PHYSICAL REVIEW D **88,** 063012 (2013)

Joint analysis of CMB temperature and lensing-reconstruction power spectra

Marcel M. Schmittfull,[1] Anthony Challinor,[2,1] Duncan Hanson,[3] and Antony Lewis[4,*]

[1]*DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*
[2]*Institute of Astronomy and Kavli Institute for Cosmology, Madingley Road, Cambridge CB3 0HA, United Kingdom*
[3]*Department of Physics, McGill University, Montreal QC H3A 2T8, Canada*
[4]*Department of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, United Kingdom*

(Received 2 August 2013; published 30 September 2013)

Gravitational lensing provides a significant source of cosmological information in modern cosmic microwave background parameter analyses. It is measured in both the power spectrum and trispectrum of the temperature fluctuations. These observables are often treated as independent, although as they are both determined from the same map, this is impossible. In this paper, we perform a rigorous analysis of the covariance between lensing power spectrum and trispectrum analyses. We find two dominant contributions coming from (i) correlations between the disconnected noise bias in the trispectrum measurement and sample variance in the temperature power spectrum and (ii) sample variance of the lenses themselves. The former is naturally removed when the dominant $N^{(0)}$ Gaussian bias in the reconstructed deflection spectrum is dealt with via a partially data-dependent correction, as advocated elsewhere for other reasons. The remaining lens-cosmic-variance contribution is easily modeled but can safely be ignored for a Planck-like experiment, justifying treating the two observable spectra as independent. We also test simple likelihood approximations for the deflection power spectrum, finding that a Gaussian with a parameter-independent covariance performs well.

DOI: 10.1103/PhysRevD.88.063012 PACS numbers: 98.70.Vc

I. INTRODUCTION

Weak gravitational lensing by large-scale structures leaves subtle imprints in the temperature anisotropies of the cosmic microwave background (CMB); see Ref. [1] for a review. These imprints can be detected in surveys with resolution better than a few arcminutes and used to reconstruct the lensing deflection field [2,3]. Since the lensing deflections depend on the growth of structure and geometry at much lower redshifts ($z \sim 2$) than the CMB last-scattering surface, lens reconstructions can be used to constrain parameters that are largely degenerate in the primary anisotropies sourced at last scattering. Examples include sub-eV neutrino masses, spatial curvature, dark energy, and modifications to gravity (e.g., Refs. [4–9]).

Lensing is an emerging frontier of observational cosmology. The first direct measurements of the deflection power spectrum were reported recently by the ACT [10,11], SPT [12], and Planck [13] teams with significances of 4.6, 6.3, and 25σ, respectively. These measurements provide the first evidence for dark energy from the CMB *alone*. Since lens reconstructions are quadratic in the temperature anisotropies, the power spectrum of the reconstruction is probing the 4-point non-Gaussianity of the CMB induced by lensing [14]. The statistical power of lens reconstructions is expected to improve rapidly with ongoing analyses of the full 2500 deg^2 SPT survey and the full-mission data from Planck, which also allow for polarization-based lensing reconstruction. Lensing also affects the power spectrum (or 2-point function) of the temperature anisotropies, smoothing the acoustic peaks and transferring power from large to small scales (e.g., Ref. [15]). The smoothing effect has been detected at nearly 10σ in current temperature power spectrum measurements [16,17].

A question that has received only limited attention to date is how one should model the likelihood of the lensed CMB anisotropies when deriving constraints on cosmological parameters. As the unlensed CMB and deflection field can be approximated as Gaussian on the scales relevant for CMB lensing, it is straightforward to write down a formal expression for the likelihood of the lensed temperature [18]. However, this is very difficult to work with directly. Indeed, working with the exact likelihood even for Gaussian fields in megapixel maps is computationally prohibitive. Instead, in the Gaussian case, at high multipoles the data are usually compressed to an empirical power spectrum (or set of cross-spectra), and an approximate likelihood is constructed based on this spectrum and its covariance. Such an approach is both computationally feasible and allows for the robust treatment of instrumental effects such as beam asymmetry. For non-Gaussian fields, like the lensed CMB, working with only the empirical power spectrum is clearly lossy. Instead, we should include further empirical connected n-point functions in our compressed data. In the context of CMB lensing, the 4-point function is the most relevant moment, and the information it carries is captured in the estimated power spectrum of the reconstructed deflection field.

*http://cosmologist.info

TABLE I. Summary of the main quantities used throughout this paper (roughly ordered after first appearance).

Symbol	Description	Definition
T	Unlensed CMB temperature	Eq. (1)
$\tilde{T}$	Lensed CMB temperature	Eq. (1)
$\phi_{\rm in}$	Input lensing potential field in simulations	
$\hat{\phi}$	Lensing reconstruction (quadratic in $\tilde{T}$)	Eq. (9)
A_L, $\tilde{g}$	Normalization and weights for $\hat{\phi}$	Eqs. (10) and (11)
C_l^{XX}	Fiducial theoretical power spectrum of $X = T$, $\tilde{T}$ or ϕ (without noise/beam)	
$C_{l,\rm expt}^{\tilde{T}\tilde{T}}$	Lensed temperature power spectrum including beam-deconvolved noise	Eq. (18)
$\hat{C}_l^{XX}$	Empirical power spectrum of a realization of $X = T$, $\tilde{T}$, $\phi_{\rm in}$ or $\hat{\phi}$	Eq. (19)
$N_l^{(0)}$	Gaussian, fully disconnected noise bias of $\hat{C}_l^{\hat{\phi}\hat{\phi}}$	Eq. (16)
$N_l^{(1)}$, $N_l^{(2)}$	Biases of $\hat{C}_l^{\hat{\phi}\hat{\phi}}$ at $\mathcal{O}(C^{\phi\phi})$ and $\mathcal{O}[(C^{\phi\phi})^2]$	Ref. [26]
$\hat{N}_l^{(0)}$	Data-dependent, empirical bias subtraction term	Eqs. (17) and (32)
$\mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\rm expt}^{\tilde{T}\tilde{T}})$	Covariance of CMB temperature and lensing reconstruction power spectra	Eq. (33)
$\mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\rm expt}^{\tilde{T}\tilde{T}})_{\rm disconn}$	Noise contribution (from Gaussian, fully disconnected CMB 6-point function)	Eqs. (35) and (36)
$\mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\rm expt}^{\tilde{T}\tilde{T}})_{[(\rm non\text{-})\rm primary]}^{\rm conn\ 4\ pt\ }{}^{X}$	Trispectrum contributions of type $X = A$, B [from (non)primary coupling]	Appendix D
$\mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\rm expt}^{\tilde{T}\tilde{T}})_{\langle \tilde{T}_1 \cdots \tilde{T}_6 \rangle_{c,\rm dom}^{(4)}}$	Matter cosmic variance contribution [from $\mathcal{O}(\phi^4)$ connected CMB 6-point function]	Eqs. (45) and (E8)
A	Overall lensing amplitude of a fiducial lensing power spectrum	$C_l^{\phi\phi} = A C_l^{\phi\phi}\vert_{\rm fid}$
$\hat{A}$	Estimator for A based on reconstruction power spectrum, i.e., CMB trispectrum	Eq. (54)
$\hat{A}'$	Estimator for A based on CMB power spectrum	Eq. (55)
n	Tilt of a fiducial lensing power spectrum $C_l^{\phi\phi}\vert_{\rm fid}$	Eq. (66)

Parameter analyses involving lens reconstruction to date have followed the route described above. However, they have simply combined estimates of the temperature power spectrum and the lensing power spectrum as if they were independent [12,13,19]. Since both power spectra are derived from the same CMB temperature map, one might question the validity of this approach, raising the concern that lensing information is inadvertently being double counted. While early lensing forecasts [4,5,20] addressed this by using *unlensed* CMB power spectra, an optimal combination of the observed lensed CMB 2- and 4-point functions should model their cross-covariance. Intuitively, we might expect two effects to be relevant. First, the statistical noise in the reconstruction, due to chance alignments in the unlensed CMB which mimic the locally anisotropic effects of lensing, is dependent on the CMB fluctuations themselves. If, due to cosmic variance, the unlensed temperature anisotropies fluctuate high at some particular scale, the noise in the lens reconstruction will also fluctuate high. The mode-coupling nature of lens reconstruction, whereby modes near the resolution limit of the observation dominate the reconstruction on much larger scales, will lead to broad correlations between the temperature and lensing power spectra. Since the reconstruction is quadratic, the correlation of the power spectra involves the CMB 6-point function, the disconnected part of which arises from the effect just described. The second effect is due to the cosmic variance of the lenses. If the lensing field fluctuates high at some scale, the reconstructed power will also fluctuate high at the same scale. In parallel, there will be more smoothing of the acoustic peaks in the measured CMB power spectrum giving an anticorrelation with the reconstucted lensing power at the location of acoustic peaks and a positive correlation at troughs. This effect is second order in the deflection power and derives from the connected part of the CMB 6-point function. As we shall show, the induced *correlations* are rather small (a few percent) for a Planck-like experiment for broadband measures of the lensing power such as a lensing power spectrum amplitude. Essentially, this is because there is a limited number of modes of the lensing power spectrum that influence the acoustic part of the temperature power spectrum, and the correlation due to cosmic variance of these modes is diluted by the significant noise due to cosmic variance of the CMB and instrumental noise (i.e., the fact that lensing measurements from the CMB 2- and 4-point functions are not limited by cosmic variance of the lenses). Moreover, the first effect mentioned above produces rather small lensing amplitude correlations since CMB modes at different scales fluctuate independently, and most of the information on peak smearing in the CMB power spectrum comes from modes near the acoustic peaks and troughs, whereas the reconstruction of (large-scale) lenses is most effective at CMB scales

where the CMB spectrum changes most rapidly, i.e., between acoustic peaks and troughs. Note also that even the unlensed CMB anisotropies are correlated with the lensing field due to the late-time integrated Sachs–Wolfe effect [21]. This mostly produces a diagonal covariance between the power spectra of the lens reconstruction and the lensed CMB, which falls rapidly at small scales ($L > 100$). One of our main aims in this paper is to quantify these arguments with detailed calculations of the 6-point function, verified with simulations, and to assess whether the correlations are small enough to be safely ignored in the likelihood.

A further issue concerns the form of the likelihood for the power spectrum estimates. This has been well studied for Gaussian fields, e.g., Refs. [22–25], and simple approximate forms are known to give reliable parameter constraints when applied on all but the largest scales. However, the lensing reconstruction is quadratic in the nearly Gaussian CMB fluctuations and therefore highly non-Gaussian. Here, we test a particularly simple form of the lensing power likelihood, a Gaussian with model-independent covariance. Constraining the amplitude and tilt of a fiducial deflection power spectrum, we demonstrate that the fiducial-Gaussian approximation performs well, returning maximum-likelihood points that scatter across simulations in a manner consistent with the width of the likelihood.

The paper is organized as follows. We review CMB lensing reconstruction in Sec. II, and we describe our simulations of lensed CMB maps and the mechanics of our reconstructions in Sec. III. Section IV surveys known results for the autocorrelations of the lensed CMB temperature power spectrum and the reconstruction power spectrum. In Sec. V we present results for the cross-correlation of these power spectra and assess the importance of correlations for estimating cosmological parameters. We test likelihood approximations for the lensing reconstruction (in isolation) in Sec. VI and conclude in Sec. VII. In Appendix A we provide intuitive arguments for the magnitude of the temperature-lensing power correlation. A series of further appendices provide calculational details and motivate some of the approaches taken in the main text. Table I summarizes the key quantities, and their definitions and symbols, used in this paper.

II. CMB LENSING RECONSTRUCTION

The lensed CMB temperature $\tilde{T}(\hat{\mathbf{n}})$ along direction $\hat{\mathbf{n}}$ is related to the unlensed temperature $T(\hat{\mathbf{n}})$ by the deflection field $\alpha(\hat{\mathbf{n}})$[1]

$$\tilde{T}(\hat{\mathbf{n}}) = T[\hat{\mathbf{n}} + \alpha(\hat{\mathbf{n}})]. \tag{1}$$

[1]The notation here is rather symbolic on the spherical sky: the point $\hat{\mathbf{n}} + \alpha(\hat{\mathbf{n}})$ is understood to be obtained from $\hat{\mathbf{n}}$ by displacing through a distance $|\alpha(\hat{\mathbf{n}})|$ along the geodesic that is tangent to α at $\hat{\mathbf{n}}$ [27].

The deflection angle can be written in the Born approximation as the angular gradient of a lensing potential, $\alpha(\hat{\mathbf{n}}) = \nabla\phi(\hat{\mathbf{n}})$, which is given by an integral along the (unperturbed) line of sight of the gravitational potentials (e.g., Ref. [1]).

On the full sky, the effect of lensing on the CMB temperature can be expressed perturbatively by Taylor expanding Eq. (1). Expanding in spherical harmonics, the multipoles of the lensed CMB, $\tilde{T}_{lm}$, are related to those of the unlensed CMB T_{lm} and the lensing potential ϕ_{lm} via [28]

$$\tilde{T}_{lm} = T_{lm} + \delta T_{lm} + \delta^2 T_{lm} + \cdots, \tag{2}$$

where changes due to lensing $\delta^n T_{lm}$ are of order $\mathcal{O}(\phi^n)$ and linear in the unlensed temperature T:

$$\delta T_{lm} = \sum_{\underline{l}_1,\underline{l}_2} \phi_{\underline{l}_1} T_{\underline{l}_2} I^{mm_1m_2}_{ll_1l_2}, \tag{3}$$

$$\delta^2 T_{lm} = \sum_{\underline{l}_1,\underline{l}_2,\underline{l}_3} \phi_{\underline{l}_1}\phi_{\underline{l}_2} T_{\underline{l}_3} J^{mm_1m_2m_3}_{ll_1l_2l_3}, \tag{4}$$

where we have introduced the notation $\underline{l} \equiv (lm)$ and I denotes an angular integral over a product of (derivatives of) spherical harmonics given by [28]

$$I^{mm_1m_2}_{ll_1l_2} = (-1)^m \begin{pmatrix} l & l_1 & l_2 \\ -m & m_1 & m_2 \end{pmatrix} F_{ll_1l_2}. \tag{5}$$

Expressions for $J^{mm_1m_2m_3}_{ll_1l_2l_3}$ can be found in Ref. [26]. The geometrical factor $F_{l_1Ll_2}$, which is symmetric in the last two indices, is given by

$$F_{l_1Ll_2} = [L(L+1) - l_1(l_1+1) + l_2(l_2+1)] \times \sqrt{\frac{(2l_1+1)(2L+1)(2l_2+1)}{16\pi}} \begin{pmatrix} l_1 & L & l_2 \\ 0 & 0 & 0 \end{pmatrix} \tag{6}$$

and describes the rotationally invariant part of the coupling between the three multipoles.

It is possible to reconstruct the lensing potential $\phi(\hat{\mathbf{n}})$ from the observed CMB by exploiting the fact that fixed lenses introduce correlations between temperature modes [2,29]. Following the nonperturbative calculations in Ref. [30], we have

$$\left\langle \frac{\partial}{\partial\phi_{LM}} (\tilde{T}_{l_1m_1}\tilde{T}_{l_2m_2}) \right\rangle \approx (-1)^M \begin{pmatrix} L & l_1 & l_2 \\ -M & m_1 & m_2 \end{pmatrix} \tilde{f}_{l_1Ll_2}, \tag{7}$$

where $\tilde{f}_{l_1Ll_2}$ is symmetric in l_1 and l_2 and contains the lensed temperature power spectrum, $C_l^{\tilde{T}\tilde{T}}$:

$$\tilde{f}_{l_1Ll_2} = C_{l_2}^{\tilde{T}\tilde{T}} F_{l_1Ll_2} + C_{l_1}^{\tilde{T}\tilde{T}} F_{l_2Ll_1}. \tag{8}$$

The angle brackets in Eq. (7) denote the expectation value over ϕ and T (and noise) and we have neglected the T–ϕ

correlation which is generally a good approximation for CMB lensing since we are usually interested in small-scale ($l > 600$) modes of the temperature where the correlation is small. More generally, we shall neglect the T–ϕ correlation throughout this paper, except where explicitly stated otherwise. Equation (7) motivates forming a quadratic estimator for the lensing potential [31],

$$\hat{\phi}_{LM} = A_L \sum_{\underline{l}_1 \underline{l}_2} (-1)^M \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \tilde{g}_{l_1 l_2}(L) \tilde{T}_{\underline{l}_1} \tilde{T}_{\underline{l}_2}. \tag{9}$$

For any choice of weights $\tilde{g}_{l_1 l_2}(L)$, we determine the normalization A_L by demanding that $\langle \partial \hat{\phi}_{LM} / \partial \phi_{L'M'} \rangle = \delta_{LL'} \delta_{MM'}$, which gives

$$A_L = (2L+1) \left(\sum_{l_1 l_2} \tilde{f}_{l_1 L l_2} \tilde{g}_{l_1 l_2}(L) \right)^{-1}. \tag{10}$$

One can determine optimal weights by minimizing the variance of the estimator to find[2]

$$\tilde{g}_{l_1 l_2}(L) = \frac{\tilde{f}_{l_1 L l_2}}{2 C^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}} C^{\tilde{T}\tilde{T}}_{l_2,\mathrm{expt}}}, \tag{11}$$

where the numerator contains the lensed temperature power $C_l^{\tilde{T}\tilde{T}}$, while the denominator involves the total power spectrum for the experiment $C^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}}$ including beam-deconvolved noise; see Eq. (18) below.

The expectation value of the simple power spectrum estimate,

$$\hat{C}_L^{\hat{\phi}\hat{\phi}} \equiv \frac{1}{2L+1} \sum_{M=-L}^{L} \hat{\phi}_{LM} \hat{\phi}^*_{LM}, \tag{12}$$

involves the 4-point function of the lensed CMB. The connected part of the 4-point function can be written in terms of the fully reduced trispectrum $\mathbb{T}^{l_1 l_2}_{l_3 l_4}(L)$ as [14]

$$\begin{aligned} &\langle \tilde{T}_{l_1 m_1} \tilde{T}_{l_2 m_2} \tilde{T}_{l_3 m_3} \tilde{T}_{l_4 m_4} \rangle_c \\ &\quad = \frac{1}{2} \sum_{LM} (-1)^M \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & M \end{pmatrix} \begin{pmatrix} l_3 & l_4 & L \\ m_3 & m_4 & -M \end{pmatrix} \\ &\qquad \times \mathbb{T}^{l_1 l_2}_{l_3 l_4}(L) + \text{perms}. \end{aligned} \tag{13}$$

By evaluating the trispectrum correct to $\mathcal{O}[(C_L^{\phi\phi})^2]$, Ref. [26] shows that the dominant terms for lensing reconstruction can be approximated by

$$\mathbb{T}^{l_1 l_2}_{l_3 l_4}(L) \approx \frac{1}{4} C_L^{\phi\phi} \tilde{f}_{l_1 L l_2} \tilde{f}_{l_3 L l_4}, \tag{14}$$

which involves the lensed $C_l^{\tilde{T}\tilde{T}}$ (via $\tilde{f}_{l_1 L l_2}$, etc.). The normalization of Eq. (10) therefore correctly normalizes the power spectrum $\hat{C}_l^{\hat{\phi}\hat{\phi}}$, avoiding a bias $[N^{(2)}]$ of around -15% on large scales that results from using the unlensed spectra, i.e., $f_{l_1 L l_2}$, in the normalization A_L.

Taking the expectation value $\langle \hat{C}_L^{\hat{\phi}\hat{\phi}} \rangle$, and using Eq. (13), gives [26,32]

$$\langle \hat{C}_L^{\hat{\phi}\hat{\phi}} \rangle = C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)} + \mathcal{O}((C^{\phi\phi})^3), \tag{15}$$

where $N_L^{(n)}$ is of order $(C^{\phi\phi})^n$ if we do not count appearances of $C^{\phi\phi}$ in the lensed temperature power spectrum. The disconnected part of the lensed temperature 4-point function leads to the Gaussian bias

$$N_L^{(0)} = \frac{2A_L^2}{2L+1} \sum_{l_1, l_2} \tilde{g}^2_{l_1 l_2}(L) C^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}} C^{\tilde{T}\tilde{T}}_{l_2,\mathrm{expt}} = A_L, \tag{16}$$

where the last equality holds only if the weights $\tilde{g}_{l_1 l_2}(L)$ are given by Eq. (11). Since this bias is present even in the absence of lensing, it corresponds to the power spectrum generated by the statistical noise of the lensing reconstruction (see Appendix C for a more detailed discussion of how this bias is generated). The $N^{(1)}$ bias is due to those permutations in Eq. (13) that mix multipoles between the primary $(l_1 m_1, l_2 m_2)$ and $(l_3 m_3, l_4 m_4)$ couplings, e.g., $1 \leftrightarrow 3$. It has been computed in Refs. [26,32]; see also Fig. 1. An unbiased estimate for the lensing potential power spectrum can be obtained by subtracting $N^{(0)} + N^{(1)}$ from $\hat{C}^{\hat{\phi}\hat{\phi}}$. Here, for $N^{(1)}$ we subtract the bias evaluated in the fiducial cosmology used for our simulations. In practice, the variation of $N^{(1)}$ with the cosmological model should be included in the likelihood analysis. However, since direct evaluation of this term is slow, a faster alternative is to include the uncertainty in $N^{(1)}$ as a correlated error in the covariance matrix for $\hat{C}_L^{\hat{\phi}\hat{\phi}}$[3].

The subtraction of the Gaussian $N^{(0)}$ bias can be done in various ways. The simplest is to subtract a fiducial model. For temperature reconstructions, $N^{(0)}$ generally exceeds the signal power on all scales, and by around two orders of magnitude at multipoles $l \sim 1000$, so that accurate $N^{(0)}$ subtraction is critical. For the idealized isotropic surveys considered here, this is perhaps not too problematic since only $C^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}}$ is required to calculate $N^{(0)}$, and this can be estimated from a smoothed version of the measured power

[2]The normalization (10) and weight (11) differ a little from that in the original works of Hu and coworkers [3,31] since $\tilde{f}_{l_1 L l_2}$ involves the power spectrum of the lensed rather than unlensed CMB. The form here is particularly suited to power spectrum estimation.

[3]Note also that the normalization A_L is dependent on the cosmological model through $\tilde{f}_{l_1 L l_2}$. However, the lensed temperature power spectrum, and hence A_L, generally varies much less across models consistent with the measured CMB power spectrum than $C_L^{\phi\phi}$ (and so $N^{(1)}$). In practical applications, if a fiducial model is assumed to normalize the 4-point function, the parameter-dependent $C_L^{\phi\phi}$ can easily be renormalized in the likelihood by the ratio of the fiducial A_L to that at the current location in parameter space [13].

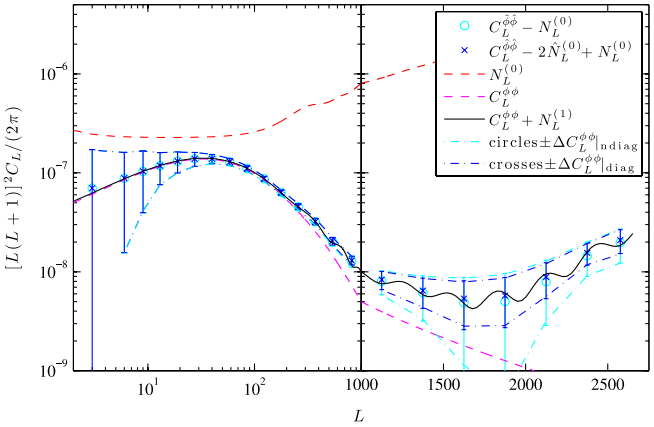


FIG. 1 (color online). Reconstructed lensing potential power spectrum $C_L^{\hat\phi\hat\phi}$ averaged over 1000 simulations, after subtracting the analytically calculated $N^{(0)}$ bias (cyan circles). The magenta line is the input lensing potential power. Blue symbols include the empirical reconstruction bias correction of Eq. (17). The error bars show the estimated standard deviation of the binned power spectrum for a single realization of the lensed CMB. Cyan error bars are only visible when they disagree from the blue error bars. Cyan dashed-dotted lines show theoretical error bars obtained by binning Eqs. (28) and (29) below. Blue dashed-dotted lines are theoretical error bars from Eq. (29) only. The biases $N^{(0)}$ (red) and $N^{(1)}$ (black, which also includes $C_L^{\phi\phi}$) are calculated analytically. Note that the left-hand panel uses a log scale for multipole L, whereas the right-hand panel uses a linear scale.

spectrum. However, in the presence of survey anisotropies, the Gaussian bias must generally be subtracted via simulations, and this requires an accurate procedure for simulating maps including all relevant real-world effects such as noise inhomogeneities and correlations, beam asymmetry, and unresolved foreground emission. A work around to these issues is to use alternative data-dependent forms of $N^{(0)}$ (see Ref. [33] and references therein) for which the expectation value either equals $N^{(0)}$ or closely approximates it. Here, we use the form (for an isotropic survey) $\hat C_L^{\hat\phi\hat\phi} - (2\hat N_L^{(0)} - N_L^{(0)})$, advocated in Refs. [26,33], where

$$\hat N_L^{(0)} = \frac{2A_L^2}{2L+1}\sum_{l_1,l_2}\tilde g_{l_1 l_2}^2(L)\hat C_{l_1,\rm expt}^{\tilde T\tilde T} C_{l_2,\rm expt}^{\tilde T\tilde T}; \qquad (17)$$

i.e., we replace one occurrence of $C_{l,\rm expt}^{\tilde T\tilde T}$ in $N^{(0)}$ with the empirical temperature power spectrum of our given sky. Note that the expectation value $\langle\hat N_L^{(0)}\rangle = N_L^{(0)}$. As well as reducing the impact of modelling errors, this form of $N^{(0)}$ subtraction has the benefit of greatly reducing correlations between the $\hat C_L^{\phi\phi}$ estimates that arise from the disconnected part of the CMB 8-point function; see Sec. IV B and Ref. [26]. As we shall see in Sec. V A, it also eliminates the Gaussian contribution to the covariance between $\hat C_L^{\hat\phi\hat\phi}$ and the measured temperature power spectrum. Further motivation for the $2\hat N_L^{(0)} - N_L^{(0)}$ construction comes from considering optimal trispectrum estimation for weakly non-Gaussian fields [34]; these arguments are discussed further in the context of CMB lensing in Appendix B.

