Impact of extragalactic foregrounds on internal delensing of the CMB *B*-mode polarization

Antón Baleato Lizancos[1,2,*] and Simone Ferraro[2,1,†]

[1]*Berkeley Center for Cosmological Physics, Department of Physics, University of California, Berkeley, California 94720, USA*
[2]*Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, California 94720, USA*

The search for primordial *B*-mode polarization of the cosmic microwave background (CMB) is limited by the sample variance of *B*-modes produced at later times by gravitational lensing. Constraints can be improved by "delensing": using some proxy of the matter distribution to partially remove the lensing-induced *B*-modes. Current and soon-upcoming experiments will infer a matter map—at least in part—from the temperature anisotropies of the CMB. These reconstructions are contaminated by extragalactic foregrounds: radio-emitting galaxies, the cosmic infrared background, or the Sunyaev-Zel'dovich effects. Using the Websky simulations, we show that the foregrounds add spurious power to the angular autospectrum of delensed *B*-modes via non-Gaussian higher-point functions, biasing constraints on the tensor-to-scalar ratio, *r*. We consider an idealized experiment similar to the Simons Observatory, with no Galactic or atmospheric foregrounds. After removing point sources detectable at 143 GHz and reconstructing lensing from CMB temperature modes $l < 3500$ using a Hu-Okamoto quadratic estimator (QE), we infer a value of *r* that is 1.5σ higher than the true $r = 0$. Reconstructing instead from a minimum-variance internal linear combination map only exacerbates the problem, bringing the bias above 3σ. When the *TT* estimator is co-added with other QEs or with external matter tracers, new couplings ensue which partially cancel the diluted bias from *TT*. We provide a simple and effective prescription to model these effects. In addition, we demonstrate that the point-source-hardened or shear-only QEs can not only mitigate the biases to acceptable levels but also lead to lower power than the Hu-Okamoto QE after delensing. Thus, temperature-based reconstructions remain powerful tools in the quest to measure *r*.

I. INTRODUCTION

The *B*-mode of polarization of the cosmic microwave background (CMB) is a powerful probe of the very early Universe. To leading order, primordial *B*-modes can only be produced by tensor fluctuations (gravitational waves) in the prerecombination plasma, so—unlike the temperature (*T*) or *E*-mode anisotropies—they are not obscured by the cosmic variance of scalar fluctuations [1,2]. *B*-modes are therefore a promising avenue to detect tensor modes, which are generically produced [3] in theories of cosmic inflation [4–7], but not in many of its alternatives [8]. A detection of primordial *B*-modes would reveal the amplitude of the power spectrum of tensor fluctuations, typically parametrized by the tensor-to-scalar ratio of primordial fluctuation power, currently constrained to be $r < 0.036$ [9] (95% confidence level, at a pivot scale of 0.05 Mpc^{-1}), and open the door to deeper investigations of inflation and fundamental physics (see, e.g., [10]).

Experimental constraints show the primordial *B*-mode signal is very small; the black, dotted curves in Fig. 1 are a theoretical calculation of its power spectrum for values of *r* compatible with observations. Making a detection therefore entails overcoming various formidable challenges—from the development of high-sensitivity detectors and readout systems, to the accurate characterization of foreground emission and the tight control of experimental systematics (see, e.g., [11]).

In addition to all these "local" challenges, there is a major one originating beyond our Galaxy. As they travel toward us from the surface of last scattering, CMB photons see their path deflected by the gravitational pull of the matter distribution of the Universe—they are gravitationally "lensed"; see [13] for a review. This converts part of the *E*-mode polarization generated by scalars into *B*-modes, an effect first detected by [14]. The angular power spectrum of these lensing *B*-modes, the black curve in Fig. 1, resembles that of white noise with $\Delta_{\rm P} \approx 5~\mu{\rm K}$ arcmin on the large angular scales where the primordial signal is expected to be strongest. Crucially, the variance associated with this lensing component
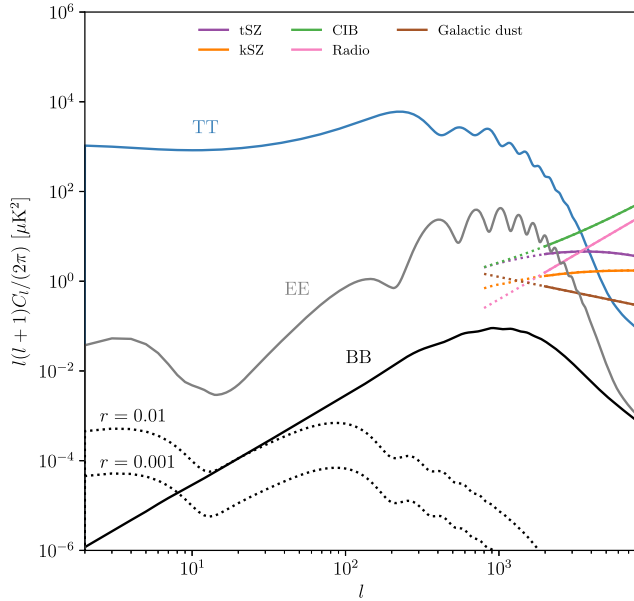
*a.baleatolizancos@berkeley.edu
†sferraro@lbl.gov

FIG. 1. Angular power spectra of the lensed CMB temperature (blue), E-mode (gray) and B-mode polarization (black) in the ΛCDM cosmological model of [12]. We emphasize two points: (i) the primordial B-modes (black, dotted) peak on degree scales, but they are obscured by the contribution from lensing (black, solid), and (ii) on small scales, extragalactic foregrounds make sizeable contributions to the temperature fluctuations (to avoid clutter, we plot them only at large l). Note that although the primordial B-mode signal is enhanced at $l < 10$—the so-called "reionization bump"—this is exceedingly difficult to measure, especially from the ground [10].

hinders searches for a primordial B-mode signal produced by tensor fluctuations [15].

The sample variance induced by lensing can be partially removed by estimating the specific realization of lensing B-modes present in the sky, and either removing them from map-level observations of the large-scale CMB polarization, or folding this information into a realization-dependent model around which to build a likelihood for r. This is known as "delensing" the B-modes [16–19]. B-mode delensing has already been demonstrated on data [20–25], and shown to improve constraints on r [25]. In fact, lensing is thought to be the single largest source of uncertainty in the most recent BICEP/Keck analysis [9]. Going forward, experiments seeking to better constrain r will rely on extensive delensing (e.g., [26–30]).

To estimate the realization of lensing B-modes in the sky, we need both high-resolution observations of the E-mode polarization and some proxy of the lensing potential [31,32]. The latter ingredient can be obtained internally, by reconstructing lensing maps from the CMB itself [19,33–35]; externally, using tracers of the large-scale structure (which correlate with CMB lensing) [31,36,37], or by combining both types of tracers (e.g., [26]). In this paper we focus on internal reconstructions, which are
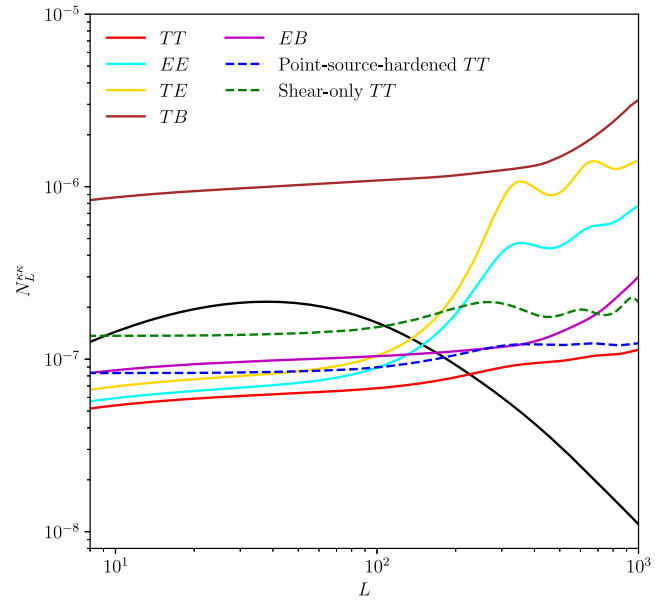

FIG. 2. Theoretical reconstruction noise of various quadratic estimators of lensing, for the experiment described in Sec. III A, similar to the 143 GHz channel of the large-aperture telescope of SO. These reconstructions use CMB anisotropies in the range $2 < l < 3000$, except for the EB estimator, for which $300 < l < 3000$, as required to avoid delensing bias [40–42]. The noise is to be compared to the signal $C_L^{\kappa\kappa}$, shown in black. For SO, the TT of Hu and Okamoto [33] has the highest signal-to-noise ratio of any quadratic estimator.

ultimately poised to return the highest-fidelity tracer of CMB lensing (see, e.g., [21]), and will be an indispensable element of upcoming delensing analyses. In particular, we restrict our investigation to quadratic estimator (QE) reconstructions obtained from the CMB temperature anisotropies, since these offer the highest signal-to-noise lensing estimates out of all possible QEs for an experiment such as the upcoming Simons Observatory (SO) [38] (see Fig. 2) and are likely to be used, albeit to a lesser extent, by other experiments such as the South Pole Observatory (see, e.g., [39]) or CMB-S4 [10]. Though QEs will ultimately be superseded by more sophisticated reconstruction algorithms which extract lensing information beyond leading order (e.g., [19,34,35]), they are near optimal for SO [38], and provide a transparent test bed for systematics that will be relevant to the more optimal—and complex—methods.

The motivating fact for this work is that, as shown in Fig. 1, a significant fraction of the temperature fluctuations in the small-scale microwave sky are produced by extragalactic "foreground" emission. These include the thermal Sunyaev-Zel'dovich effect [43] (tSZ), a distortion of the CMB frequency spectrum in the direction of massive clusters where CMB photons inverse-Compton-scatter off hot electrons; the kinetic SZ [44] (kSZ), a Doppler boost of those photons due to the bulk motion of electrons relative to the Hubble flow; the cosmic infrared background

(CIB; see, e.g., [45]), integrated emissions of interstellar dust in star-forming galaxies; and radio emission from active galactic nuclei (AGN), which appear as point sources in the CMB maps. All of these foregrounds are correlated with lensing (because they trace the matter distribution) and, importantly, obey non-Gaussian statistics. As a consequence, the auto and cross spectra of CMB lensing reconstructions can be biased at the level of several percent by the foregrounds' trispectra and their bispectra with CMB lensing [46–58]. Despite this being a well-known issue, the question of whether similar biases could affect B-mode delensing has remained unexplored to date.

In this work, we show that not accounting for the non-Gaussian statistics of the extragalactic foregrounds can severely bias the delensing pipeline of an experiment similar to the Simons Observatory if the standard TT quadratic estimator plays a part in the reconstructions. Interestingly, when the TT estimator is co-added with other quadratic estimators and/or tracers of the matter distribution, new couplings arise which partially cancel the diluted bias from TT, leading to a further reduction in the overall bias, beyond the naive expectation based on the TT weights alone. Fortunately, existing tools such as the point-source-hardened or shear-only quadratic estimators are relatively immune to foregrounds while retaining much of the delensing efficiency, and can be used to obtain overall-lower B-mode power than is possible with the standard Hu-Okamoto (HO) QE. Moreover, we show that whatever spurious power there is after delensing can be accurately modeled using a simple prescription.

This paper is structured as follows. In Sec. II, we lay out the theoretical framework behind B-mode delensing and CMB lensing reconstructions, and we use it to motivate our study of delensing biases from extragalactic foregrounds. In Sec. III, we introduce the simulations and explain in detail the delensing pipeline used to assess these biases. The results are presented in Sec. IV, including a discussion of mitigation strategies and an extension to the case of multitracer delensing. We conclude with a discussion of the implications of this work in Sec. V.

II. THEORETICAL BACKBONE

A. Lensing and delensing of B-modes

We are interested in measuring the linear polarization of the CMB, the only type produced at leading order. Though this can in principle be captured by the Q and U Stokes parameters, polarization is a spin-2 field on the sphere, so Q and U are coordinate dependent. It is therefore more convenient to work with the scalar E- and B-modes, which contain the same information and behave in unique ways under parity transformations: E-modes are even, while B-modes are odd. When the Stokes parameters

are defined on a global x-y basis, the two characterizations are related as[1]

$$(Q \pm iU)(\mathbf{x}) = -\int \frac{d^2\mathbf{l}}{(2\pi)^2} [E(\mathbf{l}) \pm iB(\mathbf{l})] e^{\pm 2i\psi_{\mathbf{l}}} e^{i\mathbf{l}\cdot\mathbf{x}}, \quad (1)$$

where $\psi_{\mathbf{l}}$ is the angle between $\mathbf{l}$ and the x direction.

Lensing causes a remapping of the primary fields by a total deflection angle, $\boldsymbol{\alpha}$, as

$$\tilde{T}(\boldsymbol{x}) = T(\boldsymbol{x} + \boldsymbol{\alpha}(\mathbf{x})),$$
$$\tilde{Q}(\boldsymbol{x}) = Q(\boldsymbol{x} + \boldsymbol{\alpha}(\mathbf{x})) \quad \text{and}$$
$$\tilde{U}(\boldsymbol{x}) = U(\boldsymbol{x} + \boldsymbol{\alpha}(\mathbf{x})), \quad (2)$$

where tildes denote lensed fields. Since the curl-like component of the deflection angle is orders of magnitude smaller than the gradient term [34,60], we may work simply with the lensing convergence, defined as $\kappa = -\frac{1}{2}\boldsymbol{\nabla} \cdot \boldsymbol{\alpha}$, and related to the lensing potential by $\kappa = -\nabla^2\phi/2$.

To leading order, the B-mode polarization induced by gravitational lensing of E-modes can be written as

$$\tilde{B}(\boldsymbol{l}) \approx \int \frac{d^2\boldsymbol{l}'}{(2\pi)^2} W(\boldsymbol{l},\boldsymbol{l}')E(\boldsymbol{l}')\kappa(\boldsymbol{l}-\boldsymbol{l}'), \quad (3)$$

where

$$W(\boldsymbol{l},\boldsymbol{l}') = \frac{2\boldsymbol{l}'\cdot(\boldsymbol{l}-\boldsymbol{l}')}{|\boldsymbol{l}-\boldsymbol{l}'|^2}\sin 2(\psi_l - \psi_{l'}). \quad (4)$$

On large angular scales, where the primordial B-mode power is expected to peak, the power spectrum of Eq. (3) is an excellent approximation to the true, nonperturbative result [59]. The reasons for this are subtle (see [32]) but they have a clear (and perhaps unintuitive) implication for delensing: a faithful lensing B-mode template can be constructed as

$$\hat{B}^{\text{lens}}(\boldsymbol{l}) = \int \frac{d^2\boldsymbol{l}'}{(2\pi)^2} W(\boldsymbol{l},\boldsymbol{l}')\mathcal{W}_{l'}^E \mathcal{W}_{|l-l'|}^\kappa \times E^{\text{obs}}(\boldsymbol{l}')\hat{\kappa}(\boldsymbol{l}-\boldsymbol{l}')$$
$$\equiv g_l[E^{\text{obs}}\hat{\kappa}], \quad (5)$$

as long as lensed—not delensed or unlensed—E-modes are used [32]. Here, E^{obs} are the observed E-modes, $\hat{\kappa}$ is an estimate of the convergence, and $\mathcal{W}^E$ and $\mathcal{W}^\kappa$ are Wiener

---

[1]There exist percent-level differences between results calculated using the flat- or spherical-sky mathematical formalisms, for example, in the power spectrum of lensing B-modes [59]. Given the uncertainties present in other parts of our analysis—particularly in the foreground models—we deem it sufficient to work with the simpler flat-sky formalism. The results we obtain should still be qualitatively correct.

filters for E^{obs} and $\hat{\kappa}$, respectively.[2] The expression above gives a lensing B-modes template that is effectively optimal until the fidelity of mass tracer and E-mode observations allow for delensing residuals of $O(1\%)$ of the original lensing B-mode power—that is, beyond the era of CMB-S4 [32].

Large-scale B-mode polarization can then be delensed by subtracting this template off from observations,[3]

$$\hat{B}^{\mathrm{del}}(\boldsymbol{l}) = \tilde{B}^{\mathrm{obs}}(\boldsymbol{l}) - \hat{B}^{\mathrm{lens}}(\boldsymbol{l}); \qquad (6)$$

the power spectrum of B-modes after delensing is then

$$C_l^{BB,\mathrm{del}} = C_l^{BB} + N_l^{BB} + C_l^{BB,\mathrm{res}} + C_l^{BB,\mathrm{fg}}, \qquad (7)$$

where C_l^{BB}, N_l^{BB}, $C_l^{BB,\mathrm{res}}$, and $C_l^{BB,\mathrm{fg}}$ are the angular power spectra of primordial/unlensed, experiment noise, residual lensing, and foreground B-modes, respectively.

The last of these components, the contribution from polarized foregrounds, is known to be a significant hurdle to searches for primordial B-modes and can come into play in a variety of ways. A well-known challenge lies in the fact that, due to synchrotron radiation and thermal dust emission, our Galaxy is the source of bright, polarized light at CMB frequencies. At their minimum frequency of 70–90 GHz and on the angular scales relevant to r-science, these foregrounds are at a level comparable to a primordial signal with $r = 0.01$–0.1, depending on sky region (see, e.g., [61] and references therein). In fact, current, high-sensitivity experiments already rely on foreground cleaning to achieve their target constraints (e.g., [62]). In parallel, polarized, extragalactic foregrounds also need to be taken into account, particularly when targeting a recombination signal with $r < 0.01$ and using small- or medium-aperture telescopes with relatively high confusion limits for point-source masking [63]. Finally, it is also possible for foreground residuals present in the large-scale B-mode maps to couple with residuals in the matter tracer and E-mode maps used to build the lensing B-mode template, biasing the delensing pipeline (e.g., [64,65]). All of these challenges are topics of active research, and we shall not consider them further in this paper. Instead, we will focus on how extragalactic foregrounds can bias lensing maps obtained from the CMB temperature anisotropies and propagate to biased inferences of r.

---

[2]As long as the fiducial spectra used in the Wiener filters are close to the truth, this deviation will only mildly reduce the delensing efficiency, for corrections to the power spectrum of delensed B-modes are second order in the error in the weights [36].

[3]It is common practice to delens by including the lensing template as an additional "channel" in a multifrequency, cross-spectral pipeline (e.g., [25]); this approach can be shown to be equivalent to a map-level subtraction.

