

# Refined approximations for the distortion visibility function and $\mu$ -type spectral distortions

J. Chluba<sup>1\*</sup>

<sup>1</sup> Department of Physics and Astronomy, Johns Hopkins University, Bloomberg Center 435, 3400 N. Charles St., Baltimore, MD 21218, USA

Accepted 2014 February 27. Received 2013 December 19

## ABSTRACT

The physical processes affecting the thermalization of cosmic microwave background spectral distortions are very simple and well understood. This allows us to make precise predictions for the distortions signals caused by various energy release scenarios, where the theoretical uncertainty is largely dominated by the physical ingredients that are used for the calculation. Here, we revisit various approximations for the *distortion visibility function* – defined using the fraction of the released energy that does not thermalize – and early  $\mu$ -type distortions. Our approach is based on a perturbative expansion, which allows us to identify and clarify the origin of different improvements over earlier approximations. It provides a better than  $\approx 0.1\% - 1\%$  description of our numerical results over a wide range of parameters. In particular, we are able to capture the high-frequency part of the  $\mu$ -distortion, which directly depends on the time derivative of the electron temperature. We also include lowest order double Compton and Compton scattering relativistic corrections, finding that because of cancelation they increase the thermalization efficiency in the tail of the distortion visibility function by only  $\approx 10\%$  (at  $z \approx 6 \times 10^6$ ), although individually their effect can reach  $\approx 20\% - 40\%$ .

**Key words:** Cosmology: CMB – spectral distortions – theory – observations

## 1 INTRODUCTION

Energy release in the early Universe can cause deviations of the cosmic microwave background (CMB) energy spectrum from a pure blackbody shape (e.g., Zeldovich & Sunyaev 1969; Sunyaev & Zeldovich 1970; Illarionov & Sunyaev 1975; Danese & de Zotti 1977). These spectral distortions can tell us about processes occurring well before photons last scattered at redshift  $z \approx 1100$ , allowing us to constrain the thermal history of our Universe looking deep into the pre-recombination plasma. This exciting possibility has recently spurred renewed theoretical interest into how spectral distortions form and evolve, showing that the distortion signals may open a new unexplored window to early-universe and particle physics (see Chluba & Sunyaev 2012; Sunyaev & Khatri 2013; Chluba 2013a, for broader overview).

The physics going into the thermalization problem – the process that restores the pure blackbody spectrum after some perturbation from thermal equilibrium – are pretty simple and well understood. For primordial spectral distortions, we are only concerned with the average CMB spectrum, so that spatial perturbations can be neglected and the Universe can be described as uniformly expanding, thermal plasma consisting of free electrons, hydrogen and helium atoms, and their corresponding ions inside a bath of photons from the CMB. We shall restrict ourselves to redshifts  $z \lesssim \text{few} \times 10^7$ , when electron-positron pairs already completely disappeared. With

these assumptions, any energy release inevitably causes a momentary distortion of the CMB spectrum. This can be understood with the following simple arguments: a pure blackbody spectrum,  $B_\nu(T)$ , is fully characterized by one number, its temperature  $T$ . Changing the energy density,  $\rho_\nu$ , of the photon field by some  $\Delta\rho_\nu/\rho_\nu \ll 1$  (e.g., from some particle decay) means that the photon number also has to be readjusted by  $\Delta N_\nu/N_\nu \approx (3/4)\Delta\rho_\nu/\rho_\nu$ . These additional photons, furthermore, have to be distributed according to  $\partial B_\nu/\partial T$  in energy, to correctly shift the initial blackbody temperature from  $T$  to  $T' \approx T + (1/4)\Delta\rho_\nu/\rho_\nu$ . In the early Universe, the double Compton (DC) and Bremsstrahlung (BR) processes are controlling the number of CMB photons, while Compton scattering (CS) allows photons to diffuse in energy. The exact interplay of these interactions between matter and radiation determines the spectrum of the CMB at any stage of its evolution. When studying different energy release mechanisms, the question thus is whether there was enough time between the energy release event and our observation to produce and redistribute those  $\Delta N_\nu/N_\nu$  of missing photons, thereby fully digesting the energy injection, rendering the *distortion visibility* tiny.

The thermalization problem has been studied thoroughly both analytically (e.g., Zeldovich & Sunyaev 1969; Sunyaev & Zeldovich 1970; Illarionov & Sunyaev 1975; Danese & de Zotti 1977; Burigana et al. 1995; Chluba 2005; Khatri & Sunyaev 2012b,a) and numerically (e.g., Burigana et al. 1991; Hu & Silk 1993; Burigana & Salvaterra 2003; Procopio & Burigana 2009; Chluba & Sunyaev 2012;

\* E-mail: jchluba@pha.jhu.edu

Chluba 2013a). From these studies, the following simple picture can be drawn: at  $z \gtrsim 2 \times 10^6$ , the thermalization process is extremely efficient and practically any distortion can be erased until today. At lower redshifts, the CMB spectrum becomes vulnerable to disturbances in the thermal history and only small amounts of energy can be ingested without violating the tight experimental bounds from COBE/FIRAS (Mather et al. 1994; Fixsen et al. 1996; Fixsen & Mather 2002) and other distortion measurements (Kogut et al. 2006; Zannoni et al. 2008; Seiffert et al. 2011).

The transition from efficient to inefficient thermalization is encoded by the distortion visibility function,  $\mathcal{J}(z, z')$ , which determines by how much the distortion amplitude is suppressed between redshift  $z$  and  $z' < z$ . The shape of the final distortion is close to a superposition of  $\mu$ - and  $y$ -type distortion (Zeldovich & Sunyaev 1969; Sunyaev & Zeldovich 1970), with a smaller residual (non- $\mu$ /non- $y$ ) which provides additional time-dependent information at  $10^4 \lesssim z \lesssim 3 \times 10^5$  (Burigana et al. 1991; Hu 1995; Chluba & Sunyaev 2012; Khatri & Sunyaev 2012a; Chluba 2013b; Chluba & Jeong 2014). How far into the *cosmic photosphere* (Bond 1996) one can view, furthermore depends on the sensitivity of the experiment and how much energy needs to be thermalized.

The distortion visibility function is independent of the precise form of the distortion and just measures how much of the released energy is left in the distorted spectrum. To determine its shape, it is sufficient to study the evolution of distortions in the  $\mu$ -era ( $3 \times 10^5 \lesssim z$ ), since later the visibility function is already extremely close to unity. This simplifies the problem immensely, since the photon distribution is brought into kinetic equilibrium with the matter within a very short time and only evolves rather slowly along a sequence of quasi-stationary stages. Due to the huge entropy of the Universe (there are  $\approx 1.6 \times 10^9$  times more photons than baryons), the DC process is furthermore the crucial source of photons in the  $\mu$ -era, so that the distortion visibility function is roughly given by  $\mathcal{J}_{\text{DC}}(z, z') = e^{-(z/z_{\text{dc}})^{5/2}} e^{(z'/z_{\text{dc}})^{5/2}}$ , with thermalization redshift  $z_{\text{dc}} \approx 1.98 \times 10^6$  (Sunyaev & Zeldovich 1970; Danese & de Zotti 1977; Hu & Silk 1993).

For the analysis of future CMB distortion data (see Kogut et al. 2011; PRISM Collaboration et al. 2013; André et al. 2014, for discussion of some experimental concepts), it is of interest to improve the approximation for  $\mathcal{J}(z, z')$ . While numerically, it is straightforward, although time consuming, to compute  $\mathcal{J}(z, z')$  precisely (e.g., using COSMOTHERM; Chluba & Sunyaev 2012), deeper physical insight can be gained analytically. Recently, Khatri & Sunyaev (2012b) [KS12, henceforth] took some steps into this direction. They showed that it is easy to analytically include the additional thermalization effect from BR, which becomes significant at  $z \lesssim 10^6$ . They also added a correction to the shape of the  $\mu$ -distortion caused by small deviations from quasi-stationarity, showing that at  $10^6 \lesssim z$  this captures the associated increase in the thermalization efficiency relative to  $\mathcal{J}_{\text{DC}}(z, z')$  seen numerically.

Here, we revisit this problem with a slightly different approach, basing our analysis on a perturbative expansion of the evolution equation. We show that several terms add at a similar order to the thermalization problem as the non-stationary correction discussed by KS12. We demonstrate that at low frequencies, small frequency-dependent variations of the photon production rate are significant in our approach. The time derivative of the critical frequency,  $x_c$ , characterizing the transition from effective to ineffective photon production also affects the shape of the distortion at the later stages ( $z \lesssim 10^6$ ). Finally, we extend the validity of the solution for the spectral distortion to high frequencies, showing that it is driven by the time derivative of the electron temperature (see Sect. 5.5),

which is in contrast to KS12, who argued that there the time derivative of the chemical potential amplitude is most important. While the latter does not affect the result for the distortion visibility function by much, as we show here, it is possible to capture this aspect in a simple way, obtaining an  $\approx 0.1\%$  description of the distortion shape at basically all relevant energies.

We fix the integration constants using energetic arguments (Sect. 2.2), finding an  $\lesssim 0.1\% - 1\%$  agreement of our approximation for the distortion visibility function with the full numerical result. We also argue that at late times ( $z \lesssim 2 \times 10^5$ ), photons produced by BR no longer are able to up-scatter efficiently and thus remain trapped at low frequencies (Sect. 3.3.4). This means that the distortion visibility function becomes very close to unity at that point. By also including this effect, we reproduce our numerical result and overcome some of the small differences with respect to the numerical solution seen by KS12 at  $z \lesssim 10^6$ .

Finally, Chluba & Sunyaev (2012) included improved approximations for the DC Gaunt factor (Chluba 2005; Chluba et al. 2007). Due to higher order temperature corrections, the DC emissivity reduces relative to the non-relativistic case (Lightman 1981; Thorne 1981), making thermalization less efficient. On the other hand, frequency-dependent corrections move the maximal emission towards slightly higher frequencies, increasing the photon production rate. These effects can be captured analytically, giving a net increase in the thermalization efficiency at early times, which is in excellent agreement with our numerical result (Sect. 6.1). This effect is partially counteracted by lowest order temperature corrections to CS, a modification that was not included by previous numerical treatments. Our perturbative approach again allows capturing this effect, providing a simple understanding for the origin of this correction (Sect. 6.2).

The paper is structured as follows: Sect. 2 and Sect. 3 provide some of the basic ingredient for computing the evolution of the distortion. Parts of this are very pedagogical and can be skipped by the expert. In Sect. 4 and Sect. 5, we discuss the solutions for the thermalization problem, assuming small distortions. We also give our discussion for the distortion visibility function, with the main result shown in Fig. 8. We close our analysis by including lowest order relativistic corrections to DC and CS in Sect. 6. Our conclusions are given in Sect. 7.<sup>1</sup>

## 2 NON-EQUILIBRIUM THERMODYNAMICS OF BOSE-EINSTEIN SPECTRA

To formulate the problem, it is useful to first go over the definitions of occupation number, number density, energy density and entropy density of a Bose-Einstein spectrum with small frequency-dependent chemical potential variable<sup>2</sup>,  $\bar{\mu}(t, x)$ . We envision conditions in the early Universe at redshift  $z \gtrsim \text{few} \times 10^5$ , when CS is still very efficient in redistributing photons over energy. Any photon

<sup>1</sup> Our reference cosmology is:  $Y_p = 0.24$ ,  $h = 0.71$ ,  $N_{\text{eff}} = 3.046$ ,  $\Omega_b = 0.044$ ,  $\Omega_{\text{cdm}} = 0.216$ ,  $\Omega_\Lambda = 0.74$  and  $T_0 = 2.726$  K.

<sup>2</sup> We call it ‘variable’ because it provides a simple parametrization of the spectrum that appears like a chemical potential, but is not generally identical with the thermodynamic chemical potential, which under equilibrium conditions vanishes for photons. In the text, we will usually drop the term ‘variable’, but it should be kept in mind.

distribution can be expressed as<sup>3</sup>

$$n(x) = \frac{1}{e^{x+\tilde{\mu}(t,x)} - 1} \approx n_{\text{pl}}(x) - G(x) \frac{\tilde{\mu}(t,x)}{x} + O(\tilde{\mu}^2), \quad (1)$$

where  $x = hv/kT_\gamma$  denotes the dimensionless frequency with temperature  $T_\gamma = T_0(1+z)$ , which is used to define a reference energy scale, and  $T_0 = 2.726$  K (Fixsen et al. 1996; Fixsen 2009). Furthermore,  $n_{\text{pl}}(x) = 1/[e^x - 1]$  is the occupation number of a blackbody at temperature  $T_\gamma$  and  $G(x) = -x\partial_x n_{\text{pl}}(x) = xe^x/[e^x - 1]^2$  describes the spectrum of a simple temperature shift: adding/removing photons with this spectral shape does not create a distortion unless higher order terms become important (i.e.,  $\Delta T/T$  becomes too large).

Importantly, with  $T_\gamma \propto (1+z)$  no redshifting term appears in the photon Boltzmann equation, which simplifies the problem significantly. Note, however, that  $T_\gamma$  generally is not identical to the effective temperature of the photon field, which for a distorted spectrum can, for instance, be defined in terms of photon number or energy density (cf. Sect. 2.2). It is also generally not identical to the Rayleigh-Jeans temperature,  $T_{\text{RJ}}$ , defined at  $x \ll 1$ , which due to photon emission and absorption processes is very close to the electron temperature,  $T_e$ , at sufficiently low frequencies.

At any moment, the total CMB spectrum is given by a blackbody part at some temperature,  $T_{\text{bb}}$ , plus a distortion relative to this. In general,  $T_{\text{bb}} \neq T_\gamma$ , and if we admit that  $T_{\text{bb}}$  can change because of energy release, we can write the parametrization

$$\tilde{\mu}(t,x) = x \left( \frac{T_\gamma(t)}{T_{\text{bb}}(t)} - 1 \right) + \mu_\infty(t) \hat{\mu}(t,x) \quad (2)$$

for the chemical potential. Here, the first term just represents the correct temperature shift of  $n_{\text{pl}}(x)$  to  $n_{\text{pl}}(xT_\gamma/T_{\text{bb}})$ . The overall amplitude of the effective chemical potential  $\mu(t,x) = \mu_\infty(t) \hat{\mu}(t,x)$  is defined by  $\mu_\infty(t)$ , with all frequency-dependent terms captured by  $\hat{\mu}(t,x)$ . Since generally both  $\mu_\infty(t)$  and  $\hat{\mu}(t,x)$  depend on time, we still need a convenient normalization condition to uniquely determine the factorization, but we shall return to this point later.

The temperature  $T_{\text{bb}}$  in principle can be chosen freely, but now its time dependence is generally unknown before the solution of the problem is obtained. Since at low frequencies the photon distribution is pushed very close to equilibrium with the electrons by the double Compton and Bremsstrahlung processes, one useful choice is  $T_{\text{bb}} \equiv T_{\text{RJ}} = T_e$ , which we shall use henceforth. This also removes any contribution  $\hat{\mu}(t,x) \propto x$  from the effective chemical potential, as we will see below.

## 2.1 Photon number and energy densities

Setting  $T_{\text{bb}} = T_{\text{RJ}} = T_e$  in the expressions from above, we can obtain the number, energy and entropy densities of the distorted photon field as (cf., Sunyaev & Zeldovich 1970)

$$N_\gamma \approx N_\gamma^{\text{pl}}(T_\gamma) \left[ 1 + 3 \frac{\Delta T_e}{T_\gamma} - \mu_\infty M_2 \right], \quad (3a)$$

$$\rho_\gamma \approx \rho_\gamma^{\text{pl}}(T_\gamma) \left[ 1 + 4 \frac{\Delta T_e}{T_\gamma} - \mu_\infty M_3 \right], \quad (3b)$$

$$s_\gamma \approx s_\gamma^{\text{pl}}(T_\gamma) \left[ 1 + 3 \frac{\Delta T_e}{T_\gamma} - \frac{3}{4} \mu_\infty M_3 \right], \quad (3c)$$

where  $\Delta T_e = T_e - T_\gamma$ , and  $N_\gamma^{\text{pl}}$ ,  $\rho_\gamma^{\text{pl}}$  and  $s_\gamma^{\text{pl}} = (4/3)\rho_\gamma^{\text{pl}}/T_\gamma$  denote the photon number, energy and entropy densities of a blackbody at temperature  $T_\gamma$ , respectively (for explicit definition of  $s_\gamma$  in terms

of photon occupation number, see Appendix B). For convenience, we also introduced the integrals  $\mathcal{M}_k$  as

$$\mathcal{M}_k = \frac{1}{\mathcal{G}_k^{\text{pl}}} \int x^{k-1} G(x) \hat{\mu}(t,x) dx \stackrel{\hat{\mu}=1}{=} \frac{k \mathcal{G}_{k-1}^{\text{pl}}}{\mathcal{G}_k^{\text{pl}}}, \quad (4a)$$

$$\mathcal{G}_k^{\text{pl}} = \int x^k n_{\text{pl}}(x) dx. \quad (4b)$$

A few important examples for  $\mathcal{G}_k^{\text{pl}}$  are  $\mathcal{G}_1^{\text{pl}} \approx 1.6449$ ,  $\mathcal{G}_2^{\text{pl}} \approx 2.4041$ ,  $\mathcal{G}_3^{\text{pl}} \approx 6.4939$ ,  $\mathcal{G}_4^{\text{pl}} \approx 24.886$  and  $\mathcal{G}_5^{\text{pl}} \approx 122.08$ . The integrals  $\mathcal{M}_k$  directly depend on the shape of the distortion  $\hat{\mu}(x,t)$ , which also introduces additional time dependence to the problem. Assuming  $\hat{\mu} = 1$  we have  $\mathcal{M}_2^c \approx 1.3684$  and  $\mathcal{M}_3^c \approx 1.1106$ . Henceforth, the superscript ‘c’ will indicate that  $\hat{\mu} = 1$  was used for the variable.

## 2.2 Effective temperatures of the photon field

One can easily show that for small distortions the effective temperatures<sup>4</sup> of the distorted spectrum with respect to photon number,  $T_N^*$ , and energy density,  $T_\rho^*$ , are respectively given by

$$T_N^* \approx T_e - \frac{1}{3} T_\gamma \mu_\infty M_2 \approx T_\gamma \left[ 1 + \frac{\Delta T_e}{T_\gamma} - 0.4561 \mu_\infty \right] \quad (5a)$$

$$T_\rho^* \approx T_e - \frac{1}{4} T_\gamma \mu_\infty M_3 \approx T_\gamma \left[ 1 + \frac{\Delta T_e}{T_\gamma} - 0.2776 \mu_\infty \right], \quad (5b)$$

where for the second approximate sign we assumed constant (independent of frequency) chemical potential. Without distortion one naturally has  $T_N^* = T_\rho^* = T_e$ . The effective temperature following from the entropy density, Eq. (3c), is identical to the one of the energy density. For positive  $\mu_\infty$  and  $\hat{\mu} = 1$  (constant chemical potential), Eq. (5) implies  $T_N^* < T_\rho^*$ , or explicitly

$$T_\rho^* - T_N^* = \frac{T_\gamma \kappa_\rho(t) \mu_\infty(t)}{4} \approx 0.1785 \mu_\infty^*(t) T_\gamma, \quad (6)$$

where we defined  $\kappa_\rho(t) = 4M_2 - 3M_3$ , which for constant chemical potential  $\mu(t,x) = \mu_\infty(t)$  gives  $\kappa_\rho^c = 4M_2^c - 3M_3^c \approx 2.1419$ . Once thermalization creates  $\mu_\infty \rightarrow 0$ , this also means that  $T_N^* \rightarrow T_\rho^*$ , restoring full thermal equilibrium.