Throughout this paper we assume that the instrumental beam has been deconvolved from the $\tilde T_{lm}$. The total power spectrum of the experiment is then of the form

$$C_{l,\rm expt}^{\tilde T\tilde T} = C_l^{\tilde T\tilde T} + \sigma_N^2 \exp[l(l+1)\sigma_{\rm FWHM}^2/(8\ln 2)], \qquad (18)$$

where σ_N^2 is the white noise power spectrum and $\sigma_{\rm FWHM}$ is the full width at half-maximum (FWHM) of the optical beam. Following Ref. [26], we will use $\sigma_N = 27\ \mu$K arcmin and $\sigma_{\rm FWHM} = 7$ arcmin, which is roughly appropriate for Planck.

III. SIMULATIONS

We use simulations to verify our analytic arguments. These are based on 1000 realizations of a flat ΛCDM cosmology with WMAP7 + BAO + H_0 parameters [35] $h = 0.704$, $\Omega_b h^2 = 0.0226$, $\Omega_c h^2 = 0.1123$, $\tau = 0.087$, $n_s = 0.963$, $\Delta_{\mathcal R}^2 = 2.441\times 10^{-9}$ at $k_0 = 0.002$ Mpc^{-1}, and three massless neutrino species. We start with realizations of the unlensed temperature and lensing potential up to $l_{\rm max}^{\rm in} = 3000$ (including the $C^{T\phi}$ correlation on large scales), which are then lensed using Lenspix [36]. The lensed temperature up to $l_{\rm max}^T = 2750$ is used to reconstruct the lensing potential up to $l_{\rm max}^\phi = 2650$ with the full-sky simulation setup of Hanson *et al.* [26]. The convolution in harmonic space in Eq. (9) is evaluated as a product in pixel space, where spin-1 spherical harmonic transforms are taken using HEALPix [37]. As a slight modification to Ref. [26], we use the reconstruction normalization A_L and weights $\tilde g$ given in Eqs. (10) and (11) to avoid the $N^{(2)}$ bias.

Figures 1 and 2 confirm that the power spectrum of the lensing reconstruction agrees with the input lensing power spectrum if the $N^{(0)}$ and $N^{(1)}$ biases are taken into account. Similar plots in Ref. [26] contain the $N^{(2)}$ bias because their simulations used reconstruction weights and normalization with unlensed instead of lensed temperature power spectra. The realization-dependent $\hat N^{(0)}$ bias correction does not change the expectation value of the reconstruction power spectrum but reduces its covariance.

IV. AUTOCORRELATIONS OF POWER SPECTRA

We will argue later that parameter estimation with the lensing reconstruction should be based on empirical power spectra defined by

$$\hat C_l^{XX} \equiv \frac{1}{2l+1}\sum_{m=-l}^{l}(-1)^m \hat X_{lm}\hat X_{l,-m}, \qquad (19)$$

where $\hat X_{lm}$ are the multipole coefficients of the realization $\hat X(\hat{\mathbf n})$ of a field on the sphere. Here, $X = \tilde T$ or ϕ. To construct a likelihood for the empirical power spectra,

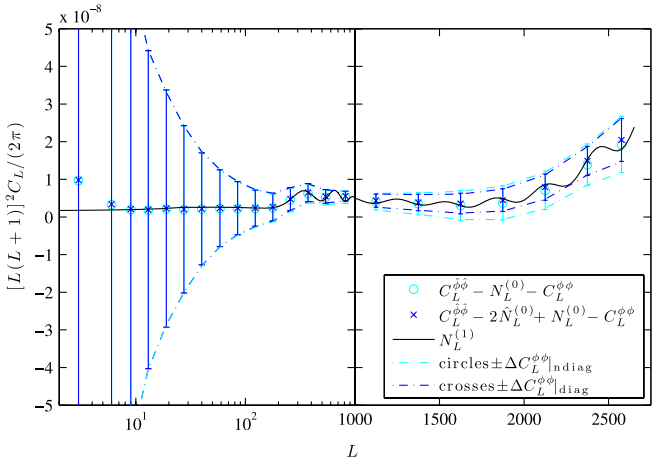


FIG. 2 (color online). Same as Fig. 1 after subtracting the theoretical lensing power spectrum $C_L^{\phi\phi}$.

we will model auto- and cross-correlations of the empirical power spectra of the observed lensed temperature and the reconstructed lensing potential. If X is a statistically isotropic, Gaussian field, then the power covariance is diagonal with

$$\mathrm{cov}(\hat C_l^{XX}, \hat C_{l'}^{XX}) = \delta_{ll'}\frac{2}{2l+1}\langle \hat C_l^{XX}\rangle^2. \tag{20}$$

In the following we will abbreviate the Gaussian variance with

$$\mathrm{var}_G(C_l^{XX}) \equiv \frac{2}{2l+1}(C_l^{XX})^2. \tag{21}$$

We demonstrate in Appendix F that for most applications, we can neglect the effect of $C^{T\phi}$ on the covariance between the power spectra of the lensed temperature and the lens reconstruction. Because the Taylor expansion of Eq. (2) is linear in the unlensed temperature, all odd n-point functions of the lensed temperature vanish.

A. Lensed temperature

The autocorrelation of the lensed temperature power spectrum has been computed at first order in $C^{\phi\phi}$ in Ref. [38] under the flat-sky approximation and in Ref. [39] on the full sky. A contribution at second order in $C^{\phi\phi}$ was recently identified in Ref. [40]. The power covariance is given by

$$\begin{aligned}&\mathrm{cov}(\hat C_{l,\mathrm{expt}}^{\tilde T\tilde T}, \hat C_{l',\mathrm{expt}}^{\tilde T\tilde T})\\ &\quad= \delta_{ll'}\mathrm{var}_G(C_{l,\mathrm{expt}}^{\tilde T\tilde T}) + \frac{1}{(2l+1)(2l'+1)}\\ &\qquad\times\sum_{m,m'}(-1)^{m+m'}\langle \tilde T_{lm}\tilde T_{l,-m}\tilde T_{l'm'}\tilde T_{l',-m'}\rangle_c,\end{aligned} \tag{22}$$

where $\langle\rangle_c$ denotes the connected part of the 4-point function, which is at $\mathcal{O}(C^{\phi\phi})$ [14]

$$\begin{aligned}&\langle \tilde T_{\underline{l}_1}\tilde T_{\underline{l}_2}\tilde T_{\underline{l}_3}\tilde T_{\underline{l}_4}\rangle_c\\ &\quad= \frac{1}{2}C_{l_2}^{TT}C_{l_4}^{TT}\sum_{LM}(-1)^M C_L^{\phi\phi}\begin{pmatrix} l_1 & l_2 & L\\ m_1 & m_2 & -M\end{pmatrix}\\ &\qquad\times\begin{pmatrix} l_3 & l_4 & L\\ m_3 & m_4 & M\end{pmatrix}F_{l_1Ll_2}F_{l_3Ll_4} + \text{all perms.}\end{aligned} \tag{23}$$

Here and in the following "all perms" denotes permutations in all noncontracted multipole indices, i.e., permutations of 1, 2, 3, and 4 in Eq. (23). If we also include the contribution at second order in $C^{\phi\phi}$ from Ref. [40], we get [39,40][4]

$$\begin{aligned}&\mathrm{cov}(\hat C_{l,\mathrm{expt}}^{\tilde T\tilde T}, \hat C_{l',\mathrm{expt}}^{\tilde T\tilde T})\\ &\quad= \delta_{ll'}\mathrm{var}_G(C_{l,\mathrm{expt}}^{\tilde T\tilde T}) + \frac{2}{(2l+1)(2l'+1)}\sum_L C_L^{\phi\phi}f_{lLl'}^2\\ &\qquad+\sum_L\frac{\partial C_l^{\tilde T\tilde T}}{\partial C_L^{\phi\phi}}\frac{2}{2L+1}(C_L^{\phi\phi})^2\frac{\partial C_{l'}^{\tilde T\tilde T}}{\partial C_L^{\phi\phi}} + \mathcal{O}[(C^{\phi\phi})^3],\end{aligned} \tag{24}$$

where $f_{l_1Ll_2}$ is from the unlensed version of Eq. (8): $f_{l_1Ll_2} = C_{l_2}^{TT}F_{l_1Ll_2} + C_{l_1}^{TT}F_{l_2Ll_1}$. The third term on the right of Eq. (24) arises from cosmic variance of the lenses. Fluctuations at lens multipole L produce fluctuations in the empirical lensed temperature power spectrum over a range of multipoles. The fluctuations in the lens power, $\Delta C_L^{\phi\phi}$, propagate to the empirical temperature power spectrum approximately as $\partial C_l^{\tilde T\tilde T}/\partial C_L^{\phi\phi}$. The power derivative here can be calculated perturbatively by noting that at $\mathcal{O}(C^{\phi\phi})$ [28]

$$C_l^{\tilde T\tilde T} = C_l^{TT}[1 - l(l+1)R] + \sum_{l_1,l_2}C_{l_1}^{\phi\phi}C_{l_2}^{TT}\frac{F_{ll_1l_2}^2}{2l+1}, \tag{25}$$

where

$$R = \frac{1}{8\pi}\sum_l(2l+1)l(l+1)C_l^{\phi\phi} \tag{26}$$

is half the mean-squared deflection. Therefore,

$$\frac{\partial C_l^{\tilde T\tilde T}}{\partial C_L^{\phi\phi}} = \sum_{l'}C_{l'}^{TT}\frac{F_{lLl'}^2}{2l+1} - \frac{L(L+1)(2L+1)}{8\pi}l(l+1)C_l^{TT}. \tag{27}$$

While Ref. [40] included higher-order corrections to this expression by taking numerical derivatives of lensed power spectra computed nonperturbatively with the CAMB

[4]The $\mathcal{O}(\phi^4)$ part of Eq. (24) was not derived in a rigorous perturbative analysis, which would imply additional corrections. For example, the unlensed $f_{lLl'}$ in the second term on the right could be replaced by its lensed counterparts, as in Eq. (14). We do not investigate such corrections to the temperature power autocovariance further because corrections to the leading Gaussian term are negligible for all applications in this paper.

code [41,42], these corrections are not expected to be important for our purposes here.

The off-diagonal contributions to the correlation between the empirical $\hat{C}^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}}$ are shown in Fig. 3 (see Refs. [39,40] for similar plots). The checkerboard structure of the $\mathcal{O}[(C^{\phi\phi})^2]$ contribution arises because fluctuations in the lensing power produce changes in the lensed temperature spectra of opposite signs at the acoustic peaks and troughs. Both of the corrections in Eq. (24) give correlations that are at most of order 10^{-4} and are rather localized in the (l, l') plane. (Note that the correlations are suppressed on small scales where noise dominates the diagonal variance.) The impact of these nondiagonal contributions is found to be negligible for all calculations in this paper; i.e., we can assume a Gaussian diagonal autocovariance of the temperature power spectrum.

B. Lensing reconstruction

The auto-correlation of the lensing reconstruction power spectrum involves the 8-point function of the lensed temperature. Hanson *et al.* [26] found that the dominant off-diagonal contributions on the full sky are given by disconnected terms that contribute as (see Ref. [32] for a similar calculation on the flat sky)

$$\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\hat{\phi}\hat{\phi}}_{L'})^{\mathrm{dom}}_{\mathrm{non\text{-}diag}} = \frac{32A_L^2A_{L'}^2}{(2L+1)(2L'+1)}\sum_{l_1}\frac{1}{2l_1+1}(C^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}})^2 \times\left[\sum_{l_2}\tilde{g}^2_{l_1l_2}(L)C^{\tilde{T}\tilde{T}}_{l_2,\mathrm{expt}}\right]\left[\sum_{l_3}\tilde{g}^2_{l_1l_3}(L')C^{\tilde{T}\tilde{T}}_{l_3,\mathrm{expt}}\right]. \quad (28)$$

This dominates over more tightly coupled terms that involve products of four weights $\tilde{g}$ that do not factor in the form $\tilde{g}^2\tilde{g}^2$ and therefore enforce a reduced summation volume. The variance ($L = L'$) on the full sky is predominantly

$$\mathrm{var}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}) = \frac{2}{2L+1}\langle C^{\hat{\phi}\hat{\phi}}_{L}\rangle^2, \quad (29)$$

with small corrections from Eq. (28) (for $L = L'$).

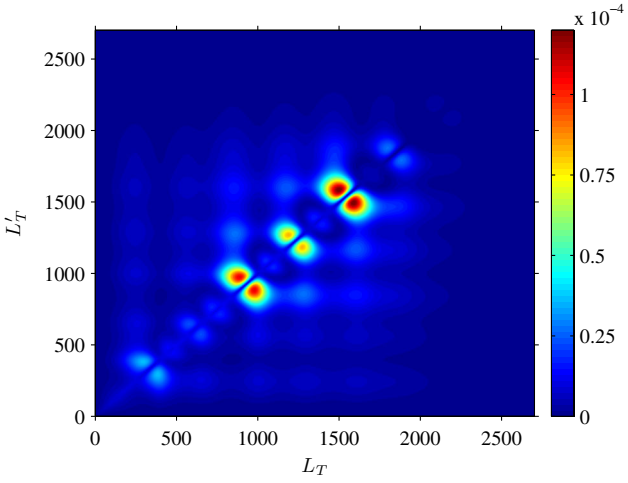
As shown in Ref. [26] the off-diagonal reconstruction power correlation can reach a level of 0.5% and is rather broadband. If the reconstructed power is binned, this can induce correlations of $\mathcal{O}(10\%)$ between different bins. Physically, these broadband correlations arise because cosmic-variance fluctuations in the CMB at a given scale produce fluctuations in $\hat{C}^{\hat{\phi}\hat{\phi}}$ that are coherent over a broad range of scales due to the mode-coupling nature of lens reconstruction (small-scale CMB fluctuations are used to reconstruct large-scale lenses). To make this physical interpretation more explicit, we note that the dominant nondiagonal covariance contribution of Eq. (28) can be written as

$$\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\hat{\phi}\hat{\phi}}_{L'})^{\mathrm{dom}}_{\mathrm{non\text{-}diag}} = \sum_{l_1}\frac{\partial(2\hat{N}^{(0)}_{L})}{\partial\hat{C}^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}}}\frac{2}{2l_1+1}(C^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}})^2\frac{\partial(2\hat{N}^{(0)}_{L'})}{\partial\hat{C}^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}}}, \quad (30)$$

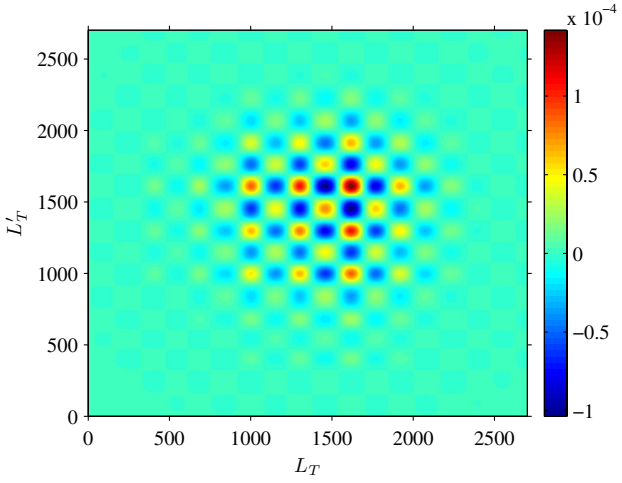
where the (realization-independent) derivative is given by

$$\frac{\partial(2\hat{N}^{(0)}_{L})}{\partial\hat{C}^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}}} = \frac{4A_L^2}{2L+1}\sum_{l_2}\tilde{g}^2_{l_1l_2}(L)C^{\tilde{T}\tilde{T}}_{l_2,\mathrm{expt}}, \quad (31)$$

which is nonzero even in the absence of lensing. Equation (31) describes the change of the Gaussian



(a) $\mathcal{O}(C^{\phi\phi})$ contribution from [39]



(b) $\mathcal{O}[(C^{\phi\phi})^2]$ contribution from [40]

FIG. 3 (color online). (a) Theoretical off-diagonal part of the correlation correl$(\hat{C}^{\tilde{T}\tilde{T}}_{L_T}, \hat{C}^{\tilde{T}\tilde{T}}_{L'_T})$ to first order in $C^{\phi\phi}$, given by the second term on the right of Eq. (24) (see Ref. [39]). The covariance is converted to a correlation by dividing by $[\mathrm{var}_G(C^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}})\mathrm{var}_G(C^{\tilde{T}\tilde{T}}_{L'_T,\mathrm{expt}})]^{1/2}$. (b) $\mathcal{O}[(C^{\phi\phi})^2]$ contribution from Ref. [40].

reconstruction noise[5] resulting from fluctuations of the observed temperature realization. In propagating these changes through to the covariance of $\hat{C}^{\hat{\phi}\hat{\phi}}$, one picks up the sample variance of the total lensed temperature power spectrum, $\mathrm{var}_G(C^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}})$.

As noted in Sec. II, using the realization-dependent $\hat{N}^{(0)}$ bias correction of Eq. (17) significantly reduces the off-diagonal covariance of the reconstruction power spectrum. To help interpret the $\hat{N}^{(0)}$ correction, we write it in the form

$$\hat{C}^{\hat{\phi}\hat{\phi}}_L - 2\hat{N}^{(0)}_L + N^{(0)}_L = \hat{C}^{\hat{\phi}\hat{\phi}}_L - \sum_l \frac{\partial(2\hat{N}^{(0)}_L)}{\partial \hat{C}^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}}} \hat{C}^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}} + N^{(0)}_L. \tag{32}$$

Therefore, for a given realization of the lensed temperature, the empirical $\hat{N}^{(0)}$ bias correction partly removes the response of the Gaussian reconstruction noise to changes in the lensed temperature realization. To see that this removes the nondiagonal power covariance of the lensing reconstruction caused by cosmic variance of the lensed temperature, note that both $\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_L, 2\hat{N}^{(0)}_{L'})$ and $\mathrm{cov}(2\hat{N}^{(0)}_L, 2\hat{N}^{(0)}_{L'})$ equal the right-hand side of Eq. (30) at $\mathcal{O}(\phi^0)$. The empirical $\hat{N}^{(0)}$ correction leads to a small reduction in the *variance* of the binned reconstructed power spectrum, as shown in Figs. 1 and 2. Any residual covariance after empirical $\hat{N}^{(0)}$ subtraction is too small to be detected in our simulations.

V. TEMPERATURE-LENSING CROSS-CORRELATION

For a joint analysis involving the empirical power spectra of the lens reconstruction and the lensed temperature anisotropies, the likelihood should model their cross-correlation to avoid potential double counting of lensing effects. In this section we calculate the cross-correlation. We recover the two main physical effects introduced in Sec. I, i.e., a "noise contribution" from the cosmic variance of the lensed temperature affecting the noise in the reconstruction over a wide range of scales and a "matter cosmic variance" contribution from cosmic variance of the lenses altering the smoothing of the acoustic peaks in the temperature power spectrum. We will show in Sec. V A that the noise contribution is due to the disconnected part of the lensed temperature 6-point function, while in Sec. V B we show that the matter cosmic variance contribution is due to the connected part of the 6-point function. Corrections from the lensed temperature trispectrum generally have a subdominant effect on parameter estimation and are discussed in Appendix D (see also Fig. 8 below).

A. Noise contribution

1. Perturbative derivation: Disconnected part of lensed temperature 4- and 6-point functions

Since the reconstructed lensing potential is quadratic in the lensed temperature, the covariance $\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_L, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})$ involves the 4- and 6-point functions of the lensed temperature. Using the definition of $\hat{\phi}$ in Eq. (9), we find

$$\begin{aligned}
&\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_L, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}) \\
&= \frac{A_L^2}{(2L+1)(2L'+1)} \sum_{\underline{l}_1,\underline{l}_2,\underline{l}_3,\underline{l}_4,M,M'} (-1)^{M+M'} \\
&\times \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \begin{pmatrix} l_3 & l_4 & L \\ m_3 & m_4 & M \end{pmatrix} \\
&\times \tilde{g}_{l_1 l_2}(L)\tilde{g}_{l_3 l_4}(L)[\langle \tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}\tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle \\
&- \langle \tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}\rangle\langle \tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle].
\end{aligned} \tag{33}$$

Since all connected terms vanish in the absence of lensing, we expect the noise contribution to come from the fully disconnected part. If we only keep disconnected terms, the second line of Eq. (33) can be replaced by

$$\begin{aligned}
&[\langle \tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}\tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle \\
&- \langle \tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}\rangle\langle \tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle] \\
&\to 4[\langle \tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_3}\rangle\langle \tilde{T}_{\underline{l}_2}\tilde{T}_{L'M'}\rangle\langle \tilde{T}_{\underline{l}_4}\tilde{T}_{L',-M'}\rangle + (M' \leftrightarrow -M')],
\end{aligned} \tag{34}$$

where we exploited symmetry under relabeling $(\underline{l}_1 \leftrightarrow \underline{l}_2)$ and/or $(\underline{l}_3 \leftrightarrow \underline{l}_4)$. We also used that the contractions $\langle \tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\rangle$ and $\langle \tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}\rangle$ do not contribute because $\langle \hat{\phi}_{LM}\rangle = 0$. The fully disconnected part of Eq. (33) is therefore

$$\begin{aligned}
&\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_L, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})_{\mathrm{disconn}} \\
&= \frac{8A_L^2}{(2L+1)(2L'+1)} (C^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})^2 \sum_{l_1} \tilde{g}^2_{l_1 L'}(L) C^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}},
\end{aligned} \tag{35}$$

where the weight $\tilde{g}_{l_1 L'}(L)$ enforces $l_1 + L + L'$ to be even.

To interpret this result, note that it can be expressed in terms of the derivative in Eq. (31) as

$$\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_L, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})_{\mathrm{disconn}} = \frac{\partial(2\hat{N}^{(0)}_L)}{\partial \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}} \frac{2}{2L'+1} (C^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})^2. \tag{36}$$

This part of the covariance is therefore due to the response of the Gaussian reconstruction noise to changes in the observed temperature realization and the resulting covariance with the observed temperature power. Based on this

[5]When squaring the reconstruction $\hat{\phi}_{LM}$ to form the reconstruction power spectrum, we pick up not only the signal power $C^{\phi\phi}_L$ but also the noise of the reconstruction. Further details on the correspondence between noise terms in the reconstruction power and the $2\hat{N}^{(0)}$ expression are provided in Appendix C.

intuition, we anticipate that the covariance can be mitigated by the realization-dependent $\hat{N}^{(0)}$ correction of the reconstruction power bias (see Sec. VA 4 for confirmation).

2. Magnitude and structure of the correlation matrix

In Fig. 4(a) we plot the power correlation resulting from the power covariance in Eq. (35) (denoting $L_\phi = L$ and $L_T = L'$ for convenience),

$$\mathrm{correl}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L_\phi}, \hat{C}^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}}) = \frac{\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L_\phi}, \hat{C}^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}})}{\sqrt{\mathrm{var}_G(C^{\phi\phi}_{L_\phi} + N^{(0)}_{L_\phi} + N^{(1)}_{L_\phi})\mathrm{var}_G(C^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}})}}. \quad (37)$$

We plot the correlation of the unbinned spectra. Note that if the covariance is broadband (i.e., roughly constant over the bin width) the *correlation* of (sufficiently finely) binned power spectra will increase roughly proportionally to the square root of the product of the two bin widths. The Gaussian variance of $C^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}}$ in the denominator of Eq. (37) contains the beam-deconvolved noisy temperature power spectrum (18), so that high temperature multipoles are suppressed.