Before we motivate our reasons for concern in Sec. II C, let us write out explicitly the terms that make up the angular power spectrum of residual lensing B-modes after delensing:

$$C_l^{BB,\mathrm{res}} = \tilde{C}_l^{BB} - 2C_l^{\tilde{B}\times\hat{B}^{\mathrm{lens}}} + C_l^{\hat{B}^{\mathrm{lens}}\times\hat{B}^{\mathrm{lens}}}, \qquad (8)$$

where $\tilde{C}^{BB}$ is the power spectrum of lensing B-modes, and we have defined the cross spectrum between the true lensing B-modes and the template,

$$
\begin{aligned}
C_l^{\tilde{B}\times\hat{B}^{\mathrm{lens}}} &= \langle \tilde{B}(\boldsymbol{l})\hat{B}^{\mathrm{lens}}(\boldsymbol{l}')\rangle' \\
&= g_l[\langle \tilde{B}E^{\mathrm{obs}}\hat{\kappa}\rangle], \qquad (9)
\end{aligned}
$$

with the prime following the angle bracket denoting that the delta function $\delta(\boldsymbol{l} - \boldsymbol{l}')$ has been removed. We have also written the auto spectrum of the template as

$$
\begin{aligned}
C_l^{\hat{B}^{\mathrm{lens}}\times\hat{B}^{\mathrm{lens}}} &= \langle \hat{B}^{\mathrm{lens}}(\boldsymbol{l})\hat{B}^{\mathrm{lens}}(\boldsymbol{l}')\rangle' \\
&= \int \frac{d^2\boldsymbol{l}_1}{(2\pi)^2} W(\boldsymbol{l},\boldsymbol{l}_1)\mathcal{W}_{l_1}^E \mathcal{W}_{|\boldsymbol{l}-\boldsymbol{l}_1|}^\kappa \\
&\quad \times \int \frac{d^2\boldsymbol{l}_2}{(2\pi)^2} W(\boldsymbol{l}',\boldsymbol{l}_2)\mathcal{W}_{l_2}^E \mathcal{W}_{|\boldsymbol{l}'-\boldsymbol{l}_2|}^\kappa \\
&\quad \times \langle E^{\mathrm{obs}}(\boldsymbol{l}_1)\hat{\kappa}(\boldsymbol{l}-\boldsymbol{l}_1)E^{\mathrm{obs}}(\boldsymbol{l}_2)\hat{\kappa}(\boldsymbol{l}-\boldsymbol{l}_2)\rangle \\
&\equiv h_l[\langle E^{\mathrm{obs}}\hat{\kappa}E^{\mathrm{obs}}\hat{\kappa}\rangle]. \qquad (10)
\end{aligned}
$$

In an ideal scenario, absent biases from foregrounds and otherwise, the residual lensing power is well modeled as [31,36,42]

$$C_l^{BB,\mathrm{res}} = \int \frac{d^2\boldsymbol{l}'}{(2\pi)^2} W^2(\boldsymbol{l},\boldsymbol{l}') C_{l'}^{EE} C_{|\boldsymbol{l}-\boldsymbol{l}'|}^{\kappa\kappa}[1 - \mathcal{W}_{l'}^E \rho_{|\boldsymbol{l}-\boldsymbol{l}'|}^2], \qquad (11)$$

where $\mathcal{W}_l^E$ is a Wiener-filter for E-modes (assumed here and throughout to be diagonal, for simplicity) and

$$\rho_L \equiv \frac{C_L^{\kappa\hat{\kappa}}}{\sqrt{C_L^{\kappa\kappa} C_L^{\hat{\kappa}\hat{\kappa}}}} \qquad (12)$$

is the correlation coefficient between the mass tracer and CMB lensing.

B. Quadratic estimators of lensing

In order to build the lensing B-mode template of Eq. (5), we need an estimate of the convergence map responsible for the deflections. In this paper, we focus on quadratic estimators derived from CMB temperature fields, as these will play an important role in current and upcoming experiments and are more prone to contamination from extragalactic foregrounds than estimators relying on polarization [66]. A general TT quadratic estimator of lensing can be written as

$$\hat{\kappa}(\boldsymbol{L}) = N^{\kappa\kappa}(\boldsymbol{L}) \int \frac{d^2\boldsymbol{l}}{(2\pi)^2} F^{\kappa}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}) T^{\text{obs}}(\boldsymbol{l}) T^{\text{obs}}(\boldsymbol{L} - \boldsymbol{l}).$$

(13)

If the normalization is chosen to be

$$[N^{\kappa\kappa}(\boldsymbol{L})]^{-1} = \int \frac{d^2\boldsymbol{l}}{(2\pi)^2} F^{\kappa}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}) f^{\kappa}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}), \quad (14)$$

then to leading order the estimator has unit response to lensing, and $N^{\kappa\kappa}$ gives the Gaussian component of the reconstruction noise[4]; this is the expression we used to calculate the various noise curves in Fig. 2. In the expression above,

$$f^{\kappa}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}) = \frac{2\boldsymbol{L}}{L^2} \cdot [\boldsymbol{l} \tilde{C}_l^{TT} + (\boldsymbol{L} - \boldsymbol{l}) \tilde{C}_{|L-l|}^{TT}], \quad (15)$$

where $\tilde{C}_l^{TT}$ is the lensed CMB temperature angular power spectrum.[5] On the other hand, one is free to choose the form of $F^{\kappa}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l})$ to satisfy some desired property. The standard HO quadratic estimator of Ref. [33] has

$$F^{\kappa,\text{HO}}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}) = \frac{f^{\kappa}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l})}{2C_l^{\text{tot}} C_{|L-l|}^{\text{tot}}}, \quad (16)$$

which results in the minimum variance estimator of lensing that is quadratic in the temperature fields.

We will be comparing this so-called "standard" QE to alternative quadratic estimators which are by design more robust to biases from extragalactic foregrounds. One of the alternative estimators we consider is the shear-only estimator of Ref. [55], defined by the weights

$$F^{\kappa,\text{shear}}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}) = \frac{\cos[2(\psi_{\boldsymbol{L}} - \psi_{\boldsymbol{l}})] \tilde{C}_l^{TT}}{2C_l^{\text{tot}} C_{|L-l|}^{\text{tot}}} \frac{d \ln \tilde{C}_l^{TT}}{d \ln l}. \quad (17)$$

The reason for this nomenclature is that in the limit where large-scale lenses are reconstructed from much smaller anisotropies the estimator can be shown to extract information only from the shear. This is a desirable property that bestows upon the estimator a high degree of immunity to extragalactic foregrounds, which are approximately azimuthally symmetric—their imprint on the CMB is therefore degenerate with the lensing magnification, but not the shear.

We also consider a bias-hardened estimator [51,52], $\hat{\kappa}_L^{\text{BH}}$, which by design has zero response to point sources at leading order; it is defined as

$$\begin{pmatrix} \hat{\kappa}_L^{\text{BH}} \\ \hat{s}_L^{\text{BH}} \end{pmatrix} = \begin{pmatrix} 1 & N_L^{\kappa} \mathcal{R}_L \\ N_L^{s} \mathcal{R}_L & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\kappa}_L \\ \hat{s}_L \end{pmatrix}, \quad (18)$$

where

$$\mathcal{R}_L = \int \frac{d^2\boldsymbol{l}}{(2\pi)^2} \frac{f^{\kappa}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}) f^{s}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l})}{2C_l^{\text{tot}} C_{|L-l|}^{\text{tot}}}, \quad (19)$$

and $\hat{s}_L$ and N_L^{s} are defined by analogy with $\hat{\kappa}_L$ and N_L^{κ}—they can be obtained from Eqs. (13), (14), and (16) by replacing κ with s in the superscripts. For point sources, $f^{s}(\boldsymbol{l}, \boldsymbol{L} - \boldsymbol{l}) = 1$.

C. Delensing bias from foreground non-Gaussianity

The foregrounds are non-Gaussian and correlated with the matter distribution that gravitationally lenses the CMB. This causes internal reconstructions of CMB lensing to be biased, in turn biasing the power spectrum of delensed B-modes. To isolate this effect, we define the bias as the difference in delensed-B-mode power between two pipelines which only differ by the statistical properties of the foregrounds that enter the QE. In this way, we can ensure that the leading-order reconstruction noise is the same in the two cases,[6] and thus that the delensing efficiency is only changed by the decorrelation induced by the foreground non-Gaussianity. Mathematically, this definition corresponds to

$$\Delta C_l^{BB,\text{res}} \equiv \langle |\tilde{B} - g_l[E^{\text{obs}} \hat{\kappa}[f_X + s_X^{\text{NG}}, f_Y + s_Y^{\text{NG}}]]|^2 \rangle$$
$$- \langle |\tilde{B} - g_l[E^{\text{obs}} \hat{\kappa}[f_X + s_X^{\text{G}}, f_Y + s_Y^{\text{G}}]]|^2 \rangle, \quad (20)$$

where s_X is the extragalactic foreground contribution to leg X of the XY QE, with the (NG) G superscript denoting that the foreground map in question is (non-)Gaussian. On the other hand, f_X refers to all the other nonforeground contributions to leg X: lensed CMB, experiment noise, etc.

In this work, we assume that extragalactic foregrounds are negligible in polarization, and set $s \equiv s_T$ while $s_E = s_B = 0$. This is a very good approximation given current sensitivity requirements because extragalactic foregrounds are expected to be polarized only at the percent level, if at all. In the future, it might be necessary to study the additional couplings that arise when the B- and E-modes in the expression above receive contributions from extragalactic foregrounds, and to investigate also the case of polarization-based reconstructions, which we do not consider here. However, the present treatment should be highly accurate for SO, particularly with the extension to

---

[4]In the idealized limit of homogeneous and isotropic noise.
[5]Reference [67] showed that $\tilde{C}_l^{TT}$ is a better approximation to the nonperturbative weights than C_l^{TT}.

[6]The Gaussian reconstruction noise is the same if the foregrounds have matching power spectra, but the non-Gaussian noise terms—$N^{(1)}$ and higher order—can in principle differ. We ignore this subtlety, since the Gaussian term dominates across the scales where the reconstruction is signal dominated (e.g., [68]).

the case of multitracer delensing that we provide in Sec. IV B.

The total bias described by Eq. (20) can be split into two parts:

$$\Delta C_l^{BB,\text{res}} = -2\Delta C_l^{\tilde{B}\times\hat{B}^{\text{lens}}} + \Delta C_l^{\hat{B}^{\text{lens}}\times\hat{B}^{\text{lens}}}, \qquad (21)$$

the bias to the cross spectrum of template and lensing B-modes, and the bias to the template auto spectrum. When delensing with a TT estimator, these can include contributions such as[7]

$$\Delta C_l^{\tilde{B}\times\hat{B}^{\text{lens}}} \supset g_l[\langle \tilde{B}\, \tilde{E}\, \hat{\kappa}^{TT}[s^{\text{NG}}, s^{\text{NG}}]\rangle_{\text{c}}], \qquad (22)$$

and

$$\Delta C_l^{\hat{B}^{\text{lens}}\times\hat{B}^{\text{lens}}} \supset$$

$$\supset 2\, h_l \left[\langle \overline{E^{\text{obs}}\hat{\kappa}^{TT}\left[\tilde{T}, \tilde{T}\right]}\, E^{\text{obs}}\hat{\kappa}^{TT}\left[s^{\text{NG}}, s^{\text{NG}}\right]\rangle_{\text{c}} \right]$$

$$+ 4\, h_l \left[\langle \overline{E^{\text{obs}}\hat{\kappa}^{TT}\left[\tilde{T}, s^{\text{NG}}\right]}\, E^{\text{obs}}\hat{\kappa}^{TT}\left[\tilde{T}, s^{\text{NG}}\right]\rangle_{\text{c}} \right]$$

$$+ h_l \left[\langle \overline{E^{\text{obs}}\hat{\kappa}^{TT}\left[s^{\text{NG}}, s^{\text{NG}}\right]}\, E^{\text{obs}}\hat{\kappa}^{TT}\left[s^{\text{NG}}, s^{\text{NG}}\right]\rangle_{\text{c}} \right]. \qquad (23)$$

In our notation, a Gaussian contraction of the fields connected by overbars is to be taken first, and this is then to be multiplied by the connected n-point function of the remaining fields inside the angle brackets. Note that all of these terms vanish if the foregrounds are Gaussian. We only show a subset of all possible contributions to $\Delta C_l^{\hat{B}^{\text{lens}}\times\hat{B}^{\text{lens}}}$, those that we believe should dominate based on the coupling structure of the weights inside the integrand.[8] It is couplings structurally identical to these we have retained (without the foregrounds) that dominate in the nominal calculation of the residual lensing B-mode spectrum and simplify to Eq. (11) [41,42]. A complete taxonomy of possible contributions is provided in Appendix A.

The couplings we have retained take a particularly simple analytic form. To leading order, they give

$$\Delta C_l^{\tilde{B}\times\hat{B}^{\text{lens}}} = \int \frac{d^2\boldsymbol{l}'}{(2\pi)^2} W^2(\boldsymbol{l},\boldsymbol{l}') \mathcal{W}_{l'}^E \mathcal{W}_{|\boldsymbol{l}-\boldsymbol{l}'|}^\kappa C_{l'}^{EE} \Delta C_{l-l'}^{\kappa\hat{\kappa}}, \quad (24)$$

and

$$\Delta C_l^{\hat{B}^{\text{lens}}\times\hat{B}^{\text{lens}}} = \int \frac{d^2\boldsymbol{l}'}{(2\pi)^2} W^2(\boldsymbol{l},\boldsymbol{l}') (\mathcal{W}_{l'}^E \mathcal{W}_{|\boldsymbol{l}-\boldsymbol{l}'|}^\kappa)^2 C_{l'}^{EE,\text{tot}} \Delta C_{l-l'}^{\hat{\kappa}\hat{\kappa}}, \qquad (25)$$

where $\Delta C^{\hat{\kappa}\hat{\kappa}}$ and $\Delta C^{\kappa\hat{\kappa}}$ are, respectively, the foreground-induced biases to the reconstruction's auto and cross spectra with the true CMB lensing convergence.

We are now in a position to make an explicit connection between the couplings in Eqs. (22) and (23) and the biases studied extensively in the context of the auto and cross spectra of CMB lensing reconstructions—e.g., [46,51,52,54–57,69]. To leading order in lensing, the trispectrum in Eq. (22) reduces to a product of the unlensed E-mode spectrum times

$$\langle \kappa\hat{\kappa}^{TT}[s^{\text{NG}}, s^{\text{NG}}]\rangle_{\text{c}}. \qquad (26)$$

This object, essentially a $\langle\kappa ss\rangle$ bispectrum, is known to bias cross correlations of CMB lensing reconstructions with any tracer of the large-scale structure. In other words, it sources $\Delta C^{\kappa\hat{\kappa}}$ in Eq. (24).

On the other hand, the terms in Eq. (23) feature a Gaussian contraction of the E-mode legs across templates, together with a connected four-point function of the fields involved in the lensing reconstruction. The latter produces a $\Delta C^{\hat{\kappa}\hat{\kappa}}$ in Eq. (25). Heuristically, the first line contains

$$\langle \hat{\kappa}^{TT}[\tilde{T}, \tilde{T}]\hat{\kappa}^{TT}[s^{\text{NG}}, s^{\text{NG}}]\rangle_{\text{c}}, \qquad (27)$$

so it is related to the "primary bispectrum bias"[9] discussed in the literature; the second line is a function of

$$\langle \hat{\kappa}^{TT}[\tilde{T}, s^{\text{NG}}]\hat{\kappa}^{TT}[\tilde{T}, s^{\text{NG}}]\rangle_{\text{c}}, \qquad (28)$$

so it can be associated with the "secondary bispectrum bias"; and the third line is sourced by

$$\langle \hat{\kappa}^{TT}[s^{\text{NG}}, s^{\text{NG}}]\hat{\kappa}^{TT}[s^{\text{NG}}, s^{\text{NG}}]\rangle_{\text{c}}, \qquad (29)$$

which is identical to the "trispectrum bias."

These biases to lensing reconstructions have been studied extensively before and are known to be at the level of several percent if unmitigated. It is therefore

---

[7]For notational economy, we drop from here on out the subscripts labeling the inputs to each QE leg. Though the new notation does not suggest this explicitly, the reader should remember that s is a placeholder for the combination of all extragalactic foregrounds—tSZ, kSZ, radio, or CIB—once the statistical properties (Gaussian or non-Gaussian) have been specified. This is an important point to encompass mixed biases sourced, for example, via the CIB-tSZ correlation.

[8]We follow the arguments in Appendix A of [42]. First, we rank terms according to their order in lensing; since the foregrounds are highly correlated with lensing, we take κ and s to be equivalent in this counting exercise. Then, we consider how different couplings affect the volume of multipole space over which the integrals are allowed to accumulate their signal. More tightly coupled integrands are likely to produce smaller results upon integration, so we rank them less highly in our priority list.

[9]To leading order in lensing, terms (27) and (28) both involve $\langle\kappa ss\rangle$ bispectra, hence the nomenclature.

pertinent to explore what impact they can have on the delensing procedure. In subsequent sections, we will use simulations to do precisely this. However, we can gain some preliminary intuition by investigating what angular-size lenses are most relevant for delensing, and seeing what the $\Delta C^{\hat{\kappa}\hat{\kappa}}$ and $\Delta C^{\kappa\hat{\kappa}}$ biases look like on those scales. Given that we are interested in B-modes with $l < 300$ and C_l^{EE} peaks near $l \sim 1500$, growing rapidly with l (see Fig. 1), the bulk of the integrals in (24) and (25) will come from the region in multipole space where $l \ll l'$. In this limit, the integrands simplify extensively, giving

$$\Delta C_l^{\tilde{B} \times \hat{B}^{\text{lens}}} \sim \frac{1}{4\pi} \int dl'\, l'\, \mathcal{W}_{l'}^E \mathcal{W}_{l'}^\kappa C_{l'}^{EE} \Delta C_{l'}^{\kappa\hat{\kappa}} \qquad (30)$$

and

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \sim \frac{1}{4\pi} \int dl'\, l'\, (\mathcal{W}_{l'}^E \mathcal{W}_{l'}^\kappa)^2 C_{l'}^{EE,\text{tot}} \Delta C_{l'}^{\hat{\kappa}\hat{\kappa}}, \qquad (31)$$

both of which are independent of l, as appropriate for the large-scale lensing B-modes, which resemble white noise. The integration kernels that accompany $C^{\kappa\hat{\kappa}}$ and $C^{\kappa\hat{\kappa}}$ in the integrands above are plotted in Fig. 3. Notice that modes of κ with $L \gtrsim 2000$ are irrelevant when delensing the large-scale B-modes; in fact, the majority of the information is coming from $L \lesssim 1000$.

This is an important insight. The effect of the bispectrum biases—Eqs. (26), (27), and (28)—is to suppress power on large scales[10] while adding it on small scales, with the transition between the two regimes happening somewhere in the multipole range $2000 \lesssim L \lesssim 2500$, the exact value being experiment dependent [69]. The bispectrum biases to lensing spectra are therefore negative across the scales where the integration kernels collect their signal. We thus expect, from Eq. (30), that $C_l^{\tilde{B} \times \hat{B}^{\text{lens}}}$ will be biased low. And once we take into account the factor of -2 preceding it in Eq. (21), we learn that it will contribute a positive bias to the power spectrum of delensed B-modes.

