

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/157003/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Zhu, Yijie, Beringue, Benjamin, Choi, Steve K., Battaglia, Nicholas, Meerburg, P. Daniel and Meyers, Joel 2022. Estimating the impact of foregrounds on the future detection of Rayleigh scattering. *Journal of Cosmology and Astroparticle Physics* 2022 (09) , 048. 10.1088/1475-7516/2022/09/048

Publishers page: <http://dx.doi.org/10.1088/1475-7516/2022/09/048>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# Estimating the Impact of foregrounds on the Future Detection of Rayleigh scattering

Yijie Zhu,<sup>1,2</sup> Benjamin Beringue,<sup>3,4</sup> Steve K. Choi,<sup>2,5</sup> Nicholas Battaglia,<sup>2</sup> P. Daniel Meerburg,<sup>6</sup> and Joel Meyers<sup>7</sup>

<sup>1</sup>*Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA*

<sup>2</sup>*Department of Physics, Cornell University, Ithaca, NY 14853, USA*

<sup>3</sup>*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, Wales CF24 3AA, UK*

<sup>4</sup>*DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, UK, CB3 0WA*

<sup>5</sup>*Department of Astronomy, Cornell University, Ithaca, NY 14853, USA*

<sup>6</sup>*Van Swinderen Institute for Particle Physics and Gravity,*

*University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands*

<sup>7</sup>*Department of Physics, Southern Methodist University, Dallas, TX 75275, USA*

Rayleigh scattering of the cosmic microwave background (CMB) by neutral hydrogen shortly after recombination leaves frequency-dependent imprints on intensity and polarization fluctuations. High signal-to-noise observations of CMB Rayleigh scattering would provide additional insight into the physics of recombination, including greater constraining power for parameters like the primordial helium fraction, the light relic density, and the sum of neutrino masses. However, such a measurement of CMB Rayleigh scattering is challenging due to the presence of astrophysical foregrounds, which are more intense at the high frequencies, where the effects of Rayleigh scattering are most prominent. Here we forecast the detectability of CMB Rayleigh scattering including foreground removal using blind internal linear combination methods for a set of near-future surveys. We show that atmospheric effects for ground-based observatories and astrophysical foregrounds pose a significant hindrance to detecting CMB Rayleigh scattering with experiments planned for this decade, though a high-significance measurement should be possible with a future CMB satellite.

## I. INTRODUCTION

The early universe was filled with a hot dense plasma composed primarily of electrons, protons, and helium nuclei in a bath of radiation tightly coupled to the plasma due to the high rate of Thomson scattering. As the universe cooled due to cosmic expansion, the plasma recombined to form a transparent gas of neutral hydrogen and helium. Shortly after recombination, the photons decoupled and have traveled mostly unimpeded ever since. These photons, which continued to redshift due to the cosmic expansion, make up the cosmic microwave background (CMB).

Observations of the CMB have provided a wealth of information about the history, contents, and evolution of the universe. The existence of a thermal bath of radiation provides very strong evidence for the Hot Big Bang model of cosmology, and measurements of the CMB temperature and polarization anisotropies have allowed for percent-level determination of the parameters that define the current  $\Lambda$ CDM model of cosmology [6]. Future CMB surveys will extract a great deal more information from the temperature and polarization anisotropies.

Looking beyond primary CMB anisotropies, additional processes after recombination leave imprints on the CMB. One such process is Rayleigh scattering of CMB photons by neutral atoms, especially by hydrogen, which achieves maximum density shortly after recombination. Rayleigh scattering is a classical scattering process of long wavelength radiation by the induced dipole of neutral species. Its impact on recombination and CMB anisotropies has been studied in Alipour et al. [7], Lewis [26], Yu et al. [37]. The spectral function for Rayleigh is well described by a cross-section that scales as  $\nu^4$ .

However, as we will show, observing Rayleigh scattering of the CMB remains challenging due to the presence of astrophysical foregrounds that act as a source of confusion at the relevant frequencies.

Physically, Rayleigh scattering causes a frequency-dependent damping of CMB anisotropies by scattering photons from hot spots out of the line-of-sight and from cold spots into the line of sight. At high frequencies the effect of Rayleigh scattering contributes to the angular power spectrum, and thus it must be consistently included in cosmological modeling.

Observing the effects of Rayleigh scattering of the CMB would provide additional information that can be used to probe the early universe, including the physics of recombination, the ionization history, the expansion rate of the early universe, and the properties of primordial perturbations. One interesting application of CMB Rayleigh scattering is to improve measurements of the light relic density ( $N_{\text{eff}}$ ), which further probes physics beyond the Standard Model in the dark sector [19, 22]. With primary CMB observations alone, measuring  $N_{\text{eff}}$  requires significantly more resources when compared to including Rayleigh scattering information [e.g., 1, 11]. However, the value of Rayleigh in constraining physics beyond the Standard Model requires a significant detection of the signal. For example, for experiments such as CCAT-prime [12], the Simons Observatory (SO) [5], and the *Planck* satellite [31], even under ideal circumstance, where foregrounds and other possible systematic effects are neglected, the expected statistical significance of the Rayleigh scattering signal would rise to the level of a detection, but would not significantly improve constraints on cosmological parameters [11].

The amplitude of the Rayleigh auto-spectrum is

very small, both in temperature and in polarization. At around  $\nu = 500$  GHz, the temperature Rayleigh scattering auto-spectrum is still around 7 orders of magnitude lower than the primary CMB spectrum [26]. Hence a detection of the auto-spectrum will remain out of reach for the foreseeable future. On the other hand, cross-correlations between the Rayleigh scattering signal and the primary temperature or polarization anisotropies are more promising [7, 11, 26]. The amplitude of the cross signals is around 3 orders of magnitude larger than the auto-spectrum of Rayleigh scattering, as shown in Fig. 2. Raw noise levels from current ground-based CMB experiments, the *Planck* satellite and future CMB satellite missions [e.g., PICO 23] suggests the Rayleigh scattering cross-correlations should be detectable [e.g., 7, 11, 26]. The main challenge is to isolate the Rayleigh scattering signal in the presence of foregrounds and the atmosphere, both of which can have a large impact on detectability. While the role of the atmosphere is only relevant for ground-based experiments, astrophysical foregrounds will affect both satellite missions and ground-based experiments. This paper aims to address the effect of foregrounds on detectability of the Rayleigh scattering signal. Indeed, the frequency dependence of some foreground emissions make their amplitude significant in the frequency range where the Rayleigh scattering is brighter. It is timely to assess how large this impact will be for future CMB experiments.

Previously, Alipour et al. [7] studied the effect of foregrounds and foregrounds removal on the Rayleigh scattering signal and concluded that the  $TE$  cross-power spectra between primary temperature and Rayleigh scattering polarization signals should be relatively unbiased by foregrounds. More recently, Beringue et al. [11] forecasted a detection of the  $TT$  cross-correlation with  $4.7\sigma$  for *Planck*, and a whopping  $715\sigma$  detection with PICO, both in the absence of foregrounds [11, 26]. Throughout this paper, we denote the signals as  $XY = \text{CMB } (T/E) \times \text{RS } (T/E)$ , where  $T$  represents the temperature fluctuations,  $E$  refers to the  $E$ -mode polarization, and RS denotes the Rayleigh scattering component. These calculations show that  $TT$  cross-correlation signal-to-noise ratio is roughly an order of magnitude larger than the  $TE$  or  $ET$  signal-to-noise. Therefore, given the large difference in signal-to-noise in the absence of foregrounds, a more comprehensive study of all Rayleigh scattering cross-correlations, including foreground removal is warranted.

We use a component separation method to quantify the impact of foreground removal on the Rayleigh scattering signal-to-noise. Several component separation methods and algorithms already exist in the literature, including *blind* methods like Spectral Matching Independent Component Analysis (SMICA) [16], which was used in *Planck* data analysis Planck Collaboration et al. [30], Fast Independent Component Analysis (FastICA) [28], the Joint Approximate Diagonalization

of Eigenmatrices method (JADE) [29], and Correlated Component Analysis (CCA) [9]. We chose to be conservative in our approach to component separation and selected the method known as Internal Linear Combination (ILC) [15], which does not assume particular parametrization of the foreground components or templates.

Since we are dealing with the cross-spectra of two signals, we use a variation of this method, the constrained Internal Linear Combination (cILC) [33]. With this method, we avoid residual bias from the CMB primary on the Rayleigh signal, and vice versa in the cross-spectra [32]. Given the conservative nature of our approach, these forecasts should be considered as lower limits on the possible signal-to-noise.

In this paper, we introduce the physical modeling and scientific significance of the Rayleigh scattering signal in Section II. We set up our modeling of the foregrounds and describe the experimental configurations we consider in Section III and Section IV; additional information about the surveys can be found in Appendix A. The CMB and Rayleigh spectra are generated using a modified version of CAMB<sup>1</sup> [27]. Details on the cILC method are presented in Section V. In Section VI we forecast the signal-to-noise for CCAT-prime, SO and *Planck*. In addition, we explore the effect of systematics: bias and gain calibration. We show how foregrounds impact our observation. Then, we perform the same analysis for PICO and explore combinations of data from LiteBIRD with CCAT-prime and *Planck*. We use the extrapolated atmosphere noise model on all frequencies from ACT [20]. The atmospheric noise for CCAT-prime and SO might differ from these estimates, so our results should be considered with this in mind. We sometimes use  $\mathcal{D}_\ell$  which is related to  $C_\ell$  as:

$$\mathcal{D}_\ell = \frac{\ell(\ell+1)}{2\pi} C_\ell. \quad (1)$$

## II. RAYLEIGH SCATTERING PHYSICS

Rayleigh scattering describes the classical scattering of electromagnetic waves by particles with size is much smaller than the wavelength of the radiation. To first order, the cross section of this scattering process is proportional to  $\nu^4$ , where  $\nu$  is the photon frequency. Its impact on the recombination history and decoupling of CMB photons has been described in details in Refs. [7, 26, 37]. We will summarise the cosmological implications of Rayleigh scattering in this section.