The unbinned power correlation shown in Fig. 4(a) is mostly constrained to a conelike region in the L_ϕ vs L_T plane, with the maximum correlation of 0.5% located at the first acoustic peak $L_T \sim 200$ and lensing reconstruction multipoles $1600 \lesssim L_\phi \lesssim 1900$. To understand the basic structure of the correlation in Fig. 4(a) we compute approximations to Eq. (35) in the limits $L_T \ll L_\phi$ and $L_\phi \ll L_T$.

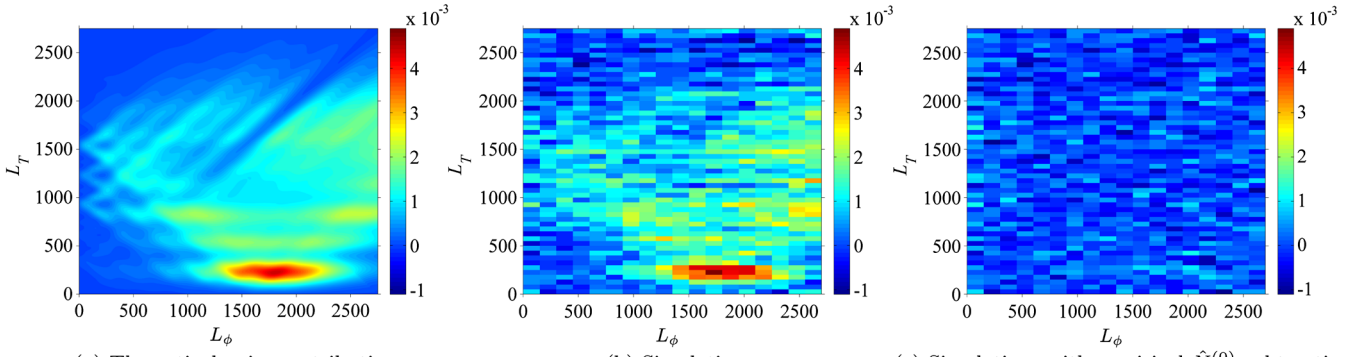
For $L_T \ll L_\phi$, the weights in Eq. (35) restrict the summation from $l_1 = L_\phi - L_T$ to $l_1 = L_\phi + L_T$. If we Taylor expand in l_1 around L_ϕ, we get

$$\mathrm{correl}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L_\phi}, \hat{C}^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}})^{L_T \ll L_\phi}_{\mathrm{disconn}} \approx \frac{[L_\phi(L_\phi+1)]^2 A_{L_\phi}}{8\pi}\sqrt{(2L_T+1)(2L_\phi+1)} \times \frac{L_T(L_T+1)C^{\tilde{T}\tilde{T}}_{L_T}}{L_\phi(L_\phi+1)C^{\tilde{T}\tilde{T}}_{L_\phi,\mathrm{expt}}}. \quad (38)$$

Recalling that $A_{L_\phi} = N^{(0)}_{L_\phi}$ for optimal weights, we see from Fig. 1 that the first term slightly increases with L_ϕ. The last term is maximized at the first acoustic peak $L_T \sim 200$ and at the reconstruction multipole $1600 \lesssim L_\phi \lesssim 1900$ where the observed temperature power $C^{\tilde{T}\tilde{T}}_{L_\phi,\mathrm{expt}}$ is minimal (for the Planck-like noise and beam considered here). In this region Eq. (38) gives a correlation of around 0.4–0.5%, which agrees with Fig. 4(a). Equation (38) also implies that lower noise in the temperature power spectrum would move the peak position to higher reconstruction multipoles L_ϕ. The cone structure in Fig. 4(a), with apex at $(L_\phi, L_T) \sim (1600\text{–}1900, 200)$ and edges $L_\phi - L_T \lesssim (1600\text{–}1900) \lesssim L_\phi + L_T$, encloses the region for which the sum over l_1 includes the maximum of $1/C^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}}$ around $l_1 = 1600\text{–}1900$. A similar argument can be applied for the cone patterns in the $L_\phi \lesssim 200$ region.

For high temperature and low reconstruction multipoles, $L_\phi \ll L_T$, we can Taylor expand Eq. (35) in l_1 around L_T (see Ref. [26] for a similar calculation):

$$\mathrm{correl}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L_\phi}, \hat{C}^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}})^{L_\phi \ll L_T}_{\mathrm{disconn}} \approx \frac{[L_\phi(L_\phi+1)]^2 A^2_{L_\phi}}{2\pi(A_{L_\phi} + C^{\phi\phi}_{L_\phi})}\left(\frac{C^{\tilde{T}\tilde{T}}_{L_T}}{C^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}}}\right)^2\sqrt{\frac{(2L_\phi+1)(2L_T+1)}{4}} \times \frac{1}{4}\left[\left(\frac{\mathrm{d}\ln(L_T^2 C^{\tilde{T}\tilde{T}}_{L_T})}{\mathrm{d}\ln L_T}\right)^2 + \frac{1}{2}\left(\frac{\mathrm{d}\ln C^{\tilde{T}\tilde{T}}_{L_T}}{\mathrm{d}\ln L_T}\right)^2\right], \quad (39)$$



(a) Theoretical noise contribution (b) Simulations (c) Simulations with empirical $\hat{N}^{(0)}$ subtraction

FIG. 4 (color online). (a) Theoretical noise contribution from Eq. (35) to the correlation of unbinned power spectra of the lensed temperature and the reconstructed lensing potential, $\mathrm{correl}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L_\phi}, \hat{C}^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}})$. The acoustic peaks of the temperature power spectrum are visible in the vertical direction. (b) Estimate of the correlation of unbinned power spectra from 1000 simulations. (c) Same as (b) after subtracting the empirical $\hat{N}^{(0)}$ bias of Eq. (17) from $\hat{C}^{\hat{\phi}\hat{\phi}}$.

where we have neglected $N^{(1)}_{L_\phi}$ and used $A_{L_\phi} = N^{(0)}_{L_\phi}$. The quadrature sum of derivatives in the last term is maximal between acoustic peaks and troughs at $L_T \sim 350, 625, 925, 1225, 1550, 1850$, etc., which agrees with the temperature multipoles where the full correlation shown in Fig. 4(a) is maximal (for $L_\phi \lesssim 200$).[6] The maximum value of the second line of Eq. (39) is around 36 (for $L_T \lesssim 2700$). If we neglect $C^{\phi\phi}_{L_\phi}$ compared to A_{L_ϕ}, which is roughly acceptable for $L_\phi \lesssim 10$, then the first term in Eq. (39) is around 2×10^{-7} (see Fig. 1); i.e., for $L_\phi \lesssim 10$,

$$\mathrm{correl}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L_\phi}, \hat{C}^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}})^{L_\phi \ll L_T}_{\mathrm{disconn}} \lesssim 2 \times 10^{-7} \times \frac{1}{2}\sqrt{(2L_\phi + 1)(2L_T + 1)} \times 36. \quad (40)$$

For example, for $L_\phi = 10$, $L_T = 1500$ the bound is 9×10^{-4} which is consistent with the full result shown in Fig. 4(a).

3. Comparison with simulations

Before assessing the relevance of the noise contribution to the covariance for parameter estimation, we compare the analytic result in Eq. (35) with the full covariance estimated from our simulations. For the latter we use

$$\widehat{\mathrm{cov}}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}) = \frac{1}{N_{\mathrm{sims}} - 1}\sum_{s=1}^{N_{\mathrm{sims}}} (\hat{C}^{\hat{\phi}\hat{\phi}}_{L,s} - \langle \hat{C}^{\hat{\phi}\hat{\phi}}_{L}\rangle_{\mathrm{sims}}) \times (\hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt},s} - \langle \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}\rangle_{\mathrm{sims}}), \quad (41)$$

where s labels different realizations and $\langle\cdot\rangle_{\mathrm{sims}}$ denotes the average over N_{sims} realizations. To reduce the noise of the estimates from the finite number of simulations, we average the measured covariance over a range of L and L' values,[7]

$$\widehat{\mathrm{cov}}(\hat{C}^{\hat{\phi}\hat{\phi}}_{\bar{L}_i}, \hat{C}^{\tilde{T}\tilde{T}}_{\bar{L}'_j,\mathrm{expt}}) = \frac{1}{\Delta L_i \Delta L'_j}\frac{1}{[\bar{L}_i(\bar{L}_i+1)]^2 \bar{L}'_j(\bar{L}'_j+1)}\sum_{L=L_i}^{L_{i+1}-1}\sum_{L'=L'_j}^{L'_{j+1}-1} \times [L(L+1)]^2 L'(L'+1)\widehat{\mathrm{cov}}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\tilde{T}\tilde{T}}_{L'}), \quad (42)$$

where L_i and L'_j are bin boundaries for lensing and temperature powers, respectively, and $\bar{L}_i$ and $\bar{L}'_j$ denote the corresponding bin centres. The bin widths are $\Delta L_i = L_{i+1} - L_i$ and $\Delta L'_j = L'_{j+1} - L'_j$. We divide out $L(L+1)$ prefactors to average over relatively slowly varying quantities. In Fig. 4(b) we plot the estimate of the correlation that *unbinned* power spectra would have. This is obtained by dividing the covariance estimate of Eq. (42) by the theoretical Gaussian variance of unbinned power spectra as in Eq. (37) [evaluated at $(\bar{L}_i, \bar{L}'_j)$]. Within the random scatter from the finite N_{sims}, the estimated correlation agrees with the theoretical noise contribution of Eq. (35). We can also conclude that the noise contribution is the dominant part of the temperature-lensing power correlation.

4. Mitigating the noise contribution with the empirical $\hat{N}^{(0)}$ bias correction

In practice, it is desirable that the temperature and reconstruction power spectra are uncorrelated so that their respective likelihoods may be simply combined. Indeed, this is the assumption that has been made in all joint analyses to date [12,13,19].

Fortunately, the empirical $\hat{N}^{(0)}$ subtraction, which was originally proposed in Ref. [26] to eliminate the nondiagonal reconstruction power autocovariance (28), also removes the noise contribution [Eq. (35)] to the temperature-lensing power cross-covariance. To see this note that if the empirical bias correction of Eq. (17) is used, the cross-covariance changes by

$$-\mathrm{cov}(2\hat{N}^{(0)}_L, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}) = -\sum_l \frac{\partial(2\hat{N}^{(0)}_L)}{\partial \hat{C}^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}}}\mathrm{cov}(\hat{C}^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}) = -\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})_{\mathrm{disconn}} + \mathcal{O}(\phi^2). \quad (43)$$

We establish a more general version of this result in Appendix B, where we show that the generalization of the $\hat{N}^{(0)}$ correction for anisotropic surveys removes the noise contribution to the covariance with any quadratic estimate (including, e.g., cross-spectra) of the temperature power spectrum. We confirm the reduction in the cross-covariance with simulations in Fig. 4(c). Corrections to Eq. (43) from the non-Gaussian terms in the covariance of the temperature power spectra [see Eq. (24)] at $\mathcal{O}(\phi^2)$ and $\mathcal{O}(\phi^4)$ reach at most 3×10^{-5} in the correlation, which is roughly 1 order of magnitude smaller than the matter cosmic variance contribution discussed below (see Fig. 5(a)). These corrections are too small to be visible in Fig. 4(c) and we neglect them in the following.

B. Matter cosmic variance contribution

1. Warmup: Power covariance of input lensing potential and lensed temperature

We expect the cosmic variance of the lenses to induce a power correlation of the lensed temperature with the

[6]The correlation at high L_T is suppressed by the noise in $C^{\tilde{T}\tilde{T}}_{L_T,\mathrm{expt}}$, which is used to normalize the covariance in Eq. (37).

[7]Note that for broadband covariances, this binning procedure does not bias the covariance estimate; i.e., the binned covariance agrees with the unbinned covariance in the limit of averaging over infinitely many simulations. However, for a finite number of simulations, the binned covariance is less noisy than the covariance estimate at a single (L, L') pair. Note also that binning the estimated covariance is equivalent to estimating the covariance of the binned spectra.

lensing reconstruction since greater lensing power in a given realization leads to additional smoothing of the empirical temperature power spectrum. As a warmup, we calculate how the same effect gives rise to a covariance between the power spectrum of the temperature and the (unobservable) power spectrum of the lensing potential, as if we were able to measure ϕ directly with no noise. This correlation can be extracted from simulations simply by measuring the correlation of the power spectra of the lensed temperature and the input lensing potential without performing any lensing reconstruction. To calculate the covariance perturbatively, note that for a fixed realization of the input lensing potential $\phi_{\rm in}$, the lensed temperature power spectrum obtained by averaging only over the unlensed CMB is given by ($L' \neq 0$),

$$\begin{aligned}
&\frac{1}{2L'+1}\sum_{M'}\langle|\tilde{T}_{L'M'}|^2\rangle_{\rm CMB} \\
&\quad = C_{L'}^{TT} + \frac{1}{2L'+1}\sum_{l_1 l_2}\hat{C}_{l_1}^{\phi_{\rm in}\phi_{\rm in}} C_{l_2}^{TT} F^2_{L'l_1l_2} - \hat{R}L'(L'+1)C_{L'}^{TT} \\
&\quad = C_{L'}^{TT} + \sum_{L''}\frac{\partial C_{L'}^{\tilde{T}\tilde{T}}}{\partial C_{L''}^{\phi\phi}}\hat{C}_{L''}^{\phi_{\rm in}\phi_{\rm in}}. \qquad (44)
\end{aligned}$$

Here $\hat{C}^{\phi_{\rm in}\phi_{\rm in}}$ denotes the empirical power spectrum of the input lensing potential realization $\phi_{\rm in}$,[8] and $\hat{R}$ is defined by replacing $C^{\phi\phi}$ in Eq. (26) by $\hat{C}^{\phi_{\rm in}\phi_{\rm in}}$. The first line of Eq. (44) can be derived from the expansion in Eq. (2) at second order in $\phi_{\rm in}$. The term linear in $\phi_{\rm in}$ vanishes because it is proportional to the monopole of $\phi_{\rm in}$. In Eq. (44) we neglected noise in the temperature measurement since this does not contribute to the covariance that we aim to calculate. The derivative in the second line of Eq. (44) is given by Eq. (27). Neglecting the T-ϕ correlation, we then have

$$\begin{aligned}
&\mathrm{cov}(\hat{C}_L^{\phi_{\rm in}\phi_{\rm in}}, \hat{C}_{L',\rm expt}^{\tilde{T}\tilde{T}}) \\
&\quad = \langle\hat{C}_L^{\phi_{\rm in}\phi_{\rm in}}\langle\hat{C}_{L',\rm expt}^{\tilde{T}\tilde{T}}\rangle_{\rm CMB}\rangle_{\rm LSS} - \langle\hat{C}_L^{\phi_{\rm in}\phi_{\rm in}}\rangle\langle\hat{C}_{L',\rm expt}^{\tilde{T}\tilde{T}}\rangle, \\
&\quad = \frac{2}{2L+1}(C_L^{\phi\phi})^2\frac{\partial C_{L'}^{\tilde{T}\tilde{T}}}{\partial C_L^{\phi\phi}}, \qquad (45)
\end{aligned}$$

where $\langle\cdot\rangle_{\rm LSS}$ denotes an average over realizations of ϕ (i.e., over large-scale structure).

2. Power covariance of reconstructed lenses and lensed temperature from the connected 6-point function

The power covariance [Eq. (33)] of the *reconstructed* lensing potential and the lensed temperature receives contributions from the connected 4- and 6-point functions, as well as the Gaussian (disconnected) part. The leading-order trispectrum is linear in $C^{\phi\phi}$ and so cannot give rise to the expected matter cosmic variance contribution calculated above. The trispectrum contributions are discussed in detail in Appendix D and are shown to have a subdominant effect on parameter estimation (see Fig. 8 below). We therefore focus here on the contribution from the connected 6-point function. Terms independent of ϕ do not contribute to the connected part, and terms linear in ϕ vanish when averaging over large-scale structure. Since the ϕ^2 contribution can be shown to vanish [43] as well as ϕ^3 terms which vanish for Gaussian ϕ, the leading-order contribution is of order ϕ^4. The $\mathcal{O}(\phi^4)$ contribution from the connected 6-point function to the temperature-lensing power covariance is calculated in detail in Appendix E. The dominant contribution there [Eq. (E8)] is exactly of the form calculated above [Eq. (45)] for the power covariance of the input lensing potential and the lensed temperature.

3. Magnitude and structure of the covariance matrix

The correlation of unbinned temperature and lensing power spectra due to Eq. (E8) is at most 0.04% (see Fig. 5(a)). Since higher lensing power lowers the acoustic peaks and increases the troughs of the temperature power spectrum, the correlation is negative at multipoles L_T where the temperature power peaks and positive where it has a trough. The correlation is small for $L_\phi \gtrsim 150$ because the acoustic peak smearing is mainly caused by large-scale lenses [38].

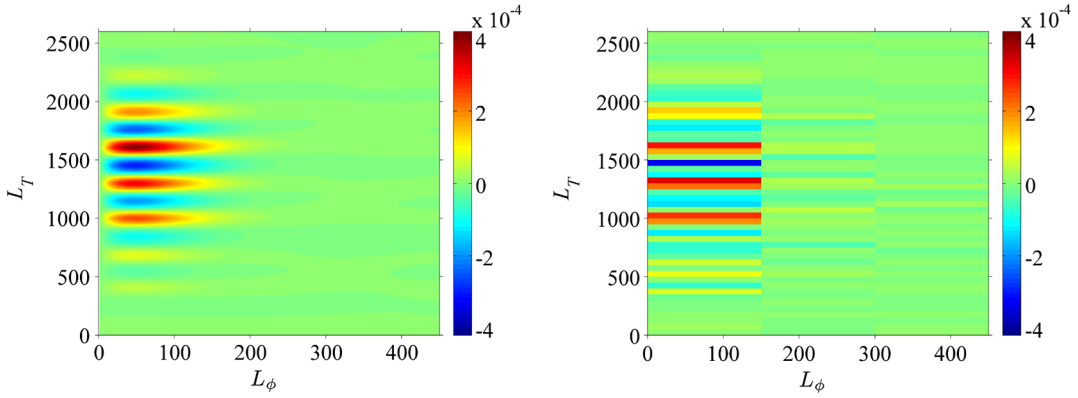
The approximate covariance matrix between $[L(L+1)]^2\hat{C}_L^{\hat{\phi}\hat{\phi}}/(2\pi)$ and $L'(L'+1)\hat{C}_{L'}^{\tilde{T}\tilde{T}}/(2\pi)$, calculated from Eq. (E8), is shown in the left-hand panel of Fig. 6. By performing a singular-value decomposition of this matrix, we find that it has a very low-rank structure. For example, retaining the multipoles of the reconstruction to $L = 500$ (and up to $L' = 2500$ for the temperature), the first singular value is 34 times larger than the second. The covariance matrix can be accurately approximated in rank-one form,

$$\mathrm{cov}\left(\frac{[L(L+1)]^2\hat{C}_L^{\hat{\phi}\hat{\phi}}}{2\pi}, \frac{L'(L'+1)\hat{C}_{L'}^{\tilde{T}\tilde{T}}}{2\pi}\right)_{\rm (E8)} \approx \lambda_1 u_L v_{L'}, \qquad (46)$$

as shown in the right-hand panel of Fig. 6. Here, the leading singular value $\lambda_1 = 8.3\times10^{-8}\ \mu\mathrm{K}^2$; the corresponding (normalized) singular vectors u_L (for the reconstructed lensing power) and $v_{L'}$ (for the temperature power) are plotted in Fig. 7. The significance of the rank-one covariance is that cosmic variance of the lenses produces correlated changes in the measured lensed temperature power spectrum with definite shape given by $v_{L'}$. Moreover, cosmic-variance fluctuations in $\hat{C}_L^{\phi\phi}$ that are orthogonal to u_L do not influence $\hat{C}_{L'}^{\tilde{T}\tilde{T}}$.

We can understand the low-rank stucture of the covariance by evaluating the derivative $\partial C_{L'}^{\tilde{T}\tilde{T}}/\partial C_L^{\phi\phi}$ with the flat-sky form of Eq. (25). Following Ref. [28] we have

[8] We denote this as $\phi_{\rm in}$, while the reconstructed lensing potential is $\hat{\phi}$. Recall that C denotes theoretical and $\hat{C}$ empirical power spectra.



(a) Theoretical matter cosmic variance contribution (b) Simulations

FIG. 5 (color online). (a) Theoretical matter cosmic variance contribution, from Eq. (E8), to the correlation of the unbinned power spectra of the reconstructed lensing potential and the lensed temperature anisotropies. The covariance is converted to a correlation following Eq. (37). (b) Measured correlation $\widehat{\text{correl}}(\hat{C}^{\phi_{\rm in}\phi_{\rm in}}_{L_\phi}, \hat{C}^{\tilde{T}\tilde{T}}_{L_T} - \hat{C}^{TT}_{L_T})$ of the power spectrum of the input lensing potential and the difference of noise-free lensed and unlensed temperature power spectra from 1000 simulations.

$$C^{\tilde{T}\tilde{T}}_{L'} = (1 - L'^2 R) C^{TT}_{L'} + \int \frac{d^2\mathbf{L}}{(2\pi)^2} [(\mathbf{L}' - \mathbf{L}) \cdot \mathbf{L}]^2 C^{\phi\phi}_L C^{TT}_{|\mathbf{L}'-\mathbf{L}|}, \quad (47)$$

where R is given by Eq. (26). It is clear from Fig. 6 that the covariance is only significant between large-scale lens modes ($L < 100$) and intermediate- and small-scale temperature modes (where lensing has a significant effect on the power spectrum). In the limit $L \ll L'$, and for L small compared to the CMB acoustic scale, we can Taylor expand $C^{TT}_{|\mathbf{L}'-\mathbf{L}|}$ in Eq. (47) to obtain

$$\frac{\partial C^{\tilde{T}\tilde{T}}_{L'}}{\partial C^{\phi\phi}_L} \approx \frac{L^5}{16 L'^2} \left[5L' \frac{d}{dL'} \left(\frac{L'(L'+1) C^{TT}_{L'}}{2\pi} \right) + 3L'^2 \frac{d^2}{dL'^2} \left(\frac{L'(L'+1) C^{TT}_{L'}}{2\pi} \right) \right] \quad (48)$$

at leading order in L/L'. Since this is a rank-one matrix, the same is true of the covariance between $\hat{C}^{\hat{\phi}\hat{\phi}}_L$ and $\hat{C}^{\tilde{T}\tilde{T}}_{L'}$. We can read off that the singular vector $v_{L'}$ in Eq. (46) is

$$v_{L'} \propto 5L' \frac{d}{dL'} \left(\frac{L'(L'+1) C^{TT}_{L'}}{2\pi} \right) + 3L'^2 \frac{d^2}{dL'^2} \left(\frac{L'(L'+1) C^{TT}_{L'}}{2\pi} \right), \quad (49)$$

which agrees well with that determined directly by singular-value decomposition (see Fig. 7). This singular vector is also similar to the change in C^{TT}_L due to lensing over the acoustic part of the spectrum, where the dominant contribution is from large-scale lenses. Finally, we note that the rank-one structure of the derivative $\partial C^{\tilde{T}\tilde{T}}_{L'} / \partial C^{\phi\phi}_L$ is consistent with the findings of Ref. [38] where it is shown that the lensed temperature power spectrum is sensitive to essentially a single mode of $C^{\phi\phi}_L$.

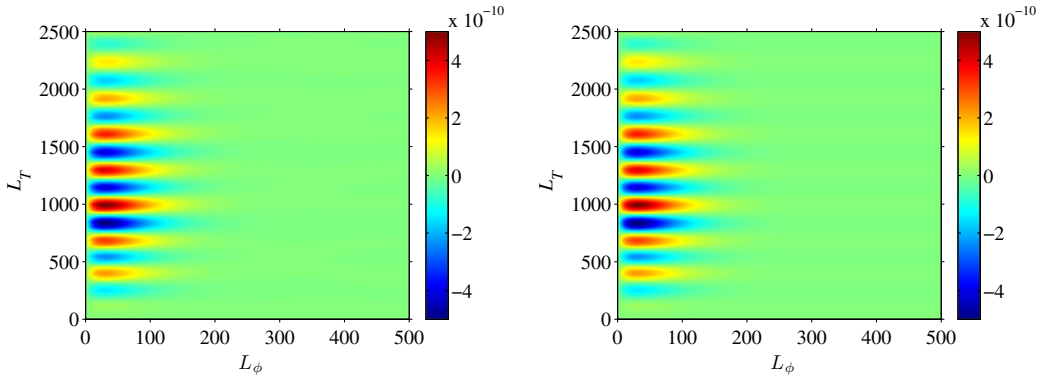


FIG. 6 (color online). *Left*: The approximate contribution to the *covariance* between $[L_\phi(L_\phi+1)]^2 \hat{C}^{\hat{\phi}\hat{\phi}}_{L_\phi}/(2\pi)$ and $L_T(L_T+1)\hat{C}^{\tilde{T}\tilde{T}}_{L_T}/(2\pi)$ from cosmic variance of the lenses, derived from Eq. (E8). *Right*: The rank-one approximation to the matrix on the left from retaining only the largest singular value.

