Department of Physics and Astronomy
University of Heidelberg

Bachelor Thesis in Physics
submitted by

Rebecca Gartner

born in Karlsruhe (Germany)

2021

Determination of the detector acceptance for heavy neutral lepton searches at LHCb

This Bachelor Thesis has been carried out by Rebecca Gartner at the
Physikalisches Institut in Heidelberg
under the supervision of
Prof. Dr. Stephanie Hansmann-Menzemer

Contents

Bestimmung der Detektorakzeptanz für die Suche nach schweren neutralen Leptonen am LHCb Experiment:

Diese Arbeit ist Teil einer Suche nach schweren neutralen Leptonen (HNLs) in B-Zerfällen. Es werden Zerfälle der Art $B \to X\mu^+ N(\to \mu^\pm\pi^\mp)$, wobei N das HNL ist, untersucht. Dabei werden Daten aus Run 2 des LHCb Experiments benutzt. In dieser Arbeit werden die Detektorakzeptanz und dazugehörige Effizienzen für Zerfälle im LHCb Detektor berechnet. RapidSim wird benutzt, um Monte Carlo Daten auf Generator Level zu erzeugen. Zuerst wird es mit den bereits existierenden, voll simulierten LHCb Daten verglichen. Als nächstes wird RapidSim verwendet, um Effizienzen für HNL Massen, die nicht in den voll simulierten Daten enthalten sind, zu berechnen und damit die Sensitivität der Analyse abzuschätzen. Darüber hinaus wird eine Schätzung der Sensitivität bei einer Suche nach HNLs in $N \to \mu^+\pi^-$ berechnet.

Determination of the detector acceptance for heavy neutral lepton searches at LHCb:

This thesis is part of a search for heavy neutral leptons (HNLs) in B decays. The investigated decays are of the type $B \to X\mu^+ N(\to \mu^\pm\pi^\mp)$, where N is the HNL. Run 2 data from the LHCb experiment is used. In this thesis the detector acceptance and corresponding efficiencies for the decays within the LHCb detector are determined. RapidSim is used to produce generator level Monte Carlo data. It is first compared to to the existing fully simulated LHCb datasets. Next, RapidSim is used to calculate the efficiencies for HNL masses which are not available in the fully simulated samples and these efficiencies are used to estimate the sensitivity of the analysis. Furthermore an estimate for the expected sensitivity of an HNL search in $N \to \mu^+\pi^-$ is calculated.

Part I:
Introduction

1 Introduction

1.1 Motivation

In today's understanding of physics, the Standard Model (SM) of particle physics is a very well tested model and provides answers to many fundamental questions. However, there are some phenomena that cannot be explained within the Standard Model and there is experimental evidence that physics beyond the Standard Model exists. Therefore a big focus in modern particle physics is the search for New Physics.

The existence of heavy neutral leptons (HNLs) might be an explanation to some observed phenomena like the mass of neutrinos, dark matter and baryon asymmetry in the universe. Several attempts to search for HNLs have been performed, but although a large fraction of the parameter space was explored, no HNLs have been found so far.

There has also been a search for HNLs by the LHCb collaboration. Using Run 1 data, an analysis in the exclusive channel $B^+ \to \mu^+(N \to \mu^+\pi^-)$, where N is the HNL, was performed. However, this exclusive analysis was not able to improve limits coming from other experiments. This thesis contributes to a new analysis that uses Run 2 data to search for HNLs in the inclusive decay channel $B \to X\mu^+(N \to \mu^+\pi^-)$, where X stands for non-reconstructed particles. This search inclusively targets HNLs from various partially reconstructed B meson decays. With this method it is expected to reach a higher sensitivity than the previous exclusive analysis.

The goal of this thesis is to study the detector acceptance for various HNL masses. The HNL mass cannot be predicted from theory and in the LHCb Monte Carlo (MC) data only a few HNL masses are fully simulated. It is important to know the detector acceptance for more than these few HNL masses to estimate the sensitivity of this analysis. The program RapidSim is used to produce generator level Monte Carlo data in the HNL mass range of 1 to 6 GeV. The efficiencies determined in this thesis are used to estimate the sensitivity of the analysis.

Moreover, RapidSim is used to determine the detector acceptance for an inclusive HNL search in $N \to \mu\pi$. With the calculated efficiencies the sensitivity of this analysis is estimated.

1.2 The LHCb experiment

1.2.1 General overview of the LHCb detector

CERN, the European Organization for Nuclear Research, is the home of the largest and most powerful circular collider ever built: the Large Hadron Collider (LHC). The LHC has a circumference of 27 kilometers. There are four main experiments observing pp collisions at the LHC. These are ATLAS, CMS, ALICE and LHCb. This thesis considers the LHCb detector, which is the smallest of the four main experiments. At LHCb mainly decays that include hadrons with beauty- or charm quarks are studied (B and D hadrons). With the LHCb experiment physicists want to perform precision measurements of loop included processes such as CP-violations and very rare decays in order to test the Standard Model and search for physics beyond the Standard Model. Until now LHCb had two runs. Run 1 took place from 2009-2013 with an integrated luminosity of 3 fb^{-1} and a center-of-mass energy of $\sqrt{s} = 7.8$ TeV. Run 2 took place from 2015-2018 with a center-of-mass energy of $\sqrt{s} = 13$ TeV and an integrated luminosity of 5.9 fb^{-1}. Currently the LHCb detector is shut down and upgraded to make it possible to operate at a five times higher instantaneous luminosity in Run 3 than in Run 2. Furthermore the experiment plans to run at 40 MHz readout. The online event selection will be done entirely by software [11]. A cross section of the LHCb detector can be found in Fig. 1.1.
The LHCb detector is a single-arm forward spectrometer. In the LHCb detector the z-axis is defined parallel to the beam axis. If the beam goes from the interaction point towards the muon stations M1-M5 (see Fig. 1.1) the beam goes downstream. If it goes in the opposite direction it goes upstream. In fig. 1.1 the y-axis points up and the x-axis, which points away from the center of the LHC accelerator ring and is parallel to the ground, points out of the figure. The pseudorapidity is defined as

$$\eta = -ln[tan(\theta/2)] \tag{1.1}$$

where θ is the horizontal angle in the z-y-plane. The LHCb detector covers a pseudorapidity range of $2 < \eta < 5$.

1.2.2 Tracking System

An important part of the detector is the Vertex Locator (VELO). The VELO is used to identify the decay vertex of B and D mesons. It covers the angular acceptance of the detector in $2 < \eta < 5$. The VELO measures the distance between the collision point, the

so-called primary vertex, and the decay point of the B/D hadrons, the secondary vertex, with a resolution of 10 μm. Further information for track reconstruction is obtained via the Trigger Tracker (TT), which gives first estimates of the particles momenta and measures the trajectories of low-momentum particles that might not be tracked in the tracking stations due to deflection in the magnets. The tracking stations (T1-T3), which are separated into Inner Tracker (IT) and Outer Tracker (OT), measure momenta of the particles. The Inner Tracker has a resolution of about 50 μm in x direction. The Outer Tracker has a spatial resolution of about 200 μm for reconstructed ionising particle beams [12].

1.2.3 Particle Identification System

At LHCb the RICH detector (RIng CHerenkov imaging detector) identifies various different particles utilizing Cherenkov radiation. The RICH system is made of two detectors, which are placed at different positions. While RICH1 is upstream of the magnet having the function to identify low momentum particles, RICH2 identifies particles with higher momenta and thus it is placed downstream of the tracking stations. In the LHCb

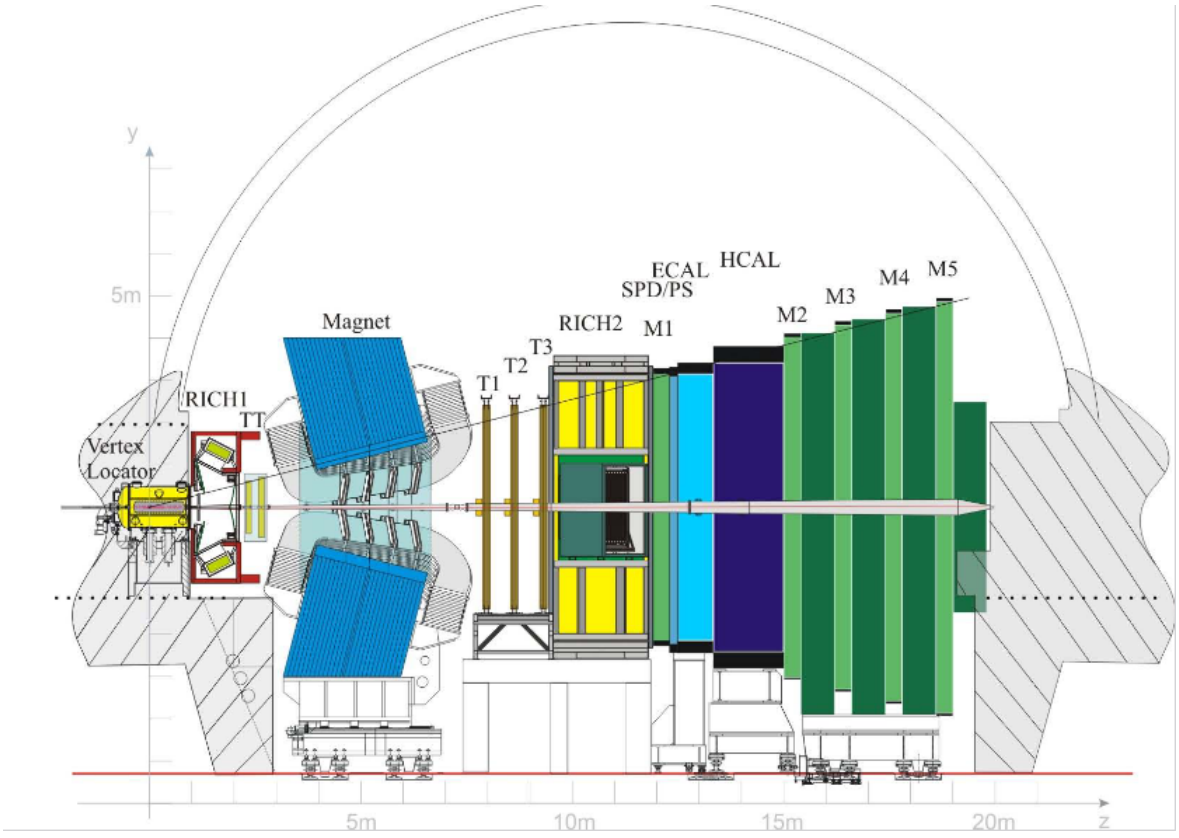


Figure 1.1: The LHCb detector [10]

calorimeter system there are the Scintillating Pad Detector (SPD), the PreShower (PS), the Electromagnetic CALorimeter (ECAL) and the Hadronic CALorimeter (HCAL). They measure the energy of electrons, photons and hadrons that have high transverse momentum. The information from ECAL and HCAL are crucial for particle identification. At the end of the detector there are the muon stations (M1 - M5). M1 has the function to improve the p_T measurement and is therefore placed upstream of the calorimeters. M2-M5 are placed downstream of the calorimeters and are used to measure the path of the muons.

1.2.4 Trigger System

The LHCb detector records a huge amount of data (ca. 1 TB per second). It is not possible to store all this data, thus triggers are needed. The first trigger, the level zero trigger, L0, is included in hardware and used to select particles which deposit energy in the calorimeters and have high transverse energy, E_T, and transverse momentum, p_T. There are two other triggers, High Level Triggers (HLT1 and HLT2), which are implemented in software. HLT1 and HLT2 fully reconstruct the events and also verify the hardware trigger decision. After the reconstruction there are two possible cases: signal candidates can be Trigger On Signal (TOS), which means that the particle that triggered the trigger decision belongs to the signal , and Trigger Independent of Signal (TIS), which means that the triggering particles do not belong to the signal decay.

1.2.5 Monte Carlo Simulations

In many cases physicists work with simulated data of certain decays before looking at the real data. In that way one can develop and test the analysis strategy beforehand and see how the data should look like if one e.g. is investigating phenomena beyond the Standard Model. In a Monte Carlo simulation one randomly produces data that should describe the data measured in the experiment. At LHCb Monte Carlo events are generated within the GAUSS framework. PYTHIA is used to produce proton-proton collisions and EVTGEN simulates the further decay of unstable particles from the proton-proton collision. GEANT4 is used to simulate the interaction with material. Other applications are used to simulate the detector and the trigger (see Fig. 1.2).