In the early universe, photons are tightly coupled with the free electrons in the primordial plasma by Thomson scattering. As the universe expands, the temperature decreases until recombination of free electrons and protons becomes thermally favored. As the fraction of

---

<sup>1</sup> <https://camb.info/>

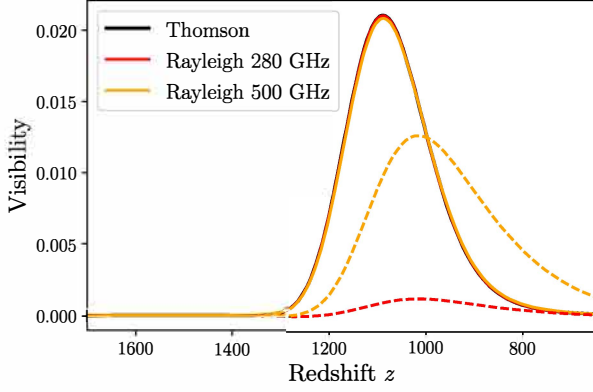


FIG. 1. Change to the visibility function  $g(\tau) = \tau e^{-\tau}$  due to Rayleigh scattering at  $\nu = 280$  GHz and  $\nu = 500$  GHz. The dashed lines represent the normalized contribution from Rayleigh scattering alone. The Rayleigh scattering signal shifts the peak of the visibility function towards lower redshifts.

free electrons drops, the Thomson scattering rate falls. Around redshift  $z \sim 1100$ , photons experience a last scattering event and start free-streaming.

However, the recombination processes also yield the production of neutral species, mainly hydrogen and helium, by which CMB photons can be scattered through Rayleigh scattering. Unlike Thomson scattering, the Rayleigh scattering cross section is frequency dependent:

$$\sigma_R(\nu) = \sigma_T \sum_{n=1} a_n \left( \frac{\nu}{\nu_0} \right)^{2n+2}, \quad (2)$$

where  $a_1 = 1$ ,  $a_2 = 2.626$ ,  $a_3 = 5.502$  [26],  $\sigma_T$  is the cross section of Thomson scattering and  $\nu_0 = 3.102 \times 10^6$  GHz is the Lyman limit frequency. After recombination, neutral species quickly dilute due to the cosmic expansion (scaling as  $\propto a^{-3}$ ) and photons are diluted and redshifted ( $\propto a^{-4}$ ), making the probability of Rayleigh scattering  $\propto a^{-7}$ . As a result, most Rayleigh scattering events occur shortly after recombination. This can be seen in the visibility function  $g(z) \equiv \tau e^{-\tau}$  where  $\tau$  is the optical depth. The visibility function describes the probability that a photon last scattered at a given redshift. The visibility functions are shown in Fig. 1 for both Thomson and Rayleigh scattering.

By increasing the comoving opacity of the plasma in a frequency-dependent way, Rayleigh scattering leaves characteristic imprints on the CMB:

- **Frequency-dependent shift of the visibility function:** As can be seen in Fig. 1, the visibility function is shifted towards lower redshifts (later time). Crucially, the position of the peak of the visibility function is made frequency dependent with the shift towards later time being larger at higher frequencies.

- **Increase of diffusion damping:** Rayleigh scattering globally increases the comoving opacity of the plasma. The mean free path of photons in the plasma is therefore reduced. The strength of diffusion damping is directly controlled by the mean free path [24]. At lower frequencies, Rayleigh scattering has negligible impact. However, at higher frequencies, the visibility function is shifted towards later time when cosmic expansion has led to a global decrease of the comoving opacity, leading to an overall increase of diffusion damping.

- **Frequency-dependent sound horizon:** As the peak of the visibility function is rendered frequency dependent by Rayleigh scattering, the size of the sound horizon at last scattering similarly inherits the same frequency dependence, being larger at higher frequencies. The size of the sound horizon dictates the location of the acoustic peaks in both the matter and the CMB power spectra, a larger sound horizon shifts the acoustic peaks towards larger scales.

- **Frequency-dependent amplitude of polarization signal:** The shift of the visibility function induced by Rayleigh scattering increases the amplitude of the local temperature quadrupole around the time of recombination, leading to a boost in  $E$  modes on large scales.

If the shift of the visibility function is small enough, the effect of Rayleigh scattering on the temperature and polarization power spectra is accurately described by a linear, frequency-dependent distortion [26]:

$$a_{\ell m}^X(\nu) = a_{\ell m}^X + \left( \frac{\nu}{\nu_0} \right)^4 \Delta a_{\ell m}^{X,4} + \dots, \quad (3)$$

where  $a_{\ell m}$  are the coefficients in the spherical harmonic transform of the temperature ( $X = T$ ) and polarization ( $X = E, B$ ) anisotropies. Note that we have neglected higher order corrections to the Rayleigh scattering cross section that would give rise to additional terms on the right hand side of Eq. (3). The power spectra are then given by

$$\begin{aligned} C_\ell^{XY}(\nu_1, \nu_2) &= \langle a_{\ell m}^X(\nu_1) a_{\ell m}^Y(\nu_2) \rangle \\ &= C_\ell^{XY} + \left( \frac{1}{\nu_0} \right)^4 \left( \nu_1^4 C_\ell^{X\Delta Y_4} + \nu_2^4 C_\ell^{\Delta X_4 Y} \right) \\ &\quad + \left( \frac{\nu_1 \nu_2}{\nu_0^2} \right)^4 C_\ell^{\Delta X_4 \Delta Y_4} + \dots \end{aligned} \quad (4)$$

The first term in this expansion is the primary CMB, sourced by Thomson scattering only. The following term contains the cross-correlation spectra between the primary CMB and the Rayleigh scattering signal. The last term is proportional to the Rayleigh scattering auto-spectrum. These terms are also ordered with decreasing



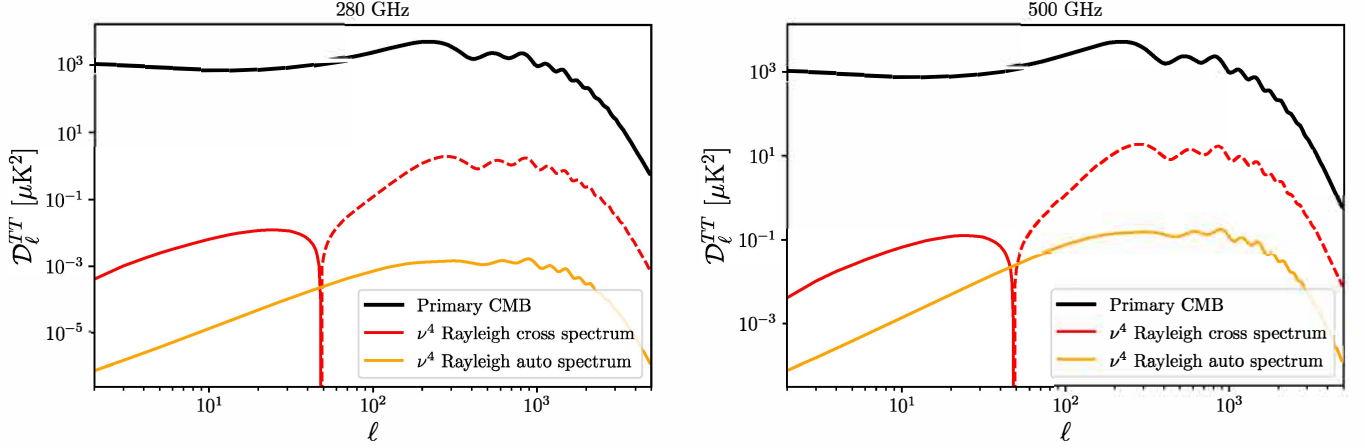


FIG. 2. Contributions from Rayleigh scattering at  $\nu = 280$  GHz and  $\nu = 500$  GHz. Red lines show the cross-spectra between the Rayleigh signal and the primary CMB. The solid lines present positive and the dotted parts are negative correlations. Yellow lines show the auto-spectra of the Rayleigh signal. The cross-spectra are roughly 3 orders of magnitude larger than the auto-spectra of the Rayleigh signal.

amplitude (see Fig. 2) making the cross-correlation the natural target to achieve a first detection.

With the low-noise and large-area maps expected from future CMB experiments, the detection of this faint Rayleigh scattering signal should be possible [11, 26]. More realistic estimates of the expected detection significance require forecasts including foregrounds and testing the robustness of these predictions after applying foreground cleaning techniques.

### III. FOREGROUND MODELING

The microwave sky foreground models we considered in this paper include the cosmic infrared background (CIB), the thermal Sunyaev-Zeldovich effect (tSZ), the kinetic Sunyaev-Zeldovich (kSZ), radio sources, and both galactic dust and synchrotron emissions. The CIB is sourced by star-forming dusty galaxies and is modelled according to Dunkley et al. [18]. The tSZ and kSZ are the two dominant effects in the scattering of CMB photons by free electrons in clusters. We model the tSZ and kSZ auto-spectra using templates from Battaglia et al. [8]. We also include the cross-spectra between tSZ and CIB in our foreground modeling, following Addison et al. [4]. In principle there is also a polarized version of the SZ effect (pSZ) and the relativistic SZ effect (rSZ), but these are too small to have an impact on our forecasts, so we do not consider them in this analysis. Radio point sources (radio galaxies) are modeled following Ade et al. [5]. The angular power spectrum in temperature and polarization for all of these foregrounds, the CMB, and Rayleigh scattering are shown in Figure 3.

The polarized dust and synchrotron foregrounds from our galaxies are modeled following Choi & Page [14]. For

the dust model the emissivity is assumed to follow:

$$\epsilon_d(\nu, \beta_d) = \left(\frac{\nu}{\nu_0}\right)^{\beta_d-2} \frac{B_\nu(T_d)}{B_{\nu_0}(T_d)}. \quad (5)$$

Here  $\beta_d = 1.59$ ,  $T_d = 19.6$  K,  $\nu_0 = 353$  GHz, and  $B_\nu$  is the Planck function.

