PHYSICAL REVIEW D **90,** 096002 (2014)

Leptogenesis and the Higgs portal

Matthias Le Dall and Adam Ritz
Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia, V8P 5C2 Canada
(Received 28 August 2014; published 12 November 2014)

We study the impact on leptogenesis of Higgs portal couplings to a new scalar singlet. These couplings open up additional *CP*-violating decay channels for the higher mass singlet neutrinos N_2 and N_3. We analyze the simplest case of two-level $N_1 - N_2$ leptogenesis, including significant mass hierarchies, in which the *CP* asymmetry is generated in part by singlet-mediated decays of N_2. For these models, provided the lightest singlet neutrino N_1 is sufficiently weakly coupled to avoid excessive washout, its mass scale is not directly constrained by the Davidson–Ibarra bound.

DOI: 10.1103/PhysRevD.90.096002 PACS numbers: 11.30.Fs, 12.60.Fr, 14.60.St, 13.15.+g

I. INTRODUCTION

The discovery of neutrino oscillations [1], and thus small neutrino masses [2], provides motivation for leptogenesis [3–7] as a simple, and seemingly generic, mechanism for producing the observed baryon asymmetry in the Universe. The simplest UV completions of the dimension-5 Weinberg operator $HLHL$ [8], which contribute to neutrino masses, naturally incorporate heavy degrees of freedom with allowed *CP*-violating couplings, of which the out-of-equilibrium decay can generate a lepton asymmetry. Standard Model (SM) sphaleron processes can then equilibrate $B - L$ above the weak scale, resulting in the required late time asymmetry in baryon number [9].

Beyond the clear possibility to test the Majorana nature of neutrinos through neutrinoless double beta decay [10], the high-scale nature of leptogenesis—characterized for example by the Davidson–Ibarra bound [11]—and the lack of model-independent links between the high-scale and low-scale manifestations of *CP* violation, renders the mechanism feasible but hard to test [12]. This has motivated continuing study of variations of this general framework which may be placed under further experimental scrutiny, particularly those that allow a lowering of the scale (see e.g. Ref. [13]). This turns out to be quite difficult for the basic reason that the asymmetries generated by loop-level decays are counteracted by similar washout processes from two-to-two scattering. Lowering the scale at which the asymmetry is generated means a lower Hubble expansion rate, and thus more scattering processes will be in equilibrium and able to efficiently cancel the asymmetry. The conclusion is that it becomes increasingly difficult to find viable scenarios which operate at or close to the weak scale; an exception is the case of resonant decays [14].

From a theoretical perspective, the simplest realization of leptogenesis, with heavy right-handed (RH), i.e. singlet, Majorana neutrinos, involves one of the few renormalizable interactions between the SM and a neutral hidden (or dark) sector. The RH neutrino coupling is described by the Langrangian

$$\mathcal{L} = \mathcal{L}_{\rm SM} + i\bar{N}_i\gamma^\mu\partial_\mu N_i - \lambda_{ji}\bar{N}_i P_L L_j \cdot H - \frac{1}{2}M_i\bar{N}_i^c N_i + \text{H.c.} \tag{1}$$

The class of UV complete relevant or marginal interactions of the type characterized by λ_{ji}, known as portals, is very small. If we require no additional states, the list contains this right-handed neutrino coupling, $\lambda\bar{N}L\cdot H$; the coupling of a scalar singlet to the Higgs, $(\beta S + \lambda S^2)H^\dagger H$; and kinetic mixing of a U(1) vector with hypercharge, $\kappa V_{\mu\nu}B^{\mu\nu}$. Since the coefficients of these operators are unsuppressed by any heavy new physics scale, they are a natural place to look for signs of new short-distance physics. Couplings to a hidden sector are also motivated by our other primary piece of empirical evidence for new physics, namely dark matter, and thus these portals have been the focus of considerable recent attention [15].

In this paper, we explore the minimal extension of "standard" leptogenesis incorporating the Higgs portal coupling, which is now subject to direct experimental probes at the LHC. The relevant and marginal interactions then include

$$\mathcal{L} = \mathcal{L}_{\rm SM} + i\bar{N}_i\gamma^\mu\partial_\mu N_i - \lambda_{ji}\bar{N}_i P_L L_j \cdot H - \frac{1}{2}M_i\bar{N}_i^c N_i - \beta S H^\dagger H - \alpha_{ij}S\bar{N}_i^c P_L N_j + \text{H.c.} + \cdots \tag{2}$$

We have not shown the quartic Higgs portal coupling here as it will not play a significant role, other than for the full scalar potential, but have added the allowed Majoron coupling α_{ij} between the scalar S and N_i. The ensuing scenario for baryogenesis will be referred to as Higgs portal leptogenesis (HPL).

As we discuss below, this minimal extension is sufficient to open up new *CP*-violating decay channels for the next-to-lightest singlet neutrino N_2 (and N_3). These new

channels also decouple the source of the CP asymmetry from the seesaw contribution to the light neutrino mass. Consequently, leptogenesis can be viable in a wider mass range, and in particular when the lightest singlet neutrino is very light, e.g. below the sphaleron threshold. To analyze the impact of the new decay channels, we study the case of two-level $N_1 - N_2$ leptogenesis in detail, paying attention to the washout induced by scattering associated with N_1 and the additional scalar. In minimal leptogenesis, these two-to-two processes are generally negligible as the neutrino abundance is Boltzmann suppressed. This is not necessarily the case for $N_{2,1}$ scattering mediated by the hidden sector, and washout would be problematic if N_1 were in equilibrium for an extended period. Nonetheless, we find that there are viable regions of parameter space in which N_1 is parametrically light, but also very weakly coupled, in which the lepton asymmetry generated by N_2 decays survives to provide the observed baryon abundance (see e.g. Ref. [16]). The possibility of effectively decoupling N_1 from its normally dominant role in leptogenesis, and having it be parametrically light and experimentally accessible [13,16], is one of the interesting features of HPL.

The rest of the paper is organized as follows. In Sec. II, we outline the model and determine the additional contributions to the CP asymmetry from the hidden sector. Since the hidden sector contributions enter only through the decays of the next-to-lightest singlet neutrinos, in Sec. III we turn to the Boltzmann evolution of the coupled $N_1 - N_2$ system and study in some detail the important role of N_1-mediated two-to-two scattering in washing out the N_2-generated asymmetry. Section IV presents a number of results for the final lepton asymmetry in different mass regimes. The new hidden sector contributions to the CP asymmetry are crucial in allowing a lowering of the overall mass scale without violating the Davidson–Ibarra bound. We conclude in Sec. V, while a series of Appendixes contains further technical details.

II. HIGGS PORTAL AND THE CP ASYMMETRY

A. Model

Our focus in this paper will be a minimal extension of conventional leptogenesis, which adds a scalar singlet S along with the right-handed singlet neutrinos N_i. This opens up the Higgs and neutrino portals and, significantly, allows for a new CP-odd source in the hidden sector. CP violation is of course a central ingredient in leptogenesis, as it is in any theory of matter genesis according to the Sakharov conditions [17]. The portal couplings in the Lagrangian (2) include

$$\mathcal{L}_{\text{portals}} = -\lambda_{ji}\bar{N}_i P_L L_j \cdot H - \beta S H^\dagger H - \left(\frac{1}{2}M_i\delta_{ij} + \alpha_{ij}S\right)\bar{N}_i^c P_L N_j + \text{H.c.} + \cdots \quad (3)$$

The Higgs portal coupling β is one part of the full scalar potential $V(H,S)$ and necessarily breaks any $S \to -S$ symmetry. Thus, determining the vacuum structure requires a separate analysis incorporating thermal corrections. This potential has been studied in detail elsewhere [18], and here we simply assume that the parameters are chosen to ensure viable electroweak symmetry breaking, and importantly that $\langle S\rangle = 0$. The possibility of a more complex behavior of $\langle S\rangle$, which modifies the effective right-handed neutrino (RHN) mass is nonetheless interesting and will be discussed further in the concluding section.

In minimal leptogenesis, the Yukawa couplings λ determine both the light neutrino mass spectrum and the CP asymmetry generated in RH neutrino decays [19]. Opening the Higgs portal allows these two physical phenomena to be decoupled, with the Majoron coupling α_{ij} providing a new CP-odd source that is unconstrained (for $\langle S\rangle = 0$) by the light neutrino mass spectrum. In the Lagrangian (3), a unitary rotation has been used to diagonalize the RH neutrino mass matrix $M_{ij} \to M_i\delta_{ij}$. The Majorana nature of $N_i = N_i^c$ ensures that M_{ij} is symmetric, and for n flavors the diagonalization leaves M_i as n real mass eigenvalues [20]. In general, the corresponding rotation simply rearranges the $n(n+1)/2$ phases in the symmetric matrix α_{ij}, which is thus a physical CP-odd source in addition to the neutrino Yukawa λ_{ij}.[1]

B. CP asymmetry

In leptogenesis, the CP asymmetry arises from RHN decays to leptons, $N \to LH$ and $N \to \bar{L}\bar{H}$, and is measured by ϵ_i,

$$\epsilon_i \equiv \frac{\Gamma(N_i \to LH) - \Gamma(N_i \to \bar{L}\bar{H})}{\Gamma_i}. \quad (4)$$

In the denominator, the N_i decay rate is calculated at tree level and reads

$$\Gamma_i \equiv \sum_{k,\alpha,\beta}\left(\Gamma(N_i \to L_k^\alpha H^\beta) + \Gamma(N_i \to \overline{L_k^\alpha H^\beta})\right) = \frac{(\lambda^\dagger\lambda)_{ii}}{8\pi}M_i, \quad (5)$$

where the lepton family index, $k = 1,2,3$, stands for the electron, muon, and tau families respectively. The $\alpha,\beta = 1,2$ indices denote the components of the SU(2) lepton and Higgs doublets $L = (\,\nu_l \quad e_l\,)^T$ and $H = (\,H^+ \quad H^0\,)^T$. If we schematically write the decay amplitude as

[1]The coupling $\alpha_{ij}S\bar{N}_i^c P_L N_j + \text{H.c}$ is more commonly used to generate the right-handed neutrino masses M_i by having S develop a vacuum expectation value $\langle S\rangle$, spontaneously breaking a global lepton number symmetry. In such cases where an explicit mass term M_{ij} is forbidden, the matrix α_{ij} can be be made real and diagonal.

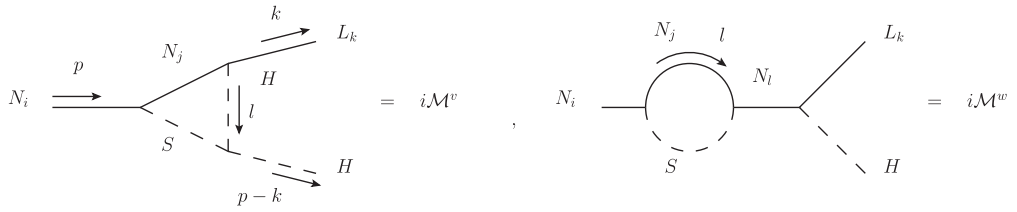


FIG. 1. Two hidden sector decay channels for RHN that contribute to the *CP* asymmetry. The superscripts "v" and "w" stand for the vertex and wave function diagrams.

$i\mathcal{M} = \Gamma_0 + \Gamma_1 I$, with $\Gamma_{0,1}$ the tree- and loop-level combinations of coupling constants and I the loop function, then the decay amplitude for the antiparticle is $i\bar{\mathcal{M}} = \Gamma_0^* + \Gamma_1^* I$, while the decay rates $\Gamma(N_i \to LH)$ and $\Gamma(N_i \to \overline{LH})$ are proportional to $|i\mathcal{M}|^2$ and $|i\bar{\mathcal{M}}|^2$ respectively. At tree level, the difference vanishes, but at the loop level, the *CP* asymmetry takes the schematic form

$$\epsilon_i \sim 2\frac{\mathrm{Im}\{\Gamma_0\Gamma_1^*\}}{|\Gamma_0|^2}\mathrm{Im}\{I\}. \tag{6}$$

Thus, at least for two-body decays, the *CP* asymmetry requires loops and a phase in the loop function itself. In standard leptogenesis, only the Yukawa λ_{ij} allows for this decay channel and can accommodate *CP* violation. In Higgs portal leptogenesis, additional lepton number-violating and *CP*-violating sources are present in the theory, specifically the α_{ij} coupling as discussed above. As a result, additional loop-induced decay channels open up, as displayed in Fig. 1.

The corresponding *CP* asymmetries will be discussed in the following subsections. We use the Majorana Feynman rules [21] for the RH neutrinos and determine the imaginary parts of the loop functions using the standard Cutkosky rules [22].

1. Vertex corrections: Two-body final states

The contribution of the diagram $i\mathcal{M}^v$ to the asymmetry is

$$\epsilon_i^v = 4\sum_j\left(\frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{ji}\beta\alpha_{ij}\}}{(\lambda^\dagger\lambda)_{ii}M_i}\mathrm{Im}\{I_{jLL}\} + \frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{ji}\beta\alpha_{ij}^*\}}{(\lambda^\dagger\lambda)_{ii}M_i}\mathrm{Im}\{I_{jRL}\}\right), \tag{7}$$

where I_{jLL} and I_{jRL} are loop function integrals. The vertex contribution splits into two halves, proportional to $\mathrm{Im}\{I_{jLL}\}$ and $\mathrm{Im}\{I_{jRL}\}$, corresponding to the mixing of the left and right chiralities of the Majorana fermions along the fermion lines, effectively leading to the two chirality chains

$$\begin{aligned}&\nu_R(\mathrm{RH}) \to \nu_R^C(\mathrm{LH}) \to L(\mathrm{LH}),\\ &\nu_R^C(\mathrm{LH}) \to \nu_R^C(\mathrm{LH}) \to L(\mathrm{LH}).\end{aligned} \tag{8}$$

Because the final leptons (assumed massless) have a definite chirality, the Yukawa coupling forces the next-to-last neutrino to be of the same chirality as the final lepton, left handed. The functions I_{jLL} and I_{jRL} correspond to the L-L-L and R-L-L chains respectively. The reason why the chirality chains do not combine owes to the fact that α_{ij} is neither real nor diagonal. In the vertex contribution, there are three possible cuts which lead to an imaginary part: cuts along the S/N lines, the H/S lines, and the H/N lines, each of which contains the two chirality chain contributions. Thus, each chirality chain function, $\mathrm{Im}\{I_{jLL}\}$ and $\mathrm{Im}\{I_{jRL}\}$, is the sum,

$$\mathrm{Im}\{I\} = \mathrm{Im}\{I\}^{S/N} + \mathrm{Im}\{I\}^{H/S} + \mathrm{Im}\{I\}^{H/N}. \tag{9}$$

It will be convenient to graphically represent the interference terms in $|\mathcal{M}|^2$ which contribute to the imaginary part in the form of *bubble* diagrams. For now we focus on the N/S cut which is shown in Fig. 2 [the full set of cuts is presented later in Fig. 4(a)]. The double line indicates external lines that are on shell by definition in $|\mathcal{M}|^2$, in this case the final lepton and Higgs, while the single line shows the Cutkosky cut. According to the Cutkosky rules, the N/S cut is given by

$$\begin{aligned}\mathrm{Im}\{I_{LL}\}^{S/N} = 2\pi^2\int\frac{d^4l}{(2\pi)^4}\frac{l\cdot k}{l^2}&\delta((l+k)^2 - M_j^2)\\ &\times\delta((l+k-p)^2 - m_S^2)\\ &\times\Theta(p^0 - l^0 - k^0)\Theta(k^0 + l^0).\end{aligned} \tag{10}$$

The delta functions impose the on-shell condition for the N/S lines, and the Heaviside functions Θ require these on-shell lines to be physical (timelike) processes. In other words, imposing positivity of the energies $p^0 - l^0 - k^0$ and $k^0 + l^0$ requires that the cut diagram corresponds to the decay $N_i \to N_j S$ followed by the scattering $N_j S \to LH$. Combining the energy constraints also restricts the individual energies $l^0 = M_i(r_{ji} - \sigma_i)/2$, and $M_i/2 > l^0 > -M_i/2$. The on-shell conditions in turn imply the quadratic constraint on the 3-momentum $\vec{l}$,

$$|\vec{l}|^2 + M_i|\vec{l}|y - \frac{M_i^2}{4}(\delta(1, r_{ji}, \sigma_i) - 1) = 0, \tag{11}$$

where we have introduced the shorthand notations

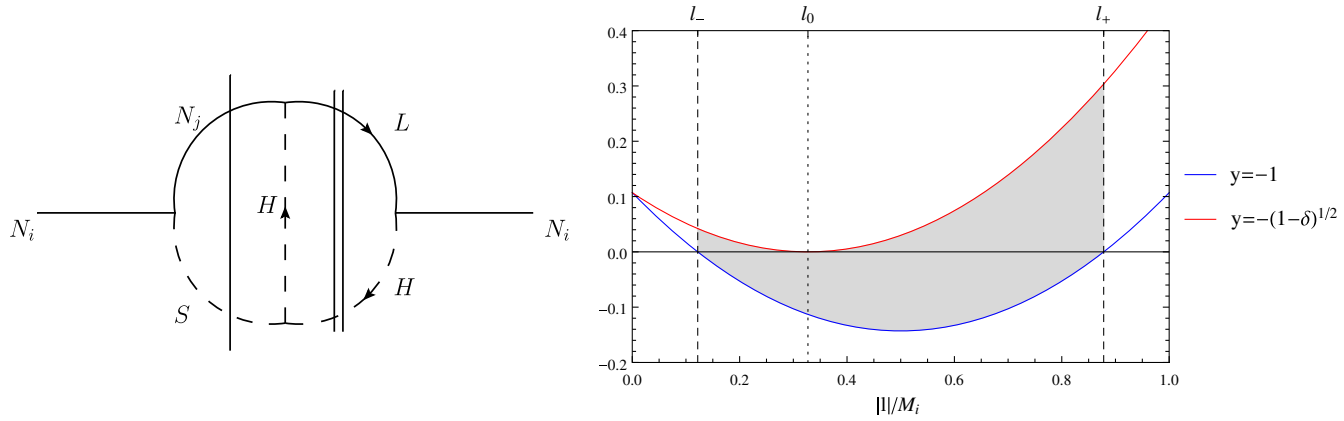


FIG. 2 (color online). The left plot graphically represents the N/S cut. The right figure exhibits the quadratic constraint in Eq. (11). The vertical dashed lines represent the momentum $|\vec{l}|$ upper and lower bounds, $|\vec{l}|_\pm = M_i(1 \pm \sqrt{\delta})/2$, which become $|\vec{l}|_0 = M_i\sqrt{1-\delta}/2$ at $y = -\sqrt{1-\delta}$. The imaginary part of the vertex correction is nonzero contributions only from the gray area.

$$r_{ji} = \frac{M_j^2}{M_i^2}, \qquad \sigma_i = \frac{m_S^2}{M_i^2},$$
$$\delta(a,b,c) = (a-b-c)^2 - 4bc, \qquad y = \cos\theta. \tag{12}$$

For simplicity below, δ without specified variables will implicitly be understood to mean $\delta(1, r_{ji}, \sigma_i)$, unless stated otherwise. The angle θ lies between the 3-momenta $\vec{l}$ and $\vec{k}$. The combined constraints on l^0 given above imply the equivalent constraint, $1 > r_{ji} - \sigma_i > -1$. Since the kinematics must allow the decay $N_i \to N_j S$, the latter constraint requires that $1 > \sqrt{r_{ji}} + \sqrt{\sigma_i}$. Importantly, we observe that the diagram only has an imaginary part for decays of the next-to-lightest neutrinos. Similarly, the imaginary part is nonvanishing provided the quadratic equation for $|\vec{l}|$ in (11) has real solutions, thus imposing the condition $\delta > 0$, here again satisfied if $1 > \sqrt{r_{ji}} + \sqrt{\sigma_i}$. The l^0 integration is trivial since its value is uniquely fixed. As usual, the remaining integration over $\vec{l}$ is split into the radial and angular parts. In spherical coordinates with $\vec{k}$ along the z axis, the azimuthal angle ϕ trivially integrates to 2π, and θ corresponds to the inclination angle of the spherical coordinate system. The leftover integrals over y and $|\vec{l}|$ are not independent because of the constraint (11). That constraint has been plotted in Fig. 2, where we see that the kinematics are constrained to the ranges $-1 \le y \le -\sqrt{1-\delta}$ and $M_i/2(1-\sqrt{\delta}) \le |\vec{l}| \le M_i/2(1+\sqrt{\delta})$. The integration is nonvanishing within this range, leading to the result,

$$\mathrm{Im}\{I_{LL}\}^{N/S} = \frac{1}{32\pi}\left[-\sqrt{\delta} + r_{ji}\log\left(\frac{\sqrt{\delta + 4r_{ji}\sigma_i} - \sqrt{\delta}}{\sqrt{\delta + 4r_{ji}\sigma_i} + \sqrt{\delta}}\right)\right],$$
$$1 \ge \sqrt{r_{ji}} + \sqrt{\sigma_i}. \tag{13}$$

Similar steps lead to the other chirality chain function, $\mathrm{Im}\{I_{jRL}\}^{N/S}$, for the N/S cut,

$$\mathrm{Im}\{I_{jRL}\}^{N/S} = \frac{\sqrt{r_{ji}}}{32\pi}\log\left(\frac{\sqrt{\delta + 4r_{ji}\sigma_i} - \sqrt{\delta}}{\sqrt{\delta + 4r_{ji}\sigma_i} + \sqrt{\delta}}\right),$$
$$1 \ge \sqrt{r_{ji}} + \sqrt{\sigma_i}, \tag{14}$$

as well as for the H/S cuts,

$$\mathrm{Im}\{I_{jRL}\}^{H/S} = \frac{\sqrt{r_{ji}}}{32\pi}\log\left|\frac{1-r_{ji}}{r_{ji}}\right|\Bigg|_{\sigma_i=0},$$
$$\mathrm{Im}\{I_{jLL}\}^{H/S} = \frac{1}{32\pi}\left(1 + r_{ji}\log\left|\frac{1-r_{ji}}{r_{ji}}\right|\right)\Bigg|_{\sigma_i=0}, \tag{15}$$

while the H/N cut gives a vanishing imaginary part $\mathrm{Im}\{I_{jLL}\}^{H/N} = \mathrm{Im}\{I_{jRL}\}^{H/N} = 0$. The notation $|_{\sigma_i=0}$ means that the imaginary part is only nonzero if $\sigma_i = 0$. Note that the N/S cut is divergent in the infrared limit $\sigma_i = 0$, where it is effectively equivalent to the decay $N_i \to N_j S$ followed by the scattering $N_j S \to LH$ mediated by the Higgs in the t channel. The divergence, due to radiating massless scalars in the infrared collinear limit, is canceled by including the appropriate three-body decays as discussed below.

2. Vertex corrections: Three-body final states

In Fig. 3, we show the two three-body final-state amplitudes for which the interference develops an imaginary part and contributes to the CP asymmetry. The three-body final-state CP asymmetry $\epsilon_i^{(3)}$ measures the difference $(\Gamma(N_i \to LHS) - \Gamma(N_i \to \overline{LH}S))/(\Gamma(N_i \to LH, LHS) + \Gamma(N_i \to \overline{LH}, \overline{LH}S))$, with the total RHN decay rate in the denominator. With the three-body final-state decay rate being subdominant due to the reduced phase space, we can approximate $\epsilon_i^{(3)}$ as

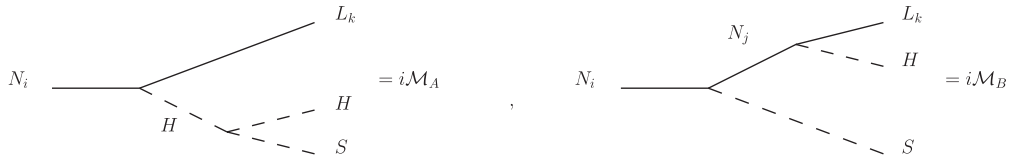


FIG. 3. Two three-body decay diagrams contributing to the *CP*-asymmetry at the same order in couplings as the loop-corrected two-body decays. As discussed in the text, their inclusion is important in ensuring that the full *CP* asymmetry is well-defined and free of infrared divergences.

