Momentum distributions of cosmic relics: Improved analysis

Kalle 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 University of Helsinki, Finland
2 Helsinki Institute of Physics, P.O.Box 64, FI-00014 University of Helsinki, Finland
3 Department of Physics, University of Jyväskylä

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 non-thermal nature of their resulting distributions. Here the effects turned out to be negligible, proving the robustness of the earlier predictions.

I. INTRODUCTION

Current cosmological observations can be accommodated within the cold dark matter (CDM) paradigm. 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.

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. 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. 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.

However, if the DM particle is not in kinetic equilibrium, the resulting suppression may be more complicated.

Earlier calculations accounting for the DM momentum distributions tend to rely on simplifying approximations. For example, in the analysis of, 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.

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 the momentum-dependent method of, which uses a generalized relaxation time approximation for the numerically expensive backreaction terms. Our results validate the approximation scheme of 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. We also apply our method in a sterile neutrino DM model including a singlet scalar mediator, first analyzed with simplified evolution equations in. 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 are robust and their results remain valid in the full solution.

The paper is organized as follows: In Sec. we describe the discretization of the collision integrals. We then apply the developed methodology to two benchmark models, first, in Sec. to the model where SM is extended with a real singlet scalar and then in Sec. 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 toward 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

\[ \frac{\partial f(p,t)}{\partial t} - H p \frac{\partial f}{\partial p} = \sum C_{\text{coll}}[f], \]

where \( C_{\text{coll}} \) 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, \( k = a p \). Then, writing the Liouville operator in terms of \( k \), the momentum derivative vanishes and we have

\[ \frac{df(k,t)}{dt} = \sum C_{\text{coll}}[f], \]

where we identified \( f(p,t) = f(k/a(t),t) \equiv \tilde{f}(k,t) \). 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 \( s/s = -3H \), where \( s = 2\pi^2\rho_{\text{rad}}T^3/45 \) is the entropy density of the Universe and \( H = (4\pi^3g_{\text{eff}}/45)^{1/2}T^2/M_P \) and \( g_{\text{eff}}(T) \) and \( h_{\text{eff}}(T) \) are the effective number of relativistic energy and entropy degrees of freedom. This implies

\[ \frac{\dot{T}}{T} = -\sqrt{\frac{4\pi^3}{45}}\frac{h_{\text{eff}}(T)}{g_{\text{eff}}(T)} \frac{T^2}{M_P}, \]

where

\[ g_{s/2}(T) \equiv \frac{h_{\text{eff}}(T)}{g_{s/2}(T)} \left(1 + \frac{T}{3h_{\text{eff}}(T)} \right). \]

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

\[ \xi \equiv \frac{k}{T_0a(t_0(T_0))} = \left(\frac{h_{\text{eff}}(T_0)}{h_{\text{eff}}(T)}\right)^{1/3} \frac{p}{T}, \]

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

\[ \frac{df(\xi,x)}{dx} = \sqrt{\frac{45}{4\pi^3}} g_{s/2} \left(\frac{m_0}{x}\right) \left(\frac{T_0}{x}\right)^3 \sum C_{\text{coll}}[f]. \]

This Boltzmann equation, written in comoving variables, can be solved numerically by discretizing in variables \( x \) and \( \xi \). 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 \( C_{\text{coll}} \).

A. Collision integral

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

\[ C_{\text{coll}}[f_i] = \frac{1}{2E_i} \int dPS_{ijk} \left( f_i f_j f_k f_\xi \right) |M_{12 \rightarrow 34}|^2, \]

where the integration measure over the phase space is

\[ dPS_{ijk} = (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4)\delta^3(p_1)\delta^3(p_2)\delta^3(p_3) \delta^3(p_4), \]

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 \( i = 1, 2, 3, 4 \). The matrix element squared is also assumed to contain all relevant symmetry factors for the initial and final states. The collision integral \( C_{\text{coll}} \) naturally splits into the backward and forward terms, given by

\[ C_{\text{BW}} \equiv \frac{1}{2E_1} \int dPS_{234} f_2 f_3 f_4 (f_1 f_1 f_2 f_2) |M_{12 \rightarrow 34}|^2 \]

and

\[ C_{\text{FW}} \equiv -\frac{1}{2E_1} \int dPS_{234} f_1 f_2 f_3 f_4 (f_1 f_1 f_2 f_2) |M_{12 \rightarrow 34}|^2. \]