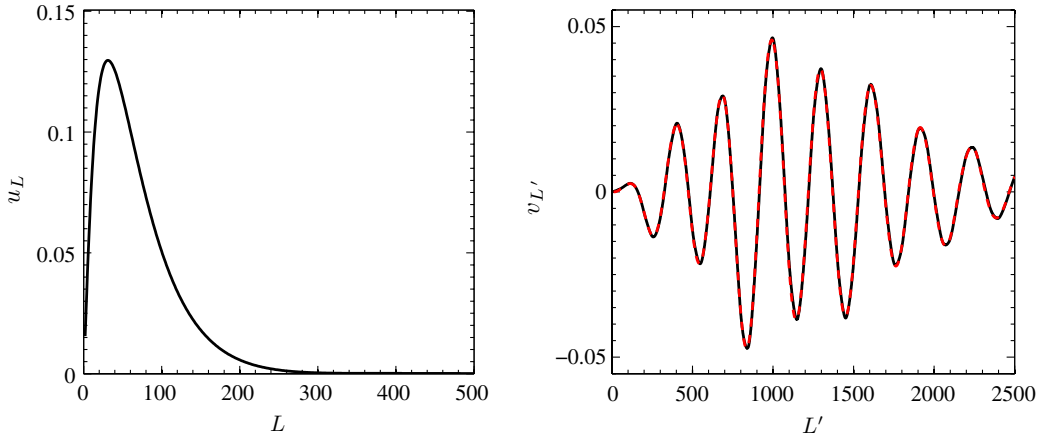


FIG. 7 (color online). Singular vectors u_L (left) and $v_{L'}$ (right, black) of the matter cosmic variance contribution to the covariance between $[L(L+1)]^2\hat{C}_L^{\hat{\phi}\hat{\phi}}/(2\pi)$ and $L'(L'+1)\hat{C}_{L'}^{\tilde{T}\tilde{T}}/(2\pi)$ corresponding to the dominant singular value; see Eq. (46). The approximation of Eq. (49) is shown in red dashed lines on the right.

4. Comparison with simulations

Before assessing the importance for parameter estimation of the matter cosmic variance contribution to the cross-covariance, we validate it against simulations. The measurements of the temperature-lensing power correlation in Fig. 4(b) are too noisy to resolve clearly the matter cosmic variance contribution. However, we can test Eq. (45) by correlating the lensed temperature with the empirical power spectrum $\hat{C}^{\phi_{\rm in}\phi_{\rm in}}$ of the realization of the input lensing potential without performing any reconstruction. To reduce the noise of the covariance estimate, we work with the empirical power of the lensed temperature without including beam effects or noise, which do not affect the matter cosmic variance effect we are looking for. Since we have so far neglected T-ϕ correlations in all calculations, we try to eliminate correlations between the lensing potential and the unlensed temperature in the simulations by calculating $\widehat{\rm cov}(\hat{C}_L^{\phi_{\rm in}\phi_{\rm in}}, \hat{C}_{L'}^{\tilde{T}\tilde{T}} - \hat{C}_{L'}^{TT})$. Here, $\hat{C}^{\phi_{\rm in}\phi_{\rm in}}$ is the empirical power of the input lensing potential and $\hat{C}^{TT}$ is the empirical power of the unlensed temperature. Additionally, subtracting the unlensed from the lensed empirical temperature power reduces the noise of the covariance estimate because it eliminates the scatter due to cosmic variance of the unlensed temperature. Otherwise, our estimate of the covariance follows the procedure described in Sec. V A 3. As shown in Fig. 5(b), these broadband estimates are consistent with the theoretical expectation from Eq. (45).

5. Mitigating the matter cosmic variance contribution

As we shall show in Sec. V D, the impact on parameter errors of ignoring the covariance between the lensed temperature and reconstruction power due to cosmic variance of the lenses is small for an experiment like Planck. Indeed, this is why the covariance is not accounted for in the current Planck likelihood [13]. However, the covariance is simple to model using Eq. (E8) and could easily be included in a joint analysis. Equivalently, the covariance could be diagonalized by appropriate modifications of the empiricial temperature or lensing power spectra. A symmetric way to do this, mirroring the empirical $N^{(0)}$ correction that we advocate applying to the empirical lensing reconstruction power, is to modify the measured temperature power spectrum by

$$\hat{C}_{l,\rm expt}^{\tilde{T}\tilde{T}} \to \hat{C}_{l,\rm expt}^{\tilde{T}\tilde{T}} - \sum_{l'} A_{ll'}(\hat{C}_{l'}^{\hat{\phi}\hat{\phi}} - 2\hat{N}_{l'}^{(0)}) + \left\langle \sum_{l'} A_{ll'}(\hat{C}_{l'}^{\hat{\phi}\hat{\phi}} - 2\hat{N}_{l'}^{(0)}) \right\rangle, \tag{50}$$

where

$$A_{ll'} = \frac{\partial C_l^{\tilde{T}\tilde{T}}}{\partial C_{l'}^{\phi\phi}} \left(\frac{C_{l'}^{\phi\phi}}{\langle \hat{C}_{l'}^{\hat{\phi}\hat{\phi}} \rangle} \right)^2. \tag{51}$$

The expectation value is introduced in Eq. (50) to preserve the mean value of $\hat{C}_{l,\rm expt}^{\tilde{T}\tilde{T}}$. The construct $(C_{l'}^{\phi\phi}/\langle \hat{C}_{l'}^{\hat{\phi}\hat{\phi}} \rangle)^2$ acts like a Wiener filter on the empirical spectrum $\hat{C}_{l'}^{\hat{\phi}\hat{\phi}} - 2\hat{N}_{l'}^{(0)}$. The effect is to "delens" the lensed temperature power spectrum with the (CMB-averaged) lensing effect from any scales in the reconstructed potential power spectrum where the signal-to-noise on the reconstruction is high. In practice, high signal-to-noise lens reconstructions are never achieved with reconstructions based only on the temperature.

An alternative procedure is suggested by Eq. (46): project out from $\hat{C}_L^{\hat{\phi}\hat{\phi}}$ the dominant singular vector u_L. Although this is lossy, it does have the virtue that the $\tilde{T}\tilde{T}$ part of the likelihood is left unchanged. We discuss the effect of such a lossy projection on the errors of amplitude estimates in Sec. V D 1.

TABLE II. Covariance types UU and VV that are picked up by power covariances $\mathrm{cov}(\hat{C}_L^{XY}, \hat{C}_{L'}^{ZW})$ according to Eq. (52). Lensed power spectra are denoted with tildes, and for $\tilde{T}\tilde{T}$ the VV covariance should be understood to correspond to $C_{l,\mathrm{expt}}^{\tilde{T}\tilde{T}}$ including noise. The notation $\hat{\phi}\hat{\phi} - 2\hat{N}^{(0)} + N^{(0)}$ stands for $\hat{C}_L^{XY} = \hat{C}_L^{\hat{\phi}\hat{\phi}} - 2\hat{N}_L^{(0)} + N_L^{(0)}$, and in this case the Gaussian variance in the first term of Eq. (52) involves $\langle \hat{C}_L^{\hat{\phi}\hat{\phi}} \rangle^2$, while the Gaussian covariance with $\hat{C}_{L'}^{\tilde{T}\tilde{T}}$ involves $(C_L^{T\phi})^2$. The symbol – means that the corresponding contribution is cancelled or can be neglected. For the temperature-lensing covariances in the bottom two rows, the leading trispectrum contribution of Eq. (D4) should be added to Eq. (52).

XY, ZW	UU	VV	Equation here	Reference
Any combination of $\tilde{T}\tilde{T}$, $\tilde{T}\tilde{E}$, $\tilde{E}\tilde{E}$	$\phi\phi$	–	(24)	[40]
$\tilde{B}\tilde{B}$, $\tilde{B}\tilde{B}$	$\phi\phi$	EE		[40]
$\hat{\phi}\hat{\phi}$, $\hat{\phi}\hat{\phi}$	–	$\tilde{T}\tilde{T}$	(29) and (30)	[26,32] and this work
$\hat{\phi}\hat{\phi} - 2\hat{N}^{(0)} + N^{(0)}$, $\hat{\phi}\hat{\phi} - 2\hat{N}^{(0)} + N^{(0)}$	–	–	(29)	[26]
$\hat{\phi}\hat{\phi}$, $\tilde{T}\tilde{T}$	$\phi\phi$	$\tilde{T}\tilde{T}$	(E8) and (36)	this work
$\hat{\phi}\hat{\phi} - 2\hat{N}^{(0)} + N^{(0)}$, $\tilde{T}\tilde{T}$	$\phi\phi$	–	(E8)	this work

C. Toward a complete model for power covariances

The power covariances of Eqs. (30), (36), and (E8) can be regarded as a natural extension of the model for temperature and polarization power covariances found in Ref. [40]. We can summarize the covariances in the unified form

$$\begin{aligned}\mathrm{cov}(\hat{C}_L^{XY}, \hat{C}_{L'}^{ZW}) = {} & \delta_{LL'}\mathrm{cov}_G(\hat{C}_L^{XY}, \hat{C}_{L'}^{ZW}) \\ & + \sum_{ll'} \frac{\partial C_L^{XY}}{\partial C_l^{UU}} \mathrm{cov}(\hat{C}_l^{UU}, \hat{C}_{l'}^{UU}) \frac{\partial C_{L'}^{ZW}}{\partial C_{l'}^{UU}} \\ & + \sum_{ll'} \frac{\partial C_L^{XY}}{\partial C_l^{VV}} \mathrm{cov}(\hat{C}_l^{VV}, \hat{C}_{l'}^{VV}) \frac{\partial C_{L'}^{ZW}}{\partial C_{l'}^{VV}}, \end{aligned} \quad (52)$$

where U and V are listed in Table II for different combinations of XY and ZW. In this context we make the identifications

$$\frac{\partial C_L^{\hat{\phi}\hat{\phi}}}{\partial C_l^{\phi\phi}} \equiv \delta_{Ll}, \qquad \frac{\partial C_L^{\hat{\phi}\hat{\phi}}}{\partial C_l^{\tilde{T}\tilde{T}}} \equiv \frac{\partial(2\hat{N}_L^{(0)})}{\partial \hat{C}_{l,\mathrm{expt}}^{\tilde{T}\tilde{T}}}. \quad (53)$$

The first term on the right of Eq. (52) is the Gaussian covariance that would arise for Gaussian fields X, Y, Z, and W.

The general formula of Eq. (52) does not include trispectrum contributions to the temperature-lensing covariance, but the dominant correction has a simple form [Eq. (D4)], which can be added straightforwardly to the covariance. Generally, terms involving $C^{T\phi}$ in Eq. (52) can be neglected. While we evaluated derivatives perturbatively in ϕ, nonperturbative corrections can be included from numerical derivatives of accurate lensed power spectra [40,41]. However, we do not expect these corrections to be significant here. All combinations listed in Table II have been verified with simulations in Refs. [26,40] or in this work. Extending the covariance model to polarization-based lensing reconstructions would be interesting but is beyond the scope of this paper.

D. Impact of correlations on parameter estimation

1. Lensing amplitude estimates

As a first step in assessing the impact of covariances between the temperature and lens-reconstruction power spectra on parameter estimation, we consider constraining an overall amplitude parameter A of a fiducial lensing power spectrum, $C_l^{\phi\phi} = AC_l^{\phi\phi}|_{\mathrm{fid}}$ [8] with all other parameters fixed. The value of $A = 1$ corresponds to lensing at the level expected in the fiducial model, while $A = 0$ corresponds to no lensing.

The lensing amplitude can be estimated from the reconstructed lensing potential with

$$\hat{A} = \frac{\sum^{l,l'} C_l^{\phi\phi} (\mathrm{cov}^{-1}_{\hat{\phi}\hat{\phi}})_{ll'} (\hat{C}_{l'}^{\hat{\phi}\hat{\phi}} - N_{l'}^{(0)} - N_{l'}^{(1)})}{\sum^{L,L'} C_L^{\phi\phi} (\mathrm{cov}^{-1}_{\hat{\phi}\hat{\phi}})_{LL'} C_{L'}^{\phi\phi}}, \quad (54)$$

where $\mathrm{cov}_{\hat{\phi}\hat{\phi}}$ denotes the autocovariance of the reconstructed lensing power given by Eqs. (28) and (29), evaluated for $A = 1$. Equation (54) is the maximum-likelihood estimator for the lensing amplitude if the likelihood is modeled as a multivariate Gaussian in the empirical power spectrum of the lensing reconstruction. This form of the likelihood will be motivated later. Alternatively, the lensing amplitude can be extracted directly from the lensed temperature power spectrum without invoking lensing reconstruction by

$$\hat{A}' = \frac{\sum^{l} (\hat{C}_{l,\mathrm{expt}}^{\tilde{T}\tilde{T}} - C_{l,\mathrm{expt}}^{TT}) (\mathrm{cov}^{-1}_{\tilde{T}\tilde{T},\mathrm{expt}})_{ll} (C_l^{\tilde{T}\tilde{T}} - C_l^{TT})}{\sum^{l'} (C_{l'}^{\tilde{T}\tilde{T}} - C_{l'}^{TT})^2 (\mathrm{cov}^{-1}_{\tilde{T}\tilde{T},\mathrm{expt}})_{l'l'}}, \quad (55)$$

where the autocovariance of the temperature power is approximated by its leading-order diagonal piece [see Eq. (24)].

Since the reconstruction-based amplitude $\hat{A}$ is linear in the empirical reconstruction power and the temperature-based amplitude $\hat{A}'$ is linear in the empirical lensed

temperature power, the covariance of $\hat{A}$ and $\hat{A}'$ involves the temperature-lensing power covariance that we computed earlier,

$$\mathrm{cov}(\hat{A},\hat{A}') = \sigma_A^2\sigma_{A'}^2 \sum_{l,l',l''} C_l^{\phi\phi}(\mathrm{cov}^{-1}_{\hat{\phi}\hat{\phi}})_{ll'}\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{l'}, \hat{C}^{\tilde{T}\tilde{T}}_{l'',\mathrm{expt}}) \times (\mathrm{cov}^{-1}_{\tilde{T}\tilde{T},\mathrm{expt}})_{l''l''}(C^{\tilde{T}\tilde{T}}_{l''} - C^{TT}_{l''}), \quad (56)$$

where the standard deviations are[9]

$$\sigma_A = \left[\sum_{l,l'} C_l^{\phi\phi}(\mathrm{cov}^{-1}_{\hat{\phi}\hat{\phi}})_{ll'}C_{l'}^{\phi\phi}\right]^{-1/2}, \qquad \sigma_{A'} = \left[\sum_{l}(C_l^{\tilde{T}\tilde{T}} - C_l^{TT})^2(\mathrm{cov}^{-1}_{\tilde{T}\tilde{T},\mathrm{expt}})_{ll}\right]^{-1/2}. \quad (57)$$

The covariance of $\hat{A}$ and $\hat{A}'$ can be measured in N_{sims} simulations with

$$\widehat{\mathrm{cov}}(\hat{A},\hat{A}') = \frac{1}{N_{\mathrm{sims}}-1}\sum_{s=1}^{N_{\mathrm{sims}}}(\hat{A}_s - \langle\hat{A}\rangle_{\mathrm{sims}})(\hat{A}'_s - \langle\hat{A}'\rangle_{\mathrm{sims}}), \quad (58)$$

where s labels different realizations. We get the correlation by dividing by the theoretical standard deviations of Eq. (57). Approximating the spectra as Gaussian variables, the variance of the estimated covariance is $\sigma_A^2\sigma_{A'}^2 + [\mathrm{cov}(\hat{A},\hat{A}')]^2$, i.e., the variance of $(\hat{A}-1)(\hat{A}'-1)$, divided by $N_{\mathrm{sims}}-1$.[10] If we ignore the (small) correlation between $\hat{A}$ and $\hat{A}'$, the theoretical standard error of the measured covariance is therefore $\sigma_A\sigma_{A'}/\sqrt{N_{\mathrm{sims}}-1}$; i.e., the theoretical error of the estimated correlation is roughly $1/\sqrt{N_{\mathrm{sims}}-1} \approx 3.2\%$ for 1000 simulations.

Figure 8 shows that the lensing amplitude correlation measured in our simulations agrees well with the theoretical correlation of Eq. (56) if all contributions to the temperature-lensing power covariance are taken into account. As one of the main results of this paper, we find that the correlation is at most 7% if the realization-dependent $\hat{N}^{(0)}$ subtraction [Eq. (17)] is used. Without this $\hat{N}^{(0)}$ subtraction, the correlation can reach up to 10% because the disconnected noise contribution of Eq. (35) is not cancelled. A plausibility argument for the relatively small level of amplitude correlations is presented in Appendix A. Briefly, the disconnected noise contribution is small since the temperature modes that bring most information to the reconstruction are in between acoustic peaks and troughs, but the temperature modes that influence $\hat{A}'$ most strongly are at the peaks and troughs. Since these disjoint modes vary independently, the amplitude correlation is suppressed. The matter cosmic variance contribution to the amplitude correlation is small since the errors in the measurements of $\hat{A}$ and $\hat{A}'$ are dominated by cosmic variance of the temperature, not the lenses.

Figure 8 also illustrates the relative importance of the individual covariance contributions derived above. The dominant effect comes from the matter cosmic variance contribution [Eq. (E8)] which induces an amplitude correlation of around 4–5% for any $l^{\phi}_{\mathrm{max}} \gtrsim 100$. The disconnected noise contribution [Eq. (35)] implies a slightly smaller amplitude correlation of 3–4% for $l^{\phi}_{\mathrm{max}} \gtrsim 300$.[11] An additional contribution to the temperature-lensing power covariance comes from the lensed temperature trispectrum discussed in Appendix D. The dominant term [Eq. (D4)] gives rise to a 2% lensing amplitude correlation. The agreement between simulations and theory in Fig. 8 gives us confidence that the power correlations modeled here include all relevant contributions for amplitude measurements.

Instead of fixing $l^{\phi}_{\mathrm{min}} = 2$ and varying l^{ϕ}_{max} it is worthwhile to consider the disjoint reconstruction bins $[l^{\phi}_{\mathrm{min}}, l^{\phi}_{\mathrm{max}}] = [40, 84], [85, 129], [130, 174], \ldots, [355, 400]$ used for the Planck analyses in Refs. [13,17]. If the realization-dependent $\hat{N}^{(0)}$ is used, the theoretical correlation of the lensing amplitude estimated from one of these bins alone with the lensing amplitude estimated from the temperature power (for $l^{T}_{\mathrm{max}} = 2002$) is 5, 3.6, and 2% for the first three reconstruction bins and decreases further for the remaining higher-l bins. This is consistent with the correlations estimated from our 1000 simulations. In particular, this result shows that the lensing amplitude estimated from the temperature power spectrum and the reconstruction amplitudes used in the Planck lensing likelihood [13] are nearly uncorrelated, which justifies neglecting this correlation in the likelihood.

[9] If no empirical $\hat{N}^{(0)}$ subtraction is used, we evaluate σ_A with nondiagonal reconstruction power autocovariance, which gives $\sigma_A \approx 2.7\%$ if $l^{\phi}_{\mathrm{max}} \geq 500$ for our noise and beam specifications. The estimated sample standard deviation of $\hat{A}$ from simulations is larger by a factor of up to 1.07 compared to the theoretical expectation. If the $\hat{N}^{(0)}$ subtraction is used, we evaluate σ_A with diagonal reconstruction power autocovariance, which yields $\sigma_A \approx 2.5$–2.6% for $l^{\phi}_{\mathrm{max}} \geq 500$. The estimated sample standard deviation from simulations is larger by a factor of at most 1.05. The modest reduction in σ_A with empirical $\hat{N}^{(0)}$ subtraction is expected given the origin of this estimator as the approximate maximum-likelihood estimator for the trispectrum (see Appendix B). For the lensing amplitude A' estimated from the temperature power spectrum, we find $\sigma_{A'} \approx 3.9\%$ for $l^{T}_{\mathrm{max}} = 2002$ and our Planck-like noise model. The estimated sample standard deviation from simulations is larger by a factor of 1.01.

[10] As a product of two approximately normally distributed variables, the random variable $(\hat{A}_s - \langle\hat{A}\rangle_{\mathrm{sims}})(\hat{A}'_s - \langle\hat{A}'\rangle_{\mathrm{sims}})$ is not normally distributed. However, the average over $N_{\mathrm{sims}} = 1000$ realizations is approximately normally distributed due to the central limit theorem.

[11] Although for the *power spectrum* cross-correlation the maximal noise contribution is about 1 order of magnitude larger than the maximal matter cosmic variance contribution, the latter can be more relevant for the correlation of amplitude estimates because of the phase argument given in the text.

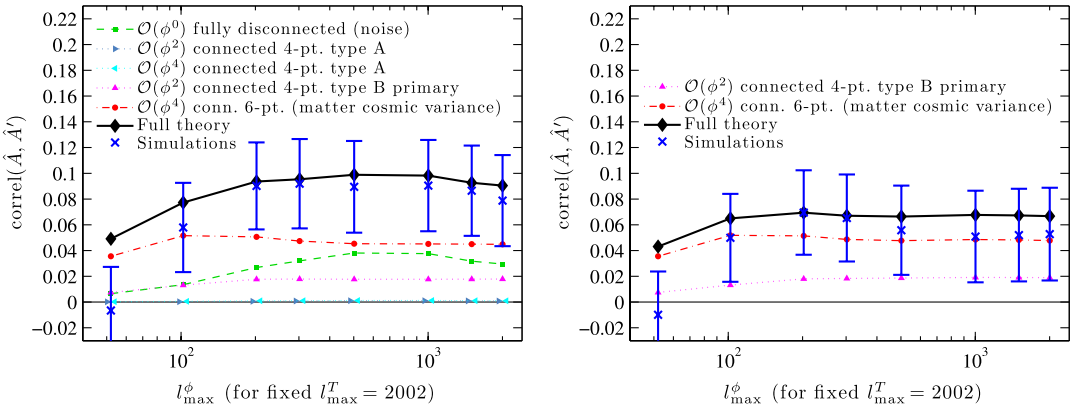


FIG. 8 (color online). (a) Correlation $\widehat{\rm correl}(\hat{A}, \hat{A}') = \widehat{\rm cov}(\hat{A}, \hat{A}')/(\sigma_A \sigma_{A'})$ of the lensing amplitude estimates of Eqs. (54) and (55) measured in 1000 simulations (blue crosses), if no empirical $\hat{N}^{(0)}$ subtraction is applied. The theoretical correlation from Eq. (56) (black) includes the noise contribution of Eq. (35), the matter cosmic variance contribution of Eq. (E8), and the dominant connected 4-point contributions from Eqs. (D2) and (D4). In the simulations the lensing potential is reconstructed from the lensed temperature power spectrum up to $l^{\rm rec}_{\rm max} = 2750$. Then, $\hat{A}$ is estimated from $\hat{C}^{\hat{\phi}\hat{\phi}}_l$ from $l^{\phi}_{\rm min} = 2$ up to the multipole $l^{\phi}_{\rm max}$, which is varied along the horizontal axis. The amplitude $\hat{A}'$ is estimated using $\hat{C}^{\hat{T}\hat{T}}_{l,\rm expt}$ up to $l^{T}_{\rm max} = 2002$. Theoretical expressions are evaluated with the same cutoffs (nondiagonal covariance matrices are cut off before inverting them). Error bars show standard errors on the measured correlation. The errors are very correlated because they all involve overlapping low-l reconstruction modes. (b) Same as (a) but with empirical $\hat{N}^{(0)}$ subtraction. Since this removes some covariance contributions, the theoretical covariance is given by Eqs. (D4) and (E8) only.

For experiments with superior noise and beam characteristics, the matter cosmic variance contribution to the temperature-lensing power covariance does not change, but the lensing amplitude errors decrease. We therefore expect the corresponding amplitude correlation to increase. For example, for a full-sky experiment with SPT-like noise and beam specificiations, $\sigma_N = 18\,\mu$K arcmin and $\sigma_{\rm FWHM} = 1$ arcmin, the amplitude correlation from the matter cosmic variance contribution alone is around 10–11% for $l^{T}_{\rm max} = 2002$ and $l^{\phi}_{\rm max} = 500$–1000.

