Weak lensing of intensity mapping: The cosmic infrared background

Emmanuel Schaan,[1,2,3,*] Simone Ferraro,[4,2] and David N. Spergel[3,5]

[1]*Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, California 94720, USA*
[2]*Berkeley Center for Cosmological Physics, University of California, Berkeley, California 94720, USA*
[3]*Department of Astrophysical Sciences, Princeton University,
Peyton Hall, Princeton, New Jersey 08544, USA*
[4]*Department of Astronomy and Miller Institute, University of California, Berkeley, California 94720, USA*
[5]*Center for Computational Astrophysics, Flatiron Institute,
162 5th Avenue, New York, New York 10010, USA*

Gravitational lensing deflects the paths of cosmic infrared background (CIB) photons, leaving a measurable imprint on CIB maps. The resulting statistical anisotropy can be used to reconstruct the matter distribution out to the redshifts of CIB sources. To this end, we generalize the cosmic microwave background (CMB) lensing quadratic estimator to any weakly non-Gaussian source field, by deriving the optimal lensing weights. We point out the additional noise and bias caused by the non-Gaussianity and the "self-lensing" of the source field. We propose methods to reduce, subtract, or model these non-Gaussianities. We show that CIB lensing should be detectable with Planck data and detectable at high significance for future CMB experiments like CCAT-Prime. The CIB thus constitutes a new source image for lensing studies, providing constraints on the amplitude of structure at intermediate redshifts between galaxies and the CMB. CIB lensing measurements will also give valuable information on the star-formation history in the Universe, constraining CIB halo models beyond the CIB power spectrum. By laying out a detailed treatment of lens reconstruction from a weakly non-Gaussian source field, this work constitutes a stepping stone toward lens reconstruction from continuum or line intensity mapping data, such as the Lyman-alpha emission, absorption, and the 21 cm radiation.

I. INTRODUCTION

Weak gravitational lensing probes the projected mass distribution between the source and the observer and is therefore sensitive to the underlying cosmology. Using multiple source and lens redshift bins further allows us to track the amplitude of structure across cosmic time. Lensing introduces subtle correlations that would be forbidden by the assumption of statistical isotropy and homogeneity, and these correlations can be used to reconstruct mass maps, which include the combined effect of dark matter, baryons, neutrinos, and all other forms of energy density. At the same time, the amplitude of these fluctuations and their angular size on the sky are determined by the expansion history and the nature of the gravitational force, making weak lensing also a sensitive probe of dark energy, modified gravity, and the masses of neutrinos. It is therefore one of the most promising cosmological tools for decades to come.

So far, two phenomenologically distinct regimes of weak lensing have been explored: on the one hand, optical surveys use images of individually resolved galaxies to measure small correlations in their ellipticities induced by lensing (see Refs. [1–3] for a review). In this case, the source plane is highly non-Gaussian,[1] and measurements of shear of individual galaxies is appropriate. Galaxy lensing has been detected in a large number of surveys, including most recently by KiDS [4], DES [5], and HSC [6].

In the opposite regime, when the source is a Gaussian random field such as the cosmic microwave background (CMB) radiation, a rich theory of estimators has been developed (see Refs. [7,8] for a review), the most commonly used being the quadratic estimator of Hu and Okamoto [9]. While only optimal among the class of estimators that are quadratic in the measured lensed fluctuations,[2] it has been shown to be close to optimal for the analysis of current experiments. Recent detections include those by the Atacama Cosmology Telescope [12], the South Pole Telescope [13], POLARBEAR [14], and the Planck mission [15]. In the case

---

[*]eschaan@lbl.gov

---

[1]Meaning that the pixel-to-pixel joint probability distribution is highly non-Gaussian.

[2]More general but much more computationally expensive methods can be used, and they have been shown to perform better than the quadratic estimator in the low-noise or small-scale regimes [10,11].

of CMB lensing, the statistical properties of the source field are very well characterized, and so is its redshift. It is possible to extract information from both the shearing of the fluctuations (which are on average round for the unlensed CMB) and a local magnification or demagnification. In fact, on large scales, the quadratic estimator can be rewritten as a minimum-variance combination of estimators of shear and dilation [16].

The cosmic infrared background (CIB) [17] is the highly blended and unresolved thermal emission from a large population of dusty star-forming galaxies over a broad range of redshifts $1 \lesssim z \lesssim 4$. The CIB is an excellent probe of star-formation history, and its fluctuations reflect the clustering properties of the underlying galaxy populations. Indeed, the CIB fluctuations are highly correlated with the CMB lensing potential, meaning that it closely traces the dark matter distribution at those intermediate redshifts [18]. Its source redshift distribution is broad, and current models predict that observations at lower frequency receive a larger contribution from higher redshifts [19,20].

A simple geometric argument shows that the sensitivity to a fixed mass lens is maximized when the lens is roughly halfway between the source and the observer. Having source images at different redshifts therefore allows one to reconstruct the mass distribution in the Universe tomographically. The CIB source distribution is at intermediate redshifts between the galaxies typically used in galaxy lensing and the CMB, thus providing a useful complement.

Because of the central limit theorem, CIB fluctuations are closer to Gaussian than images of resolved galaxies but are weakly non-Gaussian nonetheless [21]. In this paper, we generalize the CMB lensing quadratic estimator to the case of any weakly non-Gaussian source field and apply this formalism to the CIB. We discuss how the non-Gaussianity of the source modifies the usual noise biases and how new biases arise due to the large redshift span of the non-Gaussian emission. Specifically, galaxies at low redshift that source the CIB emission also act as lenses for the emission from higher-redshift CIB galaxies, an effect that we call "self-lensing." This affects both the power spectrum of the inferred CIB lensing potential and its cross-correlation with low-redshift tracers. We explore methods to mitigate this self-lensing bias.

While we take the CIB as our primary example, our formalism is general and applies to lensing of any other weakly non-Gaussian sources, such as the Lyman-alpha forest [22,23], the 21 cm radiation from either the epoch of reionization or low-redshift galaxies [24–27], or any other line intensity mapping.

Finally, any residual CIB contamination in foreground cleaned CMB maps is known to bias CMB lensing reconstruction [28,29]. However, the fact that the CIB field is itself lensed introduces an additional bias, not considered previously, which we point out and discuss in this paper.

This paper is organized as follows. We begin with a heuristic review of lensing estimation in Sec. II, in order to motivate our choice of a quadratic estimator for lensing.

In Sec. III, we describe our model for the auto- and cross-power spectra for various observables, including galaxies, the lensing convergence, and the CIB. In Sec. IV, we compute the lensing correction to the CIB power spectrum. In Sec. V, we review the quadratic estimator formalism and extend it to non-Gaussian source fields. In Sec. VI, we present forecasts for current and future experiments, showing that the signal should be detectable at high significance. In Sec. VII, we explore how the usual lensing biases are modified in the case of a non-Gaussian source and show that new biases are generated by the extended and non-Gaussian nature of the CIB. Finally, in Sec. VIII, we discuss the bias to CMB lensing given some residual level of CIB contamination to CMB maps, taking into account that the CIB itself is lensed. We conclude in Sec. IX. Appendices A–E present the details of the CIB halo model implemented in this paper, a full derivation of the lensing kernel for CIB lensing, a derivation of the lensing signal-to-noise in a simple limiting (Poissonian) regime, and the method used for our direct measurement of the CIB trispectrum on the Planck eneralized needlet internal linear combination (GNILC) maps.

Throughout the paper, we will assume a flat Lambda cold dark matter (ΛCDM cosmology with cosmological parameters from the 2015 Planck release (column 3 of Table 4 of Ref. [30]).

II. HEURISTIC REVIEW OF LENSING ESTIMATION

In this section, we present a heuristic review of the lensing estimation methods used for the CMB and galaxies. Understanding the lensing estimators in the two limiting regimes, from discrete pointlike objects to a continuum Gaussian random field, will inform us about what estimator to use in the intermediate case of the CIB. The goal of this section is therefore to provide motivation for the CIB quadratic lensing estimator we present below. Lens reconstruction consists of inferring the unlensed map and the convergence map, given the observed lensed map. In the case of CMB lensing, this problem is well posed because the statistics of the unlensed CMB is largely understood: it is a Gaussian random field with a known power spectrum. Thanks to this prior on the unlensed map, the exact posterior for the lensing map can be written explicitly [10], and numerical exploration of this posterior for realistic data sets is possible [11,31].

On the contrary, in the case of galaxy lensing, the unlensed images are non-Gaussian and complex, and a full prior on the unlensed field is not readily available. One way around this is to assume a reasonable partial prior. For instance, barring intrinsic alignments, unlensed galaxy ellipticities are assumed to be uncorrelated. One builds a catalog of the individual observed galaxy ellipticities and infers the shear. Another partial prior is that galaxy sizes and brightness are uncorrelated on large scales. Yet another one is that galaxy positions at high and low redshifts should

be uncorrelated. These priors can be used to detect magnification [32–35].

Consider now the "large-scale lens regime," where the lensing field varies on scales larger than the typical fluctuations in the unlensed image. This is the regime of galaxy lensing, where shear and convergence are coherent on the scale of several galaxies. This regime also occurs in CMB lensing, for lensing modes that are coherent over many CMB hot and cold spots. In this case, the CMB lensing quadratic estimator effectively looks for distortions of the locally measured power spectrum. Indeed, for a small patch with roughly uniform shear γ and convergence κ, the local power spectrum is modified as [16]

$$C_\ell = C_\ell^0 \left[1 + \kappa \frac{\partial \ln \ell^2 C_\ell^0}{\partial \ln \ell} + \gamma \cos(2\theta_\ell) \frac{\partial \ln C_\ell^0}{\partial \ln \ell} \right], \qquad (1)$$

where C_ℓ^0 and C_ℓ are the unlensed and lensed power spectra, respectively; κ is the convergence; and γ is the shear amplitude, assumed to be uniform on the patch where the power spectrum is measured. The angle θ_ℓ is the angle between the direction of the shear and the wave vector ℓ. As this equation shows, magnification results in a monopole distortion of the two-dimensional (2D) power spectrum, and shear produces a quadrupolar distortion. This is illustrated in Fig. 1.

In this large-scale lens regime, where the CMB quadratic estimator measures shear and magnification, it is close to optimal. However, this estimator would be suboptimal in galaxy lensing, for several reasons. For example, applying a quadratic estimator to the intensity map of a highly populated galaxy field would implicitly weight galaxies by their brightness, instead of the uncertainty on their shapes. The estimator would thus be dominated by the few brightest galaxies in the field.
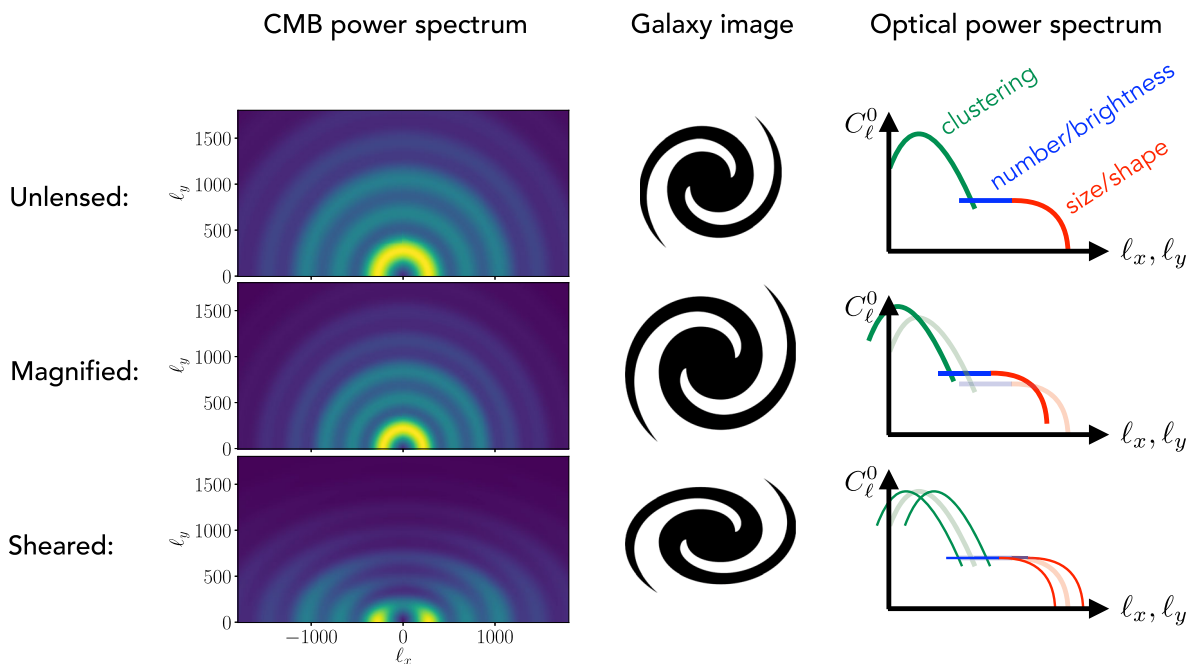

FIG. 1. Schematic illustration of the large-scale lens regime, in which the unlensed image varies on much smaller scales than the lensing convergence field. In this regime, shear and convergence are uniform on the scale of several galaxies, or several CMB hot or cold spots. The figure shows the analogy between galaxy lensing and CMB lensing estimators in this regime. Left column: The large-scale lens regime is one of the regimes in which the CMB lensing quadratic estimator operates. In this regime, the quadratic estimator can be shown to look for monopolar (magnification) and quadrupolar (shear) distortions in the local observed power spectrum [16,39]. Central column: The shear is estimated from the galaxy shape (quadrupole of the image) and in principle magnification from the galaxy size, brightness, or number density (monopole of the image). Right column: Naive schematic of the power spectrum of an optical image, on a field containing galaxies and with uniform magnification/shear. We schematically describe the power spectrum as a clustering component, plus a 1-halo or 1-galaxy term. The amplitude of this 1-halo/1-galaxy encodes both the galaxy number density and brightness, and its turnover encodes the galaxy size and shape. The effect of magnification is to rescale the multipoles ℓ_x and ℓ_y isotropically as well as the power spectrum amplitude. Magnification thus affects the local number density, brightness, and size of the galaxies, without distorting their shapes. On the other hand, the effect of shear is an anisotropic rescaling of the multipoles ℓ_x and ℓ_y, leaving the number density and brightness unchanged. This figure shows that the shape, size, brightness, and number density measured from individual objects (individual galaxies or CMB hot spots) are also encoded in the power spectrum of these objects (power spectrum of the galaxy field image or the CMB). In the large-scale lens regime, the information measured by galaxy lensing estimators on individual objects is completely analogous to that measured by the quadratic estimator on the CMB power spectrum.

The CIB is a somewhat intermediate case. Similarly to galaxy images, the unlensed CIB is a non-Gaussian field for which a full statistical description is not available. Without an explicit prior on the unlensed CIB, it is not possible to write down an explicit posterior and explore it, as is done for CMB lensing. However, the CIB is similar to the CMB in that it is a continuous field, with small fluctuations around a mean value (of order percent). Furthermore, we do have some knowledge about the non-Gaussianity of the CIB: its bispectrum has been measured [18], and halo models predicting its trispectrum exist [19,20]. The non-Gaussianity of the CIB is weak, in a sense that we shall define precisely later. Intuitively, this is expected from the central limit theorem and the fact that the CIB is the superposition of many blended galaxy emissions. It is thus natural to build upon a quadratic estimator designed for Gaussian random fields and derive the optimal weights for any weakly non-Gaussian field. We further quantify the statistical error and biases of this estimator, due to the non-Gaussianity of the CIB. On small scales ($\ell \gtrsim 1000$), the CIB is dominated by the galaxy shot noise, which is expected to be nearly Poissonian. On these scales, similar to the galaxy lensing case, the CIB multipoles become highly correlated. Accounting for these correlations is then crucial to avoid overcounting the lensing information.

