Impact of Galactic non-Gaussian foregrounds on CMB lensing measurements

Irene Abril-Cabezas,[1,2,*] Frank J. Qu,[3,4,1,2] Blake D. Sherwin,[1,2] Alexander van Engelen,[5] Niall MacCrann,[1,2] Carlos Hervías-Caimapo,[6] Omar Darwish,[7] J. Colin Hill,[8] Mathew S. Madhavacheril,[9] and Neelima Sehgal[10]

[1]*DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*
[2]*Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA, United Kingdom*
[3]*Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, California 94305, USA*
[4]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, California 94305, USA*
[5]*School of Earth and Space Exploration, Arizona State University, Tempe, Arizona 85287, USA*
[6]*Instituto de Astrofísica and Centro de Astro-Ingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860, 7820436, Macul, Santiago, Chile*
[7]*Université de Genève, Département de Physique Théorique et CAP, 24 Quai Ansermet, CH-1211 Genève 4, Switzerland*
[8]*Department of Physics, Columbia University, New York, New York 10027, USA*
[9]*Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, Pennsylvania 19104, USA*
[10]*Physics and Astronomy Department, Stony Brook University, Stony Brook, New York 11794, USA*

(Received 7 May 2025; accepted 16 June 2025; published 15 July 2025)

Weak gravitational lensing of the cosmic microwave background (CMB) has been established as a robust and powerful observable for precision cosmology. However, the impact of Galactic foregrounds, which has been studied less extensively than many other potential systematics, could in principle pose a problem for CMB lensing measurements. These foregrounds are inherently non-Gaussian and hence might mimic the characteristic signal that lensing estimators are designed to measure. We present an analysis that quantifies the level of contamination from Galactic dust in lensing measurements, focusing particularly on measurements with the Atacama Cosmology Telescope and the Simons Observatory. We employ a whole suite of foreground models and study the contamination of lensing measurements with both individual frequency channels and multifrequency combinations. We test the sensitivity of different estimators to the level of foreground non-Gaussianity and the dependence on sky fraction and multipole range used. We find that Galactic foregrounds do not present a problem for the Atacama Cosmology Telescope experiment (the bias in the inferred CMB lensing power spectrum amplitude remains below 0.3σ). For Simons Observatory, not all foreground models remain below this threshold. Although our results are conservative upper limits, they suggest that further work on characterizing dust biases and determining the impact of mitigation methods is well motivated, especially for the largest sky fractions.

DOI: 10.1103/jhmr-mg6w

I. INTRODUCTION

The cosmic microwave background (CMB) was released when the Universe was approximately 380 000 years old. As CMB photons travel to our observation point, they encounter large-scale structure that gravitationally deflects their trajectories. This phenomenon is known as weak gravitational lensing of the CMB. A measurement of this effect provides an opportunity to map the projected matter density field, probe the growth of structure, constrain the neutrino mass, learn about the geometry of our Universe, or constrain the dark energy model [1,2].

Lensing breaks the statistical isotropy of the CMB, which in turn can be used to reconstruct the lensing deflections using quadratic combinations of the CMB intensity (T) and polarization (E, B) fields [3–5]. In particular, different quadratic estimators (QEs) provide measurements with

*Contact author: ia404@cam.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

varying signal-to-noise ratios (SNRs) depending on the experiment's noise levels. The TT estimator has the highest SNR for an experiment such as Planck (with noise levels around 30 μK-arcmin in temperature) [6–8]. Moreover, the reconstructed deflection field is still dominated by temperature-only information with current CMB experiments such as the Atacama Cosmology Telescope (ACT) [9–12] or the Simons Observatory (SO) [13], all with temperature noise levels at least 3 times lower than Planck. The EB estimator becomes the most powerful and will dominate the lensing reconstruction once experiments reach noise levels well below 5 μK-arcmin [4,14].[1] Next-generation experiments, including CMB stage IV [17,18], will most likely not use quadratic estimators to reconstruct the lensing field (as they will become suboptimal) and will instead favor likelihood-based iterative methods [19–22] [as recently used by the South Pole Telescope (SPT) [23,24]]. Nevertheless, QEs (and in particular TT) will remain a useful tool to forecast and cross-check results.

Alongside lensing, foregrounds are another potential source of statistical anisotropy and non-Gaussianity in CMB maps. This is particularly important in temperature, for which extragalactic foregrounds such as the thermal Sunyaev-Zel'dovich (tSZ) effect [25,26] sourced by galaxy clusters or the millimeter-wave emission by dusty star-forming galaxies [27–29] can significantly bias CMB lensing measurements, as multiple studies (e.g. [30–35]) have shown. A common mitigation strategy in this case is to use "bias-hardened" quadratic estimators, a modified form of the standard QE that has zero response, at leading order, to the statistical anisotropy caused by these sources [30,36,37]. In polarization, extragalactic foreground biases are not expected to be significant [38–40]. Moreover, new methods are being developed to further minimize their impact [41–43].

Galactic foregrounds, on the other hand, have not been studied to the same extent in the CMB lensing literature (with only a few papers discussing the problem, e.g. [44–46]). Nevertheless, they are of particular concern for lensing studies because they are highly non-Gaussian, with interstellar medium structures correlating different scales in a highly anisotropic fashion, and hence could mimic the characteristic signal that lensing estimators are designed to measure. Thermal dust emission from our own Galaxy, in particular, is the most relevant Galactic emission mechanism at frequencies above $\nu \gtrsim 70$ GHz (see e.g. [47–50]). In polarization, Refs. [44,45] find using one non-Gaussian small-scale dust realization that dust produces non-negligible biases and suggest parametric foreground cleaning approaches that successfully mitigate them. Reference [46] performs an approximate estimate of dust biases using a small high-dust region near the Galactic plane and finds a 15% bias in the temperature reconstruction, which is reduced to acceptable levels after foreground cleaning; the authors note, however, that some caution is required in interpreting this result more broadly since the these are specific to the bright dust field they employ.

Compared to Planck, ground-based experiments have a reduced frequency coverage which hampers component-separation techniques.[2] This results in different strategies to minimize foreground contamination. While Planck used foreground-cleaned CMB maps [7,8], lensing reconstruction from ground-based experiments such as ACT [12] or SPT [24] only use CMB maps from at most three frequency bands, centered at 90, 150 and 220 GHz. These analyses performed a whole suite of null tests that ensured minimal contamination from extragalactic foregrounds (e.g. [34]). Since detailed estimates of Galactic foreground biases were not available, sky fractions were chosen very conservatively, potentially discarding a large fraction of usable data. A detailed quantification of the level of bias on these experiments is hence well motivated.

In this work, we focus on the impact of Galactic dust non-Gaussianity on lensing reconstructions performed with quadratic estimators. Estimating the level of bias has been difficult in the past because, as pointed out in [46], the dust field on small scales is not well constrained by current data, especially in polarization and at high Galactic latitudes where the dust emission is reduced. When simulating the foreground field, this issue is usually circumvented by filtering out these scales and replacing them with Gaussian realizations of a given target power spectrum that follows the trend measured on large scales (see e.g. [51,52]). However, this solution does not account correctly for non-Gaussianity on small scales. Fortunately, several groups have produced foreground simulations including non-Gaussianities on small scales in the last few years. These include magnetohydrodynamic simulations (e.g. [53,54]), phenomenological models based on temperature observations using a superposition of Galactic dust layers (e.g. [55,56]), models built from wavelet scattering transform statistics [57–59] or generative adversarial neural networks [60,61], dust models based on neutral hydrogen [62], and filament-based ones [63,64]. In particular, the PySM software [52,65] provides a suite of models spanning a range of complexities on the scales where there are little data available, hence making it an ideal tool to check for Galactic foreground biases in a variety of physical scenarios regarding the foreground properties. In this work, we will use a whole range of simulations to estimate the bias to the lensing power spectrum due to Galactic foregrounds.

[1]CMB temperature maps again become the dominant contributor to the SNR for a CMB experiment such as CMB-HD [15,16].

[2]This will soon change with the upcoming Simons Observatory, operating on six frequency bands centered around 27, 39, 93, 145, 225 and 280 GHz [13].

Previous studies in the literature have focused on previous [66] or longer-term future CMB missions [44–46]. Instead, we will explore the impact it has on lensing from current-generation experiments, in both temperature and polarization, and for a whole suite of dust models.

The paper is organized as follows. We present the general methodology of lensing reconstruction in Sec. II, with a particular focus on the approach taken in this work. In Sec. III, we present upper limits to the level of Galactic dust contamination on CMB lensing reconstruction analyses and their bias to the lensing amplitude parameter. We conclude in Sec. IV. We assume the following default cosmology: $H_0 = 67.02\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_b h^2 = 0.0222$, and $\Omega_c h^2 = 0.1203$.

II. METHODOLOGY

A. Formalism of lensing reconstruction with quadratic estimators

Weak lensing of the CMB leads to a remapping of the observed CMB sky, which can be expressed to leading order in ϕ as [2,67]

$$T(\hat{\mathbf{n}}) = \tilde{T}(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) \simeq \tilde{T}(\hat{\mathbf{n}}) + \nabla\tilde{T}(\hat{\mathbf{n}}) \cdot \nabla\phi(\hat{\mathbf{n}}), \tag{1}$$

where T is the lensed temperature field and $\tilde{T}$ is the unlensed temperature field.[3] The lensing deflection angle $\nabla\phi$ is expressed in terms of the lensing potential ϕ, itself an integral of the gravitational potential[4] along the line of sight $\hat{\mathbf{n}}$. From the equation above, we see that lensing changes the statistics of the CMB, inducing correlations between the temperature field and its gradient. When averaging over all possible unlensed CMB realizations, nonopposite $\mathbf{l}$ modes in the CMB now become correlated due to a $\phi(\mathbf{L})$ lensing mode:

$$\langle T(\mathbf{l})T(\mathbf{L}-\mathbf{l})\rangle_{\mathrm{CMB}} = f(\mathbf{l},\mathbf{L})\phi(\mathbf{L}), \qquad \mathbf{L} \neq 0, \tag{3}$$

where $f(\mathbf{l},\mathbf{L})$ is a function that depends on the angular power spectrum of the unlensed CMB temperature.[5] This mode coupling is a reflection of the fact that lensing (with a fixed lensing field) breaks translation invariance in the CMB. Quadratic estimators [4] exploit this mode-coupling property to provide an estimate for $\phi(\mathbf{L})$:

$$\hat{\phi}(\mathbf{L}) = A^{\phi}_{\mathbf{L}} \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \frac{f(\mathbf{l},\mathbf{L})T(\mathbf{l})T(\mathbf{L}-\mathbf{l})}{2C^{\mathrm{obs}}_{\ell} C^{\mathrm{obs}}_{|\mathbf{L}-\mathbf{l}|}}, \tag{4}$$

where C^{obs}_{ℓ} is the observed (lensed) CMB spectrum including noise. The quadratic estimator is normalized as follows:

$$A^{\phi}_{\mathbf{L}} = \left[\int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \frac{f^2(\mathbf{l},\mathbf{L})}{2C^{\mathrm{obs}}_{\ell} C^{\mathrm{obs}}_{|\mathbf{L}-\mathbf{l}|}}\right]^{-1}. \tag{5}$$

It can be shown that, if run on a map containing all the expected sources of power, $A^{\phi}_{\mathbf{L}}$ is equivalent to the CMB lensing power spectrum reconstruction noise, also known as the $N^{(0)}_L$ bias [3,70,71] (see also Appendix A). Below, we will perform lensing reconstruction on simulated foreground-only maps, in order to estimate the resulting bias to the lensing power spectrum. In doing this, we implicitly assume that the dust and lensed CMB maps are entirely independent, such that we only need to compute the dust trispectrum to estimate the biases to our lensing spectrum measurement. In this case, the foreground-only map has power C^{fg}_{ℓ}, rather than C^{obs}_{ℓ}. In this case, the $N^{(0)}_L$ bias takes the form [34]

$$N^{(0)}_L = (A^{\phi}_{\mathbf{L}})^2 \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \frac{f^2(\mathbf{l},\mathbf{L})C^{\mathrm{fg}}_{\ell} C^{\mathrm{fg}}_{|\mathbf{L}-\mathbf{l}|}}{2(C^{\mathrm{obs}}_{\ell} C^{\mathrm{obs}}_{|\mathbf{L}-\mathbf{l}|})^2}. \tag{6}$$

This result is derived explicitly in Appendix A. Finally, we note that one can work equivalently with the lensing potential ϕ or the convergence κ. They are related by

$$\kappa(\hat{\mathbf{n}}) = -\nabla^2\phi(\hat{\mathbf{n}})/2. \tag{7}$$

From Poisson's equation, κ measures the projected fractional matter overdensity, instead of the projected gravitational potential traced by ϕ. In terms of two-point functions,

$$C^{\kappa\kappa}_L = \left(\frac{L(L+1)}{2}\right)^2 C^{\phi\phi}_L. \tag{8}$$

The results presented in this section are derived in the flat-sky approximation to aid the reader. In practice, we use curved-sky calculations as implemented in FALAFEL [72]. Similarly, lensing deflections can be reconstructed not only from the CMB temperature field, but also using different quadratic combinations of the CMB intensity (T) and polarization (E, B). These can be combined into a minimum-variance MV estimator [5] to improve the

[3]Similar expressions can be derived for the polarization component of the CMB.

[4]More precisely, CMB lensing measurements are sensitive to the Weyl potential $\Psi_{\mathrm{W}} \equiv (\Phi + \Psi)/2$, where Ψ and Φ are the temporal and spatial distortions, respectively, of the Universe's geometry:

$$\mathrm{d}s^2 = a^2(\tau)[-(1+2\Psi)\mathrm{d}\tau^2 + (1-2\Phi)\mathrm{d}\mathbf{x}^2]. \tag{2}$$

The so-called anisotropic stress describes the difference between the time distortion and the spatial distortion of the metric. This is usually assumed to be negligible, in which case we have $\Psi_{\mathrm{W}} = \Psi = \Phi$.

[5]In practice, we use the lensed gradient power spectra $C^{X\nabla Y}_{\ell}$ for $X, Y = T, E, B$ to reduce higher-order corrections needed to match the nonperturbative result [68,69].

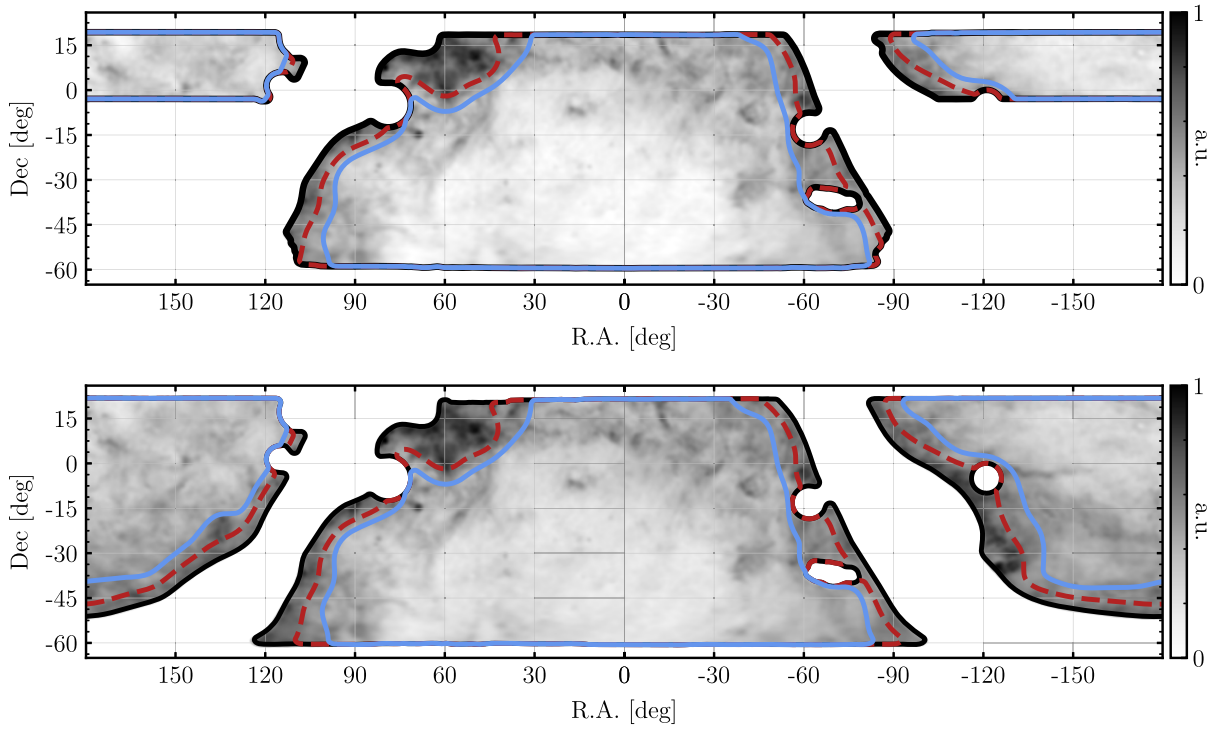


FIG. 1. Galactic masks for the two surveys explored in this work, projected in equatorial coordinates. They correspond to the intersection of the survey footprint with Planck masks labeled as XX%, where XX denotes the sky fraction of the Planck mask in percent. Their edges are apodized on 3° scales with a cosine profile, and regions with clearly visible Galactic dust clouds are also removed, following [12,76]. For reference, we also show Planck 353 GHz GNILC dust emission map [49], where the color bar (in arbitrary units, a.u.) indicates the strength of the emission. Top: ACT + Planck 60% (solid blue line), ACT + Planck 70% (dashed red line) and ACT + Planck 80% (solid black line) masks, with effective sky fractions $f_{\rm sky} = 0.28$, $f_{\rm sky} = 0.31$, and $f_{\rm sky} = 0.34$, respectively. Bottom: SO + Planck 60% (solid blue line), SO + Planck 70% (dashed red line) and SO + Planck 80% (solid black line) masks, with effective sky fractions $f_{\rm sky} = 0.37$, $f_{\rm sky} = 0.42$ and $f_{\rm sky} = 0.47$.

signal-to-noise ratio. When only polarization-derived measurements are used, the minimum-variance estimator is denoted MVPol. We use TEMPURA [73] to compute the normalizations of all estimators and perform lensing reconstructions with SO-LENSPIPE [74].