2. Combined lensing amplitude estimate

We have presented two estimators of the lensing amplitude: $\hat{A}$ is linear in the reconstruction power, and $\hat{A}'$ is linear in the CMB power. These two estimates can be combined with inverse variance weighting,

$$\hat{A}_C = \frac{1}{\sigma_A^{-2} + \sigma_{A'}^{-2}}\left(\frac{\hat{A}}{\sigma_A^2} + \frac{\hat{A}'}{\sigma_{A'}^2}\right). \tag{59}$$

This combined estimator is the maximum-likelihood estimator for the lensing amplitude if $\hat{A}$ and $\hat{A}'$ are assumed to be uncorrelated. If there is a correlation between $\hat{A}$ and $\hat{A}'$, this does not change the expectation value of $\hat{A}_C$, but it does change its variance, which is then given by[12]

$$\begin{aligned}\mathrm{var}(\hat{A}_C) = {} & \sigma_A \sigma_{A'}\left(\frac{\sigma_A}{\sigma_{A'}} + \frac{\sigma_{A'}}{\sigma_A}\right)^{-2} \\ & \times \left[\left(\frac{\sigma_A}{\sigma_{A'}} + \frac{\sigma_{A'}}{\sigma_A}\right) + 2\,\mathrm{correl}(\hat{A}, \hat{A}')\right].\end{aligned} \tag{60}$$

A correlation between $\hat{A}$ and $\hat{A}'$ therefore increases the 1σ error of the combined estimator (59) by a factor of

$$\begin{aligned}\frac{\sigma_{A_C}|_{\mathrm{cov}(A,A')\neq 0}}{\sigma_{A_C}|_{\mathrm{cov}(A,A')=0}} &= \sqrt{1 + \frac{2(\sigma_A/\sigma_{A'})}{1 + (\sigma_A/\sigma_{A'})^2}\,\mathrm{correl}(\hat{A}, \hat{A}')} \\ &\leq \sqrt{1 + \mathrm{correl}(\hat{A}, \hat{A}')}\end{aligned} \tag{61}$$

compared to the error if $\hat{A}$ and $\hat{A}'$ were uncorrelated. Since correlations between $\hat{A}$ and $\hat{A}'$ were found to be at most 7% if the empirical $\hat{N}^{(0)}$ subtraction is used, the 1σ error of the combined lensing estimate changes by at most 3.5% for Planck (5.5% for the full-sky SPT-like experiment mentioned above). Noting that this is the error on the error bar, the correlations between $\hat{A}$ and $\hat{A}'$ found above can be safely neglected when combining these two estimates of the lensing amplitude.

We briefly introduced a projection technique in Sec. V B 5 to remove the covariance between the reconstructed lensing power and the lensed temperature power spectrum due to the cosmic variance of the lenses. A simple way to perform the projection is to modify the covariance matrix $(\mathrm{cov}_{\hat{\phi}\hat{\phi}})_{ll'}$ in Eq. (54) by adding $\lambda' u_l u_{l'}$, where u_l is the dominant singular vector in Eq. (46), and taking λ' to infinity. To the extent that the

[12]To first order in the correlation, this sampling variance of the combined $\hat{A}_C$ is the same as the sampling variance of the optimal combined estimate that takes account of the correlations.

covariance between the lensing and temperature power spectra is really rank-one, this procedure removes the correlation between $\hat{A}$ and $\hat{A}'$ exactly. However, the variance of $\hat{A}$ is increased by projection: it is still given by Eq. (57) but with the modified $(\mathrm{cov}_{\hat{\phi}\hat{\phi}})_{ll'}$. For $l^{\phi}_{\mathrm{max}} = 500$, we find that σ_A is increased from 0.025 to 0.042, i.e., a 70% increase. This falls to 60% for $l^{\phi}_{\mathrm{max}} = 1000$. The reason for the large increase is that u_l is rather similar in shape to the signal whose amplitude we are trying to reconstruct. Given the large increase in the error on $\hat{A}$, and that ignoring the effect of the covariance between the lensing reconstruction and temperature power spectra is relatively harmless, we do not advocate the use of projection to remove the correlations.

3. Cosmological parameters

We naively expect the impact of power correlations on cosmological parameters to be smaller than for the lensing amplitude because the latter is directly related to the lensing potential on all scales and can therefore accumulate contributions from the full power covariances. We confirm this with a simple Fisher analysis. The covariance matrix for the joint data vector $\hat{C} = (\hat{C}^{\tilde{T}\tilde{T}}_{\mathrm{expt}}, \hat{C}^{\hat{\phi}\hat{\phi}} - 2\hat{N}^{(0)} + N^{(0)})$ is

$$\mathrm{cov}_{LL',\mathrm{joint}} \equiv \mathrm{cov}(\underline{\hat{C}}_L, \underline{\hat{C}}_{L'}) = \begin{pmatrix} \delta_{LL'}\mathrm{var}_G(C^{\tilde{T}\tilde{T}}_{L,\mathrm{expt}}) & \mathrm{cov}(\hat{C}^{\tilde{T}\tilde{T}}_{L,\mathrm{expt}}, \hat{C}^{\hat{\phi}\hat{\phi}}_{L'} - 2\hat{N}^{(0)}_{L'}) \\ \mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L} - 2\hat{N}^{(0)}_{L}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}) & \delta_{LL'}\mathrm{var}_G(\langle \hat{C}^{\hat{\phi}\hat{\phi}}_L \rangle) \end{pmatrix}. \tag{62}$$

Fisher errors are obtained by taking the square root of the diagonal entries of the inverse of the Fisher matrix

$$F_{ij} = \sum_{LL'} \frac{\partial \underline{C}_L}{\partial p_i} (\mathrm{cov}^{-1}_{\mathrm{joint}})_{LL'} \frac{\partial \underline{C}_{L'}}{\partial p_j} \tag{63}$$

for cosmological parameters $\mathbf{p} = (\Omega_b h^2, \Omega_c h^2, h, \tau, A_s, n_s, \Omega_\nu h^2, \Omega_K)$ and theoretical power spectra $\underline{C} = (C^{\tilde{T}\tilde{T}}, C^{\phi\phi})$ (assuming cosmology-independent $N^{(0)}$ for simplicity). Including the off-diagonal temperature-lensing covariances of Eqs. (D4) and (E8) in Eq. (62) increases the Fisher errors for these parameters by at most 0.7% [0.5% if only Eq. (E8) is used] compared to a completely diagonal covariance matrix.[13] The off-diagonal part of the joint covariance matrix can therefore safely be neglected for cosmological parameter estimation with a Planck-like experiment.

VI. TOWARD A LENSING LIKELIHOOD

As argued in Sec. I, dealing with the exact likelihood for the lensed CMB temperature is generally computationally prohibitive. For this reason, we have focused on a form of data compression whereby the non-Gaussian lensed CMB is represented by its 2- and 4-point functions (the latter via the lensing reconstrucion power spectrum). In computing the correlations between these spectra, we have implicitly been assuming that the likelihood takes the form of a multivariate Gaussian in the spectra. In this section, we test the accuracy of this assumption in simple parameter-estimation exercises.

[13]To obtain accurate derivatives for the Fisher matrix, we assumed a fiducial cosmology with massive neutrinos, $\mathbf{p}_{\mathrm{fid}} = (0.0226, 0.1123, 0.704, 0.087, 2.167\times10^{-9}, 0.963, 0.005, 0)$, which differs slightly from the cosmology used throughout the rest of the paper. The Fisher errors were computed for $l^{T}_{\mathrm{max}} = 2002$ and $l^{\phi}_{\mathrm{max}} = 1002$.

A. Lensing amplitude from lensing reconstruction

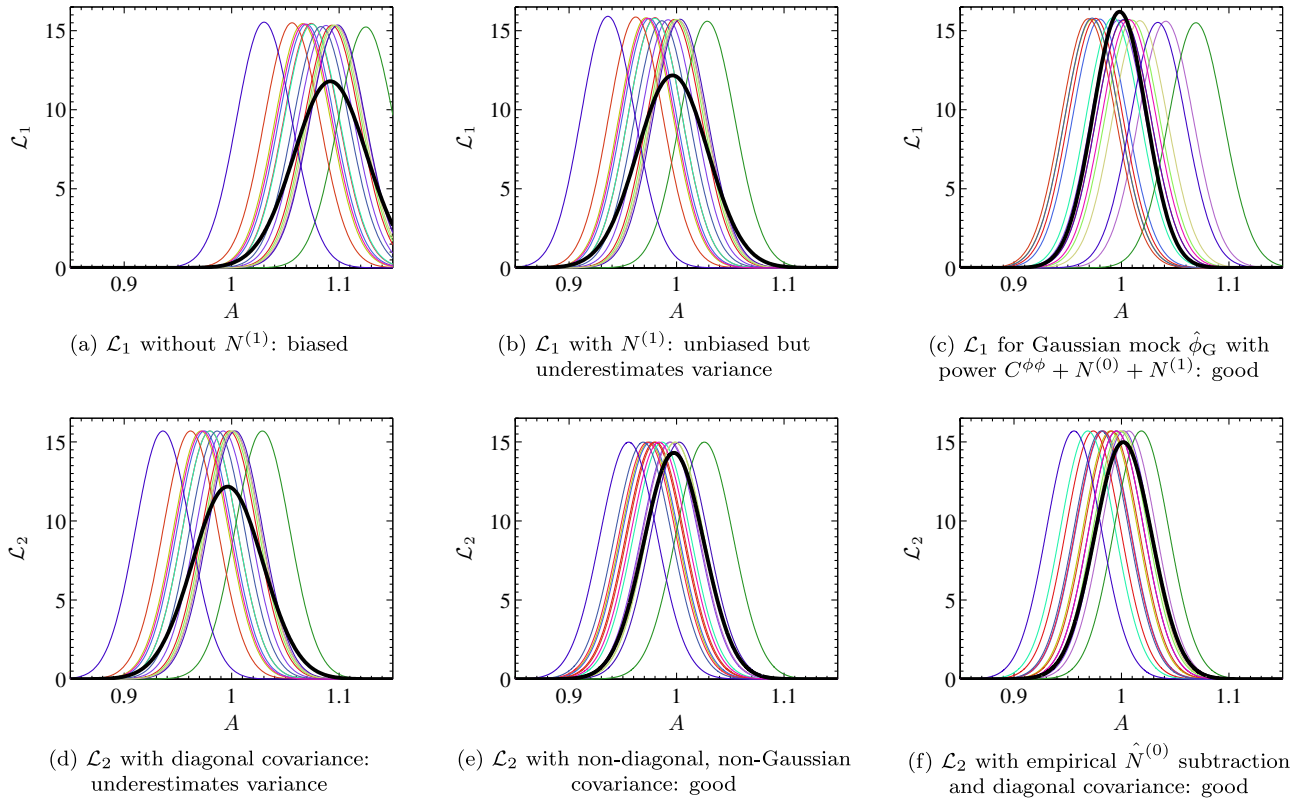
As a toy model, we first aim to constrain the lensing amplitude A from the lensing reconstruction alone. Considering an isotropic CMB survey, with Planck-like noise as described earlier, we consider two simple models for the likelihood, both of which depend only on the empirical power spectrum of the reconstruction. The first is the usual isotropic likelihood for a Gaussian field:

$$\begin{aligned} -2\ln\mathcal{L}_1(\hat{\phi}|A) &= \sum_l (2l+1)\left(\frac{\hat{C}^{\hat{\phi}\hat{\phi}}_l}{AC^{\phi\phi}_l + N_l} + \ln|AC^{\phi\phi}_l + N_l|\right) \\ &\quad + (\mathrm{const}). \end{aligned} \tag{64}$$

This would be correct if $\hat{\phi}$ were a Gaussian field. However, since the reconstruction is manifestly non-Gaussian, we do not expect this likelihood to perform well. The second is Gaussian in the empirical power spectrum $\hat{C}^{\hat{\phi}\hat{\phi}}$:

$$\begin{aligned} -2\ln\mathcal{L}_2(\hat{C}^{\hat{\phi}\hat{\phi}}|A) &= \sum_{l,l'} [\hat{C}^{\hat{\phi}\hat{\phi}}_l - (AC^{\phi\phi}_l + N_l)](\mathrm{cov}^{-1}_{\hat{\phi}\hat{\phi}})_{ll'} \\ &\quad \times [\hat{C}^{\hat{\phi}\hat{\phi}}_{l'} - (AC^{\phi\phi}_{l'} + N_{l'})] + (\mathrm{const}). \end{aligned} \tag{65}$$

The theoretical reconstruction power autocovariance $\mathrm{cov}_{\hat{\phi}\hat{\phi}}$, given by Eqs. (28) and (29), and the bias of the reconstructed lensing power, $N = N^{(0)} + N^{(1)}$, are evaluated for the fiducial amplitude $A = 1$. The empirical $\hat{N}^{(0)}$ subtraction is obtained by replacing N with $\hat{N} = 2\hat{N}^{(0)} - N^{(0)} + N^{(1)}$. For $\mathcal{L}_1$, the maximum-likelihood estimate for A (given a realization of the lensing reconstruction) is found numerically by direct evaluation of $\mathcal{L}_1$ for various A. The maximum likelihood estimator for A based on the second likelihood $\mathcal{L}_2$ is given by Eq. (54).



(a) $\mathcal{L}_1$ without $N^{(1)}$: biased

(b) $\mathcal{L}_1$ with $N^{(1)}$: unbiased but underestimates variance

(c) $\mathcal{L}_1$ for Gaussian mock $\hat{\phi}_{\rm G}$ with power $C^{\phi\phi} + N^{(0)} + N^{(1)}$: good

(d) $\mathcal{L}_2$ with diagonal covariance: underestimates variance

(e) $\mathcal{L}_2$ with non-diagonal, non-Gaussian covariance: good

(f) $\mathcal{L}_2$ with empirical $\hat{N}^{(0)}$ subtraction and diagonal covariance: good

FIG. 9 (color online). Likelihoods for the lensing amplitude for 15 different realizations (thin, colored lines), using the lensing reconstruction up to $l^{\phi}_{\rm max} = 2650$, compared with the scatter in the best-fit amplitude obtained over 1000 simulations. Each likelihood peaks at some best-fit parameter $\hat{A}$, and the mean and scatter of these best-fitting amplitudes over all 1000 simulations correspond to the Gaussian curves (thick black). The upper panels show $\mathcal{L}_1$ from Eq. (64), while the lower panels show $\mathcal{L}_2$ given by Eq. (65). All curves are normalized such that their integral over A is unity.

We compute estimates $\hat{A}$ of the lensing amplitude for 1000 realizations of the lensed CMB. The sample mean of $\hat{A}$ should be unity, and the sample variance of $\hat{A}$, i.e., the scatter of the best-fit amplitude over different realizations, should agree with the typical width of the likelihood evaluated for a single realization. Checking these two properties provides a nontrivial test of the likelihood $\mathcal{L}_1$. In contrast, for $\mathcal{L}_2$, rather than testing the accuracy of $\mathcal{L}_2$, the sample mean and sample variance of $\hat{A}$ just test our understanding of the mean and covariance of $\hat{C}^{\hat{\phi}\hat{\phi}}$.[14] This test is still useful to check for residual biases and the accuracy of our model for the reconstruction power covariance.

Figure 9 compares the likelihood evaluated for several individual realizations (colored) with a Gaussian (black) with mean and standard deviation given by sample mean and standard deviation of $\hat{A}$ averaged over all 1000 realizations (for $l^{\phi}_{\rm max} = 2650$ and our Planck-like noise model). Including the $N^{(1)}$ bias is important at high multipoles for both likelihoods: e.g., without it, $\mathcal{L}_1$ overestimates the lensing amplitude by 9%; see Fig. 9(a). Including the $N^{(1)}$ bias in $\mathcal{L}_1$ yields the correct lensing amplitude $A = 1$ in the mean, but the scatter of $\hat{A}$ over realizations is more than 30% larger than the typical width of $\mathcal{L}_1$ in a single realization; see Fig. 9(b). The likelihood $\mathcal{L}_1$ underpredicts the error of A because of the non-Gaussianity of $\hat{\phi}$. This is demonstrated in Fig. 9(c), for which we replace the reconstructions with Gaussian simulations of a field with power spectrum $C^{\phi\phi}_l + N^{(0)}_l + N^{(1)}_l$. In this case, $\mathcal{L}_1$ should be exact, and the scatter does indeed match the widths of individual realizations.

In contrast, $\mathcal{L}_2$ can partly model the non-Gaussianity of $\hat{\phi}$ through the nondiagonal reconstruction power autocovariance. We compute $\hat{A}$ based on $\mathcal{L}_2$ for different forms of the lensing covariance. Neglecting off-diagonal contributions to $\mathrm{cov}_{\hat{\phi}\hat{\phi}}$ gives likelihood-based errors for A less than 80% of the scatter of $\hat{A}$ across the simulations; see Fig. 9(d). If we include non-Gaussian, off-diagonal contributions given by Eq. (28), we find that $\mathcal{L}_2$ predicts the scatter in A to better than 5%; see Fig. 9(e). Similar

[14]This is because the estimated lensing amplitude [Eq. (54)] is linear in $\hat{C}^{\hat{\phi}\hat{\phi}}$. For example, if the true likelihood depends on the third power of $\hat{C}^{\hat{\phi}\hat{\phi}}$, it would be possible that this only shows up in the skewness of $\hat{A}$. This issue will be addressed later by considering the tilt of the lensing power spectrum, which depends nonlinearly on the reconstruction power.

results are achieved with the empirical $\hat{N}^{(0)}$ correction of Eq. (17) and diagonal (Gaussian) reconstruction power covariance; see Fig. 9(f).

B. Two-parameter likelihood tests with lensing amplitude and lensing tilt

To test the likelihood approximation $\mathcal{L}_2$, we use the lensing reconstruction additionally to constrain the lensing tilt n defined by

$$C_l^{\phi\phi} = A\left(\frac{l}{l_*}\right)^n C_l^{\phi\phi}|_{\rm fid}. \tag{66}$$

The pivot multipole $l_* = 124$ is chosen such that the Fisher matrix associated with $\mathcal{L}_2$ is diagonal [for $l^\phi_{\rm max} = \mathcal{O}(10^3)$], implying that the parameters A and n are approximately uncorrelated. The likelihoods for nine realizations are compared with the scatter of the best-fit parameters over 1000 realizations in Fig. 10. If the nondiagonal lensing power covariance of Eq. (28) is included, we find good agreement (without empiricial $\hat{N}^{(0)}$ subtraction). Note that we have binned the reconstruction power in bins with boundaries at

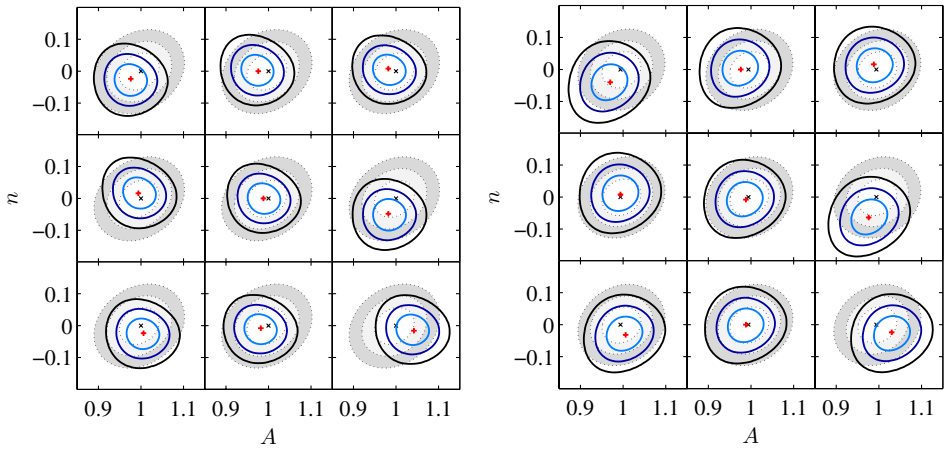
$$l = 2, 13, 35, 75, 115, 155, 195, \ldots$$
$$\text{(increasing by 40 above } l = 35). \tag{67}$$

Similar results for the unbinned case will be summarized in Fig. 11 below.

To quantify the level of agreement between the likelihoods for individual realizations and the scatter of their best-fit parameters, we compare the areas of the confidence contours shown in Fig. 10. We show in Fig. 11 the fractional deviation of the areas of the Gaussian, with the sample mean and sample covariance matched to the scatter of the best-fit parameters over realizations, from the average area of the likelihoods for individual realizations (i.e., the fractional deviation of gray background areas from average areas enclosed by the solid lines in Fig. 10).

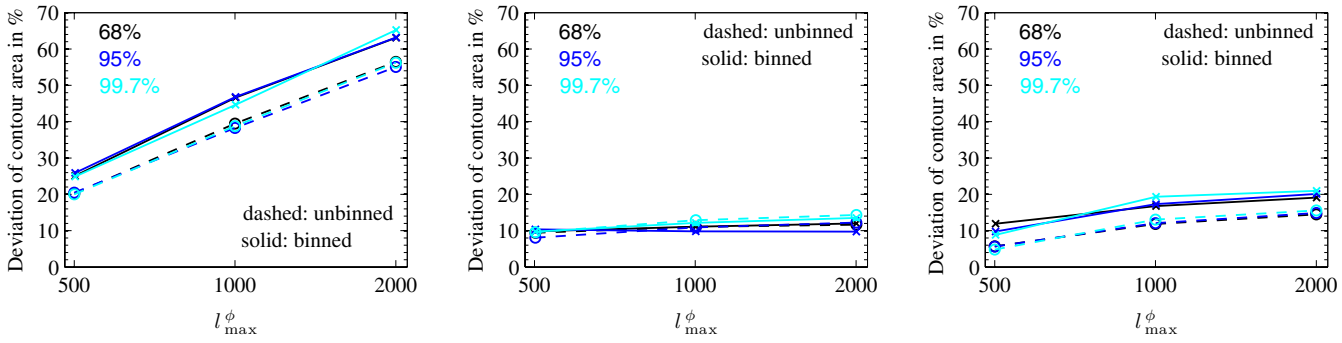
Neglecting the off-diagonal contribution to the lensing power covariance, which is largest at high reconstruction multipoles [26], gives narrow misshapen likelihoods that underestimate the scatter across simulations. This is particularly so for $l^\phi_{\rm max} \gtrsim 1000$ where the confidence areas disagree by around 40–65%. Binning does not help because it does not reduce the broadband correlations of the reconstruction power. The agreement is better when the nondiagonal reconstruction power covariance is used (the disagreement of confidence areas is at most 14%). Alternatively, if the empirical $\hat{N}^{(0)}$ bias correction and the diagonal covariance are used, the confidence areas deviate by at most 22%. If we assume circular contours, the fractional deviation of the contour radius is $\sqrt{1+d}-1 < d/2$ if d is the fractional deviation of the contour areas. Taking this as the approximate fractional error of the marginalized error bars of A or n shows that the error on the error bars is smaller than 11% if the nondiagonal reconstruction covariance or empirical $\hat{N}^{(0)}$ subtraction are used in $\mathcal{L}_2$. Therefore, these two cases provide a reasonably accurate model for the lensing likelihood in this test. If diagonal reconstruction power covariance is assumed, and no empirical $\hat{N}^{(0)}$ subtraction performed, the error on the error bars can reach 30% even if binning is used. It is also worth noting that in this last case only the confidence areas *increase* with $l^\phi_{\rm max}$; i.e., the analysis is clearly nonoptimal.

Note that in the above, ideally, we should use a histogram of the best-fit parameters instead of fitting a Gaussian to their scatter. This would test the tails of the distribution much better because it would include possible skewness,



(a) $\mathcal{L}_2$ with diagonal covariance (b) $\mathcal{L}_2$ with non-diagonal covariance

FIG. 10 (color online). Likelihood test if lensing amplitude A (horizontal axis) and tilt n (vertical axis) are varied. *Thick lines:* Contours enclosing 68, 95, and 99.7% of the probability for the likelihood $\mathcal{L}_2$ (65) evaluated for nine realizations. *Gray filled ellipses:* Contours of a Gaussian with central point and covariance matrix estimated from the scatter of the best-fit parameters over 1000 realizations. The lensing power autocovariance $\mathrm{cov}_{\hat{\phi}\hat{\phi}}$ is assumed to be diagonal in (a), while (b) also includes the nondiagonal contribution in Eq. (28). The reconstruction power is used up to $l^\phi_{\rm max} = 1002$ and is binned as described in the text.



(a) $\mathcal{L}_2$, diagonal covariance

(b) $\mathcal{L}_2$, non-diagonal covariance

(c) $\mathcal{L}_2$, empirical $\hat{N}^{(0)}$ correction, diagonal covariance

FIG. 11 (color online). Quantitative comparison of areas in the A vs n plane enclosed by contours of the $\mathcal{L}_2$ likelihood. We show the fractional deviation of the areas of Gaussians (with sample mean and sample covariance derived from the scatter of the best fit parameters over 1000 realizations) from the average area in a single realization. Positive values mean that the width of the likelihood typically underestimates the scatter of its peak across realizations. Results are shown for unbinned power spectra (dashed lines) and power spectra binned according to the scheme of Eq. (67) (solid lines) and for 68% (black), 95% (blue), and 99.7% (cyan) confidence levels. The maximum multipole of the reconstruction power is varied along the horizontal axis (only for three values which are connected by straight lines to guide the eye). The lensing power autocovariance $\mathrm{cov}_{\hat{\phi}\hat{\phi}}$ in $\mathcal{L}_2$ is assumed to be diagonal in (a), while (b) also includes the nondiagonal contribution of Eq. (28). Panel (c) is for empirical $\hat{N}^{(0)}$ bias correction and diagonal covariance. The contours in Fig. 10 correspond to the crosses at $l^{\phi}_{\rm max} = 1002$ in (a) and (b).

etc. However, we find that histograms from 1000 simulations are too noisy to be useful for this purpose, giving results that scatter significantly with changes in histogram binning widths ΔA and Δn.

