Improving small-scale CMB lensing reconstruction

Boryana Hadzhiyska, Blake D. Sherwin, Mathew Madhavacheril, and Simone Ferraro

Improving Small-Scale CMB Lensing Reconstruction

Boryana Hadzhiyska,[1, 2] Blake D. Sherwin,[2, 3] Mathew Madhavacheril,[4] and Simone Ferraro[5, 6]

[1]*Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA*
[2]*Department of Applied Mathematics and Theoretical Physics,*
University of Cambridge, Wilberforce Road, Cambridge CB1 2AD, United Kingdom
[3]*Kavli Institute for Cosmology Cambridge, University of Cambridge,*
Wilberforce Road, Cambridge CB1 2AD, United Kingdom
[4]*Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*
[5]*Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA*
[6]*Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA*

Over the past decade, the gravitational lensing of the Cosmic Microwave Background (CMB) has become a powerful tool for probing the matter distribution in the Universe. The standard technique used to reconstruct the CMB lensing signal employs the quadratic estimator (QE) method, which has recently been shown to be suboptimal for lensing measurements on very small scales in temperature and polarization data. We implement a simple, more optimal method for the small-scale regime, which involves taking the direct inverse of the background gradient. We derive new techniques to make continuous maps of lensing using this "Gradient-Inversion" (GI) method and validate our method with simulated data, finding good agreement with predictions. For idealized simulations of lensing cross- and autospectra that neglect foregrounds, we demonstrate that our method performs significantly better than previous quadratic estimator methods in temperature; at $L = 5000 - 9000$, it reduces errors on the lensing auto-power spectrum by a factor of ~ 4 for both idealized CMB-S4 and Simons Observatory-like experiments and by a factor of ~ 2.6 for cross-correlations of CMB-S4-like lensing reconstruction and the true lensing field. We caution that the level of the neglected small-scale foreground power, while low in polarization, is very high in temperature; though we briefly outline foreground mitigation methods, further work on this topic is required. Nevertheless, our results show the future potential for improved small-scale CMB lensing measurements, which could provide stronger constraints on cosmological parameters and astrophysics at high redshifts.

I. INTRODUCTION

Building on progress in CMB experimentation and theory, the gravitational lensing of the cosmic microwave background (CMB) has recently emerged as a powerful cosmological probe. The CMB lensing signal arises from the deflection of CMB photons as they pass through the matter distribution between the surface of last scattering and us; this lensing deflection induces subtle, non-Gaussian correlations of the CMB anisotropies in both temperature and polarization. Reconstructing and analyzing a lensing map from CMB data can provide a wealth of information on the sum of the neutrino masses, the equation of state of dark energy, the properties of inflation and the early universe (via delensing), and astrophysics at high redshift (through cross-correlations) [1, 2]. Rapid progress in the extraction of the lensing signal has been made only recently with the first detections in cross- and auto-correlation in 2007 [3] and 2011 respectively [4], followed by significantly improved measurements by the SPT, *Planck*, SPTpol and PO-LARBEAR collaborations [5–9]. While the field has already advanced significantly, CMB lensing science still has great potential, with large increases in the lensing signal-to-noise ratio expected from ongoing and upcoming ground-based CMB experiments such as AdvancedACTPol, SPT-3G, POLARBEAR-II, Simons Observatory and CMB-S4.

Lensing is most commonly reconstructed using the quadratic estimator (QE) proposed by Hu and Okamoto [10]. It was shown by [11–13] that in the low-noise regime in polarization, the QE is suboptimal, as it is only an approximation to the much more computationally expensive optimal (maximum likelihood) solution. While it was long thought that the QE was sufficient in temperature, we recently pointed out that this is not the case and that the QE is, in fact, suboptimal on small angular scales in temperature [12]. For small-scale reconstruction from temperature, the QE approach is hindered by the fact that the precision of the reconstruction is limited by cosmic variance of the background CMB gradient, whereas a true optimal solution, in principle, should not be. Furthermore, its errors become highly correlated on small-scales, so that measuring more modes does not offer significant improvements when one accounts for the covariance between them properly. More optimal methods suffer neither of these problems on small scales. We recently showed that on very small scales the so called "Gradient-Inversion" (GI) approach, the idea of which was first proposed by [14], can avoid the QE cosmic-variance limit and is close to optimal for reconstruction of these small-scale lenses [12]. However, in our previous work, we only discussed the relevant estimator assuming the approximation of a small, constant-gradient patch of sky, and only analyzed its impact on cluster lensing reconstruction.

In this work, we derive the most general GI lensing algorithm, allowing the reconstruction of a large-area, continuous small-scale lensing map from CMB data; we

characterize this estimator's performance for both lensing auto- and cross-spectra. Our paper is structured as follows. In section II we briefly review the QE and its limitations, contrasting it with the GI method proposed by Horowitz et al. [12]. In section III we present a new derivation of the continuous GI estimator for small-scale lensing reconstruction and describe the implementation of our algorithm. Finally, we focus on the signal-to-noise ratio improvement over the standard QE method and discuss systematic error challenges and potential applications of our methods.

II. MOTIVATION

The lensing of the CMB can be described by a remapping of the unlensed CMB tempearture field T to give a lensed field $\tilde{T}$:

$$\tilde{T}(\mathbf{x}) = T(\mathbf{x} + \nabla\phi) \approx T(\mathbf{x}) + \nabla T \cdot \nabla\phi, \qquad (1)$$

where we have expanded to first order in the lensing potential ϕ. Although this first-order approximation can be poor at the field level, it works quite well on small enough scales, where the primary temperature fluctuations are severely suppressed by diffusion damping. We can find which scales contribute to the gradient of the temperature by computing its variance up to some multipole $L_{\max}$

$$G_{\mathrm{rms}}^2(< L_{\max}) = \int_0^{L_{\max}} \frac{d^2\mathbf{L}}{(2\pi)^2} \, L^2 C_L^{TT}, \qquad (2)$$

where we are working in the flat sky approximation. As discussed in [12, 15], the gradient variance saturates at $L_G \approx 2000$. Therefore, if we restrict ourselves to small lenses where $l \gg 2000$, the first-order approximation of Eq. 1 should be very good. This is the regime we will concentrate on in this work.

Neglecting the unlensed temperature in this small-scale regime due to Silk damping, the gradient inversion (GI) solution for the lensing potential in a small patch of sky with a single (constant) background gradient was proposed by [14] and extended by [12]:

$$\hat{\phi}_{\mathrm{GI}}(\mathbf{L}) = \frac{\tilde{T}(\mathbf{L})}{i\mathbf{L} \cdot \nabla T} \, , \qquad (3)$$

where ∇T is the constant gradient of the unlensed CMB temperature in the small patch of sky considered. If we recall that $\tilde{T}$ contains instrument noise n as well as signal, i.e. that $\tilde{T} = \tilde{T}^{true} + n$ (where n has noise power spectrum N_L^{TT}), we can see that the noise level depends on the local gradient:

$$N_{\mathbf{L}}^{\phi\phi} \sim \frac{N_L^{TT}}{(\mathbf{L} \cdot \nabla T)^2}. \qquad (4)$$

We note that in the first expression for $\hat{\phi}_{\mathrm{GI}}(\mathbf{L})$, $\tilde{T}(\mathbf{L})$ and

∇T can in principle be measured with arbitrary accuracy assuming low enough noise, if we make the approximation that the effects of lensing on the large scale gradient are small, i.e. that ∇T can be measured as $\nabla\tilde{T} \approx \nabla T$.

In the small-scale limit, the QE estimator has the following form:

$$\hat{\phi}_{\mathrm{QE}}(\mathbf{L}) \approx \frac{\tilde{T}(\mathbf{L})}{L^2} \frac{\nabla T \cdot \mathbf{L}}{\frac{1}{2}|\nabla T|_{\mathrm{rms}}^2} \approx \hat{\phi}_{\mathrm{GI}}(\mathbf{L}) \frac{(\nabla T \cdot \mathbf{L})^2}{\langle(\nabla T \cdot \mathbf{L})^2\rangle_{\mathrm{CMB}}}, \qquad (5)$$

where the notation $\langle...\rangle_{\mathrm{CMB}}$ refers to an averaging over realizations of the unlensed CMB. In contrast to the GI estimator, the standard QE does something suboptimal in the low-noise limit: rather than "correctly" dividing out the gradient, it multiplies by the gradient and divides by the gradient RMS to preserve correct normalization. This implies that the QE is limited by an unnecessary cosmic variance error in the low-noise limit; in addition, the error is correlated across different modes which share a common gradient.

While, under the approximations outlined previously, the GI method is close to optimal in the small-scale limit, it has thus far only been defined for a constant gradient. In the next section we will discuss the extension to varying gradients.

III. THE CONTINUOUS GRADIENT INVERSION ESTIMATOR

The standard gradient inversion estimator was derived in the limit of a very small map with a constant gradient. In this section, we will attempt to make our derivation of the estimator form more rigorous and more generally applicable, extending our lensing reconstruction algorithm to large, real maps. Our strategy will be to relate our final target – Fourier space quantities such as power spectra – to an intermediate quantity involving real space correlation functions, where we can more easily apply approximations that are naturally made in real space, such as assuming slowly varying local gradients.

A. Deriving the estimator

We will begin by deriving a large-map GI estimator in a simplified and suboptimal manner, neglecting instrumental noise and foregrounds and weighting all regions of the map equally. (A better weighting that maximizes signal-to-noise will be derived later.) Let us define the estimator:

$$\hat{\phi}(\mathbf{x}) \equiv \tilde{T}(\mathbf{x})_{\mathrm{high\text{-}pass}} f(\nabla T(\mathbf{x})), \qquad (6)$$

where $f(\nabla T)$ is a weight function that only depends on $\mathbf{x}$ via the slowly varying gradient field (which is assumed to be well-known). Initially, our goal will be to derive f such that it gives an unbiased lensing reconstruction.