We can then express the dust and synchrotron power spectrum as:

$$\mathcal{D}_\ell^{d,s}(\nu_1, \nu_2) = \left[ \alpha_s^2 \left( \frac{\nu_1 \nu_2}{\nu_K^2} \right)^{\beta_s} + \alpha_d^2 \epsilon_d(\nu_1, \beta_d) \epsilon_d(\nu_2, \beta_d) \right] \times g(\nu_1) g(\nu_2) \left( \frac{\ell}{80} \right)^{-0.42}, \quad (6)$$

where  $\alpha_s$  and  $\alpha_d$  are two scaling constants, which we approximate as  $\alpha_s^2 = 10.3 \mu\text{K}^2$  and  $\alpha_d^2 = 0.53 \mu\text{K}^2$  in antenna temperature,  $\beta_s = -3.1$  is the synchrotron spectral index. These parameters are measured from the cleanest 45% of the sky, on angular scales with  $\ell = 50-110$  [14]. We ignore the cross-spectra between dust and synchrotron since it is small [14]. Finally,  $g(\nu)$  is the conversion function from antenna to CMB thermodynamic temperature at frequency  $\nu$ :

$$g(\nu) = \frac{\left( e^{\frac{h\nu}{k_B T_{\text{CMB}}}} - 1 \right)^2}{e^{\frac{h\nu}{k_B T_{\text{CMB}}}} \left( \frac{h\nu}{k_B T_{\text{CMB}}} \right)^2}, \quad (7)$$

where  $h$  is the Planck constant and  $k_B$  is the Boltzmann constant.

For radio point sources, we write the auto-spectrum as:

$$\mathcal{D}_\ell^{\text{ps}}(\nu_1, \nu_2) = A_{\text{ps}} \left( \frac{\nu_1 \nu_2}{\nu_r^2} \right)^{-2.5} \frac{g(\nu_1) g(\nu_2) \ell^2}{(g(\nu_r))^2 \ell_0^2}, \quad (8)$$

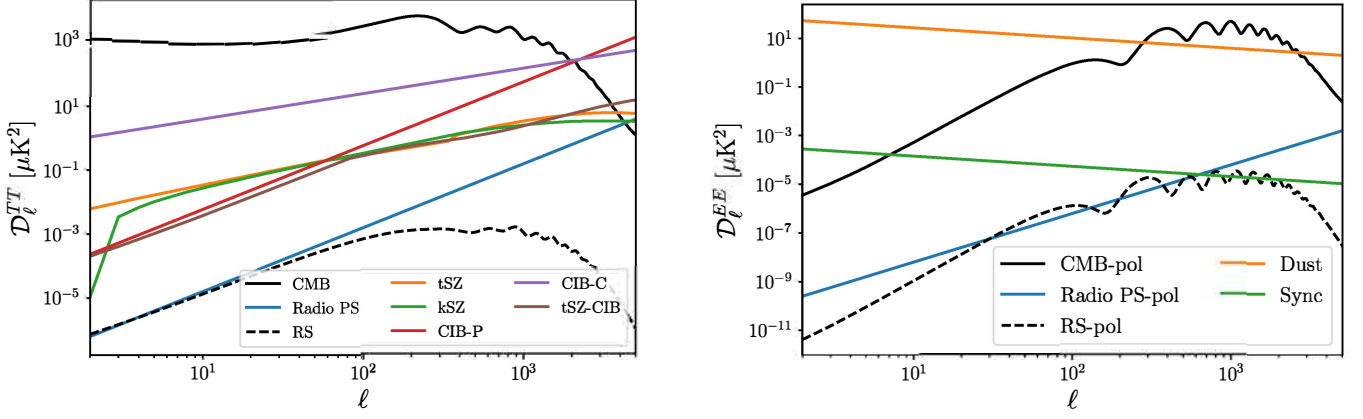


FIG. 3. Signal and foreground power spectra at  $\nu = 280$  GHz. The observed temperature includes contributions from the CMB primary signal, Rayleigh scattering auto signal, radio point sources, thermal Sunyaev–Zeldovich (tSZ), kinetic Sunyaev–Zeldovich (kSZ), cosmic infrared background (CIB), and cross correlation of tSZ and CIB; for polarization, the polarized primary CMB and polarized Rayleigh signal, dust, synchrotron, and radio point sources are included.

where  $A_{\text{ps}} = 3.1 \mu\text{K}^2$ ,  $\ell_0 = 3000$ ,  $\nu_r = 150$  GHz.

For polarized radio point sources, we modify Eq. (8) as:

$$\mathcal{D}_\ell^{\text{ps-pol}}(\nu_1, \nu_2) = A_{\text{ps-pol}} \left( \frac{\nu_1 \nu_2}{\nu_r^2} \right)^{-2.5} \frac{g(\nu_1)g(\nu_2)}{(g(\nu_r))^2} \frac{\ell^2}{\ell_0^2} \quad (9)$$

where  $A_{\text{ps-pol}} = 0.015 \mu\text{K}^2$  at  $\ell_0 = 3000$  and  $\nu_r = 150$  GHz [5]. We assume that the galactic dust and synchrotron do not depend on the position of the sky observed. We show all auto-spectra at  $\nu = 280$  GHz in Fig. 3 for temperature and polarization. We apply the best-fit value from Dunkley et al. [18] to the extragalactic foreground models.

#### IV. INSTRUMENTAL AND ATMOSPHERIC NOISE MODELS

We consistently model the noise power spectra of the surveys we considered, CCAT-prime, *Planck*, Simons Observatory (SO), PICO and LiteBIRD using a model that consists in a *red* and a *white* component:

$$n_\ell^{\text{ins}} = n_{\text{red}} \left( \frac{\ell}{\ell_{\text{knee}}} \right)^{\alpha_{\text{knee}}} + n_{\text{white}}. \quad (10)$$

Here  $n_{\text{white}}$  is the white noise due to the instrument alone,  $n_{\text{red}}$  is the red noise due to the atmosphere, where  $\ell_{\text{knee}}$  and  $\alpha_{\text{knee}}$  parameterize the scales affected by atmospheric contamination [5]. For satellite experiments such as *Planck*, we simply set  $n_{\text{red}} = 0$ . For ground-based observations, the red noise component is estimated from empirical measurements of the large angular scale noise from ACT at 150 GHz, and extrapolated to higher frequencies for SO, CCAT-prime, and CMB-S4 [1, 5, 12]. Our forecasts will include *Planck* to complement

future ground-based experiments on the largest angular scales [31].

The Simons Observatory consists in a set of telescopes under construction in Chile. Its large aperture telescope (LAT) will have a small beam allowing for precision measurement of small angular scales. We refer to the SO LAT as ‘SO’ throughout [5]. The Simons Observatory experiment also includes three small aperture telescopes (SATs), with larger beams. Because of the large beams, these SATs will be limited to  $\ell < 400$ . With polarization modulators, SATs will only produce polarization maps; hence we do not consider temperature anisotropies measurements from the SATs.

The CCAT-prime observatory, currently under construction in Chile, will consist in the Fred Young Sub-millimeter Telescope (FYST) equipped with two instruments: an imaging camera, Prime-CAM [13] and a multi-beam submillimeter heterodyne spectrometer, CHAI [36]. We will refer to the FYST + Prime-CAM configuration as ‘CCAT-prime’ for simplicity. Similar to SO, CCAT-prime will have a high sensitivity at small angular scales [13]. Compared to the SO LAT, CCAT-prime will focus on higher frequencies. Besides, the higher altitude of the site is expected to yield lower atmospheric noise.

LiteBIRD is a planned satellite mission covering a large frequency range. Due to the relatively small dish the beam will limit its resolution to  $\ell < 400$  [21].

The Probe of Inflation and Cosmic Origin (PICO) is a recently proposed space satellite [23], with extremely low noise and many frequency channels. PICO will cover a large range of multipoles. If funded, it will be capable of a high significance measurement of Rayleigh scattering [11].

All the noise parameters for the experiments mentioned above are listed in Appendix A. For the ground-based experiments, we expect some differences

between the measured red noise components and extrapolated values we use. Additionally we do not consider that the atmospheric red noise can be removed using proposed analysis techniques [17]. Thus, our forecast in this respect are conservative, since if such techniques are viable this would only improve the forecasts we present.

## V. COMPONENT SEPARATION AND RECONSTRUCTED NOISE SPECTRA

The forecasts carried out in this work follow the constrained Internal Linear Combination (cILC) framework derived in Remazeilles et al. [33]<sup>2</sup>. Since we are interested in measuring the cross correlation between the primary CMB and the Rayleigh scattering signals, we schematically need to produce two maps: one that contains only the CMB signal and no Rayleigh scattering and another one that contains only the Rayleigh scattering signal and no primary CMB. This procedure will ensure that no residual CMB contamination in the Rayleigh scattering map (and vice-versa) would bias the cross correlation. We model the total data vector at a pixel  $p$  as:

$$\mathbf{d}(p) = \mathbf{a}^{\text{CMB}} s^{\text{CMB}}(p) + \mathbf{b}^{\text{RS}} s^{\text{RS}}(p) + \mathbf{n}(p), \quad (11)$$

where  $\mathbf{a}$  is the frequency dependence of the CMB (assumed to be unity at all frequencies),  $\mathbf{b}$  is the frequency dependence of the Rayleigh scattering signal ( $\propto \nu^4$ ).  $\mathbf{n}$  captures the effect of foregrounds, instrumental and atmospheric noises. A pixel here refers to either an actual pixel in a sky map or a set of harmonic coefficients  $(\ell, m)$  of the spherical harmonic transform.

The cILC procedure yields a set of weights  $\mathbf{w}$  that, for the CMB map, satisfy:

- $\mathbf{w}^\top \cdot \mathbf{a} = 1$ .
- $\mathbf{w}^\top \cdot \mathbf{b} = 0$ .
- $\hat{s}^{\text{CMB}}(p) \equiv \mathbf{w}^\top \cdot \mathbf{d}(p)$  has minimum variance.

