

How new physics affects primordial neutrinos decoupling: Direct Simulation Monte Carlo approach

Maksym Ovchynnikov*

*Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland and
Institut für Astroteilchen Physik, Karlsruher Institut für Technologie (KIT),
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany*

Vsevolod Syvolap†

Instituut-Lorentz, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands

(Dated: September 24, 2024)

Cosmological observations from Big Bang Nucleosynthesis and the Cosmic Microwave Background (CMB) offer crucial insights into the Early Universe, enabling us to trace its evolution back to lifetimes as short as 0.01 seconds. Upcoming CMB spectrum measurements, such as those underway at the Simons Observatory, will achieve unprecedented precision, allowing for more accurate extraction of information about the properties of the primordial plasma and, in particular, primordial neutrinos. This provides an opportunity to test whether these properties align with the predictions of the standard cosmological model or indicate the presence of new physics that influenced the evolution of the MeV-temperature plasma. A key component in understanding how new physics may have affected primordial neutrinos is solving the neutrino Boltzmann equation. In this paper, we present a novel approach to solving this equation that offers model independence, transparency, and computational efficiency – features that current state-of-the-art methods lack. We demonstrate a proof-of-concept implementation and apply it to several toy scenarios, showcasing key aspects of the primordial plasma’s evolution in the presence of new physics.

I. INTRODUCTION

Primordial neutrinos are an important messenger from the Early Universe, bringing us information about the state of the Universe at times as early as $t \ll 1$ s. Their direct detection is significantly more challenging than that of primordial photons due to their tiny interaction cross-section, which is governed by weak interactions. However, numerous populating the primordial plasma, they affected a number of cosmological observables. It makes it possible to indirectly extract information about their properties from precise cosmic measurements. In particular, they contribute to the number of ultrarelativistic (UR) degrees of freedom,

$$N_{\text{eff}} = \frac{8}{7} \left(\frac{11}{4} \right)^{\frac{4}{3}} \frac{\rho_{\text{UR}} - \rho_{\gamma}}{\rho_{\gamma}} \quad (1)$$

This quantity determines the characteristic features of the Cosmic Microwave Background (CMB) and may be extracted from its measurements.

It is not only the total neutrino energy density that is important. Another essential property is the shape of the neutrino energy distribution function. It handles the neutron-to-proton conversion at MeV temperatures, which determines the onset of Big Bang Nucleosynthesis (BBN), as well as affects Baryon Acoustic Oscillations (BAO) [1, 2]. The shape of the distribution may significantly modify the cosmological neutrino mass bound [3].

Assuming the standard cosmological history, based on the Λ CDM model, N_{eff} is fully represented by neutrino, and its value is 3.043-3.044 [4–10]. The shape of the neutrino distribution is very close to the Fermi-Dirac distribution, with tiny distortions in the high-energy tail. Finally, there is no asymmetry between neutrinos and antineutrinos. Altogether, it serves as an input to the Standard Big Bang Nucleosynthesis model, which predicts the helium abundance $Y_p = 0.247 \pm 0.00017$ (see, e.g., [11, 12]). These numbers agree with the current BBN and CMB observations. In particular, the measurements performed by the Planck collaboration [13] constrain $N_{\text{eff}} = 2.99^{+0.34}_{-0.33}$ at 95% CL, whereas the primordial Helium abundance measurements are in a range 0.233-0.2573, obtained by combining the observations from the works [14–21].

However, the uncertainty window of these observations leaves room for sizeable deviations from standard neutrino properties that may potentially originate from the presence of new physics at temperatures $T \lesssim 5$ MeV, when neutrinos start decoupling. Examples of such scenarios include the presence of non-standard neutrino interactions [22, 23], a lepton asymmetry in the neutrino sector [24, 25], a change in the expansion dynamics of the Universe, and the injection of non-thermal neutrinos by hypothetical Long-Lived Particles, or LLPs [26–35]. The accuracy of the CMB measurements will be significantly improved with the future observations with Simons Observatory [36] (which has started collecting the data on June 2024) and CMB-S4 mission [37]. They will be able to measure N_{eff} with a percent precision, thus providing a unique potential to shed light on properties of the new physics or constrain it in case of the absence of deviations

* maksym.ovchynnikov@cern.ch

† v.syvolap@umail.leidenuniv.nl

from Λ CDM.

Under certain approximations of neutrino oscillations, understanding the impact of the new physics effects on the neutrino properties requires solving the Boltzmann equation on the neutrino distribution function f_{ν_α} :

$$\frac{\partial f_{\nu_\alpha}}{\partial t} - pH \frac{\partial f_{\nu_\alpha}}{\partial p} = \mathcal{I}_{\text{coll},\alpha}[f_{\nu_\alpha}] \quad (2)$$

Here, p is neutrinos' momentum, H is the Hubble factor accounting for the expansion of the Universe, and $\mathcal{I}_{\text{coll},\alpha}$ is the collision integral that takes care of the microscopic of the thermalization.

The main approach considered in literature is to reduce the integration inside $\mathcal{I}_{\text{coll}}$ analytically as much as possible and convert the complex integrodifferential equation (2) into a system of the ordinary differential equations by discretizing the grid of the comoving momenta (see the pioneering work [38] as well as later realizations [7, 39, 40], and references therein). The method has also been used to study some well-motivated scenarios with LLPs such as Heavy Neutral Leptons (HNLs) [26, 28, 32–35] and particles in late reheating scenarios [27].

However, several problems exist with this approach. First, it has a limited range of applicability, requiring analytic matrix elements for the processes and high reducibility of the dimensionality of the integration in $\mathcal{I}_{\text{coll},\alpha}$. Second, even within the case studies, its computational complexity quickly grows if high-energy neutrinos are present in the system. For instance, depending on the grid density, solving the Boltzmann equation under the presence of HNLs with masses just $\simeq 200$ MeV (injecting neutrinos with energies up to 100 MeV) may take days [32].

In addition, the method itself is very complex. The analytic reduction of the collision integral is highly non-trivial, the comoving grid density has to be adjusted to the model's parameters, and solver stability must be carefully verified. An indirect consequence of this is that there is the existing discrepancy between the predictions of various neutrino Boltzmann codes for the behavior of N_{eff} in the presence of the injection of high-energy neutrinos with energies well exceeding the plasma temperature. While some studies predict that injection of such neutrinos would increase N_{eff} , the other show the opposite [26, 28, 33–35].

In this paper, we address these issues by developing proof-of-principle of a novel approach to solving the neutrino Boltzmann equation based on the so-called Direct Simulation Monte Carlo (DSMC) [41–44]. Its basis is the numerical particle representation of the Boltzmann equation: one starts with a large number of particles obeying some initial condition in momentum and spatial spaces and then directly simulates their interactions to study the equilibration. Due to the straightforwardness of the method, DSMC directly calculates the linear functionals, e.g., the number and energy densities, velocities, etc., without any simplifications. The existing case

studies describe the implementations of DSMC that efficiently simulate collisions of a number of particles as large as 10^8 [45, 46]. The simplicity of the scheme describing the interactions and absence of momentum discretization automatically release the DSMC approach from most of the problems described above.

This work also serves as the companion to the paper [47], which presents a summary of the results.

The paper is organized as follows. In Sec. II, we review the properties of the primordial plasma around the neutrino decoupling, considering both the standard cosmological scenario and setups with new physics. Sec. III is devoted to a discussion on the existing approaches to solve the neutrino Boltzmann equation. In Sec. IV, we describe the basics of the DSMC approach and, in particular, why it may be well-applicable to studying the dynamics of primordial neutrinos. Sec. V discusses the necessary modifications to the DSMC simulation required to study the primordial MeV plasma, and how they can be implemented. In Sec. VI, we present our proof-of-principle realization of the approach and different cross-checks we performed to validate it against well-defined scenarios. In Sec. VII, we apply the developed approach to a few case studies simplifying various physics setups, highlighting the variety of the applicability of the neutrino DSMC and the importance of using full Boltzmann equations. Finally, in Sec. VIII, we make conclusions.

II. PRIMORDIAL PLASMA AT MEV TEMPERATURES

At cosmic times relevant for our discussion, primordial SM plasma consists of light particles – neutrinos $\nu, \bar{\nu}$, electromagnetically (EM) interacting light particles (electrons e^- , positrons e^+ , and photons γ), as well as baryons $B = p, n$.

The homogeneous and isotropic Universe expands with the rate $H(t) = \dot{a}(t)/a(t)$, where $a(t)$ is the scale factor, and H is the Hubble parameter. Assuming spatial flatness and neglecting the dark energy contribution, we get

$$H(t) = \frac{1}{M_{\text{Pl}}} \sqrt{\frac{8\pi}{3} \rho_{\text{Universe}}} \quad (3)$$

where M_{Pl} is the Planck mass, and ρ_{Universe} the total energy density of the Universe.

To understand the scaling of ρ_{Universe} , we need to discuss different components of the primordial plasma and, in particular, their interactions.

A. EM plasma and nucleons

Let us first consider the EM plasma. Examples of the processes are Compton scattering and electron-positron annihilation into a pair of photons. The corresponding rate well exceeds the Hubble parameter for times

$t \lesssim 10^4$ s, which includes the period we are interested in.¹ This means that the population of the EM particles can always be well described by just one quantity – the temperature of the EM plasma $T_{\text{EM}} \equiv T$.

The distribution function f_{e^\pm} of electrons and positrons is Fermi-Dirac, while for photons it is Bose-Einstein:²

$$f_{e^\pm}(p, T) = \frac{1}{\exp\left[\frac{\sqrt{p^2 + m_e^2}}{T}\right] + 1}, \quad (4)$$

$$f_\gamma(p, T) = \frac{1}{\exp\left[\frac{p}{T}\right] - 1}, \quad (5)$$

with the electron's mass $m_e \approx 0.511$ MeV. The temperature T is related to the total energy density of the EM particles ρ_{EM} by the formula

$$\rho_{\text{EM}}(T) = \rho_{e^\pm}(T) + \rho_\gamma(T) \quad (6)$$

Here, the energy densities of e^\pm, γ are

$$\rho_{e^\pm}(T) = g_{e^\pm} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \sqrt{p^2 + m_e^2} f_{e^\pm}(p, T), \quad (7)$$

$$\rho_\gamma(T) = g_\gamma \int \frac{d^3\mathbf{p}}{(2\pi)^3} p f_\gamma(p, T), \quad (8)$$

with the factors $g_{e^\pm} = 4$ and $g_\gamma = 2$ staying for the spin and charge degrees of freedom.

In terms of T , the Hubble factor (3) can be rewritten as

$$H(T) \equiv \frac{T^2}{M_{\text{pl}}^*}, \quad M_{\text{pl}}^* \approx \frac{M_{\text{pl}}}{1.66\sqrt{g_*(T)}} \quad (9)$$

where g_* the effective number of relativistic species: $g_* = \rho_{\text{Universe}}/\frac{\pi^2}{30}T^4$, with g_i being the number of spin and charge degrees of freedom.³

Finally, the number density of baryons B in the Early Universe, n_B , may be expressed in terms of the baryon-to-photon ratio η_B and the photon number density $n_B = \eta_B n_\gamma$. The value of η_B at late times $T \simeq 1$ MeV may be extracted from the CMB measurements, giving the central value $\eta_{B, \text{Planck}} = 6.09 \cdot 10^{-10}$ [13].

The relative ratio between protons and neutrons, important for BBN, is handled by their weak interactions with neutrinos and e^\pm particles, which drive the $p \leftrightarrow n$ conversion, so the baryons are coupled to the UR content of the plasma. However, because of the tiny number density and the absence of other hadrons in the plasma in the standard scenario, nucleons play a negligible role in the thermodynamics of the Universe at MeV temperatures.

B. Neutrinos

Let us now discuss neutrinos. Generically, there may be an asymmetry between neutrinos and antineutrinos, but the minimal cosmological setup assumes zero asymmetry.⁴ Neutrinos interact with themselves and e^\pm particles via the weak force. The dimensional estimate for the weak interaction rates gives

$$\Gamma_{\text{weak}} \simeq n_\nu \cdot \langle \sigma v \rangle \sim G_F^2 T^5, \quad (10)$$

where we assumed that the neutrinos have the thermal equilibrium with the EM plasma. Namely, $n_\nu \propto T^3$ is the neutrino number density, while $\langle \sigma v \rangle$ is the thermally averaged cross-section, which scales as

$$\langle \sigma v \rangle \sim G_F^2 \langle s \rangle \sim G_F^2 T^2, \quad (11)$$

and $G_F \approx 1.167 \cdot 10^{-5}$ GeV⁻² is the Fermi coupling. The important feature is that the cross-section scales with the energies of the interacting particles (we will return to it in Sec. VII).