The rest of the paper formalizes this heuristic intuition. To start, we first need to introduce our modeling of the galaxy number density field, the lensing convergence field, and the CIB temperature fluctuations. This is done in the next section.

III. MODELING TRACERS, LENSING POTENTIAL, AND THE CIB

A. Galaxy, CIB, lensing fields, and their power spectra

In this section, we define our notations for the auto- and cross-power spectra of tracers, lensing potential or convergence, and the CIB. For the number counts of galaxies (or any other tracer of the matter distribution), we denote the fractional fluctuations in number density n_g in direction $\hat{\boldsymbol{n}}$ on the sky as $\delta_g(\hat{\boldsymbol{n}}) = n_g(\hat{\boldsymbol{n}})/\bar{n}_g - 1$. Lensing is expressed interchangeably in terms of the lensing potential ϕ or convergence κ, related by[3] $\kappa = \frac{1}{2}\nabla^2\phi$. We denote by κ_{gal}, κ_{CMB}, and κ_{CIB} the lensing convergence reconstructed from galaxy shapes, the CMB, and the CIB, respectively.

---

[3]Throughout this paper, we adopt the optical lensing sign convention (e.g., Ref. [1]) rather than the CMB lensing convention (e.g., Ref. [7]). This only affects the sign of the displacement vector and of the lensing potential but not the sign of the convergence.

Each observable $A \in \{\delta_g, \kappa_{\text{gal}}, \kappa_{\text{CIB}}, \kappa_{\text{CMB}}\}$ is a projection of the total matter density contrast δ_m between the source and the observer, weighted by an efficiency kernel W_A,

$$A(\boldsymbol{n}) = \int d\chi W^A(\chi)\delta_m(\chi\hat{\boldsymbol{n}},\chi), \qquad (2)$$

where χ denotes the comoving radial distance or the comoving angular diameter distance—identical for a flat Universe.

For the projected galaxy overdensity field $\delta_g(\hat{\boldsymbol{n}})$, the efficiency kernel is

$$W^g(\chi) = b_g(z)\frac{1}{n_g}\frac{dn_g}{dz}\frac{dz}{d\chi}, \quad \text{with} \quad n_g = \int dz\frac{dn_g}{dz}, \quad (3)$$

where $b_g(z)$ is the galaxy bias and dn_g/dz is the redshift distribution of the galaxies.

In the case of lensing, the integral in Eq. (2) is taken along the unperturbed (straight) path between the source and the observer, rather than on the true (perturbed) path. This Born approximation has been extensively studied and shown to be an excellent approximation in the regime considered [36–38]. Given a source at comoving distance χ_S, the lensing efficiency is

$$W^\kappa(\chi,\chi_S) = \frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_m^0\frac{\chi}{a(\chi)}\left(1-\frac{\chi}{\chi_S}\right). \qquad (4)$$

The CMB originates from the surface of last scattering at comoving distance χ_{SLS}, corresponding to $z_{\text{SLS}} \sim 1100$. As a result, the CMB lensing kernel is given by

$$W^{\kappa_{\text{CMB}}}(\chi) = W^\kappa(\chi,\chi_{\text{SLS}}). \qquad (5)$$

For the lensing of galaxies, the sources are distributed in redshift, and the efficiency kernel is obtained by averaging over the source distribution $dn_S/d\chi_S$:

$$W^{\kappa_{\text{gal}}}(\chi) = \frac{1}{n_S}\int d\chi_S\frac{dn_S}{d\chi_S}W^\kappa(\chi,\chi_S), \quad \text{with} \quad n_S = \int d\chi\frac{dn_S}{d\chi}. \qquad (6)$$

Similarly, for the CIB, we loosely define a source distribution W^{CIB} and a resulting CIB lensing kernel $W^{\kappa_{\text{CIB}}}$. The case of CIB lensing is slightly more subtle than the cases of CMB and galaxy lensing. Indeed, one might be tempted to use the redshift distribution of the CIB monopole as the relevant CIB source distribution. However, as we explain later, we are not reconstructing the lensing from the CIB monopole but instead from the CIB fluctuations. As a result, the relevant CIB "source distribution" is related to the redshift distribution of the CIB power spectrum, not the monopole (see Fig. 2). The exact
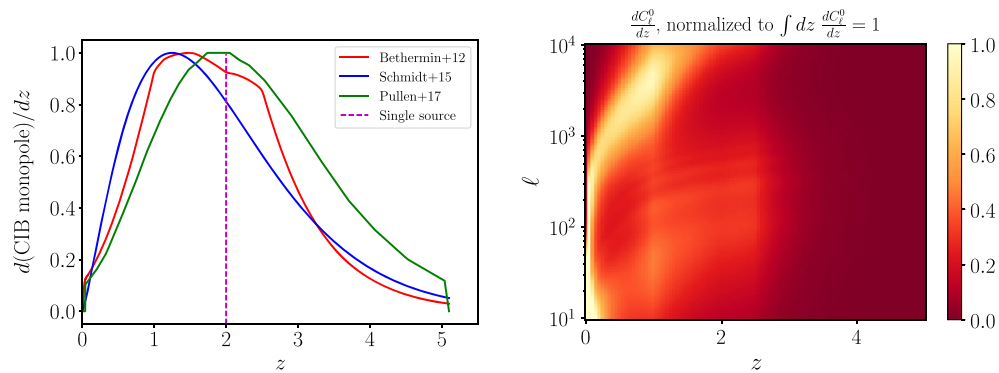
FIG. 2.   Redshift contribution to the CIB monopole (left panel) and to the CIB fluctuations (right panel) seen by Planck at 545 GHz. Left: Redshift dependence of the CIB monopole at 545 GHz from different halo models. Bethermin $+$ 12 [40] relies on the galaxy spectral energy distribution (SEDs) measured by Herschel and reproduces Herschel galaxy counts. Schmidt $+$ 15 [41] is fit to the cross-correlation of Planck High-frequency instrument (HFI) data with Sloan digital sky survey (SDSS) quasars. Pullen $+$ 17 [42] is inferred from a fit to the cross-correlation of Planck HFI data with SDSS quasars and constant mass (CMASS). This highlights a significant modeling uncertainty on the redshift distribution. Right: Since lensing is reconstructed from the CIB fluctuations rather than the monopole, we plot the redshift dependence of the CIB power spectrum, assuming the halo model of Penin $+$ 14 [20], based on the Bethermin $+$ 12 model (see the left panel). Various CIB models will differ to a similar extent as for the CIB monopole (left panel). Comparing the left and right panels, low redshifts ($z \lesssim 0.5$) make a small contribution to the CIB monopole but a large contribution to the CIB power spectrum. However, these low-redshift CIB sources are presumably known galaxies and could therefore be masked if needed.
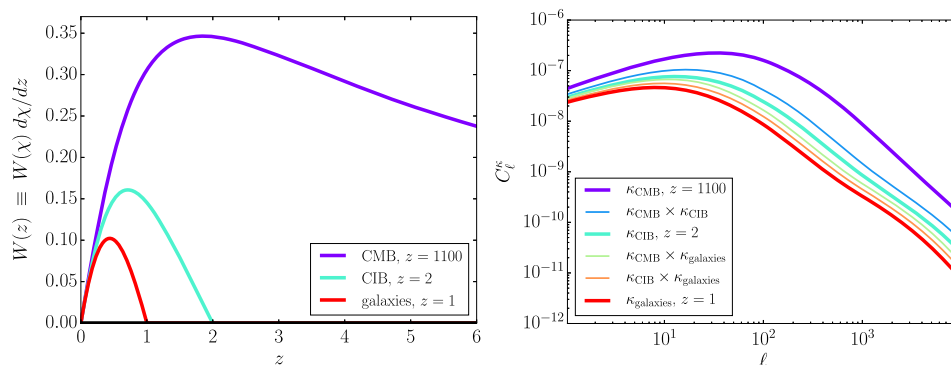

FIG. 3.   Left: Lensing efficiency kernels for the lensing of the CMB, CIB (assuming a single $z_S = 2$ source plane), and galaxies (assuming a single $z_S = 1$ source plane). CIB lensing probes an intermediate-redshift range between CMB and galaxy lensing. Right: Auto- and cross-power spectra of the lensing convergence for galaxies, the CMB, and the CIB. For lower-redshift sources, the lensing power is smaller, peaks on larger scales, and has a more visible 1-halo contribution at $\ell \gtrsim 1000$.

expression for the appropriate CIB source distribution is derived and discussed in Appendix B. Due to the very extended nature of the CIB emission, there is a non-negligible overlap between the CIB emission kernel and the CIB lensing efficiency, causing self-lensing biases that

we discuss later. Here, approximating the CIB as a single source at redshift 2 is sufficient for estimating the lensing signal. However, we will take into account the large width of the CIB redshift distribution when assessing biases in the lens reconstruction.

TABLE I.   Specifications used for Planck [48] and CCAT-Prime (see Table 2 of Ref. [49]). The Planck flux cut at 545 GHz is taken from Table 2 in Ref. [48]. For CCAT-Prime, we replaced the frequency 405 GHz by 545 GHz, in order to use the same halo model for the signal. Since the CIB emission is larger at higher frequencies, this will alter the signal-to-noise ratio in the CCAT-Prime band. We neglect this, given the large uncertainty on the actual noise level of the cleaned CIB map from CCAT-Prime. The flux cut for CCAT-Prime is obtained by assuming that point sources detected at 5σ and above via a matched filter are masked.

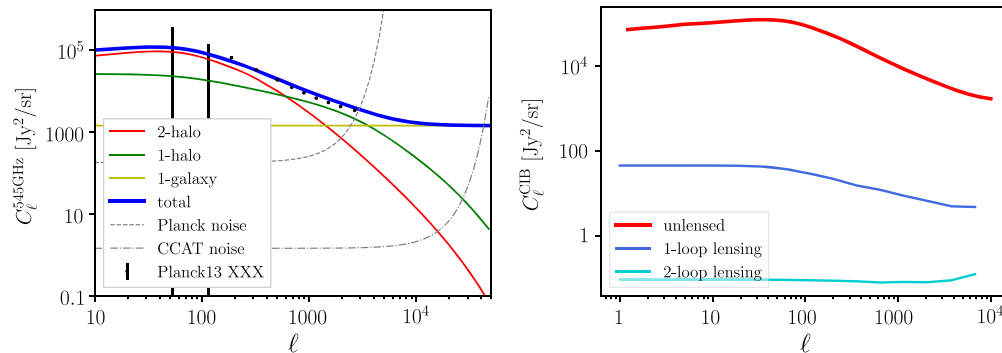|	Frequency	Beam FWHM	White noise	Flux cut	Maximum multipole	f_{sky}
Planck	545 GHz	4.8′	13.5 Jy/rad, i.e., $822\mu K'_{CMB}$	350 mJy	3000	0.4
CCAT-Prime	545 GHz	0.5′	1.2 Jy/rad, i.e., $20\mu K'_{CMB}$	73 mJy	40000	0.4

FIG. 4.   Left: Power spectrum of the CIB at 545 GHz as measured by Planck (black points with error bars) and predicted by the halo model of Penin + 14 (thick red line). The halo model calculation is the sum of the 2-halo term (dark orange line), the 1-halo term (light orange line), and the shot noise or 1-galaxy term (yellow line). Because they only start to dominate on small scales, the 1-halo and shot-noise terms are not very well constrained from the Planck power spectrum alone. The gray dashed curve shows the noise level for Planck. The gray dot-dashed curve shows the noise level for CCAT-Prime. Planck's beam and sensitivity make it a perfect CIB experiment below $\ell \lesssim 3000$, while CCAT-Prime will be a perfect CIB experiment out to $\ell \sim 50,000$. These considerations neglect the issue of component separation, and in particular the ability to distinguish the CIB from Galactic dust. Right: Comparison between the unlensed CIB power spectrum (red curve) and the corrections due to weak lensing of the CIB, from Eq. (8). The one-loop term (blue curve) corresponds to the lensing correction at first order in $C_L^{\phi\phi}$, and the two-loop term (cyan curve) corresponds to the lensing correction terms involving two powers of $C_L^{\phi\phi}$. Weak lensing thus has almost no effect on the CIB power spectrum; however, this does not mean that lensing cannot be detected, as we show below.

In the Limber and flat sky approximations, the cross-spectrum C_ℓ^{AB} between observables A and B is related to the matter power spectrum P_m via

$$C_\ell^{AB} = \int \frac{d\chi}{\chi^2} W^A(\chi) W^B(\chi) P_m \left(k = \frac{\ell + 1/2}{\chi}, \chi \right). \quad (7)$$

We show the lensing auto- and cross-power spectra in Fig. 3. For lower-redshift sources, the lensing power is smaller, peaks on larger angular scales, and has a more visible "1-halo" contribution at $\ell > 1000$. For reference, the root-mean-square lensing deflection is $1.3'$ for CIB at $z = 2$, compared to $2.3'$ for CMB at $z = 1100$.

B. Cosmic infrared background: Data and halo model

Several halo model prescriptions exist for the CIB [20,40–46]. They differ in terms of their assumptions, their level of realism, and their complexity. While most successfully reproduce the observed CIB power spectrum, their best fit parameters differ in detail (see, for instance, the redshift distributions from several models in Fig. 2).

We implement the halo model of Refs. [20,43], summarized in Appendix A, and use it throughout this paper. This model includes a halo occupation distribution, which allows to self-consistently predict the 2-halo, 1-halo, and galaxy shot-noise terms. This also allows one to compute all the halo model terms for the higher-point functions, such as the CIB trispectrum, most relevant in this paper. In this model, the average CIB intensity is determined by the galaxy flux distribution function from Ref. [40]. These galaxies are assumed to populate dark matter halos according to the halo

occupation distribution (HOD) of Ref. [47]. The CIB fluctuations are assumed to linearly trace the fluctuations in total galaxy number density, neglecting the fact that galaxy clustering is luminosity dependent.

A generic feature of the CIB halo models is that the CIB emissions at different frequencies are produced by galaxies at slightly different redshifts. Furthermore, a more sensitive instrument will resolve and typically mask more individual sources. As a result, the observed CIB depends on the frequency band and on the flux cut for the instrument. In what follows, we consider the CIB as seen by Planck [48] and by a CCAT-Prime–like experiment[4] with a configuration from Table 2 of Ref. [49]. The assumed specifications are presented in Table I. For the Planck 545 GHz CIB map, we use the 350 mJy flux cut presented in Table 2 of Ref. [48]. For CCAT-Prime, we estimate the point source detection noise with a matched filter (matched to the CCAT-Prime beam) and assume that the point sources detected at 5σ have been masked. This corresponds to a flux cut at 73 mJy.

According to the halo models (see Fig. 2), the CIB monopole has a broad redshift distribution, peaks at $z = 2$, and extends to $z \sim 4$. However, as we explain above, lensing is reconstructed from the CIB fluctuations and not the CIB monopole. The relevant redshift distribution is therefore that of the CIB power spectrum. It is much more skewed toward low redshift, especially for $\ell \lesssim 1000$. Finally, we compare the measured Planck CIB power spectrum at 545 GHz to the halo model prediction in Fig. 4.

