

Rethinking CMB foregrounds: systematic extension of foreground parameterizations

Jens Chluba^{1*}, J. Colin Hill^{2†}, Maximilian H. Abitbol^{3‡}

¹ Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK

² Department of Astronomy, Columbia University, Pupin Hall, New York, New York 10027, USA

³ Department of Physics, Columbia University, New York, NY, 10027, USA

Accepted 2017 July 31. Received 2017 January 1

ABSTRACT

Future high-sensitivity measurements of the cosmic microwave background (CMB) anisotropies and energy spectrum will be limited by our understanding and modeling of foregrounds. Not only does more information need to be gathered and combined, but also novel approaches for the modeling of foregrounds, commensurate with the vast improvements in sensitivity, have to be explored. Here, we study the inevitable effects of *spatial averaging* on the spectral shapes of typical foreground components, introducing a moment approach, which naturally extends the list of foreground parameters that have to be determined through measurements or constrained by theoretical models. Foregrounds are thought of as a superposition of individual emitting volume elements along the line of sight and across the sky, which then are observed through an instrumental beam. The beam and line of sight averages are inevitable. Instead of assuming a specific model for the distributions of physical parameters, our method identifies natural new spectral shapes for each foreground component that can be used to extract parameter moments (e.g., mean, dispersion, cross-terms, etc.). The method is illustrated for the superposition of power-laws, free-free spectra, gray-body and modified blackbody spectra, but can be applied to more complicated fundamental spectral energy distributions. Here, we focus on intensity signals but the method can be extended to the case of polarized emission. The averaging process automatically produces scale-dependent spectral shapes and the moment method can be used to propagate the required information across scales in power spectrum estimates. The approach is not limited to applications to CMB foregrounds but could also be useful for the modeling of X-ray emission in clusters of galaxies.

Key words: Cosmology: cosmic microwave background – theory – observations

1 INTRODUCTION

Decades of observations of the cosmic microwave background (CMB) temperature and polarization anisotropies have clearly helped to establish the cosmological concordance model (e.g., Bennett et al. 2003; Planck Collaboration et al. 2016b). By now, cosmologists are close to exhausting all the primordial information from the CMB temperature anisotropies. The next steps are to tease all the information out of the *E*-mode polarization signal and the quest for the even smaller primordial polarization *B*-modes (Kogut et al. 2011; André et al. 2014; Matsumura et al. 2014; Abazajian et al. 2015), as well as to extract information about the late-time universe from secondary temperature anisotropies (e.g., CMB lensing and the Sunyaev-Zeldovich effect; Abazajian et al. 2016). In addition, CMB spectral distortion measurements, e.g., with *PIXIE* (Kogut et al. 2011) or *APSEra* (Sathyanarayana Rao et al. 2015), could provide a wealth of new complementary information about

the Universe we live in (e.g., Chluba & Sunyaev 2012; Sunyaev & Khatri 2013; Chluba 2013; Tashiro 2014; Chluba 2016).

But the aforementioned signals are (very) small and buried under (much) larger foregrounds. The hope is to be able to separate the cosmological signals from foregrounds using *multi-frequency observations*. In particular for CMB spectral distortions, mainly expected to be present uniformly across the sky, this is completely uncharted territory and the anticipated improvements over previous measurements (Mather et al. 1994; Fixsen et al. 1996) in terms of the raw spectral sensitivity could be dramatic (Kogut et al. 2011, 2016). Some prior information on spatially varying foreground signals is available, e.g., from measurements of *Planck* (Planck Collaboration et al. 2014a, 2016a); however, at the required level of future foreground modeling we cannot *reliably predict* what awaits us and restrictive assumptions about the foreground model can cause biases in the deduced parameters (e.g., Errard et al. 2016; Remazeilles et al. 2016; Kogut & Fixsen 2016). This calls for more *flexible* methods, which use a minimal set of prior assumptions, relevant to both CMB spectral distortions and *B*-mode searches. Some forecasts with explicit foreground modeling for primordial CMB distortions were presented in Abitbol et al. (2017).

* E-mail: Jens.Chluba@manchester.ac.uk

† E-mail: jch@astro.columbia.edu

‡ E-mail: mha2125@columbia.edu

In this work, we approach the foreground modeling problem from a more general point of view, directly considering the averaging process of spatially varying foregrounds i) along the line of sight, ii) within the experimental beam and iii) in spherical harmonic decompositions or analyses of sky-maps (e.g., when degrading the angular resolution). The goal is to identify generalized parameterizations for the foregrounds that allow us to incorporate additional features of the underlying distributions of physical parameters (e.g., spectral indices, dust temperatures, and *all* their statistical properties) at a level that is sufficient to robustly isolate the much smaller primordial signals.

The first problem is that spatially varying foregrounds are averaged within the beam and along the line of sight. These averages cannot be avoided and lead to modifications of the spectral response even if in every *infinitesimal volume element* the signal is given by a very simple known energy distribution. For example, in every direction the CMB (to very high precision) is given by a blackbody with spatially varying temperature. For a low-resolution beam, the average over these blackbodies introduces a beam y -distortion with y -parameter $y \approx \frac{1}{2} \langle \Delta T^2 / T_0^2 \rangle$, where $\langle \Delta T^2 / T_0^2 \rangle$ is the beam-averaged variance of the blackbody temperatures (Chluba & Sunyaev 2004). Since the CMB temperature fluctuations are small, this usually only causes a very small beam spectral distortion ($y \approx 10^{-9}$), which in principle can be accurately modeled using prior information from *Planck*. Still, this example already illustrates how averaging processes can affect the required parameterization of the signal (blackbody with varying temperature versus blackbody with y -type distortion), when high accuracy is reached.

Another example is the new spectral shapes introduced to the thermal Sunyaev-Zeldovich (SZ) effect (Zeldovich & Sunyaev 1969) by line of sight variations of the electron temperature inside hot galaxy clusters (Chluba et al. 2013). In every infinitesimal volume element, the thermal SZ distortion is determined by the optical depth and electron temperature (Sazonov & Sunyaev 1998; Itoh et al. 1998); however, the average of contributions from different volume elements is no longer accurately described by one mean temperature and optical depth, but depends on higher order moments of the temperature field (Chluba et al. 2013). The precise shape of the distortion can thus in principle be used to determine those moments, $\langle (kT_e)^k \rangle$, of the electron temperature distribution. Conversely, maps of these new observables contain all the information gained from multi-frequency SZ measurements and can be used to directly constrain the detailed structure (average temperature and density profiles) of the cluster atmosphere. Similar statements hold for the all-sky SZ distortion signal (Hill et al. 2015).

From detailed CMB measurements, we also know that the standard foreground components (e.g., synchrotron and dust) vary *spectrally* across the sky (e.g., Planck Collaboration et al. 2014a; Fuskeland & et al. 2014; Planck Collaboration et al. 2016a) and similarly we expect the spectral shapes from individual contributions to vary along the line of sight. While the effect of spatial variations across the sky can be minimized using higher angular resolution (\rightarrow pencil beams), the line of sight average cannot be avoided. Thus, even the superposition of extremely simple fundamental spectral energy distributions (SEDs), e.g., power-laws, will in general not be represented by a single power-law to arbitrary precision. As we explain here, with the moment method we can parametrize the effects on the SED caused by these averages. In a similar manner, we present a moment expansion for free-free emission, gray-body and modified blackbody spectra, as examples. We focus on developing and illustrating the moment method for intensity signals, but it can be readily extended to polarized foregrounds,

which could be relevant to future B -mode searches, as we investigate in a separate paper. Similarly, we plan to use the moment method to forecast the spectral distortion sensitivity of *PIXIE* in a future work.

While the direct moment representation for a specified fundamental SED is exact, a truncated moment hierarchy (finite number of moments) may or may not allow representing all possible spectral shapes to sufficient precision (i.e., leading to residuals that are consistent or below the noise) or may describe the signal with highly correlated parameters. In the latter case, orthogonalization schemes (e.g., Gram-Schmidt or principal component decompositions) and different choices for the weighting can provide an alternative approach, introducing new spectral amplitude parameters related to linear combinations of moments. We illustrate these aspects for gray-body spectra (Sect. 5); however, we do not find orthogonalization schemes to be as beneficial and in general cases the set of orthogonal functions also depends on the average parameters and experimental settings, rendering it a less attractive procedure.

While the developed method and discussion envisions applications to CMB observations, the approach is more general and should be applicable in other regimes (e.g., X-ray modeling). One can furthermore include more complicated fundamental spectra [e.g., to parametrize the anomalous microwave emission (Draine & Lazarian 1998; Ali-Haïmoud et al. 2009; Hoang et al. 2016) or for improved models of the spectral shapes of dust, synchrotron, or even line emission]. However, we leave these investigations to future work. Overall, we find that the moment expansion can be used to parametrize some of the fundamental effects caused by spatial averaging to high precision, providing a first step towards generalized foreground separation methods relevant to future CMB anisotropy and spectroscopic studies.

2 DETAILS OF THE MOMENT METHOD FOR SPATIALLY VARYING COMPONENTS

In this section, we provide details about how to construct the moment expansion for a given underlying physical emission process. We separate the effects of line of sight and beam averages on the one hand (Sec. 2.1) from the effect of spatial averages carried out on measured sky-signals (Sec. 2.2). Even for a perfect pencil beam, the line of sight averaging is *inevitable*, but we can simply associate this averaging process with the beam average itself.

2.1 Average signal within a beam

Let us first try to understand how the average over spatially varying spectral signals along the line of sight and within the beam of an experiment affects the measured signal. We shall consider the infinitesimal SED contribution, $\delta I_\nu(\hat{\gamma}, \mathbf{p})$, in a direction $\hat{\gamma}$, which varies with frequency ν and depends on a list of spatially varying parameters $\mathbf{p}(\mathbf{r})$ at different locations, \mathbf{r} , along the line of sight. The line of sight can be parametrized by an affine parameter $s(\hat{\gamma})$ in the direction $\hat{\gamma}$, so that the line of sight averaged SED is

$$I_\nu(\hat{\gamma}) \equiv \langle I_\nu(\hat{\gamma}, \mathbf{p}) \rangle_{\text{l.o.s.}} = \int \frac{dI_\nu(\hat{\gamma}, \mathbf{p})}{ds(\hat{\gamma})} ds(\hat{\gamma}), \quad (1)$$

where $\mathbf{p} = \mathbf{p}(s)$ is used for the integration. We can also introduce the beam average of some quantity $X(\hat{\gamma}, \nu) \equiv X_\nu(\hat{\gamma})$ by

$$X(\hat{\gamma}_c, \nu_c) \equiv \langle X(\hat{\gamma}, \nu) \rangle_\Omega = \int d^2\hat{\gamma} \int d\nu W(\hat{\gamma}, \nu) X(\hat{\gamma}, \nu), \quad (2)$$

where $W(\hat{\gamma}, \nu)$ describes the frequency-dependent beam of different channels centered at frequency, $\nu_c = \langle \nu \rangle_\Omega \equiv \int \nu W(\hat{\gamma}, \nu) d^2\hat{\gamma} d\nu$,

and in the average direction $\hat{\gamma}_c = \langle \hat{\gamma} \rangle_\Omega \equiv \int \hat{\gamma} W(\hat{\gamma}, \nu) d\nu d^2\hat{\gamma}$. The specific shape of $W(\hat{\gamma}, \nu)$ depends on the experiment, as discussed below. We normalize the beam as $\int W(\hat{\gamma}, \nu) d^2\hat{\gamma} d\nu \equiv \langle 1 \rangle_\Omega \equiv 1$.

We reiterate, the guiding picture is that for any observation in an average direction $\hat{\gamma}_c$, the signal from different emitters¹ along the line of sight is picked up and then averaged in different directions within the beam. Physically, the line of sight average has the same effect as the average in different directions, leading to a mixing of the fundamental SEDs of emitters with different spectral parameters, \mathbf{p} . It therefore does not have to be distinguished in the computation, and we will use $\langle X \rangle$ to denote both the line of sight and beam average as a single beam average. The effect of redshifting along the line of sight can be captured by transforming the frequency variable and adding appropriate volume weighting factors. This affects the relative contributions of emitters in the modeling, but is not explicitly discussed here without loss of generality. The beam can also have different shapes at different frequencies *within* the band, which furthermore can vary from channel to channel.

It is now useful to perform a multi-dimensional Taylor expansion of the underlying SED in the free model parameters

$$\begin{aligned} I_\nu(\mathbf{p}) &= I_\nu(\bar{\mathbf{p}}) + \sum_i (p_i - \bar{p}_i) \partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}}) \\ &+ \frac{1}{2!} \sum_i \sum_j (p_i - \bar{p}_i)(p_j - \bar{p}_j) \partial_{\bar{p}_i} \partial_{\bar{p}_j} I_\nu(\bar{\mathbf{p}}) \\ &+ \frac{1}{3!} \sum_i \sum_j \sum_k (p_i - \bar{p}_i)(p_j - \bar{p}_j)(p_k - \bar{p}_k) \partial_{\bar{p}_i} \partial_{\bar{p}_j} \partial_{\bar{p}_k} I_\nu(\bar{\mathbf{p}}) \\ &+ \dots \end{aligned} \quad (3)$$

where $\bar{\mathbf{p}}$ (only depending on $\hat{\gamma}_c$) describes the list of beam average-parameters that need to be specified. In the notation, we suppressed the spatial dependence of $p_i = p_i(\mathbf{r})$ for convenience. We also use the notation $\partial_{\bar{p}_i} \dots \partial_{\bar{p}_j} I_\nu(\bar{\mathbf{p}}) = \partial_{p_i} \dots \partial_{p_j} I_\nu(\mathbf{p})|_{\mathbf{p}=\bar{\mathbf{p}}}$ for brevity.

To fix the average parameters, $\bar{\mathbf{p}}$, for a given spectral distribution function, we require the first order terms in the Taylor expansion to vanish, $\langle \sum_i [p_i(\mathbf{r}) - \bar{p}_i] \partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}}) \rangle = 0$. This has the solution

$$\bar{p}_i = \frac{\langle p_i(\mathbf{r}) \partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}}) \rangle}{\langle \partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}}) \rangle} \quad (4)$$

for $\langle \partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}}) \rangle \neq 0$. Generally, this does *not* imply $\bar{p}_i = \langle p_i(\mathbf{r}) \rangle$ since the frequency dependence of the beam can couple different spatial regions, where the frequency-weighting function is modified by $\partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}})$. If $W(\hat{\gamma}, \nu) = B(\hat{\gamma})F(\nu)$, one can separate the frequency and spatial averages, i.e., $\langle p_i(\mathbf{r}) \partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}}) \rangle = \langle p_i(\mathbf{r}) \rangle \langle \partial_{\bar{p}_i} I_\nu(\bar{\mathbf{p}}) \rangle$, such that we recover $\bar{p}_i \equiv \langle p_i(\mathbf{r}) \rangle$. Here, $B(\hat{\gamma})$ determines the spatial shape of the beam, which does not vary at different frequencies within each channel, and $F(\nu)$ describes the shape of the bandpass. Both can still vary from channel to channel.

We note immediately that the chosen definition for \bar{p}_i is not always the recovered best-fitting value from the analysis. This is because the spectral functions (see below) related to higher order moments are not necessarily linearly independent from the first order derivative spectra. Thus, the best-fitting values for a truncated moment expansion can receive contributions from higher order terms. In some cases, this is not a severe problem (superposition of power-laws and free-free spectra), but, for instance, in superpositions of

gray-body and modified blackbody spectra, these values can differ notably in the non-perturbative regime. Alternative weighting schemes to determine the best estimate for the obtained value of \bar{p}_i from given parameter distribution functions are discussed below.

2.1.1 Definition of Taylor-moments

Formally, the above expressions can be simplified by introducing the Taylor-moments

$$\omega_{i\dots j} = \frac{\langle [p_i(\mathbf{r}) - \bar{p}_i] \dots [p_j(\mathbf{r}) - \bar{p}_j] \partial_{\bar{p}_i} \dots \partial_{\bar{p}_j} I_\nu(\bar{\mathbf{p}}) \rangle}{\langle \partial_{\bar{p}_i} \dots \partial_{\bar{p}_j} I_\nu(\bar{\mathbf{p}}) \rangle} \quad (5)$$

and spectral functions

$$I_{i\dots j}(\nu_c, \bar{\mathbf{p}}) = \langle \partial_{\bar{p}_i} \dots \partial_{\bar{p}_j} I_\nu(\bar{\mathbf{p}}) \rangle, \quad (6)$$

where ν_c is the central frequency of the considered channel. Here the ‘ \dots ’ can contain any number and permutation of the parameter indices. **The total number of parameters that appears defines the order of the moment, and usually it is implicitly assumed that a perturbative expansion is possible, suggesting that a finite number of moments, ranked by their (derivative) order, can be used.** The spectral shapes, $I_{i\dots j}(\nu_c, \bar{\mathbf{p}})$, provide a (generally non-orthogonal and non-complete) basis for a vector space that spans the range of possible signals that can be captured by the moment expansion. With these definitions, we then find

$$\begin{aligned} \langle I_\nu(\mathbf{p}) \rangle &= I(\nu_c, \bar{\mathbf{p}}) + \sum_i \omega_i I_i(\nu_c, \bar{\mathbf{p}}) + \frac{1}{2} \sum_i \sum_j \omega_{ij} I_{ij}(\nu_c, \bar{\mathbf{p}}) \\ &+ \frac{1}{6} \sum_i \sum_j \sum_k \omega_{ijk} I_{ijk}(\nu_c, \bar{\mathbf{p}}) + \dots \end{aligned} \quad (7)$$

where $\omega_i \equiv 0$ with the definition of $\bar{\mathbf{p}}$ given in Eq. (4). With this reformulation, we have identified a set of new parameters, $\omega_{i\dots j}$, and spectral functions, $I_{i\dots j}(\nu_c, \bar{\mathbf{p}})$, which can be used to describe the beam-averaged signal. For a given theoretical model, specified by the fundamental SED, $I_\nu(\mathbf{p})$, one can simply compute these quantities and use them in the modeling of the foregrounds. Generally the moments have to be individually determined (i.e., measured), while the spectral functions can be computed using the precise knowledge of the beam’s angle-averaged frequency dependence and the underlying spectral dependence of I_ν . This provides a powerful separation of spatial and frequency-dependent functions, which we will exploit below. In particular, by studying the dependence of the spectral functions and using high-resolution measurements of average foreground parameters, one can identify and limit the contributions of different terms at varying physical scales.

