# Dynamics of metastable standard model particles from long-lived particle decays in the MeV primordial plasma

Akita, K.; Baur, G.; Ovchynnikov, M.; Schetz, T.; Syvolap, V.

# Dynamics of metastable standard model particles from long-lived particle decays in the MeV primordial plasma

Kensuke Akita,[1,*] Gideon Baur,[2,3,†] Maksym Ovchynnikov,[4,2,‡] Thomas Schwetz,[2,§] and Vsevolod Syvolap[5,‖]

[1]*Department of Physics, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan*
[2]*Institut für Astroteilchen Physik, Karlsruher Institut für Technologie (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany*
[3]*Bethe Center for Theoretical Physics, Universität Bonn, D-53115 Bonn, Germany*
[4]*Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland*
[5]*Instituut-Lorentz, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands*

(Received 12 November 2024; accepted 24 February 2025; published 24 March 2025)

We investigate the cosmological impact of hypothetical unstable new physics particles that decay in the MeV-scale plasma of the early Universe. Focusing on scenarios where the decays produce metastable species such as muons, pions, and kaons, we systematically analyze the dynamics of these particles using coupled Boltzmann equations governing their abundances. Our results demonstrate that the metastable species can efficiently annihilate or interact with nucleons, often leading to their disappearance before decay. The suppression of decay significantly alters the properties of cosmic neutrinos, impacting cosmological observables like big bang nucleosynthesis and the cosmic microwave background. To support further studies, we provide two public codes: the *Mathematica* code that traces the evolution of these metastable particles, as well as the Python-based unintegrated neutrino Boltzmann solver that uses this evolution as an input and may be applied to a broad range of scenarios. We then utilize them for studying a few particular new physics models.

DOI: 10.1103/PhysRevD.111.063542

## I. INTRODUCTION AND SUMMARY

The thermal plasma of the early Universe near the epoch of neutrino decoupling, at temperatures $T \lesssim 5$ MeV [1], serves as a crucial window into potential new physics. Any new particles or interactions present during this period can leave imprints on primordial neutrinos, affecting their abundance and energy distribution. These modifications, in turn, influence key cosmological observables, including primordial nuclear abundances [2–8], the cosmic microwave background (CMB) [5,9–20], and the cosmological implications of neutrino mass [21–24].

One intriguing scenario involves the existence of hypothetical long-lived particles (LLPs), $X$, with lifetimes $\tau_X \lesssim 1$ s. These particles can decay into Standard Model (SM) species, such as neutrinos, nucleons, electromagnetic (EM) particles—$e^\pm$ and photons, and various *metastable particles*

$$Y, \bar{Y} = \mu^\pm, \pi^\pm, K^\pm, K_L. \tag{1}$$

The specific decay channels determine how LLPs influence the neutrino population and primordial nuclear abundances. A key quantity in respect of neutrinos is the effective number of neutrino species $N_{\rm eff}$, showing how much energy is stored in the neutrino sector per the energy of the EM plasma. In the absence of other relativistic particles beyond the SM, $N_{\rm eff}$ is defined as the properly weighted ratio of neutrino and photon energy densities (see later for a precise definition).

Decays into EM particles can heat the EM plasma, thereby reducing $N_{\rm eff}$. On the other hand, decays into neutrinos have effects that depend on the energy of the injected neutrinos, $E_\nu$, relative to the thermal neutrino energy $\sim 3T$. If $E_\nu \simeq 3T$, decays heat the neutrino plasma, which leads to an increase of $N_{\rm eff}$ without substantial neutrino spectral distortions [25]. However, if $E_\nu \gg 3T$, $N_{\rm eff}$ can also decrease because of nontrivial effects related to nonthermal distortions of the neutrino momentum distributions [18,20,26,27]. Examples of LLPs that decay into $Y$ particles include Higgs-like scalars [28], generic pseudoscalars such as axionlike particles with various

*Contact author: kensuke@hep-th.phys.s.u-tokyo.ac.jp
†Contact author: gbaur@uni-bonn.de
‡Contact author: maksym.ovchynnikov@cern.ch
§Contact author: schwetz@kit.edu
‖Contact author: sivolapseva@gmail.com

coupling schemes [29–33], particles coupled to quark currents like dark photons and $B-L$ mediators [34], heavy neutral leptons [35], and neutralinos.

The metastable particles $Y$ may subsequently decay into neutrinos, other $Y$s, and EM particles. As the decays are governed by weak interactions, the $Y$s' inverse lifetimes are relatively low, $\tau_Y^{-1} \sim (10^6\text{–}10^8)\ \mathrm{s}^{-1}$, exceeding the characteristic interaction rates of $Y$ with the primordial plasma. Consequently, $Y$ particles can engage in various processes before decaying. For charged $Y$ particles at temperatures $T \gtrsim 1$ keV, frequent interactions with electrons and photons transfer their kinetic energy to the EM plasma [4,14,36,37]. This was incorporated in the studies [14,15,17,18,20,38], which examined the impact of LLP decays into $Y$ particles on neutrino properties. These works generally assumed that $Y$ particles inevitably decay.

In this paper, we highlight critical aspects of $Y$ particles dynamics that have been overlooked. Specifically, before decaying, they can undergo processes that lead to their disappearance without producing neutrinos. These processes include annihilations $Y + \bar{Y} \to \mathrm{SM}$, where the antiparticle $\bar{Y}$ is similarly produced by the decaying LLP, and interactions with nucleons $Y + \mathcal{N} \to \mathcal{N}' + \mathrm{SM}$.[1] Although the instant abundances of $Y$, $\bar{Y}$, and nucleons are small, the large interaction cross sections mediated by strong or electromagnetic forces render these processes highly efficient. Depending on the temperature, their rates can compete or even significantly exceed the decay rate, potentially preventing any neutrino injection. Consequently, the properties of cosmic neutrinos are significantly altered compared to scenarios where $Y$ decays are inevitable. We provide a public *Mathematica* code, which allows us to follow the actual evolution of $Y$ particles as well as nucleon abundances in the early Universe.[2] As a result, we find that within the lifetime range $0.01\ \mathrm{s} \lesssim \tau_X \lesssim 10\ \mathrm{s}$, the effective number of relativistic degrees of freedom $N_{\mathrm{eff}}$ and the degree of neutrino spectral distortions are substantially reduced, while in presence of the charged kaons the energy distributions of neutrinos and antineutrinos may become asymmetric.

To study the impact of the dynamics of metastable particles on the cosmic neutrinos, we have developed a Boltzmann solver for the momentum-dependent neutrino distribution functions, based on Ref. [42].[3] The code allows us to calculate the evolution of neutrinos in the presence of a broad range of new physics particles such as Higgs-like scalars, Majorons, dark photons, ALPs, and electromagnetically decaying relics. The nontrivial dynamics of intermediate metastable SM particles is considered in a self-consistent way. Hence, our work extends previous studies, see, e.g., [3,5,7,17,19,20,39,44,45], and offers a flexible public tool to be used by the community. As we will show below, in light of future precise CMB measurements, the unintegrated Boltzmann approach is needed to predict $N_{\mathrm{eff}}$ even for the simplest scenarios, such as LLPs decaying purely electromagnetically. In particular, it should replace various approximate approaches to solve the neutrino Boltzmann equations (see, e.g., [8,14,15,46–48]), which may not be accurate enough.

The findings of the given study are summarized in our companion Letter [49].

The paper is organized as follows. In Sec. II, we describe the properties of the metastable particles $Y$ and their interactions in the primordial plasma. Section III discusses our two-step scheme to calculate the metastable particles and the evolution of neutrinos. Section IV describes how we calculate the dynamics of $Y$ particles and nucleon densities. In particular, in Sec. IV B, we conduct a simplified analysis for the cases of muons and charged pions, illustrating that they prefer to disappear before decaying at MeV temperatures. In Sec. V, we describe our numeric approach to solve the neutrino Boltzmann equations in the presence of metastable particles. Section VI contains a qualitative discussion on how the dynamics of $Y$s influences neutrino properties, including $N_{\mathrm{eff}}$, the neutrino distribution function, and the neutrino-antineutrino energy asymmetry. Section VII explores a few models with LLPs, such as Higgs-like scalars and heavy neutral leptons, and analyzes how they affect the neutrino population based on the methods from the previous sections. Finally, we conclude in Sec. VIII.

## II. INTERACTIONS OF METASTABLE PARTICLES IN THE PRIMORDIAL MEV PLASMA

A brief summary of the properties and interactions of the $Y$ particles is listed in Table I, while the relevant interaction diagrams are shown in Fig. 1.

We take the information about the decay modes from PDG [50], the interaction with nucleons from Ref. [36], and the results of this section for the annihilation channels. Below, we describe them in detail for each of the particles.

### A. Muons

The muon lifetime is $\tau_\mu \approx 1.2 \times 10^{-6}\ \mathrm{s}$—the largest among all the metastable particles. The only relevant decay mode is

$$\mu^- \to e^- + \bar{\nu}_e + \nu_\mu. \tag{2}$$

The neutrino decay products may have energies as large as $E_\nu \approx m_\mu/2 \simeq 50$ MeV, which well exceeds the thermal neutrino energies at MeV temperatures.

The energy loss processes are

$$\mu + \gamma \to \mu + \gamma \quad \text{and} \quad \mu + e \to \mu + e. \tag{3}$$

[1]The meson-driven $p \leftrightarrow n$ processes have been considered in the works [4,7,8,16,36,37,39] in the context of the impact of new physics on primordial nuclear abundances. However, to the best of our knowledge, they have not been included in any previous study of the impact on neutrinos.

[2]Available on [40] and Zenodo [41].

[3]Available on [43].

TABLE I. Properties of the metastable particles in the primordial plasma. The meaning of the columns is as follows: the particle, its lifetime and decay modes, annihilation modes with their corresponding thermal-averaged cross sections, and the same for the interactions with nucleons $\mathcal{N} = n,\ p$. For the thermal-averaged cross sections, we provide the values at $T = 3$ MeV.

| Particle | Decays | Annihilations | Nucleon interactions |
|---|---|---|---|
| $\mu^{\pm}$ | $\tau = 2.2 \times 10^{-6}$ s $e^{\pm}\overset{(-)}{\nu}_e\overset{(-)}{\nu}_\mu(100\%)$ | $\langle\sigma\beta\rangle \approx 6 \times 10^{-2}$ GeV$^{-2}$<br>$\gamma\gamma(25\%)$ $e^+e^-(75\%)$ | $\mu^- p \to n\nu_\mu$ $\mu^+ n \to p\bar{\nu}_\mu$<br>Subdominant |
| $\pi^{\pm}$ | $\tau = 2.6 \times 10^{-8}$ s $\mu^{\pm}\overset{(-)}{\nu}_\mu$: 100% | $\langle\sigma\beta\rangle \approx 3$–$5$ GeV$^{-2}$ $2\pi^0(\approx 100\%)$ | $\pi^- p \to nX$ $\langle\sigma\beta\rangle \approx 4$–$4.6$ GeV$^{-2}$<br>$\pi^+ n \to pX$ $\langle\sigma\beta\rangle \approx 4$ GeV$^{-2}$ |
| $K^{\pm}$ | $\tau = 1.23 \times 10^{-8}$ s $\mu\bar{\nu}_\mu(63\%)$<br>$\pi^0 l\bar{\nu}_l(8.4\%)$ $\pi^+\pi^0(20.7\%)$<br>$3\pi(7.4\%)$ | $\langle\sigma\beta\rangle \approx 44$ GeV$^{-2}$ $\pi^+\pi^-(66.6\%)$<br>$2\pi^0(33.3\%)$ | $K^- p \to \mathcal{N}2\pi$ $\langle\sigma\beta\rangle \approx 150$ GeV$^{-2}$<br>$K^- n \to \mathcal{N}2\pi$ $\langle\sigma\beta\rangle \simeq 10^2$ GeV$^{-2}$ |
| $K_L$ | $\tau = 5.116 \times 10^{-8}$ s $\pi^{\pm} l^{\mp}\nu_l(67.6\%)$<br>$3\pi(30.6\%)$ | Same as $K^{\pm}$ | $K_L p \to \mathcal{N}2\pi$ $\langle\sigma\beta\rangle \approx 42.5$ GeV$^{-2}$<br>$K_L n \to \mathcal{N}2\pi$ $\langle\sigma\beta\rangle \approx 42.5$ GeV$^{-2}$ |
| $K_S$ | $\tau = 0.89 \times 10^{-10}$ s $2\pi^0(30.7\%)$<br>$\pi^+\pi^-(69.2\%)$ | Same as $K^{\pm}$ | Same as $K_L$ |

The overall rate has the scaling

$$\Gamma_{\rm loss} = \langle\sigma^{\mu}_{\rm loss} v\rangle n_{\rm EM} \sim \frac{\alpha^2_{\rm EM}}{m_\mu E_{\rm thermal}} T^3_{\rm EM}, \tag{4}$$

where $E_{\rm thermal} \approx 3T$ is the mean energy of thermal particles. At $T = 1$–$5$ MeV, the rate is more than 9 orders of magnitude larger than the decay rate $\Gamma^{\mu}_{\rm decay} = \hbar/\tau_\mu$. Because of this, we will assume that the muons are effectively at rest. The same conclusion holds for any other charged $Y$.

The annihilation processes are

$$\mu^+ + \mu^- \to e^+ + e^- \quad \text{and} \quad \mu^+ + \mu^- \to 2\gamma. \tag{5}$$

They are thresholdless, and given that $m_\mu \gg m_e$, the thermal average $\langle\sigma\beta\rangle$ closely matches the zero-momentum cross section:

$$\langle\sigma^{\mu}_{\rm ann} v\rangle \approx \sum_{i=ee,2\gamma} (\sigma^{\mu\to i}_{\rm ann} v)_{i,p=0} \approx \frac{4\pi\alpha^2_{\rm EM}}{m^2_\mu}. \tag{6}$$

The annihilation rate $\Gamma^{\mu}_{\rm ann} = \langle\sigma^{\mu}_{\rm ann} v\rangle n_{\bar{\mu}}$ is also suppressed compared to the energy loss rate: the cross section itself is smaller than energy loss one by the ratio $m_\mu/E_{\rm thermal} \gg 1$, and the number density $n_{\bar{\mu}}$ of antimuons (produced together with the muons) is much smaller than the thermal densities. This is because the instant $\bar{Y}$ number density, entering the annihilation rate $\Gamma_{\rm ann} = n_{\bar{Y}}\langle\sigma v\rangle$, is principally bounded from above by what can be accumulated before decays. It is $n_{\bar{Y}} \lesssim n_X \frac{\tau_Y}{\tau_X} \ll n_X$ (see a discussion in Sec. IV B), and hence is much smaller than the available $X$ number density (which is itself typically much smaller than the thermal number density). Note that generically, we assume that the same amounts of $Y$ and $\bar{Y}$ particles are produced by the $X$ decays, and the above argument on the annihilation rate applies equally to the charge-conjugated case.

Muons interact with nucleons $\mathcal{N} = n,\ p$ by

$$\mu^- + p \to n + \nu_\mu, \qquad \mu^+ + n \to p + \bar{\nu}_\mu. \tag{7}$$

They are mediated by the weak force, which, together with the tiny amount of nucleons, makes the processes irrelevant [14,37].

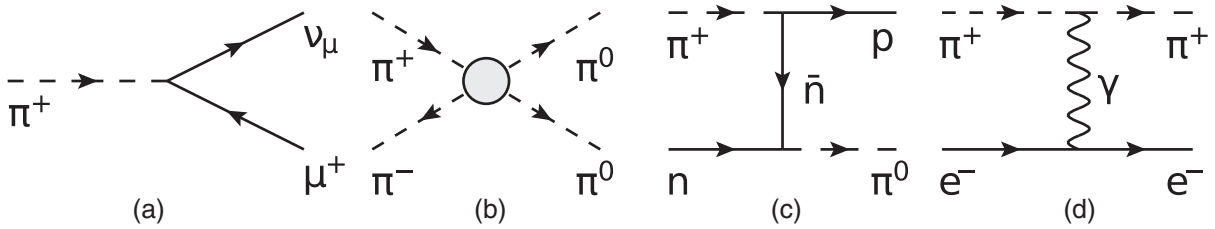


FIG. 1. Diagrams of different interaction processes with metastable particles $Y$ in the primordial plasma: decay (a), annihilation $Y + \bar{Y} \to$ SM (b), the interaction with nucleons (c), and the elastic EM scattering that leads to the deposition of the $Y$'s kinetic energy in the EM plasma (d). Processes with the pion $\pi^+$ are considered as an example. The impact of the scattering off nucleons and annihilation process is demonstrated in Figs. 2–8.

