Department of Physics and Astronomy
University of Heidelberg

Master Thesis in Physics
submitted by

Maurice Morgenthaler

born in Kehl (Germany)

2021

Search for heavy neutral leptons in inclusive B decays at the LHCb experiment

This Master Thesis has been carried out by Maurice Morgenthaler at the
Physikalisches Institut in Heidelberg
under the supervision of
Prof. Dr. Stephanie Hansmann-Menzemer

Abstract

This master thesis is dedicated to an inclusive search for heavy neutral leptons (HNL, N) in B meson decays. The analysis is performed at the LHCb experiment and is currently in a stage of blinded data. The used decay channels $B \to X_{Prim}\mu N(\to \mu\pi)$ provide a good signal signature in form of a displaced HNL decay vertex. The not reconstructed hadron is denoted as X_{Prim} while B includes all B meson flavours B^+, B^0, B_s^0 and B_c^+. A scan of the HNL lifetime of 0.1 ps to 10 ns HNL lifetime and of the HNL mass of 1.6 GeV to 6 GeV is performed to set limits on the HNL mixing with muonic neutrinos, U_μ^2. Using a large number of simulated decay channels, the analysis strategy is defined and selection efficiencies are calculated. The search is performed in 15 different categories based on the HNL decay vertex displacement, charge of the two muons and if the HNL production channel is leptonic. A cut-based preselection is applied and a multivariate analysis (MVA) is used to separate signal from background. In the area of 2 GeV HNL mass a preliminary sensitivity of 2×10^{-5} is achieved which is competitive to the current world-best upper limit.

Zusammenfassung

Diese Masterarbeit ist der inklusiven Suche von schweren neutralen Leptonen (HNL, N) in B Meson Zerfällen gewidmet. Die Analyse wird am LHCb Experiment durchgeführt und befindet sich momentan in einer Phase verblindeter Daten. Die genutzten Zerfallskanäle $B \to X_{Prim}\mu N(\to \mu\pi)$ besitzen eine gute Signalsignatur in Form einer verschobenen HNL Zerfallsvertex. Das nicht rekonstruierte Hadron wird mit X_{Prim} bezeichnet während B alle B Meson Sorten, B^+, B^0, B_s^0 und B_c^+, beinhaltet. Ein Scan über die HNL Lebenszeit von 0.1 ps bis 10 ns und HNL Masse von 1.6 GeV bis 6 GeV wird durchgeführt um Grenzen für das Mischen von HNL mit Myon-Neutrinos zu setzen. Unter Benutzung einer Vielzahl von simulierten Zerfallskanälen wird die Strategie der Analyse definiert und Effizienzen der Selektion berechnet. Die Suche wird in 15 Kategorien durchgeführt, welche auf der Verschiebung der HNL Zerfallsvertex, Ladung der zwei Myonen und darauf, ob der HNL Produktionskanal leptonisch ist, basieren. Eine schnittbasierte Selektion wird angewandt und eine multivariate Analysemethode (MVA) benutzt, um Signal von Hintergrund zu unterscheiden. Im HNL Massenbereich von 2 GeV wird eine vorläufige Sensitivität von 2×10^{-5} erreicht welche konkurrenzfähig mit der derzeitigen weltbesten Obergrenze ist.

Contents

1 Introduction

There is a multitude of fundamental physics phenomena that are currently not explained. For example why does the universe exist in its current form? More precisely, why is there an asymmetry between matter and anti-matter of that scale? Andrei Sakharov proposed in 1967 three necessary conditions to be able to explain this phenomena: The violation of the Baryon number, a large enough violation of CP symmetry as well as out of thermal equilibrium interactions in the universe [1]. Another open question is the nature of *Dark Matter*. Already in the 1930s astronomers saw that their estimation of galaxy masses based on luminous matter was off. The term of dark matter was coined by Peter Zwicky and colleagues seeing that the Coma cluster showed a large velocity dispersion [2]. With the years evidence for non-luminous matter like flat rotation curves observed by Vera Rubin in 1970 [3], mass lensing effects and more arised. Nowadays several searches for dark matter candidates are ongoing. A further problem is the non-zero mass of neutrinos. In the 1970s the *Homestake* experiment measured less electron neutrinos coming from the sun as expected [4]. In 2001, the *Sudbury Neutrino Observatory* (SNO) could prove that neutrinos oscillate between different generations which also showed that neutrinos have a mass [5]. Measuring the squared difference of neutrino mass eigenvalues Δm_{ij}^2 shows extremely small masses for neutrinos. For all those open questions so-called *heavy neutral leptons* (HNL) could provide a solution. They are right-handed neutrinos which are easy to incorporate into the Standard Model of Particle Physics. Due to their right-handedness they don't interact directly with other particles but mix with Standard Model neutrinos. For this reason HNLs are also called *sterile neutrinos*. This thesis is dedicated to the search of HNLs in B meson decays at the LHCb experiment. The LHCb experiment is one of seven experiments located at the *Large Hadron Collider* (LHC). The LHC is up to now the largest and most powerful particle accelerator and is located at the *European Organization for Nuclear Research* (CERN). Crossing over the Swiss-French border the LHC collides protons at a center-of-mass energy up to 14 TeV.

The structure of this thesis is as follows: Section 2 gives an overview over the Standard Model of Particle Physics. The elementary particles as well as their interactions are discussed. The following section 3 is about physics

beyond the Standard Model. After a short introduction about the discovery of neutrino masses it is explained how neutrino masses and heavy neutral leptons can be incorporated into the Standard Model. This includes possible production and decay channels of the HNL. An overview about the LHCb experiment is given in section 4. The layout of the detector as well as the data flow and the trigger system is described. Based on the knowledge about the HNL and the strengths of the LHCb detector an analysis strategy is formulated in section 5. Section 6 details the used data samples. It describes how the samples are simulated and which requirements (e.g stripping lines) are used to select the data used in this analysis. Several different Monte Carlo samples have to be combined into one data set, which is described section 7. This includes corrections for branching ratios and the computation of weights that scale the contributions of different decay channels to properly describe the composition. In this thesis limits are set on the mixing angle of the muon neutrino with the HNL, U_μ^2. The relevant studies are presented in section 8 which includes the evaluation of the reconstruction efficiency. To be able to scan over a large range of HNL lifetimes and masses the data sets have to be reweighted and interpolated which is also shown. With those quantities limits on the muon neutrino mixing angle U_μ^2 are set. These limits depend on the theoretical model used to predict the HNL production and decay rates. Furthermore, a model independent limit on the HNL branching ratios is estimated. At last, a discussion about systematic uncertainties is given. Section 9 is an outlook on limits set with the preliminary MVA and on the still open tasks. Finally section 10 concludes the presented work.

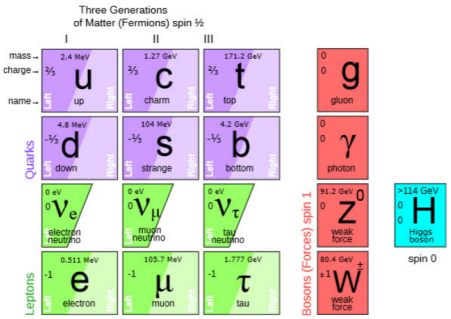


Figure 2.1: Particles in the basic Standard Model. The mass of the neutrinos is shown as zero as it can't be explained without an extension [6].

2 The Standard Model of Particle Physics

The Standard Model of Particle Physics describes all elementary particles and their interactions. In the following an overview about their properties is given.

2.1 Particle content of the Standard Model

There are four different types of particles in the Standard Model. Those are *quarks*, *leptons*, *gauge bosons* as well as the *Higgs boson*. Said four types can be seen in figure 2.1 including their mass, electric charge, handedness and spin. The mass is given in units of energy. This is coming from the convention of *natural units* setting the speed of light and Planck's constant to one: $c = 1,\ \hbar = 1$. This convention is used in the complete thesis. All particles have a corresponding antiparticle. An antiparticle has the same mass and spin as its counterpart but all other properties, so-called charges, are inverted. A particle with only neutral charges such as a photon is its own antiparticle.

2.1.1 Quarks

Quarks have a spin of $\frac{1}{2}$ and as such are fermions following Fermi-Dirac statistics. There exist six quark flavours that are ordered into pairs separated into three generations with increasing mass. They are called up (u), down (d), charm (c), strange (s), top (t), bottom (b). Up, charm and top quarks carry an electric charge of 2/3 while their partners carry electric charge -1/3.

Quarks usually don't exist in an isolated state. Instead they combine into composite particles called *hadrons*. Particles that constitute out of a quark - antiquark pair are called *mesons* while those with three or more quarks are *baryons*. Hadrons posses an integer electric charge. Examples for mesons are pions which are consisting of u and d quarks. Neutrons and protons on the other hand are baryons with a quark content of udd and uud, respectively.

2.1.2 Leptons

Leptons just like quarks are fermions therefore also have spin $\frac{1}{2}$. There are also six leptons, grouped in three generations of pairs. In comparison to quarks they have different electric charges with electrons, muons and tauons having -1 in units of elementary charges. Their respective partners the electron neutrino, muon neutrino and tau neutrino are electrically neutral. Furthermore, neutrinos are special. The Standard Model has no way to assign them a mass. This is a insufficiency of the model as neutrinos do have masses as shown by *Neutrino-Oscillation* first found by the *Super-Kamiokande* experiment [7]. Therefore the model is extended to include neutrinos with a mass. For that the *Pontecorvo–Maki–Nakagawa–Sakata* (PMNS) matrix is used as shown in equation 2.1 It illustrates how the mass and flavour eigenstates of neutrinos mix with each other.

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

Furthermore, up to this date only *left-handed* neutrinos are found which motivates the search of right-handed neutrinos discussed later.

2.1.3 Bosons

The Standard Model contains five bosons that as such have an integer spin and follow *Bose-Einstein-statistics*. The Higgs boson is the only scalar boson and its mediated Higgs-field introduces the masses of the particles. The other four are corresponding to the three forces described by the Standard Model and are their mediators. They are referred to as *gauge bosons*. The Z and W boson of the weak interaction are the only massive ones and the W is

the only electrically charged gauge boson. The photon is the exchange boson of the electromagnetic interaction and the gluon of the strong interaction.

2.2 Interactions of the Standard Model

The three interactions of the Standard Model are described by group theory in a $SU(3) \times SU(2)_L \times U(1)_Y$ symmetry. Each group describes one of the particle interactions. The generators of the groups correspond to the gauge bosons. It follows an overview about the properties of the fundamental forces and its carriers.

2.2.1 Strong interaction

The *strong interaction* is descriped by *Quantum Chrono Dynamics* (QCD) and corresponds to an SU(3) group. Its mediator are the massless gluons. The strong field couples to the property called *color* which acts as a charge like characteristic. There are three different colors called red, green and blue. Only gluons and quarks have a color and are thus the only elementary particles which interact strongly. The strong force is the interaction keeping together hadrons which consist out of quarks. As the corresponding group is SU(3) there are eight generators (gluons) representing color combinations. Gluons are the only gauge bosons that couple with themselves.

2.2.2 Electromagnetic force

This force is described by *quantum electrodynamics* and corresponds to an $U(1)_Y$ group. Its mediated by photons and couples to electric charge. An example for the electromagnetic force would be the interaction between electrons and protons in an atom. As photons don't have an electric charge they can't couple to themselves. The range of its force is infinite contrary to the other two interactions.

2.2.3 Weak force

The *weak force* is usually combined with the electromagnetic force in the *electro-weak theory*. By the Higgs-mechanism the $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken. This results in the massive mediator Z and W boson of the weak force and the photon for the electromagnetic interaction. As its mediators are massive the range is low. On the other hand the

Particle	e^-	e^+	ν_e	$\overline{\nu_e}$	μ^-	μ^+	ν_μ	$\overline{\nu_\mu}$	τ^-	τ^+	ν_τ	$\overline{\nu_\tau}$
L_e	1	-1	1	-1	0	0	0	0	0	0	0	0
L_μ	0	0	0	0	1	-1	1	-1	0	0	0	0
L_τ	0	0	0	0	0	0	0	0	1	-1	1	-1

Table 2.1: Lepton numbers split by generation. L_e, L_μ, L_τ correspond to the first, second and third particle generation respectively.

charged W boson is the only one able to change particle flavour. The charge of the weak force is the *isospin*. Every fermion of the Standard Model has a non-zero isospin meaning that they all interact weakly. Another important fact is that under SU(2) transformations left-handed particles are doublets while right-handed ones are singlets. This means that the weak interaction of the SM can transform a left-handed electron into a left-handed electron neutrino but couldn't do the right-handed equivalent. As such it at first doesn't incorporate right-handed neutrinos.

2.2.4 Conserved quantities of the SM

The Standard Model is gauge and reference-frame invariant. This is needed to have finite predictions for physical measurements independent of reference frame. As described by the *Noether theorem* each invariance is linked to a conserved quantity. In case of the three gauge invariant groups $SU(3) \times SU(2)_L \times U(1)_Y$ the conserved quantities are the charge of the corresponding force. For the strong force the color is conserved, for the electromagnetic one it's the electric charge and for the weak interaction it's the isospin. Apart from these true conserved quantities there are also such that are not grounded in known symmetries. One of them is the lepton number, L. As shown in table 2.1 every lepton has this quantity. The lepton number must be conserved for every particle generation independently. The only currently known source of lepton flavour violation is neutrino oscillation. In neutrino oscillation a neutrino can change its generation oscillating for example from a muon neutrino to an electron neutrino. As such finding processes that violate lepton number is a good sign for new physics.

3 Physics beyond the Standard Model

In the last section it is explained that the Standard Model doesn't involve right-handed neutrinos. But there are extensions of the Standard Model which include them. Depending on their mass right-handed neutrinos are often called *sterile neutrinos* or *heavy neutral leptons* (HNL). If they exist, they could solve multiple physics puzzles such as the existence of Dark Matter, the matter-asymmetry of the universe and the smallness of neutrino masses. In this section we discuss small neutrino masses and how HNLs could explain them. Furthermore, there will be a subsection about the generation of HNLs.

3.1 A short history of neutrino masses

Neutrinos have non-zero mass. This is shown by several experiments that investigated neutrino oscillation. First hints on neutrino oscillation were found by the Homestake experiment that measured the amount of solar ν_e investigating the amount of argon that was produced by interaction of the neutrinos with a medium containing chlorine:

$$\nu_e \; + \; {}^{37}Cl \; \rightarrow \; {}^{37}Ar^{+} + \; e^{-} \tag{3.1}$$

The observed number was about one third lower than the predicted amount which was a first hint towards the existence of neutrino oscillation [4]. Multiple other experiments such as the Super-Kamiokande experiment produced since then further results strengthening the finding [7]. In 2001 the *Sudbury Neutrino Observatory* was finally able to directly measure neutrino oscillation. By employing heavy water (deuterium) it was able to measure separately the rate of all neutrino species. Therefore it directly measured the increase in the abundance of one generation and the decrease of others [5]. This evidence for neutrino oscillation is a direct proof of neutrinos not being massless. Neutrino oscillations can only measure the squared difference of neutrino mass eigenvalues Δm^2_{ij} giving rise to a direct neutrino mass search. In 2021 the *KATRIN* experiment was able to measure a combined upper limit of $m_\nu < 0.8\,\mathrm{eV}$ at $90\,\%$ confidence limit [8].

3.2 Heavy neutral leptons

The *neutrino Minimal Standard Model* (νMSM) aims to solve multiple physical puzzles by adding three right-handed neutrinos to the Standard Model. Neutrinos in the SM are left-handed SU(2) doublets ν_L while the added sterile neutrinos are right-handed SU(2) singlets N_R:

$$\begin{pmatrix} \nu_l \\ l^- \end{pmatrix}_L = \nu_L \overset{CPT}{\longleftrightarrow} \nu_R = \begin{pmatrix} l^+ \\ \nu_l^c \end{pmatrix}_R \qquad N_R \overset{CPT}{\longleftrightarrow} N_L^c \tag{3.2}$$

A left-handed as well as right-handed component is needed for a *Dirac mass* term m_D in the Dirac Lagrangian $\mathcal{L}_{Dirac}$:

$$-\mathcal{L}_{Dirac} = m_D \overline{\nu}\nu \tag{3.3}$$

$$= m_D(\overline{\nu}_L N_R + \overline{N}_R \nu_L) \tag{3.4}$$

Hereby $\nu = \nu_L + N_R$ and the overline is the conjugated particle field. Neutrinos are electrically neutral which allows a Majorana mass term to be introduced. An electrically charged fermion can't acquire such a mass term as it would violate the conservation of electric charge.

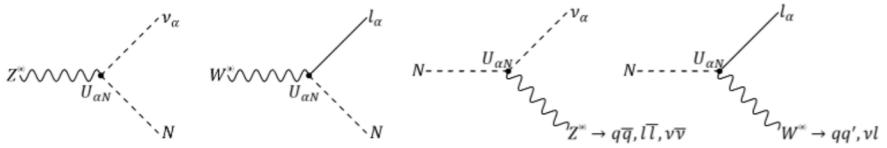
$$-\mathcal{L}_{Majorana} = \frac{1}{2} m_N \left(\overline{N}_L^c N_R \ + \overline{N}_R N_L^c \right) \tag{3.5}$$

This can be done for all three generations. To get the mass eigenvalues of the neutrino one has to get the eigenvalues λ_+, λ_- of the mass matrix M_ν:

$$M_\nu = \begin{pmatrix} 0 & m_D \\ m_D & m_N \end{pmatrix} \tag{3.6}$$

$$\lambda_+ = \frac{m_N + \sqrt{m_N^2 + 4m_D^2}}{2} \qquad \lambda_- = \frac{m_N - \sqrt{m_N^2 + 4m_D^2}}{2} \tag{3.7}$$

Using the Ansatz that $m_D << m_N$ and that $\lambda_+ \lambda_- = -M^2$ it becomes apparent that $\lambda_+ \approx m_N$ and therefore $m_\nu = \lambda_- \approx -\frac{M^2}{m_N}$. This is the *seesaw-mechanism* [9]. By introducing a heavy neutral lepton the SM, the mass of the left-handed neutrinos has to be small which explains our experimental findings.



(a) General production of an HNL *(b) General decay of an HNL*

Figure 3.1: Most general depiction of production and decay of a heavy neutral lepton. The Z and W bosons are virtual for certain channels. Dashed lines indicate for the LHCb detector invisible particles.

3.3 Production and decay of heavy neutral leptons

As mentioned beforehand HNLs are also called sterile neutrinos. As singlet under the three interactions it can't directly interact with any of the other neutrinos making HNLs *sterile*. The only way to interact is by mixing with SM neutrinos. This means that theoretically in every channel that involves neutrinos there can be HNLs. However, the mixing angle, U_α^2, could be almost as low as m_ν/m_N, which is very small. Assuming that $m_N = 2\,\text{GeV}$, as this is one of the masses used in this analysis, and assuming $m_\nu = 0.02\,\text{eV}$ this would mean that the mixing angle could be as low as $U_\mu^2 \approx 10^{-11}$.
Figure 3.1 shows how a basic Z and W boson vertex produces a neutrino that mixes with a heavy neutral lepton N via the mixing angle, U_α. For the decays involving a W boson a detectable lepton is produced. For the decay, the HNL has to return to a standard neutrino again which introduces a second factor of U_α. Therefore the production rate is proportional to U_α^2 and the HNL lifetime τ to $U_\alpha^{-2} \cdot m_N^{-5}$. The neutrino can then decay weakly to a quark pair, lepton pair or neutrino pair if decaying via a Z boson exchange. Alternatively it can decay to two quarks or a neutrino and a lepton via a W boson. Due to the possibility of a Dirac and a Majorana mass term it is possible that the two produced leptons have either *same-sign* (SS) electric charge or *opposite-sign* (OS) electric charge. The SS case would violate lepton number conservation by two units, $\Delta L = 2$. For it a neutrino has to mix with a heavy neutral neutrino. As the Majorana HNL is its own antiparticle it can then mix with the antineutrino version of the initial neutrino.

4 The LHCb experiment

The *Large Hadron Collider* (LHC) is an accelerator ring at the *European Organization for Nuclear Research* (CERN derived from *Conseil européen pour la recherche nucléaire*). The CERN main site is located in Meyrin in Switzerland while most of the accelerator ring is located on French ground. It is in terms of energy and interaction rate the most powerful proton-proton collider in the world. In its first run from 2011 to 2012 it reached an energy of 7 TeV (2011) and 8 TeV (2012) while the second run had an energy of 13 TeV (2015 - 2018). At the time of this thesis the LHC is being upgraded and prepared for Run 3 which starts in spring 2022 at a significantly higher luminosity. The ring provides collisions to several experiments with the LHCb experiment being one of the most prominent ones. As the "b" in the name suggests the experiment is dedicated to detecting particles which contain a bottom quark. The main purpose of the detector is the investigation of CP violation. CP violation is an effect occurring due to the interference of two decays involving different phases. It appears in loop mediated processes such as the box diagram of B_s^0 oscillation. Such loops are sensitive to new physics and thus are used to search for rare particles.

Apart from this main task the LHCb experiment is well suited for a multitude of analysis based on b containing particles. Heavy neutral leptons which are the subject of this master thesis are an example for this. Their existence would explain the asymmetry of matter and antimatter in our universe which CP violation cannot.

In the following a general overview of the LHCb experiment is given. A cross section of the detector can be found in figure 4.1. The different parts will be discussed going from the collision point in the Vertex Locator in positive z direction till the muon stations at the end of the detector. The x axis and y axis is perpendicular to the z axis with the x direction pointing to the center of the ring and y direction pointing upwards. This forms a right-handed coordinate system. Positions on the z axis are also often referred to as downstream and upstream. Something downstream of a position refers to a higher z value, while upstream means a lower z value. Furthermore, information about the trigger system of the LHCb experiment is given at the end of this section.

