

PREPARED FOR SUBMISSION TO JCAP

Towards a precision calculation of N_{eff} in the Standard Model IV: Impact of positronium formation

Tobias Binder,^a Marco Drewes,^b Yannis Georis,^b
Michael Klasen,^c Giovanni Pierobon^d and Yvonne Y. Y. Wong^d

^aPhysik Department T31, James-Frank-Straße 1, Technische Universität München,
D-85748 Garching, Germany

^bCentre for Cosmology, Particle Physics and Phenomenology (CP3), Université catholique
de Louvain, Chemin du Cyclotron 2, B-1348 Louvain-la-Neuve, Belgium

^cInstitut für Theoretische Physik, Universität Münster, Wilhelm-Klemm-Straße 9, D-48149
Münster, Germany

^dSydney Consortium for Particle Physics and Cosmology, School of Physics, The University
of New South Wales, Sydney NSW 2052, Australia

E-mail: tobias.binder@tum.de, marco.drewes@uclouvain.be,
yannis.georis@uclouvain.be, michael.klasen@uni-muenster.de,
g.pierobon@unsw.edu.au, yvonne.y.wong@unsw.edu.au

Abstract. We present a first assessment of how the previously unexplored effect of positronium formation can impact on the value of the effective number of neutrino species in the Standard Model, $N_{\text{eff}}^{\text{SM}}$. Adopting a Yukawa form for the electrostatic potential, we discuss two possible scenarios that differ primarily in their assumptions about entropy evolution. The first, out-of-equilibrium scenario assumes that thermal corrections to the potential such as Debye screening prevent positronium from appearing until the temperature drops below a threshold. Once the threshold is reached, entropy generated in the QED sector from the equilibration process, if instantaneous, leads to a variation in $N_{\text{eff}}^{\text{SM}}$ of at most $|\Delta N_{\text{eff}}| \sim 10^{-4}$, comparable to other uncertainties in the current benchmark value for $N_{\text{eff}}^{\text{SM}}$. A more gradual formation could however yield a larger change. The second, equilibrium scenario assumes the QED sector to stay in equilibrium at all times. In this case, we show that cancellations between the first, s -wave bound- and scattering-states contributions ensure that it is possible to evolve the system across the bound-state formation threshold without generating entropy in the QED sector. The corresponding change in $N_{\text{eff}}^{\text{SM}}$ then closely matches the $\mathcal{O}(e^2)$ perturbative result derived in previous works and the $\mathcal{O}(e^4)$ contribution is capped at $|\Delta N_{\text{eff}}| \lesssim 10^{-6}$. We also comment on the impact of deviations from a pure Yukawa potential due to the presence of a thermal width.

Contents

1	Introduction	1
2	Out-of-equilibrium positronium formation and its impact on $N_{\text{eff}}^{\text{SM}}$	4
2.1	Change in $N_{\text{eff}}^{\text{SM}}$ from instantaneous positronium production	4
2.2	Change in $N_{\text{eff}}^{\text{SM}}$ from non-instantaneous positronium production	7
3	When does positronium form?	9
3.1	Existence of the positronium at high temperatures	10
3.2	Positronium equilibration rate	13
4	Positronium as non-ideal gas correction	15
4.1	Perturbative computation of non-ideal corrections	16
4.2	Beth-Uhlenbeck formula	18
4.3	Impact on N_{eff}	20
5	Conclusions and outlook	22
A	Functions in the continuity equation	24
B	Beth-Uhlenbeck formula for a non-relativistic e^+e^- plasma	25
B.1	Charging formula	25
B.2	Non-relativistic four-point correlators	26
B.3	Integral reductions	27
C	Non-relativistic Feynman rules in momentum space	28

1 Introduction

The past few years have seen a renewed interest in the precision theoretical computation of the so-called effective number of neutrinos, N_{eff} , in the context of the Standard Model (SM) of particle physics [1–8].¹ Defined via an energy density ratio,

$$N_{\text{eff}}^{\text{SM}} \equiv \frac{8}{7} \left(\frac{11}{4} \right)^{4/3} \frac{\rho_\nu}{\rho_\gamma} \bigg|_{T/m_e \rightarrow 0, T/m_\nu \gg 0}, \quad (1.1)$$

the $N_{\text{eff}}^{\text{SM}}$ parameter quantifies the energy density contained in three ultra-relativistic SM neutrino species in the early universe after the electron-positron annihilation epoch ($T \sim 500$ keV) relative to the energy density in photons. Changes to the value of N_{eff} impact directly on the universal expansion rate at early times and are in principle observable by a multitude of cosmological probes including the primordial light element abundances and the cosmic microwave background (CMB) anisotropies. State-of-the-art calculations put the SM theoretical value at $N_{\text{eff}}^{\text{SM}} = 3.0440 \pm 0.0002$ [2, 3, 5], although some dispute remains about the value of the last but one significant digit [6–8]. This number and the accompanying

¹In this work, the term “Standard Model of particle physics” is understood to include neutrino oscillations.

Standard-Model corrections to $N_{\text{eff}}^{\text{SM}}$	Leading contribution
m_e/T_d correction	+0.04
$\mathcal{O}(e^2)$ FTQED correction to the QED EoS	+0.01
Non-instantaneous decoupling+spectral distortion	-0.006
$\mathcal{O}(e^3)$ FTQED correction to the QED EoS	-0.001
Flavour oscillations	+0.0005
Type (a) FTQED corrections (thermal mass) to the weak rates	$\lesssim 10^{-4}$
Type (d) FTQED corrections (fermion loop) to the weak rates	$\lesssim 10^{-5}$
$\mathcal{O}(e^4)$ FTQED correction to the QED EoS	3.5×10^{-6}
Electron/positron chemical decoupling	$\sim -10^{-7}$
Sources of uncertainty	
Numerical solution by FortEPiaNO	± 0.0001
Input solar neutrino mixing angle θ_{12}	± 0.0001

Table 1. Leading-digit contributions from various SM corrections, in order of importance, thus far investigated that make up $N_{\text{eff}}^{\text{SM}}$. The first five corrections have been accounted for in precision computations leading to $N_{\text{eff}}^{\text{SM}} - 3 = 0.0440 \pm 0.0002$ [2, 3, 5]; the rest are presently not part of the standard calculation as they fall below current numerical uncertainties. Type (a) FTQED corrections to the weak rates refer to thermal corrections to the electron mass which also change the electron phase space, while type (d) corrections were considered in references [7, 8] and correspond to a closed-fermion loop diagram that corrects, roughly speaking, the νe -scattering/annihilation matrix element. Electron/positron chemical decoupling refers to a small amount of entropy remaining in the electrons due to incomplete e^+e^- -annihilation after chemical decoupling at $T \simeq 20$ keV, which changes the equation of state of the QED sector [26]. Table adapted from [5, 8]; see also explanations therein.

uncertainty are to be contrasted with the current observational constraint of $N_{\text{eff}}^{\text{obs}} = 2.99 \pm 0.34$ (95% C.I.) from the Planck CMB mission [9] and a potential future sensitivity of ± 0.02 to ± 0.03 by CMB-S4 [10].

Within the SM, the physics that determines the precise value of N_{eff} takes place in the temperature range $\mathcal{O}(10)$ keV $\lesssim T \lesssim \mathcal{O}(10)$ MeV, when the universe’s energy is dominated by three families of SM neutrinos and a quantum electrodynamics (QED) sector comprising an electron/positron and photon bath. The leading digit of $N_{\text{eff}}^{\text{SM}} \simeq 3$ accounts for three neutrino families. A basic description of neutrino decoupling from the QED plasma together with the QED equation of state to $\mathcal{O}(e^2)$ suffice to predict the leading-order correction of $\Delta N_{\text{eff}} \simeq 0.04$ [11–20]. Further refinements to the computation can be made to pin down the next digits. These include finite-temperature corrections to the QED equation of state to $\mathcal{O}(e^3)$ and higher [4], distortions of the neutrino distribution from Fermi-Dirac statistics due to non-instantaneous neutrino decoupling and neutrino flavour oscillations [1–3, 5, 21–24], QED corrections to the weak interaction rates that govern neutrino decoupling [6–8], as well as the interplay between anisotropies and neutrino self-interactions [25]. See table 1 for a summary. While these refinements are likely out of the reach of the forthcoming generation of cosmological probes, it is nonetheless useful to quantify their contributions to the theoretical value of $N_{\text{eff}}^{\text{SM}}$, so as to beat down the theoretical uncertainty in cosmological calculations and, in so doing, eliminate the need for error propagation in cosmological data analyses.

In this work, we add to this list of refinements and consider for the first time the impact of positronium (Ps) formation on the theoretical value of $N_{\text{eff}}^{\text{SM}}$. Positronia, bound states of

Property	Expression at $T = 0$	Numerical value
Binding energy of energy level n	$\Delta E_n = m_e \alpha^2 / (4n^2)$	$6.8/n^2$ eV
Mass	$m_{\text{Ps}} = 2m_e - \Delta E_1$	1.022 MeV
Bohr radius	$a_0 = 2/(m_e \alpha)$	536 MeV $^{-1}$
Lifetime (for $n = 1$):		
• Para-positronium ($s = 0$)	$\tau_0 = 2/(m_e \alpha^5)$	0.1244 ns
• Ortho-positronium ($s = 1$)	$\tau_1 = 9\pi/[2(\pi^2 - 9)m_e \alpha^6]$	138.7 ns

Table 2. Physical properties of the positronium at $T = 0$ (see, e.g., [29]). The quantities $m_e \simeq 511$ keV and $\alpha = e^2/(4\pi) \simeq 1/137$ are the electron mass and the fine structure constant, respectively.

e^+e^- , may form when the electron/positron bath becomes sufficiently non-relativistic, i.e., at $T \lesssim m_e$. Although these bound states are unstable and subject to ionisation by the photon bath, should the conditions allow a *transient* population of such entities to exist and be produced and maintained in a state of *dynamical equilibrium* with the rest of the QED plasma at $\mathcal{O}(10)$ keV $\lesssim T \lesssim m_e$, they could nonetheless alter the dynamics of the universe in ways that could affect $N_{\text{eff}}^{\text{SM}}$. Specifically, given that at these temperatures the bulk of the neutrino population has long since decoupled from the QED plasma, we expect a small population of positronium to alter $N_{\text{eff}}^{\text{SM}}$ predominantly via its impact on the QED equation of state, while corrections to the neutrino interaction rates are likely negligible. Our main task in this work, therefore, is to assess under what conditions positronium can form and be maintained in equilibrium in the time frame of interest, and its likely impact on the evolution of the QED plasma and hence $N_{\text{eff}}^{\text{SM}}$.

As a final note, we emphasise that our scenario of positronium formation differs from the common understanding of, e.g., hydrogen recombination or primordial nucleosynthesis. In the latter two cases, the goal is to produce a frozen-out population of neutral hydrogen or nuclei free from ionisation or photodissociation. On the other hand, our scenario does not require the positronium to avoid ionisation altogether as long as, once it is allowed to form, its formation rate is large enough to maintain a dynamical equilibrium with the free electrons/positrons in the plasma in the presence of ionising radiation. This situation is of course also possible for neutral hydrogen and nuclei at high temperatures, but is generally of negligible consequence to the evolution of the universe because of the highly-suppressed baryon number density during radiation domination. The closest parallel we are aware of to our scenario is in fact the formation of positronium-like dark matter metastable bound states prior to dark matter freeze-out (e.g., [27, 28]), although in that case the focus is not the equation of state of the universe, but the effect of bound states on freeze-out itself.

The paper is organised as follows. We begin in section 2 by asking how large an effect we can expect on $N_{\text{eff}}^{\text{SM}}$ if a population of positronium was produced out-of-equilibrium at some low temperature $T \lesssim m_e$, while section 3 examines the conditions under which positronium can form and how fast they form. Section 4 considers the opposite limit in which positronium production takes place adiabatically and the QED plasma remains in equilibrium at all times. We discuss future directions and conclude in section 5.

2 Out-of-equilibrium positronium formation and its impact on $N_{\text{eff}}^{\text{SM}}$

Since positronium can only form when electrons and positrons are sufficiently non-relativistic, we focus our attention on the temperature epoch $10 \text{ keV} \lesssim T \lesssim m_e$. Here, the universe's energy content comprises predominantly a QED plasma of photons and electrons/positrons held in thermal equilibrium at temperature T by electromagnetic interactions, and three generations of SM neutrinos of temperature T_ν , which have largely decoupled from the QED sector. We assume as usual that the system is CP -symmetric with a vanishing chemical potential. Table 2 summarises some basic properties of the positronium at $T = 0$ that will be useful for our discussion.

We estimate first in section 2.1 the effect of instantaneous positronium production on $N_{\text{eff}}^{\text{SM}}$ via an entropy argument, before relaxing the instantaneous-production assumption in section 2.2.

2.1 Change in $N_{\text{eff}}^{\text{SM}}$ from instantaneous positronium production

Consider three epochs labelled by their scale factors, a_1 , a_2 , and a_3 , corresponding respectively to

1. the time of instantaneous neutrino decoupling, $T = T_\nu \simeq \mathcal{O}(1) \text{ MeV}$;
2. positronium production, which we assume to be instantaneous and populate both $n = 1$ singlet and triplet states, i.e., the para- and the ortho-positronium, at their equilibrium number densities at the same time; and
3. some late time when all e^+ , e^- , and Ps have annihilated to photons, where $N_{\text{eff}}^{\text{SM}}$ is defined via equation (1.1).

Except at a_1 , the QED sector and neutrino temperatures differ, i.e., $T \neq T_\nu$.

Assuming that kinetic and chemical equilibrium holds at all times for all particle species *except* at the very instant positronium forms at a_2 , entropy conservation tells us that

$$\begin{aligned} s_1 a_1^3 &= s_{2-} a_2^3, \\ s_{2+} a_2^3 &= s_3 a_3^3, \end{aligned} \tag{2.1}$$

where s denotes the physical entropy density in the QED sector, the subscripts 2- and 2+ mean just before and just after positronium formation, and we define $\delta s \equiv s_{2+} - s_{2-}$ to be the change in entropy at a_2 . Then, using the fact that the neutrino temperature scales as $T_\nu \propto a^{-1}$ after decoupling, the relations (2.1) can be combined to give

$$\frac{s_1}{T_{\nu,1}^3} \left(1 + \frac{\delta s}{s_{2-}} \right) = \frac{s_3}{T_{\nu,3}^3}, \tag{2.2}$$

and hence

$$\frac{T_{\nu,3}}{T_3} = \left[\frac{g_{*s,1}^{\text{QED}}}{g_{*s,3}^{\text{QED}}} \left(1 + \frac{\delta s}{s_{2-}} \right) \right]^{-1/3} = \left(\frac{4}{11} \right)^{1/3} \left(1 + \frac{\delta s}{s_{2-}} \right)^{-1/3}, \tag{2.3}$$

where we have used $s \propto g_{*s}^{\text{QED}} T^3$, and $g_{*s,1}^{\text{QED}} = 11/2$ and $g_{*s,3}^{\text{QED}} = 2$ denote, respectively the entropy degrees of freedom in the QED sector at neutrino decoupling and at $T/m_e \rightarrow 0$.² It follows straightforwardly that the neutrino and photon energy densities at late times are related via

$$\rho_{\nu,3} = 3 \times \frac{7}{8} \left(\frac{4}{11} \right)^{4/3} \left(1 + \frac{\delta s}{s_{2-}} \right)^{-4/3} \rho_{\gamma,3}. \quad (2.4)$$

Following the definition (1.1) and writing $N_{\text{eff}}^{\text{SM}} = 3 + \Delta N_{\text{eff}}$, we then obtain

$$\Delta N_{\text{eff}} = 3 \left[\left(1 + \frac{\delta s}{s_{2-}} \right)^{-4/3} - 1 \right] \simeq -4 \frac{\delta s}{s_{2-}} \quad (2.5)$$

for the change in $N_{\text{eff}}^{\text{SM}}$ due to instantaneous positronium production.

