Massive sterile neutrinos in the early Universe: From thermal decoupling to cosmological constraints

Leonardo Mastrototaro[1,2,*], Pasquale Dario Serpico[3,†], Alessandro Mirizzi[4,5,‡] and Ninetta Saviano[6,7,§]

[1]*Dipartimento di Fisica "E. R. Caianiello," Università degli Studi di Salerno,*
Via Giovanni Paolo II, 132—84084 Fisciano (SA), Italy
[2]*Istituto Nazionale di Fisica Nucleare—Gruppo Collegato di Salerno—Sezione di Napoli,*
Via Giovanni Paolo II, 132—84084 Fisciano (SA), Italy
[3]*LAPTh, Université Grenoble Alpes, USMB, CNRS, F-74000 Annecy, France*
[4]*Dipartimento Interateneo di Fisica "Michelangelo Merlin," Via Amendola 173, 70126 Bari, Italy*
[5]*Istituto Nazionale di Fisica Nucleare—Sezione di Bari, Via Orabona 4, 70126 Bari, Italy*
[6]*INFN—Sezione di Napoli, Complesso Università Monte S. Angelo, I-80126 Napoli, Italy*
[7]*Scuola Superiore Meridionale, Università degli studi di Napoli "Federico II,"*
Largo San Marcellino 10, 80138 Napoli, Italy

We consider relatively heavy neutrinos ν_H, mostly contributing to a sterile state ν_s, with mass in the range $10 \text{ MeV} \lesssim m_s \lesssim m_\pi \sim 135 \text{ MeV}$, which are thermally produced in the early Universe in collisional processes involving active neutrinos and freezing out after the QCD phase transition. If these neutrinos decay after the active neutrino decoupling, they generate extra neutrino radiation but also contribute to entropy production. Thus, they alter the value of the effective number of neutrino species N_{eff} as, for instance, measured by the cosmic microwave background (CMB), as well as affect primordial nucleosynthesis (BBN), notably ^{4}He production. We provide a detailed account of the solution of the relevant Boltzmann equations. We also identify the parameter space allowed by current *Planck* satellite data and forecast the parameter space probed by future stage-4 ground-based CMB observations, expected to match or surpass BBN sensitivity.

I. INTRODUCTION

Feebly interacting particles characterized by extremely suppressed interactions with the Standard Model particles have received growing interest in the past decade (see [1] for a recent review). In this context, a fourth neutrino mass state ν_H with mass $\sim\mathcal{O}(100)$ MeV, mostly contributing to an electroweak singlet neutrino state ν_s due to Z-width constraints [2], emerges rather naturally in extensions of the Standard Model, like dynamical electroweak symmetry breaking [3] or the neutrino minimal Standard Model (νMSM) [4,5]. In the latter case, such particles can be related to fundamental problems of particle physics like the

*lmastrototaro@unisa.it
†serpico@lapth.cnrs.fr
‡alessandro.mirizzi@ba.infn.it
§nsaviano@na.infn.it

origin of neutrino mass, the baryon asymmetry in the early Universe, and the nature of dark matter.

The parameter space of a fourth neutrino in this mass range is strongly constrained by collider and beam-dump experiments for a dominant mixing with either ν_e and ν_μ [6,7], but it is significantly less constrained if mixed with ν_τ, with bounds at high masses coming from searches of decays of D mesons and τ leptons [8] and SuperKamiokande data [9]. Furthermore, ν_H can be emitted by a core-collapse supernova. In this context, limits have been placed from the SN 1987A observation, requiring that the SN core may not emit too much energy in the ν_s channel, since this additional energy loss would shorten the observed neutrino burst [10–13]. Additionally, in Refs. [12,14], a possible role of heavy sterile neutrinos in enhancing supernova explosions has been discussed.

Further and complementary constraints on heavy sterile neutrinos can also be placed from cosmological arguments. Indeed, ν_s can be produced in the early Universe via collisional processes involving active neutrinos and then decay into lighter species. In particular, for masses in the range $10 \text{ MeV} \lesssim m_s \lesssim m_\pi \sim 135 \text{ MeV}$, the main decay channels are $\nu_s \rightarrow \nu_\alpha \bar{\nu}_\beta \nu_\beta$ [with a branching ratio (b.r.)

of $\mathcal{O}(60\%\text{–}90\%)$, depending on the mixing] and $\nu_s \to \nu e^+ e^-$ [with a branching ratio of $\mathcal{O}(10\%\text{–}40\%)$, depending on the mixing]. The decay products of the sterile neutrinos are injected into the primordial plasma, with the timescale of the event determining its phenomenological impact. If the decay is over when active neutrinos are still tightly coupled to the electromagnetic (e.m.) plasma dominated by photons and $e^\pm$ pairs, full equilibrium conditions are quickly established and no effect remains, but for an unobservable renormalization of the baryon to photon ratio η.[1] If significant decay happens after the active neutrinos have decoupled, part of the energy injected will end up in extra neutrino radiation and part heats the e.m. plasma up. The latter adds to the eventual heating due to $e^+ e^-$ annihilation and increases the photon to neutrino temperature beyond its standard value of $T_\gamma/T_\nu \simeq (11/4)^{1/3}$. Together with the former effect, this alters the effective number of neutrino species N_{eff}, with the two processes going in opposite directions. Also, the nonthermal ν_e and $\bar{\nu}_e$ spectra enter weak interactions, altering—together with N_{eff}—the *neutron-to-proton ratio* which rules the abundance of the primordial yields [10,11,15–18], affecting, in particular, the ⁴He abundance encoded in the primordial helium mass fraction parameter Y_p.

The aim of our paper is to perform a detailed calculation of heavy sterile neutrino decoupling in the early Universe, in particular, computing the effects on N_{eff} and Y_p and assessing the impact of the approximations presented in the seminal works of Refs. [10,11]. Compared instead to the most recent calculations such as Refs. [18,19], we put more emphasis on the low-mass sterile neutrino range, where the number of effects is limited, hopefully offering a pedagogically complete and contained treatment, besides clarifying some points of disagreement. At high masses, new channels involving pions open up and are responsible for extra effects on primordial nucleosynthesis as well as N_{eff} (some of these interesting effects have been described in Refs. [17–19]). In the same spirit, we will limit ourselves to neutrinos that decouple after the QCD phase transition, which translates on the chosen parameter space. One of our primary goals is to compare the effect on BBN with the latest constraints by the *Planck* satellite experiment as well as forecasts of future stage-4 (S4) ground-based CMB observations [20]. We shall illustrate the shifting cosmological constraining power, from a BBN-dominated one to a CMB-dominated one, expected to be basically completed by the S4 era.

The plan of our work is as follows. In Sec. II, we present the heavy sterile neutrino model we will use as a benchmark. In Sec. III, we discuss and solve the kinetic equations describing the sterile neutrino evolution in the early Universe. In Sec. IV, we characterize the impact of heavy sterile neutrino decays on active neutrinos and on N_{eff}. In Sec. V, we present the current constraints and forecasts on sterile neutrino parameter space from BBN and CMB data. Finally, in Sec. VI, we summarize our results and conclude. In Appendix A, we report useful analytical approximations of the decay and scattering rates involving sterile neutrinos. Appendix B details the steps involved in the dimensional reduction of the collision integrals used in the numerical integrations. Appendix C is devoted to a comparison of our results with others previously reported in the literature.

II. HEAVY STERILE NEUTRINO MODEL

We consider heavy sterile neutrinos with masses $10 \text{ MeV} \lesssim m_s \lesssim 135 \text{ MeV}$, mixed dominantly with one active neutrino ν_α ($\alpha = e, \mu, \tau$) as

$$
\begin{aligned}
\nu_\alpha &= \cos\theta_{\alpha s}\nu_\ell + \sin\theta_{\alpha s}\nu_H, \\
\nu_s &= -\sin\theta_{\alpha s}\nu_\ell + \cos\theta_{\alpha s}\nu_H,
\end{aligned} \tag{1}
$$

where ν_ℓ and ν_H are a light and a heavy mass eigenstate, respectively, and $\theta_{\alpha s} \ll 1$; i.e., ν_ℓ is mostly active and ν_H is mostly sterile. We can relate the mixing angle to the unitary mixing matrix U, where

$$
|U_{\alpha s}|^2 \simeq \frac{1}{4}\sin^2 2\theta_{\alpha s} \simeq \theta_{\alpha s}^2. \tag{2}
$$

Through neutral-current interactions, ν_H can decay into a ν_ℓ and a pair of other light leptons. It can also scatter with other species in the plasma. With a little abuse of notation, we shall refer to ν_H as ν_s and to ν_ℓ as ν_α. Unless stated otherwise, we shall consider neutrinos to be Dirac particles.

In vacuum, in the mass range of our interest, the decay rate of sterile neutrinos is dominated by three lepton final states. In what follows, we shall neglect terms of the order of $m_\nu/m_{W,Z}$ in the matrix elements. Their relevant decay processes and matrix elements are presented in Table I. Unless stated otherwise, we adopt the case of mixing with ν_τ as our benchmark. In some cases, we will consider mixing with ν_e in the mass range $m_s < m_e + m_\mu$ in order to compare with previous literature.[2]

The decay rate of sterile neutrinos into three neutrinos (summed over all flavors) is given by (see, e.g., [21])

---

[1]This process leads instead to observable consequences if the renormalization happens during or after primordial nucleosynthesis (BBN), creating, e.g., an effective mismatch of η_{BBN} with respect to the value extracted from cosmic microwave background (CMB) η_{CMB}. This is not relevant in the parameter space explored here, and we will ignore it in the following.

[2]This is just a simplification to avoid including the additional charged current decay channel $\nu_s \to e + \nu_\mu + \mu$.

TABLE I.   Squared matrix elements for sterile neutrino decay processes (assuming mixing with the species α, and $\beta \neq \alpha$), summed over initial and final states and divided by the spin d.o.f. of the sterile neutrino. The particles involved in each decay are enumerated as $1 \to 2 + 3 + 4$. In the last lines, $\tilde{g}_L$ is replaced by g_L in case of mixing with ν_e. The symmetry factor $S = 1/2!$ is included, when two identical particles are present in the final state (first row).

Process	$G_F^{-2}\|U_{\alpha s}\|^{-2}\|M\|^2$
$\nu_s \to \nu_\alpha + \bar{\nu}_\alpha + \nu_\alpha$	$32(p_1 \cdot p_4)(p_2 \cdot p_3)$
$\nu_s \to \nu_\alpha + \nu_\beta + \bar{\nu}_\beta$	$16(p_1 \cdot p_4)(p_2 \cdot p_3)$
$\nu_s \to \nu_\alpha + e^+ + e^-$	$64[\tilde{g}_L^2(p_1 \cdot p_4)(p_2 \cdot p_3) + g_R^2(p_1 \cdot p_3)(p_2 \cdot p_4) - \tilde{g}_L g_R m_e^2(p_1 \cdot p_3)]$

TABLE II.   Squared matrix elements for sterile neutrino scattering processes (assuming mixing with the species α, and $\beta \neq \alpha$), summed over initial and final states and divided by the two spin d.o.f. of the sterile neutrino. The particles involved in each reaction are enumerated as $1 + 2 \to 3 + 4$. In the last line, $\tilde{g}_L$ is replaced by g_L in case of mixing with ν_e. The symmetry factor $S = 1/2!$ is included, when two identical particles are present in the final state (second row).

Process	$G_F^{-2}\|U_{\tau s}\|^{-2}\|M\|^2$
$\nu_s + \bar{\nu}_\alpha \to \nu_\alpha + \bar{\nu}_\alpha$	$64(p_1 \cdot p_4)(p_2 \cdot p_3)$
$\nu_s + \nu_\alpha \to \nu_\alpha + \nu_\alpha$	$32(p_1 \cdot p_2)(p_3 \cdot p_4)$
$\nu_s + \bar{\nu}_\alpha \to \nu_\beta + \bar{\nu}_\beta$	$16(p_1 \cdot p_4)(p_2 \cdot p_3)$
$\nu_s + \bar{\nu}_\beta \to \nu_\alpha + \bar{\nu}_\beta$	$16(p_1 \cdot p_4)(p_2 \cdot p_3)$
$\nu_s + \bar{\nu}_\alpha \to e^+ + e^-$	$64[\tilde{g}_L^2(p_1 \cdot p_4)(p_2 \cdot p_3) + g_R^2(p_1 \cdot p_3)(p_2 \cdot p_4) - \tilde{g}_L g_R m_e^2(p_1 \cdot p_3)]$
$\nu_s + e^- \to \nu_\alpha + e^-$	$64[\tilde{g}_L^2(p_1 \cdot p_2)(p_3 \cdot p_4) + g_R^2(p_1 \cdot p_4)(p_2 \cdot p_3) - \tilde{g}_L g_R m_e^2(p_1 \cdot p_3)]$
$\nu_s + e^+ \to \nu_\alpha + e^+$	$64[g_R^2(p_1 \cdot p_2)(p_3 \cdot p_4) + \tilde{g}_L^2(p_1 \cdot p_4)(p_2 \cdot p_3) - \tilde{g}_L g_R m_e^2(p_1 \cdot p_3)]$

$$\sum_\beta \Gamma(\nu_s \to \nu_\tau \bar{\nu}_\beta \nu_\beta) = \frac{|U_{\tau s}|^2}{192\pi^3} G_F^2 m_s^5, \qquad (3)$$

while the decay into a neutrino plus $e^+ e^-$ pair, in the limit where $m_e/m_s \ll 1$ is neglected, is

$$\Gamma(\nu_s \to \nu_\tau e^+ e^-) = \frac{|U_{\tau s}|^2}{192\pi^3} G_F^2 m_s^5 (\tilde{g}_L^2 + g_R^2), \qquad (4)$$

where

$$g_L = \frac{1}{2} + \sin^2\theta_W,$$

$$\tilde{g}_L = g_L - 1 = -\frac{1}{2} + \sin^2\theta_W,$$

$$g_R = \sin^2\theta_W, \qquad (5)$$

and, in the case of mixing with ν_e, $\tilde{g}_L$ is replaced by g_L. As a result, the total decay width is written

$$\Gamma_{\nu_s} = \tau_s^{-1} = \frac{|U_{\tau s}|^2}{192\pi^3} G_F^2 m_s^5 (1 + \tilde{g}_L^2 + g_R^2)$$

$$= 3.90 \times 10^5 |U_{\tau s}|^2 \left(\frac{m_s}{100 \text{ MeV}}\right)^5 \text{ s}^{-1}. \qquad (6)$$

Note that, for Majorana neutrinos, the widths for the exclusive decay processes would be the same as for the Dirac case, but the inclusive decay is a factor of 2 larger, since, for each final state accessible to a Dirac particle, two

channels are present (i.e., a channel and its L-number conjugate).

