MSc in Particle Physics

Search for prompt heavy neutral lepton decays into tau leptons with the ATLAS detector

Arnau Morancho Tardà

Supervised by Assoc. Prof. Stefania Xella

May 2022

Arnau Morancho Tardà

Search for prompt heavy neutral lepton decays into tau leptons with the ATLAS detector

MSc in Particle Physics, May 2022
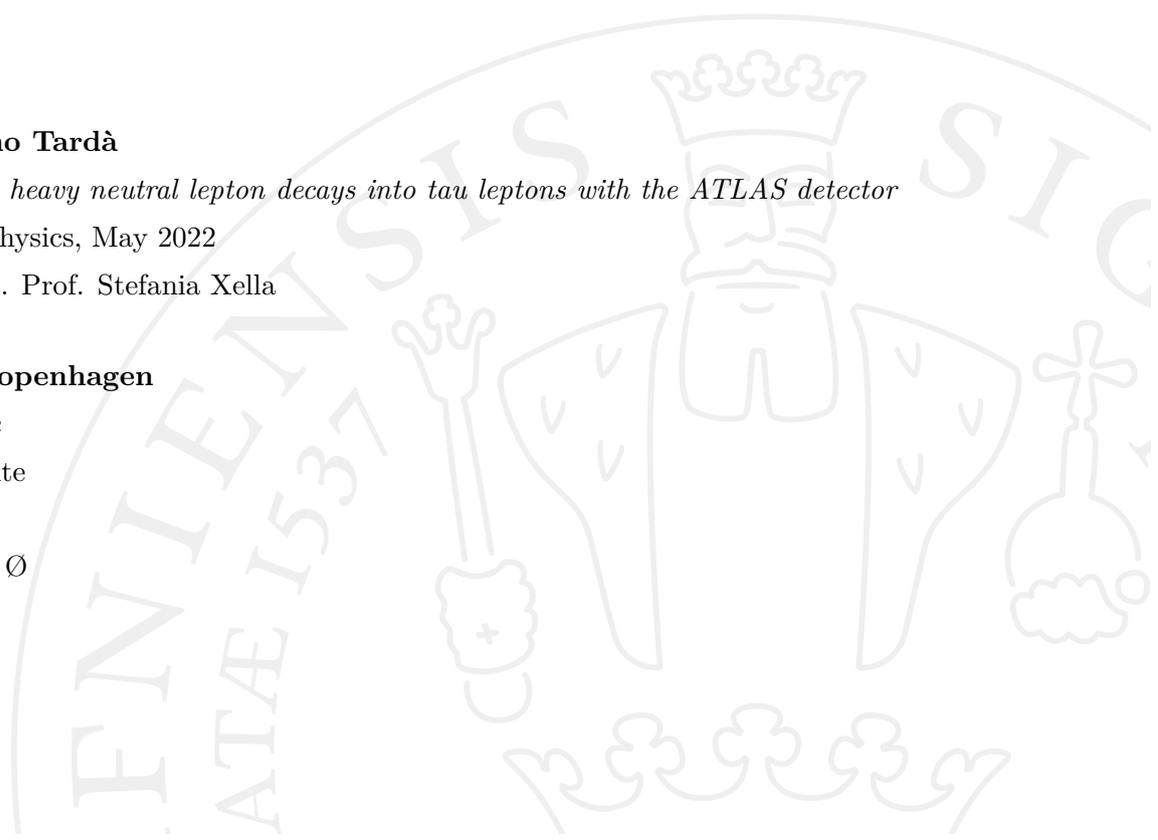
Supervisor: Assoc. Prof. Stefania Xella

University of Copenhagen

Faculty of Science

Niels Bohr Institute

Blegdamsvej 17

2100 Copenhagen Ø

Abstract

According to the Standard Model, neutrinos should not have mass, and therefore neutrino oscillations are not explained. Neutrino Minimal Standard Model is a theory that would explain the mass of the neutrinos (active neutrinos) and their oscillations by the prediction of three unobserved, sterile neutrinos, usually called *heavy neutral leptons*. These sterile neutrinos would, in principle, mix with the neutrinos described by the Standard Model. Some projects are studying the interaction of the heavy neutral lepton with electrons and muons using data from the ATLAS experiment. The channel with tau neutrinos is yet not explored, but it could be possible.

In this thesis, a prompt heavy neutral lepton generated from a W boson manifesting itself in a three lepton final state involving taus is studied using simulated data from the ATLAS collaboration. Getting the sensitivity of ATLAS to this process will tell if the exclusion limits set by other experiments can be extended. It has been found that for a specific channel involving mixing with the muon and the tau flavour, the mixing angles ATLAS is sensitive to are within the exclusion limits of previous searches.

Acknowledgements

Contents

Introduction

The Standard Model (SM) [1, 2] is the culmination of more than a century of research in particle physics and the most accurate description of nature, with all its predictions having been discovered. However, it fails to explain some observed phenomena. Therefore, it is well-known that it cannot be the ultimate model and needs to be extended.

The Neutrino Minimal Standard Model [3, 4] is a promising model that could justify three of the unexplained phenomena of the SM: the neutrino oscillation, the dark matter, and the baryon asymmetry. It introduces three right-handed neutrinos to the already known elementary particles from the Standard Model.

Neutrinos only appeared to be part of parity-violating interactions, with only those with left-handed chirality being detected [5]. Adding a right-handed partner would establish additional symmetry in the fermion region of the elementary particles. These particles could give mass to the left-handed neutrinos by the seesaw mechanism [6] if at least two right-handed neutrinos are very massive compared to the left-handed ones. Therefore in this model, the name heavy neutral leptons (HNL) is used to describe the right-handed neutrinos.

There are currently many efforts placed into detecting a process involving an HNL. In particular, the ATLAS collaboration has two groups studying displaced and prompt HNLs [7]. The prompt HNL tri-lepton final state group studies the production of the HNLs from W bosons with three light leptons as the final state. However, they do not study the possibility of mixing with tau flavour.

This project aims to study the mixing with tau flavour for prompt HNLs within the ATLAS detector, setting some exclusion limits for a specific channel.

Theory

<div style="text-align: right">2</div>

2.1 Standard Model

The Standard Model (SM) [1, 2] of particle physics has been the most successful model in particle physics. It describes how nature works in its more basic constituents unifying three of the four known forces under the same framework introducing twenty five elementary particles, twelve fermions and thirteen bosons. It is based on Quantum Field Theory (QFT) [8], constructing the most general lagrangian from known symmetries of nature.
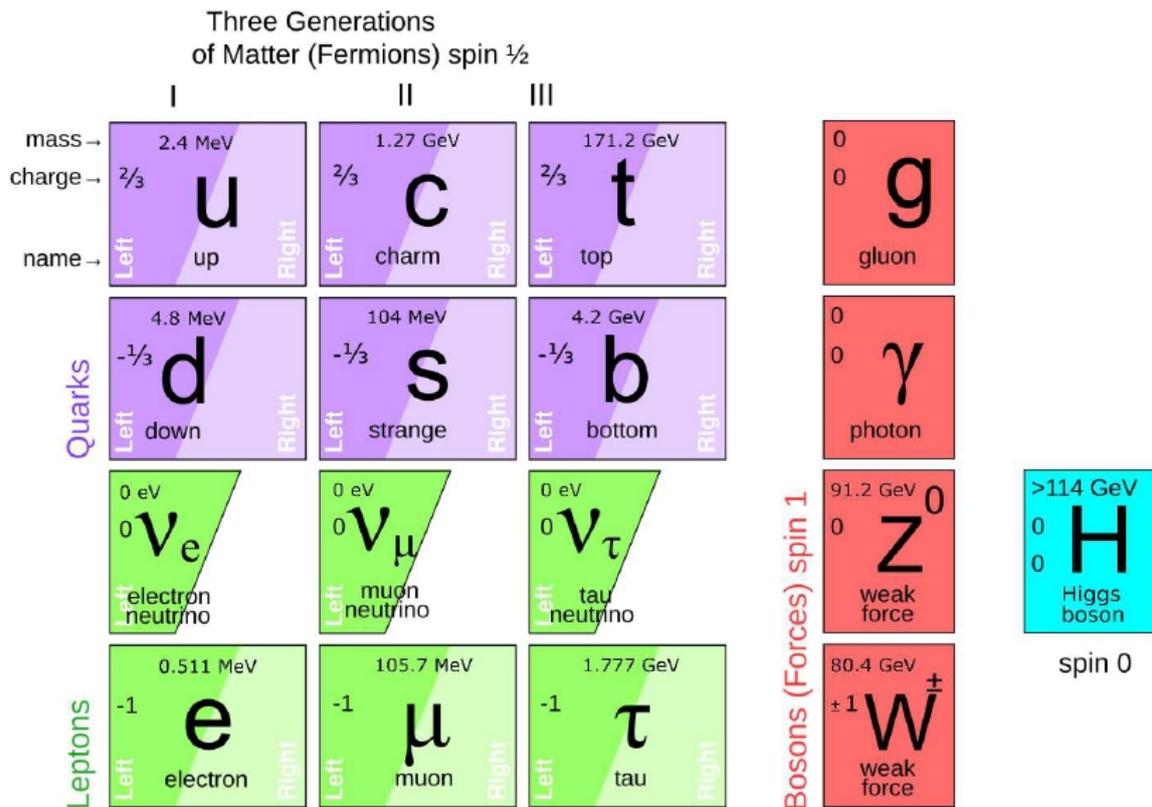

Figure 2.1: Representation of the elementary particles of the Standard Model and their properties [9].

In Fig. 2.1 there is a representation of these elementary particles. The fermions, on the left-hand side represented with purple and green color, are the constituents of the matter, and they have an odd half-integer spin, exactly $\frac{1}{2}$. Fermions interact with each other thanks to the different bosons who act as the mediators of the different interactions, also referenced as forces. Bosons, on the right-hand side and represented with red and blue color, are those particles with an integer spin number. The three forces described by the SM are the electromagnetic, the weak and the strong force. Gravity does not enter this framework, but much effort is placed into reconciling QFT with gravity.

The fermion sector can be divided in two different ways. The first divides it into three generations as a function of its mass hierarchy, the lightest being generation I. Each consists of an up-type quark with electric charge $+\frac{2}{3}$, a down-type quark with electric charge $-\frac{1}{2}$, a neutral lepton called neutrino and a charged lepton with electric charge -1.

Apart from dividing it into three generations based on the mass, it can also be divided into two major groups based on the force they can be affected by. The quarks, in purple in Fig. 2.1, are the only fermions which interact through the strong force with the gluon as the mediator boson. It is described by Quantum Chromodynamics (QCD) [10] and its most distinguishing property is the color confinement. It states that each quark has a color, but they can not be found isolated and with a distinguished color. This makes quarks clump together forming hadrons, a neutral color particle. Hadrons can be composed of two quarks (quark-antiquark) called mesons or three quarks called baryons. This force that clumps together the different quarks is called strong interaction, and it is mediated by the gluons, which are eight color charged bosons with spin 1. The strong interaction has a range of 10^{-15} m. The fermions not affected by the strong force are called leptons (in green).

The electromagnetic interaction affects all the charged fermions by interchanging a photon, a spin 1 boson. This statement includes all quarks, charged leptons and $W^{\pm}$ bosons. The range of this force is infinite, giving the freedom to photons to propagate freely. It is explained by the quantum field theory of the electromagnetic force, also known as Quantum Electrodynamics (QED) [2].

The last of the forces SM contemplates in its frame is the weak interaction, which is mediated by the spin 1 Z and $W^{\pm}$ bosons. It affects all fermions, including neutrinos, and for W boson, it also couples to photons and W itself. It has a range of 10^{-18} m,

and it can be described together with the electromagnetic interaction in the Electroweak (EW) Theory [10].

The Standard Model is described by a non-abelian gauge group, formed by the EW interaction with SU(2)xU(1) group, together with QCD with a SU(3) group, expressed in Yang-Mills theory. The S stands for special group and the U for unitary group.

At last, the Higgs boson with 0 spin is responsible for giving mass to the elementary particles thanks to the Higgs mechanism.

It is essential to clarify that all fermions have a respective antifermion, which has the same mass as the corresponding fermion but opposite electric and magnetic charge. We already mentioned that a meson is a quark and an antiquark glued together. Furthermore, all fermions have right-handed (RH) and left-handed (LH) chiralities except for the neutrinos. The chirality defines if a particle transforms in

Figure 2.2: Schematic representation of the chirality for massless particles.

a right or left representation in the Poincaré group. On the other hand, the helicity is the projection of the spin into the momentum direction [2, 10]. In the case of massless particles, helicity and chirality coincide. This allows us to represent the chirality for the massless case, Fig. 2.2.

All the fermions form a doublet with its generation partner for each RH and LH pair, except for the RH charge leptons who form a singlet. First observed by the parity violation of beta decays [5], no RH neutrino has been observed experimentally since. Later, the SM was built assuming LH neutrinos and RH antineutrinos were the only neutrinos in nature.

Some quantities are conserved in the SM. Of all these, the lepton flavour number conservation L_α will play an important role in this thesis. Each of the leptons has an associated lepton flavour number inside its generation group. For instance, ν_e and e^- have $L_e = +1$, while $\overline{\nu_e}$ and e^+ have $L_e = -1$. While this number is not always conserved in nature, the total lepton number L, which does not distinguish between the different flavours, is conserved in any process detected so far.

The discovery of the Higgs boson in 2012 by ATLAS and CMS, two experiments at CERN, showed the success of the SM almost fifty years after its prediction [11, 12]. Even with predicting many particles and characterizing them before they were discovered, it was already clear in its formulation that the SM could not be the final model explaining all particle physics. As we have already mentioned, gravity, for example, is not embedded in this description. However, some other behaviours in the Universe cannot be explained under the SM and therefore, a deep understanding of the weak points and the phenomena that the SM fails to explain is necessary.

2.2   Beyond the Standard Model

As mentioned in the previous section, the SM is an incomplete theory. These are the phenomena the SM can not explain:

- **Gravity** is well explained in large scales by general relativity. However, if we try to formalize the interaction at small scales using QFT and a hypothesized particle called graviton, we get a non-renormalizable theory [8, 13].

- **Dark energy** manifests itself as a negative repulsive energy and gives rise to the observed accelerated expansion of the Universe observed by, for example, type Ia supernovae [14, 15].

- **Neutrino oscillations** have been experimentally observed when the expected number of neutrinos of each type coming from the sun and the atmosphere was not in agreement with the observations [16]. In order to explain oscillations, neutrinos need to have not only flavour eigenstates but also mass eigenstates as known from quantum mechanics [17], hinting that the neutrinos have a mass.

- **Dark matter** accounts for a significant amount of the matter of our Universe. The necessity of predicting this type of matter comes, for instance, from observing the velocity distributions of starts when orbiting the galactic center. Only their gravitational effects can be observed since they are non interacting with the electromagnetic force, hence their name "dark" [10, 18].

- **Baryon asymmetry** accounts for the dominance of matter over antimatter found in nature. It requires thermal non-equilibrium in the early Universe together with charge-parity (CP) violation [14, 19].

Furthermore, recent experiments that were characterizing already well-known particles found an odd value compared to the prediction by the SM, even though not yet at the 5σ level needed to be considered a discovery. In particular, the most recent ones are the anomalous magnetic moment of the muon [20], the violation of lepton flavour universality on B meson decays [21], and the anomalous mass of the W boson [22]. These slight deviations of the SM give hints on where the model may be wrong, giving tools to the theoreticians to update the model.

2.3 Neutrino oscillations

Neutrino oscillations have been observed for many years now, studying the solar and the atmospheric neutrino flux measurements. Their existence directly suggests that these particles have a mass. From quantum mechanics, flavour states can oscillate if they have in addition mass eigenstates [8].

The two eigenstates are related by the following expression

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} V_{e1} & V_{e2} & V_{e3} \\ V_{\mu 1} & V_{\mu 2} & V_{\mu 3} \\ V_{\tau 1} & V_{\tau 2} & V_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \tag{2.1}$$

where ν_α with $\alpha = e, \mu, \tau$ denotes the flavour eigenstates and ν_i with $i = 1, 2, 3$ the mass eigenstates. The 3x3 matrix V is known as the PMNS matrix and describes the mixing between the different states [23]. One could note that no oscillations would occur if this matrix was diagonal.

Experiments performed in solar and atmospheric neutrinos measured two mass splittings, $\Delta m_{sun}^2 = \Delta m_{21}^2$ and $|\Delta m_{atm}^2| = |\Delta m_{32}^2|$, both being positive. While Δm_{sun}^2 results in $m_1 < m_2$, $|\Delta m_{atm}^2|$ does not distinguish between $m_2 < m_3$ or $m_2 > m_3$. Then, there are two possible hierarchies depending on the order of the neutrino masses. In Fig. 2.3, a representation of the two hierarchies is shown. A normal hierarchy (NH) would have
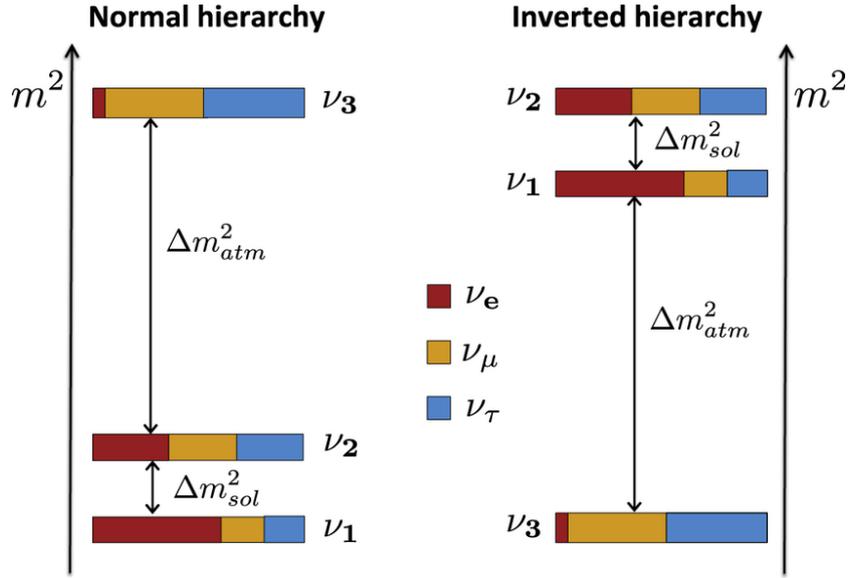
Figure 2.3: Representation of the two possible mass hierarchies based on experimental observations [24]. Each color allows seeing how much each mass eigenstate mixes with the three flavour eigenstates.

$m_1 < m_2 < m_3$, in contrast with the inverted hierarchy (IH) having $m_3 < m_1 < m_2$ [25].

2.4 Introducing right-handed neutrinos

In order to explain the theory required to cover this thesis, we will focus on one of the shortcomings of the SM. As mentioned before, neutrinos oscillations can only be explained if they are massive particles. Let us try to extend the SM so neutrinos can acquire mass.

2.4.1 Dirac mass

The Dirac mass is generated by the Yukawa interaction between massless fermion fields and the Higgs field, which after spontaneous symmetry breaking, becomes a mass term coupling left and right chiral components of the same particle. Its lagrangian term is

$$\mathcal{L}_D = m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L), \tag{2.2}$$

where ψ_L and ψ_R represent the LH and RH states respectively and $\psi^\dagger$ the hermitian conjugate, representing the "antistate".

From Eq. (2.2) it can be observed that to build this term we need four spin states, left and right chiralities for both particle and antiparticle. Since neutrinos in the SM are only LH, they can not have a Dirac mass term in the lagrangian. This mechanism allows the rest of the fermions to be massive.

Dirac particles have a clear distinction between their particles and antiparticles, conserving the number of fermions.

2.4.2 Majorana mass

Majorana fermions are characterized by being their own antiparticle, $\psi_M = \psi_M^c$ with ψ_M^c being the charge conjugate. They are only characterized by their RH and LH pair, needing only two spin states. If this particle was a lepton for example, it would violate lepton number conservation since the same state would be created twice during pair production.

Its lagrangian can be constructed as

$$\mathcal{L}_M = M(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L) = M(\psi_L^\dagger \psi_L^c + (\psi_L^\dagger)^c \psi_L). \tag{2.3}$$

Using that $\psi_L^c = \hat{C}\psi_L^* = -i\gamma^2\psi_L^* = \begin{pmatrix} 0 \\ i\sigma_2\psi_L^* \end{pmatrix}$, with the Majorana representation of the gamma matrix $\gamma^2 = \begin{pmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{pmatrix}$ and the Pauli matrix $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, we can rewrite it as

$$\mathcal{L}_M = M(\psi_L^\dagger \sigma_2 \psi_L^* - \psi_L^T \sigma_2 \psi_L) = M\overline{\psi_M}\psi_M, \tag{2.4}$$

where ψ_L^* is the complex conjugate of a state and ψ_L^T is the transposed. We have also defined $\psi_M = \begin{pmatrix} \psi_L \\ i\sigma_2\psi_L^* \end{pmatrix}$, and $\overline{\psi_M} = \psi_M^\dagger \gamma^0$ as the Dirac conjugate with $\gamma^0 = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}$ in the Majorana representation.

It is unknown which is the nature of neutrinos, respect being Dirac or Majorana. Moreover, we know that weak interaction only couples to the LH neutrinos and RH antineutrinos, as they are massless in the SM.

2.4.3 Right-handed neutrinos

Let us introduce RH neutrinos in the lagrangian giving the freedom to be Dirac or Majorana fermions, restoring some symmetry between the quark and lepton sectors as seen in Fig. 2.1. We could also consider a model with only Dirac masses, but then the Yukawa coupling of the neutrinos should be much smaller than the rest of the couplings of the SM in order to explain the small masses, something not very intuitive. Therefore we choose to introduce also a Majorana term to see if there is another mechanism that could explain neutrino mass.

From electroweak theory, the LH neutrino is part of a SU(2) gauge symmetry forming a doublet with the charged lepton, so therefore when we build its corresponding Majorana mass, it needs to be gauge invariant as well. Building the Majorana term of the lagrangian as described before gives

$$\mathcal{L}_{M_{LH}} = \frac{M_L}{2}\bar{\lambda}\lambda \qquad \text{with} \qquad \lambda = \begin{pmatrix} \nu_L \\ i\sigma_2\nu_L^* \end{pmatrix}. \tag{2.5}$$

This is not gauge invariant, but we can generalize it using the Higgs field $\tilde{H} = \frac{1}{2}\begin{pmatrix} v \\ 0 \end{pmatrix}$ and the lepton doublet $L = \begin{pmatrix} \nu_L \\ e \end{pmatrix}$. The LH neutrino term can be made gauge invariant using $\nu_\beta = (L_\beta \cdot \tilde{H})$ with β running through the three different flavours e, μ and τ. Having now the following expression

$$\mathcal{L}_{M_{LH}} = m_{\alpha\beta}\frac{\overline{(L_\alpha \cdot \tilde{H})^c}(L_\beta \cdot \tilde{H})}{v}. \tag{2.6}$$

Despite having the gauge symmetry we asked for, this expression has a dimension 5 operator, which is incompatible with a renormalizable theory. Hence there can not be any LH Majorana term in our lagrangian compatible with the SU(2) group.

Building then the following spinors

$$\text{Dirac: } \psi \equiv \begin{pmatrix} \nu_L \\ N_R \end{pmatrix}; \qquad \text{Majorana: } \chi \equiv \begin{pmatrix} -i\sigma_2 N_R^* \\ N_R \end{pmatrix}, \tag{2.7}$$

the Dirac and Majorana mass term in the lagrangian of a one flavour model would look like

$$\mathcal{L}_{DM} = m_D \bar{\psi}\psi + \frac{M_M}{2}\bar{\chi}\chi = \frac{1}{2}\begin{pmatrix} -i\sigma_2\nu_L^* \\ N_R \end{pmatrix}^\dagger \begin{pmatrix} 0 & m_D \\ m_D & M_M \end{pmatrix}\begin{pmatrix} \nu_L \\ i\sigma_2 N_R^* \end{pmatrix} + h.c., \qquad (2.8)$$

where h.c. stands for the hermitian conjugate, m_D the Dirac mass, and M_M the Majorana mass of the RH neutrino. In the last equality we have written the expression in its matrix form with a mass matrix $M_{DM} = \begin{pmatrix} 0 & m_D \\ m_D & M_M \end{pmatrix}$.

This can be generalized to three LH neutrinos and $\mathcal{N}$ RH neutrinos The matrix M_M will have dimensions $\mathcal{N} \times \mathcal{N}$ and M_{DM} $(3+\mathcal{N}) \times (3+\mathcal{N})$, including the three flavours states. Renaming the state we can rewrite the lagrangian as [26]

$$\mathcal{L}_{DM} = -\frac{1}{2}\begin{pmatrix} \nu_L^* \\ N_R^* \end{pmatrix}^\dagger \begin{pmatrix} 0 & m_D^T \\ m_D & M_M \end{pmatrix}\begin{pmatrix} \nu_L \\ N_R \end{pmatrix} + h.c., \qquad (2.9)$$

The matrix $M_{DM} = \begin{pmatrix} 0 & m_D^T \\ m_D & M_M \end{pmatrix}$ can be diagonalized using

$$V^{\nu^T} M_{DM} V^\nu = diag(m_i...) \qquad i = 1, ..., 3+\mathcal{N}, \qquad (2.10)$$

with $\begin{pmatrix} \nu_L \\ N_R \end{pmatrix} = V^\nu \nu$, where ν are the different mass eigenstates. The lagrangian can be rewritten as

$$\mathcal{L}_{DM} = -\frac{m_i}{2}(\nu_i\nu_i + \nu_i^\dagger\nu_i^\dagger). \qquad (2.11)$$

The LH neutrinos seem to have very small masses compared to the rest of the fundamental particles. This smallness can occur naturally if the Majorana mass of the RH neutrino is much larger than the Dirac mass $M_M \gg m_D$. This is known as the seesaw mechanism [6]. The light neutrino interacts with matter, whereas the heavy neutrino is effectively

decoupled from matter. In this limit, the flavour eigenstates ν_{RI} and $\nu_{L\alpha}$ can be approximated as follows with $\alpha, \beta = 1, ..., 3$ and $I, J = 1, ..., \mathcal{N}$ [27]

$$\nu_{RI} \cong N_I \tag{2.12}$$

$$\nu_{L\alpha} \cong V_{\alpha i}^{PMNS} \nu_i + U_{\alpha I} N_I \tag{2.13}$$

$$U_{\alpha I} \cong \frac{(m_D)_{\alpha I}}{M_I}, \tag{2.14}$$

where $V_{\alpha i}^{PMNS}$ is the PMNS matrix which relates the LH neutrino oscillations, and $U_{\alpha I}$ represents the mixing between RH-LH neutrinos (also referred to as "mixing angles"). $U_{\alpha I}$ could be understood as an extension of $V_{\alpha i}^{PMNS}$ for the RH neutrinos. Here, we have used the notation $(M_{DM})_{\alpha I} = (m_D)_{\alpha I}$, to make implicit that this term is the Dirac mass matrix.