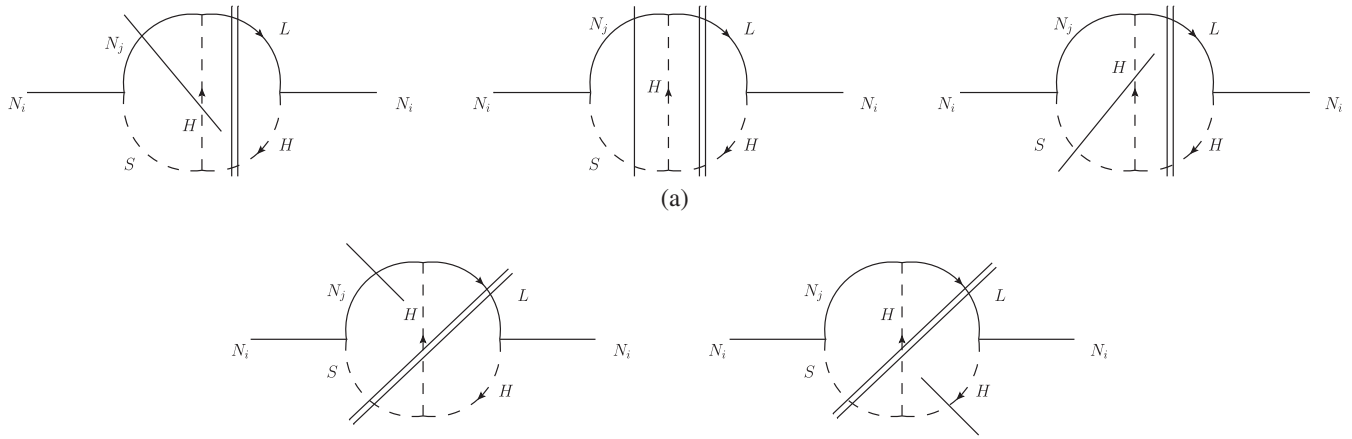


FIG. 4. The bubble diagrams enumerating the cuts that contribute to the *CP* asymmetry from both the vertex loop corrections (a) and three-body final states (b). The external state lines that are on shell by definition are cut with a double line, while the Cutkosky cuts are shown with a single line.

$$\epsilon_i^{(3)} \simeq \frac{\Gamma(N_i \to LHS) - \Gamma(N_i \to \overline{LH}S)}{\Gamma_i}. \tag{16}$$

In general, the three-body *CP*-asymmetry arises from both $i\mathcal{M}_A(i\mathcal{M}_B)^* \sim \mathcal{O}(\lambda^2\beta\alpha)$ and $|i\mathcal{M}_B|^2 \sim \mathcal{O}(\lambda^2\alpha^2)$. Only the former term, which enters at the same order as the vertex contribution, is of interest here. The two contributing cuts through the H and the N_j propagators are represented in Fig. 4(b) as bubble diagrams. The result is

$$\epsilon_i^{(3)} = 4\sum_j \left(\frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{ji}\beta\alpha_{ij}\}}{(\lambda^\dagger\lambda)_{ii}M_i}\mathrm{Im}\{\mathcal{I}_{jLL}\}^{(3)} + \frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{ji}\beta\alpha_{ij}^*\}}{(\lambda^\dagger\lambda)_{ii}M_i}\mathrm{Im}\{\mathcal{I}_{jRL}\}^{(3)} \right), \tag{17}$$

with

$$\begin{aligned}
\mathrm{Im}\{\mathcal{I}_{jLL}\}^{(3)} &= \frac{r_{ji}}{32\pi}\log\left(\frac{\sqrt{\delta + 4r_{ji}\sigma_i + 2\sigma_i} + \sqrt{\delta}}{\sqrt{\delta + 4r_{ji}\sigma_i + 2\sigma_i} - \sqrt{\delta}}\right) - \frac{1}{32\pi}\left(1 + r_{ji}\log\left|\frac{1 - r_{ji}}{r_{ji}}\right|\right)\Bigg|_{\sigma_i=0}, \\
\mathrm{Im}\{\mathcal{I}_{jRL}\}^{(3)} &= \frac{\sqrt{r_{ji}}}{32\pi}\left[-\sqrt{\delta} + \log\left(\frac{\sqrt{\delta + 4r_{ji}\sigma_i + 2\sigma_i} + \sqrt{\delta}}{\sqrt{\delta + 4r_{ji}\sigma_i + 2\sigma_i} - \sqrt{\delta}}\right)\right] - \frac{\sqrt{r_{ji}}}{32\pi}\log\left|\frac{1 - r_{ji}}{r_{ji}}\right|\Bigg|_{\sigma_i=0}.
\end{aligned} \tag{18}$$

The first terms in $\mathrm{Im}\{\mathcal{I}_{jLL}\}^{(3)}$ and $\mathrm{Im}\{\mathcal{I}_{jRL}\}^{(3)}$ come from cutting the N_j line, while the second terms come from cutting the S line, respectively combining with the N/S and H/S cuts of the tree-loop interference in (13), (14), (15), leading to the *corrected* vertex *CP* asymmetry ϵ_i^v,

$$\epsilon_i^v = \sum_j \left(\frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{ji}\beta\alpha_{ij}\}}{8\pi(\lambda^\dagger\lambda)_{ii}M_i}\mathcal{F}_{jLL}^v(r_{ji},\sigma_i) + \frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{ji}\beta\alpha_{ij}^*\}}{8\pi(\lambda^\dagger\lambda)_{ii}M_i}\mathcal{F}_{jRL}^v(r_{ji},\sigma_i) \right), \tag{19}$$

where

$$\mathcal{F}^{v}_{jLL}(r_{ji},\sigma_i) \equiv (-\sqrt{\delta} + r_{ji}\log G),$$
$$\mathcal{F}^{v}_{jRL}(r_{ji},\sigma_i) \equiv (-\sqrt{r_{ji}}\sqrt{\delta} + \sqrt{r_{ji}}\log G), \tag{20}$$

and

$$G \equiv \frac{\sqrt{\delta + 4r_{ji}\sigma_i + 2\sigma_i} + \sqrt{\delta}}{\sqrt{\delta + 4r_{ji}\sigma_i + 2\sigma_i} - \sqrt{\delta}} \frac{\sqrt{\delta + 4r_{ji}\sigma_i} - \sqrt{\delta}}{\sqrt{\delta + 4r_{ji}\sigma_i} + \sqrt{\delta}}. \tag{21}$$

Note that the infrared divergence at $\sigma_i = 0$ has disappeared, resulting from a cancellation between (13), (14), and (18). A simple graphical understanding of this cancellation emerges by comparing the N/S cut diagram of Fig. 4(a) with the N_j-line cut diagram of Fig. 4(b). In general they have different kinematics, but they coincide in the infrared limit where all the internal lines of the respective diagrams are allowed to be on shell, permitting the emission of soft particles. The inclusion of the two-body and three-body final-state contributions renders the CP asymmetry well defined.

Note also that the contributions that are nonvanishing only in the $\sigma_i = 0$ limit; i.e. those coming from the H/S cut in (15) and from the H-line cut in (18) also cancel out. This can again be understood by comparing the H/S-cut diagram of Fig. 4(a) and the H-cut of Fig. 4(b), for which the kinematics are identical.

3. Wave-function corrections

Once again, because of the Majorana nature of the right-handed neutrinos, there can be chirality mixing, leading to the following four chirality chains in the right-hand diagram of Fig. 1:

$$\begin{aligned}
&\nu_R(\mathrm{RH}) \to \nu_R(\mathrm{RH}) \to \nu^C_R(\mathrm{LH}) \to L(\mathrm{LH}),\\
&\nu_R(\mathrm{RH}) \to \nu^C_R(\mathrm{LH}) \to \nu^C_R(\mathrm{LH}) \to L(\mathrm{LH}),\\
&\nu^C_R(\mathrm{LH}) \to \nu_R(\mathrm{RH}) \to \nu^C_R(\mathrm{LH}) \to L(\mathrm{LH}),\\
&\nu^C_R(\mathrm{LH}) \to \nu^C_R(\mathrm{LH}) \to \nu^C_R(\mathrm{LH}) \to L(\mathrm{LH}).
\end{aligned} \tag{22}$$

The two first chains contain only one chirality flip, the third contains two flips, and the last contains none. Each chain will be labeled by the chiralities of the two first lines in the loop, i.e. the RR, RL, LR, and LL chains respectively. The asymmetry then takes the following form:

$$\epsilon^w_i = \sum_{l,j}\left(\frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{li}\alpha_{lj}\alpha^*_{ij}\}}{8\pi(\lambda^\dagger\lambda)_{ii}}\mathcal{F}^w_{jlLL} + \frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{li}\alpha^*_{lj}\alpha^*_{ij}\}}{8\pi(\lambda^\dagger\lambda)_{ii}}\mathcal{F}^w_{jlLR} + \frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{li}\alpha_{lj}\alpha_{ij}\}}{8\pi(\lambda^\dagger\lambda)_{ii}}\mathcal{F}^w_{jlRL} + \frac{\mathrm{Im}\{(\lambda^\dagger\lambda)_{li}\alpha^*_{lj}\alpha_{ij}\}}{8\pi(\lambda^\dagger\lambda)_{ii}}\mathcal{F}^w_{jlRR}\right). \tag{23}$$

Calculating the $\mathcal{F}^w$-loop functions is relatively simple as the imaginary part comes solely from the diagrams in which both lines in the loop are cut, which uniquely defines all the kinematics, trivializing the integrals. Thus, we will simply state the final results:

$$\mathcal{F}^w_{jlLL}(r_{ji},\sigma_i) \equiv \frac{\sqrt{\delta}}{2}\frac{\sqrt{\delta + 4r_{ji}}}{1 - r_{li}},$$
$$\mathcal{F}^w_{jlLR}(r_{ji},\sigma_i) \equiv \sqrt{\delta}\frac{\sqrt{r_{ji}}\sqrt{r_{li}}}{1 - r_{li}},$$
$$\mathcal{F}^w_{jlRL}(r_{ji},\sigma_i) \equiv \sqrt{\delta}\frac{\sqrt{r_{ji}}}{1 - r_{li}},$$
$$\mathcal{F}^w_{jlRR}(r_{ji},\sigma_i) \equiv \frac{\sqrt{\delta}}{2}\frac{\sqrt{r_{li}}\sqrt{\delta + 4r_{ji}}}{1 - r_{li}}. \tag{24}$$

As noted earlier, we have used the shorthand notation r_{ji}, and $\delta = (1 - r_{ji} - \sigma_i)^2 - 4r_{ji}\sigma_i$. The kinematic constraint remains the same as for the vertex correction, $1 > \sqrt{r_{ji}} + \sqrt{\sigma_i}$.

4. Summary

Most significantly, the kinematic constraint $1 > \sqrt{r_{ji}} + \sqrt{\sigma_i}$ prevents the lighter neutrino flavor, N_1, from having any CP-odd decays through the Higgs portal, since by definition $r_{21} > 1$. Only the heavier flavors, $N_{2,3}$, can contribute to the CP asymmetry through the hidden sector decays. For the remainder of this paper, we will generally focus on the minimal case with two heavy neutrinos N_1 and N_2, so that the hidden sector will play an important role through the decays of N_2. This presents us with the interesting possibility of taking N_1 parametrically light, where it could have other phenomenological consequences. At the same time, there is also the danger of significant washout of the asymmetry by scattering processes mediated by N_1. We will discuss the latter issue in some detail in subsequent sections.

The full CP asymmetry is obtained by combining the above results for ϵ^v_i (19) and ϵ^w_i (23). For the hierarchical $N_1 - N_2$ regime, with $M_2/M_1 \geq 10$ that will be of interest later, the CP asymmetry can be approximated by the following simple expressions (see Appendix A for details):

$$\epsilon_1 \sim \frac{3M_1\sum_\alpha m_\alpha}{8\pi v^2},$$
$$\epsilon_2 \sim \frac{3M_2\sum_\alpha m_\alpha}{16\pi v^2} + \left(\frac{\beta}{M_2} + \frac{|\alpha_{11}|}{2}(1-\sigma_2)\right)\frac{|\alpha_{21}|}{8\pi}\sqrt{\frac{M_1}{M_2}}(1-\sigma_2). \tag{25}$$

The index $\alpha = 1, 2, 3$, and m_α is the mass of the αth active neutrino. Assuming a normal hierarchy among the light neutrino masses, we set $\sum_\alpha m_\alpha \simeq m_3 \simeq \sqrt{\Delta m^2_{31}} \sim 0.05$ eV.

III. TWO-STAGE BOLTZMANN EVOLUTION

In minimal leptogenesis, the RHN sector $\lambda_{ij}N_jL_iH + M_iN_iN_i$ provides the ingredients for two of Sakharov's conditions to be satisfied; L is violated due to the presence of both M_i and λ_{ij}, while there are physical CP-odd phases in λ_{ij}. The third and final condition is satisfied dynamically as the expansion of the Universe provides a mechanism for L-violating processes to go out of equilibrium. For this to happen, the rate Γ_N of L-violating RHN decays must fall below the Hubble expansion rate H. This transition is controlled by the Gamow *equilibrium parameter*, $K = \Gamma_N/H(T = M)$ [23],[2] where the Hubble rate $H(T = M)$ sets the time scale, t_H, at which the equilibrium density becomes Boltzmann suppressed. Setting $K < 1$ ensures the particle lifetime is longer than the Hubble time, $\tau_N > t_H$, and an excess abundance develops. In that case, the rate of decays will be large compared to that of inverse decays in order for the neutrino abundance to be able to reach equilibrium, effectively putting the system out of equilibrium.

In Higgs portal leptogenesis, we require at least two Majorana neutrinos, and there are two major implications. On one hand, the CP asymmetry from N_2 decays (25) is enhanced for low masses and can in fact become the dominant contribution. This suggests the possibility of establishing a "lower energy" theory of leptogenesis, mainly controlled by N_2 physics. On the other hand, the two RHN flavors lead to a novel evolution in the total lepton asymmetry. In minimal leptogenesis, the lepton asymmetry is primarily generated in a temperature range near the lightest RHN mass, $T \sim M_1$, since the decays and scattering are out of equilibrium for lower temperatures. The difference here is that, even though most of the lepton asymmetry can be generated through N_2 decays and inverse decays at temperatures around $T \sim M_2$, the lighter neutrino flavor N_1 potentially remains in equilibrium and can mediate rapid washout of the N_2-generated asymmetry. These interactions will be studied carefully below, to identify regimes in which N_1 is sufficiently weakly coupled that these new washout processes are suppressed.

A. Boltzmann equations

In the minimal leptogenesis scenario, typically once the neutrino decays go out of equilibrium, all the scattering processes also go out of equilibrium. The new feature in HPL is the possibility of having scattering processes in equilibrium during the period that a CP asymmetry would be generated through out-of-equilibrium decays. The most significant are those L-violating scattering processes with an external N_1, the abundance of which is not Boltzmann suppressed. The scattering processes that have an external N_2 are of course suppressed by the N_2 abundance which rapidly falls off exponentially. Among the scattering processes that violate the lepton number by $\Delta L = 1$ units, we include the scattering $N_iL \leftrightarrow Q\bar{t}$ in the s channel and $N_iQ \leftrightarrow Lt$, $Nt \leftrightarrow LQ$ in the t channel. From the hidden sector, one includes the s-channel processes $N_iL \leftrightarrow HS$ mediated by a Higgs, and $N_iS \leftrightarrow LH$ mediated by a neutrino. In the t channel, one has $N_iS \leftrightarrow LH$ and $N_iH \leftrightarrow LS$ both mediated by a Higgs. A full treatment of neutrino-mediated scattering is complicated because of the λ_{ij} and α_{ij} flavor structures. For simplicity, we will ignore the flavor mixing in these processes with intermediate neutrinos, e.g. $\sigma_{N_iS\leftrightarrow LH} = \sum_j \sigma_{N_iS\leftrightarrow_j LH}$, and assume the processes are dominated by one flavor. This is sufficient for order-of-magnitude estimates. Note that, because of the α_{ij} coupling, we need to include interactions such as $N_iN_j \leftrightarrow HH$ mediated by S in the s channel, which can efficiently deplete the neutrino abundance, and in turn affect the lepton asymmetry washout [24].[3]

As for the $\Delta L = 2$ interactions, one has $LH \leftrightarrow \overline{LH}$ mediated by a neutrino in the s channel and $LL \leftrightarrow HH$, $\overline{LL} \leftrightarrow HH$ mediated by a neutrino in the t channel. We start by describing the Boltzmann equations for the lepton asymmetry, which are the most complex, and then review the neutrino abundance and the general features. Further technical details are contained in Appendixes B and C. Note that this work is concerned with the main dynamical features of the model presented above, focussing on the impact of the Higgs portal couplings. Thus, in deriving the Boltzmann equations, we study only the *total* lepton asymmetry, ignoring the often significant effects on individual lepton flavors [25]. For our purposes, it will also be sufficient to use the CP asymmetries calculated within zero-temperature field theory, although real-time thermal field theory provides a more complete formalism; see e.g. Ref. [26].

1. Lepton asymmetry

Starting with the lepton asymmetry equation, we have

[2]We will follow the literature and denote the Hubble-normalized N decay rate as K, while the modified Bessel function that generically appears in the thermal rates will consistently be written as $K_i(z)$, distinguished by the extra argument $z = M/T$.

[3]In the context of standard leptogenesis, $\Delta N = 2$ interactions, e.g. $N_1N_1 \to HH$ mediated by a lepton in the t channel, are negligible since they scale as λ^4 which is suppressed for $M_1 \lesssim 10^{15}$ GeV.

$$
\begin{aligned}
n_\gamma^{\rm eq} z_1 H \frac{\partial Y_{L-\bar{L}}}{\partial z_1} = \sum_i \Bigg[& \frac{Y_i}{Y_i^{\rm eq}} (\gamma^{\rm eq}_{N_i \to LH} - \gamma^{\rm eq}_{N_i \to \overline{LH}}) - \frac{Y_L}{Y_L^{\rm eq}} \gamma^{\rm eq}_{LH \to N_i} + \frac{Y_{\bar{L}}}{Y_L^{\rm eq}} \gamma^{\rm eq}_{\overline{LH} \to N_i} \\
& - \frac{Y_{L-\bar{L}}}{Y_L^{\rm eq}} \left(\frac{Y_i}{Y_i^{\rm eq}} \gamma^{\rm eq}_{N_i L \to Qt} + \frac{Y_i}{Y_i^{\rm eq}} \gamma^{\rm eq}_{N_i L \to HS} + 2\gamma^{\rm eq}_{N_i Q \to Lt} + 2\gamma^{\rm eq}_{N_i H \to LS} \right) \Bigg] \\
& - \frac{Y_{L-\bar{L}}}{Y_L^{\rm eq}} (\gamma^{\rm eq}_{N_1 S \to {}_1 LH} + \gamma^{\rm eq}_{N_2 S \to {}_1 LH} + \gamma^{\rm eq}_{N_2 S \to {}_2 LH}) \\
& + \frac{Y_1}{Y_1^{\rm eq}} (\gamma^{\rm eq,sub}_{N_1 S \to {}_2 LH} - \gamma^{\rm eq,sub}_{N_1 S \to {}_2 \overline{LH}}) - \frac{Y_L}{Y_L^{\rm eq}} \gamma^{\rm eq,sub}_{LH \to {}_2 N_1 S} + \frac{Y_{\bar{L}}}{Y_L^{\rm eq}} \gamma^{\rm eq,sub}_{\overline{LH} \to {}_2 N_1 S} \\
& - 2 \frac{Y_L}{Y_L^{\rm eq}} \sum_j \gamma^{\rm eq,sub}_{LH \to {}_j \overline{LH}} + 2 \frac{Y_{\bar{L}}}{Y_{\bar{L}}^{\rm eq}} \sum_j \gamma^{\rm eq,sub}_{\overline{LH} \to {}_j LH} \\
& - 2 \frac{Y_{L-\bar{L}}}{Y_L^{\rm eq}} \sum_j \gamma^{\rm eq}_{LL \to {}_j HH}.
\end{aligned} \tag{26}
$$

In this expression, we use the following notation:

$$
z_i = \frac{M_i}{T}, \qquad Y_i = \frac{n_i}{n_\gamma^{\rm eq}}, \qquad Y_i^{\rm eq} = \frac{3}{8} z_i^2 K_2(z_i), \tag{27}
$$

where $K_i(z)$ is a modified Bessel function of the second kind, along with the thermal cross sections,

$$
\begin{aligned}
\gamma^{\rm eq}_{i \to mn}(T) &= n_i^{\rm eq} \Gamma_{i \to mn} \left\langle \frac{M_i}{E} \right\rangle = n_i^{\rm eq} \frac{K_1(z_i)}{K_2(z_i)} \Gamma_{i \to mn}, \\
\gamma^{\rm eq}_{ij \to mn}(T) &= n_i^{\rm eq} n_j^{\rm eq} \langle v \sigma_{ij \to mn \cdots} \rangle \\
&= g_i g_j \frac{T^4}{32\pi^4} \int_{w_{\rm min}}^\infty dw \sqrt{w} K_1(\sqrt{w}) \hat{\sigma}_{ij \to mn} \left(w \frac{M_i^2}{z_i^2} \right),
\end{aligned} \tag{28}
$$

where $w = s/T^2$ and the reduced cross section $\hat{\sigma}$ is given by

$$
\begin{aligned}
\hat{\sigma}_{ij \to mn}(s) &= \frac{1}{s} \delta(s, m_i^2, m_j^2) \sigma_{ij \to mn}(s), \\
\delta(a,b,c) &= (a-b-c)^2 - 4bc.
\end{aligned} \tag{29}
$$

In Appendix B, we include further details about the Boltzmann equations and thermal cross sections. In writing the above equation, we have assumed CP invariance, $\gamma^{\rm eq}_{\overline{ij} \to \overline{mn}} = \gamma^{\rm eq}_{ij \to mn}$, in the scattering processes, along with CPT symmetry, $\gamma^{\rm eq}_{\overline{ij} \to \overline{mn}} = \gamma^{\rm eq}_{mn \to ij}$. CP-violating corrections appear in the scattering amplitudes at fourth order in the coupling constants, which is of higher order than we will consider here. That being said, it is necessary to make an exception when dealing with the subtracted rates as discussed below.

The superscript "*sub*" signifies that the process contains a real-intermediate-state (RIS) mediator that should be subtracted. Most famously, the process $LH \to \overline{LH}$ is mediated by a neutrino N_j in the s channel which can be on shell for a sufficiently high center-of-mass energy. The real intermediate state represents the physical process $LH \to N_j \to \overline{LH}$. However, these processes are already accounted for by decays and inverse decays and therefore need to be removed. Similarly, $N_1 S \to {}_2 LH$ has the real intermediate state $N_1 S \to N_2 \to LH$, which is also accounted for by decays and inverse decays. As it turns out, the t-channel process $N_i H \to LS$ also contains a real intermediate state, that needs to be removed. This point is explained in Appendix C, where explicit formulas for the cross sections are displayed. For simplicity, all quarks and leptons, as well as the Higgs and scalar S, are considered massless. This can be justified because leptogenesis necessarily occurs at temperatures above sphaleron decoupling, $T > T_{\rm sphaleron} \sim 160$ GeV $> m_H$ [27], though typically we shall take $M_2 > 1$ TeV. The singlet S mass is not yet stringently constrained, provided the Higgs portal coupling is not too large [28], but we will typically take it to be of the same order as the Higgs mass. With these simplifications, the t-channel processes $N_i Q \leftrightarrow Lt$ and $N_i t \to LQ$ have equal rates, and similarly for $N_i H \to LS$ and $N_i S \to LH$, which explains the factor of "2" sitting in front of these processes in Eq. (26) above.

2. Subtracted rates and real intermediate states

In this subsection, we summarize the procedure used to account for real intermediate states. The source terms in the Boltzmann equations are systematically expanded in each of the couplings, and one needs to avoid double counting the RIS contributions that appear in (naively) higher order scattering processes. Doing this consistently in standard leptogenesis requires the inclusion of all processes up to and including two-to-three scattering and the associated CP asymmetries [29]. For HPL, we will do the same, extending the analysis to account for real intermediate states coupling via both the Yukawa λ_{ij} and the singlet α_{ij} interactions. The relevant tree-level diagrams are displayed in Fig. 5,

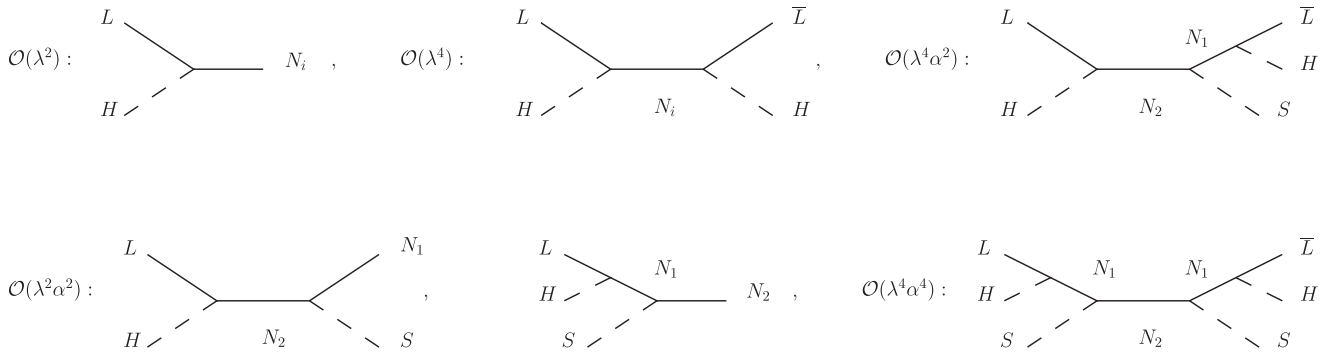


FIG. 5. Classes of decay and scattering diagrams and subdiagrams. We subtract the N_i real intermediate states coupling through α and λ order by order to avoid double counting, as discussed in the text.

although it is important to account also for loop corrections that contribute to the CP asymmetries in scattering. As already mentioned, we will focus on the impact of the additional singlet decay channel and ignore the issue of neutrino flavor mixing in scattering amplitudes, e.g. $\gamma^{\rm eq}_{LH\to\overline{LH}} = \sum_i \gamma^{\rm eq}_{LH\to_i\overline{LH}}$, which has been discussed in detail elsewhere. The RIS calculation of the s-channel cross sections is generally a nontrivial task once the flavor structure is taken into account. However, given this simplifying assumption, we can use the result $\gamma^{\rm sub} = \gamma - \gamma^{\text{on-shell}}$ [30].

