The Simons Observatory: Combining cross-spectral foreground cleaning with multitracer *B*-mode delensing for improved constraints on inflation

Emilie Hertig[1,2,*] Kevin Wolz,[3] Toshiya Namikawa,[4] Antón Baleato Lizancos,[5,6,7] Susanna Azzoni,[8]
Irene Abril-Cabezas,[9,2] David Alonso,[3] Carlo Baccigalupi,[10,11,12,13] Erminia Calabrese,[14] Anthony Challinor,[1,2,9]
Josquin Errard,[15] Giulio Fabbian,[14] Carlos Hervías-Caimapo,[16] Baptiste Jost,[17,4] Nicoletta Krachmalnicoff,[10,11,13]
Anto I. Lonappan,[18] Magdy Morshed,[15,19] Luca Pagano,[20,21,22] and Blake Sherwin[9,2]

(The Simons Observatory Collaboration)

[1]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge,
CB3 0HA, United Kingdom*
[2]*Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA, United Kingdom*
[3]*Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road,
Oxford, OX1 3RH, United Kingdom*
[4]*Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo,
Kashiwa, 277-8583, Japan*
[5]*Berkeley Center for Cosmological Physics, UC Berkeley, California 94720, USA*
[6]*Department of Physics, UC Berkeley, California 94720, USA*
[7]*Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, California 94720, USA*
[8]*Department of Astrophysical Sciences, Peyton Hall, Princeton University,
Princeton, New Jersey 08544, USA*
[9]*DAMTP, University of Cambridge, Cambridge, CB3 0WA, United Kingdom*
[10]*The International School for Advanced Studies (SISSA), via Bonomea 265, I-34136 Trieste, Italy*
[11]*The National Institute for Nuclear Physics (INFN), via Valerio 2, I-34127, Trieste, Italy*
[12]*The National Institute for Astrophysics (INAF), via Tiepolo 11, I-34143, Trieste, Italy*
[13]*The Institute for Fundamental Physics of the Universe (IFPU), Via Beirut 2, I-34151, Trieste, Italy*
[14]*School of Physics and Astronomy, Cardiff University,
The Parade, Cardiff, Wales CF24 3AA, United Kingdom*
[15]*Université Paris Cité, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France*
[16]*Instituto de Astrofísica and Centro de Astro-Ingeniería, Facultad de Física, Pontificia Universidad
Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile*
[17]*Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), UTIAS,
The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*
[18]*Dipartimento di Fisica, Universita di Roma Tor Vergata,
Via della Ricerca Scientifica, 1, 00133 Roma, Italy*
[19]*CNRS-UCB International Research Laboratory Centre Pierre Binétruy, CPB-IN2P3, Berkeley, US*
[20]*Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara,
via Saragat 1, I-44122 Ferrara, Italy*
[21]*Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, via Saragat 1, I-44122 Ferrara, Italy*
[22]*Institut d'Astrophysique Spatiale, CNRS, Univ. Paris-Sud, Univ. Paris-Saclay,
91405 Orsay cedex, France*

The Simons Observatory (SO), due to start full science operations in early 2025, aims to set tight constraints on inflationary physics by inferring the tensor-to-scalar ratio r from measurements of cosmic microwave background (CMB) polarization *B*-modes. Its nominal design including three small-aperture telescopes (SATs) targets a precision $\sigma(r = 0) \leq 0.003$ without delensing. Achieving this goal and further reducing uncertainties requires a thorough understanding and mitigation of other large-scale *B*-mode sources such as Galactic foregrounds and weak gravitational lensing. We present an analysis pipeline

*Contact author: emh83@cam.ac.uk

aiming to estimate r by including delensing within a cross-spectral likelihood, and demonstrate it for the first time on SO-like simulations accounting for various levels of foreground complexity, inhomogeneous noise and partial sky coverage. As introduced in an earlier SO delensing paper, lensing B-modes are synthesized using internal CMB lensing reconstructions as well as *Planck*-like cosmic infrared background maps and LSST-like galaxy density maps. We then extend SO's power-spectrum-based foreground-cleaning algorithm to include all auto- and cross-spectra between the lensing template and the SAT B-modes in the likelihood function. This allows us to constrain r and the parameters of our foreground model simultaneously. Within this framework, we demonstrate the equivalence of map-based and cross-spectral delensing and use it to motivate an optimized pixel-weighting scheme for power spectrum estimation. We start by validating our pipeline in the simplistic case of uniform foreground spectral energy distributions. In the absence of primordial B-modes, we find that the 1σ statistical uncertainty on r, $\sigma(r)$, decreases by 37% as a result of delensing. Tensor modes at the level of $r = 0.01$ are successfully detected by our pipeline. Even when using more realistic foreground models including spatial variations in the dust and synchrotron spectral properties, we obtain unbiased estimates of r both with and without delensing by employing the moment-expansion method. In this case, uncertainties are increased due to the higher number of model parameters, and delensing-related improvements range between 27% and 31%. These results constitute the first realistic assessment of the delensing performance at SO's nominal sensitivity level.

I. INTRODUCTION

Anisotropies of the cosmic microwave background (CMB) offer a unique window on the first instants after the birth of the Universe, probing fundamental physics at energy scales unreachable to any earthbound accelerator. After *Planck*'s cosmic-variance-limited measurements of the temperature power spectrum [1], present and future CMB experiments are directing their focus towards polarization patterns; large-scale B-modes are particularly interesting as they hold the key to a potential first detection of primordial gravitational waves (PGW). Predicted by most inflation models, these tensor fluctuations of the spacetime metric are the only ones capable of producing parity-odd polarization anisotropies, unlike the scalar (density) perturbations that exclusively create E-modes at linear order [2,3]. The detection of primordial B-modes would therefore constitute a significant breakthrough, providing unprecedented evidence in favor of the inflationary scenario, which, despite being widely accepted as a theoretical framework due to its ability to generate the right initial conditions for our Universe, still lacks direct experimental verification.

Two recent studies combining maps from BICEP/Keck, *Planck* and WMAP with baryon acoustic oscillation data allowed to set the tightest constraints to date on the amplitude of PGW, as described by the tensor-to-scalar ratio r; they respectively inferred $r < 0.036$ [4] and $r < 0.032$ [5] at 95% confidence for a pivot scale of 0.05 Mpc^{-1}. Together with the latest bounds on the scalar spectral index of the power spectrum of primordial curvature perturbations, these results favor a concave potential for the inflaton field driving inflation and have ruled out monomial models as well as natural inflation [4]. While many inflationary scenarios remain viable [6], the absence of a detection of PGW at the sensitivity level of current experiments has also motivated the development of alternative theories aiming to explain observed features of the Universe without the need for such a mechanism. These include, for example, bouncing cosmologies [7] as well as CPT-symmetric models [8]. New measurements of r with lower uncertainties will be an essential step towards discriminating between this wide range of possible descriptions of the early Universe.

The Simons Observatory (SO) [9], whose construction on Cerro Toco in the Atacama Desert is nearing completion with full science operations expected to start in early 2025, will significantly improve upon existing constraints; with deep observations of around 10% of the sky by three 0.5 m small-aperture telescopes (SATs), its nominal design aims to detect or rule out $r \geq 0.01$ at the 3σ level [10]. The planned addition of three more SATs will allow to reduce error bars even further. SO will therefore set tight bounds on currently allowed inflation scenarios such as quartic hilltop potentials and α-attractors [11]. Furthermore, the technologies and analysis pipelines developed for SO pave the way towards a new generation of experiments including CMB-S4 [12] and LiteBIRD [13], which will push limits to $\sigma(r) < 10^{-3}$ in the early 2030s. This foreshadows a particularly exciting time for early-Universe physics as several popular models, for example Starobinsky [14] and Higgs inflation [15], will then be decisively tested.

In order to reach its target precision, SO will have to overcome two main challenges. The first relates to the weak gravitational lensing of CMB photons by large-scale structures [16], which sources large-scale B-modes from intermediate- and small-scale E-modes [17]. Consequently, the deflected CMB light reaching our detectors contains an additional, lensing B-mode component sourced by scalar

fluctuations rather than PGW, whose amplitude surpasses that of the primordial tensor signal. This lensing-induced noise is larger than instrumental noise at SO's sensitivity level, and its contribution to sample variance will therefore be one of the main limiting factors affecting parameter constraints. Mitigating this effect, a process known as delensing, requires an accurate model of the specific lensing B-mode realization observed in our sky. Such a template is built by convolving measurements of inter- mediate- and small-scale E-modes with an estimator of the CMB lensing potential. In SO's case, both of these products will be extracted from high-resolution maps obtained with a 6 m large-aperture telescope (LAT); external large-scale structure tracers such as the cosmic infrared background (CIB) and galaxy surveys will also contribute to recon- structing the lensing convergence [18]. Delensing is per- formed by subtracting the B-mode template from the observed polarization maps, or by cross-correlating it with the data. This parametric power-spectrum-based approach was recently demonstrated for the first time by the BICEP/ Keck Collaboration and resulted in a 10% improvement on the inferred r constraint [19]. While the present work uses the same component separation technique, delensing of SO data is predicted to lead to a greater decrease in $\sigma(r)$ due to lower noise levels.

The second main challenge encountered by PGW searches is due to polarized Galactic foregrounds, in particular thermal dust emission and synchrotron radiation [20]. For the r values targeted by SO, the total B-modes produced by these sources dominate over the predicted inflationary tensor polarization pattern by orders of mag- nitude; on the angular scales of interest, their combined amplitude is at least equivalent to that of a primordial signal with $r \sim 0.05$ for any part of the sky [21]. Distinguishing these contaminants from the CMB is therefore an essential step towards inferring accurate cosmological information from the data [22]. Existing foreground cleaning techniques rely on the distinctive spectral energy distributions (SEDs) of Galactic emissions, which differ from the CMB blackbody spectrum and can be separated from the latter by multi- frequency observations [23]. Three independent algorithms have been designed to perform this task using SO's six frequency bands (from 27 to 280 GHz), both at the map and power spectrum levels, and have been tested on realistic SO- like simulations recently [24]. Delensing was not performed at this stage; its effect was instead estimated by using input simulations with decreased lensing B-mode power.

The present work aims to include delensing in the power- spectrum-based parametric component separation pipeline from Ref. [24], optimize its performance and characterize the subsequent reduction of statistical uncertainties on r when applied to simulations mimicking SO's noise and foreground properties. Lensing B-mode templates are obtained using the multitracer approach described in Ref. [18], where cross-spectral delensing was demonstrated

for an idealistic foreground-free case. Different levels of foreground complexity are investigated; biases related to residuals from spatial variability of foreground SEDs are mitigated by using a technique known as "moment expan- sion," which introduces additional parameters to model deviations from the sky average of the spectral properties [25–27]. This method was first demonstrated for SO B- modes in Ref. [28].

The paper is structured as follows. Section II summarizes the working principle of our analysis pipeline as well as the theoretical and mathematical foundations on which it relies. In Sec. III, we demonstrate the equivalence between map- based methods and the cross-spectral approach adopted here. Building upon this result, we optimize the delensing performance in Sec. IV by defining a new pixel-weighting scheme accounting for both instrumental noise and lensing B-mode variance. Section V describes the input simulations and the practical implementation of our pipeline, while statistical uncertainties and biases on r for all considered foreground and instrumental noise models are presented and analyzed in Sec. VI. Finally, we conclude and explore prospects for future work in Sec. VII. Further details on the likelihood derivation and tests performed with an alter- native pixel-weighting scheme are provided in Appendices A and B respectively.

II. THEORETICAL FRAMEWORK

A. Delensing: Motivation and principle

We start by outlining essential mathematical results related to CMB lensing and polarization, in order to illustrate the importance of delensing and characterize the B-mode template construction stage of our pipeline.

1. CMB lensing and cosmic variance

The linear polarization of light is quantified by the Stokes parameters Q and U, which are the components of a symmetric trace-free tensor on the sphere $\mathcal{P}_{ab}$ such that $\mathcal{P}_{\hat{\theta}\hat{\theta}} = -\mathcal{P}_{\hat{\phi}\hat{\phi}} = Q/2$ and $\mathcal{P}_{\hat{\theta}\hat{\phi}} = \mathcal{P}_{\hat{\phi}\hat{\theta}} = U/2$ in a right- handed orthonormal basis[1] $(\hat{n}, \hat{e}_\theta, \hat{e}_\phi)$. As photons traveling from the last-scattering surface are deflected by potential wells along the line of sight $\hat{n}$, $\mathcal{P}_{ab}$ is subjected to the direction-dependent remapping

$$\tilde{\mathcal{P}}_{ab}(\hat{n}) = \mathcal{P}_{ab}(\hat{n}) + \alpha^c(\hat{n})\nabla_c \mathcal{P}_{ab}(\hat{n}) \qquad (1)$$

at leading order, where the tilde refers to lensed quantities [29]. In the Born approximation, the small deflection angle $\alpha(\hat{n})$ corresponds to the gradient of the lensing potential ϕ. Using the spin-weight formalism described in [30], the Stokes parameters are related to E- and B-modes as follows:

---

[1]Here, $\hat{n}$ is the line-of-sight direction and $\hat{\theta}$ and $\hat{\phi}$ are along the θ and ϕ directions of a spherical coordinate system.

$$E_{lm} \pm iB_{lm} = -\int (Q \pm iU)(\hat{\boldsymbol{n}})_{\pm 2}Y_{lm}^*(\hat{\boldsymbol{n}})d^2\hat{\boldsymbol{n}}. \quad (2)$$

Integrating Eq. (1) allows one to derive the lensing *B*-mode contribution in the absence of primordial tensor perturbations [18]:

$$B_{lm}^{\text{lens}} = -i\sum_{l',m'}\sum_{L,M}\begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix}p^- F_{ll'L}^{(2)}E_{l'm'}^*\kappa_{LM}^*. \quad (3)$$

In Eq. (3), the factor p^- is 0 for $l + l' + L$ even and 1 for $l + l' + L$ odd, while $\kappa = -\frac{1}{2}\nabla^2\phi$ represents the lensing convergence and the spin-*s* mode coupling function $F_{ll'L}^{(s)}$ is given by

$$F_{ll'L}^{(s)} = \frac{2}{L(L+1)}[l'(l'+1) + L(L+1) - l(l+1)]$$
$$\times \sqrt{\frac{(2l+1)(2l'+1)(2L+1)}{16\pi}}\begin{pmatrix} l & l' & L \\ s & -s & 0 \end{pmatrix}. \quad (4)$$

As a result of this coupling, large-scale lensing *B*-modes receive contributions from *E*-modes at all scales, in particular from the high multipoles at which their power spectrum peaks [16].

The presence of an additional *B*-mode component in the data increases cosmic variance, thus amplifying uncertainties on inferred parameters. As an illustration, let us assume that the observed *B*-modes, B_{lm}, are a full-sky Gaussian field such that $\langle B_{lm}^*B_{l'm'}\rangle = \delta_{ll'}\delta_{mm'}C_l$, where the angle brackets denote an ensemble average over CMB and noise realizations. We fit this power spectrum with a theoretical model $rC_l^{\text{prim}} + C_l^{\text{lens}} + N_l$ consisting of the primordial signal, the lensing term[2] and a noise component. The full-sky log likelihood at a given multipole *l* corresponds to

$$-\ln\mathcal{L}(r) = \frac{1}{2}\sum_m(B_{lm}^*B_{lm}C_l^{-1} + \ln C_l)$$
$$= \frac{2l+1}{2}(\hat{C}_lC_l^{-1} + \ln C_l), \quad (5)$$

where the power spectrum estimator satisfies $\langle\hat{C}_l\rangle = C_l$. For $r = 0$, the associated statistical error can then be derived from the Fisher information:

---

[2]Since the lensing *B*-mode power spectrum is predicted very accurately within a given cosmological model and the latter is much better constrained than the foregrounds, we can afford to treat it as a fixed component. As discussed in Sec. VII, the possible inclusion of a free parameter modulating the lensing amplitude when working with real data is not expected to significantly impact $\sigma(r)$.

$$\sigma^{-2}(r)|_{r=0} = \sum_l\left\langle -\frac{\partial^2\ln\mathcal{L}}{\partial r^2}\right\rangle\Bigg|_{r=0}$$
$$= \sum_l\left(\frac{2l+1}{2}\right)\left(\frac{C_l^{\text{prim}}}{C_l^{\text{lens}} + N_l}\right)^2. \quad (6)$$

The lensing *B*-mode contribution to Eq. (6), which behaves similarly to white noise on large scales with a standard deviation of approximately $5\,\mu\text{K} - \text{arcmin}$, already surpasses SO's goal noise level of around $2\,\mu\text{K-arcmin}$ and therefore constitutes a significant limitation. While the heuristic argument above does not capture complex effects such as survey nonidealities (e.g., masking, inhomogeneous noise and foreground residuals) or the non-Gaussianity of lensing *B*-modes, it highlights the importance of delensing for SO and future PGW experiments.

