Least response method to separate CMB spectral distortions from foregrounds

J.-P. Maillard,[1] A. Mihalchenko,[2,3] D. Novikov,[2] A. Osipova,[2,4] S. Pilipenko,[2] and J. Silk[1,5,6]

[1]*Institut d'Astrophysique de Paris, UMR 7095, CNRS, Sorbonne Université,
98bis Boulevard Arago, 75014 Paris, France*
[2]*Astro-Space Center of P.N. Lebedev Physical Institute, Profsoyusnaya 84/32, Moscow 117997, Russia*
[3]*Moscow Institute of Physics and Technology, Institutskiy pereulok,
d.9, Dolgoprudny, Moscow 141701, Russia*
[4]*National Research University Higher School of Economics,
11 Pokrovsky Bulvar, Moscow 109028, Russia*
[5]*Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, Maryland 21218, USA*
[6]*Beecroft Institute for Particle Astrophysics and Cosmology, Department of Physics, University of Oxford,
Oxford OX1 3RH, United Kingdom*

We present a signal-foreground separation algorithm for filtering observational data to extract spectral distortions of the cosmic microwave background (CMB). Our linear method, called the least response method (LRM), is based on the idea of simultaneously minimizing the response to all possible foregrounds with poorly defined spectral shapes and random noise while maintaining a constant response to the signal of interest. This idea was introduced in detail in our previous paper. Here, we have expanded our analysis by taking into consideration all the main foregrounds. We draw a detailed comparison between our approach and the moment internal linear combination method, which is a modification of the internal linear combination technique previously used for CMB anisotropy maps. We demonstrate advantages of LRM and evaluate the prospects for measuring various types of spectral distortions. Besides, we show that LRM suggests the possibility of its improvements if we use an iterative approach with sequential separation and partial subtraction of foreground components from the observed signal. In addition, we estimate the optimal temperature that the telescope's optical system should have in order to detect the chemical type μ distortions. We present a design of an instrument where, according to our estimates, the optimal contrast between its thermal emission and the CMB allows us to measure such distortions.

I. INTRODUCTION

The recent space missions related to the study of relic radiation WMAP [1] and Planck [2,3] has substantially broadened our knowledge of cosmic microwave background anisotropy; the power spectrum of $\Delta T/T$ fluctuations and almost Gaussian nature of their distribution on the celestial sphere. Ultimately, these observational data helped to estimate the main cosmological parameters.

However, even more information can be gained by delving into cosmic microwave background (CMB) spectral distortions (SDs). Being one of the key goals of observational cosmology [4–10], measuring deviations of the CMB spectrum from a black body shape shall reveal a vast amount of information about the early Universe, unobtainable by other observational methods [4,11–14].

The chemical potential μ-type distortions [15] are created if either energy is injected into the cosmic plasma or photon number density is modified [6,16–21] when the redshift is less than $\sim 2 \times 10^6$. Thus, they can be used as a tool to trace the energy history of the Universe. A later type of distortions (y distortions) [11] contain information about the structure of intracluster medium [22–25]. In addition, very specific deviations in the frequency spectrum of radiation coming from galaxy clusters can be used for independent measurements of the amplitudes and orientations of low CMB anisotropy multipoles [26,27]. The COBE/FIRAS mission [28,29] measured the CMB frequency spectrum and established its Planckian shape with good accuracy. Future detection of the spectral features is expected from FIRAS-type missions [30–33], experiments with a large primary mirror with good angular resolution [34] (which is essential for y-type deviations) or even using a Moon-based telescope [35], which is revised in this paper.

An exhaustive review of the theory behind SDs, their observational challenges, computational techniques and possible new directions in this field are given in [16]. The main difficulty in the task of measuring small spectral

distortions is the presence of foregrounds of cosmic and instrumental origin, which are several orders of magnitude larger in amplitude than the signals of interest. Thus, in order to ensure detection of such signals, it is necessary not only to achieve high sensitivity of the experiment, but also to learn the most optimal way to separate SDs from the rest of the observed signal. The spectra of some astronomical foregrounds as well as the spectrum emitted by the optical system of the instrument are poorly defined and cannot be modeled and predicted with the accuracy required for reliable SDs measurements. This means that sufficiently advanced and efficient data processing methods must be used to solve such a problem. In particular, any methods used must take into account possible variations in foreground frequency spectra.

The well-known blind internal linear combination (ILC) method [36] has been successfully applied to process multifrequency analysis of CMB anisotropy maps [37–39]. However, the application of this method is limited by the presence of large foreground components that have a nonzero projection onto the signal of interest, which leads to a bias. To avoid this, the constrained ILC (cILC) [40,41] method was proposed as a modification of the ILC. This approach completely eliminates the contribution of foregrounds with well-known spectral shapes, treating the remaining components as unmodeled noise. However, to solve the problem of getting rid of all foreground components, it is necessary to take into account the spectral variations of dust, cosmic infrared background (CIB), synchrotron radiation and other sky foregrounds along the line of site and from one direction to another one, as well as time variations of the foreground contribution from the instrument emission. The spectra of these components depend on the parameters and, therefore, spectral variations are equivalent to variations in the parameters.

To take this into account, a quite effective method, called MILC (Moments ILC) [42–44], was introduced, where the foreground spectra are expanded into Taylor series in terms of parameters in the vicinity of some average reference parameter values. Data filtering, which involves zeroing out the expansion moments, ensures that the contribution from such foregrounds is eliminated. However, this method also has a rather serious disadvantage. A large number of constraints ensuring the zeroing of the foreground contribution during the filtering process inevitably leads to an unacceptably large response to unmodeled noise (including random photon noise). This is not surprising, since a large number of imposed strict conditions limits the number of degrees of freedom and increases the contribution of noise to the estimation of the signal of interest.

In work [45], another modification of the cILC method, called partially constrained ILC (pcILC), was proposed. Instead of completely eliminating the contribution from a foreground with a well-defined spectrum (as cILC does), pcILC reduces this contribution to some empirically

determinable level. As a result the strict cILC constraints are somewhat softened. This made it possible to reduce the response to noise when processing $\Delta T/T$ anisotropy data in the presence of the Sunyaev-Zel'dovich effect or CMB lensing. However, application of this approach in the presence of many different foregrounds is very complicated and its effectiveness in this case has not been proven. Besides, unlike MILC, this method does not allow for possible spectral shape variations of foregrounds.

We recently proposed a data filtering method called least response method (LRM) [46], which is quite simple and easy to implement. This approach involves the optimization of a single functional for all components of the observed signal. During the process of data filtering we minimize the response to all foregrounds and random noise simultaneously, while keeping the response to the signal of interest equal to unity. We assume that we know the following information about foreground signals:

(1) The foreground spectral parameters can change within a limited range of their possible variations.
(2) The amplitudes of the foregrounds are limited from above by known values.

Note that such information is available to us from previous observational data and characteristics of the instrument optical system for a particular experiment. In this article, we compare LRM and MILC approaches and show the advantages and prospects of LRM for measuring various spectral distortions of the relic radiation in the presence of all main foregrounds, including the signal from the instrument optics.

In addition, we show that when measuring μ distortions, the temperature of the instrument optical system should not approach the CMB temperature. Measuring such distortions requires calibration of the instrument and in this case the signal component generated by the optics becomes a part of the observed signal. This component is close in shape to a black body signal, and if its temperature is close to 2.7 K, then it becomes "poisoned" to some extent, since the device itself begins to create spectral features simulating CMB distortions. We show that, according to our estimates, the optimal temperature of the instrument for μ spectral distortion measurements should be close to 9 K.

The outline of this paper is as follows. In Sec. II, we review linear methods for observational data filtering. We describe MILC and LRM approaches for CMB spectral distortions detection in the case when the observed signal contains foregrounds with poorly defined spectral shapes. In Sec. III, we describe models of signals of interest, foregrounds and photon noise and present a detailed comparative numerical analysis of MILC and LRM methods. In this section we also estimate the instrument optics temperature that is optimal for detecting μ-type CMB spectral distortions. In the same section we give an example of an instrument for measuring the relic radiation spectral features. Brief conclusions and our suggestions about

possible improvements of the LRM approach are given in Sec. IV.

II. LINEAR DATA FILTERING METHODS

The frequency spectrum $S(\nu)$ we observe consists of the signal of interest and a set of other components that we would like to get rid of during data processing. Therefore, the total observed spectrum can be written as follows:

$$S(\nu) = a_d I_d(\nu) + \sum_{m=1}^{M} I_m(\nu), \qquad (1)$$

where I_d is the certain kind of CMB spectral distortions, which we would like to separate from other components. Index "d" can denote the μ distortions: I_μ, the Sunyaev-Zel'dovich effect (y distortions); I_{y_0}, the first or second relativistic corrections to this effect: I_{y_1} or I_{y_2}. Therefore, a_d is the frequency-independent amplitude of the signal of interest we want to estimate. The rest of the signal $S(\nu)$ consists of M foregrounds of various physical origins $I_m(\nu)$, which may include spectral distortions themselves. For example, if we are interested in signal I_μ, then signals $I_{y_0}, I_{y_1} I_{y_2}$ will be part of the total foreground.