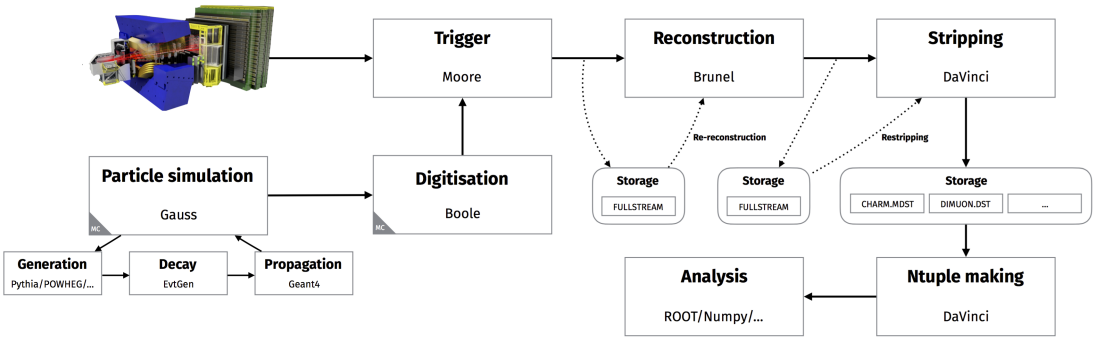


Figure 1.2: LHCb data flow [6]

1.3 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is a theory that classifies and describes the behaviour of subatomic particles and three out of the four fundamental forces: the weak interaction, strong interaction and electromagnetic interaction. It is a very well tested theory, but still there are some phenomena that cannot be explained within the Standard Model like the mass of neutrinos or the existence of dark matter.
Within the Standard Model there are two different classes of fundamental particles: Fermions and bosons. All matter and anti-matter particles are fermions and therefore have a half-integer spin. There are twelve matter particles and twelve corresponding anti-matter particles, which have the same mass, but opposite charges and quantum numbers. Six of them are quarks and the other six are leptons. Of those six leptons three are charged and three don't carry charge. In the Standard Model the forces are described by Quantum Field Theories, where the interactions are mediated by the exchange of gauge bosons. For the weak interaction this boson is the W/Z boson, for the electromagnetic it is the photon and for the strong interaction the gluon. The Higgs boson is a recent addition to the Standard Model. It is responsible for the mass of particles. The bosons all have spin 1, except for the Higgs boson, which has spin 0. In this thesis natural units are used, which means $c = \hbar = 1$. At low energies, quarks exist only in bound states due to confinement. Hadrons are particles that consist of several quarks. If the number of quarks in a hadron is odd (usually three), it is called a baryon. If the number is even (usually two - a quark-antiquark pair), it is called a meson. In this thesis, the focus lies on B mesons.

1.3.1 Neutrinos and Neutrino Oscillations

In the Standard Model, there are neutrinos in three lepton flavors: tau neutrinos (ν_τ), electron neutrinos (ν_e) and muon neutrinos (ν_μ). In all previous experiments neutrinos

were found to have a left-handed helicity, which means that their spin and momentum are anti-parallel. Usually particles exist in the left-handed and the right-handed (spin and momentum parallel) state. The heavy neutral leptons searched in the analysis to which this thesis contributes to are right-handed neutrinos. Neutrinos can oscillate. This means that e.g. a tau neutrino ν_τ doesn't have to stay a ν_τ forever, instead it can oscillate into an electron neutrino ν_e or an muon neutrino ν_μ. A first experimental hint for neutrino oscillations was found in the 1960s when the Ray Davis's Homestack experiment detected a deficit in the flux of solar neutrinos compared to predictions from the Standard Solar Model. However a clear evidence for neutrino oscillations was only found in 2001 at the Sudbury Neutrino Observatory [7]. Neutrinos oscillate because their weak eigenstate, a quantum state produced by the weak interaction, is not the same as their mass eigenstate, which means that the weak eigenstates don't have a well defined energy. In the Pontecorvo Maki Nakagawa Sakata matrix (PMNS matrix) information about the mismatch between those quantum states is given. A similar matrix is the CKM matrix, which describes the mismatch of the eigenstates of quarks when they propagate and interact with the weak interaction. While the CKM matrix being close to unity shows a clear pattern, the PMNS matrix doesn't. Neutrinos propagate as a

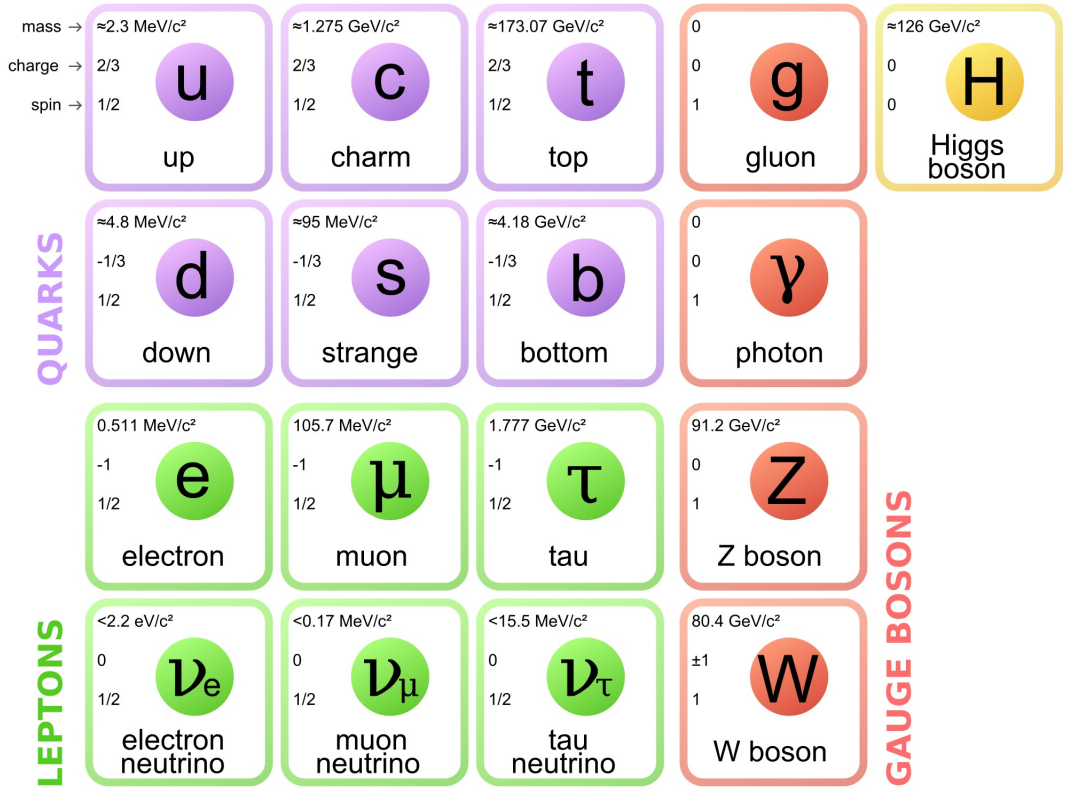


Figure 1.3: Elementary particles in the Standard Model [9]

Particle	Quark content	Mass [MeV]	Mean lifetime [s]
B^+	$u\bar{b}$	5279.2	1.6×10^{-12}
B^0	$d\bar{b}$	5279.5	1.5×10^{-12}
B_s^0	$s\bar{b}$	5366.3	1.5×10^{-12}
B_c^+	$c\bar{b}$	6276.0	4.6×10^{-13}
π^+	$u\bar{d}$	139.6	2.6×10^{-8}
π^0	$u\bar{u}/d\bar{d}$	135.0	8.4×10^{-17}
ρ^+	$u\bar{d}$	775.4	4.5×10^{-24}
ρ^0	$u\bar{u}/d\bar{d}$	775.5	4.5×10^{-24}
D^+	$c\bar{d}$	1869.6	1.0×10^{-12}
D^0	$c\bar{u}$	1864.8	4.1×10^{-13}
D^{*+}	$c\bar{d}$	2010.3	6.9×10^{-21}
D^{*0}	$c\bar{u}$	2007.0	3.1×10^{-22}
D^{*+}	$c\bar{s}$	1968.5	5.0×10^{-13}
D_s^{*+}	$c\bar{s}$	2112.3	3.4×10^{-22}

Table 1.1: Different mesons that appear in beauty decays, their quark content, mass and mean lifetime (charge conjugates are implied) [8]

linear coherent superposition of the mass eigenstates. If the mass of the neutrinos in their mass eigenstate is not the same for all mass eigenstates, a phase difference between the different components occurs which leads to the existence of neutrino oscillations. If the mass of the neutrinos in their eigenstates is not the same, the mass cannot be zero for all neutrinos. Even though the mass of the neutrino is unknown until now, upper limits can be set ($m_\nu < 0.8$ eV at 90% CL [1]). It is clear that the neutrino mass is much smaller than the masses of other SM particles. It is unknown why this is the case.

1.3.2 Heavy Neutral Leptons

Heavy neutral leptons (HNLs), also known as sterile neutrinos N, are hypothetical particles that are part of the compact Neutrino Minimal Standard Model (νMSM). If they exists, heavy neutral leptons interact with Standard Model fields the same way as the known SM neutrinos:

$$\begin{aligned}\mathcal{L}_{\text{HNL, int}} &= \frac{G_F}{2\sqrt{2}} W_\mu^+ \overline{N^c} \sum_\alpha U_{\alpha N}^* \gamma^\mu \left(1-\gamma_5\right) \ell_a^- \\ &+ \frac{G_F}{4\cos\left(\theta_W\right)} Z_\mu \overline{N^c} \sum_\alpha U_{\alpha N}^* \gamma^\mu \left(1-\gamma_5\right) \nu_\alpha + \text{ h.c.}\end{aligned} \tag{1.2}$$

The only difference is that the coupling is strongly suppressed by the mixing angles $U_{\alpha N}$. In this equation G_F is the Fermi constant and θ_W the Weinberg angle, which gives

the strength of the coupling. The HNL is $\overline{N^c}$ and $U^*_{\alpha N}$ is the matrix that includes the mixing angles between the active neutrinos and the HNL. The Dirac matrices are γ_5 and γ^μ. The interaction of the HNL with charged leptons (ℓ_a) via a W boson is described in the first part of the equation. In the second part the interaction of the HNL with uncharged leptons (ν_α) via a Z boson is described. The sum runs over all lepton flavors ($\alpha = \tau, \mu, e.$) The HNL mass m_N, the HNL lifetime τ and the mixing angles $U_{\alpha N}$ are dependent on each other:

$$\frac{1}{\tau} \propto \Gamma \propto m_N^5 U_{\alpha N}^2 \tag{1.3}$$

where Γ is the decay width. The mixing probability between active neutrinos and the HNL is given by $U^2_{\alpha N}$. It is not possible to predict the mass, $U_{\alpha N}$ or the lifetime of the HNL from theory, they are correlated input parameters to the theory[3].

Part II:
Determination of the detector acceptance

2 Determination of the detector acceptance

2.1 Definition of the fiducial volume

The analysis to which this thesis contributes to aimes to search for HNLs in $B \to X\mu^+ N(\to \mu^\pm \pi^\mp)$. To estimate and optimize the sensitivity before looking at the data, full Monte Carlo data is produced. However, these fully simulated Monte Carlo samples are only available for few HNL masses and lifetimes (see Table 2.1). The analysis is split into semileptonic decays, leptonic B_c decays and leptonic B^+ decays. An exact definition of these decays is given later.

decay channel	HNL masses [GeV]	HNL lifetimes [ps]
semileptonic decays	1.6, 2, 3, 4	100, 1000
leptonic B_c	1, 1.6, 2, 3, 4, 5, 5.5, 6	0, 10, 100, 1000
leptonic B^+	1, 2, 3, 4, 4.5, 5	0, 10, 100

Table 2.1: masses and lifetimes for which fully simulated MC samples are available

However, to estimate the sensitivity of this analysis samples at more masses and lifetimes are needed. One possible solution would be to generate more full Monte Carlo data, but this would be very time-consuming. Therefore another approach is taken: For the sensitivity estimates it is required to know what fraction of simulated signal decays is reconstructed in the detector. This is the so-called total efficiency. For this analysis it is decided to split the total efficiency into two parts: The reconstruction efficiency and the fiducial volume efficiency. The fiducial volume efficiency is defined as the region of p, p_T and η of the particles of the signal decay that can be reconstructed with high efficiency. The fiducial volume efficiency can be calculated at generator level since p, p_T and η are measured with very good precision at LHCb. A cut on the reconstructed p has the same efficiency as a cut on the true p. The fiducial volume cuts used in the HNL analysis can be found in Table 2.2. The cuts on the pseudorapidity of the particles are placed according to the geometry of the detector. The cuts on p and p_T are optimized

taking into account the existing trigger selections. The optimization was performed by Maurice Morgenthaler, a master student of the Heidelberg LHCb group.

particle	$\eta_{\min}$	$\eta_{\max}$	p_T [GeV]	p [GeV]
min 1 μ	2.0	4.5	>1.1	>10.0
both μ	2.0	4.5	>0.3	>3.0
π	2.0	4.5	>1.0	>5.0
HNL	2.0	4.5	>2.0	>28.0

Table 2.2: fiducial volume cuts

The reconstruction efficiency is defined as the number of signal decays in the fiducial volume on reconstruction level after offline and trigger cuts have been applied divided by the number of signal decays in the fiducial volume at generator level.

$$\epsilon_{reco} = \frac{\#\text{signal decays on reconstruction level after FV, offline and trigger cuts}}{\#\text{signal decays on generator level after FV cuts}} \quad (2.1)$$

The trigger and offline cuts used in this analysis are given in Table 2.3 and Table 2.4.