2. B-mode template construction

In order to reduce the lensing contribution to the variance of *r* measurements, a template matching the particular realization of the lensing *B*-modes on our sky must be computed and subtracted from the data. This is done by evaluating Eq. (3) for estimates of the *E*-modes and the lensing convergence; Ref. [18] previously investigated the lensing *B*-mode template construction process for SO. In the following, we summarize the method used to extract the *E*-modes and the lensing potential and comment on the differences between the present paper and Ref. [18]. Note that the multifrequency LAT data used at this stage will be subjected to map-based foreground cleaning before any information is drawn from it; in the present analysis, we assume foreground residuals to be negligible and do not include them in our simulations. Investigating the impact of such residuals will be the subject of future work (see Sec. VII).

Small-scale *E*-modes are obtained by coadding data from three LAT frequency channels (93, 145 and 225 GHz) with inverse-noise-variance weighting in harmonic space. Instrumental noise is then mitigated by applying a diagonal Wiener filter

$$\hat{E}_{lm}^{\text{WF}} = \frac{C_l^{EE}}{C_l^{EE} + N_l^{EE}}\hat{E}_{lm}, \quad (7)$$

where C_l^{EE} is the theoretical lensed *E*-mode power spectrum, N_l^{EE} is computed for the coadded bands using the LAT noise model and hats refer to observed quantities. In Ref. [18], the Wiener-filtered *E*-modes are obtained with the algorithm presented in Ref. [31] to account for the inhomogeneities of noise and the survey boundary. Furthermore, Ref. [18] includes SAT *E*-modes, defined over the smaller area displayed in the right panel of Fig. 1. This technique was shown to improve the correlation between the estimator and the input *E*-modes for the foreground-free situation studied in Ref. [18], but has the
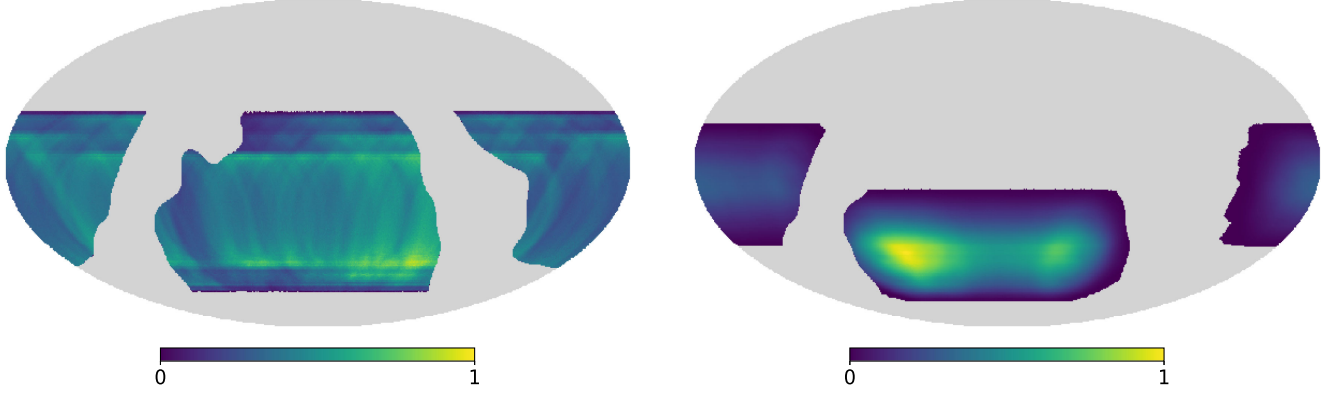
FIG. 1.   Normalized hit counts over the areas observed by the LAT (left) and SATs (right) in equatorial coordinates, multiplied by the *Planck* 70% Galactic mask. Masked pixels are shown in a lighter tone. These maps were made using the latest version of the planned SO scanning strategy. Of the pixels observed by the SAT, 96% are also seen by the LAT as a result of using slightly narrower Galactic cuts than in Ref. [24] to maximize overlap.

disadvantage of being very CPU intensive. To reduce computational costs, the present analysis will be restricted to LAT-only diagonally filtered E-modes. This condition also prevents our lensing templates from containing SAT foreground residuals and instrument systematics; these would contribute as small corrections to the cross-spectra between the template and the SAT maps, and will be the subject of future work. Note that while Eq. (3) refers to unlensed E-modes, the latter are not directly observable and our estimator is built using their lensed counterparts; gradient-order templates computed this way have actually been found to result in more efficient delensing due to cancellations between higher-order terms in the B-mode residuals [32].

Let us now briefly describe the method used to estimate the lensing convergence. As was done in Ref. [18], we combine the LAT CMB lensing map with external large-scale structure tracers. To reconstruct the CMB lensing map internally, we use quadratic estimators based on the correlation between lensed temperature (Θ) and polarization (E, B) fields, whose ensemble average over unlensed CMB fields but with a fixed realization of the κ field is given by

$$\langle \tilde{X}_{lm} \tilde{Y}_{l'm'} \rangle = \delta_{ll'} \delta_{m-m'} (-1)^m C_l^{XY}$$
$$+ \sum_{L,M} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} f_{ll'L}^{XY} \kappa_{LM}^* \quad (8)$$

at leading order [33]. The second (off-diagonal) term in Eq. (8) appears as a result of mode coupling and the response functions $f_{ll'L}^{XY}$ correspond to

$$f_{ll'L}^{\Theta\Theta} = F_{ll'L}^{(0)} C_{l'}^{\Theta\Theta} + F_{l'lL}^{(0)} C_l^{\Theta\Theta},$$
$$f_{ll'L}^{\Theta E} = p^+ F_{ll'L}^{(0)} C_{l'}^{\Theta E} + p^+ F_{l'lL}^{(2)} C_l^{\Theta E},$$
$$f_{ll'L}^{EE} = p^+ F_{ll'L}^{(2)} C_{l'}^{EE} + p^+ F_{l'lL}^{(2)} C_l^{EE},$$
$$f_{ll'L}^{EB} = -i p^- F_{l'lL}^{(2)} C_l^{EE}, \quad (9)$$

where p^+ is 1 for $l + l' + L$ even and 0 for $l + l' + L$ odd. Aiming to build a robust estimator of κ satisfying $\langle \hat{\kappa}_{LM} \rangle = \kappa_{LM}$, Ref. [34] proposed the form

$$(\hat{\kappa}_{LM}^{XY})^* = A_L^{XY} \sum_{ll'mm'} \begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix} \frac{(f_{ll'L}^{XY})^*}{\Delta^{XY}} \bar{X}_{lm} \bar{Y}_{l'm'}, \quad (10)$$

where $\Delta^{XY}|_{X \neq Y} = 1$, $\Delta^{XY}|_{X=Y} = 2$ and the fields $\bar{X}_{lm}$ and $\bar{Y}_{l'm'}$ are the inverse-variance-filtered observed fields, related to the Wiener-filtered fields in Eq. (7) by $\bar{X}_{lm} = \hat{X}_{lm}^{WF}/C_l^{XX}$. Taking the ensemble average of Eq. (10) and substituting Eq. (8) for $\langle \hat{X}_{lm} \hat{Y}_{l'm'} \rangle$, we find that the contribution from the diagonal term vanishes for $L > 0$ and the off-diagonal term allows to determine the normalization factor

$$A_L^{XY} = \left[\sum_{ll'} \frac{(2L+1)^{-1} |f_{ll'L}^{XY}|^2}{\Delta^{XY} (C_l^{XX} + N_l^{XX})(C_{l'}^{YY} + N_{l'}^{YY})} \right]^{-1}. \quad (11)$$

Again, Eq. (9) refers to unlensed fields, however the lensed ones used in practice have been found to provide a better approximation to the nonperturbative result for correlators involving $\hat{\kappa}_{LM}$ [35].

The expression (10) with A_L^{XY} given by Eq. (11) corresponds to the quadratic estimator (QE) of the lensing convergence as described in Ref. [34]. All field pairings listed in Eq. (9) ($\Theta\Theta$, ΘE, EE and EB) are included in the present work, and are coadded with weights minimizing the reconstruction noise variance at each multipole. We do not use the ΘB combination due to its low signal-to-noise ratio. As explained in Refs. [36] and [37], the EB QE becomes a source of bias if the multipole range of the fields used to build it overlaps with that of the B-modes we aim to delens. We therefore only consider modes between $l = 301$ and $l = 4096$ in our lensing reconstruction; the temperature field is further restricted to $500 < l < 3000$ in order to avoid large-scale atmospheric noise and residual small-scale

extragalactic foreground contributions [38,39]. The latter are not included in our simulations at this stage, but will be important for the analysis of real SO data. Finally, note that noise inhomogeneities and the presence of a survey mask can cause the ensemble average $\langle \hat{\kappa}_{LM}^{XY} \rangle$ to become nonzero. Following Ref. [18], we mitigate this bias for each QE by subtracting a mean-field correction computed from the average of our simulations.

Combining all possible internal tracers of κ still does not lead to optimal delensing performance. Indeed, reconstruction noise resulting from statistical fluctuations of the unlensed CMB is a significant limitation at intermediate and high multipoles which are particularly relevant for delensing. Quadratic estimators should therefore be complemented by external large-scale structure tracers such as galaxy surveys or measurements of the CIB. For future SO analysis, the latter will be extracted from *Planck* data using, for example, the GNILC algorithm [40], while the DESI Legacy Imaging Surveys [41], unWISE [42] and the upcoming LSST [43] will provide information on the spatial distribution of galaxies. Other types of internal lensing reconstruction techniques, for example likelihood-based estimators [44], are not included in this work as they are not expected to outperform QEs significantly for SO's nominal configuration [9].

Coadding all tracers $\hat{\kappa}^i$ with weights chosen to maximize the correlation between our final estimator $\hat{\kappa}^{\mathrm{comb}}$ and the true lensing convergence κ, we obtain[3] [46]

$$\hat{\kappa}_{LM}^{\mathrm{comb}} = \sum_{ij} (\rho^{-1})_L^{ij} \rho_L^{j\kappa} \sqrt{\frac{C_L^{\kappa\kappa}}{C_L^{\hat{\kappa}^i \hat{\kappa}^i}}} \hat{\kappa}_{LM}^i. \quad (12)$$

In Eq. (12), ρ_L^{ij} and $\rho_L^{i\kappa}$ represent, respectively, correlation coefficients between tracers $\hat{\kappa}^i$ and $\hat{\kappa}^j$, or between $\hat{\kappa}^i$ and the true lensing convergence. These quantities, which are the components of the correlation matrix ρ_L, were computed in Ref. [47] for the simulated tracers used in this work. As shown in Fig. 2 of Ref. [18], internal reconstruction is the most relevant tracer for $L < 250$; at smaller scales ($250 < L < 1000$), both the CIB and galaxy overdensity maps have a higher degree of correlation with the true convergence for the noise level of the SO LAT.

The lensing B-mode template, built by performing the weighted convolution of $\hat{E}_{lm}^{\mathrm{WF}}$ (with $50 < l < 2048$) and $\hat{\kappa}_{LM}^{\mathrm{comb}}$ (with $20 < L < 2048$) as indicated in Eq. (3), can then be subtracted from observations at the map level (over the regions observed by both the SATs and the LAT) or

---

[3]Here, we ignore correlations between multipoles. Further optimization would be possible by accounting for such effects as was done in Ref. [45]. However, in our setup, the external tracers are Gaussian fields with no correlations between multipoles, so the delensing efficiency from simulations is close to that obtained with idealized forecasts; extending Eq. (12) would thus not significantly improve $\sigma(r)$.

cross-correlated with the data in order to infer tighter constraints on r. While this paper focuses on the latter approach, both techniques will ultimately be used for SO and are expected to yield equivalent results (see Sec. III).

Finally, note that the ensemble-average cross-spectra of the lensing template with SAT B-modes are identical to the template autospectrum for ideal (statistically isotropic) surveys. Indeed, Eq. (12) implies

$$C_L^{\kappa \hat{\kappa}^{\mathrm{comb}}} = C_L^{\hat{\kappa}^{\mathrm{comb}} \hat{\kappa}^{\mathrm{comb}}} = C_L^{\kappa\kappa} \sum_{ij} \rho_L^{i\kappa} (\rho^{-1})_L^{ij} \rho_L^{j\kappa}. \quad (13)$$

Considering the Wiener filter in Eq. (7) and the fact that $C_l^{E\hat{E}} = C_l^{EE}$ as the noise component is uncorrelated with the underlying E-modes, we then obtain the following expression for the cross-spectrum between the template (denoted by the superscript B_t) and the true lensing B-modes:

$$C_l^{B_t B} = \sum_{l'L} |\mathcal{M}(l, l', L)|^2 \frac{C_{l'}^{EE}}{C_{l'}^{\hat{E}\hat{E}}} C_{l'}^{E\hat{E}} C_L^{\kappa \hat{\kappa}^{\mathrm{comb}}}$$

$$= \sum_{l'L} |\mathcal{M}(l, l', L)|^2 \frac{C_{l'}^{EE}}{C_{l'}^{\hat{E}\hat{E}}} C_{l'}^{EE} C_L^{\kappa \hat{\kappa}^{\mathrm{comb}}}, \quad (14)$$

where all the prefactors from Eq. (3) have been grouped into $\mathcal{M}(l, l', L)$ for legibility. Similarly, the lensing template autospectrum is given by

$$C_l^{B_t B_t} = \sum_{l'L} |\mathcal{M}(l, l', L)|^2 \left(\frac{C_{l'}^{EE}}{C_{l'}^{\hat{E}\hat{E}}} \right)^2 C_{l'}^{\hat{E}\hat{E}} C_L^{\hat{\kappa}^{\mathrm{comb}} \hat{\kappa}^{\mathrm{comb}}}$$

$$= \sum_{l'L} |\mathcal{M}(l, l', L)|^2 \frac{C_{l'}^{EE}}{C_{l'}^{\hat{E}\hat{E}}} C_{l'}^{EE} C_L^{\kappa \hat{\kappa}^{\mathrm{comb}}}. \quad (15)$$

The equality between Eqs. (14) and (15) was verified to hold approximately for our simulations, despite the anisotropic LAT noise. (We note that the external-tracer simulations are masked by the LAT footprint but are otherwise statistically isotropic.) In an effort to reduce computational costs, the power spectrum obtained from the average of simulated lensing templates was therefore also used in our component separation pipeline to model cross-spectra between the template and observed B-modes from the different SAT channels. When working with real data, the template autospectrum may contain foreground residuals and should then be modeled separately from the cross-spectra.

B. Galactic foreground model

We now introduce the parametric model used to characterize SAT Galactic foregrounds and distinguish them from the CMB signal. The present analysis focuses on thermal dust emission and synchrotron radiation, as they

are the two most prominent polarized sources over the scales of interest [21].

1. Baseline complexity

In its simplest form, our model assumes foreground properties to be well described by their sky average; the SEDs therefore depend on a restricted set of parameters that do not vary with position.

