

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Ala-Mattinen, Kalle; Heikinheimo, Matti; Kainulainen, Kimmo; Tuominen, Kimmo

Title: Momentum distributions of cosmic relics : Improved analysis

Year: 2022

Version: Published version

Copyright: © Authors, 2022

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Ala-Mattinen, K., Heikinheimo, M., Kainulainen, K., & Tuominen, K. (2022). Momentum distributions of cosmic relics : Improved analysis. *Physical Review D*, 105(12), Article 123005. <https://doi.org/10.1103/PhysRevD.105.123005>

Momentum distributions of cosmic relics: Improved analysisKalle Ala-Mattinen^{1,2,*}, Matti Heikinheimo^{1,2,†}, Kimmo Kainulainen^{3,2,‡} and Kimmo Tuominen^{1,2,§}¹*Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland*²*Helsinki Institute of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland*³*Department of Physics, University of Jyväskylä, 40500 Jyväskylä, Finland*

(Received 24 January 2022; accepted 18 May 2022; published 7 June 2022)

We solve coupled momentum-dependent Boltzmann equations for the phase space distribution of cosmic relic particles, without resorting to approximations of assuming kinetic equilibrium or neglecting backscattering or elastic interactions. Our method is amendable to precision numerical computations. To test it, we consider two benchmark models where the momentum dependence of dark matter distribution function is potentially important: a real singlet scalar extension near the Higgs resonance and a sterile neutrino dark matter model with a singlet scalar mediator. The singlet scalar example shows that the kinetic equilibrium may hold surprisingly well even near sharp resonances. However, the integrated method may underestimate the relic density by up to 40% in extreme cases. In the sterile neutrino dark matter model, we studied how the inclusion of previously ignored elastic interactions and processes with initial state sterile neutrinos could affect the nonthermal nature of their resulting distributions. Here the effects turned out to be negligible, proving the robustness of the earlier predictions.

DOI: [10.1103/PhysRevD.105.123005](https://doi.org/10.1103/PhysRevD.105.123005)**I. INTRODUCTION**

Current cosmological observations can be accommodated within the cold dark matter (CDM) paradigm [1]. This hypothesis is appealing within our present understanding of the structure of ordinary matter: extending the Standard Model (SM) of elementary particles and their interactions with dark matter particle degrees of freedom allows the abundance of CDM to be created by thermal production and decoupling of the dark matter particles in the expanding early Universe. The standard treatment for calculating this abundance of CDM particles relies on the Zel'dovich-Okun-Pikelner-Lee-Weinberg (ZOPLW) equation [2,3].

The ZOPLW equation is obtained from the Boltzmann equation by integrating over the phase space of the dark matter under the assumptions of detailed balance and kinetic equilibrium distributions [4–6]. However, the momentum distribution of dark matter may contain essential information that is neglected in this treatment. For example, if the dark matter production takes place at a

resonance region, where the DM annihilation rate is strongly momentum dependent, the elastic reactions might not be fast enough to keep kinetic equilibrium. In this setting the true annihilation rate and hence the final DM abundance may deviate from the value obtained under the equilibrium assumption [6–8]. Another example concerns warm dark matter (WDM) whose momentum distribution may directly influence the cosmic structure formation by reducing the number of DM halos at small scales compared to CDM. In kinetic equilibrium, the suppression of the matter power spectrum can be well approximated via a single scale given by the WDM mass [9]. However, if the DM particle is not in kinetic equilibrium, the resulting suppression may be more complicated [9–12].

Earlier calculations accounting for the DM momentum distributions tend to rely on simplifying approximations. For example, in the analysis of [10,11], the elastic interactions of the initial state DM particles have been neglected. In this paper, we complement these earlier analyses by presenting a numerical method based on discretization of the momentum space that allows for a completely general solution of the Boltzmann equations for the momentum distributions of multiple number of particle species, any number of which can be out of equilibrium. We note that advanced momentum-dependent methods have also been developed and used to treat neutrino oscillations in the early Universe [13–16].

We demonstrate our method in the context of the dark matter production near a sharp resonance, comparing our results with the ZOPLW approach and with

*kalle.ala-mattinen@helsinki.fi

†matti.heikinheimo@helsinki.fi

‡kimmo.kainulainen@jyu.fi

§kimmo.i.tuominen@helsinki.fi

the momentum-dependent method of [7], which uses a generalized relaxation time approximation for the numerically expensive backreaction terms. Our results validate the approximation scheme of [7] to its expected accuracy. Furthermore, we find that this scheme slightly overestimates the effect of elastic scattering channels, and this seems to be the case also with the truncated derivative methods used in [6,8]. We also apply our method in a sterile neutrino DM model including a singlet scalar mediator, first analyzed with simplified evolution equations in [10,11]. We find that neither including elastic interaction channels, nor adding new collision terms induced by a symmetry breaking changes the results appreciably. This verifies that the approximations used in [10,11] are robust and their results remain valid in the full solution.

The paper is organized as follows: In Sec. II we describe the discretization of the collision integrals. We then apply the developed methodology to two benchmark models: first, in Sec. III to the model where SM is extended with a real singlet scalar and then in Sec. IV to the model where the additional fields are a real singlet scalar and a sterile neutrino. In Sec. V we present our conclusions and outlook towards further work. Many details of the computations can be found in the Appendixes.

II. THE KINETIC EQUATION

In an expanding homogeneous and isotropic universe, the Boltzmann equation can be written as

$$\frac{\partial f(p, t)}{\partial t} - H p \frac{\partial f}{\partial p} = \sum \mathcal{C}_{\text{coll}}[f], \quad (1)$$

where $\mathcal{C}_{\text{coll}}$ are the collision terms describing the chemical and kinetic balances, $H = \dot{a}/a$ is the Hubble parameter, and $a(t)$ is the scale factor. Expansion of the Universe is best quantified by integrating along the curves of constant comoving momentum, $k = ap$. Then, writing the Liouville operator in terms of k , the momentum derivative vanishes and we have

$$\frac{df(k, t)}{dt} = \sum \mathcal{C}_{\text{coll}}[f], \quad (2)$$

where we identified $f(p, t) = f(k/a(t), t) \equiv f(k, t)$. It is more natural to work with temperature instead of time, thus we define a dimensionless variable $x \equiv m_0/T$, where m_0 is some reference scale, and T is photon temperature. To evaluate the Jacobian of this transformation, we use the adiabatic radiation era time-temperature relation $\dot{s}/s = -3H$, where $s = 2\pi^2 h_{\text{eff}} T^3/45$ is the entropy density of the Universe, $H = (4\pi^3 g_{\text{eff}}/45)^{1/2} T^2/M_{\text{P}}$, and $g_{\text{eff}}(T)$ and $h_{\text{eff}}(T)$ are the effective number of relativistic energy and entropy degrees of freedom. This implies

$$\frac{\dot{T}}{T} = -\sqrt{\frac{4\pi^3}{45}} \frac{h_{\text{eff}}(T)}{g_*^{1/2}(T)} \frac{T^2}{M_{\text{P}}}, \quad (3)$$

where

$$g_*^{1/2}(T) \equiv \frac{h_{\text{eff}}(T)}{g_{\text{eff}}^{1/2}(T)} \left(1 + \frac{T}{3h_{\text{eff}}} \frac{dh_{\text{eff}}}{dT} \right). \quad (4)$$

We want to replace also the momentum with a dimensionless variable. Using again the adiabaticity condition, one finds $(a/a_0)^3 = T_0^3 h_{\text{eff}}(T_0)/T^3 h_{\text{eff}}(T)$, which suggests to define

$$\xi \equiv \frac{k}{T_0 a(t(T_0))} = \left(\frac{h_{\text{eff}}(T_0)}{h_{\text{eff}}(T)} \right)^{1/3} \frac{p}{T}, \quad (5)$$

where k is the comoving momentum, p is the physical momentum, and a_0 is the scale factor evaluated at some reference temperature T_0 , which we set equal to the reference mass: $T_0 \equiv m_0$.

The Boltzmann equation in dimensionless variables becomes

$$\frac{df(\xi, x)}{dx} = \sqrt{\frac{45}{4\pi^3}} \frac{g_*^{1/2}(m_0/x)}{h_{\text{eff}}(m_0/x)} \frac{x M_{\text{P}}}{m_0^2} \sum \mathcal{C}_{\text{coll}}[f]. \quad (6)$$

This Boltzmann equation, written in comoving variables, can be solved numerically by discretizing in variables x and ξ . Given such a discretization, the role of the parameters m_0 and T_0 is to tune the dimensionless variables to probe the desired temperatures and physical momenta. Before describing this process in detail, we must first carefully describe the structure of the collision terms $\mathcal{C}_{\text{coll}}$.

A. Collision integral

The collision term for generic two-particle interactions $12 \leftrightarrow 34$ is given by

$$\mathcal{C}_{\text{coll}}[f_i] = \frac{1}{2E_1} \int d\text{PS}_{234} \Lambda(f_1, f_2, f_3, f_4) |M_{12 \rightarrow 34}|^2, \quad (7)$$

where the integration measure over the phase space is

$$d\text{PS}_{ijk} = (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) d^3 \tilde{p}_i d^3 \tilde{p}_j d^3 \tilde{p}_k, \quad (8)$$

with $d^3 \tilde{p} = d^3 p / [(2\pi)^3 2E]$. We always assume that labels $i = 1, 2, 3, 4$ denote all internal degrees of freedom associated with a given distribution function f_i . The phase space factor $\Lambda(f_1, f_2, f_3, f_4)$ is defined as

$$\Lambda(f_1, f_2, f_3, f_4) = f_3 f_4 (1 \pm f_1)(1 \pm f_2) - f_1 f_2 (1 \pm f_3)(1 \pm f_4), \quad (9)$$

with $+(-)$ corresponding to the boson (fermion) case. Finally, $|M_{12\rightarrow 34}|^2$ is the matrix element squared, summed or integrated over the internal degrees of freedom associated with the labels $i = 2, 3, 4$. The matrix element squared is also assumed to contain all relevant symmetry factors for the initial and final states. The collision integral $\mathcal{C}_{\text{coll}}$ naturally splits into the *backward* and *forward* terms, given by

$$\mathcal{C}_{\text{BW}} \equiv \frac{1}{2E_1} \int d\text{PS}_{234} f_3 f_4 (1 \pm f_1) (1 \pm f_2) |M_{12\rightarrow 34}|^2 \quad (10)$$

and

$$\mathcal{C}_{\text{FW}} \equiv -\frac{1}{2E_1} \int d\text{PS}_{234} f_1 f_2 (1 \pm f_3) (1 \pm f_4) |M_{12\rightarrow 34}|^2. \quad (11)$$

The phase space integration of the collision integrals has been studied in the context of neutrino astrophysics for massless neutrinos in [17] and later for nonzero neutrino masses in [18]. Similar methods were also developed, e.g., in [15, 19–22]. Here we follow the strategy of [18] to reduce the fully general, initially nine-dimensional collision integrals down to four dimensions. The momentum dependence of the matrix elements prevents making further analytic simplifications.

