# Search for heavy neutrinos and characterisation of silicon sensors for the VELO upgrade

CERN-THESIS-2020-410
17/06/2020

ELENA DALL'OCCO

Cover designed by Caterina Ferioli
Printed by Gildeprint – The Netherlands
ISBN 978-94-6402-280-3
Copyright © 2020 Elena Dall'Occo, all rights reserved.

This work is part of the research program of the Netherlands Organization for Scientific Research (NWO). It was carried out at the National Institute for Subatomic Physics (Nikhef) in Amsterdam, the Netherlands.

VRIJE UNIVERSITEIT

# Search for heavy neutrinos and characterisation of silicon sensors for the VELO upgrade

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Bètawetenschappen
op woensdag 17 juni 2020 om 13.45 uur
in een online bijeenkomst van de universiteit,
De Boelelaan 1105

door

Elena Dall'Occo

geboren te Ferrara, Italië

| | |
|---|---|
| promotor: | prof.dr. M.H.M. Merk |
| copromotoren: | dr. W.D. Hulsbergen |
| | dr. M.G. van Beuzekom |

# Contents

**Introduction** **1**

**I Heavy Neutrinos** **3**

**1 Theoretical Motivation** **5**
1.1 Neutrinos in the Standard Model . . . . . 5
1.2 Limitations of the Standard Model . . . . . 7
1.2.1 Neutrino oscillation . . . . . 8
1.3 Nature of the neutrino . . . . . 10
1.4 See-saw mechanism . . . . . 12
1.5 Implications of neutrino mass . . . . . 14
1.5.1 Models . . . . . 14
1.5.2 $\nu$MSM . . . . . 16
1.6 Current experimental status . . . . . 17
1.6.1 Cosmology and oscillation experiments . . . . . 17
1.6.2 Indirect measurements . . . . . 18
1.6.3 Direct measurements . . . . . 18
1.6.4 Future experiments . . . . . 20
1.7 Targeted mass range . . . . . 22
1.7.1 Production Mode . . . . . 22
1.7.2 Decay Mode . . . . . 23

**2 The LHCb detector** **27**
2.1 Tracking detectors . . . . . 29
2.1.1 VErtex LOcator . . . . . 29
2.1.2 Silicon trackers . . . . . 30
2.1.3 Outer tracker . . . . . 31
2.1.4 Magnet . . . . . 31
2.2 Particle identification system . . . . . 32
2.2.1 RICH . . . . . 32
2.2.2 Calorimeters . . . . . 34

2.2.3 Muon stations . . . 35
2.3 Trigger . . . 36
2.4 Offline data processing . . . 37
2.5 Event reconstruction . . . 37
2.5.1 Track reconstruction . . . 38
2.5.2 Neutral particle reconstruction . . . 39
2.5.3 Particle identification . . . 39
2.5.4 Jet reconstruction . . . 40

**3 Analysis Strategy and Data Selection 47**
3.1 Analysis strategy . . . 47
3.2 Data Samples . . . 48
3.2.1 Data . . . 49
3.2.2 Simulated Signal . . . 49
3.2.3 Simulated Backgrounds . . . 50
3.3 Signal selection . . . 51
3.3.1 Preselection . . . 51
3.3.2 Multivariate Selection . . . 54
3.3.3 Full selection . . . 64
3.3.4 Signal efficiency . . . 65
3.4 Backgrounds selection . . . 68
3.4.1 Yields prediction . . . 70
3.4.2 Background Composition per Region . . . 71
3.5 Selection of other decay modes . . . 71

**4 Results 75**
4.1 Binned maximum likelihood fit . . . 76
4.2 Determination of $W \to \mu\nu$ yield for normalisation . . . 77
4.3 Determination of the neutrino mass fit model . . . 80
4.3.1 Determination of background yields in control regions . . . 80
4.3.2 Determination of light QCD yield in the signal region . . . 82
4.3.3 Neutrino mass fit model . . . 84
4.3.4 Validation of the neutrino mass fit model . . . 85
4.4 Efficiency corrections . . . 88
4.4.1 Muon reconstruction . . . 88
4.4.2 Global event cut . . . 90
4.4.3 $\mu_W$ uBDT cut . . . 90
4.4.4 $\mu_N$ uBDT cut . . . 91
4.5 Systematic uncertainties . . . 92
4.5.1 Efficiency ratio uncertainty . . . 92
4.5.2 Background yields uncertainties . . . 97

4.5.3 Normalisation yield uncertainty . . . 98
4.6 Upper limit determination . . . 99
4.6.1 Statistical hypothesis test . . . 99
4.6.2 Unblinded result . . . 100
4.7 Outlook . . . 102

**II Silicon Sensors for the VELO Upgrade 105**

**5 The VELO Upgrade 107**
5.1 LHCb Upgrade . . . 107
5.2 VELO Constraints . . . 110
5.3 VELO Layout . . . 114
5.3.1 Sensors . . . 116
5.3.2 VeloPix . . . 116
5.3.3 Microchannel Cooling . . . 117
5.3.4 RF Box . . . 118
5.4 Simulated Performance . . . 118

**6 Silicon Sensors 121**
6.1 Sensor concepts . . . 121
6.1.1 Interaction of particles with matter . . . 121
6.1.2 The p-n junction . . . 124
6.1.3 Signal formation . . . 126
6.2 Radiation damage . . . 127
6.2.1 Microscopic level . . . 128
6.2.2 Macroscopic level . . . 129
6.2.3 Annealing . . . 132
6.2.4 Impact on detector operations . . . 133
6.3 Pixel sensors . . . 133
6.3.1 Sensor types . . . 134
6.3.2 Guard rings . . . 134

**7 Testbeam 137**
7.1 Sensors Requirements . . . 137
7.1.1 Behaviour at the Edge . . . 139
7.2 Assemblies . . . 139
7.2.1 Sensor description . . . 139
7.2.2 Timepix3 ASIC . . . 140
7.3 Timepix3 Telescope . . . 145
7.3.1 Hardware . . . 146

7.3.2 Offline reconstruction . . . 147
7.3.3 Performance . . . 149
7.4 Prototype sensor edge analysis . . . 149
7.4.1 Charge Collection Region . . . 155
7.4.2 Excess hits investigation . . . 158
7.4.3 Effect of irradiation . . . 161
7.4.4 Summary . . . 162

**8 Grazing Angles Analysis 165**
8.1 Data Selection . . . 166
8.2 Fit and Systematics . . . 167
8.3 Timewalk Correction . . . 171
8.4 Nonirradiated Results . . . 173
8.4.1 Charge Collection . . . 173
8.4.2 Time-to-Threshold . . . 174
8.5 Uniformly Irradiated Results . . . 177
8.5.1 Charge Collection . . . 178
8.5.2 Time to Threshold . . . 179
8.6 Non Uniformly Irradiated Results . . . 181
8.6.1 Fluence Profile . . . 181
8.6.2 Charge Collection and Time To Threshold . . . 182
8.6.3 Fluence Dependence . . . 184
8.7 Summary . . . 186
8.8 Outlook . . . 188
8.8.1 Possible extra studies . . . 189
8.8.2 Future sensor R&D . . . 189

**A Neutrino Mass Matrix Diagonalisation 191**

**B Background composition per region 193**

**C Normalisation channel yields 197**

**D Validation regions 199**

**References 202**

**Summary 217**

**Samenvatting 223**

**Acknowledgments 229**

# Introduction

The Standard Model of particle physics provides an accurate description of the fundamental constituents of matter and their interactions. Nevertheless, several experimental observations are not accounted for in the model. Among these, neutrino oscillations, predicted by Pontecorvo in 1957, have been observed in many experiments probing different interaction processes, in a range of neutrino sources from solar to atmospheric, accelerator and reactor experiments.
In order to oscillate, neutrinos mast be massive, opening the question of introducing a neutrino mass term in the Standard Model. Since the neutrinos are neutral particles, two possible mass terms can be added depending on the nature of the neutrino itself: the neutrinos could be Dirac fermions with a Yukawa coupling much smaller than the other fermions, or Majorana particles, that is, they are their own antiparticles. It remains a puzzle why neutrino masses are small in comparison to other fermions. Several models have been formulated that give a solution to the neutrino mass problem by introducing new particles. One of the most simple extension of the SM providing an explanation of neutrino masses, known as neutrino minimal SM, consists of the introduction of three right-handed heavy neutrinos.
Since the mass of heavy neutrinos and their couplings to the left-handed neutrinos can span a very wide range, many different experiments have attempted to establish their existence, probing different areas of the parameter space. The parameter space of the flavour mixing of the muon, electron and especially tau neutrino remain widely unexplored, in particular for heavy neutrino masses above 5 GeV. There is an ongoing effort from the LHC community in exploring the mixing of heavy neutrinos with active neutrinos. The LHCb detector, designed for the study of beauty and charmed hadrons, searched for a heavy neutrino with mass up to 5 GeV in the lepton flavour violating decay $B^- \to \mu^- N$ with $N \to \pi^+ \mu^-$ and $N$ is the heavy neutrino.
The first part of the thesis describes a new search is performed with the LHCb detector for a heavy neutrino in the mass range 5-50 GeV. In Chapter 1 the theoretical motivation for such a search is presented, as well as the current experimental status. The LHCb detector is described in Chapter 2: the different subsystems, the offline processing of the data and the event reconstruction, in particular jet reconstruction, are presented. Chapter 3 is dedicated to the first part of the analysis: the strategy and the selection of

data are described. The results, from fitting to limit setting, are presented in Chapter 4.

The LHCb detector is currently being upgraded (2019-2020) in order to enhance its physics capabilities: the instantaneous luminosity will increase of a factor five, with consequently higher pile-up, track multiplicity and radiation damage. Furthermore, the hardware trigger is removed and the subdetectors are read out at the collision rate of 40 MHz. In order to cope with such a harsh environment, all the tracking detectors need to be fully replaced and the electronics for all subdetectors upgraded.
In particular, the VErtex LOcator (VELO), which surrounds the proton-proton collision point and is dedicated to the reconstruction of primary and secondary vertices, is replaced by a pixel detector.
The sensitive elements of the new detector are hybrid pixel sensors with 55 µm pitch. The chip, called VeloPix, has been custom developed for the VELO upgrade in order to cope with the high data rate and radiation exposure. From the sensor point of view the main challenge is the high and nonuniform radiation exposure. Sensors of different geometry, bulk type and manufactures have been considered and have been characterised on a beam of minimum ionising particles to assess the best candidate for the upgrade.
The second part of the thesis is dedicated to the sensor characterisation. In Chapter 5 the upgrade of the VELO is introduced. The main principles of operation of silicon sensors are described in Chapter 6, with a particular focus on the consequences of radiation damage. In Chapter 7 the sensor characterisation is presented. First the requirements the sensor must satisfy are outlined, followed by the description of the Timepix3 telescope, used to test the prototype sensors. Several aspect of the sensor have been tested, but this thesis focuses on the behaviour of the sensor at the edge, which is the point closest to the beam in LHCb. In Chapter 8 the grazing angle technique is described, which consists of placing the sensor almost parallel to a beam of minimum ionising particles. With this technique the time and charge collection profile have been investigated in great detail, in particular to understand the effects of irradiation.

# Part I

# Heavy Neutrinos

# Chapter 1

# Theoretical Motivation

In the Standard Model (SM), discussed in Section 1.1, the neutrino is introduced without its right-handed counterpart. However it has been observed experimentally, in atmospheric, solar, reactor and accelerator experiments, that neutrinos oscillate between different flavours. Oscillations are predicted to occur if neutrinos have a mass. Therefore the introduction of right-handed neutrinos in the SM Lagrangian is required to explain the neutrino mass, mixing and the nature of the particle itself (Section 1.3). Many theoretical models introduce right-handed neutrinos, with masses ranging from order keV up to the GUT scale (Section 1.5). In particular, the neutrino minimal SM ($\nu$MSM) extension foresees a heavy neutrino which can be produced at current colliders, such as the LHC. The existence of heavy neutrinos in this mass range is investigated in this thesis. The current experimental status on heavy neutrino searches is summarised in Section 1.6. Finally, the search subject of this thesis is introduced in Section 1.7.

## 1.1 Neutrinos in the Standard Model

The Standard Model (SM) of particles describes the fundamental constituents of matter and their interactions. The particles can be distinguished according to their spin in fermions, of half integer spin, and bosons, of integer spin. Matter is composed of fermions of two types, quarks and leptons, which exist each in six different flavours arranged in three families, as schematically represented in Figure 1.1.
The interactions between fermions are mediated by the exchange of bosons. The fundamental forces included in the SM are the strong, electromagnetic and weak interactions, whereas a quantum description for gravity is not. A particle carries a specific charge only in relation to an interaction: both leptons and quarks undergo electromagnetic and weak interactions, therefore they carry electric and weak charge (weak isospin and hypercharge), while only the quarks interact strongly and carry colour charge. The electromagnetic interaction between electrically charged particles is mediated by the photon, which is massless and electrically neutral. The weak interaction is mediated by

three massive bosons, $W^{\pm}$ and $Z$: the $Z$ boson is electrically neutral and has a mass of about 91 GeV, while the $W^{\pm}$ carry electric charge and have a mass of about 80 GeV. The strong force is mediated by eight gluons, massless and electrically neutral, which carry colour charge themselves.
The interactions arise in the SM as the result of imposing symmetries to the Lagrangian of free particles: the SM is required to be locally gauge invariant under the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group, where $SU(3)_C$ describes the strong interaction ($C$ stands for colour charge) and the combined $SU(2)_L \times U(1)_Y$ groups describe the electroweak interactions ($L$ denotes the weak isospin charge and $Y$ the hypercharge). The fermion fields enter into the Lagrangian as individual left-handed and right-handed components, $\psi_{L,R} = [(1 \mp \gamma_5)/2]\psi$ and $\overline{\psi}_{L,R} = \overline{\psi}[(1 \pm \gamma_5)/2]$, because left- and right-handed fermions have different transformation properties under the electroweak gauge group: only left-handed fermions carry weak isospin charge. All the fermions are arranged in left-handed SU(2) doublets and right-handed singlets, with the exception of the massless neutrino, which does not have a right-handed counterpart in the original restricted[1] Standard Model.
Fermions and gauge bosons acquire mass via the spontaneous symmetry breaking process denoted as Brout-Englert-Higgs mechanism [1, 2]. The particles acquire mass by interacting with a new scalar field, the Higgs field, which has non-zero vacuum expectation value. The term of the Lagrangian containing the fermion-Higgs interaction for the leptons is given by the Yukawa coupling:

$$-(F_\ell \overline{L}_L \phi \ell_R + \text{h.c.}), \tag{1.1}$$

where $F_\ell$ is the coupling to the Higgs field, which is proportional to the lepton mass, $\overline{L}_L = (\nu_\ell, \ell)$ is the left-handed lepton doublet, $\ell_R$ the right-handed singlet and $\phi$ represents the Higgs field as a isospin doublet of complex scalar fields with hypercharge $Y = 1$. Both left-handed and right-handed components of a particle are needed for such a Yukawa coupling to the Higgs field.
A well known analogue relation to Eq. 1.1 holds for the quarks. In the basis of the flavour eigenstates, the quark mass matrix is in general not diagonal. The unitary matrix which allows to change to the mass eigenstate basis is known as unitary quark mixing matrix or Cabibbo-Kobayashi-Maskawa (CKM) matrix [3]. The elements of this matrix represent the mixing of each up-type quark and each down-type quark and they therefore describe every charged current weak interaction. Weak charged currents are the only tree level interactions in the Standard Model that change flavor: by emission of a $W$ an up-type quark is turned into a down-type quark, or a $\nu_\ell$ neutrino is turned into a $\ell$ charged lepton.
In particle physics there are three fundamental discrete transformations: parity ($P$),

[1] Restricted SM refers to the formulation of the SM without right-handed neutrinos.

charge conjugation ($C$) and time reversal ($T$). Parity is a space inversion $P(t, \mathbf{x}) = (t, -\mathbf{x})$, charge conjugation changes a particle in its own antiparticle, time reversal reverses time $T(t, \mathbf{x}) = (-t, \mathbf{x})$. The $CP$ combined transformation replaces a particle by its antiparticle and reverses momentum and helicity. The combination of the three symmetries, $CPT$, is an exact symmetry in any Lorentz invariant local quantum field theory [4]. $C$, $P$ and $CP$ are exact symmetries of the strong and electromagnetic interactions, while they are not conserved in weak processes. Whereas $C$ and $P$ violation each are explicitly and maximally violated in the weak interaction, $CP$ violation is subtly encoded in the Standard Model of weak interactions as a complex phase in the CKM matrix.