Using a Fourier-transform spectrometer (FTS) mounted on a space telescope, one can obtain discrete values of the total signal (or row vector) $\mathbf{S} = (S_1, ..., S_J)$ in J equally wide frequency channels in a broad frequency range from $\nu_{\min}$ to $\nu_{\max}$:

$$S_j = a_d I_d^j + \sum_m I_m^j + N_j, \quad j = 1, ..., J,$$

$$I_x^j = \int_{\nu_j - \frac{\Delta\nu}{2}}^{\nu_j + \frac{\Delta\nu}{2}} I_x(\nu) \frac{d\nu}{\Delta\nu}, \qquad (2)$$

where indices j indicate the frequency channel number, $\Delta\nu$ is the channel width and N_j is the random photon noise with zero mean and covariance matrix $[C_{ij}] = \mathbf{C} = \langle \mathbf{N^T N} \rangle$. The spectral shapes of CMB distortions I_μ, I_y, I_{y1} and I_{y2} are well-known, while the foreground spectra I_m can depend on various parameters and in fact are superpositions of spectra integrated along the line of sight or/and obtained as a result of averaging over spatial pixels of a sky map. Therefore, the foreground components can be written as

$$I_m^j = \int_\Omega a_m(\mathbf{P}) f_m(\nu_j, \mathbf{P}) d\mathbf{P},$$

$$d\mathbf{P} = dp_1 dp_2 \cdots dp_L, \qquad (3)$$

where $\mathbf{P} = p_1, ..., p_L$ is the set of L parameters, $f_m(\nu_j, \mathbf{P})$ are the functions representing the foreground spectra, Ω is the region of possible parameter variations, and a_m are the amplitudes of the foreground radiation as functions of parameters $\mathbf{P}$. Thus, if, for example, $a_m(\mathbf{P})$ has the form of a delta function $a_m(\mathbf{P}) = A_m \cdot \delta(\mathbf{P} - \mathbf{P_m})$, then the foreground spectrum with index m will have a template with well-defined parameters $\mathbf{P_m}$ and the amplitude A_m: $I_m^j = A_m \cdot f_m(\nu, \mathbf{P_m})$.

The total observed signal $\mathbf{S} = (S_1, ..., S_J)$ can be divided into three parts (three vectors):

$$\mathbf{S} = a_d \mathbf{I_d} + \mathbf{F} + \mathbf{N},$$

$$\mathbf{F} = (F_1, ..., F_J), \quad F_j = \sum_m I_m^j,$$

$$\mathbf{N} = (N_1, ..., N_J), \qquad (4)$$

where $\mathbf{I_d}$ is some particular kind of spectral distortion we want to separate from the rest of the signal, $\mathbf{F}$ is the total foreground and $\mathbf{N}$ represents the random noise.

The task of any linear algorithm is to find the optimal vector of weights $\boldsymbol{\omega} = (\omega_1, ..., \omega_J)$ for frequency channels that should have the following property:

$$\boldsymbol{\omega}\mathbf{S^T} = \sum_{j=1}^{J} \omega_j S_j \to a_d \quad for \; C_{ij} \to 0, \quad i, j = 1, ..., J. \quad (5)$$

Thus, in the ideal case, when the noise tends to zero ($C_{ij} \to 0$), it is desirable that the algorithm accurately reproduces the amplitude of the signal of interest a_d without any bias. With nonzero noise, the algorithm should minimize the deviation of the estimate of this amplitude from its true value.

Let us denote the scalar product $\boldsymbol{\omega}\mathbf{S^T} = R(\mathbf{S})$ as the response to the signal:

$$R(\mathbf{S}) = a_d R(\mathbf{I_d}) + R(\mathbf{F}) + R(\mathbf{N}). \qquad (6)$$

The first condition imposed on the weights is common to all linear algorithms and is quite obvious,

$$R(\mathbf{I_d}) = \boldsymbol{\omega}\mathbf{I_d^T} = \sum_j \omega_j I_d^j = 1. \qquad (7)$$

Thus, it is necessary to minimize the response to the rest of the signal $R(\mathbf{F}) + R(\mathbf{N})$ while maintaining the condition in Eq. (7). To denote the expected total average value of the response to foreground and noise, we use the following definition:

$$R(\mathbf{F} + \mathbf{N}) := \sqrt{R^2(\mathbf{F}) + \langle R^2(\mathbf{N}) \rangle}. \qquad (8)$$

Therefore, $R(\mathbf{F} + \mathbf{N})$ is the value that must be minimized under condition in Eq. (7).

A. Internal linear combination method

The ILC method assumes that we know the spectral energy distribution (SED) of the signal of interest, but does not assume knowledge of any information about foregrounds and noise. With this approach all components of the observed signal, apart from the signal of interest, are considered as unmodeled noise. That is, in this case in Eqs. (4) and (8) no distinction is made between the components $\mathbf{F}$ and $\mathbf{N}$. Thus, the problem is reduced to minimizing the response to such noise while maintaining a constant response to $\mathbf{I_d}$. As a result, the weights ω_j are the solution of the following system of equations:

$$
\begin{aligned}
&1.\ \boldsymbol{\omega}\mathbf{I_d^T} = 1,\\
&2.\ \partial(\boldsymbol{\omega}\mathbf{D}\boldsymbol{\omega}^T)/\partial\boldsymbol{\omega} = 0,
\end{aligned}
\tag{9}
$$

where $\mathbf{D} = \langle\mathbf{S^TS}\rangle - \langle\mathbf{S^T}\rangle\langle\mathbf{S}\rangle$ is the data covariance matrix and $\langle\rangle$ means averaging over the spatial pixels of the sky map. The solution of such a system is

$$
\boldsymbol{\omega} = \mathbf{I_d}\mathbf{D^{-1}} \cdot (\mathbf{I_d}\mathbf{D^{-1}}\mathbf{I_d^T})^{-1}.
\tag{10}
$$

This completely blind approach can be biased due to nonzero projections of foregrounds on the considered signal, $\mathbf{FI_d^T} \neq 0$. Taking into account that in our case the foreground amplitudes can exceed the amplitude of the signal of interest by several orders of magnitude, this method is not suitable.

B. Moments approach (MILC)

In order to avoid biasing in the process of signal $\mathbf{I_d}$ extraction, the constrained ILC (or cILC) method was proposed [40,41], which nullifies the response to some modeled foregrounds with known spectra. Thus, a number of constraints are added to Eq. (9) that ensure a zero response to certain foregrounds. This modification of the blind ILC approach was extended with the MILC method [42–44], where foregrounds with badly defined spectra in Eq. (3) were considered. The idea behind the MILC method is quite simple and effective. Since the spectral shapes of some signal components depend on the parameters $\mathbf{P}$, they can be expanded in a Taylor series up to some order n in the vicinity of a certain reference point corresponding to the average preestimated value $\mathbf{P_0}$:

$$
\begin{aligned}
f_m(\nu_j, \mathbf{P}) &\approx f_m(\nu_j, \mathbf{P_0}) + \sum_{n_1,\,\ldots,\,n_L}\\
&\times \left[\frac{\partial^{n_1+\cdots+n_L}}{\partial P_1^{n_1}\cdots P_L^{n_L}}f_m(\nu_j, \mathbf{P})\right]_{\mathbf{P}=\mathbf{P_0}}\\
&\times \Delta P_1^{n_1}\cdots \Delta P_L^{n_L},
\end{aligned}
\tag{11}
$$

where the summation is performed over all positive $n_1, \ldots, n_L$ satisfying the condition $0 \leq n_1 + \cdots + n_L \leq n$. In order to remove the influence of such foregrounds on data processing, it is sufficient to require a zero response to all derivatives of the spectra with respect to parameters up to the nth order. As a result, we get the following system of equations:

$$
\begin{aligned}
&1.\ \boldsymbol{\omega}\mathbf{I_d^T} = 1,\\
&2.\ \sum_j \omega_j \frac{\partial^{n_1+\ldots+n_L}}{\partial P_1^{n_1}\ldots P_L^{n_L}}f_m(\nu_j, \mathbf{P})|_{\mathbf{P}=\mathbf{P_0}} = 0,\\
&\quad m = 1, \ldots, M, \quad 0 \leq n_1 + \ldots + n_L \leq n,\\
&3.\ \partial(\boldsymbol{\omega}\mathbf{C}\boldsymbol{\omega}^T)/\partial\boldsymbol{\omega} = 0.
\end{aligned}
\tag{12}
$$

Therefore, this method ensures that the response to all foregrounds is zeroed to within

$$
R(\mathbf{F}) = O\left[\sum_j \omega_j \frac{\partial^{n+1}f_m(\nu_j, \mathbf{P})}{\partial\mathbf{P}^{n+1}}\right]_{\mathbf{P}=\mathbf{P_0}} = R_{n+1}(\mathbf{F}).
\tag{13}
$$

However, the disadvantage of MILC is that with a large number of constraints in formula 2 of Eq. (12), the response to noise $R(\mathbf{N})$ can become unacceptably large. Indeed, this formula is equivalent to the system of equations $\omega\varphi_k^T = 0$, $k = 1, \ldots K$, where the vectors φ_k correspond to the derivatives of the foregrounds with respect to the parameters $\mathbf{P}$ and the capital K denotes the total number of constraints (or derivatives). The system of vectors φ_k can be orthogonalized in such a way that the new vectors $\tilde{\varphi}_\mathbf{k}$ are a linear combinations of the original vectors φ_k and represent an orthonormal system:

$$
\begin{aligned}
&\tilde{\varphi}_\mathbf{k} = \tilde{\varphi}_\mathbf{k}(\varphi_\mathbf{1}, \ldots, \varphi_\mathbf{K}),\\
&\tilde{\varphi}_\mathbf{k}\tilde{\varphi}_\mathbf{k'}^\mathbf{T} = \delta_k^{k'}, \quad k, k' = 1, \ldots, K.
\end{aligned}
\tag{14}
$$

