

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

Bias to CMB lensing from lensed foregrounds

Nishant Mishra and Emmanuel Schaan

Phys. Rev. D **100**, 123504 — Published 3 December 2019

DOI: [10.1103/PhysRevD.100.123504](https://doi.org/10.1103/PhysRevD.100.123504)

Bias to CMB lensing from lensed foregrounds

Nishant Mishra^{1,*} and Emmanuel Schaan^{2,3,†}

¹*Department of Physics, University of California, Berkeley, CA 94720, USA*

²*Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA*

³*Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA*

Extragalactic foregrounds are known to constitute a limiting systematic in temperature-based CMB lensing with AdvACT, SPT-3G, Simons Observatory and CMB S4. Furthermore, since these foregrounds are emitted at cosmological distances, they are also themselves lensed. The correlation between this foreground lensing and CMB lensing causes an additional bias in CMB lensing estimators. In this paper, we quantify for the first time this “lensed foreground bias” for the standard CMB lensing quadratic estimator, the CMB shear and the CMB magnification estimators, in the case of Simons Observatory and in the absence of multi-frequency component separation. This percent-level bias is highly significant in cross-correlation of CMB lensing with LSST galaxies, and comparable to the statistical **uncertainty** in CMB lensing auto-spectrum. We discuss various mitigation strategies, and show that “lensed foreground bias-hardening” methods can reduce this bias at some cost in signal-to-noise. The code used to generate our theory curves is publicly available¹.

PACS numbers:

I. INTRODUCTION

Gravitational lensing of the cosmic microwave background (CMB) probes the projected mass distribution in the Universe, all the way to the surface of last scattering. This effect is measured at high significance from the WMAP satellite [1, 2], the Atacama Cosmology Telescope (ACT) [3–8], the South Pole Telescope (SPT) [9–11, 11–13], POLARBEAR [14–16], the Planck satellite [17–19] and BICEP 2 / Keck Array [20]. These measurements contain valuable information on the nature of dark energy and the masses of the neutrinos, through their effect on the growth of structure and the expansion history in the universe. Via delensing [21–23] of the CMB polarization B-modes, these lensing measurements will also be crucial in the search for primordial gravitational waves.

As the statistical signal-to-noise (SNR) in upcoming CMB lensing detections increases with AdvACT [24], SPT-3G [25], Polarbear-2 and Simons Array [26], Simons Observatory [27] and CMB S4 [28], a similar improvement in systematics control becomes necessary. In temperature-based CMB lensing reconstruction, extragalactic foregrounds such as the cosmic infrared background (CIB), the thermal and kinematic Sunyaev-Zel’dovich (tSZ and kSZ) effects and radio point sources (radio PS) constitute the main limiting systematics. If not accounted for, they are known to produce highly significant biases to CMB lensing [29–33]. These biases arise from the non-Gaussian statistics of these foregrounds, and their correlation with CMB lensing. Since these foregrounds dominate on small scales, they limit the range of temperature multipoles that can be used for lensing reconstruction to $\ell < 3000 - 3500$. Various mitigation methods exist. For quadratic CMB lensing estimators, the lensing field is reconstructed from two powers of the temperature map, one “gradient leg” and one small-scale leg. Mitigation methods typically aim at removing the foreground from one or two of these powers of the map. Masking removes the brightest point sources in the temperature map, and is an effective way to control radio PS [30]. Multi-frequency component separation [34] or the cleaned gradient method [35] can reduce the amplitude of CIB and tSZ at some cost in noise, but cannot reduce the kSZ bias. In the case of halo lensing, the gradient can also be foreground-cleaned by inpainting a Gaussian CMB realization at the localization of the halo [36]. In some cases, the foregrounds are left intact in the two legs of the quadratic estimator, and the mitigation occurs by nulling the overall response of the estimator to the foreground. “Bias-hardened” estimators reduce the contamination from foregrounds whose statistics is known (e.g., Poisson) [29, 37, 38]. Splitting the CMB lensing quadratic estimator into magnification-only and shear-only estimators [33, 39, 40] provides a useful foreground null test. Discarding the magnification-only part appears to significantly reduce the contamination from all extragalactic foregrounds, at a cost in SNR which can be compensated by including smaller temperature multipoles in the shear estimator [33]. The “multipole estimators”, which generalize this decomposition, are expected

¹ <https://github.com/EmmanuelSchaan/LensedForegroundBias>

*Electronic address: nishant.mishra@berkeley.edu

†Electronic address: eschaan@lbl.gov

to share the same property [33]. These various methods are based on different principles. Multi-frequency foreground cleaning and the cleaned gradient method partially remove the foreground from the temperature map (or one of the two, for the cleaned gradient method), so that the foregrounds do not enter the lensing estimator or do not bias it. Masking works because a large part of the non-Gaussianity of some foregrounds (radio PS, CIB, tSZ) comes from localized peaks associated with individual galaxies or halos. Masking or inpainting these foreground peaks does not completely remove the foreground from the temperature map, but it reduces its non-Gaussianity (bispectrum and trispectrum). With bias hardening and the shear/multipole estimators, the foreground is left intact in the temperature map and fed to the lensing estimator. However, the estimator is modified so as to be insensitive to the foreground bispectrum and trispectrum. The best solution to the lensing biases from the non-Gaussianity of extragalactic foregrounds likely involves combining these different methods.

However, extragalactic foregrounds produce another bias to CMB lensing, discussed in [41] but not yet been quantified, to the best of our knowledge. This bias is not a consequence of the non-Gaussianity of the foregrounds, but of the fact that extragalactic foregrounds are themselves lensed, by an amount correlated with CMB lensing [41–43]. Indeed, extragalactic foregrounds are emitted at cosmological distances, and should therefore be distorted by lensing, just like the CMB and galaxy shapes are. CMB lensing quadratic estimators will generically reconstruct not only the lensing of the CMB, but also the lensing of the foregrounds [42]. The correlation of CMB lensing and foreground lensing enhances this bias to CMB lensing. This bias is present even if the extragalactic foregrounds are perfectly Gaussian, as long as they are lensed by an amount correlated with CMB lensing. In this paper, we quantify this bias for the first time, and show that it is significant for an experiment like Simons Observatory. This bias will be even larger for the CMB S4 temperature-based lensing. However, the polarization-based quadratic estimators will carry a larger weight for CMB S4, and extragalactic foregrounds are expected to be less important there.

II. REVIEW: CMB LENSING RECONSTRUCTION WITH QUADRATIC ESTIMATORS

A. CMB lensing convergence

We consider the CMB to be emitted at a single redshift at the last scattering surface. For a source image at a single redshift and in the Born approximation, the lensing convergence is obtained as

$$\kappa_{\mathbf{L}}^{\text{CMB}} = \int_0^{\chi_{\text{CMB}}} d\chi W^\kappa(\chi, \chi_{\text{CMB}}) \delta_m(\vec{k} = \mathbf{L}/\chi, \chi), \quad (1)$$

where $W^\kappa(\chi, \chi_S)$ is the lensing kernel for a source image at distance χ_S :

$$W^\kappa(\chi, \chi_S) = \frac{3}{2} \left(\frac{H_0}{c} \right)^2 \Omega_m^0 \frac{\chi}{a(\chi)} (1 - \chi/\chi_S). \quad (2)$$

In the Limber and flat sky approximations, the CMB lensing power spectrum thus becomes:

$$C_L^{\kappa_{\text{CMB}}} = \int \frac{d\chi}{\chi^2} W^\kappa(\chi, \chi_{\text{CMB}})^2 P_m \left(k = \frac{L+1/2}{\chi}, z(\chi) \right), \quad (3)$$

where P_m is the nonlinear matter power spectrum, computed using the `Halofit` [44] implementation in `CLASS` [45].

B. Quadratic CMB lensing estimators

Throughout this paper, we adopt the flat sky approximation, and decompose the various maps (CMB, foreground, convergence) in Fourier modes rather than spherical harmonics. We denote by T^0 an unlensed map (CMB or foreground), and T the corresponding lensed map. Lensing produces off-diagonal correlations in the observed temperature map:

$$\langle T_{\ell} T_{\mathbf{L}-\ell} \rangle = f_{\ell, \mathbf{L}-\ell} \kappa_{\mathbf{L}} + \mathcal{O}(\kappa^2), \quad (4)$$

where the response function $f_{\ell, \mathbf{L}-\ell}$ is completely determined by the unlensed power spectrum:

$$f_{\ell_1, \ell_2} \equiv \left(\frac{2\mathbf{L}}{L^2} \right) \cdot [\ell_1 C_{\ell_1}^0 + \ell_2 C_{\ell_2}^0]. \quad (5)$$

In particular, we shall call this response f^{CMB} when C^0 is the unlensed CMB power spectrum, f^f when C^0 is the unlensed foreground power spectrum, and $f^{ff'}$ when C^0 is the cross-spectrum of two correlated unlensed foregrounds, such as CIB and tSZ.

This coupling of Fourier modes is used to construct unbiased CMB lensing quadratic estimators as

$$\mathcal{Q}_{\mathbf{L}}[T, T] = N_{\mathbf{L}} \int \frac{d^2\ell}{(2\pi)^2} T_{\ell} T_{\mathbf{L}-\ell} F_{\ell, \mathbf{L}-\ell}, \quad (6)$$

where the weight function $F_{\ell, \mathbf{L}-\ell}$ can in principle be chosen arbitrarily, and the corresponding normalization is then fixed in order to obtain unit response to CMB lensing:

$$N_{\mathbf{L}} = \left(\int \frac{d^2\ell}{(2\pi)^2} F_{\ell, \mathbf{L}-\ell} f_{\ell, \mathbf{L}-\ell}^{\text{CMB}} \right)^{-1}. \quad (7)$$

In particular, the response function $f_{\ell, \mathbf{L}-\ell}^{\text{CMB}}$ is that of the CMB, so that the estimator has unit response to lensing when applied to lensed CMB maps. In practice, the weights $F_{\ell, \mathbf{L}-\ell}$ are typically chosen so as to minimize the variance of the estimator [46]. These weights can be chosen differently, for example to null the response to point sources or the survey mask as in “bias hardening” [18, 19, 37, 38, 47], to discard the information in shear or magnification [33, 39, 40, 48–51], or to discard contaminated modes in one of the legs of the estimator [35]. In this paper, we will consider the following quadratic estimators:

$$F_{\ell, \mathbf{L}-\ell} = \begin{cases} \frac{f_{\ell, \mathbf{L}-\ell}^{\text{CMB}}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}} & \text{QE} \\ \frac{C_{\ell}^{\text{CMB}}}{2(C_{\ell}^{\text{total}})^2} \frac{d \ln C_{\ell}^{\text{CMB}}}{d \ln \ell} \cos(2\theta_{\mathbf{L}, \ell}) & \text{Shear} \\ \frac{C_{\ell}^{\text{CMB}}}{2(C_{\ell}^{\text{total}})^2} \frac{d \ln \ell^2 C_{\ell}^{\text{CMB}}}{d \ln \ell} & \text{Magnification} \end{cases}, \quad (8)$$

where C^{CMB} is the unlensed CMB power spectrum and C^{total} is the total map power spectrum, including lensed CMB, foregrounds and detector noise. In what follows, for convenience, we further symmetrize these lensing weights with the substitution $F_{\ell, \mathbf{L}-\ell} \rightarrow (F_{\ell, \mathbf{L}-\ell} + F_{\mathbf{L}-\ell, \ell})/2$.