It remains for us to estimate the change in entropy δs . This can be achieved by matching the energy density in the QED plasma at the time just before positronium formation begins, with the energy density just after production is complete, i.e.,

$$\rho_e(a_2, T_2) + \rho_\gamma(a_2, T_2) = \rho_e(a_2, T_{2,\text{new}}) + \rho_\gamma(a_2, T_{2,\text{new}}) + \rho_{\text{Ps}}(a_2, T_{2,\text{new}}), \quad (2.6)$$

and then solving for $T_{2,\text{new}}$, where $T_{2,\text{new}}$ is the new temperature the system settles to after an equilibrium positronium population appears instantaneously at a_2 . Because positronium formation can only occur at low temperatures $T \ll m_e$, we expect $\rho_{\text{Ps}} \ll \rho_\gamma$ and hence a small temperature change $\delta T \equiv T_{2,\text{new}} - T_2$, $|\delta T/T_2| \ll 1$. Then, equation (2.6) can be expanded in δT to give

$$\delta T \simeq - \left. \frac{\rho_{\text{Ps}}}{\frac{\partial}{\partial T} (\rho_e + \rho_\gamma + \rho_{\text{Ps}})} \right|_{T=T_2}, \quad (2.7)$$

and the change in entropy can be computed in the same limit from

$$\begin{aligned} \delta s &\simeq s_{\text{Ps}}(T_2) + \left. \frac{\partial}{\partial T} (s_e + s_\gamma + s_{\text{Ps}}) \right|_{T=T_2} \delta T \\ &= s_{\text{Ps}}(T_2) - \left. \frac{\frac{\partial}{\partial T} (s_e + s_\gamma + s_{\text{Ps}})}{\frac{\partial}{\partial T} (\rho_e + \rho_\gamma + \rho_{\text{Ps}})} \right|_{T=T_2} \rho_{\text{Ps}}(T_2). \end{aligned} \quad (2.8)$$

Noting that $s_\gamma \gg s_e, s_{\text{Ps}}$ and $\rho_\gamma \gg \rho_e, \rho_{\text{Ps}}$ at $T \ll m_e$, we use $s_\gamma \propto (4/3) T^3$ and $\rho_\gamma \propto T^4$ to further approximate

$$\delta s \simeq s_{\text{Ps}}(T_2) - \frac{\rho_{\text{Ps}}(T_2)}{T_2} = \frac{P_{\text{Ps}}(T_2)}{T_2}, \quad (2.9)$$

where P_{Ps} denotes the positronium pressure. Substituting this δs estimate into equation (2.5) then gives a change in $N_{\text{eff}}^{\text{SM}}$ of

$$\Delta N_{\text{eff}} \simeq -4 \times \frac{P_{\text{Ps}}(T_2)/T_2}{s_\gamma(T_2) + s_e(T_2)} \simeq -\frac{45}{\pi^2} \frac{P_{\text{Ps}}(T_2)}{T_2^4} < 0, \quad (2.10)$$

²The number of entropy degrees of freedom in the QED sector at neutrino decoupling is in fact less than $g_{*s,1}^{\text{QED}} = 11/2$ because of a finite electron mass and finite-temperature corrections to the QED equation of state at $T \sim \mathcal{O}(1)$ MeV. See reference [4] for discussion. This deviation is responsible for the leading-order correction to $N_{\text{eff}}^{\text{SM}}$, bringing its value from exactly 3 to 3.04. However, as a perturbative, 1%-level correction, its impact on the sub-leading correction due to positronium formation we wish to estimate in this work is negligible. Henceforth we shall always ignore this correction. Functionally, this is the same as assuming that neutrinos decouple from the QED sector at $T/m_e \rightarrow \infty$.

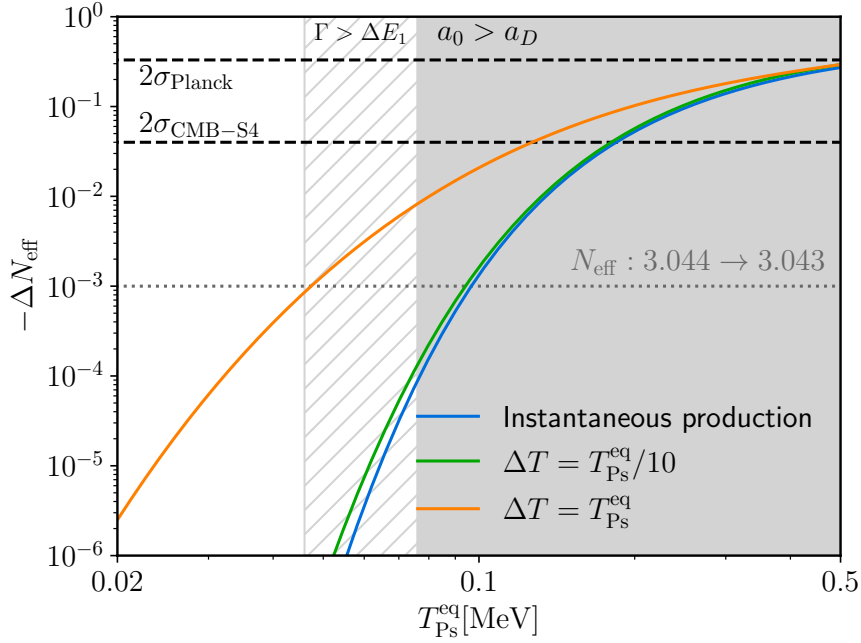


Figure 1. We display the (negative) change in $N_{\text{eff}}^{\text{SM}}$, $-\Delta N_{\text{eff}}$, on a logarithmic scale as a function of the temperature $T_{\text{Ps}}^{\text{eq}}$ at which both $n = 1$ positronium states reach equilibrium. We use two different approximations to compute ΔN_{eff} : The blue line represents estimates from the entropy argument of section 2.1 assuming instantaneous positronium equilibration. The green and orange lines denote ΔN_{eff} from non-instantaneous production modelled with the parameterisation (2.18) for the choices of $\Delta T/T_{\text{Ps}}^{\text{eq}} \in \{0.1, 1\}$, respectively. Positronium formation is prevented by Debye screening in the grey-shaded region. In the hatched region, we expect scatterings with the plasma to prevent the formation of stable bound states, but more detailed computations are required to quantify the exact impact. The three horizontal lines represent, from top to bottom, the current sensitivity of Planck at 95% C.L. [9], the planned sensitivity of CMB-S4 [10] at 95% C.L., and the last but one significant digit of the SM benchmark value $N_{\text{eff}}^{\text{SM}}$ [1–8].

which we observe is always negative. This is consistent with the expectation that the main effect of positronium formation on the neutrino decoupling era is to slow down the redshifting of the QED bath and hence increase at late times the photon temperature relative to the neutrino one.

Figure 1 shows (blue line) $-\Delta N_{\text{eff}}$ as a function of the instantaneous positronium production temperature T_2 (called $T_{\text{Ps}}^{\text{eq}}$ in the figure), assuming Bose-Einstein statistics for the positronium and an internal degree of freedom of $g_{\text{PS}} = 4$ (i.e., singlet+triplet). Observe that in order to change $N_{\text{eff}}^{\text{SM}}$ by $|\Delta N_{\text{eff}}| \sim 10^{-3}$, instantaneous positronium equilibration needs to be achieved by $T \simeq 97$ keV. This relatively high temperature (in comparison with the pertinent particle mass, m_e) is to be contrasted with the impact of bound-state formation on freeze-out dark matter phenomenology [27, 28, 30], where bound-state dynamics tends to enhance the annihilation cross-section, which delays freeze-out, thereby reducing the dark matter relic abundance. Translated to our system, an analogous cross-section enhancement would push electron/positron chemical decoupling—which takes place in the absence of bound-state dynamics at $T \simeq 20$ keV [26]—down to an even lower temperature. The consequence is a more complete e^+e^- -annihilation prior to chemical decoupling and hence a smaller contribution from the residual electrons post decoupling to the entropy in the QED sector. For $N_{\text{eff}}^{\text{SM}}$,

this means the correction due to a residual electron population would be *even smaller* in the presence of bound-state dynamics than the estimate of reference [26] – $|\Delta N_{\text{eff}}| \sim 10^{-7}$ (see also table 1), computed *without* bound-state considerations – would suggest.³

Finally, we emphasise again that, although the temperatures of interest far exceed the positronium binding energy ΔE_1 of a few eV (see table 2), the presence of a minute positronium population is not precluded from *purely thermodynamic* considerations. In fact, the positronium and free electron/positron number densities, n_{Ps} and $n_{e^\pm}$, in our scenario are fully consistent with the Saha equation (see, e.g., [32]), $n_{e^-}/n_{\text{Ps}} \propto (1/n_{e^+})(m_e T/2\pi)^{3/2} e^{-\Delta E_1/T} \sim 2.1 \times 10^2$ at $T \simeq 80$ keV and $\sim 9.7 \times 10^3$ at $T \simeq 50$ keV – indicating that the vast majority of the electrons/positrons is, expectedly, unbound in the regime of interest. This also provides a justification for treating the electrons/positrons as a thermal bath. Of course, at $T \gg \Delta E_1$ any individual positronium system is likely to be dissociated on times scales far shorter than the laboratory lifetimes given in table 2. However, for positronium to impact on the QED equation of state, it suffices that a transient population is maintained in dynamic equilibrium through efficient formation. The questions of whether or not the formation rate is adequate to maintain this equilibrium, and, indeed, if positronia are even permitted to exist by finite-temperature requirements are deferred to section 3.

2.2 Change in $N_{\text{eff}}^{\text{SM}}$ from non-instantaneous positronium production

Relaxing the assumption of instantaneous equilibration should in general increase entropy production, thereby leading to a larger $|\Delta N_{\text{eff}}|$. To account for this effect, we incorporate the contribution of positronium to the energy density of the universe directly in the continuity equation. As done in references [4, 5, 8] and following the notation of [18, 23], the continuity equation for all particle species can be expressed as

$$\frac{dz}{dx} = \frac{(\bar{\rho}_{\text{tot}} - 3\bar{P}_{\text{tot}})/x - \partial \bar{\rho}_{\text{tot}}/\partial x}{\partial \bar{\rho}_{\text{tot}}/\partial z} \quad (2.11)$$

in terms of the comoving variables $x = m_e a$, $y = p a$, and $z = T a$, the rescaled total energy density $\bar{\rho}_{\text{tot}} \equiv \bar{\rho}_e + \bar{\rho}_\gamma + \bar{\rho}_{\text{Ps}} + \bar{\rho}_\nu = \rho_{\text{tot}} a^4$, and the rescaled total pressure $\bar{P}_{\text{tot}} = P_{\text{tot}} a^4$.

In presence of photons, electrons and neutrinos only, the continuity equation (2.12) takes the standard form [8, 18]

$$\frac{dz}{dx} = \frac{\tau J_+(\tau) - 1/(2z^3) d\bar{\rho}_\nu/dx + G_1(\tau)}{\tau^2 J_+(\tau) + Y_+(\tau) + 2\pi^2/15 + G_2(\tau)}, \quad (2.12)$$

where $\tau = x/z = m_e/T$, the functions J_+ and Y_+ express the evolution of the electron/positron distribution across the annihilation epoch in the ideal gas limit, while G_1 and G_2 incorporate corrections to the QED equation of state due to finite temperature effects. Explicit expressions for J_+ , Y_+ , G_1 and G_2 can be found in appendix A. The quantity $d\bar{\rho}_\nu/dx$ corresponds to energy exchange between the neutrino and the QED sectors and can be set to zero if, as in section 2.1, we approximate neutrino decoupling to take place at $T/m_e \rightarrow \infty$. Then, under

³Reference [26] did not specify how they obtained the estimate $\Delta N_{\text{eff}} \sim -10^{-7}$. However, we find that we can reproduce this number using equation (2.5) but with $\delta s/s$ reinterpreted to mean the fractional change in the QED entropy due to electrons/positrons immediately prior to chemical decoupling at $T \simeq 20$ keV. Although the residual electrons remain kinetically coupled to photons after this time until a much lower temperature $T \lesssim 0.2$ eV, the miniscule baryon-to-photon number density ratio ($\sim 10^{-10}$) ensures that they do not alter the photon temperature evolution away from $T \propto a^{-1}$ in any appreciable way [31]. Thus, this estimate of ΔN_{eff} should remain to an extremely good approximation unchanged down to the CMB epoch.

this approximation, our variable z is equivalently the photon-to-neutrino temperature ratio, i.e., $z = T/T_\nu$,⁴ and evolving equation (2.12) from $x \rightarrow 0$ to $x \rightarrow \infty$ for an initial condition $z(x \rightarrow 0) = 1$ allows us to estimate $N_{\text{eff}}^{\text{SM}}$ from the relation

$$N_{\text{eff}}^{\text{SM}} - 3 = 3 \left[\left(\frac{z_{\text{ref}}}{z_{\text{fin}}} \right)^4 - 1 \right], \quad (2.13)$$

where we define $z_{\text{ref}} \equiv (11/4)^{1/3}$. Note that if there is non-zero energy transfer between the neutrino and the QED sectors, i.e., $d\bar{\rho}_\nu/dx \neq 0$, then z_{fin} alone does not capture the change in $N_{\text{eff}}^{\text{SM}}$ —we also need to include contributions from changes in $\bar{\rho}_\nu$.

The continuity equation (2.11) can be straightforwardly generalised to include contributions from a positronium population of energy density and pressure $\bar{\rho}_{\text{Ps}}$ and $\bar{P}_{\text{Ps}}$, which will generally need to be computed from a Boltzmann equation. For our purpose of establishing an estimate, however, we may model the evolution of $\bar{\rho}_{\text{Ps}}$ as

$$\bar{\rho}_{\text{Ps}}(x, y, z) = \sigma(\tau) \bar{\rho}_{\text{Ps}}^{\text{eq}}(x, y, z) = \sigma(\tau) \frac{g_{\text{Ps}}}{2\pi^2} \int dy y^2 \frac{\sqrt{y^2 + x_{\text{Ps}}^2}}{\exp(\sqrt{y^2 + x_{\text{Ps}}^2}/z) - 1}, \quad (2.14)$$

with $x_{\text{Ps}} = (m_{\text{Ps}}/m_e)x \simeq 2x$, and the positronium population is assumed to have a thermal distribution of the same temperature T as the QED plasma but rescaled by a smooth transfer function $\sigma(\tau) = f_{\text{Ps}}/f_{\text{Ps}}^{\text{eq}} = n_{\text{Ps}}/n_{\text{Ps}}^{\text{eq}}$ that depends only on T . The same ansatz for the positronium distribution function is assumed to model the evolution of the pressure $\bar{P}_{\text{Ps}}$. Then, we find

$$\frac{1}{2z^3} \left(\frac{\bar{\rho}_{\text{Ps}} - 3\bar{P}_{\text{Ps}}}{x} - \frac{\partial \bar{\rho}_{\text{Ps}}}{\partial x} \right) = \frac{g_{\text{Ps}}}{4} \left[\tau \left(\frac{m_{\text{Ps}}}{m_e} \right)^2 J_- \left(\frac{m_{\text{Ps}}}{m_e} \tau \right) - z W_- \left(\frac{m_{\text{Ps}}}{m_e} \tau \right) \partial_x \right] \sigma(\tau), \quad (2.15)$$

and

$$\frac{1}{2z^3} \frac{\partial \bar{\rho}_{\text{Ps}}}{\partial z} = \frac{g_{\text{Ps}}}{4} \left[\tau^2 \left(\frac{m_{\text{Ps}}}{m_e} \right)^2 J_- \left(\frac{m_{\text{Ps}}}{m_e} \tau \right) + \frac{m_{\text{Ps}}}{m_e} Y_- \left(\frac{m_{\text{Ps}}}{m_e} \tau \right) + z W_- \left(\frac{m_{\text{Ps}}}{m_e} \tau \right) \partial_z \right] \sigma(\tau), \quad (2.16)$$

where J_- and Y_- are the Bose-Einstein equivalent of J_+ and Y_+ , and $W_-(\tau)$ is a new function, all of which can be found in appendix A. Setting $m_{\text{Ps}} \simeq 2m_e$ and $g_{\text{Ps}} = 4$ and using the property $\partial_z \sigma = -(x/z) \partial_x \sigma$, we obtain for the continuity equation

$$\frac{dz}{dx} = \frac{\tau [J_+(\tau) + 4\sigma(\tau)J_-(2\tau)] - z \partial_x \sigma(\tau) W_-(2\tau)}{\tau^2 [J_+(\tau) + 4\sigma(\tau)J_-(2\tau)] + Y_+(\tau) + Y_-(2\tau) - x \partial_x \sigma(\tau) W_-(2\tau) + 2\pi^2/15}, \quad (2.17)$$

having dropped for simplicity the terms proportional to $G_{1,2}$ and to the neutrino abundance.⁵

For the transfer function $\sigma(\tau)$, we generally expect the energy density of a species far from equilibrium to scale as $\rho \propto 1 - e^{-\int \Gamma_{\text{prod}}/(Hx)dx}$, where Γ_{prod} is the production rate and H the Hubble expansion rate. We therefore adopt as a first guess the *ansatz*

$$\sigma(T) = \frac{1}{2} \left[1 + \tanh \left(\frac{T_{\text{Ps}}^{\text{eq}} - T}{\Delta T} \right) \right], \quad (2.18)$$

⁴This temperature ratio interpretation of the variable z is strictly true only when the neutrinos have completely decoupled from the QED sector.