Two-body reactions affect the sterile neutrino chemical and kinetic equilibrium. Those of interest for us are reported, together with the relevant squared matrix elements, in Table II. Our results agree with those reported in Tables 5 and 6 in Ref. [18], bearing in mind their different definition of $|M|^2$, just summed over helicities of initial and final states instead of averaged over the initial state as in our case.[3] In relation instead to Refs. [10,22], we find systematically lower matrix elements compared to their Tables 1 and 2, by a factor of 2. We believe that this is likely a typo in the quantities reported rather than in their results, since we do find agreement with the rates they compute.

III. STERILE AND ACTIVE NEUTRINO EVOLUTION IN THE EARLY UNIVERSE

A. Equations of motion

Following Ref. [23], in order to describe the time evolution of the sterile neutrino ensemble in the early Universe, it proves useful to define the following dimensionless variables which replace time, momentum, and photon temperature, respectively:

$$x \equiv ma, \qquad y = pa, \qquad z = Ta, \qquad (7)$$

---

[3]This is explicitly shown in their Eq. (D.1), where their factor of 2 larger value in $|M|^2$ is compensated by the further $1/g$ in their integral prefactor, $g = 2$ being the spin multiplicity.

where m is an arbitrary mass scale which we set equal to 1 MeV. Note that the function a can be normalized, without loss of generality, so that $z = 1$ at the largest temperature of interest here, when all particles in the plasma are in equilibrium with each other, and neutrinos also share the same temperature T. In terms of these variables, we can write the equations of motion (EOMs) for the heavy sterile neutrinos distribution function f_{ν_s} as [10,11]

$$Hx\partial_x f_{\nu_s} = I_{\nu_s}[f_{\nu_s}]. \tag{8}$$

A similar equation holds for the evolution of active neutrino species f_{ν_α}:

$$Hx\partial_x f_{\nu_\alpha} = I_{\nu_\alpha}[f_{\nu_\alpha}], \qquad \nu_\alpha = \nu_e, \nu_\mu, \nu_\tau. \tag{9}$$

In the previous expressions, H denotes the cosmic expansion Hubble rate given by the Friedmann equation as $H^2 = 8\pi\rho/(3m_{\rm Pl}^2)$, where ρ is the total energy density and $m_{\rm Pl} = G_{\rm N}^{-1/2}$ is the Planck mass in terms of the Newton constant $G_{\rm N}$. We will consider the plasma to be initially thermally populated by pions and all lighter particles, neglecting the nuclei contribution to ρ and assuming equal distributions of particles and antiparticles.

The right-hand-side (rhs) term in Eqs. (8) and (9) contains the sum of the collisional and decay terms for sterile and active neutrinos, respectively, and reads

$$I[f_\nu] = \frac{1}{2E}\int \prod_i \left(\frac{d^3 p_i}{2E_i(2\pi)^3}\right) \prod_f \left(\frac{d^3 p_f}{2E_f(2\pi)^3}\right)(2\pi)^4 \delta^{(4)}$$
$$\times \left(\sum_i p_i - \sum_f p_f\right)|M_{fi}|^2 F(f_i, f_f), \tag{10}$$

with $|M_{fi}|^2$ the sum of the squared-matrix elements over initial and final states, divided by the spin multiplicity of the state of interest. In the case of $I[f_{\nu_s}]$, it contains decay and scattering processes, as shown in Tables I and II, respectively. In the case of $I[f_{\nu_\alpha}]$, it contains scattering processes analogous to those in Table II, apart from the replacement $m_s \to 0$ and $|U_{\tau s}|^2 \to 1 - |U_{\tau s}|^2$ or 1, depending if one is dealing with the mixed flavor or not; it also contains the sterile neutrino decay source term, which is calculated referring to the processes in Table I.

We use squared-matrix elements for the collisional processes for active neutrinos consistent, e.g., with the results in Table 3 in Ref. [18], modulo our different definition of $|M_{fi}|^2$, while we find once again that the results reported, e.g., in Refs. [10,22] are a factor of 2 larger than the correct ones. The statistical factor is written, in general, as

$$F(f_i, f_f) = -\prod_i f_i \prod_f (1 \mp f_f) + \prod_i (1 \mp f_i) \prod_f f_f, \tag{11}$$

where $f_{i,f}$ are the distributions of the particles in the initial (i) or final (f) states, the ($-$) sign ("blocking") refers to fermions and ($+$) sign applies to bosons ("stimulated" effect). Only fermions, however, are present in the processes of interest for us.

Because of the different timescales of the active neutrino oscillation processes (see, e.g., [24]), we take into account oscillations among active neutrinos by postprocessing the flavor distributions via

$$I[f_\alpha] \to \sum_\beta P_{\alpha\beta} I[f_\beta], \tag{12}$$

where $P_{\alpha\beta}$ are the time-averaged transition probabilities [see, e.g., Eq. (26) in Ref. [17]]. Medium modification of the mixing parameters are also included, similarly to Ref. [18].

To get the time evolution of sterile and active neutrino distributions, we have to complete the set of equations [Eqs. (8), (9), and (23)] with the continuity equation, stating the conservation of the total energy density

$$\frac{d}{dx}\bar\rho(x) = \frac{1}{x}(\bar\rho - 3\bar P), \tag{13}$$

where $\bar\rho$ and $\bar P$ are the comoving energy density and pressure of the primordial plasma, respectively:

$$\bar\rho = \rho\left(\frac{x}{m}\right)^4,$$
$$\bar P = P\left(\frac{x}{m}\right)^4. \tag{14}$$

Note that only massive components in the plasma contribute to the rhs of Eq. (13), while (neglecting plasma corrections) relativistic species at equilibrium cancel out. From Eq. (13), by using the expressions in Ref. [23] and remembering that we consider equal particle and antiparticle distributions,

$$\bar\rho_\gamma = \frac{\pi^2}{15}z^4, \tag{15}$$

$$\bar\rho_\ell = \frac{2}{\pi^2}\sum_\ell \int_0^\infty dy\, y^2 \frac{\sqrt{m_\ell^2 x^2/m^2 + y^2}}{\exp(\sqrt{m_\ell^2 x^2/m^2 + y^2}/z) + 1}, \tag{16}$$

$$\bar P_\ell = \frac{2}{3\pi^2}\sum_\ell \int_0^\infty dy \frac{y^4}{\sqrt{m_\ell^2 x^2/m^2 + y^2}}$$
$$\times \frac{1}{\exp(\sqrt{m_\ell^2 x^2/m^2 + y^2}/z) + 1}, \tag{17}$$

$$\bar\rho_\pi = \frac{1}{2\pi^2}\sum_i \int_0^\infty dy\, y^2 \frac{\sqrt{m_i^2 x^2/m^2 + y^2}}{\exp(\sqrt{m_i^2 x^2/m^2 + y^2}/z) - 1}, \tag{18}$$

$$\bar{P}_\pi = \frac{1}{6\pi^2} \sum_i \int_0^\infty dy \frac{y^4}{\sqrt{m_i^2 x^2/m^2 + y^2}}$$
$$\times \frac{1}{\exp(\sqrt{m_i^2 x^2/m^2 + y^2}/z) - 1}, \quad (19)$$

$$\bar{\rho}_{\nu_a} = 3\bar{P}_{\nu_a} = \frac{1}{\pi^2} \int_0^\infty dy\, y^3 \sum_\alpha f_{\nu_a}(x,y), \quad (20)$$

$$\bar{\rho}_{\nu_s} = \frac{1}{\pi^2} \int_0^\infty dy\, y^2 \sqrt{m_s^2 x^2/m^2 + y^2} f_{\nu_s}(x,y), \quad (21)$$

$$\bar{P}_{\nu_s} = \frac{1}{3\pi^2} \int_0^\infty dy \frac{y^4}{\sqrt{m_s^2 x^2/m^2 + y^2}} f_{\nu_s}(x,y), \quad (22)$$

where i runs over the three pions π^+, π^-, and π^0 and ℓ is either the muon or the electron. Note that the Hubble function can be now expressed as

$$H = \sqrt{\frac{8\pi}{3}} \frac{1}{m_{\rm Pl}} \frac{m^2}{x^2} (\bar{\rho}_\gamma + \bar{\rho}_e + \bar{\rho}_\mu + \bar{\rho}_\pi + \bar{\rho}_{\nu_a} + \bar{\rho}_{\nu_s})^{1/2}. \quad (23)$$

Equation (13) gets contributions from all species and can be recast into the equation for the $z(x)$ relation. Let us specify the different contributions. For photons, one has

$$\frac{d\bar{\rho}_\gamma}{dx} - \frac{\bar{\rho}_\gamma - 3\bar{P}_\gamma}{x} = \frac{d\bar{\rho}_\gamma}{dx} = \frac{4\pi^2 z^3}{15} \frac{dz}{dx}. \quad (24)$$

For the electrons, if setting $\kappa_e \equiv m_e/m$, one finds

$$\frac{d\bar{\rho}_e}{dx} - \frac{\bar{\rho}_e - 3\bar{P}_e}{x} = \frac{2z^3}{\pi^2} \left\{ -\kappa_e^2 \frac{x}{z} F_1^+ \left(\frac{\kappa_e x}{z} \right) \right.$$
$$\left. + \frac{dz}{dx} \left[\kappa_e^2 \frac{x^2}{z^2} F_1^+ \left(\frac{\kappa_e x}{z} \right) + F_2^+ \left(\frac{\kappa_e x}{z} \right) \right] \right\}, \quad (25)$$

where the functions $F_1^\pm$ and $F_2^\pm$ are defined, respectively, as

$$F_1^\pm(\tau) \equiv \int_0^\infty d\omega\, \omega^2 \frac{\exp(\sqrt{\omega^2 + \tau^2})}{(\exp(\sqrt{\omega^2 + \tau^2}) \pm 1)^2},$$
$$F_2^\pm(\tau) \equiv \int_0^\infty d\omega\, \omega^4 \frac{\exp(\sqrt{\omega^2 + \tau^2})}{(\exp(\sqrt{\omega^2 + \tau^2}) \pm 1)^2}, \quad (26)$$

with the signs $\pm$ that take into account the boson and fermion nature of the particle, respectively. A similar expression holds for muons, with $\kappa_e \to \kappa_\mu$. For each pion species i, one has instead

$$\frac{d\bar{\rho}_i}{dx} - \frac{\bar{\rho}_i - 3\bar{P}_i}{x} = \frac{z^3}{2\pi^2} \left\{ -\kappa_i^2 \frac{x}{z} F_1^- \left(\frac{\kappa_i x}{z} \right) \right.$$
$$\left. + \frac{dz}{dx} \left[\kappa_i^2 \frac{x^2}{z^2} F_1^- \left(\frac{\kappa_i x}{z} \right) + F_2^- \left(\frac{\kappa_e x}{z} \right) \right] \right\}. \quad (27)$$

For sterile neutrinos ($\kappa_s \equiv m_s/m$),

$$\frac{d\bar{\rho}_s}{dx} - \frac{\bar{\rho}_s - 3\bar{P}_s}{x} = \frac{1}{\pi^2} \int_0^\infty dy\, y^2 \left[\sqrt{\kappa_s^2 x^2 + y^2} \frac{\partial f_{\nu_s}(x,y)}{\partial x} \right.$$
$$\left. + f_{\nu_s}(x,y) \left(\frac{\kappa_s^2 x + y^2/x - (\kappa_s^2 x + y^2/x)}{\sqrt{\kappa_s^2 x^2 + y^2}} \right) \right]$$
$$= \frac{1}{\pi^2} \int_0^\infty dy\, y^2 \sqrt{\kappa_s^2 x^2 + y^2} \frac{\partial f_{\nu_s}(x,y)}{\partial x}. \quad (28)$$

For active quasimassless neutrinos, if $\alpha = e, \mu, \tau$, one has

$$\frac{d\bar{\rho}_{\nu_a}}{dx} - \frac{\bar{\rho}_{\nu_a} - 3\bar{P}_{\nu_a}}{x} = \frac{d\bar{\rho}_{\nu_a}}{dx} = \sum_\alpha \frac{1}{\pi^2} \int_0^\infty dy\, y^3 \frac{\partial f_{\nu_a}}{\partial x}. \quad (29)$$

Down to a few MeV, the active neutrinos are coupled to the rest of the plasma, which means that at sufficiently high temperatures (low x) we can write $f_{\nu_a} = 1/(\exp(y/z) + 1)$ and

$$\frac{\partial f_{\nu_a}}{\partial x} = \frac{dz}{dx} \frac{\partial f_{\nu_a}}{\partial z} = \frac{dz}{dx} \frac{y \exp(y/z)}{z^2 (\exp(y/z) + 1)^2} \quad \text{(early times)}. \quad (30)$$

As a result, in computing the $z = z(x)$ relation we can save considerable computer time by considering two different regimes: Eq. (30) for $x < x_d$, while numerically computing f_{ν_a} from the Boltzmann Eq. (9) for $x > x_d$, where x_d represents any epoch before neutrino decoupling, but otherwise arbitrary. In terms of the step function Θ, collecting all terms for photons, electrons, pions, and active and sterile neutrinos and isolating dz/dx, Eq. (13) can be written as

$$\left[z^3 \left(\frac{4\pi^2}{15} + \frac{A(x/z)}{\pi^2} \right) + \frac{3\Theta(x_d - x)}{\pi^2 z^2} \int_0^\infty dy\, y^4 \frac{\exp(y/z)}{(\exp(y/z) + 1)^2} \right] \frac{dz}{dx}$$
$$= \frac{z^3}{\pi^2} B(x/z) - \frac{1}{\pi^2} \int_0^\infty dy\, y^2 \left(\sqrt{\kappa_s^2 x^2 + y^2} \frac{\partial f_{\nu_s}(x,y)}{\partial x} + y\Theta(x - x_d) \sum_\alpha \frac{\partial f_{\nu_a}}{\partial x} \right), \quad (31)$$
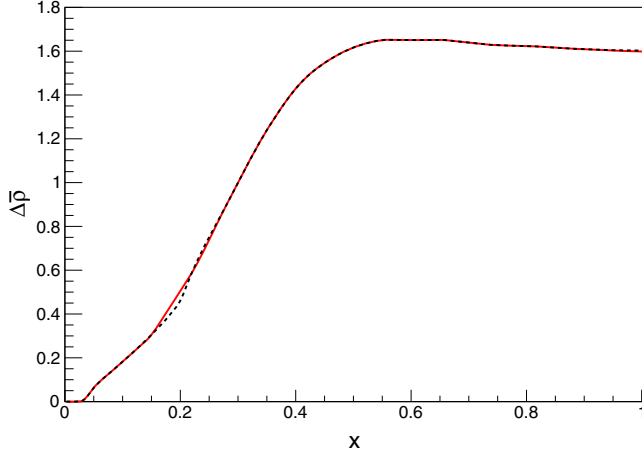
FIG. 1. Comparison between the comoving extra active neutrino energy using $x_d = 0.1$ (red curve) and $x_d = 0.2$ (black curve), for parameters $m_s = 100$ MeV and $\tau_s = 0.045$ s.

