Long lived heavy neutrinos in $W^\pm \rightarrow \mu^\pm \mu^\pm$ jet decays

by
Valeriia Lukashenko
11623314
July 2019
60 EC
October 2018 - June 2019

Supervisors:
 dr. Wouter Hulsbergen
 examined
 Elena Dall’Occo, MSc.

Examiner:
 prof. dr. ir. Paul de Jong

bfys group
Nikhef
Abstract

Valeriia LUKASHENKO

*Long lived heavy neutrinos in $W^\pm \rightarrow \mu^\pm \mu^\pm \text{jet}$ decays*

Recently, it was discovered that neutrinos have mass. Massive neutrinos can be explained by introducing a heavy right-handed Majorana neutrino via the seesaw type I mechanism. In this thesis a search for a heavy long-lived Majorana neutrino $N_R$ in the LHCb experiment is presented using 2012 data from proton-proton collisions at $\sqrt{s} = 8$ TeV. If a Majorana neutrino exists, a lepton number conservation is violated. A lepton number violating decay $W \rightarrow \mu \mu \text{jet}$ with same charge muons is explored for heavy right-handed neutrinos with masses between 5 and 50 GeV and with a lifetime of 10 ps. To isolate a muon neutrino-heavy neutrino mixing parameter, the signal channel is normalized by a $W \rightarrow \mu \nu$ channel. An expected limit on the product of branching ratio for $N_R \rightarrow \mu \text{jet}$ with a mixing parameter between a muon neutrino and a heavy neutrino at CL = 95% is set and is an order of $10^{-4}$.

*Keywords:* neutrino, Majorana, light-heavy neutrino mixing, LHCb.
Acknowledgements

First, I would like to thank Wouter Hulsbergen for welcoming me as his master student and asking tricky questions during the process. Those questions allowed me to understand deeper both physics and technical details of this analysis. I also thank him for commenting on this thesis draft.

Also, I would like to thank my daily supervisor, Elena Dall’Occo for her everyday help. Her kind guidance through the analysis techniques helped me a lot during this year. Elena’s office was always open for me to stop by with any question or request I had. Elena’s strict attitude to research and useful comments have improved the quality of my results dramatically.

I am also grateful to Marcel Merk for his encouragement and help during the tough times. His positive attitude motivated me during the entire year as this research was conducted.

My gratitude goes to the entire b-fys group at Nikhef for a welcoming atmosphere. Because of this, I always could ask someone for help, if I could not find a solution myself.

At the end I would like to thank Paul de Jong for taking time to read this thesis and provide his feedback, which I hope will help to improve the quality of my research.
Contents

Acknowledgements v

1 Introduction 1
  1.1 Introduction and motivation .............................. 1
  1.2 Standard Model overview .................................. 3
  1.3 Neutrino oscillations ..................................... 6
  1.4 Neutrino mass ............................................. 7
    1.4.1 Mixing parameter .................................... 10
  1.5 LHCb dictionary ........................................... 11

2 LHCb detector 15
  2.1 Vertex Locator ............................................ 15
  2.2 Tracking stations ......................................... 16
  2.3 Magnet .................................................... 18
  2.4 Track reconstruction and track reconstruction efficiency .... 19
  2.5 Muon stations ............................................. 22
  2.6 Ring Imagining Cherenkov detector and particle identification ... 22
  2.7 Calorimeter system ..................................... 26
  2.8 Jet reconstruction ........................................ 26
    2.8.1 Jet input particle selection .......................... 28
    2.8.2 Jet reconstruction algorithm ........................ 28
  2.9 Trigger system ........................................... 29

3 Analysis Strategy 33

4 Samples 35
  4.1 Data Samples ............................................. 35
  4.2 Signal Monte Carlo ...................................... 35
  4.3 Background Monte Carlo ................................ 36

5 Reconstruction and selection 39
  5.1 Reconstruction ........................................... 39
  5.2 W candidate preselection .................................. 39
  5.3 Signal region ............................................. 42
  5.4 Impact parameter cut ..................................... 42
  5.5 Multivariate selection .................................... 42
  5.6 Control regions .......................................... 50
  5.7 Other selections .......................................... 51
  5.8 Background in the control regions ........................ 51
### Efficiency

6.1 Global event cut correction ....................................................... 59
6.2 Muon corrections ................................................................. 63
6.3 BDT corrections .................................................................. 63
6.4 Reconstruction and stripping, and selection efficiency .......... 67

### Uncertainties

7.1 Corrections ........................................................................ 71
7.2 Momentum calibration ......................................................... 71
7.3 Jet energy scale, jet energy resolution and jet identification .... 71
7.4 Total systematic uncertainty ................................................... 72

### Results and discussion

8.1 Rolke limit estimation ........................................................... 75
8.2 Normalization channel yields ............................................... 76
8.3 Expected upper limit ............................................................ 79

### Conclusion and outlook

9.1 $\mu_N$ multiplicity ............................................................... 83

B $\mu_W$ and $\mu_N$ kinematics .................................................... 85

C Calibration of momentum ....................................................... 89
  C.1 Curvature bias momentum calibration ............................... 90

D Efficiency uncertainty .............................................................. 95

E Normalization channel fits ....................................................... 99

Bibliography .............................................................................. 103
# List of Figures

1.1 Feynman diagram for $W^\pm \rightarrow l^\pm l^\pm q\bar{q}$ .......................... 1
1.2 Limits on the muon-heavy neutrino mixing ............................ 2
1.3 Standard Model particle content ........................................ 3
1.4 Neutrino oscillation probability ........................................ 8
1.5 Decay schematics ....................................................... 13

2.1 LHCb detector schematics .............................................. 16
2.2 VELO position schematics ................................................ 17
2.3 VELO strips schematics .................................................. 17
2.4 Tracking stations hit resolution ......................................... 19
2.5 LHCb magnetic field ..................................................... 19
2.6 LHCb tracks ............................................................. 20
2.7 Track reconstruction efficiency ......................................... 21
2.8 Momentum and mass reconstruction .................................... 22
2.9 Muon stations schematics ................................................. 23
2.10 Muon chambers time resolution ........................................ 23
2.11 RICH1 and RICH2 schematics ......................................... 24
2.12 Particle ID in RICH ..................................................... 24
2.13 RICH1 and RICH2 angular resolution .................................. 25
2.14 Identification efficiency RICH ......................................... 25
2.15 Calorimeter system performance ....................................... 27
2.16 LHCb trigger system schematics for 2012 ............................ 29
2.17 L0 efficiency ............................................................ 31
2.18 HLT1 efficiency ......................................................... 32
2.19 HLT2 efficiency ......................................................... 32

5.1 Muons $p_T$ ............................................................... 40
5.2 Muons impact parameter .................................................. 42
5.3 An example of a decision tree .......................................... 43
5.4 ROC curves ............................................................. 45
5.5 Muons IT5 ............................................................... 47
5.6 Angular distance between $\mu_N$ and jet ............................. 48
5.7 Background and signal BDT classifiers distributions ................ 49
5.8 BDT efficiency .......................................................... 50
5.9 Muons $p_T$ in control regions ......................................... 52
5.10 BDT classifiers in control regions ..................................... 53
5.11 Isolation variables before selection .................................... 54
5.12 Muons IT5 in control regions ......................................... 55
5.13 $\mu_N$ IT1 and IT3 in control regions ................................ 56
5.14 Dimuon mass and missing $p_T$ in control regions .................. 57

6.1 nSPD hits as a function of nPVs ...................................... 60
6.2 Events per number of primary vertices in data and MC ............ 60
# List of Tables

2.1 Resolution and efficiency of the LHCb detector .................................. 18
2.2 LO categories ....................................................................................... 30
4.1 Data samples ....................................................................................... 35
4.2 Signal Monte Carlo samples ................................................................. 37
4.3 Background Monte Carlo samples ......................................................... 38
5.1 Stripping and trigger lines ................................................................. 40
5.2 Candidate preselection ........................................................................ 41
5.3 Control and signal regions ................................................................. 50
5.4 $Z \rightarrow \mu^+\mu^-$ and $Y(1S) \rightarrow \mu^+\mu^-$ selection ....................... 51
6.1 nSPD hits efficiency in $Z \rightarrow \mu^+\mu^-$ and new nSPD hits cut .......... 61
6.2 nSPD hits efficiency for signal and normalization channels .......... 62
6.3 Muon correction factors ..................................................................... 64
6.4 BDT classifier correction factors ....................................................... 68
6.5 Reconstruction and stripping efficiency ........................................... 69
6.6 Total efficiency ................................................................................... 69
7.1 Relative systematic uncertainties ....................................................... 73
8.1 QCD selection ..................................................................................... 77
8.2 Measured W yields ............................................................................. 79
8.3 Upper limit total table ........................................................................ 80
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDT</td>
<td>Boosted Decision Trees</td>
</tr>
<tr>
<td>ECAL</td>
<td>Electromagnetic CALorimeter</td>
</tr>
<tr>
<td>GEC</td>
<td>Global Event Cut</td>
</tr>
<tr>
<td>HCAL</td>
<td>Hadronic CALorimeter</td>
</tr>
<tr>
<td>HLT</td>
<td>High Level Trigger</td>
</tr>
<tr>
<td>LHC</td>
<td>Large Hadronic Collider</td>
</tr>
<tr>
<td>LHCb</td>
<td>Large Hadronic Collider beauty experiment</td>
</tr>
<tr>
<td>MD</td>
<td>Magnet Down</td>
</tr>
<tr>
<td>MU</td>
<td>Magnet Up</td>
</tr>
<tr>
<td>L0</td>
<td>Level-0 trigger</td>
</tr>
<tr>
<td>LO</td>
<td>Leading Order</td>
</tr>
<tr>
<td>PS</td>
<td>Pre-Shower detector</td>
</tr>
<tr>
<td>PV</td>
<td>Primary Vertex</td>
</tr>
<tr>
<td>QCD</td>
<td>Quantum Chromo-Dynamics</td>
</tr>
<tr>
<td>QFT</td>
<td>Quantum Field Theory</td>
</tr>
<tr>
<td>RICH</td>
<td>Ring Imagening CHERenkov counter</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver Operating Characteristic</td>
</tr>
<tr>
<td>SPD</td>
<td>Scintillating Pad Detector</td>
</tr>
<tr>
<td>TCK</td>
<td>Trigger Configuration Key</td>
</tr>
<tr>
<td>TIS</td>
<td>Trigger Independent of Signal</td>
</tr>
<tr>
<td>TOS</td>
<td>Trigger On Signal</td>
</tr>
<tr>
<td>TT</td>
<td>Tracking Turicensis</td>
</tr>
<tr>
<td>VELO</td>
<td>VERTex LOcator detector</td>
</tr>
<tr>
<td>n-</td>
<td>number</td>
</tr>
<tr>
<td>-s</td>
<td>in plural</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Introduction and motivation

The Standard Model does not include massive neutrinos. Nonetheless, a discovery of neutrino oscillations proved that neutrinos have nonzero mass. Two possible neutrino mass terms can be included in the Standard Model Lagrangian: Dirac mass term or Majorana mass term [1]. The Dirac particle is an ordinary Standard Model fermion. Dirac particle and anti-particle are not equivalent particles. Opposite, to the other Standard Model fermions, neutrino is neutral. As a consequence, one can not experimentally distinguish a neutrino and an anti-neutrino. For a Majorana particle, particle and antiparticle are the same. Therefore, neutrino can also be described as a Majorana particle. A Majorana particle interaction with other Standard model particles is different, specifically, it violates the lepton number conservation.

The smallness of neutrino masses can be explained by the seesaw mechanism. In the seesaw formalism two neutrino mass eigenstates are introduced using both Dirac and Majorana type mass terms in Lagrangian. One of the neutrino mass eigenstates, $N_R$, is a heavy almost right handed state and the other one, $\nu_L$, is almost left handed and very light. If both mass states exist, both lepton number conserved and lepton number violating decays should be observed. One of the possible lepton number violating decays is $W^\pm \rightarrow l^\pm l^\pm q\bar{q}'$ mediated by a heavy $N_R$, see Fig. 1.1.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{feynman_diagram.png}
\caption{Feynman diagram for $W^\pm \rightarrow l^\pm l^\pm q\bar{q}'$}
\end{figure}

The LHCb collaboration presented a search for Majorana neutrinos in the mass range 250 - 5000 MeV with lifetime from 0 to 1000 ps in lepton number violated $B^- \rightarrow \pi^+ \mu^- \mu^-$ decays at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV [3]. Previous searches by the LHCb collaboration explored other b-meson decays at $\sqrt{s} = 7$ TeV: $B^+ \rightarrow K^- \mu^+ \mu^+$, $B^+ \rightarrow \pi^- \mu^+ \mu^+$, $B^- \rightarrow \pi^- \mu^- \mu^-$, $B^- \rightarrow D^{(*)+}_s \mu^- \mu^-$, $B^- \rightarrow D^0 \pi^+ \mu^- \mu^-$ that can be mediated via a prompt Majorana neutrino in the $260 < m_N < 5000$ MeV mass range [4, 5].
Some other recent collider searches are listed below. A prompt heavy Majorana neutrino in the model with a heavy \( W_R \)\(^1\) was studied by the ATLAS collaboration in \( W_R \rightarrow ee(\mu\mu)\) jets decays for the \( 0.1 < m_N < 2.7 \) TeV mass range at \( \sqrt{s} = 7 \) TeV [6] and for the \( 100 < m_N < 500 \) GeV mass range at \( \sqrt{s} = 8 \) TeV [7]\(^2\). The same model was probed in the \( 100 < m_N < 500 \) GeV mass range at \( \sqrt{s} = 7 \) TeV in decays with same charged leptons [8]. Limits on a light-heavy neutrino mixing parameter in lepton number violated and lepton number conserved \( W \rightarrow \mu\mu\nu \) and \( W \rightarrow e\mu\nu \) decays were set by the ATLAS collaboration at 13 TeV, where a massive neutrino was assumed to be prompt or long-lived with \( 4.5 < m_N < 50 \) GeV [9].

The CMS collaboration also participated in searches for a prompt Majorana heavy neutrino. Limits on a light-heavy neutrino mixing parameter for a prompt Majorana neutrino in \( W \rightarrow ee(\mu\mu)\) jets decays at \( \sqrt{s} = 7 \) TeV (\( m_N > 90 \) GeV) and at \( \sqrt{s} = 8 \) TeV (\( 40 < m_N < 500 \) GeV) are set in [10–12]. The CMS collaboration updated a limit on a light-heavy neutrino mixing parameter for a prompt \( 20 < m_N < 1600 \) GeV heavy neutrino from lepton number violated \( W \rightarrow ee(\mu\mu)\) jet decays at \( \sqrt{s} = 13 \) TeV in [13]. In [14] an upper limit on a light-heavy neutrino mixing parameters for a Majorana neutrino with mass range \( 0.001 < m_N < 1.2 \) TeV in decays with 3 leptons (muons or electrons) at \( \sqrt{s} = 13 \) TeV is set. Two prompt Majorana neutrino searches at \( \sqrt{s} = 13 \) TeV in decays to two \( \tau \) and two jets are presented in [13, 15]. A composite Majorana neutrino model search at \( \sqrt{s} = 13 \) TeV was also done by the CMS collaboration in [16].

This thesis concentrates on an on-shell \( W^\pm \rightarrow \mu^\pm\mu^\mp q\bar{q}' \) mediated with a long-lived \( N_R \). A prompt case has been studied at LHCb in [17]. This thesis closely follows the prompt case study. The current upper limit on a mixing parameter \( |B_{\mu N}|^2 \) between second family light and heavy neutrinos as a function of \( N_R \) mass, \( M_N \), is shown in Fig. 1.2 [2]. The grey region is excluded with previous studies. LHCb detector is sensitive to the decays of heavy neutrino in the mass range of 5-50 GeV.

\(^1\)\( W_R \) is a new right-handed heavy W-like boson. \( m_{W_R} \sim O(1 \) TeV\).
\(^2\) Also a new heavy \( Z' \) boson, that can interact with a heavy neutrino was investigated.
1.2 Standard Model overview

The Standard Model is a modern particle physics theoretical model, that describes matter content, electromagnetic, weak and strong interactions and shows a great consistency with data [18, 19]. Standard Model obeys local gauge invariance and Lorentz invariance principles. Discrete symmetries: parity P, charge conjugate C, time T, as well as, CP are not conserved in the Standard Model. However, CPT is conserved. Without neutrino masses and neutrino mixing parameters Standard Model has 19 free parameters: 9 masses of fermions, 3 strength coupling, Higgs mass and vacuum expectation value, 4 CKM parameters and a QCD vacuum angle.

Matter in the Standard Model consists of spin 1/2 fermions: quarks and leptons, divided into the three generations, see Fig. 1.3. Up and down quark, electron and electron neutrino are grouped in the first generation. Charm quark, strange quark, muon and muon neutrino are the second generation. Top quark, bottom quark, τ and ντ neutrino ended up in the third generation. Quarks hadronize as 3-quarks baryons or 2-quarks mesons. Top quarks can not hadronize, due to the short lifetime.

![Standard Model of Elementary Particles](image)

**Figure 1.3:** The Standard Model particle content. In the public domain, world averages from [20].

Standard Model quarks and charged leptons obey Dirac equation and are described by the Dirac spinor.
Gauge bosons, i.e. $W^{\pm}$ bosons, $Z$ boson, gluons and photon, are mediators of interactions between fermions.

Massless spin one photon field transfers electromagnetic interactions in the Standard Model. The electromagnetic interaction conserves the hypercharge $Y$:

$$Y = 2Q - I_3$$

where $Y$ is a hypercharge; $Q$ is an electric charge; $I_3$ is $z$-component of the isospin.

The weak interaction is carried by $W^{\pm}$ and $Z$ bosons. The weak interaction is the only known source of the matter/anti-matter asymmetry in the Standard Model, since $W$ bosons interact only with left-handed (chirality) particle states (right-handed antiparticle states).

Chirality is a Lorentz invariant quantum property of a particle. It is connected with a phase of a particle wave function. Any particle can be represented as a sum of left- and right-chiral states:

$$\psi = \psi_L + \psi_R$$

$$\psi_L = \frac{1}{2}(1_4 - \gamma_5)\psi$$

$$\psi_R = \frac{1}{2}(1_4 + \gamma_5)\psi,$$

where $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$, $\gamma_i$ is a gamma matrix and $1_4$ is a $4 \times 4$ identity matrix.

In the ultra-relativistic limit chirality converges with helicity. Helicity is defined as projection of a particle spin on the momentum direction:

$$h = \frac{2\vec{s} \cdot \vec{p}}{|\vec{p}|}$$

where $\vec{s}$ is spin and $\vec{p}$ is momentum of the particle, normalized to the absolute momentum value $|\vec{p}|$. Right-handed helicity is when spin and momentum of the particle are aligned. Left-handed helicity is when they are oppositely aligned. Oppositely, to the chirality, helicity is not Lorentz invariant. Each chirality state can be described as a sum of helicity states and vice versa. For massless particles, helicity and chirality are always the same.

$W$ bosons interaction is also called the flavour changing current, because a particle can change its flavour through it. Particle flavour is particle type, see Fig. 1.3. Quarks have 6 flavours: up, down, charm, strange, top and bottom. Leptons also have 6 flavours: electron, electron neutrino, muon, muon neutrino, tau, tau neutrino. For example, an up quark can become a down quark by emitting a $W^+$ boson. $W$ bosons carry electrical charge, therefore, another name for the $W$ mediation is a charged weak current.

$Z$ boson mediates a neutral weak current. $Z$ boson does not change particle flavour. It can interfere with a photon, an electromagnetic force carrier.

Due to many similarities, electromagnetic and weak interactions can be unified as electroweak interaction.

---

$^3$Another possible source is the strong CP violation, which is not yet observed. The current limit is $\theta < 10^{-10}$ [20].
The strong Model overview

The strong interaction is mediated by massless gluons. Strong interaction conserves color charge, that is mediated by gluons. There are 3 colour charges and 3 anti-charges: red (anti-red), green (anti-green) and blue (anti-blue). Although, hadrons are colourless, gluons are colourful and change quarks colour charge. There are in total 8 different variations of gluons. Because of the gluon self-interaction there is a high ambiguity in the theoretical quantum chromodynamic (QCD) predictions.

Together, the Standard Model is a SU(3)_c × SU(2)_L × U(1)_Y theory, where electromagnetic interaction is described by U(1)_Y, weak by SU(2)_L and strong by SU(3)_c. The only scalar spin 0 particle in the Standard Model is the Higgs boson. It plays a crucial role in the emergence of bosons and fermions masses.

The simplified Standard Model Lagrangian for one particle is:

\[ \mathcal{L} = i\bar{\psi}_{L,i,k} \gamma^\mu D_\mu \psi_{L,i,k} + i\bar{\psi}_{R,j,k} \gamma^\mu D_\mu \psi_{R,j,k} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \gamma_i \psi_{L,i,k} \phi \psi_{R,j,k}^\dagger + (D^\mu \phi)^\dagger (D^\mu \phi) - V(\phi) + \text{h.c.} \]  

(1.7)

where \( \psi \) is a Dirac spinor; \( L/R \) denotes left-/right-handed components; \( \bar{\psi} \) is a Dirac adjoint spinor; \( \gamma^\mu \) are Dirac matrices; \( D \) is a covariant derivative, written separately for left- and right-handed components; \( F_{\mu\nu} \) is a field tensor; \( \gamma \) is a Yukawa coupling, that defines a coupling between Higgs boson and a fermion; \( \dagger \) and h.c. denote Hermitian conjugate; \( \phi \) is a scalar field; \( V(\phi) \) is a scalar field potential; \( \mu, \nu \) are Einstein indices; \( i, j \) are flavour indices.