The dominant contribution at low frequencies comes from synchrotron radiation, produced by the helical motion of high-energy cosmic-ray electrons around Galactic magnetic field lines. Integrating the Larmor formula over a population of electrons with a power-law energy distribution, we obtain the SED $S_\nu^s = (\nu/\nu_0^s)^{\beta_s}$ [48] in Rayleigh-Jeans units, where $\beta_s \approx -3$ and we fix $\nu_0^s = 23$ GHz. As a consequence of its negative spectral index, synchrotron radiation is subdominant above around 70 GHz where thermal emission from dust grains heated by starlight becomes the most important foreground source.

The dust component peaks in the far infrared and is polarized due to the alignment of dust particles with Galactic magnetic field lines, which arises from the action of radiative torques on irregularly shaped grains [49]. A modified blackbody function

$$I_\nu^d \propto \nu^{\beta_d} \frac{2h\nu^3}{c^2} \frac{1}{\exp(h\nu/k_B T_d) - 1} \quad (16)$$

was empirically found to be a good approximation of its intensity spectrum [50]. The SED, $S_\nu^d = I_\nu^d/I_{\nu_0^d}^d$, then depends on the positive spectral index β_d as well as on the interstellar dust temperature T_d and the reference frequency ν_0^d, which we fix at 353 GHz. Computing the sky average of the dust SED maps extracted from the *Planck* 2015 Commander analysis [51] yielded the reference values $\beta_d \approx 1.54$ and $T_d \approx 20$ K, which are consistent with subsequent results from power-spectrum-based methods [52].

Planck, WMAP and S-PASS observations have also shown that the foreground angular power spectra can be parametrized by power laws $C_l^c = A_c(l/l_0)^{\alpha_c - 2}$, where $l_0 = 80$ and $c = d$ or s for dust [52] and synchrotron [53] respectively. For each component, the slope of the angular power spectrum remains identical in all frequency channels, while its amplitude is multiplied by the product of the two relevant SEDs.

A dust-synchrotron correlation parameter ϵ_{ds} is also included in our model and contributes terms proportional to $C_l^{ds} = \epsilon_{ds}\sqrt{C_l^d C_l^s}$ to the power spectra. This brings the total number of free foreground parameters to seven: ϵ_{ds}, β_d, A_d, α_d, β_s, A_s and α_s.

Note that we do not include the Faraday rotation of primary E-modes into B-modes by Galactic or primordial magnetic fields in our foreground model. As the power

spectrum of such B-modes scales as $(\frac{\nu}{30}$ GHz$)^{-4}$ and peaks at $l \sim 1000$ [54], this effect is expected to be negligible for the frequency bands and multipole range of the SO SATs.

2. Spatially varying SEDs

While the model described above is a useful approximation for pipeline validation purposes, it does not fully capture the complexity of realistic foreground emission. Indeed, the SEDs actually exhibit spatial variations due to inhomogeneous dust temperature and grain types as well as fluctuations in the cosmic-ray energy distribution. Treating them as constants, especially over the large fractions of sky targeted by SO, has been shown to produce systematic residuals leading to significant biases in measurements of r [24].

In order to mitigate this effect, we apply the moment-expansion approach described in Ref. [28]. Assuming small spatial fluctuations $\delta\beta_c(\hat{n})$ with respect to the sky average $\bar{\beta}_c$, we can expand the foreground contributions to second order at the map level[4]:

$$m_\nu(\hat{n}) = \sum_c T_c(\hat{n}) S_\nu^c(\beta_c(\hat{n}))$$

$$= \sum_c [T_c(\hat{n})(\bar{S}_\nu^c + \delta\beta_c(\hat{n})\partial_{\beta_c}\bar{S}_\nu^c + \frac{1}{2}(\delta\beta_c)^2(\hat{n})\partial_{\beta_c}^2\bar{S}_\nu^c)]. \quad (17)$$

Here, $m_\nu(\hat{n})$ corresponds to the total foreground Stokes parameters at frequency ν, $T_c(\hat{n})$ represents the amplitudes of the Stokes parameters of a given component at the reference frequency ν_0^c and the bars denote the SED or its derivatives evaluated at $\bar{\beta}_c$.

We can now propagate the expansion in Eq. (17) to the cross-spectrum of the B-modes at frequencies ν and ν', using some simplifying assumptions justified in more detail in Ref. [28]. The leading-order term,

$$C_{l,0}^{\nu\nu'} = \bar{S}_\nu^d \bar{S}_{\nu'}^d C_l^d + \bar{S}_\nu^s \bar{S}_{\nu'}^s C_l^s + (\bar{S}_\nu^d \bar{S}_{\nu'}^s + \bar{S}_\nu^s \bar{S}_{\nu'}^d)C_l^{ds}, \quad (18)$$

corresponds to the sky-average model presented in the previous subsection. While correlations between dust and synchrotron amplitudes are accounted for in Eq. (18), we consider $\delta\beta_d$ and $\delta\beta_s$ to be independent. We further assume independence between $T_c(\hat{n})$ and $\delta\beta_c(\hat{n})$. Under these conditions, the first-order terms vanish and the second-order contribution to the power spectrum is given by

---

[4]Variations in the dust temperature T_d are almost degenerate with amplitude variations at the frequencies of interest, and contribute very little to SED variations. For this reason, we do not need to include the dust temperature in the moment expansion.

$$C_{l,2}^{\nu\nu'} = \sum_c \left[\partial_{\beta_c} \bar{S}_\nu^c \partial_{\beta_c} \bar{S}_{\nu'}^c \sum_{l_1 l_2} \frac{(2l_1+1)(2l_2+1)}{4\pi} \begin{pmatrix} l & l_1 & l_2 \\ 0 & 0 & 0 \end{pmatrix}^2 \right.$$
$$\times C_{l_1}^c C_{l_2}^{\beta_c} + \frac{1}{2} (\bar{S}_\nu^c \partial_{\beta_c}^2 \bar{S}_{\nu'}^c + \bar{S}_{\nu'}^c \partial_{\beta_c}^2 \bar{S}_\nu^c) C_l^c$$
$$\left. \times \sum_{l'} \frac{2l'+1}{4\pi} C_{l'}^{\beta_c} \right]. \qquad (19)$$

Further details of the derivation can be found in Ref. [28]; the important takeaway of Eq. (19) is that only one additional power spectrum per component, $C_l^{\beta_c}$, needs to be modeled in order to compute the effects of foreground spatial variability to second order in $\delta\beta_c(\hat{n})$. Parametrizing it as a power law, $C_l^{\beta_c} = B_c(l/l_0)^{\gamma_c}$, leads us to extend our simple model by four variables. The full set of foreground parameters to be sampled is now given by ϵ_{ds}, β_d, A_d, α_d, B_d, γ_d, β_s, A_s, α_s, B_s and γ_s.

C. Likelihood analysis

In the likelihood analysis, we use the measured B-mode cross-spectra to infer the probability distributions of the foreground and cosmological parameters. Even if the underlying fields were Gaussian, dealing with the exact likelihood for the measured spectra in the presence of survey masking and inhomogeneous noise is intractable. We therefore adopt an approximate Gaussian likelihood for the measured spectra with a fixed covariance matrix. This approximation is expected to hold for large enough sky coverage and when enough multipoles are binned together and averaged, by virtue of the central-limit theorem.

The Gaussian likelihood for the power spectra is

$$-2 \ln \mathcal{L}(\{\mathbf{X}_l\}|\{\hat{\mathbf{X}}_l\}) \approx (\hat{X} - X)^T \mathbf{M}_f^{-1} (\hat{X} - X), \qquad (20)$$

where the vectors $\hat{X}$ and X contain the measured cross-spectra and the model at all considered multipoles, respectively, and $\mathbf{M}_f$ is the fixed covariance matrix of the spectra. The covariance is precomputed from a set of simulations at fiducial parameter values and does not need to be recomputed as parameter space is explored, greatly reducing computational costs. A further advantage of working with a spectral likelihood is that non-Gaussianity of the fields (such as from lensing) can be dealt with approximately through the covariance matrix.

We also tested the approximate Hamimeche and Lewis likelihood [55], which similarly makes use of a fiducial covariance but additionally approximately captures the non-Gaussian shape of the likelihood. We compare results obtained with this likelihood and the Gaussian likelihood in Appendix A. As shown there in Fig. 9, the posterior distributions for the parameters are in excellent agreement; the Gaussian likelihood will therefore be used throughout this work as it leads to faster sampling. Note that the Gaussian approximation holds for SO due to the relatively large sky

patch allowing to average over many modes. For experiments targeting smaller areas such as BICEP/Keck, non-Gaussian likelihood approximations, such as Hamimeche and Lewis, must be used (especially at low l).

III. EQUIVALENCE OF CROSS-SPECTRAL AND MAP-BASED DELENSING

We now compare the expected performance of the power-spectrum-based approach presented above with that of map-based delensing, which will also likely be performed on future SO data. While directly subtracting the lensing B-mode template from the observed maps might seem more straightforward, cross-spectral methods have significant advantages for ground-based surveys. For example, complex filtering operations used to mitigate atmospheric disturbances are easier to take into account in harmonic space, and data splits can be used to provide an accurate estimate of the noise power spectrum (see Sec. V).

In an idealistic situation without such filtering, equivalent constraints on r can be obtained with both methods. In order to demonstrate this, let us first consider a foreground-cleaned map from which we subtract the lensing template. The B-mode harmonic coefficients are given by $B_{lm}^{\text{del}} = B_{lm} + n_{lm} - B_{t,lm}$, where the subscript t refers to the template. Assuming negligible foreground residuals, approximately Gaussian lensing B-modes and no correlations between the noise n_{lm} and the other components, the power spectrum of this map then corresponds to

$$C_l^{BB,\text{del}} = C_l^{BB} - 2C_l^{B_t B} + N_l + C_l^{B_t B_t}$$
$$= C_l^{BB} - C_l^{B_t B_t} + N_l, \qquad (21)$$

where we used the equality of Eqs. (14) and (15) (which is verified later for our simulations in Fig. 3) to simplify the result. With the likelihood in Eq. (5), the Fisher information for r at multipole l is given by

$$\mathcal{I}_l(r) = \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial r^2} \right\rangle = \frac{2l+1}{2} \left(\frac{\partial_r C_l^{BB}}{C_l^{BB} - C_l^{B_t B_t} + N_l} \right)^2. \qquad (22)$$

The error on r is the same as in Eq. (6) with C_l^{lens} replaced by its delensed counterpart.

Let us now consider the foreground-cleaned SAT map and the template separately. Again equating $C_l^{B_t B}$ and $C_l^{B_t B_t}$, the matrix $\mathbf{C}_l$ containing all auto- and cross-spectra becomes

$$\mathbf{C}_l = \begin{pmatrix} C_l^{BB} + N_l & C_l^{B_t B_t} \\ C_l^{B_t B_t} & C_l^{B_t B_t} \end{pmatrix}. \qquad (23)$$

The Fisher information can be computed from the exact, full-sky likelihood for Gaussian fields, Eq. (A1); using $\ln |\mathbf{C}_l| = \text{Tr}(\ln \mathbf{C}_l)$, we obtain

$$\mathcal{I}_l(r) = \frac{2l+1}{2} \text{Tr}(\mathbf{C}_l^{-1} \partial_r \mathbf{C}_l \mathbf{C}_l^{-1} \partial_r \mathbf{C}_l). \qquad (24)$$

We can then substitute Eq. (23) into this expression, where C_l^{BB} is the only r-dependent element (neglecting small tensor contributions to the internal lensing estimators). As expected, the result is identical to Eq. (22) and we conclude that the map-based and the cross-spectral method lead to equivalent uncertainties on r. We verify this equivalence empirically in Sec. VI. Comparing constraints obtained with both techniques will therefore be a useful cross-check for future SO data.

Note that our reasoning was demonstrated on an idealistic full-sky situation (and with no foreground residuals) for simplicity; however, the argument still holds in the presence of a survey mask, provided that the same pixel weighting is used on all maps so that the cross-spectral likelihood takes as input all of the pseudospectra required to form the pseudospectrum of the map after map-based foreground cleaning and delensing.

IV. OPTIMAL PIXEL WEIGHTING

For computational efficiency, we use local pixel weighting when measuring the power spectrum (i.e., we construct pseudo-C_l estimates) rather than full inverse-variance weighting, which is nonlocal and requires inversion of large matrices. This local weighting scheme can be chosen to minimize the variance of the measured spectra (after deconvolution of the effects of the masking and weighting). Building upon the results of Sec. III, we construct optimal local pixel weights for a foreground-cleaned and delensed SAT map; these will be applied to all SAT channels as well as to the lensing template when computing pseudo-C_l power spectrum estimates. A natural choice is the inverse-variance weighting

$$w_i \propto \frac{1}{\sigma_{i,\text{res}}^2 + \sigma_{i,\text{noise}}^2}, \qquad (25)$$

where both the lensing residuals and the noise contribution (assumed to be constant over the pixel size) are taken into account.

This scheme can be heuristically motivated as follows. Assume that any statistical anisotropy due to the survey and the pixel weighting, are slowly varying compared to the scales of interest. We can then approximate the weighting as dividing up the sky into patches, large enough that the fluctuations are approximately uncorrelated between patches but small enough that the weights and statistical properties of the fluctuations can be treated as constant within each patch. Denote these patch weights by w_i, where i labels the patch. The observed B-mode power spectrum, corrected for the weighting, can then be approximated by a sum over patches:

$$\hat{C}_l \approx \left(\sum_i w_i^2 \tilde{C}_l^i \right) / \langle w^2 \rangle_{\text{patch}}, \qquad (26)$$

where $\tilde{C}_l^i$ is the pseudo-C_l power spectrum in patch i obtained with binary weighting (one in the patch and zero outside). The expected value of $\tilde{C}_l^i$ is approximately

$$\langle \tilde{C}_l^i \rangle \approx f_{\text{sky},i}(C_l^i + N_l^i), \qquad (27)$$

where C_l^i is the signal power spectrum in patch i, which may differ between patches due to statistical anisotropy (e.g., due to variation in delensing efficiency), N_l^i is the local noise power spectrum there, and $f_{\text{sky},i}$ is the sky fraction of the patch. The normalization factor in Eq. (26),

$$\langle w^2 \rangle_{\text{patch}} \equiv \sum_i w_i^2 f_{\text{sky},i}, \qquad (28)$$

corrects for the effects of the weights and approximates the mask-deconvolution step in unbiased power-spectrum estimation.

The variance of $\hat{C}_l$ then corresponds to

$$\sigma^2(\hat{C}_l) \approx \left(\sum_i w_i^4 \sigma^2(\tilde{C}_l^i) \right) / \langle w^2 \rangle_{\text{patch}}^2, \qquad (29)$$

where the variance of the $\tilde{C}_l^i$ is [56]

$$\sigma^2(\tilde{C}_l^i) \approx \frac{2}{(2l+1)f_{\text{sky},i}} \langle \tilde{C}_l^i \rangle^2. \qquad (30)$$

Minimizing Eq. (29) with respect to w_i leads to the weights

$$w_i \propto \frac{1}{C_l^i + N_l^i}; \qquad (31)$$

as both the lensing B-modes and the noise power spectrum are approximately constant over the multipole range considered in our analysis ($30 \leq l \leq 300$), we indeed recover Eq. (25).

The lensing residual term $\sigma_{i,\text{res}}^2$ is equivalent to $(5 \text{ }\mu\text{K-arcmin})^2$ white noise multiplied by a factor A_{lens} resulting from delensing. Due to the relative homogeneity of the LAT hit-count map (see Fig. 1), we take A_{lens} to be uniform, computing it as[5]

$$A_{\text{lens}} = 1 - \overline{\left(\frac{(C_l^{B_tB})^2}{C_l^{BB} C_l^{B_tB_t}} \right)}, \qquad (32)$$

---

[5]We give this more general form for completeness; in our analysis, the simpler form $A_{\text{lens}} = 1 - \overline{C_l^{B_tB}/C_l^{BB}}$ is very nearly equivalent due to the equality of $C_l^{B_tB}$ and $C_l^{B_tB_t}$.