To begin, we may write

$$\hat{\phi}(\mathbf{x}) = \sum_j \tilde{T}(\mathbf{x})_{\text{high-pass}} f(\nabla T(\mathbf{x})) P_j(\mathbf{x}), \qquad (7)$$

where we have introduced "large-pixel" window functions P_j, which are much larger than the lensing scales of interest, but are sufficiently small that we can assume the temperature gradient is constant within the large pixels. The pixel windows P_j are zero outside the j-th pixel and equal to one inside. Since the map is high-pass filtered, we can assume that $\tilde{T}(\mathbf{x})_{\text{high-pass}} \approx \nabla\phi \cdot \nabla T$ so that

$$\hat{\phi}(\mathbf{x}) = \sum_j \nabla\phi(\mathbf{x}) \cdot \nabla T(\mathbf{x}) f(\nabla T(\mathbf{x})) P_j(\mathbf{x}). \qquad (8)$$

We can similarly write the true ϕ field in terms of the large pixelization:

$$\phi(\mathbf{x}) = \sum_j \phi(\mathbf{x}) P_j(\mathbf{x}). \qquad (9)$$

To correctly reconstruct the field ϕ using an estimator $\hat{\phi}$, our requirement is that

$$C_{\mathbf{L}}^{\phi\hat{\phi}} = C_{\mathbf{L}}^{\phi\phi}. \qquad (10)$$

We can write this power spectrum in terms of the correlation function

$$C_{\mathbf{L}}^{\phi\hat{\phi}} = \int d^2r \, e^{i\mathbf{L}\cdot\mathbf{r}} \langle \phi(\mathbf{x}+\mathbf{r})\hat{\phi}(\mathbf{x}) \rangle_{\text{all}}, \qquad (11)$$

where $\langle ... \rangle_{\text{all}}$ indicates both i) an ensemble average over realizations of small-scale CMB, lensing and noise and ii) a spatial average over $\mathbf{x}$ within a very large region with many independent gradients. However, throughout this analysis, we will assume the gradients $\nabla T(\mathbf{x})$ to be fixed and, as stated previously, assume them to be well known. In some cases it will be helpful to average over small-scale fields, denoted $\langle ... \rangle_{\text{s}}$, and $\mathbf{x}$, marked $\langle ... \rangle_{\mathbf{x}}$, in separate steps. (This implies $\langle ... \rangle_{\text{all}} = \langle \langle ... \rangle_{\text{s}} \rangle_{\mathbf{x}}$.)

Now we may evaluate the real space correlation function, which will allow us to make a number of relevant approximations. Averaging only over small-scale fields in a first step, we obtain

$$\langle \phi(\mathbf{x}+\mathbf{r})\hat{\phi}(\mathbf{x}) \rangle_{\text{s}} = \sum_{j,k} P_j(\mathbf{x}) P_k(\mathbf{x}+\mathbf{r}) \times \qquad (12)$$
$$\langle \phi(\mathbf{x}+\mathbf{r})\nabla\phi(\mathbf{x}) \cdot \nabla T(\mathbf{x}) f(\nabla T(\mathbf{x})) \rangle_{\text{s}}.$$

We may assume that, since we have high pass filtered to remove long wavelength modes in our estimator (and are, in any case, only considering lensing on small scales, which approximately corresponds to very small $\mathbf{r}$ correlations), only correlations within the same pixel $j = k$ are significantly non-zero, and we will neglect all others.

Therefore only the $j = k$ terms contribute to the sum. In addition, we may assume that within one pixel j, the temperature gradient is not a varying field, but approximately takes on a constant value ∇T_j. This gives us:

$$\langle \phi(\mathbf{x}+\mathbf{r})\hat{\phi}(\mathbf{x}) \rangle_{\text{s}} \approx \sum_j P_j^2(\mathbf{x}) f(\nabla T_j) \langle \phi(\mathbf{x}+\mathbf{r})\nabla\phi(\mathbf{x}) \rangle_{\text{s}} \cdot \nabla T_j. \qquad (13)$$

We now write the potential-potential gradient correlation function in Fourier space to give:

$$\langle \phi(\mathbf{x}+\mathbf{r})\hat{\phi}(\mathbf{x}) \rangle_{\text{s}} \approx \sum_j P_j^2(\mathbf{x}) \int \frac{d^2k}{(2\pi)^2} e^{-i\mathbf{k}\cdot\mathbf{r}} C_{\mathbf{k}}^{\phi\phi} f(\nabla T_j) i\mathbf{k} \cdot \nabla T_j. \qquad (14)$$

We can now connect this expression to the definition of the power spectrum in terms of the correlation function. For this purpose, we average Eq. (14) over $\mathbf{x}$, i.e. add $\langle ... \rangle_{\mathbf{x}}$, to give

$$\langle \phi(\mathbf{x}+\mathbf{r})\hat{\phi}(\mathbf{x}) \rangle_{\text{all}} \approx \qquad (15)$$
$$\sum_j \langle P_j^2(\mathbf{x}) \rangle_{\mathbf{x}} \int \frac{d^2k}{(2\pi)^2} e^{-i\mathbf{k}\cdot\mathbf{r}} C_{\mathbf{k}}^{\phi\phi} f(\nabla T_j) \, i\mathbf{k} \cdot \nabla T_j$$

and insert this two point correlation function into the power spectrum expression of Eq. (11), noting also that the mean value for P_j^2 over $\mathbf{x}$ is simply $1/N_{\text{pix}}$, i.e. one over the number of pixels. We thus obtain

$$C_{\mathbf{L}}^{\phi\hat{\phi}} = \qquad (16)$$
$$\frac{1}{N_{\text{pix}}} \sum_j \int \frac{d^2k}{(2\pi)^2} \int d^2r \, e^{i\mathbf{L}\cdot\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}} C_{\mathbf{k}}^{\phi\phi} f(\nabla T_j) \, i\mathbf{k} \cdot \nabla T_j.$$

Performing the $\mathbf{r}$ integral, we obtain $(2\pi)^2 \delta^{(D)}(\mathbf{L}-\mathbf{k})$; this allows us to perform the k integral to give:

$$C_{\mathbf{L}}^{\phi\hat{\phi}} = \frac{1}{N_{\text{pix}}} \sum_j C_{\mathbf{L}}^{\phi\phi} f(\nabla T_j) \, i\mathbf{L} \cdot \nabla T_j. \qquad (17)$$

In other words, our cross-correlation measurement is equal to the true cross-correlation in each large pixel, weighted by $f(\nabla T_j)(i\mathbf{L} \cdot \nabla T_j)$ and averaged over all pixels. It can clearly be seen that if we set $f(\nabla T_j) = \frac{1}{i\mathbf{L}\cdot\nabla T_j}$ we obtain

$$C_{\mathbf{L}}^{\phi\hat{\phi}} = C_{\mathbf{L}}^{\phi\phi}, \qquad (18)$$

as required, with each large-pixel region contributing equally to the cross-spectrum.

We note that for the correct reconstruction of each mode of the lensing $\phi_{\mathbf{L}}$, we require a different, specific weight function $f^{\mathbf{L}} \equiv \frac{1}{i\mathbf{L}\cdot\nabla T_j}$. We similarly denote the estimator that uses this particular weight function as $\hat{\phi}^{\mathbf{L}}(\mathbf{x})$; this estimator is correctly normalized to recover

the mode $\mathbf{L}$; i.e., if we Fourier Transform $\hat{\phi}^{\mathbf{L}}(\mathbf{x})$, then the mode $\hat{\phi}^{\mathbf{L}}(\mathbf{L}) \equiv \hat{\phi}_{\mathbf{L}}^{\mathbf{L}}$ is correctly reconstructed. Here we have the understanding that a superscripted $^{\mathbf{L}}$ implies that a spatial weighting $f^{\mathbf{L}}$ is being applied with the purpose of reconstructing ϕ at a fixed target wavenumber $\mathbf{L}$. This is to be distinguished from a subscript $_{\mathbf{L}}$, which is the true Fourier conjugate variable of the coordinate $\mathbf{x}$ describing the real space map. The two variables will be equal if the correct spatial weighting has been applied in order to reconstruct each mode, as will be shown below to be the case in all realistic applications.

In the previously derived simple estimator, all pixels of the map make equal contributions to the final lensing power spectrum. However, a more optimal estimator should up-weight pixels with a large gradient, as these have a larger signal $\nabla T \cdot \nabla \phi$. We will now derive such a more optimally weighted estimator.

Our derivation proceeds similarly to that presented previously. The small-scale temperature is now assumed to be described by $\tilde{T}(\mathbf{x})_{\text{high-pass}} = \nabla \phi \cdot \nabla T + n$, where noise n has a power spectrum N_L^{TT}. We again consider the estimator

$$\hat{\phi}^{\mathbf{L}}(\mathbf{x}) \equiv \tilde{T}(\mathbf{x})_{\text{high-pass}} f^{\mathbf{L}}(\nabla T). \quad (19)$$

As before, we obtain for the cross-correlation of the reconstruction with the input:

$$C_{\mathbf{L}}^{\phi\hat{\phi}^{\mathbf{L}}} = \frac{1}{N_{\text{pix}}} \sum_j C_{\mathbf{L}}^{\phi\phi} \times (i\mathbf{L} \cdot \nabla T_j) f^{\mathbf{L}}(\nabla T_j). \quad (20)$$

Similarly, we can obtain an expression for the reconstruction power spectrum

$$C_{\mathbf{L}}^{\hat{\phi}^{\mathbf{L}}\hat{\phi}^{\mathbf{L}}} = \frac{1}{N_{\text{pix}}} \sum_j [-C_{\mathbf{L}}^{\phi\phi} \times (i\mathbf{L}\cdot\nabla T_j)^2 + N_L^{TT}](f^{\mathbf{L}}(\nabla T_j))^2. \quad (21)$$

Given these cross- and auto-spectra, we can find the weight function $f^{\mathbf{L}}$ that minimizes the variance $\text{Var}(\hat{\phi}_{\mathbf{L}})$ of our new estimator, which we define as:

$$\langle(\hat{\phi}_{\mathbf{L}} - \phi_{\mathbf{L}})(\hat{\phi}_{\mathbf{L}'} - \phi_{\mathbf{L}'})^*\rangle_{\text{all}} \equiv \text{Var}(\hat{\phi}_{\mathbf{L}})(2\pi)^2\delta^{(D)}(\mathbf{L} - \mathbf{L}')$$
$$= (C_{\mathbf{L}}^{\hat{\phi}^{\mathbf{L}}\hat{\phi}^{\mathbf{L}}} - 2C_{\mathbf{L}}^{\phi\hat{\phi}^{\mathbf{L}}} + C_{\mathbf{L}}^{\phi\phi})(2\pi)^2\delta^{(D)}(\mathbf{L} - \mathbf{L}'). \quad (22)$$