For the mass sub-matrices we get

$$m_{\alpha\beta} \cong - \sum_{I=1,...,\mathcal{N}} \frac{(m_D)_{\alpha I}(m_D)_{\beta I}}{M_I} = - \sum_I M_I U_{\alpha I} U_{\beta I} \tag{2.15}$$

$$m_{IJ} \cong M_I \delta_{IJ}, \tag{2.16}$$

where Eq. 2.14 has been used in Eq. 2.15, and $m_{\alpha\beta}$ represents the light part of the mass matrix while m_{IJ} the heavy one.

From these expressions, we can state some essential relations. From Eq. 2.16, the RH neutrino masses m_{IJ} is approximately the Majorana mass and its value does not constraint the mixing angles $U_{\alpha I}$, only the Dirac and Majorana mass affect those (eq. 2.14). In Eq. 2.15, for $M_M \gg m_D$ the mass of the LH neutrino will be very small, or equivalently the mixing angles need to be very small to allow small LH neutrino masses.

Since the LH neutrino is part of a SU(2) group forming a doublet, we will call it an active neutrino in opposition to the RH neutrino, which forms a singlet and will be called a sterile neutrino. Furthermore, since we are approximating the Majorana mass as being much more massive than the Dirac mass in this model, we will call the RH neutrino by heavy neutral lepton (HNL).

2.5 The Neutrino Minimal Standard Model

The Neutrino Minimal Standard Model (νMSM) [3, 4] introduces three RH neutrinos in order to explain neutrino oscillations while proposing a candidate for the dark matter and offering an explanation for the baryon asymmetry (leptogenesis). Introducing two would explain the two observed mass splittings, but we need at least three for giving mass to the three LH neutrinos and get a candidate for DM. Introducing these amounts of RH neutrinos is intuitive with the number of fermionic generations that the SM accounts for, Fig. 2.4.
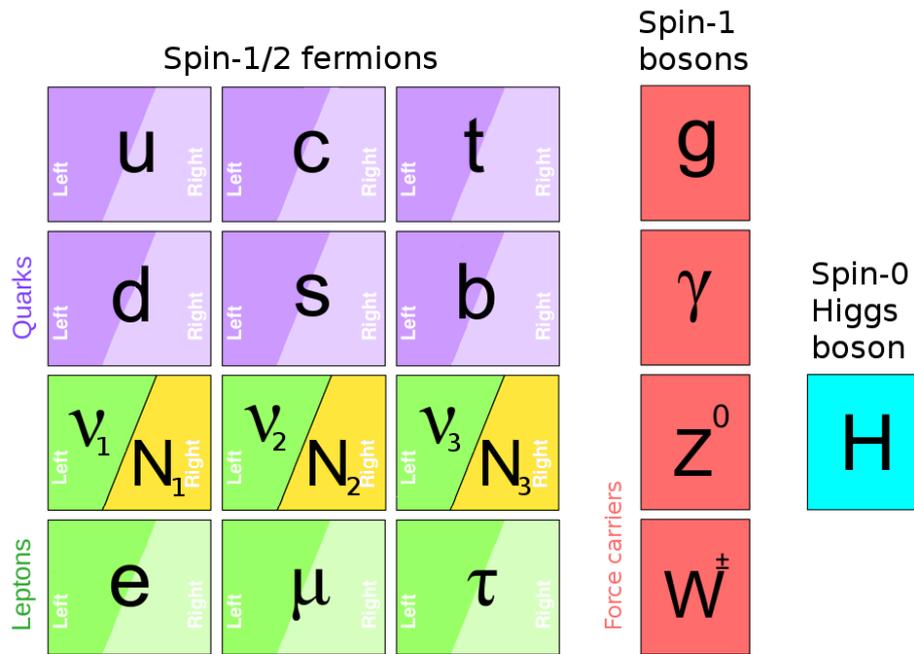
Figure 2.4: Schematic representation of Fig. 2.1, introducing three right-handed neutrinos in yellow [9].

The lightest of the HNLs, N_1 would be the DM candidate having a mass at the *keV* scale, whereas the other two $N_{2,3}$ would have a much greater value of the mass under the electroweak scale and would explain the two splitting of the active neutrino masses. In addition, provided that they have a similar mass, they could explain the baryon asymmetry.

The two heavier HNLs would form a quasi-Dirac pair, being almost degenerate in mass, $M_3 = M_2 + \delta M$ and with approximate same mixing angles, $U_{\alpha 3} \approx \pm i U_{\alpha 2}$. This pair could give rise to two different scenarios in function of the decay width of the HNL and their separation in mass δM [26].

- $\Gamma \gg \delta M$: HNL pair behaves like a single Dirac HNL

- $\Gamma \ll \delta M$: HNL pair behaves like a single Majorana HNL

A Dirac HNL has two propagators that lead to total lepton number conservation (LNC). In comparison, a single Majorana HNL has four propagators, two with total lepton number violation (LNV) processes and two more leading to LNC. The observation of LNV processes would constrain our model to Majorana HNLs, while detecting an HNL from a process with LNC would not allow to restrict the nature of HNLs.

In Fig. 2.5, an example of the two possibilities LNC/LNV is represented. We will stick with the notation that for every vertex a fermionic line enters, another fermionic line must go out. Then, the propagators including only one arrow, directly shows that the process is LNC, whereas two arrows would lead to LNV. If the arrows look at each other $\Delta L = -2$, and if the arrows look in opposite directions it gives $\Delta L = +2$.

(a)

(b)

(c)

(d)
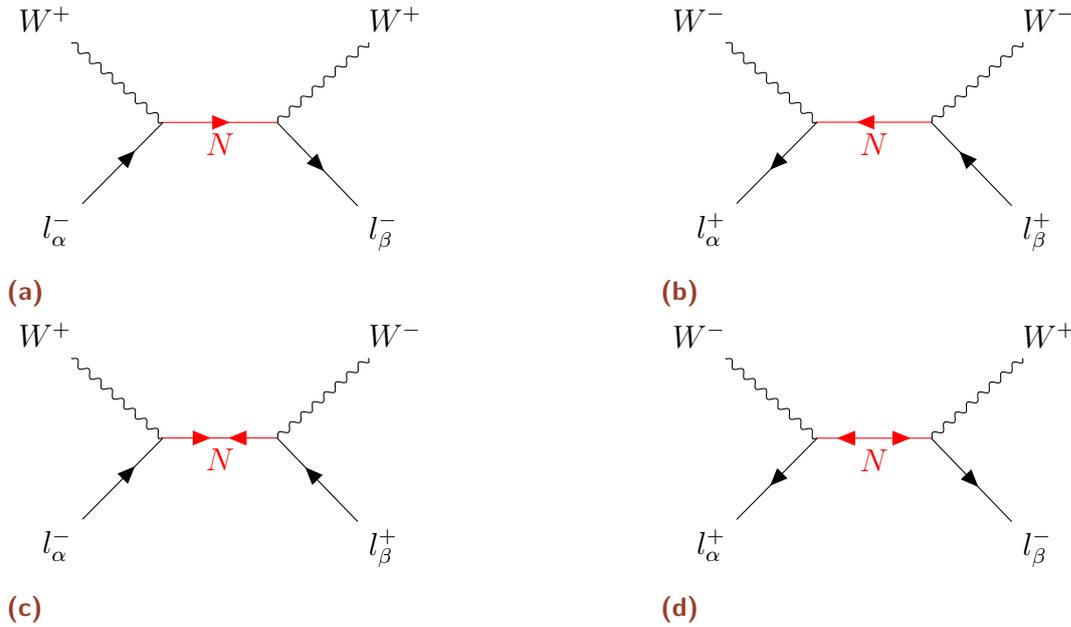
Figure 2.5: The four different propagators for a Majorana HNL, N. On the other hand, (a) and (b) represent the two propagators of a Dirac HNL. In all cases α and β can individually be $\alpha, \beta = \mu, \tau, e$.

In this model, processes that only mix with the same lepton flavour would not explain any of the problems BSM. Therefore it is interesting to study HNLs mixing with two different flavour states.

The search of HNLs at the ATLAS experiment at the Large Hadron Collider is the topic of this thesis, particularly those cases when N mixes with tau flavour neutrinos.

LHC and ATLAS

<div style="text-align: right">3</div>

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world highest energy particle collider. Built by the *Conseil européene pour la recherche nucléaire* (CERN) and located between the french and swiss border. It consists of a circumference of 27 kilometers ring shaped containing superconducting magnets. These magnets bend the trajectory of the particles inside the accelerator, while the different accelerators using interleaved electromagnetic fields accelerate the particles up to 99.999999% the speed of light.

The LHC collides two particle beams coming from both sides of the ring (head-to-head collision). These particle beams can consist of either protons or heavy ions. The beams are accelerated and collimated until they are made to collide at one of the four major detectors spread around the ring. Each of these detectors is linked to an experiment with different research topics. These are CMS, ALICE, LHCb and ATLAS.

CMS stands for *Compact Muon Solenoid*, and it is a general-purpose detector with a wide range of interests, including Higgs boson characterization, extra dimensions studies and search for dark matter candidates. ALICE, known as *A Large Ion Collider Experiment*, collides heavy-ions reproducing the primordial quark-gluon plasma for a better understanding of quantum chromodynamics (QCD) and color confinement. LHCb stands for *Large Hadron Collider Beauty*. It focuses on measuring parameters of CP violation and trying to explain the baryon asymmetry.

This project studies a specific process using the ATLAS detector, and therefore it will be explained more thoroughly in the following section.

The LHC is operative during a certain period of time called runs, where collisions are produced. Between different runs, the accelerator and detector are upgraded. LHC Run-2 ended in 2018 after reaching an energy of $\sqrt{s} = 13$ TeV for p-p collisions in the

center of mass frame, with an integrated luminosity collected from ATLAS and CMS of $\mathcal{L} = 139 \ fb^{-1}$ [28, 29].

With Run-3 start date expected in May 2022 and the next stage of LHC with the High Luminosity upgrade (HL-LHC) in 2029, the integrated luminosity will reach up to $\mathcal{L} = 3000 \ fb^{-1}$ [28].

3.2 ATLAS

A Toroidal LHC AparatuS, also known as ATLAS [30], is the largest general-purpose detector in LHC. It is designed to measure the broadest possible range of signals with the goal of being able to detect the properties of new and upcoming phenomena. It is designed based on the same scientific goals as the CMS experiment, including search and characterization of the Higgs boson, extra dimensions studies and searches for dark matter candidates. However, it uses different technical solutions and a different magnet-system design.

ATLAS is located 100 m below the ground, with a toroidal shape being 46 m long and 25 m in diameter. The detector consists of six detecting subsystems that allow reconstructing the particle trajectories, record their momenta, energies, lifetimes and charges, making it possible to identify them as individual particles. With this concentric many-layer design, each part is designed to observe specific types of particles.

Over a billion particle interactions take place every second in ATLAS. Because it would not be possible to collect information about every event, a system called trigger selects the events more likely to be potentially interesting according to their properties for further storage.

In July of 2012, ATLAS and CMS announced that the Higgs boson had been independently observed with a mass around 125 GeV [11, 12], achieving one of the goals of the experiment which was to find the only SM particle not yet observed.

More than 5500 scientists form the ATLAS collaboration from 245 institutes in 42 countries.

3.2.1 Coordinate system

The coordinate system and nomenclature will be introduced in this section using the technical reports found in [30–32].

The nominal interaction point is defined as the origin of the coordinate system. The z-axis is defined along the beam direction, and the x-y plane is transverse to the beam. The positive x-axis points to the centre of the LHC ring and the positive y-axis points upward. The polar angle θ is defined as the angle from the beam axis and the azimuthal ϕ the angle around the beam axis. See Fig. 3.1
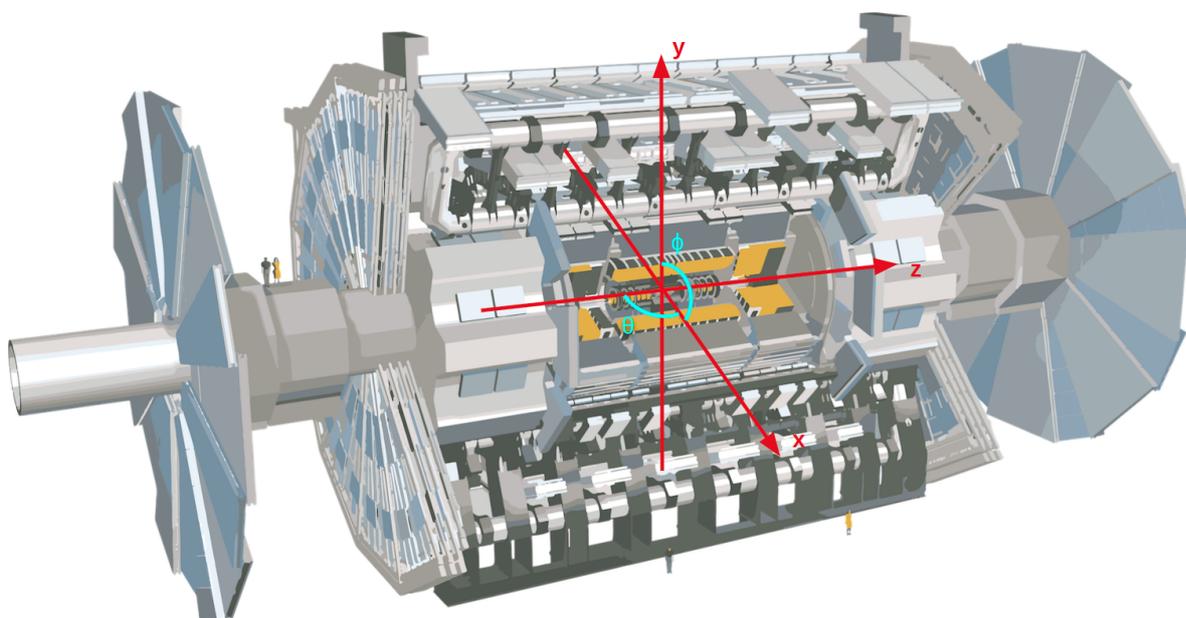
Figure 3.1: Representation of ATLAS detector with the coordinate system [33].

The transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$ is the momentum in the x-y plane. In collider physics, the pseudorapidity is commonly used instead of the polar angle, and it is defined as $\eta = -\ln\tan(\theta/2)$. An angle of $\theta = \pi/2$ would be $\eta = 0$, and $\theta = 0$ would be $\eta = \infty$. Due to its cylindrical geometry, there is a limit on how great the pseudorapidity of particles can have in order to be detected. The part of the detector which restricts it more has a coverage of $|\eta| < 2.5$. To compute the distance in angle, we use the angular distance between two particles $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.2.2 Subdetectors

ATLAS subdetectors are introduced in this part to understand where and how the particles are detected, depending on the particle properties [30–32].

The **Inner Detector** (ID) is very compact and sensitive, beginning just 3.3 cm from the beam axis and extending to a radius of 1.25 m, with a 6.2 m in length and coverage of $|\eta| < 2.5$. It measures the direction, momentum and charge of electrically-charged particles by their interaction with the material. The whole ID is immersed in a $2T$ magnetic field, which deflects charged particles. From the track, both the momentum and charge can be computed.

The ID is composed of the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The Pixel Detector consists of four layers of silicon pixels. When charged particles go through this part, they leave behind energy deposits in these pixels with a precision of 10 μm. The SCT surrounds the Pixel Detector, and it is built with narrow silicon strips rather than small pixels, still with the purpose of keeping track of the charge and direction of electrically-charge particles. The TRT is made up of thin tubes (straws) filled with a xenon-based gas mixture that becomes ionized when a charged particle interacts with it. Helped by the transition radiation we can distinguish between, for example, electrons and pions.

The **Calorimeter** is designed to absorb most of the particles that come from the collision, with layers of a dense absorbing material interspaced with a medium that measures the energy. Calorimeters would only fail to absorb muons and neutrinos, while they provide a good detection of electromagnetic and hadronic showers thanks to the two calorimeters types described below.

The Liquid Argon (LAr) calorimeter measures the energy of electrons, photons and hadrons with a coverage of $|\eta| < 3.2$. It consists of interspaced layers of an absorbent (usually tungsten, copper or lead) with cells with LAr. When the incoming particle interacts with the absorber, a shower of lower energy particles is produced. The energies of those particles are measured in the next layer of LAr when the particles ionize the medium. Combining the information of all layers, the energy of the original particle can be reconstructed. The Tile Hadronic calorimeter surrounds the LAr calorimeter. It measures the energy of those hadrons which could not deposit all of their energy in the inner calorimeter with a coverage of $|\eta| < 4.9$. In this case, we have successive layers of

steel as the absorber layer and plastic scintillators. Like before, the steel layer generates a shower of new particles. When going through the scintillators, they produce photons that are recollected and can be reconstructed according to their intensity.

The **Muon Spectrometer** (MS) curves the muons tracks with large superconducting air-core toroid magnets. It has an overall pseudorapidity coverage of $|\eta| < 2.7$, and it has two differentiated parts, a barrel toroid placed around the calorimeter and two smaller end-cap magnets placed at both ends of the barrel toroid.

The Monitored Drift Tubes (MDT) measure the deflection in most of the eta coverage, while the Cathode Strip Chambers (CSC) are used at the ends of the detector. These two components give a high-precision measurement of the track coordinates. The Resistive Plate Chambers (RPC) are placed in the barrel similarly to the MDT, whereas the Thin Gap Chambers (TGP) are in the end-cap regions. These components help cross-check the identification of the MDT and CSC and provide a precise measurement of p_T used in the trigger system.

3.2.3 Triggers

Every second there is over a billion particle interactions in the ATLAS detector. All these interactions and events can not be saved due to stored capacity. Only one in every one million collisions is tagged as potentially interesting and saved for further analysis.

The trigger and data acquisition system is responsible for tagging and saving the most interesting collisions [30, 34]. In Fig. 3.2, a schema of these parts is shown. In green, the trigger is represented and in blue, the data acquisition system. Next, we will explain how these parts work.

The Level 1 (L1) hardware trigger uses information from the calorimeters and MS. It searches for high transverse-momenta muons, electrons, photons, jets and hadronic taus (taus decaying hadronically), as well as large missing and total transversed energy. It communicates its decision of accepting the event, hence reading it out, or not to the data acquisition system. There can be up to 256 configurations of thresholds or conditions in L1 for the different particles. In less than 2.5 μs, the L1 decides if an event is interesting and is sent to the high level trigger, adding information about which are regions-of-interest (RoI). With an event rate of 40 MHz, L1 can accept up to 100kHz, which will be sent to the high level trigger.
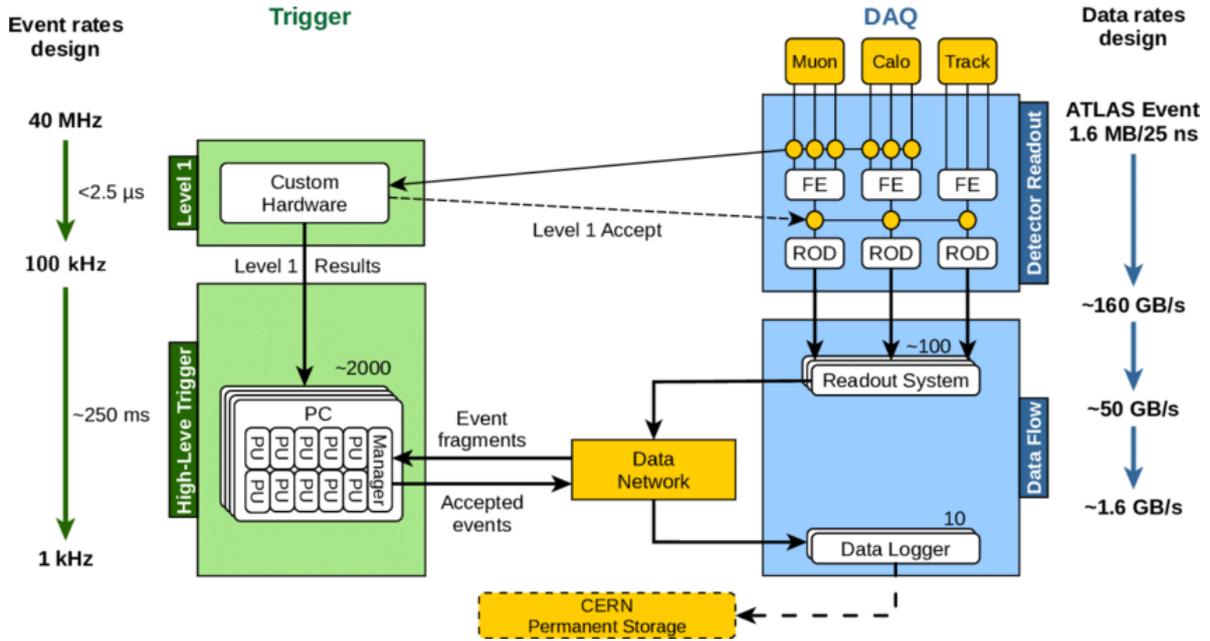
Figure 3.2: Schema of the ATLAS trigger system (in green) together with the data acquisition system (in blue) for the Run-2 [35].

The High Level Trigger (HLT) is software-based and selects the events mainly based on RoI given by L1. It uses a large farm of about 40,000 CPU cores for analysis within an average event of four seconds. The HLT algorithms allow obtaining a higher purity than L1 selection, by adding information of the tracks and combining it with more optimal calibrations of the other subdetectors. Entering 100kHz, the HLT is able to reduce this value down to 1 kHz in order to store the selected events.

During this thesis three different trigger thresholds have been used based on offline and HLT selection of the p_T [36]: single isolated μ with $p_T > 26$ GeV, single isolated e with $p_T > 26$ GeV and two hadronic τ with $p_T > 40, 30$ GeV, in addition of $|\eta| < 2.5$ requirement.

3.2.4 Identification and reconstruction

There are many processes that can give a fake signal in the detector (background). Some algorithms are used to identify the different particles better. They have different working points (WPs) that can be used on each analysis interest depending on the level of accuracy needed.

Muons

Muons are identified in the ID and the MS, with a required $p_T > 5$ GeV and pseudo-rapidity coverage of $|\eta| < 2.5$ [37]. Muons coming from light hadrons are considered background, while muons coming from W or Z bosons are part of signal. There are four different algorithms used in muon identification: combined muons (CB), segmented-tagged muons (ST), calorimeter-tagged muons (CT), and extrapolated muons (ME) [38], that combine information of the ID and MS. Three WPs are defined by combining different groups of the four different algorithms. The *Loose* WP uses all types of muons, while the *Medium* WP uses only CB and ME muons. At last, the *Tight* WP uses CB and ME muons again by maximizing selection purity.

Electrons and photons

Electrons are identified in the ID and the calorimeter, with its background originating from misidentified hadrons, non-isolated electrons from heavy-flavour decays and electrons from photon conversions [39]. Electrons are identified by a likelihood discriminant method which provides three WPs each defined as a subset of the other (*Tight* $\subset$ *Medium* $\subset$ *Loose*) [40]. The likelihood discriminant is based on giving a discriminant value to the electron candidates, having a peak at 0 for background and at 1 for signal. After an inverse sigmoid transformation, the range is more spread and for each WP, a value of the transformed discriminant is chosen. If the discriminant is larger than this value, the candidate will be considered an electron.

Identification of prompt photons, photons not originating from hadron decays, is based on the information provided by the ID and the calorimeter. They are identified from the background photons using the shape and properties of the associated electromagnetic showers and requiring them to be isolated from other particles ($\Delta R < 0.2$). Using these discriminating variables, two WPs are provided, the *Loose* and *Tight* identification [41].

Hadronic taus

Taus are the only leptons that can decay into hadrons, doing so 65% of the times [42]. Apart from hadrons, the hadronic taus τ_h will produce a tauonic neutrino that will escape from the detector without being detected. The reconstructed object will be called

visible hadronic tau $\tau_{had-vis}$ since it does not contain the neutrino. The quarks and gluons interactions can produce jets that are misidentified as hadronic taus, being the source of background.

The $\tau_{had-vis}$ reconstruction algorithms start with a jet detected in the calorimeter requiring $p_T > 10$ GeV and $|\eta| < 2.5$ [43, 44]. A signal cone with $\Delta R = 0.2$ and an isolation cone of $\Delta R = 0.4$ is defined around the jet center to limit the jets that a recurrent neural network (RNN) will study. The RNN is a machine learning algorithm trained to classify the hadronic taus with a better performance. This algorithm gives four WPs based on the background rejection and signal efficiency (*VeryLoose*, *Loose*, *Medium* and *Tight*).

Jets

Jet reconstruction uses topo-clusters formed from seed cells of the calorimeter. The jets can be divided based on their radius, with $R = 0.4$ for small-R jets and $R = 1.0$ for large-R jets. An algorithm called anti-k_t [45] is used for identifying jets. For large-R jets an additional algorithm is used for grooming the jet, throwing away any constituents the algorithm considers not to be part of the jet [46].

Research goals

This study focuses on processes that generate a heavy neutral lepton (HNL) from an on-shell W boson in proton collisions, and is detected in the ATLAS detector. The boson decays into a lepton (l1) and an HNL, which later decays into another lepton (l2) and an off-shell W, decaying into a lepton (l3) and its corresponding neutrino (ν_{l3}), see Fig. 4.1. We will focus on prompt HNLs, which means that their lifetime is short and they decay shortly after their production vertex. This allows the reconstruction of their decay products in the ID. For this analysis, the upper limit for the HNL lifetime has been set at $c\tau_N = 0.1$ mm.

Overall this process is called HNL prompt tri-lepton channel. For the first time in ATLAS, one of the three leptons is considered to be an hadronic tau. More precisely, we will study cases where the tau will be placed in l1 or l2 positions to study a mixing angle with tau, $U_{\tau I}$.