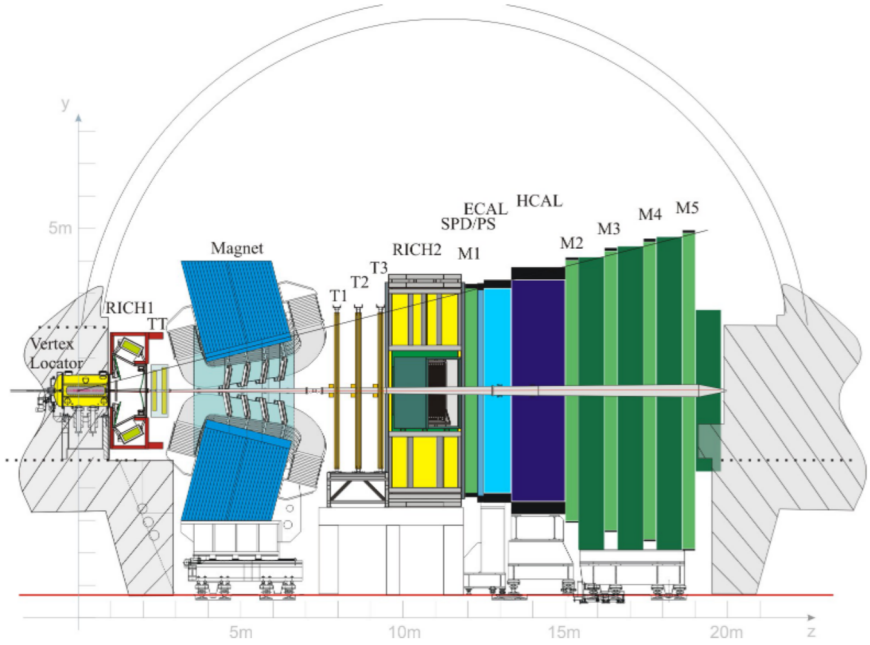


Figure 4.1: View on the LHCb detector from the x direction [10]

4.1 Detector layout

As mentioned the LHCb detector is specialized on detecting decays involving B mesons. Pairs of B mesons are boosted thus fly into the same direction with just a few degrees difference. Due to this the LHCb detector is built as a forward detector measuring just into one direction of the beam pipe. Further properties important for B meson detection are for example high efficiencies in track reconstruction and particle identification.

Vertex Locator (VELO): The *Vertex Locator* is the first part of the LHCb tracking system. Placed at the proton-proton interaction point its purpose is to reconstruct primary and secondary vertices from measured tracks [11]. It is compromised out of 42 silicon strip modules each measuring r and Φ coordinates in a cylindrical coordinate system [12]. Reconstructing displaced vertices has great importance for the high-level trigger as well as many analyses which cut on the displacement. Thus, the VELO stripes are moved as close as possible to the particle beam to improve efficiency. The modules and the beam are only 7 mm apart during data taking. The system has to be retracted during beam injection to not damage the modules [11].

Ring Imaging Cherenkov detectors (RICH): The LHCb detector incorporates two RICH detectors for the purpose of identifying charged hadrons like pions, kaons and protons. RICH 1 covers the lower momentum region from 2 to 60 GeV while RICH2 placed downstream measures between 15 GeV and 100 GeV [12]. This is done by using the effect of Cherenkov radiation which occurs if a charged particle is moving through a medium faster than the speed of light in said medium. The charged particle excites the medium causing the emission of photons when they deexcite back to the ground state. As the particle propagates faster than the photons a measurable ring of light is produced. The emission angle θ is directly proportional to particle speed v and the known speed of light in the used medium $c_{med} = c/n$ with the vacuum speed of light c and the mediums refractive index n:

$$cos(\theta) = \frac{c_{med}}{v} = \frac{c}{nv} \tag{4.1}$$

Combining this information with information from the trajectory as explained in the following enables the calculation of mass and charge of the particle and as such its identity.

Tracking modules: The tracking stations are the second part of the LHCb tracking system. There are four planar modules with the *Trigger Tracker* (TT) placed upstream of the magnet and the three other tracking stations (T1 - T3) containing four layers each directly downstream of it. The stations precisely measure particle hits and together with the VELO enable track reconstruction. With the magnet placed upstream of T1 - T3 it is furthermore possible to calculate particle momenta. The modules of TT and the *Inner Tracker* (IT) of the downstream stations are silicon microstrips. For the Outer Tracker, which measures areas of lower track density further away from the beam, gas filled straw tubes are used instead [12].

Magnet: The magnet used in the LHCb detector is a warm dipole magnet built from two coils. It provides a homogeneous field of 4 Tm needed for the downstream tracking stations mentioned before [13]. Charged particles moving through the magnetic field will experience Lorentz force. This causes a deflection of trajectory measured by T1 - T3 which is reconstructed to a radius. The Lorentz force $\overrightarrow{F}_{Lorentz}$ depends on the particle velocity $\overrightarrow{v}$,

particle charge q and the known magnetic field $\overrightarrow{B}$:

$$\overrightarrow{F}_{Lorentz} = q \cdot \left(\overrightarrow{v} \times \overrightarrow{B} \right) \tag{4.2}$$

Combining this with the equation for centrifugal force, $\overrightarrow{F}_{Centri}$, allows to calculate the particle momentum p. Thereby denotes r the reconstructed particle trajectory radius.

$$\overrightarrow{F}_{Centri} = \frac{p\,v}{r} \tag{4.3}$$

$$p = q\,r\,B \tag{4.4}$$

Calorimeter system: The calorimeter system of the LHCb detector consists of four sub-modules placed behind each other perpendicular to the beam pipe. The first module is the *Scintillator Pad Detector* (SPD). Its scintillators measure the energy of charged particles coming into the system. Following a lead foil is the *PreShower* (PS). When the foil causes an electromagnetic shower energy deposited in the PS will indicate it. The *Electromagnetic CALorimeter* (ECAL) consists out of layers of lead and scintillators. The lead once again causes electromagnetic showers. The produced electrons are measured in a grid of pixels to calculate the particle energy. Hadrons have a higher interaction length than photons and electrons and as such reach the *Hadronic CALorimeter* (HCAL). This module consists out of thicker layers of scintillators and iron and measures the energy of hadronic showers [14]. For both the ECAL and the HCAL the deposited energy is proportional to the amount of produced photons captured by photomultiplier tubes. The calorimeters are essential to identify non-charged particles as those can't be identified by the tracking system.

Muon system (M1 - M5): The *Muon system* is the last part of the LHCb detector. It measures muons which traverse through the calorimeters. The system is built from five modules, M1 - M5, where the first is placed upstream of the calorimeter system and the rest downstream. The Muon system is used to measure the momentum and transverse momentum of muons. To filter out other particles it is placed as far back in the detector as possible and has lead plates between the modules. This ensures that only the

L0 trigger	E_T or p_T threshold			SPD threshold
	2015	2016	2017	
Hadron	$> 3.60\,\text{GeV}$	$> 3.70\,\text{GeV}$	$> 3.46\,\text{GeV}$	< 450 hits
Photon	$> 2.70\,\text{GeV}$	$> 2.78\,\text{GeV}$	$> 2.47\,\text{GeV}$	< 450 hits
Electron	$> 2.70\,\text{GeV}$	$> 2.40\,\text{GeV}$	$> 2.11\,\text{GeV}$	< 450 hits
Muon	$> 2.80\,\text{GeV}$	$> 1.80\,\text{GeV}$	$> 1.35\,\text{GeV}$	< 450 hits
high p_T Muon	$> 6.00\,\text{GeV}$	$> 6.00\,\text{GeV}$	$> 6.00\,\text{GeV}$	—
Dimuon	$> 1.69\,\text{GeV}^2$	$> 2.25\,\text{GeV}^2$	$> 1.69\,\text{GeV}^2$	< 900 hits

Table 4.1: Overview over the different L0 thresholds [17]

low-interacting muons are able to reach them. Module M1 is placed further upstream to measure the transverse momentum more clearly for the trigger system explained later [15]. The modules mainly consist of *Multi Wire Proportional Chambers* (MWPC). Muons ionize the gas in the chambers proportionally to their energy. Kathodes and anodes produce an electric field in the chambers. This field causes the ionization charges to move to the read-out electronics [12].

4.2 LHCb trigger system

The trigger system of the LHCb experiment is a combination of different hardware and software triggers. In the LHCb detector pp interactions occur with a rate of 40 MHz. The purpose of the trigger system is to select interesting ones and thus reduce the interaction rate to an amount that is manageable for the readout electronics. Depicted by figure 4.2 is the general flow of the triggering process.

The first level of triggers is the so-called *L0 Hardware trigger* reducing the interaction rate to 1 MHz. Taking input from the calorimeters for non-muons and the muon system it selects candidates based on their transverse energy E_T and their transverse momentum p_T. Different types of particles have to pass different thresholds as shown in table 4.1 to be accepted. The variable nSPD is the amount of registered particles in the already mentioned scintillation pad detector of the calorimeter system. As the trigger stage has limited time to analyse the decays events with too many events to be quickly checked are discarded.

The transverse energy E_T is measured in clusters of 2x2 cells in the ECAL

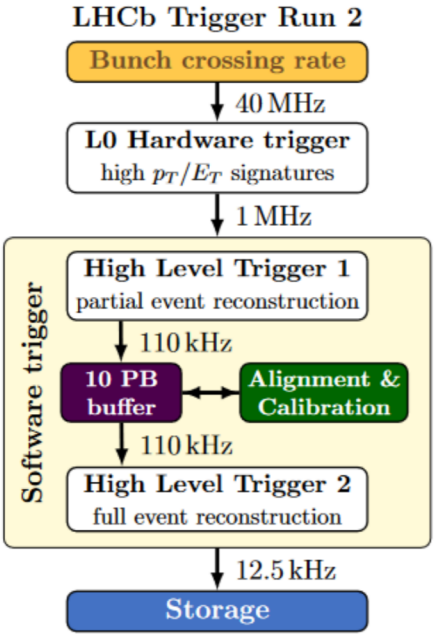


Figure 4.2: Schematic depiction of the triggering process in LHCb [16]

and HCAL defined as:

$$E_T = \sum_{i=1}^{4} E_i \sin(\theta_i) \tag{4.5}$$

where E_i is the deposited energy in each of the four cells. The angle θ_i is the one between the z axis along the beam and the line that connects the center of the corresponding cell with the pp-interaction point [16]. If the detected energy disposition is inside the HCAL it is a hadron. If the energy is only disposed in the ECAL together with a signal in the PS it is either a photon or a muon. Muons are accepted based on their transverse momentum p_T. The L0-muon trigger uses the hits in the tracking stations mentioned before to calculate the muon momentum from the curvature. If the product of the of the largest and second largest muon transverse momentum is above the dimuon threshold the event is also accepted.

After the hardware trigger the interaction rate is further reduced by two software triggers *HLT1* and *HLT2*. The first *High Level Trigger* (HLT1) reconstructs tracks of charged particles. This is done in three steps. First, the vertex is reconstructed in the VELO. Going from there the track is interpolated to the TT and if possible further to the rest of the tracking

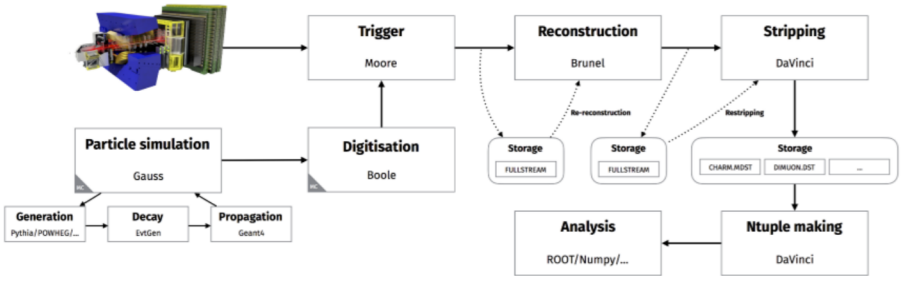


Figure 4.3: Schematic of the dataflow at the LHCb experiment [18]

stations. Those are fitted using *Kalman Filtering* to reject fake tracks [16]. Particle identification is in this step only done for muons as their clear signature in the muon system allows for simple identification. After HLT1 the rate is reduced to 110 kHz which enables the storing of selected events in a 10 PB buffer. The capacity in the buffer is large enough to store about two weeks worth of data. The buffer furthermore acts as a security in case of detector problems which delay the HLT2 step. A further usage of the buffer is the real-time calibration and alignment of the detector. This allows for an increase in reconstruction efficiency [16]. The second *High Level Trigger* (HLT2) has due to the high capacity buffer not the same time-constraints as the HLT1. This allows for full event reconstruction using the full amount of available information of the detector. As such, using RICH and the calorimeters, it is possible to not only identify muons but other particles as well. Furthermore, this produces the highest quality of reconstructed tracks. Applying further selection then results in data to be saved in storage with a rate of 12.5 kHz.

4.3 Dataflow of the LHCb experiment

The previous subsection discussed the trigger stage in detail. This subsection describes the full dataflow of the experiment. Corresponding terms and applications that appear in this analysis are explained. Figure 4.3 shows an overview of the dataflow at the LHCb experiment. Only the generation of real and full detector simulated data differs. Simulated data is also often called *Monte Carlo* (MC) data. Real events are created by pp collisions in the LHC. The generation of simulated events is handled by the framework *Gauss* which coordinates multiple external applications. The initial genera-

tion of events with the desired signal particle is most often done by *Pythia* [19]. If there are unstable particles their decay is simulated by *EvtGen* [20]. Available decay channels, their branching ratios and further configurations are given to EvtGen by *decfiles* [21][22]. Finally, *Geant4* simulates how particles propagate through the detector and how they interact with the detector [23]. In addition, a cut tool might apply additional generator level cuts to the signal events. Then, the application *Boole* turns the simulated interactions of the particles with the detector into signal which mimics the real LHCb signal output [24]. This step simulates the hardware trigger L0 for the MC data. All following steps of the dataflow are the same for real and simulated data.

Which events are selected by the triggers is decided by *trigger lines*. A trigger line is a set of requirements which select a specific type of event. For example the trigger line *HLT1TrackMuon* selects in the HLT1 trigger stage events with a well reconstructed muon track. Trigger lines allow to chose events that are interesting for an analysis. The high level trigger procedure described in the previous section is handled by *Moore* [25].

The event output of Moore is reconstructed by *Brunel*. Brunel reconstructs the full decay chain including all decay vertices and tracks [26].

In the final step of the dataflow, events are further filtered in a process called *stripping* which is done by the application *DaVinci* [27]. *Stripping lines* use the same concept as trigger lines. The stripping lines contain cuts that select events with certain signatures which are interesting for an analysis. The events are at this point fully reconstructed and can thus be selected based on more specific behaviour as in the triggering. The output is saved in usable files containing the information of each selected event.

In an analysis, even though one chooses lines which select interesting events, further filtering is done. Therefore, *preselection* cuts reduce the number of not needed events while *truthmatching* removes misreconstructed events.

The full detector simulation of events takes a long time. Thus, it is acceptable for some studies to use faster, less accurate simulation methods. One of these fast simulations is *RapidSim*. RapidSim is a Monte Carlo generator specialized on the fast simulation of b and c quark hadron decays [28]. It is based on ROOT software and especially the TGenPhaseSpace class. The kinetic properties of the hadrons are simulated using provided *fixed-order next-to-leading-logarithm* (FONLL) calculations.

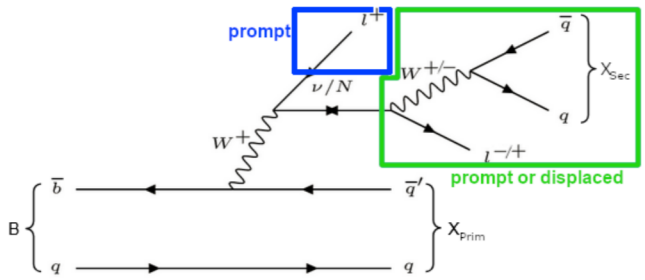


Figure 5.1: Feynman diagram of the analysed (semileptonic) channels. The label X_{Prim} stands for a not reconstructed hadron and X_{Sec} is in this analysis always a pion. The Majorana nature of the HNL allows the secondary lepton to have the same or the opposite sign of the primary lepton.

5 Analysis strategy

As described in section 3.3 technically any channel containing a neutrino can produce an HNL. Heavy neutral leptons in the GeV mass range are produced by B decays in great amount. Thus, such HNLs are searched for at the LHCb experiment which is specialized on B decay measurements. The Feynman diagram 5.1 depicts the exploited channels.

The B meson decays in the general form of $B \to X_{Prim} l\ l\ X_{sec}$. The hadrons X_{Prim} and X_{Sec} origin from the B decay vertex and the HNL decay vertex, respectively. One of the strengths of the LHCb experiment is its detection capabilities of muons. As such the search concentrates on channels that produce two muons. These channels involve weak decays via off-shell W bosons. This results in a clear signal of a prompt muon μ_{Prim} and a likely displaced vertex producing a secondary muon μ_{Sec} and a secondary hadron X_{Sec}. It is possible that no X_{Prim} particle is produced which is the case for purely leptonic decays of the charged B mesons B^+ and B_c^+. In those decays the quarks of the B meson annihilate to produce the initial W boson.

This analysis focuses on the exclusive decay of the HNL. Thus, the X_{Sec} refers to a pion which is directly produced in the HNL decay and not via the decay of an intermediate hadron. The advantage of this approach that the analysis is easier due to the peak in the spectrum of the $\pi\mu$ mass. This wouldn't be the case for an analysis of the fully inclusive channels because X_{Sec} wouldn't be reconstructed.

The investigated HNL mass range is between 1.6 GeV and 6 GeV. The

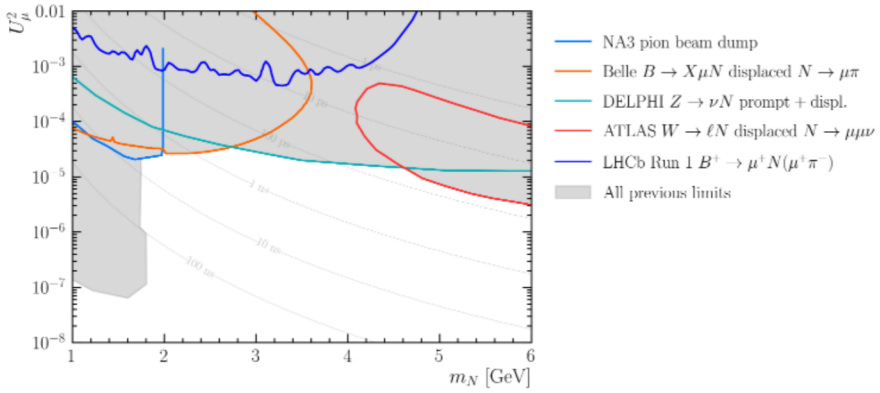


Figure 5.2: 95 % confidence level constraints on the second generation lepton mixing angle U_μ^2 as function of the HNL mass hypothesis set by different experiments. The grey area shows the combined excluded region. As U_μ^2 and the mass of the HNL are directly proportional to the HNL lifetime its contours are also shown. The shown limits are from the NA3 [29], Belle [30], DELPHI [31], ATLAS [32] and LHCb experiments [33]. Status at the start of 2021.

existing limits set by other analyses (status January 2021) are shown in figure 5.2.

The figure depicts which muon neutrino mixing angles and HNL masses are excluded at 95 % confidence level. It shows that we have to investigate HNL lifetimes from below 1 ps to at least 10 ns to test the full mass range as well as regions not excluded by previous analyses. A sensitivity of 10^{-5} on U_μ^2 in the mass range 2 - 6 GeV is required to improve the existing combined limits. For low masses between 2 GeV and 3 GeV this corresponds to a lifetime of 1 ns to 10 ns.

For the mass region of interes the branching ratios of the most important channels of HNL production are shown in figure 5.3. They are calculated following the predictions from *Bondarenko et al* [34]. The BR's are scaled with the fragmentation fraction f_q/f_u measured at LHCb of the corresponding B meson to ease the comparison among different production channels. Note that this plot only depicts theory predictions for the HNL production branching ratio without taking any efficiencies or similar into account. All channels in the box labeled "semileptonic" channels don't reconstruct the X_{Prim} particle. One can see that the leptonic channels $B^+ \to \mu^+ N$ and

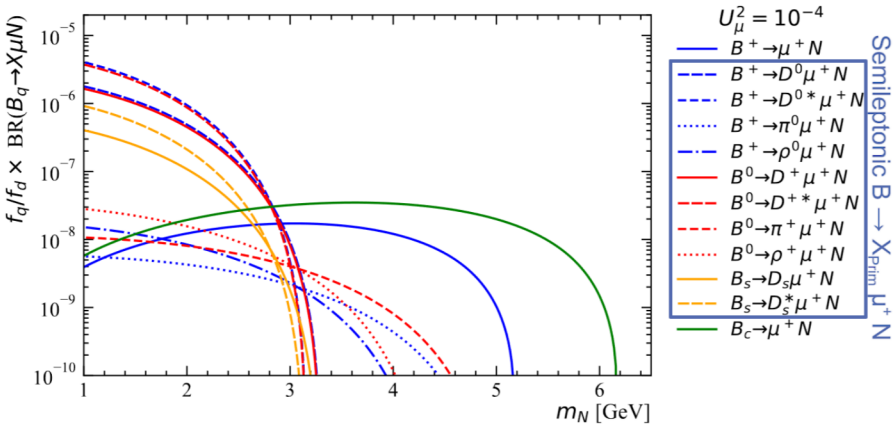


Figure 5.3: Most important hypothetical production channels of an HNL in the mass range [1, 6] GeV. The branching ratios are calculated following the predictions of Bondarenko et al [34]. The branching ratio is scaled with the normalized fragmentation fraction f_q/f_u. Marked in the box are the semileptonic channels while the other two are leptonic.

$B_c \to \mu^+ N$ dominate for high masses while for lower masses semileptonic channels containing a D meson are more sensitive.

Figure 5.4 shows the same plot as before but with contributions merged into three categories which can be either considered leptonic or semileptonic. Those categories are used in the analysis. The two leptonic decays dominate as soon as the mass of the HNL gets too large to produce an additional D meson from the original B meson: $m_N > m_B - m_D$. As there is no non-reconstructed X_{Prim} one expects a peak of the B meson in the $\pi\mu\mu$ invariant mass, which can be used to reject background. In contrast to that, the semileptonic channels with a D meson dominate when the heavy neutral lepton is light enough to allow for the creation of an (excited) D meson: $m_N < m_B - m_{D^{(*)}} - m_\mu$. Those channels have a higher amount of expected events than the leptonic ones. They are however harder to select as there will be only a $m_N = m(\pi\mu)$ peak. This results in a higher background rate.

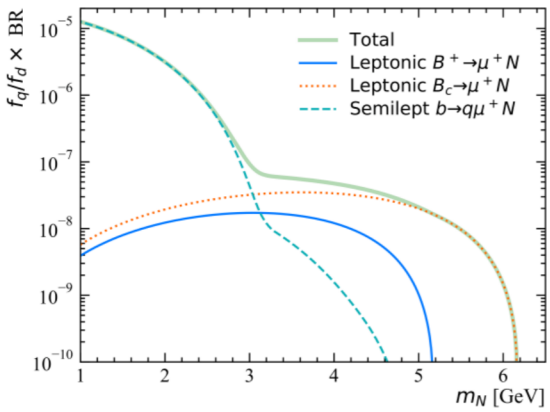


Figure 5.4: Most important production channels of an HNL in the mass range [1, 6] GeV. The channels are the same as in figure 5.3 but grouped into two leptonic and a semileptonic category.

