Probing the properties of dark matter particles with astrophysical observations

Magalich, A.

Chapter 3

Big Bang Nucleosynthesis[1]

Big Bang Nucleosynthesis (BBN) is the earliest probe of the Universe that we currently have. It is sensitive to the physics that happened from ~ 1 to ~ 200 seconds after the Big Bang. The observable quantities are abundances of light elements that were fused during nuclear reactions in the Early Universe. They are sensitive to the presence of new particles that exist or decay during BBN.

In this chapter we describe the basic idea of the primordial nucleosynthesis (Section 3.1), discuss the prediction of the Standard Model (SM) with standard cosmology (Section 3.2) and compare it with observations (Section 3.3). Next, we discuss the effects of Beyond the Standard Model physics on BBN (Section 3.4) and specifically in the case of Heavy Neutral Leptons (HNL, see Section 3.5) that we would like to constrain. To do this we developed a numerical code that is described in Section 3.6. In the final section 3.7 we present our result and give conclusions.

3.1 Origin of chemical elements

Discovery of the constituents of atomic nuclei and their ability to combine into chemical elements posed a question: how the observed distribution of elements has been generated?

We know that many elements that are present at the Earth can be created by stellar evolution. However, there are regions in the Universe that are not much influenced by star formation (see detailed discussion in Section 3.3). Measurement of abundances of chemical elements in such regions indicates the existence of large amount of Helium-4 (about 25% mass fraction) and some other light elements like Deuterium, Helium-3, Lithium. The stellar origin of Deuterium is not probable because of its depletion in stars [75, 76]. Even more striking evidence is a measurement of Helium at the time of the decoupling of

[1] Results of this chapter are presented in papers [3, 4]. The main contribution of Andrii Magalich is the development of *pyBBN*, the numerical code that models primordial nucleosynthesis in presence of Heavy Neutral Leptons with masses above the pion mass that haven't been modelled previously. Treatment of chains of meson decays has been developed with major contribution by Nashwan Sabti.

cosmic microwave background [21]. The possible explanation of these measurements is the primordial origin of these elements.

Most of the heavy chemical elements originate from stars. However, we have reasons to believe that some chemical elements are primordial – created in the Early Universe when it was hot and homogeneous.

This poses a paradox: usually we imagine the Early Universe as being in equilibrium, but then the most abundant element would be Iron, because it has the largest binding energy per nucleon (Fig. 3.2).

We observe only light primordial elements, but not Iron. This means that nuclear reactions took place in the Early Universe, but were not in the thermodynamic equilibrium. It is possible if the rate of nuclear reaction Γ_{nuc} has the same order as the rate of expansion of the Universe H at the time of BBN and the conditions for nuclear reactions to be effective existed only for a relatively short period of time.

For the beginning of nuclear reactions the electric repulsion between protons has to be overcome. Strong interactions are extremely short-range, therefore either a very high density is needed, such that the distance between nucleons is of the order of characteristic scale of strong interactions (as it happens e.g. in neutron stars) or the temperature should be large enough, such that protons have enough kinetic energy to overcome electrostatic barrier and approach each other closely enough.

For the second scenario the temperature about $T \sim 10^9$ K ~ 100 keV is needed (this is a typical temperature in the cores of stars [77]). At such a temperature, the energy density of photons is very large. To estimate the concentration of baryons, Gamov assumed that a) it should be such that the reaction rate of nuclear reactions is of the order of the expansion rate of the Universe (see the discussion above) and b) the Universe is radiation dominated, as the density of baryons will be much lower. The latter assumption can be verified by the calculation (see below).

Assuming that the Universe is radiation-dominated, the Hubble rate can be estimated as

$$H \sim \frac{T^2}{M_{\text{pl}}} \tag{3.1.1}$$

The rate of nuclear reactions is given by

$$\Gamma_{\text{nuc}} = \langle \sigma_{\text{nuc}} v \rangle n_b, \tag{3.1.2}$$

where $\sigma_{\text{nuc}} \sim 10^{-29}$ cm^2 is a capture cross section of fast neutrons in hydrogen [77], $v \sim \sqrt{T/m_p}$ and n_b is a number density of baryons. From this one can easily estimate the baryon-to-photon ratio

$$\eta_b = \frac{n_b}{n_\gamma} \sim 10^{-10} \tag{3.1.3}$$

The Early Universe was radiation-dominated with baryon-to-photon ratio $\eta_B \sim 10^{-10}$.

3.2 Standard Model Nucleosynthesis

In this section, we will discuss the Standard Model BBN. We will start with a simplified qualitative picture, that is still able to predict helium abundance with precision $\sim 10\%$ (Section 3.2.1). Next, we will discuss a method on how to predict abundances of other light elements and how to estimate helium abundance more accurately (Section 3.2.2).

3.2.1 Simplified picture of Big Bang Nucleosynthesis

As we discussed in the previous section, the Early Universe was radiation-dominated ($\eta_b \sim 10^{-10}$). At the temperature ~ 100 keV that is relevant for nucleosynthesis there are only e^-, γ, ν, p, n in the plasma. At some point protons and neutrons start to fuse into nuclei.

In the Early Universe, only nuclear reactions with 2 initial particles are effective, as the probability of 3-particle reactions is low. Also, we know that stable elements with atomic numbers $A = 5$ and $A = 8$ are absent, see Fig. 3.1. This results in the absence of elements with a mass number larger than 7 during BBN.

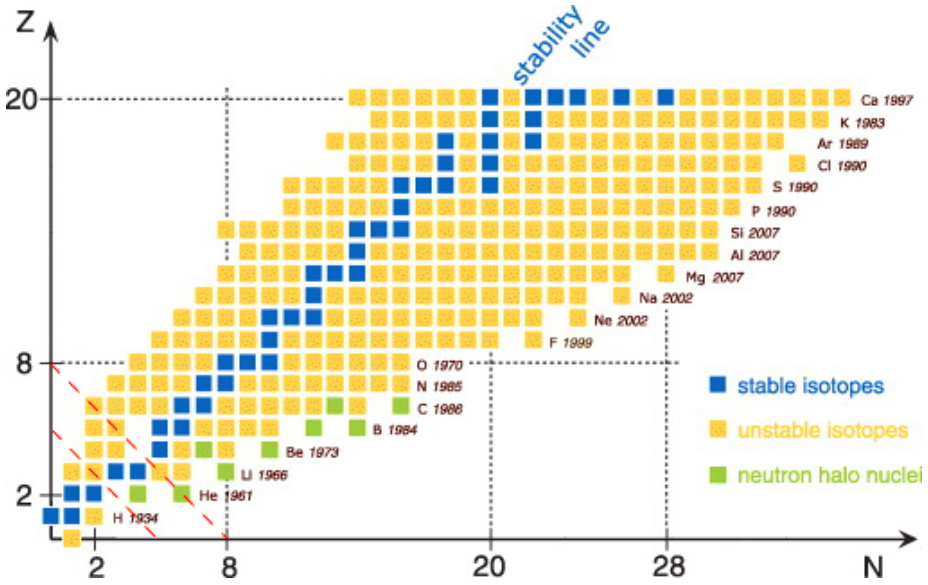


Figure 3.1: Chart of the nuclei stability. Here N is a number of neutrons, Z is a number of protons. Red dashed lines indicate the absence of a stable elements with atomic number $A = 5, 8$. Adopted from [78].

Production of the heavy elements is limited by the absence of stable nuclei with atomic numbers $A = 5$ and $A = 8$ – all elements heavier than Li^7 and B^7 are created through the 3-particle interactions in the dense medium of stars.

Among the lightest nuclei Helium-4 has the maximal binding energy per nucleon among light elements, see Fig 3.2. This means that the most probable reaction with light nuclei will be the formation of Helium-4, while reactions of helium dissociation have a much lower probability.

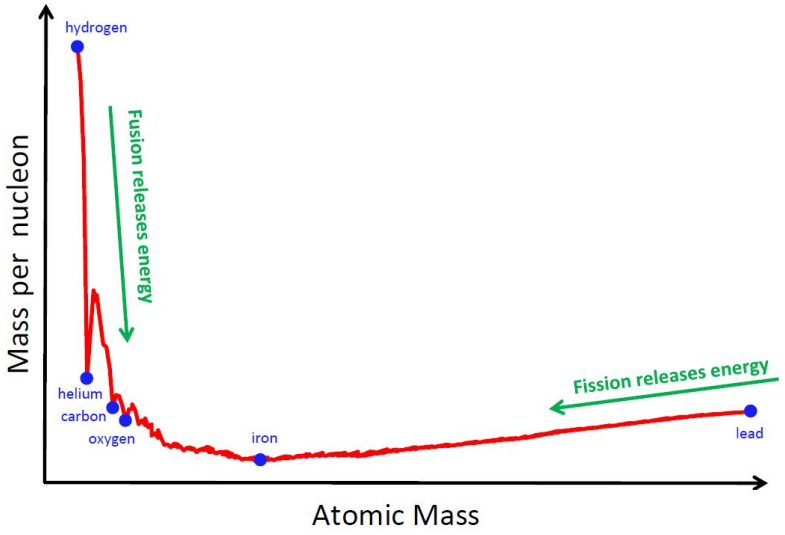


Figure 3.2: Mass per nucleon of stable nuclei. *Credit:* *oa-abruzzo.inaf.it*

Almost all neutrons and protons will fuse into Helium. The resulting amount of Helium is defined by the neutron-to-proton ratio at the start of the nuclear reactions.

Assuming that (almost) all available neutrons will be used to create helium, the mass abundance of Helium-4 can be estimated as

$$Y_{^4\mathrm{He}} \approx \frac{4m_p n_n/2}{m_p n_p + m_p n_n} = \frac{2\dfrac{n_n}{n_p}(t_{\mathrm{BBN}})}{1 + \dfrac{n_n}{n_p}(t_{\mathrm{BBN}})}, \tag{3.2.1}$$

where n_n and n_p are number densities of neutrons and protons, and t_{BBN} is a time of start of nucleosynthesis. Therefore, it is important to understand what is the value of neutron-to-proton ratio.

At large temperatures neutrons and protons were in thermal equilibrium because of weak reactions. In order to estimate $n_n/n_p(t_{\rm BBN})$, we need to establish whether the neutrons are in equilibrium at $t_{\rm BBN}$. Let us consider the effectiveness of weak $p \leftrightarrow n$ processes

$$e + p \leftrightarrow n + \nu, \quad \bar{\nu} + p \to e^+ + n \tag{3.2.2}$$

Before the decoupling density of neutrons or protons are given by thermal distribution,

$$n_A = g_A \left(\frac{m_A T}{2\pi}\right)^{3/2} e^{(\mu_A - m_A)/T}, \tag{3.2.3}$$

where

$$\mu_p + \mu_e = \mu_n + \mu_{\nu_e} \tag{3.2.4}$$

Electron chemical potential is negligibly small, since

$$\frac{\mu_e}{T} \sim \frac{\Delta n_e}{T^3} \sim \frac{n_p}{n_\gamma} \sim \eta_b \sim 10^{-9} \tag{3.2.5}$$

So neutron-to-proton ratio at the temperature of freeze out T_n is

$$\frac{n_n}{n_p}(T_n) \approx \exp\left(\frac{\mu_n - m_n}{T_n} - \frac{\mu_p - m_p}{T_n}\right) = \exp\left(-\frac{\Delta m}{T_n} - \frac{\mu_{\nu_e}}{T_n}\right), \tag{3.2.6}$$

where $\Delta m = m_n - m_p \approx 1.293$ MeV. In SM there is no reason to consider that $\mu_{\nu_e} \gg \mu_e$, so we take $\mu_{\nu_e} \approx 0$ and get

$$\frac{n_n}{n_p}(T_n) = \exp\left(-\frac{\Delta m}{T_n}\right) \tag{3.2.7}$$

The simple estimate of T_n is given by

$$\langle\Gamma_{n\to p}\rangle(T_n) \simeq H(T_n) = \frac{T_n^2}{M_{\rm Pl}^*}, \quad M_{\rm Pl}^* = \frac{M_{\rm Pl}}{1.66\sqrt{g_*}} \tag{3.2.8}$$

where $g_* \approx 10.75$ is the number of ultra-relativistic DoF at $T_n \simeq$ MeV and $\langle\Gamma_{n\to p}\rangle$ is thermally averaged neutron conversion rate,

$$\langle\Gamma_{n\to p}\rangle = \langle\sigma_{ne^+\to p\bar{\nu}_e} v_{ne^+}\rangle n_e + \langle\sigma_{n\nu_e\to pe} v_{n\nu_e}\rangle n_{\nu_e}, \tag{3.2.9}$$

where v_{ne^+} and $v_{n\nu_e}$ are relative velocities. Also we neglect neutron decays. Very rough simple estimate obtained assuming $T \gg \Delta m = m_n - m_p, m_e$ gives $\langle\Gamma_{n\to p}\rangle \simeq G_F^2 T^5$ and provides

$$T_n \simeq 1/(m_{\rm Pl}^* G_F^2)^{1/3} \simeq 1.4\ \text{MeV} \tag{3.2.10}$$

More accurate estimate for the interaction rates [79], that keep $m_e, \Delta m$ and include elec-

troweak corrections, gives

$$T_n \approx 0.72 \text{ MeV} \tag{3.2.11}$$

This temperature is larger than $T_{\text{BBN}} \simeq 100$ keV, so we need to trace changing of the n-p ratio for $T \lesssim T_{\text{BBN}}$. The neutron-to-proton ratio at this temperature T_n is

$$\left.\frac{n_n}{n_p}\right|_{T_n} \approx \frac{1}{6}. \tag{3.2.12}$$

At high temperatures, neutrons and protons are in thermal equilibrium. Because of the expansion of the Universe, the weak interactions decouple from equilibrium. This interrupts the nucleon conversion and fixes the initial neutron-to-proton ratio.

After freeze-out the ratio (3.2.7) changes only due to neutrons decays and comoving number densities change as

$$n_n(t) = n_n(T_n)e^{-t/\tau_n}, \qquad n_p(t) = n_p(T_n) + n_n(T_n)\left(1 - e^{-t/\tau_n}\right), \tag{3.2.13}$$

where τ_n is neutron lifetime and t is a time after freeze out T_n. So, the neutron-to-proton ratio at the beginning of BBN is

$$\frac{n_n}{n_p}(t_{\text{BBN}}) = \frac{\exp\left(-\frac{\Delta m}{T_n}\right) e^{-t_{\text{BBN}}/\tau_n}}{1 + \exp\left(-\frac{\Delta m}{T_n}\right)\left(1 - e^{-t_{\text{BBN}}/\tau_n}\right)} \tag{3.2.14}$$

Because of the neutron decay, the neutron-to-proton ratio at the beginning of BBN is very sensitive to the time t_{BBN} when nuclear reactions become effective. This is called the *cosmic chronometer*.

3.2.1.1 Start of nuclear reactions

The chain of 2-body nuclear reaction for the light elements is shown in Fig. 3.3. We see that to start nuclear reaction Deuterium should be formed in the first place. However, it has low binding energy and cannot be abundantly generated in the Universe full of high energy photons that can disintegrate it. The relative number of baryons to photons is very low ($\eta_b \sim 10^{-10}$), so typically there are many photons with energy $E > \Delta_D$ for each nucleus even when $T < \Delta_D$ because of the high-energy tail of photon distribution.

One can estimate the time when the abundance of Deuterium becomes large as a time when the number of the high-energy photon approximately equal to the number density of baryons,

$$n_\gamma(E > \Delta_D) \sim n_b. \tag{3.2.15}$$

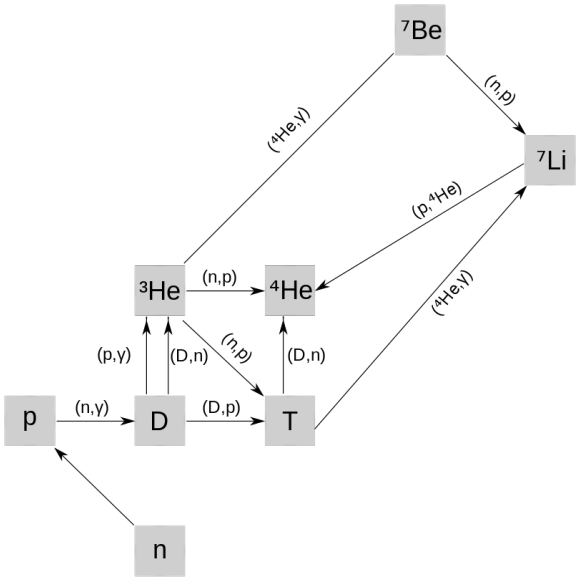


Figure 3.3: Nuclear reaction framework of nucleons and light elements. *Credit: www.aldebaran.cz*

Using Bose-Einstein distribution one can estimate number density of high-energy photons as

$$n_\gamma(E > \Delta_D) = 4\pi \int_{\Delta_D}^{\infty} \frac{p^2 dp}{e^{\frac{p}{T}} - 1} \approx 4\pi T^3 \int_{\Delta_D/T}^{\infty} e^{-x} x^2 dx = \tag{3.2.16}$$

$$= 4\pi T^3 \left(2 + 2\frac{\Delta_D}{T} + \frac{\Delta_D^2}{T^2}\right) e^{-\frac{\Delta_D}{T}} \approx 4\pi T \Delta_D^2 e^{-\frac{\Delta_D}{T}} \tag{3.2.17}$$

Hence, Deuterium becomes abundant starting from the temperature T_D satisfying

$$\eta_b \cdot \frac{\zeta(3)}{2\pi^3} \approx \frac{\Delta_D^2}{T_D^2} e^{-\frac{\Delta_D}{T_D}} \tag{3.2.18}$$

For $\eta_b = 10^{-10}$ this results in $T_D \approx 65\,\text{keV}$ (which is consistent with assumption $T_D \ll \Delta_D$).

To produce Helium (or any other element) through 2-particle interactions, nucleons need to go through Deuterium "bottleneck". Naively, its production should become effective at temperatures around the binding energy $\Delta_D = 2.2\,\text{MeV}$. However, Deuterium generation is challenged by a large amount of high-energy radiation. This delays nucleosynthesis until $T \sim 65\,\text{keV} \ll \Delta_D$

We are now ready to make a simple prediction for He abundance. The neutron-to-proton

ratio (3.2.14) is decreased to

$$n_n/n_p(t_{\rm BBN}) \approx \frac{1}{6} \times e^{-(t(t_{\rm BBN})-t(T_n))/\tau_n} \approx \frac{1}{7}. \tag{3.2.19}$$

So, the primordial abundance of Helium-4 we can be estimated from Eq. (3.2.1) as

$$\boxed{Y_{\rm He} \approx 0.227.} \tag{3.2.20}$$

This simple estimate has uncertainty of only $\simeq 10\%$.

A simple analytic prediction of He abundance based on two numbers: neutron freeze-out temperature T_n and beginning of deuterium nuclei T_D gives Helium abundance $Y_{\rm He} \approx 23\%$.

3.2.1.2 Neutron decay during Big Bang Nucleosynthesis

The intuitive picture that was used above for estimation of Helium abundance was that all available neutron at the beginning of BBN combines with protons to form Helium-4. This is correct if one can neglect neutron decay during BBN. Let us check that the rate of nuclear reaction is high in comparison to the rate of neutron decay, so the neutrons quickly become part of nuclei and do not decay.

At the start of the BBN the neutron decay competes with the neutron burning. Let us compare the decay rate

$$\Gamma_{n,\rm decay} = 1/\tau_n \approx 1.1 \cdot 10^{-3}\ {\rm s}^{-1} \tag{3.2.21}$$

with the rate of the nuclear reaction $\Gamma_{pn\to\gamma D}$. A simple estimate of the nuclear rate is given by

$$\Gamma_{pn\to\gamma D} = n_n \cdot \langle\sigma v\rangle \simeq \frac{2}{\pi^2}\frac{\eta_b}{7}\xi(3)T^3 \cdot \frac{\alpha}{m_\pi^2}\sqrt{\frac{T}{m_n}} \approx 6 \cdot 10^{-2}\left(\frac{T}{T_D}\right)^{7/2}\ {\rm s}^{-1}, \tag{3.2.22}$$

where we considered

$$\langle\sigma v\rangle \simeq \alpha_{\rm EM}/m_\pi^2\sqrt{T/m_n} \approx 1.5 \cdot 10^{-25}\ {\rm m}^2/s. \tag{3.2.23}$$

Here m_π^{-1} corresponds to the spatial range of strong interactions and α is the fine structure constant related to the creation of a photon, and used the neutron-to-proton ratio (3.2.19) and baryon-to-photon ratio to estimate $n_n = \frac{1}{7}\eta_b n_\gamma$.

Comparing (3.2.21) and (3.2.22), we conclude that $\Gamma_{pn\to\gamma D} \gg \Gamma_{n,\rm decay}$, and therefore the most of neutrons are captured in D before decaying.

As soon as BBN starts, we do not need to consider the decay of neutrons since most of them are incorporated in nuclei.

3.2.1.3 End of nuclear reactions

In this section, we will discuss when nuclear reactions stop. First, we will consider the general mechanism of freeze-out of nuclear reactions because the rate of reaction becomes smaller than the Hubble rate. However, as nuclei are positively charged their fusion can be stopped because of the Coulomb repulsion, which is known as the Coulomb barrier, so we will consider this effect below as well.

Freeze-out of nuclear reactions Let us make of conservative estimate of the nuclear reaction decoupling considering some schematic nuclear reaction $A + B \to C + D$. Similarly to the $p + n \to D + \gamma$ process, the reaction rate can be roughly approximated as

$$\Gamma_{\text{nuclear}} = n_A \cdot (\sigma v) \simeq \frac{2}{\pi^2}\eta_b \xi(3) T^3 \cdot \frac{1}{m_\pi^2}\sqrt{\frac{T}{m_C}} \approx 10^{-16}\left(\frac{T}{1\text{ MeV}}\right)^{7/2}, \tag{3.2.24}$$

where

$$\langle \sigma v \rangle \approx \frac{1}{m_\pi^2}\sqrt{\frac{T}{m_A}} \approx 1.7 \cdot 10^{-6}\text{ MeV}^{-2}\sqrt{\frac{T}{1\text{ MeV}}}, \tag{3.2.25}$$

and we considered $n_A \simeq n_n$, $m_A \simeq m_n$.

Comparing the interaction rate (3.2.24) with the expansion rate of the Universe, we can determine the decoupling temperature:

$$\frac{T^2}{m_{\text{Pl}*}} = H \simeq \Gamma_{\text{nuclear}} \tag{3.2.26}$$

The solution is given by

$$T_{\text{dec}} \simeq 1.9 \cdot 10^{-1}\text{ keV} \tag{3.2.27}$$

So we expect that at lower temperature nuclear reactions stop.