These weights are given by [33]:

$$\mathbf{w}_{\mathbf{a}=1, \mathbf{b}=0}^\top = \frac{(\mathbf{b}^\top \tilde{\mathbf{R}}_d^{-1} \mathbf{b}) \mathbf{a}^\top \tilde{\mathbf{R}}_d^{-1} - (\mathbf{a}^\top \tilde{\mathbf{R}}_d^{-1} \mathbf{b}) \mathbf{b}^\top \tilde{\mathbf{R}}_d^{-1}}{(\mathbf{a}^\top \tilde{\mathbf{R}}_d^{-1} \mathbf{a}) (\mathbf{b}^\top \tilde{\mathbf{R}}_d^{-1} \mathbf{b}) - (\mathbf{a}^\top \tilde{\mathbf{R}}_d^{-1} \mathbf{b})^2}. \quad (12)$$

Here  $\tilde{\mathbf{R}}_d$  is the covariance of the input maps. If the cILC is carried out in the harmonic domain, the weights will inherit a scale dependence:  $\mathbf{w} \rightarrow \mathbf{w}_\ell$ . Finally, we note that the same procedure yields the weights for the

Rayleigh scattering map ( $\mathbf{w}_{\mathbf{a}=0, \mathbf{b}=1}$ ) by simply inverting  $\mathbf{a}$  and  $\mathbf{b}$  in Eq. (12).

From the linear combinations  $\hat{s}^{\text{CMB}}(p) = \mathbf{w}_{\mathbf{a}=1, \mathbf{b}=0}^\top \cdot \mathbf{d}(p)$  and  $\hat{s}^{\text{RS}}(p) = \mathbf{w}_{\mathbf{a}=0, \mathbf{b}=1}^\top \cdot \mathbf{d}(p)$  we can compute the primary CMB and Rayleigh scattering auto and cross-spectra, that, given the constraints obeyed by the weights, read:

$$\begin{aligned} \hat{C}_\ell^{\text{CMB}} &= C_\ell^{\text{CMB}} + \mathbf{w}_{\mathbf{a}=1, \mathbf{b}=0}^\top \mathbf{N}_\ell \mathbf{w}_{\mathbf{a}=1, \mathbf{b}=0}, \\ &= C_\ell^{\text{CMB}} + N_\ell^{\text{CMB}}, \\ \hat{C}_\ell^{\text{RS}} &= C_\ell^{\text{RS}} + \mathbf{w}_{\mathbf{a}=0, \mathbf{b}=1}^\top \mathbf{N}_\ell \mathbf{w}_{\mathbf{a}=0, \mathbf{b}=1}, \\ &= C_\ell^{\text{RS}} + N_\ell^{\text{RS}}, \\ \hat{C}_\ell^{\text{CMB} \times \text{RS}} &= C_\ell^{\text{CMB} \times \text{RS}} + \mathbf{w}_{\mathbf{a}=0, \mathbf{b}=1}^\top \mathbf{N}_\ell \mathbf{w}_{\mathbf{a}=1, \mathbf{b}=0}, \\ &= C_\ell^{\text{CMB} \times \text{RS}} + N_\ell^{\text{CMB} \times \text{RS}}. \end{aligned} \quad (13)$$

Here  $\mathbf{N}_\ell$  is the covariance of the foregrounds, instrumental and atmospheric noises. We also note that when we consider the cross-correlation between the primary temperature CMB and the Rayleigh scattering  $E$ -mode polarization signal (or vice-versa), the noise term  $N_\ell^{\text{CMB} \times \text{RS}}$  vanishes due to our modelling of foregrounds and noise.

The variance of the reconstructed cross-spectrum is given by [25],

$$\Delta C_\ell = \frac{1}{\sqrt{f_{\text{sky}}(2\ell+1)}} \left[ (C_\ell^{\text{RS}} + N_\ell^{\text{RS}}) (C_\ell^{\text{CMB}} + N_\ell^{\text{CMB}}) + (C_\ell^{\text{CMB} \times \text{RS}} + N_\ell^{\text{CMB} \times \text{RS}})^2 \right]^{1/2}. \quad (14)$$

The calculations above pertain to a given experiment. However, we are interested in the combination of experiments to maximize our ability to measure the Rayleigh scattering signal. To obtain a combined noise on the cross-spectra for a given set of experiments, we augment the data vector  $\mathbf{d}(p)$  by taking the set of frequency channels to be:

$$f^{\text{total}} = (f^1, f^2, \dots, f^i). \quad (15)$$

The signal-to-noise of the cross-correlation is given by:

$$\text{SNR}^2 = \sum_\ell \frac{(C_\ell^{\text{CMB} \times \text{RS}})^2}{\Delta C_\ell^2} \quad (16)$$

## VI. RESULTS

The detectability for different experiments and combinations of experiments are summarized in Tab. I. In Fig. 5 we focus on the combination of CCAT-prime, SO and *Planck* and show the SNR for the four CMB-Rayleigh scattering cross-spectra with and without foregrounds. For this combined configuration we estimate SNR of  $1.1\sigma$  for the  $TT$  signal with foregrounds present. For all other combinations, we

<sup>2</sup> Technically, our analysis is not an cILC, since we are not constructing the covariance matrix from a map or data, but we will refer to it as such.

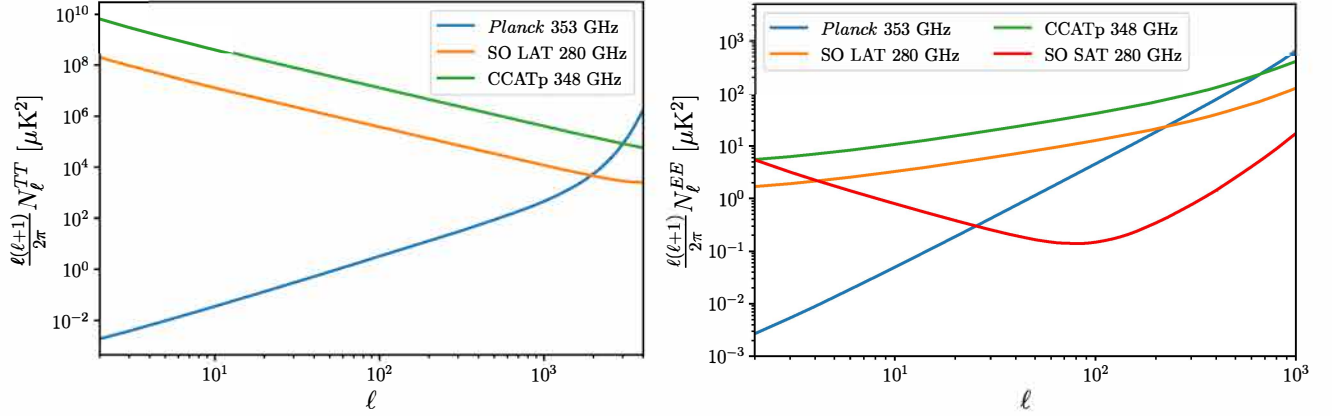


FIG. 4. Comparison of instrumental noises of *Planck*, CCAT-prime, SO LAT, and SO SAT at similar frequency channels. CCAT-prime and SO have smaller noise on small scales, while *Planck* and SAT have smaller noise on large scales. In practice, we only use data from SO SAT for  $\ell \leq 400$ .

| Configuration                         | SNR       |           |           |           |
|---------------------------------------|-----------|-----------|-----------|-----------|
|                                       | <i>TT</i> | <i>TE</i> | <i>ET</i> | <i>EE</i> |
| CCATp+SO+ <i>Planck</i>               | 1.1       | 0.3       | 0.3       | 0.6       |
| CCATp+SO+ <i>Planck</i> no Foreground | 4.6       | 0.5       | 1.6       | 1.1       |
| LiteBIRD                              | 1.9       | 0.1       | 0.9       | 0.2       |
| LiteBIRD+CCATp+ <i>Planck</i>         | 2.2       | 0.2       | 0.9       | 0.4       |
| PICO                                  | 85        | 17        | 43        | 26        |

TABLE I. Forecasted signal-to-noise ratio for various configurations. The CCATp+SO+*Planck* configuration can potentially achieve a  $1.1\sigma$  detection with foreground present. We also show the impact of foregrounds for the combination CCATp+SO+*Planck*. PICO, with its broad frequency coverage, is capable of a high-significance detection of Rayleigh scattering even after foreground removal.

find signal-to-noise below  $1\sigma$ . Without foregrounds, the combined configuration of CCAT-prime, SO and *Planck* could reach  $4.6\sigma$  in *TT*,  $1.1\sigma$  in *EE*, and  $1.6\sigma$  in *ET*. This clearly demonstrates the significant impact of foregrounds; for example the *TT* SNR is reduced by 77% by the inclusion of foregrounds and their removal. Similarly, the signal-to-noise in *ET* reduces by 81%, *EE* reduces by 45% and *TE* reduces by 40%. The latter is a result of the fact that in *TE*, foregrounds are not very significant.

The SNR without foregrounds is calculated using a modified covariance matrix with only instrumental and atmospheric noise. As a consistency check we compare our SNR without foregrounds to the Fisher forecasts from Beringue et al. [11], which did not include any foregrounds. We find good agreements between the two methods, with the SNR from the cILC being slightly smaller due to the increase in variance caused by the deprojections.

In Fig. 6 we show the  $2\sigma$  upper limit on the

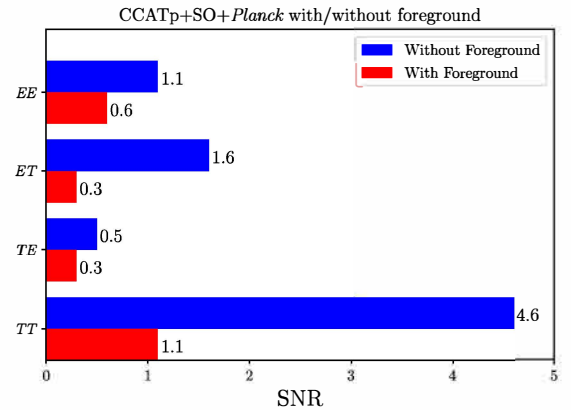


FIG. 5. Forecast of SNR with and without foregrounds for CMB Rayleigh scattering cross-spectrum using CCATp+SO+*Planck*. The *TT* SNR is significantly impacted by the foregrounds, being reduced by 77% when foregrounds are present, which demonstrates that foregrounds are an hindrance for the first detection of the Rayleigh scattering signal.

detectability of the *TT* and *EE* Rayleigh scattering signals for the CCAT-prime *Planck*, SO-LAT, and SO-SAT (*EE* only) as well as their combination. As expected we find *Planck* to dominate the constraints on large scales in temperature as ground-based experiments are heavily impacted by the atmosphere. On small scales and in polarization, the improvements in both sensitivity and resolution enabled by ground-based surveys lead to SO and CCAT-prime driving the constraints.

