

# Framework for analysis of next generation, polarized CMB data sets in the presence of Galactic foregrounds and systematic effects

Clara Vergès<sup>1</sup>, Josquin Errard<sup>1</sup> and Radek Stompor<sup>1,2</sup>

<sup>1</sup>*Université de Paris, CNRS, Astroparticule et Cosmologie, F-75006 Paris, France*

<sup>2</sup>*CNRS-UCB International Research Laboratory, “Centre Pierre Binétruy”, UMI2007, CPB-IN2P3, France*

 (Received 25 September 2020; accepted 4 February 2021; published 10 March 2021)

Reaching the sufficient sensitivity to detect primordial B-modes requires modern CMB polarization experiments to rely on new technologies, necessary for the deployment of arrays of thousands of detectors with a broad frequency coverage and operating them for extended periods of time. This increased complexity of experimental design unavoidably introduces new instrumental and systematic effects, which in turn may impact performance of the new instruments. In this work we extend the standard data analysis pipeline by including a (parametric) model of instrumental effects directly in the data model. We then correct for them in the analysis, accounting for the additional uncertainty in the final results. We embed these techniques within a general, end-to-end formalism for estimating the impact of the instrument and foreground models on constraints on the amplitude of the primordial B-mode signal. We focus on the parametric component separation approach which we generalize to allow for simultaneous estimation of instrumental and foreground parameters. We demonstrate the framework by studying in detail the effects induced by an achromatic half-wave plate (HWP), which lead to a frequency-dependent variation of the instrument polarization angle, and experimental bandpasses which define observational frequency bands. We assume a typical Stage-3 CMB polarization experiment, and show that maps recovered from raw data collected at each frequency band will unavoidably be linear mixtures of the  $Q$  and  $U$  Stokes parameters. We then derive a new generalized data model appropriate for such cases, and extend the component separation approach to account for it. We find that some of the instrumental parameters, in particular those describing the HWP, can be successfully constrained by the data themselves without need for external information, while others, like bandpasses, need to be known with good precision in advance. We also show that if the assumed model is an accurate description of the HWP and we have sufficient information about bandpasses, our approach can successfully correct for these effects. In particular, we recover the tensor-to-scalar ratio  $r$  down to the precision levels typical of Stage-3 experiments, which aim at setting a ( $1\sigma$ ) upper limit on  $r$  of  $\sim 10^{-3}$  if  $r = 0$ .

DOI: [10.1103/PhysRevD.103.063507](https://doi.org/10.1103/PhysRevD.103.063507)

## I. INTRODUCTION

Most of current and future cosmic microwave background (CMB) experiments aim at detecting large scale primordial B-modes and measuring the tensor-to-scalar ratio  $r$  as low as  $r \sim 0.001$ . This detection requires unprecedented instrumental sensitivity, calling for deployment of many thousands of photon-noise limited detectors, along with new technologies at many stages of the detection chain. This increased complexity will introduce new instrumental and systematic effects, which will need to be characterized, accounted for and most likely mitigated. Moreover, the B-modes signal is overshadowed by galactic polarised emission due to thermal dust emission and synchrotron radiation. Separating sky components and isolating the CMB signal is one of the key challenges in the search for cosmological B-modes. This requires the deployment of multifrequency experiments and multichroic focal planes to take advantage of the fact that different sky components do not have the same frequency scaling.

In this context, it is crucial to develop efficient tools for processing next generation CMB polarization experiments datasets, to ensure that these experiments achieve the required performance. It is also essential to develop reliable techniques for forecasting the performance of the future experiments, so we can be better prepared for deployment and data analysis.

CMB data analysis is a complex process composed of many steps. One of its milestones is map-making, that aims at reconstructing single frequency intensity and polarization maps from raw (time-domain) data collected by a CMB instrument. Another one is component separation, that aims at separating the CMB emission from that of galactic foregrounds. Component separation methods proposed to date, e.g., [1–5], typically use maps as produced during the map-making as inputs, and component separation operations are performed in the pixel (or equivalently harmonic) domain. However, many of the instrumental

effects can be modeled reliably only in the raw data (time) domain. In this work we therefore consider both steps, map-making and component separation, which we extend by including explicitly models of instrumental systematic effects. We then attempt to correct for them as part of the analysis process. For definiteness, we focus hereafter on a pixel-based parametric component separation technique [6,7] and implement the extensions within the previously validated framework for parametric component separation, XFORECAST [8,9]. In order to demonstrate this general approach, we study in detail the interplay between HWP frequency-dependent characteristics, frequency bandpasses, and galactic foreground emission laws, and their impact on the precision of the primordial  $B$ -modes detection.

Frequency-dependent, instrumental effects are particularly insidious for CMB polarization instruments as they are likely to have significant impact on the component separation procedure, potentially leading to large foreground residuals left in the cleaned CMB map. This could bias the derived constraints on the cosmological signal, in particular the tensor-to-scalar ratio  $r$  that we seek to measure. We note that while in our worked example we showcase the proposed methodology in the context of performance forecasting, the proposed approach is applicable to actual analysis of the forthcoming CMB datasets.

We derive a new generalized time-domain data model in Sec. II, where we also discuss its consequences for the map-making procedure and its output. In Sec. III we elaborate on the parametric component separation procedure and present the self-consistent, end-to-end framework adapted to features of the new data model. We validate and demonstrate the framework in the context of performance forecasting in Sec. IV.

## II. INSTRUMENTAL FRAMEWORK

The instrument we consider in this work is composed of a continuously rotating multilayer half-wave plate and sinuous antennas coupled to total power detectors, such as transition edge sensors (TES) bolometers. In this section, we detail mathematical models describing such an instrument, and use them to derive a corresponding data model of its measurements. We then discuss its consequences for the map-making procedure.

### A. Mueller matrix formalism

The Mueller matrix formalism is widely used to model polarizers in general, and HWP in particular (see for example [10–12]). Incoming light is described by the four Stokes parameters:

$$\mathbf{I} \equiv \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \quad (1)$$

Mueller matrices are  $(4 \times 4)$  operators acting on Stokes parameters and describing the impact of different stages of the detection chain on the state of the incoming light.

### 1. Half-wave plate

A single-layer half-wave plate (HWP) is a retarder made of a bi-refracting material that introduces a phase shift of  $\pi$  between the input and output polarization states. However, this relationship is true only for a given frequency (monochromatic light). In more generic cases where the light is poly-chromatic, a HWP introduces a phase  $\delta$  between polarization components, which depends on observing frequency. The general Mueller matrix for a single-layer HWP can be written as:

$$\mathbf{M}_{\text{layer}} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \delta & -\sin \delta \\ 0 & 0 & \sin \delta & \cos \delta \end{pmatrix}, \quad (2)$$

where the phase shift  $\delta$  is given by:

$$\delta \equiv \frac{2\pi\theta_{\text{hwp}}|n_o - n_e|\nu}{c}, \quad (3)$$

where  $c$  refers to the usual speed of light in vacuum. Through Eq. (3) defining  $\delta$ , coefficients of the HWP Mueller matrix depend on thickness of the bi-refracting layer  $\theta_{\text{hwp}}$ , optical indices of the ordinary and extraordinary axes of bi-refracting material  $n_o$  and  $n_e$ , and the frequency of observation  $\nu$ .

For typical CMB experiments observing in the millimeter range, achromatic HWP are made of several layers of bi-refracting materials, such as sapphire [13]. Each layer can be modeled as in Eq. (2).

In this work, we neglect reflections at the interface between two stacked layers, as well as effects of slant incidence angles. We thus model an achromatic HWP as a perfect stack of layers, each layer being rotated with respect to the reference frame of the instrument by an angle  $\alpha_i$ :

$$\mathbf{M}_{\text{HWP}} = \prod_{i=1}^{n_{\text{layers}}} \mathbf{R}(-2\alpha_i) \mathbf{M}_{\text{layer},i} \mathbf{R}(2\alpha_i) \quad (4)$$

where  $\mathbf{R}$  is a rotation matrix.

For the HWP is continuously spinning around its axis, we denote the rotation angle between its axis and the reference frame of the instrument by  $\varphi_t$  (see Fig. 1), and we have:

$$\mathbf{M}_{\text{rotating HWP}} = \mathbf{R}(-2\varphi_t) \mathbf{M}_{\text{HWP}} \mathbf{R}(2\varphi_t). \quad (5)$$

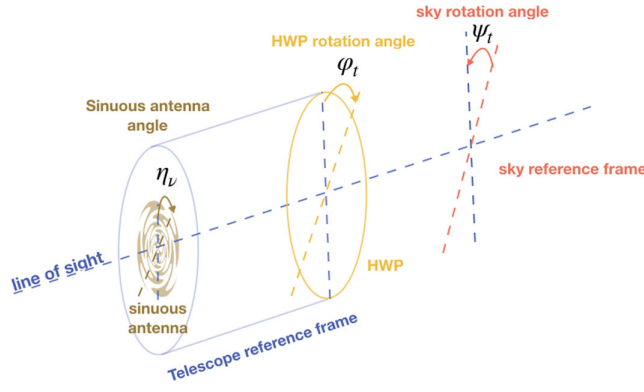


FIG. 1. Depiction of sky rotation angle,  $\psi_t$ , defined as an angle between the sky reference frame and the telescope reference frame, and of HWP rotation  $\varphi_t$  and sinuous antenna  $\eta_\nu$  angles, both assumed to be measured with respect to the telescope reference frame.

## 2. Sinuous antennas

The sky signal modulated by the HWP falls then onto an antenna coupled to a detector pair. The antenna separates the orthogonal polarization states of the incident light which are then measured by the detectors. Several future CMB experiments will deploy sinuous antennas, which are broad frequency-band, but introduce a frequency-dependent rotation of the polarization angle of the incident light [14].

Following [14], we model a sinuous antenna as a usual double-slot antenna (equivalent to a grid), rotated by a frequency-dependent angle  $\eta_\nu$  (see Figure 1), and write its Mueller matrix:

$$\mathbf{M}_{\text{antenna}} = \mathbf{R}(-2\eta_\nu)\mathbf{M}_{\text{grid}}\mathbf{R}(2\eta_\nu), \quad (6)$$

where [14]

$$\eta_\nu[\text{deg}] \simeq 4.9 \sin(12 \log(\nu) + 4.7) \quad (7)$$

and

$$\mathbf{M}_{\text{grid}} \equiv \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (8)$$

The complete Mueller matrix of this optics system can therefore be written as:

$$\mathbf{M}_{\text{optics}} = \mathbf{M}_{\text{antenna}}\mathbf{R}(-2\varphi_t)\mathbf{M}_{\text{HWP}}\mathbf{R}(2\varphi_t). \quad (9)$$

This matrix is the transfer function of the optical system in the Stokes parameter domain, applied to the incoming light represented by the four Stokes parameters, Eq. (1). In what follows, for definiteness, we assume that circular

polarization of the incident light is negligible and we set it to zero, i.e.,  $V_{\text{in}} = 0$ . This is consistent with the expectation that for the relevant cosmological, astrophysical, or atmospheric signals,  $V$  is either zero or negligible [15,16].

## B. Time domain data model

During observations, CMB telescopes scan the sky and their orientation with respect to the sky changes. This introduces a time-dependent rotation angle,  $\psi_t$ , defining the position of the instrument with respect to sky coordinates (see Fig. 1), modifying the Mueller matrix:

$$\mathbf{M}_{\text{tot}} = \mathbf{M}_{\text{optics}}\mathbf{R}(2\psi_t). \quad (10)$$

### 1. Single layer HWP

In the simple case of a monochromatic single layer rotating HWP with grid antenna, the Mueller matrix of the optics system reads:

$$\mathbf{M}_{\text{optics,mono}} = \mathbf{M}_{\text{grid}}\mathbf{R}(-2\varphi_t)\mathbf{M}_{\text{layer}}(\delta = \pi)\mathbf{R}(2\varphi_t), \quad (11)$$

what is a special case of Eq. (9). Taking into account sky rotation as in Eq. (10), the time domain data model for an instrument pointing in the direction  $\gamma$  at time  $t$ , can be written as:

$$\mathbf{m}_t(\nu) = I(\gamma_t, \nu) + \cos(4\varphi_t + 2\psi_t)Q(\gamma_t, \nu) + \sin(4\varphi_t + 2\psi_t)U(\gamma_t, \nu) + \mathbf{n}_t. \quad (12)$$

This is the standard data model of a HWP-modulated CMB measurement, as assumed in the standard map-making procedure. While all three Stokes parameter maps are mixed together in every single measurement, we can separate them from each other owing to a different time dependence of their respective coefficients: the signals modulated in Eq. (12) as  $\cos(4\varphi_t + 2\psi_t)$  are those of the  $Q$  parameter, those modulated as  $\sin(4\varphi_t + 2\psi_t)$  of  $U$ , and the nonmodulated ones of the total intensity. This separation is done during the map-making procedure which thus produces three sky maps of  $I$ ,  $Q$ , and  $U$  for each frequency band.

### 2. Multilayer HWP

In a more general case of the multilayer HWP, the detected signal can be expressed as:

$$\mathbf{m}_t(\nu) = \mathbf{M}_{00}(\nu)I(\gamma_t, \nu) + \mathbf{M}_{01}(\nu, \varphi_t, \psi_t)Q(\gamma_t, \nu) + \mathbf{M}_{02}(\nu, \varphi_t, \psi_t)U(\gamma_t, \nu) + \mathbf{n}_t \quad (13)$$

where  $\mathbf{M}_{0i}$  denotes the elements of the first row of  $\mathbf{M}_{\text{tot}}$ , Eq. (10). These elements can be represented as:

$$\mathbf{M}_{0i}(\varphi_t, \psi_t; \nu, \nu) = \sum_{k=0,4} \mathbf{C}_{0i;k}(\nu) \cos(k\varphi_t + 2\psi_t) + \sum_{k=0,4} \mathbf{S}_{0i;k}(\nu) \sin(k\varphi_t + 2\psi_t), \quad (14)$$

where we have introduced  $\mathbf{C}_{0i;k}$  (resp.  $\mathbf{S}_{0i;k}$ ), the coefficients of the cosine (resp. sine) modulated terms, which are

$$\begin{aligned} \mathbf{m}_t(\nu) \equiv & \mathbf{n}_t + \mathbf{M}_{00}(\nu)\mathbf{I}(\gamma_t, \nu) \\ & + [\mathbf{C}_{01;0}(\nu)\mathbf{Q}(\gamma_t, \nu) + \mathbf{C}_{02;0}(\nu)\mathbf{U}(\gamma_t, \nu)] \times \cos 2\psi_t \\ & + [\mathbf{S}_{01;0}(\nu)\mathbf{Q}(\gamma_t, \nu) + \mathbf{S}_{02;0}(\nu)\mathbf{U}(\gamma_t, \nu)] \times \sin 2\psi_t \\ & + [\mathbf{C}_{01;4}(\nu)\mathbf{Q}(\gamma_t, \nu) + \mathbf{C}_{02;4}(\nu)\mathbf{U}(\gamma_t, \nu)] \times \cos(4\phi_t + 2\psi_t) \\ & + [\mathbf{S}_{01;4}(\nu)\mathbf{Q}(\gamma_t, \nu) + \mathbf{S}_{02;4}(\nu)\mathbf{U}(\gamma_t, \nu)] \times \sin(4\phi_t + 2\psi_t). \end{aligned} \quad (15)$$

This expression highlights the fact that in the case of a multilayer HWP, Stokes parameters  $Q$  and  $U$  of the sky signal are not simply modulated at  $\cos(4\phi_t + 2\psi_t)$  and  $\sin(4\phi_t + 2\psi_t)$ , respectively. Instead, and in contrast to the case of the simple polarized data model in Eq. (12), the terms modulated in this way are composed of mixtures of these two Stokes parameters. We refer to those signals hereafter as *mixed-Stokes single frequency signals*. We also note the presence of two extra terms which are modulated by the sky angle only. Sky signals they correspond to constitute two additional, independent combinations of the Stokes  $Q$  and  $U$  parameters. As the coefficients  $\mathbf{C}_{0i;k}$  and  $\mathbf{S}_{0i;k}$ , depend on instrumental parameters (see Appendix A) so do the mixed-Stokes signals, which are therefore specific to a given experiment.