To proceed to discuss the subtracted rates, we first make the following definitions associated with N_2 decays:

$$\begin{aligned}
&\Gamma_{2T} = \Gamma_2 + \Gamma_{N_2\to N_1 S}, \qquad \mathrm{Br}(N_2 \to LH) = \frac{1+\epsilon_2}{2}\frac{\Gamma_2}{\Gamma_{2T}},\\
&\Gamma_2 = \Gamma_{N_2\to LH} + \Gamma_{N_2\to\overline{LH}}, \quad \Rightarrow \quad \mathrm{Br}(N_2 \to \overline{LH}) = \frac{1-\epsilon_2}{2}\frac{\Gamma_2}{\Gamma_{2T}},\\
&\epsilon_2\Gamma_2 = \Gamma_{N_2\to LH} - \Gamma_{N_2\to\overline{LH}}, \qquad \mathrm{Br}(N_2 \to N_1 S) = \frac{\Gamma_{21}}{\Gamma_{2T}} = 1 - \frac{\Gamma_2}{\Gamma_{2T}},
\end{aligned} \tag{30}$$

where Γ_{21} is shorthand notation for $\Gamma_{N_2\to N_1 S}$, while N_1 only decays to leptons, so that

$$\begin{aligned}
&\Gamma_{1T} = \Gamma_1,\\
&\Gamma_1 = \Gamma_{N_1\to LH} + \Gamma_{N_1\to\overline{LH}}, \qquad \Rightarrow \qquad \mathrm{Br}(N_1 \to LH) = \tfrac{1+\epsilon_1}{2},\\
&\epsilon_1\Gamma_1 = \Gamma_{N_1\to LH} - \Gamma_{N_1\to\overline{LH}}, \qquad\qquad \mathrm{Br}(N_1 \to \overline{LH}) = \tfrac{1-\epsilon_1}{2}.
\end{aligned} \tag{31}$$

Building up the equations order by order, we have the following contributions:

(i) $\mathcal{O}(\lambda^2)$—one-to-two decays:

$$\begin{aligned}
n^{\rm eq}_\gamma z_1 H \frac{\partial Y_{L-\bar{L}}}{\partial z_1}\bigg|_{\lambda^2} &= \sum_{i=1,2}\left[\frac{Y_i}{Y_i^{\rm eq}}(\gamma^{\rm eq}_{N_i\to LH} - \gamma^{\rm eq}_{N_i\to\overline{LH}}) - \frac{Y_L}{Y_L^{\rm eq}}\gamma^{\rm eq}_{LH\to N_i} + \frac{Y_{\bar{L}}}{Y_L^{\rm eq}}\gamma^{\rm eq}_{\overline{LH}\to N_i}\right]\\
&= \left(\frac{Y_1}{Y_1^{\rm eq}} + 1\right)\epsilon_1\gamma^{\rm eq}_{D_1} + \left(\frac{Y_2}{Y_2^{\rm eq}} + 1\right)\epsilon_2\gamma^{\rm eq}_{D_2} - \frac{Y_{L-\bar{L}}}{2Y_L^{\rm eq}}(\gamma^{\rm eq}_{D_1} + \gamma^{\rm eq}_{D_2}).
\end{aligned} \tag{32}$$

This Boltzmann equation suffers from the well-known flaw that it does not respect Sakharov conditions. Indeed, even in equilibrium $Y_i = Y_i^{\rm eq}$, a lepton asymmetry may be generated because of the nonvanishing source term. This is related to the fact that at $\mathcal{O}(\lambda^2)$ the set of interactions is incomplete because $\Delta L = 2$ rates at (naive) $\mathcal{O}(\lambda^4)$ contain real-intermediate-state contributions that are in fact of order $\mathcal{O}(\lambda^2)$ and need to be included; see below.

(ii) $\mathcal{O}(\lambda^4)$—$\Delta L = 2$ scattering (two to two):

$$\begin{aligned}
n^{\rm eq}_\gamma z_1 H \frac{\partial Y_{L-\bar{L}}}{\partial z_1}\bigg|_{\lambda^4} &= \sum_{j=1,2}\left[-2\frac{Y_L}{Y_L^{\rm eq}}\gamma^{\rm eq,sub}_{LH\underset{j}{\to}\overline{LH}} + 2\frac{Y_{\bar{L}}}{Y_{\bar{L}}^{\rm eq}}\gamma^{\rm eq,sub}_{\overline{LH}\underset{j}{\to}LH}\right]\\
&= -2\epsilon_1\gamma^{\rm eq}_{D_1} - 2\epsilon_2\frac{\Gamma_2}{\Gamma_{2T}}\gamma^{\rm eq}_{D_2} - \frac{Y_{L-\bar{L}}}{Y_L^{\rm eq}}(2\gamma^{\rm eq,sub,2}_{\Delta L=2} + 2\gamma^{\rm eq,sub,1}_{\Delta L=2}),
\end{aligned} \tag{33}$$

where we used the definitions for the subtracted $\Delta L = 2$ rates,

$$\begin{aligned}
\gamma^{\text{eq,on-shell}}_{LH\to_2\overline{LH}} &= \gamma^{\text{eq}}_{LH\to N_2}\text{Br}(N_2\to\overline{LH}) = \left(\frac{1-\epsilon_2}{2}\right)^2\frac{\Gamma_2}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2} \approx \frac{1-2\epsilon_2}{4}\frac{\Gamma_2}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2},\\
\gamma^{\text{eq,sub}}_{LH\to_2\overline{LH}} &= \gamma^{\text{eq,sub},(2)}_{\Delta L=2} - \frac{\epsilon_2}{2}\frac{\Gamma_2}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2},\\
\gamma^{\text{eq,sub},(2)}_{\Delta L=2} &\equiv \gamma^{\text{eq}}_{LH\to_2\overline{LH}} - \frac{1}{4}\frac{\Gamma_2}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2},\\
\Delta\gamma^{\text{eq,sub},(2)}_{\Delta L=2} &\equiv \gamma^{\text{eq,sub}}_{LH\to_2\overline{LH}} - \gamma^{\text{eq,sub}}_{\overline{LH}\to_2 LH} = -\epsilon_2\frac{\Gamma_2}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2},
\end{aligned} \tag{34}$$

and

$$\begin{aligned}
\gamma^{\text{eq,on-shell}}_{LH\to_1\overline{LH}} &= \gamma^{\text{eq}}_{LH\to N_1}\text{Br}(N_1\to\overline{LH}) = \left(\frac{1-\epsilon_1}{2}\right)^2\gamma^{\text{eq}}_{D_1} \approx \frac{1-2\epsilon_1}{4}\gamma^{\text{eq}}_{D_2},\\
\gamma^{\text{eq,sub}}_{LH\to_1\overline{LH}} &= \gamma^{\text{eq,sub},(1)}_{\Delta L=2} - \frac{\epsilon_1}{2}\gamma^{\text{eq}}_{D_1},\\
\gamma^{\text{eq,sub},(1)}_{\Delta L=2} &\equiv \gamma^{\text{eq}}_{LH\to_1\overline{LH}} - \frac{1}{4}\gamma^{\text{eq}}_{D_1},\\
\Delta\gamma^{\text{eq,sub},(1)}_{\Delta L=2} &\equiv \gamma^{\text{eq,sub}}_{LH\to_1\overline{LH}} - \gamma^{\text{eq,sub}}_{\overline{LH}\to_1 LH} = -\epsilon_1\gamma^{\text{eq}}_{D_1}.
\end{aligned} \tag{35}$$

The unsubtracted $\Delta L = 2$ rates are CP symmetric at $\mathcal{O}(\lambda^2)$, which implies that the subtracted rates are in fact CP asymmetric. The functions $\gamma^{\text{eq,sub},(i)}_{\Delta L=2}$ are the CP-conserving parts of the subtracted $\Delta L = 2$ rates and have been plotted in Figs. 6(a) and 6(b), showing that the CP-conserving subtracted rates are negligible.

Note that the RIS CP asymmetry that comes from the N_1-mediated $\Delta L = 2$ rate corrects the flaw above in the Boltzmann equation (32). The $\Delta L = 2$ rates mediated by N_2, however, do not fully correct the above Boltzmann equation. This is in contrast to the standard case, the reason being that $\text{Br}(N_2\to N_1 S) > 0$, so there is a second decay channel. Unlike in the standard case, for HPL, obtaining a consistent set of Boltzmann equations requires including the RIS contributions from higher order $\Delta L = 2$ scattering processes; see Eq. (36) below.

(iii) $\mathcal{O}(\lambda^4\alpha^2)$—$\Delta L = 2$ scattering (two to three):

$$\begin{aligned}
n^{\text{eq}}_\gamma z_1 H\left.\frac{\partial Y_{L-\bar{L}}}{\partial z_1}\right|_{\lambda^4\alpha^2} &= -4\frac{Y_L}{Y^{\text{eq}}_L}\gamma^{\text{eq,sub}}_{LH\to_2\overline{LH}S} + 4\frac{Y_{\bar{L}}}{Y^{\text{eq}}_L}\gamma^{\text{eq,sub}}_{\overline{LH}\to_2 LHS}\\
&= -2\epsilon_2\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2} - 4\frac{Y_{L-\bar{L}}}{Y^{\text{eq}}_L}\gamma^{\text{eq},(2-3)}_{\Delta L=2}\\
&\quad - 2\Delta\gamma^{\text{eq,sub}}_{\Delta L=1} - 2\epsilon_1\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2},
\end{aligned} \tag{36}$$

where we have defined

$$\begin{aligned}
\gamma^{\text{eq,on-shell}}_{LH\to_2\overline{LH}S} &= \gamma^{\text{eq,sub}}_{LH\to_2 N_1 S}\frac{1-\epsilon_1}{2} + \frac{(1-\epsilon_1)(1-\epsilon_2)}{4}\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2},\\
\gamma^{\text{eq,sub}}_{LH\to_2\overline{LH}S} &= \gamma^{\text{eq}}_{LH\to_2\overline{LH}S} - \gamma^{\text{eq},(2-3)}_{\Delta L=2} + \frac{1}{4}\Delta\gamma^{\text{eq,sub}}_{\Delta L=1} + \frac{\epsilon_1+\epsilon_2}{4}\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\text{eq}}_{D_2},\\
\gamma^{\text{eq},(2-3)}_{\Delta L=2} &= \gamma^{\text{eq}}_{LH\to_2\overline{LH}S} - \frac{1}{4}\gamma^{\text{eq}}_{\Delta L=1},
\end{aligned} \tag{37}$$

ignoring terms of order $\mathcal{O}(\lambda^6\alpha^2)$ and above. The functions $\gamma^{\text{eq}}_{\Delta L=1}$ and $\Delta\gamma^{\text{eq,sub}}_{\Delta L=1}$ are the CP-conserving and CP-violating parts of the $N_1 S\to_2 LH$ scattering terms, as is defined in Eq. (39) below.

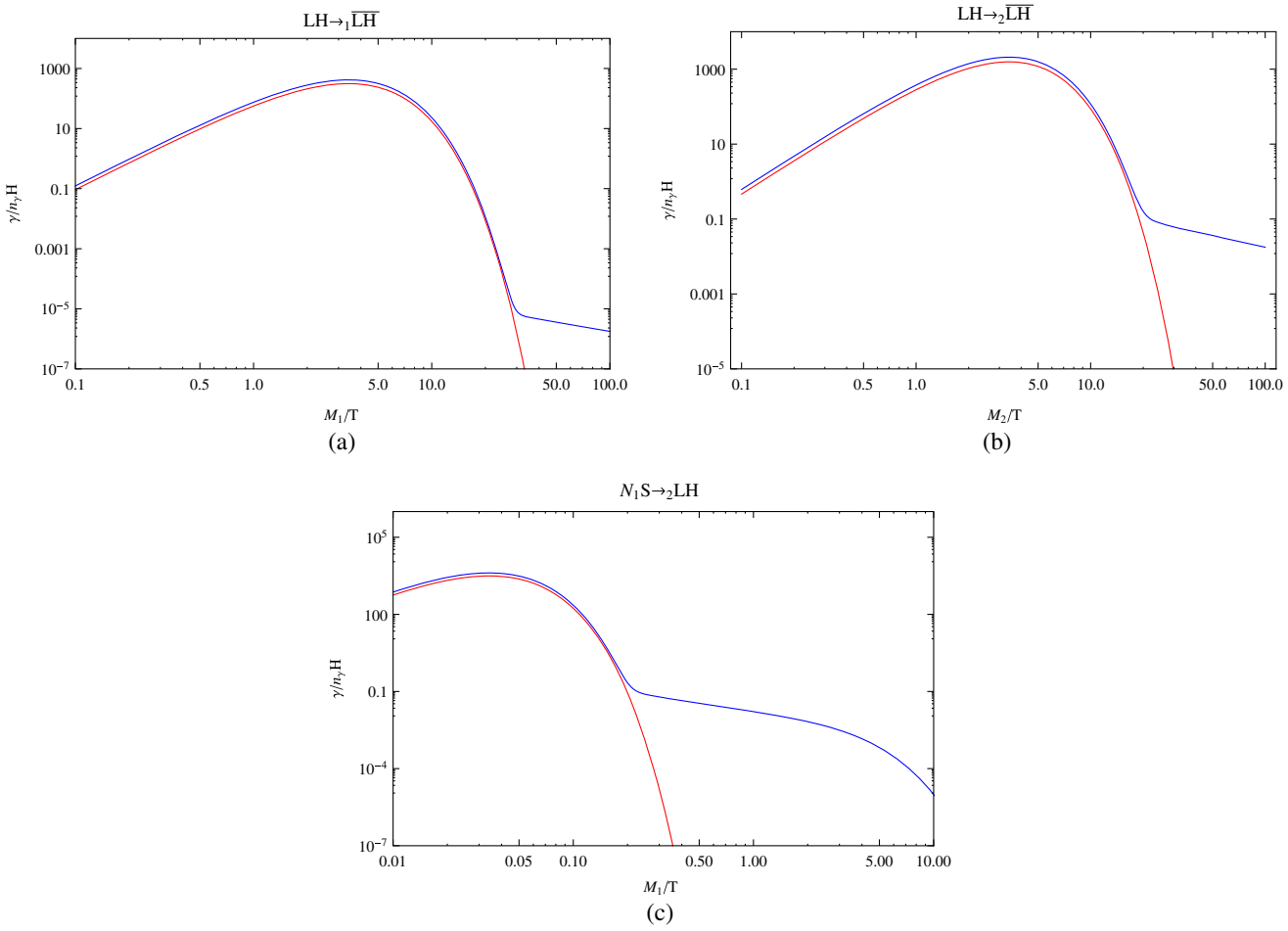


FIG. 6 (color online). Plots of the unsubtracted thermal rates (blue), along with their on-shell parts (red). The plots in (a), (b), and (c) are for $LH \underset{1}{\rightarrow} \overline{LH}$, $LH \underset{2}{\rightarrow} \overline{LH}$, and $N_1 S \underset{2}{\rightarrow} LH$ respectively, generated using $\lambda_1 \sim 0.002$, $\lambda_2 \sim 0.07$, $\alpha_{12} \sim 0.07$, $M_2 = 10^{10}$ GeV, and $M_1 = 10^8$ GeV.

In Eq. (36), the factors of 4 arise on accounting for both $LH \underset{2}{\rightarrow} \overline{LH}\, S$ and $LHS \underset{2}{\rightarrow} \overline{LH}$, which are CPT conjugates of each other. The first term on the first line is exactly what is needed to combine with the RIS term from Eq. (33) and correct the flaw in (32). The second line of Eq. (36) in fact contains terms that correspond to RIS contributions at $\mathcal{O}(\lambda^2\alpha^2)$.

(iv) $\mathcal{O}(\lambda^2\alpha^2)$—$\Delta L = 1$ scattering (two-to-two) and one-to-three decays:

As noted above, at $\mathcal{O}(\lambda^4\alpha^2)$ there are additional uncompensated terms. This is because the scattering processes at that order contain real intermediate states of $\mathcal{O}(\lambda^2\alpha^2)$, the same order as the $N_1 S \underset{2}{\rightarrow} LH$ scattering process and the $N_2 \underset{1}{\rightarrow} LHS$ three-body final-state decay rate. To include the scattering properly, it is necessary to use the subtracted rate since the on-shell piece is equivalent to $N_1 S \to N_2 \to LH$ which has already been accounted for. We have

$$
\begin{aligned}
n_\gamma^{\rm eq} z_1 H \frac{\partial Y_{L-\bar{L}}}{\partial z_1}\bigg|_{\lambda^2\alpha^2} &= \frac{Y_1}{Y_1^{\rm eq}}\left(\gamma^{\rm eq,sub}_{N_1 S \underset{2}{\rightarrow} LH} - \gamma^{\rm eq,sub}_{N_1 S \underset{2}{\rightarrow} \overline{LH}}\right) - \frac{Y_L}{Y_L^{\rm eq}}\gamma^{\rm eq,sub}_{LH \underset{2}{\rightarrow} N_1 S} + \frac{Y_{\bar{L}}}{Y_L^{\rm eq}}\gamma^{\rm eq,sub}_{\overline{LH} \underset{2}{\rightarrow} N_1 S} \\
&\quad + \frac{Y_2}{Y_2^{\rm eq}}\left(\gamma^{\rm eq}_{N_2 \rightarrow LHS} - \gamma^{\rm eq}_{N_2 \rightarrow \overline{LHS}}\right) - \frac{Y_L}{Y_L^{\rm eq}}\gamma^{\rm eq}_{N_2 \rightarrow \overline{LHS}} + \frac{Y_{\bar{L}}}{Y_L^{\rm eq}}\gamma^{\rm eq}_{N_2 \rightarrow LHS} \\
&= \left(\frac{Y_1}{Y_1^{\rm eq}} + 1\right)\Delta\gamma^{\rm eq,sub}_{\Delta L=1} - \frac{Y_{L-\bar{L}}}{2Y_L^{\rm eq}}\gamma^{\rm eq,sub}_{\Delta L=1} + \left(\frac{Y_2}{Y_2^{\rm eq}} + 1\right)\epsilon_1 \gamma_{N_2 \rightarrow N_1 S} - \frac{Y_{L-\bar{L}}}{2Y_L^{\rm eq}}\left(\gamma^{\rm eq}_{N_2 \rightarrow LHS} + \gamma^{\rm eq}_{N_2 \rightarrow \overline{LHS}}\right), \qquad (38)
\end{aligned}
$$

where

$$\gamma^{\rm eq,on\text{-}shell}_{N_1 S\underset{2}{\to} LH} = \gamma^{\rm eq}_{N_1 S\to N_2}{\rm Br}(N_2\to LH) = \frac{1+\epsilon_2}{2}\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\rm eq}_{D_2},$$
$$\gamma^{\rm eq,sub}_{N_1 S\underset{2}{\to} LH} = \gamma^{\rm eq}_{N_1 S\underset{2}{\to} LH} - \frac{1}{2}\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\rm eq}_{D_2} - \frac{\epsilon_2}{2}\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\rm eq}_{D_2},$$
$$\gamma^{\rm eq,sub}_{\Delta L=1} \equiv \gamma^{\rm eq,sub}_{N_1 S\underset{2}{\to} LH} + \gamma^{\rm eq,sub}_{N_1 S\underset{2}{\to}\overline{LH}} = \gamma^{\rm eq}_{\Delta L=1} - \frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\rm eq}_{D_2},$$
$$\Delta\gamma^{\rm eq,sub}_{\Delta L=1} \equiv \gamma^{\rm eq,sub}_{N_1 S\underset{2}{\to} LH} - \gamma^{\rm eq,sub}_{N_1 S\underset{2}{\to}\overline{LH}} = \Delta\gamma^{\rm eq}_{\Delta L=1} - \epsilon_2\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\rm eq}_{D_2}. \tag{39}$$

The function $\gamma^{\rm eq}_{\Delta L=1}/2$ determines the *CP*-conserving part of the scattering rate and has been plotted in Fig. 6(c), along with the RIS rate $\gamma^{\rm eq}_{D_2}\Gamma_{21}/(2\Gamma_{2T})$. This shows that $\gamma^{\rm eq,sub}_{\Delta L=1}$ is negligible. The *CP* asymmetry in the scattering, $\Delta\gamma^{\rm eq}_{\Delta L=1}$, is largely inherited from the on-shell part in such way that $\Delta\gamma^{\rm eq,sub}_{\Delta L=1}$ is again negligible. We can convince ourselves of this by taking the ratio $\Delta\gamma^{\rm eq,sub}_{\Delta L=1}/\gamma^{\rm eq}_{\Delta L=1} \approx \Delta\gamma^{\rm eq}_{\Delta L=1}/\gamma^{\rm eq}_{\Delta L=1} - \epsilon_2 \approx 0$, as the *CP* asymmetry in $\Delta L = 1$ scatterings is equal to that of the neutrino leptonic decays. The *CP* asymmetry in three-body final-state decay rates can only come from the kinematic point where the intermediate line is on shell, which means

$$\gamma^{\rm eq,on\text{-}shell}_{N_2\to LHS} = \gamma^{\rm eq}_{N_2\to N_1 S}\frac{1+\epsilon_1}{2},$$
$$\Delta\gamma^{\rm eq}_{N_2\to LHS} = \gamma^{\rm eq,on\text{-}shell}_{N_2\to LHS} - \gamma^{\rm eq,on\text{-}shell}_{N_2\to\overline{LH}S} = \epsilon_1\gamma_{N_2\to N_1 S}. \tag{40}$$

Equation (38) displays the same flaw as Eq. (32) in failing to follow Sakharov's criteria. Following the same logic as at $\mathcal{O}(\lambda^2)$, it is necessary to include (naively) higher order contributions, namely $\Delta L = 2$ scattering processes that contain RIS at $\mathcal{O}(\lambda^2\alpha^2)$. Indeed, the second-to-last line of Eq. (36) fully corrects this problem at the scattering level. The last line of the same equation *partially* corrects the corresponding flaw at the level of three-body final-state decays. To correct this rate completely, we need to consider yet higher order interactions, $\mathcal{O}(\lambda^4\alpha^4)$; see below.

(v) $\mathcal{O}(\lambda^4\alpha^4)$—$\Delta L = 2$ scattering (three-to-three): The three-to-three scattering process $LHS\to_2 \overline{LH}\,S$ contains a real intermediate state,

$$\gamma^{\rm eq,son\text{-}shell}_{LHS\underset{2}{\to}\overline{LH}S} \supset {\rm Br}(LH\to N_1)\gamma^{\rm eq}_{N_1 S\to N_2}{\rm Br}(N_2\to N_1 S) \times {\rm Br}(N_1\to\overline{LH}) \supset \frac{\epsilon_1}{2}\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\rm eq}_{N_2\to N_1 S}, \tag{41}$$

so that

$$n^{\rm eq}_\gamma z_1 H\frac{\partial Y_{L-\bar L}}{\partial z_1} = -2\frac{Y_L}{Y^{\rm eq}_L}\gamma^{\rm eq,sub}_{LHS\underset{2}{\to}\overline{LH}S} + 2\frac{Y_{\bar L}}{Y^{\rm eq}_L}\gamma^{\rm eq,sub}_{\overline{LH}S\underset{2}{\to}LHS} \supset -2\epsilon_1\frac{\Gamma_{21}}{\Gamma_{2T}}\gamma^{\rm eq}_{N_2\to N_1 S}, \tag{42}$$

which combines with the $-2\epsilon_1\gamma^{\rm eq}_{D_2}\Gamma_{21}/\Gamma_{2T}$ term of Eq. (36), leading to the combination $-2\epsilon_1\gamma^{\rm eq}_{N_2\to N_1 S}$ that ultimately corrects the above flaw at the three-body final-state decay level.

In conclusion, one obtains the correct Boltzmann equations at order $\mathcal{O}(\lambda^2)$, by combining Eqs. (32), (33) and (36) at order $\mathcal{O}(\lambda^2)$, $\mathcal{O}(\lambda^4)$, and $\mathcal{O}(\lambda^4\alpha^2)$ respectively. If one wishes to include the $\Delta L = 1$ scattering and decays at order $\mathcal{O}(\lambda^2\alpha^2)$, it is necessary to combine the contributions of $\mathcal{O}(\lambda^2\alpha^2)$, $\mathcal{O}(\lambda^4\alpha^2)$, and $\mathcal{O}(\lambda^4\alpha^4)$ in Eqs. (38), (36), and (42).

The need to include all these varied contributions to obtain the correct Boltzmann equations should not come as a surprise. Since there are two N_2 decay channels, whenever the decay $N_2\to LH$ is part of a scattering process, we can write down an additional scattering diagram which has the $N_2\to N_1 S\to LH$ decay chain as a subdiagram. Because both $N_2\to LH$ and $N_2\to N_1 S\to LH$ can happen on shell, they both contribute at the same order and therefore combine to provide a complete set of scattering contributions, complete in the sense that ${\rm Br}(N_2\to LH)+{\rm Br}(N_2\to N_1 S) = 1$. This explains the necessity to include all terms of both $\mathcal{O}(\lambda^4)$ and $\mathcal{O}(\lambda^4\alpha^2)$ to obtain the correct Boltzmann equations at $\mathcal{O}(\lambda^2)$.