Coulomb barrier Let us now estimate the temperature at which the nuclei reaction is not possible because of the Coulomb barrier. Due to Coulomb repulsion, a typical cross-section contains an exponential factor depending on nuclear charges [80, 81]

$$\langle \sigma v \rangle \propto e^{-\eta}, \quad \eta = \frac{Z_1 Z_2 \alpha_{\text{EM}}}{v(T)} \approx \frac{Z_1 Z_2 \alpha_{\text{EM}}}{\sqrt{T}} \cdot \frac{\sqrt{A_1 A_2}}{\sqrt{A_1} + \sqrt{A_2}} \tag{3.2.28}$$

where Z is atomic number and A is mass expressed in atomic units.

As soon as this factor becomes small, reaction inevitably will freeze out. Then, for each pair of interacting nuclei we can compute a tentative interaction freeze-out temperature:

$$\eta \simeq 1 \Rightarrow T_{\text{Coulomb}} \simeq \frac{A_1 A_2 Z_1^2 Z_2^2}{(\sqrt{A_1}+\sqrt{A_2})^2} \text{ keV} \tag{3.2.29}$$

Coulomb barrier kick-in temperatures for some reactions are given in Fig. 3.1. They are

${}^{\mathrm{A}_1}_{\mathrm{Z}_1}\mathrm{X}_1$	${}^{\mathrm{A}_2}_{\mathrm{Z}_2}\mathrm{X}_2$	T, keV
${}^{1}_{1}\mathrm{p}$	${}^{2}_{1}\mathrm{D}$	0.34
${}^{2}_{1}\mathrm{D}$	${}^{2}_{1}\mathrm{D}$	0.5
${}^{2}_{1}\mathrm{D}$	${}^{3}_{1}\mathrm{T}$	0.6
${}^{3}_{1}\mathrm{T}$	${}^{4}_{2}\mathrm{He}$	3.4
${}^{7}_{3}\mathrm{Li}$	${}^{1}_{1}\mathrm{p}$	4.7
${}^{3}_{2}\mathrm{He}$	${}^{4}_{2}\mathrm{He}$	13.8

Table 3.1: Coulomb barrier suppression temperature T for 2-body nuclear reactions with initial nuclei ${}^{\mathrm{A}_1}_{\mathrm{Z}_1}\mathrm{X}_1$ and ${}^{\mathrm{A}_2}_{\mathrm{Z}_2}\mathrm{X}_2$.

close to the scale of decoupling of nuclear reactions, but Coulomb barrier is especially important in the generation of heavier nuclei of Lithium and Beryllium.

For a complete understanding of when nuclear reactions fall out of equilibrium and stop, one has to consider a full cross-section for each reaction.

Nuclear reactions between charged nuclei cease because of the freeze-out of nuclear reactions at the temperature of about 0.2 keV. Some nuclear reaction freeze-out earlier because of the Coulomb barrier.

3.2.2 Predictions beyond ${}^4\mathrm{He}$

Above we have shown how to predict the abundance of Helium. What about other elements?

Before the start of BBN, the concentration of Deuterium is very low because of the high amount of radiation. Because of this, the rate of reactions involving Deuterium as the reactant is also low ($\Gamma \ll H$). The chain of nuclear reactions starts when Deuterium is generated in significant numbers ($n_D \sim n_n$). This system is not in thermal equilibrium and its dynamics can be described by the kinetic approach.

The abundance of Helium is easily estimated because of its high binding energy (using a simple estimate (3.2.20) one can get it with $\sim 10\%$ precision). To predict the abundances of other elements we need to use **kinetic Boltzmann equations** because rates of nuclear reactions do not exceed Hubble rate and nuclei are not in equilibrium. We write these equations for nucleons and nuclei using the laboratory information about the nuclear cross-sections.

Boltzmann equations for different particles. Naively, to describe the system kinetically, one would write all possible Boltzmann equations [82] including all possible interactions. However, electromagnetic interactions are sufficiently fast at all times, so all electrically charged particles are considered to be in dynamical (but not necessarily chemical) equilibrium. These particles have a thermal distribution that is fully characterized by temperature (equal to photon temperature) and chemical potential. Fast interactions equilibrate individual temperatures of the species with the temperature of radiation (photons). This fact greatly reduces the system of kinetic equations and the number of quantities to find solutions for.

Nuclei are heavy and charged. Because of EM interaction, they are in kinetic equilibrium, so their velocity distribution is given by Boltzmann distribution. We can integrate the Boltzmann equations for nuclei to get equations on number density n.

Then, the following system of equation arises:

$$\frac{\partial f_{\nu_i}(t,y)}{\partial t} - Hp\frac{\partial f_{\nu_i}}{\partial p} = I^f_{coll}\{f,n\} \tag{3.2.30}$$

$$\frac{\partial n_n(t)}{\partial t} + 3Hn_n = I^n_{coll}\{f,n\} \tag{3.2.31}$$

$$\frac{\partial n_X(t)}{\partial t} + 3Hn_X = I^n_{coll}\{f,n\} \tag{3.2.32}$$

These equations describe the evolution of neutrinos ν_i, neutrons n and nuclei $X = D, T, He, Li, \ldots$

Electrons and photons are in thermal equilibrium and we know their distribution functions. Nucleons and nuclei are in kinetic equilibrium with the medium, so it is enough to use for them an integrated Boltzmann equation for number density. As neutrinos freeze-out, we need to use the full Boltzmann equation for them.

Because of the tiny baryon-to-photon ratio, nuclei negligibly influence the cosmological expansion. This allows separating the nuclear reactions as a subsystem on the cosmological background.

To close the system of equations, one has to supply the Friedmann equation and condition of energy conservation:

$$H^2 = \frac{8\pi G\rho}{3} \tag{3.2.33}$$

$$\frac{d\rho}{dt} = -3H(\rho + P) \tag{3.2.34}$$

These equations contain the following independent variables: distribution functions of the particles that departed from equilibrium f_ν, number densities of the neutrons and nuclei n_n, X, temperature T and scale factor a.

3.2.2.1 Big Bang Nucleosynthesis codes

Kinetic equations described above together with an expansion of the Universe give us a closed system of equations that can be solved numerically. There are plenty of codes that compute the abundances of the chemical elements:

- KAWANO [83] (also known as NUC123; 1992) – modified version of the Fortran code by Wagoner (1972). Solves the nuclear reactions network with simple neutron-to-proton conversion rates in presence of lepton asymmetry and various cosmological parameters

- FASTBBN [84] (1999) and JAVA calculator [85] (1999) – simple codes designed primarily to impose bounds on N_{eff}.

- AlterBBN [86] (2011) – public rewrite of KAWANO. Designed to compute abundances in alternative cosmologies (extra dof, quintessence, etc).

- PArthENoPE [87] (2018) – Fortran code that builds on KAWANO and adds the following effects. Improved calculations for the neutron-to-proton reactions are implemented via new fits, not as effective corrections added *a posteriori*. They also include effects of finite nucleon mass and non-thermal neutrino spectral distortions.

- PRIMAT [79] (2018) – Mathematica code that claims to provide 10^{-4} Helium accuracy. Comes with precomputed neutron-to-proton rates in various assumptions: radiative, zerotemperature, corrections, finite nucleon mass corrections, finite temperature radiative corrections, weak-magnetism, and QED plasma effects, which are for the first time all included and calculated in a self-consistent way.

The codes all agree in the predictions of the SM BBN and give the following results for the mass abundances of Helium-4 and Deuterium:

$$Y_{\rm He} \approx 0.247, \qquad {\rm D/H} \approx 2.68 \cdot 10^{-5}. \tag{3.2.35}$$

3.2.3 Summary of Standard Model Nucleosynthesis

BBN predicts the existence of light primordial nuclei: mainly Deuterium, Helium-3/4, Lithium, and Boron. The abundance of Helium-4 depends only on the neutron-to-proton ratio at the time of the start of nuclear reactions. Simple estimates allow us to obtain this value with accuracy of $\sim 10\%$ (Sec. 3.2.1.1). To get better accuracy and predict the abundances of other elements, we need to use the kinetic approach (Sec. 3.2.2).

Standard Model predictions depend on the baryon-to-photon ratio which can be measured in some independent experiment – e.g. from CMB. Additional uncertainty comes from the contradictory measurements of the neutron lifetime [88]. However, the effect of this uncertainty on Helium abundance is quite small.

Standard nucleosynthesis theory gives specific predictions for abundances of elements with high precision. The only free parameter of the Standard BBN is the *baryon-to-photon* ratio η_b. It can be measured in some independent experiments – e.g. from CMB.

3.3 Observations

In this section, we want to discuss which observables can be deduced from the experiments and compare our theoretical predictions with measured values. Moreover, since any measurement carries also an error, available data instructs us to what precision it is required to do theoretical calculations.

3.3.1 Helium-4

Primordial helium abundance has to be measured in regions with no star formation. Below, we discuss 3 methods that utilize metal-poor extragalactic regions, intergalactic medium and CMB.

Why these targets are used to infer He abundance? At the time of formation of CMB, the Universe was homogeneous and did not contain any stars – hence no star formation by definition. Later the Helium abundance will be the closest to the primordial in metal-poor environments (by "metals", astronomers refer to all elements heavier than He), like HII regions (zones of ionized Hydrogen around hot stars that emit a lot of UV radiation) of metal-poor star-forming galaxies, and gas clouds observed in absorption against a distant quasar. These environments have a composition very close to primordial because they were virtually not affected by star formation. This is confirmed by the observation of low metallicity in these objects (100 times lower than in Sun), which is known by the extreme weakness of emission/absorption lines of metals in their spectra (dominated by H and He lines). If there was vigorous star formation occurring at some epoch in these objects, the metals lines observed in spectra, will be significantly stronger, which is not the case. Thus, extremely low metallicity indicates that these targets didn't undergo processes responsible for change of chemical composition, and that's why they are believed to have nearly original (primordial) composition.

3.3.1.1 Low-metallicity extragalactic method

Using the measurements in low-metallicity HeII regions, one can build the $Y - O/H$ relation. Extrapolating it to 0 metallicity one is able to compute the primordial Helium abundance [89, 90], see Fig. 3.4. This method gives [90] at 95% confidence

$$Y = 0.2551 \pm 0.0022 \tag{3.3.1}$$

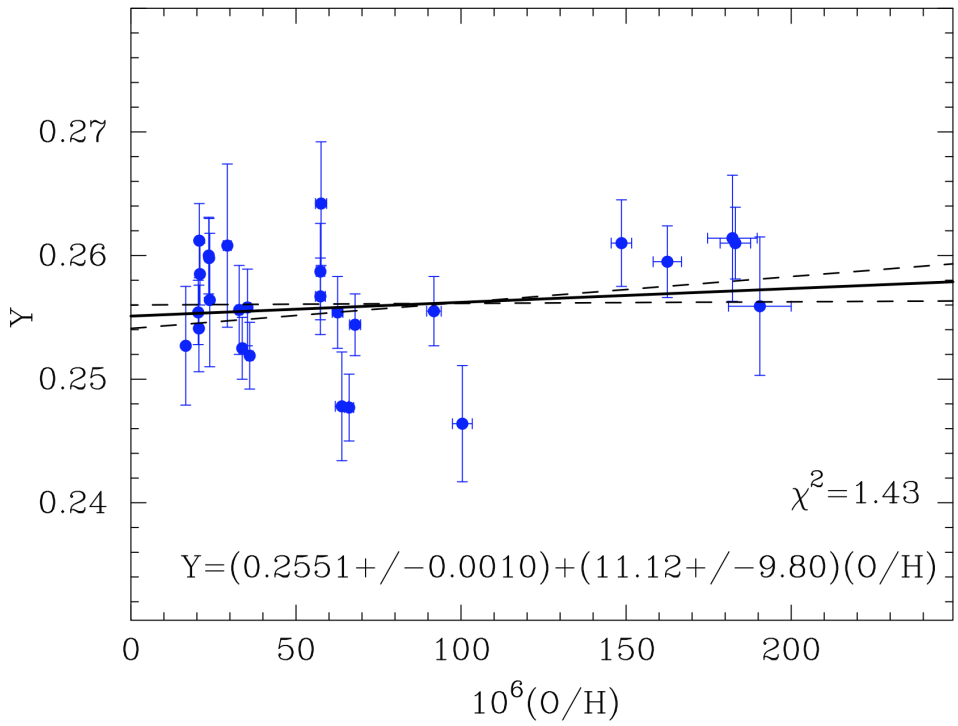


Figure 3.4: Helium abundance in 28 HII low metallicity regions as a function of Oxygen-to-Hydrogen ratio. The lines show the results of the linear regression to the data. The primordial value is measured by extrapolation of O/H to zero. *Credit: Izotov et al. [90]*

Although formal statistical error bar of this method are at sub-% level, the measurement is known to be dominated by the systematic uncertainties, see [91, Section 4.3] for a critical overview in this method.

The method of measuring Y_p from low-metallicity extragalactic regions formally gives very precise results ($\Delta Y/Y < 1\%$), but is expected to have large systematic errors.

3.3.1.2 Intergalactic medium method

The work [92] uses observational data of gas clouds with low metal content, which are observed against a quasar, located far behind the cloud (not in its vicinity). These clouds are composed mostly of Hydrogen and Helium, and very little of other elements. They are highly suitable for robust determination of the Helium abundance because the concentration of Hydrogen is high enough to produce strong absorption that can be detected, but at the same time not too high to distort the absorption line profile and make it impossible to extract exactly the H and He abundances. If the density is too high in the cloud, the atoms absorbing photons are colliding with other atoms before they de-excite. The lifetime of the electron in the excited state is reduced, and this yields a broader line profile due to the Heisenberg uncertainty principle, eventually saturating and distorting the absorption line

profile. Reconstruction of the column densities for saturated spectral lines is not reliable. On the other hand, at moderate densities, the atom is de-excited before colliding with another atom, and the resulting line profile has a pure Lorentzian shape determined only by the lifetime of the atom in the excited state ($\sim 10^{-8}$ s.) and the column density of the atom. Using pure Lorentzian profiles one can unambiguously and reliably find the column density of an element.

The neutral Hydrogen (HI) column density is inferred from the flux decrease beyond 91.2 nm wavelength (13.6 eV, the threshold energy for ionization of Hydrogen). The photons having shorter wavelengths are able to ionize Hydrogen and thus less of these photons will reach us. The flux decreased N_{obs}/N_{emit} is proportional to the column density of neutral Hydrogen n_{HI}:

$$N_{obs} = e^{-\tau} N_{emit}, \qquad \tau = L/\lambda, \qquad \lambda = \frac{1}{\sigma_{ion} n_{HI}}, \tag{3.3.2}$$

where τ is the *optical depth*, L is the distance and λ is the mean free path given by ionization cross section σ_{ion}. The Helium column density is retrieved from neutral Helium (HeI) absorption lines. The equivalent width of an absorption line (the area of the line divided by the continuum emission level) is a measure of how much flux was absorbed, and thus is proportional to the column density.

Important remark: since only neutral elements are seen in these spectra, it is necessary to estimate the ionization level of the cloud. The authors of [92] use simulations to illuminate gas with UV radiation imitating conditions (exposure due to quasars, galaxies, etc) in the observed cloud and extracting the corresponding ionization level.

The same method has been earlier applied to the determination of Deuterium [93].

$$Y = 0.250^{+0.033}_{-0.025}, \qquad \mathrm{D/H} = (2.527 \pm 0.030) \cdot 10^{-5} \qquad (68\% \text{ confidence}) \tag{3.3.3}$$

The IGM method studies the regions of lower metallicity than the extragalactic method (by $\sim 30\%$). These regions are supposedly closer to the primordial composition. However, this method has a much larger statistical error of $\Delta Y/Y \sim 10\%$.

3.3.1.3 Cosmic Microwave Background method

Planck collaboration [21] provide a Helium measurement based on the free electrons density between Helium and Hydrogen recombination. The damping tail of CMB anisotropies is sensitive to the electron density and it is possible to measure this effect when Helium is already recombined ($z \sim 2000$), but before Hydrogen recombination ($z \sim 1100$).

This effect of Helium abundance is however partially degenerate with the effective

number of relativistic degrees of freedom N_{eff}, requiring a simultaneous fit:

$$Y = 0.246 \pm 0.035 \tag{3.3.4}$$

$$N_{eff} = 2.97^{+0.58}_{-0.54} \tag{3.3.5}$$

at 95% confidence interval, using Planck TT,TE,EE+lowE+lensing+BAO [21, Section 7.6.2].

Measurement through CMB guarantees that the Helium abundance is not affected by stellar nucleosynthesis. The error of this method is about $\Delta Y/Y \sim 14\%$. However, Planck's collaboration emphasizes those systematics of the polarization spectra have not been accurately characterized.

3.3.2 Deuterium

Because of its small binding energy, Deuterium is not created by other sources and is destroyed in stars [94]. This means that any measurement of Deuterium is guaranteed to be not higher than the primordial value. The local interstellar value of $\mathrm{D/H}$ ratio is $\mathrm{D/H} = (1.56 \pm 0.40) \cdot 10^{-5}$ [95].

High-resolution absorption spectra of quasars allow measuring Deuterium in Intergalactic Medium environments with metallicities $\sim 10^3$ times smaller than Solar. Since the Lyman-series transitions of different elements are shifted, this is in principle can be used to identify absorption lines of two elements like Hydrogen and Deuterium (or Helium) and to deduce the abundance from the relative intensities. In reality, this, however, is difficult because of contamination by Hydrogen Lyman-α forest lines and requires accurate knowledge of the HI column density corresponding to absorbers.

Damped Lyman-α systems corresponding to dense regions in high-redshift galaxies allow to measure the HI density independently of the cloud model and there exists a number of systems where Deuterium line was identified and $\mathrm{D/H}$ abundance was measured:

$$\mathrm{D/H} = (2.569 \pm 0.027) \cdot 10^{-5} \tag{3.3.6}$$

This method is also used as an independent probe for $^4\mathrm{He}$.

Any measurement of Deuterium provides an upper bound on primordial abundance.

3.3.3 Helium-3

There is no reliable data on the primordial abundance of Helium-3. This isotope is only measured in the Solar system and HII regions of the Milky Way (that also have high metallicity) [96]. At this point, it is not possible to make conclusions about the effect of

stellar nucleosynthesis on the Helium-3 abundances since theoretical models are inconsistent with observations [97].

At the moment there is no $^3\mathrm{He}$ data available to constrain BBN.

3.3.4 Lithium

Lithium is best measured in the Population II metal-poor stars of the Milky Way. It is notable that Lithium abundance exhibits the *Spite plateau* – is nearly constant in stars with metallicity $\lesssim 1/30$ of the Solar value [98]. Considering this abundance as a primordial one gets

$$\mathrm{Li/H} = (1.6 \pm 0.3) \cdot 10^{-10} \tag{3.3.7}$$

However, in extremely metal-poor stars with $\mathrm{Fe/H}$ ratio $\lesssim 10^{-3}$ than Solar, no Lithium is detected. This suggests some mechanism destroying $^7\mathrm{Li}$ as well as that the abundance at the Spite plateau might also be modified from the primordial value. This, in fact, makes the method of 0-metallicity extrapolation inapplicable to Lithium. Instead, the measured abundance (3.3.7) should be considered a lower bound on the primordial value [99].

Recent observations indicate that Lithium might be destroyed in low metallicity stars and hence the measurements constitute the lower bound on primordial ratio $^7\mathrm{Li/H}$.

3.3.5 Comparison of measurements with Standard Model Nucleosynthesis predictions

Since Standard BBN predictions depend only on measurements of the baryon-to-photon ratio and neutron lifetime, abundances of the elements can be used to check the validity of the theory.

Baryon-to-photon ratio can be independently measured from CMB [100]:

$$\eta_B = (6.118 \pm 0.041) \cdot 10^{-10} \tag{3.3.8}$$

Fig. 3.5 summarizes the available measurements and predictions. This figure shows long-standing Lithium Problem: Standard BBN theory predicts Lithium abundance ~ 3 times larger than measured in metal-poor stars. However, new observational data of Lithium show that the previous measurements should be considered as a lower bound on the primordial abundance, so the Standard BBN theory is consistent with observations.

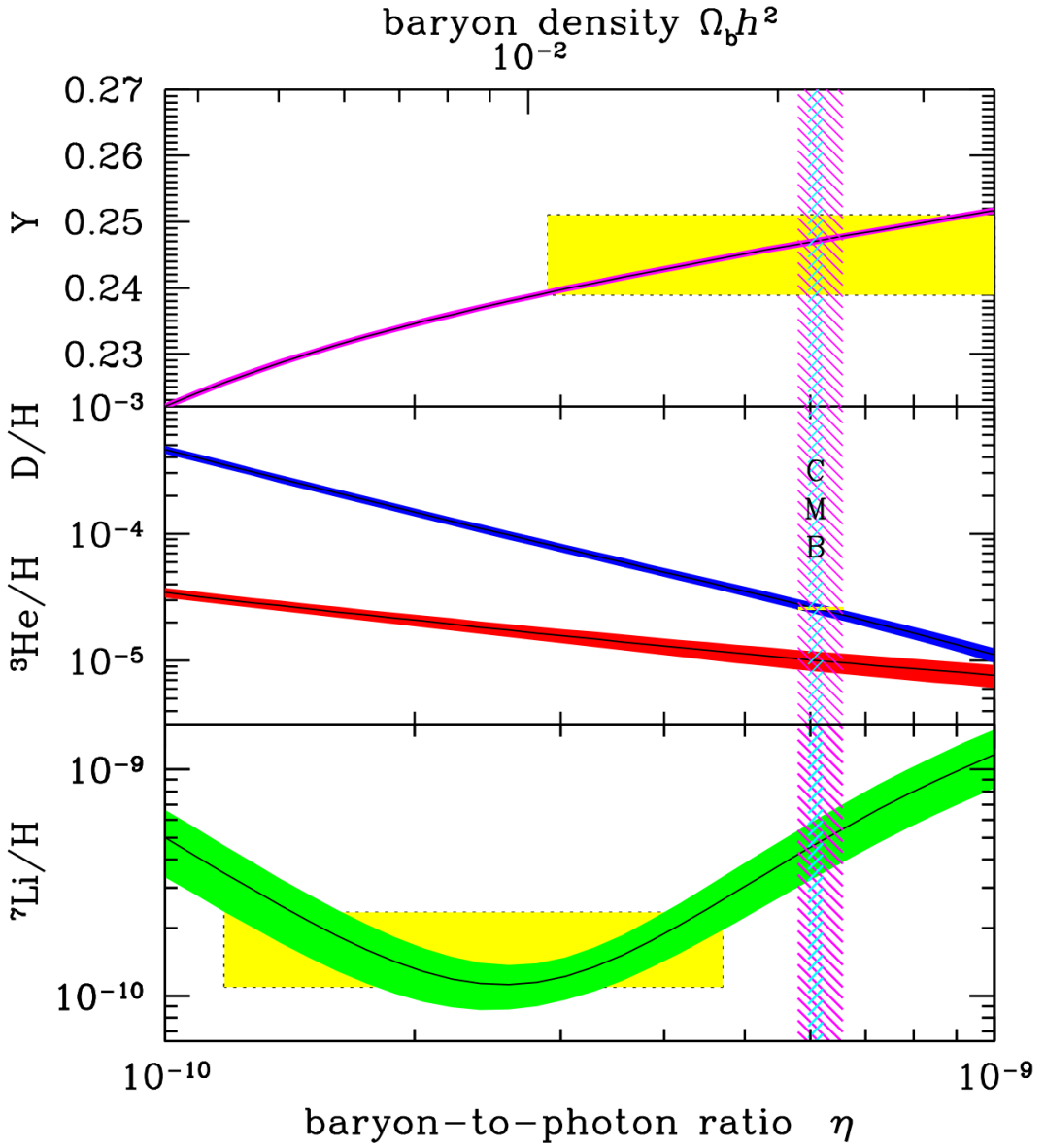


Figure 3.5: Predictions of the Standard BBN theory (with uncertainty due to error in neutron lifetime τ_n; black curves with coloured bands) compared to the measurements of primordial elements (yellow rectangles) and determination of the baryon-to-photon ratio η_B from CMB (vertical crosshatched band). *Credit: Cyburt et al. [101]*

Theoretical predictions of Standard BBN are in good agreement with the modern measurements of baryon-to-photon ratio and element abundances.