Thus, the normalized signal $\tilde{\mathbf{I}}_\mathbf{d} = \mathbf{I_d} \cdot (\mathbf{I_d}\mathbf{I_d^T})^{-\frac{1}{2}}$ can be written in the following form:

$$
\begin{aligned}
&\tilde{\mathbf{I}}_\mathbf{d} = \sum_{k=1}^{K}\gamma_k\tilde{\varphi}_\mathbf{k} + \boldsymbol{\Delta}_d,\\
&\gamma_k = \tilde{\varphi}_k \cdot \tilde{\mathbf{I}}_\mathbf{d}^\mathbf{T},
\end{aligned}
\tag{15}
$$

where $\boldsymbol{\Delta}_d$ is part of signal $\tilde{\mathbf{I}}_\mathbf{d}$ orthogonal to all modeled foregrounds: $\tilde{\varphi}_\mathbf{k}\boldsymbol{\Delta}_d^T = 0$, $k = 1, \ldots, K$. That is, the smaller the length of the vector $\boldsymbol{\Delta}_\mathbf{d}$, the more difficult it is to separate the vector $\mathbf{I_d}$ from the rest of the total signal. We can define the value of $\Gamma = (\boldsymbol{\Delta}_\mathbf{d}\boldsymbol{\Delta}_\mathbf{d}^T)^{1/2}$ as a measure of the orthogonality of the normalized signal $\tilde{\mathbf{I}}_\mathbf{d}$ to all foregrounds. It is easy to show, that $\Gamma^2 = 1 - \sum_{k=1}^{K}\gamma_k^2$. Finally, the solution of the system of Eq. (12) for weights $\boldsymbol{\omega}$ is

$$\boldsymbol{\omega} = \boldsymbol{\Delta_d}\mathbf{C}^{-1} \cdot (\boldsymbol{\Delta_d}\mathbf{C}^{-1}\mathbf{I_d^T})^{-1}, \qquad (16)$$

and the noise response can be estimated as follows:

$$\langle R^2(\mathbf{N}) \rangle = \boldsymbol{\omega}\mathbf{C}\boldsymbol{\omega}^T \sim \sigma^2/\Gamma^2,$$
$$\sigma^2 = \frac{\langle \mathbf{N}\mathbf{N^T} \rangle}{J}. \qquad (17)$$

Here, σ is equivalent to the mean sensitivity per frequency channel. Thus, a small value of Γ means a large response to random noise. In case when $\Gamma \to 0$, the signal $\mathbf{I_d}$ becomes a linear combination of foregrounds, and the separation of such a signal from them with this method becomes impossible.

C. Least response method

Recently, a new approach to the component separation problem has been proposed [46], which can be called the LRM. Unlike MILC, this method implies the availability of information not only about the possible foreground spectral shapes variations, but also about their maximum possible amplitudes. It is important to note that such information about main known foregrounds is available to us. In addition, overestimation of the upper limit of the foreground amplitudes is not critical for our approach (unlike underestimation). This helps to avoid imposing strict conditions on the weight vector $\boldsymbol{\omega}$ and thereby significantly reduce the noise response. The only fairly mild assumption about the foregrounds described in Eq. (3) is that the amplitudes a_m inside the parameter domain Ω should be less than certain (preestimated) values A_m:

$$|a_m(\mathbf{P})| \le A_m \quad for\ \mathbf{P} \in \Omega, \qquad (18)$$

and $a_m(\mathbf{P}) = 0$ otherwise.

It is easy to show, that the mean square of the response to foreground has an upper limit:

$$\langle R^2(\mathbf{F}) \rangle \le \left\langle \sum_{m=1}^{M} M \cdot a_m^2(\mathbf{P}) \left[\sum_{j=1}^{J} f_m(\nu_j, \mathbf{P}) \cdot \omega_j \right]^2 \right\rangle, \quad (19)$$

and according to Eq. (18) the following inequality is always true:

$$\langle R^2(\mathbf{F}) \rangle \le \boldsymbol{\omega}\boldsymbol{\Phi}\boldsymbol{\omega}^{\mathbf{T}}, \quad \boldsymbol{\Phi} = M\left[\sum_{m=1}^{M} A_m^2 q_{ij}^m \right],$$
$$q_{ij}^m = \frac{1}{V_\Omega} \int_\Omega f_m(\nu_i, \mathbf{P}) f_m(\nu_j, \mathbf{P}) d\mathbf{P}, \qquad (20)$$

where integrals q_{ij}^m can be precalculated for all types of foreground ($m = 1, ..., M$) numerically or in some particular cases analytically depending on the configuration of Ω.

It is worth noting that the prefactor M for the matrix $\boldsymbol{\Phi}$ in Eqs. (19) and (20) is used to ensure that these inequalities are certainly correct. In general, the prefactor for this matrix can vary from 1 (when there are no correlations between foregrounds) to M (for 100% correlation between all foregrounds). Thus, using the value M as a prefactor provides a complete guarantee of taking into account all possible correlations. In reality, foregrounds of different physical origins are weakly correlated with each other. In this case, the coefficient M can be replaced by one. In our estimates we use uncorrelated foregrounds model and, therefore, $\boldsymbol{\Phi} = [\sum_{m=1}^{M} A_m^2 q_{ij}^m]$. For a more detailed analysis, possible correlations between individual foregrounds can be taken into account in the calculation of the $\boldsymbol{\Phi}$ matrix.

Since $\langle R^2(\mathbf{N}) \rangle = \boldsymbol{\omega}\mathbf{C}\boldsymbol{\omega}^{\mathbf{T}} = \sum_{i,j} C_{ij}\omega_i\omega_j$ and the foregrounds are not correlated with noise we have

$$\langle (R(\mathbf{F}) + R(\mathbf{N}))^2 \rangle \le \boldsymbol{\omega}[\boldsymbol{\Phi} + \mathbf{C}]\boldsymbol{\omega}^{\mathbf{T}}. \qquad (21)$$

Therefore, the minimization of the response to the foreground and to the noise is achieved with weights ω_j corresponding to the minimum of the quadratic form $\boldsymbol{\omega}[\boldsymbol{\Phi} + \mathbf{C}]\boldsymbol{\omega}^{\mathbf{T}}$ under condition in Eq. (7):

$$1.\ \boldsymbol{\omega}\mathbf{I_d^T} = 1,$$
$$2.\ \partial(\boldsymbol{\omega}[\boldsymbol{\Phi} + \mathbf{C}]\boldsymbol{\omega}^{\mathbf{T}})/\partial\boldsymbol{\omega} = 0. \qquad (22)$$

The solution of the system Eq. (22) is

$$\boldsymbol{\omega} = \mathbf{I_d}[\boldsymbol{\Phi} + \mathbf{C}]^{-1} \cdot (\mathbf{I_d}[\boldsymbol{\Phi} + \mathbf{C}]^{-1}\mathbf{I_d}^T)^{-1}. \qquad (23)$$

Thus, LRM, unlike MILC, does not require complete orthogonality of the signal of interest to all foregrounds [compare Eqs. (12) and (22)]. In the next section we will demonstrate the advantages of this approach and show that it allows to detect a signal in the observational data at a much lower sensitivity.

III. METHODS COMPARISON AND PROSPECTS FOR MEASURING SPECTRAL DISTORTIONS

A. Modeling spectral distortions and foregrounds

Our numerical calculations were performed for a simulated observed frequency spectrum that included the signal of interest, all possible foregrounds with varying parameters and photon noise. We did not use any information about the spatial distribution of foreground sources in the sky. In order to carry out a numerical experiment on the application of the two methods (MILC and LRM) described in the previous section, we used the following models of signal components.

Signals of CMB origin are μ distortion I_μ, y distortion (SZ effect) I_{y_0}, first and second relativistic corrections to thermal SZ effect I_{y_1}, I_{y_2} and CMB anisotropy I_{CMBA}.

(It is assumed that the CMB monopole can be easily removed from the observational data). All these spectra have well-defined shapes that do not depend on any parameters and look as follows:

$$I_\mu(\nu) = I_0 \frac{x^4 e^x}{(e^x - 1)^2}\left(\frac{1}{b} - \frac{1}{x}\right)\mu,$$

$$I_{y_0}(\nu) = I_0 \frac{x^4 e^x}{(e^x - 1)^2}\left(x \coth\left(\frac{x}{2}\right) - 4\right)y,$$

$$I_{y_1}(\nu) = I_0 \frac{x^4 e^x}{(e^x - 1)^2}Y_1(x)\theta_e y,$$

$$I_{y_2}(\nu) = I_0 \frac{x^4 e^x}{(e^x - 1)^2}Y_2(x)\theta_e^2 y,$$

$$I_{\mathrm{CMBA}}(\nu) = \frac{2(kT_0)^3}{(hc)^2}\frac{x^4}{(e^x - 1)^2}\frac{\Delta T}{T_0}, \qquad (24)$$

where $x = h\nu/k_T \mathrm{CMB}$ and the CMB temperature is $T_{\mathrm{CMB}} = 2.72548\ K$, [28,29] and $\Delta T/T_{\mathrm{CMB}} < 10^{-4}$. The same estimated values for constants b, I_0, μ and y as in [32] are used: $I_0 = 270\ \mathrm{M\,Jy/sr}$, $b = 2.1923$, $\mu = 2 \times 10^{-8}$, $y = 1.77 \times 10^{-6}$ and $\theta_e = kT_{\mathrm{SZ}}/m_e c^2 \sim 2.44 \times 10^{-3}$. The functions $Y_1(x)$ and $Y_2(x)$ for the first and second corrections have a rather cumbersome form and analytical formulas for them can be found in [47].