Different from [18], we treat the forward and backward collision processes separately. This makes the numerical implementation more stable by avoiding the need to interpolate the unknown phase space distribution functions $f(p, t)$ in between the integration grid points. Full details of the reduction are given in Appendix A. The final result for the reduced backward term (10) is given by (A11) and for the forward term (11) by (A15).

B. Discretization

We solve Eq. (6) numerically by discretizing the momentum grid $\xi \rightarrow \xi_j$, $j = 1, \dots, N_\xi$, with uniform spacing in logarithmic scale. If the production processes spread over several orders of magnitude in temperature, this allows one to cover a sufficient range of momenta to reach the required accuracy. After discretization, the Boltzmann equation (6) becomes an initial value problem consisting of a coupled set of ordinary differential equations for $f_a(x, \xi_j) \equiv f_{aj}(x)$, over some temperature range x , that must be solved simultaneously¹ for each degree of freedom a and the momentum mode ξ_j ,

$$\frac{d}{dx} f_{aj}(x) = \alpha(x) \sum_{\text{coll.}} \mathcal{C}_{aj}(x). \quad (12)$$

¹This makes the problem highly vectorizable but not easily parallelizable. We use MATLAB and, in particular, its stiff ode15s routine.

Here $\alpha(x)$ is the prefactor given in the rhs of Eq. (6), and the sum runs over all collision terms that contribute to evolution of f_{aj} . Here we separated the degrees of freedom (labeled by a) from the discretized momentum variable. Indeed, each different particle species, and each helicity or polarization state within a species, in general has its own independent unknown distribution function, which the collision terms couple with each other. Some hierarchies between the interaction rates may allow simplifying the equation network, such as helicity equilibrium due to rapid helicity flips. This can be easily incorporated by imposing the degeneracies and introducing the corresponding averaged matrix elements. We will typically assume that initially $f_{aj} = 0$ for the dark sector particle distributions. This is justified when we start early enough in time, i.e., high enough temperature, and it allows us to track to which degree each species thermalizes before it decays or its distribution freezes.

As an example, on collision term discretization we show how the backward term (10) is implemented. Although the Boltzmann equation (6) is solved in dimensionless momentum ξ , the collision term reduction in Appendix A is done in terms of the physical momentum p . The physical momentum p_j corresponding to dimensionless momentum ξ_j at a given temperature $T = m_0/x$ is then obtained by inverting Eq. (5),

$$p_j(x) = \xi_j \frac{m_0}{x} \left(\frac{h_{\text{eff}}(\frac{m_0}{x})}{h_{\text{eff}}(T_0)} \right)^{1/3}. \quad (13)$$

Then, backward term (10) can be reduced to (A11), given in discretized form as

$$\mathcal{C}_{1j}^{\text{BW}}(x) = \frac{1}{16E_{1j}} \sum_k \sum_l \frac{\Delta p_k p_k^2 \Delta p_l p_l^2}{2\pi^2 E_{3k} 2\pi^2 E_{4l}} F_{1jkl}(x) \Lambda_{1jkl}^{\text{BW}}(x), \quad (14)$$

where the discretized backward phase space factor is

$$\Lambda_{1jkl}^{\text{BW}}(x) = f_{3k}(x) f_{4l}(x) [1 \pm f_{1j}(x)] [1 \pm f_{2jkl}(x)] \quad (15)$$

and $F_{1jkl}(x)$ is the angular integral over the matrix element squared defined in Eq. (A12). Here the superscripts 1, ..., 4 denote the particle species (in the sense described above) involved in the $12 \leftrightarrow 34$ process, $f_{ai}(x)$ is the value of distribution function of the particle “ a ” with momentum ξ_i at temperature $T = m_0/x$ and energy $E_{ai} = (p_i^2 + m_a^2)^{1/2}$. The label j refers to the species 1, whose collision term we are computing and it is not summed over. The momentum space matrix structure f_{2jkl} of species 2 follows from the four-momentum conservation.

It is essential to note that the all matrices $F_{1jkl}(x)$ can be precalculated and replaced by numerical fit functions for all relevant processes before solving the Boltzmann equations.

This fitting procedure can be done very accurately and it is pivotal for the efficiency of the numerical code. The general flow of our implementation then is as follows:

- (1) Define theory, whose masses and couplings may depend on temperature.
- (2) Create grids and define the entropy and energy degrees of freedom functions $h_{\text{eff}}, g_{\text{eff}}, g_*$.
- (3) Determine relevant interactions and compute their matrix elements and cross sections.
- (4) Precalculate the $F_{1ijk}(x_i)$ matrices for all species involved following Appendix A.
- (5) Define the initial conditions for all distributions involved in the network, Eq. (12).
- (6) Solve the Boltzmann system numerically using a suitable ordinary differential equation solver.

This formulation is generic enough to allow for dynamical changes that modify the parameters of the theory during the evolution, such as phase transitions. In the next sections, we show in detail the results of this implementation in simple hidden sector models connected with the SM via the Higgs portal.

III. FIRST BENCHMARK MODEL: THE SINGLET SCALAR EXTENSION

An extension of the SM by a real singlet scalar S , coupled with the SM Higgs doublet via the renormalizable operator $|H|^2 S^2$, the ‘‘Higgs portal,’’ provides a simple paradigm for a dark sector. Since its early introduction [23,24] this type of model building has started to gain more attention as benchmarks for experimental searches of particle dark matter [25–28]. Therefore, this model is a natural starting point for the tests of the computational method we have developed. Since the existing literature on this model is large and its phenomenology has been thoroughly exposed already, our discussion here will be brief; we will introduce only the necessary formulas and focus on the comparison of our approach with other approximate computation schemes. The singlet model is defined by the Lagrangian

$$\mathcal{L}_{\text{SSM}} = \frac{1}{2}(\partial_\mu S)^2 - V(S, H) + \mathcal{L}_{\text{SM}}, \quad (16)$$

where the scalar potential is given by

$$V(S, H) = -\mu_H^2 |H|^2 - \frac{1}{2} \mu_S^2 S^2 + \lambda_H |H|^4 + \frac{\lambda_S}{4} S^4 + \frac{\lambda_{\text{HS}}}{2} |H|^2 S^2, \quad (17)$$

and the gauge interactions of the Higgs doublet H are contained in \mathcal{L}_{SM} . The stability of the potential requires that the quartic couplings λ_S and λ_H of the singlet and the Higgs fields are positive, but the portal coupling could be negative, as the stability of the potential requires just that

$\lambda_{\text{HS}} > -2\sqrt{\lambda_S \lambda_H}$. However, here we will only consider positive values of λ_{HS} .

If the singlet scalar mass m_s is just below half of the Higgs mass $m_H = 125.25 \pm 0.17$ GeV [29], the Higgs mediated inelastic processes are resonantly enhanced but elastic processes maintaining the kinetic equilibrium are not, so one would expect the nonequilibrium effects to be relevant. Indeed, if the singlet S is required to constitute all of the dark matter and remain compatible with the current direct detection experiments, its mass is constrained to $m_s \in [56, 62.5]$ GeV [28], where precision computations are required to address the dark matter phenomenology [7].

In our current, fully momentum-dependent setup, solving the singlet scalar relic density in this region entails solving the following Boltzmann equation:

$$\partial_x f_s(\xi, x) = \mathcal{C}_{ss \leftrightarrow jj}^I + \mathcal{C}_{h \rightarrow ss}^I + \mathcal{C}_{sf \leftrightarrow sf}^E, \quad (18)$$

where s refers to the singlet, $j = \{\tau, c, b, t, h, W, Z\}$, and $f = \{\tau, c, b\}$. In this case we only need a dynamical equation for s , since all SM particles can be assumed to be in thermal equilibrium. To correctly include kinematics when the Higgs mediator in $\mathcal{C}_{ss \leftrightarrow jj}^I$ becomes on shell, we have included the on-shell Higgs decay $\mathcal{C}_{h \rightarrow ss}^I$ as a separate contribution and take all Higgs mediators in $\mathcal{C}_{ss \leftrightarrow jj}^I$ to be off shell as described in Appendix C. Resonant inelastic processes deplete and overpopulate specific momentum states, which tends to bring the distribution function f_s out of kinetic equilibrium. Elastic interactions, on the other hand, tend to restore the kinetic equilibrium, and if they are sufficiently fast, the standard thermal averaged treatment [4,5,28] suffices. However, the Higgs resonance is particularly sharp,² and one cannot *a priori* assume that the elastic processes can maintain the kinetic equilibrium to high accuracy. The issue has already been studied using momentum-dependent methods [6,7], however, using some approximations in the treatment of the collision integrals.

In left panel of Fig. 1, we show the contours of $\Omega_s h^2 = 0.1193$ in the singlet mass and the portal coupling plane using different approximations. The results from the full computation implemented in this work are shown by the red crosses, while the standard thermal averaged result is shown by the solid blue line. The yellow circles correspond to the calculation using the momentum-dependent, generalized relaxation approximation (GRA) method of Ref. [7], and finally, the calculation in Ref. [6], using truncated expansions for the elastic collision integrals, is shown by the black dots. The GRA calculation is similar to the one described in this paper, except for the numerically expensive elastic backward term $\mathcal{C}_{sf \leftarrow sf}$, given by (10).

² $\Gamma_h \simeq 4$ MeV, so the width of the resonance in the \sqrt{s} variable is $\sqrt{m_h \Gamma_h} \simeq 0.7$ GeV.