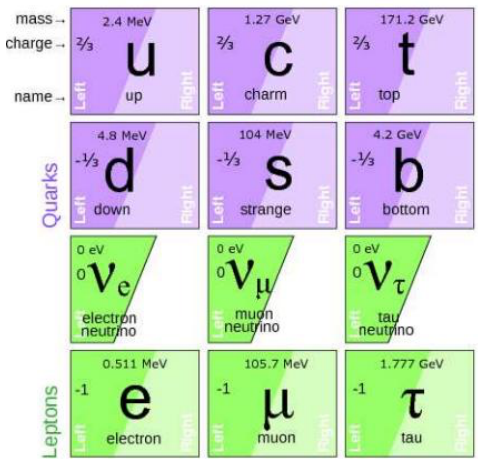



**Figure 1.1:** Standard Model particle content from [5]. To be noted, the absence of right-handed counterparts of neutrinos.

## 1.2 Limitations of the Standard Model

The restricted SM fails to explain several experimental observations in particle physics, astrophysics and cosmology, suggesting the existence of new particles and forces. Among the missing pieces of the SM, an effective mechanism for the generation of baryon asymmetry, dark matter and neutrino oscillations can be taken into account by extending the model in the leptonic sector, as discussed in Section 1.5.

The baryon asymmetry, which consists in the excess of matter over antimatter in the universe, is attributed to an excess of baryons produced in the first moments of the universe. There are three conditions, known as Sakharov conditions [6], that must be present in interactions in the early universe to generate matter and antimatter at different rates: baryon number violation, $C$ and $CP$ symmetry violation and interactions out of thermal equilibrium. Within the SM, $CP$ violation is included as the complex

phase in the CKM matrix, but the amount it accounts for is not enough to explain baryogenesis [7].
Furthermore the SM does not provide a candidate for dark matter. The existence of dark matter is established by different astrophysical and cosmological observations, starting with a larger velocity dispersion of galaxies in the same group [8] and the measurement of rotation curves in spiral galaxies [9,10]. Both observations are evidence of the presence of additional mass with respect to known baryonic matter. By studying the correlations in the angular spectrum of the cosmological microwave background, it has been concluded that the baryonic matter accounts for only 5% of the matter in the universe, while 27% is attributed to dark matter and even 68% to dark energy [11]. Dark matter could be composed of particles that have not yet been discovered. The properties of a dark matter particle candidate are in general unknown; it interacts gravitationally, but not strongly nor electromagnetically with the ordinary matter and must be stable or have a lifetime longer than the age of the universe. Dark energy refers to a form of energy different from matter and radiation that is responsible for the accelerated expansion of the universe [12].

### 1.2.1 Neutrino oscillation

The most compelling evidence for extending the leptonic sector of the SM comes from neutrino oscillations. Neutrino oscillations are transitions between neutrinos of different flavour. These have been established in solar, atmospheric, accelerator and reactor experiments. The fact that neutrinos oscillate means that the flavour states $\nu_{\alpha=e,\mu,\tau}$ are a superposition of the mass eigenstates $\nu_{i=1,2,3}$. Similarly to the CKM matrix, this can be implemented in the SM by introducing a unitary mixing matrix in the neutrino sector, called the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. In the same way as the CKM matrix, the PMNS matrix can be parameterised in terms of three mixing angles $(\theta_{12}, \theta_{23}, \theta_{13})$ and a *CP* violating phase $\delta$:

$$U_\nu = V^{(23)} U_\delta V^{(13)} U_{-\delta} V^{(12)}, \tag{1.2}$$

where $U_{\pm\delta} = \mathrm{diag}(e^{\mp i\delta/2}, 1, e^{\pm i\delta/2})$ and the matrices $V^{ij}$ are:

$$V^{(2,3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}, V^{(1,3)} = \begin{pmatrix} c_{13} & 0 & s_{13} \\ 0 & 1 & 0 \\ -s_{13} & 0 & c_{13} \end{pmatrix}, V^{(1,2)} = \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.3}$$

with $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. The oscillation probability of a neutrino depends on the three mixing angles, $\delta$ and two independent squared mass differences. If the neutrino is its own antiparticle, a so called Majorana neutrino (this is discussed in Section 1.3), two additional *CP* violating phases $\alpha_1, \alpha_2$ need to be added to the matrix

as $U_\nu \mathrm{diag}(e^{i\alpha_1/2}, e^{i\alpha_2/2}, 1)$, but they do not affect the oscillation probability.
The oscillation is observed as the appearance or disappearance of a given neutrino flavour starting from a know neutrino flux. Here the most recent observations are summarised for each of the three angles:

$\theta_{12}$
: Various experiments, first the so-called chlorine [13] and gallium [14–16] experiments followed by Kamiokande [17] and SuperKamiokande [18], observed a deficit of the amount of solar $\nu_e$. The SNO experiment [19] also reported evidence for $\nu_e \to \nu_\mu, \nu_\tau$ appearance. Evidence for disappearance of $\overline{\nu}_e$ was produced by nuclear reactors and detected by KamLAND [20].

$\theta_{23}$
: Cosmic rays interact with the nuclei in the atmosphere producing hadronic showers of mainly pions and kaons, which themselves decay giving a flux of electron and muon neutrinos and antineutrinos. The first evidence of oscillations of atmospheric neutrinos was provided by SuperKamiokande [21], followed by IceCube [22]. Both experiments observed $\nu_\mu$ disappearance with similar sensitivity. Long baseline experiments, where the detector is placed at thousands of km from the accelerator which provides the neutrino beam, confirmed the result: measurements of $\nu_\mu$ disappearance and $\nu_e$ appearance have been performed with SuperKamiokande [23], MINOS [24] and NO$\nu$A [25] detectors. Evidence of $\nu_\mu \to \nu_\tau$ appearance had been reported by the OPERA [26] collaboration.

$\theta_{13}$
: Long baseline neutrino experiments tried to access $\theta_{13}$ in appearance mode in the process $\nu_\mu \to \nu_e$, confirming a value different from zero. The $\theta_{13}$ mixing was precisely measured afterwards in $\overline{\nu}_e$ disappearance mode by the reactor experiments, first Daya Bay [27], then confirmed by RENO [28] and Double Chooze [29].

The experimental results from neutrino oscillations agree with a parametrisation of the PMNS matrix with three mixing angles and two independent squared mass differences. Some anomalies have been reported, which are discussed in Section 1.5.
Experiments have only been weakly sensitive to the *CPV* phase of the PMNS matrix so far. Global fits has been performed on the neutrino oscillation data, giving the results reported in Table 1.1. An overview of the experimental measurements can be found at [30, 31].
The oscillation neutrino experiments provide information on the mass squared difference of active neutrinos. They do not determine the absolute mass scale. As a consequence, the mass hierarchy can be interpreted in two ways, referred to as normal and inverted hierarchy. With the convention that $\Delta m^2_{\mathrm{sol}} = \Delta m^2_{21} = m_2^2 - m_1^2 > 0$ and $\Delta m^2_{\mathrm{atm}} = |\Delta m^2_{31}|$

**Table 1.1:** Best fit values obtained from a global fit to the current neutrino oscillation data from [32]. The values are reported for normal hierarchy and in brackets for inverted hierarchy.

| Parameter | best fit | $3\sigma$ |
|---|---|---|
| $\Delta m^2_{21}$ $[10^{-5}\,\mathrm{eV}^2]$ | 7.37 | 6.93 - 7.96 |
| $\Delta m^2_{31(23)}$ $[10^{-3}\,\mathrm{eV}^2]$ | 2.56 (2.54) | 2.45 - 2.69 (2.42 - 2.66) |
| $\sin^2\theta_{12}$ | 0.297 | 0.250 - 0.354 |
| $\sin^2\theta_{23}$, $\Delta m^2_{31(23)} > 0$ | 0.425 | 0.381 - 0.615 |
| $\sin^2\theta_{23}$, $\Delta m^2_{32(31)} < 0$ | 0.589 | 0.384 - 0.636 |
| $\sin^2\theta_{13}$, $\Delta m^2_{31(32)} > 0$ | 0.0215 | 0.0190 - 0.0240 |
| $\sin^2\theta_{13}$, $\Delta m^2_{32(31)} < 0$ | 0.0216 | 0.0190 - 0.0242 |
| $\delta_{CP}/\pi$ | 1.38 (1.31) | $2\sigma$ : (1.00 - 1.90) ($2\sigma$ : (0.92 - 1.88)) |

or $|\Delta m^2_{32}|$ and since from observation $\Delta m^2_{\mathrm{sol}} \ll \Delta m^2_{\mathrm{atm}}$, there are two possibilities: $m_1 < m_2 < m_3$ (normal hierarchy) or $m_3 < m_1 < m_2$ (inverted hierarchy).

## 1.3 Nature of the neutrino

Neutrino oscillation experiments proved neutrinos have mass, therefore a mass term for the neutrino must be included in the SM Lagrangian. The neutrino, differently from the other fermions in the SM, is electrically neutral. Therefore it can have a so-called Dirac or Majorana nature. If the nature of a particle is Majorana, it means that the charge conjugate field is the field itself:

$$\psi^c \equiv C\overline{\psi}^{\mathrm{T}} = \psi, \tag{1.4}$$

where $C$ represents the charge conjugation operator. The charge conjugation operator reverses the internal quantum numbers of the field it acts on, transforming a particle into its antiparticle. For this reason, to be Majorana, the field must be electrically neutral. The Majorana nature implies that the two chirality states are not independent, as for a Dirac particle, but connected by the following relations:

$$\begin{aligned} \psi_R = \psi^c_L \equiv C\overline{\psi_L}^{\mathrm{T}}, \\ \psi_L = \psi^c_R \equiv C\overline{\psi_R}^{\mathrm{T}}, \end{aligned} \tag{1.5}$$

where the subscripts $R$ and $L$ indicate right-handed and left-handed components, respectively. This makes it possible to build mass terms only with left-handed components, as shown in the following. Another consequence of the Majorana nature, is the fact that there are processes violating the lepton number by two, for example in the decay

$W \to \ell N$ with $N \to \ell X$, where $\ell$ indicates a lepton, $N$ the heavy neutrino and $X$ anything else, the two leptons can have same charge. This makes it a very clear signature from the experimental point of view.
Two possible types of mass terms can be added to the Lagrangian [33]:

Dirac mass term.
It is assumed that there are three additional right-handed singlets. The left-handed neutrinos are referred to as active, because they are involved in charge current and neutral current interactions, while the right-handed ones are defined sterile. A sterile neutrino has hypercharge zero and weak isospin zero, being a singlet under $SU(2)$ symmetry. The mass term in the flavour eigenstates base has the typical fermion mass term structure:

$$\mathcal{L}^D = -\sum_{l',l} \overline{\nu}_{l'L} M^D_{l'l} \nu_{lR} + \text{h.c.}, \tag{1.6}$$

with $l', l = e, \mu, \tau$ and $M^D$ a $3 \times 3$ complex non-diagonal matrix. The left-handed flavour eigenstates are a linear combination of the mass eigenstates, given by the PMNS mixing matrix.

Majorana mass term.
If the neutrino is a Majorana particle, because of Eq. 1.5, the mass term can contain only the left-handed active neutrino fields,:

$$\mathcal{L}^M = -\frac{1}{2} \sum_{l',l} \overline{\nu}_{l'L} M^M_{l'l} (\nu_{lL})^c + \text{h.c.}, \tag{1.7}$$

with $M^M$ a $3 \times 3$ complex non-diagonal, symmetric matrix.

The most general mass term comprises both terms:

$$\mathcal{L}^{D+M} = -\frac{1}{2} \overline{\nu}_L M^M_L (\nu_L)^c - \overline{\nu}_L M^D \nu_R - \frac{1}{2} \overline{(\nu_R)^c} M^M_R \nu_R + \text{h.c.}, \tag{1.8}$$

where the generation indices have been suppressed. In Eq. 1.8 both the left-handed active flavour fields $\nu_L$ and the sterile right-handed fields $\nu_R$ enter as:

$$\nu_L = \begin{pmatrix} \nu_{eL} \\ \nu_{\mu L} \\ \nu_{\tau L} \end{pmatrix}, \quad \nu_R = \begin{pmatrix} \nu_{eR} \\ \nu_{\mu R} \\ \nu_{\tau R} \end{pmatrix}. \tag{1.9}$$

The mass term has three components: a left-handed and a right-handed Majorana mass term and a Dirac mass term. Once the Higgs mechanism is included in the formalism, the Dirac mass matrix $M^D = Fv$, with $v$ the Higgs vacuum expectation value and $F$ the

matrix of Yukawa couplings, while the left-handed Majorana mass term can not exist in the SM. In order for a left-handed Majorana term to exist, the Higgs field would need to have weak isospin -1 and hypercharge 2 or non-renormalisable contributions. Therefore, $M_L^M$ is set to 0. In the base $(\nu_L, \nu_R^c)$, the total mass matrix results in the following symmetric matrix:

$$M^{D+M} = \begin{pmatrix} 0 & M^D \\ (M^D)^{\mathrm{T}} & M_R^M \end{pmatrix}. \tag{1.10}$$

Since the $3 \times 3$ matrices in generation space are not diagonal, the flavour fields, which couple to the $W$ and $Z$ bosons in the electroweak interaction, are a superposition of the mass eigenstates.
A natural mechanism of neutrino mass generation, called see-saw type I, is based on the Dirac and Majorana mass term of Eq. 1.8.