As mentioned earlier, the map-making procedure enables an estimation of the terms with different time dependence, assuming that those dependencies are such that they are linearly independent when limited to observations of a single sky pixel for every such a pixel. If these conditions are met, the map-making procedure applied to the data model in Eq. (15) could recover for each frequency channel a map of total intensity accompanied by four maps composed of different linear combinations of Stokes  $Q$  and  $U$  parameters. These would be those maps which could and should be considered as inputs to the next data processing stages and specifically that of the component separation. We explore some of the consequences of this fact in the remainder of this article.

One important implication of the more involved data model in Eq. (15) is that solving the complete map-making problem in this case will in general require sufficient redundancy in observations of every sky pixel, with a number of different HWP as well as sky rotation angles. This can have important impact on scan designs of the future CMB experiments.

linear combinations of elements of the Mueller matrix of the system,  $\mathbf{M}_{0i}$ . They therefore depend explicitly on instrumental, i.e., here HWP and sinuous antennas parameters, as well as observational frequency. We give explicit expressions for these coefficients in Appendix A.

Combining Eqs. (13) and (14) we can rewrite the data model by grouping together terms with the same time dependence:

In our case we will however assume a perfectly uniform sky coverage, both in terms of the number of observations of each pixel, as well as distributions of the HWP and sky angles. With these assumptions the map-making problem for the data model in Eq. (15) is not only solvable, but results in estimates of the five sky signals which have mutually uncorrelated noise. Assuming furthermore that the instrumental noise is white, the RMS of the noise in the four mixed-Stokes maps and one total intensity is the same as in the standard case, i.e.,

$$\sigma_{\text{RMS}}(\text{POL}) = \sigma_{\text{RMS}}(\text{INT})\sqrt{2} \propto \sqrt{\frac{2}{n_{\text{hits}}}} \quad (16)$$

In the following, we use these estimates as reflecting the noise levels in our mixed-Stokes maps.

For shortness, hereafter we denote the four mixed-Stokes signals as  $\mathcal{C}_0$ ,  $\mathcal{S}_0$ ,  $\mathcal{C}_4$ ,  $\mathcal{S}_4$ , and introduce an effective intensity signal,  $\mathcal{I}(\gamma_t, \nu) \equiv \mathbf{M}_{00}\mathbf{I}(\gamma_t, \nu)$ , rewriting our data model as:

$$\begin{aligned} \mathbf{m}_t(\nu) \equiv & \mathcal{I}(\gamma_t, \nu) \\ & + \mathcal{C}_0(\gamma_t, \nu) \cos(2\psi_t) \\ & + \mathcal{S}_0(\gamma_t, \nu) \sin(2\psi_t) \\ & + \mathcal{C}_4(\gamma_t, \nu) \cos(4\phi_t + 2\psi_t) \\ & + \mathcal{S}_4(\gamma_t, \nu) \sin(4\phi_t + 2\psi_t) \\ & + \mathbf{n}_t. \end{aligned} \quad (17)$$

These mixed Stokes components are related to standard Stokes components, see Eq. (15), via a linear transformation given by:

$$\begin{pmatrix} \mathcal{I}(\gamma_t, \nu) \\ \mathcal{C}_0(\gamma_t, \nu) \\ \mathcal{S}_0(\gamma_t, \nu) \\ \mathcal{C}_4(\gamma_t, \nu) \\ \mathcal{S}_4(\gamma_t, \nu) \end{pmatrix} = \underbrace{\begin{bmatrix} \mathbf{M}_{00}(\nu) & 0 & 0 \\ 0 & \mathbf{C}_{01;0}(\nu) & \mathbf{C}_{02;0}(\nu) \\ 0 & \mathbf{S}_{01;0}(\nu) & \mathbf{S}_{02;0}(\nu) \\ 0 & \mathbf{C}_{01;4}(\nu) & \mathbf{C}_{02;4}(\nu) \\ 0 & \mathbf{S}_{01;4}(\nu) & \mathbf{S}_{02;4}(\nu) \end{bmatrix}}_{\equiv \mathbf{M}(\nu)} \times \begin{pmatrix} \mathbf{I}(\gamma_t, \nu) \\ \mathbf{Q}(\gamma_t, \nu) \\ \mathbf{U}(\gamma_t, \nu) \end{pmatrix}, \quad (18)$$

where the transformation matrix,  $\mathbf{M}(\nu)$ , depends on instrumental parameters. We conclude therefore that the map-making procedure could render estimates of pure  $Q$  and  $U$  Stokes parameters only if the true values of the relevant instrumental parameters are known.

Lastly, we note that the fact that the mixed Stokes signals are combinations of only Stokes  $Q$  and  $U$  parameters is not universal. Indeed, this is, for instance, not the case whenever I-to-P leakage is present. The formalism presented hereafter can be straightforwardly extended to include such cases.

### 3. Bandpasses

The actual CMB experiments detect the incoming signal integrated over some frequency bands rather than the monochromatic one. The frequency bands are centered at desired frequencies,  $\nu_c$ , and are defined by the instrumental response functions referred to hereafter as bandpasses. These are denoted as  $\mathcal{B}(\nu, \nu_c)$ . To account for their effects we need to introduce bandpass-averaged quantities, which we distinguish with a bar over a symbol and defined as:

$$\bar{\mathbf{X}}(\nu_c) \equiv \frac{\int d\nu \mathcal{B}(\nu, \nu_c) \mathbf{X}(\nu) k(\nu)}{\int d\nu \mathcal{B}(\nu) k(\nu)}, \quad (19)$$

where  $k$  is a conversion factor which reconciles the units used for the data,  $\mathbf{X}$ , and the bandpasses,  $\mathcal{B}$ .

We can then rewrite the data model in Eq. (17) using the bandpass averaged objects as:

$$\begin{aligned} \bar{\mathbf{m}}_t(\nu_c) &\equiv \bar{\mathcal{I}}(\gamma_t, \nu_c) \\ &+ \bar{\mathcal{C}}_0(\gamma_t, \nu_c) \cos(2\psi_t) \\ &+ \bar{\mathcal{S}}_0(\gamma_t, \nu_c) \sin(2\psi_t) \\ &+ \bar{\mathcal{C}}_4(\gamma_t, \nu_c) \cos(4\phi_t + 2\psi_t) \\ &+ \bar{\mathcal{S}}_4(\gamma_t, \nu_c) \sin(4\phi_t + 2\psi_t) \\ &+ \bar{\mathbf{n}}_t. \end{aligned} \quad (20)$$

The map-making codes can be used assuming this data model in the same way as in the monochromatic case

however they will now produce bandpass-averaged maps of mixed-Stokes and effective total intensity.

We further note that going from these maps to the maps of pure Stokes parameters for each frequency band is not possible once the bandpasses are explicitly included. This is because the integration over the bandpasses does not preserve the matrix form of Eq. (18) as the pure Stokes signals,  $I(\gamma_t, \nu)$ ,  $Q(\gamma_t, \nu)$ ,  $U(\gamma_t, \nu)$ , are integrated over the frequency together with the corresponding elements of the matrix  $\mathbf{M}(\nu)$ . Consequently, the mixed-Stokes maps are the only available objects at the end of the map-making even if the instrumental parameters were perfectly known. This does not however prevent us from recovering pure Stokes parameters of maps of components of a different physical origin. We discuss this in the following.

## C. Multi-component data model

The polarized sky signal as measured at different frequencies and characterized by three Stokes parameters,  $I$ ,  $Q$ ,  $U$ , is really a mixture of the genuine CMB signal and polarized galactic foregrounds. The two most important polarized foregrounds are polarized thermal dust emission and synchrotron radiation. Separating the measured signals into signals of the sky components is the goal of component separation procedures. Those typically are performed in the pixel domain and use maps produced on the map-making step as inputs. In this section we derive a data model for the mixed-Stokes maps, Eq. (20), which in the discussed case constitute the output of the map-making procedure and therefore the input of the component separation.

### 1. Sky components

We model foregrounds by a sky template at a reference frequency  $\nu_0$ , scaled to the observation frequency  $\nu$  using foregrounds emission models. The effective signal as measured at each frequency is a mixture of these three major sky components (CMB, dust, and synchrotron), which is described by a mixing matrix  $\mathbf{A}$  containing the scaling laws for each sky component at each frequency. The scaling laws we assume in this work are fairly standard, e.g., [9], and in the  $\mu\text{K}_{\text{RJ}}$  units they are given by:

$$\mathbf{A}^{\text{sync}}(\nu, \nu_0) = \left( \frac{\nu}{\nu_0} \right)^{\beta_s} \quad (21)$$

where  $\beta_s$  is the synchrotron spectral index, and

$$\mathbf{A}^{\text{dust}}(\nu, \nu_0) = \left( \frac{\nu}{\nu_0} \right)^{\beta_d+1} \frac{\exp \frac{h\nu_0}{kT_d} - 1}{\exp \frac{h\nu}{kT_d} - 1} \quad (22)$$

where  $\beta_d$  is the dust spectral index and  $T_d$  is the dust temperature.

The CMB follows a known black body emission spectrum, with  $T_{\text{CMB}} = 2.7255\text{K}$ , leading to the assumption that  $\mathbf{A}^{\text{CMB}}(\nu, \nu_0) = 1$  in  $\mu\text{K}_{\text{CMB}}$  units.

For a given observed frequency  $\nu$ , the sky signal can be modeled as:

$$\begin{pmatrix} \mathcal{I}(\nu) \\ \mathcal{Q}(\nu) \\ \mathcal{U}(\nu) \end{pmatrix} = \sum_{\substack{\text{comp}=\text{cmb} \\ \text{dust, sync}}} \mathbf{A}^{\text{comp}}(\nu, \nu_0) \begin{pmatrix} \mathcal{I}_{\text{comp}}(\nu_0) \\ \mathcal{Q}_{\text{comp}}(\nu_0) \\ \mathcal{U}_{\text{comp}}(\nu_0) \end{pmatrix}, \quad (23)$$

with  $\mathbf{A}^{\text{comp}}$  a diagonal matrix scaling each component template to the desired frequency,  $\nu$ . The latter, as presented

here, are in  $\mu\text{K}_{\text{RJ}}$ , but for simplicity, we generate maps at all frequencies and for all components in  $\mu\text{K}_{\text{CMB}}$  units, as are the recovered maps of the CMB and foreground signals. A conversion factor is then taken into account so that units are handled consistently throughout the process.

## 2. Multicomponent sky model

To derive a complete multicomponent model of the mixed-Stokes maps, we first introduce explicitly the different sky components in the model in Eq. (18) using Eq. (23):

$$\underbrace{\begin{pmatrix} \mathcal{I}(\gamma_t, \nu) \\ \mathcal{C}_0(\gamma_t, \nu) \\ \mathcal{S}_0(\gamma_t, \nu) \\ \mathcal{C}_4(\gamma_t, \nu) \\ \mathcal{S}_4(\gamma_t, \nu) \end{pmatrix}}_{\equiv \mathbf{s}(\gamma_t, \nu)} \equiv \sum_{\substack{\text{comp}=\text{cmb} \\ \text{dust, sync}}} \underbrace{\begin{bmatrix} \mathbf{M}_{00; \cos\theta_0}(\nu) & 0 & 0 \\ 0 & \mathbf{C}_{01;0}(\nu) & \mathbf{C}_{02;0}(\nu) \\ 0 & \mathbf{S}_{01;0}(\nu) & \mathbf{S}_{02;0}(\nu) \\ 0 & \mathbf{C}_{01;4}(\nu) & \mathbf{C}_{02;4}(\nu) \\ 0 & \mathbf{S}_{01;4}(\nu) & \mathbf{S}_{02;4}(\nu) \end{bmatrix}}_{\equiv \mathbf{A}^{\text{comp}}(\nu, \nu_0)} \mathbf{A}^{\text{comp}}(\nu, \nu_0) \underbrace{\begin{pmatrix} \mathcal{I}_{\text{comp}}(\gamma_t, \nu_0) \\ \mathcal{Q}_{\text{comp}}(\gamma_t, \nu_0) \\ \mathcal{U}_{\text{comp}}(\gamma_t, \nu_0) \end{pmatrix}}_{\equiv \mathbf{c}_{\text{comp}}(\gamma_t, \nu_0)}. \quad (24)$$

We then average both sides of this equation over the bandpasses using the prescription in Eq. (19), and define component-specific, bandpass-integrated matrices,  $\bar{\mathbf{A}}$  as:

$$\bar{\mathbf{A}}^{\text{comp}}(\nu_c, \nu_0) \equiv \frac{\int d\nu \mathbf{M}(\nu) \mathcal{B}(\nu, \nu_c) \mathbf{A}^{\text{comp}}(\nu, \nu_0) k(\nu)}{\int d\nu \mathcal{B}(\nu) k(\nu)}, \quad (25)$$

where we assumed that the bandpass center frequency is  $\nu_c$ .

We can now finally write our multicomponent, bandpass-integrated data model for the mixed-Stokes maps as:

$$\begin{aligned} \bar{\mathbf{s}}(\gamma_t, \nu_c, \nu_0) &= \underbrace{[\bar{\mathbf{A}}^{\text{cmb}}(\nu_c, \nu_0), \bar{\mathbf{A}}^{\text{dust}}(\nu_c, \nu_0), \bar{\mathbf{A}}^{\text{sync}}(\nu_c, \nu_0)]}_{\equiv \bar{\mathbf{A}}(\nu_c, \nu_0)} \underbrace{\begin{pmatrix} \mathbf{c}_{\text{cmb}}(\gamma_t, \nu_0) \\ \mathbf{c}_{\text{dust}}(\gamma_t, \nu_0) \\ \mathbf{c}_{\text{sync}}(\gamma_t, \nu_0) \end{pmatrix}}_{\equiv \mathbf{c}(\gamma_t, \nu_0)} \\ &= \bar{\mathbf{A}}(\nu_c, \nu_0) \mathbf{c}(\gamma_t, \nu_0). \end{aligned} \quad (26)$$

We point out that while this equation is more involved than the usual relation between the standard observables (i.e., single frequency maps of each Stokes parameter) and the underlying components, Eq. (23), the relation in both these cases is linear as far as sky component amplitudes are concerned. However the system (mixing) matrix in the mixed-Stokes case depends on both foreground spectral parameters as well as instrumental parameters.

We explore consequences of these two observations in the follow-up sections.

## III. COMPONENT SEPARATION

In this section, we follow the formalism of [8], adapting their procedure to account for mixed Stokes components to

include instrumental parameters (HWP, sinuous antenna, bandpasses). We first elaborate the mixed component maps data model (III A), and subsequently discuss component separation and estimation of foreground and hardware parameters from these mixed-Stokes maps (III B). With the estimated parameters, we reconstruct the sky signal and estimate residuals (III C), as well as noise after component separation (III D). Finally, we evaluate the impact of the studied effects on cosmological parameters (III E).