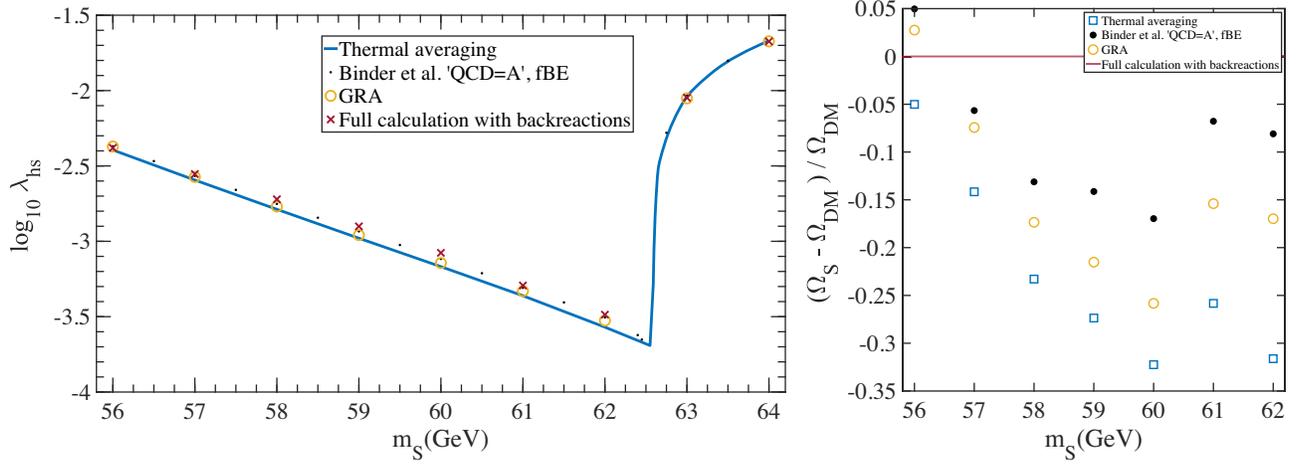


FIG. 1. Left: fixed relic densities $\Omega_S h^2 = 0.1193$ from three different methods: momentum-independent thermal averaging (solid blue line), momentum-dependent generalized relaxation approximation [7] (yellow circles), and complete momentum-dependent computation with full backreaction terms (red crosses). Black dots are the full Boltzmann solution [30] corresponding to the method introduced in Ref. [6], for the case “QCD = A.” Right: relative difference in Ωh^2 when λ_{HS} is fixed by the full calculation to yield $\Omega_S h^2 = 0.1193$ (red crosses on the left). Yellow circles correspond to relaxation approximation [7], blue squares to the momentum-independent method, and black dots correspond to the results of Ref. [6].

In the GRA method, this is treated in a simplifying approximation; for more details, see Ref. [7].

When displayed in the logarithmic scale, all calculations appear to roughly agree. Plotting on linear scale (right panel in Fig. 1) reveals the significance of the deviations. The cause for the difference between the full and the GRA calculations is seen in Fig. 2, which shows the elastic collision integrals computed exactly (red solid line) and in the GRA (yellow dashed line). At high temperatures, the GRA method works well, but around the freeze-out temperature $x \sim 10$ it starts to overestimate the elastic integral that enforces the kinetic equilibrium. Eventually the error becomes of order ~ 2 , but only well after the

freeze-out $x \sim 40$. This tendency was already noted in [7], and by construction the GRA scheme is not expected to work to a high precision for distributions that already are very close to thermal equilibrium. However, when one is close to equilibrium, the absolute magnitude of the error is already small, and GRA slightly improves on the thermal approximation. The results using a truncated expansion for the elastic collision integral from [6] are roughly comparable with the GRA.

Overall, we confirm that the effect of kinetic decoupling in the singlet model is not as dramatic as one might have initially guessed. All methods agree in the absence of resonant enhancement as they should. Even in the resonant

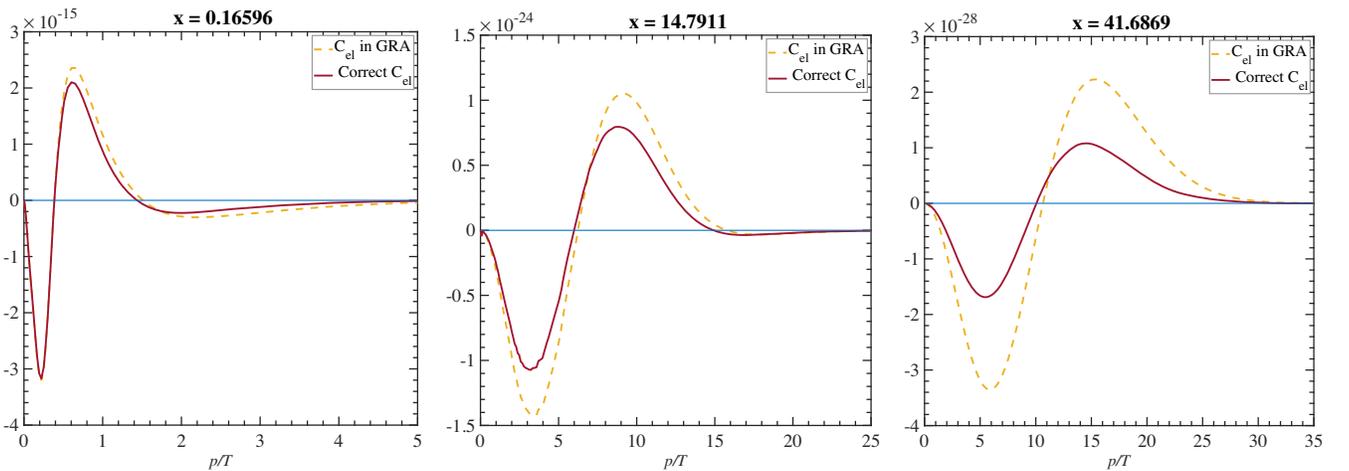


FIG. 2. Elastic integrals $(p^2/2\pi^2)C_{\text{el}}(p)$ evaluated at three different temperatures $x = m_S/T$. The dashed yellow line shows the elastic integral resulting from generalized relaxation approximation used in [7]. The red solid line shows the correct elastic collision integral containing the full backward collision term given by Eq. (A11).

region, by far the most important effect is to use the thermally averaged annihilation rate in the ZOPLW equations, first pointed out in [5]. The elastic interactions with the Standard Model particles are surprisingly efficient in keeping thermal equilibrium as observed in [7]. However, should a DM particle be identified in the resonance region, a precision calculation of its abundance requires a full momentum-dependent calculation with an exact collision integral.

Finally, let us note that the use of Maxwell-Boltzmann (MB) statistics in the *inelastic* collision integral causes about a 10% error [31] in final abundance. This is the case for all results in Fig. 1. We have checked that the corresponding error in *elastic* collision integrals only affects the final abundance by $\lesssim 1\%$.

To our knowledge, this is the first analysis of the singlet scalar in Higgs resonance with full elastic backreactions. This is undoubtedly due to the heavy numerical cost of the method. On average, we found that the computation times in this particular example (tested using a 6-core i7 laptop with 16 GB of RAM) scaled as follows in different approximations: thermal averaging runs took $\mathcal{O}(1)$ s, generalized relaxation approximation runs $\mathcal{O}(1-10)$ min, and the complete calculation with backreaction $\gtrsim 10$ h per (m_s, λ_{hs}) pair. Even faster and yet accurate methods of solving the ZOPLW equations exist [28] for the use of large scale parameter scans. Taking this hierarchy into account is obviously of paramount importance when making a choice of what method to use for a given problem.

IV. SECOND BENCHMARK MODEL: SINGLET SCALAR AND FERMION EXTENSION

As a second example, we consider an extension of the SM by a singlet scalar (S) and a singlet Dirac fermion (N). The Lagrangian of the model is

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + i\bar{N}\not{\partial}N + \frac{1}{2}(\partial_\mu S)^2 - V(S, H) + yS\bar{N}N, \quad (19)$$

where we denote the SM Higgs doublet by H . Its gauge interactions are contained in the SM Lagrangian \mathcal{L}_{SM} , while the potential terms are contained in the extended scalar potential $V(S, H)$ given by Eq. (17). In this model, the fermion N is a phenomenologically interesting candidate for cold DM [32,33]. We are interested especially in the keV mass range, where the nonequilibrium dynamics can be relevant [10,11], and the resulting nonthermal momentum distribution of DM may affect the formation of large scale structures. We focus on the question of whether highly nonthermal momentum distributions found in [10,11] survive when all elastic processes are included in the analysis.

We will assume a mass hierarchy $m_N \ll m_S$, with $m_N \sim \text{keV}$ and $m_S \sim \mathcal{O}(10-1000)$ GeV. The N -fermion mass gets a contribution from nonzero vacuum expectation value

(VEV) of the singlet scalar $m_N = \mu_N + y\langle S \rangle$. As the VEV can be quite large, we need to assume the Yukawa coupling to be tiny, $y \ll 1$, to keep m_N around keV scale. The vacuum structure is determined by the scalar sector of the theory. The field H is the usual weak doublet

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (20)$$

which has a VEV, denoted by v , along the neutral direction, $\phi_3 = v + \phi$. The VEV of the singlet field S is denoted by $\langle S \rangle \equiv w$ and we write $S = w + \sigma$. Inserting these parametrizations into Eq. (19), setting the field fluctuations to zero and extremizing the full scalar potential leads to

$$w \left(-\mu_S^2 + \frac{1}{2} \lambda_{\text{HS}} v^2 + \lambda_S w^2 \right) = 0, \quad (21)$$

$$v \left(-\mu_H^2 + \frac{1}{2} \lambda_{\text{HS}} w^2 + \lambda_H v^2 \right) = 0. \quad (22)$$

We use these conditions to eliminate μ_S^2 and μ_H^2 . This leads to the mass matrix for neutral scalars σ and ϕ ,

$$M^2 \equiv \begin{pmatrix} 2\lambda_H v^2 & \lambda_{\text{HS}} v w \\ \lambda_{\text{HS}} v w & 2\lambda_S w^2 \end{pmatrix}, \quad (23)$$

which is diagonalized by the transformation to the mass eigenbasis. We denote the mass eigenstates by h_1 and h_2 , so the explicit relation is

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \phi \\ \sigma \end{pmatrix}. \quad (24)$$

We identify h_1 with the SM Higgs field and h_2 is a heavier scalar. Consistency with LHC data on Higgs couplings then requires $\sin \theta \lesssim 0.23$ [34,35]. We therefore set $m_2 > m_1 = 125.25$ GeV and consider the physical masses m_1 and m_2 to be input parameters. We then solve the couplings λ_H and λ_S and the mixing angle θ in terms of the physical masses, the vacuum expectation values, and the portal coupling λ_{HS} as

$$\lambda_H = \frac{m_1^2 \cos^2 \theta + m_2^2 \sin^2 \theta}{2v^2}, \quad (25)$$

$$\lambda_S = \frac{m_2^2 \cos^2 \theta + m_1^2 \sin^2 \theta}{2w^2}, \quad (26)$$

$$\sin(2\theta) = \frac{2\lambda_{\text{HS}} v w}{(m_2^2 - m_1^2)}. \quad (27)$$

Requiring $\sin(2\theta)$ to be positive implies that $0 \leq \lambda_{\text{HS}} \leq \frac{m_2^2 - m_1^2}{2vw} \equiv \lambda_{\text{HS}}^{\text{max}}$.