## 1.4 See-saw mechanism

The see-saw mechanism provides an explanation for the smallness of the neutrino masses, assuming the neutrino nature is Majorana, by introducing new particles. Depending on the type of the particle, three different mechanisms can be distinguished at the tree level:

- Type I see-saw, which employs right-handed singlets;
- Type II see-saw, which employs a $SU(2)_L$ scalar triplet;
- Type III see-saw, which employs at least two $SU(2)_L$ fermion triplets.

In the following the focus is on type I see-saw, which is the simplest extension of the SM, making the leptonic sector similar to the quark sector. The type I see-saw, based on the Dirac and Majorana mass term of Eq. 1.8, follows from a few assumptions. First, the Dirac mass term is generated by the SM Higgs mechanism and therefore the elements of $M^D$ are expected to be of the same order of magnitude of quark or lepton masses. Then, the lepton number is violated with the Majorana mass term at a much larger scale than the electroweak scale, leading to eigenvalues of $M_R^M$ much larger than the ones of $M^D$. The block diagonalisation of Eq. 1.10 leads to the two mass matrices:

$$m_\nu = -\theta M_R^M \theta^T, \tag{1.11}$$

$$m_N = M_R^M + \frac{1}{2}(\theta^\dagger \theta M_R^M + M_R^{MT} \theta^T \theta^*), \tag{1.12}$$

where $\theta = M^D (M_R^M)^{-1}$ gives a rotation between active and sterile neutrinos and, following from the assumption that $M^D << M_R^M$, is very small. The matrices $m_\nu$ and $m_N$

are not diagonal. This is what leads to the neutrino oscillations. The unitary matrix $U_\nu$ that diagonalises $m_\nu$ corresponds to the PMNS leptonic mixing matrix. The analogue mixing matrix $U_N$ which diagonalises $m_N$ determines the oscillations between the sterile fields. Two sets of mass eigenstates are obtained: the light active neutrinos

$$P_L n_i = \Big(U_\nu^\dagger\Big(\Big(1 - \frac{1}{2}\theta\theta^\dagger\Big)\nu_L - \theta\nu_R^c\Big)\Big)_i, \tag{1.13}$$

and heavy sterile neutrinos

$$P_R N_I = \Big(U_N^\dagger\Big(\Big(1 - \frac{1}{2}\theta^T\theta^*\Big)\nu_R + \theta^T\nu_L^c\Big)\Big)_I. \tag{1.14}$$

Here $P_{R,L}$ are chiral projector operators and, since the neutrinos are Majorana particles, the right- and left-handed counterparts of the eigenvectors are given by the Majorana relations Eq. 1.5. The derivation of the mass matrices and of the eigenvectors is presented in Appendix A for the one generation case.
The elements of the matrix $\Theta$ determine the mixing angles between the active and sterile neutrinos that are experimentally accessible:

$$\Theta_{\alpha I} = (\theta U_N)_{\alpha I} = (M^D M^{M-1} U_N)_{\alpha I}, \tag{1.15}$$

with $\alpha = e, \mu, \tau$ and $I = 1, 2, 3$. The Eq. 1.13 and 1.14 show that, since $\theta$ is small, the active neutrino is mainly mixing of $SU(2)$ neutrino fields, while the sterile neutrino is mainly mixing of singlet fields $\nu_R$. Since the right-handed neutrinos are singlets in the SM, there can be any number $N$ of them. Assuming that the see-saw is the only mechanism that allows the neutrinos to have mass, following from [34], there must be one right-handed neutrino per active massive neutrino. Therefore, a constraint on $N$ comes from neutrino oscillation measurements: since two non zero mass difference have been observed, then at least two active neutrinos are massive and $N \geq 2$.
Once the Higgs mechanism is included, the total Lagrangian, in the basis where $M^D$ and $M^M$ are diagonal, becomes:

$$\mathcal{L} = \mathcal{L}_{SM} + i\overline{\nu_R}\partial_\mu\gamma^\mu\nu_R - \big(\overline{L_L}F\nu_R\tilde{\phi} - \frac{1}{2}\overline{\nu_R^c}M^M\nu_R + \text{h.c.}\big), \tag{1.16}$$

where $\mathcal{L}_{SM}$ is the SM lagrangian, $F$ is the matrix of Yukawa couplings ($M^D = Fv$ with $v$ the Higgs vacuum expectation value), $\nu_R$ are the right handed neutrinos, $L_L = (\nu_L, e_L)^T$ are the lepton doublets and $\phi$ is the Higgs doublet. This is the most simple extension of the SM. It will be discussed in terms of theory models (Section 1.5) and experimental constraints (Section 1.6). The Lagrangian in Eq. 1.16 has a number of additional parameters on top of the 18 free parameters of the SM. If the number of heavy neutrinos is three, then the number of additional parameters is 18. The parameters can

be chosen to be: the three sterile neutrino masses, the three active neutrino masses, three mixing angles and three phases in each of the matrices that diagonalise $m_\nu$ and $m_N$. Many models, not discussed here, add further symmetry groups to the Lagrangian in an attempt to explain the mechanism that give a mass term to the right-handed neutrino.

## 1.5 Implications of neutrino mass

The order of magnitude of both the heavy neutrino mass and the Yukawa coupling constant with ordinary neutrinos are unknown. In the SM extension via the see-saw mechanism the sterile neutrino mass is a free parameter: it can acquire any value from the order of eV up to the GUT scale. According to [35], the Yukawa coupling can span from $\sim 10^{-13}$, which would correspond to the Dirac neutrino case, up to the strong coupling. Figure 1.2 shows the allowed values in the parameter space of the Yukawa coupling and heavy neutrino mass. Many different models based on type I see-saw mechanism have been proposed, which, depending on the mass of the neutrino, address some of the SM limitations mentioned in Section 1.2. The need of explaining specific experimental observations, like the existence of dark matter, leads to constraints on the properties of the heavy neutrinos, which can be tested experimentally.

### 1.5.1 Models

Four different categories of models can be identified:

GUT scale see-saw
: Models with heavy neutrino mass $10^9 < m_N < 10^{14}$ GeV are part of Grand Unified Theories. In such models, the baryon asymmetry of the universe is explained by *CP* violating decays of heavy neutrinos [36]. This leads to an asymmetry in the leptonic sector, which can be converted into baryon asymmetry via sphaleron transitions, preserving the $B - L$ number [37,38]. No candidate for dark matter is provided. Due to the high mass of the right-handed neutrino, these models cannot be probed experimentally and impact the stability of the Higgs boson mass against radiative corrections due to loop effects.

EWSB scale see-saw
: Models with heavy neutrino mass $\sim 10^2 - 10^3$ GeV, of the order of the electroweak symmetry breaking scale, can explain the baryon asymmetry of the universe under the assumption that the heavy neutrinos are nearly degenerate in mass. The mixing among them would enhance the lepton asymmetry in the process known as resonant leptogenesis [39,40]. Such models do not provide a dark matter candidate, but can be tested experimentally at current and future colliders.

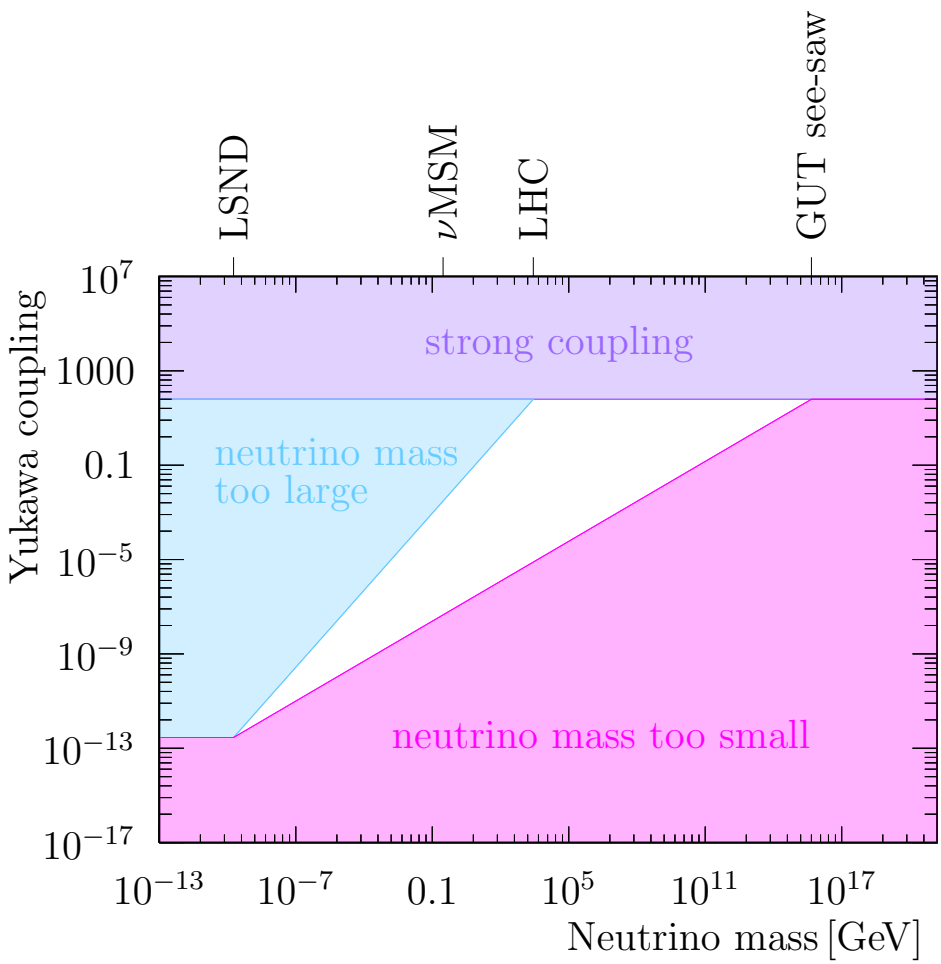



**Figure 1.2:** Visualisation of the allowed areas in the parameter space of Yukawa coupling and heavy neutrino mass, reproduced from [35]. Horizontally is the mass of the Majorana neutrino, vertically is the Yukawa coupling.

Low scale see-saw

Models with heavy neutrino masses below the electroweak scale and of the order of the masses of quarks and leptons, address many of the main issues of the SM. They can account for the baryon asymmetry in the universe [41] and, depending on the assumptions made, can even provide a candidate for dark matter [42], without the need of introducing a new energy scale. Moreover, the model predictions can be tested experimentally. The $\nu$MSM is a prominent example of such models and will be discussed below.

$\nu$ scale see-saw

Models with heavy neutrino mass $m_N$ on eV level are motivated by the anomalies observed in short baseline neutrino oscillation experiments. Four different types of experiments consistently observed a deviation from the expected oscillation rate.

The Liquid Scintillator Neutrino Detector (LSND) measured an excess of $\overline{\nu}_e$ in the process $\overline{\nu}_\mu \to \overline{\nu}_e$ [43]. The LSDN result was confirmed by the MiniBooNE experiment, which reported an excess combining the datasets from $\nu_\mu \to \nu_e$ and $\overline{\nu}_\mu \to \overline{\nu}_e$ oscillations [44]. Independently, the gallium solar neutrino calibration experiments (GALLEX and SAGE) measured a deficit in $\nu_e$ [45]. Finally, reactor neutrino experiments observed a lower flux of $\overline{\nu}_e$ than expected [46]. Among the several possible explanations, these $2-3\sigma$ deviations anomalies can be interpreted as neutrino oscillations with a heavy neutrino of mass about 1 eV and more experiments, not only short baseline neutrino, were proposed to test it (a summary of the operating and designed experiments can be found at [47]). These models can explain neither dark matter nor the baryon asymmetry of the universe.

The low scale see-saw models motivate the search presented in this thesis.

### 1.5.2 $\nu$MSM

An attractive model is the so called Neutrino Minimal Standard Model ($\nu$MSM) [48–51] due to its potential of explaining, at the same time, neutrino oscillations, dark matter and the baryon asymmetry, while applying only minimal changes to the SM. No new gauge group is added, the number of fermion families is unchanged and no new energy scale is introduced. The particle content of the model is illustrated in Figure 1.3.
The model is part of the low scale see-saw models category, with heavy neutrino masses all below the electroweak scale and therefore small Yukawa couplings, $\sim 10^{-6} - 10^{-12}$. The number of right-handed neutrinos introduced in the model is $n = 3$. The lightest right-handed neutrino is a candidate for dark matter, while the other two are responsible for the Baryon Asymmetry in the Universe (BAU) as well as for the mass generation of the active neutrinos via see-saw mechanism. The dark matter candidate has to fulfil a few conditions in order to comply with the astrophysical observations. First of all, as a viable dark matter candidate it has to interact very feebly, hence its Yukawa coupling should be relatively small, leading to the requirement of an active neutrino much lighter than the others, and the mass $M_1 \sim O(10\,\text{keV})$. To produce the baryon asymmetry in the universe the mass of the two heavier neutrinos should be in the range $150\,\text{MeV} \lesssim M_{2,3} \lesssim 100\,\text{GeV}$ and they should be nearly degenerate in mass in order to produce a sufficient amount of asymmetry in the leptonic sector in a resonant mechanism. Since the model provides clear predictions and covers a mass range accessible by colliders, it can be tested experimentally.
The search presented in this thesis is potentially sensitive to the two heavier neutrinos.

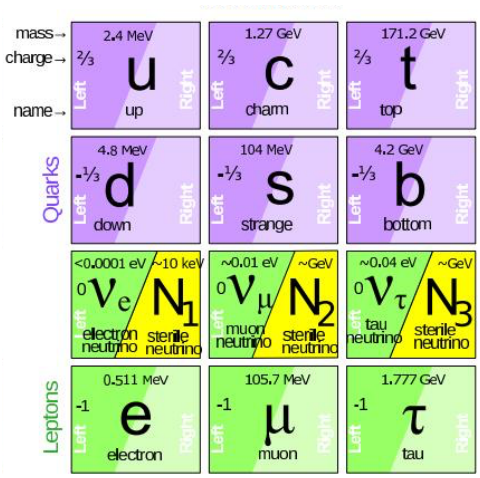



**Figure 1.3:** Particles in the neutrino minimal SM extension, taken from [5]. The sterile neutrinos shown in yellow are added to the particles of the SM.