### B. Charged pions

The lifetime of the charged pion is $\tau_\pi = 2.6 \times 10^{-8}$ s, 2 orders of magnitude smaller than for the muon. The main decay mode is

$$\pi^+ \to \mu^+ + \nu_\mu \tag{8}$$

The neutrino produced by decays of the pion at rest has a monochromatic energy $E_\nu = (m_\pi^2 - m_\mu^2)/2m_\pi \approx 29.8$ MeV, which still greatly exceeds thermal neutrino energies.

The pion's energy loss rate is similar to the muon's one, being many orders of magnitude larger than the decay rate. As decaying pions inject muons, the evolution of $\pi$s and $\mu$s is coupled.

Despite the much smaller lifetime, the processes of the annihilation and the interaction with nucleons are important for the pions: the corresponding processes are driven by the strong force, which means a much larger cross section. The dominant annihilation process is[4]

$$\pi^+ + \pi^- \to 2\pi^0. \tag{9}$$

It is close to the kinematic threshold, and the kinetic energy distribution of pions makes a non-negligible contribution to the cross section. To compute it, we use the ChPT Lagrangian as implemented in [33], and then average over thermally distributed pion energies using [[25], Eq. (A.68)]. Before averaging over energies, we find

$$\sigma_{\rm ann}^{2\pi^0}\beta = \frac{(10m_{\pi^+}^2 + 12p^2 - m_{\pi^0}^2)^2\sqrt{m_{\pi^+}^2 - m_{\pi^0}^2 + p^2}}{576\pi f_\pi^4 (m_\pi^2 + p^2)^{\frac{3}{2}}}, \tag{10}$$

where $p$ is the momentum of the interacting pion in the center-of-mass frame and $f_\pi \approx 93$ MeV is the pion decay constant. The thermal averaging increases the cross section by a factor of 2 compared to the zero-momentum limit in the temperature range $T < 5$ MeV.

Let us now discuss interactions with nucleons. Since the pions are almost stopped, the most efficient processes are thresholdless. Those are [36]

$$\pi^- + p \to n + \pi^0/\gamma, \qquad \pi^+ + n \to p + \pi^0/\gamma. \tag{11}$$

The thermal cross sections behave as

$$\langle \sigma_{p\to n}^{\pi^-}\beta\rangle \approx 3.68 \cdot F_c^\pi(T)\ \text{GeV}^{-2}, \tag{12}$$

$$\langle \sigma_{n\to p}^{\pi^+}\beta\rangle \approx 1.1\langle \sigma_{p\to n}^{\pi^-} v\rangle / F_c^\pi(T), \tag{13}$$

Here,

$$F_c^X(T) = \frac{y_X}{1 - \exp[-y_X]}, \qquad y_X = 2\pi\alpha_{\rm EM}/v_{{\rm rel},pX}, \tag{14}$$

is the Sommerfeld enhancement, occurring because of the formation of a quasibound state of the oppositely charged $X$ and $p$ particles with the relative velocity $v_{{\rm rel},pX} = |\boldsymbol{v}_p - \boldsymbol{v}_X|$.

The resulting $\langle \sigma_{p\leftrightarrow n}^\pi \beta\rangle$ is comparable to $\langle \sigma_{\rm ann}^\pi \beta\rangle$.

### C. Kaons

The case of kaons is more complicated. There are four different kaons, $K^\pm, K_L, K_S$, with $K_{L/S}$ being admixtures of $K^0$ and $\bar{K}^0$. The lifetimes are $\tau_{K^\pm} \approx 1.23 \times 10^{-8}$ s, $\tau_{K_L} \approx 5.1 \times 10^{-8}$ s, and $\tau_{K_S} \approx 0.9 \times 10^{-10}$ s. All of them, except for $K_S$, have decay modes containing neutrinos. $K_S$ decays into a pair of pions; its lifetime is very small, and it does not have time to participate in any other interactions before decaying. The neutrino energy may be as large as $m_K/2$.

$K_L$s do not lose their kinetic energy before participating in any further interaction. Here and below, we will treat them as particles-at-rest for simplicity. We argue this in the following way. If including the finite energy distribution of kaons, the decay probability decreases with the $\gamma$ factor (due to time dilation). On the other hand, the probabilities of the other processes would generically increase, as we enlarge the available scattering phase space. Therefore, our approximation would overestimate the decay probability of $K_L$. However, as we study GeV-scale LLPs, the impact of these changes would not be significant, which justifies the approach.

The dominant kaon annihilation processes are

$$K^+ + K^- \to \pi^+ + \pi^-, \qquad K^+ + K^- \to 2\pi^0 \tag{15}$$

(and the same for $K_L$, $K_S$ particles). Since $m_K - m_\pi \gg 3T$, the reactions are far from threshold, and we may safely approximate their cross sections $\sigma v$ by the zero-momentum result

$$\langle \sigma_{\rm ann}^K \beta\rangle \approx \frac{\sqrt{m_K^2 - m_\pi^2}(m_d(10m_K^2 + m_\pi^2) + m_\pi^2 m_s)^2}{3072\pi f_\pi^4 m_d^2 m_K^3}, \tag{16}$$

with the numeric value $\approx$44 GeV$^{-2}$. It is a factor of 10 larger than $\langle \sigma_{\rm ann}^\pi \beta\rangle$, because the reaction is far from the threshold.

The interaction processes with nucleons $\mathcal{N}$ are much more complicated than in the pion case. The thresholdless processes exist only for $K_L, K^-$, and go via the intermediate $\Lambda/\Sigma$ resonances [36,37]

$$K^- + \mathcal{N} \to \Lambda/\Sigma + \pi \to \mathcal{N}' + 2\pi, \tag{17}$$

$$K_L + \mathcal{N} \to \Lambda/\Sigma + \pi \to \mathcal{N}' + 2\pi. \tag{18}$$

[4]The EM process, $\pi^+ + \pi^- \to 2\gamma$, although being far from the kinematic threshold, is suppressed by two orders in magnitude.

The absence of such processes for $K^+$ follows from the fact that they would require resonances with positive baryon number and strangeness, that do not exist. The asymmetry in the evolution of $K^+, K^-$ induces an asymmetry in the energy distributions of neutrinos and antineutrinos; we will return to this phenomenon in Sec. VI B.

The thermal cross sections (here assuming that $K_L$ is at rest) are [36]

$$\langle\sigma^{K^-}_{p\to n}\beta\rangle \approx 79F_c^K(T)\ \text{GeV}^{-2}, \qquad \langle\sigma^{K^-}_{n\to p}\beta\rangle \approx 66\ \text{GeV}^{-2}, \tag{19}$$

$$\langle\sigma^{K^-}_{p\to p}\beta\rangle \approx 37F_c^K(T)\ \text{GeV}^{-2}, \qquad \langle\sigma^{K^-}_{n\to n}\beta\rangle \approx 88\ \text{GeV}^{-2}, \tag{20}$$

$$\langle\sigma^{K_L}_{p\to n}\beta\rangle \approx 18\ \text{GeV}^{-2}, \qquad \langle\sigma^{K_L}_{n\to p}\beta\rangle \approx 18\ \text{GeV}^{-2}. \tag{21}$$

Here, $F_c$ is given by Eq. (14).

Kaon decays, annihilations, and interaction with nucleons inject charged pions and/or muons, which do not transfer all their energy to the EM plasma. Therefore, the evolution of $K$, $\mu$, and $\pi$ populations is coupled.

## III. TWO-STEP APPROACH

In this section, we discuss our approach to studying the evolution of the metastable particles $Y$ in the primordial plasma and their impact on neutrinos.

We assume a generic scenario when $Y$s are injected by decays of some hypothetical LLP, denoted by $X$, at MeV temperatures. We are agnostic about the origin of $X$ and parametrize its number density as

$$n_X = n_{X,0}\left(\frac{a(t_0)}{a(t)}\right)^3 \exp\left[-\frac{t-t_0}{\tau_X}\right] \tag{22}$$

Here, $n_{X,0}$ is the number density at some initial time $t_0$, $a(t)$ is the scale factor of the Universe, and $\tau_X$ is its lifetime.

Decays into $Y$s are only possible if $m_X > m_Y \gg 3T$. This means that the LLPs we consider have to be out-of-equilibrium at the temperatures of interest; otherwise, their abundance would be exponentially suppressed. As for the LLP lifetimes, our main interest is in the range $\mathcal{O}(0.01\text{–}10)$ s. On the one hand, it covers the temperatures from the beginning of the neutrino decoupling to shortly after (in ΛCDM). On the other hand, this is also the temperature range where the metastable particles may prefer to disappear without decaying.

In general, $n_{X,0}$ is an independent parameter, but for particular models with only two parameters—mass $m_X$ and $\tau_X$—it may be uniquely fixed: $n_{X,0} = n_{X,0}(m_X, \tau_X)$. In the rest of the paper, we will explore both of these scenarios to cover as broad a range of models as possible.

In order to study the dynamics of the metastable particles and neutrinos, we follow a *two-step approach*:

(1) We trace the evolution of $Y$ particles in the expanding Universe, utilizing a simplified description of the neutrino dynamics from [25] as seed; details are described in Sec. IV (see also Appendix B).
(2) We include the calculated evolution of $Y$s from step 1 in the form of time-dependent decay probabilities in the source term of the solver of the neutrino Boltzmann equation in the momentum space. Then, we carefully trace the evolution of neutrinos and the expansion of the Universe; see Sec. V and Appendix C.

This factorization is meaningful because the evolution of $Y$s is weakly affected by details of the equilibration between neutrinos and EM plasma: it is mainly sensitive to the scale factor, which is determined by the overall energy density of the Universe.

## IV. STEP I. DYNAMICS OF METASTABLE PARTICLES

### A. System of equations

Let us now construct the system of equations for the $Y$ abundances. Most of the $Y$s are charged and, therefore, effectively at rest; given this, it is adequate to consider the system of coupled integrated Boltzmann equations on their number densities.

As we have discussed in Sec. II, we have to solve the system for all $Y$s simultaneously, given that their dynamics are coupled: heavier $Y$s produce lighter ones because of decay, annihilation, or interactions with nucleons. Then, the resulting equations for the given $Y$ and its antiparticle $\bar{Y}$ take the form

$$\begin{cases} \frac{dn_Y}{dt} + 3Hn_Y = \frac{n_X}{\tau_X}N^X_Y - \frac{n_Y}{\tau_Y} - n_Y n_{\bar{Y}}\langle\sigma^Y_{\text{ann}}v\rangle + \left(\frac{dn_Y}{dt}\right)_{\mathcal{N}} + \sum\limits_{Y'\neq Y} n_{Y'}\Gamma_{Y'\to Y}, \\ \frac{dn_{\bar{Y}}}{dt} + 3Hn_{\bar{Y}} = \frac{n_X}{\tau_X}N^X_{\bar{Y}} - \frac{n_{\bar{Y}}}{\tau_Y} - n_{\bar{Y}} n_Y\langle\sigma^Y_{\text{ann}}v\rangle + \left(\frac{dn_{\bar{Y}}}{dt}\right)_{\mathcal{N}} + \sum\limits_{Y'\neq Y} n_{Y'}\Gamma_{Y'\to \bar{Y}}. \end{cases} \tag{23}$$

The meaning of the terms is as follows.

(i) The second term on the left-hand side appears due to the expansion of the Universe. $H = \sqrt{8\pi\rho/(3m_{\text{Pl}})}$ is the Hubble parameter, with the Planck mass $m_{\text{Pl}} = 1.2\times10^{19}$ GeV and the energy density of the Universe $\rho$.

(ii) $\frac{n_X}{\tau_X} N_Y^X$ is the injection from decays of $X$. Apart from direct decays, we also include secondary contributions $X \to Z \to Y$, where $Z$ are ultra short-lived particles with $\tau_Z \ll 10^{-8}$ s: $K_S, \rho^0, \eta, \omega$, etc. $N_Y^X$ is the amount of $Y$s per $X$ decay

$$N_Y^X = \sum_i \mathrm{Br}_i \cdot N_Y^i, \tag{24}$$

with $\mathrm{Br}_i$ being the branching ratio of the given decay channel $i$, and $N_Y^i$ denoting the number of $Y$s produced per this channel.

(iii) The 2nd and 3rd terms on the rhs of Eq. (23) describe direct decays and annihilations of $Y$, respectively.

(iv) $(\frac{dn_{Y/\bar{Y}}}{dt})_{\mathcal{N}}$ is the evolution due to the interaction with nucleons $\mathcal{N} = p, n$:

$$\left(\frac{dn_{Y/\bar{Y}}}{dt}\right)_{\mathcal{N}} = -n_{Y/\bar{Y}} \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma_{\mathcal{N}}^{Y/\bar{Y}} v \rangle. \tag{25}$$

The interaction processes include the $p \leftrightarrow n$ conversion as well as the processes that do not change the $\mathcal{N}$ type.

(v) The summand $\sum_{Y' \neq Y} n_{Y'} \Gamma_{Y' \to Y}$ takes into account decay, annihilation, and nucleon interaction processes involving the metastable particles $Y' \neq Y$ with $m_{Y'} > m_Y$

$$\Gamma_{Y' \to Y} = \frac{1}{\tau_{Y'}} N_Y^{Y',\mathrm{decay}} + n_{\overline{Y}} \langle \sigma_{\mathrm{ann}}^{Y'} v \rangle N_Y^{Y',\mathrm{ann}} + \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma_{\mathcal{N}}^{Y'} v \rangle N_Y^{Y',\mathcal{N}}, \tag{26}$$

with $N_Y^{Y',\mathrm{decay}}, N_Y^{Y',\mathrm{ann}}, N_Y^{Y',\mathcal{N}}$ being the amounts of $Y$ produced per given process. We calculate them using [50] for decays, [36] for the interaction with nucleons, and this work for the annihilation.

The system (23) has to be supplemented by the equations governing the evolution of EM and neutrino populations, the scale factor, and the nucleon number densities. The first two we calculate using the method from [25,51]. It assumes that throughout their evolution, the neutrinos $\nu_\alpha$ always have Fermi-Dirac shape of the energy spectrum, parametrized by a time-dependent temperature $T_{\nu_\alpha}(t)$. Under this simplification, it is possible to integrate the neutrino Boltzmann equation in the momentum space and get the system of equations on the neutrino and EM temperatures $T_{\nu_e}, T_{\nu_\mu}, T_{\nu_\tau}, T$ and scale factor $a$ in the presence of decaying LLPs. Throughout the text, we call the approach by the *integrated* method to solve neutrino Boltzmann equations, as opposed to the unintegrated method that we will consider in Sec. V (see Appendix B discussing our implementation of the integrated approach, including incorporating neutrino oscillations). When calculating the source terms for neutrinos and the EM particles within this method, we assume that all $Y$'s energy goes to the EM plasma.[5]

Knowing the resulting dynamics of the scale factor $a$ determines the evolution of the $X$'s number density (22) and the baryon-to-photon ratio

$$\eta_B(T) = \eta_{B,\mathrm{Planck}} \cdot \left(\frac{a(T_{\mathrm{CMB}}) T_{\mathrm{CMB}}}{aT}\right)^3, \tag{27}$$

where $\eta_{B,\mathrm{Planck}} = 7.06 \times 10^{-10}$ is fixed by the CMB measurements performed with Planck [52].

For the nucleon number density, we start with the definition

$$n_{\mathcal{N}}(t) \equiv n_B(t) X_{\mathcal{N}}(t) = n_\gamma \eta_B(t) \cdot X_{\mathcal{N}}(t), \tag{28}$$

where $n_B$ is the baryon number density, and $X_{\mathcal{N}} \equiv n_{\mathcal{N}}/n_B$ is the relative fraction of the given nucleon ($X_n + X_p = 1$). The latter obeys the equation

$$\frac{dX_n}{dt} = -X_n \left(\Gamma_{n\to p}^{\nu,e} + \sum_{y=Y,\bar{Y}} n_y \langle \sigma_{n\to p}^y v \rangle\right) + (1 - X_n)\left(\Gamma_{p\to n}^{\nu,e} + \sum_{y=Y,\bar{Y}} n_y \langle \sigma_{p\to n}^y v \rangle\right), \tag{29}$$

where $\Gamma_{n\leftrightarrow p}^{\nu,e}$ are rates of the weak conversion processes with neutrinos and electrons, while $n_Y \langle \sigma_{p\leftrightarrow n}^Y v \rangle$ are those driven by the $Y$ particle. The latter processes are part of the total nucleon interaction rates $\langle \sigma_{\mathcal{N}}^Y v \rangle$:

$$\langle \sigma_{\mathcal{N}}^Y v \rangle = \langle \sigma_{\mathcal{N}\to\mathcal{N}}^Y v \rangle + \langle \sigma_{\mathcal{N}\to\mathcal{N}'}^Y v \rangle. \tag{30}$$

If $Y$ is a meson, it completely dominates the evolution of $X_n$ until the instant $Y$ population is suppressed by many orders of magnitude compared to the neutrino number density [8]. This is because of two factors. First, the meson-driven conversion cross section is 16 orders of magnitude larger than the cross section of the weak conversion. Second, at MeV temperatures, the probability of $Y$ interactions with nucleons is comparable with its decay probability, so there is no *a priori* suppression. Therefore, in practice, the weak $p \leftrightarrow n$ conversion rates may be dropped from Eq. (29).