We can also understand this conclusion at the level of unitarity and *CPT* invariance, which requires that $\sum_j|\mathcal{M}(i\to j)|^2 = \sum_j|\mathcal{M}(j\to i)|^2 = \sum_j|\mathcal{M}(\bar i\to\bar j)|^2$. At $\mathcal{O}(\lambda_1^4)$ one has

$$\begin{aligned}|\mathcal{M}(HL\to_1 X)|^2|_{\mathcal{O}(\lambda_1^4)} &= |\mathcal{M}(HL\to N_1)|^2 + |\mathcal{M}(HL\underset{1}{\to}LH)^{\rm sub}|^2 + |\mathcal{M}(HL\underset{1}{\to}\overline{HL})^{\rm sub}|^2\\ &= |\mathcal{M}(HL\to N_1)|^2 + |\mathcal{M}(HL\underset{1}{\to}LH)|^2 + |\mathcal{M}(HL\underset{1}{\to}\overline{HL})|^2\\ &\quad - |\mathcal{M}(HL\to N_1)|^2{\rm Br}(N_1\to LH) - |\mathcal{M}(HL\to N_1)|^2{\rm Br}(N_1\to\overline{LH})\\ &= |\mathcal{M}(HL\underset{1}{\to}LH)|^2 + |\mathcal{M}(HL\underset{1}{\to}\overline{HL})|^2.\end{aligned} \tag{43}$$

At this order both $|\mathcal{M}(HL\to_1 LH)|^2$ and $|\mathcal{M}(HL\to_1 \overline{HL})|^2$ are CP symmetric in which case unitarity and CPT invariance are straightforwardly satisfied. Now, at $\mathcal{O}(\lambda_2^4\alpha_{21}^2)$, it is necessary to include higher order scattering processes in order to obtain the same conclusion,

$$\begin{aligned}
|\mathcal{M}(HL\to_2 X)|^2|_{\mathcal{O}(\lambda_2^4\alpha_{21}^2)} =& |\mathcal{M}(HL\to N_2)|^2 + |\mathcal{M}(HL\underset{2}{\to} N_1 S)^{\rm sub}|^2 \\
&+ |\mathcal{M}(HL\underset{2}{\to} LH)^{\rm sub}|^2 + |\mathcal{M}(HL\underset{2}{\to} \overline{HL})^{\rm sub}|^2 \\
&+ |\mathcal{M}(HL\underset{2}{\to} LHS)^{\rm sub}|^2 + |\mathcal{M}(HL\underset{2}{\to} \overline{HL}\, S)^{\rm sub}|^2 \\
=& |\mathcal{M}(HL\to N_2)|^2 + |\mathcal{M}(HL\underset{2}{\to} N_1 S)^{\rm sub}|^2 + |\mathcal{M}(HL\underset{2}{\to} LH)|^2 + |\mathcal{M}(HL\underset{2}{\to} \overline{HL})|^2 \\
&+ |\mathcal{M}(HL\underset{2}{\to} LHS)|^2 + |\mathcal{M}(HL\underset{2}{\to} \overline{HL}\, S)|^2 \\
&- |\mathcal{M}(HL\to N_2)|^2({\rm Br}(N_2\to LH) + {\rm Br}(N_2\to \overline{LH})) \\
&- (|\mathcal{M}(HL\underset{2}{\to} N_1 S)^{\rm sub}|^2 + |\mathcal{M}(HL\to N_2)|^2 {\rm Br}(N_2\to N_1 S)) \\
=& |\mathcal{M}(HL\underset{2}{\to} LH)|^2 + |\mathcal{M}(HL\underset{2}{\to} \overline{HL})|^2 + |\mathcal{M}(HL\underset{2}{\to} LHS)|^2 + |\mathcal{M}(HL\underset{2}{\to} \overline{HL}\, S)|^2, \qquad (44)
\end{aligned}$$

where we have used the relation

$$|\mathcal{M}(HL\underset{2}{\to} LHS)^{\rm sub}|^2 = |\mathcal{M}(HL\underset{2}{\to} N_1 S)|^2 {\rm Br}(N_1\to LH) \qquad (45)$$

and further split the rate $|\mathcal{M}(HL\to_2 N_1 S)|^2$ into the subtracted and RIS parts. Recall that ${\rm Br}(N_1\to LH) + {\rm Br}(N_1\to \overline{LH}) = 1$. Thus, the unitarity+$CPT$ constraint is again consistently satisfied at this order.

We are now in a position to summarize the final Boltzmann equations. In practice, we can ignore the one-to-three decay rates, which are numerically subdominant compared to the one-to-two decays, and similarly we can ignore the two-to-three scatterings, which are subdominant compared to the two-to-two rates. The subtracted rates $\gamma^{\rm eq,sub}_{N_1 S\to LH}$ and $\gamma^{\rm eq,sub}_{\overline{LH}\to iLH}$ can also be ignored as suggested by the plots in Fig. 6.

To simplify the above discussion, we have considered the subset of interactions that contain the α and λ coupling constants. In this paper, we also consider the set of interactions involving the coupling β. However, among the set of scatterings one considers, there is no additional real intermediate state from this source, and we can directly rewrite the Boltzmann equation for the final lepton asymmetry as

$$z_1 \frac{\partial Y_{L-\bar L}}{\partial z_1} = \epsilon_1 D_1\left(\frac{Y_1}{Y_1^{\rm eq}} - 1\right) + \epsilon_2 D_2\left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right) - Y_{L-\bar L}(W_{ID_1} + W_{S_1} + W_{ID_2} + W_{S_2}), \qquad (46)$$

where we have used the notation of the decay, scattering, and washout functions D, W, and S, defined in Appendix B,

$$D_i = \frac{\gamma_{D_i}^{\rm eq}}{n_\gamma^{\rm eq} H} = K_i z_i^2 Y_i^{\rm eq}\frac{K_1(z_i)}{K_2(z_i)}, \qquad W_{ID_i} = \frac{1}{2Y_L^{\rm eq}} D_i, \qquad Y_i^{\rm eq} = \frac{3}{8} z_i^2 K_2(z_i). \qquad (47)$$

The equilibrium parameter for leptonic decays K_i is defined as

$$K_i = \frac{\Gamma_i}{H(T=M_i)} = \frac{\tilde m_i}{m_*}, \qquad (48)$$

where the effective light neutrino mass scales $\tilde m_i$ and m_* are

$$\tilde m_i = \frac{(\lambda^\dagger\lambda)_{ii} v^2}{M_i}, \qquad m_* = 8\pi v^2\sqrt{\frac{8\pi^3 g_*}{90 M_p^2}} \sim 1.05\times 10^{-3}\ {\rm eV}, \qquad (49)$$

emerging from the seesaw mechanism, with $g_* \sim 100$ the total number of degrees of freedom. We assume here the normal hierarchy among light neutrino masses. The washout functions are written in terms of the scattering function, $S_{ia\to mn} = \gamma^{\rm eq}_{ia\to mn}/(n_\gamma^{\rm eq} H)$,

$$W_{S_1} = \frac{1}{Y_L^{\rm eq}}(2S_{N_1 t\to LQ} + S_{N_1 H\to LS} + S_{N_1 S\to{}_1 LH}) + \frac{Y_1}{Y_L^{\rm eq} Y_1^{\rm eq}}(S_{N_1 L\to Qt} + S_{N_1 L\to HS}),$$
$$W_{S_2} = \frac{1}{Y_L^{\rm eq}}(2S_{N_2 t\to LQ} + S_{N_2 H\to LS} + S_{N_2 S\to{}_1 LH} + S_{N_2 S\to{}_2 LH}) + \frac{Y_2}{Y_L^{\rm eq} Y_2^{\rm eq}}(S_{N_2 L\to Qt} + S_{N_2 L\to HS}). \quad (50)$$

The functions D, S, and W have been defined to facilitate writing the Boltzmann equations in a manner that is independent of the choice of reference mass scale in the definition of the time variable. The results of integrating the Boltzmann equations can be qualitatively understood by considering the transition points where various rates go in and out of equilibrium. As discussed in Appendix B, this is conveniently tracked with the *thermal equilibrium parameters* $\mathcal{K}$. For decays $N_i \to LH + \overline{LH}$, one has $\mathcal{K}_i = \langle\Gamma_i\rangle/H = D_i/Y_i^{\rm eq}$; for inverse decays $LH + \overline{LH} \to N_i$, one has $\mathcal{K}_{ID} = \langle\Gamma_{ID}\rangle/H = D_i/2Y_L^{\rm eq}$; whereas for scattering, one has $\mathcal{K}_{ia\to mn} = n_i^{\rm eq}\langle v\sigma_{ia\to mn}\rangle/H = S_{ia\to mn}/Y_i^{\rm eq}$.

3. Neutrino abundance

The Boltzmann equations for the RHN abundances can be determined in a similar manner to the lepton asymmetry discussed above,

$$z_1\frac{\partial Y_1}{\partial z_1} = -\left(\frac{Y_1}{Y_1^{\rm eq}} - 1\right)(D_1 + D_{21} + S_1) + \left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right)D_{21} - \left(\frac{Y_1 Y_2}{Y_1^{\rm eq} Y_2^{\rm eq}} - 1\right)S_{N_1N_2\to HH} - \left(\frac{Y_1^2}{Y_1^{\rm eq2}} - 1\right)S_{N_1N_1\to HH},$$
$$z_1\frac{\partial Y_2}{\partial z_1} = -\left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right)(D_2 + D_{21} + S_2) + \left(\frac{Y_1}{Y_1^{\rm eq}} - 1\right)D_{21} - \left(\frac{Y_1 Y_2}{Y_1^{\rm eq} Y_2^{\rm eq}} - 1\right)S_{N_1N_2\to HH} - \left(\frac{Y_2^2}{Y_2^{\rm eq2}} - 1\right)S_{N_2N_2\to HH}, \quad (51)$$

with

$$S_1 = 2S_{N_1L\to Qt} + 4S_{N_1Q\to Lt} + 2S_{N_1L\to HS} + 4S_{N_1H\to LS} + 2S_{N_1S\to{}_1LH},$$
$$S_2 = 2S_{N_2L\to Qt} + 4S_{N_2Q\to Lt} + 2S_{N_2L\to HS} + 4S_{N_2H\to LS} + 2S_{N_2S\to{}_1LH} + 2S_{N_2S\to{}_2LH}. \quad (52)$$

The subtracted rate for $N_1 S \to LH$ is very small and has been ignored. The decay function for $N_2 \to N_1 S$ is defined as

$$D_{21} = K_{21} z_2^2 \frac{K_1(z_2)}{K_2(z_2)} Y_2^{\rm eq}, \qquad z_2 = \frac{M_2}{T} = \frac{M_2}{M_1} z_1, \quad (53)$$

with K_{21} defined by analogy to K_1,

$$K_{21} = \frac{\Gamma_{21}}{H(T = M_2)} = |\alpha_{21}|^2 \frac{v^2}{2m_* M_2} \sim |\alpha_{21}|^2 \left(1.44 \times \frac{10^{16}\ {\rm GeV}}{M_2}\right). \quad (54)$$

We have implicitly assumed the hierarchy $\{M_1, m_S\} \ll M_2$ in writing down K_{21} above; the exact decay rate is calculated in Appendix C. Once again, the thermal equilibrium parameter for this process is $\mathcal{K}_{N_2\to N_1 S} = D_{21}/Y_2^{\rm eq}$, while for inverse decays $N_1 S \to N_2$, one has $\mathcal{K}_{N_1 S\to N_2} = D_{21}/Y_1^{\rm eq}$.

As a summary, Fig. 7 lists the scattering processes that are relevant for the equations.

4. Physical regimes

The underlying dynamics of this system is in the end quite similar to the simpler system which only accounts for decays and inverse decays (plus RIS contributions), which will be discussed in Sec. III B. At first sight it is surprising that the two-to-two scattering processes involving N_1, which can remain in equilibrium after N_2 decays, do not have a more significant role in washing out the asymmetry. Indeed, this intuition is realized if the couplings are sufficiently large as will be seen in the next section. However, there are natural parameter regimes in which the rates involving N_1 which change lepton number can be out of equilibrium while the rates that change the number

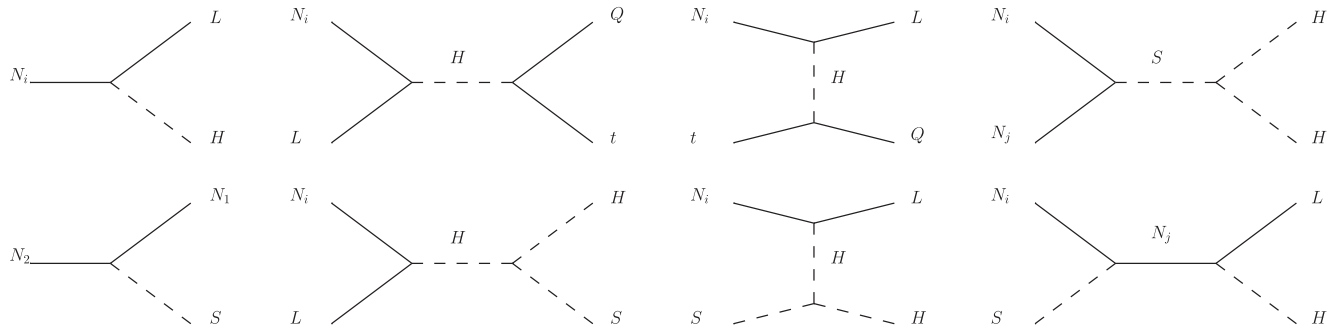


FIG. 7. A list of the decay and scattering processes included in our Boltzmann equations. This list is not exhaustive, though it includes the dominant contributions. In the limit where the quarks, Higgs, and singlet S are massless, the t-channel diagrams are counted twice since it is possible to swap $Q - t$ and $H - S$.

density remain active. This feature is crucial for realizing viable HPL and will be discussed in more detail subsequently. For now, we briefly summarize the generation of a lepton asymmetry by breaking the evolution into three distinct phases:

(1) The N_2 phase:

This phase takes place at temperatures $T \sim M_2$ and is marked by N_2 interactions going out of equilibrium, efficiently generating a primary lepton asymmetry. If there is a large mass hierarchy between N_1 and N_2, the rates involving N_1 may be sufficiently small (e.g. the decays which are proportional to the mass) that they are out of equilibrium, i.e. $D_1 \ll D_2, D_{21}$. In this case, the physics of this phase is almost that of a one-flavor system: the lepton asymmetry is generated through N_2 decays to leptons until the N_2 abundance becomes negligible. A second subdominant process can still be important, namely the mixing $N_2 - N_1$ that allows decays and inverse decays $N_2 \leftrightarrow N_1 S$ if α_{12} is sufficiently large. Because of this channel, the branching ratio of N_2 into leptons is reduced as compared to the one-flavor system, making the production of the lepton asymmetry less efficient. At the same time, this very channel populates N_1, which is then kicked out of equilibrium momentarily. Its ability to return to equilibrium depends on the rate of the inverse decay $N_1 S \to N_2$, as the decays $N_1 \to LH$ are generally out of equilibrium in this phase. The overpopulation of N_1 is not very important during this phase but will have an effect on the lepton asymmetry at the later stage when N_1 decays come into equilibrium. The resulting primary lepton asymmetry can be parametrized by an efficiency factor κ_2,

$$Y^{(2)}_{L-\bar{L}} = \epsilon_2 \kappa_2 Y^{\rm eq}_2(0). \tag{55}$$

This phase typically ends when $T \sim M_2/10$, i.e. $z_1 \sim 10 M_1/M_2$.

(2) The intermediate phase:

Given a sizeable mass hierarchy between N_1 and N_2, the second phase is marked by a large temperature gap once N_2 has effectively disappeared and before N_1 interactions come into equilibrium. Neither N_2 nor N_1 interactions are able to affect the lepton asymmetry, or the N_1 abundance, and the system effectively free streams leading to a plateau in $Y_{L-\bar{L}}$. This phase lasts for as long as the N_1 interactions remain out of equilibrium, and characteristically for a temperature range similar to the mass ratio. For example, if the decays and inverse decays dominate, the approximations discussed in Appendix B indicate that the phase ends when $z_1 \sim \sqrt{2/K_1}$. If scattering effects are also significant, then the transition to the N_1 phase can occur somewhat earlier.

(3) The N_1 phase:

This phase is marked by N_1 interactions being in equilibrium which efficiently deplete the neutrino abundance and lepton asymmetry. Assuming again a sizeable mass hierarchy, since the N_2 abundance is negligible and N_2 interactions are effectively turned off, $D_2, D_{21} \ll D_1$, the dynamics again approximates a purely one-flavor system. The distinction is that the initial lepton asymmetry is not zero, having been generated in the N_2 phase, and there is a possible N_1 overabundance $Y_1 > Y_1^{\rm eq}$ due to $N_2 \to N_1 S$ decays during the first phase. The final lepton asymmetry results from the competition between N_1-mediated processes that wash out the preexisting lepton asymmetry from the N_2 phase and those at the end of the N_1 phase that contribute to the asymmetry. The result can again be parametrized via an efficiency factor κ_1,

$$Y_{L-\bar{L}} = Y^{(2)}_{L-\bar{L}} e^{-\int_{\mathcal{Z}}^{\infty} dz' (W_1/z')} + \epsilon_1 \kappa_1 Y_1^{\rm eq}(0), \tag{56}$$

where the washout function $W_1 = W_{ID_1} + W_{S_1}$ is discussed above. The variable $\mathcal{Z}$ marks the transition

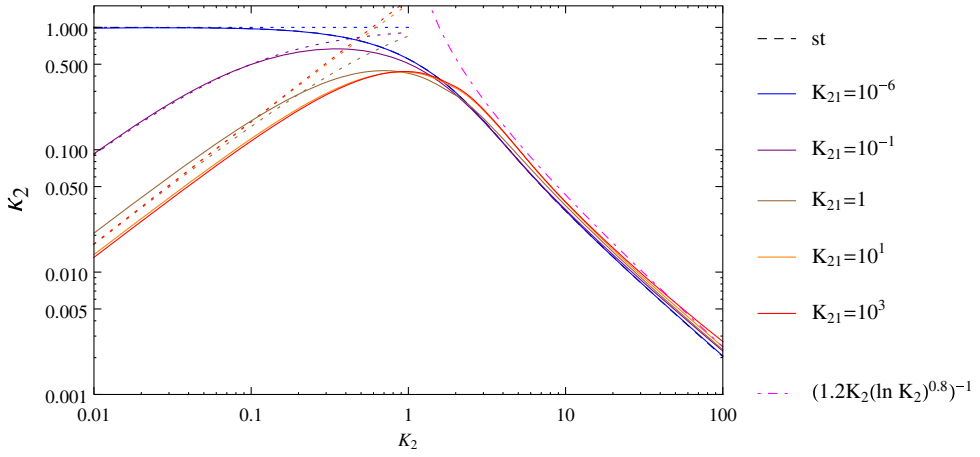


FIG. 8 (color online). Plots of the efficiency factor κ_2, as a function of K_2, for various values of K_{21}. In the limit $K_{21} \ll K_2$, the efficiency factor approaches its value in one-flavor standard leptogenesis (shown here in black, dashed, buried under the blue line). Increasing K_{21} has a significant effect on the efficiency factor at low K_2. The asymptotic behavior is also shown (dotted lines) for low K_2, from (62) and (64).

point after the N_2 phase, once the washout processes become active.

Summarizing the full two-level process, the final asymmetry resulting from the three phases can be parametrized by the two efficiency factors κ_1 and κ_2,

$$Y^f_{L-\bar{L}} = \epsilon_2 \left(\kappa_2 e^{-\int_z^\infty dz' (W_1/z')} + \frac{\epsilon_1}{\epsilon_2} \kappa_1 \right) Y_2^{\rm eq}(0). \tag{57}$$

We proceed in the next section to consider explicit examples which exhibit these features in detail. However, before considering the general case, we will first study a simplified toy model that allows some analytic understanding of the physics.

B. Toy model of the two-stage evolution

In this subsection, in order to isolate some of the dominant physical effects, we study a toy model of the two-level Boltzmann equations, accounting only for decays and inverse decays and ignoring the impact of two-to-two scattering. For simplicity, we also take the CP asymmetry to be constant, using $\epsilon_1 = \epsilon_2 = 10^{-8}$, although this constraint will be relaxed toward the end of the section. The Boltzmann equations are as written in Eqs. (46) and (51), of Sec. III A, but without scattering,

$$\begin{aligned}
z_1 \frac{\partial Y_1}{\partial z_1} &= -\left(\frac{Y_1}{Y_1^{\rm eq}} - 1\right)(D_1 + D_{21}) + \left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right) D_{21}, \\
z_1 \frac{\partial Y_2}{\partial z_1} &= -\left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right)(D_2 + D_{21}) + \left(\frac{Y_1}{Y_1^{\rm eq}} - 1\right) D_{21}, \\
z_1 \frac{\partial Y_{L-\bar{L}}}{\partial z_1} &= \epsilon_1 D_1 \left(\frac{Y_1}{Y_1^{\rm eq}} - 1\right) + \epsilon_2 D_2 \left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right) \\
&\quad - Y_{L-\bar{L}} \frac{D_1 + D_2}{2 Y_L^{\rm eq}}.
\end{aligned} \tag{58}$$

The decay functions are as defined above, and we take the initial conditions as a vanishing lepton asymmetry $Y_{L-\bar{L}} = 0$ and equilibrium initial abundances $Y_{1,2} = Y_{1,2}^{\rm eq}$. The N_2 phase ends at around $\mathcal{Z} \sim 10 M_1/M_2 < 1$, and we can numerically integrate $\int_{\mathcal{Z}}^\infty dz' W_{ID_1}/z' \sim 1.2 K_1$. Applying the general result of (57) leads to an approximate final asymmetry,

$$Y^f_{L-\bar{L}} \simeq \frac{3}{4} (\epsilon_2 \kappa_2 e^{-1.2 K_1} + \epsilon_1 \kappa_1). \tag{59}$$

1. N_2 efficiency factor: κ_2

In the N_2 phase, the scenario of interest here is characterized by having N_1 out of equilibrium with $D_1 \ll D_2, D_{21}$. Integrating the $Y_{L-\bar{L}}$ equation in (58) then leads to the efficiency factor κ_2,

$$\begin{aligned}
\kappa_2 Y_2^{\rm eq}(0) &= \int_0^{\mathcal{Z}} dz' \frac{D_2}{z'} \left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right) e^{-\int_{z'}^{\mathcal{Z}} dz'' (W_{ID_2}/z'')}, \\
&= -\frac{K_2}{K_2 + K_{21}} \int_0^{\mathcal{Z}} dz' \left(\frac{\partial Y_2}{\partial z'} - \frac{D_{21}}{z'} \left(\frac{Y_1}{Y_1^{\rm eq}} - 1 \right) \right) \\
&\quad \times e^{-\int_{z'}^{\mathcal{Z}} dz'' (W_{ID_2}/z'')}.
\end{aligned} \tag{60}$$

For concision, we have used the variable z, although strictly this is z_1. Evaluating κ_2 numerically, we obtain the contours shown in Fig. 8, as a function of K_2 for various values of K_{21}. To obtain an analytic approximation, we focus on the regime $K_2 \ll 1$, where K_{21} has the largest effect. In this limit, the washout from inverse decays is quite limited, $\exp(-\int_{z'}^{z} dz'' W_{ID_2}/z'') \sim 1$, so that the first part of the efficiency factor is trivial to integrate $\int_0^z dz' \partial Y_2/\partial z' = Y_2(z) - Y_2(0) \to -Y_2(0)$, since by definition we integrate to the point where the N_2 abundance drops to zero. Thus,

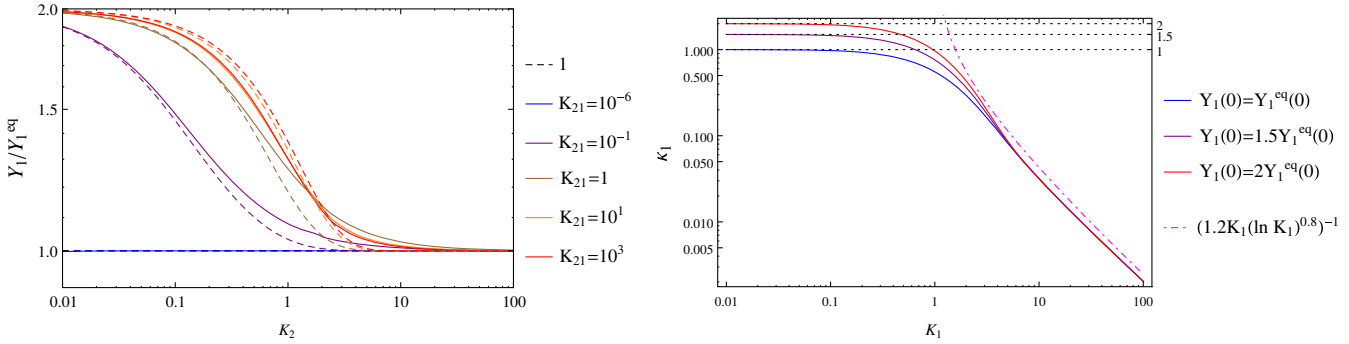


FIG. 9 (color online). The plot on the left (a) shows the N_1 abundance $Y_1(\mathcal{Z})$ offset compared to equilibrium at the end of the first phase due to the $N_2 \to N_1 S$ decay, referring to Eq. (68) along with the approximations (69). These values are also the initial conditions for the abundance at the start of the very last phase. The plot on the right (b) represents the efficiency factor κ_1 for various initial values of $Y_1(\mathcal{Z})$ at the start of the N_1 phase, as given in the left plot.