There are, however, several caveats. First, the moments, $\omega_{i\dots j}$, in general *depend* on the band in a non-trivial way. This means that within each band different sets of moments (with different weighting functions) contribute, unless the beam factorization $W(\hat{\gamma}, \nu) = B(\hat{\gamma})F(\nu)$ is possible, such that

$$\omega_{i\dots j} \equiv \langle [p_i(\mathbf{r}) - \bar{p}_i] \dots [p_j(\mathbf{r}) - \bar{p}_j] \rangle. \quad (8)$$

If this simplification is insufficient, one can in principle introduce additional parameters which are directly related to the beam properties and its coupling to different spectral shapes. By characterizing the beam carefully, this additional dependence can be modeled but it inevitably introduces additional weighting of the underlying spatially varying parameters and a steep growth in the number of independent moments. Thus, unless the effective list of parameters remains smaller than the number of channels, there is no way to

¹ Absorption can be thought of as negative emission unless the shape of the fundamental SED is affected. Otherwise, the SED parametrization has to be modified or both absorption and emission have to be treated independently.

separate the different components by using multi-frequency observations. This is indeed one of the big worries for future CMB observations, and the answer will depend on the foreground complexity and specifications of the experiments. It also renders a comparison and combination of data from different (ground-based and space-based) experiments more challenging.

A more detailed account of this problem is beyond the scope of this paper, but it has also been recognized in studies related to 21cm cosmology (Mozdzen et al. 2016). For simplicity we shall assume that the factorization $W(\hat{\gamma}, \nu) = B(\hat{\gamma})F(\nu)$ is possible, but the more general case in principle can also be parametrized using a moment expansion of the beam properties.

Second, a finite expansion in terms of the moments does not necessarily converge rapidly, especially in the non-perturbative regime. Furthermore, higher moments can be directly related to the lower moments (e.g., for a Gaussian the second moment fixes all higher moments), which implies that the effective number of degrees of freedom can be much smaller, while a large number of moments might be necessary to describe the averaged distribution. The moment method is agnostic, making very few *a priori* assumptions. This allows us to identify shortcomings of certain parameterizations, which assume a fixed form, as we illustrate here. Alternatively, we can include simplifying assumptions for the relations among the moments to return to a specific lower dimensional parameterization or reduce the parameter space (using priors) that is spanned by the spectral basis functions. This illustrates the range of possibilities when applying the moment method.

In the above, we assumed that a list of parameters (e.g., dust temperature, spectral indices, etc.) has been identified for the moment expansion. This is by no means a trivial statement and the choice of the right variable can be tricky. In addition, the weighting of different contributions has to be considered carefully, as we illustrate below for gray-body and dust spectra.

2.2 Spherical harmonic expansion and signal processing

The formalism described in the previous section can be extended to include the effect of weighted averages due to spherical harmonic expansions of sky signals or when processing maps (e.g., when degrading the angular resolution). We explicitly discuss the spherical harmonic transform of the sky. Formally, the spherical harmonic coefficients of $\langle X \rangle$ are given by

$$a_{\ell m}^X(\nu_c) = \int d^2\hat{\gamma}_c Y_{\ell m}^*(\hat{\gamma}_c) \bar{X}(\nu_c, \hat{\gamma}_c), \quad (9)$$

where $\bar{X}(\nu_c, \hat{\gamma}_c) = \langle X \rangle$ is the beam averaged quantity in the average direction, $\hat{\gamma}_c$, and channel frequency ν_c . We can see from Eq. (7) that two sources of spatial variations appear when carrying out the spherical harmonic expansion, i) those from the moments, $\omega_{i\dots j}(\hat{\gamma}_c)$, and ii) those from the variation of the mean parameters, $\bar{\mathbf{p}}(\hat{\gamma}_c)$ in different directions on the sky.

By simply following through all the steps that lead to the moment expansion, Eq. (7), we readily have the generalization

$$\begin{aligned} \bar{I}_{\nu_c}(\hat{\gamma}_c) &= I(\nu_c, \bar{\mathbf{p}}_0) + \sum_i \omega'_i(\hat{\gamma}_c) I_i(\nu_c, \bar{\mathbf{p}}_0) + \frac{1}{2} \sum_i \sum_j \omega'_{ij}(\hat{\gamma}_c) I_{ij}(\nu_c, \bar{\mathbf{p}}_0) \\ &+ \frac{1}{6} \sum_i \sum_j \sum_k \omega'_{ijk}(\hat{\gamma}_c) I_{ijk}(\nu_c, \bar{\mathbf{p}}_0) + \dots \end{aligned} \quad (10)$$

where $\bar{\mathbf{p}}_0$ is the sky-averaged parameter vector, which can be obtained by averaging Eq. (4) over all directions. Similarly, the mo-

ments follow from Eq. (5), such that overall we find

$$\begin{aligned} \bar{p}_{i,0} &= \int \frac{\langle p_i(r, \hat{\gamma}) \partial_{\bar{p}_{i,0}} I_{\nu}(\bar{\mathbf{p}}_0) \rangle}{\langle \partial_{\bar{p}_{i,0}} I_{\nu}(\bar{\mathbf{p}}_0) \rangle} \frac{d^2\hat{\gamma}_c}{4\pi} \\ \omega'_{i\dots j}(\hat{\gamma}_c) &= \frac{\langle [p_i(r, \hat{\gamma}) - \bar{p}_{i,0}] \dots [p_j(r, \hat{\gamma}) - \bar{p}_{j,0}] \partial_{\bar{p}_{i,0}} \dots \partial_{\bar{p}_{j,0}} I_{\nu}(\bar{\mathbf{p}}_0) \rangle}{\langle \partial_{\bar{p}_{i,0}} \dots \partial_{\bar{p}_{j,0}} I_{\nu}(\bar{\mathbf{p}}_0) \rangle}, \end{aligned} \quad (11)$$

where the average $\langle \dots \rangle$ connects $\hat{\gamma}$ and $\hat{\gamma}_c$ through the beam shape. The spherical harmonic coefficients of the moments are obtained using Eq. (9). The most important difference is that while by construction $\int \omega'_i(\hat{\gamma}_c) d^2\hat{\gamma}_c = 0$, in general $\omega'_i(\hat{\gamma}_c)$ no longer vanishes but fluctuates across the sky. Usually, only this term is considered, but higher order moments lead to new scale-dependent SED variations as we illustrate below. Also, in general $\omega'_{i\dots j}(\hat{\gamma}_c) \neq \omega_{i\dots j}(\hat{\gamma}_c)$, as the reference parameters were chosen using the sky average rather than the local quantity. With $W(\hat{\gamma}, \nu) = B(\hat{\gamma})F(\nu)$, we have

$$\begin{aligned} \bar{p}_{i,0} &= \int \bar{p}_i(\hat{\gamma}_c) \frac{d^2\hat{\gamma}_c}{4\pi} \\ \omega'_{i\dots j}(\hat{\gamma}_c) &= \langle [p_i(r, \hat{\gamma}) - \bar{p}_{i,0}] \dots [p_j(r, \hat{\gamma}) - \bar{p}_{j,0}] \rangle. \end{aligned} \quad (12)$$

For the first and second moments, we then find

$$\begin{aligned} \omega'_i(\hat{\gamma}_c) &= \langle [p_i(r, \hat{\gamma}) - \bar{p}_{i,0}] \rangle = \bar{p}_i(\hat{\gamma}_c) - \bar{p}_{i,0} \\ \omega'_{ij}(\hat{\gamma}_c) &= \langle [p_i(r, \hat{\gamma}) - \bar{p}_{i,0}][p_j(r, \hat{\gamma}) - \bar{p}_{j,0}] \rangle \\ &= \omega_{ij}(\hat{\gamma}_c) + [\bar{p}_i(\hat{\gamma}_c) - \bar{p}_{i,0}][\bar{p}_j(\hat{\gamma}_c) - \bar{p}_{j,0}], \end{aligned} \quad (13)$$

which directly shows how the fluctuations in \bar{p}_i enter as a new contribution to the moments. This can introduce new spectral shapes at different multipoles, even if in any direction one specific shape is present. We briefly highlight some of the effects here (Sect. 7), but discuss their relevance to the analysis of future *B*-mode polarization and distortion measurements elsewhere.

3 SUPERPOSITION OF POWER-LAW SPECTRA

Let us start with the simple example of a power-law spectral distribution, $I_{\nu} = A_0(\nu/\nu_0)^{\alpha}$. The free parameters $\mathbf{p} = \{A_0, \alpha\}$ are assumed to vary spatially. This description is typically used to model the foreground caused by synchrotron emission at low frequencies (Planck Collaboration et al. 2014a, 2016a). One common extension is to introduce curvature to the spectral index, $I_{\nu}^* = A_0(\nu/\nu_0)^{\alpha + \frac{1}{2}\beta \ln(\nu/\nu_0)}$. Let us try to see under which conditions this approximation works by using the moment method. For this, we first compute the derivatives of I_{ν} with respect to A_0 and α :

$$\partial_{A_0}^k I_{\nu}(A_0, \alpha) = \delta_{k1} (\nu/\nu_0)^{\alpha} = \delta_{k1} \frac{I_{\nu}(A_0, \alpha)}{A_0} \quad (14)$$

$$\partial_{\alpha}^k I_{\nu}(A_0, \alpha) = I_{\nu}(A_0, \alpha) \ln^k(\nu/\nu_0). \quad (15)$$

Since I_{ν} depends linearly on A_0 , only first order derivatives with respect to A_0 appear. Since the derivatives ∂_{A_0} and ∂_{α} commute, using Eq. (7), we thus find the expansion

$$\begin{aligned} \langle I_{\nu}(\mathbf{p}) \rangle &= I(\nu_c, \bar{\mathbf{p}}) + \omega_{12} I_{12}(\nu_c, \bar{\mathbf{p}}) + \frac{1}{2} \omega_{22} I_{22}(\nu_c, \bar{\mathbf{p}}) \\ &+ \frac{1}{2} \omega_{122} I_{122}(\nu_c, \bar{\mathbf{p}}) + \frac{1}{6} \omega_{222} I_{222}(\nu_c, \bar{\mathbf{p}}) \\ &+ \frac{1}{6} \omega_{1222} I_{1222}(\nu_c, \bar{\mathbf{p}}) + \frac{1}{24} \omega_{2222} I_{2222}(\nu_c, \bar{\mathbf{p}}) + \dots \end{aligned} \quad (16)$$

For the spectral functions, we then have

$$\bar{A}_0 I_{12\dots 2}(\nu_c, \bar{\mathbf{p}}) \equiv I_{2\dots 2}(\nu_c, \bar{\mathbf{p}}) = \langle \partial_{\alpha}^k I_{\nu}(\bar{\mathbf{p}}) \rangle = \langle I_{\nu}(\bar{\mathbf{p}}) \ln^k(\nu/\nu_0) \rangle. \quad (17)$$

Here, we assumed that k derivatives (indicated by the ‘...’ in the subscripts) with respect to the second parameter, α , were taken. For the moments, we then find

$$\omega_{2\dots 2} = \frac{\langle [A_0(\mathbf{r}) - \bar{A}_0][\alpha(\mathbf{r}) - \bar{\alpha}]^k \partial_{\bar{\alpha}}^k I_{\nu}(\bar{\mathbf{p}}) \rangle}{\langle \partial_{\bar{\alpha}}^k I_{\nu}(\bar{\mathbf{p}}) \rangle} \approx \langle [A_0 - \bar{A}_0][\alpha - \bar{\alpha}]^k \rangle$$

$$\omega_{2\dots 2} = \frac{\langle [\alpha(\mathbf{r}) - \bar{\alpha}]^k \partial_{\bar{\alpha}}^k I_{\nu}(\bar{\mathbf{p}}) \rangle}{\langle \partial_{\bar{\alpha}}^k I_{\nu}(\bar{\mathbf{p}}) \rangle} \approx \langle [\alpha - \bar{\alpha}]^k \rangle, \quad (18)$$

where in the second step we assumed that $W(\hat{\gamma}, \nu) \approx B(\hat{\gamma})F(\nu)$. In the considered example, one hence obtains

$$\begin{aligned} \langle I_{\nu}(\mathbf{p}) \rangle &= I(\nu_c, \bar{\mathbf{p}}) + \omega_{12} I_2(\nu_c, \bar{\mathbf{p}}) + \frac{1}{2} \left[\omega_{22} + \frac{\omega_{122}}{\bar{A}_0} \right] I_{22}(\nu_c, \bar{\mathbf{p}}) \\ &+ \frac{1}{6} \left[\omega_{222} + \frac{\omega_{1222}}{\bar{A}_0} \right] I_{222}(\nu_c, \bar{\mathbf{p}}) \\ &+ \frac{1}{24} \left[\omega_{2222} + \frac{\omega_{12222}}{\bar{A}_0} \right] I_{2222}(\nu_c, \bar{\mathbf{p}}) + \dots \end{aligned} \quad (19)$$

We therefore only need to consider the spectral functions $I_{2\dots 2}(\nu_c, \bar{\mathbf{p}})$, and the new set of moments

$$\omega_{2\dots 2}^* = \omega_{2\dots 2} + \frac{\omega_{12\dots 2}}{\bar{A}_0} = \frac{\langle A_0(\mathbf{r})[\alpha(\mathbf{r}) - \bar{\alpha}]^k \partial_{\bar{\alpha}}^k I_{\nu}(\bar{\mathbf{A}}, \bar{\alpha}) \rangle}{\bar{A}_0 \langle \partial_{\bar{\alpha}}^k I_{\nu}(\bar{\mathbf{A}}, \bar{\alpha}) \rangle}. \quad (20)$$

This shows that the amplitude $A_0(\mathbf{r})$ enters the problem as a spatial weighting function, which leaves the shape of the SED completely unchanged. With a similar redefinition one can absorb the correction term, $\propto \omega_{12} I_2(\nu_c, \bar{\mathbf{p}})$, leading to $\bar{\alpha}^* = \bar{\alpha} + \omega_{12}/\bar{A}_0$, which we have to use instead of α in the moments, $\omega_{2\dots 2}^*$. If $\omega_{12\dots 2} = 0$ [$\leftrightarrow A_0(\mathbf{r})$ and $\alpha(\mathbf{r})$ uncorrelated], one finds $\omega_{2\dots 2}^* \equiv \omega_{2\dots 2}$. This case is illustrative for constructing the moment expansion with *any* overall weighting factor that does not alter the shape of the SED.

For very narrow bands, i.e., $W(\hat{\gamma}, \nu) = B(\hat{\gamma})\delta(\nu - \nu_c)$, we have $I_{2\dots 2}(\nu_c, \bar{\mathbf{p}}) = I(\nu_c, \bar{\mathbf{p}}) \ln^k(\nu_c/\nu_0)$, such that finally

$$\begin{aligned} \langle I_{\nu}(\mathbf{p}) \rangle &= \bar{A}_0 (\nu_c/\nu_0)^{\bar{\alpha}^*} \left[1 + \frac{1}{2} \omega_{22}^* \ln^2(\nu_c/\nu_0) + \frac{1}{6} \omega_{222}^* \ln^3(\nu_c/\nu_0) \right. \\ &\left. + \frac{1}{24} \omega_{2222}^* \ln^4(\nu_c/\nu_0) + \frac{1}{120} \omega_{22222}^* \ln^5(\nu_c/\nu_0) + \dots \right]. \end{aligned} \quad (21)$$

We can now compare this expression with $I_{\nu}^* = A_0 (\nu/\nu_0)^{\alpha + \frac{1}{2}\beta \ln(\nu/\nu_0)}$ by performing a Taylor expansion in β , which yields

$$I_{\nu}^* \approx A_0 (\nu_c/\nu_0)^{\alpha} \left[1 + \frac{1}{2} \beta \ln^2(\nu_c/\nu_0) + \frac{1}{8} \beta^2 \ln^4(\nu_c/\nu_0) + \dots \right]. \quad (22)$$

This shows that $\beta \equiv \omega_{22}^*$ if with respect to the average $\langle \dots \rangle$ the distribution of $\alpha(\mathbf{r})$ is Gaussian, such that $\omega_{2\dots 2}^* = (k-1)!! (\omega_{22}^*)^{k/2}$ for even k and zero otherwise². This is quite restrictive and even in the limit of many power-law contributions hard to achieve in physical situations.

More generally, it is better to use Eq. (21) to capture the additional degrees of freedom introduced by the averaging. The free parameters can then be determined using different spectral bands. In a compact form, this can also be expressed as

$$I_{\nu}^p \approx A_0 (\nu/\nu_0)^{\alpha + \frac{1}{2}\beta \ln(\nu/\nu_0) + \frac{1}{8}\gamma \ln^2(\nu/\nu_0) + \frac{1}{24}\delta \ln^3(\nu/\nu_0) + \dots} \quad (23)$$

² Indeed $I_{\nu} = A_0 (\nu/\nu_0)^{\alpha + \frac{1}{2}\beta \ln(\nu/\nu_0)}$ is obtained when averaging power-laws, $I_{\nu} = A (\nu/\nu_0)^{\alpha}$, for a Gaussian distribution of α with variance β .

Assuming that there is no additional beam-induced frequency dependence, this expression again captures all the degrees of freedom for the superposition of power-law spectra. Here, the coefficients β , γ and δ are related to the moments $\omega_{2\dots 2}^*$. In the Rayleigh-Jeans limit, Eq. (23) is also equivalent to the commonly used parameterization, $\log T_{\text{sync}}(\nu) = \sum_{i=0}^N \alpha_i \log(\nu/\nu_0)^i$ (Pritchard & Loeb 2010) for the antenna temperature contribution due to Galactic synchrotron. However, we find that for our examples this parameterization does not perform as well as the explicit moment representation, Eq. (21), which also has a clear interpretation in terms of the physical parameters and their statistical properties.

For a Gaussian distribution of spectral indices, we can also deduce a convergence criterion for the moment expansion. If the variance, $\sigma^2 = \beta$, of the spectral indices exceeds $\beta \gtrsim 2/|\log(\nu/\nu_0)|$ a moment expansion is expected to have a small convergence radius around $\nu \simeq \nu_0$. For CMB applications, this limits $\beta \lesssim 0.3$, which seems to be fulfilled in observations (Fuskeland & et al. 2014).