The Feynman rules following from the Lagrangian (19) are tabulated in Appendix B. In the special case of $w = 0$, the treatment is more straightforward, as the fields σ and ϕ are directly the mass eigenstates of the mass matrix. Without going into further details, we simply note that the Feynman rules of Appendix B can be directly applied also in this case by letting $\theta, w \rightarrow 0$. In the limit of vanishing Yukawa coupling and singlet scalar VEV, $y, w \rightarrow 0$, the model reduces to the singlet scalar model from previous section.

With a slight abuse of notation, we denote the mass eigenstates by ϕ and σ , as this allows us to include the cases $w \neq 0$ and $w = 0$ simultaneously. Then we can summarize the above construction as follows: we have taken the masses $m_1 \equiv m_\phi$, $m_2 \equiv m_\sigma$, the portal coupling λ_{HS} , and the vacuum expectation values w and v as the input parameters, and express other Lagrangian parameters in the scalar sector in terms of these. Furthermore, we fix $v = 246$ GeV. Thus, the free parameters in this theory are $\{m_\sigma, \lambda_{\text{HS}}, w, y, m_N\}$.

A. DM production processes and coupled Boltzmann system

The Lagrangian (19) allows for various production processes for the N and σ . Processes of order $\mathcal{O}(y^2)$ are negligible and the relevant contributions under our assumptions are summarized in Table I. Because of non-zero Yukawa coupling and assumed mass hierarchy, eventually all produced σ scalars will decay into N fermions, which remains as a stable relic. Production of σ scalars is determined by the portal coupling λ_{HS} . Direct production of N fermions from a SM heat bath is allowed by a nonzero mixing angle $\sin\theta$ between the scalars, but remains subdominant for the allowed small mixing angles. Therefore, the production of N fermions proceeds mostly via σ -scalar decays, which is itself produced from a SM heat bath and whose number density can freeze (either via freeze-in or freeze-out mechanism) before it fully decays.

TABLE I. Relevant tree-level production processes for singlet fermion N and singlet scalar σ in the model (19) assuming $y \ll 1$. Each s , t , and u channel reaction can be mediated via both scalar fields: singlet σ and SM Higgs ϕ . Here $V = \{W, Z\}$ labels vector bosons and $f = \{\tau, c, b, t\}$ labels SM fermions.

Always open	Open if $w > 0$
$\sigma\sigma \leftrightarrow \phi\phi$	$\phi \rightarrow NN$
$\sigma\sigma \leftrightarrow VV$	$N\sigma \leftrightarrow N\sigma$
$\sigma\sigma \leftrightarrow ff$	$N\phi \leftrightarrow N\phi$
$\phi \rightarrow \sigma\sigma$	$Nf \leftrightarrow Nf$
$\sigma \rightarrow NN$	
$\sigma f \leftrightarrow \sigma f$	
$\sigma\sigma \leftrightarrow \sigma\sigma$	
$N\sigma \leftrightarrow N\phi$	

To obtain the momentum distribution function for σ scalar and N fermion we must solve the following set of coupled Boltzmann equations:

$$\partial_x f_\sigma(\xi, x) = \mathcal{C}_{\sigma\sigma \leftrightarrow jj}^I + \mathcal{C}_{\phi \rightarrow \sigma\sigma}^I + \mathcal{C}_{\sigma \rightarrow NN}^I + \mathcal{C}_{N\sigma \leftrightarrow N\phi}^I + \mathcal{C}_{\sigma f \leftrightarrow \sigma f}^E + \mathcal{C}_{N\sigma \leftrightarrow N\sigma}^E, \quad (28)$$

$$\partial_x f_N(\xi, x) = \mathcal{C}_{\sigma \rightarrow NN}^I + \mathcal{C}_{\phi \rightarrow NN}^I + \mathcal{C}_{N\sigma \leftrightarrow N\phi}^E + \mathcal{C}_{N\sigma \leftrightarrow N\sigma}^E + \mathcal{C}_{N\phi \leftrightarrow N\phi}^E + \mathcal{C}_{Nf \leftrightarrow Nf}^E, \quad (29)$$

where again the SM states are denoted as $j = \{\tau, c, b, t, \phi, W, Z\}$ and $f = \{\tau, c, b\}$. The form of this equation shows one obvious fact about solving the momentum-dependent kinetic equations: most of the work involved goes to definition and computation of the various collision integrals. The ϕ and σ propagators in the inelastic $2 \rightarrow 2$ collision integrals $\mathcal{C}_{\sigma\sigma \leftrightarrow jj}^I$ are taken to be off shell, as the on-shell contributions are already included separately in $\mathcal{C}_{\phi \rightarrow \sigma\sigma}^I$ and $\mathcal{C}_{\sigma \rightarrow NN}^I$. There are several different suggestions in literature as to how this real intermediate state (RIS) subtraction should be done, e.g., [36,37]. Here we are following the treatment of [38,39]; see Appendix C for more details and discussion.

Different from previous treatments, we have also accounted for the three- and four-body final states from virtual boson decays using methods described in [28], as well as the one-loop corrections for quarks in the $\sigma\sigma \leftrightarrow jj$ channel. Accounting for virtual boson decays and QCD one-loop corrections describe the SM states more accurately and slightly increase the $\mathcal{C}_{\sigma\sigma \leftrightarrow jj}^I$ contributions in Eq. (29). This is good to keep in mind when comparing our results to, e.g., Ref. [11], as in the case of σ freezing out this slight increase causes the σ to follow the SM heat bath a bit longer and slightly suppresses the final fermion distribution.

B. Results and discussion when $w = 0$

We first set the VEV of the singlet scalar to zero, so that the scalars do not mix. This leaves us with processes on the left column of Table I. This setting is equivalent to the one studied in Ref. [11], except that we have included the elastic processes $\sigma f \leftrightarrow \sigma f$, $\sigma\sigma \leftrightarrow \sigma\sigma$, and $N\sigma \leftrightarrow N\phi$, which tend to suppress the nonthermal component in the momentum distribution of N fermion. The two first processes can also lengthen the freeze-out time of σ field, thus allowing it to be Boltzmann suppressed more before it freezes out and decays, which can reduce the late time production of N fermions. The two-peaked nonthermal momentum distribution found in [11] results from N being produced at two separate temperature scales (see Ref. [12] for a comprehensive study). Hence, reducing the production at either temperature scale could prevent the momentum distribution from forming the double peak structure. The last two processes tend to restore the kinetic equilibrium by reducing the nonthermal component momentum

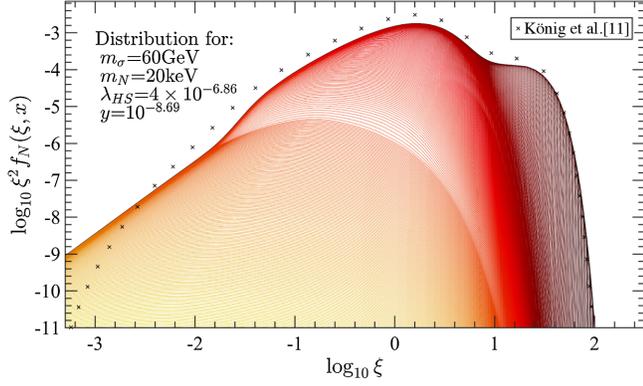


FIG. 3. Momentum distribution functions of singlet fermion obtained from our implementation versus the results in Fig. 5 of [11] (black crosses). Note that our convention to λ_{HS} differs from [11] by a factor of 4. Darker colors refer to later times and the black solid curve is the final frozen-in form. The low-end momentum tail is produced at temperatures much higher than the electroweak phase transition temperature, $T \gg T_{EWPT} = 150$ GeV. The temperature grid range is $\log_{10}(x) \in [-3.5, 3]$ and we used $m_0 = T_0 = m_h$ as the reference scales in (5) and (6).

distribution. In practice, we find their effect to be negligible.

Figure 3 shows our solution for the momentum distribution function of the fermion N for parameter values $\{m_\sigma = 60 \text{ GeV}, m_N = 20 \text{ keV}, \lambda_{HS} = 4 \times 10^{-6.86}, y = 10^{-8.69}, w = 0\}$. These values are chosen to correspond to the case presented on the right-hand panels of Fig. 5 in Ref. [11] and we have included their final result as black crosses in our Fig. 3. Clearly the elastic processes are not sufficiently fast to suppress the formation of the nonthermal momentum distribution. The reason is quite clear *a posteriori*: The second peak at high momenta forms only when σ production from the SM heat bath has already stopped and the remaining scalars decay into N fermions. At this point, the SM Higgs field is also heavily Boltzmann suppressed. The elastic channel $N\sigma \leftrightarrow N\phi$, which would most effectively thermalize the momentum distribution of N fermions, is therefore suppressed due to absence of both scalars in the thermal bath.

We have also checked the other light ($m_\sigma < m_h/2$) and heavy ($m_\sigma > m_h/2$) scalar (corresponding to feebly and weakly interacting massive particles) scenarios as discussed in [11] and found that our results agree with theirs to the same extent as in the above example. In summary, we find that the model presented in [11] is indeed inherently nonthermal in parts of its parameter space when assumptions $y \ll 1$ and $w = 0$ hold. Our numerical results for the distribution $f_N(\xi, x)$ differ noticeably from [11], but the difference does not come from elastic interactions, but from our more accurate evaluation of the inelastic interaction rates.