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 [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 \( \xi \rightarrow \xi_j, j = 1, \ldots, N_\xi \), 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 \( f_a(x, \xi_j) \equiv f_{a_j}(x) \), over some temperature range \( x \), that must be solved simultaneously for each degree of freedom \( a \) and the momentum mode \( \xi_j \):

\[
\frac{d}{dx} f_{a_j}(x) = \alpha(x) \sum_{\text{coll.}} C_{a_j}(x).
\]

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 \( f_{a_j} \). 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 \( f_{a_j} = 0 \) 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 \( \xi \), 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 \( \xi_j \) at a given temperature \( T = m_0/x \) is then obtained by inverting Eq. (5):

\[
p_j(x) = \xi_j \frac{m_0}{x} \left( \frac{\hbar^2 (\frac{m_0}{x})}{\hbar^2 (T_0)} \right)^{1/3}.
\]

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

\[
C_{1j}^{BW}(x) = \frac{1}{16E_{1j}} \sum_k \sum_l \frac{\Delta p_k p_l^2}{2\pi^2 E_{3kl}} \frac{\Delta p_l p_l^2}{2\pi^2 E_{4kl}} \times F_{1jkl}(x) A_{1jkl}^{BW}(x),
\]

where the discretized backward phase space factor is

\[
A_{1jkl}^{BW}(x) = f_{3k}(x) f_{4l}(x) [1 \pm f_{1j}(x)] [1 \pm f_{2jkl}(x)]
\]

and \( F_{1jkl}(x) \) is the angular integral over the matrix element squared defined in Eq. (A12). Here the superscripts \( 1, \ldots, 4 \) denote the particle species (in the sense described above) involved in the \( 12 \leftrightarrow 34 \) process, \( f_{1j}(x) \) is the value of distribution function of the particle "a" with momentum \( \xi_j \) at temperature \( T = m_0/x \) and energy \( E_{ai} = (p_i^2 + m_i^2)^{1/2} \). 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 \( h_{\text{eff}}, g_{\text{eff}}, g_\ast \).

3) Determine relevant interactions and compute their matrix elements and cross sections.

4) Precalculate the \( F_{1jkl}(x) \) 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

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

where the scalar potential is given by

\[ V(S, H) = -\mu_s^2 |H|^2 - \frac{1}{2} \mu_s^2 S^2 + \lambda_S |H|^4 + \frac{\lambda_H}{4} S^4 + \frac{\lambda_{HS}}{2} |H|^2 S^2, \tag{17} \]

and the gauge interactions of the Higgs doublet \(H\) are contained in \(\mathcal{L}_{\text{SM}}\). The stability of the potential requires that the quartic couplings \(\lambda_S\) and \(\lambda_H\) of the singlet and the Higgs fields are positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be positive, but the portal coupling could be negative, as the stability of the potential requires just one to be negative. However, here we will only consider positive values of \(\lambda_{HS}\).

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:

\[ \partial_x f_s(x, \xi) = C^I_{s \leftrightarrow jj} + C^I_{h \rightarrow ss} + C^E_{sf \rightarrow sf}, \tag{18} \]

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 \(s\) with all SM particles can be assumed to be in thermal equilibrium. To correctly include kinematics when the Higgs mediator in \(C^I_{s \leftrightarrow jj}\) becomes on shell, we have included the on-shell Higgs decay \(C^I_{h \rightarrow ss}\) as a separate contribution and take all Higgs mediators in \(C^I_{s \leftrightarrow jj}\) 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 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. GRA calculation is similar to the one described in this paper, except for the numerically expensive elastic backward term \(C_{sf \rightarrow sf}\), given by \(i\). In the GRA method, this is treated in a simplifying approximation; for more details, see \[7\].

