Foreground-immune CMB lensing reconstruction with polarization

Noah Sailer, Simone Ferraro, and Emmanuel Schaan

Foreground-immune CMB lensing reconstruction with polarization

Noah Sailer,[1, 2, *] Simone Ferraro,[2, 1, †] and Emmanuel Schaan[3, 4, ‡]

[1]*Berkeley Center for Cosmological Physics, Department of Physics,
University of California, Berkeley, CA 94720, USA*
[2]*Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA*
[3]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[4]*Kavli Institute for Particle Astrophysics and Cosmology and
Department of Physics, Stanford University, Stanford, CA 94305, USA*

Extragalactic foregrounds are known to generate significant biases in temperature-based CMB lensing reconstruction. Several techniques, which include "source hardening" and "shear-only estimators" have been proposed to mitigate contamination and have been shown to be very effective at reducing foreground-induced biases. Here we extend both techniques to polarization, which will be an essential component of CMB lensing reconstruction for future experiments, and investigate the "large-lens" limit analytically to gain insight on the origin and scaling of foreground biases, as well as the sensitivity to their profiles. Using simulations of polarized point sources, we estimate the expected bias to both Simons Observatory and CMB-S4 like (polarization-based) lensing reconstruction, finding that biases to the former are minuscule while those to the latter are potentially non-negligible at small scales ($L \sim 1000 - 2000$). In particular, we show that for a CMB-S4 like experiment, an optimal linear combination of point-source hardened estimators can reduce the (point-source induced) bias to the CMB lensing power spectrum by up to two orders of magnitude, at a $\sim 4\%$ noise cost relative to the global minimum variance estimator.

I. INTRODUCTION

Point sources and other foregrounds induce biases [1, 2] to standard quadratic estimators of the weak lensing potential due to the fact that foregrounds are non-Gaussian and correlated with the lensing potential. The impact of foregrounds on the CMB temperature lensing reconstruction has been explored extensively in the literature and a number of mitigation techniques have been proposed. These range from decreasing the multipole range used in the reconstruction to more sophisticated "geometric methods" that employ the different symmetries of foregrounds and lensing.

Hardening against point sources in temperature-based reconstruction has been proposed in [3–6] and has been shown to reduce the foreground biases by an order of magnitude with very modest noise cost of order 10%. A generalization to hardening against a fixed profile (profile hardening) was proposed in [6] and has been shown to be even more effective when the mean profile of tSZ halos is hardened against. An even more general "foreground-immune" estimator is the shear-only estimator [7], which suppresses the response to arbitrary azimuthally symmetric profiles. Further generalizations to "multipole" estimators have also been proposed in [7].

Moreover, frequency-based methods to reduce biases have been proposed in [8, 9], and the optimal combination of multi-frequency and geometric methods have been explored in [10]. Extragalactic foregrounds can also affect delensing of the $B-$modes produced by primordial gravitational waves, and the techniques mentioned above appear to be very effective at reducing these biases [11].

In this paper we extend two of the most promising geometric techniques used in temperature reconstruction (i.e. bias hardening and shear-only reconstruction) to polarization. We note that these can be used in addition to multi-frequency foreground reduction to further mitigate foreground biases.

CMB polarization dominates the statistical power of lensing reconstruction for map noise $\Delta_T \lesssim 5\,\mu\text{K-arcmin}$, and therefore will be the principal channel in the future. For current experiments, polarization already dominates reconstruction for SPT-3G [12] and will play an important role in Simons Observatory [13] and CMB-S4 [14]. It is therefore important to characterize and mitigate possible biases induced by foregrounds in polarization.

First of all, we note that extragalactic foregrounds in polarization are thought to be simpler than in temperature. The largest sources of extragalactic foregrounds, such as emission from the Cosmic Infrared Background (CIB) or the thermal Sunyaev-Zel'dovich effect (tSZ), are mostly unpolarized (although a small polarized signal is present at higher order [15, 16]), and therefore their effect on lensing reconstruction from polarization is likely negligible. However, bright infrared and radio point sources are known to be polarized at the $\sim$ few % level [17–19], inducing potential biases. Therefore we'll concentrate on mitigating the effect of polarized point sources, while noting that the techniques we present here have also been shown to be very effective against CIB and tSZ in temperature, and therefore are likely to also work more generally in polarization.

In this paper we focus on extragalactic foregrounds. The methods developed to mitigate these biases may also prove useful for galactic foregrounds, whose biases are potentially important and their study will be the subject of future work.

The remainder of the paper is organized as follows: in section II, we outline the theory of quadratic estimators,

* nsailer@berkeley.edu
† sferraro@lbl.gov
‡ eschaan@slac.stanford.edu

including multi-frequency and minimum-variance combinations. In section III, we extend the bias hardening techniques to polarization, while in section IV we explore generalizations to the "shear-only" estimators, highlighting important differences in polarization. In section V we study the noise properties of these new estimators and in section VI we investigate the bias reduction obtained. We conclude in section VII.

II. LINEAR RESPONSES AND QUADRATIC ESTIMATORS

Here we review the basic theory of quadratic estimators in temperature and polarization. We separately consider the cases of reconstructing the convergence from a single pair of maps and the case of summing over all pairs, and derive the minimum variance weights for each case. For simplicity we work under the flat sky approximation throughout.

A. Lensed fields and linear responses

The lensed CMB temperature $\tilde{T}_\ell$ is related to the unlensed field T_ℓ *to linear order* in the lensing convergence κ as follows [20]:

$$\tilde{T}_\ell = T_\ell - 2 \int_{\ell'} \frac{\ell' \cdot (\ell - \ell')}{(\ell - \ell')^2} \kappa_{\ell - \ell'} T_{\ell'}, \tag{1}$$

where $\int_{\ell_1 \cdots \ell_n} \equiv \int d^2\ell_1 \cdots d^2\ell_n/(2\pi)^{2n}$. Throughout we use tildes to distinguish lensed fields from their unlensed counterparts. The analogous expression for the lensed polarization fields is:

$$\tilde{E}_\ell \pm i\tilde{B}_\ell = E_\ell - 2 \int_{\ell'} \frac{\ell' \cdot (\ell - \ell')}{(\ell - \ell')^2} e^{\pm 2i(\phi_{\ell'} - \phi_\ell)} \kappa_{\ell - \ell'} E_{\ell'}, \tag{2}$$

where the angle ϕ_ℓ is defined through $\ell = \ell \cos(\phi_\ell)\hat{x}_1 + \ell \sin(\phi_\ell)\hat{x}_2$, $\{\hat{x}_1, \hat{x}_2\}$ are orthonormal basis vectors, and we've assumed that there are no primordial $B-$modes.

The lensed CMB temperature and polarization fields are statistically isotropic, and as a result $\langle \tilde{X}_\ell \tilde{Y}_{L-\ell} \rangle = 0$ for $L > 0$, where $\langle \cdots \rangle$ denotes an average over realizations of the primary CMB *and* the lensing field, and $X, Y \in \{T, E, B\}$. However, fixing the Fourier mode κ_L breaks statistical isotropy, resulting in off diagonal covariances of the form:

$$\langle \tilde{X}_\ell \tilde{Y}_{L-\ell} \rangle' = f_{\ell, L-\ell}^{XY} \kappa_L + \mathcal{O}(\kappa^2) \quad \text{for} \quad L > 0, \tag{3}$$

where $f_{\ell, L-\ell}^{XY}$ is the linear response of lensed $\tilde{X}, \tilde{Y}$ fields to the lensing convergence, and $\langle \cdots \rangle'$ denotes an average over realizations of the primary CMB and all Fourier modes of the lensing field *except*[1] for L. Note that the responses trivially satisfy $f_{\ell, \ell'}^{XY} = f_{\ell', \ell}^{YX}$.

From Eqs. (1) and (2) one obtains[2]:

$$f_{\ell, L-\ell}^{TT} = \frac{2L}{L^2} \cdot \left[\ell C_\ell^{TT} + (L - \ell) C_{|L-\ell|}^{TT} \right]$$

$$f_{\ell, L-\ell}^{TE} = \frac{2L}{L^2} \cdot \left[\ell C_\ell^{TE} \cos(2\theta_{\ell, L-\ell}) + (L - \ell) C_{|L-\ell|}^{TE} \right]$$

$$f_{\ell, L-\ell}^{TB} = \frac{2L}{L^2} \cdot \ell C_\ell^{TE} \sin(2\theta_{\ell, L-\ell})$$

$$f_{\ell, L-\ell}^{EE} = \frac{2L}{L^2} \cdot \left[\ell C_\ell^{EE} + (L - \ell) C_{|L-\ell|}^{EE} \right] \cos(2\theta_{\ell, L-\ell})$$

$$f_{\ell, L-\ell}^{EB} = \frac{2L}{L^2} \cdot \ell C_\ell^{EE} \sin(2\theta_{\ell, L-\ell})$$

$$f_{\ell, L-\ell}^{BB} = 0, \tag{4}$$

where C_ℓ^{TT}, C_ℓ^{TE}, and C_ℓ^{EE} are the unlensed power spectra, and $\theta_{\ell, \ell'} \equiv \phi_\ell - \phi_{\ell'}$ is the angle between ℓ and ℓ'.

B. Minimum variance estimators

Eq. (3) is the starting point for deriving a quadratic estimator (QE) for the lensing convergence. By dividing both sides of Eq. (3) by the linear response, one can build a simple unbiased QE $\hat{\kappa}_{\ell, L}^{XY}$ for each ℓ: $\hat{\kappa}_{\ell, L}^{XY} = \tilde{X}_\ell \tilde{Y}_{L-\ell}/f_{\ell, L-\ell}^{XY}$, such that $\langle \hat{\kappa}_{\ell, L}^{XY} \rangle' = \kappa_L$ to linear order. The most general QE is obtained by taking an arbitrary weighted average of the $\hat{\kappa}_{\ell, L}^{XY}$'s, provided that the weights sum to unity (to keep the estimator unbiased). Below we consider separately the specialized case where only a single pair of maps are used for the lensing reconstruction, and the more general scenario of summing over all available pairs. While the former is slightly suboptimal in terms of noise, it is still practically useful to have a collection of individual QEs for performing cross-checks and diagnosing systematics.

1. Single pair of maps

Given a single pair of maps $\tilde{X}, \tilde{Y}$ one can construct a generic unbiased QE for the lensing convergence:

$$\hat{\kappa}_L^{XY} = N_L^{XY} \int_\ell F_{\ell, L-\ell}^{XY} \tilde{X}_\ell \tilde{Y}_{L-\ell}$$

$$\text{where} \quad (N_L^{XY})^{-1} = \int_\ell F_{\ell, L-\ell}^{XY} f_{\ell, L-\ell}^{XY}, \tag{5}$$

where $F_{\ell, L-\ell}^{XY}$ are some weights, which are in principle arbitrary. The Gaussian component of the variance of

---

[1] We could have also defined $\langle \cdots \rangle'$ as an average over the primary

CMB with the entire lensing field fixed (not just the L'th Fourier mode) and Eq. (3) would still be valid. In III our definition of only fixing the L'th Fourier mode will be relevant for constructing the foreground QE, and we choose to adopt a common notation throughout for convenience.

[2] Note that $\phi_{-\ell} \cong \phi_\ell + n\pi$ for any odd integer n. Since all angles in Eq. (4) are multiplied by 2, we are always free to change the sign of ℓ in ϕ_ℓ, which we have done to make the responses look as symmetric as possible.

$\hat{\kappa}^{XY}$, defined such that $\langle \hat{\kappa}_{\boldsymbol{L}}^{XY} \hat{\kappa}_{\boldsymbol{L'}}^{XY} \rangle_c = (2\pi)^2 \delta_{\boldsymbol{L}+\boldsymbol{L'}}^D \mathcal{N}_{\boldsymbol{L}}^{XY}$, is given by

$$\mathcal{N}_{\boldsymbol{L}}^{XY} = (N_{\boldsymbol{L}}^{XY})^2 \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} \left[F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} \tilde{C}_{\ell}^{XX} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{YY} \right. $$
$$\left. + F_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}}^{XY} \tilde{C}_{\ell}^{XY} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{XY} \right], \quad (6)$$

where $\tilde{C}_{\ell}^{XY}$ is the *total observed* power spectrum, i.e. the power spectra of the *lensed* fields, including instrument noise and foregrounds. In the equation above we have made the simplifying assumption that the weights are real, so that $F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} = F_{-\boldsymbol{\ell},\boldsymbol{\ell}-\boldsymbol{L}}^{XY}$. From here on out we will always assume that the weights and linear responses are real.

A common and optimistic choice for the weights is to minimize $\mathcal{N}_{\boldsymbol{L}}^{XY}$. This can be accomplished using e.g. functional derivatives, giving [21]:

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} = \frac{\tilde{C}_{\ell}^{YY} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{XX} f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} - \tilde{C}_{\ell}^{XY} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{XY} f_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}}^{XY}}{\tilde{C}_{\ell}^{XX} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{YY} \tilde{C}_{\ell}^{YY} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{XX} - \left(\tilde{C}_{\ell}^{XY} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{XY} \right)^2}. \quad (7)$$

Note that in the limit where $\tilde{X}$ and $\tilde{Y}$ are uncorrelated (or when $\tilde{X} = \tilde{Y}$) the minimum variance weights reduce to $F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} \propto f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} / \tilde{C}_{\ell}^{XX} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{YY}$ up to some generic prefactor (any function of $\boldsymbol{L}$) which the QE (Eq. (5)) is insensitive to. Neglecting the cross-correlation between e.g. T and E is often a very good approximation, and doing so makes Eq. (5) amenable to FFTs. Of course, this approximation to the weights still leaves the estimator unbiased, although slightly suboptimal. In what follows we always make the approximation of neglecting the cross-correlation in the minimum variance weights for a single pair of maps.