### A. Data model

Following Eq. (26) and as shown in Fig. 2, we collect together all single-frequency mixed-Stokes maps, and we can write our data model for the entire multifrequency dataset:

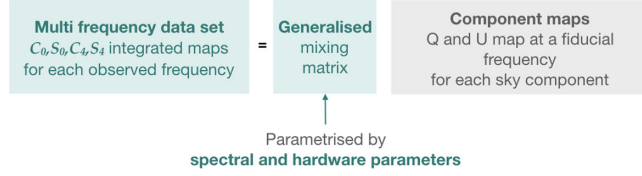


FIG. 2. Data model with generalized mixing matrix, Eq. (26).

$$\bar{\mathbf{m}} = \bar{\mathbf{s}} + \bar{\mathbf{n}} \equiv \bar{\mathcal{A}}\mathbf{c} + \bar{\mathbf{n}}, \quad (27)$$

Here the combined data vector  $\bar{\mathbf{m}}$  includes all the bandpass-integrated mixed-Stokes maps,  $\bar{\mathcal{C}}_0, \bar{\mathcal{C}}_4, \bar{\mathcal{S}}_0, \bar{\mathcal{S}}_4$ , as measured for all frequency bands which are all concatenated together in a single data vector. From now on, we focus only on the polarization and exclude total intensity from our consideration. While this is not the most general case, the generalization to all three Stokes parameters is straightforward.  $\bar{\mathbf{s}}$  is the noiseless signal in the all mixed-Stokes maps recovered at all frequencies, and  $\bar{\mathcal{A}}$  is a generalized, multifrequency mixing matrix. It is composed of single channel matrices,  $\bar{\mathcal{A}}(\nu_c, \nu_0)$  defined in Eq. (26), put on top of each other.  $\bar{\mathbf{n}}$  denotes the actual noise present in all the mixed-Stokes maps as derived by the map-making procedure.

To improve readability, as we always integrate over bandpass hereafter, we drop the bar that indicates bandpass integration in the follow-up.

For a given pixel  $p$ , we have:

$$\mathbf{m}_p = \mathbf{s}_p + \mathbf{n}_p \equiv \mathcal{A}_p(\beta_f, \beta_h)\mathbf{c}_p + \mathbf{n}_p \quad (28)$$

where  $\beta_f$  refers to foregrounds spectral parameters and  $\beta_h$  to hardware parameters. Whereas  $\beta_f$  parameters apply to a specific component (dust or synchrotron) and can vary between sky pixels,  $\beta_h$  are typically global parameters, applied to all pixels and all sky components—including CMB. For simplicity, we hereafter assume that the foreground parameters  $\beta_f$  that we fit for are not pixel-dependent (we consider a single set of  $\beta_f$  for the entire observed sky). The generalization of the proposed formalism to pixel-dependent foreground parameters is possible following the steps already outlined in [9].

We also note that the proposed formalism can be extended to allow for the variability of some of the detector properties across the detector arrays thus permitting to study their impact on the experiment performance. Such arrays are envisaged for the future generation of the CMB experiments, making this feature of the framework very timely. We leave detailed investigation of such issues to future work and in the following we give an example of the impact of a mismatch between the parameters used for generating the data and those assumed in their modeling in the case of bandpasses. The mismatch could arise in practice as a difference between effective bandpass

parameters resulting from averaging data of many detectors and a single set of idealized parameters assumed in the reconstruction.

## B. Parameter estimation

The generalized data model, although more involved than the standard model assumed by component separation methods, retains all the essential features of the latter and most of the existing component separation procedures can be adapted to account for the extensions. In this work, we focus on the parametric component separation technique as a method of choice for the component separation and in particular its two step implementation as proposed in [7] and elaborated on in [8,9,17]. We discuss below essential ingredients of the method, emphasising the new features due to the specificity of the new data model.

### 1. Ensemble average likelihood

For the data model in Eq. (28), we write an effective spectral likelihood [7]. This is obtained by replacing the standard mixing matrix  $\mathbf{A}$  (see, e.g., Eq. (6) of [8]) by the generalized mixing matrix,  $\mathcal{A}$ . This generalized log-likelihood reads therefore as:

$$\mathcal{S} = -\sum_p (\mathcal{A}_p^t \mathbf{N}_p^{-1} \mathbf{m}_p)^t (\mathcal{A}_p^t \mathbf{N}_p^{-1} \mathcal{A}_p)^{-1} \mathcal{A}_p^t \mathbf{N}_p^{-1} \mathbf{m}_p \quad (29)$$

where  $\mathbf{N}_p$  is the noise covariance matrix. A key difference with the standard case is that now the mixing matrix,  $\mathcal{A}$ , depends on both foreground and instrumental parameters. While ideally the latter are known with sufficient precision from an instrument calibration campaign, this is hardly the case for most of the parameters of interest. Here, we therefore aim at determining both these sets of parameters, i.e.,  $\beta_f$  and  $\beta_h$ , from the available dataset via the maximization of the spectral likelihood. Clearly, this may not be always possible and some external information may be required or beneficial, as detailed in Secs. III B 2 and IV B 3 b. This procedure can be applied to any specific input dataset, thus providing a basis for actual data processing framework. In the context of performance forecasting we are interested in quantities averaged over statistical ensembles of the possible input datasets. In view of this, we derive the ensemble-averaged version of the likelihood in Eq. (29) [8]:

$$\langle \mathcal{S} \rangle = -\text{tr} \sum_p \{ (\mathbf{N}_p^{-1} - \mathbf{P}_p) (\langle \mathbf{s}_p \mathbf{s}_p^t \rangle + \mathbf{N}_p) \}, \quad (30)$$

where  $\mathbf{P}_p$  is the projection operator defined as:

$$\mathbf{P}_p \equiv \mathbf{N}_p^{-1} - \mathbf{N}_p^{-1} \mathcal{A}_p (\mathcal{A}_p^t \mathbf{N}_p^{-1} \mathcal{A}_p)^{-1} \mathcal{A}_p^t \mathbf{N}_p^{-1}. \quad (31)$$

The single-pixel data covariance matrix in Eq. (30),  $\langle \mathbf{s}_p \mathbf{s}_p^t \rangle$ , reflects the properties of what we assume to be the true sky.

Unlike in the case studied in [8] it has to now explicitly account for the presence of the CMB contribution. This is due to the fact that instrumental parameters can, and do, affect the frequency scaling of the CMB signal. In order to calculate the covariance we split  $\mathcal{A}_p$  and  $\mathbf{c}_p$  into their CMB and foregrounds parts:

$$\mathbf{c}_p \equiv \begin{bmatrix} \mathbf{c}_p^{\text{cmb}} \\ \mathbf{c}_p^{\text{fg}} \end{bmatrix} \quad (32)$$

$$\mathcal{A} = [\mathcal{A}_p^{\text{cmb}}, \mathcal{A}_p^{\text{fg}}], \quad (33)$$

and we split the single-pixel data covariance matrix into its CMB and foregrounds contributions accordingly:

$$\langle \mathbf{s}_p \mathbf{s}_p^t \rangle = \mathcal{A}_p^{\text{cmb}} \langle \mathbf{c}_p^{\text{cmb}} \mathbf{c}_p^{\text{cmb},t} \rangle \mathcal{A}_p^{\text{cmb},t} + \mathcal{A}_p^{\text{fg}} \langle \mathbf{c}_p^{\text{fg}} \mathbf{c}_p^{\text{fg},t} \rangle \mathcal{A}_p^{\text{fg},t}. \quad (34)$$

The foregrounds contribution (the second term) is easy to compute as we consider foregrounds as a fixed template independent on the realization. This is not the case for the CMB contribution though, and in principle the average in the first term of the above equation should include the pixel-pixel correlations. However, if we assume that neither the pointing matrices,  $\mathcal{A}_p$ , nor the noise matrices,  $\mathbf{N}_p$ , are pixel-dependent and that the noise is white, we do not need the full multipixel covariance matrix for the CMB but only its single pixel version. Considering the two polarization states  $Q$  and  $U$ ,

$$\mathbf{c}_p^{\text{cmb}} \equiv [\mathbf{c}_{p,Q}^{\text{cmb}}, \mathbf{c}_{p,U}^{\text{cmb}}], \quad (35)$$

the single-pixel covariance of the CMB signal is a 2-by-2 matrix,  $\hat{\mathbf{S}}^{\text{cmb}} \equiv \langle \mathbf{c}_p^{\text{cmb}} \mathbf{c}_p^{\text{cmb},t} \rangle$ , given by:

$$\hat{\mathbf{S}}^{\text{cmb}} \equiv \begin{bmatrix} \sigma_{QQ}^2 & \sigma_{QU}^2 \\ \sigma_{UQ}^2 & \sigma_{UU}^2 \end{bmatrix}, \quad (36)$$

and can be straightforwardly calculated for any standard theoretical model.

We can write the total ensemble average likelihood explicitly as:

$$\mathcal{S} = -\text{tr}\{n_{\text{pix}}(\mathbf{N}^{-1} - \mathbf{P})(\mathbf{F} + \mathbf{N} + \mathcal{A}^{\text{cmb}} \hat{\mathbf{S}}^{\text{cmb}} \mathcal{A}^{\text{cmb},t})\}, \quad (37)$$

where:

$$\mathbf{F} \equiv \frac{1}{n_{\text{pix}}} \mathcal{A}^{\text{fg}} \sum_p \mathbf{c}_p^{\text{fg}} \mathbf{c}_p^{\text{fg},t} \mathcal{A}^{\text{fg},t}. \quad (38)$$

The total ensemble average likelihood therefore takes into account average over noise realizations through the noise

covariance matrix  $\mathbf{N}$ , and over CMB realizations through the single-pixel, CMB covariance matrix  $\hat{\mathbf{S}}^{\text{cmb}}$ . In what follows, we use the expression Eq. (37) for the spectral likelihood, and therefore keep the assumption that  $\mathcal{A}$  and  $\mathbf{N}$  are sky pixel independent.

## 2. Priors on instrumental parameters

As mentioned earlier we expect that not all instrumental parameters can be constrained using the observed data with sufficient precision. Whenever this is the case, we may need to introduce calibration priors on some instrumental parameters either to break degeneracies in the system and/or to ensure a tolerable level of residuals. Where appropriate, we add Gaussian priors to the ensemble average spectral likelihood as:

$$\mathcal{S}' = \mathcal{S} + \sum_{\beta_h} \frac{1}{2\sigma_{\beta_h}^2} (\beta_h - \tilde{\beta}_h)^2, \quad (39)$$

where  $\mathcal{S}$  is the likelihood of Eq. (37) and  $\sigma_{\beta_h}$  is the calibration error on parameter  $\beta_h$ . We assume the calibration measurement to be unbiased with a  $1\sigma$  error given by  $\sigma_{\beta_h}$ . Its best fit value will be typically different than the true value of the parameter, denoted with  $\tilde{\beta}_h$ , but should be within the estimated uncertainty from it,  $\sigma_{\beta_h}^2$ . Our prior in Eq. (39) is assumed to be averaged over an ensemble of calibration procedures. Consequently, it is centered at the true value of the parameter and its uncertainty is larger by a factor of 2 (in quadrature) than that of a single calibration result.

## 3. Statistical error

In this approach, the maximum-likelihood values obtained by minimizing Eq. (39) are average values, and we can compute the matrix of second derivatives at the peak of the likelihood—the Hessian matrix  $\mathcal{H}$ :

$$\mathcal{H}_{\beta\beta'} \equiv \left\langle \frac{\partial^2 \mathcal{S}'}{\partial \beta \partial \beta'} \bigg|_{\text{peak}} \right\rangle_{\text{cmb+noise}}, \quad (40)$$

where we do not assume that the peak of the likelihood corresponds to true values of parameters  $(\tilde{\beta}_f, \tilde{\beta}_h)$ . The Hessian measures the curvature of the likelihood at its peak, and is directly related to the uncertainty due to instrumental noise and the CMB:

$$\boldsymbol{\Sigma} \simeq \mathcal{H}^{-1}. \quad (41)$$

This statistical error matrix  $\boldsymbol{\Sigma}$  can be computed analytically following Eq. (A5) in [8].

If the assumed mixing matrix  $\mathcal{A}(\beta_f, \beta_h)$  corresponds to the true  $\mathcal{A}$  for a given set of parameters  $(\beta_f, \beta_h)$  derived from the minimization of the spectral likelihood, the



statistical uncertainty on the estimated parameters expressed by  $\Sigma$  will be the only source of residuals in the cleaned CMB map. Following [8] we refer to those as statistical residuals. These increase the error on the estimated cosmological parameters but do not bias their estimated values [17]. However, if the assumed and true mixing matrices do not match in average, there will be systematic differences between the estimated and true sky components leading to systematic residuals in the cleaned CMB map and therefore potentially to biases in the estimated values of cosmological parameters [8].

In the next section, we derive the expression of both systematic and statistical residuals for the case of the generalized data model.

### C. Residuals

Given a set of  $(\beta_f, \beta_h)$  parameters, we can compute the corresponding mixing matrix  $\mathcal{A}$ , and reconstruct the estimation of the noiseless component maps  $\hat{\mathbf{c}}$  [6]:

$$\hat{\mathbf{c}}_p = (\mathcal{A}'\mathbf{N}^{-1}\mathcal{A})^{-1}\mathcal{A}'\mathbf{N}^{-1}\mathbf{s}_p \equiv \mathbf{W}(\beta_f, \beta_h)\mathbf{s}_p, \quad (42)$$

where we recall that  $\mathbf{s}$  is the noiseless sky signal. The noiseless residuals are defined as the difference between the reconstructed map and the true signal:

$$\mathbf{r}_p \equiv \hat{\mathbf{c}}_p - \mathbf{c}_p = \mathbf{W}(\beta_f, \beta_h)\mathbf{s}_p - \mathbf{c}_p. \quad (43)$$

We consider two main contributions to the residuals: (1) statistical residuals, reflecting the statistical scatter due to the instrumental noise and CMB realizations; and (2) systematic residuals, reflecting the systematic (averaged) differences between the true sky and the best-fit of the assumed model. Rewriting Eq. (43) for the CMB channel only, we have:

$$\begin{aligned} \mathbf{r}_p^{\text{CMB}}(\beta_f, \beta_h) &= \underbrace{\mathbf{W}^0(\beta_f, \beta_h)\mathbf{F}_p}_{\text{foregrounds residuals}} \\ &+ \underbrace{\mathbf{W}^0(\beta_f, \beta_h)\mathbf{C}_p - \mathbf{c}_p^{\text{cmb}}}_{\text{CMB residuals}}, \end{aligned} \quad (44)$$

where the index 0 denotes the CMB part of the  $\mathbf{W}$  operator, and  $\mathbf{F}_p$  (resp.  $\mathbf{C}_p$ ) is the total foreground (resp. CMB) contribution at each frequency, gathering contribution from all mixed Stokes parameters  $(\mathcal{C}_0, \mathcal{S}_0, \mathcal{C}_4, \mathcal{S}_4)$  defined in Eq. (20). Whenever elements of the mixing matrix corresponding to the CMB component are assumed to be known, as it was the case in the original application of this framework [8,9], the second term in Eq. (44) vanishes and the recovered CMB map contains always all the CMB signal and is merely contaminated by the foreground residuals. However, this term has to be taken into account

in the present case as the instrumental effects may change the way the CMB amplitude varies between frequency bands. Here we therefore generalize the derivation of [8] explicitly accounting for this term.