We also estimate how much CCAT-prime is impacted by the atmosphere. The baseline signal-to-noise including atmosphere for *TT* signal is  $\text{SNR} = 0.07$ , while it is  $\text{SNR} = 1.79$  without atmosphere. If we combine the atmosphere-less CCAT-prime with SO and



*Planck*, we obtain  $\text{SNR} = 1.88$  for the  $TT$  channel. This suggests that the atmosphere severely limits detectability prospects for ground-based observations, thus mitigating atmospheric contributions plays an important role in the detection of Rayleigh scattering from the ground [17].

### A. Residual Foreground Bias

The cILC foreground cleaning method does not perfectly null all foregrounds, only those explicitly chosen to be nulled. As a result, residual signals from foreground not explicitly nulled remain and could possibly contaminate or bias the detection of Rayleigh scattering. For Rayleigh scattering this is especially important to check due to the relative amplitudes of foregrounds compared to Rayleigh cross-spectra. Knowing the amplitude of residuals will also inform whether further foreground cleaning is required. For this study we focus on temperature, since the foregrounds are less bright in polarization and hence less of a concern, as highlighted in Fig. 5.

We define the bias  $B_\ell$  induced by residual foreground from a component with frequency-covariance matrix  $\mathbf{S}_\ell$  to be:

$$B_\ell = \mathbf{w}_{a=0,b=1}^\top \mathbf{S}_\ell \mathbf{w}_{a=1,b=0}, \quad (17)$$

where  $\mathbf{w}_{a=0,b=1}$  and  $\mathbf{w}_{a=1,b=0}$  are defined in Eq. 12. We consider tSZ, CIB, and radio point sources for  $\mathbf{S}_\ell$ . We are especially interested in the CIB contamination since CIB becomes the dominant foreground signal at sub-millimeter wavelengths where the Rayleigh signal is stronger.

In Fig. 8 we show the level of residual bias ( $B_\ell$ ) resulting from tSZ, CIB, and radio point sources foregrounds calculated for both CCAT-prime and a combination of CCAT-prime, SO and *Planck*. Fig. 8 illustrates that for CCAT-prime alone the CIB and tSZ are 1 to 2 orders of magnitude stronger than the CMB-Rayleigh cross-spectrum at 280 GHz, while radio point source residuals are around 1 to 4 orders of magnitude lower than the signal. Between  $\ell = 2000 - 3000$ , we can see that the tSZ bias is slightly lower than the signal at 280 GHz.

Combining the experiments adds more frequency coverage, and the CIB residual is more suppressed, but the tSZ residual is less well suppressed. Given the results found here, ground-based detection of Rayleigh scattering is further complicated by CIB and tSZ residuals after the cILC cleaning. This necessitates the need to find methods to better constrain those components in the future or further constrain them in the cILC at the cost of SNR.

### B. Experimental Calibration Tolerance

Another potential source of hindrance to a first detection of Rayleigh scattering are instrumental systematic effects. In particular, we study the impact of imperfect calibration of a given experiment. We will refer to this overall calibration systematic as the gain. We model this gain systematic following Abitbol et al. [2]. Such gain calibration systematics might arise from fabrication variation in the on-chip bandpass filters between both detectors and wafers, time variation from atmospheric fluctuations, and systematic effects in the bandpass calibrations from measurements using Fourier transform spectrometers or other techniques. A formal model for gain calibration systematics would include an effective shift in frequency  $\Delta\nu$ , and a gain difference  $\Delta g$ , which describes how much of the sky emission signal the extra gain picks up [2]. We estimate an approximately  $\Delta g = 1\%$  calibration gain of the sky emission and do not consider the possible effective shift in the mean frequency to simplify our analysis. We can express the gain as

$$G_\ell = \Delta g \mathbf{w}_{a=0,b=1}^\top \mathbf{E}_\ell \mathbf{w}_{a=1,b=0}, \quad (18)$$

where the sky emission  $\mathbf{E} = \mathbf{S}_{\text{fg}} + \mathbf{S}$ . Here  $\mathbf{S}_{\text{fg}}$  represents the foregrounds, and  $\mathbf{S}$  represents the signals.

We show the sky emission in temperature for each channel for CCAT-prime in the right panel of Fig. 9. For the CCAT-prime lower frequency channels the sky emission looks like a CMB auto-spectrum, but at higher frequencies the extragalactic foreground signals are more dominant at high- $\ell$ , especially the CIB.

We calculate the gain bias for the combined configuration of CCAT-prime, SO, and *Planck* with  $\Delta g = 1\%$  and  $\Delta g = 35\%$ . These biases are shown in left panel of Fig. 9. For  $\Delta g = 1\%$  of the sky emission the gain bias is around 3 to 4 orders of magnitude smaller than the noise power spectrum and would therefore be a sufficient calibration of instrument for our purpose. This is, in principle, informative in terms of the order of magnitude of the gain. It is helpful to verify that the gain is generally 4 to 6 orders of magnitude lower than our noise power spectrum. If we raise the gain to  $\Delta g = 35\%$ , we see that the cross over happens at around  $\ell = 3000$ . This shows that if we want to set our  $\ell_{\text{max}}$  of our observation higher than  $\ell = 3000$ , a calibration around 1% would be sufficient. A more systematic study is required to establish gain calibration thresholds that contain 2 significant figures for CCAT-prime, SO, and *Planck*.

### C. The Detection of Rayleigh Scattering by Satellite Missions

Ground-based observations are heavily impacted by atmospheric noise. Given the large impact of atmosphere on the detectability of the Rayleigh scattering signal,

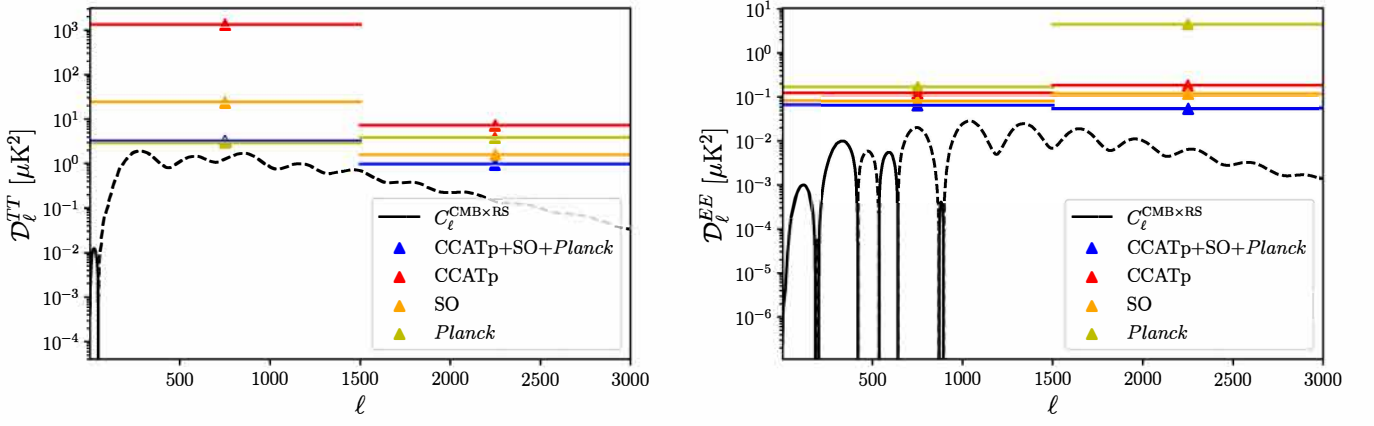


FIG. 6. Forecasted  $2\sigma$  upper limits on the Rayleigh scattering cross-spectrum with CCAT-prime, Simons Observatory LAT+SAT, *Planck*, and the combination of all of these telescopes. The SO SAT only enters for the  $EE$  signal, since it only targets polarized signals. As expected, the combined configuration has the lowest  $2\sigma$  upper limits; *Planck* is most helpful on large scales, and CCAT-prime and SO contribute mainly on small scales.