IV. EFFECT OF LENSING ON THE CIB POWER SPECTRUM

The halo model we implement in this paper [20] already includes the effect of strong lensing on the CIB. Indeed, this halo model relies on the galaxy flux distribution function from Ref. [40], which includes the effect of highly magnified dusty star-forming galaxies on the observed flux distribution (see Fig. 3 in Ref. [40]). In this section, we focus instead on the effect of weak lensing on the CIB power spectrum.

For a source image at a given redshift, gravitational lensing deflects light by an angle[5] $d = \nabla\phi$. As described in Ref. [1], the observed map T (in units of temperature, surface brightness, number counts, or any other quantity that gets lensed) at final position $\hat{n}$ is equal to the unlensed source map T^0 at position $\hat{n} - d$: $T(\hat{n}) = T^0(\hat{n} - d)$. Taylor expanding this equation in $d = \nabla\phi$ thus allows one to compute the lensed power spectrum [7].

We approximate the CIB as a single source at redshift $z = 2$. In this single-source approximation, the lensed power spectrum can be computed to "two-loop" order (i.e., to second order in C^ϕ) as [50]

$$
\begin{aligned}
C_\ell = C_\ell^0 &\left[1 - \ell^2 D^2 + \frac{1}{2}\ell^4 D^4 \right] \\
&+ \int \frac{d^2 L}{(2\pi^2)} C_L^\phi C_{\ell-L}^0 [\boldsymbol{L} \cdot (\boldsymbol{\ell} - \boldsymbol{L})]^2 [1 - |\boldsymbol{\ell} - \boldsymbol{L}|^2 D^2] \\
&+ \frac{1}{2} \int \frac{d^2 L_1 d^2 L_2}{(2\pi)^4} C_{L_1}^\phi C_{L_2}^\phi C_{\ell-L_1-L_2}^0 [\boldsymbol{L}_1 \cdot (\boldsymbol{\ell} - \boldsymbol{L}_1 - \boldsymbol{L}_2) \\
&\times \boldsymbol{L}_2 \cdot (\boldsymbol{\ell} - \boldsymbol{L}_1 - \boldsymbol{L}_2)]^2 + \mathcal{O}((C^\phi)^3),
\end{aligned} \tag{8}
$$

where $D^2 \equiv \int \frac{dL}{4\pi} L^3 C_L^\phi$ is half of the mean squared deflection. The result is shown in Fig. 4: weak lensing changes the CIB power spectrum by less than a percent for $\ell < 10^4$. As pointed out in Ref. [7], this is expected when the unlensed power spectrum is close to featureless, resembling a power law. A more precise calculation of the lensed CIB power spectrum is possible, taking account of the extended redshift distribution of the CIB emission [50]. However, given the small size of the effect, such a calculation appears unnecessary. Such a small change in the CIB power spectrum seems extremely difficult to detect, as it would require a prior knowledge of the unlensed CIB to better than 1%. However, although weak lensing does not change the mean CIB power spectrum in a detectable way, it causes detectable spatial modulations of the power spectrum, as we show below.

---

[5]Again, we follow the optical lensing sign convention $T(\hat{n}) = T^0(\hat{n} - d)$, $d = \nabla\phi$, and $\kappa = \frac{1}{2}\nabla^2\phi$ from Ref. [1], as opposed to the CMB lensing sign convention $T(\hat{n}) = T^0(\hat{n} + d)$, $d = \nabla\phi$, and $\kappa = -\frac{1}{2}\nabla^2\phi$ from Ref. [7].

V. QUADRATIC LENSING ESTIMATOR

While lensing has little effect on the power spectrum, it introduces statistical inhomogeneities in the lensed maps, which can be used to reconstruct the lensing potential. Here, we first follow the derivation and notations of the quadratic estimator of Ref. [9] and then generalize the formalism to include non-Gaussian source distributions. Expanding T^0 at position $\hat{n} - \nabla\phi$ to linear order in ϕ yields

$$
T(\hat{n}) = T^0(\hat{n}) - \nabla\phi \cdot \nabla T^0 + \mathcal{O}(\phi^2), \tag{9}
$$

or in Fourier space,

$$
T_\ell = T_\ell^0 + \int \frac{d^2 L}{(2\pi)^2} \boldsymbol{L} \cdot (\boldsymbol{\ell} - \boldsymbol{L}) \phi_L T_{\ell-L}^0 + \mathcal{O}(\phi^2). \tag{10}
$$

The unlensed map is assumed to be statistically homogenous (or isotropic for the curved sky), which translates into uncorrelated Fourier modes,

$$
\langle T_\ell^0 T_{\ell'}^0 \rangle = (2\pi)^2 \delta_{\ell+\ell'}^D C_l^0, \tag{11}
$$

where δ^D is the Dirac delta distribution. In contrast, lensing couples Fourier modes of the unlensed map, thus breaking this statistical homogeneity:

$$
\langle T_\ell T_{L-\ell} \rangle_{\text{fixed } \phi} = \phi_L \underbrace{(-\boldsymbol{L}) \cdot [\boldsymbol{\ell} C_\ell^0 + (\boldsymbol{L} - \boldsymbol{\ell}) C_{L-\ell}^0]}_{f_{\ell,L-\ell}} + \mathcal{O}(\phi^2 C^0). \tag{12}
$$

From the last equation, we see that for a fixed realization of the lensing potential the quadratic quantity

$$
\hat{\phi}_{L,\ell} \equiv \frac{T_\ell T_{L-\ell}}{f_{\ell,L-\ell}} \tag{13}
$$

is an estimator for ϕ_L, unbiased to linear order in ϕ:

$$
\left\langle \frac{T_\ell T_{L-\ell}}{f_{\ell,L-\ell}} \right\rangle_{\text{fixed } \phi} = \phi_L + \mathcal{O}(\phi^2). \tag{14}
$$

Note that $f_{\ell,L-\ell} = f_{L-\ell,\ell}$ and $\hat{\phi}_{L,\ell} = \hat{\phi}_{L,L-\ell}$ by symmetry, and $\hat{\phi}_{L,\ell}^\star = \hat{\phi}_{-L,-\ell}$. We thus obtain many estimators of ϕ_L, by fixing $\boldsymbol{L}$ and varying $\boldsymbol{\ell}$ in $\hat{\phi}_{L,\ell}$. These may be combined into the unbiased (to first order in ϕ) minimum-variance quadratic estimator for ϕ_L, with inverse-variance weighting. This minimum-variance weighting is different for a Gaussian and non-Gaussian unlensed map, as we now present.

A. Gaussian unlensed map

First, consider the case of a Gaussian unlensed map T^0, as for the CMB. For a fixed realization of the lensing field, the lensed map is then Gaussian, too, and we get

$$\langle \hat{\phi}_{L,\ell} \hat{\phi}^{\star}_{L',\ell'} \rangle_{\text{fixed } \phi} = (2\pi)^2 \delta^D_{L-L'} \frac{(2\pi)^2 (\delta^D_{\ell'-\ell} + \delta^D_{\ell'-(L-\ell)})}{2}$$
$$\times \underbrace{\frac{2C^{\text{total}}_{\ell} C^{\text{total}}_{L-\ell}}{f^2_{\ell,L-\ell}}}_{\equiv \sigma^2_{L,\ell}}. \qquad (15)$$

Note that we have fixed the lensing realization and marginalized over the unlensed map. This equation implies that the various estimators $\hat{\phi}_{L,\ell}$ are uncorrelated, and the minimum-variance unbiased (to linear order in ϕ) quadratic estimator takes the simple form

$$\hat{\phi}_L = \frac{\int \frac{d^2\ell}{(2\pi)^2} \hat{\phi}_{L,\ell} / \sigma^2_{L,\ell}}{\int \frac{d^2\ell}{(2\pi)^2} 1 / \sigma^2_{L,\ell}}. \qquad (16)$$

As usual, inverse-variance weighting optimally weights modes according to whether they are signal or noise dominated. It further upweights temperature modes ℓ that are more aligned with the lensing mode L, since these are most affected by lensing. The optimal quadratic estimator can be recast into the familiar form [8]:

$$\hat{\phi}_L = N_L \int \frac{d^2\ell}{(2\pi)^2} \underbrace{iL \cdot}_{\text{divergence}} \left[\underbrace{\frac{T_{L-\ell}}{C^{\text{total}}_{L-\ell}}}_{\substack{\text{inverse-var} \\ \text{weighted map}}} \times \underbrace{\frac{C^0_{\ell}}{C^{\text{total}}_{\ell}} i\ell T_{\ell}}_{\substack{\text{Wiener-filtered} \\ \text{gradient map}}} \right]. \qquad (17)$$

We recover the fact that the quadratic estimator for a Gaussian unlensed map is the divergence of the real-space product of the inverse variance–weighted unlensed map by its Wiener-filtered gradient. The reconstruction noise is then

$$N_L = \left[\int \frac{d^2\ell}{(2\pi)^2} \frac{1}{\sigma^2_{L,\ell}} \right]^{-1} = \left[\int \frac{d^2\ell}{(2\pi)^2} \frac{f^2_{\ell,L-\ell}}{2 C^{\text{total}}_{\ell} C^{\text{total}}_{L-\ell}} \right]^{-1}. \qquad (18)$$

In this Gaussian case, the quadratic temperature pairs are uncorrelated, so they each add independent information. This will no longer be the case for a non-Gaussian unlensed map, as we now show.

B. Non-Gaussian unlensed map

If the unlensed map is non-Gaussian, it may have a nonzero trispectrum $\mathcal{T}^0$, defined as the connected component of the four-point function: $\langle T^0_{\ell_1} T^0_{\ell_2} T^0_{\ell_3} T^0_{\ell_4} \rangle_c \equiv (2\pi)^2 \delta^D_{\ell_1+\ell_2+\ell_3+\ell_4} \mathcal{T}^0_{\ell_1,\ell_2,\ell_3,\ell_4}$. This both increases the statistical error of the quadratic lensing estimators and correlates them[6]:

$$\langle \hat{\phi}_{L,\ell} \hat{\phi}^{\star}_{L',\ell'} \rangle_{\text{fixed } \phi} = (2\pi)^2 \delta^D_{L-L'} \underbrace{\left[\frac{(2\pi)^2 (\delta^D_{\ell'-\ell} + \delta^D_{\ell'-(L-\ell)}) C^{\text{total}}_{\ell} C^{\text{total}}_{L-\ell} + \mathcal{T}^0_{\ell,-\ell',L-\ell,\ell'-L}}{f_{\ell,L-\ell} f_{\ell',L-\ell'}} \right]}_{\equiv (\Sigma_L)_{\ell,\ell'}}. \qquad (19)$$

Because the estimators $\hat{\phi}_{L,\ell}$ are now correlated for different ℓ, the inverse-variance weighting should include the non-diagonal covariance matrix $(\Sigma_L)_{\ell,\ell'}$,

$$\hat{\phi}_L = \frac{\int \frac{d^2\ell}{(2\pi)^2} \hat{\phi}_{L,\ell} \left[\int \frac{d^2\ell'}{(2\pi)^2} (\Sigma_L)^{-1}_{\ell,\ell'} \right]}{\int \frac{d^2\ell d^2\ell'}{(2\pi)^4} (\Sigma_L)^{-1}_{\ell,\ell'}}, \qquad (20)$$

where the inverse-covariance matrix in the continuum limit is defined by

$$\int \frac{d^2\ell}{(2\pi)^2} (\Sigma_L)^{-1}_{\ell_1,\ell} (\Sigma_L)_{\ell,\ell_2} = (2\pi)^2 \delta^D_{\ell_1-\ell_2}. \qquad (21)$$

The associated reconstruction-noise power spectrum is

$$N_L = \left[\int \frac{d^2\ell d^2\ell'}{(2\pi)^4} (\Sigma_L)^{-1}_{\ell,\ell'} \right]^{-1}. \qquad (22)$$

As a result, the noise in the lens reconstruction is enhanced by the presence of the trispectrum. This lensing noise constitutes a bias in the auto-power spectrum of the reconstructed lens field, as for the Gaussian case.

To add some intuition and aid comparing to the Gaussian case, we evaluate Eq. (22) in the simple case of a pure Poisson unlensed map. That is, both the power spectrum C^0 and the trispectrum $\mathcal{T}^0$ are white, i.e., independent of the multipole. In this particular case, the lensing noise in units of κ becomes (see Appendix D for a complete derivation)

$$N^{\kappa}_L = \underbrace{\frac{2\pi}{N_{\text{modes}}}}_{\text{Gaussian noise}} + \underbrace{\frac{\mathcal{T}^0}{4(C^0)^2}}_{\text{non-Gaussian correction}}, \qquad (23)$$

where $N_{\text{modes}} \equiv 4\pi \int \frac{d^2\ell}{(2\pi)^2} = \ell^2_{\text{max}} - \ell^2_{\text{min}}$. This simple formula provides several key insights. The first term is the Gaussian contribution. This term goes to zero as the number of observed multipoles increases; if the modes are Gaussian, each of them adds independent lensing information, and the number of modes is the only limit to the reconstruction precision. The second term is the

---

[6]In principle, the right-hand side should also include the self-lensing bispectrum described in Sec. VII B. In practice, neglecting this term only increases the lensing noise but does not introduce further bias.

additional lensing noise due to the non-Gaussianity of the unlensed map. It does not decrease as the number of observed modes increases. Thus, in the limit of an infinite number of modes, the Gaussian term tends to zero, and the total noise is entirely determined by the non-Gaussian term. Another important insight is the following. For the purpose of lens reconstruction, a small ratio between the trispectrum and the squared power spectrum, i.e., $\frac{\mathcal{T}^0}{(C^0)^2} \ll 1$, is not sufficient for the field to be considered Gaussian. One needs the trispectrum to be small compared to the squared power spectrum divided by the number of modes N_{modes}: $\frac{\mathcal{T}^0}{(C^0)^2/N_{\text{modes}}} \ll 1$. This is a more stringent condition and one that depends on the experiment. As we show below, it is satisfied by Planck but not by CCAT-Prime.

C. Applying Gaussian weights to a non-Gaussian map

In order to avoid inverting the large covariance matrix, or if the unlensed trispectrum is not well known, one may apply the quadratic estimator with Gaussian weights to a non-Gaussian map. If the non-Gaussianity is large, i.e., $\frac{\mathcal{T}^0}{(C^0)^2/N_{\text{modes}}} \ll 1$, the Gaussian weighting is suboptimal; it may increase significantly the reconstruction noise but does not introduce bias. When applying the Gaussian weights of Eq. (16) to a non-Gaussian map, the lensing-noise power spectrum is no longer given by Eq. (22), but instead

$$N_L = \frac{\int \frac{d^2\ell d^2\ell'}{(2\pi)^4} (\Sigma_L)_{\ell,\ell'}/\sigma^2_{L,\ell}\sigma^2_{-L,\ell'}}{[\int \frac{d^2\ell}{(2\pi)^2} 1/\sigma^2_{L,\ell}]^2} = N_L^0 + \mathcal{T}_L^0, \quad (24)$$

where the Gaussian reconstruction-noise N_L^0 is the same as in Eq. (18),

$$N_L^0 = \left[\int \frac{d^2\ell}{(2\pi)^2} \frac{1}{\sigma^2_{L,\ell}} \right]^{-1} = \left[\int \frac{d^2\ell}{(2\pi)^2} \frac{f^2_{\ell,L-\ell}}{2C^{\text{total}}_\ell C^{\text{total}}_{L-\ell}} \right]^{-1}, \quad (25)$$

and the additional contribution due to the unlensed source trispectrum is

$$\mathcal{T}_L^0 = (N_L^0)^2 \int \frac{d^2\ell d^2\ell'}{(2\pi)^4} \frac{f_{\ell,L-\ell}}{2C^{\text{total}}_\ell C^{\text{total}}_{L-\ell}} \frac{f_{\ell,-L-\ell'}}{2C^{\text{total}}_{\ell'} C^{\text{total}}_{-L-\ell'}}$$
$$\times \mathcal{T}^0_{\ell,L-\ell,\ell',-L-\ell'}. \quad (26)$$

As expected, the reconstruction noise is enhanced by the unlensed trispectrum relative to the Gaussian case.