We can rewrite the CMB contribution of Eq. (44) as:

$$\begin{aligned} \mathbf{W}^0\mathbf{C} - \mathbf{c}_{\text{cmb}} &= \mathbf{W}^0\mathcal{A}^{\text{cmb}}\mathbf{c}_{\text{cmb}} - \mathbf{c}_{\text{cmb}} \\ &= (\mathbf{W}^0\mathcal{A}^{\text{cmb}} - 1)\mathbf{c}_{\text{cmb}}. \end{aligned} \quad (45)$$

This simplifies the estimation of the CMB contribution in the residuals, that can be computed using the simplified procedure described in Appendix D of [8].

We perform a Taylor expansion of Eq. (44) around the estimated best fit values for  $\beta = \{\beta_h, \beta_f\}$ , denoted  $\hat{\beta}$ :

$$\begin{aligned} \mathbf{r}_p^{\text{CMB}}(\beta) &\simeq \mathbf{W}_p^0(\hat{\beta})(\mathbf{F}_p + \mathbf{C}_p) \\ &+ \sum_{\beta} \delta\beta \left. \frac{\partial \mathbf{W}_p^0}{\partial \beta} \right|_{\hat{\beta}} (\mathbf{F}_p + \mathbf{C}_p) \\ &+ \sum_{\beta, \beta'} \delta\beta \delta\beta' \left. \frac{\partial^2 \mathbf{W}_p^0}{\partial \beta \partial \beta'} \right|_{\hat{\beta}} (\mathbf{F}_p + \mathbf{C}_p) \\ &- \mathbf{c}_p^{\text{cmb}}, \end{aligned} \quad (46)$$

with  $\delta\beta \equiv \beta - \hat{\beta}$ .

We introduce new quantities in pixel-domain as in [8]:

$$\begin{aligned} \mathbf{y}_p &\equiv \mathbf{W}_p^0(\hat{\beta})(\mathbf{F}_p + \mathbf{C}_p) - \mathbf{c}_p^{\text{cmb}} \\ \mathbf{Y}_{p,\beta}^{(1)} &\equiv \sum_{\beta} \left. \frac{\partial \mathbf{W}_p^0}{\partial \beta} \right|_{\hat{\beta}} (\mathbf{F}_p + \mathbf{C}_p) \\ \mathbf{Y}_{p,\beta\beta'}^{(2)} &\equiv \sum_{\beta, \beta'} \left. \frac{\partial^2 \mathbf{W}_p^0}{\partial \beta \partial \beta'} \right|_{\hat{\beta}} (\mathbf{F}_p + \mathbf{C}_p), \end{aligned} \quad (47)$$

so that we can rewrite Eq. (46):

$$\mathbf{r}_p^{\text{CMB}}(\beta) \simeq \mathbf{y}_p + \sum_{\beta} \delta\beta \mathbf{Y}_{p,\beta}^{(1)} + \sum_{\beta, \beta'} \delta\beta \delta\beta' \mathbf{Y}_{p,\beta\beta'}^{(2)}. \quad (48)$$

We fit for a single value of every  $\beta$  for all considered pixels and can rewrite the same equation in the harmonic domain and express the total level of foreground residuals as:

$$\mathcal{C}_{\ell}^{\text{res}} \equiv \otimes_{\ell} (\mathbf{r}^{\text{CMB}}, \mathbf{r}^{\text{CMB}}) \quad (49)$$

$$\begin{aligned} &\simeq \otimes_{\ell} (\mathbf{y}, \mathbf{y}) + \otimes_{\ell} (\mathbf{y}, \mathbf{z}) + \otimes_{\ell} (\mathbf{z}, \mathbf{y}) \\ &+ \text{tr}[\Sigma \otimes_{\ell} (\mathbf{Y}^{(1)}, \mathbf{Y}^{(1)})] \end{aligned} \quad (50)$$

where the symbol  $\otimes$  denotes the cross-spectrum of two quantities *averaged over statistical ensemble of the CMB realizations*. We note that in the case studied in [8] this average was trivial as there was no CMB

contribution to the residuals and therefore usually omitted.  $\mathbf{z}$  is defined as:

$$\mathbf{z}_p \equiv \text{tr}[\mathbf{Y}_p^{(2)} \boldsymbol{\Sigma}]. \quad (51)$$

As outlined previously, residuals are composed of two main contributions, statistical and systematic residuals, as well as a cross-term:

$$C_\ell^{\text{sys}} \equiv \otimes_\ell (\mathbf{y}, \mathbf{y}) \quad (52)$$

$$C_\ell^{\text{stat}} \equiv \text{tr}[\boldsymbol{\Sigma} \otimes_\ell (\mathbf{Y}^{(1)}, \mathbf{Y}^{(1)})]. \quad (53)$$

$$C_\ell^{\text{cross}} \equiv \otimes_\ell (\mathbf{y}, \mathbf{z}) + \otimes_\ell (\mathbf{z}, \mathbf{y}). \quad (54)$$

We can also write the full covariance matrix of the CMB map,  $\tilde{\mathbf{c}}^{\text{cmb}} (\equiv \mathbf{c} + \mathbf{r}^{\text{CMB}} + \mathbf{n})$ , as recovered via the component separation procedure described here. It reads:

$$\mathbf{E} \equiv \langle \tilde{\mathbf{c}}^{\text{cmb}} \tilde{\mathbf{c}}^{\text{cmb}} \rangle_{\text{cmb+noise}} = \mathbf{C}^{\text{cmb}} + \mathbf{N} + \langle \mathbf{y} \mathbf{y}^\dagger \rangle_{\text{cmb}} + \langle \mathbf{z} \mathbf{y}^\dagger \rangle_{\text{cmb}} + \langle \mathbf{y} \mathbf{z}^\dagger \rangle_{\text{cmb}} + \langle \mathbf{Y}^{(1)} \boldsymbol{\Sigma} \mathbf{Y}^{(1)\dagger} \rangle_{\text{cmb}}, \quad (55)$$

where  $\mathbf{C}^{\text{cmb}}$  denotes the covariance of the CMB signal,  $\mathbf{N}$  that of the noise in the recovered map, and the angle brackets denote averaging over statistical ensemble as defined by the subscript.

#### D. Noise

We assume that the noise is homogeneous and uncorrelated for all frequency channels, and given that we have one value of all parameters  $\beta$  for the entire sky patch we can express the noise power spectrum in the cleaned CMB map as in Eq. (32) of [8]:

$$C_\ell^{\text{noise}} = [(\mathcal{A}^t \mathbf{N}_\ell^{-1} \mathcal{A})^{-1}]_{\text{CMB} \times \text{CMB}} \quad (56)$$

with  $\mathbf{N}_\ell$  describing the noise spectra of each frequency map, taking its resolution into account:

$$\mathbf{N}_\ell^{ij} \equiv (w_i)^{-1} \exp\left(\ell(\ell+1) \frac{\text{FWHM}_i^2}{8 \log 2}\right) \delta_i^j \quad (57)$$

with  $(w_i)^{-1/2}$  the sensitivity, and  $\text{FWHM}_i$  the full-width half maximum of the corresponding frequency band  $i$ .

#### E. Cosmological likelihood

Following Appendix C2 in [8], we write the cosmological parameter likelihood, averaged over instrumental noise and CMB signal realizations:

$$\langle S^{\text{cos}} \rangle \equiv \text{tr} \mathbf{C}^{-1} \mathbf{E} + \ln \det \mathbf{C} \quad (58)$$

where  $\mathbf{C}$  is the assumed multipixel covariance of the cleaned CMB signal, and  $\mathbf{E}$  is the multipixel correlation

matrix of the CMB map retrieved with the component separation procedure as defined in Eq. (55). Consequently,  $\mathbf{E}$  takes into account the presence of the residuals, Eqs. (52), (53), and (54), and all which can be computed semianalytically given the actual model of the data.

The assumed covariance  $\mathbf{C}$  expresses the state of our knowledge about the data. Ideally, we would wish that  $\mathbf{C} = \mathbf{E}$ , however in practice this is rarely the case. In the following we will consider two cases. In the first case, the assumed covariance is that of the CMB signal only, thus ignoring entirely the effects of the component separation. We have therefore:

$$\mathbf{C} = \mathbf{C}^{\text{cmb}} + \mathbf{N}. \quad (59)$$

In the second case we assume that the statistical errors can be modelled on some level. Specifically, we assume that:

$$\mathbf{C} = \mathbf{C}^{\text{cmb}} + \mathbf{C}^{\text{stat}} + \mathbf{N}, \quad (60)$$

where  $\mathbf{C}^{\text{stat}}$  is the covariance matrix of the statistical residuals  $C_\ell^{\text{stat}}$  defined in Eq. (53). We note that sufficient information allowing for effective modeling of the statistical residuals may be indeed available either internally, in some self-consistent statistical approaches, e.g., [5,7], or using some external data, e.g., [18]. Hereafter, we refer to this second case as the deprojection case. Detailed expressions are given in Appendix C.

## IV. APPLICATION

In this section, we first introduce the instrumental configuration and parametrization that we adopt in this work (IV A), then we present various sets of parameters that we consider and discuss uncertainties and degeneracies on parameters in these various cases (IV B). From these, we can finally estimate foreground residuals and forecast their impact on tensor-to-scalar ratio determination (IV C).

### A. Configuration

The framework presented here is very flexible and could be adapted to a broad range of experiment designs with different instrument models. For definiteness, we demonstrate it in the case of a typical, ground-based, Stage 3 CMB polarization experiment. The configuration and parameters we use are motivated by the publicly available Simons Observatory Small Aperture Telescopes design [19].

#### 1. Telescopes and frequency bands

As outlined in the Introduction, to enable for component separation, modern CMB polarization experiments typically deploy several instruments to trace foregrounds at different frequencies, so that we can disentangle them

from the black body CMB signal. In several ground-based experiments, these frequency bands are grouped by two on the same focal plane, hence the use of an achromatic HWP that has to accommodate two different observing frequencies in the same optics tube. Following the Simons Observatory configuration [19], we consider three telescopes, whose frequency coverage is as follows:

- (i) Low frequency (LF) telescope: frequency bands centered at 27 GHz and 39 GHz;
- (ii) Mid frequency (MF) telescope: frequency bands centered at 93 GHz and 145 GHz;
- (iii) Ultra-high frequency (UHF) telescope: frequency bands centered at 225 GHz and 280 GHz.

We model input noise covariance matrices in the pixel domain as

$$\mathbf{N}_p^{ii} = w_i^{-1}, \quad (61)$$

with  $w_i^{-1}$  the sensitivity of the corresponding frequency band  $i$ . We use the publicly available sensitivity calculator [20] to compute realistic instrumental sensitivities corresponding to the frequency coverage.

## 2. HWP

For each telescope (LF, MF, UHF), we consider a typical 3-layer achromatic HWP, for which only the central layer is rotated with respect to the reference frame of the instrument, i.e.,  $\alpha_1 = \alpha_3 = 0$ , as for example in [21]. This typical choice of angle is driven by the maximization of transmitted power and is known as Pancharatnam design [13]. This could be straightforwardly generalized to other cases, for example for an HWP with more than 3 layers, as proposed for the LiteBIRD mission which will deploy a 9 layers on its low-frequency telescope [11,22].

The Mueller matrix of the 3-layer HWP that we consider can therefore be written as:

$$\mathbf{M}_{\text{HWP}} = \mathbf{M}_{\text{layer}} \mathbf{R}(-2\alpha_2) \mathbf{M}_{\text{layer}} \mathbf{R}(2\alpha_2) \mathbf{M}_{\text{layer}}, \quad (62)$$

with  $\mathbf{M}_{\text{layer}}$  defined as in Eq. (2). The Mueller matrix of a single layer is parametrized by  $\delta$ , Eq. (3). As already pointed out in Sec. II A 1, once the birefringent material is chosen,  $\delta$  depends on observing frequency,  $\nu$ , and thickness of each layer,  $\theta_{\text{hwp}}$ . For a given HWP, we assume all layers to be identical, and parametrize  $\mathbf{M}_{\text{layer}}$  with only its thickness  $\theta_{\text{hwp}}$ .

Based on existing optimized parameters [23] and proposed designs [10,19,21], we choose the following nominal values for HWP parameters:

- (i)  $\theta_{\text{hwp}}(\text{LF}) = 14.36$  mm
- (ii)  $\theta_{\text{hwp}}(\text{MF}) = 3.8$  mm
- (iii)  $\theta_{\text{hwp}}(\text{UHF}) = 1.86$  mm
- (iv)  $\alpha_2 = 58^\circ$

We note here that  $\theta_{\text{hwp}}$  is different for each telescope since the thickness is adapted to the frequency range of observation, when  $\alpha_2$  is in principle the same. However, to account for possible fabrication and/or calibration differences between HWP, we allow  $\alpha_2$  to vary for each of the telescopes when fitting for parameters (although they have the same nominal value).

Details of  $\mu_{ij}$ , the coefficients of the HWP Mueller matrix,  $\mathbf{M}_{\text{HWP}}$ , as functions of  $\alpha_2$  and  $\theta_{\text{hwp}}$  can be found in Appendix B.

## 3. Bandpasses

We parametrize each bandpass by two parameters: its center  $\nu_0$  and its half-width  $\Delta\nu$ . For the mock data we assume the standard 30% bandpass bandwidths for all frequency bands and therefore set  $\Delta\nu = 0.15 \times \nu_0$ . In the model, we allow  $\nu_0$  and  $\Delta\nu$  to vary independently.

We model the bandpass as a top-hat function with smoothed edges as shown in Fig. 3. The analytic expression of the bandpass  $\mathcal{B}(\nu)$  as a function of  $\nu_0$  and  $\Delta\nu$  is

$$\mathcal{B}(\nu) = \exp\left[-\left(\frac{|\nu - \nu_0|}{\Delta\nu}\right)^{20}\right]. \quad (63)$$

Realistic parametric modeling of bandpass can be complex because true bandpasses are never as regular as the one shown in Fig. 3, see, e.g., [24]. The idealized model in Eq. (63) is however convenient as it permits studying the impact on the results of the two main bandpasses characteristics: the band center and its width.

Concurrently, we also explore the effects of potential deviations from this simplified bandpass form. This can reflect more complex intrinsic shape of the realistic

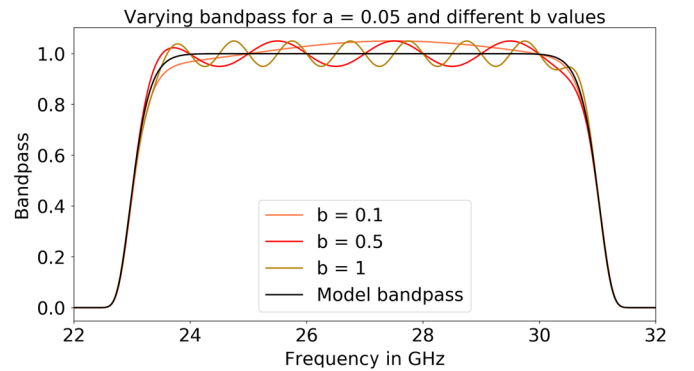


FIG. 3. Ideal and varying bandpass model for the 27 GHz channel, for  $a = 0.05$  and different  $b$  values following Eq. (64). We first use the ideal bandpass in both data and model, and we then introduce a varying bandpass in data only, to study the impact of this mismatch on systematic residuals, as detailed in Sec. IV D 2.

bandpasses but also can be due to detector-to-detector variations. To this end we propose a toy-model given by:

$$\mathcal{B}'(\nu) = \mathcal{B}(\nu) \times [1 - a \times \sin(b \times 2\pi\nu)], \quad (64)$$

and which includes deviations from the idealized, ‘‘average,’’ form in Eq. (63). An example of such bandpasses is shown in Fig. 3.