6 Data and simulation samples

This section describes all needed steps to create usable data samples for this analysis. The requirements of the triggering, stripping, preselection, truthmatch and the used fiducial volume are explained.

6.1 Generation of data samples

This analysis is based on pp collision data collected by the LHCb experiment in Run 2. The integrated luminosity of this data set is $1.6\,\mathrm{fb}^{-1}$, $1.7\,\mathrm{fb}^{-1}$ and $2.1\,\mathrm{fb}^{-1}$ for the years 2016, 2017 and 2018, respectively. The HNL search is blinded: the search for an HNL signal peak will be performed only once the analysis is ready and approved by the internal LHCb review. The search strategy and the efficiency calibration is based on Monte Carlo data produced by complete detector simulation. The simulated data without any detector response is called *generator level data*. The one with simulated detector response is called *detector level data*. All the B decay channels in figure 5.3 are produced to be used in the analysis. These semileptonic decay channels constitute the main expected contributions. As such, they provide a lower bound of the total inclusive semileptonic branching ratio. The total inclusive semileptonic BR includes further accompanying hadrons as well as decay channels with multiple hadrons. An overview over the different samples can be found in table 6.1. The data is split into three distinct categories with the following naming convention:

Semileptonic: Semileptonic samples refer to those that have an HNL production channel of the form $B \to X_{Prim}\mu N$. Hereby B refers to a B meson and N to the heavy neutral lepton. The label X_{Prim} is used for any non-reconstructed hadron that is produced alongside the muon and the HNL. Thus, this production mode is called an *inclusive* HNL production channel.

Leptonic B^+: Leptonic samples refer to those where the HNL is produced in a leptonic B decay of the form $B^+ \to \mu N$. As there is no non-reconstructed particle the channel is fully reconstructed. Thus, the mass of the B meson is fully reconstructable. This production mode is also called an *exclusive* HNL production channel.

Leptonic B_c^+: This sample is similar to the leptonic B^+ sample. However, the mother is a B_c^+ meson instead of a B^+.

The semileptonic samples are not dominant for an HNL mass hypotheses

Version	B decay	HNL decay	Muons	HNL τ [ps]	HNL m [GeV]
Sim09j	$B_s^0 \to X_{Prim}\mu N$	$N \to X_{Sec}\mu$	OS, SS	100, 1000	1.6, 2.0, 3.0
Sim09j	$B^0 \to X_{Prim}\mu N$	$N \to X_{Sec}\mu$	OS, SS	100, 1000	1.6, 2.0, 3.0, 4.0, 5.0
Sim09j	$B^+ \to X_{Prim}\mu N$	$N \to X_{Sec}\mu$	OS, SS	100, 1000	1.6, 2.0, 3.0, 4.0, 5.0
Sim09j	$B_c^+ \to \mu N$	$N \to X_{Sec}\mu$	OS, SS	100, 1000	1.6, 2.0, 3.0, 4.0, 5.0, 5.5
Sim09h	$B^+ \to \mu N$	$N \to \pi\mu$	SS	0, 10, (100)	1.0, 2.0, 3.0, 4.0, 4.5, 5.0
Sim09h	$B_c^+ \to \mu N$	$N \to \pi\mu$	SS	0, 10, (100)	1.0, 2.0, 3.0, 4.0, 5.0, 5.5, 6.0

Table 6.1: Overview over used Monte Carlo samples

of 3 GeV onward and they have no sensitivity above 4.5 GeV. Therefore samples for an HNL mass of up to 5 GeV are generated. For the leptonic samples this is increased to an upper limit of 6 GeV. For the hypotheses of the HNL lifetimes, values of 10 ps, 100 ps and 1000 ps are used covering the whole decay time acceptance of the LHCb detector. Further 0 ps samples are solely used for investigation purposes. They are **not** used for setting limits. Additionally samples for both the lepton-number conserving $\mu^+\mu^-$ (OS - Opposing Sign) and lepton-number violating $\mu^+\mu^+$ (SS - Same Sign) hypotheses are generated. For every possible combination of lifetime, mass, muon sign and B meson a different sample is generated as well as for the years 2016 to 2018 resulting in more than 100 different samples. Table 6.1 shows an overview over all the mentioned categories. The branching ratios of each simulated file are set according to the predictions used in figure 5.3. The HNL decays are generated according to a phase-space distribution. In this analysis only the exclusive decay channels of the HNL ($N \to \mu\pi$) are used. Nonetheless, the files contain also decays of the HNL to ρ^+, D_s and D_s^* which will be used in a future analysis.

6.1.1 Triggering and stripping process

Considering the triggering, the following approach is chosen. An overview over the trigger lines is shown in table 6.2. For the Level-0 (L0) hardware trigger the muon is chosen as most important particle due to its clear signal. Events that fulfill the requirements of either the *Muon* or the *DiMuon* line

from table 6.3 are accepted. One or two muon tracks are identified in the muon chambers, their transverse momentum, p_T, is roughly estimated and required to pass a certain threshold. On top of that, the particle multiplicity of the event measured by the number of hits in the SPD detector (nSPD) is required to be lower than a certain threshold to reject events that would take too much time to process in the software trigger stage. The first stage of the software trigger uses the *HLT1TrackMuon* line which requires one of the muons to be reconstructed as track and to fulfill the cuts listed in table 6.3. Only well reconstructed tracks are accepted. This is enforced by the cut on muon track χ^2/dof and on the Ghost probability. The variable χ^2 with respect to the degrees of freedom (dof) of a track measures the quality of the track fit. A Ghost track is a misidentified track. Detector hits from different particles are reconstructed to a false track. Furthermore, the line requires a minimum impact parameter (IP) of the muon with respect to the primary vertex (PV). The IP is the shortest distance of a particles track to a decay vertex.

For the second stage of the software trigger the three most efficient trigger lines are chosen. Events are accepted when they either fulfill conditions of the *Topo(Mu)TwoBody*, *ExoticaRHNu* or *MajoranaBLambdaMuDD* line. TopoTwoBody and TopoMuTwoBody are the main HLT2 lines. They are generally specialised on displaced vertices with two tracks with the TopoMuTwoBody requiring on of them to be a muon. *ExoticaRHNu* is additionally considered to improve the sensitivity in the low HNL mass region. This line also searches for displaced vertices but requires them specifically to be the inclusive $\mu\pi$ vertices. Finally, *MajoranaBLambdaMuDD* is used to identify downstream tracks contributions as efficient as possible. It targets the exclusive decay $B \to \mu N(\to \mu\pi)$ with the HNL vertex being so far displaced that it is reconstructed only downstream of the VELO. This last line is only for same sign muons and has a prescale of 20 % causing poorer results in the downstream category. Work is in progress to develop a new more precise trigger strategy for the downstream track category for the upcoming Run 3.

Offline, the data is additionally filtered through two so-called stripping lines: *Bu2LambdaSSMu* and *Bu2LambdaOSMu*, depending if the sample contains decay channels with same-sign (SS) muons or opposite-sign (OS) muons. The cuts applied by these stripping lines are shown in table 6.4. The decay tracks, meaning the tracks of the pion and the muons, need to have a certain

Trigger Level	Trigger Line	TOS on
L0	*Muon* or *DiMuon*	$\mu\mu$
HLT1	*TrackMuon*	$\mu\mu$
HLT2	*Topo(Mu)TwoBody* *ExoticaRHNu* *MajoranaBLambdaMuDD* (only SS muons, prescale 20 %)	$B \to \mu\mu\pi$

Table 6.2: Overview over the used TriggerLines.

track quality which is again checked by track χ^2/dof. These particles also need to fulfill a minimum requirement on their momentum, p, and transverse momentum, p_T. Their χ^2 on the impact parameter with respect to the primary vertex must be high which assures a reasonable PV assignment. To ensure that the misidentification chance for the muon is low it has to fulfill further cuts. A cut on the ghost probability is again used. The muon *particle identification* (PID), which is basically the probability of the muon being a muon, must be above zero and larger as the probability of it being a kaon or *rho* meson, $PID_\mu - PID_{K,\rho} > 0$. The HNL has a requirement on its mass, m, and transverse momentum, p_T. The HNL vertex also has to be of good quality as measured by the χ^2/dof of the vertex fit. Furthermore, its χ^2-distance from the primary vertex must be large. The mother B meson needs to be in a certain mass range and also has to fulfill a certain vertex quality. In addition, its *diraction angle* (DIRA) which is the cosine of the angle between its spatial trajectory and the combined momenta vector of the daughter particles must be almost one. This ensures a good reconstruction. Finally, the spatial distance on the z axis between the B vertex and HNL vertex is constrained.

6.1.2 Preselection cuts and truth matching

After the process of stripping further cuts and checks to ensure quality are applied. These cuts are split into *preselection* cuts and *truth matching*. The former is shown in table 6.5.

The preselection requires all tracks (muons and pions) to be of good quality (ghost probability $p_{Ghost} < 0.20$). Pions are not much restricted in the stripping as the triggering particles are muons. Thus the preselection cuts ensure that the pions are in acceptance range of the LHCb experiment

Trigger decision	Cut	
L0DiMuon	p_T(both μ)	$\gtrsim 1.5\,\text{GeV}$
	nSPD	< 900
L0Muon	p_T(one μ)	$\gtrsim 1.8\,\text{GeV}$
	nSPD	< 450
HLT1TrackMuon	$p(\mu)$	$> 6.0\,\text{GeV}$
	$p_T(\mu)$	$> 1.1\,\text{GeV}$
	μ Track χ^2/dof	< 3
	μ Ghost probability	< 0.2
	μ Minimum IP w.r.t. PV	> 35

Table 6.3: Cuts applied by the L0 and HLT1 trigger stage.

#	Applied to	Cut	
0	Tracks	p_T	$> 0.25\,\text{GeV}$
1		Track χ^2/dof	< 4
2	Muons	p	$> 3\,\text{GeV}$
3		Minimum IP χ^2 w.r.t. PV	> 12
4		Ghost probability	< 0.5
5		PID_μ	> 0
6		$PID_\mu - PID_{K,\rho}$	> 0
7	Pion	p	$> 2\,\text{GeV}$
8		Minimum IP χ^2 w.r.t. PV	> 10
9	HNL	m	$> 1.5\,\text{GeV}$
10		p_T	$> 0.7\,\text{GeV}$
11		Vertex χ^2/dof	< 10
12		χ^2-distance from PV	> 100
13	B_q	m	$\in (1.5, 6.5)\,\text{GeV}$
14		Vertex χ^2/dof	< 4
15		$DIRA_{PV}$	> 0.99
16		$(B_{vtx} - N_{vtx})_z$	$>$ -1 (+4) mm for SS(OS)

Table 6.4: Overview of applied stripping cuts. They come from the stripping lines Bu2LambdaSSMu and Bu2LambdaOSMu. Tracks referes to the particles with reconstructed track: muons and the pion

#	Applied to	Cut	
0	Tracks	P_{ghost}	< 0.35
1	Pion	InMuonAcc	$= 1$
2		isMuon	$= 0$
3		$ProbNN\pi$	> 0.01
4	HNL	$DIRA_{PV}$	> 0.996
5		$DIRA_B$	> 0.996

Table 6.5: Overview of preselection cuts applied offline on the nTuples.

(InMuonAcc = 1) and that they are not a misidentified muon (isMuon = 0). Furthermore, the probability on the pion being an actual pion needs to exceed a threshold of one percent. For the HNL the *DIRection Angle (DIRA)* is cut on. It is a valuable variable to check if the reconstructed B momentum and the connection of the B decay vertex to the primary vertex align. The efficiency of the preselection is 90.9 % for the same-sign muon signal events and 90.7 % for opposite-sign signal events.
The truth matching is technically not applying cuts in the sense of constraining the range of a parameter. Instead it exploits the fact that only Monte Carlo data is used. As such, full information about the true parameter values of every particle is given. With that it is made sure that the reconstructed candidate particles correspond to the generated particles. In addition, the decay channels we are considering are selected. This means channels with inclusive HNL production and exclusive HNL decay to a muon and a X_{Sec} pion as described by $B_q \to X_{Prim}N(\to \pi\mu)\mu$. The MC data for an inclusive HNL decay will be only used in future analyses. To remove background from misreconstructed candidates it is explicitly checked that all reconstructed particles are matched to the true particle, exploiting MC matching.

7 Generation of full Monte Carlo data set

One special feature of this analysis is the use of several different HNL production channels in an inclusive search. The different data sets are combined into larger sets. This section discusses corrections and additional applied cuts to ensure quality as well as the employed combining method.
Multiple methods in this thesis rely on event weights. Therefore follows a

short introduction about weighted data. In the case of unweighted data, every event has an intrinsic weight of one. Looking at for example a histogram, every event will count as one for the sake of the distribution. In the case of weighted data, every event has an assigned weight that might differ from one. As such in every occasion where the number of events is used, be it for efficiency calculation or for studying distributions, the needed quantity is not the number of unweighted events. Instead the sum of weights is to be used as shown in equation 7.1. This sum of weights, $N_{weighted}$, can be called an effective number of events. In the equation w_i is the weight that is applied to a single event with the index i.

$$N_{weighted} = \sum_i w_i \tag{7.1}$$

If multiple uncorrelated weights are applied to an event they are multiplied to a total weight:

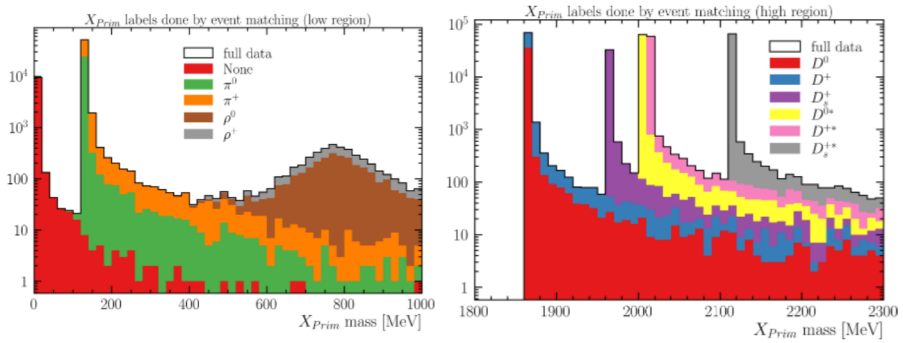
$$w_{tot,i} = \prod_j w_{j,i} \tag{7.2}$$

Where $w_{tot,i}$ represents the total weight of a single event with the index i while $w_{j,i}$ are the different applied weights.
Weighting can not produce more statistics: it merely gives more importance to certain events. As such, an uncertainty σ_w is introduced to the effective number of events. It is calculated by the square root over the sum of squared event weights:

$$\sigma_w = \sqrt{\sum_i w_{tot,i}^2} \tag{7.3}$$

7.1 Labeling of detector level MC data

One problem with the inclusive approach, analysing many decay channels together, is the need of being able to select specific subsets of channels. While the truth value of particles is also saved in the generator level data this is not the case for not reconstructed particles. Thus a method to label the not reconstructed particles X_{Prim} and X_{Sec} in the detector level events is needed. A first approach would be to use the peaks in the X_{Prim} and X_{Sec} mass distributions with added information about the particle charge gained from charge conservation. However, this approach is not feasible as the ρ mesons wide mass distribution leaks into the pion peak.



(a) Assigned X_{Prim} IDs in the lower mass range of [0, 1000] MeV

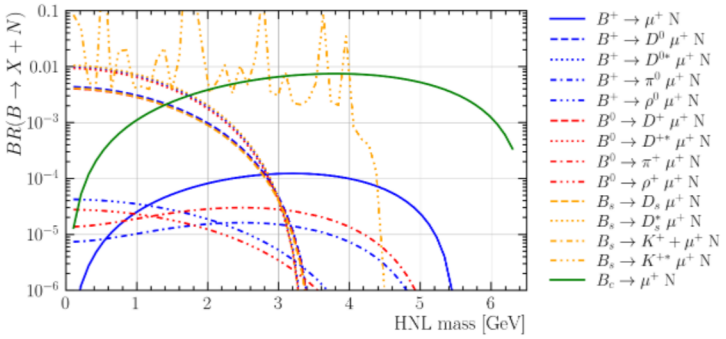
(b) Assigned X_{Prim} IDs in the higher mass range of [1800, 2300] MeV

Figure 7.1: Check of the assigned X_{Prim} IDs. The black outline shows that every event got assigned a label. The used channels have no particle in the area between 1000 MeV *and* 1800 MeV*.*

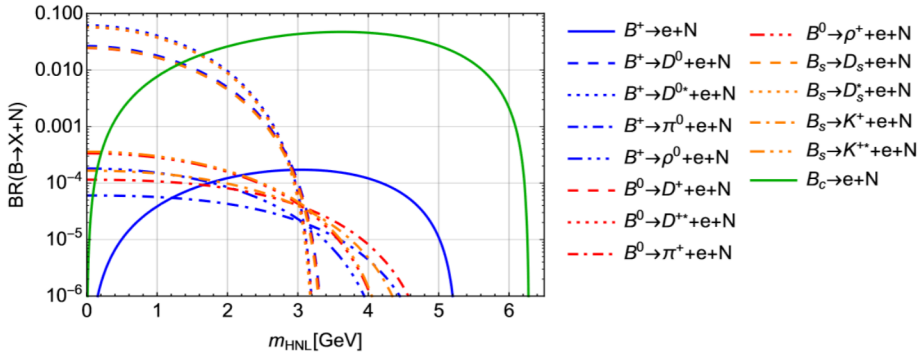
An alternative approach is employed by using the *runNumber* and *eventNumber* of events. These ID's are used by the LHCb simulation framework *Gauss* as seeds for the MC simulation [35]. Combining these two event ID's like "runNumber_eventNumber" creates an ID for every event. This ID is unique and exists both in the generator and detector level data. Using this "uniqueNumber" the missing information about the hadrons X_{Prim} and X_{Sec} can be propagated from the generator to the detector level data.
Figure 7.1 shows the result of the labeling for the X_{Prim} particles. One can clearly see the differentiation between the ρ meson and pions in the bleeding area of 400 MeV to 500 MeV mass. A further advantage of exploiting uniqueNumber is that it is possible to check for inconsistencies between the generator and detector level data, e.g. to check for events that appear in the detector level but not in the generator level data.

7.2 Branching ratio correction

Decfiles are configuration files used by the Monte Carlo event generator *EvtGen* to produce the simulated samples. They contain information on the particle production and decay channels, their branching ratios and applied cuts. As this configuration file steers the simulation of events, it is important to check that they are correct. The samples used in this thesis are produced a while ago by members of the collaboration. Checking



(a) Branching ratio for the inclusive HNL production used in simulation.



(b) Branching ratio for the inclusive HNL production. The figure is taken from a paper of Bondarenko et al [34].

Figure 7.2: Comparison of HNL production branching ratios used for the simulation (a) and the desired theoretical prediction (b).

the production branching ratios of the HNL as given by the decfiles show discrepancies to theory predictions of *Bondarenko et al* [34]. Figure 7.2 shows the comparison between the branching ratios taken from the theory paper and the branching ratios used in the decfiles. The branching ratios for various HNL production channels involving B mesons are scanned in an HNL mass range of 0 GeV to 6 GeV. The beginning and end of phasespace for the leptonic cases as well as semileptonic cases involving a D meson are accurately modeled. Apart from that, the branching ratios are all undershot and furthermore show larger discrepancies in the distribution for the other semileptonic channels.

The channels involving kaons show the largest differences but are actually not used in this analysis. With permission of *Bondarenko et al* their branch-

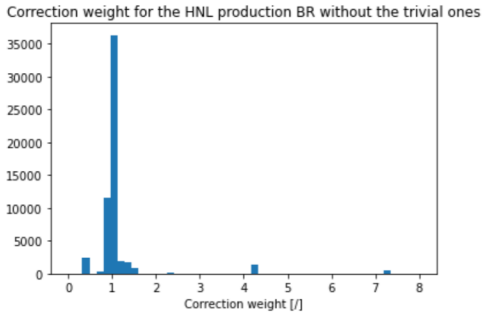


Figure 7.3: Applied weights to correct for HNL production branching ratios incorrectly simulated in MC. The weights that are one are omitted.

ing ratio calculation scripts are used to correct the MC data by applying according weights. In the following these weights will be called *BR correction weights* w_{BRcorr}.

For every decfile the production channels as well as HNL mass and lifetime are extracted. Based on those information the Bondarenko script calculates the absolute branching ratios for every decay. Those are then normalized such that their sum is one. In the same way the BRs from the decfile get normalized. The correction weight w_{BRcorr} is defined as

$$w_{BRcorr} = \frac{BR_{norm,theory}}{BR_{norm,decfile}} \tag{7.4}$$

$BR_{norm,theory}$ is the normalized branching ratio as calculated by the Bondarenko script and $BR_{norm,decfile}$ the one extracted from the decfiles. The weight is the same for all events of the same HNL production channel in the same decfile.

A histogram of the applied weights can be seen in figure 7.3. In case only one production channel for the HNL is in a decfile the normalized BR are by construction one. These trivial weights are omitted in the plot. One can see that most weights are close to one. Those with high weights are channels including a ρ meson.