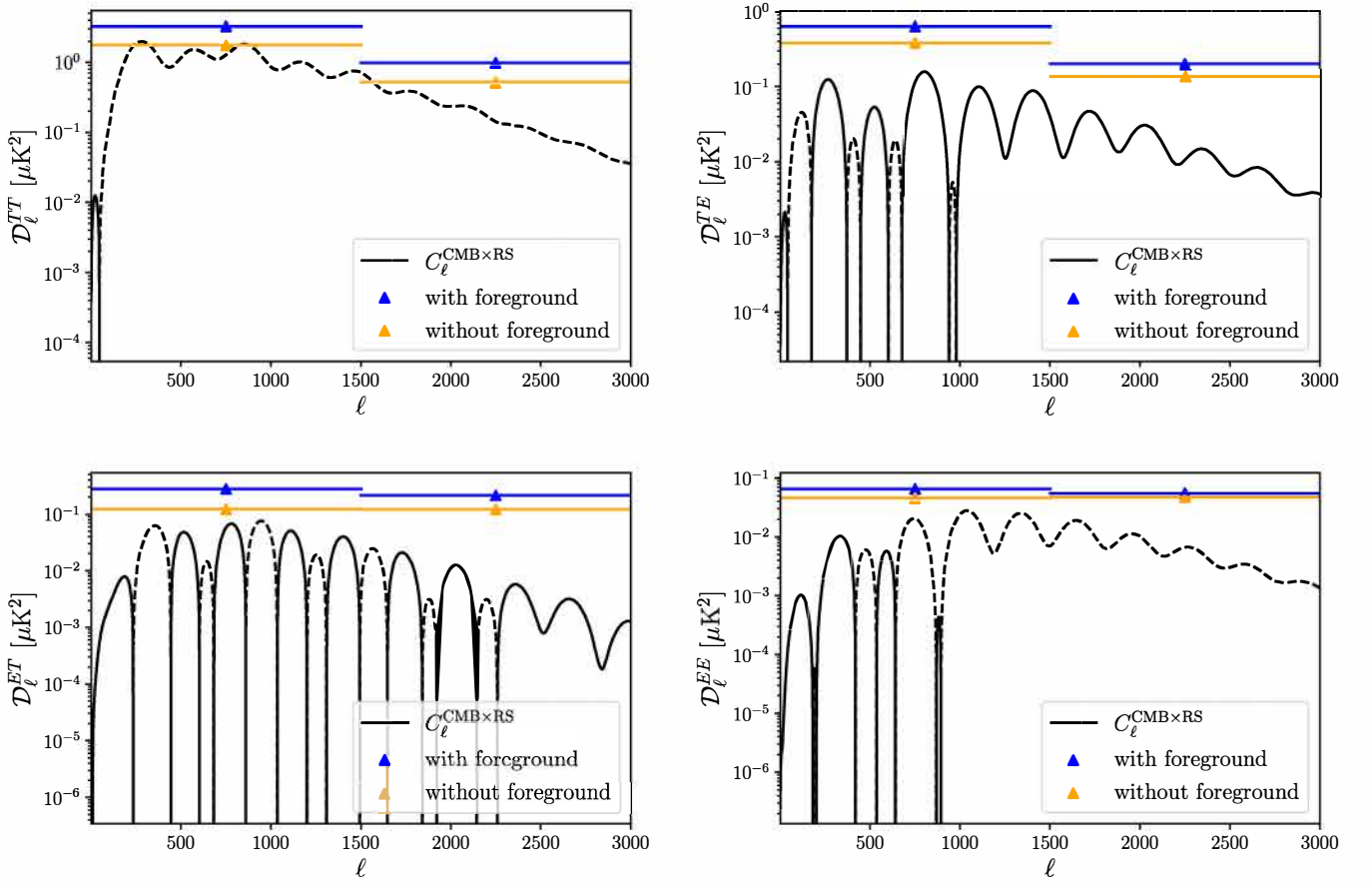


FIG. 7. Forecasted  $2\sigma$  upper limits on the  $TT$  and  $EE$  Rayleigh scattering cross-spectrum with and without foregrounds, demonstrating how the foregrounds impact detectability. The SO SAT only enters for the  $EE$  signal, since it only targets polarized signals.

ideally one would attempt such observations in space where there is no atmosphere and where large scale

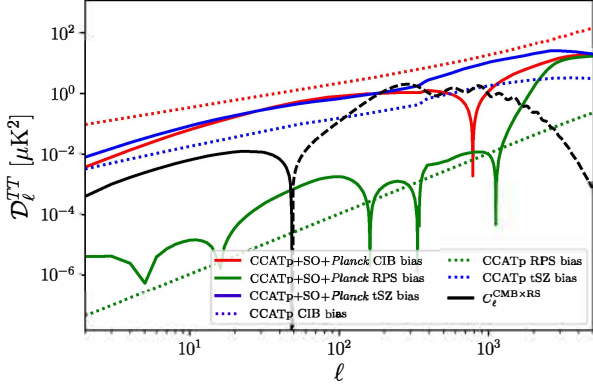


FIG. 8. CIB, tSZ, and radio point source bias for CCAT-prime and for the combined configuration. With only CCAT-prime, CIB is 1 to 2 orders of magnitude stronger than the Rayleigh scattering cross-spectrum at 280 GHz, and radio point source bias is around 1 to 4 orders of magnitude lower than the signal. The tSZ bias is comparable to the signal itself. With the combined configuration of CCAT-prime, SO and *Planck*, we the CIB is more suppressed, but tSZ is less suppressed. In general, the tSZ and CIB biases are mostly comparable or even larger than the cross Rayleigh scattering cross-spectrum at 280 GHz.

anisotropies can be efficiently recovered. We calculate the detectability of the Rayleigh scattering signal for the planned satellite mission LiteBIRD [21] and the proposed mission PICO [23].

The results are summarized in Table. I. The small aperture of LiteBIRD restricts the range of scales it can probe to  $\ell \lesssim 400$ . Including realistic foregrounds, LiteBIRD could reach a  $1.9\sigma$  detection in  $TT$ , which is better than the combined configuration of CCAT-prime, SO, and *Planck*, but insufficient for a significant first detection. PICO would lead to a significant detection even in the presence of foregrounds. With its larger aperture, low noise, and multiple frequency channels that span the millimeter to sub-millimeter wavelength range, we forecast that PICO should achieve a  $85\sigma$  detection in  $TT$  after foregrounds removal. In Fig. 10, we show the error bars for PICO and LiteBIRD. LiteBIRD has much larger error bars than PICO as expected.

The fact that PICO is forecast to significantly detect the Rayleigh scattering signal after removing foregrounds, naturally called for a careful study of the biases induced by foregrounds residuals. Following the method outlined in Sec. VIA we calculate the residual foreground bias for PICO. The biases from tSZ, CIB and radio point sources for PICO are shown in Fig. 11. We find that the CMB-Rayleigh cross-spectrum at 280 GHz is 5 to 6 orders of magnitude larger than the bias of tSZ and radio point sources. It is also 8 to 10 orders of magnitude larger than the bias of the CIB. It is clear that the multitude of frequency channels and low noise levels not only aid in PICO's ability to detect

Rayleigh scattering, but also in its ability to evade residual foreground bias.

The highly significant detection of Rayleigh scattering forecast for PICO additionally leads to the question of calibration uncertainty. Following Sec. VIB we calculate the gain bias for a  $\Delta g = 1\%$ . The results are shown in Fig. 12. At the 1% threshold the gain bias for PICO is 1 to 4 orders of magnitude lower than the reconstructed noise at  $\ell < 3 < 3000$ . This means that a 1% overall calibration is more than sufficient to safely detect Rayleigh scattering without significant biases. Overall, PICO is robust against the systematics studied here for its forecast detection of the Rayleigh scattering signal.

Given LiteBIRD's low forecast SNR we do not calculate the residual foreground and gain biases for it. However, LiteBIRD's sensitivity on large scale could still provide important information when combined with other measurements. We consider the combination LiteBird and *Planck* with CCAT-prime.

LiteBIRD contributes no information for  $\ell \gtrsim 400$ . On large scales  $\ell < 400$ , the reconstructed noise converges to LiteBIRD noise. If we limit the sky fraction to the patch observed by CCAT-prime ( $f_{\text{sky}} = 0.4$ ), we will lose essential information on large scales from LiteBIRD ( $f_{\text{sky}} = 0.7$ ). We calculate the signal-to-noise gained by each telescope in different sky fractions and add them together in quadrature. Essentially, we assign 0 weight to CCAT-prime in  $\ell \leq 400$ , and 0 weight to LiteBIRD in  $\ell > 400$ , which is consistent with the cILC method, i.e.,

$$(\text{SNR})^2 = \sum_{\ell=2}^{400} \left( \frac{(C_{\ell}^{\text{CMB} \times \text{RS}})^2}{\Delta C_{\ell}^2} \right)_{\text{LiteBIRD}} + \sum_{\ell=401}^{3000} \left( \frac{(C_{\ell}^{\text{CMB} \times \text{RS}})^2}{\Delta C_{\ell}^2} \right)_{\text{CCATp+Planck}} \quad (19)$$

With the addition of CCAT-prime and *Planck*, the SNR is 2.2 for the  $TT$  signal in the presence of foregrounds, which is 12.5% better than the individual LiteBIRD configuration.

## VII. CONCLUSION

The observation of CMB Rayleigh scattering has the potential to improve constraints on various cosmological parameters, such as the sum of neutrino masses ( $\sum m_{\nu}$ ), the primordial helium fraction ( $Y_{\text{He}}$ ), and the light relic density ( $N_{\text{eff}}$ ) [7, 11]. However, before one can utilize the full potential of the Rayleigh scattering signal for cosmological parameter inference, a first detection is needed. In this paper, we forecast the detectability of the Rayleigh signal in the presence of foregrounds, considering CCAT-prime, Simons Observatory, *Planck*, PICO and LiteBIRD, and applying a cILC to mitigate various foregrounds contamination. We explored combinations of those experiments to minimize the

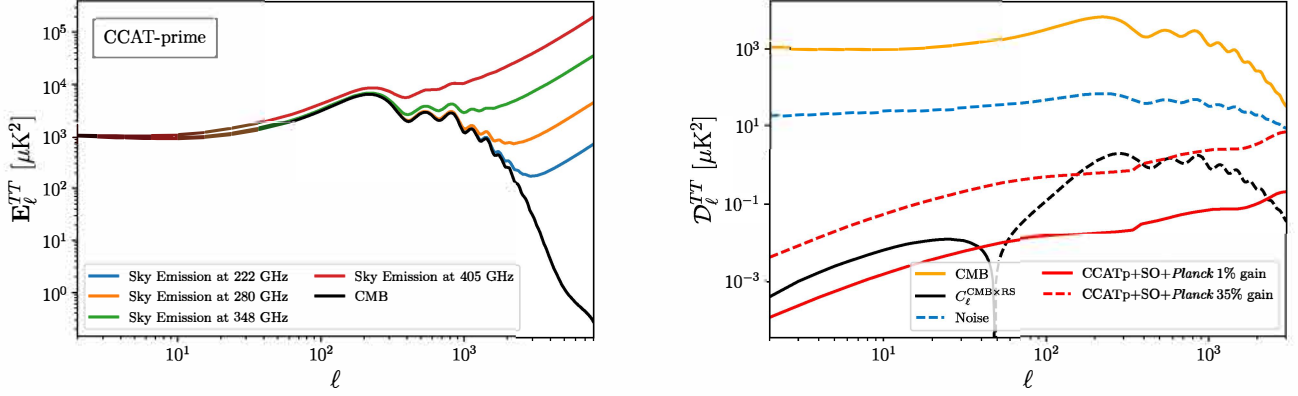


FIG. 9. The left panel shows the sky emission on each channel of CCAT-prime. We omitted the 850 GHz channel because its noise is so high that it is given a nearly zero weight by our cILC. We can see that at low frequencies, the sky emission is close to the CMB itself, but in higher frequency channels, the foreground terms start to dominate. The right panel is the gain model for the combined configuration of CCAT-prime, SO, and *Planck*. We assumed  $\Delta g = 1\%$  gain from sky emission, and we can see that it is roughly on the same scale as the Rayleigh signal. With CCAT-prime channels, in lower frequency channels, the gain looks like CMB itself, but in the higher frequency channels, the gain gradually flattens out due to the dominance of radio point sources. The gain for the combined configuration is approximately 3 to 4 orders of magnitude lower than the noise power spectrum, meaning a 1% gain calibration would be sufficient for our purpose. We also showed that a 35% gain calibration, in dashed red line, would be comparable to the noise power spectrum at high  $\ell$ .