## 1.6 Current experimental status

From the experimental point of view, direct searches for sterile neutrinos are performed through their mixing with each flavour of active neutrino independently:

$$V_{\alpha N}^2 = \sum_I \Theta_{\alpha I}\Theta_{\alpha I}^* = \sum_I \theta_{\alpha I}\theta_{\alpha I}^*, \tag{1.17}$$

where $\Theta$ is defined by Eq. 1.15, $\alpha = e, \mu, \tau$ and $I = 1, 2, 3$. Since heavy neutrinos have not been observed yet, the mixing between heavy neutrino flavours and the coupling of each flavour of sterile neutrino can not be currently probed. In addition to the mixing between each flavour of active neutrino and the sterile neutrino, another experimental observable is the mass of the neutrino. Limits are set in the parameter space of these two observables. Assuming there are $n \geqslant 2$ right-handed neutrinos that generate active neutrino masses via the see-saw mechanism, Figures 1.4, 1.5 and 1.6 respectively show the current limits and expected sensitivity of future experiments in the mass range $[100\,\text{MeV}, 80\,\text{GeV}]$ for the mixing of an electron, muon and tau neutrino with a heavy neutrino $N$. The already excluded areas of the parameter space are shaded in grey, while the lines represent the set bounds and expected sensitivity in high energy physics, neutrino physics and cosmology.

### 1.6.1 Cosmology and oscillation experiments

Lower bounds in mixing with each neutrino flavour are set by cosmology and active neutrino oscillation experiments. The most stringent cosmological constraints come

from the big bang nucleosynthesis (BBN), which sets lower bounds on the mixing of each flavour with the heavy neutrino, excluding a considerable portion of the parameter space at low mass and small mixing. The heavy neutrino lifetime should be less than one second, in order for the particle to decay before the BBN, such that the abundance of the light elements in the universe is not affected. The see-saw bound reported in the plots show which mixing would be excluded by type I see-saw: below the lower bound the mixing would not be sufficient to produce the values in Table 1.1 obtained from active neutrino oscillations experiments. Both the BBN and the see-saw limits depend on the assumption of the number of sterile neutrinos: here they have been derived assuming the existence of two heavy neutrinos, but they would become less stringent if the number of heavy neutrinos increases to three [52].

### 1.6.2 Indirect measurements

Upper bounds in the mixing with each neutrino flavour come from indirect measurement, specifically electroweak precision data (EWPD). The mass of the $Z$, the QED coupling constant $\alpha$ and the Fermi constant $G_F$ are electroweak observables measured with high accuracy. While the first two are not significantly affected by the mixing with a heavy neutrino, the Fermi constant measured in muon decays $G_\mu$ is related to $G_F$ through the heavy neutrino mixing with both the electron and muon active neutrinos. As a consequence, measurements of lepton universality, the unitarity of the CKM matrix and the invisible width of the $Z$ are all sensitive to the existence of a heavy neutrino. These observables are included in a fit which leads to a limit on the mixing [53], labelled in the plots as EWPD.

### 1.6.3 Direct measurements

Large areas of the allowed parameter space left are covered by direct searches. It is clear that some regions of the parameter space are less probed than others, due to the fact that each area poses new and different experimental challenges. The smaller the mixing and the smaller the mass, the longer the lifetime of the heavy neutrino, which requires different search strategies. The bounds set by direct searches are model independent, the only assumption being that the active neutrinos acquire mass via type I see-saw.

Mass region below $K$ mass

For both the electron and the muon mixing, the region below the $K$ mass ($\sim$ 500 MeV) is well explored. The constraints come from peak searches in the light meson decays $\pi \to \ell\nu$ and $K \to \ell\nu$ [54], where $\ell$ is a charged lepton, as well as

beam dump experiments. The first type of searches consist in measuring the energy spectrum of the emitted lepton. If a heavy neutrino is produced in the decay, an additional peak would appear in the momentum spectrum of the charged lepton. Various experiments [55–59] performed this kind of search; no heavy neutrino has been observed and upper limits have been set in the mass range probed. The most stringent bounds are given in Figures 1.4 and 1.5. Beam dump experiments probe neutrino masses up to the $B$ mass. Due to the high background near the target it is not possible to detect any particle produced in association with the neutrino, so the heavy neutrino is reconstructed from its decay products. These are powerful probes of heavy neutrino mixing with all the flavours: limits at the 90% confidence level are set by PS191 [60], NA3 [61], CHARM [62], IHEP-JINR [63], BEBC [64], FMMF [65], NuTeV [66] and NOMAD [67].

Mass region between $K$ and $B$ mass

In the mass range between the $K$ and the $B$ mass Belle and LHCb set stringent constraints on $N$ produced in heavy mesons decay. Belle looked for the inclusive decay $B \to X\ell N$ with $N \to \ell\pi$ and $\ell = e, \mu$ using $772 \times 10^6$ $B\overline{B}$ pairs collected at the $\Upsilon(4S)$ resonance. In absence of a signal, exclusion limits at 90% C.L. have been set on the mixing with both the electron and the muon neutrino as a function of mass and lifetime [68]. LHCb searched for the lepton number violating decay $B^- \to \mu^- N$ with $N \to \pi^+\mu^-$ using $3\,\mathrm{fb}^{-1}$ of integrated luminosity collected at the centre-of-mass energy of 7 and 8 TeV and set an upper limit on the mixing with the muon neutrino [69, 70]. Lepton number violating decays are forbidden in the Standard Model and therefore provide a striking signature from the experimental point of view. They also represent the only direct experimental probe for Majorana nature of the neutrino. A summary of the lepton number violating decays can be found in [71], where it is shown that the lepton number violating decay $K \to \ell\ell\pi$ with same sign leptons is expected to be the most sensitive among the lepton number violating decays on the mixing with both the electron and the muon neutrino.

Mass region between $B$ and $W$ mass

The only experimental bounds set in the range above 5 GeV up to the $W$ mass are from LEP and recently from CMS. The strongest limit up to date are from L3 [72] and DELPHI [73]: both experiments searched for a heavy neutrino produced in $Z$ decays, using 424000 and 3.3 million hadronic $Z$ decays, respectively. All visible decays of the heavy neutrino are included in the analysis ($N_I \to \nu Z^*$ with $Z^* \to \ell\overline{\ell}, q\overline{q}$ and $N_I \to \ell' W^*$ with $W^* \to \overline{\ell}\nu, q\overline{q}'$ and $\ell = e, \mu, \tau$) and different search strategies are employed for both a prompt and displaced heavy neutrino decay vertex. No signal has been observed and upper limits have been set at 95%

C.L. on the mixing for each generation. A search performed by CMS [74] using $35.9\,\mathrm{fb}^{-1}$ at the centre of mass energy of 13 TeV covers the same range as the LEP searches, ranging from 1 GeV up to 1.2 TeV. The heavy neutrino is produced in the leptonic decay of a $W$ boson, which can be on-shell or off-shell depending on the neutrino mass, and decays promptly to $W^*\ell$ with the $W$ boson decaying leptonically to $\ell\nu$ ($\ell = e, \mu$). The signature consists of three prompt charged leptons in any flavour combination of electrons and muons. Since no excess has been observed above the SM background, upper limits have been set at 95% confidence level on $|V_{eN}|^2$ and $|V_{\mu N}|^2$.

Mass region above $W$ mass

Direct searches for on-shell production of $N$ have been performed by both ATLAS and CMS at a centre of mass energy of 7 and 8 TeV, leading to 95% upper limits on the mixing with the first and second family for neutrino masses up to 500 GeV. The decay under study is the lepton number violating decay to two same sign leptons and two jets, with no missing energy, where the neutrino is assumed to decay promptly. For neutrino masses lower than the $Z$ mass the limits set are comparable to the bounds by DELPHI. The same decay has been studied by the CMS collaboration using 35.9 $\mathrm{fb}^{-1}$ of data at the centre of mass energy of 13 TeV. This analysis also accounted for the photon initiated production of a heavy neutrino $q\gamma \to N\ell q'$, increasing the sensitivity for higher neutrino masses. This allowed to probe masses up to 1.6 TeV and set the most restrictive direct limits for heavy neutrino masses above 430 GeV. In the $\tau$ sector, there are no limits from collider experiment, mainly due to the difficulties in reconstructing the $\tau$ with high efficiency.

### 1.6.4 Future experiments

Upgrades of ATLAS, CMS and LHCb as well as the $B$-factories, can increase the sensitivity on the heavy neutrino mixing with each flavour. Significant contributions are also expected from possible future experiments like SHiP, FCC-ee and the Long Baseline Neutrino Experiment (LBNE). SHiP is a fixed target experiment which is expected to cover a big portion of the parameter space in the low mass regime for the mixing of the heavy neutrino with each flavour. The projected sensitivities on the mixing are of the order of $10^{-9}$. A similar mass range will be covered by LBNE, with expected sensitivity competitive with SHiP in the $\tau$ sector, but less stringent for the mixing with the other two neutrino flavours, assuming normal hierarchy and five years data taking with the near detector. For high neutrino masses, up to 60 GeV, FCC-ee is expected to dominate, pushing the sensitivity to even smaller mixing values than SHiP, assuming $10^{12}$ $Z$ boson

decays with a 10-100 cm decay length. If these planned experiments proceed and reach the sensitivity expected, most of the parameter space between 100 MeV and 80 GeV will be probed.

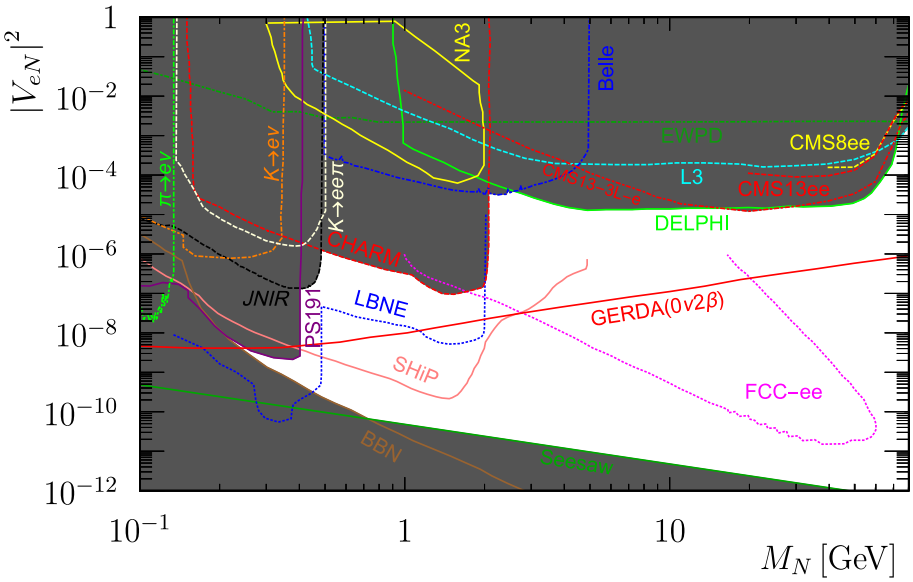



**Figure 1.4:** Limits on the mixing between a heavy neutrino and the electron neutrino in the mass range $0.1 - 80$ GeV, from [75].

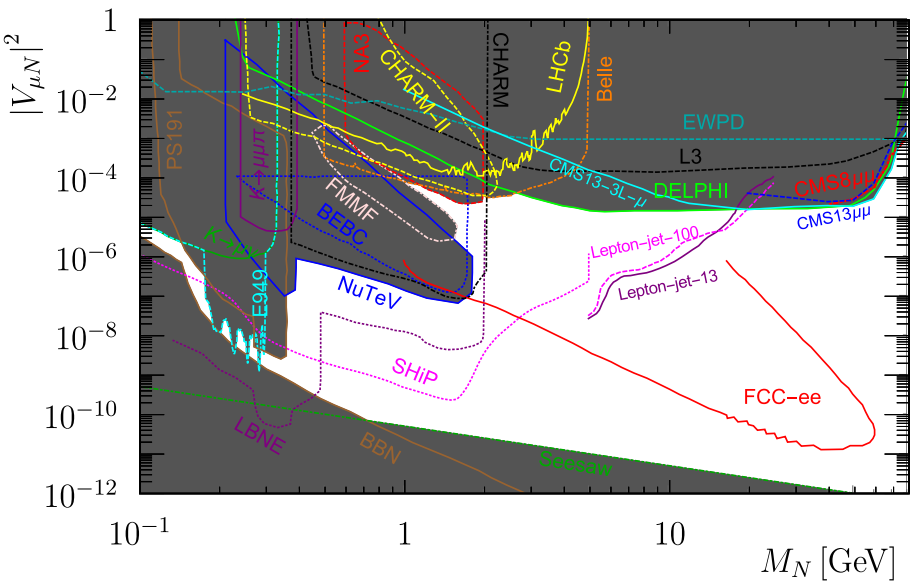



**Figure 1.5:** Limits on the mixing between a heavy neutrino and the muon neutrino in the mass range $0.1 - 80$ GeV, from [75].

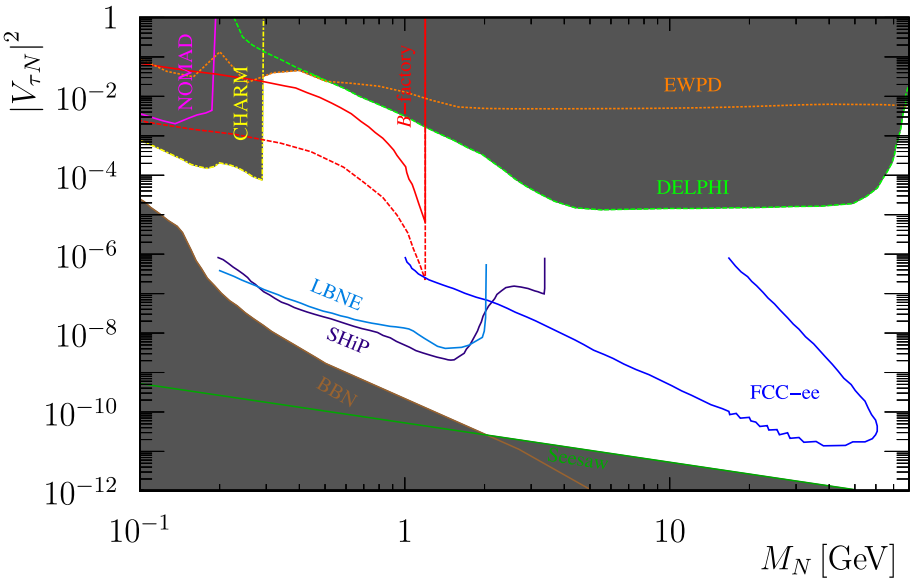



**Figure 1.6:** Limits on the mixing between a heavy neutrino and the tau neutrino in the mass range $0.1 - 80$ GeV, from [75].

## 1.7 Targeted mass range

The aim of this analysis is to look for a signal for the mixing of the sterile neutrino with the muon neutrino, taking advantage of the high reconstruction efficiency for muons at LHCb. The mass range covered is 5 to 50 GeV, aiming at reaching a better sensitivity than the DELPHI and CMS ones. In this region the heavy neutrino can be produced in gauge boson decays. The nature of the neutrino does not affect the production mode, while it has to be taken into account in the decay: the Majorana neutrino can decay into twice as many final states as the Dirac neutrino.