The rate becomes comparable to the Hubble expansion rate of the Universe already at $T \sim 1$ MeV. As a result, at these temperatures, the weak reactions are no longer able to maintain equilibrium in the neutrinos sector, and the latter gradually decouple [49]. The shape of their spectrum in Λ CDM closely follows the Fermi-Dirac one. Its temperature T_{ν_α} remains equal to the EM temperature until the annihilation of electron-positron pairs, which happens around $T \sim m_e$. Then, their relation may be found from the entropy conservation law, giving $T_{\nu_\alpha} \approx (4/11)^{1/3} T$. Extended neutrino decoupling introduces a small correction to this relation, giving Eq. (1).

Neutrino interaction processes. Let us now discuss neutrino interactions in more detail. They include elastic scatterings off neutrinos and e^\pm and annihilations:

$$\nu_\alpha + e^\pm \leftrightarrow \nu_\alpha + e^\pm, \quad \nu_\alpha + \bar{\nu}_\alpha \leftrightarrow e^- + e^+, \quad (12)$$

$$\nu_\alpha + \nu_\beta \leftrightarrow \nu_\alpha + \nu_\beta, \quad \nu_\alpha + \bar{\nu}_\alpha \leftrightarrow \nu_\beta + \bar{\nu}_\beta, \quad (13)$$

as well as charge-conjugated ones [32]. The other reactions include the electroweak corrections, such as subdominant $e^+e^- \rightarrow \nu_\alpha \bar{\nu}_\alpha \gamma$.

The MeV plasma is “flavor-asymmetric” in the sense that electrons and positrons are present in plasma, while

¹ The decoupling of EM particles happens much later, during the recombination epoch, which corresponds to a much larger timescale $t_{\text{CMB}} \sim 10^5$ years. The thermalization time is much shorter than any relevant timescale for the electromagnetic temperatures above $T \gtrsim 1$ keV. At lower temperatures, for example, by injecting high-energy e^\pm, γ s, we have a chance for them to photodisintegrate primordial nuclei before the EM particles thermalize [48].

² Due to electroneutrality of the Universe, the chemical potential is vanishingly small, $\mu_{e^\pm}/T \sim \eta_B \simeq 10^{-9}$, where η_B is the baryon-to-photon ratio.

³ Assuming the Λ CDM scenario and that all the species are in perfect equilibrium, it is $g_* \approx g_\gamma + 7/8(g_e + g_\nu) = 10.75$.

⁴ Or at the level of baryon-to-photon ratio, which is negligible.

μ and τ leptons are not. Given the structure of the charged current of weak interactions, which includes the lepton and the corresponding neutrino, the direct interaction rate of ν_e with e^\pm is larger than the rate of the corresponding scatterings but with $\nu_{\mu,\tau}$. Because of this, one can naively expect that $\nu_{\mu,\tau}$ s decouple earlier from the EM plasma, while ν_e s are kept longer in equilibrium. However, besides the interactions (13), neutrinos also experience flavor transitions called oscillations. The oscillations generically appear because the neutrino charge eigenstates do not coincide with the mass eigenstates.

In the primordial plasma, the neutrino oscillation rate is severely affected by the dense medium. Namely, neutrinos acquire a correction to the self-energy caused by interactions with electrons and positrons [49]. It effectively translates to a potential $\mathcal{V}_{\text{eff}}^{(\nu_\alpha)}$ in the Hamiltonian describing the propagation of neutrinos ν_α . The functional form of the potential is $\mathcal{V}_{\text{eff}}^{(\nu_\alpha)} = C^\alpha \frac{G_F^2 T^4 E_\nu}{\alpha_{\text{EM}}}$, where C^α is neutrino-dependent constant.

If $\mathcal{V}_{\text{eff}}^{(\nu_\alpha)}$ is higher than the energy splitting for different neutrino eigenstates $\Delta m^2/2E_\nu$, the mixing angle is effectively suppressed, and oscillations can be ignored. Therefore, the oscillations are absent at high temperatures and/or for high-energy neutrinos. In Λ CDM, they effectively turn on at $T \simeq 3$ MeV.

In total, because of oscillations, the decoupling of three neutrino flavors occurs in a similar fashion.

C. How new physics may spoil properties of primordial plasma

There is a variety of ways of introducing new physics to the primordial plasma. They will change the dynamics of the primordial plasma, in particular, departing the neutrino properties from the Λ CDM ones. To be specific, let us consider the scenario appearing in many well-motivated extensions of the Standard Model, adding LLPs (with mass $m \gg T$).

To significantly affect the Universe, such particles need to be out-of-equilibrium relics. Before decaying, they would increase the energy density of the Universe and hence modify the Hubble factor. After decaying, their influence gets split into many contributions. First, they still modify the dynamics of the Universe by introducing the dilution to the scale factor $a(t)$. It influences the behavior and value of η_B at MeV temperatures.⁵ Second, their decay products may either constitute additional species (“dark radiation”) or inject an energy into the population of neutrinos and the EM particles. The EM population gets immediately thermalized, which results in an increase of T_{EM} , while the neutrino injections

cause the spectral distortions. Since neutrinos with different energies interact at different rates, much slower than the EM particles, the distortions will not disappear, affecting the total neutrino number and energy densities, as well as the $p \leftrightarrow n$ conversion rates.

Under such scenarios, the nucleons may also be involved in the thermodynamics of the Universe in a non-trivial way. Decaying LLPs may inject relatively long-lived mesons such as π^\pm, K^\pm, K_L . Before decaying, these particles experience numerous interactions with the SM plasma particles and themselves. Scattering off nucleons surprisingly becomes very efficient – the smallness of η_B is compensated by the largeness of the nucleon interaction cross-section, driven by the strong force [31]. Because of these scatterings, the mesons change the distribution of their energy among the neutrino and EM sector [50], which leads to the impact on the time-temperature relation $t(T)$ and neutrino properties.

III. EXISTING APPROACHES TO SOLVE THE ν BOLTZMANN EQUATION

In general, to study the thermalization of neutrinos, one has to solve the quantum kinetic equations (QKEs) for the neutrino density matrix [7, 40, 51]. However, for our purposes, it may be reasonable to approximate the oscillations by the temperature-dependent oscillation probabilities, $\langle P_{\alpha\beta} \rangle(E_\nu, T)$, similarly to how this is done in [28, 32]. Then, it may be possible to reduce the complexity by converting the QKEs into the Boltzmann equations for the neutrino distribution function f_{ν_α} in the momentum space:

$$\begin{aligned} \frac{\partial f_{\nu_\alpha}(E_\nu, t)}{\partial t} - pH \frac{\partial f_{\nu_\alpha}(E_\nu, t)}{\partial p} &= \\ &= \sum_\beta \langle P_{\beta\alpha} \rangle \mathcal{I}_{\text{coll}, \nu_\beta} [f_{\nu_\alpha}, f_{\nu_\beta}, T], \end{aligned} \quad (14)$$

supplemented with the Friedmann equation describing the expansion of the Universe (and in particular H), the equation for the evolution of the EM plasma temperature T , and the equation governing the dynamics of LLPs in case they are present. Here, $p = E_{\nu_\alpha}$ is the neutrino physical momentum. $\mathcal{I}_{\text{coll}, \nu_\beta}$ is the collision integral for the neutrino of the flavor β , which in general contains a source term from new physics particles, a neutrino-neutrino interaction term, and a neutrino-EM interaction term. It has the form [39]

$$\begin{aligned} \mathcal{I}_{\text{coll}, \nu_\alpha} &= \frac{1}{2E_{\nu_\alpha}} \sum \int \prod_{i=2} \frac{d^3 p_i}{(2\pi)^3 2E_i} \prod_{f=1} \frac{d^3 p_f}{(2\pi)^3 2E_f} \\ &\times |\mathcal{M}|^2 F[f] (2\pi)^4 \delta^{(4)} \left(\sum_{i=1} p_i - \sum_{f=1} p_f \right). \end{aligned} \quad (15)$$

The first summation encompasses all potential interaction processes involving ν_α , with $i = 1$ representing the

⁵ Indeed, as its value at eV temperatures is fixed, the value at MeV temperatures must be larger than the Λ CDM value to compensate for the dilution.

neutrino itself. The integral extends over all possible states of ν_α characterized by momentum p_1 . Here, i and j denote the initial and final states of a given process, respectively. The term $|\mathcal{M}|^2$ represents the squared matrix element of the process (see Table 3 in ref. [32] for the explicit expressions of $|\mathcal{M}|^2$ relevant to neutrino processes at MeV-scale temperatures). The factor $F[f]$ accounts for the statistical distribution within the medium and is given by

$$F[f] = \prod_{i=1} (1 \mp f_i) \prod_{f=1} f_f - \prod_{i=1} f_i \prod_{f=1} (1 \mp f_f), \quad (16)$$

where $f_{i,f}$ denote the momentum distributions for the i -th and f -th particles. Finally, the factor $(1 - f)$ corresponds to Pauli blocking for fermions, whereas $(1 + f)$ corresponds to Bose enhancement for bosons. Finally, the δ function ensures the conservation of the 4-momentum in the process.

Depending on the scenario studied, there are two different state-of-the-art approaches to solving the Boltzmann equation (14). If the neutrinos injected by decays of new physics are close to thermal, $E_\nu \simeq 3.15T$, or if decays are solely electromagnetic, then it may be possible to approximate the neutrino distribution by

$$f_{\nu_\alpha}(E_\nu, t) \approx f_{\text{FD}}(E_\nu, T_{\nu_\alpha}(t)) = \frac{1}{\exp\left[\frac{E_\nu}{T_{\nu_\alpha}(t)}\right] + 1} \quad (17)$$

and consider an integrated version of the Boltzmann equations on the three neutrino temperatures $T_{\nu_\alpha}(t)$ [24, 52]. In the limit of negligible electron mass, it may be possible to represent the energy transfer rates in the system as an analytic expression. Another example includes obtaining a correction to neutrino high-energy tail caused by non-instant decoupling in the standard scenarios.⁶

In practice, once we add new physics, the assumption of the perfect thermality of the neutrino distribution is typically violated. The first reason is the energy dependence of the equilibration of neutrinos, as discussed in Sec. II. Neutrinos with different energies interact at very different rates, which leads to neutrino spectral distortions even if we simply heat the EM plasma.

To study the distortions, one needs to solve the Boltzmann equation (14) in the full generality. In the literature, this is done using the approach that we will call the discretization method. The algorithm is to analytically reduce the dimensionality of the integration in $\mathcal{I}_{\text{coll}}$ to some integer k and then discretize the comoving momentum space $y = p \cdot a(t)$.⁷ The integrodifferential Boltz-

mann equation is then converted into a system of ordinary differential equations (see [7, 38–40] and references therein).

Given the number of points N_y in the momentum grid, the complexity of the code (and hence its running time) scales as $\mathcal{O}(N_y^{k+1})$. One power of N_y comes from the number of equations on the momentum model. The power N_y^k is from the number of terms in the k -dimensional collision integral. The value of k is bounded from below by 2, which is just the scenario when we include weak $2 \rightarrow 2$ reactions with neutrinos and e^\pm [38]. Next, if we assume the linear grid, then $N_y \sim E_{\nu, \text{max}}$, where $E_{\nu, \text{max}}$ is the maximal neutrino physical energy relevant for the evolution of the system.⁸ Therefore, the complexity grows as

$$t_{\text{computation}} = \mathcal{O}(E_{\nu, \text{max}}^{k+1}) \quad (18)$$

While the approach has been successfully used in the standard cosmological scenario (see, e.g., [7, 10, 39, 40]), it has limitations when applying it to new physics scenarios. First, it requires simple analytic matrix elements in the neutrino source terms. In practice, this is not the case when we have hadronically decaying LLPs with mass $m \gg \Lambda_{\text{QCD}}$. This is because quarks and hadrons appearing in the decays undergo subsequent showering and hadronization, which results in a complicated phase space structure, which is hard to fit in the form of an analytic matrix element.

Second, even if simple analytic matrix elements do exist, the computational complexity quickly increases if we depart significantly from the standard cosmological case. For example, simply increasing the integration dimensionality from $k = 2$ to higher values may enormously increase the time of calculations. This is the case of, e.g., $2 \rightarrow 3$ scatterings with neutrinos such as the famous $e^+e^- \rightarrow \nu\bar{\nu}\gamma$. Another example is when there are n -body decays with $n > 3$, which are quite often for LLPs [53].

The computational time problem exists even in the most optimistic case $k = 2$. Consider injections of high-energy neutrinos with energy $E_\nu \gg T$. It may appear in decays of LLPs. Considering the energies $E_\nu \sim 1$ GeV would enlarge the computational time of the calculations compared to the standard cosmological case (where we assume $E_{\nu, \text{max}} = 20$ MeV) by a factor $\sim 50^3 \sim 10^5$ (remind Eq. (18)), making any applications impossible in practice. Finally, depending on the energy

⁶ It may be converted to the momentum-dependent correction to the neutrino temperature $T_\nu(p)$ that approaches T at $p \gg 3.15 \cdot T$ and vanishes at small momenta.