In the following, we first use the simplified bandpass model,  $\mathcal{B}(\nu)$ , in both the input data and the model, to assess the impact of uncertainty in the bandpass position and width on component separation. Subsequently, we use the more complex form,  $\mathcal{B}'(\nu)$ , to generate input data, but we continue fitting the simple model of Eq. (63) to the data, thus neglecting the presence of the small-scale features in the bandpasses, which are left unmodeled. As pointed out in Sec. III C, a mismatch between the data and the fitted model will typically result in systematic residuals in reconstructed maps. As mentioned earlier such a mismatch could arise due to detector-to-detector variability of the bandpasses. We therefore explore the impact on systematic residuals level and cosmological parameters of various  $a$  and  $b$  values in Sec. IV D 2, trying to understand what level of unmodeled deviations, in terms of their scale or their amplitude, can be tolerated and at what point they need to be modeled (or measured in advance).

#### 4. Summary of instrumental parameters

Table I summarizes the hardware parameters that we consider in this model and their nominal values, used to generate the input data. We recall here that all parameters are independent and can in principle all vary when we minimize the likelihood of Eq. (39).

In contrast, while we include in our data models parameters describing sinuous antennas, Sec. II A 2, those are fixed throughout this analysis. Similarly, although beam modeling and beam systematics could also be investigated within this framework, we leave this for future work and do not assume any beam smoothing of input maps. Note that beam smoothing is taken into account in the harmonic domain when we estimate noise after component separation in Eq. (57). Again, beam sizes that we adopt here correspond roughly to the Simons Observatory Small Aperture Telescopes ones [19].

TABLE I. Nominal values of instrumental parameters.

	LF		MF		UHF		
	LF1	LF2	MF1	MF2	UHF1	UHF2	
HWP	$\alpha_2$	58°	58°	58°	58°		
	$\theta_{\text{hwp}}$	14.36 mm	3.8 mm		1.86 mm		
Bandpass (GHz)	$\nu_0$	27	39	93	145	225	280
	$\Delta\nu$	4.05	5.85	13.95	21.75	33.75	42.0

#### 5. Input sky

We use the PYSM package [25] to generate foregrounds, and the HEALPY package [26,27] to generate CMB maps from fiducial CMB power spectra. We choose the following reference frequencies for the foregrounds templates:

$$\nu_{0,\text{sync}} = 70 \text{ GHz}, \quad \nu_{0,\text{dust}} = 353 \text{ GHz} \quad (65)$$

We fix the dust temperature at  $T = 19.6$  K and for spectral indices we consider the nominal values:

$$\beta_d = 1.59, \quad \beta_s = -3.1. \quad (66)$$

We take constant foreground parameters across the sky, although the effect of varying spectral parameters could be included within our framework. For details we refer the reader to [9].

PYSM uses the HEALPY wrapper of the HEALPIX pixelization scheme [26,28], and we use `NSIDE` = 256 throughout this work. We use a mask corresponding to 10% of the sky. When it comes to power spectrum reconstruction, we therefore consider multipoles from  $\ell = 30$  to  $\ell \sim 500$ , given the limitation by the mask at low  $\ell$ , and by the pixel size at high  $\ell$ .

We consider two cosmological parameters, the tensor-to-scalar ratio  $r$  which sets the amplitude of primordial B modes:

$$C_{\ell,\text{primordial}}^{BB} = r \times C_{\ell,\text{primordial}}^{BB}(r = 1), \quad (67)$$

and the lensing parameter which sets the delensing amplitude ( $A_L = 1$  corresponds to no delensing and  $A_L = 0$  corresponds to full delensing):

$$C_{\ell,\text{lensing}}^{BB} = A_L \times C_{\ell,\text{lensing}}^{BB}(A_L = 1). \quad (68)$$

We choose

$$r = 0, \quad A_L = 1, \quad (69)$$

as our reference case but also explore other cases in Sec. IV D.

Finally, as we aim at modeling a ground-based experiment, we consider that sky-only modulated terms, i.e.,  $\mathcal{C}_0$  and  $\mathcal{S}_0$  in our model, will be compromised by the atmospheric noise at long temporal modes of the collected data streams [29,30]. In what follows, we therefore use only  $\mathcal{C}_4$  and  $\mathcal{S}_4$ , in the simulated data and in the model. However, we point out that, for a space mission such as LiteBIRD, it could be possible to recover all four mixed polarized Stokes components.

#### B. Optimization of the generalized spectral likelihood

In this section, we explore the parameter space described in the previous section: 18 instrumental parameters

(6 for the HWPs and 12 for the bandpasses), as well as 2 foreground spectral parameters  $\beta_s$  and  $\beta_d$ . We start by fixing all instrumental parameters to their nominal values, and we only estimate foreground parameters,  $\beta_d$  and  $\beta_s$ . This constitutes our reference scenario for uncertainties, residuals and bias on  $r$ , and we compare results obtained when estimating instrumental parameters to this case. We then progressively free instrumental parameters in the spectral likelihood, Eq. (37). We do not consider priors in the first instance, but will introduce them for specific cases whenever it becomes necessary.

We consider the following cases, and corresponding abbreviations to identify them throughout this work:

- (i) Spectral Energy Distribution (SED) parameters (foreground spectral indices) only  $\rightarrow$  SED only
- (ii) SED + HWP central layer angle for all three HWP  $\rightarrow$  SED +  $\alpha_2$
- (iii) SED + HWP layer thickness for all three HWP  $\rightarrow$  SED +  $\theta_{\text{hwp}}$
- (iv) SED + HWP central layer angle + HWP layer thickness for all three HWP  $\rightarrow$  SED + HWP
- (v) SED + Bandcenters for all bandpasses  $\rightarrow$  SED +  $\nu_0$
- (vi) SED + Bandwidths for all bandpasses  $\rightarrow$  SED +  $\Delta\nu$
- (vii) SED + Bandcenters + Bandwidths for all bandpasses  $\rightarrow$  SED + Bandpass
- (viii) SED + all the above  $\rightarrow$  SED + All

### 1. Method

For each case, once the generalized spectral likelihood is optimized, we compute the Hessian matrix  $\mathcal{H}$  as defined in Eq. (40) at the numerically determined peak of the likelihood. We compare the one-dimensional spectral likelihood where we fix all parameters but one to their nominal value, to a Gaussian function whose variance is determined by the diagonal of the Hessian matrix:

$$\sigma_i = \frac{1}{\sqrt{\mathcal{H}_{ii}}} \quad (70)$$

for any parameter  $i$ .

We also compute the eigenvalues and eigenvectors decomposition of the Hessian matrix, to test for the presence of possible degeneracies in the considered parameter space.

Finally, we evaluate the marginalized error bar on spectral parameter using the diagonal of the inverse of the Hessian matrix,  $\Sigma$ , as defined in Eq. (41):

$$\sigma(\beta_i) = \sqrt{\Sigma_{\beta_i, \beta_i}}. \quad (71)$$

The errors on instrumental and foreground parameters are of primary importance as they determine the amplitudes of both statistical and systematic residuals which impact directly our estimates of  $r$ . Generically, we expect that

uncertainty on these parameters will increase as we increase the number of free instrumental parameters.

We emphasize however that there are fundamental differences in the way the data constrain both these types of parameters, and therefore the number of parameters that can be constrained. Indeed, the foreground parameters are usually specific to a sky component, be that dust or synchrotron, and therefore tend to affect only a single column of the generalized mixing matrix. In contrast, the instrumental parameters are rather frequency channel specific and determine the rows of the matrix. As a consequence, the constraints on the instrumental parameters do not benefit from the multifrequency information as much as do the foregrounds parameters. In particular, the quality of the constraints do not improve with the increasing number of frequency bands (while the number of parameters will typically increase). In the simple case of instrumental parameters specific to only a single frequency channel we expect that only one such parameter can be well constrained with the data. However, as we measure a few pieces of information per frequency channel, corresponding to the mixed-Stokes maps which we consider for each frequency channel, we could in principle set more than a single constraint, and up to as many as the available pieces of information. Most of these constraints, in fact all but one, will typically be of significantly lower precision and often insufficient for our purpose, rendering only one, useful constraints on a linear combination of the considered parameters. If instrumental parameters are relevant to more than a single frequency channel the number of potentially well constrained parameters increases. We note that these analytic insights are fully borne-out by our numerical results as we discuss in the following section.

### 2. HWP parameters

We first consider the following cases, involving only HWP parameters:

- (i) SED +  $\alpha_2$
- (ii) SED +  $\theta_{\text{hwp}}$
- (iii) SED + HWP

We find that all the likelihoods corresponding to these cases are very well approximated by a Gaussian, with the dispersion obtained by the corresponding inverse Hessian calculated at the peak of each likelihood. Moreover, the Hessians are very well conditioned, and all the parameters in all these cases are consequently well constrained and only weakly correlated.

These observations apply to conditional and marginalized likelihoods. We then consequently use the Gaussian approximations and the Hessians to compute the marginalized uncertainty on spectral parameters [defined in Eq. (71)] in these various cases. We find that there is no significant increase of the marginalized uncertainty on spectral indices when adding HWP parameters, as shown in Table II and Fig. 4.

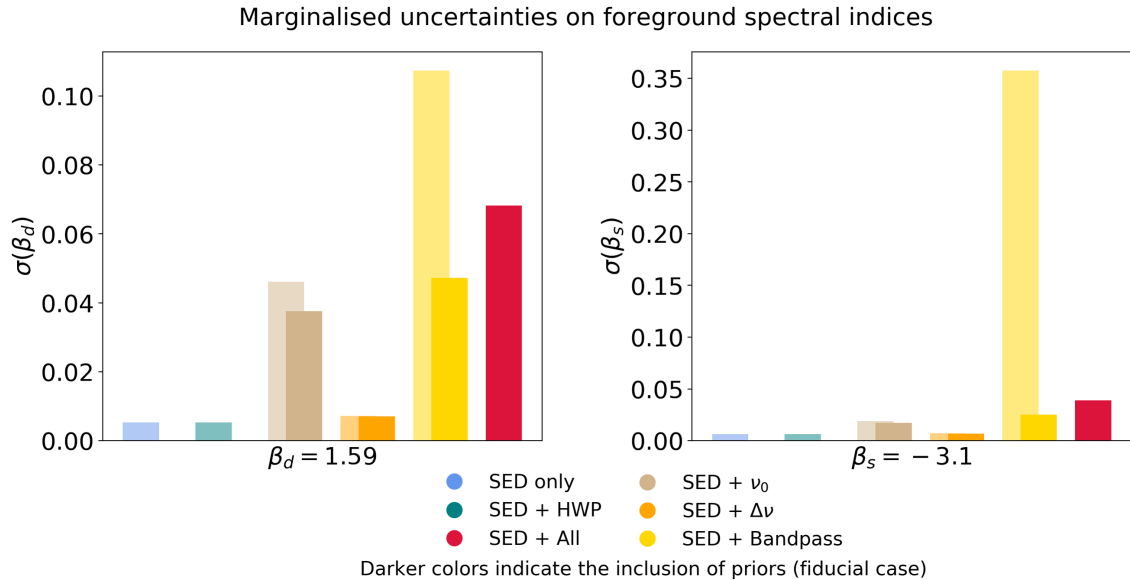


FIG. 4. Marginalized 1- $\sigma$  uncertainties on foreground spectral indices. The inclusion of HWP parameters has a negligible impact on the uncertainty on spectral parameters, which is why we do not consider the inclusion of priors. However, the effect is much more sensitive when including bandpass parameters, thus the inclusion of priors. The uncertainty on spectral parameters when considering only  $\nu_0$  or  $\Delta\nu$  does not change much when adding priors, as shown in Table II. However, we show that adding priors when all bandpass parameters are considered is required to limit uncertainty on foreground parameters and avoid degeneracies for instrumental parameters. In the case when all parameters are considered, priors are necessary because the system is otherwise degenerate.

We thus expect that when only HWP parameters are considered, the impact on foreground residuals and  $r$  estimation should be limited, as we will demonstrate in Sec. IV C. We therefore conclude that in none of these cases we need priors to properly constrain HWP parameters. Though this conclusion is drawn in the case of the specific model of the HWP and the specific experimental setup, we expect that it will hold as long as the number of HWP parameters does not exceed the number of frequency channels in the corresponding focal plane, otherwise some prior knowledge may become necessary.

### 3. Bandpass parameters

In the case of bandpasses we consider the following cases:

- (i) SED +  $\nu_0$
- (ii) SED +  $\Delta\nu$
- (iii) SED + Bandpass

*Without priors.*—In the cases with only either bandcenters or bandwidths in addition to foreground parameters allowed to change, i.e., {SED +  $\nu_0$ } and {SED +  $\Delta\nu$ }, the system is not degenerate: its Hessian matrix is still formally positive definite as all eigenvalues are strictly positive. The spectral likelihood is well approximated by a Gaussian in these cases and we use the Hessian to set the constraints. We note that though the constraints on the bandpass centers are more precise relative to the parameter values than those obtained for the bandwidths, the impact

of the former on the spectral parameters errors is significantly more pronounced. This is shown in Fig. 4. We discuss the impact of this observation on the shape of foreground residuals in the next section.

When all bandpass parameters are allowed to vary at the same time, i.e., case {SED + Bandpass}, the system becomes nearly degenerate as we have two independent parameters for each frequency channel. In this case, the uncertainty on both spectral and instrumental parameters increases significantly. The Hessian matrix is still positive definite and errors on all parameters can be formally set, however, the eigenvalues of the Hessian matrix span now a range from  $10^4$  to  $10^{-2}$ , rendering some of the constraints weak. As expected, the marginalized errors on spectral parameters largely increase in this case, as shown in Fig. 4 and Table II.

The near degeneracies in this case appear between a bandcenter and bandwidth of the same frequency band. This is illustrated by the elongated shape of the likelihood in a 2d space ( $\nu_0, \Delta\nu$ ) for the same bandpass as shown in Fig. 5. We note that the Gaussian approximation fails in these cases to reproduce the behavior of the actual likelihood, but only in the ill-constrained directions. For the well-constrained directions, the Gaussian approximation and the Hessian continue however providing a very good description.

We thus conclude that adding bandpass parameters significantly increases uncertainties on both spectral and instrumental parameters, in particular in comparison with the case of the HWP parameters, impact of which was

TABLE II. Summary of  $1\text{-}\sigma$  marginalized uncertainties on spectral parameters. Priors refer to the fiducial case: 1% on bandcenters ( $\nu_0$ ) and 5% on bandwidths ( $\Delta\nu$ ).