We must minimize the reconstruction variance $\text{Var}(\hat{\phi}_{\mathbf{L}})$ subject to the normalization constraint that $C_{\mathbf{L}}^{\phi\hat{\phi}^{\mathbf{L}}} = C_{\mathbf{L}}^{\phi\phi}$, which is equivalent to

$$A_{\mathbf{L}}[f^{\mathbf{L}}] \equiv \frac{1}{N_{\text{pix}}} \sum_j (i\mathbf{L} \cdot \nabla T_j) f^{\mathbf{L}}(\nabla T_j) = 1. \quad (23)$$

We can perform this minimization using a Lagrange multiplier λ

$$\mathcal{L}[f^{\mathbf{L}}] \equiv \quad\quad\quad\quad\quad\quad\quad\quad\quad (24)$$
$$C_{\mathbf{L}}^{\hat{\phi}^{\mathbf{L}}\hat{\phi}^{\mathbf{L}}}[f^{\mathbf{L}}] - 2C_{\mathbf{L}}^{\phi\hat{\phi}^{\mathbf{L}}}[f^{\mathbf{L}}] + C_{\mathbf{L}}^{\phi\phi} - \lambda_{\mathbf{L}}(A_{\mathbf{L}}[f^{\mathbf{L}}] - 1),$$

where $C_{\mathbf{L}}^{\hat{\phi}^{\mathbf{L}}\hat{\phi}^{\mathbf{L}}}[f^{\mathbf{L}}]$ and $C_{\mathbf{L}}^{\phi\hat{\phi}^{\mathbf{L}}}[f^{\mathbf{L}}]$ are as in Eqs. 20 and 21 above.

Setting $\delta\mathcal{L}/\delta f^{\mathbf{L}} = 0$ and then imposing the normalization condition $A_{\mathbf{L}}[f^{\mathbf{L}}] = 1$ to solve for λ, after some algebra we obtain our final weighting for our estimator

$$f^{\mathbf{L}}(\nabla T_j) = \frac{(i\mathbf{L} \cdot \nabla T_j)}{[-C_{\mathbf{L}}^{\phi\phi}(i\mathbf{L} \cdot \nabla T_j)^2 + N_L^{TT}]} \quad (25)$$
$$\times \left[\frac{1}{N_{\text{pix}}} \sum_k \frac{(i\mathbf{L} \cdot \nabla T_k)^2}{[-C_{\mathbf{L}}^{\phi\phi}(i\mathbf{L} \cdot \nabla T_k)^2 + N_L^{TT}]}\right]^{-1}.$$

Approximating this as a continuous function, we obtain

$$f^{\mathbf{L}}(\mathbf{x}) \approx \frac{(i\mathbf{L} \cdot \nabla T(\mathbf{x}))}{[-C_{\mathbf{L}}^{\phi\phi}(i\mathbf{L} \cdot \nabla T(\mathbf{x}))^2 + N_L^{TT}]} \quad (26)$$
$$\times \left[\left\langle \frac{(i\mathbf{L} \cdot \nabla T(\mathbf{x}))^2}{[-C_{\mathbf{L}}^{\phi\phi}(i\mathbf{L} \cdot \nabla T(\mathbf{x}))^2 + N_L^{TT}]} \right\rangle_{\mathbf{x}}\right]^{-1}.$$

Therefore, our final estimator for lensing (factoring out and cancelling several factors) is:

$$\hat{\phi}^{\mathbf{L}}(\mathbf{x}) \approx \frac{\tilde{T}(\mathbf{x})_{\text{high-pass}}/(i\mathbf{L} \cdot \nabla T(\mathbf{x}))}{[C_{\mathbf{L}}^{\phi\phi} + \frac{N_L^{TT}}{(\mathbf{L}\cdot\nabla T(\mathbf{x}))^2}]} \times \quad (27)$$
$$\left[\left\langle \left[C_{\mathbf{L}}^{\phi\phi} + \frac{N_L^{TT}}{(\mathbf{L} \cdot \nabla T(\mathbf{x}))^2}\right]^{-1} \right\rangle_{\mathbf{x}}\right]^{-1}.$$

To optimally recover a mode $\phi_{\mathbf{L}}$, we must simply take the Fourier transform (FT) of $\hat{\phi}^{\mathbf{L}}(\mathbf{x})$ and select the mode with wavenumber $\mathbf{L}$; i.e. we must evaluate $\hat{\phi}_{\mathbf{L}}^{\mathbf{L}} \equiv \text{FT}[\hat{\phi}^{\mathbf{L}}](\mathbf{L})$. The Fourier transform of $\hat{\phi}^{\mathbf{L}}(\mathbf{x})$, of course, contains a full array of other pixels $\hat{\phi}_{\neq\mathbf{L}}^{\mathbf{L}}$, but we discard these as our spatial weighting $f^{\mathbf{L}}(\mathbf{x})$ is not optimal or unbiased for these modes.

In addition, we note that while the estimator contains the unobservable true unlensed gradient, we approximate it by low-pass filtering the observed, lensed map (removing all power above $\ell > 2000$) and taking a real-space derivative operation. The tests described later in this work verify that this approximation does not introduce a bias. Here the second line represents the gradient-dependent normalization that corrects for the weighting of different regions.

Our estimator can be interpreted as an unbiased estimator $\tilde{T}(\mathbf{x})_{\text{high-pass}}/(i\mathbf{L} \cdot \nabla T(\mathbf{x}))$ multiplied by a weight function that increases the contribution from large-gradient, high signal-to-noise regions. Building on this intuition, our final estimator can be more compactly represented as

$$\hat{\phi}^{\mathbf{L}}(\mathbf{x}) \approx W^{\mathbf{L}}(\mathbf{x})\frac{\tilde{T}(\mathbf{x})_{\text{high-pass}}}{(i\mathbf{L} \cdot \nabla T(\mathbf{x}))} \times [\langle W^{\mathbf{L}}(\mathbf{x})\rangle_{\mathbf{x}}]^{-1}, \quad (28)$$

where we have defined the optimal spatial weighting function as $W^{\mathbf{L}}(\mathbf{x}) \equiv \left[C_{\mathbf{L}}^{\phi\phi} + \frac{N_L^{TT}}{(\mathbf{L}\cdot\nabla T(\mathbf{x}))^2}\right]^{-1}$. However, we

will justify in subsequent paragraphs why, to validate our methods using simulations, we use a slightly suboptimal weight $W^{\mathbf{L}}(\mathbf{x}) = \left[\frac{N_L^{TT}}{(\mathbf{L} \cdot \nabla T(\mathbf{x}))^2} \right]^{-1}$. It can be seen that, as expected, $W^{\mathbf{L}}$ is simply a spatial window function that upweights high signal-to-noise regions (where the gradient is large) and downweights low signal-to-noise ones.

We note that in the signal-dominated limit, Eq. (28) reduces to the naive gradient inversion estimator.

In the noisy limit, when reconstructing over a large area with many independent gradients, this appears at first glance very similar to the standard quadratic estimator. However, in fact, there is a significant difference to the quadratic estimator, namely how the estimator is normalized. For a single, constant gradient, we have argued previously that the quadratic estimator is suboptimally normalized, making an error of order $\frac{(\mathbf{L} \cdot \nabla T)^2_{\text{true}}}{\langle (\mathbf{L} \cdot \nabla T) \rangle^2_{\text{CMB}}}$. Naively, spatially averaging over x would drive the GI normalization to the RMS value used in the quadratic estimator. However, the normalization error is reduced only by the square root of the number of independent gradients that are averaged over; since the error on all lensing observables averages down by the same amount, the *relative* impact of this mis-normalization compared to the lensing spectrum error bars is not reduced by spatial averaging. We hence expect the local mis-normalization, arising from the gradient cosmic variance error, to remain a key limiting factor for the quadratic estimator. The GI estimator does not suffer from this limitation.

The GI estimator we have derived and seek to implement, as shown in Eq. (28) makes use of a spatial weight (or mask) function W; in the optimal case, this should be given by $W^{\mathbf{L}}(\mathbf{x}) = \left[C_{\mathbf{L}}^{\phi\phi} + \frac{N_L^{TT}}{(\mathbf{L} \cdot \nabla T(\mathbf{x}))^2} \right]^{-1}$, effectively a Wiener filter. However, we encountered significant difficulties in implementing this optimal weighting function for low noise levels. A likely source of our implementation problems was the fact that our estimator derivation is only valid for weights which are spatially slowly varying. For very low noise levels, however, the window function varies very rapidly: though W is typically equal to $[C_{\mathbf{L}}^{\phi\phi}]^{-1}$, along directions where the gradient is perpendicular to the wavevector $\mathbf{L}$, W very rapidly takes on a zero weight, leading to sharp edges in the window function. To regularize the weight, we modified our filtering to be a simple inverse noise filter:

$$W^{\mathbf{L}}(\mathbf{x}) \equiv \left[\frac{N_L^{TT}}{(\mathbf{L} \cdot \nabla T(\mathbf{x}))^2} \right]^{-1}. \qquad (29)$$

For noise-dominated modes, this is already (near-) optimal; for signal dominated modes, this leads to moderate loss in the area used, which typically only has a modest impact on signal-to-noise. As is visible from the form of the window function, the theoretical prediction for the noise on our estimator for a particular fixed, local gradi-

ent is given by

$$N_{\mathbf{L}}^{\hat{\phi}\hat{\phi}}(\mathbf{x}) = \frac{N_L^{TT}}{(\mathbf{L} \cdot \nabla T(\mathbf{x}))^2}. \qquad (30)$$

The modification to this expression in the case of multiple gradients is discussed in more detail in Appendix A along with other subtleties regarding the noise of our estimator.

B. Applying the estimator: summary of the algorithm for lensing reconstruction

In this section we will briefly outline the estimation procedure, Eq. 28, used for reconstruction of the lensing potential field for a given simulated lensed image using the GI method.

For each target wavevector $\mathbf{L} \equiv (L_x, L_y)$ on a discrete FFT grid given by spacing $(n_i 2\pi/S_x, n_j 2\pi/S_y)$, where S_x and S_y are the map dimensions (in this case, $S_x = S_y = 1024 \times 0.1$ arcmin) and $n_{i,j}$ are integers ranging over the number of pixels, our procedure for obtaining the map and the power spectrum is described by the following steps:

(i) We obtain the gradient field of the temperature as follows. We first filter the observed lensed temperature map, i.e. we remove the small-scale information ($L > 2000$), where the gradient saturates. We then apply a real-space numerical differentiation (using second order accurate central differences) to this filtered field in order to obtain the gradient. We denote this gradient of the map as $\nabla \tilde{T}_{\text{filt}}(\mathbf{x})$ and for the subsequent steps of the algorithm, substitute $\nabla T(\mathbf{x})$, appearing in the derivations in Section III, with it, as the true unlensed $\nabla T(\mathbf{x})$ is not observable.