where we defined

$$A(w) = 2\sum_{\ell}(\kappa_\ell^2 w^2 F_1^+(\kappa_\ell w) + F_2^+(\kappa_\ell w))$$
$$+ \frac{1}{2}\sum_{i}(\kappa_i^2 w^2 F_1^-(\kappa_i w) + F_2^-(\kappa_i w)) \quad (32)$$

and

$$B(w) = 2\sum_{\ell}\kappa_\ell^2 w F_1^+(\kappa_\ell w) + \frac{1}{2}\sum_{i}\kappa_i^2 w F_1^-(\kappa_i w). \quad (33)$$

Together with $z = 1$ as initial condition, Eq. (31) gives the "time-temperature" evolution. Provided that x_d is sufficiently small, roughly $x_d \lesssim 0.2$ (i.e., $T \gtrsim 5$ MeV), the computed behavior is insensitive to the choice of x_d, as we illustrate in Fig. 1, where the extra comoving neutrino energy density evolution is computed using $x_d = 0.1$ (red solid curve) and $x_d = 0.2$ (black dotted curve), for parameters $m_s = 100$ MeV and $\tau_s = 0.045$ s. The results are almost equal except for small numerical differences when $0.1 < x < 0.2$. In the following, we fix $x_d = 0.1$. Note that

the parentheses on the rhs of Eq. (31) describe the "heating" of the e.m. plasma due to the sterile neutrino entropy release: At early times, all of the decay products end up in the coupled plasma of photons and active neutrinos, raising z compared to standard expectations. At late times, a sizable fraction decays into (decoupled) active neutrinos; hence, the term in parentheses largely cancels out. The only nontrivial evolution of z is then due to the finite mass term of $e^\pm$, affecting their annihilation at late time.

B. Evolution of heavy sterile neutrinos

In Ref. [11], an analytical solution of Eq. (8) was provided under the following assumptions.

(i) The equilibrium distribution functions "inside" the collisional integral [Eq. (10)] are taken in the Boltzmann (classical) approximation, with Pauli blocking factors correspondingly neglected.

(ii) Electrons are considered ultrarelativistic, with terms in m_e^2 neglected throughout.

We repeated the derivation of Ref. [11] under these approximation, finding

$$I_{\text{dec}} = \frac{(1 + \tilde{g}_L^2 + g_R^2)G_F^2 m_1^5 |U_{s\tau}|^2}{192\pi^3}\frac{m_s}{E_s}(f_s^{\text{eq}} - f_s)$$
$$\equiv \frac{m_s}{E_s}\frac{1}{\tau_s}(f_s^{\text{eq}} - f_s) \quad (34)$$

and

$$I_{\text{scatt}} = \frac{G_F^2|U_{s\tau}|^2(1 + \tilde{g}_L^2 + g_R^2)T^3 m_s^2}{\pi^3}(f^{\text{eq}}(E_1) - f(E_1))$$
$$\times \left[\frac{3}{2}\zeta(3) + \frac{7T\pi^4}{72}\left(\frac{E_1}{m_s^2} + \frac{p_1^2}{3E_1 m_s^2}\right)\right]$$
$$= \frac{3 \times 2^6}{\tau_s}(f^{\text{eq}}(E_1) - f(E_1))$$
$$\times \left[\frac{3\zeta(3)}{2}\frac{T^3}{m_s^3} + \frac{7\pi^4}{72}\frac{T^4 E_1}{m_s^5}\left(1 + \frac{p_1^2}{3E_1^2}\right)\right], \quad (35)$$

in agreement with their results quoted as

$$\partial_x f_{\nu_s}(x, y) = \frac{1.48x}{\tau_s/s}\left(\sqrt{\frac{g_*}{10.75}}\right)^{1/2}\frac{(f^{\text{eq}}(E_1) - f(E_1))}{(xT)^2}\left\{\frac{m_s}{E_s} + 3 \times 2^7 T^3\left[\frac{3}{4}\frac{\zeta(3)}{m_s^3} + \frac{7T\pi^4}{144}\left(\frac{E_1}{m_s^5} + \frac{p_1^2}{3E_1 m_s^5}\right)\right]\right\}, \quad (36)$$

where $f_{\nu_s}^{\text{eq}}$ is the Fermi-Dirac equilibrium distribution of the sterile neutrinos, τ_s is the sterile neutrino lifetime, and $E_s = \sqrt{m_s^2 + (y/x)^2}$ is the sterile neutrino energy. Details of the reduction of integrals in Eq. (10) under approximations (i) and (ii) are given in Appendix A.

We also solve the sterile neutrino kinetic equations numerically, relaxing the approximations (i) and (ii) in

Eq. (10), for both sterile neutrinos and active neutrinos. Following the well-known technique developed in Ref. [25], it is possible to analytically reduce the nine-dimensional collisional integral into a two-dimensional one, which is then integrated numerically. We developed an equivalent technique for the decay processes. We report the details in Appendix B.

TABLE III. Altered cosmologies in the presence of sterile neutrinos. τ is the lifetime of the sterile neutrino considered, T_D^a is the decoupling temperature obtained with the evolution in Eq. (36), T_D^n is the decoupling temperature obtained solving the Boltzmann equation numerically, and Y_p is the estimated value of ^{4}He abundance, discussed in Sec. IV B.

m_s [MeV]	$\sin^2 \theta_{\tau 4}$	τ[s]	T_D^n [MeV]	T_D^a [MeV]	Y_p
20.0	2.6×10^{-2}	3.0×10^{-1}	4.35	4.26	0.2514
40.0	2.8×10^{-3}	8.8×10^{-2}	9.24	10.00	0.2520
60.0	5.5×10^{-4}	6.0×10^{-2}	16.83	16.20	0.2509
80.0	1.5×10^{-4}	5.0×10^{-2}	26.53	25.22	0.2628
100.0	5.8×10^{-5}	4.4×10^{-2}	37.10	37.65	0.2705
130.0	1.6×10^{-5}	4.2×10^{-2}	59.13	59.00	0.2881

Our solutions are obtained by assuming initial thermal equilibrium for all species, starting from a temperature $T = \min[2m_s, 150 \text{ MeV}]$. To compare the difference in using numerical results vs the analytical approximation, in Table III we report the sterile neutrino freeze-out temperature T_D, for a few representative points in parameter space, according to the condition $I(T_D) = H(T_D)$, with I given by Eq. (10). We find typical differences at a few percent level and in all cases below 10%. Although we are using the numerical results in the following, when requiring only moderate precision on the sterile decoupling, the analytical approximation seems largely sufficient and allows one to significantly gain in computing time.

IV. IMPACT ON COSMOLOGICAL OBSERVABLES

After the distribution functions and temperature evolution are found, we relate them to the observables N_{eff} (notably at the CMB epoch) and Y_p (notably at the BBN epoch) to derive some constraints.

A. Impact on effective number of active neutrinos N_{eff}

Heavy ν_s affect the total energy density in nonelectromagnetic species. This is usually quantified in terms of the *effective number of neutrinos* N_{eff}, which is defined from the density in all species but electromagnetically interacting ones, as (see, for instance, [24])

$$N_{\text{eff}}(x) = \frac{\rho_\gamma^{\text{inst}}}{\rho_\gamma} \sum_{i \neq \text{e.m.}} \frac{\rho_i}{\rho_{\nu_0}}$$
$$= \left(\frac{z_0(x)}{z(x)} \right)^4 \left(3 + \frac{\Delta \rho_{\nu_e}}{\rho_{\nu_0}} + \frac{\Delta \rho_{\nu_\mu}}{\rho_{\nu_0}} + \frac{\Delta \rho_{\nu_\tau}}{\rho_{\nu_0}} + \frac{\rho_{\nu_s}}{\rho_{\nu_0}} \right),$$
(37)

where the rhs is specific for our four-neutrino model, where $\Delta \rho_{\nu_\alpha}$ are the changes in the neutrino energy densities with respect to ρ_{ν_0}, the energy density in the instantaneous decoupling limit, due to the nonequilibrium effects. Note that, at early times around $T \simeq 100 \text{ MeV}$ when all species are relativistic and share the same temperature, $N_{\text{eff}} \to 4$. Asymptotically, when all sterile neutrinos have

disappeared and the $e^+ e^-$ annihilation is complete: (i) $z_0 \to (11/4)^{1/3} \simeq 1.4$, the asymptotic Standard Model photon-neutrino temperature ratio in the instantaneous decoupling limit; (ii) $z \to z_{\text{fin}}$, the actual final photon or neutrino temperature; (iii) $\rho_{\nu_s}/\rho_{\nu_0} \to 0$, since all sterile neutrinos have decayed away. At large x, we expect $N_{\text{eff}} \gtrsim 3$, since, due to the branching fraction, the contribution due to extra radiation in the neutrino sector largely compensates the entropy transfer, which would tend to lower N_{eff} below 3 via the z-dependent prefactor in the rhs of Eq. (37). Indeed, this behavior can be seen in Fig. 2, where we have plotted the N_{eff} evolution for $m_s = 30 \text{ MeV}$, $\tau_{\tau s} = 0.15 \text{ s}$ (solid red line) and $m_s = 100 \text{ MeV}$, $\tau_{\tau s} = 0.055 \text{ s}$ (dashed black line), assuming mixing with ν_τ. Note how the more massive neutrino decays deeper in its nonrelativistic regime and well after its decoupling: The initial decline is due to the still-coupled massive neutrino, experiencing "Boltzmann suppression." Soon after, it decouples and its contribution to the plasma rises, with N_{eff} that follows this trend. The growth is somewhat less steep than naively expected in the absence of decays, since this process partially counteracts the sterile neutrino energy density relative growth. This growth turns into a sharp decline around $x \simeq 0.25$, when decay takes
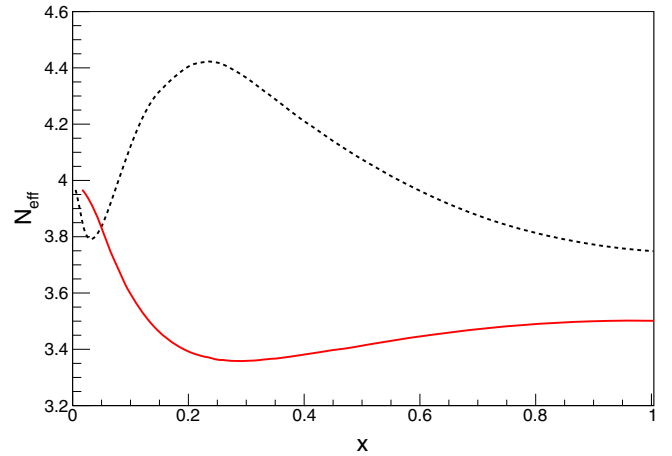
FIG. 2. N_{eff} evolution in x for $m_s = 30 \text{ MeV}$, $\tau_{\tau s} = 0.15 \text{ s}$ (red, solid line) and $m_s = 100 \text{ MeV}$, $\tau_{\tau s} = 0.055 \text{ s}$ (black, dashed line).

over. If all its entropy were transferred to the active neutrinos only, $N_{\rm eff}$ would stay constant. The partial redistribution to the e.m. plasma causes a minor decline in $N_{\rm eff}$, basically complete by $x \sim 1$. Modulo quantitative differences, the lighter and longer-lived neutrino depicted with the red line follows the same stages but shifted to the right, since it stays coupled longer and decays later.