The only significant deviation is with the measurement of Lithium which is lower than predicted value by a factor of ~ 3. However, recent observations indicate that Lithium might be destroyed in low metallicity stars and hence the measurements

constitute the lower bound on primordial ratio $^7\mathrm{Li}/\mathrm{H}$.

Any Beyond the Standard Model physics should not break this agreement.

3.4 New physics and Big Bang Nucleosynthesis

As we have seen above, the predictions of the SM nucleosynthesis are consistent with observations. This puts a requirement on any extension of the Standard Model – its predictions for nuclear abundances should be different from the SM ones only within observational error bars. This makes BBN a powerful tool to constrain new physics.

Possible effects of new physics on BBN include:

- Change of Hubble expansion rate during BBN
- Change of neutron-to-proton ratio at and after freeze-out of weak interactions
- Disintegration of nuclei due to collisions with the new particle or its decay products
- Entropy production: the creation of additional photons or electrons between BBN and CMB and, therefore a change of the baryon-to-photon ration at BBN time (as compared to the value observed e.g. by CMB).

3.4.1 Excluded domain of the parameter space

Typically a new particle is characterized by its mass M and coupling constant θ. Then the lifetime of the particle is inversely proportional to it: $\tau \propto \theta^{-2}$. This naturally gives us a bound when $\tau \ll 1sec$ – these particles vanish long before the BBN. On the other hand, there also exists a lower bound on the particle coupling, since the interaction rate becomes negligible.

New physics particles (NPPs) can contribute to each of the effects of new physics on BBN considered above. As lifetime is related to the coupling constant as $\tau \propto \theta^{-2}$, it is convenient to separate the effects of independence on the lifetime τ_X of NPP.

3.4.1.1 $\tau_X \gtrsim 0.1$ s

Sufficiently long-lived non-relativistic NPPs can contribute a major part to the energy density, thus changing the Universe from radiation dominated to matter-dominated during the BBN.

Ultrarelativistic decay products of NPPs can contribute to the energy density, thus increasing N_{eff}, neutron decoupling temperature and the neutron-to-proton ratio.

Weakly interacting particles (leptons) produced in decays of NPPs participate in *weak interactions* involving neutron and are able to decrease the neutron decoupling temperature, thus decreasing the neutron-to-proton ratio.

Heavy NPPs can decay into mesons, which can convert $p \to n$ in *strong interactions* before decaying themselves, e.g.

$$\pi(K) + p \to n + \pi(K)/\gamma \tag{3.4.1}$$

and therefore increase the neutron-to-proton ratio.

3.4.1.2 $\tau_X \gtrsim t_{\mathbf{BBN}}$

Muons and mesons from decays of heavy NPPs, $m_X \gtrsim m_\mu$, are able to dissociate primordial nuclei, e.g.

$$\pi + {}^4\text{He} \to D + 2n, \quad \pi + D \to 2n \tag{3.4.2}$$

directly changing the primordial abundances.

3.4.1.3 $\tau_X \gtrsim 10^4$ **s**

If NPPs survive down to small temperatures $T_{\text{photodiss}} \simeq$ few keV, the photons produced in its decay or in decays of its daughter particles are able to dissociate primordial nuclei.

$T_{\text{photodiss}}$ can be estimated from the requirement that the maximally possible energy of photons in SM plasma, $E_{\gamma,\text{max}} \approx m_e^2/22T$ [102],[2] becomes comparable with the binding energy of the nuclei. For example, for D and ^{4}He we have

$$E_{\gamma,\text{max}} \simeq \begin{cases} \Delta_D \approx 2.2 \text{ MeV} \to T \simeq 5 \text{ keV}, \\ \Delta_{^4\text{He}} \approx 20 \text{ MeV} \to T \simeq 0.6 \text{ keV}, \end{cases} \tag{3.4.3}$$

This translates for the requirement $\tau_X \gtrsim 5 \cdot 10^4$ s for D dissociation and $\tau_X \gtrsim 10^6$ s for ^{4}He dissociation.

> Different effects of new particles on BBN can be classified by the lifetime of a new particle. Namely, short-lived particles that decay before decoupling of weak interactions do not produce observable effects. Vice versa, particles produced in small numbers and very long-lived influence BBN negligibly. Therefore constraints from BBN usually have an upper and lower bound by particle's lifetime.

Below we provide several examples of modifications to the BBN by new physics.

Example 1: additional relativistic particles. Assume that there exists additional neutrino species. Then the number of effective degrees of freedom change as $N_{\text{eff}} \to N_{\text{eff}} + \Delta N_{\text{eff}}$. This effect increases a Hubble rate $H = T^2/M_{\text{Pl}}^* \propto g_*$, where g_* depends on

[2] The cut-off is determined by the process $\gamma + \gamma_{\text{SM}} \to e^+ + e^-$.

N_{eff} as

$$g_* = 2\cdot N_\gamma + \frac{7}{8}\cdot 2\cdot 2\cdot N_e + \frac{7}{8}\cdot 2\cdot \left(\frac{T_\nu}{T_\gamma}\right)^4 \cdot N_{\text{eff}} \approx \begin{cases} 10.83 + 1.75\Delta N_{\text{eff}}, & \text{before annihilation} \\ 3.38 + 0.45\Delta N_{\text{eff}}, & \text{after annihilation} \end{cases} \tag{3.4.4}$$

Here number of photons $N_\gamma = 1$, effective number of neutrinos in SM $N_{\text{eff}} = 3.046$ [103]. Before electron-positron annihilation ($T_\gamma \simeq m_e$) $N_e = 1$, $T_\nu = T_\gamma$. After electron-positron annihilation $N_e = 0$, $T_\nu = \left(\frac{4}{11}\right)^{1/3} T_\gamma$.

Let us estimate effect of faster Universe expansion on T_n. From Eq. (3.2.8) we have a condition

$$\langle \Gamma_{n\to p}\rangle(T_n) = H(T_n) \approx 1.66\sqrt{g_*}\frac{T_n^2}{M_{\text{Pl}}}. \tag{3.4.5}$$

Using approximate temperature dependence $\langle \Gamma_{n\to p}\rangle(T_n) \sim T_n^5$ we get

$$T_n \sim (g_*)^{1/6} \quad \Rightarrow \quad \frac{\Delta T_n}{T_n} \approx \frac{\Delta g_*}{6 g_*} \tag{3.4.6}$$

Using (3.4.6) and simple analytic estimates (3.2.14), (3.2.1) for $n_n/n_p(T_{\text{BBN}})$ and $Y_{^4\text{He}}$, we can constrain the value ΔN_{eff} from the Helium abundance. Indeed, a shift ΔT_n changes $n_n/n_p(T_{\text{BBN}})$ and, correspondingly, the $Y_{^4\text{He}}$. Corresponding corrections can be found expanding the expressions (3.2.14), (3.2.1) in series on $\Delta T_n/T_n$:

$$\Delta Y_{\text{He}} \approx \Delta\left(\frac{n_n}{n_p}\right)\frac{Y^2_{\text{He,SM}}}{2(n_n/n_p)^2_{\text{SM}}}, \quad \Delta\left(\frac{n_n}{n_p}\right) \approx e^{t_{\text{BBN}}/\tau_n + \Delta m/T_n}\left(\frac{n_n}{n_p}\right)^2_{SM}\cdot\frac{\Delta m \Delta T_n}{T_n^2} \tag{3.4.7}$$

Plugging in the numbers $T_n \approx 0.716$ MeV, $(n_n/n_p)_{\text{SM}} \approx 1/7$, $Y_{\text{He,SM}} \approx 0.227$, $t_{\text{BBN}} \approx 200$ s corresponding to analytic estimates, we obtain

$$\Delta(n_n/n_p) \approx 0.28\frac{\Delta T_n}{T_n}, \quad \frac{\Delta Y_{\text{He}}}{Y_{\text{He}}} \approx 1.54\frac{\Delta T_n}{T_n} = 0.024\Delta g_* = 0.041\Delta N_{\text{eff}} \tag{3.4.8}$$

Requiring $\Delta Y_{\text{He}}/Y_{\text{He}} < 0.05$, we get

$$\Delta g_* < 2.08, \quad \Delta N_{\text{eff}} \lesssim 0.82 \tag{3.4.9}$$

Assuming the relative error of Helium to be $\Delta Y_{He}/Y_{He} \sim 5\%$, the maximal allowed $\Delta N_{\text{eff}} = 0.82$

Example 2: additional energy density. Consider the NPP that was in thermal equilibrium, decoupled being relativistic but being non-relativistic during the BBN (i.e.,

$m_X \gg T_n$). We estimate the corresponding energy density as

$$\rho_X \approx m_X Y_X s \approx 2 \cdot 10^{-4} \frac{Y}{4 \cdot 10^{-4}} g_*(T) e^{-t/\tau_X} T^3 m_X \tag{3.4.10}$$

The ratio of ρ_X to the energy density of the SM plasma is

$$\rho_N/\rho_{\text{SM}} \approx 3.1 \cdot 10^{-4} \frac{Y_X}{4 \cdot 10^{-4}} e^{-t/\tau_X} m_X/T(t) \tag{3.4.11}$$

with the maximal value

$$\rho_X/\rho_{\text{SM}}\big|_{\text{max}} \simeq \rho_X/\rho_{\text{SM}}\big|_{t \simeq \tau_X} \approx 0.4 \frac{Y_X}{4 \cdot 10^{-4}} (m_X/1\text{ GeV}) \sqrt{\tau_X/1\text{ s}}, \tag{3.4.12}$$

where we assumed radiation dominated time-temperature relation. Having the parametric dependence of Y, τ_X on the mass of NPPs and its coupling to the SM, we can impose a constraint using the requirement

$$\rho_X/\rho_{\text{SM}} \ll 1 \tag{3.4.13}$$

3.5 Big Bang Nucleosynthesis in the presence of Heavy Neutral Leptons

We are interested in HNLs that can generate the masses of the SM neutrinos – in particular, the observed mass difference $\sqrt{\Delta m^2_{\text{atm}}} \approx 0.05$ eV. No HNLs heavier than $m_N \gtrsim 1$ GeV with lifetimes $\tau_N \gtrsim 0.1$ s are able to provide such difference.

Indeed, to provide the observed mass difference, the mixing angle must be larger than the see-saw bound [49]:

$$\theta^2 \gtrsim \theta^2_{\text{see-saw}} \simeq \sqrt{\Delta m^2_{\text{atm}}}/m_N \approx 5 \cdot 10^{-11} \left(\frac{1\text{ GeV}}{m_N} \right) \tag{3.5.1}$$

The comparison of the parameter space of HNLs with the lifetimes $\tau_N = 0.1$ s, $\tau_N = t_{\text{BBN}} \simeq 200$ s with the bound (3.5.1) is shown in Fig. 3.6. We see that the see-saw bound intersects the line $\tau_N = 0.1$ s at $m_N \simeq 1$ GeV. Based on the figure, we conclude that for heavy enough HNLs, $m_N \gtrsim m_\pi$, the only possible effects above the see-saw bound are:

- change of the Hubble expansion rate
- decrease of neutron decoupling temperature due to injection of weakly interacting particles
- increase of n-p ratio due to injection of muons or mesons

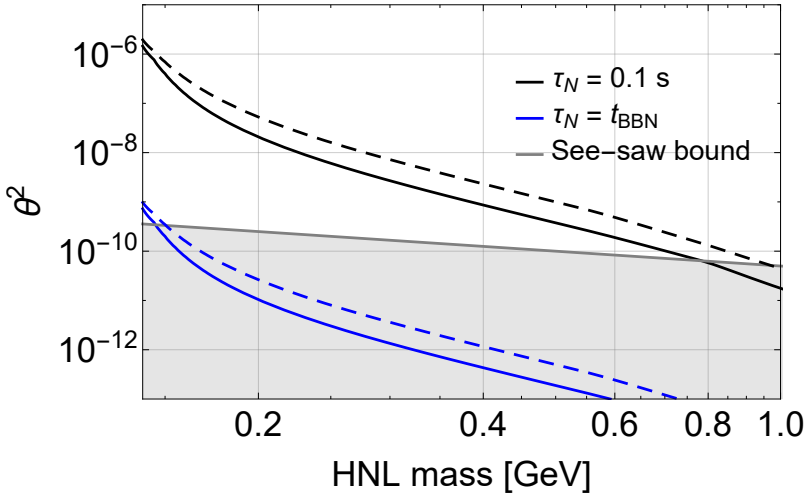


Figure 3.6: Parameter space of HNLs with lifetimes $\tau_N = 0.1$ s (black line) and $\tau_N = t_{\mathrm{BBN}}$ (blue line); solid lines corresponds to the mixing with ν_e, while dashed lines – to the mixing with ν_τ. The shaded gray region corresponds to the parameter space excluded if we consider HNLs that are able to provide observable mass difference of SM neutrinos $\sqrt{\Delta m^2_{\mathrm{atm}}} \approx 0.05$ eV.

HNLs lighter than pion mass can also affect the BBN at the later stages.

To account for the influence of HNLs we make the following changes in the equations of BBN:

- HNLs introduce additional Boltzmann equations (on f_{HNL} and their decay products – muons, pions and heavier mesons) and corresponding matrix elements
- Processes with HNLs or their decay products should be added to other Boltzmann equations
- HNLs and their decay products have pressure and energy density that modify the law of expansion

At this point, we do not consider the effect of the increase of n-p ratio due to the injection of muons and mesons.

3.6 Description of numerical methods and code

The physical system of BBN consists of the particles, some of which are in thermodynamical equilibrium with each other and some of which are not. All properties of the equilibrium plasma are given by the evolution of temperature. Non-equilibrium particles (neutrinos,

nucleons, and nuclei) evolve with the expansion of the Universe (according to Friedmann equations) and particle collisions (Boltzmann equations).

Hence the complete set of unknowns in our system of equations is:

- plasma temperature $T(t)$
- distribution functions of neutrinos $f_{\nu_i}(t, p)$
- densities of nucleons and nuclei $n_X(t)$
- distribution function of sterile neutrino and its decay products $f_N(t, p)$, $f_{\text{product,i}}(t, p)$

In the cosmological setup it is convenient to use the scale factor $a(t)$ as a parametrization of time.

Equilibrium density is defined by the temperature, while non-equilibrium density is given by distribution functions. Their evolution is given by Friedmann and Boltzmann's equations. Below we will discuss our method of their numerical solution.

Temperature evolution equation. The energy conservation equation can be used to derive the temperature evolution of the plasma. We can split the total energy density into equilibrium and non-equilibrium parts $\rho = \rho_{\text{eq}} + \rho_{\text{noneq}}$. Then the derivative of ρ_{eq} is given by $\dot{T}$ while the time evolution of non-equilibrium density depends on the Boltzmann collision integrals:

$$\begin{aligned} \dot{\rho} &= \left(\frac{\mathrm{d}\rho_{\text{eq}}}{\mathrm{d}T} \frac{\mathrm{d}T}{\mathrm{d}t} + \frac{\mathrm{d}\rho_{\text{noneq}}}{\mathrm{d}t} \right) = -3H\,(\rho + P) \\ \frac{\mathrm{d}T}{\mathrm{d}t} &= -\frac{3H(\rho + P) + \frac{\mathrm{d}\rho_{\text{noneq}}}{\mathrm{d}t}}{\frac{\mathrm{d}\rho_{\text{eq}}}{\mathrm{d}T}} \end{aligned} \tag{3.6.1}$$

where P is the total pressure and H is the Hubble rate.

Expressions for ρ_{eq} and ρ_{noneq} can be substituted in this equation to obtain an explicit formula for the temperature evolution. This is done in detail in Appendix 3.A.

Comoving coordinates Because of the expansion of the Universe, even non-interacting particles evolve non-trivially. This significantly complicates the computations and analysis.

According to Friedmann equations, particle positions and momenta behave as $\vec{r} \propto a(t)$ and $\vec{p} \propto a^{-1}(t)$. From this it is easy to conclude that particle densities $n \propto a^{-3}(t)$. Combining the latter two facts, we also find that the energy density of relativistic particles necessarily $\rho_{rel} \propto a^{-4}(t)$ while for non-relativistic ones $\rho_{nonrel} \propto m_{nonrel} a^{-3}(t)$.

It is convenient to switch to the *comoving* frame of reference that is expanding along with the Universe:

$$\vec{r} \to \vec{r}/a \tag{3.6.2}$$

$$\vec{p} \to \vec{y} = \vec{p} \cdot a \tag{3.6.3}$$

In this frame, many *comoving* quantities (denoted by ˜) become constant with respect to the expansion (e.g. photon energy density $\tilde{\rho}_\gamma \equiv \rho_\gamma a^4 = \text{const}$).

A huge additional advantage is a fact that in comoving coordinates the Boltzmann equations lose the term with the Hubble rate:

$$\frac{df(t,p(t))}{dt} = \frac{\partial f}{\partial t} - Hp\frac{\partial f}{\partial p} = aH\frac{\partial f(a,y)}{\partial a} = I_{coll} \tag{3.6.4}$$

$$\frac{\partial f(a,y)}{\partial a} = \frac{I_{coll}}{aH} \tag{3.6.5}$$

This change of variables also suggests rewriting the temperature evolution equation for the quantity $\tilde{T} \equiv a \cdot T$. For relativistic particles, temperature evolves with expansion as $T \propto a^{-1}$. Then $\dot{\tilde{T}}$ describes the change of entropy of the system due to transitions in the plasma (like electron-positron annihilation) or influence of non-equilibrium particles. Finally, we introduce the comoving mass $\tilde{m} = m \cdot a$ for particles.

BBN physics formulated in terms of comoving positions, momenta and masses look in a sense very similar to physics in non-expanding Universe, but with interaction rates decreasing with time.

> The evolution of physical quantities even in the simplest cases is non-trivial because of the expansion of the Universe. However, if we consider the *comoving* frame (expanding with the Universe), we can significantly simplify the description for numerics.

3.6.1 pyBBN: code for non-standard nucleosynthesis

Our goal is to put constraints on the Heavy Neutral Leptons from the primordial nucleosynthesis. This task has been to some extent done by [104–106]. Papers [107, 108] also consider relevant aspects of the influence of new physics on the BBN (like entropy injection). This is, however, not sufficient to describe all effects of HNLs (e.g., HNL decays produce non-equilibrium neutrinos that can directly influence the weak reactions of nucleons).

Code by [106] is based on [105] and implements the physics of HNLs with masses up to $M < m_\pi$. We extend this analysis to larger masses $M < m_\phi$ in a new code ***pyBBN***[3] [3]. Our main contribution is the treatment of hadronic decay channels of HNLs and incremental improvements and fixes to the approach of [106].

[3] https://github.com/ckald/pyBBN

General structure. Simulations are done in two stages:

1. The cosmological background physics and the rates of the reactions

$$n + e^+ \leftrightarrow p + \bar{\nu}_e, \quad n + \nu \leftrightarrow p + e, \quad n \to e + \bar{\nu}_e + p \tag{3.6.6}$$

   are computed in *pyBBN*. This involves solving the system of equations for the evolution of temperature, scale factor and distribution functions of decoupled species like neutrinos, HNLs and relevant decays products (see Sec. 3.2.2).

2. The cosmological quantities together with the aforementioned rates are tabulated and passed to an external code, the modified KAWANO code (also known as NUC123 [83]), that takes care of the nuclear part of the simulation and outputs the light element abundances. This is done to avoid reimplementing the complicated system of nuclear physics with experimental fits to the cross-sections. KAWANO is a relatively outdated code that does not implement a number of corrections introduced in PArthENoPE [87] or PRIMAT [79], but these corrections are small in SBBN and KAWANO is the easiest to modify for custom cosmological evolution and neutrino spectral distortions.

Time evolution. The system of equations step-by-step for values of scale factor evenly spaced in logarithmic scale (as opposed to linearly-spaced values). This is done to naturally increase the timestep at late times when interaction rates are low. The numerical schemes used for time-evolution equations are the *Adams-Bashforth explicit linear multistep method* for temperature evolution equation and *Adams-Moulton implicit linear multistep method* for Boltzmann equations. We decided to use these particular schemes (as opposed to the Runge-Kutta family) since they allow us to reuse the previously computed values (we use up to 5 previous steps) and proved to be stable enough for our purposes.

The implicit methods are known to help mitigate the problem of *stiffness* in Boltzmann equations when it is possible to apply them. However, the temperature evolution equation in our system can be solved only explicitly.

Boltzmann integrals for distribution functions are in general of high dimension. Typically this would suggest the usage of Monte Carlo methods. But in the particular case of the Fermi-like theory of HNLs, it is possible to reduce the number of integrations to 1 or 2 (this is described in detail in App. 3.C). The accuracy of this computation is critical to the evolution of the neutrinos which constitute a large fraction of the total energy density as well as directly influence the neutron-to-proton ratio. Hence we elected to use the adaptive integration algorithm QAG from GNU Scientific Library [109]. Integration is done using the Gaussian quadrature and the error is estimated by dividing the integration region into subintervals and repeating the procedure until the required global tolerance is met. Note that the integrand of the problem is piecewise smooth while the Gaussian quadrature is designed for the polynomials. Because of this the Gaussian quadrature might require arbitrarily large number of points to converge. However, we find that the integration using

the Gaussian quadrature of the order 60 provides the same accuracy as the adaptive method for four-particle interactions. For three-particle interactions we employ adaptive integration with Gauss-Kronrod quadrature of the order 31.

Unstable HNL decay products. Sterile neutrinos of high masses decay into short-lived mesons – e.g. π_0 lifetime is $\sim 10^{-17}$ sec and $\pi_\pm$ lifetime is $\sim 10^{-8}$ sec. These timescales are orders of magnitude smaller than any practical computational timestep for a simulation spanning 10^4 sec.

To the first approximation, short-lived particles can be immediately destroyed without populating their distributions. However, some of them are charged and thermalize with the plasma and transfer some energy.