$$\kappa_2 Y_2^{\rm eq}(0) \simeq \frac{K_2}{K_2+K_{21}} Y_2^{\rm eq}(0) + \frac{K_2}{K_2+K_{21}} \int_0^{\mathcal{Z}} dz' \frac{D_{21}}{z'} \left(\frac{Y_1}{Y_1^{\rm eq}} - 1 \right), \quad (61)$$

owing to the initial condition $Y_2(0) = Y^{\rm eq}(0)$. Taking the limit $K_{21} \ll K_2$, the equations decouple in such a way that $Y_1 \sim Y_1^{\rm eq}$, and the second term above is subdominant leading to the efficiency factor,

$$\kappa_2 \sim \frac{K_2}{K_2+K_{21}} \quad \text{when } K_{21} \ll K_2 \ll 1. \quad (62)$$

In the decoupled limit, the efficiency factor logically tends to the one-flavor value. If instead we take the limit $K_{21} \gg K_2$, the second term becomes significant, if not dominant, and we have to integrate the equations explicitly. In this limit, the branching ratio of N_2 into leptons is small, so that $D_{21} \gg D_2$. At the same time, $D_{21} \gg D_1$ in the N_2 phase. As a result, the $Y_{1,2}$ equations simplify to $\partial Y_2/\partial z_1 = -\partial Y_1/\partial z_1$, so that the total number density $Y_1 + Y_2$ is constant. This makes sense, since N_2's decay dominantly into N_1's. We thus find,

$$Y_2(\mathcal{Z}) + Y_1(\mathcal{Z}) = Y_2(0) + Y_1(0) = 2Y_1^{\rm eq}(0). \quad (63)$$

The first phase ends when $Y_2(\mathcal{Z}) = 0$, so that $Y_1(\mathcal{Z}) = 2Y_1^{\rm eq}(0)$, and the term $Y_1(\mathcal{Z})/Y_1^{\rm eq}(\mathcal{Z}) - 1 \sim 1$. The calculation of $\int_0^z dz' D_{21}/z'$ is most easily performed numerically, and for $z \geq 10$ the integral converges to 1.7. We conservatively take this result to obtain

$$\kappa_2 \sim 1.7 \frac{K_2 K_{21}}{K_2+K_{21}} \quad \text{when } K_{21} \gg K_2 \ll 1. \quad (64)$$

Figure 8 exhibits both the numerical results for κ_2 along with the approximations (62) and (64); the agreement is good in the region $K_2 \geq K_{21}$, but less so for $K_{21} \geq K_2$. The $K_2 \gg 1$ regime is similar to a one-flavor case, and the coupling $N_2 - N_1$ does not have a large effect. Thus, we can refer to the established literature [5] for an approximate expression for κ_2 in this region,

$$\kappa_2 \simeq \frac{1}{1.2K_2(\log K_2)^{0.8}} \quad \text{when } K_2 \gg 1. \quad (65)$$

2. N_1 efficiency factor: κ_1

The N_1 phase is characterized by negligible N_2 abundance, and with N_2 interactions being out of equilibrium $D_2, D_{21} \ll D_1$. Thus, the physics is once again equivalent to the one-flavor case. The efficiency factor κ_1 is obtained by integrating the Boltzmann equation for the lepton asymmetry,

$$\kappa_1 Y_1^{\rm eq}(0) = \int_{\mathcal{Z}}^{\infty} dz' \frac{D_1}{z'} \left(\frac{Y_1}{Y_1^{\rm eq}} - 1 \right) e^{-\int_{z'}^{\infty} dz''(W_{ID_1}/z'')}, = -\int_{\mathcal{Z}}^{\infty} dz' \frac{\partial Y_1}{\partial z'} e^{-\int_{z'}^{\infty} dz''(W_{ID_1}/z'')}. \quad (66)$$

In the limit $K_1 \ll 1$, the washout is minimal, leaving

$$\kappa_1 Y_1^{\rm eq}(0) = Y_1(\mathcal{Z}). \quad (67)$$

The efficiency factor depends on the value of the abundance once the third phase starts. Since Y_1 is constant in the intermediate phase, this is the same as the final abundance at the end of the first phase. In Fig. 9(a), we show numerical results for $Y_1(\mathcal{Z})$ as a function of K_2 for various values of K_{21}. In calculating κ_2, we estimated that $Y_1(\mathcal{Z}) = 2$ at the end of the first phase, which is true in the limit $K_2 \sim 0$, and $\partial(Y_1 + Y_2)/\partial z_1 = 0$. When $K_2 \ll 1$ but not zero, $Y_1 + Y_2$ is no longer constant, and instead we have

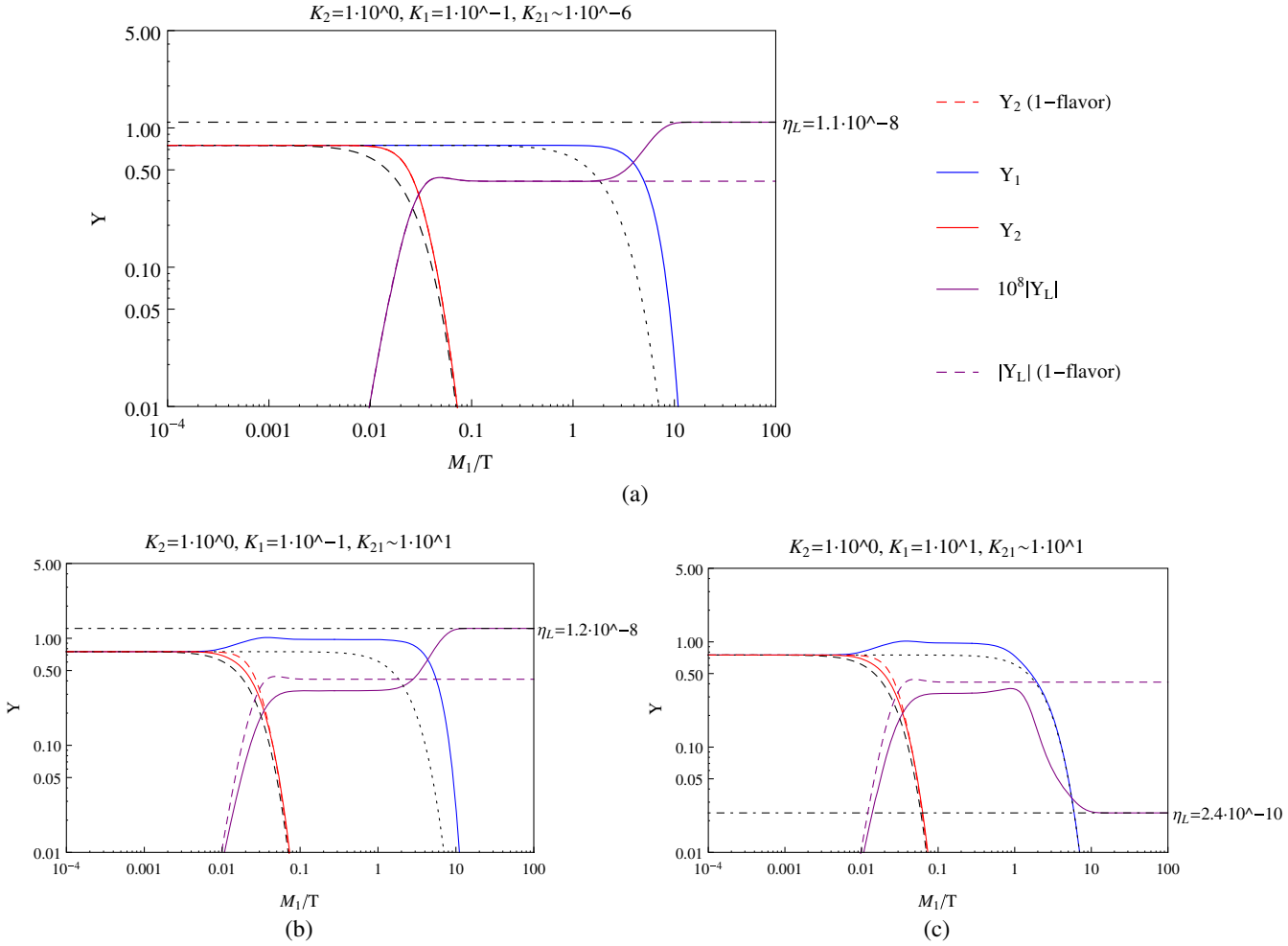


FIG. 10 (color online). Plot (a) shows the N_2 and N_1 abundances (red and blue) and $Y_{L-\bar{L}}$ (purple) in the toy two-flavor model, compared to the N_2 abundance (in dashed red) and the lepton asymmetry (in dashed purple) in the one-flavor case. Plots (b) and (c) show further examples of the two-level toy model. The coupling $N_2 \to N_1 S$ allows N_2 to cope better with expansion, which causes a reduction of the generated lepton asymmetry. In turn, this results in the N_1 abundance being further out of equilibrium. The color coding is the same in all these and subsequent plots.

$$Y_1(\mathcal{Z}) = 2Y_1^{\rm eq}(0) - \int_0^{\mathcal{Z}} dz' \frac{D_2}{z'}\left(\frac{Y_2}{Y_2^{\rm eq}} - 1\right). \tag{68}$$

This function is plotted in Fig. 9(a), along with the analytical approximations. Thus, to a good approximation, we have

$$\kappa_1 = \frac{Y_1(\mathcal{Z})}{Y_1^{\rm eq}(0)} = 1 + \frac{K_{21}}{K_2 + K_{21}} e^{-K_2} \quad \text{when } K_1 \ll 1. \tag{69}$$

For large $K_2 \gg 1$, the behavior approaches the decoupled limit. Numerical results for the efficiency factor κ_1 are displayed in Fig. 9(b) as a function of K_1 for varying initial conditions. At the other end of the spectrum, as in standard leptogenesis, the efficiency factor is [5]

$$\kappa_1 \sim \frac{1}{1.2K_1(\ln K_1)^{0.8}} \quad \text{when } K_1 \gg 1. \tag{70}$$

3. Final lepton asymmetry in the toy model

In Fig. 10(a), we show an example of a decoupled system, with the parameters $\{K_2, K_1, K_{21}\} = \{1, 0.1, 10^{-6}\}$. From Fig. 8, we find the efficiency factor $\kappa_2 \sim 0.5$, while Fig. 9(a) tells us that the overabundance at the exit of the first phase will be zero (decoupling limit), and Fig. 9(b) then implies that $\kappa_1 \sim 1$. Inserting these values into Eq. (59) leads to $Y^f_{L-\bar{L}} \sim 1.1 \times 10^{-8}$, which is in excellent agreement with the numerically determined result. Two examples of coupled system are shown in Figs. 10(b) and 10(c) with the parameters $\{K_2, K_1, K_{21}\} = \{1, 0.1, 10\}$ and $\{K_2, K_1, K_{21}\} = \{1, 10, 10\}$ respectively. Using Figs. 8 and 9, we find the efficiency factors $\{\kappa_2, \kappa_1\} = \{0.4, 1.3\}$ and $\{\kappa_2, \kappa_1\} = \{0.4, 0.03\}$ and obtain $Y^f_{L-\bar{L}} \sim 1.3 \times 10^{-8}$ and $Y^f_{L-\bar{L}} = 2.25 \times 10^{-10}$ respectively.

With this understanding of the toy model, we turn in the next section to an analysis of the full system including scattering. It should already be apparent that viable models

will be those in which N_1 is sufficiently weakly coupled that the most dangerous effect, rapid washout via N_1-mediated inverse decays and two-to-two scattering, is suppressed. With this constraint, the residual effects of scattering are generally quite small.

IV. RESULTS IN THE HIERARCHICAL REGIME

The focus in this section will be on studying the solutions to the Boltzmann equations in the hierarchical regime, $M_2/M_1 \gg 1$. As shown in Fig. 16(a) in Appendix A, the exact and large M_2/M_1 expressions for the CP asymmetry are within a factor of 2 for $M_2/M_1 > 10$, which will serve as a practical definition of this regime. Qualitatively, the physical behavior should be similar for all mass ratios outside the resonant regime [14], $M_2 - M_1 \sim \Gamma_{1,2}/2$, which we will not consider here.

The two-flavor HPL model is distinct from standard leptogenesis in at least two ways. The first difference concerns the mass dependence of the CP asymmetry. As discussed in Sec. II, and again in Appendix A, the CP asymmetry presents distinct high and the low-mass regimes. The high-mass regime, $M_2 \gtrsim 10^8$ GeV, is determined by the standard Yukawa contribution to the CP asymmetry. The low-mass regime is instead determined by the hidden sector contribution to the CP asymmetry, proportional to the trilinear coupling β/M_i. This liberates the model from the Davidson–Ibarra bound on the CP asymmetry and allows for viable low-scale scenarios.

The other significant difference with standard leptogenesis concerns the dynamics. In the minimal model, the main contribution to the CP asymmetry comes from the decays and inverse decays into leptons, and the scattering processes are largely subdominant. This is in part because as the temperature falls below $T \sim M_1$, all the L-violating scattering rates are suppressed due to Boltzmann suppression of the neutrino abundance. For HPL, the situation is different due to the emphasis on the CP asymmetry generated by N_2 decays and the importance of the evolution between $T \sim M_2$ and $T \sim M_1$. As a consequence, scattering processes involving N_1 have the potential to affect the lepton asymmetry quite significantly and need to be considered carefully.

A. Viable scenarios

We will impose two requirements on realistic scenarios, namely the ability to reproduce the observed baryon asymmetry, and similarly that they admit a consistent light neutrino mass spectrum. The first requirement translates within leptogenesis to a specific lepton asymmetry at the temperature where $B+L$-violating sphaleron processes fall out of equilibrium. For the Standard Model field content, the equilibrated lepton and baryon asymmetries are related by $\eta_L = -(51/28)\eta_B$.The additional singlet in the Higgs portal model only affects this by changing the critical temperature of the electroweak crossover. Since we assume $\langle S \rangle = 0$, then at least for relatively weak $H - S$ mixing the impact should be small [31]. We therefore require $|\eta_L| \simeq 1 \times 10^{-9}$, given the Planck result for $\Omega_b h^2$ [32], which translates to the baryon-to-photon ratio $\eta_B \simeq (6.04 \pm 0.09) \times 10^{-10}$ [33].

For the second requirement, since the seesaw mechanism is a motivating factor for leptogenesis, we also require consistency with current data on the mass squared differences, e.g. $\Delta m^2_{21} = m_2^2 - m_1^2 \simeq (7.5 \pm 0.5) \times 10^{-5}$ eV^2 [2]. The seesaw mechanism determines an *effective* light neutrino mass $\tilde{m}_i \sim \lambda_i^2 v^2/M_i$, which we can trade for the thermal equilibrium parameters $K_i = \tilde{m}_i/m_*$ from (48) and write down an effective mass squared difference, $K_2^2 - K_1^2 \simeq \mathcal{O}(\Delta m^2_{21}/m^2_*)$. However, this relation relies on the equality $\tilde{m}_i \simeq m_i$ between the effective and the physical light neutrino masses, which only holds when the neutrino flavor structure is nearly diagonal [24]. More generally, the precise relation can be relaxed, so we will consider models to be viable if the interactions are in the range $K_2 \sim \mathcal{O}(1\text{–}10)$.

In the rest of this subsection, we present example scenarios that satisfy the above constraints on the neutrino masses and the lepton asymmetry. For each case we present three figures: (i) the Boltzmann evolution of the neutrino abundances and the lepton asymmetry; (ii) the relevant thermal rates of decays, inverse decays, and scattering; and finally (iii) the CP-asymmetry "landscape" in which the theory is situated. In the following subsection, we provide further details showing the impact of varying the parameters of the theory, while relaxing the constraints imposed here on viable models. In particular, we show that the dynamics of the high-mass regime is most sensitive to $\{K_1, K_2\}$, whereas the low-mass dynamics responds to $\{K_2, K_1, \alpha, \beta\}$.

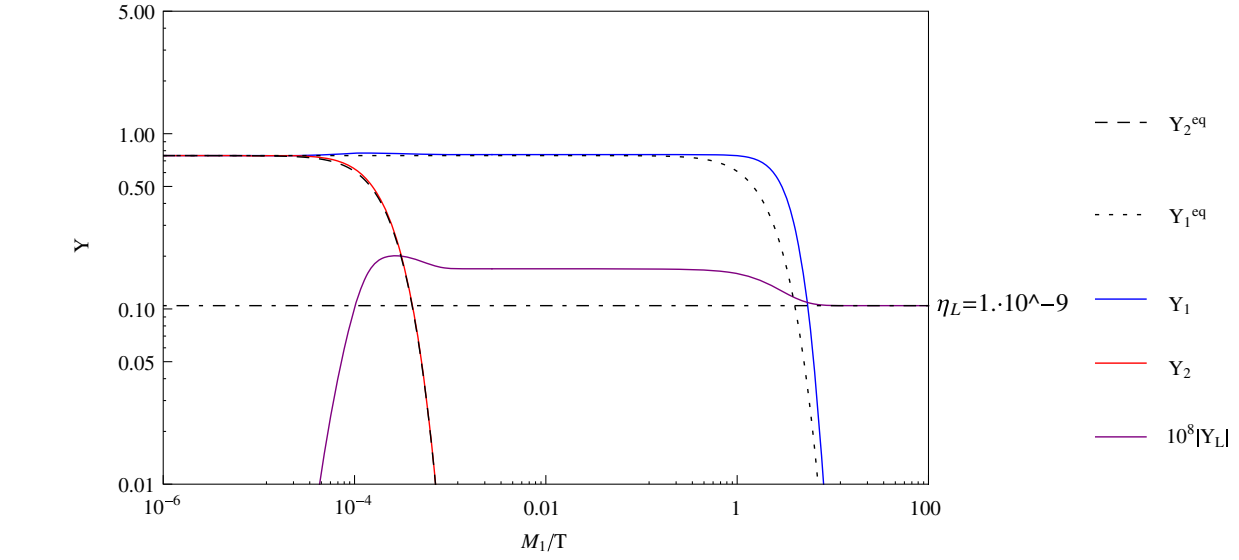
It is useful to distinguish "high" and "low" mass regimes, based primarily on the mass dependence of the N_2 CP asymmetry. We focus below on the relative impact of two-to-two scattering processes, compared to the toy model discussed above.

(i) High-scale models

An example of a high-scale scenario is shown in Fig. 11(a). The high-mass regime is marked by the dominance of the Yukawa sector in contributing to the CP asymmetry, so that $\epsilon_2 \sim (3/16\pi v^2) \sum_\alpha m_\alpha M_2 \sim \epsilon_1 (M_2/2M_1)$, as displayed in Fig. 11(c). The general result in Eq. (57) reduces to

$$Y^f_{L-\bar{L}} \sim \frac{9\sum_\alpha m_\alpha}{64\pi v^2} M_2 \left(\kappa_2 e^{-\int_z^\infty dz' (W_1/z')} + 2\frac{M_1}{M_2}\kappa_1 \right). \tag{71}$$

The lepton-number changing processes mediated by the hidden sector scale as $\sim |\lambda|^2\beta^2/M^2$ and are suppressed compared to Yukawa-mediated scattering



(a) $\{M_2, M_1, K_2, K_1, K_{21}, \beta\} = \{10^9\text{GeV}, 10^5\text{GeV}, 8, 0.2, 40, 50\text{GeV}\}$

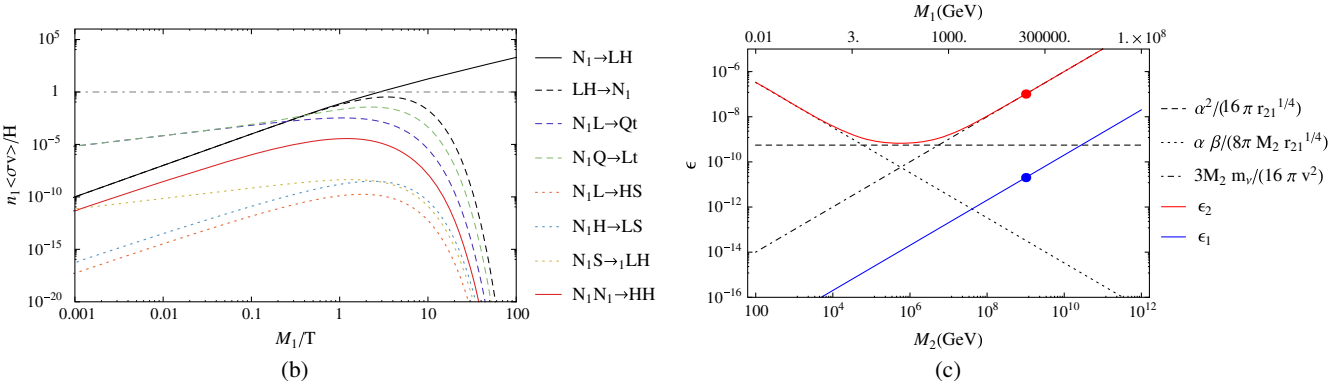
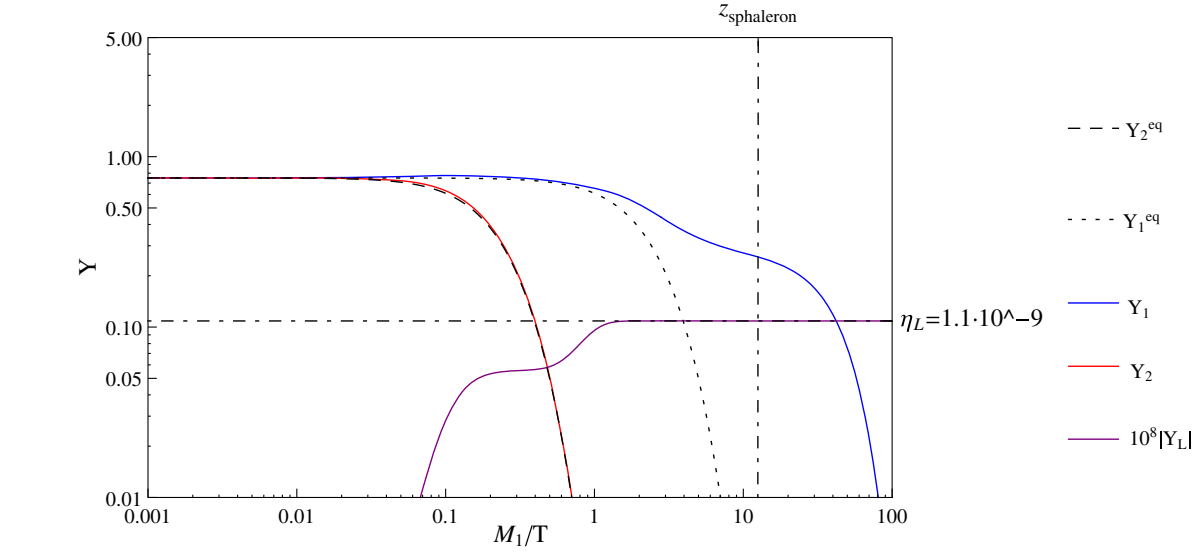


FIG. 11 (color online). These plots show a viable example of the high-scale scenario. We show the evolution of the abundances and (rescaled) asymmetry (a) and the relative rates (b) as a function of temperature. In this regime, the CP asymmetry is dominated by the standard Yukawa contribution, $\epsilon_i \propto m_\nu M_i/v^2$, and the final plot (c) displays the CP asymmetries $\epsilon_{1,2}$ as functions of M_2, for a constant mass ratio, $M_2/M_1 = 10^4$. With a large mass hierarchy, there would be plenty of time for N_1-mediated processes to wash out all of the lepton asymmetry unless the N_1 interactions are sufficiently suppressed, hence the very small Yukawa λ_1.

which scales as $\sim|\lambda|^2 m_t^2/v^2$. This is shown in Fig. 11(b). Additional scattering processes such as $N_i S \to {}_j LH$ scale as $|\alpha|^2|\lambda|^2$ and are therefore suppressed if $|\alpha|^2 \ll 1$, which holds for all the viable scenarios we consider here. The lepton-number conserving scattering process $N_i N_i \to HH$, mediated by S in the s channel, scales as $|\alpha|^2\beta^2/M^2$ and is again suppressed. Finally, the lepton-mediated t-channel $N_i N_i \to HH$ scattering is suppressed by a factor $|\lambda|^2 \ll 1$ and has no visible effect on the neutrino abundance.

It follows that the efficiency factors $\kappa_{1,2}$ determined above can be used as a reasonably good approximation here, as the scattering corrections are small. Indeed, the dominant N_i scattering processes are controlled by the same equilibrium parameter K_i, and the Boltzmann-suppressed number density for $T < M_i$ and the fast expansion rate at $T > M_i$ limits their range of activity.