The solution for $X_n$ may be obtained by setting the right-hand-side of Eq. (29) to zero (the so-called dynamic equilibrium)[6]:

[5]By varying this assumption and checking how the resulting $Y$ dynamics change, we have explicitly verified that this assumption is not important.

[6]We have validated the dynamical equilibrium solutions for $X_n$ and $n_Y$ [Eq. (35)] by computing first the exact solutions and comparing them with the approximate solution given by the dynamic equilibrium.

$$X_n \approx \frac{\sum_y n_y \langle \sigma^y_{p\to n} v\rangle}{\sum_y n_y \langle \sigma^y_{p\to n} v\rangle + \sum_y n_y \langle \sigma^y_{n\to p} v\rangle}. \quad (31)$$

Once we solve the coupled system of equations for $\mu, \pi, K, X_n$, we may compute time-dependent probabilities to decay and disappear by annihilating or interacting with nucleons:

$$P^Y_{\rm decay}(t) = \frac{\tau_Y^{-1}}{\tau_Y^{-1} + \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma^Y_{\mathcal{N}} v\rangle + n_{\bar{Y}} \langle \sigma^Y_{\rm ann} v\rangle}, \quad (32)$$

$$P^Y_{\rm ann}(t) = \frac{n_{\bar{Y}} \langle \sigma^Y_{\rm ann} v\rangle}{\tau_Y^{-1} + \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma^Y_{\mathcal{N}} v\rangle + n_{\bar{Y}} \langle \sigma^Y_{\rm ann} v\rangle}, \quad (33)$$

$$P^Y_{\mathcal{N}}(t) = \frac{\sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma^Y_{\mathcal{N}} v\rangle}{\tau_Y^{-1} + \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma^Y_{\mathcal{N}} v\rangle + n_{\bar{Y}} \langle \sigma^Y_{\rm ann} v\rangle} \quad (34)$$

These probabilities serve as an input to calculate the impact on the neutrino and EM populations of the primordial plasma. We separate annihilations and interactions with nucleons, as the latter are very important for studying the impact of $Y$ on BBN.

Assuming that we have computed the decay probability $P^Y_{\rm decay}(t)$, the number density of $Y$s available for decays is again given by the dynamical equilibrium:

$$n_Y(t) = n_X(t) N_Y^X \frac{\tau_Y}{\tau_X} P^Y_{\rm decay}(t) \quad (35)$$

We provide the implementation of this system and its solution for generic LLPs in a *Mathematica* code (see footnote 2). Details on the code may be found in Appendix A and on the GitHub repository page.

### B. Simple estimates of $Y$ evolution

Let us make a simplified analysis that allows us to understand the impact of annihilation and interaction with nucleons. First, let us neglect the influence of $X$ particles on the Hubble expansion rate. Then, we may use the standard formula $a(t) \propto \sqrt{t}$ and $H(t) = \dot{a}/a = 1/2t$ for the radiation-dominated Universe, as well as the standard cosmological value for the baryon-to-photon ratio $\eta_B(1\ \text{MeV}) \approx 1.7 \times 10^{-9}$. Next, let us assume that various $Y$s evolve independently from each other. With all these approximations, we can still qualitatively describe the dynamics of the populations of $Y$ and its antiparticle $\bar{Y}$, while presenting results in a simple form.

Similarly to the case of $X_n$, we may solve the system (23) analytically in the regime of dynamic equilibrium, when all the processes are much faster than the Hubble expansion.[7]

Assuming $n_Y = n_{\bar{Y}}$, we get

$$n_Y = \frac{\sqrt{\frac{4 n_X \langle \sigma^Y_{\rm ann} v\rangle}{\tau_X} + (\Gamma_{\mathcal{N}} + \tau_Y^{-1})^2} - \Gamma_{\mathcal{N}} - \tau_Y^{-1}}{2\langle \sigma^Y_{\rm ann} v\rangle}, \quad (36)$$

where we have defined an effective interaction rate with nucleons as

$$\Gamma_{\mathcal{N}} \equiv \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma^Y_{\mathcal{N}} v\rangle \quad (37)$$

Now, let us analyze this solution by considering two limiting cases: $n_Y \langle \sigma^Y_{\rm ann} v\rangle \gg \Gamma_{\mathcal{N}}$, meaning that annihilations dominate over the interactions with nucleons, and $n_Y \langle \sigma^Y_{\rm ann} v\rangle \ll \Gamma_{\mathcal{N}}$, which is the opposite.

For the first case, we can estimate the relative importance of decays and annihilations by considering

$$n_Y = \frac{1}{2\langle \sigma^Y_{\rm ann} v\rangle \tau_Y} \left[\sqrt{\frac{4}{\epsilon_{\rm ann}} + 1} - 1\right], \quad (38)$$

where we have used Eq. (36) in the limit $\Gamma_{\mathcal{N}} = 0$ and defined[8]

$$\epsilon_{\rm ann} = \frac{\tau_Y^{-2}}{\frac{n_X}{\tau_X}\langle \sigma^Y_{\rm ann} v\rangle} \quad (39)$$

For the second case ($n_Y \langle \sigma^Y_{\rm ann} v\rangle \ll \Gamma_{\mathcal{N}}$), we can directly compare the decay rate to the rate of the interaction with nucleons, which are both independent of the abundance $n_Y$

$$\epsilon_{\mathcal{N}} = \frac{\tau_Y^{-1}}{\Gamma_{\mathcal{N}}} \quad (40)$$

Hence, in both cases, a small value for the ratios (39) and (40) implies that $Y$ decays are much less efficient than the competing processes (annihilations or interactions with nucleons, respectively).

Let us consider the reference choice

$$n_{X,0} = 0.1 \cdot n_{\rm UR}(T_0) = 0.1 \cdot \frac{\zeta(3)}{\pi^2} T_0^3, \quad (41)$$

where $n_{\rm UR}$ is the number density of a scalar ultrarelativistic particle in equilibrium at the given temperature, and $\tau_X = 0.03$ s. The values of the quantities (39) and (40) are shown in Table II. They clearly imply that the dynamics of stopped pions, $K^\pm$, muons, and $K_L$ may be driven not by decays but by annihilations or interactions with nucleons.

[7] Note that the form of the expression (36) differs from (35). This is because in (35) we assume that the decay probability $P^Y_{\rm decay}$ has been previously computed numerically. The latter includes $n_{\bar{Y}}$, which is tightly related to $n_Y$.

[8] $\epsilon_{\rm ann}$ may be understood in the following way. Consider an instant injection of $Y$ from $n_X$ during time $\sim\tau_Y$; during this period, decays do not deplete the $Y$ population. Then, let us assume *a priori* that the annihilation does not prevent accumulating $\bar{Y}$ during this time, so $n_{\bar{Y}} \approx \frac{n_X}{\tau_X}\tau_Y$. For the ratio of $\Gamma_{\rm decay}$ and $\Gamma_{\rm ann} = n_{\bar{Y}} \langle \sigma v\rangle_{\rm ann}$, one then gets Eq. (39).

TABLE II. Ratios (39), (40) for $T = 3$ MeV, $\tau_X = 0.05$ s, and the LLP number density given by Eq. (41). The cross sections are taken from Sec. II. We assumed for simplicity $n_p(T) \approx n_B/2 \approx \eta_B(T) \cdot n_\gamma/2$, with $\eta_B(T \gg m_e) \approx 1.7 \times 10^{-9}$.

| Particle | $\epsilon_{\rm ann}$ | $\epsilon_{\mathcal{N}}$ |
|---|---|---|
| $\mu^\pm$ | $3.4 \times 10^{-4}$ | $\gg 1$ |
| $\pi^\pm$ | $4.1 \times 10^{-2}$ | 1.15 |
| $K^-$ | $1.4 \times 10^{-2}$ | $3.4 \times 10^{-2}$ |
| $K^+$ | $1.4 \times 10^{-2}$ | $\gg 1$ |
| $K_L$ | $8.6 \times 10^{-4}$ | $6.8 \times 10^{-2}$ |
| $K_S$ | $2.8 \times 10^{2}$ | 40 |

For example, the smallness of $\epsilon_{\rm ann}$ suggests that the particle prefers to annihilate rather than decay. The exceptions are short-lived $K_S$: their tiny lifetime allows them to decay before interacting.

The impact of the scattering processes significantly depends on the number density of the interacting counterparts—$\bar{Y}$ for annihilation and baryons for the nucleon interactions. Both $n_{\bar{Y}}$ and $n_B$ are suppressed at low temperatures as $a^{-3} \sim T^3$. In addition, the $\bar{Y}$ number density, entering the annihilation rate for $Y$s, gets exponentially suppressed at times $t \gg \tau_X$, so the drop in $P^Y_{\rm ann}$ would be much faster than in $P^Y_{\mathcal{N}}$. To account for these effects, we will use the solution (36) and obtain the probabilities (32)–(34) for muons and pions. They are shown in Fig. 2 for the setup (41). For the particular parameters, decays are strongly suppressed at high temperatures but become dominant at a temperature determined by the properties of $X$. A higher $X$ number density is associated with a lowering of this temperature.

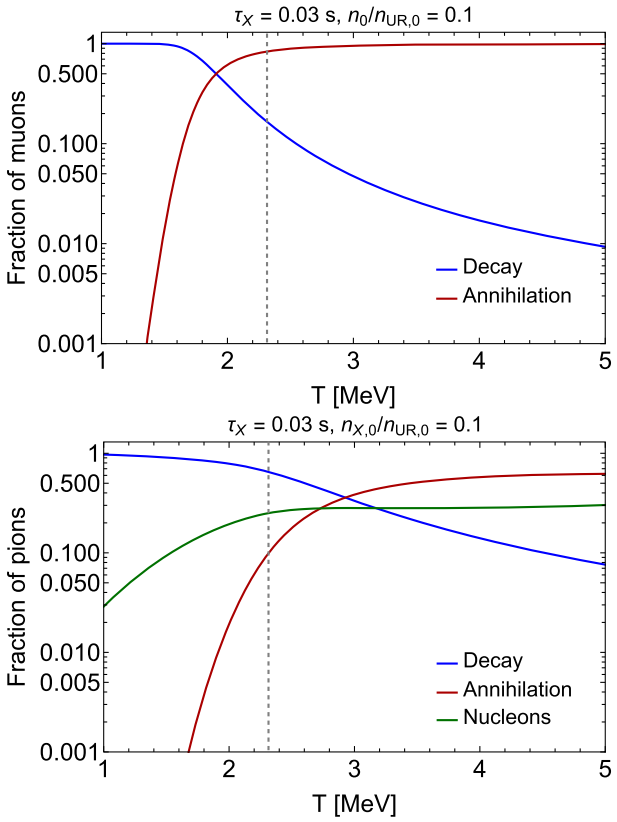


FIG. 2. The yields of muons (top) and pions (bottom) that would decay, annihilate, or interact with the nucleons [Eqs. (32)–(34)] if injected by a decaying particle $X$ with the number density given by Eq. (22) and the lifetime $\tau_X = 0.03$ s. The results are obtained using the simplified consideration presented in Sec. IV B, in order to be easily reproducible. The vertical dashed line shows the moment of time at which the comoving density of $X$ becomes 0.01 of the $n_{X,0}$, such that the dynamics of $X$ and its decay products already do not affect the Universe.

## V. STEP II. EVOLUTION OF NEUTRINOS

To trace the evolution of neutrinos and EM plasma in the presence of decaying relics, we need to solve the system of equations governing the evolution of the neutrino distribution function and the EM plasma temperature. This system must incorporate the dynamics of the metastable decay products discussed in this paper.

The most accurate way to follow the evolution of out-of-equilibrium neutrinos, including neutrino oscillations, would be solving the evolution equations for the 3-state density matrix of neutrinos, called quantum kinetic equations (QKEs) [42,53–55]. However, the accuracy of that approach is far better than needed, given the expected uncertainty of future CMB observations, and it is also computationally expensive.

At MeV temperatures, the rates of neutrino oscillations are much faster than the weak interaction rates. Because of this, one can separate the weak interactions from the oscillations and describe the neutrino evolution as quasi-classical unintegrated Boltzmann equations [1,56] on the neutrino distribution function $f_{\nu_\alpha}(p,t)$:

$$\frac{\partial f_{\nu_\alpha}}{\partial t} - Hp\frac{\partial f_{\nu_\alpha}}{\partial p} = \sum_{\beta=e,\mu,\tau} \langle P_{\alpha\beta}\rangle I_\beta[p, f_{\nu_\alpha}]. \tag{42}$$

Here, $\langle P_{\alpha\beta}\rangle(p,T)$ is the time-averaged oscillation probability depending on the temperature of the EM plasma $T$ and neutrino energy, accounting for the matter corrections to the mixing angles, expressed in terms of neutrino energy and EM plasma temperature (see Sec. 3 in Ref. [17] and Ref. [57]). $I_\alpha$ is the collision term, which describes the details of scattering and annihilation of $\nu_\alpha$, as well as the decays of LLPs

$$I_\alpha = \frac{1}{2E_{\nu_\alpha}} \sum \int \prod_{i=2} \frac{d^3p_i}{(2\pi)^3 2E_i} \prod_{f=1} \frac{d^3p_f}{(2\pi)^3 2E_f} \times S|\mathcal{M}|^2 F[f] (2\pi)^4 \delta^{(4)}\left(\sum_{i=1} p_i - \sum_{f=1} p_f\right). \tag{43}$$

The first sum runs over all possible interaction processes, including $\nu_\alpha$ as $i = 1$. The integral is performed for all

possible states for $\nu_\alpha$ with momentum $p_1$. $i, f$ denote the initial and final states for a process. $S|\mathcal{M}|^2$ is the corresponding squared matrix element times the symmetry factor $S$.[9] $F[f]$ is the quantum statistical factor to describe the population of the medium

$$F[f] = \prod_{i=1}(1 \mp f_i)\prod_{f=1} f_f - \prod_{i=1} f_i \prod_{f=1}(1 \mp f_f), \quad (44)$$

where $f_{i,f}$ are the momentum distribution for the $i$, $f$th particle. $(1 - f)$ is the Pauli-blocking factor for fermions while $(1 + f)$ is the Bose-enhancement factor for bosons.

To close the equations for the system of the plasma, it is necessary to know the evolution equation of the electromagnetic plasma (i.e., the EM temperature). It is described by the continuity equation (the energy conservation law), including the energy densities of the SM plasma and the decaying LLPs

$$\frac{d\rho}{dt} = -3H(\rho + P), \quad (45)$$

where $\rho$ and $P$ are the total energy density and pressure for the plasma, respectively. As in Ref. [42], we include thermal QED corrections in Eq. (45) following [58–61]. Further discussion on our implementation, running time, and limitations is given in Appendix C.

Our Boltzmann solver has been extensively tested against a completely independent method to calculate the evolution of neutrino distribution functions, namely the neutrino direct simulation Monte-Carlo method developed in Refs. [26,27]. Examples of the cross-checks include the evolution of the neutrino spectrum under the initial conditions of different neutrino and EM plasma temperatures, and injections of high-energy neutrinos. The agreement in the evolution of characteristic quantities, such as the energy densities, is at the level of 0.1%. Moreover, the unintegrated quantities, such as neutrino energy spectra, also agree quite well.

### A. Simple case study: Neutrino distortions matter

Using the momentum-dependent solver is necessary even in the simplest scenarios with no high-energy neutrino injections. To illustrate this point, we consider a relic $X$ decaying solely into electromagnetic particles as an example (an example of such an LLP is a light Higgs-like scalar or axionlike particle). We will compare the predictions of our code with the integrated approach to solve the neutrino Boltzmann equations we introduced in the previous Appendix B.

Figure 3 shows the quantity $N_{\rm eff}^{\rm unint} - N_{\rm eff}^{\rm int}$, i.e., the difference between the values of $N_{\rm eff}$ predicted by the unintegrated and integrated approaches, considering different mass and lifetime of the LLP. As for the initial abundance of the LLP, we use

$$n_{\rm LLP}(T = 20\ {\rm MeV}) = 0.05 \cdot n_{\rm UR}(20\ {\rm MeV}) \quad (46)$$

Naively, the purely EM decays do not disturb the thermal shape of the neutrino spectrum, and then the predictions of the unintegrated and integrated methods should match. We confirm this expectation in the lifetimes ranges $\tau \lesssim 0.02$ s (where there is no impact on $N_{\rm eff}$ at all) and $\tau \gtrsim 0.2$ s.