As shown in Fig. 2 of [33] and explained in [33], the noise power spectrum of the shear and magnification estimators shows a spike at $L \sim 3000$. In all the calculations below, the error bars indeed display this spike at that multipole, as expected. This is due to the lensing weights $F_{\ell, \mathbf{L}-\ell}$ for the shear and magnification only being optimal in the large-scale lens regime $L \ll \ell$, where large-scale lensing modes are reconstructed from small-scale temperature modes. Outside of this regime, the lensing weights $F_{\ell, \mathbf{L}-\ell}$ are effectively arbitrary, and lead to a null response to lensing at $L \sim 3000$, causing the spike in the noise power spectrum.

III. EXTRAGALACTIC FOREGROUNDS AND THEIR LENSING

Intuitively, the lensed foreground bias should depend on several properties of the foreground. First, the size of the bias depends on the amplitude of the foreground power spectrum. For example, if the foreground power spectrum is multiplied or divided by some factor, the bias to the CMB lensing estimators is also multiplied or divided by the same factor, since these estimators are quadratic in the temperature. Second, the bias also depends on the shape of the foreground power spectrum. For instance, if a foreground component had the same exact power spectrum as the CMB, then the CMB lensing estimators would have unit response not only to CMB lensing, but also to the foreground lensing. The shape of the foreground power spectrum thus determines how much the CMB lensing estimators respond to the foreground lensing. Finally, the redshift distribution of the foreground sources determines the amplitude of the foreground lensing convergence, and the size of its correlation with the CMB lensing convergence. In this section, we estimate each of these foreground properties.

A. Experimental configuration & foreground power

Throughout the paper, we consider a “CMB S3” experiment, similar to Simons Observatory, with a beam full-width at half maximum of $1.4'$ and a white noise level of $7\mu K'$ in temperature at 143 GHz. The lensing reconstruction for

the QE, shear and magnification estimators uses all Fourier modes with $\ell_{\min, T} = 30 \leq \ell \leq \ell_{\max, T} = 3,500$. We consider a temperature map at a single frequency, 143 GHz, without multi-frequency component separation. This slightly pessimistic assumption will produce slightly larger foreground biases, but it makes our results independent of the particular choice of component separation method. Furthermore, the shear estimator was shown to be robust to foreground contamination [33], so it may be used on a single-frequency temperature map. In the lensing weights defined above, the total power spectrum thus includes not only the lensed CMB and the white detector noise, but also all the foreground power spectra. This realistic choice downweights the small scales where foregrounds dominate, and will thus reduce the foreground bias. We do not include noise from atmospheric emission, as this quantity varies with observing site.

Throughout this study, we use the model from [52] for the power spectra of the extragalactic foregrounds at 143 GHz, as shown in Fig. 1. This model includes a point source mask for all objects with flux larger than 15 mJy.

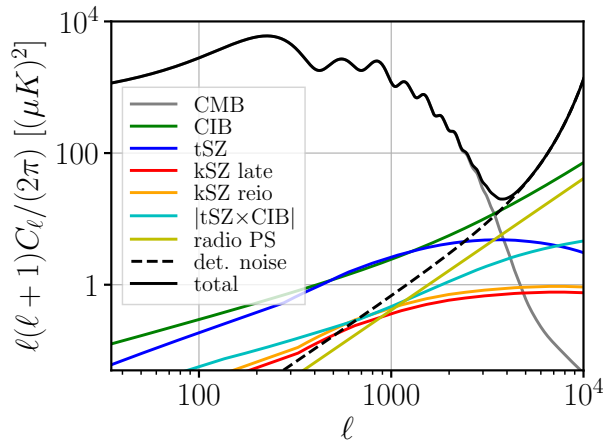


FIG. 1: Extragalactic foreground power spectra at 143GHz from [52], compared to the lensed CMB and the detector noise ($7\mu K$ white noise, $1.4'$ beam FWHM at 143 GHz). A point source cut of 15 mJy is assumed for radio PS, CIB and tSZ sources. **In the absence of foreground cleaning, the foregrounds are larger than the detector noise on scales $\ell \leq 3500$, and comparable to the lensed CMB at $\ell = 3500$, where most of the lensing information comes from.**

In order to quantify the bias from lensed foregrounds to cross-correlations of CMB lensing with tracers, we also consider the LSST gold sample of galaxies, as described in the LSST Science Book [53], chapter 3 and 13, with limiting magnitude in the i-band of $i_{\text{lim}} = 25.3$. This galaxy sample contains $n_{\text{gal}} = 46 \times 10^{0.31 \cdot (i_{\text{lim}} - 25)}$ galaxies per squared arcminute, has bias $b(z) = 1 + 0.84z$, and redshift distribution $dn/dz = n_{\text{gal}}(z/z_0)^2 e^{-z/z_0} / (2z_0)$, where $z_0 \equiv 0.0417i_{\text{lim}} - 0.744$.

B. Effective redshift distribution of foreground sources

Extragalactic foregrounds sources are associated with galaxies and clusters over a range of redshift. Intuitively, the foreground redshift distribution relevant for lensing should be related to the galaxy or cluster redshift distributions. Since the lensing estimators we consider are quadratic in the temperature map, the relevant quantity should therefore be related to the redshift distribution of the (unlensed) foreground power spectrum dC_ℓ^f/dz . However, this quantity depends on the multipole ℓ considered: generically, the large-scale power is mostly dominated by lower redshifts, because the same objects at lower redshift subtend a larger angle on the sky. As a result, the relevant source redshift distribution for foreground lensing is scale-dependent, unlike in galaxy lensing where it is simply the galaxy dn/dz . More precisely, as shown in App. B of [42], the CMB lensing estimators at multipole L respond to the following effective convergence:

$$\kappa_{\mathbf{L}}^f = \int d\chi W^{\kappa_f}(\chi, \mathbf{L}) \delta_m(\vec{k} = \mathbf{L}/\chi, \chi), \quad (9)$$

where the effective lensing kernel $W^{\kappa_f}(\chi, \mathbf{L})$ depends on the multipole L and is defined as:

$$W^{\kappa_f}(\chi, \mathbf{L}) = \int d\chi_S W^f(\chi_S, \mathbf{L}) W^\kappa(\chi, \chi_S), \quad (10)$$

where $W^\kappa(\chi, \chi_S)$ is the usual lensing kernel for a source at distance χ_S , defined in the previous section, and the effective foreground redshift distribution $W^f(\chi_S, \mathbf{L})$ is

$$W^f(\chi_S, \mathbf{L}) = \frac{dz_S}{d\chi_S} N_{\mathbf{L}} \int \frac{d^2\ell}{(2\pi)^2} F_{\ell, \mathbf{L}-\ell} f_{\ell, \mathbf{L}-\ell}^{f,z}, \quad (11)$$

with

$$f_{\ell_1, \ell_2}^{f,z} \equiv \left(\frac{2\mathbf{L}}{L^2} \right) \cdot \left[\ell_1 \frac{dC_{\ell_1}^f}{dz_S} + \ell_2 \frac{dC_{\ell_2}^f}{dz_S} \right]. \quad (12)$$

This is similar to galaxy lensing, where the lensing kernel is an integral over the source distribution. However, here, the source distribution is scale dependent. As expected, Eq. (11) involves the redshift distribution $\frac{dC_{\ell}^f}{dz_S}$ of the unlensed foreground power spectrum. One important consequence is that the noise and resolution of the CMB experiment, which determine the lensing weights $F_{\ell, \mathbf{L}-\ell}$, also determine how the $\frac{dC_{\ell}^f}{dz_S}$ terms are weighted to produce the effective foreground redshift distribution. For instance, for a higher resolution CMB experiment, the CMB lensing reconstruction relies on smaller scales, thus upweighting $\frac{dC_{\ell}^f}{dz_S}$ at high ℓ , which typically come from higher redshift sources.

For the CMB S3 experiment we consider, and assuming $l_{\max T} = 3500$, most of the CMB lensing signal-to-noise comes from temperature multipoles $\ell \sim 3000$. We therefore ignore the L -dependence of the foreground lensing kernels, and approximate them as $W^f(\chi_S, \mathbf{L}) \propto \frac{dC_{\ell=3000}^f}{d\chi_S}$. The problem then simplifies to modeling $\frac{dC_{\ell=3000}^f}{d\chi_S}$ for each foreground of interest. For the CIB, we use the value computed in [42], following the CIB halo model from [54], using the luminosity functions from [55, 56]. For tSZ, we implement the halo model in [57]. For the late time kSZ, we use the redshift distribution from Fig. 6 in [58], corresponding to their L60CSFz2 model. For the reionization kSZ, we assume a single source redshift at $z = 8$, consistent with the redshift of a step-like reionization in ([59] table A1). For the radio PS, we adopt the redshift distribution Eq. 26 in [60], describing the source sample from [61], selected in NVSS at 1.4GHz. Because the frequency dependence of the synchrotron emission is very mild, sources at 1.4GHz and 143GHz mostly coincide, and correspond to active galactic nuclei. These **approximate** effective redshift distributions are shown in Fig. 2, along with the corresponding foreground lensing kernels. Our knowledge of the foreground redshift distributions is somewhat uncertain. For this reason, the foreground lensing biases we obtain are also uncertain, and should be considered reasonable values rather than exact values.

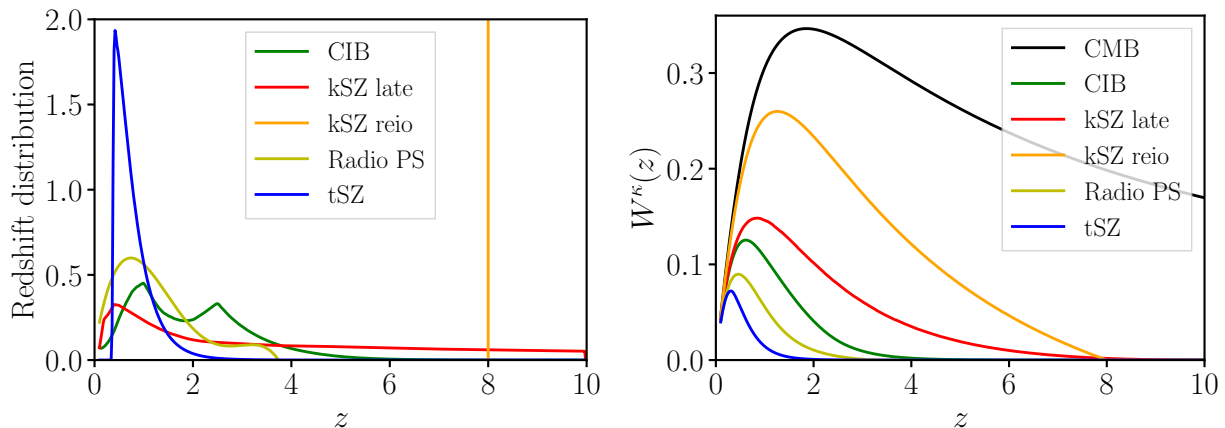


FIG. 2: Extragalactic foregrounds are themselves lensed, because they are emitted at cosmological distances. The exact foreground redshift distributions (Eq. (11)) depend on the experimental configuration (resolution, sensitivity, masking, etc.) and the lensing multipole L . We approximate these redshift distributions as $W^f(\chi_S, \mathbf{L}) \propto \frac{dC_{\ell=3000}^f}{d\chi_S}$ (left panel). From these **approximate** redshift distributions, we infer the corresponding lensing kernels (right panel) from Eq. (10), showing which redshifts contribute to the foreground lensing convergence. These lensing kernels determine the amplitude of the foreground lensing convergence and its correlation with the CMB lensing convergence. **Our approximations make these foreground lensing kernels independent of the experimental configuration and lensing multipole L .**