The possibility of having $\Delta L = 1, 2$ scattering processes mediated by N_1 that remain in equilibrium long after the lepton asymmetry $Y^{(2)}_{L-\bar{L}}$ has been generated is an important feature of HPL. In practice, these processes need to be suppressed and remain out of equilibrium in viable models to avoid too much washout. For masses chosen so that $\beta/M_1 \ll m_t/v$, the washout function is dominated by the Yukawa processes $\int_{\tilde{z}}^{\infty} dz'(W_1/z') \simeq cK_1$, where $c \sim 1.2 + 1.4(m_t/v)^2 \sim 2.5$. With $K_2 \sim 1$–10, satisfying the neutrino mass constraint, one has $\kappa_2 \sim 1/(cK_2)$, and taking a mass ratio large enough that the κ_1 term can be neglected, we find the constraint



(a) $\{M_2, M_1, K_2, K_1, K_{21}, \beta\} = \{20\text{TeV}, 2\text{TeV}, 7, 0.001, 3\cdot 10^5, 50\text{GeV}\}$

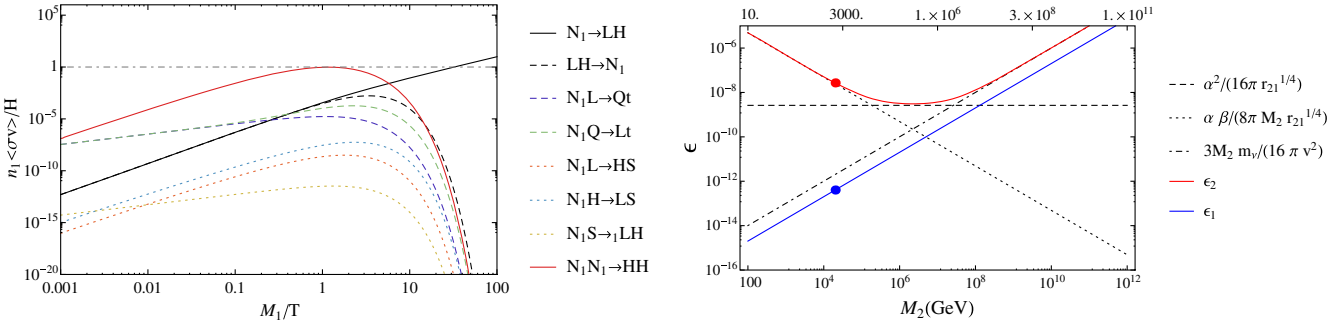


FIG. 12 (color online). These plots, in the same format as Fig. 11, show a viable example in the low-mass regime. The hidden sector-sourced *CP* asymmetry from N_2 decays is the dominant contribution. The dominant scattering processes are also mediated by the hidden sector. This is also an explicit example of a scenario leading to a lepton asymmetry enhancement, thanks to the evolution traversing the sphaleron freeze-out temperature.

$$Y^f_{L-\bar{L}} \sim \frac{9\sum_\alpha m_\alpha}{64\pi v^2} M_2 \frac{e^{-cK_1}}{cK_2} > \eta_L \simeq 10^{-9}. \quad (72)$$

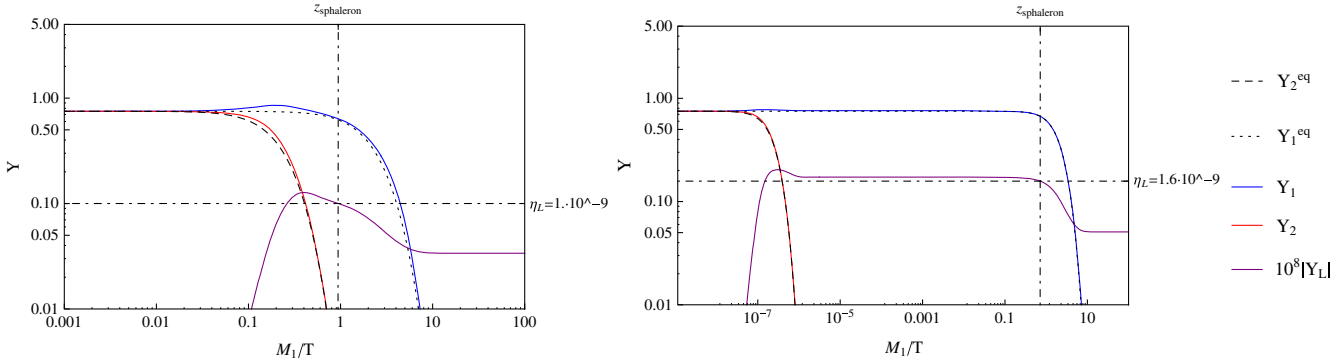
With the neutrino mass constraint $\sum_\alpha m_\alpha \sim m_3$ as input, this implies the lower bound $M_2 > (10^7\text{GeV})\cdot cK_2 e^{cK_1}$. In standard one-flavor leptogenesis, we instead obtain the simpler bound, $M_1 \gtrsim 10^7$ GeV. This is because the equilibrium parameter is constrained to satisfy $K_1 < 1$, so that $\kappa_1 \sim 1$, and because scattering processes are not in equilibrium long enough to provide any significant washout.

(ii) *Low-scale models*

A viable low-scale example is shown in Fig. 12(a). The *CP* asymmetry ϵ_2 for low values of M_2 is controlled by the hidden sector couplings $\{\alpha, \beta\}$, i.e. $\epsilon_2 \propto (\beta/M_2 + |\alpha_{21}|/2)|\alpha_{21}|/r_{21}^{1/4}$. At the same time, because the *CP* asymmetry ϵ_1 is sourced purely from the Yukawa sector, it becomes negligible at low mass, $\epsilon_1 \propto \sum_\alpha m_\alpha M_1/v^2 \ll \epsilon_2$, as exemplified by Fig. 12(c). In this case, Eq. (57) reads

$$Y^f_{L-\bar{L}} \simeq \frac{3}{32\pi}\kappa_2\left(\frac{\beta}{M_2} + \frac{|\alpha_{21}|}{2}\right)\frac{|\alpha_{21}|}{r_{21}^{1/4}} e^{-\int_z^\infty dz(W_1/z')}. \quad (73)$$

This relation relies on a hierarchical separation between the N_2 and N_1 phases, which is only marginally satisfied in Fig. 12(a) where $M_2/M_1 = 10$. Nevertheless, the above relation encodes the two competing effects, namely the enhancement of the low-mass *CP* asymmetry, and also the increased washout. Generating a larger *CP* asymmetry requires an increase in the ratio β/M, which at the same time increases the scattering rates and the washout (see the next subsection for details of the relative effects). Independent of



(a) $\{M_2, M_1, K_2, K_1, K_{21}, \beta\} = \{1500\,\text{GeV}, 155\,\text{GeV}, 1, 0.5, 10^5, 20\,\text{GeV}\}$

(b) $\{M_2, M_1, K_2, K_1, K_{21}, \beta\} = \{10^9\,\text{GeV}, 130\,\text{GeV}, 8, 0.5, 4\cdot 10^3, 1\,\text{GeV}\}$

FIG. 13 (color online). Scenarios representig low (a) and high (b) scale models which take advantage of the sphaleron cutoff temperature in order to liberate the parameters from stringent constraints.

the precise dynamics, for the relevant couplings, $K_2 \sim 1$–10, the efficiency factor will generally lie in the range $\kappa_2 \sim 0.01$–0.5 (see Fig. 8). The washout function depends on the scattering processes, and we can approximate $\int_{\mathcal{Z}}^{\infty} dz' W_1/z' \sim (3 + 2\beta/M_1)K_1$, so that $\exp(-\int_{\mathcal{Z}}^{\infty} dz' W_1/z') \sim 0.5$–$1$ for $K_1 \sim 0.01$–0.1 and $\beta/M_1 \sim 1$. As an example, taking $\kappa_2 \exp(-\int_{\mathcal{Z}}^{\infty} dz' W_1/z') \sim 0.1$ implies the following characteristic constraint on the lepton asymmetry:

$$Y^f_{L-\bar{L}} \sim \frac{3}{32\pi}\left(\frac{\beta}{M_2} + \frac{|\alpha_{21}|}{2}\right)\frac{|\alpha_{21}|}{r_{21}^{1/4}} 10^{-1} > \eta_L \simeq 10^{-9},$$
$$|\alpha_{21}|\frac{\beta}{M_2}\sqrt{\frac{M_1}{M_2}} > 3 \times 10^{-7}, \quad (74)$$

where small couplings $|\alpha_{21}| \ll \beta/M_2$ have been assumed and typically we will use $|\alpha_{21}| \sim 10^{-5}$. Given the above approximations, the light neutrino mass scale is a subleading parameter and does not appear in this bound.

A feature worth noting is the sensitivity to low-temperature boundary conditions, namely the temperature at which $B + L$-violating sphaleron transitions go out of equilibrium, $T_{\text{sphaleron}} = 159 \pm 1$ GeV [27]. The lepton asymmetry at this point effectively determines the final baryon asymmetry, while further evolution is observationally relatively unconstrained. This is phenomenologically interesting as it is usually quite difficult to find viable models with a large washout from N_1. However, if we choose M_2 above and M_1 well below the temperature at which sphaleron processes freeze out, the baryon asymmetry is generated right after N_2 falls out of equilibrium and will not be washed out at lower scales even if the lepton asymmetry is highly suppressed through N_1 processes. The N_1 sector effectively decouples in this case. A low-mass example is shown in Fig. 13. Note that $T_{\text{sphaleron}} > M_1 > m_H$ in this case, since $M_1 < m_H$ would require the inclusion of interactions such as $N_1 L \leftrightarrow H$, that have not been considered thus far.[4] In the low-mass regime, many interactions are relevant which cause significant washout of the lepton asymmetry. As a consequence, the range of parameters available for low-mass scenarios is quite limited. On the other hand, an example of a high-mass scenario which also takes advantage of the sphaleron cutoff temperature is shown in Fig. 13(b). In that situation, all of the lepton asymmetry is generated out of the first phase, from N_2 leptonic decays. In this case, with $M_1 < T_{\text{sphaleron}}$, the physics of N_1, including the hidden sector interactions, are only weakly constrained. An interesting aspect of this scenario is the possibility of exploring models where the lightest RH neutrino is so light and weakly coupled to leptons that its lifetime could be long enough to play an independent cosmological role, potentially in the form of sterile neutrino dark matter [34]. In such cases, the N_1 abundance will have to be sufficiently depleted for consistency with constraints on the dark matter abundance. This can be achieved through adjusting the $N_1 N_1 \to HH$ rate.

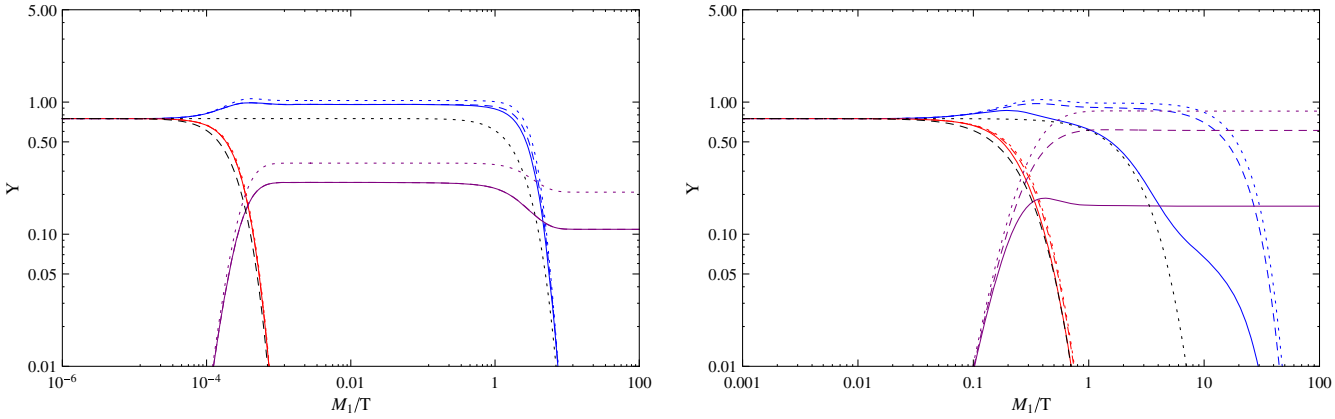
B. Aspects of the dynamics

In this subsection, we relax some of the constraints required for physical scenarios and focus on the various dynamical components that come into play: the decays and scattering processes on one hand, and the various parameters on the other.

1. Decays and inverse decays vs scattering

The impact of the decays, inverse decays, and scattering is summarized in Fig. 14. We overlay three solutions:

[4]Note that such processes may also arise on including thermal corrections to the scalar masses.



(a) $\{M_2, M_1, K_2, K_1, K_{21}, \beta\} = \{10^8\,\text{GeV}, 10^4\,\text{GeV}, 0.8, 0.4, 4\cdot 10^2, 50\,\text{GeV}\}$

(b) $\{M_2, M_1, K_2, K_1, K_{21}, \beta\} = \{2\cdot 10^3\,\text{GeV}, 200\,\text{GeV}, 0.8, 0.04, 400, 500\,\text{GeV}\}$

FIG. 14 (color online). Plots showing the impact of $\Delta L = 1$ and $\Delta N = 2$ processes on the dynamics for high scales $\{M_2, M_1\} = \{10^8, 10^4\}$ GeV (on the left) and low scales $\{M_2, M_1\} = \{2, 0.2\}$ TeV (on the right). The color coding, blue, red, purple, and black, is the same as in Figs. 11(a), 12(a), and 13(a). The dotted pattern is for decays only, dashed for decays plus $\Delta L = 1$ scattering, and plain for decays plus $\Delta L = 1$ and $\Delta N = 2$ scattering.

(i) just decays and inverse decays (dotted lines); (ii) decays, inverse decays, and $\Delta L = 1$ scattering (dashed lines); and finally (iii) decays, inverse decays, and $\Delta L = 1$ and $\Delta N = 2$ scattering (plain lines).

Starting with the high-mass regime, the hidden sector scattering processes, e.g. $N_i H \to LS$, $N_i N_i \to HH$, are subdominant compared to the Yukawa-mediated processes since $\beta/M_i \ll m_t/v$. This can be seen from Fig. 14(a), where $\Delta L = 1$ scattering has a significant effect relative to decays, whereas the $\Delta N = 2$ scatterings have no visible effect. In practice, the decays and inverse decays remain the dominant effect, and the results from Sec. III B can be reasonably well applied.

The low-mass regime is marked instead by the significant, if not dominant, effect of hidden sector scattering processes, given $\beta/M_i \sim m_t/v$. Looking at the plot in Fig. 14(b), the effect of adding $\Delta L = 1$ scattering is similar to that in the high-mass regime; however, the impact of $\Delta N = 2$ processes is greatly enhanced. Although these processes do not violate lepton number, $\Delta L = 0$, they act to maintain the neutrino abundance closer to equilibrium. As a consequence, both the $\Delta L = 1$ processes (e.g. decays) and the inverse processes (e.g. inverse decays) remain in equilibrium for longer, leading to an enhanced lepton washout in that regime.

2. Parameter dependence

The dependence of the dynamics on the parameters is displayed in Fig. 15. The subsection above considered examples of viable models that satisfy the basic constraints and focused on constraining the relevant parameters accordingly. We now ignore those constraints and instead vary the parameters $\{K_2, K_1, K_{21}, \beta\}$ to study their impact in both high- and low-mass regimes.

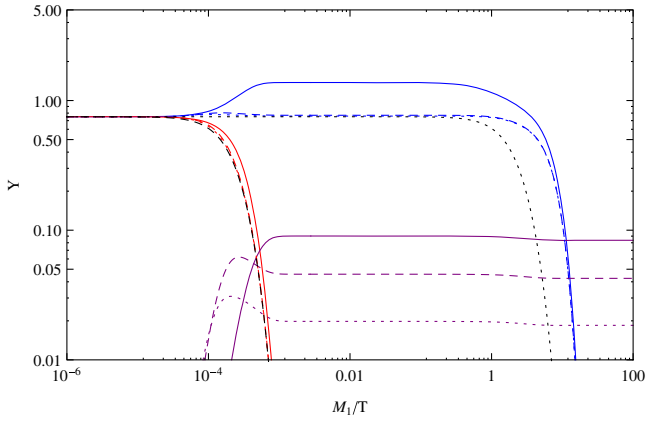
(i) $\{K_2, K_1\}$: Figures 15(a) and 15(b) exhibit the effects of K_2 and K_1 in the large mass regime. When the N_1 and N_2 phases can be hierarchically separated, they both act as independent one-flavor systems according to the respective efficiency factors $\kappa_{1,2}$. Consequently, we expect little deviation from the toy model that was studied previously.

Figures 15(c) and 15(d) exhibit the effects of K_2 and K_1 in the low-mass regime. This regime requires a smaller mass ratio to achieve sufficient CP asymmetry; therefore, the overlap of the N_2 and N_1 phases induces more intricate dynamics, though we observe that the impact of K_2 and K_1 can still be separated.

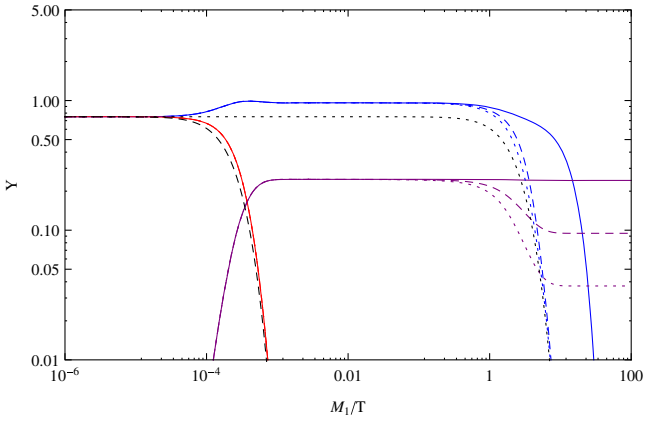
(ii) $\{K_{21}, \beta\}$: Figures 15(e) and 15(f) exhibit the effects of varying K_{21} and β in the low-mass scale regime. The primary effect is on the magnitude of the lepton asymmetry, via the impact on the CP asymmetry of N_2. However, there are also effects due to scattering. Indeed, in both cases the maximal lepton asymmetry is achieved for mid-range values, $K_{21} = 500$ and $\beta = 500$ GeV respectively. This illustrates the fact that, beyond a given threshold, increasing these parameters increases the scattering washout more significantly which more than compensates for the increase in the CP asymmetry.

V. CONCLUDING REMARKS

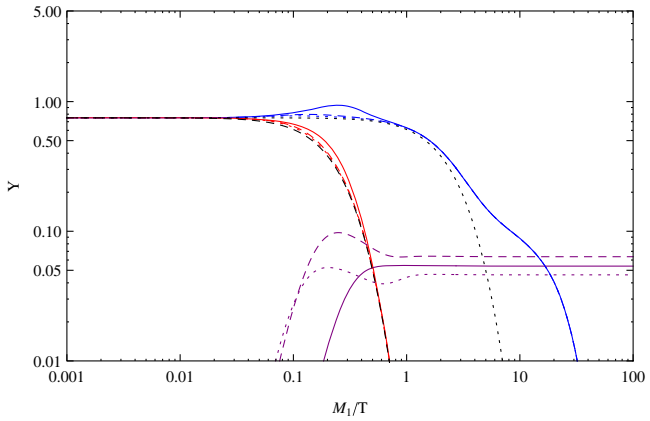
This paper has considered a minimal extension of leptogenesis that arises by opening up the Higgs portal with a new singlet scalar. This scalar can also couple at the renormalizable level to the RH neutrinos, which introduces a new (hidden sector) source of CP violation into the



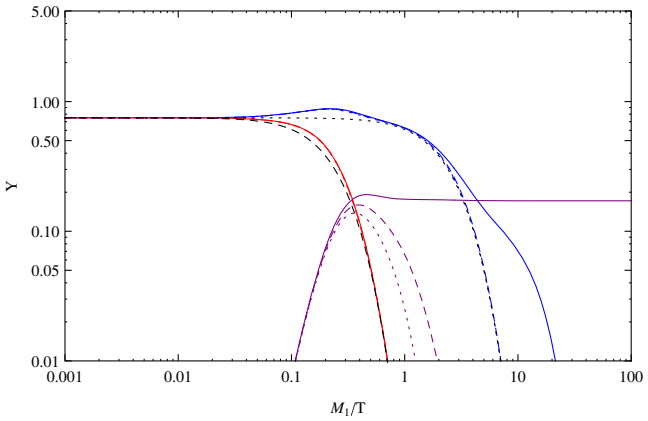
(a) $K_2 = \{0.1, 5, 10\}$ (plain, dashed, dotted)



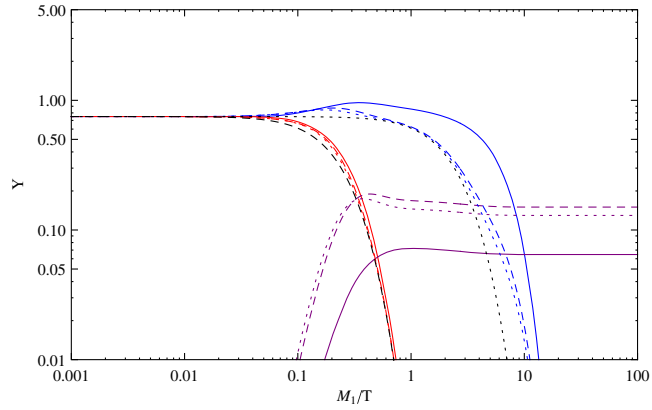
(b) $K_1 = \{0.1, 5, 10\}$ (plain, dashed, dotted)



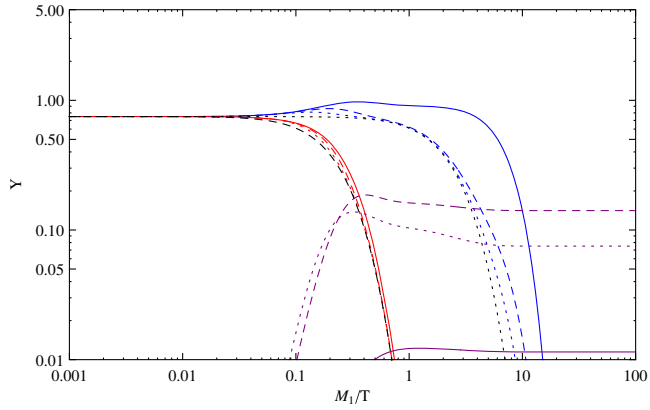
(c) $K_2 = \{0.1, 5, 10\}$ (plain, dashed, dotted)



(d) $K_1 = \{0.1, 5, 10\}$ (plain, dashed, dotted)



(e) $K_{21} = \{10, 500, 1000\}$ (plain, dashed, dotted)



(f) $\beta = \{10, 500, 1000\}$GeV (plain, dashed, dotted)

FIG. 15 (color online). The effect of $\{K_2, K_1, K_{21}, \beta\}$ on the dynamics for high $\{M_2, M_1\} = \{10^8, 10^4\}$ GeV (top two plots) and low $\{M_2, M_1\} = \{2, 0.2\}$ TeV (bottom four plots) mass regimes. The color coding, blue, red, purple, and black, is the same as Figs. 11(a), 12(a), and 13(a), while the dotted, dashed, and plain lines represent the variation of each parameter.

theory. The new RHN decay channels that are opened allow Higgs portal leptogenesis to avoid the stringent constraints of the Davidson–Ibarra bound, with viable low-scale scenarios that we have considered in detail. The new decay channels are only available for the next-to-lightest RH neutrinos, which has a number of interesting implications for phenomenology. We conclude in this section by mentioning a number of these as directions for future work.

(i) First-order leptogenesis: The new decay channels, e.g. $N_2 \to N_1 + S$, do not violate lepton number. Thus, this model falls into a general category in which the next-to-lightest RH neutrinos have both

L-violating and L-conserving decays. As has recently been emphasized [35], such models allow the original Weinberg–Nanopoulos theorem [36] to be evaded in that the loop-level amplitude can be of first-order in the L-violating vertex. This is clear from the analysis in Sec. II, and thus the HPL model is a simple example illustrating this general feature.

(ii) Light (CP-violating) sterile neutrinos: Since the new sources of the CP asymmetry arise from decays of the next-to-lightest RH neutrino states, it is possible to effectively decouple N_1 from leptogenesis. Indeed, since the normal hierarchy still allows one parametrically light (or massless) active neutrino, we can consider taking N_1 to be, for example, in the keV mass range for sterile neutrino dark matter. It would be interesting to explore whether the washout constraints on the interactions allow for viable thermal production modes in the early Universe. It is notable that, since α_{ij} contains multiple CP-odd phases, this model would generically imply some new low energy contributions (albeit suppressed) to CP-violating observables.

(iii) Dynamical seesaw scale: We assumed throughout that the scalar S was in a stable vacuum throughout the range of cosmological evolution relevant for leptogenesis. This need not be the case, and the full scalar potential $V(H,S)$ could allow for some evolution in $\langle S\rangle$, which would in turn affect the RH neutrino mass scale. Some of these issues were recently considered in Ref. [24], and it would be interesting to explore the implications of having an early epoch where, for example, the RH neutrino mass scale would pass through zero due to a phase transition in the scalar potential.[5]

ACKNOWLEDGMENTS

We would like to thank M. Pospelov for many helpful discussions. The work of M. L. and A. R. is supported in part by NSERC, Canada.

APPENDIX A: PARAMETRIZATION OF THE CP ASYMMETRY

The focus of this paper is on the *dynamics* of the Boltzmann evolution for HPL and the dependence on the masses and mass ratios. Therefore, rather than working with the full CP asymmetries computed in Sec. II, in this Appendix we will derive an order-of-magnitude estimate that will be more convenient to use in exploring the full evolution. To this end, we replace the coupling constants by their magnitudes, enabling components from the different chirality chains to be combined. Doing so effectively gives us an absolute upper bound on the asymmetries.