For illustration, we will consider the case of a muon. HNLs with masses $M_N < 105$ MeV will decay into stable particles. HNLs with higher masses will have decay products that are unstable. Some of these unstable decay products will interact with the plasma before they decay. The analysis here will be done for muons but can be applied to all other particles as well.

There are three important events to consider:

1. $\mu^\pm$ is created from HNL decay
   The distribution function of these muons is a non-thermal distribution of $f_{\rm noneq}$.

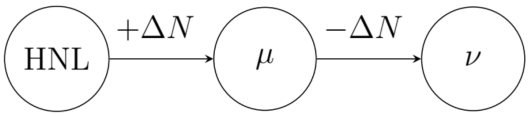
2. $\mu^\pm$ thermalizes
   The muon-photon scattering rate is higher than the muon decay rate: $\Gamma_{\gamma\mu} \sim \frac{\alpha^2}{m_\mu E_\gamma} T_\gamma^3 \sim 10^{-9}\,\mathrm{MeV}$ vs. $\Gamma_{\mu,\mathrm{decay}} \sim 10^{-16}\,\mathrm{MeV}$. This means that the muons will release their energy into the plasma and equilibrate before they decay. This process increases the temperature of the plasma and makes the muons non-relativistic. After thermalization, the muons will share the same temperature as the plasma and will have a thermal distribution $f_{\rm thermal} = e^{-\frac{m_\mu - \mu}{T}} e^{-\frac{p^2}{2m_\mu T}}$, where μ is determined by the condition that the number density before and after thermalization must be equal. The collision term corresponding to this process is then estimated as

   $$I_{\rm thermalization} \approx \frac{f_{\rm thermal} - f_{\rm noneq}}{\Delta t}\ ,$$

   with Δt the timestep of the simulation. The same procedure is followed for charged pions and charged kaons.

3. $\mu^\pm$ decays
   The main decay channel of muons is $\mu^- \to e^- + \overline{\nu_e} + \nu_\mu$. The muon has a lifetime, $\tau_\mu \sim 10^{-6}$ sec., that is much smaller than the timestep of the simulation. This poses a problem right away: when the evolution of the distribution function for the muon and active neutrinos is computed as $\Delta f = I_{\rm coll}\Delta t$, the behavior of $I_{\rm coll}$ is not resolved. It is assumed to be constant during the whole timestep Δt, which is not true; the created

muons have already decayed well within this timestep. What therefore happens is that the number of muons that have decayed and the number of neutrinos that are created are overestimated. This issue can be solved by using a dynamical equilibrium. Consider the chain The timestep Δt is much smaller than the lifetime of the HNL,



which means that there is approximately a constant inflow of muons *during* each timestep. Since the number of muons created ΔN decays almost instantaneously, the same number of active neutrinos is created: for each muon that decays, one electron neutrino and one muon neutrino is created. Now a scaling α can be introduced in Δf = $I_{\text{coll}} \Delta t \alpha$ such that $\int \mathrm{d}^3 p \Delta f/(2\pi)^3 = \Delta N$.

3.6.2 Code testing

pyBBN had been tested in multiple situations modeling both Standard BBN theory as well as in the presence of HNLs. Below we will list only the most representative selection of them.

The main areas of interest are thermodynamical properties (i.e., evolution of temperature) and computation of Boltzmann equations (collision integrals and integration of the equations themselves).

According to scope, tests can be divided into *low-level* (i.e., checking that formulas or numerical schemes are correct) and *integration tests* (i.e., those defining a toy model for which we can produce analytical results or can make a comparison with other codes).

3.6.2.1 Temperature of the Cosmic Neutrino Background

Decoupling neutrinos to the first approximation preserve their thermal distribution function with the same temperature as photons.

However, at the temperatures about $T \sim m_e = 0.511 keV$, electrons and positrons annihilate into additional photons. Since neutrinos decouple earlier, do not get heated up as well. This results in a difference in the temperature of Cosmic Microwave and Neutrino Backgrounds.

This effect is easily quantified through the comoving entropy conservation law ($sa^3 = const$):

$$\frac{T_\gamma}{T_\nu} = \frac{aT_\gamma}{aT_\nu} = \left(\frac{g_*(T_{\text{before}})}{g_*(T_{\text{after}})}\right)^{\frac{1}{3}} = \left(\frac{\frac{7}{8}\cdot 4 + 2}{2}\right)^{\frac{1}{3}} \approx 1.401 \tag{3.6.7}$$

This law is not explicitly enforced in the code, but comoving entropy is an integral of motion. Check of the Cosmic Neutrino Background temperature allows to test the thermodynamics in the code as a whole: Fig.3.7. The neutrinos are assumed to decouple instantly with any non-equilibrium dynamics beside free propagation.

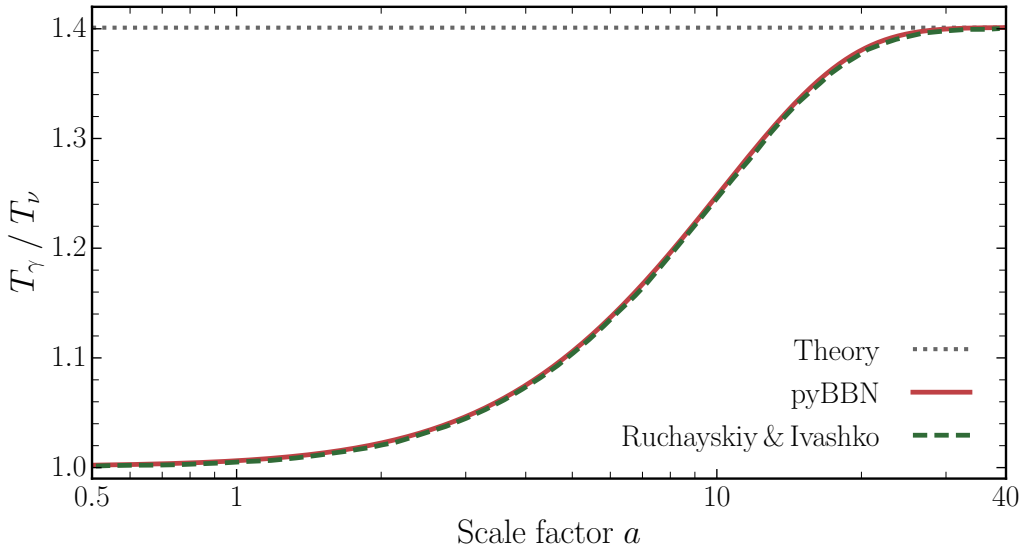


Figure 3.7: The photon temperature divided by active neutrino temperature. The increase here is due to electron-positron annihilation into photons. Dashed curve is from [106].

In the Universe filled with electrons, positron, photons and decoupled neutrino – photons and neutrinos have different temperatures after electrons-positron annihilation. The ratio of their temperatures can be predicted from the conservation of comoving entropy. Fig. 3.7 shows that the thermodynamic evolution in the code predicts this ratio with accuracy $\lesssim 0.1\%$.

3.6.2.2 Increased Hubble rate

pyBBN predicts Helium-4 abundance for the SM BBN (SBBN) to be equal to 24.8%. This value agrees with a simple analytic estimate (3.2.1) within 10%. In this test we will check that increase of a Hubble rate during BBN result in the correct qualitative behavior of $Y_{^4\mathrm{He}}$.

An increase in the Hubble parameter leads to an earlier decoupling of neutron-proton weak reactions and therefore a higher neutron-to-proton ratio. This is shown in the scenario where the Hubble rate in SBBN is increased artificially by a factor of 2, see Fig. 3.8. We check the prediction of the code using analytic estimate (3.2.1), (3.2.14) of the ^{4}He abundance. The neutron decoupling temperature obtained using $\Gamma_{n\to p}(T) = H(T)$ gives $T_n \simeq 0.92$ MeV, which translates into $n_n/n_p(T_{\mathrm{BBN}}) \approx 0.195$ and $Y_{^4\mathrm{He}} \approx 0.37$, in good agreement with the result shown in Fig. 3.8.

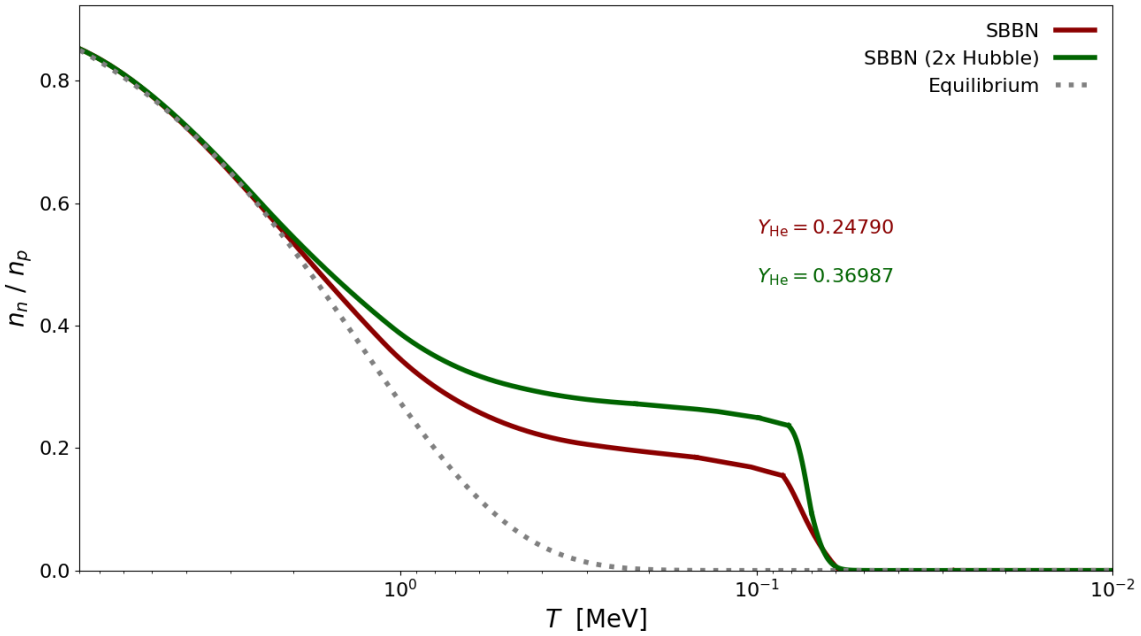


Figure 3.8: Evolution of the neutron-to-proton ratio in SBBN (red line) and SBBN + twice as large Hubble parameter (green line). The dotted line indicates the equilibrium evolution.

The prediction of the code for the value of $Y_{^4\mathrm{He}}$ in SM BBN agrees with analytic estimate. The result of the simulation with increase Hubble rate also agrees with our expectations.

3.6.2.3 Neutrino spectral distortions

In Fermi theory the cross section increases with momentum as $\sigma \propto G_F^2 p^2$, which means that neutrinos with higher momenta stay longer in equilibrium. Since these neutrinos decouple later, they will in fact briefly experience the heat-up of the plasma due to electron-positron annihilation, shown in Figure 3.7, and the corresponding increase in aT. On the other hand, oscillations of neutrinos are mixing the distributions of different flavours, bringing the spectra closer to each other.

At temperatures of $\mathcal{O}(1)$ MeV electron neutrinos interact through both charged and neutral currents, while muon and tau neutrinos only interact through the neutral current. The temperature is too low for muons and tau leptons to be present in the plasma or to be created from muon and tau neutrinos through the charged current channel. The cross section of electron neutrinos is therefore larger and they get larger spectral corrections. The results are shown in Fig. 3.9.

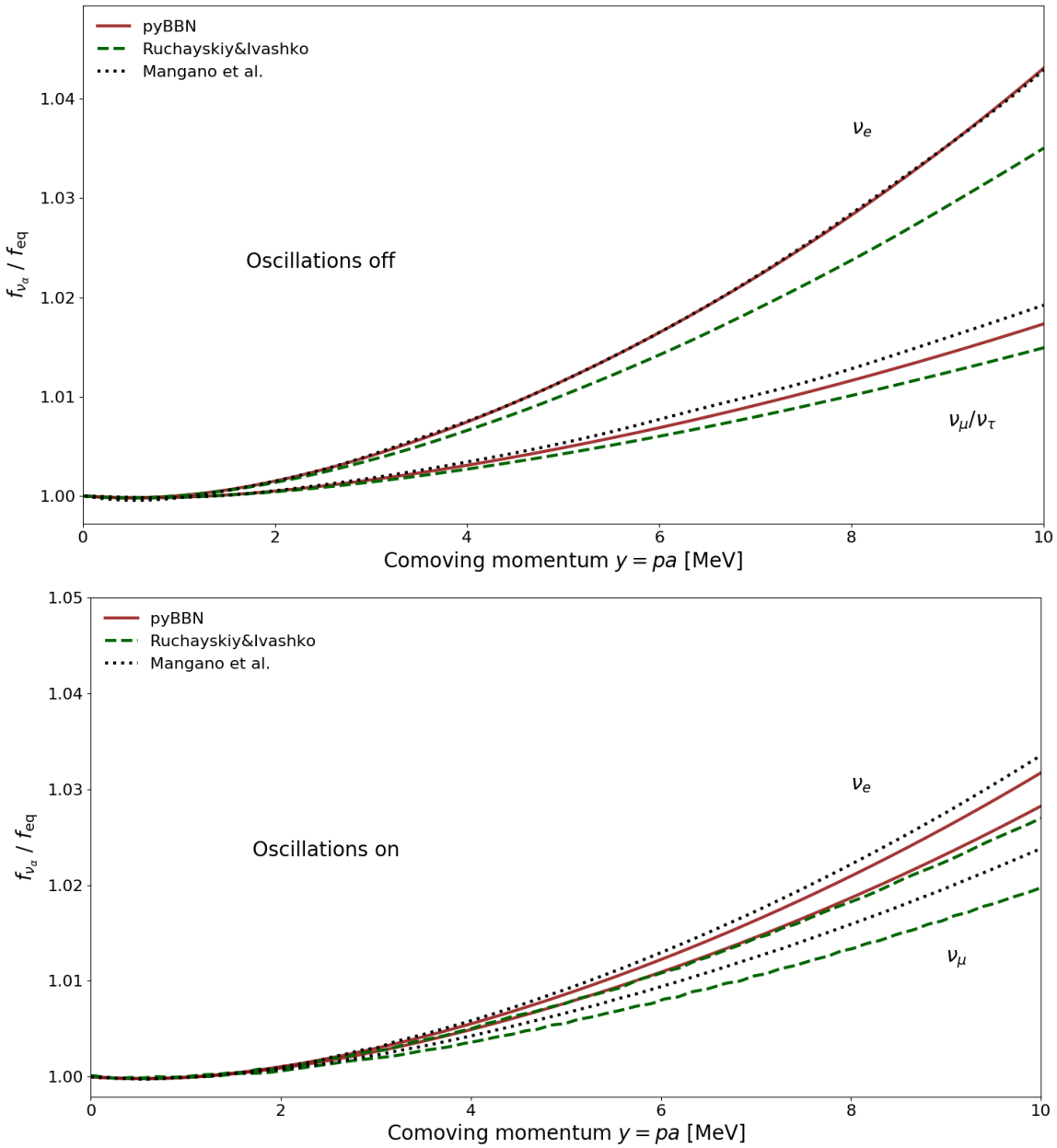


Figure 3.9: Ratio of active neutrino decoupled spectra to their equilibrium distribution before the onset of BBN. The upper curves show the distortion of the electron neutrino spectrum and the lower of muon and tau neutrinos. Dashed curves are from [106], dotted from [103].
Upper panel: no neutrino flavour oscillations. *Lower panel:* neutrino flavour oscillations with $\sin^2\theta_{12} = 0.3,\ \sin^2\theta_{13} = 0.5$ and $\theta_{23} = 0$ (rotated $\nu_\mu - \nu_\tau$ basis).

In this test we considered the spectral distortions of neutrinos during decoupling of

weak reactions. We reproduce the results of codes by [106] and [103] up to $\lesssim 1\%$ accuracy even for very sparsely populated momentum values of neutrinos ($f(p,T) \sim \exp(-p/T) \sim 10^{-3}$).

3.6.2.4 Heavy Neutral Leptons decay width

To check the computation of collision integrals and integration of Boltzmann equations, we restore the decay width of HNLs for the computed evolution of their number density using that

$$\frac{n_i - n_{i-1}}{\Delta t} + 3Hn_i \approx \Gamma_N n_i \tag{3.6.8}$$

We consider an HNL of mass $M_N < M_\mu$ that mixes only with electron neutrino through mixing angle $|\theta_e|$. There are four decay channels:

$$\begin{aligned} &N \to \nu_e + \nu_e + \overline{\nu_e} && N \to \nu_e + \nu_\mu + \overline{\nu_\mu} \\ &N \to \nu_e + \nu_\tau + \overline{\nu_\tau} && N \to \nu_e + e^+ + e^- \,, \end{aligned} \tag{3.6.9}$$

The theoretical vacuum decay width is

$$\Gamma_N = \frac{G_F^2 \left|\theta_e\right|^2 M_N^5 \left(\frac{1}{4}\left(1 + 4\sin^2\theta_W + 8\sin^4\theta_W\right) + 1\right)}{192\pi^3}. \tag{3.6.10}$$

We check that our code correctly reproduce this value. The result is shown in Fig. 3.10.

We have confirmed that decay width of HNL computed using collision integrals in our code recovers the theoretical decay width with accuracy $\sim 0.2\%$

3.6.2.5 Reheating due to neutral pion decay

As a check of energy conservation in 3-body interactions, we check the energy injected by decays of $M_N = 135$ MeV sterile neutrino.

$$N \to \pi^0 + \nu_e \tag{3.6.11}$$

$$\pi^0 \to \gamma + \gamma \tag{3.6.12}$$

For each HNL that decays, one neutral pion is created. If the HNL has a mass very close to that of the neutral pion, then the neutral pion created will be very non-relativistic. The energy injected in the plasma due to neutral pion decay during each step is therefore

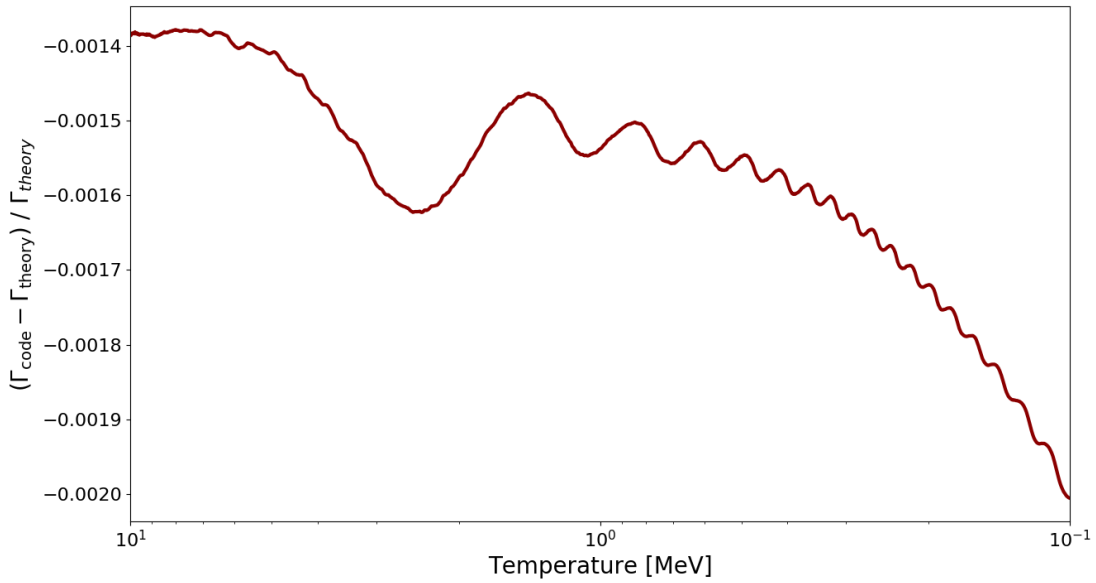


Figure 3.10: Relative difference between averaged HNL decay width calculated in pyBBN and its theoretical value (Eq. 3.6.10) of reactions Eq. 3.6.9. Here an HNL of mass $M_N = 30$ MeV and mixing angle $|\theta_e|^2 = 10^{-4}$ is chosen. The initial temperature of the simulation is $T_{\text{ini}} = 10$ MeV.

approximately

$$\Delta E \approx 2 m_{\pi^0} \Delta N \ , \tag{3.6.13}$$

with ΔN the number of HNLs that have decayed during the step. The factor of 2 comes from the fact that the charge conjugated channel also creates a neutral pion. Comoving photon energy density becomes

$$g_\gamma \frac{\pi^2}{30} (aT)^4 = \rho_{\gamma,\text{c}} = \rho_{\gamma,\text{c,old}} + a^4 \Delta E = \rho_{\gamma,\text{old}} + 2 m_{\pi^0} a^4 \Delta N \tag{3.6.14}$$

and

$$(aT)_{\text{new}} = \left(\frac{30}{g_\gamma \pi^2} \rho_{\gamma,\text{c}} \right)^{1/4} . \tag{3.6.15}$$

The result of this test is shown in Fig. 3.11.

Sterile neutrinos with masses $m_N > m_\pi$ have qualitatively new decay channels into 2 particles. This involves a different expression of the collision integral that haven't been tested before. We have confirmed that 3-body interactions conserve particle number, energy and comoving entropy up to $\lesssim 0.1\%$.

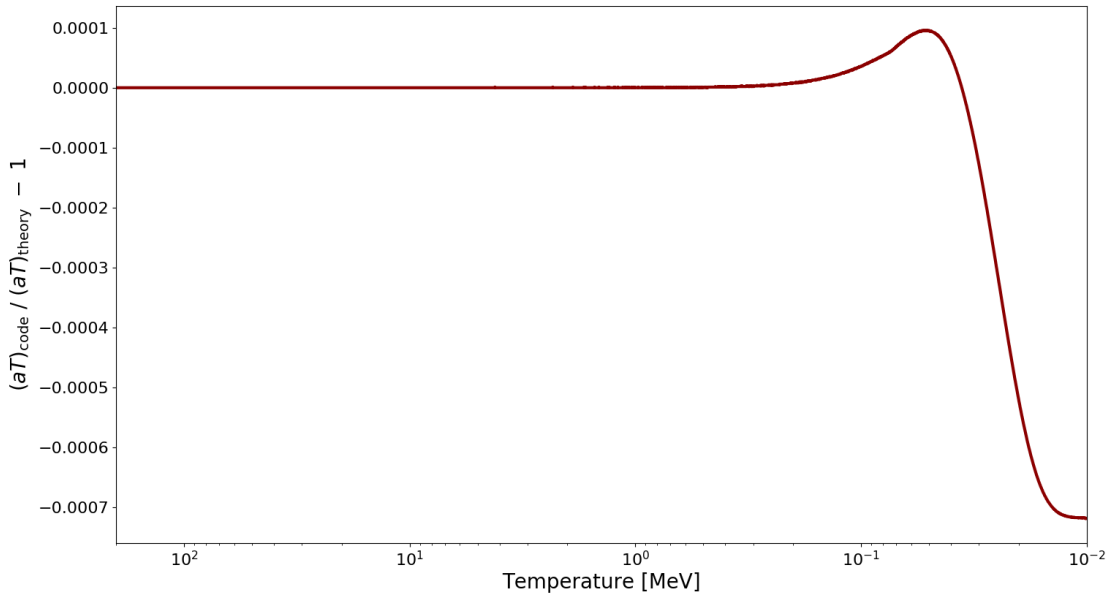


Figure 3.11: Reheating of plasma due to decays 3.6.11 and 3.6.12. Theoretical value aT_{theory} is given by Eq. 3.6.15. Here an HNL of mass $M_N = 135$ MeV with $|\theta_e|^2 = 10^{-4}$ is taken.

