

and the inverse-Fourier transform of a harmonic-space function $B(\mathbf{l})$ as

$$\mathcal{F}^{-1}[B(\mathbf{l})](\hat{\mathbf{n}}) \equiv \int \frac{d^2\mathbf{l}}{(2\pi)^2} B(\mathbf{l}) e^{i\mathbf{l}\cdot\hat{\mathbf{n}}}. \quad (6)$$

III. HU AND OKAMOTO ESTIMATORS

We now briefly rederive the HO02 and OH03 quadratic estimators for the lensing potential, setting the stage for our subsequent derivation of the global-minimum-variance estimator.

The approach of HO02 consists in constructing the single-pair estimators $\hat{\phi}_\alpha(\mathbf{L})$ separately for each pair $\alpha = TT, TE, EE, TB, EB$, and then combining them together to form their minimum variance estimator.

OH03 moreover derive efficient full-sky single-pair estimators in configuration space. They are identical to the HO02 estimators, except for the TE estimator, which is slightly sub-optimal. We give explicit expressions for these estimators in the flat-sky limit in Section III B. Here again, the final OH03 estimator is obtained by combining these single-pair estimators. Note that HO02 and OH03 used unlensed spectra in the response functions rather than the lensed gradient spectra. We still refer to the estimators constructed with the lensed gradient spectra as the HO02 and OH03 estimators, given that the procedure is identical.

A. Single-pair minimum-variance quadratic estimators in harmonic space

We start by constructing quadratic estimators out of a single pair XY , of the form

$$\hat{\phi}_{XY}(\mathbf{L}) \equiv \int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} X(\mathbf{l}_1) Y(\mathbf{l}_2) F_{XY}(\mathbf{l}_1, \mathbf{l}_2). \quad (7)$$

For the estimator to be unbiased, the weights F_{XY} must satisfy the constraint

$$\int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} f_{XY}(\mathbf{l}_1, \mathbf{l}_2) F_{XY}(\mathbf{l}_1, \mathbf{l}_2) = 1. \quad (8)$$

The noise of this estimator (defined as in Eq. (3)) is

$$N_{XY}(L) = \int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} F_{XY}(\mathbf{l}_1, \mathbf{l}_2) \left(F_{XY}(\mathbf{l}_1, \mathbf{l}_2) C_{l_1}^{XX} C_{l_2}^{YY} + F_{XY}(\mathbf{l}_2, \mathbf{l}_1) C_{l_1}^{XY} C_{l_2}^{XY} \right). \quad (9)$$

1. All pairs except TE

For all pairs except TE , either $X = Y$ or $C_l^{XY} = 0$. As a consequence, the variance of the estimator takes the form

$$N_{XY}(L) = (1 + \delta_{XY}) \int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} C_{l_1}^{XX} C_{l_2}^{YY} [F_{XY}(\mathbf{l}_1, \mathbf{l}_2)]^2 \quad (10)$$

Minimizing this variance under the constraint (8) results in the following coefficients

$$F_{XY}(\mathbf{l}_1, \mathbf{l}_2) = \lambda_{XY}(L) \frac{f_{XY}(\mathbf{l}_1, \mathbf{l}_2)}{(1 + \delta_{XY}) C_{l_1}^{XX} C_{l_2}^{YY}}, \quad (11)$$

$$\lambda_{XY}(L) \equiv \left[\int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \frac{[f_{XY}(\mathbf{l}_1, \mathbf{l}_2)]^2}{(1 + \delta_{XY}) C_{l_1}^{XX} C_{l_2}^{YY}} \right]^{-1}. \quad (12)$$

Inserting back into Eq. (10), we find the corresponding minimum variance $N_{XY}(L) = \lambda_{XY}(L)$.

2. Special case of $XY = TE$

We may decompose F_{TE} into a symmetric and anti-symmetric piece:

$$F_{TE}(\mathbf{l}_1, \mathbf{l}_2) = F_{TE}^+(\mathbf{l}_1, \mathbf{l}_2) + F_{TE}^-(\mathbf{l}_1, \mathbf{l}_2), \quad (13)$$

$$F_{TE}^\pm(\mathbf{l}_1, \mathbf{l}_2) \equiv \frac{1}{2} (F_{TE}(\mathbf{l}_1, \mathbf{l}_2) \pm F_{TE}(\mathbf{l}_2, \mathbf{l}_1)). \quad (14)$$

For each pair $(\mathbf{l}_1, \mathbf{l}_2)$, we further define the 2-dimensional vector

$$\mathbf{F}(\mathbf{l}_1, \mathbf{l}_2) \equiv (F_{TE}^+(\mathbf{l}_1, \mathbf{l}_2), F_{TE}^-(\mathbf{l}_1, \mathbf{l}_2)). \quad (15)$$

After some algebra, and only keeping even functions of $(\mathbf{l}_1, \mathbf{l}_2)$ in the integral, Eq. (9) can be rewritten as

$$N_{TE}(L) = \int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \mathbf{F}(\mathbf{l}_1, \mathbf{l}_2) \cdot \mathbf{M}(\mathbf{l}_1, \mathbf{l}_2) \cdot \mathbf{F}(\mathbf{l}_1, \mathbf{l}_2), \quad (16)$$

where for each pair $(\mathbf{l}_1, \mathbf{l}_2)$, the 2 by 2 matrix $\mathbf{M}(\mathbf{l}_1, \mathbf{l}_2)$ is given by

$$\mathbf{M}(\mathbf{l}_1, \mathbf{l}_2) = \begin{pmatrix} C_{(l_1) (l_2)}^{TT} C_{l_2}^{EE} + C_{l_1}^{TE} C_{l_2}^{TE} & C_{[l_1] [l_2]}^{TT} C_{l_2}^{EE} \\ C_{[l_1] [l_2]}^{TT} C_{l_2}^{EE} & C_{(l_1) (l_2)}^{TT} C_{l_2}^{EE} - C_{l_1}^{TE} C_{l_2}^{TE} \end{pmatrix}, \quad (17)$$

where $A_{(l_1 l_2)} \equiv (A_{l_1 l_2} + A_{l_2 l_1})/2$ and $A_{[l_1 l_2]} \equiv (A_{l_1 l_2} - A_{l_2 l_1})/2$ are the symmetric and antisymmetric parts of $A_{l_1 l_2}$.

Similarly, we may define the symmetric and antisymmetric parts of the correlation coefficients $f_{TE}^{\pm}(\mathbf{l}_1, \mathbf{l}_2)$ and the two-dimensional vector $\mathbf{f} = (f_{TE}^+, f_{TE}^-)$, for each pair $(\mathbf{l}_1, \mathbf{l}_2)$, and rewrite the constraint (8) as

$$\int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \mathbf{F}(\mathbf{l}_1, \mathbf{l}_2) \cdot \mathbf{f}(\mathbf{l}_1, \mathbf{l}_2) = 1. \quad (18)$$

Minimizing the variance (16) under this constraint leads to the solution

$$\mathbf{F}(\mathbf{l}_1, \mathbf{l}_2) = \lambda(L) \mathbf{M}^{-1}(\mathbf{l}_1, \mathbf{l}_2) \cdot \mathbf{f}(\mathbf{l}_1, \mathbf{l}_2), \quad (19)$$

where the Lagrange multiplier λ is obtained from the constraint (18):

$$\lambda(L) = \left(\int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \mathbf{f}(\mathbf{l}_1, \mathbf{l}_2) \cdot \mathbf{M}^{-1}(\mathbf{l}_1, \mathbf{l}_2) \cdot \mathbf{f}(\mathbf{l}_1, \mathbf{l}_2) \right)^{-1}. \quad (20)$$

The 2 by 2 matrix $\mathbf{M}(\mathbf{l}_1, \mathbf{l}_2)$ is easily invertible, and after re-expressing Eq. (19) in terms of the original $F_{XY}(\mathbf{l}_1, \mathbf{l}_2)$ and $f_{TE}(\mathbf{l}_1, \mathbf{l}_2)$, one recovers the HO02 optimal weights for TE , namely, with our notation,

$$F_{TE}(\mathbf{l}_1, \mathbf{l}_2) = \lambda_{TE}(L) \frac{C_{l_1}^{EE} C_{l_2}^{TT} f_{TE}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{TE} C_{l_2}^{TE} f_{TE}(\mathbf{l}_2, \mathbf{l}_1)}{C_{l_1}^{TT} C_{l_2}^{EE} C_{l_1}^{EE} C_{l_2}^{TT} - (C_{l_1}^{TE} C_{l_2}^{TE})^2}, \quad (21)$$

$$\lambda_{TE}(L) \equiv \left[\int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} f_{TE}(\mathbf{l}_1, \mathbf{l}_2) \frac{C_{l_1}^{EE} C_{l_2}^{TT} f_{TE}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{TE} C_{l_2}^{TE} f_{TE}(\mathbf{l}_2, \mathbf{l}_1)}{C_{l_1}^{TT} C_{l_2}^{EE} C_{l_1}^{EE} C_{l_2}^{TT} - (C_{l_1}^{TE} C_{l_2}^{TE})^2} \right]^{-1}. \quad (22)$$

Inserting Eq. (19) into Eq. (16), we see that the noise of the minimum-variance estimator is just $N_{TE}(L) = \lambda_{TE}(L)$.

