Rethinking CMB foregrounds: systematic extension of foreground parametrizations

Jens Chluba,1* James Colin Hill2 and Maximilian H. Abitbol3

1Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
2Department of Astronomy, Columbia University, Pupin Hall, New York, NY 10027, USA
3Department of Physics, Columbia University, New York, NY 10027, USA

Accepted 2017 August 1. Received 2017 July 31; in original form 2017 January 6

ABSTRACT

Future high-sensitivity measurements of the cosmic microwave background (CMB) anisotropies and energy spectrum will be limited by our understanding and modelling of foregrounds. Not only does more information need to be gathered and combined, but also novel approaches for the modelling 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, grey-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 modelling of X-ray emission in clusters of galaxies.

Key words: cosmic background radiation – cosmological parameters – early Universe – cosmology: theory.

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 XIII 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–Zel’dovich, SZ effect; Abazajian et al. 2016). In addition, CMB spectral distortion measurements, e.g. with $\text{PIXIE}$ (Kogut et al. 2011) or $\text{APSERa}$ (Sathyarayana Rao et al. 2015), could provide a wealth of new complementary information about the Universe we live in (e.g. Chluba & Sunyaev 2012; Chluba 2013, 2016; Sunyaev & Khatri 2013; Tashiro 2014).

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 multifrequency 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 XII 2014a; Planck Collaboration IX 2016a); however, at the required level of future foreground modelling we cannot reliably predict what awaits us and restrictive assumptions about the foreground model can cause biases in the deuced parameters (e.g. Errard et al. 2016; Kogut & Fixsen 2016; Remazeilles et al. 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 modelling for primordial CMB distortions were presented in Abitbol et al. (2017).

In this work, we approach the foreground modelling 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 parametrizations 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 \( y \approx \frac{1}{2} \left( \frac{\Delta T^2}{T_0^2} \right) \), where \( \Delta T^2 \) 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 \( (y \approx 10^{-5}) \), which in principle can be accurately modelled using prior information from Planck. Still, this example already illustrates how averaging processes can affect the required parametrization 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 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 (Itoh, Kohyama \\& Nozawa 1998; Sazonov \\& Sunyaev 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, \( \langle k T e \rangle \), of the electron temperature distribution. Conversely, maps of these new observables contain all the information gained from multifrequency 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. Fuskeland et al. 2014; Planck Collaboration XII 2014a; Planck Collaboration IX 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 (→ 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, grey-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 grey-body spectra (Section 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 modelling). One can furthermore include more complicated fundamental spectra [e.g. to parametrize the anomalous microwave emission, AME (Draine \\& Lazarian 1998; Ali-Ha¨ımoud, Hirata \\& Dickinson 2009; Hoang, Vnth \\& Quynh Lan 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 (Section 2.1) from the effect of spatial averages carried out on measured sky signals (Section 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, \( \delta I(\hat{\mathbf{r}}, p) \), in a direction \( \hat{\mathbf{r}} \), which varies with frequency \( v \) and depends on a list of spatially varying...
parameters \( p(r) \) at different locations, \( r \), along the line of sight. The line of sight can be parametrized by an affine parameter \( s(\hat{y}) \) in the direction \( \hat{p} \), so that the line-of-sight-averaged SED is

\[
I_s(\hat{y}) \equiv (I_s(\hat{y}, p))_{\text{los.}} = \int \frac{dI_s(\hat{y}, p)}{ds(\hat{y})} \, ds(\hat{y}),
\]

(1)

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

\[
X(\hat{y}_c, v_c) = \langle X(\hat{y}, v) \rangle = \frac{1}{V} \int d^2 \hat{y} \, W(\hat{y}, v) \, X(\hat{y}, v),
\]

(2)

where \( W(\hat{y}, v) \) describes the frequency-dependent beam of different channels centred at frequency, \( v_c = \langle v \rangle \equiv \int v \, W(\hat{y}, v) \, dv \, d^2 \hat{y} \), and in the average direction \( \hat{y}_c = \langle \hat{y} \rangle \equiv \int \hat{y} \, W(\hat{y}, v) \, dv \, d^2 \hat{y} \). The specific shape of \( W(\hat{y}, v) \) depends on the experiment, as discussed below. We normalize the beam as \( \int W(\hat{y}, v) \, d^2 \hat{y} \, dv \equiv \langle 1 \rangle = 1 \).

We reiterate, the guiding picture is that for any observation in an average direction \( \hat{y}_c \), the signal from different emitters\(^1\) 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, \( p \). 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 averages 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 modelling, 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 multidimensional Taylor expansion of the underlying SED in the free model parameters \( I_0(p) = I_s(\hat{p}) + \sum_i (p_i - \bar{p}_i) \partial_{\bar{p}_i} I_s(\hat{p}) \)

\[
+ \frac{1}{2!} \sum_{i,j} (p_i - \bar{p}_i)(p_j - \bar{p}_j) \partial_{\bar{p}_i} \partial_{\bar{p}_j} I_s(\hat{p})
\]

\[
+ \frac{1}{3!} \sum_{i,j,k} (p_i - \bar{p}_i)(p_j - \bar{p}_j)(p_k - \bar{p}_k)
\]

\[
\times \partial_{\bar{p}_i} \partial_{\bar{p}_j} \partial_{\bar{p}_k} I_s(\hat{p}) + \ldots
\]

(3)

where \( \hat{p} \) (only depending on \( \hat{y}_c \)) describes the list of beam average parameters that need to be specified. In the notation, we suppressed the spatial dependence of \( p_i = p_i(r) \) for convenience. We also use the notation \( \partial_{\bar{p}_i} \ldots \partial_{\bar{p}_j} I_s(\hat{p}) = \partial_{\bar{p}_i} \ldots \partial_{\bar{p}_j} I_s(\hat{p}) \)

By fixing the average parameters, \( \bar{p} \), for a given spectral distribution function, we require the first-order terms in the Taylor expansion to vanish, \( \langle \sum_i (p_i(r) - \bar{p}_i) \partial_{\bar{p}_i} I_s(\hat{p}) \rangle = 0 \). This has the solution

\[
\bar{p}_i = \langle p_i(r) \partial_{\bar{p}_i} I_s(\hat{p}) \rangle
\]

(4)

for \( \langle \partial_{\bar{p}_i} I_s(\hat{p}) \rangle \neq 0 \). Generally, this does not imply \( \bar{p}_i = \langle p_i(r) \rangle \) since the frequency dependence of the beam can couple different spatial regions, where the frequency-weighting function is modified by \( \partial_{\bar{p}_i} I_s(\hat{p}) \).

\(^1\) 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.

modified by \( \partial_{\bar{p}_i} I_s(\hat{p}) \). If \( W(\hat{y}, v) = B(\hat{y}) F(v) \), one can separate the frequency and spatial averages, i.e. \( \langle p_i(r) \partial_{\bar{p}_i} I_s(\hat{p}) \rangle = \langle p_i(r) \rangle \langle \partial_{\bar{p}_i} I_s(\hat{p}) \rangle \), such that we recover \( \bar{p}_i = \langle p_i(r) \rangle \). Here, \( B(\hat{y}) \) determines the spatial shape of the beam, which does not vary at different frequencies within each channel, and \( F(v) \) 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 grey-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 \( p_i \) 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

\[
o_{i,j} = \langle (\bar{p}_i(r) - \bar{p}_i) \ldots (\bar{p}_j(r) - \bar{p}_j) \partial_{\bar{p}_i} \ldots \partial_{\bar{p}_j} I_s(\hat{p}) \rangle
\]

(5)

and spectral functions

\[
I_{i,j}(v_c, \bar{p}) = \langle \partial_{\bar{p}_i} \ldots \partial_{\bar{p}_j} I_s(\bar{p}) \rangle,
\]

(6)

where \( v_c \) is the central frequency of the considered channel. Here the ‘…’ 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, \( I_{i,j}(v_c, \bar{p}) \), 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

\[
\langle I_s(\hat{p}) \rangle = I_s(v_c, \bar{p}) + \sum_i \omega_i I_s(v_c, \bar{p}) + \frac{1}{2} \sum_{i,j} \omega_{ij} I_{i,j}(v_c, \bar{p})
\]

\[
+ \frac{1}{6} \sum_{i,j,k} \omega_{ijk} I_{i,j,k}(v_c, \bar{p}) + \ldots
\]

(7)

where \( \omega_i \equiv 0 \) with the definition of \( \bar{p} \) given in equation (4). With this reformulation, we have identified a set of new parameters, \( \omega_{i,j} \), and spectral functions, \( I_{i,j}(v_c, \bar{p}) \), which can be used to describe the beam-averaged signal. For a given theoretical model, specified by the fundamental SED, \( I_s(p) \), one can simply compute these quantities and use them in the modelling 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 \( I_s \). 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, \( \omega_{i,j} \), in general depend on the band on a non-trivial way. This means
that within each band different sets of moments (with different weighting functions) contribute, unless the beam factorization \(W(\hat{\nu}, \nu) = B(\hat{\nu}) F(\nu)\) is possible, such that
\[
\omega_m = \{p_1(r) - \bar{p}_1 \ldots [p_1(r) - \bar{p}_1]\}.
\]
(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 modelled, 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 separate the different components by using multifrequency observations. This is indeed one of the big worries for future CMB observations. This is indeed one of the big worries for future CMB observations. It also renders a comparison

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 parametrizations, 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 parametrization 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 grey-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 \(X\)  are given by
\[
\omega_{lm}(\nu_c) = \int \hat{d}^2 \hat{\gamma}_c \ Y^*_{lm}(\hat{\gamma}_c) \bar{X}(\nu_c, \hat{\gamma}_c),
\]
(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 \(\nu_c\). We can see from equation (7) that two sources of spatial variations appear when carrying out the spherical harmonic expansion, (i) those from the moments, \(\omega_{lm}(\hat{\gamma}_c)\), and (ii) those from the variation of the mean parameters, \(\bar{p}(\hat{\gamma}_c)\) in different directions on the sky.