TABLE I.   Noise and beam specifications used to produce the simulations in this work, corresponding to the goal/optimistic case of Refs. [9] and [24]. Note that the white-noise levels are given for the central area of the map assuming 5 years of observations, and are weighted according to the hit counts when generating the final noise maps. We use $l_{knee} = 700$ and $\alpha_{knee} = -1.4$ for all LAT channels.

| Frequency [GHz] | SAT | | | | | LAT | |
	FWHM [arcmin]	$\sigma_{\text{white},\nu}$ [μK-arcmin]	ℓ_{knee}	α_{knee}		FWHM [arcmin]	$\sigma_{\text{white},\nu}$ [μK-arcmin]
27	91	33	15	−2.4		7.4	74
39	63	22	15	−2.4		5.1	38
93	30	2.5	25	−2.5		2.2	8.2
145	17	2.8	25	−3.0		1.4	8.9
225	11	5.5	35	−3.0		1.0	21
280	9	14	40	−3.0		0.9	52

where the bar refers to an average over *l*, on the whole area of overlap between the surveys with the SATs and the LATs [18]. For future full-sky surveys such as LiteBIRD, which may be combined with a range of partially overlapping external tracers, the template properties might not be as uniform [45,57]. In this case, A_{lens} could be treated as a spatially varying quantity. The noise term $\sigma^2_{i,\text{noise}}$ corresponds to the minimal white-noise variance in the deepest area of the map (hereafter referred to as the central white-noise variance σ^2_{white}) modulated by the relative number of hits:

$$\sigma^2_{i,\text{noise}} = \frac{\sigma^2_{\text{white}}}{N_i^{\text{hit}}}. \qquad (33)$$

In the case of noise-dominated data, Eq. (25) therefore reduces to the hit-count mask used in Ref. [24]; however, this is suboptimal for SO as the residual lensing variance surpasses the central white noise variance of the mid-frequency SAT channels. Similarly, the uniform weighting used in Ref. [18] is only appropriate for areas in which noise inhomogeneities are negligible; mitigating the strongly increasing noise variance towards the outer regions of the map then requires aggressive apodization resulting in a loss of data. The strategy proposed here naturally transitions between these two extremes when moving from the edges to the center of the survey area.

With the lensing term characterized by Eq. (32), the final step towards building our mask is the determination of the central noise variance σ^2_{white} for a foreground-cleaned coaddition of SAT multifrequency maps. For this purpose, we approximate the map-based foreground cleaning assuming known SEDs for all components.

In detail, we express the data $d_\nu(\hat{\boldsymbol{n}})$ at frequency ν as a sum of a noise contribution $n_\nu(\hat{\boldsymbol{n}})$ and three components $s_c(\hat{\boldsymbol{n}})$ (the CMB, dust and synchrotron) multiplied by their respective SEDs, which make up the mixing matrix **A** [58]:

$$d_\nu(\hat{\boldsymbol{n}}) = \sum_c A_{\nu c} s_c(\hat{\boldsymbol{n}}) + n_\nu(\hat{\boldsymbol{n}}). \qquad (34)$$

Assuming the noise to be Gaussian with covariance matrix $\mathbf{N} = \text{diag}(\sigma^2_{\text{white},1}, \ldots, \sigma^2_{\text{white},6})$, where the central levels for individual channels are listed in Table I, we get the likelihood function

$$-2 \log \mathcal{L} = (\boldsymbol{d} - \mathbf{A}\boldsymbol{s})^T \mathbf{N}^{-1} (\boldsymbol{d} - \mathbf{A}\boldsymbol{s}). \qquad (35)$$

We aim to compute the error on the element of $\boldsymbol{s}$ corresponding to the CMB [59]. The associated Fisher information is given by

$$\mathcal{I}(s_{\text{CMB}}) = \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial s^2_{\text{CMB}}} \right\rangle = (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})_{\text{CMB,CMB}}, \qquad (36)$$

leading to the variance

$$\sigma^2(s_{\text{CMB}}) = \sigma^2_{\text{white}} = (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})^{-1}_{\text{CMB,CMB}}, \qquad (37)$$

which we can then substitute into Eq. (25). Considering the goal values presented in Table I, Eq. (37) indicates a central white noise level of 2.5 μK-arcmin for the foreground-cleaned map.

V. PIPELINE INPUTS AND LOGISTICS

A. Simulations

Full-sky CMB *Q* and *U* polarization maps are simulated as lensed Gaussian realizations of CMB power spectra computed with CAMB [60] for the *Planck* best-fit cosmological parameters listed in Ref. [61]. One set of input simulations is produced with $r = 0$ (no primordial *B*-modes) and the other one with $r = 0.01$. The lensing operation is performed using the pixell package. These maps are generated at a high resolution ($N_{\text{side}} = 2048$ in HEALPix [62]), then convolved with symmetrical Gaussian beams whose FWHMs are listed in Table I.

Our noise model accounts for detector white noise as well as a $1/f$ component related to timestream–level filtering (which mainly targets the mitigation of atmospheric noise), leading to the following angular power spectrum:

$$N_l = N_{\text{white}} \left[1 + \left(\frac{l}{l_{\text{knee}}} \right)^{\alpha_{\text{knee}}} \right]. \qquad (38)$$

The amplitude N_{white} is adjusted for each instrument and frequency band in order for the central variance in the noise maps to match the values listed in Table I. Current or past experiments, such as QUIET [63] near SO's observing site and BICEP/Keck [64] at the South Pole, have allowed to calibrate l_{knee} and α_{knee} for the six SAT frequency channels (see Table I). For the LAT, we set $l_{\text{knee}} = 700$ and $\alpha_{\text{knee}} = -1.4$ [9]; these estimates based on ACT [65] noise levels are currently used for all LAT bands, but future work will aim to characterize the frequency dependency of these parameters. Noise maps are built by generating Gaussian realizations of the power spectrum in Eq. (38) (with $l_{\text{max}} = 6144$ and 1536 for the LAT and SAT, respectively), then multiplying each pixel i by $1/\sqrt{N_i^{\text{hit}}}$, where N_i^{hit} refers to the LAT or SAT normalized hit counts shown in Fig. 1.

The inputs of the internal lensing reconstruction stage simply consist of the CMB realizations described above convolved with LAT beams for three different frequency channels (93, 145 and 225 GHz), to which LAT-like noise is added. These maps do not contain any Galactic or extragalactic foregrounds, and therefore do not require point-source masks.

In addition to LAT data, the template construction process relies on external lensing tracers. The present work follows Ref. [18] and uses Gaussian simulations of the CIB as well as of a galaxy density field mimicking the forecasted LSST gold sample [43]. The latter is split into six redshift bins with edges $z = [0, 0.5, 1, 2, 3, 4, 7]$. Relevant auto- and cross-spectra were computed in Ref. [47] assuming linear galaxy bias. In order to ensure accurate correlation between the external tracers and the corresponding realization of the CMB lensing potential, we implement the algorithm described in Appendix F of Ref. [66]. The LAT binary mask derived from the hit counts in the left panel of Fig. 1 is applied to all input maps, both internal and external, throughout the template construction step.

The next stage of the pipeline consists of extracting cosmological information from SAT polarization data. To simulate such observations, we use the same lensed CMB realizations as for the LAT-like maps, downgrade the resolution to $N_{\text{side}} = 512$, convolve them with SAT beams in six frequency channels and add noise following Eq. (38) with the parameters listed in the left columns of Table I. Our choice to use the goal/optimistic noise levels from Ref. [9] is motivated by the recent approval of the SO:UK and SO:Japan projects, which are constructing three additional SATs; indeed, even if the nominal design of SO were to fall slightly short of these targets, our forecasts would still be conservative in light of the expected performance of the extended configuration. When analyzing real SAT observations, noise biases in the spectra will be mitigated

by the use of data splits (see Sec. V B), corresponding to maps built from different time periods throughout the survey duration. In this paper, we use four splits (as was done for the ACT DR4 maps [67]); this process is simulated by generating four uncorrelated noise realizations for each CMB map and multiplying the amplitudes from Table I by $\sqrt{N_{\text{splits}}} = 2$.

In order to assess the impact of Galactic foregrounds on delensing performance, our SAT-like simulations contain both dust and synchrotron emission. While past experiments have provided valuable information on these contaminants, the fluctuations of their SEDs across the sky have not been fully characterized yet. We therefore investigate three foreground models based on the PySM package templates [68] and including different degrees of complexity. All three models rely on dust and synchrotron amplitude maps established, respectively, by *Planck* at 353 GHz [51] and WMAP at 23 GHz [69]. The Commander component separation code also allowed to estimate the spatial distribution of the dust spectral index β_d and temperature T_d from the aforementioned *Planck* data set. In the case of synchrotron radiation, spectral properties were inferred by combining WMAP data with the Haslam map at 408 GHz [70,71]. The d1s1 foreground model used in our analysis is built upon this information.

A simpler model with uniform SEDs, referred to as d0s0, is derived from the d1s1 templates by averaging the spectral parameters over the whole sky. The resulting values are $T_d = 20$ K, $\beta_d = 1.54$ and $\beta_s = -3$. While it does not capture the full complexity of realistic foregrounds, this model is a very useful approximation for pipeline validation purposes as it allows the component separation process to be performed at a significantly lower computational cost.

On the other hand, more recent data from S-PASS at 2.3 GHz [72] seems to indicate that the synchrotron spectral index varies more significantly than in the original PySM template. A more complex model called dmsm is therefore constructed by rescaling the d1s1 map of the fluctuations $\beta_s - \bar{\beta}_s$, using an amplification factor of 1.6. In this model, the spectral parameter templates for both dust and synchrotron are also smoothed at an angular resolution of 2 deg.. in order to mitigate instrumental noise residuals, which might have affected the d1s1 maps. Gaussian fluctuations are then artificially added below this scale.

Foreground maps are convolved with the SO SAT-like Gaussian beams listed in Table I before being combined with the CMB and noise simulations, and are not subjected to bandpass integration. The choice of working with delta-function bandpasses simply aims to reduce computational costs here and does not significantly affect the results. (The analysis pipeline is also equipped to deal with realistic SO bandpasses.)

Finally, note that the foreground templates described above are informed by real data and are therefore only

available in one realization. In order to evaluate the fiducial covariance matrix required to compute the likelihoods in Eqs. (20) and (A6), 500 random Gaussian foreground simulations are generated using the angular power spectra mentioned in Sec. II B 1 with reference parameters obtained from the d0s0 dust and synchrotron maps. The best-fit power-law parameters depend on the considered masking scheme, and thus exhibit slight variations with the value of A_{lens} in the optimal weights of Eq. (25). As the shifts observed in our case do not exceed 10%, we use the same set of Gaussian simulations with and without delensing, and average the best-fit parameters obtained with $A_{\text{lens}} = 1$ and $A_{\text{lens}} = 0.35$. The resulting reference values are $A_d = 44.6$, $\alpha_d = -0.16$, $A_s = 1.07$ and $\alpha_s = -0.78$. The suitability of Gaussian simulations for covariance matrix estimation has been verified through χ^2 tests in Appendix A of Ref. [24]; furthermore, Ref. [73] showed that modifying the covariance matrix to account for dust non-Gaussianity did not impact constraints on r for an experiment with SO's characteristics.

B. Workflow

The general working principle and successive stages of the analysis pipeline are summarized in Fig. 2. When
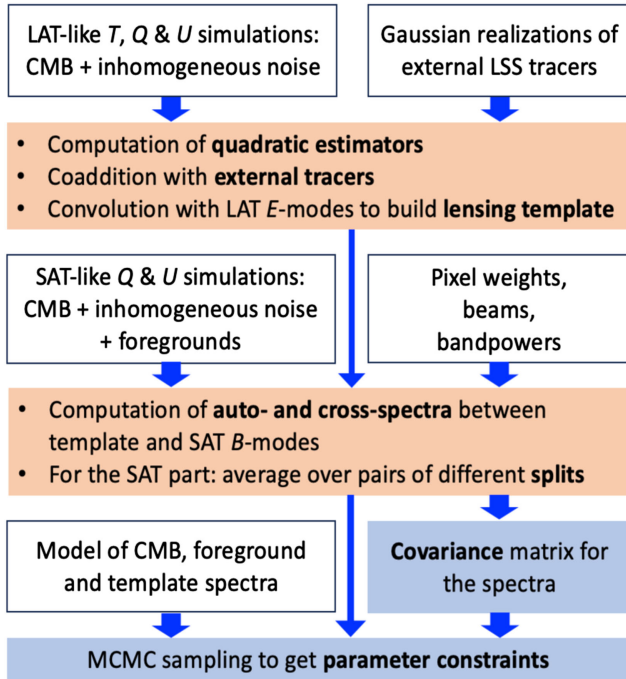
FIG. 2. Flowchart of the three main pipeline stages: lensing template construction, computation of power spectra, and component separation. Inputs are listed in the white boxes. The same CMB realizations are used to generate the LAT and SAT maps, and the Gaussian realizations of external LSS tracers are appropriately correlated with the CMB lensing convergence. Orange boxes represent stages carried out on both the mock data and the 500 fiducial simulations, while the blue-colored operations are only applied to the mock data.

running the algorithm, one input map with either d0s0, d1s1 or dmsm foregrounds is selected as mock data to run the parameter inference on, and the 500 simulations with Gaussian foregrounds are only used to compute the fiducial covariance matrix.

In the first stage, a binary mask based on the hit counts in Fig. 1 and apodized by 5 deg is applied to all LAT-like simulations. The masked maps at 93, 145 and 225 GHz are then converted to harmonic space and E-mode coefficients are combined with inverse-noise-variance weighting. Finally, these E-modes are diagonally Wiener filtered as indicated in Eq. (7), where the coadded noise power spectrum N_l^{EE} is obtained from noise-only simulations over the observed area.

The internal lensing convergence reconstruction process uses unfiltered E- and B-modes as well as temperature information extracted from the same LAT-like input maps. Diagonal inverse-variance filtering is directly incorporated in Eq. (10), which is applied to four pairs of observed fields ($\Theta\Theta$, ΘE, EE and EB) in order to compute the QEs.

Correlation coefficients between the minimum-variance combination of QEs and the external mass tracers, as well as between each tracer and the input lensing convergence are evaluated for all simulations. These values are averaged over the 500 CMB realizations and used in Eq. (12) to coadd our internal reconstruction with the masked external-tracer maps. Convolving this combined estimator with the Wiener-filtered E-modes as indicated in Eq. (3) yields the harmonic coefficients of the lensing B-mode template, which we then convert into real-space Q and U maps. Note that the method used here to estimate the optimal coaddition of mass tracers will not be applicable to real data, for which only one realization is available and the true lensing convergence is not known. Instead, we obtain the coefficients of Eq. (12) by fitting theoretical models to the measured auto- and cross-spectra of our tracers. Uncertainties on these spectra were investigated in Ref. [18] and found not to have a significant impact on $\sigma(r)$ at SO's sensitivity level.

Throughout the second stage of the pipeline, both the SAT-like maps and the real-space Q, U template are weighted according to Eq. (25), with the lensing and noise variance estimated as described in Sec. IV. Furthermore, the analysis region is restricted to the overlap between the SAT and the LAT survey areas, and an apodization length of 10 deg is applied. This additional smoothing is necessary for the B-mode purification process described below.