3.6.2.6 Chain of decays with short-lived particles

Sterile neutrinos of high masses can decay into heavy mesons that themselves trigger the decay chains. Since mesons have very short lifetimes compared to the timescales of the simulation, their decays are treated by a different numerical scheme than neutrinos or nucleons. Eventually, all mesons decay into photons, leptons and neutrinos. For a given mass of HNL we can compute the total number of decay products and compare it with the code.

Consider the chain of decays

$$N \to \nu_e + \phi \tag{3.6.16}$$
$$\phi \to \pi^0 + \rho^0 \tag{3.6.17}$$
$$\pi^0 \to \gamma + \gamma \; ; \; \rho^0 \to \pi^+ + \pi^- \tag{3.6.18}$$
$$\pi^- \to \mu^- + \overline{\nu_\mu} \tag{3.6.19}$$
$$\mu^- \to e^- + \overline{\nu_e} + \nu_\mu \tag{3.6.20}$$

The total comoving number density of electron and muon neutrinos created is

$$n_{\nu_e,\text{c}} = (1 + 2\text{BR}_{\phi\to\pi^0\rho^0})n_{\text{ini}} \tag{3.6.21}$$
$$n_{\nu_\mu,\text{c}} = 4\text{BR}_{\phi\to\pi^0\rho^0} n_{\text{ini}} \tag{3.6.22}$$

with $\mathrm{BR}_{\phi\to\pi^0\rho^0}$ the branching ratio of the corresponding decay and n_{ini} the initial comoving number density of HNLs. The factor of 2 comes from the fact that the charge conjugated channel $\overline{N} \to \overline{\nu_e} + \phi$ also creates a phi meson. Results are shown in Fig. 3.12.

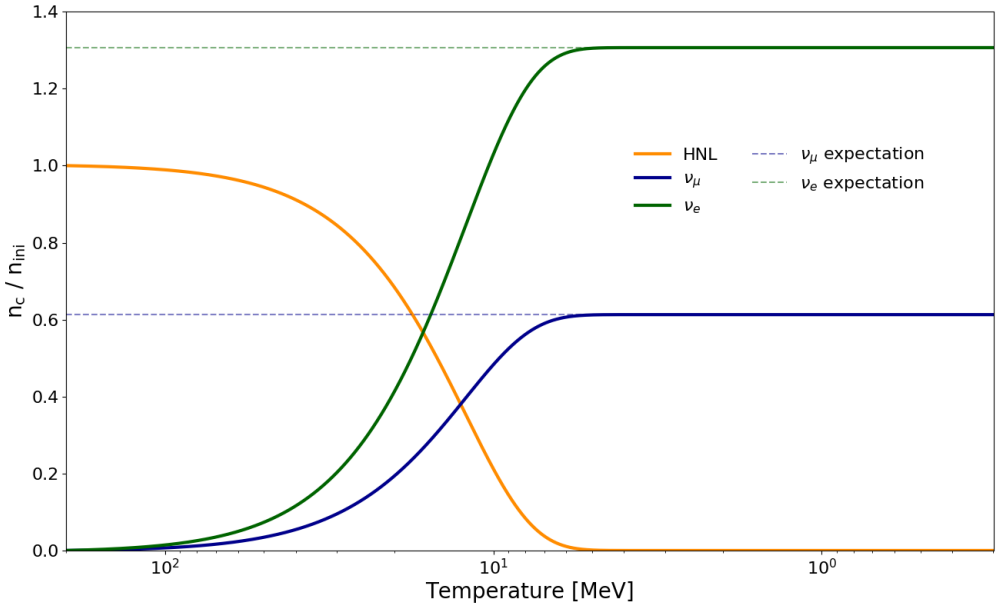


Figure 3.12: Comoving number density n_{c} of HNL, electron neutrino and muon neutrino normalized by the initial HNL comoving number density n_{ini} when the decays 3.6.16 - 3.6.20 are considered. The mass of the HNL here is $M_N = 1200$ MeV and $|\theta_e|^2 = 10^{-8}$. The initial densities of all particles except for HNLs, electrons and photons are taken 0 for convenience. The dynamical equilibrium mechanism discussed in Subsection 3.6.1 is used here.

Even for particles with lifetimes much smaller than the timestep of the simulation $\tau \ll \Delta t$, we correctly reproduce the expected concentrations of the decay products.

3.6.2.7 Helium-4 and Deuterium abundance in Standard Model Nucleosynthesis

We present the predictions for Helium and Deuterium in Standard BBN as computed by *pyBBN* in Fig. 3.13. Our result is consistent with state-of-the art precision BBN code PArthENoPE [87]. In both cases we used the neutron lifetime $\tau_n = 880.2$ sec from [1].

3.6.2.8 Non-equilibrium neutrinos vs thermal neutrinos

The current generation of high-precision BBN codes (e.g. PArthENoPE [87] and PRIMAT [79]) include many corrections not included in KAWANO. However, they

Code	$Y_{^4\mathrm{He}}$	D/H
PArthENoPE2.0 [87]	0.24676	$0.26800 \cdot 10^{-4}$
pyBBN	0.24678	$0.26776 \cdot 10^{-4}$
Relative difference	-0.008%	0.09%

Figure 3.13: Comparison of *pyBBN* with PArthENoPE

were designed for the SBBN and treat neutrinos as thermal-like particles, ignoring possible spectral distortions. We show that in presence of HNLs the spectral distortions can have the leading effect on the neutron-to-proton ratio.

In modern BBN codes (e.g. PArthENoPE [87] and PRIMAT [79]), neutrinos are assumed to have a thermal-like Fermi-Dirac distribution at all times. The neutrino temperature T_ν is determined such that

$$g_\nu \int \mathrm{d}^3 p_\nu \frac{E_\nu}{e^{\frac{E_\nu}{T_\nu}} + 1} = \rho_\nu \tag{3.6.23}$$

with ρ_ν the energy density.

The energy density is computed by the evolution equation

$$\frac{\partial \rho_\nu}{\partial t} + 4H\rho_\nu = C \tag{3.6.24}$$

with C the source function (integrated collision integral). In this way they account for the gravitational effect of incomplete neutrino decoupling.

Spectral distortions are ignored. In order to track spectral distortions, it is necessary to use the full machinery of the Boltzmann equation. On the other hand, the authors of [110] have compared their results with the authors of [111] (where spectral distortions are taken into account) and found a correction of only order 10^{-3}.

Since we expect production of neutrinos after decoupling, we expect a larger influence on the neutron-to-proton rates. Therefore, we have performed paired simulations: one using the non-equilibrium neutrino distribution in the calculation of the neutron-to-proton rates, while the other uses a thermal distribution with temperature that gives the same energy density as the non-equilibrium distribution. Note: this is only done at the level of the computation of the neutron-proton weak rates, in order to understand whether spectral distortions can be neglected. The evolution of the neutron-to-proton ratio is shown in Figure 3.14 and the final $Y_{^4\mathrm{He}}$ abundances are displayed in the table below.

	Non-equilibrium	Thermal	Rel. diff.
SBBN	0.24790	0.24743	0.19%
HNL ($M_N = 30\,\mathrm{MeV}, \tau = 0.075\,\mathrm{s}$)	0.24824	0.24714	0.44%
HNL ($M_N = 105\,\mathrm{MeV}, \tau = 0.11\,\mathrm{s}$)	0.29266	0.24405	16.6%

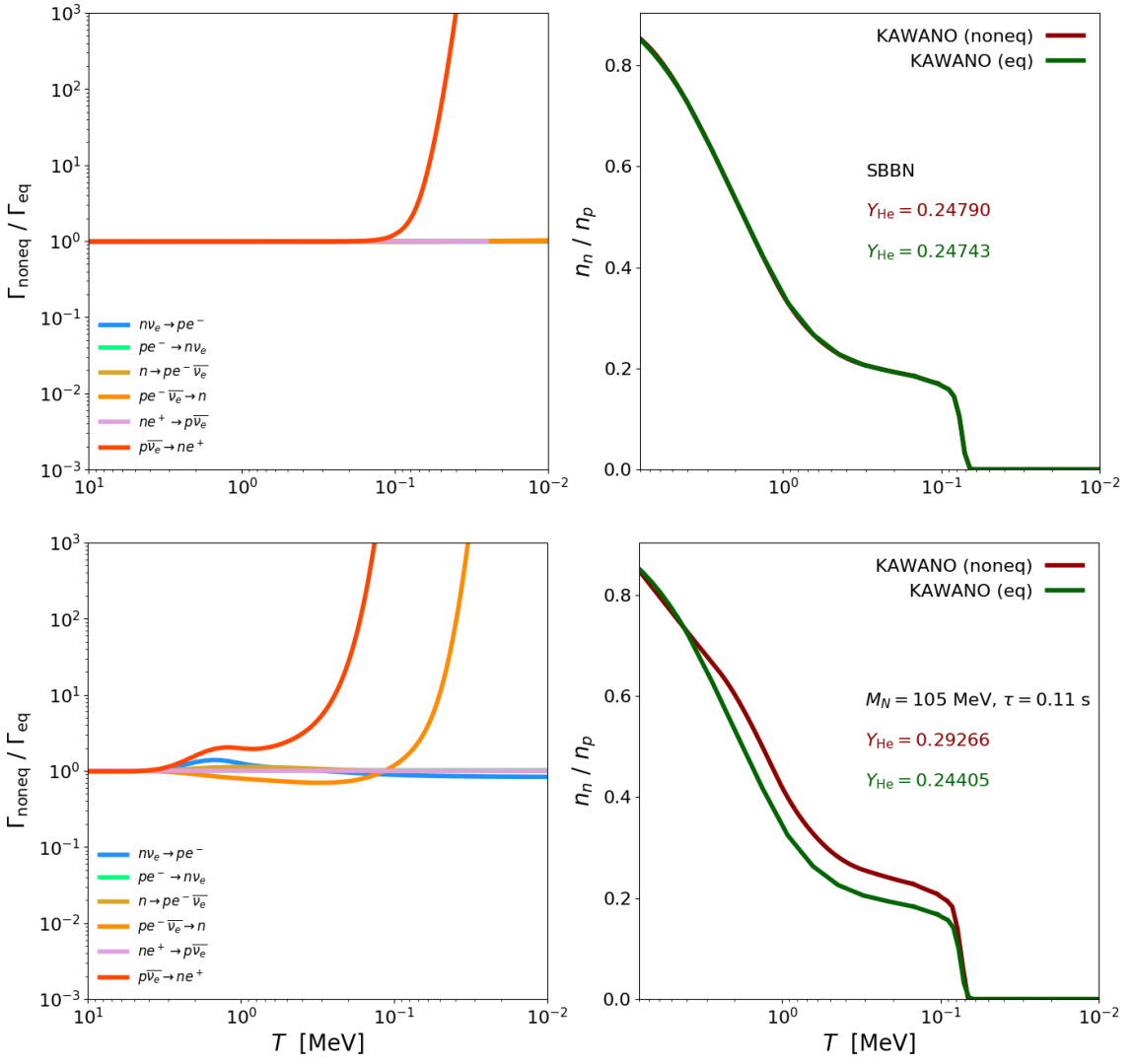


Figure 3.14: Comparison of nucleon reaction rates and neutron-to-proton ratio in SBBN (upper row) and in a model including HNL of mass $105\,\mathrm{MeV}$ and lifetime $0.11\,\mathrm{s}$ (lower row). *"noneq"* quantities are computed using the electron neutrino distribution function as obtained from the Boltzmann equation , while *"eq"* quantities are given under assumption of a thermal-like electron neutrino distribution (with the same energy density).

Fig. 3.14 indicates that spectral distortions can be the dominant effect that change the neutron-to-proton ratio in presence of sterile neutrinos.

3.6.2.9 Summary on tests

We have performed a large number tests, verifying the correctness of the equations, numerical schemes and conservation of basic integrals of motion. We payed special attention to the most computationally intensive and difficult part of the code – the evaluation of Boltzmann integrals. We checked

- integration of the expansion laws
- behaviour of the thermodynamical quantities under expansion
- energy transfer between non-equilibrium particles and plasma
- non-equilibrium corrections to particle species

We have implemented most of the tests previously done in literature. However, since physics of sterile neutrinos with masses $m_N > m_\pi$ have not been implemented elsewhere, we have created a number of sanity checks which we believe to give us reasonable control.

3.7 Results and conclusions

We have extended the constraints on sterile neutrinos from pion mass up to phi-meson mass. To achieve this, we have developed a numerical code *pyBBN* that we use to predict abundances of light chemical elements.

We implemented hadronic decay modes of HNLs, subsequent hadronic decays and their influence on the BBN. We validated the code comparing it with other codes and analytic estimates, and making consistency checks.

The comparison between bounds on the HNL parameter space provided by our code and codes from [105, 106] for HNLs in the mass range $m_N < m_\mu \approx 105$ MeV is shown in Fig. 3.15. Our results are in good agreement with previous studies in the mass range $m_N < 105$ MeV considered there. In the Figures 3.16 and 3.17 we show the predictions of our code for HNLs with masses up to $m_N \simeq 1$ GeV, comparing it with the simple constraint $\tau_N \gtrsim 0.1$ s from [49].

So far we have not implemented the effect of the interaction of massive decay products of HNLs – mesons and muons – with nucleons. Even if the lifetimes of these particles are much shorter than characteristic times of BBN, their production rate can be large enough to give a significant effect on n-p ratio [81]. To take into account this effect, we need to add in

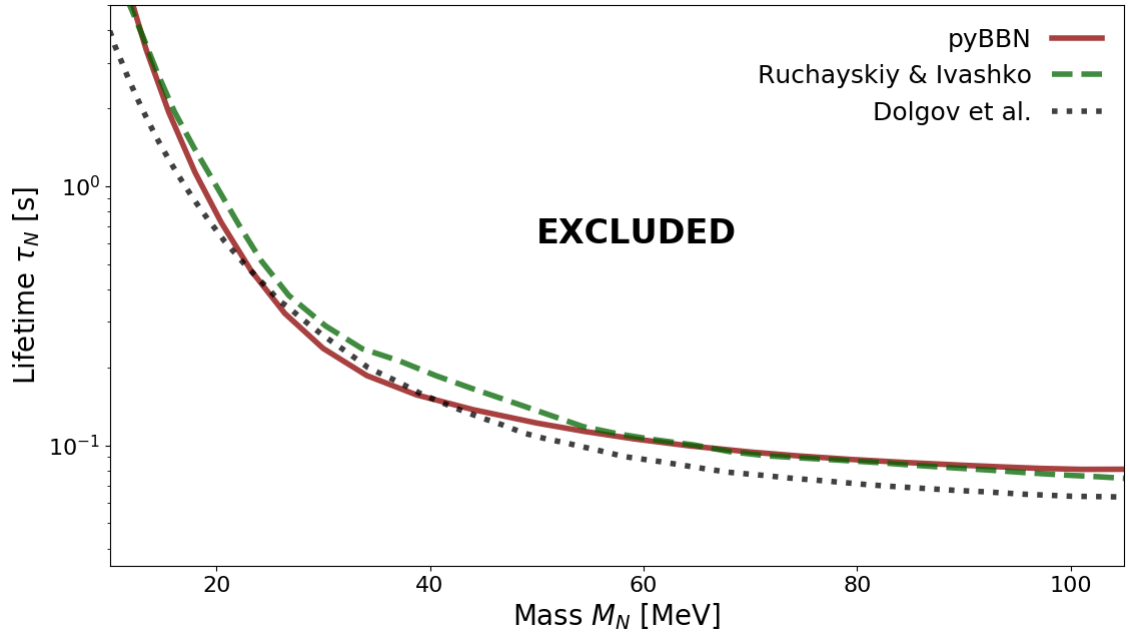


Figure 3.15: Constraints from BBN on the lifetime of HNLs up to the muon mass $M \approx 105$ MeV. Mixing with electron neutrino only is assumed here. Dashed curve is from [106], dotted from [105].

Boltzmann equations for nucleon number densities additional terms describing not decay of mesons, but their scattering on nucleons. This effect increases the neutron-to-proton ratio and hence the Helium abundance. Therefore, our constraints obtained without this effect are conservative. We will investigate this effect in the future.

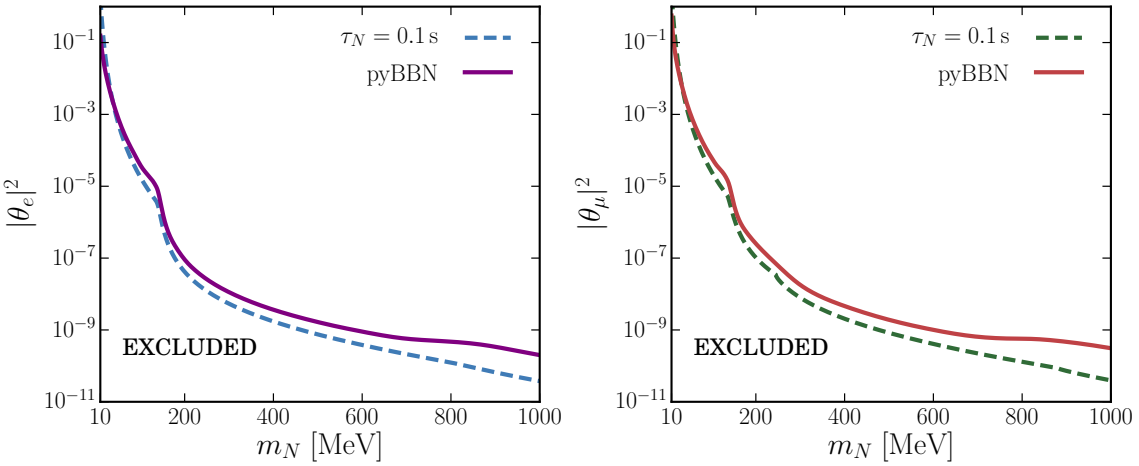


Figure 3.16: Bounds on HNL mixing angles. *Left:* mixing only with electron neutrinos. *Right:* mixing only with muon neutrinos.

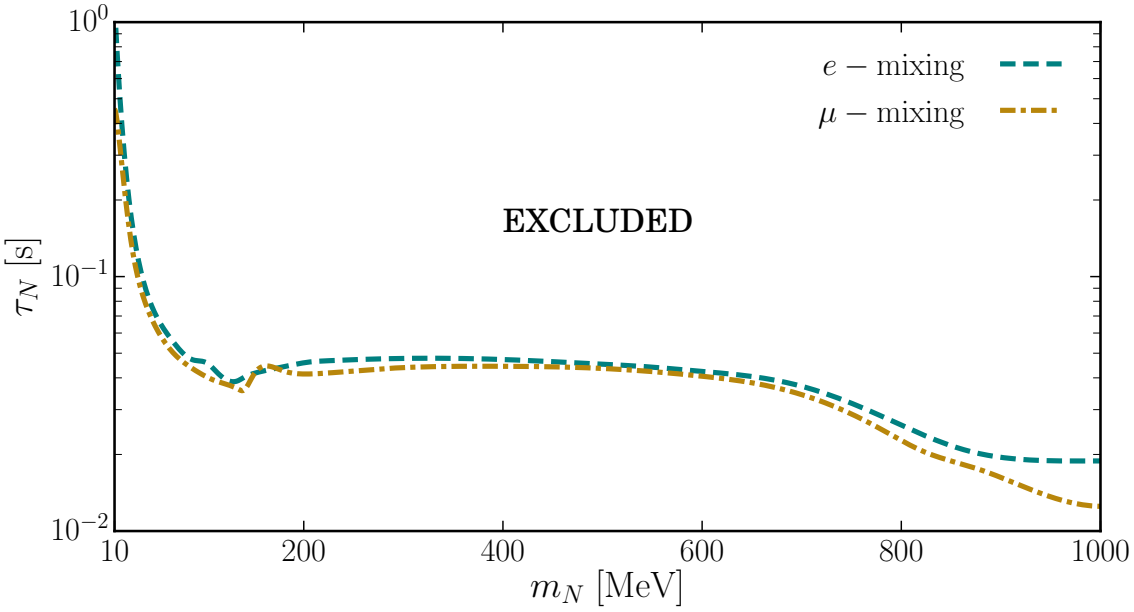


Figure 3.17: Constraints from BBN on the lifetime of HNLs up to the phi-meson mass $M \approx 1$ GeV (for electron and muon mixing).