⁷ In the Standard cosmological scenario case, the comoving grid is convenient since it “freezes” the neutrino distribution: e.g., the peak of the energy distribution corresponds to the same y at different times.

⁸ In principle, one may consider a different grid structure, e.g., logarithmic. However, it may cause problems with the energy conservation and stability of the solution throughout the evolution. It is also heavily model-dependent: injected neutrinos typically have a complicated spectrum, which requires adapting the grid in a non-trivial way to achieve good accuracy. In fact, even two-body decays into neutrinos would correspond to different comoving momenta $y = pa$ (depending on the time when they are injected), as the scale factor would be different. By varying, e.g., the lifetime τ of the decaying X , one would change the domain of y populated by the neutrino injections.

density of X , it may sizeably contribute to the Universe’s energy density. For the same temperature range, the scale factor would be larger than in the Standard model case, which does not allow fixing the maximal comoving momentum in the grid y_{\max} .

To summarize, there is no adequate approach to study the dynamics of primordial neutrinos in the presence of new physics while maintaining model independence, efficiency, and transparency.

IV. BASICS OF DSMC

Consider the Liouville equation for the N -particle probability distribution density $F_N(\mathcal{R}, \mathcal{V}, t)$ with a short-range potential $\Phi_{i,j}$ of binary interactions:

$$\frac{\partial F_N}{\partial t} + \sum_{i=1}^N \mathbf{v}_i \frac{\partial F_N}{\partial \mathbf{r}_i} + \sum_{1 \leq i < j \leq N} \Phi_{i,j} F_N = 0 \quad (19)$$

The DSMC approach approximately solves it using the following scheme (see [44, 54, 55] and references therein):

1. Apply the $N-1$ space variable reduction $F_N \rightarrow \tilde{F}_N = \int F_N \prod_{s=2}^N d\mathbf{r}_s$.
2. Switch to the iteration scheme by considering the equation on the time intervals $(t; t + \Delta t)$.
2. Decompose the space domain \mathcal{D} onto disconnected sub-domains $\mathcal{D} = \cup_{l=1}^M \mathcal{D}^{(l)}$ (“cells”), populated by fixed amounts of particles during Δt .
3. Split the evolution into three successive procedures within each time step: ballistic motion (free-streaming in the absence of collisions), binary collisions within each $\mathcal{D}^{(l)}$, and then interchanging particles between cells as a result of the first two steps.

Under an assumption that the system obeys ergodic conditions, the DSMC approach may be converted to an analog of the Bogoliubov–Born–Green–Kirkwood–Yvon hierarchy for $3 + 3N$ phase space, which reduces to the Boltzmann equation in the limit $N \rightarrow \infty$ and assuming the molecular chaos (i.e., that the velocities of colliding particles are statistically independent).

A. No-Time-Counter scheme

The central part of the DSMC approach is to simulate the evolution of particles within an individual cell. There are various methods [44, 56–60]. Some of which have $\mathcal{O}(N_{\text{cell}})$ computational complexity, where N_{cell} is the number of particles per cell. Examples of the latter are No-Time-Counter (NTC), Majorant collision frequency, Simplified Bernoulli Trial, and others [60]. Here, we will discuss the NTC method, proposed in [59], which we will adapt for our purposes.

First, one defines the timestep of the simulation Δt . It must be sufficiently small to resolve the characteristic interaction time in the system. It may be calculated as

$$\Delta t = \left(\frac{(\chi_{\text{particle}} \cdot \sigma v)_{\max} \cdot N}{V_{\text{system}}} \right)^{-1} \quad (20)$$

Here, χ_{particle} is the particles’ weight – how many physical particles are represented by each particle; N is the total number of interacting particles; V_{system} is the system’s volume; σ is the interaction cross-section; v is the relative velocity. The subscript “max” denotes finding the maximal value among the system.

Next, consider splitting the system’s volume into cells. Let us assume that there are n_{cells} cells, each having the volume $V_{\text{cell}} = V_{\text{system}}/n_{\text{cells}}$. In the standard DSMC application cases, particular cells contain N_{cell} particles as low as $\mathcal{O}(10-20)$ and even lower, which is enough for simulating the evolution properly. Within a particular cell, one samples randomly

$$N_{\text{sampled}} = \frac{N_{\text{cell}}(N_{\text{cell}} - 1)}{2} \frac{\omega_{\text{cell,max}} \Delta t}{V_{\text{cell}}} \quad (21)$$

pairs of particles to interact. Here, $\omega_{\text{cell,max}} = (\chi_{\text{particle}} \sigma v)_{\text{cell,max}}$ is the estimate of the maximum interaction cross-section within the cell.

For each sampled pair, one accepts its interaction with the probability

$$P_{\text{acc}} = \frac{\omega}{\omega_{\text{cell,max}}}, \quad \omega = (\chi_{\text{particle}} \sigma v)_{\text{pair}} \quad (22)$$

If the interaction is accepted, one simulates the possible final states for the given pair and its scattering kinematics.

The complexity of the NTC scheme grows as $\mathcal{O}(N_{\text{cell}})$ [46]. This is achieved by the fact that $\omega_{\text{cell,max}} \Delta t / V_{\text{cell}}$ in the number of sampled events is typically $\ll 1$. The systems with the total number of particles $N \gg 10^6$ may be simulated within minutes, even on ordinary laptops. Such large values are already enough to reach the precision required in our studies.

The NTC method has been tested for various systems, including relativistic ones [61–65], which demonstrates its flexibility and coverage of the wide range of scenarios.

V. DSMC FOR NEUTRINOS

Let us now discuss how to apply the DSMC approach to study the evolution of primordial neutrinos.

As in the case of the state-of-the-art methods, we will first utilize the simplification coming from the properties of the Early Universe at the times of interest – its homogeneity and isotropy. Because of this, we may drop the spatial degrees of freedom and view the system as effectively zero-dimensional, with all interactions occurring at one point. Splitting the system into cells is then a formal

step to maintain the performance. We will also neglect any cells' boundary interactions.

To accurately trace the thermalization of neutrinos, we represent their population by a set of individual particles characterized by the 4-momentum, flavor, and particle-antiparticle type. Every interaction involving neutrino (remind Eq. (13)) would modify its properties. Namely, it may change its 4-momentum (if the interaction is elastic) and/or flavor (if it is the annihilation of the type $\nu_\alpha \bar{\nu}_\alpha \rightarrow \nu_\beta \bar{\nu}_\beta$). Finally, there are annihilation processes $\nu_\alpha \bar{\nu}_\alpha \leftrightarrow e^+ e^-$, which may lead to a change in the number of neutrinos.

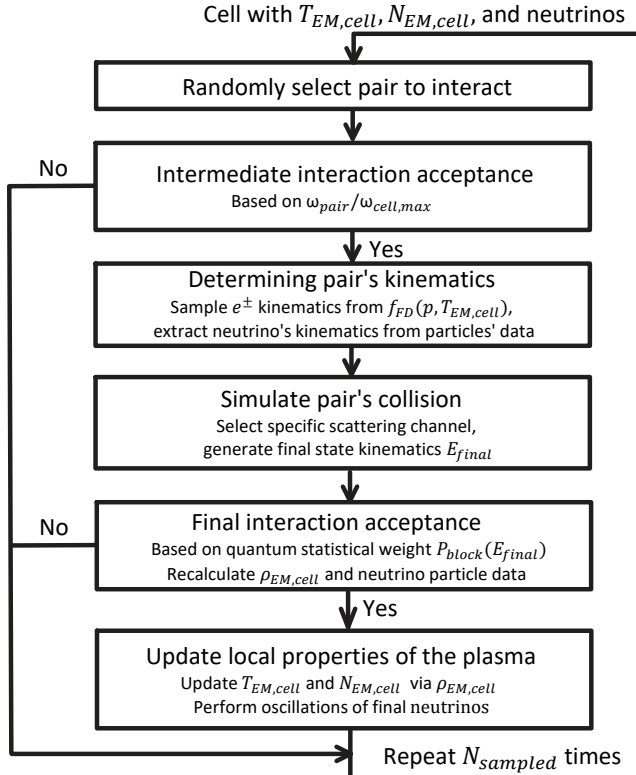


FIG. 1. The modification of the No-Time-Counter scheme, used to simulate the interactions within the system's cells within the Direct Simulation Monte Carlo approach, for describing interactions in the MeV primordial plasma. First, we sample N_{sampled} pairs to interact, Eq. (21). For each pair, we compute its interaction weight and make an intermediate decision on whether it will interact using the criterion (22). Then, we sample the kinematics of the interacting particles, generate the final states resulting from the collision, and make the final decision of whether the interaction takes place from the Pauli principle (25). Finally, we update the local properties of the plasma: the EM plasma temperature and the number of EM particles, as well as neutrino flavor distributions by the oscillation probabilities, Eq. (26).

Proceeding with the DSMC method in its current form is impossible, as it does not incorporate fundamental features of the Early Universe's plasma. These include the expansion of the Universe, the hierarchy between the

equilibration rates in the neutrino and EM sectors, the Pauli principle, neutrino oscillations, and the presence of decaying particles. Below, we discuss these features and how we address them in detail (see also Fig. 1, showing the modification of the NTC scheme).

1. *Expansion of the Universe.* From the DSMC's point of view, it simply represents an external force acting on the particles of the system, with an additional modification that it also increases the system's (and cells') volume. These two effects may be simply accounted for by redshifting the total volume of the system V_{system} (and hence the cell's volume) as well as the individual energies E_i of the particles $\{i\}$, applied at each step of the simulation. Namely, at the beginning of the timestep Δt , we calculate the Hubble factor H using Eq. (3), and then make use of the relation

$$V_{\text{system}} \rightarrow V_{\text{system}}(1 + 3H\Delta t), \quad E_i \rightarrow \frac{E_i}{1 + H\Delta t} \quad (23)$$

provided that $H\Delta t \ll 1$. To account for this requirement, we modify the definition (20):

$$\Delta t = \min \left[0.01H^{-1}, \left(\frac{(X_{\text{particle}} \cdot \sigma v)_{\text{max}} \cdot N}{V_{\text{system}}} \right)^{-1} \right] \quad (24)$$

Here, 0.01 is the arbitrary small factor.

2. *Properties of the EM plasma.* As we discussed previously in Sec. II, the reactions involving solely EM particles are orders of magnitude faster than those where neutrinos participate. As we have chosen the timestep Δt comparable to the neutrino interaction rates, the EM particles may be viewed as a part of perfectly thermal plasma characterized by one parameter – temperature T . However, we then need to implement the response of any single interaction involving the EM particles on T .

To reach this, at the beginning of the iteration, we characterize the EM plasma with the energy density ρ_{EM} , both globally (for the whole system) and locally (at the level of the individual cell). The global and cells' temperature T is related to ρ_{EM} by Eq. (6).

During the NTC routine, the local number of electrons and positrons $N_{e^\pm, \text{cell}}$ per cell is calculated from the relation between T_{EM} and the number density. The kinematics of any e^\pm selected within the NTC algorithm is sampled from the Fermi-Dirac distribution $f_{\text{FD}}(p, T_{\text{EM}})$. The change $\rho_{\text{EM, cell}}$ resulting from the accepted interaction leads to the update in $T_{\text{EM, cell}}$ and $N_{e^\pm, \text{cell}}$.

At the global level, once the simulations for all cells are performed, $\rho_{\text{EM, cell}}$ s are merged into the total energy density $\rho_{\text{EM, system}}$, which allows obtaining the global temperature of the EM plasma.

3. *Quantum statistics.* It enters the binary part of the collision integral (14) with fermionic final states F_1, F_2 having energies $E_{F_{1,2}}$ via multiplicative Pauli blocking factors

$$P_{\text{block}} = (1 - f_{F_1}(E_{F_1})) \times (1 - f_{F_2}(E_{F_2})), \quad (25)$$

where f is the energy distribution of the given final state. Thus, it suppresses interactions where the final states would occupy the high-populated part of the energy distribution (e.g., $E \lesssim T$ for the equilibrium shape distribution with the temperature T). To implement this, one should consider the local energy distributions for both EM particles and neutrinos and calculate P_{block} . A possible simplification is, when calculating the blocking factor, to describe neutrino's distribution by the Fermi-Dirac function $f_{\text{FD}}(T_{\nu_\alpha})$, where T_{ν_α} is the local effective neutrino temperature obtained in a way similar as we do for the EM plasma.⁹

4. *Neutrino oscillations.* We incorporate them at the end of the iteration timestep by changing each of the neutrino flavors according to the formula

$$\nu_\alpha(E_\nu) \rightarrow \sum_\beta \langle P_{\alpha\beta} \rangle(E_\nu, T) \nu_\beta(E_\nu), \quad (26)$$

where $\langle P_{\alpha\beta} \rangle(E_\nu, T)$ are averaged neutrino oscillation probabilities (remind also Eq. (2)).