In the usual pseudo-C_l formalism, the presence of such a mask/weighting leads to a mixing of the E- and B-modes as well as to the appearance of a coupling between different multipoles when computing auto- and cross-spectra between the observed SAT B-modes and the lensing template. Let us define the masked polarization vector $\boldsymbol{P}_{lm}^w = (E_{lm}^w, B_{lm}^w)^T$, which contains the harmonic coefficients of the sky maps multiplied by the weights w_i at the

pixel level. The power spectrum estimator then corresponds to $\hat{\mathbf{C}}_l^w = (2l+1)^{-1} \sum_m \mathbf{P}_{lm}^w \mathbf{P}_{lm}^{w\dagger}$, and the relation to the true spectrum (for statistically isotropic signals) is given by [74]

$$\text{vec}(\hat{\mathbf{C}}_l^w) = \sum_{l'} \mathbf{M}_{ll'}^{22} \text{vec}(\hat{\mathbf{C}}_{l'}), \qquad (39)$$

where the notation $\text{vec}(\mathbf{C}_l)$ refers to the vector of auto- and cross-power spectra. The mode-coupling matrix $\mathbf{M}_{ll'}^{22}$, which can be computed directly from the pixel weights w_i, is generally not invertible due to the loss of information resulting from masking. This issue is resolved by grouping sets of Δl multipoles into bandpowers and inverting the smaller binned form of $\mathbf{M}_{ll'}^{22}$. A bin width of $\Delta l = 10$ is used in the present work.

As a consequence of the CMB E-mode power being significantly larger than that of B-modes, the aforementioned decoupling technique remains suboptimal. Indeed, E-mode leakage increases the variance of $\hat{C}_l^{BB}$, and this additional source of uncertainty needs to be mitigated at the map level [30]. This process, referred to as B-mode purification, is described in Ref. [75] and implemented in the NaMaster package [76]. The Master algorithm [74] also allows to perform beam deconvolution and invert the mode-coupling matrix $\mathbf{M}_{ll'}^{22}$, whose expression for purified pseudo-B-modes is derived in Ref. [76].

After mitigating all mask-related artifacts, we still need to remove significant noise biases from our SAT B-mode autospectra (and, to a lesser extent, from cross-spectra between different frequency channels where noise may be at least partially correlated, for example due to atmospheric and instrumental systematic residuals). Our pipeline is designed to estimate the noise component directly from the data instead of relying on an instrumental model which might not be sufficiently precise. As briefly mentioned in Sec. V A, this is facilitated by the use of data splits, i.e., maps assembled from nonoverlapping observations such that they do not share the same instrumental white noise realization. The final $\hat{C}_l^{BB}$ estimator between frequencies ν and ν' is then obtained by averaging the cross-spectra measured over all pairs of unequal splits s_i, s_j:

$$(\hat{C}_l^{BB})_{\nu\nu'} = \frac{1}{N_{\text{splits}}(N_{\text{splits}} - 1)} \sum_{i \neq j} (\hat{C}_l^{BB})_{\nu\nu'}^{s_i s_j}. \qquad (40)$$

Due to the use of Wiener-filtered E-modes and lensing convergence tracers, noise-dominated multipoles are down-weighted in the power spectra (14) and (15) involving the lensing template, and no significant noise biases are observed for our simulations. Our template is therefore built from full-survey maps instead of being subdivided into data splits, and its cross-spectrum with SAT B-modes at frequency ν is averaged as follows:

$$(\hat{C}_l^{BB_t})_\nu = \frac{1}{N_{\text{splits}}} \sum_i (\hat{C}_l^{BB_t})_\nu^{s_i}. \qquad (41)$$

Template splits may, however, have to be introduced when working with real data, similar to their use in the recent ACT DR6 lensing analysis [77].

Once all cross-spectra are calculated for the 500 fiducial model simulations, they are used to compute the sample covariance matrix $\mathbf{M}_f$ required for the likelihood in Eq. (20). This matrix is conditioned to be block-diagonal, with all correlations between multipoles separated by more than the bin width Δl being set to zero. Such a precaution is necessary in order to remove spurious off-diagonal elements related to Monte Carlo noise, a consequence of estimating the covariance from a finite number of simulations, which has been shown to result in bias and underestimated uncertainties when inferring r [78].

Example power spectra are displayed in Fig. 3 for one specific CMB realization with d0s0 foregrounds, using optimal pixel weights. The 1σ standard deviation on the spectra, represented by the blue shaded areas, is given by the square root of the diagonal covariance elements. In the two left-hand columns, corresponding to the 23 and 39 GHz bands, large noise fluctuations appear at high multipoles. This amplified variance is a consequence of beam deconvolution in the frequency channels with the lowest resolution. The shape of the lensing B-mode spectrum (dotted line) is mostly recognizable at 93 and 145 GHz, while the two highest frequency bands are heavily dominated by dust emission. For $l > 40$, all panels display a good agreement between the measured power spectra and the fiducial model described in Sec. II B 1 (with the best-fit parameters cited at the end of Sec. V A). The steepening of the power spectra on large scales, which is not observed in Ref. [24], is a consequence of the less conservative survey mask used in the present work. With narrower Galactic cuts, the foregrounds contained in our simulations are brighter and more complex, deviating from our simple power-law fiducial model. A curvature term may be added to the exponent in future work in order to obtain a better fit. Alternatively, assumptions regarding the shape of the power spectrum may be removed by using an l-wise parametrization, or by working with empirically determined spectral templates. When working with real data, different combinations of Galactic cuts and model parametrizations will be tested in order to determine the optimal analysis settings. While a more conservative mask may end up being selected at this stage, we intentionally keep the wider survey area here in order to demonstrate the proper functioning of our pipeline in the presence of complex foreground features.

In the subplots in Fig. 3 involving the lensing template, the black continuous line represents the mean of $\hat{C}_l^{B_t B_t}$ over the 500 CMB realizations. This closely matches the mean of the cross-spectra of the template with SAT B-modes
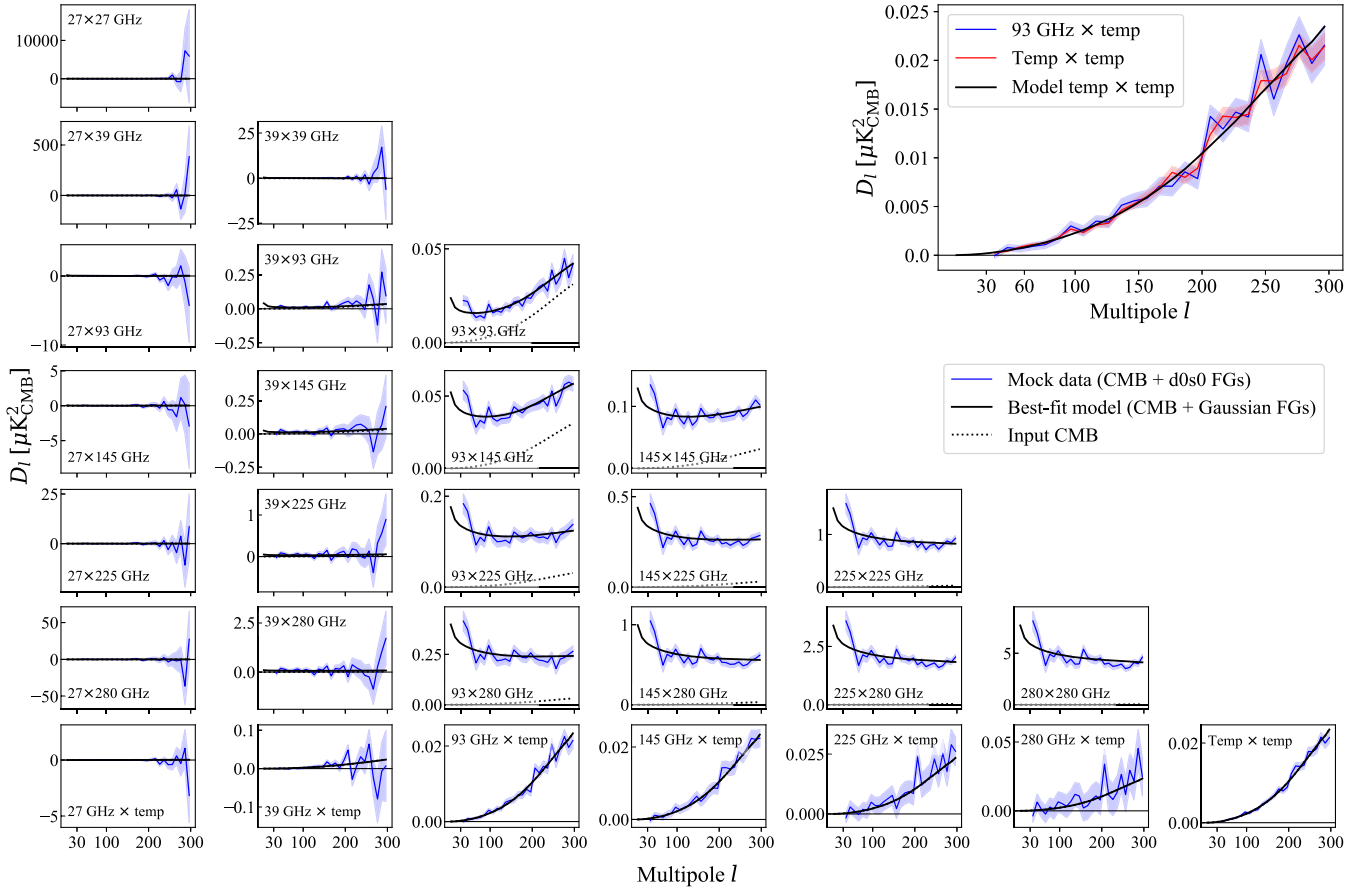
FIG. 3. Power spectra computed for one realization of SAT B-modes in all frequency channels and the corresponding lensing template, with $D_l = l(l+1)C_l/2\pi$. The input maps include the CMB signal, SAT-like noise at goal/optimistic levels and `d0s0` foregrounds. The blue shaded areas represent the 1σ standard deviation calculated over 500 simulations with Gaussian foregrounds. The lensing template autospectrum and its cross-spectrum with SAT B-modes at 93 GHz are highlighted in the top-right corner.

$\hat{C}_l^{BB_t}$, thus explicitly verifying that the equivalence discussed in Sec. III still holds in the presence of inhomogeneous LAT noise. This conclusion is further supported by the highlighted subplot in the top-right corner of Fig. 3, where correlated fluctuations are visible between $\hat{C}_l^{B_t B_t}$ and $\hat{C}_l^{BB_t}$ at 93 GHz. We therefore use the mean of $\hat{C}_l^{B_t B_t}$, computed for the fiducial simulations using the apodized optimal weighting mask, as our model for all spectra involving the lensing template.

In the final stage of the pipeline, the measured cross-spectra, their covariance matrix and our parametric model are substituted into the likelihood function, which can either be the Gaussian likelihood in Eq. (20) or the Hamimeche and Lewis approximation (A6). Note that we only select multipoles between 30 and 300; this range targets the scales at which primordial B-modes are expected to be the most significant while preventing our analysis from being affected by effective large-scale noise induced by timestream filtering. A Markov chain Monte Carlo (MCMC) algorithm implemented in the emcee package is then run for 8000 iterations in order to sample the posterior distributions of r and our foreground model parameters. The number of walkers is set to 24 for `d0s0` input simulations, and increased to 48 for the more complex `d1s1` and `dmsm` models. In order to ensure convergence,

TABLE II. Parameter priors for our MCMC analysis. The numbers listed here correspond to the center value and standard deviation in the Gaussian (G) case, or the lower and upper bounds for top-hat (TH) priors. Note that B_d, B_s, γ_d and γ_s only appear in the model when performing the moment expansion for spatially varying foregrounds.

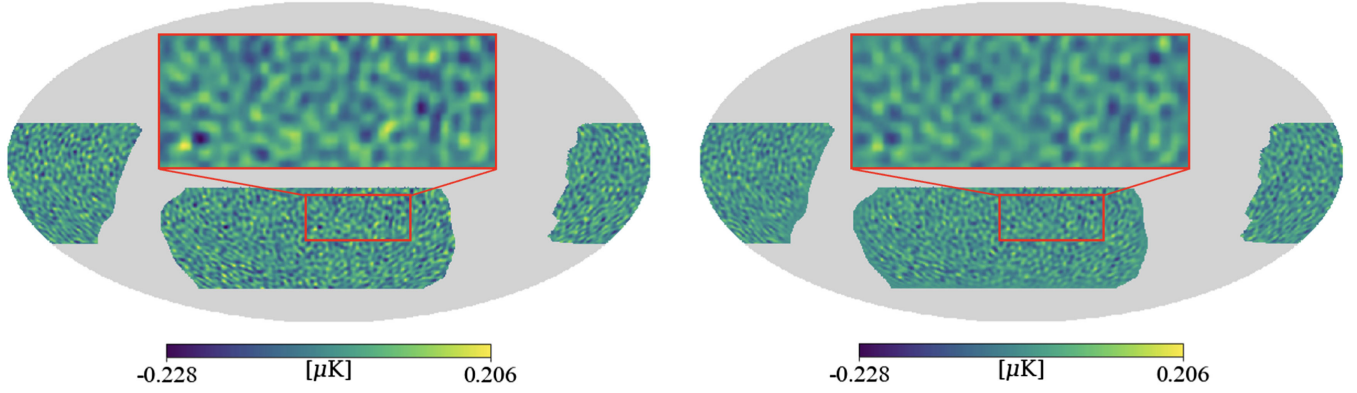| Parameter	r	ϵ_{ds}	β_d	B_d	γ_d	A_d	α_d	β_s	B_s	γ_s	A_s	α_s
Prior	TH	TH	TH	TH	TH	TH	TH	G	TH	TH	TH	TH
Bounds	$[-0.1, 0.1]$	$[-1, 1]$	$[1.3, 1.8]$	$[0, 10]$	$[-6, -2]$	$[20, 60]$	$[-0.5, 0.5]$	-3 ± 0.3	$[-5, 5]$	$[-6, -2]$	$[0, 4]$	$[-3, 0]$

FIG. 4. Left: lensing B-modes in one of the input CMB maps, projected as a scalar field onto the overlap area between the SAT and LAT masks. Right: corresponding lensing B-mode template projected in the same way. In both panels, we only show multipoles between $20 \leq l \leq 128$ and include a zoomed-in region to highlight the visual correlation between the two maps.

we estimate the integrated autocorrelation time (IAT) for r and obtain 81 for the baseline analysis and 191 for the moment expansion method, thus confirming that the length of our chains is sufficient (between 40 and 100 times the IAT). Table II lists the priors used for parameter estimation; most of them follow Ref. [24], with the exception of β_d for which the Gaussian prior is replaced by a top hat. This choice, physically motivated by the β_d map extracted from *Planck* data and shown in Ref. [79], was found to lead to faster convergence in the presence of complex foregrounds. It is worth mentioning that the prior on r will be restricted to positive values when analyzing real data, but is deliberately kept open for now in order to check for potential biases.

VI. RESULTS AND DISCUSSION

A. Lensing template

The lensing B-mode template obtained from one realization of our LAT-like simulations and the associated (correlated) realization of the external LSS tracers is displayed in the right panel of Fig. 4, and compared to the input lensing B-modes in the same CMB realization. Both subplots only contain a restricted set of large-scale multipoles ($20 \leq l \leq 128$) for legibility purposes, and show the region of overlap between the LAT and SAT surveys. The mask is different from the one presented in Ref. [18] as updated versions of the hit-count maps are used here. In particular, Galactic cuts are less conservative in the present work than in Ref. [18] in order to maximize the analysis area.

The correlation between the template and the original B-modes is clearly visible in Fig. 4, especially in the zoomed-in rectangles. A certain degree of attenuation is noticeable when comparing the two panels of the figure. This is a direct consequence of the use of Wiener-filtered E-modes and lensing convergence tracers in Eq. (3); indeed, Eqs. (7)

and (12) imply that any noise in the LAT maps results in a decreased amplitude for $\hat{E}_{lm}^{\mathrm{WF}}$ and $\hat{\kappa}_{LM}^{\mathrm{comb}}$.