The green part of (1.7) describes fermion kinematics (for a massless fermion) and interaction with bosons, except Higgs. Color indices \( k \) and \( n \) are equal (\( k=n \)) for all interaction, except for the color-changing strong interaction. Flavour indices \( i, j \) are equal, unless the interaction is the charged-current weak interaction. In the explicit equation for the \( D_{L,R} \) the strong generator \( \lambda_{kn} \) describes the color change and the second term in the left-handed part contains the charged-current weak interaction:

\[ D_L = \partial_\mu - ig \frac{Y}{2} a_\mu - ig_W \tau \cdot b_\mu - ig_s \frac{\lambda_{kn}}{2} \cdot G^n_\mu \]  

(1.8)

\[ D_R = \partial_\mu - ig \frac{Y}{2} a_\mu - ig_s \frac{\lambda_{kn}}{2} \cdot G^n_\mu \]  

(1.9)

\( g \) is the electromagnetic interaction coupling constant; \( Y \) is the hypercharge; \( a_\mu \) is a gauge field corresponding to \( U(1)_Y; g_W \) is the weak interaction coupling constant; \( \tau = (\tau_1, \tau_2, \tau_3) \) are Pauli matrices; \( b = (b_1, b_2, b_3) \) are gauge fields corresponding to \( SU(2)_L; g_S \) is the strong coupling constant; \( \lambda^a_{kn} = (\lambda^1_{kn}, \lambda^2_{kn}, \lambda^3_{kn}) \) are Gell-Mann matrices; \( G^n_\mu = (G^n_1, G^n_2, G^n_3) \) are gluons (SU(3)_c) fields; \( \mu \) and \( k \) are color indices.

The violet part of (1.7) is a kinematic term for a free massless boson field, except Higgs. The field tensor \( F_{\mu\nu} \) for the \( a_\mu \) field:

\[ F_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu \]  

(1.10)

---

4SU(N) is a unitary Lie group of degree N, that representations are N×N unitary matrices. SU(N) is a special unitary Lie group of degree N, that representations are N×N unitary matrices M with \( det(M) = 1 \).

5\( \bar{\psi} = (\psi^\dagger, \gamma^\mu) \).

6Note, this is not a “canonical” way of writing the Lagrangian.

7Photon is a combination of \( a_\mu \) and \( b_3 \) in the electroweak theory.

8The \( W^\pm \) bosons are combination of \( b_1 \) and \( b_2 \). Z boson is a combination or \( b_3 \) and \( a_\mu \).
Chapter 1. Introduction

The Field tensor for the $b_\mu$ fields:

$$F_{\mu\nu} = \partial_\mu b_\nu - \partial_\nu b_\mu - 2g_W (b_\mu \times b_\nu) \quad (1.11)$$

The Field tensor for the $G_\mu$ fields:

$$F^{a}_{\mu\nu} = \partial_\mu G^{a}_\nu - \partial_\nu G^{a}_\mu - 2g_s f_{bc} (G^{b}_\mu G^{c}_\nu) \quad (1.12)$$

where $a$ is a label of gluon field (8 in total); $f_{bc}$ is a color factor, that defines strength of attraction/repulsion in the (anti-)quarks-(anti-)quarks interaction depending on the color quarks color states.

The orange part in (1.7) is the Yukawa coupling term, which describes fermion mass. Yukawa couplings are proportional to the fermion mass. The scalar field $\phi$ in this term is the Higgs field.

The blue part in (1.7) describes the Higgs field physics. The first term describes its kinematics in interaction of bosons with the Higgs field, therefore, boson mass. And the second term is the Higgs potential.

1.3 Neutrino oscillations

Neutrinos were assumed to be massless and left-handed in the Standard Model, where other leptons are massive and can be both left- and right-handed. Neutrinos have no electrical or colour charge and can interact in the Standard Model only through the weak force, making neutrino detection challenging. In the collider experiments the typical signature of a neutrino is missing energy in the event.

Nowadays, it is known, that neutrinos oscillate from one flavour state to another flavour state. The first evidence for atmospheric neutrinos oscillation was announced in 1998 by Super-Kamiokande collaboration [21]. Solar neutrino oscillations were discovered in 2001 by Sudbury neutrino observatory, see [22]. These two discoveries lead to the 2015 Noble Prize in Physics. Further experiments, proved the existence of the neutrino oscillations in the neutrino beam experiments [23–25]. Neutrino oscillations can be explained only if neutrinos have mass states, i.e. mass.

Neutrinos can be described in terms of flavour states: electron, muon and $\tau$, or in terms of mass states: first, second and third. These states mix through the unitary matrix, called Pontecorvo-Maki-Nakagawa-Sakata matrix, which matches flavour neutrino states to the combination of the mass neutrino states and describes neutrino oscillations [20]:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = V_{PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}$$

(1.13)

where $\nu_{e,\mu,\tau}$ are flavour states of neutrino; $\nu_{1,2,3}$ are mass states of neutrino with masses $m_1, m_2, m_3$; $V_{PMNS}$ is Pontecorvo-Maki-Nakagawa-Sakata matrix or PMNS matrix.

$$V_{PMNS} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix}$$

(1.14)

where $U_{fi}$ is a mixing parameter between flavour $f$ and mass state $i$. 
1.4. Neutrino mass

Assuming the Dirac nature of a neutrino, the PMNS matrix can be parameterized in terms of the mixing angles between neutrino states:

\[ V_{PMNS} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23}s_{13} + s_{12}s_{23}e^{i\delta} & s_{23}s_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \]  

(1.15)

where \( c_{ij} = \cos(\theta_{ij}) \) is a cosine of the mixing angle between \( i \) and \( j \) mass state; \( s_{ij} = \sin(\theta_{ij}) \) is a sine of the mixing angle between \( i \) and \( j \) mass state; \( \delta \) is a Dirac CP violation phase.

Mixing angles can be expressed in terms of the elements of (1.14):

\[
\begin{align*}
c_{12}^2 &= \frac{|U_{11}|^2}{1 - |U_{13}|^2} \\
c_{23}^2 &= \frac{|U_{23}|^2}{1 - |U_{33}|^2} \\
s_{12}^2 &= \frac{|U_{12}|^2}{1 - |U_{33}|^2} \\
s_{23}^2 &= \frac{|U_{33}|^2}{1 - |U_{33}|^2} \\
s_{13}^2 &= |U_{13}|^2 \\
\end{align*}
\]

(1.16) to (1.21)

Neutrino oscillation probabilities are \( \theta_{ij} \) and \( \Delta m_{ij}^2 = (m_i^2 - m_j^2) \) dependent. The most general probability of neutrino/anti-neutrino oscillation in the vacuum can be written in the form:

\[
P(v_l \rightarrow v_{l'}) = \sum_j |U_{lj}|^2 |U_{lj'}|^2 + 2 \sum_{j > k} |U_{lj}|U_{lj}^*U_{lk}U_{lk}^* \cos \left( \frac{\Delta m_{jk}^2 L}{2p} - \phi_{l',l;j,k} \right) \\
P(\bar{v}_l \rightarrow \bar{v}_{l'}) = \sum_j |U_{lj}|^2 |U_{lj'}|^2 + 2 \sum_{j > k} |U_{lj}|U_{lj}^*U_{lk}U_{lk}^* \cos \left( \frac{\Delta m_{jk}^2 L}{2p} + \phi_{l',l;j,k} \right)
\]

(1.22) to (1.23)

where \( l, l' \) are flavour indices; \( L \) is a length of the neutrino path; \( p = (p_j + p_k)/2 \) is 3-momentum amplitude, which under assumption of the negligible neutrino masses is equal to the neutrino energy; \( \phi_{l',l;j,k} = \arg(U_{lj}|U_{lj}^*U_{lk}U_{lk}^*|) \)

In reality of course, neutrinos also pass some matter, which makes the above equations more complicated. In Fig. 1.4 one can see how the neutrino oscillation probability in matter and vacuum changes with \( E/\Delta m_{13}^2 \).

1.4 Neutrino mass

The current upper limit on the light neutrino mass from tritium decay is \( m < 2\text{eV} \) [20]. Standard Model has to be adjusted accordingly to incorporate neutrino masses and has to be extended with neutrino mass term. As was mentioned above, neutrinos can be described as Dirac particle and as a Majorana particle. Therefore, the Standard Model has to be extended with a Dirac-type mass term or with a Majorana-type mass term.
If it is a Dirac term, than it is similar to the other fermion mass terms, see orange term in 1.7. However, since, the current upper limit on the neutrino mass is as low as 2 eV, the neutrino Yukawa coupling must be extra tiny: $< 10^{-12}$ [27]. Although, it is a possible scenario, it is unclear, why such a little parameter would be allowed in the Standard Model or in other words, this parameter would be “unnatural” if it is true.

Let’s assume one has a fermion:

$$\psi = \psi_L + \psi_R \tag{1.24}$$

where $\psi_L$ is a left-handed and $\psi_R$ is a right-handed components.

Charge conjugation operator changes particle to anti-particle, i.e. flips particles charges:

$$\psi^C = C\bar{\psi}^T = i\gamma_0\gamma_2\bar{\psi}^T, \tag{1.25}$$

where $\phi^C$ is an antiparticle spinor; $C$ is a charge conjugation operator; $T$ denotes a transpose operation; $\gamma_0$ and $\gamma_2$ are Dirac matrices:

$$\gamma_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.26}$$

$$\gamma_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}. \tag{1.27}$$

Charge conjugation changes left-handness to the right-handness and vice verse.
Dirac particle obeys the Dirac equation:
\[(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{1.28}\]
where \(\gamma^\mu\) are complex matrices, that satisfy the next condition:
\[\{\gamma^\nu, \gamma^\mu\} = 4\eta^{\nu\mu}, \tag{1.29}\]
where \(\{A, B\} = AB + BA\) is an anti-commutation relation; \(\eta^{\nu\mu}\) is Minkowski metric.

To satisfy (1.28) the spinor \(\psi\) has to be complex. A charge conjugate of the Dirac complex field is not equal to the original Dirac field:
\[\psi^C \neq \psi. \tag{1.30}\]

Therefore, the particle is not same as an antiparticle for a Dirac fermion.

If the field is real, then its charge conjugate field is the same as its original field.

The real field satisfies the Majorana equation [28]:
\[(i\tilde{\gamma}^\mu \partial_\mu - m)\psi = 0, \tag{1.31}\]
where \(\tilde{\gamma}^\mu\) are real \(4 \times 4\) Majorana matrices (also satisfy (1.29)) and \(\psi\) is a real field.

Majorana particle satisfies (1.31) and is its own antiparticle:
\[\psi^C = \psi. \tag{1.32}\]

For a Dirac particle Lagrangian mass term looks like:
\[-L^m_D = m_D^f \bar{\psi} \psi = m_D^f (\bar{\psi}_L + \bar{\psi}_R)(\psi_L + \psi_R) = m_D^f (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) \tag{1.33}\]
where \(m_D^f\) is a Dirac fermion mass.

However, for a Majorana particle \(\psi_R = C\bar{\psi}_L^T\) and \(\bar{\psi}_R = \psi_L^T C\):
\[-L^m_M = \frac{m_M^f}{2} \bar{\psi} \psi = \frac{m_M^f}{2} (\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R) = \frac{m_M^f}{2} (\psi_L^T C \psi_L + \bar{\psi}_L C \bar{\psi}_L^T) \tag{1.34}\]
where \(m_M^f\) is Majorana fermion mass and \(1/2\) comes to account for the Hermitian conjugate. A similar Majorana mass term can be written using the right-handed components.

The most general neutrino mass term [29]:
\[-L_{mass} = m_D (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) + \frac{m_L^M}{2} (\psi_L^T C \psi_L + \bar{\psi}_L C \bar{\psi}_L^T) + \frac{m_R^M}{2} (\psi_R^T C \psi_R + \bar{\psi}_R C \bar{\psi}_R^T) \tag{1.35}\]

By defining new Majorana fields \(N_1, N_2\) one can rewrite (1.35):
\[-L_{mass} = \begin{bmatrix} N_1 & N_2 \end{bmatrix} \begin{bmatrix} m_L & m_D \\ m_D & m_R \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} \tag{1.36}\]
where \(N_1 = \frac{\psi_L + \psi_R^C}{\sqrt{2}}\) and \(N_2 = \frac{\psi_R + \psi_L^C}{\sqrt{2}}\).

A few scenarios are possible:

1. \(m_D \neq 0; m_L = m_R = 0\): neutrino is a Dirac particle;
2. \(m_L \neq 0; m_D = m_R = 0\): neutrino is a Majorana left-handed particle;
3. $m_L \neq 0; m_D \neq 0; m_R \neq 0$: if $m_D \gg m_R, m_L$, then neutrino is mainly Dirac particle; if $m_D \approx m_R, m_L$, then neutrino is a mix of Dirac and Majorana particle; if $m_R, m_L \gg m_D$, then neutrino is a Majorana particle.

In case, $m_R \gg m_D$, the neutrino mass mechanism is called a seesaw mechanism. If $m_L = 0; m_R \gg m_D$, then it is a type I seesaw mechanism. After diagonalizing the mass matrix in (1.36) one can obtain two eigenstates: a light almost left-handed $\nu'_L$ neutrino and a heavy almost right-handed $N_R$ neutrino. Those neutrino masses are:

$$m_\nu = -\frac{m_D^2}{m_R} \quad (1.37)$$
$$m_N \approx m_R \quad (1.38)$$

where $m_\nu$ is a light neutrino mass and $m_N$ is a heavy neutrino mass. The negative value for the mass eigenvalue in (1.37) becomes positive if one redefines the light neutrino mass eigenstate as $\nu_L = \gamma_3 \nu'_L$.

As it was mentioned before, neutrinos are neutral, but they carry the lepton number $L$:

$$L = L_e + L_\mu + L_\tau, \quad (1.39)$$

where $L$ is a lepton number; $L_e, L_\mu, L_\tau$ are electron, muon and tau-lepton numbers, that are equal to 1 when the particle is electron, muon or tau flavored and -1 is it’s an electron, muon, tau flavored anti-particle. Therefore, a neutrino will always have $L = 1$ and anti-neutrino $L = -1$. For Dirac particles, lepton number is conserved in the interaction, i.e. the sum of the lepton numbers in the process is 0. However, if neutrino is a Majorana particle, the lepton number conservation is violated on $\Delta L = 2$.

If neutrino is a Majorana particle, the PMNS matrix (1.13) will have two extra Majorana CP violation phases, [20]:

$$V_{PMNS} = V_D \times \text{diag}(1, e^{i\alpha_2/2}, e^{i\alpha_3/2}) \quad (1.40)$$

where $V_D$ is a matrix from (1.15); $\alpha_{21/31}$ are Majorana CP violation phases.

### 1.4.1 Mixing parameter

The number of events is defined as:

$$N = \mathcal{L} \varepsilon \sigma BR \quad (1.41)$$

where $N$ is a number of events of interest; $\mathcal{L}$ is an integrated luminosity, that is proportional to the number of collisions inside the detector; $\varepsilon$ is an efficiency; $\sigma$ is a production cross-section, i.e. relative amount of produced mother particles; $BR$ identifies the ratio of the specific decay of the mother particle.

In a specific case of $W \rightarrow \mu N_R$ with $N_R \rightarrow \mu jet$:

$$N_N = \mathcal{L} \varepsilon_N \sigma_W BR(W \rightarrow \nu \mu) BR(N_R \rightarrow \mu jet) |B_{\mu N}|^2 \quad (1.42)$$

where $\varepsilon_N$ is a signal efficiency; $\sigma_W$ is a W production cross section; $BR(W \rightarrow \nu \mu)$ is a branching ratio of $W$ decaying into muon and muon neutrino; $BR(N_R \rightarrow \mu jet)$ is a branching ratio of $N_R$ decaying into muon and jet; $|B_{\mu N}|$ is a muon-heavy neutrino mixing parameter.

To get the results proportional to the $|B_{\mu N}|$ a ratio to the normalization channel $W \rightarrow \mu \nu$ is calculated. The number of the normalization channel decays is defined
as:

\[ N_W = \mathcal{L}\epsilon_W\sigma_W\text{BR}(W \to \nu\mu) \quad (1.43) \]

where \( \epsilon_W \) is the normalization channel efficiency. Then the product of a branching ratio of a heavy neutrino decaying into a muon and a jet and a mixing parameter can be defined as:

\[ \text{BR}(N_R \to \mu\text{jet})|B_{\mu N}|^2 = \frac{N_N\epsilon_W}{N_W\epsilon_N} \quad (1.44) \]

1.5 LHCb dictionary

Below some LHCb terms are described in the alphabetical order.

**Candidate**

is a reconstructed particle, that is not tagged as a signal or a background event yet.

**Cross section**

is a measure of a probability for an event to happen, calculated using quantum field theory.

**Daughter particles**

are products of a decay, see **Mother particle**.

**Event**

is a certain particle decay or a set of decays, that happened during one proton-proton collision.

**Exclusive measurement**

is a measurement of a process with a specific final state. For example, particle \( Y \) decays into final states \( X_1, X_2, X_3, Z_1, Z_2, Z_3 \). A measure of \( Y \to X \), where \( X = X_1, X_2, X_3 \) is an inclusive measurement. See also **Inclusive measurement**.

**Integrated luminosity**

is a luminosity integrated over time. See **Luminosity** for luminosity definition.

**Inclusive measurement**

is a set of possible processes with a shared certain characteristics and usually same mother particle (see **Mother particle**), measured all together. For example, a particle \( Y \) can decay into final states \( X_1, X_2, X_3, Z_1, Z_2, Z_3 \). A measure of \( Y \to X \), where \( X = X_1, X_2, X_3 \) is an inclusive measurement. See also **Exclusive measurement**.

**Interaction point**

is where the proton-proton collisions take place.
Impact parameter

is a closest distance between the original vertex and the track, see Fig. 1.5.

Luminosity

is a detected events rate normalized to the corresponding cross-section:

\[ \mathcal{L} = \frac{dN_{\text{det}}}{dt} \cdot \frac{1}{\sigma} \] (1.45)

where \( N_{\text{det}} \) is number of detected events and \( \sigma \) is a cross-section. See also integrated luminosity and Cross section.

Monte Carlo generation

is a random simulation of any physics process. In this thesis, it is usually referred to as a set of simulations: a simulation of the quark-level process itself, a simulation of quarks hadronization, a simulation of the detector and a simulation of the read-out electronics.

Mother particle

is a decaying particle. Products of decaying mother particle are often called daughter particles, see Daughter particles.

Pile Up

is a set of all other proton collisions that do not contribute to the decay of interest. Pile up events play a background role.

Pseudo-rapidity

is a kinematic variable, that is defined by the angle between emerging particle and beam axis:

\[ \eta = -\ln\left(\tan\frac{\theta}{2}\right) \] (1.46)

where \( \theta \) is an angle between particle 3-momentum, \( \vec{p} \) and beam axis.

Stripping

is a part of a preselection, where data is separated into separate streams. Each stream is defined by a dedicated selection, to facilitate the user access for specific analyses.

Track

is a reconstructed charged particle trajectory.

Vertex (decay vertex)

is defined as a common origin daughter particles, see Fig. 1.5. Primary vertex is an actual proton-proton collision point. Secondary vertex is a vertex of a particle decay.
1.5. LHCb dictionary

Figure 1.5: A schematics of a decay with vertices and particles indicated.
LHCb detector

LHCb experiment is b-physics dedicated experiment at LHC. Although LHCb research concentrates on CP violation studies, specifically b-hadrons physics, LHCb also participates in the search for the rare decays, lepton universality violation, etc. Since the main targets of the experiment are b-hadrons decays, which have a very distinctive high-pseudorapidity kinematics, the geometry of the detector is unique among all other LHC experiments. The detector geometry covers $2 < \eta < 5$ range, where $\eta$ is pseudo-rapidity, that corresponds to $10 < \theta < 250$ (300) mrad in the non-bending (bending) plane [30].

The LHCb detector consists of a several subdetectors. As can be seen in Fig. 2.1, the closest to the interaction point is a vertex detector, VELO (“Vertex Locator”). VELO is the first part of the charged particles tracking system, that also includes four tracking stations: Tracker Turicensis (TT), Tracker 1 (T1), Tracker 2 (T2), Tracker 3 (T3). TT is located right after the ring imagining Cherenkov detector RICH1 and before the LHCb magnet. RICH is used for the particle identification. T1, T2 and T3 are located after the magnet. Together with the magnet the tracking system allows to measure charged particles momentum. Another ring imagining Cherenkov detector, RICH2, is installed right after T3. Right next to it, the first muon station, M1, a part of a bigger muon detection system, that includes 5 muon stations M1-5, is installed. After the first muon station, the calorimeter system is placed. The calorimeter system measures the energy deposit of particles. It consists of the scintillating pad detector (SPD), preshower detector (PS), electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). The calorimeter system and the muon system are used for hardware Level-0 or L0 trigger, the first trigger in the trigger system. The trigger system consists of the mentioned already hardware L0 trigger and two subsequent software triggers: high level trigger 1 or HLT1 and high level trigger 2 or HLT2. HLT1 and HLT2 use information from different parts of the tracking system to trigger events.

General information about the LHCb detector resolution and efficiency can be found in Tab. 2.1.

2.1 Vertex Locator

VELO is a silicon strip vertex locator detector mounted at the LHCb experiment [32, 33]. The VELO main task is to reconstruct the vertices originating near the interaction point. Therefore, it is required to be placed close to the beam pipe. When the LHC beam is stable, the VELO sensors are as close as 8.2 mm\(^1\) to the beam. During the beam fill, to prevent the damage from the unstable LHC beam, VELO stations

\(^1\)Depending on the momentum an average b-meson will propagate (450-490)$\mu m\ p/m$, where mass is approximately 5300 MeV. For a 100 GeV $B^0$-meson it is about 8.6 mm [20]
have to be relocated 3 cm away [34]. To protect LHC vacuum from VELO detector vacuum, VELO is surrounded with an RF-foil box. The RF-box also protects VELO electronics from the beam induced RF noise and guides the beam mirror current\(^2\).

The VELO detector 21 stations were divided into 2 halves: left and right, to move the VELO before and after the physics runs, see Fig. 2.2. Each VELO module consists of a silicon r-sensor and a $\phi$-sensor, named after the strips orientation, which is azimuth and quasi-radial respectively, see Fig. 2.3. The signal hit is caught by the strip and is then collected via the metal routing line, that transfers signal to the read-out electronics. The VELO read-out electronics is situated away from the radiation intense interaction point.

VELO has a vertex resolution of 71 $\mu$m along the beam direction and 13 $\mu$m resolution perpendicular to the beam direction for a primary vertex with 25 tracks [33]. For the tracks with $p_T > 1$ GeV the impact parameter resolution is better than 35 $\mu$m.