(ii) Construct the unweighted estimator $\hat{\phi}^{\mathbf{L}}(\mathbf{x})_{\text{raw}} = \frac{\tilde{T}(\mathbf{x})_{\text{high-pass}}}{i\mathbf{L} \cdot \nabla \tilde{T}_{\text{filt}}(\mathbf{x})}$ as defined in Eq. 6. At a given position $\mathbf{x}$, the value of the estimator depends only on the local gradient $\nabla \tilde{T}_{\text{filt}}(\mathbf{x})$ and the value of the lensed temperature $\tilde{T}(\mathbf{x})$. (Note that zeros of the gradient are not problematic after the weighting in the next step is applied.)

(iii) Apply an inverse noise weighting $W^{\mathbf{L}}(\mathbf{x}) = [N_{\mathbf{L}}^{\hat{\phi}\hat{\phi}}(\mathbf{x})]^{-1}$ to it, i.e. compute $W^{\mathbf{L}}(\mathbf{x}) \times \hat{\phi}^{\mathbf{L}}(\mathbf{x})_{\text{raw}}$, where the window function is given in Eq. 29. The motivation behind choosing this form for the filtering function is that it upweights large-gradient, high signal regions; it also downweights the modes perpendicular to the gradient direction, as they have infinite variance. Thus, we obtain a real-space map of the lensing potential weighted for this specific wavevector.

(iv) Fourier transform the estimator and keep only the value of the transformed estimator at the target

wavenumber $\mathbf{L}$, i.e. $\hat{\phi}^{\mathbf{L}}(\mathbf{x}) \xrightarrow{FT} \hat{\phi}^{\mathbf{L}}(\mathbf{L}) \, (= \hat{\phi}^{\mathbf{L}}_{\mathbf{L}})$. The reason for this step is that the weighting we apply to the reconstructed map in the previous step is by construction optimal for the target wavevector, $\mathbf{L}$, we chose initially (Eq. 18). To obtain the reconstructed map, normalize by $1/\langle W^{\mathbf{L}}(\mathbf{x})\rangle$, where averages are over real space (i.e., $\mathbf{x}$). This step returns the normalized estimator derived in Eq. 28. For each target (L_x, L_y) wavevector pixel, we repeat this procedure to build up the entire reconstructed map in Fourier space. This requires us to perform a full spatial weighting and Fourier transform of the map for each pixel, which becomes more computationally expensive as the size of the maps is increased (though the algorithm could potentially be sped up by applying the same weighting to blocks of similar (L_x, L_y)).

(v) Compute the 2-dimensional power spectrum. The the cross power spectrum (between the reconstructed and the true fields) is given by $C^{\hat{\phi}^{\mathbf{L}}\phi}_{\mathbf{L}} = \frac{(\hat{\phi}^{\mathbf{L}}_{\mathbf{L}})^* \phi_{\mathbf{L}}}{\langle W^{\mathbf{L}}(\mathbf{x})\rangle_{\mathbf{x}}}$, while auto power spectrum (of the reconstructed field) can be obtained as $C^{\hat{\phi}\hat{\phi}}_{\mathbf{L}} = \frac{(\hat{\phi}^{\mathbf{L}}_{\mathbf{L}})^* \hat{\phi}^{\mathbf{L}}_{\mathbf{L}}}{\langle (W^{\mathbf{L}}(\mathbf{x}))^2\rangle_{\mathbf{x}}}$, where for the normalization we have used the standard expressions for calculating pseudo-C_ℓ power spectra given a spatially dependent slowly varying window function W [16].

(vi) Finally, bin the power spectrum with appropriately chosen maximal multipole $L_{\max}$ and number of bins n_{bins}. In the subsequent sections, we show the computed power spectra with $L_{\max} = 20000$ and $n_{\mathrm{bins}} = 10$ for an ultra-low noise experiment and $L_{\max} = 12000$ and $n_{\mathrm{bins}} = 12$ for a CMB-S4-like experiment (although much of the signal-to-noise will be at lower multipoles). The small-scale regime is where we expect our method to outperform the QE most significantly.

IV. RESULTS AND DISCUSSION

In order to test our method, we apply it to simulations mimicking data from CMB temperature measurements with different instrumental noise levels. We simulate the lensing deflection using the following algorithm, described in [17]. First, we generate a 2-dimensional Gaussian random field realization from a theoretical CMB unlensed temperature power spectrum with a pixel size of 0.05 arcmins on a 2048×2048 pixel map. The CMB maps are remapped (using fifth order spline interpolation to more accurately apply the deflection) to produce a lensed temperature map from a Gaussian realization of the convergence field as predicted from the theoretical power spectrum. The lensed map is then convolved with a beam of appropriate size and instrumental white noise is added. It is then downsampled to 1024×1024 pixel maps in Fourier space in order to avoid the need to use a pixel window function and to speed up the lensing reconstruction. We thus obtain lensed temperature map realizations with simulated instrumental noise of area $1.707 \times 1.707 \ \mathrm{deg}^2$ and a pixel size of 0.1 arcmins.

In this section, we consider three measures of instrumental noise for future experiments:

- **Ultra-low (UL) noise experiment:** Futuristic experiment with a noise factor of $\Delta_{\mathrm{T}} = 0.1 \ \mu$K-arcmin and a beam size of $\theta_{\mathrm{FWHM}} = 0.3$ arcmin.

- **CMB-S4-like experiment:** CMB-S4-like experiment with a noise factor of $\Delta_{\mathrm{T}} = 1\mu$K-arcmin and a beam size of $\theta_{\mathrm{FWHM}} = 1$arcmin.

- **SO-like experiment:** Simons Observatory (SO)-like experiment with a noise level of $\Delta_{\mathrm{T}} = 6\mu$K-arcmin and a beam size of $\theta_{\mathrm{FWHM}} = 1.4$arcmin.

The simulations are highly simplified and optimistic in that they contain no Sunyaev-Zel'dovich signal, Cosmic Infrared Background emission or any other extragalactic foreground; we revisit the topic of foreground contamination later. We note that these experiments are highly signal-dominated on large scales in temperature, which means that the local gradient can be measured with high accuracy.

Throughout this section, we will be comparing the GI method with the QE one to test whether one indeed observes an improvement on small scales, as we expect based on our discussion in Section III. The QE convergence field estimate is obtained from the standard TT estimator [18] calculated using the modes $\ell_{\min} = 1000$ to $\ell_{\max} = 20000$. However, we have checked that the QE reconstruction maps give nearly identical results if $\ell_{\min}$ is decreased below 1000 (even to 0). The gradient leg is then low-pass filtered to remove the $L > 2000$ modes. For the autospectrum calculation, we, furthermore, apply a realization-dependent N_0 subtraction to the QE results (which reduces off-diagonal bandpower correlations) [19–21]. Note that our method is not expected to work on scales with L below $L \sim 4000$, so we apply a high-pass filter to the lensed temperature map before performing a GI reconstruction on the simulations. We further note that the statistical power of the QE is reduced in the small-scale regime compared to naive forecasts. This effect is a consequence of the fact that the errors on the different modes are highly correlated [12] – a fact not accounted for in the theoretical model of its reconstruction noise.

A. Reconstructed Maps

In Fig. 1, we show a comparison between the true (input) and the GI-reconstructed lensing convergence fields. We denote the convergence field by $\kappa(\mathbf{x})$ and it is related to the lensing potential by $\kappa(\mathbf{x}) = -\frac{1}{2}\nabla^2 \phi(\mathbf{x})$. To aid

with visualization of the reconstruction, the input lensing convergence field has been modified with the addition of large, randomly scattered massive point sources on top of the CMB lensing convergence map. We expect that modes with variation nearly perpendicular to the gradient direction or patches where the magnitude of the gradient is small will be reconstructed with very large noise and will, thus, be downweighted by applying the inverse noise filtering function, $W^{\mathbf{L}}(\mathbf{x}) = [N^{\hat{\phi}\hat{\phi}}(\mathbf{L}, \mathbf{x})]^{-1}$, described in the previous section.

Indeed, in regions where the magnitude of the gradient is large, the values of the reconstructed maps appear to be larger as well. This is a consequence of the fact that the inverse noise filter applied on the reconstructed maps downweights patches with small gradients, and they end up with values close to zero in the reconstructed plots. In particular, it is easy to see that some of the artificial point sources have been recovered quite well by the GI method, while others are not at all discernible in the reconstruction map. By glancing at the magnitude of the gradient map, it becomes evident why that is the case – the point sources located in positions of the map where the gradient is large have been correctly identified by the GI estimator, while those in regions with small gradient are missing from the reconstruction. The same qualitative behavior can in fact be observed in the QE-reconstructed maps as well; however, the weighting of regions with different gradients is expected to be suboptimal for the QE (see end of Subsection III A for discussion).

B. Estimator Validation and Power Spectra

In Fig. 2, we show the auto- and cross-power spectra (blue solid and dashed-dotted lines, respectively) obtained through the GI method from 360 simulations of a small CMB temperature patch as measured by a CMB-S4-like experiment and compare them with the auto power spectrum of the true lensing signal (black dashed curve). The cross power spectrum of the GI-reconstructed lensing matches the true input lensing power to high accuracy, demonstrating that the GI reconstruction algorithm is unbiased and works well for $L > 4000$. No bias subtraction is required in this case, which makes this cross-correlation measurement a highly robust validation of our algorithm.