B. Single-pair efficient configuration-space estimators

1. All pairs except TE

The response coefficients $f_{XY}(\mathbf{l}_1, \mathbf{l}_2)$ can all be written as linear combinations of products of functions of \mathbf{l}_1 with functions of \mathbf{l}_2 , with coefficients depending on L . From Eq. (11), we see that this property transfers to the optimal weights F_{XY} for all pairs except TE . As a consequence, all single-pair estimators except TE can be written as sums of convolutions of functions of \mathbf{l}_1 with functions of \mathbf{l}_2 . This implies that they can be written as a sum of products of functions of configuration space – they are “separable” in configuration space. This allows to use Fast Fourier Transforms (FFTs) (or fast harmonic transforms for full-sky expressions [33]) to compute them efficiently.

Similar to OH03, we define the following bilinear operator of harmonic-space functions:

$$\mathcal{P}[A(\mathbf{l}_1), B(\mathbf{l}_2)](\hat{\mathbf{n}}) \equiv \nabla \mathcal{F}^{-1}[A(\mathbf{l}_1)](\hat{\mathbf{n}}) \times \mathcal{F}^{-1}[B(\mathbf{l}_2)](\hat{\mathbf{n}}). \quad (23)$$

The single-pair estimators can all be written in the form

$$\hat{\phi}_{XY}(\hat{\mathbf{n}}) = -\nabla \cdot \mathcal{F}^{-1}[\lambda_{XY}(L) \mathcal{F}[\psi_{XY}(\hat{\mathbf{n}})]], \quad (24)$$

with

$$\psi_{TT} = \mathcal{P} \left[\frac{\tilde{C}_{l_1}^{T\nabla T}}{C_{l_1}^{TT}} T(\mathbf{l}_1), \frac{T(\mathbf{l}_2)}{C_{l_2}^{TT}} \right], \quad (25)$$

$$\psi_{EE} = \frac{1}{2} \sum_{\epsilon=\pm 1} \mathcal{P} \left[e^{2i\epsilon\varphi_{l_1}} \frac{\tilde{C}_{l_1}^{E\nabla E}}{C_{l_1}^{EE}} E(\mathbf{l}_1), e^{-2i\epsilon\varphi_{l_2}} \frac{E(\mathbf{l}_2)}{C_{l_2}^{EE}} \right], \quad (26)$$

$$\psi_{TB} = \frac{1}{2i} \sum_{\epsilon=\pm 1} \epsilon \mathcal{P} \left[e^{2i\epsilon\varphi_{l_1}} \frac{\tilde{C}_{l_1}^{T\nabla E}}{C_{l_1}^{TT}} T(\mathbf{l}_1), e^{-2i\epsilon\varphi_{l_2}} \frac{B(\mathbf{l}_2)}{C_{l_2}^{BB}} \right] \quad (27)$$

$$\psi_{EB} = \frac{1}{2i} \sum_{\epsilon=\pm 1} \epsilon \mathcal{P} \left[e^{2i\epsilon\varphi_{l_1}} \frac{\tilde{C}_{l_1}^{E\nabla E}}{C_{l_1}^{EE}} E(\mathbf{l}_1), e^{-2i\epsilon\varphi_{l_2}} \frac{B(\mathbf{l}_2)}{C_{l_2}^{BB}} \right] \quad (28)$$

These expressions are the flat-sky limit of the OH03 full-sky expressions.

2. Case of $XY = TE$

The separability property is not satisfied by the TE estimator, due to the non-factorizable term in the denominator of F_{TE} in Eq. (21). Instead of the optimal F_{TE} , one can use a slightly suboptimal coefficient, obtained by setting $C_l^{TE} = 0$ in Eq. (21), namely

$$F_{TE}^{\text{eff}}(\mathbf{l}_1, \mathbf{l}_2) = \lambda_{TE}^{\text{eff}}(L) \frac{f_{TE}(\mathbf{l}_1, \mathbf{l}_2)}{C_{l_1}^{TT} C_{l_2}^{EE}}, \quad (29)$$

$$\lambda_{TE}^{\text{eff}}(L) \equiv \left[\int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \frac{[f_{TE}(\mathbf{l}_1, \mathbf{l}_2)]^2}{C_{l_1}^{TT} C_{l_2}^{EE}} \right]^{-1}. \quad (30)$$

The resulting suboptimal estimator $\hat{\phi}_{TE}^{\text{eff}}$ also takes the form Eq. (24), with

$$\begin{aligned} \psi_{TE}^{\text{eff}} = & \frac{1}{2} \sum_{\epsilon=\pm 1} \mathcal{P} \left[e^{2i\epsilon\varphi_{l_1}} \frac{\tilde{C}_{l_1}^{T\nabla E}}{C_{l_1}^{TT}} T(\mathbf{l}_1), e^{-2i\epsilon\varphi_{l_2}} \frac{E(\mathbf{l}_2)}{C_{l_2}^{EE}} \right] \\ & + \mathcal{P} \left[\frac{\tilde{C}_{l_2}^{T\nabla E}}{C_{l_2}^{EE}} E(\mathbf{l}_2), \frac{T(\mathbf{l}_1)}{C_{l_1}^{TT}} \right]. \end{aligned} \quad (31)$$

C. Optimal combination of single-pair estimators

Given the five single-pair estimators $\hat{\phi}_\alpha(\mathbf{L})$ constructed for each $\alpha \in \{TT, TE, EE, TB, EB\}$, HO02 combine them to form the estimator

$$\hat{\phi}_{\text{HO02}}(\mathbf{L}) = \sum_{\alpha} w_{\alpha}(\mathbf{L}) \hat{\phi}_{\alpha}(\mathbf{L}), \quad (32)$$

where the optimal weights $w_{\alpha}(\mathbf{L})$ are obtained by minimizing the variance of the linear combination with the constraint that they sum up to unity i.e. $\sum_{\alpha} w_{\alpha} = 1$. Subject to this constraint, one gets

$$w_{\alpha} = \frac{\sum_{\beta} (\mathbf{N}^{-1})_{\alpha\beta}}{\sum_{\beta\gamma} (\mathbf{N}^{-1})_{\beta\gamma}}, \quad (33)$$

where for each L , $N_{\alpha\beta}(L)$ is the covariance matrix of the separate estimators $\hat{\phi}_{\alpha}$, whose elements are obtained by the generalization of Eq. (3) to the cross-correlation of two estimators. The overall noise of this estimator is then $N_{\text{HO}} \equiv \left(\sum_{\alpha\beta} (\mathbf{N}^{-1})_{\alpha\beta} \right)^{-1}$.

The final HO02 estimator thus takes the form

$$\hat{\phi}_{\text{HO02}}(\mathbf{L}) = \int_{\mathbf{l}_1+\mathbf{l}_2=\mathbf{L}} \sum_{XY} F_{XY}^{\text{HO02}}(\mathbf{l}_1, \mathbf{l}_2) X(\mathbf{l}_1) Y(\mathbf{l}_2), \quad (34)$$

where the sum runs over the five unique pairs $XY = TT, TE, EE, TB, EB$, and the weights are proportional to the single-pair optimal weights, each with a different proportionality coefficient:

$$F_{XY}^{\text{HO02}}(\mathbf{l}_1, \mathbf{l}_2) = w_{XY}(\mathbf{L}) F_{XY}(\mathbf{l}_1, \mathbf{l}_2). \quad (35)$$

The same procedure can be carried with the single-pair separable estimators of OH03. These estimators are all identical to the minimum-variance estimators of HO02, except for $\hat{\phi}_{TE}^{\text{eff}}$, which is slightly suboptimal relative to $\hat{\phi}_{TE}$. Upon combining all five estimators, the OH03 estimator also takes the form of Eq. (34), with weights

$$F_{XY}^{\text{OH03}}(\mathbf{l}_1, \mathbf{l}_2) = w_{XY}^{\text{eff}}(\mathbf{L}) \lambda_{XY}^{\text{eff}}(\mathbf{L}) \frac{f_{XY}(\mathbf{l}_1, \mathbf{l}_2)}{(1 + \delta_{XY}) C_{l_1}^{XX} C_{l_2}^{YY}}. \quad (36)$$

IV. GLOBAL MINIMUM-VARIANCE QUADRATIC ESTIMATOR

It is easy to see that the final HO02 estimator is a linear combination of terms quadratic in T, E, B . Rather than splitting the optimization process in two steps, we instead directly seek the global minimum variance quadratic estimator, in one single step. By doing so, we can account for the correlations between different XY pairs for each $(\mathbf{l}_1, \mathbf{l}_2)$, rather than only after integrating over $(\mathbf{l}_1, \mathbf{l}_2)$, as done in the HO02 estimator. The estimator built this way is therefore necessarily less noisy than the HO02 estimator, as we will show explicitly.

A. Harmonic-space expression

We start by deriving the global-minimum-variance (hereafter GMV) estimator in harmonic space, following the steps of Appendix A of Hirata & Seljak [25].