### 2.2 Tracking stations

TT is located before the magnet and T1-3 are located after the magnet [30, 35]. The entire TT area and around 1.2% of the T1-3 stations area are covered with silicon 200 $\mu$m wide strips [36]. The silicon detectors are intentionally installed close to the beam, since majority of produced particles will get a relativistic boost in the forward direction [37]. The 98.8% of the T1-3 stations is the gaseous 5 mm diameter strawtube detector filled with Ar/CO\(_2\) (70\%:30\%) gas mixture.

\(^2\)Beam mirror current appears, because charged protons attract charge. Attracted charge, called mirror charge, then follows the proton beam in the LHC.
2.2. Tracking stations

Figure 2.2: Above: the position of 21 VELO stations. Below: the left and right modules when the LHC beam is stable (VELO “closed”) and unstable (VELO “open”).

Figure 2.3: The schematics of the VELO strip r- and φ-detectors. The black lines indicate silicon strips and the light green lines indicate the routing lines.
momentum
ECAL
impact parameter
invariant mass
\[ \frac{\Delta p}{p} = 0.5 - 1\% \]
\[ 1\% + 10\% / \sqrt{E [\text{GeV}]} \]
\[ 15 + 29 / p_T [\text{GeV}] [\mu \text{m}] \]
\[ \sim 8 \text{ MeV for } B \to J/\psi X \]
\[ J/\psi \text{ mass constrained } \sim 22 \text{ MeV} \]
in two-body B decays
\[ \sim 100 \text{ MeV for } B_s \to \phi \gamma \]
(photonic dominated)
\[ \sim 45 \text{ fs} \]
\[ (B \to J/\psi \phi, B \to D_s \pi) \]

decay time

| subdetectors working channels | \sim 99\% |
| data taking | \sim 90\% |
| trigger | \sim 90\% |
| (dimuon channels) | \sim 30\% |
| (multi-body hadronic final states) | \sim 96\% long tracks \footnote{Track is considered long if it has hits in all parts of the tracking system: VELO, TT, T1-T3.} |
| electron ID | \sim 90\% |
| (for \sim 5\% probability to misidentify electron as hadron) | \sim 95\% |
| kaon ID | \sim 97\% |
| (for \sim 5\% probability to misidentify pion as kaon) | \sim 97\% |
| muon ID | \sim 97\% |
| (for 1\% - 3\% probability to misidentify pion as muon) |

\begin{table}
\centering
\begin{tabular}{|l|c|}
\hline
resolution & \frac{\Delta p}{p} = 0.5 - 1\% \\
& \begin{align*}
1\% + 10\% / \sqrt{E [\text{GeV}]} \\
15 + 29 / p_T [\text{GeV}] [\mu \text{m}]
\end{align*} \\
& \sim 8 \text{ MeV for } B \to J/\psi X \\
& J/\psi \text{ mass constrained } \sim 22 \text{ MeV} \\
in two-body B decays & \sim 100 \text{ MeV for } B_s \to \phi \gamma \\
& (photonic dominated) \sim 45 \text{ fs} \\
& \begin{align*}
(B \to J/\psi \phi, B \to D_s \pi)
\end{align*} \\
\hline
efficiency & \sim 99\% \\
& \sim 90\% \\
& \sim 90\% \\
& \sim 30\% \\
& \sim 96\% long tracks \footnote{Track is considered long if it has hits in all parts of the tracking system: VELO, TT, T1-T3.} \\
& electron ID \sim 90\% \\
& (for \sim 5\% probability to misidentify electron as hadron) kaon ID \sim 95\% \\
& (for \sim 5\% probability to misidentify pion as kaon) muon ID \sim 97\% \\
& (for 1\% - 3\% probability to misidentify pion as muon) \\
\hline
\end{tabular}
\end{table}

The layers of the tracking stations are rotated around z-axis with respect to each other to exclude hit position ambiguity. The first and the last layer are positioned along y-axis and the intermediate layers are rotated $\pm 5^\circ$ with respect to the first and the last layer.

The hit resolution of the TT detector is approximately 40-50 $\mu$m, see Fig. 2.4a taken from \cite{37}. The hit resolution is determined from the comparison of the difference between the expected position from track reconstruction and the measured position of the hit. The inner silicon parts of T1-T3 have a comparable hit resolution of about 50 – 60$\mu$m depending on the layer of the detector and the geometrical location of strips, see Fig. 2.4b taken from \cite{37}. The single hit resolution of the gaseous detector part of T1-T3 is 160 $\mu$m \cite{38}.

\subsection*{2.3 Magnet}

The LHCb dipole magnet is located between 3 and 7 m away from the interaction point. The magnet together with the tracking system are used for charged particles momentum measurements from the track curvature in the magnetic field. It is
2.4. Track reconstruction and track reconstruction efficiency

(A) TT hit resolution  
(B) IT hit resolution

Figure 2.4: The hit resolution of silicon tracking detectors taken from [37]. X1, X2 are the first and the last of the four layers. U, V are the intermediate layers. C-side, A-side, Central, Top and Bottom mark the geometrical position of the detector.

Figure 2.5: Magnetic field inside LHCb detector along the beam axis (z-axis), “Magnet Down” configuration [39].

designed to retrieve the maximum magnetic field possible between the TT and T1 tracking stations\(^3\) and at the same time to have negligible influence on the RICH hybrid photo detectors [30]. The LHCb magnet integrated field is 4T·m for 10 m long tracks. To reduce the detector alignment uncertainty, the data at LHCb is usually taken with two magnetic field configurations: “Magnet Up” (MU) and “Magnet Down” (MD), corresponding to the positive/negative \(B_y\) \(^4\) respectively. Magnetic field distribution in the LHCb detector is shown in Fig. 2.5, taken from [39].

2.4 Track reconstruction and track reconstruction efficiency

A track reconstruction is the core part of a particle reconstruction. It includes not only determination of the momentum, matching the vertex, matching with the calorimeter cluster and, therefore, measuring the energy, and combining particles for invariant mass calculation.

The LHCb collaboration track classification [37, 40, 41]:

\(^3\)Some field is still left in the tracking stations for the curvature measurements.

\(^4\)\(B_y\) is a magnetic field along y-axis, which is perpendicular to the beam line (z-axis) in the left-handed convention
1. VELO track: track leaves hits only in the VELO.
2. T track: track leaves hits only in the T1-3 stations.
3. Downstream track: track leaves hits in the TT and T1-3.
4. Upstream track: track leaves hits in the VELO and TT.
5. Long track: track leaves hits in the VELO and T1-3 (also can have TT hits).

In Fig. 2.6 one can see the schematics of the LHCb with different types of tracks sketched.

The most informative track will be stored after reconstruction for each particle in case there are more than one possible track types for a particle. For example, if a long track will be reconstructed and any other type track is also reconstructed for the same particle, the other type will be thrown away and only the long track will be saved. To be reconstructed as a VELO track, particle has to leave minimum 3 hits in both r- and $\phi$-sensors of the VELO. The VELO track can then be combined with at least 3 TT hits to reconstruct an upstream track or with T1-3 hits to reconstruct a long track. From a combination of a VELO track and hits in the T1-3 stations, the particle curvature in the magnet is reconstructed and the particle momentum is determined. If one hit in the T1-3 station matches with a VELO track, further reconstruction of a long track is done along the direction defined by matching VELO track and one T1-3 station hit. TT hits are added to a long track if possible, to improve the momentum resolution and reduce the fake track rate. Reconstructed T tracks should be matched with at least 3 hits in the TT station to be reconstructed as a downstream track. After matching the hits, the track is fitted with a Kalman fitter [39]. Track fitting is necessary for an accurate track parameter estimation. In the Kalman fitter formalism each track is represented with a set of nodes. Track node consists of a track state in the measurement plane and a measurement. Track states are straight lines between track hits. The track state parametrization allows to store information about the position of the track at each given $z$-coordinate and the curvature of the track. Noise

---

5 Measurement plane is a surface where hit was registered

6 $z$-axis is oriented along the beam line
2.4. Track reconstruction and track reconstruction efficiency

Figure 2.7: Track reconstruction efficiency as a function of momentum, $p$; pseudorapidity, $\eta$; number of reconstructed tracks in the event, $N_{\text{track}}$; number of reconstructed primary vertices in the event, $N_{\text{PV}}$. Taken from [37].

Effects, like multiple scattering and energy loss, are as well taken into account. The Kalman filter is a $\chi^2$-minimization fit. Measurements are added sequentially to the fit. Each time the measurement is added, the track state and $\chi^2$ are recomputed. Fake tracks are not produced by a real particle, but can appear due to some mismatching of hits before and after the magnet. Fake track rate varies from 6.5% to 20% depending on the particles multiplicity and is reduced using multivariate analysis. The track reconstruction efficiency is the probability of an existing long track that passes all tracking stations to be reconstructed. It is estimated using the tag-and-probe method in $J/\psi \rightarrow \mu^+\mu^-$ channel as described in [37] and [41]. The track reconstruction efficiency is about 95%. The track efficiency as a function of the particle momenta, rapidity, number of reconstructed tracks in the event and number of reconstructed primary vertices is shown in Fig. 2.7.

From curvature and magnetic field one can calculate particle momentum. By applying the standard reconstruction with two long tracks the momentum resolution of 16 MeV can be achieved in $J/\psi \rightarrow \mu^+\mu^-$ channel. The momentum resolution as a function of tracks momentum determined from $J/\psi \rightarrow \mu^+\mu^-$ channel is shown in Fig. 2.8a. For the low momentum tracks it is 0.5%. For the high momentum tracks resolution reaches 1-1.1%.

Similarly to the momentum resolution, the mass resolution increases as the mass of the particle increases. The relative mass resolution as a function of reconstructed dimuon mass is shown in Fig. 2.8b. For the low dimuon mass muons dimuon peak has resolution of about 0.5% and rises to 2% for the high dimuon mass.
2.5 Muon stations

The muon stations detect and identify muons passing through the LHCb. There are 5 muons stations: M1, M2, M3, M4, M5. In total they consist of 1380 multi-wire proportional chambers filled with Ar/CO$_2$/CF$_4$ (40:55:5) [30, 42]. A multi-wire proportional chamber is a gaseous tracking detector, where signal is produced via the gas ionization by the penetrating charged particle. The new electrons and ions are accelerated with voltage, applied to the chamber plates, and are collected with a set of wires. The muon system covers $16 < \theta < 306 \text{ (258) mrad}$ in the bending (non-bending) plane, which allows to detect about 20% of all muons coming from inclusive semileptonic b-decays. To reduce the background level muon tracks are composed out of the muon station tracks only. Between the M2-5 stations 80 cm wide iron plates are installed to stop hadrons, see schematics in Fig. 2.9.

The operating voltages of the muon chambers are 2500-2700 V. The gas gap is 5 mm wide and wires are 2 mm away from each other to reduce the cross-talk between the channels. The gas gain is about $10^5$ at 2650 V.

The muon chambers have a time resolution from 2.5 ns to 4 ns depending on the geometrical position [43]. The time resolution is shown in Fig. 2.10.

The muon stations information is used for the L0 trigger. The LHCb muon trigger is 95% efficient.

The muon stations are used for muon identification. As a result of the muon identification by M2-5, hadron misidentification decreases to 1% level [44]. Muons that hit M2-5 are tagged with an isMuon tag. To get the isMuon tag a 3-6 GeV muon has to leave hits in M2 and M3; a 6-10 GeV in M2, M3 and M4 or M5; a muon with momentum above 10 GeV in M2, M3, M4 and M5. The isMuon flag provides a first level loose selection of muons that can be further strengthened by comparison of muon and non-muon hypothesis. By adding the RICH and calorimeter information, a combined likelihood can be constructed for the muon identification.

2.6 Ring Imagining Cherenkov detector and particle identification

Two Ring Imagining Cherenkov detectors are installed at the LHCb: RICH1, located before the magnet and RICH2, located after the magnet [45–47]. RICH1 is used to identify particles, specifically charged $\pi$, K and p, in the momentum range 2-60 GeV
2.6. Ring Imaging Cherenkov detector and particle identification

**Figure 2.9:** M1-5 schematics, taken from [30]. The muon filters are 80 cm wide iron plates. R1-4 identify the muon stations geometrical regions.

**Figure 2.10:** Muon chambers time resolution, taken from [43].
Chapter 2. LHCb detector

Figure 2.11: Schematics of the RICH1 and RICH2 detectors.

Figure 2.12: Cherenkov angle as a function of a particle momentum.

and RICH 2 identifies particles in the momentum range 15-100 GeV. The RICH particle identification is based on the emission of the Cherenkov radiation by passing particles with a speed higher than the speed of light in the RICH gas mixture. RICH1 is filled with Aerogel and C4F10 and RICH2 uses CF4. A special system of mirrors, sketched in Fig. 2.11, is made to collect the produced light and direct it to the Hybrid Photon Detectors, HPD.

A Cherenkov radiation photon reflected from the mirrors, is caught by one of the HPDs photocathodes, where it is converted into a photoelectron. A 10-20 kV amplifying electric field is applied to accelerate the photoelectron. The photoelectron is detected by the silicon detector opposite to the photocathode. The received signal is amplified with approximately 5000 electron-hole pairs produced in silicon. The silicon detector has 1024 pixels, each 2.5×2.5 mm. From a ring of light, emitted by a passing particle, a Cherenkov angle can be reconstructed. By plotting the Cherenkov radiation angle as a function of momentum, particles can be identified, see Fig. 2.12.

The Cherenkov angle resolution measured as described in [48] as a function of time (i.e. run number) is shown in Fig. 2.13. Both RICH1 and RICH2 angular resolutions are stable in time. RICH1 resolution reaches 1.62 mrad, while RICH2 resolution is approximately 0.66 mrad.

RICH particles identification efficiency and misidentification rate, measured in [48] as a function of momentum, is shown in Fig. 2.14.
2.6. Ring Imaging Cherenkov detector and particle identification

Figure 2.13: RICH1-2 angular resolution in a time scale [48].

(A) RICH1 angular resolution  
(B) RICH2 angular resolution

Figure 2.14: Kaon and proton identification efficiency and pion misidentification rate for two $\Delta \log L$ requirements [48].
Chapter 2. LHCb detector

2.7 Calorimeter system

The LHCb calorimeter system consists of 4 components: a preshower, a scintillating pad detector, an electromagnetic calorimeter and a hadron calorimeter [30, 49]. Charged pions are defined by the signal in the PS and the ECAL. High $E_T$ neutral pions do not leave a trail in the SPD, which is used to identify charged particles.

The ECAL has a "shashlik" structure, composed of 2 mm thick lead and 4 mm thick scintillator plates. The ECAL energy resolution is $1\%+10\%/\sqrt{E}$. The above resolution is enough for efficient electron/hadron distinction. Particles lose energy via producing shower of daughter particles in the lead plate. This shower is then measured in the scintillating plate by capturing the amount of light produced. The ECAL is divided into three regions with different segment size: $4 \times 4 \text{ cm}^2$, $6 \times 6 \text{ cm}^2$ and $12 \times 12 \text{ cm}^2$, to compensate for the occupancy variation. The ECAL is 25 radiation length long, so it can fully absorb high energy photons and electrons.

The five radiation lengths long HCAL is made of 16 mm iron and 4 mm scintillator pads and has only two regions with the different cell size along the surface. The HCAL energy resolution is $10\%+80\%/\sqrt{E}\%$

The SPD and the PS are scintillator pads, that are separated by a 2.5 radiation length (15 mm) lead converter [30, 49, 50]. The PS/SPD pair copies the same cell size regions division, as ECAL, with smaller cells closer to the beam pipe and increasing cell size to the corner.

From a high energy photon reconstruction in the $B^0 \rightarrow K^{0\ast} \gamma$ sample the PS/ECAL cell-to-cell calibration accuracy was measured to be 2% [37]. The ECAL mass resolution for $p_T < 2 \text{ GeV}$ neutral pions is about 8 MeV and declines to 30 MeV for $p_T > 2 \text{ GeV}$ neutral pions. The high $p_T$ neutral pion decrease in the mass resolution is a consequence of ECAL cell dimensions and a small opening angle of photons coming from the high $p_T$ neutral pion.

The calorimeter system provides particle identification information for photons, electron and neutral pions. Charged particles are recognized by matching deposits in calorimeters with tracks. Neutral particles are defined as a calorimeter response without a matched track. Photon or neutral pion have similar signals and are distinct by applying a $\Delta \log L$ criteria to the calorimeter shower shape. Efficiency of the photon identification with respect to the neutral pion rejection rate in the $B^0 \rightarrow K^{0\ast} \gamma$ sample is shown in Fig 2.15a. Information for the electron identification is taken from the entire calorimeter system. The electron identification efficiency and misidentification rate in the $B^+ \rightarrow J/\psi K^\pm$ decay with $J/\psi \rightarrow e^+ e^-$ are shown in Fig. 2.15b and Fig. 2.15c.

By combining information from the RICH, the calorimeters and the muon system a combined $\Delta \log L$ can be used to identify particles.

2.8 Jet reconstruction

A jet is a spray of energetic particles produced by a scattering process that includes partons in the final state. It can appear for example in a $e^+ e^- \rightarrow q \bar{q}$ process through QCD hadronization of the final quarks.

Jet reconstruction starts with selecting input particles for a jet reconstruction algorithm.
2.8. Jet reconstruction

Figure 2.15: (A) Photon identification efficiency with respect to neutral pion rejection rate for $B^0 \rightarrow K^{*0}\gamma$ data and simulation. (B) Electron identification efficiency as a function of track momentum for $B^\pm \rightarrow J/\psi K^\pm$ decay with $J/\psi \rightarrow e^+e^-$ 2011 data under different $\Delta \log L$ requirements. (C) Electron misidentification rate as a function of track momentum for $B^\pm \rightarrow J/\psi K^\pm$ decay with $J/\psi \rightarrow e^+e^-$ 2011 data under different $\Delta \log L$ requirements. All taken from [37].
2.8.1 Jet input particle selection

Charged particles are reconstructed using the track and calorimeter information, whereas neutral particles are reconstructed using the calorimeter information only\(^7\) [51]. The input selection for jets is done by the ParticleFlow4Jets algorithm. It selects particles that can be a part of a jet and reduces a non-jet noise.

First, ParticleFlow4Jets finds a charged composition of a jet by matching tracks and calorimeter clusters (both ECAL and HCAL, if applicable). Tracks used in this step should pass track quality criteria described in [51] and have a curvature \(\frac{\epsilon_T}{q/p} > 10\). To avoid double counting as a charged and a neutral particle, a track with all connected to it clusters are not used in a search for a neutral component. Using different PID hypothesizes neutral particles are classified as \(\pi^0\) or \(\gamma\). The HCAL clusters, not connected with charged particles tracks, are identified as jet neutrals. Since ECAL and HCAL clusters can overlap, an additional constraint on matching a cluster and a track is applied: for a charged particle \(\chi^2_{\text{ECAL/TRACK}} < 25\) and for a neutral particle \(\chi^2_{\text{ECAL/TRACK}} > 16\) for tracks with \(E > 5\) GeV. An advantage of using ParticleFlow4Jets algorithm is that it reduces dependence on the calorimeter resolution and uses an advantage of a high-resolution tracking of the LHCb experiment [52].

2.8.2 Jet reconstruction algorithm

For reconstruction of jets a default LHCb procedure is: the anti-\(k_T\) algorithm with \(R = 0.5\), where \(R\) is a jet cone radius with respect to the hardest particle in a jet\(^8\). The anti-\(k_T\) algorithm is a sequential recombination jet algorithm\(^9\), which is infrared and collinear safe\(^10\) [54]. The anti-\(k_T\) algorithm results in cone-shaped jets with regular boundaries, not influenced by the infrared radiation\(^11\). Cone-shaped jets are especially advantageous for the LHCb detector. They allow to reconstruct jet energy using only a part of a jet, which is needed since LHCb detector acceptance is small [55]. The algorithm itself is based on calculating the following measures:

\[
\begin{align*}
    d_{ij} &= \min(k_{T_i}^{-2}, k_{T_j}^{-2}) \frac{\Delta^2_{ij}}{R^2} \quad (2.1) \\
    d_{iB} &= k_{T_i}^{-2}, \quad (2.2)
\end{align*}
\]

where \(d_{ij}\) is a distance measure between \(i\)- and \(j\)-particle, \(d_{iB}\) is a distance measure between \(i\)- particle and the beam, \(k_T\) is the transverse momentum, \(R\) is a jet radius parameter. \(\Delta_{ij}\) in (2.2) is a geometrical scale defined as:

\[
\Delta_{ij} = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2, \quad (2.3)
\]

where \(\eta\) is a pseudorapidity and \(\phi\) is an azimuth angle [52].

\(^7\)Neutral particles do not leave tracks in the detector.

\(^8\)\(R = \sqrt{\Delta \eta^2 + \Delta \phi^2}\).

\(^9\)Other examples of sequential recombination algorithms are \(k_T\) and Cambridge/Aachen algorithms. One can also use a completely different class of jet reconstruction algorithms, called cone-type algorithms (Iterative Cone, SisCone).

\(^10\)Infrared and collinear safety means that adding soft radiation or a collinear split of a hard particle to an input jet particles will not change the jets size, energy, multiplicity, etc. [53].

\(^11\)As a result, the area of a jet cone is the same for point-like and diffuse radiation.
2.9 Trigger system

The LHC produces collisions at the 40 MHz rate, i.e. a bunch is crossing the interaction point every 25 ns. LHCb trigger system has to efficiently select different final states. The trigger system is divided into three levels: hardware L0, software HLT1 and software HLT2 [56]. See Fig. 2.16 for a schematics.

The L0 trigger has to reduce the 40 MHz rate to the 1 MHz read-out rate\(^\text{12}\). This is done by using the muon stations and the calorimeter system information. The L0 roughly reconstructs high-\(p_T\) muons and high-\(E_T\) photons, hadrons and electrons in the calorimeters. The 1 MHz rate is decomposed to the 450 kHZ of charged hadron candidates (L0Hadron), the 400 kHz of muons candidates (L0Muon/L0DiMuon) and the 150 kHz of photon/electron candidates (L0Photon/L0Electron), see Fig. 2.16.

---

\(^{12}\)All the rates are approximate, due to the slightly changing collision rate.
There are three separate triggers included into L0 trigger: L0 Muon trigger, L0 Calorimeter trigger and L0 Pile-Up trigger.