However, there is a sizeable deviation in the intermediate domain. Its origin is accumulating distortions in the neutrino sector during the equilibration with the EM plasma (see also the discussion in Ref. [26]). Namely, the neutrino-EM interaction rate is a growing function of the neutrino energy, so high-energy neutrinos equilibrate better. Distortions are mainly accumulated in the temperature range 2 MeV $\lesssim T \lesssim$ 5 MeV (translating to the lifetime range mentioned above) when neutrinos are already partially decoupled but still interact efficiently with the EM particles. They destroy the main approximation of the integrated approach, which leads to the discrepancy (see also [17]).

The magnitude of the deviations between the unintegrated and integrated methods may go beyond the error bars in the determination of $N_{\rm eff}$ by future CMB observations, such as Simons Observatory. Therefore, even such relatively simple scenarios with LLPs require accurate numeric studies of how the neutrino population is affected.

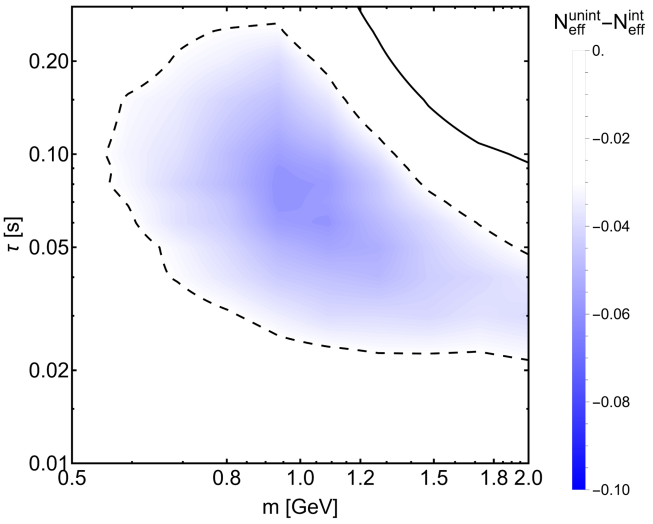


FIG. 3. The behavior of $N_{\rm eff}$ under the scenario with an LLP decaying solely into the EM particles (see text for details). The plot shows the behavior of the deviation $N_{\rm eff}^{\rm unint} - N_{\rm eff}^{\rm int}$, where $N_{\rm eff}^{\rm unint}$ has been obtained using the unintegrated method to solve the neutrino Boltzmann equation described in Sec. V, whereas $\Delta N_{\rm eff}^{\rm int}$ is calculated utilizing the integrated approach from [25] including neutrino oscillations (see Appendix B). The solid black line indicates zero deviation, whereas the dashed black line shows $N_{\rm eff}^{\rm unint} - N_{\rm eff}^{\rm int} = -0.03$.

[9]See Table 3 in Ref. [17] for the specific formula of $S|\mathcal{M}|^2$ for relevant processes of neutrinos at MeV temperature.

## VI. QUALITATIVE IMPACT OF METASTABLE PARTICLES ON NEUTRINOS AND BBN

Let us now qualitatively analyze the impact of the evolution of $Y$ particles on the dynamics of the MeV plasma. We will consider several aspects: properties of primordial neutrinos—$N_{\rm eff}$, neutrino spectral distortions, the asymmetry in the energy distribution between neutrinos and antineutrinos, and the neutron-to-proton conversion, which sets the initial condition for BBN.

The purpose of this analysis is to understand the impact of metastable particles' dynamics on the population of neutrinos in simple terms. This way, it accompanies the numerical study of the evolution of the neutrino distribution function (Sec. V), used to obtain our main results—Figs. 3, 5, and 7.

### A. $N_{\rm eff}$ and neutrino spectral distortions

The effective number of relativistic neutrino species, $N_{\rm eff}$, is defined as

$$N_{\rm eff} = \frac{8}{7}\left(\frac{11}{4}\right)^{\frac{4}{3}} \left.\frac{\rho_{\rm UR}-\rho_\gamma}{\rho_\gamma}\right|_{m_\nu \ll T \ll m_e}, \tag{47}$$

where $\rho_{\rm UR}$ and $\rho_\gamma$ represent the energy densities of ultrarelativistic particles and photons, respectively. Under the assumption that neutrinos follow an equilibrium (Fermi-Dirac) distribution, $N_{\rm eff}$ effectively characterizes the neutrino population. However, deviations from thermal equilibrium can lead to a nonthermal neutrino distribution function, $f_\nu(p,t)$ and break this degeneracy.

In the ΛCDM framework, the value of $N_{\rm eff}$ is $N_{\rm eff}^{\Lambda\rm CDM} \approx 3.04$ [42,61–67], and the neutrino distribution closely resembles a Fermi-Dirac distribution with temperature $T_\nu \approx (4/11)^{1/3} T_\gamma$. Variations in $N_{\rm eff}$ and $f_\nu(p,t)$ influence the Universe's expansion rate and the neutron-to-proton conversion rates. Specifically, energetic neutrinos can efficiently convert protons to neutrons, thereby increasing the neutron-to-proton ratio beyond the ΛCDM prediction and enhancing primordial helium abundance. Additionally, distortions break the degeneracy between the neutrino energy and number densities, which may be important in the epoch when they become nonrelativistic.

Without decays into metastable particles, there are two distinct scenarios:

(1) *LLPs decaying solely into EM particles*: In this scenario, the evolution of the neutrino population may be approximately described in terms of the evolution of its temperature [25] (see a discussion in Sec. V). The resulting deviation in the effective number of relativistic degrees of freedom is $\Delta N_{\rm eff} = N_{\rm eff} - N_{\rm eff}^{\Lambda\rm CDM} < 0$, with the neutrino temperature $T_\nu$ being lower than in the standard case, $T_\nu < T_\nu^{\Lambda\rm CDM}$ due to the heating of the EM plasma by the energy injection from the $X$ decays.

(2) *LLPs decaying solely into neutrinos*: Decays into neutrinos with thermal energies $E_\nu \simeq 3T$ have the opposite effect compared to the pure EM decays: heating the neutrino plasma and increasing $N_{\rm eff}$. Decays into high-energy neutrinos ($E_\nu \gg 3T$) in MeV plasma have a qualitatively different impact. As detailed in [18,26,27], they reduce $N_{\rm eff}$, which arises from two main effects:

(a) Spectral distortions: high-energy neutrinos interact with thermal neutrinos, enhancing the high-energy tail and depleting the low-energy part of the spectrum. It is important since the rates of the neutrino-EM interaction grow with the energy of the particles.

(b) Instant thermalization of the EM plasma: any energy injection to the EM sector instantly thermalizes. Without distortions in the spectrum of $e^\pm$ particles, the net energy flow is shifted to the EM sector even when the ratio of the energy densities $\rho_\nu/\rho_{\rm EM}$ reaches equilibrium value. As a result, this shift leads to $\Delta N_{\rm eff} < 0$, analogous to pure EM plasma heating.

As the LLP lifetime $\tau_X$ increases, high-energy neutrinos interact less with the EM plasma, reducing the energy transfer and mitigating the negative impact on $N_{\rm eff}$. For sufficiently large $\tau_X$, $\Delta N_{\rm eff}$ becomes positive.

The scenario of LLPs decaying into metastable particles is even more nuanced. Given that $m_Y \gg 3T$, neutrinos from $Y$ decays are typically energetic and resemble the second scenario. However, at MeV temperatures, the decay probability $P^Y_{\rm decay}$ is a lot smaller than unity, meaning that $Y$ particles are more likely to annihilate or interact with nucleons before decaying, effectively suppressing neutrino injection. This behavior mimics pure EM heating. At lower temperatures, as $P^Y_{\rm decay} \to 1$, the situation transitions toward the mix between scenarios 1 and 2.

To qualitatively characterize the impact of the varying $P^Y_{\rm decay}$ on $N_{\rm eff}$, we define the ratio

$$r_\nu = \left.\frac{\rho_{{\rm inj},\nu}}{\rho_{\rm inj}}\right|_{t=\infty}, \tag{48}$$

which represents the fraction of the LLP's total injected energy $\rho_{\rm inj}$ allocated to neutrinos. $\rho_{\rm inj}$ and the energy density injected into neutrinos $\rho_{{\rm inj},\nu}$ evolve according to

$$\frac{d\rho_I}{dt} + 4H\rho_I = \left(\frac{d\rho_I}{dt}\right)_{\rm source}, \tag{49}$$

where

$$\left(\frac{d\rho_I}{dt}\right)_{\rm source} = \frac{m_X n_X}{\tau_X} \times \begin{cases} 1, & \rho_I = \rho_{\rm inj} \\ \xi_{X\to\nu} + \sum_{y=Y,\bar{Y}} \frac{n_y}{n_X} P^y_{\rm decay} \xi_{y\to\nu}, & \rho_I = \rho_{{\rm inj},\nu} \end{cases} \tag{50}$$

Here, $\xi_{A\to\nu}$ denotes the fraction of the $A$'s energy injected into the neutrino sector per decay:

$$\xi_{A\to\nu} = \frac{1}{m_X}\sum_j \mathrm{Br}_{A,j}\langle E_\nu^{(j)}\rangle, \tag{51}$$

with $\mathrm{Br}_{A,j}$ denoting the branching ratio of the $j$th decay mode of the particle $A$, and $\langle E_\nu^{(j)}\rangle$ mean energy of neutrinos injected in this decay. When calculating it, we assume that all metastable particles do not decay. As an example, for the decay channel $K^+ \to \mu^+ + \nu_\mu$, only the neutrino energy is accounted for, whereas the muon is dropped.

As is seen from the definition of $r_\nu$, we drop any interactions between neutrinos and electromagnetic particles. This is done for qualitative studies—to concentrate on the impact of the dynamics of metastables.

The minimum value of $r_\nu$ occurs when $P^Y_{\mathrm{decay}} = 0$, implying that only direct $X$'s decays into neutrinos would contribute. Conversely, the maximum value is achieved when $P^Y_{\mathrm{decay}} = 1$, meaning that all mesons and muons decay:

$$r_{\nu,0} = \frac{1}{m_X}\sum_j \mathrm{Br}_j \cdot \overline{\langle E_\nu^{(j)}\rangle}, \tag{52}$$

where, unlike Eq. (51), we include the contribution from inevitable decays when calculating the mean neutrino energy, $\overline{\langle E_\nu^{(j)}\rangle}$. As a cross-check, the expression (50) (and hence $r_\nu$) should give exactly the same results as Eq. (52) in the case $P^y_{\mathrm{decay}} = 1$. We confirm this in Figs. 4, 6, 8 in the limit of large $X$ lifetimes.

When neutrinos from the decay of $Y$ particles effectively decouple, the sign of $\Delta N_{\mathrm{eff}} = N_{\mathrm{eff}} - N_{\mathrm{eff}}^{\Lambda\mathrm{CDM}}$ is determined by whether $r_{\nu,0}$ exceeds the ratio of neutrino to total energy densities in standard cosmology, which for temperatures $T \gtrsim m_e$ is

$$q_\nu = \frac{\rho_\nu}{\rho_\nu + \rho_{\mathrm{EM}}} = \frac{21}{43}. \tag{53}$$

If $r_{\nu,0} > q_\nu$, then $\Delta N_{\mathrm{eff}}$ increases as $\tau_X$ becomes large ($\tau_X \gtrsim 1$ s). Consequently, $\Delta N_{\mathrm{eff}}$ transitions from negative to positive values as $r_\nu$ approaches $r_{\nu,0}$.[10]

### B. Neutrino-antineutrino energy asymmetry

Generically, the evolution (23), (31) is not $Y - \bar{Y}$ symmetric due to the term describing the interactions with nucleons. The reason is that there are no antinucleons, which means that the generic interaction rate of $Y$ and $\bar{Y}$ does not have charge conjugation symmetry. This implies

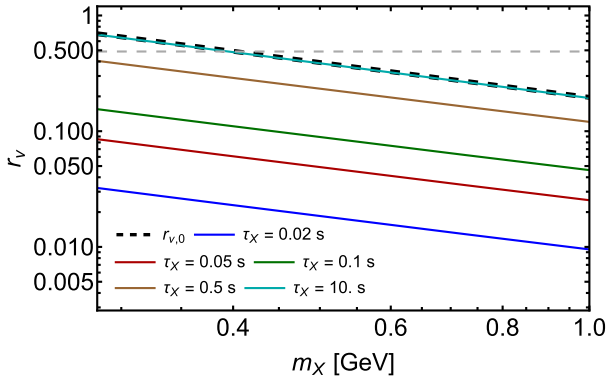


FIG. 4. The *qualitative* impact of the dynamics of metastable particles on neutrinos for the model of a hypothetical particle $X$ decaying solely into a pair of the charged pions (see text for details). The plot shows the evolution of the fraction $r_\nu$ of the $X$'s energy directly injected into neutrinos, Eq. (48), as a function of the $X$'s mass $m_X$ and lifetime $\tau_X$ (shown by different solid lines). The gray dashed line shows the value $r_\nu = 21/43$, for which $\Delta N_{\mathrm{eff}} = 0$ in the absence of the neutrino-EM interactions [see a discussion around Eq. (53)]. For the fixed $X$ mass, $r_\nu(\tau_X) \to 0$ at lifetimes $\tau_X \ll 1$ s. Once lifetime increases, it gradually grows and, for $\tau_X \gg 1$ s, approaches the value $r_{\nu,0}$ (the dashed black line), which is when all pions and muons inevitably decay [Eq. (59)]. The pattern occurs since $\mu$s' and $\pi$s' scattering and annihilation processes, preventing them from releasing energy into neutrinos, become less efficient at lower temperatures (remind Sec. IV B). The slope of the lines represents the increasing kinetic energy of the pion as a function of the LLP mass; it gets immediately transferred to the EM sector independently of the lifetime.

that, in general, $n_Y \neq n_{\bar{Y}}$, i.e., metastable particles and antiparticles evolve differently. This asymmetry translates to an asymmetry between neutrinos and antineutrinos via their decays. Let us discuss its qualitative aspects.

The asymmetry may be in number ($n_\nu \neq n_{\bar{\nu}}$) and energy distributions (meaning in particular that $\rho_\nu \neq \rho_{\bar{\nu}}$). In the first case, the net lepton charge $L_\nu \propto n_\nu - n_{\bar{\nu}}$ is generated in the neutrino sector, and the opposite charge $L_l = -L_\nu$ in the electron-positron sector. No sizeable $L_\nu$ is induced because we assume that the initial $X$ particle is electrically neutral and the baryon number is conserved. Indeed, the electric charge conservation means that independently of the microscopics of the $Y, \bar{Y}$ evolution, $L_l$ may occur only because of changing the yield of protons. The baryon number conservation implies that this change is bounded by $\eta_B \sim 10^{-9}$. Therefore, we may just assume that $n_\nu = n_{\bar{\nu}}$.

However, the magnitude of the energy asymmetry is not bounded by this argument. First, even if conserving the number of neutrinos, decays of different $Y$s inject neutrinos with different energies. Namely, decays of kaons release neutrinos with energies as large as $E_{\nu,\max} \approx m_K/2$, decays of pions result in the neutrinos with energy $E_{\nu,\max} \approx 29$ MeV, whereas the maximal neutrino energy from muons decays is $E_{\nu,\max} \approx m_\mu/2$. Second, some $Y$s, such as kaons, may

[10]Note that $q_\nu$ decreases after electron-positron annihilation, allowing for an additional sign change if the LLP decays at $T \lesssim m_e$.

interact with both protons and neutrons, as well as may or may not convert them, meaning that $\rho_\nu - \rho_{\bar{\nu}}$ may easily exceed the bound $n_B \times E_\nu$ coming from the number asymmetry.