B. Experimental setup

We explore the impact of Galactic dust on lensing reconstruction for two different experimental configurations. Firstly, we focus on an ACT-like experiment, comprising two frequency bands centered at 90 and 150 GHz. The footprint for such analysis is constructed by intersecting the ACT footprint with Planck masks,[6] and further removing regions with clearly visible Galactic dust clouds, as described in [12]. Planck masks are labeled as XX%, where XX denotes their sky fraction (in percent). The resulting masks are ACT + Planck 60%, ACT + Planck 70% and ACT + Planck 80%. For short, we refer to them as ACT 60%, ACT 70% and ACT 80%. They have effective sky fractions $f_{\rm sky} = 0.28$, $f_{\rm sky} = 0.31$, and $f_{\rm sky} = 0.34$, respectively, after apodizing the edges on 3° scales with a cosine profile. These are shown in the top panel of Fig. 1.

We take the Simons Observatory experiment, as described in [13], as the second experimental setup that we explore. We assume six SO frequency bandpasses, centered around 27 and 39 GHz [low-frequency bands (LF)], 93 and 145 GHz (midfrequency bands), and 225 and 280 GHz (ultrahigh-frequency bands). We focus on the Large Aperture Telescope (LAT) survey of SO. Similar to ACT, we construct three Galactic masks for this configuration by intersecting the SO-LAT footprint[7] with *Planck* 60%, 70% and 80% masks and further removing the regions flagged by ACT [12]. Their edges are also apodized on 3° scales with a cosine profile, and the resulting sky fractions are $f_{\rm sky} = 0.37$, $f_{\rm sky} = 0.42$ and $f_{\rm sky} = 0.47$, respectively. The masks are shown in the bottom panel of Fig. 1. Similar

[6]HFI_Mask_GalPlane-apo0_2048_R2.00.fits, available on the Planck Legacy Archive [75]. This is the most straightforward way to remove Galactic contamination.

[7]LAT_hits_2048_galcoords.fits, available from NERSC upon request.

to the labeling used for ACT masks, we refer to SO masks (for short) as SO 60%, SO 70% and SO 80%.

In this work, for the purposes of filtering and normalizing the quadratic estimator as above [Eq. (4)], we will make the simplifying assumption that the total power in the maps can be expressed as

$$C_\ell^{\rm obs} = C_\ell + \Delta^2 e^{\ell(\ell+1)\sigma^2_{\rm FWHM}/8\ln 2}, \tag{9}$$

where C_ℓ is the theoretical lensed power spectrum (computed using CAMB [77]). Δ is the white-noise level of the ground-based experiment. We consider two different white-noise levels. In temperature, $\Delta_T = 12$ $\mu{\rm K_{CMB}}$-arcmin (ACT-like) or $\Delta_T = 6$ $\mu{\rm K_{CMB}}$-arcmin (SO-like). Both have angular resolution $\sigma_{\rm FWHM} = 1.4'$. In polarization, we assume that the white-noise levels are a factor $\sqrt{2}$ larger than in temperature.[8]

C. Non-Gaussian Galactic sky model

Before summarizing the foreground models used in this work, we will introduce the common empirical relations employed to describe the Galactic dust component. These same relations are used when simulating the dust field. We work in thermodynamic temperature units, $\mu{\rm K_{CMB}}$, for which spectral energy distribution (SED) of the CMB as function of frequency ν is a constant by definition:

$$S_\nu^{\rm CMB} = 1. \tag{10}$$

The SED of thermal dust emission (d) is often modeled as a modified blackbody spectrum (see e.g. [49]):

$$S_\nu^d = \frac{g(x_{\nu_0^d})}{g(x_\nu)}\left(\frac{\nu}{\nu_0^d}\right)^{\beta_d}\frac{B_\nu(\Theta_d)}{B_{\nu_0^d}(\Theta_d)}, \tag{11}$$

where β_d is the dust spectral index, $\nu_0^d = 353$ GHz is the pivot frequency, and $B_\nu(\Theta)$ is the Planck blackbody spectrum, which expresses the spectral radiance of a blackbody for frequency ν at temperature Θ:

$$B_\nu(\Theta) = \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{h\nu}{k\Theta}\right) - 1\right]^{-1}, \tag{12}$$

and where k and h are the Boltzmann and Planck constants, respectively. $g(x_\nu)$ is the conversion factor between Rayleigh-Jeans brightness temperature units (commonly used for foregrounds) and thermodynamic temperature units [78]:

$$g(x_\nu) = e^x\left(\frac{x}{e^x-1}\right)^2, \qquad x = \frac{h\nu}{k\Theta_{\rm CMB}}, \tag{13}$$

where $\Theta_{\rm CMB} = 2.7255K$ is the CMB temperature [79].

Regardless of the exact values that Θ_d and β_d take within the different dust models, a simple evaluation of Eq. (11) yields that dust intensity at 150 GHz is ∼3 times brighter than in the 90 GHz channel. As a result, we expect biases in the lensing reconstruction to be almost 2 orders of magnitude larger in the 150 GHz channel than in the 90 GHz one, since we are working with a four-point function of the dust amplitude.

Empirically, the dust (d) autospectrum is often parametrized as a power law:

$$\frac{\ell(\ell+1)}{2\pi}C_\ell^{dd} = A_d\left(\frac{\ell}{\ell_0}\right)^{\alpha_d}, \tag{14}$$

where A_d is the amplitude, α_d the tilt, and $\ell_0 = 80$ the pivot scale. Planck measures these spectra up to $\ell \sim 1000$ [47–50]; hence, the models of Galactic foreground emission that we use in this work are driven by the existing data on those scales. Planck small-scale measurements are noise dominated, especially in polarization, and hence unable to inform any modeling.[9] Different models then have different strategies to generate the (non-Gaussian) small scales, which we now describe.[10]

(i) PySM d9. PySM models are based on Planck dust templates [49,88,89], which are constructed using the generalized needlet internal linear combination (GNILC) algorithm [90]. The models are calibrated to observations at 353 GHz in the Planck GAL097 mask.[11] PySM d9 is the lowest complexity model provided by [65]. It introduces non-Gaussian small-scale fluctuations in A_d by modulating the small-scale emission by the large-scale signal. The spectral parameters remain uniform across the sky ($\beta_d = 1.54, \Theta_d = 20K$), resulting in 100% correlation across frequency channels.

(ii) PySM d10. Similar model construction as to PySM d9. This model presents an additional complexity, where small-scale fluctuations are present both in dust amplitude and spectral parameters. The latter leads to frequency decorrelation.

(iii) PySM d12. PySM implementation of [56], where the final model is comprised of a superposition of six Galactic dust layers. The dust amplitudes on large

[8] Any experiment sensitive to polarization will always need ×2 more observations to achieve the same level of sensitivity in temperature than in polarization, since polarization measurements require solving for two (Q, U) fields.

[9] Nevertheless, several works have studied different approaches to quantify the level of non-Gaussianity and statistical anisotropy of Galactic foregrounds; see e.g. [80–82].

[10] We also considered including TIGRESS [83–85] synthetic dust models, or those produced by [86] or [87]. However, they did not have enough resolution for our purposes.

[11] Planck masks are labeled as GALXXX, where XXX denotes the sky fraction in percent.

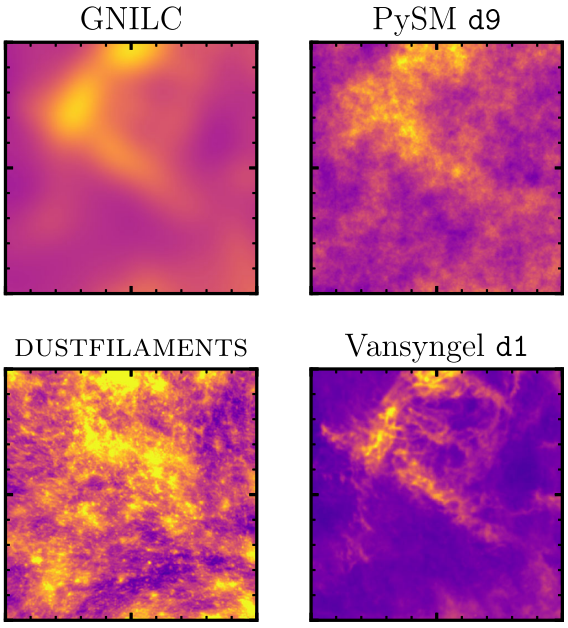


FIG. 2. 10×10 deg^2 enlarged regions centered at $(\mathrm{RA}, \mathrm{DEC}) = (-42°, -5°)$ for PySM d9, DUSTFILAMENTS and Vansyngel d1 dust emission models in temperature at 353 GHz (in arbitrary units). They all trace the Planck GNILC map [88] (upper left corner) on the largest scales and include small-scale features unresolved with current data.

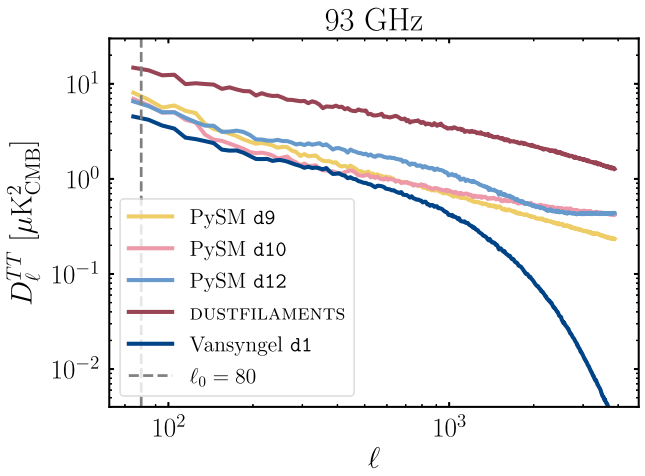


FIG. 3. TT power spectra at 93 GHz, computed on the SO 70% mask, for the variety of dust models used in this work. Even though all models are calibrated using 353 GHz Planck data at the pivot scale $\ell_0 = 80$, we see a spread at the level of 0.4dex in power. This is mainly due to differences in both the sky footprint and template used to calibrate the models, and the spectral energy distribution assumed to extrapolate between frequencies (see text for further details). The power in Vansyngel + 17 is truncated above $\ell \sim 1000$ due to the presence of a Gaussian beam with $\sigma_{\rm FWHM} = 30'$. Consequently, the constraints derived from this model will be slightly optimistic.

scales follow Planck data [50,88] and the spectral parameters are generated randomly within each layer [56].

(iv) DUSTFILAMENTS [64,91,92]. Galactic dust exhibits a filamentary structure that cannot be reproduced from random Gaussian statistics. These filaments have been observed at millimeter wavelengths by Planck (e.g. [93,94]), among others, and also found to be traced by Galactic neutral hydrogen emission [62,95–97]. Filaments and their misalignment with respect to the Galactic magnetic field can quantitatively reproduce the properties of Galactic dust as observed by Planck at 353 GHz [50,63,98]. Based on this idea, DUSTFILAMENTS constructs a full-sky simulation of the Galactic dust at millimeter frequencies by integrating the emission of millions of filaments in a predefined volume. By construction, the filament population is calibrated to match the dust power spectrum on the SO-LAT footprint that intersects with Planck 80% Galactic mask, as described in [64,92], with spectral parameters derived from the GNILC best-fit β_d and Θ_d maps [88].

(v) Vansyngel d1 [55]. Phenomenological model using a superposition of (seven) Galactic dust layers. Every layer follows the same intensity observations provided by Planck GNILC maps [88]. The maps are calibrated to match the Large Region 33 ($f_{\rm sky} = 0.33$) Galactic mask in [50], as described in [55]. In temperature, there is a loss of power above $\ell \sim 1000$ due to the presence of a $30'$ FWHM beam. As a result, the biases derived from this model will be slightly optimistic. This model also introduces EB correlations in the dust field at the level compatible with existing data [98].

For illustrative purposes, Fig. 2 shows 10×10 deg^2 enlarged regions centered at $(\mathrm{RA}, \mathrm{DEC}) = (-42°, -5°)$ for PySM d9, DUSTFILAMENTS and Vansyngel d1 dust emission models (in arbitrary units). They all trace the Planck GNILC map [88] (upper left corner) on the largest scales and include small-scale features unresolved with current data. In Fig. 3, we present the temperature power spectra at 93 GHz for the whole variety of dust models used in this work. These are computed on the SO 70% mask. Even though all models are calibrated using 353 GHz Planck data at the pivot scale $\ell_0 = 80$, we see a spread at the level of 0.4dex in power. This is due to both differences in the sky footprint used to calibrate their models and the spectral energy distribution assumed to extrapolate between frequencies. In addition, the PySM models are calibrated primarily with the Planck GNILC dust template [88,89], while DUSTFILAMENTS is calibrated with the spectra of the 353 GHz Planck NPIPE maps (with the best-fit CMB subtracted). The latter case has more power, which could originate from the cosmic infrared background (CIB), as evidenced by Fig. 4 of [65]. We do

not correct for these differences when reporting the biases that different models produce.

As described earlier, Galactic foregrounds are dominated by dust emission at frequencies above $\nu \gtrsim 70$ GHz (see e.g. [48]). At low $\nu \lesssim 70$ GHz frequencies, synchrotron emission overwhelms the CMB signal. As a result, we must consider synchrotron emission models for SO LF channels. We consider PySM [65] models s4, s5 and s7. These offer varying levels of complexity regarding the modeling of small scales in terms both of amplitude and spectral parameters. They can be combined with PySM d9, d10 and d12, respectively. We follow the PySM convention and denote d9s4, d10s5 and d12s7 as the PySM low-, medium- and high-complexity models, respectively. Vansyngel + 17 [55] also provides a synchrotron realization, produced in a similar procedure as the one used for dust. We denote the sum of those dust and synchrotron simulations as Vansyngel d1s1.

D. Estimation of foreground biases on lensing reconstruction measurements

We run the different quadratic estimators on foreground-only maps and measure the autospectrum of the resulting lensing reconstructions. The resulting measurement is $C_L^{\hat{\phi}\hat{\phi},\mathrm{fg}}$. We treat the foreground-only maps in the same way as CMB maps. In particular, we filter the maps (both in temperature and polarization) with the two-point power defined in Eq. (9), which does not include the power sourced by the Galactic foreground field. We do this to mimic a real lensing analysis, where Galactic foregrounds are typically not included when constructing the filters. This does not cause a significant loss of optimality, since the level of dust power predicted by the sky models we use is below the experiment's observed signal and noise power for all relevant multipoles in the analysis.

Because it includes reconstruction noise power, $C_L^{\hat{\phi}\hat{\phi},\mathrm{fg}}$ constitutes an upper limit to the bias that Galactic non-Gaussian foregrounds pose to CMB lensing reconstruction measurements. In particular, current analyses further remove the well-known lensing power spectrum biases, including the $N_L^{(0)}$ and $N_L^{(1)}$ biases; our results are upper limits because we do not subtract any of those biases in our baseline analysis. We investigated different avenues to subtract the dominant $N_L^{(0)}$ bias but concluded that we cannot trust our ability to remove it reliably in our estimates (see Appendix B for more details). Hence, in the main text we present upper bounds on the bias produced on the CMB lensing power spectrum amplitude, and, as such, the maximum impact that Galactic foregrounds could produce in our lensing analyses. We follow [12,34] and employ the statistic ΔA_{lens} to quantify the bias in the inferred lensing power spectrum amplitude A_{lens} due to the presence of Galactic foregrounds:

$$\Delta A_{\mathrm{lens}} = \frac{\sum_L \sigma_L^{-2} C_L^{\hat{\phi}\hat{\phi},\mathrm{fg}} C_L^{\phi\phi}}{\sum_L (C_L^{\phi\phi}/\sigma_L)^2}, \tag{15}$$

where we assume a fiducial lensing power spectrum signal $C_L^{\phi\phi}$ computed in the default cosmology and a diagonal covariance matrix (with elements σ_L) representing the error bars of an actual lensing measurement.

Since the uncertainty on A_{lens} is given by $(\sum_L (C_L^{\phi\phi}/\sigma_L)^2)^{-1/2}$, we can also report the bias as a fractional error:

$$\frac{\Delta A_{\mathrm{lens}}}{\sigma(A_{\mathrm{lens}})} = \frac{\sum_L \sigma_L^{-2} C_L^{\hat{\phi}\hat{\phi},\mathrm{fg}} C_L^{\phi\phi}}{\sqrt{\sum_L (C_L^{\phi\phi}/\sigma_L)^2}}. \tag{16}$$

Since the variance of the estimator is given by the normalization, we will approximate the noise on our lensing bandpowers as

$$N_L^{\kappa\kappa} = A_L^{\phi} \left(\frac{L(L+1)}{2} \right)^2, \tag{17}$$

and we will use the Knox formula [99–101] to compute the diagonal covariance:

$$\sigma_L^2 = \frac{2}{(2L+1)\Delta L f_{\mathrm{sky}}} (C_L^{\kappa\kappa} + N_L^{\kappa\kappa})^2, \tag{18}$$

where ΔL is the width of the bandpowers used and $C_L^{\kappa\kappa}$ is our fiducial CMB lensing power spectrum. For ACT, this approach leads to an underestimation of $\sigma(A_{\mathrm{lens}})$, predicting overly optimistic SNRs (∼90) that have not been achieved in practice, due to various reasons including filtering and noise anisotropies. We therefore introduce a common fudge factor of 1.3 to all our analytical predictions of $\sigma(A_{\mathrm{lens}})$ for ACT to recover the SNR values (∼70) measured in practice. For SO, this empirical degradation factor is not necessary as $\sigma(A_{\mathrm{lens}})$ is in line with previous forecasts. We use the binning and L range from the latest ACT lensing analysis [12]: 13 bins distributed within the range $40 < L < 1300$ with bin edges $L = [40, 66, 101, 145, 199, 264, 339, 426, 526, 638, 763, 902, 1100, 1300]$.