L0 Muon trigger uses muon stations hits to detect highest $p_T$ muons [57]. Hits are checked for aligning along the same line across M1-5. Events that pass this line are tagged L0Muon or L0DiMuon, depending on the number whether trigger threshold was set on the highest $p_T$ only (threshold: $p_T > 1.76$ GeV in 2012) or two highest $p_T$ (threshold: $p_T^1 \times p_T^2 > (1.60 \text{ GeV})^2$ in 2012).

The L0 Calorimeter trigger selects highest $E_T$ clusters in the calorimeters. Clusters are categorized based on the calorimeter system response. Categories for 2012 are described in Tab. 2.2.

<table>
<thead>
<tr>
<th>category</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0Hadron</td>
<td>$E_T &gt; 3.7$ GeV HCAL+ECAL cluster, 600 SPD hits</td>
</tr>
<tr>
<td>L0Photon</td>
<td>$E_T &gt; 3$ GeV HCAL cluster, PS hits, no SPD hits in front of ECAL shower</td>
</tr>
<tr>
<td>L0Electron</td>
<td>$E_T &gt; 3$ GeV HCAL cluster, PS hits, 600 SPD hits</td>
</tr>
</tbody>
</table>

Table 2.2: L0 Calorimeter trigger categories for 2012 [57, 58].

L0 Pile-Up trigger is used to determine a relative luminosity. It recognizes an event multiplicity and provides an independent monitoring of the luminosity.

A random L0 triggers also select “No Biased” events. The collective L0 trigger decision goes to the readout supervisor board.

Then events are selected with a software trigger, divided into two levels: HLT1 and HLT2. The total HLT2 output read-out rate is about 5 kHz. In 2012 approximately 2 kHz of the output read-out was coming from inclusive topological trigger lines, that include b-hadron candidates, 2 kHz from charm trigger lines and 1 kHz from muon/dimuon trigger lines. HLT1 and HLT2 allow for programming different trigger lines. Each trigger configuration is assigned with a unique Trigger Configuration Key, TCK, that defines the software version, the selection and the algorithm for trigger [58].

HLT1 reduces the data rate to about 40 kHz using a simplified track reconstruction. HLT1 then decides which tracks are likely to come from a signal, based on the lowest impact parameter to any primary vertex\(^\dagger\). To match the VELO track with a muon stations track in case of L0Muon and L0DiMuon events, HLT1 extrapolates VELO track through magnet to M3. At least one other hit has to be found along the defined trajectory, in muons stations, except M3. The muon momentum has to be higher than 6 GeV in order to be reconstructed as a muon track. The combined track is then fitted and required to have $\chi^2/ndof < 25$. Using the forward tracking procedure (VELO + T track), described [57], the momentum is reconstructed for tracks reconstructed in HLT1. HLT level $J/\psi \rightarrow \mu^+\mu^-$ mass resolution is 3% worse than determined in offline reconstruction\(^\ddagger\) for the Run I data.

HLT2 adds particle identification information. The topological trigger line, charm trigger line and muon trigger line are applied in HLT2.

The trigger efficiency is estimated using the TISTOS method. The TISTOS method is based on two event classifiers: triggered independent of signal (TIS) and trigger on signal (TOS). The TIS tag is assigned if trigger flag fired not on the signal candidate. The TOS tag is assigned if the trigger flag fired on the signal candidate. TOS

\(^\dagger\)If no muon is included in the HLT1 line.

\(^\ddagger\)Offline reconstruction is a reconstruction done with trigger selected and stored event.
efficiency is defined as:
\[ \varepsilon_{TOS} = \frac{N(TOS \cap TIS)}{N(TIS)} \]  (2.4)

where \( \varepsilon_{TOS} \) is a TOS efficiency; \( N(TIS) \) is a number of TIS events and \( N(TOS \cap TIS) \) is a number of events classified as TIS and TOS.

The trigger efficiency is then calculated as:
\[ \varepsilon_{trig} = \frac{N_{sel}}{N_{TOS}} \cdot \varepsilon_{TOS} \]  (2.5)

where \( N_{sel} \) is a total number of events that passed the trigger; \( N_{TOS} \) is a number of TOS events.

See more on the trigger efficiency estimation from offline reconstructed data using the TISTOS method in [57, 59].

The L0Muon and L0DiMuon TOS efficiency from \( B^\pm \to J/\psi K^\pm \) with \( J/\psi \to \mu^+\mu^- \) (2012, data) as a function of the B momenta is shown in Fig. 2.17a [60]. The L0Hadron TOS efficiency is evaluated from the beauty decays (\( B^0 \to K^\pm\pi^\mp \), \( B^0 \to D^\pm\pi^\mp \)) and the charm decays (\( D^0 \to K^-\pi^+ \), \( D^\pm \to K^\mp\pi^\pm\pi^\pm \) and \( D^{*\pm} \to D^0\pi^\pm \)). As a function of the mother hadron transverse momentum the L0Hadron TOS efficiency is shown in Fig. 2.17b.

The muon HLT1 lines TOS efficiency as a function of the mother hadron momentum can be seen in Fig. 2.18a. The HLT1 inclusive track line efficiency is show in Fig. 2.18b. HLT2 lines TOS efficiency as a function of the mother hadron momentum is shown in Fig. 2.19.

For an extended description of the trigger lines see [57, 60].
Figure 2.18: HLT1 trigger lines TOS efficiency from different decays, taken from [37], as a function of a mother particle momentum.

Figure 2.19: HLT2 trigger lines TOS efficiency from different decays, taken from [37], as a function of a mother particle momentum.
Chapter 3

Analysis Strategy

The aim of the analysis is to discover a new neutrino-like heavy particle in the mass range 5-50 GeV with a lifetime of 10 ps. This analysis extends a prior analysis for a prompt heavy neutrinos ([17]) with focus on the improvement for the long-lived heavy neutrino case. The signal channel is $W \rightarrow \mu WN_R$, with subsequent $N_R \rightarrow \mu NX$ decay, where $\mu_W$ and $\mu_N$ have the same charge. The main backgrounds are $W \rightarrow \mu\nu$, $Z \rightarrow \mu^+\mu^-$, $W \rightarrow \tau\nu$, $Z \rightarrow \tau^+\tau^-$, $b\bar{b}$ and $c\bar{c}$. $W \rightarrow \mu\nu$ is used as a normalization channel.

The analysis strategy consists of the five main steps:

1. data reconstruction and Monte Carlo generation,
2. applying preselection and selection to retrieve the yields,
3. computing efficiency,
4. estimating systematic uncertainties,
5. comparison of a measured and an estimated upper limit.

The core of the analysis is the identification of background contributions and reducing background by optimizing the selection, resulting in a better upper limit on the light-heavy neutrino coupling as a function of mass. Using the prompt heavy neutrino results from [17] the upper limit can be interpolated also as a function of lifetime.

The first step of the analysis is the data reconstruction and the Monte Carlo generation. Data and Monte Carlo samples are described in the Chapter 4.

Then the preselection and the selection have to be applied. The preselection and the selection cuts, described in Chapter 5, closely follow [17]. Originally implemented for the prompt case, the selection was adjusted for the long-lived case by using a different impact parameter cut. The impact parameter is also used to tag muons on the reconstruction level, alternatively to the muon and jet invariant mass tag used in [17]. Boosted decision tree classifiers are trained to improve the selection. The BDT input variables were specifically chosen to be lifetime independent [17] to extend the original prompt analysis to the multiple heavy neutrino lifetime points.

There are three contributing efficiency: generator cut efficiency, reconstruction and stripping efficiency and selection efficiency, that are calculated in Chapter 6. The total efficiency is corrected to ensure that the efficiency estimated from the Monte Carlo, describes the actual efficiency in data adequately. These corrections include muon reconstruction, identification and trigger corrections, a global event cut correction and a boosted decision tree classifiers correction. Muon corrections were taken from [61]. Other corrections were specially calculated for this analysis.
In Chapter 7 systematic uncertainties are described: a boosted decision tree classifiers correction uncertainty, a jet energy scale and resolution uncertainty, a jet identification uncertainty, uncertainties that come from muon efficiency corrections, an uncertainty from the global event cut correction.

In Chapter 8 results on the expected upper limit are shown and further improvements are discussed in Chapter 9.
Chapter 4

Samples

In this chapter the data and the Monte Carlo samples used in the analysis are described. Signal Monte Carlo samples are generated for six different masses of heavy neutrino with a lifetime of 10 ps. Produced Monte Carlo background samples are $W \rightarrow \mu \nu$ sample, $Z \rightarrow \mu^+\mu^-$ sample, $W \rightarrow \tau\nu$ sample, $Z \rightarrow \tau^+\tau^-$ sample, $b\bar{b}$ sample and $c\bar{c}$ sample.

4.1 Data Samples

The data corresponds to the 8 TeV 2012 LHCb data. The data samples properties per magnet polarity are shown in Tab. 4.1. The stripping line WMuLine from Stripping21 is applied. It requires at least one $p_T > 20$ GeV muon from StdAllLooseMuons.

I report that one event (run number: 133523 , event number: 2070595238) with 21 candidates in the Magnet Down data sample was found corrupted and is not used in the following analysis.

<table>
<thead>
<tr>
<th>polarity</th>
<th>$\mathcal{L}$ [pb$^{-1}$]</th>
<th>Fraction of total events recorded in 2012</th>
<th>events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>966.33</td>
<td>46.46 %</td>
<td>33369968</td>
</tr>
<tr>
<td>Down</td>
<td>960.67</td>
<td>46.18 %</td>
<td>34811836</td>
</tr>
<tr>
<td>Total</td>
<td>1927.00</td>
<td>92.64 %</td>
<td>68181804</td>
</tr>
</tbody>
</table>

TABLE 4.1: Properties of the data samples.

4.2 Signal Monte Carlo

The signal Monte Carlo samples leading order (LO) parton level generation was done using MadGraph5. The hadronization and reconstruction with Reco14-Stripping20NoPrescalingFlagged were done using Pythia 8 (Sim08). Both $W^\pm \rightarrow \mu^\pm\mu^\mp jet$ (muons have same charge) and $W^\pm \rightarrow \mu^\pm\mu^\mp jet$ (muons have opposite charge) samples were produced in the ratio 1:1. Different heavy neutrino with 10 ps lifetime masses were simulated: 5 GeV, 10 GeV, 15 GeV, 20 GeV, 30 GeV, 50 GeV. The properties of the signal simulated samples can be found in Tab. 4.2. The generation efficiency in Tab. 4.2 is defined as:

$$\varepsilon_{\text{gen}} = \frac{N(\text{passed generation cut})}{N(\text{generated})},$$

(4.1)
where $\varepsilon_{\text{gen}}$ is the generation efficiency; $N(\text{passed generation cut})$ is a number of events that passed the generation cuts; $N(\text{generated})$ is a number of events, that were generated.

### 4.3 Background Monte Carlo

The Monte Carlo background samples were generated using the Pythia 8 (Sim08). In the reconstruction Reco14-Stripping20NoPrescalingFlagged and trigger TCK0x409f0045 were applied. Information about produced background simulation samples can be found in Tab. 4.3.
<table>
<thead>
<tr>
<th>$m_N$ [GeV]</th>
<th>$\tau$ [ps]</th>
<th>event type</th>
<th>generator level cuts</th>
<th>events MU</th>
<th>events MD</th>
<th>$\varepsilon_{gen}$ MU</th>
<th>$\varepsilon_{gen}$ MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>42912011</td>
<td>2 muons with $p &gt; 2$ GeV, $\eta &gt; 2$, $\theta &lt; 400$ mrad</td>
<td>24425</td>
<td>24900</td>
<td>20.98 ± 0.12 %</td>
<td>20.65 ± 0.12 %</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>42912013</td>
<td>2 muons with $p &gt; 2$ GeV, $\eta &gt; 2$, $\theta &lt; 400$ mrad</td>
<td>24999</td>
<td>25000</td>
<td>19.26 ± 0.22 %</td>
<td>19.46 ± 0.11 %</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>42912015</td>
<td>2 muons with $p &gt; 2$ GeV, $\eta &gt; 2$, $\theta &lt; 400$ mrad</td>
<td>24750</td>
<td>24900</td>
<td>18.78 ± 0.11 %</td>
<td>18.63 ± 0.11 %</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>42912017</td>
<td>2 muons with $p &gt; 2$ GeV, $\eta &gt; 2$, $\theta &lt; 400$ mrad</td>
<td>24950</td>
<td>24799</td>
<td>17.93 ± 0.10 %</td>
<td>17.90 ± 0.10 %</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>42912019</td>
<td>2 muons with $p &gt; 2$ GeV, $\eta &gt; 2$, $\theta &lt; 400$ mrad</td>
<td>17300</td>
<td>24899</td>
<td>16.84 ± 0.10 %</td>
<td>16.69 ± 0.11 %</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>42912021</td>
<td>2 muons with $p &gt; 2$ GeV, $\eta &gt; 2$, $\theta &lt; 400$ mrad</td>
<td>24900</td>
<td>24799</td>
<td>14.99 ± 0.09 %</td>
<td>14.87 ± 0.18 %</td>
</tr>
</tbody>
</table>

**Table 4.2:** Properties of the simulated in MadGraph5 at LO and Pythia (Sim08) at different mass points signal samples.
<table>
<thead>
<tr>
<th>%</th>
<th>%</th>
<th>0.01</th>
<th>0.02</th>
<th>575029</th>
<th>511229</th>
<th>514348</th>
<th>512052</th>
<th>49011003</th>
<th>49011004</th>
<th>q\bar{q}</th>
<th>pT &gt; 10 GeV lepton</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>%</td>
<td>0.09</td>
<td>0.88</td>
<td>201149</td>
<td>201149</td>
<td>301549</td>
<td>301549</td>
<td>4221092</td>
<td>4221092</td>
<td>Z → µ⁺µ⁻</td>
<td>pT &gt; 4 GeV lepton</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>13.47</td>
<td>36.83</td>
<td>1072163</td>
<td>1072163</td>
<td>1086627</td>
<td>1086627</td>
<td>42300001</td>
<td>42300001</td>
<td>W → τν</td>
<td>pT &gt; 10 GeV muon</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>36.91</td>
<td>36.73</td>
<td>2039966</td>
<td>2039966</td>
<td>1994991</td>
<td>1994991</td>
<td>42210021</td>
<td>42210021</td>
<td>Z → τ⁺τ⁻</td>
<td>pT &gt; 4 GeV lepton</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>4.21</td>
<td>4.21</td>
<td>247286</td>
<td>247286</td>
<td>425730</td>
<td>425730</td>
<td>42311003</td>
<td>42311003</td>
<td>W → lν</td>
<td>pT &gt; 10 GeV lepton</td>
</tr>
<tr>
<td>event type</td>
<td>generator level cut</td>
<td>events</td>
<td>MD</td>
<td>events</td>
<td>MD</td>
<td>events</td>
<td>MD</td>
<td>events</td>
<td>MD</td>
<td>events</td>
<td>MD</td>
</tr>
</tbody>
</table>
Chapter 5

Reconstruction and selection

5.1 Reconstruction

The data is reconstructed using DaVinci v44r7.

At the level of reconstruction the sum of daughter muon candidates mass has to be bigger than 10 GeV. The consequential $N \rightarrow \mu_N X$ candidates must have mass bigger than 0.01 GeV and smaller than 100 GeV. Here and below $N_R$ is denoted as $N$.

The heavy neutrino candidate is required to have $0.01 < M < 80$ GeV and $p_T > 10$ GeV. The combined W candidate is required to have $20 < M < 200$ GeV.

A muon with the lowest impact parameter in the event is tagged as $\mu_W$ candidate. Other muons in the event are tagged as $\mu_N$ candidates and are not uniquely defined. This result in multiple $\mu_N$ candidates for each $\mu_W$. It is disfavored to choose an unique $\mu_N$ candidate too, because of the ambiguity in the definition of the best quality $\mu_N$ at the reconstruction stage. The distribution of number of $\mu_N$ candidates per $\mu_W$ candidate before applying preselection can be found in App. A. After the full selection was applied the multiplicity was found to be exclusively 1 for all signal Monte Carlo samples.

The Stripping21 line is used, that requires $p_T > 20$ GeV and $\text{isMuon}$ for $\mu_W$ candidate.

The further analysis concentrates on the same charge muons candidates. For the same charge muons $W^\pm \rightarrow \mu^\pm \mu^\pm$jet decay the Standard Model backgrounds are suppressed.

In Fig. 5.1 $p_T$ distributions for 4 heavy neutrino mass points: 5 GeV, 15 GeV, 30 GeV and 50 GeV from Monte Carlo samples are shown. $\mu_W$ is a high-$p_T$ particle for a 5 GeV heavy neutrino, coming from heavy 80.38 GeV W-boson. It becomes softer, as the mass of heavy neutrino increases. $\mu_N$ is a soft particle for a 5 GeV heavy neutrino, that hardens as heavy neutrino mass rises.

In App. B one can find distribution of reconstructed $\mu_W$ and $\mu_N$ in bins of $p_T$ and $\eta$ for different mass points of the heavy neutrino signal Monte Carlo.

5.2 W candidate preselection

The W candidate is formed from selected $\mu_W$, $\mu_N$ and jet candidates.

The $\mu_W$ has to pass all three muon trigger lines: L0Muon, Hlt1SingleMuonHighPT and Hlt2SingleMuonHighPT lines, which are described in Tab. 5.1. Both muons are required to be in LHCb pseudorapidity range. Muons must have $|q/p| > 10$ to ensure that high-$p_T$ muons charge can be reconstructed precisely. $\mu_N$ is required to be identified as a muon with isMuon tag. The soft $\mu_N$ is required to have $p_T > 3$ GeV.

On the other hand, $\mu_W$ is high-$p_T$ and boosted, so $20 < p_T(\mu_W) < 70$ GeV cut
Chapter 5. Reconstruction and selection

Figure 5.1: Muons $p_T$ distributions for different heavy neutrino mass Monte Carlo samples.

is applied. The $\mu_W$ track must have at least 1 TT-station hit and $P(\chi^2) > 0.01$ to minimize fake tracks contribution. To reduce the misidentification rate of $\pi$ and $K$, as $\mu_W$ in a calorimeter signal, $(E_{ECAL} + E_{HCAL})/p < 4\%$ is applied for $\mu_W$. The dimuon mass has to be in the range $20 < M(\mu_W\mu_N) < 70$ GeV to reduce the number of muons coming from a $Z \rightarrow \mu^+\mu^-$ peak. The $\mu_W$ impact parameter has to be lower than 0.04 mm, which insures that it is prompt. Opposite, $\mu_N$ is a non-prompt muon, so its an impact parameter is required to be bigger than 0.04 mm.

<table>
<thead>
<tr>
<th>stripping line</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stripping21</td>
<td>$p_T &gt; 20$ GeV, isMuon</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>trigger line</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0MuonDecision</td>
<td>nSPDhits &lt; 600, $p_T &gt; 1.5$ GeV.</td>
</tr>
<tr>
<td>Hlt1SingleMuonHighPTDecision</td>
<td>L0Muon, $p_T &gt; 4.8$ GeV, $p &gt; 8$ GeV, $\chi^2/ndf &lt; 4$</td>
</tr>
<tr>
<td>Hlt2SingleMuonHighPTDecision</td>
<td>$p_T &gt; 10$ GeV</td>
</tr>
</tbody>
</table>

Table 5.1: Stripping and trigger lines cuts.

The jet is reconstructed with the anti-$k_T$ algorithm, described in Sec.2.8.2, with a cone radius of 0.5. The jet $p_T$ has to be higher than 10 GeV and include at least one $p_T > 1.2$ GeV track, that passed all tracking stations. The charged $p_T$ fraction (cpf) in a jet has to be more than 0.1. The maximum $p_T$ of an individual track in the jet must be bigger than 1.2 GeV. Muons with $p_T$ higher than 2 GeV are excluded from the jet search.

Combined $\mu_N$ and jet candidates form a heavy neutrino N candidate, which is required to have $M(N) < 80$ GeV. The heavy neutrino candidate transverse momentum has to be bigger than 10 GeV.

The N candidate itself is combined with the $\mu_W$ candidate to get a W candidate. The W candidate is only required to be in the W mass window: $60 < M(W) < 100$ GeV.
### 5.2. W candidate preselection

<table>
<thead>
<tr>
<th><strong>μN candidates</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>isMuon</code></td>
<td></td>
</tr>
<tr>
<td>$2 &lt; \eta &lt; 4.5$</td>
<td></td>
</tr>
<tr>
<td>$p_T(\mu_N) &gt; 3$ GeV</td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>q/p</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>μW candidates</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>isMuon, p_T &gt; 20</code> GeV</td>
<td></td>
</tr>
<tr>
<td>$2 &lt; \eta &lt; 4.5$</td>
<td></td>
</tr>
<tr>
<td>$20 &lt; p_T(\mu_W) &lt; 70$ GeV</td>
<td></td>
</tr>
<tr>
<td>$(E_{ECAL} + E_{HCAL})/p &lt; 4%$</td>
<td></td>
</tr>
<tr>
<td>$P(\chi^2) &gt; 0.01$</td>
<td></td>
</tr>
<tr>
<td><code>NumTThits &gt; 0</code></td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>q/p</td>
</tr>
<tr>
<td>$20 &lt; M(\mu_W\mu_N) &lt; 70$ GeV</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>jet candidates</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 0.5$</td>
<td></td>
</tr>
<tr>
<td>$p_T &gt; 10$ GeV</td>
<td></td>
</tr>
<tr>
<td>1 track $p_T &gt; 1.2$ GeV, all tracking stations</td>
<td></td>
</tr>
<tr>
<td>fraction of charged particles &gt; 10%</td>
<td></td>
</tr>
<tr>
<td>maximum $p_T$ of a track &gt; 1.2 GeV</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>N candidates</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$M(N) &lt; 80$ GeV</td>
<td></td>
</tr>
<tr>
<td>$p_T(N) &gt; 10$ GeV</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>W candidates</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$60 &lt; M(W) &lt; 100$ GeV</td>
<td></td>
</tr>
</tbody>
</table>

**Table 5.2**: Candidate preselection.
The preselection cuts for $\mu_W$, $\mu_N$, jets, heavy neutrinos and W are summarized in Tab.5.2.