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 \(T \sim 10\) it starts to overestimate the elastic integral that enforces the kinetic equilibrium. Eventually the error becomes of order \(\sim 2\), but only well after the freeze-out \(T \sim 40\). 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 are already 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 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 resonating region, a precision calculation of its abundance

\[ 2\ \text{Γ}_h \approx 4\ \text{MeV}, \text{ so the width of the resonance in the} \sqrt{s} \text{ variable is} \sqrt{m_i \Gamma}\approx 0.7\ \text{GeV}. \]
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 \(O(1)\) s, generalized relaxation approximation runs \(O(1 - 10)\) min, and the complete calculation with backreaction \(\gtrsim 10\) h per \((m_s, \lambda_{hs})\) 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...
leads to the mass matrix for neutral scalars

\[ \mathcal{L} = \mathcal{L}_{\text{SM}} + i\overline{N}\phi N + \frac{1}{2} (\partial_\mu S)^2 - V(S,H) + yS\overline{N}N, \quad (19) \]

where we denote the SM Higgs doublet by \( H \). Its gauge interactions are contained in the SM Lagrangian \( \mathcal{L}_{\text{SM}} \), 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 \( m_S \sim \mathcal{O}(10-1000) \text{ GeV} \). The \( N \)-fermion mass gets a contribution from nonzero vacuum expectation value (VEV) of the singlet scalar \( m_N = \mu_N + y(S) \).

As the VEV can be quite large, we need to assume the Yukawa coupling to be tiny, \( y \ll 1 \). Thus, the treatment is more straightforward, as the fields \( \sigma \) and \( \phi \) are directly the mass eigenstates of the mass matrix. Without going into further details, we simply note that the Feynman rules of Appendix I can be directly applied also in this case by setting \( \theta, w \to 0 \). In the limit of vanishing Yukawa coupling and singlet scalar VEV, \( y, w \to 0 \), the model reduces to the singlet scalar model from previous section.

With a slight abuse of notation, we denote the mass eigenstates by \( \sigma \) and \( \phi \), 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 \( m_1 \equiv m_\sigma, m_2 \equiv m_\phi \), the portal coupling \( \lambda_{\text{HS}} \), 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 \( v = 246 \text{ GeV} \). 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 \( \sigma \). Processes of order \( \mathcal{O}(y^2) \) are negligible and the relevant contributions under our assumptions are summarized in Table I. Because of nonzero Yukawa coupling and assumed mass hierarchy, eventually all produced \( \sigma \) scalars will decay into \( N \) fermions, which remains as a stable relic. Production of \( \sigma \) scalars is determined by the portal coupling \( \lambda_{\text{HS}} \). 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 \( \sigma \)-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.

To obtain the momentum distribution function for \( \sigma \) scalar and \( N \) fermion we must solve the following set of
TABLE I. Relevant tree-level production processes for singlet fermion \(N\) and singlet scalar \(\sigma\) in the model \[10\] assuming \(y \ll 1\). Each \(s, t,\) and \(u\) channel reaction can be mediated via both scalar fields: singlet \(\sigma\) and SM Higgs \(\phi\). Here \(V = \{W, Z\}\) labels vector bosons and \(f = \{\tau, c, b, t\}\) labels SM fermions.

| \(\sigma\) reaction | \(\phi\) reaction |
|---------------------|------------------|
| \(\sigma \leftrightarrow \phi\) | \(\phi \rightarrow NN\) |
| \(\sigma \leftrightarrow VV\) | \(N\sigma \leftrightarrow N\sigma\) |
| \(\sigma \leftrightarrow ff\) | \(N\phi \leftrightarrow N\phi\) |
| \(\phi \rightarrow \sigma\) | \(Nf \leftrightarrow Nf\) |
| \(\sigma \rightarrow NN\) | \(\sigma f \leftrightarrow \sigma f\) |
| \(\sigma \sigma \leftrightarrow \sigma\) | \(N\sigma \leftrightarrow N\phi\) |

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 \(\sigma f \leftrightarrow \sigma f\), \(\sigma \sigma \leftrightarrow \sigma\), 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 \(\sigma\) 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 \[12\] 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 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 \textit{a posteriori}: The second peak at high momenta forms only when \(\sigma\) production from the SM heat bath has already stopped and the remaining scalars decay into \(N\) fermions. At this point, the SM Higgs 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 fee-
ably 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 $w = 0$. 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 evalua-
tion 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 sym-
metry is given by Eq. (27). It increases the $\sigma$ 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. 3 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 suppress-
ion 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 $\sigma$ scalar and SM heat
bath strongly increase for increasing values of $w$, as il-
lustrated in Fig. 5. As a result, the number density of $\sigma$
particles becomes more and more Boltzmann sup-
pressed 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 in-
tegrated yields of the singlet scalar fields (dotted lines) as
a function of $m_0/T$ for the same choices of parameters as
in Fig. 3. 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 $\sigma$-field abundance.