The situation is less clear cut for $\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}}$. Though the primary and secondary bispectrum biases are guaranteed to make a negative contribution, this will be offset to some extent by the trispectrum bias. Being a four-point function of the foreground amplitudes, the trispectrum bias is particularly sensitive to the extent to which individual sources can be removed [69]—it most certainly will

---

[10]Heuristically, this can be understood as follows. In regions of negative convergence, the CMB power spectrum is shifted to smaller angular scales [70]—i.e., to the right. From Fig. 1, it is clear than this results in more power at $l \sim 3000$ relative to the unlensed scenario. This is at the core of how a QE extracts the lensing signal: it interprets any excess of power at large l as evidence of there being below-average κ in the region. When the excess power is due to foregrounds, the estimator will return a value of $\hat{\kappa}$ that is biased low.
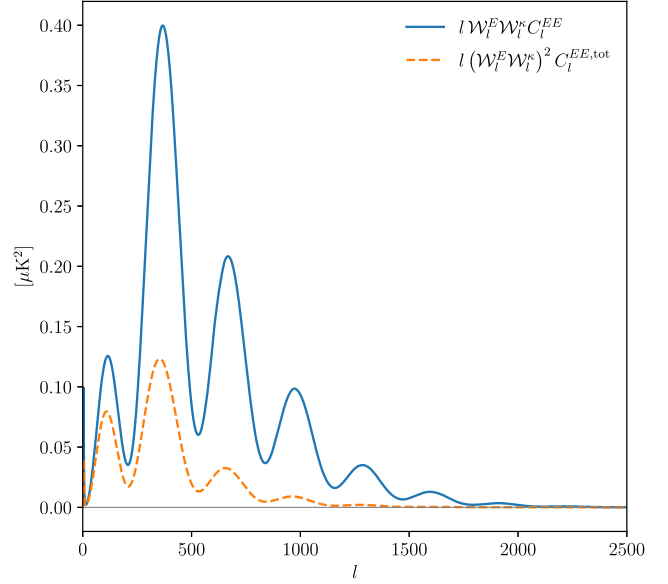

FIG. 3. Integration kernels in Eqs. (30) (blue) and (31) (orange), highlighting what scales of the lensing convergence are most important when delensing large-scale B-modes.

dominate unless the removal is extensive. This sensitivity to masking, compounded with uncertainties in foreground modeling, makes it difficult to predict the relative amplitude of the different contributions (though analytic work is ongoing [71]), and thus the sign of $\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}}$. Failing a categorical prediction, we at least expect to see evidence of cancellations between terms in the form of a somewhat reduced bias amplitude.

For completeness, we note that the TE reconstruction is in principle vulnerable to a bias analogous to the secondary bispectrum bias,

$$\Delta C_l^{BB,\text{res}} = h_l \left[\langle \overline{E^{\text{obs}} \hat{\kappa}^{\text{TE}} \left[s^{\text{NG}}, \tilde{E} \right]}\, E^{\text{obs}} \hat{\kappa}^{\text{TE}} \left[s^{\text{NG}}, \tilde{E} \right] \rangle_c \right]. \qquad (32)$$

There are also additional contributions from coupling arrangements that feature one or both of the template-E-mode legs in the trispectrum; though we omit them here, they are described in Appendix B. As we shall soon see, the bias terms associated with TE reconstructions are negligibly small.

The breakdown of bias terms in Eqs. (22), (23), and (32) is only provided for illustration purposes. In what follows, we will evaluate Eq. (20) directly using simulations, thus including all possible contributions.

III. COMPUTATIONAL METHODS

A. Simulations

We measure these effects using the Websky simulation [72]: a single, full-sky realization of the microwave sky. At

the lowest level, it relies on the efficient peak-patch algorithm [73] to construct halo catalogs with a mass resolution of 10^{12} $M_\odot$ at various redshifts, in the cosmology best fitting the Planck 2018 data [12]. Emission from various processes is then assigned to these halos based on prescriptions from astrophysical models, and this is then projected along the line of sight. The cornerstone of the tSZ model is the "Battaglia" pressure profile of Ref. [74], while for the CIB, pointlike galaxies are distributed in halos according to the CIB halo model of Ref. [75] with the best-fit parameters of Ref. [76]. Testament to the effectiveness of these models is that the tSZ-CIB correlation in the simulation is in good agreement with measurements by Planck [77]. In addition to the baseline Websky products, we also include the catalogs of radio sources developed by [78], which match the realization of the large-scale structure in Websky. Other astrophysical effects, such as CMB lensing or the kinetic SZ effect, are produced by fluctuations on scales so large that they are not bound in halos. In those cases, WebSky projects the emission from the "field"—calculated using Lagrangian perturbation theory—in addition to the contribution from halos. Note that we do not consider the integrated Sachs-Wolfe effect [79], since this only affects large angular scales of the temperature field which are unimportant for the sake of lensing reconstruction. On small scales, the Rees-Sciama effect [80] has been shown to be subdominant to the other extragalactic foregrounds [81], so we ignore it as well.

In later sections, we will often refer to the "unmitigated foreground intensity." This should be taken to mean the combination of all foreground intensity maps at 143 GHz—one of the main "science channels" of SO, where the CMB is brightest relative to the foregrounds. A less noisy lensing reconstruction can be obtained by applying the estimators to an internal linear combination (ILC; see, e.g., [82]) of observations at different frequencies, instead of just the 143 GHz map. To approximate this ILC procedure, we combine mock maps of the extragalactic foregrounds and beam-deconvolved white noise at $\{27.3, 41.7, 93.0, 143.0, 225.0, 278.0\}$ GHz, at or near the nominal central frequencies of the observation channels of the large-aperture telescope of SO,[11] for which we assume the "goal" noise and beam properties described in [38]. We use the publicly available code BasicILC[12] to calculate the harmonic-space ILC weights that minimize the variance of the co-added map. When doing so, we take into account not just CMB, extragalactic foregrounds (before point-source removal) and white noise, but also atmospheric noise and Galactic dust emission, as modeled in [83].

We remove point sources from the simulations to the extent that the SO LAT (Large-Aperture Telescope) is expected to be able to identify them at 143 GHz. In the case of radio sources, we simply avoid including those with flux above 5 mJy when building maps from the catalogs of [78]. An advantage of doing this at the catalog level is that the number of holes that we need to "drill" in the mask is minimized, simplifying later analyses. For the other foregrounds, for which we have access to map-level simulations rather than the catalogs, we identify point sources detected at 5σ confidence after applying a matched filter to the maps (see, e.g., [84]). We then set to zero all pixels located within a circle of radius $3'$ around the point source; this entails masking a mere 0.13% of pixels due to tSZ clusters, and 0.33% because of CIB point sources, which suggests that the bias due to lensing-mask correlations is likely to be negligible [85]. Doing this to the foreground maps, as opposed to the combination of all components, has the benefit that the components not being masked "inpaint" the hole once all maps are combined, thus avoiding spurious artifacts in the lensing reconstruction [86]. It is worth noting that it might be possible to achieve more extensive removal of clusters—and thus tSZ emission—by searching for sources at the level of the Compton-y maps instead of the individual frequency channels. Similarly, radio sources are more easily detected at frequencies lower than 143 GHz, and dusty star-forming galaxies at frequencies higher than that. Since the lensing biases are known to be highly sensitive to the extent of point source removal[13] [50], upcoming experiments aiming to mitigate delensing biases should explore the impact of various masking schemes on the biases presented in this work.

At this point, we are ready to project the various full-sky emission components onto smaller, flat patches. We do this because publicly accessible implementations of some of the quadratic estimators we will be testing are only available in the flat-sky limit. From a single full-sky simulation of intensity and polarization with NSIDE = 2048 in the HEALPIX pixelization,[14] we extract 48 flat, square, non-overlapping patches measuring 23.7 deg on a side. These patches are distributed in a homogeneous way across the celestial sphere; in order to avoid excessive distortions when projecting from spherical to planar geometries, we rotate the full-sky maps so that the center of each patch aligns with the origin of equatorial coordinates before projecting it. We pixelize the projected patch into a grid of square pixels, 1 arcmin on a side, such that there are 1424 of them along each of the two dimensions. A sample

---

[11]Whenever the SO central frequencies do not match the frequency of the simulated maps, we scale the latter on a pixel-by-pixel basis using the model frequency dependence in [83].

[12]See https://github.com/EmmanuelSchaan/BasicILC.

[13]Note however that, as pointed out in [56], point-source masking reduces the trispectrum bias of lensing by a much larger fraction than the bispectrum biases. This can have the unfortunate consequence of spoiling the low-L cancellation between these terms, leading to an overall increase in bias if masking is too aggressive.

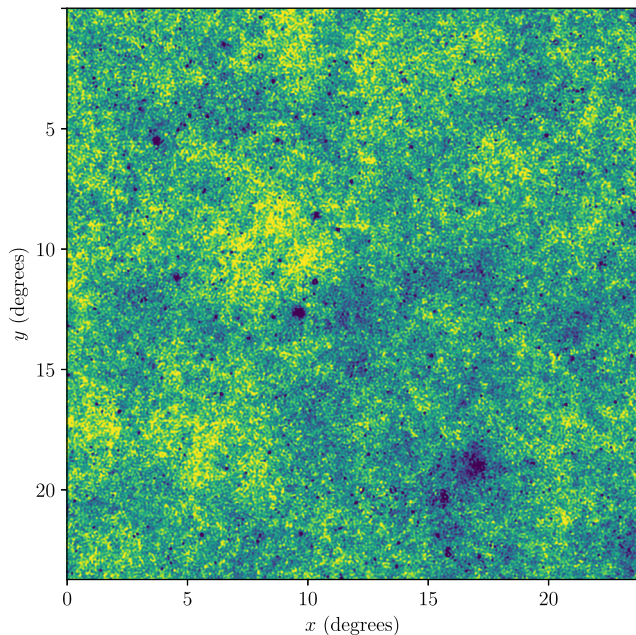[14]See http://healpix.sourceforge.net.

FIG. 4.   Sample realization of the temperature anisotropies at 143 GHz produced by the CIB, tSZ, and kSZ effects, and radio galaxies. Point sources detectable with 5σ confidence in the 143 GHz channel of the SO LAT have been removed from the individual foreground maps before co-adding. The intensity scale ranges from $-25\ \mu$K (dark blue) to 25 μK (yellow).

realization of the unmitigated foreground component is shown in Fig. 4.

The projection is applied to both temperature and polarization maps, with the latter requiring a little extra care. In order to avoid E-to-B leakage due to finite sky coverage [87], we separately project two sets of CMB-only polarization maps: a QU pair which contains only lensed E-modes (those from which we will build the lensing template), and another such pair containing only lensed B-modes (those that we would like to delens). The resulting, projected fields correctly reproduce the fiducial CMB spectra upon averaging over the 48 patches, up to sample variance.

In parallel, we produce an equal number of Gaussian foreground temperature maps. When simulating the unmitigated foreground component, we do so directly on flat, periodic patches with the same footprint and pixelization properties as those described above. When considering the ILC-cleaned scenario, on the other hand, we produce full-sky realizations at all the relevant frequencies and form the ILC prior to projection, thus reducing the number of expensive projection operations that are required. In either case, we draw harmonic coefficients from a Gaussian distribution with fiducial power spectrum drawn from a smooth fit to the power spectrum measured from the combination of all foregrounds in the full-sky Websky simulation.

Finally, we combine the CMB-only temperature map with the other components. We take two slightly different routes depending on whether we are dealing with the case of unmitigated or ILC-cleaned foregrounds. In the former case, we combine the projected CMB-only and foreground-only temperature maps, convolve them with a Gaussian beam with $\theta_{\mathrm{FWHM}} = 1.5$ arcmin and add white noise to the pixels with a standard deviation of $\Delta_{\mathrm{T}} = 6\ \mu$K arcmin. In the foreground-cleaned scenario, on the other hand, the linear combination of all foreground-plus-experiment-noise maps at different frequencies already took place prior to projection, so there is no need to convolve with the beam or add noise again. These steps are the same whether the foregrounds are Gaussian or non-Gaussian; in fact, in order to cancel sample variance, we add the same realization of the noise in both cases.

The reader might have noticed that we do not include atmospheric noise in our simulations. The atmosphere is the dominant source of noise when observing large-scale $l < 1000$ temperature anisotropies, so it can increase reconstruction noise—the reconstructions in [26], for example, are obtained after discarding T modes below $l \leq 500$—and degrade delensing efficiency. However, the only way that atmospheric noise can affect the amplitude of the biases we are concerned with in this work is via a relatively small change to the Wiener filter that is applied to the lensing reconstructions (see Sec. III C), as this modulates the amplitude of the foreground modes. Since characterizations of the atmospheric noise are experiment dependent and still uncertain, we choose not to factor this effect into our analysis except when calculating the ILC weights (where atmospheric noise does play a leading role at low and intermediate l). However, we note that our neglect of this contribution will result in slightly more extensive delensing than was forecasted in [26].

We apply the same beam convolution process to the Q and U maps used to generate the CMB E-modes, but add white noise with $\Delta_{\mathrm{P}} = \sqrt{2}\Delta_{\mathrm{T}}$. These values are intended to approximate the characteristics of the SO LAT at 145 GHz (in the "goal" scenario). The B-modes are left free of noise to reduce their variance and better isolate the biases of interest.

B. Lensing reconstructions

We now explain how our internal reconstructions of CMB lensing are performed. We use the three quadratic estimators presented in Sec. II B—the standard HO QE, the point-source-hardened (PSH) QE, and the shear-only QE—as implemented in the publicly available code SYMLENS.[15]

These estimators take as input the projected temperature maps described in the previous section, which contain lensed CMB, foregrounds and experiment noise, possibly after ILC cleaning. The input fields are inverse-variance-filtered (under the assumption of diagonal covariance)

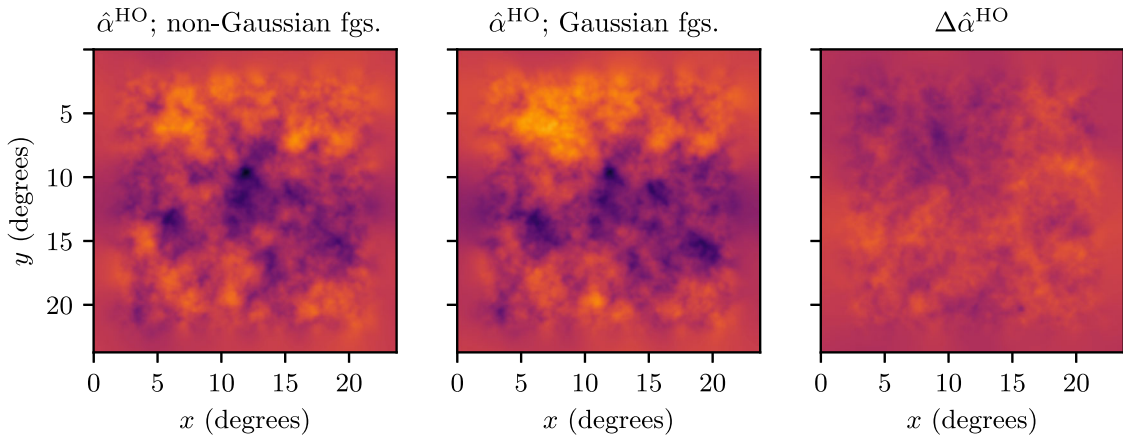[15]See https://github.com/simonsobs/symlens.

FIG. 5. Wiener-filtered magnitude of the deflection angle ($\hat{\alpha} = 2\hat{\kappa}/l$) reconstructed using a Hu-Okamoto TT QE applied to mock SO LAT observations at 143 GHz up to $l_{\max} = 3500$ and featuring unmitigated, non-Gaussian (left: same realization as in Fig. 4) or Gaussian (center) extragalactic foregrounds; also shown is the difference between the two (right). The realization of lensed CMB and noise is the same in both pipelines. The intensity scale ranges from -0.0025 (dark blue) to 0.0025 (yellow).

using a smooth fit to the total power in the full-sky simulation, be it the raw map or after foreground cleaning.

Since we are also interested in measuring the bias to TE reconstructions—e.g., Eq. (32)—we implement the Hu-Okamoto TE estimator [33]. In this case, the E-modes contain CMB and noise, but no foregrounds. We follow [22] and inverse-variance-filter T and E independently, ignoring their correlation.

The input fields are beam deconvolved, restricted to a bandpass of $2 < l < l_{\max}$ (we consider $l_{\max} = 3000, 3500$), and masked with a cosine apodization window with a width of 200 pixels (3.33 deg). The finite sky coverage gives rise to a mean-field contribution to the reconstructed CMB lensing maps. We estimate it as the mean of the map-level lensing reconstructions, averaged over the entire simulation set; the mean-field subtraction only increases the Gaussian reconstruction noise by $(100/N_{\text{sim}})\%$ [86], where N_{sim} is the number of realizations used in the computation, in this case 48.

In Fig. 5, we compare Wiener-filtered reconstructions of the magnitude of the lensing deflection angle ($\hat{\alpha} = 2\hat{\kappa}/l$) obtained by applying the standard Hu-Okamoto TT estimator to a simulated patch at 143 GHz, featuring unmitigated, non-Gaussian (left panel) or Gaussian (center panel) extragalactic foregrounds. There is a significant difference between the two, as shown in the right-most panel.

C. Lensing B-mode templates

The next step is to construct a lensing B-mode template in the style of Eq. (5). This takes as input the internal CMB lensing reconstructions obtained in the previous section, and a mock observation of the E-modes; the latter comes from projecting the full-sky, lensed-CMB-only Websky E-modes onto tiles, and adding experiment noise as explained in Sec. III A.

These input fields must be Wiener filtered prior to building the template. The E-mode filter,

$$\mathcal{W}_l^E \equiv \frac{\tilde{C}_l^{EE}}{\tilde{C}_l^{EE} + N_l^{EE}}, \tag{33}$$

is calculated from a fiducial noise model that assumes a Gaussian beam and white noise levels appropriate for the SO LAT at 145 GHz (see Sec. III A and [38]).

On the other hand, the Wiener filter for the convergence is

$$\mathcal{W}_L^\kappa \equiv \frac{C_L^{\kappa\kappa}}{C_l^{\hat{\kappa}\hat{\kappa}}}. \tag{34}$$

We calculate the numerator from the same fiducial model used to generate the Websky simulations. For the denominator, $C_l^{\hat{\kappa}\hat{\kappa}}$, we simply add to $C_L^{\kappa\kappa}$ an analytic estimate of the Gaussian reconstruction noise, computed from Eq. (14).

We verify that calculating $\mathcal{W}^\kappa$ analytically is only marginally less optimal—in the sense of how tight the error bars on our bias estimates eventually are—than doing so in a realization-dependent manner, in which $C^{\hat{\kappa}\hat{\kappa}}$ is measured from the simulations. The advantage of the analytic route, however, is that when the filter is the same for both pipelines—with Gaussian or non-Gaussian foregrounds—we ensure that, when differencing the delensed B-mode power obtained from each, we are not simply picking up effects coming from Gaussian terms calculated with different filtering functions.