In Fig. 4, we show the fiducial CMB lensing power spectrum (solid black line) compared to the lensing reconstruction noise levels corresponding to different estimators. As discussed in Sec. I, we see that the TT estimator only offers moderate gain going from an ACT-like experiment to an SO-like one, since the CMB temperature is already signal dominated on the relevant scales. Consequently, for a given frequency channel, any increase in foreground bias in temperature-only reconstructions will be primarily driven to the wider area accessed by SO in comparison to ACT. The vast majority of the improvement

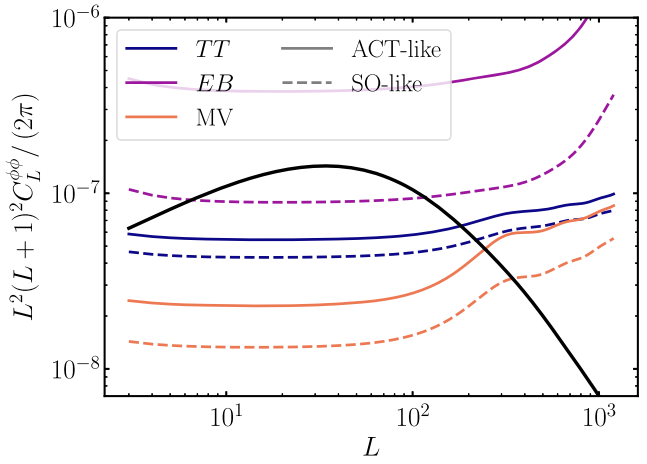


FIG. 4. Fiducial CMB lensing power spectrum (solid black line) compared to the lensing reconstruction noise levels corresponding to different estimators. The TT estimator only offers moderate gain going from an ACT-like experiment (with a white noise level of 12 μK_{CMB}-arcmin) to an SO-like one (6 μK_{CMB}-arcmin). The vast majority of the improvement comes from estimators making use of polarization channels, such as EB. As a result, for a given frequency channel, any increase in foreground bias in temperature-only reconstructions will be primarily driven by the wider area accessed by SO in comparison to ACT.

comes from estimators making use of polarization channels, such as EB.

III. RESULTS

A. Temperature-only reconstruction

In Fig. 5, we show the results of running the TT quadratic estimator on different non-Gaussian dust foreground models and as a function of sky fraction. The top panel corresponds to ACT-like footprints and noise levels, whereas the bottom panel refers to SO-like ones. We also show the fiducial CMB lensing power spectrum for reference as a solid black line. These curves present an upper bound on dust contamination. We have not subtracted any known biases, and hence their shape at low L is contaminated by the mean-field bias produced by the mask. On scales much smaller than the ones shown in these figures, we see the characteristic shape produced by the $N_L^{(0)}$ bias. An explanation of these is presented in Fig. 13 (Appendix A). It is not straightforward to estimate and subtract those biases because, on the one hand, the two-point power spectrum used to compute the analytical $N_L^{(0)}$ bias is not representative of the local two-point power spectrum across the map. On the other hand, we do not have enough realizations of the dust field for the different models to estimate the mean field. In Appendix B, we explore a possible strategy based on a patch-by-patch reconstruction, which has previously been proposed in [36,46]. This approach had previously been used in [102], following [103], as a validation method of their main reconstruction pipeline.

The white regions in Fig. 5 correspond to the scales used in [12] to derive cosmological parameters, namely $40 < L < 1300$, and hence the scales on which we compute the upper limit to the bias $\Delta A_{\rm lens}$ on the lensing power spectrum amplitude $A_{\rm lens}$. We report these in Fig. 6, in terms of the measurement uncertainty on $A_{\rm lens}$. On the 70% masks, $\sigma(A_{\rm lens})$ is 0.016 (ACT) and 0.009 (SO). In Figs. 5 and 6, we focus on two different midfrequency channels where the CMB is brightest. Since the dust intensity at 150 GHz is ∼3 times brighter than in the 90 GHz channel (see Sec. II C), we see that the biases in the lensing reconstruction are almost 2 orders of magnitude larger in the 150 GHz channel than in the 90 GHz one.[12] Finally, we perform lensing reconstruction on different footprints. As expected, the biases increase with sky fraction since we are accessing regions closer to the Galactic plane, where Galactic foreground emission dominates. At 90 GHz, we find that the shift induced on $A_{\rm lens}$ remains below 0.3σ for ACT (limit marked as a vertical dashed gray line in Fig. 6). It is well below the 0.5σ level for SO (limit marked as a thick gray line in Fig. 5), except for when using the 80% mask for which biases become non-negligible. As a result (unless there are further improvements to our conservative upper limits in future work), we advise against using Planck 80% dust mask for CMB lensing reconstruction analyses.[13]

1. Mitigation strategies: Profile hardening

A common mitigation strategy of extragalactic foregrounds in CMB lensing analyses is the use of bias-hardened quadratic estimators. This is a geometric method that modifies the form of the standard QE such that, at leading order, it has zero response to the statistical anisotropy caused by these foregrounds [30,36,37,42]. In particular, [12,34] use a tSZ-like profile to bias harden against tSZ clusters, although the method performs well in the presence of other foregrounds with different profiles, such as the CIB [42]. We test whether the tSZ-profile-hardened "source" (s) estimator also successfully mitigates non-Gaussian dust foregrounds. Since the Galactic dust distribution contains both sourcelike clumps and a large-scale masklike modulation in its amplitude, we might expect hardening approaches, which reduce the impact of both point sources and mask effects, to reduce the level of dust bias.

[12]Fortunately for ACT lensing, the 90 GHz frequency channel has more statistical weight than the 150 GHz one, as we discuss in Sec. III C.

[13]We also note that [102] found discrepancies at low L when using $f_{\rm sky} = 0.8$, which was suspected to be result of a residual dust contamination.

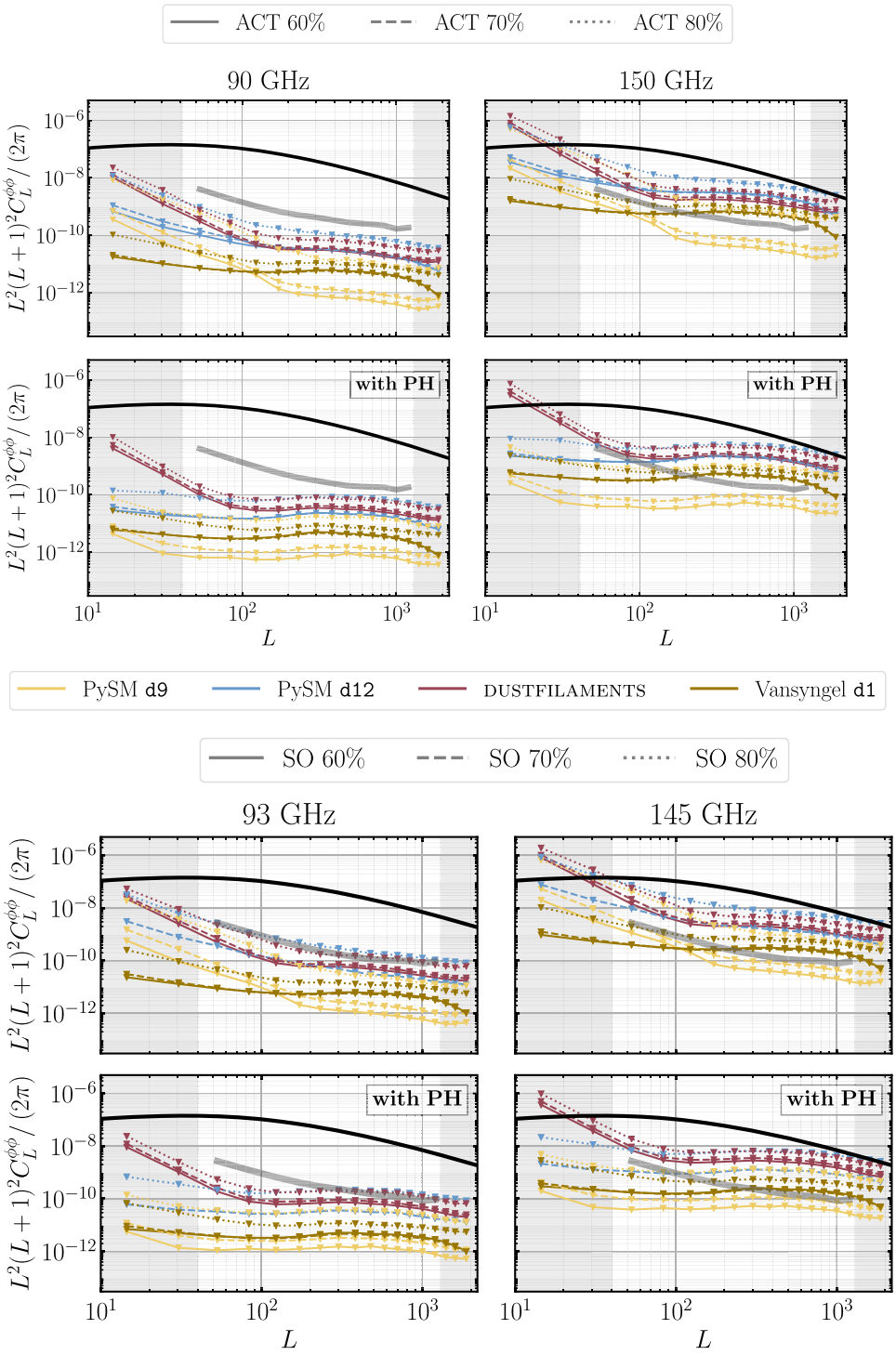


FIG. 5. Upper limits to the level of Galactic dust contamination on CMB lensing reconstruction measurements in temperature, for a subset of foreground models studied in this work (see Fig. 17 in Appendix E for the full version of this figure). We use the scales $600 < \ell_T < 3000$ for the reconstruction. For reference, we also show the fiducial CMB lensing powers spectrum (black line) and mark as a thick gray line the level of foreground that would produce a 0.5σ shift on the inferred CMB lensing power spectrum amplitude, which varies depending on the experimental configuration and the estimator used. The white regions correspond to the scales usually used to derive cosmological parameters, namely $40 < L < 1300$. The SED of dust is such that biases in the lensing reconstruction are almost 2 orders of magnitude larger in the 150 GHz channel (right column) than in the 90 GHz one (left column). Biases also increase as a function of sky fraction, denoted with different line styles (for clarity, we have also omitted PySM d12 results on the SO 60% mask in this figure). The top (bottom) four panels correspond to results with an experimental setup similar to ACT (SO). The panels labeled as "with PH" include profile hardening (tSZ-optimized) as a mitigation strategy, although this does not decrease of the level of contamination in the scales of interest significantly.

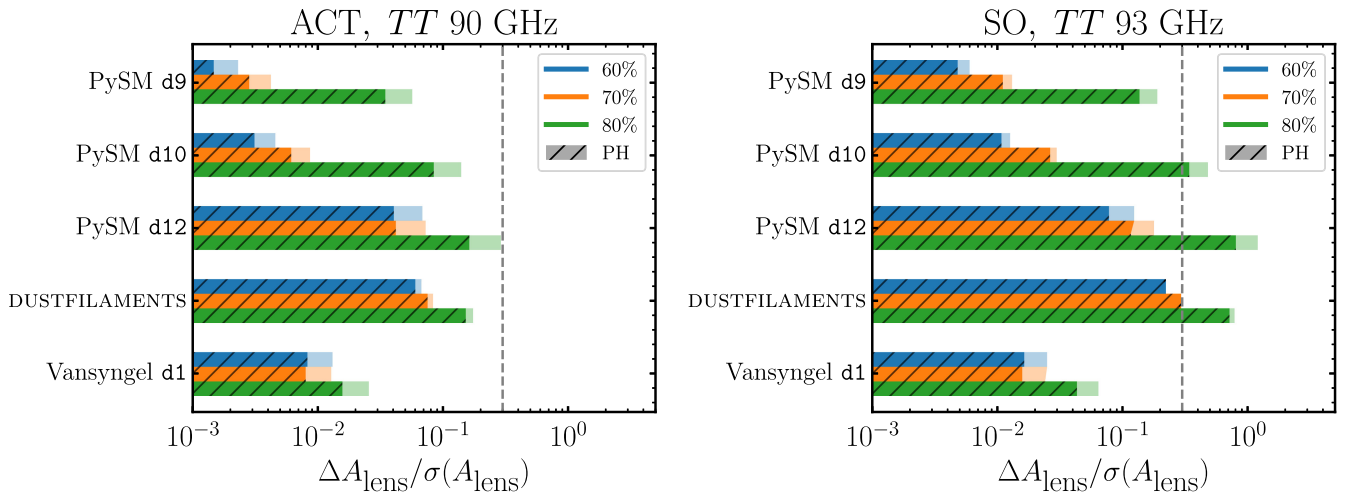


FIG. 6. Summary of upper limits to the bias of the CMB lensing power spectrum amplitude temperature-only reconstruction induced by the different dust models on ACT (left panel) and SO (right panel). We show as different colors the results for different sky fractions (60%, blue; 70%, orange; 80%, green). The hatched bars correspond to results obtained on the same setting but with profile hardening (PH) added as a mitigation strategy. This helps in all models except on DUSTFILAMENTS. For reference, with profile hardening, the uncertainty on $A_{\rm lens}$ on the 70% masks is 0.019 (ACT) and 0.011 (SO). We find that the shift induced on $A_{\rm lens}$ remains well below 0.3σ (vertical dashed line) for ACT and also for SO when excluding the 80% mask. We show results at 90 GHz. At 150 GHz, the biases are in all cases almost a hundred times larger.

In the presence of this foreground, Eq. (3) becomes

$$\langle T_{\mathbf{l}} T_{\mathbf{L}-\mathbf{l}} \rangle = f^{\phi}_{\mathbf{l},\mathbf{L}} \phi_{\mathbf{L}} + f^{s}_{\mathbf{l},\mathbf{L}} s_{\mathbf{L}}, \tag{19}$$

where the average is taken at fixed lensing and source field. The bias-hardened estimator for $\phi_{\mathbf{L}}$ in the presence of the source field $s_{\mathbf{L}}$ is [30,36,37]

$$\phi^{\mathrm{BH},s}_{\mathbf{L}} = \frac{1}{1 - \mathcal{R}^{\phi s}_{\mathbf{L}} \mathcal{R}^{s\phi}_{\mathbf{L}}} (\hat{\phi}_{\mathbf{L}} - \mathcal{R}^{\phi s}_{\mathbf{L}} \hat{s}_{\mathbf{L}}), \tag{20}$$

where $\hat{s}_{\mathbf{L}}$ is an estimator for the source field, constructed following the same procedure as that for the lensing field $\phi_{\mathbf{L}}$, and $\mathcal{R}^{\phi s}_{\mathbf{L}}$ is the response function [36]:

$$R^{a,b}_{\mathbf{L}} \sim A^{a}_{\mathbf{L}} \int \mathrm{d}^2\mathbf{l} f^{a}(\mathbf{l},\mathbf{L}) f^{b}(\mathbf{l},\mathbf{L}), \tag{21}$$

which we define to absorb the normalization of the estimator. There is a moderate noise cost for bias hardening (BH), where the noise in the reconstruction [cf. Eq. (17)] now takes the form [30,34,36,37]

$$N^{\kappa\kappa,\mathrm{BH}}_{L} = N^{\kappa\kappa}_{L} (1 - N^{\kappa\kappa}_{L} N^{ss}_{L} \mathcal{R}^{2}_{L})^{-1}, \tag{22}$$

where $\mathcal{R}_L$ is the response of the estimator to the presence of the foreground we harden against and N^{ss}_L is the noise of the source estimator. With profile hardening, our estimate of $\sigma(A_{\rm lens})$ increases by 20%: 0.019 (ACT) and 0.011 (SO).