Level	Trigger line	TOS (see sec. 1.2.4) on
L0	Muon or DiMuon	$\mu\mu$
HLT1	TrackMuon	$\mu\mu$
HLT2	Topo(Mu)2Body	$B \to \mu\mu\pi$
	ExoticaRHNu	
	MajoranaBLambdaMuDD (only SS muons, prescale 20%)	

Table 2.3: trigger cuts

Tracks	p_t >0.25 GeV, χ^2_{track} <4, P_{ghost} <0.35
Muons	p >3 GeV, χ^2_{IP} >12, PID$_\mu$ >0, PID$_\mu$-PID$_{K,p}$ >0
Pion	p >2 GeV, χ^2_{IP} >10, InMuonAcc = 1, isMuon = 0, ProbNNμ >0.1
$N \to \mu\pi$	m >1.5 GeV, p_T >0.7 GeV, χ^2_{vtx}/dof <10, χ^2_{fromPV} >100, DIRA$_{PV}$ >0.996, DIRA$_B$ >0.996
$B \to \mu\mu\pi$	m <6.5 GeV, χ^2_{vtx} <4, DIRA$_{PV}$ >0.99, (Bvtx - Nvtx)$_z$ >-1(+4) mm for SS(OS)

Table 2.4: offline cuts

The offline cuts select HNL candidates with a momentum pointing back to the primary vertex and the B decay vertex. The cuts on the pion and the muon guarantee well reconstructed trajectories in the tracking system and good particle identification.
The fiducial volume efficiency is the number of signal decays at generator level after the fiducial volume cuts divided by the number of signal decays without any cuts.

$$\epsilon_{FV} = \frac{\#\text{signal decays on generator level after FV cuts}}{\#\text{signal decays on generator level}} \tag{2.2}$$

Multiplying the reconstruction efficiency and the fiducial volume efficiency leads to the total efficiency of the reconstructed signal decays in the fiducial volume:

$$\epsilon_{total} = \epsilon_{FV} \cdot \epsilon_{reco} = \frac{N_{FV}}{N_{total}} \cdot \frac{N_{reco}}{N_{FV}} \tag{2.3}$$

Here N is the number of signal decays.
The reason for splitting the total efficiency into reconstruction efficiency and fiducial volume efficiency is that the total efficiency is mass dependent. Since only data for few HNL masses is given in the full Monte Carlo simulation it would be imprecise to just interpolate the total efficiency between these masses. However, since the definition of the reconstruction efficiency, ϵ_{reco}, is designed to be approximately flat as a function of HNL mass, it is possible to interpolate the reconstruction efficiency between the HNL masses of the available full Monte Carlo samples. The fiducial volume efficiency is dependent on the HNL mass, therefore more Monte Carlo data is needed for a precise interpolation between the HNL masses. However, as mentioned earlier generator level simulation is fully efficient to determinate the fiducial volume efficiency ϵ_{FV}. The goal of this thesis is to produce this additional Monte Carlo data with a fast Monte Carlo generator called RapidSim and to study the fiducial volume efficiency.
Firstly RapidSim needs to be evaluated to reproduce the full simulation. For this InLHCbAcceptance cuts (see Table 2.5) are applied to the MC data generated with RapidSim, because the full MC data is only available with the InLHCbAcceptance cuts.

particle	η_{min}	η_{max}
both μ	1.59	5.3
π	1.59	5.3

Table 2.5: InLHCbAcceptance cuts

In order to compare full and fast MC simulation the fiducial volume efficiency in LHCb acceptance, $\epsilon_{FVinLHCb}$, is determined. In contrast to the total fiducial volume efficiency, the fiducial volume efficiency in LHCb acceptance has the InLHCbAcceptance cuts applied in the denominator.

$$\epsilon_{FVinLHCb} = \frac{\#\text{signal decays on generator level after FV cuts}}{\#\text{signal decays on generator level after InLHCbAcceptance cuts}} \tag{2.4}$$

2.2 RapidSim

RapidSim is a fast Monte Carlo generator. It is specialized for the generation of B hadron decays. It is based upon the TGenPhaseSpace class from ROOT. RapidSim is much faster than usual Monte Carlo generators because it generates only the decaying particle and not the underlying event, e.g. for the B production it doesn't simulate the *pp*-collision, but instead takes the B meson spectra from a histogram, in which the B production spectrum data, that was simulated with PYTHIA, is stored [5]. Whenever the term "fast Monte Carlo" is used in this thesis, this refers to Monte Carlo data generated with RapidSim. The documentation of RapidSim can be found here [4].

2.3 Crosscheck of full Monte Carlo data

Producing the generator level Monte Carlo data with RapidSim is a good opportunity to crosscheck the full Monte Carlo data. In the full Monte Carlo simulation all decay channels are simulated separately, thus the data sets need to be weighted according to their contribution. Therefore it is useful to check if RapidSim produces Monte Carlo samples that are similar to the reweighted full Monte Carlo samples. The first step in this project is to produce fast Monte Carlo data with the same fiducial volume cuts than in the full MC samples. Thus the decay $B^0 \to \pi^-\mu^+(N \to \pi^-\mu^+)$ is used. Fast MC data was generated successfully and the distributions of the different variables match (see Fig. 2.2, 2.3 and section 5). In this crosscheck it was discovered that in the full Monte Carlo samples there is the cut $1.59 < \eta < 5.3$ applied on all charged final state particles, which means that in the investigated decay channel $B \to X_{prim}\ \mu^+(N \to \mu^+\pi^-)$ the cut is also applied on some of the particles that are not reconstructed, X_{prim} (see Fig. 2.1). This cut was mistakenly applied in the production of the full Monte Carlo samples. This will have to be taken into account in the final analysis. The effect of this mistake is studied in this thesis (see section 2.5). To be able to compare the full MC data and the fast MC data the η-cut was applied to all X_{prim} particles in RapidSim as well, but was removed for the fiducial volume calculations.

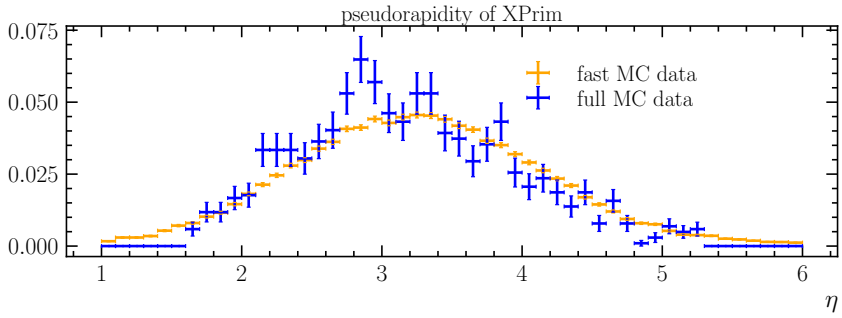


Figure 2.1: η distribution of the X_{prim} from full and fast MC data

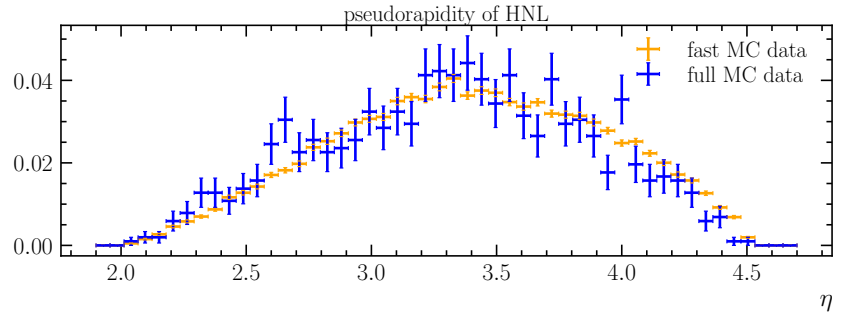


Figure 2.2: η distribution of the HNL from full and fast MC data

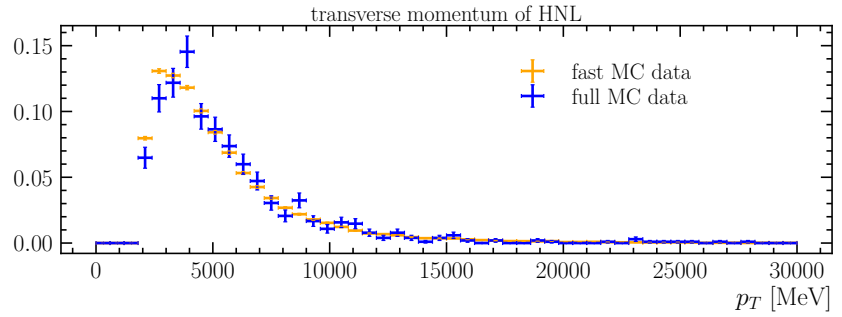
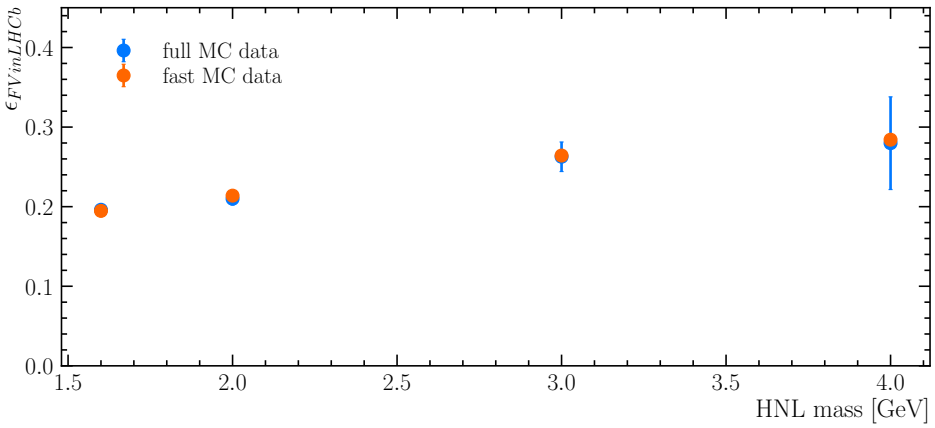


Figure 2.3: p_T of the HNL distributions from full and fast MC data

2.4 The decay $B^0 \to \pi^- \mu^+ (N \to \pi^- \mu^+)$

After checking that RapidSim reproduces the distributions of the full simulation rather well, first only the fiducial volume efficiency in LHCb acceptance for the decay $B^0 \to \pi^- \mu^+ (N \to \pi^- \mu^+)$ is studied. To determine how much the efficiencies from full and fast Monte Carlo data differ, the relative difference Δ_r was calculated for every HNL mass



mass [GeV]	1.6	2	3	4
relative difference [%]	0.7	1.8	0.6	1.5
pull [σ]	0.25	0.5	0.08	0.07

Figure 2.4: in plot: fiducial volume efficiencies in LHCb acceptance for decay $B^0 \to \pi^-\mu^+(N \to \pi^-\mu^+)$ calculated with full MC data and with fast MC data. In table: relative differences and pull at different masses for the same decay

that is included in the full MC sample as follows

$$\Delta_r = \frac{\epsilon^{fast}_{FVinLHCb} - \epsilon^{full}_{FVinLHCb}}{\epsilon^{fast}_{FVinLHCb}} \tag{2.5}$$

In this formula $\epsilon^{full}_{FVinLHCb}$ is the fiducial volume efficiency in LHCb acceptance calculated from the full MC samples and $\epsilon^{fast}_{FVinLHCb}$ is the one calculated from the fast MC data. If the relative difference is smaller than 5% the HNL search will not be limited by this uncertainty, thus it has been shown that RapidSim is valid to evaluate the fiducial volume efficiencies in this channel. The results of the calculation of the fiducial volume efficiency in LHCb acceptance is shown in Figure 2.4.

2.5 Semileptonic decays

Given that for the decay $B^0 \to \pi^-\mu^+(N \to \pi^-\mu^+)$ the fiducial volume efficiencies in LHCb acceptance, $\epsilon_{FVinLHCb}$, from the fast and full MC simulation fit very well together, more decays are included. Since the analysis is split in semileptonic decays, leptonic B_c decays and leptonic B^+ decays, the efficiencies are also calculated separately. The reason for the splitting is to gain a higher sensitivity. The leptonic decays are fully reconstructable and therefore have a much smaller background. The semileptonic decays are dominant at low HNL masses and the leptonic decays are dominant at higher masses.