Finally, we restrict both the E-mode and convergence fields to angular scales $2 < l < 3000$, since by the upper end of this range the signal-to-noise is already saturated.
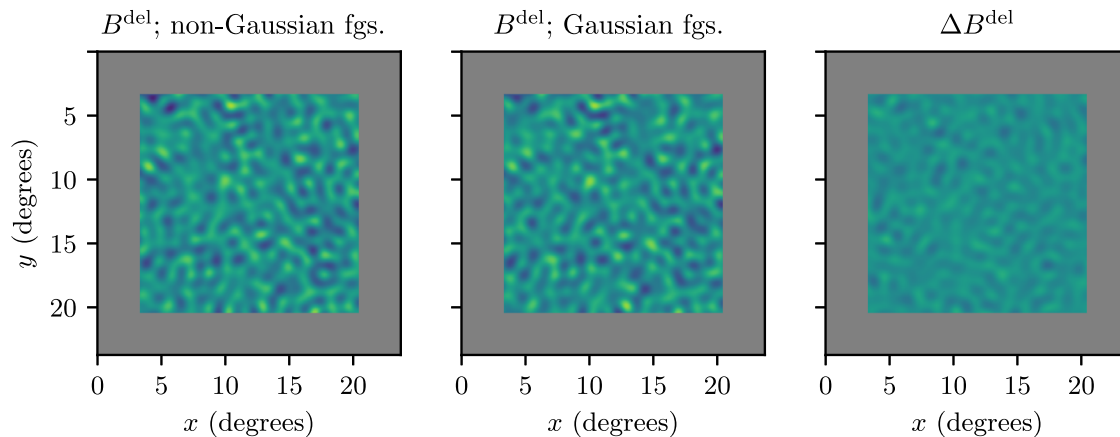
FIG. 6.   Real-space scalar fields associated with B-modes after delensing noiseless, foreground-free "observations" using the reconstructions in Fig. 5—that is, reconstructions obtained using a Hu-Okamoto QE applied to temperature fields up to $l_{\mathrm{max}} = 3500$ featuring unmitigated, non-Gaussian (left) or Gaussian (center) extragalactic foregrounds. The difference between the two (right) has noticeable structure on degree scales—modes above $l > 300$ have been removed to highlight the degree-scale pattern of interest to primordial B-mode searches. Recall that the realization of lensed CMB and noise is the same in both pipelines. Note also that the shaded region corresponds to pixels that are set to zero when measuring spectra. The intensity scale is the same across panels, ranging from $-0.5\ \mu\mathrm{K}$ (dark blue) to $0.5\ \mu\mathrm{K}$ (yellow).

D. Measuring spectra of delensed B-modes

Once the lensing B-mode template has been built, we can finally delens. We do this by subtracting the template from noiseless, foreground-free B-mode maps, as in Eq. (6).

We apply to the delensed fields a mask that is zero up to 200 pixels from the edges—avoiding in this way any overlap with the apodized regions of the temperature fields going into the lensing estimator—and transitions to one over the next 200 pixels in a smooth way, following a cosine curve.

A sample realization of delensed B-mode maps is shown in Fig. 6. The panel on the left is obtained using a HO TT estimator applied to 143 GHz temperature maps featuring non-Gaussian foregrounds, while the foregrounds are Gaussian for the panel in the center. The differences between the two are significant and become stark in the right-most panel, where we restrict the comparison to $l < 300$, the scales of interest to primordial B-mode searches.

Since our observations cover only a fraction of the sky, they can suffer B-to-E leakage in polarization.[16] To circumvent this challenge, we extract pure B-modes using NAMASTER [88]. We also use this code to calculate the pseudo-C_l's of the delensed field, and deconvolve the mode coupling induced by the mask. We verify that when the foregrounds are Gaussian the power spectrum of delensed

B-modes is well modeled by Eq. (11), up to sample variance. We also ensure that the measured band powers scatter consistently with the cosmic variance expected of Gaussian fields with their same power spectrum, sky footprint and binning scheme. Note, however, that in the next section we will be able to beat much of this cosmic variance—and better infer the delensing bias—by harnessing the fact that the CMB and experiment noise realizations are the same for a given patch and estimator.

IV. RESULTS

A. Bias when delensing with temperature-based reconstructions

The machinery described above allows us to evaluate Eq. (20) directly from simulations, thus isolating the bias to the power spectrum of residual lensing B-modes, $\Delta C_l^{BB,\mathrm{res}}$, that ensues after delensing. This contribution will go unmodeled in any analysis of the data that ignores the foreground non-Gaussianity, misleading parameter constraints obtained from B-mode spectra.

Before we quote any results, let us briefly explain how we will translate modeling inaccuracies to bias on the tensor-to-scalar ratio, r. To do this, we will use [89]

$$\Delta \hat{r} = \left(\sum_{l_i} \frac{[C_{l_i}^{BB}(r=1)]^2}{\mathrm{Var}(C_{l_i}^{BB,\mathrm{del}})} \right)^{-1}$$

$$\times \sum_{l_i} \frac{C_{l_i}^{BB,\mathrm{unmodeled}} C_{l_i}^{BB}(r=1)}{\mathrm{Var}(C_{l_i}^{BB,\mathrm{del}})}, \qquad (35)$$

derived from the maximum-likelihood expression for $\hat{r}$. Here, l_i indexes the ith bin; we employ three bins in the

---

[16]As discussed in Sec. III A, our lensing B-modes are measured on tiles extracted from full-sky QU maps containing only B-modes (in the full sky, there is no ambiguity between E and B, so we can separate them perfectly). On the other hand, our lensing template contains only B-modes by construction. There are therefore no E-modes in the unmasked sky in our setup, so we are susceptible to B-to-E rather than E-to-B leakage.

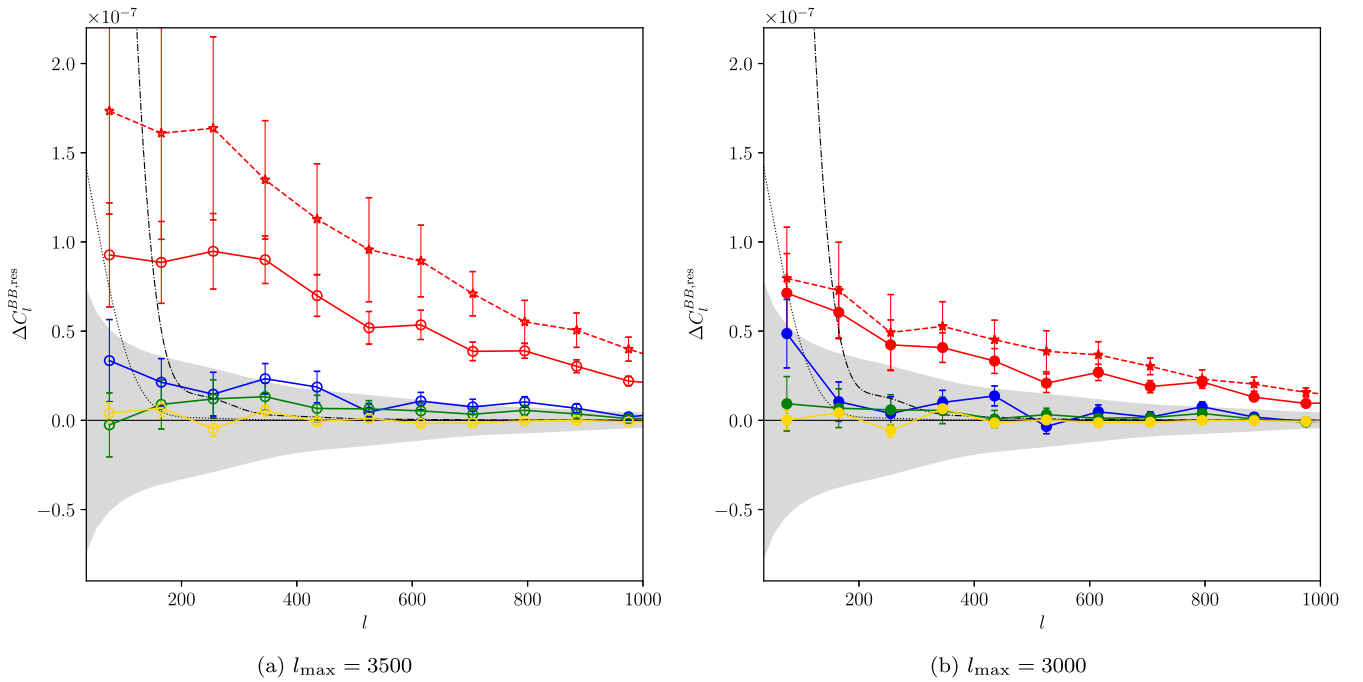(a) $l_{max} = 3500$ (b) $l_{max} = 3000$

FIG. 7.   Bias to the power spectrum of B-modes after delensing with various lensing reconstructions obtained from an experiment similar to SO. We compare the Hu-Okamoto TT (red), TE (yellow), point-source-hardened TT (blue), and shear-only TT (green) quadratic estimators. In the left panel, reconstructions are obtained from CMB modes up to $l_{max} = 3500$, and up to $l_{max} = 3000$ in the right. These input maps either include unmitigated foregrounds as they appear at 143 GHz (solid curves), or they come from a minimum-variance temperature ILC (dashed). In all cases, point sources detected with 5σ confidence at 143 GHz are removed. The curves show the mean bias calculated from 48 patches, each covering $O(1\%)$ of the sky, while the error bars show $\pm 1\sigma$ of the sample mean. (We combine the measurements between $20 < l < 1200$ into bins of width $\Delta l = 70$.) For comparison, the dashed (dotted) curve shows a tensor contribution with $r = 0.01$ ($r = 0.001$). The shaded region denotes the $\pm 1\sigma$ fluctuation interval expected of Gaussian fields with the same power spectrum as the residual lensing B-modes leftover after delensing with the Hu-Okamoto TT QE (as predicted by theory in the limit of unmitigated Gaussian foregrounds), when observed over 10% of the sky to approximate the SO SAT footprint, and under the same binning scheme as above.

range $30 \leq l \leq 300$, matching [38], with $\Delta l = 90$—the first three bins in Fig. 7. $\mathrm{Var}(C_{l_i}^{BB,\mathrm{del}})$ is the variance of the binned power spectrum of delensed-B-modes (ignoring polarized foregrounds, but including a primordial component), for a sky fraction of 10%, similar to that covered by the SO SATs; we approximate this variance by its Gaussian component. By setting $C_{l_i}^{BB,\mathrm{unmodeled}} = \Delta C_l^{BB,\mathrm{res}}$, we will be able to estimate the Δr shift caused by the foreground non-Gaussianity as a function of the true value of r.

Figure 7(a) shows the mean bias associated with each of the lensing estimators we consider, along with the standard deviation on this sample mean computed from the scatter of the simulations. For comparison, the shaded gray region shows the ± 1-σ fluctuation interval expected of the residual lensing B-modes and Small-Aperture Telescope (SAT) polarization noise of SO (taking both of these components to be Gaussian) after delensing with an unbiased HO TT QE reconstruction obtained from 143 GHz maps.

We learn that when delensing with a HO TT QE that takes in temperature modes up to $l_{max} = 3500$ measured at 143 GHz, the power spectrum is biased high by $\sim 6\%$–7%

at $l < 300$. In order to understand the source of these biases, we plot in Fig. 8 the shifts to $C^{\tilde{B} \times \hat{B}^{\mathrm{lens}}}$ and $C^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}}$ individually,[17] before they combine via Eq. (21) to give $\Delta C_l^{BB,\mathrm{res}}$. As expected from the arguments in Sec. II C, $\Delta C^{\tilde{B} \times \hat{B}^{\mathrm{lens}}}$, which can only be produced by the bispectrum bias to $\Delta C^{\hat{\kappa}\kappa}$, is negative, so it contributes a positive bias to the delensed B-mode spectrum. In parallel, $C^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}}$ is also biased low, albeit by a smaller amount—this suggests that the bispectrum biases to $\Delta C^{\hat{\kappa}\hat{\kappa}}$ are dominating over the trispectrum bias on the scales relevant for delensing. Conveniently, this means that the bispectrum contributions to $\Delta C^{\tilde{B} \times \hat{B}^{\mathrm{lens}}}$ and $\Delta C^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}}$ cancel each other out partially in the calculation of $\Delta C_l^{BB,\mathrm{res}}$.

Biases to the power spectrum of residual lensing B-modes translate—via Eq. (35)—to biases on r: the spurious, additional power leads to a systematic overestimation of the true

---

[17]This might in fact be the more relevant presentation for delensing pipelines based on cross-spectral approaches (e.g., [25,26]).
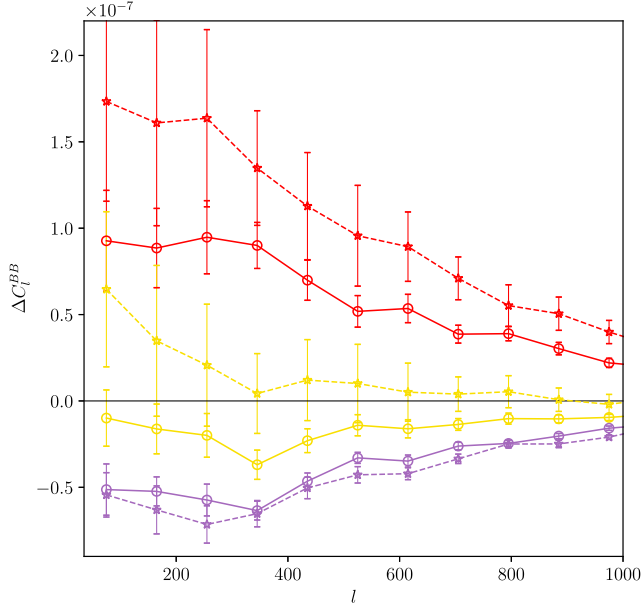
FIG. 8. Breakdown of contributions to the B-mode power spectrum bias when delensing with a HO TT QE applied to 143 GHz maps (solid) or an ILC of SO-LAT-like observations (dashed), using CMB modes $l < 3500$. The cross-correlation between template and lensing B-modes, $C^{\tilde{B} \times \hat{B}^{\mathrm{lens}}}$, is always biased low (magenta), whereas the bias to the template auto-spectrum, $C^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}}$, (yellow) can be negative when the bispectrum biases dominate on the relevant scales of $C^{\hat{\kappa}\hat{\kappa}}$, or positive when the trispectrum term does instead, as is the case for the ILC scenario. In each case, subtracting twice the magenta curve from the yellow gives the bias to the power spectrum of delensed B-modes (red), and we recover the corresponding curve in Fig. 7(a).

r. This is illustrated in Fig. 9(a), where we compare the mean inferred values of the tensor-to-scalar ratio to the truth, for various values of r. In the null scenario of $r = 0$, a naive delensing of SO-like data using a HO TT QE with $l_{\mathrm{max}} = 3500$ infers $\hat{r} = 1.7 \times 10^{-3}$, 1.5σ away from zero (though any measure of the quality of the fit would likely return a poor value), understanding σ to be the fluctuation expected of the B-mode spectrum after delensing—including residual lensing and SAT noise, but ignoring the contribution of polarized foregrounds. Note that our estimate of Δr is only the mean value measured from the simulations. This estimate is uncertain due to both sample variance and the simulations being constructed around incomplete astrophysical models of the foregrounds; due to the difficulty in quantifying the latter element, we do not assign error bars to Δr.

For completeness, we also investigate the impact of extragalactic foregrounds when delensing with a TE quadratic estimator. In Sec. II C, we warned that this pipeline is in principle vulnerable to biases associated with foreground non-Gaussianity in T; in particular, we identified a possible secondary bispectrum coupling in Eq. (32). Fortunately, the yellow curves in Figs. 7 and 9 suggest that

for the SO-like experiment we consider the bias is always consistent with zero both at the level of $C^{BB,\mathrm{res}}$ and of $\hat{r}$.

We now consider four ways of mitigating the delensing bias: (i) multifrequency cleaning, (ii) geometric methods, (iii) reducing l_{max}, and (iv) modeling it away. Except for the last one, all of these approaches entail a tradeoff between lensing reconstruction noise—and thus delensing efficiency—and bias. In order to better compare the methods on the basis of this tradeoff, we will refer to plots such as Fig. 10, where we compare Δr to $\sigma(r)$ in the null scenario where $r = 0$.

1. Multifrequency cleaning

To gauge the impact of foreground cleaning, we consider the case where the HO TT QE reconstruction is derived from a minimum-variance (MV) ILC of temperature maps at frequencies near the centers of the nominal channels of the SO LAT; see Sec. III A for details. This multifrequency combination reduces the map-level noise and consequently allows for more precise lensing reconstructions. Furthermore, Ref. [56] found that it was a useful step in the direction of mitigating biases to the auto- and cross-correlations of temperature-based CMB lensing reconstructions. This is in contrast to what we see here: Fig. 7(a) shows that the procedure actually worsens the delensing bias relative to the single-frequency scenario described above: $\Delta C_{l<300}^{BB,\mathrm{res}}$ is now at the level of $\sim 12\%$ when $l_{\mathrm{max}} = 3500$. Naturally, this translates to a larger bias on r—see the dashed curve in Fig. 9(a)—a full 3σ away from $r = 0$ in the null scenario.