C. Results and discussion when $w > 0$

We now let the singlet obtain a nonzero VEV $w > 0$. The nonzero mixing angle induced by the broken symmetry is

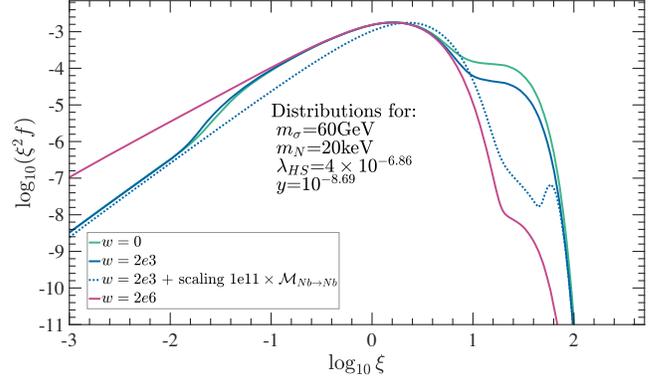


FIG. 4. Frozen-in momentum distributions for the fermion N for varying w . The green solid line corresponds to the case in Fig. 3, and the blue dotted line shows an example of forced thermalization when the elastic rates are increased by hand, which we show only as a proof of concept to show that the kinematics work as intended and the system then tends towards thermal distribution.

given by Eq. (27). It increases the σ and N particle production rates and opens new elastic channels for the N field. All relevant reaction channels are given in Table I, where also the channels on the right column now contribute.

In Fig. 4 we compare our results in the $w > 0$ case with the previous $w = 0$ case (shown with the green curve) to see the effect of additional channels and the increased interaction rates on the nonthermal double peak structure of the momentum distribution of N fermions. We have chosen conservative values for the singlet VEV: $w = 2 \times 10^3$ GeV (blue curve) and $w = 2 \times 10^6$ GeV (red curve). In both cases, we find a strong suppression on the amplitude of the nonthermal second peak, which strongly increases for increasing w . However, this is again not due to elastic processes being effective in redistributing the quanta, but due to the fact that the inelastic interactions between the σ scalar and SM heat bath strongly increase for increasing values of w , as illustrated in Fig. 5. As a result, the number density of σ particles becomes more and more Boltzmann suppressed before they finally decay to N fermions. This then suppresses or removes the high-momentum out-of-equilibrium tail from the N distribution. The mechanism is clearly illustrated in Fig. 6, where we show the integrated yields of the singlet scalar fields (dotted lines) as a function of m_0/T for the same choices of parameters as in Fig. 4. The effect on integrated neutrino distributions (dashed lines) is much smaller than to their momentum dependence, but it shows the right tendency as function of the σ -field abundance.

We found that the elastic interactions had negligible effect on the N distribution in comparison with the Boltzmann suppression discussed above. Increasing the VEV of the singlet does result in stronger elastic rates for the fermion N , but they still fail to restore the thermal equilibrium. This is partly because the having $w > 0$ also

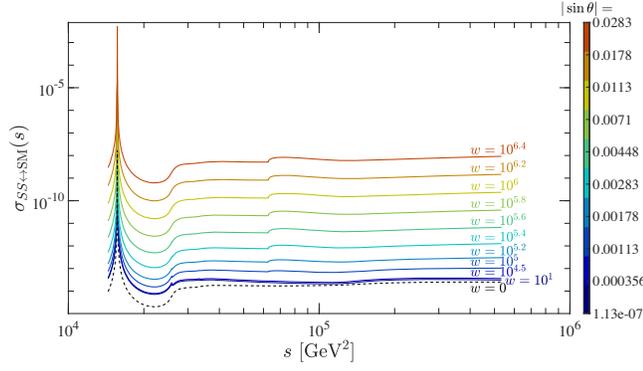


FIG. 5. Cross section between the singlet scalars $S = w + \sigma$ and SM particles plotted over a representative range of incoming energies with varying singlet VEV w (mixing angle $|\sin \theta|$). Other parameters are $\{m_S = 60 \text{ GeV}, \lambda_{HS} = 4 \times 10^{-6.86}\}$.

increases the overall production of N fermions via inelastic and decay channels, which is the root cause for the non-thermal distributions. We then find that even in the $w > 0$ case the elastic processes are never sufficiently fast to thermalize the fermion momentum distribution and the model (19) is inherently nonthermal when $y \ll 1$ and $m_N \ll m_S$. However, if a phase transition takes place in the singlet sector before the singlet field freezes out, the predictions for the amplitude of the out-of-equilibrium component in N distribution change dramatically for a given set of parameters, in comparison to the $w = 0$ case studied in [11]. However, this change is not due to increased elastic interactions, but due to change in inelastic rates, and to discover this effect the high-accuracy Boltzmann codes developed in this work are, in fact, not necessary.

The inefficacy of the elastic rates in the examples studied above results from the smallness of the adopted values for the y coupling and this does not imply that elastic interactions were irrelevant in general. Instead of constructing a model just to make the elastic rates important, as a

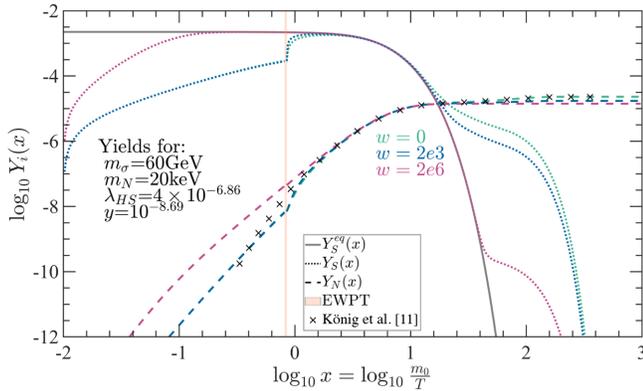


FIG. 6. Yields $Y \equiv \frac{n_i(x)}{s(x)}$ for singlet scalar S (dotted lines) and fermion N (dashed lines). $m_0 = m_H$ and colors show results from different choices of the singlet VEV $w = \{0, 2 \times 10^3, 2 \times 10^6\} \text{ GeV}$. Crosses show data from Fig. 5 of [11] corresponding to $w = 0$ case.

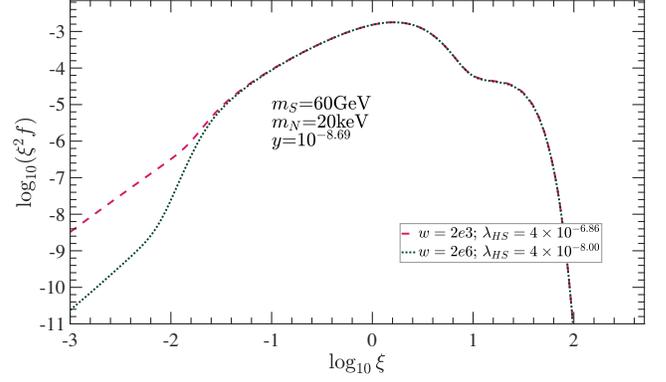


FIG. 7. The final frozen-in momentum distribution of the fermion N depends on the parameters $\{w, \lambda_{HS}\}$ so that increasing one can be countered by decreasing the other.

proof of concept, we simply increase the elastic process $Nf \leftrightarrow Nf$ by hand to see how large it must be for a significant thermalization by elastic channels to take place. For $\{m_\sigma = 60 \text{ GeV}, m_N = 20 \text{ keV}, \lambda_{HS} = 4 \times 10^{-6.86}, y = 10^{-8.69}, w = 2 \times 10^3\}$ we scaled the matrix element $\mathcal{M}_{Nf \rightarrow Nf} \rightarrow 10^{11} \times \mathcal{M}_{Nf \rightarrow Nf}$. The nonscaled result is shown as a solid blue curve and the result after scaling as a dotted blue curve in Fig. 4. Thus, in this case, at the level of cross sections, the elastic rates are roughly $\sim 10^{-22}$ times too small to significantly thermalize the system.

Let us finally note that, while the introduction of the phase transition changes the predictions significantly for a given portal coupling, the same out-of-equilibrium distribution can be obtained in the spontaneously broken case for a different portal coupling. That is, there are degenerate subspaces in the $(\lambda_{HS}, y; w)$ parameter space, where the same nonthermal momentum distribution can be obtained. In particular, for a fixed Yukawa y , essentially only the interaction rate between the singlet scalar and SM heat bath is important in determining the degree to which the momentum distribution becomes nonthermal. As this rate is determined by a combination of λ_{HS} and w , we can always find such a (w', λ'_{HS}) pair that the original nonthermal behavior for $(\lambda_{HS}, y; w = 0)$ parameters is obtained with $(\lambda'_{HS}, y; w' > 0)$. This is shown in Fig. 7, where we plot the distributions with $y = 10^{-8.69}$ for the cases $\{w = 2 \times 10^3 \text{ GeV}, \lambda_{HS} = 4 \times 10^{-6.86}\}$ and $\{w = 2 \times 10^6 \text{ GeV}, \lambda_{HS} = 4 \times 10^{-8}\}$, respectively. In each case, we find exactly the same momentum distribution, apart from small deviations in the very small momenta.

V. CONCLUSIONS

We have presented a computational method that is generally applicable for solving the coupled set of Boltzmann equations for phase space distribution functions of cosmic relics. Similar techniques have been developed earlier in the context of the neutrino astrophysics [13–16], but until now they have not been carefully adapted to be

used in the dark matter abundance calculations. For earlier implementations that use different levels of approximations for the collision integrals, see Refs. [6,7]. One of the main results is the compilation of these methods into a simple and generic form that provides a suitable starting point for their wider utilization in cases, where the standard thermal averaging methods fail and more accurate predictions in DM models are needed.

We demonstrated our method in two models known to be sensitive to the momentum dependency of the phase space distribution. Our first example concerned the freeze-out of a singlet scalar dark matter coupled to SM via the Higgs portal near Higgs resonance. In this case, the kinetic equilibrium approximation required by the usual ZOPLW equation cannot be assumed *a priori*. We performed full momentum calculations, comparing our results to the ZOPLW approximation and two earlier momentum-dependent calculations, where further approximations of the form of the elastic collision integrals were made.