The contribution of the sterile neutrino to this dynamics is more clearly visible in Fig. 3, while Fig. 4 shows the entropy dilution effect entering z, which partially counteracts the extra energy density in neutrinos and also affects $N_{\rm eff}$. All other parameters being the same, the effect on z is more pronounced, as expected due to the larger branching ratio in e.m. species. Note that a further, "standard" enhancement in z, due to $e^{\pm}$ annihilation, happens at $x \gg 1$ and is not visible in the figure. In Fig. 5, we show the evolution of the active neutrino energy densities. Note
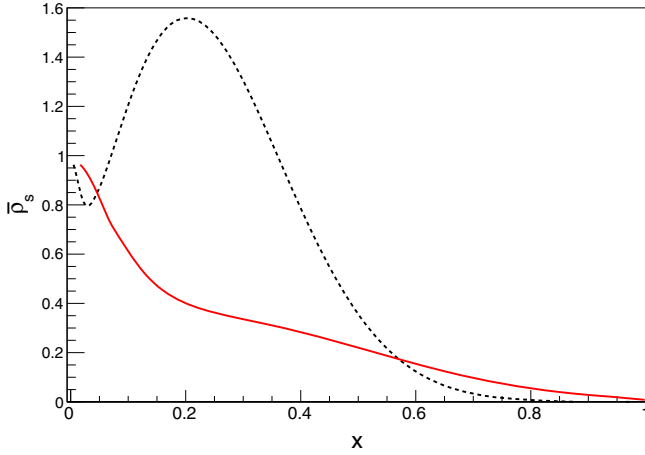
FIG. 3. Comoving energy density evolution for the sterile neutrino vs x for $m_s = 30$ MeV, $\tau_{\tau s} = 0.15$ s (red, solid line) and $m_s = 100$ MeV, $\tau_{\tau s} = 0.055$ s (black, dashed line).
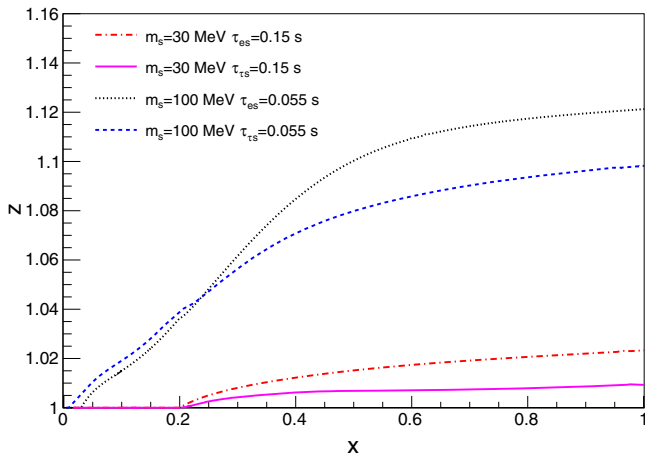

FIG. 4. Evolution of the dimensionless temperature z vs x for $m_s = 30$ MeV, $\tau_{es} = 0.15$ s (red, dash-dotted line) $\tau_{\tau s} = 0.15$ s (purple, solid line) and $m_s = 100$ MeV, $\tau_{es} = 0.055$ s (blue, dashed line), and $\tau_{\tau s} = 0.055$ s (black, dotted line).
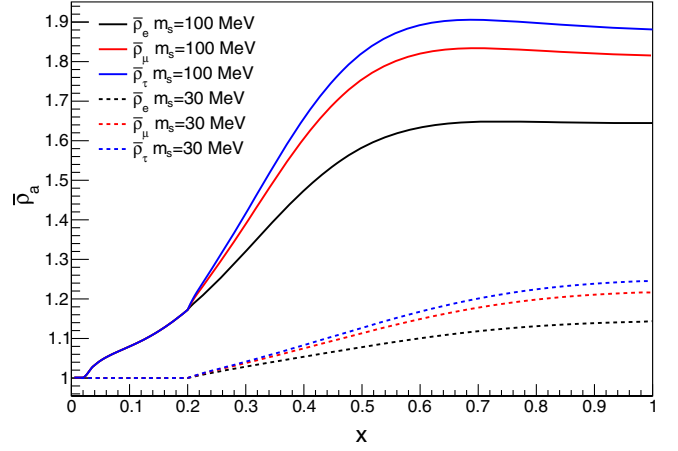

FIG. 5. Comoving energy density evolution for the three active neutrinos vs x for $m_s = 30$ MeV, $\tau_{\tau s} = 0.15$ s (dashed lines) and $m_s = 100$ MeV, $\tau_{\tau s} = 0.055$ s (solid lines). Each sets of curves, from top to bottom, represents the ν_τ, ν_μ, and ν_e energy density evolution, respectively, reflecting the assumed sterile mixing with ν_τ and the active neutrino mixing matrix. Here, $x_d = 0.2$ is assumed.

how they already depart from equilibrium somewhat by $x \sim 0.2$ (for $m_s = 100$ MeV), when the major enhancement happens due to the bulk of the decays.

B. Impact on Y_p

Another important parameter affected by a massive sterile neutrino scenario is the Y_p value, i.e., a proxy for the primordial ^{4}He mass fraction. The effect arises due to the modified expansion history, i.e., via the role that $N_{\rm eff}$ and $z(x)$ play in the Hubble function $H(x)$, but above all via the distortions to the electron (anti)neutrino distribution entering the isospin changing reactions between neutrons and protons. In Fig. 6, we illustrate the typical ν_τ and ν_e spectra (ν_μ being intermediate between the two) associated to heavy sterile neutrino decays, in the case of mixing with ν_τ: Despite mixing among the active species, the largest distortion remains in the ν_τ species; also, heavier neutrinos lead to more energetic residual distortions. Alterations due to heavy sterile neutrino decays also affect other nuclei such as deuterium, but these are subleading compared to the effect on Y_p (see, e.g., Fig. 3 in Ref. [18]), and for simplicity we will limit ourselves to model the modifications on Y_p. Both CMB and BBN are sensitive to Y_p, but astrophysical determinations of Y_p and, thus, a comparison with primordial nucleosynthesis predictions are currently more constraining.

A precise standard model calculation $Y_{p,\rm SM}^{\rm prec}$ for the best-fit cosmological parameter $\Omega_b h^2 = 0.02225$ [26] with a number of subtle effects included (see [27,28] for details) is available from PArthENoPE [29,30]. This result is rescaled via the ratio of the Born estimate of the Y_p for the ν_s model, $Y_{p,\nu_s}^{\rm Born}$, over the Born standard model calculation $Y_{p,\rm SM}^{\rm Born}$, as
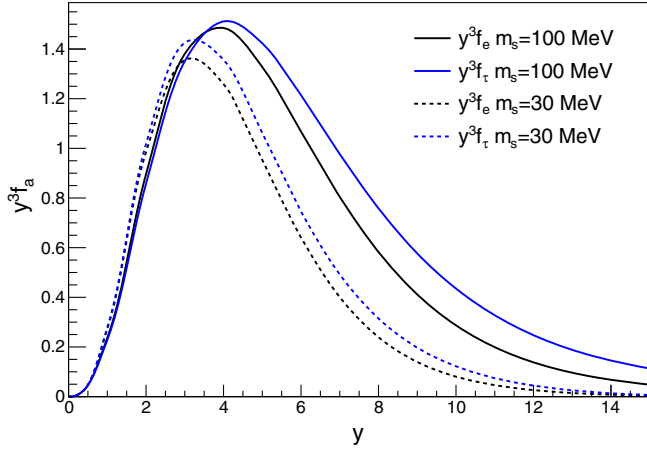
FIG. 6. ν_τ and ν_e distribution functions vs y at $x = 1$ for the same cases reported in Fig. 5.

$$Y_p = Y_{p,\text{SM}}^{\text{prec}} \frac{Y_{p,\nu_s}^{\text{Born}}}{Y_{p,\text{SM}}^{\text{Born}}}. \tag{38}$$

Each term of the fraction on the rhs can be estimated as (see, e.g., [31])

$$Y_p = 2X_n(t_{\text{on}})e^{-t_{\text{on}}/\tau_n}, \qquad X_n = \frac{n_n}{n_p + n_n}, \tag{39}$$

$$\frac{dX_n}{dx} = \frac{[\omega_B(p \to n)(1 - X_n) - \omega_B(n \to p)X_n]}{xH}, \tag{40}$$

where $t_{\text{on}} \simeq 180$ s corresponds to the onset of the BBN (i.e., deuterium bottleneck opening around $T \simeq 0.08$ MeV),[4] τ_n is the neutron lifetime, and ω_B are the rates in the Born approximation of the processes in Table IV (with $\Delta \equiv m_n - m_p \simeq 1.29$ MeV) that can be written as

$$\omega_B = \frac{G_F^2(C_V^2 + 3C_A^2)}{2\pi^3} \int_0^\infty dp\, p^2 q_0^2 \theta(q_0) F. \tag{41}$$

The distributions entering in F, as well as $H(x)$, are taken from the numerical solutions of our system of equations. In the last column in Table III, we show the results for some relevant sterile neutrino parameters.

---

[4]Note that, while strictly speaking t_{on} is altered in the nonstandard scenario considered here, the bulk of the change in Y_p comes from the prefactor $\propto X_n$ in Eq. (39). A simple estimate yields the scaling $t_{\text{on}} \simeq 180\text{s}\sqrt{H_{\text{SM}}/H}$, where H_{SM} and H are the Hubble parameter values at the beginning of BBN in the Standard Model and the case under exam, respectively. For a typical allowed modification of $N_{\text{eff}} = 3.2$, the effect on t_{on} is 2%, propagating to a 0.3% effect on Y_p, about one order of magnitude below the 2σ observational error on Y_p considered in the following; see Eq. (42). As a consequence, neglecting the change of t_{on} does not lead to appreciably different results.

TABLE IV. Relevant quantities for $n - p$ reactions.

Process	F	q_0
$\nu_e + n \to e^- + p$	$f_\nu(q_0)(1 - f_e(p_0))$	$-\Delta + p_0$
$e^- + p \to \nu_e + n$	$f_e(p_0)(1 - f_\nu(q_0))$	$-\Delta + p_0$
$e^+ + p \to \bar\nu_e + p$	$f_e(p_0)(1 - f_\nu(q_0))$	$\Delta + p_0$
$\bar\nu_e + p \to e^+ + n$	$f_\nu(q_0)(1 - f_e(p_0))$	$\Delta + p_0$
$n \to e^- + \bar\nu_e + p$	$(1 - f_\nu(q_0))(1 - f_e(p_0))$	$\Delta - p_0$
$e^- + \bar\nu_e + p \to n$	$f_\nu(q_0)f_e(p_0)$	$\Delta - p_0$

V. CONSTRAINTS AND FORECASTS

In order to obtain constraints on heavy sterile neutrinos, we compare our results on N_{eff} and on the modification on Y_p with both the latest CMB and BBN measurements. For BBN, we use the current bound at 2σ [32]:

$$Y_p = 0.245 \pm 0.006. \tag{42}$$

Concerning CMB, if limiting oneself to N_{eff}, the latest measurements of the Planck Collaboration provide a value $N_{\text{eff}} = 2.99 \pm 0.17$ [26]. Therefore, we could exclude at 2σ extra radiation leading to $\Delta N_{\text{eff}} > 0.33$. In practice, the massive sterile neutrino model under consideration here leads to changes in both N_{eff} and Y_p, and the CMB is sensitive to both (albeit much less to Y_p than BBN, at the moment). Hence, we infer the CMB constraints using a reduced Gaussian likelihood matrix involving N_{eff} and Y_p, of the form [18]

$$\chi_{\text{CMB}}^2 = (\Theta - \Theta_{\text{obs}})\Sigma_{\text{CMB}}^{-1}(\Theta - \Theta_{\text{obs}})^T, \tag{43}$$

$$\Theta = (N_{\text{eff}}, Y_p), \tag{44}$$

$$\Theta_{\text{obs}} = (2.97, 0.246), \tag{45}$$

$$\Sigma_{\text{CMB}} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 \end{pmatrix}, \tag{46}$$

$$(\sigma_1, \sigma_2) = (0.2650, 0.0177), \tag{47}$$

$$\rho_{12} = -0.845. \tag{48}$$

To obtain results at 2σ, we have to consider a value of $\chi^2 = 6.18$, value obtained by requiring that the integral of the χ^2 distribution with two d.o.f. is equal to 0.9545. Our results from CMB measurements and from BBN based on Y_p are shown in Fig. 7 for the constraints on the decay time and in Fig. 8 for the constraints on the mixing parameter, for the most interesting case of mixing with ν_τ. As a side result and in order to allow for cross-checks with past literature, in Figs. 9 and 10 we also report the corresponding results for a mixing with ν_e. Besides the current constraints, we also show the sensitivity forecast of the future CMB-S4 observations, with uncertainties $(\sigma_1, \sigma_2) = (0.062, 0.0053)$ according to Ref. [33], considering the same Θ_{obs} as in Eq. (45).
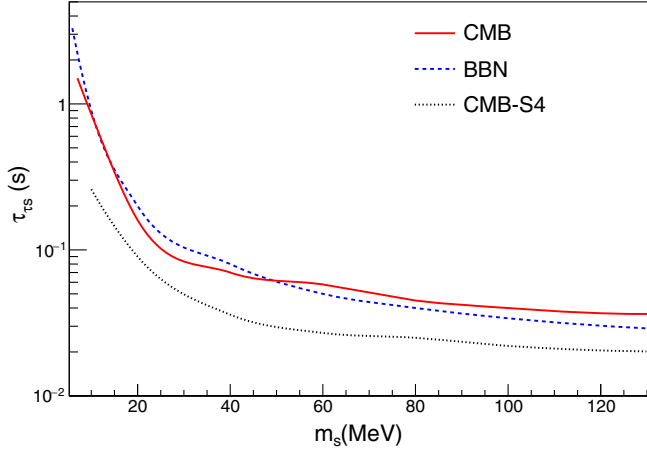
FIG. 7.   Bounds in the plane (m_s, τ_s) obtained from CMB (red curve) and BBN-Y_p (blue curve), as well as forecast sensitivity of CMB-S4 (black curve), for a sterile neutrino mixed with ν_τ (or ν_μ). The 2σ excluded region is the one above the curves.
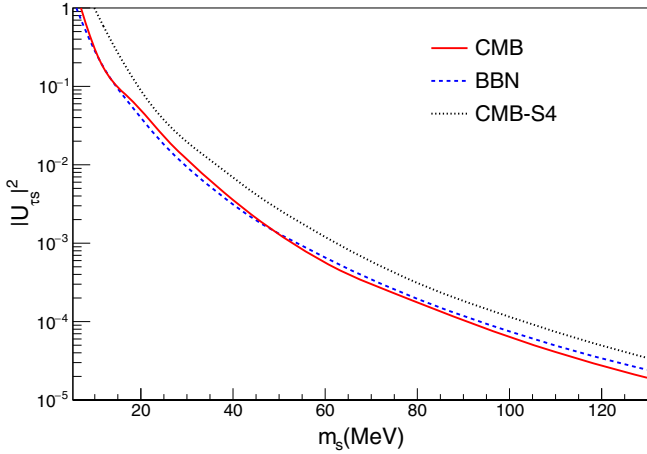

FIG. 9.   Bounds in the plane (m_s, τ_s) obtained from CMB (red curve) and BBN-Y_p (blue curve), as well as forecast sensitivity of CMB-S4 (black curve), for a sterile neutrino mixed with ν_e. The 2σ excluded region is the one above the curves.

FIG. 8.   Bounds in the plane $(m_s, \theta_{\tau s})$ obtained from CMB (red curve) and BBN-Y_p (blue curve), as well as forecast sensitivity of CMB-S4 (black curve), for a sterile neutrino mixed with ν_τ (or ν_μ). The 2σ excluded region is the one under the curves.

FIG. 10.   Bounds in the plane (m_s, θ_{es}) obtained from CMB (red curve) and BBN-Y_p (blue curve), as well as forecast sensitivity of CMB-S4 (black curve), for a sterile neutrino mixed with ν_e. The 2σ excluded region is the one under the curves.