5.3 Signal region

5.4 Impact parameter cut

An impact parameter cut is used to separate prompt $\mu_W$ muon and non-prompt $\mu_N$.

In Fig. 5.2a the impact parameter distribution for $\mu_N$ and $\mu_W$ from the signal Monte Carlo with 15 GeV heavy neutrino is shown. The $\mu_W$ impact parameter has to be less then 0.04 mm, opposite to $\mu_N$, that are required to have IP $> 0.04$ mm.

The $\mu_N$ impact parameter from the prompt heavy neutrino case, see [17], is compared with a long-lived case in Fig. 5.2b.

![Graph](image)

**Figure 5.2:** Muons impact parameter distributions from signal m=15 GeV Monte Carlo.

5.5 Multivariate selection

A multivariate analysis improves dramatically analysis sensitivity and is widely used in particle physics. There are many different tools available: rectangular cut optimization, multidimensional likelihood estimator, linear discriminant analysis (incl. Fischer discriminant), deep neural networks, deep learning, support vector machine, boosted decision trees, etc [55].

In this analysis the boosted decision trees method is used to improve signal/background classification. If an event was wrongly classified, than its importance is increased, by increasing event weight to it, i.e. boosting an event [61]. A decision tree consists of nodes, each of which represents a split in the data based on a particular variable, i.e. classifier, as can be seen in Fig. 5.3.

At each node a candidate fulfills a criteria or not. The last leaf is marked as one of the classes, based on the event majority of the class. Decision trees are used in the high energy physics to classify signal and background events. At each node the most optimal variable to maximize separation of signal and background events is used.
5.5. Multivariate selection

**Figure 5.3:** An example of a decision tree, where $\text{Var}_i$ stands for some variable and $c_i$ stands for some cut value. $S$ stands for signal. $B$ stands for background.
Chapter 5. Reconstruction and selection

The same variable can be used several times at different stages of a tree. A huge advantage of a decision tree is that a signal event can be separated from background events based on a set of the most efficient cuts in a few regions on the parameter space, that were found to be "signal-like". Since decision trees suffer from the over-training, due to the unconfined growing, a deepness of a tree is defined, i.e. number of allowed splits\(^1\). Decision trees can be mislead by the statistical fluctuations in the training data sample. To increase the decision tree resistance to the random fluctuations a forest of trees can be grown, each trained on the same data sample, but with different weights. The weights are assign through boosting, giving a name to the entire technique: boosted decision trees (BDT). Except for an increased stability, a boosted decision tree also increases a separation power of "weak" classifiers, i.e. classifiers are not powerful independently. Non-discriminating input variables are not important for the overall performance of BDT. As well as, a decision tree, BDT should be restrained in their growth to avoid the over-training.

To define the best classifier, i.e. a variable to cut on and the best cut at a particular node, the difference between a separation criteria of the parent node and a sum of separation criteria of two daughter nodes should be maximized:

\[
\max(I_{\text{parent}} - (I_{\text{daughter}_1} + I_{\text{daughter}_2})),
\]

(5.1)

where \(I_{\text{parent}}\) is a parent node separation criteria and \(I_{\text{daughter}_i}\) is an ith daughter node separation criteria. A few separation criteria, used for optimization during the tree growth, are \([55, 62]\):

\[
I = p \cdot (1 - p)
\]

(5.2)

Gini Index

\[
I = -p \cdot \ln(p) - (1 - p) \cdot \ln(1 - p)
\]

(5.3)

Cross entropy

\[
I = 1 - \max(p, 1 - p)
\]

(5.4)

Misclassification error

\[
I = \frac{S}{\sqrt{S + B}}
\]

(5.5)

Statistical significance

\[
I = \frac{S}{\sqrt{B}}
\]

(5.6)

Signal-noise ratio

\[
I = \frac{\varepsilon}{(\sqrt{B} + \sigma/2)}
\]

(5.7)

Punzi criteria

where \(p\) is a purity of a sample\(^2\); \(\varepsilon\) is a signal efficiency; \(\sigma\) is a number of sigma in a Gaussian test, corresponding to a specific significance; \(S\) is a number of signal events; \(B\) is a number of background events.

Although BDT discriminators provide a simple and beautiful solution for events classification, they suffer from a non-uniform signal efficiency as a function of parameter of interest. By applying such discriminators, a risk of getting a biased to a specific parameter of interest hypothesis increases dramatically and, therefore, the

\(^1\)Also known as a stop condition.

\(^2\)The purity is defined as a ratio between a number of signal events and a total number of events in a sample at a particular node.
5.5. Multivariate selection

discriminator is not trustworthy anymore. A special discriminator for BDT was developed called “uBoost”. Classical BDT will optimize themselves by rising wrong identified events weights. “uBoost” BDT (uBDT) additionally assign weights to events based on the signal efficiency uniformity with respect to the parameter of interest. This feature allows to safely use “uBoost” BDT for a multi-hypothesis parameter of interest measurement.

Three uBDT classifiers were trained that they can be used for both prompt and non-prompt heavy neutrino analysis. Therefore, no life-time dependent input variables were chosen. The training is described in [17]. Here I only describe the classifiers.

Two classifiers, Nmu_mc_uBoost and Wmu_uBoost, identify the heavy neutrino daughter muon and the W daughter muon. The third classifier, global_mc_uBoost, exploits candidates kinematics. In Fig. 5.4 a receiver operating characteristic (ROC) curve, a dependence of a background rejection rate and a signal efficiency\(^3\), for the classifiers trained in [17] are shown.

\(^3\)The best performance is achieved, when background rejection rate does not decrease as a function of signal efficiency.

\begin{figure}[h]
\centering
\begin{subfigure}{0.4\textwidth}
\centering
\includegraphics[width=\textwidth]{global_mc_uBoost ROC}
\caption{global_mc_uBoost}
\end{subfigure}
\begin{subfigure}{0.4\textwidth}
\centering
\includegraphics[width=\textwidth]{Nmu_mc_uBoost ROC}
\caption{Nmu_mc_uBoost}
\end{subfigure}
\begin{subfigure}{0.4\textwidth}
\centering
\includegraphics[width=\textwidth]{Wmu_uBoost ROC}
\caption{Wmu_uBoost}
\end{subfigure}
\caption{ROC curves for global_mc_uBoost, Nmu_mc_uBoost and Wmu_uBoost classifiers, taken from [17].}
\end{figure}
Chapter 5. Reconstruction and selection

To train uBDTs, training samples were divided into 5 samples, where 4 would be used for training and one for testing. To avoid biased results, all samples were a testing sample once.

The W\textsubscript{mu} uBoost classifier was trained using as background the b\bar{b} region in data sample and the prompt\textsuperscript{4} signal Monte Carlo sample with different mass points merged as signal. The b\bar{b} region was defined as will be further described in Sec. 5.7 without BDT classifiers cuts applied. The N\textsubscript{mu} mc uBoost and global mc uBoost are trained using weighted background Monte Carlo samples as the background and the prompt signal Monte Carlo sample with different mass points merged as the signal.

For N\textsubscript{mu} mc uBoost the energy over momentum in ECAL and HCAL, three isolation variables IT5, IT3, IT1, the \mu\textsubscript{N} transverse momentum and the \mu\textsubscript{N} PID are used as an input.

Isolation variables are defined as a fraction of the muon transverse momentum in the cone with radius R around the muon:

\begin{equation}
IT1 = \frac{p_{T}^{\mu\textsubscript{N}}}{p_{T}^{cone}} \text{ with } R = 0.1
\end{equation}

\begin{equation}
IT3 = \frac{p_{T}^{\mu\textsubscript{N}}}{p_{T}^{cone}} \text{ with } R = 0.3
\end{equation}

\begin{equation}
IT5 = \frac{p_{T}^{\mu\textsubscript{N}}}{p_{T}^{cone}} \text{ with } R = 0.5,
\end{equation}

where \(p_{T}^{\mu\textsubscript{N}}\) is neutrino muon momentum and \(p_{T}^{cone}\) is a transverse momentum of sum of all particles in the cone with radius R\textsuperscript{5}. Tracks and neutrals from the ParticleFlow are taken as IT isolation variables input. The bigger the value of the isolation variable, the more isolated the muon is.

Input variables of W\textsubscript{mu} uBoost are the energy over momentum in ECAL and HCAL, the \mu\textsubscript{W} PID and the \mu\textsubscript{W} isolation IT5, which is called 0.50_IT. Although, isolation variable follows (5.11), the name is different to account for the different particle input: long tracks and photons.

In Fig. 5.5 isolation variables distributions for different heavy neutrinos are shown. For the \mu\textsubscript{W} isolation and the \mu\textsubscript{N} isolation IT1 no mass dependence is observed. The shape of the \mu\textsubscript{N} IT3 distribution for a 5 GeV heavy neutrino is quite distinctive, compared to other mass points. For higher masses than 5 GeV, \mu\textsubscript{N} are more isolated in terms of IT3. The \mu\textsubscript{N} IT5 is extremely mass dependent. As heavy neutrino gets more massive, the \mu\textsubscript{N} becomes more isolated in terms of IT5. The reason is the following: for the low mass heavy neutrino, muon is likely to emerge closely to the jet and, therefore, the muon is more likely to be considered poorly isolated (low IT3 and IT5). In Fig. 5.6 one can see the angular distance between the muon and the jet for different masses. Angular distance between \mu\textsubscript{N} and jet distribution is defined as:

\begin{equation}
\Delta R(\mu\textsubscript{jet}) = \sqrt{\Delta \phi(\mu\textsubscript{jet}) + \Delta \eta(\mu\textsubscript{jet})}
\end{equation}

\textsuperscript{4}It does not matter, that the prompt sample was used for BDT training and then the BDT was used for non-prompt analysis, since all BDT input variables are lifetime independent.

\textsuperscript{5}First, the scalar sum of all momenta in the cone is computed. Then it’s transverse part is calculated. The cone momentum includes the muon momentum.
where $\phi(jet)$ and $\eta(jet)$ are defined from the direction of the total jet momentum. The small angular distance corresponds to closely spaced $\mu_N$ and jet. Indeed, for the low heavy neutrino mass sample, muon are closely spaced with jets. Note, that in case of IT1, no such effects is observed, since IT1 is a loose isolation variable\(^6\).

Figure 5.5: Muons isolation distribution at the different heavy neutrino masses in the signal Monte Carlo.

For the `global_mc_uBoost` the cosine between two muons in the rest frame of neutrino, the jet transverse momentum, the reconstructed $W$ mass, the missing transverse momentum, the angular distance between $\mu_N$ and jet (see Fig. 5.6), and the invariant mass of two muons are used as input. The missing transverse momentum is defined as a transverse part of a sum of all momentum in the event. According to the momentum conservation law the sum of all particles momentum should be 0. Therefore, the transverse momentum is used to understand whether some particles are lost in the detector. If the missing transverse momentum is non-zero, than there is some particle missing.

\(^6\)Same muon is always more isolated in term of IT1, than in terms of other isolation variables, since the cone size is smaller.
In Fig. 5.6 Angular distance between $\mu_N$ and jet distribution for different mass heavy neutrinos in signal Monte Carlo.

In Fig. 5.7 BDT classifiers distribution for the background and the heavy neutrino $m = 15$ GeV signal are shown. The $W_{\mu u}$ uBoost provides no separation between the $W \to \mu\nu$ background and the signal, contrary to $N_{\mu mc}$ uBoost and $global_{mc}$ uBoost, because both signal and $W \to \mu\nu\mu$ have the muon originating from $W$ decay.

For the signal region the $W_{\mu u}$ Booster has to be over 0.55, i.e. has a good $\mu_W$.

The $N_{\mu mc}$ uBoost and $global_{mc}$ uBoost have to be over 0.61, i.e. $\mu_N$ is good and event is signal-like. The above cuts are applied in case of all heavy neutrino mass hypotheses.

Although, the “uBoost” should have a uniform signal efficiency distribution as a function of the parameter of interest (heavy neutrino mass), it is not expected to be true, since in the list of the input variables many are heavy neutrino mass dependent.

The efficiency is defined as:

$$\epsilon_{\text{classifier}} = \frac{N(\text{passed cut})}{N},$$

where $\epsilon_{\text{classifier}}$ is a classifier cut efficiency; $N(\text{passed cut})$ is a number of passed candidates and $N$ is a total number of candidates. The uncertainty is evaluated as a normal approximation binomial confidence interval, see App. D (D.2).

In Fig. 5.8 classifiers cut efficiency, i.e. “uBoost” signal efficiency, is shown as a function of heavy neutrino mass. For all classifiers efficiency is stable at the low neutrino masses. However, as mass rises to 30 GeV both muon classifiers, the $W_{\mu u}$ uBoost and the $N_{\mu mc}$ uBoost, efficiency changes. The $W_{\mu u}$ uBoost efficiency decreases on about 10%. The $N_{\mu mc}$ uBoost efficiency increases from about 40% for a 20 GeV heavy neutrino to 65%. The $global_{mc}$ uBoost efficiency also experiences some changes in the range of $\sim 45\%$-$55\%$. The total classifiers efficiency, however,
5.5. Multivariate selection

Figure 5.7: Background and signal BDT classifiers distributions prior trigger preselection, IP cuts or BDT cuts applied. For backgrounds corresponding Monte Carlo are used. As a signal the heavy neutrino $m = 15$ GeV Monte Carlo is used.
Figure 5.8: “uBoost” classifiers efficiency as a function of a parameter of interest, i.e. heavy neutrino mass.

 stays relatively the same.

5.6 Control regions

Two control regions are used to verify the analysis procedure, understand background distributions and extrapolate background yields to the signal region:

- $W \rightarrow \mu\nu$ region
- $b\bar{b}$ region.

The control regions selections were taken from [17]. Both regions have same the preselection as the $W$ preselection (see Sec. 5.2), but different cuts on the impact parameters of muons and the uBoost classifiers.

Control regions separating cuts, including the signal region cuts, are listed in Tab. 5.3.

<table>
<thead>
<tr>
<th>signal</th>
<th>$b\bar{b}$</th>
<th>$W \rightarrow \mu\nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP($\mu_W$) &lt; 0.04 [mm]</td>
<td>IP($\mu_W$) &gt; 0.04 [mm]</td>
<td>IP($\mu_W$) &lt; 0.04 [mm]</td>
</tr>
<tr>
<td>IP($\mu_N$) &gt; 0.04 [mm]</td>
<td>IP($\mu_N$) &gt; 0.04 [mm]</td>
<td>IP($\mu_N$) &lt; 0.04 [mm]</td>
</tr>
<tr>
<td>$W_{mu\text{-uBoost}} &gt; 0.55$</td>
<td>$W_{mu\text{-uBoost}} &lt; 0.55$</td>
<td>$W_{mu\text{-uBoost}} &gt; 0.55$</td>
</tr>
<tr>
<td>$N_{mu\text{-mc\text{-uBoost}}} &gt; 0.61$</td>
<td>$N_{mu\text{-mc\text{-uBoost}} &lt; 0.61$</td>
<td>$N_{mu\text{-mc\text{-uBoost}} &lt; 0.61$</td>
</tr>
<tr>
<td>$global_{mc\text{-uBoost}} &gt; 0.61$</td>
<td>$global_{mc\text{-uBoost}} &lt; 0.61$</td>
<td>$global_{mc\text{-uBoost}} &lt; 0.61$</td>
</tr>
</tbody>
</table>

Table 5.3: Control and signal regions separating cuts.
5.7 Other selections

Normalization channel selection and preselection are similar to the \( W \rightarrow \mu \nu \) region selection, see Tab. 5.2 and Tab. 5.3. However, no \( W \) mass window cut is applied, no impact parameter cut is required for the \( \mu_N \) and only \( \mu_{\text{mu} \text{uBoost}} \) classifier cut is applied, compared to \( W \rightarrow \mu \nu \) region selection.

To evaluate the efficiency in the entire kinematic range of both muons two additional channels are used: \( Z \rightarrow \mu^+ \mu^- \) and \( \Upsilon(1S) \rightarrow \mu^+ \mu^- \).

One of the muon candidates from \( Z \rightarrow \mu^+ \mu^- \) channel has to be TOS on three trigger lines: \( \text{L0Muon} \), \( \text{Hlt1SingleMuonHighPT} \) and \( \text{Hlt2SingleMuonHighPT} \). Both muons must have \( p_T \) bigger than 20 GeV and be in the LHCb acceptance (\( 2 < \eta < 4.5 \)). One of the muons must have a low impact parameter (less than 0.04 mm). The combined \( Z \) candidate mass has to be in the mass window of 60 GeV to 120 GeV.

Muon candidates from \( \Upsilon(1S) \rightarrow \mu^+ \mu^- \) channel must have \( p_T \) bigger than 3 GeV and \( 2 < \eta < 4.5 \). The mass window of \( \Upsilon(1S) \) is from 9.3 GeV to 9.7 GeV.

The \( Z \rightarrow \mu^+ \mu^- \) channel and \( \Upsilon(1S) \rightarrow \mu^+ \mu^- \) channel selections are summarized in Tab. 5.4.

| \( Z \rightarrow \mu^+ \mu^- \) | \( \text{L0Muon} \)  
| | \( \text{Hlt1SingleMuonHighPT} \)  
| | \( \text{Hlt2SingleMuonHighPT} \)  
| | \( 2 < \eta(\mu) < 4.5 \)  
| | \( p_T(\mu) > 20 \text{ GeV} \)  
| | \( 60 < M(Z) < 120 \text{ GeV} \)  
| | \( \text{IP}(\mu_N) < 0.04 \text{ mm} \)  

| \( \Upsilon(1S) \rightarrow \mu^+ \mu^- \) | \( 2 < \eta(\mu) < 4.5 \)  
| | \( p_T(\mu) > 3 \text{ GeV} \)  
| | \( 9.3 < M(\Upsilon(1S)) < 9.7 \text{ GeV} \)  

Table 5.4: Preselection and selection cuts for the normalization channel, \( Z \rightarrow \mu^+ \mu^- \) and \( \Upsilon(1S) \rightarrow \mu^+ \mu^- \).

5.8 Background in the control regions

The considered backgrounds are: \( W \rightarrow \mu \nu \), \( Z \rightarrow \mu^+ \mu^- \), \( W \rightarrow \tau \nu \), \( Z \rightarrow \tau^+ \tau^- \), semi-leptonic decays from \( b \bar{b} \) and \( c \bar{c} \). Selection and BDT input variables background compositions were checked in the control regions at different stages of the selection to verify understanding of backgrounds behavior. Monte Carlo background samples, described in Ch. 4, are used to draw the background distributions. For the signal distribution the Monte Carlo sample for 15 GeV heavy neutrino is used.

The \( p_T(\mu_N) \) and \( p_T(\mu_W) \) distributions in the \( b \bar{b} \) and \( W \rightarrow \mu \nu \) control regions are show in Fig. 5.9. In the \( b \bar{b} \) region mainly \( b \bar{b} \) events are left with a tiny bit of a \( c \bar{c} \) contribution. Both \( p_T(\mu_W) \) and \( p_T(\mu_N) \) distributions in \( b \bar{b} \) region closely describe data distributions. The \( W \rightarrow \mu \nu \) region consists of electroweak contributions: mainly \( W \rightarrow \mu \nu \), a bit less of \( W \rightarrow \tau \nu \) and \( Z \rightarrow \mu^+ \mu^- \). Similarities between the signal and the \( W \rightarrow \mu \nu \) channel result in the survival of some signal events in the \( W \rightarrow \mu \nu \)
Chapter 5. Reconstruction and selection

region. The $p_T(\mu_W)$ and $p_T(\mu_N)$ background distributions look similarly to the data shape.

![Graphs showing $p_T$ distributions in control regions](image)

**Figure 5.9:** $p_T$ distribution in the control regions.

The BDT classifiers distributions are checked in Fig. 5.10.

The $N_{\mu\text{mc\_uBoost}}$ background composition is reasonably matching the data. There are some small fluctuations in the tails of the $N_{\mu\text{mc\_uBoost}}$ distribution in both control regions.

$W_{\mu\text{uBoost}}$ distributions in the $b\bar{b}$ and the $W \rightarrow \mu\nu$ regions are very different compare to each other, because both regions cut exactly the opposite ranges of $W_{\mu\text{uBoost}}$. Some fluctuations are seen in $W_{\mu\text{uBoost}}$ distribution in $W \rightarrow \mu\nu$ region around middle values of the $W_{\mu\text{uBoost}}$.

For the $\text{global\_mc\_uBoost}$ backgrounds are in good agreement with data. BDT input variables are also checked. The IT5, IT3, IT1, $M(\mu_1\mu_2)$ and missing $p_T$ distributions are shown in Fig. 5.12, 5.13, 5.14.

In Fig. 5.11 one can see background distribution of isolation variables prior trigger, BDT and IP cuts. Distributions give an idea of how well they separate the background from the signal. It is easy to see from Fig. 5.11, which isolation variables are
5.8. Background in the control regions

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figs/figure5_10.png}
\caption{BDT classifiers distribution in the control regions.}
\end{figure}
likely to contribute more into the final classifier. In the signal region for the 15 GeV heavy neutrino, both muons have to be highly isolated\textsuperscript{7}. In the $W \rightarrow \mu\nu$ region $\mu_W$ should be isolated, however, $\mu_N$ should not, since it is in some sense “a background” muon, that should not be in the $W \rightarrow \mu\nu$. In the $b\bar{b}$ both muons should not be well isolated.

The $\mu_W$ IT5 isolation (0.50 IT), as expected, has a small separation power between the signal and $W \rightarrow \mu\nu$, and a clear separation of $b\bar{b}$ with respect to the signal and $W \rightarrow \mu\nu$.

The $\mu_N$ IT5 and the $\mu_N$ IT3 separate well backgrounds and signal. The signal is much more isolated, compared to the backgrounds.

For the $\mu_N$ IT1 signal is still highly isolated, but since, the IT1 is a looser isolation variable, backgrounds with a poorly isolated $\mu_N$, have a higher measured isolation.

\textbf{Figure 5.11:} Background and signal isolation variables distributions prior trigger preselection, IP cuts or BDT cuts applied. As a signal heavy neutrino $m = 15$ GeV Monte Carlo is used.