(a)   (b)
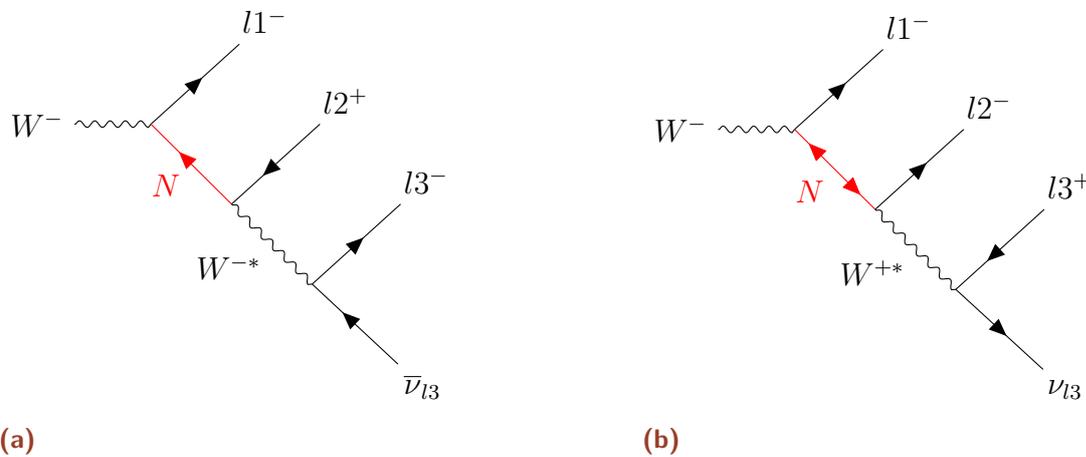
Figure 4.1: General Feynman diagram of the process studied with the HNL propagator N in red. On the left, there is total lepton number conservation (LNC) with $\Delta L = 0$, in contrast with total lepton number violation (LNV) on the right, $\Delta L = +2$. The star on the W boson stands for of-shell W.

Fig. 4.1 shows how the HNL is created, and the process is sensitive to that a ν_{l1} is created in the decay vertex of the on-shell W, that mixes or "oscillates" with N, that later mixes with ν_{l2}.

The research consists of two parts. First, a Monte Carlo simulation of all the possible channels [1], validating the simulation and studying the preselection cuts and triggers is done, to pick which of the channels is more plausible to be observed. For the most promising one, a thorough study on the kinematics will be performed to preselect the signal in both LNC and LNV cases. In the process, we will inevitably be picking background, events that are seen for the detector as signal. We will quantify how visible the signal is compared to the background for a given integrated luminosity using machine learning.

4.1 Current status search of HNL

It is clear that if the HNL has a large mixing angle, it would be common to have processes involving these particles and therefore generating them. For a specific analysis, a sensitivity study can be performed, where the exclusion bounds on the mixing angle squared per mass are computed. In Fig. 4.2, we show the exclusion limits of the mixing angle squared with muon and tau separately. The exclusion limit is the area painted, and it tells us that no HNL has been found in that part of the parameter space. If HNL exists, it needs to be outside the painted area.

It can be observed that whereas mixing angles with light leptons (in this case only μ is shown, yellow line in Fig. 4.2a [7]) have been already studied in the ATLAS experiment, no analysis involving tau has been done yet. For displaced HNLs, a group is currently studying the possibility of detecting these particles mixing with the three flavours in the HL-LHC (green line in Fig. 4.2 [47]). Having displaced HNLs allows to reduce the background having better sensitivity and therefore reaching a small mixing angle, but it comprises smaller m_N that what we intend to study.

This motivates our project to look at the current sensitivity of the ATLAS detector for prompt HNLs mixing with tau flavour and compare it to the previous DELPHI limit (purple line in Fig. 4.2) and the displaced HNL in the future ATLAS runs.

---

[1]Combination of leptons in the different positions l1, l2 and l3
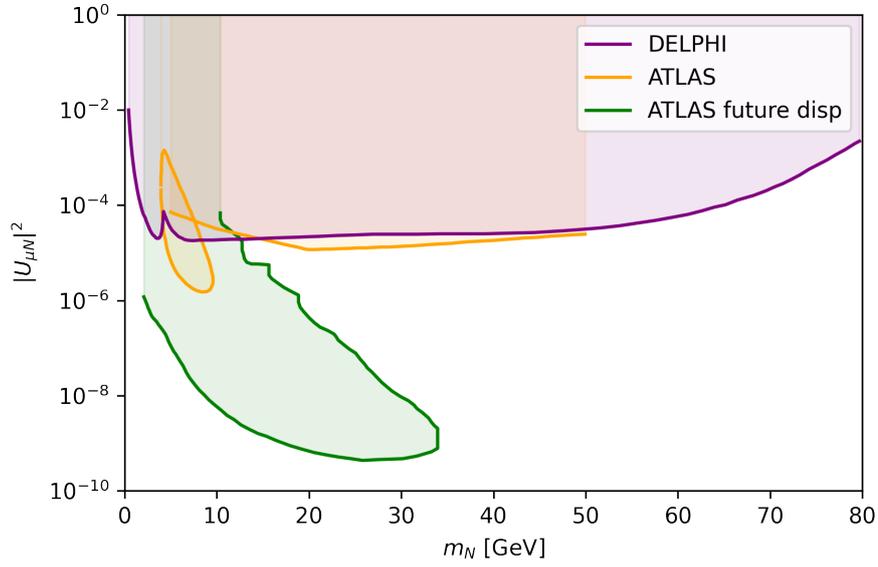
(a)
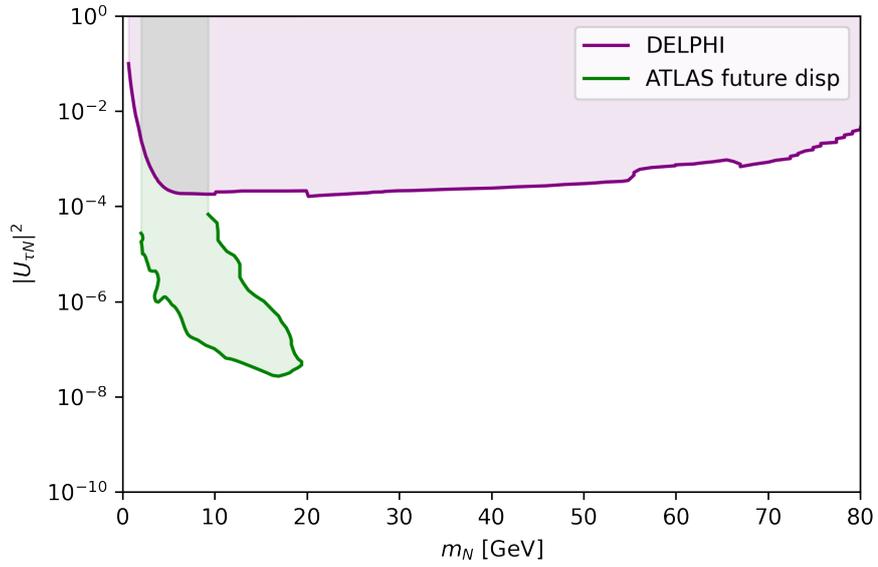
(b)

Figure 4.2: Exclusion limits on the mixing angle squared in function of the HNL mass, for μ flavour in (a), and τ in (b). In purple there are the previous constraints set by DELPHI [48], in yellow the current studies of HNL in ATLAS [7] and in green the predicted future limits for displaced HNL in ATLAS are shown [47].

Simulation

<div style="text-align: right">5</div>

At the time of starting the project, no simulation with the processes we are interested in was available. In order to simulate the events, we first need to generate a Monte Carlo sample containing the relevant physics signal we want to study. A description of the different software used is explained ahead.

5.1 MadGraph

MadGraph (MG) is a Monte Carlo event generator for studying proton-proton collisions. MadGraph5_aMC@NLO framework provides the elements to study SM and BSM phenomenology [49].

MG uses Standard Model by default, but one can download and use other models including new physics. Each model corresponds to a number of Feynman rules which include the physics of the model. For this simulation the HeavyN model [50, 51] has been used, specifically the SM_HeavyN_Gen3Mass_NLO model, which includes the production and decay of three HNLs with Majorana mass, assuming massive bottom quark and tau lepton and a diagonal CKM (no mixing between different quark flavours).

The information of the generation is set in the cards, being the parameter and run cards the most important ones. On the parameter card there is information about the physics, like the mass and decay width of the particles, the mixing angle $U_{\alpha I}$ of the HNLs with each of the three flavours and some physical constants. The run card fixes information about the run itself: the number of events requested, the energy on the beams and some preliminary cuts on the p_T and η.

5.2 Pythia 8

Pythia 8 is a program for the generation of high-energy physics collision events [52]. It can also be used as an interface with MG. We will use it to include parton showers, hadronization and tau decay in our events. The partons are quarks and gluons that emit further gluons or produce quark-antiquark forming a parton shower. When the strong interaction rises, hadronization takes place where the partons are bound into colourless hadrons, Pythia 8 handles tau decay taking into account the spin correlations of the hadronic currents, based on the prior work in Tauola and Herwig++ [53].

5.3 GEANT4

GEANT4, GEometry ANd Tracking, is a toolkit that simulates the passage of particles through matter using Monte Carlo methods [54–56]. It will be useful because it simulates how the detector sees the events generated from MG and Pythia 8.

The information of the event simulator that comes from Pythia 8 will be true objects called TRUTH, whereas reconstructed objects found from detector information created via GEANT4 will be labelled as RECO.

5.4 Athena

Athena is a collision-event processing software framework used by the ATLAS experiment for event simulation and reconstruction, detector simulation and data analysis [57].

Next, we will go through the different steps in order to analyse the TRUTH and RECO events using the Athena framework.

- Event generation: simulation of the collisions of particles, with the subsequent parton showering and hadronization, as well as all the decays into stable particles. The output format is EVNT. (MadGraph and Pythia8)

- Detector simulation: simulates how the particles from the generator interact with the material from the detector and which energy deposit is left in each element. Now the output is HITS. (GEANT4)

- Digitization: turns the energy deposits into detector response, with RDO format.

- Reconstruction: identification and reconstruction of the different particles, giving xAOD outcome.

- Derivation: the same format as reconstruction step but reduced in size by choosing a "TRUTH" format depending on the level of detail needed and some filtering on the trigger or reconstruction objects properties. The format is DAOD.

These steps are schematically represented in Fig. 5.1. Following all green boxes will give TRUTH and RECO objects, while following the "validation" will only give TRUTH in an appropriate format for analysis.

Expanding in the TRUTH definition, there are four kinds of formats that can be checked with an ATLAS account in [58]. The specific formats used will be TRUTH1 for validation and TRUTH3 for analysis, the only difference being the information it stores. ATLAS recommends using TRUTH3 for the final analysis since it contains all the necessary information while reducing its size for better manipulation of the samples.
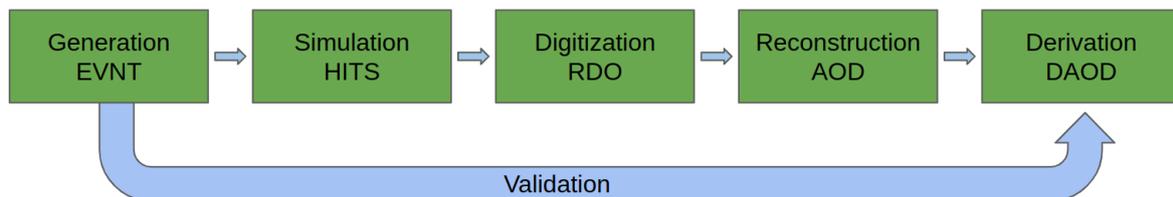
Figure 5.1: Representation of the steps to get from the generation to objects containing TRUTH and RECO for further analysis.

Validation

<div style="text-align: right; font-size: 3em;">6</div>

The model used in this project, SM_HeavyN_Gen3Mass_NLO [50, 51] was released recently when we started using it. An explicit validation of the physical variables and interactions was carried out.

6.1 Validate the model

The model works in leading order (LO) or next-to-leading order(NLO) which includes loops in the considered Feynman diagrams. We have used LO with the possibility of having up to two jets from the proton-proton interaction (LO+0,1,2 jets).

In this stage, a truth level study has been carried out following the bottom arrow of the generation steps described in Fig. 5.1. For the validation, 5,000 events have been generated using AthGeneration,21.6.73 with MadGraph 2.9,3 and Pythia 8.245, the versions can be accessed with an ATLAS account in [59]. The analysis was done using DAOD containing TRUTH1 information.

A first check consisted of comparing the ratio of the cross section [1] of W^-/W^+ production with the ratio of $\tau^- N/\tau^+ N$ [2]. In the following expression, "pp" stands for proton-proton collision and the arrow "$\rightarrow$" for the products of the collision. Both ratios are equal within the uncertainty

$$\frac{\sigma(pp \rightarrow W^-)}{\sigma(pp \rightarrow W^+)} = 0.719 \pm 0.001 \tag{6.1}$$

$$\frac{\sigma(pp \rightarrow \tau^- N)}{\sigma(pp \rightarrow \tau^+ N)} = 0.718 \pm 0.002. \tag{6.2}$$

---

[1]MadGraph and Pythia 8 gives a value of the cross section.
[2]N represents one of the HNLs. We can use any of the three allowed by the model.

In addition, we want to check if the particles are decaying properly, conserving the four momentum. For this reason, more checks were performed by adding the decay of the N and τ, as shown in Fig. 4.1. For these processes, the invariant mass and the transverse momentum, p_T, of the W, N and τ have been compared to the invariant mass and p_T of the sum of the decay products.

Once these initial checks were performed, we studied ten different channels involving tau mixing. For easy naming of the different channels, we will write the three leptons in order, for example, $\mu\tau e$ would mean the l1 is a μ, l2 a τ and l3 an e based on Fig. 4.1. The ten channels are: $e\tau e$, $e\tau\mu$, $\mu\tau\mu$, $\mu\tau e$, $\tau\mu e$, $\tau\mu\mu$, $\tau e\mu$, τee, $\tau\tau e$ and $\tau\tau\mu$.

For each of these processes, we asked for a lifetime of $c\tau_N = 0.1$ mm, a range of masses of $m_N = (20, 30, 40, 50)$ GeV, and both LNC and LNV scenarios (Fig. 4.1). We will plot and study the invariant mass, p_T and η of W and N, and p_T and η of the leptons. In addition, we plot the invariant masses between pairs of leptons using the Mandelstam variable s, and the displacement vertex in the transverse plane of the HNL decay. We can see in Fig. 6.1, 6.2 and 6.3 an example of these plots showing $\mu\tau e$ with an LNV scenario.

In Fig. 6.1a there is represented the on-shell W. On the left plot, we can see a peak centered at 80 GeV corresponding to the mass of W and a decay width of $\Gamma_W = 2.085$ GeV. In the middle, even though the W is initially generated without p_T, the jets and the parton shower in Pythia 8 make the W win some p_T. On the right plot, we can see clearly that the pseudorapidity has a bimodal distribution, with the maxima around ± 3.5, and that it is highly suppressed at zero (direction perpendicular to the beam).

Fig. 6.1b shows now the variables of the HNL, N. The invariant mass has a narrow peak according to the reference decay width in MadGraph of $\Gamma_{ref,N} = 1.97 \cdot 10^{-12}$ GeV. The p_T shows how the transverse momentum of N decreases when increasing the m_N, since the W produces a heavier particle and as a result, it will have less additional transverse momentum. For the pseudorapidity, we can observe that for high masses of the HNL the perpendicular direction is more suppressed. When the HNL is heavier, it is produced nearly at rest in the W frame and both η distributions would look alike. Whereas when the HNL is light, it will have a significant transverse momentum compared to W, which will balance the HNL distribution of η, not having a bimodal distribution anymore.

On the other hand, in Fig. 6.2a there is the p_T distributions of the three leptons l1, l2 and l3, in this case l1 $= \mu$, l2 $= \tau$ and l3 $= e$. The plot from l1 resembles the HNL

Figure 6.1: Histograms of $\mu\tau e$ LNV for different masses: blue for $m_N = 20$ GeV, red $m_N = 30$ GeV, green $m_N = 40$ GeV and black $m_N = 50$ GeV. In (a), there is the invariant mass m_W, the transverse momentum p_T^W and the pseudorapidity η_W of W. In (b), the same variables are represented for the HNL, N.

Figure 6.2: Histograms of $\mu\tau e$ LNV for different masses: blue for $m_N = 20$ GeV, red $m_N = 30$ GeV, green $m_N = 40$ GeV and black $m_N = 50$ GeV. In (a), there is the transverse momenta of the three leptons in order p_T^{l1}, p_T^{l2} and p_T^{l3}. In (b), the pseudorapidity η is shown for the three leptons, with the same order as before.
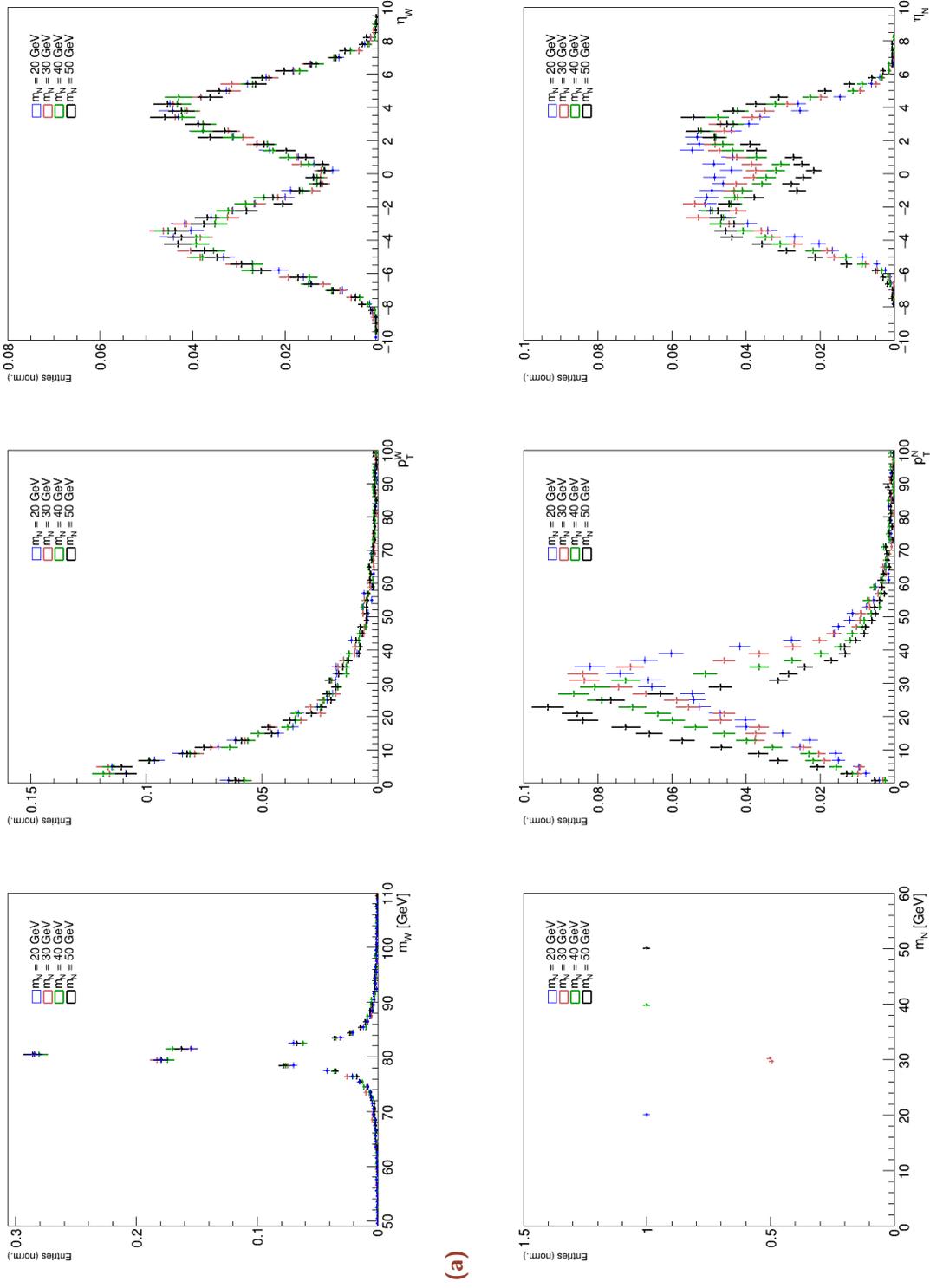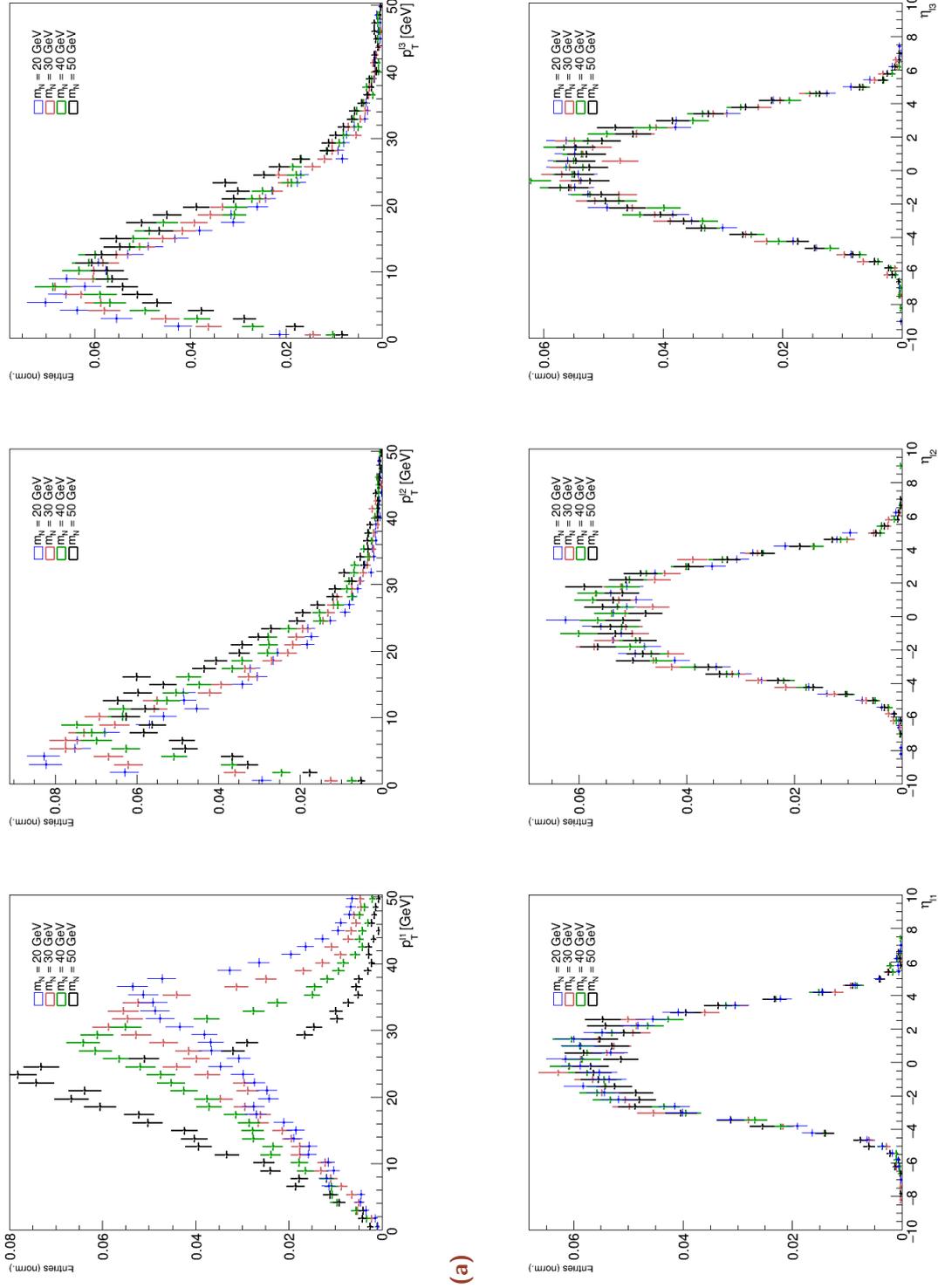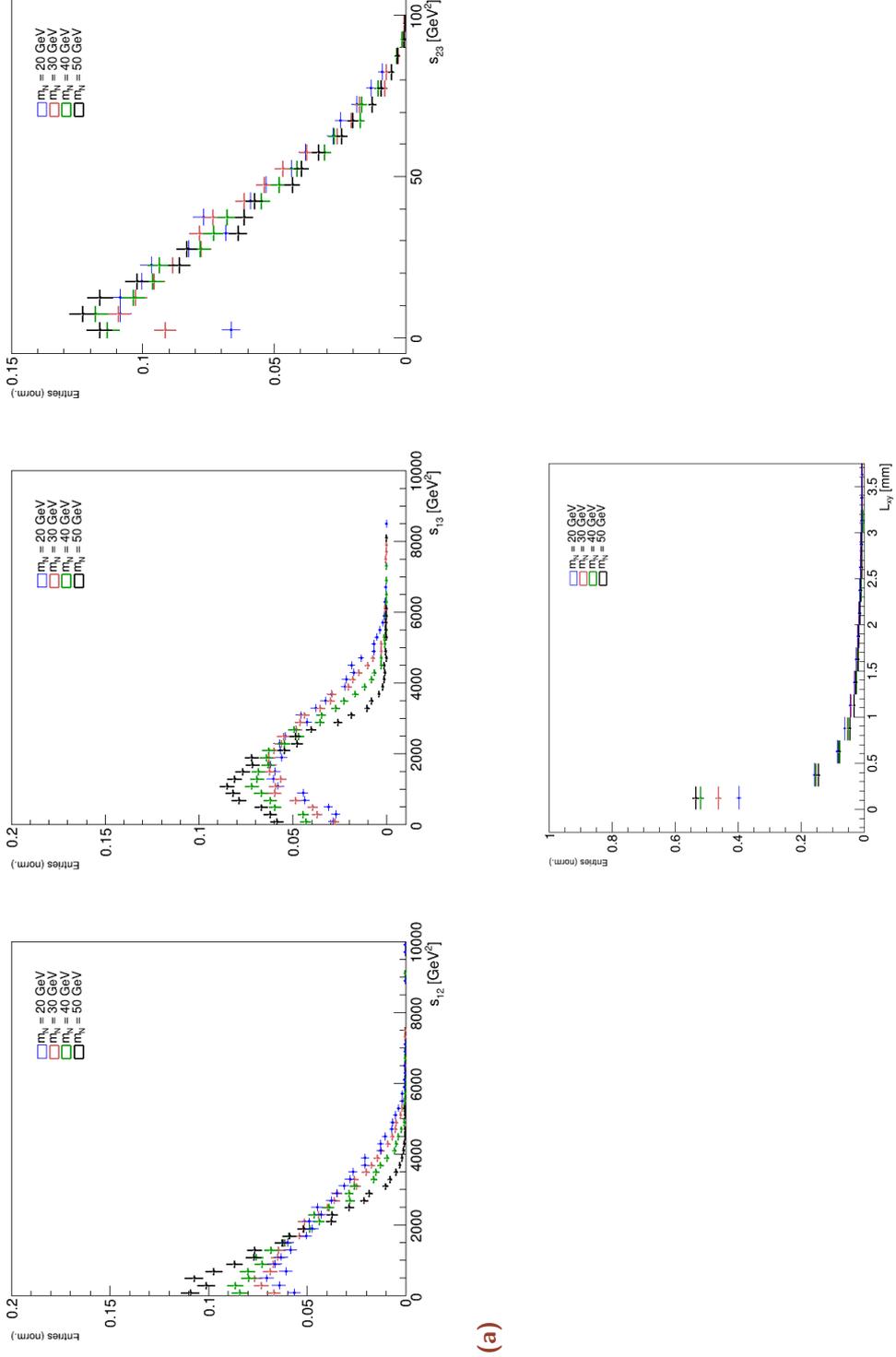
Figure 6.3: Histograms of $\mu\tau e$ LNV for different masses: blue for $m_N = 20$ GeV, red $m_N = 30$ GeV, green $m_N = 40$ GeV and black $m_N = 50$ GeV. In (a), there is the Mandelstam variable s for the different pairs of leptons. In (b), the transverse displaced vertex is shown.

distribution, decreasing its p_T when increasing the mass. Whereas l2 and l3 win slightly some p_T when increasing the HNL mass because even if HNL has less p_T, it has more energy and it can be distributed within its products. These p_T plots of l2 and l3 are sensitive to LNC/LNV processes, as can be seen in Fig. 6.4.