By simply following through all the steps that lead to the moment expansion (equation 7), we readily have the generalization
\[
\bar{I}_\nu(\hat{\gamma}_c) = I(\nu_c, \bar{p}_\circ) + \sum_m \omega_m(\hat{\gamma}_c) I_\nu(\nu_c, \bar{p}_\circ)
\]
\[
+ \frac{1}{2} \sum_l \sum_{m} \omega_{lm}(\hat{\gamma}_c) I_{lm}(\nu_c, \bar{p}_\circ)
\]
\[
+ \frac{1}{6} \sum_{lm} \sum_{jk} \omega_{lk}(\hat{\gamma}_c) I_{lmk}(\nu_c, \bar{p}_\circ) + \ldots\]
(10)

where \(\bar{p}_\circ\) is the sky-averaged parameter vector, which can be obtained by averaging equation (4) over all directions. Similarly, the moments follow from equation (5), such that overall we find
\[
\bar{p}_{\nu,0} = \int \frac{\omega_c(\hat{\nu}, \hat{\gamma}_c) d^2 \hat{\gamma}_c}{4\pi} = \langle \frac{\delta_{\nu,0} \bar{I}_\nu(\nu_c, \bar{p}_\circ)}{\delta_{\nu,0} \bar{I}_\nu(\nu_c, \bar{p}_\circ)} \rangle,
\]
(11)
where the average \(\langle \ldots \rangle\) connects \(\hat{\gamma}_c\). Through the beam shape. The spherical harmonic coefficients of the moments are obtained using equation (9). The most important difference is that while by construction \(\int \omega_c(\hat{\nu}, \hat{\gamma}_c) d^2 \hat{\gamma}_c = 0\), in general \(\omega_c(\hat{\nu}, \hat{\gamma}_c)\) no longer vanishes but fluctuates across the sky. Usually, this term is considered, but higher order moments lead to new scale-dependent SED variations as we illustrate below. Also, in general \(\omega_{lm}(\hat{\gamma}_c) \neq \omega_{lm}(\hat{\nu})\), as the reference parameters were chosen using the sky average rather than the local quantity. With \(W(\hat{\nu}, \nu) = B(\hat{\nu}) F(\nu)\), we have
\[
\bar{p}_{\nu,0} = \int \bar{p}(\hat{\nu}) \frac{d^2 \hat{\gamma}_c}{4\pi} = \langle \{p_0(r, \hat{\gamma}_c) - \bar{p}_0, 0\} \rangle.
\]
(12)
For the first and second moments, we then find
\[
\omega_{0,0}(\hat{\gamma}_c) = \langle \{p_0(r, \hat{\gamma}_c) - \bar{p}_0, 0\} \rangle = \bar{p}(\hat{\gamma}_c) - \bar{p}_0
\]
\[
\omega_{0,0}(\hat{\gamma}_c) = \langle \{p_0(r, \hat{\gamma}_c) - \bar{p}_0, 0\} \rangle = \omega_{0,0}(\hat{\nu}) + \bar{p}(\hat{\gamma}_c) - \bar{p}_0
\]
(13)
which directly shows how the fluctuations in \(\bar{p}_\circ\) 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 (Section 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 \(p = \{A_0, \alpha\}\) are assumed to vary spatially. This description is typically used to model the foreground caused by synchrotron emission at low frequencies (Planck Collaboration XII 2014a; Planck Collaboration IX 2016a). One common extension is to introduce curvature to the spectral index, \(I_\nu = A_0 (\nu/\nu_0)^{\alpha} + \beta (\nu/\nu_0)^{\alpha/2}\). Let us try to see under which conditions this approximation works by using the moment method. For this, we first compute the derivatives of \(I_\nu\) with respect to \(A_0\) and \(\alpha\):
\[ \frac{\partial^k}{\partial \alpha^k} I_0(A_0, \alpha) = I_0(A_0, \alpha) \ln^k(v/v_0). \]  
(15)

Since \( I_0 \) depends linearly on \( A_0 \), only first-order derivatives with respect to \( A_0 \) appear. Since the derivatives \( \partial A_0 \) and \( \partial \alpha \) commute, using equation (7), we thus find the expansion

\[ \langle I_0(p) \rangle = I_0(v, \bar{p}) + \omega_{12} \omega_{22} I_{22}(v, \bar{p}) + \frac{1}{2} \omega_{222} I_{222}(v, \bar{p}) + \frac{1}{6} \omega_{2222} I_{2222}(v, \bar{p}) + \ldots \]  
(16)

For the spectral functions, we then have

\[ \bar{A}_0 \omega_{12}(v, \bar{p}) \equiv I_{22}(v, \bar{p}) = \langle \partial^k_\alpha I_0(p) \rangle = \langle I_0(p) \ln^k(v/v_0) \rangle. \]  
(17)

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

\[ \omega_{12} = \frac{\langle A_0 \rangle - \bar{A}_0}{\langle \alpha \rangle - \bar{\alpha}} \approx \frac{\langle A_0 \rangle - \bar{A}_0}{\langle \alpha \rangle - \bar{\alpha}} \]  
(18)

\[ \omega_{22} = \frac{\langle \alpha \rangle - \bar{\alpha} \rangle}{\langle \alpha \rangle - \bar{\alpha}} \approx \frac{\langle \alpha \rangle - \bar{\alpha} \rangle}{\langle \alpha \rangle - \bar{\alpha}} \]  
(19)

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

\[ \langle I_0(p) \rangle = I_0(v, \bar{p}) + \omega_{12} I_2(v, \bar{p}) + \frac{1}{2} \omega_{22} + \frac{1}{2} \omega_{222} I_{222}(v, \bar{p}) + \ldots \]  
(20)

We therefore only need to consider the spectral functions \( I_{22}(v, \bar{p}) \), and the new set of moments

\[ \omega_{12} = \frac{\omega_{12} + \omega_{12} \bar{A}_0}{A_0} = \frac{\langle A_0 \rangle \langle \alpha \rangle - \bar{\alpha} \rangle}{\langle \alpha \rangle - \bar{\alpha}} \approx \frac{\langle \alpha \rangle - \bar{\alpha} \rangle}{\langle \alpha \rangle - \bar{\alpha}} \]  
(21)

This shows that the amplitude \( A_0(\alpha) \) 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, \( \propto \omega_{12} I_2(v, \bar{p}) \), leading to \( \bar{\alpha} \approx \bar{\alpha} + \omega_{12} \bar{A}_0 \), which we have to use instead of \( \alpha \) in the moments, \( \omega_{12} \). If \( \omega_{12} = 0 \), the correlation term \( \omega_{12} \approx \omega_{12} \) is not relevant. 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. \( W(\bar{\gamma}, v) = B(\bar{\gamma}) \delta(v - v_0) \), we have

\[ I_{22}(v, \bar{p}) = I_0(v, \bar{p}) \ln^k(v/v_0) \]  
(22)

\[ \omega_{22} = \frac{\omega_{22} \bar{A}_0}{A_0} = \frac{\omega_{22}}{A_0} \approx \frac{\omega_{22}}{A_0} \]  
(23)

This shows that \( \beta \), which is the distribution of \( \alpha(\sigma) \) Gaussian, such that \( \omega_{22} \approx (k - 1)! (\omega_{22})^{k/2} \) for even \( k \) and zero otherwise.\(^2\) 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 equation (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

\[ I_{22} \approx A_0(v, v_0)^\beta \ln^k(v/v_0) + \frac{1}{6} \omega_{22} I_{222}(v, \bar{p}) + \ldots \]  
(24)

\[ I_{22} \approx A_0(v, v_0)^\beta \ln^k(v/v_0) + \frac{1}{6} \omega_{22} I_{222}(v, \bar{p}) + \ldots \]  
(25)