C. Foreground lensing power spectra

We use again the Limber and flat sky approximations, and obtain the auto and cross-spectra of the foreground lensing convergences:

$$C_L^{\kappa_1 \kappa_2} = \int \frac{d\chi}{\chi^2} W^{\kappa_1}(\chi) W^{\kappa_2}(\chi) P_m(k = \frac{L+1/2}{\chi}, z(\chi)). \quad (13)$$

How small is foreground lensing compared to CMB lensing? As we demonstrate in the next section, the relevant quantity to assess the bias in CMB lensing auto-spectrum is $C_L^{\kappa_f \kappa_{\text{CMB}}} / C_L^{\kappa_{\text{CMB}} \kappa_{\text{CMB}}}$, shown in the left panel of Fig. 3. To assess the bias in cross-correlation with a tracer g , the relevant quantity is $C_L^{\kappa_f g} / C_L^{\kappa_{\text{CMB}} g}$, shown in the right panel of Fig. 3 for the LSST gold sample. These plots show that the amplitude of foreground lensing for $L = 100 - 1000$ ranges between 5% (for tSZ) and 85% (for the reionization kSZ) of that of CMB lensing, depending on the foreground considered.

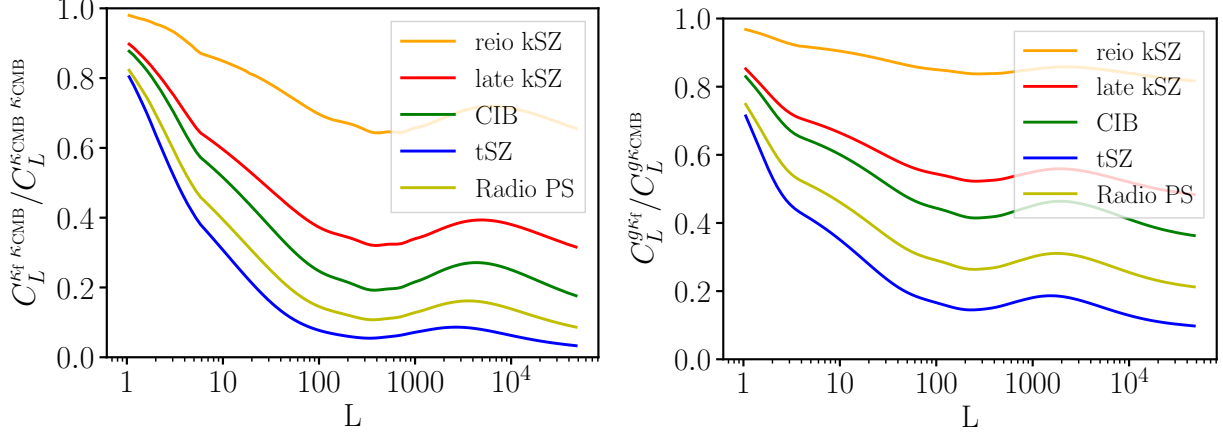


FIG. 3: Since foreground sources lie at lower redshift than the CMB, the lensing they experience is less than that of the CMB, although of the same order of magnitude. The relevant reduction factor is $C_L^{\kappa_f \kappa_{\text{CMB}}} / C_L^{\kappa_{\text{CMB}} \kappa_{\text{CMB}}}$ for the primary bias on the CMB lensing auto-spectrum (left panel, from Eq. (19)) and $C_L^{\kappa_f g} / C_L^{\kappa_{\text{CMB}} g}$ for the bias in cross-correlation with LSST galaxies (right panel, from Eq. (17)).

IV. BIAS TO CMB LENSING FROM LENSED FOREGROUNDS

In this section, we use the foreground power spectra and redshift distributions from the previous section to predict the bias to CMB lensing from lensed foregrounds. Throughout, we make the simplifying approximation that each foreground map f is lensed by one lensing convergence field κ_f , computed from the redshift distribution of the foreground. In reality, each redshift slice of the foreground is lensed separately by the convergence field corresponding to that source redshift. We believe that this approximation is sufficient, given the uncertainty in the foreground redshift distributions, and leave its exploration to future work.

Furthermore, since the lensing convergence and the foregrounds are both produced by the matter distribution in the universe, they are correlated. The correlation between foregrounds and the CMB lensing convergence is the origin of the usual foreground biases to CMB lensing [29, 30, 32, 33], and we shall therefore not discuss it further. The correlation between foregrounds and the foreground lensing convergence contributes additional bias in principle. However, this will be smaller than the previous one, and we shall only mention it in this section.

We compute all the bias terms described in this section in two independent ways, as described in App. B, and find a good agreement everywhere.

A. Response of CMB lensing estimators to lensed contaminants

Following Sec. VIII in [42], any quadratic combination of lensed foreground maps effectively constitutes a (potentially biased and suboptimal) foreground lensing estimator. In particular, the CMB lensing estimators of the form Eq. (6)

are such quadratic combinations. As a result, they generically have non-zero response to the foreground convergence κ_f . Indeed, when applied to a lensed foreground map f , the quadratic estimators above have the following response to foreground lensing:

$$\langle \mathcal{Q}_{\mathbf{L}}[f, f] \rangle = \mathcal{R}_{\mathbf{L}}^f \kappa_{\mathbf{L}}^f, \text{ with } \mathcal{R}_{\mathbf{L}}^f = \frac{\int \frac{d^2 \ell}{(2\pi)^2} F_{\ell, \mathbf{L}-\ell} f_{\ell, \mathbf{L}-\ell}^f}{\int \frac{d^2 \ell}{(2\pi)^2} F_{\ell, \mathbf{L}-\ell} f_{\ell, \mathbf{L}-\ell}^{\text{CMB}}}, \quad (14)$$

where the expectation value is at fixed convergence field, varying the realization of the unlensed foreground map, like in the derivation of the standard quadratic estimator [46].

If the unlensed power spectra of the foreground and CMB were identical, the response $\mathcal{R}_{\mathbf{L}}^f$ would be unity on all scales, and the CMB lensing estimator would also be an unbiased foreground lensing estimator. If the foreground power spectrum C^f is reduced by a factor α , e.g., from multi-frequency component separation, then the response $\mathcal{R}_{\mathbf{L}}^f$ is also reduced by the same factor α . It is thus clear that the value of the response $\mathcal{R}_{\mathbf{L}}^f$ depends both on the amplitude and shape of the foreground power spectrum C^f , and can in principle take any value and sign. In practice, Fig. 4 shows that $\mathcal{R}_{\mathbf{L}}^f$ is of order one percent for $L \lesssim 1000$, with different signs depending on the foreground and choice of estimator (QE, shear, magnification). The shear estimator does not reduce the response to foreground lensing compared to the QE. This is not surprising since foreground lensing is a true lensing effect, which the shear is built to include. Furthermore, the lensing responses for shear and magnification have opposite signs. As a result, comparing shear and magnification provides a useful null test to detect the presence of lensed foreground bias.

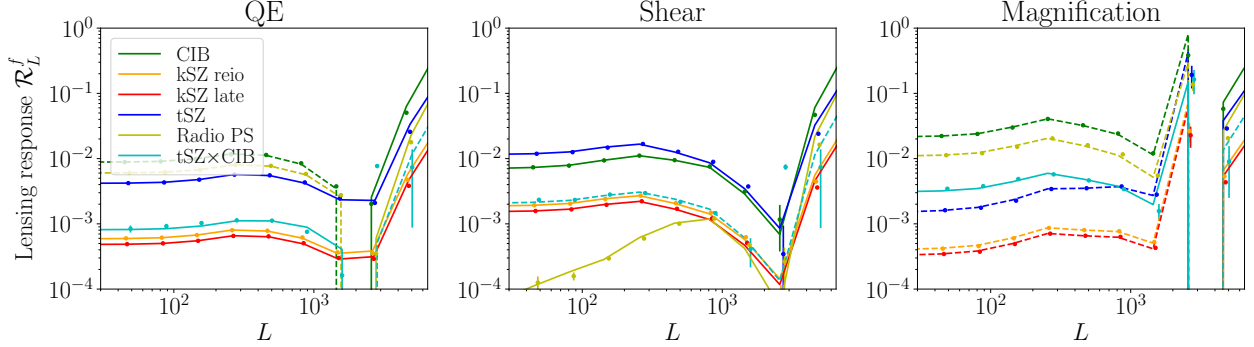


FIG. 4: The lensing responses $\mathcal{R}_{\mathbf{L}}^f$ from Eq. (14) for the QE (left), shear (center), and magnification (right) estimators show what fraction of the foreground lensing convergence leaks into the CMB lensing estimator. A response of unity would mean that the CMB lensing estimator has a bias equal to one times the foreground lensing convergence. In practice, the responses for the various foregrounds and estimators are found to be of order percent for $L \lesssim 1000$. The sign varies depending on the multipole, the foreground and the estimator (dashed lines represent negative responses). The points with error bars are the simulation results, and the lines are the analytical calculations, binned like the simulations. The points and curves for each foreground are slightly shifted horizontally to improve the visibility of error bars.

As a result, if a lensed foreground is present in the temperature map, the quadratic estimator will inevitably partially reconstruct the lensing of the foreground:

$$\langle \mathcal{Q}_{\mathbf{L}}[T + f, T + f] \rangle = \kappa_{\mathbf{L}}^{\text{CMB}} + \mathcal{R}_{\mathbf{L}}^f \kappa_{\mathbf{L}}^f + \mathcal{O}(\kappa_{\mathbf{L}}^{\text{CMB}2}, \kappa_f^2, \kappa_{\text{CMB}} \kappa_f), \quad (15)$$

where again the expectation value is at fixed lensing fields ($\kappa^{\text{CMB}}, \kappa^f$) but marginalizing over the unlensed fields (T^0, f^0). This term produces biases in CMB lensing auto and cross-correlation, as we explain below.

Furthermore, two distinct unlensed foreground components f and f' may have a significant correlation. This is the case for example for CIB and tSZ. In this case, additional biases to CMB lensing occur, coming from terms of the form

$$\mathcal{Q}_{\mathbf{L}}[f, f'] + \mathcal{Q}_{\mathbf{L}}[f', f] = \mathcal{R}_{\ell}^{ff'} \left[\kappa_{\mathbf{L}}^f + \kappa_{\mathbf{L}}^{f'} \right], \quad (16)$$

where the additional response $\mathcal{R}_{\ell}^{ff'}$ is computed as in Eq. (14), except that f^f is replaced by $f^{ff'}$. This response is typically smaller, reduced by a factor of order the correlation coefficient between f and f' . We quantify these extra terms below, in the case of CIB and tSZ.