For each Fourier mode \mathbf{l} , we define the three-dimensional vector $\mathbf{X}(\mathbf{l}) = [T(\mathbf{l}), E(\mathbf{l}), B(\mathbf{l})]$. We seek an estimator of the form

$$\hat{\phi}(\mathbf{L}) = \int_{\mathbf{l}_1+\mathbf{l}_2=\mathbf{L}} X^i(\mathbf{l}_1) \Xi_{ij}(\mathbf{l}_1, \mathbf{l}_2) X^j(\mathbf{l}_2), \quad (37)$$

where we use the Einstein summation convention. Without loss of generality, we may assume $\Xi_{ji}(\mathbf{l}_2, \mathbf{l}_1) = \Xi_{ij}(\mathbf{l}_1, \mathbf{l}_2)$, as only the part of the integrand symmetric under exchange of $(\mathbf{l}_1, \mathbf{l}_2)$ contributes to the integral.

For $\mathbf{l}_1 + \mathbf{l}_2 \neq \mathbf{0}$, we define $f_{ij}(\mathbf{l}_1, \mathbf{l}_2)$ through

$$\left\langle \frac{\delta}{\delta\phi(\mathbf{L})} (X^i(\mathbf{l}_1) X^j(\mathbf{l}_2)) \right\rangle = \delta(\mathbf{l}_1 + \mathbf{l}_2 - \mathbf{L}) f_{ij}(\mathbf{l}_1, \mathbf{l}_2). \quad (38)$$

In other words, if $i = 1, 2, 3$ correspond to $X^i = T, E, B$, we have $f_{11}(\mathbf{l}_1, \mathbf{l}_2) = f_{TT}(\mathbf{l}_1, \mathbf{l}_2)$, $f_{12}(\mathbf{l}_1, \mathbf{l}_2) = f_{TE}(\mathbf{l}_1, \mathbf{l}_2) = f_{21}(\mathbf{l}_2, \mathbf{l}_1)$, etc... Here again, we have $f_{ji}(\mathbf{l}_2, \mathbf{l}_1) = f_{ij}(\mathbf{l}_1, \mathbf{l}_2)$.

Requiring the estimator to be unbiased thus leads the constraint equation

$$\int_{\mathbf{l}_1+\mathbf{l}_2=\mathbf{L}} \Xi_{ij}(\mathbf{l}_1, \mathbf{l}_2) f_{ij}(\mathbf{l}_1, \mathbf{l}_2) = 1. \quad (39)$$

Using the symmetry properties of Ξ , the variance of this estimator is then

$$N(\mathbf{L}) = 2 \int_{\mathbf{l}_1+\mathbf{l}_2=\mathbf{L}} \Xi_{ij}(\mathbf{l}_1, \mathbf{l}_2) \Xi_{pq}(\mathbf{l}_1, \mathbf{l}_2) C_{l_1}^{ip} C_{l_2}^{jq}. \quad (40)$$

Minimizing this variance under the constraint (39) leads to

$$C_{l_1}^{ip} \Xi_{pq}(\mathbf{l}_1, \mathbf{l}_2) C_{l_2}^{jq} = \frac{\lambda(\mathbf{L})}{2} f_{ij}(\mathbf{l}_1, \mathbf{l}_2), \quad (41)$$

where $\lambda(\mathbf{L})$ is a Lagrange multiplier. This equation is more easily solved in matrix form. For each l , we define

the three by three symmetric matrix $[C_l]$ with elements C_l^{ij} ; similarly, for each pair $(\mathbf{l}_1, \mathbf{l}_2)$, we define the three by three matrices $[\Xi(\mathbf{l}_1, \mathbf{l}_2)]$ and $[\mathbf{f}(\mathbf{l}_1, \mathbf{l}_2)]$. Equation (41) then has the solution

$$[\Xi(\mathbf{l}_1, \mathbf{l}_2)] = \frac{\lambda(L)}{2} [C_{l_1}]^{-1} [\mathbf{f}(\mathbf{l}_1, \mathbf{l}_2)] [C_{l_2}]^{-1}. \quad (42)$$

Inserting back into the constraint equation, we obtain

$$\lambda(L)^{-1} = \int_{\substack{\mathbf{l}_1 + \mathbf{l}_2 \\ = \mathbf{L}}} \frac{1}{2} \text{Tr} \left([C_{l_1}]^{-1} [\mathbf{f}(\mathbf{l}_1, \mathbf{l}_2)] [C_{l_2}]^{-1} [\mathbf{f}(\mathbf{l}_2, \mathbf{l}_1)] \right). \quad (43)$$

The noise of the minimum-variance estimator is then simply $N(L) = N_{\text{GMV}}(L) = \lambda(L)$.

Putting everything together, the GMV estimator takes the final form

$$\hat{\phi}_{\text{GMV}}(\mathbf{L}) = \int_{\substack{\mathbf{l}_1 + \mathbf{l}_2 \\ = \mathbf{L}}} \sum_{XY} F_{XY}^{\text{GMV}}(\mathbf{l}_1, \mathbf{l}_2) X(\mathbf{l}_1) Y(\mathbf{l}_2), \quad (44)$$

where the sum runs over the five unique pairs $XY = TT, TE, EE, TB, EB$. Explicitly, the weights are

$$F_{TT}^{\text{GMV}}(\mathbf{l}_1, \mathbf{l}_2) = \frac{N_{\text{GMV}}(L)}{2D_{l_1}D_{l_2}} \times \left[C_{l_1}^{EE} C_{l_2}^{EE} f_{TT}(\mathbf{l}_1, \mathbf{l}_2) + C_{l_1}^{TE} C_{l_2}^{TE} f_{EE}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{EE} C_{l_2}^{TE} f_{TE}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_2}^{EE} C_{l_1}^{TE} f_{TE}(\mathbf{l}_2, \mathbf{l}_1) \right], \quad (45)$$

$$F_{EE}^{\text{GMV}}(\mathbf{l}_1, \mathbf{l}_2) = \frac{N_{\text{GMV}}(L)}{2D_{l_1}D_{l_2}} \times \left[C_{l_1}^{TE} C_{l_2}^{TE} f_{TT}(\mathbf{l}_1, \mathbf{l}_2) + C_{l_1}^{TT} C_{l_2}^{TT} f_{EE}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{TE} C_{l_2}^{TT} f_{TE}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_2}^{TE} C_{l_1}^{TT} f_{TE}(\mathbf{l}_2, \mathbf{l}_1) \right], \quad (46)$$

$$F_{TE}^{\text{GMV}}(\mathbf{l}_1, \mathbf{l}_2) = \frac{N_{\text{GMV}}(L)}{D_{l_1}D_{l_2}} \times \left[-C_{l_1}^{TE} C_{l_2}^{EE} f_{TT}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{TE} C_{l_2}^{TT} f_{EE}(\mathbf{l}_1, \mathbf{l}_2) + C_{l_1}^{EE} C_{l_2}^{TT} f_{TE}(\mathbf{l}_1, \mathbf{l}_2) + C_{l_1}^{TE} C_{l_2}^{TE} f_{TE}(\mathbf{l}_2, \mathbf{l}_1) \right], \quad (47)$$

$$F_{TB}^{\text{GMV}}(\mathbf{l}_1, \mathbf{l}_2) = \frac{N_{\text{GMV}}(L)}{D_{l_1}C_{l_2}^{BB}} \times \left[C_{l_1}^{EE} f_{TB}(\mathbf{l}_1, \mathbf{l}_2) - C_{l_1}^{TE} f_{EB}(\mathbf{l}_1, \mathbf{l}_2) \right], \quad (48)$$

$$F_{EB}^{\text{GMV}}(\mathbf{l}_1, \mathbf{l}_2) = \frac{N_{\text{GMV}}(L)}{D_{l_1}C_{l_2}^{BB}} \times \left[-C_{l_1}^{TE} f_{TB}(\mathbf{l}_1, \mathbf{l}_2) + C_{l_1}^{TT} f_{EB}(\mathbf{l}_1, \mathbf{l}_2) \right], \quad (49)$$

where $D_l \equiv C_l^{TT} C_l^{EE} - [C_l^{TE}]^2$.

These explicit expressions should make it very clear that the GMV estimator is different from the HO02 estimator. Put differently, the first step of the likelihood-based iterative technique of Ref. [25] is *not* equivalent to the HO02 estimator. Indeed, in the HO02 estimator, the weight F_{XY} of each pair XY is proportional to f_{XY} only

(times a function of L), even after combining all single-pair estimators; in contrast, for the GMV estimator, the weight of each pair is a linear combination of the response coefficients from *all* pairs. The weights in the GMV estimator would not separately minimize the variance of an individual XY pair, but they provide a global optimum when combining all the pairs together.

These expressions also show that the weights are all sums of products of function of \mathbf{l}_1 with functions of \mathbf{l}_2 , including F_{TE}^{GMV} . In other words, the GMV estimator is separable without requiring any additional approximation, making it well adapted for efficient computations, as we now discuss.

As a side note, let us point out that the GMV estimator (just like the HO02 and OH03 estimators) can be split into two pieces, built from $\{TT, TE, EE\}$ and $\{TB, EB\}$, respectively, which are uncorrelated as $C_l^{TB} = C_l^{EB} = 0$. Our publicly available Python code `GlobalLensQuest` first computes these two separate uncorrelated estimators $\hat{\phi}_a \equiv \hat{\phi}_{\{TT, TE, EE\}}$ and $\hat{\phi}_b \equiv \hat{\phi}_{\{TB, EB\}}$, and then combines them with inverse variance weighting to obtain the GMV estimator.