Appendix

3.A Temperature evolution

Consider a plasma consisting of four particle species representing the contents of the Universe. For example, the four species here are photons, electrons, active neutrinos and HNLs. The addition of other species (e.g. muons) will then follow a similar procedure. In what follows: $\widetilde{T} = aT,\ \widetilde{E} = aE$. The total energy density and total pressure are given by

$$\rho_{\rm tot} = \rho_\gamma + \rho_e + \rho_\nu + \rho_N \qquad P_{\rm tot} = P_\gamma + P_e + P_\nu + P_N$$

$$\rho_\gamma = g_\gamma \frac{\pi^2}{30} \frac{1}{a^4} \widetilde{T}^4 \qquad P_\gamma = \frac{1}{3}\rho_\gamma$$

$$\rho_e = \frac{g_e}{2\pi^2} \frac{1}{a^4} \int \mathrm{d}y y^2 \frac{\widetilde{E_e}}{e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}+1} \qquad P_e = \frac{g_e}{6\pi^2} \frac{1}{a^4} \int \mathrm{d}y \frac{y^4}{\widetilde{E_e}} \frac{1}{e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}+1}$$

$$\rho_\nu = \frac{g_\nu}{2\pi^2} \frac{1}{a^4} \int \mathrm{d}y y^3 f_\nu \qquad P_\nu = \frac{1}{3}\rho_\nu$$

$$\rho_N = \frac{g_N}{2\pi^2} \frac{1}{a^4} \int \mathrm{d}y y^2 \sqrt{y^2 + a^2 m_N^2} f_N \qquad P_N = \frac{g_N}{6\pi^2} \frac{1}{a^4} \int \mathrm{d}y \frac{y^4}{\widetilde{E_N}} f_N$$

The energy conservation law is

$$\frac{\mathrm{d}\rho_{\rm tot}}{\mathrm{d}\ln a} \frac{\mathrm{d}\ln a}{\mathrm{d}t} + 3H\left(\rho_{\rm tot} + P_{\rm tot}\right) = 0 \Longrightarrow \frac{\mathrm{d}\rho_{\rm tot}}{\mathrm{d}\ln a} + 3\left(\rho_{\rm tot} + P_{\rm tot}\right) = 0$$

Taking derivatives:

$$\frac{\mathrm{d}\rho_\gamma}{\mathrm{d}\ln a} = g_\gamma \frac{\pi^2}{30}\left(-4\frac{1}{a^4}\widetilde{T}^4 + 4\frac{1}{a^4}\widetilde{T}^3\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}\ln a}\right) = -4\rho_\gamma + 4\frac{\rho_\gamma}{\widetilde{T}}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}\ln a}$$

$$\begin{aligned}\frac{\mathrm{d}\rho_e}{\mathrm{d}\ln a} &= -4\rho_e + \frac{g_e}{2\pi^2}\frac{1}{a^4}\int \mathrm{d}y y^2 \left[\frac{\left(e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}+1\right)\frac{a^2m_e^2}{\widetilde{E_e}} - \widetilde{E_e}e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}\left(\frac{a^2m_e^2}{\widetilde{T}\widetilde{E_e}} - \frac{\widetilde{E_e}}{\widetilde{T}^2}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}\ln a}\right)}{\left(e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}+1\right)^2}\right] \\ &= -4\rho_e + \frac{g_e}{2\pi}\frac{1}{a^4}\int \mathrm{d}y y^2 \left[\frac{a^2m_e^2}{\widetilde{E_e}}\frac{1}{e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}+1} - \left(\frac{a^2m_e^2}{\widetilde{T}} - \frac{\left(\widetilde{E_e}\right)^2}{\widetilde{T}^2}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}\ln a}\right)\frac{e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}}{\left(e^{\frac{1}{\widetilde{T}}\widetilde{E_e}}+1\right)^2}\right]\end{aligned}$$

$$\frac{\mathrm{d}\rho_\nu}{\mathrm{d}\ln a} = -4\rho_\nu + \frac{g_\nu}{2\pi^2}\frac{1}{a^4}\int \mathrm{d}y y^3 \frac{\mathrm{d}f_\nu}{\mathrm{d}\ln a} = -4\rho_\nu + \frac{g_\nu}{2\pi^2}\frac{1}{a^4}\int \mathrm{d}y y^3 \frac{1}{H} I_\nu$$

$$\begin{aligned}\frac{\mathrm{d}\rho_N}{\mathrm{d}\ln a} &= -4\rho_N + \frac{g_N}{2\pi^2}\frac{1}{a^4}\int \mathrm{d}y y^2 \left[\frac{a^2m_N^2}{\widetilde{E_N}} + \widetilde{E_N}\frac{\mathrm{d}f_N}{\mathrm{d}\ln a}\right] \\ &= -4\rho_N + \frac{g_N}{2\pi^2}\frac{1}{a^4}\int \mathrm{d}y y^2 \left[\frac{a^2m_N^2}{\widetilde{E_N}} + \widetilde{E_N}\frac{1}{H} I_N\right]\end{aligned}$$

The equation is also valid for individual species:

$$\begin{aligned}\frac{\mathrm{d}\rho_\gamma}{\mathrm{d}\ln a} + 3\left(\rho_\gamma + P_\gamma\right) &= 4\frac{\rho_\gamma}{\widetilde{T}}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}\ln a} \\ \frac{\mathrm{d}\rho_e}{\mathrm{d}\ln a} + 3\left(\rho_e + P_e\right) &= \frac{1}{a^4}\left[-\frac{a^2m_e^2}{\widetilde{T}}R_1 + \frac{1}{\widetilde{T}^2}\left\{R_2 + a^2m_e^2R_1\right\}\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}\ln a}\right] \\ \frac{\mathrm{d}\rho_\nu}{\mathrm{d}\ln a} + 3\left(\rho_\nu + P_\nu\right) &= \frac{g_\nu}{2\pi^2}\frac{1}{a^4}\int \mathrm{d}y y^3 \frac{1}{H} I_\nu \\ \frac{\mathrm{d}\rho_N}{\mathrm{d}\ln a} + 3\left(\rho_N + P_N\right) &= \frac{g_N}{2\pi^2}\frac{1}{a^4}\int \mathrm{d}y y^2 \widetilde{E_N}\frac{1}{H} I_N\ ,\end{aligned}$$

with

$$R_1 = \frac{g_e}{2\pi^2}\int \mathrm{d}y y^2 \frac{e^{\frac{\widetilde{E_e}}{\widetilde{T}}}}{\left(e^{\frac{\widetilde{E_e}}{\widetilde{T}}}+1\right)^2}$$

$$R_2 = \frac{g_e}{2\pi^2}\int \mathrm{d}y y^4 \frac{e^{\frac{\widetilde{E_e}}{\widetilde{T}}}}{\left(e^{\frac{\widetilde{E_e}}{\widetilde{T}}}+1\right)^2}$$

Adding all these terms together and solving for the temperature derivative gives:

$$\frac{\mathrm{d}\widetilde{T}}{\mathrm{d}\ln a} = \frac{\frac{a^2 m_e^2}{\widetilde{T}} R_1 - \frac{g_\nu}{2\pi^2}\int \mathrm{d}y y^3 \frac{1}{H} I_\nu - \frac{g_N}{2\pi^2}\int \mathrm{d}y y^2 \widetilde{E_N} \frac{1}{H} I_N}{\frac{2\pi^2 g_\gamma}{15}\widetilde{T}^3 + \frac{1}{\widetilde{T}^2} R_2 + \frac{a^2 m_e^2}{\widetilde{T}^2} R_1} \tag{3.A.1}$$

3.B Relevant matrix elements

The matrix elements listed here are not averaged over any helicities. Subsection 3.B.1 contains the reactions involving SM particles only, Subsection 3.B.2 the reactions involving HNLs above QCD-scale and Subsection 3.B.3 the reactions involving HNLs below QCD-scale. HNL decay channels with a branching ratio of at least 1% for some mass below $\sim$1 GeV are considered in this work (see Figure 3.18). The results for HNLs do not take into account charge conjugated channels, which are possible if they are Majorana particles.

The explicit determination of matrix elements involving multiple mesons can be extremely challenging. Therefore, an approximation has been used by assuming the matrix element to be constant and using the definition of decay width,

$$\Gamma = \frac{1}{2gM}\int \left(\prod_i \frac{\mathrm{d}^3 y_i}{(2\pi)^3 2E_i}\right) |\mathcal{M}|^2 (2\pi)^4 \delta^4 (P - \sum_i P_i) \,,$$

together with its measured value (from e.g. [88]) to solve for $|\mathcal{M}|^2$. For three-particle reactions this method gives the exact matrix element.

The values of the meson decay constants used in Subsection 3.B.3 are from [74] and summarized below.

f_{π^0}	$f_{\pi^\pm}$	f_η	f_{ρ^0}	$f_{\rho^\pm}$	f_ω	$f_{\eta'}$	f_ϕ	
130.2	130.2	81.7	208.9	208.9	195.5	-94.7	229.5	MeV

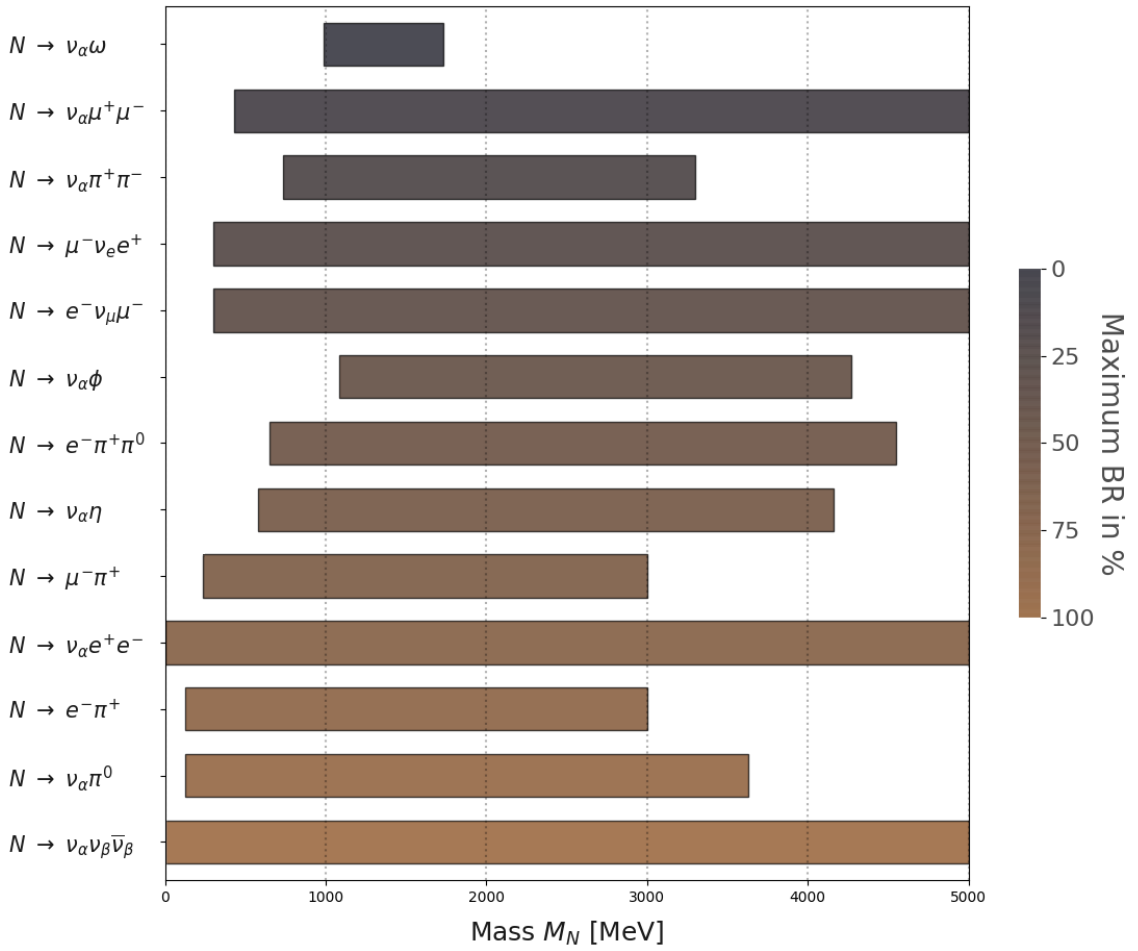


Figure 3.18: List of HNL decay channels with branching ratios more than 1% for some HNL mass below $\sim$ 1 GeV. The left border indicates the HNL mass where the branching ratio exceeds 1%, the right border when it falls below the 1 % threshold. In this plot a model was assumed where all three mixing angles are equal to each other.

3.B.1 Matrix elements in the Standard Model

3.B.1.1 Four-particle processes with leptons

Process $(1+2 \to 3+4)$	S	$SG_{\mathrm{F}}^2 a^{-4} \left\|\mathcal{M}\right\|^2$
$\nu_\alpha + \nu_\beta \to \nu_\alpha + \nu_\beta$	1	$32\,(Y_1 \cdot Y_2)\,(Y_3 \cdot Y_4)$
$\nu_\alpha + \overline{\nu_\beta} \to \nu_\alpha + \overline{\nu_\beta}$	1	$32\,(Y_1 \cdot Y_4)\,(Y_2 \cdot Y_3)$
$\nu_\alpha + \nu_\alpha \to \nu_\alpha + \nu_\alpha$	$\frac{1}{2}$	$64\,(Y_1 \cdot Y_2)\,(Y_3 \cdot Y_4)$
$\nu_\alpha + \overline{\nu_\alpha} \to \nu_\alpha + \overline{\nu_\alpha}$	1	$128\,(Y_1 \cdot Y_4)\,(Y_2 \cdot Y_3)$
$\nu_\alpha + \overline{\nu_\alpha} \to \nu_\beta + \overline{\nu_\beta}$	1	$32\,(Y_1 \cdot Y_4)\,(Y_3 \cdot Y_2)$
$\nu_e + \overline{\nu_e} \to e^+ + e^-$	1	$128\Big[g_L^2\,(Y_1 \cdot Y_3)\,(Y_2 \cdot Y_4) + g_R^2\,(Y_1 \cdot Y_4)\,(Y_2 \cdot Y_3)$ $+ g_L g_R a^2 m_e^2\,(Y_1 \cdot Y_2)\Big]$
$\nu_e + e^- \to \nu_e + e^-$	1	$128\Big[g_L^2\,(Y_1 \cdot Y_2)\,(Y_3 \cdot Y_4) + g_R^2\,(Y_1 \cdot Y_4)\,(Y_3 \cdot Y_2)$ $- g_L g_R a^2 m_e^2\,(Y_1 \cdot Y_3)\Big]$
$\nu_e + e^+ \to \nu_e + e^+$	1	$128\Big[g_L^2\,(Y_1 \cdot Y_4)\,(Y_3 \cdot Y_2) + g_R^2\,(Y_1 \cdot Y_2)\,(Y_3 \cdot Y_4)$ $- g_L g_R a^2 m_e^2\,(Y_1 \cdot Y_3)\Big]$
$\nu_{\mu/\tau} + \overline{\nu_{\mu/\tau}} \to e^+ + e^-$	1	$128\Big[\widetilde{g_L}^{\,2}\,(Y_1 \cdot Y_3)\,(Y_2 \cdot Y_4) + g_R^2\,(Y_1 \cdot Y_4)\,(Y_2 \cdot Y_3)$ $+ \widetilde{g_L} g_R a^2 m_e^2\,(Y_1 \cdot Y_2)\Big]$
$\nu_{\mu/\tau} + e^- \to \nu_{\mu/\tau} + e^-$	1	$128\Big[\widetilde{g_L}^{\,2}\,(Y_1 \cdot Y_2)\,(Y_3 \cdot Y_4) + g_R^2\,(Y_1 \cdot Y_4)\,(Y_3 \cdot Y_2)$ $- \widetilde{g_L} g_R a^2 m_e^2\,(Y_1 \cdot Y_3)\Big]$
$\nu_{\mu/\tau} + e^+ \to \nu_{\mu/\tau} + e^+$	1	$128\Big[\widetilde{g_L}^{\,2}\,(Y_1 \cdot Y_4)\,(Y_3 \cdot Y_2) + g_R^2\,(Y_1 \cdot Y_2)\,(Y_3 \cdot Y_4)$ $- \widetilde{g_L} g_R a^2 m_e^2\,(Y_1 \cdot Y_3)\Big]$

Table 3.2: Squared matrix elements for weak processes involving active neutrinos and electrons/positrons. S is the symmetry factor and $\alpha, \beta \in \{e, \mu, \tau\}$, where $\alpha \neq \beta$. Here: $g_R = \sin^2\theta_{\mathrm{W}},\ g_L = 1/2 + \sin^2\theta_{\mathrm{W}}$ and $\widetilde{g_L} = -1/2 + \sin^2\theta_{\mathrm{W}}$, with θ_{W} the Weinberg angle.

3.B.1.2 Three-particle and four-particle meson decays

Process ($1 \to 2+3$)	S	$\lvert\mathcal{M}\rvert^2$
$\pi^0 \to \gamma + \gamma$	1	$\alpha_{\mathrm{em}}^2 m_\pi^4 \left[2\pi^2 f_\pi^2\right]^{-1}$
$\pi^+ \to \mu^+ + \nu_\mu$	1	$2G_{\mathrm{F}}^2 \lvert V_{ud}\rvert^2 f_\pi^2 m_\mu^4 \left[\frac{m_\pi^2}{m_\mu^2} - 1\right]$

Table 3.3: Squared matrix elements for pion decays.

Process ($1 \to 2+3+4$)	$\lvert\mathcal{M}\rvert^2$	Process ($1 \to 2+3$)	$\lvert\mathcal{M}\rvert^2$ [MeV^2]
$K^+ \to \pi^0 + e^+ + \nu_e$	$1.42906 \cdot 10^{-13}$	$K^+ \to \pi^+ + \pi^0$	$3.28177 \cdot 10^{-10}$
$K^+ \to \pi^+ + \pi^- + \pi^+$	$1.85537 \cdot 10^{-12}$	$K^+ \to \mu^+ + \nu_\mu$	$8.78918 \cdot 10^{-10}$
$K_L^0 \to \pi^\pm + e^\mp + \nu_e$	$2.80345 \cdot 10^{-13}$	$K_S^0 \to \pi^+ + \pi^-$	$1.53713 \cdot 10^{-7}$
$K_L^0 \to \pi^\pm + \mu^\mp + \nu_\mu$	$3.03627 \cdot 10^{-13}$	$K_S^0 \to \pi^0 + \pi^0$	$6.71800 \cdot 10^{-8}$
$K_L^0 \to \pi^0 + \pi^0 + \pi^0$	$1.05573 \cdot 10^{-12}$	$\eta \to \gamma + \gamma$	$1.42174 \cdot 10^{1}$
$K_L^0 \to \pi^+ + \pi^- + \pi^0$	$8.26989 \cdot 10^{-13}$	$\rho^0 \to \pi^+ + \pi^-$	$1.86839 \cdot 10^{7}$
$\eta \to \pi^0 + \pi^0 + \pi^0$	$8.70984 \cdot 10^{-2}$	$\rho^+ \to \pi^+ + \pi^0$	$1.86390 \cdot 10^{7}$
$\eta \to \pi^+ + \pi^- + \pi^0$	$6.90629 \cdot 10^{-2}$	$\omega \to \pi^0 + \gamma$	$8.55086 \cdot 10^{4}$
$\eta \to \pi^+ + \pi^- + \gamma$	$4.66530 \cdot 10^{-3}$	$\eta' \to \rho^0 + \gamma$	$8.04463 \cdot 10^{3}$
$\omega \to \pi^+ + \pi^- + \pi^0$	$1.14569 \cdot 10^{3}$	$\phi \to K^+ + K^-$	$1.28798 \cdot 10^{6}$
$\eta' \to \pi^+ + \pi^- + \eta$	$4.38880 \cdot 10^{1}$	$\phi \to K_L^0 + K_S^0$	$1.03471 \cdot 10^{6}$
$\eta' \to \pi^0 + \pi^0 + \eta$	$2.00986 \cdot 10^{1}$	$\phi \to \rho^0 + \pi^0$	$2.86706 \cdot 10^{5}$

Table 3.4: Squared matrix elements for meson decays, where the constant matrix element approximation is used. For Majorana particles that can also decay through the charge-conjugated channel, the factor of 2 in the decay width is already taken into account here.

3.B.2 Matrix elements for Heavy Neutral Leptons above $\Lambda_{\rm QCD}$

3.B.2.1 Four-particle processes with leptons only

Process $(1+2 \to 3+4)$	S	$SG_{\rm F}^{-2}a^{-4}\left\|\mathcal{M}\right\|^2$
$N+\nu_\beta \to \nu_\alpha+\nu_\beta$	1	$32\left\|\theta_\alpha\right\|^2(Y_1\cdot Y_2)(Y_3\cdot Y_4)$
$N+\overline{\nu_\beta} \to \nu_\alpha+\overline{\nu_\beta}$	1	$32\left\|\theta_\alpha\right\|^2(Y_1\cdot Y_4)(Y_2\cdot Y_3)$
$N+\nu_\alpha \to \nu_\alpha+\nu_\alpha$	$\frac{1}{2}$	$64\left\|\theta_\alpha\right\|^2(Y_1\cdot Y_2)(Y_3\cdot Y_4)$
$N+\overline{\nu_\alpha} \to \nu_\alpha+\overline{\nu_\alpha}$	1	$128\left\|\theta_\alpha\right\|^2(Y_1\cdot Y_4)(Y_2\cdot Y_3)$
$N+\overline{\nu_\alpha} \to \nu_\beta+\overline{\nu_\beta}$	1	$32\left\|\theta_\alpha\right\|^2(Y_1\cdot Y_4)(Y_3\cdot Y_2)$
$N+\overline{\nu_e} \to e^+ + e^-$	1	$128\left\|\theta_e\right\|^2\Big[g_L^2(Y_1\cdot Y_3)(Y_2\cdot Y_4)+g_R^2(Y_1\cdot Y_4)(Y_2\cdot Y_3)$ $+g_Lg_Ra^2m_e^2(Y_1\cdot Y_2)\Big]$
$N+e^- \to \nu_e + e^-$	1	$128\left\|\theta_e\right\|^2\Big[g_L^2(Y_1\cdot Y_2)(Y_3\cdot Y_4)+g_R^2(Y_1\cdot Y_4)(Y_3\cdot Y_2)$ $-g_Lg_Ra^2m_e^2(Y_1\cdot Y_3)\Big]$
$N+e^+ \to \nu_e + e^+$	1	$128\left\|\theta_e\right\|^2\Big[g_L^2(Y_1\cdot Y_4)(Y_3\cdot Y_2)+g_R^2(Y_1\cdot Y_2)(Y_3\cdot Y_4)$ $-g_Lg_Ra^2m_e^2(Y_1\cdot Y_3)\Big]$
$N+\overline{\nu_{\mu/\tau}} \to e^+ + e^-$	1	$128\left\|\theta_{\mu/\tau}\right\|^2\Big[\widetilde{g_L}^2(Y_1\cdot Y_3)(Y_2\cdot Y_4)+$ $g_R^2(Y_1\cdot Y_4)(Y_2\cdot Y_3)$ $+\widetilde{g_L}g_Ra^2m_e^2(Y_1\cdot Y_2)\Big]$
$N+e^- \to \nu_{\mu/\tau} + e^-$	1	$128\left\|\theta_{\mu/\tau}\right\|^2\Big[\widetilde{g_L}^2(Y_1\cdot Y_2)(Y_3\cdot Y_4)+$ $g_R^2(Y_1\cdot Y_4)(Y_3\cdot Y_2)$ $-\widetilde{g_L}g_Ra^2m_e^2(Y_1\cdot Y_3)\Big]$
$N+e^+ \to \nu_{\mu/\tau} + e^+$	1	$128\left\|\theta_{\mu/\tau}\right\|^2\Big[\widetilde{g_L}^2(Y_1\cdot Y_4)(Y_3\cdot Y_2)+$ $g_R^2(Y_1\cdot Y_2)(Y_3\cdot Y_4)$ $-\widetilde{g_L}g_Ra^2m_e^2(Y_1\cdot Y_3)\Big]$

Table 3.5: Squared matrix elements for weak processes involving HNLs and leptons. S is the symmetry factor and $\alpha,\beta \in \{e,\mu,\tau\}$, where $\alpha \neq \beta$. Here: $g_R = \sin^2\theta_{\rm W},\ g_L = 1/2+\sin^2\theta_{\rm W}$ and $\widetilde{g_L} = -1/2+\sin^2\theta_{\rm W}$, with $\theta_{\rm W}$ the Weinberg angle.