2. Summing over all pairs

Given a collection of lensed maps $\{\tilde{M}^i\}$ one can build the general QE:

$$\hat{\kappa}_{\boldsymbol{L}} = N_{\boldsymbol{L}}^{\kappa} \sum_{ij} \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} \tilde{M}_{\boldsymbol{\ell}}^i \tilde{M}_{\boldsymbol{L}-\boldsymbol{\ell}}^j$$
$$\text{where} \quad (N_{\boldsymbol{L}}^{\kappa})^{-1} = \sum_{ij} \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij}, \quad (8)$$

where the linear response $f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij}$ to the lensing convergence is defined in an analogous way as Eq. (3). Note that in Eq. (8) we are summing over all pairs (e.g. both ij and ji) and that the weights satisfy $F_{\boldsymbol{\ell},\boldsymbol{\ell'}}^{ij} = F_{\boldsymbol{\ell'},\boldsymbol{\ell}}^{ji}$. The index i is quite general. For example, i could simply index T, E and B; or $\tilde{M}^i$ could be a temperature map with i indexing the frequency; or even more generally i could index a pair of indices (X, ν) where $X \in \{T, E, B\}$ and ν is a frequency.

The variance of Eq. (8) is:

$$\mathcal{N}_{\boldsymbol{L}}^{\kappa} = 2(N_{\boldsymbol{L}}^{\kappa})^2 \sum_{ijmn} \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{mn} \tilde{C}_{\ell}^{im} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{jn}, \quad (9)$$

where the covariance matrix $\tilde{C}_{\ell}^{ij}$ is defined as $\langle \tilde{M}_{\boldsymbol{\ell}}^i \tilde{M}_{\boldsymbol{L}-\boldsymbol{\ell}}^j \rangle = (2\pi)^2 \delta_{\boldsymbol{L}}^D \tilde{C}_{\ell}^{ij}$.

The weights which minimize $\mathcal{N}_{\boldsymbol{L}}^{\kappa}$ are given by:

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} = \frac{1}{2} \sum_{mn} \left(\tilde{\boldsymbol{C}}_{\ell}^{-1} \right)^{im} \left(\tilde{\boldsymbol{C}}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{-1} \right)^{jn} f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{mn}, \quad (10)$$

where we have chosen the normalization of the weights so that $\mathcal{N}_{\boldsymbol{L}}^{\kappa} = N_{\boldsymbol{L}}^{\kappa}$. Note that unlike for the single pair case, here the true minimum variance QE is always FFT-able. We provide explicit expressions for these weights in Appendix B.

Note that in the case where $\tilde{M}^i$ is a temperature map with i indexing a frequency, Eq. (10) reduces to the familiar form: $w_{\ell}^i w_{|\boldsymbol{L}-\boldsymbol{\ell}|}^j f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TT} / 2 \tilde{C}_{\ell}^{\text{ILC}} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{\text{ILC}}$, where $\tilde{C}_{\ell}^{\text{ILC}}$ is the power spectrum of the minimum variance internal linear combination (MVILC) of temperature maps, and w_{ℓ}^i are the MVILC weights. In the case where i indexes $T, E,$ and B (at a single frequency) Eq. (10) recovers the global minimum variance estimator (GMV), originally derived in [22]. As was pointed out in [23], the GMV has a slightly lower noise than the minimum variance linear combination of the MV single-pair estimators $\hat{\kappa}_{\boldsymbol{L}}^{XY}$.

III. BIAS HARDENED ESTIMATORS

The most general QE takes the form of Eq. (8), which is unbiased when the maps $\tilde{M}^i$ are composed of a lensed CMB signal and Gaussian noise, so that $\langle \tilde{M}_{\boldsymbol{\ell}}^i \tilde{M}_{\boldsymbol{L}-\boldsymbol{\ell}}^j \rangle' = f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} \kappa_{\boldsymbol{L}}$. If instead the maps are contaminated by a foreground s^i, the covariance $\langle \tilde{M}_{\boldsymbol{\ell}}^i \tilde{M}_{\boldsymbol{L}-\boldsymbol{\ell}}^j \rangle'$ picks up additional off-diagonal contributions, which to lowest order in κ and s^i take the form[3] [4, 6]:

$$\langle \tilde{M}_{\boldsymbol{\ell}}^i \tilde{M}_{\boldsymbol{L}-\boldsymbol{\ell}}^j \rangle' = f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} \kappa_{\boldsymbol{L}} + \sum_k g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij,k} s_{\boldsymbol{L}}^k, \quad (11)$$

where as before f^{ij} is the linear response to lensing in the absense of foregrounds, while $g^{ij,k}$ is the linear response to the foreground in the absence of lensing[4]: $\langle s_{\boldsymbol{\ell}}^i s_{\boldsymbol{L}-\boldsymbol{\ell}}^j \rangle' = \sum_k g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij,k} s_{\boldsymbol{L}}^k$. The additional term in Eq. (11) induces an additive bias to Eq. (8) which is proportional to the foreground:

$$\langle \hat{\kappa}_{\boldsymbol{L}} \rangle' = \kappa_{\boldsymbol{L}} + \sum_k \underbrace{\left(N_{\boldsymbol{L}}^{\kappa} \sum_{ij} \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij,k} \right)}_{\equiv \mathcal{R}_{\boldsymbol{L}}^{\kappa,k}} s_{\boldsymbol{L}}^k, \quad (12)$$

where we have defined the response $\mathcal{R}_{\boldsymbol{L}}^{\kappa,i}$ of the lensing estimator to the foreground in the i'th map.

---

[3] Note that in the average $\langle \cdots \rangle'$ we are fixing the $\boldsymbol{L}$'th Fourier mode of both the lensing convergence and the foreground.

[4] Note that one could equivalently define $g^{ij,k}$ as the linear response of $\langle \tilde{M}^i \tilde{M}^j \rangle'$ to s^k in the absence of lensing.

The intuition behind bias hardening goes as follows: just as we can construct a convergence QE $\hat{\kappa}$ we can also construct a foreground QE $\hat{s}^i$ which can in turn be used to subtract off the bias to Eq. (12). The only ingredients that we need are the foreground linear response and a set of weights $G^{ij,k}$:

$$\hat{s}_{\boldsymbol{L}}^i = N_{\boldsymbol{L}}^i \sum_{jk} \int_{\boldsymbol{\ell}} G_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk,i} \tilde{M}_{\boldsymbol{\ell}}^j \tilde{M}_{\boldsymbol{L}-\boldsymbol{\ell}}^k$$

$$\text{where} \quad (N_{\boldsymbol{L}}^i)^{-1} = \sum_{jk} \int_{\boldsymbol{\ell}} G_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk,i} g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk,i}. \tag{13}$$

When evaluated on the contaminated maps, $\hat{s}^i$ picks up field level biases proportional to the foreground in the other maps ($m \neq i$), and to κ:

$$\langle \hat{s}_{\boldsymbol{L}}^i \rangle' = \sum_m \left(N_{\boldsymbol{L}}^i \sum_{jk} \int_{\boldsymbol{\ell}} G_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk,i} g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk,m} \right) s_{\boldsymbol{L}}^m$$

$$+ \left(N_{\boldsymbol{L}}^i \sum_{jk} \int_{\boldsymbol{\ell}} G_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk,i} f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk} \right) \kappa_{\boldsymbol{L}} \tag{14}$$

$$\equiv \sum_m \mathcal{R}_{\boldsymbol{L}}^{i,m} s_{\boldsymbol{L}}^m + \mathcal{R}_{\boldsymbol{L}}^{i,\kappa} \kappa_{\boldsymbol{L}},$$

where $\mathcal{R}_{\boldsymbol{L}}^{i,i} = 1$. By recasting Eqs. (12) and (14) in a matrix form one can construct a suitable linear combination of $\hat{\kappa}$ and $\hat{s}^i$ to null the linear order field level biases:

$$\begin{pmatrix} \hat{\kappa}_{\boldsymbol{L}}^{\text{BH}} \\ \hat{s}_{\boldsymbol{L}}^{\text{BH},1} \\ \vdots \\ \hat{s}_{\boldsymbol{L}}^{\text{BH},n} \end{pmatrix} = \begin{pmatrix} 1 & \mathcal{R}_{\boldsymbol{L}}^{\kappa,1} & \cdots & \mathcal{R}_{\boldsymbol{L}}^{\kappa,n} \\ \mathcal{R}_{\boldsymbol{L}}^{1,\kappa} & 1 & \cdots & \mathcal{R}_{\boldsymbol{L}}^{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{R}_{\boldsymbol{L}}^{n,\kappa} & \mathcal{R}_{\boldsymbol{L}}^{n,1} & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\kappa}_{\boldsymbol{L}} \\ \hat{s}_{\boldsymbol{L}}^1 \\ \vdots \\ \hat{s}_{\boldsymbol{L}}^n \end{pmatrix}. \tag{15}$$

By taking an average of Eq. (15) and plugging in Eqs. (12) and (14), we see that the bias hardened estimators satisfy: $\langle \hat{\kappa}_{\boldsymbol{L}}^{\text{BH}} \rangle' = \kappa_{\boldsymbol{L}}$ and $\langle \hat{s}_{\boldsymbol{L}}^{\text{BH},i} \rangle' = s_{\boldsymbol{L}}^i$.

In practice we choose the foreground weights G and lensing weights F to minimize the variance of $\hat{s}^i$ and $\hat{\kappa}$ respectively. In Appendix D we show that this choice of weights is equivalent to the minimum variance lensing estimator constrained to have zero response to foregrounds (i.e. the optimal bias hardened estimator).

For the case of using a single pair of maps for reconstruction we therefore take $G_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij,k} = g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij,k}/\tilde{C}_{\boldsymbol{\ell}}^{ii}\tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{jj}$ and $F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij} = f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij}/\tilde{C}_{\boldsymbol{\ell}}^{ii}\tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{jj}$, where we have neglected the cross-correlation $\tilde{C}^{ij}$ to make the estimators FFT-able. As an illustrative example we explicitly bias harden the TE estimator in Appendix F using the foreground linear responses derived in section III B.

A. Foreground model

The only input required to construct the foreground QE is the linear response. Below we outline our model for the foregrounds. We use this model to compute the linear response in the following subsection.

We assume that the foreground s is a sum of N_{sources} sources with identical profiles, and further assume that polarization fractions and polarization angles do not vary significantly over a single source, so that the intensity and polarization of the foreground take the form:

$$s^I(\boldsymbol{x}) = \sum_i s_i^I u(\boldsymbol{x} - \boldsymbol{x}_i)$$

$$s^Q(\boldsymbol{x}) \pm i s^U(\boldsymbol{x}) = \sum_i s_i^I p_i e^{\pm 2i\psi_i} u(\boldsymbol{x} - \boldsymbol{x}_i) \tag{16}$$

where u is the profile, and $s_i^I, p_i, \psi_i, \boldsymbol{x}_i$ is the flux, polarization fraction, polarization angle, and position of the i'th source respectively. For simplicity we will assume that the sources emit no circular polarization[5]. From the definition of the E and B fields [20, 26]

$$E_{\boldsymbol{\ell}} \pm i B_{\boldsymbol{\ell}} = -\left[Q_{\boldsymbol{\ell}} \pm i U_{\boldsymbol{\ell}}\right] e^{\mp 2i\phi_{\boldsymbol{\ell}}} \tag{17}$$

the T, E, and B fields for the sources take the form:

$$s_{\boldsymbol{\ell}}^T = \sum_i s_i e^{-i\boldsymbol{\ell}\cdot\boldsymbol{x}_i} u_{\boldsymbol{\ell}}$$

$$s_{\boldsymbol{\ell}}^E \pm i s_{\boldsymbol{\ell}}^B = -\sum_i s_i p_i e^{\pm 2i(\psi_i - \phi_{\boldsymbol{\ell}})} e^{-i\boldsymbol{\ell}\cdot\boldsymbol{x}_i} u_{\boldsymbol{\ell}}, \tag{18}$$

where s_i is the flux in the appropriate units.

To compute the responses we must make some assumptions about the joint PDF of the sources $P(\{s_i, p_i, \psi_i, \boldsymbol{x}_i\})$. As we will see in the next subsection, this is easily accomplished with the following simple set of assumptions:

- Each source is statistically independent.

- Positions and polarization angles are statistically independent, both of each other, and independent of polarization fractions and fluxes.

- Positions and polarization angles are uniformly distributed. This key assumption neglects the clustering of sources, and is appropriate for the polarized radio and infrared sources we consider here (on the small scales relevant for lensing reconstruction).

These assumptions force the joint PDF to factorize as[6]:

$$P(\{s_i, p_i, \psi_i, \boldsymbol{x}_i\}) = P(N_{\text{sources}})$$

$$\times \prod_{i=1}^{N_{\text{sources}}} \mathcal{U}(\boldsymbol{x}_i)\mathcal{U}(\psi_i)\text{LF}(s_i)P(p_i|s_i). \tag{19}$$

---

[5] While the primary CMB is expected to have a negligible level of circular polarization in absence of new physics [24], physical processes such as synchrotron emission (which is responsible for radio point sources) can have a circularly-polarized component, which may be potentially detectable by future experiments at low frequencies ($\nu \lesssim 10$ GHz) [25]. It appears that circular polarization should be negligible at CMB frequencies, and it will not be explored further in this work.

[6] In practice, we do not need to assume a specific distribution for the number of sources N_{sources} to compute the response. However, as soon as the sources are independent, the number of sources in any subpatch of the map will be Poisson distributed.

Here, "LF" is the source luminosity function, normalized to integrate to unity, and $\mathcal{U}$ denotes a uniform distribution (in coordinate or angle).

Note that by forcing each source to be independent and uniformly distributed on the sky, any n-point function of the foreground will only involve one independent sum over sources, since

$$\langle \text{Exp}[-i(\boldsymbol{\ell}_1 \cdot \boldsymbol{x}_{i_1} + \cdots + \boldsymbol{\ell}_n \cdot \boldsymbol{x}_{i_n})]\rangle$$

vanishes unless $i_1 = \cdots = i_n$ and $\sum_i \boldsymbol{\ell}_i = 0$, assuming that no non-trivial subset of the $\boldsymbol{\ell}_i$'s sum to zero, which will always be the case for the responses below. That is, our assumptions create the usual shot noise n-point functions.

B. Foreground linear response

Our goal is to compute the linear response $g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY,Z}$ of the covariance $\langle s_{\boldsymbol{\ell}}^X s_{\boldsymbol{L}-\boldsymbol{\ell}}^Y \rangle'$ to the field $s_{\boldsymbol{L}}^Z$ using the simple foreground model outlined in the previous subsection.