### 1.7.1 Production Mode

Sterile neutrinos interact only via the mixing with active neutrinos, implying that they can take part in any process involving SM neutrinos. The interaction between the active and sterile neutrino is determined by the mixing: the matrix element is the product of the matrix element for the SM process and the mixing $|V_{\alpha N_I}|^2$. The dominant production mechanism above the $B$ mass is via the decay of on-shell gauge bosons, $W \to \ell^{\pm} N$ and $Z \to \nu N$.

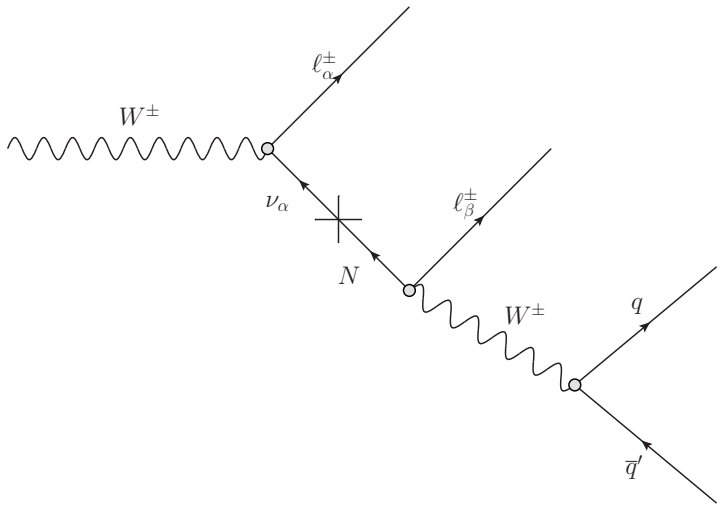



**Figure 1.7:** Feynman diagram for the production of a heavy neutrino via mixing with the active neutrino in the decay of a $W$ boson and semileptonic decay of the heavy neutrino into a lepton and two quarks.

### 1.7.2 Decay Mode

The heavy neutrino can decay leptonically and semileptonically via off-shell $W$ or $Z$ bosons. The branching ratios of the heavy neutrino are here computed based on the inclusive approach described in [76,77]. The inclusive approach approximates the semileptonic decays of the heavy neutrino by its decays into quark-antiquark pairs $N \to l(\nu)q_1\bar{q}_2$, since the masses probed in this analysis are sufficiently above the QCD confinement scale, whereas in the literature are mostly presented only decay widths of heavy neutrinos to bound states. In the decay the heavy neutrino mixes again with the SM model one, so the heavy neutrino decay width is proportional to the second power of the mixing between the active and sterile neutrino.
The decay modes relevant in the mass range under study, for charged (CC) and neutral current interactions (NC), are the following:

- leptonic charge current interaction

$$\Gamma(N \to l_1^- l_2^+ \nu_{l_2}) = |V_{l_1 N}|^2 \frac{G_F^2}{192\pi^3} m_N^5 I_1(y_{l_1}, y_{\nu l_2}, y_{l_2})(1 - \delta_{l_1 l_2}), \tag{1.18}$$

  where $y_i = m_i/m_N$.

- semileptonic charge current interaction

$$\Gamma(N \to l_1^- u\bar{d}) = |V_{l_1 N}|^2 |V_{ud}|^2 \frac{G_F^2}{64\pi^3} m_N^5 I_1(y_{l_1}, y_u, y_d), \tag{1.19}$$

  where $u = u, c, t$ and $d = d, s, b$. The expression is the same as for the previous

channel multiplied by the CKM matrix element and by a factor three for colour.

- leptonic charged and neutral current interaction

$$\Gamma(N \to \nu_{l_1} l_2^- l_2^+) = |V_{l_1 N}|^2 \frac{G_F^2}{192\pi^3} m_N^5 [(g_L^l g_R^l + \delta_{l_1 l_2} g_R^l) I_2(y_{\nu l_1}, y_{l_2}, y_{l_2}) + ((g_L^l)^2 + (g_R^l)^2 + \delta_{l_1 l_2}(1 + 2g_L^l)) I_1(y_{\nu l_1}, y_{l_2}, y_{l_2})], \quad (1.20)$$

where $g_L^l = -\frac{1}{2} + x_w$ , $g_R^l = x_w$ and $x_w = \sin^2\theta_w$ =0.231 , with $\theta_w$ Weinberg angle.

- semileptonic neutral current interaction

$$\Gamma(N \to \nu_{l_1} q\bar{q}) = |V_{l_1 N}|^2 \frac{G_F^2}{64\pi^3} m_N^5 [g_L^q g_R^q I_2(y_{\nu l_1}, y_q, y_q) + ((g_L^q)^2 + (g_R^q)^2) I_1(y_{\nu l_1}, y_q, y_q)], \quad (1.21)$$

where $g_L^u = \frac{1}{2} - \frac{2}{3}x_w$ , $g_L^d = -\frac{1}{2} + \frac{1}{3}x_w$ , $g_R^u = -\frac{2}{3}x_w$ , $g_R^d = \frac{1}{3}x_w$. The expression is again the same as for the previous channel multiplied by a factor three for colour, but no delta terms appear because there is no interference between NC and CC interactions.

- leptonic neutral current interaction (decay to invisible)

$$\Gamma(N \to \nu_{l_1} \nu_{l_2} \bar{\nu_{l_2}}) = |V_{l_1 N}|^2 (1 + \delta_{l_1 l_2}) \frac{G_F^2}{768\pi^3} m_N^5. \quad (1.22)$$

The kinematic functions are the following:

$$\begin{aligned} I_1(x,y,z) &= 12 \int_{(x+y)^2}^{(1-z)^2} \frac{ds}{s} (s - x^2 - y^2)(1 + z^2 - s) \lambda^{\frac{1}{2}}(s, x^2, y^2) \lambda^{\frac{1}{2}}(1, s, z^2), \\ I_2(x,y,z) &= 24yz \int_{(y+z)^2}^{(1-x)^2} \frac{ds}{s} (1 + x^2 - s) \lambda^{\frac{1}{2}}(s, y^2, z^2) \lambda^{\frac{1}{2}}(1, s, x^2), \end{aligned} \quad (1.23)$$

where $\lambda(a,b,c) = a^2 + b^2 + c^2 - 2ab - 2bc - 2ca$. The resonant decays have not been listed here since they are only relevant at lower heavy neutrino masses.
The total width has been computed as the sum of the partial decay width of charged current interactions and neutral current interactions. For a Majorana neutrino each partial decay width is multiplied by a factor two to account for decays to both charge conjugate final states. The contribution of CC would be 68% and NC 32% for a mass of 15 GeV.
In Figure 1.8 the branching ratio for decay modes with a muon in the final state have been plotted. The difference between the decay modes to quarks is only due to CKM matrix elements. The quark masses play a role only at low neutrino masses, where

overlapping curves slightly diverge (see for example $N \to \mu us$ and $N \to \mu cd$). The decay to $\mu\mu\nu$ is about two order of magnitude smaller than the one to jets, due to negative interference between CC and NC interactions.
From the experimental point of view, leptonic decays give a very distinctive signature, having only leptons in the final state, but the mass of the heavy neutrino can not be reconstructed due to the missing energy of the neutrino in the final state. On the contrary, in semileptonic decays the heavy neutrino can be fully reconstructed from its decay products.
In Figure 1.9, the lifetime is plotted as function of heavy neutrino mass assuming a coupling of $10^{-5}$, which is the current bound in the mass range of interest (Figure 1.5). In the low mass regime, the lifetime is of the order of 10 ps, while at higher masses the lifetime is so small that the decay can be considered prompt. Therefore two different signatures need to be considered, prompt and displaced, especially in view of the fact that for smaller coupling the lifetime will be longer.
In this thesis the process $W \to N\mu$ with prompt $N \to \mu q\overline{q}$ decay is investigated. Such a search exploits the larger decay width to $q\overline{q}$ pairs, the high muon reconstruction efficiency of LHCb and the full reconstruction of the on-shell $W$ from its decay products. The analysis is developed such that it can be easily extended to account for a long lived heavy neutrino, as described in more details in Chapter 3.

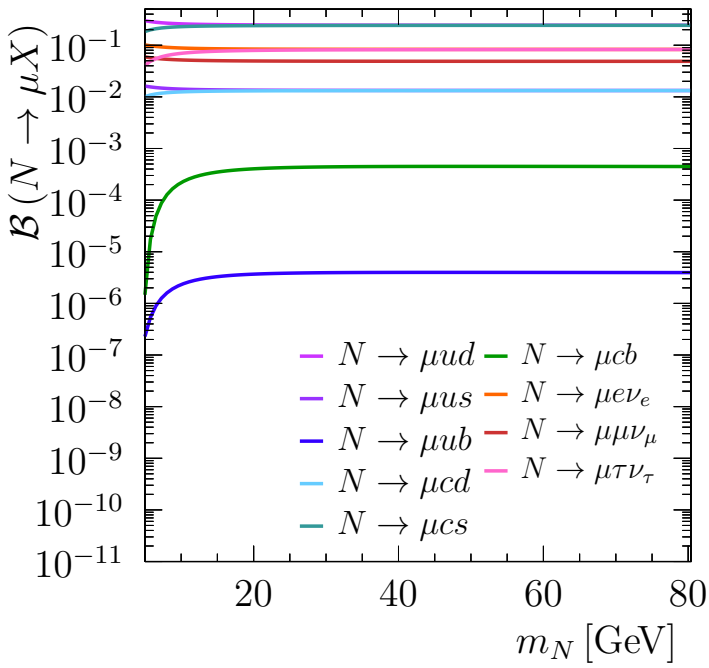



**Figure 1.8:** BR of heavy neutrino in the mass range of interest.

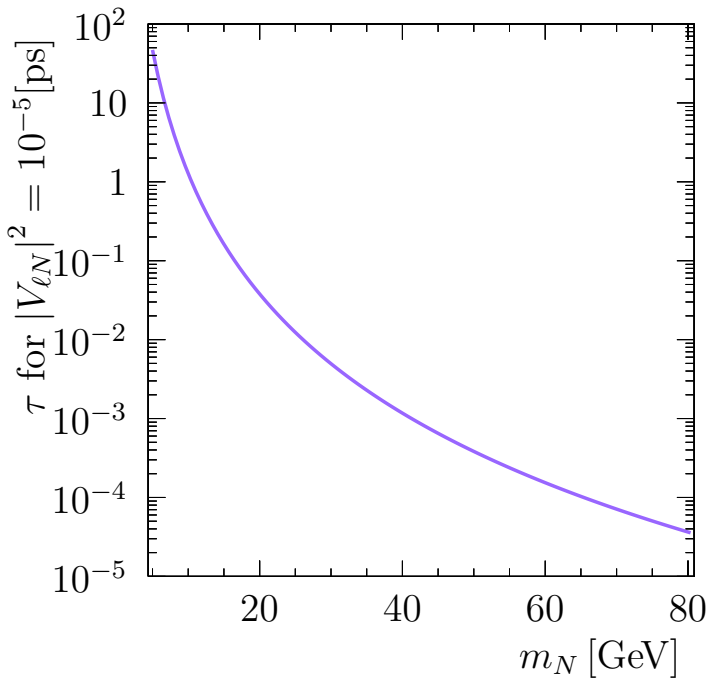



**Figure 1.9:** Lifetime of heavy neutrino in the mass range of interest assuming a coupling of the order $10^{-5}$ (current limit).

# Chapter 2

# The LHCb detector

The Large Hadron Collider beauty (LHCb) detector [78, 79] is one of the four major experiments at the CERN Large Hadron Collider (LHC). The main purpose of the experiment is to perform precision measurements of CP violation and detect rare decays of beauty and charm hadrons. Hence, the layout of the detector has been optimised considering in particular the $b\bar{b}$ production mechanism. At high energy $b\bar{b}$ pairs are produced in the same forward or backward cone due to the unequal momenta of the colliding partons. The detector is therefore designed in the forward direction, covering 10-300 mrad in the horizontal plane and 10-250 mrad in the vertical plane. The physics program is however not restricted to beauty and charm physics, but extends to electroweak measurements and exotica searches, as well as to heavy-ion physics.
Differently from ATLAS and CMS the luminosity in LHCb is levelled by defocusing and partially colliding the beams in the transverse plane, in order to maintain a lower and constant instantaneous luminosity. In this way, the average number of visible interactions per bunch crossing is about 1.1, with reduced occupancy and radiation damage in the subdetectors, as well as faster event reconstruction processing time.
The LHCb detector became operational in 2010. A first period of operations, the Run I, lasted two years with $3\,\mathrm{fb}^{-1}$ of integrated luminosity collected, of which $1\,\mathrm{fb}^{-1}$ at the centre-of-mass energy of 7 TeV (in 2011) and $2\,\mathrm{fb}^{-1}$ at 8 TeV (in 2012). The analysis presented here exploits the data collected during 2012. After a long shutdown (LS2), a second run started in 2015 and lasted until the end of 2018 with collisions at the centre-of-mass energy of 13 TeV, increasing the integrated luminosity to a total of more than $9\,\mathrm{fb}^{-1}$. The evolution of the integrated luminosity over Run I and Run II of data taking is illustrated in Figure 2.1. After this, a second long shutdown (LS2) of the LHC machine, from the end of 2018 until 2020, is used for a major upgrade of the experiment, described more in detail in Section 5.
The layout of the LHCb experiment is shown in Figure 2.2. A right-handed reference frame is adopted with the $z$ axis along the beam and the $y$ axis pointing up. The direction of a particle is typically indicated in terms of azimuthal angle $\phi$ and pseudorapidity

$\eta$, defined as $\eta \equiv -\ln(\tan(\theta/2))$ with $\theta$ polar angle with respect to the beam axis. The pseudorapidity is zero for a particle track perpendicular to the beam direction and approaches infinity as the track is produced along the $z$ axis. In terms of $\eta$, the LHCb detector covers the range $2 - 4.5$.
This chapter is dedicated to the description of the LHCb detector. Firstly, the subdetectors are introduced: in Section 2.1 the tracking system and in Section 2.2 the particle identification systems. Secondly, the trigger is described, followed by the offline processing of the data. Finally, the event reconstruction is outlined, with particular focus on jet reconstruction.