In particular, for a white trispectrum (but an arbitrary power spectrum), this non-Gaussian noise simplifies to

$$\mathcal{T}_L^0 = (N_L^0)^2 \left[\int \frac{d^2\ell}{(2\pi)^2} \frac{f_{\ell,L-\ell}}{2C^{\text{total}}_\ell C^{\text{total}}_{L-\ell}} \right]^2 \mathcal{T}^0. \quad (27)$$

In Sec. VI, we evaluate this expression to quantify the effect of the shot-noise trispectrum on the CIB lensing estimator.

If the power spectrum and trispectrum are both white, e.g., in the shot-noise regime of the CIB, Eq. (27) further simplifies to $\mathcal{T}_L^0 = \mathcal{T}^0/(L^4 C^2)$ (Appendix D). In this very special case, despite the potentially large non-Gaussianity, we recover the corresponding result for the non-Gaussian weights. Indeed, if both the power spectrum and the trispectrum are white, then all temperature modes are equivalent and will be weighted equally by both the Gaussian and non-Gaussian weights.

D. CIB trispectrum: Halo model prediction and map-based measurement

As we have shown, the unlensed CIB trispectrum contributes to the lensing noise. Furthermore, as we shall demonstrate shortly, the quadratic CIB lensing estimator best reconstructs the low lensing multipoles L and relies mostly on the high temperature multipoles ℓ. In other words, the relevant trispectrum configuration is the collapsed one: $\mathcal{T}^{\text{CIB}}_{\ell,L-\ell,\ell',-L-\ell'}$, with $L \ll \ell$. In what follows, we therefore focus exclusively on this configuration. This configuration can be measured from the data without implementing a complete trispectrum estimator (see Appendix E), and it also simplifies the halo model calculations (see Appendix A).

In Fig. 5, we show estimates of the CIB trispectrum from the halo model, together with a map-based measurement. As shown in the left panel, the halo model prediction for the CIB trispectrum appears to be extremely sensitive to the low-redshift CIB sources, as defined in Appendix A 2. This makes it difficult to reliably estimate the CIB trispectrum from the halo model, given current observables. For this reason, we also measured the collapsed trispectrum from the Planck GNILC CIB map at 545 GHz. The details of the method are presented in Appendix E. Our measurement from the Planck maps extends to $\ell = 1500$. Over this range of multipoles, we do not detect a collapsed trispectrum but place an upper limit. This upper limit excludes most halo model components below $\ell = 1500$, except for the 4-halo and shot-noise terms, This implies that the "excluded" terms are highly overestimated in the halo model we implemented. Although our measurement does not constrain the higher multipoles directly, it suggests that the shot-noise trispectrum is the only dominant term for $\ell \gtrsim 1500$, where most of the lensing information originates. Indeed, all other halo model terms decrease rapidly with ℓ, so if their amplitudes were consistent with the upper limit at $\ell = 1500$, they would be small compared to the shot noise at $\ell \gtrsim 1500$. As shown in Fig. 6, the non-Gaussian lensing noise due to the shot-noise trispectrum is negligible compared to the Gaussian contribution for Planck and dominant for CCAT-Prime. Thus, the Gaussian lensing weights are close to optimal for Planck and likely suboptimal for CCAT-Prime.

The halo model term which is least uncertain is probably the shot-noise trispectrum, since it comes from measured
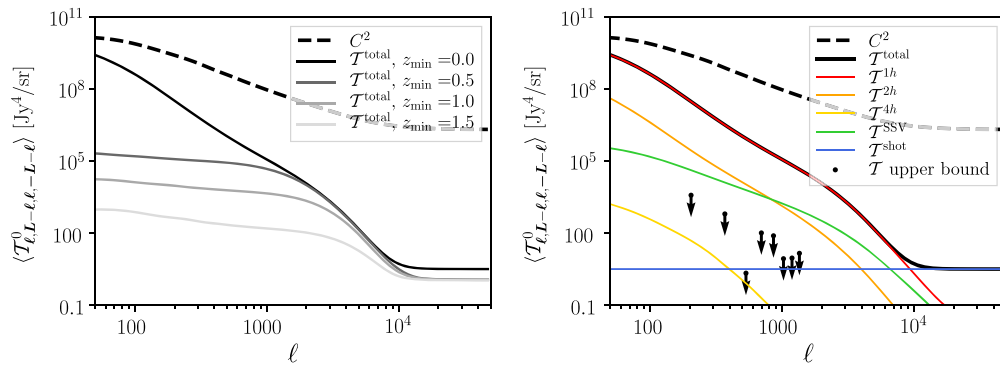
FIG. 5. Measure of the CIB non-Gaussianity relevant for the lensing-noise power spectrum: comparison between the angle-averaged collapsed CIB trispectrum (solid black line) and the squared CIB power spectrum (dashed black line) Left panel: We show the large dependence of the halo model trispectrum on the minimum-redshift cut z_{min} in the halo model trispectrum, Eq. (A11). We vary this value from 0 to 1.5 and show the corresponding CIB trispectra in shades of gray. Rigt panel: Keeping $z = 0$ as the minimum-redshift cut, we compare the various halo model terms (colored lines) to our trispectrum measurement from the Planck GNILC map at 545 GHz. The measurement procedure is presented in Appendix E. We find no significant evidence for a nonzero collapsed trispectrum and show one-sided 2σ upper bounds as downward triangles. The nondetection of a collapsed trispectrum with the optimal estimator for Planck implies that CIB non-Gaussianity can be safely ignored for the CIB lens reconstruction with Planck. The measurement only extends to $\ell = 1500$ and therefore does not directly constrain the higher multipoles. However, the fact that most halo model terms are excluded at low multipoles suggests that only the shot-noise trispectrum will be significant at higher multipoles. We include it in the Planck and CCAT-Prime forecasts, which rely primarily on these high multipoles.

luminosity functions. As we show below, CIB lensing from CCAT-Prime relies mostly on temperature multipoles above $\ell = 10,000$, where the shot noise is the only dominant term. As a result, our forecast for the non-Gaussian lensing noise for CCAT-Prime should be robust.

Since the CIB non-Gaussianity increases the lensing noise for CCAT-Prime, one natural question is whether the CIB can be "Gaussianized." An efficient way of achieving this is to mask rare, bright objects, which contribute little to the power

spectrum but a lot to the trispectrum. For instance, the deeper CCAT-Prime flux cut reduces the shot-noise power spectrum by less than 10% (compared to the 350 mJy flux cut of Planck) but lowers the shot-noise trispectrum by an order of magnitude. Masking known infrared sources from galaxy surveys should further reduce the CIB trispectrum, while marginally affecting the CIB power spectrum.

On the other hand, the apparent discrepancy between the halo model and the data also suggests that the higher-point
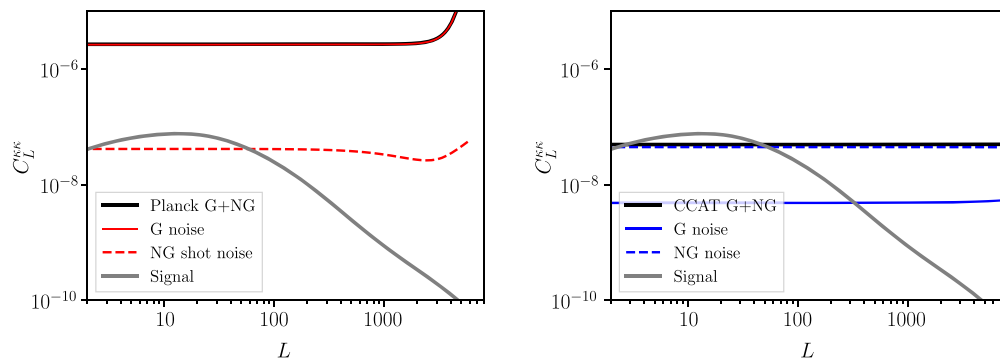

FIG. 6. Noise per multipole on the CIB lens reconstruction, using Gaussian weights for the quadratic estimator. Solid black lines show the lensing signal and the lensing noise. The solid red curve shows the lensing noise expected if the CIB were perfectly Gaussian. The dashed red curve shows the additional noise contribution from the shot-noise trispectrum. Left panel: Planck CIB lensing. As explained in the main text, the CIB trispectrum does not alter the lensing noise in a detectable manner, since it is not detected by our collapsed trispectrum estimator. In particular, the lensing noise due to the CIB shot-noise trispectrum is negligible compared to the Gaussian noise. Right panel: CCAT-Prime CIB lensing. The lensing noise is dominated by the non-Gaussian contribution, i.e., the CIB shot-noise trispectrum. The lensing modes between a few and 50 are signal dominated. CIB lensing from CCAT-Prime relies on higher multipoles than for Planck (see Fig. 7). While the shot-noise trispectrum is smaller for CCAT-Prime than for Planck, due to the lower flux cut, the CIB power spectrum is also smaller at these higher multipoles. As a result, the non-Gaussian lensing noise is similar for CCAT-Prime and Planck, scaling roughly as $\frac{\mathcal{T}^0}{4(C^0)^2}$.

functions of the CIB contain useful information to distinguish between different halo models of the CIB and thus constrain the history of star formation in the Universe.

VI. FORECASTS

We forecast the signal-to-noise ratio for the CIB lensing auto- and cross-correlations, for Planck and CCAT-Prime. Our specifications are presented in Table I. Because the CIB trispectrum is very uncertain, we assume Gaussian weights for the quadratic estimator for both Planck and CCAT-Prime. As discussed above, this should be optimal for Planck, but not for CCAT-Prime. In both cases, we compute the noise contribution due to the CIB shot-noise trispectrum. The resulting lensing noises for Planck and CCAT-Prime are shown in Fig. 6, in the left and right panels, respectively. In Fig. 7, we show that the lens reconstruction is dominated by the highest signal-dominated multipoles available in the experiment: $\ell \sim 1000$ for Planck and $\ell \sim 10,000$ for CCAT-Prime. This justifies retaining only the shot-noise term in the CIB trispectrum.

We then forecast the signal-to-noise ratio on the auto- and cross-correlations between CIB lensing and various tracers of the matter density. The results are summarized in Table II. For Planck CMB lensing, we consider CMB maps with $5'$ beam full width at half maximum (FWHM), $60\ \mu\mathrm{K}'$ noise in temperature, and $60\sqrt{2}\ \mu\mathrm{K}'$ noise in E- and B-mode polarizations at 143 GHz [48]. We use $\ell_{\max\,\mathrm{T}} = \ell_{\max\,\mathrm{P}} = 2000$. This reproduces the lensing noise measured by the Planck Collaboration [15]. For CMB lensing from a stage 4 CMB experiment (CMB S4), we assume a $1'$ FWHM beam, $1\mu\mathrm{K}'$ noise in temperature, and
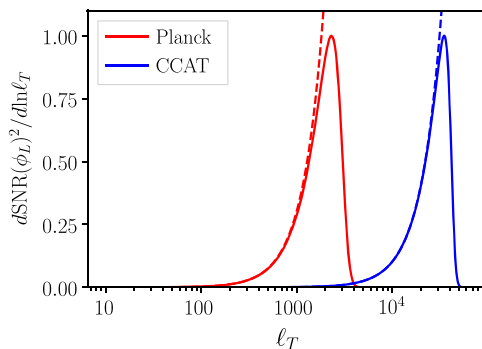

FIG. 7. Contribution from the various ℓ-modes of the CIB temperature map to the lensing potential at $L = 100$, for Planck (solid red curve) and CCAT-Prime (solid blue curve). To show both curves on the same axis, they have been normalized to a maximum of unity. The dashed lines show the expectation for Planck and CCAT-Prime when including only sample variance mode counting [$N_L^\kappa = 2\pi/N_{\mathrm{modes}}$ from Eq. (23)]. They show that most of the lensing information comes from the highest multipoles that are signal dominated in the CIB map, since they contribute the largest number of modes. For these high multipoles, the trispectrum can be approximated by the shot-noise term.
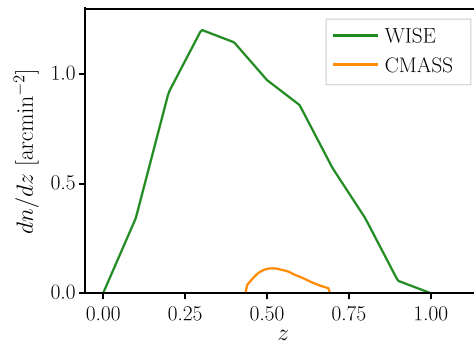

FIG. 8. Redshift distribution for WISE (green curve) and CMASS (orange curve). As lens galaxies, these may be correlated with the reconstructed CIB lensing potential.

$\sqrt{2}\mu\mathrm{K}'$ noise in E- and B-mode polarization at 143 GHz [51,52] and assume $\ell_{\max\,\mathrm{T}} = 3000$ and $\ell_{\max\,\mathrm{P}} = 5000$. For the CMASS galaxy sample, we assume a galaxy number density of $0.02\ \mathrm{arc\,min}^{-2}$ and bias of 2 over 24% of the sky. For the WISE galaxy catalog, we assume 0.6 galaxies per squared arc min, with a bias of 1.13, over 40% of the sky. The redshift distributions of CMASS and WISE are shown in Fig. 8.

As shown in Table II, the auto-power spectrum of CIB lensing is not detectable with existing Planck CIB data. However, the cross-correlations of Planck CIB lensing with tracers should be significant, including the cross-correlations with CMB lensing and galaxy catalogs like WISE or CMASS. With CCAT-Prime data, CIB lensing will be detectable both in auto- and cross-correlation at high significance. CIB lensing will therefore provide an independent measurement of the amplitude of density fluctuations. However, the cosmological interpretation of this measurement will be affected by the uncertainty on the redshift distribution of the CIB. On the other hand, one can use CIB lensing measurements and the current best estimates of cosmological parameters to constrain the CIB redshift distribution. Moreover, if CIB lensing

TABLE II. Forecasts for the signal-to-noise ratio of each auto-/cross-correlation. These values of the signal-to-noise ratio include the Gaussian and non-Gaussian lensing noises as well as the cosmic variance. They thus correspond to the final error on the measured amplitude of the signal. The signal-to-noise ratio values for detection alone, i.e., without the cosmic variance, are therefore slightly larger than those quoted in this table. These signal-to-noise ratios are statistical only and therefore do not include the biases discussed in Sec. VII.