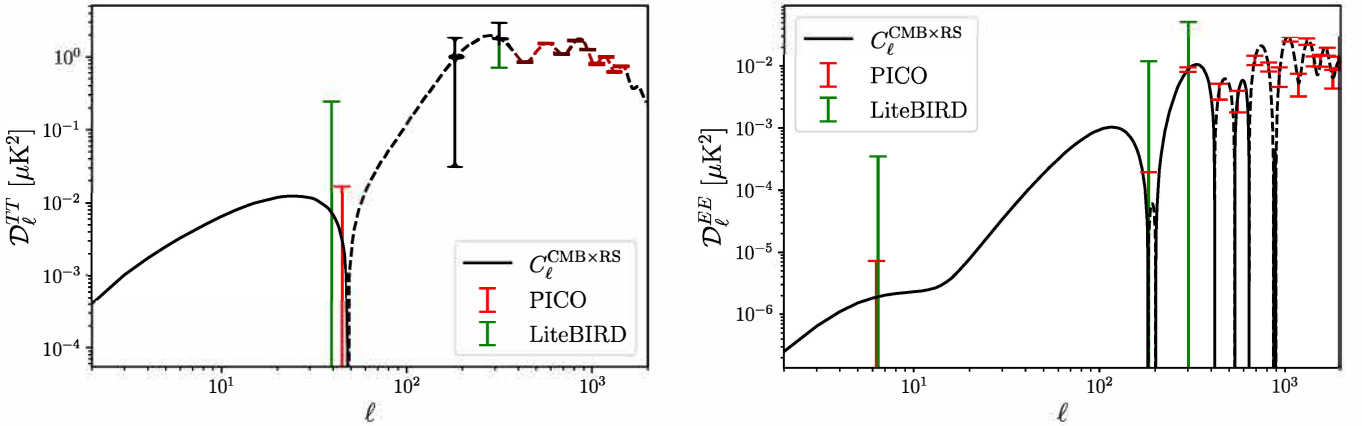


FIG. 10. Noise comparison for PICO and LiteBIRD. PICO has much tighter error bars, especially on small scales, which will enable a high significance measurement of Rayleigh scattering. LiteBIRD has larger error bars, but can still contribute information on large scales.

reconstructed noise and determine the effectiveness of these configurations.

Our analysis shows that in the presence of foregrounds a combination of CCAT-prime, SO and *Planck* can reach an SNR of 1.1 for the  $TT$  CMB-Rayleigh cross-spectrum. As expected, this is below previous estimates which were derived in the absence of foregrounds. This is a clear demonstration that foreground removal, as well as atmospheric contamination, have a large impact on detectability. We compared the capability of these ground-based experiments to future satellite

missions. PICO is capable of an  $85\sigma$  detection of the  $TT$  CMB-Rayleigh cross-spectrum even in the presence of foregrounds. LiteBIRD can potentially achieve an SNR of 1.9 of the same cross signal. When combined with *Planck* and CCAT-prime, LiteBIRD's SNR is improved to 2.2. We found that the signal-to-noise of  $EE$  cross-spectrum is typically the second highest among all four cross-correlations. While the  $EE$  spectrum has the weakest signal strength, it also has the lowest cILC reconstruction noise, since polarized foregrounds are weaker. This also explains why the  $EE$  signal-to-noise

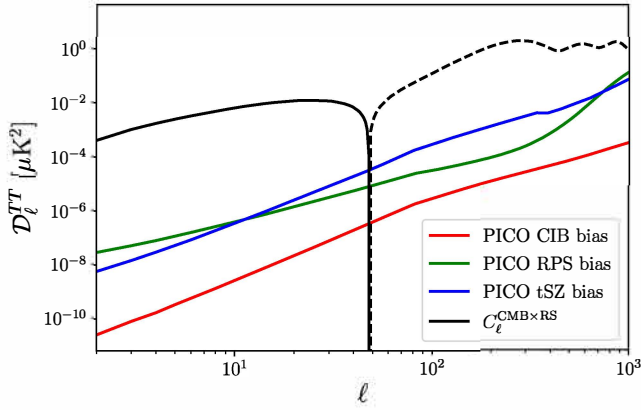


FIG. 11. Bias of tSZ, CIB, and radio point sources for PICO. The bias of tSZ, CIB and radio point sources are all lower than the CMB-Rayleigh cross-spectrum at 280 GHz by several orders of magnitude.

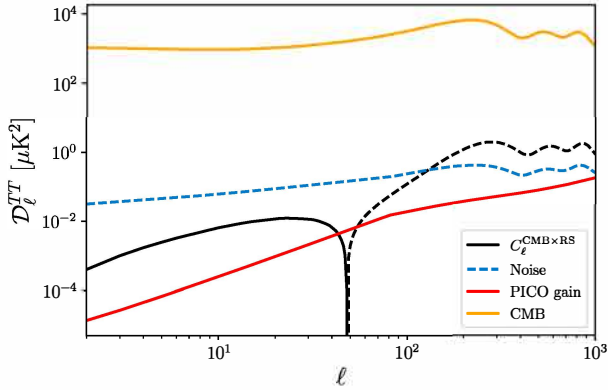


FIG. 12. Gain of PICO. The gain of  $\Delta g = 1\%$  of sky emission is around 3 to 4 orders of magnitude lower than the noise power spectrum. This means that a 1% gain calibration would be sufficient for PICO.

is least impacted by the presence of foregrounds.

Looking beyond the SNR of Rayleigh scattering, we calculated the biases introduced by tSZ, CIB, and radio point source residuals for the combined configuration of CCAT-prime, SO, and *Planck* and for PICO. For CCAT-prime, SO, and *Planck* the biases from tSZ and CIB have a similar magnitude or are larger than the Rayleigh signal at  $\nu = 280$  GHz. Thus a detection of Rayleigh scattering requires further foreground cleaning, for example the deprojection of additional foreground components. Similarly, a denser and broader frequency coverage would help. Indeed for PICO, we found that none of those foregrounds introduce a significant bias which would impact a detection.

Furthermore, we estimate the required instrumental calibration to prevent errors in the gain that may impact the detectability of Rayleigh scattering. We found that an overall gain calibration error of  $\Delta g =$

1% is sufficient for CCAT-prime, SO, and *Planck* and PICO. Nevertheless, a more comprehensive study of the calibration of CCAT-prime, SO, and PICO is still warranted and practical for upcoming analysis.

With CCAT-prime, SO, and *Planck*, a first detection is unlikely to be achievable by relying solely on constrained-ILC methods. Due to the blind nature of such methods, they do not allow us to leverage the scale dependence of the Rayleigh scattering signal. Indeed, for a fixed cosmology, the Rayleigh scattering signal is straightforward to model, which would be an incentive to explore more parametric approaches such as maximum likelihood algorithm for parametric component separation [3, 35]. More realistic analysis of instrumental calibration requirements will also be necessary and will be the topic of future investigations. Telescopes not considered here, such as SPT [10], CMB-S4 [1] and CMB-HD [34], are guaranteed to help in an attempt to make a first detection of CMB Rayleigh scattering.

The detection of the Rayleigh signal is an exciting prospect which will eventually allow us to observe a frequency-dependent last-scattering surface. With advancing telescope technology and new data analysis techniques, there is hope to make a first detection of Rayleigh signal this decade. Eventually a significant detection will allow us to use the cosmological information hidden in the Rayleigh signal to improve parameter constraints beyond those achievable from the primary CMB alone.

## ACKNOWLEDGEMENTS

We thank Jens Chluba, Tom Crawford, and Blake Sherwin for their useful comments on our manuscript. NB acknowledges support from NSF grant AST-1910021 and NASA grants 21-ADAP21-0114 and 21-ATP21-0129. PDM acknowledges support from the Netherlands organisation for scientific research (NWO) VIDI grant (dossier 639.042.730). JM is supported by the US Department of Energy under Grant DE-SC0010129. BB acknowledges postdoctoral support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 849169). SKC acknowledges support from NSF award AST-2001866.

## Appendix A: Experimental configurations

The following tables summarize the noise models for all experiments considered in this work. More details can be found in Sec. IV.



SO LAT ( $f_{\text{sky}} = 0.4$ ,  $\ell_{\text{knee}}^T = 1000$ ,  $\alpha_{\text{knee}}^T = -3.5$ ,  
 $\ell_{\text{knee}}^P = 700$ ,  $\alpha_{\text{knee}}^P = -1.4$ )

| Freq (GHz) | Beam (arcmin) | $N_{\text{white}}(\mu\text{K}^2)$ | $N_{\text{red}}(\mu\text{K}^2\text{s})$ |
|------------|---------------|-----------------------------------|---|
| 27         | 7.4           | $2.3 \times 10^{-4}$              | 100                                     |
| 39         | 5.1           | $6.2 \times 10^{-5}$              | 39                                      |
| 93         | 2.2           | $2.8 \times 10^{-6}$              | 230                                     |
| 145        | 1.4           | $3.6 \times 10^{-6}$              | 17000                                   |
| 225        | 1.0           | $1.9 \times 10^{-5}$              | 17000                                   |
| 280        | 0.9           | $1.16 \times 10^{-4}$             | 31000                                   |

TABLE II. Survey configuration of Simons Observatory Large Aperture Telescope (SO LAT) taken from Ade et al. [5].