Dust and CIB foregrounds are considered together and modeled as a modified blackbody radiation with two floating parameters: temperature T and spectral index β,

$$I_{\mathrm{Dust,CIB}}(\nu, T, \beta) = \tau_{DC}(\nu/\nu_{DC})^\beta B(\nu, T),$$

$$B(\nu, T) = \frac{2(kT)^3}{(hc)^2}\frac{x^3}{e^x - 1}, \qquad (25)$$

where $\nu_{DC} = 353$ GHz. The boundaries of the parameters (T, β) domain were determined in [46] using Planck data [48,49]. The probability distribution function for these parameters was calculated for 10 degrees circular sky part centered at $l = 13.731^o$, $b = -73.946^o$, see Fig. 1(a). Two isocontour solid lines limit the region $\Omega(T, \beta)$ of possible parameters variations for both; dust and CIB. The probability to find parameters outside these two spots is less than 0.0002. The maximum allowable value of emissivity τ_{DC} for the data we used does not exceed 10^{-6}.

Synchrotron radiation is modeled according to [50] and its spectrum has a power-law form with the single free parameter β_s:

$$I_{\mathrm{sync}}(\nu, \beta_s) = A_s(\nu/\nu_s)^{-\beta_s}, \qquad (26)$$

where $\nu_s = 30$ GHz and $A_s < 1000$ Jy/sr. In accordance with the results of [51], β_s can vary from 0.9 to 1.4.

Free-free emission is given by the following formula in [52]:

$$I_{ff}(\nu) = A_{ff}\left(1 + \ln\left[1 + \left(\frac{\nu_{ff}}{\nu}\right)^{\sqrt{3}/\pi}\right]\right), \qquad (27)$$

where, $A_{ff} < 500$ Jy/sr, $\nu_{ff} = 255.33(\frac{T_e}{1000K})^{3/2}$ GHz. According to [52], the parameter $T_e = 7000K \pm 3K$ does not vary across the sky strongly enough to noticeably make any difference in the shape of the spectrum. Therefore, in our calculations we consider free-free spectrum as well-defined one without any parameter variations.

The instrument optics emission is considered in this paper as a graybody radiation with varying temperature T_{opt}:

$$I_{\mathrm{opt}} = \tau_{\mathrm{opt}}B(\nu, T_{\mathrm{opt}}). \qquad (28)$$

Temperature variations can depend on the stability of the cooling system, design features and quality of the optical system. In our numerical calculations we use a possible range of temperature T_{opt} variations as $T_{\mathrm{min}} \leq T_{\mathrm{opt}} \leq T_{\mathrm{max}}$ and $T_{\mathrm{max}} - T_{\mathrm{min}} \leq 2K$. The emissivity τ_{opt} depends on the quality of the polishing of the reflecting surfaces of the optical system. We use three different upper limits for emissivity: $\tau_{\mathrm{opt}} \leq 0.001$, $\tau_{\mathrm{opt}} \leq 0.01$ and $\tau_{\mathrm{opt}} \leq 0.05$.

The photon noise comes from the CMB radiation, a number of main foregrounds and emission from the instrument. All these sources of fluctuations can be characterized by the noise equivalent power (NEP) which is measured in units of W/Hz$^{1/2}$. For the FTS, intensity of the noise (i.e. the 1σ sensitivity) is given by the following equation [53]:

$$\sigma = 0.61 \frac{NEP}{\Delta\nu\sqrt{tG}}, \qquad (29)$$

where t is the integration time and G is the throughput of the system. The value G is the solid angle of the entrance pupil seen from the detector multiplied by the area of the instrument feedhorn. For the diffraction-limited instrument it can be estimated as

$$G_{\mathrm{dif}} = (0.61\pi c/\nu_{\mathrm{min}})^2. \qquad (30)$$

However, the instruments we analyze are not necessarily diffraction-limited, so generally it is possible to have $G \gg G_{\mathrm{dif}}$.

The NEP is computed as

$$NEP^2 = NEP_{\mathrm{det}}^2 + NEP_F^2, \qquad (31)$$

where NEP_{det} is the intrinsic NEP of the detector while NEP_F is created by M various foregrounds [see Eq. (1)]. We assume that $NEP_{\mathrm{det}} \leq 10^{-19}\ \mathrm{W \cdot Hz}^{-1/2} \ll NEP_F$ in our estimates. Therefore,
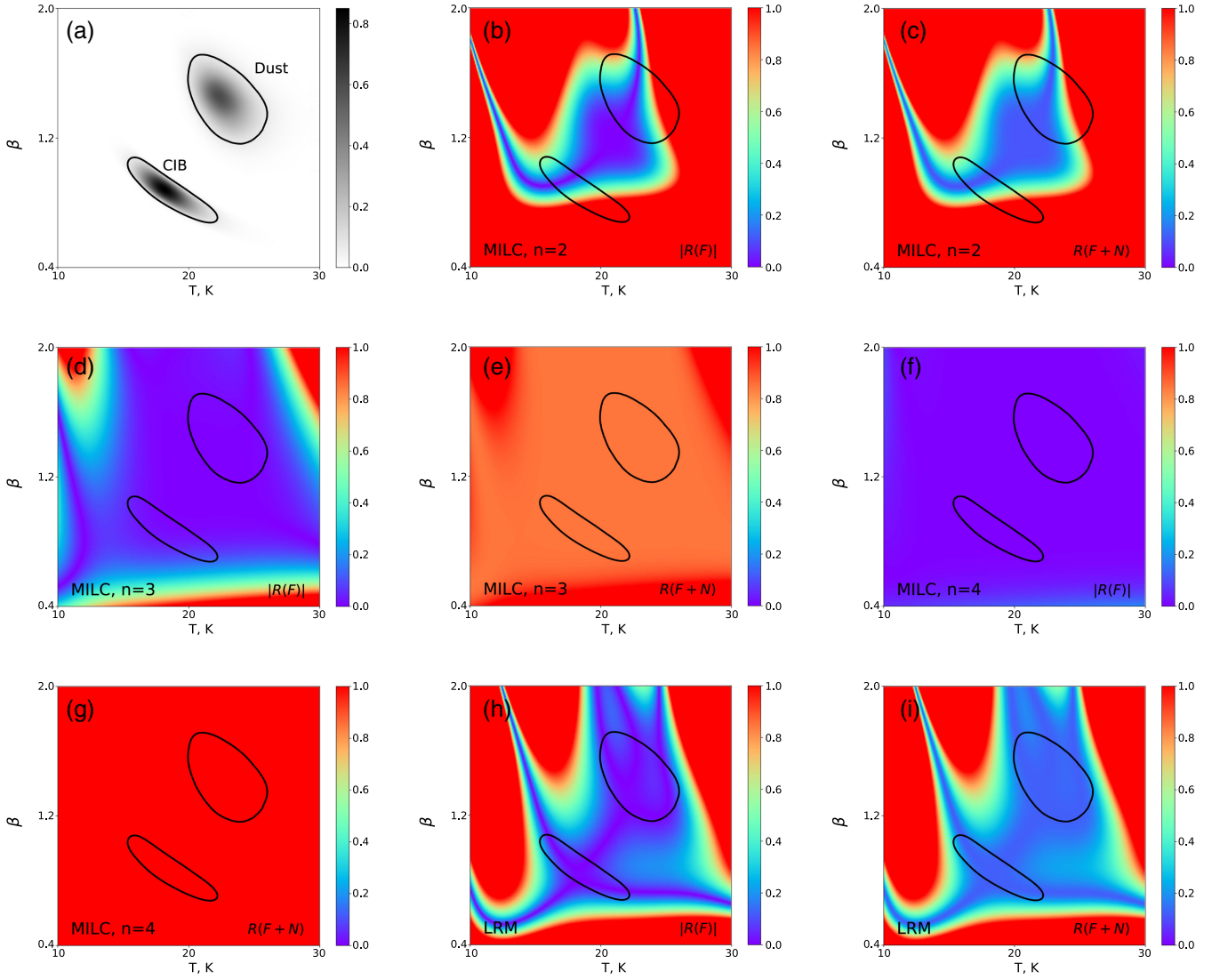
FIG. 1.   MILC and LRM methods for separating μ distortions when only dust and CIB are taken as foregrounds. Panel (a) shows the probability distribution function for T and β parameters. Isocontour lines limit the Ω region of parameter variations. Panel (b) shows MILC response to foreground $|R(\mathbf{F})|$ if n = 2. Dark red indicates the region where the response to the foreground exceeds the response to the μ signal: $|R(\mathbf{F})| \geq 1$. Panel (c) shows the total MILC noise + signal response $R(\mathbf{F}+\mathbf{N})$ if n = 2. Panel (d) $|R(\mathbf{F})|$ for MILC, n = 3. Panel (e) $R(\mathbf{F}+\mathbf{N})$ for MILC, n = 3. Panel (f) $|R(\mathbf{F})|$ for MILC, n = 4. Panel (g) $R(\mathbf{F}+\mathbf{N})$ for MILC, n = 4. Panel (h) $|R(\mathbf{F})|$ for LRM. Panel (i) $R(\mathbf{F}+\mathbf{N})$ for LRM.

$$NEP^2 \approx NEP_F^2 = \sum_{m=1}^{M} NEP_m^2,$$

$$NEP_m^2 = \int_{\nu_{\min}}^{\nu_{\max}} \frac{4Gh^2}{c^2} \nu^4 n_m(\nu)[1 + n_m(\nu)]d\nu, \quad (32)$$

where $n_m(\nu) = \frac{c^2}{2h\nu^3} I_m(\nu)$ is the photon concentration in phase space for mth foreground component. Here we assume ideal optical and main beam efficiency.