	Planck κ_{CIB}	CCAT-Prime–like κ_{CIB}
Auto-power, $f_{\mathrm{sky}} = 0.4$	1	37
Planck κ_{CMB}, $f_{\mathrm{sky}} = 0.4$	5	28
CMB S4 κ_{CMB}, $f_{\mathrm{sky}} = 0.4$	12	80
WISE, $f_{\mathrm{sky}} = 0.4$	18	129
CMASS, $f_{\mathrm{sky}} = 0.24$	6	40

measurements are combined with galaxy or CMB lensing, one can effectively use distance ratios (also referred to as "cosmography") to determine the CIB redshift distribution, without requiring a cosmological model.

VII. BIASES TO THE AUTO- AND CROSS-CORRELATION

A. Lens reconstruction biases: Overview

In addition to the usual noise biases present in CMB lensing reconstruction [53,54], the non-Gaussian nature of CIB emission over a broad redshift range introduces new forms of bias. These affect both the $\hat{\phi}$ power spectrum and its cross-correlation with tracers in slightly distinct ways:

(i) Standard noise biases:

   It is well known that the power spectrum of the standard quadratic estimator $\hat{\phi}$ is subject to noise biases, even in the case of the Gaussian source and lensing potential. Schematically, we can write

   $$\langle C_L^{\hat{\phi}\hat{\phi}} \rangle = N_L^{(0)} + C_L^{\phi\phi} + N_L^{(1)} + \mathcal{O}([C_L^{\phi\phi}]^2). \quad (28)$$

   The $N_L^{(0)}$ term is present even when applying the quadratic estimator to a Gaussian unlensed source field and represents the disconnected part of the temperature four-point function. Following the convention in the literature, we denote the noise biases by $N^{(n)}$, where n is the power of $C_L^{\phi\phi}$ appearing explicitly in the evaluation of the bias term.

   At lowest order, the *connected* part of the $\hat{\phi}$ power spectrum expansion contains the true signal $C_L^{\phi\phi}$ and another contraction $N_L^{(1)}$ at the same order (but typically smaller than the signal, at least for CMB lensing) [53,54]. At higher order, additional biases $N_L^{(2)}$, $N_L^{(3)}$, etc., may become important for small-scale or low-noise experiments. In cross-correlation with tracers, $N_L^{(2)}$ is the lowest order bias present [54].

   The higher order biases are partially due to the first order Taylor expansion used in deriving the quadratic estimator:

   $$T(\hat{n}) = T^0(\hat{n} - \nabla\phi) \approx T^0(\hat{n}) - \nabla\phi \cdot \nabla T^0. \quad (29)$$

   This linear truncation of the expansion is inexact and misses all the terms $\mathcal{O}(\phi^n)$ for $n \geq 2$. As a result, the quadratic estimator $\hat{\phi}$ is biased by terms $\mathcal{O}(\phi^2)$. An estimator including higher order Taylor expansion terms, or a maximum likelihood estimator where the lensed field contains the correct pixel to pixel remapping by lensing, would avoid some part (but not all) of the biases $N_L^{(1)}$, $N_L^{(2)}$, etc.

(ii) Non-Gaussian unlensed map T^0:

   Each noise term $N_L^{(n)}$ involves a four-point function of the unlensed temperature map and is therefore enhanced by the unlensed trispectrum. Schematically, when using Gaussian weights, $N_L^{(n)} \rightarrow N_L^{(n)} + \mathcal{T}_L^{(n)}$, where $\mathcal{T}_L^{(n)}$ is an integral over the unlensed trispectrum $\mathcal{T}^0$ weighted by $2n$ powers of ϕ. The $n = 0$ case is given by Eq. (26).

   As discussed previously, the lowest order bias in cross-correlation is $N^{(2)}$. This makes the cross-correlation more robust to uncertainties in the modeling of the trispectrum, which can be hard to disentangle from the lensing contribution.

(iii) Non-Gaussian lensing potential ϕ:

   As the lensing potential is sourced by large-scale structure, it is affected by the nonlinear evolution under gravity. This makes the lensing potential non-Gaussian on small scales. A consequence is the appearance of new bias terms $N^{(j)}$ that involve both odd and even powers of ϕ (as opposed to even powers only if ϕ is Gaussian). For example, terms like $N^{(3/2)}$ or $N^{(5/2)}$ appear and involve the lensing potential bispectrum or five-point function, respectively [55]. The other biases $N^{(n)}$ for integer values of n are also enhanced by the connected $(2n)$-point function of ϕ. For example, the $N^{(2)}$ bias is enhanced by the four-point function of ϕ. These biases also appear in CMB lensing and are subject to active research.

(iv) Self-lensing:

   When the source distribution is broad in redshift, as for the CIB, the lower-redshift sources can act as lenses for the higher-redshift sources and therefore introduce a bias in both the auto- and cross-correlation. We call this effect self-lensing. This is analogous to the dominant contribution of some of the foreground-induced biases to CMB lensing [28,29,56], in which part of the foregrounds leaks into the temperature map. Given the importance of this term, which affects both the auto- and cross-correlation, we explore it in detail in the next section.

(v) Foreground contamination:

   Analogously to CMB lensing, the imperfect removal of foregrounds such as Galactic dust or extragalactic thermal Sunyaev Zel'dovich (tSZ) can lead to biases in the lens reconstruction [28,29,56]. While the tSZ signal can be accurately cleaned due to its unique frequency dependence, Galactic dust is highly degenerate with the CIB, especially at lower frequencies, and residual amounts can lead to a bias for the lensing reconstruction.

(vi) Magnification bias:

   Magnification bias occurs in galaxy catalogs with a fixed detection threshold. In magnified regions, galaxies appear brighter, and therefore more galaxies will be detected. This effect partially compensates the expected dilution of the galaxy number density due to magnification.

However, as surface brightness is conserved in lensing, intensity maps with no detection threshold, e.g., in CIB or 21 cm intensity mapping [57], do not suffer from magnification bias. In practice, when masking a bright point source in a realistic CIB analysis, some magnification bias is introduced. Because point sources are relatively rare and have a small effect on the power spectrum (compared to the trispectrum), we expect this bias to be negligible.

B. Self-lensing contribution

The CIB self-lensing bias is simpler to evaluate in cross-correlation with tracers than in auto-correlation. We discuss them separately in this section.

1. Cross-correlation

Suppose that we cross-correlate the reconstructed $\hat{\phi}$ with a low-redshift tracer δ_g. For concreteness, we consider the cross-correlation with galaxies, but the same formalism applies to any tracer of matter, such as galaxy lensing or CMB lensing. Since these galaxies produce some IR emission that contributes to the CIB, we decompose the unlensed CIB emission T^0 into a high-redshift component $T^{\mathrm{high}-z}$ (uncorrelated with the tracer) and a component T_g originating from the tracers themselves, or objects correlated with them. Thus, $T^0 = T_{\mathrm{high}-z} + T_g$. Since the $\hat{\phi}$ estimator is quadratic in T, $\hat{\phi} \sim TT$, the cross-correlation $\langle \hat{\phi} \delta_g \rangle$ contains a term of the form $\langle T_g T_g \delta_g \rangle \propto B_{\delta_g T_g T_g}$, proportional to the hybrid bispectrum of two powers of T_g and one power of δ_g.

Following the Appendix of Ref. [56], the self-lensing bias to the cross-correlation $C_L^{\hat{\phi} \times g}$ is given by

$$(\Delta C_L^{\hat{\phi} \times g})_{\mathrm{SL}} = N_L \int d\chi \frac{W^g(\chi)(W^{\mathrm{CIB}}(\chi))^2}{\chi^2} \mathcal{B}(\mathbf{k} = \boldsymbol{L}/\chi; \chi),$$

$$(30)$$

with

$$\mathcal{B}(\mathbf{k}; \chi) \equiv \int \frac{d^2\mathbf{q}}{(2\pi)^2} \frac{f_{\mathbf{q}\chi, \mathbf{k}\chi - \mathbf{q}\chi}}{2 C_{\mathbf{q}\chi}^{\mathrm{total}} C_{\mathbf{k}\chi - \mathbf{q}\chi}^{\mathrm{total}}} B_{\delta T_g T_g}(\mathbf{k}, \mathbf{q}, -\mathbf{k} - \mathbf{q}; \chi).$$

$$(31)$$

In the expression above, the momenta $\mathbf{k}$, $\mathbf{q}$ and $-\mathbf{k} - \mathbf{q}$ lie on a plane perpendicular to the line of sight.

2. Auto-correlation

Since the true lensing potential ϕ is also a tracer of the matter at low redshift, the previous discussion implies a bias to the auto-correlation of the reconstructed lensing potential,

$$(\Delta C_L^{\hat{\phi} \hat{\phi}})_{\mathrm{SL}} \approx 2(\Delta C_L^{\hat{\phi} \phi})_{\mathrm{SL}} + \text{secondary contractions}, \quad (32)$$

where the combinatorial factor of 2 arises from the two possible choices of $\hat{\phi}$ as a tracer. From Eq. (30), we have

$$(\Delta C_L^{\hat{\phi} \phi})_{\mathrm{SL}} = N_L \int d\chi \frac{W_L^{\phi_{\mathrm{CIB}}}(\chi)(W^{\mathrm{CIB}}(\chi))^2}{\chi^2} \mathcal{B}(\mathbf{k} = \boldsymbol{L}/\chi; \chi),$$

$$(33)$$

with $W_L^{\phi_{\mathrm{CIB}}}(\chi) = -2 W^{\kappa_{\mathrm{CIB}}}(\chi)/L^2$ and $W^{\kappa_{\mathrm{CIB}}}$ as defined in Sec. III.

In addition to the term (33), secondary contractions appear in Eq. (32). To understand the origin of such terms, we once again separate the CIB emission into high- and low-z components. Schematically, the auto-correlation is given by $\langle \hat{\phi} \hat{\phi} \rangle \sim \langle (TT)(TT) \rangle$. Expanding the lensed T, we get "primary contractions" of the form $\langle (T_g T_g)(T_{\mathrm{high}-z} T_{\mathrm{high}-z}) \rangle$, which correspond to the ones in Eq. (32), since $\langle (T_{\mathrm{high}-z} T_{\mathrm{high}-z}) \rangle \sim \phi$. In addition, we also get "secondary contractions" of the form $\langle (T_g T_{\mathrm{high}-z})(T_g T_{\mathrm{high}-z}) \rangle$, which are the same order in the expansion as the primary contractions but more involved to calculate, since the integrals do not separate as in Eqs. (30) and (31). Similar contractions also appear in CMB lensing reconstruction (e.g., in the $N^{(1)}$ bias), and are highly suppressed compared to the primary contractions. However, this suppression is not guaranteed in our case. For example, secondary contractions of foreground contamination of CMB lensing have been estimated to be of the same order of magnitude as the primary ones [29,56].

C. Mitigation strategies

Controlling the biases presented above is crucial, in order to provide a convincing measurement of the lensing auto- or cross-correlation. In this section, we propose several mitigation approaches.

1. Cross-correlation

First, in cross-correlation with low-z tracers, the $N_L^{(0)}$ and $N_L^{(1)}$ terms only act as extra noise but not as a bias. In practice, this noise can usually be evaluated from the data themselves. As a result, the most concerning bias in cross-correlation is the self-lensing bispectrum. We propose four methods to quantify or reduce this self-lensing bispectrum.

Emission template: For a given tracer sample (e.g., galaxy catalog), we can build a template of the emission from the tracer sample at CIB frequencies T_g. Because the bispectrum $B_{\delta T_g T_g}$ sourcing the self-lensing bias involves two powers of T_g, having the correct galaxy flux fluctuations is important. The optical or infrared brightness of the galaxies can inform us about their relative brightness at CIB frequency. The overall normalization of the template map is

obtained by cross-correlating the template with the CIB map. We can then apply the quadratic estimator to this template map and cross-correlate the result with the tracer map to estimate the self-lensing bias. Alternatively, the template can be subtracted directly from the CIB map before applying the lensing quadratic estimator.

Source hardening: Techniques similar to "source hardening" [29] or the related "bias hardening" [58] used in CMB lensing are applicable to CIB lensing. If we assume a known shape for the source bispectrum $B_{\delta T_g T_g}$, an optimal estimator for its amplitude can be derived [29,59]. Given an amplitude for the source bispectrum, it can be propagated to the bias to CMB lensing by evaluating Eq. (30). The source estimator and the lensing estimator are not orthogonal but can be combined into unbiased estimates of the lensing potential and amplitude of sources, by inverting a response matrix as in Ref. [29].

Separate shear and dilation estimation: As discussed in the Introduction (and Ref. [16]), the quadratic estimator on large scales combines estimates of shear and dilation to reconstruct the lensing potential. It is, however, possible to split it into shear-only and dilation-only estimators [16,39]. While the statistical power of the shear-only and dilation-only estimators is approximately equal, it is reasonable to assume that point-source contamination, such as in self-lensing, should mostly affect the dilation estimator but not the shear estimator. We leave investigating this conjecture for future work and note that this could be a very effective technique to isolate contamination in both CIB and CMB lensing.

Using disjoint sets of multipoles: In this case, as for self-lensing, lensing reconstruction and cross-correlation using disjoint ℓ can help. For example, we could reconstruct $\hat{\phi}_L$ only using CIB multipoles $\ell > \ell_{\text{cut}}$ and then perform the cross-correlation only for $L < \ell_{\text{cut}}$. This ensures that the map used in the reconstruction $T(\ell > \ell_{\text{cut}})$ is uncorrelated with the tracer map on the scales used for the cross-correlation. This eliminates the equilateral self-lensing bispectrum configuration and can thus limit the impact of contaminants that are common to the tracers and CIB map, e.g., residual Galactic dust.

2. Auto-correlation

Similar mitigation techniques apply for the auto-power spectrum; we summarize these below. In auto-correlation, $N^{(0)}$ and $N^{(1)}$ constitute a bias, contrary to the cross-correlation case. These biases are usually computed with simulations and subtracted from the $\hat{\phi}$ power spectrum. However, this requires correctly simulating the intrinsic trispectrum of the CIB to much better precision than the lensing-induced trispectrum. This can be challenging, since the intrinsic trispectrum is not well constrained, and even though it can be predicted in the halo model, achieving the required accuracy may not be straightforward. For this reason, auto-correlation analyses may be

more challenging. Most mitigation strategies presented for the cross-correlation also apply in auto-correlation. Here, we focus on the differences.

Making an *emission template* from low-redshift galaxies is likely to be more challenging in auto-correlation than in cross-correlation; one would need a flux-weighted template of every source contributing to the low-z part of the CIB, not only the sources from a given tracer catalog. To the extent that the emission traces matter and that the redshift-dependent IR luminosity function is known, a combination of galaxy surveys covering most of the CIB redshift distribution might provide the required template.

Source hardening seems a more promising avenue. In this case, we must assume a known shape for the intrinsic source bispectrum and trispectrum and measure their amplitudes with appropriate estimators [29]. This procedure is complicated by the presence of secondary contractions.

Disjoint ℓ reconstruction can be used in the auto-correlation as well. This method can be generalized and made more optimal by partitioning Fourier space into annuli and evaluating the ϕ power spectrum only using the cross-correlation between estimates reconstructed from modes belonging to different annuli [60].

Finally, we note that part of the non-Gaussianity of the CIB is due to nonlinear gravitational evolution. In particular, nonlinear coupling between long and short wavelength modes produces a modulation in the small-scale power spectrum that is proportional to the long wavelength mode. The part of the trispectrum generated by this coupling can be undone by *tidal reconstruction* [61].