B. Compact configuration-space expression

We may write the GMV estimator in even more compact form by defining the inverse-covariance-weighted fields

$$\bar{\mathbf{X}}(\mathbf{l}) \equiv [C_l]^{-1} \mathbf{X}(\mathbf{l}), \quad (50)$$

and write

$$\bar{T}(\mathbf{l}_1) \equiv \bar{\mathbf{X}}^1(\mathbf{l}_1) = [C_{l_1}^{EE} T(\mathbf{l}_1) - C_{l_1}^{TE} E(\mathbf{l}_1)] / D_{l_1}, \quad (51)$$

and similarly $\bar{E}(\mathbf{l}_1) = \bar{\mathbf{X}}^2$, $\bar{B}(\mathbf{l}_1) = \bar{\mathbf{X}}^3$. We then have

$$\hat{\phi}_{\text{GMV}}(L) = \frac{\lambda(L)}{2} \int_{\substack{\mathbf{l}_1 + \mathbf{l}_2 \\ = \mathbf{L}}} \bar{\mathbf{X}}^i(\mathbf{l}) f_{ij}(\mathbf{l}_1, \mathbf{l}_2) \bar{\mathbf{X}}^j(\mathbf{l}_2). \quad (52)$$

We moreover define the Wiener-filtered fields

$$X_{\text{WF}}^i(\mathbf{l}) \equiv \tilde{C}_l^{ij} \bar{\mathbf{X}}^j, \quad (53)$$

where $\tilde{C}_l^{11} \equiv \tilde{C}_l^{T\nabla T}$, $\tilde{C}_l^{12} = \tilde{C}_l^{21} \equiv \tilde{C}_l^{T\nabla E}$, etc., and write

$$T_{\text{WF}}(\mathbf{l}) \equiv X_{\text{WF}}^1(\mathbf{l}) \equiv \tilde{C}_l^{T\nabla T} \bar{T}(\mathbf{l}) + \tilde{C}_l^{T\nabla E} \bar{E}(\mathbf{l}), \quad (54)$$

and similarly for $E_{\text{WF}}(\mathbf{l}) \equiv X_{\text{WF}}^2(\mathbf{l})$.

In terms of these fields, the GMV estimator takes the particularly simple form

$$\hat{\phi}_{\text{GMV}}(\mathbf{L}) = \frac{\lambda(L)}{2} \int_{\substack{\mathbf{l}_1 + \mathbf{l}_2 \\ = \mathbf{L}}} (\mathbf{L} \cdot \mathbf{l}_1) \left[2T_{\text{WF}}(\mathbf{l}_1) \bar{T}(\mathbf{l}_2) + \sum_{\epsilon=\pm 1} E_{\text{WF}}(\mathbf{l}_1) e^{-2i\epsilon\varphi_{l_1}} (\bar{E}(\mathbf{l}_2) + i\epsilon\bar{B}(\mathbf{l}_2)) e^{2i\epsilon\varphi_{l_2}} \right]. \quad (55)$$

This form is well adapted for efficient evaluation, as it is the sum of convolutions of functions of \mathbf{l}_1 with functions of \mathbf{l}_2 . To see this, let us define

$$\pm_2 E_{\text{WF}}(\mathbf{l}) \equiv E_{\text{WF}}(\mathbf{l}) e^{\pm 2i\varphi_{\mathbf{l}}}, \quad (56)$$

$$\pm_2 \bar{P}(\mathbf{l}) \equiv \frac{1}{2} (\bar{E}(\mathbf{l}) \pm i\bar{B}(\mathbf{l})) e^{\pm 2i\varphi_{\mathbf{l}}}. \quad (57)$$

We may then express the GMV estimator in terms of the configuration-space versions of these fields (i.e. their inverse-Fourier transforms):

$$\hat{\phi}_{\text{GMV}}(\hat{\mathbf{n}}) = -\nabla \cdot \mathcal{F}^{-1} [\lambda(L) \mathcal{F}[\psi_{\text{GMV}}(\hat{\mathbf{n}})]], \quad (58)$$

where

$$\begin{aligned} \psi_{\text{GMV}}(\hat{\mathbf{n}}) &= \nabla T_{\text{WF}}(\hat{\mathbf{n}}) \bar{T}(\hat{\mathbf{n}}) \\ &+ \sum_{s=\pm 2} \nabla [-_s E_{\text{WF}}(\hat{\mathbf{n}})] {}_s \bar{P}(\hat{\mathbf{n}}). \end{aligned} \quad (59)$$

This expression is the flat-sky equivalent of Eq. (3) in Ref. [10], derived in Ref. [37]. A similar expression is derived in Ref. [38], in terms of (T, Q, U) rather than (T, E, B) ; nevertheless, it is also incorrectly stated in that paper that this estimator is identical to an estimator built out of single-pair estimators, i.e. the HO02 estimator.

V. SUBOPTIMAL QUADRATIC ESTIMATOR (SQE) USED IN RECENT DATA ANALYSES

While the full expression for the configuration-space GMV estimator was already known (although it was not known that it differs from the HO02 estimator) [10, 37], in practice only an approximate version was used for data analyses thus far. Instead of using the full covariance matrix $[C_l]$ in Eq. (50), the Planck collaboration [9, 10] and SPT collaboration [13] approximate it as diagonal by setting $C_l^{TE} = 0$ – note that [10] still use the exact response coefficients $f_{XY}(\mathbf{l}_1, \mathbf{l}_2)$. This simplification allows to deal with a cut-sky setup with a lower computational cost; it moreover preserves the configuration space separability. We denote this suboptimal quadratic estimator SQE. Explicitly, the weights of this estimator are

$$F_{XY}^{\text{SQE}}(\mathbf{l}_1, \mathbf{l}_2) = \lambda_{\text{SQE}}(L) \frac{f_{XY}(\mathbf{l}_1, \mathbf{l}_2)}{(1 + \delta_{XY}) C_{l_1}^{XX} C_{l_2}^{YY}}, \quad (60)$$

$$\lambda_{\text{SQE}}(L) \equiv \left(\int_{\mathbf{l}_1+\mathbf{l}_2=\mathbf{L}} \sum_{XY} \frac{f_{XY}(\mathbf{l}_1, \mathbf{l}_2)^2}{(1 + \delta_{XY}) C_{l_1}^{XX} C_{l_2}^{YY}} \right)^{-1} \quad (61)$$

where again the sum runs of the five distinct pairs XY .

By definition, this estimator is suboptimal relative to the GMV estimator. Furthermore, it should be clear from Eq. (36) that it is also noisier than the OH03 estimator (and as we will see, also noisier than the HO02 estimator). Indeed, the OH03 estimator accounts for the covariances between different single-pair estimators, which depend on C_l^{TE} ; as a consequence, the L -dependent proportionality constant in Eq. (36) is different for each pair.

In contrast, the SQE amounts to neglecting correlations between single-pair estimators, and simply using their inverse-variance combination, leading to the same coefficient $\lambda_{\text{SQE}}(L)$ for all weights in Eq. (60). Given that the SQE estimator is effectively a linear combination of single-pair estimator, and that the OH03 weights represent the optimal linear combination of single-pair estimators, we conclude that the SQE estimator must be suboptimal to the OH03 estimator.

To compute the noise of this estimator, we must account for $C_l^{TE} \neq 0$. We find

$$\begin{aligned} N_{\text{SQE}}(L) &= \lambda_{\text{SQE}}^2(L) \int_{\mathbf{l}_1+\mathbf{l}_2=\mathbf{L}} \sum_{XY} \frac{f_{XY}(\mathbf{l}_1, \mathbf{l}_2)}{(1 + \delta_{XY}) C_{l_1}^{XX} C_{l_2}^{YY}} \\ &\sum_{UV} \left[\frac{f_{UV}(\mathbf{l}_1, \mathbf{l}_2) C_{l_1}^{XU} C_{l_2}^{YV}}{(1 + \delta_{UV}) C_{l_1}^{UU} C_{l_2}^{VV}} + \frac{f_{UV}(\mathbf{l}_2, \mathbf{l}_1) C_{l_1}^{XV} C_{l_2}^{YU}}{(1 + \delta_{UV}) C_{l_2}^{UU} C_{l_1}^{VV}} \right] \end{aligned} \quad (62)$$

The SQE estimator enables faster evaluation with a cut-sky, at the cost of only $\sim 3\%$ increase in the reconstruction noise for *Planck* [10], as we confirm in Fig. 2. However, we will see that for more sensitive experimental setups, the reconstruction noise penalty can be more than 10% and thus a full joint filtering analysis of the temperature and polarization maps would be beneficial in future.

VI. QUANTITATIVE COMPARISON OF DIFFERENT QUADRATIC ESTIMATORS

A. Experimental setups

In order to evaluate the variance of different quadratic estimators, we use three different setups which correspond to the *Planck*, SO-like, and CMBS4-like experiments. In Table II, we provide the adopted specifications for these setups. The Gaussian random noise of the detector is calculated as

$$C_\ell^{T,E/B}|_{\text{noise}} = (\Delta_{T,P})^2 e^{\ell(\ell+1)\sigma^2/8 \ln 2} \quad (63)$$

where $\Delta_{T,P}$ denote the white-noise of the detector in μK -radian, and σ is the full width at half maximum (FWHM) of the beam in arcmin.