If the noise is determined by a weak constant foreground $I(\nu) = $ const with $n_m \ll 1$, from (29) and (32) one can obtain,

$$\sigma \propto \sqrt{\frac{I}{t} \frac{\nu_{\max} - \nu_{\min}}{\Delta\nu}},$$

which coincides with the often used estimate of FTS sensitivity, see, e.g. [54].

As one can see from Eqs. (29)–(32), the noise amplitude depends on the intensity of radiation coming from the sky and from the all optical system, on the FTS frequency range and on the spectral resolution. The main sources of the noise are CMB monopole, dust radiation, CIB and the radiation emitted by the instrument optics. In this section we use a white noise model. Therefore, the noise covariance matrix has a diagonal form $\mathbf{C} = \sigma^2 \mathbf{E}$, where σ is the

sensitivity. This kind of noise distribution corresponds to a single frequency band of the FTS instrument.

We consider the single band receiver for 384 channels (7.5 GHz each) from 15 GHz to 2895 GHz. The throughput is computed as for a diffraction-limited system, i.e. from Eq. (30).

B. Numerical results

1. Dust and CIB foregrounds

We start our analysis by comparing the efficiency of the MILC and LRM methods for $I_\mu(\nu)$ distortion detection with a simple example when only dust and infrared radiation are taken into account as foregrounds. Noise level (or sensitivity) in this example is $\sigma = 1$ Jy/sr per single channel. The MILC method was applied for three different versions of the Taylor series expansion of the modified black body in terms of parameters T and β [see Eq. (25)] to the second, third and fourth order respectively in the vicinity of the reference value $\mathbf{P_0}$ of the vector $\mathbf{P}(T,\beta)$: $\mathbf{P_0} = (T_0, \beta_0)$, $T_0 = 21.2$ K $\beta_0 = 1.11$ in Eq. (12). So the number of constraints zeroing derivatives over two variables up to the nth order are as follows: 6 constraints for $n = 2$, 10 constraints for $n = 3$ and 15 constraints for $n = 4$. The results are shown in Fig. 1.

For $n = 2$, the response to foreground $|R(\mathbf{F})|$ and the total response to foreground and noise $R(\mathbf{F} + \mathbf{N})$ are shown in Figs. 1(b) and 1(c) respectively in comparison with the response to the μ signal $R(\mathbf{I}_\mu) = 1$. Obviously, the constraints imposed in Eq. (12) for $n = 2$ are not enough to clean the area Ω from the influence of the foreground, that

is, to ensure the condition $|R(\mathbf{F})| \ll R(\mathbf{I}_\mu)$ inside two spots limited by isocontour lines.

For $n = 3$, the foreground response is quite low compared to the signal response, although the noise response becomes relatively large and as a result $R(\mathbf{F} + \mathbf{N}) \approx 0.8R(\mathbf{I}_\mu)$, see Figs. 1(d) and 1(e).

In the case when $n = 4$ [Figs. 1(f) and 1(g)], the region Ω is excessively clean from the influence of the foreground, but at the same time 15 constraints in Eq. (12) lead to an extremely high response to random noise for the given sensitivity, $|R(\mathbf{N})| \gg R(\mathbf{I}_\mu)$.

Thus, the choice of $n = 3$ is the optimal number of moments in the decomposition of dust and CIB spectra when using the MILC method. Figure 2 shows the dependence of responses to foreground and noise on the number of constraints imposed. As this number increases, the response to the foreground gets smaller and, at the same time, the response to random noise rises due to a decrease in the index Γ which is a measure of orthogonality between foreground and μ signal [see Eq. (17)]. So the total response to the foreground + noise reaches its minimum at a certain number of constraints.

LRM, in contrast to MILC, assumes one single condition on foregrounds and noise and provides us with the optimal weight vector $\boldsymbol{\omega}$ to minimize the ratio (foreground + noise)/signal. The result can be seen in Figs. 1(h) and 1(i) for foreground and foreground + noise responses, respectively and in Fig. 2.

An interesting approach is to study the responses to the foreground and random noise on the sensitivity of the experiment: $R(\mathbf{F}, \sigma)$ and $R(\mathbf{N}, \sigma)$. This makes it possible to
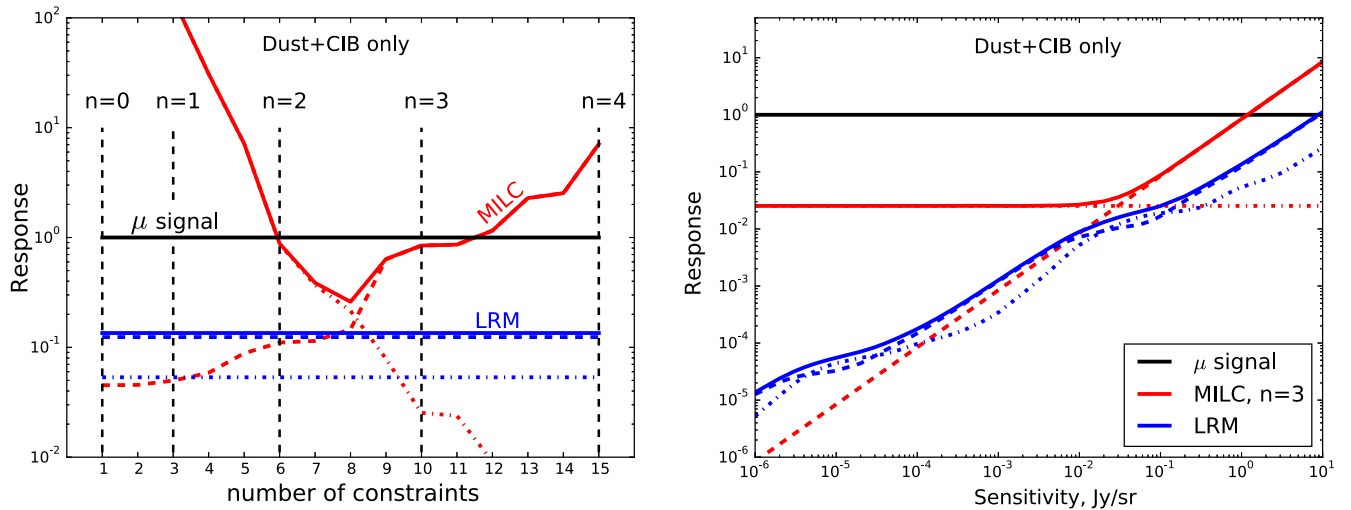

FIG. 2.   Left panel: red solid line: total response to foreground + noise for MILC as a function of the number of constraints. Red dashed line: Noise response for MILC. Red dash-dotted line: MILC response to foreground. The blue solid, dashed, and dash-dotted lines show LRM responses to foreground + noise, noise, and foreground, respectively. Right panel: responses to noise and foreground as functions of sensitivity for MILC (n = 3) and LRM approaches. Solid, dashed, and dash-dotted lines represent the same as in the left panel.

estimate the necessary sensitivity for reliable detection of certain spectral distortions in the presence of foregrounds. In addition, this dependence allows one to compare the effectiveness of different methods of data cleaning.

Figure 2 (right panel) shows the response to foreground and noise as a function of σ for the MILC and LRM methods. It is important to note that for the MILC method, the weight vector $\boldsymbol{\omega}$ does not depend on the sensitivity. Thus, the dependence of the response to noise on σ is linear $\langle R^2(\mathbf{N})\rangle^{\frac{1}{2}} = \sigma(\boldsymbol{\omega}\cdot\boldsymbol{\omega}^T)^{\frac{1}{2}}$. Since the response to the foreground is constant and given by Eq. (13), the overall response to noise and signal as a function of σ has a simple analytical form:

$$R_{\text{MILC}}(\mathbf{F}+\mathbf{N},\sigma) = \left[(\boldsymbol{\omega}\cdot\boldsymbol{\omega}^T)\sigma^2 + R_{n+1}^2(\mathbf{F})\right]^{\frac{1}{2}}. \quad (33)$$

Unlike the MILC method, when applying the LRM filtering, the weights ω_j depend on the magnitude of the photon noise: $\boldsymbol{\omega} = \boldsymbol{\omega}(\sigma)$. As can be seen from Fig. 2, the total response to noise and foreground for the LRM approach is always significantly less than for the MILS and provides a better signal/(foreground + noise) ratio:

$$R_{\text{LRM}}(\mathbf{F}+\mathbf{N},\sigma) < R_{\text{MILC}}(\mathbf{F}+\mathbf{N},\sigma). \quad (34)$$

This inequality is true for any sensitivity σ. The reason for this is that LRM, unlike MILC, does not require complete orthogonality of the investigated signal to the foreground.