VIII. BIAS FROM CIB LENSING TO CMB LENSING RECONSTRUCTION

In this section, we switch focus and consider CIB lensing not as the signal of interest but as a contaminant to CMB lensing. CMB lensing reconstruction typically operates on a cleaned temperature map. Imperfect subtraction of the CIB (or any other foreground) biases the auto- and cross-spectra of the reconstructed CMB lensing potential. One factor that is often neglected is that some of the foregrounds lie at cosmological distance from us and are therefore lensed. In this section, we present the extra bias to CMB lensing due to residual lensed CIB.

Suppose that the cleaned CMB map T contains not only the true lensed CMB T_{CMB} but also some lensed CIB T_{CIB}. This lensed CIB is either the full CIB emission or the reduced CIB residual if some foreground reduction is applied,

$$\begin{aligned} T &= T_{\text{CMB}} + T_{\text{CIB}} \\ &= T_{\text{CMB}}^0 - \boldsymbol{\nabla}\phi_{\text{CMB}} \cdot \boldsymbol{\nabla}T_{\text{CMB}}^0 + T_{\text{CIB}}^0 \\ &\quad - \boldsymbol{\nabla}\phi_{\text{CIB}} \cdot \boldsymbol{\nabla}T_{\text{CIB}}^0 + \cdots \end{aligned} \tag{34}$$

We carefully distinguish the CMB lensing potential ϕ_{CMB}, which receives contributions from redshift zero to the surface of last scattering, from the CIB lensing potential ϕ_{CIB}, which is only sourced by lower-redshift structures. As shown in Refs. [28,29], neglecting the T_{CIB}^0 contribution results in a biased estimation of ϕ_{CMB}. As before, we treat the auto- and cross-correlations separately.

A. Cross-correlation

Suppose we ignore the presence of CIB contamination and apply the quadratic estimator to T, to get an estimate of $\hat{\phi}_{\mathrm{CMB}} = \mathcal{Q}_{\mathrm{CMB}}[T]$, where the subscript "CMB" on the quadratic estimator $\mathcal{Q}$ denotes that the weights used in the estimator are those appropriate to the CMB lensing reconstruction. The cross-correlation with tracers δ_g then contains terms of the form $\langle \mathcal{Q}_{\mathrm{CMB}}[T_{\mathrm{CIB}}]\delta_g \rangle$. When replacing the lensed T_{CIB} by the unlensed T_{CIB}^0, we get bispectrum terms $\langle \mathcal{Q}_{\mathrm{CMB}}[T_{\mathrm{CIB}}^0]\delta_g \rangle \propto B_{T_g T_g \delta_g}$, which have been discussed in previous work. But we also note that $\mathcal{Q}_{\mathrm{CMB}}[T_{\mathrm{CIB}}] - \mathcal{Q}_{\mathrm{CMB}}[T_{\mathrm{CIB}}^0]$ is effectively a reconstruction of ϕ_{CIB}. This reconstruction is both suboptimal and biased because the CMB rather than the CIB weights have been used.[7] Nonetheless, if we write $\langle \mathcal{Q}_{\mathrm{CMB}}[T_{\mathrm{CIB}}](L)\rangle = R(L)\phi_{\mathrm{CIB}}(L)$ in terms of a response function R, then the cross-correlation with tracers is further biased by a factor $R(L)\langle \phi_{\mathrm{CIB}}\delta_g \rangle$.

B. Auto-correlation

Similarly, the auto-correlation will be enhanced by terms of the form $R(L)\langle \phi_{\mathrm{CIB}}\phi_{\mathrm{CMB}}\rangle(L)$, $R^2(L)\langle \phi_{\mathrm{CIB}}\phi_{\mathrm{CIB}}\rangle(L)$, as well as the trispectrum term $\mathcal{T}_L$ of the form of Eq. (26), but with the CIB lensing weights replaced with the CMB lensing ones. Quantifying this effect is an interesting and important problem, but it is outside the scope of this paper.

IX. CONCLUSIONS

Intensity mapping experiments such as 21 cm, Hα, Lyman-α, carbon monoxyde (CO), or ionized carbon (CII) [62] will probe large volumes of the Universe. Lensing reconstruction from these intensity maps has the potential to become a powerful cosmological tool, by enabling measurements of the amplitude of structure at intermediate redshifts between galaxy surveys and the CMB. This constitutes a new regime for lens reconstruction. While the source field is continuous with small fluctuations (like in CMB lensing), it is also generically non-Gaussian, and a full statistical prior is not readily available (like in galaxy lensing). Furthermore, intensity maps may contain some redshift information, allowing the full three-dimensional gravitational potential to be estimated.

Because of smooth foregrounds in frequency space (e.g., due to Galactic dust in 21 cm [63] or continuum fitting in the Lyman-α forest), wave vectors with low radial components may be unusable, which complicates the lens reconstruction.

In this paper, we have addressed the issue of the non-Gaussianity of the source field. We have generalized the quadratic estimator formalism commonly used in CMB lensing to any weakly non-Gaussian source field by deriving the optimal weights. We have computed the effect of non-Gaussianity on the lensing noise and shown that the relevant measure of non-Gaussianity is the ratio of the trispectrum to the squared power spectrum, divided by the number of modes.

Beyond the statistical error, the non-Gaussianity of the unlensed map also produces systematic biases in the lens reconstruction, both in auto- and cross-correlation. We have explored in detail the non-Gaussian noise biases, the effect of foregrounds, and self-lensing due to the lensing of high-redshift sources by the halos low-redshift sources. We have proposed various means to reduce the non-Gaussianity of the source (e.g., by masking the rare bright point sources), to avoid it (e.g., using disjoint multipoles or shear/dilation estimators), and to subtract it (e.g., using emission templates or source/bias hardening). Overall, these biases appear easier to mitigate in cross-correlation than in auto-correlation.

Since high-quality CIB intensity maps already exist, we have applied our formalism to the case of CIB lensing. We forecast that CIB lensing is detectable with current Planck data and will be measured at high precision in future submillimeter experiments like CCAT-Prime. Because the CIB is at lower redshift than most proposed intensity mapping surveys, the non-Gaussianity is expected to be larger. Furthermore, the redshift distribution of the CIB is complex (dependent on the multipole considered), somewhat uncertain, and extremely wide, extending all the way to redshift zero. This produces a larger self-lensing bias, compared to line intensity maps where the redshifts are known precisely. CIB lensing is therefore a very useful stepping stone toward lensing reconstruction of intensity mapping experiments.

CIB lensing will also provide a measurement of the amplitude of structure at intermediate redshifts, although its cosmological interpretation may at first be affected by uncertainties in the CIB redshift distribution. Turning this around, CIB lensing can be used to constrain the redshift distribution of the CIB, by measuring distance ratios (sometimes referred to as cosmography) between CIB lensing and CMB or galaxy lensing or by assuming a known cosmology. In other words, measuring CIB lensing can provide useful insight into star-formation history.

Finally, because residual CIB will be present in cleaned CMB maps, the lensing of the CIB will be a bias in CMB lensing. We have described these biases but leave their evaluation to future work.

---

[7]The optimal and unbiased reconstruction of ϕ_{CIB} would have been $\mathcal{Q}_{\mathrm{CIB}}[T_{\mathrm{CIB}}]$.

APPENDIX A: CIB HALO MODEL

1. CIB monopole and power spectrum

We follow the halo model implementation from Refs. [20,43]. Consider a galaxy with luminosity L_ν, i.e., such that the power emitted in frequency band $d\nu$ is $L_\nu d\nu$ and the observed flux per unit frequency is

$$S_\nu = \frac{aL_{\nu/a}}{4\pi\chi^2}\,[\text{power/area/frequency}], \qquad (A1)$$

where a is the scale factor and χ is the comoving distance. Note this equation does not involve the usual luminosity distance because we are not considering the bolometric flux but a band-limited flux. The observed CIB specific surface brightness I_ν is the sum of the fluxes from all galaxies within an observed solid angle $d\Omega$:

$$I_\nu = \int \frac{dN_{\text{gal}}}{d\Omega} S_\nu \,[\text{power/area/solid angle/frequency}].$$

$$(A2)$$

Introducing the comoving volume $dV = d\Omega\chi^2 d\chi$ and the galaxy number density contrast $\delta_{\text{gal}}(S_\nu)$, this can be expressed as

$$\delta I_\nu = \int d\chi\chi^2 \int dS_\nu \frac{d\bar{N}_{\text{gal}}}{dS_\nu dV} S_\nu \delta_{\text{gal}}(S_\nu). \qquad (A3)$$

The flux distribution functions of galaxies $d\bar{N}_{\text{gal}}/dS_\nu dV$ for the Planck frequency bands are taken from the "two star-forming modes" model in Ref. [40]. We apply the flux cuts relevant for Planck (see Table 1 in Ref. [21]) and for CCAT-Prime. The model we implement makes the simplifying but unphysical assumption that all galaxies cluster identically, regardless of their luminosity. With this assumption, the equation above becomes

$$\delta I_\nu = \int d\chi\chi^2 \bar{j}_\nu \delta_{\text{gal}}, \qquad (A4)$$

where we have introduced

$$\bar{j}_\nu = \int dS_\nu \frac{d\bar{N}_{\text{gal}}}{dS_\nu dV} S_\nu. \qquad (A5)$$

Note that our definition differs by a factor a/χ^2 from other sources in the literature.

Hence, the CIB intensity power spectrum

$$C_\ell^{I_\nu I_{\nu'}} = \int d\chi\chi^2 \bar{j}_\nu \bar{j}_{\nu'} P_{\text{gal}}^{3d}\left(k = \frac{\ell + 1/2}{\chi}, z\right). \qquad (A6)$$

The power spectrum of galaxies is then evaluated from a halo model, which produces a 2h-term, 1h-term (or halo shot noise) and a 1-galaxy term (or galaxy shot noise). The clustering of galaxies (2-halo and 1-halo terms) is computed using a HOD from Ref. [47],

$$\begin{cases} P_{\text{gal}}^{1h}(k) = I_{02}(k) \\ P_{\text{gal}}^{2h}(k) = I_{11}^2(k) P^{\text{lin}}(k), \end{cases} \qquad (A7)$$

where

$$I_{ij}(k_1,...,k_j) \equiv \int dm \frac{dn}{dm} b_i \left(\frac{N_{\text{gal}}(m)}{\bar{n}_{\text{gal}}}\right)^j u(k_1)...u(k_j).$$

$$(A8)$$

Here, $N_{\text{gal}}(m)$ and $\bar{n}_{\text{gal}}$ define the HOD, and $u(k)$ is the normalized Fourier transform of the halo profile [47]. Following Ref. [20], we vary the HOD depending on the frequency band observed. The shot-noise power spectrum is computed as

$$dC_\ell^{\text{gal shot}} = \frac{dN}{d\Omega} S_\nu,$$

$$\text{so} C_\ell^{\text{gal shot}} = \int d\chi\chi^2 \underbrace{\int dS_\nu \frac{d\bar{N}_{\text{gal}}}{dS_\nu dV} S_\nu^2}_{\equiv \bar{j}_{\nu,2}}. \qquad (A9)$$

Finally, our model for the CIB power spectrum becomes

$$C_\ell^{I_\nu I_{\nu'}} = \int d\chi\chi^2 [\bar{j}_\nu \bar{j}_{\nu'} P_{\text{gal}}^{1h+2h} + \bar{j}_{\nu,2}]. \qquad (A10)$$

Several limitations should be noted. First, the CIB emission is assumed independent of galaxy mass. Second, for this approach to be consistent, there has to exist a mass-flux

relation such that the HOD in Ref. [47] produces the flux counts of Ref. [43], which is not guaranteed. Finally, for the mass integrals to converge, one needs to integrate to very low halo masses, $\simeq 10^{10}\,M_\odot$, where the HOD may be uncertain.

2. CIB trispectrum

The power spectrum of the reconstructed lensing potential is a four-point function of the temperature. Therefore, the unlensed CIB trispectrum contributes to the lensing power spectrum. Similarly to the case of the power spectrum, the halo model relates the CIB trispectrum to the galaxy trispectrum and eventually to the matter power spectrum, bispectrum, and trispectrum:

$$\mathcal{T}^{\mathrm{CIB}}_{\ell_1,\ell_2,\ell_3,\ell_4} = \int d\chi \chi^2 [\bar{j}_\nu^4 \mathcal{T}^{1h+2h+3h+4h}_{\mathrm{gal}} + \bar{j}_{\nu,4}], \quad (A11)$$

with

$$\bar{j}_{\nu,4} = \int dS_\nu \frac{d\bar{N}_{\mathrm{gal}}}{dS_\nu dV} S_\nu^4. \quad (A12)$$

As shown in Fig. 5, our halo model trispectrum is very sensitive to the minimum redshift (or distance) in the integral above. In order to evaluate Eq (A11), we need to compute the galaxy trispectrum $\mathcal{T}_{\mathrm{gal}}$ in the halo model [64],

$$
\begin{cases}
\mathcal{T}^{1h}_{\mathrm{gal}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) = I_{04}(k_1,k_2,k_3,k_4) \\
\mathcal{T}^{2h}_{\mathrm{gal}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) = P_m(k_{12})I_{12}(k_1,k_2)I_{12}(k_3,k_4) + 2\ \mathrm{perm.} + P_m(k_1)I_{11}(k_1)I_{13}(k_2,k_3,k_4) + 3\ \mathrm{perm.} \\
\mathcal{T}^{3h}_{\mathrm{gal}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) = B_m(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_{34})I_{11}(k_1)I_{11}(k_2)I_{12}(k_3,k_4) + 5\ \mathrm{perm.} \\
\mathcal{T}^{4h}_{\mathrm{gal}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4) = T_m(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4)I_{11}(k_1)I_{11}(k_2)I_{11}(k_3)I_{11}(k_4),
\end{cases}
\quad (A13)
$$

where P_m, B_m, and T_m are the matter power spectrum, bispectrum, and trispectrum, respectively.

When applying the quadratic estimator, the low lensing multipoles L are best reconstructed, and this reconstruction relies mostly on the high map multipoles ℓ. The trispectrum configuration of interest is therefore "almost collapsed," $\mathcal{T}^{\mathrm{CIB}}_{\ell,L-\ell,\ell',-L-\ell'}$, with $L \ll \ell$. This is convenient for two reasons. First, the almost-collapsed trispectrum can be measured from the map without having to implement a general trispectrum estimator (see Appendix E). Second, in this limit, the halo model expressions simplify. Indeed, for the galaxy number density or any general three-dimensional (3D) field, there exists a consistency relation between the almost-collapsed trispectrum, the exactly collapsed trispectrum $\mathcal{T}^{\mathrm{gal}}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}')$, and the response of the power spectrum to a matter overdensity:

$$
\underbrace{\mathcal{T}^{\mathrm{gal}}(\mathbf{k},\mathbf{K}-\mathbf{k},\mathbf{k}',-\mathbf{K}-\mathbf{k}')}_{\text{almost-collapsed}}
$$

$$
\simeq \underbrace{\mathcal{T}^{\mathrm{gal}}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}')}_{\text{exactly-collapsed}} + \underbrace{\frac{\partial P^{\mathrm{gal}}(k)}{\partial \delta_m}\frac{\partial P^{\mathrm{gal}}(k)}{\partial \delta_m}P^{\mathrm{lin}}(K)}_{\text{power spectrum response}}.
$$

$$(A14)$$

This relation was derived in the context of the covariance of measured power spectra, and the second term on the right-hand side is then interpreted as the supersample variance [64–66]. In Ref. [64], it was shown that the halo model satisfies this consistency relation. This is convenient because it means that we only need to evaluate the halo model trispectrum in the exactly collapsed limit.