B. Bias in cross-correlation

The response of CMB lensing quadratic estimators to the foreground convergence naturally leads to a bias in cross-correlation with any tracer g (e.g., galaxy or cluster number density, or galaxy shear):

$$\frac{\delta C_L^{\kappa_{\text{CMB}}g}}{C_L^{\kappa_{\text{CMB}}g}} = \mathcal{R}_L^f \frac{C_L^{\kappa_f g}}{C_L^{\kappa_{\text{CMB}}g}}. \quad (17)$$

When two correlated foregrounds are present, such as CIB and tSZ, an additional bias from the cross term arises:

$$\frac{\delta C_L^{\kappa_{\text{CMB}}g}}{C_L^{\kappa_{\text{CMB}}g}} = \mathcal{R}_L^{ff'} \frac{[C_L^{\kappa_f g} + C_L^{\kappa_{f'} g}]}{C_L^{\kappa_{\text{CMB}}g}}. \quad (18)$$

Fig. 5 shows that these biases are typically of order one percent of the cross-power spectrum $C_L^{\kappa_{\text{CMB}}g}$, i.e. significantly larger than the statistical uncertainty (lensing reconstruction noise plus cosmic variance) for Simons Observatory lensing and the LSST gold galaxy sample.

C. Bias in auto-correlation

In auto-correlation, several bias terms arise. These are derived by expanding $\langle \mathcal{Q}[T+f, T+f] \mathcal{Q}[T+f, T+f] \rangle$, using the bilinearity of the quadratic estimators.

We call “primary” bias the terms where one quadratic estimator is applied to two foreground maps and the other is applied to two CMB maps: $\langle \mathcal{Q}[f, f] \mathcal{Q}[T, T] \rangle + f \leftrightarrow T$. These terms are analogous to the cross-correlation case, with an additional combinatorial factor 2:

$$\text{Primary bias: } \frac{\delta C_L^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}}}{C_L^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}}} = 2 \mathcal{R}_L^f \frac{C_L^{\kappa_f \kappa_{\text{CMB}}}}{C_L^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}}}. \quad (19)$$

Again, when two correlated foregrounds like CIB and tSZ are present, an additional primary bias arises from the cross term:

$$\frac{\delta C_L^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}}}{C_L^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}}} = 2 \mathcal{R}_L^{ff'} \frac{[C_L^{\kappa_f \kappa_{\text{CMB}}} + C_L^{\kappa_{f'} \kappa_{\text{CMB}}}]}{C_L^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}}}. \quad (20)$$

These terms are shown to be percent level biases in the CMB lensing auto-spectrum in Fig. 6.

In auto-correlation, an additional “secondary” bias is present, due to the terms where each quadratic estimator is applied to one CMB map and one foreground map: $\langle \mathcal{Q}[T, f] \mathcal{Q}[T, f] \rangle + \text{perm.}$. The corresponding secondary bias is derived in App. A:

$$\begin{aligned} \text{Secondary bias: } \delta C_{L_0}^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}} &= 8N_L^2 \int \frac{d^2 \mathbf{L}}{(2\pi)^2} \int \frac{d^2 \boldsymbol{\ell}}{(2\pi)^2} F_{\boldsymbol{\ell}, L_0 - \boldsymbol{\ell}} F_{\boldsymbol{\ell} - \mathbf{L}, L_0, L - \boldsymbol{\ell}} \alpha_{\mathbf{L}, \boldsymbol{\ell} - \mathbf{L}} C_L^{\kappa_{\text{CMB}}\kappa_f} \\ &\times \left[\alpha_{-\mathbf{L}, \boldsymbol{\ell} - L_0} C_{L_0 - \boldsymbol{\ell}}^{\text{CMB}} C_{\boldsymbol{\ell} - \mathbf{L}}^f + \alpha_{-\mathbf{L}, L_0 + \mathbf{L} - \boldsymbol{\ell}} C_{\boldsymbol{\ell} - \mathbf{L}}^{\text{CMB}} C_{\boldsymbol{\ell} - L_0}^f \right] \end{aligned} \quad (21)$$

where we used the notation $\alpha_{\mathbf{L}, \boldsymbol{\ell} - \mathbf{L}} = -2 \frac{\mathbf{L}}{L^2} \cdot (\boldsymbol{\ell} - \mathbf{L})$, and C^{CMB} and C^f represent the unlensed CMB and foreground power spectra. Again, when correlated foregrounds such as CIB and tSZ are present, an additional secondary bias appears, obtained by substituting $C^{ff'}$ to C^f and $[C^{\kappa_{\text{CMB}}\kappa_f} + C^{\kappa_{\text{CMB}}\kappa_{f'}}]$ to $C^{\kappa_{\text{CMB}}\kappa_f}$ in Eq. (21). Naïvely, this secondary bias should be of the same order of magnitude as the primary bias. However, in the limit $L, L_0 \ll \ell$, i.e. when reconstructing large-scale lensing modes from small-scale temperature modes, the terms in the square bracket cancel exactly. As a result, the secondary bias is negligible compared to the primary bias at low lensing multipoles. However, the secondary bias dominates over the primary bias at lensing multipoles of a few thousand. This is shown in Fig. 7. Overall, the sum of primary and secondary lensed foreground biases constitutes a percent-level bias in CMB lensing auto-spectrum, comparable with the statistical uncertainty (lensing reconstruction noise plus cosmic variance) for a stage III CMB experiment.

Finally, a “4-point” bias is also present, where both quadratic estimators are evaluated on two foreground maps:

$$\text{4-point bias: } \delta C_L^{\kappa_{\text{CMB}}\kappa_{\text{CMB}}} = \left(\mathcal{R}_L^f \right)^2 C_L^{\kappa_f \kappa_f} + N^{(1)\text{-like term}}. \quad (22)$$

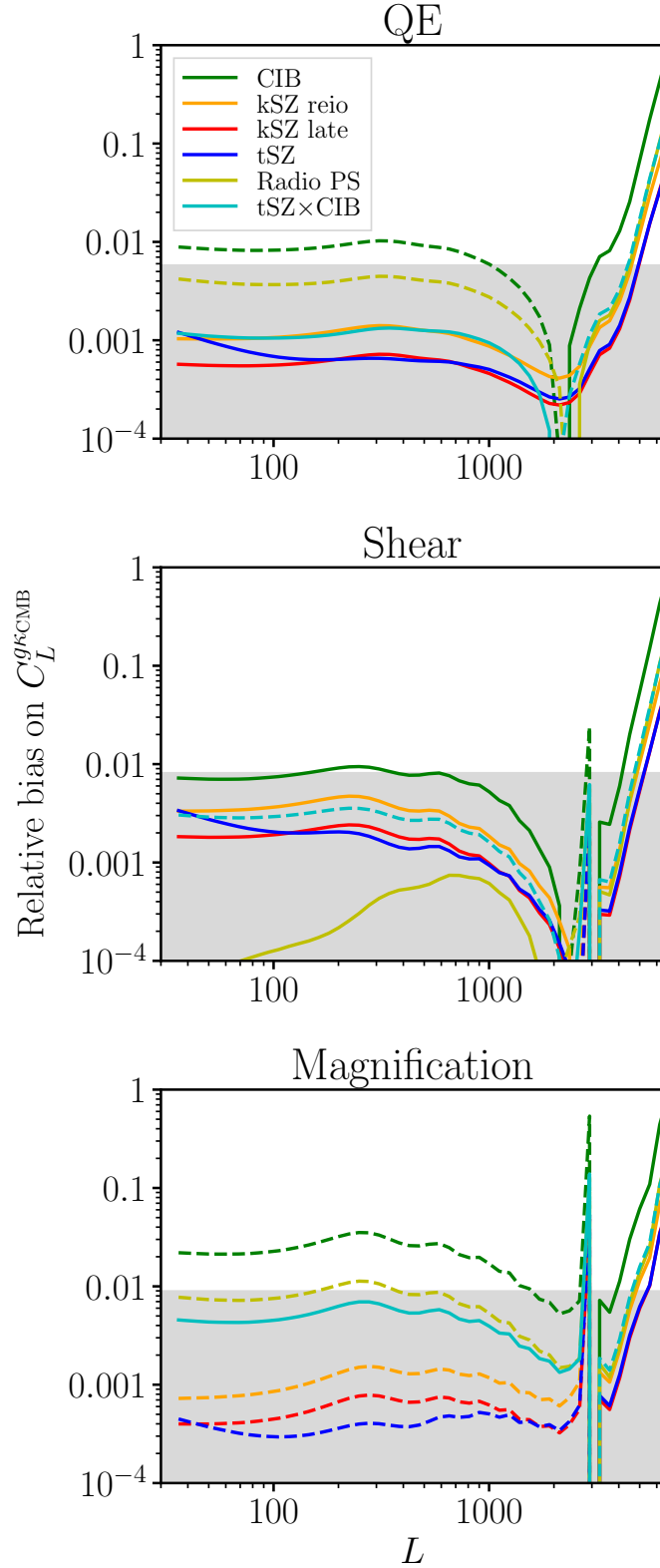


FIG. 5: The relative bias from lensed foregrounds on the cross correlation of CMB lensing from Simons Observatory and the LSST gold galaxies is one percent or larger, depending on the foreground, when using the QE (top), shear (center) and magnification (bottom) estimators. This percent-level systematic bias is larger than the statistical uncertainty (grey shaded areas) of 0.6%, 0.8% and 0.9% on the amplitude of the cross-correlation for the QE, shear and magnification respectively (assuming $f_{\text{sky}} = 40\%$, $L_{\text{max}, \kappa} = 1000$ and including cosmic variance). Dashed lines represent negative biases. The spikes at $L \simeq 3000$ for the shear and magnification estimators are due to spikes in their noise power spectra at this multipole, as shown in Fig. 2 of [33].

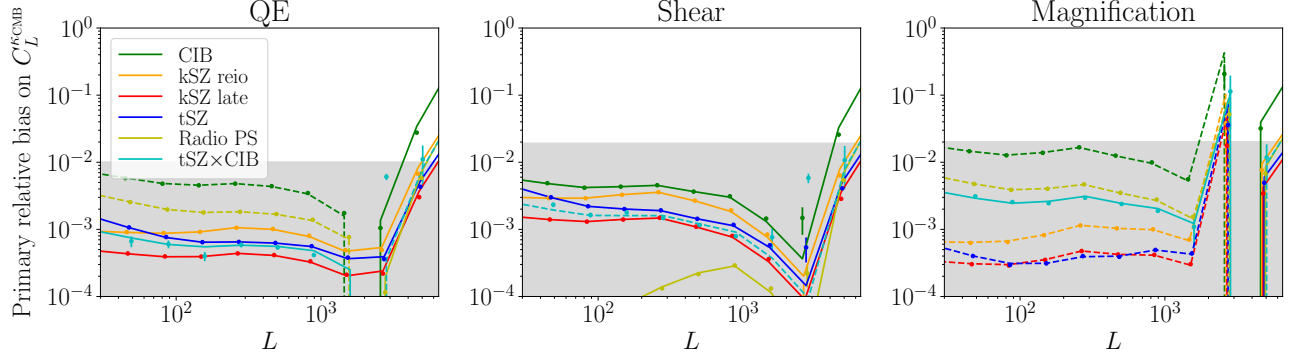


FIG. 6: The primary relative bias from lensed foregrounds on the power spectrum of CMB lensing from Simons Observatory is of order one percent or less, when using the QE (left), shear (center) and magnification (right) estimators. This percent-level systematic bias can be as large as the statistical uncertainty (grey shaded areas) of 1%, 2%, and 2% on the amplitude of the CMB lensing power spectrum for the QE, shear and magnification (assuming $f_{\text{sky}} = 40\%$ and including cosmic variance). Dashed lines represent negative biases. The points with error bars are the simulation results, and the lines are the analytical calculations, binned like the simulations. The points and curves for each foreground are slightly shifted horizontally to improve the visibility of error bars.