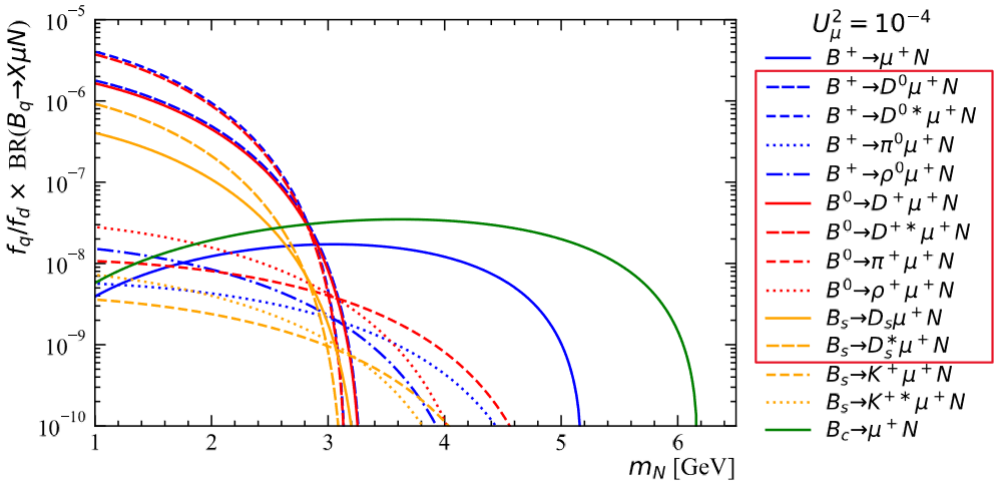
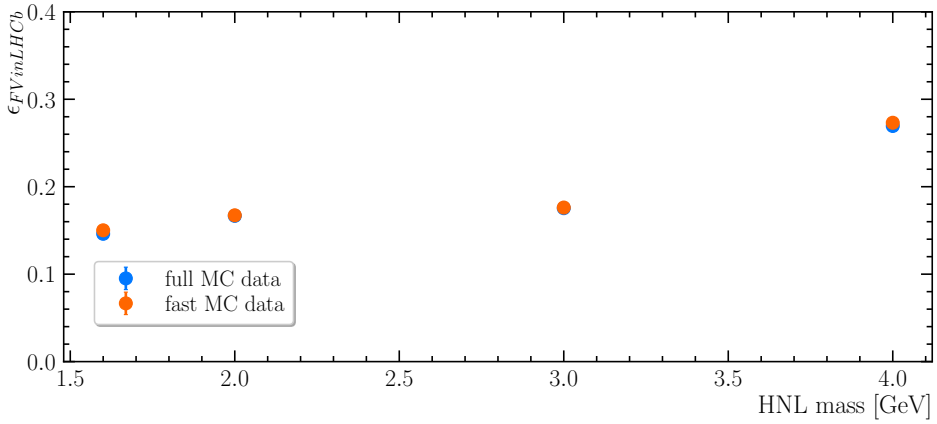


Figure 2.5: The decays in the red box are the semileptonic decays used in this analysis



mass [GeV]	1.6	2	3	4
relative difference [%]	2.6	0.4	0.4	1.4
pull [σ]	4.4	0.7	0.7	2.6

Figure 2.6: in plot: fiducial volume efficiencies in LHCb acceptance for semileptonic decays calculated with full MC data and with fast MC data. in table: relative differences and pulls at different masses for semileptonic decays

For simplicity, only the 10 dominant semileptonic decay channels are considered in this analysis. The list can be seen in Fig. 2.5.

The fiducial volume efficiencies in LHCb acceptance, $\epsilon_{FVinLHCb}$ are calculated successfully and they fit to efficiencies calculated from the full MC data.

As one can see in Figure 2.6 there is a pull of 4.4 σ at 1.6 GeV. Although the pull is high, this is considered acceptable since the relative difference of 2.6 % is still small enough. The high pull comes from the high number of events that are available in the full and fast MC samples at 1.6 GeV, which leads to a very small uncertainty.

After seeing the agreement between the fiducial volume efficiencies in LHCb acceptance in full and fast MC samples, more fiducial volume efficiencies in LHCb acceptance for the masses in between the HNL masses from the full MC samples are calculated. The result can be seen in Fig. 2.7. The shape of Fig. 2.7 can be explained by the branching ratios of the different semileptonic decays. Up to 3 GeV there are several decay channels that include a D meson as X_{prim}. D mesons have masses of ~ 1.9 GeV, while B mesons have masses of ~ 5.3 GeV (see Table 1.1). In the decay $B \to D\mu(N \to \mu\pi)$ the mass of the D meson already takes up a big part of the phase space, therefore there is not much energy left for the primary muon so it does not always reach the threshold for the fiducial volume cuts. Above a HNL mass of 3 GeV there are no semileptonic decays that have a D meson as X_{prim} possible. The leftover decays have π and ρ as X_{prim}. These particles are much lighter than the D meson. The primary muon now has more energy left than before and can more often reach the threshold for the fiducial volume cuts.Thus the efficiency increases.
In order to compare full and fast MC samples the η cut on X_{prim} that was mistakenly applied to full MC simulation was also applied in the fast MC simulation. The effect of this additional cut is investigated. For this investigation the fiducial volume efficiency was calculated, once with the η cut on X_{prim} applied in the numerator and once without the η cut on X_{prim} in the numerator. The result can be seen in Fig. 2.8. The effect is small, but visible. With the η cut on X_{prim} applied the fiducial volume efficiency is slightly lower than the fiducial volume efficiency without the η cut on X_{prim}. If the cut is not applied there will be a higher number of signal decays in the fiducial volume and the

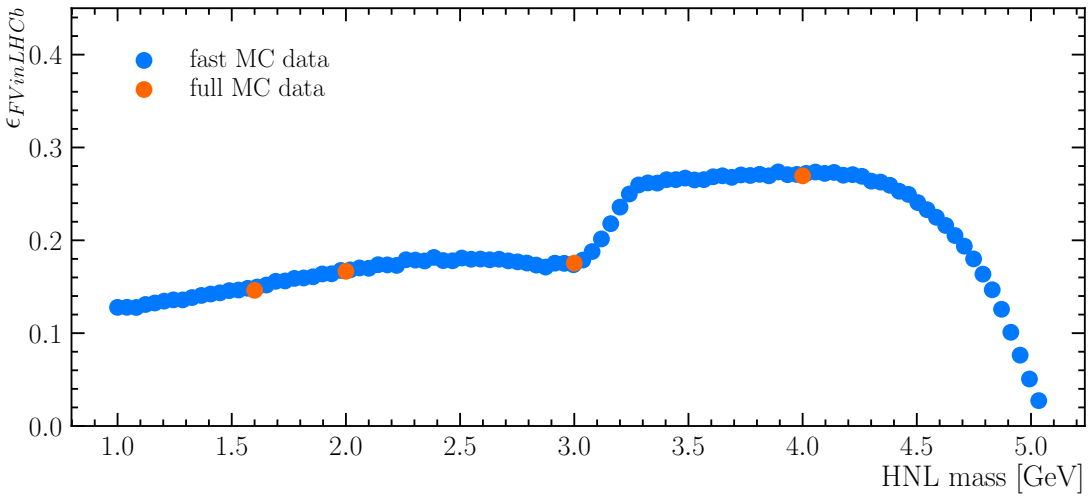


Figure 2.7: fiducial volume efficiencies in LHCb acceptance calculated with full and fast MC data for semileptonic decays

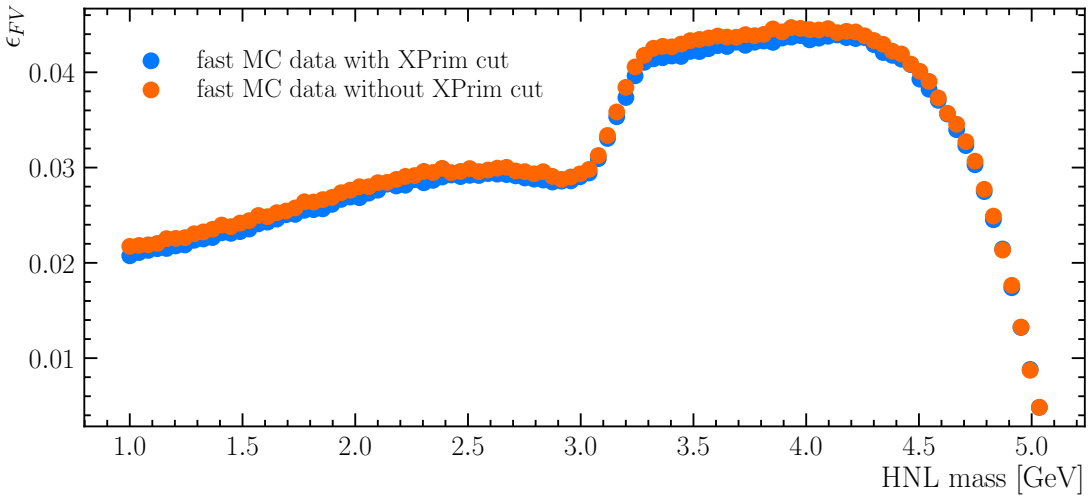
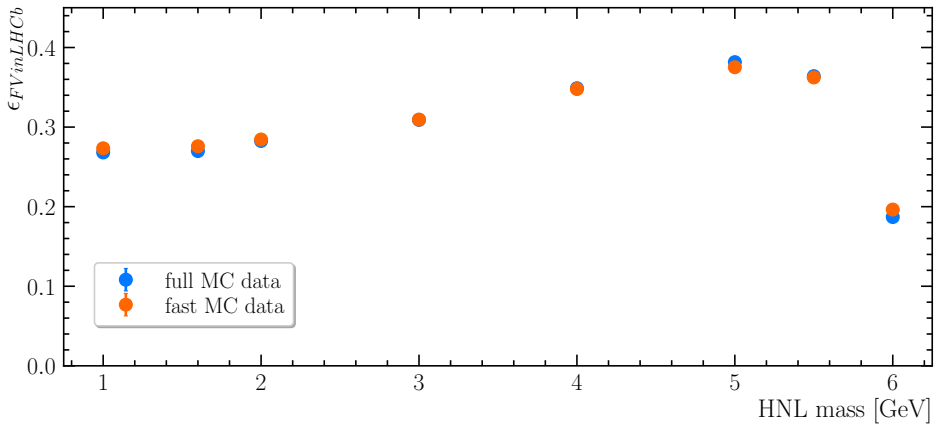


Figure 2.8: fiducial volume efficiencies calculated from full and fast MC samples for semileptonic decays with and without X_{prim} cut

efficiency increases. The relative differences between the fiducial volume efficiencies with and without the η cut on X_{prim} in the numerator increases by a maximum of 4.7%, but the increase has a mean of 2.2%. Also the fiducial volume efficiency in LHCb acceptance with and without the η cuts on X_{prim} are compared to the fiducial volume efficiencies in LHCb acceptance calculated from full MC samples. The result is that if the η cuts on X_{prim} are not applied the relative differences between the efficiencies from full and fast MC samples increase by 1.4% to 4.0% (mean: 2.2%).



mass [GeV]	1	1.6	2	3	4	5	5.5	6
relative difference [%]	1.90	2.18	0.64	0.09	0.26	1.69	0.43	4.77
pull [σ]	4.26	4.96	1.53	0.25	0.71	4.94	1.25	9.11

Figure 2.9: in plot: fiducial volume efficiencies in LHCb acceptance calculated from full and fast MC samples for B_c decays. In table: relative differences and pulls at different masses for B_c decays

2.6 B_c leptonic decays

The fiducial volume efficiency in LHCb acceptance is also calculated for the B_c decay $B_c^+ \to \mu^+(N \to \pi^-\mu^+)$.

For 1, 1.6, 5 and 6 GeV the pull is larger than 3 σ and at 6 GeV almost 5% relative difference is reached. However, for other HNL masses the relative differences are low. With the relative differences mostly being low, these high pulls are acceptable. It is known that the B_c p_T spectrum is not correctly simulated in RapidSim.

More full Monte Carlo data is produced in order to investigate the effect of the incorrect B_c p_T spectrum since MC data without any cuts is needed for this. However, the new full MC sample is relatively small, containing only 50000 events. The difference between the B_c p_T spectrum from full and fast MC data can be seen in Fig 2.10.