We found that the elastic interactions had negligible ef-
fect on the $N$ distribution in comparison with the Boltz-
mann 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 equi-
lbrium. This is partly because the having $w > 0$ also
increases the overall production of $N$ fermions via inelas-
tic and decay channels, which is the root cause for the
nonthermal distributions. We then find that even in the

\begin{table}
\centering
\begin{tabular}{|c|c|c|}
\hline
$w$ & $f_N(\xi,x)$ & Notes \\
\hline
$0$ & & \\
$10^3$ & & \\
$10^6$ & & \\
\hline
\end{tabular}
\end{table}

\[ \text{FIG. 5. Cross section between the singlet scalars } S = w + \sigma \text{ and SM particles plotted over a representative range of incoming energies with varying singlet VEV } w \text{ (mixing angle } |\sin \theta|). \text{ Other parameters are } \{m_S = 60 \text{ GeV, } \lambda_{HS} = 4 \times 10^{-6.86}\}. \]
interactions were irrelevant in general. Instead of constructing a model just to make the elastic rates important, as a proof of concept, we simply increase the elastic process \( N_f \leftrightarrow N_f \) 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 \( M_{N_f \rightarrow N_f} \) 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 phase space distribution. Our first example concerned a singlet sector. We found that this rate is determined by a combination of \( \lambda_{HS} \) and \( w \), we can always find such a \( (w', \lambda_{HS}) \) 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 [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 a 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 preevaluated prior to the solution of the Boltzmann equation.

1. 2-2 Scatterings

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 \(dp_2\), which leaves out a delta-function \(\delta(p_2^2 - m_2^2)\), where

\[
p_2^2 - m_2^2 = Q + 2(p_3 \cdot p_4 - p_1 \cdot p_3 - p_1 \cdot p_4),
\]

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

\[
\begin{align*}
p_1 &= p_1 (0, 0, 1), \\
p_3 &= p_3 (0, \sin \theta, \cos \theta), \\
p_4 &= p_4 (\sin \alpha \sin \beta, \sin \alpha \cos \beta, \cos \alpha).
\end{align*}
\]

The remaining integrals can now be readily reduced to

\[
\begin{align*}
d^3p_3 &= 2\pi p_3^2 dp_3 \cos \theta, \\
d^3p_4 &= p_4^2 dp_4 \cos \alpha d\beta.
\end{align*}
\]

The integral over \(\beta\) can be done [18] using

\[
\delta(p_2^2 - m_2^2) = \delta(f(\beta)) = \sum_i |f'(\beta_i)|^{-1} \delta(\beta - \beta_i),
\]

where \(f'(\beta) = 2p_3p_4 \sin \alpha \sin \theta \sin \beta\) and \(\beta_i\) are the roots of \(f(\beta) = 0\). It is easy to see that the \(\delta\) function is symmetric in \(\beta\), 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

\[
f'(\beta_i)^2 = (2p_3p_4 \sin \alpha \sin \theta \sin \beta_i)^2 \geq 0,
\]

so that

\[
\int_0^{2\pi} d\beta \delta(f(\beta)) = 2f'(\beta_i)|^{-1}\Theta(|f'(\beta_i)|^2).
\]

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

\[
|f'(\beta_i)| = \sqrt{a \cos^2 \alpha + b \cos \alpha + c},
\]