In Fig. 6.2b, we observed that the pseudorapidity of the three leptons is not affected by the change of m_N. In all cases there is a maximum around 0 with a half-width of $\approx \pm 4$. Zero is the direction perpendicular to the beam, so the majority of the leptons are very spread in $|\eta| < 4$ [3].

Fig. 6.3a contains the invariant mass squared (Mandelstam variable s) computed for the pair of leptons as $s_{12} = (p_{l1} + p_{l2})^2$, $s_{13} = (p_{l1} + p_{l3})^2$, $s_{23} = \frac{(p_{l2} + p_{l3})^2}{m_N^2/100}$, where p_α is the four momentum vector. The first and second plot distributions depend on the process being LNC or LNV (Fig. 6.4 for seeing the differences). On the right, the invariant mass squared of l2 and l3 is normalized with the HNL mass to check if the shape changes with m_N. Nevertheless, we can see that the invariant mass squared distribution looks the same for all masses when normalizing. This variable does not depend on having an LNC or LNV process because s_{23} gives the invariant mass squared of the visible HNL[4].

In Fig. 6.3b, the displacement in transversed plane of the HNL is plotted and it has been computed using $vtx_N = \sqrt{x_N^2 + y_N^2 + z_N^2}$. There is a maxima around $L_{xy} = 0.1$ mm, in agreement with the lifetime requested in the parameter card on MadGraph.

As mentioned before, some variables show different distributions depending on whether the process conserves the total leptonic number (LNC) or violates it (LNV). Fig. 6.4 explicitly shows the comparison of these variables for $\mu\tau e$.

In Fig. 6.4a the p_T^{l2} does not change much with mass, while we observe that the p_T increases significantly with m_N for l3. Comparing it to the LNV process, Fig. 6.4b, it looks like the l2 and l3 are swapped. This behaviour is expected because the charges and thus the chiralities of these leptons are swapped between LNC and LNV.

The Mandelstam variable s is sensitive to the effect of spin correlation between particles. For l2 and l3 the correlation is the same between the LNC scenario and the LNV (note that electric charge is opposite in both cases, $Q_{l2} \neq Q_{l3}$). However, for the pair s_{12} and

---

[3]Recall the logarithmic expression of η in Section 3.2.1.

[4]The neutrino decaying from HNL scape the detector. We call visible HNL the sum of the visible products (l2 and l3).
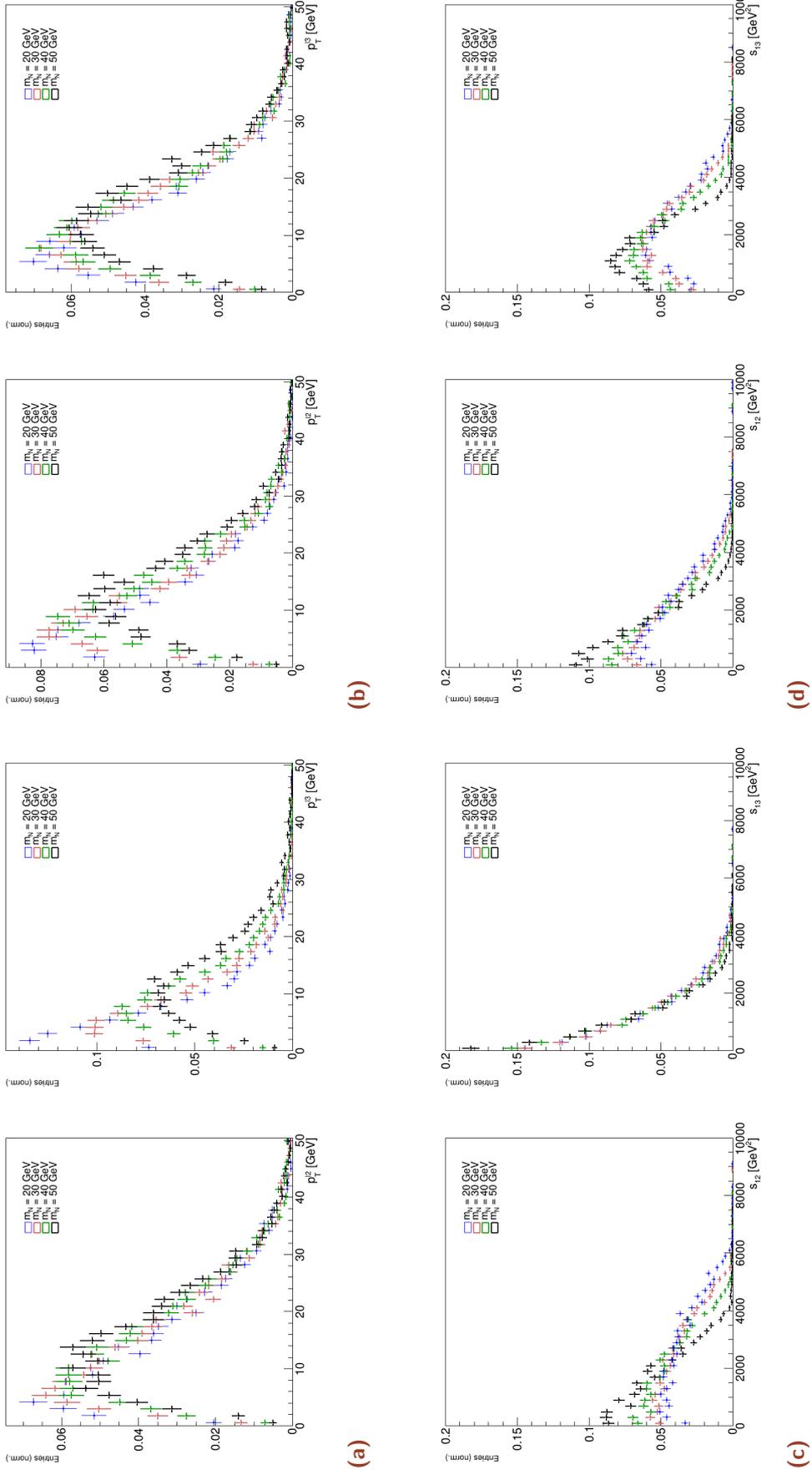
Figure 6.4: Comparison of LNC/LNV scenarios for $\mu\tau e$. In (a), p_T of l1 and l2 for LNC scenario compared to (b), where there is the LNV scenario. In (c), one can find the invariant mass squared between l1-l2 and l1-l3 for LNC, while (d) shows it for LNV.

s_{13} this is not the case. Now, there is spin correlation between LNC and LNV processes. In Fig. 6.4c and 6.4d, this difference is shown.

All the physical interactions and kinematics observed in the validation have been interpreted and explained. This model is then suitable for our analysis and will be used for the generation of all the events requested.

6.2 Trigger efficiency estimation

Before proceeding with the search, we want to estimate which process involving HNL mixed with tau flavour is most likely to produce enough events to be visible at the ATLAS detector. This starts at the trigger level. So first, we want to estimate the efficiency of the triggers available in ATLAS. The efficiency ε is defined as the number of events passing a cut divided by the number of events before the cut.

Since at this stage we do not have yet a proper trigger simulation, only Monte Carlo truth information, we emulate the trigger selection by applying some p_T and η cuts: we compute how many events we are left over after imposing a cut of $|\eta| < 2.5$, due to the geometry of the tracking detector, and a cut on the leptons p_T.

The goal is to establish which final state for W production of HNLs can provide new data to set limits, or observe, tau flavour neutrino mixing with HNLs. One can have a process sensitive to mixing angle to tau squared $U_{\tau I}^2$ (W decaying to τ and ν_τ mixing to HNL and back to ν_τ) or to mixing angle to tau and to another flavour (for example, W decaying to μ and ν_μ, mixing to HNL and back to ν_τ). We compute it for different channels with lepton number conservation (LNC) with $m_N = 20, 50$ GeV and depending on how the tau decays, either leptonically or hadronically. In all the cases, we have at least one tau decaying hadronically.

Comparing Tab. 6.1 and 6.2 we can get an idea of which process could be more promising to study. We will explain them from left to right comparing both HNL masses.

For $\tau\mu e$, since the μ decays directly from the HNL (l2) while the e comes from the off-shell W (l3), the μ carries more momentum, therefore, having more ε for the muon trigger than the electron one. When increasing m_N, even with the p_T of the HNL

---

[5]Double tau trigger, one tau with $p_T > 40$ GeV and the other with $p_T > 30$ GeV.

	$\tau\mu e$ (h)	$\mu\tau e$ (h)	$e\tau e$ (h)	$\tau\tau e$ (hh)	$\tau\tau e$ (he)	$\tau\tau e$ (hμ)
Total N	6481	6475	6425	419	2340	2285
$|\eta|<2.5$	3146	3241	3173	2140	1165	1115
$p_T^e > 26$ GeV	133	118	2419	67	129	46
ε_e	0.021	0.018	0.376	0.016	0.055	0.020
$p_T^\mu > 26$ GeV	512	2418				89
ε_μ	0.079	0.373				0.039
$p_T^{\tau_{had-vis}} > 30, 40$ GeV [5]				8		
ε_τ				0.002		

Table 6.1: Efficiencies ε for $m_N = 20$ GeV of three different triggers for a representative choice of processes for LNC scenario. The efficiencies are computed from N, the total number of events for a given final state (indicated in the red cells).

	$\tau\mu e$ (h)	$\mu\tau e$ (h)	$e\tau e$ (h)	$\tau\tau e$ (hh)	$\tau\tau e$ (he)	$\tau\tau e$ (hμ)
Total N	6468	6405	6564	4315	2253	2233
$|\eta|<2.5$	3178	3182	3335	2231	1148	1127
$p_T^e > 26$ GeV	370	364	1237	243	134	128
ε_e	0.057	0.057	0.188	0.056	0.059	0.057
$p_T^\mu > 26$ GeV	556	955				35
ε_μ	0.086	0.149				0.016
$p_T^{\tau_{had-vis}} > 30, 40$ GeV				10		
ε_τ				0.002		

Table 6.2: Efficiencies ε for $m_N = 50$ GeV of three different triggers for a representative choice of processes for LNC scenario. The efficiencies are computed from N, the total number of events for a given final state (indicated in the red cells).

decreasing, the p_T of the HNL decay products increases, so we have bigger efficiencies with higher mass.

In the $\mu\tau e$ case, μ is now in the position of l1, the prompt lepton, having high momentum because it comes directly from the on-shell W and therefore, many events will be selected by the muon trigger. When increasing the mass, the p_T of l1 decreases, and so does its efficiency. As in the previous process, p_T of l2 and l3 increases with HNL mass, improving the electron trigger.

The third column, $e\tau e$, shows how since there is no muon on the process, no events are selected due to the muon trigger. Instead, the electrons that go through the trigger come from both l1 and l3, being more significant for the trigger the l1. When increasing m_N, the efficiency decreases because p_T of l1 decreases more than the increase of p_T of l3.

The process $\tau\tau e$ has been divided into three possibilities of the different tau decays: both taus decaying hadronically (hh) or one tau decaying hadronically and the other

leptonically, either to a muon (hμ) or to an electron (he). The fourth column shows the efficiencies when both taus decay hadronically, $\tau\tau e$(hh). This case is the only one where we could use the double tau trigger, but the big constraints of high p_T make the efficiency to be very low. The electron trigger improves with m_N with similar efficiencies than $\tau\mu e$ and $\mu\tau e$, since the e is placed in the same place (l3).

The next column, $\tau\tau e$(he), takes into account one of the taus decaying leptonically into an electron. The efficiency does not change significantly because even if the l3 and l2 gain p_T with mass, the l1 loses p_T, compensating its numbers.

For the last column, $\tau\tau e$(hμ), the electron trigger should resemble the values we had for $\tau\tau e$(hh) since all the electrons come from l3. In addition, we observe a decrease on ε_μ with the mass of the HNL hinting that most of the events that go through the cut come from the τ of l1 decaying into a μ, instead of coming from the l2. If it was the other way around, we would observe an increase in efficiency with m_N.

With these preliminary studies of the processes, we can already see that the double tau trigger is very inefficient for a process with two taus, and in addition, $\tau\tau e$ has low efficiency in general for all the triggers. Actually, studying a process with an HNL only mixing to the same flavour would not help us explain any Beyond Standard Model physics.

We observed that those processes with only a τ in l2 are the ones surviving best. Therefore, for the purpose of this thesis, we will focus on one of these channels. Between all the possibilities of having a τ in l2, $\mu\tau e$ has been selected because having three different flavour leptons might help in reducing the background. Both electron and muon triggers will be more investigated since for high masses both triggers show comparable efficiencies.

Clearly, other processes would give us the same final state, a muon, a hadronic tau and an electron. For instance, $e\tau\mu$ or $\tau\tau\mu$(he), and even if not containing an HNL mixing with tau, $\mu e\tau$ could also be selected from our triggers as signal. However, we will proceed with only considering $\mu\tau e$ as the process originating the signal.

Data and Monte Carlo samples

Focusing our efforts on studying the channel $\mu\tau e$, we will first explicitly write the Feynman diagram of the process, Fig. 7.1. In addition, the charged conjugate of the processes shown also needs to be taken into account.

Writing down all the possibilities

$$pp \to W^{\pm} + X \quad \text{with} \quad W^{\pm} \to \mu^{\pm} + N \quad \text{followed by} \quad N \to \tau^{\mp} + e^{\pm} + \overset{(-)}{\nu_e} \quad (7.1)$$

$$pp \to W^{\pm} + X \quad \text{with} \quad W^{\pm} \to \mu^{\pm} + N \quad \text{followed by} \quad N \to \tau^{\pm} + e^{\mp} + \overset{(-)}{\nu_e}, \quad (7.2)$$

where X represents the jets from 0 up to 2, and N once again represents the HNL. Eq. 7.1 shows a process where the total lepton number is conserved $\Delta L = 0$ (LNC), in comparison with Eq. 7.2 that violates total lepton number by two units $\Delta L = \pm 2$ (LNV).

(a) (b)
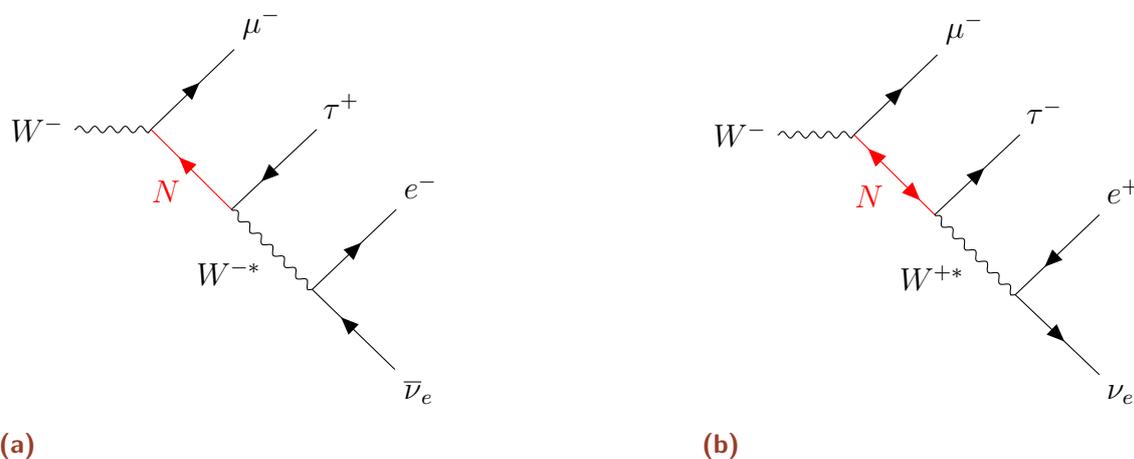
Figure 7.1: Feynman diagram examples of the process $\mu\tau e$, LNC on the left and LNV on the right. A part from these diagrams, we have their respective charge conjugation processes and the possibilities to have jets coming from the proton quarks.

The simulation steps are represented schematically in Fig. 7.2. Until this moment, all simulations were generated in a level truth of the events (following the bottom blue arrow). For analysing the sensitivity of ATLAS to this channel, a complete reconstruction

of the objects (RECO) needs to be done, meaning our samples need to go through all the steps in Fig. 7.2. Having the truth information at the same time is crucial to compare how much the detected signal differs from the truth and how much signal we lose.

Once the samples are ready, we need to build a specific strategy to select the process $\mu\tau e$. All the samples used from this point onwards come from a large request to the ATLAS Monte Carlo Production team, who generated, simulated, digitized and reconstructed the events giving all the information we need to prepare our analysis strategy. Once the request was completed all the files were in an AOD format. We made another request to derive the samples into the desired format in order to study them more easily, since DAOD format has a much smaller size. The derivation was performed by the ATLAS SUSY DPD Production team, using DAOD_SUSY3 format, which is designed for hadronic tau analyses, asking for at least one reconstructed tau with $p_T^\tau > 10$ GeV and $|\eta| < 2.6$. More information on DAOD_SUSY3 can be accessed with an ATLAS account in [60]. The generation and derivation was performed by MadGraph (v.2.9.3.atlas), Pythia8 (v.245p3.lhcb7) and EvtGen(v.1.7.0), with AthDerivation_21.2.130.0, these releases can be accessed with an ATLAS account in [59, 61].
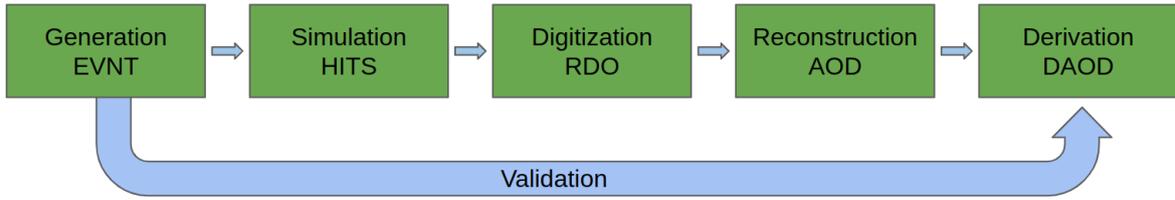
Figure 7.2: Representation of the steps to get from the generation to objects containing TRUTH and RECO for further analysis.

We did a first request of $\mu\tau e$ for LNC asking for $180,000$ events split between three different periods of Run-2 (mc16a, mc16d and mc16e) and for each $m_N = 20, 30, 40, 50$ GeV. After a first glance at the results, we wanted to have more statistics, so we requested for the same channel but for LNV with ten times more statistics, $1,800,000$ events, again for the three same data periods.

From now on, we will label the signal final state $\mu\tau e$, which involves an HNL, as the signal. But when analysing real data, there will be some processes that are seen by the detector to have the same final state as the process we want to study, in this case $\mu\tau e$. Processes well defined in the Standard Model giving a similar final state but without involving an HNL will be called background [1].

---

[1]Other processes of BSM like the ones shown in chapter 6 are not considered as background for our analysis.

Our aim is to build a smart strategy to select as much signal as possible while rejecting the background.

A complete analysis would require a thorough study of all the background source for our signal, with $W^{\pm} \to l_{\alpha}^{\pm} \overset{(-)}{\nu_{\alpha}} + X$, $Z \to l_{\alpha}^{\pm} l_{\alpha}^{\mp} + X$ and $t\bar{t}$ being the most important ones. For the purpose of this thesis, we will only study $W^{\pm} \to l_{\alpha}^{\pm} \overset{(-)}{\nu_{\alpha}} + X$, more specifically $W^{\pm} \to \mu^{\pm} \overset{(-)}{\nu_{\mu}} + X$ and $W^{\pm} \to e^{\pm} \overset{(-)}{\nu_{e}} + X$. In particular, we have chosen the datasets that have a p_T between 0 and 70 GeV for the W boson and that is filtered in respect of bottom or charm quarks also, called light jets samples [2]. The background containing bottom or charm quarks in the final state has not been considered because it is expected that a request of no jets passing bottom or charm tagging would reject them to a large extent. The background not containing these quarks is truly more dangerous and should be considered first.

The background simulation was performed in 2019, generated by Sherpa(v2.2.1) with AthDerivation_21.2.79.0, for the same data period as the signal request (mc16a, mc16d and mc16e).

	σ [fb]	Filter ε	Total events	Events filtered for analysis
$W \to \mu\nu$	$19.2 \cdot 10^6$	0.82465	97,624,850	6,209,924
$W \to e\nu$	$19.2 \cdot 10^6$	0.82463	21,278,251	11,747,676
$\mu\tau e$ 20 LNC	$5.820 \cdot 10^{11}$	1	162,000	26,879
$\mu\tau e$ 50 LNC	$3.692 \cdot 10^{13}$	1	180,000	37,895
$\mu\tau e$ 20 LNV	$5.819 \cdot 10^{11}$	1	1,770,000	422,822
$\mu\tau e$ 50 LNV	$3.696 \cdot 10^{13}$	1	1,100,000	272,091

Table 7.1: Table with the cross section reported by Atlas Metadata Interface [63], the efficiency of the generator filter, the number of total events generated and the number of events converted into DAOD.

In Tab. 7.1 there is a summary of the data sample information. The cross section reported by ATLAS Metadata Interface (AMI) [63], comes in the case of signal from the cross section given by MadGraph + Pythia 8. For the background, the cross section reported by AMI has been computed by the Physics Modelling Group.

For the signal and a specific mass, we expect the cross section of LNC and LNV processes to match because having an on-shell HNL allows us to separate the production and the decay using the narrow-width approximation [64], and the same branching ratios will be

---

[2]More information about weak boson processes can be accessed with an ATLAS account in [62].

applied to both processes, raising the same cross section. This can be seen in chapter 8, more precisely in Eq. 8.2.

The error of the cross section is a complex quantity to compute, having many contributions like the scale or the PDF uncertainties. The systematical error of the numerical integration, performed by MadGraph in order to calculate the cross section, is the only error we could estimate. Its value of approximate 0.5%, is negligible compared to the statistical error taken into account in chapter 9. In any case, the error provided by MadGraph has been used to give the significant digits of the calculations that follows from the cross section.

The σ given by MadGraph + Pythia 8 for the signal is not the physical cross section, because we have used an unphysical pair of values for the mixing angles, $U_{\alpha,I}$, and decay width of N, Γ_N, see chapter 8 for a detailed explanation.

In the signal case we did not ask for any cut in the generator, therefore, the generator efficiency is 1. More information of the datasets used can be found in Appendix A.

Signal prediction

<div style="text-align: right; font-size: 2em; color: #c00;">8</div>

The cross section, σ, is an important parameter in particle physics as it tells us how probable it is that two particles collide and interact in a certain way. On the other hand, the integrated luminosity, $\mathcal{L}$, gives us an idea of how packed the particles are in a beam for a period of time, for instance, the integrated luminosity in Run-2 was 139 fb^{-1} in ATLAS [29]. The efficiency, ε, takes into account the limitations of the detector and the analysis telling us how likely it is that we will detect the event once created. Multiplying all these variables gives the expected number of a specific event for a period of time

$$Y = \mathcal{L}\sigma\varepsilon, \tag{8.1}$$

where we have used Y instead of the usual N for not confusing it with the HNL.

The cross section is often given by the Monte Carlo generators used like MadGraph (MG) or Pythia 8. For leading order processes, it is computed by generating tree-level matrix elements using Feynman rules for the different diagrams and performing the numerical integrations [65]. For the signal, it can also be computed experimentally using information known from SM processes. In the case of $\mu\tau e$

$$\sigma(pp \to \mu(N \to \tau e\nu_e)) = \sigma(pp \to W) \cdot Br(W \to \mu\nu_\mu) \cdot |U_\mu|^2 \cdot Br(N \to \tau e\nu_e)^1, \tag{8.2}$$

where Br denotes the branching ratio defined by the partial width of an specific decay divided by the total width of that particle,

$$Br(N \to \tau e\nu_e) = \frac{\Gamma(N \to \tau e\nu_e)}{\Gamma_N}. \tag{8.3}$$

Whereas the cross section for the background processes is very well defined in the ATLAS experiment for specific conditions of the detector, getting the cross section for our signal is more complex and the value given by MG needs to be rescaled. This is because we used an unphysical pair of values for the HNL decay width $\Gamma_N^{ref} = 1.97 \cdot 10^{-12}$ GeV and

---

[1] A phase space correction keeping track of the mass of the HNL should also be taken into account.

the mixing angles $|U_\mu^{ref}| = |U_\tau^{ref}| = 1$, where we use the subscript "ref" to make explicit that these values are the reference given to MG.

Using the method described in [66], we compute the physical cross section σ_p for a specific m_N and mixing angles, using the σ^{ref} given by MG, and Γ_N^{ref} and $|U_\alpha^{ref}|$ given to MG as

$$\sigma_P\left(m_N, U_e, U_\mu, U_\tau\right) = \sigma^{\text{ref}} \times \frac{|U_\alpha|^2 |U_\beta|^2}{|U^{\text{ref}}|^4} \times \frac{\Gamma_N^{\text{ref}}}{\Gamma_N\left(m_N, U_e, U_\mu, U_\tau\right)}, \qquad (8.4)$$

where $\beta, \alpha = e, \mu, \tau$ flavour. At the same time the decay width of HNL depends on the mixing angles as follows

$$\Gamma_N\left(m_N, U_e, U_\mu, U_\tau\right) = \sum_{\beta = e, \mu, \tau} \frac{|U_\beta|^2}{|U^{ref}|^2} \times \hat{\Gamma}_\beta\left(m_N\right). \qquad (8.5)$$

To understand Eq. 8.4 let us recall Eq. 8.2 and 8.3. The branching ratio is proportional to the mixing angle squared and inversely proportional the the total decay width of HNL, therefore $\sigma \propto \frac{|U_\alpha|^2 |U_\beta|^2}{\Gamma_N}$. If we want to rescale it with other values for mixing angles and decay width, Eq. 8.4 is consistent.