7.3 Merging of the data samples

One important step in this analysis is the creation of a single inclusive MC "cocktail" containing the different produced samples. By checking and cor-

recting the branching ratios of the decfiles as described beforehand every initial sample has the correct mix of events. Next, one has to combine the initial samples with correct weights with respect to each other. This is done via weights named *cocktail weights* w_{Cockt}. The starting point for this procedure can be seen in equation 7.5. The expected number of events produced in the LHC $N_{ExpGenEvents}$ is derived from known features. Given the integrated luminosity L, the b quark pair production cross section $\sigma(b\bar{b})$ and the fragmentation fraction f_q one gets the number of expected produced B mesons of a given species. The q of f_q stands for the B meson species. Thus, in this analyses f_u, f_d, f_c and f_s are used for B^+, B^0, B_c^+ and B_s^0, respectively. Using every possible combination between an HNL production and decay channel as defined in a decfile one can deduce how many events should on average be described by this specific decfile. For this it is important not to use the normalized ratios of the file itself but to use the absolute branching ratio.

$$N_{ExpGenEvents} = L{\cdot}\sigma(b\bar{b}){\cdot}2{\cdot}f_q{\cdot}\left(\left(\sum_i BR_{Prod,i}\right)\cdot\left(\sum_j BR_{Dec,j}\right)\right) \tag{7.5}$$

The cocktail weight has to make sure that the different samples contribute correctly with respect to each other. Normalizing the expected events, $N_{EexpGenEvents}$, by dividing it with the number of generated events, $N_{GenEvents}$, per decfile results in the cocktail weight. This results in the same weight for all events of the same decfile. The number of generator level data in this case doesn't directly correspond to the number of data found in the initial files, $N_{DetecEvents}$. In the creation process generator level cuts are applied so their efficiency ϵ_{GenCut} has to be taken into account as shown in equation 7.6. The resulting weight can be seen in equation 7.7:

$$N_{GenEvents} = \frac{N_{DetecEvents}}{\epsilon_{GenCut}} \tag{7.6}$$

$$w_{Cockt} = \frac{\frac{f_q}{f_u}\cdot\left(\left(\sum_i BR_{Prod,i}\right)\cdot\left(\sum_j BR_{Dec,j}\right)\right)}{N_{GenEvents}} \tag{7.7}$$

The constant values of integrated luminosity L, the $b\bar{b}$ cross section $\sigma_{b\bar{b}}$, and the factor two can be neglected as they'll drop out in calculated efficiencies.

f_u/f_u	f_d/f_u	f_c/f_u	f_s/f_u
1	1	$2 \cdot (3.78 \pm 0.89) \times 10^{-3}$ [36]	$2 \cdot (0.122 \pm 0.006)$ [37]

Table 7.1: Fragmentation fractions for the used B mesons normalized to the one of B^+. The factor two comes in as the fragmentation fraction is given as $\frac{f_q}{f_u+f_d}$. Due to isospin symmetry one can assume $f_u = f_d$.

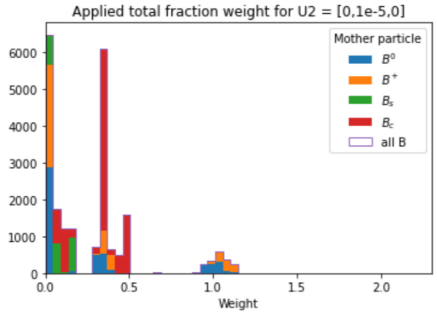


Figure 7.4: Histogram of the total weights for a subsample. One can clearly see the majority to lie in range [0, 1] due to the scaling by 1e13.

Reason being that the cocktail weight appears both for the detector level data in the nominator as well as the generator level data in the denominator of said efficiencies. For the same reason the fragmentation fraction f_q is replaced by the one normalized to the fraction of B^+ as these values are often given in that form. The fragmentation fractions can be taken from table 7.1. As a side note the Bondarenko script that calculates the absolute branching ratios needed for the cocktail weight takes the muon neutrino mixing angle U_μ^2 as input. The mixing angle is proportional to the HNL lifetime which begs the question if a conditional weight is needed for an HNL lifetime scan. However, this is not the case as the branching ratio scales linearly with the lifetime in case of only muon neutrino mixing. The resulting dependency drops also out for the same reason as before. Therefore a flat mixing angle distribution of $U^2 = [0,\ 10^{-5},\ 0]$ is used.

At the end the two weights, cocktail and BR correction weight, are combined to a single weight and scaled by a factor of 10^{13}. With that this weight is roughly between zero and one which ensures easy use and no problems with rounding. It can be seen in figure 7.4 split into samples with different mother B meson.

Applied to	η [/]	p_T [GeV]	p [GeV]
one μ	(2, 4.5)	> 1.1	> 6
Tracks	(2, 4.5)	> 0.3	> 3
HNL	(2, 4.5)	> 1.0	> 15

Table 7.2: Overview over the fiducial volume cuts which are applied to the data. Tracks references all particles with a reconstructed track meaning the muons and the pion.

7.4 Investigation of fiducial volume cuts

Detectors sometimes don't have the same measurement sensitivity in all parts. If this is the case often a *fiducial volume* (FV) is introduced. It's an area in the detector that produces reliable data which is then analyzed. In the case of some experiments like LHCb, ATLAS and others this concept is also widened to include kinematic properties of data. For this analysis a FV according to the latter definition is defined. The fiducial volume cuts are designed based on knowledge gained during studies of this thesis. As later shown, the semileptonic SS category at 2 GeV HNL mass is the most potent sample. Therefore, the cuts are designed mainly around this data. The final version of these FV cuts can be found in table 7.2. Kinematic cuts on the momentum p and transverse momentum p_T of multiple particles as well as a cut on the pseudorapidity η are used. The cuts are applied to the reconstructed final state particles which are the muons and the pion. The X_{Prim} hadron isn't cut on as it is not reconstructed. The cut on the pseudorapidity is a sanity cut that orients itself on the geometric acceptance of the LHCb detector of $2 < \eta < 5$. The cut on transverse and general momenta is oriented on the initial HLT1 trigger for muons as discussed before. Especially the requirement that one of the two muons has to fulfill both the p and p_T requirement is an enforcement of the HLT1 trigger cuts. The cuts of the FV are set on the turn-on curves of the distributions which are caused by different effects. The turn-on and turn-off of the pseudorapidity is caused by the limited geometrical acceptance of the LHCb detector. For the momentum the reason is the triggering stage and the selection cuts. The relatively low transverse momentum resolution at trigger level causes a broader turn-on curve. Furthermore, the tracking of particles performs worse at low particle momentum. This is shown in the figures of 7.5. The distributions have no applied cuts at this point except for those that get

immediately applied by the simulation and tuple generation. This means the events are triggered, reconstructed and pass through the stripping. Furthermore, the cuts are set loosely so that they don't drop below an efficiency of 90 % with respect to the detector level MC data. This is done so that a later applied *Multi Variate Analysis* (MVA) can perform the signal event selection without already losing many signal events beforehand. The preliminary MVA is discussed in section 9. The efficiency of 90 % is checked on semileptonic events with an HNL mass of 2 GeV. The subset of 2 GeV semileptonic events is chosen based on insight gained during the analysis. As shown in later sections this analysis has its best limits on the muon neutrino mixing angle in this specific subset. Therefore a focus is put on it to ensure its quality. The final FV efficiency of the same-sign muon data is 91.4 % while for opposite-sign muons it is 92.0 %.

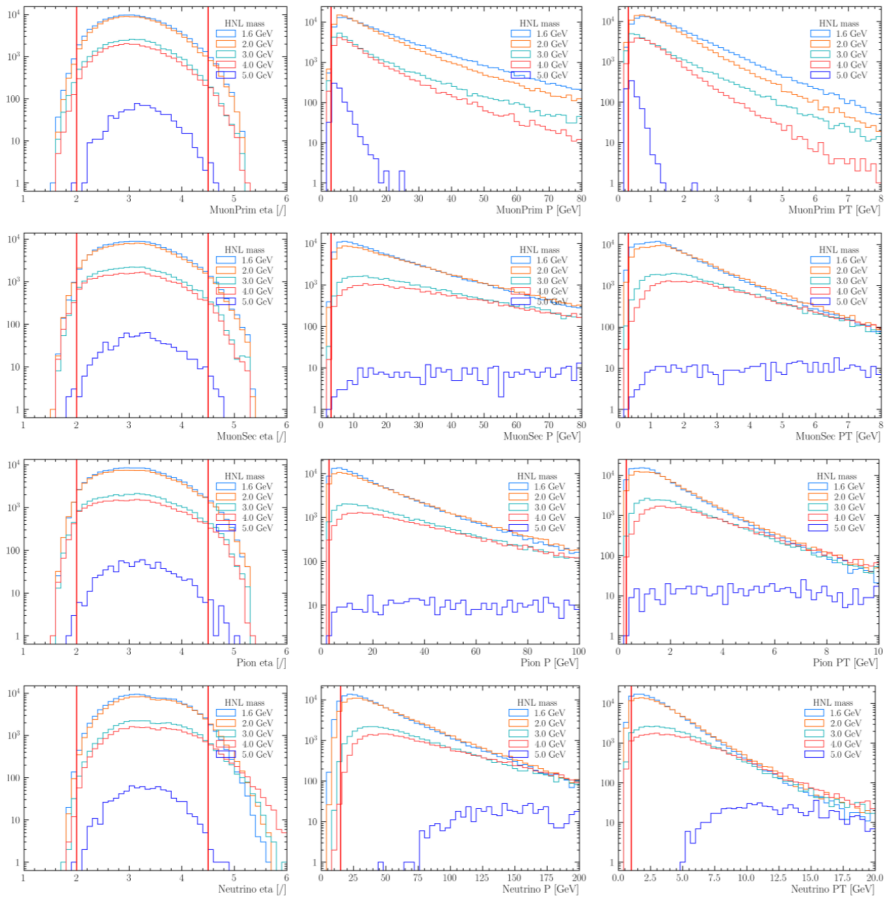


Figure 7.5: Overview of the FV cuts of table 7.2 on detector level data. Tracks corresponds to μ_{Prim}, μ_{Sec} and π. The "or" cut the one muon has to fulfill can't be properly shown in this way.

8 Exclusion limit setting

As the previous chapter discussed the creation of the used MC sample for the analysis this part will discuss the investigation of expected events and the setting of preliminary expected limits. For the setting of limits, the expected amount of events produced and detected by the LHCb detector N_{decay} has to be estimated, which is done fully based on the MC data if not said otherwise. This can be done in similar fashion to the needed expected amount of events in the cocktail weight beforehand.

$$N_{decays} = B_{prod} \cdot HNL_{prod/dec} \cdot \epsilon_{tot} \tag{8.1a}$$

$$B_{prod} = L \cdot \frac{\sigma_{pp\to B^{\pm}X}}{\epsilon_{pp\to B^{\pm}X}} \cdot \frac{f_q}{f_u} \tag{8.1b}$$

$$HNL_{prod/dec} = BR(B_q \to X\mu^+ N) \cdot BR(N \to \mu^+\pi^-) \tag{8.1c}$$

$$\epsilon_{tot} = \epsilon_{FV} \cdot \epsilon_{rec} \cdot \epsilon_{MVA} \tag{8.1d}$$

The corresponding equation 8.1 can be separated into three parts. First B_{prod} is the amount of B mesons produced by the LHC pp collisions. As before $L = (5.1 \pm 0.3)\,\mathrm{fb}^{-1}$ is the integrated luminosity, the amount of pp collisions that occurred during measurement time of the detector with respect to a cross section [38]. The ratio $\sigma_{pp\to B^{\pm}X}/\epsilon_{pp\to B^{\pm}X}$ is the cross section of pp collisions resulting in B mesons over the fraction of those events ending up in the detectable range of the LHCb detector. The detectable range is for transverse momentum $p_T(B^{\pm}) \in [0,\ 40]$ GeV and for pseudorapidity $\eta(B^{\pm}) \in [2.0, 4.5]$. The cross section is measured as $\sigma_{pp\to B^{\pm}X} = (86.6 \pm 6.4)\,\mu\mathrm{b}$ at $\sqrt{s} = 13\,\mathrm{TeV}$ [39]. The efficiency $\epsilon_{pp\to B^{\pm}X}$ is coming from RapidSim studies. The fragmentation fraction f_q/f_u is the same as found in table 7.1. The q in f_q once more refers to the B meson species. This means $q \in [u, d, c, s]$ for B^+, B^0, B_c^+ and B_s^0, respectively.

The second part of the equation $HNL_{prod/dec}$ is the fraction of B mesons decaying into a heavy neutral lepton and which is then decaying via the exclusive channel. As such, $BR(B_q \to X\mu^+ N)$ is the HNL production and $BR(N \to \mu^+\pi^-)$ the exclusive decay to $\mu^+\pi^-$.

The third part ϵ_{tot} is the total efficiency of various methods applied to the

data. The fiducial volume efficiency ϵ_{FV} is derived with the simulation program *RapdiSim* and is evaluated by Rebecca Gartner. A short introduction of her work can be found in subsection 8.1. The reconstruction efficiency ϵ_{rec} is the efficiency between generator level data and detector level data with all applied corrections and cuts. Finally the last efficiency is the one of an applied *multivariate analysis* (MVA). The MVA is worked on by Serhii Cholak a PhD student of EPFL and an overview over it can be found in subsection 9.1. It has to be noted that instead of an MVA efficiency distribution added to the amount of expected events the inclusion of the MVA is done more directly. The MVA cut is applied to the data and as such is later bundled into the reconstruction efficiency. It is clearly said when this is the case. In equation 8.1 it is separated for clarity.

8.1 Evaluation of the fiducial volume efficiency

The evaluation of the fiducial volume efficiency with respect to the mass of the HNL is done by Rebecca Gartner [40]. One problem in investigating the fiducial volume efficiency is that due to its kinematic cuts it has a dependence on the mass of the HNL. As this analysis uses a large amount of different samples it is not possible to fully generate more data than what is already available for the HNL mass points between 1.6 GeV and 6 GeV. Those points are not sufficient to extrapolate between them in a manner that doesn't lose the structures of the efficiency distribution. Therefore, another method is developed. The used approach exploits that the LHCb experiment has an excellent momentum resolution $\Delta p/p$ for charged particles between 0.5 % for low momenta and 1 % for momenta of about 200 GeV [41]. As the fiducial volume cuts are all based on the momentum it is therefore possible to use generator level data instead of detector level data for the FV efficiency. This allows to use the fast Monte Carlo generator *RapidSim* and to define the FV efficiency ϵ_{FV} as:

$$\epsilon_{FV} = \frac{\text{Amount of generator level events after FV cuts}}{\text{Amount of generator level events before FV cuts}} \tag{8.2}$$

RapidSim allows for the quick production of large MC samples. In this case several large samples are produced for the HNL mass range [1.6, 6] GeV and further investigated. The studies demonstrated that RapidSim is reproducing the full simulation well. RapidSim data is therefore used to evaluate the

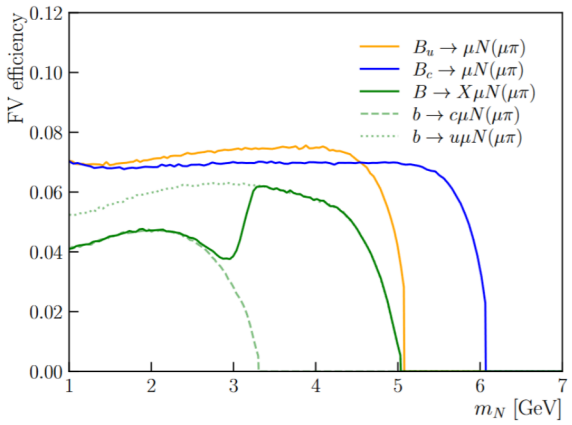


Figure 8.1: Fiducial volume efficiency with respect to the mass of the heavy neutral lepton. The semileptonic green curve is split into the contributions of X_{Prim} containing a c quark or a u quark.

fiducial volume efficiency. Figure 8.1 shows the results of the studies applied to the FV mentioned in table 7.5.

The results are split into the three categories used in the analysis. In orange and blue the two leptonic production channels are shown. The distributions end when the mass of the HNL gets too large for an additional muon. The green line shows the efficiency of the semileptonic sample and presents a large structure around an HNL mass of 3 GeV. By splitting the contributions into the part where X_{Prim} contains a c quark (is a D meson) or a u quark (is a pion or ρ meson) shows the reason. For low masses the decay is dominated by the channels containing a D meson. At around 3.5 GeV this phase space closes and instead the channels for the ρ meson and the pions dominate which have a higher efficiency. In general the channels with a lighter or no X_{Prim} hadron have a higher FV efficiency as more of the kinetic energy is available for the other particles. As a lower cut on the momentum, p, and transverse momentum, p_T, is set these events are more likely to pass the selection.

8.2 Evaluation of the reconstruction efficiency

The second ingredient needed for the total efficiency is the *reconstruction efficiency* ϵ_{rec}. The efficiency shows how many events are properly reconstructed in the detector and pass the following quality assuring offline cuts.

As such, the efficiency is described by equation 8.3 which will be explained more thoroughly after first needed information.
In the following the initial 3×2 categories semileptonic, leptonic B_c^+, leptonic B^+ and same-sign muons (SS), opposite-sign muons (OS) are used. They are split further in three categories based on the displacement of the HNL decay vertex and the reconstruction of the tracks.

Prompt: The prompt category contains all events where the vertex of the HNL decay is less than 20 mm apart from the B meson vertex on the z-axis along the particle beam: $(N_{vtx} - B_{vtx})_z < 20\,\text{mm}$. The 20 mm are chosen based on studies from Serhii Cholak of EPFL This category contains about 85 % of the background before the MVA is applied. A further important information is that the stripping of the OS events includes a cut that forces the vertices in question to be at least 5 mm apart: $(N_{vtx} - B_{vtx})_z > 5\,\text{mm}$. As such, the category called prompt in the following fulfills these conditions but is not prompt in the usual LHCb meaning.

Displaced long track / Long Long (LL): The displaced long track category enforces the inversion of the prompt category. The vertices of B meson and HNL must be at least 20 mm apart: $(N_{vtx} - B_{vtx})_z > 20\,\text{mm}$. Furthermore, the decay of the HNL, $N \to \mu\pi$, must happen inside the vertex locator. If thus the track of the produced pion or muon is fully detected it is called a *long track*, hence the name of the category. The category explicitly checks for long tracks. Long tracks in general have the best performance due to being detected by all systems of the LHCb detector. Therefore this is the category with the best sensitivity.

Displaced downstream / Down Down (DD): Like the other displaced category the displacement of $(N_{vtx} - B_{vtx})_z > 20\,\text{mm}$ is enforced. In contrast to the long track category these tracks did not leave a signature in the VELO. They are reconstructed based on tracker information further downstream, hence the name *downstream track*. In general, this category performs worse than the long track one as information from the VELO is missing making vertex reconstruction less reliable. A further problem is that the HLT2 trigger line for this category has a prescale of 20 % which means only a fifth of the events is actually saved. Moreover, the line only works for

SS muons events. Nevertheless, this line isn't dismissed due to importance for the semileptonic category as discussed later.

The needed variables only appear in the detector level MC data. Thus, it is not possible to distinguish the categories in the generator level MC data. As an example when a reconstruction efficiency is given for the semileptonic SS prompt events it means that all used generator and detector level events belong to the semileptonic SS category while only reconstructed events belong to the prompt category.

The calculation of the reconstruction efficiency is done in a large range of mass points like the fiducial volume efficiency. RapidSim can't be used for this task as it can't properly generate full detector simulation. For that reason another method to interpolate between mass points is used. The reconstruction efficiency further depends on the lifetime of the HNL. Just as for the masses there are only at most three different lifetime points in every category. Therefore a reweighting method is employed to enable the scan over a large range. This method is only applied to the detector level for reasons explained in the corresponding section.

With this information in mind the efficiency is described by equation 8.3. Hereby N_{rec} is the amount of detector level candidates and N_{gen} the amount of generator level candidates. Both MC data sets have the fiducial volume cuts as well as the truth matching applied as those ensure a general quality of the data independent of the reconstruction process. The offline cuts on the detector level data corresponds to the preselection cuts discussed earlier in table 6.5. Both datasets have their full branching ratio correction and cocktail weights applied. As such, the number of effective events is calculated according to equation 7.1 by summing over the total weights applied to each event. As a reminder the total weight w_{tot} is the product of the branching ratio correction w_{BRcorr} and cocktail weight w_{cockt} for the generator level data. The detector level has furthermore the mentioned lifetime reweighting which is labeled here as w_τ. The lifetime reweighting is explained in subsection 8.2.1.

$$\epsilon_{rec} = \frac{N_{rec}(\text{with offline and FV cuts applied})}{N_{gen}(\text{with FV cuts applied})} \tag{8.3a}$$

$$= \frac{\sum_i w_{rec,tot,i}}{\sum_j w_{gen,tot,j}} \tag{8.3b}$$

$$= \frac{\sum_i w_{rec,BRcorr,i} \cdot w_{rec,cockt,i} \cdot w_{rec,\tau,i}}{\sum_j w_{gen,BRcorr,i} \cdot w_{gen,cockt,i}} \tag{8.3c}$$

For the uncertainty on the efficiency ϵ_{rec} the standard method for weighted data is used. Using uncertainty propagation on poissonian statistics shows that the uncertainty is the square root of the sum over squared weights [42].

$$\sigma(\epsilon_{rec}) = \frac{\sqrt{\sum_i w^2_{rec,tot,i}}}{\sqrt{\sum_j w^2_{gen,tot,j}}} \tag{8.4a}$$

$$= \frac{\sqrt{\sum_i (w_{rec,BRcorr,i} \cdot w_{rec,cockt,i} \cdot w_{rec,\tau,i})^2}}{\sqrt{\sum_j (w_{gen,BRcorr,i} \cdot w_{gen,cockt,i})^2}} \tag{8.4b}$$

At this point it is important to discuss the matter of correlation between the generator and detector level data. The uncertainty according to equation 8.4 assumes no correlation between the nominator and denominator even though all reconstructed events are included in the generator level set as well. The reason for this assumption is the following. To be able to have no correlation between the samples one could split the generator level data into two subgroups: Events that are reconstructed N_{rec} and those that are missed $N_{miss} = N_{gen} - N_{rec}$. This would mean the efficiency would look like equation 8.5

$$\epsilon'_{rec} = \frac{N_{rec}}{N_{rec} + N_{miss}} \tag{8.5a}$$

$$= \frac{\sum_i w_{rec,tot,i}}{\sum_i w_{rec,tot,i} + (\sum_j w_{gen,tot,j} - \sum_i w_{gen,tot,i})} \tag{8.5b}$$

$$\neq \frac{\sum_i w_{rec,tot,i}}{\sum_j w_{gen,tot,j}} \tag{8.5c}$$

If summing over i means summing over the events that will be reconstructed

one can see that this method fails due to the lifetime weighting only being applied to the detector level data, $w_{rec,tot,i} \neq w_{gen,tot,i}$. Using the weights $w_{rec,tot,i}$ from the reconstructed and therefore lifetime reweighted events results not in the same sum as what would be left over of the not lifetime reweighted events with weight $w_{gen,tot,i}$. For this reason this can't be done and the assumption of no correlation is used. To further motivate this one can also look at first reconstruction efficiencies that are calculated. They showed an efficiency of at most 15 % with a tendency to be lower. This means that one can say that the effect of left out correlation is low as the amount of not reconstructed events is about an order of magnitude higher than the amount of reconstructed ones.