Process ($1+2 \to 3+4$)	S	$SG_{\mathrm{F}}^{-2} a^{-4} \left\|\mathcal{M}\right\|^2$
$N + \overline{\nu_\mu} \to e^- + \mu^+$	1	$128 \left\|\theta_e\right\|^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3)$
$N + \overline{\nu_e} \to e^+ + \mu^-$	1	$128 \left\|\theta_\mu\right\|^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4)$
$N + e^- \to \nu_e + \mu^-$	1	$128 \left\|\theta_\mu\right\|^2 (Y_1 \cdot Y_2) (Y_3 \cdot Y_4)$
$N + e^+ \to \nu_\mu + \mu^+$	1	$128 \left\|\theta_e\right\|^2 (Y_1 \cdot Y_4) (Y_3 \cdot Y_2)$
$N + \overline{\nu_\mu} \to \mu^+ + \mu^-$	1	$128 \left\|\theta_\mu\right\|^2 \Big[g_L^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4) + g_R^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3) + g_L g_R a^2 m_\mu^2 (Y_1 \cdot Y_2) \Big]$
$N + \overline{\nu_{e/\tau}} \to \mu^+ + \mu^-$	1	$128 \left\|\theta_{e/\tau}\right\|^2 \Big[\widetilde{g_L}^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4) + g_R^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3) + \widetilde{g_L} g_R a^2 m_\mu^2 (Y_1 \cdot Y_2) \Big]$

Process ($1 \to 2+3+4$)	S	$SG_{\mathrm{F}}^{-2} a^{-4} \left\|\mathcal{M}\right\|^2$
$N \to \nu_\alpha + \nu_\beta + \overline{\nu_\beta}$	1	$32 \left\|\theta_\alpha\right\|^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3)$
$N \to \nu_\alpha + \nu_\alpha + \overline{\nu_\alpha}$	$\frac{1}{2}$	$64 \left\|\theta_\alpha\right\|^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3)$
$N \to \nu_e + e^+ + e^-$	1	$128 \left\|\theta_e\right\|^2 \Big[g_L^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4) + g_R^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3) + g_L g_R a^2 m_e^2 (Y_1 \cdot Y_2) \Big]$
$N \to \nu_{\mu/\tau} + e^+ + e^-$	1	$128 \left\|\theta_{\mu/\tau}\right\|^2 \Big[\widetilde{g_L}^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4) + g_R^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3) + \widetilde{g_L} g_R a^2 m_e^2 (Y_1 \cdot Y_2) \Big]$
$N \to \nu_\mu + e^- + \mu^+$	1	$128 \left\|\theta_e\right\|^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3)$
$N \to \nu_e + e^+ + \mu^-$	1	$128 \left\|\theta_\mu\right\|^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4)$
$N \to \nu_\mu + \mu^+ + \mu^-$	1	$128 \left\|\theta_\mu\right\|^2 \Big[g_L^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4) + g_R^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3) + g_L g_R a^2 m_\mu^2 (Y_1 \cdot Y_2) \Big]$
$N \to \nu_{e/\tau} + \mu^+ + \mu^-$	1	$128 \left\|\theta_{e/\tau}\right\|^2 \Big[\widetilde{g_L}^2 (Y_1 \cdot Y_3) (Y_2 \cdot Y_4) + g_R^2 (Y_1 \cdot Y_4) (Y_2 \cdot Y_3) + \widetilde{g_L} g_R a^2 m_\mu^2 (Y_1 \cdot Y_2) \Big]$

Table 3.6: Squared matrix elements for weak processes involving HNLs and leptons. Note: low temperatures are assumed here. At high temperatures, reactions such as $N + \mu^- \to e^- + \nu_\mu$ are possible. The corresponding matrix elements can be trivially deduced from the ones given above.

3.B.2.2 Four-particle processes with leptons and quarks

Process $(1+2 \to 3+4)$	S	$SG_{\mathrm{F}}^{-2} a^{-4} \left\|\mathcal{M}\right\|^2$
$N + \ell_\alpha^+ \to U + \overline{D}$	1	$128 \left\|\theta_\alpha\right\|^2 \left\|V_{ud}\right\|^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3)$
$N + D \to \ell_\alpha^- + U$	1	$128 \left\|\theta_\alpha\right\|^2 \left\|V_{ud}\right\|^2 (Y_1 \cdot Y_2)(Y_3 \cdot Y_4)$
$N + \overline{U} \to \ell_\alpha^- + \overline{D}$	1	$128 \left\|\theta_\alpha\right\|^2 \left\|V_{ud}\right\|^2 (Y_1 \cdot Y_4)(Y_3 \cdot Y_2)$
$N + \overline{\nu_\alpha} \to \overline{U} + U$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[16 g_R^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3) + (3-4g_R)^2 (Y_1 \cdot Y_3)(Y_2 \cdot Y_4) + 4 g_R \theta_W (4g_R - 3) a^2 m_U^2 (Y_1 \cdot Y_2) \Big]$
$N + U \to \nu_\alpha + U$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[16 g_R^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3) + (3-4g_R)^2 (Y_1 \cdot Y_2)(Y_3 \cdot Y_4) - 4 g_R (4g_R - 3) a^2 m_U^2 (Y_1 \cdot Y_3) \Big]$
$N + \overline{U} \to \nu_\alpha + \overline{U}$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[16 g_R^2 (Y_1 \cdot Y_2)(Y_3 \cdot Y_4) + (3-4g_R)^2 (Y_1 \cdot Y_4)(Y_3 \cdot Y_2) - 4 g_R (4g_R - 3) a^2 m_U^2 (Y_1 \cdot Y_3) \Big]$
$N + \overline{\nu_\alpha} \to \overline{D} + D$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[4 g_R^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3) + (3-2g_R)^2 (Y_1 \cdot Y_3)(Y_2 \cdot Y_4) + 2 g_R (2g_R - 3) a^2 m_D^2 (Y_1 \cdot Y_2) \Big]$
$N + D \to \nu_\alpha + D$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[4 g_R^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3) + (3-2g_R)^2 (Y_1 \cdot Y_2)(Y_3 \cdot Y_4) - 2 g_R (2g_R - 3) a^2 m_D^2 (Y_1 \cdot Y_3) \Big]$
$N + \overline{D} \to \nu_\alpha + \overline{D}$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[4 g_R^2 (Y_1 \cdot Y_2)(Y_3 \cdot Y_4) + (3-2g_R)^2 (Y_1 \cdot Y_4)(Y_3 \cdot Y_2) - 2 g_R (2g_R - 3) a^2 m_D^2 (Y_1 \cdot Y_3) \Big]$

Table 3.7: Squared matrix elements for weak scattering processes involving HNLs, leptons and quarks. Here: U are up-type quarks, D down-type quarks and $g_R = \sin^2 \theta_{\mathrm{W}}$.

Process ($1 \to 2+3+4$)	S	$SG_{\mathrm{F}}^{-2} a^{-4} \left\|\mathcal{M}\right\|^2$
$N \to \ell_\alpha^- + U + \overline{D}$	1	$128 \left\|\theta_\alpha\right\|^2 \left\|V_{ud}\right\|^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3)$
$N \to \nu_\alpha + \overline{U} + U$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[16 g_R^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3) + (3-4g_R)^2 (Y_1 \cdot Y_3)(Y_2 \cdot Y_4) + 4g_R(4g_R-3) a^2 m_U^2 (Y_1 \cdot Y_2)\Big]$
$N \to \nu_\alpha + \overline{D} + D$	1	$\frac{32}{9} \left\|\theta_\alpha\right\|^2 \Big[4 g_R^2 (Y_1 \cdot Y_4)(Y_2 \cdot Y_3) + (3-2g_R)^2 (Y_1 \cdot Y_3)(Y_2 \cdot Y_4) + 2g_R(2g_R-3) a^2 m_D^2 (Y_1 \cdot Y_2)\Big]$

Table 3.8: Squared matrix elements for weak decay processes involving HNLs, leptons and quarks. Here: U are up-type quarks, D down-type quarks and $g_R = \sin^2\theta_{\mathrm{W}}$.

3.B.3 Matrix elements for Heavy Neutral Leptons below Λ_{QCD}

In addition to interactions with leptons, HNLs will also decay into mesons.

3.B.3.1 Three-particle processes with single mesons

Process ($1 \to 2+3$)	S	$SG_{\mathrm{F}}^{-2} M_N^{-4} \left\|\mathcal{M}\right\|^2$
$N \to \nu_\alpha + \pi^0$	1	$\left\|\theta_\alpha\right\|^2 f_\pi^2 \left[1 - \frac{m_\pi^2}{M_N^2}\right]$
$N \to \ell_\alpha^\mp + \pi^\pm$	1	$2\left\|\theta_\alpha\right\|^2 \left\|V_{ud}\right\|^2 f_\pi^2 \left[\left(1 - \frac{m_{\ell_\alpha}^2}{M_N^2}\right)^2 - \frac{m_\pi^2}{M_N^2}\left(1 + \frac{m_{\ell_\alpha}^2}{M_N^2}\right)\right]$
$N \to \nu_\alpha + \eta$	1	$\left\|\theta_\alpha\right\|^2 f_\eta^2 \left[1 - \frac{m_\eta^2}{M_N^2}\right]$
$N \to \nu_\alpha + \rho^0$	1	$\left\|\theta_\alpha\right\|^2 \left(1 - 2\sin^2\theta_{\mathrm{W}}\right)^2 f_\rho^2 \left[1 + 2\frac{m_\rho^2}{M_N^2}\right]\left[1 - \frac{m_\rho^2}{M_N^2}\right]$
$N \to \ell_\alpha^\mp + \rho^\pm$	1	$2\left\|\theta_\alpha\right\|^2 \left\|V_{ud}\right\|^2 f_\rho^2 \left[\left(1 - \frac{m_{\ell_\alpha}^2}{M_N^2}\right)^2 + \frac{m_\rho^2}{M_N^2}\left(1 + \frac{m_{\ell_\alpha}^2}{M_N^2}\right) - 2\frac{m_\rho^4}{M_N^4}\right]$
$N \to \nu_\alpha + \omega$	1	$\left\|\theta_\alpha\right\|^2 \left(\frac{4}{3}\sin^2\theta_{\mathrm{W}}\right)^2 f_\omega^2 \left[1 + 2\frac{m_\omega^2}{M_N^2}\right]\left[1 - \frac{m_\omega^2}{M_N^2}\right]$
$N \to \nu_\alpha + \eta'$	1	$\left\|\theta_\alpha\right\|^2 f_{\eta'}^2 \left[1 - \frac{m_{\eta'}^2}{M_N^2}\right]$
$N \to \nu_\alpha + \phi$	1	$\left\|\theta_\alpha\right\|^2 \left(\frac{4}{3}\sin^2\theta_{\mathrm{W}} - 1\right)^2 f_\phi^2 \left[1 + 2\frac{m_\phi^2}{M_N^2}\right]\left[1 - \frac{m_\phi^2}{M_N^2}\right]$

Table 3.9: Squared matrix elements for HNL decays into mesons.

3.C Collision integrals

Consider the Boltzmann equation in comoving coordinates:

$$\frac{\mathrm{d}f_1}{\mathrm{d}t} = \frac{\mathrm{d}f_1}{\mathrm{d}\ln a}\frac{\mathrm{d}\ln a}{\mathrm{d}t} = \frac{\mathrm{d}f_1}{\mathrm{d}\ln a}H = \sum_{\text{reactions}} I_{\text{coll}} \,, \tag{3.C.1}$$

with

$$I_{\text{coll}} = \frac{a^{7-2Q}}{2g\widetilde{E_1}} \sum_{\text{in,out}} \int \left(\prod_{i=2}^{Q} \frac{\mathrm{d}^3 y_i}{(2\pi)^3 2\widetilde{E_i}} \right) S|\mathcal{M}|^2 F[f] (2\pi)^4 \delta^4(Y_{\text{in}} - Y_{\text{out}}) \tag{3.C.2}$$

The delta function can be rewritten as

$$\delta^4(Y_{\text{in}} - Y_{\text{out}}) = \delta^4(s_1 Y_1 + s_2 Y_2 + ... + s_Q Y_Q) \,, \tag{3.C.3}$$

with $s_i = \{-1, 1\}$ if particle i is on the {left, right}-hand side of the reaction. The $Y_i = aP_i$ here are the comoving four-momenta.

3.C.1 Three-particle collision integral

$$I_{\text{coll}} = \frac{a}{2\widetilde{E_1}} \int \frac{\mathrm{d}^3 y_2 \mathrm{d}^3 y_3}{(2g\pi)^6 2\widetilde{E_2} 2\widetilde{E_3}} S|\mathcal{M}|^2 F[f] (2\pi)^4 \delta^4(s_1 Y_1 + s_2 Y_2 + s_3 Y_3) \tag{3.C.4}$$

3.C.1.1 Case $y_1 \neq 0$

Since a homogeneous and isotropic universe is assumed, only absolute values of momenta are relevant. Moreover, the matrix element in three particle interactions is independent of the four-momenta. The collision integral becomes:

$$I_{\text{coll}} = \frac{S|\mathcal{M}|^2 a}{8(2\pi)^2 g\widetilde{E_1}} \int \frac{\mathrm{d}y_2 \mathrm{d}y_3 \mathrm{d}\Omega_2 \mathrm{d}\Omega_3 y_2^2 y_3^2}{\widetilde{E_2}\widetilde{E_3}} F[f] \delta^4(s_1 Y_1 + s_2 Y_2 + s_3 Y_3) \tag{3.C.5}$$

Using the identity

$$\delta^3(s_1 \mathbf{y_1} + s_2 \mathbf{y_2} + s_3 \mathbf{y_3}) = \frac{1}{(2\pi)^3} \int \mathrm{d}\lambda \mathrm{d}\Omega_\lambda \lambda^2 e^{i(s_1 \mathbf{y_1} + s_2 \mathbf{y_2} + s_3 \mathbf{y_3}) \cdot \boldsymbol{\lambda}} \tag{3.C.6}$$

gives

$$
\begin{aligned}
I_{\text{coll}} =& \frac{S|\mathcal{M}|^2 a}{8(2\pi)^5 g\widetilde{E_1}} \int \frac{\mathrm{d}y_2 \mathrm{d}y_3 y_2^2 y_3^2}{\widetilde{E_2}\widetilde{E_3}} F[f]\delta(s_1\widetilde{E_1} + s_2\widetilde{E_2} + s_3\widetilde{E_3}) \cdot \\
& \cdot \int \mathrm{d}\lambda\lambda^2 \int \mathrm{d}\Omega_\lambda e^{is_1 y_1 \lambda \cos\theta_\lambda} \int \mathrm{d}\Omega_2 e^{is_1 y_2 \lambda \cos\theta_2} \int \mathrm{d}\Omega_3 e^{is_1 y_3 \lambda \cos\theta_3} \\
=& \frac{S|\mathcal{M}|^2 a}{8(2\pi)^5 g\widetilde{E_1}} \int \frac{\mathrm{d}y_2 \mathrm{d}y_3 y_2^2 y_3^2}{\widetilde{E_2}\widetilde{E_3}} F[f]\delta(s_1\widetilde{E_1} + s_2\widetilde{E_2} + s_3\widetilde{E_3}) \cdot \\
& \cdot \int \mathrm{d}\lambda\lambda^2 \left(4\pi \frac{\sin(y_1\lambda)}{y_1\lambda}\right)\left(4\pi \frac{\sin(y_2\lambda)}{y_2\lambda}\right)\left(4\pi \frac{\sin(y_3\lambda)}{y_3\lambda}\right) \\
=& \frac{S|\mathcal{M}|^2 a}{(2\pi)^2 g\widetilde{E_1} y_1} \int \frac{\mathrm{d}y_2 \mathrm{d}y_3 y_2 y_3}{\widetilde{E_2}\widetilde{E_3}} F[f]\delta(s_1\widetilde{E_1} + s_2\widetilde{E_2} + s_3\widetilde{E_3}) \cdot \\
& \cdot \int \frac{\mathrm{d}\lambda}{\lambda} \sin(y_1\lambda)\sin(y_2\lambda)\sin(y_3\lambda)
\end{aligned}
\tag{3.C.7}
$$

Rewrite the delta function of energies as

$$
\begin{aligned}
\int \frac{\mathrm{d}y_3 y_3}{\widetilde{E_3}} \delta(s_1\widetilde{E_1} + s_2\widetilde{E_2} + s_3\widetilde{E_3}) &= \int \mathrm{d}y_3 \frac{y_3}{\widetilde{E_3}} \frac{\delta\left(y_3 - y_3^*\right)}{\frac{y_3^*}{\widetilde{E_3^*}}} \theta\left(\left((s_1\widetilde{E_1} + s_2\widetilde{E_2}\right))^2 - a^2 m_3^2\right) \\
&= \int \mathrm{d}y_3 \frac{y_3}{\widetilde{E_3}} \frac{\widetilde{E_3^*}}{y_3^*} \delta(y_3 - y_3^*)\, \theta\left(\left(\widetilde{E_3^*}\right)^2 - a^2 m_3^2\right) ,
\end{aligned}
\tag{3.C.8}
$$

where $\left(\widetilde{E_3^*}\right)^2 = \left(y_3^*\right)^2 + \frac{x^2}{M^2} m_3^2 = \left(s_1\widetilde{E_1} + s_2\widetilde{E_2}\right)^2$ and $y_3^* = \sqrt{(s_1\widetilde{E_1} + s_2\widetilde{E_2})^2 - a^2 m_3^2}$.

Plugging Eq. (3.C.8) in Eq. (3.C.7) above:

$$
I_{\text{coll}} = \frac{S|\mathcal{M}|^2 a}{(2\pi)^2 g\widetilde{E_1} y_1} \int \frac{\mathrm{d}y_2 y_2}{\widetilde{E_2}} F[f] \int \frac{\mathrm{d}\lambda}{\lambda} \sin(y_1\lambda)\sin(y_2\lambda)\sin(y_3^*\lambda)\theta\left(\left(\widetilde{E_3^*}\right)^2 - a^2 m_3^2\right)
\tag{3.C.9}
$$

Now, the integral over λ is equal to

$$
\mathcal{X} = \frac{\pi}{8}\left(-\mathrm{Sgn}[y_1 - y_2 - y_3^*] + \mathrm{Sgn}[y_1 + y_2 - y_3^*] + \mathrm{Sgn}[y_1 - y_2 + y_3^*] - 1\right) , \tag{3.C.10}
$$

with Sgn the signum function and where $y_1 \geq y_2 \geq y_3$ is assumed.
The final form is then

$$
I_{\text{coll}} = \frac{S|\mathcal{M}|^2 a}{(2\pi)^2 g\widetilde{E_1} y_1} \int \frac{\mathrm{d}y_2 y_2}{\widetilde{E_2}} \mathcal{X}\theta\left(\left(s_1\widetilde{E_1} + s_2\widetilde{E_2}\right)^2 - a^2 m_3^2\right)\left(F[f]\right)\Big|_{y_3 = y_3^*}
\tag{3.C.11}
$$

3.C.1.2 Case $y_1 = 0$

$$
\begin{aligned}
I_{\text{coll}} =& \frac{S|\mathcal{M}|^2 a}{8(2\pi)^2 g a m_1} \int \frac{\mathrm{d}^3 y_2 \mathrm{d}^3 y_3}{\widetilde{E_2}\widetilde{E_3}} F[f] \delta\left(s_1 a m_1 + s_2 \widetilde{E_2} + s_3 \widetilde{E_3}\right) \delta^3\left(s_2 \mathbf{y_2} + s_3 \mathbf{y_3}\right) \\
=& \frac{S|\mathcal{M}|^2}{8(2\pi)^2 g m_1} \int \mathrm{d}^3 y_2 F[f] \delta\left(s_1 a m_1 + s_2 \sqrt{y_2^2 + (a m_2)^2} + s_3 \sqrt{y_2^2 + (a m_3)^2}\right) \cdot \\
& \cdot \left(\sqrt{y_2^2 + (a m_2)^2} \sqrt{y_2^2 + (a m_3)^2}\right)^{-1} \\
=& \frac{S|\mathcal{M}|^2}{8\pi g m_1} \int \mathrm{d} y_2 y_2^2 F[f] \delta\left(y_2 - y_2^*\right) \left| \frac{s_2 y_2^*}{\sqrt{(y_2^*)^2 + a^2 m_2^2}} + \frac{s_3 y_2^*}{\sqrt{(y_2^*)^2 + a^2 m_3^2}} \right|^{-1} \cdot \\
& \cdot \left(\sqrt{y_2^2 + (a m_2)^2} \sqrt{y_2^2 + (a m_3)^2}\right)^{-1} \theta\left(\left(s_1 a m_1 + s_2 \widetilde{E_2}\right)^2 - a^2 m_3^2\right) \\
=& \frac{S|\mathcal{M}|^2}{8\pi g m_1} y_2^* \left| s_2 \sqrt{(y_2^*)^2 + (a m_3)^2} + s_3 \sqrt{(y_2^*)^2 + (a m_2)^2} \right|^{-1} \theta\left(\left(\widetilde{E_3^*}\right)^2 - a^2 m_3^2\right) \cdot \\
& \cdot \left(F[f]\right) \Big|_{y_1 = 0,\, y_2 = y_2^*,\, y_3 = -y_2^*} \\
=& \frac{S|\mathcal{M}|^2}{8\pi g m_1} y_2^* \left|s_1 s_2 s_3 a m_1\right|^{-1} \theta\left(\left(\widetilde{E_3^*}\right)^2 - a^2 m_3^2\right) \left(F[f]\right) \Big|_{y_1 = 0,\, y_2 = y_2^*,\, y_3 = -y_2^*} \\
=& \frac{S|\mathcal{M}|^2}{8\pi g m_1^2} \frac{y_2^*}{a} \theta\left(\left(\widetilde{E_3^*}\right)^2 - a^2 m_3^2\right) \left(F[f]\right) \Big|_{y_1 = 0,\, y_2 = y_2^*,\, y_3 = -y_2^*} ,
\end{aligned}
\tag{3.C.12}
$$

with $\left(\widetilde{E_3^*}\right)^2 = \left(s_1 a m_1 + s_2 \widetilde{E_2}\right)^2$ and $y_2^* = a\sqrt{\frac{\left(m_1^2 - m_2^2 - m_3^2\right)^2 - 4 m_2^2 m_3^2}{4 m_1^2}}$.

3.C.2 Four-particle collision integral

$$
I_{\text{coll}} = \frac{1}{2g\widetilde{E_1}} \frac{1}{a} \int \frac{\mathrm{d}^3 y_2 \mathrm{d}^3 y_3 \mathrm{d} y_4^3}{(2\pi)^9 8 \widetilde{E_2}\widetilde{E_3}\widetilde{E_4}} S|\mathcal{M}|^2 F[f] (2\pi)^4 \delta^4(s_1 Y_1 + s_2 Y_2 + s_3 Y_3 + s_4 Y_4)
\tag{3.C.13}
$$

As can be seen in Appendix 3.B, $|\mathcal{M}|^2$ can be written as

$$
|\mathcal{M}|^2 = \frac{1}{a^4} \sum_{i \neq j \neq k \neq l} \left[K_1 (Y_i \cdot Y_j)(Y_k \cdot Y_l) + K_2 a^2 m_i m_j (Y_k \cdot Y_l)\right]
\tag{3.C.14}
$$

A similar procedure as with the three-particle case is followed here.