At the same time Eq. 8.5 comes from the definition of the decay width and its relation with the total mixing angle $U^2 = \sum_\alpha |U_\alpha^2|$ with $\alpha = e, \mu, \tau$, that far from the QCD and the electroweak scale (5 GeV $\ll m_{HNL} \ll$ 80 GeV) can be approximated to [47, 67]

$$c\tau_N^{-1} = \Gamma_N \simeq 11.9 \times \frac{G_F^2}{96\pi^3} U^2 m_N^5, \qquad (8.6)$$

where $c\tau$ is the lifetime and G_F the Fermi constants.

Now that Eq. 8.4 and 8.5 are understood, let us use it to get some physical cross sections for specific mixing angles and HNL mass.

For computing $\hat{\Gamma}_\beta\left(m_N\right)$, we have run the process $n1 >$ *all all all* [2] in MadGraph, for $|U^{ref}| = 1$ for each flavour and m_N individually. Since in the process studied our HNL mixes with μ and τ, we do not need to compute $\hat{\Gamma}_e$. In Tab. 8.1, the $\hat{\Gamma}_\beta(m_N)$ has been

---

[2]Notation used by MadGraph to express the generation of HNL decaying into all possibilities.

computed. Once more the error given by MadGraph appears to be negligible with the statistical error calculated in chapter 9, being in the worst case 0.2%.

m_N [GeV]	$\hat{\Gamma}_\mu$ [GeV]	$\hat{\Gamma}_\tau$ [GeV]
20	$1.807 \cdot 10^{-6}$	$1.740 \cdot 10^{-6}$
50	$2.259 \cdot 10^{-4}$	$2.244 \cdot 10^{-4}$

Table 8.1: Decays widths of the HNL computed in MG for two different HNL masses $m_N = 20, 50$ GeV, and using a reference mixing angle of $|U_\mu| = 1$ while setting the rest to zero in the second column, and in the third column setting all the mixing angles to zero while $|U_\tau| = 1$.

Once Eq. 8.5 is used to get the total decay width of the HNL, the values can be compared to the approximation given in Eq. 8.6 and represented in Fig. 8.1. We can see that for HNLs with longer lifetimes, the mixing angle is smaller because it is harder for it to decay. For a specific lifetime, we can observe that the greater the HNL mass, the smaller the $|U^2|$ is due to its dependence on the mass in the phase space integral.
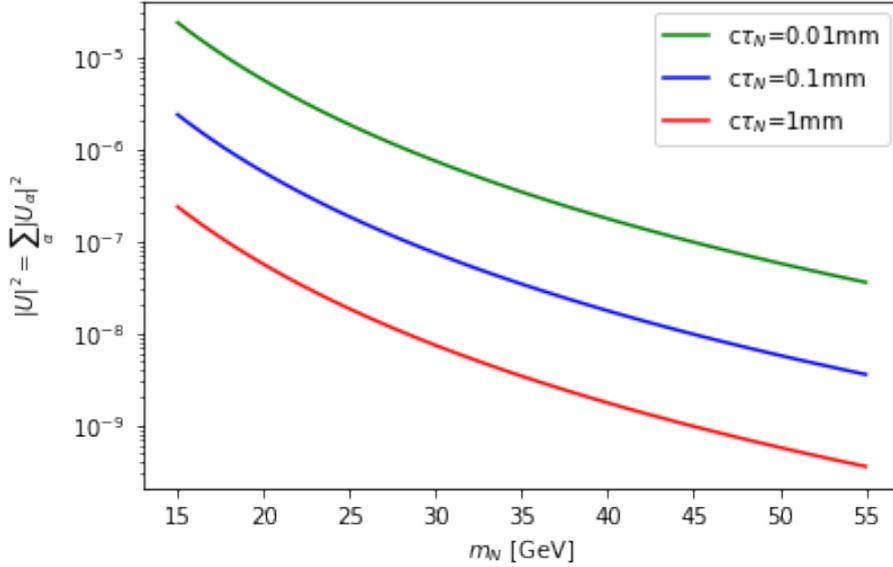
Figure 8.1: Relation of the total mixing angle squared $|U|^2$ depending on the HNL mass for different lifetimes $c\tau_N$, where $|U_\alpha|$ denotes the specific mixing angle.

From Fig. 8.1, we can get the lower bound for a prompt HNL that can be investigated, set to $c\tau_N = 0.1$ mm. Any lifetime longer than 0.1 mm requires special reconstruction, not subject of this project. Note that for this lifetime, the total mixing angle goes down to $|U(m_N = 20)|^2 \simeq 5.58 \cdot 10^{-7}$ and $|U(m_N = 50)|^2 \simeq 5.71 \cdot 10^{-9}$, but we need to compute if we are sensible to so small mixing angles.

Now that we know how to compute the cross section and the mixing angles for a specific lifetime, we need to select smart choices of $|U_\alpha|$. Using the neutrino oscillation data

we can look for combinations of $|U_\alpha|^2/|U|^2$ that are consistent with the data, as shown in Fig. 8.2 [66]. Since the studied process mixes with μ and τ flavour, we choose to work with benchmark points 4 and 6, corresponding to a normal hierarchy. This is just a choice, and of course, one should also consider other scenarios for a more complete overview of sensitivity to HNL.
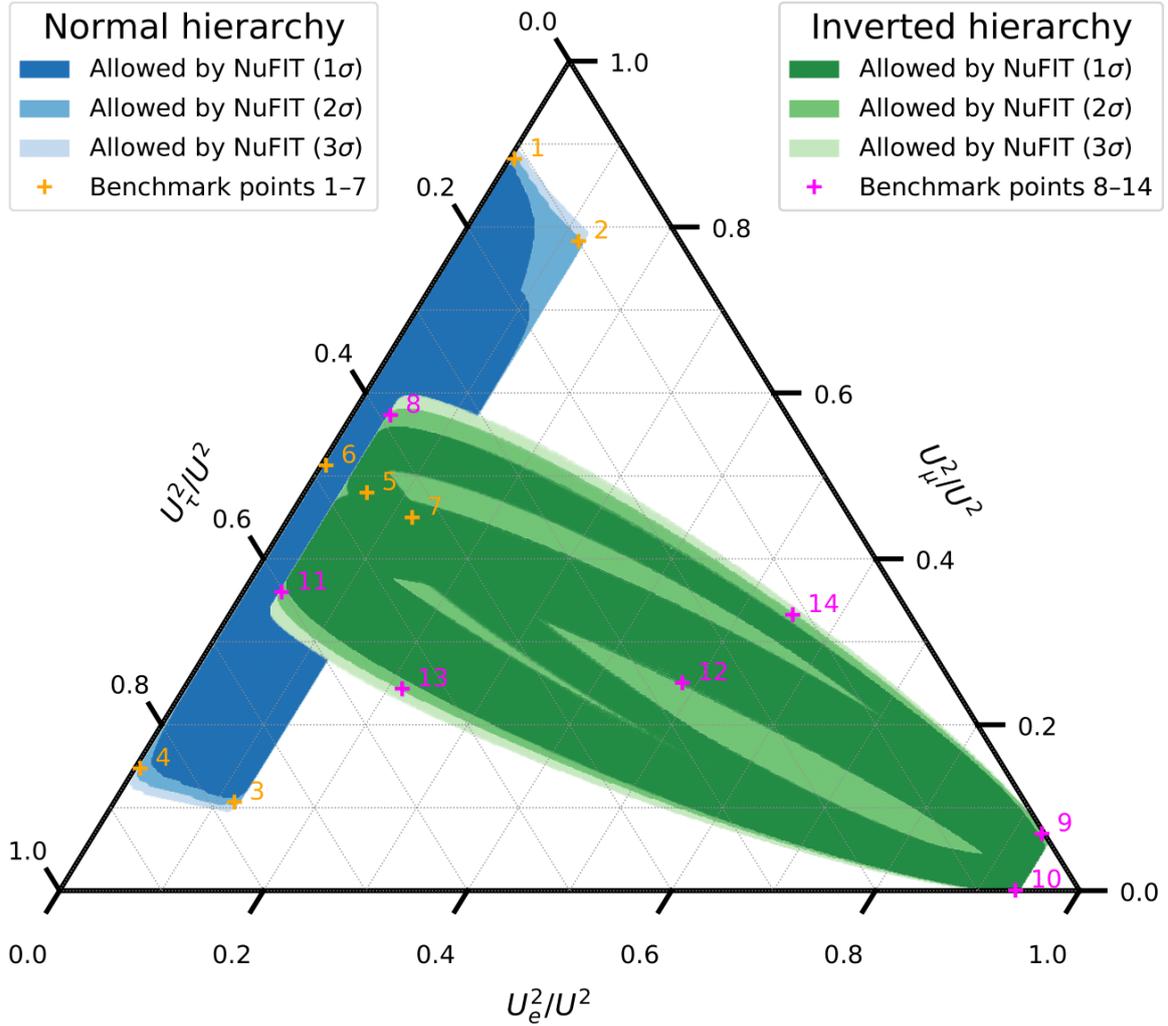
Figure 8.2: Ternary plot from [66], showing the combinations of mixing angles U_α^2/U^2 with $\alpha = e, \mu, \tau$ flavour, compatible with the neutrino oscillation data [68, 69].

In Tab. 8.2 and 8.3, we showed the physical cross section computed from the benchmark points 4 and 6 and for different lifetimes for LNC case, using Eq. 8.5 and 8.4. Since the referece cross section between LNC and LNV are alike in Tab. 7.1, similar results for the physical cross sections are found in LNV case.

m_N [GeV]	$c\tau_N$ [mm]	$|U|^2$	$|U_\mu| = |U_\tau|$	Γ_N [GeV]	σ [fb]
20	0.1	$5.58 \cdot 10^{-7}$	$5.28 \cdot 10^{-4}$	$9.90 \cdot 10^{-13}$	$9.04 \cdot 10^{-2}$
20	0.01	$5.58 \cdot 10^{-6}$	$1.67 \cdot 10^{-3}$	$9.90 \cdot 10^{-12}$	$9.04 \cdot 10^{-1}$
20	0.001	$5.58 \cdot 10^{-5}$	$5.28 \cdot 10^{-3}$	$9.90 \cdot 10^{-11}$	9.04
50	0.1	$5.71 \cdot 10^{-9}$	$5.34 \cdot 10^{-5}$	$1.286 \cdot 10^{-12}$	$4.62 \cdot 10^{-4}$
50	0.01	$5.71 \cdot 10^{-8}$	$1.69 \cdot 10^{-4}$	$1.286 \cdot 10^{-11}$	$4.62 \cdot 10^{-3}$
50	0.001	$5.71 \cdot 10^{-7}$	$5.34 \cdot 10^{-4}$	$1.286 \cdot 10^{-10}$	$4.62 \cdot 10^{-2}$
50	0.0001	$5.71 \cdot 10^{-6}$	$1.69 \cdot 10^{-3}$	$1.286 \cdot 10^{-9}$	$4.62 \cdot 10^{-1}$
50	0.00001	$5.71 \cdot 10^{-5}$	$5.34 \cdot 10^{-3}$	$1.286 \cdot 10^{-8}$	4.62

Table 8.2: Table containing for $m_N = 20, 50$ GeV and a specific lifetime of HNL its allowed total mixing angle squared $|U^2|$ (Eq. 8.6), the mixing angle to muons and taus, the total decay width (Eq. 8.5) and cross section (Eq. 8.4) for the **benchmark point 6** in LNC.

m_N [GeV]	$c\tau_N$ [mm]	$|U|^2$	$|U_\mu|$	$|U_\tau|$	Γ_N [GeV]	σ [fb]
20	0.1	$5.58 \cdot 10^{-7}$	$2.89 \cdot 10^{-4}$	$6.89 \cdot 10^{-4}$	$9.77 \cdot 10^{-13}$	$4.67 \cdot 10^{-2}$
20	0.01	$5.58 \cdot 10^{-6}$	$9.15 \cdot 10^{-4}$	$2.18 \cdot 10^{-3}$	$9.77 \cdot 10^{-12}$	$4.67 \cdot 10^{-1}$
20	0.001	$5.58 \cdot 10^{-5}$	$2.89 \cdot 10^{-3}$	$6.89 \cdot 10^{-3}$	$9.77 \cdot 10^{-11}$	4.67
50	0.1	$5.71 \cdot 10^{-9}$	$2.93 \cdot 10^{-5}$	$6.97 \cdot 10^{-5}$	$1.283 \cdot 10^{-12}$	$2.361 \cdot 10^{-4}$
50	0.01	$5.71 \cdot 10^{-8}$	$9.25 \cdot 10^{-5}$	$2.20 \cdot 10^{-4}$	$1.283 \cdot 10^{-11}$	$2.361 \cdot 10^{-3}$
50	0.001	$5.71 \cdot 10^{-7}$	$2.93 \cdot 10^{-4}$	$6.97 \cdot 10^{-4}$	$1.283 \cdot 10^{-10}$	$2.361 \cdot 10^{-2}$
50	0.0001	$5.71 \cdot 10^{-6}$	$9.25 \cdot 10^{-4}$	$2.20 \cdot 10^{-3}$	$1.283 \cdot 10^{-9}$	$2.361 \cdot 10^{-1}$
50	0.00001	$5.71 \cdot 10^{-5}$	$2.93 \cdot 10^{-3}$	$6.97 \cdot 10^{-3}$	$1.283 \cdot 10^{-8}$	2.361

Table 8.3: Table containing for $m_N = 20, 50$ GeV and a specific lifetime of HNL its allowed total mixing angle squared $|U^2|$ (Eq. 8.6), the mixing angle of muons and taus, the total decay width (Eq. 8.5) and cross section (Eq. 8.4) for the **benchmark point 4** in LNC.

Inside the same table, comparing m_N for the same lifetime we see how for 50 GeV the values of σ are two orders of magnitude less than for 20 GeV. This already comes from Tab. 8.1, when we computed the intermediate step for $\hat{\Gamma}_N$.

Given the current limits shown in Fig. 4.2, choosing a mixing angle squared of at least $|U|^2 = 10^{-5}$ is what is interesting to see if we are sensitive to, since everything higher is already excluded. Then, for $m_N = 20$ GeV we will have at least $c\tau_N = 10^{-3}$ and $c\tau_N = 10^{-5}$ for $m_N = 50$ GeV.

Signal selection

9.1 Preselection

All the events that have been saved from a specific run will go through our analysis aiming at selecting the signal we are studying. In this process, it is inevitable to be selecting background as well. In order to minimize the background as much as possible, we define a preselection strategy. We are interested in getting the efficiencies of the individual cuts (for validation) and the overall cut flow.

Cut flow
Acceptance: $p_T^{\mu,e} > 5$ GeV, $p_T^{\tau} > 10$ GeV, $|\eta_{\mu,\tau,e}| < 2.5$
Trigger: $p_T^{\mu} > 26$ GeV
Identification: Muon Tight
Identification: Electron Medium
Identification: Tau Medium
$\Delta R_{e-\tau} > 0.2$
$\Delta R_{\mu-\tau} > 0.6$
Charge: LNC/LNV configuration

Table 9.1: Summary of our cut flow for $\mu\tau e$ process.

In Tab. 9.1 there is a summary of the choices when preselecting the events. Next, a detailed explanation of these choices is described. We have studied the different possibilities by using the lepton number conservation (LNC) sample of the signal.

The events produced with MadGraph are converted to a DAOD where they contain the information of both the TRUTH and the RECO objects (Fig. 7.2). Precisely, the DAOD_SUSY3 format only saves those events with at least one reconstructed hadronic tau with a minimum $p_T > 10$ GeV and $|\eta| < 2.5$, condition coming from the tau reconstruction and the detector acceptance limitations. In order to always choose a final state needed by the search, we will impose some additional conditions on p_T and η for three leptons. These conditions are all summarized in Tab. 9.2.

Lepton flavour	p_T[GeV] $>$	$|\eta| <$
μ	5	2.5
τ	10	2.5
e	5	2.5

Table 9.2: Requirements of p_T and η for each lepton flavour.

The events with at least one lepton of each type described in Tab. 9.2 need to be selected further. For this reason, we have to consider the sources of our background already and take decisions on how we can reduce the background without losing signal events.

The events we are interested in will not be recorded if they are not selected as relevant by one of the actual triggers on the ATLAS detector (section 3.2.3). For the specific channel $\mu\tau e$, we could use either the muon or the electron trigger, as explained in section 6.2, which imposes $p_T^\mu > 26$ GeV or $p_T^e > 26$ GeV respectively. For the background we will take into account Standard Model production $W \to \mu\nu_\mu$ and $W \to e\nu_e$, expected to get through the trigger filter due to its prompt lepton.

These backgrounds have a prompt lepton (the one decaying from W), another fake lepton and a jet that gives a τ. This is why both backgrounds end up giving the same final state as our signal.

In the preliminary studies that can be revisited in section 6.2, we can see that the electron trigger gets a more important role for the signal, the higher the HNL mass is. Because of this, we tried to use two different strategies: for low HNL masses using only the μ trigger, and for higher HNL masses using both μ and e triggers. If only the muon trigger is used, $W \to e\nu_e$ will not be triggered because the fake muons do not have enough p_T (for example, see Tab. 9.3). Whereas for high mass if we use both triggers, we need to consider $W \to \mu\nu_\mu$ background, but also $W \to e\nu_e$.

For $W \to \mu\nu_\mu$, some events will go through our preselection because there is a prompt μ, a fake e and a jet that fakes a τ, but we end up selecting the same events if we use both triggers or only the muon, meaning that the fake electrons have not significantly enough p_T. When using both triggers for $W \to e\nu_e$, we get an efficiency of the same order of magnitude as $W \to \mu\nu_\mu$ due to the electron trigger.

The effect on the signal of using only the muon trigger or the electron and muon triggers does not vary significantly for $m_N = 20$ GeV, and is a small increase for $m_N = 50$ GeV, while bringing on twice the background.

In Tab. 9.3, there is a summary of the total efficiencies for background and signal when using both triggers or only the muon one. We can see how using only the muon trigger we are getting rid of $W \to e\nu_e$ while keeping similar efficiencies for the signal and $W \to \mu\nu_\mu$. With these numbers we decided to proceed using only the muon trigger.

Triggers	$W \to \mu\nu$	$W \to e\nu$	$\mu\tau e$ 20 LNC	$\mu\tau e$ 50 LNC
μ or e	$5.35{\cdot}10^{-6}$	$2.40{\cdot}10^{-6}$	0.00251	0.00217
μ	$5.31{\cdot}10^{-6}$	0	0.00246	0.00172

Table 9.3: Preselection efficiencies when using μ and e trigger or only using μ for the background and the signal using the rest of the cut flow described in Tab. 9.1.

In section 3.2.4 there is a description of how the identification of the different particles is made in ATLAS and which working points (WPs) can be used to choose the different kind of particles.

For the hadronic tau $\tau_{had-vis}$, we have decided to work with a *Medium* WP that has a signal efficiency of 75% and 60% for 1-prong and 3-prong hadronic taus respectively and a background rejection[1] of 70 and 240 for 1-prong and 3-prong [44]. In the case of muons where they will be very energetic due to the trigger, we choose to identify them using a *Tight* WP which maximizes the selection purity and gave a reconstruction efficiency for a $t\bar{t}$ MC sample of $\varepsilon_\mu = 89.9\%$ for $4 < p_T < 20$ GeV and $\varepsilon_\mu = 91.8\%$ for $20 < p_T < 100$ GeV [38]. Since the electrons have low p_T in this analysis and there are plenty of low p_T pions in jets for background events, we choose a *Medium* WP to keep the background under control. This WP shows a signal efficiency for $Z \to ee$ of 88.26% at fixed background rejection, and a rejection of 37,700 for fixed signal efficiency [70].

During the analysis of our signal, we plotted the invariant mass of the pair l2 and l3, τ and e in our case, to see which distribution it had and we found there was a peak at $m_{e,\tau} = 0$ GeV. This was due to a double identification of true electrons and hadronic taus, meaning that the same particle was identified twice. This problem was fixed by asking the tau and electron to be separated in angular space $\Delta R_{e-\tau} > 0.2$, so the same particle can not be selected twice by two different identifications. This was not a problem with the muons because they are prompt and very well-identified by the ID and MS.

Fig. 9.1 shows the angular distance $\Delta R_{e-\tau}$ between the electron and the muon and the invariant mass $m_{e,\tau}$ of these two leptons before and after imposing the cut $\Delta R_{e-\tau} > 0.2$. After imposing the cut, we end up having a reasonable invariant mass without the peak at 0 GeV.

---

[1]The rejection R is usually given as one over the efficiency, $\varepsilon = \frac{1}{R}$.
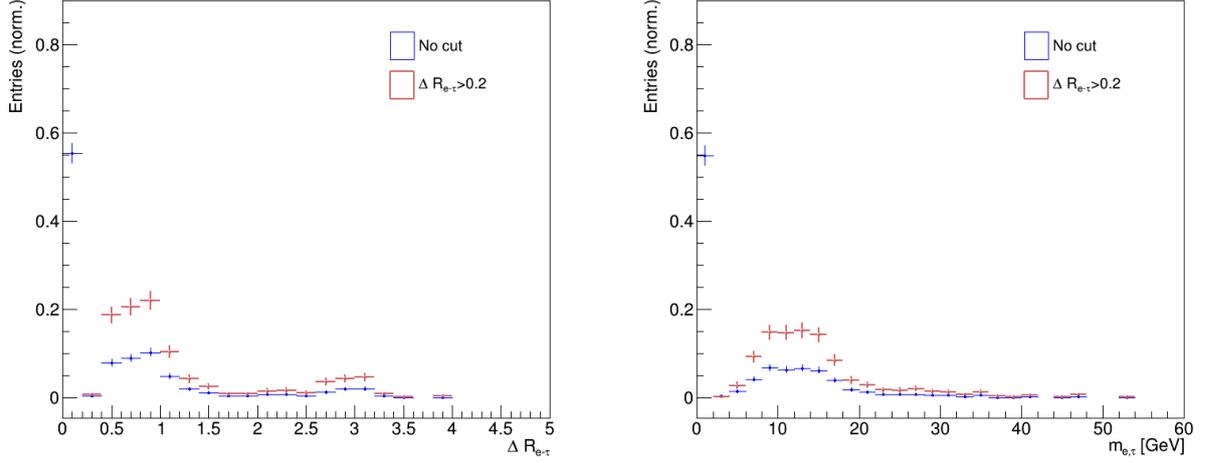
Figure 9.1: On the left, ΔR between e and τ is shown, and on the right, the invariant mass of the two leptons $m_{e,\tau}$, before and after imposing the condition of $\Delta R_{e-\tau} > 0.2$ for $m_N = 20$ GeV.

An additional requirement has been used between μ and τ, $\Delta R_{\mu-\tau} > 0.6$. Since we noticed that this cut was making us lose a significant amount of background while the signal was not affected much (see for example Tab. 9.4 and 9.6).

In addition, for the signal we asked for a minimum $\Delta R_\tau < 1$ between the reconstructed and the truth tau, since we now have the TRUTH information and we did not want to select the incorrect tau when training the boosted decision tree subsequently. It has only been asked for the τ because it is the worst identified lepton, as it can be seen in Fig. 9.3, where it is represented the angular distance between the truth tau and the selected one. There, we can see how some taus have $\Delta R_\tau > 1$ and we do not want to consider those for training the boosted decision tree. Obviously, this is not a cut we would apply when analysing real data.

Once we have imposed cuts on p_T and η due to detector limitations and the triggers, and we have asked for an identification of the leptons, we have to consider that sometimes we will have more than one lepton of each type that survives this early preselection. In Fig. 9.2 we can see how many leptons of each type survive, being the taus the ones that have more candidates.

Some searches for the best way to select the best candidate have been done. For instance, choosing based on the angular distribution between the three particles or selecting the candidate with greater p_T. However, none of the approaches shows any significant difference and therefore, we have decided to choose the candidate with the maximum
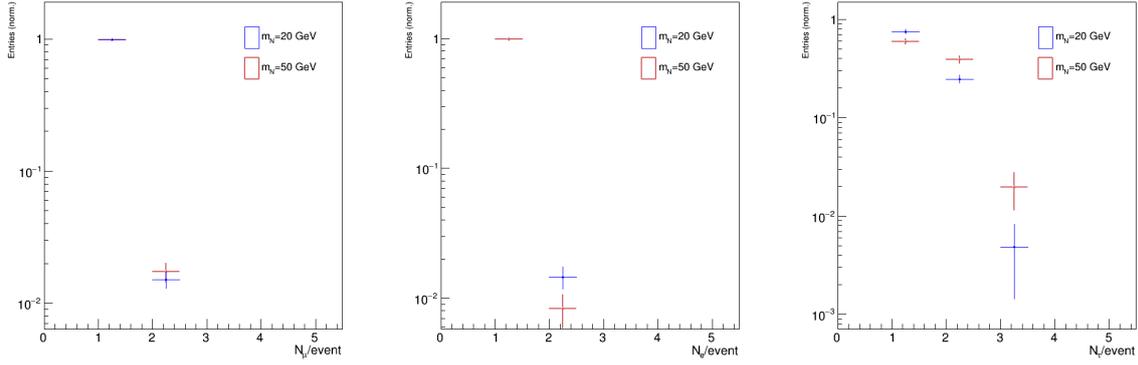
Figure 9.2: Histograms with the number of leptons left after all the preselection has been done. From left to right there is muons, electrons and taus.

transverse momentum, p_T. In Fig. 9.3 we plot the ΔR between the true tau and the tau selected based on the p_T. Bigger ΔR here mean we are not picking the right tau. Hence this plot is important and helpful to justify that maximum p_T is a good choice for selecting the remaining candidates. We can see that even if ΔR has some values bigger than 1, it only accounts for approximate 1% for $m_N = 20$ GeV and a 3% for $m_N = 50$ GeV.
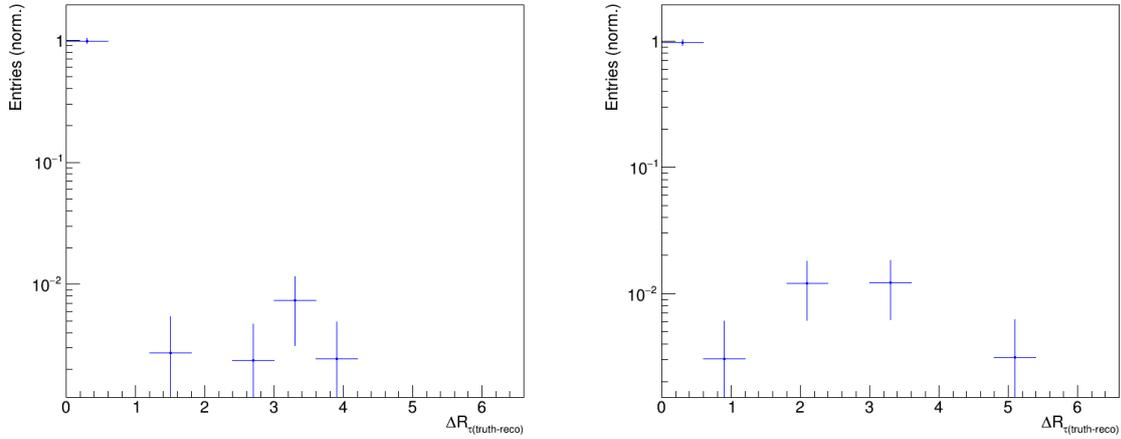
Figure 9.3: Histograms with the angular distance ΔR between the true and the selected tau. On the left we represent it for $m_N = 20$ GeV and on the right for $m_N = 50$ GeV.