8.2.1 Lifetime reweighting

A lifetime reweighting is done to determine the reconstruction efficiency for different HNL lifetimes than the ones initially simulated in the Monte Carlo data. Particles follow the law of exponential decay describing the decay time as exponential function with a lifetime parameter, τ. As shown in equation 8.6 one can use this as an Ansatz to define a lifetime dependent weight $w'_{\tau_{gen}}(\tau_{new}, t)$. The weight for an event with the decay time t is the ratio between the normalized exponential function of the initial lifetime τ_{gen} and the to be investigated lifetime τ_{new}. For infinite statistics the sum of events doesn't change: $\sum w'(\tau_{new}, t) = N_{gen}$.

$$w'_{\tau_{gen}}(\tau_{new}, t) = \frac{\tau_{gen}}{\tau_{new}} \frac{exp(-t/\tau_{new})}{exp(-t/\tau_{gen})} \tag{8.6}$$

However, this is only true for infinite statistics. In reality biases are introduced if one tries to weight from lower to higher lifetimes. The reason is depicted in figure 8.2. Longer lifetimes have a less steep exponential function. This means that the reweighting has to use large weights for few events at high decay time to try to reproduce the distribution. At some point the original sample has no events left to reweight. As such, the new distribution ends abruptly causing a bias as shown in figure 8.3.
This figure is a result of toy studies. Several toys with $N_{gen} = 10^4$ each are produced according to an exponential function with the lifetime $\tau = 1\,\text{ps}$. Those are reweighted according to equation 8.6. Next, the number of reweighted events in form of the sum of their lifetime weights, $\sum w'(\tau_{new}, t)$,

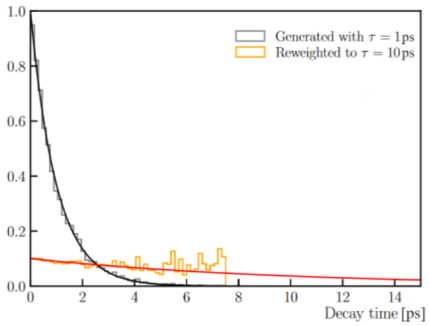


Figure 8.2: Depiction of an exponential decay distribution with lifetime $\tau = 1$ ps reweighted to $\tau = 10$ ps. As no events are left in higher decay times there is nothing to be reweighted which causes the cutoff.

is compared to the initial number of events, N_{gen}. The mean and standard deviation of the toy samples are plotted in red and orange, respectively. As one can see the reweighting to lower lifetimes introduces no bias but just a larger standard deviation. In contrast to that, for higher lifetimes the effect of finite statistics can be seen. As there are no events for high enough decay times the normalization doesn't keep the number of events constant. The new number of events undershoots the initial one. Additionally, the standard deviation corresponding to the standard sum of squared weights is plotted in blue. The toy spread is well approximated for reweighting to lower lifetimes. For higher lifetimes the spread diverges from the standard deviation. The fact that reweighting from lower to higher lifetimes introduces a bias proportional to the difference between initial and new lifetime motivated the search for a method to lower this effect. Otherwise the mixed sample with the three lifetimes 10 ps, 100 ps, 1000 ps would introduce a bias on the limit setting.

Firstly, only the detector level MC data is lifetime reweighted. This can be done as the reweighted data is used to calculate the reconstruction efficiency according to equation 8.3 which uses the effective number of events. By construction, the number of generated events is unchanged by the weighting function since the latter is normalised. As its effective number stays constant so does the denominator of the reconstruction efficiency and no generator level reweighting is needed. Furthermore, it is to be made clear that a weight doesn't influence which events are filtered by the cuts. Instead the weight only changes the importance of certain events.

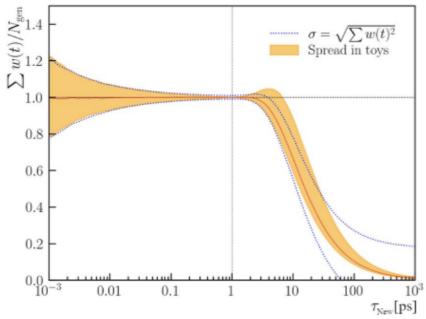


Figure 8.3: Reweighting of a toy of $N_{gen} = 10^4$ events distributed according to a lifetime $\tau = 1\,\text{ps}$ to different lifetimes τ_{new}. Shown is the ratio between the sum of weights and the initial number of events, N_{gen}, which in case of no bias should stay one. For reweighting to lower lifetimes only the uncertainties rise (orange). For higher lifetimes a bias by the attempted reweighting of nonexistent events of higher decay times is introduced.

Secondly, a further normalization factor is introduced. Every reconstructed MC category (e.g. the prompt same-sign muons semileptonic category) is weighted independently. For each of these categories, samples with different lifetimes are available. For example, in the prompt same-sign muons semileptonic category there are samples with 10 ps, 100 ps and 1000 ps HNL lifetime. For each of these samples the highest lifetime weight is determined and used as conditional normalization factor as shown in equation 8.7. Furthermore, even though the generator level MC data doesn't get reweighted it also gets the normalisation factor as shown by N_{gen} in equation 8.8. Like this samples with a lifetime very different from the target lifetime get downweighted. Using this method guarantees that the sample which has the lifetime closest to the one which is investigated has the largest contribution.

$$w_{\tau_{gen}}(\tau_{new}, t) = \frac{1}{max(w'_{\tau_{gen}}(\tau_{new}, t))} \frac{\tau_{gen}}{\tau_{new}} \frac{exp(-t/\tau_{new})}{exp(-t/\tau_{gen})} \tag{8.7}$$

$$N_{rec}(\tau_{new}) = \sum_{\tau_{gen}} \left(\sum_{t} w_{\tau_{gen}}(\tau_{new}, t) \right) \tag{8.8a}$$

$$N_{gen}(\tau_{new}) = \sum_{\tau_{gen}} max(w'_{\tau_{gen}}(\tau_{new}, t)) N_{\tau_{gen}} \tag{8.8b}$$

where t sums over all events in one sample and τ_{gen} sums over all generated samples.
To further illustrate the lifetime reweighting an example is discussed. Assume the prompt same-sign muons semileptonic category with lifetimes $\tau_{gen} \in [10,\ 100,\ 1000]$ ps is reweighted to a 100 ps lifetime. The detector level events are reweighted according to equation 8.7. For each of the three samples the highest weight is used to normalize the weights:

- Case 10 ps $\rightarrow$ 100 ps: The exponential distribution gets flatter. The events with low decay time have to be reweighted to have a lower contribution while the events with high decay time events get a lower weight. The event with the highest decay time has the largest weight which causes a small factor. Thus, the importance of the sample is low.

- Case 100 ps $\rightarrow$ 100 ps: By construction all the weights are one. As such, no penalty in form of the normalization factor is applied and the sample has the highest importance.

- Case 1000 ps $\rightarrow$ 100 ps: In opposition to the first case the low decay time events get a large weight while the high decay time ones get a weight smaller than one. The highest weight is at $t = 0$ ps and the importance of this sample is scaled down accordingly.

As one can see on this example the importance of a subset is higher the closer its original lifetime it is to the investigated one. By this method the smallest possible bias is introduced for each new HNL lifetime. This method works especially well on the displaced category for the reason that in case of high to low lifetime weighting the largest weight is for the event with the lowest decay time. In the displaced category there are no reconstructed events at low lifetimes therefore the penalty is less.
To test the method performance the reconstruction efficiencies of all possible categories and HNL masses are calculated. Each category with its different lifetime samples is reweighted to the lifetimes it is composed of and the efficiency is calculated for every HNL mass. To check the performance of the reweighting the reconstruction efficiency of data of the same category is used. This reference data is not lifetime reweighted as it only contains events with the tested HNL lifetime. For example samples with 10 ps, 100 ps

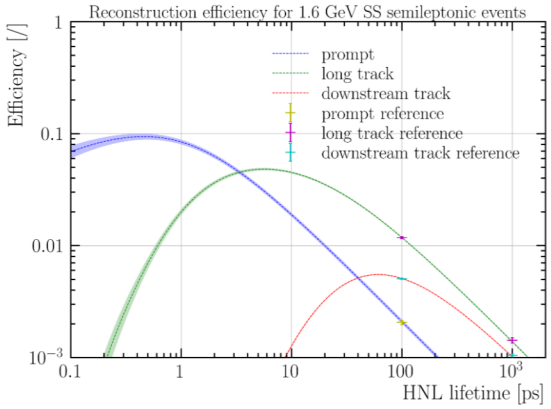


Figure 8.4: Example reconstruction efficiency of all three displacement categories of the 1.6 GeV *same-sign muons semileptonic events set. The reference efficiencies are also marked.*

and 1000 ps are reweighted to 100 ps. The reference MC data in this case are the non-reweighted 100 ps events. In case of unbiased reweighting the ratio between the efficiencies should be one. The full set can be found in the appendix in table 10.1 to 10.3. An overview over the performance of the improved weighting method can be seen in table 8.1. There one can see which fraction of the ratios are more than 1, 2 or 3 σ away from the desired ratio of one. The row labeled "Ansatz" refers to the weighting without the maximum weight normalization (eq. 8.6). "Normalized" refers to the final used version.

Weighting method	$> 1\,\sigma$	$> 2\,\sigma$	$> 3\,\sigma$
Ansatz	36.31 %	14.01 %	6.37 %
Normalized	8.28 %	0.64 %	0.64 %

Table 8.1: Comparison between the reference checks for the initial reweighing Ansatz and the one with the normalization by the highest weight. It shows how many ratios between the reference and reweighted reconstruction efficiencies are more than one, two or three sigma away from one. The correlation between the original and the reweighted sample are not taken into account.

8.2.2 Mass interpolation

As already mentioned the large amount of different samples used in this analysis means that only about six HNL mass points are generated. For the fiducial volume efficiency RapidSim is used to generate more points for a scan over the mass range. As RapidSim cannot be used to determine the reconstruction efficiencies interpolation is used.
For the interpolation between the mass points *cubic splines* are used. The corresponding fitting algorithm uses four knots and low-degree polynomials between them as fit function. This Ansatz is chosen because single high degree polynomials tend to oscillate between points. The spline fitting takes the initial mass points reweighted to a given lifetime. It interpolates only between the given mass points not over the highest or below the lowest as reasonable behaviour cannot be ensured there. To also interpolate uncertainties of the reconstruction efficiency 5000 toys are generated. Each toy contains one efficiency per initial mass point. The efficiencies are drawn from its according Gaussian distribution described by the initial reconstruction efficiency and uncertainty. Every toy set is fitted with the cubic spline algorithm resulting in 5000 fit functions. From these 5000 functions the mean is taken as the interpolated efficiency function depending on the HNL mass. The standard deviation of the fitted functions is the corresponding uncertainty.
To study the behaviour over the mass range the reconstruction efficiencies of different categories are compared. As there is no HNL lifetime which is contained in every category the MC data for all mass points is reweighted to 10 ps. Thus, all events are reweighted to a lower or the same lifetime which is shown to introduce almost no bias. This allows to investigate without potential influence from lifetime reweighting. Figure 8.5 shows the overview for the long track displacement category.
Depicted are the interpolated efficiencies with their mean as dashed line and their $1\,\sigma$ uncertainty, as shaded area. In addition the reference points of available mass points are added. This illustrates the quality of the interpolation fitting. Multiple trends are shown in this study. In case of the prompt and long track categories the efficiency shows a maximum in the range between 3 GeV and 4 GeV HNL mass while falling off towards low and high masses. A reason for this is the nature of the selection. Low HNL masses mean that the secondary muon coming from its decay is soft,

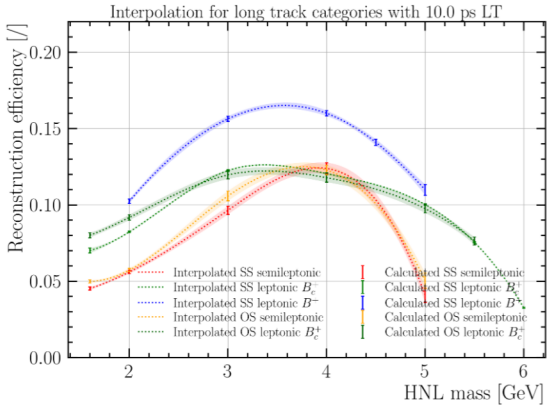


Figure 8.5: Reconstruction efficiency of the long track categories interpolated over the masses. The used MC data is reweighted to 10 ps *HNL lifetime. The reference efficiencies at the available mass points are shown. For the interpolation the dashed line is the mean while the shaded area the* $1\,\sigma$ *range.*

meaning it has a low momentum. On the other side HNLs with high masses take most of the available center-of-mass energy resulting in soft primary muons. Looking back at the used triggers and stripping lines, at all stages a muon momentum threshold is required to be passed. This results in a worse efficiency in the extreme mass areas.

For the downstream categories (Fig. 8.6) the efficiency also has the fall-off towards higher masses but for low HNL masses it rises instead. To explain this one has to keep in mind that the downstream category has its HNL decay vertex outside of the VELO. This causes a worse reconstruction efficiency for the secondary muon. Without this secondary muon the HLT1 absolutely needs the primary muon to have a high enough momentum. Therefore, for increasing HNL mass the softer primary muon decreases the efficiency. As such, the downstream category contributes the most at low masses. One can further see that for the low efficiency downstream OS categories the algorithm has problems fitting the points. For the OS semileptonic category this appears as an oscillation. In addition, the interpolated efficiency of the OS leptonic B_c^+ category briefly drops below zero. This is taken into account during the limit setting. A second trend can be seen regarding the efficiency of the leptonic categories. Generally the leptonic B^+ events perform better than the B_c^+ ones. One reason for this can be found in the lifetime of the

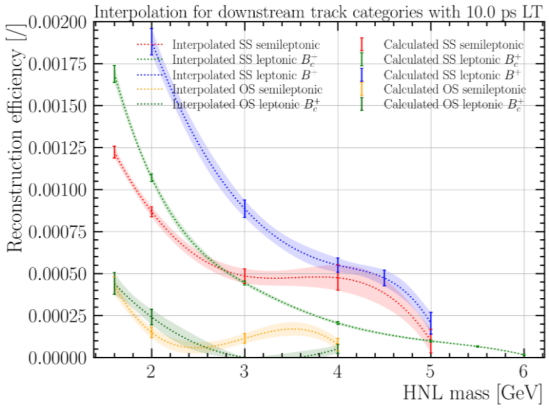


Figure 8.6: Reconstruction efficiency of the downstream categories interpolated over the masses with the same considerations as figure 8.5

initial B meson. B^+ has a lifetime of $\tau_{B^+} = (1.638 \pm 0.004)\,\text{ps}$ while B_c^+ has a lifetime of $\tau_{B_c^+} = (0.510 \pm 0.009)\,\text{ps}$ [43] resulting in, on average, a lower displacement of the B decay vertex in case of the B_c^+. The selection includes minimum requirements on the impact parameters of tracks and displacement of the B vertex and therefore is less effective at selecting B_c^+ decays.
Finally, one can see an efficiency drop in figure 8.7 for the prompt OS leptonic B_c^+ category at 4 GeV HNL mass. There is no physical reason why the shape should vary that strongly from it SS counterpart which is also shown. Due to the large uncertainties it is most likely to be a statistical fluctuation.

8.3 Reconstruction efficiency map creation

Using the two previously discussed scanning techniques a full scan over the region of interest is performed. The HNL lifetime range of 1 ps to 10 ns is scanned in logarithmic fashion by reweighting the MC data to it. For every lifetime step then the mass is interpolated as described and the efficiencies calculated. With this approach it is possible to create 2D maps of the reconstruction efficiency. In figure 8.8 two examples are depicted. A full set of maps for all categories can be found in the appendix. The maps confirm trends seen in the sections of mass and lifetime interpolation already. In low efficiency areas, especially for low lifetime downstream tracks, the interpolation is not fully flawless. Investigation of this showed that this has no influence on the ability to set a limit. As already seen beforehand the

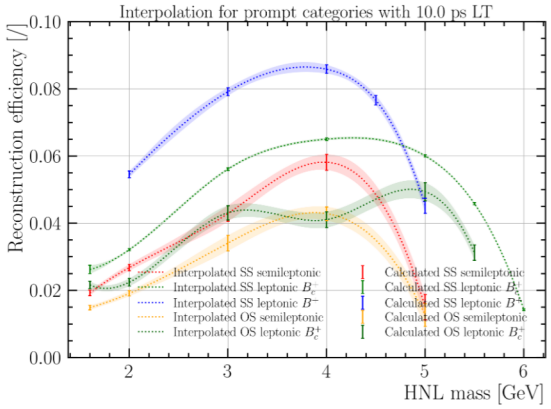


Figure 8.7: Reconstruction efficiency of the prompt categories interpolated over the masses with the same considerations as figure 8.5

most interesting areas are for high lifetimes rendering such imperfections negligible. Table 8.2 shows the highest efficiency per category. The given uncertainty is the one introduced by the weights as discussed. It has to be noted that the highest efficiency doesn't necessarily correspond to a strong limit later on. The reason is that it is possible that the highest efficiency is at an HNL mass and lifetime combination which is not relevant for a competitive limit.

Further tests include the investigation of the prescaled HLT2 trigger line *MajoranaBLambdaMuDD* mentioned in subsection 6.1.1. This line is included to enrich the downstream category but due to its prescale its efficiency is low. To see if a future use of a not prescaled version is justified an additional map is created with events upscaled to their proper rate. Events that are only triggered by a single prescaled line, in this analysis MajoranaBLambdaMuDD, received an additional weight of five counteracting the 20 % prescale. If the vast majority of the events in this category being triggered by this specific line the efficiency should rise by a factor of five. Figure 8.9 shows the reconstruction efficiency map for the downstream SS semileptonic category. Hereby the left side is the one with the prescale still applied while the right one has the counteracting weight. It's visible that as expected the general shape of the efficiencies doesn't change as only a further constant weight is applied. To have a closer look at the increase figure 8.10 shows the ratio between the MC data set with and without the

Category			Reconstruction efficiency [%]
Prompt	SS	semileptonic	15.80 ± 0.80
Prompt	SS	leptonic B_c^+	16.72 ± 0.12
Prompt	SS	leptonic B^+	24.80 ± 0.50
Prompt	OS	semileptonic	8.30 ± 0.40
Prompt	OS	leptonic B_c^+	9.10 ± 0.50
Long track	SS	semileptonic	12.62 ± 0.32
Long track	SS	leptonic B_c^+	12.87 ± 0.05
Long track	SS	leptonic B^+	16.83 ± 0.21
Long track	OS	semileptonic	12.64 ± 0.27
Long track	OS	leptonic B_c^+	12.51 ± 0.23
Downstream	SS	semileptonic	0.55 ± 0.01
Downstream	SS	leptonic B_c^+	0.73 ± 0.01
Downstream	SS	leptonic B^+	0.87 ± 0.18
Downstream	OS	semileptonic	0.10 ± 0.01
Downstream	OS	leptonic B_c^+	0.14 ± 0.01

Table 8.2: Reconstruction efficiencies of the different categories. The given value is the highest reconstruction efficiency found in the entire mass and lifetime range.

prescale. A rather flat increase by a factor of four is seen with a drop off above 4 GeV. It isn't the full factor five expected by the prescale but it shows the worth of having a new dedicated stripping line for this category.

8.4 Investigation of zero-background exclusion limits

Having investigated the fiducial volume as well as the reconstruction efficiency the calculation of a total efficiency misses only the efficiency of an applied multivariate analysis (MVA). To construct exclusion limits from the total efficiency also requires a background rate. In this case the background is assumed to be zero. At the point of this analysis only a preliminary MVA is available. Limits set with it can be seen in the section 9. Before that, the sensitivity is studied without the MVA to determine future search strategies. Note that at this point the analysis is still fully blinded. The following expected limit is a projection of the analysis capabilities. For this the expected number of events including the efficiencies assuming an MVA with 80 % efficiency are calculated using equation 8.1 and the previous findings. For the case of the inclusive semileptonic production of the HNL the expected number of events for all different production channels in

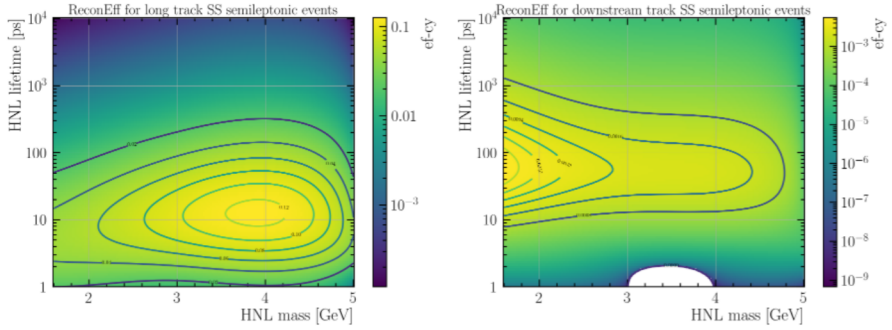


Figure 8.8: Two examples for reconstruction efficiency maps. The left one shows a map without any problems. On the right hand the white area is a low efficiency part which the interpolation caused to be negative. Such unstable areas of low efficiency don't contribute to the limit setting and are excluded. The appendix contains larger versions of all categories.

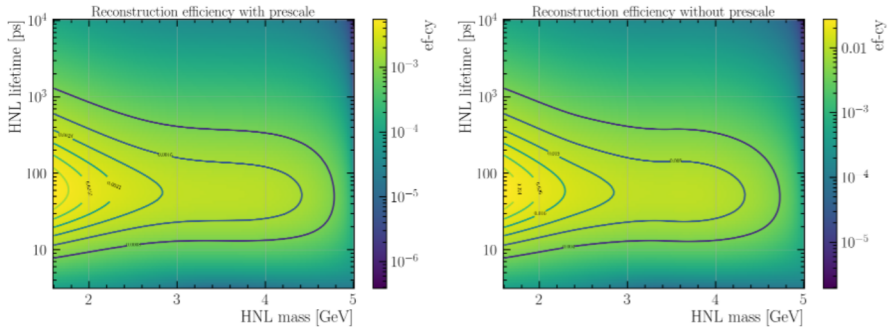


Figure 8.9: Comparison of the downstream SS semileptonic reconstruction efficiency map with (left) and without (right) the prescale of 0.2.

the MC data are calculated and summed up. The sensitivity is investigated with respect to the mass of the HNL and the mixing angle U_μ^2. Similar to the reconstruction efficiency maps the range $m(HNL) \in [1.6,\ 6]$ GeV and $U_\mu^2 \in [10^{-2}, 10^{-8}]$ is tested. The HNL lifetime used in the reconstruction efficiency maps are translated into the corresponding mixing angles. The mentioned zero-background assumption is reasonable considering the displaced vertex of the downstream and long track categories as well as the preliminary MVA rejecting 99.9 % of background. For the prompt category it is slightly optimistic given that our studies showed that without filtering about 85 % of the background will be found there. Nonetheless it can be

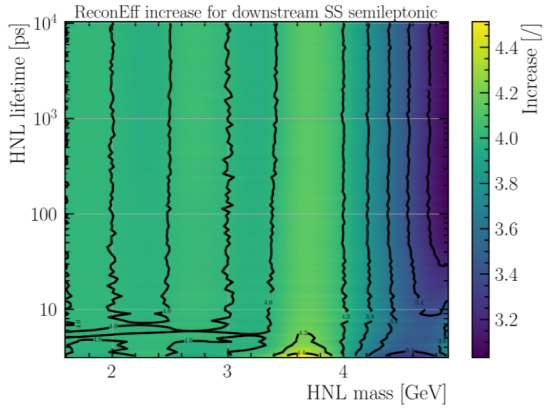


Figure 8.10: Increase in the reconstruction efficiency due to removal of the prescale shown in figure 8.9. The used sample is again downstream SS semileptonic.

used to gain insight on the usefulness of the category. The expected exclusion limits are set as a standard single sided 95 % confidence interval of a Poisson distribution. Considering the zero-background assumption this corresponds to a threshold of $N_{95\,\%} = 3.0$ expected events [44]. Therefore, for this estimate, we assume that the region of parameter space which shows more than three expected event is excluded at 95 % CL.