The results are shown in the panels labeled as "with PH" in Fig. 5. We also quote the bias in the lensing amplitude as hatched bars in Fig. 6. We find that profile hardening helps reduce the amplitude of the bias on the PySM and Vansyngel + 17 models, although the improvement is not very large. This result is somewhat surprising since the bias produced by dust "blobs" should be absorbed by profile hardening, as noted earlier. We find the improvement to be mainly driven by a suppression of the power spectrum amplitude at low L, usually enhanced by the presence of a mask. We can see this explicitly by computing the expected value of the bias-hardening source estimator:

$$\langle \phi^{\mathrm{BH},s}_{\mathbf{L}} \rangle = \frac{1}{1 - \mathcal{R}^{\phi s}_{\mathbf{L}} \mathcal{R}^{s\phi}_{\mathbf{L}}} (\langle \hat{\phi}_{\mathbf{L}} \rangle - \mathcal{R}^{\phi s}_{\mathbf{L}} \langle \hat{s}_{\mathbf{L}} \rangle), \tag{23}$$

in the presence of both foreground s and a mask $w(\hat{\mathbf{n}})$, that is,

$$\langle T_{\mathbf{l}} T_{\mathbf{L}-\mathbf{l}} \rangle = f^{\phi}_{\mathbf{l},\mathbf{L}} \phi_{\mathbf{L}} + f^{s}_{\mathbf{l},\mathbf{L}} s_{\mathbf{L}} + f^{M}_{\mathbf{l},\mathbf{L}} M_{\mathbf{L}}, \tag{24}$$

where we have introduced the complementary function of the mask in Fourier space, $M_\ell = \delta_\ell - W_\ell$ [36]. We introduce (24) in (4) and substitute in (23). Making use of (21), we find

$$\langle \phi^{\mathrm{BH},s}_{\mathbf{L}} \rangle = \phi_{\mathbf{L}} + \frac{\mathcal{R}^{\phi M}_{\mathbf{L}} - \mathcal{R}^{\phi s}_{\mathbf{L}} \mathcal{R}^{sM}_{\mathbf{L}}}{1 - \mathcal{R}^{\phi s}_{\mathbf{L}} \mathcal{R}^{s\phi}_{\mathbf{L}}} M_{\mathbf{L}}. \tag{25}$$

That is, the prefactor quantifying the contribution to the bias due to the presence of a mask (originally $\mathcal{R}^{\phi M}_{\mathbf{L}}$) has now been reduced. This reduction is proportional to how similarly the two fields respond to the presence of ϕ.

In the limit $s \sim M$, $\mathcal{R}_{\mathbf{L}}^{sM} = 1$, and the contribution from the mask is completely suppressed. We compute explicitly the ratio of this new prefactor obtained when we bias harden against point sources with respect to the original response $\mathcal{R}_{\mathbf{L}}^{\phi M}$. The reduction is typically 20%, with values ranging from 30% at $L \sim 10$ to 10% at $L \sim 1000$.

This idea is now linked to the fact that both PySM and Vansyngel + 17 models rely on modulating small scales by a large-scale field $t(\hat{\mathbf{n}})$ to generate non-Gaussianities. This operation can straightforwardly be absorbed by the presence of a mask $\tilde{w}(\hat{\mathbf{n}}) = w(\hat{\mathbf{n}})t(\hat{\mathbf{n}})$. Therefore, using profile hardening will also mitigate against biases arising from this new mask (M_ℓ now becomes $\tilde{M}_\ell = \delta_\ell - \tilde{W}_\ell$ in the discussion above). DUSTFILAMENTS does not synthesize small scales following the modulation approach. We can speculate that since DUSTFILAMENTS neither looks like a mask modulation nor looks like individual sources, instead appearing as very elongated structures (as can be seen in Fig. 2), profile hardening has minimal impact.

In order to decrease the bias at all multipoles, new profiles based on the foreground's bispectrum and power spectrum [37] are needed. For example, these could be chosen to reflect the typical size of filaments or dust clumps. We leave this line of research to future work.

B. Galactic foreground biases in both intensity and polarization

As CMB experiments reach higher sensitivities, polarization measurements will contribute to the majority of SNR in lensing analyses [24,104] and quadratic estimators will become suboptimal [19–22]. The improvements from optimal methods applied to experiments such as ACT or

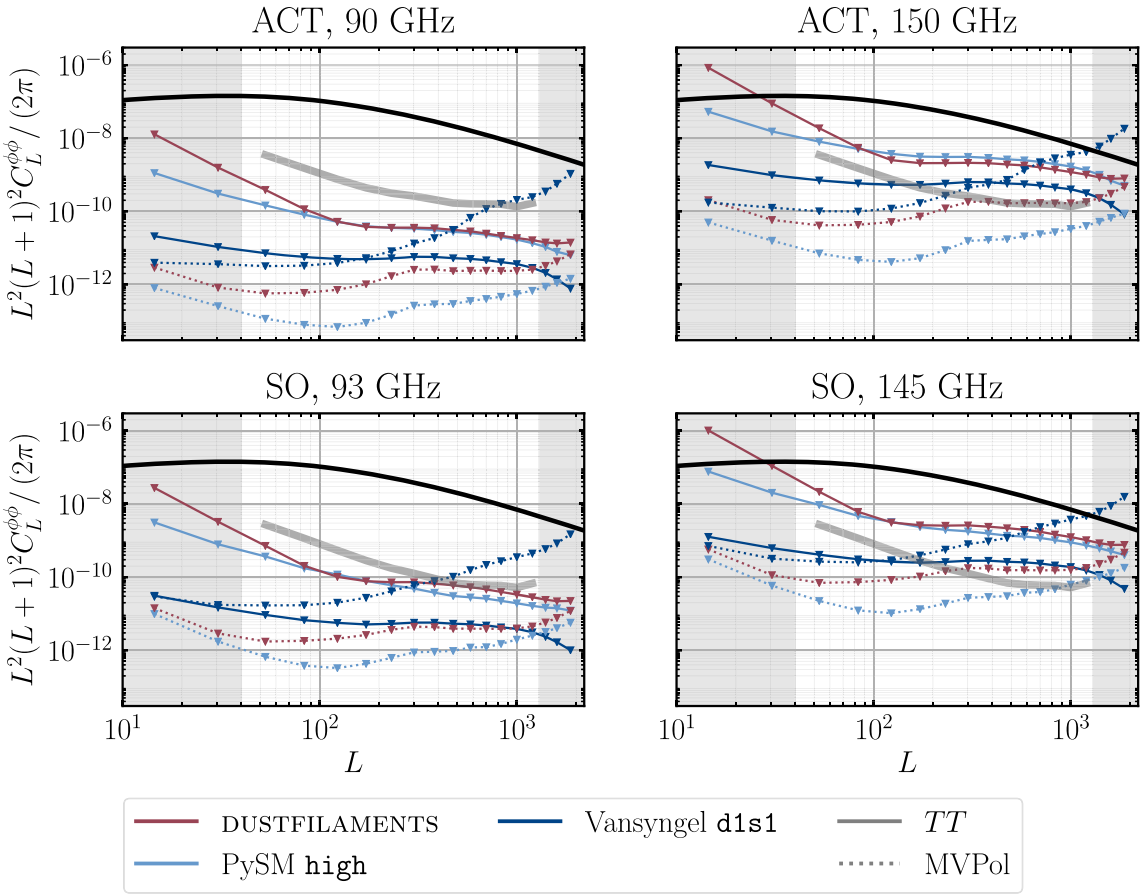


FIG. 7. Upper limit to the level of Galactic foreground contamination on CMB lensing reconstruction measurements for the temperature and polarization-only quadratic estimators (marked with different line styles). The top panels show results for ACT on the ACT 70% footprint ($f_{\rm sky} = 0.31$), and the bottom panels correspond to SO on the SO 70% footprint ($f_{\rm sky} = 0.42$). The plots on the left (right) correspond to the 90 GHz (150 GHz) channels. Different colors correspond to different foreground models. The reconstruction is performed with multipoles $600 < \ell < 3000$, and we shade the L regions excluded on cosmological analyses. For reference, we mark as a thick gray line the level of foreground that would produce a 0.5σ shift on the inferred CMB lensing power spectrum amplitude for the different configurations with the minimum-variance estimator. On PySM and DUSTFILAMENTS maps, estimators that make use of polarization measurements are less biased, whereas Vansyngel d1s1 model presents a higher bias in polarization than in temperature. At 150 GHz, the biases show similar trends than at 90 GHz but are almost a hundred times larger, as was discussed in Sec. III A.

SO are very modest. We do expect, however, polarization measurements to contribute significantly to the reconstruction SNR, especially for SO. As a result, we now explore the impact of non-Gaussian Galactic foregrounds in both intensity and polarization within the framework of quadratic estimators. We now include synchrotron within our foreground modeling (PySM low-, medium- and high-complexity models, as discussed in Sec. II C), even though its contribution on frequencies above 70 GHz is negligible. We further check that synchrotron produces no shift in the CMB lensing power spectrum bias on those channels.

In Fig. 7, we show upper limits to the CMB lensing power spectrum bias for the temperature and polarization-only quadratic estimators. The top panel shows results for ACT channels on the ACT 70% footprint ($f_{\rm sky} = 0.31$), and the bottom panel corresponds to the SO mid-frequency channels on the SO 70% footprint ($f_{\rm sky} = 0.43$). We only include the most complex PySM model (PySM high), together with DUSTFILAMENTS and Vansyngel d1s1 for clarity. Firstly, we find that on PySM and DUSTFILAMENTS maps, estimators that make use of polarization measurements are less biased. This is in agreement with previous works [46], and it is not a result of the lack of *EB* correlation present on these models. Any nonzero *EB* signal cannot produce a mean-field bias in the *EB* estimator. However, it could generate some nonzero bias at the trispectrum level, as pointed out by [45]. We show in Appendix D that the level of *EB* correlations compatible with existing data is too small to have any detectable contribution to the trispectrum.

Finally, we also find that Vansyngel + 17 models present a higher bias in polarization than in temperature. We believe this is due to the loss of power above $\ell \sim 1000$ present in these simulations, which we explore further in the next section (Sec. III B 1). In a nutshell, the temperature-only estimator receives most of its contribution from high multipoles [105], but there is no signal on those small scales on this model.

Previous papers have found that the biases present in polarization due to non-Gaussian dust foregrounds can be mitigated well below the experiment's requirements after including the dust power spectrum in the filters for the CMB fields [Eq. (9)]. This strategy did not work for the temperature-only estimator since the CMB component dominates the temperature power spectrum CMB [45,46]. We are in the same position here. The level of dust predicted by the sky models we use is below the experiment's observed signal (including noise) for all relevant multipoles in the analysis, so including it does not serve as a mitigation strategy.

1. Mitigation strategies: Choice of multipole range used in the reconstruction

We follow [44,45] and show the sensitivity of the bias to changing the multipole range used in the lensing reconstruction. We explore changes in both $\ell_{\rm min}$ and $\ell_{\rm max}$ with respect to the default analysis range, $600 < \ell < 3000$ [12]. In particular, Galactic foreground biases are known to worsen when decreasing the minimum multipole.

We show in Fig. 8 the upper limit of dust foreground biases in the SO-like experiment (70% mask)

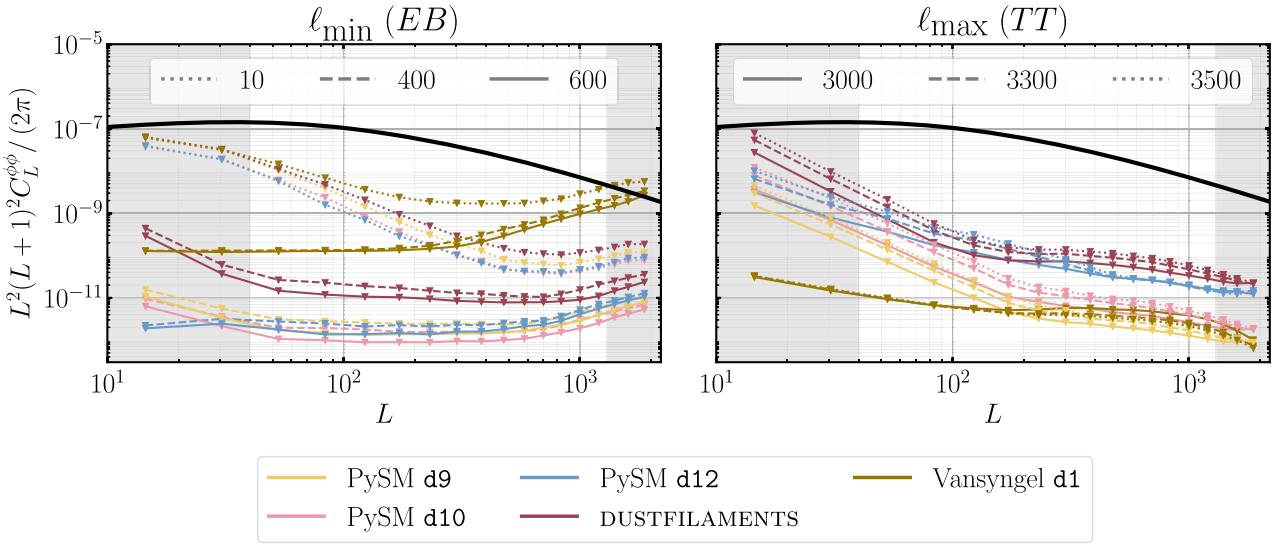


FIG. 8. Upper limit on the raw lensing power spectrum due to non-Gaussian dust foregrounds in an experiment such as SO (70% mask, $f_{\rm sky} = 0.42$) at 93 GHz, as a function of the multipole range used in the reconstruction, for *TT* and *EB* estimators. Left: as $\ell_{\rm min}$ is varied (fixed $\ell_{\rm max} = 3000$), the *EB* estimator presents significant shifts in the amplitude of the bias because it has important contributions from low-ℓ modes. Right: increasing $\ell_{\rm max}$ (while keeping $\ell_{\rm min} = 600$ fixed) in the temperature-only estimator also leads to an increase in the bias, although the shifts are small.

configuration. The left panel shows the sensitivity to lowering $\ell_{\rm min}$ in polarization, while the right panel corresponds to changes in $\ell_{\rm max}$ in a temperature-only reconstruction. We report changes in $A_{\rm lens}$ always with respect to the baseline $\ell_{\rm min} = 600$, $\ell_{\rm max} = 3000$. In polarization, the bias on $A_{\rm lens}$ increases from 0.8σ at $\ell_{\rm min} = 600$ up to 6.8σ for $\ell_{\rm min} = 10$ within the Vansyngel d1 dust model.[14] In temperature, increasing $\ell_{\rm max}$ increases the amplitude of the bias, although the shift is minimal (from 0.3σ to 0.8σ, corresponding to the DUSTFILAMENTS model as one goes from $\ell_{\rm max} = 3000$ to $\ell_{\rm max} = 3500$). We see no effect of changing the minimum multipole used in the reconstruction for the TT estimator or the maximum multipole for the EB estimator, so these are not shown.

These results can be explained as follows: the contribution of CMB multipoles $\ell \lesssim 700$ to the signal-to-noise ratio of the lensing reconstruction in temperature is very small [105], so we only see appreciable changes when we vary the maximum multipole instead. On the smallest scales, however, Galactic foregrounds become subdominant and produce a minimal change if included. The EB estimator, on the other hand, has significant contributions from lower ℓ [106], the scales at which Galactic foregrounds become more important.

Similar behavior was found in [44], who suggest low-ℓ filtering as a useful robustness check for the effect of diffuse foreground biases. As a general guidance, CMB lensing analyses should always ensure the consistency of different scale cuts.

In this section, we have stressed how high-ℓ multipoles contribute the most to the signal-to-noise in temperature-only reconstructions [105]. These scales, however, are not present in the Vansyngel + 17 model due to the presence of a $30'$ beam. If this was not the case, we would expect this model to present a larger bias in temperature than in polarization, as it is the case for the other sky models explored in this paper. This is apparent in Fig. 8: both solid lines in the two plots correspond to the same multipole range used in the reconstruction ($600 < \ell < 3000$), and we see that TT presents a larger bias in temperature than in polarization in all models except the Vansyngel + 17 one. As a result, the constraints derived with the Vansyngel + 17 model [55] for both temperature and minimum variance estimators are slightly optimistic.

C. Biases on coadded frequency channels

In practice, CMB lensing measurements are not performed on individual frequency maps. The first step is often to combine the data gathered by the experiment at the different frequencies into a single CMB map. This has several advantages, including foreground mitigation. We now discuss biases measured on coadded frequency channels. This method is similar to the foreground-cleaning approaches discussed in [44–46], which were found to significantly help reduce Galactic foreground biases.

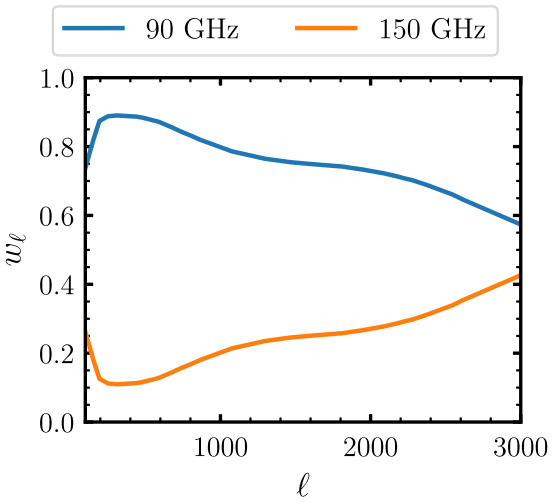


FIG. 9. Inverse-variance weights used to combine ACT 90 and 150 GHz channels in temperature. They are derived from the actual ACT noise power spectra [107] and are normalized to unity at every multipole ℓ. Similar weights are computed to combine the polarization (E, B) components. The 90 GHz channel provides the largest contribution to the coadded map.

For the ACT-like experiment, we perform an inverse-variance coaddition of the foreground maps at the different frequencies ("fcoadd"). We follow the same procedure as in ACT [12] to combine the maps from both 90 and 150 GHz frequency channels into a single "coadded" map $d^{\rm coadd}_{\ell m}$. The coaddition is done in harmonic space:

$$d^{\rm coadd}_{\ell m} = \sum_f w^f_\ell d^f_{\ell m}, \qquad f = \{90\ {\rm GHz}, 150\ {\rm GHz}\}, \quad (26)$$

where the weights are derived from the real ACT noise at each frequency channel [107] and are normalized to sum up to unity for every multipole ℓ. We show in Fig. 9 the resulting weights for the temperature component of the foreground fields. The weights for the E and B components are derived from the same procedure and follow the same trend: the 90 GHz channel provides the largest contribution to the coadded map.