Though this behavior might at first come as a surprise, it is consistent with the understanding we developed in Sec. II C, particularly in light of the effect that the ILC is known to have on the individual foreground components and the biases to lensing reconstructions: while the MV ILC suppresses the CIB extensively [56], it simultaneously boosts the tSZ component in the maps by a factor of $X \sim \sqrt{2}$–$\sqrt{3}$ on the scales relevant to lensing reconstruction (see Fig. 1 of [56]). The trispectrum is very sensitive to this boost because it scales as the fourth power of the maps, so we expect it to grow by a factor of X^4. Since the trispectrum bias is positive and now better able to cancel the negative contributions from the bispectrum biases, $\Delta C^{\hat{\kappa}\hat{\kappa}}$ moves in the positive direction. For the configuration we have chosen, this makes $\Delta C^{\hat{\kappa}\hat{\kappa}}$ smaller than it was in the single-frequency scenario, in agreement with [56]. However, this is actually detrimental to our goals because, as shown by the dashed curves in Fig. 8, a less negative $\Delta C^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}}$ spoils the cancellation with $-2\Delta C^{\tilde{B} \times \hat{B}^{\mathrm{lens}}}$ that we were seeing in the single-frequency scenario. This means that $\Delta C^{BB,\mathrm{res}}$ is larger when $\hat{\kappa}$ is reconstructed from ILC maps, despite $\Delta C^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}}$ being smaller in absolute terms and $\Delta C^{\tilde{B} \times \hat{B}^{\mathrm{lens}}}$ being practically unchanged. Moreover, the scatter in $\Delta C^{BB,\mathrm{res}}$ is larger

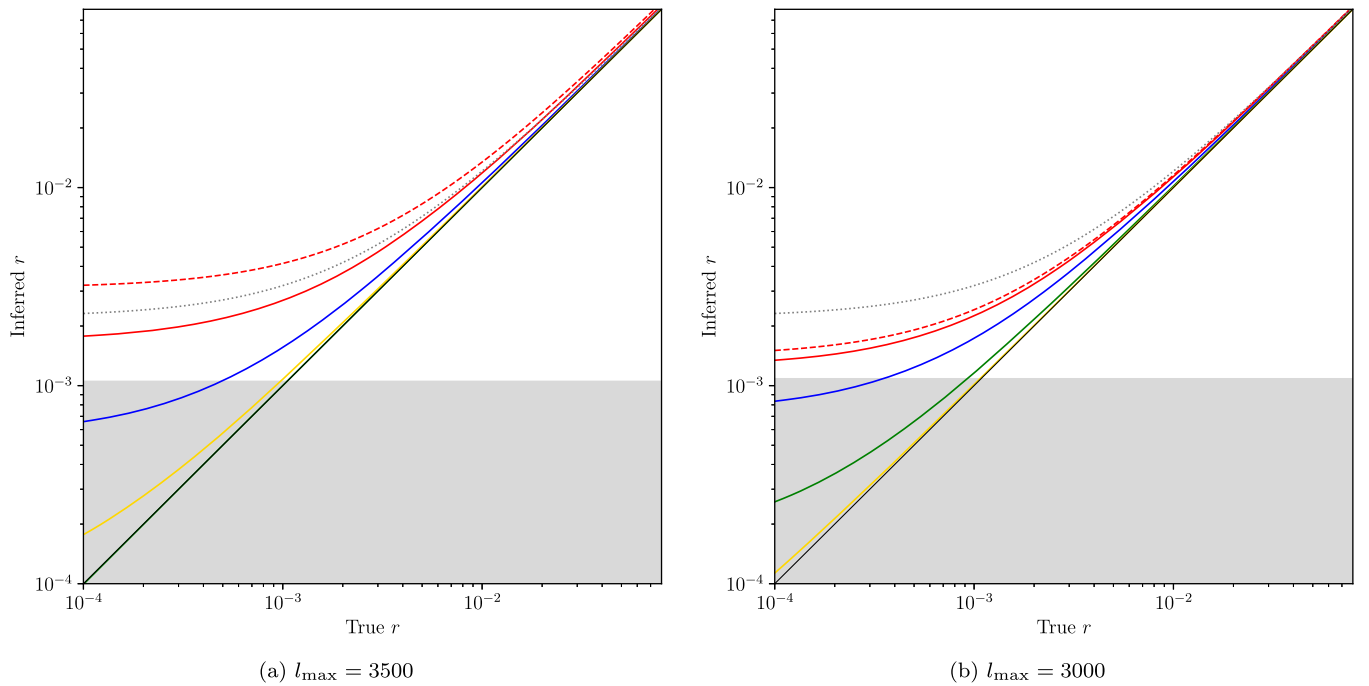(a) $l_{\max} = 3500$ (b) $l_{\max} = 3000$

FIG. 9.   Mean inferred value of r in our simulated patches vs the true input value, after delensing with the Hu-Okamoto TT (red), Hu-Okamoto TE (yellow), point-source-hardened TT (blue), or shear-only TT (green) quadratic estimators applied to CMB fields featuring non-Gaussian extragalactic foregrounds in temperature as they appear at 143 GHz (solid) or to an MV-ILC-cleaned map (dashed). Reconstructions are obtained from CMB modes up to $l_{\max} = 3500$ (left) or $l_{\max} = 3000$ (right). Inferences are biased high for all pipelines, though this is only potentially significant for the Hu-Okamoto estimator: the shaded region shows the 1σ uncertainty afforded by delensing SO data with the Hu-Okamoto TT estimator, including residual lensing B-modes and experimental noise, but no foregrounds—for reference, the dotted line shows the same calculation but featuring also Galactic foregrounds. The biases on r shown here are derived using Eq. (35).

when working with ILC-cleaned maps: this is because $\Delta C^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}}$ and $\Delta C^{\tilde{B} \times \hat{B}^{\text{lens}}}$ are correlated (they are both dominated by very similar bispectrum biases) so when they are allowed to cancel out, both the mean and the variance of the bias decrease.[18]

Given that the worsening in delensing bias when reconstructing from a MV ILC can be attributed to an increase in the tSZ trispectrum, it is tempting to consider an alternative ILC where the tSZ component is explicitly deprojected—that is, nulled by construction [90]. Though this possibility should be examined in detail in future work, we suspect it will not be all that advantageous for the case at hand. In addition to increasing the noise, tSZ deprojection is known to boost the CIB component by a factor of $\sim\sqrt{10}$ at the map level [56]. This could significantly increase the bispectrum bias, to which we are very sensitive via $-2\Delta C^{\tilde{B} \times \hat{B}^{\text{lens}}}$—despite this partially cancelling with the bispectrum bias to $C^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}}$, the latter will grow more slowly. On the other hand, a joint deprojection of both tSZ and CIB is likely to incur too large a penalty in terms of

reconstruction noise [56] and delensing efficiency. The optimal solution might result from a compromise between variance reduction and bias mitigation [56,91], or from a combination of multifrequency cleaning with geometric methods [57].

2. Geometric methods

The class of "geometric methods" comprises quadratic estimators constructed using carefully chosen weights which differ from those of the standard HO QE and afford the new estimator more immunity to foreground contamination. In Sec. II B, we introduced the point-source-hardened and shear-only estimators, which have proven their worth in mitigating biases to CMB lensing spectra [52,55,57] and appear to be similarly effective against delensing biases.

The blue curve in Fig. 7(a) demonstrates that when the delensing pipeline involves a point-source-hardened estimator applied to 143 GHz temperature maps, $\Delta C^{BB,\text{res}}$ is reduced very substantially relative to the case where the standard HO QE is used. Consequently, the bias to r is also significantly reduced and is now well within the 1-σ level, as seen in Fig. 9(a). Moreover, as shown in Fig. 11, the total

---

[18]We are grateful to Marius Millea for directing our attention to the question of variance.

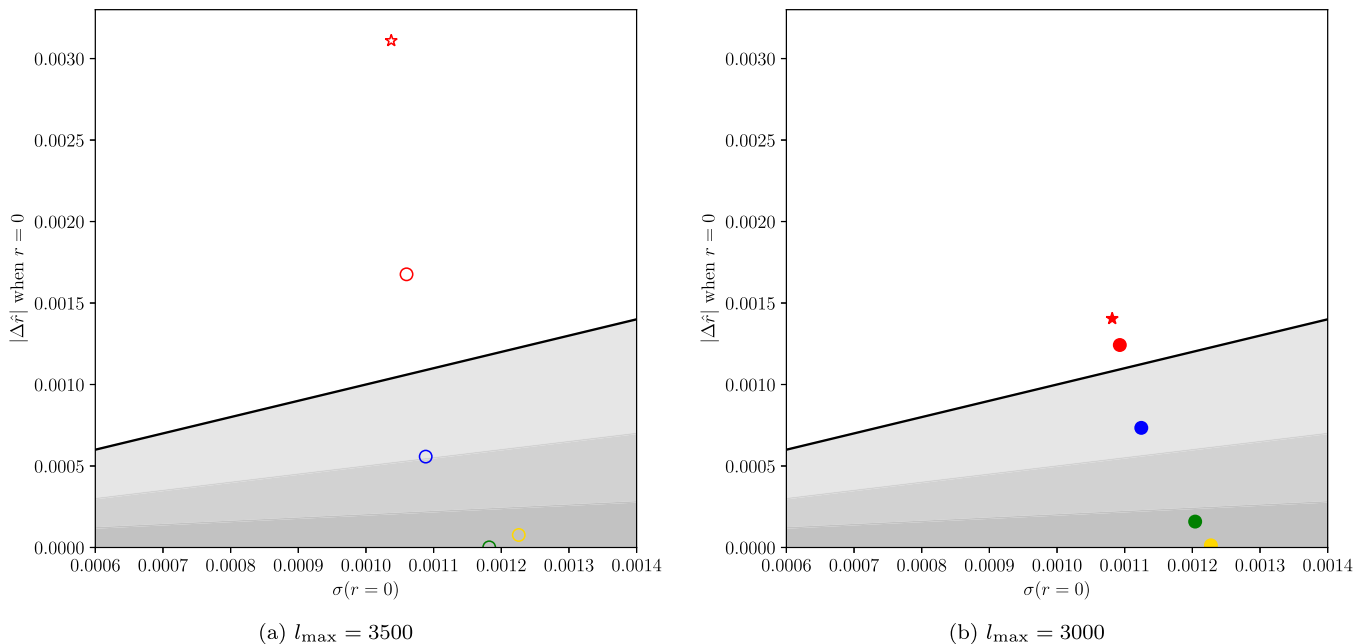(a) $l_{\max} = 3500$ (b) $l_{\max} = 3000$

FIG. 10.   Bias to r vs statistical uncertainty postdelensing for pipelines involving various quadratic estimators of lensing, in the null scenario where $r = 0$. Here, $\sigma(r)$ is as for the SO SAT's 143 GHz channel white noise levels and approximate sky coverage of 10%, with the residual lensing afforded by each estimator, but no polarized foregrounds. The transition between regions of different shading intensity happen at (from bottom to top) $|\Delta\hat{r}|/\sigma^{\mathrm{HO}}(r) = 1/5, 1/2, 1$, where σ^{HO} is appropriate for the HO TT QE applied to 143 GHz maps. The colors are as in Figs. 7 and 9, with the difference that the star now denotes the MV-ILC-cleaned pipeline. We see that mitigation strategies based on geometric arguments are most effective at suppressing the bias, while only incurring a moderate degradation in delensing efficiency: they allow us to walk an almost vertical, downward line on the Δr-$\sigma(r)$ plane.

amount of B-mode power after delensing is lower than when the HO QE is used: although the PSH reconstruction is slightly noisier (see Fig. 2) and thus results in somewhat less extensive removal of lensing B-modes, this is more than compensated for by the lower amount of spurious power that it receives from foreground non-Gaussianity. All in all, this estimator stands out for its ability to mitigate the bias to acceptable levels at very little cost in terms of delensing efficiency: in Fig. 10(a), we see that it reduces Δr by a factor of 3 relative to the HO QE, while only degrading $\sigma(r = 0)$ by a negligible amount.

We also consider the shear-only estimator, shown in green in the figures, which is even better at suppressing the bias and comes close to completely neutralizing it. It is actually not surprising that $\Delta C^{BB,\mathrm{res}}$ is not exactly zero despite the foregrounds in the simulation being azimuthally symmetric. The estimator is only immune to them in the regime where large-scale lenses are reconstructed from CMB fluctuations with much smaller angular sizes, but B-mode delensing hinges on relatively small-scale reconstructed lenses, so the separation of scales required for exact immunity is not fully satisfied. At any rate, this estimator is the least-biased out of all the ones we have explored, yielding a Δr in Fig. 9a that is consistent with zero. The downside is that it comes with greater degradation in delensing efficiency, encapsulated by a larger $\sigma(r)$ in Fig. 10(a).

3. Lowering $l_{\max}$

Finally, we consider restricting the $l_{\max}$ of the CMB temperature fields from which lensing is reconstructed. This is expected to limit the contamination from extragalactic foregrounds because the latter have a bluer angular spectrum than the CMB—see Fig. 1—exceeding it in amplitude beyond $l \sim 3000$ (in the science channels).

Figure 7(b) shows the bias to $\Delta C^{BB,\mathrm{res}}$ for various estimators in the case of $l_{\max} = 3000$. Notice that it is reduced in all cases relative to the results in Fig. 7(a), for which $l_{\max} = 3500$. The same is true for the biases to r in Figs. 9(b) and 10(b), although it is still above the 1-σ level for the standard HO QE, in both the ILC-cleaned and unmitigated scenarios. Since lowering $l_{\max}$ discards information that could otherwise have been used to reconstruct lensing, the delensing efficiency is worsened in all cases, yielding values of $\sigma(r)$ in Fig. 10(b) higher than what we were seeing in Fig. 10(a).

4. Modeling

To the extent that the terms in Eqs. (22) and (23) are the main sources of bias, the effect of the extragalactic foregrounds can be incorporated into models for the power spectrum of delensed B-modes—and the bias removed—simply by using empirically calibrated $C^{\kappa\hat{\kappa}}$ and $C^{\hat{\kappa}\hat{\kappa}}$ in Eqs. (24) and (25). In this approach, $C^{\hat{\kappa}\hat{\kappa}}$ would come from
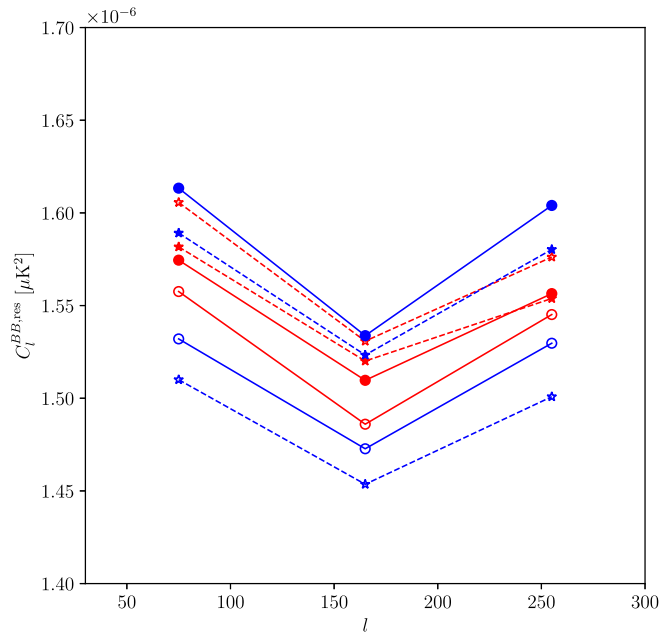
FIG. 11. Mean delensed-B-mode band powers for pipelines involving different QEs applied to simulations featuring non-Gaussian foregrounds. (Note that the delensed spectrum is in fact rather flat, though the restricted y scale does not make it seem that way.) We consider the Hu-Okamoto (red) and point-source-hardened (blue) QEs, applied to 143 GHz (circle; solid lines) or MV ILC maps (star; dashed lines), using temperature modes $l < 3000$ (filled symbol) or $l < 3500$ (empty symbol). All things being equal, the HO TT allows for more removal of lensing B-modes, but it is also more prone to receiving additional power from the foreground non-Gaussianity. Consequently, among the options we consider, the largest suppression of B-mode power is obtained by delensing with a point-source-hardened QE applied to MV ILC maps up to $l_{\max} = 3500$.

a smooth fit to the lensing reconstruction auto-spectrum, while $C^{\kappa\hat{\kappa}}$ would be the result of applying a similar procedure to the cross-correlation of polarization-only and TT reconstructions.[19]

Figure 12 demonstrates that a model built this way tracks the simulated data very accurately. The solid curves show the difference in model predictions when $C^{\kappa\hat{\kappa}}$ and $C^{\hat{\kappa}\hat{\kappa}}$ are measured from simulations featuring Gaussian versus non-Gaussian foregrounds (we take the average lensing spectra measured across all 48 patches). Notice that the prediction for $\Delta C^{\tilde{B}\times\hat{B}^{\mathrm{lens}}}$, $\Delta C^{\hat{B}^{\mathrm{lens}}\times\hat{B}^{\mathrm{lens}}}$, and $\Delta C^{BB,\mathrm{res}}$ are all in excellent agreement with the data. This also implies that the terms we highlighted in Eqs. (22) and (23) do indeed capture most of the non-Gaussian effects produced by the foregrounds.

An important benefit of the modeling approach is that it relaxes the requirement to mitigate the bias—usually at the cost of delensing efficiency—given that we are now able to

---

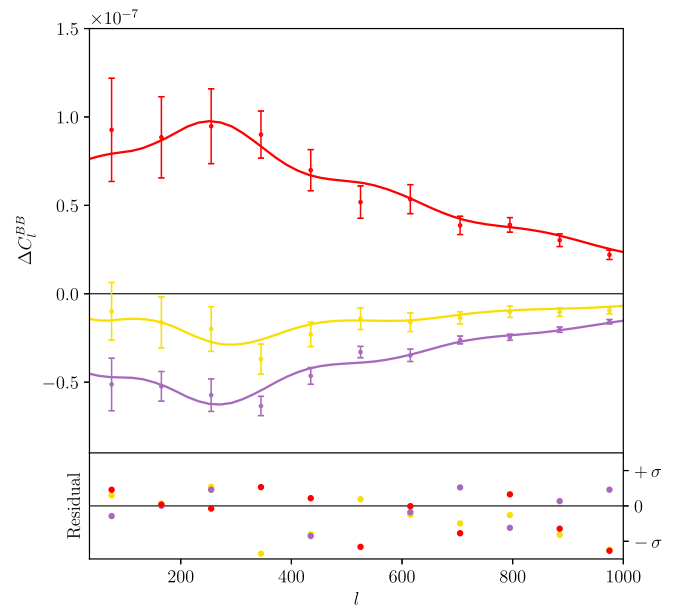[19]We are grateful to Anthony Challinor for suggesting this way of obtaining $C^{\kappa\hat{\kappa}}$.

FIG. 12. Measured impact of foreground non-Gaussianity on delensed-B-mode spectra compared to predictions from empirically calibrated models. The data points show the mean bias to $C^{\tilde{B}\times\hat{B}^{\mathrm{lens}}}$ (yellow), $C^{\hat{B}^{\mathrm{lens}}\times\hat{B}^{\mathrm{lens}}}$ (magenta), and $C^{BB,\mathrm{res}}$ (red) arising from a delensing pipeline relying on a HO TT QE applied to unmitigated maps ($l_{\max} = 3500$), with error bars denoting $\pm 1\sigma$ on the sample mean. (These data are the same as the solid curves in Fig. 8.) The curves, on the other hand, show the difference in model prediction when the $C^{\kappa\hat{\kappa}}$ and $C^{\hat{\kappa}\hat{\kappa}}$ used to evaluate Eqs. (24) and (25) are measured from simulations featuring non-Gaussian versus Gaussian foregrounds. The bottom panel shows the residuals between data and theory when both are binned the same way, normalized to the standard deviation of each datum. This demonstrates that a very accurate model for the delensed data can be built as long as $C^{\kappa\hat{\kappa}}$ and $C^{\hat{\kappa}\hat{\kappa}}$ are determined empirically and are subject to the effects of foreground non-Gaussianity.

model the combination of residual lensing and foreground effects. The focus can then be shifted towards determining what analysis choices minimize the power spectrum of B-modes after delensing, whatever its composition in terms of lensing or foreground contributions. These choices include what estimator is used, whether or not a foreground-cleaning procedure is applied, how extensively point sources are masked, what $l_{\max}$ is used for the reconstructions, etc. Unfortunately, the possibility of getting a reduction in power (and thus variance) "for free" is dispelled by the discussion in Sec. II C, which suggests that the foregrounds will inevitably add power to the power spectrum of delensed B-modes.

This last point also questions the assumption that delensing with a HO QE will result in lower power than doing so with any other QE. While it is true that, all other things being equal, the HO QE should allow the most extensive removal of lensing B-mode power, it is also the one most susceptible to receiving additional power from the foreground non-Gaussianity. Figure 11 illustrates the importance of this tradeoff. Among the estimators we

consider, the point-source-hardened estimator applied to MV ILC temperature maps up to $l_{\max} = 3500$ is in fact the one that ultimately results in the lowest amount of B-mode power after delensing. Even though we can model equally well the delensed B-mode spectrum resulting from all pipelines, this one is the most powerful one when searching for a primordial component, because of the lower variance.

B. Bias to a multitracer delensing pipeline

1. Theory

In the previous section, we quantified biases to the delensing procedure when lensing is reconstructed using only TT or TE quadratic estimators. In practice, it is unlikely that these reconstructions will be used in isolation. Instead, the $\hat{\kappa}$ estimate will probably be constructed as a combination of different estimators—other quadratic combinations of CMB fields [33,92], such as EE or EB, but also maps of tracers of the large-scale structure that are "external" to the CMB [21,26,31,36,37,93]. The advantage of this multitracer approach, illustrated in Fig. 13, is that the co-added tracer maintains a higher degree of correlation with CMB lensing than do any of the tracers individually, and the correlation can be relatively high across the scales of importance to degre-scale-B-mode delensing (see Fig. 3). In this section, we investigate what happens to the delensing bias when temperature-based reconstructions are co-added with other tracers of the mass distribution—as we shall soon see, the problem is not trivial.