2. All foregrounds

Below we perform an analysis of two approaches to the problem of signal separation, taking into account all the main foregrounds, adding them to the already considered components of dust and infrared radiation. As mentioned above, the signals of relic origin and free-free emission are considered as well-defined components of the spectrum without parameter variations. They are only subject to restrictions from above in terms of amplitudes.

The model of synchrotron radiation depends on one parameter β_s which varies from 0.9 to 1.4. Empirically, it was found that three derivatives of the expansion in a Taylor series in the vicinity of the reference point $\beta_s = 1.15$ are sufficient to get rid of the response to this component for MILC method. Thus, the synchrotron requires four constraints for the weights $\boldsymbol{\omega}$ (the function itself and three derivatives).

The spectrum emitted by the optical system of the instrument in our example depends only on its temperature in the vicinity of $T_{\text{opt}} = 10$ K and can vary between 9 K and 11 K. To reliably get rid of the influence of this radiation on data analysis, it is necessary to zero the response to the function $B(\nu, T_{\text{opt}})$ and seven derivatives over temperature: $\partial^n B(\nu, T_{\text{opt}})/\partial T_{\text{opt}}^n|_{T_{\text{opt}}=10\,\text{K}}$, $n \leq 7$ (eight constraints in total).

The results for μ signal extraction for both methods are shown in Fig. 3 (left panel). The more components are taken into account, the larger the foreground + noise response becomes for both methods. Nevertheless LRM always shows an acceptable result, while MILC gives almost infinitely large response to noise when trying to
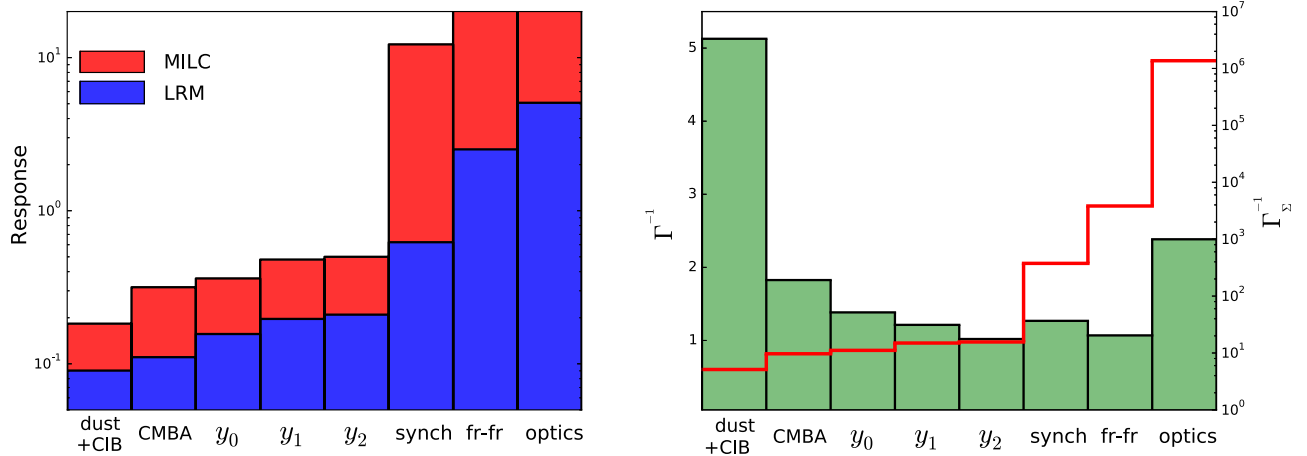

FIG. 3. Left panel: the histogram represents the response to foreground + noise with the sequential addition (one by one from left to right) of various components to the studied signal. Thus, the leftmost column shows the response to foreground + noise in the presence of only dust and CIB radiation as foreground, and the rightmost column shows this response if all the components listed along the horizontal axis are taken into account. Calculations were made for MILC and LRM to separate μ distortions and correspond to the sensitivity $\sigma = 1$ Jy/sr. Right panel: the green histogram shows the measure of orthogonality Γ_c of the μ signal to each individual component. The red step line is the measure of orthogonality Γ_Σ of μ distortion to all components to the left of the step in question (similar to the left panel).
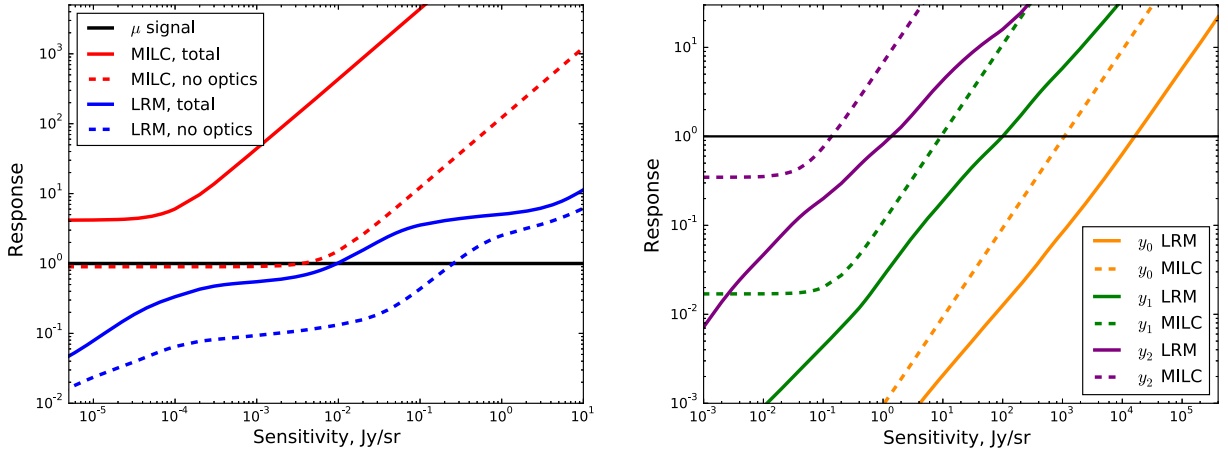
FIG. 4. Left panel: the result of applying the MILC and LRM methods to separate the μ signal. The dependencies of the total noise + foreground response on the sensitivity of the experiment in the presence of all the main foregrounds are shown. The dashed lines show results for an "ideal" experiment with no influence from the instrument optical system. Right panel shows the results of extracting y_0, y_1 and y_2 signals for the LRM and for MILC approaches. We show the responses to noise + foreground when taking into account all the main foregrounds except the instrument optics (sky difference should be used).

get rid of all foregrounds. In the same figure (right panel) we show the orthogonality measures of individual components to the μ signal: $\Gamma_m = \Gamma_{\text{dust+CIB}}, \Gamma_{\text{CMBA}}, \ldots, \Gamma_{\text{optics}}$ along with a orthogonality measure of the aggregate of components Γ_Σ. It is important to note that there is no direct analytical relation between Γ_Σ and Γ_m since the foreground signals are not orthogonal to each other. It is easy to see, that $R_{\text{MILC}}(\mathbf{F} + \mathbf{N}) \sim \Gamma_\Sigma^{-1}$ as it should be.

Finally, Fig. 4 shows the dependence of the response $R(\mathbf{F} + \mathbf{N}, \sigma)$ on the sensitivity of the experiment in the presence of all foregrounds for both methods for μ signal detection (left panel). The right panel shows the same functions for the LRM approach in cases when the signals of interest are $I_{y_0}, I_{y_1}, I_{y_2}$. In this case, the foreground created by the instrument is not taken into account, since when observing y distortions, it is possible to use the sky difference, which automatically excludes this component from the observed signal.

3. Optimal temperature for an instrument's optical system

For most space experiments, it is preferable to cool the telescope's primary mirror and other mirrors as much as possible to avoid creating additional photon noise and degrading the sensitivity. Below we will demonstrate that cooling the system of mirrors to too much low temperature, that is, close to the CMB temperature ~ 3 K, can significantly worsen the result if the signal of interest is μ-type distortion. Indeed, on the one hand, a decrease in the instrument optics temperature leads to a decrease in photon noise, on the other hand, the approach of this temperature to the temperature of CMB reduces the degree of orthogonality of the μ signal to the foreground created by the optics.