We first note that the source cross-power spectra of T, E and B vanish, because the polarization angles are uniform and independent:

$$\langle s_{\boldsymbol{L}}^X s_{-\boldsymbol{L}}^Y \rangle \propto \delta_{XY}^K. \tag{20}$$

A more rigorous argument for this orthogonality goes as follows: when $X = T$ and $Y = E$ or B, the LHS of Eq. (20) vanishes since $\langle e^{\pm 2i\psi_i} \rangle = 0$. When $X = E$ and $Y = B$, we average over $\langle \cos(2\psi_i - 2\phi_{\boldsymbol{L}}) \sin(2\psi_i - 2\phi_{-\boldsymbol{L}}) \rangle \propto \sin(2\phi_{\boldsymbol{L}} - 2\phi_{-\boldsymbol{L}}) = 0$. Eq. (20) follows from these two observations.

With Eq. (20) at our disposal, it is straightforward to compute the responses. We start with the definition of the response g: $\langle s_{\boldsymbol{\ell}}^X s_{\boldsymbol{L}-\boldsymbol{\ell}}^Y \rangle' = \sum_Z g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY,Z} s_{\boldsymbol{L}}^Z$, we multiply both sides by $s_{-\boldsymbol{L}}^{Z'}$, and take an average:

$$g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY,Z} = \frac{\langle s_{\boldsymbol{\ell}}^X s_{\boldsymbol{L}-\boldsymbol{\ell}}^Y s_{-\boldsymbol{L}}^Z \rangle}{\langle s_{\boldsymbol{L}}^Z s_{-\boldsymbol{L}}^Z \rangle}. \tag{21}$$

Using this expression and the foreground PDF (Eq. (19)), we can explicitly compute all unique 18 responses, which are given in Table I. Similar responses to s^Q and s^U were recently derived in Appendix E of [27].

As an illustrative example we will calculate the response of TE to B for the specialized case of point sources ($u_{\boldsymbol{\ell}} = 1$). We first compute the bispectrum $\langle s_{\boldsymbol{\ell}}^T s_{\boldsymbol{L}-\boldsymbol{\ell}}^E s_{-\boldsymbol{L}}^B \rangle$, given by:

$$\sum_i \langle s_i^3 p_i^2 \rangle \int \frac{d\psi_i}{2\pi} \cos(2(\psi_i - \phi_{\boldsymbol{L}-\boldsymbol{\ell}})) \sin(2(\psi_i - \phi_{-\boldsymbol{L}}))$$
$$= \frac{\langle N_{\text{sources}} \rangle}{2} \langle s_i^3 p_i^2 \rangle \underbrace{\sin(2(\phi_{\boldsymbol{L}-\boldsymbol{\ell}} - \phi_{-\boldsymbol{L}}))}_{= \sin(2\theta_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{L}})}. \tag{22}$$

	$Z = T$	E	B
$XY = TT$	$\mathcal{A}$	0	0
TE	0	$\mathcal{B}\cos(2\theta_{\boldsymbol{L},\boldsymbol{L}-\boldsymbol{\ell}})$	$\mathcal{B}\sin(2\theta_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{L}})$
TB	0	$\mathcal{B}\sin(2\theta_{\boldsymbol{L},\boldsymbol{L}-\boldsymbol{\ell}})$	$\mathcal{B}\cos(2\theta_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{L}})$
EE	$\mathcal{C}\cos(2\theta_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}})$	0	0
EB	$\mathcal{C}\sin(2\theta_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}})$	0	0
BB	$\mathcal{C}\cos(2\theta_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}})$	0	0

Table I. The linear response $g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY,Z}$ for all unique combinations of XY (rows) and Z (columns). Note that all entries in the table should be multiplied by a geometric factor $u_{\boldsymbol{\ell}} u_{\boldsymbol{L}-\boldsymbol{\ell}}/u_{\boldsymbol{L}}$. For brevity we have defined $\mathcal{A} \equiv \langle s_i^3 \rangle / \langle s_i^2 \rangle$, $\mathcal{B} \equiv \langle p_i^2 s_i^3 \rangle / \langle p_i^2 s_i^2 \rangle$, and $\mathcal{C} \equiv \langle p_i^2 s_i^3 \rangle / 2\langle s_i^2 \rangle$.

A similar calculation gives the power spectrum $\langle s_{\boldsymbol{L}}^B s_{-\boldsymbol{L}}^B \rangle$:

$$\sum_i \langle s_i^2 p_i^2 \rangle \int \frac{d\psi_i}{2\pi} \sin(2(\psi_i - \phi_{\boldsymbol{L}})) \sin(2(\psi_i - \phi_{-\boldsymbol{L}}))$$
$$= \frac{\langle N_{\text{sources}} \rangle}{2} \langle s_i^2 p_i^2 \rangle \underbrace{\cos(2(\phi_{\boldsymbol{L}} - \phi_{-\boldsymbol{L}}))}_{= 1}. \tag{23}$$

The response $g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TE,B}$ is given by the ratio of Eq. (22) to Eq. (23), corresponding to the value reported in Table I. Note that the factor $\langle N_{\text{sources}} \rangle$ cancels, which is why we do not explicitly need to assume the distribution of the number of sources.

C. Picking the wrong profile

Here we explore the expected bias to a bias hardened estimator (hardened against a profile $u_{\boldsymbol{\ell}}$) from a foreground with profile $w_{\boldsymbol{\ell}} \neq u_{\boldsymbol{\ell}}$. For brevity we will consider the case of temperature-only reconstruction in the large-lens limit ($L \to 0$).

For the case of temperature-only reconstruction the optimal bias hardened lensing weights take the form:

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\text{BH}} = F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\text{MV}} - \frac{\int_{\boldsymbol{\ell}'} F_{\boldsymbol{\ell}',\boldsymbol{L}-\boldsymbol{\ell}'}^{\text{MV}} u_{\boldsymbol{\ell}'} u_{|\boldsymbol{L}-\boldsymbol{\ell}'|}}{\int_{\boldsymbol{\ell}'} G_{\boldsymbol{\ell}',\boldsymbol{L}-\boldsymbol{\ell}'}^{\text{MV}} u_{\boldsymbol{\ell}'} u_{|\boldsymbol{L}-\boldsymbol{\ell}'|}} G_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\text{MV}}, \tag{24}$$

as can be shown with Eq. (15). In Eq. (24) we have assumed that the input profile $u_{\boldsymbol{\ell}}$ is isotropic, and F^{MV} (G^{MV}) are the minimum variance temperature-only weights for the lensing convergence (foreground):

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\text{MV}} = f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TT}/2\tilde{C}_{\ell}^{TT} C_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}$$
$$G_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\text{MV}} = \mathcal{A} u_{\boldsymbol{\ell}} u_{|\boldsymbol{L}-\boldsymbol{\ell}|}/u_{\boldsymbol{L}} 2\tilde{C}_{\ell}^{TT} C_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}. \tag{25}$$

As an aside, note that the bias hardened weights (Eq. (24)) are insensitive to an overall normalization of G^{MV}. Hence bias hardening the TT estimator requires no knowledge of the source amplitude $\mathcal{A}$, or the normalization of the profile. In Appendix C 1 we show that this is always the case when bias hardening any single-pair lensing estimator, but is no longer true when bias hardening the global minimum variance estimator.

Suppose that the true foreground has an isotropic profile $w_\ell \neq u_\ell$. In the $L \to 0$ limit the response of a generic lensing estimator $\hat{\kappa}$ (with temperature weights F^κ and normalization N^κ) to the true foreground is:

$$
\begin{aligned}
\mathcal{R}_{\mathbf{0}}^{\kappa,w} &= N_{\mathbf{0}}^\kappa \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},-\boldsymbol{\ell}}^\kappa g_{\boldsymbol{\ell},-\boldsymbol{\ell}}^{TT,T} \\
&= \mathcal{A} \int \frac{d\ell}{(2\pi)^2} w_\ell^2 \left[N_{\mathbf{0}}^\kappa \ell \int d\theta_{\boldsymbol{L},\boldsymbol{\ell}} F_{\boldsymbol{\ell},-\boldsymbol{\ell}}^\kappa \right],
\end{aligned}
\tag{26}
$$

such that the field level bias to $\langle \hat{\kappa}_{\mathbf{0}} \rangle'$ is $\mathcal{R}_{\mathbf{0}}^{\kappa,w} s_{\mathbf{0}}^T$. Here we have assumed that the profile is normalized so that $\lim_{\ell \to 0} u_\ell = 1$.
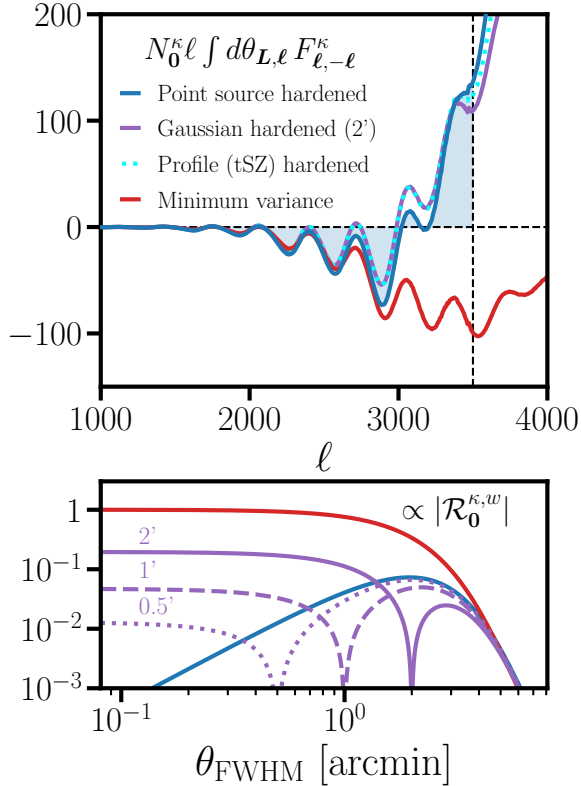
Figure 1. *Top:* The angular averaged temperature-only lensing weights $(\times 2\pi N_{\mathbf{0}}^\kappa \ell)$ in the large-lens limit $(L \to 0)$ when hardening against a point source (blue), hardening against a Gaussian profile with a 2′ FHWM (purple) and when choosing the weights to minimize the variance (red). Note that since the point source hardened estimator has zero response to point-sources (when $w_\ell = 1$) the area under the blue curve vanishes. *Bottom:* The response to a Gaussian foreground w_ℓ with a full width half max θ_{FHHM}, normalized to the response of the MV estimator to point sources. In both plots we take $\ell_{\mathrm{max},T} = 3500$. See section V for details regarding the instrumental assumptions.

In the bottom panel of Fig. 1 we show the values of $|\mathcal{R}_{\mathbf{0}}^{\kappa,w}|$ when the true foreground profile w_ℓ is a Gaussian with a full width half maximum (FWHM) given by θ_{FWHM} (in arcmin), and consider three cases for F^κ: minimum variance weights (red), hardening against a Gaussian profile (purple), and hardening against a point source

(blue). The angular averaged values of these weights are plotted in the top panel of Fig. 1.

We find that point source hardening reduces the response to foregrounds smaller than ~ 2 arcmin by an order of magnitude relative to the MV weights. By instead hardening against a Gaussian with a non-zero width, one can further reduce (or precisely null) the response to foregrounds whose angular size is within $\sim 40\%$ of the size of the profile. In particular the tSZ-profile hardened estimator [6] (cyan), which is well approximated by the Gaussian-hardened estimator with a 2 arcmin FWHM, reduces the response to point sources by roughly a factor of 5 relative to the MV estimator.

Finally, if the foreground is larger than a few arcmin, the response is always small regardless of the assumed profile or technique, since the primary CMB dominates dominates on these scales (i.e. the lensing weights have no support for $\ell < 1500$).

IV. LARGE-LENS SUPPRESSED ESTIMATORS

In this section we study the "large-lens" or "squeezed" limit $(L \ll \ell)$ analytically. A similar study for the temperature-only case has given rise to "shear-only" estimators [7], which are highly effective at reducing foreground biases in this limit. Here we extend this study to polarization, highlighting important differences compared to temperature.

As in section III, define $\mathcal{R}_{\boldsymbol{L}}^{XY,Z}$ as the response of a single-pair lensing estimator to a foreground in the Z'th map:

$$
(N_{\boldsymbol{L}}^{XY})^{-1} \mathcal{R}_{\boldsymbol{L}}^{XY,Z} = \int \frac{\ell d\ell}{(2\pi)^2} \int d\theta_{\boldsymbol{L},\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY,Z},
\tag{27}
$$

so that the field level bias to $\langle \hat{\kappa}_{\boldsymbol{L}}^{XY} \rangle'$ due to $s_{\boldsymbol{L}}^Z$ is $\mathcal{R}_{\boldsymbol{L}}^{XY,Z} s_{\boldsymbol{L}}^Z$. The inverse of the normalization can similarly be expressed as:

$$
(N_{\boldsymbol{L}}^{XY})^{-1} = \int \frac{\ell d\ell}{(2\pi)^2} \int d\theta_{\boldsymbol{L},\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY}.
\tag{28}
$$

The large-lens suppressed (LLS) estimators are derived by choosing the weights to suppress the field level bias by $\mathcal{O}((L/\ell)^2)$ relative to the lensing signal, where $\ell \sim 3000$ is evaluated in the domain where the lensing estimator receives the majority of its signal (see the top panel of Fig. 1). This can be accomplished by forcing Eq. (27) to be suppressed by $\mathcal{O}((L/\ell)^2)$ while simultaneously retaining a $\mathcal{O}(1)$ scaling in the inverse normalization.

We first consider the temperature-only case. By Taylor expanding the angular-dependent piece of foreground linear response $g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TT,T} = \mathcal{A} u_\ell u_{\boldsymbol{L}-\boldsymbol{\ell}}/u_L$ in powers of L/ℓ (see Appendix A) we find:

$$
\begin{aligned}
g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TT,T} = \mathcal{A} \frac{u_\ell}{u_L} \Big\{ & u_\ell - L u_\ell' \cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) \\
& + \frac{1}{4} \left(\frac{L}{\ell} \right)^2 \left[D_\ell^+ + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) D_\ell^- \right] u_\ell + \mathcal{O}\left((L/\ell)^3 \right) \Big\},
\end{aligned}
\tag{29}
$$

with corrections going as $\mathcal{O}\left((L/\ell)^3\right)$, and we have defined the operators $D_\ell^\pm \equiv \ell^2 d^2/d\ell^2 \pm d/d\ln\ell$. Likewise, the linear response of the CMB to lensing also factorizes into a nice multipole form [7, 28]:

$$f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TT} = C_\ell^{TT}\left[\frac{d\ln\ell^2 C_\ell^{TT}}{d\ln\ell} + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})\frac{d\ln C_\ell^{TT}}{d\ln\ell}\right] \quad (30)$$

with corrections going as $\mathcal{O}(L/\ell)$.