Be they internal or external, the tracers can all be combined in an optimal manner as $\hat{\kappa}_{LM}^{\mathrm{comb}} = \sum_i c_L^i \hat{\kappa}_{LM}^i$, where [36]

$$c_L^i = \sum_j (\rho^{-1})_L^{ij} \rho_L^{j\kappa} \sqrt{\frac{C_L^{\kappa\kappa}}{C_L^{\hat{\kappa}^i \hat{\kappa}^i}}}. \tag{36}$$

Here, $\rho_L^{i\kappa}$ is the cross-correlation coefficient, at multipole L, between tracer $\hat{\kappa}^i$ and the true convergence; ρ_L^{ij} is the cross-correlation between tracers $\hat{\kappa}^i$ and $\hat{\kappa}^j$; and $C_L^{\hat{\kappa}^i \hat{\kappa}^i}$ is the angular power spectrum of tracer $\hat{\kappa}^i$. By design, these weights maximize the cross-correlation between the co-added tracer and the true convergence; in the limit of a single tracer, they reduce to the standard Wiener filter in Eq. (34).

We will use the weights above to combine TT and TE QEs with external tracers and with other QEs. Reference [92] recently pointed out that forming the minimum-variance
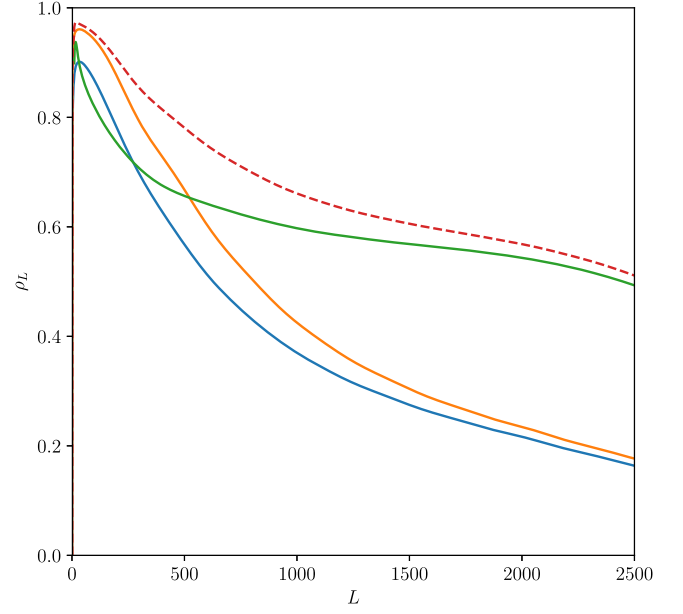

FIG. 13. Correlation with CMB lensing of various matter tracers: a HO TT QE (blue), a minimum-variance combination of all the HO QEs (orange), a Planck-like measurement of the CIB at 545 GHz (green), and the optimal combination of them all (dashed red). All internal reconstructions are derived from CMB modes up to $l_{\max} = 3500$. Notice that the co-added tracer is dominated by internal reconstructions on large scales, where the correlation with κ is very high, and by the CIB on smaller scales. When combined, the resulting tracer is highly correlated with lensing across the scales that are relevant to large-scale-B-mode delensing; cf. Fig. 3.

combination of separate QEs—the approach of Hu and Okamoto [33], implicit in Eq. (36)—is in principle less optimal than constructing a single, "global minimum variance" (GMV) estimator that finds the combination of pairs of CMB fields that affords the least variance at each scale. Our decision to work with the Hu-Okamoto approach stems from the fact that the difference between methods is expected to be small for SO, and the approach presented here offers more analytic transparency into a problem that is expected to affect the GMV algorithm as well.

Once again, we can use Eq. (20) to estimate the delensing bias due to non-Gaussian, extragalactic foreground emission in the observed temperature maps; this time, to a multitracer pipeline. Assuming the primordial CMB is Gaussian, the bias terms are, schematically,

$$
\begin{aligned}
\Delta C_l^{BB,\mathrm{res}} = &-2g_l[c^{TT}\langle \tilde{B}\,\tilde{E}\,\hat{\kappa}^{TT}[f_T + s^{\mathrm{NG}}, f_T + s^{\mathrm{NG}}]\rangle] + h_l[(c^{TT})^2 \langle |E^{\mathrm{obs}}\hat{\kappa}^{TT}[f_T + s^{\mathrm{NG}}, f_T + s^{\mathrm{NG}}]|^2\rangle] - (s^{\mathrm{NG}} \to s^{\mathrm{G}}) \\
&-2g_l[c^{TE}\langle \tilde{B}\,\tilde{E}\,\hat{\kappa}^{TE}[f_T + s^{\mathrm{NG}}, f_E]\rangle] + h_l[(c^{TE})^2 \langle |E^{\mathrm{obs}}\hat{\kappa}^{TE}[f_T + s^{\mathrm{NG}}, f_E]|^2\rangle] - (s^{\mathrm{NG}} \to s^{\mathrm{G}}) \\
&+2h_l[c^{TT}c^{TE}\langle E^{\mathrm{obs}}\hat{\kappa}^{TT}[f_T + s^{\mathrm{NG}}, f_T + s^{\mathrm{NG}}]E^{\mathrm{obs}}\hat{\kappa}^{TE}[f_T + s^{\mathrm{NG}}, f_E]\rangle] - (s^{\mathrm{NG}} \to s^{\mathrm{G}}) \\
&+2h_l\left[c^{TT}\left\langle E^{\mathrm{obs}}\sum_{i \neq TT,TE} c^i \hat{\kappa}^i E^{\mathrm{obs}}\hat{\kappa}^{TT}[f_T + s^{\mathrm{NG}}, f_T + s^{\mathrm{NG}}]\right\rangle\right] - (s^{\mathrm{NG}} \to s^{\mathrm{G}}).
\end{aligned}
\tag{37}
$$

The last term in each line represents a duplicate of the previous terms in the line with non-Gaussian foregrounds replaced by their Gaussian version.

The first line is essentially the TT delensing bias in Eq. (21), diluted by the corresponding multitracer weight c^{TT}. Similarly, the second line is the diluted version of the TE bias. The third and fourth lines, on the other hand, are new.

The third line is a cross-term between TT and TE reconstructions which produces a plurality of contributions. The two couplings that are likely to stand out among these in terms of significance are

$$\Delta C_l^{BB,\mathrm{res}} \supset$$
$$\supset 2\, h_l \left[c^{TT} c^{TE} \langle \overline{E^{\mathrm{obs}} \hat{\kappa}^{TT} \left[s^{\mathrm{NG}}, s^{\mathrm{NG}} \right]} E^{\mathrm{obs}} \hat{\kappa}^{TE} \left[\tilde{T}, \tilde{E} \right] \rangle_c \right]$$
$$+ 4\, h_l \left[c^{TT} c^{TE} \langle \overline{E^{\mathrm{obs}} \hat{\kappa}^{TT} \left[\tilde{T}, s^{\mathrm{NG}} \right]} E^{\mathrm{obs}} \hat{\kappa}^{TE} \left[s^{\mathrm{NG}}, \tilde{E} \right] \rangle_c \right],$$

(38)

that is, terms where we take the Gaussian contraction of the E-modes that appear explicitly in the template. The first term above is similar in nature to the primary bispectrum bias of CMB lensing spectra; the second is more similar to the secondary bispectrum bias.

The fourth line produces contributions such as[20]

$$\Delta C_l^{BB,\mathrm{res}} \supset 2h_l \left[c^{TT} c^i \langle \overline{E^{\mathrm{obs}} \hat{\kappa}^i} E^{\mathrm{obs}} \hat{\kappa}^{TT} [s^{\mathrm{NG}}, s^{\mathrm{NG}}] \rangle_c \right],$$

(39)

where $\hat{\kappa}^i$ is any matter tracer except for the TT and TE QEs. Cross-terms like this appear in two qualitatively different, though highly related scenarios. The first case is when the matter tracer features the foregrounds explicitly; for example, when the CIB or the galaxy positions in an imaging survey are used as tracers of the matter (the galaxy maps can feature star-forming galaxies that also make up the CIB, or "radio galaxies" whose emission reaches the microwave range of the spectrum). When this is the case the term is a function of the foreground bispectrum $\langle sss \rangle$.

On the other hand, a contribution will arise even if the tracers are foreground-free, merely due to the fact that tracers and foregrounds are correlated because they both trace the same underlying matter distribution. This will be the case for the EE, EB, or TB reconstructions. (More complicated, though in principle possible, is the case when the weights used to correct for systematic effects in galaxy surveys have large-scale-structure residuals; see, e.g., [94]). The term is then a function of the mixed lensing-foreground

bispectrum $\langle \kappa ss \rangle$, a coupling similar to the primary bispectrum bias described in Sec. II C; as we saw there, we expect this contribution to be negative.

Note that the couplings we have highlighted can be modeled in a manner similar to what we described in Sec. IV A 4 for a single tracer. The diluted TT and TE contributions can be modeled following the arguments there, but including the multitracer weights in the relevant integrands. Term (39) can be calculated by modifying Eq. (25) to include the multitracer weights c^{TT} and c^i in the integrand, and replacing $C^{\hat{\kappa}\hat{\kappa}}$ with (a smoothed version of) the measured cross-spectrum between a TT reconstruction and the tracer in question—a CIB map, or perhaps the EB, EE, or TB QE reconstructions. The same goes for the terms in Eq. (38), though this time the multitracer weights are c^{TT} and c^{TE}, and $C^{\hat{\kappa}\hat{\kappa}}$ is calibrated against the measured cross-correlation between TT and TE reconstructions.

Alternatively, terms like Eq. (39) above can be mitigated by reducing the overlap in multipole space of the multitracer weights. Though impractical when striving to mitigate the bias from the cross-correlation of $\hat{\kappa}^{TT}$ with other internal reconstructions, this approach is certainly feasible when targeting the correlation with external tracers, which are complementary to internal reconstructions in terms of the scales that they target: while internal reconstructions can only accurately reconstruct the largest lenses, external tracers are good at providing the high-L ones; see Fig. 13. Such cuts have already been explored in the context of CIB delensing, either to reduce biases from Galactic dust [37,65], or because of mode loss during component separation [95]; it appears that the delensing efficiency is not majorly affected even if fairly stringent cuts such as the removal of $L < 100$ are put in place. Moreover, it has recently been shown that when internal reconstructions are prioritized in the provision of the largest-angular-scale lenses, any possible residual in the power spectrum of delensed B-modes tends to flatten out and can be easily marginalized over [26].

2. Methods

Now that we have laid out the theory of what couplings could potentially bias a multitracer delensing pipeline, we set out to quantify them for an experiment similar to the Simons Observatory, with characteristics described in Sec. III A. In what follows, we co-add internal reconstructions from various QEs with CIB maps intended to mimic Planck's observations, thus testing the impact of all the bias terms in Eq. (37).

We carry out the TT and TE reconstructions as described in Sec. III B; on the other hand, and for the sake of simplicity, we approximate the TB, EE, and EB reconstructions as the sum of the input convergence and noise realizations drawn from a Gaussian distribution with the same power spectrum as the $N^{(0)}$ noise corresponding to that estimator, computed using Eq. (14). In all cases, we set $l_{\mathrm{max}} = 3500$, and apply the estimators to the "unmitigated"

---

[20]Note that we have assumed that the primordial CMB is Gaussian, in which case there are no similar contributions associated with the TE reconstruction.
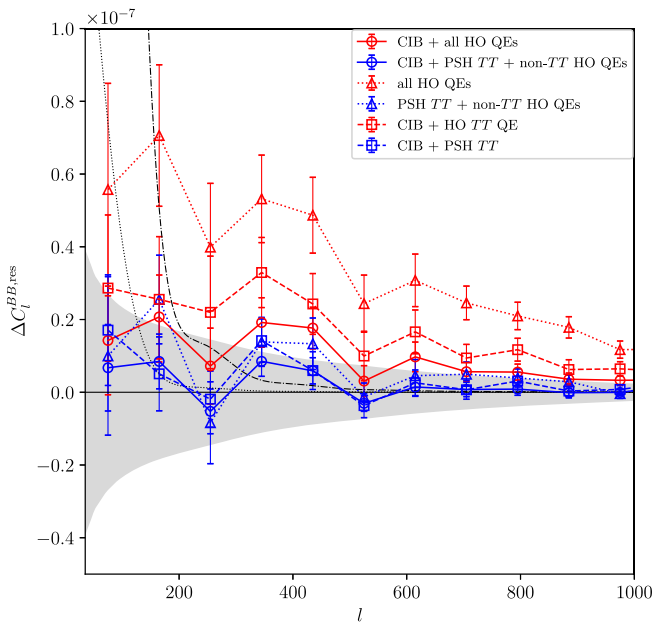
FIG. 14. Bias to the power spectrum of B-modes after delensing with various multitracer pipelines. (Note that PSH stands for the point-source-hardened QE.) The shaded region denotes the $\pm 1\sigma$ fluctuation interval expected of Gaussian fields with the same power spectrum as the residual lensing B-modes leftover after delensing with all the Hu-Okamoto QEs and a Planck-like CIB map at 545 GHz (as predicted by theory in the limit of unmitigated Gaussian foregrounds), when observed over 10% of the sky to approximate the SO SAT footprint, and under the same binning scheme as above. Note that the y scale differs from Fig. 7.
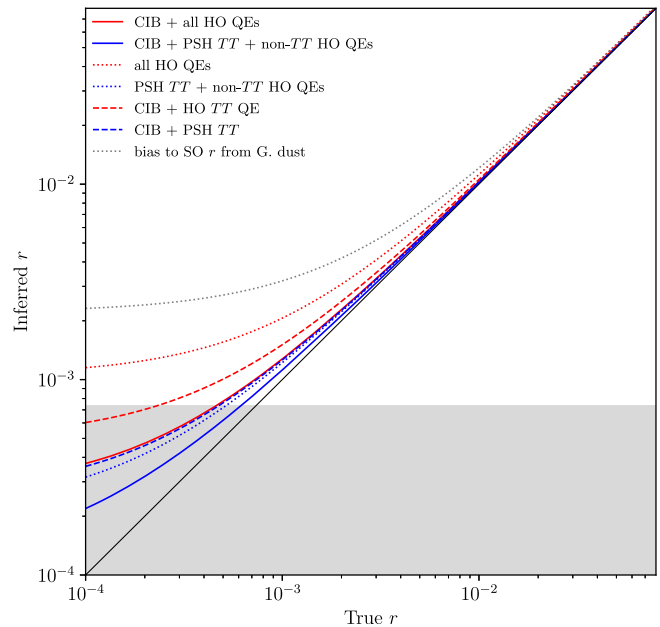

FIG. 15. Mean inferred value of r in our simulated patches vs the true input value, after multitracer delensing. These biases in r are to be attributed to the power spectrum biases in Fig. 14. The shaded region shows the 1σ uncertainty afforded by delensing SO data with a combination of the CIB plus all the HO QEs (including residual lensing B-modes and experimental noise, but not foregrounds—for reference, the dotted black line shows the same calculation but featuring also Galactic foregrounds).

143 GHz maps—that is, we do not perform any foreground cleaning except for removing point sources. We then take the Websky CIB map at 545 GHz as an approximation to the CIB map that would be used as a matter tracer for delensing, making sure to add detector noise as appropriate for the equivalent channel of Planck, and removing point sources that fall above the detection limit of 350 mJy by setting the pixels to the mean value of the remaining, unmasked pixels.

In order to calculate the multitracer weights of Eq. (36) and combine these tracers optimally, we need all the auto- and cross-spectra of the tracers involved. We measure the auto-correlation of the noiseless 545 GHz CIB map, as well as its cross-correlation with the input convergence map, on the full sky. We then add an idealized noise power spectrum to the auto-spectrum (using the instrument characterization described in [96]), and fit smooth functions to the resulting auto- and cross-spectra.[21] The remaining ingredients, once the CIB angular spectra have been measured, are all the auto- and cross-spectra of the internal reconstructions. To simplify matters, we assume that the noise covariance

matrix between different estimators is diagonal,[22] and model their cross-correlation as the true convergence power spectrum; on the other hand, we approximate the individual auto-spectra as a sum of the true convergence power spectrum and the $N^{(0)}$ noise biases.[23] From these spectra, we can also calculate the correlation of the various tracers with CMB lensing; this is shown in Fig. 13.

We then build the template following Sec. III C, subtract it from lensing B-mode maps, and measure the power spectrum of the resulting, delensed B-modes using the methodology in Sec. III D. By applying this pipeline to temperature fields featuring either Gaussian or non-Gaussian foregrounds, we can evaluate Eq. (37) for our co-added tracer and isolate the relevant delensing biases.

3. Results

Figure 14 shows the bias to the power spectrum of B-modes delensed using various multitracer pipelines. These translate to the biases on the inferred value of r shown in Fig. 15.

---

[21]Reference [37] showed that constructing the multitracer weights from smooth spectra avoids biasing the delensed B-mode spectrum.

[22]Different HO quadratic estimator do in fact have correlated noise (with the exception of TT and EB), but this is below the 10% level [33].

[23]For the EB estimator, we calculate the reconstruction noise assuming that the B-modes going into the estimator are masked below $l < 300$; this would be required when delensing at the map level in order to avoid bias due to the overlap in modes with the target B-modes [40,42].
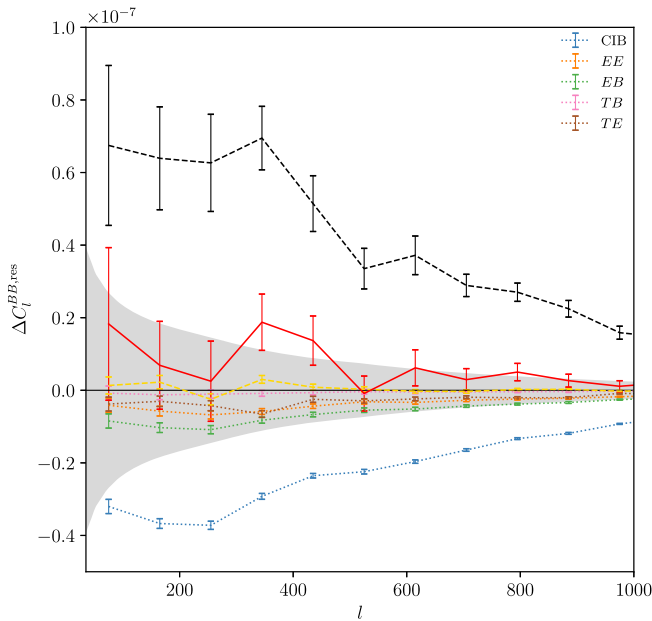
FIG. 16. Breakdown of contributions to $\Delta C_l^{BB,\mathrm{res}}$ when delensing with a combination of TT, TE, TB, EE, and EB HO QEs and a Planck-like map of the CIB at 545 GHz. The shaded region shows the $\pm 1\sigma$ (Gaussian) uncertainty due to residual lensing B-mode power when observed over 10% of the sky, and the binning shown. The diluted bias from TT reconstruction (dashed, black) partially cancels with the cross-terms in Eq. (37) (dotted; the legend shows which tracer is involved)—especially, with the coupling involving the CIB. Also shown are the TT-TE cross-term (dotted, brown) and the diluted TE bias (dashed, yellow). All of these effects combine to give the red solid curve, which matches the total bias (red, solid curve in Fig. 14) up to sample variance.