The fractional lensing B-mode power residual after map-level delensing with our template is displayed in Fig. 5. Considering the minimum-variance delensed B-modes[6] [18]

$$B_{lm}^{\mathrm{del}} = B_{lm} - \frac{C_l^{BB_t}}{C_l^{B_t B_t}} B_{t,lm}, \tag{42}$$

the average of the ratio $C_l^{BB,\mathrm{del}}/C_l^{BB}$ over l corresponds to the A_{lens} parameter defined in Eq. (32). In practice, we evaluate this quantity by computing the pseudo-C_l of the noise-free input lensing B-modes (C_l^{BB}) and of the lensing template ($C_l^{B_t B_t}$), as well as their cross-spectrum $C_l^{BB_t}$, using the binary SAT/LAT overlap mask with an apodization length of 5 deg.

Figure 5 illustrates the slight dependence of $C_l^{\tilde{B}\tilde{B},\mathrm{del}}/C_l^{\tilde{B}\tilde{B}}$ on angular scale, and is consistent with the results obtained in Ref. [18] for LAT-only E-modes with the previous version of the survey masks. The fraction of lensing B-mode power remaining after delensing averages out at 35%. As shown in Fig. 9 of Ref. [18] for an idealistic foreground-free situation, including SAT E-modes in the template construction stage may allow to bring this ratio further down towards 30%. However, applying this operation to real data would also require a careful treatment of foreground residuals, which we leave for future work.

B. Pixel weighting

We now present a series of tests performed in order to check the proper functioning of our pipeline and validate the optimal pixel weights discussed in Sec. IV. The `d0s0` foreground model is used in our input maps throughout this

---

[6]We give the general form here, involving the Wiener-filtered template, although in our analysis the filter is very close to unity since $C_l^{BB_t} \approx C_l^{B_t B_t}$.
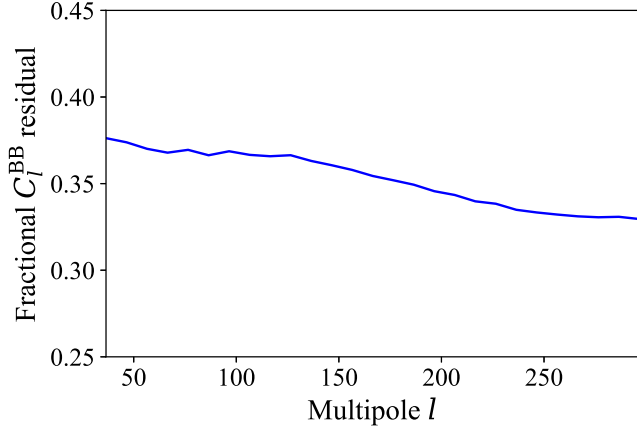
FIG. 5. Ratio between the residual and input lensing B-mode power, obtained as in Eq. (32) but without averaging over l, in the overlap area between the SAT and the LAT surveys. Using LAT-only E-modes combined with a multitracer κ estimator and considering the goal LAT noise level, the delensing efficiency of our template reaches approximately 65%.

section, as it is simple enough to incur low computational costs but realistic enough to provide informative results. We use 500 simulations containing the CMB, SAT-like noise and Gaussian foregrounds to evaluate the covariance matrix in each variant of the analysis.

Our aim is to verify that the pixel weights in Eq. (25) do lead to tighter cosmological constraints than the uniform or inverse-noise-variance ($\propto N_i^{\text{hit}}$) weights used in Refs. [18] and [24], respectively. This is done by analyzing the same input map with different weighting schemes and recomputing all cross-spectra as well as the corresponding covariance matrix each time. The apodization length is set to 10 deg for the optimized and hit-count weights; we increase it to 25 deg for the uniform weights which do not otherwise fall off smoothly near the edge of the survey area. We use

$A_{\text{lens}} = 1$ to compute the optimal weights without delensing, and change it to 0.35 when delensing is performed.

The marginal posterior distributions for r are obtained in each case for one realization and are displayed in Fig. 6, confirming that our optimal weights (solid lines) indeed perform better than the inverse-noise and uniform cases. The uniform weighting is clearly suboptimal, yielding the MCMC standard deviation $\sigma(r) = 2.9 \times 10^{-3}$ before delensing and 2.3×10^{-3} after. The larger statistical errors and the modest improvement from delensing (only 21%) are a consequence of the high effective noise level in the outer regions of this mask, which leads N_l to dominate in Eq. (6). On the other hand, the distributions obtained with the hit-count mask and the optimal pixel weights are very similar. Their respective standard deviations are $\sigma(r) = 2.1 \times 10^{-3}$ and 1.9×10^{-3} with the baseline analysis; once delensing is applied, both values decrease to about 1.2×10^{-3}. At SO's nominal sensitivity, the difference between the two methods is therefore not very significant. As expected, it is more apparent without delensing and practically invisible in the $A_{\text{lens}} = 0.35$ case, where instrumental noise still dominates over lensing residuals. However, our new weighting scheme will become increasingly relevant for future experiments with lower noise levels [for which the lensing B-mode variance in Eq. (25) will be comparatively more important], and in particular for LiteBIRD where the residual lensing B-modes will vary across the sky [45].

The derivation presented in Sec. IV assumes instrumental noise to be well described by a white power spectrum. To remove this assumption, we also investigated an alternative technique inspired by the hybrid C_l estimator described in Ref. [80], which combines power spectra obtained with uniform and hit-count weights. Due to issues related to its very high computational costs, this procedure is not yet suitable for use in SO data analysis. A more detailed discussion of the hybrid-weighting method and of the
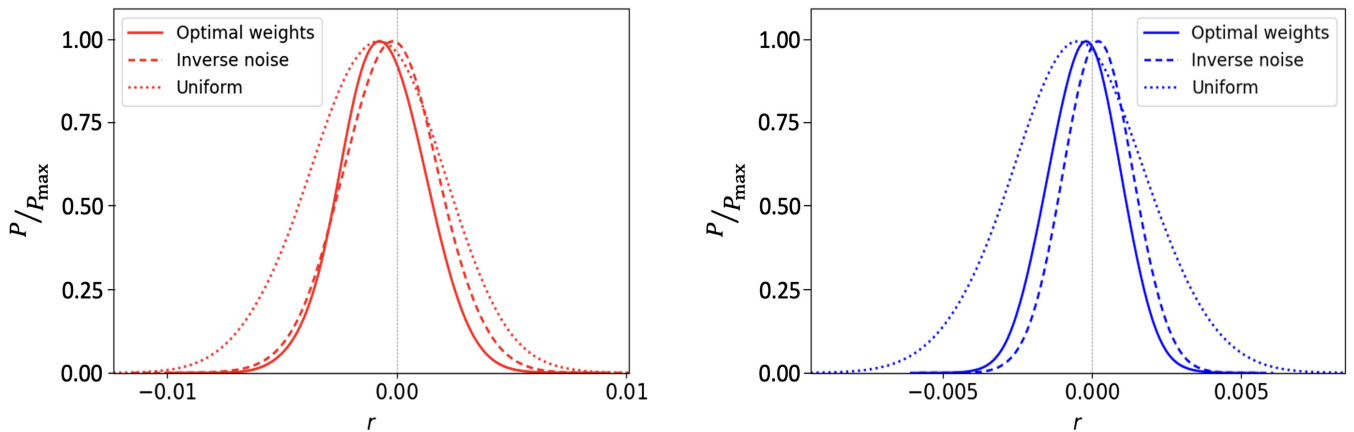

FIG. 6. Marginalized posterior distributions for the tensor-to-scalar ratio r obtained by sampling the Gaussian likelihood (20) for a single input realization. The input simulation has $r = 0$ and d0s0 foregrounds, and cross-spectra are computed using uniform (dotted lines), hit-count (dashed lines) and optimal (solid lines) pixel weights, respectively. Left: no delensing is applied. Right: cross-spectral delensing is performed using the template characterized in Sec. VI A.
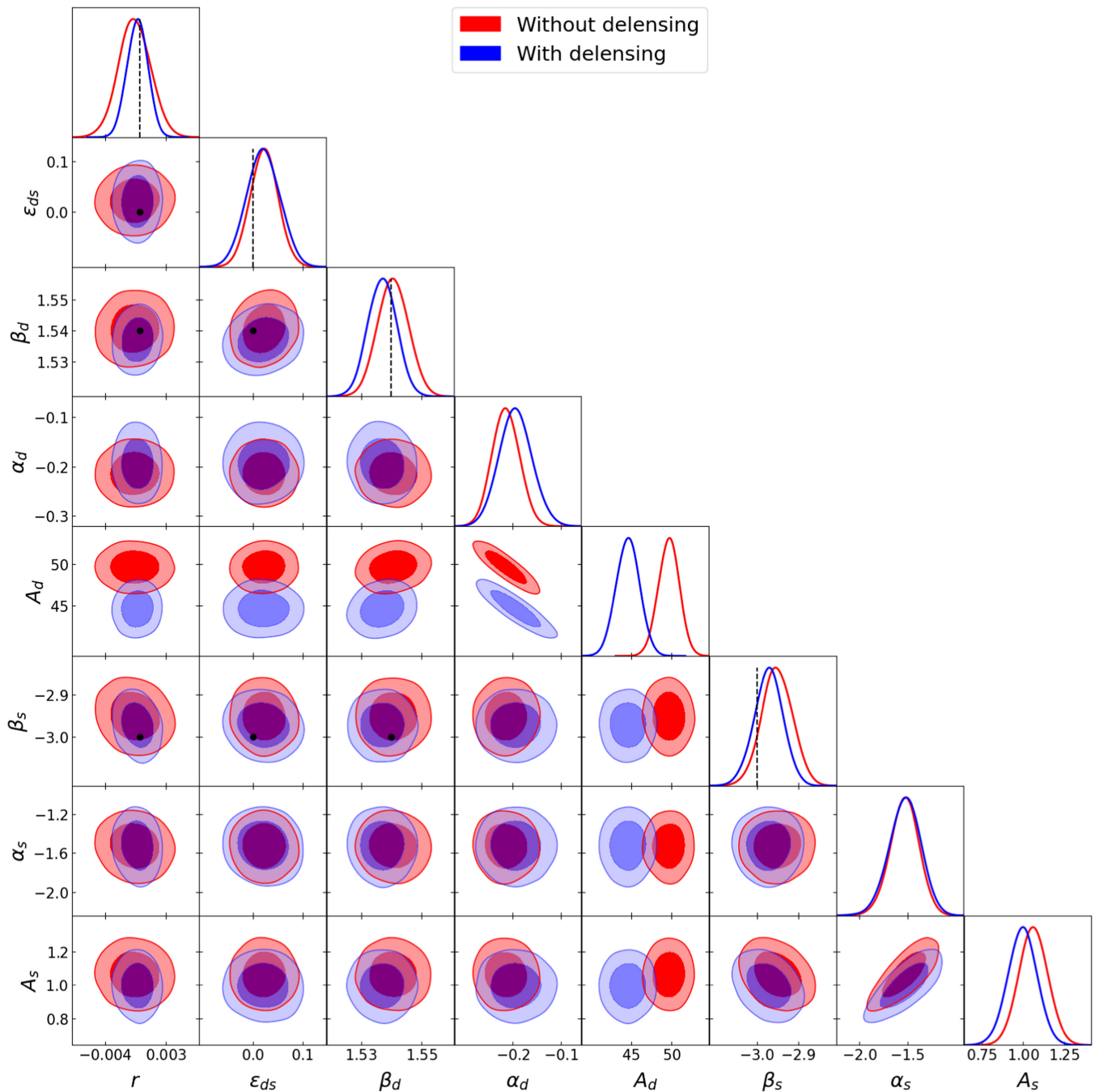
FIG. 7. Triangle plot displaying the posterior distributions obtained for all model parameters with (blue) and without (red) delensing, using one input map realization with d0s0 foregrounds. The moment expansion is not performed here. For r, ϵ_{ds}, β_d and β_s, the dashed lines and dots represent the true values used to generate our simulations. The inferred parameters for the foreground angular power spectra depend on pixel weights (as visualized by the shift of the A_d posterior); therefore we do not plot ground truths for A_d, α_d, A_s and α_s.

problems arising when testing it on SO-like simulations can be found in Appendix B. The optimized weights in Eq. (25) are systematically used for all results that follow in the main text.

All final outputs of our analysis for a given input simulation are shown in Fig. 7, where posterior distributions for r and the foreground parameters can be seen with and without delensing. As the d0s0 foreground map used here does not include SED spatial variability, we consider the simplest form of our parametric model and do not perform the moment expansion. The r distribution visibly tightens as a result of delensing, with $\sigma(r)$ dropping from 1.9×10^{-3} to 1.2×10^{-3} (37% improvement).

TABLE III.   Tensor-to-scalar ratio statistics averaged over 100 simulations for three increasingly complex foreground models and $r = 0$. The moment expansion of the SEDs is performed for `d1s1` and `dmsm`. The quantities $\hat{r}$ and $\sigma(r)$ refer to the maximum of the r posterior distribution and the SD of the corresponding MCMC samples, respectively. Values in the last column are obtained by decreasing the lensing B-mode power in the input simulations and running the original pipeline (without cross-spectral delensing).

	Without delensing			With delensing			$A_{\mathrm{lens}} = 0.35$
FG model	Mean of $\hat{r}$ ($\times 10^3$)	Mean of $\sigma(r)$ ($\times 10^3$)	SD of $\hat{r}$ ($\times 10^3$)	Mean of $\hat{r}$ ($\times 10^3$)	Mean of $\sigma(r)$ ($\times 10^3$)	SD of $\hat{r}$ ($\times 10^3$)	$\sigma(r)$ ($\times 10^3$)
`d0s0`	0.30	1.9	1.7	0.20	1.2	1.1	1.2
`d1s1`	−0.35	3.2	2.6	−0.53	2.2	2.1	2.4
`dmsm`	−0.73	3.0	2.5	−0.34	2.2	2.0	2.3

The only foreground parameter whose posterior significantly changes after delensing is the dust amplitude A_d; this is easily explained by the dependence of the weights in Eq. (25) on A_{lens}. Indeed, without delensing, the optimized mask is slightly closer to uniform than with $A_{\mathrm{lens}} = 0.35$. Pixels near the Galactic plane are then weighted more heavily, resulting in a higher effective dust amplitude. The `d0s0` best-fit estimates are respectively $A_d = 47.2$ without delensing and 42.1 with, averaging to 44.6. The observed shift of the A_d maximum posterior by about 5 units is therefore consistent with our expectations. The synchrotron amplitude and the power-law exponents exhibit less significant shifts; in all cases, priors are chosen wide enough to hold for both values of A_{lens}.

Note that the α_s posterior in Fig. 7 peaks significantly lower than the reference value mentioned in Sec. V A ($\alpha_s = -0.78$); this is due to the fact that the synchrotron angular power spectrum is not a perfect power law and has a steeper slope on large scales. As the noise level is strongly amplified by beam deconvolution above $l \approx 200$ at 27 and 39 GHz, where synchrotron emission dominates, the SATs are mostly sensitive to α_s at low multipoles. This results in our pipeline inferring a steeper power-law index than the best-fit value obtained from the `s0`-only map over the full $30 \leq l \leq 300$ range (which is how the fiducial value is obtained). Figure 7 does not indicate any significant degeneracy between r and α_s; our cosmological constraints are therefore unlikely to be affected; however a more complex synchrotron model including a curved power-law index may be used in future work.

C. Delensing performance with realistic foregrounds

Having demonstrated the pipeline on the simplest of our sky models, we now assess the impact of delensing in the presence of more complex foregrounds as well as nonzero tensor modes. In each case considered, we also check for biases and verify the robustness of our statistics by computing the maximum-posterior estimate $\hat{r}$ and the width of the distribution $\sigma(r)$ for 100 different input maps. These simulations correspond to random realizations of the CMB and noise, with a fixed PySM foreground template. For `d1s1` and `dmsm`, we perform the moment expansion

described in Sec. II B 2 in order to account for the spatially varying SEDs. Following Ref. [24], we use the same fiducial covariance matrix when changing the foreground model; however it must be modified to account for the variance of tensor modes in the $r = 0.01$ case.