VII. CONCLUSIONS

To include the CMB lensing reconstruction power spectrum in a joint likelihood analysis with the power spectrum of the temperature anisotropies requires knowledge of the cross-covariance of the two spectra. We computed this cross-covariance between the CMB 4-point and 2-point functions perturbatively, identifying two physical contributions. The disconnected part of the 6-point function of the lensed temperature leads to a noise contribution which can be interpreted as the response of the statistical noise in the lens reconstruction to fluctuations in the underlying CMB temperature field. The connected $\mathcal{O}(\phi^4)$ piece of the 6-point function gives rise to a second contribution attributable to the cosmic variance of the lenses, which causes the power spectrum of the lens reconstruction and the smoothing effect in the anisotropy power spectrum to covary. The temperature-lensing power covariance can therefore be written as

$$\begin{aligned}&\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}) \\ &\quad= \frac{\partial(2\hat{N}^{(0)}_{L})}{\partial \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}}} \frac{2}{2L'+1} (C^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})^2 \left[1 + 2\frac{C^{\phi\phi}_{L}}{A_L}\right] \\ &\qquad+ \frac{2}{2L+1} (C^{\phi\phi}_{L})^2 \frac{\partial C^{\tilde{T}\tilde{T}}_{L'}}{\partial C^{\phi\phi}_{L}},\end{aligned} \tag{68}$$

where perturbative expressions for the derivatives are given in Eqs. (27) and (31). Both contributions were confirmed with simulations. The second term in the square brackets represents the leading correction from the connected 4-point function; see Eq. (D4). The $C^{T\phi}$ correlation gives a diagonal contribution to the temperature-lensing power correlation, which is less than 5% for Planck-like specifications and falls rapidly with L. This generally has a negligible impact on parameter constraints derived jointly from the CMB 2- and 4-point functions.

We showed that correcting for the Gaussian $N^{(0)}$ bias in the reconstruction power with the data-dependent $\hat{N}^{(0)}$, advocated by Ref. [26] to remove autocovariances of the lensing reconstruction power spectrum, also removes the noise contribution to the temperature-lensing power correlation, and we provided an intuitive interpretation of this result.

For Planck-like specifications, estimates of the lensing amplitude A based on the lensing reconstruction or the peak smearing of the lensed temperature power spectrum can be correlated at around the 10% level due to the power correlations. If the correlations are ignored, this gives a misestimate of the error on a joint amplitude estimate of only 5%, which should be negligible. The data-dependent $\hat{N}^{(0)}$ bias correction reduces the amplitude correlation further to 7% and the error of the error to 3.5%. Intuitively, we can understand the smallness of the correlation (found perturbatively and with simulations) by noting that (i) covariance of the amplitude estimates due to cosmic variance of the lenses is limited by the small number of modes of $C^{\phi\phi}$ that influence the acoustic region of the temperature power spectrum, and is diluted significantly by CMB cosmic variance (and noise), and (ii) roughly disjoint scales in the CMB contribute to the amplitude determination from peak smearing and to the lens reconstruction limiting the correlation due to CMB cosmic variance.

(See Appendix A for further details of these arguments.) For a joint analysis of the power spectrum of a temperature-based CMB lensing reconstruction and the power spectrum of the temperature anisotropies themselves, the likelihoods for these two observables can therefore be simply combined for a Planck-like experiment (as was the case for the 2013 Planck analysis [13]).

Non-Gaussianity of the lensing reconstruction complicates the construction of a likelihood. We showed that the usual likelihood for isotropic Gaussian fields does not perform well for lens reconstruction in simple parameter tests, significantly underestimating the scatter seen in the best-fitting parameters across simulations. We obtained better results with simple likelihoods that are Gaussian in the measured spectra (with fiducial covariance matrix) provided that power spectrum covariances were properly modeled or data-dependent $N^{(0)}$ subtraction included. In two-parameter tests based on the amplitude and tilt of a fiducial lensing power spectrum, the widths of these Gaussian likelihoods reproduce the scatter in parameters across simulations at the 10% level.

With polarization-based reconstructions becoming feasible with current observations, it will be important to extend the analysis presented here to polarization (see Ref. [44] for work in this direction). While we expect that many of our results can be simply applied to reconstructions based on the temperature *and* polarization, the correlations are likely to be much more significant and particularly so for the most powerful *EB*-based reconstructions. We leave this to future work.

ACKNOWLEDGMENTS

We would like to thank Uros Seljak for illuminating discussions, particularly related to Appendix A 2. We also thank Helge Gruetjen, Simon Su, and Oliver Zahn for helpful discussions. The numerical calculations for this paper were performed on the COSMOS supercomputer, part of the DiRAC HPC Facility jointly funded by STFC and the Large Facilities Capital Fund of BIS. We are grateful to Andrey Kaliazin for computational support. We acknowledge use of LensPix [36] and HEALPix [37]. M. M. S. was supported by STFC, DAMTP Cambridge and St John's College Cambridge, and he thanks the University of Zurich, Institut d'astrophysique de Paris and Berkeley Center for Cosmological Physics for hospitality and the opportunity to present this work.

APPENDIX A: WHY ARE THE LENSING AMPLITUDE CROSS-CORRELATIONS SO SMALL?

The calculations in this paper give a rigorous derivation of the cross-correlation of the power spectra of the lensing reconstruction and the temperature anisotropies. In this appendix we present simple physical arguments for why the correlation of the lensing amplitudes estimated from the reconstruction power and the anisotropy power are so small. We present these arguments first for the correlation due to CMB cosmic variance and then for the correlation due to cosmic variance of the lenses.

1. Cosmic variance of the CMB

Due to the smoothing effect of lensing, most of the constraint on the lensing amplitude $\hat{A}'$ estimated from the CMB power spectrum comes from the CMB on scales of the acoustic peaks and troughs. This can be seen directly from the contribution $s_l^{\tilde{T}\tilde{T}}$ to the total signal-to-noise squared $[(S/N)^2]$ associated with the CMB cosmic variance at multipole l (see the blue curve in Fig. 12):

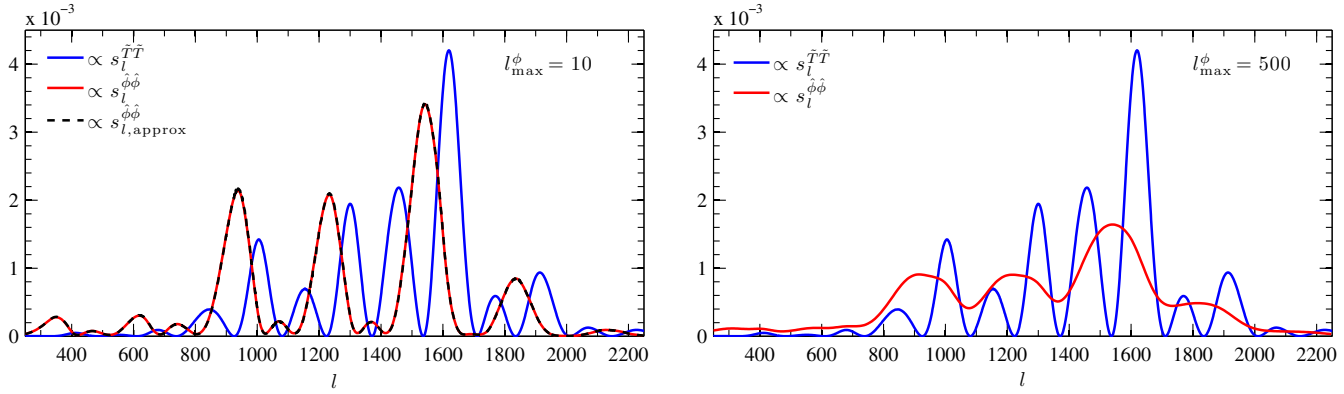
$$\left(\frac{S}{N}\right)^2_{A'} = \frac{1}{\sigma^2_{A'}} = \sum_{l=l^T_{\rm min}}^{l^T_{\rm max}} \frac{(C_l^{\tilde{T}\tilde{T}} - C_l^{TT})^2}{\underbrace{{\rm var}_G(C^{\tilde{T}\tilde{T}}_{l,{\rm expt}})}_{s_l^{\tilde{T}\tilde{T}}}}. \tag{A1}$$

In contrast, in the limit of very large-scale lenses, and as argued in more detail below, the reconstruction combines local convergence and shear measurements, for which scales in the CMB where the power spectrum changes rapidly are most informative. For large-scale lenses, the $(S/N)^2$ on the reconstruction-based amplitude estimate $\hat{A}$ is thus expected to be dominated by CMB modes *between* acoustic peaks and troughs. Therefore, the lensing amplitude estimates $\hat{A}$ and $\hat{A}'$ are determined by rather disjoint CMB modes with independent CMB cosmic-variance fluctuations. We therefore expect the amplitude correlation due to CMB cosmic variance to be suppressed (in case this correlation is not mitigated by the empirical $\hat{N}^{(0)}$ subtraction anyway).

To make this point more quantitative, note that the reconstruction power $\hat{C}^{\hat{\phi}\hat{\phi}}$ is affected by CMB cosmic variance through the disconnected CMB 4-point contribution $N^{(0)}$. Keeping the estimator normalization A_L and weights $\tilde{g}$ fixed in Eq. (16), the contribution from the CMB at multipole l to the $(S/N)^2$ of the reconstruction-based amplitude estimate $\hat{A}$ is monitored by

$$\begin{aligned} s_l^{\hat{\phi}\hat{\phi}} &= \left| \frac{\delta \sigma_A^{-2}}{\delta \ln C^{\tilde{T}\tilde{T}}_{l,{\rm expt}}} \right|_{{\rm fix}A_L,\,\tilde{g}} \\ &= \sum_{L=l^\phi_{\rm min}}^{l^\phi_{\rm max}} \left| \frac{-2(C_L^{\phi\phi})^2}{\frac{2}{2L+1}(N_L^{(0)} + C_L^{\phi\phi} + N_L^{(1)})^3} \frac{4A_L^2}{2L+1} \right. \\ &\quad \left. \times \sum_{l_1} \tilde{g}^2_{l_1 l}(L) C^{\tilde{T}\tilde{T}}_{l_1,{\rm expt}} \right| C^{\tilde{T}\tilde{T}}_{l,{\rm expt}}, \end{aligned} \tag{A2}$$

where we used σ_A from Eq. (57) and kept only the dominant diagonal part of the reconstruction power autocovariance. As shown in Fig. 12, $s_l^{\hat{\phi}\hat{\phi}}$ (red) is out of phase compared to



(a) $l^{\phi}_{\rm max} = 10$ (b) $l^{\phi}_{\rm max} = 500$

FIG. 12 (color online). Contribution of different CMB multipoles l to the $(S/N)^2$ of the lensing amplitude for estimators based on the CMB power spectrum [blue; Eq. (A1)] and on the lensing reconstruction [red; Eq. (A2)] for $l^{\phi}_{\rm max} = 10$ (left) and $l^{\phi}_{\rm max} = 500$ (right). The approximation in Eq. (A3) for large-scale lenses is shown in black dashed in the left plot. Since we are not interested in the total S/N but only in its distribution over different CMB scales, all curves are normalized such that their integral over l is unity.

$s_l^{\tilde{T}\tilde{T}}$ (blue). To understand this structure, we separate the sums over L and l_1 in Eq. (A2) by restricting ourselves to very large-scale lenses, $l^{\phi}_{\rm max} \ll l$, and using the large-scale approximation for $N^{(0)}$ derived in Ref. [26] [see its Eq. (19)], to find

$$s_l^{\hat{\phi}\hat{\phi}} \approx s^{\hat{\phi}\hat{\phi}}_{l,\rm approx} = f(l^{\phi}_{\rm min}, l^{\phi}_{\rm max}) \frac{1}{{\rm var}_G(C^{\tilde{T}\tilde{T}}_{l,\rm expt})} \left[\left(C_l^{\tilde{T}\tilde{T}} \frac{{\rm d}\ln(l^2 C_l^{\tilde{T}\tilde{T}})}{{\rm d}\ln l} \right)^2 + \frac{1}{2} \left(C_l^{\tilde{T}\tilde{T}} \frac{{\rm d}\ln C_l^{\tilde{T}\tilde{T}}}{{\rm d}\ln l} \right)^2 \right], \qquad [l^{\phi}_{\rm max} \lesssim \mathcal{O}(10)]. \quad \text{(A3)}$$

Here, the prefactor f depends on the minimum and maximum reconstruction multipole but not on the CMB multipole l. The terms in square brackets have the form of the quadrature sum of the information in convergence and shear. Convergence changes locally the angular scale of the CMB anisotropies and so would contribute nothing to the $(S/N)^2$ for a scale-invariant spectrum, $l^2 C_l = {\rm const}$, while shear contributes nothing for a white-noise spectrum, $C_l = {\rm const}$.[15] Thus, for large-scale lenses, the $(S/N)^2$ for the reconstruction-based amplitude gets most contributions from CMB scales where the gradient of the CMB power spectrum is maximal, i.e., between acoustic peaks and troughs (see red and black curves in Fig. 12(a)).

[15]The relation between large-scale lenses and the induced local convergence and shear is discussed in detail by Ref. [45], where the authors also find agreement between the $(S/N)^2$ of a combined convergence and shear estimate with the large-scale limit of the $(S/N)^2$ of the trispectrum reconstruction. This correspondence has also been used to approximate the squeezed limit of the lensing-Integrated-Sachs-Wolfe (ISW) bispectrum [30].

In reality, temperature multipoles that are not precisely at peaks or troughs and not precisely in between them will affect both amplitude estimates, which implies a small amplitude correlation. Intermediate- and small-scale lenses can mix CMB modes over multipole ranges comparable to the acoustic peak separation so that they are affected by wider ranges of CMB multipoles than argued above (see the red curve in Fig. 12(b)), which implies a somewhat larger amplitude correlation. However, since the CMB scales that are most important for the reconstruction still have negligible impact on the amplitude estimated from the temperature power, we expect the correlation of the amplitudes to stay rather small.

2. Cosmic variance of the lenses

We now consider the contribution of cosmic variance of the lenses to the covariance of the lensing amplitude estimates given in Eq. (56). It is instructive to consider a toy model where the reconstruction "noise" power is proportional to $C^{\phi\phi}$, i.e., $N_l^{(0)} + N_l^{(1)} = \lambda C_l^{\phi\phi}$. Taking the limit $\lambda \to 0$ is equivalent to being able to observe ϕ directly with no measurement error, while $\lambda \to \infty$ corresponds to there being no information in the reconstruction. With $N_l^{(0)} + N_l^{(1)} = \lambda C_l^{\phi\phi}$, the weighting of the reconstruction power spectrum in $\hat{A}$ is the same as for an ideal reconstruction (i.e., one with no $N^{(0)}$ and $N^{(1)}$ noise). Provided we then determine $\hat{A}$ from all those ϕ modes that influence the temperature power spectrum, the contribution to the amplitude covariance from cosmic variance of the lenses simplifies significantly to give[16]

[16]Note that $\sum_{l=1}^{l^{\phi}_{\rm max}} (\partial C_{l'}^{\tilde{T}\tilde{T}} / \partial C_l^{\phi\phi}) C_l^{\phi\phi} = C_{l'}^{\tilde{T}\tilde{T}} - C_{l'}^{TT}$ for sufficiently large $l^{\phi}_{\rm max}$.

$$\mathrm{cov}(\hat{A}, \hat{A}') = \sigma^2_{A,\mathrm{ideal}}. \tag{A4}$$

Here, $\sigma^2_{A,\mathrm{ideal}} = [\sum_l (l+1/2)]^{-1} \approx 2/(l^\phi_{\mathrm{max}})^2$ is the variance of the reconstruction-based amplitude in the ideal limit $\lambda \to 0$ using all modes up to l^ϕ_{max}. Since only a few (large-scale) lensing modes affect $\hat{A}'$, including more lensing modes in the reconstruction dilutes the covariation of $\hat{A}$ and $\hat{A}'$ over different realizations of the lenses because there are increasingly more lensing modes in $\hat{A}$ whose fluctuations do not enter $\hat{A}'$. The amplitude covariance falls inversely as the number of modes in the reconstruction since the weight in $\hat{A}$ given to those (few) modes of $C^{\phi\phi}$ that influence the temperature power spectrum falls as the total number of modes. Note that the covariance is independent of the weighting of the measured temperature power spectrum in $\hat{A}'$, provided $\hat{A}'$ is appropriately normalized, and it is also independent of additional contributions to the reconstruction noise (e.g., from CMB cosmic variance, for fixed l^ϕ_{max}). The variance of $\hat{A}$ does depend on the reconstruction noise level, with

$$\begin{aligned}\sigma_A^{-2} &\approx \sum_l \frac{2l+1}{2}\left(\frac{C_l^{\phi\phi}}{C_l^{\phi\phi}+N_l^{(0)}+N_l^{(1)}}\right)^2 \\ &= \frac{1}{(1+\lambda)^2}\sum_l \frac{2l+1}{2} = \frac{1}{(1+\lambda)^2}\,\sigma^{-2}_{A,\mathrm{ideal}}.\end{aligned} \tag{A5}$$

The result $\mathrm{cov}(\hat{A}, \hat{A}') = \sigma^2_{A,\mathrm{ideal}}$ for ideal weighting is necessary to ensure that the lensed CMB spectrum adds no further information on the lensing amplitude when combined with an ideal measurement of ϕ itself on all scales that are relevant for peak smearing of the temperature power spectrum. To see this, note that we can combine the amplitude estimates $\hat{A}$ and $\hat{A}'$ optimally into a single estimate $\hat{A}_{\mathrm{opt}}$, properly taking account of their correlation. If we do this, the inverse variance of the optimal estimate is given by contracting the inverse covariance matrix of the estimates:

$$\sigma^{-2}_{A,\mathrm{opt}} = \mathrm{cov}^{-1}(\hat{A}, \hat{A}) + 2\mathrm{cov}^{-1}(\hat{A}, \hat{A}') + \mathrm{cov}^{-1}(\hat{A}', \hat{A}'). \tag{A6}$$

This evaluates to

$$\sigma^{-2}_{A,\mathrm{opt}} = \frac{\sigma^2_A + \sigma^2_{A'} - 2\sigma^2_{A,\mathrm{ideal}}}{\sigma^2_A\sigma^2_{A'} - \sigma^4_{A,\mathrm{ideal}}}, \tag{A7}$$

on using Eq. (A4) for the covariance. In the ideal case, taking the limit $\lambda \to 0$, we have $\sigma_A = \sigma_{A,\mathrm{ideal}}$ so that $\sigma_{A,\mathrm{opt}} = \sigma_{A,\mathrm{ideal}}$. This is as it must be—the observation of the peak smearing in the power spectrum adds no new information to that obtained from the ideal measurement of ϕ. In the opposite limit, $\lambda \to \infty$, we have $\sigma_A \to \infty$ and $\sigma_{A,\mathrm{opt}} = \sigma_{A'}$ and all information is coming from the temperature power spectrum.

The correlation induced by matter cosmic variance,

$$\mathrm{correl}(\hat{A}, \hat{A}') = \frac{\sigma^2_{A,\mathrm{ideal}}}{\sigma_A\sigma_{A'}} = \frac{\sigma_A}{(1+\lambda)^2\sigma_{A'}}, \tag{A8}$$

reaches its maximal value of $\sigma_{A,\mathrm{ideal}}/\sigma_{A'}$ if the variance of the reconstruction power spectrum is only due to matter cosmic variance, $\lambda = (N_L^{(0)} + N_L^{(1)})/C_L^{\phi\phi} \to 0$; and it falls monotonically with increasing λ, tending to zero as $\lambda \to \infty$ (when CMB cosmic variance dominates the reconstruction uncertainty). This is expected since we assume matter and CMB fluctuations to be independent. More generally, σ_A is determined by the number of high S/N modes in the reconstruction, but $\sigma_{A'}$ depends not only on the number of CMB modes but also the fractional size of the power spectrum corrections from lensing relative to the total spectrum, $C^{\tilde{T}\tilde{T}}_{l,\mathrm{expt}}$. The result is that both factors $\sigma_{A,\mathrm{ideal}}/\sigma_A$ and $\sigma_{A,\mathrm{ideal}}/\sigma_{A'}$ in Eq. (A8) are less than 1, diluting the amplitude correlation.

For our Planck-like parameters, the power spectrum corrections from lensing are only ever a few percent of the total spectrum, and so cosmic variance of the CMB limits $\sigma_{A'} \approx 0.04$. Statistical noise in the lens reconstruction limits $\sigma_A \approx 0.025$. It is clear from Fig. 1 that a constant λ is not a good approximation for lens reconstruction, but we can crudely limit $\lambda \gtrsim 2$, in which case we expect the amplitude correlation to be less than $(0.025/0.04)/3^2 \sim 0.06$, which is close to the value plotted in Fig. 8.

To summarize, the correlation of the lensing amplitudes due to the cosmic variance of the lenses is generally small since there are a limited number of modes of $C^{\phi\phi}$ that influence the acoustic part of the temperature power spectrum (so the covariance for an ideal reconstruction scales inversely as the number of reconstruction modes), and the small covariance [less than $\mathcal{O}(10^{-4})$] is diluted by cosmic variance of the CMB (and noise), which dominates the error on $\hat{A}'$ and contributes significantly to the error on $\hat{A}$. We emphasize that these conclusions assume that the temperature power spectrum at multipoles $l \gtrsim 3000$, where the lensing-induced power from small-scale lenses dominates the unlensed power, does not influence the amplitude estimate (i.e., the spectrum is limited by noise or foregrounds there).

APPENDIX B: OPTIMAL TRISPECTRUM ESTIMATION AND $N^{(0)}$ SUBTRACTION

In this appendix we show that the lensing power spectrum estimator advocated in Sec. II, which includes a data-dependent $N^{(0)}$ subtraction to remove the (disconnected) Gaussian bias, follows naturally from optimal trispectrum estimation (e.g., Ref. [34]). Moreover, we show that in the Gaussian limit, this estimator is uncorrelated with any quadratic estimate of the temperature power spectrum, generalizing the result established in Sec. V A 4.

We start by considering the Gram-Charlier expansion for the probability density function (PDF) of a weakly non-Gaussian, zero-mean temperature field T.[17] Writing the field in an arbitrary basis (e.g., a pixelized map or multipole coefficients) as T_i, with covariance $\langle T_i T_j \rangle = C_{ij}$, the expansion truncated at the trispectrum level can be written in the form [34]

$$P(T) = \frac{e^{-T_i C^{-1}_{ij} T_j/2}}{\sqrt{\det(2\pi C)}}\Big(1 + \frac{1}{24}\langle T_i T_j T_k T_l\rangle_c (\bar{T}_i \bar{T}_j \bar{T}_k \bar{T}_l - [C^{-1}_{ij}\bar{T}_k\bar{T}_l + C^{-1}_{kl}\bar{T}_i\bar{T}_j - C^{-1}_{ij}C^{-1}_{kl}] - [C^{-1}_{ik}\bar{T}_j\bar{T}_l + C^{-1}_{jl}\bar{T}_i\bar{T}_k - C^{-1}_{ik}C^{-1}_{jl}] - [C^{-1}_{il}\bar{T}_j\bar{T}_k + C^{-1}_{jk}\bar{T}_i\bar{T}_l - C^{-1}_{jk}C^{-1}_{il}])\Big). \quad \text{(B1)}$$

Here, $\bar{T}_i \equiv C^{-1}_{ij} T_j$ is the field filtered by the inverse covariance, and summation over repeated indices is implicit. Note that $\langle \bar{T}_i \bar{T}_j \rangle = C^{-1}_{ij}$. We have ignored bispectrum contributions (due to the lensing-ISW correlation), which is a good approximation on intermediate and small scales. Using Wick's theorem, it is straightforward to show from Eq. (B1) that the mean and bispectrum of T vanish, while the covariance is C_{ij}, and the trispectrum is $\langle T_i T_j T_k T_l \rangle_c$, as intended. That the covariance is C_{ij} implies that the data combination associated with the trispectrum term in Eq. (B1) is orthogonal to any quadratic function of T with respect to the Gaussian weight function. In other words, if T were Gaussian, this data combination is uncorrelated with any quadratic function of T.

We can form an estimator for $C_L^{\phi\phi}$ by maximizing the PDF with respect to $C_L^{\phi\phi}$. In practice, we can approximate this by one step of a Newton–Raphson scheme starting at $C_L^{\phi\phi} = 0$. The estimator is then proportional to $\partial \ln P(T)/\partial C_L^{\phi\phi}$ evaluated at $C_L^{\phi\phi} = 0$. The lensing power spectrum enters the PDF explicitly through the connected 4-point function and also implicitly through the lensed $\tilde{C}_l^{TT}$ in the covariance C_{ij}. The contribution to the $C_L^{\phi\phi}$ estimate from the 4-point function is of the form

$$\hat{C}_L^{\phi\phi} \sim \frac{\partial}{\partial C_L^{\phi\phi}}\langle T_i T_j T_k T_l\rangle_c (\bar{T}_i \bar{T}_j \bar{T}_k \bar{T}_l - [C^{-1}_{ij}\bar{T}_k\bar{T}_l + C^{-1}_{kl}\bar{T}_i\bar{T}_j - C^{-1}_{ij}C^{-1}_{kl}] - [C^{-1}_{ik}\bar{T}_j\bar{T}_l + C^{-1}_{jl}\bar{T}_i\bar{T}_k - C^{-1}_{ik}C^{-1}_{jl}] - [C^{-1}_{il}\bar{T}_j\bar{T}_k + C^{-1}_{jk}\bar{T}_i\bar{T}_l - C^{-1}_{jk}C^{-1}_{il}]). \quad \text{(B2)}$$

The terms involving the data in this equation are symmetric under permutations, and so we need to retain only primary couplings of the trispectrum, i.e., the first term in Eq. (13).