SO SAT ( $f_{\text{sky}} = 0.1$ )

| Freq (GHz) | Beam (arcmin) | $N_{\text{white}} = N_{\text{red}}(\mu\text{K}^2)$ | $\ell_{\text{knee}}^P$ | $\alpha_{\text{knee}}^P$ |
|------------|---------------|--|------------------------|--------------------------|
| 27         | 91            | $5.3 \times 10^{-5}$                               | 15                     | -2.4                     |
| 39         | 63            | $2.4 \times 10^{-5}$                               | 15                     | -2.4                     |
| 93         | 30            | $3.0 \times 10^{-7}$                               | 25                     | -2.5                     |
| 145        | 17            | $3.7 \times 10^{-7}$                               | 25                     | -3.0                     |
| 225        | 11            | $1.5 \times 10^{-6}$                               | 35                     | -3.0                     |
| 280        | 9             | $8.5 \times 10^{-6}$                               | 40                     | -3.0                     |

TABLE III. Survey configuration of Simons Observatory Small Aperture Telescope (SO SAT) taken from Ade et al. [5].

CCAT-prime ( $f_{\text{sky}} = 0.4$ ,  $\ell_{\text{knee}}^T = 1000$ ,  $\alpha_{\text{knee}}^T = -3.5$ ,  
 $\ell_{\text{knee}}^P = 700$ ,  $\alpha_{\text{knee}}^P = -1.4$ )

| Freq (GHz) | Beam (arcsec) | $N_{\text{white}}(\mu\text{K}^2)$ | $N_{\text{red}}(\mu\text{K}^2)$ |
|------------|---------------|-----------------------------------|---------------------------------|
| 220        | 57            | $1.8 \times 10^{-5}$              | $1.6 \times 10^{-2}$            |
| 280        | 45            | $6.4 \times 10^{-5}$              | $1.1 \times 10^{-1}$            |
| 350        | 35            | $9.3 \times 10^{-4}$              | $2.7 \times 10^0$               |
| 410        | 30            | $1.2 \times 10^{-2}$              | $1.7 \times 10^1$               |
| 850        | 14            | $2.8 \times 10^4$                 | $6.1 \times 10^6$               |

TABLE IV. Survey configuration of CCAT-prime taken from CCAT-Prime collaboration et al. [12].

Planck ( $f_{\text{sky}} = 0.6$ )

| Freq (GHz) | Beam (arcmin) | $N_{\text{white}}(\mu\text{K}^2)$ |
|------------|---------------|-----------------------------------|
| 100        | 9.7           | $5.07 \times 10^{-4}$             |
| 143        | 7.2           | $9.21 \times 10^{-5}$             |
| 217        | 4.9           | $1.85 \times 10^{-4}$             |
| 353        | 4.9           | $2.00 \times 10^{-3}$             |

TABLE V. Survey configuration of Planck taken from Planck Collaboration et al. [31].

LiteBIRD ( $f_{\text{sky}} = 0.7$ )

| Freq (GHz) | Beam (arcmin) | $N_{\text{white}}(\mu\text{K}^2)$ |
|------------|---------------|-----------------------------------|
| 40         | 69            | $2.38 \times 10^{-4}$             |
| 50         | 56            | $9.75 \times 10^{-5}$             |
| 60         | 48            | $6.70 \times 10^{-5}$             |
| 68         | 43            | $4.44 \times 10^{-4}$             |
| 78         | 39            | $3.08 \times 10^{-5}$             |
| 89         | 35            | $2.32 \times 10^{-5}$             |
| 100        | 29            | $1.43 \times 10^{-5}$             |
| 119        | 25            | $9.77 \times 10^{-6}$             |
| 140        | 23            | $5.89 \times 10^{-6}$             |
| 166        | 21            | $7.15 \times 10^{-6}$             |
| 195        | 40            | $5.69 \times 10^{-6}$             |
| 235        | 19            | $1.00 \times 10^{-5}$             |
| 280        | 24            | $2.95 \times 10^{-5}$             |
| 337        | 20            | $6.44 \times 10^{-5}$             |
| 402        | 17            | $2.38 \times 10^{-4}$             |

TABLE VI. Survey configuration of LiteBIRD taken from Ghigna et al. [21].

PICO ( $f_{\text{sky}} = 0.7$ )

| Freq (GHz) | Beam (arcmin) | $N_{\text{white}}(\mu\text{K}^2)$ |
|------------|---------------|-----------------------------------|
| 21         | 38.4          | $1.21 \times 10^{-5}$             |
| 25         | 32.0          | $7.15 \times 10^{-6}$             |
| 30         | 28.3          | $3.20 \times 10^{-6}$             |
| 36         | 23.6          | $1.33 \times 10^{-6}$             |
| 43         | 22.2          | $1.33 \times 10^{-6}$             |
| 52         | 18.4          | $6.77 \times 10^{-7}$             |
| 62         | 12.8          | $6.11 \times 10^{-7}$             |
| 75         | 10.7          | $3.81 \times 10^{-7}$             |
| 90         | 9.5           | $1.69 \times 10^{-7}$             |
| 108        | 7.9           | $1.08 \times 10^{-7}$             |
| 129        | 7.4           | $9.52 \times 10^{-8}$             |
| 159        | 6.2           | $7.15 \times 10^{-8}$             |
| 186        | 4.3           | $3.32 \times 10^{-7}$             |
| 223        | 3.6           | $4.33 \times 10^{-7}$             |
| 268        | 3.2           | $2.05 \times 10^{-7}$             |
| 321        | 2.6           | $3.81 \times 10^{-7}$             |
| 385        | 2.5           | $4.33 \times 10^{-7}$             |
| 462        | 2.1           | $1.73 \times 10^{-6}$             |
| 555        | 1.5           | $4.44 \times 10^{-5}$             |
| 666        | 1.3           | $6.61 \times 10^{-4}$             |
| 799        | 1.1           | $2.32 \times 10^{-2}$             |

TABLE VII. Survey configuration of PICO taken from Hanany et al. [23].

- [2] Abitbol, M. H. et al. 2021, *Journal of Cosmology and Astroparticle Physics*, 2021, 032
- [3] Abylkairov, Y. S., Darwish, O., Hill, J. C., & Sherwin, B. D. 2021, *Physical Review D*, 103, 103510
- [4] Addison, G. E., Dunkley, J., & Spergel, D. N. 2012, *MNRAS*, 427, 1741
- [5] Ade, P. et al. 2019, *J. Cosmology Astropart. Phys.*, 2, 056
- [6] Aghanim, N., et al. 2020, *Astron. Astrophys.*, 641, A6, [Erratum: *Astron. Astrophys.* 652, C4 (2021)]
- [7] Alipour, E., Sigurdson, K., & Hirata, C. M. 2015, *Phys. Rev. D*, 91, 083520
- [8] Battaglia, N., Bond, J. R., Pfrommer, C., Sievers, J. L., & Sijacki, D. 2010, *ApJ*, 725, 91
- [9] Bedini, L., Herranz, D., Salerno, E., Baccigalupi, C., Kuruoglu, E. E., & Tonazzini, A. 2005, *EURASIP Journal on Applied Signal Processing*, 2005, 2400
- [10] Benson, B. A. et al. 2014, in *Proc. SPIE*, Vol. 9153, *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, 91531P
- [11] Beringue, B., Meerburg, P. D., Meyers, J., & Battaglia, N. 2020, *arXiv e-prints*, arXiv:2008.11688
- [12] CCAT-Prime collaboration et al. 2021, *arXiv e-prints*, arXiv:2107.10364
- [13] Choi, S. K. et al. 2020, *Journal of Low Temperature Physics*, 199, 1089–1097
- [14] Choi, S. K., & Page, L. A. 2015, *J. Cosmology Astropart. Phys.*, 2015, 020
- [15] Delabrouille, J., Cardoso, J. F., Le Jeune, M., Betoule, M., Fay, G., & Guilloux, F. 2009, *A&A*, 493, 835
- [16] Delabrouille, J., Cardoso, J. F., & Patanchon, G. 2003, *MNRAS*, 346, 1089
- [17] Dibert, K., et al. in prep
- [18] Dunkley, J. et al. 2013, *J. Cosmology Astropart. Phys.*, 7, 025
- [19] Dvorkin, C., et al. 2022, in *2022 Snowmass Summer Study*
- [20] Dünner, R. et al. 2012, *The Astrophysical Journal*, 762, 10
- [21] Ghigna, T., Matsumura, T., Patanchon, G., Ishino, H., & Hazumi, M. 2020, *J. Cosmology Astropart. Phys.*, 2020, 030
- [22] Green, D. et al. 2019, in *BAAS*, Vol. 51, 159
- [23] Hanany, S. et al. 2019, *arXiv:1902.10541*, *arXiv:1902.10541*
- [24] Hu, W., Sugiyama, N., & Silk, J. 1997, *Nature*, 386, 37–43
- [25] Knox, L. 1995, *Physical Review D*, 52, 4307–4318
- [26] Lewis, A. 2013, *Journal of Cosmology and Astro-Particle Physics*, 2013, 053
- [27] Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, 538, 473
- [28] Maino, D. et al. 2002, *MNRAS*, 334, 53
- [29] Pires, S., Juin, J. B., Yvon, D., Moudden, Y., Anthoine, S., & Pierpaoli, E. 2006, *A&A*, 455, 741
- [30] Planck Collaboration et al. 2014, *A&A*, 571, A12
- [31] —. 2018, *arXiv e-prints*, arXiv:1807.06209
- [32] Remazeilles, M., & Chluba, J. 2018, *MNRAS*, 478, 807
- [33] Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, *MNRAS*, 410, 2481
- [34] Sehgal, N. et al. 2019, in *Bulletin of the American Astronomical Society*, Vol. 51, 6
- [35] Stomp, R., Leach, S., Stivoli, F., & Baccigalupi, C. 2009, *Monthly Notices of the Royal Astronomical Society*, 392, 216
- [36] Terry, H. et al. 2019, in *Bulletin of the American Astronomical Society*, Vol. 51, 213
- [37] Yu, Q., Spergel, D. N., & Ostriker, J. P. 2001, *ApJ*, 558, 23