	Priors	$\sigma(\beta_d = 1.59)$	$\sigma(\beta_s = -3.1)$
SED only	No	$5.291 \times 10^{-3}$	$6.420 \times 10^{-3}$
SED + $\alpha_2$	No	$5.292 \times 10^{-3}$	$6.423 \times 10^{-3}$
SED + $\theta_{\text{hwp}}$	No	$5.292 \times 10^{-3}$	$6.422 \times 10^{-3}$
SED + HWP	No	$5.300 \times 10^{-3}$	$6.444 \times 10^{-3}$
SED + $\nu_0$	No	$4.609 \times 10^{-2}$	$1.890 \times 10^{-2}$
SED + $\nu_0$	Yes	$3.761 \times 10^{-2}$	$1.736 \times 10^{-2}$
SED + $\Delta\nu$	No	$7.155 \times 10^{-3}$	$7.129 \times 10^{-3}$
SED + $\Delta\nu$	Yes	$7.014 \times 10^{-3}$	$6.538 \times 10^{-3}$
SED + Bandpass	No	$1.074 \times 10^{-1}$	$3.577 \times 10^{-1}$
SED + Bandpass	Yes	$4.721 \times 10^{-2}$	$2.493 \times 10^{-2}$
SED + All	No	Degenerate	Degenerate
SED + All	Yes	$6.825 \times 10^{-2}$	$3.903 \times 10^{-2}$

found to be very minor. We therefore consider the inclusion of calibration priors on bandpass parameters to alleviate this issue.

*With priors.*—To remove bandpass-associated degeneracies in the parameter space, we introduce Gaussian priors on  $\nu_0$  and  $\Delta\nu$  as proposed in Eq. (39).

In general, the eigenvalues decomposition of the Hessian matrix provides good guidance on how to introduce priors in a most efficient and economic way. We start from the lowest eigenvalues and candidate parameters are those which are most aligned with the corresponding eigenvector. In the specific case here, the lowest eigenvalues of the Hessian matrix are associated with eigenvectors with dominant components corresponding to the bandpass

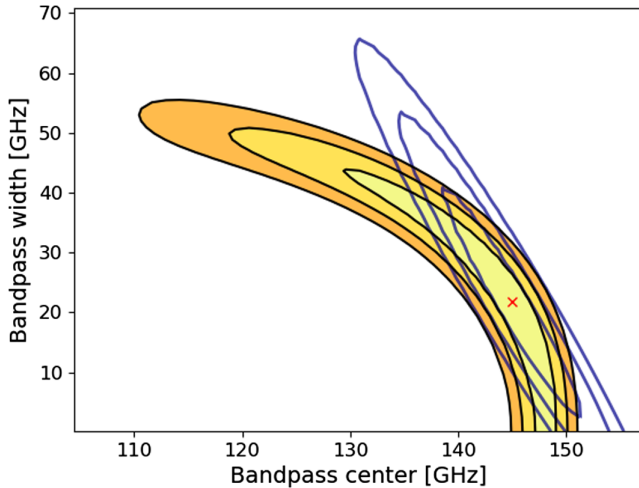


FIG. 5. Two-dimensional gridding of the generalized spectral likelihood, Eq. (37), estimated from 100 random sky pixels. We overplot in blue the Gaussian approximation given by the Hessian, and show that in this case it fails to describe the system. The red cross indicate true values of parameters.

 TABLE III. Marginalized  $1\text{-}\sigma$  uncertainties on spectral parameters when considering bandpass parameters for different choices of priors.

	Priors	$\sigma(\beta_d = 1.59)$	$\sigma(\beta_s = -3.1)$
SED + Bandpass	No	0.108	0.358
	Pessimistic	0.0690	0.0419
	Fiducial	0.0472	0.0249
	Optimistic	0.0281	0.0161

parameters of the same frequency band. This is illustrated in Fig. 5. As we found that the eigenvectors components in the directions of the band center and widths are comparable, we decided to set prior constraints on both of them for each band. We note that while other priors can also be considered here and may be helpful, the bandpass parameter degeneracies are the only degeneracies found in the studied problem and therefore the priors on these parameters are the most efficient and have the most pronounced impact.

In the absence of calibration data, we test three levels of the priors, in order to best illustrate the impact of them on the conclusions. These are

- (i) Pessimistic: 5% on  $\nu_0$  and 8% on  $\Delta\nu$ ;
- (ii) Fiducial: 1% on  $\nu_0$  and 5% on  $\Delta\nu$ ;
- (iii) Optimistic: 0.5% on  $\nu_0$  and 1% on  $\Delta\nu$ .

The corresponding uncertainties on spectral parameters for these different choices are listed on Table III.

We also show uncertainties with and without priors on Fig. 4 for the fiducial case. The priors are efficient to reduce uncertainties in all considered cases, but their effect is most pronounced whenever all bandpass parameters are allowed to vary. These are indeed the cases when the system is nearly degenerate—and the Hessian nearly singular. Introducing the priors reduces the range of eigenvalues by at least one or two orders of magnitude as does the condition number of the Hessian matrix.

When it comes to the level of priors, even not very accurate priors allow for a significant reduction of marginalized uncertainties on spectral parameters. However, the change in eigenvalues is not significant between pessimistic and fiducial cases, although it is better than in the no prior case. In what follows, unless otherwise is specified, we choose to keep the fiducial values as our nominal level of the priors, as they are closer to currently achieved calibration performance.

This choice of priors is not unrealistic at least in the cases of limited number of detectors, for instance compared to calibration performance currently achieved with a Fourier-transform spectrometer (FTS) on a typical CMB experiment. Accuracy of such measurements has been demonstrated to be at the level of 1 GHz for the POLARBEAR experiment [31] for example. Future experiments such as LiteBIRD plan on even higher accuracy on bandpass calibration, up to 0.2 GHz [32]. The priors at the level

we assume here should therefore be available for future analyses.

Overall, this shows that the inclusion of the priors at the level we propose is necessary and efficient to reduce marginalized uncertainties and lift degeneracies in the multiparameter space. We also show in the next section that it reduces foreground residuals and bias on  $r$ .

#### 4. All parameters

We consider here the case of

(i) SED + All,

when all instrumental and foreground parameters are allowed to vary. In this case, if we do not include priors, the Hessian matrix is nearly singular with some of the eigenvalues numerically zero. This is expected given the large number of parameters. We therefore once again include priors on bandpass parameters as discussed in the previous section, which resolves the degeneracy as shown in Table II. In the follow-up, unless otherwise specified, in the case {SED + All} we always consider fiducial priors on bandpass parameters.

#### 5. Summary

In this first part of our work, we have studied parameter space degeneracies as well as statistical uncertainties on spectral parameters estimated by the Hessian matrix at the peak of the spectral likelihood. We showed that we can estimate HWP parameters from the actual data with no significant increase of marginalized uncertainties, and without adding priors. This is due to the fact that HWP parameters apply to more than one frequency band, and therefore we have enough leverage to constrain them. However, when it comes to bandpass parameters, we

showed that it is necessary to introduce priors on both bandcenters and bandwidths to avoid degeneracies in the parameter space and maintain marginalized error bars on spectral parameters low enough. This is due to the fact that we have two bandpass parameters (bandwidth and bandcenter) for only one frequency band, and therefore these two parameters are degenerate. In what follows, unless otherwise is specified, we keep the following priors on bandpass parameters anytime these parameters are estimated: 1% on  $\nu_0$  and 5% on  $\Delta\nu$ .

Overall, and as it could have been expected, we show that the more parameter we consider in the spectral likelihood, the higher statistical uncertainties. However, different instrumental parameters affect the uncertainties on spectral parameters to a different degree. All results of this section are summarized in Table II. All these factors determine the level and shape of foreground residuals and the estimation of  $r$ , as we will now demonstrate.

### C. Residuals

#### 1. Statistical residuals

At first, we assume the same model underlying the dataset as it is used for its analysis. In this case, we expect no systematic residuals, and we can therefore limit ourselves to statistical residuals as defined in Eqs. (53). We compute foreground residuals for all cases studied in the previous section as shown in Fig. 6.

As expected, the level of statistical residuals scales directly as the statistical uncertainties on recovered foreground parameters: the more parameters, the greater the uncertainties and the higher the statistical residuals. Including priors on bandpass parameters lower the level of residuals as expected.

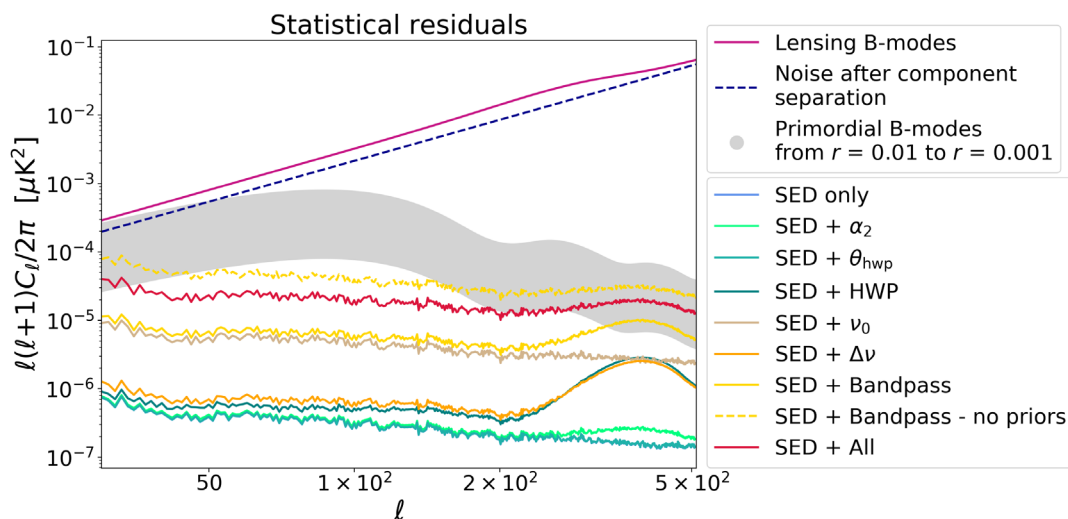


FIG. 6. Foreground statistical residuals on B-modes, for various instrumental parameter sets. As expected, statistical residuals scale as statistical uncertainty on the estimated parameters. The bump at high  $\ell$  is due to statistical uncertainty dominated by uncertainty on instrumental parameters, resulting into a CMB E- to B-mode leakage.



We note that, despite being very similar in shape at low  $\ell$ , the various residuals curves differ occasionally at higher  $\ell$ : the bump is due to the leakage of CMB polarization signal (E-modes) to the total residuals, referred to as the CMB residuals in Eq. (44). The relative amplitude of both residual terms (foregrounds and CMB) depends on the instrumental parameters. For the parameters which do not affect much the determination of the spectral parameters, the additional foreground residuals and the CMB residuals scale the same way, as they both are proportional to the uncertainties on these parameters. As the CMB E-mode power dominates that of the foregrounds at small angular scales the CMB residuals may then become dominant at high- $\ell$  s. If the instrumental parameters are strongly coupled to the spectral ones, there is an extra increase of the foreground residuals because of the larger errors on the spectral parameters themselves and the foreground residuals may therefore dominate over the entire range of angular scales considered here. In such cases no high- $\ell$  bump is seen in the residuals spectra.

We validate these expectations by computing residuals assuming no CMB and show that this distinctive feature disappears. This is illustrated in Fig. 7.

As we observed in Sec. IV B 3 (see Fig. 4), marginalized uncertainties on spectral parameters in the  $\{\text{SED} + \nu_0\}$  case are much higher than in the  $\{\text{SED only}\}$  case, while there is only a negligible increase in the  $\{\text{SED} + \Delta\nu\}$  case. This is reflected in Fig. 6, where at low  $\ell$  the statistical residuals in the  $\{\text{SED} + \nu_0\}$  case are higher than in the  $\{\text{SED} + \Delta\nu\}$  case. Moreover, in the former case, unlike in the latter, there is no high- $\ell$  bump in the residuals, and the residuals are dominated by the foreground term. We also note that including either the bandpass widths,

$\{\text{SED} + \Delta\nu\}$ , or the HWP parameters,  $\{\text{SED} + \text{HWP}\}$ , gives rise to residuals which are very much comparable across the entire range of the considered angular scales. This is because in both these cases the spectral parameters uncertainties are not significantly affected by these instrumental parameters, whereas the instrumental parameters themselves are estimated with similar precision. Therefore the levels of the foreground and CMB residuals are expected to be indeed comparable.

In Fig. 7 we show the impact of priors on the residuals. We see that more stringent priors indeed lower the overall level of the residuals but also that the relative import of the CMB residual progressively decreases.

## 2. Systematic residuals

We now introduce a mismatch between simulated data and the model. As outlined in Sec. IV A 3, we investigate the effect of slowly varying, more realistic bandpass in the data, while we keep on using a simple smoothed top-hat model in the analysis. We consider the following cases:

- (i) SED only
- (ii) SED + HWP
- (iii) SED + Bandpass
- (iv) SED + All

For each case, we compute residuals when introducing a discrepancy between bandpass data and model, with  $a = 0.01$  and  $b = 1$  in the varying bandpass. We compare them to residuals in the same instrumental configuration, but with no bandpass variation in the input data. For both statistical [Eq. (53)] and systematic residuals [Eq. (52)], we verify that they behave as expected:

- (i) no notable difference between statistical residuals, as they are unaffected by discrepancy between data and model;

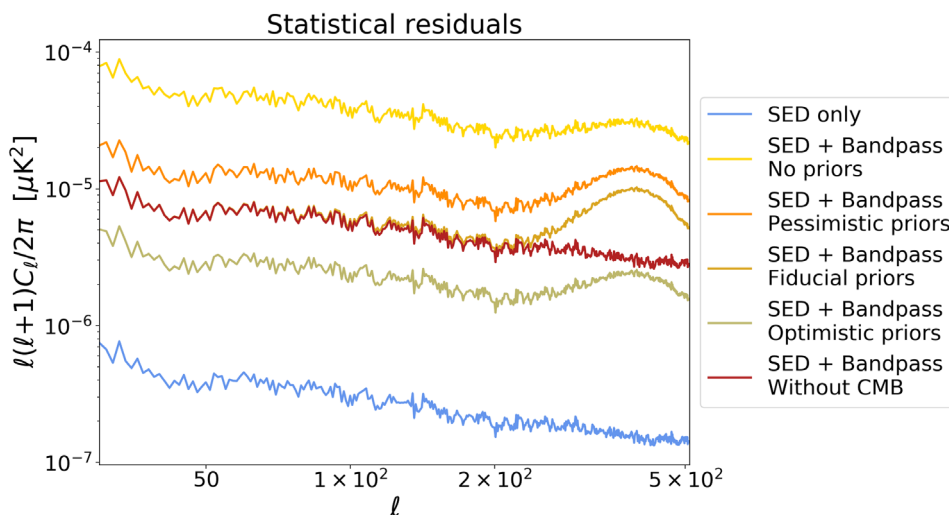


FIG. 7. Foreground statistical residuals on B-modes, with and without including CMB contribution. We also compare different level of priors on bandwidths, to show the effective reduction of statistical residuals by choosing more stringent priors. The case without CMB is computed with fiducial priors.