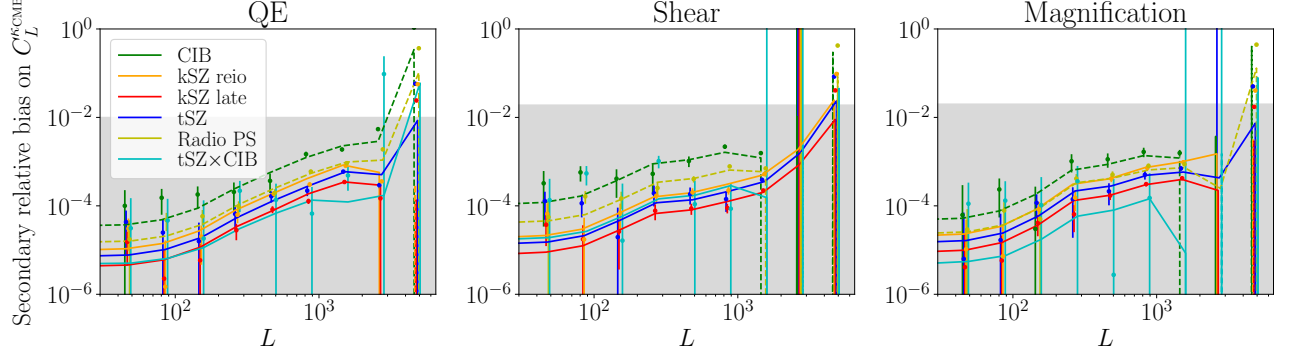


FIG. 7: The secondary relative bias from lensed foregrounds on the power spectrum of CMB lensing is suppressed at low lensing multipoles due to the cancellation in Eq. (21). However, for lensing multipoles of a few thousand, the secondary bias dominates over the primary bias. The various panels show this for the QE (left), shear (center) and magnification (right) estimators. The grey shaded areas represent the statistical uncertainty on the amplitude of the CMB lensing power spectrum (1%, 2% and 2% for the QE, shear and magnification, assuming $f_{\text{sky}} = 40\%$ and including cosmic variance), and dashed lines represent negative terms. The points with error bars are the simulation results, and the lines for QE are the analytical calculations, evaluated (not binned) at the simulation points. The lines for shear and magnification are simply connecting the points, not analytical calculations. The points and curves for each foreground are slightly shifted horizontally to improve the visibility of error bars.

Since the response to foreground lensing is small for $L \lesssim 1000$, i.e. $\mathcal{R}_L^f \ll 1$, this 4-point bias is expected to be much smaller than the primary and secondary biases on these scales, and we will not discuss it further in this paper. However, it may be large on smaller scales $L \gtrsim 1000$.

All the biases due to lensed foregrounds discussed so far are present whether or not the foreground component is a Gaussian random field. This is in contrast with the non-Gaussian foreground biases usually discussed in the literature, which are caused by the non-Gaussianity of the foregrounds, and not the fact that they are lensed. There exist additional biases due to the foregrounds being both non-Gaussian and lensed. To lowest order in κ_f , these terms are of the form:

$$\begin{aligned}
 \langle \mathcal{Q}[f_1, f_2] \mathcal{Q}[f_3, f_4] \rangle &\sim \langle \kappa_{f_i} f_1 f_2 f_3 f_4 \rangle \sim \langle \kappa_{f_i} f_1 \rangle \langle f_2 f_3 f_4 \rangle \leftarrow \text{foreground bispectrum} \\
 &+ \langle \kappa_{f_i} f_1 f_2 \rangle \langle f_3 f_4 \rangle \leftarrow \text{lensing-foreground-foreground bispectrum} \\
 &+ \langle \kappa_{f_i} f_1 f_2 f_3 f_4 \rangle_c \leftarrow \text{connected 5-point function}
 \end{aligned} \tag{23}$$

We expect these terms to be smaller than the usual non-Gaussian foreground biases [29, 30, 32, 37, 38] by the

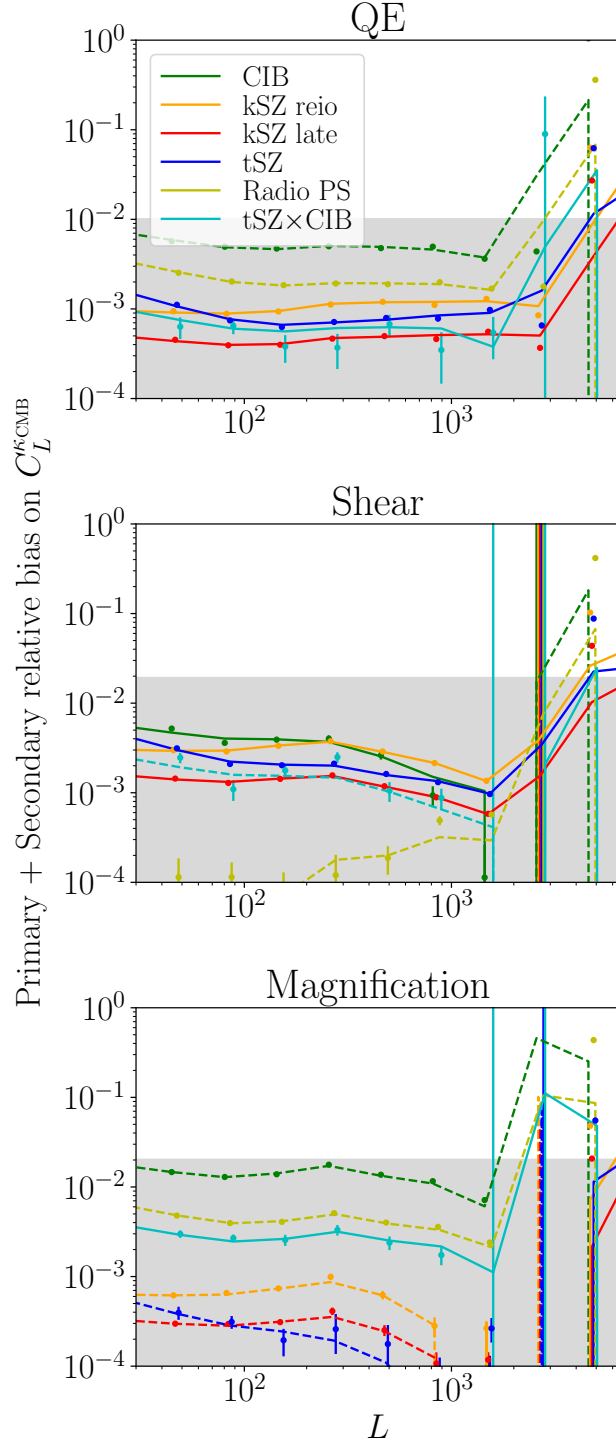


FIG. 8: The total (primary + secondary) relative bias from lensed foregrounds on the power spectrum of CMB lensing from Simons Observatory is of order one percent, when using the QE (top), shear (center) and magnification (bottom) estimators. It is dominated by the primary bias for $L \lesssim 1000$, then by the secondary bias on smaller scales. The grey shaded areas represent the statistical uncertainty on the amplitude of the CMB lensing power spectrum (1%, 2% and 2% for the QE, shear and magnification, assuming $f_{\text{sky}} = 40\%$ and including cosmic variance), and dashed lines represent negative terms. The points with error bars are the simulation results, and the lines for QE are the analytical calculations, evaluated (not binned) at the simulation points. The lines for shear and magnification are simply connecting the points, not analytical calculations. The points and curves for each foreground are slightly shifted horizontally to improve the visibility of error bars. The spikes at $L \simeq 3000$ for the shear and magnification estimators are due to spikes in their noise power spectra at this multipole, as shown in Fig. 2 of [33].

foreground lensing response $\mathcal{R}_{\mathbf{L}}^f$, but we leave the evaluation of these terms to future work.

V. CONCLUSION

In this paper, we quantified for the first time the bias to CMB lensing auto and cross-correlations due to the presence of lensed foregrounds in the observed temperature map. For an experiment similar to Simons Observatory, and in the absence of multi-frequency foreground cleaning, this lensed foreground bias is a percent-level effect for both the CMB lensing power spectrum and for the cross-correlation of CMB lensing with LSST galaxies. This bias is thus marginally significant in auto-correlation, and highly significant in cross-correlation. For future polarization-dominated CMB lensing experiments like CMB S4, we expect the lensed foreground bias to be a lesser problem, since extragalactic foregrounds are expected to be smaller in polarization than in temperature.

For the standard quadratic estimator and the magnification estimators, the standard non-Gaussian foreground biases are typically more important than the lensed foreground bias. Any method that successfully controls the former will thus automatically control the latter. On the other hand, the shear estimator is mostly insensitive to the standard foreground bias. As a result, the lensed foreground bias is dominant. It needs to be reduced in order to provide an unbiased lensing measurement.

Any mitigation method that reduces the level of foregrounds in CMB temperature maps will also reduce the lensed foreground bias. This is the case of multi-frequency component separation, scale cuts in the temperature map, masking or inpainting point sources (below the 15 mJy assumed used in this paper), and the cleaned-gradient estimators. This suggests that a combination of such methods may be the best approach.

On the other hand, mitigation methods that rely on the non-Gaussian structure of foregrounds will not in general reduce the lensed foreground bias, since the lensed foreground bias is present even if the foreground of interest is perfectly Gaussian. Examples are the standard foreground bias hardening and the shear estimator. Indeed, we have shown that the shear estimator, designed to distinguish the spatial symmetry of the lensing shear (quadrupole) from those of the non-Gaussian foregrounds (monopole), is sensitive to the lensed foreground bias. However, the lensed foreground biases in shear and magnification have opposite signs. Comparing shear and magnification estimators therefore still provides a useful null test.

In App. C, we suggest a “lensed foreground bias-hardening” to reduce the lensed foreground bias. Contrary to the usual bias hardening, this lensed foreground bias hardening does not assume any knowledge of the often uncertain non-Gaussianity of the foreground. However, it relies on a knowledge of the power spectrum and source redshift distribution of the foreground. The first estimator we derive has zero response to the lensing of a given foreground component, at the cost of an increased noise. To avoid a potentially large noise cost, we derive a second estimator, which instead minimizes the total variance from lensing noise plus residual lensed foreground bias. We leave the exploration of these forms of bias hardening to future work.