Despite the fact that LRM does not assume complete orthogonality to the foregrounds (including orthogonality to the signal created by the optics), the combination of CMB related signals together with a poorly defined signal from the optics can, to a certain extent, mimic the desired μ signal. This inevitably increases the response to foreground + noise and, thereby, reduces the sensitivity of the experiment to μ distortion measurements. Figure 5 demonstrates this effect. Shown here are the responses $R(\mathbf{F} + \mathbf{N}, T_{\text{opt}})$ as a functions of the temperature of the
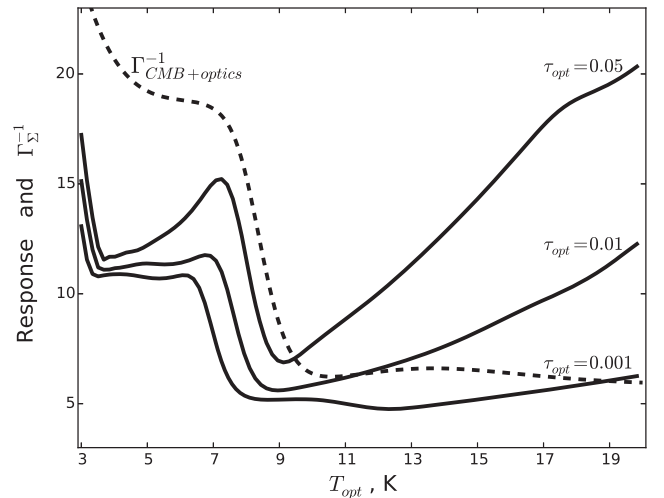

FIG. 5. Solid lines show the response to foreground + noise as a function of the temperature of the instrument optics for different emissivity values. The dashed line corresponds to the value which is inverse to the measure of orthogonality between I_μ and a set of signals associated with CMB and optics: $I_{\text{CMBA}}, I_{y_0}, I_{y_1}, I_{y_2}, I_{\text{optics}}$.

optics for three different values of optics emissivity $\tau_{\rm opt}$. The responses are normalized in such a way that for all three considered $\tau_{\rm opt}$, a sensitivity of $1 Jy/sr$ is achieved at $T_{\rm opt} = 10$ K. In the same figure we show the orthogonality measure of the I_μ to the combination of CMB related signals and optics: $I_{\rm CMBA}, I_{y_0}, I_{y_1}, I_{y_2}, I_{\rm opt}$. We denote this measure as $\Gamma_{\rm CMB+optics}$. It is assumed that the temperature can vary within $T_{\rm opt} \pm 1$ K. At low temperatures, this response behaves in a similar way as $\Gamma_{\rm CMB+optics}^{-1}$ and, therefore, increases strongly if $T_{\rm opt} \rightarrow T_{\rm CMB}$. At high temperatures the response increases due to an increase in photon noise. This effect obviously depends on the emissivity of the optics $\tau_{\rm opt}$. Nevertheless, even for small emissivity, cooling the mirror to temperatures close to $T_{\rm CMB}$ is undesirable. Finally the minimum response is reached when $T_{\rm opt} \sim 9$ K.

It should be noted that the dependence we calculated will change for a different FTS configuration, but the effect of increasing the response to foreground + noise at $T_{\rm opt} \sim T_{\rm CMB}$ will remain relatively strong in any case.

C. The instrument

Below we present the example of a possible space mission for measuring CMB spectral distortions of μ type. Figure 6 shows the top view of the instrument called SIMBAD (Spectroscopic Interferometer for Microwave BAckground Distortions). The concept, originally developed to be placed at L2, has been then proposed to be deposited in a permanently shadowed region (PSR) present in some polar lunar craters [35] since they could offer a very low temperature, passive cooling for the instrument. The property of a few spots within the PSRs which reached 18 K has been put forward, as providing a much lower temperature than is obtained at L2 for JWST [55] to the
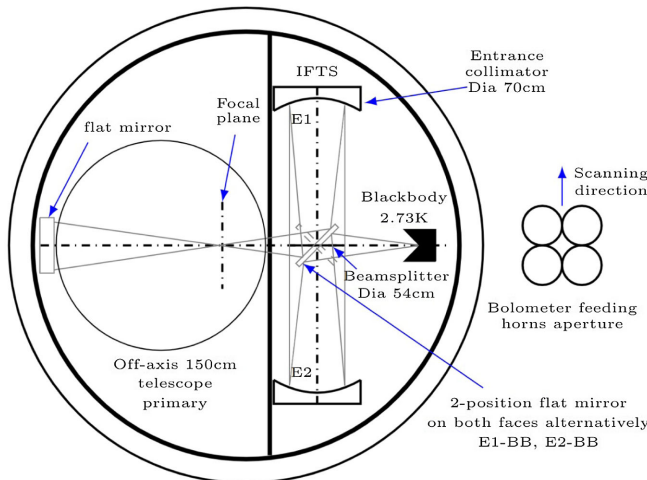

FIG. 6.   SIMBAD instrument.

price of a huge sunscreen. In this version of SIMBAD (Fig. 6) the diameter of the telescope primary mirror is 1.5 m, feeding a dual-input imaging FTS [54].

All the instrument is placed in a cryostat designed to fit into the internal fairing of the largest cargo launcher in development to go to the Moon [56]. To bend the beam, the telescope primary mirror and the following flat mirror must be actively cooled at 9 K according to our results (Fig. 5). But the IFTS behind (right part of the cryostat) is actively cooled at a lower temperature, close to the CMB temperature, since its contribution to the foregrounds is cancelled by the differential concept of the instrument, reducing the photon noise budget. If the SIMBAD cryostat is installed on one of the cold lunar spots at 18 K, as described in [35], then such a low temperature passive cooling could reduce the power needed for the active cooling of the whole instrument.

At the two IFTS outputs there are 2×2 bolometers with their feeding horn as shown on Fig. 6, each of $1.5°$ aperture on the sky, providing a total field coverage of $3° \times 3°$ for each data acquisition. The spectral coverage of the IFTS is going from 15 to 2895 GHz, to cover the full CMB domain. The maximum motion of the moving mirror of the IFTS (6 mm) is adjusted to reach a 7.5 GHz resolution on all the frequency domain, providing 384 equal frequency channels. To be as simple as possible, the telescope has no tracking motion. The direction of pointing is determined by three adjustable legs which support the SIMBAD cryostat on the lunar ground. The scanning of a full ring of $3°$ width is simply obtained by the rotation of the Moon. Thanks to the IFTS [54], the four bolometers are integrating simultaneously, making a gain of a factor 2 in sensitivity for μ compared to an FTS with a single detector. In paper [35] the sensitivity of 1 Jy/sr is reached by a SIMBAD instrument with a frequency resolution of 15 GHz and a 72 cm telescope diameter, but already a total field-of-view of $3° \times 3°$ obtained by four detectors observing simultaneously for a total integration time of 4 years. With the new SIMBAD with a frequency resolution divided by 2 and a collecting area multiplied by 4, from Eq. (29), the same total integration time is needed. As proposed in [35] the sensitivity and the field of view can be improved by placing in the same lunar crater several identical SIMBAD units pointing on adjacent directions scanning a wider ring.

It is important to note that the instrument requirements for measuring μ, y signals and recombination lines vary. For a μ signal, it is not important for us to have good angular resolution, but we need reliable calibration of the instrument and the temperature of its optical system should be about 9 K. At the same time, y distortions and especially relativistic corrections y_1, y_2 have a small angular size. For their observations, it is desirable to have good angular resolution, but there is no need to have a contrast between the CMB temperature and the temperature of the

instrument. Besides the calibration of the instrument for this task is not required. To detect extremely weak signals from hydrogen and helium recombination lines, it is desirable to cool the instrument optics as much as possible to reduce photon noise.

Thus, it is necessary to look for a very complex compromise to combine all these requirements in one device if the goal of the mission is to simultaneously solve all three problems. SIMBAD is mostly targeted for μ signal measurements.

IV. CONCLUSIONS AND DISCUSSION

Detecting CMB spectral distortions is one of the key goals of modern observational cosmology. Hidden in foregrounds of cosmic and instrumental origin, these deviations of the CMB spectrum from a perfect black body shape contain a wealth of information unobtainable by other observational methods. To extract CMB frequency spectral distortions, a signal-foreground separation algorithm was introduced. The key feature of this algorithm is its weak sensitivity to the foregrounds spectral shapes.

Given the areas of foreground parameters variation, which are relatively easy to determine, and the upper bounds of the foreground amplitudes, whose overestimation would not skew our results, we are able to minimize the response both to the foregrounds and to noise. With this in mind, we have compared the performance of our LRM algorithm with the MILC method.

Starting off with a limited number of foregrounds (dust and CIB) we define the minimal number of moments/constraints optimal for the MILC analysis. Since LRM allows for the investigated signal to be nonorthogonal to the foreground, the total response to noise and foreground for our method is significantly less than for the MILC for any sensitivity value. Subsequently, we broaden the scope of our analysis by including CMB anisotropy, y-distortions, synchrotron and free-free emission and radiation created by the instrument optics. Even though adding more components increases the total response for both methods, the response to noise obtained with MILC exceeds LRM results by a wide margin. In this paper we did not take into account Zodiacal light and CO lines as foregrounds in our modeling.

It is important to note that our LRM method allows for an improvement by subtracting foreground components one by one from the observed signal. Applying our method sequentially to all foreground components, one can estimate the real foreground amplitudes a_m. By subtracting foregrounds with estimated amplitudes from the total signal, new restrictions can be made on the maximum possible amplitudes of foreground residuals A_m. Reducing the upper limits of foreground amplitudes leads to a decrease in the response to foreground + noise and, therefore, more accurate estimation of the spectral distortion amplitudes. Implementation of such improvement is out of the scope of this paper.

Finally, we have established that cooling the instrument down to the temperature close to the temperature of relic radiation spoils the results drastically when measuring μ type distortions. This occurs mainly due to a decrease in the measure of orthogonality of the μ signal to the instrument optics component. If the system of mirrors is cooled to the temperature of the relic, it itself begins to create distortions close in shape to the CMB spectral distortions. The optimal result is reached when the temperature of the mirrors is around 9 K.