3.1 Examples for sums of power-law spectra

To develop some intuition for how the moment expansion can be applied, let us consider the simple sum of two power-law spectra, $\mathbf{p} = \{A_1, \alpha_1, A_2, \alpha_2\}$. This is a low dimensional parameter case, which can be used to illustrate the limitations and strength of the moment method. It is also an extreme example for the distribution of parameters, since it clearly is far from ‘Gaussian’. In this case, we have the moments

$$\begin{aligned} \bar{A} &= A_1 + A_2 \\ \bar{\alpha} &= \frac{A_1 \alpha_1 + A_2 \alpha_2}{A_1 + A_2} = f \alpha_1 + (1-f) \alpha_2 \\ \omega_{2\dots 2} &= \frac{A_1 (\alpha_1 - \bar{\alpha})^k + A_2 (\alpha_2 - \bar{\alpha})^k}{A_1 + A_2} = [f(f-1)^k + (1-f)f^k] \Delta \alpha^k \end{aligned} \quad (24)$$

with $f = A_1/(A_1 + A_2)$ and $\Delta \alpha = \alpha_2 - \alpha_1$. It is easy to see that for $f = 1/2$ (i.e., $A_1 = A_2$) all odd moments vanish, while the even ones are $\omega_{2\dots 2} = \Delta \alpha^k / 2^k$. For general f , we have $\omega_{22} = f(1-f) \Delta \alpha^2 \geq 0$, which illustrates that for a *sum*³ of power-law spectra one always obtains a *convex* foreground spectrum.

The expressions in Eq. (24) also illustrate why thinking of the problem in terms of moments of distribution functions makes a lot of sense. The probability of finding an SED contribution with spectral index α_1 in the beam is $f = A_1/(A_1 + A_2)$ and the one for α_2 is $\bar{f} = 1 - f$. We thus have the probability distribution function (PDF), $P(\alpha) = f \delta(\alpha - \alpha_1) + \bar{f} \delta(\alpha - \alpha_2)$. Then the average spectral index is, $\bar{\alpha} = \int P(\alpha) \alpha d\alpha = f \alpha_1 + \bar{f} \alpha_2$, and similarly $\omega_{2\dots 2} = \int P(\alpha) (\alpha - \bar{\alpha})^k d\alpha = f(\alpha_1 - \bar{\alpha})^k + \bar{f}(\alpha_2 - \bar{\alpha})^k$, as also found in Eq. (24). Only that in general the PDF is a much more complicated multi-dimensional function.

We now discuss the performance of different parameterizations. For a two power-law foreground, we choose the two illustrative cases $P_1 = \{A_1, \alpha_1, A_2, \alpha_2\} = \{C/2, -2.9, C/2, -0.9\}$ and $P_2 = \{A_1, \alpha_1, A_2, \alpha_2\} = \{C/2, -0.9, C/2, 0.1\}$ with pivot frequency $\nu_0 = 300$ GHz. Here C is a normalization constant with dimension Jy/sr. The latter model has a low-frequency behavior that is similar to that of synchrotron emission in our Galaxy (Planck Collaboration et al. 2016a). Varying the fraction f is similar to changing ν_0 and we tried different fractions but the conclusions did not change

³ This is no longer true when *difference* of power-law spectra is allowed.

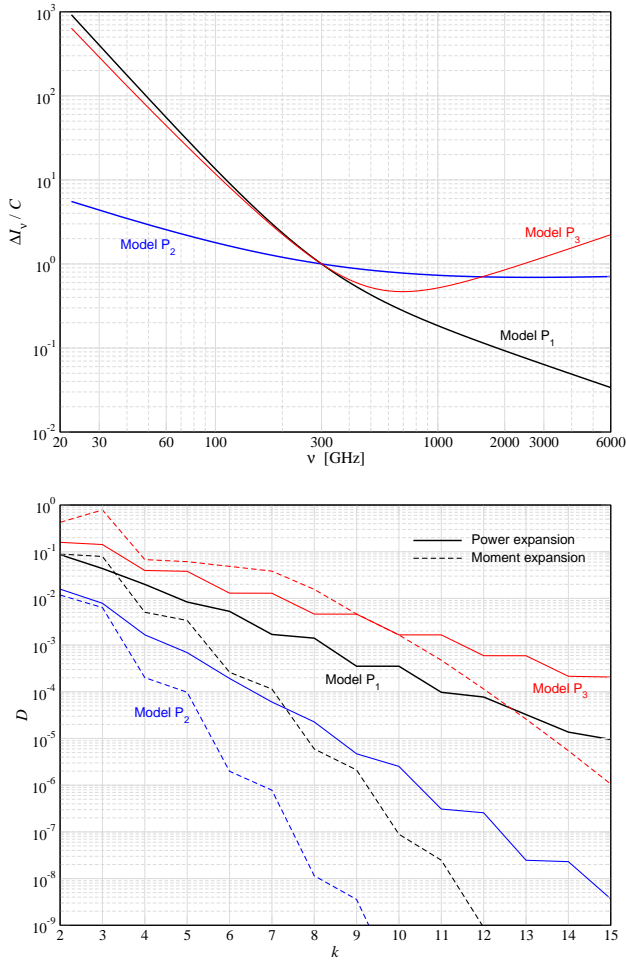


Figure 1. Intensity for a superposition of power-law spectra as a function of frequency (top panel) and the convergence of the foreground parameterizations (lower panel) for different models, P_i . The two parameterizations Eq. (21) and Eq. (23) are used with k denoting the maximal moment index ($k = 2$ for ω_{22} and β , $k = 3$ for ω_{33} and γ , etc).

much. We now simply perform a least-square fit in logarithmic coordinates⁴ (both in ν and $\Delta\nu$) for data points between 30 GHz and 6000 GHz with $\Delta\nu = 15$ GHz (this is similar to *PIXIE*), considering noise-less reconstructions for illustration of the main effects. We then compute the convergence of the two parameterizations, Eq. (21) and Eq. (23), varying the number of terms. To assess the convergence, we use the simple criterion

$$\mathcal{D} = \sqrt{\frac{1}{N_{\text{ch}}} \sum_i (I_{\nu_i}^{\text{model}} / I_{\nu_i}^{\text{input}} - 1)^2}, \quad (25)$$

which shows how large the average relative deviation of the fit from the input is. Here, $N_{\text{ch}} \simeq 400$ is the number of channels. Since the CMB signals are typically several orders of magnitude smaller than the foregrounds, values for $\mathcal{D} \simeq 10^{-4} - 10^{-3}$ need to be achieved. This is obtained with $\simeq 3-4$ moments for model P_2 and $\simeq 5-6$ moments for P_1 . Model P_3 , which has the largest change of the spectral index from low to high frequencies, shows the slowest convergence of the three considered cases.

⁴ This is not a crucial choice.

Table 1. Comparison of first few moments for power-law model P_2 . In the fitting procedure (last three columns), the number of parameters was varied.

Parameter	Theory	2 parameters	3 parameters	5 parameters
\bar{A}/C	1	1.039	0.977	1.000
$\bar{\alpha}$	-0.4	-0.169	-0.399	-0.400
ω_{22}	0.25	—	0.301	0.248
ω_{222}	0	—	—	-0.001
ω_{2222}	0.0625	—	—	0.070

In Fig. 1 we show the intensity for the sum of power-laws and the convergences of the foreground parameterizations as a function of the included terms. For the models P_1 and P_2 , the convergence is more rapid with the moment expansion, Eq. (21). We varied the combination of the spectral indices and also f but found that this parameterization generally performed better than Eq. (23). Of course by directly assuming a two power-law model the convergence is perfect already for four parameters. However, if the underlying foreground is composed of a distribution of power-laws (physically very plausible), then it should be better to use the moment expansion, Eq. (21), to allow more freedom.

To further illustrate this point, we chose a 5-power-law model, $P_3 = \{A_1, \alpha_1, \dots, A_5, \alpha_5\} = \{0.3 C, -2.9, 0.35 C, -2.1, 0.15 C, -1.5, 0.05 C, -1.1, 0.15 C, 0.9\}$. As Fig. 1 shows, even if the overall convergence is rather slow, also for this case the moment expansion performs better than Eq. (23) when including many terms ($k \gtrsim 9 - 10$). For comparison, we also performed explicit (finite number) power-law superpositions. Assuming a two power-law approximation, one is able to represent the model P_3 with 1% – 10% precision. This does not improve much for a three or four power-law approximation, highlighting that only for a sum of five power-laws (identical to the 10 parameter input model) convergence can be achieved. Also, a simple sum of power-laws exhibits significant parameter degeneracies, which make finding a solution harder. This problem is at least partially mitigated with the moment method.

We also mention that the convergence criterion, Eq. (25), is quite demanding, requiring an overall relative accuracy in the full frequency domain. When studying the performance for fixed channel sensitivity, we find a lower number of moments suffices in many cases. This is particularly true if the average spectrum is monotonic.

3.1.1 Comparison of the best-fitting parameters with the theoretical moment values

In the exercise for Fig. 1, we simply used the expressions in Eq. (21) and Eq. (23) to represent the different examples and determined the free parameters using a noiseless χ^2 minimization. For the moment expansion, Eq. (21), we can explicitly compare the obtained best-fitting values with the theoretical values, Eq. (24). In Table 1, we show this comparison for model P_2 when varying the number of parameters that are included in the fit. As alluded to in the introduction, the recovered values from the fit depend on the included number of parameters in the representation. This is because the spectral shapes of higher moments are not completely independent of the lower-moment shapes. However, by including more moments, the values converge rapidly towards the theoretical moment values. In a real analysis, this behavior can be used as a diagnostic to determine at what level the moment expansion can be truncated, a procedure that is limited by the number of channels.

4 SUPERPOSITION OF FREE-FREE SPECTRA

Another CMB foreground, present at low-frequencies, is caused by the free-free emission of thermal electrons ($T_e \approx 7000$ K). The plasma is assumed to be optically thin at the relevant frequencies ($\nu \geq 1$ GHz), so that the fundamental spectral shape is given by

$$\delta I_\nu^{\text{ff}} \approx \frac{hc\delta t \alpha}{\sqrt{6\pi^3} \theta_e^{1/2}} \sum_i Z_i^2 N_i g_{\text{ff}}(Z_i, T_e, \nu) \frac{e^{x-x_e} - 1}{e^x - 1} \quad (26)$$

with $x = x_e T_e/T_0 = h\nu/kT_0$, $\theta_e = kT_e/m_e c^2$ and where g_{ff} denotes the thermally averaged free-free Gaunt-factor. The factor, $f(x, x_e) = (e^{x-x_e} - 1)/(e^x - 1)$, accounts for stimulated emission⁵ and absorption of ambient CMB photons. This can be directly obtained from the radiative transfer equation (see Appendix A). At $\nu \lesssim 1$ THz, it can usually be neglected, but it does suppress the emission at higher frequency (see Fig. 2).

The Gaunt-factor can be obtained from detailed tables (e.g., Itoh & et. al. 2000); however, to demonstrate the important effects we shall use the simpler approximation⁶ (Draine 2011)

$$g_{\text{ff}}(Z_i, T_e, \nu) = 1 + \ln \left[1 + \left(\frac{\nu_{\text{ff}}(Z_i, T_e)}{\nu} \right)^{\frac{\sqrt{3}}{\pi}} \right] \quad (27)$$

$$\nu_{\text{ff}}(Z_i, T_e) \approx \frac{255.33 \text{ GHz}}{Z_i} \left[\frac{T_e}{10^3 \text{ K}} \right]^{3/2},$$

where Z_i is the charge of the nucleus of ion i . Neglecting the high-frequency suppression [$f(x, x_e) = 1$], we can perform a simple moment expansion⁷ in $\xi = \ln \nu_{\text{ff}}$. With $\gamma = \frac{\sqrt{3}}{\pi}$, $\eta = (\nu_{\text{ff}}/\nu)^\gamma$ and

$$\begin{aligned} \partial_\xi g_{\text{ff}} &= \frac{\gamma \eta}{(1+\eta)}, & \partial_\xi^2 g_{\text{ff}} &= \frac{\gamma^2 \eta}{(1+\eta)^2} \\ \partial_\xi^3 g_{\text{ff}} &= \frac{\gamma^3 \eta (1-\eta)}{(1+\eta)^3}, & \partial_\xi^4 g_{\text{ff}} &= \frac{\gamma^4 \eta (1-4\eta+\eta^2)}{(1+\eta)^4} \end{aligned} \quad (28)$$

for the first few derivatives of g_{ff} with respect to ξ , we then have

$$\begin{aligned} \langle I_\nu^{\text{ff}} \rangle &\approx \bar{\epsilon} g_{\text{ff}}(\bar{\nu}_{\text{ff}}, \nu) + \bar{\epsilon} \sum_{k=2}^{\infty} \frac{g_{\text{ff}}^{(k)}(\bar{\nu}_{\text{ff}}, \nu)}{k!} \omega_{\text{ff}}^{(k)} \\ g_{\text{ff}}^{(k)}(\bar{\nu}_{\text{ff}}, \nu) &= \partial_\xi^k g_{\text{ff}} \Big|_{\nu_{\text{ff}}=\bar{\nu}_{\text{ff}}} \\ \bar{\epsilon} &= \left\langle \sum_i \epsilon_i \right\rangle, \quad \ln \bar{\nu}_{\text{ff}} = \frac{\langle \sum_i \epsilon_i \ln \nu_{\text{ff}} \rangle}{\bar{\epsilon}}, \quad \omega_{\text{ff}}^{(k)} = \frac{\langle \sum_i \epsilon_i \ln^k (\nu_{\text{ff}}/\bar{\nu}_{\text{ff}}) \rangle}{\bar{\epsilon}}. \end{aligned} \quad (29)$$

Here, we introduced the parameter $\epsilon_i = hc\delta t \alpha Z_i^2 N_i / [\sqrt{6\pi^3} \theta_e^{1/2}]$, which just modifies the relative proportion of emission from different ions and can be kept as one single weight parameter. In this case, no separate moment expansion in T_e and Z_i has to be performed, since variations in Z_i can be modeled as variations in $T_e^* = T_e/Z_i^{2/3}$, or equivalently, variations of $\ln \nu_{\text{ff}}$. This moment expansion converges very rapidly; however, it fails to accurately capture the high-frequency suppression at $\nu \gtrsim 1$ THz.

⁵ Without stimulated free-free emission due to the CMB we would have $f(x, x_e) \rightarrow e^{-x_e}$, which decays much faster towards high frequencies, but is usually also neglected.

⁶ We treat these expressions as exact, given that the Gaunt factor can be precisely computed if necessary (Karzas & Latter 1961).

⁷ We also tried an expansion in T_e and $1/T_e$, but found this to converge quite slowly. Using $\ln \nu_{\text{ff}}$ as a parameter is also motivated by the fact that most of the emission is found at low frequencies where one has $g_{\text{ff}}(Z_i, T_e, \nu) \approx 1 + \frac{\sqrt{3}}{\pi} \ln \left(\frac{\nu_{\text{ff}}(Z_i, T_e)}{\nu} \right)$.

4.1 Inclusion of the high-frequency suppression

To also include the high-frequency suppression caused by ambient CMB photons, we have to add the separate dependence on Z_i . For $f(x, x_e)$, we need T_e , which can be expressed as $T_e = \kappa \nu_{\text{ff}}^{2/3} Z_i^{2/3}$, with $\kappa \approx 24.851 \text{ K GHz}^{-2/3}$. We can then define \bar{T}_e using $\ln \bar{\nu}_{\text{ff}}$ and \bar{Z} to determine $\bar{x}_e = x T_0/\bar{T}_e$. Since $\rho_e = T_0/\bar{T}_e \ll 1$, we can write

$$\begin{aligned} f(x, x_e) &= f(x, \bar{x}_e) + \sum_{k=1}^{\infty} \frac{J_k(x, \bar{x}_e)}{k!} \frac{(\rho_e - \bar{\rho}_e)^k}{\bar{\rho}_e^k} \\ J_k(x, x') &= (-1)^k \frac{x'^k e^{x-x'}}{e^x - 1}. \end{aligned} \quad (30)$$

For CMB applications, it should be sufficient to include the first two terms in $\Delta \rho_e = \rho_e - \bar{\rho}_e$. To perform the moment expansion in $\ln \nu_{\text{ff}}$ and Z_i , we define $\zeta_i = Z_i^{-2/3}$. This yields

$$\begin{aligned} \langle I_\nu^{\text{ff}} \rangle_{\text{high}} &= \langle I_\nu^{\text{ff}} \rangle f(x, \bar{x}_e) + \langle \Delta I_\nu^{\text{ff}} \rangle \\ \bar{\zeta} &= \frac{\langle \sum_i \epsilon_i Z_i^{-2/3} \rangle}{\bar{\epsilon}} \equiv \frac{1}{\bar{Z}^{2/3}}, \quad \bar{T}_e = \kappa \bar{\nu}_{\text{ff}}^{2/3} / \bar{\zeta}, \end{aligned} \quad (31)$$

where the additional correction $\langle \Delta I_\nu^{\text{ff}} \rangle$ is obtained from a moment expansion of $\Delta I_\nu^{\text{ff}} = \epsilon_i g_{\text{ff}}(\nu_{\text{ff}}, \nu) [f(x, x_e) - f(x, \bar{x}_e)]$. At zeroth order in $\Delta \rho_e$, the high frequency suppression is simply captured by multiplying $\langle I_\nu^{\text{ff}} \rangle$ in Eq. (29) with $f(x, \bar{x}_e)$. However, variations in the temperature and composition do require a more detailed treatment. The derivations are straightforward but cumbersome. Defining the frequency-dependent functions

$$G_k^{(l,m)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) = \frac{\bar{\zeta}^m \partial_\xi^l \partial_{\zeta_i}^m}{\bar{\rho}_e^k k! l! m!} [g_{\text{ff}}(\rho_e - \bar{\rho}_e)^k]_{\nu_{\text{ff}}=\bar{\nu}_{\text{ff}}, \zeta_i=\bar{\zeta}} \quad (32)$$

we can write

$$\begin{aligned} \langle \Delta I_\nu^{\text{ff}} \rangle &= \bar{\epsilon} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \left\{ \sum_{k=1}^{\infty} J_k(x, \bar{x}_e) G_k^{(l,m)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \right\} \sigma_{\text{ff}}^{(l,m)} \\ \sigma_{\text{ff}}^{(k,m)} &= \frac{\langle \sum_i \epsilon_i \ln^k (\nu_{\text{ff}}/\bar{\nu}_{\text{ff}}) (\zeta_i - \bar{\zeta})^m \rangle}{\bar{\epsilon}}. \end{aligned} \quad (33)$$

We note that $\sigma_{\text{ff}}^{(k,0)} \equiv \omega_{\text{ff}}^{(k)}$, $G_k^{(0,0)} = 0$ and $\sigma_{\text{ff}}^{(1,0)} = \sigma_{\text{ff}}^{(0,1)} = 0$ by construction. The first few functions necessary for a fourth order expansion in $\ln \nu_{\text{ff}}$ and Z_i and second order in $\bar{\rho}_e$ are given in Appendix B. At low frequencies, $f(x, x_e) \rightarrow 1$. In this case, no new spectral shapes are added through $\langle \Delta I_\nu^{\text{ff}} \rangle$ and the number of free parameters, equivalent to the number of independent spectral functions, can be reduced to only using $\bar{\epsilon}$, $\bar{\nu}_{\text{ff}}$ and $\omega_{\text{ff}}^{(k)}$.