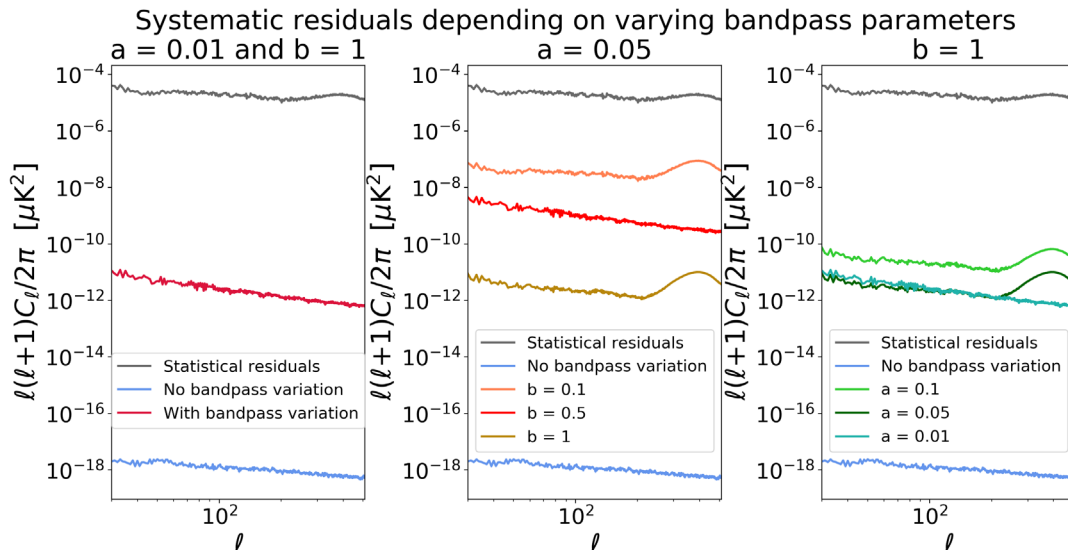


FIG. 8. Left: Statistical and systematic residuals, showing the increase of systematic residuals when introducing a mismatch between bandpass in data and model when considering the  $\{\text{SED} + \text{All}\}$  case. Center and right: scaling of systematic residuals depending on bandpass variation, compared to statistical residuals. We consider here all 20 parameters, i.e., the  $\{\text{SED} + \text{All}\}$  case. The systematic residuals in the case with no bandpass variation should be zero, the light blue curve defines therefore numerical precision of our computations.

- (ii) increase of systematic residuals as they are precisely sensitive to the mismatch between data and its model.

We demonstrate that this is true for all cases and, for reference, show the result for the  $\{\text{SED} + \text{All}\}$  case in the left panel of Fig. 8.

The level of systematic residuals depends on parameters  $a$  and  $b$  (defined in Sec. IV A 3), which set the amplitude and frequency of bandpass variations. We investigate this dependence focusing on the  $\{\text{SED} + \text{All}\}$  case, as this is the case where we expect the biggest impact. We consider various  $\{a, b\}$  combinations and compute corresponding residuals. We find that, for a fixed value of  $a$ , i.e., the amplitude of the bandpass variation, a smaller  $b$ , corresponding to variations on larger scales, leads to an increase in the systematic residuals. On the contrary, the effects of small-scale features in the bandpass are quickly becoming very small, suggesting that characterizing the bandpasses with a  $\sim 1$  GHz sampling as this is most often the case in practice [31], should indeed be sufficient. Moreover, for a given value of  $b$ , the systematics residuals increase with increasing value of  $a$ . This is an expected effect, since  $a$  sets the amplitude of bandpass variation. These two effects are demonstrated on Fig. 8. In Sec. IV D 2 hereafter, we comment on the impact of the systematic residuals on  $r$  estimation.

#### D. Constraints on tensor-to-scalar ratio

With the residuals estimated for various cases, we now estimate their impact on the determination of the tensor-to-scalar ratio  $r$ . We use the cosmological likelihood in

Eq. (58) and consider two models for the assumed covariance of the CMB signal,  $\mathbf{C}$ , as described in Sec. III E. These correspond to two extreme cases when we either assume complete ignorance of the residuals present in the cleaned CMB maps, Eq. (59), or when the statistical residuals are modeled and deprojected consistently, Eq. (60). We focus only on the B-mode polarization and the CMB covariance includes only the BB power spectra parametrized by the tensor-to-scalar ratio  $r$  as elaborated on in Sec. IV A 5.

##### 1. Cases with statistical residuals only

In this case the model underlying the data coincides with that assumed in the analysis and the only residuals are statistical. We first assume that the actual value of  $r$  is zero.

In this first approach we ignore the residuals present in the recovered CMB map and therefore expect that any residual, be it statistical or systematic, will result in a bias on  $r$  which in turn will depend on their level. The results are shown in Fig. 9 and the corresponding values of  $r$  and  $\sigma(r)$  for all instrumental parameter sets are collected in Table IV.

As expected, the higher the residuals, the bigger the bias on  $r$ . We can therefore suppress the bias by suppressing the residuals. This can be achieved with help of additional priors on instrumental parameters. As shown in Fig. 9 and Table IV, this way is indeed efficient in reducing the bias on  $r$  and ensuring that it does not lead to an erroneous but statistically significant detection. Its downside is that it may call for high precision priors. This can be particularly demanding if a large number of relevant instrumental effects is present.

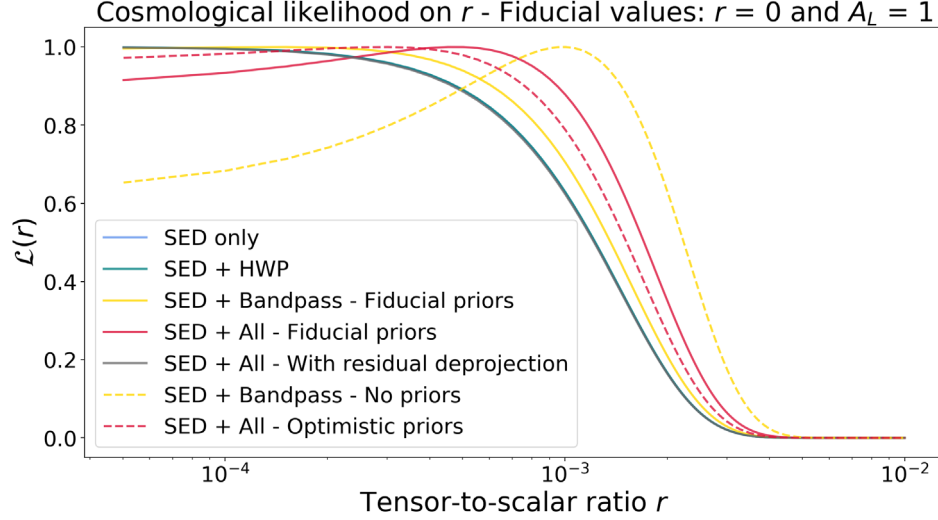


FIG. 9. Cosmological likelihood Eq. (58) as a function of the tensor-to-scalar ratio  $r$ , with its fiducial value assumed to be 0. In all cases, but one as marked in the legend, no deprojection is applied. In these cases the priors are shown to be efficient to alleviate the bias on  $r$ . The deprojection of the residuals not only resolves the issue of the bias (due to the statistical residuals) but also recovers the lowest, theoretically possible uncertainty on  $r$ .

The statistical residuals can be effectively marginalised, or deprojected, by using the model covariance in Eq. (60), therefore minimizing potential bias of the measured value of  $r$ . In this case, we may hope that the need for high precision priors can at least be partly alleviated. This expectation is indeed confirmed in the cases studied here, as it is shown with a gray solid line in Fig. 9. The loss of accuracy and precision seems to be in this case negligible. In more realistic cases and in the presence of a potentially significant mismatch between the assumed sky and instrument models, and the true ones, the deprojection may not be however as successful as in the simple cases studied here. We expect therefore that the calibration priors will continue playing a key role in the component separation process either in order to suppress the residuals, as in the first method described above, or to demonstrate that the deprojection was indeed performed successfully.

TABLE IV. Best-fit values for  $r$  and  $\sigma(r)$  without the residuals deprojection. Corresponding cosmological likelihoods for selected cases are shown in Figure 9.

	$r$ ( $\times 10^{-5}$ )	$\sigma(r)$ ( $\times 10^{-3}$ )
SED only	0.819	1.47
SED + $\alpha_2$	0.870	1.47
SED + $\theta_{\text{hwp}}$	0.823	1.47
SED + HWP	1.15	1.47
SED + $\nu_0$ —fiducial priors	8.40	1.47
SED + $\Delta\nu$ —fiducial priors	1.38	1.47
SED + Bandpass—no priors	98.6	1.51
SED + Bandpass—fiducial priors	14.1	1.48
SED + All—fiducial priors	47.6	1.49
SED + All—optimistic priors	29.1	1.48

We also test our framework with other values of cosmological parameters considering for the tensor-to-scalar ratio a value of  $r = 0.01$  consistent with what typical Stage 3 experiment plan to achieve and assuming partial delensing with  $A_L = 0.5$  in addition to the no-delensing case  $A_L = 1$ . We limit ourselves to the case with all instrumental parameters and include priors, {SED + All}. Results without the residual deprojection are shown in Fig. 10. We show that in all considered cases, our estimates of the value of  $r$  are consistent with the true values with precision much better than the  $1 - \sigma(r)$  uncertainty, even if we do not deproject statistical residuals. As shown in Fig. 10 and Table V, the uncertainty on  $r$  is reduced when we assume partial delensing, as expected.

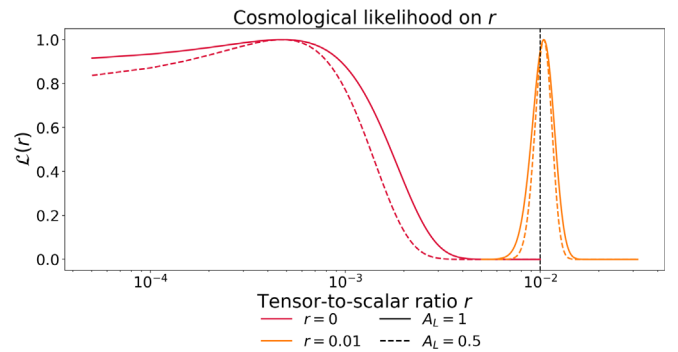


FIG. 10. Cosmological likelihood on  $r$  without the residual deprojection for two different fiducial values of the tensor-to-scalar ratio,  $r = 0$  or  $r = 0.01$ , and assuming either no,  $A_L = 1$ , or partial, 50% delensing,  $A_L = 0.5$ . The shown results are for the case of {SED + All}, thus including priors on bandpass parameters.

TABLE V. Uncertainties on  $r$  without residual deprojection, with varying tensor-to-scalar ratio  $r$  and lensing amplitude  $A_L$ , in the case of SED + All, including priors on bandpasses.

	$A_L = 1$	$A_L = 0.5$
$r = 0$	$\sigma(r) = 1.49 \times 10^{-3}$	$\sigma(r) = 1.05 \times 10^{-3}$
$r = 0.01$	$\sigma(r) = 1.96 \times 10^{-3}$	$\sigma(r) = 1.51 \times 10^{-3}$

Overall, these results demonstrate that the proposed framework allows to correct for the effects related to the presence of the HWP and bandpasses without compromising the precision of the constraints which can be set on the tensor-to-scalar ratio,  $r$ , and does so down to the precision levels expected for typical Stage-3 experiments. In particular, the limits we derive within this framework are in line with Simons Observatory forecast, as shown in e.g., Tables 4 and 5 of [19]. However, we emphasize that in our analysis, we have not fully accounted for all characteristics of the Simons Observatory instrument model, nor have we included  $1/f$  noise.

## 2. Impact of bandpass imperfections

As we have shown previously, the mismatch in the bandpass parameters unavoidably leads to the presence of the systematic residuals. These were found to be subdominant with respect to the statistical residuals in the all the cases studied in this work, Sec. IV C 2. As long as no deprojection is applied both types of the residuals impact the estimates of  $r$  in the same way. Consequently, in all our cases the effect on  $r$  due to the systematic residuals is found to be subdominant and negligible. This remains to be the case also when the statistical residuals are deprojected. The impact of the systematic residuals on the value of  $r$  is then very minor and much smaller than the statistical error on  $r$ . This is due to the low absolute level of the systematic residuals present in the studied cases. This result is consistent with the complementary analysis [33] which in addition to bandpasses also includes other hardware parameters, and performs component separation in the power spectrum domain.

## V. CONCLUSIONS

In this work we have extended the standard CMB data analysis pipeline to include explicitly a treatment of instrumental effects. We focused on two key data analysis stages, map-making, and component separation, and considered, as an example, instrumental effects related to the presence of a broadband HWP in the instrument optical design, and bandpasses defining the frequency bands of the observations.

We have subsequently implemented the proposed framework as part of the performance forecasting tool XFORECAST [8,9], extending it to account for instrumental

effects. This has allowed us to propagate the impact of the instrumental effects all the way to cosmological constraints.

We have applied the method in the context of a modern CMB experiment of the 3rd generation, modeled on the small aperture telescopes of Simons Observatory [19] assessing its performance in the light of setting constraints on the tensor-to-scalar ratio  $r$ . We have discussed the role and impact of the instrumental parameters on the cosmological constraints. The constraints on cosmological parameters that we derive, as well as the level of instrumental priors that we suggest are consistent with results presented in [33], which have been derived using a different approach and are more specific to the Simons Observatory. This shows that our method is reliable in forecasting experimental performance.

We have shown how calibration information can be incorporated in the analysis and how to determine which instrumental parameters may need such external prior information in order to not compromise the analysis results. We have demonstrated that in the studied cases with help of either suitable priors or data analysis techniques we can efficiently suppress the systematic biases and control the statistical uncertainties. In turn, these results can provide insights about the precision level for calibration of various instrumental parameters required given predefined science goals.

The cases studied here were clearly overidealized. This concerns both the models assumed for the foregrounds and the instruments. This is because our main purpose was to describe, validate, and demonstrate the proposed approach. However, more complex foreground models can be included following the procedures of [9], as can more realistic instrument models. As an example we have studied the case of the bandpasses where in addition we allowed for a mismatch between the underlying bandpass model and the one used for the analysis. We leave a more thorough and exhaustive exploration of other possible effects to future work.

The key feature of the proposed approach is the assumption that efficient, parametric models of the instrumental effects can be devised and then used to mitigate their impact. We emphasize that such models may be merely approximate and phenomenological, as in our bandpass example. The proposed forecasting tool allows us to evaluate efficiency (and sufficiency) of a model given more complex real instruments which could be characterized by nonparametric models derived from actual measurements.

Overall, the method we propose allows a more comprehensive approach of component separation, taking into account various instrumental systematic effects. We have demonstrated both the feasibility of the method, and its ability to reproduce results derived using other methods in the context of a typical Stage-3 CMB polarization experiment. The flexibility and reliability of the framework opens the possibility to apply it to various CMB experiments, in

particular those which will deploy continuously rotating HWP such as the LiteBIRD mission [22,34].

### ACKNOWLEDGMENTS

We acknowledge the use of HEALPY [27], PYSM [35] and FGBUSTER [36] software packages, as well as NUMPY [37] and SCIPY [38]. We thank Maximilian Abitbol, David Alonso, and Davide Poletti for useful comments and discussions. This work was supported by the French National Research Agency (ANR) Grants No. ANR-B3DCMB, (No. ANR-17-CE23-0002), and No. ANR-BxB (No. ANR-17-CE31-0022).