5. *Presence of LLPs X and new interactions.* Let us start with discussing LLPs. Decaying either into the EM plasma particles or neutrinos, they would heat the EM plasma temperature or distort the properties of the neutrino sea. Having the initial condition for the X 's abundance, on the language of DSMC, we introduce the equivalent amount n of X particles to the system. Further, we assume that LLPs are non-relativistic and decoupled, which ideally matches the scope of this study. Per each timestep Δt , provided that it is much smaller than the LLP's lifetime τ , their population is evolved by decaying the fraction

$$\Delta n = n(t + \Delta t) - n(t) \approx \frac{\Delta t}{\tau} n(t) \quad (27)$$

For each decay, it is possible to obtain the energies of resulting neutrinos and calculate the amount of the EM energy using Monte-Carlo simulations – the baseline approach for particle physics. This is a natural choice if one wants to maintain the model independence, as it is maximally general and may describe

any process. In particular, exclusive decays (where we have well-defined “fixed” final states, e.g., $X \rightarrow 3\pi$) may be simulated on-flight by sampling the phase space of decay products using the analytic matrix element of the process. The phase space of hadronic decays in the LLP mass range $m \gg \Lambda_{\text{QCD}}$ (such as $X \rightarrow q\bar{q}\nu$, where q is a quark) may be obtained by simulating them in PYTHIA8 [66] for a grid of masses and subsequently using the output particle's data in the form of events inside the DSMC code.

The Monte Carlo sampler must incorporate the interactions of the decay products with the primordial plasma, which may substantially redistribute their energy between the neutrino and EM sectors compared to the vacuum case. Namely, all electrically charged particles with lifetimes $\tau \gtrsim 10^{-10}$ s, such as muons, charged pions, and kaons, appearing in the MeV plasma may undergo kinetic energy loss via EM interactions, annihilation, interactions with nucleons before decaying [50]. This evolution may again be implemented probabilistically, in the spirit of Monte Carlo simulations.

Absolutely similarly, it is possible to sample the energies for non-standard scattering processes, e.g., for the $2 \rightarrow 3$ scatterings $e^+e^- \rightarrow \nu_\alpha \bar{\nu}_\alpha \gamma$.

In order to finish the discussion of the approach, let us address the question of the number of particles per cell, N_{cell} , entering Eq. (21). In our system, it is

$$N_{\text{cell}} = N_{e^\pm} + 2 \sum_\alpha N_{\nu_\alpha} \quad (28)$$

Since statistical quantities, such as temperatures, are involved in simulating the interactions, it is not possible to use small $N_{\text{cell}} \sim 10$, as it is typically done in the DSMC simulations. Instead, the values as large as $N_{\text{cell}} = \mathcal{O}(100)$ should be considered. As a bonus, such a large number also allows for avoiding various stochastic problems of the NTC method, including repeated interaction of the same pair [43].

VI. CURRENT IMPLEMENTATION

We have implemented a simplified version of the DSMC method described above, which serves as proof-of-principle.¹⁰

The first approximation is that we have neglected the electron mass when describing the population of the EM particles; this is done in order to simplify the sampling of electron and positron particles. The second simplification concerns the absence of implemented continuous decays of LLPs – only instant injections of decay products are included.

⁹ The actual neutrino distribution is, of course, non-thermal, and we use this approximation only when calculating P_{block} . Since the deviations from the thermality we study are not very large without loss of generality, we believe that the approximation is accurate.

¹⁰ The code may be provided upon request.

None of these limitations are fundamental; they will be added with the development of the approach. Namely, by neglecting the electron mass, we restrict the temperature range covered in the study by $T \gtrsim m_e$, but this is already enough to study interesting scenarios, as temperatures of interest are those when neutrinos start decoupling, which occurs at $T \lesssim 5$ MeV. As for the instant decays, although being an artificial scenario, they already provide insights about both the performance of the DSMC approach and physics case studies, such as the impact on N_{eff} .

The implementation is written in `Mathematica`. It allows combining moderate performance¹¹ with symbolic calculations, which are needed when dealing with describing kinematics and deriving the matrix elements of various processes. Also, it makes it possible to use existing realizations of Monte-Carlo sampling of decays of LLPs, such as `SensCalc` [67].

The typical number of particles per neutrino flavor we consider in the setup is $(1 - 5) \cdot 10^5$, which results in the total number of particles at the level of a $N = \text{few} \times 10^6$. The standard number of particles per cell we have chosen is 400, meaning that the cell number is $\mathcal{O}(10^3)$. The running time required to produce most of the plots below was < 30 minutes; it varied only mildly depending on the setup, including the energies of the neutrinos included in the system. In particular, in order to produce the neutrino distributions shown in Fig. 5, we spent only $\simeq 5$ minutes. We expect significant improvement, possibly by an order of magnitude, in the running time after optimizing the code and/or rewriting some of its modules in native C++. Finally, with the implementation, we maintain the approximate linear scaling of the computational time with N , as expected from the basics of the NTC approach.

To validate the developed neutrino DSMC, we have studied its predictions in the case of well-established scenarios, including:

1. *Approaching thermal equilibrium.* In the absence of Universe expansion, independently of the initial conditions, neutrinos have to reach thermal equilibrium with the EM particles. In particular, their differential distribution in the number and energy densities, which we will plot throughout the paper, must be

$$\frac{dn_\nu}{dE_\nu} = \frac{g_\nu}{2\pi^2} f_{\text{FD}}(E_\nu, T_\nu) \times E_\nu^2, \quad (29)$$

$$\frac{d\rho_\nu}{dE_\nu} = \frac{g_\nu}{2\pi^2} f_{\text{FD}}(E_\nu, T_\nu) \times E_\nu^3, \quad (30)$$

where $T_\nu = T$ is the neutrino temperature, two powers of E_ν come from the phase space, and one

in Eq. (30) from the definition of ρ_ν . Finally, f_{FD} is the Fermi-Dirac distribution (remind Eq. (5)), with g_ν being the lepton charge degree of freedom (remind Eq. (9)).

Eq. (30) automatically implies that in equilibrium, the ratio of the energy densities of the neutrino and EM plasmas is

$$\left(\frac{\rho_\nu}{\rho_{\text{EM}}}\right)_{\text{eq}} = \frac{7/8 \cdot g_\nu}{7/8 \cdot g_e + g_\gamma} = \frac{21}{22}, \quad (31)$$

where we have used Eq. (6) and assumed $T \gg m_e$.

2. *Energy transition rates.* Consider the initial setup where the distribution function of neutrinos is fixed by f_{FD} , parametrized with the temperature $T_{\nu_\alpha} \neq T$. During the equilibration and in the absence of expansion, the energy transition rates between the neutrino and EM sectors must match the well-known analytic result from [24] (where we turn off the expansion as well).
3. *Expansion and decoupling.* If including the expansion of the Universe in the previous setup, we should consistently recover the decoupling of neutrinos, which prevents their population from full thermalization, as well as reproduce the results of [24].

Details may be found in Appendix A. In addition, we have performed tests that are not present in the paper. Those include the evolution of neutrinos and antineutrinos (the evolution must preserve the lepton symmetry up to Monte Carlo fluctuations) and independence on the exact simulation setup (*e.g.* Number of simulation cells, the total number of particles, etc.). We believe that it proves that our approach fulfills the requirements to be accepted as a valid method for treating the evolution of neutrinos.

VII. CASE STUDIES

To demonstrate the potential of various implications of the DSMC method, we will consider several toy case studies specified by the initial conditions on the neutrino distribution functions. These setups have two applications. On the one hand, they mimic distinct scenarios with new physics and thus provide useful insights into the dynamics of the primordial plasma. On the other hand, they will comprehensively demonstrate the performance and flexibility of the neutrino DSMC approach.

Firstly, we investigate the evolution of a system where neutrinos initially possess an equilibrium energy distribution with a temperature $T_{\nu_\alpha} \neq T$ (see Sec. VII A). This setup encompasses two distinct scenarios. The first scenario arises when energy is injected exclusively into the electromagnetic (EM) sector, resulting in $T > T_{\nu_\alpha}$.

¹¹ Low-level routines, such as simulations of interactions and manipulations with cells, are compiled in C++.

The second scenario occurs when nearly thermal neutrinos are introduced into the neutrino sector, as explored in [24]. These cases can be analyzed using the integrated Boltzmann equation developed in [24, 52]. Nevertheless, we will demonstrate that even within these simplified setups, deviations from the thermal shape of the neutrino distribution emerge, leading to discrepancies between the solutions of the unintegrated and integrated approaches to the neutrino Boltzmann equation, particularly in the determination of N_{eff} .

Second, we will consider injections of high-energy monochromatic neutrinos (Sec. VII B). This scenario represents the case of two-body decays of heavy LLPs, such as neutrinophilic scalars [68], majorons [69], $B-L_\alpha$ mediators [70], and relics in late reheating scenarios [27]. We will consider high injection temperatures, $T \gtrsim 1$ MeV. We will show that in the case of sufficiently large neutrino energy, such that $E_\nu \gg T$, these injections would result in a decrease in the neutrino-to-EM energy densities ratio compared to the standard cosmological scenarios. This setup will also serve to demonstrate that the performance of the DSMC does not depend on the neutrino energy (supporting the initial expectations) and to cross-check it by comparing the neutrino evolution with the predictions of the discretization codes.

Finally, we will study injections of neutrinos from decays of different long-lived SM particles, such as muons, charged pions, and kaons (Sec. VII C). This case corresponds to a common scenario of LLPs with complex decay chains, which may not decay into neutrinos directly but instead decay into such heavy states. Examples are, e.g., a decay of the Higgs-like scalars into $\pi^+\pi^-/K^+K^-$, the dark photon decay into $2\pi/3\pi/4\pi$, and decays of HNLs into $\pi\mu$ [71]. Another illustrative case is the decay into quarks, where we have a high multiplicity of meson states. We will show that independently of the decaying particle (or the fraction of their energy placed to the neutrino plasma right after decay), the ratio (32) decreases below the equilibrium value. This case also study demonstrates the flexibility of our approach, which may handle any decay chain with complicated kinematics.

To make the illustrative analysis for this and other studies performed in this paper, we introduce the quantity

$$\delta\rho_\nu = \left(\frac{\rho_\nu}{\rho_{\text{EM}}}\right)_{\text{eq}}^{-1} \frac{\rho_\nu}{\rho_{\text{EM}}} - 1 \quad (32)$$

A. From equilibrium spectral shapes to distortions

Let us consider a system with neutrinos having an equilibrium shape of energy distributions, but the temperatures of these distributions differ from the EM plasma temperature.

We will study how the equilibration of this initial condition evolves in time, to identify the possible deviations from the description dynamics of the equilibration fol-

lowing Ref. [24], where we turn off the electron mass in order to compare apples with apples. These deviations genuinely appear from the non-thermal distortions in the neutrino sector (invisible within the method of [24]). It is because the interaction rates of different parts of the neutrino spectrum are energy-dependent (remind Sec. II).

We will consider the particular initial condition where neutrinos have the same temperature $T_{\nu_\alpha} = 3.2$ MeV, and the EM plasma has a lower temperature $T = 3$ MeV.

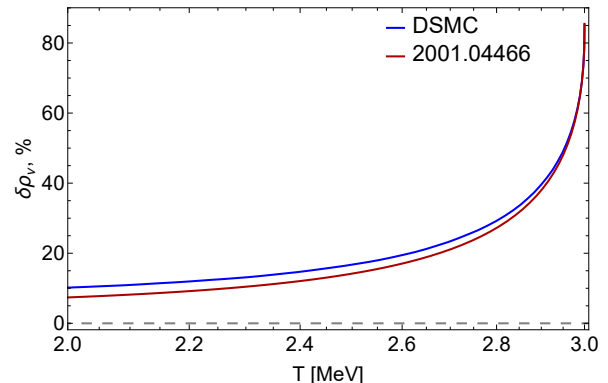


FIG. 2. The evolution of the neutrino-to-EM energy densities ratio $\delta\rho_\nu$, given by Eq. (32), under the scenario where the neutrino distribution shape is thermal (Eq. (17)), but has temperature T_ν different from the EM plasma T_{EM} . For the initial setup, we consider $T_\nu = 3.2$ MeV and $T_{\text{EM}} = 3$ MeV. The blue line shows the result of our DSMC approach, whereas the red line is obtained using the method of integrated neutrino Boltzmann equations from [24], which assumes that the shape of the neutrino distribution is perfectly thermal throughout the whole evolution.