Consequently, only the high frequency tail ($\nu \gtrsim 1$ THz) can be used to obtain direct information about \bar{Z} , which specifies the chemical composition of the medium and its ionization state, through the moments $\sigma_{\text{ff}}^{(k,l)}$. But how many coefficients $\sigma_{\text{ff}}^{(k,l)}$ can in principle be constrained independently? The functions $G_k^{(l,m)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu)$ are all just linear combinations of g_{ff} and its derivatives, $g_{\text{ff}}^{(j)}$, up to $j = 3$ (see Appendix B). These would be degenerate with simple variations of ν_{ff} , having the same spectral dependence as in the expansion, Eq. (29). However, the functions $J_k(x, \bar{x}_e)$ introduce new spectral shapes that can in principle be used to break the degeneracy. Up to fourth order in $\ln \nu_{\text{ff}}$ and Z_i ($l+m \leq 4$) and second order in $\bar{\rho}_e$ one thus finds 7 linearly independent spectral shapes. The function $J_1(x, \bar{x}_e) G_1^{(0,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu)$ fixes $\bar{\zeta}$. Since $\sigma_{\text{ff}}^{(k,0)} \equiv \omega_{\text{ff}}^{(k)}$ (3 parameters) are determined through $\langle I_\nu^{\text{ff}} \rangle f(x, \bar{x}_e)$, only 6 additional $\sigma_{\text{ff}}^{(k,m)}$ remain unspecified. This means that all $\sigma_{\text{ff}}^{(k,m)}$ can in principle be extracted from detailed measurements of the free-free spectrum at low and sufficiently high frequencies. The dimensionality of the problem reduces at low frequencies, when only

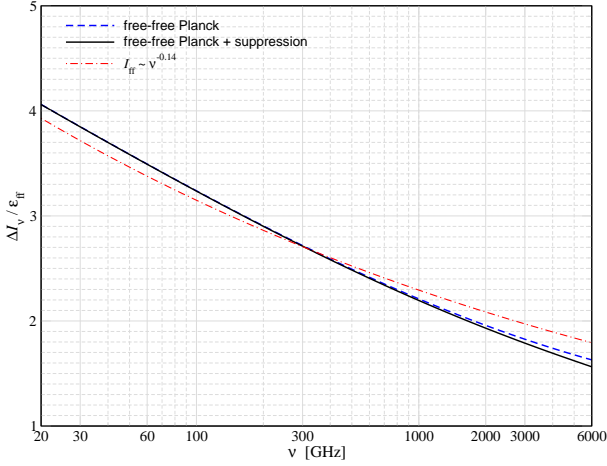


Figure 2. Free-free intensity (in units of $\epsilon_{\text{ff}} \approx 150 \text{ Jy/sr}$) for different cases.

$J_1 \approx \text{const}$ contributes, as mentioned above. In addition, instrumental noise and contributions from other foregrounds will reduce our ability to distinguish different chemical compositions and temperature distributions so that ultimately a much lower number of parameters can be independently inferred.

4.2 Simple approximations for single-temperature plasmas

For *Planck* measurements (Planck Collaboration et al. 2016a), the free-free emission is modeled using a single-temperature approximation with $\nu_{\text{ff}}(1, T_e) \approx 4727 \text{ GHz}$ for $T_e = (7000 \pm 11) \text{ K}$. The high-frequency suppression is neglected, such that one can write⁸

$$I_{\nu}^{\text{ff,P}} \approx \epsilon_{\text{ff}} g_{\text{ff}}(\nu_{\text{ff}}, \nu), \quad (34)$$

where $\epsilon_{\text{ff}} \approx 150 \text{ Jy/sr}$ (Planck Collaboration et al. 2016a). To include the high-frequency suppression, we simply multiply by $f(x, x_c)$. These two cases are shown in Fig. 2 together with a simple power-law fit $I_{\nu}^{\text{ff,P}} \propto \nu^{-0.14}$ ($\equiv T_b \propto \nu^{-2.14}$), which is close to the template used by Thorne et al. (2016). The latter only captures the overall trend but fails to reproduce the curvature of the spectrum.

A more accurate alternative representation of the free-free emission can be obtained using a power-law moment expansion, Eq. (21). We find

$$I_{\nu}^{\text{ff,2}} \approx \frac{3.229}{\nu_{100}^{0.1518}} \epsilon_{\text{ff}} \left[1 - 6.129 \times 10^{-3} \ln^2 \nu_{100} \right] \quad (35)$$

with $\nu_{100} \equiv \frac{\nu}{100 \text{ GHz}}$ to work to $\lesssim 0.5\%$ precision in the frequency range $1 \text{ GHz} \lesssim \nu \lesssim 6 \text{ THz}$. For better than $\approx 0.06\%$ precision we obtain the fourth order power-law moment expansion

$$I_{\nu}^{\text{ff,4}} \approx \frac{3.235}{\nu_{100}^{0.1527}} \epsilon_{\text{ff}} \left[1 - 7.116 \times 10^{-3} \ln^2 \nu_{100} + 8.701 \times 10^{-5} \ln^3 \nu_{100} + 6.157 \times 10^{-5} \ln^4 \nu_{100} \right]. \quad (36)$$

Both expressions include the high-frequency suppression effect and were obtained by noiseless fits using power-law moments as free parameters. This shows that the free-free emission for a single-temperature component can be fully represented using a superposition of power-laws. Thus, in the simplest CMB foreground scenarios it does not seem necessary to add any extra parameters when

Table 2. Comparison of first few moments for the free-free model P_1 . In the fitting procedure (last three columns), the number of parameters was varied. For all cases we set $Z_i = 1$.

Parameter	Theory	2 parameters	3 parameters	5 parameters
$\bar{\epsilon}/\epsilon_{\text{ff}}$	1	0.8991	1.001	0.999
$\bar{\nu}_{\text{ff}}$ [GHz]	653.18	1600.4	651.88	658.32
$\omega_{\text{ff}}^{(2)}$	3.9176	–	3.659	3.835
$\omega_{\text{ff}}^{(3)}$	0	–	–	–0.276
$\omega_{\text{ff}}^{(4)}$	15.348	–	–	11.054

considering both synchrotron and free-free emission with spatial variations. On the full sky, they apparently can be treated as one component with slightly more complex frequency structure. The spatial-spectral variations are then captured by the spatial variations of the moment values. In CMB analysis, the main motivation for modeling free-free and synchrotron independently is thus really driven by external priors but not because the data distinguishes them easily. However, this is not expected to be sufficient in more general cases, which could be relevant to non-CMB applications.

4.3 Free-free models with multiple components

In general, the Galactic free-free emission has contributions from multiple components. In this case, a single-temperature representation of the spectrum will no longer work and the moment expansions, Eq. (29) and (31), are expected to perform better, while also capturing the more general cases with distributions of temperatures and chemical composition.

To illustrate the effects, we start with a two-temperature case, $P_1 = \{\bar{\epsilon}_1, T_{e,1}, \bar{\epsilon}_2, T_{e,2}\} = \{0.5 \epsilon_{\text{ff}}, 7000 \text{ K}, 0.5 \epsilon_{\text{ff}}, 500 \text{ K}\}$, representing a mix of hot and cold components. For now we assume $Z_i = 1$. The two components are generally not in the same region but could be situated at different locations along the line of sight. The relatively large fractional contribution from the low temperature component is physically slightly difficult to achieve, since at low temperatures the ionization degree (of hydrogen) is low. However, the effective weight scales as $\epsilon_i \approx 1/\sqrt{T_e}$ and also there is typically more gas in the colder phase, so that we still consider this model for illustrative purposes. Model P_1 is illustrated in Fig. 3. Approximating it with a single temperature spectrum⁹, $I_{\nu}^{\text{ff,1}} \approx \bar{\epsilon} g_{\text{ff}}(\bar{\nu}_{\text{ff}}, \nu) f(x, x_c)$, we find $\bar{\nu}_{\text{ff}} \approx 1600.4 \text{ GHz}$ ($T_e \approx 3400 \text{ K}$) and $\bar{\epsilon} \approx 0.8992 \epsilon_{\text{ff}}$ (red line in Fig. 3). This representation works to $\lesssim 6\%$ precision in the considered frequency range (see Fig. 3). Explicitly using the free-free moment expansion, Eq. (29) and (31), with $\bar{\epsilon}$, $\bar{\nu}_{\text{ff}}$ and $\omega_{\text{ff}}^{(2)}$ as free parameters, we obtain the values given in Table 2. This approximation already represents the free-free spectrum to $\lesssim 0.3\%$ precision at $1 \text{ GHz} \lesssim \nu \lesssim 6 \text{ THz}$. Adding also $\omega_{\text{ff}}^{(3)}$ and $\omega_{\text{ff}}^{(4)}$, we obtain the values given in the last column of Table 2. This approximation already works to better than 0.005% precision. While in the latter case, the spectrum is described by one additional parameter than the input model (5 versus 4 parameters), this approximation is more general, not implicitly assuming a two-temperature model and allowing one to capture the effect of more general spatial variations in the temperature of the electrons. This again demonstrates the potential of the moment expansion.

⁸ The Rayleigh-Jeans temperature is obtained as $T_b = c^2 I_{\nu}^{\text{ff,P}} / (2k\nu^2)$.

⁹ We use 100 log-spaced frequency bins in the range $1 \text{ GHz} \lesssim \nu \lesssim 6 \text{ THz}$.

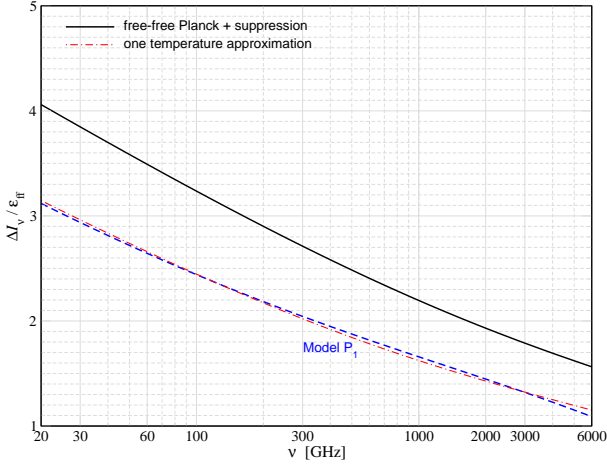


Figure 3. Free-free intensity (in units of ϵ_{ff}) for different models. For comparison we also show the $T_e = 7000$ K *Planck* model, Eq. (34), with the high-frequency suppression added. The change in the normalization is absorbed when determining ϵ_{ff} , but changes in the spectral shape remain. The red line represents a single temperature approximation. The moment approximations are not shown as they basically coincide with the input model.

We can again compare the obtained best-fitting values, with the theoretical predictions for the moments. For a general two-component case, $P = \{C_{\text{ff}}(1 - f), T_{e,1}, C_{\text{ff}}f, T_{e,2}\}$, we have

$$\begin{aligned} \bar{\epsilon} &= C_{\text{ff}} \\ \ln \bar{\nu}_{\text{ff}} &= (1 - f) \ln \nu_{\text{ff}}(T_{e,1}) + f \ln \nu_{\text{ff}}(T_{e,2}) \\ \omega_{\text{ff}}^{(k)} &= (1 - f) \ln^k[\nu_{\text{ff}}(T_{e,1})/\bar{\nu}_{\text{ff}}] + f \ln^k[\nu_{\text{ff}}(T_{e,2})/\bar{\nu}_{\text{ff}}] \\ \bar{\rho}_e &= T_0/[\kappa \bar{\nu}_{\text{ff}}^{2/3}] \approx 1.097 \times 10^{-3} \left[\frac{\bar{\nu}_{\text{ff}}}{10^3 \text{ GHz}} \right]^{-2/3} \\ \sigma_{\text{ff}}^{(k,0)} &= \omega_{\text{ff}}^{(k)}, \quad \sigma_{\text{ff}}^{(k,m)} = 0 \quad \text{for } m > 0 \end{aligned} \quad (37)$$

For model P_1 , we find $\bar{\nu}_{\text{ff}} = 653.18$ GHz, $\bar{\rho}_e = 1.4571 \times 10^{-3}$, $\omega_{\text{ff}}^{(2)} = 3.9176$, $\omega_{\text{ff}}^{(4)} = 15.348$, and $\omega_{\text{ff}}^{(6)} = 60.126$ (odd moments vanish for $f = 0.5$). Comparing this with the values obtained from different model in the fit, we again see that with increasing order, the recovered values approach the theoretical one quite fast (see Table 2). This behavior is similar to that of the power-law moment expansion (Table 1).

We also illustrate how well the theoretical moments can be used to model the shape of the spectrum. Including only $\bar{\nu}_{\text{ff}}$ and $\bar{\rho}_e$ (\equiv single-temperature model) in the moment expansion represents the low-frequency free-free spectrum very well, while underestimating the high-frequency spectrum by $\approx 7\%$ at $\nu \approx 1$ THz. Including $\omega_{\text{ff}}^{(2)} = 3.9176$ as additional parameter, the discrepancy decreases to $\lesssim 0.6\%$ at $1 \text{ GHz} \lesssim \nu \lesssim 3 \text{ THz}$. Adding the next non-vanishing moment, $\omega_{\text{ff}}^{(4)} = 15.348$, the moment expansion captures the full spectrum at the $\lesssim 0.02\%$ level in the same range.

When also varying f for the two-temperature case, the approximation also requires inclusion of $\omega_{\text{ff}}^{(3)} \neq 0$, which no longer vanishes. For the chosen extreme difference in temperatures, this representation is only accurate at the $\approx 0.1\%$ level. This can be improved by adding higher order moments. However, with much smaller spread in the temperature even a second order expansion is found to converge very rapidly. Our study indicates that the free-free moment expansion allows us to accurately represent the free-free emission from more general temperature distributions with a minimal number of assumptions.

4.4 Variations in the chemical composition and ionization

As alluded to above, when adding variations in the chemical composition, the low and high frequency spectra become more independent. This may be relevant to detailed modeling of the continuum X-ray component from clusters (Ponente et al. 2011), where the gas is found in highly-ionized states with significant contributions from metals. For the modeling of emission in our galaxy, at temperatures $T \lesssim 10^4$ K the dominant contributions are from singly-ionized hydrogen and helium, which both have $Z = 1$. Thus, variations in Z are not expected to enter at a significant level to the spectral shape.

When varying the distribution of temperatures, e.g., adding more components with different weights, we find that even a small addition of low-temperature plasma causes a single-temperature approximation to depart from the free-free spectrum at the 1% level. In contrast, adding high temperature components the spectrum remains well described by a single temperature approximation. This can be understood since for high temperatures one has

$$g_{\text{ff}}(Z_i, T_e, \nu) \approx 1 + \ln \left(\frac{\nu_{\text{ff}}(Z_i, T_e)}{\nu} \right)^{\frac{\sqrt{3}}{\pi}} = 1 + \frac{\sqrt{3}}{\pi} [\ln \nu_{\text{ff}}(Z_i, T_e) - \ln \nu]$$

and $f(x, x_e) \approx 1$ at CMB frequencies. This means that adding multiple high temperature components, in the CMB regime [$\nu \lesssim \nu_{\text{ff}}(Z_i, T_e) \approx O(10^3 \text{ GHz})$] only leads to an overall change in the zero offset [$\approx 1 + \frac{\sqrt{3}}{\pi} \ln \nu_{\text{ff}}(Z_i, T_e)$] but no change in the spectral dependence, $\approx -\frac{\sqrt{3}}{\pi} \ln \nu$. Thus, a sufficient approximation can indeed be obtained with a single-temperature model. The inferred effective temperature is close to

$$\ln \bar{T}_e^* = \frac{\langle \sum \epsilon_i \ln T_e^* \partial_{\ln T_e^*} \ln \nu_{\text{ff}}(T_e^*) \rangle}{\langle \sum \epsilon_i \partial_{\ln T_e^*} \ln \nu_{\text{ff}}(T_e^*) \rangle} \equiv \frac{\langle \sum \epsilon_i \ln T_e^* \rangle}{\bar{\epsilon}},$$

with $T_e^* = T_e \zeta_i$. This expression explicitly shows how the chemical composition enters into the problem and that the weighted average has to be carried out over $\epsilon_i \ln(T_e \zeta_i)$. However, at high frequencies ($\nu \gtrsim 1 \text{ THz}$), corrections can become noticeable at $\approx 0.1\% - 1\%$ level even in this situation.

This highlights that for CMB applications the shape of the free-free spectrum is quite insensitive to the detailed temperature structure and composition. Thus, simpler parameterizations such as a power-law moment expansion should be applicable, in particular for CMB applications. We also highlight that in the presence of instrumental noise and other foregrounds (which we neglected here), physically distinct components are expected to become spectrally indistinguishable. This can lead to significant model-dependence (in form of prior choices), which can also cause biases.