In a similar manner one can expand the lensing weights into multipoles:

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TT} = \sum_{m=0}^{\infty}\left(A_{L,\ell}^{(m)}\cos(m\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + B_{L,\ell}^{(m)}\sin(m\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right),$$
$$(31)$$

and in turn express the coefficients $A^{(m)}$ and $B^{(m)}$ as a Taylor expansion in L/ℓ. From Eq. (29), we see that setting $A^{(0)} = \mathcal{O}((L/\ell)^2)$, $A^{(1)} = \mathcal{O}(L/\ell)$ forces the response to foregrounds (27) to be suppressed by $\mathcal{O}((L/\ell)^2)$. From Eq. (30), we see that retaining a $\mathcal{O}(1)$ scaling in Eq. (28) requires choosing $A^{(2)} = \mathcal{O}(1)$. The remaining coefficients are chosen to minimize the noise of the LLS estimator in the large-lens limit. The large-lens noise (Eq. (6) as $L \to 0$) of the temperature-only LLS estimator takes the form:

$$\mathcal{N}_0^{TT} = \frac{\int\frac{\ell d\ell}{2\pi}(C_\ell^{TT})^2\left[\sum_{m=2}^{\infty}\left(A_{0,\ell}^{(m)}\right)^2 + \sum_{m=1}^{\infty}\left(B_{0,\ell}^{(m)}\right)^2\right]}{\left[\int\frac{\ell d\ell}{4\pi}A_{0,\ell}^{(2)}\frac{dC_\ell^{TT}}{d\ln\ell}\right]^2}.$$
$$(32)$$

Note that the $A^{(2)}$-term is the only term contributing to the normalization (since we set $A_{0,\ell}^{(0)} = 0$). Thus a non-zero value for the remaining coefficients ($A^{(m)}$ for $m > 2$ and all $B^{(m)}$) only increases the noise of the LLS estimator. To minimize the large-lens noise, we therefore choose the remaining coefficients to be $\mathcal{O}(L/\ell)$, and solve for the value of $A_{0,\ell}^{(2)}$ which minimizes Eq. (32). Doing so forces the LLS estimator to take the form:

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{LLS},TT} = \frac{1}{2(\tilde{C}_\ell^{TT})^2}\frac{dC_\ell^{TT}}{d\ln\ell}\cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \mathcal{O}(L/\ell), \quad (33)$$

recovering the shear-only estimator [7]. As was shown in [29], while the shear-only estimator suppresses the bias by $\mathcal{O}((L/\ell)^2)$, the bias *is not precisely nulled* for a source with a generic profile, even if the profile is azimuthally symmetric. We further note that Eq. (29) also contains derivatives of the (Fourier-transformed) profile, which additionally suppress biases for small sources. In the special case of point sources, for which these derivatives vanish, the shear-only estimator is non-perturbatively unbiased.

We explicitly generalize this procedure for all pairs of temperature and polarization maps in Appendix E. Our results are summarized by:

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{LLS},TE} = \frac{C_\ell^{TE}}{\tilde{C}_\ell^{TT}\tilde{C}_\ell^{EE}}\frac{d\ln\ell^2|C_\ell^{TE}|}{d\ln\ell} + \mathcal{O}(L/\ell)$$

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{LLS},TB} = \mathcal{O}(L/\ell)$$

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{LLS},EE} = \frac{1}{2(\tilde{C}_\ell^{EE})^2}\frac{dC_\ell^{EE}}{d\ln\ell}\cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \mathcal{O}(L/\ell) \quad (34)$$

$$F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{LLS},EB} = F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{MV},EB} = \frac{f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{EB}}{\tilde{C}_\ell^{EE}\tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{BB}}.$$

One can continue this procedure to obtain LLS estimators to any order in L/ℓ. Including higher order corrections will in principle make the estimators less noisy, however, by construction the approximations in Eqs. (33) and (34) are only slightly suboptimal in the regime where the large-lens expansion is valid.

As indicated by Eq. (34), we find that there is no leading order LLS TB estimator. As shown in Appendix A and E, this is because $f^{TB} \propto g^{TB,E} \propto \sin(2\theta)$ to lowest order in the large-lens limit, i.e. foregrounds and lensing are indistinguishable to lower order. On the contrary we find that the MV EB estimator already satisfies the LLS requirements.

	$\mathcal{O}$(MV biases)	$\mathcal{O}$(LLS suppression)
TT	1	$L^2 u_\ell''/u_\ell$, $L^2 u_\ell'/\ell u_\ell$
TE	$p_{\mathrm{source}}/p_{\mathrm{CMB}}$	$L^2 u_\ell''/u_\ell$, $L^2 u_\ell'/\ell u_\ell$
TB	$p_{\mathrm{source}}/p_{\mathrm{CMB}}$	$-$
EE	$(p_{\mathrm{source}}/p_{\mathrm{CMB}})^2$	$(L/\ell)^2$
EB	$(L/\ell)^2 \times (p_{\mathrm{source}}/p_{\mathrm{CMB}})^2$	$-$

Table II. In the first column we show scaling of the field level biases for each single-pair MV estimator relative to the MV TT field-level bias. Here p_{source} refers to the polarization fraction of the foreground, while $p_{\mathrm{CMB}} \sim \sqrt{C_\ell^{EE}/C_\ell^{TT}}$, where ℓ should be evaluated where the lensing weights receive the majority of their support ($\ell \sim 3000$). In the second column we show the scaling for the field-level bias suppression when the MV estimators are substituted for LLS estimators. Note that there is no LLS TB estimator to leading order in L/ℓ, while the LLS EB estimator is equivalent to the MV EB estimator.

We show the scaling of the field-level biases for all single-pair estimators in Table II. In the first column we show the scaling of the minimum variance (MV) field-level bias (relative to the MV TT bias) with the polarization fractions of the foreground and CMB. In the second column we show the scaling of the field-level bias suppression for each LLS estimator. Note that while all LLS estimators reduce the bias by $\mathcal{O}((L/\ell)^2)$, the TT and TE estimators additionally reduce the bias by factors involving derivatives of the Fourier transform of the profile (e.g. $\ell u_\ell' \equiv \ell\, du_\ell/d\ell$). For the special case of point sources ($u_\ell = 1$) these derivatives vanish, and the (TT, TE) LLS estimators are unbiased to all orders in L/ℓ.
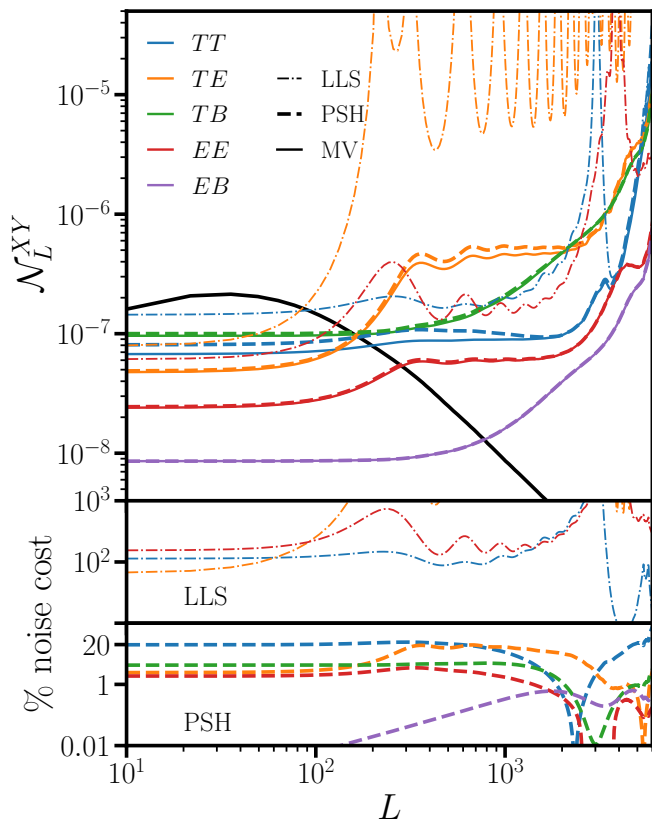
Figure 2. The top panel shows the reconstruction noise (per Fourier mode) when reconstructing on a single pair of maps. The colors denote the pair XY. Solid, dashed and dashed-dotted curves represent minimum variance (MV), point source hardened (PSH) or large-lens suppressed (LLS) estimators respectively. $C_L^{\kappa\kappa}$ is shown in black. The bottom two panels show the percent noise cost for the LLS and PSH estimators relative to the appropriate MV estimator.

V. NOISE PROPERTIES

Here we compute the variance of the estimators discussed in sections II, III and IV. We consider a full-sky CMB experiment at a single frequency (150 GHz) with stage-4 like sensitivity ($\Delta_T = 1$ μK-arcmin, $\Delta_P = \sqrt{2}\Delta_T$), and assume a Gaussian beam with a 1.4 arcmin FWHM. Throughout we take $\ell_{\mathrm{max},T} = 3500$ and $\ell_{\mathrm{max},P} = 5000$.

In our total temperature power spectrum $\tilde{C}_\ell^{TT}$ we include the lensed CMB, instrument noise, and contributions from extragalactic foregrounds (tSZ, kSZ, CIB and radio point sources) using `LensQuEst`[7], which adopts templates for the extragalactic foreground power spectra from [30]. With the exception of tSZ × CIB, we neglect the cross-correlation of the components listed above. For polarization-only spectra and for cross-correlations of temperature and polarization, we include the lensed CMB

---

[7] https://github.com/EmmanuelSchaan/LensQuEst

and instrument noise (where appropriate) in our total power spectra, but neglect extragalactic foregrounds. We use `symlens`[8] to implement our estimators and to compute noise curves.

We first consider the variance of single-pair estimators, given by Eq. (6) and plotted in Fig. 2. We see that point source hardening comes at a $1-20\%$ noise cost (relative to the MV estimators) for most pairs of maps, and that bias hardening polarization-based estimators generically comes at a lower noise cost than the temperature-only case. In particular, the noise cost for hardening EB against point sources is less than 0.1% for all signal dominated scales ($L < 1000$). By contrast, the single-pair LLS estimators come at more modest factor of ~ 2 noise cost for low L. As shown in Appendix E, all three non-trivial LLS estimators (TT, TE, EE) are only sensitive to one of the two terms (monopole and quadrupole) in the lensing linear response in the large lens limit. Since these two terms are comparable in magnitude, the LLS estimators are only sensitive to roughly half of the lensing signal, providing some intuition for the approximate factor of 2 cost in noise[9].
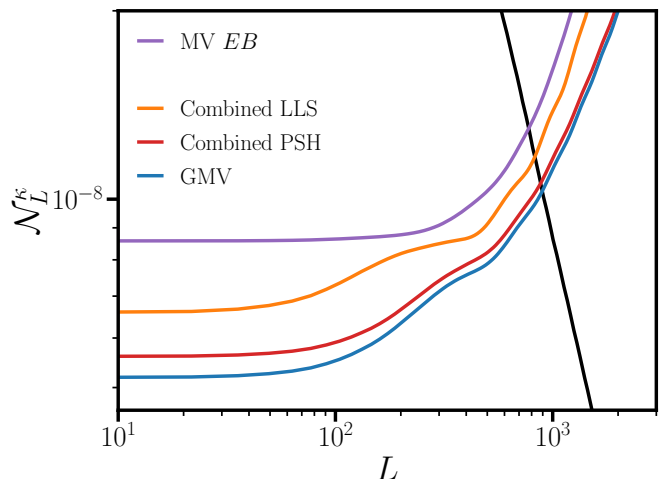

Figure 3. Lensing reconstruction noise for the optimal linear combination of single-pair LLS (orange) and point source hardened (PSH, red) estimators. The noise for the global MV (GMV) estimator is shown in blue. For reference we show the MV EB noise (purple) and $C^{\kappa\kappa}$ (black).

We next consider the reconstruction noise for combined estimators, shown in Fig. 3. In blue we show the reconstruction noise for the Global Minimum Variance estimator (i.e. Eq. (9)), whose weights are explicitly given by Eq. (B2). In orange and red we show the reconstruction noises for optimal linear combinations (see Appendix C) of the single-pair LLS and point source hardened (PSH) estimators respectively. In particular, we find that near

---

[8] https://github.com/simonsobs/symlens
[9] That is, from Eq. (32) we see that the large-lens noise very roughly scales as (# of multipoles)/(# of multipoles)^2, assuming that each multipole is of a similar size (and sign).

the transition of signal- to noise-domination ($L \sim 1000$) the combined PSH estimator's noise is only 4% larger than the GMV's. The combined LLS estimator comes at more modest $\sim 20\%$ noise penalty.

In Appendix C we explore the feasibility of constructing global forms of the LLS and PSH estimators, however, we find that an optimal linear combination of single-pair estimators is both more practical and nearly optimal in both cases.