In Eq. (6), we introduced the effective amplitude of the chemical potential,  $\mu_\infty^*(t) = \hat{\kappa}_\rho(t) \mu_\infty(t)$ , with  $\hat{\kappa}_\rho(t) = \kappa_\rho(t)/\kappa_\rho^c$ . If  $\hat{\mu} = 1$ , we have  $\hat{\kappa}_\rho(t) = 1$  and hence  $\mu_\infty^*(t) \equiv \mu_\infty(t)$ , but generally  $\hat{\kappa}_\rho(t) \neq 1$ . Since we still have the freedom to normalize  $\hat{\mu}$ , for convenience we choose its normalization such that  $\hat{\kappa}_\rho(t) \equiv 1$  at all times. This simplifies all the expressions discussed below and we can use  $\mu_\infty^*(t) = \mu_\infty(t)$  and  $\kappa_\rho(t) = \kappa_\rho^c$  without loss of generality.

## 2.3 Compton equilibrium temperature

It is well known that in the early Universe after a very short time the free electrons approach a temperature that is dictated by the shape of the (high-frequency) photon distribution and the energy exchange through Compton scattering (Zeldovich & Levich 1970):

$$T_e \approx T_e^{\text{eq}} = T_\gamma \frac{\int n(1+n)x^4 dx}{4 \int nx^3 dx}. \quad (7)$$

For a constant chemical potential  $\mu(t,x)$ , one has  $n(1+n) = -\partial_x n$  and after integration by parts one finds  $T_e^{\text{eq}} = T_\gamma \equiv T_{\text{RJ}}$ . This is no

<sup>3</sup> See Appendix A for general values of  $\tilde{\mu} = \text{const}$ .

<sup>4</sup> These are defined by equating the true number and energy densities with the one of a blackbody at the corresponding effective temperature.

longer exactly true once photon production starts to revert the distorted spectrum to a blackbody at  $x \ll 1$ . In this case,  $T_e$  is pushed away from  $T_\gamma$  and the global energetics have to be considered.

#### 2.4 Initial state after very short burst of energy release

After energy release ceases, the comoving energy density of the photon-baryon system remains constant. Assuming that there was no time to produce photons, but that Compton scattering already brought electrons and photons into kinetic equilibrium, we can compute the initial photon temperature and chemical potential. Defining the photon temperature before the energy release as  $T_\gamma^{\text{in}}$  and assuming that a total of  $\Delta\rho_\gamma/\rho_\gamma \ll 1$  of energy was injected at redshift  $z_{\text{in}}$ , with Eq. (5), assuming  $\mu = \mu_\infty^{\text{st}} = \text{const}$  (so that  $T_e = T_\gamma = T_{\text{RJ}}$ ), we find (cf., Sunyaev & Zeldovich 1970)

$$T_\gamma^{\text{st}} \approx T_\gamma^{\text{in}} [1 + 0.4561 \mu_\infty] \equiv T_\gamma = T_0(1 + z_{\text{in}}) \quad (8a)$$

$$\mu_\infty^{\text{st}} \approx \frac{3}{\kappa_p^c} \frac{\Delta\rho_\gamma}{\rho_\gamma} \approx 1.401 \frac{\Delta\rho_\gamma}{\rho_\gamma}. \quad (8b)$$

Without additional photon production by the DC and BR processes and without subsequent energy release, this would describe the state of the CMB spectrum even today because the adiabatic expansion of the Universe leaves the photon distribution unaltered once Compton scattering brought things into kinetic equilibrium<sup>5</sup>.

Expression (8a) for  $T_\gamma^{\text{st}}$  also defines the initial Rayleigh-Jeans temperature of the photon field. Once photon production starts and transport of photons to higher frequencies reduces the value of  $\mu_\infty$ , one finds  $T_\gamma^{\text{st}} > T_{\text{RJ}} = T_e$ . On the other hand, the temperature following from the energy density is  $T_\rho^* \approx T_\gamma^{\text{in}}(1 + \frac{1}{4}\Delta\rho_\gamma/\rho_\gamma) \approx T_\gamma^{\text{st}}(1 - 0.2776\mu_\infty^{\text{st}}) < T_\gamma^{\text{st}}$ . This determines the final temperature of the blackbody (corrected for the effect of redshifting), which is slowly approached as the thermalization process completes. In this process, a total of  $\Delta N_\gamma/N_\gamma \approx (3/4)\Delta\rho_\gamma/\rho_\gamma$  photons have to be replenished by BR and DC emission relative to the initial blackbody at temperature  $T_\gamma^{\text{in}}$ , at least if there was a sufficient amount of time to completely thermalize the distortions.

#### 2.5 Evolution equations for $N_\gamma$ and $\rho_\gamma$

The equations (3) are similar to the expressions given by Sunyaev & Zeldovich (1970); however, we do not make the assumption that the integrals  $\mathcal{M}_k$  are constant and can be computed for  $\hat{\mu} = \text{const}$ . Combining the comoving time derivatives of Eq. (3), we can thus write the more general evolution equations

$$\frac{\kappa_p^c}{\mathcal{M}_2\mathcal{M}_3} \frac{d T_e}{dt} \approx \frac{d \ln a^4 \rho_\gamma}{\mathcal{M}_3 dt} - \frac{d \ln a^3 N_\gamma}{\mathcal{M}_2 dt} + \mu_\infty \frac{d}{dt} \ln \left( \frac{\mathcal{M}_3}{\mathcal{M}_2} \right) \quad (9a)$$

$$\frac{d\mu_\infty}{dt} \approx \frac{3}{\kappa_p^c} \frac{d \ln a^4 \rho_\gamma}{dt} - \frac{4}{\kappa_p^c} \frac{d \ln a^3 N_\gamma}{dt}, \quad (9b)$$

where  $a = (1 + z)^{-1}$  is the scale factor normalized to unity at  $z = 0$ . As explained in Sect. 2.2, we chose the normalization of  $\hat{\mu}(t, x)$  such that  $\kappa_p(t) = \kappa_p^c = \text{const}$ . The only additional assumption at this point is that the overall amplitude of the distortion is small and only linear order terms need to be considered.

The terms on the r.h.s. of Eq. (9b) determine real changes of

<sup>5</sup> This is not entirely true since the slow Hubble expansion leads to extra Compton cooling of photons caused by the difference in the adiabatic indices of baryons and photons and hence a change of the spectrum by additional Compton scattering (Chluba & Sunyaev 2012).

the photon energy and number density. These need to be obtained from the photon Boltzmann equation, which includes the effect of electron scattering and emission/absorption. After energy release ends, only  $d \ln a^4 N_\gamma/dt$  drives the evolution of  $\mu_\infty(t)$  and  $T_e(t)$ . To compute this term, assuming  $\mu(t, x) \approx \mu_\infty(t)$  is *insufficient*, since at low frequencies, where most of the photon production happens, emission and absorption processes return the photon distribution to a blackbody very quickly (i.e.,  $\hat{\mu}(t, x) \rightarrow 0$  at  $x \ll 1$ ), reducing  $d \ln a^4 N_\gamma/dt$  to a finite (!) value that critically depends on the shape of the low-frequency spectrum. Knowing the exact time dependence of this term is therefore crucial for describing the overall evolution of the distortion (see Sect. 5).

#### 2.6 Consistency relations

The way the problem is set up, we have three unknown functions:  $T_e(t)$ ,  $\mu_\infty(t)$  and  $\hat{\mu}(t, x)$ . By construction,  $\mu_\infty(t)$  and  $\hat{\mu}(t, x)$  define the deviation of the CMB spectrum with respect to the Rayleigh-Jeans temperature,  $T_{\text{RJ}} \equiv T_e$ . Assuming that only a single injection of  $\Delta\rho_\gamma/\rho_\gamma$  occurs at  $t_{\text{in}}$ , we already know the initial state of the photon distribution from Eq. (8) with  $\hat{\mu}(t_{\text{in}}, x) = 1$ . As thermalization proceeds,  $\mu$  and  $T_{\text{RJ}} \equiv T_e$  slowly approach  $\mu = 0$  and  $T_e(z) = T_\gamma^{\text{in}}(1 + \frac{1}{4}\Delta\rho_\gamma/\rho_\gamma)(1 + z)/(1 + z_{\text{in}})$  after full thermalization.

In this picture, a few consistency relations ought to be fulfilled by the solution at any time. First, if only one episode of energy release occurs, then at any time  $t > t_{\text{in}}$  we should find

$$\frac{\Delta T_e(t)}{T_\gamma} \approx \frac{\mathcal{M}_3(t)}{4} \mu_\infty(t) - \frac{3}{4} \frac{\mathcal{M}_3^c}{\kappa_p^c} \frac{\Delta\rho_\gamma}{\rho_\gamma}, \quad (10)$$

reflecting conservation of the comoving energy density<sup>6</sup> by the system. The coefficient reads  $(3/4)(\mathcal{M}_3^c/\kappa_p^c) \approx 0.3889$ .

Similarly, the total number of photons that have been created by DC and BR emission since the energy release should at any stage of the evolution be given by

$$\frac{\Delta N_\gamma(t)}{N_\gamma} \approx \frac{3}{4} \left( \frac{\Delta\rho_\gamma}{\rho_\gamma} - \frac{\kappa_p^c \mu_\infty(t)}{3} \right), \quad (11)$$

where  $\Delta N_\gamma(t)/N_\gamma$  is computed with respect to the initial blackbody spectrum before the energy release. This expression already shows that  $\kappa_p^c \mu_\infty(t)/3$  can be interpreted as the energy density carried by the non-blackbody part of the spectrum, or explicitly

$$\mu_\infty(t) \equiv \frac{3}{\kappa_p^c} \frac{\Delta\rho_\gamma(t)}{\rho_\gamma} - \frac{4}{\kappa_p^c} \frac{\Delta N_\gamma(t)}{N_\gamma}. \quad (12)$$

Here, the energy density and number densities are momentary values computed directly from the distorted photon field. This expression is independent of which temperature is used for the pure blackbody part, as terms  $\hat{\mu} \propto x$  automatically cancel. This is because the distortion is defined with respect to photon deficit (for positive  $\mu_\infty$ ) relative to the photon energy density, also reflected by Eq. (9b). This is a very sensible interpretation of the distortion, since scattering processes leave the number of photons unchanged. In addition, we need two boundary conditions for  $\hat{\mu}(t, x)$  to close the problem. By construction, we have  $\hat{\mu}(t, x) \rightarrow 0$  for  $x \ll 1$ . This also implies that we need to ensure that  $T_{\text{RJ}} = T_e$ . The final condition follows from  $\kappa_p = 4\mathcal{M}_2 - 3\mathcal{M}_3 = \kappa_p^c$ , which fixes the overall normalization for  $\hat{\mu}$  as explained in Sect. 2.2.

<sup>6</sup> We neglect the tiny heat capacity of ordinary matter here.

### 3 THE PHOTON BOLTZMANN EQUATION

In this section, we give the Boltzmann equation for photons writing out first-order relativistic corrections to the Compton and DC processes. Although slightly technical, we need these equations to obtain the solutions for the chemical potential. We linearize the expressions assuming small spectral distortions. Discussion of the solutions is presented in Sect. 5.

#### 3.1 General form

In the early Universe, photons undergo many interactions with the electrons. At redshifts  $z \gtrsim 4 \times 10^5$ , the most important processes are CS and DC scattering: DC serves as a source of soft photons at low frequencies, whereas CS efficiently redistributes photons over frequency. At lower redshifts ( $z \lesssim 4 \times 10^5$ ), BR starts taking over the photon production; however, at that stage, the up-scattering of photons becomes inefficient. It is convenient to express all time-scales in units of the Thomson scattering time-scale,  $t_C = 1/\sigma_T N_e c \approx 2.3 \times 10^{20}(1+z)^{-3}$  sec. Then the Boltzmann equation describing the time evolution of the photon phase space density  $n_\gamma(\tau, x)$  in the expanding Universe is given by

$$\frac{\partial n}{\partial \tau} = \left. \frac{\partial n}{\partial \tau} \right|_{\text{CS}} + \left. \frac{\partial n}{\partial \tau} \right|_{\text{DC}} + \left. \frac{\partial n}{\partial \tau} \right|_{\text{BR}} + \mathcal{S}(\tau, x), \quad (13)$$

where we introduced the optical depth  $d\tau = dt/t_C$  to Thomson scattering as dimensionless time variable. The terms on the r.h.s. of this equation, respectively, describe the effect of CS, DC scattering, BR and additional sources of photons with the source function,  $\mathcal{S}(\tau, x)$ . This source function, for example, could include the effect of dissipation of acoustic modes in the early Universe (e.g., see Chluba et al. 2012b) or photons produced by decaying or annihilating particles. As explained in the previous section, in  $x = hv/kT_\gamma$  we chose  $T_\gamma = T_0(1+z)$  and describe the distortion with respect to a blackbody at the temperature of the electrons  $T_{\text{bb}} = T_e \equiv T_{\text{RJ}}$ .

#### 3.2 Compton scattering

The contribution of CS by thermal electrons to the r.h.s. of Eq. (13), including first-order relativistic corrections, may be compactly written as (compare Challinor & Lasenby 1998; Itoh et al. 1998; Sazonov & Sunyaev 1998; Chluba 2005):

$$\left. \frac{dn}{d\tau} \right|_{\text{CS}} = \frac{\theta_e}{x^2} \frac{\partial}{\partial x} x^4 \left[ F + \theta_e \left\{ \frac{5}{2} F + \frac{21}{5} x \frac{\partial F}{\partial x} + \frac{7}{10} x^2 \left[ \phi F_- (\phi - 6 F_-) + \frac{\partial^2 F_+}{\partial x^2} \right] \right\} \right] + \mathcal{O}(\theta_e^2). \quad (14)$$

Here,  $\theta_e = kT_e/m_e c^2$ ,  $\phi = T_\gamma/T_e$ , and we defined the photon flux functions  $F = \partial_x n_\gamma + \phi n(n+1)$  and  $F_\pm = F \pm \phi n(n+1)$ . The first term in the bracket represents the Kompaneets equation (Kompaneets 1956), whereas the terms proportional to  $\theta_e^2$  arise from first-order relativistic corrections, i.e., higher order Doppler effect [ $\approx \mathcal{O}(\theta_e^2)$ ], recoil effect [ $\approx \mathcal{O}((hv/m_e c^2)^2)$ ] and cross terms [ $\approx \mathcal{O}(\theta_e hv/m_e c^2)$ ]. We assumed that the photon distribution is isotropic even if it deviates from the pure blackbody shape. One can readily verify that for a blackbody spectrum at temperature  $T_\gamma \equiv T_e$  the flux  $F$  vanishes; also the combination of terms related to  $F_\pm$  vanishes, so that in equilibrium CS leaves the photon distribution unaffected.

We now linearize Eq. (14) with respect to  $\mu \ll 1$ . The effect of CS on the blackbody,  $n_{\text{pl}}(x T_\gamma/T_e)$ , which we chose as a reference (see previous section), vanishes identically. For terms that involve

$\mu(\tau, x) = \mu_{\text{co}}(\tau) \hat{\mu}(\tau, x)$ , as defined in Eq. (2), with  $F \rightarrow -G\mu/x$ , the linearized Compton collision term reads:

$$\left. \frac{dn_\gamma}{d\tau} \right|_{\text{CS}} \approx -\theta_\gamma \frac{G}{x} \left[ x^2 \mu'' + 2g_1(x) x \mu' \right] \left( 1 + \frac{5}{2} \theta_\gamma \right) - \theta_\gamma^2 \frac{G}{x} \left[ -\frac{42}{5} f_1(x) x \mu' + \frac{21}{5} f_2(x) x^2 \mu'' \right] - \theta_\gamma^2 \frac{G}{x} \left[ \frac{28}{5} f_3(x) x^3 \mu''' + \frac{7}{10} x^4 \mu'''' \right], \quad (15a)$$

where  $\theta_\gamma = hv/kT_\gamma$ . Here, primes denote derivatives with respect to  $x$  and we defined the frequency-dependent functions

$$g_1(x) = -\frac{Y_{\text{SZ}}}{2G} \approx 1 - \frac{x^2}{12} + \frac{x^4}{720} \quad (15b)$$

$$f_1(x) = xY_{\text{SZ}} - xG(1 + C_x - 5S_x) \approx 1 + \frac{5x^2}{12} - \frac{29x^4}{720} \quad (15c)$$

$$f_2(x) = 5 + \frac{x^2}{6} + 2xG(1 - 3S_x) \approx 1 - \frac{x^2}{2} + \frac{x^4}{60} \quad (15d)$$

$$f_3(x) = \frac{1}{2} [1 + g_1(x)] \approx 1 - \frac{x^2}{24} + \frac{x^4}{1440}, \quad (15e)$$

with  $C_x = x \coth(x/2)$ ,  $S_x = x/\sinh(x/2)$  and the y-distortion shape  $Y_{\text{SZ}}(x) = G(x)[x \coth(x/2) - 4]$ . For the functions  $g_1$  and  $f_i$ , we also gave the limits  $x \ll 1$ . These functions are defined such that to leading order in  $x \ll 1$  they are all equal to unity.

If we look at the terms in Eq. (15) in this limit, we can see that they all are of similar order in  $x$ . However, the higher derivative terms are suppressed by an extra factor  $\theta_\gamma \ll 1$ . This allows us to first consider only the lowest order solution, neglecting terms  $\propto \theta_\gamma^2$ . Relativistic correction can then be added as perturbations to the non-relativistic solution (Sect. 6).

#### 3.2.1 Effect on the photon energy density

By integrating Eq. (14) over  $x^2 dx$ , it is trivial to confirm that the collision terms for CS conserve the photon number. To compute the energy exchange between electrons and photons via Compton scattering, we integrate Eq. (14) over  $x^3 dx$ . This yields (see Sazonov & Sunyaev 2001; Chluba 2005, for similar expressions):

$$-\frac{1}{a^4 \rho_\gamma^{\text{pl}}} \left. \frac{da^4 \rho_\gamma}{d\tau} \right|_{\text{CS}} \approx \frac{\mathcal{I}_4}{\mathcal{G}_3} \theta_\gamma - 4\theta_e - 10\theta_e^2 - \frac{21}{5} \frac{\mathcal{G}_5}{\mathcal{G}_3} \theta_\gamma^2 + \frac{\mathcal{I}_4}{\mathcal{G}_3} \left( \frac{47}{2} - \frac{21}{5} \frac{\mathcal{H}_6}{\mathcal{I}_4} \right) \theta_e \theta_\gamma, \quad (16)$$

where the integrals  $\mathcal{G}_i$  are defined by Eq. (4b) with  $n_{\text{pl}}(x) \rightarrow n(x)$ . For convenience, we also introduced  $\mathcal{I}_i = \int x^i n(1+n) dx$  and  $\mathcal{H}_i = \int x^i (\partial_x n)^2 dx$ . The first two terms in Eq. (16) determine the usual contributions in the non-relativistic limit, while the other terms capture relativistic corrections. From this expression, the Compton equilibrium temperature can be given by

$$\theta_{\text{eq}} \approx \theta_{\text{eq}}^{\text{nr}} \left[ 1 - \frac{21}{5} \theta_{\text{eq}}^{\text{nr}} \left\{ \frac{4\mathcal{G}_3\mathcal{G}_5}{\mathcal{I}_4^2} + \frac{\mathcal{H}_6}{\mathcal{I}_4} - 5 \right\} \right] \approx \theta_\gamma [1 + \delta\gamma_1 - \theta_\gamma \delta\gamma_2], \quad (17a)$$

using the fact that the non-relativistic Compton equilibrium temperature is small, i.e.,  $\theta_{\text{eq}}^{\text{nr}} = \theta_\gamma \mathcal{I}_4/4\mathcal{G}_3 \ll 1$ , and keeping terms up to second order of  $\theta_{\text{eq}}^{\text{nr}}$  only. In the second line, we linearized the problem for small distortion. The corrections to the integrals thus

take the form

$$\begin{aligned} \delta\gamma_1 &= \frac{\delta\mathcal{I}_4}{4\mathcal{G}_3^{\text{PI}}} - \frac{\delta\mathcal{G}_3}{\mathcal{G}_3^{\text{PI}}} \quad (17b) \\ \delta\gamma_2 &= \frac{21}{20} \left[ \left( \frac{\delta\mathcal{G}_3}{\mathcal{G}_3^{\text{PI}}} - \frac{\mathcal{G}_5^{\text{PI}}}{\mathcal{G}_3^{\text{PI}}} \frac{\delta\mathcal{G}_3}{\mathcal{G}_3^{\text{PI}}} \right) + 4 \left( \frac{\delta\mathcal{H}_6}{\mathcal{I}_4^{\text{PI}}} - \frac{\mathcal{H}_6^{\text{PI}}}{\mathcal{I}_4^{\text{PI}}} \frac{\delta\mathcal{I}_4}{\mathcal{I}_4^{\text{PI}}} \right) \right] - 4\pi^2\gamma_1, \end{aligned}$$

where we only have to compute the terms for the spectral distortions, as indicated by ‘ $\delta$ ’. These do, however, generally include the difference in the spectrum due to  $T_{\text{RJ}} = T_e \neq T_\gamma$ , so that for example  $\delta\gamma_1 \approx \Delta T_e/T_\gamma + \delta\gamma_1^h$ , where now the second term only includes contributions  $\propto \hat{\mu}$ . Note also that  $\mathcal{I}_4^{\text{PI}} \equiv 4\mathcal{G}_3^{\text{PI}}$ .