To develop some intuition for how the moment expansion can be applied, let us consider the simple sum of two power-law spectra, \( p = [A_1, \alpha_1, A_2, \alpha_2] \). 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

\[ \bar{A} = A_1 + A_2 \]  
(26)

\[ \bar{\alpha} = \frac{A_1 \alpha_1 + A_2 \alpha_2}{A_1 + A_2} = f \alpha_1 + (1 - f) \alpha_2 \]  
(27)

\[ \omega_{22} = \frac{A_1 \alpha_1 - \bar{\alpha}}{A_1 + A_2} \]  
(28)

\[ \omega_{12} = \frac{A_1 \alpha_1 - \bar{\alpha}}{A_1 + A_2} \]  
(29)

This is no longer true when difference of power-law spectra is allowed.
The expressions in equation (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 $\alpha_i$ in the beam is $f = A_i/(A_1 + A_2)$ and the one for $\alpha_2$ is $\bar{f} = 1 - f$. We thus have the probability distribution function (PDF), $P(\alpha) = \bar{f} \delta(\alpha - \alpha_2) + f \delta(\alpha - \alpha_1)$. Then the average spectral index is, $\bar{\alpha} = \int P(\alpha) \alpha \, d\alpha = f \alpha_1 + \bar{f} \alpha_2$, and similarly $\alpha_{2,\bar{f}} = \int P(\alpha) (\alpha - \bar{\alpha}) \, d\alpha = f(\alpha_1 - \bar{\alpha}) + \bar{f}(\alpha_2 - \bar{\alpha})$, as also found in equation (24). Only that in general the PDF is a much more complicated multidimensional function. We now discuss the performance of different parametrizations. For a two power-law model, the convergence is perfect already for four parameters. How- ever, the converged of the foreground parametrizations as a function of frequency (top panel) and the convergence of the foreground parametrizations (lower panel) for different models, $P_i$. The two parametrizations equations (21) and (23) are used with $k$ denoting the maximal moment index ($k = 2$ for $\alpha_{2,\bar{f}}$ and $\beta$, $k = 3$ for $\alpha_{3,\bar{f}}$ and $\gamma$, etc.).

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 (equation 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 equations (21) and (23) to represent the different examples and determined the free parameters using a noiseless $\chi^2$ minimization. For the moment expansion (equation 21), we can explicitly compare the obtained best-fitting values with the theoretical values (equation 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

---

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 | Theor | Two parameters | Three parameters | Five parameters |
|-----------|-------|----------------|------------------|----------------|
| $A/C$     | 1     | 1.039          | 0.977            | 1.000          |
| $a$       | $-0.4$| $-0.169$       | $-0.399$         | $-0.400$       |
| $\omega_{22}$ | 0.25  | $-0.301$       | 0.248            |                |
| $\omega_{222}$ | 0    | $-0.001$       |                 | 0.070          |

use the simpler approximation (Draine 2011)

$$
\begin{align*}
g_{\theta}(Z_i, T_e, v) & = 1 + \ln \left[ 1 + \left( \frac{\nu_{th}(Z_i, T_e)}{v} \right)^2 \right] \\
v_{th}(Z_i, T_e) & \approx \frac{255.33 \text{ GHz}}{Z_i} \left( \frac{T_e}{10^5 \text{ K}} \right)^{3/2},
\end{align*}
$$

(27)

where $Z_i$ is the charge of the nucleus of ion $i$. Neglecting the high-frequency suppression $[f(x, x_c) = 1]$, we can perform a simple moment expansion in $\xi = \ln \nu_{th}$. With $\gamma = \sqrt{x_c} / \xi$, $\eta = (\nu_{th}/v)^2$ and

$$
\begin{align*}
\delta \xi g_{\theta} & = -\gamma \eta (1 + \eta)^2 \\
\delta^2 \xi g_{\theta} & = -\gamma^2 \eta (1 + \eta)^2,
\end{align*}
$$

(28)

for the first few derivatives of $g_{\theta}$ with respect to $\xi$, we then have

$$
\begin{align*}
\langle f_{\theta}^{(k)}(\nu) \rangle & \approx \langle f_{\theta}(\nu_{th}, v) \rangle + \sum_{k=2}^{\infty} \frac{\langle f_{\theta}^{(k)}(\nu_{th}, v) \rangle}{k!} \omega_{\theta}^{(k)} \xi \omega_{\theta}^{(k)} \\
\omega_{\theta}^{(k)} & = \left( \sum_i \epsilon_i \ln \nu_{th} \right), \quad \nu_{th} = \left( \sum_i \epsilon_i \ln \nu_{th} \right) / \epsilon,
\end{align*}
$$

(29)

Here, we introduced the parameter $\epsilon_i = h c \delta \alpha Z_i^2 N_i / (\sqrt{6} \pi^3 \epsilon / m_e)^{1/2}$, which just modifies the relativistic emission of the different ions and can be kept as a 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 modelled as variations in $T_e$ $T_e = Z_i / Z_i^{(2)}$, or equivalently, variations in $\nu_{th}$. This moment expansion converges very rapidly; however, it fails to accurately capture the high-frequency suppression at $v \gtrsim 1$ THz.

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 $f(x, x_c)$, we need $T_e$, which can be expressed as $T_e = \kappa \nu_{th}^{2/3} Z_i^{1/3}$, with $\kappa \approx 24.851$ K GHz$^{-2/3}$. We can then define $T_e$ using $\ln \nu_{th}$ and $Z_i$ to determine $\bar{x}_e = x T_e / T_i$. Since $\rho_c = \nu_{th} / T_e \ll 1$, we can write

$$
\begin{align*}
f(x, x_c) & = f(x, \bar{x}_e) + \sum_{k=2}^{\infty} \frac{J_k(x, \bar{x}_e) (\rho_c - \bar{\rho}_c)^k}{k!} \\
J_k(x, x_c) & = (-1)^k \frac{x^{k-1} e^{-x}}{e^{x_c} - 1}.
\end{align*}
$$

(30)

For CMB applications, it should be sufficient to include the first two terms in $\Delta \rho_c$. To perform the moment expansion in

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

\footnote{We also tried an expansion in $T_e$ and $1/T_e$, but found this to converge quite slowly. Using $\ln \nu_{th}$ as a parameter is also motivated by the fact that most of the emission is found at low frequencies where one has $g_{\theta}(Z_i, T_e, v) \approx 1 + \sqrt{2} \ln \left( \nu_{th}/v \right)$.}
In \( v_\text{eff} \) and \( Z_\text{e} \), we define \( \zeta = Z_\text{e}^{-2/3} \). This yields
\[
\langle \Delta I^\nu \rangle_{\text{high}} = \langle \Delta I^\nu \rangle f(x, \bar{x}) + \langle \Delta I^\nu \rangle \frac{1}{Z^{2/3}}, \quad \bar{\zeta} = \frac{1}{Z^{2/3}}.
\]
where the additional correction \( \langle \Delta I^\nu \rangle \) is obtained from a moment expansion of \( \Delta I^\nu \). The high-frequency suppression is simply captured by multiplying \( \langle \Delta I^\nu \rangle \) in equation (29) with \( f(x, \bar{x}) \). However, variations in the temperature and composition do require a more detailed treatment. The derivations are straightforward but cumbersome. Defining the frequency-dependent functions
\[
G^{(l,m)}(\nu, \bar{\zeta}, \bar{x}) = \frac{\bar{x}^m}{\rho^l} \sum_{k} \frac{\partial^l A^m_{\nu}(\rho^l - \bar{x}^l)^k}{k! l! m!}, \quad G^{(l,m)}(\nu, \bar{\zeta}, \bar{x}) = \langle \Delta I^\nu \rangle_{\text{high}}(\nu, \bar{\zeta}, \bar{x}) \sigma^{(l,m)}_{\nu},
\]
we write
\[
\langle \Delta I^\nu \rangle = \sum_{l,m} \int_{0}^{\infty} \sum_{k} \int_{0}^{\infty} \int_{0}^{\nu} G^{(l,m)}(\nu, \bar{\zeta}, \bar{x}) \sigma^{(l,m)}_{\nu} \frac{1}{\bar{\zeta}}.
\]
We note that \( \sigma^{(l,m)}_{\nu} \equiv \sigma_{\nu}^{(l,m)} \) and \( \sigma^{(0)}_{\nu} = \sigma^{(1)}_{\nu} = 0 \) by construction. The first few functions necessary for a fourth-order expansion in \( \ln \nu \) and \( Z_\text{e} \) are listed in Appendix B. At low frequencies, \( f(x, \bar{x}) \rightarrow 1 \). In this case, no new spectral shapes are added through \( \langle \Delta I^\nu \rangle \) and the number of free parameters, equivalent to the number of independent spectral functions, can be reduced to only \( \bar{\zeta}, \bar{\eta}, \bar{\xi} \).