VI. BIAS REDUCTION

Biases to the CMB lensing power spectrum can be decomposed into primary, secondary and trispectrum components [2, 4]. For simplicity consider the case of a single-pair estimator $\hat{\kappa}_{L}^{XY}$ which is used to measure $\hat{C}_{L}^{\kappa\kappa} \propto \langle \hat{\kappa}_{L}^{XY} \hat{\kappa}_{-L}^{XY} \rangle$ from a set of maps containing the lensed CMB $(\tilde{X}, \tilde{Y})$ and the foregrounds of interest (s^X, s^Y). Schematically, let $\hat{\kappa}[X, Y]$ denote reconstruction with the estimator $\hat{\kappa}_{L}^{XY}$ on the maps X_{ℓ}, $Y_{L-\ell}$. Using this notation the primary, secondary and trispectrum biases take the form:

$$
\begin{aligned}
\text{Primary} &= 2\langle \kappa\, \hat{\kappa}[s^X, s^Y] \rangle \\
\text{Secondary} &= 2\langle \hat{\kappa}[\tilde{X}_0, s^Y]\hat{\kappa}[\tilde{X}_1, s^Y] \rangle \\
&\quad + 2\langle \hat{\kappa}[s^X, \tilde{Y}_0]\hat{\kappa}[s^X, \tilde{Y}_1] \rangle \\
&\quad + 2\langle \hat{\kappa}[s^X, \tilde{Y}_0]\hat{\kappa}[\tilde{X}_1, s^Y] \rangle \\
&\quad + 2\langle \hat{\kappa}[s^X, \tilde{Y}_1]\hat{\kappa}[\tilde{X}_0, s^Y] \rangle \\
\text{Trispectrum} &= \langle \hat{\kappa}[s^X, s^Y]\hat{\kappa}[s^X, s^Y] \rangle_c,
\end{aligned}
\tag{35}
$$

where we have expanded $\tilde{X} = \tilde{X}_0 + \tilde{X}_1 + \cdots$ in powers of κ to remove noise bias in our calculation of the secondary [7]. The Gaussian component of the foreground four-point function is subtracted analytically to obtain the trispectrum, as indicated by the c subscript.

A. From field-level to power spectrum bias suppression

Both bias hardened and LLS estimators aim to suppress the field-level response to foregrounds, which schematically take the form $\int Fg$, where F is the lensing weights and g is the foreground linear response. Here we relate these field-level responses to the primary and trispectrum biases to the CMB lensing power spectrum.

The primary bias ultimately arises from a non-zero $\langle s_{\ell} s_{L-\ell} \kappa_{-L} \rangle$ bispectrum. Since it is trivially the case that $\langle s_{\ell} s_{L-\ell} \kappa_{-L} \rangle = \langle \langle s_{\ell} s_{L-\ell} \rangle' \kappa_{-L} \rangle$, one can expand this bispectrum in terms of the linear foreground response and the foreground-lensing cross-power spectrum:

$$
\langle s_{\ell}^{X} s_{L-\ell}^{Y} \kappa_{-L} \rangle \propto \sum_{Z} g_{\ell, L-\ell}^{XY,Z} \langle s_{L}^{Z} \kappa_{-L} \rangle.
\tag{36}
$$

As a consequence the primary bias is proportional to the field-level response. Thus choosing the weights F to null the field-level response also nulls the primary bias.

Likewise, for a foreground obeying the assumptions discussed in Section III A, we also find that the trispectrum can be expressed as a sum over power spectra, this time weighted by two factors of the foreground linear response:

$$
\langle s_{\ell}^{X} s_{L-\ell}^{Y} s_{\ell'}^{M} s_{-L-\ell'}^{N} \rangle \propto \sum_{Z} g_{\ell, L-\ell}^{XY,Z}\, g_{\ell', -L-\ell'}^{MN,Z}\, \langle s_{L}^{Z} s_{-L}^{Z} \rangle,
\tag{37}
$$

when at least one of the maps is a temperature map. Thus, for these cases, the trispectrum bias is proportional to the field-level response squared. Polarization-only trispectra also contain a term proportional to either $\cos(2\Phi)$ or $\sin(2\Phi)$, where $\Phi \equiv \phi_{\ell} + \phi_{L-\ell} - \phi_{\ell'} - \phi_{-L-\ell'}$, in addition to those in Eq. (37). In practice, we find that these extra terms are highly subdominant to those in Eq. (37), as shown in Fig. 5.

One can also obtain expressions analogous to Eq. (36) for the secondary bias, which is sourced by the same $\langle ss\kappa \rangle$ bispectrum. However, the angular structure of the secondary bias differs significantly from that of the primary, and as a result there is no simple relation between the field-level response and the secondary bias. Thus we expect nulling the field-level bias to null both the primary and trispectrum, but not the secondary. We outline a new approach to null the secondary in Appendix D, but leave further development to future work.
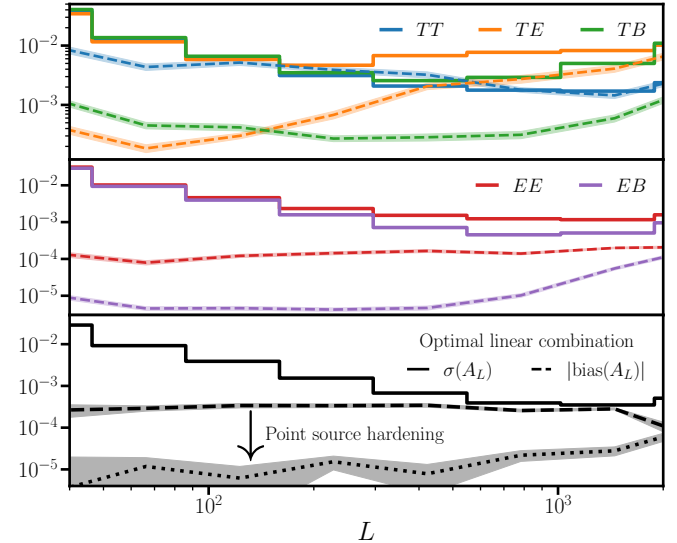

Figure 4. *Top two panels:* Solid lines show the 1σ error of the lensing amplitude ($A_L \equiv C_L^{\kappa\kappa,\text{measured}}/C_L^{\kappa\kappa,\text{true}}$) for each single-pair minimum variance (MV) estimator. Colors denote pairs of temperature and polarization. Dashed lines show the simulated (absolute) biases to the MV estimators due to radio point sources, with the shaded regions denoting the 1σ scatter of the mean. *Bottom:* Error (solid) and bias (dashed) to the lensing amplitude for the optimal linear combination of MV single-pair estimators. The dotted curve shows the bias for an optimal combination of point source hardened estimators. See sections V and VI for details regarding the experimental configuration and simulations respectively.
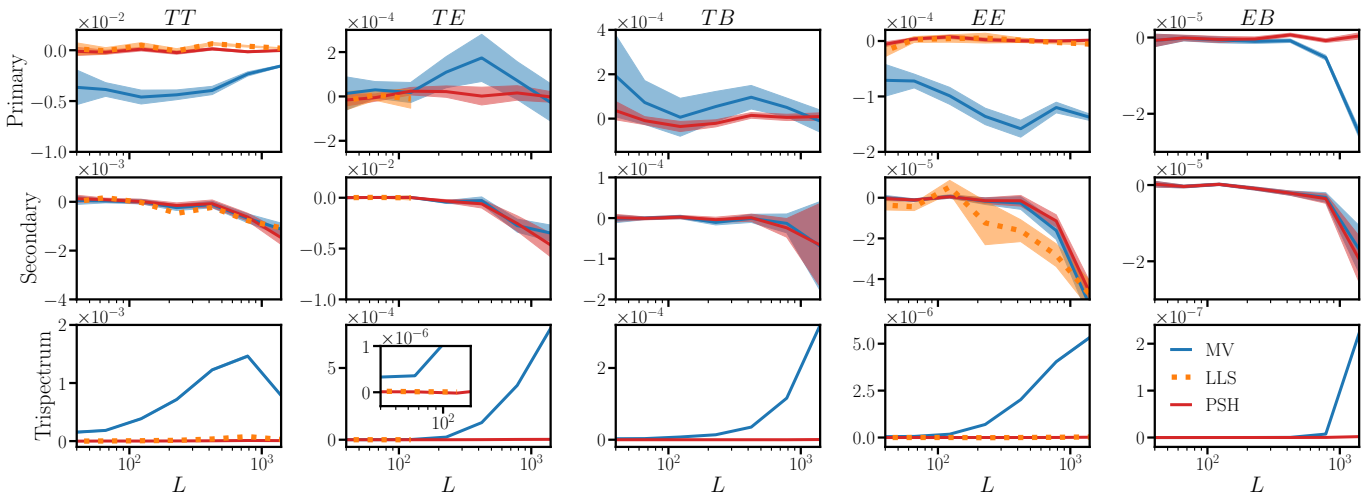
Figure 5. Simulated relative biases to the CMB lensing power spectrum from point sources (assuming a 5 mJy flux cut and 3% polarization fraction) for single-pair lensing estimators with $\ell_{\mathrm{max},T} = 3500$ and $\ell_{\mathrm{max},P} = 5000$, and CMB-S4 like sensitivity. Columns correspond to different pairs of temperature and polarization, while the rows indicate primary, secondary or trispectrum biases. Biases for the Minimum Variance (MV), Large-Lens Suppressed (LLS) and Point Source Hardened (PSH) estimators are shown in blue, orange and red respectively. Note that for the TE LLS estimator we only show biases for $L < 120$ since this estimator's noise dramatically increases at smaller scales (see Fig. 2), and that the TE, TB primary biases vanish analytically. The assumed L-binning is depicted in Fig. 4, and the shaded regions denote the 1σ scatter of the mean biases across 40 cutouts. We stress that the TT estimator here only shows the contribution from point sources: the overall bias is significantly larger when contributions from the CIB, tSZ and kSZ are included [6].

B. Power spectrum biases: simulations

To numerically estimate the biases in Eq. (35) we use the WebSky simulations of both the lensing convergence [31] and radio galaxies [17] (in temperature). These simulations accurately reproduce number counts measured from current CMB and radio telescopes, and have a realistic correlation between the galaxies and lensing convergence. Following [6, 7] we mask sources brighter than 5 mJy by fulling excluding the pixels associated with those sources.

We generate mock polarization data by assigning random polarization angles (sampled from a uniform distribution) to each source. For simplicity we assume that all sources have the same 3% polarization fraction. The measured biases presented below tend to scale simply with moments of the polarization fraction distribution, e.g. the EE trispectrum bias $\propto \langle p^4 \rangle$. Thus one can scale our results to any desired polarization fraction distribution.

We divide a subsample of the WebSky simulation into 40 square 15 degree cutouts using `pixell`[10]. For each cutout we generate mock Gaussian CMB maps with `LensQuEst` and lens these maps using the corresponding WebSky convergence cutout. We run the lensing reconstruction for each single-pair estimator under the flat sky approximation using `symlens`, assuming the S4-like configuration described in section V.

First, in top two panels of Fig. 4 we show the sensitivity[11] to the lensing amplitude of each single-pair minimum variance (MV) estimator, along with the total bias (primary + secondary + trispectrum) from polarized point sources. We note that these biases are significant ($\sim 0.5 - 1\sigma$) for both the TT and TE estimators, while for the remaining single-pair estimators they can approach $\sim 0.1\sigma$ at small ($L > 500$) scales. The bias to the optimal linear combination of single-pair MV estimators (bottom panel) is $\sim 1\sigma$ for $L > 500$, and non-negligible on larger scales. These biases are suppressed by up to two orders of magnitude when one instead takes an optimal linear combination of point source hardened estimators, ensuring bias $< \sigma/50$ for $L < 500$ and $< \sigma/20$ for $500 < L < 1000$. We show an analogous figure for a Simons Observatory-like survey in Appendix G, and again find that the biases to TT and TE are significant ($\sim 1\sigma$), while for the remaining estimators the biases are negligible ($< \sigma/100$).

In Fig. 5 we show the individual (relative) primary, secondary and trispectrum biases for the Minimum Variance (MV; blue), Point Source Hardened (PSH; orange), and Large-Lens Suppressed estimators (LLS; red). We find that both point source hardening and large-lens suppression effectively null the primary and trispectrum biases for all single-pair estimators, but have little impact on the secondary bias, as expected. Since the total TT, TB, EE, and EB biases receive sizeable contributions from the primary and trispectrum terms, one signif-

---

[10] https://github.com/simonsobs/pixell

---

[11] The sensitivities quoted here are for the assumed $L-$binning in Fig. 4, and would be more significant if broader bins were used.

icantly reduces the overall bias for these estimators. However, the TE bias is dominated by the secondary $\langle s^T s^T \kappa \rangle$ contribution, and thus the TE bias is only marginally reduced using these techniques. If necessary one can null this secondary contribution using the technique outlined in Appendix D. As indicated in the bottom panel of Fig. 4, the TE secondary is a subdominant contribution to the bias from the optimal combination of PSH estimators, and we therefore expect that an optimal linear combination of PSH estimators is sufficient for a robust and precise measurement of $C^{\kappa\kappa}$ with future surveys.

We note that there is another population of "Infrared" (IR) point sources which are dominant at higher frequency and have not been included in the simulations described here, since their polarization fraction is expected to be smaller than for radio sources [18, 19]. The same techniques discussed here should also mitigate the impact of polarized IR sources with similar effectiveness.

VII. DISCUSSION AND CONCLUSIONS

In this paper we generalized previous work to polarization, showing that geometric methods are effective at reducing the bias to CMB lensing in both temperature and polarization, even on single-frequency maps. We used state-of-the art simulations of radio point sources to estimate the size of these biases and the effectiveness of our mitigation strategies. We have found that the point-source bias in polarization is relatively small for upcoming experiments such as Simons Observatory (Appendix G), but it can become potentially significant on small scales for future low-noise experiments (Fig. 4). We have shown that point source hardened estimators, which come at a negligible noise cost, can very effectively control these biases for all experiments in the foreseeable future. We find that LLS estimators are equally as effective at reducing these biases, however, generically come at a significantly higher noise cost.

We note that patchy reionization is also expected to give a contribution to the temperature fluctuations (through the kSZ effect) and to polarization, through patchy screening and scattering [32]. While we expect the biases induced by reionization to be subdominant (this has been shown explicitly for temperature in [33]), we note that that the majority of the reionization bubbles will be sub-beam in size, and therefore we expect PSH to decrease any possible bias. A full study of this is left to future work.

In this study we have not explored the impact of *galactic* foregrounds and potential mitigation strategies. In [34], the authors show that foreground-induced biases can be significant in polarization, and consider multi-frequency cleaning as a partial mitigation strategy. While we note that the simple geometry of point sources allowed us to analytically write estimators with zero response to them, some of the same techniques presented here may prove useful when dealing with galactic foregrounds. The lensing-induced distortion provides a consistency between the different angular multipoles discussed in Section IV. Due to their complex geometry, galactic foregrounds are unlikely to match such a consistency and therefore they may be distinguishable using geometric methods such as the ones discussed in this paper. Finally, it's worth noting that the same dust seen in emission in the CMB maps also causes extinction at optical wavelengths, potentially modulating the observed galaxy field and creating biases to cross-correlations. This further motivates this line of research, which is left to a future investigation.