We conclude that the CMB provides already the best constraints for $m_s \lesssim 50$ MeV, while BBN takes over at larger masses. However, we expect that CMB-S4 will attain leading constraining power in the whole range of parameter space considered here, if performing close to expectations. Qualitatively similar considerations apply for mixing with ν_e, although the transition mass is around $m_s \sim 20$ MeV, and the future improvement of CMB-S4 over BBN is less significant. Note that, in particular, for the CMB, at the same mass and lifetime the bounds are more stringent for a mixing with ν_τ than one with ν_e: This is due to the fact that the bound is dominated by $N_{\rm eff}$ and, due to the larger b.r. in neutrinos for the case of mixing with ν_τ, the growth of neutrino density via nonthermal injection is only mildly compensated by the entropy effect. For the case of mixing with ν_e, there is instead a substantial compensation via the growth of z.
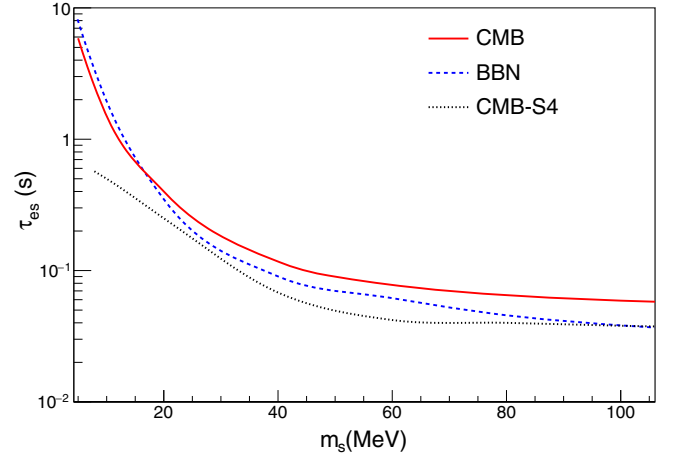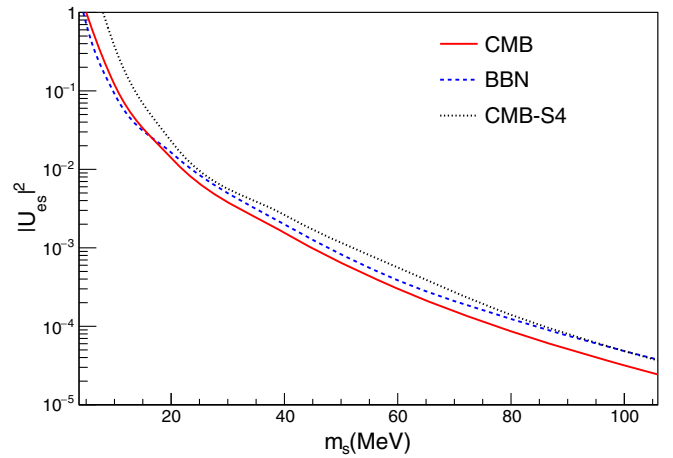
In the case of the BBN bound, however, the leading effect is due to ν_e distortions, which are larger when the mixing is with ν_e; the effect of $N_{\rm eff} > 3$ altering H is, however, more relevant when the mixing is with ν_τ, so that the two constraints are closer to each other in this case.

Our BBN constraints are largely consistent with recent calculations presented in Ref. [18], while our CMB constraints are consistent with theirs at low masses while more stringent than theirs at high masses. A more detailed comparison and discussion are reported in Appendix C, where we identify the origin of the difference in the estimate of $\Delta N_{\rm eff}$. In particular, contrarily to the results of Ref. [18], we always obtain $\Delta N_{\rm eff} > 0$ for the parameter space of interest. We have supplemented our numerical calculations with a qualitative study of the Boltzmann

equation in the analytical approximation of Ref. [11] to further support our conclusions.

VI. CONCLUSIONS

Heavy sterile neutrinos with masses $\mathcal{O}(\text{MeV} - \text{GeV})$ are predicted in extensions of the Standard Model such as the neutrino minimal Standard Model (νMSM). Besides affecting collider and supernovae observables, their mass and mixing angle parameters can be also constrained with cosmological observables, notably CMB and BBN.

We have numerically studied the evolution of sterile neutrinos with $10 \text{ MeV} \lesssim m_s \lesssim 135 \text{ MeV}$ in the early Universe and set constraints on the mixing angles or lifetimes using the N_{eff} and Y_p observables. In order to achieve these results, we have solved the exact Boltzmann equation for sterile and active neutrino evolution while taking into account the temperature evolution of electrons and photons. Also, we checked the correctness of analytical approximations in the literature and verified that they are adequate to describe sterile neutrino decoupling at better than 10% level.

For the least constrained (and, thus, phenomenologically most interesting) sector of mixing with ν_τ, at $m_s \gtrsim 50$ MeV these cosmological bounds surpass the traditional benchmark of 0.1 s lifetime often considered in the literature, up to about 0.03 s for the highest masses considered. While currently CMB is more constraining at low masses and BBN dominates at high masses, we expect the future CMB-S4 experiments to yield the dominant constraining power by the end of the decade, unless the systematic error affecting the astrophysical determinations of Y_p can be significantly reduced.

APPENDIX A: REDUCTION OF THE DECAY AND COLLISIONAL INTEGRALS

In this Appendix, we show how, under the approximations adopted in Ref. [11], we can analytically solve the integrals appearing on the rhs in Eq. (8), namely,

$$I_{\text{coll}} + I_{\text{dec}} = \frac{1}{2E} \int \prod_i \left(\frac{d^3 p_i}{2E_i(2\pi)^3} \right) \prod_f \left(\frac{d^3 p_f}{2E_f(2\pi)^3} \right)$$
$$\times (2\pi)^4 \delta^{(4)} \left(\sum_i p_i - \sum_f p_f \right) |M_{fi}|^2 F(f_i, f_f),$$
(A1)

with $|M_{fi}|^2$ the sum of the squared-matrix elements for the decay and scattering processes and

$$F(f_i, f_f) = -\prod_i f_i \prod_f (1 - f_f) + \prod_i (1 - f_i) \prod_f f_f. \quad \text{(A2)}$$

Please note that our results are obtained with a numerical integration of the collisional integrals in Eq. (A1), after the dimensional reduction recapped in Appendix B, not with the analytical approximations. Also, none of the approximations reported below is new and can be skipped unless one is interested in reproducing the analytical approximations. However, for pedagogical purposes, we thought it is useful to report them here in great detail, not to force readers to go back to the decades-old original literature.

1. Decay integral

Following the approximations of Ref. [11], we assume that in Eq. (A1) the energy distributions are represented by Maxwell-Boltzmann distributions instead of Fermi-Dirac ones. Moreover, we neglect the Pauli blocking factor, assuming $(1 - f_i) \simeq 1$. Then one gets

$$I_{\text{dec}} = \frac{1}{2E_1(2\pi)^5} \int \frac{d^3 p_2}{2E_2} \frac{d^3 p_3}{2E_3} \frac{d^3 p_4}{2E_4} |M_{fi}|^2$$
$$\times [-f_1 + e^{-E_2/T} e^{-E_3/T} e^{-E_4/T}] \delta^{(4)}(p_1 - p_2 - p_3 - p_4),$$
(A3)

where the label 1 indicated the sterile neutrino and $|M_{fi}|^2$ is the sum over the dominant decay processes

$$\nu_s \to \nu_\tau \nu_\alpha \bar{\nu}_\alpha,$$
$$\nu_s \to \nu_\tau e^+ e^-,$$

with $\alpha = e, \mu, \tau$. Performing the integral over $d^3 p_4$ in Eq. (A3) using the delta function enforcing $E_1 = E_2 + E_3 + E_4$, one obtains that $e^{-E_2/T} e^{-E_3/T} e^{-E_4/T} = e^{-E_1/T} = f_1^{\text{eq}}$. Moreover, using the property of the delta function in Eq. (B1), we can write

$$I_{\text{dec}} = (f_1^{\text{eq}} - f_1) \frac{1}{2E_1 (2\pi)^5} \int \frac{d^3 p_2}{2E_2} \frac{d^3 p_3}{2E_3} |M_{fi}|^2 \delta((p_1 - p_2 - p_3)^2) = (f_1^{\text{eq}} - f_1) \frac{m_1}{E_1} \Gamma_D, \quad \text{(A4)}$$

where the sterile neutrino decay rate Γ_D is given by [10]

$$\Gamma_D \equiv \frac{1}{2m_1 (2\pi)^5} \int \frac{d^3 p_2}{2E_2} \frac{d^3 p_3}{2E_3} \frac{d^3 p_4}{2E_4} |M_{fi}|^2 \delta(p_1 - p_2 - p_3 - p_4)$$

$$= \frac{1}{2m_1 (2\pi)^5} \int \frac{d^3 p_2}{2E_2} \frac{d^3 p_3}{2E_3} |M_{fi}|^2 \delta((p_1 - p_2 - p_3)^2) = \frac{(1 + \tilde{g}_L^2 + g_R^2) G_F^2 m_1^5 |U_{s\tau}|^2}{192\pi^3}. \quad \text{(A5)}$$

Thus, we end up with

$$I_{\text{dec}} = \frac{m_s}{E_s} \frac{1}{\tau_s} (f_s^{\text{eq}} - f_s), \quad \text{(A6)}$$

where $\tau_s = 1/\Gamma_D$ is the sterile neutrino lifetime.

We could also write the corresponding source term for an active species at the same level of approximation (i.e., neglecting the blocking factor and the inverse decay) following the treatment detailed in Ref. [13], which we direct the reader to for details. This yields

$$Hx \frac{df_a}{dx}(x, y) \simeq \sum_i \frac{B_i}{\tau_s} \int d\cos\theta \int_0^\infty dy_s \frac{y_s^2}{y^2} (f_s(x, y_s) - f_s^{\text{eq}}(x, y_s)) \mathcal{F}_{a,i} \left(\frac{y}{\gamma(1 + \beta\cos\theta)}, \cos\theta \right), \quad \text{(A7)}$$

where B_i is the branching ratio of the ith exclusive reaction, $\gamma = \sqrt{1 + (y_s m)^2 / (x m_s)^2}$, $\beta = y_s / \sqrt{(m_s x/m)^2 + y_s^2}$, and $\mathcal{F}_{a,i}$ is the double-differential distribution (with respect to y and to the angular variable θ) of the daughter particle a in the reaction i in the sterile neutrino rest frame, normalized to 1. The integral kernel in the integral above acts as a constraint picking the "right" momentum for the daughter neutrino, weighting it for the occupation factor of the parent sterile species. In practice, we always use the full numerical integration rather than this approximate expression.

2. Collisional integral

In order to evaluate the collisional integral in Eq. (A1), we assume below that they mix only with ν_τ. The relevant collisional processes are shown in Table II and have a squared interaction matrix given by

$$|M|^2 = 4C[(p_1 \cdot p_2)(p_3 \cdot p_4) + 2(p_1 \cdot p_4)(p_2 \cdot p_3)], \quad \text{(A8)}$$

where $C = 2^4 G_F^2 |U_{\tau4}|^2 (1 + \tilde{g}_L^2 + g_R^2)$ (remember that in our definition of $|M|^2$ we sum over all the degrees of freedom and include the average over the relevant state). We indicate with 1 the sterile neutrino state and evaluate the quantities in the center of momentum frame; since particles 2, 3, and 4 are relativistic, $\mathbf{p}_1 = -\mathbf{p}_2$, $\mathbf{p}_3 = -\mathbf{p}_4$, $E_3 = E_4$, and $E_1 + E_2 = E_3 + E_4 = 2E_3 = 2E_4$. We have

$$p_1 \cdot p_2 = E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2 = E_1 E_2 + \mathbf{p}_1^2$$

$$= E_1 E_2 + \frac{\mathbf{p}_1^2 + \mathbf{p}_2^2}{2} = E_2(E_1 + \mathbf{p}_1) \quad \text{(A9)}$$

and

$$p_3 \cdot p_4 = E_3 E_4 - \mathbf{p}_3 \cdot \mathbf{p}_4 = 2E_3 E_4 = 2E_3^2. \quad \text{(A10)}$$

Thus,

$$p_1 \cdot p_2 - p_3 \cdot p_4 = E_1 E_2 + \frac{\mathbf{p}_1^2 + \mathbf{p}_2^2}{2} - 2 \frac{(E_1 + E_2)^2}{4}$$

$$= \frac{\mathbf{p}_1^2 - E_1^2}{2} + \frac{\mathbf{p}_2^2 - E_2^2}{2} = -\frac{m_s^2}{2}. \quad \text{(A11)}$$

Hence,

$$(p_1 \cdot p_2)(p_3 \cdot p_4) = (p_1 \cdot p_2)^2 + \frac{m_s^2}{2}(p_1 \cdot p_2)$$

$$= (p_1 \cdot p_2)^2 + \frac{m_s^2}{2}[(p_3 \cdot p_2) + (p_4 \cdot p_2)]. \quad \text{(A12)}$$

Similarly,

$$p_1 \cdot p_4 - p_2 \cdot p_3 = E_1 E_4 - E_2 E_3 - \mathbf{p}_1 \cdot \mathbf{p}_4 + \mathbf{p}_2 \cdot \mathbf{p}_3$$

$$= E_4(E_1 - E_2) + \mathbf{p}_3 \cdot (\mathbf{p}_1 + \mathbf{p}_2)$$

$$= \frac{(E_1 + E_2)}{2}(E_1 - E_2) + \mathbf{0}$$

$$= \frac{(E_1^2 - E_2^2)}{2} = \frac{m_s^2}{2}. \quad \text{(A13)}$$

Hence,

$$(p_1 \cdot p_4)(p_2 \cdot p_3) = (p_2 \cdot p_3)^2 + \frac{m_s^2}{2}(p_2 \cdot p_3). \tag{A14}$$

These equalities among Lorentz invariants hold in any frame. As a result, let us write

$$|M|^2 = 4C\left[(p_1 \cdot p_2)^2 + 2(p_2 \cdot p_3)^2 + \frac{m_s^2}{2}[3(p_3 \cdot p_2) + (p_4 \cdot p_2)]\right] \equiv I[\mathsf{A}] + I[\mathsf{B}] + I[\mathsf{C}]. \tag{A15}$$