### APPENDIX A: HWP OPTICS

We give here the detailed expressions of  $\mathbf{C}_{0i;k}$  and  $\mathbf{S}_{0i;k}$  coefficients introduced in Eq. (15), as a function of HWP Mueller matrix elements  $\mu_{ij}$ . We note that  $\mu_{0i}$  and  $\mu_{i0}$  elements are zeros by design of the HWP, so we have:

$$\mathbf{C}_{01;0}(\nu) = \frac{1}{2}(\mu_{11} + \mu_{22}) \cos(2\eta_\nu) \quad (\text{A1})$$

$$\mathbf{C}_{01;4}(\nu) = \frac{1}{2}[(\mu_{11} - \mu_{22}) \cos(2\eta_\nu) - (\mu_{12} + \mu_{21}) \sin(2\eta_\nu)] \quad (\text{A2})$$

$$\mathbf{S}_{01;0}(\nu) = -\frac{1}{2}(\mu_{11} + \mu_{22}) \sin(2\eta_\nu) \quad (\text{A3})$$

$$\mathbf{S}_{01;4}(\nu) = \frac{1}{2}[(\mu_{12} + \mu_{21}) \cos(2\eta_\nu) + (\mu_{11} - \mu_{22}) \sin(2\eta_\nu)] \quad (\text{A4})$$

$$\mathbf{C}_{02;0}(\nu) = -\frac{1}{2}(\mu_{11} + \mu_{22}) \sin(2\eta_\nu) \quad (\text{A5})$$

$$\mathbf{C}_{02;4}(\nu) = \frac{1}{2}[(\mu_{12} + \mu_{21}) \cos(2\eta_\nu) + (\mu_{11} - \mu_{22}) \sin(2\eta_\nu)] \quad (\text{A6})$$

$$\mathbf{S}_{02;0}(\nu) = -\frac{1}{2}(\mu_{11} + \mu_{22}) \cos(2\eta_\nu) \quad (\text{A7})$$

$$\mathbf{S}_{02;4}(\nu) = -\frac{1}{2}[(\mu_{11} - \mu_{22}) \cos(2\eta_\nu) - (\mu_{12} + \mu_{21}) \sin(2\eta_\nu)] \quad (\text{A8})$$

We have explicitly included the frequency dependence of the coefficients, as Mueller matrix elements of the HWP [through  $\delta$  defined in Eq. (3)] and sinuous antenna depend on observing frequency. This allows us to naturally take into account the instrumental frequency-dependent effects in the model.

### APPENDIX B: COEFFICIENTS IN THE 3-LAYER HWP CASE

In Eqs. (A1)–(A8) in Appendix A, we expressed the coefficients of the full optics chain Mueller matrix, as a function of HWP Mueller matrix coefficients  $\mu_{ij}$ . These coefficients can be computed from Eq. (4) for any HWP configuration, or obtained from measurements.

We give here the full analytical expression of these coefficients in the instrumental configuration that we detailed in Sec. IVA 2. We recall that we have two parameters per HWP:

- (i)  $\alpha_2$ : central layer rotation angle;
- (ii)  $\theta_{\text{hwp}}$ : thickness of one layer.

For clarity reasons, we use  $\delta$  instead of  $\theta_{\text{hwp}}$ , defined as in Eq. (3):

$$\delta = \frac{2\pi\theta_{\text{hwp}}|n_o - n_e|\nu}{c}. \quad (\text{B1})$$

We can then express HWP Mueller matrix coefficients as:

$$\begin{aligned} \mu_{11} &= \cos^2(2\alpha_2) + \cos(\delta) \sin^2(2\alpha_2) \\ \mu_{12} &= \mu_{21} = \sin(2\alpha_2) \cos(2\alpha_2) [\cos^2(\delta) - \cos(\delta)] \\ &\quad - \sin^2(\delta) \sin(2\alpha_2) \\ \mu_{22} &= \cos^2(\delta) \sin^2(2\alpha_2) + \cos^3(\delta) \cos^2(2\alpha_2) \\ &\quad - [2 \sin^2(\delta) \cos(\delta) \cos(2\alpha_2) + \sin^2(\delta) \cos(\delta)]. \end{aligned} \quad (\text{B2})$$

### APPENDIX C: COSMOLOGICAL LIKELIHOOD

Our implementation of the cosmological likelihood is based on the one proposed in Appendix C of [8]. The cosmological likelihood Eq. (58) is split into three terms:

$$\langle S^{\text{cos}} \rangle = \text{tr} \mathbf{C}^{-1} \hat{\mathbf{C}} + \text{tr} \mathbf{C}^{-1} (\mathbf{E} - \hat{\mathbf{C}}) + \ln \det \mathbf{C}, \quad (\text{C1})$$

where  $\mathbf{E}$  is the true (observed) signal covariance matrix. Note that when there are no systematic residuals (our main case in this work, except when we consider bandpass variation), the second term of the sum actually vanishes. As outlined in Sec. III E, we consider two cases for the assumed signal covariance matrix: the no deprojection case, where  $\mathbf{C} = \mathbf{C}^{\text{cmb}}$ ; and the deprojection case, where  $\mathbf{C} = \mathbf{C}^{\text{cmb}} + \mathbf{C}^{\text{stat}}$ .

#### 1. No deprojection case

In the no deprojection case, the three terms of the cosmological likelihood are written as:

$$\text{tr} \mathbf{C}^{-1} \hat{\mathbf{C}} = \sum_{\ell} \left[ \frac{(2\ell + 1)}{C_{\ell}} (\hat{C}_{\ell} + \text{tr}[\Sigma \otimes_{\ell} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{Y}}^{(1)})]) \right] \quad (\text{C2})$$

$$\text{tr } \mathbf{C}^{-1}(\mathbf{E} - \hat{\mathbf{C}}) = \sum_{\ell} \left[ \frac{(2\ell + 1)}{C_{\ell}} (\otimes_{\ell} (\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) + \otimes_{\ell} (\tilde{\mathbf{z}}, \tilde{\mathbf{y}}) + \otimes_{\ell} (\tilde{\mathbf{y}}, \tilde{\mathbf{z}})) \right] \quad (\text{C3})$$

$$\ln \det \mathbf{C} = \ln \det \mathbf{C}^{\text{cmb}} \quad (\text{C4})$$

## 2. Deprojection case

In the deprojection case, we consider that we have a model for the statistical residuals, that we include in the modeled covariance matrix  $\mathbf{C}$ . The three terms of the cosmological likelihood then reads as (note that this case corresponds to previous implementation of the formalism, and thus the following equations are exactly equations (C9), (C10), and (C12) of [8]):

$$\begin{aligned} \text{tr } \mathbf{C}^{-1} \hat{\mathbf{C}} = & \sum_{\ell} \left[ (2\ell + 1) \frac{\hat{C}_{\ell}}{C_{\ell}} (1 - C_{\ell}^{-1} \text{tr}[\mathbf{U} \otimes_{\ell} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{Y}}^{(1)})]) + \frac{(2\ell + 1)}{C_{\ell}} \text{tr}[\Sigma \otimes_{\ell} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{Y}}^{(1)})] \right] \\ & - \sum_{\ell, \ell'} \frac{(2\ell + 1)(2\ell' + 1)}{C_{\ell} C_{\ell'}} \text{tr}[\mathbf{U} \otimes_{\ell'} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{Y}}^{(1)}) \Sigma \otimes_{\ell} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{Y}}^{(1)})] \end{aligned} \quad (\text{C5})$$

$$\begin{aligned} \text{tr } \mathbf{C}^{-1}(\mathbf{E} - \hat{\mathbf{C}}) = & \sum_{\ell} \left[ \frac{2\ell + 1}{C_{\ell}} (\otimes_{\ell} (\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) + \otimes_{\ell} (\tilde{\mathbf{z}}, \tilde{\mathbf{y}}) + \otimes_{\ell} (\tilde{\mathbf{y}}, \tilde{\mathbf{z}})) \right] \\ & - \sum_{\ell, \ell'} \frac{(2\ell + 1)(2\ell' + 1)}{C_{\ell} C_{\ell'}} \\ & \times \text{tr}[\mathbf{U} (\otimes_{\ell'} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{y}}) \otimes_{\ell} (\tilde{\mathbf{y}}, \tilde{\mathbf{Y}}^{(1)}) + \otimes_{\ell'} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{y}}) \otimes_{\ell} (\tilde{\mathbf{z}}, \tilde{\mathbf{Y}}^{(1)}) + \otimes_{\ell'} (\tilde{\mathbf{Y}}^{(1)}, \tilde{\mathbf{z}}) \otimes_{\ell} (\tilde{\mathbf{y}}, \tilde{\mathbf{Y}}^{(1)})] \end{aligned} \quad (\text{C6})$$

$$\ln \det \mathbf{C} = \ln \frac{\det \mathbf{C}^{\text{cmb}}}{\det \mathbf{U}}, \quad (\text{C7})$$

where we have introduced  $\mathbf{U}$  as defined in Eq. (C1) in [8]:

$$\mathbf{U} \equiv (\Sigma^{-1} + \tilde{\mathbf{Y}}^{(1)\dagger} \mathbf{C}^{\text{cmb}} \tilde{\mathbf{Y}}^{(1)})^{-1}. \quad (\text{C8})$$

- 
- [1] S. M. Leach, J. F. Cardoso, C. Baccigalupi, R. B. Barreiro, M. Betoule, J. Bobin, A. Bonaldi, J. Delabrouille, G. de Zotti, C. Dickinson *et al.*, *Astron. Astrophys.* **491**, 597 (2008).
- [2] H. K. Eriksen, A. J. Banday, K. M. Górski, and P. B. Lilje, *Astrophys. J.* **612**, 633 (2004).
- [3] J.-F. Cardoso, M. Le Jeune, J. Delabrouille, M. Betoule, and G. Patanchon, *IEEE J. Sel. Top. Signal Process.* **2**, 735 (2008).
- [4] K. Ichiki, H. Kanai, N. Katayama, and E. Komatsu, *Prog. Theor. Exp. Phys.* **2019**, 033E01 (2019).
- [5] H. K. Eriksen, J. B. Jewell, C. Dickinson, A. J. Banday, K. M. Górski, and C. R. Lawrence, *Astrophys. J.* **676**, 10 (2008).
- [6] F. Stivoli, J. Grain, S. M. Leach, M. Tristram, C. Baccigalupi, and R. Stomp, *Mon. Not. R. Astron. Soc.* **408**, 2319 (2010).
- [7] R. Stomp, S. Leach, F. Stivoli, and C. Baccigalupi, *Mon. Not. R. Astron. Soc.* **392**, 216 (2009).
- [8] R. Stomp, J. Errard, and D. Poletti, *Phys. Rev. D* **94**, 083526 (2016).
- [9] J. Errard, S. M. Feeney, H. V. Peiris, and A. H. Jaffe, *J. Cosmol. Astropart. Phys.* **03** (2016) 052.
- [10] M. Salatino, J. Lashner, M. Gerbino, S. M. Simon, J. Didier, A. Ali, P. C. Ashton, S. Bryan, Y. Chinone, K. Coughlin *et al.*, in *Proceedings of SPIE* (2018), Vol. 10708 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 1070848 [arXiv:1808.07442].
- [11] K. Komatsu, T. Matsumura, H. Imada, H. Ishino, N. Katayama, and Y. Sakurai, *J. Astron. Telesc. Instrum. Syst.* **5**, 044008 (2019).
- [12] S. A. Bryan, T. E. Montroy, and J. E. Ruhl, *Appl. Opt.* **49**, 6313 (2010).
- [13] S. Pancharatnam, *Proc. Indian Acad. Sci. A* **41**, 130 (1955).

- [14] A. Suzuki, Ph.D. thesis, University of California, Berkeley, 2013.
- [15] I. L. Padilla, J. R. Eimer, Y. Li, G. E. Addison, A. Ali, J. W. Appel, C. L. Bennett, R. Bustos, M. K. Brewer, M. Chan *et al.*, *Astrophys. J.* **889**, 105 (2020).
- [16] T. A. Enßlin, S. Hutschenreuter, V. Vacca, and N. Oppermann, *Phys. Rev. D* **96**, 043021 (2017).
- [17] J. Errard, F. Stivoli, and R. Stompor, *Phys. Rev. D* **84**, 069907 (2011).
- [18] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday *et al.*, *Astron. Astrophys.* **571**, A15 (2014).
- [19] The Simons Observatory Collaboration, *J. Cosmol. Astropart. Phys.* **02** (2019) 056.
- [20] SO\_Noise\_Calculator\_Public.py available at <https://simonsobservatory.org/publications.php>.
- [21] C. A. Hill, S. Beckman, Y. Chinone, N. Goeckner-Wald, M. Hazumi, B. Keating, A. Kusaka, A. T. Lee, F. Matsuda, R. Plambeck *et al.*, *Design and Development of an Ambient-Temperature Continuously-Rotating Achromatic Half-Wave Plate for CMB Polarization Modulation on the POLAR-BEAR-2 Experiment*, Vol. 9914 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (SPIE, 2016), p. 99142U, <https://doi.org/10.1117/12.2232280>.
- [22] Y. Sakurai, T. Matsumura, N. Katayama, K. Komatsu, R. Takaku, S. Sugiyama, Y. Nomura, T. Toda, T. Ghigna, T. Iida *et al.*, in *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy X*, edited by J. Zmuidzinas and J.-R. Gao, International Society for Optics and Photonics (SPIE, 2020), Vol. 11453, pp. 743–762, <https://doi.org/10.1117/12.2560289>.
- [23] T. Matsumura, S. Hanany, P. Ade, B. R. Johnson, T. J. Jones, P. Jonnalagadda, and G. Savini, *Appl. Opt.* **48**, 3614 (2009).
- [24] J. T. Ward, D. Alonso, J. Errard, M. J. Devlin, and M. Hasselfield, *Astrophys. J.* **861**, 82 (2018).
- [25] B. Thorne, D. Alonso, S. Naess, and J. Dunkley, <http://ascl.net/1704.007>.
- [26] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, *Astrophys. J.* **622**, 759 (2005).
- [27] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. Gorski, *J. Open Source Softw.* **4**, 1298 (2019).
- [28] <https://healpix.sourceforge.io/>.
- [29] A. Kusaka, T. Essinger-Hileman, J. W. Appel, P. Gallardo, K. D. Irwin, N. Jarosik, M. R. Nolta, L. A. Page, L. P. Parker, S. Raghunathan *et al.*, *Rev. Sci. Instrum.* **85**, 024501 (2014).
- [30] S. Takakura, M. Aguilar, Y. Akiba, K. Arnold, C. Baccigalupi, D. Barron, S. Beckman, D. Boettger, J. Borrill, S. Chapman *et al.*, *J. Cosmol. Astropart. Phys.* **05** (2017) 008.
- [31] F. Matsuda, L. Lowry, A. Suzuki, M. Aguilar Fáunderz, K. Arnold, D. Barron, F. Bianchini, K. Cheung, Y. Chinone, T. Elleflot *et al.*, *Rev. Sci. Instrum.* **90**, 115115 (2019).
- [32] T. Ghigna, T. Matsumura, G. Patanchon, H. Ishino, and M. Hazumi, *J. Cosmol. Astropart. Phys.* **11** (2020) 030.
- [33] M. H. Abitbol, D. Alonso, S. M. Simon, J. Lashner, K. T. Crowley, A. M. Ali, S. Azzoni, C. Baccigalupi, D. Barron, M. L. Brown *et al.*, [arXiv:2011.02449](https://arxiv.org/abs/2011.02449).
- [34] H. Sugai, P. A. R. Ade, Y. Akiba, D. Alonso, K. Arnold, J. Aumont, J. Ausermann, C. Baccigalupi, A. J. Banday, R. Banerji *et al.*, *J. Low Temp. Phys.* **199**, 1107 (2020).
- [35] <https://pysm3.readthedocs.io/en/latest/#>.
- [36] <https://fgbuster.github.io/fgbuster/index.html>.
- [37] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith *et al.*, *Nature (London)* **585**, 357 (2020).
- [38] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, *Nat. Methods* **17**, 261 (2020).