The first thing to note is that the hierarchy in bias amplitude of the different pipelines involving the HO TT QE cannot be explained just in terms of the dilution of the TT reconstruction bias. This is a hint that the other couplings in Eq. (37) are playing a significant role. Take, as a case study, a multitracer pipeline built around a combination of the $\{TT, TE, TB, EE, EB\}$ QEs of Hu and Okamoto, combined with a Planck-like map of the CIB at 545 GHz. In Fig. 16, we dissect the total bias resulting from this pipeline into its various constituent contributions.[24] Notice that not only are the cross-terms in Eq. (39) significant, but they also partially cancel with the diluted bias from TT reconstruction, yielding a lower bias (red, solid) than would be expected merely on the grounds of dilution (black, dashed). Naturally, the coupling is particularly strong, and the cancellation more extensive, when the CIB is used as an external tracer for delensing, but the effect is qualitatively the same for any other tracer of the matter distribution, including the EE, EB, and TB reconstructions. The TT-TE cross-term, Eq. (38)

and similar, also contribute to the cancellation, albeit by a smaller amount.

These cancellations are good news from the point of view of delensing. They suggest that it might be possible to use the more contaminated—but more effective—HO TT QE down to smaller scales (higher l_{max}) than one would be inclined to pursue based on the findings in Sec. IV A. Consider Fig. 17, where we compare Δr to $\sigma(r)$ after delensing with various multitracer pipelines, in the null scenario where $r = 0$. When the HO TT QE is co-added with all the other HO QEs, the bias is above 1σ. Coadding instead with the CIB not only improves the delensing effciency, but it also gives a lower bias on r thanks to the cancellations brought about by a strong TT-CIB cross-term from Eq. (39). Finally, when all the HO QEs are combined with the CIB, the delensing efficiency,[25] the dilution, and the cancellations are all at a

FIG. 17. Bias to r vs statistical uncertainty after delensing with various multitracer pipelines, in the null scenario where $r = 0$. Here, $\sigma(r)$ is as for the SO SAT's 143 GHz channel white noise levels and approximate sky coverage of 10%, with the residual lensing afforded by each pipeline, but no polarized foregrounds. The transition between regions of different shading intensity happen at (from bottom to top) $|\Delta \hat{r}|/\sigma_0(r) = 1/5, 1/2, 1$, where σ_0 is appropriate for the combination of a HO TT QE with all the other HO QEs (applied to 143 GHz maps) and a Planck-like map of the CIB at 545 GHz. The pipelines are the same as in Figs. 14 and 15, and the colors and symbols also match.

---

[24]We compute the last line in Eq. (37) directly from Eq. (39), with nothing but foregrounds in the inputs to the QE. This explains the smaller error bars shown by these terms in Fig. 16 relative to other contributions.
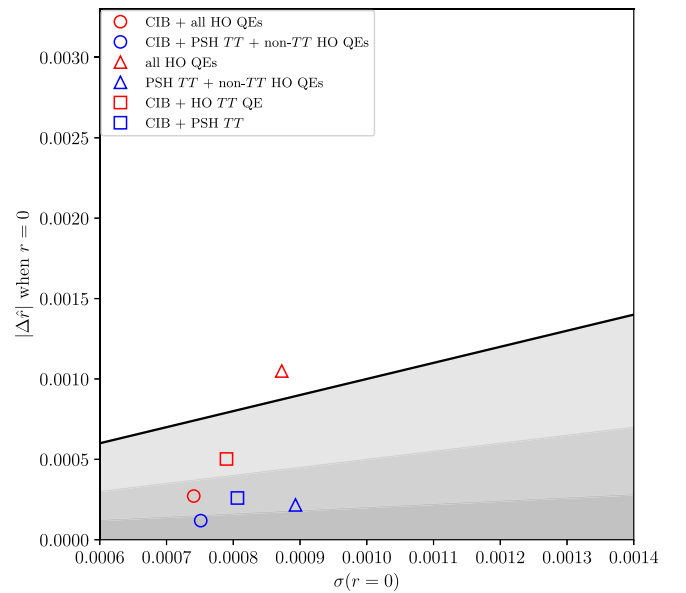
[25]Note that throughout this paper we see more extensive delensing than was forecasted in [26]. There are several reasons for this. First, we ignore atmospheric noise, which led [26] to discard temperature modes below $l < 500$. Second, we constrain r over the BB range $30 < l < 300$, matching [38], whereas [26] used $50 < l < 200$. Third, we often report results for $l_{\mathrm{max}} = 3500$, while they only ever consider $l_{\mathrm{max}} = 3000$. And last, but certainly not least, we assume the SAT observes 10% of the sky, as in [38], whereas [26] works with a smaller patch covering 5% of the sky.

maximum (among the cases we consider), and the bias is below $\sigma/2$.

If needed, the bias could be suppressed even further by resorting to the methods described in Sec. IV A: reducing $l_{\max}$ for the TT estimator, carefully removing foregrounds (bearing in mind that foreground removal is not guaranteed to improve the situation, as seen above), or adopting alternative, more robust estimators such as those described in Sec. II B. As an example of the latter approach, Fig. 17 shows that replacing the HO TT QE with the point-source-hardened estimator reduces the bias while only incurring a very modest when it comes to removing lensing B-modes.

V. CONCLUSIONS

Delensing is already an essential ingredient in any effort to constrain primordial gravitational waves from the B-mode polarization of the CMB. In this work, we characterized, for the first time, a potential source of bias to delensing pipelines: the non-Gaussian nature of extragalactic foregrounds. We show that the effect can be major when delensing hinges on internal reconstructions of lensing derived from the CMB temperature anisotropies, as will be the case for SO. However, it will be relevant more generally, including for the South Pole Observatory and CMB-S4, whenever temperature information is used to improve the quality of the lensing recontruction (as in, e.g., [39]), even if polarization dominates.

The problem is that the foreground non-Gaussianity induces higher-point contributions that add power to the power spectrum of delensed B-modes, potentially confusing inference pipelines that ignore the foregrounds or approximate them as being Gaussian. In particular, the cross-correlation of a lensing B-mode template constructed from a standard TT QE with the true B-modes is biased low, whereas the auto-spectrum of such a template can be biased either low or high, depending on the relative strength of the bispectrum and trispectrum biases of lensing. When the former dominates, convenient cancellations appear in the power spectrum of delensed B-modes between biases to the template auto- and cross-spectrum. This has surprising implications: when the TT estimator is applied to a MV ILC of SO-LAT-like observations, the bias to the power spectrum of delensed B-modes is larger and has higher variance than in the case of no foreground cleaning, despite the bias to the template auto-spectrum being smaller and the bias to the cross-spectrum being practically unchanged—this is a consequence of the cancellation between the two being spoilt.

We use the Websky simulations to study in detail the case of SO—ignoring polarized foregrounds—and find that when delensing with a HO TT estimator applied to CMB modes $l < 3500$, the bias can be as large as 1.5σ before foreground cleaning, or 3σ when reconstructing from a MV ILC of temperature maps. Though it is in principle also possible for the TE estimator to be biased by

the foreground non-Gaussianity, we find this effect to be negligible for SO.

In real analyses, the TT estimator will likely be combined with other tracers—other internal reconstructions, or perhaps external tracers. We have shown that when this is the case, the power spectrum of delensed B-modes receives new, nontrivial couplings beyond a simple dilution of the spurious power found in the case of a single tracer. Fortunately, these new terms appear to work in our favor, as they partially cancel with the diluted ones. For example, when the CIB is used in combination with TT and other QEs to delens SO data, the cross-term cancels most of the diluted $\Delta C^{BB,\mathrm{res}}$, and allows the $l_{\max}$ in the TT reconstruction to be larger than one would naively expect before receiving too large amounts of additional power and either incurring too large a bias, or degrading the delensing efficiency more than desired.

The impact of foreground non-Gaussianity can be modeled very accurately using simple analytic expressions, as long as the lensing spectra these calculations rely on— the auto-spectrum of TT lensing reconstructions, and their cross-correlation with CMB lensing—come from smooth fits to measurements that are themselves affected by the foreground non-Gaussianity. Once the non-Gaussian effects can be modeled, we are no longer obligated to prioritize bias mitigation over reconstruction efficiency (and thus lensing B-mode removal), and the goal becomes to determine what combination of analysis choices will result in the least amount of B-mode power after delensing, be it due to residual lensing or foreground-related contributions.

In this context, alternative TT QEs such as the point-source-hardened or the shear-only estimator are likely to be crucial, as they are by design more robust to foregrounds than the Hu-Okamoto QE. On the one hand, they return reconstructions that lead to negligible delensing bias, even when the non-Gaussian effects are not modeled explicitly. On the other hand, when delensing with pipelines built around these estimators, the contribution from foreground non-Gaussianity to the power spectrum of delensed B-modes is much smaller than in the HO TT case. As a result, the point-source-hardened estimator, which in general comes close to the HO TT in terms of reconstruction efficiency, can ultimately yield the lowest amount of B-mode power after delensing. This is thanks to being able to extract information from smaller scales with little penalty in the form of additional foreground contributions, but also due to the fact that it can be applied to MV ILC maps without seeing its performance worsened.

Given the nuanced interplay of effects that are taking place, with cancellations that can be highly beneficial, but which depend on the experimental configuration and can be easily spoilt by masking choices and foreground cleaning, it will be important to be able to estimate the biases ahead of any application on real data, such that the most optimal

analysis choices can be made. This is particularly important whenever modeling the bias away is not an option—in such case, the tradeoff between bias and noise of different mitigation methods should be thoroughly explored in a way that is specific to the experiment in question. Alternatively, when taking the modeling route, estimating the bias ahead of time would help determine the analysis strategy that minimizes the total *B*-mode power after delensing, as it is not clear *a priori* at what point the extra delensing efficiency obtained by a more audacious pipeline will be countered by the additional power sourced by foreground non-Gaussianity. This will entail validating pipelines on realistic simulations that accurately reproduce the foreground non-Gaussianity. Alternatively, it would be useful to have flexible analytic tools that can predict the biases as a function of the experimental and analysis parameters—work on this front is in fact ongoing [71].

Either approach will benefit from a better understanding of the extragalactic foregrounds. This is a critical stage of the path ahead: the uncertainty in the foreground modeling is likely to be the largest term in our error budget, but it is unclear how to account for it.[26] The bias amplitudes we report are likely to be underestimated due to halo mass resolution of Websky, which misses contributions to the CIB from halos lighter than 10^{12} $M_\odot$—note, for instance, that the correlation between the 545 GHz CIB map of Websky with CMB lensing, shown in Fig. 13, is a few percent lower than what was seen by [36].

In conclusion, we have shown that the challenge posed by non-Gaussian, extragalactic foregrounds can be overcome through modeling or bias mitigation strategies. Thanks to this, the rich amounts of information encoded in the temperature anisotropies of the CMB can be leveraged to delens *B*-modes and improve constraints on the amplitude of primordial gravitational waves, a key open question in our quest to understand the Universe.

ACKNOWLEDGMENTS

APPENDIX A: *B*-MODE DELENSING WITH A *TT* QE—POSSIBLE BIASES

Whenever lensing reconstructions are contaminated by non-Gaussian, extragalactic foregrounds, the angular spectra of *B*-modes delensed using those reconstructions will receive numerous new contributions. In Eqs. (22) and (23), we highlighted a subset of those term, the ones that we expect to be most significant based on arguments we will outline below. In this appendix, we provide the full list of contributions that are zero or first order in κ.

For the sake of clarity, we take several notational shortcuts. In our presentation, the Gaussian contraction over fields denoted with the superscript "obs" and linked by overbars is to be taken first, followed by the connected *n*-point function of the remaining *n* fields inside the expectation value.

The *n*-point functions can themselves be broken into different couplings. To explore these, we introduce the notation $\tilde{X}[X, \kappa]$, where X is either T or E, to represent the functional dependence of $\tilde{X}$ on the unlensed field X and κ; recall that $\tilde{X}$ is linear in X, so where κ is uncontracted, the unlensed field X is implied. We then use another set of overbars to denote which unlensed fields are coupled together inside the *n*-point function. On the other hand, and in order to highlight their importance in our investigation, we use bars under the expression to identify the foreground and convergence fields that are coupled together in bispectra and trispectra.

We begin by considering the cross-correlation of lensing template and observed *B*-modes. At leading order, the only possible bias term is a function of the fully connected $\langle \tilde{B} \tilde{E} \hat{\kappa}^{TT} \rangle$ trispectrum,

$$\Delta C_l^{\bar{B} \times \hat{B}^{\text{lens}}} \supset g_l \left[\langle \tilde{B}[E, \kappa] \tilde{E}[E, \kappa] \hat{\kappa}^{TT} [s^{\text{NG}}, s^{\text{NG}}] \rangle \right].$$

(A1)

This is analogous to the bispectrum bias studied in the context of CMB lensing reconstructions (e.g., [50,56]), and can be modeled using Eq. (24).

---

[26]This uncertainty is not expected to affect the accuracy of our empirically calibrated prescription for modeling the bias to the delensed *B*-mode spectrum. It could only do so if the simulations grossly misrepresented the relative importance of the different couplings in Appendix A—this seems unlikely.

[27]See https://github.com/simonsobs/pixell.

When it comes to the auto-spectrum of the lensing B-modes template, many more bias couplings are in principle possible. First, there is the terms that feature a Gaussian contraction of the E-modes across templates. These are

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[s^{\text{NG}}, s^{\text{NG}} \right] E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[s^{\text{NG}}, s^{\text{NG}} \right] \rangle \right] , \tag{A2}$$

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 2 h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[\tilde{T}[T,\kappa], \tilde{T}[T,\kappa] \right] E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[s^{\text{NG}}, s^{\text{NG}} \right] \rangle \right] , \tag{A3}$$

and

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 4 h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[\tilde{T}[T,\kappa], s^{\text{NG}} \right] E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[\tilde{T}[T,\kappa], s^{\text{NG}} \right] \rangle \right] . \tag{A4}$$

Based on their coupling structure, we expect these to be the dominant bias terms. The reason is that this is the coupling arrangement in which the weights in the integrands are the least-tightly coupled, leaving them free to explore and accumulate signal over the largest volume of multipole space (see the discussion in Appendix A of [42]). This view is supported by the fact that terms with this general structure are the only connected contributions to the template autospectrum that are included in the standard calculation of the residual lensing B-mode spectrum, Eq. (11) [42].

Terms (A2)–(A4) can be modeled using Eq. (25). The connection with the lensing reconstruction biases is then explicit: Eq. (A2) is related to the trispectrum bias, Eq. (A3) to the primary bispectrum bias, and (A4) to the secondary bispectrum bias.

There are also terms where the disconnected coupling is between the E-modes in one template and an observed T field in the QE in that same template:

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 8 h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[T^{\text{obs}}, s^{\text{NG}} \right] \tilde{E}[E,\kappa] \hat{\kappa}^{\text{TT}} \left[\tilde{T}[T,\kappa], s^{\text{NG}} \right] \rangle \right] , \tag{A5}$$

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 8 h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[T^{\text{obs}}, s^{\text{NG}} \right] \tilde{E}[E,\kappa] \hat{\kappa}^{\text{TT}} \left[\tilde{T}[T,\kappa], s^{\text{NG}} \right] \rangle \right] , \tag{A6}$$

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 4 h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[T^{\text{obs}}, \tilde{T}[T,\kappa] \right] \tilde{E}[E,\kappa] \hat{\kappa}^{\text{TT}} \left[s^{\text{NG}}, s^{\text{NG}} \right] \rangle \right] , \tag{A7}$$

and

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 4 h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TT}} \left[T^{\text{obs}}, \tilde{T}[T,\kappa] \right] \tilde{E}[E,\kappa] \hat{\kappa}^{\text{TT}} \left[s^{\text{NG}}, s^{\text{NG}} \right] \rangle \right] . \tag{A8}$$

Although these terms are more tightly coupled than those in (A2) through (A4), Refs. [41,42] showed that couplings like these constitute the leading, pure-lensing corrections to Eq. (11); they therefore have the potential to be somewhat relevant.

Similarly, there are terms where the E-modes feature in a Gaussian contraction with the T in the QE that appears in the *other* leg of the correlation:

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 8\, h_l \left[\langle E^{\mathrm{obs}} \hat{\kappa}^{\mathrm{TT}} \big[\tilde{T}[T,\kappa], s^{\mathrm{NG}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, s^{\mathrm{NG}} \big] \rangle \right], \tag{A9}$$

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 8\, h_l \left[\langle E^{\mathrm{obs}} \hat{\kappa}^{\mathrm{TT}} \big[\tilde{T}[T,\kappa], s^{\mathrm{NG}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, s^{\mathrm{NG}} \big] \rangle \right], \tag{A10}$$

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 4\, h_l \left[\langle E^{\mathrm{obs}} \hat{\kappa}^{\mathrm{TT}} \big[s^{\mathrm{NG}}, s^{\mathrm{NG}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, \tilde{T}[T,\kappa] \big] \rangle \right] \tag{A11}$$

and

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 4\, h_l \left[\langle E^{\mathrm{obs}} \hat{\kappa}^{\mathrm{TT}} \big[s^{\mathrm{NG}}, s^{\mathrm{NG}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, \tilde{T}[T,\kappa] \big] \rangle \right]. \tag{A12}$$

Finally, the Gaussian contraction can be among T fields, either inside the same QE, as in

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 2\, h_l \left[\langle \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, T^{\mathrm{obs}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[s^{\mathrm{NG}}, s^{\mathrm{NG}} \big] \rangle \right], \tag{A13}$$

and

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 2\, h_l \left[\langle \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, T^{\mathrm{obs}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[s^{\mathrm{NG}}, s^{\mathrm{NG}} \big] \rangle \right]; \tag{A14}$$

or between QEs in different legs of the correlator,

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 8\, h_l \left[\langle \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, s^{\mathrm{NG}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TT}} \big[T^{\mathrm{obs}}, s^{\mathrm{NG}} \big] \rangle \right]. \tag{A15}$$

Terms (A9)–(A15) are all very tightly coupled, so we expect them to make only small contributions to the total bias.