In the case of the SO-like experiment, we perform a harmonic internal linear combination (HILC) (e.g. [90,108–110]), similar to Eq. (26) where the weights are now computed directly from the data covariance $\mathbb{C}$ (a 6×6 matrix encoding the auto- and cross-power spectra at the SO frequencies):

[14]Nevertheless, in this case the bias can be reduced to a 0.8σ bias by simply including the foreground power spectra in the filters. When multipoles below $\ell \sim 100$ in polarization are considered, the dust power spectrum is greater than the observed CMB signal (including noise), so it would be suboptimal not to include them in the filters.

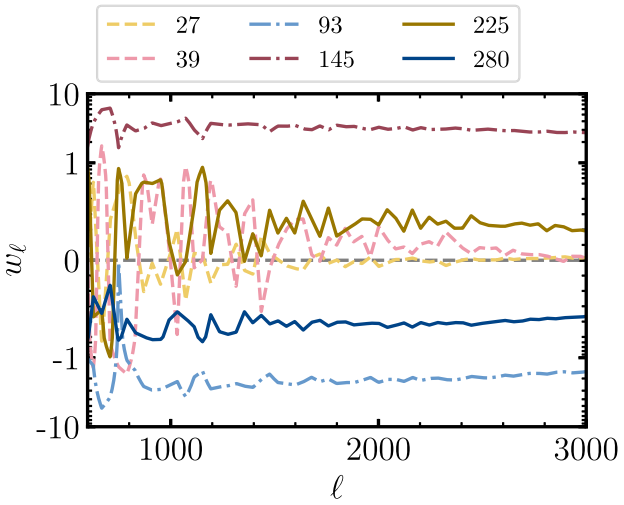


FIG. 10. HILC weights used to combine all six SO frequency channels in temperature (labeled in GHz units). They are computed directly from the data covariance, encoding all auto- and cross-spectra (including the lensed TT CMB spectra, thermal and kinetic SZ effects, CIB and radio point source emission [13]). For clarity, we smooth the resulting weights before plotting them. Similar weights are computed to combine the polarization (E, B) components. The midfrequency channels provide the dominant contribution to the HILC map.

$$\boldsymbol{w} = \frac{\mathbf{1}^\top \mathbb{C}^{-1}}{\mathbf{1}^\top \mathbb{C}^{-1}\mathbf{1}}, \tag{27}$$

where both $\boldsymbol{w}$ and $\mathbf{1}$ are six-dimensional vectors (each entry corresponding to a different frequency channel).[15] The resulting weights $\boldsymbol{w}$ are shown in Fig. 10. In the limit where the data are completely noise dominated, both "fcoadd" and HILC methods are equivalent. To construct the covariance, we use the auto- and cross-frequency temperature power spectra used in SO forecasting [13] and discussed in detail in the aforementioned paper. As a summary, they include lensed TT CMB spectra, thermal and kinetic SZ effects, CIB and radio point source emission, in addition to synchrotron, free-free and anomalous microwave emission from Planck Commander models [111] and thermal dust emission maps ("model 8" [112] in Sehgal *et al.* 2010 [113] simulations). In polarization, we add to the lensed CMB mock power spectra that include SO-like noise [114] and Galactic foreground emission. We measure the power spectra of PySM d9 and s4 models at 353 and 23 GHz, respectively. We extrapolate to the different SO channels using a modified blackbody spectrum for dust ($\beta_d = 1.59$, $\Theta_d = 19.6$ [47]) and a simple blackbody distribution for synchrotron ($\beta_s = -3.00$ [49,115,116]). As expected, we find that the weights are such that SO midfrequency channels provide the dominant contribution to the HILC map across all multipoles.

We present the results (with and without profile hardening) in Fig. 11. These correspond to reconstructions performed with the minimum-variance (MV) estimator. We report the resulting upper limits to the bias on the CMB power spectrum amplitude $A_{\rm lens}$ in Fig. 12 (left panel). We find that within the ACT 70% mask, the shift in $A_{\rm lens}$ lies below 0.3σ (dashed gray vertical line). For SO, however, it is unclear whether the bias remains at those levels, as not all models lie within this range when the SO 70% mask is used. Finally, we find that extending the CMB analysis footprint to the Planck 80% dust mask is potentially too risky if no further mitigation strategies are in place.

In Fig. 12 (right panel), we also quantify the bias on the lensing power spectrum amplitude for other estimators. We find that temperature-only reconstructions continue suffering from Galactic foreground biases. For ACT, this is expected as we are now including (although small; see Fig. 9) contributions from the 150 GHz channel. For SO, we see that the HILC method performs very well for the PySM low-complexity model. We believe this due to the fact that the spectral parameters on this model remain uniform across the sky. The absence of decorrelation and the use of six frequency channels makes the HILC method ideal to remove the non-Gaussian Galactic contamination. However, we find larger biases with more complex models (reaching the 3σ level for Vansyngel d1s1). Additional foreground complexity, in addition to the fact that SO HILC relies much more on the 150 GHz frequency channel (see Fig. 10), is possibly the origin of these larger biases in comparison to the results obtained in Sec. III A. In polarization, biases are smaller than in temperature. Therefore, quadratic estimators making use of polarization channels will be more robust against these biases. However, as we reach even lower noise levels and make use of iterative estimators, it is unclear what the impact of non-Gaussian Galactic foregrounds will be, and therefore future work is needed.

As mentioned earlier, the results presented here constitute an upper limit to the contamination produced by non-Gaussian Galactic foregrounds since we do not remove any of the expected lensing power spectrum biases. In Appendix B, we explore a possible strategy to remove the well-known mean-field and Gaussian biases. The approach is based on a patch-by-patch reconstruction (e.g. [103]) and had previously been proposed in [36,46]. However, we find this avenue unsuccessful and therefore choose not to include it when deriving our main results.

On the other hand, current analyses make use of the realization-dependent $N_L^{(0)}$ algorithm (RDN0 [7,36]), which is insensitive to errors in the power spectra of the simulations used to compute this bias to first order in the fractional error. This algorithm is known to reduce

[15]We use the ILC method to isolate the CMB. We work in thermodynamic temperature units, for which the SED is constant across frequencies [cf. Eq. (10)]; hence, we have a six-dimensional unit vector.

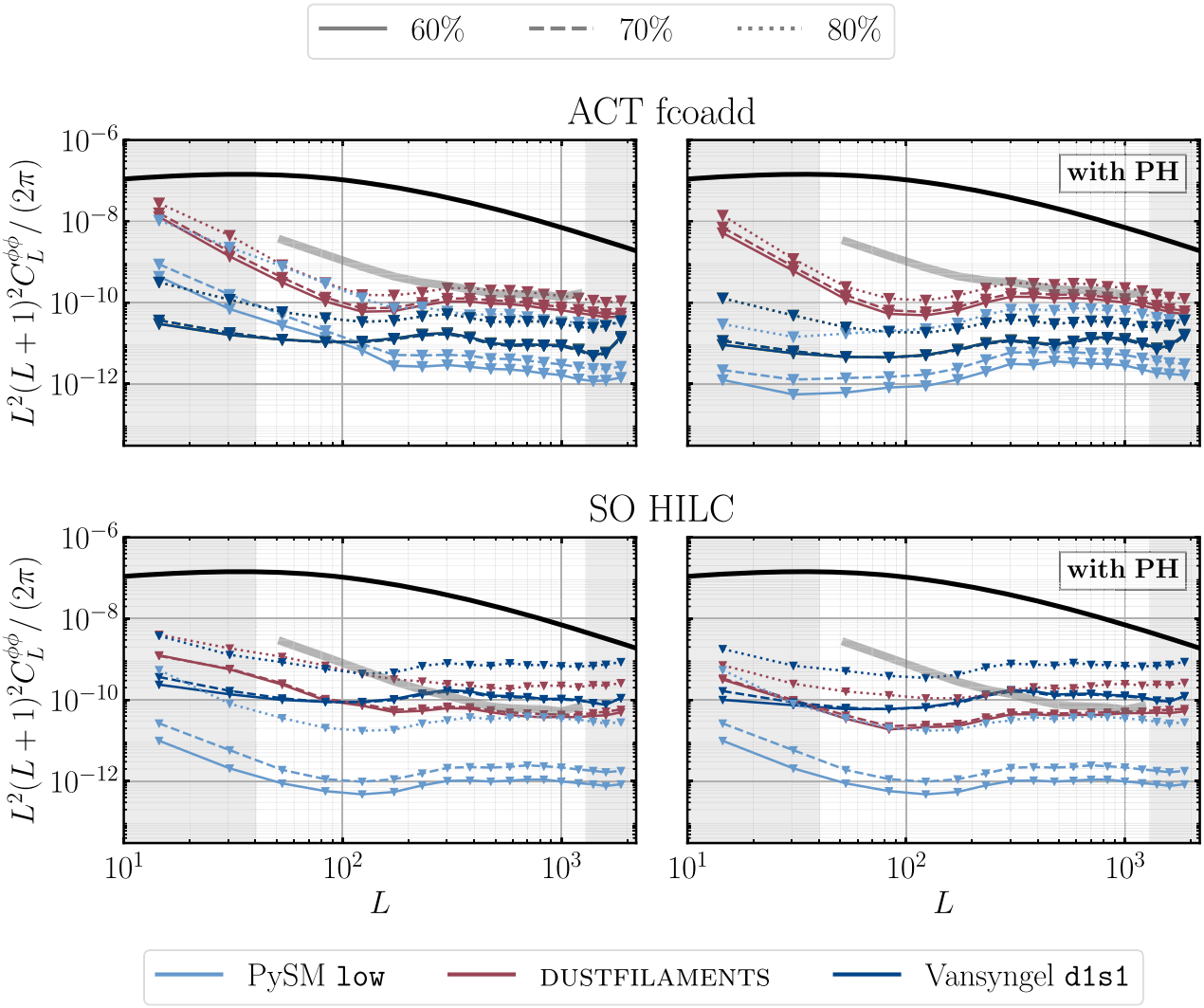


FIG. 11. Upper limits on the amplitude of the non-Gaussian Galactic foreground bias using the minimum-variance quadratic estimator within ACT and SO, for a subset of the foreground models studied in this work and as a function of sky fraction (see Fig. 18 in Appendix E for the full version of this figure). We include profile hardening (tSZ-optimized) in the panels marked as "with PH." For ACT, we use inverse-variance weighting coaddition of frequencies. For SO, we use the HILC method to coadd all six frequency channels. After coadding, and within the 70% mask region, these upper limits on the Galactic foregrounds biases are at the subpercent level for the ACT experiment. We quantify the bias on $A_{\rm lens}$ on the left panel of Fig. 12. For reference, we mark as a thick gray line the level of foreground that would produce a 0.5σ shift on the inferred CMB lensing power spectrum amplitude for the different configurations.

off-diagonal L correlations [16]. In Appendix C, we test whether it is successful at mitigating unaccounted non-Gaussian Galactic foreground biases with a connected four-point function. In summary, we find that bias subtraction does not appear to qualitatively change the conclusions from our simpler upper limit, with the caveat that those results are not fully converged due to limited simulation availability.

IV. CONCLUSION

Weak gravitational lensing of the CMB is a mature observable for precision cosmology. However, a particular concern for CMB lensing studies might be Galactic foregrounds. These are inherently non-Gaussian and hence might mimic the characteristic signal that lensing estimators are designed to measure [45,46]. As a result, any unknown bias in the inferred lensing power spectrum amplitude can directly impact cosmological inferences.

In this work, we presented an analysis that quantifies the impact of Galactic dust non-Gaussianity on lensing reconstructions performed with quadratic estimators in both temperature and polarization on the Atacama Cosmology Telescope (two frequency bands centered at 90 and 150 GHz with a white noise level of 12 $\mu K_{\rm CMB}$-arcmin

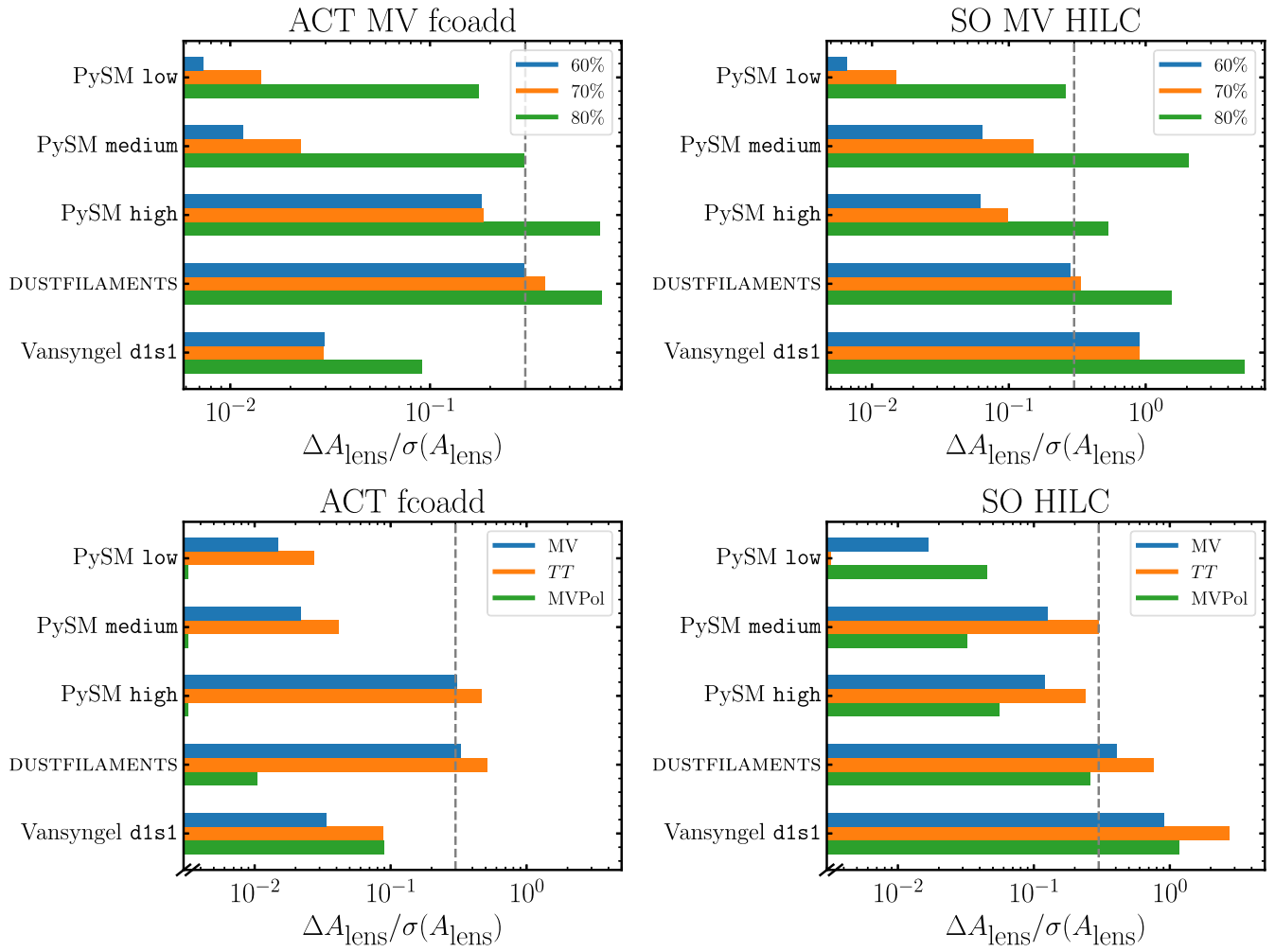


FIG. 12. Similar to Fig. 6, summary of upper limits to the bias of the CMB lensing power spectrum amplitude reconstruction induced by the different dust models, for both ACT (left column) and SO (right column) experiments, in units of $\sigma(A_{\rm lens})$. We run the estimators on coadded frequency channels (inverse-variance weighted "fcoadd" for ACT, harmonic internal linear combination "HILC" for SO). Top: reconstruction performed with the minimum variance MV estimator, with profile hardening. We show the bias as a function of sky fraction (60%, blue; 70%, orange; 80%, green). Within the ACT 70% mask, the shift in $A_{\rm lens}$ lies is below 0.3σ (vertical dashed gray line). For SO, however, it is unclear whether the bias remains at those levels, as not all models lie within this range. Bottom: upper limit to the bias as a function of the temperature or polarization channels used in the reconstruction: both (MV) in blue, temperature-only (TT) in orange and polarization-only (MVPol) in green. We use the 70% mask and no profile hardening for this test. In general, we find much larger biases in temperature-only reconstructions in comparison to those using polarization channels.

in temperature) and Simons Observatory (six frequency channels and 6 $\mu K_{\rm CMB}$-arcmin white-noise level in temperature) experiments.

We applied the relevant quadratic estimators to a comprehensive suite of foreground simulations. The resulting measurement of the lensing power spectrum $C_L^{\hat{\phi}\hat{\phi},{\rm fg}}$ constitutes an upper limit to the bias that Galactic non-Gaussian foregrounds pose to CMB lensing reconstruction measurements, and we used it to compute the bias on the inferred lensing power spectrum amplitude $A_{\rm lens}$. The dust SED is such that we found biases in the lensing reconstruction almost 2 orders of magnitude larger in the 150 GHz channel than in the 90 GHz one.

In general, we found polarization measurements to present smaller biases than temperature ones. We also found, in line with previous works, that $\ell_{\rm min}$ is the most sensitive to the presence of Galactic foregrounds and recommend that all analyses continue ensuring consistency of different scale cuts. Biases also increased with sky fraction, and, although our results are conservative upper limits, to be cautious we advise against using a Planck 80% dust mask for CMB lensing reconstruction analyses without any further mitigation strategy in place.