Consequently, only the high-frequency tail \( (\nu \gtrsim 1 \text{ GHz}) \) can be used to obtain direct information about \( Z_\text{e} \), which specifies the chemical composition of the medium and its ionization state, through the moments \( \sigma^{(l,m)}_{\nu} \). But how many coefficients \( \sigma^{(l,m)}_{\nu} \) can be principle be constrained independently? The functions \( G^{(l,m)}(\nu, \bar{\zeta}, \bar{x}) \) are all just linear combinations of \( g_\nu \) and its derivatives, \( g_\nu^{(j)} \), up to \( j = 3 \) (see Appendix B). These would be degenerate with simple variations of \( v_\text{eff} \), having the same spectral dependence as in the expansion (equation 29). However, the functions \( J_k(x, \bar{x}) \) introduce new spectral shapes that can in principle be used to break the degeneracy. Up to fourth order in \( \ln \nu \) and \( Z_\text{e} \) \( (l + m \leq 4) \) and second order in \( \bar{\zeta}, \bar{\eta}, \bar{\xi} \), one thus finds seven linearly independent spectral shapes. The function \( J_1(x, \bar{x}) G^{(1,1)}(\nu, \bar{\zeta}, \bar{x}) \) fixes \( \bar{\zeta} \). Since \( \sigma^{(0,0)}_{\nu} \equiv \sigma^{(0,1)}_{\nu} \equiv \sigma^{(1,0)}_{\nu} \equiv \sigma^{(1,1)}_{\nu} \) (three parameters) are determined through \( \langle I^\nu \rangle f(x, \bar{x}) \), only six additional \( \sigma^{(l,m)}_{\nu} \) remain unspecified. This means that all \( \sigma^{(l,m)}_{\nu} \) 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 \( J_1 \approx \text{const} \) 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 IX 2016a), the free–free emission is modelled using a single-temperature approximation with \( v_\text{eff}(T_\text{e}) \approx 4727 \text{ GHz} \) for \( T_\text{e} = (7000 \pm 11) \text{ K} \). The high-frequency suppression is neglected, such that one can write
\[
I^\nu_{\nu} \approx \epsilon_{\nu} g_{\nu}(v_\text{eff}, v),
\]
where \( \epsilon_{\nu} \approx 150 \text{ Jy sr}^{-1} \) (Planck Collaboration IX 2016a). To include the high-frequency suppression, we simply multiply by \( f(x, \bar{x}) \). These two cases are shown in Fig. 2 together with a simple power-law fit \( I^\nu_{\nu} \approx v^{-0.14} \) (and \( T_\text{e} \approx v^{-2.14} \)), 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 (equation 21). We find
\[
I^\nu_{\nu} \approx \frac{3.229}{v_{100}^{1515}} \epsilon_{\nu} \left[ 1 - 6.129 \times 10^{-3} \ln^2 v_{100} \right]
\]
with \( v_{100} = \frac{\nu}{100 \text{ GHz}} \) to work to \( \lesssim 0.5 \) per cent precision in the frequency range \( 1 \text{ GHz} \lesssim \nu \lesssim 6 \text{ GHz} \). For better than \( \approx 0.06 \) per cent precision, we obtain the fourth-order power-law moment expansion
\[
I^\nu_{\nu} \approx \frac{3.235}{v_{100}^{1527}} \epsilon_{\nu} \left[ 1 - 7.116 \times 10^{-3} \ln^2 v_{100} + 8.701 \times 10^{-5} \ln^3 v_{100} + 6.157 \times 10^{-3} \ln^4 v_{100} \right].
\]
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 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 modelling free–free and synchrotron independently is thus really driven by external priors but not because the data distinguish 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 (equations 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, \( P_1 = \{ \bar{\epsilon}_1, \bar{T}_{\text{e},1}, \bar{\epsilon}_2, \bar{T}_{\text{e},2} \} = \{ 0.5 \epsilon_{\nu}, 7000 \text{ K}, 0.5 \epsilon_{\nu}, 500 \text{ K} \} \), representing a mix of hot and cold components. For now, we assume \( Z_{\text{e}} = 1 \). 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 \( \epsilon_{\nu} \approx 1/\sqrt{T_{\text{e}}} \) 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\(^8\) Rayleigh–Jeans temperature is obtained as \( T_{\text{e}} = c^2 I^\nu_{\nu} / (2k \nu^2) \).

\(^8\) The Rayleigh–Jeans temperature is obtained as \( T_{\text{e}} = c^2 I^\nu_{\nu} / (2k \nu^2) \).
Figure 3. Free–free intensity (in units of \(\epsilon_{\|}\)) for different models. For comparison, we also show the \(T_e = 7000\) K Planck model (equation 34), with the high-frequency suppression added. The change in the normalization is absorbed when determining \(\epsilon_{\|}\), 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.

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 \(Z_{e,T} = 1\).

| Parameter   | Theory | Two parameters | Three parameters | Five parameters |
|-------------|--------|----------------|-----------------|----------------|
| \(\bar{\nu}\) (GHz) | \(1\) | 0.8991 | 1.001 | 0.999 |
| \(\nu_{\|}\) | 653.18 | 1600.4 | 651.88 | 658.32 |
| \(\omega_{\|}^{(2)}\) | 3.9176 | - | 3.659 | 3.835 |
| \(\omega_{\|}^{(3)}\) | 0 | - | - | -0.276 |
| \(\omega_{\|}^{(4)}\) | 15.348 | - | - | 11.054 |

We can again compare the obtained best-fitting values, with the theoretical predictions for the moments. For a general two-component case, \(P = \{C_0(1-f), T_{e,1}, C_0 f, T_{e,2}\}\), we have

\[
\bar{\epsilon} = C_{\|} \ln \bar{\nu}_{\|} = (1-f) \ln \nu_{\|}(T_{e,1}) + f \ln \nu_{\|}(T_{e,2})
\]

\[
\omega_{\|}^{(k)} = (1-f) \ln^k \nu_{\|}(T_{e,1})/\bar{\nu}_{\|} + f \ln^k \nu_{\|}(T_{e,2})/\bar{\nu}_{\|}
\]

\[
\bar{\nu}_{\|} = \frac{\nu_{\|}(T_{e,1})}{e^{\epsilon_{\|}(T_{e,1})} - 1} + \frac{\nu_{\|}(T_{e,2})}{e^{\epsilon_{\|}(T_{e,2})} - 1}
\]

\[\epsilon_{\|} = T_{\text{Planck}}/k\]

\[
\bar{\nu}_{\|}^{2/3} \approx 1.097 \times 10^{-3} \left[\frac{\bar{\nu}_{\|}}{10^2 \text{ GHz}}\right]^{-2/3}
\]

\[
\sigma_{\|}^{(0,k)} = \omega_{\|}^{(k)}, \sigma_{\|}^{(l,m)} = 0 \quad \text{for} \ m > 0
\]

For model \(P_1\), we find \(\nu_{\|} = 653.18\) GHz, \(\bar{\nu}_{\|} = 1.04571 \times 10^{-3}\), \(\omega_{\|}^{(2)} = 3.9176\), \(\omega_{\|}^{(4)} = 15.348\) and \(\omega_{\|}^{(6)} = 60.126\) (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 behaviour 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 \(\nu_{\|}\) and \(\bar{\nu}_{\|}\) (i.e., single-temperature model) in the moment expansion represents the low-frequency free–free spectrum very well, while underestimating the high-frequency spectrum by \(\approx 7\) per cent at \(\nu \approx 1\) THz. Including \(\omega_{\|}^{(2)} = 3.9176\) as additional parameter, the discrepancy decreases to \(\lesssim 0.6\) per cent at 1 GHz \(\lesssim \nu \lesssim 3\) THz. Adding the next non-vanishing moment, \(\omega_{\|}^{(4)} = 15.348\), the moment expansion captures the full spectrum at the \(\lesssim 0.02\) per cent level in the same range.