Finally, one may remove the lensed foreground bias by subtracting their theory predictions, as computed in this paper. An uncertainty on the theory prediction of the lensed foreground bias of order ten percent would be acceptable. Quantifying the uncertainty in the foreground source distributions assumed in this paper would be useful to assess whether such accuracy is reachable. Furthermore, our study makes an important simplifying approximation, by assuming that each foreground map f is lensed as a whole by a single convergence field κ_f , determined by the foreground redshift distribution. In reality, each redshift slice of the foreground emission is lensed separately by a slightly different convergence field, determined by the redshift of that slice. We believe that the error due to this approximation is comparable to the uncertainty in the foreground redshift distributions. A complete analysis of foreground biases to CMB lensing including both the effect of foreground non-Gaussianity and of realistic foreground lensing, would be a worthwhile endeavor. This could be achieved by lensing realistic non-Gaussian simulations such as the *WebSky*¹ mocks [62].

Finally, we suggest that foreground lensing may be considered a signal rather than a bias [41–43].

VI. ACKNOWLEDGMENTS

We are indebted to Marcelo Alvarez for making this collaboration possible, and to Sukhdeep Singh and Yu Feng for their help. We warmly thank Simone Ferraro, Alex van Engelen and Uros Seljak for many insightful discussions which inspired this project. We thank Rahul Datta for his insight on the nature and redshift distribution of radio

¹ https://mocks.cita.utoronto.ca/index.php/WebSky_Extragalactic_CMB_Mocks

point sources. We thank Vanessa Boehm, Simone Ferraro, Colin Hill, Ben Horowitz, Srinivasan Raghunathan and Martin White for their comments on an earlier version of this paper.

ES is supported by the Chamberlain fellowship at Lawrence Berkeley National Laboratory.

This work used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

-
- [1] K. M. Smith, O. Zahn, and O. Doré, *Phys. Rev. D* **76**, 043510 (2007), 0705.3980.
 - [2] C. M. Hirata, S. Ho, N. Padmanabhan, U. Seljak, and N. A. Bahcall, *Phys. Rev. D* **78**, 043520 (2008), 0801.0644.
 - [3] S. Das, B. D. Sherwin, P. Aguirre, J. W. Appel, J. R. Bond, C. S. Carvalho, M. J. Devlin, J. Dunkley, R. Dünner, T. Essinger-Hileman, et al., *Physical Review Letters* **107**, 021301 (2011), 1103.2124.
 - [4] S. Das, T. Louis, M. R. Nolte, G. E. Addison, E. S. Battistelli, J. R. Bond, E. Calabrese, D. Crichton, M. J. Devlin, S. Dicker, et al., *J. Cosmology Astropart. Phys.* **4**, 014 (2014), 1301.1037.
 - [5] M. Madhavacheril, N. Sehgal, R. Allison, N. Battaglia, J. R. Bond, E. Calabrese, J. Caligiuri, K. Coughlin, D. Crichton, R. Datta, et al., *Physical Review Letters* **114**, 151302 (2015), 1411.7999.
 - [6] A. van Engelen, B. D. Sherwin, N. Sehgal, G. E. Addison, R. Allison, N. Battaglia, F. de Bernardis, J. R. Bond, E. Calabrese, K. Coughlin, et al., *ApJ* **808**, 7 (2015), 1412.0626.
 - [7] B. D. Sherwin, A. van Engelen, N. Sehgal, M. Madhavacheril, G. E. Addison, S. Aiola, R. Allison, N. Battaglia, D. T. Becker, J. A. Beall, et al., *Phys. Rev. D* **95**, 123529 (2017), 1611.09753.
 - [8] R. Allison, S. N. Lindsay, B. D. Sherwin, F. de Bernardis, J. R. Bond, E. Calabrese, M. J. Devlin, J. Dunkley, P. Gallardo, S. Henderson, et al., *MNRAS* **451**, 849 (2015), 1502.06456.
 - [9] A. van Engelen, R. Keisler, O. Zahn, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, et al., *ApJ* **756**, 142 (2012), 1202.0546.
 - [10] D. Hanson, S. Hoover, A. Crites, P. A. R. Ade, K. A. Aird, J. E. Austermann, J. A. Beall, A. N. Bender, B. A. Benson, and L. E. Bleem, *Phys. Rev. Lett.* **111**, 141301 (2013), 1307.5830.
 - [11] E. J. Baxter, R. Keisler, S. Dodelson, K. A. Aird, S. W. Allen, M. L. N. Ashby, M. Bautz, M. Bayliss, B. A. Benson, L. E. Bleem, et al., *ApJ* **806**, 247 (2015), 1412.7521.
 - [12] K. T. Story, D. Hanson, P. A. R. Ade, K. A. Aird, J. E. Austermann, J. A. Beall, A. N. Bender, B. A. Benson, L. E. Bleem, J. E. Carlstrom, et al., *ApJ* **810**, 50 (2015), 1412.4760.
 - [13] Y. Omori, R. Chown, G. Simard, K. T. Story, K. Aylor, E. J. Baxter, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, et al., *ApJ* **849**, 124 (2017), 1705.00743.
 - [14] P. A. R. Ade, Y. Akiba, A. E. Anthony, K. Arnold, M. Atlas, D. Barron, D. Boettger, J. Borrill, S. Chapman, Y. Chinone, et al., *Physical Review Letters* **113**, 021301 (2014), 1312.6646.
 - [15] T. Namikawa, Y. Chinone, H. Miyatake, M. Oguri, R. Takahashi, A. Kusaka, N. Katayama, S. Adachi, M. Aguilar, H. Aihara, et al., *arXiv e-prints* (2019), 1904.02116.
 - [16] M. Aguilar Faundez, K. Arnold, C. Baccigalupi, D. Barron, D. Beck, F. Bianchini, D. Boettger, J. Borrill, J. Carron, K. Cheung, et al., *arXiv e-prints* (2019), 1903.07046.
 - [17] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, et al., *A&A* **571**, A17 (2014), 1303.5077.
 - [18] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, et al., *ArXiv e-prints* (2015), 1502.01591.
 - [19] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, et al., *arXiv e-prints* (2018), 1807.06210.
 - [20] T. Keck Array, BICEP2 Collaborations, :, P. A. R. Ade, Z. Ahmed, R. W. Aikin, K. D. Alexander, D. Barkats, S. J. Benton, C. A. Bischoff, et al., *ArXiv e-prints* (2016), 1606.01968.
 - [21] A. Manzotti, K. T. Story, W. L. K. Wu, J. E. Austermann, J. A. Beall, A. N. Bender, B. A. Benson, L. E. Bleem, J. J. Bock, and J. E. Carlstrom, *ApJ* **846**, 45 (2017), 1701.04396.
 - [22] J. Carron, A. Lewis, and A. Challinor, *Journal of Cosmology and Astro-Particle Physics* **2017**, 035 (2017), 1701.01712.
 - [23] J. Carron, *Phys. Rev. D* **99**, 043518 (2019), 1808.10349.
 - [24] S. W. Henderson, R. Allison, J. Austermann, T. Baildon, N. Battaglia, J. A. Beall, D. Becker, F. De Bernardis, J. R. Bond, E. Calabrese, et al., *Journal of Low Temperature Physics* **184**, 772 (2016), 1510.02809.
 - [25] B. A. Benson, P. A. R. Ade, Z. Ahmed, S. W. Allen, K. Arnold, J. E. Austermann, A. N. Bender, L. E. Bleem, J. E. Carlstrom, C. L. Chang, et al., in *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII* (2014), vol. 9153 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 91531P, 1407.2973.
 - [26] A. Suzuki, P. Ade, Y. Akiba, C. Aleman, K. Arnold, C. Baccigalupi, B. Barch, D. Barron, A. Bender, D. Boettger, et al., *Journal of Low Temperature Physics* **184**, 805 (2016), 1512.07299.
 - [27] P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali, D. Alonso, M. A. Alvarez, K. Arnold, P. Ashton, J. Austermann, et al., *Journal of Cosmology and Astro-Particle Physics* **2019**, 056 (2019), 1808.07445.
 - [28] K. N. Abazajian, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso, K. S. Arnold, C. Baccigalupi, J. G. Bartlett, N. Battaglia,

- B. A. Benson, et al., arXiv e-prints arXiv:1610.02743 (2016), 1610.02743.
- [29] S. J. Osborne, D. Hanson, and O. Doré, *Journal of Cosmology and Astro-Particle Physics* **2014**, 024 (2014), 1310.7547.
 - [30] A. van Engelen, S. Bhattacharya, N. Sehgal, G. P. Holder, O. Zahn, and D. Nagai, *ApJ* **786**, 13 (2014), 1310.7023.
 - [31] S. Raghunathan, S. Patil, E. J. Baxter, F. Bianchini, L. E. Bleem, T. M. Crawford, G. P. Holder, A. r. Manzotti, and C. L. Reichardt, *Journal of Cosmology and Astro-Particle Physics* **2017**, 030 (2017), 1705.00411.
 - [32] S. Ferraro and J. C. Hill, *Phys. Rev. D* **97**, 023512 (2018), 1705.06751.
 - [33] E. Schaan and S. Ferraro, *Phys. Rev. Lett.* **122**, 181301 (2019), 1804.06403.
 - [34] S. Raghunathan, S. Patil, E. Baxter, B. A. Benson, L. E. Bleem, T. L. Chou, T. M. Crawford, G. P. Holder, T. McClintock, and C. L. Reichardt, *ApJ* **872**, 170 (2019), 1810.10998.
 - [35] M. S. Madhavacheril and J. C. Hill, arXiv e-prints arXiv:1802.08230 (2018), 1802.08230.
 - [36] S. Raghunathan, G. P. Holder, J. G. Bartlett, S. Patil, C. L. Reichardt, and N. Whitehorn, arXiv e-prints arXiv:1904.13392 (2019), 1904.13392.
 - [37] T. Namikawa, D. Hanson, and R. Takahashi, *MNRAS* **431**, 609 (2013), 1209.0091.
 - [38] T. Namikawa and R. Takahashi, *MNRAS* **438**, 1507 (2014), 1310.2372.
 - [39] M. Bucher, C. S. Carvalho, K. Moodley, and M. Remazeilles, *Phys. Rev. D* **85**, 043016 (2012), 1004.3285.
 - [40] H. Prince, K. Moodley, J. Ridl, and M. Bucher, *Journal of Cosmology and Astro-Particle Physics* **2018**, 034 (2018), 1709.02227.
 - [41] A. Cooray, *New Astronomy* **9**, 173 (2004), astro-ph/0309301.
 - [42] E. Schaan, S. Ferraro, and D. N. Spergel, *Phys. Rev. D* **97**, 123539 (2018), 1802.05706.
 - [43] C. Feng and G. Holder, arXiv e-prints arXiv:1905.02084 (2019), 1905.02084.
 - [44] A. J. Mead, J. A. Peacock, C. Heymans, S. Joudaki, and A. F. Heavens, *MNRAS* **454**, 1958 (2015), 1505.07833.
 - [45] D. Blas, J. Lesgourgues, and T. Tram, *Journal of Cosmology and Astro-Particle Physics* **2011**, 034 (2011), 1104.2933.
 - [46] W. Hu and T. Okamoto, *ApJ* **574**, 566 (2002), astro-ph/0111606.
 - [47] S. Foreman, P. D. Meerburg, A. van Engelen, and J. Meyers, *Journal of Cosmology and Astro-Particle Physics* **2018**, 046 (2018), 1803.04975.
 - [48] T. Lu and U.-L. Pen, *MNRAS* **388**, 1819 (2008), 0710.1108.
 - [49] M. Zaldarriaga and U. Seljak, *Phys. Rev. D* **59**, 123507 (1999), astro-ph/9810257.
 - [50] U.-L. Pen, *New Astronomy* **9**, 417 (2004), astro-ph/0305387.
 - [51] T. Lu, U.-L. Pen, and O. Doré, *Phys. Rev. D* **81**, 123015 (2010), 0905.0499.
 - [52] J. Dunkley, E. Calabrese, J. Sievers, G. E. Addison, N. Battaglia, E. S. Battistelli, J. R. Bond, S. Das, M. J. Devlin, R. Dünner, et al., *Journal of Cosmology and Astro-Particle Physics* **2013**, 025 (2013), 1301.0776.
 - [53] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, et al., arXiv e-prints arXiv:0912.0201 (2009), 0912.0201.
 - [54] A. Pézin, O. Doré, G. Lagache, and M. Béthermin, *A&A* **537**, A137 (2012), 1110.0395.
 - [55] M. Béthermin, E. Daddi, G. Magdis, M. T. Sargent, Y. Hezaveh, D. Elbaz, D. Le Borgne, J. Mullaney, M. Pannella, V. Buat, et al., *ApJ* **757**, L23 (2012), 1208.6512.
 - [56] M. Béthermin, L. Wang, O. Doré, G. Lagache, M. Sargent, E. Daddi, M. Cousin, and H. Aussel, *A&A* **557**, A66 (2013), 1304.3936.
 - [57] J. C. Hill and E. Pajer, *Phys. Rev. D* **88**, 063526 (2013), 1303.4726.
 - [58] L. D. Shaw, D. H. Rudd, and D. Nagai, *ApJ* **756**, 15 (2012), 1109.0553.
 - [59] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, et al., arXiv e-prints arXiv:1807.06209 (2018), 1807.06209.
 - [60] G. de Zotti, M. Massardi, M. Negrello, and J. Wall, *A&A Rev.* **18**, 1 (2010), 0908.1896.
 - [61] M. H. Brookes, P. N. Best, J. A. Peacock, H. J. A. Röttgering, and J. S. Dunlop, *MNRAS* **385**, 1297 (2008), 0802.1158.
 - [62] G. Stein, M. A. Alvarez, and J. R. Bond, *MNRAS* **483**, 2236 (2019), 1810.07727.
 - [63] E. Anderes and J. Guinness, arXiv e-prints arXiv:1603.03496 (2016), 1603.03496.