Figure 8.11 visualizes the results for all categories split between the same-sign muons categories on the top and the OS categories on the bottom. Every color type references a certain category, either one of the leptonic or the semileptonic ones. The same line style is always corresponding to one of the displacement categories. Furthermore, the contours of the HNL lifetime equivalence and the previous limits of other searches are shown in the background. As a note the limits stop at 1.6 GeV HNL mass as this is the lowest available mass point in the simulation. For the same reason it stops for the leptonic B^+ category at 2.0 GeV HNL mass. In addition the downstream limits are stopped at lifetimes before they become unreliable. Finally the combined limit of each of the main B categories is shown. The limits are combined by summing up the expected events of the combined categories and comparing it with the event threshold of $N_{95\%} = 3.0$.

As expected the long track categories produce the lowest absolute limit compared to the other displacement categories. Therefore all of the long track

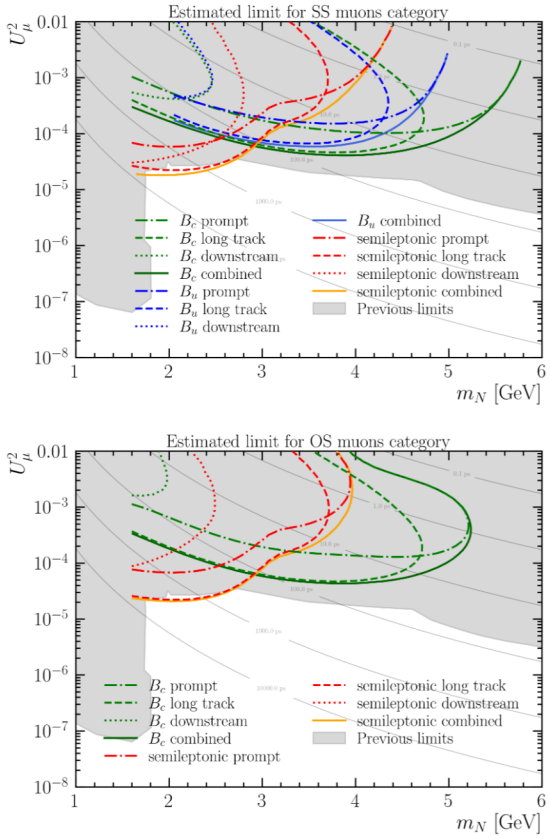


Figure 8.11: Zero-background exclusion limit estimation with assumed 80 % *MVA efficiency for all 15 categories. The upper plot depicts the same-sign muons categories while the lower plot shows the OS categories. For reference the contours of corresponding HNL lifetimes are also shown.*

categories are kept for further usage. The prompt categories locally outperform the others in case of high masses above 4.5 GeV. In those areas the limits are much worse than other previous experiments making them less of interest. Furthermore, studies showed that around 80 % of background events are inside the prompt categories. As such, the final limit on this category will be worse due to remaining background. The prompt categories are therefore dropped and instead the 20 mm displacement cut is optimized to enhance the performance of the displaced categories. Downstream performs worse than the rest as expected. The only relevant downstream category is the semileptonic SS subset which contributes to the combined limit for

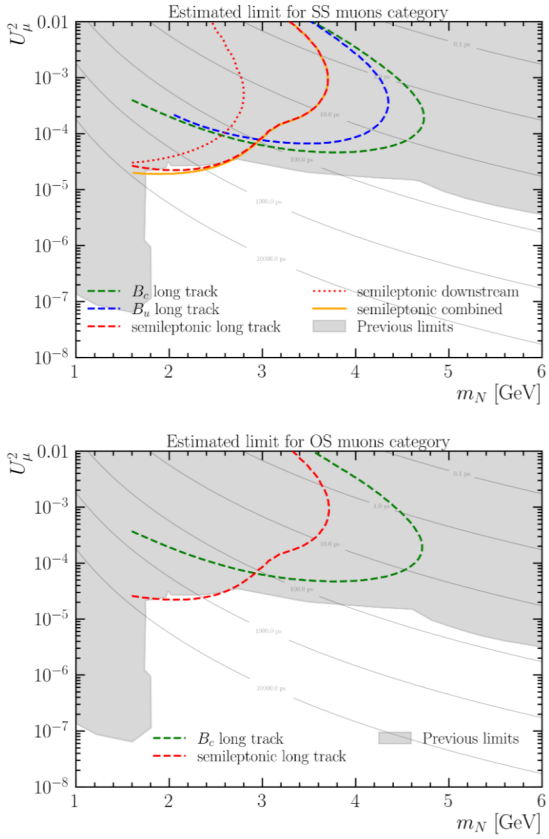


Figure 8.12: Zero-background limit estimation with assumed 80 % MVA efficiency for selected categories based on performance. The upper plot depicts the same-sign muon categories while the lower plot shows the OS categories. For reference the contours of corresponding HNL lifetimes are also shown.

masses below 2 GeV. As already mentioned one reason for this are the missing VELO information for downstream tracks reducing the reconstruction quality. In general it is seen that the semileptonic category in the area around 2 GeV is the best performing with a chance to set a new world-best limit. Based on those observations six categories are selected to be kept while the others are not used for the further analysis. These categories are all the long track categories as well as the semileptonic SS downstream tracks as shown in figure 8.12.

Projecting the sensitivity reveals the importance of optimizing the semileptonic long track SS category especially around the 2 GeV mass point. With

this knowledge the fiducial volume is optimized as described in the earlier chapter table 7.2. The mixing angle limits produced with an assumed MVA efficiency of 80 % demonstrate the importance of an efficient neural network to preserve the chance on a new world-best limit.

8.5 Model independent limit

The setting of the previous limits has an intrinsic dependency on the model of the HNL. This is introduced by the calculation of HNL production and decay branching ratios based on a theoretical framework. In this section an model independent limit is set. It is not an upper limit on the muon neutrino mixing angle U_μ^2. Instead it is an upper limit on the branching ratio of decay channels containing an HNL. For the limit the independent expected number of events N'_{decays} are calculated according to equation 8.9. This is identical to the calculation via equation 8.1 except without the model dependent branching ratios. This new number of expected events is equivalent to assuming a 100 % HNL decay channel branching ratio. The total fragmentation fraction is estimated as the sum of those for B^0, B^+ and B_s^0: $f_{semi}/f_u = f_u/f_u + f_d/f_u + f_s/f_u = 2 + f_s/f_u$ where isospin symmetry $f_u = f_d$ is used and the values from table 7.1 are used. The $B_q \to X_{Prim}\mu N$ branching ratios are assumed to be the same for the three mesons.

$$N'_{decays} = B_{prod} \cdot \epsilon_{tot} \tag{8.9a}$$

$$B_{prod} = L \cdot \frac{\sigma_{pp\to B^\pm X}}{\epsilon_{pp\to B^\pm X}} \cdot \frac{f_q}{f_u} \tag{8.9b}$$

$$\epsilon_{tot} = \epsilon_{FV} \cdot \epsilon_{rec}(\cdot\epsilon_{MVA}) \tag{8.9c}$$

Just as for the previous limit a single-sided 95 % confidence level interval of a Poissonian, corresponding to $N_{95\%} = 3.0$ events, is used. Dividing these three events by the number of expected events $3/N'_{decays}$ is equivalent to setting upper limits on the HNL decay channel branching ratios.
Figure 8.13 shows the model independent limits for the chosen six categories for different HNL lifetimes over the full mass range. In the semileptonic case the long and downstream track categories are combined to a displaced category by adding up the expected events. For this reason a second plot

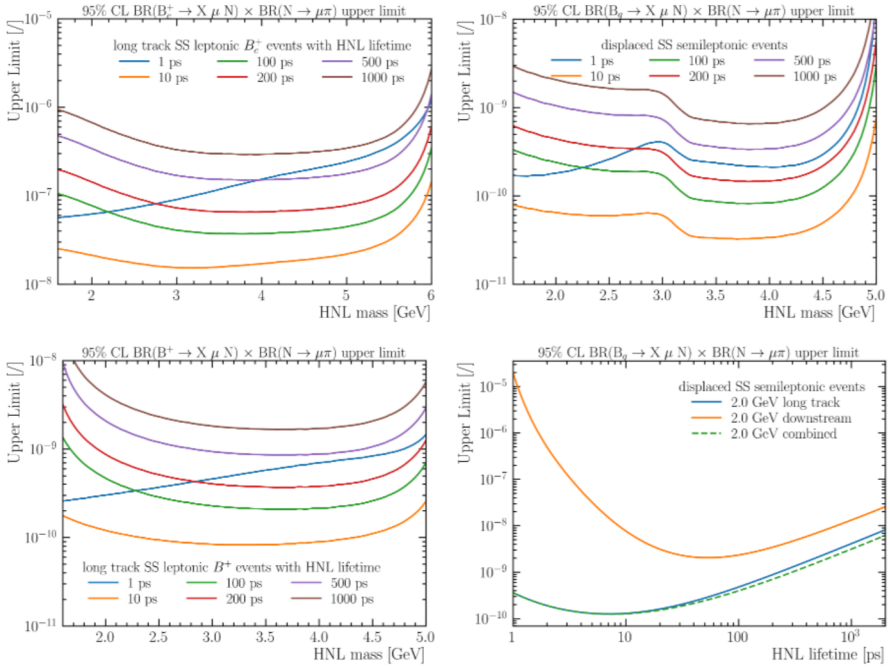


Figure 8.13: Model independent limits for the selected categories. Displaced on the right-hand side refers to the combination of downstream and long tracks. For the semileptonic case on the right side the limit with respect to the HNL lifetime is also shown. B_q in the semileptonic case refers to all B mesons in the semileptonic sample: B^+, B^0 and B_s^0

shows the contribution of the two displacement categories at the previously seen most sensitive region of 2 GeV HNL mass.

Investigating the independent limit shows expected results. Taking the reconstruction efficiency maps into account about 10 ps is the most sensitive region for the long track category. Going below or above that soon worsens the sensitivity as shown by the 1 ps and higher lifetime lines. Another trend that can be seen is that for lifetimes above optimum the shape of the expected limits stays the same. Only a constant offset is introduced. This is not the case for lower lifetimes which is also shown in the reconstruction efficiency maps. In the semileptonic case a clear sensitivity step is seen at 3 GeV coming from the step seen in the fiducial volume efficiency when channels containing pions and rho become dominant. Scanning over the lifetime further confirms again that the highest sensitivity for long track is slightly under 10 ps HNL lifetime while for downstream tracks it is at around 40 ps.

Downstream here especially contributes above this mark. Looking at the maximum sensitivity the B_c^+ meson category is the worst with a limit of about 2×10^{-8} up until an HNL mass of 5.4 GeV where it starts to rise strongly. It's followed by the B^+ category with a limit slightly below 10^{-10} in the HNL mass range $m(HNL) \in [2.4,\ 4.3]$ GeV. Best category stays the semileptonic one. Between 2 GeV and 3 GeV HNL mass the limit is 6×10^{-11}. Above 3 GeV up until 4.2 GeV HNL mass the best limit is about 4×10^{-11}.

8.6 Systematic uncertainties

It is once more important to emphasize that this analysis is still blinded. As such, the presented limits on the muon neutrino mixing angle are expected limits. Furthermore, the selection MVA has not been optimally developed. Additionally, the limit contains multiple sources of systematic uncertainties which are summarized in table 8.3. For the discussion of these uncertainties one has to once more take a look at the equation for expected HNL events:

$$N_{decays} = B_{prod} \cdot HNL_{prod/dec} \cdot \epsilon_{tot} \tag{8.10a}$$

$$B_{prod} = L \cdot \frac{\sigma_{pp \to B^{\pm} X}}{\epsilon_{pp \to B^{\pm} X}} \cdot \frac{f_q}{f_u} \tag{8.10b}$$

$$HNL_{prod/dec} = BR(B_q \to X \mu^+ N) \cdot BR(N \to \mu^+ \pi^-) \tag{8.10c}$$

$$\epsilon_{tot} = \epsilon_{FV} \cdot \epsilon_{rec} \cdot \epsilon_{MVA} \tag{8.10d}$$

Beginning with the B meson production term B_{prod} the integrated luminosity is given as $(5.1 \pm 0.3)\,\mathrm{fb}^{-1}$ [38]. This results in a relative systematic uncertainty of 5.9 % for all categories. The B meson production cross section $\sigma_{pp \to B^{\pm} X} = (86.6 \pm 6.4)\,\mu\mathrm{b}$ and its uncertainties is also constant for all categories [39]. The cross section contributes with a relative systematic uncertainty of 7.4 %. The fragmentation fraction f_q/f_u has different systematic uncertainties dependent on the mother B meson and therefore varies for the different analysis categories. In case of the leptonic B^+ category there is no systematic uncertainty because the fragmentation fraction is normalized to the B^+ and B^0 mesons. The leptonic B_c^+ category contains decays of B_c^+ mesons. Its relative systematic uncertainty is 23.5 % coming from the

corresponding fragmentation fraction $f_c/f_u = (2.61 \pm 0.62) \times 10^{-3}$. For the semileptonic category one has to determine the contribution of the the decay channels with different mother B mesons. The branching ratios for the different HNL production channels in the MC sample are calculated. This is done as follows. The full HNL mass range is scanned and for every step the mixing angle is evaluated that corresponds to the needed three expected events. If this is the case for more than one mixing angle the one producing the strongest limit is taken. For the resulting parameter pairs (mass, mixing angle), the ratio between the sum of decay branching ratios from a certain B meson and the branching ratios of all B mesons are calculated. Taking into account these contribution fractions the total fragmentation fraction and the corresponding systematic uncertainty is determined. In case of HNL masses below $m(HNL) = m(B_s^0) - m(D_s^+)$ the B_s^0 meson contribution is significant resulting in a systematic uncertainty of 0.5 %. Above this threshold the phase space of this D meson containing channel closes. The B_s^0 meson contribution vanishes and the fragmentation fractions of the B^+ and B^0 meson lower this systematic uncertainty to zero.

The efficiency of produced B mesons being in the LHCb acceptance is denoted by $\epsilon_{pp\to B^\pm X}$. The efficiency value is derived from *RapidSim* simulations. As this fast simulation generator can produce millions of events in short time the statistical uncertainty of the efficiency goes to zero. A systematic uncertainty is introduced by the model that RapidSim uses for its simulation. At this point the systematic uncertainty introduced by RapidSim is yet to be investigated. One can argue that the source of systematic uncertainty is the same for the LHCb acceptance and FV efficiency which means the ratio of them should cancel most bias.

Production and decay branching ratios of the different decay channels are calculated with theory predictions. No systematic uncertainties are assigned to these numbers.

The total efficiency ϵ_{tot} consists of the efficiencies of the fiducial volume and the MVA as well as the reconstruction efficiency. Since the MVA has not been completed at this point, no systematic uncertainty can yet be assigned. The systematic uncertainty for the reconstruction efficiency is discussed at the beginning of this section. As a reminder the reconstruction efficiency ϵ_{rec} is defined as the ratio of detector level events to generator level events. The effective number of events for both nominator and denominator are

the sum of weights which are applied. Weights are bound to decrease the statistical strength of a sample by giving more importance to certain events and disregarding others. The uncertainty on this efficiency is derived from the sum of squared weights. Therefore, the uncertainty of the reconstruction efficiency is equivalent to the systematic uncertainty introduced by Monte Carlo statistics. To estimate the effect on the derived limits, the HNL mass is scanned. For every mass step the reconstruction efficiency corresponding to the strongest limit is extracted as already described. The systematic uncertainty is then given by a range defined by the lowest and the highest reconstruction efficiency uncertainty. For this a wide range is seen for the long track SS B^+ category. A value of about 10 % occurs at 2.5 GeV HNL mass and then falls of again for higher masses. The mentioned trend can be seen in figure 8.14. The figure shows for all categories the distribution of their systematic uncertainties with respect to the HNL mass. The mentioned trend is explained by the used MC data. For the leptonic B^+ category only MC data with a lifetime of 10 ps is available. As the exclusion limits for low HNL masses are at high lifetimes the uncertainty is higher. Currently additional MC data with an HNL lifetime of 100 ps and 1000 ps for this category is requested. The uncertainty will therefore decrease in the future. Another trend seen is that for the semileptonic categories the MC statistics uncertainty rises for HNL masses higher than 3.2 GeV. This is the point at which the channels including a D meson stops to dominate the semileptonic category as previously seen in figure 5.3 and 5.4.

Apart from the ingredients which enter the prediction of the number of expected events systematic uncertainties can be introduced by the uncertainty on weights themselves. In this analysis three different types of weights are used: The lifetime weights that weights the MC data to different HNL lifetimes, the BR correction weights correcting the branching ratios given in the initial decfiles and the cocktail weights that scale the contribution of the different MC samples correctly. The lifetime weights and BR correction weights are calculated from fully known parameters such as the HNL lifetime. They therefore have no uncertainty themselves. Looking at the cocktail weight from equation 7.7 though shows that it's also based on the fragmentation fraction. This is not relevant for the leptonic categories, though. As these categories only have one decay channel their cocktail weight is one. The semileptonic channel includes decays of B^0, B^+ and B_s^0 mesons. The

former two mesons have a fragmentation fraction of one so only B_s^0 channels contribute to a systematic uncertainty. With a fragmentation fraction $f_s/f_u = 0.244 \pm 0.012$ and knowing from before that the semileptonic fragmentation fraction has at most 0.5 % uncertainty one can argue that the introduced systematic uncertainty is of second order and neglect it.
The last systematic uncertainty that can be investigated is the mass interpolation model. For the interpolation a cubic spline is used which needs four mass points as input. Another possible interpolation shape would've been a quadratic spline. To estimate the systematic uncertainty for the usage of the cubic spline the reconstruction efficiency maps of subsection 8.3 are used. They are produced a second time but with the quadratic spline instead. The systematic uncertainty that is given per category is the lowest and highest difference in reconstruction efficiency again along the lifetime-mass pairs that correspond the the strongest limit. Generally the interpolation performs better at areas with less difference between the efficiencies. This means that for high masses where the efficiency often decreases rapidly due to closing phase-spaces the interpolation behaves most differently for two alternative fit models. This can be seen as well in figure 8.14 especially for the semileptonic categories. The reason is again the closing D meson phasespace for HNL masses larger than 3.2 GeV.
The discussed systematic uncertainties can be seen in table 8.3. For every category they are combined into a total systematic uncertainty. As they all contribute directly to the number of expected events and hence the limit they are combined by squared summing. For the given range of total systematic uncertainty the lowest and highest of each range are not simply combined. Instead the uncertainties corresponding to the same mass step inside a category are combined. As such, the highest total systematic uncertainties don't correspond to the squared sum of the highest individual uncertainties. The same applies for the lowest values.

Source	Contribution [%]					
	long track SS			long track OS		downstream
	semi	B^+	B_c^+	semi	B_c^+	SS semi
Integrated luminosity	5.9					
B meson cross section	7.4					
Fragmentation fraction	None - 0.5	None	23.5	None - 0.5	23.5	0.4 - 0.5
LHCb acceptance efficiency	negligible					
Monte Carlo statistics	1.6 - 4.0	0.9 - 10.5	0.4 - 2.0	1.6 - 3.2	1.4 - 3.3	1.3 - 3.6
Mass interpolation	< 0.1 - 16.1	< 0.1 - 0.2	< 0.1 - 0.5	0.1 - 6.0	< 0.1 - 0.5	< 0.1 - 7.6
FV efficiency	negligible					
MVA efficiency	not yet available					
Total	9.6 - 19.1	9.5 - 14.1	25.4 - 25.4	9.6 - 11.6	25.4 - 25.6	9.5 - 12.7

Table 8.3: Overview over different systematic uncertainties.

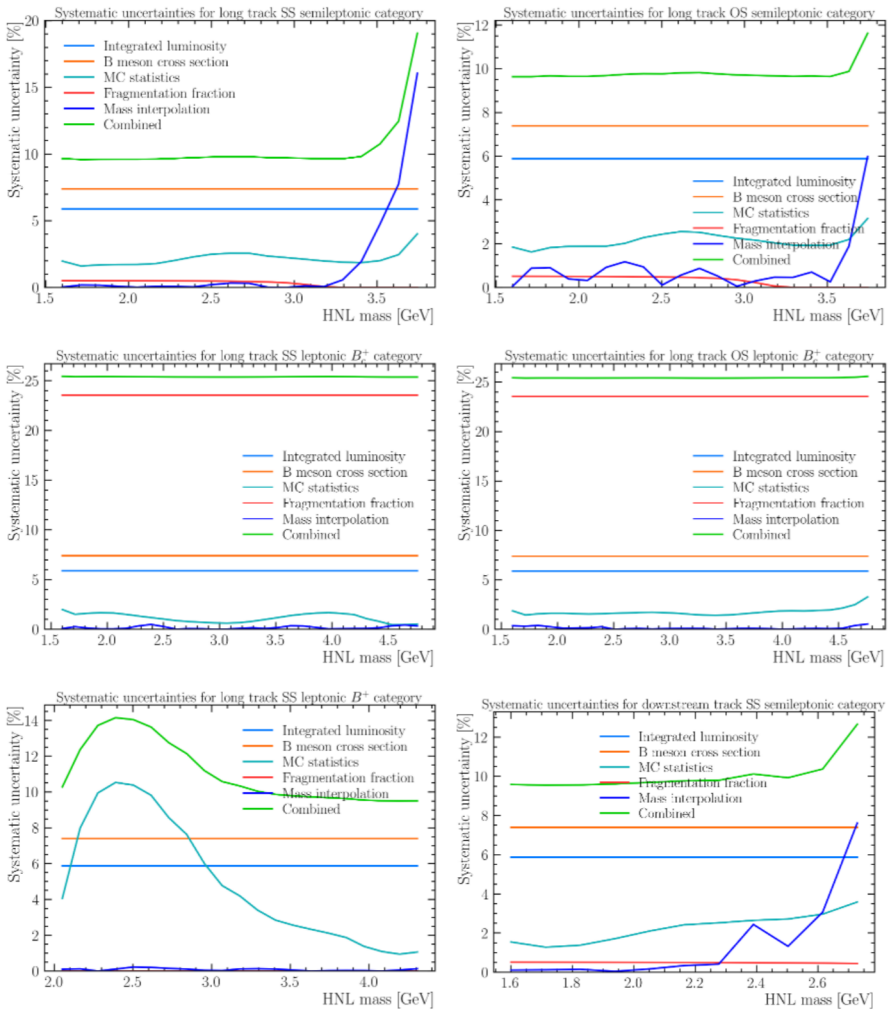


Figure 8.14: Systematic uncertainties for the selected data categories. The shown mass range is where a limit is found for the investigated mixing angles.