\textsuperscript{7}Note, that this is not the case for a 5 GeV heavy neutrino, where the $\mu_N$ is not well separated from the jet.
5.8. Background in the control regions

Isolation variables after the full selection is applied are shown in Fig. 5.12 and Fig. 5.13.

In the $b\bar{b}$ region the $\mu_W$ 0.5_IT background composition is biased to the right compared to the actual data, see Fig. 5.12. This is a hint of the differences in the Monte Carlo and the data. If one looks at the Fig. 5.10, this difference does not seem to affect the resulting $\mu_{\text{w Boost}}$ classifier, which is actually used in the selection. However, it should be understood in the follow-up research. Also, the $\mu_W$ 0.5_IT changes shape in the $W \rightarrow \mu\nu$ region compare to the $b\bar{b}$ region. This is expected, since the $\mu_{\text{w Boost}}$ cut isolates $b\bar{b}$ and $W \rightarrow \mu\nu$ regions very well and one sees the corresponding background peaks from Fig. 5.11 separated.

The $\mu_N$ IT5 background contribution adequately follows the actual data in the $b\bar{b}$ region with some tail fluctuations. However, in the $W \rightarrow \mu\nu$ region backgrounds peak is narrower compared to the one observed in data. This also should be studied.

The $\mu_W$ IT5, $b\bar{b}$ region

The $\mu_N$ IT3 and IT1 isolation variables are shown in Fig. 5.13.

There are quite some $\mu_N$ IT1 fluctuations in the $b\bar{b}$ region and the $W \rightarrow \mu\nu$ region.

There is underestimation in the background of the muons with IT1 close to 1 in both

![Figure 5.12: IT5 distribution in the control regions.](image-url)

(A) $\mu_W$ IT5, $b\bar{b}$ region  
(B) $\mu_W$ IT5, $W \rightarrow \mu\nu$ region

(C) $\mu_N$ IT5, $b\bar{b}$ region  
(D) $\mu_N$ IT5, $W \rightarrow \mu\nu$ region
Chapter 5. Reconstruction and selection

There is a clear discrepancy in the tails of $\mu_N$ IT3 distributions, especially in the $b\bar{b}$ region. This indicates some background (around 12%) in the high IT3, that shares some similarity with $b\bar{b}$ and has not been yet understood.

![Graphs showing distributions of $\mu_N$ IT1 and IT3 in control regions.](image)

**Figure 5.13:** IT1 and IT3 distribution in the control regions.

Distributions of $M(\mu\mu)$ and missing $p_T$ are shown in Fig. 5.14. These variables enter the `global_mc_uBoost` classifier.

$M(\mu\mu)$ background distributions are fairly close to the data distribution and have similar shapes in both control regions.

In the $W \rightarrow \mu\nu$ the missing $p_T$ corresponds to the neutrino $p_T$ in the event. In the $b\bar{b}$ region the missing $p_T$ is naturally lower, than in the $W \rightarrow \mu\nu$ region, where one expects a neutrino-like candidate. Matching of background distributions with data is fairly good in both regions.
5.8. Background in the control regions

![Graphs showing M(µµ) and missing p_T distributions in control regions](image)

**Figure 5.14:** $M(\mu\mu)$ and missing $p_T$ distribution in the control regions.
Chapter 6

Efficiency

Generation efficiency, reconstruction and stripping efficiency and selection efficiency are taken into account in total efficiency. Generation efficiency can be found in Tab. 4.2. To correct for the data and Monte Carlo discrepancies, correction factor are calculated: a global event cut correction, muon reconstruction, muon identification and $\mu_W$ trigger corrections and muon “uBoost” classifiers corrections. The discussion on the efficiency uncertainty calculation can be found in the App. D.

6.1 Global event cut correction

A charged particle leaves a hit in the scintillator pad detector (SPD). All events in data were forced to have less than 600 SPD hits on the L0 level. The same number of SPD hits cut in Monte Carlo has a different efficiency. To correct for this difference between data and Monte Carlo, one computes a correction factor. It is computed as following. First the efficiency of $n_{SPD}$ hits cut in the L0 trigger ($n_{SPDhits} < 600$) is computed in the $Z \rightarrow \mu^+\mu^-$ data sample with a Z-selection applied, excluding trigger and IP cuts (see Tab. 5.4). Using the $Z \rightarrow \mu^+\mu^-$ Monte Carlo simulation a new $n_{SPD}$ hits cut: $n_{SPDhits} < N$, that corresponds to the same efficiency as in the data, is found. This new cut is then applied to both the normalization channel ($W \rightarrow \mu\nu$) and to the signal channel ($W^{\pm} \rightarrow \mu^{\pm}\mu^{\pm} + jet$). Then a correction factor, $c_{\text{GEC}}$, is calculated:

$$c_{\text{GEC}} = \frac{\varepsilon(n_{SPDhits} < N)}{\varepsilon(n_{SPDhits} < 600)}$$

where $c_{\text{GEC}}$ is a global event cut correction, $\varepsilon(n_{SPDhits} < 600)$ is an efficiency of the original L0 number of SPD hits cut and $\varepsilon(n_{SPDhits} < N)$ is an efficiency of a new SPD hits cut, defined as described above.

The number of SPD hits distribution for events with particular number of the primary vertices is shown in Fig. 6.1a. As expected, the mean of the distribution shifts to the right with the increase of the number of primary vertices per event.

The $n_{SPD}$ hits increases proportionally to the increasing of number of primary vertices and on average a hundred SPD hits per one extra primary vertex are added, as can be seen in Fig. 6.1b. The mean number of SPD hits starts to deviate from the linear dependence on the number of the PVs for the high number of PVs. This is a consequence of the looser cut $n_{SPDhits} < 900$, that comes from the DiMuon trigger lines which also trigger on $Z \rightarrow \mu^+\mu^-$ signal. It cuts the tails of high PV distributions. Another effect, that influences the high number of primary vertices results is the saturation of the SPD.

The distribution of events per number of PV for the data and the Monte Carlo is the same, as can be seen from Fig. 6.2. Only 4 % of a sample has events with
(A) nSPD hits in $Z \rightarrow \mu^+\mu^-$ data with different nPVs before selection.

(B) Mean nSPD hits as a function of nPVs in $Z \rightarrow \mu^+\mu^-$ data before selection.

**Figure 6.1**: Number of SPD hits dependence on the number of primary vertices.

**Figure 6.2**: Proportional distribution of events with different number of primary vertices (nPVs) for the $Z \rightarrow \mu^+\mu^-$ data sample.
6.1. Global event cut correction

Efficiency of the original nSPD hits cut and a new determined cut is shown in Tab. 6.2.

<table>
<thead>
<tr>
<th>nPVs</th>
<th>$\varepsilon$(nSPDhits &lt; 600)</th>
<th>nSPD hits cut</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PV</td>
<td>99.982 ± 0.007</td>
<td>407±8</td>
</tr>
<tr>
<td>2 PV</td>
<td>98.777 ± 0.047</td>
<td>398±1</td>
</tr>
<tr>
<td>3 PV</td>
<td>92.566 ± 0.136</td>
<td>398±1</td>
</tr>
<tr>
<td>4 PV</td>
<td>78.664 ± 0.329</td>
<td>402±1</td>
</tr>
<tr>
<td>5 PV</td>
<td>59.329 ± 0.753</td>
<td>408±1</td>
</tr>
<tr>
<td>&gt; 5 PV</td>
<td>39.374 ± 1.505</td>
<td>427±2</td>
</tr>
<tr>
<td>All</td>
<td>94.026 ± 0.061</td>
<td>403±1</td>
</tr>
</tbody>
</table>

Table 6.1: Efficiency of the number of SPD hits original L0 cut, retrieved from $Z \rightarrow \mu^+ \mu^-$ data and a new cut, that is found by requiring the same efficiency in the $Z \rightarrow \mu^+ \mu^-$ Monte Carlo, as in the $Z \rightarrow \mu^+ \mu^-$ data.

An uncertainty on the new nSPD hits cut has been calculated from the uncertainty on the efficiency $\Delta \varepsilon$ by finding the cut that corresponds to $\varepsilon \pm \Delta \varepsilon$.

The dependence of a new cut on the nSPD hits and the number of SPD hits is shown in Fig. 6.3.

From the calculation without PV separation the correspondent cut in $Z \rightarrow \mu^+ \mu^-$ Monte Carlo was found to be 403±1. Since the Fig. 6.2 showed consistency between data and Monte Carlo number of primary vertices distributions, the new nSPD hits cut is taken without primary vertex separation. This cut is used to evaluate the final correction. The corrections can be found in Tab. 6.2.
### Table 6.2: Efficiency of the new number of SPD hits cut (see 6.1 and correction for the normalization W and the signal \( W \rightarrow \mu \nu \) and the signal \( W \rightarrow \mu \mu \) jet channel.

<table>
<thead>
<tr>
<th>Correction</th>
<th>0.000 ( \pm ) 0.014%</th>
<th>0.000 ( \pm ) 0.014%</th>
<th>0.000 ( \pm ) 0.014%</th>
<th>0.000 ( \pm ) 0.014%</th>
<th>0.000 ( \pm ) 0.014%</th>
<th>0.000 ( \pm ) 0.014%</th>
<th>0.000 ( \pm ) 0.014%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>62.492 ( \pm ) 0.637%</td>
<td>62.492 ( \pm ) 0.637%</td>
<td>62.492 ( \pm ) 0.637%</td>
<td>62.492 ( \pm ) 0.637%</td>
<td>62.492 ( \pm ) 0.637%</td>
<td>62.492 ( \pm ) 0.637%</td>
<td>62.492 ( \pm ) 0.637%</td>
</tr>
<tr>
<td>5 PV</td>
<td>42.342 ( \pm ) 4.623%</td>
<td>42.342 ( \pm ) 4.623%</td>
<td>42.342 ( \pm ) 4.623%</td>
<td>42.342 ( \pm ) 4.623%</td>
<td>42.342 ( \pm ) 4.623%</td>
<td>42.342 ( \pm ) 4.623%</td>
<td>42.342 ( \pm ) 4.623%</td>
</tr>
<tr>
<td></td>
<td>83.817 ( \pm ) 1.370%</td>
<td>83.817 ( \pm ) 1.370%</td>
<td>83.817 ( \pm ) 1.370%</td>
<td>83.817 ( \pm ) 1.370%</td>
<td>83.817 ( \pm ) 1.370%</td>
<td>83.817 ( \pm ) 1.370%</td>
<td>83.817 ( \pm ) 1.370%</td>
</tr>
<tr>
<td>4 PV</td>
<td>63.429 ( \pm ) 2.575%</td>
<td>63.429 ( \pm ) 2.575%</td>
<td>63.429 ( \pm ) 2.575%</td>
<td>63.429 ( \pm ) 2.575%</td>
<td>63.429 ( \pm ) 2.575%</td>
<td>63.429 ( \pm ) 2.575%</td>
<td>63.429 ( \pm ) 2.575%</td>
</tr>
<tr>
<td></td>
<td>40.708 ( \pm ) 4.623%</td>
<td>40.708 ( \pm ) 4.623%</td>
<td>40.708 ( \pm ) 4.623%</td>
<td>40.708 ( \pm ) 4.623%</td>
<td>40.708 ( \pm ) 4.623%</td>
<td>40.708 ( \pm ) 4.623%</td>
<td>40.708 ( \pm ) 4.623%</td>
</tr>
<tr>
<td>3 PV</td>
<td>58.416 ( \pm ) 0.164%</td>
<td>58.416 ( \pm ) 0.164%</td>
<td>58.416 ( \pm ) 0.164%</td>
<td>58.416 ( \pm ) 0.164%</td>
<td>58.416 ( \pm ) 0.164%</td>
<td>58.416 ( \pm ) 0.164%</td>
<td>58.416 ( \pm ) 0.164%</td>
</tr>
<tr>
<td></td>
<td>27.509 ( \pm ) 7.239%</td>
<td>27.509 ( \pm ) 7.239%</td>
<td>27.509 ( \pm ) 7.239%</td>
<td>27.509 ( \pm ) 7.239%</td>
<td>27.509 ( \pm ) 7.239%</td>
<td>27.509 ( \pm ) 7.239%</td>
<td>27.509 ( \pm ) 7.239%</td>
</tr>
<tr>
<td>2 PV</td>
<td>96.047 ( \pm ) 0.613%</td>
<td>96.047 ( \pm ) 0.613%</td>
<td>96.047 ( \pm ) 0.613%</td>
<td>96.047 ( \pm ) 0.613%</td>
<td>96.047 ( \pm ) 0.613%</td>
<td>96.047 ( \pm ) 0.613%</td>
<td>96.047 ( \pm ) 0.613%</td>
</tr>
<tr>
<td></td>
<td>82.566 ( \pm ) 0.738%</td>
<td>82.566 ( \pm ) 0.738%</td>
<td>82.566 ( \pm ) 0.738%</td>
<td>82.566 ( \pm ) 0.738%</td>
<td>82.566 ( \pm ) 0.738%</td>
<td>82.566 ( \pm ) 0.738%</td>
<td>82.566 ( \pm ) 0.738%</td>
</tr>
<tr>
<td>1 PV</td>
<td>65.356 ( \pm ) 1.370%</td>
<td>65.356 ( \pm ) 1.370%</td>
<td>65.356 ( \pm ) 1.370%</td>
<td>65.356 ( \pm ) 1.370%</td>
<td>65.356 ( \pm ) 1.370%</td>
<td>65.356 ( \pm ) 1.370%</td>
<td>65.356 ( \pm ) 1.370%</td>
</tr>
<tr>
<td></td>
<td>43.158 ( \pm ) 0.533%</td>
<td>43.158 ( \pm ) 0.533%</td>
<td>43.158 ( \pm ) 0.533%</td>
<td>43.158 ( \pm ) 0.533%</td>
<td>43.158 ( \pm ) 0.533%</td>
<td>43.158 ( \pm ) 0.533%</td>
<td>43.158 ( \pm ) 0.533%</td>
</tr>
<tr>
<td>0 PV</td>
<td>49.708 ( \pm ) 1.370%</td>
<td>49.708 ( \pm ) 1.370%</td>
<td>49.708 ( \pm ) 1.370%</td>
<td>49.708 ( \pm ) 1.370%</td>
<td>49.708 ( \pm ) 1.370%</td>
<td>49.708 ( \pm ) 1.370%</td>
<td>49.708 ( \pm ) 1.370%</td>
</tr>
<tr>
<td></td>
<td>22.759 ( \pm ) 7.239%</td>
<td>22.759 ( \pm ) 7.239%</td>
<td>22.759 ( \pm ) 7.239%</td>
<td>22.759 ( \pm ) 7.239%</td>
<td>22.759 ( \pm ) 7.239%</td>
<td>22.759 ( \pm ) 7.239%</td>
<td>22.759 ( \pm ) 7.239%</td>
</tr>
<tr>
<td>All</td>
<td>86.978 ( \pm ) 0.636%</td>
<td>86.978 ( \pm ) 0.636%</td>
<td>86.978 ( \pm ) 0.636%</td>
<td>86.978 ( \pm ) 0.636%</td>
<td>86.978 ( \pm ) 0.636%</td>
<td>86.978 ( \pm ) 0.636%</td>
<td>86.978 ( \pm ) 0.636%</td>
</tr>
</tbody>
</table>
6.2 Muon corrections

To correct for the muon identification, reconstruction and trigger difference in the data and Monte Carlo, corresponding correction factors were taken into account. The correction factor is defined as:

\[ c = \frac{\varepsilon_{data}}{\varepsilon_{MC}} \quad (6.2) \]

where \( c \) is a correction factor, \( \varepsilon_{data} \) is a data-driven efficiency and \( \varepsilon_{MC} \) is a Monte Carlo computed efficiency.

The total muon correction factor is:

\[ c_{\text{muon}} = c_{\text{rec}} \cdot c_{\text{ID}} \cdot c_{\text{trigger}} \quad (6.3) \]

where \( c_{\text{muon}} \) is a total muon correction factor, \( c_{\text{rec}} = c_{\text{rec}}(\mu_N) \cdot c_{\text{rec}}(\mu_W) \) is a reconstruction correction factor for both muons, \( c_{\text{ID}} = c_{\text{ID}}(\mu_N) \cdot c_{\text{ID}}(\mu_W) \) is an identification correction factor for both muons and \( c_{\text{trigger}} \) is a trigger correction factor for \( \mu_W \).

Corrections are computed using the tag-and-probe method in \( Z \rightarrow \mu^+ \mu^- \) channel, as described in [63]. In Tab. 6.3 one can find the corrections applied to the efficiency for the signal channel and for the normalization channel.

6.3 BDT corrections

The BDT might perform differently on the data and on the Monte Carlo. To correct for this effect a correction is applied. For both classifiers the correction factor is defined as:

\[ c_{BDT} = \frac{\varepsilon_{BDT}(\text{data})}{\varepsilon_{BDT}(\text{MC})} \quad (6.4) \]

where \( \varepsilon_{BDT}(\text{MC}) \) is a Monte Carlo computed classifier cut efficiency and \( \varepsilon_{BDT}(\text{data}) \) is a data-driven classifier cut efficiency. A classifier cut efficiency is defined as:

\[ \varepsilon_{BDT} = \frac{N(\text{presel} + \text{BDT cut})}{N(\text{presel})} \quad (6.5) \]

where \( N(\text{presel} + \text{BDT cut}) \) is a number of events that passed the signal region classifier cut and channel preselection; \( N(\text{presel}) \) is a number of events that passed channel preselection.

To account for the kinematic differences between muons different decay channels are used to estimate the above classifier cut efficiency.

The correction estimation procedure for the \( \tilde{\nu}_\mu \_u\text{Boost} \) classifier follows [17]. For the \( \tilde{\nu}_\mu \_u\text{Boost} \) classifier \( Z \rightarrow \mu^+ \mu^- \) channel is used. \( Z \rightarrow \mu^+ \mu^- \) has a high signal purity, so the cut and count approach was adopted. Only the \( Z \rightarrow \mu^+ \mu^- \) preselection, described in Tab. 5.4 is applied to the data and Monte Carlo samples. The correction is calculated in bins of \( \eta \) and \( p_T \). The total correction is weighted according to \( \mu_W \) kinematics. Corrections in bins of \( \eta \) and \( p_T \) for \( \tilde{\nu}_\mu \_u\text{Boost} \) classifier are shown in Fig. 6.4. The total weighted corrections can be found in Tab. 6.4.

For the \( \tilde{\nu}_\mu \_m\text{c}_\mu \_u\text{Boost} \) classifier a correction is computed using two kinematic samples. To cover the low-\( p_T \) range \( Y(1S) \rightarrow \mu^+ \mu^- \) channel is used and \( Z \rightarrow \mu^+ \mu^- \) is used for the high-\( p_T \) range. Results from \( Z \rightarrow \mu^+ \mu^- \) channel are shown in Fig. 6.7a.
<table>
<thead>
<tr>
<th>Energy (GeV)</th>
<th>Signal 1</th>
<th>Signal 2</th>
<th>Signal 3</th>
<th>Signal 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.977 ± 0.002</td>
<td>0.977 ± 0.002</td>
<td>0.977 ± 0.002</td>
<td>0.977 ± 0.002</td>
</tr>
<tr>
<td>0.500</td>
<td>1.062 ± 0.004</td>
<td>1.062 ± 0.004</td>
<td>1.062 ± 0.004</td>
<td>1.062 ± 0.004</td>
</tr>
<tr>
<td>1.000</td>
<td>1.000 ± 0.004</td>
<td>1.000 ± 0.004</td>
<td>1.000 ± 0.004</td>
<td>1.000 ± 0.004</td>
</tr>
<tr>
<td>1.500</td>
<td>1.058 ± 0.003</td>
<td>1.058 ± 0.003</td>
<td>1.058 ± 0.003</td>
<td>1.058 ± 0.003</td>
</tr>
<tr>
<td>2.000</td>
<td>1.057 ± 0.003</td>
<td>1.057 ± 0.003</td>
<td>1.057 ± 0.003</td>
<td>1.057 ± 0.003</td>
</tr>
<tr>
<td>2.500</td>
<td>1.056 ± 0.003</td>
<td>1.056 ± 0.003</td>
<td>1.056 ± 0.003</td>
<td>1.056 ± 0.003</td>
</tr>
<tr>
<td>3.000</td>
<td>1.055 ± 0.003</td>
<td>1.055 ± 0.003</td>
<td>1.055 ± 0.003</td>
<td>1.055 ± 0.003</td>
</tr>
<tr>
<td>3.500</td>
<td>1.054 ± 0.003</td>
<td>1.054 ± 0.003</td>
<td>1.054 ± 0.003</td>
<td>1.054 ± 0.003</td>
</tr>
<tr>
<td>4.000</td>
<td>1.053 ± 0.003</td>
<td>1.053 ± 0.003</td>
<td>1.053 ± 0.003</td>
<td>1.053 ± 0.003</td>
</tr>
<tr>
<td>4.500</td>
<td>1.052 ± 0.003</td>
<td>1.052 ± 0.003</td>
<td>1.052 ± 0.003</td>
<td>1.052 ± 0.003</td>
</tr>
<tr>
<td>5.000</td>
<td>1.051 ± 0.003</td>
<td>1.051 ± 0.003</td>
<td>1.051 ± 0.003</td>
<td>1.051 ± 0.003</td>
</tr>
</tbody>
</table>

**Table 6.3:** Muon correction factors for the signal and for the normalization channels.
3. BDT corrections

Y(1S) → μ⁺μ⁻ preselection from Tab. 5.4 is applied. Unfortunately, Y(1S) → μ⁺μ⁻ has a significant background contribution, therefore, a simple cut and count technique is not applicable. To retrieve signal yields from the Y(1S) → μ⁺μ⁻ data and Monte Carlo samples the M(μμ) mass spectrum is fitted in the mass window defined by the Y(1S) → μ⁺μ⁻ preselection, see Tab. 5.4.

The signal is fitted with the Ipatia2 shape \[64\]. It is a double Crystal Ball-like function with a hyperbolic core, that allows to fit a mass spectrum with unknown/different per-event mass uncertainties.