Let us now discuss the energy asymmetry in more detail. If only muons are injected, the nucleon interaction term may be neglected (see a discussion in Sec. II A). If, in addition, the $X$ particle decays into the charged pions, it is important, and we need to analyze it further. Both $\pi^+, \pi^-$ interact with nucleons; in addition,

$$\langle \sigma^{\pi^\pm}_{\mathcal{N}} v \rangle = \langle \sigma^{\pi^\pm}_{\mathcal{N}\to\mathcal{N}'} v \rangle, \tag{54}$$

i.e., pions interact with nucleons solely via converting them [recall Eq. (11)]. Using this and utilizing the expression for the nucleon abundance $X_n$ from Eq. (31), we find that the nucleon interaction terms for $\pi^+, \pi^-$ are actually identical:

$$\begin{aligned} n_{\pi^-} \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma^{\pi}_{\mathcal{N}} v \rangle &= n_{\pi^+} \sum_{\mathcal{N}} n_{\mathcal{N}} \langle \sigma^{Y}_{\mathcal{N}} v \rangle \\ &= \frac{n_{\pi^-} n_{\pi^+} \langle \sigma^{\pi^-}_{p\to n} v \rangle \langle \sigma^{\pi^+}_{n\to p} v \rangle}{n_{\pi^-} \langle \sigma^{\pi^-}_{p\to n} v \rangle + n_{\pi^+} \langle \sigma^{\pi^+}_{n\to p} v \rangle} \end{aligned} \tag{55}$$

The situation is different when charged kaons are injected as well. There is an explicit asymmetry due to the interaction with nucleons: $K^+$ does not interact with nucleons in the MeV plasma, while $K^-$ participates in various processes with them: interacting with both $n$ and $p$, converting $n \leftrightarrow p$ as well as keeping the nucleon type the same (remind Sec. II C). As a result, more $K^-$ would disappear before decaying than $K^+$. Decays of $K^+$ would directly produce muon neutrinos and not antineutrinos. On the other hand, it means that we have more $\pi^+, \mu^+$ particles, that decay into antineutrinos.

Overall, this decay asymmetry may induce sizeable differences in the energy distributions of neutrinos and antineutrinos. The energy asymmetry may be split into the ranges $E_\nu > m_\mu/2$, to which only the $K$ decays contribute, and $E_\nu < m_\mu/2$, where the main sources are decays of muons and pions. The first domain is overabundant for neutrinos, whereas the second is for antineutrinos. We leave the quantitative study of this intriguing asymmetry development question for future work.

### C. Evolution of the $n/p$ ratio

As was mentioned in Sec. IV, injecting mesons into the primordial plasma significantly modifies the dynamics of the $n/p$ ratio $n_n/n_p$. Overall, the effect of the meson-driven $p \leftrightarrow n$ conversion is well-known [4,8,36,37], but let us describe it shortly. In ΛCDM, the $n/p$ ratio is suppressed by the Boltzmann exponent as far as weak interactions maintain chemical equilibrium between the neutrons and protons:

$$\frac{n_n}{n_p} \approx \exp\left[-\frac{m_n - m_p}{T}\right] \tag{56}$$

Once mesons are injected, they increase the ratio above the value (56). This is mainly because meson-driven $p \leftrightarrow n$ conversion is thresholdless. The BBN constraint on LLP lifetimes may be imposed from the requirement on this enlarged ratio to relax to the ΛCDM value within the margin determined by the error in the primordial helium measurements [8]. The meson-driven $p \leftrightarrow n$ conversion cross section is orders of magnitude higher than the one for the weak conversion, and even exponentially suppressed amounts of mesons (at times $t \gg \tau_X$) completely drive the dynamics of the $n/p$ ratio. Because of this, the resulting constraint on the LLP's lifetime depends on the LLP's initial number density and the yield of mesons available for the conversion only logarithmically [[8], Eq. (11)].

Because of the same reason, the meson-driven $p \leftrightarrow n$ conversion typically dominates over other effects of LLPs on the dynamics of the $n/p$ ratio, including the modified expansion of the Universe and neutrino properties. For example, consider heavy neutral leptons with lifetimes $\tau_N \simeq 0.02$ s and heavy enough to decay into mesons. While modifying $N_{\rm eff}$ at a percent level, they induce a huge change in the $n/p$ ratio due to mesons [8,17].

The only modification of this picture due to our study comes from adding the meson annihilation processes. They suppress the yield of mesons available for the $p \leftrightarrow n$ conversion, Eq. (34). However, the suppression is maximum a factor of few (remind Fig. 2), which would modify the BBN constraint in a minor way as it enters the logarithm.

## VII. CASE STUDIES

In this section, we consider the impact of the evolution of $Y$s on neutrinos for three models with LLPs $X$: A toy model adding a particle with constant abundance decaying into charged pions (Sec. VII A), Higgs-like scalars (Sec. VII B), and heavy neutral leptons (HNLs) (Sec. VII C). We will discuss the mass and lifetime dependence utilizing two approaches: the qualitative one, by calculating the fraction of $X$'s energy directly injected into neutrinos, Eq. (48), and the quantitative one, for which we will accurately trace the evolution of neutrinos using the unintegrated Boltzmann approach. The former serves the purpose of demonstrating the impact of varying decay probability on neutrino injections (Figs. 4, 6, 8), while the latter is used to obtain the main results of this paper—Figs. 5 and 7.

### A. Toy model: LLPs decaying into pions

Consider a toy model with the LLP $X$ decaying solely into charged pions. It means that in Eq. (23)

$$N^X_{\pi^\pm} = 1, \qquad N^X_{\mu,K} = 0 \tag{57}$$

As a result, no kaons are involved, but there still would be muons originating from the pion decay, recall Eq. (8). To make the analysis as transparent as possible, the $X$'s

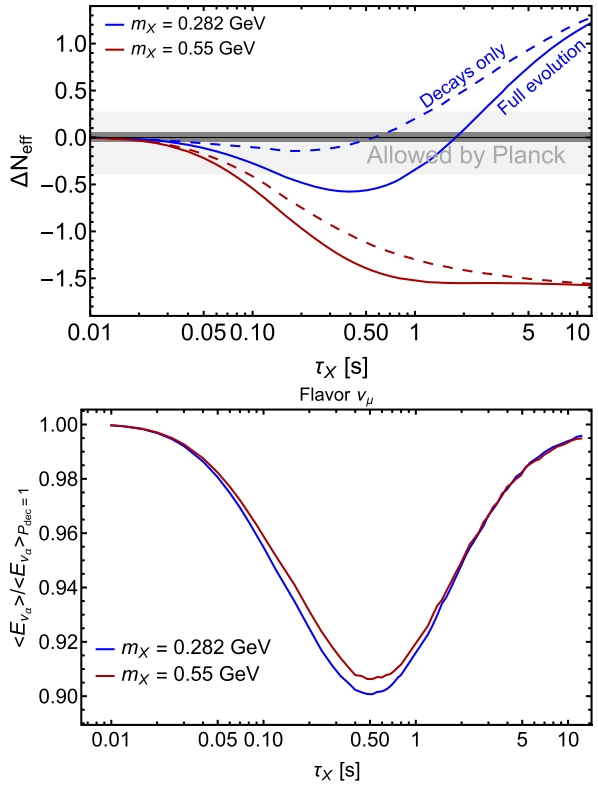


FIG. 5. The effect of the evolution of $Y$s on the properties of neutrinos for the same LLP model as in Fig. 4. The results are obtained using the unintegrated neutrino Boltzmann equation solver incorporating the dynamics of metastable particles, which we describe in Sec. V. Two LLP masses are considered—$m_X = 282$ and 550 MeV. They represent the cases when the correction $\Delta N_{\rm eff} = N_{\rm eff} - N_{\rm eff}^{\Lambda\rm CDM}$ tends, correspondingly, to a positive and negative value in the limit of large lifetimes (see a discussion in Sec. VII A). To highlight the importance of the annihilation and interactions with nucleons, we consider two setups for each of the masses: the one that includes annihilation and interactions with nucleons (the realistic setup) and the one that includes solely decays and kinetic energy loss, which corresponds to the assumption $P^Y_{\rm decay} = 1$ used in all previous studies. Top panel: the correction $\Delta N_{\rm eff}$. The gray band represents the Planck 95% CL constraints $N_{\rm eff} = 2.99^{+0.33}_{-0.34}$ [52], whereas the black band shows the forecast of the accuracy of the measurements performed by the Simons Observatory, which we assume to be centered at $\Delta N_{\rm eff} = 0$ [68]. Bottom panel: the ratio of the mean energies of the muon neutrinos in the realistic setup case to the setup $P_{\rm decay} = 1$, as a function of the $X$'s lifetime. The numerical noise in the domain of large lifetimes is caused by the precision limit of the Boltzmann solver.

abundance is chosen to be a constant

$$\mathcal{Y}_{\rm LLP} \equiv \left(\frac{n_{\rm LLP}}{s}\right)_{T=10\ \rm MeV} = 2\times 10^{-3} \tag{58}$$

It corresponds to the scenario where the particle $X$ was in thermal equilibrium and decoupled while still being relativistic in a broad range of masses and lifetimes. We consider the masses above the dipion decay threshold $m_X > 2m_\pi \approx 0.28$ GeV. Regarding the lifetimes, following the discussion in Sec. III, we test the range $0.01\text{s} < \tau_X < 10$ s.

Let us first discuss how $X$ would distribute its energy among the neutrino and the EM sectors. Upon decay, it will produce a pair of pions whose nontrivial evolution has been discussed before. Injection into the neutrino sector would occur only in the case of decay of $\pi^\pm$ producing a muon and a muon (anti)neutrino, Eq. (8); the resulting muon may then decay into neutrinos as well [Eq. (2)].

If the pion and the subsequently produced muon would inevitably decay, roughly $r^\pi_{\nu,0} \approx 70\%$ of the pion mass would go to the neutrino sector. In this case, the injection into the EM sector is composed of the initial kinetic energy of the pion $(m_X - 2\cdot m_\pi)/2$ and the EM part of the muon decay, which is approximately 30% of its mass. This gives us the maximal possible fraction of the energy of the $X$ particle directly injected in the neutrino sector:

$$r_{\nu,0} \approx \frac{2m_\pi \cdot r^\pi_{\nu,0}}{m_X} \approx q_\nu \cdot \frac{388\ \rm MeV}{m_X}, \tag{59}$$

where $q_\nu$ is the $\Lambda$CDM ratio (53). Provided that there are no interactions of neutrinos with the EM plasma, if $r_{\nu,0}$ exceeds this ratio (i.e., $m_X < 388$ MeV), the correction $\Delta N_{\rm eff}$ would be positive in the limit of large lifetimes (remind the discussion in the previous section).

The disappearance of the pions and muons because of annihilation and interaction with nucleons spoils this picture. Let us first estimate the impact of these effects qualitatively. Namely, we utilize the discussion of Sec. VI and calculate the quantity $r_\nu(m_X, \tau_X)$, see Fig. 4. At small lifetimes $\tau_X \lesssim 0.5$ s, $r_\nu \ll r_{\nu,0}$; this is because pions and muons produced in $X$ decays would prefer to disappear before decaying at $T \gtrsim 1$ MeV, remind Sec. IV B. Therefore, decays into pions and muons affect the neutrino bath very similar to the solely electromagnetic decays. With the increase of the lifetime, more and more $Y$s would decay, and $r_\nu$ tends to the maximal possible value $r_{\nu,0}$.

Now, let us calculate $N_{\rm eff}$ as a function of the LLP's mass and lifetime for this toy model. For this purpose, we switch to the full unintegrated Boltzmann approach described in Sec. V. The plot with $\Delta N_{\rm eff}$ as a function of $X$ mass and lifetime is shown in the upper panel of Fig. 5. Two representative choices for the mass of $X$ are considered: $m_X = 282$ MeV, for which $r_{\nu,0} > q_\nu$ (and hence $\Delta N_{\rm eff}$ would be positive at large lifetimes), and $m_X = 550$ MeV, for which $r_{\nu,0} < q_\nu$. To highlight the impact of the $Y$ evolution on the properties of neutrinos, we show the results for two setups—the one assuming $P^Y_{\rm decay} = 1$ (i.e., when the decays are inevitable), and the one including the full evolution of $Y$s, i.e. accounting for annihilations and

interactions with nucleons (which we will call below the realistic setup).

The behavior of the curves in Fig. 5 agrees with the qualitative discussion in Sec. VI and in this section. For the lifetimes $\tau_X \ll 10$ s and both masses, there are severe differences in $\Delta N_{\rm eff}$ between the two setups. The realistic setup corresponds to a lower $\Delta N_{\rm eff}$; this is expected since the decay of $Y$ particles injects more energy directly into the EM sector. In the limit of large lifetimes $\tau_X \to 10$ s, the two results match, as annihilation and interactions with nucleons become irrelevant.

Note that our approach predicts a *decrease* of $N_{\rm eff}$ in the presence of the LLPs with the lifetimes $\tau_X \lesssim 1$ s, decaying into neutrinos of high energies $E_\nu \gg T$ at MeV temperatures. These results are aligned with the previous studies utilizing different unintegrated approaches [18,20,44] (see also a discussion in Refs. [26,27]).

To investigate the impact of the $Y$ disappearance further, let us consider the ratio of the mean neutrino energies after the electron-positron annihilation for these two setups; see the lower panel of Fig. 5. The setup with $P_{\rm decay} = 1$ leads to higher neutrino energies, which is expected, as we have a more abundant high-energy neutrino tail.

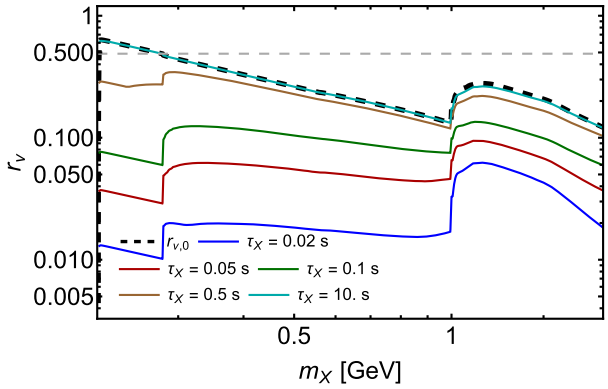


FIG. 6. The fraction of the energy directly injected into neutrinos in decays of Higgs-like scalar, see Eq. (48) (remind also Fig. 4). The solid lines correspond to different scalar lifetimes, whereas the dashed black line is obtained under the usual assumption that all the metastable particles decay. The rapid change at the masses $m_X = 2m_\mu, 2m_\pi, 2m_K$ is caused by the opening of the decay into the pair of corresponding particles. The horizontal gray dashed line denotes the value of $r_{\nu,0}$ for which the injections would increase the neutrino-to-EM energy density ratio if assuming no interactions in the primordial plasma [see a discussion around Eq. (53)].

### B. Higgs-like scalars

Let us now consider a particular model of long-lived particles. We start with Higgs-like scalars $S$ [28]. We will concentrate on the minimal model with the effective Lagrangian

$$\mathcal{L} = \theta m_h^2 h S + \mathcal{L}_{\rm kinetic} \tag{60}$$

Here, $h$ is the Higgs boson, and $\theta$ is the mixing angle, with $|\theta| \ll 1$. Due to the mass mixing, the scalars have a similar interaction pattern as $h$ (so Yukawa couplings to the SM fermions), with the couplings additionally suppressed by $\theta$.

The main decay modes of these scalars in the GeV mass range are two-body decays into particle-antiparticle pairs

$$S \to e^+e^-/\mu^+\mu^-/\pi^+\pi^-/2\pi^0/K^+K^-/K_L K_S, \tag{61}$$

with the decays into heavier particles dominating once they become kinematically possible. The fraction of energy injected into neutrinos $r_{\nu,0}$ by the scalar decays is shown in Fig. 6. It is exactly zero for masses $m_S < 2m_\mu$, because the only available scalar decay modes are into the EM particles. Then, it gets rapidly enhanced at $m_S = 2m_\mu$ and $m_S = 2m_K$—the mass thresholds where decays into two muons and kaons open up. In the domain of intermediate masses, it gradually decreases as the decay products have more and more kinetic energy that gets stored in the EM plasma.

The cosmological production and constraints of $S$ have been studied in [14,15]. Among the cosmological effects of the scalars, it studied the impact of the Higgs-like scalars on neutrinos. The analysis was simplified by considering a version of the integrated neutrino Boltzmann equation and assuming that $Y = \mu$, $\pi$, $K$ decay after thermalizing their kinetic energy. Under this approximation, $\Delta N_{\rm eff}$ is determined by whether $r_{\nu,0}$ exceeds the quantity $q_\nu$ during decays of the scalar. This is the case in the region $2m_\mu < m_S \lesssim 2m_\pi$.