Experiment	ℓ_{max}	Δ_T $\mu\text{K-arcmin}$	Δ_P $\mu\text{K-arcmin}$	σ arcmin
<i>Planck</i>	3000	35.0	60.0	5.0
SO	3000	8.0	$8.0\sqrt{2}$	1.4
CMBS4	3000	1.0	$1.0\sqrt{2}$	1.0

TABLE II: Experimental specifications used in this work.

It has been shown that extra-galactic foregrounds can bias the CMB lensing reconstruction from temperature

maps and different strategies have been proposed to mitigate this issue [e.g. 34, 35, 39–41]. We neglect the foregrounds in this study, and simply choose $\ell_{\max} = 3000$ in both temperature and polarization. Although it is possible to go for a much higher ℓ_{\max} in polarization than in temperature due to lack of strongly polarized foregrounds, for simplicity we take $\ell_{\max}^T = \ell_{\max}^P$. We have checked our results with different ℓ_{\max} ranges and found no drastic difference in our results.

B. Comparison of reconstruction noises

We start by comparing the OH03 and HO02 quadratic estimators in Figure 1. We find that the HO02 estimator is systematically less noisy than the OH03 estimator for large angular scales (by less than 0.5%). Interestingly, the OH03 estimator becomes slightly less noisy than the HO02 estimator for L larger than several hundreds. We have checked that our numerical integrals are converged to better than 0.01% relative accuracy up to $L \approx 2000$, and to better than 0.03% for $L \lesssim 3000$ which we also show in Fig. 4. This gives us confidence that the $\sim 0.08\%$ improvement of OH03 estimator over HO02 seen for a CMBS4-like experiment is real and not a numerical artifact.

The lower noise of OH03 at small angular scales may appear surprising at first, given that this estimator uses a TE estimator suboptimal to that of HO02. However, the suboptimal TE estimator does not guarantee that the overall OH03 estimator (obtained by optimally combining the five single-pair estimators) is noisier than the overall HO02 estimator: indeed, if the OH03 TE estimator happens to be more correlated with the TT and EE estimators than the HO02 TE estimator which is the case here, the overall combination of OH03 estimators can be less noisy than that of HO02.

In Fig. 2, we show the ratios of the reconstruction noise of the GMV and HO02 estimators to that of the SQE estimator for different experimental setups. As expected, we find that the variance of the GMV quadratic estimator is lower than that of the HO02 and SQE estimators. Also, the variance of HO02 estimator is smaller than the SQE estimator. For a *Planck*-like experimental setup, on large angular scales ($L \lesssim 500$), the difference between the SQE and GMV estimators is of the order of 3% and can reach $\sim 6\%$ around $L \sim 2000$. However, for more sensitive experiments, this difference reaches $\sim 11\%$ and $\sim 12\%$ at $L \lesssim 100$ and $L \sim 2000$ for SO- or CMBS4-like experiments, respectively. This result may motivate using the full covariance matrix $[C_l]$ in Eq. (50) instead of assuming $C_l^{TE} = 0$, in order to obtain more precise results in future data analyses.

From Fig. 2, we can also see that for a *Planck*-like experimental setup, at small angular

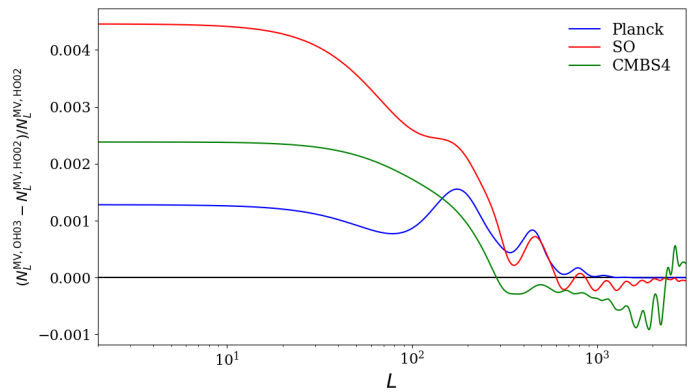


FIG. 1: Fractional difference between the variance of the OH03 and HO02 estimators. Different colors correspond to the different experimental setups described in Sec. VI A.

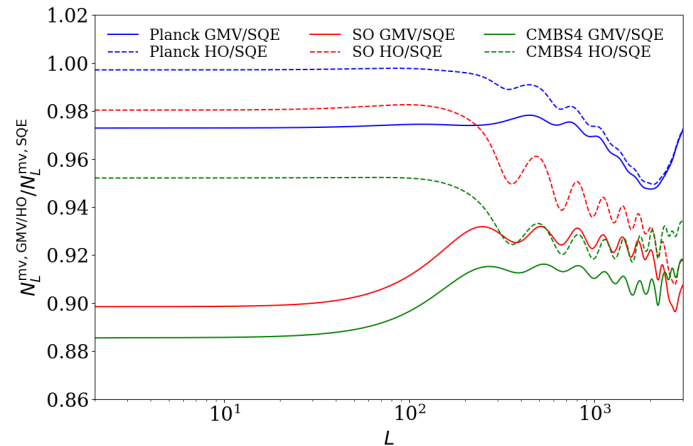


FIG. 2: Ratio of the minimum variance reconstruction noise of the GMV and SQE estimators, and HO02 and SQE estimators for different experimental setups.

scales ($L \gtrsim 1000$) the GMV and HO02 estimators perform almost equally well, while for SO- and CMBS4-like experiment GMV outperforms HO02 everywhere. On large angular scales, the difference between the GMV and HO02 is much more significant for all the experiments considered here. For *Planck*, polarization noise is significant, so the difference in lensing estimators is quite small on all scales. However, although CMBS4-like experiments are *EB*-dominated for the purpose of lensing reconstruction, the improvement of the GMV over HO02 is driven by a significant difference in the TT-TE-EE part of the minimum-variance estimator (rather than improved filtering of *E* improving *EB*; see Appendix C). The effect on the combined MV estimator is largest at $L \lesssim 200$ where the MV estimator has the largest contribution from TT-TE-EE.

VII. CONCLUSIONS

Quadratic estimators (QEs) are widely used to reconstruct the CMB lensing potential from CMB temperature and polarization maps. In fact, up until the very recent POLARBEAR [17] and SPTPol [14] results, all maps of the CMB lensing potential have been constructed using QEs. In this work, we present a clear comparison between different QEs, both in terms of explicit equations, and quantitatively, by comparing their reconstruction noise. Importantly, we show that the Hu-Okamoto [30] (HO02) optimization method, consisting in first constructing optimal single-pair quadratic estimators, and then their optimal linear combination, does not lead to the absolute minimum-variance QE. Instead, we derive the global-minimum-variance (GMV) QE, which minimizes the variance of quadratic temperature and polarization combinations in one single step.

Interestingly, the GMV estimator derived here had been hiding in plain sight in previous works. It is equivalent to the first step (or weak-signal limit) of likelihood-based methods [25, 28, 38], which is therefore *not* equivalent to the HO02 estimator, contrary to what was previously thought (although technically the GMV estimator with non-perturbative lensed gradient weights presented here is a modification to the first-step of likelihood based estimators so that the result is non-perturbatively unbiased). Our work is the first to note that the HO02 estimator is not the global-optimum quadratic estimator, and make this point sharply clear through explicit expressions, as well as numerical comparisons. Indeed, we show that the reconstruction noise of the GMV estimator is lower than that of the HO02 estimator by up to $\sim 9\%$ for a SO-like experiment.

We also study the suboptimal QE used in the 2018 Planck [9, 10, 28] and recent SPT [13] lensing analyses (SQE), which is obtained from the GMV quadratic estimator (appropriately generalized to account for beam and pixel convolution), with the additional approximation of neglecting $C_l^{TE} = 0$ in the inverse filter matrix. We show that this approximation makes the SQE suboptimal not only relative to the GMV estimator, but also relative to the HO02 and OH03 estimators. We evaluate the reconstruction noise of the different estimators for ongoing and planned CMB experimental setups and find that while the improvement in the reconstruction noise between the SQE and HO02 estimator is of order $\sim 1-8\%$, the difference between the SQE and GMV estimators is $\sim 9-12\%$ for more sensitive experiments, especially on large angular scales $L \lesssim 10^2$ and scales around $L \sim 2000$. This improvement amounts to achieving a better sensitivity for the same experiment at no additional cost. This should motivate overcoming the added complexity associated with joint filtering of cut-sky temperature and polarization maps, in order to be able to use the GMV estimator in future lensing data analyses. Our

Python code `GlobalLensQuest`¹ to compare and compute the noise variances of the HO02, OH03, GMV, and SQE estimators and Julien Carron’s codes `LensIt`² and `plancklens`³ which can perform the optimal GMV operation with anisotropic noise and cut-sky are publicly available.