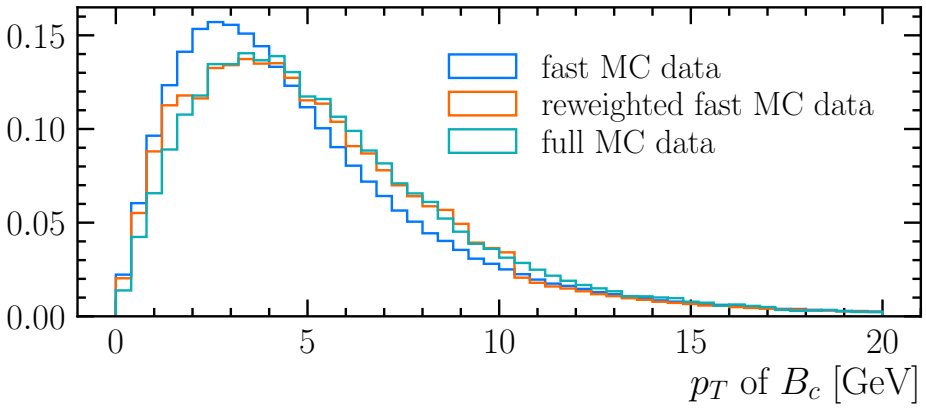
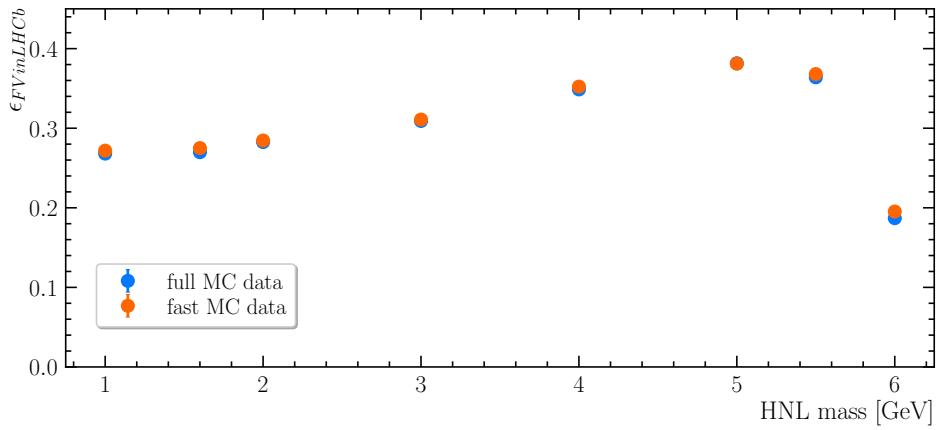


Figure 2.10: p_T distribution of B_c from full, fast and reweighted fast MC samples

Using this new fully simulated MC data without cuts and fast MC data from RapidSim, 2D histograms where obtained from the $log(p_T)$ and η distributions of the full and fast Monte Carlo samples. By dividing the number of events per bin from the histograms obtained with full and fast MC data, a weight can be assigned to p_T and η regions for which enough events are available. By taking the sum of the weights before and after the fiducial volume cuts, the fiducial volume efficiency can be calculated. This is done for the HNL masses given in the full MC sample. As one can see in Fig. 2.11, the reweighted efficiencies do not vary significantly from the unweighted efficiencies. Therefore the effect of the incorrectly simulated B_c p_T spectrum is negligible.



mass [GeV]	1	1.6	2	3	4	5	5.5	6
relative difference [%]	1.4	1.9	0.7	0.6	1.0	0.1	1.1	4.3
pull [σ]	3.1	4.3	1.7	1.6	2.8	0.3	3.3	8.3

Figure 2.11: in plot: fiducial volume efficiencies in LHCb acceptance calculated from full and fast MC samples for B_c decays with reweighted fast MC samples. In table: relative differences and pulls at different masses for B_c decays

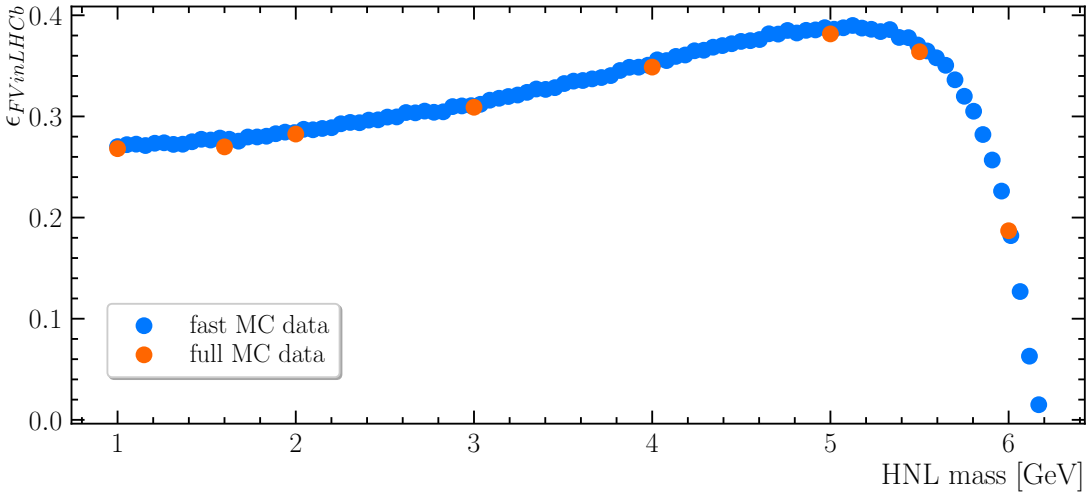
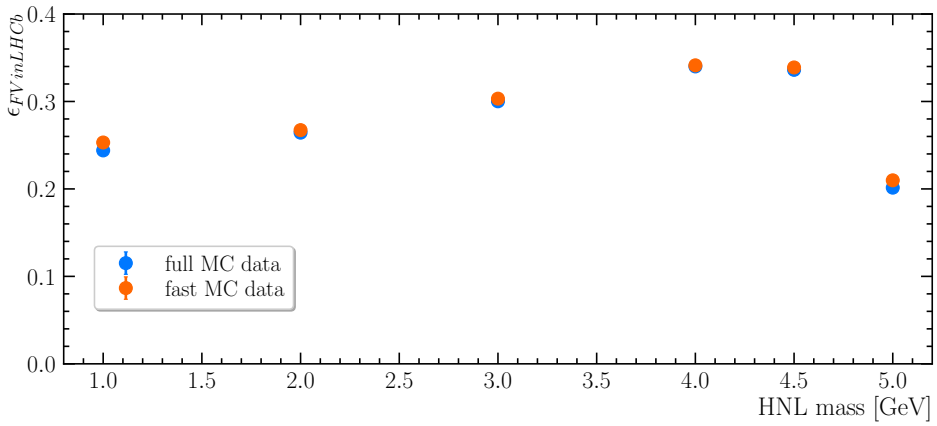


Figure 2.12: Fiducial volume efficiencies in LHCb acceptance from full and fast MC samples for B_c decays

The fiducial volume efficiencies in LHCb acceptance are calculated for 100 HNL masses. The result can be seen in Fig. 2.12.
The shape of this distribution can be explained. With increasing HNL mass the decay products of the HNL, μ and π, will become more energetic and will more often reach the threshold needed for the fiducial volume cuts and the efficiency increases as well. But if the HNL mass gets close to the B_c mass there is less energy left for the primary muon so it stops to reach the threshold for the fiducial volume cuts more and more and the efficiency decreases.

2.7 B^+ leptonic decays

For the decay $B^+ \to \mu^+(N \to \pi^-\mu^+)$ the fiducial volume efficiencies in LHCb acceptance are calculated as well. A similar result as for the B_c decay is obtained. The relative differences are always reasonably low, though sometimes a pull of >3 σ is reached. Similar to the B_c case this is not a problem. The higher pulls come from a high number of signal decays which leads to a very small uncertainty.



mass [GeV]	1	2	3	4	4.5	5
relative difference [%]	3.58	1.09	0.99	0.38	0.80	3.99
pull [σ]	5.83	1.98	1.92	0.76	1.52	4.41

Figure 2.13: in plot: fiducial volume efficiencies in LHCb acceptance calculated from full and fast MC samples for B^+ decays. In table: relative differences and uncertainties at different masses for B^+ decays

Also for B^+ the fiducial volume efficiencies in LHCb acceptance are calculated for various HNL masses in between HNL masses available in the full MC simulation. The shape of the distribution can be explained with the same argumentation as for the B_c channel: When the HNL mass increases, the decay products of the HNL will become more energetic and the efficiency rises. But with a heavy HNL the primary muon gets less energy and the efficiency drops again when the HNL gets heavier.

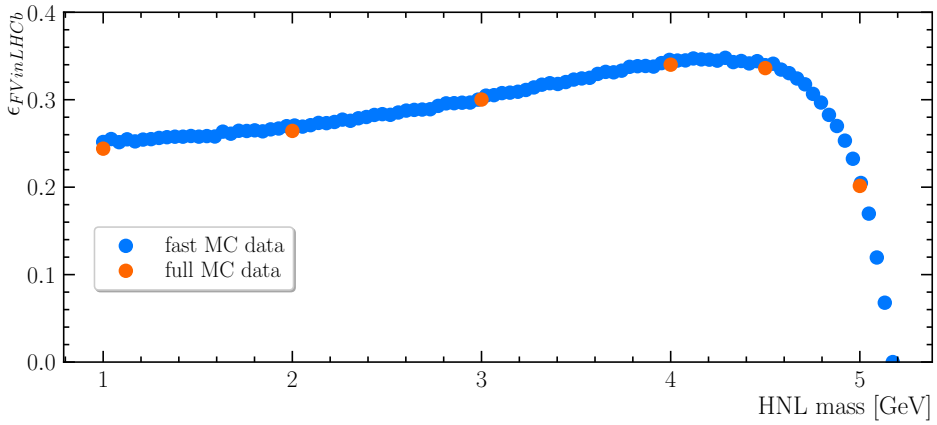


Figure 2.14: Fiducial volume efficiencies in LHCb acceptance from full and fast MC samples for B^+ decays

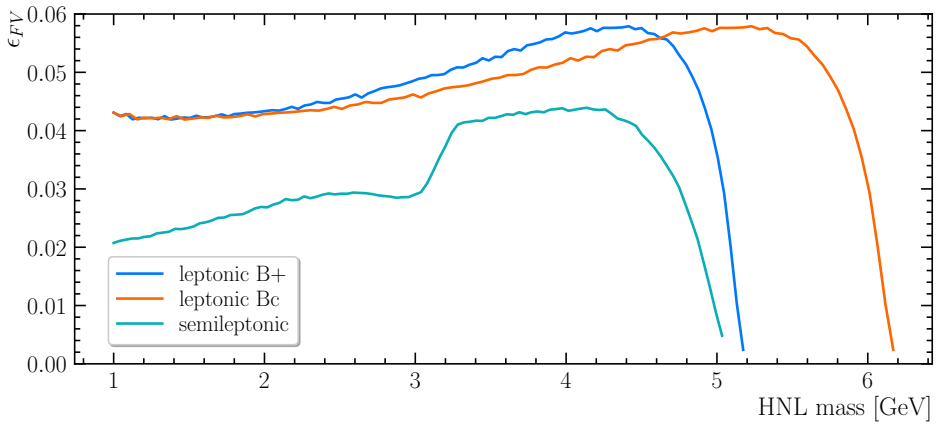


Figure 2.15: fiducial volume efficiencies for semileptonic decays, leptonic B_c decays and leptonic B^+ decays

2.8 Fiducial volume efficiencies

Given that the fiducial volume efficiencies in LHCb acceptance, $\epsilon_{FVinLHCb}$, calculated from full and fast MC samples agree very well for all channels, the fiducial volume efficiencies, ϵ_{FV}, are calculated. As expected, the fiducial volume efficiencies are much lower than the fiducial volume efficiencies in LHCb acceptance, because there are much more signal decays in the denominator since there are no cuts anymore. The shape of the efficiency distribution stays the same. The fiducial volume efficiencies for all the channels can be found in Fig. 2.15.

2.9 Expected number of signal decays including an HNL

With the previous calculated fiducial volume efficiencies one can now calculate the number of signal decays that include an HNL in the fiducial volume acceptance. The number of signal decays is calculated as follows:

$$N_{\text{decays}} = L \cdot \frac{\sigma_{pp\to B^{\pm}X}}{\epsilon_{pp\to B^{\pm}X}} \cdot \frac{f_q}{f_u} \cdot \text{BR}\left(B_q \to X\mu^+ N\right) \cdot \text{BR}\left(N \to \mu^+\pi^-\right) \cdot \epsilon_{\text{FV}} \tag{2.6}$$

with L being the luminosity, $\sigma(pp \to B^{\pm}X)$ the $pp \to B^{\pm}X$ cross section, $BR(b \to X\mu N)$ the branching ratio of the B meson decay, in which the HNL is produced and $BR(N \to \mu\pi)$ the branching ratio of the decaying HNL. The branching ratios are taken from a paper by Bondarenko et al. [2]. The cross section is given in a B_c p_T range of

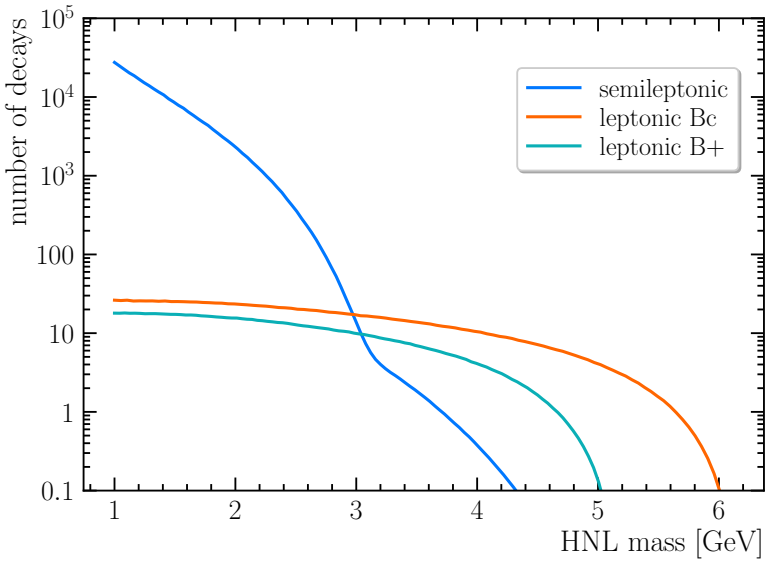


Figure 2.16: Number of signal decays in the fiducial volume that include an HNL

0-40 GeV and an η range of 2-4.5, therefore a normalization with the efficiency of this parameter space $\epsilon_{pp\to B^{\pm}X}$ is needed. Since only the decays that are actually within the fiducial volume are of interest, a multiplication with the efficiency of the fiducial volume is needed. The cross section is given for $pp \to B^{\pm}X$, but not only B^+ mesons are used in this analysis. Therefore a reweighting with the fragmentation fraction $\frac{f_q}{f_u}$ is needed. As one can see in Fig. 2.16 at low HNL masses the semileptonic decays have a really high number of signal decays. In general the leptonic B_c decays always have higher numbers of decays than the leptonic B^+ decays.
This fits to what is expected from theory. In Fig. 2.17 one can see the branching ratio of the decays. The shape is similar to the shape of Fig. 2.16.