After coadding the two frequency channels, and within the 70% mask region, Galactic foregrounds do not present a problem for the ACT experiment (the shift in $A_{\rm lens}$ remains

below 0.3σ). In SO, not all foreground models remain below this threshold after performing a HILC. Although our results are conservative upper limits, the wider area accessed by SO, together with the larger weight on the 150 GHz channel, suggests that further work on characterizing dust biases and determining the impact of mitigation methods is well motivated, especially for the largest sky fractions.

The non-Gaussian nature of Galactic foregrounds poses a challenge to upcoming CMB lensing analyses. This is relevant not only for precision cosmology with the CMB lensing power spectrum, but it also impacts other probes such as the detection of inflationary B modes [117,118] or post-Born lensing curl B modes [119,120]. This work was an early analysis to quantify the level of contamination from Galactic emission on current ground-based CMB experiments and motivates further investigation to develop mitigation strategies for this foreground. In particular, ongoing work is studying the impact of component separation methods that use positional information, such as wavelets or needlets (see e.g. [121,122]), on the Galactic foreground problem for lensing. In addition, synergies with upcoming space telescopes such as the LiteBIRD satellite [123,124] will better inform our component separation strategy.

This research has also made extensive use of ORPHICS [125] and the ASTROPY [126], NumPy [127], and SciPy [128] packages. We also acknowledge use of the Matplotlib [129] package to produce the plots in this paper.

ACKNOWLEDGMENTS

We thank Toshiya Namikawa and Anthony Challinor for helpful input and discussions and Zach Atkins for providing the core code to plot masks in CAR projection. I. A.-C. acknowledges support from Fundación Mauricio y Carlota Botton and the Cambridge International Trust. C. H.-C. acknowledges funding by Agencia Nacional de Investigación y Desarrollo, Fondo Nacional de Desarrollo Científico y Tecnológico Postdoc fellowship 3220255 and by Agencia Nacional de Investigación y Desarrollo, BASAL, FB210003. B. D. S. acknowledges support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 851274). N. S. acknowledges support from DOE Grant No. DE-SC0025309. M. S. M. acknowledges support from NSF Grants No. AST-2307727 and No. AST-2153201 and NASA Grant No. 21-ATP21-0145. O. D. acknowledges support from a SNSF Eccellenza Professorial Fellowship (No. 186879). J. C. H. acknowledges support from the Sloan Foundation and the Simons Foundation. This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service [130], provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (Capital Grant No. EP/T022159/1), and DiRAC funding from the Science and Technology Facilities Council [131]. In particular, this work used the DiRAC Data Intensive service. The DiRAC component of CSD3 at Cambridge was funded by BEIS, UKRI and STFC capital funding and STFC operations grants. DiRAC is part of the UKRI Digital Research Infrastructure. This is not an official Simons Observatory Collaboration paper.

DATA AVAILABILITY

The data that support the findings of this article are openly available [51,64,65].

APPENDIX A: A PEDAGOGICAL INTRODUCTION TO QUADRATIC ESTIMATORS IN CMB LENSING

For simplicity, we continue the discussion working in the flat-sky approximation and focusing only on temperature fields. Lensing breaks translation invariance in the CMB. As described in the main text, when averaging over all possible unlensed CMB realizations, different modes in the CMB now become correlated due to lensing [Eq. (3)]. That expression suggests the following estimator for $\phi(\mathbf{L})$:

$$\hat{\phi}(\mathbf{L}) \sim \frac{T(\mathbf{l})T(\mathbf{L}-\mathbf{l})}{f(\mathbf{l},\mathbf{L})}, \tag{A1}$$

where we can combine all different pairs of modes applying inverse-variance weighting[16]:

$$\hat{\phi}(\mathbf{L}) = A^{\phi}_{\mathbf{L}} \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \frac{1}{\mathrm{Var}\left(\frac{T(\mathbf{l})T(\mathbf{L}-\mathbf{l})}{f(\mathbf{l},\mathbf{L})}\right)} \frac{T(\mathbf{l})T(\mathbf{L}-\mathbf{l})}{f(\mathbf{l},\mathbf{L})}, \tag{A2}$$

where $A^{\phi}_{\mathbf{L}}$ is inversely proportional to the total sum of inverse-variance weights,

$$A^{\phi}_{\mathbf{L}} = \left[\int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \frac{1}{\mathrm{Var}\left(\frac{T(\mathbf{l})T(\mathbf{L}-\mathbf{l})}{f(\mathbf{l},\mathbf{L})}\right)}\right]^{-1}. \tag{A3}$$

Using Wick's theorem,

$$\mathrm{Var}(T(\mathbf{l})T(\mathbf{L}-\mathbf{l})) = 2C^{\mathrm{obs}}_{\ell} C^{\mathrm{obs}}_{|\mathbf{L}-\mathbf{l}|}, \tag{A4}$$

[16]Given a set of measurements $\{\hat{X}_i\}$, they can be combined into a single estimate for X as $\hat{X} = \sum_i w_i \hat{X}_i$. The weights that provide an unbiased estimate for $\hat{X}$ (expressed with the condition $\sum_i w_i = 1$) while minimizing its variance are $w_k = (\sum_i 1/\sigma_i^2)^{-1} \cdot 1/\sigma_k^2$, where σ_i^2 is the variance of each measurement.

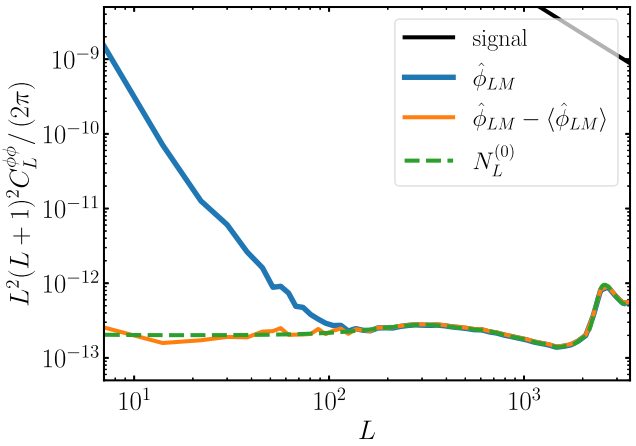


FIG. 13. Pedagogical introduction to CMB lensing biases. We run the TT quadratic estimator on a masked Gaussian (unlensed) field. The recovered signal (solid blue line) is equivalent to the reconstruction noise on small scales, which can be computed analytically and is shown as a dashed green line. The mask is also a source of statistical anisotropy and generates the so-called mean-field bias. This bias can be computed from simulations and then subtracted from the measured signal (solid orange line). See text for further details.

where $C_\ell, C_{|\mathbf{L}-\mathbf{l}|}$ are the observed (lensed) CMB spectra including noise. Substituting this expression in Eqs. (A2) and (A3), we obtain Eqs. (4) and (5), respectively.

When the temperature fields are combined using inverse-variance weights computed from the total power in the temperature map, as is the case in Eq. (4), the weights are so-called optimal. However, if those same weights, or filters, are used on a different temperature map (e.g. a foreground map), those weights will no longer be optimal. In that case, the variance of the reconstructed field will be given by[17] [4,34]:

$$N_L^{(0)} = (A_{\mathbf{L}}^{\phi})^2 \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \frac{f^2(\mathbf{l},\mathbf{L}) C_\ell^{\mathrm{fg}} C_{|\mathbf{L}-\mathbf{l}|}^{\mathrm{fg}}}{2(C_\ell^{\mathrm{obs}} C_{|\mathbf{L}-\mathbf{l}|}^{\mathrm{obs}})^2}. \quad \text{(A6)}$$

The normalization of the quadratic estimator is equivalent to the lensing power spectrum reconstruction noise. This is clear when running the lensing reconstruction on a Gaussian field. In this case, no lensing signal should be present. The output of the reconstruction is simply the reconstruction noise, sourced by the disconnected four-point function of the field. This is pictured in Fig. 13, where we run the temperature quadratic estimator on a Gaussian field generated from a given power spectrum. The measured signal (solid blue line) perfectly agrees with the analytical expectation for the $N_L^{(0)}$ bias (dashed green line) at high L. At low L, there is the so-called mean-field bias, which is due to the mask. The mask is a source of statistical anisotropy, present even in the absence of lensing. This is picked up by the quadratic estimator. In order to remove this contribution, we run the estimator on several masked Gaussian simulations with the same power spectrum as our fiducial map and take the average $\langle\hat{\phi}_{LM}\rangle$. We remove the mean-field bias $\langle\hat{\phi}_{LM}\rangle$ from the reconstructed $\hat{\phi}_{LM}$ before taking the power spectrum. The result is shown as the solid orange line in Fig. 13. With this correction, the reconstructed signal agrees with the Gaussian bias on all scales.

[17]This problem is equivalent to computing $\mathrm{Var}(\hat{X})$:

$$\mathrm{Var}(\hat{X}) = \mathrm{Var}\left(A\sum_i w_i \hat{X}_i\right) = A^2 \sum_i w_i^2 \mathrm{Var}(\hat{X}_i), \quad \text{(A5)}$$

where w_i are derived from the total (observed) CMB map [Eq. (A4)], whereas $\mathrm{Var}(\hat{X}_i)$ is proportional to the total foreground power.

APPENDIX B: LOCAL LENSING RECONSTRUCTIONS

To fully capture the impact of non-Gaussian dust on lensing spectrum measurements, the steps of mean-field subtraction and realization-dependent bias subtraction would need to be simulated accurately. However, this is challenging because, to accurately capture these steps, CMB signal and instrument noise would need to be added to the dust simulations, which would make it difficult to obtain converged estimates of very small dust biases. Despite these challenges, we will attempt this approach in the Appendix C.

In this appendix, however, we aim to obtain a simple, approximate (and more converged) estimate of the impact of mean-field and $N_L^{(0)}$ bias subtraction. Note that simply using the mean dust two-point function in the map to compute these mean-field and bias terms is not sufficient, since this does not capture position dependence in the dust power; correspondingly, we find that this approach underestimates the bias present in the reconstruction for all models beyond a simple Gaussian dust field.

The results provided in the main text correspond to upper limits to the total bias on the CMB lensing power spectrum amplitude due to the presence of non-Gaussian Galactic foregrounds. We found that the bias should lie below percent level. We confirm this result by running the full ACT lensing pipeline on two different simulations, which include a non-Gaussian dust realization from DUSTFILAMENTS [64] and PySM d12, respectively, using the MV estimator.

A possible workaround is to divide the footprint in different patches, perform the lensing reconstruction locally, and then sum over all different patches [36,46,103]. This is an approximation to the performance of the realization-dependent bias subtraction algorithm, making our results less conservative (and perhaps overly optimistic) when compared with the ones presented in the main text. With this approach, the two-point function within each patch will

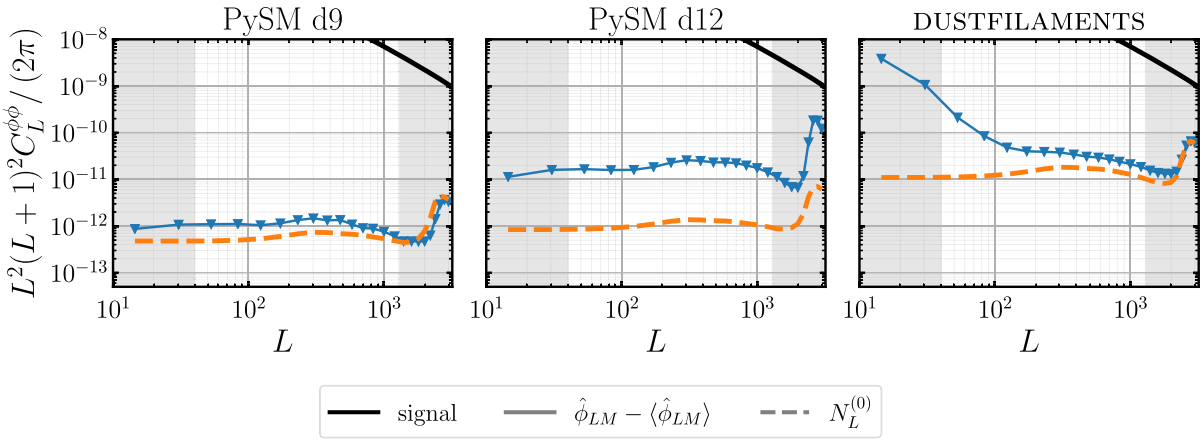


FIG. 14. Lensing reconstruction performed on different dust models, where we divide the footprint in different 50 deg^2 patches, perform the lensing reconstruction locally, and then sum over different patches. With this approach, we can estimate the mean-field and $N_L^{(0)}$ bias (dashed orange line), although with varying levels of success. The results here correspond to the 90 GHz ACT channel on the ACT 60% mask, for the temperature-only estimator including profile hardening.

represent the power within each patch more closely. We perform this exercise for the profile-hardened temperature-only estimator at 90 GHz on the ACT 60% mask and noise levels. We choose individual patches of area 50 deg^2.

In Fig. 14, we show our estimate for the $N_L^{(0)}$ bias computed this way [following Eq. (A6)] for three different models (dashed lines). We compare this to the power spectrum of the reconstructed $\hat{\phi}_{LM}$, where we have previously removed an estimate of the mean-field bias within each patch following the approach described in Appendix A (solid lines). We find that this approach is the most successful with the DUSTFILAMENTS model. For PySM d9, the $N_L^{(0)}$ bias is slightly overestimated, which leads to a negative lensing power spectrum at high L. In the case of PySM d12, the $N_L^{(0)}$ bias is greatly underestimated. We believe this is due to the presence of localized dust blobs in some patches, which makes the two-point function not representative of the power across the patch even for a reduced sky area. These results are also captured in Table I, where we derive upper limits to the shift induced on the lensing power spectrum amplitude with this approach. In the next section, we bypass the aforementioned issues by performing the full reconstruction (including bias subtraction) with the machinery of modern lensing analyses.

TABLE I. Upper limit to the bias of the lensing power spectrum amplitude for different dust models in a configuration similar to the ACT experiment in the 60% mask for the 90 GHz channel. We report results obtained for the TT estimator with profile hardening. We compare performing the reconstruction locally and then summing over the different patches ("patches") versus simply performing the reconstruction on the full footprint without any bias subtraction ("full-sky"). The former allows one to compute estimates for the mean-field and $N_L^{(0)}$ biases, which we subtract. In general, this approach allows for a reduction of the bias but it is not perfect, as can be seen in the accompanying Fig. 14.

$100 \times \Delta A_{\rm lens}$	Patches	Full-sky
PySM d9	0.002 (0.001σ)	0.003 (0.002σ)
PySM d12	0.084 (0.054σ)	0.081 (0.053σ)
DUSTFILAMENTS	0.067 (0.043σ)	0.120 (0.078σ)

APPENDIX C: RDN0

We show the reconstructed bandpowers in the top panel of Fig. 15. Note that both simulations include a map-based realistic noise realization computed with MNMS [107], and the reconstruction is done employing the cross-correlation-based quadratic estimator [132] which is immune to noise bias. Moreover, we subtract all relevant biases following the same procedure as [12]. This includes the realization-dependent $N_L^{(0)}$ algorithm, which absorbs small changes in the data due to the presence of dust. We estimate the lensing amplitude parameter $A_{\rm lens}$ by fitting the measured bandpowers to the baseline cosmology prediction. In the range $40 < L < 763$, we find

$$\text{No dust}: A_{\rm lens} = 1.021 \pm 0.021,$$
$$\text{DUSTFILAMENTS}: A_{\rm lens}^{\rm fg} = 1.026 \pm 0.021,$$
$$\text{PySM d12}: A_{\rm lens}^{\rm fg} = 1.023 \pm 0.021.$$

In all cases, the shift in $A_{\rm lens}$ is less than 0.3σ, where σ is the measurement error on $A_{\rm lens}$. Within the extended range $40 < L < 1300$, we find a slight increase in the bias, at the 0.4σ level:

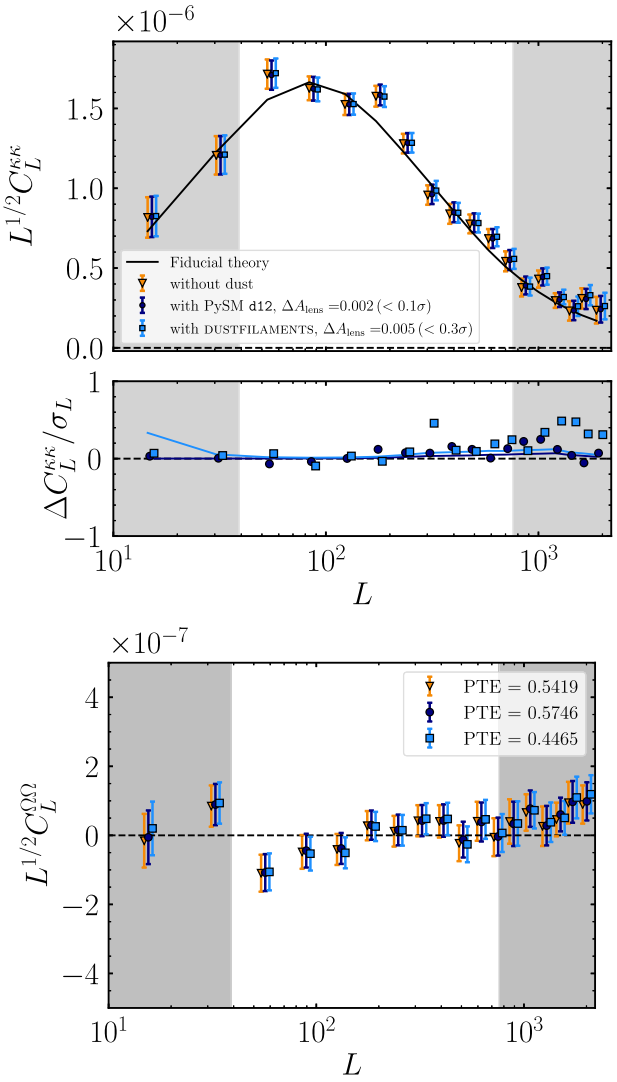


FIG. 15. We run the full ACT lensing pipeline on two simulations that include a non-Gaussian dust realization: DUSTFILAMENTS (light blue squares) and PySM d12 (dark blue circles). We compare them to a reconstruction done on a simulation that does not include dust (orange triangles). The results shown correspond to the MV estimator. Top: we confirm that the bias on the CMB lensing power spectrum amplitude due to the presence of a non-Gaussian dust field is consistent with producing a negligible shift in $A_{\rm lens}$ in comparison to its measurement uncertainty σ and in agreement with the results found on the main text (solid lines). Bottom: the curl component of the reconstruction is compatible with zero. We are not sensitive to any dust bias.

$$\text{No dust:}\ A_{\rm lens} = 1.023 \pm 0.020,$$
$$\text{DUSTFILAMENTS:}\ A_{\rm lens}^{\rm fg} = 1.031 \pm 0.020,$$
$$\text{PySM d12:}\ A_{\rm lens}^{\rm fg} = 1.027 \pm 0.020.$$

This bound is in agreement with what we found in the main text (Sec. III C, Figs. 11 and 12), where we also reported an upper limit for the bias of 0.3σ–0.4σ for these models. Finally, in the bottom panel of Fig. 15 we show the curl component of both reconstructions, where we also compute the probability to exceed with respect to a zero curl. No curl is significantly detected.