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

\[
\begin{align*}
a &= p_3^2(-4\kappa + 8\epsilon), \\
b &= p_4(-p_1 + \epsilon/p_1)(8\gamma + 4Q + 8\epsilon), \\
c &= 4p_3^2p_4^2 \sin^2 \theta - (2(\gamma + \epsilon) + Q)^2,
\end{align*}
\]

with \(\gamma = E_3E_4 - E_1E_3 - E_1E_4\) and \(\epsilon = p_1p_3 \cos \theta\) and \(\kappa = p_1^2 + p_3^2\). Then, after a trivial \(d\mu\) integral, the source term becomes

\[
C_{BW} = \frac{2}{(2\pi)^4} \frac{1}{2E_1} \int_f^1 dp_3p_4^2 dp_3 \left[ \frac{2}{2E_3} \frac{2}{2E_4} F(p_1, p_3, p_4) \times f_3 f_4 [1 \pm f_1][1 \pm f_2],
\]

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,

\[
F(p_1, p_3, p_4) = \int_{-1}^{1} d\cos \theta \int_{\min(1, z^+)}^{\max(-1, z^-)} dz \times \frac{|M(s, t)|^2}{\sqrt{a(z - z^-)(z - z^+)}},
\]

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 \(s = (p_3 + p_1)^2\) and \(t = (p_1 - p_3)^2\), which depend on the angles and momenta as follows:

\[
\begin{align*}
t &= m_1^2 + m_3^2 - 2E_1E_3 + 2p_1p_3 \cos \theta, \\
s &= 2E_1E_3 + 2E_1E_4 - 2p_1p_3 \cos \theta - 2p_1p_4 \cos \alpha + m_2^2 - m_1^2.
\end{align*}
\]
In the particular case where the matrix element in \(\text{A12}\) does not depend on \(\cos \alpha\) (a pure \(t\)-channel process), the \(d \cos\) integral can be reduced to a one-dimensional integral,

\[
F(p_1, p_3, p_4) = \frac{1}{2} \int \frac{d \cos \theta |M(t)|^2 \pi}{\sqrt{-\alpha}} \Theta(b^2 - 4ac) .
\]  
(A14)

**Forward term.** In the forward term, given by Eq. \(\text{A11}\), we integrate first over \(d^3p_4\), which leaves us with the delta function \(\delta(p_4^2 - m_4^2)\). Paralleling the backward term reduction, we eventually obtain

\[
C_{FW} = -\frac{2}{(2\pi)^3} \left(\frac{1}{2E_1} \int \frac{p_2^2 dp_2}{2E_2} \frac{p_3^2 dp_3}{2E_3} F'(p_1, p_2, p_3) \times f_1 f_2 [1 \pm f_3][1 \pm f_4],
\]  
(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 \(\text{A12}\), where one replaces everywhere \((a, b, c) \rightarrow (a', b', c')\) with

\[
a' = p_2^2(-4\kappa' + 8\epsilon'),
\]  
(A16)

\[
b' = p_2 (p_1 - \epsilon'/p_1) (8\gamma' + 4Q' + 8\epsilon'),
\]  
(A17)

\[
c' = 4p_2^2 p_3^2 \sin^2 \theta - (2(\gamma' + \epsilon') + Q')^2,
\]  
(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 \(\kappa' = p_1^2 + p_2^2\).

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

\[
s = m_1^2 + m_2^2 + 2E_1 E_2 - 2p_1 p_2 \cos \alpha,
\]  
\[
t = m_1^2 + m_3^2 - 2E_1 E_3 + 2p_1 p_3 \cos \theta.
\]  
(A19)

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

2. Special case: Inelastic 2-2 scattering toward/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 \(\text{A11}\), 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 \(\text{A9}\) as

\[
\Lambda(f_1, f_2) = f_1^{eq} f_2^{eq} - f_1 f_2 .
\]  
(A20)

The collision can then be written as

\[
C_i(p_1, t) = \frac{1}{2E_1} \int d^3p_2 \Lambda(f_1, f_2) \sum_n F_i(s) \sigma_i^{(n)}(s) ,
\]  
(A21)

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

\[
F_i(s) = 4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} ,
\]  
\[= 2\lambda^{1/2}(s, m_1^2, m_2^2) = 4E_1 E_2 v_{\text{MB}},
\]  
(A22)

where \(\lambda(a, b, c) = (a - b - c)^2 - 4bc\) is the Källén kinetic function. As explained in \(\text{A7}\), one can reduce the integral over \(d^3p_2\) to a single integral over \(s\),

\[
C_i(p_1, t) = \frac{1}{2\pi^2} \int_0^{\infty} d\tau \sigma_i^{(n)}(s) ,
\]  
(A23)

where

\[
\sigma_i^{(n)}(p_1, p_2) = \frac{1}{16p_1 p_2 E_1 E_2} \int_{s_{\min}}^{s_{\max}} ds O(s) .
\]  
(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.