5 SUPERPOSITION OF GRAY-BODY SPECTRA

We can also apply the moment method to spatially varying gray-body spectra, with intensity $I_\nu(A_0, T) = A_0 \nu^3 / (e^{h\nu/kT} - 1)$. This case is not directly relevant to CMB applications but illustrates how to use the moment expansion. It is convenient to perform the expansion in terms of $\beta^{10} = 1/T$, since in this case a closed form for the derivatives of the distribution function with good convergence properties can be given (see Appendix C). The factor A_0 is again a

¹⁰ We will discuss this choice in Sect. 5.2.1.

simple weight factor which does not affect the spectral shape. For the first few derivatives with respect to β , we find

$$\begin{aligned}\beta \partial_\beta I_\nu(A_0, T) &= -I_\nu(A_0, T) \frac{x e^x}{e^x - 1} \\ \beta^2 \partial_\beta^2 I_\nu(A_0, T) &= I_\nu(A_0, T) \frac{x e^x}{e^x - 1} x \coth(x/2) \\ \beta^3 \partial_\beta^3 I_\nu(A_0, T) &= -I_\nu(A_0, T) \frac{x e^x}{e^x - 1} x^2 \frac{\cosh(x) + 2}{\cosh(x) - 1} \\ \beta^4 \partial_\beta^4 I_\nu(A_0, T) &= I_\nu(A_0, T) \frac{x e^x}{e^x - 1} \frac{x^3}{2} \frac{\cosh(x) + 5}{\sinh^2(x/2)} \coth(x/2) \\ \beta^5 \partial_\beta^5 I_\nu(A_0, T) &= -I_\nu(A_0, T) \frac{x e^x}{e^x - 1} \frac{x^4}{8} \frac{33 + 26 \cosh(x) + \cosh(2x)}{\sinh^4(x/2)}\end{aligned}\quad (38)$$

with $x = h\nu/kT$. The same arguments as in the previous section apply, so that Eq. (19) carries over with $\alpha \rightarrow 1/T$ in the moments. For narrow spectral bands, we then find the moment representation

$$\begin{aligned}\langle I_\nu \rangle &= \frac{\bar{A}_0 \nu^3}{e^x - 1} \left\{ 1 + \frac{1}{2} \omega_{22}^g Y_2(x) + \frac{1}{6} \omega_{222}^g Y_3(x) \right. \\ &\quad \left. + \frac{1}{24} \omega_{2222}^g Y_4(x) + \frac{1}{120} \omega_{22222}^g Y_5(x) + \dots \right\} \\ \frac{1}{\bar{T}} &= \frac{\langle A_0(\mathbf{r})/T(\mathbf{r}) \rangle}{\bar{A}_0} \\ \omega_{2\dots 2}^g &= \bar{T}^k \omega_{2\dots 2}^* = \frac{\langle A_0(\mathbf{r}) [\bar{T}/T(\mathbf{r}) - 1]^k \rangle}{\bar{A}_0},\end{aligned}\quad (39)$$

where now $x = h\nu/k\bar{T}$ and we defined $Y_k = [(-\beta)^k \partial_\beta^k I_\nu]/I_\nu$ (e.g., $Y_1(x) = x e^x/(e^x - 1)$ and $Y_2(x) = Y_1(x) x \coth(x/2)$, etc.). The first correction term is similar to a Compton y -type distortion (Zeldovich & Sunyaev 1969), with effective y -parameter $y = \frac{1}{2} \omega_{22}^g$, but generalized to gray-body spectra. The effective moments, $\omega_{2\dots 2}^g$, are obtained using Eq. (20), with $\alpha \rightarrow 1/T$.

Below we illustrate how to apply the gray-body moment expansion for several cases. However, Eq. (39) can also be used to model the correction to the average CMB spectrum caused by the temperature fluctuations of the CMB itself. In this case $A_0 = (2h/c^2)$, so that the moments simply characterize the fractional temperature variations around the average CMB blackbody without any albedo variation. Since $\Delta T/T \approx 10^{-5} - 10^{-4}$, the overall effect is captured by adding only a y -parameter, $y \approx 10^{-9}$ (Chluba & Sunyaev 2004; Chluba et al. 2012a), which can be neglected. The variation caused by the CMB dipole gives rise to $y = (2.525 \pm 0.012) \times 10^{-7}$ (Chluba & Sunyaev 2004; Chluba 2016), which can be accurately removed.

5.1 Convergence at high frequencies

The functions Y_k are all positive. At low frequencies, they scale like $Y_k \approx 1$ [so that $\langle I_\nu(\mathbf{p}^*) \rangle \approx \text{const} \times \nu^2$], while for $x \gg 1$, they all approach $Y_k \approx x^k$. Higher order moments thus become distinguishable mainly at high frequencies. This leads to rather slow convergence of the expansion in the non-perturbative regime, implying that without additional prior knowledge about the relations between the moments, alternative treatments, e.g., directly based on the temperature distribution function, may be needed. This problem reappears for the superposition of spatially varying thermal dust emission, and is also known in connection with relativistic corrections to the Sunyaev-Zeldovich effect (Sazonov & Sunyaev 1998; Itoh et al. 1998; Nozawa & et al. 2006; Chluba et al. 2012b). This also implies that it becomes difficult to use the high-frequency spectrum to improve the modeling of the low-frequency spectrum.

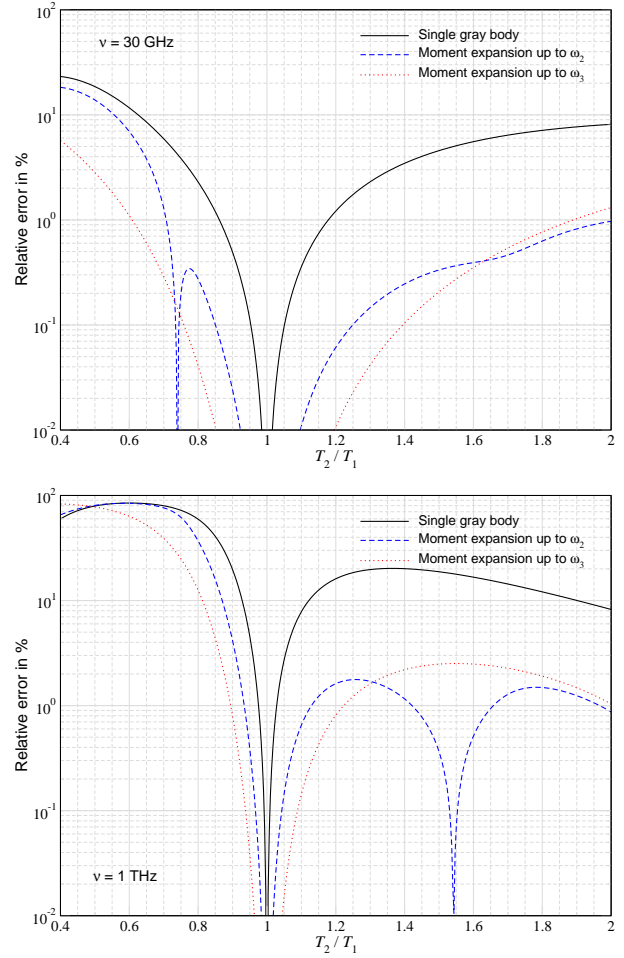


Figure 4. Relative departure from the input model for the sum of two gray-body spectra with $A_1 = 0.3A$, $T_1 = 3$ K, $A_2 = 0.7A$ and varying T_2 at two representative frequencies.

5.2 Examples for sums of two gray-body spectra

Let us consider the simple case of two gray-body spectra, with parameters A_i, T_i . We then have

$$\begin{aligned}\bar{A} &= A_1 + A_2, \quad \bar{\beta} = \frac{f}{T_1} + \frac{1-f}{T_2} \\ \omega_{2\dots 2} &= \left[f(f-1)^k + (1-f)f^k \right] \frac{(T_1 - T_2)^k}{[(1-f)T_1 + fT_2]^k}\end{aligned}\quad (40)$$

with $f = A_1/(A_1 + A_2)$. For $A_1 = A_2 = A$ one has the moments $\omega_{2\dots 2} = (T_1 - T_2)^k/(T_1 + T_2)^k$ for even k and zero otherwise. At $x \gg 1$, with $Y_k(x) \approx x^k$ as expected we thus find $\langle I_\nu \rangle \approx 2A \nu^3 e^{-x} \cosh[x(T_1 - T_2)/(T_1 + T_2)] \equiv A \nu^3 [e^{-x_1} + e^{-x_2}]$ for $x_i = h\nu/kT_i$. To achieve full convergence at high frequencies, an infinite number of moments is required, while at low frequencies, the convergence is much more rapid. This demonstrates that in the non-perturbative regime ($\Leftrightarrow T_1$ and T_2 differ strongly or x become very large) it is difficult to use the high-frequency spectrum to constrain the low-frequency model (see also Sect. 5.1).

We start with an example in the perturbative regime, choosing $P_1 = \{A_1, T_1, A_2, T_2\} = \{0.3A, 3 \text{ K}, 0.7A, 4 \text{ K}\}$. A single gray-body spectrum with $T^{\text{fit}} = 3.8336$ K and amplitude $A^{\text{fit}} = 0.9329 A$ represents the average SED at the level of $\lesssim 3\% - 4\%$ at $\nu \lesssim 550$ GHz. The given best-fitting values are not far away from what is ob-

tained with the moment expressions in Eq. (40), which imply $\bar{T} = 3.6363$ K and amplitude¹¹ $\bar{A} = A$. This approximation fails, however, at the level of $\gtrsim 10\%$ at $\nu \gtrsim 550$ GHz, exceeding $\approx 50\%$ at about 2 THz. Adding the first moment, ω_2 , as a free parameter, the spectrum is represented at a level better than $\approx 1.5\%$ at all frequencies $\nu \lesssim 1$ THz and to better than $\approx 0.25\%$ at all frequencies $\nu \lesssim 600$ GHz. Although not perfect, this level of precision is reached with *fewer* parameters than the input model, without assuming a specific number of components, which again illustrates the potential of the method.

Here we immediately mention a caveat of the moment representation for gray-body spectra. The approximation was obtained by fitting the average SED assuming zero noise. The best-fitting parameters were $R_1 = \{A^{\text{fit}}, T^{\text{fit}}, \omega_2^{\text{fit}}\} = \{0.9932 A, 3.6667 \text{ K}, 0.01341\}$. Comparing this to what is expected from the expressions in Eq. (40), $M_1 = \{A, 3.6363 \text{ K}, 0.01928\}$, shows that the best-fitting parameters do not match perfectly. This is most obvious for the value of ω_2 , which deviates by $\approx 68\%$. We have seen a similar behavior for the superposition of power-laws and free-free spectra (Sect. 3.1.1 and 4.3). The source of this difference is that the spectral function $\propto \bar{A}_0 \nu^3 Y_2(x)/(e^x - 1)$ is not linearly independent of $\propto \bar{A}_0 \nu^3/(e^x - 1)$. The recovered values of a truncated moment expansion thus are expected to differ from the theoretical values and only approach these when a sufficiently large number of moments is evaluated. This can be used as a *diagnostic* for how many moments may be needed to describe a given spectral shape in real experimental situations, as closer to convergence the moment values will not vary anymore when adding more components. A cure for this issue can be obtained using orthogonalization schemes, as we briefly discuss below (Sect. 5.3). However, in this case the interpretation of the obtained parameters in terms of the underlying temperature moments $\omega_{2...2}$, becomes more cumbersome.

While we found the moment representation to work very well for the model chosen above, let us investigate how the convergence changes with increasing temperature difference and varied fractional contribution. For this, we determined the best-fitting parameters for a given input model using varying representations in the analysis. In Fig. 4, we show the results for $A_1 = 0.3A$, $T_1 = 3$ K and $A_2 = 0.7A$ varying T_2 . We illustrate the relative difference with respect to the input spectrum at $\nu = 30$ GHz, representative of the behavior in the Rayleigh-Jeans part of the spectrum, and $\nu = 1$ THz, for the Wien part of the spectrum. For $T_1 = T_2$, all approximations naturally show no departure from the input spectrum. The convergence radius around $T_1 \approx T_2$ is larger at low frequencies than at high frequencies, owing to the asymptotic behavior of the moment expansion in this regime. For the chosen example, the convergence is faster at low than at high frequencies. Similarly, we find that for $T_2 > T_1$, the departure is smaller than for $T_2 < T_1$. In the latter case one thus needs more moments to reach the same level of precision. This observation depends on the chosen value for f , which up-weights T_2 here. A small contribution of a colder component in the Rayleigh-Jeans part of the average spectrum thus seems easier to incorporate than a small contribution of a hotter component in the Wien-tail. We find that the curves become relatively symmetric around $T_1 = T_2$ when fixing the height of the maxima for each component to be equal. All these cases are simply to illustrate the behavior of the moment expansion for gray-body spectra.

5.2.1 Best-fitting single gray-body parameters

We already saw that the best-fitting values for the parameters in a moment expansion depend on the number of terms that are included. The same is true when we simply approximate the superposition of gray-body spectra with a single gray-body spectrum. The best-fitting temperature and amplitude are not simply given by the average parameters $\langle A_i \rangle$ or $\langle T_i \rangle$. In our moment expansion we used $\beta_i = T_i^{-1}$ as temperature parameter, and again, the best-fitting value is usually not determined by $T^* = \langle \beta_i \rangle^{-1}$ or $T^* = \langle A_i \rangle / \langle A_i \beta_i \rangle$ [see Eq. (39)]. How can we estimate the best-fitting parameters for a single gray-body approximation? The answer is: by using moments of the frequency-integrated spectrum.

Like for a blackbody spectrum, we can define the number and energy density of the fundamental spectral shape. These are simply $N_\gamma = A_i c_N T_i^3$ and $\rho_\gamma = A_i c_\rho T_i^4$, where the values of two constants are not important. We then have two equations $A^* (T^*)^3 = \langle A_i T_i^3 \rangle$ and $A^* (T^*)^4 = \langle A_i T_i^4 \rangle$ when assuming a single gray-body approximation for the superposition. This yields

$$A^* = \langle A_i T_i^3 \rangle^4 / \langle A_i T_i^4 \rangle^3, \quad T^* = \langle A_i T_i^4 \rangle / \langle A_i T_i^3 \rangle. \quad (41)$$

For our example, $P_1 = \{0.3A, 3 \text{ K}, 0.7A, 4 \text{ K}\}$, we then find $A^* = 0.9292 A$ and $T^* = 3.8469$, which is in very good agreement with the values of a single gray-body approximation recovered from an explicit fit ($T^{\text{fit}} = 3.8336$ K and amplitude $A^{\text{fit}} = 0.9329 A$). As mentioned above, this departs notably from the moment values ($\bar{T} = 3.6363$ K and $\bar{A} = A$) that are obtained using Eq. (39).

The difference between the best-fitting values and those obtained with Eq. (39) can become much more dramatic. Choosing $P_2 = \{0.5A, T_0, 0.7A, 1.5 T_0, 0.4A, T_0/1.5, 0.2A, T_0/3.5\}$ with $T_0 = 2.726$ K, we find $\bar{A} = 1.8 A$ and $\bar{T} = 2.165$ K using the moment definitions, Eq. (39). From Eq. (41), we have $A^* = 1.1329 A$ and $T^* = 3.765$ K and a fit gives $A^{\text{fit}} = 1.1245 A$ and $T^{\text{fit}} = 3.743$ K, which clearly illustrates the point.¹² This example shows that the distribution of temperatures becomes so wide that a perturbative expansion requires many moments to converge. The weighting ($A_i T_i^3$ instead of A_i) and choice of variables (T_i instead of $1/T_i$) furthermore can lead to improved convergence properties (higher level of independence between moments). In spite of all this, the moment representation, Eq. (39), can still be used to approximate the average SED in a meaningful way, even without explicitly rearranging the moments and spectral functions. This is because the moment expansion naturally captures the dominant new degrees of freedom (higher order derivatives) of each moment order. A different weighting scheme or change of variables (which ultimately is equivalent to a different weighting scheme) does not alter this statement, but only changes the interpretation of the moment values in terms of the underlying temperature distribution function.

For completeness, we also give the moment expansion using weight $B = A T^3$ and $p = T$. In this case, the fundamental SED reads $I_\nu(B, T) = B (\bar{T}/T)^3 x^3 / (e^{x\bar{T}/T} - 1)$, with the derivatives

$$\begin{aligned} T \partial_T I_\nu(B, T) &= I_\nu(B, T) [Y_1(x) - 3] \\ T^2 \partial_T^2 I_\nu(B, T) &= I_\nu(B, T) [Y_2(x) - 8Y_1(x) + 12] \\ T^3 \partial_T^3 I_\nu(B, T) &= I_\nu(B, T) [Y_3(x) - 15Y_2(x) + 60Y_1(x) - 60] \\ T^4 \partial_T^4 I_\nu(B, T) &= I_\nu(B, T) [Y_4(x) - 24Y_3(x) + 180Y_2(x) \\ &\quad - 480Y_1(x) + 360]. \end{aligned} \quad (42)$$

¹² In the presence of noise, other foreground components, and differences in the frequency coverage, the answers can furthermore differ significantly.

¹¹ We will address the cause of the small difference in Sect. 5.2.1.

Defining the functions $Z_k = [T^k \partial_T^k I_\nu(B, T)]/I_\nu(B, T)$, we find the alternative moment expansion

$$\langle I_\nu \rangle = \frac{\bar{B} x^3}{e^x - 1} \left\{ 1 + \frac{1}{2} \kappa_{22}^g Z_2(x) + \frac{1}{6} \kappa_{222}^g Z_3(x) + \frac{1}{24} \kappa_{2222}^g Z_4(x) + \dots \right\} \quad (43)$$

$$\bar{B} = \langle B(\mathbf{r}) \rangle \equiv \langle A(\mathbf{r}) T^3(\mathbf{r}) \rangle, \quad \bar{T} = \frac{\langle B(\mathbf{r}) T(\mathbf{r}) \rangle}{\langle B(\mathbf{r}) \rangle} \equiv \frac{\langle A(\mathbf{r}) T^4(\mathbf{r}) \rangle}{\langle A(\mathbf{r}) T^3(\mathbf{r}) \rangle}$$

$$\kappa_{2...2}^g = \frac{\langle B(\mathbf{r}) [T(\mathbf{r})/\bar{T} - 1]^k \rangle}{\bar{B}} \equiv \frac{\langle A(\mathbf{r}) T^3(\mathbf{r}) [T(\mathbf{r})/\bar{T} - 1]^k \rangle}{\langle A(\mathbf{r}) T^3(\mathbf{r}) \rangle}.$$

This expression is close to optimal regarding the independence of individual moment coefficients (see Sect. 5.3).