⁵We have checked explicitly that adopting a more realistic modelling that includes $G_{1,2}$ and $d\bar{\rho}_\nu/dx$ has no substantial impact on our estimate of ΔN_{eff} from a transient positronium population.

where the width of the tanh function ΔT parameterises the effective positronium equilibration timescale, and $T = m_e/\tau = m_e(z/x)$ in terms of the dimensionless variables. We show in figure 1 (green and orange lines) the expected ΔN_{eff} for two choices of $\Delta T \in \{T_{\text{Ps}}^{\text{eq}}/10, T_{\text{Ps}}^{\text{eq}}\}$, which serve as benchmarks for a fast and a slow equilibration respectively. As expected, the faster the equilibration, the more closely the change in $N_{\text{eff}}^{\text{SM}}$ approaches the estimates from entropy arguments (section 2.1); in fact, taking the limit $\sigma(\tau) \rightarrow \Theta[(T_{\text{Ps}}^{\text{eq}} - m_e/\tau)/\Delta T]$, where Θ is a Heaviside function, and adjusting the temperature to account for the sudden increase in entropy at $T = T_{\text{Ps}}^{\text{eq}}$ allow us to recover the estimates of section 2.1 exactly. We have also tested other functional forms of σ , and find that as long as the model contains an exponential dependence on $(T_{\text{Ps}}^{\text{eq}} - T)/\Delta T$, it will yield a ΔN_{eff} similar to the model (2.18).

3 When does positronium form?

We have seen in figure 1 that the change in $N_{\text{eff}}^{\text{SM}}$ due to out-of-equilibrium positronium formation is a rapidly changing number with respect to the positronium equilibration temperature $T_{\text{Ps}}^{\text{eq}}$. We now need to estimate $T_{\text{Ps}}^{\text{eq}}$ and how fast positronium equilibration can be achieved in the presence of the QED plasma.

Generally speaking, the spectrum and properties of bound states are encoded in the dressed spectral functions in the medium, which evolve dynamically (along with the occupation numbers) and are heavily affected by the surrounding plasma. To rigorously compute these spectral functions, a first-principles approach rooted in the Schwinger-Keldysh (or closed time path) formalism could be used. However, applying this formalism to our system is technically and computationally challenging, as it entails solving a coupled set of integro-differential equations for dynamical two- and four-point correlation functions. Moreover, while steps in this direction have been taken in the non-relativistic and dilute regime in, e.g., reference [33], a precise understanding of how to treat bound states in the relativistic regime remains missing. Therefore, before committing ourselves to overcoming these non-trivial hurdles, we first use physical arguments to assess whether the expected effect on ΔN_{eff} justifies the effort.

Since the evolution is close to equilibrium and the electrons/positrons are mostly non-relativistic, one may (as also discussed in [33]) recast the problem in terms of a finite-temperature effective potential $V(r)$, which in the hard thermal loop (HTL) limit takes the form [33, 34]

$$V(r) = -\alpha m_D - \frac{\alpha}{r} e^{-m_D r} - i\alpha T \phi(m_D r), \quad (3.1)$$

with

$$\phi(y) = 2 \int_0^{+\infty} du \frac{u}{(u^2 + 1)^2} \left(1 - \frac{\sin(uy)}{uy} \right), \quad (3.2)$$

where m_D is the Debye mass, and $\alpha = e^2/(4\pi)$ the electromagnetic fine structure constant. The HTL approximation assumes relativistic electrons ($m_e = 0$) as well as soft momenta ($< \alpha T$) for the electrons, both of which assumptions are evidently inconsistent at $T \lesssim m_e$. Nonetheless, physically it seems sensible to expect the functional form of (3.1) to hold also in the $T \lesssim m_e$ regime, although the Debye mass m_D appearing in the expression may need to be modified from its standard HTL form $m_D^{\text{HTL}} = eT/\sqrt{3}$. Here, our strategy is to substitute m_D with a value extracted from the *full* photon propagator without HTL approximation—to be elaborated on later.

Inspecting the effective potential (3.1), the usual quasiparticle picture applies when the imaginary part is small. In this limit, one may obtain the energies of bound states from $\text{Re}[V(r)]$, while $\text{Im}[V(r)]$ gives their thermal width. This understanding then allows us to factorise the problem of bound-state formation into two separate questions:

- (i) At which temperatures can positronium exist in the early universe plasma?
- (ii) If positronia are allowed to form, how fast do they equilibrate?

We shall examine these two questions in the following subsections.

3.1 Existence of the positronium at high temperatures

It is well known that bound states should generally cease to exist at high temperatures—a phenomenon called “melting”—because of

- (i) Debye screening of the mediator, and
- (ii) Scatterings between the mediator and particles in the plasma.

The two can be roughly associated with the real and imaginary parts of $V(r)$, respectively, in the quasiparticle limit where the latter is small. The precise theoretical calculation of the melting temperature is intricate [34, 35]. Here, we use simple criteria to quantify the impact of these two effects to obtain a first realistic estimate of the positronium melting temperature.

Debye screening of the photon. At $T > 0$ the photon develops a thermal mass from forward-scattering with electrons present in the early universe plasma. Following the original argument by Matsui & Satz [36], i.e., neglecting for now $\text{Im}[V(r)]$, we see immediately that this thermal mass modifies the electrostatic potential (second term in equation (3.1)) to

$$V(r) \sim \frac{e^{-r/a_D}}{r}, \quad (3.3)$$

suppressing it exponentially at large distances $r \gtrsim a_D$, where $a_D \equiv 1/m_D$ is the photon’s Debye screening length. As discussed earlier, while the shape of the effective potential (3.1) has, strictly speaking, been obtained in the HTL limit, a general estimate of a_D applicable also in the non-relativistic (i.e., $T \lesssim m_e$) regime can be extracted from the 00-component of the retarded longitudinal photon self-energy, $\Pi_{L,00}^R$, via

$$a_D = \frac{1}{\sqrt{-\text{Re} \Pi_{L,00}^R(q_0 = 0, |\mathbf{q}| \rightarrow 0)}}, \quad (3.4)$$

and we note that the full 1PI-resummed retarded photon propagator has been computed in reference [8]; see also [37, 38] for different analytical approximations. Equivalently, a_D can be obtained from the QED equation of state via the relation (see, e.g., [39, 40])

$$\text{Re} \Pi_{L,00}^R(q_0 = 0, |\mathbf{q}| \rightarrow 0) = -e^2 \frac{\partial^2 P(\mu, T)}{\partial \mu^2}, \quad (3.5)$$

where $P(\mu, T)$ is the pressure of the $e^+e^-\gamma$ system with chemical potential μ , given in the ideal gas limit by

$$P^{(0)}(\mu, T) = \frac{T}{\pi^2} \int_0^\infty dp p^2 \log \left[\frac{(1 + e^{-(E_e - \mu)/T})^2}{1 + e^{-p/T}} \right], \quad (3.6)$$

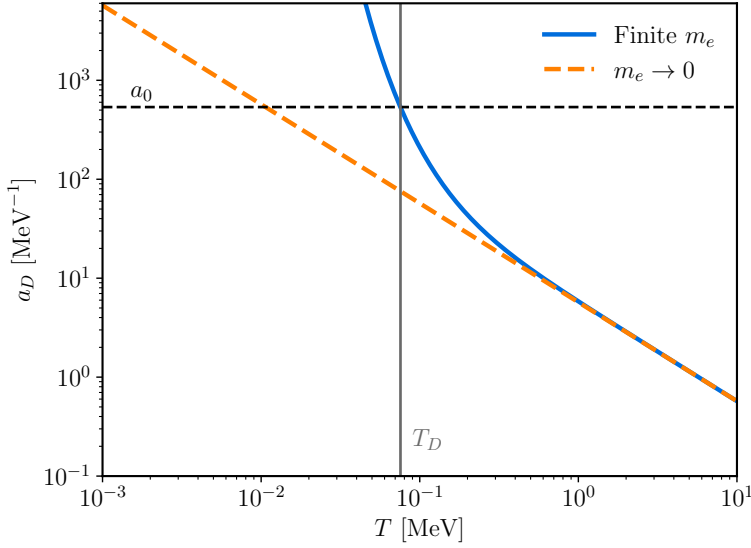


Figure 2. Debye screening length as a function of the temperature computed from the 1PI-resummed photon propagator (blue) and the HTL approximation (orange). Observe that while the two estimates coincide at $T \gtrsim m_e$, the former is significantly larger at $T \lesssim m_e$. We compare a_D to the Bohr radius a_0 of the positronium state in vacuum (black dashed horizontal line). The vertical line represents T_D , the temperature at which $a_0 = a_D$ and Debye screening starts to be effective at $T > T_D$.

with $E_e = \sqrt{p^2 + m_e^2}$. Evaluating (3.5) in the large T limit ($m_e \rightarrow 0$) yields the well-known HTL result $\text{Re} \Pi_{L,00}^{R,\text{HTL}} = -e^2 T^2/3$ [40] and hence $a_D^{\text{HTL}} = \sqrt{3}/(eT)$.

As a first estimate of the positronium melting temperature, we follow the arguments presented in, e.g., [35, 36] and assume that melting occurs once the Debye screening length a_D drops below the most likely radius of the positronium, $r_\ell \approx a_0 n^2$, where

$$a_0 = \frac{1}{m_r \alpha} = \frac{8\pi}{m_e e^2} \simeq 536 \text{ MeV}^{-1} \quad (3.7)$$

is the positronium's Bohr radius, and $m_r = m_e/2$ is the system's reduced mass. Note that a_0 does not depend on the spin (singlet versus triplet) of the positronium configuration, nor is the running of α significant in the temperature range of interest. We do caution however that where Debye screening of the electrostatic potential is effective at high temperatures, the Bohr radius as given in equation (3.7) might not be a fully accurate way to assess the size of the positronium; for a first estimate, we shall neglect this detail.

Figure 2 shows the Debye screening length a_D as a function of temperature, extracted from the 1PI-resummed photon self-energy [8] (or, equivalently, from equation (3.5)), assuming a vanishing chemical potential $\mu = 0$ for the electrons. For completeness, we also plot the HTL approximation a_D^{HTL} . Applying the criterion that bound states only exist when⁶ $a_D > r_\ell$,

⁶For a standard Yukawa potential, numerical analyses [36, 41, 42] of the eigenstates of the Schrödinger equation found a slightly more constraining condition, $a_D \gtrsim a_0/0.84$, for $n = 1$. Applying this condition to our case shifts the melting temperature to 72 keV. We shall however neglect this effect for a first estimate, as the non-trivial momentum dependence of the 1PI-resummed photon propagator might in any case spoil this numerical estimate based on the Yukawa potential.

we see immediately that $n = 1$ positronium formation may be possible at temperatures

$$T \lesssim T_D \simeq 75.6 \text{ keV}, \quad (3.8)$$

while the $n = 2$ state has a melting temperature of 55 keV. Temperatures exceeding the constraint (3.8) are marked in figure 1 by a grey-shaded region; no positronium exists at these temperatures within this approximation. Where positronium formation is possible, the corresponding change in $N_{\text{eff}}^{\text{SM}}$ can be up to $|\Delta N_{\text{eff}}| \simeq 10^{-4}$ assuming instantaneous equilibration (see figure 1). Non-instantaneous equilibration generally results in a larger effect. Using for example the kernel $\sigma(T) = 1 - \exp[(a_D/a_0)^k]$, mimicking a progressive positronium production around T_D , yields $\Delta N_{\text{eff}} = -2.25 \times 10^{-3}, -2.61 \times 10^{-4}$ for $k = 2, 4$ respectively.

Note that had we used the HTL approximation a_D^{HTL} , we would have arrived at a melting temperature of 10.7 keV—much lower than the estimate (3.8) and much too low to impact $N_{\text{eff}}^{\text{SM}}$ in an appreciable way. Physically, the higher temperature (3.8) can be understood as a result of the Boltzmann-suppressed densities of real electrons and positrons at $T \lesssim m_e$ (such suppression is absent in the HTL limit), leading to less screening. However, even substituting the more correct estimate (3.5) of a_D for a_D^{HTL} may still fail to capture the highest temperature (3.8) at which bound states can exist for the following reasons. Firstly, in addition to the Yukawa-like term (3.3), the real part of the full potential (3.1) contains the so-called Salpeter term, $-\alpha m_D$, which directly impacts the bound state’s spectral function. Secondly, the imaginary part of $V(r)$ [34, 43–46], which was neglected in the potential (3.3), implies that the transition to the regime in which bound states can exist is not a sharp one, as discussed in the following. Lastly, the non-trivial momentum dependence of the 1PI-resummed photon propagator could generate sizeable deviations from the potential (3.1), obtained in the HTL limit; see also footnote 6.

Dissociation from scatterings: Beyond Debye screening of the electric field, dissociation can also be driven by the photon’s real scatterings with the rest of the QED bath [34, 35], an effect often referred to as Landau damping. Scattering has even been shown to be the dominant melting mechanism in the case of heavy quarkonia [43, 44, 47]. In terms of the effective potential (3.1), dissociation due to scatterings can be ascribed to $\text{Im}[V(r)]$,⁷ which gives the bound states a finite width. As the temperature increases, this width increases, progressively bridging the gap between the bound state and the continuum. When the thermal width is much larger than the positronium’s binding energy, the bound-state contribution to the spectral function is not well-separated from the continuum, and the positronium is expected to dissociate into an electron-positron pair.

To estimate this thermal width Γ , we use the results from reference [34] and approximate

$$\Gamma = \alpha T \int d^3\mathbf{r} \phi(r/a_D) |\Psi_{1s}^{\text{Ps}}(\mathbf{r})|^2, \quad (3.9)$$

where Ψ_{1s}^{Ps} represents the ground state of positronium, and $\phi(r/a_D)$, with $r = |\mathbf{r}|$, has been defined in equation (3.2). Again, we note that, strictly speaking, both equations (3.2) and (3.9) have been derived using the HTL form of the photon propagator. But, as with our estimate of T_D , we take equation (3.9) at face value and evaluate it using the Debye screening length a_D derived from the full 1PI-resummed photon propagator. Then, as a simple criterion

⁷Generally, the real and imaginary corrections to the photon propagator can be related respectively to the real and imaginary parts of the refractive index, with the latter having its physical origin in scatterings with particles in the medium.

to determine the dissociation temperature T_{dis} , we define it to be the temperature at which the width Γ equals the binding energy ΔE_n , i.e.,

$$\Gamma(T_{\text{dis}}) = \Delta E_n = \frac{m_e \alpha^2}{4n^2}. \quad (3.10)$$

Assuming the positronium wavefunction takes its zero-temperature form (a reasonable approximation when $a_D \gg a_0$ is satisfied), we find from evaluating (3.9) for the $n = 1$ state

$$T_{\text{dis}} \simeq 45.8 \text{ keV}, \quad (3.11)$$

well below the melting temperature $T_D \simeq 75.6 \text{ keV}$ from the Debye screening argument. (Again, we do not consider excited states, as the same argument forbids their production to begin before, e.g., $T \simeq 40.5 \text{ keV}$ for $n = 2$.) The dissociation temperatures of the hydrogen and the muonic hydrogen atoms have been computed in [45, 48] using a similar approach but with a different *ansatz* for the resummed photon propagator; we have checked that our method reproduces the results of [45, 48] to within 10% for these two systems.