Let us now include the effects of mesons and muons evolution as well as momentum-dependent (unintegrated) Boltzmann equations for the neutrino distributions. In Fig. 6, we show the mass-lifetime dependence of $r_\nu$ including the impact of annihilation and interactions. Similarly to the case of the toy model decaying into pions, the generic pattern is that $r_\nu(\tau_S) \to 0$ for small scalar lifetimes and reaches $r_{\nu,0}$ for the lifetimes $\simeq 10$ s. In particular, the ratio $r_\nu$ becomes less than 5% (and so most of the scalar's energy goes to the EM sector) for the lifetimes $\tau_S \lesssim 0.05$ s. $r_\nu$ jumps at $m_S = 2m_\pi$, which is caused by the opening of the dipion decay channel. The pions (the main decay products in the mass range $2m_\pi < m_S < 2m_K$) have a larger decay probability than the muons, which means that they have a higher chance to release energy into neutrinos than muons. The behavior of $\Delta N_{\rm eff}$ is shown in Fig. 7. In the mass range $2m_\mu < m_S \lesssim 2m_\pi$, increasing the scalar lifetime, we see the transition between negative and positive changes in $N_{\rm eff}$. It is caused by tending $r_\nu \to r_{\nu,0} > q_\nu$. At higher masses, $r_{\nu,0} < q_\nu$, so in any case, $\Delta N_{\rm eff}$ remains negative.

### C. Heavy neutral leptons

Let us now consider heavy neutral leptons. The Lagrangian of HNLs has the form

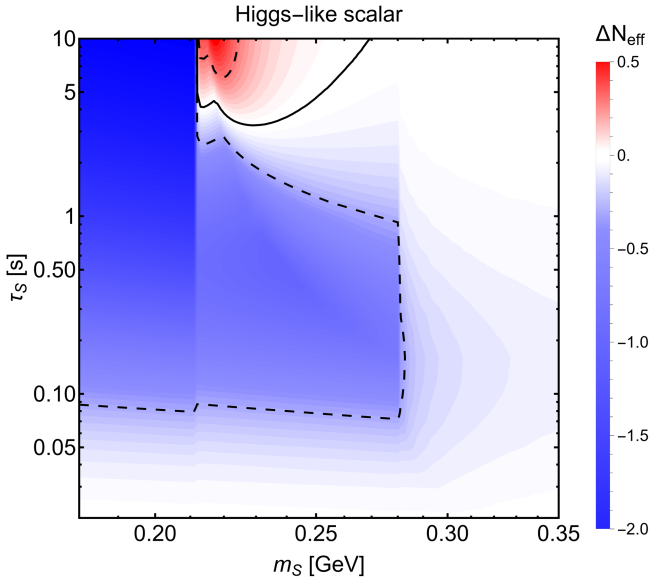


FIG. 7. The effect of the presence of the Higgs-like scalars on the correction $\Delta N_{\rm eff} = N_{\rm eff} - N_{\rm eff}^{\Lambda\rm CDM}$, obtained by using the unintegrated Boltzmann equations from Sec. V. The solid black line shows the parameter space where $\Delta N_{\rm eff} = 0$, whereas the dashed black lines denote the domain where $\Delta N_{\rm eff}$ are beyond the lower and upper bounds of the $N_{\rm eff}$ measurements as extracted from Planck measurements [69]. The change in the sign of $\Delta N_{\rm eff}$ is driven by the dynamics of metastable particles produced by $S$'s decays. The decrease of the magnitude of $|\Delta N_{\rm eff}|$ with the scalar mass is caused by the scaling of the scalar abundance $\mathcal{Y}_S(m_S) \propto \Gamma_S^{-1}(m_S, \theta = 1)$, where $\Gamma_S$ is the scalar decay width (see [14] for details).

$$\mathcal{L} = y_\alpha \bar{L}_\alpha \tilde{H} \mathrm{HNL} + \mathrm{H.c.}, \tag{62}$$

where $\alpha$ denotes the SM lepton generation, $L_\alpha$ is the corresponding left doublet, $y_\alpha$ is the Yukawa interaction coupling, while $\tilde{H} = i\sigma_2 H^*$ is the Higgs doublet in the conjugated representation. Effectively, HNLs interact as heavy neutrinos, with the interaction coupling being suppressed by the mixing angle $U_\alpha \simeq y_\alpha v_H / m_{\rm HNL}$, where $v_H$ is the Higgs VEV [35]. We will consider the case of HNLs mixing with the muon neutrinos $\nu_\mu$, keeping in mind that the other cases are similar.

Let us briefly discuss the production of HNLs. In high-temperature plasma, the mixing angle gets modified because of the thermal neutrino self-energy correction. In particular, in the plasma without the lepton asymmetry at temperatures $T \gtrsim 1$ GeV, the effective mixing angle is given by

$$U_{\rm m}^2(T) \approx \frac{U^2}{\left[1 + 9.6 \times 10^{-24} \left(\frac{T}{1\ \rm MeV}\right)^6 \left(\frac{m_{\rm HNL}}{150\ \rm MeV}\right)^{-2}\right]^2}, \tag{63}$$

where $m_{\rm HNL}$ is the HNL mass. The scaling of the HNL production rate with temperature is $\Gamma_{\rm int} \sim G_F^2 T^5 U_{\rm m}^2$, with $G_F$ being the Fermi coupling. Comparing the HNL interaction rate with the Hubble rate $H$, we may establish whether HNLs entered the thermal equilibrium. Namely, the ratio $\Gamma_{\rm int}/H$ is $\ll 1$ at high temperatures $T$ because of the suppression of $U_m(T)$, then reaches the peak value at $T_{\rm peak} \approx 12\ \text{GeV}\ (m_{\rm HNL}/(1\ \text{GeV}))^{1/3}$ GeV, and then starts decreasing, since $\Gamma_{\rm int}$ drops with $T$ faster than $H$. If the rate-to-Hubble ratio at $T_{\rm peak}$ is $< 1$, HNLs never entered thermal equilibrium.

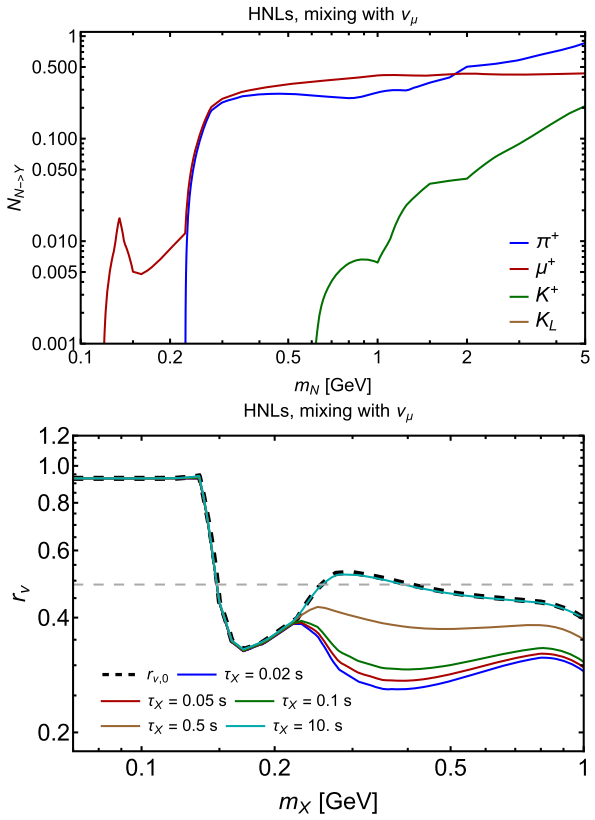


FIG. 8. The qualitative impact of the metastable particle dynamics on neutrinos for the case of HNLs coupled to the muon neutrinos. Top panel: the fractions of the metastable particles $Y = \mu^\pm, \pi^\pm, K^\pm$ per HNL decay. Bottom panel: the dependence of $r_\nu$ on the HNL lifetime. The minimal value of $r_\nu$ corresponds to the situation when all mesons and muons disappear without decaying; then, $r_\nu$ is saturated solely by direct decays into neutrinos.

We have calculated the HNL abundance following the approach similar to the one used in Ref. [8]. To compute the kinematics of HNL decay products, we used the SensCalc package [70], which we have modified to account for transferring of all the kinetic energy of the charged metastable particles to the EM plasma and forbid the mesons and muons to decay. We used the exclusive decays below the HNL mass $m_{\rm HNL} \simeq 1$ GeV and decays into jets above this mass, with showering and hadronization performed using Pythia8 [71]. The amounts of the charged pions, muons, and kaons per HNL mass are shown in Fig. 8 (top panel).

Using the machinery described above, we computed the quantity $r_\nu$; see the same figure (bottom panel). Unlike the models considered above, the decay palette of the HNLs includes processes directly injecting neutrinos. They are

$$\text{HNL} \to \nu_\alpha \bar{\nu}_\alpha \nu_\beta, \qquad \text{HNL} \to \text{hadrons} + \nu_\beta, \quad (64)$$

where "hadrons" denote either a single meson such as $\pi^0$ or a multimeson state, depending on the HNL mass [35]. Therefore, even if all the mesons and muons disappear without decaying, the quantity $r_\nu$ would be nonzero even for short HNL lifetimes $\tau_{\text{HNL}} \lesssim 0.1$ s. However, the fraction to be injected by meson decays is still large, depending on the HNL mass.

Because of the presence of the direct decays into neutrinos, computing the impact of HNLs on the primordial neutrinos is much more complicated: the traditional approaches of solving the neutrino Boltzmann equation based on the discretization of the comoving momentum space would take a too large amount of time to evolve the neutrino distribution function. We will return to this in future work. Nevertheless, Fig. 8 shows the importance of careful tracing of the evolution of the metastable particles in the HNL case.

## VIII. CONCLUSIONS

Many new physics scenarios introduce long-lived heavy particles $X$, that decay into *metastable* Standard Model (SM) particles $Y$, such as muons, charged pions, and kaons. Examples of the $X$ particles include Higgs-like scalars, dark photons, axionlike particles, and others. The lifetimes of the $Y$ particles are sufficiently long to allow numerous interactions with components of the primordial plasma, including electromagnetic particles and nucleons. These interactions significantly modify the evolution of $Y$ abundances and, consequently, affect the properties of primordial neutrinos.

In this work, we conducted a detailed study of $Y$ particle evolution, incorporating processes such as annihilation with antiparticles, interactions with nucleons, elastic electromagnetic scattering, and decays (see Sec. II). Notably, annihilation processes are examined here for the first time, while interactions with nucleons have previously been considered only regarding big bang nucleosynthesis (BBN) and not when studying the impact on neutrinos.

We have outlined a two-step scheme to trace the dynamics of $Y$ particles and their impact on neutrinos in Sec. III. The first step is to analyze the coupled dynamics of $Y$ particles and nucleon densities. We have developed a systematic approach based on the system of the integrated Boltzmann equations on their number densities (Sec. IV) and incorporated the method in a public *Mathematica* code. We have demonstrated that at MeV temperatures, $Y$ particles predominantly annihilate or interact with nucleons rather than decay (Sec. IV B).

As step two, we have incorporated the $Y$ dynamics in an unintegrated Boltzmann solver for the neutrino momentum distribution functions, see Sec. V. It may be used to study a broad range of LLPs, including those decaying into metastable particles, as well as neutrinos or purely electromagnetic particles. Being public and simple in use, it allows the scientific community to robustly trace the evolution of neutrinos in the presence of new physics, which is especially important in light of future CMB observations.

The nontrivial dynamics of $Y$ particles substantially alters the influence of new physics on neutrino properties, which we discuss in qualitative terms in Sec. VI. Specifically, if $Y$ particles decay, a significant fraction of their mass energy is transferred to the neutrino sector, inducing spectral distortions. Conversely, if they disappear without decaying, their energy is instead fully injected into the electromagnetic sector. Additionally, the differential decay rates of kaons and antikaons lead to asymmetries in the energy distributions of neutrinos and antineutrinos, which may persist if injections occur during neutrino decoupling (Sec. VI B). A comprehensive analysis of this intriguing question is left for future work.

We applied the combined methodology of Secs. V and VI to specific models with LLPs decaying into metastable particles: a toy model with pions (Sec. VII A), Higgs-like scalars (Sec. VII B), and heavy neutral leptons (Sec. VII C). Our findings reveal significant deviations from previous studies assuming the inevitability of $Y$ decays, including changes in both the magnitude and sign of $\Delta N_{\text{eff}}$ and alterations in the neutrino distribution functions, as illustrated in Figs. 5 and 7.

In summary, our results provide a deeper understanding of how long-lived particles influence the neutrino population in the early Universe.

## ACKNOWLEDGMENTS

This work has received support by the European Union's Framework Programme for Research and Innovation Horizon 2020 under Grant No. H2020-MSCA-ITN-2019/860881-HIDDeN, JSPS Grant-in-Aid for Scientific Research KAKENHI Grant No. 24KJ0060. K. A. is grateful to Maksym Ovchynnikov and Thomas Schwetz for the hospitality during the stay at the Institute for Astroparticle Physics, KIT. The authors thank Miguel Escudero for carefully reading the manuscript and for providing useful comments.

## DATA AVAILABILITY

The data that support the findings of this article are openly available [40,41,43], embargo periods may apply.

## APPENDIX A: *Mathematica* CODE FOR THE EVOLUTION OF METASTABLE PARTICLES

In this appendix, we discuss the *Mathematica* code that traces the evolution of the metastable particles in the presence of the decaying LLPs $X$; it is available on [40] and Zenodo [41]. The `Zenodo` repository also contains precomputed data for some LLP models.

The central notebook is `main.nb`. Once launching its in`itialization cells, it first calls secondary notebooks with all the necessary definitions. The secondary notebooks are located in the folder *codes/Secondary particles evolution*. They are: `parameters-functions.nb`, defining various parameters, analytic functions, and the list of the metastables `Yparticleslist`; `cross-sections.nb`, containing the calculations of various interaction rates involving $Y$s; `universe-simplified-dynamics.nb`, containing a simplified description of the thermodynamics following the approach of [25]; `evolution-Ys.nb`, which brings all the processes with $Y$s altogether and defines the system of the Boltzmann equations on the $Y$ s' number densities, depending of various options and properties of the $X$ particles; and `final-system.nb`, which uses these codes to calculate the impact of the decaying LLPs with $Y$ decay products. Apart from that, the folder *SM_Rates* contains useful definitions such as effective Lagrangians, tabulated energy densities of electrons, and oscillation probabilities.

Once all secondary codes are called, users may use the main notebook to study various physics cases. As an input for the model, the code requires various properties. The input for the implemented models is stored in the section *LLP input*. Each of its subsections is dedicated to a separate model.

For the given model `LLP` and `mass`, $\tau$ staying for the LLP's mass and lifetime, the main definitions are:

(i) $\tau$`LLP [LLP,mass,coupling]`, which describes the dependence of the lifetime on LLP's mass and coupling;

(ii) `nLLPini[LLP, mass, τ]`, which is the number density of the LLP in the units of $\mathrm{GeV}^3$ at $T = 20$ MeV, $n_{\mathrm{X,ini}}$; the code assumes that the LLPs are already decoupled at that epoch.

(iii) {$\xi$ `to`$\nu$ `[LLP,` $\nu_e$`, mass],` $\xi$`to`$\nu$ `[LLP,` $\nu_\mu$`, mass],` $\xi$`to`$\nu$ `[LLP,` $\nu_\tau$`, mass]`}, which are the mass-dependent fractions of the LLP's mass injected *directly* in the neutrino sector, the flavor $\nu_\alpha$.

(iv) `NtoY[LLP,Y, mass]`—the amount of the $Y$ particles produced per LLP's decay. It is defined by Eq. (24).

(v) `BrTo2`$\nu$ `[LLP, "`$\nu_e$`," mass], BrTo2`$\nu$ `[LLP, "`$\nu_\mu$`," mass], BrTo 2`$\nu$ `[LLP, "`$\nu_\tau$`," mass]`, which is the branching ratio of decays $X \to \nu\bar{\nu}$.

(vi) `EnergyFractionsTo`$\nu$`[LLP, "Total," mass]`, which is the *total* fraction injected into neutrinos, assuming that all $Y$ particles inevitably decay.

Note that the code is more generic than the unintegrated Boltzmann code. For example, it may consider any LLP with any decay mode into neutrinos, $K_L$s, etc.