The resulting evolution of $\delta\rho_\nu$, as predicted by the DSMC approach and the method from [24], is shown in Fig. 2. From the figure, we see that the two descriptions match at the initial stages, while the deviations appear once the system develops. They get frozen throughout the evolution because of the decoupling of neutrinos. The same conclusion holds in the opposite case of the initial condition $T > T_\nu$.

Therefore, we conclude that the integrated Boltzmann approach may provide insufficient accuracy even in cases where there are no direct distortions of the neutrino spectrum (see further discussion of this point in Ref. [50]).

B. Instant neutrino injection

Let us now proceed to a more generic case in which there are injections of non-thermal neutrinos with $E_\nu \gg T$. For this setup, the integrated Boltzmann approach is completely inapplicable, as high-energy neutrinos have a much larger rate of interactions than their thermal counterparts, and severely influence the dynamics of the thermalization even if their amount is low.

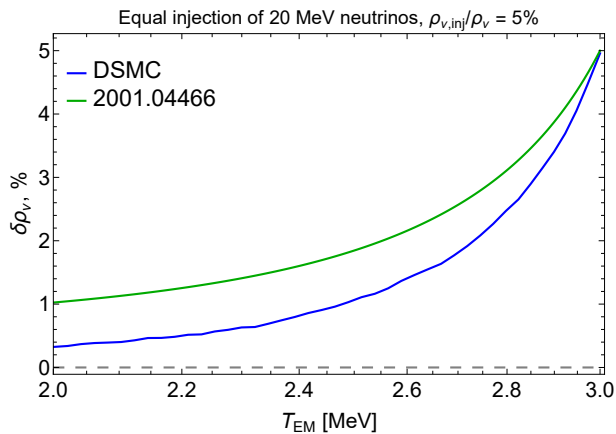


FIG. 3. The behavior of the ratio (32) under the injection of 20 MeV neutrinos equally to all neutrino flavors at the temperature $T = 3$ MeV. The total injected energy density is $\rho_{\nu,\text{inj}}/\rho_{\nu,\text{total}} = 5\%$. The blue line shows the prediction of the DSMC method, whereas the green one corresponds to the integrated Boltzmann approach from [24].

Let us assume, without loss of generality, the injection of monochromatic neutrinos with energy $E_{\nu,\text{inj}}$ at temperature $T = 3$ MeV, and consider three different values $E_{\nu,\text{inj}} = 20, 70, 500$ MeV. We will analyze both the evolution of $\delta\rho_\nu$ and the neutrino spectrum shape. The option $E_{\nu,\text{inj}} = 20$ MeV primarily serves to demonstrate the necessity of using the unintegrated Boltzmann approach in case of non-thermal distortions. The second setup is central – it will show the qualitative impact of large neutrino energies on N_{eff} . We will use it to compare with the discretization codes from [7, 33, 34], which predict contradictory behavior of the sign of $N_{\text{eff}} - N_{\text{eff}}^{\Lambda\text{CDM}}$ in presence of high-energy neutrinos. Finally, the last choice highlights the performance of our setup – the running time and precision are almost independent of the neutrino energy.

1. Injection of 20 MeV neutrinos

Consider the injection of 20 MeV neutrinos. Without loss of generality, we assume equal injection among the three neutrino flavors, with the total injected energy density $\rho_{\nu,\text{inj}}/\rho_{\nu,\text{total}} = 5\%$. Here and below, we include the Hubble expansion of the Universe, but turn off the neutrino oscillations.

The evolution of the resulting $\delta\rho_\nu$ is shown in Fig. 3, where we, as usual, also include the prediction of the integrated Boltzmann approach. Both approaches predict a monotonic decrease of $\delta\rho_\nu$. In particular, at late temperatures, when the expansion prevents equilibrating, we end up with the value of $\delta\rho_\nu$ close to 0. However, the rate of decrease of $\delta\rho_\nu$ predicted the neutrino DSMC is much faster. This is explained by the fact that, compared to thermal particles, the injected high-energy neutrinos have a larger probability of interacting with the EM sec-

tor and, hence, transporting their energy.

2. Injection of 70 MeV neutrinos

Let us now proceed with the 70 MeV injection. We will consider several setups here. The first one is with equal injection among the neutrino flavors and a large $\rho_{\nu,\text{inj}}/\rho_{\nu,\text{total}} = 30\%$. It serves as a very illustrative demonstration of the qualitative features of the evolution of $\delta\rho_\nu$. The two others are with the smaller injected energy $\rho_{\nu,\text{inj}}/\rho_{\nu,\text{total}} = 5\%$ and two different injection patterns: equal energy distribution among the flavors, and the injection solely into the sector of electron neutrinos. We will use them to compare with the predictions of different discretization codes from the literature.

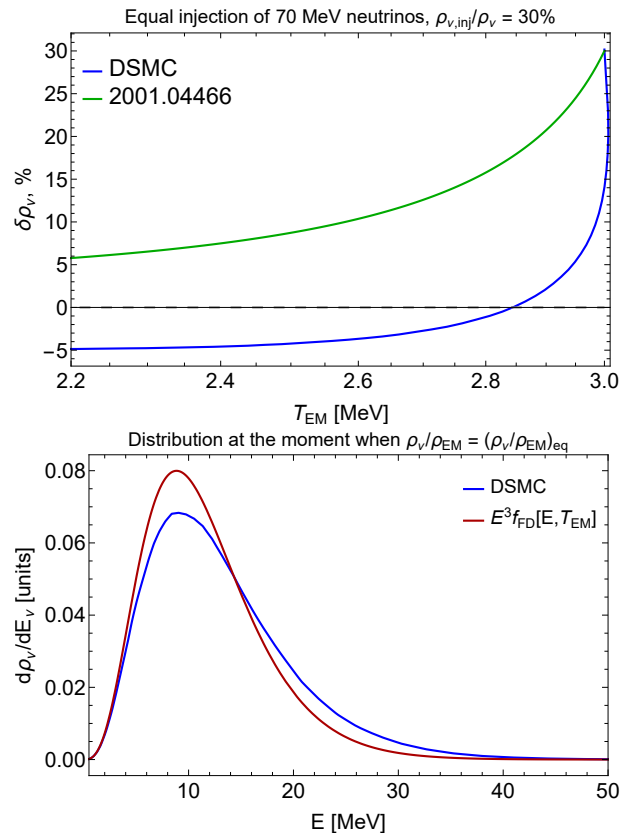


FIG. 4. The evolution of the primordial plasma after the injection of neutrinos with energies $E_\nu = 70$ MeV with the amount $\rho_{\nu,\text{inj}}/\rho_{\nu,\text{total}} = 30\%$. The other parameters of the setup are similar to the one considered in Fig. 3. *Top panel*: the behavior of $\delta\rho_\nu$ with temperature. The DSMC curve twice intersects the value $T = 3$ MeV, which is because of the reheating of the EM plasma by neutrinos immediately after the injection. *The bottom panel*: comparison of the shape of the neutrino energy distribution at the moment when $\delta\rho_\nu = 0$ during the equilibration, as obtained with the DSMC simulation (the blue curve) and assuming the equilibrium neutrino spectrum (the green curve).

Fig. 4, upper panel, shows the evolution of $\delta\rho_\nu$ for the

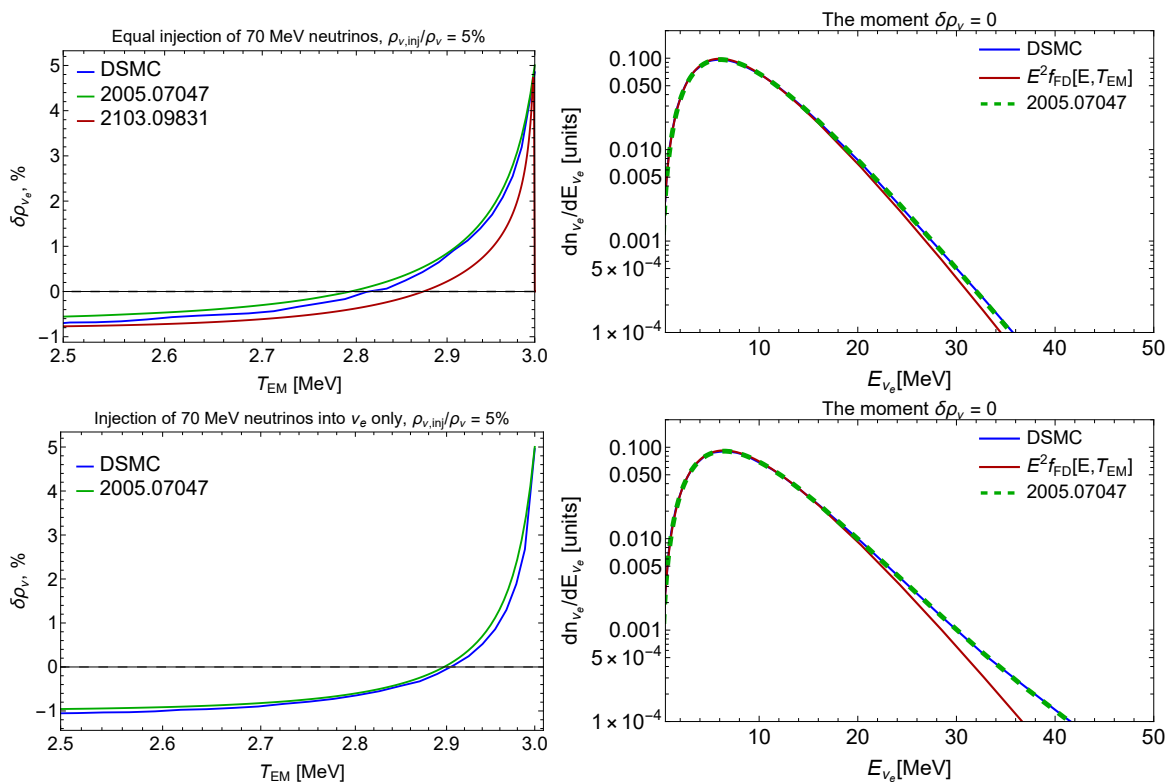


FIG. 5. Comparison of the DSMC approach with the discretization codes for the setup of injection of 70 MeV neutrinos at $T = 3$ MeV. Two configurations are considered: equal injection among the flavors (the top panels) and the injection solely into ν_e (the bottom panels). In both cases, the injected energy fraction is $\rho_{\nu,\text{inj}}/\rho_{\nu,\text{total}} = 5\%$. The left plots show the evolution of $\delta\rho_\nu$, given by Eq. (32), whereas the right plots are snapshots of the electron neutrino distribution spectrum at the temperature when $\delta\rho_\nu = 0$. The blue lines are the DSMC predictions, the green lines denote the calculation by the discretization approach from [7] (see also [50]), whereas the red line is the result obtained in [33].

30% injection setup. Now, there is a qualitative difference in its behavior between the integrated and DSMC approaches. The former results in the naively expected monotonic decrease of $\delta\rho_\nu$, whereas according to the latter, it first rapidly drops below zero, where it then freezes out. Without the expansion of the Universe, it would have been a decrease of $\delta\rho_\nu$ to negative values, and then a slow monotonic reaching $\delta\rho_\nu \rightarrow 0$ from below.

To understand this counter-intuitive result, let us remind Sec. II and highlight two important properties of the plasma: (i) EM particles instantly equilibrate between themselves, and (ii) weak interaction rates grow with the invariant mass of colliding particles. Because of this, the injected non-thermal neutrinos quickly “knock out” thermal neutrinos by the interactions

$$\nu_{\text{inj}}\nu_{\text{thermal}} \rightarrow e^+e^-, \quad \nu_{\text{inj}}\nu_{\text{thermal}} \rightarrow \nu^+\bar{\nu}^- \quad (33)$$

The first process pumps the injected energy and a fraction of the energy of the thermal population to the EM sector. The rate of these processes is much higher than the rate of the same processes when only thermal particles are involved. Knocking out thermal neutrinos determines the shape of the neutrino spectrum during these interactions: compared to the equilibrium spectrum f_{FD} ,

it is underabundant in small energies and overabundant in large energies.

The snapshot of the neutrino spectrum at the moment when $\delta\rho_\nu = 0$ is shown in the lower panel of Fig. 4. Then, we have equilibrium amounts of energies in the EM and neutrino sectors. However, while the EM plasma has a perfect thermal spectrum, the neutrino spectrum has a shift to higher energies.

The further dynamics of $\delta\rho_\nu$ depends on the balance between the energy transfer rates $\nu \rightarrow \text{EM}$ and $\text{EM} \rightarrow \nu$. Because of the energy dependence of the weak processes’ rate, the overabundance of the high-energy neutrino leads to the faster transfer $\nu \rightarrow \text{EM}$ than $\text{EM} \rightarrow \nu$, where we have thermal electrons. As a result, $\delta\rho_\nu$ continues falling below zero until neutrino-induced heating of the EM plasma temperature and/or the expansion of the Universe turn the negative energy transfer from the neutrino sector to zero.