The auto-power spectrum of the raw reconstructed convergence field, which includes a reconstruction noise bias, is somewhat larger than the signal on large scales and on small scales $L > 15000$, i.e., it becomes noise dominated. Even so, a measurement of the power spectrum is possible because of the large number of modes. We note that the reconstruction noise bias present here could be characterized by simulation and subtracted, estimated and removed using data-derived methods (see discussion later), or reduced by cross-correlating different splits of the data. We defer a detailed discussion of robust autospectrum bias subtraction to future work.
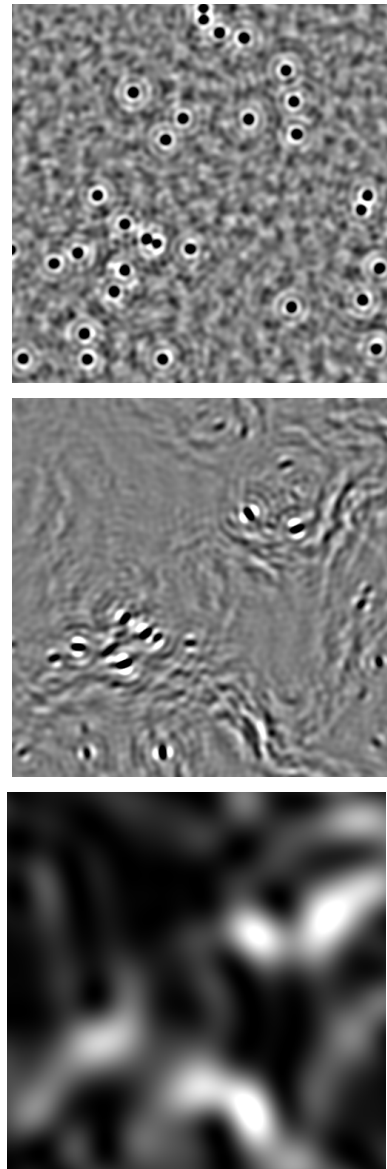

FIG. 1. Map of the input (*upper panel*) and the GI-reconstructed convergence field (*center panel*), as well as the magnitude of the temperature gradient (*lower panel*) for a simulated small patch (1.707×1.707 deg^2) of a CMB temperature ultra-low noise experiment with very bright galaxy clusters to better illustrate the reconstruction. A Wiener filter has been applied to the first two. A correlation between the convergence field maps is noticeable by eye as well corresponding to the regions where the magnitude of the gradient is largest.

C. Simulated Estimator Performance

In this subsection, our goal is to characterize the performance of the GI estimator and to compare it with that of the QE estimator, the current standard for reconstructing lensing from CMB temperature.

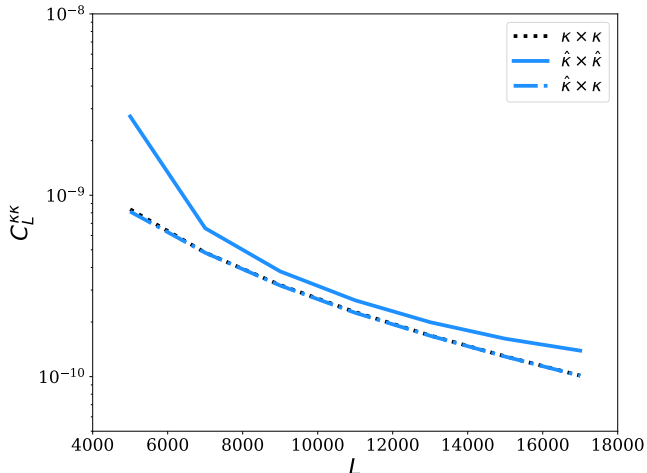For this purpose, we will use simulations to compute

FIG. 2. Auto and cross power spectra of the reconstructed convergence field for an experiment with ultra-low noise from 360 simulations. Here we do not show the $L < 4000$ modes, where we do not expect our method to work well (note our bin width is $\Delta L = 2000$ here). We see that the cross-correlation with the input is very accurate which gives credibility to our method. The auto power spectrum remains only slightly biased up to large wave vector values, $L \sim 15000$. The auto power spectrum shown above includes noise bias, which should be subtracted in a realistic analysis.

the errors on measurements of auto- and cross-power spectra with the reconstructed lensing field. Using the 360 simulations of lensing reconstruction described in the previous section, we measure and bin the lensing power spectrum and the lensing cross-correlation with the input convergence field; from these simulated measurements, we determine the covariance matrices and standard deviations for cross- and auto-spectra, for both GI and QE. For the autospectra, we assume that reconstruction noise biases will be characterized by monte-carlo simulations; since subtracting a mean bias does not affect fluctuations, we do not subtract the reconstruction noise bias when determining covariances. We show our results in Fig. 3.

As seen in the top panels of Fig. 3, the standard deviation (error) of the measured **auto**-power spectrum for the ultra-low noise experiment (top right) using the GI estimator is significantly smaller than the result from the QE simulations across all L-modes under consideration, reaching a full order of magnitude in difference around $L \sim 10000$. For an experiment with CMB-S4-like noise (top left), the error in the GI auto-correlation measurements is smaller by about a factor of 5 at $L \sim 6500$ compared with the QE, and seems to decrease the smaller the scale probed, leveling off and remaining non-negligible.

The bottom two panels of Fig. 3 shows the standard deviation of the measured **cross**-power spectrum. In both the case of an ultra-low noise experiment and a

CMB-S4-like experiment, the GI estimator appears to have smaller error bars than the QE on small scales ($L \gtrsim 4000$). In particular, for the UL-noise case (light-red and light-green curves), on scales around $L \gtrsim 10000$, our method outperforms the QE significantly, improving it by a factor of ~ 4. The CMB-S4 case (bottom left panel) also sees improvement when the GI method is applied, but it is somewhat more modest, at about a factor of ~ 2 over all scales shown.

We can also compare our simulated results with approximate theoretical predications. In Section III, we presented the theoretical framework for forcasting the noise of the GI approach. In particular, we showed that the noise on our estimator (in terms of the lensing potential field) approximately takes the form given in Eq. 30 (see a more detailed discussion of the GI noise in Appendix A). We assume Gaussian covariance [22] to convert this reconstruction noise power to errors $\sigma^2[C_L^{AB}]$:

$$\sigma^2[C_L^{AB}] = \frac{1}{\Delta L(2L+1)f_{\mathrm{sky}}}(C_L^{AA}C_L^{BB} + [C_L^{AB}]^2), \quad (31)$$

where the power spectra C_L^{AB} can include cross-spectra $C_L^{\hat{\kappa}\kappa}$ and autospectra $C_L^{\hat{\kappa}\hat{\kappa}}$, which include the reconstruction noise power. Following our results in Fig. 2, we assume $\langle C_L^{\hat{\kappa}\kappa} \rangle = C_L^{\kappa\kappa}$. We thus obtain theoretical predictions for the errors which are shown in Fig. 3 using thin lines.

We expect our theoretically derived expression for the GI noise to be accurate in the limit $L > 4000$ from Fig. 2, where the primary CMB is highly suppressed by diffusion damping. This is indeed what is observed in Fig. 3: for ($L \gtrsim 6500$), GI theory and simulation are in good agreement with each other for the case of CMB-S4-like experiment (dark-blue thin vs. thick curves). For the UL-noise experiment the agreement becomes more evident at $L \gtrsim 9000$. On larger scales, $L \lesssim 4000$, the primordial temperature fluctuations are not negligible, and many assumptions made in deriving the GI estimator on small scales break down. Therefore, it is perhaps not surprising that the errors from the simulations do not match the approximate theoretical calculation on these scales. Regarding the match of simulated results to simple theoretical predictions for the QE, as shown in Nguyen et al. [17] for bandpowers (and in Horowitz et al. [12] for clusters) , we confirm that the QE indeed has a much larger error on small scales than expected by the "naive" prediction which neglects mode coupling. This is especially true of the UL-noise experiment, where the difference is of about an order of magnitude on all scales. This suggests that if one wants to forecast the QE noise more correctly, a more careful treatment of its reconstruction noise has to be applied. On larger scales ($L < 1000$), we have verified that QE theory and simulation agree well, as expected. We emphasize that the large improvements of the GI estimator we have found are relative to the simulated QE, not the QE forecasts (which are overly optimistic).
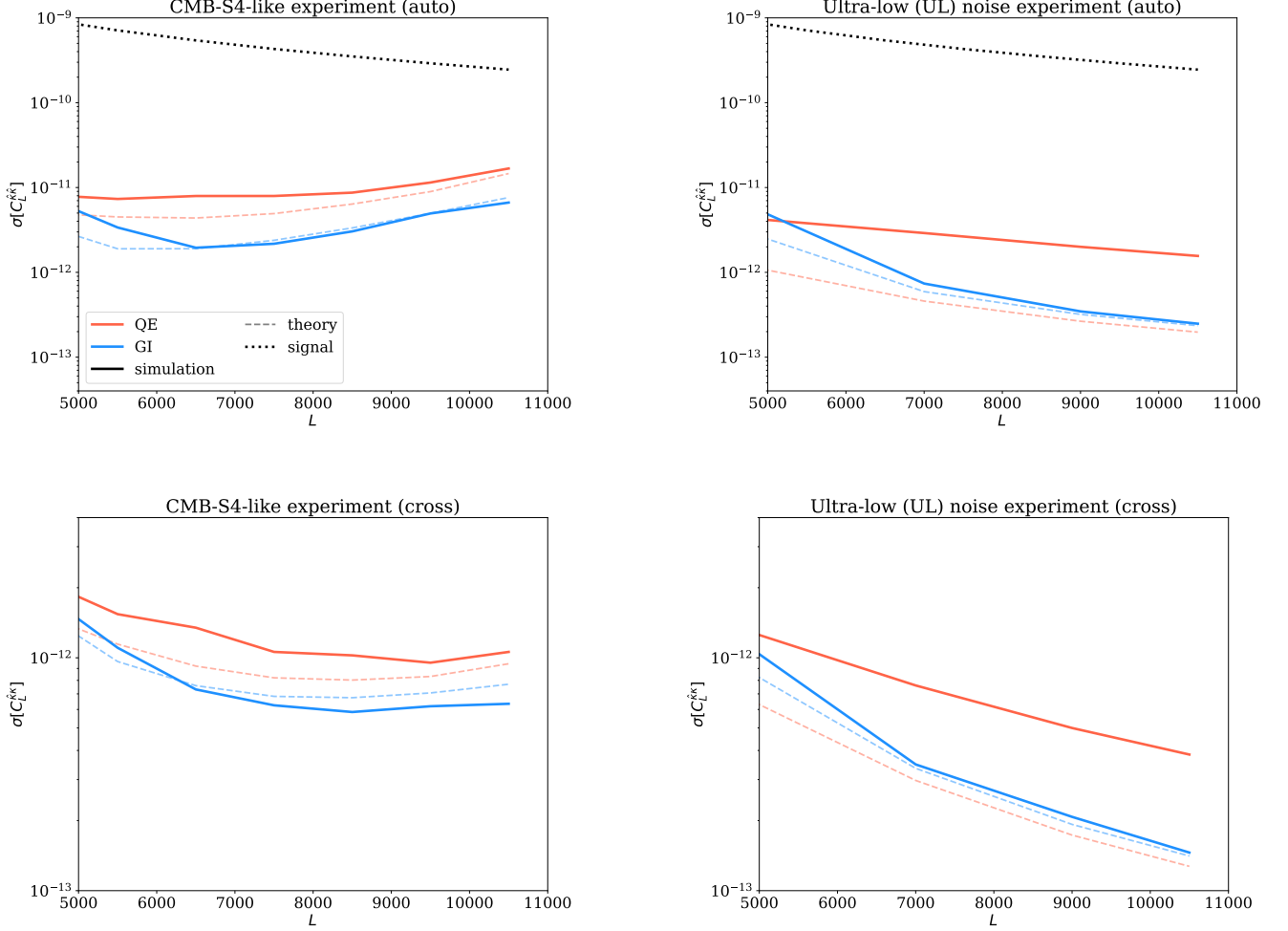
FIG. 3. Standard deviation of the auto (*upper panels*) and cross (*lower panels*) power spectra from multiple simulations for an experiment with CMB-S4-like (*left plots*) and ultra-low noise (*right plots*). QE results are in red, while GI results are in blue. The thick lines denote simulation, while the fainter dashed ones are derived from a theoretical prediction of the noise. Here we discard the $L < 5000$ modes, where our method is expected to be suboptimal. The error bars are scaled such that they simulate a sky coverage of 40%. The QE theory curves are "naive" predictions that ignore mode correlations, while the GI theory, which shows good agreement with the simulations results, curves take into account the effective area left unmasked by the inverse-noise filter. In (*dotted black*) we denote the power spectrum of the convergence field (signal). The bin width here is $\Delta L = 2000$.