While in this work we have chosen to present relevant equations in the flat-sky limit for conciseness, it is straightforward to generalize our results to the full-sky case. The flat-sky fields $X(\mathbf{l})$ are to be replaced by the full-sky harmonic coefficients $X_{\ell m}$; the generalization of Eq. (2) and of the coupling coefficients $f_{XY}(\mathbf{l}, \mathbf{l}')$ are then provided in Ref. [33]. While [33] provide the full-sky expressions for the HO02 estimator, [37] provide the full-sky expressions for the GMV estimator incorporating the instrumental beam response and anisotropic noise. We expect a comparable improvement over the full-sky version of the HO02 estimator [33] when using the GMV estimator [37].

The approach presented here for the GMV estimator would also apply to any other joint estimator constructed from similar linear combinations of other estimators. For example, the foreground-immune hybrid QE of Ref. [35] splits the TT lensing estimator into magnification-only and shear-only estimators, and then form a hybrid estimator through a minimum-variance linear combination of these two estimators. Their hybrid estimator can be further optimized by following the logic presented here, i.e. searching for the global-minimum-variance shear and magnification estimator, accounting for correlations for each $(\mathbf{l}_1, \mathbf{l}_2)$, rather than after integration over $\mathbf{l}_1, \mathbf{l}_2$.

For SO- and CMBS4-like experiments, on large angular scales ($L < 100$) the reconstruction is expected to be signal dominated [20, 22]. The power spectrum $C_L^{\phi\phi}$ uncertainty is therefore dominated by cosmic variance and using the GMV estimator rather than the HO02 or SQE estimators would not drastically affect the measurement of $C_L^{\phi\phi}$ on these large angular scales. However, the reduction in the noise of the reconstructed ϕ field on the signal-dominated large angular scales will be beneficial for science goals which involve cross-correlation of the ϕ field with other tracers of large-scale structure [5, 20], utilizing the sample variance cancellation through cross-correlations. Also, lensing induced B-modes act as a source of noise and limit the measurement of the primordial B-modes [42] which is a major scientific goal for CMB experiments. These modes can be removed using map-level estimates of both the primordial E-modes and lensing potential ϕ with a technique called delensing and depend on the estimate of the particular realization of ϕ in the given patch of the sky [20, 22, 43]. Lower noise estimates of the ϕ field will therefore be crucial for such

¹ <https://github.com/abhimaniyar/GlobalLensQuest>

² <https://github.com/carronj/LensIt>

³ <https://github.com/carronj/plancklens>

operations and motivate using the GMV estimator instead of other QEs.

Even if future CMBS4 experiments will likely use likelihood-based iterative methods to reconstruct the lensing potential, QEs will remain very useful as a forecasting and cross-checking tool. More immediately, QEs are still the primary lensing reconstruction tool for current and near-future CMB experiments. The GMV estimator will therefore be a useful tool to harvest even more information out of the CMB data.

ACKNOWLEDGMENTS

We acknowledge useful discussions with Colin Hill, Emmanuel Schaan, and Simone Ferraro. YAH is supported by NSF award No 1820861 and NASA grant No 80NSSC20K0532. JC acknowledges support from a SNSF Eccellenza Professorial Fellowship (No. 186879). AL has support from the UK STFC grant ST/T000473/1.

Appendix A: Explicit form for the GMV $N^{(1)}(L)$

As pointed out in Sec. II, in the main text we only optimize relative to the reconstruction noise $N^{(0)}(L)$. Here we give an explicit expression for the additional $N^{(1)}(L)$ bias [36] for the GMV estimator, for which an explicit expression has not been provided in the literature.

The covariance of the GMV estimator given by Eq. (37) becomes

$$\langle \hat{\phi}(\mathbf{L})\hat{\phi}(\mathbf{L}') \rangle = \int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \int_{\substack{\mathbf{l}'_1+\mathbf{l}'_2 \\ =\mathbf{L}'}} \langle X^i(\mathbf{l}_1)X^j(\mathbf{l}_2)X^p(\mathbf{l}'_1)X^q(\mathbf{l}'_2) \rangle \Xi_{ij}(\mathbf{l}_1, \mathbf{l}_2) \Xi_{pq}(\mathbf{l}'_1, \mathbf{l}'_2). \quad (\text{A1})$$

Evaluating the expectation value in the brackets

$$\begin{aligned} \langle X^i(\mathbf{l}_1)X^j(\mathbf{l}_2)X^p(\mathbf{l}'_1)X^q(\mathbf{l}'_2) \rangle &= (2\pi)^4 [C_{l_1}^{ij}C_{l'_1}^{pq}\delta_D(\mathbf{L})\delta_D(\mathbf{L}') + C_{l_1}^{ip}C_{l'_1}^{jq}\delta_D(\mathbf{l}_1 + \mathbf{l}'_1)\delta_D(\mathbf{l}_2 + \mathbf{l}'_2) \\ &\quad + C_{l_1}^{iq}C_{l'_1}^{jp}\delta_D(\mathbf{l}_1 + \mathbf{l}'_2)\delta_D(\mathbf{l}_2 + \mathbf{l}'_1)] + (2\pi)^2 T^{ijpq}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}'_1, \mathbf{l}'_2)\delta_D(\mathbf{L} + \mathbf{L}'), \end{aligned} \quad (\text{A2})$$

where the first term in the square brackets disappears because $\mathbf{L} \neq 0$ and rest of the terms in the square bracket represent $N^{(0)}(L)$ and $T^{ijpq}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}'_1, \mathbf{l}'_2)$, the trispectrum containing terms that contribute to the lensing power spectrum signal and the signal-dependent $N^{(1)}(L)$ bias. Following [36], the trispectrum term can be written in terms of $f_{ij}(\mathbf{l}_1, \mathbf{l}_2)$ to first order in explicit $C_L^{\phi\phi}$ as

$$T^{ijpq}(\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}'_1, \mathbf{l}'_2) = C_{|\mathbf{l}_1+\mathbf{l}_2|}^{\phi\phi} f_{ij}(\mathbf{l}_1, \mathbf{l}_2) f_{pq}(\mathbf{l}'_1, \mathbf{l}'_2) + C_{|\mathbf{l}_1+\mathbf{l}'_1|}^{\phi\phi} f_{ip}(\mathbf{l}_1, \mathbf{l}'_1) f_{jq}(\mathbf{l}_2, \mathbf{l}'_2) + C_{|\mathbf{l}_1+\mathbf{l}'_2|}^{\phi\phi} f_{iq}(\mathbf{l}_1, \mathbf{l}'_2) f_{jp}(\mathbf{l}_2, \mathbf{l}'_1). \quad (\text{A3})$$

Substituting Eqs. (A2) and (A3) in Eq. (A1) we have

$$\langle \hat{\phi}(\mathbf{L})\hat{\phi}(\mathbf{L}') \rangle = (2\pi)^2 \delta_D(\mathbf{L} + \mathbf{L}') [C_L^{\phi\phi} + N^{(0)}(L) + N^{(1)}(L)], \quad (\text{A4})$$

where

$$N^{(1)}(L) \equiv \int_{\substack{\mathbf{l}_1+\mathbf{l}_2 \\ =\mathbf{L}}} \int_{\substack{\mathbf{l}'_1+\mathbf{l}'_2 \\ =\mathbf{L}'}} \Xi_{ij}(\mathbf{l}_1, \mathbf{l}_2) \Xi_{pq}(\mathbf{l}'_1, \mathbf{l}'_2) \times [C_{|\mathbf{l}_1+\mathbf{l}'_1|}^{\phi\phi} f_{ip}(\mathbf{l}_1, \mathbf{l}'_1) f_{jq}(\mathbf{l}_2, \mathbf{l}'_2) + C_{|\mathbf{l}_1+\mathbf{l}'_2|}^{\phi\phi} f_{iq}(\mathbf{l}_1, \mathbf{l}'_2) f_{jp}(\mathbf{l}_2, \mathbf{l}'_1)], \quad (\text{A5})$$

The optimal weight matrix $\Xi_{ij}(\mathbf{l}_1, \mathbf{l}_2)$ was determined in order to minimize the variance while only considering $N^{(0)}(L)$.

In the left panel of Fig. 3, we show the comparison between the $N^{(0)}(L)$ and $N^{(1)}(L)$ curves for the GMV estimator for our given experimental configurations. We find that for *Planck*-like experiments $N^{(1)}(L)$ is a couple of orders of magnitude smaller than $N^{(0)}(L)$, while for more sensitive (less noisy) SO- and CMBS4-like experiments, $N^{(1)}(L)$ is a factor of few to an order of magnitude smaller than $N^{(0)}(L)$. On small scales it can however become comparable to the signal spectrum and is important to model in any likelihood analysis. It would be straightforward to apply the perturbative likelihood approximation of [9, 10] (which accounts consistently for the signal dependence of $N^{(1)}(L)$) to the GMV estimator. In the right panel of Fig. 3, we show the ratio of the $N^{(1)}(L)$ for SQE and GMV estimators for different experiments considered here. For SO- and CMBS4-like experiments, $N^{(1)}(L)$ bias for the GMV estimator is smaller than for the SQE estimator for $L \lesssim 1800$. For *Planck*-like experiments, apart from the large angular scales $L \lesssim 200$ where GMV estimator gives smaller $N^{(1)}(L)$ bias than the SQE estimator, both the estimators have almost the same $N^{(1)}(L)$ bias.