With Eq. (17a) one can recast Eq. (16) into the form

$$\begin{aligned} -\frac{d \ln a^4 \rho_\gamma}{d\tau} \Big|_{\text{CS}} &\approx 4(\theta_{\text{eq}} - \theta_e) \left[ 1 + \theta_{\text{eq}}^{(0)} \left( \frac{21}{5} \frac{\mathcal{H}_6}{\mathcal{I}_4} - \frac{37}{2} \right) \right] \\ &\approx 4(\theta_{\text{eq}} - \theta_e) [1 - 17.239\theta_\gamma], \quad (18) \end{aligned}$$

where in the second line we again linearized the problem for small distortion, with  $2\pi^2 - 5/2 \approx 17.239$ . This shows that at the lowest order, the distortions only affect the Compton equilibrium temperature. Furthermore, the time-scale on which Compton equilibrium is achieved is increased by  $\approx [1 - 17.239\theta_\gamma]^{-1}$ , which at  $z \approx 2 \times 10^6$  implies an  $\approx 2\%$  effect. The change in the equilibration time-scale and even the exact value for the equilibrium temperature are, however, not relevant to the final solution for the chemical potential, and only enter the problem at higher perturbation order.

### 3.3 Double Compton and Bremsstrahlung emission

The contribution of DC scattering and BR to the r.h.s. of the photon Boltzmann equation (13) can be written in the form (cf., Rybicki & Lightman 1979; Lightman 1981; Thorne 1981):

$$\frac{dn_\gamma}{d\tau} \Big|_{\text{DC+BR}} = \frac{e^{-x}}{x^3} [1 - n_\gamma(e^{\phi x} - 1)] \Lambda(x, \theta_\gamma, \theta_e) \quad (19)$$

where the emission coefficient  $\Lambda$  is given by the sum of the contribution due to double Compton scattering and Bremsstrahlung,  $\Lambda = \Lambda_{\text{DC}} + \Lambda_{\text{BR}}$ . As this expression shows, the leading order emission term scales  $\propto x^{-3}$ . This is the reason why at low frequencies the spectrum returns to a blackbody in equilibrium with the electrons after a very short time.

At lowest order in the distortion, we have

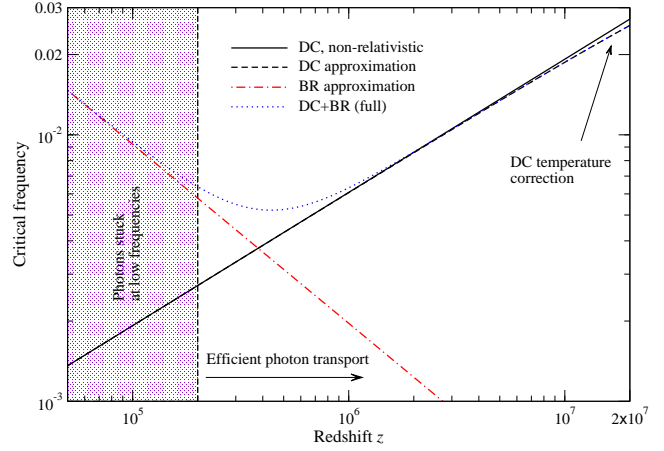
$$\frac{dn_\gamma}{d\tau} \Big|_{\text{DC+BR}} \approx \frac{\Lambda}{x^4} (1 - e^{-x}) \mathcal{G} \mu \quad (20)$$

where we inserted Eq. (1) and (2). Note that the emission coefficient,  $\Lambda$ , is evaluated at the background level, i.e., it just depends on  $x$  and  $\theta_\gamma$ . At low frequencies, its dependence on  $x$  is rather weak, a fact that can be used to simplify the problem. Also, since we chose the temperature of the electrons as reference ( $T_{\text{bb}} = T_e$ ), no additional emission/absorption term related to the blackbody part arises, one of the reasons for this definition.

#### 3.3.1 Double Compton scattering

Due to the large entropy of the Universe, DC emission dominates over BR at high redshifts ( $z \gtrsim 4 \times 10^5$ ). The DC scattering emission coefficient can be expressed as

$$\Lambda_{\text{DC}}(x, \theta_\gamma) = \frac{4\alpha}{3\pi} \theta_\gamma^2 e^x g_{\text{dc}}(x, \theta_\gamma), \quad (21)$$



**Figure 1.** Critical frequency,  $x_c$ , defined by Eq. (25) as a function of  $z$ . Photon transport is inefficient below  $z \approx 2 \times 10^5$  so that the distortion visibility function quickly approaches unity. DC temperature corrections become noticeable at  $z \gtrsim 10^6$ . The approximations are from Eq. (26) and (27).

where  $\alpha$  is the fine structure constant and  $g_{\text{dc}}(x, \theta_\gamma)$  is the effective DC Gaunt factor. It has the form (Chluba 2005; Chluba et al. 2007; Chluba & Sunyaev 2012):

$$g_{\text{dc}}(x, \theta_\gamma) \approx \frac{\mathcal{I}_4^{\text{PI}} H_{\text{dc}}^{\text{PI}}(x)}{1 + 14.16\theta_\gamma}, \quad (22)$$

where  $\mathcal{I}_4^{\text{PI}} = \int x^4 n_{\text{PI}}(n_{\text{PI}} + 1) dx = 4\pi^4/15 \approx 25.976$ , and

$$H_{\text{dc}}^{\text{PI}}(x) \approx e^{-2x} \left[ 1 + \frac{3}{2}x + \frac{29}{24}x^2 + \frac{11}{16}x^3 + \frac{5}{12}x^4 \right]. \quad (23)$$

Here, we included the first-order relativistic correction in the photon temperature. At low frequencies,  $\Lambda_{\text{DC}}(x, \theta_\gamma) \propto (1 + \frac{1}{2}x)$ . In the non-relativistic limit, we have  $H_{\text{dc}}^{\text{PI}}(x) \approx e^{-x}$  and  $g_{\text{dc}}(x, \theta_\gamma) \approx \mathcal{I}_4^{\text{PI}} e^{-x}$ .

#### 3.3.2 Bremsstrahlung

At low redshifts ( $z \lesssim 4 \times 10^5$ ), Bremsstrahlung becomes the main source of soft photons. One can define the Bremsstrahlung emission coefficient by (cf. Burigana et al. 1991; Hu & Silk 1993)

$$\Lambda_{\text{BR}}(x, \theta_e) = \frac{\alpha \lambda_e^3}{2\pi \sqrt{6\pi}} \frac{\theta_e^{-7/2} e^{x(1-\phi)}}{\phi^3} \sum_i Z_i^2 N_i g_{\text{ff}}(Z_i, x, \theta_e). \quad (24)$$

Here,  $\lambda_e = h/m_e c$  is the Compton wavelength of the electron, and  $Z_i$ ,  $N_i$  and  $g_{\text{ff}}(Z_i, x, \theta_e)$  are the charge, the number density and the BR Gaunt factor for a nucleus of the atomic species  $i$ , respectively.

Various simple analytical approximations exist (Rybicki & Lightman 1979), but nowadays more accurate fitting formulae, valid over a wide range of temperatures and frequencies, may, for example, be found in Nozawa et al. (1998) and Itoh et al. (2000). We find, however, that the differences introduced by the various approximations for  $\Lambda_{\text{BR}}$  are not very important, both for the shape of the distortion and the distortion visibility function. It is pretty straightforward to include them consistently, so that for any of the computations we just use the expressions of Itoh et al. (2000).

#### 3.3.3 Critical frequency

For the computations below, we need the critical frequency determined by (compare Sunyaev & Zeldovich 1970; Danese & de Zotti

1982; Burigana et al. 1991; Hu & Silk 1993)

$$x_c(\tau) \equiv \sqrt{\Lambda(x_c)/\theta_\gamma}. \quad (25)$$

Close to  $x_c$ , the rate of Compton scattering equals that of photon emission/absorption, roughly defining the maximum of the photon production, with some corrections which are important for the late-time evolution (see Sect. 5.2). For DC alone, one has

$$x_c^{\text{DC},0} \approx \sqrt{\frac{4\alpha}{3\pi} \theta_\gamma I_4^{\text{PI}}} \approx 8.60 \times 10^{-3} \left[ \frac{1+z}{2 \times 10^6} \right]^{1/2} \quad (26a)$$

$$x_c^{\text{DC}} \approx x_c^{\text{DC},0} \left( 1 + \frac{1}{2} x_c^{\text{DC},0} \right)^{1/2} / \left( 1 + 14.16 \theta_\gamma \right)^{1/2}, \quad (26b)$$

where for  $x_c^{\text{DC},0}$  we neglected relativistic corrections to the DC Gaunt factor, while we included them for  $x_c^{\text{DC}}$ .

The critical frequency for DC is shown in Fig. 1. One can see that at high redshifts, the temperature correction reduces the critical frequency notably. As we will see below, this implies that thermalization should be less efficient, since the frequency at which most photons are produced decreases. At the thermalization redshift  $z \approx 2 \times 10^6$ , the temperature correction to the critical frequency  $x_c$  is roughly 0.5% and it reaches  $\approx 1\%$  at  $z \approx 4 \times 10^6$ . Although this appears to be small, since the critical frequency enters the problem through an integral, the cumulative effect matters so that the correction is amplified and hence significant (see below). At  $z \gtrsim \text{few} \times 10^7$ , the number of electrons and positrons becomes comparable to the number of photons, so that there the thermalization efficiency increases vastly. A treatment of the thermalization problem in that era is, however, beyond the scope of this paper.

For BR alone, we determined the critical frequency numerically using the expressions from Itoh et al. (2000) and assuming a helium mass fraction of  $Y_p = 0.24$ . We find<sup>7</sup>

$$x_c^{\text{BR}} \approx 1.23 \times 10^{-3} \left[ \frac{1+z}{2 \times 10^6} \right]^{-0.672} \quad (27)$$

to work very well. Comparing with Eq. (26a), we can see that at the thermalization redshift  $z \approx 2 \times 10^6$  BR contributes about 10% to the value of the critical frequency. However, the contribution drops rapidly towards higher redshifts (Fig. 1). To percent precision, the total critical frequency is  $x_c^{\text{c}} \approx (x_c^{\text{DC}})^2 + (x_c^{\text{BR}})^2$  (Hu & Silk 1993).

### 3.3.4 Transport of photons towards higher frequencies

Photons, produced by DC and BR at low frequencies, can only help thermalizing the distortion if they actually reach the high-frequency part of the spectrum before being reabsorbed or before no time to up-scatter is left. We thus need to estimate until when the photon redistribution by Compton scattering is efficient. The evolution of a narrow line within an ambient blackbody radiation field can be computed for any value of the Compton  $y$ -parameter  $y_\gamma = \int \theta_\gamma d\tau$  using the analytic solutions from Chluba & Sunyaev (2008). These are valid as long as electron recoil remains negligible. They include the extra drift of photons towards lower frequencies caused by stimulated effects. The average photon energy thus increases like  $\nu \approx \nu_0 e^{2y_\gamma}$  (the classical result of Zeldovich & Sunyaev 1969, does not include stimulated effects and thus gives  $\nu \approx \nu_0 e^{4y_\gamma}$ ).

Assuming that photons start their journey around  $x \approx x_c \ll 1$ , we can estimate the time it takes for the average photon distribution

to reach  $x \approx 1$ . This then implies that we need  $y_\gamma \gtrsim (1/2) \ln x_c^{-1}$ . Estimating the total  $y$ -parameter from some initial redshift  $z$  until today using  $y_\gamma \approx 4.3[(1+z)/3 \times 10^5]^2$  (radiation domination) then implies that at  $z \lesssim 10^5 \sqrt{\ln x_c^{-1}}$  photons can no longer be sufficiently up-scattered. This implies that around  $z \approx 2 \times 10^5$ , photon transport no longer is efficient enough to replenish the high-frequency photon deficit. At that moment, the low- and high-frequency parts of the photon distribution practically decouple, and the effective chemical potential (or more precisely the high-frequency photon number) freezes in. Below  $z \approx 2 \times 10^5$ , basically all the released energy remains in the distortion, and the distortion visibility function thus becomes unity. This statement is in fact irrespective of the shape of the distortion at late times ( $\mu$ ,  $y$  and residual distortion). Our numerical computations confirm this statement (see Fig. 9).

### 3.3.5 Effect on number and energy density

In contrast to Compton scattering, DC and BR change the photon number density. To compute the effective photon production and associated change in the photon energy density caused by this, we can simply integrate Eq. (20) over  $x^2 dx$  and  $x^3 dx$ , finding:

$$\begin{aligned} \frac{d \ln a^3 N_\gamma}{d\tau} \Big|_{\text{DC,BR}} &\approx \frac{1}{\mathcal{G}_2^{\text{PI}}} \int \frac{\Lambda(x)}{x[e^x - 1]} \mu dx \\ \frac{d \ln a^4 \rho_\gamma}{d\tau} \Big|_{\text{DC,BR}} &\approx \frac{1}{\mathcal{G}_3^{\text{PI}}} \int \frac{\Lambda(x)}{e^x - 1} \mu dx = \mathcal{H}_{\text{em}}. \end{aligned} \quad (28)$$

In general these integrals have to be performed numerically once the solution for the low-frequency distortion is known. In particular, the non-trivial frequency dependence of the BR Gaunt factor renders it difficult to find accurate analytic approximations. This is because at the percent level the integrals picks up small contributions even at higher frequencies and the assumption of constant  $\Lambda$  is not well justified. The integrand scales like  $\Lambda(x)\mu(x)/x^2$  at low frequencies, so that for  $\Lambda(x) \approx \text{const}$  we need  $\mu$  to vanish faster than  $\approx x^2$  to obtain a finite result for  $d \ln a^3 N_\gamma / d\tau$ . For  $d \ln a^3 \rho_\gamma / d\tau$ , we find  $\mu \propto x$  is sufficient, at least to first order in perturbations. This shows that the integrals in Eq. (28) both diverge for constant  $\mu$ , but the lowest order solution (see Sect. 5.2)  $\mu = e^{-x/x}$  is sufficient to regularize these expressions.

We also mention a slight inconsistency of the formulation that is present in all treatments of the problem so far. The energy needed for the production of DC photons is taken partially (in the limit of resting electron in which the DC Gaunt factor is derived in fact fully!) from the photon field itself: the incoming high-energy photons scatter and redistribute in energy, giving rise to a small correction to the Compton process and heat exchange with the electrons in addition to the photon emission. This implies that the emission integral  $\mathcal{H}_{\text{em}}$  should be slightly smaller; we are, however, going to neglect this effect, given that the total energy used up by photon production is small. For additional discussion, see Chluba (2005).

## 4 EVOLUTION OF THE MATTER TEMPERATURE

At high redshifts, electrons and photons rapidly exchange energy via Compton scattering, while Coulomb interactions keep electrons and baryons in equilibrium at one temperature,  $T_m = T_e$ . The electrons also cool by BR and DC emission; in addition, the matter (electrons plus baryons) in the Universe cools because of the adiabatic expansion of the medium. And finally, energy release can heat the matter and thereby increase its temperature.

<sup>7</sup> Note that again we can evaluate  $\Lambda$  assuming  $T_e = T_\gamma$ .

Including all these processes, the evolution equation for the matter temperature can be written as (e.g., Chluba & Sunyaev 2012)

$$\frac{d\theta_e/\theta_\gamma}{d\tau} = \frac{\dot{Q}_e}{\alpha_h\theta_\gamma} + \frac{4\tilde{\rho}_\gamma^*}{\alpha_h\theta_\gamma}[\theta_{\text{eq}} - \theta_e] - \frac{\tilde{\rho}_\gamma}{\alpha_h\theta_\gamma}\mathcal{H}_{\text{em}} - H t_C \frac{\alpha_h^{\text{nr}} - \xi_h}{\alpha_h\theta_\gamma}\theta_e,$$

$$\alpha_h = \alpha_h^{\text{nr}} + \xi_h, \quad \xi_h = \frac{15}{4}\theta_e \frac{1 - \frac{1}{2}Y_p}{1 - Y_p} N_e \approx 4.342 N_e \theta_e. \quad (29)$$

Here,  $k\alpha_h^{\text{nr}} = \frac{3}{2}k[N_e + N_H + N_{\text{He}}]$  denotes the matter heat capacity, and  $\xi_h$  defines its lowest order temperature correction<sup>8</sup>. The number densities,  $N_i$ , are for free electrons ( $i='e'$ ), total number of hydrogen ( $i='H'$ ) and helium ( $i='He'$ ) nuclei, and  $Y_p \approx 0.24$  is the helium mass fraction. Furthermore, we introduced the energy injection term,  $\dot{Q}_e \equiv (m_e c^2)^{-1} dQ_e/d\tau$ , which for example could be caused by some decaying or annihilating particles; the energy density of the photon field in units of electron rest mass is defined as  $\tilde{\rho}_\gamma = \kappa_\gamma \theta_\gamma^4 \mathcal{G}_3$  with  $\kappa_\gamma = 8\pi\lambda_e^{-3} \approx 1.760 \times 10^{30} \text{ cm}^{-3}$  and  $\lambda_e$  denoting the Compton wavelength of the electron. Finally, the Compton equilibrium temperature,  $\theta_{\text{eq}}$ , is given by Eq. (17a), which includes first-order relativistic corrections to the problem and the effect of spectral distortions. We also used expression (18) for the Compton energy exchange between electrons and photons. This is indicated by the asterisk on  $\tilde{\rho}_\gamma$ , which means  $\tilde{\rho}_\gamma^* = \tilde{\rho}_\gamma[1 - 17.239\theta_\gamma]$ . The cooling by DC and BR is defined as  $\mathcal{H}_{\text{em}} \equiv d \ln a^4 \rho_\gamma / d\tau|_{\text{DC, BR}}$ .