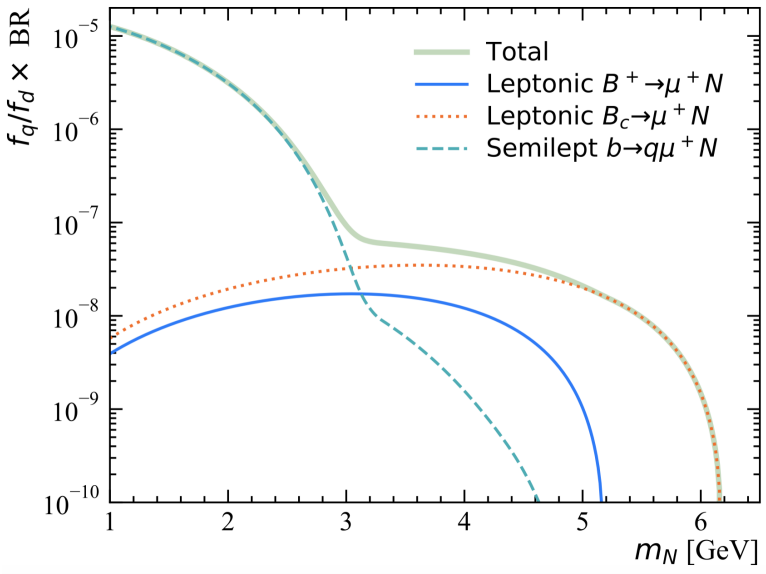


Figure 2.17: Branching ratios of different decay channels, data from [2], plot by Dr. Martino Borsato

2.10 Expected sensitivity

Furthermore, RapidSim is used to calculate an estimate of the expected sensitivity of the analysis. The number of reconstructed HNL decays $N_{HNL,reco}$ is calculated as a function of the HNL mass and lifetime via

$$N_{HNL,reco} = N_{decays} \cdot \epsilon_{reco} \tag{2.7}$$

where N_{decays} is the expected number of signal decays including an HNL for a given coupling $U_{\alpha N}$ and mass m_N and ϵ_{reco} the reconstruction efficiency. For the reconstruction efficiency an estimate based on Maurice Morgenthalers calculations is used. When the number of reconstructed HNL decays is bigger than three, there will be a peak in the invariant mass of the decay products of the HNL ($\mu\pi$) which will be detected with 95% CL. However, if zero signal decays are observed in a specific m_N-$U_{\alpha N}$ area in the data and the expected number of signal decays is bigger than three, this area can be excluded. Previous experiments have already set stringent limits in a huge area of the parameter space. The expected limits are calculated separately for each decay channel and under the assumption of zero background. The results of previous limits and the expected sensitivity calculated here can be seen in Fig. 2.18. As one can see, the leptonic B_c and B^+ decay channels will most likely not set a new world-best limit. However, for the

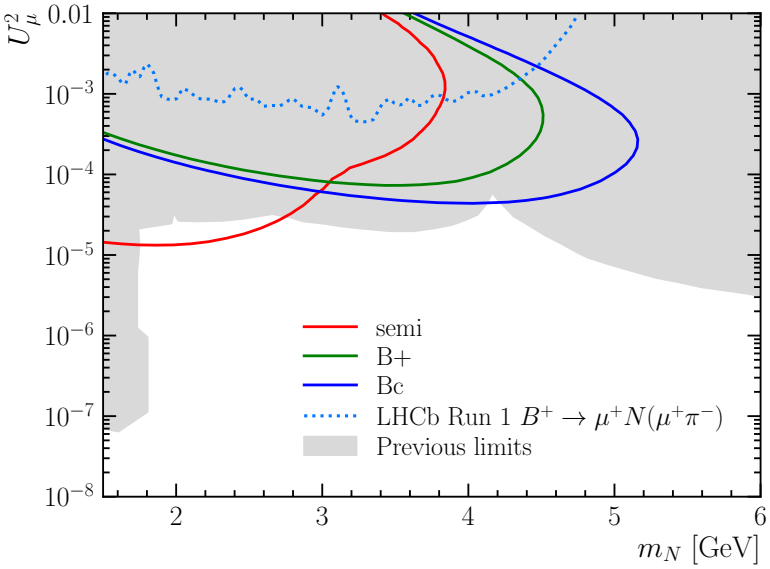


Figure 2.18: Expected sensitivity for semileptonic decays, leptonic B_c decays and leptonic B^+ decays

semileptonic decay channel at masses below 3 GeV the sensitivity is good. Depending on the background a new limit might be set in this area. The background is expected to be small because the signal signature is well distinguishable from Standard Model processes due to the $\mu\pi$ peak from the decaying HNL, the displacement of the HNL vertex and the required mass of the $\mu\pi$ ($m > 1.5$ GeV). The analysis distinguishes between signal decay in which both muons have the same sign (SS) and the opposite sign (OS). In the same sign case the background can be reduced even more. In leptonic B_c and leptonic B^+ decays the background is further minimized because the B meson can be reconstructed from its decay products, $\mu\mu\pi$.

Even though the background might be too large to set a new limit, this analysis will have a much better sensitivity than a previous exclusive HNL search at LHCb that was done on $B^+ \longrightarrow \mu^+ N(\longrightarrow \mu^-\pi^+)$ in 2013 (see Fig 2.18). Furthermore, in this analysis only Run 2 data is used. It is expected that the sensitivity of this analysis can be further improved when Run 3 data is available due to the higher data rate.

Part III:
Towards an inclusive search for heavy neutral leptons in $N \to \mu^+\pi^-$

3 Towards an inclusive search for heavy neutral leptons in $N \to \mu^+\pi^-$

It is also possible to search for HNLs only using as signature a displaced $N \to \mu^+\pi^-$. This means that the HNL can come from all decays that include an HNL. This analysis has not been done so far. In the low HNL mass range 1-6 GeV the dominant decays that include an HNL are the previously studied B decays. In the following, an estimate of the sensitivity of this search is presented.
Like before, the total efficiency is split into fiducial volume efficiency and reconstruction efficiency. In Table 3.1 one can see the used fiducial volume cuts. They are tighter than in the current analysis in order to reduce the background. Since this analysis will be even more inclusive than the current analysis, there will be more background. For the reconstruction efficiency the same estimate as in the current analysis is used, because full MC samples have not been produced yet. As a result, the used reconstruction efficiency is smaller than it should be. A solution for this would be to produce full MC samples with the cuts needed for the $N \to \mu\pi$ analysis, but this exceeds the time limits of this thesis.

particle	η_{min}	η_{max}	p[GeV]	p_T[GeV]
μ	2	5	$>$10	$>$0.5
π	2	5	$>$10	$>$0.5
HNL	-	-	-	$>$1

Table 3.1: Fiducial volume cuts

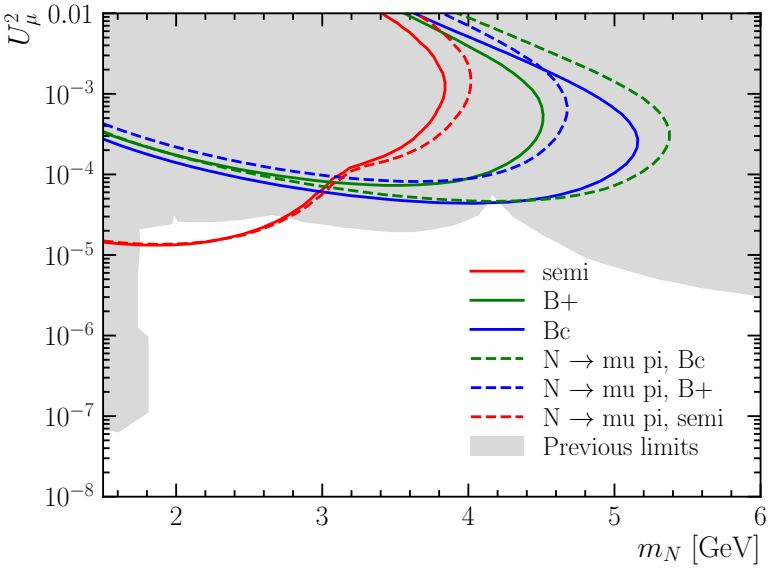


Figure 3.1: Expected sensitivities for current analysis and possible future $N \to \mu\pi$ analysis

As one can see in Fig. 3.1 the sensitivity of the possible future $N \to \mu\pi$ analysis is the most areas slightly better than in the current analysis. However, in the relevant m_N area of less then 3 GeV of the semileptonic decay channel, the sensitivities for the current analysis and the possible future $N \to \mu\pi$ analysis are almost the same. As mentioned above, the reconstruction efficiency for the possible future $N \to \mu\pi$ analysis is probably too low, but at the same time the background rate is expected to be higher. Therefore it is possible that this analysis will reach a better sensitivity as the current analysis, but more research is needed to make more precise predictions about this.

Part IV:
Conclusion

4 Conclusion

In this thesis fast Monte Carlo samples are set up to study the LHCb sensitivity to heavy neutral leptons produced in B decays.

The simulated data fits very well to the full Monte Carlo data. A cut that was mistakenly applied in the full Monte Carlo data is found and its effect is studied. The calculated fiducial volume efficiencies in LHCb acceptance fit well to the fiducial volume efficiencies in LHCb acceptance that are calculated from full Monte Carlo data at the given HNL masses.

The data produced with RapidSim is used to calculate the fiducial volume efficiencies for many masses in the mass range between 1 and 6 GeV. Detailed plots of the behaviour of these efficiencies are now available. Additionally, it is calculated how many signal decays that include an HNL are expected to be in the fiducial volume region.

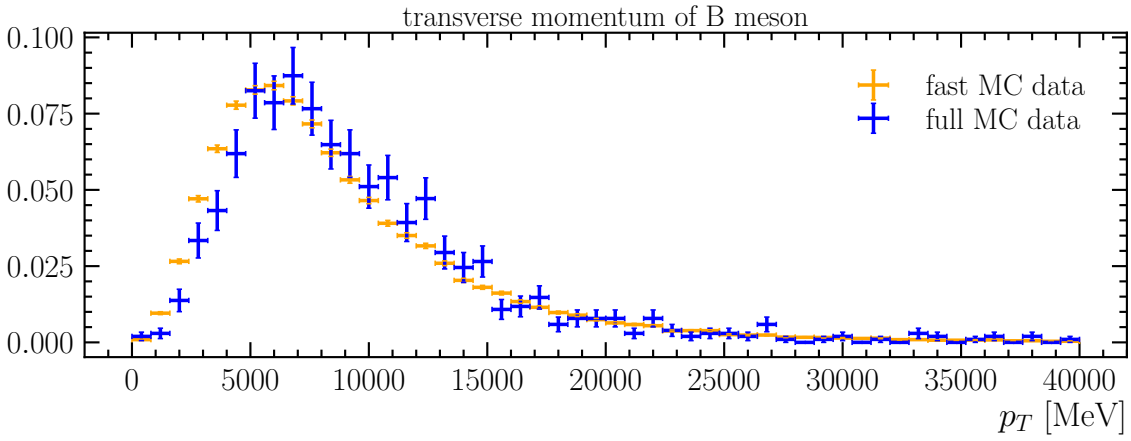
Furthermore, the fiducial volume efficiencies are used to calculate the expected sensitivity that this analysis will possibly reach. Although the sensitivity of this study might not be good enough to set a new world-best limit, a huge improvement in comparison to the previous HNL search at LHCb has been made. In addition to that, RapidSim is used to calculate the expected sensitivity for an HNL search in $N \to \mu^+\pi^-$, which is a possible future analysis. The results show that it is likely that this analysis will lead to better sensitivities as the current one. More research needs to be done in order to give more precise answers to this.