The last requirement to select the events is to ask for the charge expected for the leptons. In the LNC case, we know that the charges must be $+-+$ or $-+-$, respectively for μ, τ and e, see Fig. 7.1. On the other hand, for LNV processes we expect $++-$ or $--+$. This requirement allows us to get rid of a lot of background.

9.2 Efficiency of preselection

We will show the final efficiencies once the preselection has been chosen.

On Tab. 9.4 and 9.5 efficiencies of the signal for LNC and LNV preselection, respectively are listed. In Tab. 9.6 and 9.7 we found the efficiencies for the background, again when asking for a LNC and LNV preselection.

For the signal we can find the following intermediate steps: generator filter efficiency indicated by GenFiltEff, events generated, events filtered for analysis (DAOD_SUSY3), acceptance cut applied to truth leptons, p_T mu trigger at the truth level, μ preselection, μ trigger, e preselection, τ preselection, $\Delta R_{e-\tau} > 0.2$, $\Delta R_{\mu-\tau} > 0.6$, $\Delta R_\tau < 1$ between truth and preselected τ, and charge requirement. The reason for considering taus only matching the true taus is to pick the best objects for the later machine learning algorithm.

For the background we use only the reconstructed objects, asking for the same requirements except for the last one on true tau matching.

In the following tables (Tab. 9.4, 9.5, 9.6 and 9.7), we compute the efficiency for signal and background for each of the requirements and the total efficiency of this preselection, including the generator filter efficiency.

To estimate the statistical error of the efficiency, the error of a Poisson distribution was computed to get the intermediate steps derived from [71]. Considering n the total events before the cut with an error $\delta n = \sqrt{n}$, and k the events after the cut with a corresponding error of $\delta k = \sqrt{k}$, the error efficiency is

$$\delta\varepsilon = \sqrt{\frac{k(n+k)}{n^3}}. \tag{9.1}$$

In the limit where $k \ll n$, Poisson error is not a reasonable consideration. Therefore, for the total efficiency a binomial error is used

$$\delta\varepsilon' = \frac{1}{n}\sqrt{k(1 - k/n)}. \tag{9.2}$$

	$\mu\tau e$ 20 LNC		$\mu\tau e$ 50 LNC	
	# events	ε	# events	ε
GenFiltEff		1		1
Total events	162,000		180,000	
Events DAOD_SUSY3	26,879	0.166 ± 0.001	37,895	0.211 ± 0.001
Truth acceptance: $\tau_{had} + p_T + \eta$	3,632	0.135 ± 0.002	6,881	0.182 ± 0.002
Truth trigger μ	2,956	0.81 ± 0.02	2,254	0.328 ± 0.008
μ: Tight, highest p_T	2,912	0.99 ± 0.03	2,219	0.98 ± 0.03
Trigger μ	2,874	0.99 ± 0.03	2,215	1.00 ± 0.03
e: Medium	1,739	0.61 ± 0.02	1,491	0.67 ± 0.02
τ: Medium	1,057	0.61 ± 0.02	993	0.67 ± 0.03
$\Delta R_{e-\tau} > 0.2$	1,025	0.97 ± 0.04	803	0.81 ± 0.04
$\Delta R_{\mu-\tau} > 0.6$	407	0.40 ± 0.02	324	0.40 ± 0.03
$\Delta R_{\tau(truth-sel)} < 1$	401	0.99 ± 0.07	312	0.96 ± 0.08
Q: $Q_{\mu\neq\tau}$, $Q_{\tau\neq e}$ $(+-+/-+-)$	398	0.99 ± 0.07	309	0.99 ± 0.08
Total		0.0025 ± 0.0002		0.0017 ± 0.0001

Table 9.4: Table with the efficiencies ε of each preselection step for the signal $\mu\tau e$ in LNC preselection.

	$\mu\tau e$ 20 LNV		$\mu\tau e$ 50 LNV	
	# events	ε	# events	ε
GenFiltEff		1		1
Total events	1,770,000		1,100,000	
Events DAOD_SUSY3	422,822	0.2389 ± 0.0004	272,091	0.2474 ± 0.0005
Truth acceptance: $\tau_{had} + p_T + \eta$	44,147	0.1044 ± 0.0005	44,805	0.1647 ± 0.0008
Truth trigger μ	36,382	0.824 ± 0.006	14,111	0.315 ± 0.003
μ: Tight, highest p_T	35,821	0.985 ± 0.007	13,839	0.98 ± 0.01
Trigger μ	35,417	0.989 ± 0.007	13,198	0.95 ± 0.01
e: Medium	23,529	0.664 ± 0.006	9,500	0.72 ± 0.01
τ: Medium	14,481	0.615 ± 0.007	6,673	0.70 ± 0.01
$\Delta R_{e-\tau} > 0.2$	10,339	0.714 ± 0.009	4,746	0.71 ± 0.01
$\Delta R_{\mu-\tau} > 0.6$	3,909	0.378 ± 0.007	1,886	0.40 ± 0.01
$\Delta R_{\tau(truth-sel)} < 1$	3,763	0.96 ± 0.02	1,791	0.950 ± 0.03
Q: $Q_{\mu=\tau}$, $Q_{\tau\neq e}$ $(++-/--+)$	3,673	0.98 ± 0.02	1,760	0.983 ± 0.03
Total		0.00208 ± 0.00003		0.00160 ± 0.00004

Table 9.5: Table with the efficiencies ε of each preselection step for the signal $\mu\tau e$ in LNV preselection.

9.3 Signal and background production after preselection

Using the computed cross section in Tab. 8.2 and 8.3 and efficiencies from Tab. 9.4 to 9.7, we can compute the signal rate for a choice of mixing angles and an integrated luminosity, and a background rate for an integrated luminosity. The higher the luminosity, the easier it will be to detect an HNL with smaller mixing angles. As an example, we will use the integrated luminosity of Run-2 $\mathcal{L} = 139 \ fb^{-1}$, but when the Run-3 and the

	$W \to \mu\nu_\mu$		$W \to e\nu_e$	
	# events	ε	# events	ε
GenFiltEff		0.82465		0.825
Total events	97,624,850		21,278,251	
Events DAOD_SUSY3	6,209,924	0.06361 ± 0.00003	11,747,676	0.4700 ± 0.0002
μ: Tight, highest p_T	5,918,124	0.9530 ± 0.0005	74,654	0.00635 ± 0.00002
trigger μ	5,064,550	0.8558 ± 0.0005	154	0.0021 ± 0.0002
e: Medium	64,993	0.0128 ± 0.0001	121	0.8 ± 0.1
τ: Medium	23,129	0.356 ± 0.003	83	0.69 ± 0.01
$\Delta R_{e-\tau} > 0.2$	13,170	0.57 ± 0.01	5	0.06 ± 0.03
$\Delta R_{\mu-\tau} > 0.6$	1,821	0.079 ± 0.003	1	0.3 ± 0.3
Q: $Q_{\mu \neq \tau}$, $Q_{\tau \neq e}$ $(+-+/-+-)$	628	0.05 ± 0.02	0	0
Total		$(5.3 \pm 0.3) \cdot 10^{-6}$		0

Table 9.6: Table with efficiencies of each preselection step for the background in LNC preselection.

	$W \to \mu\nu_\mu$		$W \to e\nu_e$	
	# events	ε	# events	ε
GenFiltEff		0.82465		0.825
Total events	97,624,850		21,278,251	
Events DAOD_SUSY3	6,209,924	0.06361 ± 0.00003	11,747,676	0.4700 ± 0.0002
μ: Tight, highest p_T	5,918,124	0.9530 ± 0.0005	74,654	0.00635 ± 0.00002
trigger μ	5,064,550	0.8558 ± 0.0005	154	0.0021 ± 0.0002
e: Medium	64,993	0.0128 ± 0.0001	121	0.8 ± 0.1
τ: Medium	23,129	0.356 ± 0.003	83	0.69 ± 0.01
$\Delta R_{e-\tau} > 0.2$	13,170	0.57 ± 0.01	5	0.06 ± 0.03
$\Delta R_{\mu-\tau} > 0.6$	1,821	0.079 ± 0.003	1	0.3 ± 0.3
Q: $Q_{\mu = \tau}$, $Q_{\tau \neq e}$ $(++-/--+)$	350	0.03 ± 0.01	0	0
Total		$(3.0 \pm 0.2) \cdot 10^{-6}$		0

Table 9.7: Table with efficiencies of each preselection step for the background in LNV preselection.

HL-LHC starts one will explore a bigger range of the parameter space $|U_{tot}(m_N)|^2$ since the integrated luminosity will increase to $\mathcal{L} = 3000\ fb^{-1}$.

Tab. 9.8 shows us the rate for the two benchmarks (4 and 6) selected using Tab. 8.2, with the integrated luminosity of Run-2 with the minimum decay width in order to obtain signal. We can observe that even having nine orders of magnitude less for σ in signal than in background, the preselection allows us to reduce the difference to four orders of magnitude for the rate. Also, we can see how the benchmark 6 gives a slightly better rate for the signal than the benchmark 4.

In this stage, we still have three to four orders of magnitude more background than signal. In the following chapter, we will apply a boosted decision tree to distinguish the signal from the background, hoping to improve the rates shown in Tab. 9.8.

	$|U_{\alpha,N1}|$	$c\tau_N$ [mm]	σ [fb]	$\varepsilon_{preselection}$	$Y = \sigma\mathcal{L}\varepsilon$
$\mu\tau e$ 20 LNC	$|U_\tau| = 2.18 \cdot 10^{-2}$; $|U_\mu| = 9.15 \cdot 10^{-3}$	10^{-4}	46.7	$(2.5 \pm 0.2) \cdot 10^{-3}$	16 ± 2
	$|U_\mu| = |U_\tau| = 1.67 \cdot 10^{-2}$	10^{-4}	90.4	$(2.5 \pm 0.2) \cdot 10^{-3}$	31 ± 3
$\mu\tau e$ 50 LNC	$|U_\tau| = 2.20 \cdot 10^{-2}$; $|U_\mu| = 9.25 \cdot 10^{-3}$	10^{-6}	23.61	$(1.7 \pm 0.1) \cdot 10^{-3}$	5.6 ± 0.4
	$|U_\mu| = |U_\tau| = 1.69 \cdot 10^{-2}$	10^{-6}	46.2	$(1.7 \pm 0.1) \cdot 10^{-3}$	11.0 ± 0.7
$W \to \mu\nu$			$19.2 \cdot 10^6$	$(5.3 \pm 0.3) \cdot 10^{-6}$	$14,157 \pm 801$
$\mu\tau e$ 20 LNV	$|U_\tau| = 2.18 \cdot 10^{-2}$; $|U_\mu| = 9.15 \cdot 10^{-3}$	10^{-4}	46.7	$(2.08 \pm 0.03) \cdot 10^{-3}$	13.5 ± 0.2
	$|U_\mu| = |U_\tau| = 1.67 \cdot 10^{-2}$	10^{-4}	90.4	$(2.08 \pm 0.03) \cdot 10^{-3}$	26.1 ± 0.4
$\mu\tau e$ 50 LNV	$|U_\tau| = 2.20 \cdot 10^{-2}$; $|U_\mu| = 9.25 \cdot 10^{-3}$	10^{-6}	23.6	$(1.60 \pm 0.04) \cdot 10^{-3}$	5.3 ± 0.2
	$|U_\mu| = |U_\tau| = 1.69. \cdot 10^{-2}$	10^{-6}	46.3	$(1.60 \pm 0.04) \cdot 10^{-3}$	10.3 ± 0.3
$W \to \mu\nu$			$19.2 \cdot 10^6$	$(3.0 \pm 0.2) \cdot 10^{-6}$	$7,890 \pm 534$

Table 9.8: Table with mixing angle $U_{\alpha,N1}$, decay width of HNL Γ_N, cross section σ, total efficiency ε and rate Y using a luminosity of $\mathcal{L} = 139\ fb^{-1}$. The data is for both LNC and LNV preselection.

Boosted Decision Tree

<div style="text-align: right">10</div>

10.1 Introduction

A Boosted Decision Tree (BDT) is one of the simplest machine learning techniques based on a decision tree architecture. It is often used in physics for classification problems, such as distinguishing the signal from the background, due to its simplicity, performance and interpretation. LightGBM [72] has been used to help us select our signal.

Decision trees [73] divide the parameter space separating the data into classes in a problem (for example if we are trying to classify signal and background, these will be our classes). It asks questions about the value of features, splitting the data until a specific region in this parameter space is all of the same class. This is a basic structure where internal nodes ask a question regarding the value of some variables of our data. The nodes at the end of the tree are called leaves, and it is where we stop dividing our data any further. In Fig. 10.1 the nodes are represented in red and the leaves in blue or orange, depending on the predicted class they are part of.
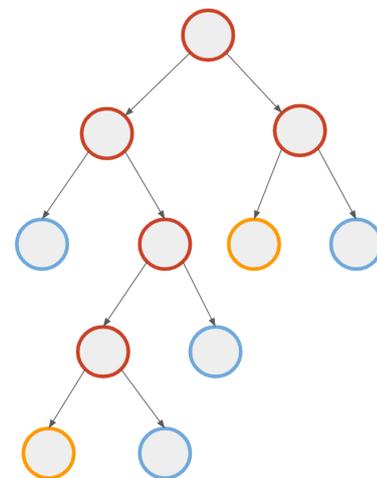
Figure 10.1: Schematic representation of a Decision Tree.

Once the tree is completed, we can create another tree by giving more importance, or weight, to those regions where the error has been bigger, boosting the performance of the algorithm by focusing on the more problematic parts of the parameter space. We rerun until the classifier performance is increased, hence the name of boosted decision tree [73]. In the end, we will have an ensemble of classifiers or forest and we will classify according to what the majority of classifiers vote.

In gradient BDTs [73, 74] the loss is minimized as the model is fit. Each tree can give a continuous score on each leaf, when all trees are compared one takes into account how sure each tree is about its prediction. This method was introduced by the model XGBoost [75] and used later for the model we have used, LightGBM.

LightGBM [72] or Light Gradient Boosted Machine has many similarities with XGBoost but it has a different approach to construct the trees. It chooses the leaf which decreases the loss the most, rather than implementing new rows of nodes. We then can have a tree with different longitude of "branches". This allows the algorithm not to waste time on unnecessary extra splits.

LightGBM has been used for its speed while maintaining high accuracy and for its scalability, the amount of data can be increased without significant changes. This would be very useful for further analysis when considering more background types.

10.2 Signal vs Background

We have trained four different models for $m_N = 20, 50$ GeV and for LNC and LNV preselection for each case. For the training we have given the same number of signal and background events to each algorithm, Tab. 10.1. We do not give more background than signal so the algorithm does not learn due simply to the fact of having more background than signal. For each case, we have given the most number of events available. For LNC preselection, the signal was limiting the available events. On the other hand, the background was the limiting sample for LNV. For this last case, we used both backgrounds coming from LNC and LNV preselection, after carefully studying that there were no kinematic correlations between them.

	LNC		LNV	
	$m_N = 20$ GeV	$m_N = 50$ GeV	$m_N = 20$ GeV	$m_N = 50$ GeV
# Signal	398	309	978	978
# Backgound	398	309	978	978

Table 10.1: Number of events used to train the models for signal and for background.

The models have been trained with twelve variables: p_T, η and ϕ of the three leptons (τ, μ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and missing energy transversed (E_T^{miss})[1].

---

[1]The transverse energy need to be conserved in the collision, the missing transverse energy accounts for this giving an idea of the transverse energy carried by the neutrinos.

10.2.1 Choice of the variables

It is important that we understand which variables can give information to the BDT and use only those. Even if the architecture is simple, it becomes a tedious process to understand how the BDT splits the data when having a large amount of variables. For this, we have plotted all the variables which could give us some information in scatter plots. First, we show in Tab. 10.2 the variables that present more correlation between them and are good candidates to help select the signal. With this plot in mind one can go to Fig. 10.2 and 10.3 to explicitly observe these correlations when the signal was selected for LNV preselection. In the Appendix B, we show the same plots for the case with LNC preselection. We have added this case in the appendix since the difference between the two cases is that the sample corresponding to LNV had more data and allowed us to identify the correlations easier.

	p_T^μ	η_μ	ϕ_μ	p_T^τ	η_τ	ϕ_τ	p_T^e	η_e	ϕ_e	$m_{\tau,e}$	$m_{\mu,\tau,e}$	E_T^{miss}
p_T^μ										★	★	★
η_μ					★			★				
ϕ_μ						●			●			
p_T^τ										★	★	
η_τ		★						★				
ϕ_τ			●						●			
p_T^e										★	★	
η_e		★			★							
ϕ_e			●			●						
$m_{\tau,e}$	★			★			★			★	★	★
$m_{\mu,\tau,e}$	★			★			★			★	★	★
E_T^{miss}	★									★	★	★

Table 10.2: Schema of the variables that present more correlation. "★" shows high correlation for $m_N = 20, 50$ GeV, while "●" only shows high correlation for $m_N = 20$ GeV.

For Fig. 10.2, the scatter plots between pseudorapidities of different leptons have a clear predominance in the diagonal of the parameter space for signal, hinting that they have a similar direction in η while the background is spread through all the space. A similar distribution happens between the azimuthal angles of tau and the electron ϕ_τ-ϕ_e. On the other hand, for ϕ_μ-ϕ_τ and ϕ_μ-ϕ_e the signal tends to be off-diagonal, meaning that they are separated in space, while the background is spread through all the space. It is important to mention also how helpful the histogram of E_T^{miss} is, having two distinct peaks.

In the case of higher masses $m_N = 50$ GeV, Fig. 10.3, there is no correlation for the azimuthal angle between the pair of leptons as there was for $m_N = 20$ GeV. We also observe that for $m_N = 50$ GeV, the histogram of invariant mass $m_{\tau,e}$ has a broader peak for signal, more alike than the background. These tells us that the higher the mass, the more difficult it will be to distinguish the signal from the background.

We can observe that all the variables shown here are relevant and can help the BDT to distinguish between signal and background. Therefore they will be used as input features for the BDT.

10.3 Train the models

The hyperparameters of the model are those we can set before training and that define the architecture of our algorithm. The hyperparameters have been set using a Bayesian Optimization [76], to get the optimal ones that minimize our loss function. The basic idea is that instead of randomly checking a set of values for the hyperparameters, every evaluation of the model keeps track of the past evaluations so that the minima can be more easily reached. We have optimized the "num_leaves", the "max_depth" and the "learning_rate", while trying to minimize a binary log-loss function. The "num_leaves" sets a maximum of the number of leaves, while "max_depth" imposes a maximum of how deep a branch can be. At last, "learning_rate" refers to the step size for approaching the minimum of a loss function. If the "learning_rate" is too big is possible that we skip the minimum, while if the value is too small we will approach the minimum too slowly.

Algorithm	num_leaves	max_depth	learning_rate
20 LNC	124	187	0.05526
50 LNC	82	42	0.03918
20 LNV	122	174	0.05808
50 LNV	45	98	0.04323

Table 10.3: List of the three hyperparameters optimized for the four different algorithms. We have used that 20 and 50 are $m_N = 20$ GeV and $m_N = 20$ GeV respectively, and LNC and LNV is the preselection we have done in the events.

Four different algorithms will be trained for the four different conditions we are considering, the signal could have LNC or LNV preselection, and could have $m_N = 20, 50$ GeV. For each model, we have listed in Tab. 10.3 the optimized hyperparameters using Bayesian Optimization.
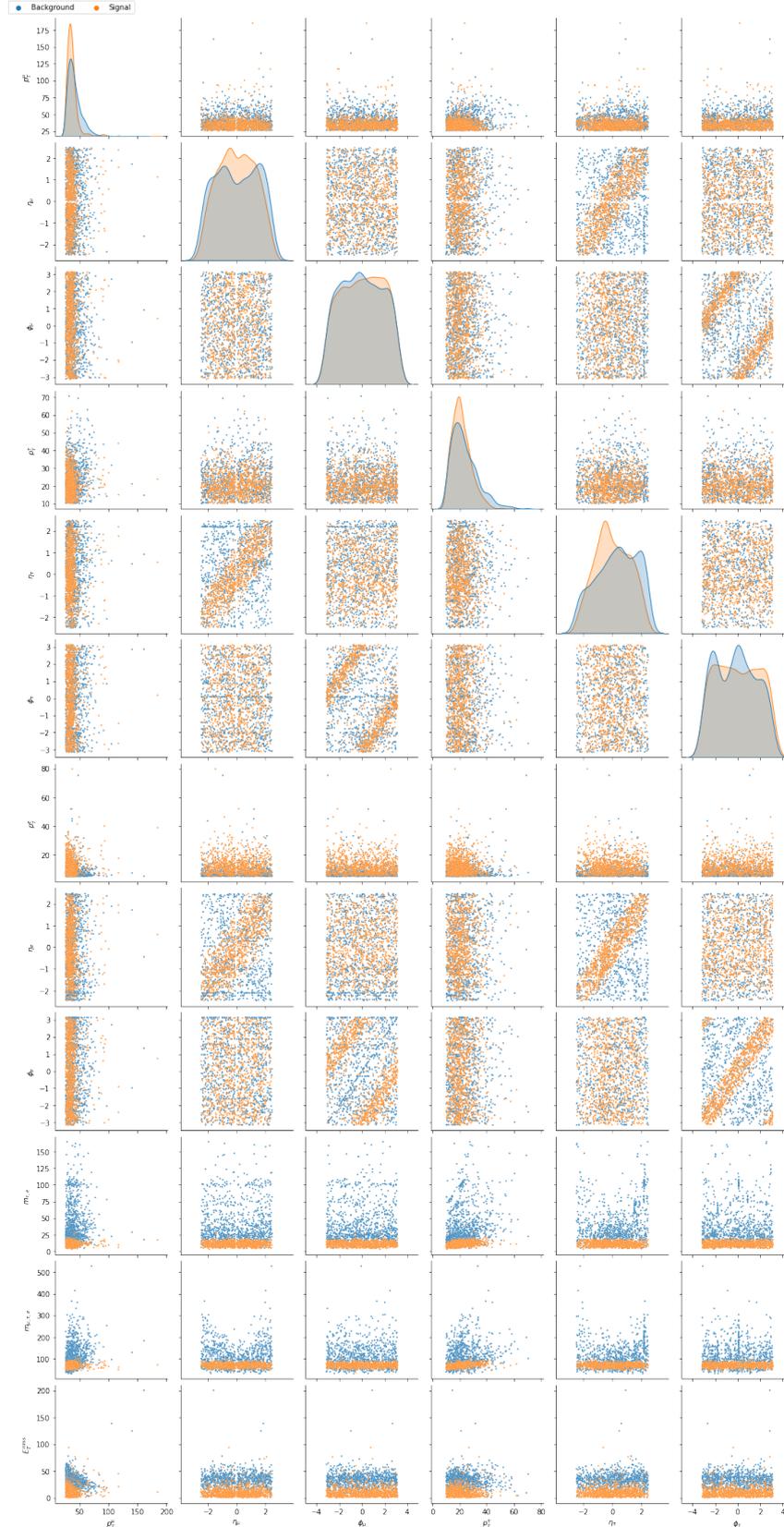
Figure 10.2: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $\boldsymbol{m_N = 20}$ **GeV LNV** preselection. The figure is split in two for better visibility, continuing in next page.

Figure 10.2: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $\boldsymbol{m_N = 20}$ **GeV LNV** preselection. (continuation)
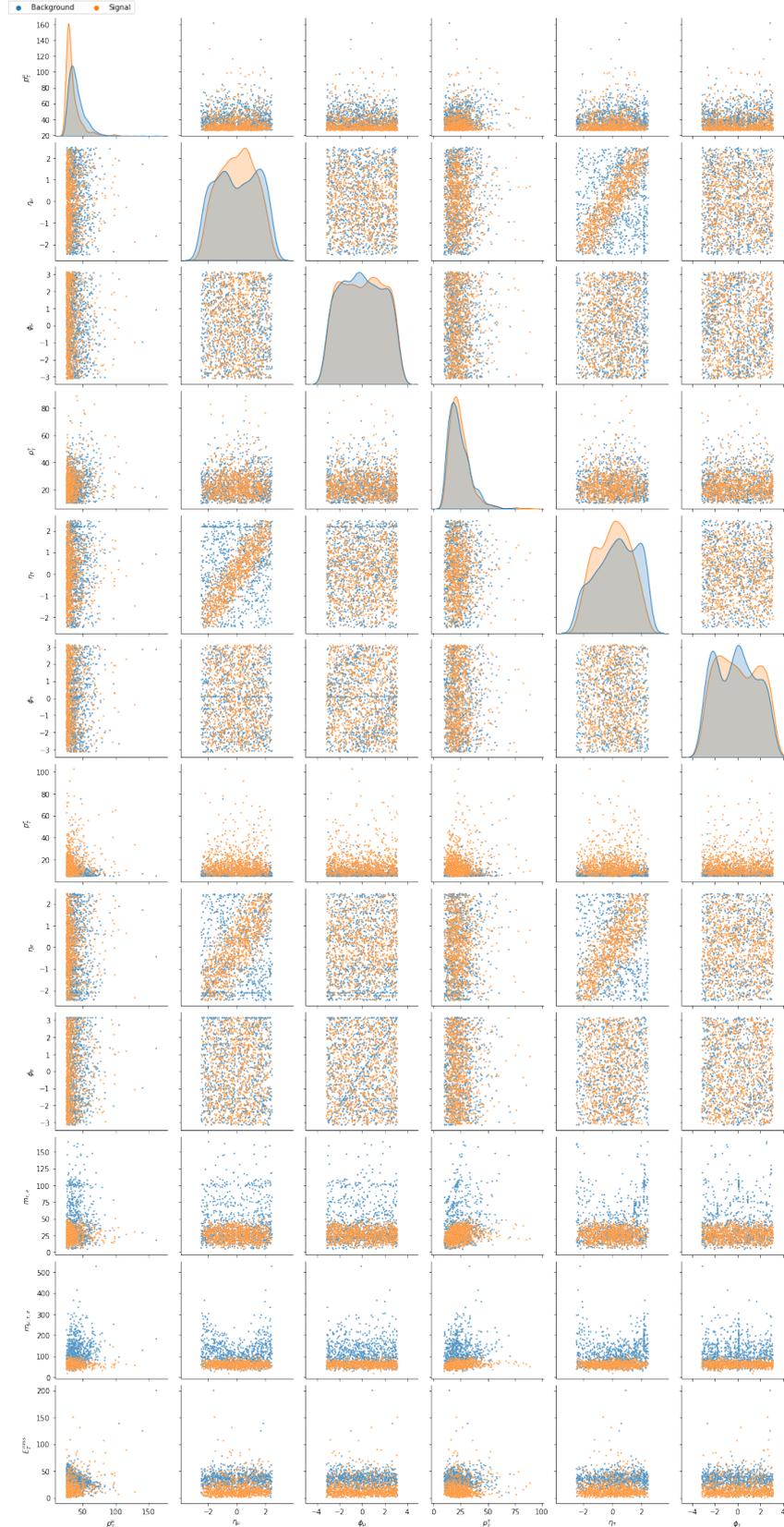
Figure 10.3: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $\boldsymbol{m_N = 50}$ **GeV LNV** preselection. The figure is split in two for better visibility, continuing in next page.