#### 4.1 Perturbative solution for $T_e$

Without knowing the exact solution of the CMB spectrum, it is still possible to obtain a solution for  $T_e$ , because for conditions in the early Universe it evolves along a sequence of quasi-stationary stages. Assuming that only the matter is heated directly by the energy release, this then allows us to eliminate  $d \ln a^4 \rho_\gamma / d\tau$  in Eq. (9b). Setting  $d(\theta_e/\theta_\gamma)/d\tau \approx 0$ , from Eq. (29) we thus find

$$\theta_e^{(0)} \approx \theta_{\text{eq}}^{(0)} + \frac{\dot{Q}_e}{4\rho_\gamma^*} - \frac{\tilde{\rho}_\gamma}{4\tilde{\rho}_\gamma^*}\mathcal{H}_{\text{em}}^{(0)} - \frac{H t_C \alpha_h^{\text{nr}} \theta_\gamma}{4\tilde{\rho}_\gamma^*} [1 - \lambda_h \theta_\gamma]. \quad (30)$$

with  $\lambda_h \theta_\gamma = \xi_h(\theta)/\alpha_h^{\text{nr}} \approx 1.39\theta_\gamma$ . This means that at the lowest order in perturbations, the Compton energy-exchange term reads

$$\left. \frac{d \ln a^4 \rho_\gamma^{(0)}}{d\tau} \right|_{\text{CS}} \approx \frac{\dot{Q}_e}{\rho_\gamma} - \mathcal{H}_{\text{em}}^{(0)} - \frac{H t_C \alpha_h^{\text{nr}} \theta_\gamma}{\tilde{\rho}_\gamma} [1 - \lambda_h \theta_\gamma]. \quad (31)$$

Therefore the net change in the photon energy density is

$$\left. \frac{d \ln a^4 \rho_\gamma^{(0)}}{d\tau} \right|_{\text{CS}} \approx \left. \frac{d \ln a^4 \rho_\gamma}{d\tau} \right|_{\text{CS}} + \mathcal{H}_{\text{em}}^{(0)} \approx \frac{\dot{Q}_e}{\rho_\gamma} - \frac{H t_C \alpha_h^{\text{nr}} \theta_\gamma}{\tilde{\rho}_\gamma} [1 - \lambda_h \theta_\gamma].$$

The emission integral and also the precise value of the Compton equilibrium temperature dropped out of the problem. At lowest order perturbation theory, the spectral distortion therefore does not directly affect the heat exchange between electrons and photons. Higher order temperature corrections only enter the thermalization problem via the heat capacity of the electrons. This reduces the distortion created by the adiabatic cooling of matter (Chluba & Sunyaev 2012) by a small amount. However, the effect only reaches 1 percent at  $z \approx 1.6 \times 10^7$ , where the distortion visibility is already extremely small (see next Section). This effect can therefore be safely neglected.

Neglecting higher order corrections to the adiabatic cooling

<sup>8</sup> Only the contribution from electrons matters here, since the terms for the baryons are suppressed by ratios of the masses (Chluba 2005).

term, and realizing that the corrections to  $\mathcal{H}_{\text{em}}$  are always canceled by the corresponding photon emission term (a consequence of energy conservation), we find the correction to the electron-photon energy-exchange term  $d \ln a^4 \rho_\gamma^{(1)}/d\tau \approx -(\alpha_h \theta_\gamma / \tilde{\rho}_\gamma) \partial_\tau (T_e^{(0)}/T_\gamma)$ . Knowing the solution for the spectral distortion at lowest order in perturbation theory, one can compute  $d(T_e^{(0)}/T_\gamma)/d\tau$  using Eq. (9a) to close the system of equations. This in principle allows us to obtain the next-order correction to the photon distribution. As we will see below,  $d(T_e^{(0)}/T_\gamma)/d\tau \approx \mathcal{O}(x_c)$ . In comparison with any of the external heating terms,  $\dot{Q}_e$ , the correction is suppressed by an additional factor of the photon-to-baryon ratio, so that from the practical point of view this correction again can be neglected. Henceforth, we will thus simply write the photon heating term as  $d \ln a^4 \rho_\gamma / d\tau \approx \dot{Q}_e / \rho_\gamma - [H t_C + d(T_e^{(0)}/T_\gamma)/d\tau] \alpha_h \theta_\gamma / \tilde{\rho}_\gamma \equiv \dot{Q}_e^* / \rho_\gamma$ . This defines the effective photon heating rate caused by energy release that initially only affects the temperature of ordinary matter in the Universe. The second term will be neglected in our discussion.

## 5 SOLUTION FOR THE SPECTRAL DISTORTIONS IN THE LIMIT OF SMALL CHEMICAL POTENTIAL

In this section, we develop a perturbative treatment for the approximation to the chemical potential,  $\mu(\tau, x)$ . Before adding higher order corrections, we briefly recap the classical solution obtained by Sunyaev & Zeldovich (1970). The basic ansatz is that time- and frequency-dependent parts of the solution can be approximately separated:  $\mu(\tau, x) \approx \mu_\infty(\tau) \hat{\mu}(x)$ . This means that  $\mu(\tau, x)$  only evolves very *slowly* with time, moving along a sequence of quasi-stationary stages, with the main time dependence being captured by an overall amplitude factor, while the *shape* of the distortion is fixed to  $\approx \hat{\mu}(x)$ . Here, we go beyond this approximation. It is furthermore clear, that DC and BR emission are effective only at rather small  $x$ , so that one can expect to find the main frequency dependence of the solution there, while at much higher frequencies the solution varies only slowly with  $x$ . This introduces an energy scale, suggesting the scaling  $x \rightarrow x_c \xi$ , where  $x_c$  is a critical frequency at which the spectrum changes rapidly. This provides a natural perturbation parameter,  $x_c$ , with corrections being ranked by their order in  $x_c \ll 1$ .

### 5.1 Integral solution

With the general ansatz  $\mu(\tau, x) = \mu_\infty(\tau) \hat{\mu}(\tau, x)$ , we can already write an integral solution to Eq. (9b). In contrast to the original works, we do not assume that  $\hat{\mu}(\tau, x) \approx \hat{\mu}(x)$ , but explicitly include slow time dependence in the shape of the distortion. Scaling out the main terms, with Eq. (28) the photon production rate is given by

$$\left. \frac{d \ln a^3 N_\gamma}{d\tau} \right|_{\text{DC, BR}} \approx \frac{\theta_\gamma x_c}{\mathcal{G}_2^{\text{PI}}} \mu_\infty \mathcal{I}_{\hat{\mu}}$$

$$\mathcal{I}_{\hat{\mu}}(\tau) = \int \frac{\Lambda(x)}{\Lambda(x_c)} \frac{x_c \hat{\mu}(\tau, x)}{x[e^x - 1]} dx \quad (32)$$

Generally,  $\mathcal{I}_{\hat{\mu}}$  is very close to unity with corrections of  $\mathcal{O}(x_c)$  [i.e., higher order terms]. Inserting this into Eq. (9b), with the definition of the effective heating rate  $\dot{Q}_e^*$  given in Sect. 4.1 we find

$$\frac{d\mu_\infty}{d\tau} \approx \gamma_\rho \frac{\dot{Q}_e^*}{\rho_\gamma} - \gamma_N \theta_\gamma x_c \mathcal{I}_{\hat{\mu}} \mu_\infty$$

$$\gamma_\rho = 3/\kappa_\rho^c \approx 1.401, \quad \gamma_N = 4/(\mathcal{G}_2^{\text{PI}} \kappa_\rho^c) \approx 0.7769, \quad (33)$$



where  $\kappa_p^c = 2.1419$ . Then, by introducing the optical depth

$$\begin{aligned} \tau_\mu(z) &= \tau_{\mu,0}(z) + \Delta\tau_\mu(z), \\ \tau_{\mu,0}(z) &= \gamma_N \int_0^\infty \frac{\theta_\gamma x_c}{t_c} \frac{dz'}{H(1+z')}, \end{aligned} \quad (34a)$$

$$\Delta\tau_\mu(z) = \gamma_N \int_0^\infty \frac{\theta_\gamma x_c}{t_c} (\mathcal{I}_{\hat{\mu}} - 1) \frac{dz'}{H(1+z')}, \quad (34b)$$

and assuming that there is no initial distortion at very early times, we can finally write

$$\mu_\infty(z) \approx 1.401 \int_z^\infty \frac{\dot{Q}_e^* e^{-\tau_\mu(z',z)} dz'}{\rho_\gamma H(1+z')} \quad (35)$$

with  $\tau_\mu(z, z') = \tau_\mu(z) - \tau_\mu(z')$ . Equation (35) gives the formal solution for  $\mu_\infty(z)$  at high redshifts, including all important effects that affect the integrated energy and number density of the photon field when heating the ordinary matter by some process occurs. The optical depth,  $\tau_\mu$ , is affected by (i) the shape of the spectral distortion at low frequencies (which enters  $\mathcal{I}_{\hat{\mu}}$ ; see Fig. 7) and (ii) the precise redshift dependence of the critical frequency,  $x_c$  (see Fig. 1). At lowest order in  $x_c$ , we have  $\mathcal{I}_{\hat{\mu}} \approx 1$ , so that  $\Delta\tau_\mu(z) \approx 0$ .

### 5.1.1 Total photon emission

Inserting the solution for  $\mu$  into Eq. (32), we can directly compute the total change in the number of photons over the energy release history. This yields the simple expression

$$\begin{aligned} \frac{\Delta N_\gamma}{N_\gamma} &\approx \frac{3}{4} \int_z^\infty \frac{\partial \tau_\mu(z')}{\partial z'} \left( \int_{z'}^\infty \frac{\dot{Q}_e^* e^{-\tau_\mu(z'',z')}}{\rho_\gamma H(1+z'')} dz'' \right) dz' \\ &= \frac{3}{4} \int_z^\infty \frac{\partial e^{-\tau_\mu(z')}}{\partial z'} \left( \int_{z'}^\infty \frac{\dot{Q}_e^* e^{-\tau_\mu(z'',z')}}{\rho_\gamma H(1+z'')} dz'' \right) dz' \\ &= \frac{3}{4} \int_z^\infty \frac{\dot{Q}_e^*}{\rho_\gamma} (1 - e^{-\tau_\mu(z',z)}) \frac{dz'}{H(1+z')} \\ &= \frac{3}{4} \left( \frac{\Delta \rho_\gamma}{\rho_\gamma} - \frac{\Delta \rho_\gamma}{\rho_\gamma} \Big|_{\text{dist}} \right) \equiv \frac{3}{4} \frac{\Delta \rho_\gamma}{\rho_\gamma} \Big|_T, \end{aligned} \quad (36)$$

where we used the definition of  $\tau_\mu$  given in Eq. (34) and identified the energy density changes

$$\frac{\Delta \rho_\gamma}{\rho_\gamma} = \int_z^\infty \frac{\dot{Q}_e^*}{\rho_\gamma} \frac{dz'}{H(1+z')} \quad (37a)$$

$$\frac{\Delta \rho_\gamma}{\rho_\gamma} \Big|_{\text{dist}} = \int_z^\infty \frac{\dot{Q}_e^* e^{-\tau_\mu(z',z)}}{\rho_\gamma} \frac{dz'}{H(1+z')} \quad (37b)$$

$$\frac{\Delta \rho_\gamma}{\rho_\gamma} \Big|_T = \frac{\Delta \rho_\gamma}{\rho_\gamma} - \frac{\Delta \rho_\gamma}{\rho_\gamma} \Big|_{\text{dist}}. \quad (37c)$$

These expressions give the following picture: the total energy release branches into temperature shift and distortion. At any moment, part of the total energy release,  $\Delta \rho_\gamma / \rho_\gamma$ , is stored in the distortion (non-blackbody) and carries an energy density change  $\Delta \rho_\gamma / \rho_\gamma |_{\text{dist}}$ . The remainder  $\Delta \rho_\gamma / \rho_\gamma |_T$  is carried by the blackbody and is associated with a shift of the initial blackbody temperature  $T_\gamma^{\text{in}}$  by  $(T_N^* - T_\gamma^{\text{in}}) / T_\gamma^{\text{in}} = \frac{1}{3} \Delta N_\gamma / N_\gamma = \frac{1}{4} \Delta \rho_\gamma / \rho_\gamma |_T$ . The branching ratio between the distortion and temperature parts depends on time and the efficiency of the thermalization process. This defines the distortion visibility function,  $\mathcal{J}(z', z) = e^{-\tau_\mu(z', z)}$ , which determines the energy branching ratio at  $z$  given that the heating occurred at  $z'$ .

### 5.1.2 Single energy release

Assuming that the distortion caused by the adiabatic cooling of matter is negligible, from Eq. (35) we find (cf.

Sunyaev & Zeldovich 1970)

$$\mu_\infty(z) \approx 1.401 \frac{\Delta \rho_\gamma}{\rho_\gamma} e^{-\tau_\mu(z_h, z)} = \mu_\infty^{\text{st}} \mathcal{J}(z_h, z) \quad (38)$$

for a single energy release of  $\Delta \rho_\gamma / \rho_\gamma$  at heating redshift  $z_h$ . Here, we used  $\mu_\infty^{\text{st}} = 1.401 \Delta \rho_\gamma / \rho_\gamma$ . The factor  $e^{-\tau_\mu(z_h, z)}$  is the spectral distortion visibility between the heating redshift  $z_h$  and  $z$ . Explicitly, the amount of energy stored in distortions (independent of its specific shape actually) at any moment is

$$\frac{\Delta \rho_\gamma}{\rho_\gamma} \Big|_{\text{dist}} \equiv \frac{\Delta \rho_\gamma}{\rho_\gamma} e^{-\tau_\mu(z_h, z)}. \quad (39)$$

In the above formulation, only  $\tau_\mu(z_h, z)$  has to be computed precisely to obtain approximations for  $\mu_\infty(z)$ . Both quantities require a solution for  $\hat{\mu}$ , which we shall discuss below. According to Eq. (11), the added number of photons therefore is

$$\frac{\Delta N_\gamma(t)}{N_\gamma} \approx \frac{3}{4} \frac{\Delta \rho_\gamma}{\rho_\gamma} (1 - e^{-\tau_\mu(z_h, z)}) \equiv \frac{3}{4} \frac{\Delta \rho_\gamma(t)}{\rho_\gamma} \Big|_T, \quad (40)$$

again with respect to the initial blackbody at temperature  $T_\gamma^{\text{in}} < T_\gamma$ . From Eq. (10), it also immediately follows that

$$\frac{\Delta T_e(t)}{T_\gamma} \approx \frac{3}{4} \frac{M_3^c}{\kappa_p^c} \left( \frac{M_3(t)}{M_3^c} e^{-\tau_\mu(z_h, z)} - 1 \right) \frac{\Delta \rho_\gamma}{\rho_\gamma}. \quad (41)$$

Initially  $\Delta T_e(t) / T_\gamma = 0$  and as thermalization proceeds,  $T_e = T_{\text{RJ}}$  decreases until  $\Delta T_e(t) / T_\gamma \approx -0.3889 \Delta \rho_\gamma / \rho_\gamma$ , so that in the final stage  $T_N^* = T_\rho^* = T_\gamma^{\text{in}} (1 + \frac{1}{4} \Delta \rho_\gamma / \rho_\gamma)$ . These relations show that  $\mathcal{J}(z_h, z)$  and  $\hat{\mu}(t, x)$  fully characterize the solution of the thermalization problem in terms of photon energy and number density.

### 5.1.3 Numerical computation of $\mathcal{J}(z', z)$

Equations (38)–(40) suggest a procedure to numerically compute the distortion visibility function *without* directly relying on the amplitude of the chemical potential at different frequencies: in the picture given above,  $\mathcal{J}(z_h, z)$  determines how much of the injected energy is available for distortions at redshift  $z$ . Numerically, this means: (i) compute the effective temperature of the photon distribution with respect to the photon number density,  $T_N^*$  [Eq. (5a)] and (ii) subtract the corresponding energy density based on this temperature from the total injected energy density,  $\Delta \rho_\gamma / \rho_\gamma$ . The remainder determines the total amount of energy that went into distortions,  $\Delta \rho_\gamma / \rho_\gamma |_{\text{dist}}$ , and from the final ratio of the injected energy densities one can obtain

$$\mathcal{J}(z_h, z) \approx \frac{\Delta \rho_\gamma / \rho_\gamma |_{\text{dist}}}{\Delta \rho_\gamma / \rho_\gamma}. \quad (42)$$

Since the high-frequency spectrum varies logarithmically with  $x$  [see Eq. (70)], this approach provides a more robust (valid even after the  $\mu$ -era) definition for the visibility function, removing possible ambiguities introduced by comparing the chemical potential at some fixed frequency and different times to construct the visibility function. We will now derive analytic expressions for  $\mathcal{J}(z_h, z)$  in different limits and then compare them with numerical results.

## 5.2 Classical solution of Sunyaev & Zeldovich (1970)

We already mentioned that for the evolution of  $\mu_\infty(\tau)$ , we need to obtain an approximation for  $\ln a^3 N_\gamma / d\tau$  in terms of  $\hat{\mu}(x)$ . From

Eq. (15) and (20) with  $dn/d\tau = -(G/x)\partial_\tau\mu$ , we have the general evolution equation for the chemical potential

$$\frac{\partial\mu}{\partial\tau} - x\partial_\tau\frac{T_e}{T_\gamma} \approx \theta_\gamma [x^2\mu'' + 2g_1(x)x\mu'] - \frac{\Lambda}{x^3}(1 - e^{-x})\mu. \quad (43)$$

We discuss terms of  $\mathcal{O}(\theta_\gamma^2)$  in Sect. 6. Following Sunyaev & Zeldovich (1970), we set the l.h.s. of this equation to zero and go to the limit  $x \ll 1$ , finding

$$0 \approx x^2\mu'' + 2x\mu' - \frac{\Lambda/\theta_\gamma}{x^2}\mu = \partial_x x^2 \partial_x \mu - \frac{\Lambda/\theta_\gamma}{x^2}\mu. \quad (44)$$

Since at low frequencies,  $\Lambda$  varies only slowly with  $x$  (cf. Sect. 3.3), we can replace it by a constant,  $\Lambda(x) \approx \Lambda(x_c) = \theta_\gamma x_c^2$ . The idea is that  $\Lambda(x)$  evaluated at  $x_c$  roughly determines the maximum of the emission. The lowest order solution found by Sunyaev & Zeldovich (1970) therefore reads  $\mu^{(0)}(\tau, x) = \mu_\infty(\tau) e^{-x_c(\tau)/x}$ . This solution becomes constant at high frequencies and vanishes at low frequencies. It does, however, not follow our normalization condition  $\kappa_\rho = \kappa_\rho^c$ , but the deviation is of higher order in  $x_c$  and thus is neglected now.<sup>9</sup>

### 5.2.1 Lowest order solution for $\mu_\infty(\tau)$

To lowest order in  $x_c$ , we can set  $\mathcal{I}_{\hat{\mu}} \approx 1$ . Assuming that energy release occurs only at one single heating redshift,  $z_h$ , we then have  $\tau_{\mu,0}(z_h, 0) \approx \gamma_\mu \int_0^{z_h} (1+z)x_c(z) dz$ . Using  $H(1+z)/\sigma_T N_e c \approx 4.79$  for  $H \approx 2.09 \times 10^{-20} (1+z)^2 \text{sec}^{-1}$  [radiation-dominated era], the coefficient  $\gamma_\mu$  is given by

$$\gamma_\mu \approx \gamma_N \frac{cN_{e,0}\sigma_T}{H_{\text{rad},0}} \frac{kT_0}{m_e c^2} \approx 7.45 \times 10^{-11}, \quad (45)$$

where  $N_{e,0} \approx 1.12 \times 10^{-5} \Omega_b h^2 (1 - Y_p/2) \text{cm}^{-3}$  is the electron number density at  $z = 0$  and  $H_{\text{rad},0} = \Omega_{\text{rel}}^{1/2} H_0 \approx 2.09 \times 10^{-20} \text{sec}^{-1}$ . Here,  $\Omega_{\text{rel}}$  is the density parameter of relativistic species (radiation + neutrinos) and  $H_0$  denotes the Hubble parameter today. Neglecting BR [i.e.  $x_c \propto (1+z)^{1/2}$ ] then yields

$$\mu_\infty(z=0) \approx 1.401 \frac{\Delta\rho_\gamma}{\rho_\gamma} e^{-(z_h/z_{\text{dc}})^{5/2}}, \quad (46)$$

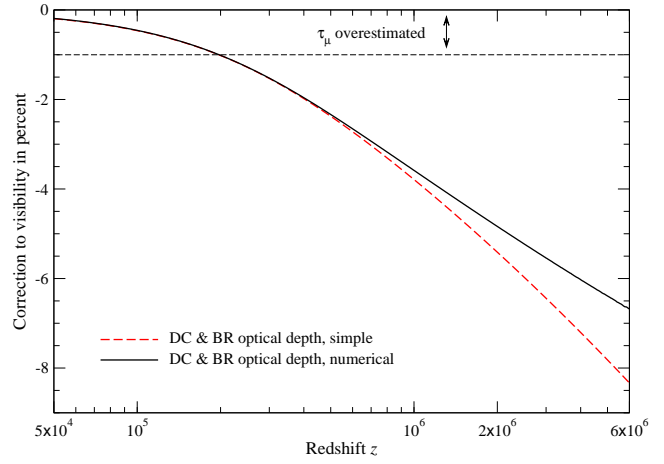
or  $\tau_{\mu,0}^{\text{DC}}(z) = (z/z_{\text{dc}})^{5/2}$ , where the DC thermalization redshift is given by  $z_{\text{dc}} = [(2/5)\gamma_\mu x_c^{\text{DC},0}(z=0)]^{-2/5} \approx 1.98 \times 10^6$ . The exponential factor is the distortion visibility function,  $\mathcal{F}_{\text{DC}} = e^{-(z_h/z_{\text{dc}})^{5/2}}$ , discussed above. Using  $x_c = x_c^{\text{BR}}$  from Eq. (27) instead, we find

$$\tau_{\mu,0}^{\text{BR}}(z) \approx \gamma_\mu \int_0^z (1+z)x_c^{\text{BR}} dz \approx \left( \frac{1+z}{5.27 \times 10^6} \right)^{1.328}. \quad (47)$$

In the classical result, given first by Sunyaev & Zeldovich (1970), the power-law coefficient is  $5/4 = 1.25$  because a different approximation for the BR Gaunt factor was utilized. Similar expressions were also given by Danese & de Zotti (1982) and Hu & Silk (1993). This shows that the thermalization redshift is significantly higher when only BR is included. In addition, the distortion visibility function is less steep at  $z \gtrsim 5.27 \times 10^6$ .