Appendix A: Derivation of the secondary bias to CMB lensing from lensed foregrounds

We consider the auto-spectrum of $\mathcal{Q}_{\mathbf{L}}[T + f, T + f]$, where T now represents the lensed CMB and f the lensed foreground. The quadratic estimators are by definition bilinear in their arguments, and can thus be expanded. We further Taylor expand $T = T^0 + T^1$ and $f = f^0 + f^1$ to first order in the lensing convergence as follows:

$$T_{\ell} = T_{\ell}^0 + T_{\ell}^1 + \mathcal{O}(\kappa^2), \quad (\text{A1})$$

where

$$T_{\ell}^1 = \int \frac{d^2 \mathbf{L}}{(2\pi)^2} \alpha_{\mathbf{L}, \ell - \mathbf{L}} \kappa_{\mathbf{L}} T_{\ell - \mathbf{L}}^0 \quad (\text{A2})$$

and $\alpha_{\ell_1, \ell_2} = -2 \frac{\ell_1}{\ell_1^2} \cdot \ell_2$. These expressions are valid both for the lensed CMB and a lensed foreground. Finally, we keep only the terms that are first order in $\kappa_f \times \kappa_{\text{CMB}}$.

We thus get:

$$\begin{aligned}\delta C_{L_0}^{\kappa^{\text{CMB}}} &= 8\langle \mathcal{Q}_{\mathbf{L}_0} [f^0, f^1] \mathcal{Q}_{\mathbf{L}_0} [T^0, T^1] \rangle' \\ &\quad + 8\langle \mathcal{Q}_{\mathbf{L}_0} [T^0, f^0] \mathcal{Q}_{-\mathbf{L}_0} [T^1, f^1] \rangle' \\ &\quad + 8\langle \mathcal{Q}_{\mathbf{L}_0} [T^0, f^1] \mathcal{Q}_{-\mathbf{L}_0} [T^1, f^0] \rangle',\end{aligned}\tag{A3}$$

where $\langle \dots \rangle' \equiv \langle \dots \rangle / (2\pi)^2 \delta^D(\vec{0})$. The first line corresponds to the primary bias, and we refer to the additional terms on line 2 and 3 as secondary bias. They can be expressed explicitly as:

$$\begin{aligned}8\langle \mathcal{Q}_{\mathbf{L}_0} [T^0, f^0] \mathcal{Q}_{-\mathbf{L}_0} [T^1, f^1] \rangle &= \\ 8N_{\mathbf{L}}^2 \int \frac{d^2 \mathbf{L}}{(2\pi)^2} \int \frac{d^2 \ell}{(2\pi)^2} F_{\ell, \mathbf{L}_0 - \ell} F_{\ell - \mathbf{L} - \mathbf{L}_0, \mathbf{L} - \ell} \alpha_{\mathbf{L}, \ell - \mathbf{L}} \alpha_{-\mathbf{L}, \ell - \mathbf{L}_0} C_{\mathbf{L}_0 - \ell}^{\text{CMB } 0} C_{\ell - \mathbf{L}}^{f0} C_{\mathbf{L}}^{\kappa^{\text{CMB } \kappa^f}}\end{aligned}\tag{A4}$$

and

$$\begin{aligned}8\langle \mathcal{Q}_{\mathbf{L}_0} [T^0, f^1] \mathcal{Q}_{-\mathbf{L}_0} [T^1, f^0] \rangle &= \\ 8N_{\mathbf{L}}^2 \int \frac{d^2 \mathbf{L}}{(2\pi)^2} \int \frac{d^2 \ell}{(2\pi)^2} F_{\ell, \mathbf{L}_0 - \ell} F_{\ell - \mathbf{L}, \ell - \mathbf{L} - \mathbf{L}_0} \alpha_{\mathbf{L}, \ell - \mathbf{L}} \alpha_{-\mathbf{L}, \mathbf{L}_0 + \mathbf{L} - \ell} C_{\ell - \mathbf{L}}^{\text{CMB } 0} C_{\ell - \mathbf{L} - \mathbf{L}_0}^{f0} C_{\mathbf{L}}^{\kappa^{\text{CMB } \kappa^f}}.\end{aligned}\tag{A5}$$

Eq. (21) is then obtained by adding these two terms together, and makes apparent their exact cancellation in the limit $L, L_0 \ll \ell$, where large-scale lensing modes are reconstructed from small-scale temperature modes.

Appendix B: Comparison of analytical and numerical methods

1. Analytical evaluation methods

For the bias in CMB lensing cross-correlation, and for the primary and secondary biases to CMB lensing, we evaluate the response $\mathcal{R}_{\mathbf{L}}^f$ from Eq. (14) by using the Fast Fourier Transform (FFT). Indeed, each integral is a sum of products and convolutions, which can be computed efficiently by successive products and FFT steps. For the secondary bias, we use the python package `vegas`² to compute the 4d integral with a Monte Carlo method. All these calculations are performed in the publicly available repository <https://github.com/EmmanuelSchaan/LensedForegroundBias>, building upon the `ForQuE` module³.

2. Simulations

In addition to the analytical calculations describe above, we also evaluate the foreground lensing responses, primary and secondary biases using simulated maps of the CMB, foregrounds and their lensing convergences. These calculations build upon the `LensQuEst` module⁴. We generate Gaussian random fields (GRF) for the unlensed CMB T^0 and the unlensed foreground f^0 . We simulate the CMB lensing convergence κ_{CMB} and foreground lensing convergence κ_f as correlated GRF. Using GRFs for these simulated maps is sufficient, since the lensed foreground biases we consider in this paper depend only on the foreground power spectra, and not on the detail of their non-Gaussian statistics.

We generate correlated GRFs as follows. To generate GRF maps m_1 and m_2 with auto-spectra C_L^{11} and C_L^{22} and cross-spectrum C_L^{12} , we first generate a GRF m_1 with power spectrum C_L^{11} , then define $m_2 \equiv \alpha m_1 + m_{\perp}$, where $\alpha_L = C_L^{12}/C_L^{11}$ and m_{\perp} is a GRF with power spectrum $C_L^{22} - (C_L^{12})^2/C_L^{11}$. This produces GRF maps m_1 and m_2 with the correct auto and cross-spectra.

At this point, we could simply lens the CMB and foreground maps with their respective convergences, add them, and apply the lensing estimators to the resulting map. However, one would then need to perform N^0 subtraction, and the result would also be affected by a large noise from the lensing reconstruction. Measuring the percent-level bias due to lensed foregrounds, with a precision of say a few percent, would then require a large amount of simulations.

² <https://pypi.org/project/vegas/>

³ <https://github.com/EmmanuelSchaan/ForQuE>

⁴ <https://github.com/EmmanuelSchaan/LensQuEst>

Instead, we use the trick of lensing the CMB and foreground maps only to first order. In other words, we Taylor expand the lensed temperature map in powers of the convergence field: $T = T^0 + T^1 + \mathcal{O}(\kappa^2)$, where T^0 is the unlensed map, and $T^1 \equiv \vec{\nabla}\phi \cdot \vec{\nabla}T^0$ is the first order lensing correction. We perform the same operation for the foreground map f .

Applying the lensing estimators directly on the total map, CMB plus foreground, lensed to first order, would still require a N^0 subtraction, and would actually make the noise worse [63]. Instead, we evaluate the lensing responses, primary and secondary biases as follows:

$$\text{Response: } \mathcal{R}_L^f = \frac{\langle (\mathcal{Q}[f_0, f_1] + \mathcal{Q}[f_1, f_0]) \kappa_f \rangle}{\langle \kappa_f \kappa_f \rangle} \quad (\text{B1})$$

$$\text{Primary bias: } \frac{\delta C_L^{\kappa_{\text{CMB}} \kappa_{\text{CMB}}}}{C_L^{\kappa_{\text{CMB}} \kappa_{\text{CMB}}}} = 2 \frac{\langle (\mathcal{Q}[f_0, f_1] + \mathcal{Q}[f_1, f_0]) \kappa_{\text{CMB}} \rangle}{\langle \kappa_{\text{CMB}} \kappa_{\text{CMB}} \rangle} \quad (\text{B2})$$

$$\begin{aligned} \text{Secondary bias: } \frac{\delta C_L^{\kappa_{\text{CMB}} \kappa_{\text{CMB}}}}{C_L^{\kappa_{\text{CMB}} \kappa_{\text{CMB}}}} &= 2 \frac{\langle (\mathcal{Q}[f_0, T_0] + \mathcal{Q}[T_0, f_0]) (\mathcal{Q}[f_1, T_1] + \mathcal{Q}[T_1, f_1]) \rangle}{\langle \kappa_{\text{CMB}} \kappa_{\text{CMB}} \rangle} \\ &+ 2 \frac{\langle (\mathcal{Q}[f_1, T_0] + \mathcal{Q}[T_0, f_1]) (\mathcal{Q}[f_0, T_1] + \mathcal{Q}[T_1, f_0]) \rangle}{\langle \kappa_{\text{CMB}} \kappa_{\text{CMB}} \rangle} \end{aligned} \quad (\text{B3})$$

Here \mathcal{Q} represents any quadratic lensing estimator, such as the QE, shear and magnification estimators. This method was found to be much less noisy than the naïve approach of simply adding the lensed CMB and foregrounds and applying the lensing estimators to the sum. Furthermore, it does not involve any auto-spectrum, so no noise bias subtraction is needed. This trick relies on the fact that the quadratic estimators are by construction unbiased when applied to T^0 and T^1 .