We follow Ref. [64] for the expressions of the halo model trispectrum. The shot-noise and 1-halo trispectra are easy to evaluate, depending only on the mass function, HOD, and halo profile. The 2-halo further depends on the matter power spectrum and is also straightforward to compute. In the exactly collapsed limit, the 3-halo term $\mathcal{T}^{3h}$ cancels because the matter bispectrum is zero in the collapsed configuration. The 4-halo trispectrum is related to the matter trispectrum.

For the purpose of lens reconstruction, we focus on the collapsed trispectrum $\langle\mathcal{T}^{\mathrm{gal}}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}')\rangle$, averaged over the angle between $\mathbf{k}$ and $\mathbf{k}'$, and with $k'=k$. Indeed, we mostly reconstruct the low-L lensing modes and from only high-ℓ temperature modes, which justifies the collapsed limit. Furthermore, as can be seen from Fig. 7, the range of multipoles which contribute to the (Gaussian) signal-to-noise ratio is quite narrow, which justifies considering $k'=k$. Finally, the power spectrum of the reconstructed lensing at fixed L is effectively an average over the angle between ℓ and ℓ'. It is true that this average involves some angular weights, but this is not expected to change the result by more than a factor of a few, which is sufficient here. In this very specific limit, the matter trispectrum takes a simple form, $\langle\mathcal{T}_{\mathrm{m}}(\mathbf{k},-\mathbf{k},\mathbf{k}',-\mathbf{k}')\rangle \simeq (232/441)P_m(k)^3$ [67], and the halo model terms become

$$\begin{cases} \langle \mathcal{T}_{\text{gal}}^{1h}(\mathbf{k}, -\mathbf{k}, \mathbf{k}', -\mathbf{k}') \rangle = I_{04}(k, k, k, k) \\ \langle \mathcal{T}_{\text{gal}}^{2h}(\mathbf{k}, -\mathbf{k}, \mathbf{k}', -\mathbf{k}') \rangle = 4P_m(k)I_{11}(k)I_{13}(k, k, k) + 2P_m(k)I_{12}(k, k)^2 \\ \langle \mathcal{T}_{\text{gal}}^{3h}(\mathbf{k}, -\mathbf{k}, \mathbf{k}', -\mathbf{k}') \rangle = 0 \\ \langle \mathcal{T}_{\text{gal}}^{4h}(\mathbf{k}, -\mathbf{k}, \mathbf{k}', -\mathbf{k}') \rangle = \frac{232}{441} P_m(k)^3 I_{11}(k)^4. \end{cases} \tag{A15}$$

Finally, we follow Ref. [65] for the expression of the power spectrum response $\frac{\partial P^{\text{gal}}(k)}{\partial \delta_m}$ in the halo model.

APPENDIX B: DERIVING THE LENSING KERNEL FOR THE CIB LENSING QUADRATIC ESTIMATOR

In this Appendix, we derive the lensing efficiency kernel for the CIB lensing quadratic estimator. In other words, we wish to derive the redshift range probed by the CIB lensing estimator.

In the case of CMB lensing, this question is trivial, since the CMB is sourced at a single redshift of 1100. In the case of galaxy lensing, there is a well-defined number density of galaxies dn/dz. The shear estimator is linear in the individual galaxy shapes. If the signal-to-noise ratio on each galaxy shape were the same, regardless of galaxy redshift, then the observed shear would be simply the average of the shears defined at each galaxy redshift, weighted by dn/dz.

In the case of CIB lensing, this question is less straightforward. Indeed, depending on the multipole, the unlensed CIB is sourced by different redshift ranges: the 2-halo, 1-halo, and 1-galaxy terms in the CIB power spectrum have different redshift distributions (see Fig. 12 in Ref. [68]). The quadratic estimator mixes the CIB multipoles to reconstruct the lensing potential, in a way that depends on the multipole of lensing to be reconstructed. Thus, different multipoles of the CIB lensing quadratic estimator probe different redshifts.

We decompose the unlensed CIB emissions into statistically independent redshift slices:

$$T_{\ell}^0 = \int dz T_{\ell}^z, \quad \text{with} \quad \langle T_{\ell}^z T_{\ell'}^{z'} \rangle \propto \delta_{z-z'}^D. \tag{B1}$$

As a result, the unlensed CIB power spectrum can be decomposed as

$$C_{\ell}^0 = \int dz \frac{dC_{\ell}^0}{dz},$$

$$\text{where } \langle T_{\ell}^z T_{\ell'}^{z'} \rangle = \delta_{z-z'}^D (2\pi)^2 \delta_{\ell+\ell'}^D \frac{dC_{\ell}^0}{dz}. \tag{B2}$$

We then define ϕ^z to be the lensing potential for a source at redshift z. The unlensed CIB emission at redshift z is lensed by the potential ϕ^z. Thus,

$$\langle T_{\ell}^z T_{\boldsymbol{L}-\ell}^z \rangle = \phi_{\boldsymbol{L}}^z f_{\ell,\boldsymbol{L}-\ell}^z + \mathcal{O}(\phi^2 C^0), \tag{B3}$$

where

$$f_{\ell,\boldsymbol{L}-\ell}^z = -\boldsymbol{L} \cdot \left[\boldsymbol{\ell} \frac{dC_{\ell}^0}{dz} + (\boldsymbol{L} - \boldsymbol{\ell}) \frac{dC_{\boldsymbol{L}-\ell}^0}{dz} \right]. \tag{B4}$$

In particular,

$$f_{\ell,\boldsymbol{L}-\ell} = \int dz f_{\ell,\boldsymbol{L}-\ell}^z. \tag{B5}$$

As a result, the basic quadratic estimator measures

$$\langle \hat{\phi}_{\boldsymbol{L},\ell} \rangle = \int dz \phi_{\boldsymbol{L}}^z \frac{f_{\ell,\boldsymbol{L}-\ell}^z}{f_{\ell,\boldsymbol{L}-\ell}}. \tag{B6}$$

We can finally express the quadratic estimator as a weighted sum of the lensing potential over the various redshifts slices, with the weights

$$\langle \hat{\phi}_{\boldsymbol{L}} \rangle \propto \int dz \phi_{\boldsymbol{L}}^z \int \frac{d^2\ell}{(2\pi)^2} \frac{f_{\ell,\boldsymbol{L}-\ell}^z}{f_{\ell,\boldsymbol{L}-\ell}} \frac{1}{\sigma_{\boldsymbol{L},\ell}^2}. \tag{B7}$$

In conclusion, the measured CIB lensing convergence can be written as

$$\kappa_{\text{CIB}\boldsymbol{L}} = \int d\chi W^{\kappa_{\text{CIB}}}(\chi, \boldsymbol{L}) \delta_m(\vec{k} = \boldsymbol{L}/\chi, \chi), \tag{B8}$$

where

$$W^{\kappa_{\text{CIB}}}(\chi, \boldsymbol{L}) = \int d\chi_S W^{\text{CIB}}(\chi_S, \boldsymbol{L}) W^{\kappa}(\chi, \chi_S), \tag{B9}$$

$$W^{\kappa}(\chi, \chi_S) = \frac{3}{2} \left(\frac{H_0}{c}\right)^2 \Omega_m^0 \frac{\chi}{a(\chi)} (1 - \chi/\chi_S), \tag{B10}$$

and

$$W^{\mathrm{CIB}}(\chi_S, \boldsymbol{L}) \propto \frac{dz}{d\chi_s} \int \frac{d^2\ell}{(2\pi)^2} \frac{f_{\ell,\boldsymbol{L}-\ell}^z}{f_{\ell,\boldsymbol{L}-\ell}} \frac{1}{\sigma_{L,\ell}^2},$$

$$\text{normalized by } \int d\chi_S W^{\mathrm{CIB}}(\chi_S, \boldsymbol{L}) = 1.$$

$$(B11)$$

A few comments are in order. First, the lensing kernel $W^{\kappa_{\mathrm{CIB}}}(\chi, \boldsymbol{L})$ depends on the multipole $\boldsymbol{L}$ considered. This is not the case in CMB lensing, due to fact that the unlensed CMB comes from a single redshift. This is also different from the galaxy lensing case.

Second, the lensing kernel for the CIB quadratic estimator is determined by the redshift distribution of the unlensed CIB power spectrum $\frac{dC_\ell^0}{dz}$ as well as the noise power spectrum of the CIB map. Intuitively, the minimum-variance quadratic estimator weighs the various CIB multipoles differently depending on their signal-to-noise ratio. Since different redshifts contribute differently to the various CIB multipoles, this weighting will upweight or downweight certain redshift ranges.

APPENDIX C: NON-GAUSSIANITY AND POWER SPECTRUM SIGNAL-TO-NOISE RATIO

Before looking at the case of lensing, for a first step, we compute the angular power spectrum covariance in both the Gaussian and Poisson cases. For multipoles above $\ell \sim 1000$, the CIB is dominated by the 1-halo or 1-galaxy terms. In this regime, the statistics of the field are close to a uniform Poissonian sampling, rather than a Gaussian random field. Below, we compute and interpret intuitively the reduction in signal-to-noise ratio due to the Poisson sampling. This intuition is useful for understanding the effect of non-Gaussianity on the lensing signal-to-noise ratio, as described in the next section.

1. Gaussian random field

For a Gaussian random field, we know that the various ℓ-modes are independent. This leads to the well-known formula for the power spectrum covariance,

$$\text{cov}[C_\ell, C_{\ell'}] = \delta_{\ell,\ell'}^K \frac{2C_\ell^2}{N_{\mathrm{modes}\ell}}, \qquad (C1)$$

where $N_{\mathrm{modes}\ell} \approx 2f_{\mathrm{sky}}\ell\Delta\ell$ is the number of observed Fourier modes in a bin centered at ℓ and of width $\Delta\ell$. As a result, the signal-to-noise ratio for the power spectrum is

$$\text{SNR}_{\mathrm{Gaussian}}(C_\ell) = \sqrt{\frac{N_{\mathrm{modes}\ell}}{2}}. \qquad (C2)$$

Intuitively, each mode contributes a signal-to-noise ratio of $1/\sqrt{2}$, and the modes are all independent, so $\text{SNR} \propto \sqrt{N_{\mathrm{modes}\ell}}$. As expected, the signal-to-noise ratio

becomes infinite in the limit of infinite number of modes; the only limit to the signal-to-noise is the number of modes.

2. Non-Gaussian random field

For a non-Gaussian random field, the situation is different. The various multipoles are no longer independent, which translates into correlations across power spectrum bins,

$$\text{cov}[C_\ell, C_{\ell'}] = \delta_{\ell,\ell'}^K \frac{2C_\ell^2}{N_{\mathrm{modes}\ell}} + \frac{\mathcal{T}_{\ell,-\ell,\ell',-\ell'}}{V}, \quad (C3)$$

where V is the 3D volume or 2D area of the survey. The trispectrum of the field thus adds a nondiagonal covariance term. As a result, the signal-to-noise is reduced compared to the Gaussian case. We explore this in more detail in the special case of the uniform Poisson sampling in the following.

3. Uniform Poisson sampling

A uniform Poisson sampling describes a collection of unclustered point sources well. In this case, the power spectrum and trispectrum take simple forms that are independent of multipole,

$$\begin{cases} C_\ell = \frac{\langle s^2 \rangle}{\bar{n}}, \\ \mathcal{T}_{\ell,-\ell,\ell',-\ell'} = \frac{\langle s^4 \rangle}{\bar{n}^3}, \end{cases} \qquad (C4)$$

where each object has a flux s and the mean number density of objects is $\bar{n}$.

In this simple case, the power spectrum is entirely determined by these two numbers. The most natural way of extracting this information would be to estimate $\bar{n}$ and $\langle s^2 \rangle$ directly, without computing the power spectrum. One could simply count the number of objects. Intuitively, the signal-to-noise ratio on this measurement should scale as $\sqrt{N_{\mathrm{gal}}} = \sqrt{\bar{n}V}$. One would then estimate $\langle s^2 \rangle$ directly from the galaxy catalog. The signal-to-noise ratio on the power spectrum will be determined by the amount of object-to-object fluctuations in the flux; i.e., it will scale as $\langle s^2 \rangle / \sqrt{\langle s^4 \rangle}$.

While measuring the power spectrum is a good way to extract the information for the Gaussian case, it is a clumsy approach for the Poisson case. First, we can see that the various C_ℓ bins are correlated due to the trispectrum term. However, they are not 100% correlated, so extracting all the information requires measuring all the available multipoles. Indeed, the non-Gaussian formula for the signal-to-noise ratio simplifies to

$$\text{SNR}_{\text{Poisson}}(C_\ell) = \sqrt{\frac{N_{\text{modes}\ell}}{2}} \frac{1}{\sqrt{1 + \frac{N_{\text{modes}\ell}\mathcal{T}}{2C^2 V}}}$$

$$= \sqrt{\frac{N_{\text{modes}\ell}}{2}} \frac{1}{\sqrt{1 + \frac{N_{\text{modes}\ell}}{2N_{\text{gal}}} \frac{\langle s^4 \rangle}{\langle s^2 \rangle^2}}}. \quad (C5)$$

In particular, in the limit of infinite number of modes, the signal-to-noise ratio is not infinite but asymptotes to

$$\text{SNR}_{\text{Poisson}}(C_\ell, N_{\text{modes}\ell} \to \infty) = \sqrt{N_{\text{gal}}} \frac{\langle s^2 \rangle}{\sqrt{\langle s^4 \rangle}}. \quad (C6)$$

We thus obtain the expression expected for the information content of a Poisson sampling.

APPENDIX D: NON-GAUSSIANITY AND LENSING SIGNAL-TO-NOISE

We go a step further and compute the total noise on the lens reconstruction when the unlensed source image is a uniform Poisson sampling. In this simple limiting case, the power spectrum and trispectrum are independent of the multipole. This allows us to evaluate Eqs. (22) and (24) and gain some intuition for the magnitude of the non-Gaussian lensing noise.