Ipatia2 is defined as:

\[
G(\mu, \sigma, \lambda, \xi, \beta) = \begin{cases} 
G(m, \mu, \sigma, \lambda, \xi, \beta) & -\alpha_1 < \frac{m - \mu}{\sigma} < \alpha_2 \\
\frac{G(m, \mu, \sigma, \lambda, \xi, \beta)}{(1 - m/(n_1 - \mu, \sigma, \lambda, \xi, \beta))^{n_1}} & \frac{m - \mu}{\sigma} \leq -\alpha_1 \\
\frac{G(m, \mu, \sigma, \lambda, \xi, \beta)}{(1 - m/(n_2 - \mu, \sigma, \lambda, \xi, \beta))^{n_2}} & \frac{m - \mu}{\sigma} \geq \alpha_2
\end{cases}
\]

(6.6)

where \(I2(m, \mu, \sigma, \lambda, \xi, \beta, \alpha_1, n_1, \alpha_2, n_2)\) is the Ipatia2 function; \(N\) is a normalization; \(G(m, \mu, \sigma, \lambda, \xi, \beta)\) is a generalized hyperbolic function; \(\alpha_1\) and \(n_1\) are left-side tail parameters; \(\alpha_2\) and \(n_2\) are right-side tail parameters.

A generalized hyperbolic distribution \(G(m, \mu, \sigma, \lambda, \xi, \beta)\) is defined as:

\[
G(m, \mu, \sigma, \lambda, \xi, \beta) \propto ((m - \mu)^2 + A^2(\xi)\sigma^2)^{\frac{1}{2}} e^{\beta(m - \mu) \lambda} A^{-\frac{1}{2}}(\xi) \sqrt{1 + \frac{(m - \mu)}{A^2(\xi)\sigma^2}}
\]

(6.7)

**FIGURE 6.4:** Wmu_boost correction in bins of η with kinematics evaluated using Z → μ⁺μ⁻ channel.
where \( m \) is mean; \( \mu \) is a location shift of the mean; \( \sigma \) is a rms in case of a symmetric function; \( K_{\lambda - \frac{1}{2}} \) is a modified Bessel function of a second kind; \( \lambda \) is a Bessel function order; \( \zeta \) = \( a \delta \), where \( \delta \) is a scale parameter and \( \beta \) is an asymmetry parameter.

The background contribution is fitted with a 2nd order polynomial:

\[
f(x) = a + bx + cx^2. \tag{6.8}
\]

The resulting fits are shown in Fig. 6.5 for both data and Monte Carlo \( Y(1S) \rightarrow \mu^+ \mu^- \) unbinned sample.

Analogously, fits are performed in the bins of \( p_T \) and \( \eta \). Then the correction is calculated.

Per \( p_T \) and \( \eta \) statistics in \( Y(1S) \rightarrow \mu^+ \mu^- \) Monte Carlo and data with the \( \text{Nmu mc uBoost} \) cut applied is shown in Fig. 6.6. Bins with less than 10 events left, are considered low statistics and are not used for correction estimation.

The resulting corrections for the low-\( p_T \) bins can be found in Fig. 6.7b.
The $\mu\text{mc}_\text{uBoost}$ classifier corrections from $Y(1S) \rightarrow \mu^+\mu^-$ and $Z \rightarrow \mu^+\mu^-$ are then weighted according to the $\mu N$ kinematics. The total weighted $\mu\text{mc}_\text{uBoost}$ corrections can be found in Tab. 6.4.

### 6.4 Reconstruction and stripping, and selection efficiency

The reconstruction and stripping efficiency is defined as:

$$\epsilon_{\text{reco+strip}} = \frac{N(\text{recon})}{N(\text{gen})},$$

(6.9)

where $\epsilon_{\text{reco+strip}}$ is the reconstruction and stripping efficiency; $N(\text{recon})$ and $N(\text{gen})$ are a number of events after reconstruction and a number of generated events.

The selection efficiency is defined as:

$$\epsilon_{\text{sel}} = \frac{N(\text{sel})}{N(\text{tot})},$$

(6.10)

where $N(\text{sel})$ and $N(\text{tot})$ are number of events after and before the selection and the preselection, described in Tab. 5.2.

Results for both reconstruction and stripping, and selection efficiency can be found in Tab. 6.5.

The total efficiency is defined as:

$$\epsilon_{\text{tot}} = \epsilon_{\text{gen}} \cdot \epsilon_{\text{reco+strip}} \cdot \epsilon_{\text{sel}} \cdot \epsilon_{\text{GEC}} \cdot \epsilon_{\text{muon}} \cdot \epsilon_{\text{BDT}}.$$

(6.11)

For the generation efficiency $\epsilon_{\text{gen}}$ see Tab. 4.2.

The total efficiency, including the corrections, is summarized in Tab. 6.6. The total efficiency uncertainty is defined as summed in the quadrature uncertainties of component efficiency, excluding corrections uncertainty, which is used as systematic uncertainty.
<table>
<thead>
<tr>
<th>Signal [GeV]</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>30</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>W → µν</td>
<td>0.985 ± 0.002</td>
<td>0.985 ± 0.003</td>
<td>0.985 ± 0.002</td>
<td>0.984 ± 0.002</td>
<td>0.980 ± 0.003</td>
<td>0.963 ± 0.008</td>
</tr>
<tr>
<td>Nμν mc uBoost</td>
<td>0.923 ± 0.007</td>
<td>0.992 ± 0.008</td>
<td>0.958 ± 0.007</td>
<td>0.966 ± 0.007</td>
<td>1.011 ± 0.008</td>
<td>0.909 ± 0.007</td>
</tr>
</tbody>
</table>

Table 6.4: Classifiers correction factors for signal and normalization channels.
6.4. Reconstruction and stripping, and selection efficiency

<table>
<thead>
<tr>
<th></th>
<th>reco+strip</th>
<th>selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal 5 [GeV]</td>
<td>11.497 ± 0.144 %</td>
<td>5.030 ± 0.413 %</td>
</tr>
<tr>
<td>Signal 10 [GeV]</td>
<td>10.958 ± 0.140 %</td>
<td>8.370 ± 0.527 %</td>
</tr>
<tr>
<td>Signal 15 [GeV]</td>
<td>10.649 ± 0.138 %</td>
<td>10.286 ± 0.585 %</td>
</tr>
<tr>
<td>Signal 20 [GeV]</td>
<td>10.494 ± 0.137 %</td>
<td>9.562 ± 0.572 %</td>
</tr>
<tr>
<td>Signal 30 [GeV]</td>
<td>10.619 ± 0.150 %</td>
<td>9.657 ± 0.623 %</td>
</tr>
<tr>
<td>Signal 50 [GeV]</td>
<td>10.644 ± 0.138 %</td>
<td>4.881 ± 0.433 %</td>
</tr>
<tr>
<td>Normalization</td>
<td>60.782 ± 0.022 %</td>
<td>61.290 ± 0.028 %</td>
</tr>
</tbody>
</table>

**Table 6.5:** Reconstruction and stripping efficiency.

<table>
<thead>
<tr>
<th></th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal 5 [GeV]</td>
<td>0.095 ± 0.008 %</td>
</tr>
<tr>
<td>Signal 10 [GeV]</td>
<td>0.151 ± 0.010 %</td>
</tr>
<tr>
<td>Signal 15 [GeV]</td>
<td>0.167 ± 0.010 %</td>
</tr>
<tr>
<td>Signal 20 [GeV]</td>
<td>0.146 ± 0.009 %</td>
</tr>
<tr>
<td>Signal 30 [GeV]</td>
<td>0.146 ± 0.010 %</td>
</tr>
<tr>
<td>Signal 50 [GeV]</td>
<td>0.061 ± 0.005 %</td>
</tr>
<tr>
<td>Normalization</td>
<td>8.798 ± 0.890 %</td>
</tr>
</tbody>
</table>

**Table 6.6:** Total efficiency with corrections included.
Chapter 7

Uncertainties

Systematic uncertainties come from the intrinsic imperfectness of a detection system or method used in the analysis. Important sources of efficiency systematic uncertainties include uncertainties from efficiency corrections, momentum calibration systematic uncertainty, jet energy scale, jet energy resolution and jet identification. A systematic uncertainty on the observable $Y = f(X_1, X_2, ..., X_i, ..., X_n)$, where each $X_i$ represents an input quantity, can be defined as following, assuming that $Y$ is linearly dependent on $X_i$:

$$\sigma^2_{sys}(Y) = \sum_i \Delta_i^2 + 2 \sum_{l<m} \rho_{lm} \Delta_l \Delta_m \tag{7.1}$$

where $\Delta_i$ is a symmetric $^1$ $Y$ variation due to the variation of $X_i$ on $\pm 1\sigma$ and $\rho_{lm}$ is a correlation coefficient between $X_l$ and $X_m$. In case $X_i$ are uncorrelated the second term in (7.1) is zero.

7.1 Corrections

Errors from efficiency corrections described in Ch.6 are treated as systematic uncertainty. Those corrections include: global event cut correction, muon ID, muon reconstruction, muon trigger correction, $W_{\mu} uBoost$ classifier correction and $N_{\mu mc} uBoost$ classifier correction. Systematic uncertainty from muon corrections is taken from the total results in Tab. 6.3. Systematic uncertainties from each correction are shown in Tab. 7.1.

7.2 Momentum calibration

In 2012 data, a global curvature bias lead to the incorrect momentum reconstruction [65]. This affects $\mu$ momentum distribution, which is used later to extract the $W \rightarrow \mu\nu$ yields. The momentum can be calibrated to account for the above affect, as described in App. C. A conservative systematic uncertainty is derived from the difference on the $W$ yield with and without applied momentum calibration. The total uncertainty is 2.65 %.

7.3 Jet energy scale, jet energy resolution and jet identification

The jet energy is measured in the calorimeter system. Unfortunately, the jet energy measurement is affected by a few effects that can change the final quantity,

$^1$If $\Delta_i^{+1\sigma} \neq \Delta_i^{-1\sigma}$ then in 7.1 $\Delta_i^2$ is replaced with an averaged $\bar{\Delta}_i^2$, where $\bar{\Delta} = (\Delta_i^{+1\sigma} + \Delta_i^{-1\sigma})/2$
like calorimeter undetected particles (neutrals). These problems lead to the discrepancy between the reconstructed jet energy, $E_{\text{rec}}^\text{jet}$, and the true energy of the original quark that generated the jet, $E_{\text{true}}^\text{jet}$.

The jet energy scale, $R$, is a correction factor (systematic bias) applied to the reconstructed jet energy to get the true energy of the jet:

$$E_{\text{rec}}^\text{jet} = R \cdot E_{\text{true}}^\text{jet} \quad (7.2)$$

The measured jet energy also has an uncertainty associated with the calorimeter energy resolution. The jet energy scale and resolution uncertainty are taken from [17] and are shown in Tab. 7.1.

Another systematic uncertainty originates from the jet identification. The jet identification uncertainty is taken from [66] and equals 1.7%.

### 7.4 Total systematic uncertainty

The total systematic uncertainty is summarized in Tab. 7.1. The main contribution comes from the jet energy scale uncertainty.
7.4. Total systematic uncertainty

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>global event cut</td>
<td>0.795 %</td>
<td>0.794 %</td>
<td>0.798 %</td>
<td>0.804 %</td>
<td>0.918 %</td>
<td>0.918 %</td>
<td>-</td>
</tr>
<tr>
<td>total muon</td>
<td>0.915 %</td>
<td>0.916 %</td>
<td>0.916 %</td>
<td>0.916 %</td>
<td>0.913 %</td>
<td>0.878 %</td>
<td>0.483 %</td>
</tr>
<tr>
<td>$W_{\mu u\text{Boost}}$</td>
<td>0.203%</td>
<td>0.304%</td>
<td>0.203%</td>
<td>0.203%</td>
<td>0.306%</td>
<td>0.831%</td>
<td>0.204%</td>
</tr>
<tr>
<td>$N_{\mu u\text{Boost}}$</td>
<td>0.758%</td>
<td>0.806%</td>
<td>0.731%</td>
<td>0.724%</td>
<td>0.791%</td>
<td>0.770%</td>
<td>-</td>
</tr>
<tr>
<td>momentum calibration</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.65 %</td>
</tr>
<tr>
<td>jet energy scale</td>
<td>6.05%</td>
<td>9.78%</td>
<td>8.64%</td>
<td>13.02%</td>
<td>11.33%</td>
<td>9.41%</td>
<td>-</td>
</tr>
<tr>
<td>jet energy resolution</td>
<td>2.87%</td>
<td>0.87%</td>
<td>1.83%</td>
<td>7.37%</td>
<td>3.87%</td>
<td>4.71%</td>
<td>-</td>
</tr>
<tr>
<td>jet energy identification</td>
<td>1.7%</td>
<td>1.7%</td>
<td>1.7%</td>
<td>1.7%</td>
<td>1.7%</td>
<td>1.7%</td>
<td>-</td>
</tr>
<tr>
<td>total</td>
<td>7.131%</td>
<td>10.117%</td>
<td>9.163%</td>
<td>15.168%</td>
<td>12.230%</td>
<td>10.865%</td>
<td>3.19 %</td>
</tr>
</tbody>
</table>

Table 7.1: Relative systematic uncertainty.
Chapter 8

Results and discussion

8.1 Rolke limit estimation

To set a limit on the muon-heavy neutrino mixing parameter the next procedure is adopted. First, the limit on the number of the signal events is estimated using the TRo1ke procedure, as defined in ROOT implementation of [67], called TRo1ke [68]. The TRo1ke procedure allows to set limits in a presence of nuisance parameters.

The TRo1ke method is based on the profile likelihood ratio test statistics used for hypothesis testing in the frequentist approach. The signal model is always assumed to be Poisson. Background models can be Poisson or Gaussian. Efficiency can be Binomial or Gaussian. The Gaussian models were chosen for both background and efficiency. Only a statistical uncertainty on the background is used. To correctly propagate the efficiency uncertainty to the limit, the efficiency ratio total uncertainty is used in the limit estimation, see (1.44). Since the efficiency ratio is just a scaling factor in (1.44), it is set to 1 in the TRo1ke.

The likelihood ratio, $\lambda$, is:

$$\lambda = \frac{L(\mu, \hat{b}, \hat{e}|x, b, \sigma_b, \sigma_e)}{L(\hat{\mu}, \hat{b}, \hat{e}|x, b, \sigma_b, \sigma_e)}, \quad (8.1)$$

where $\mu$ is a number of signal events and a parameter of interest, $b$ is a background estimate, $e$ is an efficiency, $\sigma_b$ is a background estimate standard deviation, $\sigma_e$ is a signal estimate standard deviation, $x$ is a total number of measured events, $\hat{\cdot}$ denotes maximum likelihood estimator as a function of $\mu$ and $L$ is a likelihood function:

$$L(\mu, \hat{b}, \hat{e}|x, b, \sigma_b, \sigma_e) = p(x, b, e|\mu, \hat{b}, \hat{e}, \sigma_b, \sigma_e), \quad (8.2)$$

where $p$ is a probability density function.

Assuming the Gaussian background and efficiency models, the probability density function is:

$$p(x, b, e|\mu, \hat{b}, \hat{e}, \sigma_b, \sigma_e) = \frac{(\hat{e} \mu + \hat{b})^x}{x!} e^{-(\hat{e} \mu + \hat{b})} \cdot \frac{1}{\sigma_b \sqrt{2\pi}} e^{-\frac{(\mu - \mu_0)^2}{2\sigma_b^2}} \cdot \frac{1}{\sigma_e \sqrt{2\pi}} e^{-\frac{(e - e_0)^2}{2\sigma_e^2}}. \quad (8.3)$$

According to the Wilk’s theorem [69], $-2\log \lambda$ test statistics follows a $\chi^2$-distribution with a number of degrees of freedom $k$, that equals to the difference in the number of free parameters in two tested hypotheses. In the above case, $k = 1$, since fixed or free $\mu$ is the only difference between the two tested hypotheses.

The limit at the confidence level CL is defined as $(1-CL)\%$ percentile difference of $\chi^2(k = 1)$ in $-2\log \lambda$ distribution from the $-2\log \lambda$ minimum. For 95% CL the $\chi^2(k = 1)$ percentile is 3.84. So, the limit on the number of signal events equals to
Chapter 8. Results and discussion

the number of signal events, for which \(-2\log \lambda\) distribution changes on 3.84, compared to the minimum. For the upper limit the percentile difference is searched to the right from the \(-2\log \lambda\) minimum. For the lower limit the same is done, but to the left of the minimum. The cases of negative values of \(\mu\), that can arise when the lower limit is estimated, are banned by TrRolke.

Background is extrapolated from the control regions to the aimed signal region with a transfer factor. For the expected upper limit, the hypothesis is that no signal is found. Therefore, the number of measured events is equal to the estimated number of background.

A pre-study for different efficiency errors and different number of measured signal events was done. A pseudo signal selection was used as an aimed one, where all BDT classifiers cuts were set to be 0.55 and all other cuts remained untouched in comparison with a signal selection (see Tab. 5.2 and Tab. 5.3). The limit was estimated for the positive and negative charge signal. Positive/negative charge background is defined according to the \(\mu\) charge. For \(W \rightarrow \mu \nu\) more positive \(\mu_W\) are expected and, therefore, positive \(W^1\). For \(Z \rightarrow \mu^+\mu^-\) there is an equal amount of positive and negative muons, therefore, an equal amount of the background.

The expected upper limit for the number of signal events as a function of efficiency standard deviation in the range from \(10^{-4}\) to 0.5 is shown in Fig. 8.1a. The number of estimated positively charged background is higher than of negatively charged background and the upper limit on the number of signal events behaves accordingly. At \(\sigma_\varepsilon \leq 0.3\) upper limit is quite stable. At \(\sigma_\varepsilon \geq 0.3\) the upper limit blows up due to the high uncertainty of the \(\varepsilon\).

As a number of “measured” signal events an upper limit is shown in Fig. 8.1b. The number of “measured” signal events is set to the chosen trial values in the range from 0 to 50 to study the upper limit behavior. The standard deviation is set to be 0.1. As expected, the upper limit on the number of signal events depends linearly on the number of “measured” signal events. The respective behavior of the different charge samples is the same as was seen in Fig. 8.1a.

8.2 Normalization channel yields

The normalization channel yield is taken from a fit of \(\mu_W\) transverse momentum spectrum separately for \(W^+ \rightarrow \mu^+\nu\) and \(W^- \rightarrow \mu^-\nu\), as shown in Fig. 8.2 and Fig. 8.3 for \(2 < \eta < 2.25\). Other \(\eta\) bins fits can be found in App. E. Results for positively and negatively charged background are then added together and used in the upper limit calculation.

Fit is done using the RooFit::HistFactory tool that allows to use shape templates as fitting models [70].

All background, except for QCD background, use the corresponding Monte Carlo to get the shape after the normalization channel selection is applied. QCD shape is taken from normalization channel data sample by applying selection described in Tab. 8.1.

Due to the ambiguity of choosing the correct QCD shape, as an alternative, the exponential QCD background shape is planned to be tested. The QCD shape ambiguity is another source of a systematic uncertainty, which was not considered yet.

Some discrepancies between fit and data arise in the 30-45 GeV region. The source is the incorrect \(p_T\) shape in the \(W \rightarrow \mu\nu\) Monte Carlo. A correction procedure is going to be applied to fix this.

\(^{1}\)Since the system in LHC is biased to the positive charge (proton is positive), a positive charged particles production is enhanced.
8.2. Normalization channel yields

\[8.2.\text{ Normalization channel yields}\]

\[m = 15 \text{ [GeV]}, x-b = 0\]

\[\begin{align*}
\text{charge: +} \\
\text{charge: -}
\end{align*}\]

\[m = 15 \text{ [GeV]}, \sigma \epsilon = 0.1\]

\[\begin{align*}
\text{charge: +} \\
\text{charge: -}
\end{align*}\]

\(\sigma_{\epsilon} = 0.1\)

\(F_{IGURE\ 8.1}:\) Upper limit on the number of signal events evaluated using TRolke method.

<table>
<thead>
<tr>
<th>QCD selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0Muon</td>
</tr>
<tr>
<td>Hlt1SingleMuonHighPT</td>
</tr>
<tr>
<td>Hlt2SingleMuonHighPT</td>
</tr>
<tr>
<td>2 &lt; \eta(\mu) &lt; 4.5</td>
</tr>
<tr>
<td>(</td>
</tr>
<tr>
<td>(P(\chi^2) &gt; 0.01)</td>
</tr>
<tr>
<td>NumTTHits &gt; 0</td>
</tr>
<tr>
<td>20 &lt; (p_T(\mu) &lt; 70) GeV</td>
</tr>
<tr>
<td>IP(\mu) &lt; 0.04 mm</td>
</tr>
<tr>
<td>(E_{\text{ECAL}} + E_{\text{HCAL}})/p &lt; 4%</td>
</tr>
<tr>
<td>Wmu_uBoost &lt; 0.55</td>
</tr>
</tbody>
</table>

\(T\ABLE\ 8.1: QCD\ selection\ for\ getting\ the\ QCD\ shape\ from\ data.\)
Chapter 8. Results and discussion

**Figure 8.2:** Fit of $PT(\mu^+_W)$ in $2 < \eta < 2.25$ bin.

**Figure 8.3:** $PT(\mu^-_W)$ in $2 < \eta < 2.25$ fit.
8.3. Expected upper limit

The expected upper limit on the product of a BR($N \rightarrow \mu jet$) with a muon-heavy neutrino mixing parameter calculated using (1.44) and TRo1ke procedure at CL=95% is shown in Tab. 8.3 and in Fig. 8.4 as a function of a heavy neutrino mass. As an aimed selection a full signal selection is used (see. Tab. 5.2 and Tab. 5.3).

\[ \text{Table 8.2: Measured W yields in bins of } \eta \text{ and total W yields.} \]

<table>
<thead>
<tr>
<th>$\Delta \eta$</th>
<th>$W^+ \rightarrow \mu^+ \nu$</th>
<th>$W^- \rightarrow \mu^- \nu$</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00-2.25</td>
<td>175416 ± 623</td>
<td>98847 ± 481</td>
<td></td>
</tr>
<tr>
<td>2.25-2.50</td>
<td>202175 ± 662</td>
<td>111754 ± 520</td>
<td></td>
</tr>
<tr>
<td>2.50-2.75</td>
<td>179653 ± 600</td>
<td>103883 ± 498</td>
<td></td>
</tr>
<tr>
<td>2.75-3.00</td>
<td>153592 ± 545</td>
<td>97633 ± 484</td>
<td></td>
</tr>
<tr>
<td>3.00-3.25</td>
<td>128812 ± 499</td>
<td>94261 ± 481</td>
<td></td>
</tr>
<tr>
<td>3.25-3.50</td>
<td>95073 ± 430</td>
<td>81505 ± 453</td>
<td></td>
</tr>
<tr>
<td>3.50-4.00</td>
<td>107407 ± 501</td>
<td>123286 ± 587</td>
<td></td>
</tr>
<tr>
<td>4.00-4.50</td>
<td>21201 ± 244</td>
<td>27827.2 ± 282</td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>1063329 ± 16375</td>
<td>738997 ± 12255</td>
<td>1802325 ± 40788</td>
</tr>
</tbody>
</table>

\[ \text{Figure 8.4: Expected upper limit on the muon-heavy neutrino mixing as a function of the heavy neutrino mass.} \]
with a mixing parameter between a muon neutrino and a heavy neutrino at CL = 95%.