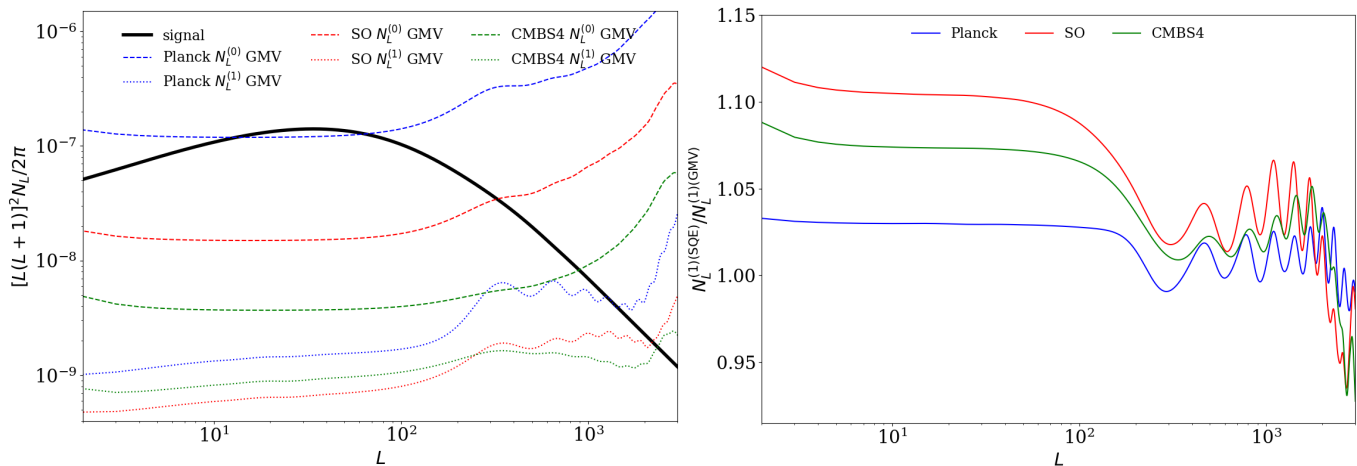


FIG. 3: Comparison of the $N^{(0)}(L)$ and $N^{(1)}(L)$ curves for the GMV estimator (left) and ratio of the $N^{(1)}(L)$ for SQE and GMV estimators (right) for given experimental configurations.

Appendix B: Numerical convergence

In Fig. 4, we show the result of the convergence test we perform for our Python code. It shows the % change in the noise calculation of HO02 (dashed curves) and OH03 (solid curves) estimators when we double the number of steps in the angular part of the integration for the given ℓ_{\max} and $L \lesssim 3000$. The dashed HO02 curves mostly overlap with the solid OH03 curves and thus are not distinctly visible. This shows that our numerical integrals are converged to better than 0.01% relative accuracy up to $L \approx 2000$, and to better than 0.03% for $L \lesssim 3000$, and thus makes us confident that the improvement observed for OH03 over HO02 for small angular scales is not a numerical artifact.

Appendix C: GMV to HO02 comparison

We perform the following exercise to compare the GMV and HO02 estimators. As mentioned at the end of Sec. IV A, the GMV estimator can be split into two independent estimators $\hat{\phi}_{\{TT,TE,EE\}}^{\text{GMV}}$ and $\hat{\phi}_{\{TB,EB\}}^{\text{GMV}}$. We compare the minimum variance reconstruction noise of these individual estimators with their HO02 counterparts i.e. $\hat{\phi}_{\{TT,TE,EE\}}^{\text{HO02}}$ and $\hat{\phi}_{\{TB,EB\}}^{\text{HO02}}$. This is shown in Fig. 5, where we plot the ratio of the reconstruction noises for the GMV and HO02 versions of these two estimators. As we can see, HO02 performs almost equally well as the GMV estimator for $\{TB,EB\}$ pair. The overall improvement of the GMV estimator over HO02 estimator is thus mainly driven by $\{TT,TE,EE\}$, especially for more sensitive SO- and CMBS4-like experiments. We also show an ideal case setup in Fig. 5 which corresponds to a noise-less experiment with the same multipole ranges as other experiments considered. HO02 estimator for a CMBS4-like setup considered here performs almost as well as the GMV estimator for $\{TB,EB\}$ pair and very slightly under-performs for the $\{TT,TE,EE\}$ set when compared to the ideal case setup.

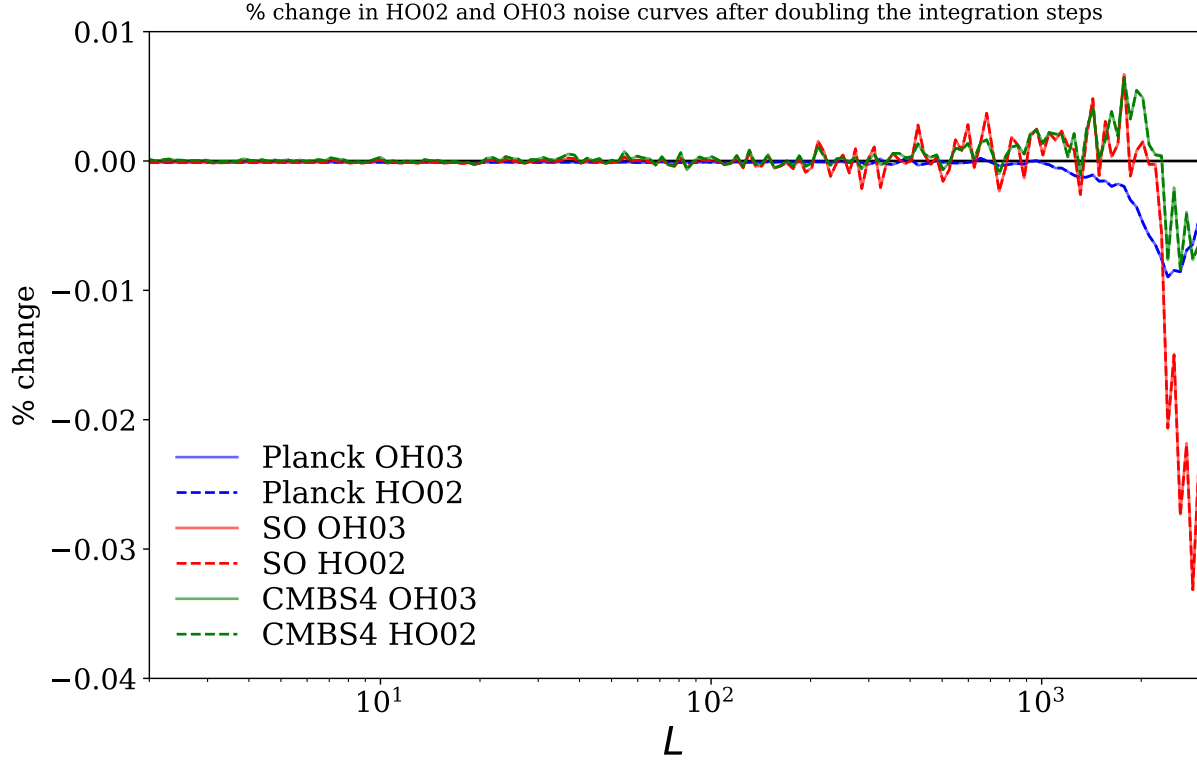


FIG. 4: % change in the noise curves for HO02 and OH03 estimators when we double the number of integration steps in our Python code.