**Decay from (fusion to) equilibrium** \(f_A^{eq} \leftrightarrow f_1 f_2\). Assume we are tracking the species labeled as "1", while the "2" species is arbitrary and "A" species follows the equilibrium. Then,

\[
C_{A-12}^{eq}(p_1, t) = \frac{1}{2E_1} \int d\tau \Lambda(f_A^{eq}, f_1, f_2) |M_{A-12}|^2 ,
\]  
(A25)

where the phase space integration is denoted by

\[
d\tau = d^3p_A d^3p_{f_1} d^3p_{f_2} (2\pi)^4 \delta^{(4)}(p_A - p_1 - p_2) ,
\]  
(A26)

and the distribution factor \(\Lambda\) is given by

\[
\Lambda = f_A^{eq} [1 \pm f_1] [1 \pm f_2] - f_1 f_2 [1 \pm f_A^{eq}] ,
\]  
\[
\rightarrow f_1^{eq} f_2^{eq} - f_1 f_2 ,
\]  
(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

$$C_{A-12}^{eq}(p_1,t) = \frac{|M_{A-12}|^2}{16\pi E_{1} p_1} \left( f_1^{eq} \chi[f_2^{eq}] - f_1 \chi[f_2] \right), \quad (A28)$$

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

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

with

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

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

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

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

where $v = \lambda^{1/2}(m_1^2, m_2^2, m_A^2)/m_1^2$ 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. (1) for the new/modified interaction vertices in the model with a singlet scalar $S = \sigma + w$ and a singlet fermion $N$.

$$\lambda_{\phi\phi\phi} = -6i\lambda_{HS}\phi^4 + 6s_H^2 (\lambda_{HS} \phi^2 + \lambda_s \phi^2),$$

$$\lambda_{\sigma\sigma\sigma} = -6i\lambda_{HC} \phi + 6s_H \phi (\lambda_{HC} \phi + \lambda_w \phi),$$

$$\lambda_{\phi\sigma\sigma} = -(i/4) \left( \lambda_{HS} [1 + 3c_H] + (3\lambda_H + 3\lambda_w) [1 - c_H] \right),$$

$$\lambda_{\phi\phi\sigma} = -6i\lambda_{HC} \phi^3 + 3s_H (\lambda_{HC} \phi^2 - \lambda_{HS} \phi^2),$$

$$\lambda_{\phi\sigma\sigma} = -6i\lambda_{HC} \phi^2 + 3s_H (\lambda_{HC} \phi + \lambda_w \phi),$$

$$\lambda_{\phi\phi\phi} = -i\lambda_{HC} \phi^3 + s_H (\lambda_{HC} \phi^2 - \lambda_{C1} \phi),$$

$$\lambda_{\sigma\sigma\sigma} = -i\lambda_{HC} \phi^3 + s_H (\lambda_{C2} \phi^2 + \lambda_{HS} \phi^2),$$

$$\lambda_{\sigma\sigma\sigma} = -2i(M^2/v)s_H \phi \sigma \sigma \nu \nu,$$

$$\lambda_{\sigma\sigma\nu} = -2i(M^2/v)c_H \phi \sigma \sigma \nu \nu,$$

$$\lambda_{\sigma\sigma\sigma} = -i(m_f/v)c_H,$$

$$\lambda_{\sigma\sigma\nu} = -i(m_f/v)s_H,$$

$$\lambda_{\sigma\sigma\sigma} = i(y/2)c_H. \quad (B1)$$

Here $\phi$ denotes the SM Higgs field with VEV $v$ and $\nu$ and $f$ denote SM vector bosons and fermions; we used the following shorthand notations: $s_H \equiv \sin \theta$, $c_H \equiv \cos \theta$, where $\theta$ is the mixing angle between the two scalars $\phi$ and $\sigma$, 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 (RIS) subtraction.