Figure 10.3: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $\boldsymbol{m_N = 50}$ **GeV LNV** preselection. (continuation)

Once the model is trained, we will perform a validation with 25% of the data for each model. We will get a score for each of the events, the closer to 1 it is, the surest the algorithm is that the event is signal. We can decide which minimum score will be the cut value. From the truth information of which group each event is part of, we can study the performance of the model.
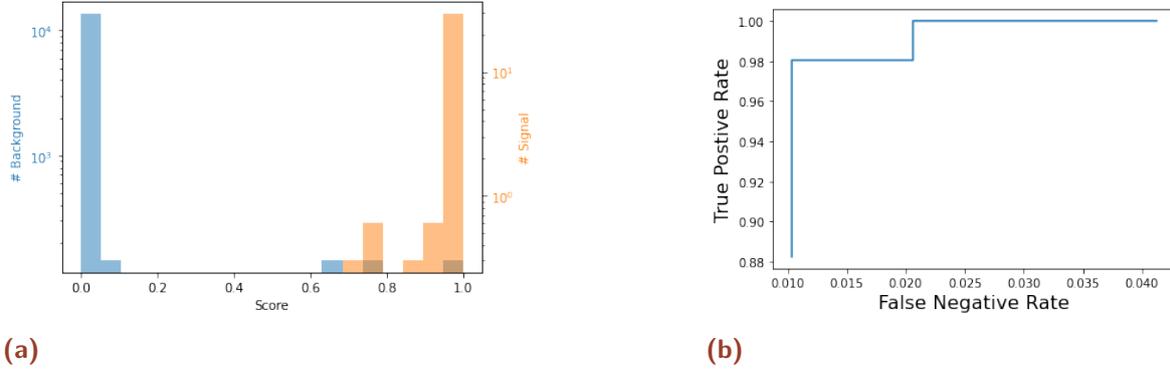
(a)

(b)

Figure 10.4: For $m_N = 20$ **GeV LNC** preselection, we have on the left the score given by the BDT, having the background in blue and the signal in orange; and on the right, rate of true positive compared to false negative when we scan between 0.5 and 1 the minimum value of BDT score required to be classified as signal.

In Fig. 10.4a we can see the performance of the BDT after hyperparameter optimization and training of $m_N = 20$ GeV for LNC case. We have scaled both background and signal using Tab. 9.8 numbers for Y, so a realistic comparison between signal and background can be performed. In Fig. 10.4b we have scanned the BDT score to see how much efficiency we have in selecting signal (True Positive Rate or ε_s) and how many background we select by mistake (False Negative Rate or ε_b). For example, if we cut at 0.90 in score: $\varepsilon_s = 0.961$ and then $\varepsilon_b = 0.010$. The low statistics for LNC preselection case does not show us how the efficiencies vary slowly, having only three possible values for ε_b. A better understanding of the results can be obtained from the case of LNV preselection.

In Fig. 10.5a, for $m_N = 50$ GeV, the score is more spread out, hence the algorithm does not have as clear as before the distinction between background and signal. We can also see this in Fig. 10.5b where now with a cut at 0.90 in score the performance is: $\varepsilon_s = 0.667$ and $\varepsilon_b = 0.024$.

Comparing Fig.10.4a and 10.5a, we can observe that the BDT for 20 GeV performs better in distinguishing the signal from the background. This is expected by the relations we found previously between masses looking at the clustering plots (Fig. 10.2 and 10.3).
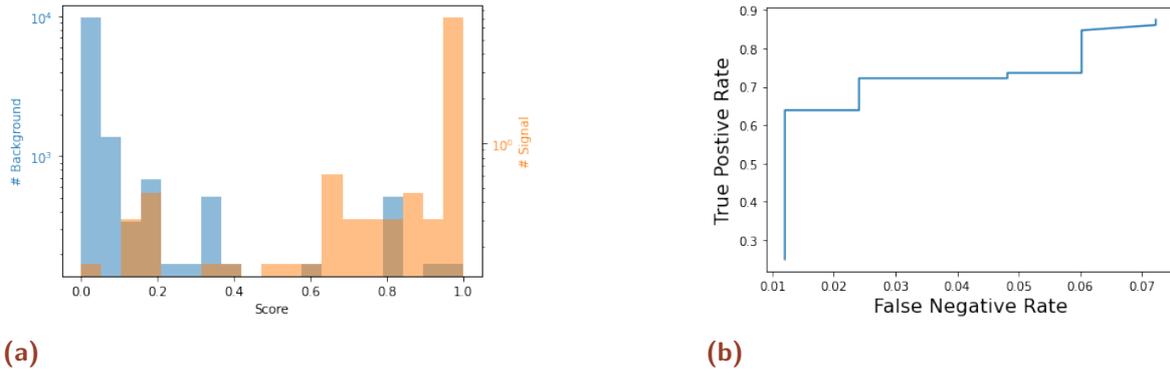

Figure 10.5: For $m_N = 50$ **GeV LNC** preselection, we have on the left the score given by the BDT, having the background in blue and the signal in orange; and on the right, rate of true positive compared to false negative when we scan between 0.5 and 1 the minimum value of BDT score required to be classified as signal.

When the HNL mass is higher, the three leptons do not have such a clear distribution in space.

In Fig. 10.6 and 10.7 the same plots are represented by the case where we are training the models from samples that have an LNV preselection. Even with better statistics, similar results are found for these algorithms, showing that LightGBM is suitable for small data and can train the model with little data.
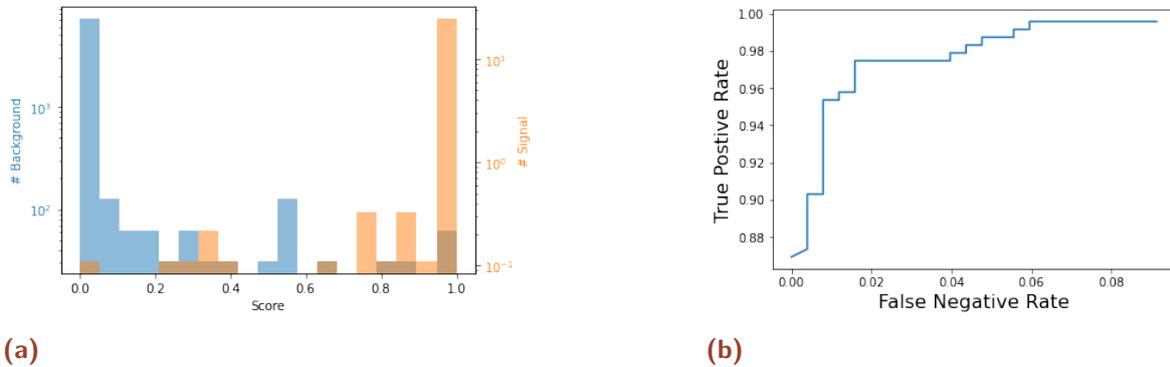

(a)   (b)

Figure 10.6: For $m_N = 20$ **GeV LNV** preselection, we have on the left the score given by the BDT, having the background in blue and the signal in orange; and on the right, rate of true positive compared to false negative when we scan between 0.5 and 1 the minimum value of BDT score required to be classified as signal.

10.4 Efficiencies

In Tab. 10.4, there is a list of signal and background efficiencies for the four algorithms using two different working points (WPs), 0.90 and 0.99. From these numbers, we can
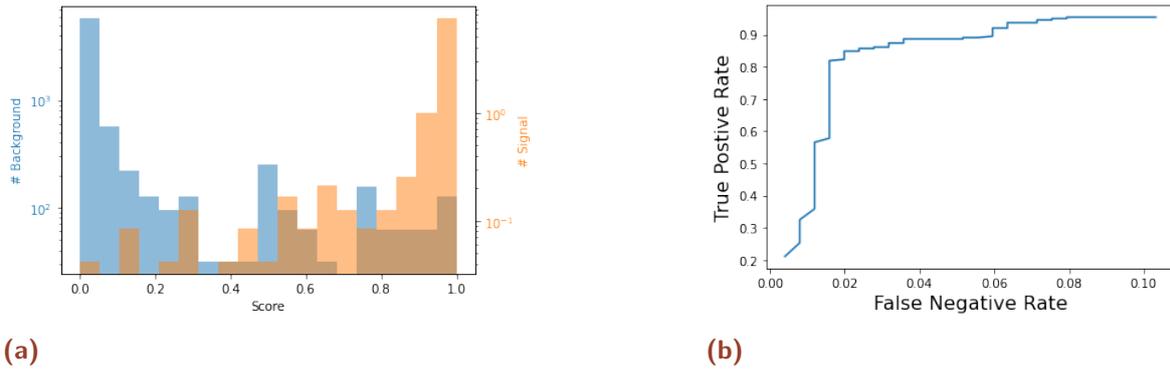
(a)   **(b)**

Figure 10.7: For $m_N = 50$ **GeV LNV** preselection, we have on the left the score given by the
BDT, having the background in blue and the signal in orange; and on the right,
rate of true positive compared to false negative when we scan between 0.5 and 1
the minimum value of BDT score required to be classified as signal.

see how the algorithm works similarly for LNC and LNV cases, even with more events
to train in LNV.

	LNC				LNV			
	$m_N = 20$ GeV		$m_N = 50$ GeV		$m_N = 20$ GeV		$m_N = 50$ GeV	
WP	0.90	0.99	0.90	0.99	0.90	0.99	0.90	0.99
ε_{Signal}	0.961	0.873	0.667	0.222	0.941	0.857	0.848	0.143
$\varepsilon_{Background}$	0.010	0.010	0.024	0.012	0.008	0.000	0.024	0.004

Table 10.4: Efficiencies with different cuts in the required score (WP) when selecting signal
on the different BDTs depending on the m_N and LNC/LNV preselections.

The results obtained for $m_N = 20$ GeV are very promising for distinguishing the signal
from the background. It performs better than $m_N = 50$ GeV, expected from having less
correlation between particles (see Tab. 10.2). Actually, it is for the case of $m_N = 50$
GeV, where the algorithms differ the most between LNC and LNV preselections.

The WP of 0.99 in score of the BDT shows a very bad efficiency, mostly for high mass
but also for lower. For this reason we choose to work with 0.90 for the next chapter,
where we will put in context the results obtained.

Discussion

11

11.1 Conclusion

Let us summarise here the efficiencies from the preselection (chapter 9) and the ones from the selection of the BDT (chapter 10). We can then compare the exclusion limit using our signal, for different integrated luminosities of ATLAS upgrades, with other experiments to show if we would exclude a wider range of the mixing angle as a function of the HNL mass.

In Tab. 11.1 an example of the rate Y of events selected of signal and background is summarized for the four different algorithms using the integrated luminosity of Run-2 $\mathcal{L} = 139 \ fb^{-1}$ for the benchmark 6. The LNV preselection shows a big background reduction mainly due to the low preselection efficiency. In addition, the prompt muon of the signal and the different correlation between signal and background for $m_N = 20$ GeV (Fig. 10.6) make this case the most favourable for detecting HNLs.

	$|U_\mu| = |U_\tau|$	$c\tau_N$ [mm]	σ_{LNC} [fb]	$\varepsilon_{preselection}$	ε_{BDT}	$Y = \sigma\mathcal{L}\varepsilon$
$\mu\tau e$ 20 LNC	$1.67 \cdot 10^{-2}$	10^{-4}	90.4	$(2.5 \pm 0.4) \cdot 10^{-3}$	0.961	30 ± 3
$W \to \mu\nu$			$19.2 \cdot 10^6$	$(5.3 \pm 0.2) \cdot 10^{-6}$	0.010	142 ± 8
$\mu\tau e$ 50 LNC	$1.69 \cdot 10^{-2}$	10^{-6}	46.2	$(1.7 \pm 0.3) \cdot 10^{-3}$	0.667	7.4 ± 0.5
$W \to \mu\nu$			$19.2 \cdot 10^6$	$(5.3 \pm 0.2) \cdot 10^{-6}$	0.024	340 ± 20
$\mu\tau e$ 20 LNV	$1.67 \cdot 10^{-2}$	10^{-4}	90.4	$(2.08 \pm 0.09) \cdot 10^{-3}$	0.941	24.6 ± 0.4
$W \to \mu\nu$			$19.2 \cdot 10^6$	$(3.0 \pm 0.1) \cdot 10^{-6}$	0.008	63 ± 5
$\mu\tau e$ 50 LNV	$1.69 \cdot 10^{-2}$	10^{-6}	46.3	$(1.6 \pm 0.1) \cdot 10^{-3}$	0.848	8.7 ± 0.3
$W \to \mu\nu$			$19.2 \cdot 10^6$	$(3.0 \pm 0.1) \cdot 10^{-6}$	0.024	189 ± 13

Table 11.1: Table showing the rate of events of each type generated with $\mathcal{L} = 139 \ fb^{-1}$ using the efficiency we get from the BDT for the benchmark 6. Under each of the signals (in orange), there is the corresponding background (in blue) that has been selected for the same preselection and after the BDT.

The numbers presented in Tab. 11.1 show that the search of prompt HNLs with muon and tau flavour mixing with three lepton final state is not hopeless for this value of mixing angles, having only one order of magnitude difference at most for the amount of signal and background selected. More specifically, we now want to find which exclusion limits we can set for a given luminosity with this selection strategy.

Run-3 and High Luminosity LHC (HL-LHC) upgrades will increase the integrated luminosity in ATLAS up to $\mathcal{L} = 3000 \ fb^{-1}$ [28], allowing to push down the sensitivity to HNLs. Both rates (signal and background) will increase proportionally with integrated luminosity, but it will give us sensitivity for lower mixing angles. Using Tab. 11.1, the amount of signal and background expected can be computed increasing the luminosity according to $Y = \sigma \mathcal{L} \varepsilon$.

In order to estimate the exclusion limits, we have used the significance $S(b|b + s) \geq 2$, with b and s being the amount of background and signal expected respectively. To compute the significance we have used [47],

$$S(n|h) = \sqrt{-2 \ln \frac{P(n|h)}{P(n|n)}} \quad \text{with} \quad P(n|h) = \frac{h^n}{n!} e^{-h}. \tag{11.1}$$

Having the number of events expected of background for each of the cases, we can get the minimal number of signal events needed in order to have a $S(b|b + s) \geq 2$. We can therefore get the mixing angle squared of a specific benchmark using the efficiency values of Tab. 11.1, and Eq. 8.4 and 8.5. The exclusion limits found for $\mathcal{L} = 139 \ fb^{-1}$ can be seen in Tab. 11.2 and 11.3, for benchmark 6 and 4 respectively.

	LNC		LNV	
	$m_N = 20$ GeV	$m_N = 50$ GeV	$m_N = 20$ GeV	$m_N = 50$ GeV
$|U_\mu|^2 = |U_\tau|^2$	$2.4 \cdot 10^{-4}$	$1.5 \cdot 10^{-3}$	$2.0 \cdot 10^{-4}$	$9.5 \cdot 10^{-4}$

Table 11.2: Exclusion limits of the mixing angle squared in function of each mass studied in the case of LNC and LNV preselections for $\mathcal{L} = 139 \ fb^{-1}$ using benchmark 6.

	LNC		LNV	
	$m_N = 20$ GeV	$m_N = 50$ GeV	$m_N = 20$ GeV	$m_N = 50$ GeV
$|U_\mu|^2$	$1.4 \cdot 10^{-4}$	$8.9 \cdot 10^{-4}$	$1.2 \cdot 10^{-4}$	$5.6 \cdot 10^{-4}$
$|U_\tau|^2$	$7.9 \cdot 10^{-4}$	$5.1 \cdot 10^{-3}$	$6.7 \cdot 10^{-4}$	$3.2 \cdot 10^{-3}$

Table 11.3: Exclusion limits of the mixing angle squared in function of each mass studied in the case of LNC and LNV preselections for $\mathcal{L} = 139 \ fb^{-1}$ using benchmark 4.

The same steps can be done using the amount of background expected when the luminosity of HL-LHC is reached ($\mathcal{L} = 3000 \ fb^{-1}$), giving the results from Tab. 11.4 and 11.5, for benchmark 6 and 4 respectively.

By comparing our results with previous and future bounds, we can see where our search is in the parameter space of the mixing angle squared, and check if we could exclude a new area.

	LNC		LNV	
	$m_N = 20$ GeV	$m_N = 50$ GeV	$m_N = 20$ GeV	$m_N = 50$ GeV
$|U_\mu|^2 = |U_\tau|^2$	$4.8 \cdot 10^{-5}$	$3.2 \cdot 10^{-4}$	$4.0 \cdot 10^{-5}$	$2.0 \cdot 10^{-4}$

Table 11.4: Exclusion limits of the mixing angle squared in function of each mass studied in the case of LNC and LNV preselections for $\mathcal{L} = 3000\ fb^{-1}$ using benchmark 6.

|	LNC		LNV	
	$m_N = 20$ GeV	$m_N = 50$ GeV	$m_N = 20$ GeV	$m_N = 50$ GeV
$|U_\mu|^2$	$2.8 \cdot 10^{-5}$	$1.8 \cdot 10^{-4}$	$2.3 \cdot 10^{-5}$	$1.2 \cdot 10^{-4}$
$|U_\tau|^2$	$1.6 \cdot 10^{-4}$	$1.0 \cdot 10^{-3}$	$1.3 \cdot 10^{-4}$	$6.6 \cdot 10^{-4}$

Table 11.5: Exclusion limits of the mixing angle squared in function of each mass studied in the case of LNC and LNV preselections for $\mathcal{L} = 3000\ fb^{-1}$ using benchmark 4.

To compare the mixing angles squared obtained using the significance, we should use only studies that use the same benchmark we have used from Fig. 8.2. Unfortunately, the exclusion limits we have available from other projects only account for HNL mixing with the same flavour neutrino $|U|^2 = |U_\alpha|^2$. Therefore the comparison should not be done directly, or at least it needs to be further interpreted. We have decided to represent it anyway in the same plot to compare the different mixing angles squared reached, commenting on how the results can be interpreted. In Fig. 11.1 and Fig. 11.2, we have added the lowest exclusion limit for each mass with our search at benchmark 6 and 4 respectively (in both cases this was for LNV preselection), to the bound already presented in Fig. 4.2.

Let us first focus on benchmark 6 results where $|U_\mu|^2/|U^2| = |U_\tau|^2/|U^2| = 0.5$. Fig. 11.1a (mixing with muon flavour) shows that for both luminosity periods we are inside the already set exclusion zone. On the other hand, in Fig. 11.1b (mixing with tau flavour), we are inside the limit set by DELPHI for Run-2 luminosity while being just outside for HL-LHC. These could, a priori, seem promising results, but we have to be careful here.

The HNL of the channel studied, $\mu\tau e$, first mixes with muon neutrino and then it decays to tau neutrino. In order to detect the signal, we need the HNL to mix with muon and tau neutrino with greater mixing angle than the limits set in our search.

The parameter space for the muon mixing obtained is already excluded by previous searches. Therefore, even if in Fig. 11.1b it seems that the limits are expanded, we will not be sensible to the lower mixing angle with muon.
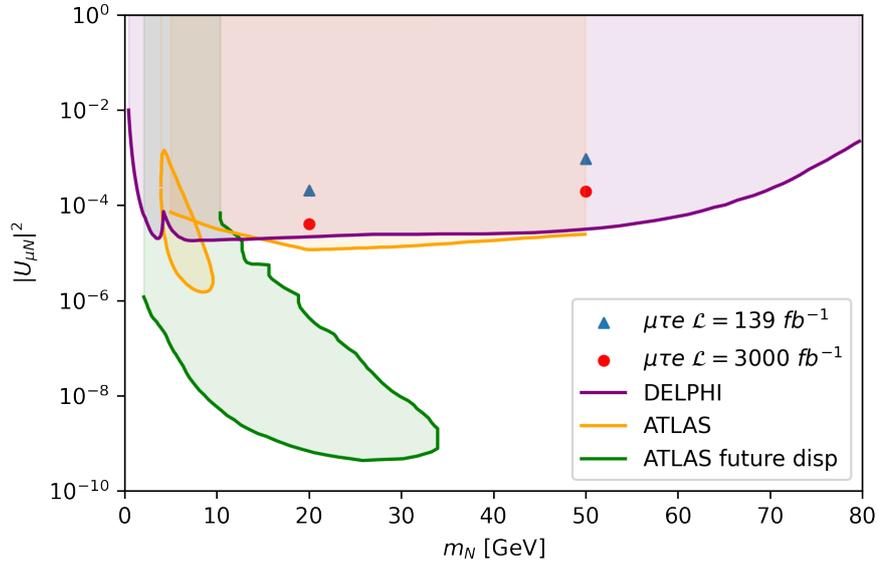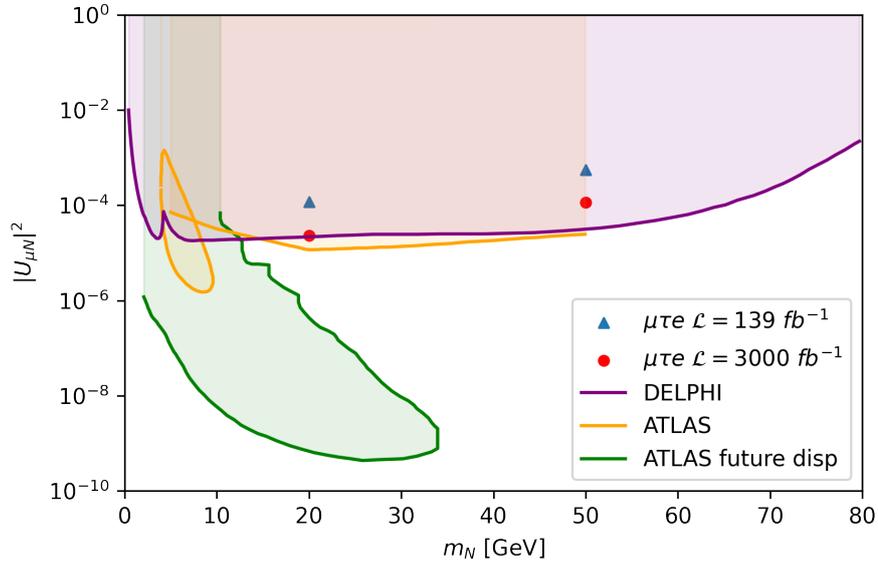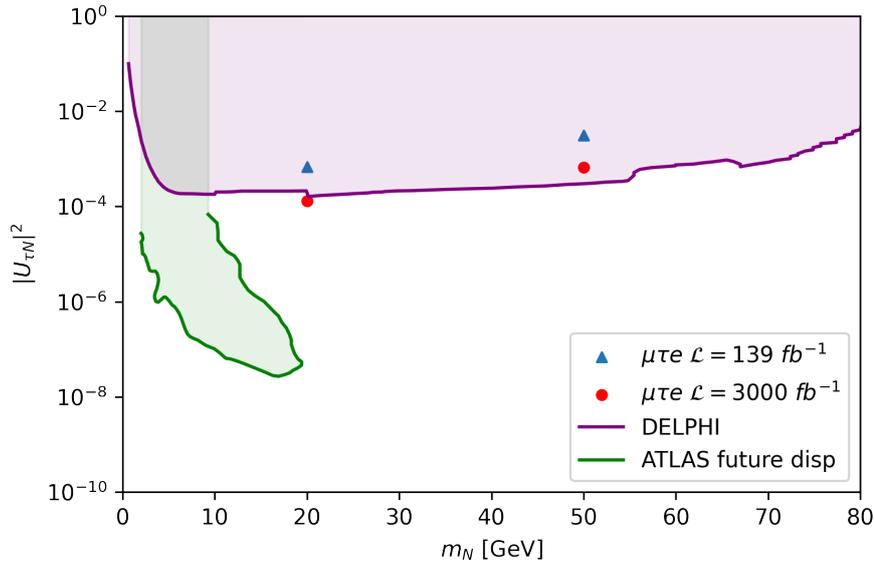
(a)
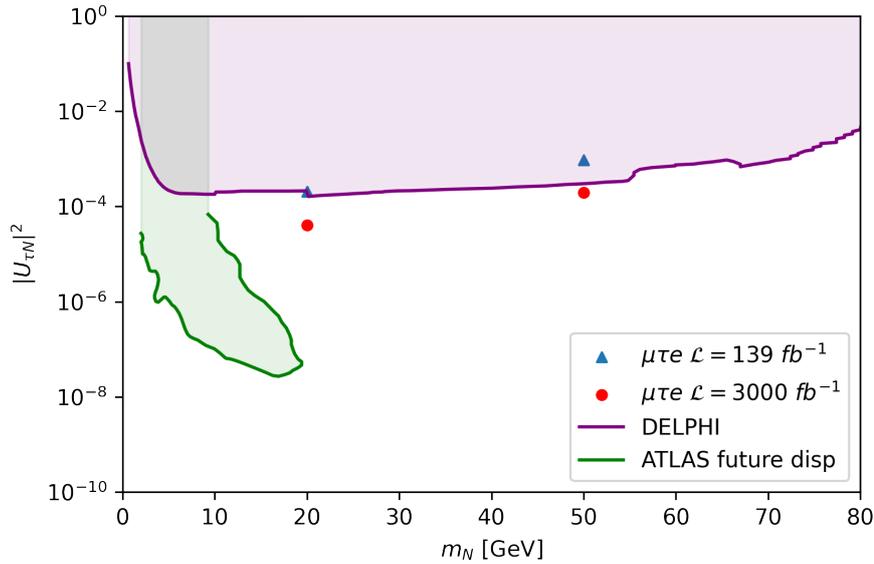
(b)

Figure 11.1: Bounds on the mixing angle squared in function of the HNL mass, for μ flavour in (a), and τ in (b). The signal $\mu\tau e$ for LNV with $\mathcal{L} = 139\ fb^{-1}$ is represented in blue and with $\mathcal{L} = 3000\ fb^{-1}$ in red, for the benchmark 6. In purple the current studies in DELPHI [48], in yellow the current studies of HNL in ATLAS [7] and in green the future constraints for displaced HNL in ATLAS [47], all considering $|U|^2 = |U_\mu|^2$ for (a), and $|U|^2 = |U_\tau|^2$ for (b).