When also varying \(f\) for the two-temperature case, the approximation also requires inclusion of \(\omega_{\|}^{(3)} \neq 0\), which no longer vanishes. For the chosen extreme difference in temperatures, this representation is only accurate at the \(\approx 0.1\) per cent 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 modelling 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 modelling 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 per cent 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

\[
g_{\|}(Z_{e,T}, \nu) \approx 1 + \ln \left(\frac{\nu_{\|}(Z_{e,T}, \nu)}{\nu_{\|}}\right) \frac{\nu_{\|}}{\nu} = 1
\]

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

$$\ln \tilde{T}_e = \left( \frac{\sum \epsilon_i T_e^\nu \ln Y_{\nu}(T_e^\nu)} {\sum \epsilon_i \ln T_e} \right) = \left( \frac{\sum \epsilon_i \ln T_e} {\epsilon} \right),$$

with $T_e = T_e \tilde{C}_o$. This expression explicitly shows how the chemical composition enters into the problem and that the weighted average has to be carried out over $\epsilon_i \ln (T_e \tilde{C}_o)$. However, at high frequencies ($v \gtrsim 1$ THz), corrections can become noticeable at $\lesssim 0.1$ per cent 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 parametrizations 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 GREY-BODY SPECTRA

We can also apply the moment method to spatially varying grey-body spectra, with intensity $I_{\nu}(A_0, T) = A_0 \nu^\gamma / (\exp(\nu/kT) - 1)$. 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 $\beta = 1/T$, 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 simple weight factor which does not affect the spectral shape. For the first few derivatives with respect to $\beta$, we find

$$\beta \partial_\beta I_{\nu}(A_0, T) = -I_{\nu}(A_0, T) \frac{xe^x}{e^x - 1} x \coth(x/2),$$

$$\beta^2 \partial_\beta^2 I_{\nu}(A_0, T) = I_{\nu}(A_0, T) \frac{xe^x}{e^x - 1} x \left[ x \cosh(x) + 2 \right] \cosh(x) - 1,$$

$$\beta^3 \partial_\beta^3 I_{\nu}(A_0, T) = I_{\nu}(A_0, T) \frac{xe^x}{e^x - 1} x \left[ x \cosh(x) + 5 \right] \cosh(x/2),$$

$$\beta^4 \partial_\beta^4 I_{\nu}(A_0, T) = I_{\nu}(A_0, T) \frac{xe^x}{e^x - 1} x \left[ x \cosh(x) + 33 + 26 \cosh(x) \sinh^2(x/2) \right].$$

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

$$\langle I_{\nu} \rangle = \frac{A_0 \nu^\gamma}{e^x - 1} \left[ 1 + \frac{1}{2} \omega_{22} Y_2(x) + \frac{1}{6} \omega_{222} Y_3(x) \right] + \frac{1}{24} \omega_{2222} Y_4(x) + \frac{1}{120} \omega_{22222} Y_5(x) + \ldots \right\} \frac{1}{T} = \left( \frac{A_0(t)/T(t)} {A_0} \right) \omega_{22} \omega_{22} \right) = \frac{A_0(t)(T/T(t) - 1)^2} {A_0},$$

where now $x = \nu/kT$ and we defined $Y_2 = [(e-x)/e-x] I_1 / I_e$ (e.g. $Y_1(x) = e^x/(e+1) - 1$ 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 $y = \mu \omega_{22}$, but generalized to grey-body spectra. The effective moments, $\omega_{22} \omega_{22}$, are obtained using equation (20), with $\alpha \rightarrow 1/T$.

Below we illustrate how to apply the grey-body moment expansion for several cases. However, equation (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 $A_0 = (2\pi/c^2)$, 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, Khati & Sunyaev 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 $I_{\nu}(p^\gamma) \approx \text{const} \times \nu^\gamma$, 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 SZ 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 modelling of the low-frequency spectrum.

5.2 Examples for sums of two grey-body spectra

Let us consider the simple case of two grey-body spectra, with parameters $A_1, T_1$. We then have

$$\bar{A} = A_1 + A_2, \quad \bar{\beta} = \frac{fT_1 + 1 - f}{T_2},$$

$$\omega_{22} = \frac{[f(f - 1)^2 + (1 - f)^2]}{[(1 - f)T_1 + f T_2]^2}$$

with $f = A_1/(A_1 + A_2)$. For $A_1 = A_2 = A$ one has the moments $\omega_{22} = (T_1 - T_2)^2/(T_1 + T_2)^2$ for even $k$ and zero otherwise. At $x \gg 1$, with $Y_k(x) \approx x^k$ as expected we thus find $A \approx 2A \nu^\gamma e^{-x^2} \sinh(x(T_1 - T_2)/T_1 + T_2)^2$ for even $k$ and zero otherwise. At $x \gg 1$, with $Y_k(x) \approx x^k$ as expected we thus find $A \approx 2A \nu^\gamma e^{-x^2} \sinh(x(T_1 - T_2)/T_1 + T_2)^2$. 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 ($\rightarrow T_1$ and $T_2$ diverge strongly or $x$ become very large), it is difficult to use the high-frequency spectrum to constrain the low-frequency model (see also Section 5.1). We start with an example in the perturbative regime, choosing $T_1 = \{A_1, T_1, A_2, T_2\} = \{0.3A, 3K, 0.7A, 4K\}$. A single grey-body spectrum with $T_{\text{fit}} = 3.833K$ and amplitude $A_{\text{fit}} = 0.9329$ A represents the average SED at the level of $\lesssim 1$ per cent. The given best-fitting values are not far away from what is obtained with the moment expressions in equation (40), which imply...
\( \bar{T} = 3.6363 \) K and amplitude\(^{11} \) \( \bar{A} = A \). This approximation fails, however, at the level of \( \gtrsim 10 \) per cent at \( v \gtrsim 550 \) GHz, exceeding \( \gtrsim 50 \) per cent at about 2 THz. Adding the first moment, \( \omega_2 \), as a free parameter, the spectrum is represented at a level better than \( \gtrsim 1.5 \) per cent at all frequencies \( v \lesssim 1 \) THz and to better than \( \gtrsim 0.25 \) per cent at all frequencies \( v \lesssim 600 \) 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 grey-body spectra. The approximation was obtained by fitting the average SED assuming zero noise. The best-fitting parameters were \( R_1 = \{ A^0, \bar{T}, \omega_2^0 \} = \{ 0.9932 A, 3.6667 K, 0.01341 \} \). Comparing this to what is expected from the expressions in equation (40), \( M_i = \{ A, 3.6363 K, 0.01928 \} \), shows that the best-fitting parameters do not match perfectly. This is most obvious for the value of \( \omega_2 \), which deviates by \( \gtrsim 68 \) per cent. We have seen a similar behaviour for the superposition of power laws and free–free spectra (Sections 3.1.1 and 4.3). The source of this difference is that the best-fitting parameters for a given input model using varying representations in the analysis. In Fig. 4, we show the results for \( A_1 = 0.3 A, T_1 = 3 K \) and \( A_2 = 0.7 A \) varying \( T_2 \). We illustrate the relative difference with respect to the input spectrum at \( v = 30 \) GHz, representative of the behaviour in the Rayleigh–Jeans part of the spectrum, and \( v = 1 \) 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 behaviour 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 difference 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 upweights \( T_2 \) 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 behaviour of the moment expansion for grey-body spectra.

\(^{11}\) We will address the cause of the small difference in Section 5.2.1.

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

5.2.1 Best-fitting single grey-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 grey-body spectra with a single grey-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 equation 39). How can we estimate the best-fitting parameters for a single grey-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 \( N_i = A_i c_i T_i^3 \) and \( \rho_i = A_i c_{\gamma_i} T_i^4 \), where the values of two constants are not important. We then have two equations \( \langle A^* \rangle = \langle A, T_i^3 \rangle \) and \( A^* \langle T_i^4 \rangle = \langle A, T_i^3 \rangle \) when assuming a single grey-body approximation for the superposition. This yields \( A^* = \langle A, T_i^3 \rangle / \langle A, T_i^3 \rangle, \quad T^* = \langle A, T_i^3 \rangle / \langle A, T_i^3 \rangle \). (41)

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

The difference between the best-fitting values and those obtained with equation (39) can become much more dramatic. Choosing \( P_r = \{0.5 A, T_0, 0.7 A, 1.5 T_0, 0.4 A, T_0/1.5, 0.2 A, T_0/3.5\} \) with \( T_0 = 2.726 \) K, we find \( A = 1.8 A \) and \( \bar{T} = 2.165 \) K using the moment definitions (equation 39). From equation (41), we have \( A^\dagger = 1.1329 A \) and \( T^\dagger = 3.765 \) K and a fit gives \( A^{\dagger^2} = 1.1245 A \) and \( T^{\dagger^2} = 3.743 \) 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 (\( A, T_i \) instead of \( A_i \)) and choice of variables (\( T_i \) instead of \( 1/T \)) furthermore can lead to improved convergence properties (higher level of independence between moments). In spite of all this, the moment representation (equation 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 \( B = AT^2 \) and \( p = T \). In this case, the fundamental SED reads \( I_s(B, T) = B (\bar{T}/T)^{3} x^{3}/(e^{x/T} - 1) \), with the derivatives

\[
T \frac{\partial}{\partial T} I_s(B, T) = I_s(B, T) \left[ Y_1(x) - 3 \right]
\]

\[
T^2 \frac{\partial^2}{\partial T^2} I_s(B, T) = I_s(B, T) \left[ Y_3(x) - 8Y_1(x) + 12 \right]
\]

\[
T^3 \frac{\partial^3}{\partial T^3} I_s(B, T) = I_s(B, T) \left[ Y_5(x) - 15Y_3(x) + 60Y_1(x) - 60 \right]
\]

\[
T^4 \frac{\partial^4}{\partial T^4} I_s(B, T) = I_s(B, T) \left[ Y_4(x) - 24Y_2(x) + 180Y_1(x) - 360 \right]
\]