[17] In this appendix, to avoid undue clutter in our expressions, we shall suppress the tildes that are used to denote the lensed temperature in the body of this paper.

The primary coupling has the form

$$\frac{\partial}{\partial C_L^{\phi\phi}}\langle T_i T_j T_k T_l\rangle_c^{\text{primary}} \propto \sum_M (-1)^M X^{ij}_{LM} X^{kl}_{L-M}, \quad \text{(B3)}$$

where, in multipole space,

$$X^{l_1 m_1 l_2 m_2}_{LM} = \tilde{f}_{l_1 L l_2}\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & M \end{pmatrix}. \quad \text{(B4)}$$

The power spectrum estimator in Eq. (B2) generalizes that introduced in Sec. II. It reduces to the simple form given there for an isotropic survey. To see this, note that for an isotropic survey, $\bar{T}_{l_1 m_1} = T^*_{l_1 m_1}/C^{\tilde{T}\tilde{T}}_{l_1,\text{expt}}$, and $X^{l_1 m_1 l_2 m_2}_{LM}\bar{T}_{l_1 m_1}\bar{T}_{l_2 m_2}$ is proportional to the estimator $\hat{\phi}_{LM}$ for the lensing potential in Eq. (9). The first set of data-dependent terms in Eq. (B2) can be rewritten as

$$\bar{T}_i\bar{T}_j\bar{T}_k\bar{T}_l - [C^{-1}_{ij}\bar{T}_k\bar{T}_l + C^{-1}_{kl}\bar{T}_i\bar{T}_j - C^{-1}_{ij}C^{-1}_{kl}] = (\bar{T}_i\bar{T}_j - \langle\bar{T}_i\bar{T}_j\rangle)(\bar{T}_k\bar{T}_l - \langle\bar{T}_k\bar{T}_l\rangle). \quad \text{(B5)}$$

Each factor on the right combines with an X_{LM} to give an unnormalized $\hat{\phi}_{LM} - \langle\hat{\phi}_{LM}\rangle$. The "mean-field" term $\langle\hat{\phi}_{LM}\rangle$ vanishes (except for $L = 0$) for a full-sky survey with homogeneous noise. More generally, in the presence of beam asymmetry, realistic noise, and masking, it can require careful subtraction on large scales (see, e.g., Refs. [13,33,46]). The sum of the remaining terms in Eq. (B2) has an expectation value that, for an isotropic survey, is simply $-N_L^{(0)}$ [see Eq. (16)]. However, rather than removing the Gaussian (disconnected) bias from $\hat{C}_L^{\phi\phi}$ with the data-independent $N_L^{(0)}$, the optimal trispectrum estimator debiases with additional data-dependent terms that are equivalent to the $\hat{N}_L^{(0)}$ of Eq. (17). For example,

$$\begin{aligned}\sum_M (-1)^M X^{ij}_{LM} X^{kl}_{L-M} C^{-1}_{ik}\bar{T}_j\bar{T}_l &= 4\sum_M (-1)^{m_1+M}\tilde{g}_{l_1 l_2}(L)\tilde{g}_{l_1 l_4}(L)\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & M\end{pmatrix}\begin{pmatrix} l_1 & l_4 & L \\ -m_1 & m_4 & -M\end{pmatrix} C^{\tilde{T}\tilde{T}}_{l_1,\text{expt}} T^*_{l_2 m_2} T^*_{l_4 m_4} \\ &= 4\sum_{l_1 l_2}\tilde{g}^2_{l_1 l_2}(L) C^{\tilde{T}\tilde{T}}_{l_1,\text{expt}}\Big(\frac{1}{2l_2+1}\sum_{m_2} T_{l_2 m_2} T^*_{l_2 m_2}\Big) \\ &= 4\sum_{l_1 l_2}\tilde{g}^2_{l_1 l_2}(L) C^{\tilde{T}\tilde{T}}_{l_1,\text{expt}}\hat{C}^{\tilde{T}\tilde{T}}_{l_2,\text{expt}}, \end{aligned} \quad \text{(B6)}$$

where summation over (l_1, m_1), (l_2, m_2), and (l_4, m_4) is implicit in the first line.

We showed in Sec. V A 4 that, for an isotropic survey, the data-dependent $N^{(0)}$ subtraction removes the disconnected covariance between $\hat{C}_L^{\hat{\phi}\hat{\phi}}$ and the empirical power spectrum of the measured temperature anisotropies, $\hat{C}^{\tilde{T}\tilde{T}}_{l,\text{expt}}$. For the power spectrum estimator in Eq. (B2), this result

generalizes to anisotropic surveys and also to arbitrary quadratic power spectrum estimates of the temperature anisotropies. The latter are often estimated with combinations of cross-spectra, each of which are quadratic forms constructed from maps $T^{(1)}$ and $T^{(2)}$ for which the noise is independent. Each cross-spectrum therefore involves terms like $T^{(1)}_p T^{(2)}_q$, where p and q are pixel (or multipole) indices, and the fields $T^{(1)}$ and $T^{(2)}$ are correlated with the field that enters the lensing power spectrum estimate at least through the common temperature anisotropies. By using Wick's theorem, it is straightforward to show that the complete set of data-dependent terms in Eq. (B2) has vanishing covariance with $T^{(1)}_p T^{(2)}_q$ in the Gaussian limit.

APPENDIX C: $\hat{N}^{(0)}$ SUBTRACTION AS EFFICIENT MITIGATION OF CHANCE ALIGNMENTS

In this appendix we provide further details to understand why the empirical $\hat{N}^{(0)}$ bias subtraction cancels not only the nondiagonal reconstruction power autocovariance but also the noise contribution to the temperature-lensing power cross-covariance. After identifying chance-aligned terms in the empirical reconstruction power spectrum in Eq. (C2) below (see also Ref. [47]), we will show that avoiding these terms naturally cancels the noise contribution to the auto- and cross-covariance. We then show that empirical $\hat{N}^{(0)}$ subtraction achieves the same but in a faster way (due to the specific multipole couplings that are relevant for the covariances).

Given a CMB realization $\tilde{T}_{lm}$, we split the *empirical* reconstruction power spectrum into two contributions,

$$\begin{aligned}\hat{C}^{\hat{\phi}\hat{\phi}}_L &= \frac{A_L^2}{2L+1}\sum_{\underline{l}_1\cdots\underline{l}_4 M}(-1)^M \begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M\end{pmatrix} \\ &\quad\times\begin{pmatrix} l_3 & l_4 & L \\ m_3 & m_4 & M\end{pmatrix}\tilde{g}_{l_1 l_2}(L)\tilde{g}_{l_3 l_4}(L)\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4} \\ &= \hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}} + \hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{rest}},\end{aligned} \tag{C1}$$

where the noise term contains the part of the sum over l_i where CMB multipoles are chance-aligned,[18]

$$\begin{aligned}\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}} &= \frac{2A_L^2}{2L+1}\sum_{\underline{l}_1\underline{l}_2 m_3 m_4 M}(-1)^M \\ &\quad\times\begin{pmatrix} l_1 & l_2 & L \\ m_1 & m_2 & -M\end{pmatrix}\begin{pmatrix} l_1 & l_2 & L \\ m_3 & m_4 & M\end{pmatrix} \\ &\quad\times\tilde{g}^2_{l_1 l_2}(L)s_{l_1 l_2}\tilde{T}_{l_1 m_1}\tilde{T}_{l_2 m_2}\tilde{T}_{l_1 m_3}\tilde{T}_{l_2 m_4},\end{aligned} \tag{C2}$$

[18]It may be possible to modify the specific splitting of Eq. (C1) chosen here without significantly changing the conclusions below (e.g., by coupling m_3 and m_4 to m_1 and m_2). Our choice allows for a relatively simple analytical assessment of the mean and covariance properties and is sufficiently accurate for the discussion here.

taking only half of the terms for $l_1 = l_2$ by defining $s_{l_1 l_2} = 1 - \delta_{l_1 l_2}/2$. The expectation values are as desired, at subpercent-level accuracy[19],

$$\langle\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}\rangle \approx N^{(0)}_L, \qquad \langle\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{rest}}\rangle \approx C^{\phi\phi}_L + N^{(1)}_L. \tag{C7}$$

To assess the covariance properties of the noise-corrected reconstruction power spectrum $\hat{C}^{\hat{\phi}\hat{\phi}} - \hat{C}^{\hat{\phi}\hat{\phi}}_{\mathrm{noise}}$, note that the $\mathcal{O}(\phi^0)$ noise contribution (35) to the temperature-lensing power cross-covariance is sourced by contractions

$$\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_4} \;\big|\; \tilde{T}_{L'M'}\tilde{T}_{L',-M'}, \tag{C8}$$

which imply $l_1 = l_3$ and $l_2 = l_4$. Since all terms with coinciding multipoles of this form are contained in the noise term (C2), this covariance is cancelled if $\hat{C}^{\hat{\phi}\hat{\phi}} - \hat{C}^{\hat{\phi}\hat{\phi}}_{\mathrm{noise}}$ is used instead of $\hat{C}^{\hat{\phi}\hat{\phi}}$. To see this explicitly, note that

[19]The disconnected part is $\langle\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}\rangle_{\mathrm{disconn}} = N^{(0)}_L$. Using Eq. (14) and simplifying the product of a $3j$- and a $6j$-symbol [48], we have

$$\begin{aligned}\langle\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}\rangle_{\mathrm{conn}} &= \frac{2A_L^2}{(2L+1)^2}\sum_{l_1 l_2}s_{l_1 l_2}\tilde{g}^2_{l_1 l_2}(L)\tilde{f}^2_{l_1 L l_2}C^{\phi\phi}_L + \frac{A_L^2}{16(2L+1)} \\ &\quad\times\sum_{m' l_1 l_2 L'}(-1)^{m'}C^{\phi\phi}_{L'}\frac{s_{l_1 l_2}}{(C^{\tilde{T}\tilde{T}}_{l_1,\mathrm{expt}}C^{\tilde{T}\tilde{T}}_{l_2,\mathrm{expt}})^2} \\ &\quad\times[(\tilde{f}^{(\mathrm{u})}_{l_1 L l_2})^2\tilde{f}^{(\mathrm{u})}_{l_1 L' l_2}\tilde{f}^{\mathrm{no3j}}_{l_1 L' l_2}\mathcal{G}^{l_1 L l_2}_{0-m'm'}\mathcal{G}^{l_1 L l_2}_{-m'm'0}\mathcal{G}^{l_1 L' l_2}_{m'0-m'} \\ &\quad+\tilde{f}^{\mathrm{no3j}}_{l_1 L l_2}\tilde{f}^{(\mathrm{u})}_{l_1 L l_2}\tilde{f}^{(\mathrm{u})}_{l_1 L' l_1}\tilde{f}^{(\mathrm{u})}_{l_2 L' l_2}\mathcal{G}^{l_1 L l_2}_{-m'0m'}\mathcal{G}^{l_1 L' l_1}_{0-m'm'}\mathcal{G}^{l_2 L' l_2}_{-m'm'0}],\end{aligned} \tag{C3}$$

where we defined the Gaunt coefficients, $\tilde{f}^{(\mathrm{u})}$ with unsymmetric terms and $\tilde{f}^{\mathrm{no3j}}$ with no $3j$-symbols, by

$$\begin{aligned}\mathcal{G}^{l_1 l_2 l_3}_{m_1 m_2 m_3} &\equiv \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \\ &\quad\times\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3\end{pmatrix},\end{aligned} \tag{C4}$$

$$\begin{aligned}\tilde{f}^{(\mathrm{u})}_{l_1 L l_2} &\equiv [L(L+1) - l_1(l_1+1) + l_2(l_2+1)]C^{\tilde{T}\tilde{T}}_{l_2} + (l_1 \leftrightarrow l_2), \\ \tilde{f}^{\mathrm{no3j}}_{l_1 L l_2} &\equiv \sqrt{\frac{(2l_1+1)(2L+1)(2l_2+1)}{16\pi}}\tilde{f}^{(\mathrm{u})}_{l_1 L l_2}.\end{aligned} \tag{C5}$$

We find that Eq. (C3) is smaller than $C^{\phi\phi}_L + N^{(1)}_L$ by a factor of more than 400 for any L, so that [with Eq. (15), $q_L \lesssim 0.002$]

$$\begin{aligned}\langle\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}\rangle &= N^{(0)}_L + q_L(C^{\phi\phi}_L + N^{(1)}_L), \\ \langle\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{rest}}\rangle &= (1-q_L)(C^{\phi\phi}_L + N^{(1)}_L).\end{aligned} \tag{C6}$$

$$\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})_{\mathrm{disconn}} = \mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})_{\mathrm{disconn}}. \tag{C9}$$

The dominant nondiagonal contribution of Eq. (28) to the autocovariance of the uncorrected reconstruction power $\hat{C}^{\hat{\phi}\hat{\phi}}$ is due to couplings (see Eq. (46c) in Ref. [26])

$$\overbrace{\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_3}}\,\overbrace{\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_4}\;|\;\tilde{T}_{\underline{l}_7}\tilde{T}_{\underline{l}_5}}\,\overbrace{\tilde{T}_{\underline{l}_6}\tilde{T}_{\underline{l}_8}}, \tag{C10}$$

which implies $l_1 = l_3$ and $l_6 = l_8$. Writing out the contractions following Ref. [26] and summing over m_1 and M in Eq. (42) of Ref. [26] enforces $l_2 = l_4$, and therefore also $l_5 = l_7$. Since all terms of the reconstruction power (C1) with these coinciding multipoles are in the noise term (C2), the autocovariance of $\hat{C}^{\hat{\phi}\hat{\phi}} - \hat{C}^{\hat{\phi}\hat{\phi}}_{\mathrm{noise}}$ does not contain the nondiagonal contribution (28). In contrast, the covariance contributions that involve a product of two trispectra and lead to the dominant diagonal variance (29) [26] are still present, so that

$$\begin{aligned}&\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L} - \hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}, \hat{C}^{\hat{\phi}\hat{\phi}}_{L'} - \hat{C}^{\hat{\phi}\hat{\phi}}_{L',\mathrm{noise}})\\ &\quad\approx \delta_{LL'}\frac{2}{2L+1}\langle C^{\hat{\phi}\hat{\phi}}_{L}\rangle^2.\end{aligned} \tag{C11}$$

Thus, the chance-aligned noise terms (C2), which lead to the $N^{(0)}$ bias, are responsible for both the dominant nondiagonal reconstruction power autocovariance (28) and the noise contribution (35) to the temperature-lensing power cross-covariance. It is therefore desirable to avoid these noise terms. In practice, a brute-force way to achieve this would be to subtract $\hat{C}^{\hat{\phi}\hat{\phi}}_{\mathrm{noise}}$ from $\hat{C}^{\hat{\phi}\hat{\phi}}$ directly or, equivalently, to restrict the summation over the l_i in Eq. (C1) appropriately. A more efficient method is obtained from a "partial-averaging" procedure, which will turn out to be equivalent to the empirical $\hat{N}^{(0)}$ subtraction.[20] Guided by the fact that both covariances (28) and (25) are due to disconnected terms containing the contraction

$$\overbrace{\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_3}} \tag{C12}$$

[see Eqs. (C8) and (C10)], we take the corresponding expectation value already at the level of evaluating Eq. (C2), before computing covariances, by defining the partial averaging operation, $\mathcal{R}$, by

$$\mathcal{R}[\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}] \equiv \langle\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\rangle\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4} + 5\ \text{perms}; \tag{C13}$$

i.e., we average out two of the four modes while keeping the unaveraged realization of the other two modes. This leaves contractions of the form (C8) in the disconnected (4 + 2)-point function and contractions of the form (C10) in the disconnected (4 + 4)-point function invariant in the sense that

$$\begin{aligned}&\langle\mathcal{R}[\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}]\tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle^{(\mathrm{C8})\ \mathrm{terms}}_{\mathrm{disconn}}\\ &\quad= \langle\tilde{T}_{\underline{l}_1}\tilde{T}_{\underline{l}_2}\tilde{T}_{\underline{l}_3}\tilde{T}_{\underline{l}_4}\tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle^{(\mathrm{C8})\ \mathrm{terms}}_{\mathrm{disconn}}\end{aligned} \tag{C14}$$

and

$$\begin{aligned}&\langle\mathcal{R}[\tilde{T}_{\underline{l}_1}\cdots\tilde{T}_{\underline{l}_4}]\mathcal{R}[\tilde{T}_{\underline{l}_5}\cdots\tilde{T}_{\underline{l}_8}]\rangle^{(\mathrm{C10})\ \mathrm{terms}}_{\mathrm{disconn}}\\ &\quad= \langle\mathcal{R}[\tilde{T}_{\underline{l}_1}\cdots\tilde{T}_{\underline{l}_4}](\tilde{T}_{\underline{l}_5}\cdots\tilde{T}_{\underline{l}_8})\rangle^{(\mathrm{C10})\ \mathrm{terms}}_{\mathrm{disconn}}\\ &\quad= \langle\tilde{T}_{\underline{l}_1}\cdots\tilde{T}_{\underline{l}_8}\rangle^{(\mathrm{C10})\ \mathrm{terms}}_{\mathrm{disconn}}.\end{aligned} \tag{C15}$$

Therefore, both covariance contributions (28) and (35) are still eliminated if instead of subtracting the full noise term (C2) from the reconstruction power we subtract the partial average of this noise term, i.e., if we consider (up to realization-independent bias mitigation terms) $\hat{C}^{\hat{\phi}\hat{\phi}}_{L} - \mathcal{R}[\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}]$. We find that this partially averaged noise mitigation reduces to the empirical $\hat{N}^{(0)}$ subtraction,

$$\mathcal{R}[\hat{C}^{\hat{\phi}\hat{\phi}}_{L,\mathrm{noise}}] = 2\hat{N}^{(0)}_{L}. \tag{C16}$$

Thus, the empirical $\hat{N}^{(0)}$ subtraction can be interpreted as an efficient method to mitigate disconnected (auto- and cross-) covariance contributions generated by the chance-aligned noise terms (C2) in the reconstruction power spectrum. In contrast to the noise terms (C2), the empirical $\hat{N}^{(0)}$ defined in Eq. (17) can be evaluated very efficiently because the empirical temperature power spectrum is isolated in the sum.

APPENDIX D: TEMPERATURE-LENSING POWER COVARIANCE FROM THE CMB TRISPECTRUM

We will show here that contributions from the lensed CMB trispectrum to the temperature-lensing power covariance have a subdominant effect on parameter estimation compared to the fully disconnected $\mathcal{O}(\phi^0)$ contribution and the $\mathcal{O}(\phi^4)$ contribution from the connected 6-point function. The connected 4-point function contributes to Eq. (33) with couplings of the form $\langle\tilde{T}_1\tilde{T}_3\rangle\langle\tilde{T}_2\tilde{T}_4\tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle_c$ ("type A") and $\langle\tilde{T}_1\tilde{T}_{L'M'}\rangle\times\langle\tilde{T}_2\tilde{T}_3\tilde{T}_4\tilde{T}_{L',-M'}\rangle_c$ ("type B"). Other couplings either cancel in Eq. (33) or vanish because $\langle\hat{\phi}_{LM}\rangle = 0$. The contribution from type A can be expressed nonperturbatively in terms of the derivative in Eq. (31) and the connected 4-point contribution to the temperature power autocovariance as

$$\begin{aligned}&\mathrm{cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_{L}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})^{\mathrm{conn\ 4\ pt\ A}}\\ &\quad= \sum_{l_2}\frac{\partial(2\hat{N}^{(0)}_{L})}{\partial\hat{C}^{\tilde{T}\tilde{T}}_{l_2,\mathrm{expt}}}\mathrm{cov}(\hat{C}^{\tilde{T}\tilde{T}}_{l_2,\mathrm{expt}}, \hat{C}^{\tilde{T}\tilde{T}}_{L',\mathrm{expt}})^{\mathrm{conn\ 4\ pt}}.\end{aligned} \tag{D1}$$

Perturbatively, up to $\mathcal{O}(\phi^4)$, we have

[20] An alternative would be "phase randomization" following Ref. [10]. Alternatively, one could split $\tilde{T}_{lm}$ into in- and out-annuli [47], which, however, reduces the signal-to-noise [12]. Note that while the goal of Ref. [47] was to avoid the Gaussian $N^{(0)}$ noise bias for any CMB realization, our goal is to simplify the reconstruction power auto- and cross-covariance.

$$\mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}})^{\text{conn 4 pt A}} = \sum_{l_2} \frac{\partial(2\hat{N}_L^{(0)})}{\partial \hat{C}_{l_2,\mathrm{expt}}^{\tilde{T}\tilde{T}}} [\mathrm{cov}(\hat{C}_{l_2,\mathrm{expt}}^{\tilde{T}\tilde{T}}, \hat{C}_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}})|_{\mathcal{O}(\phi^2)} + \mathrm{cov}(\hat{C}_{l_2,\mathrm{expt}}^{\tilde{T}\tilde{T}}, \hat{C}_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}})|_{\mathcal{O}(\phi^4)}], \tag{D2}$$

where the perturbative temperature covariances are given by Eq. (24). The covariance (D2) can be interpreted as the correction to the noise contribution of Eq. (36) due to the nondiagonal $\mathcal{O}(\phi^2)$ and $\mathcal{O}(\phi^4)$ parts of the temperature power autocovariance (24). The correlation corresponding to Eq. (D2) is at most 5×10^{-5}, which is 2 orders of magnitude smaller than the dominant noise contribution. The induced correlation of the lensing amplitude estimates $\hat{A}$ and $\hat{A}'$ is less than 0.15% (see Fig. 8); i.e., the effect of (D2) is negligible.

The type-A contribution to the covariance is removed by the empirical $\hat{N}^{(0)}$ correction introduced in Sec. II. As for the disconnected contribution to the covariance (see Appendix B), this is actually a more general result that applies for anisotropic surveys and for an arbitrary quadratic estimate of the temperature power spectrum. To see this, we generalize to the $\hat{C}_L^{\phi\phi}$ estimator of Eq. (B2) and to quadratic temperature power spectrum estimates of the form $T_p^{(1)} T_q^{(2)}$. The part of the $\hat{C}_L^{\phi\phi}$ estimator that is quartic in the data has a type-A, 4-point covariance with $T_p^{(1)} T_q^{(2)}$ given by

$$\mathrm{cov}(\bar{T}_i \bar{T}_j \bar{T}_k \bar{T}_l, T_p^{(1)} T_q^{(2)})^{\text{conn 4 pt A}} = \langle \bar{T}_i \bar{T}_j T_p^{(1)} T_q^{(2)} \rangle_c \langle \bar{T}_k \bar{T}_l \rangle + 5 \text{ terms}, \tag{D3}$$

where the other five terms are inequivalent permutations of i, j, k, and l. The part of the $\hat{C}_L^{\phi\phi}$ estimator that is quadratic in the data is composed of six terms of the form $-\langle \bar{T}_i \bar{T}_j \rangle \bar{T}_k \bar{T}_l$. The contribution from the connected 4-point function to the covariance of these terms with $T_p^{(1)} T_q^{(2)}$ exactly cancels with the terms of the right-hand side of Eq. (D3), removing the type-A covariance.

Terms of coupling type B have contributions from the primary trispectrum term, where sums over m_i simplify due to orthogonality relations of $3j$ symbols, and from nonprimary trispectrum terms, where sums over m_i lead to a nontrivial $6j$ symbol:

$$\mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}})^{\text{conn 4 pt B}}_{\text{primary}} = 2 \frac{C_L^{\phi\phi}}{A_L} \frac{\partial(2\hat{N}_L^{(0)})}{\partial \hat{C}_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}}} \frac{2}{2L'+1} (C_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}})^2 = 2 \frac{C_L^{\phi\phi}}{A_L} \mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}})_{\mathrm{disconn}}, \tag{D4}$$

$$\mathrm{cov}(\hat{C}_L^{\hat{\phi}\hat{\phi}}, \hat{C}_{L',\mathrm{expt}}^{\tilde{T}\tilde{T}})^{\text{conn 4 pt B}}_{\text{non-primary}} = \sum_{\substack{m'L'' \\ l_2 l_3 l_4}} \frac{A_L^2 (-1)^{m'} C_{L''}^{\phi\phi} \tilde{f}^{\text{no3j}}_{L'Ll_2} \tilde{f}^{(\mathrm{u})}_{l_3 Ll_4} \tilde{f}^{(\mathrm{u})}_{L'L''l_3} \tilde{f}^{(\mathrm{u})}_{l_2 L'' l_4} \mathcal{G}^{l_3 L l_4}_{-m'0m'} \mathcal{G}^{L'L''l_3}_{0-m'm'} \mathcal{G}^{l_2 L'' l_4}_{0m'-m'}}{4(2L+1)(2L'+1) C_{l_2,\mathrm{expt}}^{\tilde{T}\tilde{T}} C_{l_3,\mathrm{expt}}^{\tilde{T}\tilde{T}} C_{l_4,\mathrm{expt}}^{\tilde{T}\tilde{T}}} + (l_3 \leftrightarrow l_4). \tag{D5}$$

Here we used Eq. (14) and expressed the product of a $6j$ symbol with a $3j$ symbol as a sum over a product of three $3j$ symbols [48]. The $\tilde{f}$ factors are defined in Appendix C. Neither contributions in Eqs. (D4) and (D5) is cancelled by empirical $\hat{N}^{(0)}$ subtraction.