We found that the approximation methods of Refs. [6,7] are in good agreement with our full results. Both methods present some improvement over the ZOPLW results even very near equilibrium, although they both slightly (the former a little more) overestimate the elastic rates.

As another example, we considered the model presented in [10,11]. Using our methodology we extended their calculations to include the elastic processes and possible mixing between the two scalar states as a result of a phase transition in the singlet sector. We found that this could significantly alter the predicted size of the nonequilibrium distribution, the more so the larger the VEV of the singlet field. However, this difference was mainly due to changes in the inelastic rates, caused by the phase transition. Elastic rates turned out to be inefficient and even when they were included, the momentum distribution of the singlet fermion remains inherently nonthermal. Overall, we find that, in a vast majority of cases, the momentum averaged methods work surprisingly well.

ACKNOWLEDGMENTS

We thank T. Bringmann for correspondence and sharing data related to Fig. 1. This work was supported by the Academy of Finland Grants No. 310130, No. 342777, and No. 318319. K. A.-M. is funded by the doctoral program in Particle Physics and Universe Sciences in University of Helsinki.

Note added.—Recently, a paper presenting similar methods, also based on [18], appeared in [40]. Our results agree qualitatively with theirs.

APPENDIX A: COLLISION TERM KINEMATICS

Here we give details of the integration of collision terms following closely the treatment of Ref. [18]. Our goal is to

clearly isolate the general structures in the collision terms that can be pre-evaluated prior to the solution of the Boltzmann equation.

1. 2-2 Scatterings

a. Backward term

Different from [18], we treat the forward and backward collision integrals separately. We start from the backward term. In Eq. (10), we use the four-momentum conservation to integrate over $d^3 p_2$, which leaves out a delta function $\delta(p_2^2 - m_2^2)$, where

$$p_2^2 - m_2^2 = Q + 2(p_3 \cdot p_4 - p_1 \cdot p_3 - p_1 \cdot p_4), \quad (\text{A1})$$

where $Q = m_1^2 - m_2^2 + m_3^2 + m_4^2$. Measuring angles relative to the external momentum p_1 , we define the coordinate system such that

$$\begin{aligned} p_1 &= p_1(0, 0, 1), \\ p_3 &= p_3(0, \sin \theta, \cos \theta), \\ p_4 &= p_4(\sin \alpha \sin \beta, \sin \alpha \cos \beta, \cos \alpha). \end{aligned} \quad (\text{A2})$$

The remaining integrals can now be readily reduced to

$$\begin{aligned} d^3 p_3 &= 2\pi p_3^2 dp_3 d\cos \theta, \\ d^3 p_4 &= p_4^2 dp_4 d\cos \alpha d\beta. \end{aligned} \quad (\text{A3})$$

The integral over β can be done [18] using

$$\delta(p_2^2 - m_2^2) = \delta(f(\beta)) = \sum_i |f'(\beta_i)|^{-1} \delta(\beta - \beta_i), \quad (\text{A4})$$

where $f'(\beta) = 2p_3 p_4 \sin \alpha \sin \theta \sin \beta$ and β_i are the roots of $f(\beta) = 0$. It is easy to see that the δ function is symmetric in β , so we can multiply by 2 and integrate over $[0, \pi]$. Thus, $\cos \beta_i \in [-1, 1]$ which demands $\sin^2 \beta_i \geq 0$ for the root to exist within the integration interval. This can be equivalently written as

$$f'(\beta_i)^2 = (2p_3 p_4 \sin \alpha \sin \theta \sin \beta_i)^2 \geq 0, \quad (\text{A5})$$

so that

$$\int_0^{2\pi} d\beta \delta(f(\beta)) = 2|f'(\beta_i)|^{-1} \Theta(|f'(\beta_i)|^2). \quad (\text{A6})$$

Using the equation $f(\beta_i) = 0$, one can rewrite the derivative as

$$|f'(\beta_i)| = \sqrt{a \cos^2 \alpha + b \cos \alpha + c}, \quad (\text{A7})$$

where, again paralleling the notation of [17,18],

$$a = p_4^2(-4\kappa + 8\epsilon), \quad (\text{A8})$$

$$b = p_4(-p_1 + \epsilon/p_1)(8\gamma + 4Q + 8\epsilon), \quad (\text{A9})$$

$$c = 4p_3^2 p_4^2 \sin^2 \theta - (2(\gamma + \epsilon) + Q)^2, \quad (\text{A10})$$

with $\gamma \equiv E_3 E_4 - E_1 E_3 - E_1 E_4$ and $\epsilon \equiv p_1 p_3 \cos \theta$ and $\kappa \equiv p_1^2 + p_3^2$. Then, after a trivial $d\mu$ integral, the source term becomes

$$\begin{aligned} C_{\text{BW}} = & \frac{2}{(2\pi)^4} \frac{1}{2E_1} \iint \frac{dp_3 p_3^2}{2E_3} \frac{dp_4 p_4^2}{2E_4} F(p_1, p_3, p_4) \\ & \times f_3 f_4 [1 \pm f_1][1 \pm f_2], \end{aligned} \quad (\text{A11})$$

where $f_i \equiv f_i(E_i)$ with $E_2 = E_4 + E_3 - E_1$, and the function $F(p_1, p_3, p_4)$ contains the squared matrix element integrated over the angles and fixing kinematics,

$$\begin{aligned} F(p_1, p_3, p_4) = & \int_{-1}^1 d \cos \theta \int_{\max(-1, z^+)}^{\min(1, z^-)} dz \\ & \times \frac{|M(s, t)|^2}{\sqrt{a(z - z^-)(z - z^+)}} \Theta(b^2 - 4ac), \end{aligned} \quad (\text{A12})$$

where we set $z = \cos \alpha$ and $z^\pm \equiv (-b \pm \sqrt{b^2 - 4ac})/2a$. This function contains all process specific dynamical information and since it is independent of the distribution functions, it can be computed and fitted before solving the Boltzmann equations. In general, the matrix element squared is a function of $s = (p_3 + p_4)^2$ and $t = (p_1 - p_3)^2$, which depend on the angles and momenta as follows:

$$\begin{aligned} t &= m_1^2 + m_3^2 - 2E_1 E_3 + 2p_1 p_3 \cos \theta, \\ s &= 2E_1 E_3 + 2E_1 E_4 - 2p_1 p_3 \cos \theta \\ &\quad - 2p_1 p_4 \cos \alpha + m_2^2 - m_1^2. \end{aligned} \quad (\text{A13})$$

In the particular case where the matrix element in (A12) does not depend on $\cos \alpha$ (a pure t -channel process), the $d \cos$ integral can be reduced to a one-dimensional integral,

$$F(p_1, p_3, p_4) = \int_{-1}^1 d \cos \theta |M(t)|^2 \frac{\pi}{\sqrt{-a}} \Theta(b^2 - 4ac). \quad (\text{A14})$$

b. Forward term

In the forward term, given by Eq. (11), we integrate first over $d^3 p_4$, which leaves us with the delta function $\delta(p_4^2 - m_4^2)$. Paralleling the backward term reduction, we eventually obtain

$$\begin{aligned} C_{\text{FW}} = & -\frac{2}{(2\pi)^4} \frac{1}{2E_1} \iint \frac{p_2^2 dp_2}{2E_2} \frac{p_3^2 dp_3}{2E_3} F'(p_1, p_2, p_3) \\ & \times f_1 f_2 [1 \pm f_3][1 \pm f_4], \end{aligned} \quad (\text{A15})$$

where f_4 is evaluated at $E_4 = E_1 + E_2 - E_3$ and $F'(p_1, p_2, p_3)$ has an identical expression to the right-hand side of (A12), where one replaces everywhere $(a, b, c) \rightarrow (a', b', c')$ with

$$a' = p_2^2(-4\kappa' + 8\epsilon'), \quad (\text{A16})$$

$$b' = p_2(p_1 - \epsilon'/p_1)(8\gamma' + 4Q' + 8\epsilon'), \quad (\text{A17})$$

$$c' = 4p_2^2 p_3^2 \sin^2 \theta - (2(\gamma' + \epsilon') + Q')^2, \quad (\text{A18})$$

where $Q' = m_1^2 + m_2^2 + m_3^2 - m_4^2$ and, moreover, $\gamma' = E_1 E_2 - E_1 E_3 - E_2 E_3$ and $\epsilon' = p_1 p_3 \cos \theta$ and $\kappa' = p_1^2 + p_3^2$.

Now the Mandelstam variables must be written as $s = (p_1 + p_2)^2$ and $t = (p_1 - p_3)^2$ so that

$$\begin{aligned} s &= m_1^2 + m_2^2 + 2E_1 E_2 - 2p_1 p_2 \cos \alpha, \\ t &= m_1^2 + m_3^2 - 2E_1 E_3 + 2p_1 p_3 \cos \theta. \end{aligned} \quad (\text{A19})$$

If the matrix element is again independent of $\cos \alpha$, the result (A14) applies also as such, after replacing $(a, b, c) \rightarrow (a', b', c')$. Note that the forward and backward collision integrals (A11) and (A15) are valid for general Bose-Einstein and Fermi-Dirac statistics.