(42)

Defining the functions \( Z_k = T^k \frac{\partial}{\partial T} I_s(B, T)/I_s(B, T) \), we find the alternative moment expansion

\[
\langle I_s \rangle = \frac{\bar{B}}{e^x - 1} \left\{ 1 + 2 \kappa_{22}^g Z_2(x) + \frac{1}{6} \kappa_{2222}^g Z_3(x) + \frac{1}{24} \kappa_{22222}^g Z_4(x) + \ldots \right\}
\]

\[
\bar{B} = \langle B(r) \rangle \equiv \left\langle \frac{(A(r))T^3(r)}{(B(r))} \right\rangle \quad \bar{T} = \langle \frac{(B(r)T(r))}{B(r)} \rangle = \left\langle \frac{(A(r))T^3(r)}{(A(r))} \right\rangle \quad \kappa_{2222}^g = \left\langle \frac{(B(r)T(r)\bar{T} - 1)}{B} \right\rangle = \left\langle \frac{(A(r))T^3(r)}{A(r)T^3(r)} \right\rangle \quad \kappa_{22222}^g = \left\langle \frac{(B(r)(T(r)\bar{T} - 1)^4}{B} \right\rangle = \left\langle \frac{(A(r))T^3(r)(T(r)\bar{T} - 1)}{A(r)T^3(r)} \right\rangle.
\]

(43)

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

5.3 Orthonormal basis functions for grey-body spectra

Since the functions \( Y_k \) are 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 \( x = \hbar \nu / kT \), we can remove the explicit dependence on \( T \). Before the orthonormalization, we then have the basic spectral functions

\[
F_i(x) = \frac{x^3}{e^{x} - 1} Y_i(x)
\]

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

\[
(F_i(x) F_m(x)) = \int_0^\infty F_i(x) F_m(x) \, dx.
\]

(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 \( |F_0| = 2.5447 \), \( |F_1| = 9.4417 \), \( |F_2| = 44.469 \), \( |F_3| = 253.91 \) and \( |F_4| = 1703.81 \). Within a limited space (including fourth-order moment terms only), we then find the orthonormal basis

\[
H_0(x) = (0.39297, 0, 0, 0) \cdot F(x)
\]

\[
H_1(x) = (-1.11687, 0.31911, 0, 0) \cdot F(x)
\]

\[
H_2(x) = (2.25944, -1.31301, 0.16413, 0) \cdot F(x)
\]

\[
H_3(x) = (-3.88531, 3.38761, -0.85407, 0.06326) \cdot F(x)
\]
$H_0(x) = (6.07788, -7.03848, 2.65812, -0.39518, 0.01976)^T \cdot F(x)$, (46)

where we defined $F = (F_0, F_1, F_2, F_3, F_4)$. These functions are also illustrated in Figure 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 equation (46) with equation (43). From the coefficients, we have $|F_0| H_0(x) = F_0(x)$ and $|k!| F_k | H_k(x) \approx F_0(x) Z_k(x)$. This shows that the moment expansion in equation (43), with modified weighting, results in a quasi-orthogonal (or optimal) re-summation of terms. We will discuss a similar approach for the superposition of modified blackbody spectra (Section 6.2).

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

$$
(I_{\nu}) = H_{0}(x) + \sum_{k=2}^{k_{\text{max}}} \sigma_k^\nu H_k(x)
$$

for the superposition of grey-body spectra, with overall amplitude $A^\nu = |F_0| A_0 (k T / h)^3$. In an idealized case, the coefficients $\sigma_k^\nu$ are uncorrelated and directly determine the underlying moments $\alpha_k^\nu$ 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 offers a comparison of the results from different experiments easier.

While for a superposition of grey-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 grey-body spectra by using the transformation $\nu \to x$. 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 modelled. 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 parametrization for a single-temperature dust spectrum, $I_{\nu}(A_0, \alpha, T) = A_0 (\nu / \nu_0)^\alpha \nu^q / (\exp(k \nu T) - 1)$, which is referred to as a modified blackbody spectrum. This model for the fundamental SED can be motivated by considering the physical properties of dust grains in the vicinity of stars (Draine & Lazarian 1998; Finkbeiner, Davis & Schlegel 1999; Draine 2003; Planck Collaboration XII 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 grey-body spectra discussed in Section 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 $p = [A_0, \alpha, \beta = 1 / T]$. Alternative parametrizations and weighting schemes are briefly discussed below (Section 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 $\alpha$ and $\beta = 1 / T$. Since all the derivatives commute, we only have to deal with moments of the form $\alpha_{12}, \alpha_{13}, \alpha_{23}, \alpha_{123}$ and $\alpha_{123}$. The relevant spectral functions are $I_{\nu_{ij}}(v_{ij}, p) = I_{\nu_{ij}}(v_{ij}, p)/A_0$ with $i, j \in \{2, 3\} \equiv \{\alpha, \beta\}$, so that the moments $\alpha_{12}, \alpha_{23}$ and $\alpha_{123}$ can again be absorbed by defining $\alpha_{123} = \alpha_{12} + \alpha_{23} + \alpha_{123} / A_0$. Assuming narrow bands and $W(\hat{\nu}, v) = B(\hat{\nu}) F(v)$, we then find

$$
\langle I_{\nu_{ij}} \rangle = A_0 \left( \frac{\nu_{ij} v_{ij}}{e^v - 1} \right)^2 \left\{ \frac{1}{2} \left( 1 + \frac{1}{2} a_2^\nu \ln(v_{ij}/v_0) \right) + \frac{1}{6} a_2^{\nu \nu_{ij}} \ln(v_{ij}/v_0) Y_1(x) + \frac{1}{2} a_2^{\nu Y_1} \ln(v_{ij}/v_0) Y_1(x) + \frac{1}{6} a_3^{\nu Y_1} Y_2(x) + \frac{1}{2} a_3^{\nu Y_2} Y_3(x) + \ldots \right\}
$$

$$
\bar{a} = \left( \frac{A_0 (\nu / \nu_0)^\alpha}{A_0} \right), \frac{1}{\bar{f}} = \left( \frac{\langle A_0 (\nu / \nu_0)^\alpha \rangle}{A_0} \right), \frac{a_{23}^\nu}{A_0} = \left( \frac{\langle A_0 (\nu / \nu_0)^\alpha \rangle / \langle A_0 \rangle}{\langle T / T(r) - 1 \rangle} \right)^m.
$$

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 grey-body spectra would be insufficient, since the cross terms $\alpha \ln(v_{ij}) Y_0(x)$, would not have independent coefficients to take into account possible correlations between $\alpha$ and $T$. For example, the term for $k = 1$ and $m = 1$ would be absent.

#### 6.1 Behaviour in the Rayleigh–Jeans limit

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

$$
I_{\nu_{ij}}^{\nu_{ij}}(A_0, \alpha, T) \approx A_0 (\nu / \nu_0)^{\nu_{ij} + 2} \left[ 1 - \frac{x}{2} + \frac{x^2}{12} - \frac{x^4}{720} + \ldots \right].
$$

This represents a superposition of power-law spectra, which in principle can be approximated with equation (21). Since $x = h \nu / k T$, 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 $1/e^{x^2 - 1}$ 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 (Section 6.3.1).

#### 6.2 Alternative parametrizations and weighting schemes

In Section 5.2.1, we illustrated how the choice of parameters and weighting affects the moment expansion for grey-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 grey-body spectra, by setting $v_0 = k T / h$
To illustrate the application of the moment expansion, let us consider two-temperature dust models.

The general behaviour of the moment expansion for different cases is similar to that of grey-body spectra, as mentioned above. The additional parameter, $\alpha$, causes the number of moments per perturbation order to increase strongly. For example, adding all second-order moment terms means three extra parameters instead of 1 for the grey-body spectrum; at third order we need seven instead of 4.