Since the sign of $\delta\rho_\nu$ is associated with the sign of the correction ΔN_{eff} , we conclude that the injection of such high-energy neutrinos is associated with a decrease in N_{eff} below its ΛCDM value. This conclusion holds in the case when the EM plasma temperature is high enough during the neutrino injection, such that the interactions

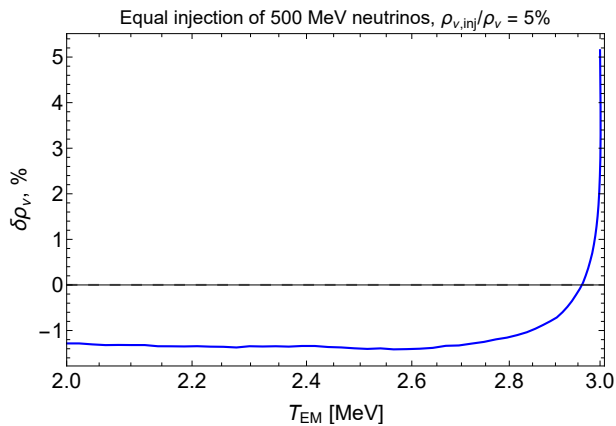


FIG. 6. The same setup as in Fig. 3 but under an injection of 500 MeV neutrinos.

between the neutrinos and the EM plasma are possible.

A similar result has been obtained in our previous work [33], which considered a setup with the injection of 70 MeV neutrinos but with a smaller amount within the discretization approach. The same behavior has been observed when considering the cosmological impact of HNLs decaying mainly into neutrinos (see also Refs. [28, 35]). These results, however, contradicted Ref. [34] (see also [26]), which studied the same setup with HNLs with masses below the pion mass and found that N_{eff} may only increase. Given that all these studies are based on the discretization method, the discrepancy became an open question. Our approach is completely independent and, therefore, resolves it.

We finish this discussion by directly comparing our method with the discretization codes. Let us consider the setup when we inject 70 MeV neutrinos with the amount $\rho_{\nu, \text{inj}}/\rho_{\nu, \text{tot}} = 5\%$. Fig. 5 shows the evolution of $\delta\rho_\nu$ and neutrino spectra snapshot according to DSMC and the discretization codes from [7, 33], where for the latter we take the results shown in Fig. 7 from Appendix A. In the discretization codes, the electron mass effects are included.

We see a very similar behavior of the evolution predicted by DSMC and [7], both in terms of $\delta\rho_\nu$ and the spectrum. The tiny discrepancy may be explained by the fact that we have neglected the electron mass in the DSMC calculations. On the other hand, the discrepancy between DSMC and Ref. [33] is somewhat larger. This may be explained by the fact that the caption of Fig. 7 in [33] does not correspond to the setup used to obtain the evolution.

3. Injection of 500 MeV neutrinos

Let us finalize this case study by considering the injection of 500 MeV neutrinos. The behavior of $\delta\rho_\nu$ is shown in Fig. 6; it resembles the features shown in the case of the injection of 70 MeV neutrinos.

The more important point is the performance of the DSMC setup. The running time required to simulate this setup was roughly the same as the running time for simulating 20 MeV and 70 MeV neutrinos. The 500 MeV case is already unrealistic to consider with the discretization codes, as the running time would grow by a factor of $> (500/70)^3 \simeq 400$.

C. Decays of long-lived SM particles

Let us now proceed with a more complicated case, when neutrinos are not injected directly in the decay chain but emerge via the evolution of heavy primary decay products Y , which may be muons or long-lived mesons such as π^\pm, K^\pm, K_L .

In the primordial plasma, Y s experience a non-trivial evolution once being injected. The interactions they are involved in include kinetic energy loss, interactions with nucleons, annihilation with themselves, and decays, see [50] for more details. The decay products generically involve neutrinos. This evolution influences their energy distribution among the neutrino and EM sectors.

Our approach for simulating this redistribution is the following. We first inject these particles into the plasma and then decay them using Monte Carlo techniques. For the case of charged decay products, we transfer all the kinetic energy to the EM plasma and then decay them at rest. This is because the energy loss rate is much faster than any other relevant process in the MeV plasma. This simplified description follows the state-of-the-art studies [29, 32]; the rest of the interactions discussed above will be added in the future. To simulate the phase space of the decay chain, we use `SensCalc` [67], a tool calculating the event rate with the decaying LLPs at various laboratory experiments. It contains a module handling LLP decay chains and, in particular, the decays of different SM particles. We have modified it to incorporate the evolution of mesons and muons in the primordial plasma. In absolutely the same way, it may be used to simulate decays of the LLPs, with these mesons appearing among the final states.

The neutrinos from Y s decays have a non-trivial spectrum. For instance, the neutrino distribution from decays of neutral kaons $K_L + K_S$ is shown in Fig. 7. They have the following main decay modes:

$$K_S \rightarrow 2\pi^0, \quad K_S \rightarrow \pi^+\pi^-, \quad (34)$$

$$K_L \rightarrow 3\pi^0, \quad K_L \rightarrow \pi^+\pi^-\pi^0, \quad K_L \rightarrow \pi^\pm l^\mp \nu_l \quad (35)$$

The neutral pions instantly decay into photons, just heating the EM plasma, whereas π^\pm, μ^\pm particles lose kinetic energy before decaying:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu, \quad \mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \quad (36)$$

The spectrum of neutrinos from all these particles has the high-energy part with $E_\nu \gg T = \mathcal{O}(1 \text{ MeV})$, and we

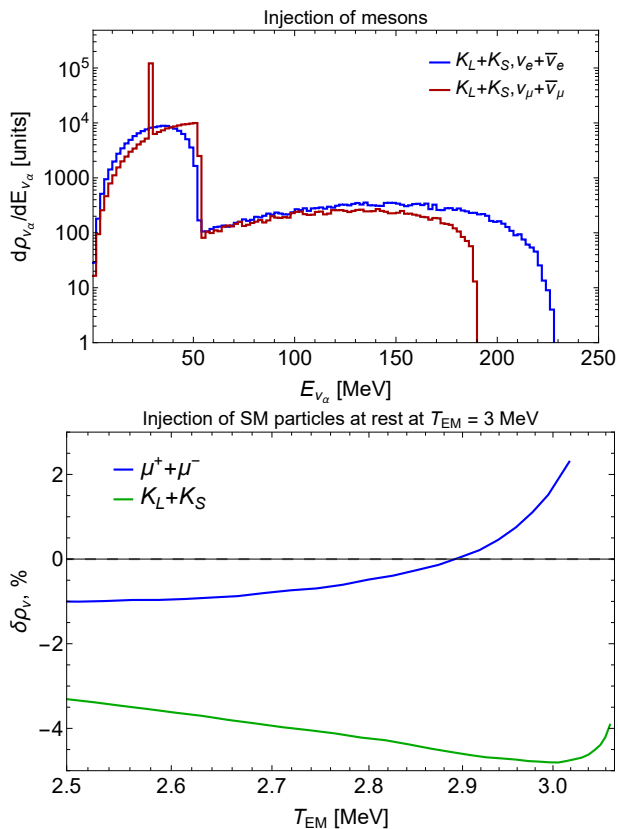


FIG. 7. Impact of injection of heavy SM particles in the primordial plasma. *Top panel:* the distribution of electron and muon neutrinos produced by decays of $K_L + K_S$ particles. When simulating their decay, we used the module of `SensCalc` tool [67]. For the chain of the decay products, we account for instant kinetic energy loss by charged particles. The continuous extension of the spectrum to $\simeq 200$ MeV is caused by the direct decay of kaons into neutrinos. The increase at $E_\nu = 50$ MeV is caused by decays of secondary muons stopped in the plasma, whereas the sharp increase at $E_\nu \approx 34$ MeV follows from decays of secondary pions. *Bottom panel:* The evolution of the quantity $\delta\rho_\nu$ under the injection of $\mu^+ + \mu^-$ (the blue curve) and $K_L + K_S$ (the green curve) in the primordial plasma at temperature $T = 3$ MeV.

expect the same behavior of $\delta\rho_\nu$ as in the case of the injections of high-energy neutrinos. Clarifying this question is important since the past studies [25, 29, 30] treated these injections using the semi-analytic integrated Boltzmann approach.

The evolution of $\delta\rho_\nu$ under the injection of $\mu^+ + \mu^-$ and $K_L + K_S$ is shown in Fig. 7. Let us start with the case of the muons. They inject 1/3 of their energy into the EM plasma, with the rest going to the non-thermal neutrino population. Completely similar to the instant neutrino injection case, $\delta\rho_\nu$, being initially positive, instantly decreases below the Λ CDM value. This finding contradicts the studies [29, 30], which considered the scenario of decays of Higgs-like scalars into two muons and found that it increases N_{eff} even in the regime of small

scalar lifetimes $\mathcal{O}(0.1 \text{ s})$.

The $K_L + K_S$ case is also interesting. Decaying, they put most of their energy into the EM plasma sector, so we start with a negative $\delta\rho_\nu$. However, the presence of very high-energy neutrinos with $E_\nu = 100 - 200$ MeV leads to a further slight drop of $\delta\rho_\nu$, and then it tries to approach the equilibrium.

VIII. CONCLUSIONS

Upcoming CMB observations will reach unprecedented precision, which may be used to discover or constrain new physics that was present in the primordial plasma at temperatures as large as a few MeV. To reach this goal, we have to understand the dynamics of the Early Universe in the presence of new physics. It requires solving the neutrino Boltzmann equation across a variety of scenarios, including long-lived relics, non-standard neutrino interactions, and lepton asymmetry in the neutrino sector.

Current state-of-the-art methods are limited in scope and face computational challenges when neutrino evolution deviates significantly from the standard scenario. These limitations arise from the complex phase space of interactions, the presence of high-energy neutrinos, and the lack of analytic matrix elements – features that are common in systems with new physics. Furthermore, the complexity of implementing these methods makes it difficult to extend them to include various new physics models, even within the range of applicability.

In this paper, we have presented an approach that is potentially free from all these limitations. It is based on the Direct Simulation Monte Carlo method to solve the Boltzmann equation, see Sec. IV. The DSMC approach has been previously applied to rarefied gases and, in its existing form, cannot be applied to studying the Early Universe. Fundamental modifications are required, such as including the Universe expansion, the hierarchy between weak and electromagnetic interaction rates, the Pauli principle, neutrino oscillations, and the presence of decaying particles. We have discussed these features and how to include them in the DSMC in Sec. V.

In Sec. VI, we have described our current proof-of-principle implementation of the DSMC approach for neutrinos that incorporates these modifications. We have validated it by conducting cross-checks within well-understood physics scenarios (see also Appendix A). In Sec. VII, we have applied the prototype to several toy scenarios that mimic real case studies: the equilibration of the neutrinos and EM plasma initially having different temperatures (Sec. VIIA), injection of high-energy neutrinos (Sec. VIIB), and decays of particles Y with complicated decay chain, including muons, pions, and kaons (Sec. VIIC). In particular, we have studied the evolution of the plasma in the presence of neutrinos with energies $E_\nu \gg T \simeq 1$ MeV. We have found that they decrease the neutrino-to-electromagnetic energy density

ratio, which leads to a negative change in N_{eff} compared to the Standard cosmological scenario (Fig. 5). This finding resolves the previously existing discrepancy between different state-of-the-art approaches in predictions about the dynamics of N_{eff} in the presence of high-energy neutrinos. The same conclusion was the case for metastable muons and mesons appearing among the decay products of various classes of long-lived particles (Sec. VII C).

Our current neutrino DSMC code is rather proof-of-principle, limited by the efficiency of the implementation and some approximations. Once these problems are overcome, it will result in a powerful independent method of solving neutrino Boltzmann equations. We leave this for

future work.

ACKNOWLEDGEMENTS

We thank Stefan Stefanov for the in-depth review of the implementation of the proof-of-principle DSMC approach for neutrinos, and Fabio Peano, Luís Olivera e Silva, and Kyrylo Bondarenko for discussions at the early stages of this project. MO received support from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 860881-HIDDeN.

Appendix A: Cross-checks

1. Approaching thermal equilibrium

To test whether the DSMC simulation brings the system of neutrinos and EM particles to the dynamical equilibrium defined by Eqs. (30) and (31), we will use the following setup:

- The Universe content is neutrinos and anti-neutrinos of all flavors together with electrons, positrons, and photons.
- The expansion of the Universe is not taken into account. Therefore, the total energy density of the system is constant, and the energy/momenta of particles are subject only to their interactions.
- The initial distribution function of neutrinos consists of two components that are the same for all flavors:
 1. The equilibrium component, which has Fermi-Dirac distribution with the temperature $T_{\nu}^{\text{ini}} = 3\text{MeV}$.
 2. The non-equilibrium component - neutrinos with an arbitrary energy distribution, with the energy density constituting some fraction $\ll 1$ of the equilibrium energy density.