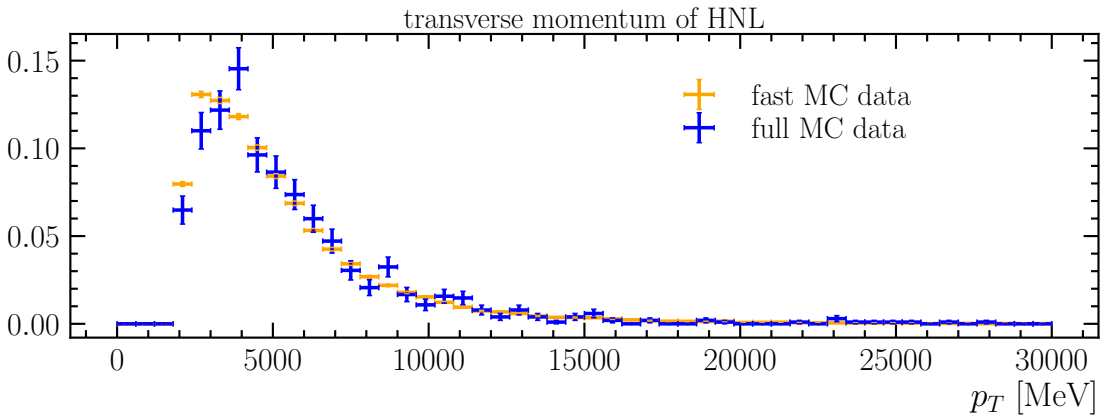
This thesis shows that RapidSim is a good tool that can also be used for possible future studies. In suitable cases one can avoid producing fully simulated Monte Carlo data which is often time-consuming by using RapidSim.

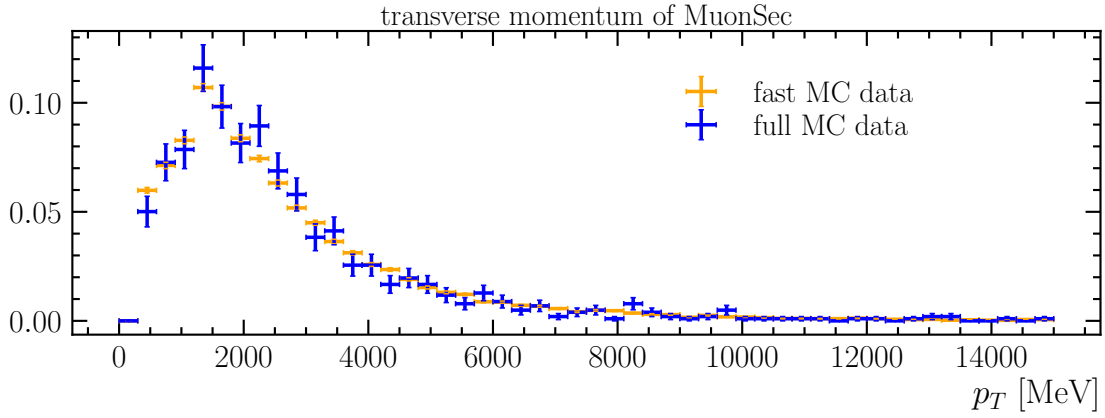
Part V:
Appendix

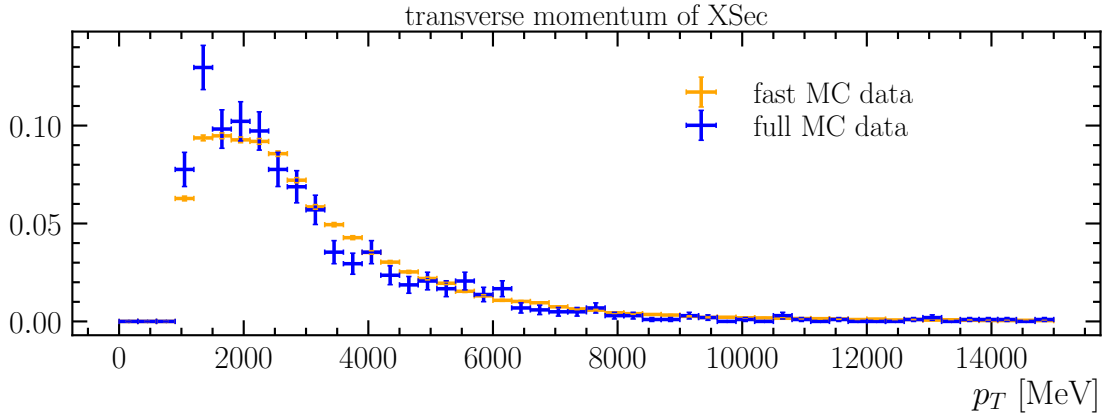
5 Appendix

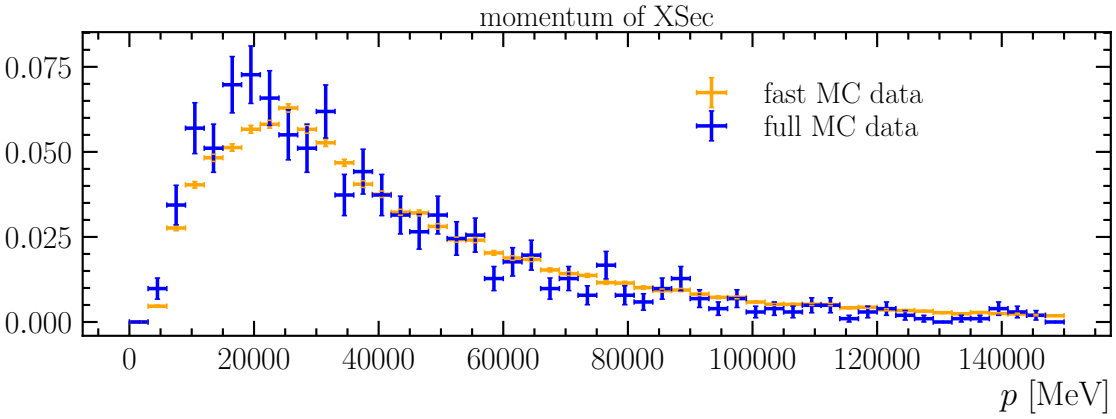
Plots from the comparison between full MC data and fast MC data for the decay $B^0 \to \pi^- \mu^+ (N \to \mu^+ \pi^-)$

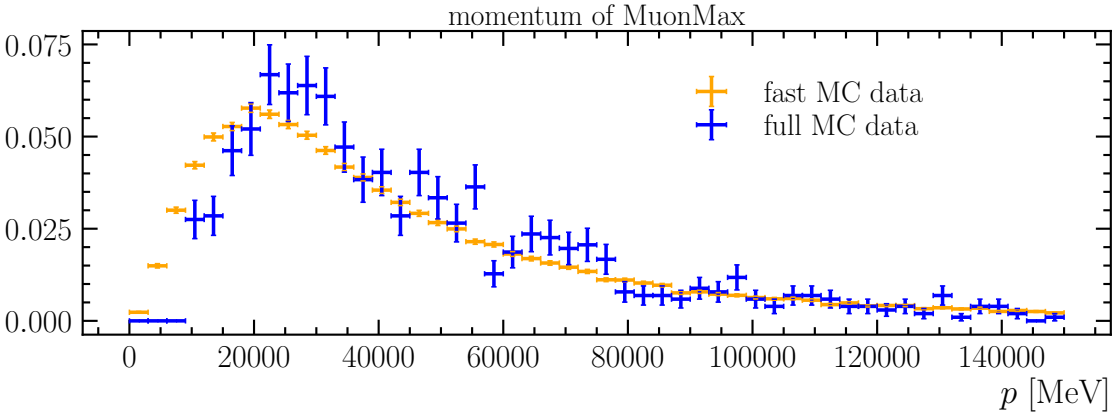


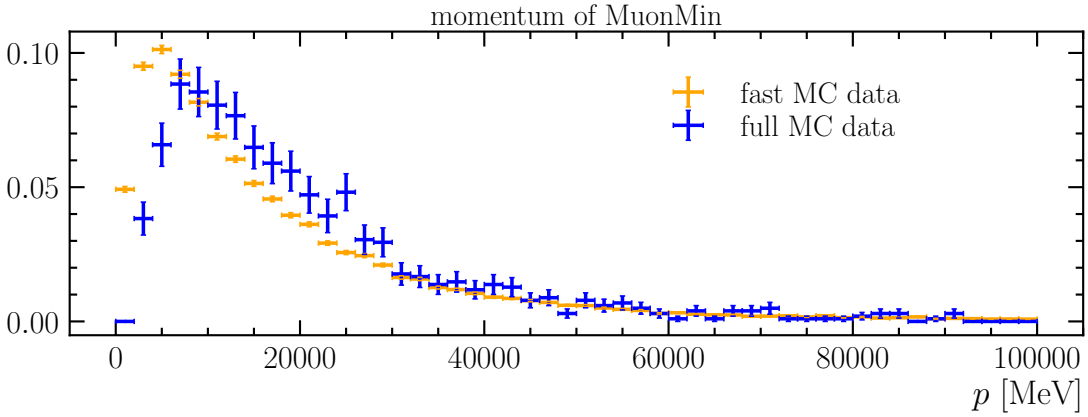


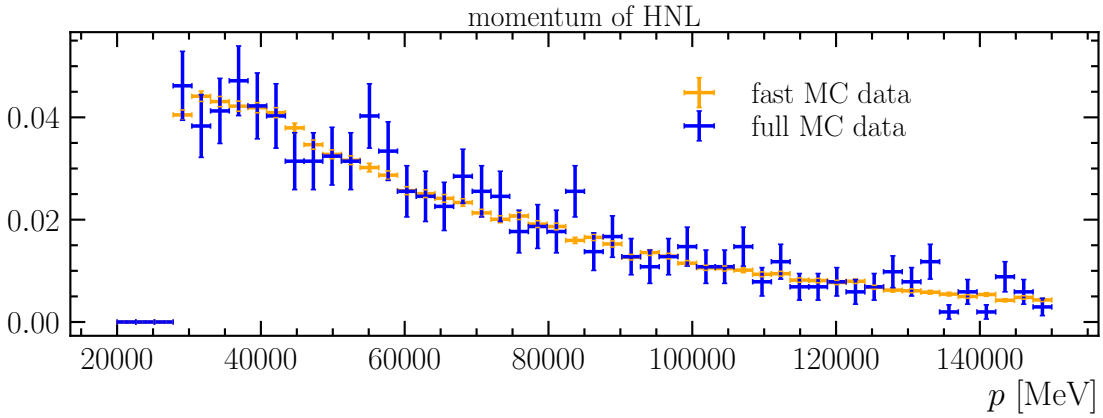

transverse momentum of MuonSec
fast MC data
full MC data
0.10
0.05
0.00
0
2000
4000
6000
8000
10000
12000
14000
p_T [MeV]

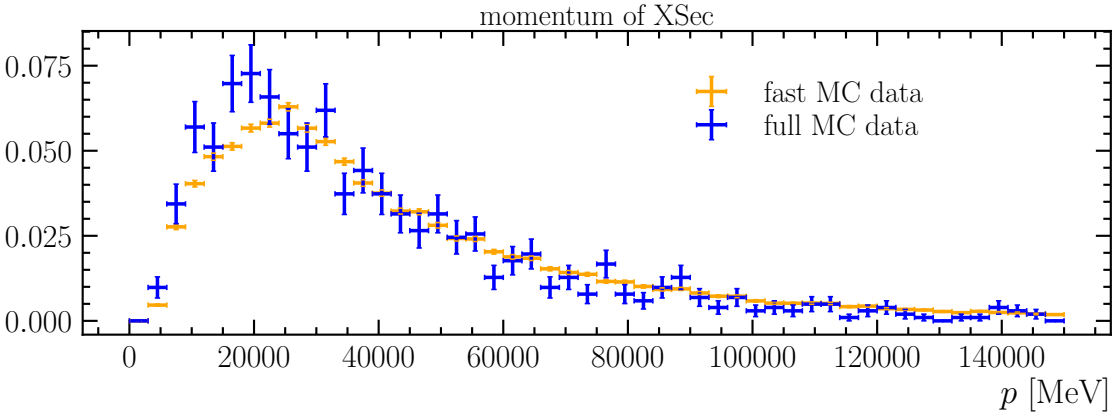

transverse momentum of XSec
fast MC data
full MC data
0.10
0.05
0.00
0
2000
4000
6000
8000
10000
12000
14000
p_T [MeV]