Since for energy release at high redshifts the photon distribution evolves through both the DC- and BR-era, we need to take the

<sup>9</sup> A small improvement is in principle possible here. Using  $\mu^{(0)}(x, \tau)$  one can show that for nearly frequency independent  $\Lambda(x)$  the maximum emission arises around  $x_m \approx x_c/2$  instead of  $x_c$ . Therefore, instead of Eq. (25) one should use  $\Lambda(x_c/2)(1 - e^{-x_c/2})/(x_c/2) = \theta_\gamma x_c^2$  to determine  $x_c$ . This aspect becomes noticeable during the BR-era, but for simplicity we shall treat the associated difference as a correction.



**Figure 2.** Change in the distortion visibility,  $\mathcal{F} = e^{-\tau_{\mu,0}(z,0)}$ , when using the total optical depth including DC and BR. The dotted line is for the simple approximation Eq. (48), while the solid line is obtained by evaluating Eq. (34a) numerically. We compared to  $\mathcal{F}_{\text{DC}} = \exp(-[z/z_{\text{dc}}]^{5/2})$  with thermalization redshift  $z_{\text{dc}} \approx 1.98 \times 10^6$ . The thermalization optical depth correction was computed between  $z$  and  $z = 0$ , so that the correction is slightly overestimated (see text).

full expression for  $x_c$  into account when computing  $\tau_\mu(z, z')$ . With  $x_c \approx [(x_c^{\text{DC}})^2 + (x_c^{\text{BR}})^2]^{1/2}$ , the integral is rather simple and given by

$$\tau_{\mu,0}(z) \approx \frac{5}{5-\gamma} \left( \frac{z}{z_{\text{dc}}} \right)^{5/2} \sqrt{1 + (z/z_{\text{br}})^{-\gamma}} \times \left[ 1 - \frac{\gamma(z/z_{\text{br}})^\gamma}{5+\gamma} {}_2F_1 \left( 1, 1 + \frac{5}{2\gamma}, \frac{3}{2} + \frac{5}{2\gamma}, -(z/z_{\text{br}})^\gamma \right) \right], \quad (48)$$

where we introduced  $\gamma = 2 \times 0.672 + 1 = 2.344$ ,  $z_{\text{dc}} = 1.98 \times 10^6$ ,  $z_{\text{br}} = [x_c^{\text{BR},0}(z=0)/x_c^{\text{DC},0}(z=0)]^{2/\gamma} \approx 3.81 \times 10^5$  and  ${}_2F_1(a, b, c, x)$  is the hypergeometric function. Here, we neglected DC temperature and frequency corrections. All the coefficients just follow from the expressions for the critical frequency.

In Fig. 2, we illustrate the effect on the visibility function, comparing with  $\mathcal{F}_{\text{DC}}$ . Close to the thermalization redshift  $z_{\text{dc}} \approx 2 \times 10^6$ , the visibility of spectral distortions is reduced by  $\approx 5\%$  in comparison to the DC only approximation. This is in good agreement with the recent findings of KS12. We only show the correction to the visibility function up to  $\sim 3$  times the thermalization redshift, since there  $\mathcal{F} \approx 1.7 \times 10^{-7}$ , which for  $\Delta\rho/\rho \approx 1\%$  could still lead to a detectable  $\mu$ -distortion for PRISM. We see that at high redshifts the approximation, Eq. (48), starts to break down. This is because especially around  $z \approx 4 \times 10^5$  the true critical frequency deviates slightly from the approximation  $x_c \approx [(x_c^{\text{DC}})^2 + (x_c^{\text{BR}})^2]^{1/2}$ , which leads to degradation of the total integral for very large  $z_h$ . Full numerical determination of  $x_c$  is trivial and Eq. (48) will thus only be used for estimates.

We also already explained in Sect. 3.3.4 that photon transport ceases at  $z \lesssim 2 \times 10^5$ . Photons produced by BR below this redshift are stuck at low frequencies and no longer help thermalizing the full spectrum. The total DC thermalization optical depth between  $z \approx 2 \times 10^5$  and  $z = 0$  is only  $\tau_\mu \approx 0.003$ , but when including BR, from Eq. (47) we find  $\tau_\mu \approx 0.01$ . This leads to an  $\approx 1\%$  overestimation of the thermalization efficiency and hence a similar error in the distortion visibility function. Corrections to the thermalization optical depth should thus only be computed at  $2 \times 10^5 \lesssim z$ , a modification that is straightforward to include but was omitted before.

### 5.2.2 Compton equilibrium temperature

To check the consistency of the solution, we briefly turn to the condition  $T_e = T_{\text{RJ}}$ . Using  $\mu^{(0)} = \mu_\infty(\tau) e^{-x_c(\tau)/x}$ , we can readily compute the Compton equilibrium temperature in the distorted radiation field. From Eq. (17a), neglecting terms  $\mathcal{O}(\theta_\gamma^2)$  it is given by

$$T_e^{\text{eq.}(0)} \approx T_{\text{RJ}}^{(0)} - \frac{T_\gamma}{4\mathcal{G}_3^{\text{PI}}} \int_0^\infty x^2 Y_{\text{SZ}}(x) \mu^{(0)}(\tau, x) dx. \quad (49)$$

For frequency independent  $\mu^{(0)}(\tau, x)$ , this immediately implies  $T_e^{\text{eq.}(0)} = T_{\text{RJ}}^{(0)}$  because  $\int x^2 Y_{\text{SZ}}(x) dx$  vanishes; however, when inserting the lowest order solution, the integral no longer vanishes and  $T_e^{\text{eq.}(0)}$  deviates from  $T_{\text{RJ}}^{(0)}$  by  $\mathcal{O}(x_c/10)$ . Since after the energy release stopped, at lowest perturbation order one should find  $T_e^{(0)} = T_e^{\text{eq.}(0)} \equiv T_{\text{RJ}}^{(0)}$ , this means that the solution is slightly inconsistent, but the discrepancy is indeed of higher order in  $x_c$ .

### 5.3 First-order corrections

At this point, we have not included any additional physics but simply kept all lowest order terms, consistent with  $\mathcal{O}(x_c)$  in the evolution equation for  $\mu$  and definition of  $\tau_\mu$ . The derived corrections were already present in previous numerical calculations and even without improvements of the BR and DC Gaunt factors introduced in COSMOTHERM could have been obtained with little effort. The next step is to obtain the corrections to  $\hat{I}_\mu$  and then evaluate the changes to the optical depth term,  $\Delta\tau_\mu(z)$ , given by Eq. (34b), adding terms of  $\mathcal{O}(x_c^2)$  to the integrand.

#### 5.3.1 Correction to $\hat{\mu}(\tau, x)$ at order $\mathcal{O}(x_c)$

Since the photon production integral, Eq. (32), is of  $\mathcal{O}(x_c)$ , we have to improve the solution of  $\hat{\mu}(\tau, x)$  to order  $\mathcal{O}(x_c)$ . From the lowest order solution, we know already that the time derivatives  $\partial_\tau \mu$  and  $\partial_\tau(T_e/T_\gamma)$  are both of order  $\mathcal{O}(x_c)$ . Similarly,  $\Lambda(x)/\theta_\gamma \approx x_c^2$ . We furthermore already understand that the main frequency dependence of the solution is found around  $x \approx x_c \ll 1$ . Defining  $\sigma = x_c y = x_c \int \theta_\gamma d\tau$ , and re-scaling the frequency as  $x = x_c \xi$ , we have the evolution equation:

$$x_c \frac{\partial \mu}{\partial \sigma} - x_c^2 \xi \partial_\sigma \frac{T_e}{T_\gamma} \approx \xi^2 \mu'' + 2g_1(x_c \xi) \xi \mu' - \frac{\lambda(x_c \xi) (1 - e^{-x_c \xi})}{\xi^2} \frac{1}{x_c \xi} \mu,$$

where primes now denote derivatives with respect to  $\xi$  and we defined  $\lambda = \Lambda(x)/[\theta_\gamma x_c^2] = 1 + \Delta\lambda$  [for double Compton  $\Delta\lambda = 0$  when neglecting frequency corrections to the Gaunt factor]. Scaling the equations in this way shows that both the temperature term on the l.h.s. of this equation and the higher order Compton correction [ $\propto 2(g_1 - 1) \approx -x^2/6[1 - x^2/60]$  for  $x \ll 1$ ] enter the problem at higher order in perturbation theory, so that we neglect them for now<sup>10</sup>. For the correction to the emission term, we have  $\lambda(x_c \xi)(1 - e^{-x_c \xi})/(x_c \xi) \approx 1 - x_c \xi/2 + \mathcal{O}(x_c^2)$  in the DC-era. During the BR-era, matters are complicated by the logarithmic dependence of the Gaunt factor on  $x$ . This means that in this case deviations from  $\Lambda(x) = \text{const}$  enter at order  $\approx x_c \ln(x_c)$ . For simplicity we define

$$\alpha_{\text{em}}(x) = \Lambda(x)(1 - e^{-x})/(\theta_\gamma x_c^2 x) - 1, \quad (50)$$

<sup>10</sup> As we will see in Sect. 5.5, this is too naive and the Compton terms give  $\mathcal{O}(x_c)$  contributions at intermediate frequencies.

and include all frequency correction terms simultaneously at the first perturbation order. For the time derivative of  $\mu^{(0)}(x, \tau)$ , we have

$$\frac{\partial \mu^{(0)}(x_c(\tau)\xi, \tau)}{\partial y} \approx \mu^{(0)} \left[ \frac{\partial \ln \mu_\infty^{(0)}(\tau)}{\partial y} - \frac{x_c}{x} \frac{\partial \ln x_c(\tau)}{\partial y} \right]. \quad (51)$$

The first term was recently considered by KS12, although there it was treated as  $\mu^{(0)} \partial_y \ln \mu_\infty^{(0)}(\tau) \rightarrow \mu \partial_y \ln \mu_\infty^{(0)}(\tau)$ , giving a modified Bessel function solution for  $\mu$ . The second term leads to a small correction at high redshifts ( $z \gtrsim 10^6$ ), but in our approach it does become significant later.

Put together, this then determines the evolution equation for the first correction  $\mu^{(1)}(\tau, x) = \mu_\infty^{(0)}(\tau) \hat{\mu}^{(1)}(\tau, x) + \mu_\infty^{(1)}(\tau) \hat{\mu}^{(0)}(\tau, x)$ :

$$S^{(1)}(x) = \left[ \frac{\partial \ln \mu^{(0)}}{\partial y} + \alpha_{\text{em}}(x) \right] \hat{\mu}^{(0)} \approx \partial_x x^2 \partial_x \hat{\mu}^{(1)} - \frac{x_c^2}{x^2} \hat{\mu}^{(1)}. \quad (52)$$

The general solution of this equation reads

$$\hat{\mu}^{(1)} = C_1 e^{-x_c/x} + C_2 e^{x_c/x} + \int_0^x \sinh\left(\frac{x_c}{x'} - \frac{x_c}{x}\right) S^{(1)}(x') \frac{dx'}{x_c}, \quad (53)$$

where  $C_1$  and  $C_2$  are fixed by the boundary conditions. Inserting the expression for  $S^{(1)}(x)$ , then gives the first-order correction to the chemical potential as

$$\hat{\mu}^{(1)} \approx C_1 e^{-x_c/x} + \ln(x/x_c) e^{-x_c/x} \partial_y \ln \mu_\infty^{(0)} + D_\mu(x_c/x) \left[ \partial_y \ln \mu_\infty^{(0)} + \frac{1}{2} \partial_y \ln x_c \right] + D_{\text{em}}(x_c, x_c/x) x_c. \quad (54)$$

The new integration constant can be fixed by requiring  $\kappa_\rho = \kappa_\rho^c$  (see Sect. 5.3.2). We directly absorbed any contribution from the source term leading to an asymptotic behavior  $\propto e^{-x_c/x}$  at  $x \ll 1$  into this integration constant. The functions  $D_\mu$  and  $D_{\text{em}}$  are defined as

$$D_\mu(\zeta) = e^{-\zeta} \left[ \ln(2\zeta) + \gamma_E - e^{2\zeta} \text{Ei}(-2\zeta) \right] \\ D_{\text{em}}(x_c, \zeta) = F_{\text{em}}(x_c, 0) e^{-\zeta} - F_{\text{em}}(x_c, \zeta) \quad (55)$$

$$F_{\text{em}}(x_c, \zeta) = \frac{e^{-\zeta}}{2x_c} \int_0^x \zeta' \alpha_{\text{em}}(x') \left[ e^{2(\zeta-\zeta')} - 1 \right] \frac{dx'}{x'},$$

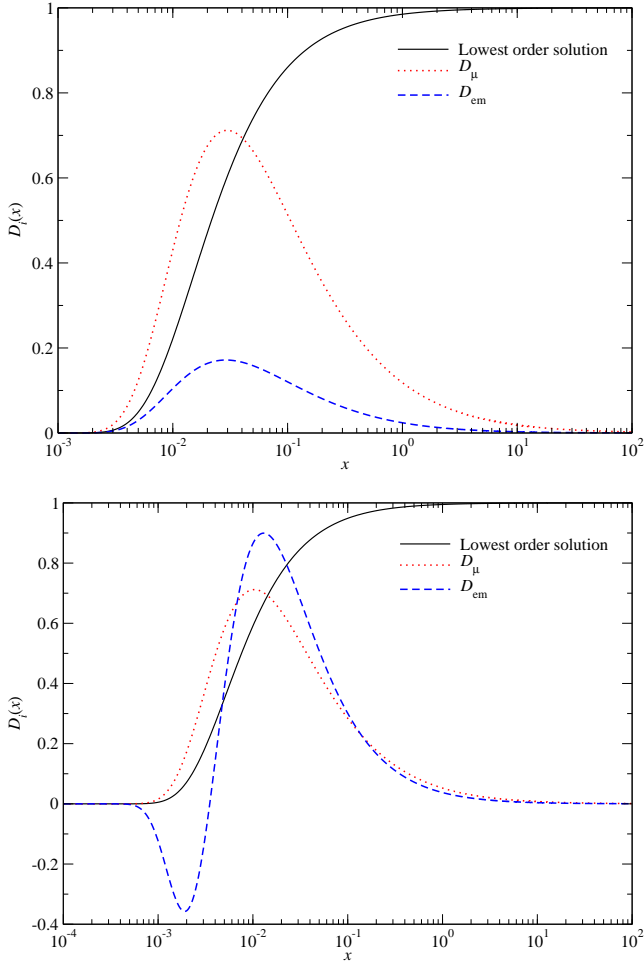
where in the expression for  $F_{\text{em}}$  we use  $x = x_c/\zeta$  and  $x' = x_c/\zeta'$ . The integral for  $F_{\text{em}}$  can be carried out numerically very efficiently.

We scaled the two correction function so that they are comparable in amplitude. Their shapes are illustrated in Fig. 3. The main correction appears around  $x \approx 2x_c$ , where also most of the photon emission comes from. This also implies that, while these two correction functions change little for the total energetics, they directly affect the thermalization efficiency. In particular at  $z \lesssim 10^6$ , when BR starts dominating, the correction related to  $D_{\text{em}}$  becomes significant, giving rise to a nontrivial dependence on frequency. This is partially due to the small mismatch of  $x_c$  with the real position of the emission maximum caused by the frequency dependence of the DC and BR Gaunt factors, but also the  $\approx (1 - e^{-x})/x$  modulation of the emission term in Eq. (20).

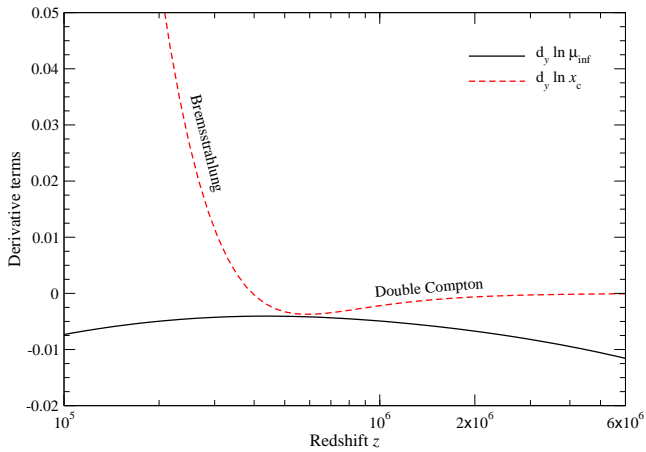
To obtain the final solution, we can use  $\partial_y \ln \mu_\infty^{(0)} \approx -0.7769 x_c$ , which follows from Eq. (33). The derivative of the critical frequency with respect to  $y$  is approximately given by

$$\frac{\partial_\tau x_c}{2x_c \theta_\gamma} \approx \frac{2.01 \times 10^{-2}}{1+z} \frac{\partial_\tau z}{x_c^2} \left( 1 - 2.75 \times 10^{-2} \left[ \frac{1+z}{2 \times 10^6} \right]^{-2.344} \right), \quad (56)$$

where we used the expression given in Sect. 3.3.3, but neglected relativistic corrections. During the radiation-dominated era ( $z \gtrsim 3300$ ), we have  $\partial_\tau z = -H(1+z)/\sigma_T N_e c \approx -4.80$ . The second term in parentheses arises because of BR, which can be neglected at high redshifts. At  $z = 2 \times 10^6$ , we find  $\frac{1}{2} \partial_y \ln x_c \approx -6 \times 10^{-4} \approx -0.07 x_c$ ; however, at  $z \approx 2 \times 10^5$  we have  $\frac{1}{2} \partial_y \ln x_c \approx 6 \times 10^{-2} \approx 9.4 x_c$  (see



**Figure 3.** Frequency dependence of the correction functions  $D_\mu$  and  $D_{\text{em}}$  in comparison to  $\hat{\mu}^{(0)} = e^{-x_c/x}$ . The upper panel illustrates the case for  $x_c \approx 0.015$  ( $z \approx 6 \times 10^6$ ).  $D_\mu$  and  $D_{\text{em}}$  have a very similar shape at low frequencies, although rescaling them to coincide around the maximum reveals small differences at  $x \gtrsim 1$ . In the lower panel, we show the functions for  $x_c \approx 5.3 \times 10^{-3}$  ( $z \approx 5 \times 10^5$ ), for which, due to the logarithmic dependence of the BR Gaunt factor on frequency,  $D_\mu$  and  $D_{\text{em}}$  differ significantly.



**Figure 4.** Contributions to the time derivative of the lowest order solution to  $\mu_\infty$ . At late times, the low-frequency spectrum changes mostly because of emission and absorption processes ( $\partial_y \ln x_c$  becomes significant), so that scattering-driven quasi-stationarity no longer is a good assumption.