The section *Generating the evolution of Ys* is devoted to generating the data for the grid of masses and lifetimes of the given LLPs: `MassGrid[LLP]`, `lifetimeGrid[LLP]`, defined in section *Mass and lifetime grids* for each LLP. Subsection *Launching for mass and lifetime grids* launches the system of equations for the given LLP model, mass and lifetime grids. This is done with the help of the routine `exportBlockFullData[IfWritingOutputForBoltzmann]`, where the parameter `IfWritingOutputForBoltzmann` may be `True` (if one wants to prepare output for the unintegrated Boltzmann solver) or `False` (if does not). For each mass and lifetime, this routine launches

`mergedFunction[LLP, mass,` $\tau$`, DecayOnly]`

where the parameter `DecayOnly` (`True` or `False`) controls whether including annihilation and interactions with nucleons, and returns the following data:

{`LLP parameters, tabulated decay probabilities of Ys`}

where `LLP parameters` is the row

$$\{\texttt{mass}, \tau, n_{\mathrm{X,ini}}, N_{\mathrm{eff}}, r_1, r_\nu, r_3, N^X_{\mu^+}, N^X_{\pi^+}, N^X_{K^+}, N^X_{K_L}, r_{\nu,0}, \mathrm{Br}_{X\to\nu_e\bar{\nu}_e}, \mathrm{Br}_{X\to\nu_\mu\bar{\nu}_\mu}, \mathrm{Br}_{X\to\nu_\tau\bar{\nu}_\tau}\}$$

Here, the value of $N_{\mathrm{eff}}$ is obtained via the integrated approach of Ref. [25], $r_1$ is the cumulative fraction of the total energy density injected by LLP to the total energy density of the Universe, $r_3$ is the ratio of the energy density injected into the neutrino sector to the total neutrino energy density. The quantities $N^X_Y$ are defined around Eq. (23), while $r_\nu$, $r_{\nu,0}$ are given by Eqs. (48), (52). The tabulated decay probabilities are provided in the form {`time in seconds,` $P^Y_{\mathrm{decay}}$`(time)`}.

The routine `exportOutputForCluster[LLP]` prepares the data for the unintegrated Boltzmann solver for the neutrino Boltzmann equation from Sec. V.

There are also notebooks `plot-numbers.nb` and `plot-distributions.nb`. The former imports the datasets prepared by the integrated and unintegrated approaches and makes plots (subsection *Plots*). The imported data has the names `OutputLLPintegrated[LLP]` and `OutputLLPUnintegrated[LLPsel]`, correspondingly for the output of the notebook and the Boltzmann solver. The latter is precomputed for some models and may be found in the associated `Zenodo` repository. Each subsection (e.g., $r_\nu$), contains definitions needed to make a plot for each model (say, {`mminPlot [LLP, "`$r_\nu$`"], mmaxPlot [LLP, "`$r_\nu$`"]`} defines the LLP mass range for the $r_\nu$ plot), as well as the code making the plot itself.

The notebook `plots-distributions.nb` makes plots with the neutrino momentum distributions and average energies calculated using the unintegrated code.

## APPENDIX B: APPROACH OF INTEGRATED BOLTZMANN EQUATIONS

According to the integrated approach to solve the neutrino Boltzmann equations [25,51], we need to solve the following system in order to trace the neutrino distribution

$$\begin{cases} \frac{d\rho_{\nu_\alpha}}{dt}+4H\rho_{\nu_\alpha}=\left(\frac{d\rho_{\nu_\alpha}}{dt}\right)_X+\left(\frac{d\rho_{\nu_\alpha}}{dt}\right)_{\mathrm{EM}\leftrightarrow\nu_\alpha}, \\ \frac{d\rho_{\mathrm{EM}}}{dt}+3H(p_{\mathrm{EM}}+\rho_{\mathrm{EM}})=\left(\frac{d\rho_{\mathrm{EM}}}{dt}\right)_X-\sum\limits_\alpha\left(\frac{d\rho_{\nu_\alpha}}{dt}\right)_{\mathrm{EM}\leftrightarrow\nu_\alpha}, \\ H=\frac{\dot{a}}{a}=\sqrt{\frac{8\pi G}{3}(\rho_X+\rho_\gamma+\rho_e+\sum\limits_\alpha\rho_{\nu_\alpha})}, \\ \rho_X\approx\left(\frac{a(t_0)}{a(t)}\right)^3 n_{X,0}m_X\exp\left[-\frac{t-t_0}{\tau_X}\right], \end{cases} \tag{B1}$$

Here, $\rho_{\mathrm{EM}}=\rho_\gamma+\rho_e$ is the energy density of the EM particles, with the electron-positron and photon components

$$\rho_\gamma=\frac{1}{\pi^2}\int dE\frac{E^3}{\exp\left(\frac{E}{T}\right)-1},$$
$$\rho_e=\frac{2}{\pi^2}\int dE\frac{E^2\sqrt{E^2-m_e^2}}{\exp\left(\frac{E}{T}\right)+1} \tag{B2}$$

$T$ is the temperature of the EM plasma. The pressure of the EM plasma is $p_{\mathrm{EM}}=p_e+p_\gamma$, where

$$p_e=\frac{2}{3\pi^2}\int dE\frac{(E^2-m_e^2)^{\frac{3}{2}}}{\exp\left(\frac{E}{T}\right)+1},$$
$$p_\gamma=\frac{1}{3\pi^2}\int dE\frac{E^3}{\exp\left(\frac{E}{T}\right)-1}=\frac{\rho_\gamma}{3} \tag{B3}$$

$\rho_{\nu_\alpha}$ is similar for $\rho_e$, but with the extra prefactor $1/2$ and in the limit of zero mass, so that there is a simple analytic relation between the energy density and temperature $\rho_{\nu_\alpha}=\frac{7\pi^2T_{\nu_\alpha}^4}{120}$. Next, $(d\rho_{\nu_\alpha}/dt)_X$, $(d\rho_{\mathrm{EM}}/dt)_X$ are the source terms due to injections of the particles by $X$'s decays.

$(d\rho_{\nu_\alpha}/dt)_{\mathrm{EM}\leftrightarrow\nu_\alpha}$ is the energy exchange rate between neutrinos and the EM plasma [25]

$$\left(\frac{d\rho_{\nu_\alpha}}{dt}\right)_{\mathrm{EM}\leftrightarrow\nu_\alpha}=\sum_\beta\langle P_{\beta\alpha}\rangle(T_{\mathrm{EM}},\langle E_\nu\rangle)\frac{\delta\rho_{\nu_\beta}}{\delta t}, \tag{B4}$$

with $\langle P_{\beta\alpha}\rangle$ being the neutrino oscillation probabilities from Eq. (42) but evaluated for the mean neutrino energy $\langle E_\nu\rangle=3.15T_\nu$.

$\frac{\delta\rho_{\nu_\beta}}{\delta t}$ is the evolution rate of the individual neutrino flavor; we discuss it in the subsection below.

Finally, $t_0$ is some initial moment of time at which we define the $X$'s initial number density $n_{X,0}\equiv n_X(t_0)$. It is taken as $T(t_0)=20$ MeV.

Now, let us specify the source terms and energy transfer rates. Having the decay probabilities of the mesons and muons (32)–(34), we may define the following source term for neutrinos:

$$\left(\frac{d\rho_{\nu_\alpha}}{dt}\right)_{\mathrm{source}}=\frac{n_X}{\tau_X}\sum_\beta\langle P_{\beta\alpha}\rangle(T_{\mathrm{EM}},\langle E_{\nu_\beta}\rangle)\times\left(\langle E_{X\to\nu_\beta}\rangle+\sum_{y=Y,\bar{Y}}\frac{n_y}{\tau_y}P^y_{\mathrm{decay}}\langle E_{y\to\nu_\beta}\rangle\right), \tag{B5}$$

where

$$\langle E_{A\to\nu_\alpha}\rangle=\sum_i\mathrm{Br}(A\to\nu_\alpha)\cdot\langle E^{(i)}_{A\to\nu_\alpha}\rangle \tag{B6}$$

is the mean energy injected in the neutrino sector by decays of a particle $A$, while $\langle E_{y\to\nu_\alpha}\rangle$ is the mean energy per $y$ decay assuming that all $y'$ appearing per its decay are stable. To calculate it, we take into account only primary decays; i.e., for $K^+\to\mu^++\nu_\mu$ we do not subsequently decay $\mu^+$, as it may disappear. We evaluate these energy fractions using the decay module of the code SensCalc [70]. The EM source term has the form

$$\left(\frac{d\rho_{\mathrm{EM}}}{dt}\right)_{\mathrm{source}}=\frac{n_X}{\tau_X}(m_X-\langle E_{X\to\nu}\rangle), \tag{B7}$$

assuming that the $X$ particle decays at rest.

The neutrino energy transfer rates are

$$\frac{\delta\rho_{\nu_e}}{\delta t}\approx\frac{G_F^2}{\pi^5}[4(g_{eL}^2+g_{eR}^2)F(T_\gamma,T_{\nu_e})+F(T_{\nu_\mu},T_{\nu_e})+F(T_{\nu_\tau},T_{\nu_e})], \tag{B8}$$

$$\frac{\delta\rho_{\nu_\mu}}{\delta t}\approx\frac{G_F^2}{\pi^5}[4(g_{\mu L}^2+g_{\mu R}^2)F(T_\gamma,T_{\nu_\mu})-F(T_{\nu_\mu},T_{\nu_e})+F(T_{\nu_\tau},T_{\nu_\mu})], \tag{B9}$$

$$\frac{\delta\rho_{\nu_\tau}}{\delta t}\approx\frac{G_F^2}{\pi^5}[4(g_{\mu L}^2+g_{\mu R}^2)F(T_\gamma,T_{\nu_\tau})-F(T_{\nu_\tau},T_{\nu_e})-F(T_{\nu_\tau},T_{\nu_\mu})], \tag{B10}$$

with $G_F\approx1.167\times10^{-5}$ GeV$^{-2}$ being the Fermi's constant,

$$g_{eL} = 0.727, \qquad g_{eR} = 0.233,$$
$$g_{\mu L} = -0.273, \qquad g_{\mu R} = 0.233, \tag{B11}$$

and the function

$$F(T_x, T_{\nu_\alpha}) = 32 f_a^{\mathrm{FD}} \mathcal{G}^{\nu_\alpha}_{x,\mathrm{ann}}(T_x)(T_x^9 - T_2^9) + 56 f_s^{\mathrm{FD}} \mathcal{G}^{\nu_\alpha}_{x,\mathrm{scatt}} \times (T_x) T_x^4 T_{\nu_\alpha}^4 (T_x - T_{\nu_\alpha}) \tag{B12}$$

describes the temperature-dependent part of the energy exchange itself. Its analytic part is computed using the Maxwell-Boltzmann approximation and neglects the electron mass. To account for the Fermi-Dirac statistics, the following factors are introduced: $f_a^{\mathrm{FD}} = 0.884$, $f_s^{\mathrm{FD}} = 0.829$. To account for the electron mass, we take the interpolations of the corrections $\mathcal{G}^{\nu_\alpha}_{x,\mathrm{ann}}(T_x)$, $\mathcal{G}^{\nu_\alpha}_{x,\mathrm{scatt}}(T_x)$ from the repository accompanying [25]; they are nonunit only if $x$ represents the EM particle.

## APPENDIX C: DETAILS OF UNINTEGRATED BOLTZMANN EQUATION IMPLEMENTATION

We incorporate the system (42), (45) in a Python code based on Ref. [42] (see footnote 3). As an input, it requires the LLP mass and lifetime, the number density at some initial temperature, and the tabulated decay probabilities. Then, it evolves the neutrino population for a broad range of LLPs $X$ decaying into muons, pions, and neutrinos via a 2-body decay $X \to \nu\bar{\nu}$. These include Higgs-like scalars [28], generic pseudoscalars such as axionlike particles with various coupling schemes [29–33], particles coupled to quark currents like dark photons and $B - L$ mediators [34], and majorons. Other models, such as HNLs and neutralinos, will be added in the future.

We numerically solve Eqs. (42) and (45) with the following dimensionless variables

$$x = m_e a, \qquad y = pa, \qquad z = aT, \tag{C1}$$

approximately normalizing $z \to 1\,(a \to 1/T)$ at high-temperature limit (slight deviations from $z_{\mathrm{in}} = 1$ are due to the entropy conservation of electromagnetic plasma, neutrinos, and antineutrinos [72]). $x, y, z$ characterize time, momentum, and the EM temperature, respectively.

To solve the system of ODEs (42), we use the `RK45` method in `solve_ivp` distributed in SciPy [73]. We linearly discretize the neutrino momentum grid $y_i$ using 301 grid points with $y_{\mathrm{min}} = 0.01$ and $y_{\mathrm{max}} = \max[a_{\mathrm{stop}} E_{\nu,\mathrm{max}}, 40\ \mathrm{MeV}]$. Here, $E_{\nu,\mathrm{max}}$ is the maximal energy of injected neutrinos (depending on the LLP, it may, for example, be half the LLP mass or half the muon mass); $a_{\mathrm{stop}}$ is the scale factor after which further injections are neglected. It is chosen to correspond to a simulated time of $10\tau_X$, when it is reasonable to assume that essentially all LLPs will have decayed. Since $a_{\mathrm{stop}}$ needs to be known *a priori*, it is calculated using a tabulated relation $a(t)$ based on the case when LLPs are absent. Finally, the integration is performed via a summed Simpson's rule. When we compute the integrations inside Eq. (45), we linearly discretize the EM grid $y^{\mathrm{QED}} \in [0.0120]$ using 81 grid points. It is enough, as the EM plasma is rapidly thermalized.

The dynamics of the metastable particles, muons and pions, is incorporated by relating their instant number densities to the number density of LLP and utilizing the decay probabilities (32) as obtained in this study. Namely, in the source term, apart from the direct LLP decays into neutrinos (see, e.g., [17]), we include (see Appendix D for the derivation)

$$\left(\frac{df_{\nu_\alpha}(p)}{dt}\right)_{\mathrm{source}} = \frac{2\pi^2}{p^2}\left|\frac{dn_X}{dt}\right| \Bigg\{ N_\pi^X P^\pi_{\mathrm{decay}} \Bigg[ \langle P_{\alpha\mu}\rangle \cdot \delta\left(E_\nu - \frac{m_\pi^2 - m_\mu^2}{2m_\pi}\right) + P^\mu_{\mathrm{decay}}\left(\langle P_{\alpha\mu}\rangle \mathcal{F}_{\nu_\mu}(p) + \langle P_{\alpha e}\rangle \mathcal{F}_{\nu_e}(p)\right)\Bigg] + N_\mu^X P^\mu_{\mathrm{decay}}\left[\langle P_{\alpha\mu}\rangle \mathcal{F}_{\nu_\mu}(p) + \langle P_{\alpha e}\rangle \mathcal{F}_{\nu_e}(p)\right]\Bigg\}, \tag{C2}$$

Here

$$\left|\frac{dn_X}{dt}\right| = \frac{n_X(t)}{\tau_X} \tag{C3}$$

is the differential rate of LLP's decay and $n_X(t)$ is given by Eq. (22). $\delta(\ldots)$ is the Dirac $\delta$ function. The first squared brackets describe the contribution from LLPs decaying into pions. It includes the direct decay $\pi^+ \to \mu^+\nu_\mu$ into muon neutrinos (resulting the energy distribution to a $\delta$ function) and the secondary muon decay, which is $\mu^+ \to e^+\nu_e\bar{\nu}_\mu$; the functions $\mathcal{F}_{\mu_\alpha}(p)$ are energy distributions of the neutrinos. The second brackets are from LLPs decaying into muons. The decay probabilities entering Eq. (C2) are taken as input from the tabulated files generated by our *Mathematica* code. Further details of usage may be found on the GitHub page.

Let us briefly discuss the limitations of the method. It utilizes the discretization approach from [56], and hence, its performance heavily depends on the maximal neutrino energy in the system, $E_{\nu,\mathrm{max}}$. The scaling of the computational time is

$$t_{\mathrm{computation}} \propto E^4_{\nu,\mathrm{max}}, \tag{C4}$$

Here, one power of $E_{\nu,\mathrm{max}}$ comes from the scaling of the number of time steps required to resolve the thermalization of neutrinos, another power is due to the number of bins (scaling as $\propto E_{\nu,\mathrm{max}}$), and two powers are due to the

dimensionality of the analytically reduced collision integral (see Ref. [26] for more discussions). In the case of new physics decaying into muons and pions, when $E_{\nu,\max} \approx m_\mu/2$, the computational time needed to obtain $N_{\rm eff}$ for the single LLP setup is $\mathcal{O}(1\ h)$. However, it increases by orders of magnitude if, e.g., kaons are present (as then $E_{\nu,\max} \approx m_K/2$), making the solver inapplicable in practice.

Therefore, in this study, we restrict ourselves to scenarios of LLPs decaying into pions, muons, or neutrinos with energies $E_\nu \lesssim 100$ MeV. The cases of decays into kaons, as well as neutrinos with higher energies, may be studied using different approaches, such as the direct simulation Monte Carlo from Refs. [26,27].

## APPENDIX D: DERIVATION OF THE SOURCE TERM IN THE UNINTEGRATED BOLTZMANN EQUATIONS

In this appendix, we derive the source term for neutrinos, Eq. (C2), i.e., the collision term from metastable particle decays to neutrinos, in the unintegrated Boltzmann equations.

We consider the decay process

$$X \to Y \to \nu_\alpha + \ldots, \tag{D1}$$

where, as usual, $X$ denotes the LLP, while $Y$ stays for metastable SM particles decaying into neutrinos. We can easily generalize the method to include other decay processes such as $X \to Y \to Z \to \nu_\alpha + \ldots$ and $X \to \nu_\alpha + \ldots$, etc.