We thus have

$$
\begin{aligned}
\frac{I[\mathsf{A}]}{4C} &= \int \frac{d^3 p_2 d^3 p_3 d^3 p_4}{(2\pi)^9 2^4 E_1 E_2 E_3 E_4}(p_1 \cdot p_2)^2 \delta(p_1 + p_2 - p_3 - p_4)(2\pi)^4 (f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) \\
&= \int \frac{d^3 p_2 d^3 p_3}{(2\pi)^5 2^3 E_1 E_2 E_3}(p_1 \cdot p_2)^2 \delta((p_1 + p_2 - p_3)^2)(f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) \\
&= \int \frac{d^3 p_2 d^3 p_3}{(2\pi)^5 2^3 E_1 E_2 E_3}(p_1 \cdot p_2)^2 \delta(m_s^2 + 2 p_1 \cdot p_2 - 2 p_2 \cdot p_3 - 2 p_3 \cdot p_1)(f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) \\
&= \int \frac{d^3 p_2 d^3 p_3}{(2\pi)^5 2^4 E_1 E_2 E_3}(p_1 \cdot p_2)^2 (f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) \\
&\qquad \delta\left(\frac{m_s^2}{2} + (p_1 \cdot p_2) - E_2 E_3(1 - \cos\theta_{32}) - E_1 E_3(1 - v_1 \cos\theta_{31})\right) \\
&= \int \frac{d^3 p_2 dE_3 d\cos\theta_{31} d\mu_3 E_3}{(2\pi)^5 2^4 E_1 E_2}(p_1 \cdot p_2)^2 \frac{1}{E_2(1 + \cos\theta_{31}) + E_1(1 - v_1 \cos\theta_{31})} \\
&\qquad \delta\left(E_3 - \frac{(p_1 \cdot p_2) + m_s^2/2}{E_2(1 + \cos\theta_{31}) + E_1(1 - v_1 \cos\theta_{31})}\right)(f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) \\
&= \int \frac{d^3 p_2}{(2\pi)^4 2^4 E_1 E_2}\left[(p_1 \cdot p_2)^3 + (p_1 \cdot p_2)^2 \frac{m_s^2}{2}\right](f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) \\
&\qquad \times \int \frac{d\mu_3}{2\pi} \int d\cos\theta_{31} \frac{1}{[E_2(1 + \cos\theta_{31}) + E_1(1 - v_1 \cos\theta_{31})]^2} \\
&= \int \frac{d^3 p_2}{(2\pi)^4 2^4 E_1 E_2} \frac{2(p_1 \cdot p_2)^3 + (p_1 \cdot p_2)^2 m_s^2}{2(p_1 \cdot p_2) + m_s^2}(f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) \\
&\quad \int \frac{d^3 p_2}{(2\pi)^4 2^4 E_1 E_2}(p_1 \cdot p_2)^2 (f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2), \tag{A16}
\end{aligned}
$$

where θ and μ are the azimuth and polar angle, respectively, $\cos\theta_{32} = -\cos\theta_{31}$ due to the fact we evaluated the integral in the center of momentum frame, and we used Eq. (A29), with $a = E_1 + E_2$ and $b = E_2 - |\mathbf{p}_1|$, as well as

$$\delta(\kappa x) = \frac{1}{|\kappa|}\delta(x). \tag{A17}$$

Let us evaluate the integral over particle 2 (expressed in terms of Lorentz invariants) by the explicit replacement $p_1 \cdot p_2 = E_1 E_2 - p_1 p_2 \cos\theta_{12}$:

$$
\begin{aligned}
\frac{I[\mathsf{A}]}{4C} &= (f^{\mathrm{eq}}(E_1) - f(E_1)) \int \frac{dE_2 d\cos\theta_{12} d\mu_2}{(2\pi)^4 2^4 E_1} E_2^3 (E_1 - |\mathbf{p}_1| \cos\theta_{12})^2 f^{\mathrm{eq}}(E_2) \\
&= \frac{(f^{\mathrm{eq}}(E_1) - f(E_1))}{(2\pi)^3 2^4 E_1} 2\left(E_1^2 + \frac{|\mathbf{p}_1|^2}{3}\right) \int dE_2 E_2^3 f^{\mathrm{eq}}(E_2). \tag{A18}
\end{aligned}
$$

Similarly,

$$\frac{I[\mathsf{B}]}{8C} = \int \frac{d^3 p_2 d^3 p_3 d^3 p_4}{(2\pi)^9 2^4 E_1 E_2 E_3 E_4} (p_2 \cdot p_3)^2 \delta(p_1 + p_2 - p_3 - p_4)(2\pi)^4 (f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2)$$

$$= \int \frac{d^3 p_2}{(2\pi)^4 2^4 E_1 E_2} \frac{(p_1 \cdot p_2)^2}{3} (f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2) = \frac{1}{3} \frac{I[\mathsf{A}]}{4C} \Rightarrow I[\mathsf{B}] = \frac{2}{3} I[\mathsf{A}]. \tag{A19}$$

Finally, taking into account that the integrals for the C term are symmetric under the relabeling $3 \leftrightarrow 4$, we have with a similar procedure of Eq. (A16)

$$\frac{I[\mathsf{C}]}{8Cm_s^2} = \int \frac{d^3 p_2 d^3 p_3 d^3 p_4}{(2\pi)^9 2^4 E_1 E_2 E_3 E_4} (p_2 \cdot p_3)\delta(p_1 + p_2 - p_3 - p_4)(2\pi)^4 (f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2)$$

$$= \int \frac{d^3 p_2}{(2\pi)^4 2^4 E_1 E_2} \frac{(p_1 \cdot p_2)}{2} (f^{\mathrm{eq}}(E_1) - f(E_1)) f^{\mathrm{eq}}(E_2). \tag{A20}$$

Hence,

$$\frac{I[\mathsf{C}]}{8Cm_s^2} = (f^{\mathrm{eq}}(E_1) - f(E_1)) \int \frac{dE_2 d\cos\theta_{12} d\mu_2}{(2\pi)^4 2^4} \frac{E_2^2}{2}(1 - v_1 \cos\theta_{12}) f^{\mathrm{eq}}(E_2)$$

$$= \frac{(f^{\mathrm{eq}}(E_1) - f(E_1))}{(2\pi)^3 2^4} \int dE_2 E_2^2 f^{\mathrm{eq}}(E_2). \tag{A21}$$

Summing all contributions, we have

$$\frac{I_{\mathrm{coll}}}{C} = \frac{(f^{\mathrm{eq}}(E_1) - f(E_1))}{(2\pi)^3} \left[\frac{5}{6} \left(E_1 + \frac{|\mathbf{p}_1|^2}{3E_1} \right) \int dE_2 E_2^3 f^{\mathrm{eq}}(E_2) + \frac{m_s^2}{2} \int dE_2 E_2^2 f^{\mathrm{eq}}(E_2) \right]. \tag{A22}$$

Assuming in this last step Fermi-Dirac distribution for particle 2, following the procedure in Ref. [11],

$$\frac{I_{\mathrm{coll}}}{C} = \frac{(f^{\mathrm{eq}}(E_1) - f(E_1))}{(2\pi)^3} \left[\left(E_1 + \frac{|\mathbf{p}_1|^2}{3E_1} \right) \frac{7\pi^4}{144} T^4 + m_s^2 \frac{3\zeta(3)}{4} T^3 \right], \tag{A23}$$

which implies

$$I_{\mathrm{coll}} = G_F^2 |U_{\tau 4}|^2 (1 + \tilde{g}_L^2 + g_R^2)(f^{\mathrm{eq}}(E_1) - f(E_1)) T^3 \left[\left(E_1 + \frac{|\mathbf{p}_1|^2}{3E_1} \right) \frac{7\pi^4}{72} T + \frac{3\zeta(3)}{2\pi^3} m_s^2 \right]. \tag{A24}$$

The above result agrees with that reported in Ref. [11] in the same limit:

$$I_{\mathrm{coll}} = \frac{4G_F^2 |U_{s\tau}|^2 (1 + \tilde{g}_L^2 + g_R^2) T^3 m_s^2}{\pi^3} (f^{\mathrm{eq}}(E_1) - f(E_1)) \left[\frac{3}{4}\zeta(3) + \frac{7T\pi^4}{144} \left(\frac{E_1}{m_s^2} + \frac{p_1^2}{3E_1 m_s^2} \right) \right]$$

$$= \frac{3 \times 2^6}{\tau_s} (f^{\mathrm{eq}}(E_1) - f(E_1)) \left[3\zeta(3) \frac{T^3}{m_s^3} + \frac{7\pi^4}{36} \frac{T^4 E_1}{m_s^5} \left(1 + \frac{p_1^2}{3E_1^2} \right) \right], \tag{A25}$$

with τ_s given in Eq. (6).

Finally, we report below some relations used in the numerous integrations:

$$\int \frac{d^3 p}{2E} \cdots = \int \frac{d^3 p dp_0}{2E} \delta(p^0 - E) = \int d^4 p \delta(p \cdot p - m^2)\Theta(p^0)\cdots \tag{A26}$$

so that, when $p_1^2 = m_s^2$ and $p_2^2 = p_3^4 = p_4^2 = 0$,

$$\int \frac{d^3 p_4}{2E_4} F(p_1, p_2, p_3, p_4)\delta(p_1 + p_2 - p_3 - p_4)$$

$$= \int d^4 p_4 \delta(p_4^2)\delta(p_1 + p_2 - p_3 - p_4)F(p_1, p_2, p_3, p_4)$$

$$= F(p_1, p_2, p_3, p_1 + p_2 - p_3)$$
$$\times \delta(m_s^2 + 2p_1 \cdot p_2 - 2p_1 \cdot p_3 - 2p_2 \cdot p_3). \quad (A27)$$

Also, we used some notable integrals:

(i)

$$\int dx \frac{1}{[a+bx]^2} = -\frac{1}{b(a+bx)} + \text{const}, \quad (A28)$$

implying that

$$\int_{-1}^{+1} dx \frac{1}{[a+bx]^2} = \frac{2}{a^2 - b^2}. \quad (A29)$$

(ii)

$$\int dx \frac{(1+x)^2}{[a+bx]^4}$$
$$= -\frac{a^2 + a(3bx+b) + b^2(3x^2 + 3x + 1)}{3b^3(a+bx)^3} + \text{const},$$
$$(A30)$$

implying that

$$\int_{-1}^{+1} dx \frac{(1+x)^2}{[a+bx]^4}$$
$$= \frac{8}{3(a-b)(a+b)^3} = \frac{8}{3} \frac{1}{(a^2-b^2)(a+b)^2}. \quad (A31)$$

(iii)

$$\int dx \frac{(1+x)}{[a+bx]^3} = -\frac{a+2bx+b}{2b^2(a+bx)^2} + \text{const}, \quad (A32)$$

implying that

$$\int_{-1}^{+1} dx \frac{(1+x)}{[a+bx]^3} = \frac{2}{(a-b)(a+b)^2}. \quad (A33)$$

APPENDIX B: NUMERICAL REDUCTION OF THE DECAY AND COLLISIONAL INTEGRALS

In this Appendix, we show how the integrals in Eq. (A1) can be reduced from nine to three dimensions (four in the case of the decay processes) using the procedure reported in Ref. [25]. Although the procedure is not new, we recall it here for completeness.

Using the property

$$\frac{d^3 p_4}{2E_4} = d^4 p_4 \delta(p_4^2 - m_4^2)\Theta(p_4^0), \quad (B1)$$

the integral over p_4 is done using the delta in Eq. (A1). For the scattering processes, we obtain

$$p_4 = p_1 + p_2 - p_3. \quad (B2)$$

Introducing the following angles:

$$\cos(\alpha) = \frac{\mathbf{p_1} \cdot \mathbf{p_2}}{p_1 p_2}, \quad (B3)$$

$$\cos(\theta) = \frac{\mathbf{p_1} \cdot \mathbf{p_3}}{p_1 p_3}, \quad (B4)$$

$$\cos(\alpha') = \frac{\mathbf{p_2} \cdot \mathbf{p_3}}{p_2 p_3} = \cos\alpha\cos\theta + \sin\alpha\sin\theta\cos\beta, \quad (B5)$$

we can write

$$d^3 p_2 = p_2^2 dp_2 d\cos\alpha d\beta, \quad (B6)$$

$$d^3 p_3 = p_3^2 dp_3 d\cos\theta d\mu, \quad (B7)$$

with β and μ the azimuthal angles for $\mathbf{p_2}$ and $\mathbf{p_3}$, respectively. The integration over $d\beta$ is carried out using the δ function:

$$p_4^2 - m_4^2 = f(\beta). \quad (B8)$$

We use the relation for the δ:

$$\int d\beta \delta(f(\beta)) = \sum_i \int d\beta \frac{1}{|df(\beta)/d\beta|_{\beta=\beta_i}} \delta(\beta - \beta_i), \quad (B9)$$

where the β_i are the roots of $f(\beta) = 0$. Using the previously introduced angles

$$\frac{df(\beta)}{d\beta} = 2p_2 p_3 \sin\alpha\sin\theta\sin\beta, \quad (B10)$$

$\sin\beta_i$ is found as $\pm(1 - \cos^2\beta_i)^{1/2}$, where

$$\cos\beta_i = \frac{2E_2 E_3 - 2p_2 p_3 \cos\alpha\cos\theta - Q - 2E_1 E_2 + 2p_1 p_2 \cos\alpha + 2E_1 E_2 - 2p_1 p_3 \cos\theta}{2p_2 p_3 \sin\alpha\sin\theta} \quad (B11)$$

and $Q \equiv m_1^2 + m_2^2 + m_3^2 - m_4^2$. The equation for $\cos \beta$ has two solutions, but we can account for them by multiplying by 1 and using as integration interval $[0, \pi]$. The limits of integration in $d \cos \alpha$ come from demanding that $\cos^2 \beta \leq 1$, meaning that

$$(2p_2 p_3 \sin \alpha \sin \theta \sin \beta)^2 \geq 0. \tag{B12}$$

This is the same requirement that $(df(\beta)/d\beta)^2 \geq 0$. Therefore, we can write

$$\int_0^{2\pi} d\beta \delta(f(\beta)) = 2\frac{1}{|df(\beta)/d\beta|_{\beta=\beta_i}^2} \Theta\left(\left|\frac{df(\beta)}{d\beta}\right|_{\beta=\beta_i}^2\right). \tag{B13}$$

Introducing the following definitions:

$$\gamma = E_1 E_2 - E_1 E_3 - E_2 E_3;$$
$$\epsilon = p_1 p_3 \cos \theta;$$
$$k = p_1^2 + p_3^2;$$
$$a = p_2^2(-4k + 8\epsilon);$$
$$b = p_2(p_1 - \epsilon/p_1)(8\gamma + 4Q + 8\epsilon);$$
$$c = -4\gamma^2 - 4\gamma Q - Q^2 - 8\gamma\epsilon - 4Q\epsilon - 4\epsilon^2$$
$$+ 4p_2^2 p_3^2 (1 - \cos \theta)^2,$$

the derivative can be written as

$$|\frac{df(\beta)}{d\beta}|_{\beta=\beta_i} = \sqrt{a \cos^2 \alpha + b \cos \alpha + c}. \tag{B14}$$

All possible matrix elements include only products of the four-momenta. All the products are analytically integrable over $d \cos \alpha$ and can be carried out by using these relations:

$$\int \frac{1}{\sqrt{ax^2 + bx + c}} \Theta(ax^2 + bx + c)dx = \frac{\pi}{\sqrt{-a}}\Theta(b^2 - 4ac);$$

$$\int \frac{x}{\sqrt{ax^2 + bx + c}} \Theta(ax^2 + bx + c)dx = -\frac{b}{2a}\frac{\pi}{\sqrt{-a}}\Theta(b^2 - 4ac);$$

$$\int \frac{x^2}{\sqrt{ax^2 + bx + c}} \Theta(ax^2 + bx + c)dx = \left(\frac{3b^2}{8a^2} - \frac{c}{2a}\right)\frac{\pi}{\sqrt{-a}}\Theta(b^2 - 4ac).$$

The step function comes from demanding a real integration interval. This also ensures that the roots of $ax^2 + bx + c$ are not outside the fundamental integration interval of $[-1, 1]$. Integration over $d\mu$ is trivial, because there is no dependence on this parameter.