Fig. 4 for more details), which shows that at low frequencies the quasi-stationary approximation starts to break down in this regime. This behavior is expected, since the efficiency of Compton scattering decreases with redshift, so that full kinetic equilibrium between photons and electrons can no longer be established. We thus do not expect to obtain very accurate analytic approximations at  $z \lesssim 2 \times 10^5$  using our perturbative approach.

At  $z \lesssim 2 \times 10^5$ , photons produced by BR are furthermore stuck at low frequencies and no longer up-scatter strongly (see Sect. 3.3.4). In this regime, the shape of the low-frequency spectrum is fully determined by photon emission and absorption, and the effect of scattering can be added as a perturbation. Since here we are mainly interested in the  $\mu$ -era, we leave a more detailed discussion of this problem for another paper.

### 5.3.2 Fixing the integration constant $C_1$

To give the full first-order solution, we still need to determine the integration constant  $C_1$  in Eq. (54). It simply follows from our normalization condition  $\kappa_\rho \equiv \kappa_\rho^c$ . Since the lowest order solution is  $\hat{\mu} = e^{-x_c/x}$ , we can just set  $1 + C_1 = A(x_c)$  in Eq. (54) and directly discuss the full solution up to first order in  $x_c$ . Just using  $\hat{\mu} = A_0(x_c) e^{-x_c/x}$ , we find that

$$A_0(x_c) \approx [1 - 1.85 x_c^{0.83} / (1 + 2.44 x_c)]^{-1} \quad (57)$$

fulfills the normalization condition for  $10^{-3} \leq x_c \leq 0.05$  very well.

For  $\hat{\mu} = A(x_c) e^{-x_c/x} + \ln(x/x_c) e^{-x_c/x} \partial_y \ln \mu_\infty^{(0)}$ , we obtain the correction

$$\Delta A_{\text{in}}(x_c) \approx 1.06 A_0(x_c) (1 + 0.98 \ln x_c) \partial_y \ln \mu_\infty^{(0)} \quad (58)$$

to  $A_0(x_c)$ . Since the term associated with  $D_\mu$  can become significant at late times, we have the additional contribution

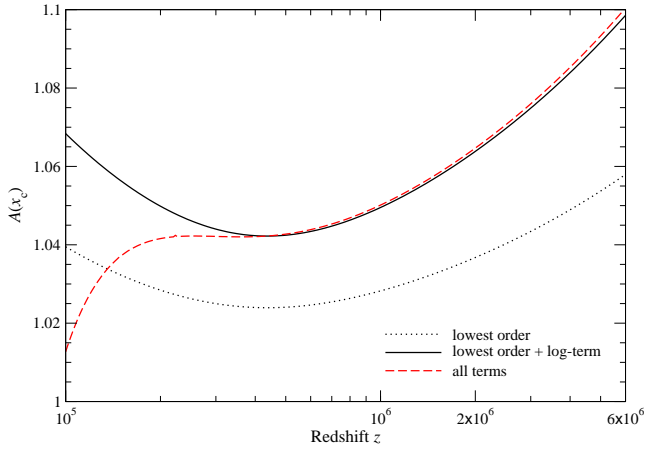
$$\Delta A_\mu(x_c) \approx 1.15 x_c A_0(x_c) [1 + 22.7(1 - 0.97 x_c^{0.029}) \ln x_c] \alpha(\tau), \quad (59)$$

with  $\alpha(\tau) = \partial_y [\ln \mu_\infty^{(0)} + (1/2) \ln x_c]$ , which works well when using  $\hat{\mu} = A(x_c) e^{-x_c/x} + \ln(x/x_c) e^{-x_c/x} \partial_y \ln \mu_\infty^{(0)} + D_\mu(x_c/x) \alpha(\tau)$ . Note that in terms of perturbations,  $\Delta A_\mu(x_c)$  formally is of higher order in  $x_c$  at  $z \gtrsim 4 \times 10^5$ ; however, at late times it becomes pretty significant, so that we generally include it.

In Fig. 5, we show the numerical results for  $A(x_c)$  in different cases. The changes in the normalization are  $\approx 1\% - 10\%$  in the shown redshift range. The  $\ln(x/x_c) e^{-x_c/x}$  term is clearly important at all times, and especially at late times the contribution from  $D_\mu(x_c/x)$  becomes large. The correction due to  $D_{\text{em}} x_c$  is not as crucial and the normalization constant for the full solution is well represented by adding all terms from Eq. (57)–(59).

The solution for  $\hat{\mu}$  is illustrated in Fig. 6 for two redshifts. The largest corrections are due to the renormalization factor,  $A(x_c)$ , and the logarithmic term in Eq. (54), which are significant at both  $x \approx x_c$  and large  $x$ . These two terms capture the main behavior of the expression given by KS12. Notice that they normalized their solution using the condition  $\mu_\infty = \mu(x = 0.5)$ . They furthermore neglected the correction caused by the  $\approx \mathcal{O}(x_c)$  emission terms leading to  $D_{\text{em}}$ , as well as the time derivative of  $x_c$ , which does become significant at the later stages (see Fig. 4). However, the difference due to this does not seem as crucial.

In Fig. 6, we show a comparison with the numerical result obtained with COSMOTHERM. Our approximation captures the shape of the solution very well. Only at high frequencies, which matters for the overall energetics of the solution but not as much for the photon production, the approximation Eq. (54) deviates noticeably



**Figure 5.** Normalization constant  $A(x_c)$  for the three cases discussed in Sect. 5.3.2. The change in the normalization when adding  $D_{\text{em}}(x)$  is very small and can generally be neglected.

from the full numerical result. This is expected since our perturbative approach is meant to work at  $x \lesssim 1$  only. We can further improve the solution by matching with the high-frequency solution obtained in Sect. 5.5.1, reaching agreement at all relevant frequencies to  $\approx 0.1\% - 1\%$  at  $z \gtrsim \text{few} \times 10^5$ .

### 5.3.3 Compton equilibrium temperature

By construction, we should find  $T_e = T_{\text{RJ}}$  at least to  $O(x_c^2 \ln x_c)$ , which implies the condition

$$0 = \int x^2 \hat{\mu}(t, x) Y_{\text{SZ}}(x) dx = \langle \hat{\mu} \rangle_{\text{SZ}}, \quad (60)$$

like in Sect. 5.2.2. For this, we only need to worry about the first two terms of Eq. (54), as the other are energetically much less important. We find

$$\langle e^{-x_c/x} \rangle_{\text{SZ}} \approx -0.127 x_c \quad (61a)$$

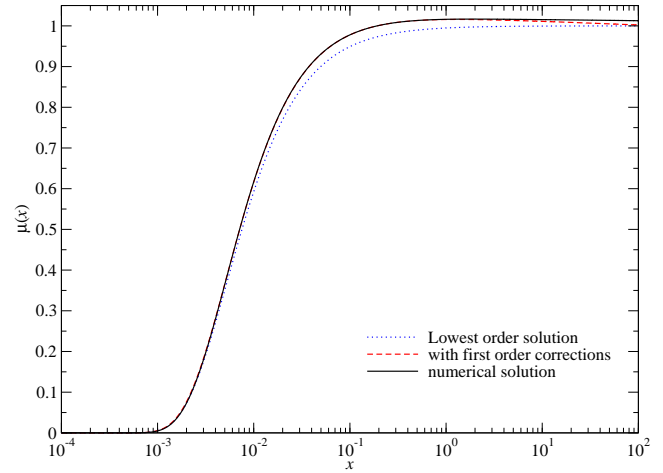
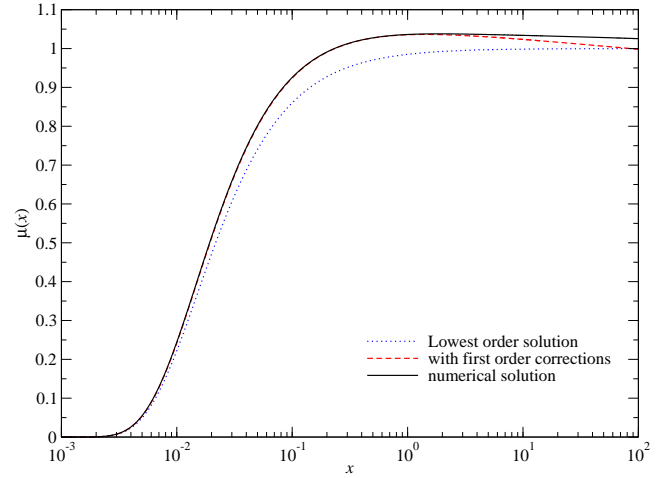
$$\langle \ln(x/x_c) e^{-x_c/x} \rangle_{\text{SZ}} \approx -0.277(1 + 0.336 x_c^{0.598}), \quad (61b)$$

which with  $\partial_y \ln \mu_\infty^{(0)} \approx -0.7769 x_c$  implies an imbalance of order  $\approx 0.1 x_c$ . This means that our solution is slightly inconsistent, but at this point we have no freedom left to ‘fix’ this discrepancy. The situation is improved a bit once we correct the high-frequency solution using the asymptotic behavior determined in Sect. 5.5. However, a small difference larger than  $O(x_c^2)$  remains. For the photon production rate, this discrepancy does not seem to matter and will be neglected below.

### 5.3.4 Change in the photon production rate

In Fig. 7, we show how  $\Sigma = \bar{J}_\mu - 1$  changes for different approximations. This is the relevant quantity for the optical depth correction,  $\Delta\tau_\mu$ , defined by Eq. (34b). Just using  $\hat{\mu} = e^{-x_c/x}$  already gives a significant correction; however, for consistency the renormalization  $A(x_c) \neq 1$  has to be included. We note also that during the BR-era the logarithmic dependence of the Gaunt factor is very important, and assuming  $\Lambda(x) = \text{const}$  gives incorrect results at  $z \lesssim 10^6$ .

If we neglect the contributions from  $D_\mu$  and  $D_{\text{em}}$  in the approximation Eq. (54), we obtain the dashed red line, where the difference to the previous case is only caused by the logarithmic term,  $\propto \ln(x/x_c) e^{-x_c/x}$ . For the violet double-dash dotted curve, we used

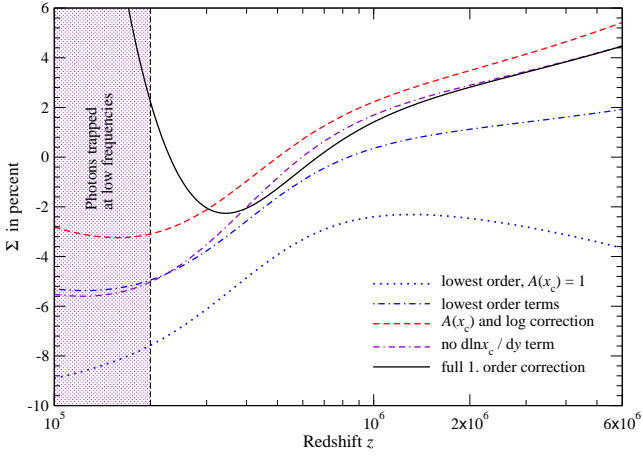


**Figure 6.** Comparison of the lowest order solution  $\hat{\mu}^{(0)} = e^{-x_c/x}$  and Eq. (54) with the numerical result obtained with COSMOTHERM. We scaled the numerical solution for  $\mu(t, x)$  by  $\mu_\infty(t) = [3\Delta\rho_\gamma(t)/\rho_\gamma - 4\Delta N_\gamma(t)/N_\gamma]/\kappa_\rho^c$  in agreement with our normalization condition for  $\hat{\mu}$ . The upper panel shows the solution for  $x_c = 0.015$  ( $z \approx 6 \times 10^6$ ), while in the lower panel we have  $x_c = 5.3 \times 10^{-3}$  ( $z \approx 4.8 \times 10^5$ ). The difference at high frequencies can be captured by matching with the high-frequency limit of the photon Boltzmann equation, giving extremely good agreement with the numerical result over the full range of frequencies (Sect. 5.5.1 and Fig. 11).

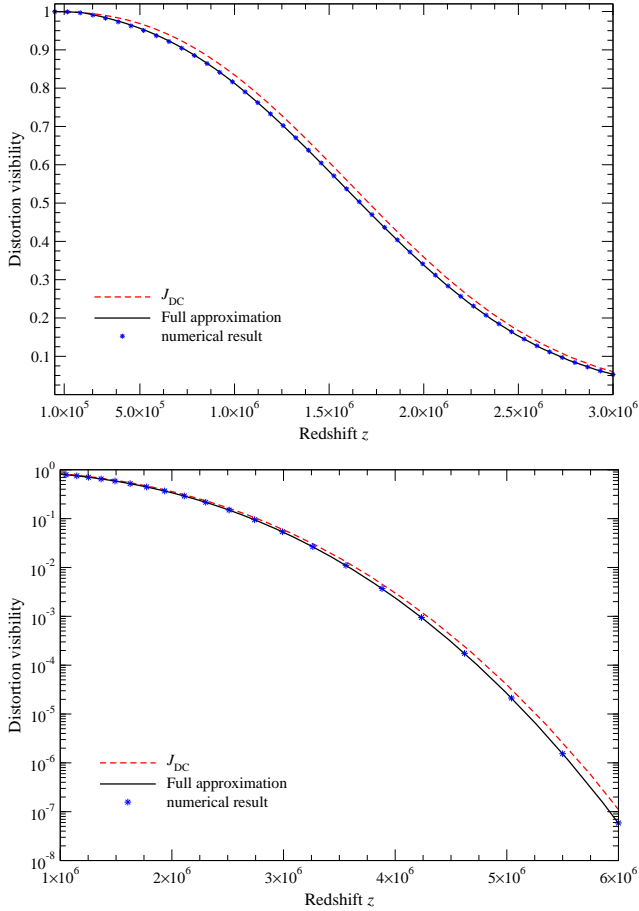
the full first-order expression, Eq. (54), for  $\hat{\mu}$ , but neglected the contribution from  $\partial_y \ln x_c$ , which becomes large at low redshifts. At  $z \lesssim 5 \times 10^5$ , neglecting the extra emission terms brings  $\Sigma$  again closer to the lowest order case, while the change is much smaller at earlier times. When also adding the contribution from  $\partial_y \ln x_c$  we find a large change of  $\Sigma$  during the BR-era. This difference is not as important eventually, since the total optical depth contribution becomes rather small at late times. Also, at  $z \lesssim \text{few} \times 10^5$ , the approximation is not expected to be as accurate (we find that it works well until  $z \approx 3 \times 10^5$ ), but the overall effect on the visibility function remains small.

### 5.3.5 Effect on the distortion visibility function

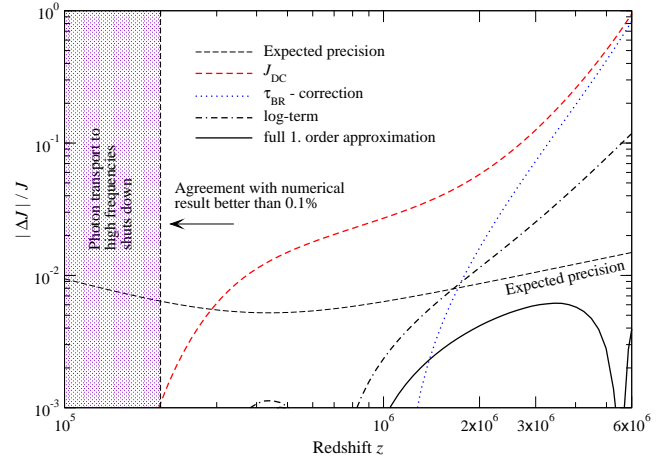
With Fig. 7, we can now compute the visibility function for different approximations. In Fig. 8, we show the comparison of  $\mathcal{J}_{\text{DC}}$  and our approximation with the numerical result obtained with COSMOTHERM. The agreement of our approximation is extremely good,



**Figure 7.** Changes of  $\Sigma = \bar{J}_{\hat{\mu}} - 1$  for different approximations of  $\hat{\mu}$  discussed in Sect. 5.3.2. For the lowest order solution, we used  $\hat{\mu} = A(x_c) e^{-x_c/x}$  with  $A(x_c) = 1$  and  $A(x_c) = A_0(x_c)$  defined by Eq. (57). For the dashed red line, we used  $\hat{\mu} = A(x_c) e^{-x_c/x} + \ln(x/x_c) e^{-x_c/x} \partial_y \ln \mu_{\infty}^{(0)}$ , while for the solid black line we used the full first-order expression, Eq. (54), each with their corresponding normalization constants,  $A(x_c)$ . The double-dash dotted curve also gives the result using Eq. (54), but when neglecting the contribution from  $\partial_y \ln x_c$ , which becomes large at low redshifts. The shaded region indicates where the high-frequency photon number freezes out.



**Figure 8.** Distortion visibility function at different redshifts. The red dashed curve shows  $\mathcal{J}_{DC} = e^{-(z/z_{dc})^{5/2}}$  with  $z_{dc} = 1.98 \times 10^6$ . The solid black line gives our approximation based on Eq. (54), with all terms included. The numerical result was obtained using COSMOTHERM.



**Figure 9.** Corrections to the distortion visibility function at different redshifts. For all curves, the numerical result obtained using COSMOTHERM was used as reference. The dashed red line shows  $\mathcal{J}_{DC} = e^{-(z/z_{dc})^{5/2}}$  with  $z_{dc} = 1.98 \times 10^6$ . When only including the BR correction to the optical depth (Sect. 5.2), we obtain the dotted blue line. Only adding the  $\ln(x/x_c) e^{-x_c/x}$  term, we improve the agreement at early times. The solid black line gives our approximation based on Eq. (54), with all terms included, showing precision below the level expected in terms of perturbation order  $\approx x_c$ .

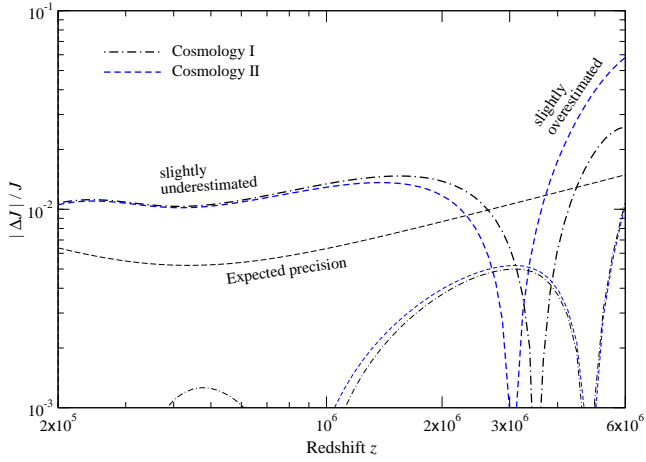
without any matching with the numerical solution being carried out. Also, evaluation of the simple integrals over the emission term and the optical depth integrals take no more than a few seconds as opposed to a couple of hours for the full numerical calculation, giving a huge improvement of the performance. We note that the full effect of the distortion visibility function and the full shape of the distortion are also captured by the efficient Green's function method introduced earlier (Chluba 2013b).

In Fig. 9, we illustrate more clearly which terms actually matter most. The simplest approximation,  $\mathcal{J}_{DC} = e^{-(z/z_{dc})^{5/2}}$ , shows excellent agreement with the numerical result until  $z \approx 2 \times 10^5$ , when low-frequency photons produced by BR start reaching the high-frequency domain. In particular at  $z \gtrsim 10^6$ , the visibility is significantly lower than estimated with  $\mathcal{J}_{DC}$ . Adding the BR correction to the optical depth, significantly improves the solution below  $z \lesssim 10^6$  even to the sub-0.1% level. Clearly, by calculating the full optical depth integral and realizing that at  $z \approx 2 \times 10^5$  photon transport to high frequencies shuts down, one can improve the approximation significantly. All the physics of this correction were already included by the early treatments (Sunyaev & Zeldovich 1970; Burigana et al. 1991; Hu & Silk 1993), but since at  $z \lesssim 10^6$ , also  $\mathcal{J}_{DC}$  already has  $\lesssim 3\%$  precision, it was previously not of much interest and only added recently by KS12 in preparation for high-precision spectral distortion measurements.