APPENDIX D: LENSING RECONSTRUCTION IS NOT SENSITIVE TO *EB* POWER SPECTRA

In Fig. 16, we see that estimators that make use of polarization measurements are less biased. This is not a result of the lack of *EB* correlation present on these models. Any nonzero *EB* signal cannot produce a mean-field bias in the *EB* estimator. However, it could generate some nonzero bias at the trispectrum level, as pointed out by [45]. In Fig. 16, we show that the level of *EB* correlations compatible with existing data is too small to have any detectable contribution to the trispectrum.

We run the quadratic estimator on two sets of DUSTFILAMENTS simulations, one of which includes *EB* correlations [92] at the level compatible with existing data [50,98]. We only find differences in the temperature-only reconstructions (that are then propagated into the MV estimator). This is due to small differences in the temperature component of the two simulations, present in order to achieve the nonzero parity-violating correlation: the orientation of the filaments with respect to the underlying magnetic field is manipulated, rendering populations of filaments with different orientations and therefore slightly different temperature fields [92].

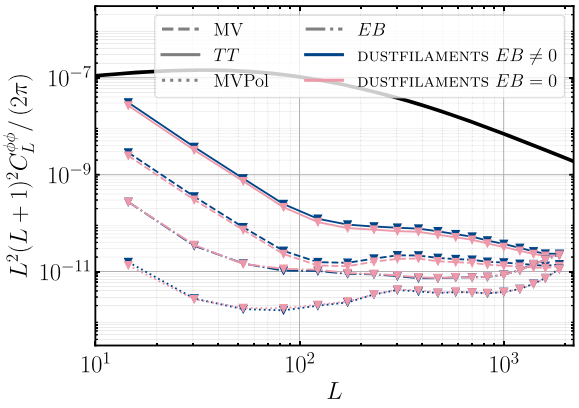


FIG. 16. The absence of *EB* correlation on PySM models [65] is not the cause for ACT/SO experiments to be less sensitive to non-Gaussian Galactic foreground biases in polarization. Here we show the difference in the reconstruction on DUSTFILAMENTS [64] simulations that include and do not include *EB* correlations [92], for a SO-like configuration at 93 GHz on the SO 70% mask. We find that the level of *EB* correlations compatible with existing data are too small to have any detectable contribution to the trispectrum. The small differences seen in temperature are due to the way these simulations are constructed.

APPENDIX E: EXTENSION OF FIGURES SHOWN IN THE MAIN TEXT

We show extended versions of Figs. 5 and 11, in order to showcase the results for the wider variety of foreground models studied in this work, in Figs. 17 and 18, respectively.

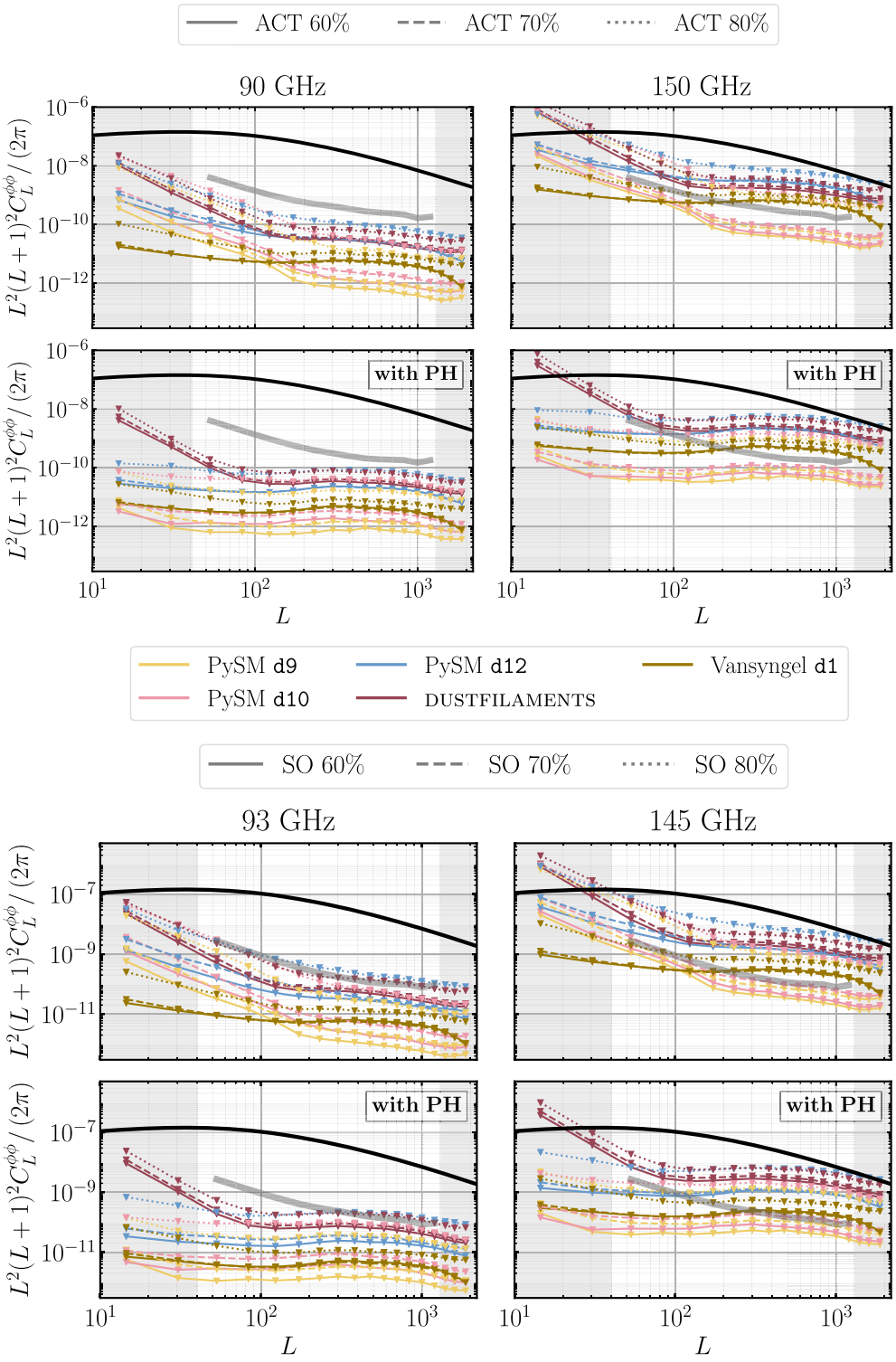


FIG. 17. Upper limits to the level of Galactic dust contamination on CMB lensing reconstruction measurements in temperature, for a variety of foreground models studied in this work (extension of Fig. 5).

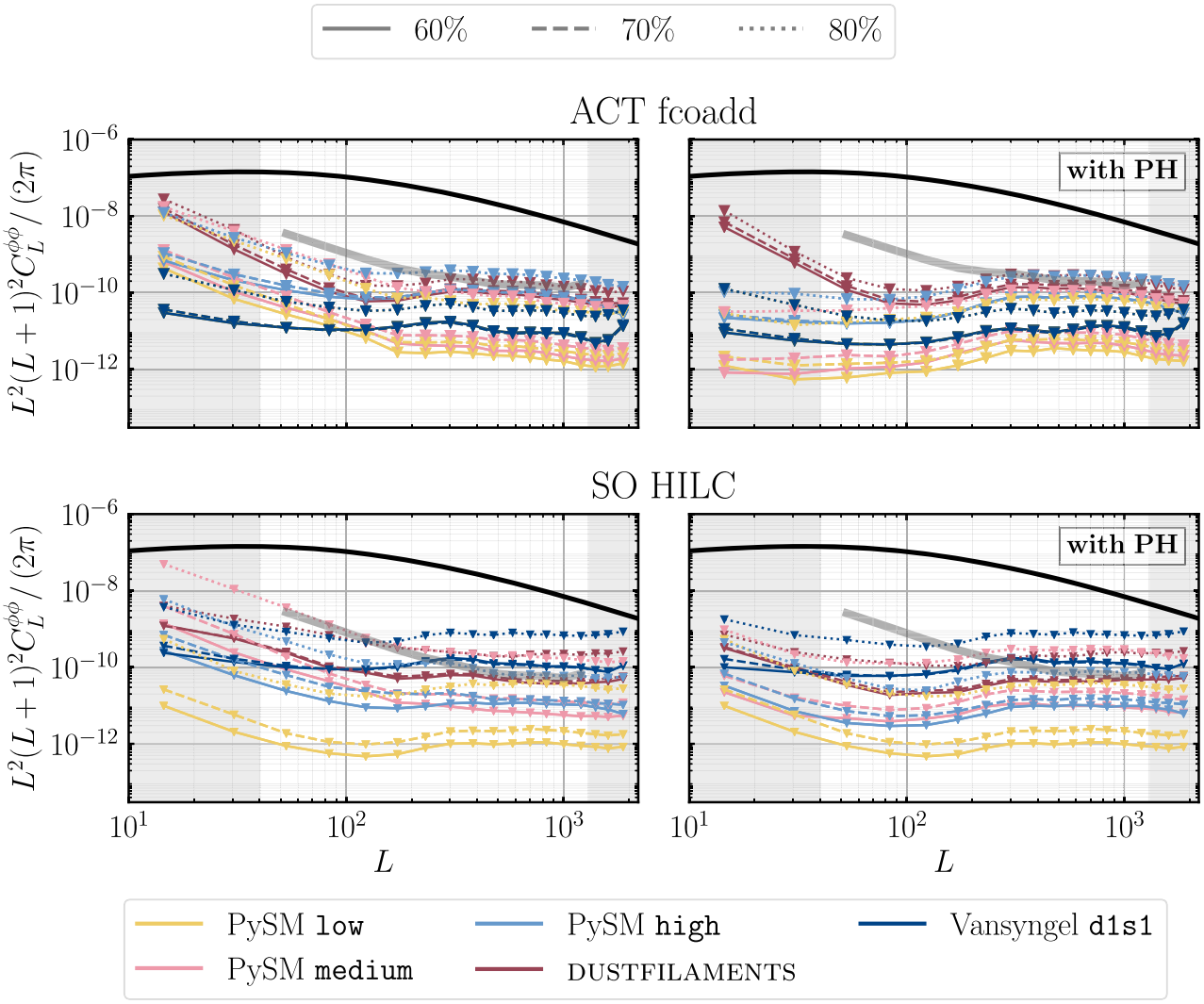


FIG. 18. Upper limits on the amplitude of the non-Gaussian Galactic foreground bias using the minimum-variance quadratic estimator within ACT and SO, for different foreground models and sky fractions (extension of Fig. 11).

[1] A. Blanchard and J. Schneider, Gravitational lensing effect on the fluctuations of the cosmic background radiation, Astron. Astrophys. **184**, 1 (1987).

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

[3] W. Hu, Mapping the dark matter through the cosmic microwave background damping tail, Astrophys. J. Lett. **557**, L79 (2001).

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

[5] T. Okamoto and W. Hu, Cosmic microwave background lensing reconstruction on the full sky, Phys. Rev. D **67**, 083002 (2003).

[6] Planck Collaboration, Planck 2018 results. X. Constraints on inflation, Astron. Astrophys. **641**, A10 (2020).

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

[8] J. Carron, M. Mirmelstein, and A. Lewis, CMB lensing from Planck PR4 maps, J. Cosmol. Astropart. Phys. 09 (2022) 039.

[9] J. W. Fowler *et al.*, Optical design of the Atacama Cosmology Telescope and the millimeter bolometric array camera, Appl. Opt. **46**, 3444 (2007).
[10] R. J. Thornton *et al.*, The Atacama Cosmology Telescope: The polarization-sensitive ACTPol instrument, Astrophys. J. Suppl. Ser. **227**, 21 (2016).
[11] S. W. Henderson *et al.*, Advanced ACTPol cryogenic detector arrays and readout, J. Low Temp. Phys. **184**, 772 (2016).
[12] F. J. Qu *et al.*, The Atacama Cosmology Telescope: A measurement of the DR6 CMB lensing power spectrum and its implications for structure growth, Astrophys. J. **962**, 112 (2024).
[13] Simons Observatory, The Simons Observatory: science goals and forecasts, J. Cosmol. Astropart. Phys. 02 (2019) 056.
[14] M. Zaldarriaga and U. Seljak, Gravitational lensing effect on cosmic microwave background polarization, Phys. Rev. D **58**, 023003 (1998).
[15] N. Sehgal *et al.*, CMB-HD: An ultra-deep, high-resolution millimeter-wave survey over half the sky, Bull. Am. Astron. Soc. **51**, 6 (2019).
[16] H. N. Nguyên, N. Sehgal, and M. S. Madhavacheril, Measuring the small-scale matter power spectrum with high-resolution CMB lensing, Phys. Rev. D **99**, 023502 (2019).
[17] M. H. Abitbol *et al.*, CMB-S4 technology book, first edition, arXiv:1706.02464.
[18] D. R. Barron *et al.*, Conceptual design of the modular detector and readout system for the CMB-S4 survey experiment, Proc. SPIE Int. Soc. Opt. Eng. **12190**, 164 (2022).
[19] C. M. Hirata and U. Seljak, Reconstruction of lensing from the cosmic microwave background polarization, Phys. Rev. D **68**, 083002 (2003).
[20] J. Carron and A. Lewis, Maximum *a posteriori* CMB lensing reconstruction, Phys. Rev. D **96**, 063510 (2017).
[21] L. Legrand and J. Carron, Lensing power spectrum of the cosmic microwave background with deep polarization experiments, Phys. Rev. D **105**, 123519 (2022).
[22] L. Legrand and J. Carron, Robust and efficient CMB lensing power spectrum from polarization surveys, Phys. Rev. D **108**, 103516 (2023).
[23] B. A. Benson *et al.*, SPT-3G: A next-generation cosmic microwave background polarization experiment on the south pole telescope, Proc. SPIE Int. Soc. Opt. Eng. **9153**, 91531P (2014).
[24] Z. Pan *et al.*, Measurement of gravitational lensing of the cosmic microwave background using SPT-3G 2018 data, Phys. Rev. D **108**, 122005 (2023),
[25] Y. B. Zeldovich and R. A. Sunyaev, The interaction of matter and radiation in a hot-model universe, Astrophys. Space Sci. **4**, 301 (1969).
[26] R. A. Sunyaev and Y. B. Zeldovich, Small-scale fluctuations of relic radiation, Astrophys. Space Sci. **7**, 3 (1970).
[27] J. L. Puget, A. Abergel, J. P. Bernard, F. Boulanger, W. B. Burton, F. X. Desert, and D. Hartmann, Tentative detection of a cosmic far-infrared background with COBE, Astron. Astrophys. **308**, L5 (1996).
[28] E. Dwek *et al.*, The COBE diffuse infrared background experiment search for the cosmic infrared background. IV. Cosmological implications, Astrophys. J. **508**, 106 (1998).
[29] D. J. Fixsen, E. Dwek, J. C. Mather, C. L. Bennett, and R. A. Shafer, The spectrum of the extragalactic far-infrared background from the COBE FIRAS observations, Astrophys. J. **508**, 123 (1998).
[30] S. J. Osborne, D. Hanson, and O. Doré, Extragalactic foreground contamination in temperature-based CMB lens reconstruction, J. Cosmol. Astropart. Phys. 03 (2014) 024.
[31] A. van Engelen, S. Bhattacharya, N. Sehgal, G. P. Holder, O. Zahn, and D. Nagai, CMB lensing power spectrum biases from galaxies and clusters using high-angular resolution temperature maps, Astrophys. J. **786**, 13 (2014).
[32] M. S. Madhavacheril and J. C. Hill, Mitigating foreground biases in CMB lensing reconstruction using cleaned gradients, Phys. Rev. D **98**, 023534 (2018).
[33] O. Darwish *et al.*, The Atacama Cosmology Telescope: A CMB lensing mass map over 2100 square degrees of sky and its cross-correlation with BOSS-CMASS galaxies, Mon. Not. R. Astron. Soc. **500**, 2250 (2021).
[34] N. MacCrann *et al.*, The Atacama Cosmology Telescope: Mitigating the impact of extragalactic foregrounds for the DR6 cosmic microwave background lensing analysis, Astrophys. J. **966**, 138 (2024).
[35] M. Doohan, M. Millea, S. Raghunathan, F. Ge, L. Knox, K. Prabhu, C. L. Reichardt, and W. L. K. Wu, Quantifying bias due to non-Gaussian foregrounds in an optimal reconstruction of CMB lensing and temperature power spectra, arXiv:2502.20801.
[36] T. Namikawa, D. Hanson, and R. Takahashi, Bias-hardened CMB lensing, Mon. Not. R. Astron. Soc. **431**, 609 (2013).
[37] N. Sailer, E. Schaan, and S. Ferraro, Lower bias, lower noise CMB lensing with foreground-hardened estimators, Phys. Rev. D **102**, 063517 (2020).
[38] K. M. Smith *et al.*, CMBPol mission concept study: Gravitational lensing, in *CMB Polarization Workshop: Theory and Foregrounds: CMBPol Mission Concept Study*, American Institute of Physics Conference Series Vol. 1141, edited by S. Dodelson, D. Baumann, A. Cooray, J. Dunkley, A. Fraisse, M. G. Jackson, A. Kogut, L. Krauss, M. Zaldarriaga, and K. Smith (AIP, New York, 2009), pp. 121–178.
[39] A. Hall and A. Challinor, Detecting the polarization induced by scattering of the microwave background quadrupole in galaxy clusters, Phys. Rev. D **90**, 063518 (2014).
[40] C. Feng and G. Holder, Polarization of the cosmic infrared background fluctuations, Astrophys. J. **897**, 140 (2020).
[41] T. Namikawa and R. Takahashi, Bias-hardened CMB lensing with polarization, Mon. Not. R. Astron. Soc. **438**, 1507 (2014).
[42] N. Sailer, S. Ferraro, and E. Schaan, Foreground-immune CMB lensing reconstruction with polarization, Phys. Rev. D **107**, 023504 (2023).
[43] F. J. Qu, M. Millea, and E. Schaan, Impact and mitigation of polarized extragalactic foregrounds on Bayesian cosmic microwave background lensing, Phys. Rev. D **111**, 023503 (2025).