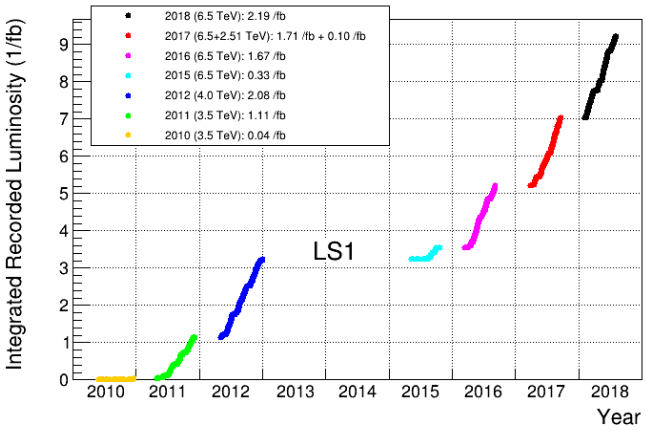



**Figure 2.1:** Integrated luminosity per year of data taking of the LHCb experiment.

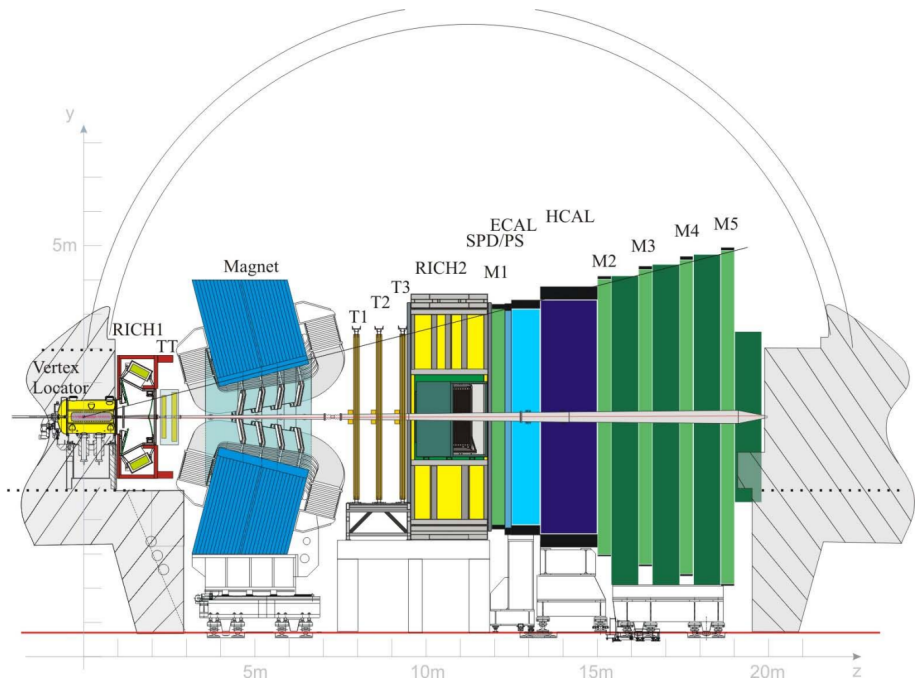



**Figure 2.2:** Side view of the LHCb experiment from [78].

## 2.1 Tracking detectors

Starting from the proton-proton collision point ($z = 0$ in Figure 2.2) and proceeding downstream, the LHCb tracking system consists of the VErtex LOcator, four tracking layers upstream the magnet, named Tracker Turicensis (TT), and 12 tracking layers downstream the magnet, the T stations. Since both the TT and the internal part of the T stations (IT) are silicon microstrips detectors, these have been developed together and are collectively called Silicon Tracker (ST), while the outer part of the T stations (OT) uses straw drift tubes technology.

### 2.1.1 VErtex LOcator

The VELO is a silicon microstrip detector that surrounds the proton-proton collision point. It represents the first part of the tracking system and is dedicated to the reconstruction of primary and secondary vertices. The VELO consists of two retractable halves that are moved out by 35 mm during the beam injection and closed when the beams are stable. During operation of the detector, the distance from the beam to the first sensitive element is 8.2 mm. Each half is formed by 21 modules arranged perpendicularly to the beam direction, where each module consists of two semicircular 300 µm thick sensors, one of which provides the $\phi$ angle, so called $\phi$ sensor, and the other one, named $R$ sensor, the radial position of the track at the $z$ position of the sensor. All sensors are n-on-n type, with the exception of two n-on-p sensors, installed to be studied in view of the upgrade. The front-end electronics is placed external to the semicircle and the signals are routed to the readout chip by a conductive double metal layer. The two different types of sensors are illustrated in Figure 2.3. The $R$ sensor consists of semicircular strips in segments of 45°, while the $\phi$ sensors are formed by radial strips arranged in an inner and an outer region. The strip pitch varies between about 40 µm and about 100 µm across the sensor. The layout of the modules along the beam axis has

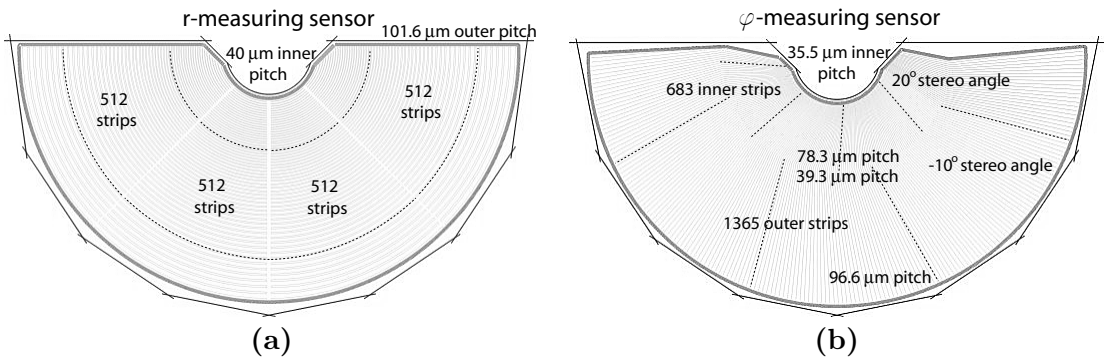



**Figure 2.3:** Sketch illustrating the geometry of $R$ (left) and $\phi$ (right) sensors. From [80].

been chosen based on the angular coverage and such that there is a minimum of six hits per track for all tracks with $2 < \eta < 4.5$. In order to be close to the interaction point and to minimise the material traversed by tracks, it has been decided not to install a beam pipe at the interaction region, but instead separate the vacuum of the beam from the one of the modules with a light-weight aluminium box called Radio Frequency (RF) box. Two RF-boxes are needed, each of which encloses a VELO half. The prime role of the RF box is to preserve the modules from RF pick up of the beam. The thin walls of the boxes that are facing the beam, referred to as RF foils, have an average thickness of 300 µm and are corrugated to minimise the material seen by tracks before they reach the sensitive elements.
Due to the vicinity to the beam, the highest fluence that the VELO has two withstand is of $5 \times 10^{13} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}\,fb^{-1}}$ for the closest sensor to the interaction point. The main origin of particle radiation is the prompt production of hadrons (pions, protons and neutrons). The irradiation, which is highly non uniform along the radius of the same sensors, causes a progressive degradation of the detector and is mitigated by increasing the operation voltage from 20 V up to 300 V, for some sensors, at the end of 2018. This allowed to operate the detector with fully depleted modules until the end of Run II. More details on the radiation damage on silicon detectors are discussed in Section 6.2. During operation the sensors are maintained at a temperature of about $-10°$ to mitigate the effects of radiation damage. The cooling is provided by a biphase $CO_2$ system, where the heat is transferred by evaporative cooling [81].

### 2.1.2 Silicon trackers

The sensitive elements of the silicon tracker detectors, collectively know as ST, are silicone microstrips of p-on-n type. The ST includes one station upstream of the magnet, the Tracker Turicensis (TT), and the central part of the three T stations downstream of the magnet, the Inner Tracker (IT). The former covers the acceptance of the LHCb detector, while the latter has a cross-like shape covering the high occupancy region around the beam. Each station consists of four detector layers placed in the stereo configuration $x$-$u$-$v$-$x$, where in the $x$ layers the strips are installed perpendicular to the beam and in the $u$ and $v$ layers are tilted of $+5°$ and $-5°$ with respect to the vertical direction, respectively.
The TT sensors are 500 µm thick with 183 µm pitch. Depending on the distance from the beam and therefore on the hit-occupancy, four different readout sectors are created where one, two, three, or four sensors are bonded together, as indicated in Figure 2.4. The readout is outside the acceptance, at the top or at the bottom of the module: the most external readout sectors are directly bonded to the readout, the internal sectors are connected via flex cables.

The geometry of the IT, shown in Figure 2.5b, reflects the particle occupancy. Each

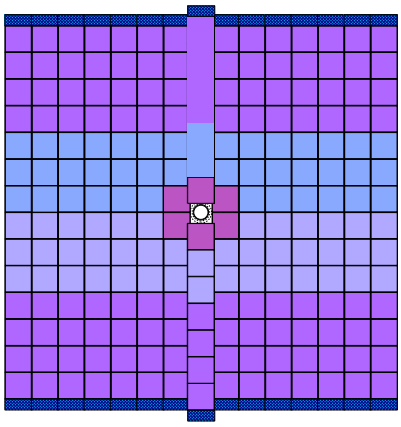


**Figure 2.4:** Sketch illustrating the geometry of a TT layer. The different colours indicate the different readout sectors. From [78].

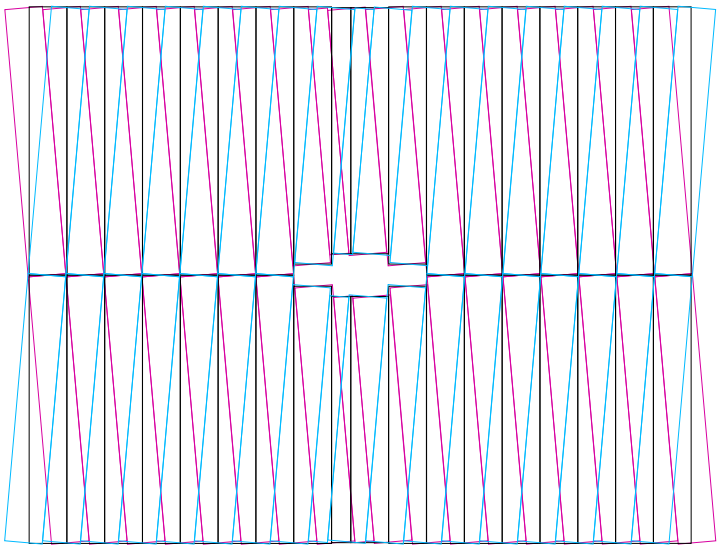
station consists of four detection boxes placed around the beam pipe: one row of sensors is in the top and bottom boxes, while in the side boxes two rows of sensors are bonded together. Both TT and IT are cooled using $C_6F_{14}$. The sensors are operated at a temperature of about 8° when the electronics is switched on. The highest fluence expected per $\text{fb}^{-1}$ is $4 \cdot 10^{12}$ for the TT and $2.5 \cdot 10^{12}$ for the IT.

### 2.1.3 Outer tracker

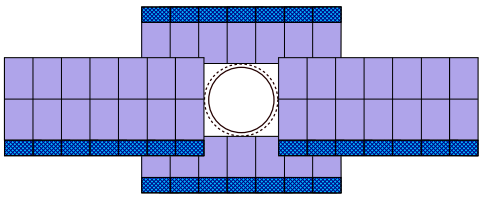
The remaining area of the T stations is covered by the outer tracker (OT), shown in Figure 2.5a. Since in this region the occupancy is lower due to the larger distance from the beam pipe, the requirement in terms of granularity and resolution is less demanding. The technology adopted consists of gas filled drift tubes: ionising particles traversing the straw tubes liberate charges which are collected by the anode wire in the middle of each tube and the arrival time is recorded. The gas used is a mixture of argon (70%), carbon dioxide (28.5%) and oxygen (1.5%), which allows for a fast drift velocity, resulting in a maximum drift time smaller than 50 ns. Each layer is composed by modules, where each module consists of two staggered layers of straw tubes enclosed in a gas-tight box.

### 2.1.4 Magnet

A dipole magnet is used to bend charged particles such that their momenta can be measured. It provides an integrated magnetic field 4T m for tracks of 10 m length. The magnetic field is directed along the $y$ axis, such that particle are bent in the $x$-$z$ plane. During data taking the polarity of the magnet has been reversed every couple of weeks in order to control detection asymmetries.



(a)



(b)

**Figure 2.5:** In (a) schematic illustrating the outer tracker from [82], with different colours for different layers of a station. In (b) sketch of a layer of the inner tracker from [78].

## 2.2 Particle identification system

The components of the PID system at LHCb are the two Ring Imaging CHerenkov detectors (RICH), the calorimeter system and the muon stations. The main role of the RICH system is to identify the charged hadrons $\pi$, $K$ and $p$, but also to contribute to the identification of electrons and muons, providing complementary information with respect to the calorimeter and muon systems, respectively. The calorimeter system provides the identification of photon, electron and $\pi^0$ candidates. The muon system is dedicated to the muon identification. The combination of the particle identification (PID) information from the different subdetectors is used to construct a likelihood adopted in the identification of a particle type.

### 2.2.1 RICH

When a particle traverses a medium with a velocity greater than the speed of light in that medium, it emits light in a cone, the so-called Cherenkov radiation. The cone opening angle depends on the particle velocity and the refractive index of the medium.

Combining the angle measurement from the RICH with the momentum measurement provided by the tracking detectors, the likelihood for a certain particle hypothesis can be built.

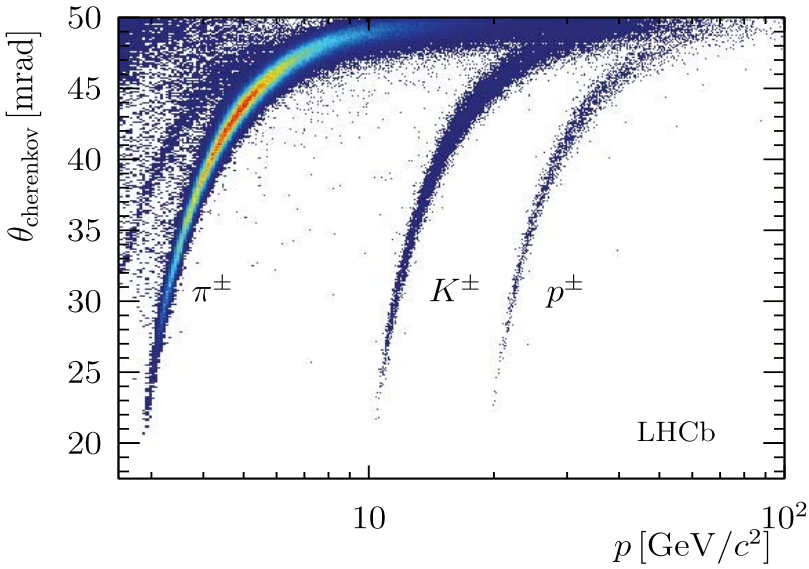



**Figure 2.6:** Reconstructed Cherenkov angle as a function of momentum of the track in $C_4F_{10}$ radiator. From [83].

RICH detectors are optimised for a range of track momenta by tuning the refractive index of the medium. In Figure 2.6 the reconstructed Cherenkov angle as a function of track momentum in the $C_4F_{10}$ radiator is shown. It can be seen that the separation power between different particle type is limited within a certain momentum range: a minimum velocity, and therefore mass dependent momentum, is needed to emit Cherenkov radiation, while at high momenta the opening angle saturates. The higher the refractive index is, the lower the momentum threshold for a given particle type, but also the lower the momenta at which the separation power is lost. For this reason, there are two RICH detectors in LHCb covering different momentum range. The RICH-1 detector is placed upstream the dipole magnet, between the VELO and the TT, and covers the low momentum charged particle range, about 1-60 GeV. The radiators used are aerogel[1] and $C_4F_{10}$ gas. The RICH2 detector is placed downstream the dipole magnet, between the last tracking station and the first muon station, and covers the high momentum charged particle range, from about 15 GeV to and beyond 100 GeV. The radiator used in the RICH2 is $CF_4$ gas. In both RICH detectors, spherical and flats mirrors are used in a

[1]The areogel has been removed in Run II due to degradation in performance in high track multiplicity environment.

particular configuration in order to focus the Cherenkov light. Cherenkov photons are detected in the wavelength range 200-600 nm by Hybrid Photon Detectors (HPD), which are shielded from external magnetic fields up to 50 mT by an iron cover.