As a first sanity check, we verify (using our $r = 0$ simulations) that the standard deviation (SD) of the 100 best-fit estimates $\hat{r}$ is broadly consistent with the average $\sigma(r)$ obtained from the SD of the MCMC samples for each realization (see Table III). Note that the 100 simulations used as mock data contain the same foreground map; we therefore expect the SD estimated from this set of inputs to be lower than the MCMC $\sigma(r)$ value determined by the covariance matrix, which accounts for the effects of foreground variance. For this reason, we choose the latter as the preferred statistic (it is also the only one we will have access to when working with real data).

We then check that $\sigma(r)$ does not scatter significantly; its SD over the 100 simulations is of the order of 10^{-5} for `d0s0` and 10^{-4} for `d1s1` and `dmsm`. This result is expected, as error bars are mostly determined by the covariance matrix (which remains unchanged) and should not strongly depend on the input realization. Finally, we compare the SD of the mean of the posterior r_{mean} to that of $\hat{r}$. With a difference of less than 2.5% in all foreground cases, the two values are in good agreement both with and without delensing, as expected for Gaussian posteriors.

The mean $\hat{r}$ and $\sigma(r)$ values quoted in Table III are summarized in Fig. 8, which also displays results for input maps with $r = 0.01$. We do not observe any significant biases compared to the statistical error for a single realization. The small fluctuations in $\hat{r}$ appearing with the spatially varying foregrounds are related to volume effects in the distributions of the moment expansion parameters, and will be mitigated by using different priors in future work (see Sec. VII).

The key point of Fig. 8 is the comparison between the results obtained with and without delensing. For all foreground models and input r values, the average of $\hat{r}$ remains consistent in both cases, thus explicitly verifying that our implementation of delensing does not introduce any additional bias into the parameter estimates. This does not
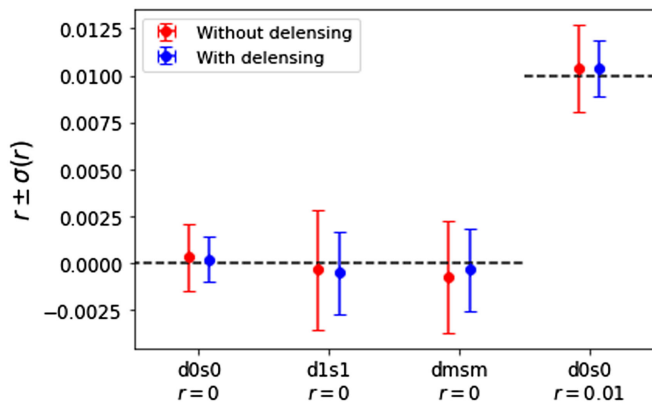
FIG. 8. Comparison of the mean of $\hat{r}$ and the MCMC standard deviation $\sigma(r)$ (represented by the error bars) with and without delensing. We average over 100 simulations for each case. The moment-expansion method is used for `d1s1` and `dmsm` foregrounds, resulting in larger error bars. The presence of tensor modes is successfully detected by the pipeline and also leads to increased uncertainties. Delensing causes the constraints to tighten by 27% to 37% depending on the characteristics of the input maps.

exclude small shifts in $\hat{r}$ occurring for any single realization, as expected given the shrink in the width of the posterior for r. Figure 8 also illustrates the clear error reduction after delensing, with the mean of $\sigma(r)$ decreasing by 37%, 31% and 27% for `d0s0`, `d1s1` and `dmsm`, respectively, in the absence of primordial B-modes. In the latter two cases, the four additional sampled parameters result in wider distributions and a more modest delensing-related improvement. Despite the degradation of uncertainties associated with the use of the moment-expansion method, SO's goal precision of $\sigma(r) = 0.003$ is achieved even with the most complex of our foreground models. Our results without delensing are broadly consistent with Ref. [24]; the slightly lower $\sigma(r)$ we obtain with the `d0s0` model can be attributed to the use of optimized pixel weights, while the slightly higher uncertainties and biases with spatially varying foregrounds are probably related to the wider survey mask. As expected, delensing improvements observed in the present work exceed the $A_{\mathrm{lens}} = 0.5$ forecasts in Ref. [24], due to our templates achieving a delensing efficiency of 65%.

With $r = 0.01$ and `d0s0` foregrounds, the pipeline successfully detects the primordial signal, and the additional cosmic variance due to the presence of tensor modes leads to larger error bars. Without delensing, $\sigma(r)$ grows by 16% compared to the $r = 0$ case, reaching a value of 2.2×10^{-3}. This increase is consistent with the results of Ref. [24]. After delensing, we obtain $\sigma(r) = 1.5 \times 10^{-3}$, corresponding to a 32% improvement. Note that we do not modify the pixel weights compared to the $r = 0$ case [i.e., we do not add the primordial B-mode variance to the denominator in Eq. (25)], as the value of r will not be

known when analyzing real data and the $r = 0$ weights will be used by default.

The last column of Table III contains $\sigma(r)$ values obtained after applying the foreground-cleaning pipeline (without delensing) to $r = 0$ input maps in which the lensing B-mode power was artificially reduced by 65%. This procedure approximately simulates the effect of delensing at the map level. The good agreement observed between these results and our average uncertainties after cross-spectral delensing is consistent with the theoretical development in Sec. III and attests to the robustness of our forecasts.

VII. CONCLUSION

In this work, we implemented delensing in SO's power-spectrum-based foreground-cleaning pipeline and obtained the first performance forecasts for this technique on realistic SO-like maps including Galactic foregrounds and inhomogeneous noise. Our lensing templates, built by optimally coadding quadratic convergence estimators with Gaussian external LSS tracer simulations, allowed to remove 65% of the lensing B-mode power. By treating the template as an additional input channel and including its autospectrum as well as all its cross-spectra with the SAT-like maps in the likelihood function, we observed no significant bias and a clear tightening of the posterior distributions for the tensor-to-scalar ratio r. An inverse-variance weighting accounting for the lensing B-mode power as well as for the noise level in each pixel was applied when computing power spectra, and was shown to outperform the uniform- and hit-count-weighting schemes used previously.

Our parametric foreground model was adjusted to account for the different levels of complexity present in the input simulations. With $r = 0$ and uniform dust and synchrotron SEDs, errors on r decreased by 37% after delensing. When introducing SED spatial variability, the additional parameters required by the moment-expansion method resulted in wider posteriors. Delensing then yielded a 27% improvement in $\sigma(r)$ and a final value of 2.2×10^{-3} for our most complex model, well within SO's target sensitivity. Tensor modes at the $r = 0.01$ level were successfully detected by the pipeline and led to a 16% increase in $\sigma(r)$ due to the additional sample variance. We finally validated our results by comparing them to the uncertainties obtained with artificially delensed input maps, thus verifying the equivalence between map-based and cross-spectral delensing.

The most exciting future prospect for this work is of course the application to early SO data as it becomes available within the next year. The main challenge at this stage will be related to timestream filtering of the SAT data, which we will need to account for at the power spectrum level through the use of transfer functions. For the LAT, maximum-likelihood maps will be obtained with a

weighting scheme proportional to the Fourier-diagonal inverse detector-detector covariance matrix. This process will also have to be considered in the first step of the pipeline when working with real LAT data. However, as the LAT's first light is not expected to happen before 2025, early demonstrations of delensing will likely be carried out using *B*-mode templates built from external mass tracers and preexisting ACT or *Planck* CMB maps. Finally, another aspect we will need to be mindful of when analyzing real data is the choice of cosmology used to generate our theoretical model. An additional free parameter A_L modulating the lensing *B*-mode power spectrum amplitude may be introduced in order to avoid small biases on *r* related to uncertainties on cosmological parameters. At SO's sensitivity level, the signal-to-noise ratio on the CMB lensing power spectrum is expected to be greater than 100 [9] and such uncertainties should therefore be subpercent. This could be used as a prior on A_L and any inflation of the errors on *r* should be negligible as a result.

Other future extensions of this paper will include the use of more realistic LAT-like simulations containing Galactic and extragalactic foregrounds, which have been shown to impact the delensed *B*-mode power spectrum and lead to biases on *r*, both when using internal [39,81] and external [66] tracers of the lensing convergence. By performing map-based foreground cleaning on the LAT-like mock data and quantifying the remaining biases, we will check that these effects can be successfully mitigated as described in Refs. [39,66,81]. Furthermore, increasing the realism of our analysis will require including more complexity in our external mass tracers, which are currently only generated as Gaussian realizations of theoretical power spectra assuming linear bias.

We may also attempt to combine LAT and SAT *E*-modes when constructing the lensing template, which was shown to improve delensing efficiency by more than 5% in an idealistic foreground-free situation [18]. With realistic input maps, this will require a careful investigation of SAT foreground residuals in the cross-spectra between the lensing template and the observed *B*-modes. Regarding our treatment of variable foreground SEDs, one possible improvement would be the implementation of Jeffreys' priors for the moment parameters in order to mitigate volume effects in their posterior distributions. It would also be interesting to implement delensing in the hybrid component separation pipeline described in Ref. [82]. This technique, which starts by removing foregrounds at the map level assuming spatially uniform SEDs and subsequently models the power spectra of the residuals, was found to lead to tighter constraints than the moment expansion used here; upcoming efforts will be dedicated to determining whether this conclusion remains true with delensing.

Finally, it is important to note that all results presented here assume the nominal SO design with three SATs;

however, three additional SATs are planned, which will lead to lower noise levels in the sky maps and greater delensing-related improvements on $\sigma(r)$ as indicated by Eq. (6). Furthermore, efficient delensing will be essential for upcoming projects such as CMB-S4 [12], which requires removing more than 90% of the lensing *B*-mode power to achieve its target constraints on the tensor-to-scalar ratio. This will only be possible with dedicated high-sensitivity, high-resolution polarization measurements and more optimal, likelihood-based internal-reconstruction techniques [57]. Delensing will thus be crucial in reaching the full potential of these extremely sensitive future experiments and hopefully unveiling the elusive physics of the early Universe.

APPENDIX A: HAMIMECHE AND LEWIS LIKELIHOOD

In this Appendix we introduce an alternative likelihood to the simple Gaussian approximation used in the main text, following Ref. [55]. This improved approximation aims to capture the non-Gaussian dependence of the likelihood on the model power spectra. We briefly review the construction of the likelihood and present the results from applying it to one of our simulations with d0s0 foregrounds.

The Hamimeche and Lewis likelihood approximation takes as its starting point the exact likelihood for an ideal, full-sky survey of several correlated Gaussian fields. Considering n fields $\boldsymbol{a}_{lm} = (a_{lm}^1, …, a_{lm}^n)^T$ (typically seven, with the B-modes from each SAT channel and the lensing template), the log-likelihood is given by the matrix generalization of Eq. (5):

$$-\ln P(\{\boldsymbol{a}_{lm}\}|\mathbf{C}_l) = \frac{2l+1}{2}[\text{Tr}(\hat{\mathbf{C}}_l\mathbf{C}_l^{-1}) + \ln|\mathbf{C}_l|] \qquad (A1)$$

up to a constant, where $\mathbf{C}_l$ is the model covariance depending on a set of parameters and $\hat{\mathbf{C}}_l = (2l+1)^{-1}\sum_{m=-l}^{l}\boldsymbol{a}_{lm}\boldsymbol{a}_{lm}^\dagger$. Note that $\boldsymbol{a}_{lm}$ contains noise contributions that must be included in the model. Equation (A1) is the log-likelihood at a given multipole l; the full log-likelihood is obtained by summing over l.

In the ideal full-sky likelihood, Eq. (A1), the fields only enter through their power spectra $\hat{\mathbf{C}}_l$. Integrating over the $\boldsymbol{a}_{lm}$ at fixed $\hat{\mathbf{C}}_l$, gives the sampling distribution $P(\hat{\mathbf{C}}_l|\mathbf{C}_l)$ for the measured power spectra. The likelihood function for the $\mathbf{C}_l$ given the measured $\hat{\mathbf{C}}_l$ is $\mathcal{L}(\mathbf{C}_l|\hat{\mathbf{C}}_l) = P(\hat{\mathbf{C}}_l|\mathbf{C}_l)$.

Normalizing the log-likelihood to zero at $\mathbf{C}_l = \hat{\mathbf{C}}_l$, we have [55]

$$-\ln\mathcal{L}(\mathbf{C}_l|\hat{\mathbf{C}}_l) = \frac{2l+1}{2}[\text{Tr}(\hat{\mathbf{C}}_l\mathbf{C}_l^{-1}) - \ln|\hat{\mathbf{C}}_l\mathbf{C}_l^{-1}| - n]. \qquad (A2)$$

Introducing a fiducial model with cross-spectra $\mathbf{C}_{fl}$, Eq. (A2) can be rewritten as

$$-\ln\mathcal{L}(\mathbf{C}_l|\hat{\mathbf{C}}_l) = \frac{2l+1}{4}\text{Tr}[(\mathbf{C}_{fl}^{-1/2}\mathbf{C}_{gl}\mathbf{C}_{fl}^{-1/2})^2] \qquad (A3)$$

with $\mathbf{C}_{gl} = \mathbf{C}_{fl}^{1/2}\mathbf{g}(\mathbf{C}_l^{-1/2}\hat{\mathbf{C}}_l\mathbf{C}_l^{-1/2})\mathbf{C}_{fl}^{1/2}$. Here, the matrix function $\mathbf{g}$ acts on symmetric, positive-definite matrices by application of the function

$$g(x) = \text{sign}(x-1)\sqrt{2(x-\ln x - 1)} \qquad (A4)$$

to its eigenvalues. We can now rearrange the upper triangular portion of $\mathbf{C}_{gl}$ into a vector X_{gl} in order to express Eq. (A3) as a quadratic form

$$-2\ln\mathcal{L}(\mathbf{C}_l|\hat{\mathbf{C}}_l) = X_{gl}^T\mathbf{M}_{fl}^{-1}X_{gl}, \qquad (A5)$$

where $\mathbf{M}_{fl}$ is the covariance of the measured power spectra in the fiducial model.