3.C.2.1 Case $y_1 \neq 0$

$$
\begin{aligned}
I_{\rm coll} =& \frac{S}{16(2\pi)^5 g\widetilde{E_1} a}\int \frac{\mathrm{d}y_2\mathrm{d}y_3\mathrm{d}y_4 y_2^2 y_3^2 y_4^2}{\widetilde{E_2}\widetilde{E_3}\widetilde{E_4}} F[f]\delta\left(s_1\widetilde{E_1}+s_2\widetilde{E_2}+s_3\widetilde{E_3}+s_4\widetilde{E_4}\right)\cdot \\
&\cdot\int \mathrm{d}\Omega_2\mathrm{d}\Omega_3\mathrm{d}\Omega_4\left|\mathcal{M}\right|^2\left|\delta^3\left(s_1\mathbf{y_1}+s_2\mathbf{y_2}+s_3\mathbf{y_3}+s_4\mathbf{y_4}\right)\right. \\
=& \frac{S}{64\pi^3 g\widetilde{E_1} y_1 a^5}\int \frac{\mathrm{d}y_2\mathrm{d}y_3\mathrm{d}y_4 y_2 y_3 y_4}{\widetilde{E_2}\widetilde{E_3}\widetilde{E_4}} F[f]\delta\left(s_1\widetilde{E_1}+s_2\widetilde{E_2}+s_3\widetilde{E_3}s_4\widetilde{E_4}\right)\cdot \\
&\cdot\mathcal{D}(Y_1,Y_2,Y_3,Y_4)\ ,
\end{aligned}
\tag{3.C.15}
$$

with

$$
\begin{aligned}
\mathcal{D}(Y_1,Y_2,Y_3,Y_4) =& \frac{y_1y_2y_3y_4}{64\pi^5}\int \mathrm{d}\Omega_2\mathrm{d}\Omega_3\mathrm{d}\Omega_4\left|\mathcal{M}\right|^2\left|\delta^3\left(s_1\mathbf{y_1}+s_2\mathbf{y_2}+s_3\mathbf{y_3}+s_4\mathbf{y_4}\right)\right. \\
=& \frac{y_1y_2y_3y_4}{64\pi^5}\int \mathrm{d}\lambda\lambda^2\int \mathrm{d}\Omega_\lambda e^{is_1\mathbf{y_1}\cdot\boldsymbol{\lambda}}\int \mathrm{d}\Omega_2 e^{is_2\mathbf{y_2}\cdot\boldsymbol{\lambda}}\int \mathrm{d}\Omega_3 e^{is_3\mathbf{y_3}\cdot\boldsymbol{\lambda}}\cdot \\
&\cdot\int \mathrm{d}\Omega_4 e^{is_4\mathbf{y_4}\cdot\boldsymbol{\lambda}}\sum_{i\neq j\neq k\neq l}\left[K_1(Y_i\cdot Y_j)(Y_k\cdot Y_l)+K_2a^2m_im_j(Y_k\cdot Y_l)\right] \\
=& \frac{y_1y_2y_3y_4}{64\pi^5}\sum_{i\neq j\neq k\neq l}\int \mathrm{d}\lambda\lambda^2\int \mathrm{d}\Omega_\lambda e^{is_iy_i\lambda\cos\theta_i}\int \mathrm{d}\Omega_j e^{is_jy_j\lambda\cos\theta_j}\cdot \\
&\cdot\int \mathrm{d}\Omega_k e^{is_ky_k\lambda\cos\theta_k}\int \mathrm{d}\Omega_l e^{is_ly_l\lambda\cos\theta_l}\left[K_1(Y_i\cdot Y_j)(Y_k\cdot Y_l)+\right. \\
&\left.+\ K_2a^2m_im_j(Y_k\cdot Y_l)\right]
\end{aligned}
\tag{3.C.16}
$$

Working out the inner products

$$
\begin{aligned}
Y_i\cdot Y_j =& \widetilde{E_i}\widetilde{E_j}-\mathbf{y_i}\cdot\mathbf{y_j}=\widetilde{E_i}\widetilde{E_j}-y_iy_j\cos\theta_{ij} \\
=& \widetilde{E_i}\widetilde{E_j}-y_iy_j\left(\cos\theta_i\cos\theta_j+\cos(\phi_i-\phi_j)\sin\theta_i\sin\theta_j\right)\ ,
\end{aligned}
\tag{3.C.17}
$$

where θ_{ij} is the angle between vectors $\mathbf{y_i}$ and $\mathbf{y_j}$, and using that

$$
\int_0^\pi\int_0^{2\pi}\mathrm{d}\theta_i\mathrm{d}\phi_i e^{is_iy_i\lambda\cos\theta_i}\sin^2\theta_i\cos(\phi_i-\phi_j)=0 \tag{3.C.18}
$$

gives

$$\mathcal{D}(Y_1,Y_2,Y_3,Y_4) = \frac{y_1y_2y_3y_4}{64\pi^5}\sum_{i\neq j\neq k\neq l}\int \mathrm{d}\lambda\lambda^2\int \mathrm{d}\theta_i\mathrm{d}\phi_i\sin\theta_i e^{is_iy_i\lambda\cos\theta_i}\cdot \tag{3.C.19}$$
$$\cdot\int \mathrm{d}\theta_j\mathrm{d}\phi_j\sin\theta_j e^{is_jy_j\lambda\cos\theta_j}\int \mathrm{d}\theta_k\mathrm{d}\phi_k\sin\theta_k e^{is_ky_k\lambda\cos\theta_k}\cdot$$
$$\cdot\int \mathrm{d}\theta_l\mathrm{d}\phi_l\sin\theta_l e^{is_ly_l\lambda\cos\theta_l}\left[K_1\left(\widetilde{E_i}\widetilde{E_j}-y_iy_j\cos\theta_i\cos\theta_j\right)\cdot\right.$$
$$\left.\cdot\left(\widetilde{E_k}\widetilde{E_l}-y_ky_l\cos\theta_k\cos\theta_l\right)+K_2a^2m_im_j\left(\widetilde{E_k}\widetilde{E_l}-y_ky_l\cos\theta_k\cos\theta_l\right)\right]$$

The integrals over the angles are given by

$$\int_0^\pi\int_0^{2\pi}\mathrm{d}\theta\mathrm{d}\phi\sin\theta e^{isy\lambda\cos\theta} = 4\pi\frac{\sin(y\lambda)}{y\lambda} \tag{3.C.20}$$
$$\int_0^\pi\int_0^{2\pi}\mathrm{d}\theta\mathrm{d}\phi\sin\theta\cos\theta e^{isy\lambda\cos\theta} = \frac{4\pi}{isy\lambda}\left[\cos(y\lambda)-\frac{\sin(y\lambda)}{y\lambda}\right] \tag{3.C.21}$$
$$\tag{3.C.22}$$

and working out all the brackets gives

$$\mathcal{D}(Y_1,Y_2,Y_3,Y_4) = \sum_{i\neq j\neq k\neq l}\left[K_1\left\{\widetilde{E_1}\widetilde{E_2}\widetilde{E_3}\widetilde{E_4}D_1\left(y_1,y_2,y_3,y_4\right)+\widetilde{E_i}\widetilde{E_j}D_2\left(y_i,y_j,y_k,y_l\right)+\right.\right.$$
$$+\left.\widetilde{E_k}\widetilde{E_l}D_2\left(y_k,y_l,y_i,y_j\right)+D_3\left(y_1,y_2,y_3,y_4\right)\right\}+$$
$$\left.+K_2a^2m_im_j\left\{\widetilde{E_k}\widetilde{E_l}D_1\left(y_1,y_2,y_3,y_4\right)+D_2\left(y_i,y_j,y_k,y_l\right)\right\}\right], \tag{3.C.23}$$

with

$$D_1\left(y_i,y_j,y_k,y_l\right) = \frac{4}{\pi}\int\frac{\mathrm{d}\lambda}{\lambda^2}\sin(y_i\lambda)\sin(y_j\lambda)\sin(y_k\lambda)\sin(y_l\lambda) \tag{3.C.24}$$
$$D_2\left(y_i,y_j,y_k,y_l\right) = s_ks_l\frac{4y_ky_l}{\pi}\int\frac{\mathrm{d}\lambda}{\lambda^2}\sin(y_i\lambda)\sin(y_j\lambda)\left[\cos(y_k\lambda)-\frac{\sin(y_k\lambda)}{y_k\lambda}\right]\cdot$$
$$\cdot\left[\cos(y_l\lambda)-\frac{\sin(y_l\lambda)}{y_l\lambda}\right] \tag{3.C.25}$$
$$D_3\left(y_i,y_j,y_k,y_l\right) = s_is_js_ks_l\frac{4y_iy_jy_ky_l}{\pi}\int\frac{\mathrm{d}\lambda}{\lambda^2}\left[\cos(y_i\lambda)-\frac{\sin(y_i\lambda)}{y_i\lambda}\right]\left[\cos(y_j\lambda)-\frac{\sin(y_j\lambda)}{y_j\lambda}\right]\cdot$$
$$\cdot\left[\cos(y_k\lambda)-\frac{\sin(y_k\lambda)}{y_k\lambda}\right]\left[\cos(y_l\lambda)-\frac{\sin(y_l\lambda)}{y_l\lambda}\right] \tag{3.C.26}$$

All these three functions are symmetric under the exchange $y_i \leftrightarrow y_j$ and $y_k \leftrightarrow y_l$, which then allows to take $y_i > y_j$ and $y_k > y_l$. Integrating out λ gives the functions in terms of

polynomials for all possible cases (factors $s_k s_l$ and $s_i s_j s_k s_l$ omitted):

- $y_i > y_j + y_k + y_l$ or $y_k > y_i + y_j + y_l$:

$$D_1 = D_2 = D_3 = 0$$

- $y_i + y_j > y_k + y_l$ and $y_i + y_l < y_j + y_k$:

$$\begin{aligned} D_1 =& y_l \\ D_2 =& \frac{1}{3} y_l^3 \\ D_3 =& \frac{1}{30} y_l^3 \left[5\left(y_i^2 + y_j^2 + y_k^2\right) - y_l^2\right] \end{aligned}$$

- $y_i + y_j > y_k + y_l$ and $y_i + y_l > y_j + y_k$:

$$\begin{aligned} D_1 =& \frac{1}{2}(y_j + y_k + y_l - y_i) \\ D_2 =& \frac{1}{12}\left[(y_i - y_j)\left\{(y_i - y_j)^2 - 3\left(y_k^2 + y_l^2\right)\right\} + 2\left(y_k^3 + y_l^3\right)\right] \\ D_3 =& \frac{1}{60}\left[y_i^5 - y_j^5 - y_k^5 - y_l^5 + 5\left(-y_i^3 y_j^2 + y_i^2 y_j^3 - y_i^3 y_k^2 + y_i^2 y_k^3\right.\right. \\ & \left.\left. - y_i^3 y_l^2 + y_i^2 y_l^3 + y_j^3 y_k^2 + y_j^2 y_k^3 + y_j^3 y_l^2 + y_j^2 y_l^3 + y_k^3 y_l^2 + y_k^2 y_l^3\right)\right] \end{aligned}$$

- $y_i + y_j < y_k + y_l$ and $y_i + y_l > y_j + y_k$:

$$\begin{aligned} D_1 =& y_j \\ D_2 =& \frac{1}{6} y_j \left[3\left(y_k^2 + y_l^2 - y_i^2\right) - y_j^2\right] \\ D_3 =& \frac{1}{30} y_j^3 \left[5\left(y_i^2 + y_k^2 + y_l^2\right) - y_j^2\right] \end{aligned}$$

- $y_i + y_j < y_k + y_l$ and $y_i + y_l < y_j + y_k$:

$$\begin{aligned} D_1 =& \frac{1}{2}(y_i + y_j + y_l - y_k) \\ D_2 =& -\frac{1}{12}\left[(y_i + y_j)\left\{(y_i + y_j)^2 - 3\left(y_k^2 + y_l^2\right)\right\} + 2\left(y_k^3 - y_l^3\right)\right] \\ D_3 =& \frac{1}{60}\left[y_k^5 - y_i^5 - y_j^5 - y_l^5 + 5\left(-y_k^3 y_l^2 + y_k^2 y_l^3 - y_k^3 y_i^2 + y_k^2 y_i^3\right.\right. \\ & \left.\left. - y_k^3 y_j^2 + y_k^2 y_j^3 + y_l^3 y_i^2 + y_l^2 y_i^3 + y_l^3 y_j^2 + y_l^2 y_j^3 + y_i^3 y_j^2 + y_i^2 y_j^3\right)\right] \end{aligned}$$

Going back to the collision integral, the same trick as before can be applied to the delta

function of energies, which then gives:

$$I_{\text{coll}} = \frac{S}{64\pi^3 g\widetilde{E_1} y_1 a^5} \int \mathrm{d}y_2 \mathrm{d}y_3 \frac{y_2 y_3}{\widetilde{E_2}\widetilde{E_3}} \mathcal{D}(Y_1, Y_2, Y_3, Y_4) \cdot \\ \cdot \theta\left(\left((s_1\widetilde{E_1} + s_2\widetilde{E_2} + s_3\widetilde{E_3}\right)^2 - a^2 m_4^2\right)(F[f])\Big|_{y_4 = y_4^*} , \tag{3.C.27}$$

with $y_4^* = \sqrt{\left(s_1\widetilde{E_1} + s_2\widetilde{E_2} + s_3\widetilde{E_3}\right)^2 - a^2 m_4^2}$.

3.C.2.2 Case $y_1 = 0$

$$I_{\text{coll}} = \frac{S}{64\pi^3 g a m_1} \frac{1}{a^5} \int \frac{\mathrm{d}y_2 \mathrm{d}y_3 \mathrm{d}y_4 y_2 y_3 y_4}{\widetilde{E_2}\widetilde{E_3}\widetilde{E_4}} F[f] \delta\left(s_1 a m_1 + s_2\widetilde{E_2} + s_3\widetilde{E_3} + s_4\widetilde{E_4}\right) \cdot \\ \cdot \mathcal{B}(Y_1, Y_2, Y_3, Y_4) , \tag{3.C.28}$$

with

$$\begin{aligned} \mathcal{B}(Y_1, Y_2, Y_3, Y_4) =& \frac{y_2 y_3 y_4}{64\pi^5} \int \mathrm{d}\Omega_2 \mathrm{d}\Omega_3 \mathrm{d}\Omega_4 \left|\mathcal{M}\right|^2 \left|\delta^3\left(s_2 \mathbf{y_2} + s_3 \mathbf{y_3} + s_4 \mathbf{y_4}\right)\right. \\ =& \frac{y_2 y_3 y_4}{64\pi^5} \int \mathrm{d}\lambda \lambda^2 \mathrm{d}\Omega_\lambda \int \mathrm{d}\theta_2 \mathrm{d}\phi_2 \sin\theta_2 e^{i s_2 y_2 \lambda \cos\theta_2} \cdot \\ & \cdot \int \mathrm{d}\theta_3 \mathrm{d}\phi_3 \sin\theta_3 e^{i s_3 y_3 \lambda \cos\theta_3} \int \mathrm{d}\theta_4 \mathrm{d}\phi_4 \sin\theta_4 e^{i s_4 y_4 \lambda \cos\theta_4} \cdot \\ & \cdot \sum_{i \neq j \neq k \neq l} \left[K_1 \left(\widetilde{E_i}\widetilde{E_j} - y_i y_j \cos\theta_i \cos\theta_j\right) \cdot \left(\widetilde{E_k}\widetilde{E_l} - y_k y_l \cos\theta_k \cos\theta_l\right) + \right. \\ & \left. + K_2\, a^2 m_i m_j \left(\widetilde{E_k}\widetilde{E_l} - y_k y_l \cos\theta_k \cos\theta_l\right)\right] \end{aligned} \tag{3.C.29}$$

Consider the case that $i = 1$ in one of the terms of $|\mathcal{M}|^2$. Then the $\mathcal{B}-$function can be written as:

$$\begin{aligned} \mathcal{B}_{i=1}(Y_1, Y_2, Y_3, Y_4) =& \frac{y_2 y_3 y_4}{64\pi^5} 4\pi \int \mathrm{d}\lambda \lambda^2 \sum_{j \neq k \neq l} \int \mathrm{d}\theta_j \mathrm{d}\phi_j \sin\theta_j e^{i s_j y_j \lambda \cos\theta_j} \cdot \\ & \cdot \int \mathrm{d}\theta_k \mathrm{d}\phi_k \sin\theta_k e^{i s_k y_k \lambda \cos\theta_k} \int \mathrm{d}\theta_l \mathrm{d}\phi_l \sin\theta_l e^{i s_l y_l \lambda \cos\theta_l} \cdot \\ & \cdot \left[K_1 a m_1 \widetilde{E_j} \cdot \left(\widetilde{E_k}\widetilde{E_l} - y_k y_l \cos\theta_k \cos\theta_l\right) + \right. \\ & \left. + K_2\, a^2 m_1 m_j \left(\widetilde{E_k}\widetilde{E_l} - y_k y_l \cos\theta_k \cos\theta_l\right)\right] \end{aligned} \tag{3.C.30}$$

$$\mathcal{B}_{i=1}(Y_1,Y_2,Y_3,Y_4) = K_1 a m_1 \sum_{j\neq k\neq l}\left[\widetilde{E_j}\widetilde{E_k}\widetilde{E_l}B_1\left(y_j,y_k,y_l\right)+\widetilde{E_j}B_2\left(y_j,y_k,y_l\right)\right]+$$
$$+K_2 a m_1 \sum_{j\neq k\neq l} a m_j\left[\widetilde{E_k}\widetilde{E_l}B_1\left(y_j,y_k,y_l\right)+B_2\left(y_j,y_k,y_l\right)\right], \tag{3.C.31}$$

with $B_1\left(y_j,y_k,y_l\right)$ given by Eq. (3.C.10) and

$$B_2\left(y_j,y_k,y_l\right) = s_k s_l \frac{4y_k y_l}{\pi}\int\frac{\mathrm{d}\lambda}{\lambda}\sin(y_j\lambda)\left[\cos(y_k\lambda)-\frac{\sin(y_k\lambda)}{y_k\lambda}\right]\left[\cos(y_l\lambda)-\frac{\sin(y_l\lambda)}{y_l\lambda}\right]$$
$$=\begin{cases}\frac{1}{2}\left[y_k^2+y_l^2-y_j^2\right], & y_j+y_k\geq y_l \;\&\; y_j+y_l\geq y_k \;\&\; y_k+y_l\geq y_j\\ 0, \quad \text{otherwise}\end{cases} \tag{3.C.32}$$

This procedure can be done for all the other terms in $|\mathcal{M}|^2$. If $j=1$, the result is the same, but with $i\leftrightarrow j$. Note that if $k=1$ or $l=1$, there is no B_2-term in the part with K_2. The collision integral then becomes

$$I_{\text{coll}} = \frac{S}{64\pi^3 g m_1 a^6}\int \mathrm{d}y_2\mathrm{d}y_3\frac{y_2y_3}{\widetilde{E_2}\widetilde{E_3}}\mathcal{B}(Y_1,Y_2,Y_3,Y_4)\cdot$$
$$\cdot\,\theta\left(\left(\left(s_1m_1+s_2\widetilde{E_2}+s_3\widetilde{E_3}\right)^2-a^2m_4^2\right)\left(F[f]\right)\right|_{y_4=y_4^*}, \tag{3.C.33}$$

with $y_4^*=\sqrt{\left(s_1am_1+s_2\widetilde{E_2}+s_3\widetilde{E_3}\right)^2-a^2m_4^2}$.

3.D Low-level code checks

Units correctness is a very important sanity check of the code. In practice not a single programming language supports this out of the box, but many of them allow to implement more or less complete units algebra with compile time or runtime verification. Typically this restricts the structure of the code or is too verbose and has a big runtime performance overhead.

Fortunately, there is a very economic way to gain a reasonable control of the units. It is based on the idea that units can be treated as some free constants. For example,

$$m=\frac{E}{c^2} \tag{3.D.1}$$

$$E=0.511\,\text{MeV} \qquad c=3\cdot10^8\,\text{m/s} \tag{3.D.2}$$

$$m=\frac{0.511\,\text{MeV}}{9\cdot10^{16}\,(\text{m/s})^2}=5.67\cdot10^{-18}\,\frac{\text{MeVs}^2}{\text{m}^2} \tag{3.D.3}$$

Some units are related to each other:

$$\frac{eV}{J} = 1.602 \cdot 10^{-19} \qquad J = \frac{\mathrm{kg\ m^2}}{\mathrm{s^2}} \tag{3.D.4}$$

$$m = 5.67 \cdot 10^{-18} \cdot 1.602 \cdot 10^{-13} \frac{\mathrm{J\ s^2}}{\mathrm{m^2}} \approx 9 \cdot 10^{-31}\,\mathrm{kg} \tag{3.D.5}$$

To employ this idea, one can write a program in the following way:

```
class UNITS:
    # Base units
    eV = 1.653e-2  # arbitrary constants
    m = 5.27e3
    s = 6.24
    # Derived units
    MeV = 1e6 * eV
    J = 6.242e18 * eV
    kg = J * s**2 / m**2

class CONST:
    c = 3e8 * UNITS.m / UNITS.s

energy = 0.511 * UNITS.MeV
electron_mass = energy / CONST.c**2
print "Electron mass is", electron_mass / UNITS.kg, "kg"
```

Figure 3.19: Sample unit-preserving code in Python

Basically, some arbitrary (possibly random) constants should be assigned to all base units, all derived units are defined in terms of base units and all dimensionful quantities must be created with correct unit multipliers. Then, right before the output, one divides quantity by the desired units constants and obtains the correct answer.

As long as units are used properly, change of the base units constants does not affect the answer (except for corner cases of rounding errors). By running the code twice with different definitions of base units verifies the correctness of the code.

The advantages of this approach are:

- Simplicity and zero overhead: for each occurrence of the unit, there is a simple multiplication operation. There are additional constructs in the code and an obvious recipe for proper usage
- Easy opt-out: all units can be set equal to 1 such that they have no way to influence the computation
- Free unit conversion

On the other hand this approach is runtime-only and does not produce any exceptions during the compilation or runs of the code.

This idea is implemented in the **numericalunits**[4] package that we use in conjunction with simpler code suitable for natural units and similar to the listing 3.19 to enforce units handling in our code.

Verification of the matrix elements. Our computations involve tens of different quantum processes for which we have computed the matrix elements – App. 3.B. It is not uncommon to find mistakes in papers with similar computations. We have checked our matrix elements using Mathematica symbolic algebra scripts that automate the necessary transformations for Fermi theory interactions.

[4] https://pypi.python.org/pypi/numericalunits