9 Outlook

This section gives an overview over tasks that still remain to be done. Included is the multivariate analysis that is planned to be used for the final selection.

9.1 Limit setting with preliminary MVA

Currently (status December 2021) the multivariate analysis needed for the complete limits of the muon neutrino mixing angle U_μ^2 is still being optimized. This work is done by Serhii Cholak of the *École Polytechnique Fédérale de Lausanne* (EPFL). It follows a short overview about the preliminary MVA and a sensitivity projection with it.
The MVA classifier is trained to suppress combinatorial background while having a high signal efficiency. The training is performed separately for the six selected event categories. To train the MVA a sample for the signal and one for the background is chosen. For the signal sample the full Monte Carlo data set with all available mass points is taken. Furthermore, the lifetime of the sample is reweighted in such a way that it populates the needed category. To train the background discrimination real data is used. One can be sure that the sideband above the B meson mass corresponding to the category doesn't contain signal events. Therefore it is used as the background training sample. Standard neural networks are not optimal when they are trained to classify a particle with a wide range of possible masses. One would have to train a new network for every investigated mass and even then they are not able to interpolate between the mass points. For this reason an approach is used where the true mass of the particle is used as a training feature [45]. Thus, the classifier response is parameterized with the true HNL mass which is given in the MC data. In case of the background sample the true HNL mass is no parameter of importance. As such for this background sample the HNL mass parameter is randomly sampled from the signal distribution. With this approach the performance at intermediate masses is greatly increased.
The used parameters for the training are the same for all categories:

- **Direction angle:** $log_{10}(1 - DIRA(N))$
- **Vertex χ^2:** $log_{10}(\chi^2_{vtx}(N)/dof),\ log_{10}(\chi^2_{vtx}(B_q)/dof)$

- **Impact parameter χ^2:** $ln\ \chi^2_{IP}(\pi),\ ln\ \chi^2_{IP}(\mu_{Sec}),\ ln\ \chi^2_{IP}(N)$
- **Transverse momentum:** $ln\ p_T(\pi),\ ln\ p_T(\mu_{Sec}),\ ln\ p_T(\mu_{Prim})$
- **Displacement:** $ln\ FD(B_q),\ ln\ \tau(B_q)$
- **Only leptonic categories:** $log_{10}(1 - DIRA(B_q))$
- **Parameterization:** m_N^{True}

The training parameters are standard choices. The vertex parameter, χ^2_{vtx}, checks if a good vertex can be formed. Transverse momentum, p_T, of the decay products is well measured therefore a good training parameter. Flight distance, FD, and lifetime, τ, of the initial B meson performs especially well for the displaced categories. For the leptonic categories the direction angle is also used as in these categories the B momentum is reconstructed.
For the resulting MVA response a cut value has to be chosen. This is done with a Punzi figure of merit (f.o.m.) as shown in equation 9.1 [46].

$$f.o.m. = \frac{\epsilon}{\sqrt{B} + 3/2} \tag{9.1}$$

In this Punzi figure of merit ϵ is the signal efficiency taken from the MC data. The background yield B is taken from the $\pm 2\ \sigma_m$ area around the peak in the $\mu\pi$ mass distribution. From the maximum of the f.o.m. the MVA decision cut is chosen for each category. Figure 9.1 shows the signal efficiency and background rejection efficiency for the first three mass points in the semileptonic SS long track category. At the most sensitive mass point of an HNL mass of 2 GeV the signal efficiency is about 55 %. The background rejection efficiency is above 99.8 %. This validates the zero background approach for the muon neutrino mixing angle U^2_μ limit setting. It is very likely that the beforehand assumed 80 % MVA signal efficiency can be reached with further optimization.
Applying the MVA decision cut to the data allows to perform the same limit setting as described in subsection 8.4. Figure 9.2 shows the estimated exclusion limits on the muon neutrino mixing angle for the categories selected in subsection 8.4. One can see at the combined orange limit that the same-sign category already performs well in the low mass areas. In the area $m(HNL) \in [1.6, 2.3]$ GeV the strongest limit is about 2×10^{-5} rivaling the current world best limits of other experiments.

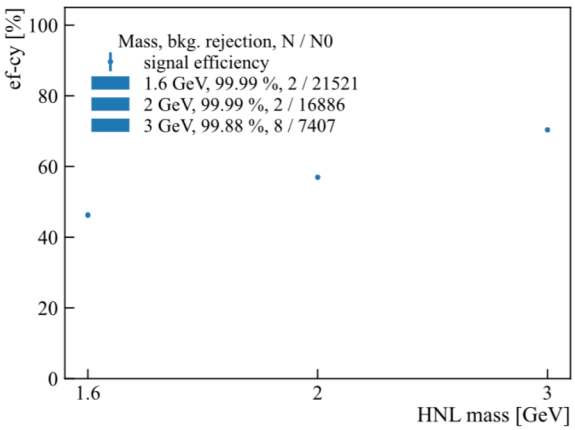


Figure 9.1: Plot by Serhii Cholak: Signal efficiency of the MVA for the semileptonic SS long track category. For the given mass points the background rejection is given in the legend.

9.2 Remaining tasks

This section discusses the remaining main tasks that need to be done for this analysis. As already mentioned the MVA is being worked on. The goal is to optimize it so that the signal efficiency is higher than the initial 50 % - 70 % efficiency. Furthermore, peaking background coming from misidentified D^0 and J/Ψ mesons are worked on. In addition, combinatorial background is to be investigated. Studies are done as to see if it comes from real or misidentified muons. After these studies the peak search in the $m(\mu\pi)$ mass distribution has to be updated. The current planned approach is doing a fit with three components: The signal, the combinatorial background and the peaking background from misidentified D^0 mesons. For the D^0 component in the fit the so-called ABCD method is employed [47]. Data from around the current optimal MVA response cut of 0.4 is taken. This corresponds to all Run 2 data within the range of 0.4 ± 0.2 MVA response cut. The combinatorial and the D^0 background are fit in four regions. The first two are the prompt category below 3.4 GeV of the $m(\mu\mu\pi)$ mass spectrum where the X_{Prim} particle in the signal channels can't be a D^0 and above this threshold. The other two regions are with an HNL vertex displacement of above the usual 20 mm for the same two mass regions. Using this regions the D^0 background shape and normalization is extracted and later used in the peak search fit. If at a later point this is done it is planned to also go

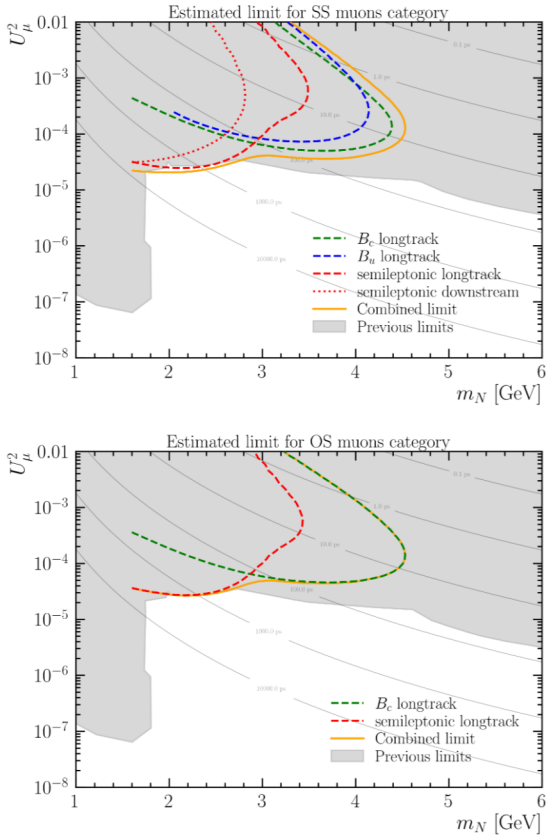


Figure 9.2: Estimated limit of the muon neutrino mixing angle U_μ^2 with preliminary MVA. The orange line is the combined limit.

back and optimize the 20 mm displacement cut.
Another task is further corrections to the Monte Carlo data. This includes the correction of the nSPD variable and the *particle identification* (PID) variables. The nSPD is used in the trigger stage of data. As mentioned in section 4 the variable nSPD is the number of hits in the *Scintillating Pad Detector* (SPD), which is part of the *electromagnetic calorimeter* (ECAL). This variable directly depends on the event multiplicity which is hard to simulate. Therefore it can make a difference in the number of events accepted by the trigger when this is corrected. The PID variables share a similar problem. The particle identification depends on the response of many different parts of the LHCb detector like the RICH detectors and the muon stations. For this reason the PIDCalib software and its successor PIDCalib2 are developed [48]. The packages use internal real data samples to correct the distributions of MC samples. It has to be discussed how this is applied to the used MC data set as it consists out of many different data samples.

10 Conclusion

A search of heavy neutral leptons from B meson decays is performed at the LHCb experiment. The used decay channels are of inclusive nature and described by $B_q \to X_{Prim} N(\to \pi\mu)\mu$. For most of the parameter space investigated, the HNL decays with a displaced vertex producing a pion and a muon. The full final state consists of a pion, two muons and a not reconstructed particle X_{Prim}. The used data set corresponds to pp collisions recorded during LHC Run 2 in the years 2016 to 2018. This corresponds to an integrated luminosity of $1.6\,\text{fb}^{-1}$, $1.7\,\text{fb}^{-1}$ and $2.1\,\text{fb}^{-1}$ for 2016, 2017 and 2018, respectively. The corresponding Monte Carlo data includes samples for different potential HNL masses ranging from 1.6 GeV to 6 GeV as well as HNL lifetimes of 10 ps, 100 ps and 1000 ps. The samples are combined by scaling their contribution correctly. Expected limits on the muon neutrino mixing angle U_μ^2 are calculated in the mass range $m(HNL) \in [1.6, 6]$ GeV and the HNL lifetime range $\tau(HNL) \in [0.1, 10^4]$ ps. The presented analysis is performed in categories dependent on the following characteristics:

- **Muon charge:** Do the primary and secondary muon have the same sign (SS category) or opposing signs (OS category)?
- **Leptonic or semileptonic:** Does the HNL production channel have a not reconstructed X_{Prim} hadron (semileptonic category)? If not, is the mother particle a B^+ or B_c^+ meson (leptonic B^+ or B_c^+ category, respectively)?
- **Vertex displacement:** Is the HNL vertex displaced by less than 20 mm (prompt category), more than 20 mm and the HNL decay tracks reconstructed as long tracks (long track category) or more than 20 mm and the decay products are reconstructed as downstream tracks (downstream category)?

To set expected limits on the muon neutrino mixing angle, U_μ^2, as a function of HNL mass, the efficiencies for event reconstruction, the applied fiducial volume and the applied MVA have to be known over the full investigated range. Efficiencies are modelled using MC over the full range of lifetimes and masses. HNL lifetimes different from those generated in the MC are obtained by reweighting. The weighting procedure is optimized to reduce the potential bias from reweighting from lower to higher lifetimes.

Efficiencies for mass hypotheses that are not generated are obtained by interpolation using cubic spline functions. They have the advantage of higher stability than polynomials. To propagate uncertainty introduced by weights toys have been evaluated. Furthermore, several sources of systematic uncertainties are considered and estimated.
Taking into account the integrated luminosity, the B meson productions cross sections, branching ratios from theoretical calculations and the total selection efficiency, the yields of expected signal events is calculated for a large range of neutrino mixing parameters U_μ^2 and HNL masses. Assuming the precise signal selection will reduce the background to a negligible level, 95 % CL expected limits are calculated and compared to those set by previous experiments.
These estimates are used to simplify the analysis strategy by dropping several signal categories that turn out to contribute with negligible sensitivity. The prompt categories perform poorly as their highest efficiencies are at low HNL lifetimes that are already excluded by previous experiments. They furthermore would contain most background in case of real data. Bad performance of the downstream track categories is caused by worse reconstruction efficiency based on missing VELO information and no dedicated trigger line. The combined expected limit of the remaining six categories are shown in figure 10.1. One can see that the long track SS semileptonic events category has the best performance. Its lowest limit is in the range of HNL masses of 1.6 GeV to 2.4 GeV. The limits are compared with the previous limits of other experiments in figure 10.2. During the work on this analysis a new result by the CMS experiment was published. This result over-performs the expected limit of this analysis for HNL masses larger than 2 GeV while this analysis performs better at HNL masses below 2 GeV.
Next, the analysis has to be performed on real data. There is a good chance to improve the world best limit at the 2 GeV HNL mass. Furthermore, this analysis only used the exclusive decay of HNL to a muon and a pion. Monte Carlo data for a possible future analysis with the inclusive decay $N \rightarrow X_{Sec}\ \mu$ is available. Adding these decays as well as dedicated trigger lines for LHC Run 3 will improve the sensitivity of this analysis further.

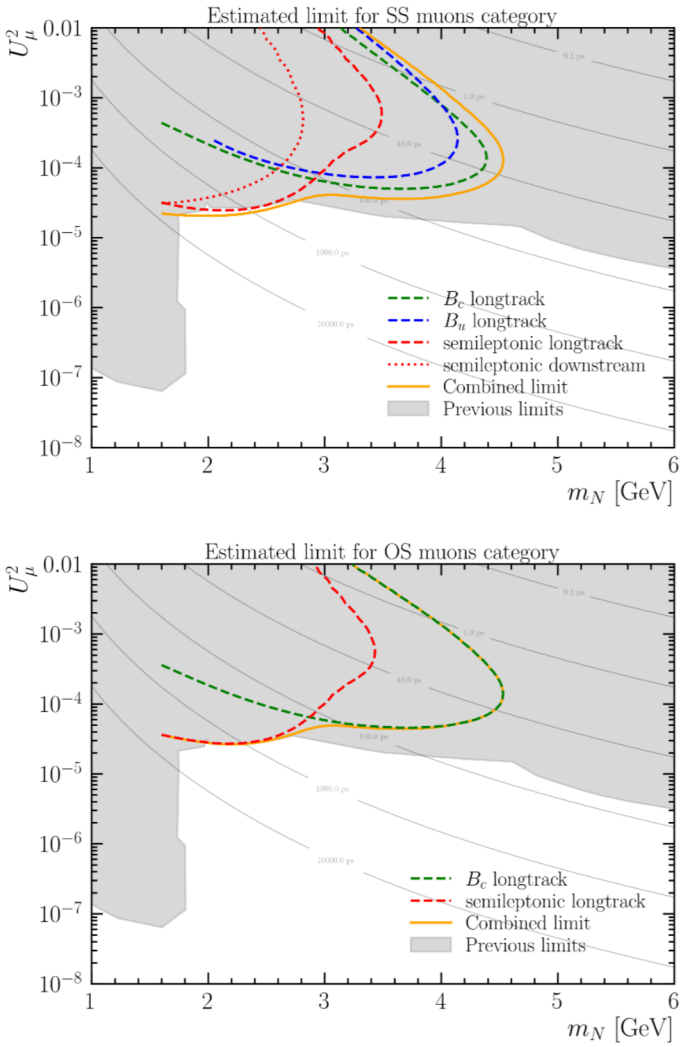


Figure 10.1: Combined expected exclusion region for the selected data categories in the parameter space of mixing angle U_μ^2 and HNL mass. The limit assumes the search will have a zero background. Excluded regions correspond to a 95 % upper limit on the signal yield. The upper figure shows the same-sign muons categories while to lower are the opposing-sign muons categories.

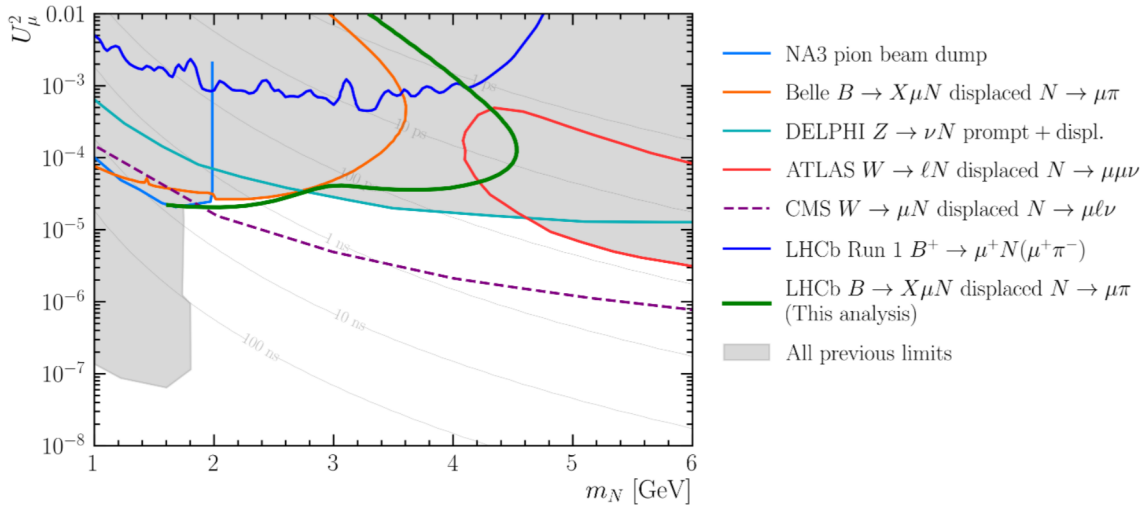


Figure 10.2: Comparison of the expected combined limit of the selected SS categories with the results of other HNL searches. The limits of the other experiments are also for the same-sign muons case. Depicted are the results of NA3 [29], Bele [30], DELPHI [31], ATLAS [32], LHCb Run 1 [33] and CMS [49].

Appendix A: Results of lifetime reweighting checks

Category: prompt			Reweighted HNL lifetime τ		
HNL mass	μ sign	Part	10 ps	100 ps	1000 ps
1.6 GeV	SS	BcLept	—	0.994 ± 0.068	1.066 ± 0.185
		semilept	—	1.017 ± 0.062	0.862 ± 0.116
	OS	BcLept	—	0.985 ± 0.072	1.183 ± 0.228
		semilept	—	0.995 ± 0.068	1.056 ± 0.177
2.0 GeV	SS	BcLept	1.000 ± 0.014	0.997 ± 0.045	1.025 ± 0.144
		BuLept	1.000 ± 0.026	—	—
		semilept	—	0.994 ± 0.051	1.068 ± 0.134
	OS	BcLept	—	0.980 ± 0.070	1.256 ± 0.241
		semilept	—	1.001 ± 0.060	0.987 ± 0.144
3.0 GeV	SS	BcLept	1.000 ± 0.009	1.020 ± 0.019	0.862 ± 0.114
		BuLept	1.000 ± 0.021	—	—
		semilept	—	0.995 ± 0.080	1.050 ± 0.160
	OS	BcLept	—	0.993 ± 0.071	1.073 ± 0.186
		semilept	—	0.989 ± 0.099	1.127 ± 0.259
4.0 GeV	SS	BcLept	1.000 ± 0.008	0.963 ± 0.045	1.009 ± 0.152
		BuLept	1.000 ± 0.021	—	—
		semilept	—	1.011 ± 0.059	0.907 ± 0.121
	OS	BcLept	—	0.991 ± 0.081	1.089 ± 0.218
		semilept	—	0.994 ± 0.068	1.066 ± 0.174
4.5 GeV	SS	BuLept	1.000 ± 0.026	—	—
5.0 GeV	SS	BcLept	1.000 ± 0.009	0.940 ± 0.048	0.942 ± 0.151
		BuLept	1.000 ± 0.070	—	—
		semilept	—	0.989 ± 0.353	1.120 ± 1.154
	OS	BcLept	—	1.013 ± 0.080	0.887 ± 0.156
		semilept	—	1.046 ± 0.407	0.693 ± 0.531
5.5 GeV	SS	BcLept	1.000 ± 0.010	1.006 ± 0.021	0.921 ± 0.177
	OS	BcLept	—	1.014 ± 0.106	0.882 ± 0.203
6.0 GeV	SS	BcLept	1.000 ± 0.025	—	—

*Table 10.1: Ratio between **prompt** MC data of a certain HNL lifetime and the full category MC data reweighted to said lifetime. No ratio means there is no MC sample for this specific configuration. Categories which have no data for a certain mass are omitted.*

Category: long track			Reweighted HNL lifetime τ		
HNL mass	μ sign	Part	10 ps	100 ps	1000 ps
1.6 GeV	SS	BcLept	—	1.003 ± 0.028	0.960 ± 0.061
		semilept	—	1.003 ± 0.026	0.971 ± 0.055
	OS	BcLept	—	1.004 ± 0.026	0.963 ± 0.056
		semilept	—	0.990 ± 0.024	1.088 ± 0.063
2.0 GeV	SS	BcLept	1.000 ± 0.009	0.996 ± 0.023	1.026 ± 0.055
		BuLept	1.000 ± 0.019	—	—
		semilept	—	1.008 ± 0.022	0.940 ± 0.044
	OS	BcLept	—	1.006 ± 0.023	0.955 ± 0.048
		semilept	—	1.001 ± 0.022	0.993 ± 0.049
3.0 GeV	SS	BcLept	1.000 ± 0.006	1.000 ± 0.009	1.028 ± 0.050
		BuLept	1.000 ± 0.015	—	—
		semilept	—	0.991 ± 0.032	1.068 ± 0.075
	OS	BcLept	—	0.998 ± 0.024	1.012 ± 0.053
		semilept	—	1.005 ± 0.032	0.965 ± 0.070
4.0 GeV	SS	BcLept	1.000 ± 0.006	0.995 ± 0.025	1.024 ± 0.058
		BuLept	1.000 ± 0.016	—	—
		semilept	—	1.003 ± 0.029	0.953 ± 0.082
	OS	BcLept	—	0.999 ± 0.026	1.011 ± 0.055
		semilept	—	1.006 ± 0.027	0.970 ± 0.071
4.5 GeV	SS	BuLept	1.000 ± 0.020	—	—
5.0 GeV	SS	BcLept	1.000 ± 0.007	0.991 ± 0.027	1.074 ± 0.067
		BuLept	1.000 ± 0.046	—	—
		semilept	—	1.053 ± 0.154	0.667 ± 0.175
	OS	BcLept	—	0.993 ± 0.028	1.030 ± 0.064
		semilept	—	0.966 ± 0.122	1.490 ± 0.527
5.5 GeV	SS	BcLept	1.000 ± 0.008	0.999 ± 0.011	1.144 ± 0.084
	OS	BcLept	—	1.006 ± 0.036	0.965 ± 0.071
6.0 GeV	SS	BcLept	1.000 ± 0.017	—	—