### 2.2.2 Calorimeters

Downstream from the RICH-2 there is the calorimeter system, which identifies electron, photons and hadrons. It also selects high transverse energy hadron, electron and photon candidates for the hardware trigger and measures energy and position of the identified particles. The calorimeter system consists of two scintillating layers, the Scintillator Pad Detector (SPD) and the preshower (PS) detector, followed by two calorimeters, an Electromagnetic Calorimeter (ECAL) and an Hadronic Calorimeter(HCAL). The sensitive medium consists of scintillating material for all the subdetectors. The scintillation light is collected by Wavelength Shifting Fibres (WLS) and then transmitted to Photo Multipliers (PMTs) in the ECAL and HCAL, or to multianode photomultipliers (MAPMTs) in the SPD and PS.
Different particles have different penetration power, as illustrated in Figure 2.7. A thin lead layer separates SPD and PS to initiate the photon shower and hence distinguishes between electrons and photons: the former deposit energy in both subdetectors, while the latter only in the PS. Hadrons are not likely to deposit energy in SPD and PS . The ECAL measures electromagnetically interacting particles. The structure of the

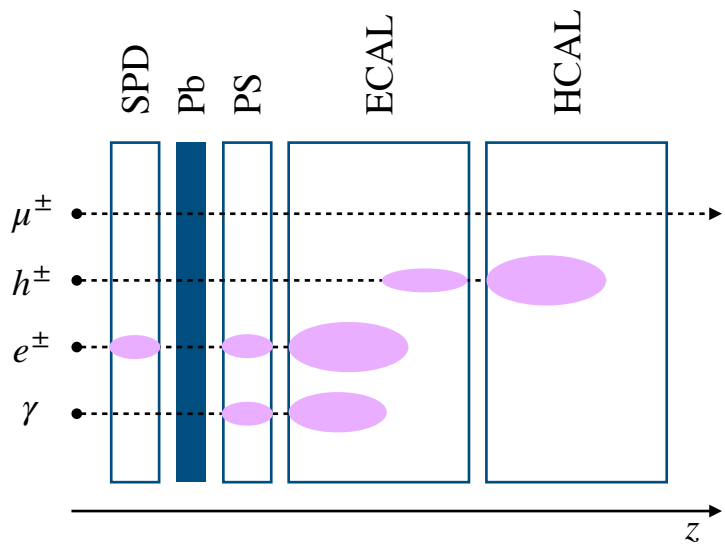



**Figure 2.7:** Sketch of the energy deposited by different types of particles in the calorimeter system. Reproduced from [84].

ECAL is formed by alternating layers of absorbers (2 mm thick lead layers) and active material. The ECAL has sufficient radiation length thickness to contain the shower

from high energy photons and hence to achieve an optimal resolution, measured to be $\frac{\sigma_E}{E} = 1\% \oplus \frac{10\%}{\sqrt{E(\,\mathrm{GeV})}}$ [85]. The HCAL has the same layers structure of the ECAL, but with lead layers twice as thick. Differently from the ECAL, the HCAL thickness is not enough to contain the hadronic showers. The energy resolution measures $\frac{\sigma_E}{E} = \frac{69\%}{\sqrt{E}(\,\mathrm{GeV})} \oplus 9\%$ [85]. Since the main purpose of the HCAL is to identify hadrons in the trigger, the limited energy resolution is not critical. Both the ECAL and HCAL detectors have regions with different granularity, as illustrated in Figure 2.8.

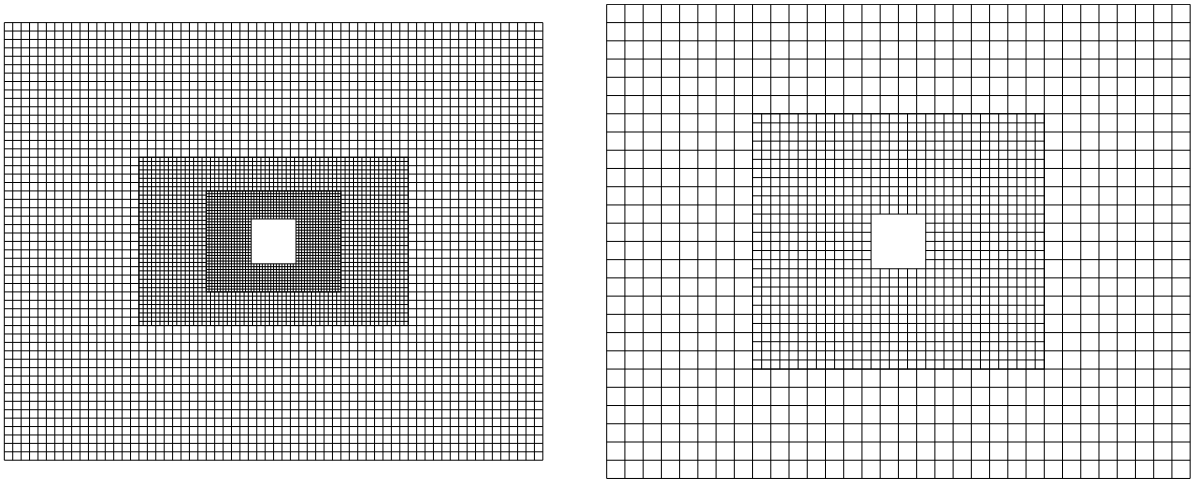


**Figure 2.8:** Schematic of ECAL (left) and HCAL (right) detectors, from [82].

### 2.2.3 Muon stations

The muon system plays a fundamental role in particle identification and in the hardware trigger, since it provides an estimate of the transverse momentum of the track which allows to identify and select high-$p_{\mathrm{T}}$ muons. The muon system consists of five stations placed along the beam axis. One station, M1 is placed upstream of the calorimeters to improve the transverse momentum measurement in the trigger. The other four stations, M2-M5, are located downstream of the calorimeters and are separated by iron absorbers 80 cm thick in order to select highly penetrating muons. The dimensions of the stations scale with the distance from the interaction point. Each muon station consists of four concentric regions, whose linear dimensions and segmentations scale as 1:2:4:8 such that the particle flux is about the same in every region. Two different technologies are adopted: gas multiplier foils in the inner region of M1, where the particle flux is the highest, and multi-wire proportional chambers (MWPC) in the outer region of M1 and in the four outermost stations. The minimum momentum for a muon to cross all five stations is 6 GeV.

## 2.3 Trigger

The trigger [86] system consists of three stages that collectively reduce the event rate from 40 MHz to a few kHz in order to store only the interesting events for further offline processing. The first stage of the trigger is implemented in dedicated electronics hardware, which reduces the readout rate to about 1 MHz. The events that survive are processed via the High Level Trigger (HLT), which is itself composed by two phases referred to as HLT1 and HLT2. In between HLT1 and HLT2 the data are buffered in disk storage such that preliminary alignment and calibration and partial reconstruction can be performed (including PID) and more analysis specific selections can be performed at the HLT2 level.

#### Level 0

The Level 0 trigger (L0) applies thresholds on the energy deposited in the calorimeters, on the transverse momentum of the track and on the number of hits in the SPD. The L0 lines[2] that select hadrons, photons and electrons apply requirements on the transverse energy of the reconstructed ECAL and HCAL clusters, while the single muon and di-muons lines apply them on the $p_{\mathrm{T}}$ of the track.
In the L0 trigger the transverse energy, deposited in clusters of $2 \times 2$ cells in ECAL and HCAL, is defined as:

$$E_{\mathrm{T}} = \sum_{i=1}^{4} E_i \sin\theta_i, \tag{2.1}$$

where $i$ identifies the cell, $E_i$ is the energy deposited in the cell $i$ and $\theta_i$ is the angle between the beam axis and the line between the centre of the cell and the average proton-proton collision point. The information from PS and SPD subdetectors allow to separate electron, hadron and photon candidates.
A muon candidate track is reconstructed as a straight line through the five muon stations, where the slope of the line is used to estimate the transverse momentum of the track assuming the muon originated from the interaction point. A threshold is then applied on the $p_{\mathrm{T}}$ of the one (single muon line) or two muon (di-muon line) candidates with the highest $p_{\mathrm{T}}$ in the event, where the transverse momentum has a resolution of about 25%. In the analysis of this thesis the single muon L0 line is used. A requirement for the number of hits in SPD to be lower than 600 is introduced to remove events with high occupancy, which would lead to a long processing time in the HLT. The dimuon trigger line has a looser cut on the number of SPD hits, at 900, since the rate is lower.

[2] As is common in LHCb, the set of requirements for a single trigger path is called trigger line.

**HLT**

Both HLT1 and HLT2 lines used in this thesis select high-$p_{\mathrm{T}}$ muons.
In addition to the calorimeter and muon systems, the HLT1 exploits also the information of the VELO and tracking detectors. Primary vertices are reconstructed from at least five VELO tracks and further muon identification is performed for events passing the muon or di-muon L0 lines by searching for hits around the extrapolation point of the track in the muon stations. The tracks satisfying IP requirements or identified as muons are then matched to IT and OT hits and fitted with a Kalman filter. The difference with respect to the offline fitting procedure is that a simplified material description is used and fewer iterations of the fit are performed. Specific cuts on impact parameter, momentum and mass can subsequently be applied. The HLT1 line used in this analysis implements requirements on the momentum, transverse momentum and quality of the track.
At the HLT2 level, forward tracking is performed for all VELO tracks and the information from all subdetectors is exploited. Furthermore, muon identification is applied for all the tracks and they are also matched to energy deposits in the calorimeter system. The HLT2 selects interesting final states according to inclusive and exclusive criteria. The high-$p_{\mathrm{T}}$ muon HLT2 line requires the $p_{\mathrm{T}}$ of the track to be larger than 10 GeV.

## 2.4 Offline data processing

Events that pass an HLT2 line are stored in a raw data format, which retains the full event information including information about the trigger selection. The events are processed with full detector alignment and calibration in order to reconstruct tracks and neutral particles and compute the particle identification likelihoods, as described in the next section. A process called *stripping* is then performed, which divides the data into different event streams. It consists in hundreds of sets of selections known as stripping lines which select individual events of interest for specific physics channels.

## 2.5 Event reconstruction

In this section an overview of the offline event reconstruction is presented, starting with the description of charged tracks and neutral particles reconstruction. Subsequently the particle identification procedure is discussed. For the work in this thesis the muon identification is most important and is described in more detail. Finally, since the decay under study $N \to \mu$jet has a jet among the decay products, it is illustrated how the jet reconstruction is performed in LHCb.

### 2.5.1 Track reconstruction

Tracks are reconstructed from the hits in the tracking subdetectors (VELO, TT, IT, OT). As shown in Figure 2.9, several track types are defined depending on which subdetector hits are used.

Long tracks are reconstructed using two different algorithms, called forward tracking

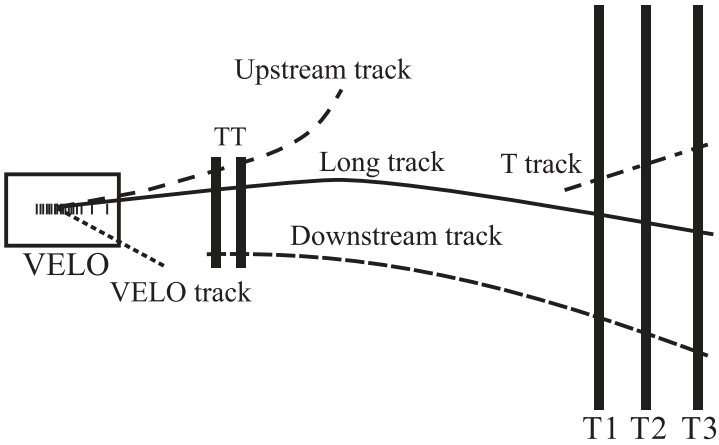



**Figure 2.9:** Sketch of track types reconstructed with the LHCb detector. From [79].

and track matching. For both algorithms, first VELO segments are reconstructed as straight line trajectories from at least three hits in the $R$ sensors and three hits in the $\phi$ sensors. In the forward tracking VELO tracks are extended to the T stations while in the track matching algorithm track segments in the T stations are reconstructed and matched to the VELO segments. The two set of candidates are combined removing duplicates. The TT hits that can be associated to the extrapolated trajectory are added to the track to improve the momentum determination and reduce the fraction of ghost tracks. Since hits from all the subdetectors are used, long tracks provide the best momentum resolution. For downstream and upstream tracks, T tracks and VELO tracks, respectively, are extrapolated to the TT station and compatible hits are added to the track. Upstream tracks correspond to particles bent out of the acceptance of the detector, while downstream tracks are used to reconstruct the decay products of long-lived particles decaying outside of the VELO acceptance. VELO segments that have been produced at large angles or in the backward direction are used for the vertex reconstruction. The reconstructed tracks are fit with a Kalman filter, which takes into account multiple scattering and corrects for the energy loss due to ionisation.

The track reconstruction efficiency is about 96% for tracks with $5 < p < 200\,\text{GeV}$ and $2 < \eta < 4.5$. The momentum resolution varies between 0.5% for low momentum tracks to about 1% for tracks with high momentum [87], as illustrated in Figure 2.10a.

The impact parameter of a track (IP) is defined as the distance of closest approach of the track from the primary vertex and is used to discriminate between prompt and

long-lived particles. Typically, the decay products of a prompt particle have smaller IP then the ones from a long-lived particle. This is extensively used in this analysis to select the prompt signal and reject the long-lived backgrounds from $b$ particles decays. The IP resolution as a function of $1/p_{\mathrm{T}}$ for long tracks can be seen in Figure 2.10. More details on the detector characteristics which affect the IP resolution are described in Section 5.2.