To summarize the improvements expected from the GI estimator, we now calculate the cumulative signal-to-noise on lensing observables for different experiments. In Table I, we show the signal-to-noise ratio,

$$\text{SNR}^2 = \sum_{ij} C_L^{i\ \kappa\kappa} \widehat{\text{Cov}}_{ij}^{-1} C_L^{j\ \kappa\kappa}, \qquad (32)$$

computed over a range of L-modes for three noise levels, mimicking the fiducial ultra-low noise and Stage-IV-like experiments, with i and j indexing the bandpower bins and $\widehat{\text{Cov}}$ denoting the covariance of the measured power spectra, $C_L^{\hat{\kappa}\kappa}$ (cross) and $C_L^{\hat{\kappa}\hat{\kappa}}$ (auto). Notice that here we loosen the assumption of Gaussian covariance to obtain

more accurate estimates of the SNR, as indeed we observe non-negligible correlations between the small-scale modes of the QE-reconstructed power spectra. This is less true in the GI case, as expected. We have also added a forecast for an experiment similar to Simons Observatory (SO) , assuming instrumental noise with beam $\theta_{\text{FWHM}} = 1.4$ arcmin and noise factor $\Delta_{\text{T}} = 5\mu$K-arcmin, as it is the fastest approaching CMB observational project with potential to go to small enough scales in temperature. This calculation also includes the effect of off-diagonal bandpower covariances. We infer an improvement of the GI estimator over the QE on the order of $\sim 5-6$ times for the UL-noise case and about ~ 2.5 and ~ 3 for SO-like and CMB-S4, respectively, on small

SNR	UL	S4-like	SO-like
L-range	$4000 - 18000$	$5000 - 11000$	$5000 - 9000$
Cross QE	710	550	195
Cross GI	4100	1440	270
Auto QE	205	100	7
Auto GI	1515	360	30

TABLE I. Signal-to-noise ratio (SNR) computed from simulations ($\sim$300 for GI and 1000 for QE) of experiments with three different noise-levels: ultra-low noise (UL), CMB-Stage-IV-like (S4) and Simons-Observatory-like (SO).

scales $L = 5000 - 9000$.

We note that applying a more optimal window function, such as the full Wiener filter in Eq. 28, could, to some extent, further increase the precision of the GI lensing measurements on small-scales.

In the small-scale regime, we caution that there are significant contributions from foreground contaminants which still need to be properly accounted for (see Subsection IV E). Though there are several possibilities for mitigating these with multifrequency cleaning and estimator modification, we note that our forecasts are therefore likely optimistic. Nevertheless, especially at moderate L, the GI estimator appears promising for improving lensing signal-to-noise from future experiments beyond what was thought possible using the QE estimator.

D. Effective Reconstruction Noise

Another useful metric for comparing estimator performance is to calculate the reconstruction noise power spectrum, or, equivalently, the noise per lensing mode. However, due to inhomogeneous spatial weighting, the GI estimator can have a different effective area than the QE, so that a naive comparison is difficult. We proceed instead as follows. We define the effective reconstruction noise $N_{L,\mathrm{eff}}^{\kappa\kappa}$ as the level of noise that would give the same power spectrum error bars given a simple forecast with uniform weighting across the map. We can therefore calculate it by inverting the Gaussian covariance formula defined above (Eq. 31) for the lensing autospectrum to yield

$$N_{L,\mathrm{eff}}^{\kappa\kappa} \equiv \hat{\sigma}[C_L^{\hat{\kappa}\hat{\kappa}}]\sqrt{\Delta L (L+\tfrac{1}{2}) f_{\mathrm{sky}}} - C_L^{\kappa\kappa}, \qquad (33)$$

where $\hat{\sigma}[C_L^{\hat{\kappa}\hat{\kappa}}]$ is the standard deviation of the auto power spectra measured from multiple simulations and $C_L^{\kappa\kappa}$ is the theoretical power spectrum of the convergence field.

In Fig. 4, we show the QE and GI effective noise curves, where we use statistical results derived from simulations (360 GI reconstructions for UL noise; 300 GI reconstructions for SO-like and S4-like noise; 1000 QE reconstructions for each noise level). The graphs suggest that the GI noise of the estimator for an ultra-
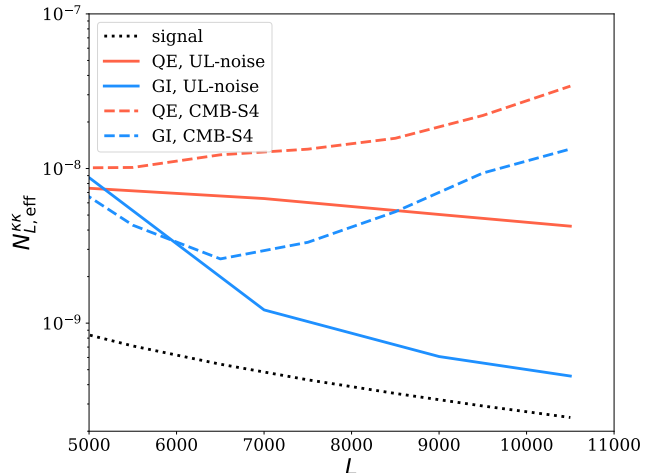

FIG. 4. Effective noise curves for an ultra-low noise and CMB-S4-like experiment from simulations. The largest improvement occurs for wave modes around $L \sim 10000$ of the order of ~ 10 for the UL-noise case and ~ 5 around $L \sim 6500$ for the CMB-S4 case. The effective noise is inferred from the scatter in the bandpowers of the auto-spectrum of each each estimator as in Eq. 33.

low noise experiment is on the order of 10 times smaller than that of the QE for $L \sim 10000$. For an experiment with CMB-S4-like instrumental noise, the improvement is more modest, peaking at $L \approx 6500$, where the ratio between the two curves is about 5. In both cases, the GI estimator results in a significantly more accurate measurement of the lensing field.

E. Systematic Errors from Foregrounds

Though the methods we have described for improved small-scale lensing reconstruction appear promising, small-scale extragalactic foregrounds will likely be a significant limiting factor for such analyses.

This is especially true for temperature reconstruction, which is the focus of this paper and is expected to have higher signal-to-noise. Polarization reconstruction should have minimal small-scale foregrounds, and we will briefly discuss the extension of our methods to polarization in Appendix B, but we will defer a detailed treatment to future work.

In temperature, as previously mentioned, several extragalactic foregrounds are of primary concern on small angular scales: Cosmic Infrared Background (CIB) emission and the thermal and kinematic Sunyaev-Zel'dovich effects (tSZ and kSZ). Other contaminants such as radio source emission are likely subdominant. CIB and tSZ foregrounds have a distinctive frequency dependence; this, in principle, allows them to be separated from the blackbody CMB and lensing signals using multi-

frequency measurements of the small-scale CMB from upcoming experiments.

The kSZ effect, on the other hand, cannot be separated using multifrequency data, as it has a blackbody frequency dependence. At first glance it therefore presents a more serious challenge, with the potential to significantly bias small-scale lensing measurements. However, the kSZ differs in an important way from lensing: the lensing signal is correlated with the background CMB gradient, whereas the kinetic SZ effect and other foregrounds are not. (This, of course, assumes that the gradient itself has a negligible kSZ contribution; we will briefly revisit this assumption below.) Building on this statistical difference, we can construct simple methods to avoid any kSZ bias to lensing measurements.

We will here briefly outline such a method to remove kSZ bias in autospectrum measurements using the GI lensing estimator. To calculate the bias from kSZ (in addition to that arising from instrument noise and other sources), we simply evaluate the GI estimator with the gradient coordinates offset by a small vector $\mathbf{r}$,

$$\hat{\phi}_{bias}^{\mathbf{L}}(\mathbf{x}) \approx \tag{34}$$
$$W^{\mathbf{L}}(\mathbf{x} + \mathbf{r}) \frac{\tilde{T}(\mathbf{x})_{\text{high-pass}}}{(i\mathbf{L} \cdot \nabla T(\mathbf{x} + \mathbf{r}))} \times \left[\langle W^{\mathbf{L}}(\mathbf{x} + \mathbf{r}) \rangle_{\mathbf{x}} \right]^{-1} .$$

Here $\mathbf{r}$ is chosen to be longer than the correlation lengths of the CMB fluctuations and the lensing, but small enough that the noise properties are unchanged. Since the kSZ, noise and other effects and foregrounds are not correlated with the large scale gradient [23–25], we can determine the non-lensing foreground and noise bias by simply calculating the power spectrum of $\hat{\phi}_{bias}^{\mathbf{L}}$. This can be subtracted off from the raw reconstruction power to calculate the lensing power spectrum, taking $(C_{\mathbf{L}}^{\hat{\phi}\hat{\phi}} - C_{\mathbf{L},bias}^{\hat{\phi}\hat{\phi}}) \times N_{\mathbf{L}}$. The normalization function $N_{\mathbf{L}}$ (which can be obtained via simulation or analytics) is required because $C_{\mathbf{L},bias}^{\hat{\phi}\hat{\phi}}$ also contains some lensing signal, though it is very sub-optimally weighted and thus significantly reduced in amplitude; re-normalization is needed to restore an unbiased measurement of the lensing power spectrum.