Once we also add the  $\ln(x/x_c) e^{-x_c/x}$  term to the expression for  $\hat{\mu}$ , we further improve the agreement at  $z \gtrsim 10^6$ . The slight disagreement introduced at lower redshifts is cancelled mostly when all terms are added to the approximation. This shows, the importance of both  $D_{\mu}$  and  $D_{em}$  at  $z \gtrsim 10^6$ ; in our approach these terms need to be included to obtain an approximations below the expected level of precision which is comparable to  $\approx x_c$ .

#### 5.4 Comparison with Khatri & Sunyaev 2012

In Fig. 10, we compare our numerical results directly with the approximations for the distortion visibility function given by KS12.



**Figure 10.** Comparison of the approximation given by KS12 with our numerical result from COSMOTHERM. We compare for the cosmology used in KS12 (blue dashed) and the one used here (black dash-dotted). Their simple expression (heavy lines) works very well overall. Our approximation (thin lines) represents our numerical result below the expected precision  $\approx x_c$  at all redshifts and giving  $\leq 0.1\%$  precision at  $z \lesssim 10^6$ . In Sect. 5.4.1, we briefly discuss the possible explanations for difference with KS12.

We included DC relativistic corrections, because KS12 applied the expressions from Chluba & Sunyaev (2012) in their numerical computations, which included these aspects (see Sect. 6 for more discussion). Overall, their approximation captures the full numerical result very well. We give the comparison for two slightly different cosmologies, showing that their expression represents our result for the distortion visibility to a few percent precision. Our approximation performs a little better, representing our numerical result below the expected precision  $\approx x_c$  at all redshifts and giving much higher precision at  $z \lesssim 10^6$ . Also, the cosmology dependence is not as pronounced. This is reassuring, demonstrating that our perturbative approach works very well; the achieved level of precision is, however, generally very futuristic, although the computational cost is also very small.

At low redshift ( $z \lesssim 10^6$ ), the approximation of KS12 slightly underestimates the true distortion visibility function, an effect that is also visible in their Fig. 7. This is because they did not take into account that photon transport from low to high frequencies stops below  $z \approx 2 \times 10^5$  (see Sect. 3.3.4). Also, at those epochs, it becomes difficult even numerically to define the amplitude of  $\mu_\infty$  without using energetic arguments, because the shape of the distortions starts departing from a pure  $\mu$ -distortion. Our approach avoids this complication (see discussion below).

#### 5.4.1 Possible causes for the small differences with KS12

Although pretty small, the differences between our numerical result and the approximations of KS12 are larger than the stated precision of their formulae. In particular, at  $z \gtrsim 10^6$ , they obtain sub-percent agreement with their numerical solution. What could be the possible causes for the differences?

One possibility is simply the numerical treatment. This is however unlikely, since both KS12 and our approximate solutions provide an approach that is independent of the more delicate partial differential equation solving, finding excellent agreement internally. The next possibility is the included physics. Again, this seems unlikely, since they also base their physical setup on

Chluba & Sunyaev (2012) and what went into COSMOTHERM. The only small problem could be related to the fact that KS12 did not explicitly separate the physics of DC relativistic corrections, possibly explaining some part of the cosmology dependence we find.

The most plausible cause of the differences is the normalization condition. In the derivation of Eq. (9b), it was explicitly required that  $\partial_\tau \kappa_\rho = 0$ . In our formulation, this is directly achieved using the normalization condition  $\kappa_\rho(t) = \kappa_\rho^c \approx 2.1419$  to fix the free integration constant in Eq. (54). In contrast, KS12 just normalized their solution at one fixed frequency. This generally gives  $\partial_\tau \kappa_\rho \neq 0$ , so that the equivalent of Eq. (9b) reads

$$\frac{d(\hat{\kappa}_\rho \mu_\infty)}{dt} \approx \frac{3}{\kappa_\rho^c} \frac{d \ln a^4 \rho_\gamma}{dt} - \frac{4}{\kappa_\rho^c} \frac{d \ln a^3 N_\gamma}{dt}. \quad (62)$$

This adds another small time-dependent term to the problem, which can be thought of as an equivalent of the effective heat capacity for the distorted photon field. This term was not discussed by KS12, while we absorbed it in the definition of  $\mu_\infty^*$ . Physically, this probably implies that the distortion visibility function of KS12 does not exactly represent the fraction of energy that is stored by the distortion at a given moment. However, since the difference is small, we address this question in some future work.

### 5.5 High-frequency matching and corrections due to other neglected terms

While in terms of perturbation theory, we have already included all contributions  $\mathcal{O}(x_c)$  into the analysis of Sect. 5.3, it is interesting to understand the role of higher order corrections in  $x$ . These are expected to become relevant at intermediate frequencies  $x \approx 1$ , reaching similar amplitudes as the other terms. We start by using the solution obtained in the high-frequency limit and match it with the low-frequency solution discussed in Sect. 5.3. We then proceed by adding the temperature drift term and higher order Compton corrections in frequency.

#### 5.5.1 High-frequency solution matching

Our numerical results show that the high-frequency behavior is not well represented by extrapolating the low-frequency solution Eq. (54). Instead, we should separately consider the high-frequency limit of the photon Boltzmann equation and then match the solutions at some matching point  $x_m$ .

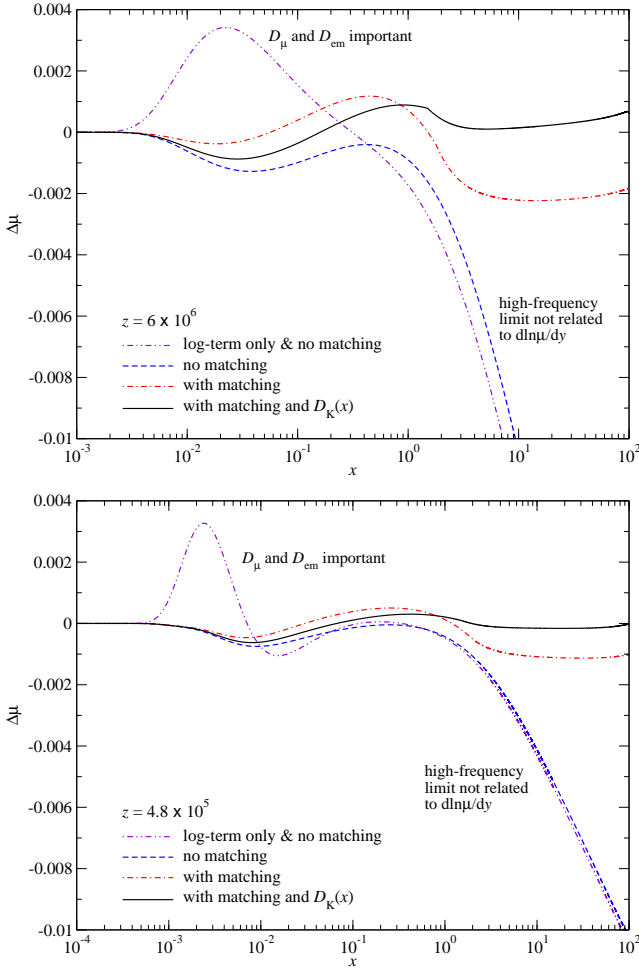
At high frequencies, emission and absorption terms can be neglected and we only need to worry about the effect of Compton scattering. This gives the evolution equation

$$\partial_y \mu - x \partial_y (T_e/T_\gamma) \approx x^2 \mu'' - x^2 \mu'. \quad (63)$$

Assuming that the time-derivative terms can be treated as perturbations, we find the lowest order solution  $\mu_{\text{high}}^{(0)}(x) \approx C_{\text{high}}$ , which is consistent with the low-frequency solution  $\mu = C_1 e^{-x_c/x}$ . In the next iteration, we find

$$\mu_{\text{high}}(x) \approx C_{\text{high}} + \ln(x) \partial_y (T_e^{(0)}/T_\gamma), \quad (64)$$

which shows that energetically the time derivative of the electron temperature plays the most important role at high frequencies. Physically, this makes a lot of sense as well, since the energy exchange is dominated by the high-energy spectrum, which is mainly driven by the Compton process and hence directly related to the electron temperature and its rate of change. From Eq. (9a),



**Figure 11.** Difference of the analytic approximation for  $\hat{\mu}$  with respect to the numerical solution obtained from COSMOTHERM at  $z = 6 \times 10^6$  and  $z = 4.8 \times 10^5$ . The approximation, Eq. (54), shown as dashed blue line captures the behavior well at low frequencies, while at high frequencies it deviates from the numerical solution at the level of a few percent. Neglecting the contributions from  $D_\mu$  and  $D_{em}$  (violet dash-dot-dotted line) degrades the solution at low frequencies. Matching with the high-frequency solution, Eq. (64), gives sub-percent agreement. Also adding the Compton scattering correction  $D_K$  to the low-frequency solution and using Eq. (70) for the high-frequency part further improves the agreement.

we find  $\partial_y(T_e^{(0)}/T_\gamma) \approx (M_3/4)\partial_y\mu_\infty^{(0)} \approx -0.2157 x_c\mu_\infty^{(0)}$ , which implies that the contribution of the log-term is reduced roughly 3 times with respect to the low-frequency solution. Looking at Fig. 6 suggests that this goes into the right direction. Note, however, that according to Eq. (9a) also  $(M_2M_3/\kappa_p^c)d\ln(M_3/M_2)/dy \approx -0.25x_c(1 + 0.88\ln x_c)(1 + 1.81x_c)d\ln x_c/dy$  contributes to the derivative  $\partial_y(T_e^{(0)}/T_\gamma)$  at late times.

As the next step, we should continuously match the two limiting solutions at some frequency  $x_m$ . We know that the terms  $D_\mu$  and  $D_{em}$  in Eq. (54) are rather small at  $x \geq 1$ , so that for the matching condition we can use

$$A(x_c) + \ln(x_m/x_c)\partial_y \ln \mu_\infty^{(0)} \approx C_{\text{high}} + \ln(x_m/x_c) \frac{\partial_y(T_e^{(0)}/T_\gamma)}{\mu_\infty^{(0)}} \quad (65)$$

Note that we scaled the whole solution relative to  $\mu_\infty^{(0)}$ . To determine the best value for  $x_m$ , we require smooth derivatives of the solution.

This then implies  $C_{\text{high}} \approx A(x_c) - 0.56 x_c \ln(x_m/x_c)$ . The matching point is always close to  $x_m \approx 1.8$ .

We find that this procedure improves the agreement of the numerical and analytical solutions significantly, in particular capturing the high-frequency scaling (cf. Fig. 11). However, since we leave the low-frequency solution practically unchanged, for the computation of the visibility function this modification can be omitted. In addition, a small correction to the normalization arises but it is  $\approx O(x_c^2)$ , and can also be neglected.

In terms of the photon spectrum, Eq. (64) implies that the high-frequency solution for the photon occupation number is

$$n(t, x_e) \approx \alpha(t)x_e^\gamma(t)e^{-x_e}, \quad (66)$$

with power-law coefficient  $\gamma = -\partial_y(T_e^{(0)}/T_\gamma)$  and  $x_e = h\nu/kT_e$ . This shows that due to the lack of high-frequency photons, the spectrum only slowly reaches a pure Wien-spectrum at the electron temperature,  $n(t, x_e) \approx e^{-x_e}$ . If the electron temperature changes, the shape of the spectrum is determined by the transport of photons between low and high frequencies. In particular, this indicates that the chemical potential generally remains non-zero at high frequencies. This limiting behavior is not captured by the solution given by KS12.

### 5.5.2 Temperature drift term

Let us consider the term  $-x\partial_y(T_e/T_\gamma)$  on the left-hand side of Eq. (43). As before we shall treat it as a source term in Eq. (53). Carrying out the integrals and absorbing any contribution  $\propto e^{x_c/x}$  (these can be absorbed by the boundary condition at  $x \rightarrow 0$ ) and  $\propto e^{-x_c/x}$  at  $x \gg 1$ , we find:

$$\hat{\mu}_T(\tau, x) = [x_c D_T(x_c/x) - x] \frac{\partial_y(T_e^{(0)}/T_\gamma)}{2\mu_\infty^{(0)}} \quad (67)$$

$$D_T(\zeta) = \frac{1}{2} \left[ e^{-\zeta} \text{Ei}(\zeta) - e^\zeta \text{Ei}(-\zeta) \right].$$

The frequency dependence of  $D_T$  is illustrated in Fig. 12. While the corresponding correction is  $O(x_c^2)$ , we find that  $D_T$  peaks slightly below  $\approx 2x_c$ , with a long tail towards lower frequencies, making its fractional contribution rather significant in comparison to the corrections discussed in the previous section (compare Fig. 3). Still, we neglect this second-order correction, as we expect other terms to contribute at similar order.

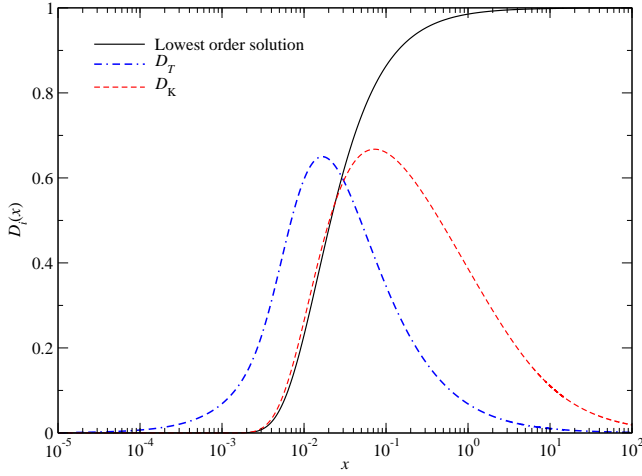
The second correction is physically more interesting, exhibiting a  $\hat{\mu} \propto x$  scaling. This term has to be interpreted as a shift in the electron temperature  $\Delta T_e^{(1)}/T_\gamma \propto \frac{1}{2}\partial_y(T_e^{(0)}/T_\gamma)$ . Since for  $\hat{\mu} \propto x$ , the photon emission integral diverges (see Sect. 3.3.5), this term eventually does not appear as contribution to the distortion and is absorbed as small correction to  $T_e = T_{\text{RJ}}$ , again regularizing the expression. Overall, the corrections due to the temperature drift term should be neglected at  $O(x_c)$  and all frequencies.

### 5.5.3 Compton scattering corrections

Earlier we argued that correction caused by  $g_1(x)$  is of second-order in  $x_c$ . While this is true at very low frequencies, it turns out to be incorrect at  $x_c \leq x \leq 1$ . We can again include the effect by simply adding the associated corrections as source term to the lowest order solution. For the Compton scattering corrections, it reads

$$S_K^{(1)}(x, \mu^{(0)}) = 2[1 - g_1(x)]x\partial_x\mu^{(0)} = \mu^{(0)} \frac{x_c}{x} \left[ x \frac{1 + e^{-x}}{1 - e^{-x}} - 2 \right].$$





**Figure 12.** Comparison of  $\hat{\mu} = e^{-x_c/x}$  with  $D_T(x_c/x)$  and  $D_K(x_c, x_c/x)$  for  $x_c = 0.015$  ( $z \approx 6 \times 10^6$ ).

Inserting this into the solution Eq. (53), we find

$$\begin{aligned} \hat{\mu}_K(\tau, x) &= x_c \ln(x/x_c) e^{-x_c/x} + 6 D_K(x_c, x_c/x) x_c \\ D_K(x_c, \zeta) &= \frac{1}{6} [F_K(x_c, 0) e^{-\zeta} - F_K(x_c, \zeta)] \\ F_K(x_c, \zeta) &= \ln(x) e^{-\zeta} + \frac{e^{-\zeta}}{x_c} \int_0^x \left[ \frac{x'}{2} \frac{1 + e^{-x'}}{1 - e^{-x'}} - 1 \right] [e^{2(\zeta - \zeta')} - 1] \frac{dx'}{x'}. \end{aligned} \quad (68)$$

The frequency dependence of  $F_K(x)$  is illustrated in Fig. 12. It has most of its contributions at frequencies  $x > 2x_c$ , so that in comparison with the previous correction functions it dominates in this range. In particular, the typical amplitude of the correction is *not*  $\approx O(x_c^2)$  but rather  $O(x_c)$ . At intermediate frequencies, this is the dominant correction we missed in our treatment above.

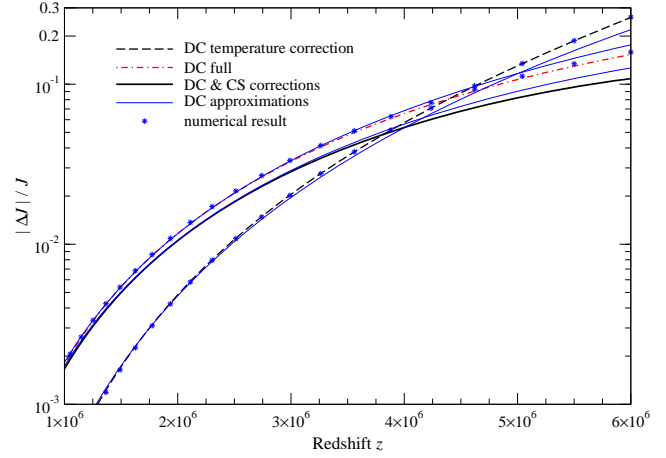
We furthermore see that the correction  $\hat{\mu}_K$  picks up a contribution  $\approx x_c \ln(x/x_c) e^{-x_c/x}$ , changing the extrapolated behavior of the total low-frequency solution to  $\hat{\mu}(\tau, x) \approx A(x_c) + 0.233x_c \ln(x/x_c)$  rather than  $\hat{\mu}(\tau, x) \approx A(x_c) - 0.777x_c \ln(x/x_c)$  at high redshifts. This behavior suggests a problem with matching the low- and high-frequency solutions smoothly, since from Eq. (64) we find a *negative* derivative for  $\hat{\mu}_{\text{high}}$  with respect to  $x$ . The problem is solved when including the next-order corrections in  $x$  for the high-frequency limit, giving the evolution equation

$$\partial_y \mu - x \partial_y (T_e/T_\gamma) \approx x^2 \mu'' + (4 - x)x\mu'. \quad (69)$$

and hence

$$\begin{aligned} \hat{\mu}_{\text{high}}(x) &\approx C_{\text{high}} + \ln(x) \frac{\partial_y (T_e^{(0)}/T_\gamma)}{\mu_\infty^{(0)}} \\ &+ \left( \partial_y \ln \mu_\infty^{(0)} - 3 \frac{\partial_y (T_e^{(0)}/T_\gamma)}{\mu_\infty^{(0)}} \right) \frac{1 + \frac{1}{x} + \frac{2}{3x}}{x}. \end{aligned} \quad (70)$$

Matching with this high-frequency approximation is no problem, with the typical matching frequency  $x_m \approx 1.5$ , from requiring smooth first derivatives. Including the Compton correction term in Eq. (68), we find agreement of the analytic solution with the numerical solution at the level  $\leq 0.1\%$  at  $0.1x_c \leq x \leq 100$  and  $z \gtrsim 3 \times 10^5$  (e.g., see Fig. 11). That is without any direct fitting to the full numerical result, underlining the advantages of our approach.



**Figure 13.** DC and CS corrections to the distortion visibility function. The full non-relativistic result is used as a reference. The thin solid blue lines give the simple approximations discussed in Sect. 6, while the other lines were obtained by modifying our perturbation expansion, Eq. (54), accordingly. For the DC corrections, we also show the full numerical result for the correction obtained with COSMOTHERM.

## 6 FIRST-ORDER RELATIVISTIC CORRECTIONS

We finish our analysis for the early  $\mu$ -distortions including lowest order relativistic corrections to the DC and Compton processes. Part of the corrections were studied analytically in Chluba (2005). The DC corrections were also already included numerically by Chluba & Sunyaev (2012) as part of COSMOTHERM, but no more detailed discussion was given.