The exact full-sky likelihood function corresponds to the sum of Eq. (A5) over all considered multipoles. While exact on the full sky, it can also be used as an approximate likelihood for an anisotropic survey substituting for mask-deconvolved measured power spectra in the construction of X_{gl} and their covariance for $\mathbf{M}_{fl}$. Mask-induced couplings between multipoles can be accounted for in the covariance,
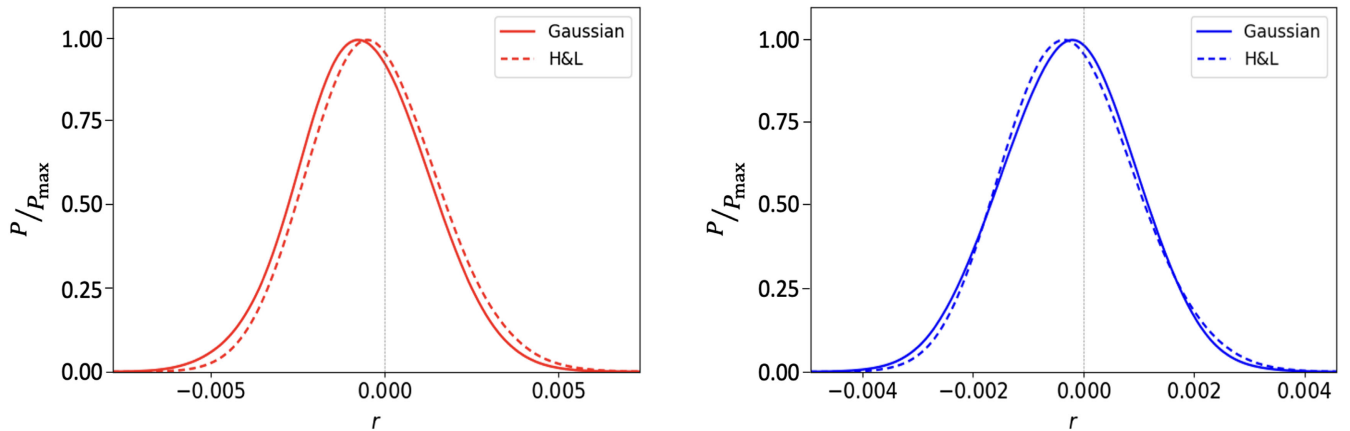

FIG. 9.   Marginalized posterior distributions for the tensor-to-scalar ratio r, obtained by sampling the full Hamimeche and Lewis likelihood (A6) (dashed lines) and its Gaussian approximation (20) (solid lines) for a single input realization. The input simulation has $r = 0$ and d0s0 foregrounds, and cross-spectra are computed using the optimal pixel weights. Left: no delensing is applied. Right: cross-spectral delensing is performed using the template characterized in Sec. VI A.

leading to the following final result (known as the Hamimeche and Lewis likelihood):

$$-2 \ln \mathcal{L}(\{\mathbf{C}_l\}|\{\hat{\mathbf{C}}_l\}) \approx \sum_{ll'} [X_g]_l^T [\mathbf{M}_f^{-1}]_{ll'} [X_g]_{l'}. \quad (A6)$$

In Fig. 9, we compare Eq. (A6) to its Gaussian approximation (20) for one realization of our input maps with $r = 0$ and d0s0 foregrounds. All cross-spectra are computed with the optimal pixel weights in Eq. (25), and 500 SAT-like simulations with Gaussian foregrounds are used to estimate the covariance matrix. The marginalized posterior distributions obtained for r by sampling these two likelihoods are in good agreement both with and without delensing.

APPENDIX B: HYBRID WEIGHTING METHOD

As an alternative to the pixel weights discussed in Sec. IV, we investigated a hybrid masking method using both uniform and inverse-noise-variance weighting schemes. To do so, we generated two distinct sets of input maps containing the same realizations of the sky signals (CMB plus Galactic foregrounds) and noise, and then applied the uniform mask to one set and the hit-count mask to the other. The same treatment was applied to the lensing B-mode template, and the cross-spectra between all pairs of maps (including pairs with different masks) were incorporated in the likelihood. With six frequency bands and the lensing template, the number of coadded spectra therefore increased to 105, compared to the 28 shown in Fig. 3.

In Ref. [80], this approach was used to combine pseudo-C_l estimators computed with uniform and inverse-noise-variance weighting into a single hybrid estimator. The advantage of this approach is that it effectively selects different weights on different scales, where isotropic signal or anisotropic noise may dominate. Applying this technique to our analysis would allow us to drop the assumption of white noise, on which the argument in Sec. IV relies, and potentially lead to better results on r.

However, issues related to the increased number of cross-spectra appeared in practice: with a covariance matrix obtained from 500 simulations, parameter distributions for an input map containing d0s0 foregrounds were unrealistically tight, indicating an inaccurate estimation of the covariance. This problem was shown to stem from the insufficient number of simulations compared to the size of the data vector by performing a simple test with foreground-free input maps (containing only the CMB signal and noise). Restricting the analysis to the three central frequency channels, we obtained a reasonable SD of $\sigma(r) = 0.0013$ without delensing (21 cross-spectra) for the hybrid method; using the optimal weights of Sec. IV led to a nearly identical result. With the hybrid method, including all six channels (78 cross-spectra) nearly divided this value by 2, even though the low- and ultrahigh-frequency bands are not expected to contribute significant information in the absence of foregrounds. To obtain accurate error bars with so many cross-spectra, we would need to increase the number of simulations to about 2000 for the covariance estimation. As the results with both weighting methods were equivalent in the simple three-channel test, we concluded that there is little to be gained from the hybrid method for SO over the optimal weights. Given the prohibitive computational cost of accurate covariance estimation for the hybrid method, we decided not to pursue it further. It might, however, be worth revisiting this technique if a future experiment exhibits a clearly nonwhite noise power spectrum.

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

[2] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Statistics of cosmic microwave background polarization, Phys. Rev. D **55**, 7368 (1997).

[3] U. Seljak and M. Zaldarriaga, Signature of gravity waves in the polarization of the microwave background, Phys. Rev. Lett. **78**, 2054 (1997).

[4] BICEP/Keck Collaboration, BICEP/Keck XIII: Improved constraints on primordial gravitational waves using Planck, WMAP, and BICEP/Keck observations through the 2018 observing season, Phys. Rev. Lett. **127**, 151301 (2021).

[5] M. Tristram *et al.*, Improved limits on the tensor-to-scalar ratio using BICEP and Planck, Phys. Rev. D **105**, 083524 (2022).

[6] J. Martin, C. Ringeval, and V. Vennin, Encyclopaedia inflationaris, Phys. Dark Universe **5–6**, 75 (2014).

[7] A. Ijjas and P. J. Steinhardt, Bouncing cosmology made simple, Classical Quantum Gravity **35**, 135004 (2018).

[8] L. Boyle, K. Finn, and N. Turok, *CPT*-symmetric universe, Phys. Rev. Lett. **121**, 251301 (2018).

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

[10] The Simons Observatory Collaboration, The Simons Observatory: Astro2020 decadal project whitepaper, Bull. Am. Astron. Soc. **51,** 147 (2019).

[11] A. Linde, Single-field α-attractors, J. Cosmol. Astropart. Phys. 05 (2015) 003.

[12] K. Abazajian *et al.*, CMB-S4: Forecasting constraints on primordial gravitational waves, Astrophys. J. **926,** 54 (2022).

[13] LiteBIRD Collaboration, Probing cosmic inflation with the LiteBIRD cosmic microwave background polarization survey, Prog. Theor. Exp. Phys. **2023,** 042F01 (2023).

[14] A. A. Starobinskii, Spectrum of relict gravitational radiation and the early state of the universe, JETP Lett. **30,** 682 (1979).

[15] F. L. Bezrukov and M. E. Shaposhnikov, The standard model Higgs boson as the inflaton, Phys. Lett. B **659,** 703 (2008).

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

[17] D. Hanson *et al.*, Detection of B-mode polarization in the cosmic microwave background with data from the south pole telescope, Phys. Rev. Lett. **111,** 141301 (2013).

[18] T. Namikawa *et al.*, The Simons Observatory: Constraining inflationary gravitational waves with multi-tracer B-mode delensing, Phys. Rev. D **105,** 023511 (2022).

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

[20] S. K. Choi and L. A. Page, Polarized Galactic synchrotron and dust emission and their correlation, J. Cosmol. Astropart. Phys. 12 (2015) 020.

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

[22] J. Delabrouille and J.-F. Cardoso, Diffuse source separation in CMB observations, arXiv:astro-ph/0702198.

[23] S. M. Leach *et al.*, Component separation methods for the Planck mission, Astron. Astrophys. **491,** 597 (2008).

[24] K. Wolz *et al.*, The Simons Observatory: Pipeline comparison and validation for large-scale B-modes, Astron. Astrophys. **686,** A16 (2024).

[25] M. Tegmark, Removing real-world foregrounds from CMB maps, Astrophys. J. **502,** 1 (1998).

[26] J. Chluba, J. C. Hill, and M. H. Abitbol, Rethinking CMB foregrounds: Systematic extension of foreground parameterizations, Mon. Not. R. Astron. Soc. **472,** 1195 (2017).

[27] L. Vacher, J. Chluba, J. Aumont, A. Rotti, and L. Montier, High precision modeling of polarized signals: Moment expansion method generalized to spin-2 fields, Astron. Astrophys. **669,** A5 (2023).

[28] S. Azzoni, M. H. Abitbol, D. Alonso, A. Gough, N. Katayama, and T. Matsumura, A minimal power-spectrum-based moment expansion for CMB B-mode searches, J. Cosmol. Astropart. Phys. 05 (2021) 047.

[29] A. Challinor and G. Chon, Geometry of weak lensing of CMB polarization, Phys. Rev. D **66,** 127301 (2002).

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

[31] H. K. Eriksen, I. J. O'Dwyer, J. B. Jewell, B. D. Wandelt, D. L. Larson, K. M. Gorski, S. Levin, A. J. Banday, and P. B. Lilje, Power spectrum estimation from high-resolution maps by Gibbs sampling, Astrophys. J. Suppl. Ser. **155,** 227 (2004).

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

[33] T. Namikawa and R. Nagata, Lensing reconstruction from a patchwork of polarization maps, J. Cosmol. Astropart. Phys. 09 (2014) 009.

[34] T. Okamoto and W. Hu, CMB Lensing Reconstruction on the full sky, Phys. Rev. D **67,** 083002 (2003).

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

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

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

[38] 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).

[39] A. Baleato Lizancos and S. Ferraro, The impact of extragalactic foregrounds on internal delensing of CMB B-mode polarization, Phys. Rev. D **106,** 063534 (2022).

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

[41] A. Dey *et al.*, Overview of the DESI legacy imaging surveys, Astron. J. **157,** 168 (2019).

[42] E. F. Schlafly, A. M. Meisner, and G. M. Green, The unWISE catalog: Two billion infrared sources from five years of WISE imaging, Astrophys. J. Suppl. Ser. **240,** 30 (2019).

[43] Z. Ivezić *et al.*, LSST: From science drivers to reference design and anticipated data products, Astrophys. J. **873,** 111 (2018).

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

[45] T. Namikawa *et al.*, LiteBIRD science goals and forecasts: Improving sensitivity to inflationary gravitational waves with multitracer delensing, J. Cosmol. Astropart. Phys. 06 (2024) 010.

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

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

[48] G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics* (Wiley-VCH, Hoboken, New Jersey, 1986).

[49] A. Lazarian, B.-G. Andersson, and T. Hoang, Grain alignment: Role of radiative torques and paramagnetic relaxation, in *Polarimetry of Stars and Planetary Systems*, edited by L. Kolokolova, J. Hough, and A.-C. Levasseur-Regourd (Cambridge University Press, Cambridge, England, 2015), pp. 81–113.

[50] 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).

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

[52] Planck Collaboration, Planck 2018 results. XI. Polarized dust foregrounds, Astron. Astrophys. **641**, A11 (2020).

[53] U. Fuskeland *et al.*, Constraints on the spectral index of polarized synchrotron emission from WMAP and Faraday-corrected S-PASS data, Astron. Astrophys. **646**, A69 (2021).

[54] S. De, L. Pogosian, and T. Vachaspati, CMB Faraday rotation as seen through the Milky Way, Phys. Rev. D **88**, 063527 (2013).

[55] S. Hamimeche and A. Lewis, Likelihood analysis of CMB temperature and polarization power spectra, Phys. Rev. D **77**, 103013 (2008).

[56] L. Knox, Determination of inflationary observables by cosmic microwave background anisotropy experiments, Phys. Rev. D **52**, 4307 (1995).

[57] S. Belkner, J. Carron, L. Legrand, C. Umiltà, C. Pryke, and C. Bischoff, CMB-S4: Iterative internal delensing and r constraints, Astrophys. J. **964**, 148 (2024).

[58] J. Errard, S. M. Feeney, H. V. Peiris, and A. H. Jaffe, Robust forecasts on fundamental physics from the foreground-obscured, gravitationally-lensed CMB polarization, J. Cosmol. Astropart. Phys. 03 (2016) 052.

[59] R. Stompor, S. Leach, F. Stivoli, and C. Baccigalupi, Maximum likelihood algorithm for parametric component separation in CMB experiments, Mon. Not. R. Astron. Soc. **392**, 216 (2009).

[60] 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).

[61] E. Calabrese *et al.*, Cosmological parameters from pre-Planck CMB measurements: A 2017 update, Phys. Rev. D **95**, 063525 (2017).

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

[63] QUIET Collaboration, The QUIET instrument, Astrophys. J. **768**, 9 (2013).

[64] R. W. Ogburn IV *et al.*, BICEP2 and Keck Array operational overview and status of observations, in *Proceedings of SPIE* (2012), Vol. 8452, arXiv:1208.0638.

[65] M. Mallaby-Kay *et al.*, The Atacama Cosmology Telescope: Summary of DR4 and DR5 data products and data access, Astrophys. J. Suppl. Ser. **255**, 11 (2021).

[66] A. Baleato Lizancos, A. Challinor, B. D. Sherwin, and T. Namikawa, Delensing the CMB with the cosmic infrared

[67] background: The impact of foregrounds, Mon. Not. R. Astron. Soc. **514**, 5786 (2022).

[67] S. Aiola *et al.*, The Atacama Cosmology Telescope: DR4 maps and cosmological parameters, J. Cosmol. Astropart. Phys. 12 (2020) 047.

[68] 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).

[69] C. L. Bennett *et al.*, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Final maps and results, Astrophys. J. Suppl. Ser. **208**, 20 (2012).

[70] C. G. T. Haslam *et al.*, A 408 MHz all-sky continuum survey. I—Observations at southern declinations and for the North Polar region, Astron. Astrophys. **100**, 209 (1981).

[71] C. G. T. Haslam *et al.*, A 408-MHZ All-Sky continuum survey. II. The atlas of contour maps, Astron. Astrophys. Suppl. Ser. **47**, 1 (1982).

[72] N. Krachmalnicoff *et al.*, The S-PASS view of polarized Galactic synchrotron at 2.3 GHz as a contaminant to CMB observations, Astron. Astrophys. **618**, A166 (2018).

[73] I. Abril-Cabezas, C. Hervías-Caimapo, S. von Hausegger, B. D. Sherwin, and D. Alonso, Impact of Galactic dust non-Gaussianity on searches for B-modes from inflation, Mon. Not. R. Astron. Soc. **527**, 5751 (2023).

[74] E. Hivon, K. M. Gorski, C. Barth Netterfield, B. P. Crill, S. Prunet, and F. Hansen, MASTER of the CMB anisotropy power spectrum: A fast method for statistical analysis of large and complex CMB data sets, Astrophys. J. **567**, 2 (2002).

[75] K. M. Smith, Pseudo-C_ℓ estimators which do not mix E and B modes, Phys. Rev. D **74**, 083002 (2006).

[76] D. Alonso, J. Sanchez, and A. Slosar, A unified pseudo-C_ℓ framework, Mon. Not. R. Astron. Soc. **484**, 4127 (2019).

[77] 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).

[78] D. Beck, A. Cukierman, and W. L. Kimmy Wu, Bias on tensor-to-scalar ratio inference with estimated covariance matrices, Mon. Not. R. Astron. Soc. **515**, 229 (2022).

[79] A. Zonca, B. Thorne, N. Krachmalnicoff, and J. Borrill, The Python sky model 3 software, J. Open Source Softwaare **6**, 3783 (2021).

[80] G. Efstathiou, Myths and truths concerning estimation of power spectra, Mon. Not. R. Astron. Soc. **349**, 603 (2004).

[81] 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.

[82] S. Azzoni, D. Alonso, M. H. Abitbol, J. Errard, and N. Krachmalnicoff, A hybrid map-C_ℓ component separation method for primordial CMB B-mode searches, J. Cosmol. Astropart. Phys. 03 (2023) 035.

[83] A. Lewis, A. Challinor, and A. Lasenby, Efficient computation of CMB anisotropies in closed FRW models, Astrophys. J. **538**, 473 (2000).

[84] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. Gorski, healpy: Equal area pixelization

and spherical harmonics transforms for data on the sphere in Python, J. Open Source Softwaare **4**, 1298 (2019).

[85] T. Namikawa, cmblensplus: A tool to analyze cosmic microwave background anisotropies, Astrophysics Source Code Library, record ascl:2104.021 (2021).

[86] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman *et al.*, emcee: The MCMC Hammer, Astron. Soc. Pac. Conf. Ser. **125**, 306 (2013).

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

[88] P. Virtanen *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in Python, Nat. Methods **17**, 261 (2020).

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