The correlation of unbinned power spectra from the primary term of Eq. (D4) is shown in Fig. 13. It is at most 0.08%, which is almost a factor of two larger than the maximum of the matter cosmic variance contribution [Eq. (E8)] shown in Fig. 5(a). The structure is very similar to that of the disconnected (noise) contribution [Eq. (35)] shown in Fig. 4(a), but with the additional signal-to-noise factor of $C_L^{\phi\phi}/A_L$ that falls off rapidly for $L > 200$. (Recall $A_L = N_L^{(0)}$ for our choice of optimal weights.) Therefore, the contribution of the primary type-B covariance to correlations between the lensing amplitudes $\hat{A}$ and $\hat{A}'$ is suppressed compared to that of the disconnected noise contribution, reaching at most 2% (see Fig. 8). It is also suppressed compared to the matter cosmic variance contribution because $\hat{A}'$ gives most weight to CMB modes at the acoustic peaks and troughs where the primary type-B power correlation is small.

While the measured correlation of $\hat{A}$ and $\hat{A}'$ is consistent with our theoretical expectations, the measurements are too noisy to test subdominant terms such as that in Eq. (D4). In Fig. 4(c), the lowest L_ϕ pixels show hints of the structure plotted in Fig. 13, but the measurements are, again, too noisy to provide a conclusive test of Eq. (D4). The simulations do, however, imply that there cannot be significantly larger covariance contributions than the ones we model analytically.

Although the summation in the expression (D5) for the nonprimary type-B terms is restricted by triangle inequalities, its evaluation is still numerically challenging. We could only evaluate Eq. (D5) for 30 (L, L') pairs at $L \lesssim 400$, which is relevant for the current Planck lensing likelihood [13]. At these points the power correlation from Eq. (D5) is at most around 4×10^{-5}, which is 1 order of magnitude smaller than the maximum correlation from the primary term [Eq. (D4)]. The structure of the correlation matrix seems similar to that of the primary term in the L_T direction but seems to peak toward higher L_ϕ (but we cannot assess the structure reliably from the small number

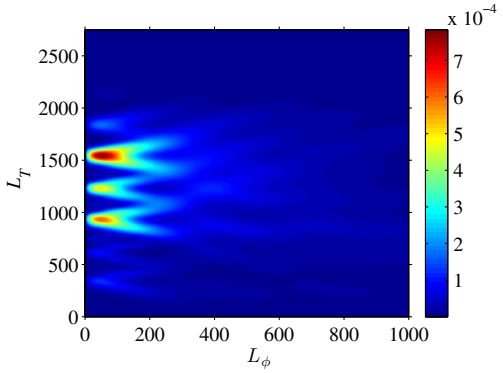


FIG. 13 (color online). Theoretical contribution from the primary type-B connected 4-point function to the correlation of the unbinned power spectra of the lens reconstruction and the (lensed) temperature power spectrum. The covariance (D4) is converted to a correlation using the same conversion factor as in Eq. (37). Correlations for $L_\phi > 1000$ are too small to be visible in this plot.

of evaluation points). Moreover, given that our simulations do not show any significant excess covariance, we expect the nonprimary term of Eq. (D5) to be negligible.

APPENDIX E: TEMPERATURE-LENSING POWER COVARIANCE FROM THE CONNECTED CMB 6-POINT FUNCTION

The $\mathcal{O}(\phi^2)$ contribution to the connected part of the lensed temperature 6-point function vanishes [43]. Here we will compute the $\mathcal{O}(\phi^4)$ terms. There are five types of terms, which involve $\delta^4 T$, $\delta^3 T \delta T$, $\delta^2 T \delta^2 T$, $\delta^2 T \delta T \delta T$, or $\delta T \delta T \delta T \delta T$ with the appropriate number of factors of the unlensed CMB. Only the last two types of terms contribute to the connected 6-point function; the remaining terms contribute to the full 6-point function but are cancelled when the connected 4-point and fully disconnected parts are removed.

Writing out only the two relevant types of terms, the full 6-point function at $\mathcal{O}(\phi^4)$ is ($\tilde{T}_i \equiv \tilde{T}_{l_i m_i}$)

$$\langle \tilde{T}_1 \tilde{T}_2 \tilde{T}_3 \tilde{T}_4 \tilde{T}_5 \tilde{T}_6 \rangle^{(4)} = \frac{1}{12}[\langle \delta^2 T_1 \delta T_2 \delta T_3 T_4 T_5 T_6 \rangle + \text{all perms}] + \frac{1}{48}[\langle \delta T_1 \delta T_2 \delta T_3 \delta T_4 T_5 T_6 \rangle + \text{all perms}]. \tag{E1}$$

From this we must subtract the following terms involving the connected 4-point function,

$$\begin{aligned} &\frac{1}{48}[\langle \tilde{T}_1 \tilde{T}_2 \tilde{T}_3 \tilde{T}_4 \rangle_c \langle \tilde{T}_5 \tilde{T}_6 \rangle + \text{all perms}] \\ &\quad = \frac{1}{48}[6\langle \delta T_1 \delta T_2 T_3 T_4 \rangle \langle \delta T_5 \delta T_6 \rangle + 12\langle \delta T_1 \delta T_2 T_3 T_4 \rangle \langle \delta^2 T_5 T_6 \rangle + 12\langle \delta^2 T_1 \delta T_2 \delta T_3 T_4 \rangle \langle T_5 T_6 \rangle \\ &\qquad + \langle \delta T_1 \delta T_2 \delta T_3 \delta T_4 \rangle \langle T_5 T_6 \rangle - (9\langle \delta T_1 \delta T_2 \rangle + 24\langle \delta^2 T_1 T_2 \rangle) \langle \delta T_3 \delta T_4 \rangle \langle T_5 T_6 \rangle + \text{all perms}] + \cdots, \end{aligned} \tag{E2}$$

and the fully disconnected part,

$$\begin{aligned} \langle \tilde{T}_1 \tilde{T}_2 \tilde{T}_3 \tilde{T}_4 \tilde{T}_5 \tilde{T}_6 \rangle^{(4)}_{\text{disc}} &= \frac{1}{2}[\langle \delta^2 T_1 \delta T_2 \rangle \langle \delta T_3 T_4 \rangle \langle T_5 T_6 \rangle + \text{all perms}] + \frac{1}{4}[\langle \delta^2 T_1 T_2 \rangle \langle \delta T_3 \delta T_4 \rangle \langle T_5 T_6 \rangle + \text{all perms}] \\ &\quad + \frac{1}{2}[\langle \delta^2 T_1 T_2 \rangle \langle \delta T_3 T_4 \rangle \langle \delta T_5 T_6 \rangle + \text{all perms}] + \frac{1}{16}[\langle \delta T_1 \delta T_2 \rangle \langle \delta T_3 \delta T_4 \rangle \langle T_5 T_6 \rangle + \text{all perms}] \\ &\quad + \frac{1}{4}[\langle \delta T_1 \delta T_2 \rangle \langle \delta T_3 T_4 \rangle \langle \delta T_5 T_6 \rangle + \text{all perms}] + \cdots. \end{aligned} \tag{E3}$$

To evaluate these expressions, we use, for example,

$$\langle \delta T_1 \delta T_2 T_3 T_4 \rangle = \langle \delta T_1 \delta T_2 \rangle \langle T_3 T_4 \rangle + \langle \langle \delta T_1 T_3 \rangle_{\text{CMB}} \langle \delta T_2 T_4 \rangle_{\text{CMB}} \rangle_{\text{LSS}} + \langle \langle \delta T_1 T_4 \rangle_{\text{CMB}} \langle \delta T_2 T_3 \rangle_{\text{CMB}} \rangle_{\text{LSS}}, \tag{E4}$$

where $\langle\rangle_{\text{CMB}}$ denotes averaging over unlensed CMB realizations and $\langle\rangle_{\text{LSS}}$ denotes averaging over realizations of the lensing potential ϕ. Rewriting the other 4- and 6-point functions in a similar way gives the following final expression for the connected 6-point function:

$$\begin{aligned} \langle \tilde{T}_1 \tilde{T}_2 \tilde{T}_3 \tilde{T}_4 \tilde{T}_5 \tilde{T}_6 \rangle^{(4)}_c &= \frac{1}{2}\text{cov}_{\text{LSS}}(\langle \delta^2 T_1 T_2 \rangle_{\text{CMB}}, \langle \delta T_3 T_4 \rangle_{\text{CMB}} \langle \delta T_5 T_6 \rangle_{\text{CMB}}) + \text{all perms} \\ &\quad + \frac{1}{4}\text{cov}_{\text{LSS}}(\langle \delta T_1 \delta T_2 \rangle_{\text{CMB}}, \langle \delta T_3 T_4 \rangle_{\text{CMB}} \langle \delta T_5 T_6 \rangle_{\text{CMB}}) + \text{all perms} \\ &= \frac{1}{16}\text{cov}_{\text{LSS}}(\langle \tilde{T}_1 \tilde{T}_2 \rangle^{(2)}_{\text{CMB}}, \langle \tilde{T}_3 \tilde{T}_4 \rangle^{(1)}_{\text{CMB}} \langle \tilde{T}_5 \tilde{T}_6 \rangle^{(1)}_{\text{CMB}}) + \text{all perms}. \end{aligned} \tag{E5}$$

The result is written in terms of the covariance over different realizations of the lensing potential, $\text{cov}_{\text{LSS}}(X, Y) \equiv \langle XY \rangle_{\text{LSS}} - \langle X \rangle_{\text{LSS}} \langle Y \rangle_{\text{LSS}}$.

We expect only a few of the 720 permutations in Eq. (E5) to be relevant for the lensing-temperature power covariance [Eq. (33)]. The weights $\tilde{g}$ in Eq. (33) impose triangle conditions on l_1, l_2, L and l_3, l_4, L, which constrain the summation volume (especially at low L where the lensing power dominates). Terms that couple, e.g., $\langle\tilde{T}_1\tilde{T}_3\rangle_{\rm CMB}$ impose an additional triangle constraint on l_1, l_3, and a lensing multipole $\tilde{L}$, which reduces the summation volume further. This is not the case for couplings of the type $\langle\tilde{T}_1\tilde{T}_2\rangle_{\rm CMB}$ and $\langle\tilde{T}_3\tilde{T}_4\rangle_{\rm CMB}$. For this reason we expect the dominant terms to come from the couplings 12, 34, and $(L', M'; L', -M')$ in Eq. (33), i.e., terms which factor most under the weights [26]. The dominant contribution of the connected 6-point function at $\mathcal{O}(\phi^4)$ to the lensing-temperature power covariance is therefore expected to come from

$$(-1)^{M'}\langle\tilde{T}_1\tilde{T}_2\tilde{T}_3\tilde{T}_4\tilde{T}_{L'M'}\tilde{T}_{L',-M'}\rangle^{(4)}_{c,\rm dom} = {\rm cov}_{\rm LSS}(\langle\tilde{T}_{L'M'}\tilde{T}^*_{L'M'}\rangle^{(2)}_{\rm CMB}, \langle\tilde{T}_1\tilde{T}_2\rangle^{(1)}_{\rm CMB}\langle\tilde{T}_3\tilde{T}_4\rangle^{(1)}_{\rm CMB}) + 2{\rm cov}_{\rm LSS}(\langle\tilde{T}_1\tilde{T}_2\rangle^{(2)}_{\rm CMB}, \langle\tilde{T}_{L'M'}\tilde{T}^*_{L'M'}\rangle^{(1)}_{\rm CMB}\langle\tilde{T}_3\tilde{T}_4\rangle^{(1)}_{\rm CMB}). \quad \text{(E6)}$$

Here, we have used the result that permutations of the form $1 \leftrightarrow 2$, $3 \leftrightarrow 4$, $(12) \leftrightarrow (34)$ and $(L', M') \leftrightarrow (L', -M')$ all lead to the same contribution to Eq. (33).

The $\langle\tilde{T}_1\tilde{T}_2\rangle^{(1)}_{\rm CMB}$ term on the right of Eq. (E6) evaluates to

$$\langle\tilde{T}_1\tilde{T}_2\rangle^{(1)}_{\rm CMB} = \sum_{L_1M_1}(-1)^{M_1}\begin{pmatrix} l_1 & l_2 & L_1 \\ m_1 & m_2 & -M_1\end{pmatrix} f_{l_1L_1l_2}\phi_{L_1M_1}, \quad \text{(E7)}$$

which combines with the weights, normalization, and one of the $3j$ symbols in Eq. (33) to give $\phi_{LM} + \mathcal{O}(\phi^2)$. Similarly, the $\langle\tilde{T}_3\tilde{T}_4\rangle^{(1)}_{\rm CMB}$ term returns $\phi_{L,-M}$. Equation (E7) also shows that the second term on the right of Eq. (E6) does not contribute to the power covariance since $\langle\tilde{T}_{L'M'}\tilde{T}^*_{L'M'}\rangle^{(1)}_{\rm CMB}$ returns the monopole of ϕ on summing over M'.

Putting these pieces together, and using Eq. (44), finally gives

$$\begin{aligned}{\rm cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_L, \hat{C}^{\tilde{T}\tilde{T}}_{L'})_{\langle\tilde{T}_1\cdots\tilde{T}_6\rangle^{(4)}_{c,\rm dom}} &= \sum_{L''}\frac{\partial C^{\tilde{T}\tilde{T}}_{L'}}{\partial C^{\phi\phi}_{L''}}{\rm cov}_{\rm LSS}(\hat{C}^{\phi\phi}_{L''}, \hat{C}^{\phi\phi}_L) + \mathcal{O}(\phi^5) \\ &= \frac{2}{2L+1}(C^{\phi\phi}_L)^2\frac{\partial C^{\tilde{T}\tilde{T}}_{L'}}{\partial C^{\phi\phi}_L} + \mathcal{O}(\phi^5). \end{aligned} \quad \text{(E8)}$$

APPENDIX F: EFFECT OF $C^{T\phi}$

In this appendix we discuss the contribution of the ISW-induced large-scale $C^{T\phi}$ correlation on the temperature-lensing power covariance. We first compute corrections due to the ISW-lensing bispectrum and then the ones due to corrections of the lensed temperature power spectrum.

The lensed temperature 6-point function in the temperature-lensing power covariance of Eq. (33) involves the following 3-point terms,

$$\langle\tilde{T}_{l_1}\tilde{T}_{l_2}\tilde{T}_{l_3}\tilde{T}_{l_4}\tilde{T}_{L'M'}\tilde{T}_{L'-M'}\rangle|_{3\text{-pt},C^{T\phi}} = \langle\tilde{T}_{l_1}\tilde{T}_{l_2}\tilde{T}_{l_3}\rangle^{\rm ISW}\langle\tilde{T}_{l_4}\tilde{T}_{L'M'}\tilde{T}_{L'-M'}\rangle^{\rm ISW} + 9\,\text{perms}, \quad \text{(F1)}$$

where the nonperturbative ISW-lensing bispectrum is approximately given by [30,49]

$$\langle\tilde{T}_{l_1}\tilde{T}_{l_2}\tilde{T}_{l_3}\rangle^{\rm ISW} = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3\end{pmatrix} B^{\rm ISW}_{l_1l_2l_3} = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3\end{pmatrix} C^{T\phi}_{l_1}\tilde{f}_{l_2l_1l_3} + 2\,\text{perms}. \quad \text{(F2)}$$

While the contribution to Eq. (33) from the coupling $(123)(4L'L')$, which is written out explicitly in Eq. (F1), vanishes on summing over M', the couplings of type $(12L')(34L')$ yield the diagonal covariance contribution[21]

$$\begin{aligned}{\rm cov}(\hat{C}^{\hat{\phi}\hat{\phi}}_L, \hat{C}^{\tilde{T}\tilde{T}}_{L',\rm expt})|^{(12L'),(34L')}_{3\text{-pt},C^{T\phi}} &= \delta_{LL'}\frac{2A_L^2}{(2L+1)^3}\left[\sum_{l_1l_2}\tilde{g}_{l_1l_2}(L)B^{\rm ISW}_{l_1l_2L}\right]^2 \\ &\approx \delta_{LL'}\frac{2}{2L+1}(C^{T\phi}_L)^2. \end{aligned} \quad \text{(F3)}$$

The approximation on the right is simply the covariance between the power spectra of the input lensing potential and the lensed temperature, an intuitive result that we might have anticipated. For our Planck-like parameters, the corresponding power correlation is always less than 5% and rapidly decreases with increasing multipoles, being less than 0.1% for $L \geq 60$ (see Fig. 14). The induced correlation of the lensing amplitudes A and A' is less than 10^{-5} and therefore negligible. This small correlation arises because most of the information on the lensing amplitude from the temperature power spectrum

[21]To obtain the approximate result, in the last step we use Eq. (10) and neglect all terms depending on $C^{T\phi}_{l_1}$ or $C^{T\phi}_{l_2}$ because they enforce l_1 or l_2 to be small, which reduces the remaining summation volume due to the triangle condition on l_1, l_2, and L. We find that using the approximation instead of the full expression leads a temperature-lensing correlation which is wrong by at most 3×10^{-6} for $L \leq 300$.

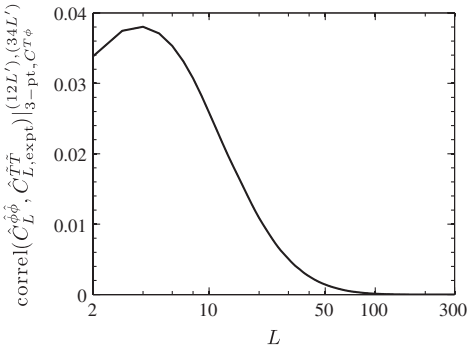


FIG. 14. ISW-lensing bispectrum contribution of Eq. (F3) to the diagonal unbinned temperature-lensing power correlation. The difference between the full expression and the approximation in Eq. (F3) is too small to be visible in this plot.

comes from small scales. The power correlation due to the ISW-lensing effect is small on these scales, and, additionally, there is limited information in the lensing reconstruction on such scales as the signal-to-noise is very low there. We expect that couplings of the type $(13L')(24L')$ in Eq. (F1), which do not factor under the weights, are further suppressed as they limit the summation volume.

The ISW-induced change in the lensed temperature power spectrum,

$$\Delta C_L^{\tilde{T}\tilde{T}} = C_L^{\tilde{T}\tilde{T}} - C_L^{\tilde{T}\tilde{T}}|_{C^{T\phi}=0} = \frac{1}{2L+1}\sum_{L_1,L_2} F^2_{LL_1L_2} C^{T\phi}_{L_1} C^{T\phi}_{L_2}, \tag{F4}$$

leads to a correction of the disconnected noise contribution [Eq. (35)] to the temperature-lensing power covariance, which is also second order in $C^{T\phi}$. The correction has a similar structure to the noise contribution itself (shown in Fig. 4(a)), but it is around 10^6 times smaller and induces a lensing amplitude correlation of $\mathcal{O}(10^{-9})$ that is totally negligible.

[1] A. Lewis and A. Challinor, Phys. Rep. **429**, 1 (2006).
[2] M. Zaldarriaga and U. Seljak, Phys. Rev. D **59**, 123507 (1999).
[3] W. Hu, Astrophys. J. Lett. **557**, L79 (2001).
[4] W. Hu, Phys. Rev. D **65**, 023003 (2001).
[5] M. Kaplinghat, L. Knox, and Y.-S. Song, Phys. Rev. Lett. **91**, 241301 (2003).
[6] L. Verde and D. N. Spergel, Phys. Rev. D **65**, 043007 (2002).
[7] V. Acquaviva, C. Baccigalupi, and F. Perrotta, Phys. Rev. D **70**, 023515 (2004).
[8] E. Calabrese, A. Slosar, A. Melchiorri, G. F. Smoot, and O. Zahn, Phys. Rev. D **77**, 123531 (2008).
[9] A. C. Hall and A. Challinor, Mon. Not. R. Astron. Soc. **425**, 1170 (2012).
[10] S. Das *et al.*, Phys. Rev. Lett. **107**, 021301 (2011).
[11] S. Das *et al.*, arXiv:1301.1037.
[12] A. van Engelen *et al.*, Astrophys. J. **756**, 142 (2012).
[13] P. A. R. Ade *et al.* (Planck Collaboration), arXiv:1303.5077.
[14] W. Hu, Phys. Rev. D **64**, 083005 (2001).
[15] U. Seljak, Astrophys. J. **463**, 1 (1996).
[16] K. T. Story *et al.*, arXiv:1210.7231.
[17] P. A. R. Ade *et al.*, (Planck Collaboration), arXiv:1303.5076.
[18] C. M. Hirata and U. Seljak, Phys. Rev. D **67**, 043001 (2003).
[19] B. D. Sherwin *et al.*, Phys. Rev. Lett. **107**, 021302 (2011).
[20] J. Lesgourgues, L. Perotto, S. Pastor, and M. Piat, Phys. Rev. D **73**, 045021 (2006).
[21] D. M. Goldberg and D. N. Spergel, Phys. Rev. D **59**, 103002 (1999).
[22] J. R. Bond, A. H. Jaffe, and L. Knox, Astrophys. J. **533**, 19 (2000).
[23] L. Verde *et al.*, Astrophys. J. Suppl. Ser. **148**, 195 (2003).
[24] S. Smith, A. Challinor, and G. Rocha, Phys. Rev. D **73**, 023517 (2006).
[25] S. Hamimeche and A. Lewis, Phys. Rev. D **77**, 103013 (2008).
[26] D. Hanson, A. Challinor, G. Efstathiou, and P. Bielewicz, Phys. Rev. D **83**, 043005 (2011).
[27] A. Challinor and G. Chon, Phys. Rev. D **66**, 127301 (2002).
[28] W. Hu, Phys. Rev. D **62**, 043007 (2000).
[29] W. Hu and T. Okamoto, Astrophys. J. **574**, 566 (2002).
[30] A. Lewis, A. Challinor, and D. Hanson, J. Cosmol. Astropart. Phys. 03 (2011) 018.
[31] T. Okamoto and W. Hu, Phys. Rev. D **67**, 083002 (2003).
[32] M. Kesden, A. Cooray, and M. Kamionkowski, Phys. Rev. D **67**, 123507 (2003).
[33] T. Namikawa, D. Hanson, and R. Takahashi, Mon. Not. R. Astron. Soc. **431**, 609 (2013).
[34] D. M. Regan, E. P. S. Shellard, and J. R. Fergusson, Phys. Rev. D **82**, 023520 (2010).
[35] NASA, http://lambda.gsfc.nasa.gov/.
[36] A. Lewis, Phys. Rev. D **71**, 083008 (2005).
[37] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. **622**, 759 (2005).
[38] K. M. Smith, W. Hu, and M. Kaplinghat, Phys. Rev. D **74**, 123002 (2006).
[39] C. Li, T. L. Smith, and A. Cooray, Phys. Rev. D **75**, 083501 (2007).
[40] A. Benoit-Lévy, K. M. Smith, and W. Hu, Phys. Rev. D **86**, 123008 (2012).
[41] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000).

[42] A. Challinor and A. Lewis, Phys. Rev. D **71**, 103010 (2005).
[43] M. Kesden, A. Cooray, and M. Kamionkowski, Phys. Rev. D **66**, 083007 (2002).
[44] O. Zahn *et al.* (work in progress).
[45] M. Bucher, C. S. Carvalho, K. Moodley, and M. Remazeilles, Phys. Rev. D **85**, 043016 (2012).
[46] D. Hanson, G. Rocha, and K. Górski, Mon. Not. R. Astron. Soc. **400**, 2169 (2009).
[47] B. D. Sherwin and S. Das, arXiv:1011.4510.
[48] NIST Digital Library of Mathematical Functions, Release 1.0.5 of 2012-10-01, online companion to [50], http://dlmf.nist.gov/.
[49] D. Hanson, K. M. Smith, A. Challinor, and M. Liguori, Phys. Rev. D **80**, 083004 (2009).
[50] *NIST Handbook of Mathematical Functions*, edited by F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (Cambridge University Press, Cambridge, England, 2010), print companion to Ref. [48].