1. Hidden sector CP asymmetry

Starting with the hidden sector, and using the forms (19) and (23) for the vertex and wave function corrections, we write

$$|\epsilon_i^v| \sim \sum_j \frac{|(\lambda^\dagger\lambda)_{ji}|}{(\lambda^\dagger\lambda)_{ii}} \frac{|\alpha_{ij}|}{8\pi}\frac{\beta}{M_i}(\mathcal{F}^v_{jLL}+\mathcal{F}^v_{jRL}),$$
$$|\epsilon_i^w| \sim \sum_{l,j} \frac{|(\lambda^\dagger\lambda)_{ji}|}{(\lambda^\dagger\lambda)_{ii}} \frac{|\alpha_{lj}\alpha_{il}|}{8\pi}(\mathcal{F}^w_{jlLL}+\mathcal{F}^w_{jlRL}+\mathcal{F}^w_{jlLR}+\mathcal{F}^w_{jlRR}). \tag{A1}$$

Assuming the standard seesaw mechanism, we can replace the Yukawa couplings by the light active neutrino masses through the following relations [37]:

$$m_\nu = v^2\lambda^* M^{-1}\lambda^\dagger, \qquad \lambda = \frac{1}{v} U D_{\sqrt{m}} R^\dagger D_{\sqrt{M}}, \tag{A2}$$

where the Higgs vacuum expectation value is $v = 174$ GeV, R is a (complex) orthogonal matrix, and the diagonal Majorana and active neutrino mass matrices are respectively $D_M = \mathrm{diag}(M_1, M_2, M_3)$ and $D_m = \mathrm{diag}(m_1, m_2, m_3)$. The matrix U is the so-called unitary Pontecorvo-Maki-Nakagawa-Sakata matrix. Using this notation, we can write

$$(\lambda^\dagger\lambda)_{ji} = \frac{\sqrt{M_i M_j}}{v^2}\sum_b m_b R_{jb} R^*_{ib},$$
$$(\lambda^\dagger\lambda)_{ii} = \frac{M_i}{v^2}\sum_b m_b |R_{ib}|^2. \tag{A3}$$

The Schwartz inequality $|\sum_b m_b R_{jb} R^*_{ib}|^2 \le \sum_\alpha m_\alpha |R_{j\alpha}|^2 \sum_\beta m_\beta |R_{i\beta}|^2$ then allows the couplings to be bounded from above,

$$\frac{|(\lambda^\dagger\lambda)_{ji}|}{(\lambda^\dagger\lambda)_{ii}} \lesssim \sqrt{\frac{M_j\sum_b m_b|R_{jb}|^2}{M_i\sum_b m_b|R_{ib}|^2}} \sim \sqrt{\frac{M_j}{M_i}}. \tag{A4}$$

The final approximation assumes all the entries of the orthogonal matrix R are similar in magnitude. When R is real, the orthogonality condition $R^T R = \mathbf{1}$ ensures that its elements satisfy $R_{ij} \le 1$. When R is complex, we have fewer constraints, but bounding $|R_{ij}|$ by unity is a sufficient condition for orthogonality, and to obtain a characteristic estimate below, we will simply assume $|R_{ij}| \sim 1$.

Because of kinematic constraints, we have $\epsilon_1 = 0$. For $i = 2$, the neutrino sum runs over $j = 1, 3$. However, the asymmetry vanishes when $j = 3$, since $M_3 > M_2$ is kinematically forbidden for the cut loop. The sum thus reduces to $j = 1$ only. For the wave-function corrections, we have

[5] We thank Maxim Pospelov for suggesting this possibility and related discussions.

two sums over $j, l = 1, 3$. The index j denotes the neutrino inside the loop and is therefore constrained to $j = 1$ by the cut kinematics, but $l = 3$ is not forbidden. However, by assumption $r_{32} \gg 1$, and in that limit the resulting CP asymmetry is negligible. The sums therefore collapse to $j = l = 1$. Hence, we arrive at the parametric estimate,

$$\epsilon_2^v \sim \frac{|\alpha_{21}|}{8\pi}\frac{\beta}{M_2}\sqrt{\frac{M_1}{M_2}}(\mathcal{F}^v_{jLL} + \mathcal{F}^v_{jRL}),$$
$$\epsilon_2^w \sim \frac{|\alpha_{11}\alpha_{21}|}{8\pi}\sqrt{\frac{M_1}{M_2}}(\mathcal{F}^w_{jLL} + \mathcal{F}^w_{jlRL} + \mathcal{F}^w_{jlLR} + \mathcal{F}^w_{jlRR}). \tag{A5}$$

The functions $\mathcal{F}$ depend upon the variable $1/r_{21}$, which tends to zero in the hierarchical regime, $r_{21} \gg 1$. In that limit, we find the asymptotic behavior

$$\epsilon_2^v \sim \frac{|\alpha_{21}|}{8\pi}\frac{\beta}{M_2}\sqrt{\frac{M_1}{M_2}}(1-\sigma_2),$$
$$\epsilon_2^w \sim \frac{|\alpha_{11}\alpha_{21}|}{16\pi}\sqrt{\frac{M_1}{M_2}}(1-\sigma_2)^2. \tag{A6}$$

In Fig. 16(a), we compare the above approximate functions with the exact functions calculated in the main text; cf. Eqs. (19) and (23). The approximations are excellent for large M_2/M_1, but only deviate from the exact answer by a factor of 2 for $M_2/M_1 \gtrsim 10$.

2. Standard Yukawa *CP* asymmetry

We also recall the conventional contribution to the CP asymmetry, ϵ_0 [19]. In general, the loop-induced vertex and wave-function contributions from N_i decays are

$$\epsilon^v_{0i} = \sum_{j\neq i}\frac{\mathrm{Im}\{(\lambda^\dagger\lambda)^2_{ji}\}}{8\pi(\lambda^\dagger\lambda)_{ii}}\mathcal{F}^v_0(r_{ji}),$$
$$\epsilon^w_{0i} = \sum_{j\neq i}\frac{\mathrm{Im}\{(\lambda^\dagger\lambda)^2_{ji}\}}{8\pi(\lambda^\dagger\lambda)_{ii}}\mathcal{F}^w_0(r_{ji}), \tag{A7}$$

with

$$\mathcal{F}^v_0(r_{ji}) = \sqrt{r_{ji}}\left[1-(1+r_{ji})\log\left(\frac{1+r_{ji}}{r_{ji}}\right)\right],$$
$$\mathcal{F}^w_0(r_{ji}) = \frac{\sqrt{r_{ji}}}{(1-r_{ji})}. \tag{A8}$$

Using Eq. (A3) and the Schwartz inequality, we can again relate the Yukawa coupling to the light neutrino masses and bound from above the magnitude of the CP asymmetry. Taking $|R_{ij}| \sim 1$, we find

$$|\epsilon_{0i}| \sim \frac{\sum_\alpha m_\alpha}{8\pi v^2}\sum_{j\neq i} M_j|\mathcal{F}^v_0(r_{ji}) + \mathcal{F}^w_0(r_{ji})|. \tag{A9}$$

The sum over the active neutrino masses is constrained by cosmological data, with bounds in the range $\sum_\alpha m_\alpha <$ 0.2–1 eV [32,38]. The additional assumption of a normal hierarchy leads to a stronger constraint $\sum_\alpha m_\alpha \gtrsim m_3 \simeq \sqrt{\Delta m^2_{31}} \simeq 0.04$–$0.05$ eV at the 3σ level [2]. Throughout this paper, we assume a normal hierarchy for the light neutrinos, taking $\sum_\alpha m_\alpha \sim m_3 \sim 0.05$ eV. When the internal RHN is much heavier than the external neutrino, so that $r_{ji} \gg 1$, the loop function has the limit $|\mathcal{F}^v_0 + \mathcal{F}^w_0| \sim 3/(2\sqrt{r_{ji}})$. At the other end of the spectrum, when the internal RHN is much lighter, $r_{ji} \ll 1$, we find $|\mathcal{F}^v_0 + \mathcal{F}^w_0| \sim \sqrt{r_{ji}}|2 + \log(r_{ji})|$, as shown in Fig. 16(b). The CP asymmetry from N_1 decays receives contributions from internal $j = 2, 3$ heavy neutrinos, which imply $\sum_{j=2,3} M_j|\mathcal{F}^v_0(r_{j1}) + \mathcal{F}^w_0(r_{j1})| \sim 3M_1$. The CP asymmetry from N_2 decays receives a contribution from $j = 3$ giving $3M_2/2$ and a contribution from $j = 1$ giving

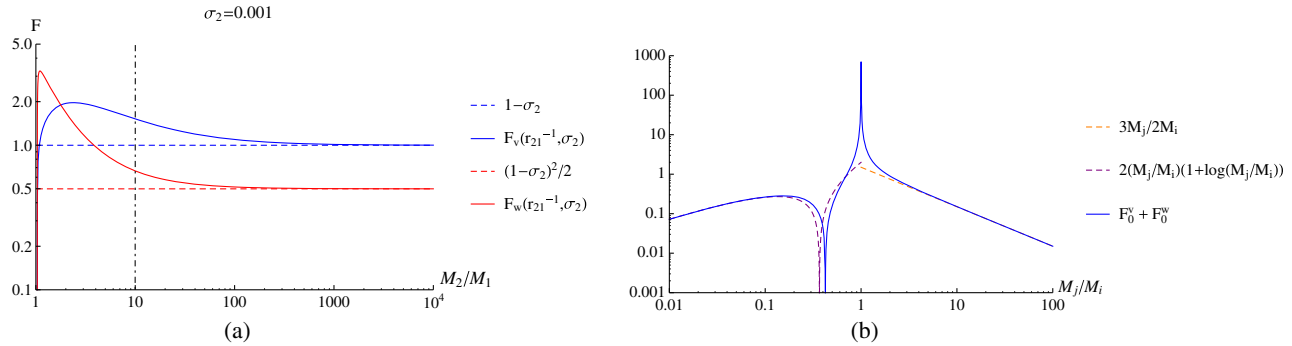


FIG. 16 (color online). Plots of the full loop functions $\mathcal{F}$ (solid) and the approximations (dashed) discussed in this section, as a function of the RHN mass hierarchy.

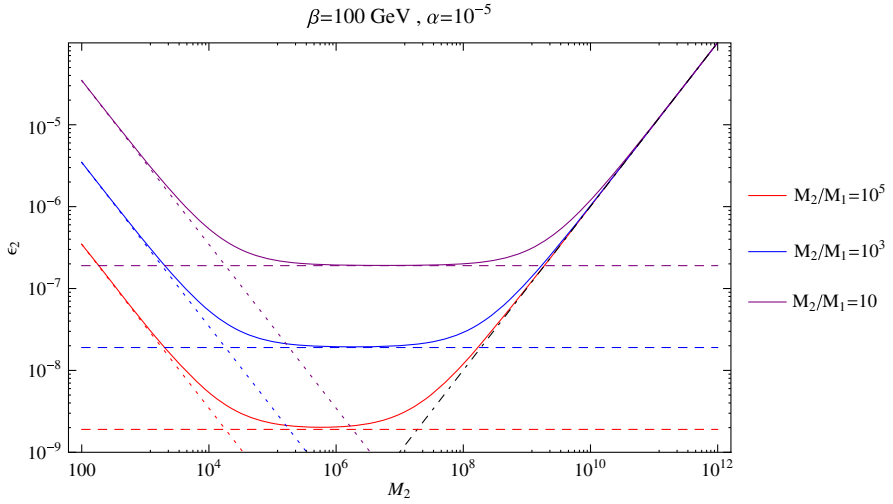


FIG. 17 (color online). The differing contributions to the total CP asymmetry ϵ_2 for various mass ratios M_2/M_1. The high-mass ratio is dominated by the Yukawa-sourced CP asymmetry, $\epsilon_0 \propto \sum_\alpha m_\alpha M_2/v^2$ (black, dotdashed line). The low-mass CP asymmetry is proportional to the β coupling via the vertex correction, $\epsilon_{(\beta)} \sim \sqrt{M_1/M_2}|\alpha\beta|/M_2$ (dotted lines). In the intermediate mass range, the CP asymmetry is dominantly generated by the wave function correction $\epsilon_{(\alpha)} \sim |\alpha|^2\sqrt{M_1/M_2}$ (dashed lines).

$M_2(M_1^2/M_2^2)|\log(M_2^2/M_1^2)| \ll M_2$, which is neglected. In total, we obtain the standard parametric scaling of the CP asymmetry,

$$|\epsilon_{01}| \sim \frac{3M_1\sum_\alpha m_\alpha}{8\pi v^2}, \qquad |\epsilon_{02}| \sim \frac{3M_2\sum_\alpha m_\alpha}{16\pi v^2}. \tag{A10}$$

This is analogous to the Davidson–Ibarra bound [11], though somewhat less strict as it depends on $\sum_\alpha m_\alpha$ rather than $m_3 - m_1$, which is a consequence of taking $|R_{ij}| \sim 1$. In the standard case, one can go further as the CP asymmetry is sensitive to the $(\lambda^\dagger\lambda)^2_{ij}$ elements which depend on R^2_{ij}, as can be seen from the representation in (A3). Thus, the orthogonality condition $\sum_k R^2_{ki} = 1$ can be used directly, leading to the conventional Davidson–Ibarra bound. The hidden sector CP asymmetry on the other hand, only depends on $(\lambda^\dagger\lambda)_{ij}$, in which case the orthogonality condition is less constraining, and we have taken $|R_{ij}| \sim 1$ to obtain a characteristic magnitude. For consistency, we have also used this approach to obtain the above magnitudes for the standard CP asymmetries. In practice, the scaling is very similar for the normal hierarchy, where $m_3 - m_1 \sim m_3 \sim \sum_\alpha m_\alpha$.

Combining both the standard and hidden sector contributions gives the total $\epsilon_{1,2}$ CP asymmetries,

$$|\epsilon_1| \sim \frac{3M_1\sum_\alpha m_\alpha}{8\pi v^2},$$
$$|\epsilon_2| \sim \frac{3M_2\sum_\alpha m_\alpha}{16\pi v^2} + \left(\frac{\beta}{M_2} + \frac{|\alpha_{11}|}{2}(1-\sigma_2)\right)\frac{|\alpha_{21}|}{8\pi}\sqrt{\frac{M_1}{M_2}}(1-\sigma_2). \tag{A11}$$

The ϵ_2 function is exhibited for a given set of parameters $\{\beta, \alpha\}$ and various mass ratios in Fig. 17.

APPENDIX B: BOLTZMANN EQUATIONS AND EQUILIBRIA

1. General considerations

For completeness, in this Appendix we review the relativistic formulation of the classical Boltzmann equations for the evolution of ensemble phase space densities in curved spacetime, and specifically for the Friedmann-Lemaître-Robertson-Walker metric. Recall that in a thermal quantum field theoretic context the classical Boltzmann evolution implicitly assumes that the effects of quantum coherence are negligible. In that case, the Boltzmann equation for the number densities n_i for a particle species "i" takes the form [39]

$$\frac{1}{a^3}\frac{\partial}{\partial t}(a^3 n_i) = \frac{\partial n_i}{\partial t} + 3Hn_i = \gamma_i, \qquad n_i = g_i\int\frac{d^3p}{(2\pi)^3}f_i, \tag{B1}$$

and

$$n_i^{\rm eq} = \frac{g_i m_i^3}{2\pi^2}\frac{T}{m_i}K_2\left(\frac{m_i}{T}\right), \qquad n_\gamma^{\rm eq}(z_i) = \frac{g_\gamma}{\pi^2}T^3,$$
$$H(T) = T^2\sqrt{\frac{8\pi^3 g_*}{90M_p^2}} \sim 1.66\sqrt{g_*}\frac{T^2}{M_p}. \tag{B2}$$

In these expressions H is the Hubble parameter, while the parameter g_i counts the number of degrees of freedom of the particle i' $g_N = 2$ for the RHN, $g_L = 2$ for components of the SU(2) doublets, $g_{e_R} = 1$ for the SU(2) singlets, and $g_H = 2$ for the SU(2) Higgs doublet. The parameter g_*

counts the effective number of degrees of freedom in the theory and typically within the extension of the SM one considers for standard Leptogenesis, $g_* \sim 100$. The left-hand side of (B1) incorporates information about the cosmology, while the right-hand side is the collision term and dictates how the interactions affect the number densities. Boltzmann expressed the collistion term via the *Stosszahlansatz* (collision number hypothesis),

$$\gamma_i = -\sum_{m,n}(\gamma_{i\to mn} - \gamma_{mn\to i}) - \sum_a \sum_{m,n}(\gamma_{ia\to mn} - \gamma_{mn\to ia}) + \cdots \tag{B3}$$

For one particular interaction $ia \to mn\ldots$, the collision term is written as

$$\gamma_{ia\to mn\cdots} = \int d\Pi_{ia} d\tilde{\Pi}_{mn\cdots} |i\mathcal{M}_{ia\to mn\cdots}|^2 f_i f_a (1 \pm f_m)(1 \pm f_n)\cdots, \tag{B4}$$

where $i\mathcal{M}_{ia\to mn\cdots}$ is an S-matrix element, and the phase space integrals are

$$\begin{aligned} \int d\Pi_{ia} &\equiv g_i g_a \int \frac{d^3 p_i}{(2\pi)^3 2E_i} \frac{d^3 p_a}{(2\pi)^3 2E_a}, \\ \int d\tilde{\Pi}_{mn\cdots} &\equiv \int \frac{d^3 p_m}{(2\pi)^3 2E_m} \frac{d^3 p_n}{(2\pi)^3 2E_n} \cdots (2\pi)^4 \delta(p_i + p_a - p_m - p_n - \cdots). \end{aligned} \tag{B5}$$

The "+" sign in $(1 \pm f)$ is for bosons, and the "−" is for fermions. These are the induced emission and Pauli blocking factors respectively [40]. However, we will assume that the gas of particles is dilute enough to use the classical Maxwell–Boltzmann approximation, $1 \pm f \approx 1$. Also, under the assumption that the scattering processes are fast enough to maintain kinetic equilibrium, the phase space densities and number densities are related by $n/n^{\rm eq} = f/f^{\rm eq}$ [39],

$$\gamma_{ia\to mn} = \frac{n_i}{n_i^{\rm eq}} \frac{n_a}{n_a^{\rm eq}} \gamma^{\rm eq}_{ij\to mn}, \tag{B6}$$

where $\gamma^{\rm eq}$ is given for $f_i = f_i^{\rm eq}$. The thermal cross sections $\gamma^{\rm eq}$ are given specifically for two-to-two scatterings $ia \to mn$ and decays $i \to mn$ by

$$\begin{aligned} \gamma^{\rm eq}_{i\to mn}(T) &= n_i^{\rm eq} \Gamma_{i\to mn} \left\langle \frac{M_i}{E} \right\rangle = n_i^{\rm eq} \frac{K_1(z_i)}{K_2(z_i)} \Gamma_{i\to mn}, \qquad z_i = \frac{m_i}{T}, \\ \gamma^{\rm eq}_{ia\to mn}(T) &= n_i^{\rm eq} n_a^{\rm eq} \langle v \sigma_{ia\to mn\cdots} \rangle = g_i g_a \frac{T^4}{32\pi^4} \int_{w_{\rm min}}^{\infty} dw \sqrt{w} K_1(\sqrt{w}) \hat{\sigma}_{ij\to mn}\left(w \frac{m_i^2}{z_i^2}\right), \end{aligned} \tag{B7}$$

where $w = s/T^2$ and $K_{1,2}(z)$ are the modified Bessel functions of the second kind. The decay rate $\Gamma_{i\to mn}$ in the above equations is calculated in the center-of-mass frame of particle i, while the ratio $\langle m_i/E_i \rangle = K_1(z_i)/K_2(z_i)$ is the thermal average of the Lorentz factor between the center-of-mass frame and any other frame [5,39,41]. We have also used above the reduced cross section, defined as

$$\hat{\sigma}_{ia\to mn}(s) = \frac{1}{s} \delta(s, m_i^2, m_a^2) \sigma_{ia\to mn}(s), \qquad \delta(a,b,c) = (a-b-c)^2 - 4bc. \tag{B8}$$

It is convenient to switch to the comoving system of variables,

$$z_i = \frac{m_i}{T}, \qquad Y_i = \frac{n_i}{n_\gamma^{\rm eq}}, \qquad Y_i^{\rm eq} = \frac{n_i^{\rm eq}}{n_\gamma^{\rm eq}} = \frac{3}{8} z_i^2 K_2(z_i). \tag{B9}$$

With these variables, the left-hand side of the Boltzmann equation transforms into $\partial n_i/\partial t + 3Hn_i = n_\gamma^{\rm eq} z_i H \partial Y_i/\partial z_i$, and the full equation now reads

$$n_\gamma^{\rm eq} z_i H \frac{\partial Y_i}{\partial z_i} = -\sum_{m,n}\left(\frac{Y_i}{Y_i^{\rm eq}} \gamma^{\rm eq}_{i\to mn} - \frac{Y_m}{Y_m^{\rm eq}} \frac{Y_n}{Y_n^{\rm eq}} \gamma^{\rm eq}_{mn\to i}\right) - \sum_a \sum_{m,n}\left(\frac{Y_i}{Y_i^{\rm eq}} \frac{Y_a}{Y_a^{\rm eq}} \gamma^{\rm eq}_{ia\to mn} - \frac{Y_m}{Y_m^{\rm eq}} \frac{Y_n}{Y_n^{\rm eq}} \gamma^{\rm eq}_{mn\to ia}\right). \tag{B10}$$

In addition, we can define the decay and scattering functions in the following way:

$$D_{i\to mn} \equiv \frac{\gamma^{\rm eq}_{i\to mn}}{n^{\rm eq}_\gamma H} = z_i^2 Y_i^{\rm eq}\frac{K_1(z_i)}{K_2(z_i)} K_{i\to mn}, \qquad K_{i\to mn} \equiv \frac{\Gamma_{i\to mn}}{H_i},$$
$$S_{ia\to mn} \equiv \frac{\gamma^{\rm eq}_{ia\to mn}}{n^{\rm eq}_\gamma H} = \frac{g_i g_a}{g_\gamma}\frac{m_i}{H_i}\frac{1}{32\pi^2} z_i \int_{w_{\rm min}}^\infty dw\sqrt{w}K_1(\sqrt{w})\hat{\sigma}_{ij\to mn}\left(w\frac{m_i^2}{z_i^2}\right). \tag{B11}$$

The Hubble rate $H_i = H(T = m_i)$. Through the Hubble time $t_i = 1/H_i$, we have a notion of the time scale before the equilibrium density of the massive particle i is Boltzmann suppressed. This is to be compared with the natural time scale set by the particle lifetime $\tau_i = 1/\Gamma_{i\to mn}$. If the lifetime is larger than the Hubble time, $\tau_i > t_i$, we anticipate a number density excess relative to equilibrium. Thus, the equilibrium parameter $K_{i\to mn} \equiv \Gamma_{i\to mn}/H_i$, as defined above, characterizes Sakharov's nonequilibrium condition. As noted in the main text, we follow the literature in using the notation K for the equilibrium parameters, to be distinguished from the modified Bessel function $K_i(z)$ by the presence in the latter of the argument z.

The above definition of the equilibrium parameter is not fully consistent since the decay rate is calculated at zero temperature. More precisely, we define thermal equilibrium parameters $\mathcal{K}_{i\to mn} = \langle\Gamma_{i\to mn}\rangle/H(T)$ for the decay rate, and similarly $\mathcal{K}_{ia\to mn} = n_i^{\rm eq}\langle v\sigma_{ia\to mn}\rangle/H(T)$ for the scattering processes. It turns out that the thermal equilibrium parameters are related to the decay and scattering functions D, S, defined earlier,

$$\mathcal{K}_{i\to mn} = \frac{\langle\Gamma_{i\to mn}\rangle}{H} = \frac{\gamma^{\rm eq}_{i\to mn}}{n_i^{\rm eq}H} = \frac{D_{i\to mn}}{Y_i^{\rm eq}},$$
$$\mathcal{K}_{ia\to mn} = \frac{n_i^{\rm eq}\langle v\sigma_{ia\to mn}\rangle}{H} = \frac{\gamma^{\rm eq}_{ia\to mn}}{n_a^{\rm eq}H} = \frac{S_{ia\to mn}}{Y_a^{\rm eq}}. \tag{B12}$$

Under the assumption of CP and CPT conservation, energy conservation implies that $\gamma^{\rm eq}_{ia\to mn} = \gamma^{\rm eq}_{mn\to ia}$. It is possible to relate the above equilibrium parameters to the parameters for the inverse processes, as follows:

$$\mathcal{K}_{mn\to i} = \frac{\gamma^{\rm eq}_{mn\to i}}{n_n^{\rm eq}H} = \frac{Y_i^{\rm eq}}{Y_n^{\rm eq}}\mathcal{K}_{i\to mn} = \frac{1}{Y_n^{\rm eq}}D_{i\to mn},$$
$$\mathcal{K}_{mn\to ia} = \frac{\gamma^{\rm eq}_{mn\to ia}}{n_n^{\rm eq}H} = \frac{Y_a^{\rm eq}}{Y_n^{\rm eq}}\mathcal{K}_{ia\to mn} = \frac{1}{Y_n^{\rm eq}}S_{ia\to mn}. \tag{B13}$$

Note that alternative definitions appear in the literature. For example, the decay function D in Ref. [5] corresponds here to $D_{i\to mn}/z_i Y_i^{\rm eq}$. The difference is a consequence of the present need to consider two-level leptogenesis. In Ref. [5], only terms of the form $1/Y_i^{\rm eq}$ appear in the Boltzmann equations, since only one RHN flavor is accounted for. Instead, we take into account scattering processes with external states mixing several RHN flavors, leading to terms like $1/Y_1^{\rm eq}Y_2^{\rm eq}$ or $(1/Y_i^{\rm eq})^2$. In addition, the definition $z_i = m_i/T$ explicitly brings in an arbitrary choice of reference mass scale, m_i. Therefore, we choose to keep the left-hand side of the Boltzmann equations in the form $z_i\partial Y_i/\partial z_i$ in order for it to be independent of that choice. Therefore, removing the $z_i Y_i^{\rm eq}$ from the definition of the decay functions seems more appropriate to the present case.