[1] A. van Engelen *et al.*, Astrophys. J.**786**, 13 (2014).

[2] S. Ferraro and J. C. Hill, Phys. Rev. D**97**, 023512 (2018).

[3] T. Namikawa, D. Hanson, and R. Takahashi, Mon. Not. R. Astron. Soc. **431**, 609 (2013).

[4] S. J. Osborne, D. Hanson, and O. Doré, J. Cosm. Astrop. Phys. **2014**, 024 (2014).

[5] T. Namikawa and R. Takahashi, Mon. Not. Roy. Astron. Soc. **438**, 1507 (2014).

[6] N. Sailer, E. Schaan, and S. Ferraro, Phys. Rev. D **102**, 063517 (2020).

[7] E. Schaan and S. Ferraro, Phys. Rev. Lett.**122**, 181301 (2019).

[8] M. S. Madhavacheril and J. C. Hill, Phys. Rev. D**98**, 023534 (2018).

[9] N. Sailer *et al.*, Phys. Rev. D **104**, 123514 (2021).

[10] O. Darwish *et al.*, arXiv e-prints arXiv:2111.00462 (2021).

[11] A. Baleato Lizancos and S. Ferraro, Phys. Rev. D **106**, 063534 (2022).

[12] B. A. Benson *et al.*, **9153**, 91531P (2014).

[13] P. Ade *et al.*, J. Cosm. Astrop. Phys. **2019**, 056 (2019).

[14] K. Abazajian *et al.*, arXiv e-prints arXiv:2203.08024 (2022).

[15] C. Feng and G. Holder, Astrophys. J. **897**, 140 (2020).

[16] A.-S. Deutsch, M. C. Johnson, M. Münchmeyer, and A. Terrana, JCAP **04**, 034 (2018).

[17] Z. Li, G. Puglisi, M. S. Madhavacheril, and M. A. Alvarez, JCAP **08**, 029 (2022).

[18] R. Datta *et al.*, Mon. Not. Roy. Astron. Soc. **486**, 5239 (2019).

[19] N. Gupta *et al.*, Mon. Not. Roy. Astron. Soc. **490**, 5712 (2019).
[20] A. Lewis and A. Challinor, Phys. Rep. **429**, 1 (2006).
[21] W. Hu and T. Okamoto, Astrophys. J.**574**, 566 (2002).
[22] C. M. Hirata and U. Seljak, Phys. Rev. D **68**, 083002 (2003).
[23] A. S. Maniyar *et al.*, Phys. Rev. D **103**, 083524 (2021).
[24] P. Montero-Camacho and C. M. Hirata, JCAP **08**, 040 (2018).
[25] S. King and P. Lubin, Phys. Rev. D **94**, 023501 (2016).
[26] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997).
[27] P. A. R. Ade *et al.*, arXiv e-prints arXiv:2210.08038 (2022).
[28] H. Prince, K. Moodley, J. Ridl, and M. Bucher, J. Cosm. Astrop. Phys. **2018**, 034 (2018).
[29] F. J. Qu, A. Challinor, and B. D. Sherwin, arXiv e-prints arXiv:2208.14988 (2022).
[30] J. Dunkley *et al.*, J. Cosm. Astrop. Phys. **2013**, 025 (2013).
[31] G. Stein *et al.*, JCAP **10**, 012 (2020).
[32] C. Dvorkin and K. M. Smith, Phys. Rev. D **79**, 043003 (2009).
[33] H. Cai, M. S. Madhavacheril, J. C. Hill, and A. Kosowsky, Phys. Rev. D **105**, 043516 (2022).
[34] D. Beck, J. Errard, and R. Stompor, JCAP **06**, 030 (2020).

Appendix A: Expansions in the large-lens limit

Any function of $|\boldsymbol{L} - \boldsymbol{\ell}|$ can be Taylor expanded in powers of L/ℓ to give:

$$u_{|\boldsymbol{L}-\boldsymbol{\ell}|} = u_\ell - Lu'_\ell \cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \frac{1}{4}\left(\frac{L}{\ell}\right)^2 \left[D_\ell^+ u_\ell + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})D_\ell^- u_\ell\right] \tag{A1}$$

with corrections going as $\mathcal{O}((L/\ell)^3)$, where in Eq. (A1) we have defined the operators

$$D_\ell^\pm \equiv \ell^2 \frac{d^2}{d\ell^2} \pm \frac{d}{d\ln\ell}. \tag{A2}$$

We can similarly Taylor expand the following geometric factors to quadratic order:

$$\begin{aligned}
\cos(2\theta_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}) &= 1 - \left(\frac{L}{\ell}\right)^2 + \left(\frac{L}{\ell}\right)^2 \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) \\
\sin(2\theta_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}) &= 2\frac{L}{\ell}\sin(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \left(\frac{L}{\ell}\right)^2 \sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) \\
\cos(2\theta_{\boldsymbol{L},\boldsymbol{L}-\boldsymbol{\ell}}) &= \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \frac{L}{\ell}\left[\cos(3\theta_{\boldsymbol{L},\boldsymbol{\ell}}) - \cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right] + \left(\frac{L}{\ell}\right)^2\left[\cos(4\theta_{\boldsymbol{L},\boldsymbol{\ell}}) - \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right] \\
\sin(2\theta_{\boldsymbol{L},\boldsymbol{L}-\boldsymbol{\ell}}) &= \sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \frac{L}{\ell}\left[\sin(3\theta_{\boldsymbol{L},\boldsymbol{\ell}}) - \sin(\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right] + \left(\frac{L}{\ell}\right)^2\left[\sin(4\theta_{\boldsymbol{L},\boldsymbol{\ell}}) - \sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right]
\end{aligned} \tag{A3}$$

Using these expressions one can Taylor expand the linear responses in Eq. (4):

$$f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TT} = C_\ell^{TT}\left[\frac{d\ln\ell^2 C_\ell^{TT}}{d\ln\ell} + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})\frac{d\ln C_\ell^{TT}}{d\ln\ell}\right] - \frac{1}{2}\frac{L}{\ell}\left[D_\ell^+ C_\ell^{TT} + 4\frac{dC_\ell^{TT}}{d\ln\ell}\cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})D_\ell^- C_\ell^{TT}\right]$$

$$f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TE} = C_\ell^{TE}\left[\frac{d\ln\ell^2 |C_\ell^{TE}|}{d\ln\ell} + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})\frac{d\ln C_\ell^{TE}}{d\ln\ell}\right] - \frac{1}{2}\frac{L}{\ell}\left[(D_\ell^+ + 4)C_\ell^{TE} + 4\frac{dC_\ell^{TE}}{d\ln\ell}\cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})(D_\ell^- - 4)C_\ell^{TE}\right]$$

$$f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{TB} = 2\sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})C_\ell^{TE} + \frac{L}{\ell}C_\ell^{TE}\left[\sin(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \sin(3\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right]$$

$$f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{EE} = C_\ell^{EE}\left[\frac{d\ln\ell^2 C_\ell^{EE}}{d\ln\ell} + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})\frac{d\ln C_\ell^{EE}}{d\ln\ell}\right] - \frac{1}{2}\frac{L}{\ell}\left[D_\ell^+ C_\ell^{EE} + 4\frac{dC_\ell^{EE}}{d\ln\ell}\cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})D_\ell^- C_\ell^{EE}\right]$$

$$f_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{EB} = 2\sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})C_\ell^{EE} + \frac{L}{\ell}C_\ell^{EE}\left[\sin(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \sin(3\theta_{\boldsymbol{L},\boldsymbol{\ell}}),\right]$$

$$\tag{A4}$$

with corrections going as $\mathcal{O}((L/\ell)^2)$. The colors here denote terms which are orthogonal to the relevant foreground linear response functions to $\mathcal{O}((L/\ell)^2)$ (blue), and those which aren't (red). That is, within the LLS approach we wish to retain sensitivity to the (clean) blue terms, while killing sensitivity to the red.

Likewise, one can Taylor expand the angular dependent pieces of the foreground linear responses (Table I) to give:

$$g^{TT,T}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \mathcal{A}\frac{u_\ell}{u_L}\left\{u_\ell - Lu'_\ell\cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \frac{1}{4}\left(\frac{L}{\ell}\right)^2\left[D^+_\ell + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})D^-_\ell\right]u_\ell\right\}$$

$$g^{TE,E}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \mathcal{B}\frac{u_\ell}{u_L}\left\{u_\ell\cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \frac{L}{\ell}\left[(u_\ell - \ell u'_\ell/2)\cos(3\theta_{\boldsymbol{L},\boldsymbol{\ell}}) - (u_\ell + \ell u'_\ell/2)\cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right]\right.$$
$$\left. + \frac{1}{4}\left(\frac{L}{\ell}\right)^2\left[D^-_\ell u_\ell/2 + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})(D^+_\ell u_\ell - 4u_\ell) + (4u_\ell + D^-_\ell u_\ell/2)\cos(4\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right]\right\}$$

$$g^{TE,B}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = -\mathcal{B}\frac{u_\ell}{u_L}\left\{u_\ell\sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \frac{L}{\ell}\left[(u_\ell - \ell u'_\ell/2)\sin(3\theta_{\boldsymbol{L},\boldsymbol{\ell}}) - (u_\ell + \ell u'_\ell/2)\sin(\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right]\right. \tag{A5}$$
$$\left. + \frac{1}{4}\left(\frac{L}{\ell}\right)^2\left[\sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})(D^+_\ell u_\ell - 4u_\ell) + \sin(4\theta_{\boldsymbol{L},\boldsymbol{\ell}})(4u_\ell + D^-_\ell u_\ell/2)\right]\right\}$$

$$g^{EE,T}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \mathcal{C}\frac{u_\ell}{u_L}\left\{u_\ell - Lu'_\ell\cos(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \frac{1}{4}\left(\frac{L}{\ell}\right)^2\left[(D^+_\ell - 4)u_\ell + \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})(D^-_\ell + 4)u_\ell\right]\right\}$$

$$g^{EB,T}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \mathcal{C}\frac{u_\ell}{u_L}\left\{2\frac{L}{\ell}u_\ell\sin(\theta_{\boldsymbol{L},\boldsymbol{\ell}}) - \left(\frac{L}{\ell}\right)^2\left[\ell u'_\ell + u_\ell\right]\sin(2\theta_{\boldsymbol{L},\boldsymbol{\ell}})\right\},$$

while the remaining two responses satisfy $g^{TB,E}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = -g^{TE,B}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}$ and $g^{TB,B}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = g^{TE,E}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}$. Terms in $\{\cdots\}$'s are accurate to $\mathcal{O}((L/\ell)^3)$.

Appendix B: Explicit expressions for the global minimum variance weights

Let $\boldsymbol{F}^{\mathrm{GMV}}$ denote the minimum variance set of weights (Eq. (10)) for a quadratic estimator taking the form of Eq. (8), where the index i runs over T, E, B at a single frequency:

$$\boldsymbol{F}^{\mathrm{GMV}}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \begin{pmatrix} F^{\mathrm{GMV},TT}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} & F^{\mathrm{GMV},TE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} & F^{\mathrm{GMV},TB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} \\ F^{\mathrm{GMV},TE}_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}} & F^{\mathrm{GMV},EE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} & F^{\mathrm{GMV},EB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} \\ F^{\mathrm{GMV},TB}_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}} & F^{\mathrm{GMV},EB}_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}} & 0 \end{pmatrix}. \tag{B1}$$

Assuming $\tilde{C}^{TB} = \tilde{C}^{EB} = 0$, we find:

$$F^{\mathrm{GMV},TT}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \left[\tilde{C}^{EE}_\ell\tilde{C}^{EE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TT}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} + \tilde{C}^{TE}_\ell\tilde{C}^{TE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{EE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} - \tilde{C}^{EE}_\ell\tilde{C}^{TE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} - \tilde{C}^{TE}_\ell\tilde{C}^{EE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TE}_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}}\right]\frac{1}{2\mathcal{D}_\ell\mathcal{D}_{|\boldsymbol{L}-\boldsymbol{\ell}|}}$$

$$F^{\mathrm{GMV},TE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \left[\tilde{C}^{EE}_\ell\tilde{C}^{TT}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} + \tilde{C}^{TE}_\ell\tilde{C}^{TE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TE}_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}} - \tilde{C}^{TE}_\ell\tilde{C}^{TT}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{EE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} - \tilde{C}^{EE}_\ell\tilde{C}^{TE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TT}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}\right]\frac{1}{2\mathcal{D}_\ell\mathcal{D}_{|\boldsymbol{L}-\boldsymbol{\ell}|}}$$

$$F^{\mathrm{GMV},TB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \left[\tilde{C}^{EE}_\ell f^{TB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} - \tilde{C}^{TE}_\ell f^{EB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}\right]\frac{1}{2\mathcal{D}_\ell\tilde{C}^{BB}_{|\boldsymbol{L}-\boldsymbol{\ell}|}} \tag{B2}$$

$$F^{\mathrm{GMV},EE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \left[\tilde{C}^{TE}_\ell\tilde{C}^{TE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TT}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} + \tilde{C}^{TT}_\ell\tilde{C}^{TT}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{EE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} - \tilde{C}^{TE}_\ell\tilde{C}^{TT}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TE}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} - \tilde{C}^{TT}_\ell\tilde{C}^{TE}_{|\boldsymbol{L}-\boldsymbol{\ell}|}f^{TE}_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}}\right]\frac{1}{2\mathcal{D}_\ell\mathcal{D}_{|\boldsymbol{L}-\boldsymbol{\ell}|}}$$

$$F^{\mathrm{GMV},EB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} = \left[\tilde{C}^{TT}_\ell f^{EB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}} - \tilde{C}^{TE}_\ell f^{TB}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}\right]\frac{1}{2\mathcal{D}_\ell\tilde{C}^{BB}_{|\boldsymbol{L}-\boldsymbol{\ell}|}}$$

where we have defined $\mathcal{D}_\ell \equiv \tilde{C}^{TT}_\ell\tilde{C}^{EE}_\ell - (\tilde{C}^{TE}_\ell)^2$. We note that our convention of summing over both permutations of pairs, e.g. both TE and ET, differs from that used in [23]. As a result our TE, TB and EB weights differ by a conventional factor of 2. Additionally we find that the last term of $F^{\mathrm{GMV},TE}$ differs from that in [23], where the arguments $\boldsymbol{\ell}$ and $\boldsymbol{L}-\boldsymbol{\ell}$ have been flipped.