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 GHz $\lesssim \nu \lesssim 2$ THz, we obtain $\alpha = 1.412$, $T = 14.84$ K, $\alpha_{d1} = -0.2793$, $\alpha_{d2} = 0.05132$ and $\alpha_{d3} = 0.2191$. This representation improves the fit at low frequencies, with the departure decreasing to $\lesssim 8$ percent at $\nu \approx 100$ 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 modelling of foregrounds for primordial $B$ modes. Incorrect and incomplete (truncated) parametrizations 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 $\alpha = 2.057$, $T = 13.33$ K, $\alpha_{d1} = 0.002363$, $\alpha_{d2} = -0.4189$, $\alpha_{d3} = 0.3789$, $\alpha_{d4} = -0.004326$, $\alpha_{d5} = 0.5326$, $\alpha_{d6} = -0.2657$ and $\alpha_{d7} = 0.08050$ at 100 GHz $\lesssim \nu \lesssim 2$ THz.

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 $\nu \lesssim 3$ THz. As expected from our discussion of grey-body spectra, in this case higher order moments are required to compensate for the asymptotic behaviour of the basis functions. For a given finite order and distribution of temperatures, this behaviour 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 PDFs (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.
corrections related to the Taylor series of $1/\nu$ spectra, relevant to CMB applications, also works well with $N$ moments (six parameters in total). A sum of synchrotron and dust $K_1$, but already provides a good approximation for $K_2$ (see Fig. 7).

Clearly, this representation fails to approximate the SED for model considered frequency range. For $K_2$, the six parameter moment expansion should be preferred. The resultant SEDs are shown in Fig. 7.

4.3.1 Rayleigh–Jeans approximation

We saw in Section 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 (six parameters in total). A sum of synchrotron and dust spectra, relevant to CMB applications, also works well with $N \approx 6$ moments (eight 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 $1/(e^\nu - 1)$ 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.

4.4 Distributions of temperatures

To demonstrate the potential of the dust moment representation, we analyse two more general dust models considered by Kogut & Fixsen (2016), one assuming a sum of two Gaussians for the dust temperature ($K_1$), the other using a transient heating model $^{13}$ ($K_2$). 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 (equation 50). Fitting in the range $100 \text{GHz} \lesssim \nu \lesssim 2 \text{THz}$, we find $\alpha = 0.9379$ and $T = 18.15 \text{K}$ ($K_1$) and $\alpha = 1.671$ and $T = 21.69 \text{K}$ ($K_2$) assuming a single-temperature modified blackbody spectrum. The overall amplitude, $A_0$, is be determined using $\langle I \rangle = 156.61 \text{Jy sr}^{-1}$ ($K_1$) and $\langle I \rangle = 220.11 \text{Jy sr}^{-1}$ ($K_2$) at $\nu = 100 \text{GHz}$. Clearly, this representation fails to approximate the SED for model $K_1$, but already provides a good approximation for $K_2$ (see Fig. 7).

A second-order moment expansion (six parameters) significantly improves the fit in both cases, leaving residuals at the level of $|\Delta I| \lesssim 50–200 \text{Jy sr}^{-1}$ ($K_1$) and $|\Delta I| \lesssim 5 \text{Jy sr}^{-1}$ ($K_2$) in the considered frequency range. For $K_2$, the six 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 $\text{PIXIE}$ sensitivity using only six parameters without any prior 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 $\text{PIXIE}$ in both cases. For $K_1$, we have $\alpha = 2.327$, $T = 14.56 \text{K}$, $\omega_{22}^d = -0.2983$, $\omega_{23}^d = -0.1718$, $\omega_{24}^d = -0.1396$, $\omega_{222}^d = 0.03299$, $\omega_{223}^d = 0.5685$, $\omega_{233}^d = -0.07951$ and $\omega_{2223}^d = 0.03106$ (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 modelling (e.g. see Martínez-Soláche, Karacci & Delabrouille 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 Section 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

$$I_\nu^{\text{CMB}}(\hat{\nu}) = \frac{2h}{c^2} \frac{\nu^3}{e^{\nu/kT} - 1} \approx \frac{2h}{c^2} \frac{\nu^3}{e^{\nu/kT} - 1} \text{times} \left[ 1 + G(x)(\Delta + \Delta^2) + \frac{1}{2} Y(x) \Delta^2 + O(\Delta^3) \right],$$

(52)

where $x = \hbar \nu/k T$, with $T$ denoting the full-sky average temperature and $\Delta = T(\hat{\nu})/T - 1$. 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 $T(\hat{\nu})$ in every direction, $\hat{\nu}$; thus, even the Taylor series (equation 52) still simply represents a pure blackbody spectrum in every direction. If

---

13 We cordially thank Alan Kogut for providing the average SEDs to us.
we compute the all-sky average intensity, we find

\[ I_{ν}^{\text{CMB}} \approx \frac{2h}{c^2} \frac{ν^3}{e^ν - 1} \left[ 1 + 2y_{\text{sup}} G(x) + y_{\text{sup}} Y(x) \right], \]  

(53)

where \( y_{\text{sup}} = \int \frac{2y_{\text{sup}}(ν)}{ν} d^2 \hat{p} \), which in general does not vanish. The frequency dependence of the average intensity spectrum is no longer a pure blackbody, but has a \( ν \)-type distortion (Chluba & Sunyaev 2004). One can absorb the temperature shift term \( \propto 2y_{\text{sup}} G(x) \) by redefining the average temperature to \( T' = T(1 + 2y_{\text{sup}}) \), 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 CMB from a pure blackbody in each direction to a distorted blackbody (Chluba & Sunyaev 2004).

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

\[ \Delta T^{\text{CMB}}(ν, \hat{p}) = \frac{c^2}{2h} \frac{ν^3}{ν^2 G(ν)} \left[ I_{ν}^{\text{CMB}} - I_{ν}^{\text{CMB}} \right], \]  

(54)

the multipoles for \( ℓ > 0 \) are given by

\[ \Delta T_{ℓm}^{\text{CMB}}(ν) \approx \Delta_{ℓm} + \left[ Δ_{ℓm}^2 \right] + \frac{Y(ν)}{2G(ν)} \left[ Δ_{ℓm}^2 \right], \]  

(55)

where \( [X]_m = \int Y_{ℓm}^*(\hat{p}) X(\hat{p}) d^2 \hat{p} \) is the multipole coefficient of \( X(\hat{p}) \). This shows two features of the spherical harmonic intensity map: (i) the derived antenna temperature fluctuations are frequency dependent, even if the thermodynamic temperature fluctuations, \( Δ_{ℓm} \), 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 \( ℓ > 1 \) (e.g. Chluba & Sunyaev 2004); however, our motion with respect to the CMB rest frame causes a relative leakage of power \( ^{−3} \) 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 XXVII 2014b). With the moment expansion, here defined by \( Δ_{ℓm} \) and \( 2y_{ℓm} = [Δ_{ℓm}^2] \), this effect can be trivially included and propagated. The CMB sky is thus described by maps of \( Δ_{ℓm} \) and \( y_{ℓm} \), 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 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 us 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, \( C_ℓ \propto ℓ^{−2Δ} \), matching recent Planck measurements (Planck Collaboration XLVIII 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 XLVIII 2016c) and does not vary over the sky. The spectral index \( α \), however, is varied on the sky. We generate maps of \( α \) in which the value in each pixel is drawn from a Gaussian of mean \( α_{\text{fid}} = 1.6 \) and standard deviation 0.1, matching recent Planck full-sky results (Planck Collaboration XLVIII 2016c).

Using these ingredients, we construct maps at 10 linearly spaced frequencies from 20 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 \( 2 < ℓ < 400 \). 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 \( D_{ℓ} \equiv ℓ(ℓ+1)C_{ℓ}/4π(ℓ_{0}) \) 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. \( ℓ \approx 200–300 \) (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 towards the input mean SED (e.g. for the lowest bandpower centred at \( ℓ = 30 \), the behaviour 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 behaviour than seen here. None the less, 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 behaviour, 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.

8 THE MODELLING 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 IX 2016a, for some overview). At low frequencies, AME, possibly produced by spinning dust grains in the intergalactic medium, contributes significantly, showing a broad maximum around \( v \approx 10–30 \) GHz (e.g. Draine & Lazarian 1998; Ali-Haimoud 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 IX 2016a). Significant new information about this component is now becoming available (e.g. with C-BASS, Irfan et al. 2015 and QUIJOTE, Génova-Santos et al. 2015). This will help in understanding the complexity of the AME and a moment expansion may provide one
et al. 2016) show that in agreement with previous model estimates, Monteagudo & Sunyaev 2008; De Zotti et al. 2016; Mashian, Loeb cumulative CO rotational line emission from star-forming galaxies approach for incorporating it in future CMB analyses. Future large a moment expansion in the main parameters may provide a viable and require refined theoretical studies; however, also in this case, uncertainties in the modelling of this component are significant full potential of future CMB spectral distortion measurements. The this component will have to be considered carefully to reach the common modes for the different variables.orthogonalization scheme may be beneficial to reduce the number of common modes for the different variables.