The constraint from scatterings is denoted by the hatched region in figure 1. Taken at face value along with the assumption of instantaneous positronium equilibration, such a low formation temperature as indicated by (3.11) would yield a negligible change in $N_{\text{eff}}^{\text{SM}}$ of $\Delta N_{\text{eff}} \simeq -3 \times 10^{-8}$. However, we caution that there are many uncertainties in this calculation at this stage. To begin with, because $T_{\text{dis}} < T_D$, there is a temperature window in which the quasiparticle picture does not hold and we cannot write down an on-shell Boltzmann equation to directly evaluate the positronium equilibration temperature and rate. Nonetheless, dissociation due to scatterings is expected to be a more continuous process arising from the progressive increase of the thermal width Γ with temperature, in contrast to melting from Debye screening where a distinct (though temperature-dependent) mediator mass sets the scale above which bound states cannot exist.⁸ In terms of its impact on ΔN_{eff} , a reasonable expectation is that dissociation will lead to a slower approach to equilibrium in the hatched region than is suggested by the instantaneous equilibration approximation, an effect that may to an extent be captured by a larger $\Delta T/T_{\text{Ps}}^{\text{eq}}$ in our phenomenological modelling of positronium equilibration.⁹ This would in turn drive up ΔN_{eff} from the instantaneous equilibration estimate. Modelling the effect precisely is however beyond the scope of this work, as an accurate description would require that we study the form of the positronium spectral functions, such as in, e.g., [33], which we leave for future work.

3.2 Positronium equilibration rate

Having established an estimate of the upper limit on the positronium formation temperature, namely, $T_D \simeq 80 \text{ keV}$, and supposing that Boltzmann equations provide a valid description

⁸We have reached this conclusion via a very simple understanding of the thermal positronium spectral function, namely, by comparing the separation between the positronium ground state energy in vacuum and the continuum of scattering states with energy $E > 0$, to the estimate (3.9) of the positronium's thermal width. In reality, the positronium spectral function at $T > 0$ is continuous (and non-zero) also for $E < 0$, and the position of the quasiparticle peaks changes with T . While integrals over this spectral function (as they appear in collision terms) tend to be dominated by the pronounced quasiparticle peaks when they exist, sum rules ensure that there are always non-zero contributions from the melted peak for both $E < 0$ and $E > 0$ for $T > 0$. This suggests a gradual increase of the impact of bound states in the T_{dis} -to- T_D window (and likely even at $T > T_D$, implying that even T_D should strictly speaking not be viewed as a sharp upper limit).

⁹Because the quasiparticle picture/standard Boltzmann equations do not apply in the transition region, it is not straightforward to relate ΔT to the microphysical interaction rates.

of positronium production, we now turn to the question of whether this formation can happen quickly enough to produce a sizeable impact on $N_{\text{eff}}^{\text{SM}}$. Properties of positronium have been under scrutiny for decades; see, e.g., [49] and references therein. Positronium-like dark matter bound states have also been the subject of intensive study because of their potential impact on dark matter freeze-out and relic abundance as well as on indirect dark matter searches (e.g., [27, 50–52]). The Boltzmann equations describing the evolution of positronium-like states and the associated production rates in a non-relativistic plasma are very well known in the literature (e.g., [27, 28]).

In the situation under consideration, given that electrons remain in kinetic equilibrium in the temperature range relevant for $N_{\text{eff}}^{\text{SM}}$, the Boltzmann equation for the positronium number density n_{Ps} takes the form

$$\frac{d}{dx}(n_{\text{Ps}}x^3) = -\frac{1}{Hx} [\Gamma^{\text{dec}} + \Gamma^{\text{ion}}(T)] (n_{\text{Ps}}x^3 - n_{\text{Ps}}^{\text{eq}}x^3), \quad (3.12)$$

where $n_{\text{Ps}}^{\text{eq}}$ is the equilibrium positronium number density. In the non-relativistic regime, the ionisation rate Γ^{ion} and the decay rate Γ^{dec} are given by [28]

$$\Gamma^{\text{ion}}(T) = \frac{m_r \alpha^5}{8\pi} \int_0^{+\infty} \frac{d\xi \xi^{-4}}{\exp\left[\left(\frac{1}{n^2} + \frac{1}{\xi^2}\right) \frac{m_e \alpha^2}{4T}\right] - 1} S_{\text{Ps}}^{\text{BSF}}(\xi), \quad (3.13a)$$

$$\Gamma^{\text{dec}} = \begin{cases} m_r \alpha^5, & \text{for } s = 0, \\ \frac{4(\pi^2 - 9)}{9\pi} m_r \alpha^6, & \text{for } s = 1, \end{cases} \quad (3.13b)$$

where $s = 0, 1$ denote, respectively, the singlet and triplet positronium states, and

$$S_{\text{Ps}}^{\text{BSF}}(\xi) = \frac{2^9}{3} \frac{\xi^4 e^{-4\xi \cot^{-1}(1/\xi)}}{(\xi^2 + 1)^2} \frac{2\pi\xi}{1 - e^{-2\pi\xi}}. \quad (3.14)$$

To establish the positronium abundance as a function of temperature T , one needs in principle to solve two Boltzmann equations (3.12) using the rates (3.13) to keep track of the $s = 0$ and $s = 1$ states separately. However, even without an explicit numerical solution, it is straightforward to verify that, at $T \simeq 80$ keV, the decay and ionisation rates are well in excess of the Hubble expansion rate, i.e.,

$$\left. \frac{\Gamma^{\text{dec}} + \Gamma^{\text{ion}}}{H} \right|_{T=T_D} \simeq 10^{16} \gg 1. \quad (3.15)$$

This indicates that, *in the absence of the melting and dissociation effects* discussed in section 3.1, the equilibration of positronium would be instantaneous.

Finally, we remark again that the rates (3.13) apply in the non-relativistic regime only. While this should be a good approximation at $T \lesssim 80$ keV, one may wonder if it overestimates contributions from relativistic electrons and positrons if blindly applied to a thermal distribution. We therefore consider also a worst-case scenario estimate for the rates by (i) neglecting the decay contribution, and, for the ionisation rate, (ii) selecting only electron momenta corresponding to kinetic energies smaller than the positronium binding energy, i.e.,

$$\begin{aligned} \Gamma_{\text{extr}} &= \Gamma^{\text{ion}} \left(\frac{n_e^{\text{eq}}(p < p_{\text{lim}})}{n_e^{\text{eq}}} \right)^2 \\ &= \Gamma^{\text{ion}} \left(\frac{\int_0^{p_{\text{lim}}} p^2 f_{\text{FD}}(p/T, m_e/T) dp}{\int_0^{+\infty} p^2 f_{\text{FD}}(p/T, m_e/T) dp} \right)^2, \end{aligned} \quad (3.16)$$

where p_{lim} is the largest momentum for which the electron-positron relative velocity v_{rel} satisfies $v_{\text{rel}} < \alpha$. But, even with this extremely conservative estimate of the production rate Γ_{extr} , we still find $\Gamma_{\text{extr}}/H \simeq 3 \times 10^2$ at $T \lesssim 80$ keV, indicating again a rapid approach to equilibrium as soon as formation is allowed. We therefore conclude that the formation of positronium is not limited by the production rates, but only by the behaviour of the system between T_D and T_{dis} . Since the overall change in $N_{\text{eff}}^{\text{SM}}$ in this region is exponentially sensitive to the positronium equilibration temperature, further investigations will be necessary to characterise with certainty the ultimate impact of positronium formation on $N_{\text{eff}}^{\text{SM}}$.

4 Positronium as non-ideal gas correction

Up to now we have considered a scenario in which the production of positronium at $T_{\text{Ps}}^{\text{eq}} \lesssim 80$ keV is accompanied by a net increase in the total entropy of the QED plasma at production time. The corresponding correction to $N_{\text{eff}}^{\text{SM}}$ is directly linked to this increase; see, e.g., equation (2.5). Having established in section 3.2, however, that the positronium equilibration rate is in fact very large, it is also useful to consider the opposite limit in which the appearance of bound states at $T_{\text{Ps}}^{\text{eq}}$ does not generate entropy in the QED sector *overall*. That is, the QED sector remains in equilibrium within itself at all times after neutrino decoupling. In this limit, any changes to $N_{\text{eff}}^{\text{SM}}$ due to bound states will depend solely on their contribution to the entropy of the QED plasma *at neutrino decoupling*, which, in the instantaneous-decoupling approximation, can be established from non-ideal corrections to the QED equation of state at $T = T_d \simeq 1.3$ MeV.

Given that $T_d \gg T_{\text{Ps}}^{\text{eq}}$, one might naively conclude that the absence of bound states at $T \simeq T_d$ must incur a negligible effect on $N_{\text{eff}}^{\text{SM}}$ in this scenario. This simplified picture may however be misleading for the following reason. The same long-range electron-positron interaction that gives rise to bound states at low temperatures also implies that resonant interactions amongst the scattering-state electrons and positrons are possible prior to bound-state formation. These non-perturbative effects alter the QED equation of state (EoS) away from the ideal-gas behaviour and could persist at some level even at much higher temperatures, e.g., around $T = T_d$, thereby impacting on $N_{\text{eff}}^{\text{SM}}$. Indeed, as we shall show below, in this entropy-conserving scenario, the entropy residing in bound states—once they are allowed to form—can be envisaged as having been “stolen” from the resonant e^+e^- -scattering states, whose own entropy decreases accordingly.

To estimate ΔN_{eff} in this scenario quantitatively, we note first of all that no fully non-perturbative EoS formula exists in the *relativistic* and non-dilute regime, i.e., $T \gg m_e, \mu_e$, with μ_e the electron chemical potential. However, the *non-relativistic* limit $T \ll m_e$ is well described by the so-called Beth-Uhlenbeck formula [53]. This formula consistently incorporates both bound and interacting scattering states in the non-ideal pressure. Importantly, as a fully non-perturbative formula, Beth-Uhlenbeck contains also perturbative contributions to the QED EoS in the non-relativistic limit that, in the context of $N_{\text{eff}}^{\text{SM}}$, have already been accounted for elsewhere (e.g., [4]), and care needs to be taken to avoid double counting.

Our strategy in this section, therefore, is as follows. We first isolate in section 4.1 the leading-order perturbative correction to the EoS from a non-relativistic partition function that forms the basis of the Beth-Uhlenbeck formula. Comparison of this leading correction with the fully non-perturbative formula in section 4.2 then allows us to demonstrate that the system remains perturbative even at the time of bound-state formation and hence provides a means for us to consistently extract the higher-order corrections—including non-perturbative

corrections from bound states—contained in the formula. Non-technical readers wishing to skip these details may jump directly to section 4.3, where we extrapolate the extraction approach to $T = T_d$ to obtain an estimate of (or, more correctly, an upper limit on) ΔN_{eff} due to bound-state formation in this entropy-conserving scenario.

4.1 Perturbative computation of non-ideal corrections

The goal of sections 4.1 and 4.2 is to demonstrate that perturbation theory can be applied to the system at hand in a controlled way, even at times when bound states begin to form. To this end, our first task is to determine the leading-order non-ideal correction in the coupling e to the QED EoS in the non-relativistic limit. We begin by noting for future reference that, in equilibrium, the energy density ρ and the entropy s are related to the pressure density via

$$P = \frac{T}{\mathcal{V}} \ln Z, \quad \rho = \frac{T^2}{\mathcal{V}} \partial_T \ln Z = -P + T \partial_T P, \quad s = \frac{1}{\mathcal{V}} \partial_T (T \ln Z) = \frac{\rho + P}{T}, \quad (4.1)$$

where $\mathcal{V}$ is the spatial volume, $Z = \text{Tr}[e^{-\beta H}]$ the partition with Hamiltonian $H = H_0 + H_{\text{int}}$, $\beta = 1/T$, and we have set the chemical potentials to zero here for simplicity.

In *relativistic* QED, the standard textbook procedure is to treat interaction term $\mathcal{L}_{\text{int}} = -e\bar{\psi}\gamma^\mu\psi A_\mu$ as a perturbation [39], and expand the partition function in powers of the QED coupling e

$$\ln Z = \ln Z^{(0)} + \ln Z^{(2)} + \dots, \quad (4.2)$$

where $\ln Z^{(2)}$ constitutes the first, i.e., $\mathcal{O}(e^2)$, non-ideal correction. The corresponding $\mathcal{O}(e^2)$ correction to the pressure, $P^{(2)}$, can be written as a sum over three parts, $P^{(2)} = P^{(2a)} + P^{(2b)} + P^{(2c)}$, where the individual contribution reads

$$P^{(2a)} = \frac{e^2 m_e^2}{4\pi^4} \int \int dp d\tilde{p} \frac{p\tilde{p}}{E_p E_{\tilde{p}}} \ln \left| \frac{p+\tilde{p}}{p-\tilde{p}} \right| f_F(E_p) f_F(E_{\tilde{p}}) \stackrel{\text{NR}}{\simeq} \frac{e^2 m_e^2 T^2}{8\pi^3} e^{-\beta 2m_e}, \quad (4.3a)$$

$$P^{(2b)} = -\frac{e^2}{2\pi^4} \left(\int dp \frac{p^2}{E_p} f_F(E_p) \right)^2 \stackrel{\text{NR}}{\simeq} -\frac{e^2 m_e T^3}{4\pi^3} e^{-\beta 2m_e}, \quad (4.3b)$$

$$P^{(2c)} = -\frac{e^2 T^2}{6\pi^2} \int dp \frac{p^2}{E_p} f_F(E_p) \stackrel{\text{NR}}{\simeq} -\frac{e^2}{3} \left(\frac{m_e T^7}{(2\pi)^3} \right)^{1/2} e^{-\beta m_e}, \quad (4.3c)$$

with $f_F(E) = 1/(e^{E/T} + 1)$ denoting the standard Fermi-Dirac distribution. For later reference, we have evaluated each contribution in the non-relativistic limit; in this limit, $P^{(2a)}$ is the leading correction in the squared number density $n_e^2 \sim m_e^3 T^3 e^{-2\beta m_e}$, while $P^{(2c)}$ is linear in n_e .

The pressure correction in the non-relativistic limit can also be obtained directly from the non-relativistic interaction Hamiltonian

$$H_{\text{NR}}^{\text{int}} = -\frac{1}{2} \sum_{\sigma, \sigma'} \int d^3\mathbf{x} d^3\mathbf{y} V(|\mathbf{x} - \mathbf{y}|) \left[\eta_\sigma^\dagger(\mathbf{x}) \eta_{\sigma'}^\dagger(\mathbf{y}) \eta_{\sigma'}(\mathbf{y}) \eta_\sigma(\mathbf{x}) \right. \\ \left. + \xi_\sigma(\mathbf{x}) \xi_{\sigma'}(\mathbf{y}) \xi_{\sigma'}^\dagger(\mathbf{y}) \xi_\sigma^\dagger(\mathbf{x}) \right. \\ \left. - 2\eta_\sigma^\dagger(\mathbf{x}) \xi_{\sigma'}(\mathbf{y}) \xi_{\sigma'}^\dagger(\mathbf{y}) \eta_\sigma(\mathbf{x}) \right], \quad (4.4)$$

where V is the potential, the field operator $\eta^\dagger$ creates an electron and ξ a positron, and summation is over the spin indices σ, σ' . Performing a real-time computation yields an $\mathcal{O}(e^2)$ non-relativistic correction to the pressure given by

$$P_{\text{NR}}^{(2)} = \frac{T}{\mathcal{V}} \ln Z_{\text{NR}}^{(2)} = \frac{1}{\mathcal{V}} \frac{1}{2} \sum_{\sigma\sigma'} \int d^3\mathbf{x} d^3\mathbf{y} V(|\mathbf{x} - \mathbf{y}|) \left[G_{\eta\eta,0}^{++--}(t\mathbf{x}, \mathbf{y}, t\mathbf{x}, \mathbf{y}) + G_{\xi\xi,0}^{++--} - 2G_{\eta\xi,0}^{++--} \right], \quad (4.5)$$

where the four-point equal-time correlators are in normal order according to (4.4), e.g.,

$$G_{\eta\eta,0}^{++--}(t\mathbf{x}, \mathbf{y}, t\mathbf{x}, \mathbf{y}) \equiv \langle \eta_\sigma^\dagger(\mathbf{x}) \eta_{\sigma'}^\dagger(\mathbf{y}) \eta_{\sigma'}(\mathbf{y}) \eta_\sigma(\mathbf{x}) \rangle_0, \quad (4.6)$$

and $\langle \dots \rangle_0$ denotes a thermal average with respect to the free Hamiltonian H_0 .