1. Non-Gaussian weights

To evaluate Eq. (22), we need to be careful about double counting modes. Noting that $\hat{\phi}_{L,\ell} = \hat{\phi}_{L,L-\ell}$, we can replace the covariance matrix by

$$(\Sigma_L)_{\ell,\ell'} = \frac{1}{f_{\ell,L-\ell}f_{\ell',L-\ell'}}$$
$$\times [2(2\pi)^2 \delta^D_{\ell'-\ell} C^{\text{total}}_\ell C^{\text{total}}_{L-\ell} + \mathcal{T}^0_{\ell,-\ell',L-\ell,\ell'-L}]. \quad (D1)$$

If power spectrum and trispectrum are independent of ℓ, and $C^{\text{total}} \simeq C^0$, the Sherman-Morrison formula allows one to invert the covariance matrix as

$$(\Sigma_L)^{-1}_{\ell,\ell'} = \frac{f_{\ell,L-\ell}f_{\ell',L-\ell'}}{2(C^0)^2}$$
$$\times \left[(2\pi)^2 \delta^D_{\ell'-\ell} - \frac{\mathcal{T}^0}{2(C^0)^2 + \mathcal{T}^0 (\int \frac{d^2\ell}{(2\pi)^2})} \right], \quad (D2)$$

which we then integrate:

$$\int \frac{d^2\ell d^2\ell'}{(2\pi)^4} (\Sigma_L)^{-1}_{\ell,\ell'} = L^4 \frac{(\int \frac{d^2\ell}{(2\pi)^2})(C^0)^2}{2(C^0)^2 + (\int \frac{d^2\ell}{(2\pi)^2})\mathcal{T}^0}. \quad (D3)$$

Therefore, the reconstruction noise on ϕ [Eq. (22)] becomes

$$N_L = \left[\int \frac{d^2\ell d^2\ell'}{(2\pi)^4} (\Sigma_L)^{-1}_{\ell,\ell'} \right]^{-1}$$
$$= \underbrace{\frac{2}{L^4(\int \frac{d^2\ell}{(2\pi)^2})}}_{\text{Gaussian}N_L} + \underbrace{\frac{\mathcal{T}^0}{L^4(C^0)^2}}_{\text{non-Gaussian correction}}, \quad (D4)$$

or in terms of κ,

$$N_L^\kappa = \underbrace{\frac{2\pi}{N_{\text{modes}}}}_{\text{Gaussian noise}} + \underbrace{\frac{\mathcal{T}^0}{4(C^0)^2}}_{\text{non-Gaussian correction}}, \quad (D5)$$

where $N_{\text{modes}} \equiv 4\pi \int \frac{d^2\ell}{(2\pi)^2} = \ell^2_{\text{max}} - \ell^2_{\text{min}}$ and ℓ_{min} and ℓ_{max} define the range of observed multipoles used for the lens reconstruction. The first term recovers the noise relevant for a Gaussian (instead of Poisson) white noise. This term goes to zero as the number of observed multipoles increases; if the modes are Gaussian, each of them adds independent lensing information, and the number of modes is the only limit to the reconstruction precision. The second term is due to the non-Gaussian nature of the unlensed map. Interestingly, it is independent of the number of observed multipoles.

In the limit of an infinite number of observed multipoles, the reconstruction noise thus asymptotes to $N_L \to \frac{\mathcal{T}^0}{L^4 C^2}$. This is the noise on the reconstructed lensing potential ϕ. The corresponding noise on the lensing convergence κ is

$$N_L^\kappa = \frac{\mathcal{T}^0}{4C^2} = \frac{1}{4} \frac{\langle s^4 \rangle}{\langle s^2 \rangle^2} \frac{1}{\bar{n}}. \quad (D6)$$

This shows that the noise on the reconstructed $\hat{\phi}_L$ is independent of the multipole L. This is the same L-dependence as the shape noise of galaxy shear. As stated above, because the unlensed power spectrum is white (C_ℓ is independent of ℓ), shear has no effect on the map, and all the lensing information is reconstructed from magnification. Hence, this noise does not originate from fluctuations in the shapes of galaxies but rather in fluctuations in the local number density and brightness of galaxies. Locally, these fluctuations are degenerate with a local magnification and are therefore a source of noise for the lens reconstruction.

A more intuitive derivation is the following. Schematically, the observed surface brightness map can be described as ns, where n is the number count of galaxies and s is the galaxy brightness. Lensing conserves the surface brightness. Indeed, if the convergence is nonzero, the observed number counts are diluted as $\bar{n} \to \bar{n}/(1+2\kappa)$, but the individual galaxy brightnesses are enhanced as $\langle s \rangle \to \langle s \rangle(1+2\kappa)$, thus producing no net effect. On the other hand, the power spectrum of the map, which scales as $C^0_\ell = \langle s^2 \rangle/\bar{n}$, is modified as $C_\ell = C^0_\ell(1+2\kappa)$. As a result, the reconstruction noise on the convergence is directly related to the intrinsic fluctuations in the power spectrum,

$$\sigma_\kappa^2 = \frac{1}{4} \frac{1}{\text{SNR}_{\text{Poisson}}}, \qquad (D7)$$

where $\text{SNR}_{\text{Poisson}} = \frac{\mathcal{T}^0}{C^2}$ is the signal-to-noise ratio on the measured power spectrum, computed above. Finally, since this noise on the reconstruction of the local κ is uncorrelated between positions, we get

$$N_L^\kappa = \sigma_\kappa^2 = \frac{1}{4} \frac{1}{\text{SNR}_{\text{Poisson}}} = \frac{\mathcal{T}^0}{4C^2}, \qquad (D8)$$

consistent with the result above.

2. Gaussian weights

When the power spectrum and trispectrum are independent of the multipole, Eq. (24) for the lens reconstruction noise takes the simple form

$$N_L^\kappa = \frac{2\pi}{N_{\text{modes}}} + \frac{\mathcal{T}^0}{4(C^0)^2}. \qquad (D9)$$

Notice that the reconstruction noise in this case (Gaussian weights on a non-Gaussian unlensed map) is the same as

when using the more optimal non-Gaussian weights. This coincidence occurs for the following reason: since the power spectrum and trispectrum are independent of the multipole, both the non-Gaussian and Gaussian weights give equal weight to every temperature mode, resulting in the same estimator. This is not the case for a general power spectrum and trispectrum.

APPENDIX E: MEASURING THE TRISPECTRUM IN THE COLLAPSED LIMIT

Here, we present the estimator for the CIB trispectrum $\mathcal{T}^0_{\ell, L-\ell, \ell, -L-\ell}$ in the collapsed limit $L \ll \ell$ that we have used to derive the upper limits in Fig. 5. Let us define a filtered map $T_f(\ell) = W_f(\ell) T_\ell$, with the filter normalizing such that

$$\int \frac{d^2\ell}{(2\pi)^2} W_f^2(\ell) = 1. \qquad (E1)$$

Let us further define the field $K(\mathbf{x})$ in real space as $K(\hat{\mathbf{n}}) = T_f^2(\hat{\mathbf{n}})$. The power spectrum of K is then

$$\langle K(\mathbf{L}) K(\mathbf{L}') \rangle = \int \frac{d^2\ell \, d^2\ell'}{(2\pi)^4} W_f(\ell) W_f(|\mathbf{L} - \ell|) W_f(\ell') W_f(|\mathbf{L}' - \ell'|) \langle T_\ell T_{\mathbf{L}-\ell} T_{\ell'} T_{\mathbf{L}'-\ell'} \rangle$$

$$= (2\pi)^2 \delta^D_{\mathbf{L}+\mathbf{L}'} \times \left(2 \int \frac{d^2\ell}{(2\pi)^2} W_f^2(\ell) W_f^2(|\mathbf{L} - \ell|) C_\ell^{\text{total}} C_{|\mathbf{L}-\ell|}^{\text{total}} \right.$$

$$\left. + \int \frac{d^2\ell \, d^2\ell'}{(2\pi)^4} W_f(\ell) W_f(|\mathbf{L} - \ell|) W_f(\ell') W_f(|\mathbf{L}' - \ell'|) \mathcal{T}^0_{\ell, \mathbf{L}-\ell, \ell', \mathbf{L}'-\ell'} \right). \qquad (E2)$$

In this limit, and for a narrow filter centered around ℓ_*, we thus obtain

$$C_L^{KK} = \left[\underbrace{2 \left(\int \frac{d^2\ell}{(2\pi)^2} W_f^4(\ell) \right) (C_{\ell_*}^{\text{total}})^2}_{\text{Gaussian part}} + \underbrace{\langle \mathcal{T}^0_{\ell_*, \mathbf{L}-\ell_*, \ell_*, -\mathbf{L}-\ell_*} \rangle}_{\text{angle-averaged trispectrum}} \right]. \qquad (E3)$$

To estimate the trispectrum, we subtract the Gaussian part of C_L^{KK} measured on Gaussian simulations with the same power spectrum as the CIB.

[1] M. Bartelmann and P. Schneider, Phys. Rep. **340**, 291 (2001).
[2] P. Schneider, arXiv:astro-ph/0509252.
[3] M. Kilbinger, Rep. Prog. Phys. **78**, 086901 (2015).
[4] H. Hildebrandt et al., Mon. Not. R. Astron. Soc. **465**, 1454 (2017).
[5] M. R. Becker et al., Phys. Rev. D **94**, 022002 (2016).
[6] R. Mandelbaum et al., arXiv:1705.06745.
[7] A. Lewis and A. Challinor, Phys. Rep. **429**, 1 (2006).
[8] D. Hanson, A. Challinor, and A. Lewis, Gen. Relativ. Gravit. **42**, 2197 (2010).
[9] W. Hu and T. Okamoto, Astrophys. J. **574**, 566 (2002).
[10] C. M. Hirata and U. Seljak, Phys. Rev. D **68**, 083002 (2003).
[11] J. Carron and A. Lewis, Phys. Rev. D **96**, 063510 (2017).
[12] B. D. Sherwin et al., Phys. Rev. D **95**, 123529 (2017).
[13] Y. Omori et al., Astrophys. J. **849**, 124 (2017).
[14] P. A. R. Ade et al., Phys. Rev. Lett. **113**, 021301 (2014).

[15] P. A. R Ade *et al.* (Planck Collaboration) Astron. Astrophys. **594,** A15 (2016).

[16] M. Bucher, C. S. Carvalho, K. Moodley, and M. Remazeilles, Phys. Rev. D **85,** 043016 (2012).

[17] J.-L. Puget *et al.*, Astron. Astrophys. **308,** L5 (1996).

[18] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **571,** A18 (2014).

[19] F. Lacasa, A. Pénin, and N. Aghanim, Mon. Not. R. Astron. Soc. **439,** 123 (2014).

[20] A. Pénin, F. Lacasa, and N. Aghanim, Mon. Not. R. Astron. Soc., **439,** 143 (2014).

[21] P. A. R. Ade *et al.* (Planck Collaboration) Astron. Astrophys. **571,** A30 (2014).

[22] R. A. C. Croft, A. Romeo, and R. B. Metcalf, arXiv: 1706.07870.

[23] R. B. Metcalf, R. A. C. Croft, and A. Romeo, Mon. Not. R. Astron. Soc. **477,** 2841 (2018).

[24] O. Zahn and M. Zaldarriaga, Astrophys. J. **653,** 922 (2006).

[25] A. Pourtsidou and R. B. Metcalf, Mon. Not. R. Astron. Soc. **448,** 2368 (2015).

[26] T. Lu and U.-L. Pen, Mon. Not. R. Astron. Soc. **388,** 1819 (2008).

[27] T. Lu, U.-L. Pen, and O. Doré, Phys. Rev. D **81,** 123015 (2010).

[28] A. van Engelen, S. Bhattacharya, N. Sehgal, G. P. Holder, O. Zahn, and D. Nagai, Astrophys. J. **786,** 13 (2014).

[29] S. J. Osborne, D. Hanson, and O. Doré, J. Cosmol. Astropart. Phys. 03 (2014) 024.

[30] P. A. R. Ade *et al.* (Planck Collaboration) Astron. Astrophys. **594,** A13 (2016).

[31] M. Millea, E. Anderes, and B. D. Wandelt, arXiv: 1708.06753.

[32] X. Yang and P. Zhang, Mon. Not. R. Astron. Soc. **415,** 3485 (2011).

[33] F. Schmidt, A. Leauthaud, R. Massey, J. Rhodes, M. R. George, A. M. Koekemoer, A. Finoguenov, and M. Tanaka, Astrophys. J. **744,** L22 (2012).

[34] R. Scranton *et al.*, Astrophys. J. **633,** 589 (2005).

[35] E. M. Huff and G. J. Graves, Astrophys. J. **780,** L16 (2014).

[36] G. Pratten and A. Lewis, J. Cosmol. Astropart. Phys. 08 (2016) 047.

[37] G. Fabbian, M. Calabrese, and C. Carbone, J. Cosmol. Astropart. Phys. 02 (2018) 050.

[38] A. Petri, Z. Haiman, and M. May, Phys. Rev. D **95,** 123503 (2017).

[39] H. Prince, K. Moodley, J. Ridl, and M. Bucher, J. Cosmol. Astropart. Phys. 01 (2018) 034.

[40] M. Béthermin *et al.*, Astrophys. J. **757,** L23 (2012).

[41] S. J. Schmidt, B. Ménard, R. Scranton, C. B. Morrison, M. Rahman, and A. M. Hopkins, Mon. Not. R. Astron. Soc. **446,** 2696 (2015).

[42] A. R. Pullen, P. Serra, T.-C. Chang, O. Dore, and S. Ho, arXiv:1707.06172.

[43] A. Pénin, O. Doré, G. Lagache, and M. Béthermin, Astron. Astrophys. **537,** A137 (2012).

[44] M. Béthermin, H. Dole, G. Lagache, D. Le Borgne, and A. Penin, Astron. Astrophys. **529,** A4 (2011).

[45] G. E. Addison, J. Dunkley, and J. R. Bond, Mon. Not. R. Astron. Soc. **436,** 1896 (2013).

[46] H.-Y. Wu and O. Doré, Mon. Not. R. Astron. Soc. **466,** 4651 (2017).

[47] J. L. Tinker and A. R. Wetzel, Astrophys. J. **719,** 88 (2010).

[48] R. Adam *et al.* (Planck Collaboration) Astron. Astrophys. **594,** A1 (2016).

[49] A. Mittal, F. de Bernardis, and M. D. Niemack, J. Cosmol. Astropart. Phys. 02 (2018) 032.

[50] A. Cooray, New Astron. **9,** 173 (2004).

[51] K. N. Abazajian *et al.*, arXiv:1610.02743.

[52] E. Schaan, E. Krause, T. Eifler, O. Doré, H. Miyatake, J. Rhodes, and D. N. Spergel, Phys. Rev. D **95,** 123512 (2017).

[53] M. Kesden, A. Cooray, and M. Kamionkowski, Phys. Rev. D **67,** 123507 (2003).

[54] D. Hanson, A. Challinor, G. Efstathiou, and P. Bielewicz, Phys. Rev. D **83,** 043005 (2011).

[55] V. Böhm, M. Schmittfull, and B. D. Sherwin, Phys. Rev. D **94,** 043519 (2016).

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

[57] A. Hall, C. Bonvin, and A. Challinor, Phys. Rev. D **87,** 064026 (2013).

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

[59] K. M. Smith, O. Zahn, and O. Doré, Phys. Rev. D **76,** 043510 (2007).

[60] B. D. Sherwin and S. Das, arXiv:1011.4510.

[61] H.-M Zhu, U.-L. Pen, Y. Yu, and X. Chen, arXiv: 1610.07062.

[62] E. D. Kovetz *et al.*, arXiv:1709.09066.

[63] H.-J. Seo and C. M. Hirata, Mon. Not. R. Astron. Soc. **456,** 3142 (2016).

[64] M. Takada and W. Hu, Phys. Rev. D **87,** 123504 (2013).

[65] C.-T. Chiang, C. Wagner, F. Schmidt, and E. Komatsu, J. Cosmol. Astropart. Phys. 05 (**2014**) 048.

[66] E. Schaan, M. Takada, and D. N. Spergel, Phys. Rev. D **90,** 123523 (2014).

[67] R. Scoccimarro, M. Zaldarriaga, and L. Hui, Astrophys. J. **527,** 1 (1999).

[68] M. Béthermin, L. Wang, O. Doré, G. Lagache, M. Sargent, E. Daddi, M. Cousin, and H. Aussel, Astron. Astrophys. **557,** A66 (2013).