At intermediate CMB frequencies (30 GHz ≲ ν ≲ 200 GHz), the cumulative CO rotational line emission from star-forming galaxies throughout the Universe becomes important (Righi, Hernández-Monteagudo & Sunyaev 2008; De Zotti et al. 2016; Mashian, Loeb & Sternberg 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 modelling 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 modelling 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 which can be described using a halo-model approach (e.g. Desjacques et al. 2015). At frequencies above the dust peak (ν ≃ 1.52 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 parametrization for the modelling 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 (Section 3), free–free (Section 4), grey-body (Section 5) and modified blackbody (Section 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 parametrizations. 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 (equation 21) to converge quite rapidly. For the modelling of synchrotron emission, we expect N ≃ 3–4 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 Section 3.1).

For free–free, we obtained two moment expansions (equations 29 and 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 (ν ≥ 1 THz) and is captured by the second representation (equation 31). Variations in the chemical compositions and ionization degree are not expected to be as relevant to CMB foreground modelling, but applications of the moment method to the continuum X-ray emission from clusters could require a more detailed treatment.

14 For some of the models presented in Imara & Loeb (2016), we already confirmed this statement, finding only N = 6 parameters to suffice.
We also find that the free–free SED can be described using a simple power-law moment expansion with $N \simeq 3$–$4$ moments. This suggests that the combined emission from synchrotron and free–free could be modelled using one parametrization. Only at very high frequencies ($ν \gtrsim 1$ 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 $α \simeq −0.14$ rather than $α \simeq −0.9$ for synchrotron).

For the superposition of grey-body spectra, we developed two moment expansions with individual weighting (equations 39 and 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 (Section 5.2.1). We also derived an orthonormal set of basis functions for the grey-body moment expansion (Section 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 (equation 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 modelling suggested that once general distributions of temperature and dust opacities are allowed, about $\simeq 6$–$10$ 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.

In Section 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 parametrizing and propagating these effects in real maps. This is achieved by separating spatial from spectral variations though 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 modelling 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 Section 8) and to CMB polarization. Given the physically motivated extensions of foreground parametrizations, the moment expansion can furthermore be used as a diagnostic, testing the robustness of the foreground modelling 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.

ACKNOWLEDGEMENTS

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., 2015, Astropart. Phys., 63, 66
Abazajian K. N. et al., 2016, preprint (arXiv:1610.02743)
Abitbol M. H., Chluba J., Hill J. C., Johnson B. R., 2017, MNRAS, preprint (arXiv:1705.01534)
Ali-Haïmoud Y., Hirata C. M., Dickinson C., 2009, MNRAS, 395, 1055
Arendt L., Catena R., Masina I., Notari A., Quartin M., Quercellini C., 2011, J. Cosmol. Astropart. Phys., 7, 27
André P. et al., 2014, J. Cosmol. Astropart. Phys., 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., Sunyaev R. A., 2004, A&A, 424, 389
Chluba J., Sunyaev R. A., 2012, MNRAS, 419, 1294
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., Switzer E., Nelson K., Nagai D., 2013, MNRAS, 430, 3054
Chluba J., Dai L., Grin D., Amin M. A., Kamionkowski M., 2015, MNRAS, 446, 2871
De Zotti G., Negrello M., Castex G., Lapi A., Bonato M., 2016, J. Cosmol. Astropart. Phys., 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, Princeton, NJ
Draine B. T., Lazarian A., 1998, ApJ, 494, L19
Errard J., Feeney S. M., Peiris H. V., Jaffe A. H., 2016, J. Cosmol. Astropart. Phys., 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., Wehus I. K., Eriksen H. K., Næs S. K., 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, Phys. Rev. Lett., 115, 261301
Hoang T., Vinh N.-A., Quyhn Lan N., 2016, ApJ, 824, 18
Hu W., 1995, preprint (astro-ph/9508126)
Imara N., Loeb A., 2016, ApJ, 825, 130
Irfan M. O. et al., 2015, MNRAS, 448, 3572
Itoh N., Sakanoto T., Kusano S., Nozawa S., Kohyama Y., 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., Fixsen D. J., 2016, ApJ, 826, 101
Kogut A. et al., 2011, ApJ, 734, 4
Kogut A., Chluba J., Fixsen D. J., Meyer S., Spergel D., 2016, Proc. SPIE, 9904, 99040W
Martínez-Solà G., Karakci A., Delabrouille J., 2017, preprint (arXiv:1706.04162)

MNRAS 472, 1195–1213 (2017)
APPENDIX A: FREE–FREE EMISION IN THE OPTICALLY THIN REGIME

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

\[
\frac{\partial n(x_e)}{\partial \tau} = \kappa(x_e, T_e) \frac{e^{x_e} - e^{-x_e}}{x_e^2} \left[ 1 - n(x_e)(e^{x_e} - 1) \right]. \tag{A1}
\]

Here, \( \tau = \int N_N \sigma_T \, dl \) 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 temperature, 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 \( x = h\nu/kT_0 \), we have

\[
\frac{\delta n(x_e)}{\delta \tau} \approx \frac{\kappa(x_e, T_e) e^{-x_e} - e^{-x_e}}{x_e^3} \frac{e^{-x_e} - e^{-x_e}}{e^{-x_e} - 1} = \frac{\delta \tau}{x_e^3} \frac{\kappa(x_e, T_e) e^{x_e} - 1}{e^{x_e} - 1}, \tag{A2}
\]

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

APPENDIX B: BASIS FOR FREE–FREE SPECTRA

Here, we give the first few functions, \( G^{(f,m)}_k \), defined by equation (32). They read

\[
G^{(1,0)}_1(\bar{\nu}_f, \bar{\xi}, \nu) = \left( -\frac{4}{81}, 0, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(2,0)}_1(\bar{\nu}_f, \bar{\xi}, \nu) = \left( \frac{2}{9}, -\frac{20}{9}, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(2,0)}_2(\bar{\nu}_f, \bar{\xi}, \nu) = \left( \frac{2}{9}, 0, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(3,0)}_3(\bar{\nu}_f, \bar{\xi}, \nu) = \left( -\frac{4}{27}, 0, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(3,0)}_4(\bar{\nu}_f, \bar{\xi}, \nu) = \left( -\frac{4}{27}, -\frac{4}{3}, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(4,0)}_5(\bar{\nu}_f, \bar{\xi}, \nu) = \left( -\frac{4}{27}, -\frac{4}{3}, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(0,1)}_1(\bar{\nu}_f, \bar{\xi}, \nu) = (1, 0, 0, 0)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(1,1)}_2(\bar{\nu}_f, \bar{\xi}, \nu) = \left( -\frac{2}{3}, 1, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(2,1)}_3(\bar{\nu}_f, \bar{\xi}, \nu) = \left( \frac{2}{9}, -\frac{2}{3}, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(3,1)}_4(\bar{\nu}_f, \bar{\xi}, \nu) = \left( -\frac{4}{27}, -\frac{4}{3}, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(0,2)}_5(\bar{\nu}_f, \bar{\xi}, \nu) = \left( -\frac{2}{3}, 1, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(1,2)}_6(\bar{\nu}_f, \bar{\xi}, \nu) = \left( \frac{4}{9}, 0, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu)
\]

\[
G^{(2,2)}_7(\bar{\nu}_f, \bar{\xi}, \nu) = \left( \frac{4}{9}, 0, 0, 0 \right)^T \cdot \mathbf{g}_f(\bar{\nu}_f, \bar{\xi}, \nu), \tag{B1}
\]

where we defined \( \mathbf{g}_f = (g_{f1}, g_{f2}, g_{f3}) \). These are needed for a fourth-order moment expansion. We also have \( G^{(0,0)}_1 = G^{(0,0)}_2 = G^{(0,0)}_3 = 0 \).

APPENDIX C: DERIVATIVES OF PLANCKIAN

In this section, we give a closed form for the derivatives of a Planckian, \( n^0(\nu, \beta) = (e^{\beta h\nu/k} - 1)^{-1} \) with respect to \( \beta = 1/T \). Since \( x = h\nu/k \) we have \( \partial x/\partial k = h\nu/k \), we can simply write

\[
\beta^j \partial^j_x n^0(\nu, \beta) = x^j \partial^j_x n^0(x) = x^j \partial^j_x n^0(x) \left( \frac{e^{x^j}}{(1 - e^{-x})} \sum_{m=0}^{\infty} \left( \begin{array}{c} j \\ m \end{array} \right) \right) e^{-mx}, \tag{C1}
\]

where we used the result of Chluba et al. (2012b) for \( x^j \partial^j_x n^0(x) \).

Here, \( \left( \begin{array}{c} j \\ m \end{array} \right) \) denotes the Eulerian numbers, which can be computed using recursion relations (compare Chluba et al. 2012b).

\[\text{This paper has been typeset from a \LaTeX/\bibtex} file prepared by the author.}\]