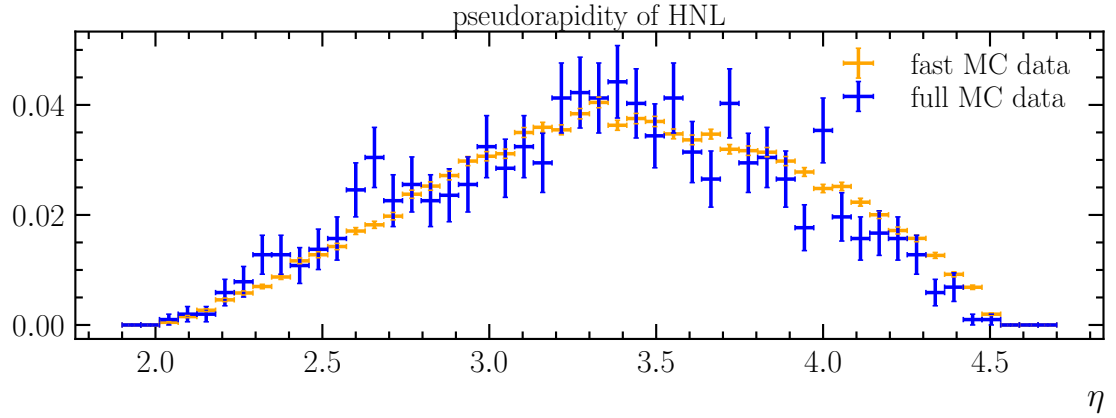

momentum of XSec
fast MC data
full MC data
0.075
0.050
0.025
0.000
0
20000
40000
60000
80000
100000
120000
140000
p [MeV]

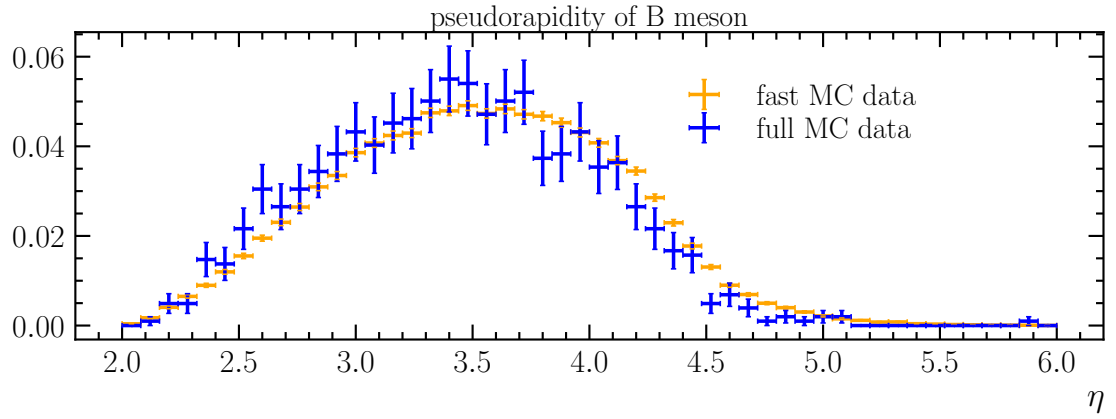

momentum of MuonMax
0.075
0.050
0.025
0.000
0
20000
40000
60000
80000
100000
120000
140000
p [MeV]
fast MC data
full MC data

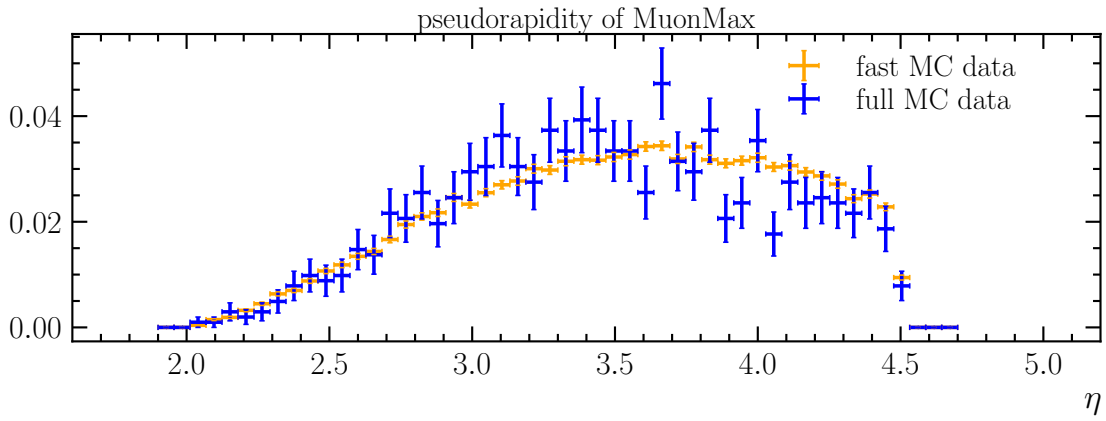

momentum of MuonMin
0.10
0.05
0.00
0
20000
40000
60000
80000
100000
p [MeV]
fast MC data
full MC data

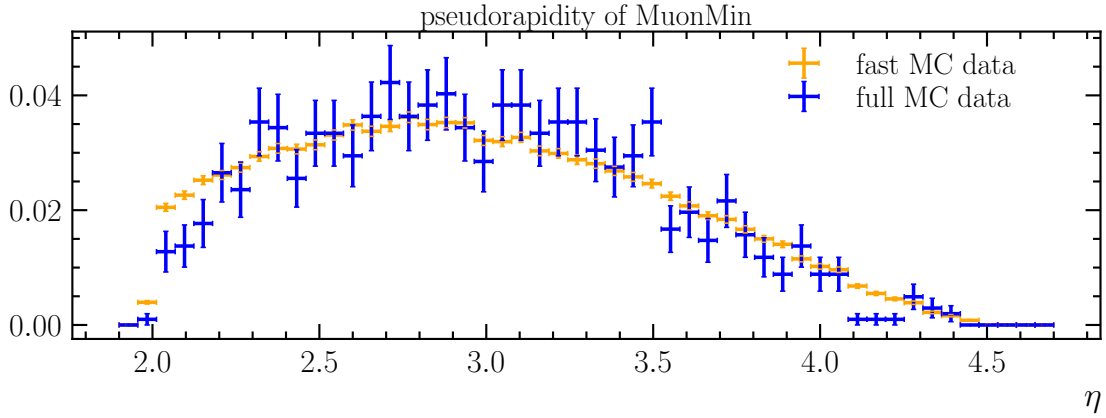

momentum of HNL
0.04
0.02
0.00
20000
40000
60000
80000
100000
120000
140000
p [MeV]
fast MC data
full MC data

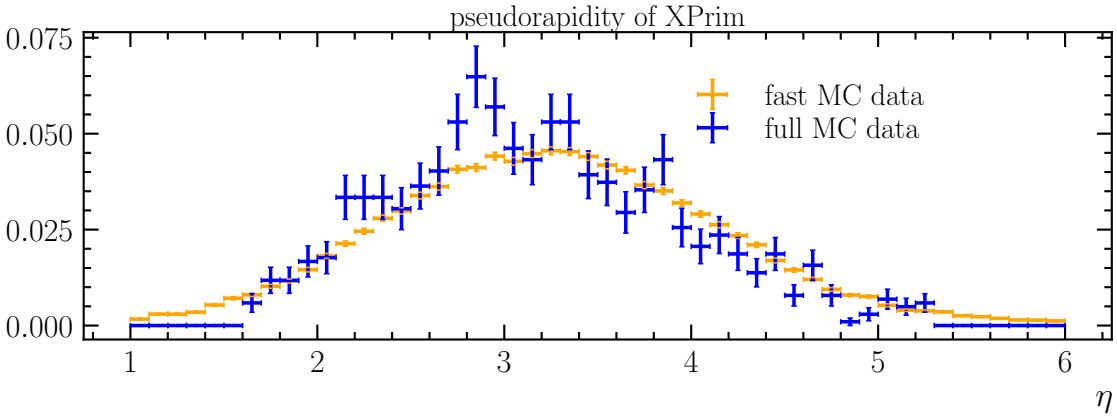

momentum of XSec
fast MC data
full MC data
0.075
0.050
0.025
0.000
0
20000
40000
60000
80000
100000
120000
140000
p [MeV]

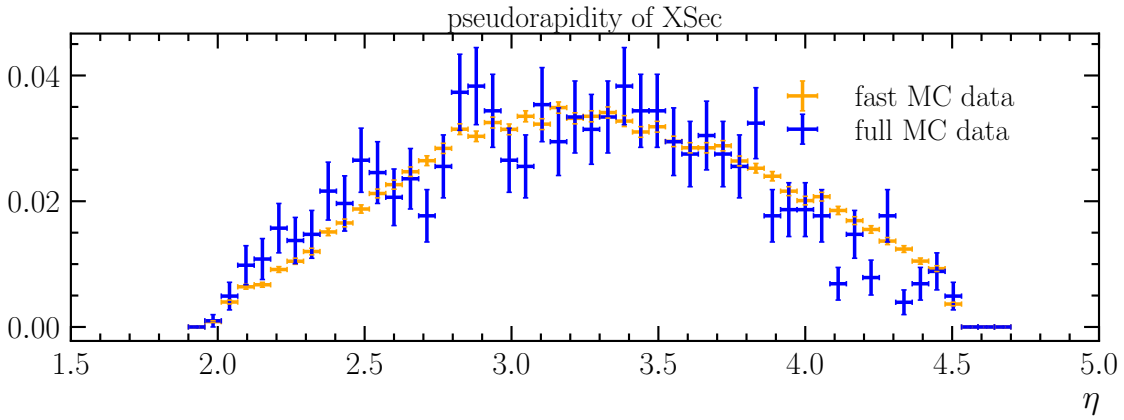

pseudorapidity of HNL
fast MC data
full MC data
0.04
0.02
0.00
2.0
2.5
3.0
3.5
4.0
4.5
η


pseudorapidity of B meson
fast MC data
full MC data
0.06
0.04
0.02
0.00
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
η


pseudorapidity of MuonMax
fast MC data
full MC data
0.04
0.02
0.00
2.0
2.5
3.0
3.5
4.0
4.5
5.0
η


pseudorapidity of MuonMin
fast MC data
full MC data
0.04
0.02
0.00
2.0
2.5
3.0
3.5
4.0
4.5
η


pseudorapidity of XPrim
fast MC data
full MC data
0.075
0.050
0.025
0.000
1
2
3
4
5
6
η


pseudorapidity of XSec
fast MC data
full MC data
0.04
0.02
0.00
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
η

Bibliography

[1] M. Aker, K. Altenmüller, M. Arenz, M. Babutzka, J. Barrett, S. Bauer, M. Beck, A. Beglarian, J. Behrens, T. Bergmann, and et al. Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN. *Physical Review Letters*, 123(22), Nov 2019.

[2] Bondarenko, K., Boyarsky, A., Gorbunov, D. et al. Phenomenology of GeV-scale heavy neutral leptons. *J. High Energ. Phys.*, 2018.

[3] A. Boyarsky, M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy. Sterile neutrino dark matter. *Progress in Particle and Nuclear Physics*, 104:1–45, 2019.

[4] G. Cowan. Rapidsim. `https://github.com/gcowan/RapidSim`, 2019.

[5] G. Cowan, D. Craik, and M. Needham. RapidSim: An application for the fast simulation of heavy-quark hadron decays. *Computer Physics Communications*, 214:239â246, May 2017.

[6] LHCb Starterkit. The LHCb data flow . 2015-2019.

[7] Mark Thomson. Modern particle physics.

[8] P: A. Zyla et al. Review of particle physics. 2020.

[9] Pauline Gagnon. Quantum Diaries: The Standard Model.

[10] Rolf Lindner. LHCb layout_2. LHCb schema_2. LHCb Collection. . Feb 2008.

[11] O. Steinkamp and on behalf of the LHCb Collaboration. LHCb upgrades. *Journal of Physics: Conference Series*, 1271:012010, jul 2019.

[12] The LHCb Collaboration. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, aug 2008.

6 Acknowledgements

First of all, I would like to thank Prof. Dr. Stephanie Hansmann-Menzemer for allowing me to work on this thesis in her group and also giving me the possibility to work at the institute in times of Corona. Huge thanks to Dr. Martino Borsato for always taking time to answer my questions and helping me and for creating an environment in which one feels free to ask all kinds of questions. Also thanks to Maurice Morgenthaler for taking time to explain the analysis and code to me, especially in the beginning of my work. I appreciate that I got the opportunity to get to know the nice working group of Stephanie and Ulli and the people at EPFL, who work with us on the analysis. I'm also grateful for every one of the "Feynfrauen" for going through the highs and lows of studying physics with me since day one. Special thanks to Thea for being my Zettelpartner and Erklärbär in almost everything. Thanks to Carlos Brandl and my family for being by my side.
Finally, I would like to thank the Heinrich-Böll-Stiftung for the new friends, new ideas and the financial support they gave me.

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

Rebecca Gartner

Heidelberg, den 17.08.2021