The first sub-scenario we consider is where there are no non-equilibrium neutrinos, so the system is initially in the fully equilibrium state. If at least one component of the DSMC simulation is implemented incorrectly, the system will escape the equilibrium, tending to the false ground state. A prominent example is when the cross-sections are taken to be velocity-independent; then, the distribution of the system tends to the fake-equilibrium spectrum $d\rho_{\nu}/dE_{\nu} \sim E_{\nu}^2 \times f_{\text{FD}}$ instead of the correct $E_{\nu}^3 f_{\text{FD}}$ (see a discussion for relativistic particles with Boltzmann statistics has been made in [65]). Another issue may be if the maximal interaction weight ω_{max} in the acceptance criterion of the pair’s interaction (22) is not actually the maximal one. Then, the system falls into the state with $\delta\rho_{\nu} < 0$.

Our DSMC implementation passes this test, see Fig. 8.

Next, we include the non-equilibrium neutrinos. An example of such a simulation is shown in Fig.9. Its results are in perfect agreement with the theoretical expectations.

2. Energy transition rates

In this scenario, we will reproduce the semi-analytical result of [24, 52], where the evolution of neutrinos in the Early Universe was studied under an assumption that every moment of time, the shape of their energy distribution is thermal. The energy transition rates were calculated analytically in terms of the temperatures of neutrinos and EM plasma $T_{\nu_{\alpha}}, T$. The Boltzmann equations are reduced to the simple system of differential equations on $T_{\nu_{\alpha}}, T$. For our simulation, the following setup will be used:

- The Universe’s content is neutrinos and anti-neutrinos of all flavors together with electrons, positrons, and photons.
- The simulation is altered such that neutrino distributions always have the shape (30) at each simulation step. Basically, we treat neutrinos in exactly the same way as the EM particles in the full DSMC simulation.
- The expansion of the Universe is not included, to concentrate on the energy exchange rates.

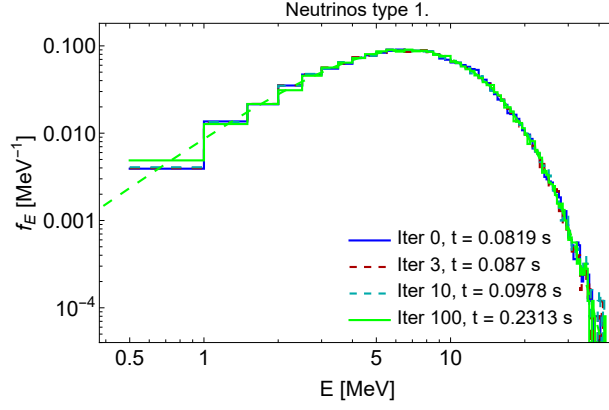


FIG. 8. The evolution of the neutrino distribution $d\rho_\nu/dE_\nu$ under the assumption of fully equilibrium initial conditions (30) and (31). No significant changes are developed throughout the simulation. The minor changes are related to the quality of the sampler of the kinematics of the electrons via the Fermi-Dirac distribution.

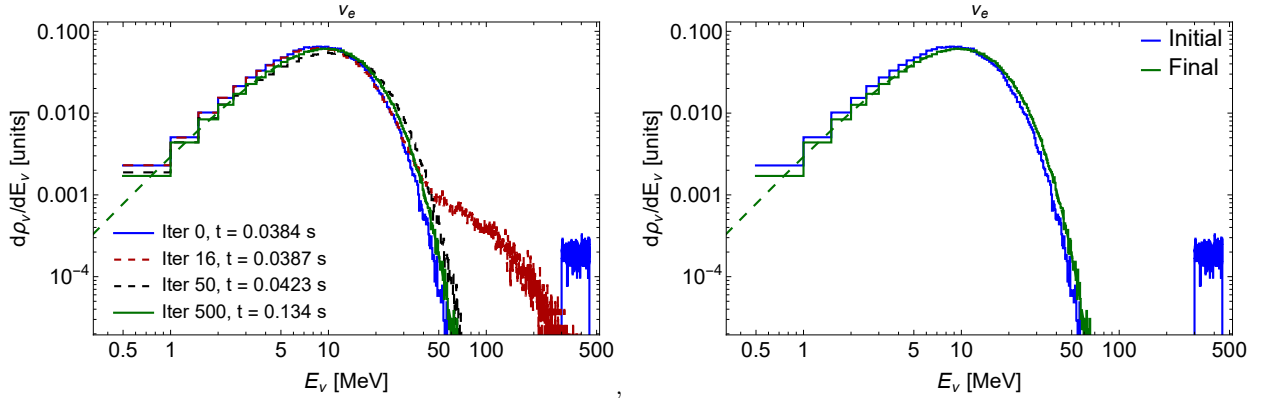


FIG. 9. Evolution of the neutrino distribution function under the initial setup with equilibrium neutrinos and EM plasma at temperature $T = 3$ MeV and non-equilibrium neutrinos with energies uniformly distributed in the range $300\text{MeV} < E_\nu < 450\text{MeV}$. Their total energy density is related to the total energy of equilibrium part as $\rho_{\nu\alpha}^{\text{non-eq}}/\rho_{\nu\alpha}^{\text{eq}} = 0.15$. The non-equilibrium part of the spectra rapidly loses its energy in the first steps of simulation, leading to the distortions of the spectra at high energies which are eventually equilibrated. *Left plot*: snapshots of the binned neutrino distribution function as obtained at different iterations of the DSMC simulation. The iteration 0 corresponds to the initial setup, while the iteration 200 is the final state. *Right plot*: the comparison of the neutrino distribution function between the initial (the blue line) and final (the green line) iterations. The final distribution approaches the analytic Fermi-Dirac neutrino distribution with the temperature equal to the temperature of the electromagnetic plasma T_{EM} , which we shown by the dashed green line.

The example of the resulting evolution of the energy density of the neutrino plasma is presented in Fig.10, where the almost perfect correspondence between theoretical predictions and simulation can be seen. Such reproduction of the energy evolution behavior confirms that averaged energy transition rates are computed correctly.

3. Expansion and decoupling

In the third cross-check, we will follow the previous setup, but with the expansion of the Universe included. Due to initial difference between temperatures of neutrino and EM plasma, we expect some remaining inequality between them, since the start of the simulation occurs close to the temperature of the neutrino decoupling. In similar terms,

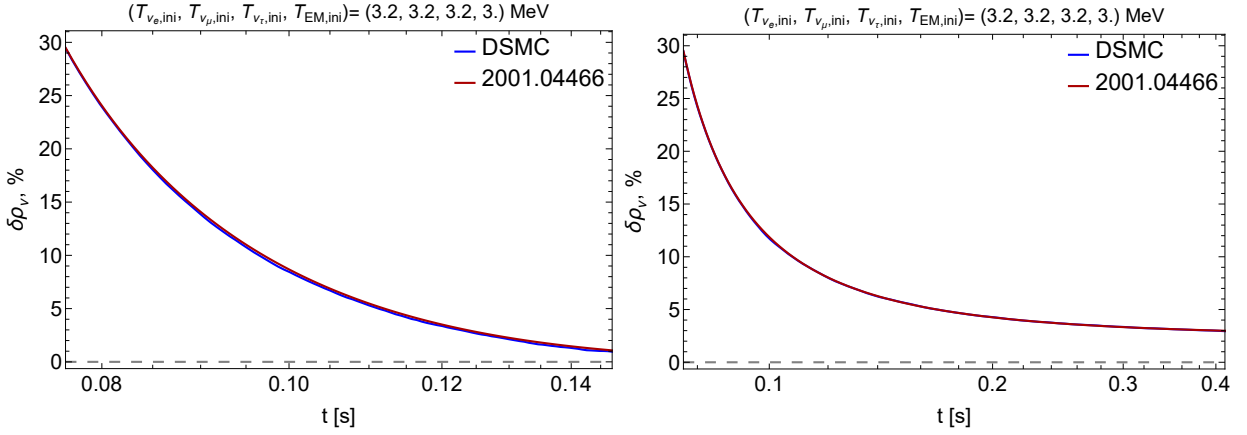


FIG. 10. The evolution of the ratio of the neutrino energy density to the EM energy density in DSMC simulation compared to the theoretical prediction from [24], under an assumption that the shape of the neutrino distribution function is always thermal at each step of the simulation. The initial conditions for the setup are $T_{\nu_i} = 3.2$ MeV for every flavor and the temperature of the EM plasma is $T_{EM} = 3$ MeV. *Left panel:* not including the expansion of the Universe. Due to the absence of expansion, the ratio approaches to the exact SM value. *Right panel:* expansion included.

we present the example of such comparison in Fig. 10.

-
- [1] S. Bashinsky and U. Seljak, “Neutrino perturbations in CMB anisotropy and matter clustering,” *Phys. Rev. D* **69** (2004) 083002, [arXiv:astro-ph/0310198](#).
 - [2] D. Baumann, F. Beutler, R. Flauger, D. Green, A. Slosar, M. Vargas-Magaña, B. Wallisch, and C. Yèche, “First constraint on the neutrino-induced phase shift in the spectrum of baryon acoustic oscillations,” *Nature Phys.* **15** (2019) 465–469, [arXiv:1803.10741 \[astro-ph.CO\]](#).
 - [3] J. Alvey, M. Escudero, N. Sabti, and T. Schwetz, “Cosmic neutrino background detection in large-neutrino-mass cosmologies,” *Phys. Rev. D* **105** (2022) no. 6, 063501, [arXiv:2111.14870 \[hep-ph\]](#).
 - [4] G. Mangano, G. Miele, S. Pastor, and M. Peloso, “A Precision calculation of the effective number of cosmological neutrinos,” *Phys. Lett. B* **534** (2002) 8–16, [arXiv:astro-ph/0111408](#).
 - [5] J. J. Bennett, G. Buldgen, M. Drewes, and Y. Y. Y. Wong, “Towards a precision calculation of the effective number of neutrinos N_{eff} in the Standard Model I: the QED equation of state,” *JCAP* **03** (2020) 003, [arXiv:1911.04504 \[hep-ph\]](#). [Addendum: *JCAP* **03**, A01 (2021)].
 - [6] 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_{eff} in the Standard Model II: Neutrino decoupling in the presence of flavour oscillations and finite-temperature QED,” *JCAP* **04** (2021) 073, [arXiv:2012.02726 \[hep-ph\]](#).
 - [7] K. Akita and M. Yamaguchi, “A precision calculation of relic neutrino decoupling,” *JCAP* **08** (2020) 012, [arXiv:2005.07047 \[hep-ph\]](#).
 - [8] J. Froustey, C. Pitrou, and M. C. Volpe, “Neutrino decoupling including flavour oscillations and primordial nucleosynthesis,” *JCAP* **12** (2020) 015, [arXiv:2008.01074 \[hep-ph\]](#).
 - [9] M. Cielo, M. Escudero, G. Mangano, and O. Pisanti, “Neff in the Standard Model at NLO is 3.043,” *Phys. Rev. D* **108** (2023) no. 12, L121301, [arXiv:2306.05460 \[hep-ph\]](#).
 - [10] M. Drewes, Y. Georis, M. Klasen, L. P. Wiggnering, and Y. Y. Y. Wong, “Towards a precision calculation of N_{eff} in the Standard Model III: Improved estimate of NLO corrections to the collision integral,” [arXiv:2402.18481 \[hep-ph\]](#).
 - [11] O. Pisanti, A. Cirillo, S. Esposito, F. Iocco, G. Mangano, G. Miele, and P. D. Serpico, “PARthENoPE: Public Algorithm Evaluating the Nucleosynthesis of Primordial Elements,” *Comput. Phys. Commun.* **178** (2008) 956–971, [arXiv:0705.0290 \[astro-ph\]](#).
 - [12] C. Pitrou, A. Coc, J.-P. Uzan, and E. Vangioni, “Precision big bang nucleosynthesis with improved Helium-4 predictions,” *Phys. Rept.* **754** (2018) 1–66, [arXiv:1801.08023 \[astro-ph.CO\]](#).
 - [13] **Planck** Collaboration, N. Aghanim *et al.*, “Planck 2018 results. I. Overview and the cosmological legacy of Planck,” *Astron. Astrophys.* **641** (2020) A1, [arXiv:1807.06205 \[astro-ph.CO\]](#).
 - [14] **Planck** Collaboration, N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters,” *Astron. Astrophys.* **641** (2020) A6, [arXiv:1807.06209 \[astro-ph.CO\]](#).
 - [15] Y. I. Izotov, T. X. Thuan, and N. G. Guseva, “A new determination of the primordial He abundance using the He i $\lambda 10830$ Å emission line: cosmological implications,” *Mon. Not. Roy. Astron. Soc.* **445** (2014) no. 1, 778–793, [arXiv:1408.6953 \[astro-ph.CO\]](#).