Using the equations of motion (B.8a) and (B.8b) in integral form, the four-point correlators can be expressed in terms of free two-point functions as

$$\begin{aligned} G_{\eta\eta,0}^{++--}(t\mathbf{x}, \mathbf{y}, t\mathbf{x}, \mathbf{y}) &= G_{\eta,0}^{+-}(t\mathbf{x}, t\mathbf{x}) G_{\eta,0}^{+-}(t\mathbf{y}, t\mathbf{y}) - G_{\eta,0}^{+-}(t\mathbf{x}, t\mathbf{y}) G_{\eta,0}^{+-}(t\mathbf{y}, t\mathbf{x}), \\ G_{\xi\xi,0}^{++--}(t\mathbf{x}, \mathbf{y}, t\mathbf{x}, \mathbf{y}) &= G_{\xi,0}^{+-}(t\mathbf{x}, t\mathbf{x}) G_{\xi,0}^{+-}(t\mathbf{y}, t\mathbf{y}) - G_{\xi,0}^{+-}(t\mathbf{x}, t\mathbf{y}) G_{\xi,0}^{+-}(t\mathbf{y}, t\mathbf{x}), \\ G_{\eta\xi,0}^{++--}(t\mathbf{x}, \mathbf{y}, t\mathbf{x}, \mathbf{y}) &= G_{\eta,0}^{+-}(t\mathbf{x}, t\mathbf{x}) G_{\xi,0}^{+-}(t\mathbf{y}, t\mathbf{y}), \end{aligned} \quad (4.7)$$

where note that the degenerate cases (i.e., the $\eta\eta$ and $\xi\xi$ correlators) have two contributions each. Inserting (4.7) into (4.5), figure 3 shows the diagrammatic representation of the $\mathcal{O}(e^2)$ non-relativistic correction to the pressure. The “dumbbell” diagrams correspond to the Hartree corrections, which cancel in a symmetric plasma, i.e., where the electron and positron number densities are equal, such that $G_{\eta,0}^{+-}(t\mathbf{x}, t\mathbf{x}) = G_{\xi,0}^{+-}(t\mathbf{x}, t\mathbf{x})$. The remainder is the exchange or “oyster” diagram, also known as the Hartree-Fock correction, which evaluates using Wigner coordinates to

$$\begin{aligned} P_{\text{NR}}^{(2)} &= (-2) \times \frac{1}{2} \sum_{\sigma\sigma'} \int d^3\mathbf{r} V(r) G_{\eta,0}^{+-}(0, \mathbf{r}) G_{\eta,0}^{+-}(0, -\mathbf{r}) \\ &= -2 \int \frac{d^3\mathbf{p}_1}{(2\pi)^3} \frac{d^3\mathbf{p}_2}{(2\pi)^3} V(\mathbf{p}_1 - \mathbf{p}_2) e^{-\beta(2m_e + \mathbf{p}_1^2/(2m_e) + \mathbf{p}_2^2/(2m_e))} \\ &= \frac{e^2 m_e^2 T^2}{8\pi^3} e^{-\beta 2m_e}. \end{aligned} \quad (4.8)$$

Here, we have used the real-time Feynman rules in momentum space for the free non-relativistic two-point functions in equilibrium provided in appendix C, and the Coulomb potential in momentum space is $V(\mathbf{q}) = -e^2/\mathbf{q}^2$. Comparing this result (4.8) with equation (4.3), we see immediately that $P_{\text{NR}}^{(2)}$ coincides with $P^{(2a)}$ in the non-relativistic limit, while the sub-leading term $P^{(2b)}$ and the correction $P^{(2c)}$ linear in $n_e \sim e^{-\beta m_e}$ are absent.

Pressure corrections proportional to $n_e^2 \sim e^{-\beta 2m_e}$ correspond to the *second* virial coefficient in a virial expansion, and the normal-ordered Hamiltonian (4.4) is commonly used in the literature [54, 55] to compute this quantity. Indeed, for our purpose of estimating the effects of bound states on the QED EoS, the second virial coefficient is the quantity of interest, as it accounts for bound states when computed non-perturbatively. These considerations lead us to conclude that, for a fixed-order perturbation theory in e to be under control, higher-order/non-perturbative corrections (due to, e.g., bound states) to P_{NR} must remain small relative to $P_{\text{NR}}^{(2)}$ computed perturbatively in equation (4.8). We turn to the non-perturbative evaluation of P_{NR} below.

$$\left[\begin{array}{c} \text{---} \circ^{e^+} \\ | \\ \text{---} \circ^{e^+} \end{array} - \text{---} \circ^{e^+} \text{---} \circ^{e^+} \right] + \left[\begin{array}{c} \text{---} \circ^{e^-} \\ | \\ \text{---} \circ^{e^-} \end{array} - \text{---} \circ^{e^-} \text{---} \circ^{e^-} \right] - 2 \begin{array}{c} \text{---} \circ^{e^+} \\ | \\ \text{---} \circ^{e^-} \end{array} = -2 \text{---} \circ^{e^\pm} \text{---} \circ^{e^\pm}$$

Figure 3. Feynman diagrammatic representation of equation (4.5). In a symmetric plasma, the “dumbbells” diagrams cancel, leaving us with the “oyster” diagram to the right of the equal sign.

4.2 Beth-Uhlenbeck formula

The Beth-Uhlenbeck formula [53] is a non-perturbative expression for the non-relativistic pressure that accounts for both bound and interacting scattering states, and can be derived from the partition function of the non-relativistic Hamiltonian (4.4); see, e.g., [54, 55] and appendix B. The formula takes the form

$$\beta(P_{\text{NR}} - P_{\text{NR}}^{(0)}) = 4 \left(\frac{2m_e T}{2\pi} \right)^{3/2} e^{-\beta 2m_e} \sum_{\ell=0}^{\infty} (2\ell+1) b_\ell, \quad (4.9)$$

with

$$b_\ell = \sum_{n \geq \ell+1} e^{-\beta E_{n\ell}} + \frac{1}{\pi} \int_0^\infty dE e^{-\beta E} \frac{d}{dE} \left\{ \delta_\ell^{e^+e^-} + \left(\frac{1}{2} - \frac{(-1)^\ell}{4} \right) (\delta_\ell^{e^+e^+} + \delta_\ell^{e^-e^-}) \right\}, \quad (4.10)$$

where $E_{n\ell}$ is the binding energy of the positronium bound state with principle quantum number n and angular momentum ℓ , and δ_ℓ denotes the quantum mechanical phase shift of the scattering states, with $\delta_\ell^{e^+e^+} = \delta_\ell^{e^-e^-}$ for the repulsive contributions.

We proceed by first applying perturbation theory to (4.9) to identify the leading correction (4.8). In the Born approximation of (4.10), the first, bound-state term is zero, while the phases in the integral are given by

$$\delta_{\ell, \text{Born}}^{e^+e^-/e^+e^+} = \mp \frac{m_e}{k} \int_0^\infty dr V(r) [kr j_\ell(kr)]^2, \quad (4.11)$$

with $k = \sqrt{m_e E}$ and $j_\ell(kr)$ is a spherical Bessel function of degree ℓ . Observe that the phases $\delta_{\ell, \text{Born}}^{e^+e^-}$ and $\delta_{\ell, \text{Born}}^{e^+e^+}$ differ only by a sign. This means that $\delta_{\ell, \text{Born}}^{e^+e^-}$ in equation (4.10) cancels entirely against $(\delta_{\ell, \text{Born}}^{e^+e^+} + \delta_{\ell, \text{Born}}^{e^-e^-})/2$, resulting in

$$\sum_{\ell=0}^{\infty} (2\ell+1) b_\ell^{\text{Born}} = -\frac{1}{2} \frac{1}{\pi} \int_0^\infty dE e^{-\beta E} \frac{d}{dE} \underbrace{\sum_{\ell=0}^{\infty} (2\ell+1) (-1)^\ell \delta_{\ell, \text{Born}}^{e^+e^+}}_{=k f_{\text{Born}}^{e^+e^+}(\theta=\pm\pi)}, \quad (4.12)$$

where the sum over all partial waves can be identified with the scattering amplitude $f_{\text{Born}}^{e^+e^+}(\theta = \pm\pi)$. For a Yukawa potential with screening mass m_D , the scattering amplitude is given by

$$f_{\text{Born, Yukawa}}^{e^+e^-/e^+e^+}(\theta) = \pm \frac{\alpha m_e}{m_D^2 + 4k^2 \sin^2(\theta/2)}, \quad (4.13)$$

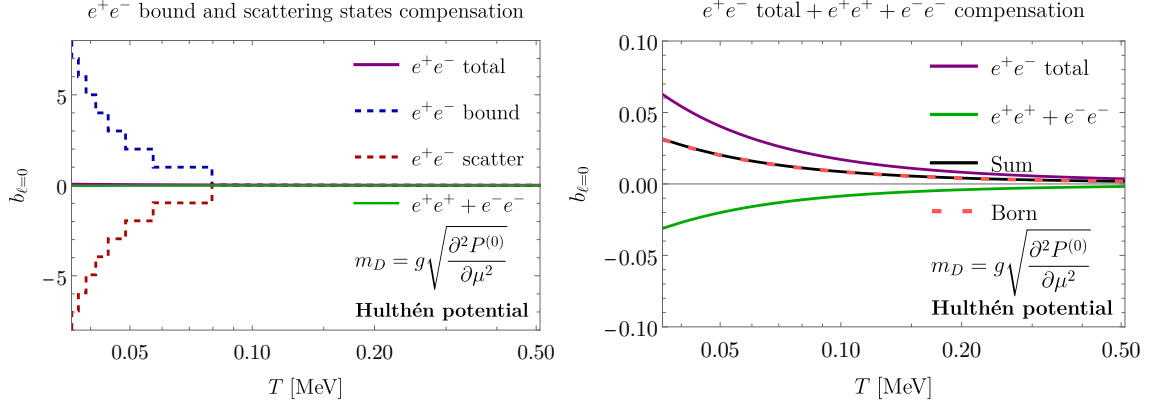


Figure 4. Breakdown of the various contributions to the coefficient $b_{\ell=0}$ in equation (4.10) as a function of temperature.

and is not divergent at $\theta = \pm\pi$ in the Coulomb (i.e., $m_D \rightarrow 0$) limit. Then, the energy integral in (4.12) can be performed analytically. Taking the Coulomb limit at the end of the integration yields

$$\left(P_{\text{NR}} - P_{\text{NR}}^{(0)}\right)|_{\text{Born}} = \frac{e^2 m_e^2 T^2}{8\pi^3} e^{-\beta 2m_e}, \quad (4.14)$$

consistent with perturbative correction (4.8). Thus, the Beth-Uhlenbeck formula (4.9) reproduces the leading, $\mathcal{O}(e^2)$ result.

Next, we turn to a non-perturbative evaluation of (4.10). Analytic expressions of the the binding energy and non-perturbative phase shifts do not exist for the screened Yukawa potential. We therefore approximate the Yukawa potential by the Hulthén potential,

$$V_H(r) = -\frac{\alpha\kappa m_D e^{-\kappa m_D r}}{1 - e^{-\kappa m_D r}}, \quad (4.15)$$

where the choice of the parameter $\kappa = \pi^2/6$ is known to approximate the Yukawa potential particularly well in the vicinity of the first resonance (i.e., the appearance of the first $n = 1, \ell = 0$ zero-energy bound state) [56]. Even so, exact analytic expressions exist only for the s -wave¹⁰ (e.g., [58, 59]),

$$E_{n,\ell=0} = -\frac{\alpha^2 m_e}{4n^2} (1 - \kappa\epsilon_\phi n^2)^2, \quad \text{for } m_D \leq \frac{\alpha m_e}{\kappa n^2}, \quad (4.16)$$

and

$$\delta_{\ell=0} = \text{Arg} \frac{i\Gamma(\lambda_+ + \lambda_- - 2)}{\Gamma(\lambda_+)\Gamma(\lambda_-)}, \quad \text{where } \lambda_\pm = \begin{cases} 1 + i \frac{\epsilon_v}{\kappa\epsilon_\phi} \pm \frac{\sqrt{\kappa\epsilon_\phi - \epsilon_v^2}}{\kappa\epsilon_\phi}, & \text{for } e^+e^- \\ 1 + i \frac{\epsilon_v}{\kappa\epsilon_\phi} \pm i \frac{\sqrt{\kappa\epsilon_\phi + \epsilon_v^2}}{\kappa\epsilon_\phi}, & \text{for } e^+e^+, e^-e^- \end{cases} \quad (4.17)$$

with dimensionless quantities $\epsilon_v^2 = E/(\alpha^2 m_e)$ and $\epsilon_\phi = m_D/(\alpha m_e)$.

Using equations (4.16) and (4.17), we evaluate $b_{\ell=0}$ in (4.10) semi-analytically for the Hulthén potential. The results are shown in figure 4 for temperatures close to but less than

¹⁰For higher partial waves, approximations of the centrifugal term are necessary to obtain analytic expressions. These exist, but are rather inaccurate because of quasi-bound states for $\ell > 0$ [57].

the electron mass. In the left panel, the blue dashed line denotes the contribution from bound states, which exhibits a discrete jump every time a bound state is permitted to exist. Indeed, as the screening mass m_D decreases exponentially with decreasing temperature, the number of permissible bound states also increases rapidly. Similar jumps can be seen in the attractive scattering-state contributions—represented in figure 4 by the magenta dashed line—but in the opposite direction. Most interestingly, these discontinuities compensate precisely with those in the bound-state contribution, such that $b_{\ell=0}$ and hence the pressure and all other thermodynamic quantities in fact remain *continuous*. The right panel zooms in on the sum of the e^+e^- -scattering and the bound-state contributions (purple solid line), which is indeed continuous and positive-valued. In the high temperature limit (i.e., $\beta \rightarrow 0$), the e^+e^- bound- and scattering-state contributions cancel each other,

$$\lim_{\beta \rightarrow 0} b_\ell = n_\ell + \frac{1}{\pi} [\delta_\ell^{e^+e^-}(E = \infty) - \delta_\ell^{e^+e^-}(E = 0)] = 0, \quad (4.18)$$

where n_ℓ is the number of permissible bound states for a fixed ℓ , and we have used Levinson’s theorem.

Observe also the green solid line in the right panel of figure 4. This denotes the total contribution from the repulsive scattering states and runs negative, such that when summing up all contributions to $b_{\ell=0}$ in (4.10), there occurs another cancellation between the total attractive and total repulsive components, represented by the black solid line. This line is very close to the Born approximation computed using equation (4.11) for the Hulthén potential (4.15) indicated by the red dashed line. We have checked that $b_{\ell=0}^{\text{Born}}$ for the Hulthén potential differs from the screened Yukawa case (4.13) by at most 3% in the temperature range shown in figure 4. As the latter agrees with the $\mathcal{O}(e^2)$ perturbative computation of the pressure correction (see (4.8) versus (4.14)), we therefore conclude that perturbation theory is indeed under control if the potential is screened.

4.3 Impact on N_{eff}

Having now established that bound-state effects on the QED EoS is perturbatively small even at $T \ll m_e$, we are now in a position to estimate ΔN_{eff} . We adopt the usual approach of demanding comoving entropy conservation (e.g., [4]; see also equation (2.5) and accompanying text), and estimate the change in the QED entropy,

$$\delta s = \partial_T (P_{\text{NR}} - P_{\text{NR}}^{(0)}), \quad (4.19)$$

from a non-ideal pressure correction $P_{\text{NR}} - P_{\text{NR}}^{(0)}$ using the non-perturbative Beth-Uhlenbeck formula (4.9) for the Hulthén potential (4.15) and only s -wave contributions.