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

$$iD(p^2) \equiv \frac{i}{p^2 - m^2 + i\Gamma} = \frac{i(p^2 - m^2)}{(p^2 - m^2)^2 + m^2 \Gamma^2}) + \frac{m\Gamma}{(p^2 - m^2)^2 + m^2 \Gamma^2},$$

where $\Gamma$ 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 $A(p^2)$. 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

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

in all scattering calculations. This corresponds to removing the pole contribution at the level of the matrix element as proposed in [35 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

$$|D(p^2)|^2 \rightarrow |D(p^2)|^2 - \frac{\pi \delta(p^2 - m^2)}{m\Gamma}. \quad (C3)$$
One can argue for this prescription by noting that, for a simple matrix element with no mixing between different channels,

$$\vert M \vert^2 \sim D(p^2)D^*(p^2) = D_H^2(p^2) + \mathcal{A}^2(p^2).$$

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

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

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.

---

**References**

[1] P. A. R. Ade et al. (Planck), Planck 2015 results. XIII. Cosmological parameters, Astron. Astrophys. **594**, A13 (2016) [arXiv:1502.01589 [astro-ph.CO]]

[2] Y. Zel’dovich, L. Okun, and S. Pikelner, Quarks: astrophysical and physicochemical aspects, Sov. Phys. Uspekhi. 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). [Erratum: Phys.Rev.D 101, 099901 (2020)], arXiv:1706.07433 [astro-ph.CO].

[7] K. Ala-Mattinen and K. Kainulainen, Precision calculations of dark matter relic abundance, JCAP **09**, 040 [arXiv:1912.02870 [hep-ph]].

[8] T. Abe, Early kinetic decoupling and a pseudo-Nambu-Goldstone dark matter model, Phys. Rev. D 104, 035025 (2021) [arXiv:2106.01956 [hep-ph]].

[9] M. Drewes et al., A White Paper on keV Sterile Neutrino Dark Matter, JCAP **01**, 025 [arXiv:1602.04816 [hep-ph]].

[10] A. Merle and M. Totzauer, keV Sterile Neutrino Dark Matter from Singlet Scalar Decays: Basic Concepts and Subtle Features, JCAP **1506**, 011 [arXiv:1502.01011 [hep-ph]].

[11] J. König, A. Merle, and M. Totzauer, keV Sterile Neutrino Dark Matter from Singlet Scalar Decays: The Most General Case, JCAP **11**, 038 [arXiv:1609.01289 [hep-ph]].

[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) [arXiv:2001.02193 [astro-ph.CO]].

[13] K. Kainulainen and A. Sorri, Oscillation induced neutrino asymmetry growth in the early universe, JHEP **02**, 019, [arXiv:1506.05266 [hep-ph]].

[14] J. Gligleri and M. Laine, Improved determination of sterile neutrino dark matter spectrum, JHEP **11**, 171 [arXiv:1506.06752 [hep-ph]].

[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, JCAP **1508**, 019, [arXiv:1506.05266 [hep-ph]].

[16] D. Bodeker and A. Klaus, Sterile neutrino dark matter: Impact of active-neutrino opacities, JHEP **07**, 218 [arXiv:2005.03093 [hep-ph]].

[17] W. R. Yueh and J. R. Buchler, Scattering functions for neutrino transport, Astrophysics and Space Science **39**, 429 (1976).

[18] S. Hannestad and J. Madsen, Neutrino decoupling in the early universe, Phys. Rev. D52, 1764 (1995) [arXiv:astro-ph/9506015 [astro-ph]].

[19] D. V. Semikoz and I. I. Tkachev, Condensation of bosons in kinetic regime, Phys. Rev. D55, 489 (1997) [arXiv:hep-ph/9507306 [hep-ph]].

[20] A. D. Dolgov, S. H. Hansen, and D. V. Semikoz, Nonequilibrium corrections to the spectra of massless neutrinos in the early universe, Nucl. Phys. B503, 426 (1997) [arXiv:hep-ph/9703315 [hep-ph]].