5.3 Orthonormal basis functions for gray-body spectra

Since the functions $Y_k \approx 1$ at low frequencies, all moments become degenerate in this regime. It is thus useful to think about a new set of functions that minimize the correlations between different moments. This can be achieved using a Gram-Schmidt orthogonalization scheme, ordering the functions with respect to their moment order, which provides a perturbative ordering. By defining $x = h\nu/k\bar{T}$, we can remove the explicit dependence on \bar{T} . Before the orthonormalization we then have the basic spectral functions

$$F_k(x) = \frac{x^3}{e^x - 1} Y_k(x) \quad (44)$$

with $Y_0(x) = 1$. These functions are shown in Fig. 5 for $k \leq 4$ and fully determine the moment expansion for the superposition of gray-body spectra. With increasing k , the maximum of F_k moves towards higher frequencies and smaller contributions are visible at low frequencies. We now define the dot-product

$$\langle F_k(x) F_m(x) \rangle = \int_0^\infty F_k(x) F_m(x) dx. \quad (45)$$

This could be modified by adjusting the range of integration to match some experimental setting, but we do not explore this possibility any further. For the norms of the first few functions we then have $|F_0| = 2.5447$, $|F_1| = 9.4417$, $|F_2| = 44.469$, $|F_3| = 253.91$, $|F_4| = 1703.81$. Within a limited space (including fourth order moment terms only) we then find the orthonormal basis

$$\begin{aligned} H_0(x) &= (0.39297, 0, 0, 0, 0)^T \cdot \mathbf{F}(x) \\ H_1(x) &= (-1.11687, 0.31911, 0, 0, 0)^T \cdot \mathbf{F}(x) \\ H_2(x) &= (2.25944, -1.31301, 0.16413, 0, 0)^T \cdot \mathbf{F}(x) \\ H_3(x) &= (-3.88531, 3.38761, -0.85407, 0.06326, 0)^T \cdot \mathbf{F}(x) \\ H_4(x) &= (6.07788, -7.03848, 2.65812, -0.39518, 0.01976)^T \cdot \mathbf{F}(x), \end{aligned} \quad (46)$$

where we defined $\mathbf{F} = (F_0, F_1, F_2, F_3, F_4)$. These functions are also illustrated in Fig. 5 and show an increasing number of oscillations with k at quasi-constant overall amplitude. This nicely illustrates why the moment expansion can exhibit asymptotic convergence, with both high and low frequencies being affected.

We can also compare Eq. (46) with Eq. (43). From the coefficients, we have $|F_0| H_0(x) = F_0(x)$ and $k!|F_0| H_k(x) \approx F_0(x) Z_k(x)$. This shows that the moment expansion in Eq. (43), with modified weighting, $B = A T^3$, results in a quasi-orthogonal (\leftrightarrow optimal) re-summation of terms. We will discuss a similar approach for the superposition of modified blackbody spectra (Sect. 6.2).

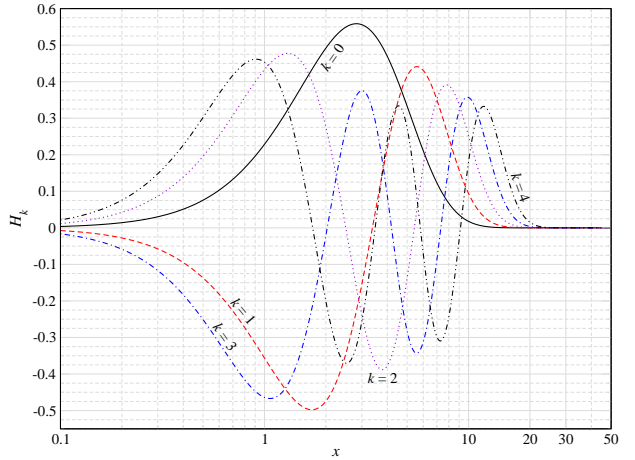
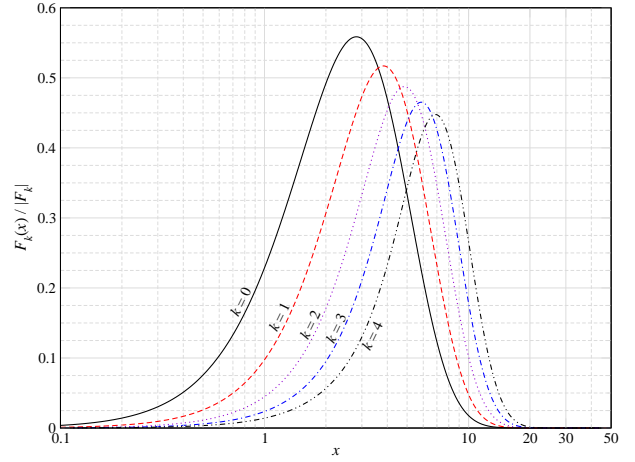


Figure 5. Gray-body spectral functions, $F_k(x)$ and $H_k(x)$ for $k \leq 4$.

With the basis functions, $H_k(x)$, we can also write

$$\langle I_\nu \rangle = A_0^* \left\{ H_0(x) + \sum_{k=2}^{k_{\max}} \sigma_k^g H_k(x) \right\} \quad (47)$$

for the superposition of gray-body spectra, with overall amplitude $A_0^* = |F_0| A_0 (k\bar{T}/h)^3$. In an idealized case, the coefficients σ_k^g are uncorrelated and directly determine the underlying moments $\omega_{2...2}^g$ up to a given order. However, the presence of other signals (which we did not project out) and truncation in the frequency domain or bandpass effects can re-introduce correlations. This is not necessarily a severe limitation and applying the idealized expansion renders a comparison of the results from different experiments easier.

While for a superposition of gray-body spectra a Gram-Schmidt orthogonalization leads to an efficient re-summation of higher order moments, more generally this approach does not improve matters as much. In particular, generally it will be impossible to define a basis that is independent of the leading order average spectral parameter values. This was avoided for the gray-body spectra by using the transformation $\nu \rightarrow x$. However, for a superposition of power-law spectra this already is not possible and the set of basis functions would depend on the average spectral index, $\bar{\alpha}$. This limits the benefits of orthogonalization schemes.

6 SUPERPOSITION OF THERMAL DUST SPECTRA

We now investigate how a superposition of thermal dust spectra can be modeled. Here, we first develop a general expansion in terms of Taylor moments. Our examples will then be guided by models that may be relevant to future CMB analyses. We start with the commonly used parameterization for a single-temperature dust spectrum, $I_\nu(A_0, \alpha, T) = A_0 (\nu/\nu_0)^\alpha \nu^3 / (e^{h\nu/kT} - 1)$, which is referred to as a modified blackbody spectrum. This model for the fundamental spectral energy distribution can be motivated by considering the physical properties of dust grains in the vicinity of stars (Draine & Lazarian 1998; Finkbeiner et al. 1999; Draine 2003; Planck Collaboration et al. 2014a). Physically speaking, this shape is certainly not expected to be exact, as many different types of dust grains can contribute. Especially at high frequencies, the spectral shape can be much more rich (e.g., Draine 2003). Nevertheless, as a first step we can extend the widely used dust models using the moment expansion to incorporate the effect of spatial variations parametrically.

The superposition of modified blackbody spectra shares many of the features found for the superposition of gray-body spectra discussed in Sect. 5. In particular, at high frequencies the convergence of a moment expansion is again expected to be slow in the non-perturbative regime. This highlights the fact that the high-frequency spectrum of the dust SED is quite sensitive to the detailed shape of the underlying parameter distribution functions and thus hard to use to constrain the low-frequency spectrum.

To obtain the moment expansion, we will use the parameters $\mathbf{p} = \{A_0, \alpha, \beta = 1/T\}$. Alternative parameterizations and weighting schemes are briefly discussed below (Sect. 6.2). In the model, $I_\nu(A_0, \alpha, T)$, no second derivatives with respect to A_0 appear, however, many additional spectral shapes are created due to mixed derivatives with respect to α and $\beta = 1/T$. Since all the derivatives commute, we only have to deal with moments of the form $\omega_{12\dots 2}$, $\omega_{13\dots 3}$, $\omega_{12\dots 23\dots 3}$, $\omega_{2\dots 23\dots 3}$, $\omega_{2\dots 2}$ and $\omega_{3\dots 3}$. The relevant spectral functions are $I_{1\dots j}(\nu_c, \bar{\mathbf{p}}) = I_{1\dots j}(\nu_c, \bar{\mathbf{p}})/\bar{A}_0$ with $i, j \in \{2, 3\} \equiv \{\alpha, \beta\}$, so that the moments $\omega_{12\dots 2}$, $\omega_{13\dots 3}$ and $\omega_{12\dots 23\dots 3}$ can again be absorbed by defining $\omega_{1\dots ij\dots j}^d = \omega_{1\dots ij\dots j} + \omega_{1\dots ij\dots j}/\bar{A}_0$. Assuming narrow bands and $W(\hat{\gamma}, \nu) = B(\hat{\gamma})F(\nu)$, we then find

$$\begin{aligned} \langle I_\nu \rangle &= \frac{\bar{A}_0 (\nu/\nu_0)^{\bar{\alpha}} \nu^3}{e^{\bar{x}} - 1} \left\{ 1 + \frac{1}{2} \omega_{22}^d \ln^2(\nu/\nu_0) \right. \\ &\quad + \omega_{23}^d \ln(\nu/\nu_0) Y_1(x) + \frac{1}{2} \omega_{33}^d Y_2(x) \\ &\quad + \frac{1}{6} \omega_{222}^d \ln^3(\nu/\nu_0) + \frac{1}{2} \omega_{223}^d \ln^2(\nu/\nu_0) Y_1(x) \\ &\quad \left. + \frac{1}{2} \omega_{233}^d \ln(\nu/\nu_0) Y_2(x) + \frac{1}{6} \omega_{333}^d Y_3(x) + \dots \right\} \\ \bar{\alpha} &= \frac{\langle A_0(\mathbf{r}) \alpha(\mathbf{r}) \rangle}{\bar{A}_0}, \quad \frac{1}{\bar{T}} = \frac{\langle A_0(\mathbf{r})/T(\mathbf{r}) \rangle}{\bar{A}_0} \\ \omega_{2\dots 23\dots 3}^d &= \frac{\langle A_0(\mathbf{r})[(\alpha(\mathbf{r}) - \bar{\alpha})^k ((\bar{T}/T(\mathbf{r}) - 1)^m)] \rangle}{\bar{A}_0}. \end{aligned} \quad (48)$$

Higher order terms can be easily added in a similar way. This expression again captures all the degrees of freedom introduced by the averaging inside the beam and along the line of sight. Note that the simple product assuming individual superpositions of power-law and gray-body spectra would be *insufficient*, since the cross-terms $\propto \ln^k(\nu/\nu_0) Y_m(x)$, would not have independent coefficients to take into account possible correlations between α and T . For example, the term for $k = 1$ and $m = 1$ would be absent.

6.1 Behavior in the Rayleigh-Jeans limit

In the Rayleigh-Jeans limit ($h\nu \ll kT$), we can re-express the modified blackbody spectrum as

$$I_\nu^{\text{RJ}}(A_0, \alpha, T) \approx A_0 (\nu/\nu_0)^{\alpha+2} \left[1 - \frac{x}{2} + \frac{x^2}{12} - \frac{x^4}{720} + \dots \right]. \quad (49)$$

This represents a superposition of power-law spectra, which in principle can be approximated with Eq. (21). Since $x = h\nu/kT$, the temperature variations of the dust itself enter as a weight-factor for the power-law moment amplitudes, introducing specific correlations among the moments. The Taylor series of $1/(e^x - 1)$ furthermore can only be recovered when allowing *differences* between power-laws instead of just sums as considered previously. We will discuss the applicability of this approximation below (Sect. 6.3.1).

6.2 Alternative parameterizations and weighting schemes

In Sect. 5.2.1, we illustrated how the choice of parameters and weighting affects the moment expansion for gray-body spectra. In particular, we found that the recovered best-fitting parameters for truncated moment expansions are sensitive to these choices. In a similar manner as for the gray-body spectra, by setting $\nu_0 = kT/h$ and rewriting $I_\nu(A_0, \alpha, T) = A_0 (\nu/\nu_0)^\alpha \nu^3 / (e^{h\nu/kT} - 1)$ as

$$I_\nu = A_0 \left(\frac{kT}{h} \right)^3 \frac{x^{3+\alpha}}{e^x - 1} = B(T, \alpha) \frac{\bar{x}^{3+\alpha} (\bar{T}/T)^{3+\alpha}}{e^{\bar{x}\bar{T}/T} - 1} \quad (50)$$

we can derive an alternative moment expansion for the modified blackbody spectra with new weighting $B(T, \alpha) = A_0 (kT/h)^{3+\alpha}$. However, we find that this approach does not significantly change the capabilities of the moment expansion to represent different SEDs. It only leads to a re-summation of terms up to a given moment order. We therefore did not consider this approach any further.

6.3 Two-temperature dust models

To illustrate the application of the moment expansion, let us consider a simple two-temperature dust model, with SED parameters $\mathbf{p} = \{A_1, A_2, \alpha_1, \alpha_2, T_1, T_2\}$. We then have $\bar{A} = A_1 + A_2$, $\bar{\alpha} = (A_1\alpha_1 + A_2\alpha_2)/\bar{A}$, $\bar{T} = \bar{A}/(A_1/T_1 + A_2/T_2)$ and the moments

$$\omega_{2\dots 23\dots 3}^d = \sum_{a=1}^2 \frac{A_a}{\bar{A}} (\alpha_a - \bar{\alpha})^k (\bar{T}/T_a - 1)^m. \quad (51)$$

The general behavior of the moment expansion for different cases is similar to that of gray-body spectra, as mentioned above. The additional parameter, α , causes the number of moments per perturbation order to increase strongly. For example, adding all second order moment terms means 3 extra parameters instead of 1 for the gray-body spectrum; at third order we need 7 instead of 2 and so on. The required number of parameters can thus grow quickly for complicated dust distribution functions.

To illustrate the method, let us consider a concrete example, setting the model parameters to $P_1 = \{A_1, A_2, \alpha_1, \alpha_2, T_1, T_2\} = \{A f_1, A(1 - f_1), 1.63, 2.82, 9.75 \text{ K}, 15.7 \text{ K}\}$ with $f_1 = 0.34188$ and $\nu_0 = 3 \text{ THz}$, based on recent modeling of CMB data (Meisner & Finkbeiner 2015). We fix A such that $\langle I_\nu \rangle \approx 3388 \text{ Jy sr}^{-1}$ at $\nu = 100 \text{ GHz}$, but for illustrations of the method, the absolute value does not matter, as the spectral shape is fixed without the amplitude. We assume 200 frequency bins in $\log-\nu$ of different ranges and perform a simple χ^2 -fit using Eq. (50) setting $\nu_0 = k\bar{T}/h$. This choice of the pivot frequency leads to faster convergence. Representing the resultant spectrum (see Fig. 6) with a single-temperature modified

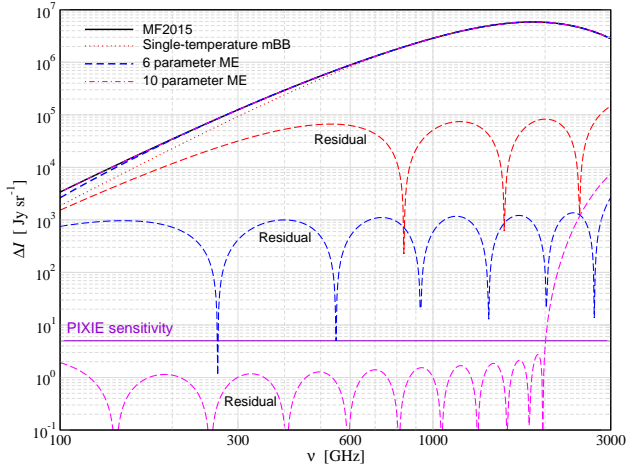


Figure 6. Representation of a two-temperature modified blackbody spectrum from Meisner & Finkbeiner (2015) using a varying number of moments (see text for details). The long-dashed lines show the residuals in comparison with the estimated absolute sensitivity of *PIXIE*.

blackbody in the range $100 \text{ GHz} \lesssim \nu \lesssim 3 \text{ THz}$ yields $\alpha = 1.819$ and $T = 18.45 \text{ K}$. We find that at low frequencies this approximation shows a large departure from the true SED, underestimating the emission by a factor of two (see Fig. 6). Adding the first order moment corrections, we find $\alpha = 1.270$, $T = 15.27 \text{ K}$, $\omega_{22}^d = -0.4438$, $\omega_{23}^d = 0.1695$ and $\omega_{33}^d = 0.2094$, which provides an approximation that is better than $\approx 0.4\%$ at $300 \text{ GHz} \lesssim \nu \lesssim 3 \text{ THz}$, departing only notably at low frequencies, reaching $\approx 20\%$ at $\nu \approx 100 \text{ GHz}$ (see Fig. 6). This representation uses the same number of parameters (six in total) as the input model, but without assuming a two-temperature case, and can in principle already capture a broader range of distributions in temperature and spectral indices.

The convergence of the moment representation also depends on the frequency range that is assumed in the fitting process. For a second order moment representation at $100 \text{ GHz} \lesssim \nu \lesssim 2 \text{ THz}$, we obtain $\alpha = 1.412$, $T = 14.84 \text{ K}$, $\omega_{22}^d = -0.2793$, $\omega_{23}^d = 0.05132$ and $\omega_{33}^d = 0.2191$. This representation improves the fit at low frequencies, with the departure decreasing to $\lesssim 8\%$ at $\nu \approx 100 \text{ GHz}$. This illustrates that in precision measurements over a wide range of frequencies, the high frequency part of the dust spectrum can drive the solution away from the optimal solution at low frequencies. This could affect the ability to recover primordial distortion signals in high precision CMB spectroscopy, but also is relevant to the modeling of foregrounds for primordial *B*-modes. Incorrect and incomplete (truncated) parameterizations for the foregrounds can thus yield biased results. Convergence properties of the moment expansion can again be tested by varying the moment order, which provides a powerful diagnostic, but is limited by the available number of channels.