### 6.1 DC corrections

The effect of relativistic corrections caused by DC scattering is straightforward to include in our perturbation approach. The main effect is driven by a shift in the critical frequency of a few percent (see Fig. 1) and a change in the emission integral,  $\mathcal{I}_{\hat{\mu}}$ , caused by the frequency dependence of the DC Gaunt factor. These effects can be estimated relative to the standard DC visibility function.

Neglecting the additional frequency dependence of the DC Gaunt factor,  $\Lambda_{\text{DC}}(x, \theta_\gamma) \approx \Lambda_{\text{DC}}(0, \theta_\gamma)$ , the dominant effect can be captured by re-evaluating the optical depth integral for  $\tau_{\mu,0}$  with modified  $x_c$  [obtained from Eq. (26b)], giving (cf. Chluba 2005)

$$\Delta\tau_\mu^{\text{DC}}(z, 0) \approx -5.06 \theta_\gamma (z/z_{\text{dc}})^{5/2}, \quad (71)$$

which at  $z \approx 6 \times 10^6$  implies a  $\Delta\mathcal{J}/\mathcal{J}_{\text{DC}} \approx 5.06 \theta_\gamma \approx 22\%$  visibility increase relative to  $\mathcal{J}_{\text{DC}}$ . The effect is illustrated in Fig. 13, where we compared with the full non-relativistic result instead of  $\mathcal{J}_{\text{DC}}$ . At high redshifts, the correction-to-correction is noticeable, and the simple expression, Eq. (71), slightly underestimates the effect; however, our full perturbation approximation agrees very well with the numerical result obtained with COSMOTHERM, taking only a few seconds to evaluate rather than hours.

When also accounting for the frequency dependence of the DC Gaunt factor, both Eq. (34a) and (34b) have to be re-evaluated, where we simply insert  $\hat{\mu} \approx e^{-x_c/x}$ . We find  $\Delta I_{\hat{\mu}} \approx 3.70 \times 10^{-3} + 1.46x_c$  captures the correction to  $I_{\hat{\mu}}$  pretty well, where we used  $\Lambda_{\text{DC}}(x, \theta_\gamma)/\Lambda_{\text{DC}}(0, \theta_\gamma) \approx 1 + x/2$ . Evaluating the optical depth integrals then gives

$$\Delta\tau_\mu^{\text{DC}}(z, 0) \approx [3.70 \times 10^{-3} + 1.43x_c^{\text{DC},0}] (z/z_{\text{dc}})^{5/2}. \quad (72)$$

The contribution to  $\tau_{\mu,0}(z)$ , Eq. (34a), just coming from the shift in the critical frequency by  $\Delta x_c \approx x_c/4$  [cf. Eq. (26b)] is  $\approx 0.21x_c^{\text{DC},0}$ , while re-evaluation of Eq. (34b) gave rise to the rest. This correction cancels the DC temperature correction, leading to a net change  $\Delta\mathcal{J}/\mathcal{J}_{\text{DC}} \approx -17\%$  at  $z \approx 6 \times 10^6$ . The full result is illustrated in Fig. 13. This time, our simple approximation overestimates the effect slightly, due to corrections-to-corrections that were not included. Our full perturbation approximation again agrees very well with the numerical result obtained with `CosmoTHERM`.

We also mention, that at lowest order in  $x$ , the frequency modulation of the emission term, Eq. (20), caused by the factor  $(1 - e^{-x})/x \approx 1 - x/2$  and  $\Lambda_{\text{DC}}(x, \theta_\gamma) \propto 1 + x/2$  cancel identically. Since previously we corrected for the effect of  $(1 - e^{-x})/x$ , inclusion of the frequency correction to  $\Lambda_{\text{DC}}(x, \theta_\gamma)$  reverses this correction. In our perturbation approach, this is easy to account for, and for the correction Eq. (72) we also included it.

## 6.2 CS temperature corrections

To include CS temperature correction, we return to Eq. (15) and insert  $\hat{\mu}^{(0)} = e^{-x_c/x}$ . By keeping only terms at lowest order in  $x \ll 1$ , we find the additional source term

$$S_{\text{CS}}^{(1)}(x) = \theta_\gamma \left( \frac{17}{10} + \frac{14}{5} \frac{x_c}{x} - \frac{7}{10} \frac{x_c^2}{x^2} \right) \frac{x_c^2}{x^2} \hat{\mu}^{(0)}. \quad (73)$$

Inserting this into the integral of Eq. (53), we obtain the following frequency dependent correction (cf. Chluba 2005)

$$\hat{\mu}_{\text{CS}}(x) \approx C_{\text{CS}} e^{-x_c/x} - \theta_\gamma D_{\text{CS}}(x_c/x) \quad (74)$$

$$D_{\text{CS}}(\zeta) = \frac{\zeta}{2} \left( \frac{11}{4} + \frac{21}{20}\zeta - \frac{7}{30}\zeta^2 \right) e^{-\zeta}. \quad (75)$$

We absorbed any term  $\propto e^{-x_c/x}$  into the constant and also ensured  $\mu \rightarrow 0$  for small  $x$ . Since  $D_{\text{CS}}(\zeta) > 0$  around  $x \approx 2x_c$ , the main effect of CS temperature corrections is to move the critical frequency of the solution towards slightly higher values. The mean shift of the photon energy per scattering is given by  $\Delta\nu/\nu \approx 4\theta_e[1 + (5/2)\theta_e]$  (e.g., Sazonov & Sunyaev 2000), which makes CS win the upper hand over DC at slightly higher frequencies, but this time decreasing the effective photon production rate, because according to  $\Lambda_{\text{DC}}/[\theta_\gamma(1+(5/2)\theta_\gamma)] = x_c^2$  the effective critical frequency decreases by  $\Delta x_c \approx -(5/4)x_c\theta_\gamma$ . Inserting the correction function  $D_{\text{CS}}(\zeta)$  into the emission integral Eq. (32), we find

$$\Delta I_{\hat{\mu}}^{\text{CS}} \approx -[1.72 - 0.82x_c]\theta_\gamma, \quad (76)$$

which is in agreement with the argument given above. The effect is slightly larger than expected from the simple estimate  $\Delta I_{\hat{\mu}}^{\text{CS}} \approx -(5/4)\theta_\gamma x_c$ . This is likely due to the higher derivative terms and the precise shape of  $D_{\text{CS}}(\zeta)$ . The final correction to the thermalization optical depth thus is

$$\Delta\tau_{\hat{\mu}}^{\text{CS}}(z, 0) \approx -1.23\theta_\gamma (z/z_{\text{dc}})^{5/2}, \quad (77)$$

which is roughly  $\approx 4$  times smaller than the DC temperature correction, but it goes into the same direction. This is in good agreement with the estimates of Chluba (2005).

Adding the CS correction to our perturbation treatment, we obtain the thick solid black line in Fig. 13. Since the correction due to CS appears to be relatively small, we did not go through the trouble of implementing the effect numerically for `CosmoTHERM`. It is possible to iteratively include the correction using a Compton kernel approach, similar to how it was done in connection with refined helium recombination calculations (Chluba et al. 2012a). However,

we leave a numerical confirmation of the CS scattering correction to some future work.

For completeness, by redetermining the normalization of our full solution (see Sect. 5.3.2), we find a small negative correction  $\Delta A_{\text{CS}} \approx x_c A_0(x_c)[0.138 + 1.06 \ln x_c]\theta_\gamma$ , which usually is negligible.

## 7 CONCLUSION

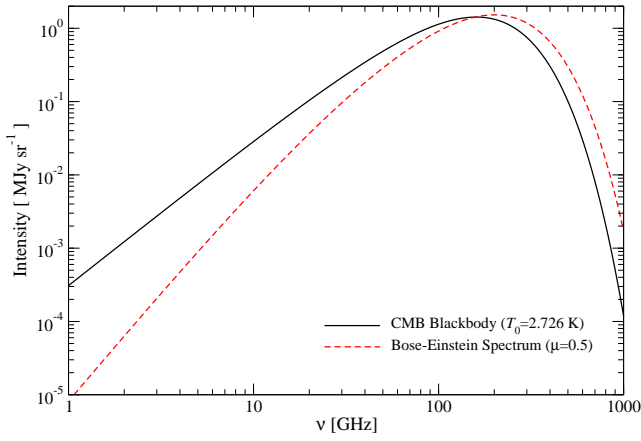
We carried out a systematic study of approximations for the distortion visibility function and early  $\mu$ -distortions, basing our analysis on a perturbative expansion of the solution in terms of the critical frequency,  $x_c \ll 1$ . Our approximations for both the distortion visibility function and the  $\mu$ -type distortions, over a wide range of redshifts and frequencies, agree very well with the numerical solutions obtained with `CosmoTHERM`. Only a few simple integrals have to be evaluated numerically, speeding the computation up from several hours<sup>11</sup> down to seconds.

We demonstrate that the high-frequency chemical potential scales like  $\mu(x) \propto \text{const} + \partial_\gamma(T_e/T_\gamma) \ln x$  (see Sect. 5.5.1). The shape of the high-frequency spectrum is thus driven by the evolution of the electron temperature, giving rise to an  $n(x) \approx x^\gamma e^{-x}$  dependence of the photon occupation number at large  $x = h\nu/kT_\gamma$ . The non-stationary correction caused by the time derivative of the chemical potential amplitude discussed by KS12 is noticeable at low frequencies, although a slightly larger term  $\propto \ln(x/x_c)e^{-x_c/x}$  arises due to Compton scattering frequency corrections [see Eq. (68)]. At intermediate frequencies  $x \approx 1$ , additional modifications due to the Compton process become noticeable (Sect. 5.5.3). These corrections allow us to smoothly match the solutions for the  $\mu$ -distortion obtained in the high- and low-frequency limit, giving an accurate description of the distortion at  $z \gtrsim 3 \times 10^5$ , with no extra calibration of constants relative to the numerical result required (Fig. 11). This extends the validity for the  $\mu$ -distortion approximations in comparison to the approach of KS12 at practically no additional cost.

Overall our results for the distortion visibility function agree well with the approximations of KS12. Being slightly more elaborate, our approach seems to agree a bit better with our numerical result (see Fig. 10). We argue (see Sect. 5.4.1) that part of the difference could be related to the precise definition of what the distortion visibility function really is, which in our case is ensured to represent the momentary fraction of energy that is stored by the distortion. We will investigate the source of the deviations in a future work. A simple code for computing the distortion visibility function will be made available at [www.Chluba.de/CosmoTherm](http://www.Chluba.de/CosmoTherm).

Finally, in Sect. 6 we explain how DC and CS relativistic corrections affect the distortion visibility at  $10^6 \lesssim z$ . DC and CS temperature corrections decrease the thermalization efficiency, with the effect reaching  $\Delta\mathcal{J}/\mathcal{J}_{\text{DC}} \approx 27\%$  at  $z \approx 6 \times 10^6$ . This is canceled by DC frequency-dependent corrections to the Gaunt factor, giving rise to a net  $\Delta\mathcal{J}/\mathcal{J}_{\text{DC}} \approx -10\%$  at  $z \approx 6 \times 10^6$  (see Fig. 13). Including all corrections discussed here, at  $z \approx 6 \times 10^6$  the distortion visibility function thus is about a factor of  $\approx 2$  smaller than the simplest approximation,  $\mathcal{J}_{\text{DC}}(z) = e^{-(z/z_{\text{dc}})^{2.5}}$ , a modification that is important for the interpretation of future spectral distortion data (see Figs. 8 and 9 for more details).

<sup>11</sup> For sampling the visibility function at  $\approx 60$  redshifts.



**Figure A1.** Shape of the CMB spectrum with large chemical potential. For the considered case, the crossover frequency is at  $\nu \approx 158\text{GHz}$ . Number changing processes at low frequencies were neglected, but would restore the blackbody shape at  $\nu \lesssim 1\text{GHz}$ .

## ACKNOWLEDGMENTS

JC thanks Liang Dai, Donghui Jeong, Marc Kamionkowski and Josef Pradler for stimulating discussions of the problem. He is also grateful to Geoff Vasil for useful discussions on multi-scale perturbation theory, and Rishi Khatri and Rashid Sunyaev for their helpful feedback on the paper. Use of the GPC supercomputer at the SciNet HPC Consortium is acknowledged. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto. This work was supported by DoE SC-0008108 and NASA NNX12AE86G.

## APPENDIX A: BOSE-EINSTEIN SPECTRUM FOR FIXED NUMBER AND ENERGY DENSITY

Assuming that the photon occupation number is given by a Bose-Einstein spectrum, we can determine the precise shape from the number and energy density of the distribution. Using the ansatz,  $n = 1/(e^{x+\mu} - 1)$  [ $\phi$  is needed to fix the correct number density and  $\mu > 0$  is constant], we can write

$$\phi = \left( \frac{2 \text{Li}_3(e^{-\mu})}{G_2^{\text{Pl}}} \right)^{1/3} \approx 1 - 0.4561\mu - 0.137\mu^2 \ln \mu, \quad (\text{A1})$$

where  $\text{Li}_n(x)$  is the polylogarithm. We assumed that the number density of the photon distribution did not change. With this solution, one can obtain the correct Bose-Einstein spectrum as a function of  $x$  and  $\mu$  (see Fig. A1). Fixing the energy density, we find that

$$1 + \frac{\Delta\rho_\gamma}{\rho_\gamma} = \frac{6 \text{Li}_4(e^{-\mu})}{\rho^4 G_3^{\text{Pl}}} \approx 1 + 0.7140\mu + (0.815 + 0.555 \ln \mu)\mu^2 \quad (\text{A2})$$

can be used to determine the value of  $\mu$ . Evidently, at lowest order one has  $\mu \approx 1.401\Delta\rho_\gamma/\rho_\gamma$ , as expected.

One interesting aspect is that for larger values of  $\mu$ , the zero crossing of the distortion with respect to the blackbody increases. The crossover frequency is roughly given by  $\nu_{\text{cr}} \approx 124\text{GHz}(1 - 0.304\mu \ln \mu)$ , so that even for very large values of  $\mu \approx 0.01$  the zero does not change dramatically.

## APPENDIX B: ENTROPY OF A NON-EQUILRIUM BOSE-EINSTEIN SPECTRUM

In terms of the photon occupation number,  $n = 1/(e^{x+\mu} - 1)$ , the photon entropy density can be written as (Landau & Lifshitz 1980)

$$\begin{aligned} s_\gamma &= 8\pi k \left( \frac{kT_\gamma}{hc} \right)^3 \int x^2 [(1+n) \ln(1+n) - n \ln n] dx \\ &= 8\pi k \left( \frac{kT_\gamma}{hc} \right)^3 \int x^2 [\ln(1+n) + n(x+\mu)] dx \\ &= \frac{4}{3} \frac{\rho_\gamma}{T_\gamma} - \frac{8\pi k}{3} \left( \frac{kT_\gamma}{hc} \right)^3 \int x^3 \mu \partial_x n dx \\ &\stackrel{\mu \ll 1}{\approx} \frac{4}{3} \frac{\rho_\gamma^{\text{Pl}}(T_\gamma)}{T_\gamma} \left[ 1 + 3 \frac{\Delta T_e}{T_\gamma} \right] - \frac{\rho_\gamma^{\text{Pl}}(T_\gamma)}{T_\gamma} \mu_\infty \mathcal{M}_3 \\ &= \frac{4\mathcal{G}_3^{\text{Pl}}}{3\mathcal{G}_2^{\text{Pl}}} kN_\gamma + \frac{\kappa_\rho \rho_\gamma^{\text{Pl}}(T_\gamma)}{3 T_\gamma} \mu_\infty \approx 3.601kN_\gamma [1 + 0.5355\mu_\infty^*], \quad (\text{B1}) \end{aligned}$$

where we used  $\rho_\gamma^{\text{Pl}}(T) = (\mathcal{G}_3^{\text{Pl}}/\mathcal{G}_2^{\text{Pl}})kTN_\gamma^{\text{Pl}}(T) \approx 2.701kTN_\gamma^{\text{Pl}}(T)$  and the effective chemical potential  $\mu_\infty^* = \bar{k}_\rho \mu_\infty$ .

## REFERENCES

- André P. et al., 2014, JCAP, 2, 6  
 Bond J. R., 1996, in Cosmology and Large Scale Structure, Schaeffer R., Silk J., Spiro M., Zinn-Justin J., eds., p. 469  
 Burigana C., Danese L., de Zotti G., 1991, A&A, 246, 49  
 Burigana C., de Zotti G., Danese L., 1995, A&A, 303, 323  
 Burigana C., Salvaterra R., 2003, MNRAS, 342, 543  
 Challinor A., Lasenby A., 1998, ApJ, 499, 1  
 Chluba J., 2005, PhD thesis, LMU München  
 Chluba J., 2013a, MNRAS, 436, 2232  
 Chluba J., 2013b, MNRAS, 434, 352  
 Chluba J., Fung J., Switzer E. R., 2012a, MNRAS, 423, 3227  
 Chluba J., Jeong D., 2014, MNRAS, 438, 2065  
 Chluba J., Khatri R., Sunyaev R. A., 2012b, MNRAS, 425, 1129  
 Chluba J., Sazonov S. Y., Sunyaev R. A., 2007, A&A, 468, 785  
 Chluba J., Sunyaev R. A., 2008, A&A, 488, 861  
 Chluba J., Sunyaev R. A., 2012, MNRAS, 419, 1294  
 Danese L., de Zotti G., 1977, Nuovo Cimento Rivista Serie, 7, 277  
 Danese L., de Zotti G., 1982, A&A, 107, 39  
 Fixsen D. J., 2009, ApJ, 707, 916  
 Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576  
 Fixsen D. J., Mather J. C., 2002, ApJ, 581, 817  
 Hu W., 1995, arXiv:astro-ph/9508126  
 Hu W., Silk J., 1993, Phys.Rev.D, 48, 485  
 Illarionov A. F., Sunyaev R. A., 1975, Soviet Astronomy, 18, 413  
 Itoh N., Kohyama Y., Nozawa S., 1998, ApJ, 502, 7  
 Itoh N., Sakamoto T., Kusano S., Nozawa S., Kohyama Y., 2000, ApJS, 128, 125  
 Khatri R., Sunyaev R. A., 2012a, JCAP, 9, 16  
 Khatri R., Sunyaev R. A., 2012b, JCAP, 6, 38  
 Kogut A. et al., 2006, New Astronomy Reviews, 50, 925  
 Kogut A. et al., 2011, JCAP, 7, 25  
 Kompaneets A., 1956, Sov.Phys. JETP, 31, 876  
 Landau L. D., Lifshitz E. M., 1980, Statistical physics. Pt.1, Pt.2  
 Lightman A. P., 1981, ApJ, 244, 392  
 Mather J. C. et al., 1994, ApJ, 420, 439  
 Nozawa S., Itoh N., Kohyama Y., 1998, ApJ, 507, 530  
 PRISM Collaboration et al., 2013, ArXiv:1306.2259  
 Procopio P., Burigana C., 2009, A&A, 507, 1243  
 Rybicki G. B., Lightman A. P., 1979, Radiative processes in astrophysics. New York, Wiley-Interscience, 1979. 393 p.  
 Sazonov S. Y., Sunyaev R. A., 1998, ApJ, 508, 1  
 Sazonov S. Y., Sunyaev R. A., 2000, ApJ, 543, 28  
 Sazonov S. Y., Sunyaev R. A., 2001, Astronomy Letters, 27, 481

- Seiffert M. et al., 2011, *ApJ*, 734, 6  
Sunyaev R. A., Khatri R., 2013, *International Journal of Modern Physics D*, 22, 30014  
Sunyaev R. A., Zeldovich Y. B., 1970, *ApSS*, 7, 20  
Thorne K. S., 1981, *MNRAS*, 194, 439  
Zannoni M., Tartari A., Gervasi M., Boella G., Sironi G., De Lucia A., Passerini A., Cavaliere F., 2008, *ApJ*, 688, 12  
Zel'dovich Y. B., Levich E. V., 1970, *Soviet Journal of Experimental and Theoretical Physics Letters*, 11, 35  
Zeldovich Y. B., Sunyaev R. A., 1969, *ApSS*, 4, 301