While this method should remove the leading source of kSZ bias, an additional, smaller non-Gaussian bias could arise because the measured CMB gradient also contains kSZ at a low level; this should be explored further in future work.

We also note that, even with reduced foreground biases, the kSZ contribution could still inflate errors and lower signal-to-noise in temperature, especially at very high L. Possible ways to reduce kSZ foreground contributions to errors (aside from simply relying on polarization-based reconstructions) might include subtracting a reconstructed kSZ map derived from galaxy surveys or quadratic kSZ reconstructions. We will defer a detailed exploration of foreground issues to future work.

V. CONCLUSIONS

The standard tool used for reconstruction of the gravitational lensing field is the quadratic estimator, but it has recently been shown to be suboptimal on small angular scales. The most rigorous way to overcome this limitation is to adopt an optimal approach (e.g., [11, 13, 26]), but this is not necessarily the most computationally efficient one, and in any case, algorithms that rapidly converge on small scales have not yet been presented. In this small-scale regime, a simple gradient-inversion estimator (GI) should approach the optimal solution.

In this paper, we have derived a more general GI algorithm, one that is now capable of reconstructing lensing over large fields. We have applied it to simulations of the lensed CMB temperature field to validate the method and characterize its performance. For idealized, foreground-free simulations, we have shown that the GI-reconstructed auto and cross correlation error bars are significantly smaller than those derived using the QE. In particular, we found that when using our method on small scales for the computation and analysis of the cross-correlations between these next-generation experiments, one obtains a much more precise result, with measurement errors which are approximately ~ 2.6 times smaller for CMB-S4 and ~ 1.4 for SO-like than the corresponding ones for QE at $L = 5000 - 9000$. The improvement of the auto-correlation power spectrum is even more substantial: a factor of ~ 4 for both CMB-S4 and SO-like in the highly optimistic scenario of removed foregrounds and measured small-scale anisotropies. We argue that the origin of these improvements lies in the fact that, unlike the GI, the QE is limited by gradient cosmic variance.

We expect that the small-scale lensing measurements could, to some extent, be even further improved by fully implementing the weighting derived in Eq.28 without any approximations.

Though we have presented simple ideas for mitigating or removing foreground biases to the estimator, we caution that the levels of extragalactic foregrounds are large, and that a full examination of the impact of foregrounds has been left to future work. However, as sketched out in Appendix B, an analogous version of the GI estimator can also easily be derived for the case of polarization; the polarization GI estimator has the advantage that it does not suffer from comparable foreground contamination.

Small-scale lensing reconstruction has many potential applications to cosmology such as constraining cosmological parameters, distinguishing between dark matter models (e.g. [17],[27]), validating galaxy weak lensing shear measurements, and constraining high-redshift astrophysics. If further work on polarization, foregrounds and systematics is successful, the improved small-scale lensing reconstruction algorithms presented here could allow new scientific possibilities in these areas.

ACKNOWLEDGMENTS

Appendix A: Noise Discussion

In this appendix, we discuss the theoretical prediction of the reconstruction noise in more detail. As discussed in Section III, the noise on our estimator for a local gradient is given by

$$N_{\mathbf{L}}^{\hat{\phi}\hat{\phi}}(\mathbf{x}) = \frac{N_L^{TT}}{(\mathbf{L} \cdot \nabla T(\mathbf{x}))^2}. \tag{A1}$$

We average over space and over different gradients to obtain the noise power spectrum. If the gradient modes are uncorrelated, then the minimum variance weights are given by

$$N_L^{\hat{\phi}\hat{\phi}} = \left[\int \frac{d^2\mathbf{x}}{A} \frac{1}{N_{\mathbf{L}}^{\hat{\phi}\hat{\phi}}(\mathbf{x})} \right]^{-1} = \frac{2N_L^{TT}}{L^2 |\nabla T|_{\rm rms}^2}, \tag{A2}$$

where A is the total area over which we are averaging.

In our initial approximations we have assumed that the noise on the CMB temperature on small scales is merely the instrumental noise; in reality, it can be more accurately modeled by "noise" power coming from other first-order contributions to the lensed temperature power spectrum (see Eq. 4.16 in Ref. [28])

$$N_L^{TT} = C_L^{\rm noise} + C_L^{\tilde{T}\tilde{T}} - \frac{1}{2}L^2 C_L^{\phi\phi}|\nabla T|_{\rm rms}^2. \tag{A3}$$

The second and third terms here are an attempt to isolate the contributions to the lensed power spectrum that do not arise from a particular lensing mode. The instrumental noise is estimated using the conventional parametrization [22] to be

$$C_L^{\rm noise} = \Delta_{\rm T}^2 \exp\left[\frac{L(L+1)\theta_{\rm FWHM}^2}{8\ln 2}\right], \tag{A4}$$

where the noise factor, $\Delta_{\rm T}$, is measured in μK-rad and the beam full-width half maximum (FWHM), $\theta_{\rm FWHM}$, in radians (given in arcmin in the paper for convenience).

Appendix B: Polarization Estimator

Similarly to the temperature estimators considered earlier, we can straightforwardly extend the GI method to polarization. In this appendix, we outline the approach one would take, assuming a constant gradient within the patch under consideration. This approximation is in fact not completely unreasonable in the limit of small-scale lensing we are interested in. A proper treatment of combining different patches with different gradient magnitudes and directions would follow in an analogous manner to what was developed in Section III.

As shown in [28], we can write the lensed E and B fields on small scales as:

$$\tilde{E}(\mathbf{L}) \approx n(\mathbf{L}) + E(\mathbf{L}) - \phi(\mathbf{L})\mathbf{L} \cdot (\nabla Q \cos(2\varphi_{\mathbf{L}}) + \nabla U \sin(2\varphi_{\mathbf{L}})) \tag{B1}$$

$$\tilde{B}(\mathbf{L}) \approx n(\mathbf{L}) + B(\mathbf{L}) - \phi(\mathbf{L})\mathbf{L} \cdot (-\nabla Q \sin(2\varphi_{\mathbf{L}}) + \nabla U \cos(2\varphi_{\mathbf{L}})). \tag{B2}$$

We then have two estimators for the deflection potential with two filter functions:

$$\hat{\phi}_1(\mathbf{L}) \equiv \tilde{E}(\mathbf{L})F_1(\mathbf{L}) \tag{B3}$$

$$\hat{\phi}_2(\mathbf{L}) \equiv \tilde{B}(\mathbf{L})F_2(\mathbf{L}). \tag{B4}$$

Multiplying Eq. B1 and Eq. B2 by $F_1(\mathbf{L})$ and $F_2(\mathbf{L})$, we average over CMB realizations of the polarization map, requiring that the estimators are unbiased, to obtain:

$$\phi(\mathbf{L}) = \langle\hat{\phi}_1(\mathbf{L})\rangle_{\rm CMB} = \langle\tilde{E}(\mathbf{L})F_1(\mathbf{L})\rangle_{\rm CMB} = $$
$$= \langle\phi(\mathbf{L})\mathbf{L} \cdot (\nabla Q \cos(2\varphi_{\mathbf{L}}) + \nabla U \sin(2\varphi_{\mathbf{L}}))F_1(\mathbf{L})\rangle_{\rm CMB}$$

$$\phi(\mathbf{L}) = \langle\hat{\phi}_2(\mathbf{L})\rangle_{\rm CMB} = \langle\tilde{B}(\mathbf{L})F_2(\mathbf{L})\rangle_{\rm CMB} = $$
$$= \langle\phi(\mathbf{L})\mathbf{L} \cdot (-\nabla Q \sin(2\varphi_{\mathbf{L}}) + \nabla U \cos(2\varphi_{\mathbf{L}}))F_2(\mathbf{L})\rangle_{\rm CMB}.$$

It follows that:

$$F_1(\mathbf{L}) = \frac{1}{\mathbf{L} \cdot (\nabla Q \cos(2\varphi_{\mathbf{L}}) + \nabla U \sin(2\varphi_{\mathbf{L}}))} \tag{B5}$$

$$F_2(\mathbf{L}) = \frac{1}{\mathbf{L} \cdot (-\nabla Q \sin(2\varphi_{\mathbf{L}}) + \nabla U \cos(2\varphi_{\mathbf{L}}))}. \tag{B6}$$

We, thus, end up with the two estimators:

$$\hat{\phi}_1(\mathbf{L}) = \tilde{E}(\mathbf{L})F_1(\mathbf{L})$$
$$= \frac{n(\mathbf{L}) + E(\mathbf{L}) + \mathbf{L} \cdot (\nabla Q \cos(2\varphi_{\mathbf{L}}) + \nabla U \sin(2\varphi_{\mathbf{L}}))}{\mathbf{L} \cdot (\nabla Q \cos(2\varphi_{\mathbf{L}}) + \nabla U \sin(2\varphi_{\mathbf{L}}))}$$

$$\hat{\phi}_2(\mathbf{L}) = \tilde{B}(\mathbf{L})F_2(\mathbf{L})$$
$$= \frac{n(\mathbf{L}) + \mathbf{L} \cdot (-\nabla Q \sin(2\varphi_{\mathbf{L}}) + \nabla U \cos(2\varphi_{\mathbf{L}}))}{\mathbf{L} \cdot (-\nabla Q \sin(2\varphi_{\mathbf{L}}) + \nabla U \cos(2\varphi_{\mathbf{L}}))},$$

with noise power spectra:

$$N_L^{1,\hat{\phi}\hat{\phi}} = \langle(\hat{\phi}_1-\phi)^2\rangle = \frac{N_L^{EE}}{[\mathbf{L}\cdot(\nabla Q\cos(2\varphi_{\mathbf{L}})+\nabla U\sin(2\varphi_{\mathbf{L}}))]^2} \tag{B7}$$

$$N_L^{2,\hat{\phi}\hat{\phi}} = \langle(\hat{\phi}_2-\phi)^2\rangle = \frac{N_L^{BB}}{[\mathbf{L}\cdot(-\nabla Q\sin(2\varphi_{\mathbf{L}})+\nabla U\cos(2\varphi_{\mathbf{L}}))]^2}. \tag{B8}$$

We thus effectively get two new estimators for the lensing potential which weighted appropriately, can provide a better estimator for the lensing potential. Polarization measurements, in addition, have significantly lower foreground contamination compared to temperature measurements. The generalization for many CMB patches of different gradients follows the approach outlined in Section III and will be addressed in a subsequent paper.