2. Example with leptonic RHN decay

To be specific, we will consider how these formulas apply to the case of leptonic RHN decays and inverse decays, $N_i \leftrightarrow LH$, directly relevant for leptogenesis. In equilibrium, we have

$$Y_i^{\rm eq} = \frac{n_i^{\rm eq}}{n_\gamma^{\rm eq}} = \frac{3}{8}z_i^2 K_2(z_i), \qquad Y_L^{\rm eq} = \frac{3}{4}. \tag{B14}$$

Note that the derivation of the Boltzmann equations in the form (B10) assumed Maxwell–Boltzmann statistics for all the particles. This assumption should technically lead to $Y_i^{\rm eq} = z_i^2 K_2(z_i)/2$. However, we know that relativistic fermions have an abundance $Y_L^{\rm eq} = 3/4$. To reproduce the relativistic RHN abundance at $T \gg M_i$, we multiply by an overall 3/4, so that $Y_i^{\rm eq}(z_i \to 0) = 3/4$ [6,40]. The thermal equilibrium parameters take the form

$$\mathcal{K}_{N_i\to LH} = K_{N_i\to LH}z_i^2\frac{K_1(z_i)}{K_2(z_i)},$$
$$\mathcal{K}_{LH\to N_i} = K_{N_i\to LH}z_i^2\frac{K_1(z_i)}{K_2(z_i)}\frac{Y_i^{\rm eq}}{Y_L^{\rm eq}}. \tag{B15}$$

Of course, it is a matter of convention to use $1/Y_L^{\rm eq}$ rather than $1/Y_H^{\rm eq}$ in the definition of the thermal equilibrium parameter for inverse decays. These two functions have been plotted in Fig. 18(a) for various parameters $K_{N_i\to LH}$. The equilibrium conditions for the decays and inverse decays are given as $\mathcal{K}_{N_i\to LH} > 1$ and $\mathcal{K}_{LH\to N_i} > 1$ respectively. The first condition translates to a restriction on $z_i > z_d$, while the second translates to the condition $z_{id}^+ > z_i > z_{id}^-$. The solutions z_d and $z_{id}^\pm$ have been numerically solved and plotted in Fig. 18(b).

APPENDIX C: DECAY RATES AND SCATTERING CROSS SECTIONS

This Appendix compiles the relevant decay rates and scattering cross sections used in the HPL Boltzmann equations. We summarize the relevant scattering processes in Fig. 19.

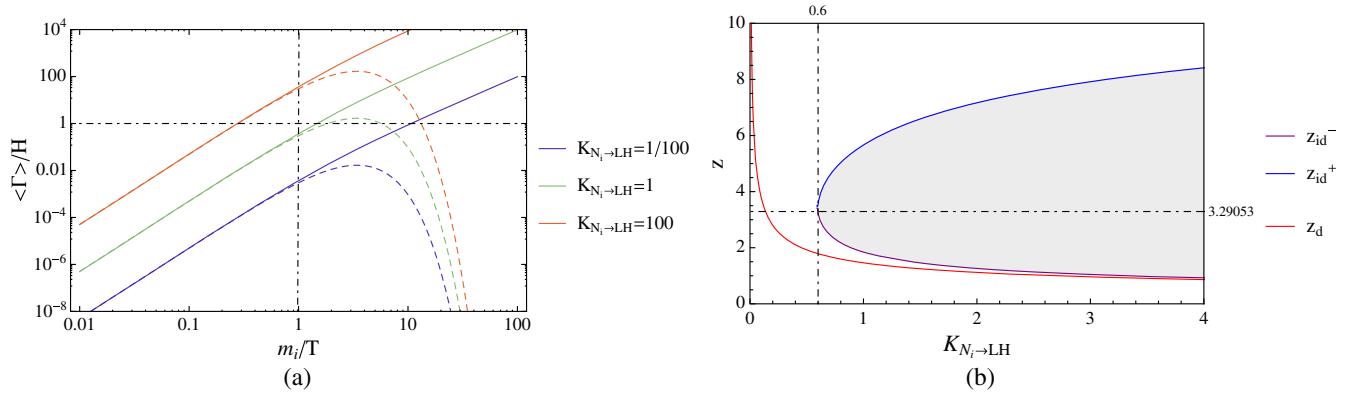


FIG. 18 (color online). The left plot shows the thermal equilibrium parameter $\mathcal{K}_{N_i\leftrightarrow LH}$ for both decays (plain) and inverse decays (dashed), for several equilibrium parameters $K_{N_i\leftrightarrow LH}$. The right plot shows the solutions z_d and $z_{id}^{\pm}$ to the equilibrium condition for decays and inverse decays respectively. Decays and inverse decays are both in equilibrium in the gray area, which only happens if $K \geq 0.6$.

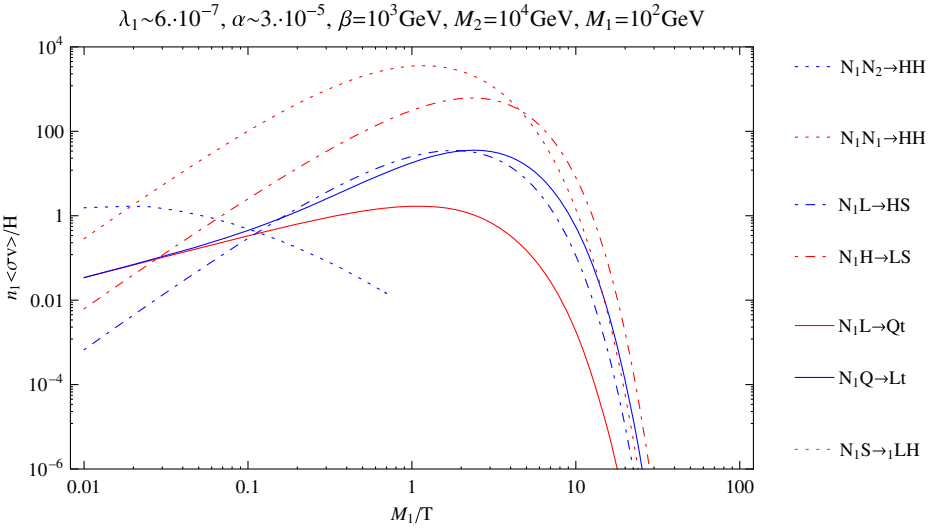


FIG. 19 (color online). This plot is a summary of the important scattering processes to account for, in order to properly manage the effect of N_1 on the lepton asymmetry. The corresponding Feynman diagrams are shown in the main text, in Fig. 7.

1. Decays and inverse decays

At tree level, one has the RHN decay rates to leptons,

$$\Gamma(N_i \to LH) = \Gamma(N_i \to \overline{LH}) = \sum_{k,\alpha,\beta} \Gamma(N_i \to L_k^\alpha H^\beta) = \frac{(\lambda^\dagger\lambda)_{ii}}{16\pi} M_i,$$
$$\Gamma_i = \Gamma(N_i \to LH) + \Gamma(N_i \to \overline{LH}) = \frac{(\lambda^\dagger\lambda)_{ii}}{8\pi} M_i. \tag{C1}$$

The indices $\alpha, \beta = 1, 2$ refer to the SU(2) doublets and enumerate the neutrino-type and the electron-type leptons. The index $k = 1, 2, 3$ runs through the three families, electron, muon, and tau. One of the hidden sector decays we encounter is $N_2 \to N_1 S$,

$$\Gamma(N_2 \to N_1 S) = \frac{|\alpha_{12}|^2 M_2}{16\pi} \left[\left(1 + \frac{M_1}{M_2}\right)^2 - \frac{m_S^2}{M_2^2}\right] \sqrt{\left(1 - \frac{M_1^2}{M_2^2} - \frac{m_S^2}{M_2^2}\right)^2 - 4\frac{M_1^2}{M_2^2}\frac{m_S^2}{M_2^2}}, \tag{C2}$$

where we used the approximation that $\alpha_{ij} = \mathrm{Re}\{\alpha_{ij}\}$, which is not generally true because of the complex phases contained in α, but the CP-odd contributions are not relevant here as the decay is L conserving.

2. Visible sector scattering

These processes involve the SM quarks and leptons in external states.

(i) s channel: the $N_i L \leftrightarrow Qt$ cross section and reduced cross section read

$$\sigma(N_i L \to Q\bar{t}) = \sum_{\alpha,k} \sigma(N_i L_k^\alpha \to Q\bar{t}) = \frac{(\lambda^\dagger\lambda)_{ii}}{8\pi}\frac{m_t^2}{v^2}\frac{1}{s},$$
$$\hat{\sigma}(N_i L \to Q\bar{t}) = \frac{(\lambda^\dagger\lambda)_{ii}}{8\pi}\frac{m_t^2}{v^2}\frac{(s-M_i^2)^2}{s^2}. \quad \text{(C3)}$$

(ii) t channel: We consider $N_i Q \leftrightarrow Lt$ and $N_i t \leftrightarrow LQ$. Because the leptons and quarks are assumed massless, these two channels are in fact equal. We have the cross section and reduced cross sections:

$$\sigma(N_i Q \to Lt) = \frac{(\lambda^\dagger\lambda)_{ii}}{8\pi}\frac{m_t^2}{v^2}\frac{1}{s}\left[\frac{s-2M_i^2+2m_h^2}{s-M_i^2+m_h^2} + 2\frac{M_i^2-m_h^2}{s-M_i^2}\ln\left(\frac{s-M_i^2+m_h^2}{m_h^2}\right)\right],$$
$$\hat{\sigma}(N_i Q \to Lt) = \frac{(\lambda^\dagger\lambda)_{ii}}{8\pi}\frac{m_t^2}{v^2}\frac{s-M_i^2}{s}\left[\frac{s-2M_i^2+2m_h^2}{s-M_i^2+m_h^2} + 2\frac{M_i^2-m_h^2}{s-M_i^2}\ln\left(\frac{s-M_i^2+m_h^2}{m_h^2}\right)\right]. \quad \text{(C4)}$$

Throughout this paper, we take the approximation that the zero-temperature Higgs mass $m_h = 0$, but in this limit, these cross sections are infrared divergent. The regulator to use, however, is the thermal mass which can potentially be quite large at leptogenesis temperatures. In practice, the cross section is only logarithmically sensitive to the regulator, and we therefore make the conventional choice $m_h/M_i = 10^{-5}$.

3. Hidden sector scattering

(i) s channel: $N_i L \leftrightarrow HS$ through the βHHS vertex. The cross section and reduced cross section are

$$\sigma(N_i L \to HS) = \sum_{\alpha,k}\sigma(N_i L_k^\alpha \to HS) = \frac{(\lambda^\dagger\lambda)_{ii}\beta^2}{8\pi}\frac{s-m_S^2}{s^3},$$
$$\hat{\sigma}(N_i L \to HS) = \frac{(\lambda^\dagger\lambda)_{ii}\beta^2}{8\pi}\frac{(s-m_S^2)(s-M_i^2)^2}{s^4}. \quad \text{(C5)}$$

(ii) s channel: $N_i S \leftrightarrow LH$, through the hidden sector vertex $\alpha_{ij} N_i N_j S$. This process is mediated by N_j, and the amplitude should thus be summed over all flavors. However, we shall simplify the discussion by considering only one internal flavor. In the the limit of massless S, the cross section takes the form

$$\sigma(N_i S \to LH) = \sum_j \sum_{\alpha,k} \sigma(N_i S \to L_k^\alpha H) = \sum_j \frac{(\lambda^\dagger\lambda)_{jj}|\alpha_{ij}|^2}{16\pi}\frac{(s+M_i^2)(s+M_j^2)-4sM_iM_j}{\sqrt{\delta(s,M_i^2,m_s^2)}((s-M_j^2)^2+\mathcal{E}_j^2)},$$
$$\hat{\sigma}(N_i S \to LH) = \sum_j \frac{(\lambda^\dagger\lambda)_{jj}|\alpha_{ij}|^2}{16\pi}\frac{(s-M_i^2)}{s}\frac{(s+M_i^2)(s+M_j^2)-4sM_iM_j}{(s-M_j^2)^2+\mathcal{E}_j^2}, \qquad \mathcal{E}_j = M_j\Gamma_j. \quad \text{(C6)}$$

(iii) s channel: $N_i N_j \leftrightarrow HH$ mediated by S. Taking the notation, $\delta(s,M_i,M_j) = (s-M_i^2-M_j^2)^2 - 4M_i^2M_j^2$, the cross section and reduced cross section are given by

$$\sigma(N_i N_j \to HH) = \frac{|\alpha_{ij}|^2\beta^2}{32\pi}\frac{s-(M_i+M_j)^2}{\sqrt{\delta(s,M_i,M_j)}(s-m_S^2)^2},$$
$$\hat{\sigma}(N_i N_j \to HH) = \frac{|\alpha_{ij}|^2\beta^2}{32\pi}\sqrt{\delta(s,M_i,M_j)}\frac{s-(M_i+M_j)^2}{s(s-m_S^2)^2}. \quad \text{(C7)}$$

(iv) t channel: $N_i H \leftrightarrow LS$ and $N_i S \leftrightarrow LH$, both mediated by a Higgs. Care is needed in computing these cross sections, because the Higgs mediator can be produced on shell. This is true even if the Higgs has a small but finite mass. Thus,

these processes will almost always be divergent, as they include the kinematic regime where the RH neutrino decays on shell to LH. It is therefore important to note that the cross section is regulated by the external neutrino decay width. This subtlety has been noted previously in a different context [42]. Starting with $N_i H \to LS$, the amplitude squared takes the form

$$\sum_{\alpha,\beta,k}\sum_{\text{spins}} |i\mathcal{M}(N_i H^\beta \to L_k^\alpha S)|^2 = 2\beta^2\left[\frac{(\lambda^\dagger\lambda)_{ii}M_i^2}{t^2+\mathcal{E}_i^2} - (\lambda^\dagger\lambda)_{ii}\frac{t}{t^2+\mathcal{E}_i^2}\right]. \tag{C8}$$

Upon integration over the transfer momentum, the second term leads to a logarithmic divergence. The first term is naively more problematic because it leads to a linear divergence, $1/\mathcal{E}_i$. However, in the narrow width limit, upon integration this term gives the delta function $\delta(t)$. This is the signature of an on-shell mediator, which splits the scattering into the two on-shell subprocesses $N \to LH$ followed by $HH \to S$. The first part is already accounted for in the Boltzmann equations and should be subtracted in order to avoid double counting [43]. In effect, this is a t-channel RIS. The subtracted scattering cross section we use is then

$$\begin{aligned}\sigma(N_i H \to LS) &= \frac{(\lambda^\dagger\lambda)_{ii}\beta^2}{16\pi(s-M_i^2)^2}\log\left(\frac{s^2(s-M_i^2-m_S^2)^2+s^2\mathcal{E}_i^2}{M_i^4 m_S^4+s^2\mathcal{E}_i^2}\right),\\ \hat{\sigma}(N_i H \to LS) &= \frac{(\lambda^\dagger\lambda)_{ii}\beta^2}{16\pi s}\log\left(\frac{s^2(s-M_i^2-m_S^2)^2+s^2\mathcal{E}_i^2}{m_S^4+s^2\mathcal{E}_i^2}\right), \qquad \mathcal{E}_i = M_i\Gamma_i.\end{aligned} \tag{C9}$$

The decay rate to account for in $\mathcal{E}_i$ is the total rate, that is $\mathcal{E}_1 = M_i\Gamma_1$ for N_1 but $\mathcal{E}_2 = M_2(\Gamma_2+\Gamma_{21})$ for N_2. This is model dependent, though we can safely assume that $\lambda^2, \alpha^2 \leq 10^{-5}$, which inspires our choice $\mathcal{E}_i/M_i^2 = 10^{-5}$. The cross section is only weakly dependent on the prescription, as the residual divergence is logarithmic.

[1] R. Mohapatra *et al.*, Rep. Prog. Phys. **70**, 1757 (2007); A. de Gouvea, in *Physics in D>=4: TASI 2004, Proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, CO, USA, 2004*, edited by J. Terning, C. E. M. Wagner, D. Zeppenfeld (World Scientific, Singapore, 2006), 197 M. Gonzalez-Garcia and Y. Nir, Rev. Mod. Phys. **75**, 345 (2003).

[2] F. Capozzi, G. L. Fogli, E. Lisi, A. Marrone, D. Montanino, and A. Palazzo, Phys. Rev. D **89**, 093018 (2014); G. L. Fogli, E. Lisi, A. Marrone, D. Montanino, A. Palazzo, and A. M. Rotunno, Phys. Rev. D **86**, 013012 (2012); M. Gonzalez-Garcia, M. Maltoni, J. Salvado, and T. Schwetz, J. High Energy Phys. 12 (2012) 123; D. Forero, M. Tortola, and J. Valle, Phys. Rev. D **86**, 073012 (2012).

[3] M. Fukugita and T. Yanagida, Phys. Lett. B **174**, 45 (1986).

[4] M. Luty, Phys. Rev. D **45**, 455 (1992).

[5] W. Buchmuller, P. Di Bari, and M. Plumacher, Ann. Phys. (Amsterdam) **315**, 305 (2005).

[6] S. Davidson, E. Nardi, and Y. Nir, Phys. Rep. **466**, 105 (2008).

[7] S. Blanchet and P. Di Bari, New J. Phys. **14**, 125012 (2012).

[8] S. Weinberg, Phys. Rev. Lett. **43**, 1566 (1979).

[9] F. R. Klinkhamer and N. Manton, Phys. Rev. D **30**, 2212 (1984); V. Kuzmin, V. Rubakov, and M. Shaposhnikov, Phys. Lett. **155B**, 36 (1985); P. B. Arnold and L. D. McLerran, Phys. Rev. D **36**, 581 (1987); **37**, 1020 (1988); J. A. Harvey and M. S. Turner, Phys. Rev. D **42**, 3344 (1990).

[10] G. Racah, Nuovo Cimento **14**, 322 (1937); W. Furry, Phys. Rev. **56**, 1184 (1939); J. Vergados, H. Ejiri, and F. Simkovic, Rep. Prog. Phys. **75**, 106301 (2012); S. Bilenky and C. Giunti, Mod. Phys. Lett. A **27**, 1230015 (2012).

[11] S. Davidson and A. Ibarra, Phys. Lett. B **535**, 25 (2002).

[12] G. Branco, R. G. Felipe, and F. Joaquim, Rev. Mod. Phys. **84**, 515 (2012).

[13] T. Hambye, Nucl. Phys. **B633**, 171 (2002); C. S. Fong, M. Gonzalez-Garcia, E. Nardi, and E. Peinado, J. High Energy Phys. 08 (2013) 104; T. Tsuyuki, Phys. Rev. D **90**, 013007 (2014); E. Ma, N. Sahu, and U. Sarkar, J. Phys. G **32**, L65 (2006).J. Racker, J. Cosmol. Astropart. Phys. 03, (2014) 025

[14] A. Pilaftsis and T. E. Underwood, Nucl. Phys. **B692**, 303 (2004); Phys. Rev. D **72**, 113001 (2005).

[15] R. Essig *et al.*, arXiv:1311.0029; F. S. Queiroz, K. Sinha, Phys. Lett. B **735**, 69 (2014)

[16] G. Engelhard, Y. Grossman, E. Nardi, and Y. Nir, Phys. Rev. Lett. **99**, 081802 (2007); O. Vives, Phys. Rev. D **73**, 073006 (2006); P. Di Bari, Nucl. Phys. **B727**, 318 (2005); S. Blanchet and P. Di Bari, J. Cosmol. Astropart. Phys. 06 (2006) 023.

[17] A. Sakharov, Pis'ma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).

[18] S. Profumo, M. J. Ramsey-Musolf, and G. Shaughnessy, J. High Energy Phys. 08 (2007) 010; A. Ahriche, Phys. Rev. D **75**, 083522 (2007); A. Biswas and D. Majumdar, Pramana **80**, 539 (2013).

[19] L. Covi, E. Roulet, and F. Vissani, Phys. Lett. B **384**, 169 (1996); E. Roulet, L. Covi, and F. Vissani, Phys. Lett. B **424**, 101 (1998); W. Buchmuller and M. Plumacher, Phys. Lett. B **431**, 354 (1998); A. Pilaftsis, Phys. Rev. D **56**, 5431 (1997).

[20] I. I. Bigi and A. Sanda, Cambridge Monogr. Part. Phys., Nucl. Phys., Cosmol. **9**, 1 (2000); M. Doi, T. Kotani, and E. Takasugi, Prog. Theor. Phys. Suppl. **83**, 1 (1985).
[21] E. I. Gates and K. L. Kowalski, Phys. Rev. D **37**, 938 (1988); A. Denner, H. Eck, O. Hahn, and J. Kublbeck, Phys. Lett. B **291**, 278 (1992); Nucl. Phys. **B387**, 467 (1992).
[22] R. Cutkosky, J. Math. Phys. (N.Y.) **1**, 429 (1960).
[23] E. W. Kolb and M. S. Turner, Front. Phys. **69**, 1 (1990).
[24] D. Aristizabal Sierra, M. Tortola, J. Valle, and A. Vicente, J. Cosmol. Astropart. Phys. 07 (2014) 052; M. Plumacher, Z. Phys. C **74**, 549 (1997).J. Racker and E. Roulet, J. High Energy Phys. 03 (2009) 065
[25] R. Barbieri, P. Creminelli, A. Strumia, and N. Tetradis, Nucl. Phys. **B575**, 61 (2000); A. Abada, S. Davidson, F.-X. Josse-Michaux, M. Losada, and A. Riotto, J. Cosmol. Astropart. Phys. 04 (2006) 004; E. Nardi, Y. Nir, E. Roulet, and J. Racker, J. High Energy Phys. 01 (2006) 164; T. Endoh, T. Morozumi, and Z.-h. Xiong, Prog. Theor. Phys. **111**, 123 (2004); A. Pilaftsis, Phys. Rev. Lett. **95**, 081602 (2005); P. S. B. Dev, P. Millington, A. Pilaftsis, and D. Teresi, Nucl. Phys. **B886**, 569 (2014).
[26] C. Kiessig and M. Plumacher, J. Cosmol. Astropart. Phys. 07 (2012) 014; B. Garbrecht and M. J. Ramsey-Musolf, Nucl. Phys. **B882**, 145 (2014); A. Pilaftsis and D. Teresi, Nucl. Phys. **B874**, 594 (2013); M. Beneke, B. Garbrecht, C. Fidler, M. Herranen, and P. Schwaller, Nucl. Phys. **B843**, 177 (2011); C. Kiessig and M. Plumacher, J. Cosmol. Astropart. Phys. 09 (2012) 012; T. Frossard, M. Garny, A. Hohenegger, A. Kartavtsev, and D. Mitrouskas, Phys. Rev. D **87**, 085009 (2013); C. P. Kiessig, M. Plumacher, and M. H. Thoma, Phys. Rev. D **82**, 036007 (2010).
[27] M. D'Onofrio, K. Rummukainen, and A. Tranberg, arXiv:1404.3565.
[28] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **112**, 201802 (2014).
[29] A. Abada, S. Davidson, A. Ibarra, F.-X. Josse-Michaux, M. Losada, and A. Riotto, J. High Energy Phys. 09 (2006) 010; E. Nardi, J. Racker, and E. Roulet, J. High Energy Phys. 09 (2007) 090.
[30] E. W. Kolb and S. Wolfram, Nucl. Phys. **B172**, 224 (1980); A. Strumia, arXiv:hep-ph/0608347.
[31] A. Pilaftsis, Phys. Rev. D **78**, 013008 (2008).
[32] P. Ade *et al.* (Planck Collaboration), arXiv:1303.5076.
[33] A. Coc, J.-P. Uzan, and E. Vangioni, arXiv:1403.6694.
[34] A. Kusenko, Phys. Rep. **481**, 1 (2009); K. Petraki and A. Kusenko, Phys. Rev. D **77**, 065014 (2008).
[35] A. Bhattacharya, R. Gandhi, and S. Mukhopadhyay, Phys. Rev. D **89**, 116014 (2014).
[36] D. V. Nanopoulos and S. Weinberg, Phys. Rev. D **20**, 2484 (1979).
[37] J. Casas and A. Ibarra, Nucl. Phys. **B618**, 171 (2001).
[38] M. Moresco, L. Verde, L. Pozzetti, R. Jimenez, and A. Cimatti, J. Cosmol. Astropart. Phys. 07 (2012) 053; J.-Q. Xia, B. R. Granett, M. Viel, S. Bird, L. Guzzo, M. G. Haehnelt, J. Coupon, H. Joy McCracken, and Y. Mellier, J. Cosmol. Astropart. Phys. 06 (2012) 010; E. Giusarma, R. de Putter, S. Ho, and O. Mena, Phys. Rev. D **88**, 063515 (2013).
[39] W. Buchmuller and M. Plumacher, Int. J. Mod. Phys. A **15**, 5047 (2000).
[40] F. Hahn-Woernle, M. Plumacher, and Y. Wong, J. Cosmol. Astropart. Phys. 08 (2009) 028.
[41] W. Buchmuller, R. Peccei, and T. Yanagida, Annu. Rev. Nucl. Part. Sci. **55**, 311 (2005).
[42] I. Ginzburg, arXiv:hep-ph/9509314; K. Melnikov and V. Serbo, Nucl. Phys. **B483**, 67 (1997).
[43] G. Giudice, A. Notari, M. Raidal, A. Riotto, and A. Strumia, Nucl. Phys. **B685**, 89 (2004).