*Table 10.2: Ratio between **long track** MC data of a certain HNL lifetime and the full category MC data reweighted to said lifetime. No ratio means there is no MC sample for this specific configuration. Categories which have no data for a certain mass are omitted.*

Category: downstream track			Reweighted HNL lifetime τ		
HNL mass	μ sign	Part	10 ps	100 ps	1000 ps
1.6 GeV	SS	BcLept	—	1.004 ± 0.020	0.985 ± 0.038
		semilept	—	1.005 ± 0.018	0.972 ± 0.031
	OS	BcLept	—	1.001 ± 0.105	0.955 ± 0.190
		semilept	—	0.965 ± 0.093	1.300 ± 0.283
2.0 GeV	SS	BcLept	1.000 ± 0.031	1.008 ± 0.023	0.968 ± 0.041
		BuLept	1.000 ± 0.059	—	—
		semilept	—	0.997 ± 0.021	1.011 ± 0.040
	OS	BcLept	—	1.000 ± 0.126	1.005 ± 0.242
		semilept	—	0.978 ± 0.154	1.142 ± 0.407
3.0 GeV	SS	BcLept	1.000 ± 0.041	0.999 ± 0.016	1.090 ± 0.081
		BuLept	1.000 ± 0.083	—	—
		semilept	—	0.993 ± 0.064	1.028 ± 0.123
	OS	BcLept	—	0.880 ± 0.249	1.765 ± 1.297
		semilept	—	0.951 ± 0.208	1.578 ± 0.831
4.0 GeV	SS	BcLept	1.000 ± 0.061	0.997 ± 0.066	1.003 ± 0.122
		BuLept	1.000 ± 0.109	—	—
		semilept	—	0.997 ± 0.074	1.016 ± 0.104
	OS	BcLept	—	1.521 ± 0.773	0.362 ± 0.196
		semilept	—	1.000 ± 0.472	0.000 ± 0.000
4.5 GeV	SS	BuLept	1.000 ± 0.140	—	—
5.0 GeV	SS	BcLept	1.000 ± 0.087	0.982 ± 0.093	1.008 ± 0.184
		BuLept	1.000 ± 0.447	—	—
		semilept	—	1.000 ± 1.001	0.000 ± 0.000
	OS	BcLept	—	0.000 ± 0.000	0.000 ± 0.000
		semilept	—	0.000 ± 0.000	0.000 ± 0.000
5.5 GeV	SS	BcLept	1.000 ± 0.095	0.985 ± 0.046	0.843 ± 0.206
	OS	BcLept	—	0.000 ± 0.000	0.000 ± 0.000
6.0 GeV	SS	BcLept	1.000 ± 0.324	—	—

*Table 10.3: Ratio between **downstream track** MC data of a certain HNL lifetime and the full category MC data reweighted to said lifetime. No ratio means there is no MC sample for this specific configuration. Categories which have no data for a certain mass are omitted.*

Appendix B: Reconstruction efficiency maps

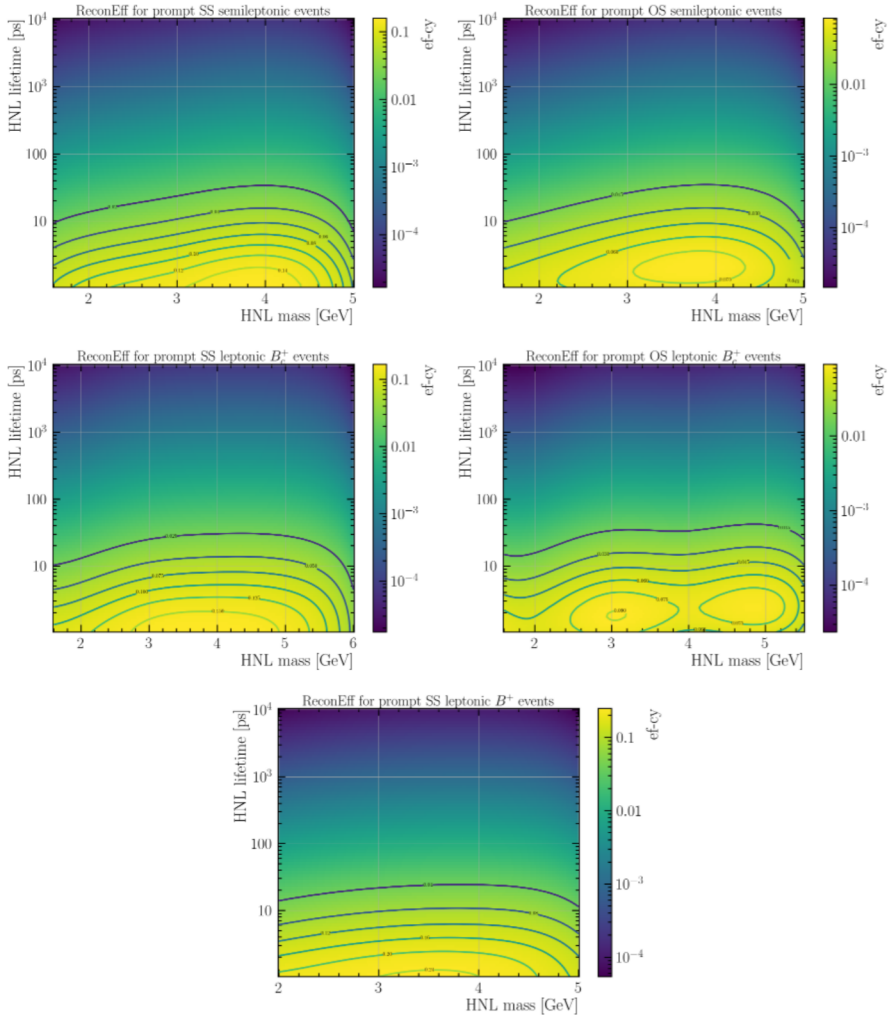


Figure 10.3: Reconstruction efficiency maps for the long track categories.

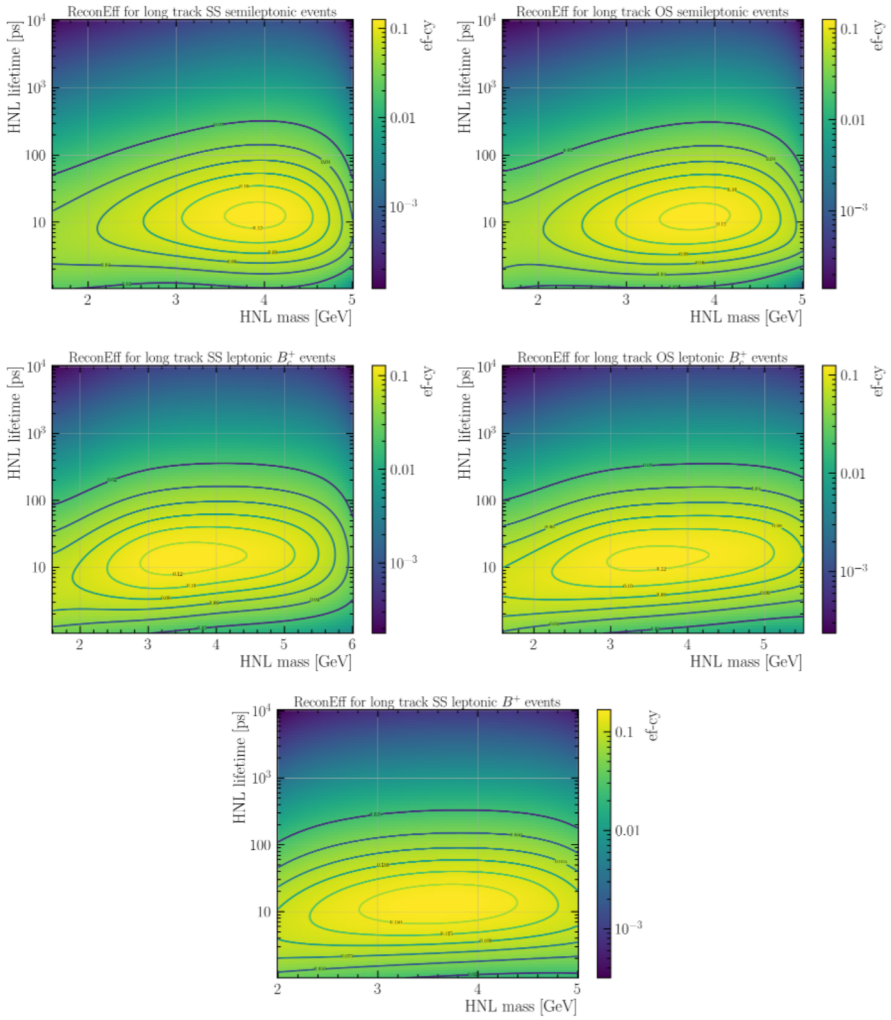


Figure 10.4: Reconstruction efficiency maps for the long track categories.

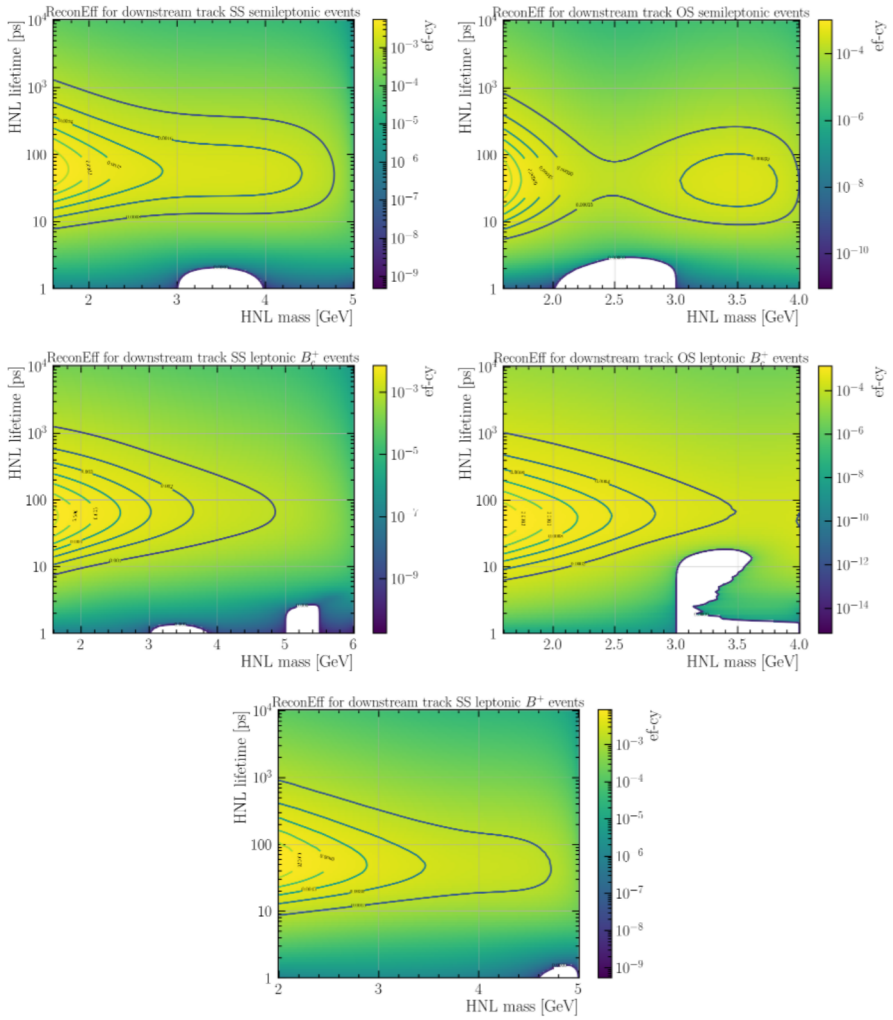


Figure 10.5: Reconstruction efficiency maps for the long track categories.

References

[1] A. D. Sakharov, "*Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*," *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32–35.

[2] F. Zwicky, "*On the Masses of Nebulae and of Clusters of Nebulae*," *Astrophys. J.* **86** (1937) 217–246.

[3] V. C. Rubin and W. K. Ford, Jr., "*Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*," *Astrophys. J.* **159** (1970) 379–403.

[4] B. T. Cleveland, T. Daily, J. Raymond Davis, J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, "*Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector*," *The Astrophysical Journal* **496** no. 1, (Mar, 1998) 505–526. https://doi.org/10.1086/305343.

[5] **SNO** Collaboration, Q. R. Ahmad *et al.*, "*Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory*," *Phys. Rev. Lett.* **87** (2001) 071301, arXiv:nucl-ex/0106015.

[6] S. N. Gninenko, D. S. Gorbunov, and M. E. Shaposhnikov, "*Search for GeV-scale sterile neutrinos responsible for active neutrino oscillations and baryon asymmetry of the Universe*," *Adv. High Energy Phys.* **2012** (2012) 718259, arXiv:1301.5516 [hep-ph].

[7] **Super-Kamiokande** Collaboration, Y. Fukuda *et al.*, "*Evidence for oscillation of atmospheric neutrinos*," *Phys. Rev. Lett.* **81** (1998) 1562–1567, arXiv:hep-ex/9807003.

[8] M. Aker *et al.*, "*First direct neutrino-mass measurement with sub-eV sensitivity*," arXiv:2105.08533 [hep-ex].

[9] P. Langacker, "*Models of neutrino mass*." https://www.physics.upenn.edu/~pgl/neutrino/now2000/node5.html. Accessed: 2021-11-08.

[10] "*Webpage: LHCb detector along the bending plane*." https://commons.wikimedia.org/wiki/File:Lhcbview.jpg. Accessed: 2019-01-11.

[11] R. Aaij *et al.*, "*Performance of the LHCb Vertex Locator*," *JINST* **9** (2014) P09007, arXiv:1405.7808 [physics.ins-det].

[12] **LHCb** Collaboration, R. Aaij *et al.*, "*LHCb Detector Performance*," *Int. J. Mod. Phys. A* **30** no. 07, (2015) 1530022, arXiv:1412.6352 [hep-ex].

[13] **LHCb** Collaboration, L. Collaboration, "*LHCb magnet: Technical design report*,".

[14] C. Abellán Beteta *et al.*, "*Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*," arXiv:2008.11556 [physics.ins-det].

[15] F. P. Albicocco *et al.*, "*Long-term Operation of the Multi-Wire-Proportional-Chambers of the LHCb Muon System*," *JINST* **14** (2019) P11031, arXiv:1908.02178 [physics.ins-det].

[16] **LHCb** Collaboration, R. Aaij *et al.*, "*Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*," *JINST* **14** no. 04, (2019) P04013, arXiv:1812.10790 [hep-ex].

[17] R. Aaij *et al.*, "*Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*," arXiv:1812.10790 [hep-ex].

[18] "*The LHCb data flow — LHCb Starterkit Lessons documentation.*" https://lhcb.github.io/starterkit-lessons/first-analysis-steps/dataflow.html. Accessed: 2021-12-04.

[19] T. Sjostrand, S. Mrenna, and P. Z. Skands, "*A Brief Introduction to PYTHIA 8.1*" *Comput. Phys. Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[20] D. J. Lange, "*The EvtGen particle decay simulation package*," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462** no. 1, (2001) 152–155. https://www.sciencedirect.com/science/article/pii/S0168900201000894. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[21] "*The DECFILES Package.*" http://lhcbdoc.web.cern.ch/lhcbdoc/decfiles/. Accessed: 2021-12-04.

[22] "*Controlling the decay: DecFiles — LHCb Starterkit Lessons documentation.*" https://lhcb.github.io/starterkit-lessons/second-analysis-steps/simulation-dkfiles.html. Accessed: 2021-12-04.

[23] S. Agostinelli *et al.*, "*Geant4—a simulation toolkit*," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators,*

Spectrometers, Detectors and Associated Equipment **506** no. 3, (2003) 250–303. https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[24] "*The BOOLE Project*." http://lhcbdoc.web.cern.ch/lhcbdoc/boole/. Accessed: 2021-12-04.

[25] "*The MOORE Project*." http://lhcbdoc.web.cern.ch/lhcbdoc/moore/. Accessed: 2021-12-04.

[26] "*The BRUNEL Project*." http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/. Accessed: 2021-12-04.

[27] "*The DAVINCI Project*." http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/. Accessed: 2021-12-04.

[28] G. A. Cowan, D. C. Craik, and M. D. Needham, "*RapidSim: an application for the fast simulation of heavy-quark hadron decays*," *Comput. Phys. Commun.* **214** (2017) 239–246, arXiv:1612.07489 [hep-ex].

[29] J. Badier, C. Bemporad, J. Boucrot, O. Callot, P. Charpentier, A. Cnops, M. Crozon, P. Delpierre, J. Detoeuf, G. Giannini, *et al.*, "*Mass and lifetime limits on new long-lived particles in 300 GeV/c π-interactions*," *Zeitschrift für Physik C Particles and Fields* **31** no. 1, (1986) 21–32.

[30] **Belle** Collaboration, D. Liventsev *et al.*, "*Search for heavy neutrinos at Belle*," *Phys. Rev. D* **87** no. 7, (2013) 071102, arXiv:1301.1105 [hep-ex]. [Erratum: Phys.Rev.D 95, 099903 (2017)].

[31] D. collaboration *et al.*, "*Search for neutral heavy leptons produced in Z decays*," *Zeitschrift für Physik C Particles and Fields* **74** no. 1, (1997) 57–71.

[32] **ATLAS** Collaboration, G. 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*," *JHEP* **10** (2019) 265, arXiv:1905.09787 [hep-ex].

[33] **LHCb** Collaboration, R. Aaij *et al.*, "*Search for Majorana neutrinos in $B^- \to \pi^+\mu^-\mu^-$ decays*," *Phys. Rev. Lett.* **112** no. 13, (2014) 131802, arXiv:1401.5361 [hep-ex].

[34] K. Bondarenko, A. Boyarsky, D. Gorbunov, and O. Ruchayskiy, "*Phenomenology of GeV-scale Heavy Neutral Leptons*," *JHEP* **11** (2018) 032, arXiv:1805.08567 [hep-ph].

[35] "*GaussInGeneralFAQ ¡ LHCb/FAQ ¡ TWiki*." https://twiki.cern.ch/twiki/bin/view/LHCb/FAQ/GaussInGeneralFAQ.

[36] **LHCb** Collaboration, R. Aaij *et al.*, "*Measurement of b hadron fractions in 13 TeV pp collisions*," *Phys. Rev. D* **100** no. 3, (2019) 031102, arXiv:1902.06794 [hep-ex].

[37] **LHCb** Collaboration, R. Aaij *et al.*, "*Measurement of the B_c^- meson production fraction and asymmetry in 7 and 13 TeV pp collisions*," *Phys. Rev. D* **100** no. 11, (2019) 112006, arXiv:1910.13404 [hep-ex].

[38] **LHCb** Collaboration, R. Aaij *et al.*, "*Searches for low-mass dimuon resonances*," *JHEP* **10** (2020) 156, arXiv:2007.03923 [hep-ex].

[39] **LHCb** Collaboration, R. Aaij *et al.*, "*Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s} = 7$ and 13 TeV*," *JHEP* **12** (2017) 026, arXiv:1710.04921 [hep-ex].

[40] R. Gartner, "*Determination of the detector acceptance for heavy neutral lepton searches at LHCb*," Aug, 2021. not published.

[41] "*LHCb Performance Numbers*." https://lhcb.web.cern.ch/speakersbureau/html/PerformanceNumbers.html.

[42] R. Wischnewski, "*Errors in weighted histograms*," 8, 2000. https://www.zeuthen.desy.de/~wischnew/amanda/discussion/wgterror/working.html.

[43] **Particle Data Group** Collaboration, P. Zyla *et al.*, "*Review of Particle Physics*," *PTEP* **2020** no. 8, (2020) 083C01.

[44] P. D. Group *et al.*, "*Review of particle physics*," *Progress of Theoretical and Experimental Physics* **2020** no. 8, (2020) 1–2093.

[45] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, "*Parameterized neural networks for high-energy physics*," *Eur. Phys. J. C* **76** no. 5, (2016) 235, arXiv:1601.07913 [hep-ex].

[46] G. Punzi, "*Sensitivity of searches for new signals and its optimization*," *eConf* **C030908** (2003) MODT002, arXiv:physics/0308063 [physics]. [,79(2003)].

[47] W. Buttinger, "*Background Estimation with the ABCD Method Featuring the TRooFit Toolkit*,". https://twiki.cern.ch/twiki/pub/Main/ABCDMethod/ABCDGuide_draft18Oct18.pdf.

[48] L. Anderlini *et al.*, "*The PIDCalib package*," Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016. https://cds.cern.ch/record/2202412.

[49] **CMS Collaboration** Collaboration, C. Collaboration, "*Search for long-lived heavy neutral leptons with displaced vertices in pp collisions at* $\sqrt{s} = 13\,\text{TeV}$ *with the CMS detector*," tech. rep., CERN, Geneva, 2021. https://cds.cern.ch/record/2777047.

Acknowledgments

First, I want to thank Stephanie Hansmann-Menzemer for giving me the opportunity to work on this analysis and introducing me to Martino Borsato. I want to thank Martino for the fruitful discussions while working on this project as well as the helpful feedback he gave me. Also a big thank you to the members of EPFL which we are working with on this analysis: Lesya Shchutska, Federico Leo Redi and Serhii Cholak. The weekly status update discussions with all of them made it possible to overcome many problems and gave incentive to look at things from another perspective. I want to thank Finja Reichardt for taking her time to support me even when she was also busy. Especially in this time of Corona where it's barely possible to socialize with someone. As such I'm also grateful for my friends of "das Lerngruppe". Our virtual meetings and game sessions help me be less lonely in these times of isolation. Of course I also want to thank my family. Thank you for standing by my side my whole life. I can't express my gratitude enough for giving me all the support you can in every way possible.

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.
Heidelberg, den 04.12.2021,