All the possible products of these momenta are calculated below:

$$p_1 \cdot p_2 = E_1 E_2 - p_1 p_2 \cos \alpha,$$
$$p_1 \cdot p_3 = E_1 E_3 - p_1 p_3 \cos \theta,$$
$$p_1 \cdot p_4 = m_1^2 + (E_1 E_2 - p_1 p_2 \cos \alpha) - (E_1 E_3 - p_2 p_3 \cos \theta),$$
$$p_2 \cdot p_3 = (E_1 E_2 - p_1 p_2 \cos \alpha) - (E_1 E_3 - p_1 p_3 \cos \theta) + \frac{Q}{2},$$
$$p_2 \cdot p_4 = (E_1 E_3 - p_1 p_3 \cos \theta) + m_2^2 - \frac{Q}{2},$$
$$p_3 \cdot p_4 = (E_1 E_2 - p_1 p_2 \cos \alpha) - m_3^2 + \frac{Q}{2}.$$

To integrate over $d \cos \theta$, the solutions of $b^2 - 4ac$ are important for the integration interval:

$$\cos \theta = \frac{-2\gamma - 2p_2^2 - Q \pm 2p_2\sqrt{2\gamma + p_1^2 + p_2^2 + p_3^2 + Q}}{2p_1 p_2}. \tag{B15}$$

If there is to be a real integration interval, both of these solutions must be real, and we will refer to them as $\cos \theta_{min}$ and $\cos \theta_{max}$. The real integration limits are $\alpha = \sup[-1, \cos \theta_{min}]$ and $\beta = \inf[+1, \cos \theta_{max}]$ with $\alpha \leq \beta$. Finally, with these conditions, it is possible to calculate numerically the collision integral left:

$$C_{coll}[f] = \frac{2}{(2\pi)^4}\frac{1}{2E_1}\int_0^\infty \int_0^{p_1+p_2}\int_\alpha^\beta \frac{p_2^2 dp_2}{2E_2}\frac{p_3^2 dp_3}{2E_3}d\cos\theta F(f_i, f_f)\Lambda(p_1, p_2, p_3)\Theta(A), \tag{B16}$$

where A is the parameter space allowed, Λ comes from the following analytical integral:

$$\Lambda(p_1, p_2, p_3) \equiv \int \frac{|M|^2}{\sqrt{a\cos^2\alpha + b\cos\alpha + c}}$$
$$\Theta(a\cos^2\alpha + b\cos\alpha + c)d\cos\alpha, \quad (B17)$$

and $F(f_i, f_f)$ is the expression defined in Eq. (11) and evaluated for p_4 given in Eq. (B2).

Finally, for the decay processes, it is possible to follow a similar procedure modulo the fact that we do not integrate analytically over $\cos\alpha$ and we define

$$p_4 = p_1 - p_2 - p_3, \quad (B18)$$

$$\gamma_{\text{dec}} \equiv E_2 E_3 - E_2 E_1 - E_3 E_1, \quad (B19)$$

$$b_{\text{dec}} \equiv 4p_2(-p_1 + \epsilon/p_1)(2\gamma + Q + 2e). \quad (B20)$$

APPENDIX C: COMPARISON WITH PREVIOUS RESULTS

To validate our code, we have first compared our results with pioneering results obtained in Ref. [11], finding excellent agreement with respect to several outputs; see, e.g., Fig. 11 showing the comoving density of the sterile species.

The most recent constraints on heavy decaying sterile neutrinos have been reported in Ref. [18]. In Figs. 12 and 13, we compare BBN bounds from Ref. [18] with our BBN bounds, while in Figs. 14 and 15 we compare CMB bounds from Fig. 11 in Ref. [18] with CMB results from our code, obtained using the same value of Θ_{Obs} as a benchmark.

While there is always a qualitative agreement, the quantitative agreement between the results is rather good only for the BBN case, while it shows some discrepancy in the CMB case for high masses. From inspection of Ref. [18] (e.g., Sec. 4.1.2, Appendix B) we infer that the
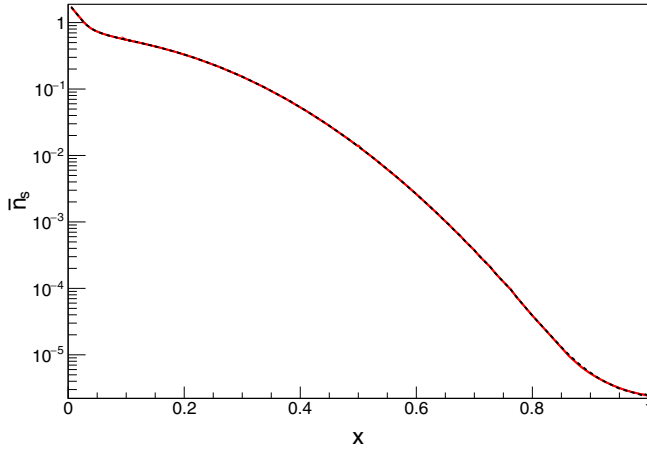

FIG. 11. Comparison between our results (red, solid line) and Ref. [11] (black, dashed line) on the comoving density of the sterile species for $m_s = 100$ MeV and $\tau_s = 0.055$ s.
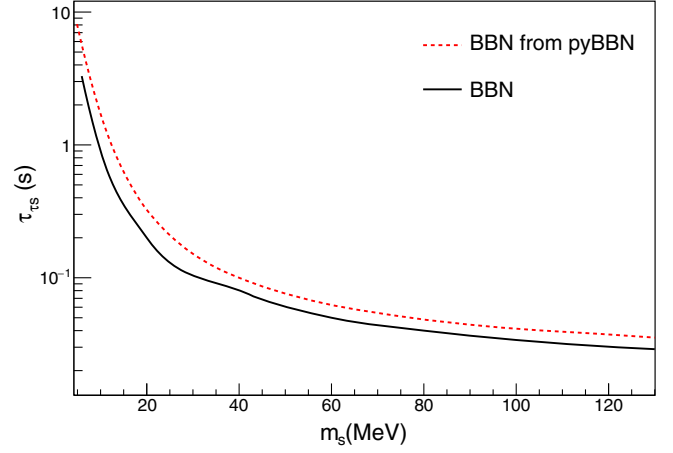

FIG. 12. Comparison of results from the BBN constraints for the decay time of a sterile neutrino mixed only with active tauonic (or muonic) neutrino in Refs. [11,18] and the one from our code.
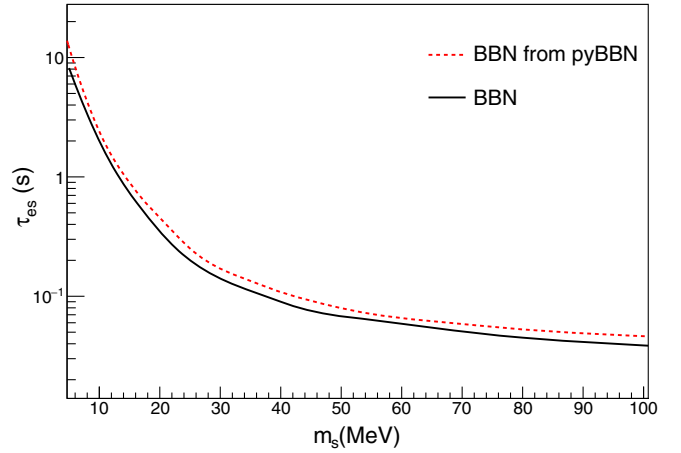

FIG. 13. Comparison of results from the BBN constraints for the decay time of a sterile neutrino mixed only with active electron neutrino in Refs. [11,18] and the one from our code.

authors find a systematically lower value of N_{eff} than us when masses are significantly large than ~ 10 MeV and that their N_{eff} can also attain values below 3, while we always find $\Delta N_{\text{eff}} \gtrsim 0$ (see, e.g., Fig. 2). This has been confirmed by private communication with the corresponding author of Ref. [18]. This behavior is also found in Ref. [17] (right panel in Fig. 3) and is indeed traced back by the authors of Ref. [18] to the solution scheme provided by the PYBBN code. As a result, in Ref. [18], CMB bounds at large masses are dominated by Y_p (to which CMB is less sensitive) rather than by N_{eff} and are weaker than ours. Since BBN bounds from Y_p depend mostly on spectral distortions of the electron-type neutrinos, it is not surprising that the agreement is much better in this observable channel. For the case of mixing with ν_τ, where N_{eff} plays a slightly bigger role, the agreement is not as excellent while remaining good. Note that the inclusion of deuterium
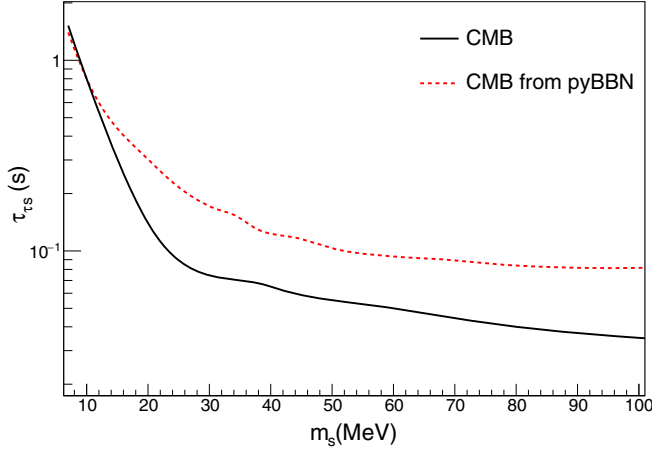
FIG. 14. Comparison of results from the CMB constraints for the decay time of a sterile neutrino mixed only with active tauonic (or muonic) neutrino in Ref. [18] and the one from our code obtained using their same value of Θ_{Obs} as a benchmark.
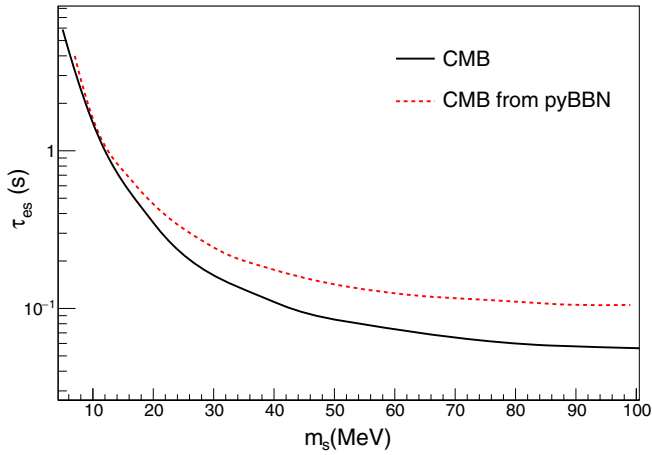

FIG. 15. Comparison of results from the CMB constraints for the decay time of a sterile neutrino mixed only with an active electron neutrino in Ref. [18] and the one from our code obtained using their same value of Θ_{Obs} as a benchmark.

constraints is subleading to the effect on Y_p (see Fig. 3 in Ref. [18]), so neglecting it has no major impact on our BBN bounds.

In the recent Ref. [34], a physical interpretation of the effect on N_{eff} found in Ref. [18] is discussed: It is claimed that $\Delta N_{\text{eff}} < 0$ is a quasigeneric outcome of injection of energy after neutrino decoupling, even for decay modes mostly in neutrinos, as a result of a dominant entropy transfer to $e^{\pm}$ via nonthermal neutrino interactions with particles of the thermal bath; see Eqs. (3)–(5) in Ref. [34]. Although the authors of Ref. [34] draw some analogy of their $\Delta N_{\text{eff}} < 0$ effect with the results of the pioneering work [35] on a low-reheating scenario, we believe that this is not very instructive, since the thermal neutrino bath is obviously suppressed if the Universe starts at temperatures

comparable or lower than the neutrino decoupling, a situation very different from the one under study.

Since we could not confirm numerically these surprising results, we thought it useful to discuss their plausibility with a qualitative study of the Boltzmann equation, in the analytical approximation derived above, matching, e.g., the one in Ref. [11]. If denoting with f_1 the nonthermal neutrino distribution under exam, it obeys an equation of the form

$$x\partial_x f_1 = \frac{I[f_1]}{H} = \frac{1}{H}\left[S(x) + \varsigma^2 G_F^2 (f^{\text{eq}} - f_1) T^4 E_1\right]. \quad (C1)$$

The first term $S \propto (f_s - f_s^{\text{eq}})/\tau_s \simeq f_s/\tau_s$ on the rhs is always positive, since it describes the injection due to decays of the sterile neutrinos. More precisely, S is an integral where $(f_s - f_s^{\text{eq}})$ enters as a kernel; see, e.g., Eq. (A7). The second (collisional) term, where ς^2 is a positive numerical constant, is initially zero, since neutrinos are at equilibrium, but as a result of the source term S, f_1 grows above the equilibrium value and leads to a negative value of the collisional term, linear in $(f^{\text{eq}} - f_1)$. A depletion of the neutrino distribution to "subthermal" values, as apparently found in Ref. [34], would imply reversing the sign of the collisional term, i.e., obtaining a *positive* value of $(f^{\text{eq}} - f_1)$. However, this second term being controlled by the source term S, it can at most grow negative to the point of compensating the first term, thus reaching an equilibrium between injection and collisional redistribution of the energy. At that moment, the derivative of f_1 is driven to zero and f_1 becomes constant. In practice, unless none of the rates is fast compared to the Hubble expansion H, the evolution follows one of the following paths.