Appendix C: Cleaning combined estimators

Here we explore the feasibility of cleaning the GMV estimator via bias hardening and large-scale suppression. We find that for both cases one should in practice take an optimal linear combination of the relevant single-pair estimators,

rather than to harden (or suppress) a GMV estimator.

1. Bias hardening

In section III C we found that the optimal temperature-only bias hardened estimator is insensitive to the amplitude $\mathcal{A}$ of the foregrounds. Here we show that this is always be the case when bias hardening a single-pair lensing estimator.

Proof: In general, the optimal bias hardened weights for an estimator constructed from a single pair X, Y takes the (schematic) form:

$$F^{XY} + \alpha^T G^{XY,T} + \alpha^E G^{XY,E} + \alpha^B G^{XY,B}, \tag{C1}$$

where F^{XY} is the MV lensing weight and $G^{XY,Z}$ is the MV weight for the foreground in the Z'th map. WLOG we will assume that the weights $G^{XY,Z} \propto g^{XY,Z}$ have been rescaled to remove the unknown prefactor ($\mathcal{A}$, $\mathcal{B}$ or $\mathcal{C}$) in the foreground response g, which can be thought of as being absorbed into coefficients ($\alpha^T, \alpha^E, \alpha^B$). These coefficients are chosen so that the bias hardened estimator has zero response to s^T, s^E and s^B. This is enforced by integrating Eq. (C1) against $g^{XY,T}$, and forcing this integral to vanish (and similarly for $g^{XY,E}$ and $g^{XY,B}$). Since these integral constraints are set to zero, we are always free to divide out the the unknown prefactor introduced by the foreground response $g^{XY,Z}$ in all three constraint equations. We can then solve for the coefficients ($\alpha^T, \alpha^E, \alpha^B$) without any knowledge of ($\mathcal{A}, \mathcal{B}, \mathcal{C}$). Thus Eq. (C1) is independent of ($\mathcal{A}, \mathcal{B}, \mathcal{C}$). $\square$

This is no longer true, however, when bias hardening the GMV estimator. In this case, the optimal bias hardened weights take the same form as Eq. (C1), but with F^{XY} replaced by the GMV lensing weights, and similarly for the foreground weights (e.g. the GMV weights for s^T takes the same form as Eq. (10), but with f^{ij} replaced by $g^{XY,T}$, see Appendix D). From Table I note that the GMV weights for s^E and s^B are proportional to $\mathcal{B}$. However, the GMV weights for s^T contain a linear combination of $\mathcal{A}$ and $\mathcal{C}$. Consequentially, one cannot bias harden the GMV κ estimator without assuming the ratio $\mathcal{C}/\mathcal{A} \simeq \langle p_i^2 \rangle / 2$.

To avoid assuming this ratio, one can simply take an optimal linear combination of the single-pair bias hardened estimators, which takes the form $\sum_i w_L^i \hat{\kappa}_L^{\mathrm{BH},i}$, where $i \in \{TT, TE, TB, EE, EB\}$ and $w_i = \sum_j C_{ji}^{-1} / \sum_{ij} C_{ij}^{-1}$, where C_{ij} is the covariance between $\hat{\kappa}^{\mathrm{BH},i}$ and $\hat{\kappa}^{\mathrm{BH},j}$:

$$C_L^{XY,MN} = N_L^{XY} N_L^{MN} \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{XY} \left[F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{MN} \tilde{C}_\ell^{XM} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{YN} + F_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}}^{MN} \tilde{C}_\ell^{XN} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{YM} \right]. \tag{C2}$$

The noise of this optimal linear combination is given by $1/\sum_{ij} C_{ij}^{-1}$.

This simple linear combination of bias hardened estimators is very close to optimal (while remaining practical) for the following reason. The GMV estimator can be separated into two statistically independent GMV estimators constructed from $\{TT, TE, EE\}$ and $\{TB, EB\}$ [23]. Following the same argument as above one cannot bias harden the former without assuming the ratio $\mathcal{C}/\mathcal{A}$. To get around this, one can further decompose the set into $\{TT, TE\}$ and $\{EE\}$. In practice the GMV estimator constructed from $\{TT, TE\}$ is not FFT-able. To make this estimator FFT-able one can neglect the cross-correlation $\tilde{C}^{TE}$, however, in this approximation the $\{TT, TE\}$ GMV estimator reduces to the optimal linear combination of $\hat{\kappa}^{TT}$ and $\hat{\kappa}^{EE}$. Thus, to bias harden the $\{TT, TE, EE\}$ GMV estimator in an FFT-able way without assuming $\mathcal{C}/\mathcal{A}$ one must fully decompose $\{TT, TE, EE\}$ into an optimal linear combination of TT, TE and EE.

While one can in principle bias harden the GMV estimator constructed from $\{TB, EB\}$ in practice the gains for doing so (as opposed to an optimal linear combination) are expected to be minimal. We have checked that the reconstruction noise for the GMV constructed from $\{TB, EB\}$ and the optimal linear combination of TB and EB differ by less than half a percent for $L < 5000$.

2. Large-lens suppression

One can generalize the single-pair LLS estimator to a global form in a straightforward way. However, as we show below, we find that an optimal linear combination of single-pair LLS estimators is practically more useful.

From Eqs. (33) and (34) a natural ansatz for the global LLS weights is:

$$\boldsymbol{F}_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{LLS}} = \begin{pmatrix} A_{L,\ell} \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) & B_{L,\ell} & 0 \\ B_{L,|\boldsymbol{L}-\boldsymbol{\ell}|} & C_{L,\ell} \cos(2\theta_{\boldsymbol{L},\boldsymbol{\ell}}) & F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{MV},EB} \\ 0 & F_{\boldsymbol{L}-\boldsymbol{\ell},\boldsymbol{\ell}}^{\mathrm{MV},EB} & 0 \end{pmatrix}, \tag{C3}$$

with corrections going as $\mathcal{O}(L/\ell)$, where the matrix is in the basis $\{T, E, B\}$. The coefficients (A, B, C) should be chosen to minimize Eq. (9) in the large-lens limit:

$$
\mathcal{N}_{\mathbf{0}}^{\kappa} = 2(N_{\mathbf{0}}^{\kappa})^2 \int \ell \, d\ell \left[\frac{1}{2}(\tilde{C}_{\ell}^{TT})^2 A_{0,\ell}^2 + \frac{1}{2}(\tilde{C}_{\ell}^{EE})^2 C_{0,\ell}^2 + 2((\tilde{C}_{\ell}^{TE})^2 + \tilde{C}_{\ell}^{TT}\tilde{C}_{\ell}^{EE})B_{0,\ell}^2 + (\tilde{C}_{\ell}^{TE})^2 A_{0,\ell}C_{0,\ell} \right]
$$
$$
+ 2(N_{\mathbf{0}}^{\kappa})^2 \int_{\boldsymbol{\ell}} 2\tilde{C}_{\ell}^{EE}\tilde{C}_{\ell}^{BB}(F_{\boldsymbol{\ell},-\boldsymbol{\ell}}^{\mathrm{MV},EB})^2,
\tag{C4}
$$

where

$$
(N_{\mathbf{0}}^{\kappa})^{-1} = \int_{\boldsymbol{\ell}} 2f_{\boldsymbol{\ell},-\boldsymbol{\ell}}^{EB} F_{\boldsymbol{\ell},-\boldsymbol{\ell}}^{\mathrm{MV},EB} + \int \ell \, d\ell \left[\frac{1}{2}\frac{dC_{\ell}^{EE}}{d\ln\ell}C_{0,\ell} + 2C_{\ell}^{TE}\frac{d\ln\ell^2 C_{\ell}^{TE}}{d\ln\ell}B_{0,\ell} + \frac{1}{2}\frac{dC_{\ell}^{TT}}{d\ln\ell}A_{0,\ell}. \right]
\tag{C5}
$$

By varying A, B, C individually and setting $\delta\mathcal{N}^{\kappa} = 0$ we find:

$$
A_{L,\ell} = \frac{(\tilde{C}_{\ell}^{EE})^2\frac{dC_{\ell}^{TT}}{d\ln\ell} - (\tilde{C}_{\ell}^{TE})^2\frac{dC_{\ell}^{EE}}{d\ln\ell}}{2\left[(\tilde{C}_{\ell}^{TT})^2(\tilde{C}_{\ell}^{EE})^2 - (\tilde{C}_{\ell}^{TE})^4\right]} \qquad C_{L,\ell} = \frac{(\tilde{C}_{\ell}^{TT})^2\frac{dC_{\ell}^{EE}}{d\ln\ell} - (\tilde{C}_{\ell}^{TE})^2\frac{dC_{\ell}^{TT}}{d\ln\ell}}{2\left[(\tilde{C}_{\ell}^{TT})^2(\tilde{C}_{\ell}^{EE})^2 - (\tilde{C}_{\ell}^{TE})^4\right]}
$$
$$
B_{L,\ell} = \frac{C_{\ell}^{TE}\frac{d\ln\ell^2|C_{\ell}^{TE}|}{d\ln\ell}}{2\left[\tilde{C}_{\ell}^{TT}\tilde{C}_{\ell}^{EE} + (\tilde{C}_{\ell}^{TE})^2\right]}.
\tag{C6}
$$

up to $\mathcal{O}(L/\ell)$ corrections.

As shown in Fig. 6, the global LLS estimator (GLLS) negligibly improves upon the optimal linear combination of single-pair LLS estimators at large scales. At small scales where the $L/\ell \ll 1$ approximation breaks down the GLLS estimator is no longer guaranteed to be optimal. In fact we find that the the optimal linear combination of LLS estimators outperforms (or is nearly identical to) the GLLS for all relevant scales.
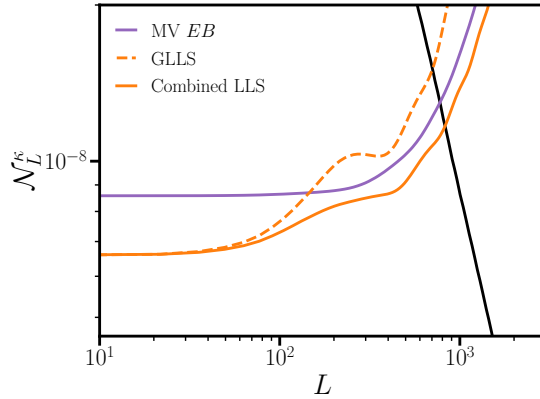

Figure 6. Lensing reconstruction noise for the optimal linear combination of single-pair LLS estimators (solid orange) and for the Global LLS estimator (GLLS, dashed orange). For reference we show the reconstruction noise for the MV EB estimator (purple) and the convergence power spectrum (black).

Appendix D: Bias hardening as constrained minimization

Here we will derive an optimal bias hardened estimator using Lagrange multipliers. Let's suppose that we have a foreground s^i that we wish to harden a generic quadratic estimator (Eq. (8)) against, with the relevant linear responses given by $g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij,k}$. To construct an optimal bias hardened estimator, we choose our lensing weights F^{ij} to minimize the noise of our lensing estimator under the constraint of having no response to the foreground. That is, we wish to minimize the Lagrangian:

$$
\mathcal{L}_{\boldsymbol{L}}[F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{ij}, \lambda_{\boldsymbol{L}}^i] = \mathcal{N}_{\boldsymbol{L}}^{\kappa} + 2\sum_i \lambda_{\boldsymbol{L}}^i \int_{\boldsymbol{\ell}} F_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk} g_{\boldsymbol{\ell},\boldsymbol{L}-\boldsymbol{\ell}}^{jk,i}
\tag{D1}
$$

where λ_L^i are some Lagrange multipliers. We first vary the weights, which gives

$$\delta \mathcal{L}_L = 4(N_L^\kappa)^2 \sum_{mn} \int_{\boldsymbol{\ell}} \left[\sum_{ij} F_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{ij} \tilde{C}_{\ell}^{im} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{jn} - \frac{1}{2} \frac{\mathcal{N}_L^\kappa}{N_L^\kappa} f_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{mn} + \frac{1}{2} \sum_k \sigma_L^k g_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{mn,k} \right] \delta F_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{mn} \tag{D2}$$

where $\sigma_L^k = \lambda_L^k / (N_L^\kappa)^2$ are a renormalized set of Lagrange multipliers. After setting $\delta \mathcal{L}_L = 0$ and fixing the normalization of the weights ($\mathcal{N}_L^\kappa = N_L^\kappa$), we find that the optimal zero-response weights satisfy

$$\begin{aligned} F_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{ij} &= \frac{1}{2} \sum_{mn} \left(\tilde{C}_\ell^{-1} \right)^{im} \left(\tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{-1} \right)^{jn} \left[f_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{mn} - \sum_k \sigma_L^k g_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{mn,k} \right] \\ &= F_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{\mathrm{GMV}, ij} - \sum_k \sigma_L^k G_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{ij,k} \end{aligned} \tag{D3}$$

where $F^{\mathrm{GMV}, ij}$ are the weights for the GMV κ estimator and $G^{ij,k}$ are the GMV weights for s^k. That is, the true minimum variance estimator with zero response to the foregrounds is a some linear combination of the GMV estimators for κ and s^i. The linear combination which nulls the response of $\hat{\kappa}$ to s^i is unique, given by Eq. (15).