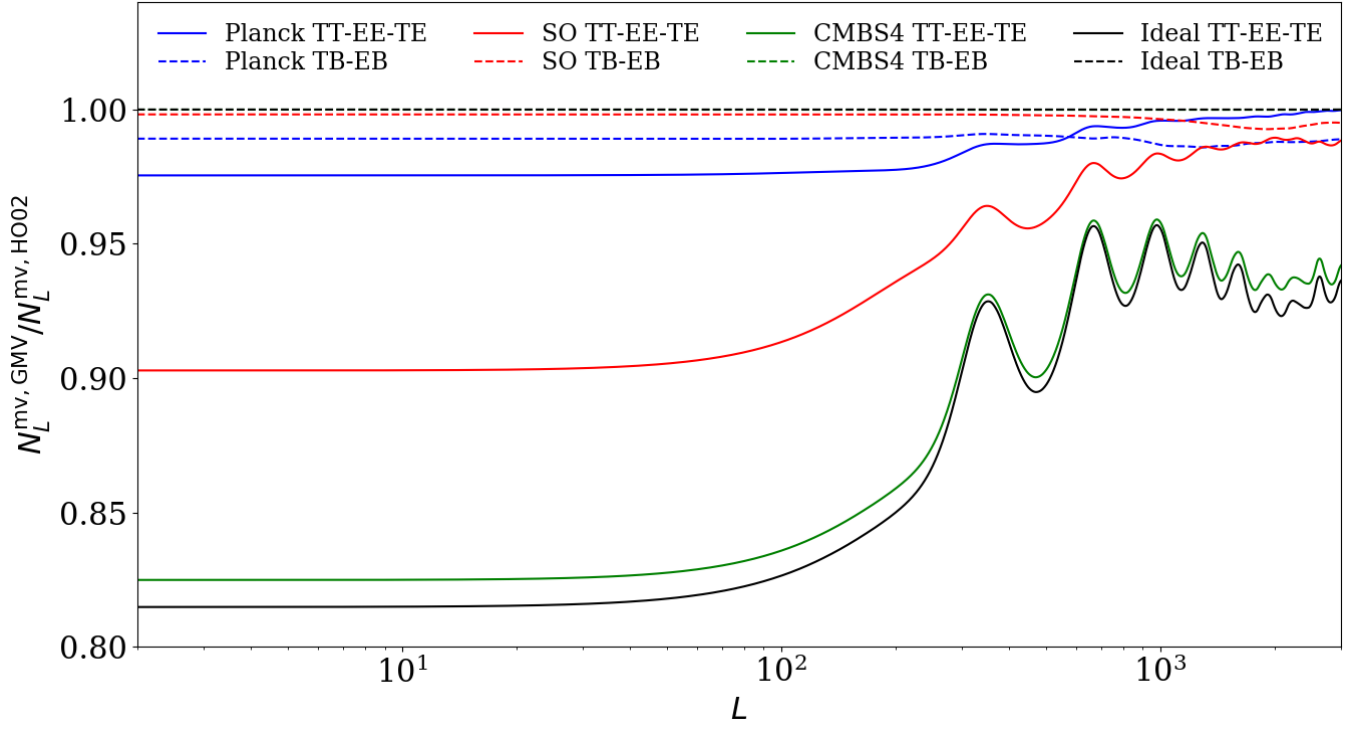


FIG. 5: Ratio of the minimum variance reconstruction noise of the GMV and HO02 estimators for $\hat{\phi}_{\{TT,TE,EE\}}$ and $\hat{\phi}_{\{TB,EB\}}$ estimators for different experimental setups. The ideal setup corresponds to a noise-less case.

-
- [1] A. Blanchard and J. Schneider, *Astron. Astrophys.* **184**, 1 (1987).
- [2] A. Lewis and A. Challinor, *Phys. Rep.* **429**, 1 (2006), [arXiv:astro-ph/0601594 \[astro-ph.CO\]](#).
- [3] U. Seljak and M. Zaldarriaga, *Phys. Rev. Lett.* **82**, 2636 (1999), [arXiv:astro-ph/9810092 \[astro-ph\]](#).
- [4] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, *Phys. Rev. D* **92**, 123535 (2015), [arXiv:1509.07471 \[astro-ph.CO\]](#).
- [5] M. Schmittfull and U. Seljak, *Phys. Rev. D* **97**, 123540 (2018), [arXiv:1710.09465 \[astro-ph.CO\]](#).
- [6] S. Das *et al.*, *Phys. Rev. Lett.* **107**, 021301 (2011), [arXiv:1103.2124 \[astro-ph.CO\]](#).
- [7] B. D. Sherwin *et al.*, *Phys. Rev. D* **95**, 123529 (2017), [arXiv:1611.09753 \[astro-ph.CO\]](#).
- [8] A. van Engelen *et al.*, *Astrophys. J.* **756**, 142 (2012), [arXiv:1202.0546 \[astro-ph.CO\]](#).
- [9] P. A. R. Ade *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A15 (2016), [arXiv:1502.01591 \[astro-ph.CO\]](#).
- [10] N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A8 (2020), [arXiv:1807.06210 \[astro-ph.CO\]](#).
- [11] Y. Omori *et al.*, *Astrophys. J.* **849**, 124 (2017), [arXiv:1705.00743 \[astro-ph.CO\]](#).
- [12] K. T. Story *et al.*, *Astrophys. J.* **810**, 50 (2015), [arXiv:1412.4760 \[astro-ph.CO\]](#).
- [13] W. L. K. Wu *et al.*, *Astrophys. J.* **884**, 70 (2019), [arXiv:1905.05777 \[astro-ph.CO\]](#).
- [14] M. Millea *et al.*, (2020), [arXiv:2012.01709 \[astro-ph.CO\]](#).
- [15] P. A. R. Ade *et al.* (BICEP2 Collaboration and Keck Array Collaboration), *Astrophys. J.* **833**, 228 (2016), [arXiv:1606.01968 \[astro-ph.CO\]](#).
- [16] P. A. R. Ade *et al.* (Polarbear Collaboration), *Phys. Rev. Lett.* **113**, 021301 (2014), [arXiv:1312.6646 \[astro-ph.CO\]](#).
- [17] S. Adachi *et al.* (Polarbear Collaboration), *Phys. Rev. Lett.* **124**, 131301 (2020), [arXiv:1909.13832 \[astro-ph.CO\]](#).
- [18] S. W. Henderson *et al.*, *Journal of Low Temperature Physics* **184**, 772 (2016), [arXiv:1510.02809 \[astro-ph.IM\]](#).
- [19] B. A. Benson *et al.*, in *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9153, edited by W. S. Holland and J. Zmuidzinas (2014) p. 91531P, [arXiv:1407.2973 \[astro-ph.IM\]](#).
- [20] P. Ade *et al.* (Simons Observatory Collaboration), *JCAP* **2019**, 056 (2019), [arXiv:1808.07445 \[astro-ph.CO\]](#).
- [21] K. N. Abazajian *et al.*, arXiv e-prints (2016), [arXiv:1610.02743 \[astro-ph.CO\]](#).
- [22] K. Abazajian *et al.* (CMB-S4 Collaboration), arXiv e-prints (2019), [arXiv:1907.04473 \[astro-ph.IM\]](#).
- [23] M. Zaldarriaga and U. Seljak, *Phys. Rev. D* **59**, 123507 (1999), [arXiv:astro-ph/9810257 \[astro-ph\]](#).
- [24] W. Hu, *Astrophys. J. Lett.* **557**, L79 (2001), [arXiv:astro-ph/0105424 \[astro-ph\]](#).
- [25] C. M. Hirata and U. Seljak, *Phys. Rev. D* **68**, 083002 (2003), [arXiv:astro-ph/0306354 \[astro-ph\]](#).
- [26] B. Hadzhiyska, B. D. Sherwin, M. Madhavacheril, and S. Ferraro, *Phys. Rev. D* **100**, 023547 (2019), [arXiv:1905.04217 \[astro-ph.CO\]](#).
- [27] C. M. Hirata and U. Seljak, *Phys. Rev. D* **67**, 043001 (2003), [arXiv:astro-ph/0209489 \[astro-ph\]](#).
- [28] J. Carron and A. Lewis, *Phys. Rev. D* **96**, 063510 (2017), [arXiv:1704.08230 \[astro-ph.CO\]](#).
- [29] M. Millea, E. Anderes, and B. D. Wandelt, *Phys. Rev. D* **102**, 123542 (2020), [arXiv:2002.00965 \[astro-ph.CO\]](#).
- [30] W. Hu and T. Okamoto, *Astrophys. J.* **574**, 566 (2002), [arXiv:astro-ph/0111606 \[astro-ph\]](#).
- [31] A. Lewis, A. Challinor, and D. Hanson, *JCAP* **2011**, 018 (2011), [arXiv:1101.2234 \[astro-ph.CO\]](#).
- [32] G. Fabbian, A. Lewis, and D. Beck, *JCAP* **2019**, 057 (2019), [arXiv:1906.08760 \[astro-ph.CO\]](#).
- [33] T. Okamoto and W. Hu, *Phys. Rev. D* **67**, 083002 (2003), [arXiv:astro-ph/0301031 \[astro-ph\]](#).
- [34] S. Ferraro and J. C. Hill, *Phys. Rev. D* **97**, 023512 (2018), [arXiv:1705.06751 \[astro-ph.CO\]](#).
- [35] E. Schaen and S. Ferraro, *Phys. Rev. Lett.* **122**, 181301 (2019), [arXiv:1804.06403 \[astro-ph.CO\]](#).
- [36] M. Kesden, A. Cooray, and M. Kamionkowski, *Phys. Rev. D* **67**, 123507 (2003), [arXiv:astro-ph/0302536 \[astro-ph\]](#).
- [37] J. Carron, (2019), [arXiv:1908.02016 \[astro-ph.CO\]](#).
- [38] J. Peloton, M. Schmittfull, A. Lewis, J. Carron, and O. Zahn, *Phys. Rev. D* **95**, 043508 (2017), [arXiv:1611.01446 \[astro-ph.CO\]](#).
- [39] M. S. Madhavacheril and J. C. Hill, *Phys. Rev. D* **98**, 023534 (2018), [arXiv:1802.08230 \[astro-ph.CO\]](#).
- [40] O. Darwish *et al.*, *Mon. Not. Roy. Ast. Soc.* **500**, 2250 (2021), [arXiv:2004.01139 \[astro-ph.CO\]](#).
- [41] A. van Engelen, S. Bhattacharya, N. Sehgal, G. P. Holder, O. Zahn, and D. Nagai, *Astrophys. J.* **786**, 13 (2014), [arXiv:1310.7023 \[astro-ph.CO\]](#).
- [42] M. Zaldarriaga and U. Seljak, *Phys. Rev. D* **58**, 023003 (1998), [astro-ph/9803150](#).
- [43] U. Seljak and C. M. Hirata, *Phys. Rev. D* **69**, 043005 (2004), [arXiv:astro-ph/0310163 \[astro-ph\]](#).