Figure 5 show $\delta s/T^3$ as a function of T . The left panel displays a breakdown of the contributions from various components: attractive e^+e^- contributions from bound states (blue dashed) and scattering states (magenta dashed), as well as the total repulsive contribution from e^+e^+ and e^-e^- (green solid). The bound-state contribution appears at $T \simeq 80$ keV with a peak value of $\delta s/T^3 \sim 0.0005$, corresponding to the entropy of one Maxwellian-distributed species of $g_{\text{Ps}} = 4$ internal degrees of freedom. Analogously to figure 4, the contribution from attractive scattering states is almost as large in magnitude as their bound-state counterpart, but carries the opposite sign. These two attractive contributions compensate each other almost exactly in summation (purple solid). Furthermore, the total attractive contribution itself cancels to a good extent with the total repulsive contribution, leading to the sum

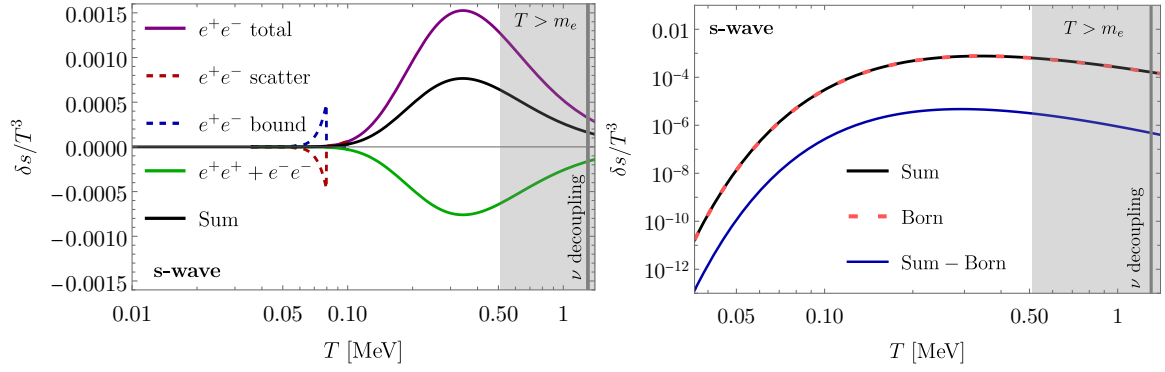


Figure 5. Change to the entropy, δs , as a function of temperature T . *Left:* Breakdown of the contributions from various components. *Right:* The sum of all contributions as derived from a full non-perturbative computation (black solid) and from the Born approximation (red dashed). The difference between the non-perturbative and Born results (blue solid) remains substantially below the individual result, suggesting that perturbation theory is under control.

total represented by the black solid line. As in figure 4, this total fully-non-perturbative δs is remarkably close to what we would obtain from (4.9) in the Born approximation (red dashed). Indeed, as shown on the right panel of figure 5, subtracting the Born result from the total δs , the remainder—in blue, representing higher-order contributions—is two orders of magnitude smaller than the Born result.

To estimate ΔN_{eff} due to bound states from these results, we note first of all that the $\mathcal{O}(e^2)$ correction to the QED EoS and its impact on $N_{\text{eff}}^{\text{SM}}$ have previously been considered in, e.g., [4]. As we have established in sections 4.1 and 4.2, in the non-relativistic limit, this correction coincides with the Born approximation of the Beth-Uhlenbeck formula. Therefore, to avoid double counting, the true contribution of bound states to $N_{\text{eff}}^{\text{SM}}$ needs to be estimated from the fully non-perturbative Beth-Uhlenbeck formula *minus* the Born piece—that is, the blue line in figure 5.

Secondly, following entropy conservation arguments, changes to $N_{\text{eff}}^{\text{SM}}$ in our scenario are ultimately linked to corrections to the QED entropy at neutrino decoupling, i.e.,

$$\Delta N_{\text{eff}} \simeq -4 \left. \frac{\delta s}{s} \right|_{T=T_d} \simeq -\frac{180}{11\pi^2} \left. \frac{\delta s}{T^3} \right|_{T=T_d} \quad (4.20)$$

where we have approximated $s \simeq (11\pi^2/45)T^3$ at the second equality, assuming massless electrons and positrons. Given that $T_d \simeq 1.3 \text{ MeV} > m_e$, evaluation of (4.20) for bound state effects necessitates that we extrapolate to the relativistic regime outside of the validity region of the Beth-Uhlenbeck formula. Reading $\delta s/T^3$ off figure 5 at $T = T_d$, we find $\Delta N_{\text{eff}} \sim -10^{-6}$ —similar in magnitude to the relativistic $\mathcal{O}(e^4)$ correction identified in [4] but with the opposite sign. We have also checked numerically that implementing the same EoS correction into the continuity equation (2.12) and solving it alongside a momentum-averaged Boltzmann equation (for the neutrinos) in the damping approximation yields a similarly small $\Delta N_{\text{eff}} \sim -6 \times 10^{-7}$.¹¹

¹¹This EoS correction can be straightforwardly implemented by making the following replacements in equation (2.12): $G_1 \rightarrow G_1 - (\tau^3/(2m_e^4)) [3(\bar{\rho}^{(4)} + \bar{P}^{(4)} + \tau \partial_\tau \bar{\rho}^{(4)})]$, and $G_2 \rightarrow G_2 - (\tau^4/(2m_e^4)) \tau \partial_\tau \bar{\rho}^{(4)}$. Here, the rescaled energy density and pressure, $\bar{\rho}^{(4)}$ and $\bar{P}^{(4)}$, refer to the contribution that results after summing over all bound, scattering, and repulsive (i.e., $e^\pm e^\pm$) contributions and subtracting from it the already-known, per-

Lastly, we note that for a perturbation theory that is under control, it is in principle possible to further subtract from the blue line in figure 5 the $\mathcal{O}(e^4)$ contribution, potentially yielding an even smaller $\delta s/T^3$ and hence ΔN_{eff} due to bound states. In this sense, our estimate of $\Delta N_{\text{eff}} \sim -10^{-6}$ above serves merely as an upper limit on the possible effect of bound states in the equilibrium scenario. However, to perform such a subtraction would entail a consistent derivation of the non-relativistic limit of the relativistic QED pressure, as well as an expansion of the Beth-Uhlenbeck formula to higher orders, both of which are beyond the scope of the present work. We also leave for future work a more complete and self-consistent determination of the full range of corrections due to bound states, including thermal width and the Salpeter correction, Bose-enhancement factors in the non-dilute case close to $T \lesssim m_e$, as well as summation over partial waves, all of which are expected to be necessary in the temperature range of interest.

5 Conclusions and outlook

We have presented in this work a first assessment of how positronium states could affect the Standard Model prediction of the effective number of neutrinos, $N_{\text{eff}}^{\text{SM}}$. We anticipate positronium to impact on $N_{\text{eff}}^{\text{SM}}$ primarily through modifications to the QED EoS. Their impact via changes to the weak and electromagnetic particle scattering rates (e.g., enhanced annihilation cross section) is expected to be negligible.

A key challenge in quantifying the role of positronium in the QED EoS lies in the fact that the spectrum of bound states changes dynamically over the pertinent temperature regime because of in-medium effects such as screening and scatterings. This does not only shift the bound-state energies, but also dresses the states with thermal widths which make the transition to the continuum of scattering states gradual. To our knowledge, there is currently no known controlled approximation scheme that can reliably describe all potentially relevant effects across the entire temperature range of interest, covering the transition from the relativistic to the non-relativistic regime. In particular, we do not know to what degree the emergence of bound states implies a deviation from equilibrium in the early universe.

To systematically assess what different assumptions about equilibrium imply for $N_{\text{eff}}^{\text{SM}}$, we adopt a simple model of the medium by modifying the Coulomb potential to a Yukawa potential via the introduction of a Debye mass m_D for the photon. In this approximation positronia can be treated as on-shell particles for which one can, e.g., unambiguously write down a set of Boltzmann equations to track their evolution in the early universe. We can then conceive of two possible positronium formation scenarios which differ mainly in their treatment of entropy evolution.

Out-of-equilibrium scenario We consider first the viewpoint that positronium appears instantaneously in the spectrum of quantum states once the temperature drops to a certain value $T = T_{\text{Ps}}^{\text{eq}}$. The blue line in figure 1 shows the expected change in $N_{\text{eff}}^{\text{SM}}$, ΔN_{eff} , under the assumption that the occupation number is then instantaneously driven towards equilibrium, as suggested by the large rate found in section 3.2. The sudden transition from vanishing to equilibrium positronium occupation numbers at a single moment represents a sizeable deviation from thermal equilibrium and generates significant entropy. Clearly, the corresponding ΔN_{eff} is exponentially dependent on $T_{\text{Ps}}^{\text{eq}}$ due to Boltzmann suppression of the positronium

turbative $\mathcal{O}(e^2)$ correction. For the neutrinos, the momentum-averaged Boltzmann equation in the damping approximation can be found in [8].

number density. Identifying $T_{\text{Ps}}^{\text{eq}}$ with the highest temperature at which the real Yukawa potential can host a bound state, $T \simeq 75.6$ keV, would lead us to conclude that $|\Delta N_{\text{eff}}| \lesssim 10^{-4}$, below the uncertainties of current precision $N_{\text{eff}}^{\text{SM}}$ computations.

It is however possible that Debye screening is overcome in a more gradual manner, e.g., because of the finite thermal width of the quantum state. We mimic the resulting effect on ΔN_{eff} by implementing a large transition parameter $\Delta T/T_{\text{Ps}}^{\text{eq}}$ symmetrically around $T_{\text{Ps}}^{\text{eq}}$ in our phenomenological modelling. Because this effectively creates a small population of positronium at $T > T_{\text{Ps}}^{\text{eq}}$, a large transition parameter tends to drive up entropy production and hence ΔN_{eff} as indicated, for example, by the orange line in figure 1.

Equilibrium scenario Here, we consider the case in which the QED sector *overall* is assumed to remain in thermal equilibrium at all times, even when crossing the existence threshold temperature of a positronium state. In this limit, the primary contribution of positronium states to $N_{\text{eff}}^{\text{SM}}$ should arise through non-perturbative corrections to the QED EoS around neutrino decoupling, $T_d \simeq 1$ MeV, i.e., a significantly higher temperature than the nominally highest temperature, $T \simeq 75.6$ keV, at which a bound state can exist.

To investigate this scenario, we first use the Beth-Uhlenbeck formula, a non-relativistic but non-perturbative expression for the pressure that accounts for photon-mediated interactions (“ladder diagrams”) to all orders in all two-particle bound and scattering states, to study thermodynamic quantities near the threshold temperature of the first s -wave positronium states. Figure 5 shows that, while the bound-state and the attractive scattering-state contributions to the entropy each features a discontinuous jump at $T \simeq 75.6$ keV, their sum (as well as the total including repulsive contributions) varies smoothly across this threshold temperature. This smoothness affirms that the appearance of bound states at low temperatures *need not* be accompanied by a sudden change in the overall QED entropy (although we emphasise that it *does not preclude* overall entropy production if bound state formation turns out to be disruptive, as in the out-of-equilibrium scenario).

Comparing the Born limit of the Beth-Uhlenbeck formula with the leading-order non-ideal gas correction computed using relativistic finite-temperature QED, we find close agreement at $O(e^2)$ in the non-relativistic regime, even at the formation threshold of the first s -wave positronium states (when summed over *all* two-body states). In particular, the Born approximation captures the effects of scattering states well, as shown in the right panel of figure 5, indicating that perturbation theory can be trusted. This latter insight is important, as it supports the validity of the perturbative approach used in previous studies of non-ideal corrections to the QED EoS. Subtracting the Born approximation from the Beth-Uhlenbeck formula and extrapolating to the neutrino decoupling temperature, we find that in the equilibrium scenario the change in $N_{\text{eff}}^{\text{SM}}$ from the s -wave contribution should be limited to $|\Delta N_{\text{eff}}| \lesssim 10^{-6}$. In view of the proximity of this number in magnitude to the $\mathcal{O}(e^4)$ correction found in [4], it is probable that the result of [4] already subsumes the relevant effects.

Importantly, the discussion above—in both the out-of-equilibrium and the equilibrium scenario—ignores modifications to the bound-state properties due to scattering with the plasma constituents, such as Landau damping. In the static limit, such effects can be captured by an imaginary component in the effective potential (3.1). Assuming that the sole effect of the imaginary potential is to push the bound-state formation threshold down to an even lower temperature, the out-of-equilibrium model for instantaneous production suggests $|\Delta N_{\text{eff}}| \simeq 3 \times 10^{-8}$, a number much smaller than the hard Debye cut-off estimate, $|\Delta N_{\text{eff}}| \simeq 10^{-4}$,

discussed above. In the equilibrium scenario, a self-consistent Beth-Uhlenbeck-like formulation that includes Landau damping (as well as the Salpeter correction) does not yet exist: we leave its exploration for future work.

In summary, a minute and transient population of positronium in dynamical equilibrium with the rest of the QED plasma could in principle have existed in roughly the same time frame during which the Standard-Model value of N_{eff} is set. Uncertainties in relation to the dynamical evolution of bound-state properties and interaction rates in a QED plasma in the temperature range $\mathcal{O}(10)$ keV $\lesssim T \lesssim \mathcal{O}(100)$ keV prevent us from pinning down its exact contribution to $N_{\text{eff}}^{\text{SM}}$. However, our current estimates suggest that the effect is likely to be 10^{-4} or smaller, i.e., not larger than other theoretical uncertainties, and an order of magnitude below the anticipated experimental sensitivity of CMB-S4.

Acknowledgements

We are grateful to Georg Raffelt for triggering this research with an insightful question about the role of positronium. We also thank Iason Baldes, Lorenzo de Ros, Mathias Garny, Jacopo Ghiglieri, and Georg Raffelt for useful discussions and comments on the manuscript. MaD acknowledges the hospitality of the Technical University of Munich in general and Björn Garbrecht in particular during part of the period when this research was performed. YG acknowledges the support of the French Community of Belgium through the FRIA grant No. 1.E.063.22F and thanks the UNSW School of Physics for its hospitality during part of this project. MK acknowledges financial support by DFG under grant KL 1266/11-1. GP and Y³W are supported by the Australian Research Council’s Discovery Project (DP240103130) scheme. This work has been partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - SFB 1258 - 283604770.

A Functions in the continuity equation

The continuity equation (2.12), expressed in terms of the dimensionless variables x, y, z , contains the functions $J_{\pm}, Y_{\pm}, W_{\pm}$, and $G_{1,2}$. We provide here at $\mathcal{O}(e^2)$ their explicit forms [18]:

$$J_{\pm}(\tau) = \frac{1}{\pi^2} \int d\omega \omega^2 \frac{\exp(\sqrt{\tau^2 + \omega^2})}{\left[\exp(\sqrt{\tau^2 + \omega^2}) \pm 1\right]^2}, \quad (\text{A.1a})$$

$$Y_{\pm}(\tau) = \frac{1}{\pi^2} \int d\omega \omega^4 \frac{\exp(\sqrt{\tau^2 + \omega^2})}{\left[\exp(\sqrt{\tau^2 + \omega^2}) \pm 1\right]^2}, \quad (\text{A.1b})$$

$$W_{\pm}(\tau) = \frac{1}{\pi^2} \int_0^\infty d\omega \omega^2 \frac{\sqrt{\omega^2 + \tau^2}}{\exp(\sqrt{\omega^2 + \tau^2}) \pm 1}, \quad (\text{A.1c})$$

$$G_1(\tau) = 2\pi\alpha \left[\frac{1}{\tau} \left(\frac{K(\tau)}{3} + 2K^2(\tau) - \frac{J(\tau)}{6} - K(\tau)J(\tau) \right) + \left(\frac{K'(\tau)}{6} - K(\tau)K'(\tau) + \frac{J'(\tau)}{6} + J'(\tau)K(\tau) + J(\tau)K'(\tau) \right) \right], \quad (\text{A.1d})$$

$$G_2(\tau) = -8\pi\alpha \left(\frac{K(\tau)}{6} + \frac{J(\tau)}{6} - \frac{1}{2}K^2(\tau) + K(\tau)J(\tau) \right) \quad (\text{A.1e})$$

$$+ 2\pi\alpha \left(\frac{K'(\tau)}{6} - K(\tau)K'(\tau) + \frac{J'(\tau)}{6} + J'(\tau)K(\tau) + J(\tau)K'(\tau) \right),$$

where $' \equiv \partial/\partial\tau$, and the $+/-$ signs pertain to the Fermi-Dirac and the Bose-Einstein distributions respectively. Expressions for $G_{1,2}(\tau)$ at $\mathcal{O}(e^3)$ can be found in reference [4].