APPENDIX B: *B*-MODE DELENSING WITH A *TE* QE—POSSIBLE BIASES

We can use the notation and arguments described in Appendix A to dissect the possible biases that appear when delensing with a *TE* estimator. The least tightly coupled—and thus most concerning—term is

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 2\, h_l \left[\langle E^{\mathrm{obs}} \hat{\kappa}^{\mathrm{TE}} \big[s^{\mathrm{NG}}, \tilde{E}[E,\kappa] \big] E^{\mathrm{obs}} \hat{\kappa}^{\mathrm{TE}} \big[s^{\mathrm{NG}}, \tilde{E}[E,\kappa] \big] \rangle \right]. \tag{B1}$$

This is essentially a "secondary bispectrum" bias that can be modeled via Eq. (31) once the bias to the $\hat{\kappa}^{TE}$ auto-spectrum is known. Other, likely smaller, contributions are

$$\Delta C_l^{\hat{B}^{\mathrm{lens}} \times \hat{B}^{\mathrm{lens}}} \supset 2\, h_l \left[\langle E^{\mathrm{obs}} \hat{\kappa}^{\mathrm{TE}} \big[s^{\mathrm{NG}}, E^{\mathrm{obs}} \big] \tilde{E}[E,\kappa] \hat{\kappa}^{\mathrm{TE}} \big[s^{\mathrm{NG}}, \tilde{E}[E,\kappa] \big] \rangle \right], \tag{B2}$$

and

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 2\, h_l \left[\langle E^{\text{obs}} \hat{\kappa}^{\text{TE}} \big[s^{\text{NG}}, E^{\text{obs}} \big] \tilde{E}[E, \kappa] \hat{\kappa}^{\text{TE}} \big[s^{\text{NG}}, \tilde{E}[E, \kappa] \big] \rangle \right] , \tag{B3}$$

in which the Gaussian contraction is between E-modes that feature explicitly in the template and E-modes that are input into the TE QE in that same template. In addition to these, there is also

$$\Delta C_l^{\hat{B}^{\text{lens}} \times \hat{B}^{\text{lens}}} \supset 4\, h_l \left[\langle \tilde{E}[E, \kappa] \hat{\kappa}^{\text{TE}} \big[s^{\text{NG}}, E^{\text{obs}} \big] \tilde{E}[E, \kappa] \hat{\kappa}^{\text{TE}} \big[s^{\text{NG}}, E^{\text{obs}} \big] \rangle \right] , \tag{B4}$$

where the Gaussian contraction is between E-modes across the two TE QEs.

[1] U. Seljak and M. Zaldarriaga, Signature of Gravity Waves in Polarization of the Microwave Background, Phys. Rev. Lett. **78,** 2054 (1997).

[2] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Probe of Primordial Gravity Waves and Vorticity, Phys. Rev. Lett. **78,** 2058 (1997).

[3] A. G. Polnarev, Polarization and anisotropy induced in the microwave background by cosmological gravitational waves, Sov. Astron. **29,** 607 (1985).

[4] A. H. Guth, Inflationary universe: A possible solution to the horizon and flatness problems, Phys. Rev. D **23,** 347 (1981).

[5] A. Linde, A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems, Phys. Lett. **108B,** 389 (1982).

[6] A. Albrecht and P. J. Steinhardt, Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking, Phys. Rev. Lett. **48,** 1220 (1982).

[7] A. A. Starobinskii, Spectrum of relict gravitational radiation and the early state of the universe, Sov. J. Exp. Theor. Phys. Lett. **30,** 682 (1979).

[8] J. Khoury, P. J. Steinhardt, and N. Turok, Inflation Versus Cyclic Predictions for Spectral Tilt, Phys. Rev. Lett. **91,** 161301 (2003).

[9] BICEP/Keck Collaboration, Improved Constraints on Primordial Gravitational Waves Using Planck, WMAP, and BICEP/Keck Observations Through the 2018 Observing Season, Phys. Rev. Lett. **127,** 151301 (2021).

[10] CMB-S4 collaboration, CMB-S4 science book, first edition, arXiv:1610.02743.

[11] M. Kamionkowski and E. D. Kovetz, The quest for B modes from inflationary gravitational waves, Annu. Rev. Astron. Astrophys. **54,** 227 (2016).

[12] Planck Collaboration, Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641,** A6 (2020).

[13] A. Lewis and A. Challinor, Weak gravitational lensing of the CMB, Phys. Rep. **429,** 1 (2006).

[14] D. Hanson and SPT Collaboration, Detection of B-Mode Polarization in the Cosmic Microwave Background with Data from the South Pole Telescope, Phys. Rev. Lett. **111,** 141301 (2013).

[15] M. Zaldarriaga and U. Seljak, Gravitational lensing effect on cosmic microwave background polarization, Phys. Rev. D **58,** 023003 (1998).

[16] L. Knox and Y.-S. Song, Limit on the Detectability of the Energy Scale of Inflation, Phys. Rev. Lett. **89,** 011303 (2002).

[17] U. Seljak and C. M. Hirata, Gravitational lensing as a contaminant of the gravity wave signal in CMB, Phys. Rev. D **69,** 043005 (2004).

[18] J. Carron, Optimal constraints on primordial gravitational waves from the lensed CMB, Phys. Rev. D **99,** 043518 (2019).

[19] M. Millea, E. Anderes, and B. D. Wandelt, Sampling-based inference of the primordial CMB and gravitational lensing, Phys. Rev. D **102,** 123542 (2020).

[20] J. Carron, A. Lewis, and A. Challinor, Internal delensing of Planck CMB temperature and polarization, J. Cosmol. Astropart. Phys. 05 (2017) 035.

[21] A. Manzotti and SPTpol Collaboration, CMB polarization B-mode delensing with SPTpol and herschel, Astrophys. J. **846,** 45 (2017).

[22] Planck Collaboration, Planck 2018 results. VIII. Gravitational lensing, Astron. Astrophys. **641,** A8 (2020).

[23] POLARBEAR Collaboration, Internal Delensing of Cosmic Microwave Background Polarization B -Modes with the POLARBEAR Experiment, Phys. Rev. Lett. **124,** 131301 (2020).

[24] D. Han and ACT Collaboration, The atacama cosmology telescope: Delensed power spectra and parameters, J. Cosmol. Astropart. Phys. 01 (2021) 031.

[25] BICEP/Keck Collaboration and SPTpol Collaboration, A demonstration of improved constraints on primordial gravitational waves with delensing, Phys. Rev. D **103,** 022004 (2021).

[26] The Simons Observatory Collaboration, Simons observatory: Constraining inflationary gravitational waves with multitracer B -mode delensing, Phys. Rev. D **105,** 023511 (2022).

[27] D. Adak, A. Sen, S. Basak, J. Delabrouille, T. Ghosh, A. Rotti, G. Martínez-Solaeche, and T. Souradeep, *B*-mode forecast of CMB-Bhārat, Mon. Not. R. Astron. Soc. **514,** 3002 (2022).

[28] P. Diego-Palazuelos, P. Vielva, E. Martínez-González, and R. B. Barreiro, Comparison of delensing methodologies and assessment of the delensing capabilities of future experiments, J. Cosmol. Astropart. Phys. 11 (2020) 058.

[29] CMB-S4 Collaboration, CMB-S4: Forecasting constraints on primordial gravitational waves, Astrophys. J. **926,** 54 (2022).

[30] BICEP/Keck Collaboration, The latest constraints on inflationary B-modes from the BICEP/Keck telescopes, arXiv:2203.16556.

[31] K. M. Smith, D. Hanson, M. LoVerde, C. M. Hirata, and O. Zahn, Delensing CMB polarization with external datasets, J. Cosmol. Astropart. Phys. 06 (2012) 014.

[32] A. Baleato Lizancos, A. Challinor, and J. Carron, Limitations of CMB B -mode template delensing, Phys. Rev. D **103,** 023518 (2021).

[33] W. Hu and T. Okamoto, Mass reconstruction with cosmic microwave background polarization, Astrophys. J. **574,** 566 (2002).

[34] C. M. Hirata and U. Seljak, Reconstruction of lensing from the cosmic microwave background polarization, Phys. Rev. D **68,** 083002 (2003).

[35] J. Carron and A. Lewis, Maximum a posteriori CMB lensing reconstruction, Phys. Rev. D **96,** 063510 (2017).

[36] B. D. Sherwin and M. Schmittfull, Delensing the CMB with the cosmic infrared background, Phys. Rev. D **92,** 043005 (2015).

[37] B. Yu, J. C. Hill, and B. D. Sherwin, Multitracer CMB delensing maps from Planck and WISE data, Phys. Rev. D **96,** 123511 (2017).

[38] Simons Observatory Collaboration, The Simons Observatory: Science goals and forecasts, J. Cosmol. Astropart. Phys. 02 (2019) 056.

[39] W. L. K. Wu and SPTPol Collaboration, A measurement of the cosmic microwave background lensing potential and power spectrum from 500 deg^2 of SPTpol temperature and polarization data, Astrophys. J. **884,** 70 (2019).

[40] W.-H. Teng, C.-L. Kuo, and J.-H. Proty Wu, Cosmic microwave background delensing revisited: Residual biases and a simple fix, arXiv:1102.5729.

[41] T. Namikawa, CMB internal delensing with general optimal estimator for higher-order correlations, Phys. Rev. D **95,** 103514 (2017).

[42] A. Baleato Lizancos, A. Challinor, and J. Carron, Impact of internal-delensing biases on searches for primordial B-modes of CMB polarisation, J. Cosmol. Astropart. Phys. 03 (2021) 016.

[43] R. A. Sunyaev and Y. B. Zeldovich, The observations of relic radiation as a test of the nature of x-ray radiation from the clusters of galaxies, Comments Astrophys. Space Phys. **4,** 173 (1972).

[44] R. A. Sunyaev and Y. B. Zeldovich, The velocity of clusters of galaxies relative to the microwave background—The possibility of its measurement, Mon. Not. R. Astron. Soc. **190,** 413 (1980).

[45] M. G. Hauser and E. Dwek, The cosmic infrared background: Measurements and implications, Annu. Rev. Astron. Astrophys. **39,** 249 (2001).

[46] A. Amblard, C. Vale, and M. White, Weak lensing of the CMB by large-scale structure, Nature (London) **9,** 687 (2004).

[47] K. M. Smith, O. Zahn, and O. Doré, Detection of gravitational lensing in the cosmic microwave background, Phys. Rev. D **76,** 043510 (2007).

[48] L. E. Bleem and SPT Collaboration, A measurement of the correlation of galaxy surveys with CMB lensing convergence maps from the south pole telescope, Astrophys. J. Lett. **753,** L9 (2012).

[49] A. van Engelen *et al.*, A measurement of gravitational lensing of the microwave background using South Pole Telescope data, Astrophys. J. **756,** 142 (2012).

[50] A. van Engelen, S. Bhattacharya, N. Sehgal, G. P. Holder, O. Zahn, and D. Nagai, CMB lensing power spectrum biases from galaxies and clusters using high-angular resolution temperature maps, Astrophys. J. **786,** 13 (2014).

[51] S. J. Osborne, D. Hanson, and O. Doré, Extragalactic foreground contamination in temperature-based CMB lens reconstruction, J. Cosmol. Astropart. Phys. 03 (2014) 024.

[52] N. Sailer, E. Schaan, and S. Ferraro, Lower bias, lower noise CMB lensing with foreground-hardened estimators, Phys. Rev. D **102,** 063517 (2020).

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

[54] S. Ferraro and J. C. Hill, Bias to CMB lensing reconstruction from temperature anisotropies due to large-scale galaxy motions, Phys. Rev. D **97,** 023512 (2018).

[55] E. Schaan and S. Ferraro, Foreground-Immune Cosmic Microwave Background Lensing with Shear-Only Reconstruction, Phys. Rev. Lett. **122,** 181301 (2019).

[56] N. Sailer, E. Schaan, S. Ferraro, O. Darwish, and B. Sherwin, Optimal multifrequency weighting for CMB lensing, Phys. Rev. D **104,** 123514 (2021).

[57] O. Darwish, B. D. Sherwin, N. Sailer, E. Schaan, and S. Ferraro, Optimizing foreground mitigation for CMB lensing with combined multifrequency and geometric methods, arXiv:2111.00462.

[58] D. Han and N. Sehgal, Mitigating foreground bias to the CMB lensing power spectrum for a CMB-HD survey, Phys. Rev. D **105,** 083516 (2022).

[59] A. Challinor and A. Lewis, Lensed CMB power spectra from all-sky correlation functions, Phys. Rev. D **71,** 103010 (2005).

[60] G. Pratten and A. Lewis, Impact of post-born lensing on the CMB, J. Cosmol. Astropart. Phys. 08 (2016) 047.

[61] N. Krachmalnicoff, C. Baccigalupi, J. Aumont, M. Bersanelli, and A. Mennella, Characterization of foreground emission on degree angular scales for CMB B-mode observations. Thermal dust and synchrotron signal from Planck and WMAP data, Astron. Astrophys. **588,** A65 (2016).

[62] BICEP2/Keck Array Collaboration and Planck Collaboration, Joint Analysis of BICEP2/Keck Array and Planck Data, Phys. Rev. Lett. **114,** 101301 (2015).

[63] G. Lagache, M. Béthermin, L. Montier, P. Serra, and M. Tucci, Impact of polarised extragalactic sources on the measurement of CMB B-mode anisotropies, Astron. Astrophys. **642,** A232 (2020).

[64] D. Beck, G. Fabbian, and J. Errard, Lensing reconstruction in post-Born cosmic microwave background weak lensing, Phys. Rev. D **98,** 043512 (2018).

[65] A. Baleato Lizancos, A. Challinor, B. D. Sherwin, and T. Namikawa, Delensing the CMB with the cosmic infrared background: The impact of foregrounds, Mon. Not. R. Astron. Soc. **514,** 5786 (2022).

[66] K. M. Smith *et al.*, Gravitational lensing, arXiv:0811.3916.

[67] A. Lewis, A. Challinor, and D. Hanson, The shape of the CMB lensing bispectrum, J. Cosmol. Astropart. Phys. 03 (2011) 018.

[68] D. Hanson, A. Challinor, G. Efstathiou, and P. Bielewicz, CMB temperature lensing power reconstruction, Phys. Rev. D **83,** 043005 (2011).

[69] A. van Engelen and ACT Collaboration, The atacama cosmology telescope: Lensing of CMB temperature and polarization derived from cosmic infrared background cross-correlation, Astrophys. J. **808,** 7 (2015).

[70] S. Dodelson, *Gravitational Lensing* (Cambridge University Press, Cambridge, England, 2017).

[71] A. Baleato Lizancos, W. Coulton, A. Challinor, and B. Sherwin, CosmoBLENDER: Fast modelling of biases to CMB lensing power spectra and cross-correlations sourced by galaxies and clusters (unpublished).

[72] G. Stein, M. A. Alvarez, J. R. Bond, A. van Engelen, and N. Battaglia, The websky extragalactic CMB simulations, J. Cosmol. Astropart. Phys. 10 (2020) 012.

[73] G. Stein, M. A. Alvarez, and J. R. Bond, The mass-Peak Patch algorithm for fast generation of deep all-sky dark matter halo catalogues and its N-body validation, Mon. Not. R. Astron. Soc. **483,** 2236 (2019).

[74] N. Battaglia, J. R. Bond, C. Pfrommer, and J. L. Sievers, On the cluster physics of Sunyaev-Zel'dovich and x-ray surveys. II. Deconstructing the thermal SZ power spectrum, Astrophys. J. **758,** 75 (2012).

[75] C. Shang, Z. Haiman, L. Knox, and S. P. Oh, Improved models for cosmic infrared background anisotropies: New constraints on the infrared galaxy population, Mon. Not. R. Astron. Soc. **421,** 2832 (2012).

[76] M. P. Viero and HerMES Collaboration, HerMES: Cosmic infrared background anisotropies and the clustering of dusty star-forming galaxies, Astrophys. J. **772,** 77 (2013).

[77] Planck Collaboration, Planck 2015 results. XXIII. The thermal Sunyaev-Zeldovich effect-cosmic infrared background correlation, Astron. Astrophys. **594,** A23 (2016).

[78] Z. Li, G. Puglisi, M. S. Madhavacheril, and M. A. Alvarez, Simulated catalogs and maps of radio galaxies at millimeter wavelengths in Websky, J. Cosmol. Astropart. Phys. 08 (2022) 029.

[79] R. K. Sachs and A. M. Wolfe, Perturbations of a cosmological model and angular variations of the microwave background, Astrophys. J. **147,** 73 (1967).

[80] M. J. Rees and D. W. Sciama, Large-scale density inhomogeneities in the universe, Nature (London) **217,** 511 (1968).

[81] S. Ferraro, E. Schaan, and E. Pierpaoli, Is the Rees-Sciama effect detectable by the next generation of cosmological experiments?, arXiv:2205.10332.

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

[83] J. Dunkley and ACT Collaboration, The atacama cosmology telescope: likelihood for small-scale CMB data, J. Cosmol. Astropart. Phys. 07 (2013) 025.

[84] M. G. Haehnelt and M. Tegmark, Using the Kinematic Sunyaev-Zeldovich effect to determine the peculiar velocities of clusters of galaxies, Mon. Not. R. Astron. Soc. **279,** 545 (1996).

[85] G. Fabbian, J. Carron, A. Lewis, and M. Lembo, Lensed CMB power spectrum biases from masking extragalactic sources, Phys. Rev. D **103,** 043535 (2021).

[86] A. Benoit-Lévy, T. Déchelette, K. Benabed, J. F. Cardoso, D. Hanson, and S. Prunet, Full-sky CMB lensing reconstruction in presence of sky-cuts, Astron. Astrophys. **555,** A37 (2013).

[87] A. Lewis, A. Challinor, and N. Turok, Analysis of CMB polarization on an incomplete sky, Phys. Rev. D **65,** 023505 (2001).

[88] D. Alonso, J. Sanchez, A. Slosar, and LSST Dark Energy Science Collaboration, A unified pseudo-C_ℓ framework, Mon. Not. R. Astron. Soc. **484,** 4127 (2019).

[89] A. Roy, G. Kulkarni, P. D. Meerburg, A. Challinor, C. Baccigalupi, A. Lapi, and M. G. Haehnelt, Revised estimates of CMB B-mode polarization induced by patchy reionization, J. Cosmol. Astropart. Phys. 01 (2021) 003.

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

[91] Y. Sultan Abylkairov, O. Darwish, J. C. Hill, and B. D. Sherwin, Partially constrained internal linear combination: A method for low-noise CMB foreground mitigation, Phys. Rev. D **103,** 103510 (2021).

[92] A. S. Maniyar, Y. Ali-Haïmoud, J. Carron, A. Lewis, and M. S. Madhavacheril, Quadratic estimators for CMB weak lensing, Phys. Rev. D **103,** 083524 (2021).

[93] K. S. Karkare, Delensing degree-scale B-mode polarization with high-redshift line intensity mapping, Phys. Rev. D **100,** 043529 (2019).

[94] DES Collaboration, Dark energy survey year 3 results: Galaxy clustering and systematics treatment for lens galaxy samples, Mon. Not. R. Astron. Soc. **511,** 2665 (2022).

[95] Planck Collaboration, Planck intermediate results. XLVIII. Disentangling Galactic dust emission and cosmic infrared background anisotropies, Astron. Astrophys. **596,** A109 (2016).

[96] Planck Collaboration, Planck 2018 results. I. Overview and the cosmological legacy of Planck, Astron. Astrophys. **641,** A1 (2020).

[97] C. R. Harris *et al.*, Array programming with NumPy, Nature (London) **585,** 357 (2020).

[98] P. Virtanen and SciPy 1.0 Contributors, SciPy 1.0: Fundamental algorithms for scientific computing in python, Nat. Methods **17,** 261 (2020).

[99] J. D. Hunter, Matplotlib: A 2d graphics environment, Comput. Sci. Eng. **9,** 90 (2007).

[100] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, HEALPix—A framework for high resolution discretization, and fast analysis of data distributed on the sphere, Astrophys. J. **622,** 759 (2005).

[101] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. Gorski, healpy: Equal area pixelization and spherical harmonics transforms for data on the sphere in python, J. Open Source Softwaare **4,** 1298 (2019).