2. Special case: Inelastic 2-2 scattering towards/from equilibrium in MB statistics

As stated in the main text, we are using the MB statistics throughout. This is not needed for our computation of the elastic rates, but to keep the computation time associated with the large number of inelastic interactions with the SM states manageable. In this case, the final states are in equilibrium, and we can reduce the nine-dimensional integral down to one dimension. Overall, using the MB statistics amounts to about 10% error in the overall magnitude of the elastic collision integral [31], which should have but a very small effect on the final abundance. Indeed, we checked that scaling the elastic collision terms by a factor 0.9–1.1 caused only a 0.7% change in the final abundance. Then, working under the assumption $f \ll 1$ and enforcing the detailed balance, we can write the phase space factor (9) as

$$\Lambda(f_1, f_2) = f_1^{\text{eq}} f_2^{\text{eq}} - f_1 f_2. \quad (\text{A20})$$

The collision can then be written as

$$C_1(p_1, t) = \frac{1}{2E_1} \int d^3 \tilde{p}_2 \Lambda(f_1, f_2) \sum_n F_1(s) \sigma_1^{(n)}(s), \quad (\text{A21})$$

where n runs over different equilibrium states, $\sigma_1^{(n)}(s)$ is the corresponding cross section, and $F_1(s)$ is the Lorentz invariant flux factor,

$$\begin{aligned} F_1(s) &= 4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}, \\ &= 2\lambda^{1/2}(s, m_1^2, m_2^2) \equiv 4E_1 E_2 v_{M\phi 1}, \end{aligned} \quad (\text{A22})$$

where $\lambda(a, b, c) = (a - b - c)^2 - 4bc$ is the Källén kinetic function. As explained in [7], one can reduce the integral over $d^3\tilde{p}_2$ to a single integral over s ,

$$C_I(p_1, t) = \frac{1}{2\pi^2} \int_0^\infty dp_2 p_2^2 \Lambda(f_1, f_2) \sum_n [F_1 \sigma_1^{(n)}], \quad (\text{A23})$$

where

$$[\mathcal{O}](p_1, p_2) \equiv \frac{1}{16p_1 p_2 E_1 E_2} \int_{s_-}^{s_+} ds \mathcal{O}(s), \quad (\text{A24})$$

with $s_\pm = m_1^2 + m_2^2 + 2E_1 E_2 \pm 2p_1 p_2$. Again, this function can be evaluated and fitted for each collision channel before one attempts to solve the dynamical Boltzmann equations, which gives a dramatic boost in numerical efficiency.

3. 1-2 decays and fusions

A similar reduction that was carried out above for the $2 \leftrightarrow 2$ scatterings, can be performed for the $1 \leftrightarrow 2$ processes. We shall assume that either the decaying particle or the decay products are in thermal equilibrium.

a. Decay from (fusion to) equilibrium $f_A^{\text{eq}} \leftrightarrow f_1 f_2$

Assume we are tracking the species labeled as “1,” while the “2” species is arbitrary and “A” species follows the equilibrium. Then,

$$C_{A-12}^{\text{eq}}(p_1, t) = \frac{1}{2E_1} \int d\text{PS}_{A2} \Lambda(f_A^{\text{eq}}, f_1, f_2) |M_{A-12}|^2, \quad (\text{A25})$$

where the phase space integration is denoted by

$$d\text{PS}_{A2} = d^3\tilde{p}_A d^3\tilde{p}_2 (2\pi)^4 \delta^{(4)}(p_A - p_1 - p_2), \quad (\text{A26})$$

and the distribution factor Λ is given by

$$\begin{aligned} \Lambda &= f_A^{\text{eq}} [1 \pm f_1] [1 \pm f_2] - f_1 f_2 [1 \pm f_A^{\text{eq}}], \\ &\rightarrow f_1^{\text{eq}} f_2^{\text{eq}} - f_1 f_2, \end{aligned} \quad (\text{A27})$$

where we again assumed that $f \ll 1$ and applied the detailed balance for the equilibrium state. A similar procedure as in the previous section eventually gives

$$C_{A-12}^{\text{eq}}(p_1, t) = \frac{|M_{A-12}|^2}{16\pi E_1 p_1} (f_1^{\text{eq}} \chi[f_2^{\text{eq}}] - f_1 \chi[f_2]), \quad (\text{A28})$$

where we used the fact that the matrix element for the decay process is a constant and defined

$$\chi[f] \equiv \int_0^\infty dp_2 \frac{p_2}{E_2} f(p_2) \Theta(1 - |\cos \theta_0|), \quad (\text{A29})$$

with

$$\cos \theta_0 = \frac{|m_1^2 + m_2^2 - m_A^2 + 2E_1 E_2|}{2p_1 p_2}. \quad (\text{A30})$$

In this case, one only needs to compute the matrix element as a function of the masses of particles involved.

b. Decay to (fusion from) equilibrium $f_1 \leftrightarrow f_A^{\text{eq}} f_B^{\text{eq}}$

Now assume we track the species 1, while the arbitrary species A follows the equilibrium. After similar steps as above, we get an even simpler expression

$$C_{1 \rightarrow AB}^{\text{eq}}(p_1, t) = \frac{|M_{1-AB}|^2}{16\pi E_1} (f_1^{\text{eq}} - f_1) v \Theta(v^2), \quad (\text{A31})$$

where $v = \lambda^{1/2}(m_1^2, m_A^2, m_B^2)/m_1^2$ and where we used the fact that the matrix element squared (in tree level) is always a constant.

APPENDIX B: FEYNMAN RULES

Here we summarize the Feynman rules that follow from the Lagrangian in Eq. (19) for the new/modified interaction vertices in the model with a singlet scalar $S = \sigma + w$ and a singlet fermion N ,

$$\begin{aligned} \lambda_{\phi\phi\phi\phi} &= -6i\lambda_{\text{H}} c_\theta^4 + 6s_\theta^2 (\lambda_{\text{HS}} c_\theta^2 + \lambda_{\text{S}} s_\theta^2), \\ \lambda_{\sigma\sigma\sigma\sigma} &= -6i\lambda_{\text{S}} c_\theta^4 + 6s_\theta^2 (\lambda_{\text{HS}} c_\theta^2 + \lambda_{\text{H}} s_\theta^2), \\ \lambda_{\phi\phi\sigma\sigma} &= -(i/4) (\lambda_{\text{HS}} [1 + 3c_{4\theta}] + (3\lambda_{\text{H}} + 3\lambda_{\text{S}}) [1 - c_{4\theta}]), \\ \lambda_{\phi\phi\phi\phi} &= -6i\lambda_{\text{H}} v c_\theta^3 - 3s_\theta (\lambda_{\text{HS}} w c_\theta^2 - \lambda_{\text{HS}} v s_\theta c_\theta + 2\lambda_{\text{S}} w s_\theta^2), \\ \lambda_{\sigma\sigma\sigma} &= -6i\lambda_{\text{S}} w c_\theta^3 + 3s_\theta (\lambda_{\text{HS}} v c_\theta^2 + \lambda_{\text{HS}} w s_\theta c_\theta + 2\lambda_{\text{H}} v s_\theta^2), \\ \lambda_{\phi\phi\sigma} &= -i\lambda_{\text{HS}} v c_\theta^3 - s_\theta (\lambda_{\text{C}2} w c_\theta^2 - \lambda_{\text{C}1} v s_\theta c_\theta + \lambda_{\text{HS}} w s_\theta^2), \\ \lambda_{\phi\sigma\sigma} &= -i\lambda_{\text{HS}} w c_\theta^3 + s_\theta (\lambda_{\text{C}1} v c_\theta^2 + \lambda_{\text{C}2} w s_\theta c_\theta + \lambda_{\text{HS}} v s_\theta^2), \\ \lambda_{\phi V V} &= -2i(M_{\tilde{V}}^2/v) s_\theta g^{\mu\nu}, \\ \lambda_{\sigma V V} &= -2i(M_{\tilde{V}}^2/v) c_\theta g^{\mu\nu}, \\ \lambda_{\phi f f} &= -i(m_f/v) c_\theta, \\ \lambda_{\sigma f f} &= -i(m_f/v) s_\theta, \\ \lambda_{\phi N N} &= -i(y/2) s_\theta, \\ \lambda_{\sigma N N} &= i(y/2) c_\theta. \end{aligned} \quad (\text{B1})$$

Here ϕ denotes the SM Higgs field with VEV v , and V and f denote SM vector bosons and fermions; we used the following shorthand notations: $s_\theta \equiv \sin \theta$, $c_\theta \equiv \cos \theta$, where θ is the mixing angle between the two scalars ϕ and σ , and finally $\lambda_{C1} \equiv 6\lambda_H - 2\lambda_{HS}$ and $\lambda_{C2} \equiv 6\lambda_S - 2\lambda_{HS}$.

APPENDIX C: RIS SUBTRACTION

A well-known problem with Boltzmann equation networks concerns the double counting of pole contributions to scattering processes. The problem arises from imposing the on-shell limit for the dynamical two-point functions, when deriving the Boltzmann theory from the Kadanoff-Baym (KB) equations. In KB equations, where self-energies are defined by the two-particle irreducible (2PI)-effective action, the one-loop diagrams contain both on- and off-shell information. However, the on-shell limit spoils the 2PI hierarchy and removes the off-shell information at one-loop level. To regain it, one has to introduce by hand higher-order 1PI-self-energy diagrams, whose cuts then create most scattering processes in Boltzmann theory. This division does not automatically respect unitarity and, in particular, the poles of the 2-2-scattering processes encode information already included in the fusion processes onto and in the decays of the unstable ‘‘on-shell’’ state. Removing this double counting is usually dubbed real intermediate state subtraction.

To see how this works, we split the Breit-Wigner propagator as follows:

$$\begin{aligned} iD(p^2) &\equiv \frac{i}{p^2 - m^2 + im\Gamma} \\ &= \frac{i(p^2 - m^2)}{(p^2 - m^2)^2 + m^2\Gamma^2} + \frac{m\Gamma}{(p^2 - m^2)^2 + m^2\Gamma^2} \\ &= iD_H(p^2) + \mathcal{A}(p^2), \end{aligned} \quad (\text{C1})$$

where Γ is the decay width of the propagating particle with mass m . The propagator is resonant at $p^2 = m^2$ with a resonance of width $m\Gamma$, isolated in the spectral function $\mathcal{A}(p^2)$. This observation clearly suggests to remove the on-shell contribution from the propagator by simply dropping the spectral part and retaining only the Hermitian part

$$iD(p^2) \rightarrow iD_H(p^2) \quad (\text{C2})$$

in all scattering calculations. This corresponds to removing the pole contribution at the level of the matrix element as proposed in [38,39]. This is the approach taken in this paper.