[21] F. Hahn-Woernle, M. Plumacher, and Y. Y. Y. Wong,
Full Boltzmann equations for leptogenesis including scattering, JCAP **0908**, 028 [arXiv:0907.0205 [hep-ph]]

[22] I. M. Oldengott, C. Rampf, and Y. Y. Y. Wong, Boltzmann hierarchy for interacting neutrinos I: formalism, JCAP **1504** (04), 016, arXiv:1409.1577 [astro-ph.CO].

[23] V. Silveira and A. Zee, SCALAR PHANTOMS, Phys. Lett. B **161**, 136 (1985).

[24] J. McDonald, Gauge singlet scalars as cold dark matter, Phys. Rev. D **50**, 3637 (1994) [arXiv:hep-ph/0702143]

[25] C. P. Burgess, M. Pospelov, and T. ter Veldhuis, The Minimal model of nonbaryonic dark matter: A Singlet scalar, Nucl. Phys. B **619**, 709 (2001) [arXiv:hep-ph/0011335].

[26] I. M. Oldengott, C. Rampf, and Y. Y. Y. Wong, Boltzmann hierarchy for interacting neutrinos I: formalism, JCAP **1504** (04), 016, arXiv:1409.1577 [astro-ph.CO].

[27] V. Silveira and A. Zee, SCALAR PHANTOMS, Phys. Lett. B **161**, 136 (1985).

[28] J. McDonald, Gauge singlet scalars as cold dark matter, Phys. Rev. D **50**, 3637 (1994) [arXiv:hep-ph/0702143]

[29] C. P. Burgess, M. Pospelov, and T. ter Veldhuis, The Minimal model of nonbaryonic dark matter: A Singlet scalar, Nucl. Phys. B **619**, 709 (2001) [arXiv:hep-ph/0011335].

[30] I. M. Oldengott, C. Rampf, and Y. Y. Y. Wong, Boltzmann hierarchy for interacting neutrinos I: formalism, JCAP **1504** (04), 016, arXiv:1409.1577 [astro-ph.CO].

[31] V. Silveira and A. Zee, SCALAR PHANTOMS, Phys. Lett. B **161**, 136 (1985).

[32] J. McDonald, Gauge singlet scalars as cold dark matter, Phys. Rev. D **50**, 3637 (1994) [arXiv:hep-ph/0702143]

[33] C. P. Burgess, M. Pospelov, and T. ter Veldhuis, The Minimal model of nonbaryonic dark matter: A Singlet scalar, Nucl. Phys. B **619**, 709 (2001) [arXiv:hep-ph/0011335].

[34] V. Silveira and A. Zee, SCALAR PHANTOMS, Phys. Lett. B **161**, 136 (1985).

[35] J. McDonald, Gauge singlet scalars as cold dark matter, Phys. Rev. D **50**, 3637 (1994) [arXiv:hep-ph/0702143]

[36] C. P. Burgess, M. Pospelov, and T. ter Veldhuis, The Minimal model of nonbaryonic dark matter: A Singlet scalar, Nucl. Phys. B **619**, 709 (2001) [arXiv:hep-ph/0011335].

[37] V. Silveira and A. Zee, SCALAR PHANTOMS, Phys. Lett. B **161**, 136 (1985).

[38] J. McDonald, Gauge singlet scalars as cold dark matter, Phys. Rev. D **50**, 3637 (1994) [arXiv:hep-ph/0702143]

[39] C. P. Burgess, M. Pospelov, and T. ter Veldhuis, The Minimal model of nonbaryonic dark matter: A Singlet scalar, Nucl. Phys. B **619**, 709 (2001) [arXiv:hep-ph/0011335].

[40] V. Silveira and A. Zee, SCALAR PHANTOMS, Phys. Lett. B **161**, 136 (1985).

[41] J. McDonald, Gauge singlet scalars as cold dark matter, Phys. Rev. D **50**, 3637 (1994) [arXiv:hep-ph/0702143]

[42] C. P. Burgess, M. Pospelov, and T. ter Veldhuis, The Minimal model of nonbaryonic dark matter: A Singlet scalar, Nucl. Phys. B **619**, 709 (2001) [arXiv:hep-ph/0011335].