This approach can in principle be extended to null not only the field-level bias (and hence the primary and tripectrum biases, see Section VI), but also the secondary bias. Consider the temperature-only case for simplicity. In this scenario the secondary bias is proportional to:

$$\mathrm{Secondary} \propto \int_{\boldsymbol{\ell}\boldsymbol{\ell}'} F_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{TT} F_{\boldsymbol{\ell}', -\boldsymbol{L}-\boldsymbol{\ell}'}^{TT} f_{\boldsymbol{\ell}, -\boldsymbol{L}-\boldsymbol{\ell}'}^{TT} g_{\boldsymbol{\ell}', \boldsymbol{L}-\boldsymbol{\ell}}^{TT,T}. \tag{D4}$$

We null the secondary by adding a second term to the Lagrangian, yielding:

$$F_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{TT} = \frac{f_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{TT}}{2\tilde{C}_\ell^{TT} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}} + \sigma_L \frac{g_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{TT,T}}{2\tilde{C}_\ell^{TT} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}} + \frac{\lambda_L}{2\tilde{C}_\ell^{TT} \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}} \int_{\boldsymbol{\ell}'} F_{\boldsymbol{\ell}', -\boldsymbol{L}-\boldsymbol{\ell}'}^{TT} f_{\boldsymbol{\ell}, -\boldsymbol{L}-\boldsymbol{\ell}'}^{TT} g_{\boldsymbol{\ell}', \boldsymbol{L}-\boldsymbol{\ell}}^{TT,T} \tag{D5}$$

where σ and λ are chosen to null the field-level and secondary biases respectively. In practice one would have to solve for these weights numerically, which could potentially be accomplished in an iterative fashion with the bias-hardened weights as an initial guess.

Appendix E: Derivation of polarization-based LLS estimators

In a similar manner to IV we expand our lensing weights into multipoles:

$$F_{\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}}^{XY} = \sum_{m=0}^\infty A_{L,\ell}^{(m)} \cos(m\theta_{\boldsymbol{L},\boldsymbol{\ell}}) + \sum_{m=1}^\infty B_{L,\ell}^{(m)} \sin(m\theta_{\boldsymbol{L},\boldsymbol{\ell}}), \tag{E1}$$

The noise of the lensing estimator estimator in the large lens limit is then

$$\mathcal{N}_0 \propto \frac{\int \ell \, d\ell \, \tilde{C}_\ell^{XX} \tilde{C}_\ell^{YY} \left[2\left(A_{0,\ell}^{(0)}\right)^2 + \sum_{m=1}^\infty \left(A_{0,\ell}^{(m)}\right)^2 + \sum_{m=1}^\infty \left(B_{0,\ell}^{(m)}\right)^2 \right]}{\left[\sum_{m=0}^\infty A_{0,\ell}^{(m)} f_{c,\ell}^{XY,m} + \sum_{m=1}^\infty B_{0,\ell}^{(m)} f_{s,\ell}^{XY,m} \right]^2}. \tag{E2}$$

where we have neglected the cross-correlation $\tilde{C}^{XY}$, and we have defined

$$f_{c,\ell}^{XY,m} = \int \frac{d\theta_{\boldsymbol{L},\boldsymbol{\ell}}}{2\pi} \cos(m\theta_{\boldsymbol{L},\boldsymbol{\ell}}) f_{\boldsymbol{\ell}, -\boldsymbol{\ell}}^{XY} \qquad f_{s,\ell}^{XY,m} = \int \frac{d\theta_{\boldsymbol{L},\boldsymbol{\ell}}}{2\pi} \sin(m\theta_{\boldsymbol{L},\boldsymbol{\ell}}) f_{\boldsymbol{\ell}, -\boldsymbol{\ell}}^{XY}. \tag{E3}$$

Varying $A^{(m)}$, $B^{(m)}$ individually and setting $\delta \mathcal{N}_0 = 0$ then gives

$$A_{L,\ell}^{(0)} \propto \frac{f_{c,\ell}^{XY,0}}{2\tilde{C}_\ell^{XX} \tilde{C}_\ell^{YY}} \qquad A_{L,\ell}^{(m)} \propto \frac{f_{c,\ell}^{XY,m}}{\tilde{C}_\ell^{XX} \tilde{C}_\ell^{YY}} \qquad B_{L,\ell}^{(m)} \propto \frac{f_{s,\ell}^{XY,m}}{\tilde{C}_\ell^{XX} \tilde{C}_\ell^{YY}}, \tag{E4}$$

up to $\mathcal{O}(L/\ell)$, where the (common) proportionality constant is any function of L. That is, we can read off the optimal values of $A^{(m)}, B^{(m)}$ directly from the multipole expansion of the linear response f^{XY} in the large-lens limit. To

suppress the response to foregrounds we will manually set the coefficients which are sensitive to foregrounds to zero. We do so below for each single-pair estimator, making frequent use of Eqs. (A4) and (A5).

TE: Note that $g^{TE,E}$ contains a $\cos(2\theta)$ term which is $\mathcal{O}(1)$, while $g^{TE,B}$ has a $\mathcal{O}(1)\sin(2\theta)$ term. We therefore manually set $A^{(2)} = B^{(2)} = 0$ to suppress responses to s^E and s^B. Both $g^{TE,E}$ and $g^{TE,E}$ only have a monopole at $\mathcal{O}((L/\ell)^2)$. Thus we can retain the monopole $A^{(0)}$, whose optimal value takes the form

$$A^{(0)}_{L,\ell} = \frac{C^{TE}_\ell}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_\ell} \frac{d\ln \ell^2 |C^{TE}_\ell|}{d\ln \ell}, \tag{E5}$$

where we have made us of Eqs. (E4) and (A4).

TB: Likewise $g^{TB,E}$ contains a $\mathcal{O}(1)\sin(2\theta)$ term while $g^{TE,B}$ has a $\mathcal{O}(1)\cos(2\theta)$ term. We again set $A^{(2)} = B^{(2)} = 0$ to suppress responses to s^E and s^B. Both $g^{TE,E}$ and $g^{T\tilde{E},E}$ only have a monopole at $\mathcal{O}((L/\ell)^2)$. Thus we can retain the monopole $A^{(0)}$, however, the response f^{TB} only has a $\sin(2\theta)$ term to leading order, and so we set $A^{(0)} = 0$. There is no LLS TB estimator to leading order.

EE: This case is identical to TT. We only retain the $\cos(2\theta)$ term.

EB: Note that $F^{\mathrm{MV},EB} = f^{EB}/\tilde{C}^{EE}\tilde{C}^{BB}$ only has a $\sin(2\theta)$ term at $\mathcal{O}(1)$, while at $\mathcal{O}(L/\ell)$ it has a $\cos(\theta)$, $\sin(\theta)$ and $\sin(3\theta)$ term. Note also that $g^{EB,T}$ only has a $\sin(\theta)$ term at $\mathcal{O}(L/\ell)$. Thus the foreground response of the MV EB estimator is already suppressed by $\mathcal{O}((L/\ell)^2)$.

Appendix F: Worked example: explicitly hardening the TE estimator

From Table I we see that the TE has a non-zero linear response to s^E and s^B in addition to lensing. Thus to bias harden $\hat\kappa$ against s^E and s^B we must construct (optimal) estimators for κ, s^E and s^B, which take the form:

$$\hat\kappa_{\boldsymbol L} = N^\kappa_{\boldsymbol L} \int_{\boldsymbol\ell} \frac{f^{TE}_{\boldsymbol\ell,\boldsymbol L-\boldsymbol\ell}}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \tilde T_{\boldsymbol\ell} \tilde E_{\boldsymbol L-\boldsymbol\ell} \quad \text{where} \quad (N^\kappa_{\boldsymbol L})^{-1} = \int_{\boldsymbol\ell} \frac{(f^{TE}_{\boldsymbol\ell,\boldsymbol L-\boldsymbol\ell})^2}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}}$$

$$\hat s^E_{\boldsymbol L} = N^E_{\boldsymbol L} \int_{\boldsymbol\ell} \frac{\cos(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \tilde T_{\boldsymbol\ell} \tilde E_{\boldsymbol L-\boldsymbol\ell} \quad \text{where} \quad (N^E_{\boldsymbol L})^{-1} = \int_{\boldsymbol\ell} \frac{\cos^2(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \tag{F1}$$

$$\hat s^B_{\boldsymbol L} = N^B_{\boldsymbol L} \int_{\boldsymbol\ell} \frac{\sin(2\theta_{\boldsymbol L-\boldsymbol\ell,\boldsymbol L})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \tilde T_{\boldsymbol\ell} \tilde E_{\boldsymbol L-\boldsymbol\ell} \quad \text{where} \quad (N^B_{\boldsymbol L})^{-1} = \int_{\boldsymbol\ell} \frac{\sin^2(2\theta_{\boldsymbol L-\boldsymbol\ell,\boldsymbol L})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}},$$

where $\tilde{C}^{XX}_\ell$ is the total power spectrum, including the lensed field $\tilde X$ as well as instrumental noise and foregrounds. Here we have neglected the cross-correlation between T and E in the minimum variance weights to make the estimators FFT-able, and we have assumed that the foreground is composed of point sources ($u_\ell = 1$). Note that we have neglected the normalization $\mathcal{B}$ of the foregrounds. This makes $\hat s^E$ an estimator of $\mathcal{B} s^E$, and similarly for $\hat s^B$. As we show in Appendix C 1, the bias hardened TE lensing estimator is insensitive to $\mathcal{B}$, which we will also show below explicitly.

The field-level biases to the minimum variance estimators above take form:

$$\begin{pmatrix} \langle \hat\kappa_{\boldsymbol L} \rangle' \\ \langle \hat s^E_{\boldsymbol L} \rangle' \\ \langle \hat s^B_{\boldsymbol L} \rangle' \end{pmatrix} = \begin{pmatrix} 1 & \mathcal{R}^{\kappa E}_{\boldsymbol L} & \mathcal{R}^{\kappa B}_{\boldsymbol L} \\ \mathcal{R}^{E\kappa}_{\boldsymbol L} & 1 & \mathcal{R}^{EB}_{\boldsymbol L} \\ \mathcal{R}^{B\kappa}_{\boldsymbol L} & \mathcal{R}^{BE}_{\boldsymbol L} & 1 \end{pmatrix} \begin{pmatrix} \kappa_{\boldsymbol L} \\ \mathcal{B}\, s^E_{\boldsymbol L} \\ \mathcal{B}\, s^B_{\boldsymbol L} \end{pmatrix}, \tag{F2}$$

where we have defined six response functions:

$$\mathcal{R}^{\kappa E}_{\boldsymbol L} = N^\kappa_{\boldsymbol\ell} \int_{\boldsymbol\ell} \frac{f^{TE}_{\boldsymbol\ell,\boldsymbol L-\boldsymbol\ell} \cos(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \qquad \mathcal{R}^{E\kappa}_{\boldsymbol L} = N^E_{\boldsymbol\ell} \int_{\boldsymbol\ell} \frac{f^{TE}_{\boldsymbol\ell,\boldsymbol L-\boldsymbol\ell} \cos(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}}$$

$$\mathcal{R}^{\kappa B}_{\boldsymbol L} = N^\kappa_{\boldsymbol\ell} \int_{\boldsymbol\ell} \frac{f^{TE}_{\boldsymbol\ell,\boldsymbol L-\boldsymbol\ell} \sin(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \qquad \mathcal{R}^{B\kappa}_{\boldsymbol L} = N^B_{\boldsymbol\ell} \int_{\boldsymbol\ell} \frac{f^{TE}_{\boldsymbol\ell,\boldsymbol L-\boldsymbol\ell} \sin(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \tag{F3}$$

$$\mathcal{R}^{EB}_{\boldsymbol L} = N^E_{\boldsymbol\ell} \int_{\boldsymbol\ell} \frac{\cos(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell}) \sin(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}} \qquad \mathcal{R}^{BE}_{\boldsymbol L} = N^B_{\boldsymbol\ell} \int_{\boldsymbol\ell} \frac{\cos(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell}) \sin(2\theta_{\boldsymbol L,\boldsymbol L-\boldsymbol\ell})}{\tilde{C}^{TT}_\ell \tilde{C}^{EE}_{|\boldsymbol L-\boldsymbol\ell|}}.$$

The bias hardened estimators are simply:

$$\begin{pmatrix} \hat\kappa^{\mathrm{BH}}_{\boldsymbol L} \\ \hat s^{\mathrm{BH},E}_{\boldsymbol L} \\ \hat s^{\mathrm{BH},B}_{\boldsymbol L} \end{pmatrix} = \begin{pmatrix} 1 & \mathcal{R}^{\kappa E}_{\boldsymbol L} & \mathcal{R}^{\kappa B}_{\boldsymbol L} \\ \mathcal{R}^{E\kappa}_{\boldsymbol L} & 1 & \mathcal{R}^{EB}_{\boldsymbol L} \\ \mathcal{R}^{B\kappa}_{\boldsymbol L} & \mathcal{R}^{BE}_{\boldsymbol L} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat\kappa_{\boldsymbol L} \\ \hat s^E_{\boldsymbol L} \\ \hat s^B_{\boldsymbol L} \end{pmatrix}. \tag{F4}$$

Appendix G: Expected biases to Simons Observatory

In Fig. 7 we show the expected sensitivity to the lensing amplitude, and the bias from radio point sources, for Simons Observatory-like sensitivity. The assumptions regarding the experimental setup and simulations are identical to those in the main text, but with $\Delta_T = 6$ μK-arcmin instead of $\Delta_T = 1$ μK-arcmin. We see that for polarization-based reconstruction the point source-induced bias is negligible.
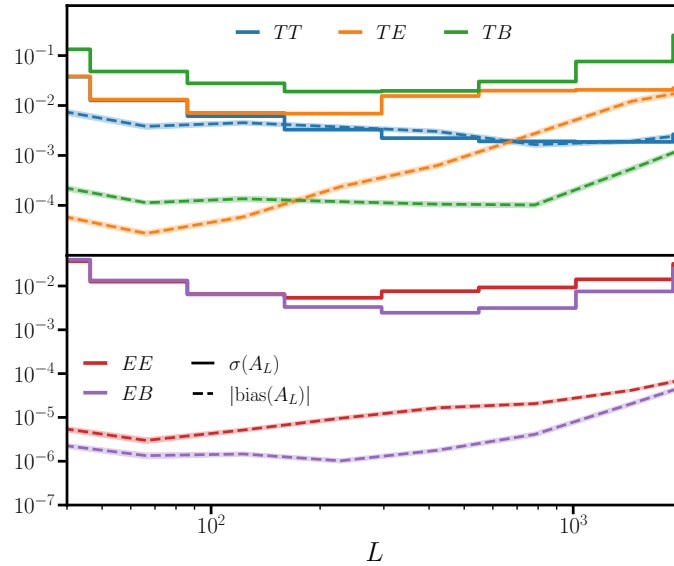

Figure 7. The same as the top two panels of Fig. 4, but with Simons Observatory-like sensitivity instead of CMB-S4.