For each foreground, we simulate 8060 flat square maps, of 30 degrees and 1200 pixels on the side. For the biases from lensed tSZ \times CIB, we ran only 6200 such simulations, due to limits in computing time. This high resolution ensures that Fourier modes up to $2\ell_{\text{max, T}}$ are correctly sampled in the maps, in order to avoid aliasing when nonlinear operations such as the first order lensing or the quadratic estimators are performed on the maps.

Appendix C: Bias hardening against a lensed foreground

In this section, we derive a modification to the standard CMB lensing quadratic estimator which nulls its response to foreground lensing. We do this in the case of the standard quadratic estimator, but the same method can be generalized to the shear and magnification estimators. As in the rest of the paper, we describe foreground lensing as a single foreground map being lensed by a single convergence map. This ignores the fact that each redshift slice of the foreground map is lensed by a slightly different lensing convergence field.

For a given lensing multipole \mathbf{L} , the quantity $\hat{\kappa}_{\mathbf{L}, \ell} \equiv \frac{\ell^T \mathbf{L} \cdot \ell}{f_{\text{CMB}}^{\text{CMB}}}$ is an unbiased estimator of $\kappa_{\mathbf{L}}^{\text{CMB}}$, with variance $\sigma_{\mathbf{L}, \ell}^2 = \frac{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}}{f_{\text{CMB}}^{\text{CMB}}}$. We denote $\sum_{\ell} \equiv \int \frac{d^2\ell}{(2\pi)^2}$, and consider estimators of the form $\hat{\kappa} = \sum_{\ell} w_{\ell} \hat{\kappa}_{\ell}$.

1. Standard quadratic estimator

The standard quadratic estimator of [46] is simply the minimum variance unbiased linear combination of these estimators. Indeed, we look for weights $w_{\mathbf{L}, \ell}$ satisfying:

$$\begin{cases} \sum_{\ell} w_{\mathbf{L}, \ell} = 1 & \text{(unit response to CMB lensing)} \\ \sum_{\ell} w_{\mathbf{L}, \ell}^2 \sigma_{\mathbf{L}, \ell}^2 \text{ is minimal} & \text{(minimum variance)} \end{cases} \quad (\text{C1})$$

This problem of minimization under constraints can be solved with Lagrange multipliers. We thus minimize the following quantity with respect to the lensing weights $w_{\mathbf{L},\ell}$ and the Lagrange multiplier α :

$$\mathcal{L} = \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell}^2 \sigma_{\mathbf{L},\ell}^2 \right)}_{\text{min. var.}} + \alpha \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell} - 1 \right)}_{\text{unit response to } \kappa_{\mathbf{L}}^{\text{CMB}}}. \quad (\text{C2})$$

This gives:

$$\hat{\kappa}_{\mathbf{L}} \equiv \frac{\sum_{\ell} \hat{\kappa}_{\mathbf{L},\ell} / \sigma_{\mathbf{L},\ell}^2}{\sum_{\ell} 1 / \sigma_{\mathbf{L},\ell}^2} = \frac{\int \frac{d^2\ell}{(2\pi)^2} T_{\ell} T_{\mathbf{L}-\ell} \frac{f_{\ell,\mathbf{L}-\ell}^{\text{CMB}}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}}}{\int \frac{d^2\ell}{(2\pi)^2} \frac{f_{\ell,\mathbf{L}-\ell}^{\text{CMB}^2}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}}}, \quad (\text{C3})$$

which is indeed the standard quadratic estimator of [46].

2. Nulling the response to foreground lensing

In the presence of a lensed foreground, the estimator $\hat{\kappa}_{\mathbf{L},\ell}$ acquires a bias $\mathcal{R}_{\mathbf{L},\ell}^f \kappa_{\mathbf{L}}^f$, where $\mathcal{R}_{\mathbf{L},\ell}^f = \frac{f_{\ell,\mathbf{L}-\ell}^f}{f_{\ell,\mathbf{L}-\ell}^{\text{CMB}}}$. Here we would like the combined estimator to have zero response to $\kappa_{\mathbf{L}}^f$. We thus look for lensing weights $w_{\mathbf{L},\ell}$ such that:

$$\begin{cases} \sum_{\ell} w_{\mathbf{L},\ell} = 1 & \text{(unit response to CMB lensing)} \\ \sum_{\ell} w_{\mathbf{L},\ell}^2 \sigma_{\mathbf{L},\ell}^2 \text{ is minimal} & \text{(minimum variance)} \\ \sum_{\ell} w_{\mathbf{L},\ell} \mathcal{R}_{\mathbf{L},\ell}^f = 0 & \text{(zero response to foreground lensing)} \end{cases} \quad (\text{C4})$$

We thus minimize the following quantity with respect to the lensing weights $w_{\mathbf{L},\ell}$ and the Lagrange multipliers α and β :

$$\mathcal{L} = \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell}^2 \sigma_{\mathbf{L},\ell}^2 \right)}_{\text{min. var.}} + \alpha \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell} - 1 \right)}_{\text{unit response to } \kappa_{\mathbf{L}}^{\text{CMB}}} + \beta \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell} \mathcal{R}_{\mathbf{L},\ell}^f \right)}_{\text{zero response to } \kappa_{\mathbf{L}}^f}. \quad (\text{C5})$$

The solution is:

$$w_{\mathbf{L},\ell} \equiv \frac{1}{(S_0 S_2 - S_1^2)} \frac{(S_2 - S_1 \mathcal{R}_{\mathbf{L},\ell}^f)}{\sigma_{\mathbf{L},\ell}^2}, \quad (\text{C6})$$

where the quantities S_n are the following functions of L :

$$S_n \equiv \int \frac{d^2\ell}{(2\pi)^2} \left(\mathcal{R}_{\mathbf{L},\ell}^f \right)^n \frac{f_{\ell,\mathbf{L}-\ell}^{\text{CMB}^2}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}}. \quad (\text{C7})$$

More explicitly, the lensed foreground bias hardened estimator is:

$$\hat{\kappa}_{\mathbf{L}} \equiv \frac{1}{(S_0 S_2 - S_1^2)} \int \frac{d^2\ell}{(2\pi)^2} T_{\ell} T_{\mathbf{L}-\ell} \frac{f_{\ell,\mathbf{L}-\ell}^{\text{CMB}}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}} \left(S_2 - S_1 \mathcal{R}_{\mathbf{L},\ell}^f \right), \quad (\text{C8})$$

This estimator indeed has unit response to CMB lensing, zero response to the foreground lensing, and reduces to the standard quadratic estimator [46] if the response $\mathcal{R}_{\mathbf{L},\ell}^f$ to foreground lensing is identically zero.

However, this lensed foreground bias hardened estimator has a larger noise power spectrum than the QE, given by:

$$N_{\hat{\mathbf{L}}}^{\kappa} = \frac{1}{(S_0 S_2 - S_1^2)^2} \int \frac{d^2 \ell}{(2\pi)^2} \frac{f_{\ell, \mathbf{L}-\ell}^{\text{CMB}^2}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}} \left(S_2 - S_1 \mathcal{R}_{\mathbf{L},\ell}^f \right)^2. \quad (\text{C9})$$

In principle, this approach could increase the noise by a large amount, in order to subtract the small bias due to foreground lensing.

3. Minimizing the total variance from noise plus lensed foreground bias

An alternative approach is thus not to require the response to foreground lensing to be exactly zero, but instead to minimize the total variance of the lensing estimator, including the additional variance due to the foreground lensing. This would make sure that we are not subtracting a $\sim 1\%$ foreground lensing bias at a large cost in signal-to-noise. In other words, we look for weights $w_{\mathbf{L},\ell}$ such that:

$$\begin{cases} \sum_{\ell} w_{\mathbf{L},\ell} = 1 \text{ (unit response to CMB lensing)} \\ \sum_{\ell} w_{\mathbf{L},\ell}^2 \sigma_{\mathbf{L},\ell}^2 + 2 \left(\sum_{\ell} w_{\mathbf{L},\ell} \mathcal{R}_{\mathbf{L},\ell}^f \right) \left(\sum_{\ell} w_{\mathbf{L},\ell'} \right) C_L^{\kappa_f \kappa_{\text{CMB}}} \text{ is minimal} \end{cases} \quad (\text{C10})$$

The second line minimizes the sum of the noise variance plus the primary lensed foreground bias. This ignores the secondary lensed foreground bias, which is important on the smaller lensing scales. Again, we minimize the quantity:

$$\mathcal{L} = \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell}^2 \sigma_{\mathbf{L},\ell}^2 \right)}_{\text{min. noise var.}} + 2 \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell} \mathcal{R}_{\mathbf{L},\ell}^f \right) \left(\sum_{\ell} w_{\mathbf{L},\ell'} \right) C_L^{\kappa_f \kappa_{\text{CMB}}}}_{\text{primary lensed foreground bias}} + \alpha \underbrace{\left(\sum_{\ell} w_{\mathbf{L},\ell} - 1 \right)}_{\text{unit response to } \kappa_{\mathbf{L}}^{\text{CMB}}}. \quad (\text{C11})$$

The solution is:

$$w_{\mathbf{L},\ell} \equiv \frac{1}{S_0 \sigma_{\mathbf{L},\ell}^2} \left[1 + C_L^{\kappa_f \kappa_{\text{CMB}}} \left(S_0 \mathcal{R}_{\mathbf{L},\ell}^f - S_1 \right) \right] \quad (\text{C12})$$

The corresponding estimator is then:

$$\hat{\kappa}_{\mathbf{L}} \equiv \frac{1}{S_0} \int \frac{d^2 \ell}{(2\pi)^2} T_{\ell} T_{\mathbf{L}-\ell} \frac{f_{\ell, \mathbf{L}-\ell}^{\text{CMB}}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}} \left(S_0 \mathcal{R}_{\mathbf{L},\ell}^f - S_1 \right). \quad (\text{C13})$$

Then estimator indeed has unit response to CMB lensing. It reduces to the standard quadratic estimator [46] when $\mathcal{R}_{\mathbf{L},\ell}^f = 0$. Its noise power spectrum is:

$$N_{\hat{\mathbf{L}}}^{\kappa} = \frac{1}{S_0^2} \int \frac{d^2 \ell}{(2\pi)^2} \frac{f_{\ell, \mathbf{L}-\ell}^{\text{CMB}^2}}{2C_{\ell}^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}}} \left(S_0 \mathcal{R}_{\mathbf{L},\ell}^f - S_1 \right)^2. \quad (\text{C14})$$