---

[1] K. M. Smith, W. Hu, and M. Kaplinghat, Phys. Rev. D **74**, 123002 (2006), arXiv:astro-ph/0607315 [astro-ph].

[2] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, L. P. L. Colombo, C. Combet, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, R. Fernandez-Cobos, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, A. Melchiorri, A. Mennella, M. Migliaccio, M. A. Miville-Deschênes, D. Molinari, A. Moneti, L. Montier, G. Morgante, A. Moss, P. Natoli, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, C. Sirignano, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, J. Valiviita, B. Van Tent, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca, arXiv e-prints , arXiv:1807.06210 (2018), arXiv:1807.06210 [astro-ph.CO].

[3] K. M. Smith, O. Zahn, and O. Dore, Phys. Rev. **D76**, 043510 (2007), arXiv:0705.3980 [astro-ph].

[4] S. Das, B. D. Sherwin, P. Aguirre, J. W. Appel, J. R. Bond, C. S. Carvalho, M. J. Devlin, J. Dunkley, R. Dünner, T. Essinger-Hileman, J. W. Fowler, A. Ha-

jian, M. Halpern, M. Hasselfield, A. D. Hincks, R. Hlozek, K. M. Huffenberger, J. P. Hughes, K. D. Irwin, J. Klein, A. Kosowsky, R. H. Lupton, T. A. Marriage, D. Marsden, F. Menanteau, K. Moodley, M. D. Niemack, M. R. Nolta, L. A. Page, L. Parker, E. D. Reese, B. L. Schmitt, N. Sehgal, J. Sievers, D. N. Spergel, S. T. Staggs, D. S. Swetz, E. R. Switzer, R. Thornton, K. Visnjic, and E. Wollack, Physical Review Letters **107**, 021301 (2011), arXiv:1103.2124.

[5] A. van Engelen, R. Keisler, O. Zahn, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, J. Dudley, E. M. George, N. W. Halverson, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, M. Joy, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, J. J. Mohr, T. E. Montroy, T. Natoli, S. Padin, T. Plagge, C. Pryke, C. L. Reichardt, J. E. Ruhl, J. T. Sayre, K. K. Schaffer, L. Shaw, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. Story, K. Vanderlinde, J. D. Vieira, and R. Williamson, Astrophys. J. **756**, 142 (2012), arXiv:1202.0546 [astro-ph.CO].

[6] Y. Omori, R. Chown, G. Simard, K. T. Story, K. Aylor, E. J. Baxter, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, W. B. Everett, E. M. George, N. W. Halverson, N. L. Harrington, G. P. Holder, Z. Hou, W. L. Holzapfel, J. D. Hrubes, L. Knox, A. T. Lee, E. M. Leitch, D. Luong-Van, A. Manzotti, D. P. Marrone, J. J. McMahon, S. S. Meyer, L. M. Mocanu, J. J. Mohr, T. Natoli, S. Padin, C. Pryke, C. L. Reichardt, J. E. Ruhl, J. T. Sayre, K. K. Schaffer, E. Shirokoff, Z. Staniszewski, A. A. Stark, K. Vanderlinde, J. D. Vieira, R. Williamson, and O. Zahn, Astrophys. J. **849**, 124 (2017), arXiv:1705.00743 [astro-ph.CO].

[7] P. A. R. Ade *et al.* (Planck), Astron. Astrophys. **594**, A15 (2016), arXiv:1502.01591 [astro-ph.CO].

[8] K. T. Story, D. Hanson, P. A. R. Ade, K. A. Aird, J. E. Austermann, J. A. Beall, A. N. Bender, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. C. Chiang, H. M. Cho, R. Citron, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, W. Everett, J. Gallicchio, J. Gao, E. M. George, A. Gilbert, N. W. Halverson, N. Harrington, J. W. Henning, G. C. Hilton, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, N. Huang, J. Hubmayr, K. D. Irwin, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, D. Li, C. Liang, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, L. Mocanu, T. E.

Montroy, T. Natoli, J. P. Nibarger, V. Novosad, S. Padin, C. Pryke, C. L. Reichardt, J. E. Ruhl, B. R. Saliwanchik, J. T. Sayre, K. K. Schaffer, G. Smecher, A. A. Stark, C. Tucker, K. Vand erlinde, J. D. Vieira, G. Wang, N. Whitehorn, V. Yefremenko, and O. Zahn, Astrophys. J. **810**, 50 (2015), arXiv:1412.4760 [astro-ph.CO].

[9] P. A. R. Ade *et al.* (POLARBEAR), Phys. Rev. Lett. **113**, 021301 (2014), arXiv:1312.6646 [astro-ph.CO].

[10] W. Hu and T. Okamoto, Astrophys. J. **574**, 566 (2002), arXiv:astro-ph/0111606 [astro-ph].

[11] C. M. Hirata and U. Seljak, Phys. Rev. **D67**, 043001 (2003), arXiv:astro-ph/0209489 [astro-ph].

[12] B. Horowitz, S. Ferraro, and B. D. Sherwin, ArXiv e-prints (2017), arXiv:1710.10236.

[13] J. Carron and A. Lewis, Phys. Rev. **D96**, 063510 (2017), arXiv:1704.08230 [astro-ph.CO].

[14] U. Seljak and M. Zaldarriaga, Astrophys. J. **538**, 57 (2000), arXiv:astro-ph/9907254 [astro-ph].

[15] W. Hu, S. DeDeo, and C. Vale, New J. Phys. **9**, 441 (2007), arXiv:astro-ph/0701276 [astro-ph].

[16] G. Efstathiou, Mon. Not. R. Astron. Soc **349**, 603 (2004), arXiv:astro-ph/0307515 [astro-ph].

[17] H. N. Nguyen, N. Sehgal, and M. Madhavacheril, Phys. Rev. **D99**, 023502 (2019), arXiv:1710.03747 [astro-ph.CO].

[18] W. Hu, S. DeDeo, and C. Vale, New Journal of Physics **9**, 441 (2007).

[19] T. Namikawa, D. Hanson, and R. Takahashi, Mon. Not. R. Astron. Soc **431**, 609 (2013), arXiv:1209.0091 [astro-ph.CO].

[20] B. D. Sherwin, A. van Engelen, N. Sehgal, M. Madhavacheril, G. E. Addison, S. Aiola, R. Allison, N. Battaglia, D. T. Becker, J. A. Beall, J. R. Bond, E. Calabrese, R. Datta, M. J. Devlin, R. Dünner, J. Dunkley, A. E. Fox, P. Gallardo, M. Halpern, M. Hasselfield, S. Henderson, J. C. Hill, G. C. Hilton, J. Hubmayr, J. P. Hughes, A. D. Hincks, R. Hlozek, K. M. Huffenberger, B. Koopman, A. Kosowsky, T. Louis, L. Maurin, J. McMahon, K. Moodley, S. Naess, F. Nati, L. New-

burgh, M. D. Niemack, L. A. Page, J. Sievers, D. N. Spergel, S. T. Staggs, R. J. Thornton, J. Van Lanen, E. Vavagiakis, and E. J. Wollack, Phys. Rev. D **95**, 123529 (2017).

[21] H. N. Nguyên, N. Sehgal, and M. S. Madhavacheril, Phys. Rev. D **99**, 023502 (2019), arXiv:1710.03747 [astro-ph.CO].

[22] L. Knox, Phys. Rev. **D52**, 4307 (1995), arXiv:astro-ph/9504054 [astro-ph].

[23] S. Ferraro and J. C. Hill, Phys. Rev. **D97**, 023512 (2018), arXiv:1705.06751 [astro-ph.CO].

[24] K. M. Smith and S. Ferraro, Phys. Rev. Lett. **119**, 021301 (2017), arXiv:1607.01769 [astro-ph.CO].

[25] A. Amblard, C. Vale, and M. J. White, New Astron. **9**, 687 (2004), arXiv:astro-ph/0403075 [astro-ph].

[26] M. Millea, E. Anderes, and B. D. Wandelt, ArXiv e-prints (2017), arXiv:1708.06753.

[27] N. Sehgal, H. N. Nguyen, J. Meyers, M. Munchmeyer, T. Mroczkowski, L. Di Mascolo, E. Baxter, F.-Y. Cyr-Racine, M. Madhavacheril, B. Beringue, G. Holder, D. Nagai, S. Dicker, C. Dvorkin, S. Ferraro, G. M. Fuller, V. Gluscevic, D. Han, B. Jain, B. Johnson, P. Klaassen, D. Meerburg, P. Motloch, D. N. Spergel, A. van Engelen, P. Adshead, R. Armstrong, C. Baccigalupi, D. Barron, K. Basu, B. Benson, F. Beutler, J. R. Bond, J. Borrill, E. Calabrese, O. Darwish, S. L. Denny, K. A. Douglass, T. Essinger-Hileman, S. Foreman, D. Frayer, M. Gerbino, S. G. A. Gontcho, E. B. Grohs, N. Gupta, J. C. Hill, C. M. Hirata, S. Hotinli, M. C. Johnson, M. Kamionkowski, E. D. Kovetz, E. T. Lau, M. Liguori, T. Namikawa, L. Newburgh, B. Partridge, F. Piacentni, B. Rose, G. Rossi, B. Saliwanchik, E. Schaan, H. Shan, S. Simon, A. Slosar, E. R. Switzer, H. Trac, W. Xu, M. Zaldarriaga, and M. Zemcov, arXiv e-prints , arXiv:1903.03263 (2019), arXiv:1903.03263 [astro-ph.CO].

[28] A. Lewis and A. Challinor, Phys. Rept. **429**, 1 (2006), arXiv:astro-ph/0601594 [astro-ph].