A common practice in the literature has been [36,37,41–46] to subtract the divergence at the level of the matrix element squared, according to prescription

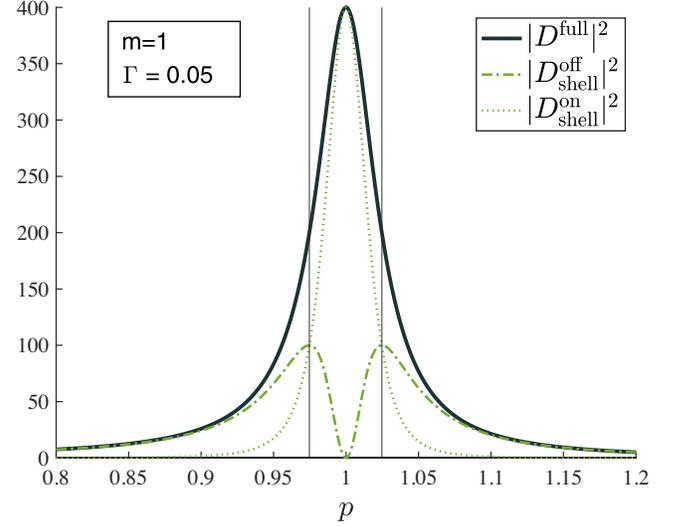


FIG. 8. Shown are the Breit-Wigner propagator split into on- and off-shell part definitions used in [39]. The split is done for the propagator and the resulting difference shown for the propagator squared. Full propagator $D(p^2)$ is shown with the black solid line, the off-shell part $D_H(p^2)$ by the dash-dotted line, and the on-shell part $\mathcal{A}(p^2)$ by dash-dotted lines. The vertical lines mark the resonance width $p = \sqrt{m^2 \pm m\Gamma}$, which coincides with the maxima of the off-shell propagator.

$$|D(p^2)|^2 \rightarrow |D(p^2)|^2 - \frac{\pi\delta(p^2 - m^2)}{m\Gamma}. \quad (\text{C3})$$

One can argue for this prescription by noting that, for a simple matrix element with no mixing between different channels,

$$|\mathcal{M}|^2 \sim D(p^2)D^*(p^2) = D_H^2(p^2) + \mathcal{A}^2(p^2). \quad (\text{C4})$$

To the lowest order in a small but finite Γ , the square of the spectral function can be replaced by

$$\mathcal{A}^2(p^2) \approx \frac{\pi\delta(p^2 - m^2)}{m\Gamma}. \quad (\text{C5})$$

The prescription (C3) is thus approximately the same as (C2). However, this argument fails in the presence of interference terms and the prescription (C3) has been found to give superficial negative cross sections [38], which never happens in the subtraction scheme (C2). In Fig. 8 we show the split of the square of the propagator function into the on- and off-shell contributions according to (C2) in a representative case.

- [1] P. A. R. Ade *et al.* (Planck Collaboration), Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, A13 (2016).
- [2] Y. Zel'dovich, L. Okun, and S. Pikelner, Quarks: Astrophysical and physicochemical aspects, *Sov. Phys. Usp.* **8**, 702 (1966).
- [3] B. W. Lee and S. Weinberg, Cosmological Lower Bound on Heavy Neutrino Masses, *Phys. Rev. Lett.* **39**, 165 (1977).
- [4] P. Gondolo and G. Gelmini, Cosmic abundances of stable particles: Improved analysis, *Nucl. Phys.* **B360**, 145 (1991).
- [5] K. Griest and D. Seckel, Three exceptions in the calculation of relic abundances, *Phys. Rev. D* **43**, 3191 (1991).
- [6] T. Binder, T. Bringmann, M. Gustafsson, and A. Hryczuk, Early kinetic decoupling of dark matter: When the standard way of calculating the thermal relic density fails, *Phys. Rev. D* **96**, 115010 (2017); **101**, 099901(E) (2020).
- [7] K. Ala-Mattinen and K. Kainulainen, Precision calculations of dark matter relic abundance, *J. Cosmol. Astropart. Phys.* **09** (2020) 040.
- [8] T. Abe, Early kinetic decoupling and a pseudo-Nambu-Goldstone dark matter model, *Phys. Rev. D* **104**, 035025 (2021).
- [9] M. Drewes *et al.*, A white paper on keV sterile neutrino dark matter, *J. Cosmol. Astropart. Phys.* **01** (2017) 025.
- [10] A. Merle and M. Totzauer, keV sterile neutrino dark matter from singlet scalar decays: Basic concepts and subtle features, *J. Cosmol. Astropart. Phys.* **06** (2015) 011.
- [11] J. König, A. Merle, and M. Totzauer, keV sterile neutrino dark matter from singlet scalar decays: The most general case, *J. Cosmol. Astropart. Phys.* **11** (2016) 038.
- [12] K. R. Dienes, F. Huang, J. Kost, S. Su, and B. Thomas, Deciphering the archaeological record: Cosmological imprints of nonminimal dark sectors, *Phys. Rev. D* **101**, 123511 (2020).
- [13] K. Kainulainen and A. Sorri, Oscillation induced neutrino asymmetry growth in the early Universe, *J. High Energy Phys.* **02** (2002) 020.
- [14] J. Ghiglieri and M. Laine, Improved determination of sterile neutrino dark matter spectrum, *J. High Energy Phys.* **11** (2015) 171.
- [15] S. Hannestad, R. S. Hansen, T. Tram, and Y. Y. Y. Wong, Active-sterile neutrino oscillations in the early Universe with full collision terms, *J. Cosmol. Astropart. Phys.* **08** (2015) 019.
- [16] D. Bodeker and A. Klaus, Sterile neutrino dark matter: Impact of active-neutrino opacities, *J. High Energy Phys.* **07** (2020) 218.
- [17] W. R. Yueh and J. R. Buchler, Scattering functions for neutrino transport, *Astrophys. Space Sci.* **39**, 429 (1976).
- [18] S. Hannestad and J. Madsen, Neutrino decoupling in the early Universe, *Phys. Rev. D* **52**, 1764 (1995).
- [19] D. V. Semikoz and I. I. Tkachev, Condensation of bosons in kinetic regime, *Phys. Rev. D* **55**, 489 (1997).
- [20] A. D. Dolgov, S. H. Hansen, and D. V. Semikoz, Non-equilibrium corrections to the spectra of massless neutrinos in the early Universe, *Nucl. Phys.* **B503**, 426 (1997).
- [21] F. Hahn-Woernle, M. Plumacher, and Y. Y. Y. Wong, Full Boltzmann equations for leptogenesis including scattering, *J. Cosmol. Astropart. Phys.* **08** (2009) 028.
- [22] I. M. Oldengott, C. Rampf, and Y. Y. Y. Wong, Boltzmann hierarchy for interacting neutrinos I: Formalism, *J. Cosmol. Astropart. Phys.* **04** (2015) 016.
- [23] V. Silveira and A. Zee, Scalar Phantoms, *Phys. Lett.* **161B**, 136 (1985).
- [24] J. McDonald, Gauge singlet scalars as cold dark matter, *Phys. Rev. D* **50**, 3637 (1994).
- [25] C. P. Burgess, M. Pospelov, and T. ter Veldhuis, The minimal model of nonbaryonic dark matter: A singlet scalar, *Nucl. Phys.* **B619**, 709 (2001).
- [26] V. Barger, P. Langacker, M. McCaskey, M. J. Ramsey-Musolf, and G. Shaughnessy, LHC phenomenology of an extended standard model with a real scalar singlet, *Phys. Rev. D* **77**, 035005 (2008).
- [27] M. Farina, D. Pappadopulo, and A. Strumia, CDMS stands for constrained dark matter singlet, *Phys. Lett. B* **688**, 329 (2010).
- [28] J. M. Cline, K. Kainulainen, P. Scott, and C. Weniger, Update on scalar singlet dark matter, *Phys. Rev. D* **88**, 055025 (2013); **92**, 039906(E) (2015).
- [29] P. A. Zyla *et al.* (Particle Data Group), Review of particle physics, *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).
- [30] T. Bringmann (private communication).
- [31] A. D. Dolgov and K. Kainulainen, Fermi-Dirac corrections to the relic abundances, *Nucl. Phys.* **B402**, 349 (1993).
- [32] S. B. Roland and B. Shakya, Cosmological imprints of frozen-in light sterile neutrinos, *J. Cosmol. Astropart. Phys.* **05** (2017) 027.
- [33] A. Boyarsky, M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy, Sterile neutrino dark matter, *Prog. Part. Nucl. Phys.* **104**, 1 (2019).
- [34] V. Khachatryan *et al.* (CMS Collaboration), Searches for invisible decays of the Higgs boson in pp collisions at $\sqrt{s} = 7, 8$, and 13 TeV, *J. High Energy Phys.* **02** (2017) 135.
- [35] G. Aad *et al.* (ATLAS Collaboration), Combined measurements of Higgs boson production and decay using up to 80 fb⁻¹ of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment, *Phys. Rev. D* **101**, 012002 (2020).
- [36] E. W. Kolb and S. Wolfram, Baryon number generation in the early Universe, *Nucl. Phys.* **B172**, 224 (1980); **B195**, 542(E) (1982).
- [37] G. F. Giudice, A. Notari, M. Raidal, A. Riotto, and A. Strumia, Towards a complete theory of thermal leptogenesis in the SM and MSSM, *Nucl. Phys.* **B685**, 89 (2004).
- [38] J. M. Cline, K. Kainulainen, and K. A. Olive, Protecting the primordial baryon asymmetry from erasure by sphalerons, *Phys. Rev. D* **49**, 6394 (1994).
- [39] J. M. Cline, K. Kainulainen, and D. Tucker-Smith, Electroweak baryogenesis from a dark sector, *Phys. Rev. D* **95**, 115006 (2017).
- [40] Y. Du, F. Huang, H.-L. Li, Y.-Z. Li, and J.-H. Yu, Revisit Dark matter freeze-in and freeze-out through phase-space distribution, *J. Cosmol. Astropart. Phys.* **04** (2022) 012.
- [41] M. A. Luty, Baryogenesis via leptogenesis, *Phys. Rev. D* **45**, 455 (1992).

- [42] M. Plumacher, Baryogenesis and lepton number violation, *Z. Phys. C* **74**, 549 (1997).
- [43] W. Buchmuller and M. Plumacher, Neutrino masses and the baryon asymmetry, *Int. J. Mod. Phys. A* **15**, 5047 (2000).
- [44] A. Pilaftsis and T. E. J. Underwood, Resonant leptogenesis, *Nucl. Phys.* **B692**, 303 (2004).
- [45] M. Frigerio, T. Hambye, and E. Masso, Sub-GeV Dark Matter as Pseudo-Goldstone from the Seesaw Scale, *Phys. Rev. X* **1**, 021026 (2011).
- [46] N. Bernal, C. Cosme, T. Tenkanen, and V. Vaskonen, Scalar singlet dark matter in non-standard cosmologies, *Eur. Phys. J. C* **79**, 30 (2019).