(i) The first term on the rhs in Eq. (C1) dominates over the collisional term, which means that f_1 grows significantly larger than f^{eq}, with the scattering incapable of fully compensating it.

(ii) The second term on the rhs in Eq. (C1) is dominant: As a result, f_1 tends to f^{eq}, annihilating the collisional term or, more precisely, settling to a slightly larger-than-thermal value to compensate for the injection.

Hence, we conclude that either f_1 freezes out at a value larger than the equilibrium one (implying $\Delta N_{\text{eff}} > 0$) or in the "worst" case it tends to the equilibrium distribution ($\Delta N_{\text{eff}} \to 0^+$), in contradiction with the conclusions of Ref. [34]. In the above discussion, we neglected the effect of the growing temperature as a result of the transfer of entropy from the sterile neutrino decays: At this level of approximation, however, its effect is to make the collisional term increase via the T^4 factor, as well as to increase f^{eq}, so that the equilibrium distribution the particles are driven to is not the same as the initial one.

For the sake of the argument, let us assume that, as a result of collisions, the second term on the rhs in Eq. (C1)
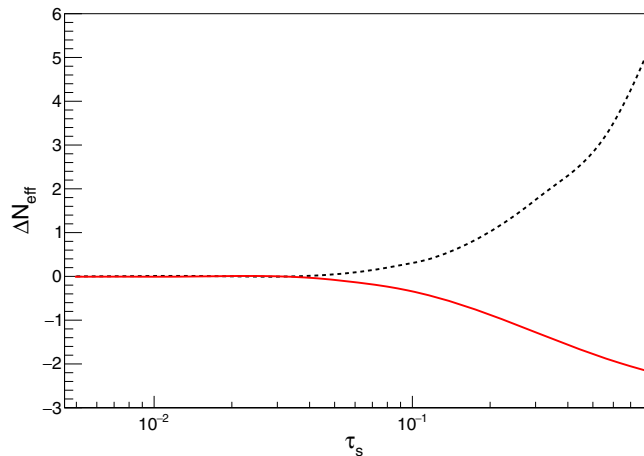
FIG. 16. ΔN_{eff} vs τ_s for $m_s = 50$ MeV in the limit of zero b.r. into active neutrino states (solid red line) and zero b.r. into electromagnetic sector (black dashed line).

starts becoming positive at some instant, i.e., $f^{\text{eq}} > f_1$: Then, f_1 would start growing again (because the rhs would be positive), and, hence, the rhs is brought closer to zero. The feedback is such that any collisionally induced depletion of neutrinos would be immediately compensated, preventing a depletion of f_1 in the circumstances under exam. For that to happen, one needs a situation in which f^{eq} grows due to the growth of temperature, with a minimal impact on f_1. This can be easily obtained if the injection of energy from decay is dominantly in the e.m. sector and the collisional terms is negligible, i.e., a situation of type (i) but in the e.m. sector, as naively expected. We checked that this is indeed the case; namely, one obtains $\Delta N_{\text{eff}} < 0$ when artificially pushing the b.r. into neutrinos to subdominant values, and the decay happens sufficiently late. These regimes are illustrated in Fig. 16: If the sterile state decays early, the collisional term is capable of restoring equilibrium and $\Delta N_{\text{eff}} \rightarrow 0$, as argued in situation (ii) described above, independently of the channel in which energy is injected. For a decay happening later and later, collisions are less and less efficient in restoring equilibrium and $\Delta N_{\text{eff}} > 0$ for a dominant b.r. into neutrino states, while $\Delta N_{\text{eff}} < 0$ for a dominant b.r. into the e.m. channel.

All our numerical results qualitative agree with these conclusions that can also be drawn from Eq. (C1). Since Eq. (C1) is an approximation, one may wonder if the results in Ref. [34] are due to some features not captured by Eq. (C1) (and, for some unknown reason, also missed by our numerical calculation). We believe that this is not the case as we argue in the following.

(i) Equation (C1) neglects quantum statistics effects, which appear, however, irrelevant to the arguments inf Ref. [34] and are anyway present for both electrons and neutrinos, without causing a qualitative change in one sector compared to the other.

(ii) Equation (C1) assumes $T_\nu = T$. However, this is not crucial to the conclusions above. The presence of a "two-temperature background" is essentially equivalent to split the term $(f^{\text{eq}} - f_1)$ into a linear combination of $(f_\nu^{\text{eq}} - f_1)$ and $(f_e^{\text{eq}} - f_1)$, each weighted by a positive factor. Both these functions are negative under the effect of neutrino energy injection: One has *a fortiori* $(f_\nu^{\text{eq}} - f_1) < 0$, since $f_\nu^{\text{eq}} < f_e^{\text{eq}}$ when $T > T_\nu$ (and electrons are relativistic), while $(f_e^{\text{eq}} - f_1) < 0$ must be satisfied if we ask for T to grow larger than T_ν as a result of injection of energy in the neutrino sector. The T evolution equation is indeed controlled by the *opposite* of the neutrino collisional term, i.e., goes as $(f_1 - f_e^{\text{eq}})$, as can be checked via Eq. (31).

(iii) Equation (C1) *does not* account for the quadratic terms involving the nonthermal parts of the neutrino distributions, i.e., depending from $(f_\nu - f_\nu^{\text{eq}})^2$, since thermal distributions have been used in the kernels of the collisional integrals; see Appendix A. Note, however, that these *are not* the processes claimed in Ref. [34] to be responsible for the effect on N_{eff}, since their Eqs. (3)–(5) explicitly indicate reactions of the nonthermal neutrinos on thermal background species, which *are included* in the above treatment. Yet, let us entertain the possibility that the results outlined by the authors of Ref. [34] are physical and due to these nonlinear effects and that it is only their interpretation or attribution to be incorrect. One should then expect that $\Delta N_{\text{eff}} < 0$ shows a "threshold" behavior with respect to the amount of energy injected in neutrinos: The less nonthermal energy is injected, the less likely these nonlinear interactions among nonthermal particles should become when compared to interactions with the thermal background. Hence, the authors of Ref. [34] should have found that the effect kicks in only above some fraction of the background energy injected in the medium, growing quadratically above this value. Instead, not only do they claim an effect also in the idealized case of "single neutrino injection" (see their Fig. 1), but clearly show an effect that is roughly linear in the injected nonthermal energy (see Fig. 7, left: The change is from slightly below −0.2% to about −0.8% when moving from an injection of 1% to 5%), inconsistent with this hypothetical explanation.

In conclusion, to the best of our knowledge, we attribute our departure from the results in Ref. [18] on the CMB constraints at large masses to some unidentified systematics, probably the same effects responsible for the "universal" $\Delta N_{\text{eff}} < 0$ outcome described in Ref. [34] which appears unphysical, for the reasons detailed above. The toy model advocated in Ref. [34] to support their findings is

also untrustworthy, since it does not account for reverse reactions (only reactions transferring energy from neutrinos into e.m. particles are included) and does not implement the physical requirement that only excess energy (above the thermal value) can be effectively transferred in collisions. As a consequence, it is, for instance, incapable of predicting equilibration and $\Delta N_{\mathrm{eff}} \to 0$ when energy is injected at early times.

[1] P. Agrawal, M. Bauer, J. Beacham, A. Berlin, A. Boyarsky, S. Cebrian, X. Cid-Vidal, D. d'Enterria, A. De Roeck, and M. Drewes *et al.*, Feebly-interacting particles: FIPs 2020 workshop report, arXiv:2102.12143.

[2] S. Schael *et al.* (ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group and SLD Heavy Flavour Group), Precision electroweak measurements on the *Z* resonance, Phys. Rep. **427,** 257 (2006).

[3] T. Appelquist and R. Shrock, Neutrino masses in theories with dynamical electroweak symmetry breaking, Phys. Lett. B **548,** 204 (2002).

[4] T. Asaka, S. Blanchet, and M. Shaposhnikov, The nuMSM, dark matter and neutrino masses, Phys. Lett. B **631,** 151 (2005).

[5] T. Asaka and M. Shaposhnikov, The nuMSM, dark matter and baryon asymmetry of the universe, Phys. Lett. B **620,** 17 (2005).

[6] S. Alekhin *et al.*, A facility to search for hidden particles at the CERN SPS: The SHiP physics case, Rep. Prog. Phys. **79,** 124201 (2016).

[7] E. J. Chun, A. Das, S. Mandal, M. Mitra, and N. Sinha, Sensitivity of lepton number violating meson decays in different experiments, Phys. Rev. D **100,** 095022 (2019).

[8] J. Orloff, A. N. Rozanov, and C. Santoni, Limits on the mixing of tau neutrino to heavy neutrinos, Phys. Lett. B **550,** 8 (2002).

[9] P. Coloma, P. Hernández, V. Muñoz, and I. M. Shoemaker, New constraints on heavy neutral leptons from Super-Kamiokande data, Eur. Phys. J. C **80,** 235 (2020).

[10] A. D. Dolgov, S. H. Hansen, G. Raffelt, and D. V. Semikoz, Cosmological and astrophysical bounds on a heavy sterile neutrino and the KARMEN anomaly, Nucl. Phys. **B580,** 331 (2000).

[11] A. D. Dolgov, S. H. Hansen, G. Raffelt, and D. V. Semikoz, Heavy sterile neutrinos: Bounds from big bang nucleosynthesis and SN1987A, Nucl. Phys. **B590,** 562 (2000).

[12] G. M. Fuller, A. Kusenko, and K. Petraki, Heavy sterile neutrinos and supernova explosions, Phys. Lett. B **670,** 281 (2009).

[13] L. Mastrototaro, A. Mirizzi, P. D. Serpico, and A. Esmaili, Heavy sterile neutrino emission in core-collapse supernovae: Constraints and signatures, J. Cosmol. Astropart. Phys. 01 (2020) 010.

[14] T. Rembiasz, M. Obergaulinger, M. Masip, M. Á. Pérez-García, M. Á. Aloy, and C. Albertus, Heavy sterile

neutrinos in stellar core-collapse, Phys. Rev. D **98,** 103010 (2018).

[15] M. Drewes and B. Garbrecht, Combining experimental and cosmological constraints on heavy neutrinos, Nucl. Phys. **B921,** 250 (2017).

[16] V. Domcke, M. Drewes, M. Hufnagel, and M. Lucente, MeV-scale seesaw and leptogenesis, J. High Energy Phys. 01 (2021) 200.

[17] O. Ruchayskiy and A. Ivashko, Restrictions on the lifetime of sterile neutrinos from primordial nucleosynthesis, J. Cosmol. Astropart. Phys. 10 (2012) 014.

[18] N. Sabti, A. Magalich, and A. Filimonova, An extended analysis of heavy neutral leptons during big bang nucleosynthesis, J. Cosmol. Astropart. Phys. 11 (2020) 056.

[19] A. Boyarsky, M. Ovchynnikov, O. Ruchayskiy, and V. Syvolap, Improved BBN constraints on heavy neutral leptons, arXiv:2008.00749 [Phys. Rev. D (to be published)].

[20] K. Abazajian, G. Addison, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso, M. Alvarez, A. Anderson, K. S. Arnold, C. Baccigalupi *et al.*, CMB-S4 science case, reference design, and project plan, arXiv:1907.04473.

[21] K. Bondarenko, A. Boyarsky, D. Gorbunov, and O. Ruchayskiy, Phenomenology of GeV-scale heavy neutral leptons, J. High Energy Phys. 11 (2018) 032.

[22] A. D. Dolgov, S. H. Hansen, and D. V. Semikoz, Non-equilibrium corrections to the spectra of massless neutrinos in the early universe, Nucl. Phys. **B503,** 426 (1997).

[23] S. Esposito, G. Miele, S. Pastor, M. Peloso, and O. Pisanti, Nonequilibrium spectra of degenerate relic neutrinos, Nucl. Phys. **B590,** 539 (2000).

[24] G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti, and P. D. Serpico, Relic neutrino decoupling including flavor oscillations, Nucl. Phys. **B729,** 221 (2005).

[25] S. Hannestad and J. Madsen, Neutrino decoupling in the early universe, Phys. Rev. D **52,** 1764 (1995).

[26] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641,** A6 (2020).

[27] P. D. Serpico, S. Esposito, F. Iocco, G. Mangano, G. Miele, and O. Pisanti, Nuclear reaction network for primordial nucleosynthesis: A detailed analysis of rates, uncertainties and light nuclei yields, J. Cosmol. Astropart. Phys. 12 (2004) 010.

[28] F. Iocco, G. Mangano, G. Miele, O. Pisanti, and P. D. Serpico, Primordial nucleosynthesis: From precision cosmology to fundamental physics, Phys. Rep. **472,** 1 (2009).

[29] O. Pisanti, A. Cirillo, S. Esposito, F. Iocco, G. Mangano, G. Miele, and P. D. Serpico, PArthENoPE: Public algorithm evaluating the nucleosynthesis of primordial elements, Comput. Phys. Commun. **178,** 956 (2008).

[30] R. Consiglio, P. F. de Salas, G. Mangano, G. Miele, S. Pastor, and O. Pisanti, PArthENoPE reloaded, Comput. Phys. Commun. **233,** 237 (2018).

[31] S. Esposito, G. Mangano, G. Miele, and O. Pisanti, Precision rates for nucleon weak interactions in primordial nucleosynthesis and He-4 abundance, Nucl. Phys. **B540,** 3 (1999).

[32] M. Tanabashi *et al.* (Particle Data Group), Review of particle physics, Phys. Rev. D **98,** 030001 (2018).

[33] D. Baumann, D. Green, J. Meyers, and B. Wallisch, Phases of new physics in the CMB, J. Cosmol. Astropart. Phys. 01 (2016) 007.

[34] A. A. Boyarsky, M. Ovchynnikov, N. Sabti, and V. Syvolap, When FIMPs decay into neutrinos: The N_{eff} story, arXiv:2103.09831.

[35] S. Hannestad, What is the lowest possible reheating temperature?, Phys. Rev. D **70,** 043506 (2004).