(a)

(b)

Figure 11.2: Bounds on the mixing angle squared in function of the HNL mass, for μ flavour in (a), and τ in (b). The signal $\mu\tau e$ for LNV with $\mathcal{L} = 139\ fb^{-1}$ is represented in blue and with $\mathcal{L} = 3000\ fb^{-1}$ in red, for the benchmark 4. In purple the current studies in DELPHI [48], in yellow the current studies of HNL in ATLAS [7] and in green the future constraints for displaced HNL in ATLAS [47], all considering $|U|^2 = |U_\mu|^2$ for (a), and $|U|^2 = |U_\tau|^2$ for (b).

For the benchmark 4 where $|U_\mu|^2/|U^2| = 0.15$ and $|U_\tau|^2/|U^2| = 0.85$, we would expect our exclusion limits to go down for muon mixing, while going up for tau mixing. Fig. 11.2 shows how we are in both cases closer to the previous exclusion limits but still inside, still not being possible to extend the exclusion limits already set.

11.2 Outlook

This project intends to be the starting point of a complete analysis of all the different channels described in chapter 6. Even if the estimation of the triggers were not as promising as the channel selected by us, all the channels could contribute to having more sensitivity for detecting the HNL. These are: $e\tau e$, $e\tau\mu$, $\mu\tau\mu$, $\mu\tau e$, $\tau\mu e$, $\tau\mu\mu$, $\tau e\mu$, τee, $\tau\tau e$ and $\tau\tau\mu$. The channel that has been studied in more detail has been included in the list, $\mu\tau e$, because as mentioned beforehand, we did not perform a complete analysis.

A complete analysis will include another significant difference with respect to our project.

A full background study should be done, including all the possible backgrounds that give a fake signal. Starting with $W^\pm \to l_\alpha^\pm \overset{(-)}{\nu_\alpha} + X$, the samples with higher p_T for the W boson and those containing charm and bottom quark in the jets need to be examined. In addition, other processes could be as significant or even more. For instance, $Z \to l_\alpha^\pm l_\alpha^\mp + X$ and $t\bar{t}$ could be very relevant.

In any case, this thesis shows that even by using one of the channels with better trigger efficiency and only taking into account one source of background, we are far from improving the lower bounds of the exclusion limit. That makes us believe that a prompt HNL from $\mu\tau e$ process could not be detected due to ATLAS detector limitations.

Bibliography

[1] Paul Langacker. "Introduction to the Standard Model and Electroweak Physics". In: *The Dawn of the LHC Era*. WORLD SCIENTIFIC, Aug. 2010. DOI: `10.1142/9789812838360_0001`.

[2] Matthew D Schwartz. "Quantum field theory and the standard model". In: (2014).

[3] Takehiko Asaka and Mikhail Shaposhnikov. "The νMSM, dark matter and baryon asymmetry of the universe". In: *Physics Letters B* 620.1-2 (July 2005). DOI: `10.1016/j.physletb.2005.06.020`.

[4] Takehiko Asaka, Steve Blanchet, and Mikhail Shaposhnikov. "The νMSM, dark matter and neutrino masses". In: *Physics Letters B* 631.4 (Dec. 2005). DOI: `10.1016/j.physletb.2005.09.070`.

[5] C. S. Wu *et al.* "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957). DOI: `10.1103/PhysRev.105.1413`.

[6] O.G. Miranda and J.W.F. Valle. "Neutrino oscillations and the seesaw origin of neutrino mass". In: *Nuclear Physics B* 908 (2016). Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015. ISSN: 0550-3213. DOI: `https://doi.org/10.1016/j.nuclphysb.2016.03.027`.

[7] Georges Aad *et al.* "Search for heavy neutral leptons in decays of W bosons produced in 13 TeV pp collisions using prompt and displaced signatures with the ATLAS detector". In: *Journal of high energy physics* 2019.10 (2019). DOI: `10.1007/jhep10(2019)265`.

[8] Mark Srednicki. "Quantum field theory". In: (2007).

[9]   S. Gninenko, D. Gorbunov, and M. Shaposhnikov. "Search for GeV-Scale Sterile Neutrinos Responsible for Active Neutrino Oscillations and Baryon Asymmetry of the Universe". In: *Advances in High Energy Physics* 2012 (Jan. 2013). DOI: `10.1155/2012/718259`.

[10]   Mark Thomson. "Modern particle physics". In: (2013).

[11]   G. Aad and T. Abajyan. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012). DOI: `10.1016/j.physletb.2012.08.020`.

[12]   S. Chatrchyan and. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012). DOI: `10.1016/j.physletb.2012.08.021`.

[13]   Assaf Shomer. "A pedagogical explanation for the non-renormalizability of gravity". In: *arXiv preprint arXiv:0709.3555* (2007). DOI: `10.48550/ARXIV.0709.3555`.

[14]   David Griffiths. "Introduction to elementary particles". In: (2008).

[15]   Kazuharu Bamba *et al.* "Dark energy cosmology: the equivalent description via different theoretical models and cosmography tests". In: *Astrophysics and Space Science* 342.1 (2012). DOI: `10.1007/s10509-012-1181-8`.

[16]   Raymond Davis Jr, Don S Harmer, and Kenneth C Hoffman. "Search for neutrinos from the sun". In: *Physical Review Letters* 20.21 (1968). DOI: `10.1103/PhysRev Lett.20.1205`.

[17]   Jun John Sakurai and Eugene D Commins. "Modern quantum mechanics, revised edition". In: (1995).

[18]   Neta A. Bahcall. "Dark matter universe". In: *Proceedings of the National Academy of Sciences* (2015). DOI: `10.1073/pnas.1516944112`.

[19]   A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967). DOI: `10.1070/pu1991v 034n05abeh002497`.

[20]   Babak Abi *et al.* "Measurement of the positive muon anomalous magnetic moment to 0.46 ppm". In: *Physical Review Letters* 126.14 (2021), p. 141801. DOI: `10.1103 /physrevlett.126.141801`.

[21]   Roel Aaij *et al.* "Test of lepton universality in beauty-quark decays". In: *arXiv preprint arXiv:2103.11769* (2021). DOI: `10.48550/ARXIV.2103.11769`.

[22]   CDF Collaboration *et al.* "High-precision measurement of the W boson mass with the CDF II detector". In: *Science* 376.6589 (2022). DOI: `10.1126/science.abk17 81`.

[23]   Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the unified model of elementary particles". In: *Progress of Theoretical Physics* 28.5 (1962). DOI: `10.1143/PTP.28.870`.

[24]   Silvia Adrian-Martinez *et al.* "Letter of intent for KM3NeT 2.0". In: *Journal of Physics G: Nuclear and Particle Physics* 43.8 (2016). DOI: `10.1088/0954-3899/43/8/084001`.

[25]   Thomas Schwetz, Mariam Tó rtola, and José W F Valle. "Three-flavour neutrino oscillation update". In: *New Journal of Physics* (Nov. 2008). DOI: `10.1088/1367-2630/10/11/113011`.

[26]   Jean-Loup Tastet. "Searching for Heavy Neutral Leptons at CERN". English. PhD thesis. 2021.

[27]   J.-L. Tastet and I. Timiryasov. "Dirac vs. Majorana HNLs (and their oscillations) at SHiP". In: *Journal of High Energy Physics* 2020.4 (Apr. 2020). DOI: `10.1007/jhep04(2020)005`.

[28]   J. T. Boyd. "LHC Run-2 and Future Prospects". In: (Jan. 2020). DOI: `10.48550/ARXIV.2001.04370`. arXiv: `2001.04370 [hep-ex]`.

[29]   ATLAS Collaboration. "Luminosity determination in *pp* collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: (June 2019). URL: `https://cds.cern.ch/record/2677054`.

[30]   G. Aad *et al.* "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008). DOI: `10.1088/1748-0221/3/08/S08003`.

[31]   A. Airapetian *et al.* "ATLAS: Detector and physics performance technical design report. Volume 1". In: (May 1999). URL: `https://cds.cern.ch/record/391176`.

[32]   A. Airapetian *et al.* "ATLAS: Detector and physics performance technical design report. Volume 2". In: (May 1999).

[33]   Joao Pequenao. "Computer generated image of the whole ATLAS detector". Mar. 2008. URL: `https://cds.cern.ch/record/1095924`.

[34]   ATLAS Collaboration. "Expected performance of the ATLAS experiment-detector, trigger and physics". In: (2008). DOI: `10.48550/ARXIV.0901.0512`.

[35]   Tommaso Colombo, ATLAS Collaboration, *et al.* "Data-flow performance optimisation on unreliable networks: the ATLAS data-acquisition case". In: *Journal of Physics: Conference Series.* Vol. 608. 1. IOP Publishing. 2015. DOI: `10.1088/1742-6596/608/1/012005`.

[36]   ATLAS Collaboration. "Trigger menu in 2018". In: (Oct. 2019). URL: `https://cds.cern.ch/record/2693402`.

[37] Sebastien Rettie. "Muon identification and performance in the ATLAS experiment". In: (June 2018). DOI: 10.22323/1.316.0097.

[38] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 76.5 (May 2016). DOI: 10.1140/epjc/s10052-016-4120-y.

[39] Morad Aaboud *et al.* "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 79.8 (2019). DOI: 10.1140/epjc/s10052-019-7140-6.

[40] ATLAS Collaboration. "Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data". In: (June 2016). Ed. by Etienne Augé, Jacques Dumarchez, and Jean Tran Thanh Van. URL: https://cds.cern.ch/record/2157687.

[41] Morad Aaboud *et al.* "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016". In: *The European Physical Journal C* 79.3 (2019). DOI: 10.1140/epjc/s10052-019-6650-6.

[42] Particle Data Group and. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104.

[43] Camilla Galloni. "Hadronic tau reconstruction and identification performance in ATLAS and CMS". In: *PoS* LHCP2018 (2018). DOI: 10.22323/1.321.0228.

[44] ATLAS Collaboration. "Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment". In: (2019). URL: https://cds.cern.ch/record/2688062.

[45] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-kt jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008). DOI: 10.1088/1126-6708/2008/04/063.

[46] Steven Schramm. "ATLAS Jet Reconstruction, Calibration, and Tagging of Lorentz-boosted Objects". In: *EPJ Web Conf.* 182 (2018). Ed. by Y. Aharonov, L. Bravina, and S. Kabana. DOI: 10.1051/epjconf/201818202113.

[47] Marco Drewes and Jan Hajer. "Heavy Neutrinos in displaced vertex searches at the LHC and HL-LHC". In: *Journal of High Energy Physics* 2020.2 (2020). DOI: 10.1007/jhep02(2020)070.

[48]   DELPHI Collaboration. "Search for neutral heavy leptons produced in Z decays".
       In: *Zeitschrift für Physik C Particles and Fields* 74 (1997). DOI: 10.1007/s00288
       0050370.

[49]   J. Alwall *et al.* "The automated computation of tree-level and next-to-leading order
       differential cross sections, and their matching to parton shower simulations". In:
       Journal of High Energy Physics 2014.7 (July 2014). DOI: 10.1007/jhep07(2014
)079.

[50]   Daniel Alva, Tao Han, and Richard Ruiz. "Heavy Majorana neutrinos from $W\gamma$
       fusion at hadron colliders". In: *Journal of High Energy Physics* 2015.2 (2015). DOI:
       10.1007/jhep02(2015)072.

[51]   Céline Degrande *et al.* "Fully automated precision predictions for heavy neutrino
       production mechanisms at hadron colliders". In: *Physical Review D* 94.5 (2016).
       DOI: 10.1103/physrevd.94.053002.

[52]   Christian Bierlich *et al.* "A comprehensive guide to the physics and usage of
       PYTHIA 8.3". English. In: *SciPost Physics* (2022). ISSN: 2542-4653. DOI: 10.4855
       0/ARXIV.2203.11601.

[53]   Philip Ilten. "Tau decays in PYTHIA 8". In: *Nuclear Physics B-Proceedings
       Supplements* 253 (2014). DOI: 10.1016/j.nuclphysbps.2014.09.019.

[54]   Sea Agostinelli *et al.* "GEANT4—a simulation toolkit". In: *Nuclear instruments
       and methods in physics research section A: Accelerators, Spectrometers, Detectors
       and Associated Equipment* 506.3 (2003). DOI: 10.1016/S0168-9002(03)01368-8.

[55]   J Allison *et al.* "Recent developments in Geant4". In: *Nuclear Instruments and
       Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors
       and Associated Equipment* 835 (2016). DOI: 10.1016/j.nima.2016.06.125.

[56]   J. and Allison. "Geant4 developments and applications". In: *IEEE Transactions
       on Nuclear Science* 53.1 (2006). DOI: 10.1109/TNS.2006.869826.

[57]   *"Athena Introduction by ATLAS Software Documentation"*. Visit the following
       URL: https://atlassoftwaredocs.web.cern.ch/athena/athena-intro/.
       Accessed: 20/05/2022.

[58]   *"TruthDAOD"*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/Tr
       uthDAOD. Accessed: 20/05/2022.

[59]   *"Overview of MC generator software in ATLAS"*. https://twiki.cern.ch/twik
       i/bin/view/AtlasProtected/PmgMcSoftware. Accessed: 20/05/2022.

[60]   *"SUSY xAOD Derivations in Release 19"*. https://twiki.cern.ch/twiki/bin
       /view/AtlasProtected/SUSYxAODDerivationsr19. Accessed: 20/05/2022.

[61] *"DerivationProductionTeam"*. Visit the following URL: `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/DerivationProductionTeam`. Accessed: 20/05/2022.

[62] *"Weak Boson Processes"*. `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PmgWeakBosonProcesses`. Accessed: 20/05/2022.

[63] *"ATLAS Metadata Interface"*. `https://ami.in2p3.fr/`. Accessed: 20/05/2022.

[64] D. Berdine, N. Kauer, and D. Rainwater. "Breakdown of the Narrow Width Approximation for New Physics". In: *Phys. Rev. Lett.* 99 (11 Sept. 2007). DOI: `10.1103/PhysRevLett.99.111601`.

[65] Johan Alwall *et al.* "MadGraph 5: going beyond". In: *Journal of High Energy Physics* 2011.6 (June 2011). DOI: `10.1007/jhep06(2011)128`.

[66] J.-L. Tastet, O. Ruchayskiy, and I. Timiryasov. "Reinterpreting the ATLAS bounds on heavy neutral leptons in a realistic neutrino oscillation model". In: *Journal of High Energy Physics* 2021.12 (Dec. 2021). DOI: `10.1007/jhep12(2021)182`.

[67] Dmitry Gorbunov and Mikhail Shaposhnikov. "How to find neutral leptons of the νMSM?" In: *Journal of High Energy Physics* 2007.10 (2007). DOI: `10.1088/1126-6708/2007/10/015`.

[68] Ivan Esteban *et al.* "The fate of hints: updated global analysis of three-flavor neutrino oscillations". In: *Journal of High Energy Physics* 2020.9 (Sept. 2020). DOI: `10.1007/jhep09(2020)178`.

[69] Ivan Esteban and M.C. Gonzalez-Garcia and Michele Maltoni and Thomas Schwetz and Albert Zhou. *NuFIT 5.0 (2020)*. `http://www.nu-fit.org`. Online; accessed 29 January 2014.

[70] ATLAS Collaboration. "Reconstruction and identification of electrons in ATLAS". In: *CERN Yellow book* 4 (2008). URL: `http://www.physics.smu.edu/~scalise/SMUpreprints/SMU-HEP-08-21.pdf`.

[71] Marc Paterno. "Calculating efficiencies and their uncertainties". In: (Dec. 2004). DOI: `10.2172/15017262`.

[72] Guolin Ke *et al.* "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". In: 2017. URL: `https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf`.

[73] Trevor Hastie *et al.* "The elements of statistical learning: data mining, inference, and prediction". In: 2 (2009).

[74] Jerome H Friedman. "Greedy function approximation: a gradient boosting machine". In: *Annals of statistics* (2001). DOI: `10.1214/aos/1013203451`.

[75] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: Aug. 2016. DOI: 10.1145/2939672.2939785.

[76] Jia Wu *et al.* "Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization". In: *Journal of Electronic Science and Technology* 17.1 (2019). ISSN: 1674-862X. DOI: https://doi.org/10.11989/JEST.1674-862X.80 904120.

Appendices

Appendix A Data samples

Information of the data samples used for the analysis of signal $\mu\tau e$ and background.

Dataset number	Dataset short name	Period Run2	Total events	Events DAOD
364156	Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto	a, d, e	97624850	6209924
364170	Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto	a	24997000	11747676
510416	MGPy8EG_A14N23LO_HNL20_ctau01_mutae_LNC	a, d, e	162000	26879
510446	MGPy8EG_A14N23LO_HNL50_ctau01_mutae_LNC	a, d, e	180000	37895
512067	MGPy8EG_A14N23LO_HNL20_ctau01_mutae_LNV	a, d, e	1770000	422822
512070	MGPy8EG_A14N23LO_HNL50_ctau01_mutae_LNV	a, d	1100000	272091

Table 11.6: Table of the samples used for background and signal, containing the dataset number, the dataset short name, the period of Run-2 used, the total events of generation and the number of events selected by the DAOD_SUSY3.

Dataset number	σ by AMI	Filter ε	Generator	ATLAS release
364156	19.2 nb	0.82465	Sherpa(v2.2.1)	AthDerivation_21.2.79.0
364170	19.2 nb	0.82463	Sherpa(v2.2.1)	AthDerivation_21.2.79.0
510416	582 μb	1	MadGraph(v.2.9.3.atlas)+Pythia8	AthDerivation_21.2.130.0 (v.245p3.lhcb7)+EvtGen(v.1.7.0)
510446	36.9 mb	1	MadGraph(v.2.9.3.atlas)+Pythia8	AthDerivation_21.2.130.0 (v.245p3.lhcb7)+EvtGen(v.1.7.0)
512067	581 μb	1	MadGraph(v.2.9.5.atlas2)+Pythia8	AthDerivation_21.2.130.0 (v.245p3.lhcb7)+EvtGen(v.1.7.0)
512070	37.0 mb	1	MadGraph(v.2.9.5.atlas2)+Pythia8	AthDerivation_21.2.130.0 (v.245p3.lhcb7)+EvtGen(v.1.7.0)

Table 11.7: Table of the samples used for background and signal, containing the dataset number, the cross section given by AMI, the generator filter efficiency, the generator used and the ATLAS release.

Appendix B   Correlation variables for BDT in LNC

The histograms and scatter plots of the different variables used in the boosted decision tree (chapter 10) are presented for the case where we have performed an LNC preselection, Fig. B.1 and B.2. These plots are comparable with the ones for LNV preselection in Fig. 10.2 and 10.3, having these last more statistics.
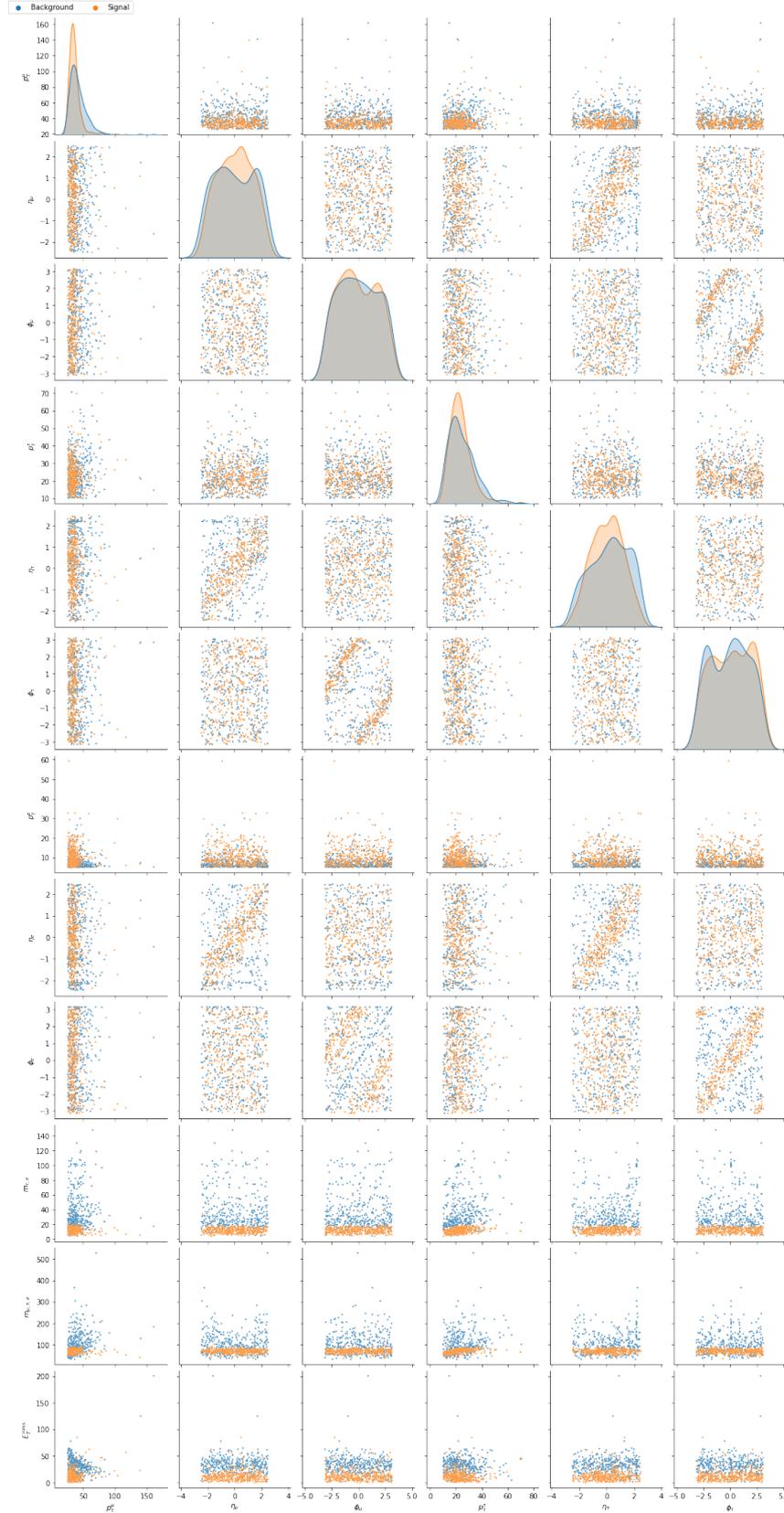
Figure B.1: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $\boldsymbol{m_N = 20\ \mathbf{GeV}\ \mathbf{LNC}}$ preselection. The figure is split in two for better visibility, continuing in next page.

Figure B.1: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $\boldsymbol{m_N = 20\ \mathbf{GeV}}$ **LNC** preselection. (continuation)
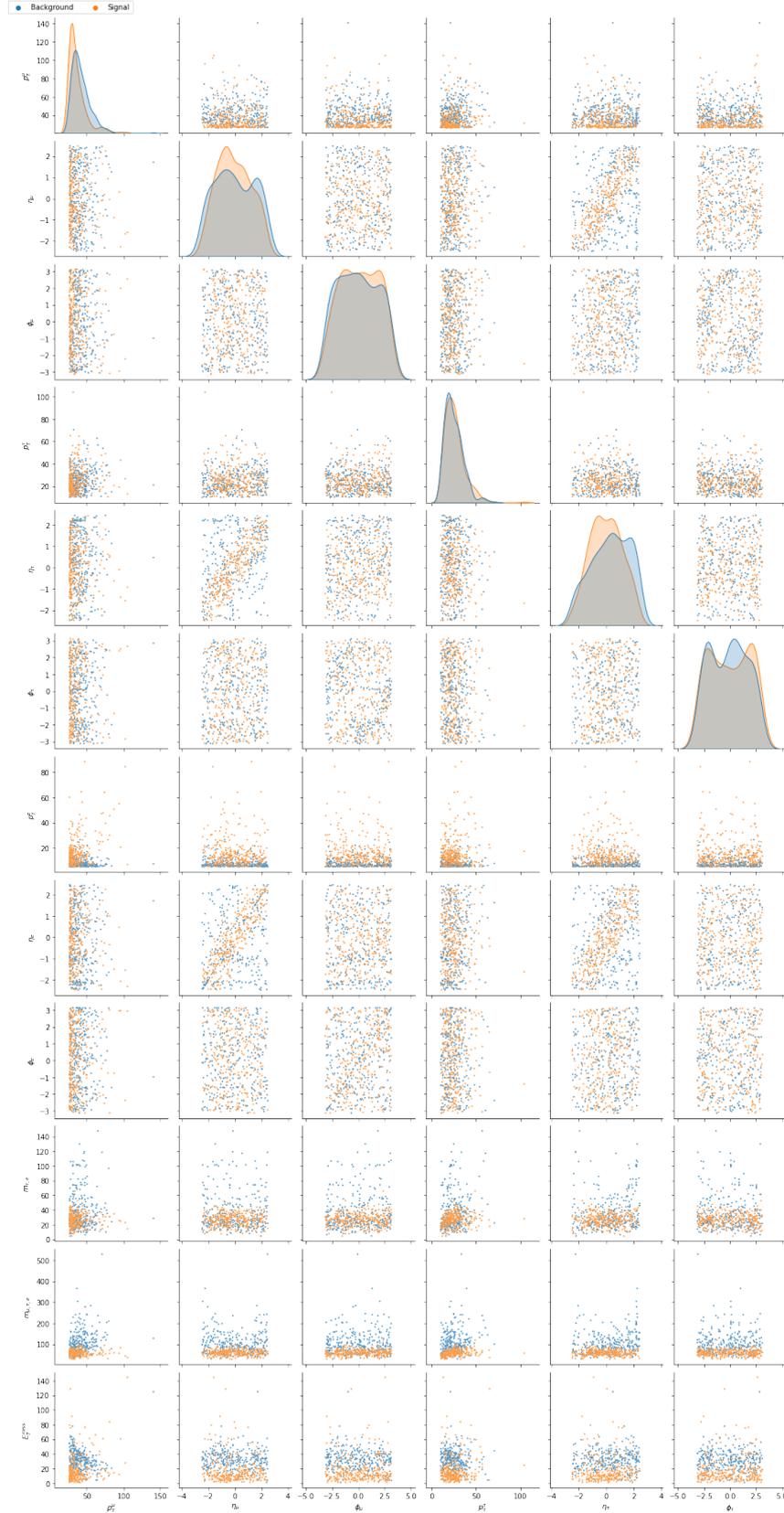
Figure B.2: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $\boldsymbol{m_N = 50}$ **GeV LNC** preselection. The figure is split in two for better visibility, continuing in next page.

Figure B.2: Diagonal 1-D histograms and off-diagonal 2-D scattering plots of the following variables: p_T, η and ϕ of the three leptons (μ, τ and e in order); invariant mass of μ and e; invariant mass of the three leptons; and E_T^{miss}. This is the order from top to bottom on y-axis and from left to right on x-axis. In blue the background and in orange the signal with $m_N = 50$ **GeV LNC** preselection. (continuation)