B Beth-Uhlenbeck formula for a non-relativistic e^+e^- plasma

We summarise here a derivation of the Beth-Uhlenbeck formula (4.9); see also [54, 55].

B.1 Charging formula

Given the non-relativistic Hamiltonian (4.4), all thermodynamics quantities relevant for ΔN_{eff} can be established from the associated partition function Z and hence pressure $P = T \ln Z$. To compute the partition function, we make use of the charging formula,

$$PV - P^{(0)}\mathcal{V} = T[\ln Z - \ln Z_0] = T \int_0^1 d\lambda \partial_\lambda \ln Z_\lambda = - \int_0^1 d\lambda \frac{\text{Tr}[e^{-\beta(H_0 + \lambda H_{\text{int}})} H_{\text{int}}]}{Z_\lambda}, \quad (\text{B.1})$$

which is a formal trick to evaluate the pressure in a non-perturbative way. Here, $Z_\lambda = \text{Tr}[e^{-\beta(H_0 + \lambda H_{\text{int}})}]$, and we have omitted the sub- and superscripts “NR” for notational simplicity. In the language of the real-time formalism—also called the closed-time path (CTP) formalism—one can identify from the structure of H_{int} in equation (4.4) three equal-time equilibrium correlators,

$$G_{\eta\xi,\lambda}^{++--}(t\mathbf{x}, \mathbf{y}; t\mathbf{x}, \mathbf{y}) \equiv \langle \eta_\sigma^\dagger(\mathbf{x}) \xi_{\sigma'}(\mathbf{y}) \xi_{\sigma'}^\dagger(\mathbf{y}) \eta_\sigma(\mathbf{x}) \rangle_\lambda, \quad (\text{B.2a})$$

$$G_{\eta\eta,\lambda}^{++--}(t\mathbf{x}, \mathbf{y}; t\mathbf{x}, \mathbf{y}) \equiv \langle \eta_\sigma^\dagger(\mathbf{x}) \eta_{\sigma'}^\dagger(\mathbf{y}) \eta_{\sigma'}(\mathbf{y}) \eta_\sigma(\mathbf{x}) \rangle_\lambda, \quad (\text{B.2b})$$

$$G_{\xi\xi,\lambda}^{++--}(t\mathbf{x}, \mathbf{y}; t\mathbf{x}, \mathbf{y}) \equiv \langle \xi_\sigma(\mathbf{x}) \xi_{\sigma'}(\mathbf{y}) \xi_{\sigma'}^\dagger(\mathbf{y}) \xi_\sigma^\dagger(\mathbf{x}) \rangle_\lambda, \quad (\text{B.2c})$$

where $\langle \dots \rangle_\lambda = \text{Tr}[e^{-\beta(H_0 + \lambda H_{\text{int}})} \dots] / Z_\lambda$, all field operators on the right are evaluated at the same time t , and the superscript “++--” denotes a certain operator ordering (of a more general correlator on the CTP contour; the interested reader may consult [33] for the convention, although it is of no relevance here).

In terms of the correlators (B.2), the pressure (B.1) can now be written as

$$PV - P^{(0)}\mathcal{V} = \int_0^1 d\lambda \frac{1}{2} \sum_{\sigma,\sigma'} \int d^3\mathbf{x} d^3\mathbf{y} V(|\mathbf{x} - \mathbf{y}|) \left[G_{\eta\eta,\lambda}^{++--}(t\mathbf{x}, \mathbf{y}; t\mathbf{x}, \mathbf{y}) + G_{\xi\xi,\lambda}^{++--} - 2G_{\eta\xi,\lambda}^{++--} \right], \quad (\text{B.3})$$

where the spin indices of the correlators are implicit and set by equations (B.2). We can further connect these correlators to spectral functions via the Kubo-Martin-Schwinger (KMS) relation. Taking the dilute limit and expressed in Wigner coordinates and momentum space, these relations are (see, e.g., [33], chapter V)

$$G_{\bullet\bullet,\lambda}^{++--}(t, \mathbf{x}, \mathbf{y}; t, \mathbf{x}, \mathbf{y}) = e^{-\beta 2m_e} \int \frac{d^3\mathbf{K}}{(2\pi)^3} e^{-\beta \mathbf{K}^2/(4m_e)} \int_{-\infty}^{\infty} \frac{dE}{(2\pi)} e^{-\beta E} G_{\bullet\bullet,\lambda}^\rho(E; \mathbf{r}, \mathbf{r}), \quad (\text{B.4})$$

where $\bullet = \eta, \xi$, K is the centre-of-mass momentum, and $\mathbf{r}$ is the relative distance. Observe that E runs from $-\infty$ to $+\infty$: a positive value of $E = m_e v^2/4$ denotes the relative kinetic energy with relative velocity v , while a negative value corresponds to the binding energy once the spectral function G^ρ has support.

Then, recasting (B.3) in terms of (B.4) and performing the integration over $\mathbf{K}$ and the centre-of-mass position (the latter integration yields the volume $\mathcal{V}$), the charging formula (B.1) now reads

$$P - P^{(0)} = \left(\frac{2m_e T}{2\pi} \right)^{3/2} e^{-\beta 2m_e} \frac{1}{2} \sum_{\sigma, \sigma'} \int_0^1 d\lambda \int d^3\mathbf{r} V(r) \times \int_{-\infty}^{\infty} \frac{dE}{(2\pi)} e^{-\beta E} \left[G_{\eta\eta, \lambda}^\rho(E; \mathbf{r}, \mathbf{r}) + G_{\xi\xi, \lambda}^\rho(E; \mathbf{r}, \mathbf{r}) - 2G_{\eta\xi, \lambda}^\rho(E; \mathbf{r}, \mathbf{r}) \right]. \quad (\text{B.5})$$

We have not yet specified the form of the spectral functions G^ρ ; these can be obtained from the imaginary part of the retarded correlator, to be derived next.

B.2 Non-relativistic four-point correlators

Our starting point to estimate the retarded four-point correlators is non-relativistic QED [60], where the non-relativistic effective action $S = S^0 + S^{\text{int}}$ in vacuum is composed of

$$S^0 = \sum_{\sigma} \int d^4x \eta_{\sigma}^{\dagger}(x) \left[i\partial_t + \frac{\nabla^2}{2m_e} \right] \eta_{\sigma}(x) + \xi_{\sigma}^{\dagger}(x) \left[i\partial_t - \frac{\nabla^2}{2m_e} \right] \xi_{\sigma}(x), \\ S^{\text{int}} = \frac{1}{2} \sum_{\sigma, \sigma'} \int d^4x d^3\mathbf{y} V(|\mathbf{x} - \mathbf{y}|) \times \left[\eta_{\sigma}^{\dagger}(\mathbf{x}) \eta_{\sigma}(\mathbf{x}) \eta_{\sigma'}^{\dagger}(\mathbf{y}) \eta_{\sigma'}(\mathbf{y}) + \xi_{\sigma}^{\dagger}(\mathbf{x}) \xi_{\sigma}(\mathbf{x}) \xi_{\sigma'}^{\dagger}(\mathbf{y}) \xi_{\sigma'}(\mathbf{y}) + 2\eta_{\sigma}^{\dagger}(\mathbf{x}) \eta_{\sigma}(\mathbf{x}) \xi_{\sigma'}^{\dagger}(\mathbf{y}) \xi_{\sigma'}(\mathbf{y}) \right], \quad (\text{B.6})$$

and we focus on potential exchanges only. Note the different operator ordering compared to equation (4.4).

We wish to derive from the action (B.6) the equations of motion for the four-point correlators in vacuum,

$$G_{\eta\eta}(x, y; z, w) = \langle T \eta_{\sigma_1}(x) \eta_{\sigma_2}(y) \eta_{\sigma_3}^{\dagger}(w) \eta_{\sigma_4}^{\dagger}(z) \rangle, \\ G_{\eta\xi}(x, y; z, w) = \langle T \eta_{\sigma_1}(x) \xi_{\sigma_2}^{\dagger}(y) \xi_{\sigma_3}(w) \eta_{\sigma_4}^{\dagger}(z) \rangle. \quad (\text{B.7})$$

In integral form, these equations of motion are given by

$$G_{\eta\eta}(t\mathbf{x}\mathbf{y}, t'\mathbf{z}\mathbf{w}) = G_{\eta}(t\mathbf{x}, t'\mathbf{z}) G_{\eta}(t\mathbf{y}, t'\mathbf{w}) - G_{\eta}(t\mathbf{x}, t'\mathbf{w}) G_{\eta}(t\mathbf{y}, t'\mathbf{z}) \\ + i \int d\bar{t} d^3\mathbf{a} d^3\mathbf{b} G_{\eta}(t\mathbf{x}, \bar{t}\mathbf{a}) G_{\eta}(t\mathbf{y}, \bar{t}\mathbf{b}) V(\mathbf{a} - \mathbf{b}) G_{\eta\eta}(\bar{t}\mathbf{a}\mathbf{b}, t'\mathbf{z}\mathbf{w}), \quad (\text{B.8a})$$

$$G_{\eta\xi}(t\mathbf{x}\mathbf{y}, t'\mathbf{z}\mathbf{w}) = G_{\eta}(t\mathbf{x}, t'\mathbf{z}) G_{\xi}(t\mathbf{y}, t'\mathbf{w}) \\ - i \int d\bar{t} d^3\mathbf{a} d^3\mathbf{b} G_{\eta}(t\mathbf{x}, \bar{t}\mathbf{a}) G_{\xi}(t\mathbf{y}, \bar{t}\mathbf{b}) V(\mathbf{a} - \mathbf{b}) G_{\eta\eta}(\bar{t}\mathbf{a}\mathbf{b}, t'\mathbf{z}\mathbf{w}), \quad (\text{B.8b})$$

where we have truncated the correlator hierarchy at the four-point function level. In differential form and approximating the two-point functions by their free counterparts, they appear in Wigner coordinates as

$$\left[E + i\epsilon + \frac{\nabla_{\mathbf{r}}^2}{m_e} + V(\mathbf{r}) \right] G_{\eta\eta}(E; \mathbf{r}, \mathbf{r}') = i\delta_{\sigma_1\sigma_4} \delta_{\sigma_2\sigma_3} \delta^{(3)}(\mathbf{r} - \mathbf{r}') - i\delta_{\sigma_1\sigma_3} \delta_{\sigma_2\sigma_4} \delta^{(3)}(\mathbf{r} + \mathbf{r}'), \quad (\text{B.9a})$$

$$\left[E + i\epsilon + \frac{\nabla_{\mathbf{r}}^2}{m_e} - V(\mathbf{r}) \right] G_{\eta\xi}(E; \mathbf{r}, \mathbf{r}') = i\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}\delta^{(3)}(\mathbf{r} - \mathbf{r}') , \quad (\text{B.9b})$$

and admit Green's function solutions

$$G_{\eta\eta}(E; \mathbf{r}, \mathbf{r}') = i \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}\psi_{\mathbf{p}}(\mathbf{r})\psi_{\mathbf{p}}^*(\mathbf{r}') - \delta_{\sigma_1\sigma_3}\delta_{\sigma_2\sigma_4}\psi_{\mathbf{p}}(\mathbf{r})\psi_{\mathbf{p}}^*(-\mathbf{r}')}{E - p^2/m_e + i\epsilon} , \quad (\text{B.10a})$$

$$G_{\eta\xi}(E; \mathbf{r}, \mathbf{r}') = i\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3} \oint_{\gamma} \frac{\psi_{\gamma}(\mathbf{r})\psi_{\gamma}^*(\mathbf{r}')}{E - E_{\gamma} + i\epsilon} , \quad (\text{B.10b})$$

expressed here in spectral representation. Observe that the wavefunctions ψ are the solution of the homogenous (Schrödinger) equation and fulfil the completeness relation and normalisation

$$\oint_{\gamma} \psi_{\gamma}(\mathbf{r})\psi_{\gamma}^*(\mathbf{r}') = \delta^{(3)}(\mathbf{r} - \mathbf{r}') , \quad (\text{B.11a})$$

$$\int d^3\mathbf{r} \psi_{\mathbf{p}}(\mathbf{r})\psi_{\mathbf{p}'}^*(\mathbf{r}) = (2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{p}') , \quad (\text{B.11b})$$

$$\int d^3\mathbf{r} \psi_{n\ell m}(\mathbf{r})\psi_{n'\ell' m'}^*(\mathbf{r}) = \delta_{nn'}\delta_{\ell\ell'}\delta_{mm'} . \quad (\text{B.11c})$$

Starting from equation (B.11a), it is straightforward to verify that equations (B.10a) and (B.10b) are the solutions to (B.9a) and (B.9b), respectively.

B.3 Integral reductions

To evaluate the charging formula (B.5), we use the solutions (B.10a) and (B.10b) to reduce the number of integrals. We perform first a partial integration,

$$\int dE e^{-\beta E} \int_0^1 d\lambda \int d^3\mathbf{r} V(\mathbf{r}) G_{\bullet\bullet,\lambda}^{\rho}(E; \mathbf{r}, \mathbf{r}) = T \int dE e^{-\beta E} \int_0^1 d\lambda \int d^3\mathbf{r} V(\mathbf{r}) \partial_E G_{\bullet\bullet,\lambda}^{\rho}(E; \mathbf{r}, \mathbf{r}). \quad (\text{B.12})$$

Here, the spectral function G^{ρ} can be computed from $G_{\bullet\bullet,\lambda}^{\rho} = G_{\bullet\bullet,\lambda}^R - G_{\bullet\bullet,\lambda}^A$, where G^R and G^A are the retarded and the advanced correlators, respectively. In vacuum, the retarded correlator $G_{\bullet\bullet}^R$ is equivalent to the time-ordered correlators (B.10a) or (B.10b) for $\bullet\bullet = \eta\eta$ and $\eta\xi$, respectively, while its advanced counterpart has $-i\epsilon$ instead of $+i\epsilon$ in the denominator.

Noting that the energy derivative acting on the retarded correlator can be written as

$$\partial_E \langle \mathbf{r} | \frac{iV}{E - h_0 \mp \lambda V + i\epsilon} | \mathbf{r} \rangle = \mp \partial_{\lambda} \langle \mathbf{r} | \frac{i}{E - h_0 \mp \lambda V + i\epsilon} | \mathbf{r} \rangle , \quad (\text{B.13})$$

and analogously for the advanced correlator, the λ -integration in (B.12) evaluates trivially to

$$\int_0^1 d\lambda \int d^3\mathbf{r} V(\mathbf{r}) \partial_E G_{\bullet\bullet,\lambda}^{\rho}(E; \mathbf{r}, \mathbf{r}) = \mp \int d^3\mathbf{r} \left[G_{\bullet\bullet}^{\rho}(E; \mathbf{r}, \mathbf{r}) - G_{\bullet\bullet,0}^{\rho}(E; \mathbf{r}, \mathbf{r}) \right] , \quad (\text{B.14})$$

where $G_{\bullet\bullet,0}^{\rho}$ and $G_{\bullet\bullet}^{\rho}$ are the free and the fully interacting ($\lambda = 1$) spectral function, respectively, and the “ $-$ ” (“ $+$ ”) sign applies to the attractive (repulsive) case. Then, we arrive at

$$\beta P - \beta P^{(0)} = \left(\frac{2m_e T}{2\pi} \right)^{3/2} e^{-\beta 2m_e} \int_{-\infty}^{\infty} \frac{dE}{(2\pi)} e^{-\beta E} \times \quad (\text{B.15})$$

$$\sum_{\sigma,\sigma'} \int d^3\mathbf{r} \left[G_{\eta\xi}^\rho(E; \mathbf{r}, \mathbf{r}) - G_{\eta\xi,0}^\rho(E; \mathbf{r}, \mathbf{r}) + \frac{1}{2} \left(G_{\eta\eta}^\rho - G_{\eta\eta,0}^\rho + G_{\xi\xi}^\rho - G_{\xi\xi,0}^\rho \right) \right]$$

as an intermediate result.