<table>
<thead>
<tr>
<th>Signal</th>
<th>5 GeV</th>
<th>90.36</th>
<th>15.00%</th>
<th>1.425 ± 0.674</th>
<th>1.425 ± 0.674</th>
<th>3.914</th>
<th>1802325 ± 40788</th>
<th>2.116 · 10^−4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>10 GeV</td>
<td>57.23</td>
<td>14.86%</td>
<td>1.425 ± 0.674</td>
<td>1.425 ± 0.674</td>
<td>3.914</td>
<td>1802325 ± 40788</td>
<td>1.340 · 10^−4</td>
</tr>
<tr>
<td>Signal</td>
<td>15 GeV</td>
<td>51.40</td>
<td>15.29%</td>
<td>1.425 ± 0.674</td>
<td>1.425 ± 0.674</td>
<td>3.914</td>
<td>1802325 ± 40788</td>
<td>1.204 · 10^−4</td>
</tr>
<tr>
<td>Signal</td>
<td>20 GeV</td>
<td>58.79</td>
<td>16.83%</td>
<td>1.425 ± 0.674</td>
<td>1.425 ± 0.674</td>
<td>3.916</td>
<td>1802325 ± 40788</td>
<td>1.378 · 10^−4</td>
</tr>
<tr>
<td>Signal</td>
<td>30 GeV</td>
<td>58.79</td>
<td>17.33%</td>
<td>1.425 ± 0.674</td>
<td>1.425 ± 0.674</td>
<td>3.917</td>
<td>1802325 ± 40788</td>
<td>1.378 · 10^−4</td>
</tr>
<tr>
<td>Signal</td>
<td>50 GeV</td>
<td>140.71</td>
<td>17.43%</td>
<td>1.425 ± 0.674</td>
<td>1.425 ± 0.674</td>
<td>3.917</td>
<td>1802325 ± 40788</td>
<td>3.298 · 10^−4</td>
</tr>
</tbody>
</table>

Table 8.3: Normalization to signal efficiency ratio, its uncertainty, the number of expected background events, the total number of events, the upper limit on the expected signal, the W yield and the expected upper limit on the product of branching ratio for $N \rightarrow \mu \text{jet}$ with a mixing parameter between a muon neutrino and a heavy neutrino at CL = 95%.
Chapter 9

Conclusion and outlook

Results of the search for the non-prompt heavy neutrino in $W^\pm \rightarrow \mu^\pm \mu^\pm jet$ were presented. The expected upper limit on the muon-heavy neutrino mixing for 5-50 GeV heavy neutrino with 10 ps lifetime is found to be:

| Signal [GeV] | $\text{BR}(N_R \rightarrow \mu jet) | B_{\mu N} |^2$ |
|--------------|-----------------------------------|
| 5            | $2.116 \cdot 10^{-4}$             |
| 10           | $1.340 \cdot 10^{-4}$             |
| 15           | $1.204 \cdot 10^{-4}$             |
| 20           | $1.378 \cdot 10^{-4}$             |
| 30           | $1.378 \cdot 10^{-4}$             |
| 50           | $3.298 \cdot 10^{-4}$             |

Discrepancies in the isolation variables distribution in the control regions between background Monte Carlo and data might indicate an unconsidered background. A further investigation is needed.

Originally, this analysis was planned to be a complimentary measurement for the [17]. Selection was left to be as similar as possible to the [17]. However, the low efficiency of the $\mu N$ IP cut is found to be unsatisfactory, so the $\mu N$ IP cut will be separately optimized to improve the expected limit.

Known issues with $W \rightarrow \mu \nu$ Monte Carlo shapes will be solved to improve the quality of the normalization channel fit.

After answering the above questions the data will be unblinded.

These results are going to be combined with prompt results from [17] to retrieve lifetime dependent interpolated limit. The limit will also be interpolated between the mass points.

Finally, the entire Run I and Run II data should be analyzed.
Appendix A

\[ \mu_N \text{ multiplicity} \]
Appendix A. $\mu_N$ multiplicity

\begin{figure}[h]
\centering
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{example_a}
\caption{$m = 5$ [GeV]}
\end{subfigure} \hfill
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{example_b}
\caption{$m = 10$ [GeV]}
\end{subfigure}
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{example_c}
\caption{$m = 15$ [GeV]}
\end{subfigure} \hfill
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{example_d}
\caption{$m = 20$ [GeV]}
\end{subfigure}
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{example_e}
\caption{$m = 30$ [GeV]}
\end{subfigure} \hfill
\begin{subfigure}{0.49\textwidth}
\centering
\includegraphics[width=\textwidth]{example_f}
\caption{$m = 50$ [GeV]}
\end{subfigure}
\caption{Distribution of $\mu_N$ candidate multiplicity per $\mu_W$ for different mass Monte Carlo samples.}
\end{figure}
Appendix B

\( \mu_W \) and \( \mu_N \) kinematics
Appendix B. $\mu_W$ and $\mu_N$ kinematics

![Histograms showing kinematics of $\mu_W$ in bins of $p_T$ and $\eta$ for different mass Monte Carlo samples.](image)

Figure B.1: Kinematics of $\mu_W$ in bins of $p_T$ and $\eta$ for different mass Monte Carlo samples.

(A) $m = 5$ [GeV]

(B) $m = 10$ [GeV]

(C) $m = 15$ [GeV]

(D) $m = 20$ [GeV]

(E) $m = 30$ [GeV]

(F) $m = 50$ [GeV]
Appendix B. $\mu_W$ and $\mu_N$ kinematics

Figure B.2: Kinematics of $\mu_N$ in bins of $p_T$ and $\eta$ for different mass Monte Carlo samples.
Appendix C

Calibration of momentum

The run I data has an imperfectness of the momentum reconstruction that leads to shift of momenta [65]. The effect can be easily seen in the $Z \rightarrow \mu^+\mu^-$ data when plotting mean dimuon mass with respect to phi, see Fig. C.1. In Fig. C.1 one can see sin-like behavior. Positive and negative tracks experience influence of the momentum shift differently at the same $\phi$. If positive tracks momentum shifts to the higher values, negative tracks momentum shifts to the lower values in the same $\phi$ bin.

This effect can not be explained by magnetic field effects or scale precision of a detector in z-axis. Sin-shape of a bias advocates that the bias is a result of a misalignment proportional to tangent $t_y = \frac{dy}{dx}$. Such misalignment can be a rotation of subdetectors after magnet with respect to the ones before the magnet (i.e. closer to the VELO) along z-axis.

The resulting momentum bias effect broads the width of a Z-mass peak in $Z \rightarrow \mu^+\mu^-$ data and decreases the momentum/mass resolution.

It is possible to calculate the bias that lead to the above effect and correct for it. Below a procedure for the curvature bias momentum calibration developed in [65] is described.
Appendix C. Calibration of momentum

C.1 Curvature bias momentum calibration

A curvature of a track is defined as following:

\[ \omega = \frac{q}{|\vec{p}_{xz}|}, \]  

where \( \omega \) is a curvature, \( q \) is a track charge and \( |\vec{p}_{xz}| \) is an absolute momentum in xz-plane.

The bias is defined separately for positive and negative tracks. A dimuon invariant mass is defined as:

\[ p^2(\mu\mu) = M^2(\mu\mu)^1 = (p(\mu_1) + p(\mu_2))^2 \]  

\[ M(\mu\mu) = \sqrt{2|\vec{p}(\mu_1)||\vec{p}(\mu_2)|(1 - \cos \theta)}, \]  

where muon masses were neglected and \( |\vec{p}(\mu)| \) is a muon momentum modulus, \( \theta \) is an angle between two muon 3-momenta.

From (C.3) the dimuon invariant mass variation is defined as:

\[ \frac{dM(\mu\mu)}{d\omega(\mu)} = \frac{dM(\mu\mu)}{d|\vec{p}(\mu)|} \frac{d|\vec{p}(\mu)|}{d\omega}, \]  

\[ \frac{d|\vec{p}|}{d\omega_1} = -\frac{M(\mu\mu)}{2\omega_1}, \]  

\[ \delta M(\mu\mu) \equiv M - M_Z = -\frac{M(\mu\mu)}{2\omega_i} \delta \omega_i, \]  

where \( \delta \omega_i \) is a curvature bias in ith bin.

If \( M(\mu\mu) \) is close to the Z peak we can assume that \( M(\mu\mu) = M_Z \) then

\[ \langle \delta M(\mu\mu) \rangle_i = \langle M \rangle_i - M_Z = -\frac{M_Z}{2\langle \omega_i \rangle} \delta \omega_i. \]  

Then the curvature bias can be defined from (C.7) as

\[ \delta \omega_i = -2q \left( \frac{\langle M \rangle_i}{M_Z} - 1 \right) \frac{1}{\langle |\vec{p}| \rangle}. \]  

Using a \( \chi^2 \) method for charge split tracks:

\[ \chi^2 = \sum_{\text{events} k} \frac{(M_k - M_Z)^2}{\sigma_k^2}, \]  

where \( M_k \) is a dimuon mass of kth event and \( \sigma_k^2 = \Gamma_k^2 + \sigma_k^2(M_k) \). \( \Gamma_k \) is a Z decay width and \( \sigma(M_k) \) is a statistical uncertainty of the reconstructed dimuon mass.

To minimize \( \chi^2 \) curvature \( \omega_i \) has to be shifted on \( \delta \omega_i \)

\[ \Delta \omega_i = -\left( \sum_{\text{events} k} \frac{d^2 \chi^2_k}{d\omega_i \omega_j} \right)^{-1} \left( \sum_{\text{events} k} \frac{d\chi^2_k}{d\omega_i} \right), \]  

\(^1\)By definition \( p^2 \equiv M^2 \)
C.1. Curvature bias momentum calibration

where

\[ \frac{1}{2} \frac{d \chi_k^2}{d \omega_i} = - \frac{(M_k - M_Z)M_Z}{\sigma_k^2} \left( \delta_{i+} \frac{1}{q^+} |\vec{p}^+| + \delta_{i-} \frac{1}{q^-} |\vec{p}^-| \right), \] (C.11)

\[ \frac{1}{2} \frac{d^2 \chi_k^2}{d \omega_i d \omega_j} = \frac{M_Z^2}{4 \sigma_k^2} \left( \delta_{i+} \delta_{j+} |\vec{p}^+|^2 + \delta_{i-} \delta_{j-} |\vec{p}^-|^2 - |\vec{p}^+| |\vec{p}^-| (\delta_{i+} \delta_{j-} + \delta_{i-} \delta_{j+}) \right). \] (C.12)

In the above \( \delta_{i+}(-) \) takes into the account charge of a track and \( q^+ = 1, q^- = -1 \). The latter was used in the (C.12). It is important to keep the derivation explicit to include correlation of curvature between positive and negative tracks. From Fig. C.2 one can see that the curvature bias in bins of \( \phi \) is only slightly different for opposite charges.

Curvature bias defined from the above is then used to correct \( |\vec{p}_{xz}| \):

\[ \frac{q}{|\vec{p}_{xz}|_{\text{new}}} = \frac{q}{|\vec{p}_{xz}|_{\text{old}}} + \delta \omega, \] (C.13)

where \( |\vec{p}_{xz}| = |\vec{p}| \cdot \left( \frac{1+t_x^1}{t_x^1+t_y^1} \right)^2 \).

Results of the correction in bins of \( \phi \) and \( \eta \) are shown in Fig. C.3 and in Fig. C.4. In terms of \( \phi \) bins a huge improvement is observed. The \( \eta \) distributions also flattens. The improvements on the mass resolution of the Z peak are shown in Fig. C.5.

\[ t_x = \frac{p_x}{p_z}, \quad t_y = \frac{p_y}{p_z} \]
Appendix C. Calibration of momentum

\begin{figure}[h]
\centering
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{figureA}
\caption{$\mu^+$, Magnet Down}
\end{subfigure}
\hspace{0.05\textwidth}
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{figureB}
\caption{$\mu^-$, Magnet Down}
\end{subfigure}
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{figureC}
\caption{$\mu^+$, Magnet Up}
\end{subfigure}
\hspace{0.05\textwidth}
\begin{subfigure}{0.45\textwidth}
\includegraphics[width=\textwidth]{figureD}
\caption{$\mu^-$, Magnet Up}
\end{subfigure}
\caption{Mean dimuon mass in bins of $\phi$ before and after applying curvature bias momentum correction. $Z \rightarrow \mu^+\mu^-$ data.}
\end{figure}
C.1. Curvature bias momentum calibration

Figure C.4: Mean dimuon mass in bins of $\eta$ before and after applying curvature bias momentum correction. $Z \rightarrow \mu^+\mu^-$ data.
Appendix C. Calibration of momentum

Figure C.5: Mass resolution improvement in the $Z \rightarrow \mu^+\mu^-$ mass peak after the momentum calibration is applied.
Appendix D

Efficiency uncertainty

A true efficiency is an unknown parameter of our system. Both frequentist and
Bayesian statistics can be used for the estimation of the true efficiency interval, which
we will assign as a ranges of the measured efficiency uncertainties. For an ongoing
discussion refer to [71].

The frequentist and Bayesian approaches define the interval in a different way. In
the frequentist approach one uses a confidence interval and finds an interval which
includes a true efficiency \( \alpha \) times, where \( \alpha \) denotes the confidence level. The probability for a particle to pass a trigger cut (success) follows a Binomial distribution:

\[
P(k; \varepsilon, n) = \frac{n!}{k!(n-k)!} \varepsilon^k (1-\varepsilon)^{(n-k)}, \quad (D.1)
\]

where \( k \) is a number of candidates that successfully passed a cut, \( n \) is a total
number of candidates and \( \varepsilon \) is a true efficiency.

However, while constructing a confidence interval for a binomial distribution
one runs into the problem that an integral of binomial with respect to \( k \) is not analyt-
ical. For this case a normal approximation binomial confidence interval is computed:

\[
\varepsilon \pm z \frac{k(n-k)}{n}, \quad (D.2)
\]

where \( z \) is a \( 1 - \frac{1-\alpha}{2} \) quantile of a normal distribution \( N(0,1) \) for a confidence level \( \alpha \).

This definition runs into a problem when \( k = 0 \) or \( (n-k) = 0 \). In both cases one
assigns a 0 value for the uncertainty.

Another disadvantage of binomial confidence interval is that it is a symmetric
interval, which means, that in case of \( \varepsilon = 100\% \ (\varepsilon = 0\%) \), upper (lower) uncertainty
is wrong, since efficiency can not be greater (less) than 100\% (0\%).

An alternative approach is Bayesian, where the trick is to choose a reasonable
prior for the efficiency. Since we do not know anything about the true efficiency,
except that it is ranged from 0 to 1, the best prior we can choose is a uniform flat
prior in [0, 1] range. Then according to the Bayes theorem one can write:

\[
P(\varepsilon; k, n) = \frac{P(k; \varepsilon, n)P(\varepsilon; n)}{C}, \quad (D.3)
\]

where \( P(\varepsilon; n) \) is a prior for efficiency, chosen to be a uniform distribution in [0,1],
\( P(k; \varepsilon, n) \) is a likelihood of data, \( C = 1/(n+1) \) is a normalization constant [71] and
\( P(\varepsilon; k, n) \) is a posterior distribution.

In our case likelihood is defined as a Binomial distribution:

\[
P(k; \varepsilon, n) = \frac{n!}{k!(n-k)!} \varepsilon^k (1-\varepsilon)^{(n-k)}. \quad (D.4)
\]
Putting all together one gets:

\[ P(\varepsilon; k, n) = \frac{(n + 1)!}{k!(n-k)!} \varepsilon^k (1 - \varepsilon)^{(n-k)} \]  
(D.5)

in [0, 1] range.

Now one can formally derive a credential interval:

\[ \int_a^0 d\varepsilon P(\varepsilon; k, n) = 0.1585, \]  
(D.6a)

\[ \int_0^1 d\varepsilon P(\varepsilon; k, n) = 0.8415, \]  
(D.6b)

\[ \int d\varepsilon P(\varepsilon; k, n) = \beta_r(a, k + 1, n - k + 1), \]  
(D.6c)

where first integral correspond to the lower one-tail credible interval and 0.1585 corresponds to the 1\(\sigma\)/2 from the left in \(N(0, 1)\); second integral correspond to the higher one-tail credible interval and 0.8415 corresponds to the 1\(\sigma\)/2 from the right in \(N(0, 1)\); \(\beta_r\) is a regularized beta-function as defined in ROOT:

\[ \beta_r(x, b, a) = \int_0^x u^{a-1} (1 - u)^{b-1} du \]  
\(\beta(a, b)\) is a beta function.

Using the above one can find a credible interval for an asymmetric 1\(\sigma\) interval.

This procedure, however, is time consuming.

Luckily, in case of a flat prior the Bayesian results should converge with frequentest. The convergence of both methods is checked in Fig. D.1 by evaluating errors of the efficiency of one of the L0-level trigger cuts:

\[ \varepsilon(n_{SPDhits} < 600) = \frac{N(n_{SPDhits} < 600)}{N} \]  
(D.8)

where \(n_{SPDhits}\) is a number of SPD hits; \(\varepsilon(n_{SPDhits} < 600)\) is an efficiency of the L0-level cut on the number of SPD hits; \(N(n_{SPDhits} < 600)\) is a number of events that passed the cut; \(N\) is a total number of events.

In Fig. D.1a the resulting distribution of efficiency is shown. Uncertainty was computed twice: as a binomial confidence interval and as a Bayesian credential interval. There is almost no difference in the efficiency uncertainties between confidence and credential interval in the absolute values.

To closer compare two approaches a relative difference between intervals is calculated as following:

\[ \frac{(\Delta \varepsilon(Binomial) - \Delta \varepsilon(Bayesian))}{\Delta \varepsilon(Binomial)}, \]  
(D.9)

where \(\Delta \varepsilon(Binomial)\) is a binomial confidence interval and \(\Delta \varepsilon(Bayesian)\) is a Bayesian credential interval. Since Bayesian interval is asymmetric, the above difference is studied for lower and upper uncertainty separately. Results can be found in Fig. D.1b. Binomial confidence interval overestimates the Bayesian credential interval in most cases by 1-4\%. For the 1 primary vertex events Bayesian interval overestimates the lower uncertainty compare to Binomial interval for 40\%, but binomial confidence interval overestimates on 30\% the upper uncertainty. The Binomial distribution is
Appendix D. Efficiency uncertainty

(A) Efficiency with different errors as a function of number of primary vertices (nPVs).

(B) Difference between confidence and credential intervals with respect to the confidence interval.

FIGURE D.1: Comparison of the efficiency errors evaluated for efficiency of nSPDhits < 600 cut in $Z \rightarrow \mu^+ \mu^-$ data.

more skewed for efficiency close to 0 or 1. For the probabilities close to 1 the Binomial uncertainty distribution has a longer tail to the left and rapid to the right. Binomial uncertainty is symmetric and is not sensitive to the non-zero skewness. On the other hand, Bayesian uncertainty allows asymmetry and is based on the fact that the probabilistic coverage is similar in both tails. Therefore, a discrepancy appears. Since computing of the binomial confidential interval is less time consuming and more conservative, it was used to compute the uncertainty of efficiency.
Appendix E

Normalization channel fits
Appendix E. Normalization channel fits

Figure E.1: $PT(\mu_W)$ fit.
Appendix E. Normalization channel fits

Figure E.2: $PT(\mu W)$ fit.
Appendix E. Normalization channel fits

Figure E.3: $PT(\mu_W)$ fit.
Bibliography

7. Sirunyan, A. M. et al. Search for heavy Majorana neutrinos in $e^\pm e^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. JHEP 120, 221801. 20 p (2018).
8. Khachatryan, V. et al. Search for heavy Majorana neutrinos in $e^\pm e^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. Phys. Lett. B 748, 144. 23 p (2015).
10. Chatrchyan, S. et al. Search for heavy Majorana neutrinos in $e^\pm e^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. JHEP 04, 169. 42 p (2016).
11. Khachatryan, V. et al. Search for heavy Majorana neutrinos in $e^\pm e^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. JHEP 10, 107. 21 p (2011).
14. Chatrchyan, S. et al. Search for heavy Majorana neutrinos in $\mu^\pm \mu^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. Phys. Lett. B 748, 144. 23 p (2015).
15. Khachatryan, V. et al. Search for heavy Majorana neutrinos in $e^\pm e^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. JHEP 10, 107. 21 p (2011).
18. Chatrchyan, S. et al. Search for heavy Majorana neutrinos in $\mu^\pm \mu^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. Phys. Lett. B 748, 144. 23 p (2015).
19. Sirunyan, A. M. et al. Search for heavy Majorana neutrinos in $e^\pm e^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. JHEP 10, 107. 21 p (2011).
20. Khachatryan, V. et al. Search for heavy Majorana neutrinos in $\mu^\pm \mu^\pm$ + jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV. Phys. Lett. B 748, 144. 23 p (2015).


65. Hulsbergen, W. *Momentum calibration for \(Z \rightarrow \mu^+\mu^-\) events 2017*.

66. Aaij, R. et al. Measurement of forward \(W\) and \(Z\) boson production in association with jets in proton-proton collisions at \(\sqrt{s} = 8\) TeV. *JHEP* 05, 131. 27 p (2016).