[44] Y. Fantaye, C. Baccigalupi, S. M. Leach, and A. P. S. Yadav, CMB lensing reconstruction in the presence of diffuse polarized foregrounds, J. Cosmol. Astropart. Phys. 12 (2012) 017.

[45] D. Beck, J. Errard, and R. Stompor, Impact of polarized galactic foreground emission on CMB lensing reconstruction and delensing of B-modes, J. Cosmol. Astropart. Phys. 06 (2020) 030.

[46] A. Challinor *et al.*, Exploring cosmic origins with CORE: Gravitational lensing of the CMB, J. Cosmol. Astropart. Phys. 04 (2018) 018.

[47] Planck Collaboration, Planck intermediate results. XXII. Frequency dependence of thermal emission from Galactic dust in intensity and polarization, Astron. Astrophys. **576**, A107 (2015).

[48] Planck Collaboration, Planck 2015 results. X. Diffuse component separation: Foreground maps, Astron. Astrophys. **594**, A10 (2016).

[49] Planck Collaboration, Planck 2018 results. IV. Diffuse component separation, Astron. Astrophys. **641**, A4 (2020).

[50] Y. Akrami *et al.*, Planck 2018 results. XI. Polarized dust foregrounds, Astron. Astrophys. **641**, A11 (2020).

[51] C. Hervías-Caimapo, A. Bonaldi, and M. L. Brown, A new model of the microwave polarized sky for CMB experiments, Mon. Not. R. Astron. Soc. **462**, 2063 (2016).

[52] B. Thorne, J. Dunkley, D. Alonso, and S. Næss, The Python sky model: Software for simulating the Galactic microwave sky, Mon. Not. R. Astron. Soc. **469**, 2821 (2017).

[53] A. G. Kritsuk, R. Flauger, and S. D. Ustyugov, Dust-polarization maps for local interstellar turbulence, Phys. Rev. Lett. **121**, 021104 (2018).

[54] C. G. Kim, S. K. Choi, and R. Flauger, Dust polarization maps from TIGRESS: E/B power asymmetry and TE correlation, Astrophys. J. **880**, 106 (2019).

[55] F. Vansyngel *et al.*, Statistical simulations of the dust foreground to cosmic microwave background polarization, Astron. Astrophys. **603**, A62 (2017).

[56] G. Martínez-Solaeche, A. Karakci, and J. Delabrouille, A 3D model of polarized dust emission in the Milky Way, Mon. Not. R. Astron. Soc. **476**, 1310 (2018).

[57] E. Allys, F. Levrier, S. Zhang, C. Colling, B. Regaldo-Saint Blancard, F. Boulanger, P. Hennebelle, and S. Mallat, The RWST, a comprehensive statistical description of the non-Gaussian structures in the ISM, Astron. Astrophys. **629**, A115 (2019).

[58] B. Regaldo-Saint Blancard, F. Levrier, E. Allys, E. Bellomi, and F. Boulanger, Statistical description of dust polarized emission from the diffuse interstellar medium. A RWST approach, Astron. Astrophys. **642**, A217 (2020).

[59] L. Mousset, E. Allys, M. A. Price, J. Aumont, J. M. Delouis, L. Montier, and J. D. McEwen, Generative models of astrophysical fields with scattering transforms on the sphere, Astron. Astrophys. **691**, A269 (2024).

[60] N. Krachmalnicoff and G. Puglisi, ForSE: A GAN-based algorithm for extending CMB foreground models to subdegree angular scales, Astrophys. J. **911**, 42 (2021).

[61] J. Yao, N. Krachmalnicoff, M. Foschi, G. Puglisi, and C. Baccigalupi, FORSE+: Simulating non-Gaussian CMB foregrounds at 3 arcmin in a stochastic way based on a generative adversarial network, Astron. Astrophys. **686**, A290 (2024).

[62] S. E. Clark and B. S. Hensley, Mapping the magnetic interstellar medium in three dimensions over the full sky with neutral hydrogen, Astrophys. J. **887**, 136 (2019).

[63] K. M. Huffenberger, A. Rotti, and D. C. Collins, The power spectra of polarized, dusty filaments, Astrophys. J. **899**, 31 (2020).

[64] C. Hervías-Caimapo and K. M. Huffenberger, Full-sky, arcminute-scale, 3D models of galactic microwave foreground dust emission based on filaments, Astrophys. J. **928**, 65 (2022).

[65] The Pan-Experiment Galactic Science Group, Full-sky models of galactic microwave emission and polarization at sub-arcminute scales for the Python sky model, arXiv:2502.20452.

[66] L. Perotto, J. Bobin, S. Plaszczynski, J. L. Starck, and A. Lavabre, Reconstruction of the cosmic microwave background lensing for Planck, Astron. Astrophys. **519**, A4 (2010).

[67] D. Hanson, A. Challinor, and A. Lewis, Weak lensing of the CMB, Gen. Relativ. Gravit. **42**, 2197 (2010).

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

[69] G. Fabbian, A. Lewis, and D. Beck, CMB lensing reconstruction biases in cross-correlation with large-scale structure probes, J. Cosmol. Astropart. Phys. 10 (2019) 057.

[70] M. Kesden, A. Cooray, and M. Kamionkowski, Lensing reconstruction with CMB temperature and polarization, Phys. Rev. D **67**, 123507 (2003).

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

[72] https://github.com/simonsobs/falafel

[73] https://github.com/simonsobs/tempura

[74] https://github.com/simonsobs/so-lenspipe

[75] https://pla.esac.esa.int/

[76] M. S. Madhavacheril *et al.*, The Atacama Cosmology Telescope: DR6 gravitational lensing map and cosmological parameters, Astrophys. J. **962**, 113 (2024).

[77] A. Lewis, A. Challinor, and A. Lasenby, Efficient computation of cosmic microwave background anisotropies in closed Friedmann-Robertson-Walker models, Astrophys. J. **538**, 473 (2000).

[78] Planck Collaboration, Planck 2013 results. IX. HFI spectral response, Aston. Astrophys. **571**, A9 (2014).

[79] D. J. Fixsen, The temperature of the cosmic microwave background, Astrophys. J. **707**, 916 (2009).

[80] M. Kamionkowski and E. D. Kovetz, Statistical diagnostics to identify galactic foregrounds in B-mode maps, Phys. Rev. Lett. **113**, 191303 (2014).

[81] A. Rotti and K. Huffenberger, Isotropy-violation diagnostics for B-mode polarization foregrounds to the cosmic microwave background, J. Cosmol. Astropart. Phys. 09 (2016) 034.

[82] O. H. E. Philcox, B. D. Sherwin, and A. van Engelen, Detection and removal of B-mode dust foregrounds with

signatures of statistical anisotropy, Mon. Not. R. Astron. Soc. **479**, 5577 (2018).
[83] C. G. Kim and E. C. Ostriker, Three-phase interstellar medium in galaxies resolving evolution with star formation and supernova feedback (TIGRESS): Algorithms, fiducial model, and convergence, Astrophys. J. **846**, 133 (2017).
[84] C.-G. Kim, S. K. Choi, and R. Flauger, Dust polarization maps from TIGRESS: E/B power asymmetry and TE correlation, Astrophys. J. **880**, 106 (2019).
[85] S. E. Clark, C.-G. Kim, J. C. Hill, and B. S. Hensley, The origin of parity violation in polarized dust emission and implications for cosmic birefringence, Astrophys. J. **919**, 53 (2021).
[86] M. O. Irfan, J. Bobin, M.-A. Miville-Deschênes, and I. Grenier, Determining thermal dust emission from Planck HFI data using a sparse, parametric technique, Astron. Astrophys. **623**, A21 (2019).
[87] J. M. Delouis, E. Allys, E. Gauvrit, and F. Boulanger, Non-Gaussian modelling and statistical denoising of Planck dust polarisation full-sky maps using scattering transforms, Astron. Astrophys. **668**, A122 (2022).
[88] N. Aghanim *et al.*, Planck intermediate results. XLVIII. Disentangling Galactic dust emission and cosmic infrared background anisotropies, Astron. Astrophys. **596**, A109 (2016).
[89] N. Aghanim *et al.*, Planck 2018 results. XII. Galactic astrophysics using polarized dust emission, Astron. Astrophys. **641**, A12 (2020).
[90] M. Remazeilles, J. Delabrouille, and J.-F. Cardoso, Foreground component separation with generalized internal linear combination, Mon. Not. R. Astron. Soc. **418**, 467 (2011).
[91] https://github.com/chervias/DustFilaments
[92] C. Hervías-Caimapo, A. J. Cukierman, P. Diego-Palazuelos, K. M. Huffenberger, and S. E. Clark, Modeling parity-violating spectra in Galactic dust polarization with filaments and its applications to cosmic birefringence searches, Phys. Rev. D **111**, 083532 (2025).
[93] Planck Collaboration, Planck intermediate results. XXXIII. Signature of the magnetic field geometry of interstellar filaments in dust polarization maps, Astron. Astrophys. **586**, A136 (2016).
[94] Planck Collaboration, Planck intermediate results. XXXVIII. E- and B-modes of dust polarization from the magnetized filamentary structure of the interstellar medium, Astron. Astrophys. **586**, A141 (2016).
[95] S. E. Clark, J. E. G. Peek, and M. E. Putman, Magnetically aligned H I fibers and the rolling hough transform, Astrophys. J. **789**, 82 (2014).
[96] S. E. Clark, J. C. Hill, J. E. G. Peek, M. E. Putman, and B. L. Babler, Neutral hydrogen structures trace dust polarization angle: Implications for cosmic microwave background foregrounds, Phys. Rev. Lett. **115**, 241302 (2015).
[97] G. Halal, S. E. Clark, A. Cukierman, D. Beck, and C.-L. Kuo, Filamentary dust polarization and the morphology of neutral hydrogen structures, Astrophys. J. **961**, 29 (2024).
[98] Planck Collaboration, Planck intermediate results. XXX. The angular power spectrum of polarized dust emission at intermediate and high Galactic latitudes, Astron. Astrophys. **586**, A133 (2016).
[99] L. Knox, Determination of inflationary observables by cosmic microwave background anisotropy experiments, Phys. Rev. D **52**, 4307 (1995).
[100] L. Knox, Cosmic microwave background anisotropy observing strategy assessment, Astrophys. J. **480**, 72 (1997).
[101] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, Towards a cosmological neutrino mass detection, Phys. Rev. D **92**, 123535 (2015).
[102] Planck Collaboration, Planck 2013 results. XVII. Gravitational lensing by large-scale structure, Astron. Astrophys. **571**, A17 (2014).
[103] S. Plaszczynski, A. Lavabre, L. Perotto, and J. L. Starck, A hybrid approach to cosmic microwave background lensing reconstruction from all-sky intensity maps, Astron. Astrophys. **544**, A27 (2012).
[104] F. Ge *et al.*, Cosmology from CMB lensing and delensed EE power spectra using 2019–2020 SPT-3G polarization data, Phys. Rev. D **111**, 083534 (2025).
[105] M. M. Schmittfull, A. Challinor, D. Hanson, and A. Lewis, Joint analysis of CMB temperature and lensing-reconstruction power spectra, Phys. Rev. D **88**, 063012 (2013).
[106] R. Pearson, B. Sherwin, and A. Lewis, CMB lensing reconstruction using cut sky polarization maps and pure B modes, Phys. Rev. D **90**, 023539 (2014).
[107] Z. Atkins *et al.*, The Atacama Cosmology Telescope: Map-based noise simulations for DR6 2023, J. Cosmol. Astropart. Phys. 11 (2023) 073.
[108] C. L. Bennett *et al.*, First-year Wilkinson microwave anisotropy probe (WMAP) observations: Preliminary maps and basic results, Astrophys. J. Suppl. Ser. **148**, 1 (2003).
[109] H. K. Eriksen, A. J. Banday, K. M. Górski, and P. B. Lilje, On foreground removal from the Wilkinson microwave anisotropy probe data by an internal linear combination method: Limitations and implications, Astrophys. J. **612**, 633 (2004).
[110] F. McCarthy and J. C. Hill, Component-separated, CIB-cleaned thermal Sunyaev-Zel'dovich maps from Planck PR4 data with a flexible public needlet ILC pipeline, Phys. Rev. D **109**, 023528 (2024).
[111] Planck Collaboration, Planck 2015 results. XXV. Diffuse low-frequency Galactic foregrounds, Astron. Astrophys. **594**, A25 (2016).
[112] D. P. Finkbeiner, M. Davis, and D. J. Schlegel, Extrapolation of galactic dust emission at 100 microns to cosmic microwave background radiation frequencies using FIRAS, Astrophys. J. **524**, 867 (1999).
[113] N. Sehgal, P. Bode, S. Das, C. Hernandez-Monteagudo, K. Huffenberger, Y.-T. Lin, J. P. Ostriker, and H. Trac, Simulations of the microwave sky, Astrophys. J. **709**, 920 (2010).
[114] Generated using https://github.com/simonsobs/V3_calc.
[115] J. Dunkley *et al.*, Five-year Wilkinson microwave anisotropy probe (WMAP) observations: Bayesian estimation of cosmic microwave background polarization maps, Astrophys. J. **701**, 1804 (2009).

[116] Planck Collaboration, Planck 2015 results. XXV. Diffuse low-frequency Galactic foregrounds, Astron. Astrophys. **594**, A25 (2016).
[117] T. Namikawa and B. D. Sherwin, New probe of inflationary gravitational waves: Cross-correlations of lensed primary CMB B-modes with large-scale structure, Phys. Rev. Lett. **131**, 131001 (2023).
[118] T. Namikawa, I. Abril-Cabezas, and B. D. Sherwin, Probing inflationary gravitational waves with cross-correlations: Improved forecasting and validation with simulations, arXiv:2502.20741.
[119] M. Robertson and A. Lewis, How to detect lensing rotation, J. Cosmol. Astropart. Phys. 08 (2023) 048.
[120] M. Robertson, G. Fabbian, J. Carron, and A. Lewis, Detectable signals of post-Born lensing curl B-modes, J. Cosmol. Astropart. Phys. 02 (2025) 034.
[121] W. Coulton *et al.*, Atacama Cosmology Telescope: High-resolution component-separated maps across one third of the sky, Phys. Rev. D **109**, 063530 (2024).
[122] S. Goldstein, I. Abril-Cabezas *et al.* (to be published).
[123] LiteBIRD Collaboration, LiteBIRD: JAXA's new strategic L-class mission for all-sky surveys of cosmic microwave background polarization, Proc. SPIE Int. Soc. Opt. Eng. **11443**, 114432F (2020).
[124] LiteBIRD Collaboration, Probing cosmic inflation with the LiteBIRD cosmic microwave background polarization survey, Prog. Thoer. Exp. Phys. **2023**, 042F01 (2023).
[125] https://github.com/msyriac/orphics
[126] Astropy Collaboration, Astropy: A community Python package for astronomy, Astron. Astrophys. **558**, A33 (2013).
[127] S. van der Walt, S. C. Colbert, and G. Varoquaux, The NumPy array: A structure for efficient numerical computation, Comput. Sci. Eng. **13**, 22 (2011).
[128] P. Virtanen *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, Nat. Methods **17**, 261 (2020).
[129] J. D. Hunter, Matplotib: A 2D graphics environment, Comput. Sci. Eng. **9**, 90 (2007).
[130] www.csd3.cam.ac.uk
[131] www.dirac.ac.uk
[132] M. S. Madhavacheril, K. M. Smith, B. D. Sherwin, and S. Naess, CMB lensing power spectrum estimation without instrument noise bias, J. Cosmol. Astropart. Phys. 05 (2021) 028.