To perform the radial integration, consider first the $E < 0$ case. Here, only the interacting attractive case is non-vanishing, giving the positronium bound states. The integrand on the right-hand side of (B.14) is therefore

$$\left[G_{\eta\xi}^\rho(E; \mathbf{r}, \mathbf{r}) - G_{\eta\xi,0}^\rho(E; \mathbf{r}, \mathbf{r}) \right] \Big|_{E < 0} = G_{\eta\xi}^\rho(E; \mathbf{r}, \mathbf{r}) \Big|_{E < 0} = \delta_{\sigma\sigma} \delta_{\sigma'\sigma'} \sum_{\mathcal{B}} |\psi_{\mathcal{B}}(\mathbf{r})|^2 (2\pi) \delta(E - E_{\mathcal{B}}) . \quad (\text{B.16})$$

Then, inserting (B.16) into equation (B.15) and performing the radial and energy integration (noting that all bound-state wavefunctions are normalised to unity under radial integration) and spin summation yield the bound-state part of the Beth-Uhlenbeck formula (4.9).

For $E > 0$ scattering states, decomposing the Green's functions into partial waves,

$$G_{\bullet\bullet}(E; \mathbf{r}, \mathbf{r}') = \sum_{\ell} \frac{2\ell+1}{4\pi} P_{\ell}(\cos \theta) G_{\bullet\bullet,\ell}(E, r, r') , \quad (\text{B.17})$$

leaves us with the radial integral

$$\int_0^\infty dr [u_{k\ell}(r)u_{k\ell}(r) - u_{k\ell}^0(r)u_{k\ell}^0(r)] = \frac{1}{\pi} \frac{dE}{dk} \frac{d}{dE} \delta_{\ell}(E) , \quad (\text{B.18})$$

where we have used the Schrödinger equation for the radially reduced wavefunction $u_{k\ell}$ and its asymptotic behaviour at large r ; see chapter 3.5 in [55] for details. Equation (B.18) is the origin of the phase shifts δ_{ℓ} in the Beth-Uhlenbeck formula (4.9). The non-trivial factor $(1 - (-1)^{\ell}/2)/2$ preceding the phases of the repulsive degenerate states follows on the other hand from the spin structure in (B.10a) and the partial wave decomposition of $\psi_{\mathbf{p}}(\mathbf{r})\psi_{\mathbf{p}}^*(-\mathbf{r}')$, which yields $P_{\ell}(-\cos \theta) = (-1)^{\ell} P_{\ell}(\cos \theta)$.

C Non-relativistic Feynman rules in momentum space

The two-point function of non-relativistic electrons on the CTP contour $\mathcal{C}$ is defined as

$$G_{\eta}(x, y) \equiv \langle T_{\mathcal{C}} \eta_{\sigma}(x) \eta_{\sigma'}^{\dagger}(y) \rangle . \quad (\text{C.1})$$

In Wigner coordinates, the corresponding Fourier-transformed correlator is given by

$$G_{\eta}(x - y) = \int \frac{d^4 p}{(2\pi)^4} e^{-ip(x-y)} G_{\eta}(p) . \quad (\text{C.2})$$

Of special interest is the free “+−” component, which in equilibrium can be expressed via the KMS relation as

$$\begin{aligned} G_{\eta,0}^{+-}(p) &= G_{\eta,0}^{\rho}(p) e^{-\beta(m_e + p^0)} , \\ G_{\eta,0}^{\rho}(p) &= \delta_{\sigma\sigma'} 2\pi \delta(p^0 - \mathbf{p}^2/(2m_e)) , \end{aligned} \quad (\text{C.3})$$

where $G_{\eta,0}^{\rho}$ is the free single electron spectral function. These expressions are used in equation (4.8).

References

- [1] S. Gariazzo, P.F. de Salas and S. Pastor, *Thermalisation of sterile neutrinos in the early Universe in the 3+1 scheme with full mixing matrix*, *JCAP* **07** (2019) 014 [[1905.11290](#)].
- [2] K. Akita and M. Yamaguchi, *A precision calculation of relic neutrino decoupling*, *JCAP* **08** (2020) 012 [[2005.07047](#)].
- [3] J. Froustey, C. Pitrou and M.C. Volpe, *Neutrino decoupling including flavour oscillations and primordial nucleosynthesis*, *JCAP* **12** (2020) 015 [[2008.01074](#)].
- [4] J.J. Bennett, G. Buldgen, M. Drewes and Y.Y.Y. Wong, *Towards a precision calculation of the effective number of neutrinos N_{eff} in the Standard Model I: the QED equation of state*, *JCAP* **03** (2020) 003 [[1911.04504](#)].
- [5] J.J. Bennett, G. Buldgen, P.F. De Salas, M. Drewes, S. Gariazzo, S. Pastor et al., *Towards a precision calculation of N_{eff} in the Standard Model II: Neutrino decoupling in the presence of flavour oscillations and finite-temperature QED*, *JCAP* **04** (2021) 073 [[2012.02726](#)].
- [6] M. Cielo, M. Escudero, G. Mangano and O. Pisanti, *N_{eff} in the Standard Model at NLO is 3.043*, *Phys. Rev. D* **108** (2023) L121301 [[2306.05460](#)].
- [7] G. Jackson and M. Laine, *QED corrections to the thermal neutrino interaction rate*, *JHEP* **05** (2024) 089 [[2312.07015](#)].
- [8] M. Drewes, Y. Georis, M. Klasen, L.P. Wiggering and Y.Y.Y. Wong, *Towards a precision calculation of N_{eff} in the Standard Model. Part III. Improved estimate of NLO contributions to the collision integral*, *JCAP* **06** (2024) 032 [[2402.18481](#)].
- [9] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [[1807.06209](#)].
- [10] CMB-S4 collaboration, *CMB-S4 Science Book, First Edition*, [1610.02743](#).
- [11] D.A. Dicus, E.W. Kolb, A.M. Gleeson, E.C.G. Sudarshan, V.L. Teplitz and M.S. Turner, *Primordial Nucleosynthesis Including Radiative, Coulomb, and Finite Temperature Corrections to Weak Rates*, *Phys. Rev. D* **26** (1982) 2694.
- [12] S. Dodelson and M.S. Turner, *Nonequilibrium neutrino statistical mechanics in the expanding universe*, *Phys. Rev. D* **46** (1992) 3372.
- [13] A.F. Heckler, *Astrophysical applications of quantum corrections to the equation of state of a plasma*, *Phys. Rev. D* **49** (1994) 611.
- [14] S. Hannestad and J. Madsen, *Neutrino decoupling in the early universe*, *Phys. Rev. D* **52** (1995) 1764 [[astro-ph/9506015](#)].
- [15] N. Fornengo, C.W. Kim and J. Song, *Finite temperature effects on the neutrino decoupling in the early universe*, *Phys. Rev. D* **56** (1997) 5123 [[hep-ph/9702324](#)].
- [16] R.E. Lopez and M.S. Turner, *An Accurate Calculation of the Big Bang Prediction for the Abundance of Primordial Helium*, *Phys. Rev. D* **59** (1999) 103502 [[astro-ph/9807279](#)].
- [17] S. Esposito, G. Miele, S. Pastor, M. Peloso and O. Pisanti, *Nonequilibrium spectra of degenerate relic neutrinos*, *Nucl. Phys. B* **590** (2000) 539 [[astro-ph/0005573](#)].
- [18] G. Mangano, G. Miele, S. Pastor and M. Peloso, *A Precision calculation of the effective number of cosmological neutrinos*, *Phys. Lett. B* **534** (2002) 8 [[astro-ph/0111408](#)].
- [19] J. Birrell, C.-T. Yang and J. Rafelski, *Relic Neutrino Freeze-out: Dependence on Natural Constants*, *Nucl. Phys. B* **890** (2014) 481 [[1406.1759](#)].
- [20] E. Grohs, G.M. Fuller, C.T. Kishimoto, M.W. Paris and A. Vlasenko, *Neutrino energy transport in weak decoupling and big bang nucleosynthesis*, *Phys. Rev. D* **93** (2016) 083522 [[1512.02205](#)].

- [21] A.D. Dolgov, S.H. Hansen and D.V. Semikoz, *Nonequilibrium corrections to the spectra of massless neutrinos in the early universe*, *Nucl. Phys. B* **503** (1997) 426 [[hep-ph/9703315](#)].
- [22] A.D. Dolgov, S.H. Hansen and D.V. Semikoz, *Nonequilibrium corrections to the spectra of massless neutrinos in the early universe: Addendum*, *Nucl. Phys. B* **543** (1999) 269 [[hep-ph/9805467](#)].
- [23] G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti and P.D. Serpico, *Relic neutrino decoupling including flavor oscillations*, *Nucl. Phys. B* **729** (2005) 221 [[hep-ph/0506164](#)].
- [24] P.F. de Salas and S. Pastor, *Relic neutrino decoupling with flavour oscillations revisited*, *JCAP* **07** (2016) 051 [[1606.06986](#)].
- [25] R.S.L. Hansen, S. Shalgar and I. Tamborra, *Neutrino flavor mixing breaks isotropy in the early universe*, *JCAP* **07** (2021) 017 [[2012.03948](#)].
- [26] L.C. Thomas, T. Dezen, E.B. Grohs and C.T. Kishimoto, *Electron-Positron Annihilation Freeze-Out in the Early Universe*, *Phys. Rev. D* **101** (2020) 063507 [[1910.14050](#)].
- [27] B. von Harling and K. Petraki, *Bound-state formation for thermal relic dark matter and unitarity*, *JCAP* **12** (2014) 033 [[1407.7874](#)].
- [28] C. Vasilaki and K. Petraki, *Radiation back-reaction during dark-matter freeze-out via metastable bound states*, *JCAP* **06** (2024) 027 [[2402.13069](#)].
- [29] S.G. Karshenboim, *Precision physics of simple atoms: QED tests, nuclear structure and fundamental constants*, *Phys. Rept.* **422** (2005) 1 [[hep-ph/0509010](#)].
- [30] J.L. Feng, M. Kaplinghat, H. Tu and H.-B. Yu, *Hidden Charged Dark Matter*, *JCAP* **07** (2009) 004 [[0905.3039](#)].
- [31] J.A.D. Diacounis and Y.Y.Y. Wong, *Trading kinetic energy: how late kinetic decoupling of dark matter changes N_{eff}* , *JCAP* **01** (2019) 001 [[1811.05601](#)].
- [32] T. Binder, A. Filimonova, K. Petraki and G. White, *Saha equilibrium for metastable bound states and dark matter freeze-out*, *Phys. Lett. B* **833** (2022) 137323 [[2112.00042](#)].
- [33] T. Binder, L. Covi and K. Mukaida, *Dark Matter Sommerfeld-enhanced annihilation and Bound-state decay at finite temperature*, *Phys. Rev. D* **98** (2018) 115023 [[1808.06472](#)].
- [34] M. Laine, O. Philipsen, P. Romatschke and M. Tassler, *Real-time static potential in hot QCD*, *JHEP* **03** (2007) 054 [[hep-ph/0611300](#)].
- [35] S. Kim and M. Laine, *On thermal corrections to near-threshold annihilation*, *JCAP* **01** (2017) 013 [[1609.00474](#)].
- [36] T. Matsui and H. Satz, *J/ψ Suppression by Quark-Gluon Plasma Formation*, *Phys. Lett. B* **178** (1986) 416.
- [37] E. Braaten and D. Segel, *Neutrino energy loss from the plasma process at all temperatures and densities*, *Phys. Rev. D* **48** (1993) 1478 [[hep-ph/9302213](#)].
- [38] H. Scherer and K. Schutz, *Photon self-energy at all temperatures and densities in all of phase space*, *JHEP* **11** (2024) 139 [[2405.18466](#)].
- [39] J.I. Kapusta and C. Gale, *Finite-temperature field theory: Principles and applications*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (2011), [10.1017/CBO9780511535130](#).
- [40] M.L. Bellac, *Thermal Field Theory*, Cambridge Monographs on Mathematical Physics, Cambridge University Press (3, 2011), [10.1017/CBO9780511721700](#).
- [41] R. Sachs and M. Goeppert-Mayer, *Calculations on a new neutron-proton interaction potential*, *Physical Review* **53** (1938) 991.

- [42] F. Rogers, H. Graboske Jr and D. Harwood, *Bound eigenstates of the static screened coulomb potential*, *Physical Review A* **1** (1970) 1577.
- [43] N. Brambilla, J. Ghiglieri, A. Vairo and P. Petreczky, *Static quark-antiquark pairs at finite temperature*, *Phys. Rev. D* **78** (2008) 014017 [[0804.0993](#)].
- [44] A. Vairo, *Effective field theories for heavy quarkonium at finite temperature*, *PoS CONFINEMENT8* (2008) 002 [[0901.3495](#)].
- [45] M.A. Escobedo and J. Soto, *Non-relativistic bound states at finite temperature (I): The Hydrogen atom*, *Phys. Rev. A* **78** (2008) 032520 [[0804.0691](#)].
- [46] S. Biondini, N. Brambilla, G. Qerimi and A. Vairo, *Effective field theories for dark matter pairs in the early universe: cross sections and widths*, *JHEP* **07** (2023) 006 [[2304.00113](#)].
- [47] F. Dominguez and B. Wu, *On dissociation of heavy mesons in a hot quark-gluon plasma*, *Nucl. Phys. A* **818** (2009) 246 [[0811.1058](#)].
- [48] M.A. Escobedo and J. Soto, *Non-relativistic bound states at finite temperature (II): the muonic hydrogen*, *Phys. Rev. A* **82** (2010) 042506 [[1008.0254](#)].
- [49] M.A. Strosio, *Positronium: Review of the Theory*, *Phys. Rept.* **22** (1975) 215.
- [50] M. Pospelov and A. Ritz, *Astrophysical Signatures of Secluded Dark Matter*, *Phys. Lett. B* **671** (2009) 391 [[0810.1502](#)].
- [51] J.D. March-Russell and S.M. West, *WIMPosium and Boost Factors for Indirect Dark Matter Detection*, *Phys. Lett. B* **676** (2009) 133 [[0812.0559](#)].
- [52] W. Shepherd, T.M.P. Tait and G. Zaharijas, *Bound states of weakly interacting dark matter*, *Phys. Rev. D* **79** (2009) 055022 [[0901.2125](#)].
- [53] E. Beth and G. Uhlenbeck, *The quantum theory of the non-ideal gas. II. Behaviour at low temperatures*, *Physica* **4** (1937) 915.
- [54] D. Kremp, M. Schlanges and W.-D. Kraeft, *Quantum statistics of nonideal plasmas*, vol. 25, Springer Science & Business Media (2005).
- [55] W. Ebeling, W. Kraeft and D. Kremp, *Theory of Bound States and Ionization Equilibrium in Plasmas and Solids*, Akademie-Verlag (01, 1976).
- [56] S. Cassel, *Sommerfeld factor for arbitrary partial wave processes*, *J. Phys. G* **37** (2010) 105009 [[0903.5307](#)].
- [57] M. Beneke, T. Binder, L. De Ros and M. Garny, *Enhancement of p-wave dark matter annihilation by quasi-bound states*, *JHEP* **06** (2024) 207 [[2403.07108](#)].
- [58] S. Tulin, H.-B. Yu and K.M. Zurek, *Beyond Collisionless Dark Matter: Particle Physics Dynamics for Dark Matter Halo Structure*, *Phys. Rev. D* **87** (2013) 115007 [[1302.3898](#)].
- [59] A. Mitridate, M. Redi, J. Smirnov and A. Strumia, *Cosmological Implications of Dark Matter Bound States*, *JCAP* **05** (2017) 006 [[1702.01141](#)].
- [60] W.E. Caswell and G.P. Lepage, *Effective Lagrangians for Bound State Problems in QED, QCD, and Other Field Theories*, *Phys. Lett. B* **167** (1986) 437.