- [16] E. Aver, K. A. Olive, and E. D. Skillman, “The effects of He I λ 10830 on helium abundance determinations,” *JCAP* **1507** (2015) no. 07, 011, [arXiv:1503.08146 \[astro-ph.CO\]](#).
- [17] A. Peimbert, M. Peimbert, and V. Luridiana, “The primordial helium abundance and the number of neutrino families,” *Rev. Mex. Astron. Astrofis.* **52** (2016) 419, [arXiv:1608.02062 \[astro-ph.CO\]](#).
- [18] V. Fernández, E. Terlevich, A. I. Díaz, R. Terlevich, and F. F. Rosales-Ortega, “Primordial helium abundance determination using sulphur as metallicity tracer,” *Monthly Notices of the Royal Astronomical Society* **478** (2018) no. 4, 5301–5319. <http://dx.doi.org/10.1093/mnras/sty1206>.
- [19] M. Valerdi, A. Peimbert, M. Peimbert, and A. Sixtos, “Determination of the Primordial Helium Abundance Based on NGC 346, an H II Region of the Small Magellanic Cloud,” *Astrophys. J.* **876** (2019) no. 2, 98, [arXiv:1904.01594 \[astro-ph.GA\]](#).
- [20] E. Aver, D. A. Berg, A. S. Hirschauer, K. A. Olive, R. W. Pogge, N. S. J. Rogers, J. J. Salzer, and E. D. Skillman, “A comprehensive chemical abundance analysis of the extremely metal poor Leoncino Dwarf galaxy (AGC 198691),” *Mon. Not. Roy. Astron. Soc.* **510** (2021) no. 1, 373–382, [arXiv:2109.00178 \[astro-ph.GA\]](#).
- [21] A. Matsumoto *et al.*, “EMPRESS. VIII. A New Determination of Primordial He Abundance with Extremely Metal-poor Galaxies: A Suggestion of the Lepton Asymmetry and Implications for the Hubble Tension,” *Astrophys. J.* **941** (2022) no. 2, 167, [arXiv:2203.09617 \[astro-ph.CO\]](#).
- [22] M. Archidiacono and S. Hannestad, “Updated constraints on non-standard neutrino interactions from Planck,” *JCAP* **07** (2014) 046, [arXiv:1311.3873 \[astro-ph.CO\]](#).
- [23] F. Forastieri, M. Lattanzi, and P. Natoli, “Constraints on secret neutrino interactions after Planck,” *JCAP* **07** (2015) 014, [arXiv:1504.04999 \[astro-ph.CO\]](#).
- [24] M. Escudero Abenza, “Precision early universe thermodynamics made simple: N_{eff} and neutrino decoupling in the Standard Model and beyond,” *JCAP* **05** (2020) 048, [arXiv:2001.04466 \[hep-ph\]](#).
- [25] 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,” *JCAP* **09** (2020) 051, [arXiv:2005.06721 \[hep-ph\]](#).
- [26] A. D. Dolgov, S. H. Hansen, G. Raffelt, and D. V. Semikoz, “Heavy sterile neutrinos: Bounds from big bang nucleosynthesis and SN1987A,” *Nucl. Phys. B* **590** (2000) 562–574, [arXiv:hep-ph/0008138](#).
- [27] S. Hannestad, “What is the lowest possible reheating temperature?,” *Phys. Rev. D* **70** (2004) 043506, [arXiv:astro-ph/0403291](#).
- [28] O. Ruchayskiy and A. Ivashko, “Restrictions on the lifetime of sterile neutrinos from primordial nucleosynthesis,” *JCAP* **10** (2012) 014, [arXiv:1202.2841 \[hep-ph\]](#).
- [29] A. Fradette and M. Pospelov, “BBN for the LHC: constraints on lifetimes of the Higgs portal scalars,” *Phys. Rev. D* **96** (2017) no. 7, 075033, [arXiv:1706.01920 \[hep-ph\]](#).
- [30] 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** (2019) no. 7, 075004, [arXiv:1812.07585 \[hep-ph\]](#).
- [31] A. Boyarsky, M. Ovchinnikov, O. Ruchayskiy, and V. Syvolap, “Improved big bang nucleosynthesis constraints on heavy neutral leptons,” *Phys. Rev. D* **104** (2021) no. 2, 023517, [arXiv:2008.00749 \[hep-ph\]](#).
- [32] N. Sabti, A. Magalich, and A. Filimonova, “An Extended Analysis of Heavy Neutral Leptons during Big Bang Nucleosynthesis,” *JCAP* **11** (2020) 056, [arXiv:2006.07387 \[hep-ph\]](#).
- [33] A. Boyarsky, M. Ovchinnikov, N. Sabti, and V. Syvolap, “When feebly interacting massive particles decay into neutrinos: The Neff story,” *Phys. Rev. D* **104** (2021) no. 3, 035006, [arXiv:2103.09831 \[hep-ph\]](#).
- [34] 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** (2021) no. 1, 016026, [arXiv:2104.11752 \[hep-ph\]](#).
- [35] 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** (2022) no. 8, 083513, [arXiv:2109.11176 \[hep-ph\]](#).
- [36] **Simons Observatory** Collaboration, P. Ade *et al.*, “The Simons Observatory: Science goals and forecasts,” *JCAP* **02** (2019) 056, [arXiv:1808.07445 \[astro-ph.CO\]](#).
- [37] **CMB-S4** Collaboration, K. N. Abazajian *et al.*, “CMB-S4 Science Book, First Edition,” [arXiv:1610.02743 \[astro-ph.CO\]](#).
- [38] S. Hannestad and J. Madsen, “Neutrino decoupling in the early universe,” *Phys. Rev. D* **52** (1995) 1764–1769, [arXiv:astro-ph/9506015](#).
- [39] E. Grohs, G. M. Fuller, C. T. Kishimoto, M. W. Paris, and A. Vlasenko, “Neutrino energy transport in weak decoupling and big bang nucleosynthesis,” *Phys. Rev. D* **93** (2016) no. 8, 083522, [arXiv:1512.02205 \[astro-ph.CO\]](#).
- [40] S. Gariazzo, P. F. de Salas, and S. Pastor, “Thermalisation of sterile neutrinos in the early Universe in the 3+1 scheme with full mixing matrix,” *JCAP* **07** (2019) 014, [arXiv:1905.11290 \[astro-ph.CO\]](#).
- [41] G. A. Bird, “Monte carlo simulation of gas flows,” *Annual Review of Fluid Mechanics* **10** (1978) no. 1, 11–31.
- [42] G. Bird, “Monte carlo simulation of gas flows,” *Annu Rev Fluid Mech* **10** (2003) 11–31.
- [43] E. Roohi and S. Stefanov, “Collision partner selection schemes in dsmc: From micro/nano flows to hypersonic flows,” *Physics Reports* **656** (2016) 1–38. <https://www.sciencedirect.com/science/article/pii/S0370157316302344>.
- [44] S. Stefanov, “On the basic concepts of the direct simulation monte carlo method,” *Physics of Fluids* **31** (2019) no. 6, .
- [45] R. Fonseca *et al.*, “Lect. notes comp. sci. 2331, 342 (2002); ra fonseca et al,” *Plasma Phys. Control. Fusion* **50** (2008) 124034.
- [46] M. A. Gallis, J. R. Torczynski, S. J. Plimpton, D. J. Rader, and T. Koehler, “Direct simulation monte carlo: The quest for speed,” tech. rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2014.

- [47] M. Ovchinnikov and V. Syvolap, “Primordial neutrinos and new physics: novel approach to solving neutrino Boltzmann equation,”.
- [48] M. Kawasaki, K. Kohri, and T. Moroi, “Big-Bang nucleosynthesis and hadronic decay of long-lived massive particles,” *Phys. Rev. D* **71** (2005) 083502, [arXiv:astro-ph/0408426](#).
- [49] A. D. Dolgov, “Neutrinos in cosmology,” *Phys. Rept.* **370** (2002) 333–535, [arXiv:hep-ph/0202122](#).
- [50] K. Akita, G. Baur, M. Ovchinnikov, and V. Syvolap, “New physics decaying into muons and mesons in MeV primordial plasma: what happens next?,” *to appear* (2024), [arXiv:2410.XXXXX](#) [[hep-ph](#)].
- [51] A. Vlasenko, G. M. Fuller, and V. Cirigliano, “Neutrino Quantum Kinetics,” *Phys. Rev. D* **89** (2014) no. 10, 105004, [arXiv:1309.2628](#) [[hep-ph](#)].
- [52] M. Escudero, “Neutrino decoupling beyond the Standard Model: CMB constraints on the Dark Matter mass with a fast and precise N_{eff} evaluation,” *JCAP* **02** (2019) 007, [arXiv:1812.05605](#) [[hep-ph](#)].
- [53] J. Beacham *et al.*, “Physics Beyond Colliders at CERN: Beyond the Standard Model Working Group Report,” *J. Phys. G* **47** (2020) no. 1, 010501, [arXiv:1901.09966](#) [[hep-ex](#)].
- [54] G. Bird, “Direct simulation and the boltzmann equation,” *Physics of Fluids* **13** (1970) no. 11, 2676–2681.
- [55] M. Ivanov and S. Rogasinskii, “Theoretical analysis of traditional and modern schemes of the dsmc method,” *Rarefied gas dynamics* (1991) 629–642.
- [56] G. Bird, “The velocity distribution function within a shock wave,” *Journal of Fluid Mechanics* **30** (1967) no. 3, 479–487.
- [57] K. Koura, “Transient couette flow of rarefied binary gas mixtures,” *The Physics of Fluids* **13** (1970) no. 6, 1457–1466.
- [58] K. Koura, “Null-collision technique in the direct-simulation monte carlo method,” *The Physics of fluids* **29** (1986) no. 11, 3509–3511.
- [59] G. Bird, “Perception of numerical methods in rarefied gasdynamics,” *Progress in Astronautics and Aeronautics* **117** (1989) 211–226.
- [60] E. Roohi and S. Stefanov, “Collision partner selection schemes in dsmc: From micro/nano flows to hypersonic flows,” *Physics Reports* **656** (2016) 1–38.
- [61] D. P. Schmidt and C. J. Rutland, “A new droplet collision algorithm,” *Journal of Computational Physics* **164** (2000) no. 1, 62–80.
- [62] J.-S. Wu and K.-C. Tseng, “Analysis of micro-scale gas flows with pressure boundaries using direct simulation monte carlo method,” *Computers & Fluids* **30** (2001) no. 6, 711–735.
- [63] J.-S. Wu and Y.-Y. Lian, “Parallel three-dimensional direct simulation monte carlo method and its applications,” *Computers & Fluids* **32** (2003) no. 8, 1133–1160.
- [64] A. Venkattraman, A. A. Alexeenko, M. Gallis, and M. Ivanov, “A comparative study of no-time-counter and majorant collision frequency numerical schemes in dsmc,” in *AIP Conference Proceedings*, vol. 1501, pp. 489–495, American Institute of Physics. 2012.
- [65] F. Peano, M. Marti, L. Silva, and G. Coppa, “Statistical kinetic treatment of relativistic binary collisions,” *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* **79** (2009) no. 2, 025701.
- [66] C. Bierlich *et al.*, “A comprehensive guide to the physics and usage of PYTHIA 8.3,” *SciPost Phys. Codeb.* **2022** (2022) 8, [arXiv:2203.11601](#) [[hep-ph](#)].
- [67] M. Ovchinnikov, J.-L. Tastet, O. Mikulenko, and K. Bondarenko, “Sensitivities to feebly interacting particles: Public and unified calculations,” *Phys. Rev. D* **108** (2023) no. 7, 075028, [arXiv:2305.13383](#) [[hep-ph](#)].
- [68] K. J. Kelly and Y. Zhang, “Mononeutrino at DUNE: New Signals from Neutrinophilic Thermal Dark Matter,” *Phys. Rev. D* **99** (2019) no. 5, 055034, [arXiv:1901.01259](#) [[hep-ph](#)].
- [69] J. E. Kim, “Light Pseudoscalars, Particle Physics and Cosmology,” *Phys. Rept.* **150** (1987) 1–177.
- [70] P. Ilten, Y. Soreq, M. Williams, and W. Xue, “Serendipity in dark photon searches,” *JHEP* **06** (2018) 004, [arXiv:1801.04847](#) [[hep-ph](#)].
- [71] K. Bondarenko, A. Boyarsky, D. Gorbunov, and O. Ruchayskiy, “Phenomenology of GeV-scale Heavy Neutral Leptons,” *JHEP* **11** (2018) 032, [arXiv:1805.08567](#) [[hep-ph](#)].