The source term for $Y \to \nu_\alpha + \ldots$ in the neutrino Boltzmann equations is

$$\left(\frac{df_{\nu_\alpha}}{dt}\right)_{\rm source} = \langle P_{\alpha\beta}\rangle I_{Y\to\nu_\beta+\cdots}, \tag{D2}$$

where $I_{Y\to\nu_\beta+\ldots}$ is the corresponding collision term,

$$I_{Y\to\nu_\beta+\cdots} = \frac{1}{2E_{\nu_\beta}}\int \frac{d^3p_Y}{(2\pi)^3 2E_Y}\prod_{f=2}\frac{d^3p_f}{(2\pi)^3 2E_f} S|\mathcal{M}_{Y\to\nu_\beta+\cdots}|^2 \times (2\pi)^3\delta^{(4)}(p_Y - p_{\nu_\beta} - \cdots) f_Y(p_Y), \tag{D3}$$

where $f = 1$ corresponds $\nu_\beta$, $f_Y(p_Y)$ is the distribution function for $Y$, and we neglect the Pauli-blocking/Bose-enhancement factors. $Y$ becomes nonrelativistic before their decay due to the strong EM scattering. In this case, the collision term is

$$I_{Y\to\nu_\beta+\cdots} = \frac{n_Y}{2E_{\nu_\beta}2m_Y}\int \prod_{f=2}\frac{d^3p_f}{(2\pi)^3 2E_f} S|\mathcal{M}_{Y\to\nu_\beta+\cdots}|^2 \times (2\pi)^3\delta^{(4)}(m_Y - p_{\nu_\beta} - \cdots). \tag{D4}$$

The decay rate for $Y \to \nu_\beta + \ldots$ at the rest frame of $Y$ is

$$\Gamma_Y = \frac{1}{2m_Y}\int \frac{d^3p_{\nu_\beta}}{(2\pi)^3 2E_{\nu_\beta}}\prod_{f=2}\frac{d^3p_f}{(2\pi)^3 2E_f} S|\mathcal{M}_{Y\to\nu_\beta+\cdots}|^2 \times (2\pi)^3\delta^{(4)}(m_Y - p_{\nu_\beta} - \cdots). \tag{D5}$$

The collision term is rewritten in terms of $d\Gamma_Y/dp_{\nu_\beta}, \Gamma_Y$ and $\tau_Y = \Gamma_Y^{-1}$,

$$I_{Y\to\nu_\beta+\cdots} = \frac{n_Y}{\tau_Y}\frac{1}{\Gamma_Y}\frac{d\Gamma_Y}{dp_{\nu_\beta}}\frac{2\pi^2}{p_{\nu_\beta}^2}, \tag{D6}$$

where $n_Y/\tau_Y$ is the production rate of neutrinos from Y decays and $\frac{1}{\Gamma_Y}\frac{d\Gamma_Y}{dp_{\nu_\beta}} \equiv \mathcal{F}_{\nu_\beta}(p_{\nu_\beta})$ is the momentum distribution (normalized to be unity) for a neutrino by a decay. The number density for $Y$ produced by the $X$ decay is given in Eq. (35),

$$n_Y(t) = n_X(t) N_Y^X \frac{\tau_Y}{\tau_X} P^Y_{\rm decay}(t), \tag{D7}$$

where $n_X(t)$ is given by Eq. (22). Finally we obtain the source term for $X \to Y \to \nu_\beta + \ldots$,

$$\left(\frac{df_{\nu_\alpha}(p)}{dt}\right)_{\rm source} = \langle P_{\alpha\beta}\rangle \times \frac{n_X(t)}{\tau_X} N_Y^X \mathcal{F}_{\nu_\beta}(p) P^Y_{\rm decay}(t), \tag{D8}$$

where $\frac{n_X(t)}{\tau_X}N_Y^X$ is the production rate of $Y$ from $X$ decays. Following Sec. IV, we obtain the decay probability for $Y$, $P^Y_{\rm decay}$, and then we can compute the source term. Generalizing the decay processes, we obtain the source term for the pion decay to neutrinos (C2).


[1] A. D. Dolgov, Neutrinos in cosmology, Phys. Rep. **370**, 333 (2002).

[2] S. Sarkar, Big bang nucleosynthesis and physics beyond the standard model, Rep. Prog. Phys. **59**, 1493 (1996).

[3] 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).

[4] K. Kohri, Primordial nucleosynthesis and hadronic decay of a massive particle with a relatively short lifetime, Phys. Rev. D **64**, 043515 (2001).
[5] S. Hannestad, What is the lowest possible reheating temperature?, Phys. Rev. D **70**, 043506 (2004).
[6] M. Pospelov and J. Pradler, Big bang nucleosynthesis as a probe of new physics, Annu. Rev. Nucl. Part. Sci. **60**, 539 (2010).
[7] M. Kawasaki, K. Kohri, T. Moroi, and Y. Takaesu, Revisiting big-bang nucleosynthesis constraints on long-lived decaying particles, Phys. Rev. D **97**, 023502 (2018).
[8] A. Boyarsky, M. Ovchynnikov, O. Ruchayskiy, and V. Syvolap, Improved big bang nucleosynthesis constraints on heavy neutral leptons, Phys. Rev. D **104**, 023517 (2021).
[9] J. R. Ellis, J. E. Kim, and D. V. Nanopoulos, Cosmological gravitino regeneration and decay, Phys. Lett. **145B**, 181 (1984).
[10] T. Moroi, H. Murayama, and M. Yamaguchi, Cosmological constraints on the light stable gravitino, Phys. Lett. B **303**, 289 (1993).
[11] M. Kawasaki and T. Moroi, Gravitino production in the inflationary universe and the effects on big bang nucleosynthesis, Prog. Theor. Phys. **93**, 879 (1995).
[12] G. F. Giudice, E. W. Kolb, and A. Riotto, Largest temperature of the radiation era and its cosmological implications, Phys. Rev. D **64**, 023508 (2001).
[13] T. Kanzaki, M. Kawasaki, K. Kohri, and T. Moroi, Cosmological constraints on neutrino injection, Phys. Rev. D **76**, 105017 (2007).
[14] A. Fradette and M. Pospelov, BBN for the LHC: Constraints on lifetimes of the Higgs portal scalars, Phys. Rev. D **96**, 075033 (2017).
[15] A. Fradette, M. Pospelov, J. Pradler, and A. Ritz, Cosmological beam dump: Constraints on dark scalars mixed with the Higgs boson, Phys. Rev. D **99**, 075004 (2019).
[16] T. Hasegawa, N. Hiroshima, K. Kohri, R. S. L. Hansen, T. Tram, and S. Hannestad, MeV-scale reheating temperature and thermalization of oscillating neutrinos by radiative and hadronic decays of massive particles, J. Cosmol. Astropart. Phys. 12 (2019) 012.
[17] 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.
[18] A. Boyarsky, M. Ovchynnikov, N. Sabti, and V. Syvolap, When feebly interacting massive particles decay into neutrinos: The Neff story, Phys. Rev. D **104**, 035006 (2021).
[19] L. Mastrototaro, P. D. Serpico, A. Mirizzi, and N. Saviano, Massive sterile neutrinos in the early Universe: From thermal decoupling to cosmological constraints, Phys. Rev. D **104**, 016026 (2021).
[20] H. Rasmussen, A. McNichol, G. M. Fuller, and C. T. Kishimoto, Effects of an intermediate mass sterile neutrino population on the early Universe, Phys. Rev. D **105**, 083513 (2022).
[21] J. Alvey, M. Escudero, and N. Sabti, What can CMB observations tell us about the neutrino distribution function?, J. Cosmol. Astropart. Phys. 02 (2022) 037.
[22] J. Alvey, M. Escudero, N. Sabti, and T. Schwetz, Cosmic neutrino background detection in large-neutrino-mass cosmologies, Phys. Rev. D **105**, 063501 (2022).
[23] M. Escudero, T. Schwetz, and J. Terol-Calvo, A seesaw model for large neutrino masses in concordance with cosmology, J. High Energy Phys. 02 (2023) 142; 06 (2024) 119(A).
[24] D. Naredo-Tuero, M. Escudero, E. Fernández-Martínez, X. Marcano, and V. Poulin, Living at the edge: A critical look at the cosmological neutrino mass bound, Phys. Rev. D **110**, 123537 (2024).
[25] M. Escudero Abenza, Precision early universe thermodynamics made simple: $N_{\rm eff}$ and neutrino decoupling in the standard model and beyond, J. Cosmol. Astropart. Phys. 05 (2020) 048.
[26] M. Ovchynnikov and V. Syvolap, How new physics affects primordial neutrinos decoupling: Direct Simulation Monte Carlo approach, Phys. Rev. D **111**, 063527 (2025).
[27] M. Ovchynnikov and V. Syvolap, Primordial neutrinos and new physics: Novel approach to solving neutrino Boltzmann equation, Phys. Rev. Lett. **134**, 101003 (2025).
[28] I. Boiarska, K. Bondarenko, A. Boyarsky, V. Gorkavenko, M. Ovchynnikov, and A. Sokolenko, Phenomenology of GeV-scale scalar portal, J. High Energy Phys. 11 (2019) 162.
[29] J. Beacham *et al.*, Physics beyond colliders at CERN: Beyond the standard model working group report, J. Phys. G **47**, 010501 (2020).
[30] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, The low-energy effective theory of axions and ALPs, J. High Energy Phys. 04 (2021) 063.
[31] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, Flavor probes of axion-like particles, J. High Energy Phys. 09 (2022) 056.
[32] D. Aloni, Y. Soreq, and M. Williams, Coupling QCD-scale axionlike particles to gluons, Phys. Rev. Lett. **123**, 031803 (2019).
[33] G. Dalla Valle Garcia, F. Kahlhoefer, M. Ovchynnikov, and A. Zaporozhchenko, Phenomenology of axionlike particles with universal fermion couplings revisited, Phys. Rev. D **109**, 055042 (2024).
[34] P. Ilten, Y. Soreq, M. Williams, and W. Xue, Serendipity in dark photon searches, J. High Energy Phys. 06 (2018) 004.
[35] K. Bondarenko, A. Boyarsky, D. Gorbunov, and O. Ruchayskiy, Phenomenology of GeV-scale heavy neutral leptons, J. High Energy Phys. 11 (2018) 032.
[36] M. H. Reno and D. Seckel, Primordial nucleosynthesis: The effects of injecting hadrons, Phys. Rev. D **37**, 3441 (1988).
[37] M. Pospelov and J. Pradler, Metastable GeV-scale particles as a solution to the cosmological lithium problem, Phys. Rev. D **82**, 103514 (2010).
[38] G. B. Gelmini, M. Kawasaki, A. Kusenko, K. Murai, and V. Takhistov, Big bang nucleosynthesis constraints on sterile neutrino and lepton asymmetry of the Universe, J. Cosmol. Astropart. Phys. 09 (2020) 051.
[39] M. Kawasaki, K. Kohri, and T. Moroi, Big-bang nucleosynthesis and hadronic decay of long-lived massive particles, Phys. Rev. D **71**, 083502 (2005).
[40] M. Ovchynnikov, Metastable-dynamics (2025), https://github.com/maksymovchynnikov/Metastable-dynamics.
[41] M. Ovchynnikov, Metastable-dynamics (2025), 10.5281/zenodo.14020343.

[42] K. Akita and M. Yamaguchi, A precision calculation of relic neutrino decoupling, J. Cosmol. Astropart. Phys. 08 (2020) 012.

[43] G. Baur, Nudec_LLP_Solver (2025), https://github.com/baugid/Nudec_LLP_Solver.

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

[45] 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, J. Cosmol. Astropart. Phys. 07 (2019) 014.

[46] W. Chao, J.-J. Feng, and M.-J. Jin, Neff from an excited dark matter state, Phys. Rev. D **107**, 015022 (2023).

[47] Y.-M. Chen and Y. Zhang, BBN constraint on heavy neutrino production and decay, arXiv:2410.07343.

[48] S. Chang, S. Ganguly, T. H. Jung, T.-S. Park, and C. S. Shin, Constraining MeV to 10 GeV Majorons by big bang nucleosynthesis, Phys. Rev. D **110**, 015019 (2024).

[49] K. Akita, G. Baur, M. Ovchynnikov, T. Schwetz, V. Syvolap, companion Letter, New physics decaying into metastable particles: Impact on cosmic neutrinos, Phys. Rev. Lett. **134**, 121001 (2025).

[50] R. L. Workman *et al.* (Particle Data Group Collaboration), Review of particle physics, Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).

[51] M. Escudero, Neutrino decoupling beyond the Standard Model: CMB constraints on the dark matter mass with a fast and precise $N_{\rm eff}$ evaluation, J. Cosmol. Astropart. Phys. 02 (2019) 007.

[52] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641**, A6 (2020); **652**, C4(E) (2021).

[53] A. Vlasenko, G. M. Fuller, and V. Cirigliano, Neutrino quantum kinetics, Phys. Rev. D **89**, 105004 (2014).

[54] C. Volpe, D. Väänänen, and C. Espinoza, Extended evolution equations for neutrino propagation in astrophysical and cosmological environments, Phys. Rev. D **87**, 113010 (2013).

[55] J. Serreau and C. Volpe, Neutrino-antineutrino correlations in dense anisotropic media, Phys. Rev. D **90**, 125040 (2014).

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

[57] A. Strumia and F. Vissani, Neutrino masses and mixings and..., arXiv:hep-ph/0606054.

[58] A. F. Heckler, Astrophysical applications of quantum corrections to the equation of state of a plasma, Phys. Rev. D **49**, 611 (1994).

[59] N. Fornengo, C. W. Kim, and J. Song, Finite temperature effects on the neutrino decoupling in the early universe, Phys. Rev. D **56**, 5123 (1997).

[60] J. I. Kapusta and C. Gale, *Finite-Temperature Field Theory: Principles and Applications*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England, 2011).

[61] J. J. Bennett, G. Buldgen, M. Drewes, and Y. Y. Y. Wong, Towards a precision calculation of the effective number of neutrinos $N_{\rm eff}$ in the Standard Model I: The QED equation of state, J. Cosmol. Astropart. Phys. 03 (2020) 003; 03 (2021) A01.

[62] G. Mangano, G. Miele, S. Pastor, and M. Peloso, A precision calculation of the effective number of cosmological neutrinos, Phys. Lett. B **534**, 8 (2002).

[63] J. J. Bennett, G. Buldgen, P. F. De Salas, M. Drewes, S. Gariazzo, S. Pastor, and Y. Y. Y. Wong, Towards a precision calculation of $N_{\rm eff}$ in the standard model II: Neutrino decoupling in the presence of flavour oscillations and finite-temperature QED, J. Cosmol. Astropart. Phys. 04 (2021) 073.

[64] J. Froustey, C. Pitrou, and M. C. Volpe, Neutrino decoupling including flavour oscillations and primordial nucleosynthesis, J. Cosmol. Astropart. Phys. 12 (2020) 015.

[65] M. Cielo, M. Escudero, G. Mangano, and O. Pisanti, Neff in the standard model at NLO is 3.043, Phys. Rev. D **108**, L121301 (2023).

[66] G. Jackson and M. Laine, QED corrections to the thermal neutrino interaction rate, J. High Energy Phys. 05 (2024) 089.

[67] M. Drewes, Y. Georis, M. Klasen, L. P. Wiggering, and Y. Y. Y. Wong, Towards a precision calculation of $N_{\rm eff}$ in the standard model III: Improved estimate of NLO corrections to the collision integral, J. Cosmol. Astropart. Phys. 06 (2024) 032.

[68] P. Ade *et al.* (Simons Observatory Collaboration), The Simons Observatory: Science goals and forecasts, J. Cosmol. Astropart. Phys. 02 (2019) 056.

[69] N. Aghanim *et al.* (Planck Collaboration), Planck 2018 results. I. Overview and the cosmological legacy of Planck, Astron. Astrophys. **641**, A1 (2020).

[70] M. Ovchynnikov, J.-L. Tastet, O. Mikulenko, and K. Bondarenko, Sensitivities to feebly interacting particles: Public and unified calculations, Phys. Rev. D **108**, 075028 (2023).

[71] C. Bierlich *et al.*, A comprehensive guide to the physics and usage of Pythia 8.3, SciPost Phys. Codebases **2022**, 8 (2022).

[72] 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. **B543**, 269 (1999).

[73] P. Virtanen *et al.* (SciPy 1.0 Contributors), SciPy 1.0: Fundamental algorithms for scientific computing in Python, Nat. Methods **17**, 261 (2020).