These reasonings do not apply to the measurement of y perturbations, since in this case the signal difference can be used, which leads to the automatic exclusion of the component created by the instrument optics from the observational data. Also, these arguments are not relevant to the measurement of recombination lines because their shape is not related to possible distortions of the blackbody spectrum. Note that data processing for any physical experiment with poorly defined foregrounds can use the LRM approach.

[1] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern *et al.*, Astrophys. J. Suppl. Ser. **208,** 20 (2013).

[2] 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.* (Planck Collaboration), Astron. Astrophys. **571,** A15 (2014).

[3] Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak *et al.* (Planck Collaboration), Astron. Astrophys. **641,** A4 (2020).

[4] J. Chluba, M. H. Abitbol, N. Aghanim, Y. Ali-Haïmoud, M. Alvarez, K. Basu, B. Bolliet, C. Burigana, P. de Bernardis, J. Delabrouille *et al.*, Exp. Astron. **51,** 1515 (2021).

[5] J. Silk, in *APS April Meeting Abstracts* (APS, 2021), Vol. 2021, p. B21.003.

[6] J. Chluba and R. A. Sunyaev, Mon. Not. R. Astron. Soc. **419,** 1294 (2012).

[7] J. Silk and J. Chluba, Science **344,** 586 (2014).

[8] G. De Zotti, M. Negrello, G. Castex, A. Lapi, and M. Bonato, J. Cosmol. Astropart. Phys. 03 (2016) 047.

[9] J. Chluba, Mon. Not. R. Astron. Soc. **460,** 227 (2016).

[10] H. Tashiro, Prog. Theor. Exp. Phys. **2014,** 06B107 (2014).

[11] Y. B. Zel'dovich and R. A. Sunyaev, Astrophys. Space Sci. **4,** 301 (1969).

[12] C. Burigana, L. Danese, and G. de Zotti, Astron. Astrophys. **246,** 49 (1991).

[13] J. Chluba, A. Kogut, S. P. Patil, M. H. Abitbol, N. Aghanim, and Y. Ali-Haïmoud, M. A. Amin, J. Aumont, N. Bartolo, K. Basu *et al.*, Bull. Am. Astron. Soc. **51,** 184 (2019).

[14] T. Nakama, B. Carr, and J. Silk, Phys. Rev. D **97,** 043525 (2018).

[15] R. A. Sunyaev and Y. B. Zel'dovich, Astrophys. Space Sci. **7,** 20 (1970).

[16] H. Fu, M. Lucca, S. Galli, E. S. Battistelli, D. C. Hooper, J. Lesgourgues, and N. Schöneberg, J. Cosmol. Astropart. Phys. 12 (2021) 050.

[17] R. A. Sunyaev and Y. B. Zeldovich, Astrophys. Space Sci. **9,** 368 (1970).

[18] R. A. Daly, Astrophys. J. **371,** 14 (1991).

[19] W. Hu, D. Scott, and J. Silk, Astrophys. J. Lett. **430,** L5 (1994).

[20] A. Ota, T. Takahashi, H. Tashiro, and M. Yamaguchi, J. Cosmol. Astropart. Phys. 10 (2014) 029.

[21] S. K. Acharya and J. Chluba, Mon. Not. R. Astron. Soc. **515,** 5775 (2022).

[22] S. Colafrancesco and P. Marchegiani, Astron. Astrophys. **535,** A108 (2011).

[23] S. Colafrancesco, P. Marchegiani, and R. Buonanno, Astron. Astrophys. **527,** L1 (2011).

[24] T. A. Enßlin and C. R. Kaiser, Astron. Astrophys. **360,** 417 (2000).

[25] P. Marchegiani and S. Colafrancesco, Mon. Not. R. Astron. Soc. **452,** 1328 (2015).

[26] I. G. Edigaryev, D. I. Novikov, and S. V. Pilipenko, Phys. Rev. D **98,** 123513 (2018).

[27] D. I. Novikov, S. V. Pilipenko, M. De Petris, G. Luzzi, and A. O. Mihalchenko, Phys. Rev. D **101,** 123510 (2020).

[28] J. C. Mather, E. S. Cheng, R. E. Eplee, Jr., R. B. Isaacman, S. S. Meyer, R. A. Shafer, R. Weiss, E. L. Wright, C. L. Bennett, N. W. Boggess *et al.*, Astrophys. J. Lett. **354,** L37 (1990).

[29] D. J. Fixsen, Astrophys. J. **707,** 916 (2009).

[30] A. Kogut, D. Fixsen, D. Chuss, J. Dotson, E. Dwek, M. Halpern, G. Hinshaw, S. Meyer, S. Moseley, M. Seiffert *et al.*, J. Cosmol. Astropart. Phys. 07 (2011) 025.

[31] A. Kogut, J. Chluba, D. J. Fixsen, S. Meyer, and D. Spergel, in *Space Telescopes and Instrumentation 2016: Optical,*

Infrared, and Millimeter Wave, edited by H. A. MacEwen, G. G. Fazio, M. Lystrup, N. Batalha, N. Siegler, and E. C. Tong, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9904 (Society of Photo-Optical Instrumentation Engineers (SPIE), Bellingham, 2016), p. 99040W.

[32] M. H. Abitbol, J. Chluba, J. C. Hill, and B. R. Johnson, Mon. Not. R. Astron. Soc. **471,** 1126 (2017).

[33] A. Kogut, D. Fixsen, N. Aghanim, J. Chluba, D. T. Chuss, J. Delabrouille, B. S. Hensley, J. C. Hill, B. Maffei, A. R. Pullen *et al.*, J. Cosmol. Astropart. Phys. 07 (2023) 057.

[34] I. D. Novikov, S. F. Likhachev, Y. A. Shchekinov, A. S. Andrianov, A. M. Baryshev, A. I. Vasyunin, D. Z. Wiebe, T. d. Graauw, A. G. Doroshkevich, I. I. Zinchenko *et al.*, Phys. Usp. **64,** 386 (2021).

[35] J.-P. Maillard, Phil. Trans. R. Soc. A **379,** 20200212 (2021).

[36] G. B. Rybicki and W. H. Press, Astrophys. J. **398,** 169 (1992).

[37] M. Tegmark, A. de Oliveira-Costa, and A. J. Hamilton, Phys. Rev. D **68,** 123523 (2003).

[38] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern *et al.*, Astrophys. J. Suppl. Ser. **148,** 97 (2003).

[39] H. K. Eriksen, A. J. Banday, K. M. Górski, and P. B. Lilje, Astrophys. J. **612,** 633 (2004).

[40] M. Remazeilles, J. Delabrouille, and J.-F. Cardoso, Mon. Not. R. Astron. Soc. **418,** 467 (2011).

[41] M. Remazeilles and J. Chluba, Mon. Not. R. Astron. Soc. **494,** 5734 (2020).

[42] V. Stolyarov, M. P. Hobson, A. N. Lasenby, and R. B. Barreiro, Mon. Not. R. Astron. Soc. **357,** 145 (2005).

[43] J. Chluba, J. C. Hill, and M. H. Abitbol, Mon. Not. R. Astron. Soc. **472,** 1195 (2017).

[44] A. Rotti and J. Chluba, Mon. Not. R. Astron. Soc. **500,** 976 (2020).

[45] Y. S. Abylkairov, O. Darwish, J. C. Hill, and B. D. Sherwin, Phys. Rev. D **103,** 103510 (2021).

[46] D. I. Novikov and A. O. Mihalchenko, companion paper, Phys. Rev. D **107,** 063506 (2023).

[47] A. D. Challinor, M. T. Ford, and A. N. Lasenby, Mon. Not. R. Astron. Soc. **312,** 159 (2000).

[48] A. Abergel, P. A. R. Ade, N. Aghanim, M. I. R. Alves, G. Aniano, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela *et al.* (Planck Collaboration), Astron. Astrophys. **571,** A11 (2014).

[49] N. Aghanim, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak *et al.* (Planck Collaboration), Astron. Astrophys. **596,** A109 (2016).

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

[51] E. de la Hoz, R. B. Barreiro, P. Vielva, E. Martínez-González, J. A. Rubiño-Martín, B. Casaponsa, F. Guidi, M. Ashdown, R. T. Génova-Santos, E. Artal *et al.*, Mon. Not. R. Astron. Soc. **519,** 3504 (2023).

[52] R. Adam, P. A. R. Ade, N. Aghanim, M. I. R. Alves, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi,

A. J. Banday *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A10 (2016).

[53] P. de Bernardis, S. Colafrancesco, G. D'Alessandro, L. Lamagna, P. Marchegiani, S. Masi, and A. Schillaci, Astron. Astrophys. **538**, A86 (2012).

[54] J. P. Maillard, L. Drissen, F. Grandmont, and S. Thibault, Exp. Astron. **35**, 527 (2013).

[55] J. P. Gardner, J. C. Mather, M. Clampin, R. Doyon, M. A. Greenhouse, H. B. Hammel, J. B. Hutchings, P. Jakobsen, S. J. Lilly, K. S. Long *et al.*, Space Sci. Rev. **123**, 485 (2006).

[56] NASA Space Launch System, 0080-sls-fact-sheet-jun2023-508-1.pdf.