If we now include all moments up to third order ($3+3+4=10$ parameters), we find $\alpha = 2.057$, $T = 13.33 \text{ K}$, $\omega_{22}^d = 0.002363$, $\omega_{23}^d = -0.4189$, $\omega_{33}^d = 0.3789$, $\omega_{222}^d = -0.004326$, $\omega_{223}^d = 0.5326$, $\omega_{233}^d = -0.2657$ and $\omega_{333}^d = 0.08050$ at $100 \text{ GHz} \lesssim \nu \lesssim 2 \text{ THz}$. We find that this approximation represents the dust spectrum at $100 \text{ GHz} \lesssim \nu \lesssim 2 \text{ THz}$ to better than $\approx 0.06\%$ precision (see Fig. 6). We can see that the higher order moments start decreasing rapidly, showing that the chosen example is in the perturbative regime and converges fairly quickly. The achieved level of precision would be about 4–5 times better than the expected sensitivity of *PIXIE* for

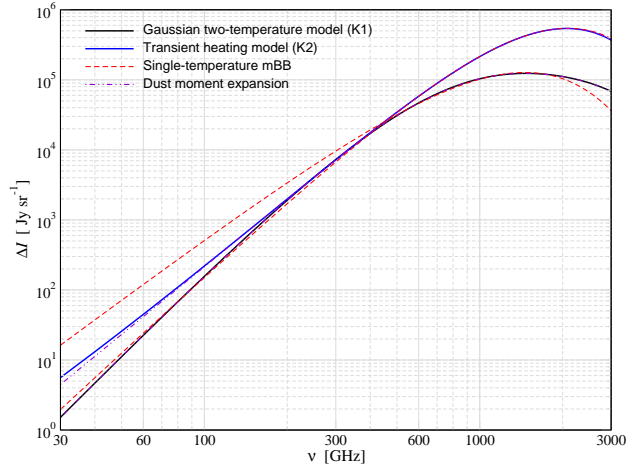


Figure 7. Moment representation of dust models recently discussed by Kogut & Fixsen (2016). The fits were obtained in the frequency range $100 \text{ GHz} \lesssim \nu \lesssim 2 \text{ THz}$. The red lines assume a single-temperature dust model. The violet (double-dotted dashed) lines assume a 10/6 parameter moment expansion for K1/K2. Notice that the 10 parameter moment expansion for K1 covers the input model curve even outside the fit domain.

absolute CMB spectroscopy ($\Delta I_\nu \approx 5 \text{ Jy/sr}$; Kogut et al. 2011). Again, the parameterization is more general than simply assuming a two-temperature model, capturing more general distributions of dust temperatures and spectral indices.

The overall representation of the average spectrum with the third order moment expansion degrades significantly (by more than one order of magnitude in terms of absolute precision) when increasing the upper frequency to $\nu = 3 \text{ THz}$. As expected from our discussion of gray-body spectra, in this case higher order moments are required to compensate for the asymptotic behavior of the basis functions. For a given finite order and distribution of temperatures, this behavior is inevitable with the moment expansion. In this case, alternative approaches that directly assume representations of the temperature distribution functions could be used; however, a precise description of the underlying probability distribution functions (with a significant number of parameters) is still required. Similarly, one could use two hierarchies of moment expansions, forming a two-temperature basis to improve the local convergence. We leave a more detailed discussion of these ideas to future work.

6.3.1 Rayleigh-Jeans approximation

We saw in Sect. 6.1, at low frequencies the superposition of dust spectra can be thought of as a superposition of power-laws. It turns out that using a simple power-law expansion to represent the average dust spectrum for the model discussed above works quite well at $30 \text{ GHz} \lesssim \nu \lesssim 500 \text{ GHz}$, when including only $N \approx 4$ moments (6 parameters in total). A sum of synchrotron and dust spectra, relevant to CMB applications, also works well with $N \approx 6$ moments (8 parameters in total, which is one less than the input model) at $30 \text{ GHz} \lesssim \nu \lesssim 200 \text{ GHz}$. However, at higher frequencies corrections related to the Taylor series of $1/(e^x - 1)$ become too large so that many power-law moments ($N \gg 6$) have to be included for a sufficient representation of the dust spectrum. In this case, the general dust moment expansion should be preferred.

6.4 Distributions of temperatures

To demonstrate the potential of the dust moment representation, we analyze two more general dust models considered by Kogut & Fixsen (2016), one assuming a sum of two Gaussians for the dust temperature (K1), the other using a transient heating model¹³ (K2). The resultant SEDs are shown in Fig. 7.

Without knowing the details of the temperature model, we can represent these spectra using the dust moment expansion, Eq. (50). Fitting in the range $100 \text{ GHz} \lesssim \nu \lesssim 2 \text{ THz}$, we find $\alpha = 0.9379$ and $T = 18.15 \text{ K}$ (K1) and $\alpha = 1.671$ and $T = 21.69 \text{ K}$ (K2) assuming a single-temperature modified blackbody spectrum. The overall amplitude, A_0 , is determined using $I_\nu = 156.61 \text{ Jy/sr}$ (K1) and $I_\nu = 220.11 \text{ Jy/sr}$ (K2) at $\nu = 100 \text{ GHz}$. Clearly, this representation fails to approximate the SED for model K1 but already provides a good approximation for K2 (see Fig. 7).

A second order moment expansion (6 parameters) significantly improves the fit in both cases, leaving residuals at the level of $|\Delta I_\nu| \lesssim 50 - 200 \text{ Jy/sr}$ (K1) and $|\Delta I_\nu| \lesssim 5 \text{ Jy/sr}$ (K2) in the considered frequency range. For K2, the 6 parameter moment expansion is shown in Fig. 7. We find that at $100 \text{ GHz} \lesssim \nu \lesssim 2 \text{ THz}$, the transient heating model can be fully represented down to the *PIXIE* sensitivity using only 6 parameters without a priori assumptions about the temperature model, while outside this range higher order moments are required. We also find that for both models, a two-temperature dust model leads to a similar performance, albeit being less general.

Using a third order moment representation (10 parameters), we obtain residuals that are more than one order of magnitude below the sensitivity of *PIXIE* in both cases. For K1, we have $\alpha = 2.327$, $T = 14.56 \text{ K}$, $\omega_{22}^d = -0.2983$, $\omega_{23}^d = -0.1718$, $\omega_{33}^d = -0.1396$, $\omega_{222}^d = 0.03299$, $\omega_{223}^d = 0.5685$, $\omega_{233}^d = -0.07951$ and $\omega_{333}^d = 0.03106$ (see Fig. 7). This shows that the dust moment expansion is able to handle more complicated dust spectra with few a priori assumptions about the underlying distribution functions. It also indicates that a significantly larger number of foreground parameters has to be considered. This means that the overall dust foreground exhibits a richer spectral-spatial morphology related to each new parameter. With approaches based on probability distribution function modeling (e.g., see Martínez-Solache et al. 2017, for a recent example), this aspect cannot be easily incorporated, which highlights another benefit of the moment approach.

7 EFFECT OF SPATIAL VARIATIONS ON THE ENERGY DISTRIBUTION AT DIFFERENT ANGULAR SCALES

As mentioned in Sect. 2.2, the moment method could provide a simple way to propagate scale-dependent changes to the SED from small to large scales. This point can be illustrated using the spatially varying CMB temperature and the effect on the SED of different multipoles. A similar effect occurs due to the spatial variations of the dust temperature and spectral index, as discussed below and in more detail in a future paper (Hill et al., in preparation).

7.1 CMB blackbody superposition

A spherical harmonic transformation of the (ideal CMB-only) sky intensity leads to distortions (\leftrightarrow modified frequency dependence) of the multipole coefficients due to the weighted average (\leftrightarrow superposition) of blackbodies with different temperatures. This is an

artefact of the map making procedure, which (under idealized assumptions) could be completely avoided by using maps of the thermodynamic temperature instead. To illustrate this point, let us consider the CMB blackbody intensity in different directions

$$I_\nu^{\text{CMB}}(\hat{\gamma}) = \frac{2h}{c^2} \frac{\nu^3}{e^{h\nu/kT(\hat{\gamma})} - 1} \quad (52)$$

$$\approx \frac{2h}{c^2} \frac{\nu^3}{e^x - 1} \left[1 + G(x)(\Delta + \Delta^2) + \frac{1}{2}Y(x)\Delta^2 + \mathcal{O}(\Delta^3) \right],$$

where $x = h\nu/k\bar{T}$, with \bar{T} denoting the full-sky average temperature and $\Delta(\hat{\gamma}) = T(\hat{\gamma})/\bar{T} - 1$. In the second step we assumed $\Delta \ll 1$ and introduced $G(x) = xe^x/(e^x - 1)$ and $Y(x) = G(x)[x \coth(x/2) - 4]$. We neglected higher order terms; however, these can become important at very high frequencies (Chluba & Sunyaev 2004).

We started with a blackbody at thermodynamic temperature $T(\hat{\gamma})$ in every direction, $\hat{\gamma}$; thus, even the Taylor series, Eq. (52), still simply represents a pure blackbody spectrum in every direction. If we compute the all-sky average intensity, we find

$$\bar{I}_\nu^{\text{CMB}} \approx \frac{2h}{c^2} \frac{\nu^3}{e^x - 1} \left[1 + 2y_{\text{sup}} G(x) + y_{\text{sup}} Y(x) \right], \quad (53)$$

where $y_{\text{sup}} = \int \frac{\Delta^2(\hat{\gamma})}{8\pi} d^2\hat{\gamma}$, which in general does not vanish. The frequency dependence of the average intensity spectrum is no longer a pure blackbody, but has a y-type distortion (Chluba & Sunyaev 2004). One can absorb the temperature shift term $\propto 2y_{\text{sup}} G(x)$ by redefining the average temperature to $T' = \bar{T}(1 + 2y_{\text{sup}})$, but the last term cannot be removed in this way. This already demonstrates how fluctuations at the smallest scales propagate to the largest, changing the SED from a pure blackbody in each direction to a distorted blackbody (Chluba & Sunyaev 2004).

Similarly, for the multipole coefficients with $\ell > 0$, a distortion is created when carrying out the spherical harmonic transformation of the intensity. Defining the antenna temperature fluctuation as

$$\Delta T^{\text{CMB}}(\nu, \hat{\gamma}) = \frac{c^2}{2h} \frac{(e^x - 1)}{\nu^3 G(x)} [I_\nu^{\text{CMB}} - \bar{I}_\nu^{\text{CMB}}], \quad (54)$$

the multipoles for $\ell > 0$ are given by

$$\Delta T_{\ell m}^{\text{CMB}}(\nu) \approx \Delta_{\ell m} + [\Delta^2]_{\ell m} + \frac{1}{2} \frac{Y(x)}{G(x)} [\Delta^2]_{\ell m}, \quad (55)$$

where $[X]_{\ell m} = \int Y_{\ell m}^*(\hat{\gamma}) X(\hat{\gamma}) d^2\hat{\gamma}$ is the multipole coefficient of $X(\hat{\gamma})$. This shows two features of the spherical harmonic transformation of the intensity map: i) the derived antenna temperature fluctuations are frequency dependent, even if the thermodynamic temperature fluctuations, $\Delta_{\ell m}$, per definition are independent of frequency; ii) independent thermodynamic temperature multipoles couple at second order due to the weighted average (mixing). Both of these effects are small for the primordial CMB temperature fluctuations at multipoles $\ell > 1$ (e.g., Chluba & Sunyaev 2004); however, our motion with respect to the CMB rest frame causes a relative leakage of power $\mathcal{O}(v/c) \simeq 10^{-3}$ between adjacent multipoles (e.g., Challinor & van Leeuwen 2002; Amendola et al. 2011; Chluba 2011), which causes a larger effect that needs to be taken into account (Planck Collaboration et al. 2014b). With the moment expansion, here defined by $\Delta_{\ell m}$ and $2y_{\ell m} = [\Delta^2]_{\ell m}$, this effect can be trivially included and propagated. The CMB sky is thus described by maps of $\Delta_{\ell m}$ and $y_{\ell m}$, which have slightly different spatial morphology and statistical properties.

We reiterate that these issues could be completely avoided by converting to thermodynamic temperature *before* the spherical harmonic transformation, requiring a different map making procedure. In the presence of foregrounds, this step is non-trivial, but the total

¹³ We cordially thank Alan Kogut for providing the average SEDs to us.

SEDs and the power spectra of different moment values can be used to handle these issues. We also mention that a similar distortion is created due to the superposition of blackbodies of different temperatures for polarized contributions (Chluba et al. 2015), which could be relevant to high precision E - and B -mode searches. A more detailed consideration of these problems is left to future work.

7.2 Thermal dust superposition

As a second illustration, we consider a simple example in which only the spectral index of a thermal dust SED is allowed to vary over the sky, asking how a spherical harmonic transformation affects the SED at different multipoles. First, we generate a reference template dust map at 353 GHz, consisting of a Gaussian random field with a power spectrum described by a pure power-law, $C_\ell \propto \ell^{-2.7}$, matching recent *Planck* measurements (Planck Collaboration et al. 2016c). Second, we assume that this template can be rescaled to other frequencies according to a modified blackbody SED. The dust temperature of this SED is fixed to $T_d = 19$ K (Planck Collaboration et al. 2016c) and does not vary over the sky. The spectral index α , however, is allowed to vary on the sky. We generate maps of α in which the value in each pixel is drawn from a Gaussian of mean $\alpha_{\text{fid}} = 1.6$ and standard deviation 0.1, matching recent *Planck* full-sky results (Planck Collaboration et al. 2016c). The amplitude in each direction on the sky is set by the reference 353 GHz template. We generate the reference template at HEALPix resolution $N_{\text{side}} = 256$ (pixel size ≈ 0.052 deg²), while the spectral index maps are generated at $N_{\text{side}} = 64$ (pixel size ≈ 0.84 deg²).

Using these ingredients, we construct maps at ten linearly spaced frequencies from 20 GHz to 353 GHz, generating 100 independent realizations (i.e., 100 spectral index maps and 100 reference template maps). We do not add noise or beam smoothing for this simple illustration. We compute the angular power spectrum for each frequency map and average the results in seven multipole bins linearly spaced between $2 < \ell < 400$. We average over all 100 realizations at each frequency, and compute error bars from the scatter among these 100 realizations.

The results are shown in Fig. 8, plotted as the effective SED for each bandpower in terms of $D_\ell \equiv \ell(\ell+1)C_\ell/(2\pi)$ at each frequency, normalized to the 353 GHz result. The effective SED deviates most significantly from the input mean SED (i.e., the square of the modified blackbody described above) on the scales at which the spectral index has been specified to vary, i.e., $\ell \approx 200$ –300 (corresponding to the pixel size of the spectral index map, as given above). On much larger or much smaller scales, the effective SED tends back toward the input mean SED (e.g., for the lowest bandpower centered at $\ell = 30$, the behavior is nearly indistinguishable from the input mean SED). Of course, in reality the dust properties likely vary over a range of scales, which would lead to more complex behavior than seen here. Nonetheless, this simple example already illustrates the key point: averaging over populations with differing SEDs (even if drawn from the same fundamental SED shape, in this case a modified blackbody) will generically produce a multipole-dependent SED in the power spectrum. The moment method provides a way to efficiently capture this behavior, by introducing new moment maps with varying statistical properties and spatial morphology which combine through new spectral functions related to higher derivatives with respect to the fundamental SED parameters.

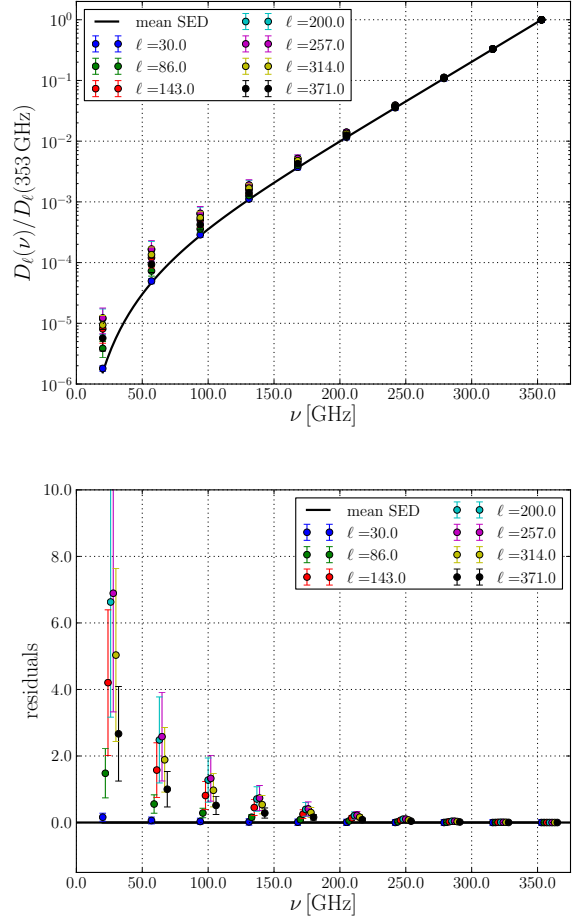


Figure 8. Effective SEDs of angular power spectra of modified blackbody simulations (top panel) and fractional residuals with respect to the input mean SED (bottom panel). The solid black curve shows the input mean SED, a modified blackbody with spectral index $\alpha_{\text{fid}} = 1.6$ and temperature $T_d = 19$ K. The amplitude is set by normalizing all results to the 353 GHz reference template. In the bottom panel, the points are slightly offset in the horizontal direction for additional clarity. The effective SED shows the largest departure from the assumed form on the scales where the spectral index variations are largest ($\ell \approx 200$ –300, corresponding to the pixel size of the generated spectral index map). In reality, the SED properties will vary over a range of scales and can include both temperature and spectral index variations, yielding more complex behavior than in this simple example.

8 THE MODELING OF OTHER SPECTRAL COMPONENTS IN CMB APPLICATIONS

In this paper, we focused on a few common examples appearing in CMB foreground analysis. There are several additional known and less known components (see De Zotti et al. 2016; Planck Collaboration et al. 2016a, for some overview). At low frequencies, anomalous microwave emission (AME), possibly produced by spinning dust grains in the intergalactic medium, contributes significantly, showing a broad maximum around $\nu \approx 10$ GHz – 30 GHz (e.g., Draine & Lazarian 1998; Ali-Haïmoud et al. 2009; Hoang et al. 2016). This component exhibits a rich phenomenology of relatively broad spectral shapes, depending on the grain sizes and optical properties, temperature and ambient stellar light, to name a few of the ingredients. It is usually treated using spectral templates with a free amplitude and position (Planck Collaboration et al. 2016a).

Significant new information about this component is now becoming available [e.g., with *C-BASS* (Irfan et al. 2015) and *QUIJOTE* (G  nova-Santos et al. 2015)]. This will help in understanding the complexity of the AME and a moment expansion may provide one avenue towards a refined treatment in component separation analyses, capturing variations in the amount of spinning dust and typical emissivity but also other physical properties. In this case, an orthogonalization scheme may be beneficial to reduce the number of common modes for the different variables.

At intermediate CMB frequencies ($30 \text{ GHz} \lesssim \nu \lesssim 200 \text{ GHz}$), the cumulative CO rotational line emission from star-forming galaxies throughout the Universe becomes important (Righi et al. 2008; De Zotti et al. 2016; Mashian et al. 2016). Recent measurements with *ALMA* (Carilli et al. 2016) show that in agreement with previous model estimates, this component will have to be considered carefully to reach the full potential of future CMB spectral distortion measurements. The uncertainties in the modeling of this component are significant and require refined theoretical studies; however, also in this case, a moment expansion in the main parameters may provide a viable approach for incorporating it in future CMB analyses. Future large galaxy surveys may provide additional constraints to inform the theoretical modeling of the CO emission.

We also mention that at high frequencies additional dust-like components are present. One is due to the cosmic infrared background (CIB) which can be described using a halo-model approach (e.g., Desjacques et al. 2015). At frequencies above the dust peak ($\nu \approx 1.5\text{--}2 \text{ THz}$), the SED in this case is typically much flatter than that of a modified blackbody spectrum, while at lower frequencies a dust moment expansion is expected to provide an efficient representation. This further highlights that measurements at THz frequencies will not necessarily help in constraining the low frequency tails of the dusty foregrounds.

Another dusty component is caused by the cumulative emission of intergalactic dust (Imara & Loeb 2016). Since the fundamental SED in these models is assumed to be given by a modified blackbody spectrum, a dust moment expansion is expected to efficiently represent this contribution.¹⁴

9 CONCLUSIONS

We developed and illustrated a new generalized parameterization for the modeling of spatially varying foreground components relevant to CMB observations. The method relies on a minimal set of assumptions about the underlying distributions of physical parameters that determine the SED of the considered component. The parameter list is extended by moments of the underlying distribution functions, which fix the amplitude of new spectral shapes (related to derivatives of the fundamental SED) that are caused by line of sight and beam averages. We applied the method to superpositions of power-law (Sect. 3), free-free (Sect. 4), gray-body (Sect. 5), and modified blackbody (Sect. 6) spectra, illustrating how these can be used to describe different foreground models.

All our examples assume noiseless reconstructions for a single type of component, meant to illustrate the main features and limitations of different parameterizations. Our analysis clearly suggests that even simply incorporating the inevitable effects of spatial variations of SED parameters calls for more general methods with a

larger number of degrees of freedom being required. This provides an indication for the true challenges that are awaiting us in future CMB B-mode and spectral distortion searches.

We find power-law moment expansions [Eq. (21)] to converge quite rapidly. For the modeling of synchrotron emission, we expect $N \approx 3\text{--}4$ moments to provide sufficient spectral freedom to reach sensitivity levels required for absolute CMB spectroscopy. However, for power-law superpositions with stronger variation of the spectral index, the convergence of the power-law moment expansion can become quite slow (see Sect. 3.1).

For free-free we obtained two moment expansions [Eq. (29) and Eq. (31)]. The first represents the commonly used model for the optically thin free-free spectrum; however, it neglects free-free absorption of CMB photons and stimulated free-free emission, which leads to a small suppression at high frequencies ($\nu \geq 1 \text{ THz}$) and is captured by the second representation, Eq. (31). Variations in the chemical compositions and ionization degree are not expected to be as relevant to CMB foreground modeling, but applications of the moment method to the continuum X-ray emission from clusters could require a more detailed treatment.

We also find that the free-free SED can be described using a simple power-law moment expansion with $N \approx 3\text{--}4$ moments. This suggests that the combined emission from synchrotron and free-free could be modeled using one parametrization. Only at very high frequencies ($\nu \geq 1 \text{ THz}$) do these two components become spectrally distinguishable once the effects of spatial variations are included. In a similar manner, we expect a power-law moment expansion to capture the contributions from radio-point sources in low-angular resolution observations, when masking cannot be applied. This could greatly reduce the number of independent parameters that are required to model these low-frequency foregrounds, with the different components mainly being distinguished via physical priors (e.g., by assuming that free-free has a typical spectral index of $\alpha \approx -0.14$ rather than $\alpha \approx -0.9$ for synchrotron).

For the superposition of gray-body spectra, we developed two moment expansions with individual weighting [Eq. (39) and Eq. (43)]. We showed that the convergence of the moment expansion is slow at high frequencies (in the Wien tail), which can lead to biases at low frequencies. The two considered weighting schemes are physically equivalent; however, an interpretation of the obtained best-fitting moments in truncated moment expansions in terms of the underlying parameter distribution functions can be affected significantly (Sect. 5.2.1). We also derived an orthonormal set of basis functions for the gray-body moment expansion (Sect. 5.3), however, our discussion does not suggest that this will greatly improve the applicability of the moment parametrization in this case, as the number of free parameters is not reduced.

Finally, we provide a moment expansion for modified blackbody spectra [Eq. (50)]. We blindly applied this parametrization to recent models for the dust emission discussed by Kogut & Fixsen (2016), showing that without many a priori assumptions about the temperature distribution functions we obtain accurate representations of the average SEDs (Fig. 7). Our modeling suggested that once general distributions of temperature and dust opacities are allowed, about $\approx 6\text{--}10$ parameters are required to model the individual dust components down to the level of sensitivity required in absolute CMB spectroscopy. We expect a similar number to be necessary for future CMB polarization studies; with more restrictive models biases in the deduced cosmological parameters (in particular the tensor-to-scalar ratio) could be produced. The capabilities of future CMB anisotropy experiments have to be explicitly tailored towards this challenge.

¹⁴ For some of the models presented in Imara & Loeb (2016), we already confirmed this statement, finding only $N = 6$ parameters to suffice.

In Sect. 7, we also briefly discussed the effect of spatial variations on the SED at different scales. Specifically, the dust SED can show significant scale-dependent modifications caused by spatial variations of the spectral index (see Fig. 8). The moment expansion provides a simple way for parameterizing and propagating these effects in real maps. This is achieved by separating spatial from spectral variations through the moment expansion, where each moment map can have a new spatial morphology. For SZ clusters, this was illustrated in Chluba et al. (2013) and for CMB foreground modeling we plan to study this thoroughly in the future.

We close by saying that one of the important assumptions of the moment method is that a detailed form for the fundamental SED is known. Nature may not comply with the spectral forms used here (and in standard CMB foreground analyses); however, improvements of the fundamental SEDs can be easily implemented. We also expect the moment method to be extendable to other foreground components (see Sect. 8) and to CMB polarization. Given the physically motivated extensions of foreground parameterizations, the moment expansion can furthermore be used as a diagnostic, testing the robustness of the foreground modeling and derived cosmological parameters. External data sets can also be used to place priors on the moment values and the relations of different moments. Overall, the moment method provides a much more general framework for the treatment of CMB foregrounds.

ACKNOWLEDGMENTS

JC is supported by the Royal Society as a Royal Society University Research Fellow at the University of Manchester, UK. This work was partially supported by a Junior Fellow award from the Simons Foundation to JCH.

REFERENCES

- Abazajian K. N. et al., 2016, ArXiv:1610.02743
 Abazajian K. N. et al., 2015, *Astroparticle Physics*, 63, 66
 Abitbol M. H., Chluba J., Hill J. C., Johnson B. R., 2017, *MNRAS*
 Ali-Haïmoud Y., Hirata C. M., Dickinson C., 2009, *MNRAS*, 395, 1055
 Amendola L., Catena R., Masina I., et al., 2011, *JCAP*, 7, 27
 André P. et al., 2014, *JCAP*, 2, 6
 Bennett C. L. et al., 2003, *ApJS*, 148, 1
 Carilli C. L. et al., 2016, *ApJ*, 833, 73
 Challinor A., van Leeuwen F., 2002, *Phys.Rev.D*, 65, 103001
 Chluba J., 2011, *MNRAS*, 415, 3227
 Chluba J., 2013, *MNRAS*, 436, 2232
 Chluba J., 2016, *MNRAS*, 460, 227
 Chluba J., Dai L., Grin D., Amin M. A., Kamionkowski M., 2015, *MNRAS*, 446, 2871
 Chluba J., Khatri R., Sunyaev R. A., 2012a, *MNRAS*, 425, 1129
 Chluba J., Nagai D., Sazonov S., Nelson K., 2012b, *MNRAS*, 426, 510
 Chluba J., Sunyaev R. A., 2004, *A&A*, 424, 389
 Chluba J., Sunyaev R. A., 2012, *MNRAS*, 419, 1294
 Chluba J., Switzer E., Nelson K., Nagai D., 2013, *MNRAS*, 430, 3054
 De Zotti G., Negrello M., Castex G., Lapi A., Bonato M., 2016, *JCAP*, 3, 047
 Desjacques V., Chluba J., Silk J., de Bernardis F., Doré O., 2015, *MNRAS*, 451, 4460
 Draine B. T., 2003, *ARA&A*, 41, 241
 Draine B. T., 2011, *Physics of the Interstellar and Intergalactic Medium*. Princeton University Press
 Draine B. T., Lazarian A., 1998, *ApJL*, 494, L19
 Errard J., Feeney S. M., Peiris H. V., Jaffe A. H., 2016, *JCAP*, 3, 052
 Finkbeiner D. P., Davis M., Schlegel D. J., 1999, *ApJ*, 524, 867
 Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, *ApJ*, 473, 576
 Fuskeland U., et al., 2014, *ApJ*, 790, 104
 Génova-Santos R. et al., 2015, *MNRAS*, 452, 4169
 Hill J. C., Battaglia N., Chluba J., Ferraro S., Schaan E., Spergel D. N., 2015, *Physical Review Letters*, 115, 261301
 Hoang T., Vinh N.-A., Quynh Lan N., 2016, *ApJ*, 824, 18
 Hu W., 1995, arXiv:astro-ph/9508126
 Imara N., Loeb A., 2016, *ApJ*, 825, 130
 Irfan M. O. et al., 2015, *MNRAS*, 448, 3572
 Itoh N., et al., 2000, *ApJS*, 128, 125
 Itoh N., Kohyama Y., Nozawa S., 1998, *ApJ*, 502, 7
 Karzas W. J., Latter R., 1961, *ApJS*, 6, 167
 Kogut A., Chluba J., Fixsen D. J., Meyer S., Spergel D., 2016, in *Proc.SPIE*, Vol. 9904, SPIE Conference Series, p. 99040W
 Kogut A., Fixsen D. J., 2016, *ApJ*, 826, 101
 Kogut A. et al., 2011, *ApJ*, 734, 4
 Martínez-Solache G., Karakci A., Delabrouille J., 2017, ArXiv:1706.04162
 Mashian N., Loeb A., Sternberg A., 2016, *MNRAS*, 458, L99
 Mather J. C. et al., 1994, *ApJ*, 420, 439
 Matsumura T. et al., 2014, *Journal of Low Temperature Physics*, 176, 733
 Meisner A. M., Finkbeiner D. P., 2015, *ApJ*, 798, 88
 Mozdzen T. J., Bowman J. D., Monsalve R. A., Rogers A. E. E., 2016, *MNRAS*, 455, 3890
 Nozawa S., et al., 2006, *Nuovo Cimento B Serie*, 121, 487
 Planck Collaboration et al., 2016a, *A&A*, 594, A9
 Planck Collaboration et al., 2014a, *A&A*, 571, A12
 Planck Collaboration et al., 2016b, *A&A*, 594, A13
 Planck Collaboration et al., 2014b, *A&A*, 571, A27
 Planck Collaboration et al., 2016c, *A&A*, 596, A109
 Ponente P. P., Diego J. M., Sheth R. K., Burigana C., Knollmann S. R., Ascasibar Y., 2011, *MNRAS*, 410, 2353
 Pritchard J. R., Loeb A., 2010, *Phys.Rev.D*, 82, 023006
 Remazeilles M., Dickinson C., Eriksen H. K. K., Wehus I. K., 2016, *MNRAS*, 458, 2032
 Righi M., Hernández-Monteagudo C., Sunyaev R. A., 2008, *A&A*, 489, 489
 Rybicki G. B., Lightman A. P., 1979, *Radiative processes in astrophysics*. New York, Wiley-Interscience, 1979. 393 p.
 Sathyanarayana Rao M., Subrahmanyam R., Udaya Shankar N., Chluba J., 2015, *ApJ*, 810, 3
 Sazonov S. Y., Sunyaev R. A., 1998, *ApJ*, 508, 1
 Sunyaev R. A., Khatri R., 2013, *IJMPD*, 22, 30014
 Tashiro H., 2014, *Prog. of Theo. and Exp. Physics*, 2014, 060000
 Thorne B., Dunkley J., Alonso D., Naess S., 2016, ArXiv:1608.02841
 Zeldovich Y. B., Sunyaev R. A., 1969, *ApSS*, 4, 301

APPENDIX A: FREE-FREE EMISSION IN THE OPTICALLY THIN REGIME

The radiative transfer equation for the photon occupation number, $n(x_e = hv/kT_e)$, under free-free emission and absorption in an isotropic medium takes the simple form (Rybicki & Lightman 1979; Hu 1995)

$$\frac{\partial n(x_e)}{\partial \tau} = \frac{\kappa(x_e, T_e) e^{-x_e}}{x_e^3} [1 - n(x_e)(e^{x_e} - 1)]. \quad (A1)$$

Here, $\tau = \int N_e \sigma_T dl$ is the Thomson optical depth along the photon path and $\kappa(x_e, T_e)$ describes the free-free emissivity of the plasma, which depends on the electron temperature¹⁵, ionization degree, composition and weakly on frequency (through the Gaunt-factor). Assuming that the ambient radiation field is given by the CMB blackbody, $n_0 \approx 1/(e^x - 1)$ with $x = hv/kT_0$, we have

$$\delta n(x_e) \approx \delta \tau \frac{\kappa(x_e, T_e) e^{-x_e}}{x_e^3} \frac{e^x - e^{x_e}}{e^x - 1} = \delta \tau \frac{\kappa(x_e, T_e)}{x_e^3} \frac{e^{x-x_e} - 1}{e^x - 1}, \quad (A2)$$

which implies a suppression factor $f(x, x_e) = (e^{x-x_e} - 1)/(e^x - 1)$ from stimulated CMB emission and CMB absorption.

¹⁵ A single temperature plasma is assumed

APPENDIX B: BASIS FOR FREE-FREE SPECTRA

Here we give the first few functions, $G_k^{(l,m)}$, defined by Eq. (32). They read

$$\begin{aligned}
 G_1^{(1,0)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{2}{3}, 0, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_1^{(2,0)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{2}{9}, -\frac{2}{3}, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_1^{(3,0)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{4}{81}, \frac{2}{9}, -\frac{1}{3}, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \quad (\text{B1}) \\
 G_1^{(4,0)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{2}{243}, -\frac{4}{81}, \frac{1}{9}, -\frac{1}{9}\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(2,0)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{2}{9}, 0, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(3,0)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{4}{27}, \frac{2}{9}, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(4,0)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{14}{243}, -\frac{4}{27}, \frac{1}{9}, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_1^{(0,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= (1, 0, 0, 0)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_1^{(1,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{2}{3}, 1, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_1^{(2,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{2}{9}, -\frac{2}{3}, \frac{1}{2}, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_1^{(3,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{4}{81}, \frac{2}{9}, -\frac{1}{3}, \frac{1}{6}\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(1,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{2}{3}, 0, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(2,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{2}{3}, -\frac{2}{3}, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(3,1)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{28}{81}, \frac{2}{3}, -\frac{1}{3}, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(0,2)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{1}{2}, 0, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(1,2)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(-\frac{2}{3}, \frac{1}{2}, 0, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) \\
 G_2^{(2,2)}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu) &= \left(\frac{4}{9}, -\frac{2}{3}, \frac{1}{4}, 0\right)^T \cdot \mathbf{g}_{\text{ff}}(\bar{\nu}_{\text{ff}}, \bar{\zeta}, \nu),
 \end{aligned}$$

where we defined $\mathbf{g}_{\text{ff}} = (g_{\text{ff}}^{(1)}, g_{\text{ff}}^{(2)}, g_{\text{ff}}^{(3)})$. These are needed for a fourth order moment expansion. We also have $G_1^{(0,0)} = G_2^{(0,0)} = G_2^{(1,0)} = G_2^{(0,1)} = 0$.

APPENDIX C: DERIVATIVES OF PLANCKIAN

In this section we give a closed form for the derivatives of a Planckian, $n^{\text{pl}}(\nu, \beta) = (e^{h\nu\beta/k} - 1)^{-1}$ with respect to $\beta = 1/T$. Since for $x = h\nu\beta/k$ we have $\partial_\beta x = h\nu/k$, we can simply write

$$\beta^j \partial_\beta^j n^{\text{pl}}(\nu, \beta) = x^j \partial_x^j n^{\text{pl}}(x) = n^{\text{pl}}(\nu, \beta) \frac{(-x)^j}{(1 - e^{-x})^j} \sum_{m=0}^{j-1} \left\langle \begin{matrix} j \\ m \end{matrix} \right\rangle e^{-mx} \quad (\text{C1})$$

where we used the result of Chluba et al. (2012b) for $x^j \partial_x^j n^{\text{pl}}(x)$. Here $\left\langle \begin{matrix} j \\ m \end{matrix} \right\rangle$ denotes the Eulerian numbers, which can be computed using recursion relations (compare Chluba et al. 2012b).