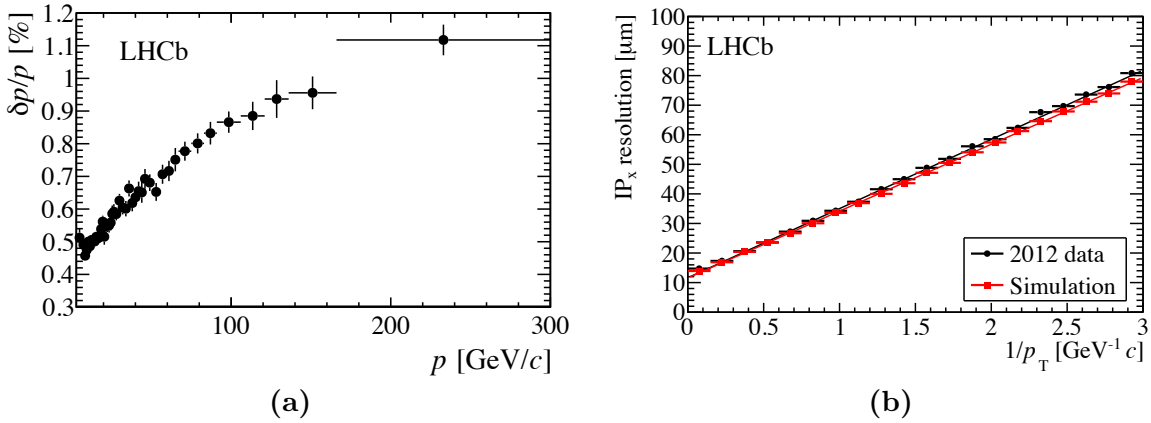



**Figure 2.10:** In (a) momentum resolution as a function of momentum and in (b) impact parameter resolution in $x$ as a function of the inverse of the transverse momentum for 2012 data (black) and simulation (red). From [79].

### 2.5.2 Neutral particle reconstruction

Neutral particles are reconstructed from energy deposits in SPD, PS, ECAL and HCAL that can not be associated to a charged track. The photon energy is obtained by adding the reconstructed energy deposited in the ECAL and in the PS. The photon direction is defined by the assumed primary vertex origin of the photon and the energy weighted position in the calorimeter system. Neutral pions $\pi^0 \to \gamma\gamma$ are reconstructed in two different ways depending on their momentum: as separated photons for low momentum $\pi^0$ and as a unique cluster for high momentum $\pi^0$, due to the finite ECAL granularity. The former are known as resolved $\pi^0$, the latter as merged $\pi^0$.

### 2.5.3 Particle identification

Particle identification variables are obtained by combining the information provided by the calorimeter, muon and RICH systems. Each of the subdetectors reconstruction algorithms provides a likelihood ratio between particle hypotheses for the reconstructed track, which are multiplied to obtain the combined likelihood $\mathcal{L}_{combi} = \mathcal{L}_{RICH} \cdot \mathcal{L}_{CALO} \cdot \mathcal{L}_{MUON}$, with:

- The calorimeter system likelihood $\mathcal{L}_{CALO}$, providing electrons, photons and $\pi^0$ identification. Energy deposits in the calorimeter that are not associated to tracks are due to neutrals. The shape of the neutral deposits is the main factor that allows for the distinction between photons and $\pi^0$. When the energy deposit is associated to a track instead, the likelihood for the electron hypothesis is computed to distinguish electron from hadrons. The likelihood is obtained by combining the likelihoods from ECAL, PS and HCAL, obtained from the energy deposits in each of the subdetectors.

- The RICH system likelihood $\mathcal{L}_{RICH}$, providing electron, muon, pion, kaon or proton identification. It is based on the reconstructed Cherenkov photons.

- The muon system likelihood $\mathcal{L}_{MUON}$, providing muon identification. The muon hypothesis and the not-muon hypothesis are computed based on the number of hits in the muon stations that can be associated to the extrapolated trajectory. Both the search windows around the extrapolation point and the number of hits required depend on the momentum of the track at each stations. Figure 2.11 shows the performance of the muon identification evaluated from data as a function of muon momentum and in bins of muon transverse momentum. The misidentification of pions and kaons is mainly due to decays in flight, while the misidentification of protons is due to the spurious hits in the muon stations along the proton direction. With increasing momentum the requirement on the search window becomes tighter and the number of hits to be associated larger, leading to a drop of misidentification rate for tracks with momentum between 3 and 30 GeV.
  Overall, the muon identification is about 97% efficient with 1-3% pion misidentification probability.

The variables labelled as $\mathrm{PID}_X$ provide a measure of how likely the mass hypothesis under consideration $X$ relative to the pion hypothesis is:

$$PID_X \equiv \Delta \log \mathcal{L}_{X\pi} = \log \mathcal{L}(X) - \log \mathcal{L}(\pi), \tag{2.2}$$

where $X$ can be a electron, muon, kaon or proton mass hypothesis.

### 2.5.4 Jet reconstruction

A jet is a collimated spray of charged and neutral particles produced by the fragmentation of quarks and gluons. Two different approaches are typically used to select the inputs for the reconstruction algorithms. In experiments such as ATLAS, with a good calorimeter resolution, the energy deposits in the calorimeters are the only inputs needed. Since in the LHCb detector the calorimeter cells are coarse and saturate at high energy,

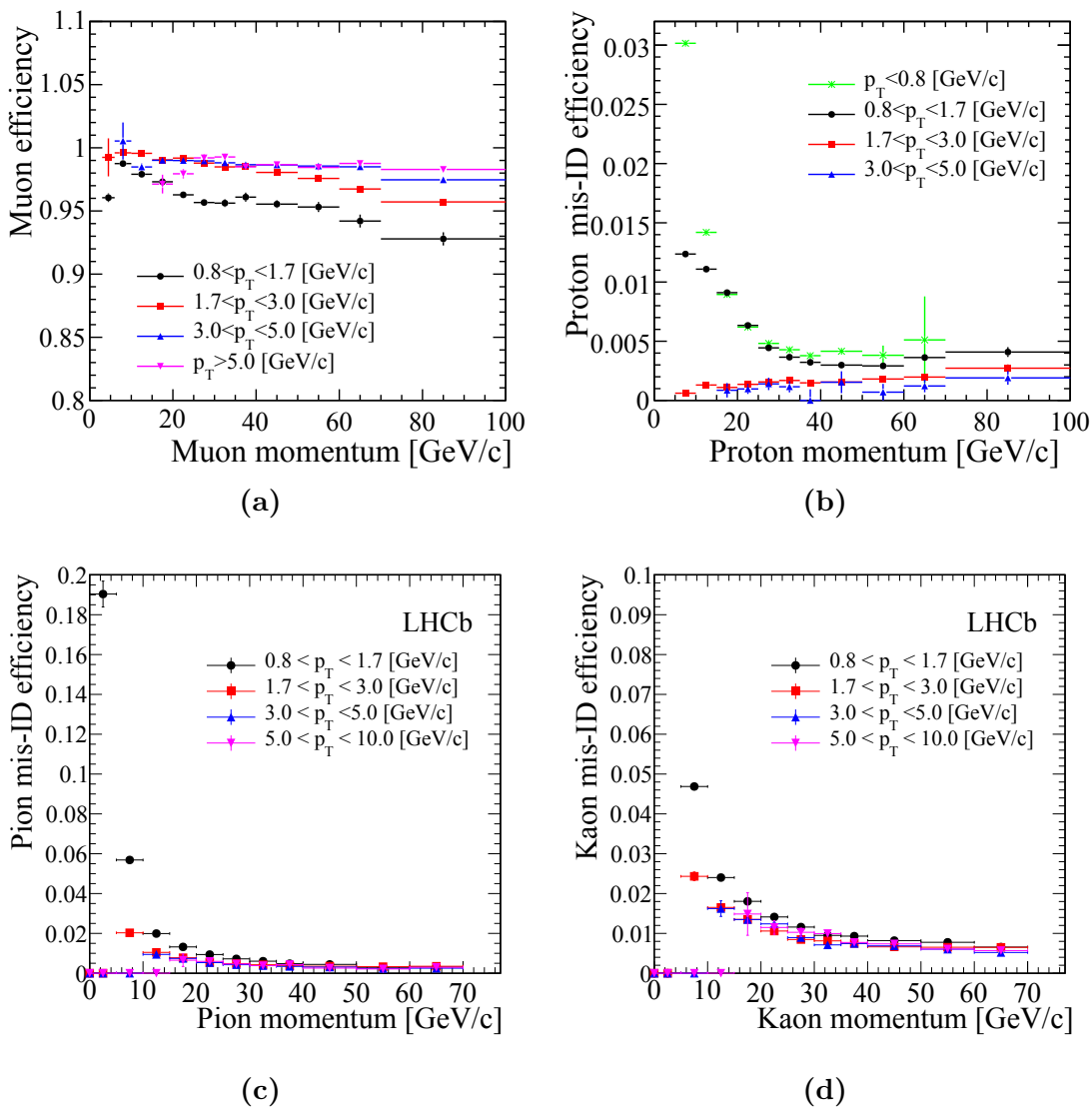



**Figure 2.11:** (a) muon identification efficiency, (b) proton misidentification probability, (c) pion misidentification probability, (d) kaon misidentification probability, as a function of the track momentum for four different transverse momentum ranges. From [79].

also the high momentum resolution tracking system is exploited.
In the following jet reconstruction with the LHCb detector is presented. First, the selection of the inputs for the jet reconstruction as performed in the so called Particle Flow is described, followed by the anti-$k_t$ algorithm for the clusterisation. Then jet energy scale correction and jet energy resolution are outlined.

### Particle Flow

Jets are reconstructed at LHCb using a particle flow algorithm [88,89]. Given that about 60% of the reconstructed jet $p_{\mathrm{T}}$ is carried by charged particles [90] and that the tracking system has excellent performance, the inputs are based on tracks for charged particles and calorimeter deposits for neutral particles. Both long and downstream tracks are selected as inputs. Since the momentum estimate of upstream tracks is quite poor, while they have good directionality, they are included among the inputs, but with null momentum in order to contribute only to the improvement of the jet direction. Tracks that can be combined in a neutral resonance are excluded from the inputs as individual tracks, while the resonance is included. There is also the possibility of banning specific particles from inclusion in the jets, for example in the case of this analysis muons with $p_{\mathrm{T}} > 2\,\mathrm{GeV}$, in order to avoid signal muon candidates to be included. Energy deposits in the ECAL that are not associated to a track and are classified as isolated (meaning that they are sufficiently far from any track) are used as inputs for neutral particles. There is also the possibility of recovering energy deposited by neutrals from ECAL clusters associated to tracks, in a procedure known as neutral recovery. To do this, first a particle hypothesis is assigned to the reconstructed track according to the information provided by the particle identification subdetectors. Once the track is identified either as a charged hadron, an electron or a muon, the predicted energy deposited is subtracted from the ECAL energy. If a significant amount of energy is left, this is classified as energy deposited by neutrals.

### Clusterisation

Several algorithms are available to perform the so called jet clusterisation, that is the combination of the inputs into jets. The main characteristics that a jet algorithm needs to satisfy are:

- collinear safety: the algorithm should be insensitive to the number of final state particles carrying the energy;
- infra-red safety: the algorithm should be insensitive to the emission of a soft gluon.

Infra-red and collinear safety are usually collectively referred to as IRC safety.
The jet size and shape determine the jet area, which should be large enough to capture non-perturbatively hadronised particles but not to include particles from the underlying event or pile-up events. The jet reconstruction algorithms can be distinguished into two main categories: cone algorithms and sequential clustering algorithms. Cone algorithms define a jet as a cone around the hardest seed. They provide smooth circular boundaries, but they generally are not IRC safe [91]. Sequential clustering algorithms start with the hardest particles and cluster other particles around them based on the measured distance

between them, which depends on the energy and spatial distance between the particles. Two distances are defined:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \tag{2.3}$$

$$d_{iB} = k_{ti}^{2p}, \tag{2.4}$$

where $\Delta R_{ij} = \Delta\eta^2 + \Delta\phi^2$ with $\eta$ and $\phi$ the pseudorapidity and azimuthal angle, respectively, $k_t$ is the transverse momentum, $R$ is the maximum cone radius and $p$ is an integer described below. The distance $d_{ij}$ is between the entities $i$ and $j$(particles and pseudojets) and $d_{iB}$ is the distance between the entity $i$ and the beam line $B$. If $d_{ij} < d_{iB}$ the entity $j$ is combined with $i$, while on the contrary, if $d_{ij} > d_{iB}$ $i$ is identified as a jet and removed from the list of entities. The procedure starts with the particles with smallest $d_{ij}$ and then it is iterated until all the entities have been assigned. The difference between several sequential clustering algorithms lies in the value of the exponent $p$. For $p = 1$, the so called $k_t$ [92] algorithm, the distance is dominated by soft particles, while for $p = 0$, known as Cambridge/Aachen [93] algorithm, the distance is independent of the transverse momentum of the particle. Both these algorithms are IRC safe, but the soft radiation brings irregularities to the jet boundaries. The currently mostly used algorithm for jet reconstruction is the so called anti-$k_t$ [91] algorithm, for which $p = -1$. In this case, the shape of the jet is not influenced by soft radiation while maintaining IRC safety. In the anti-$k_t$ algorithm, the distance between a hard and a soft particle is determined by the hard particle, hence soft particles tend to cluster first among hard particles than among themselves. The shape of the jet is only modified by hard particles:

- when another hard particle is present at a distance $R < \Delta R_{ij} < 2R$, two jets are produced: if there is an unbalance in transverse momentum, the hardest jet is conical and the other one partly conical; if the transverse momenta of the two jets are equal, the overlapping region is split between the two jets and neither of them is conical.

- when two hard particles 1 and 2 are within the maximum cone size, they are combined in one jet centred on the highest transverse momentum particle if the transverse momentum $k_1 \gg k_2$, while the shape is a bit more complicated if the two particles have similar transverse momenta $k_1 \sim k_2$.

The anti-$k_t$ jet algorithm within the FastJet package [94] is adopted in the LHCb collaboration. Typically it is required that the jet $p_{\mathrm{T}} > 10\,\mathrm{GeV}$, while the radius of the jet varies between 0.4 and 0.7.

### Energy scale correction

Several factors affect the reconstructed energy of the jet, such as the finite resolution of the calorimeter, pile-up, charged particles which are outside the acceptance of the detector and not-reconstructed neutrals. Therefore, a correction is applied to the reconstructed jet four-momentum to estimate the true jet momentum. The correction is a multiplicative factor evaluated from simulated $Z + \text{jet}$ events and is typically of order 10%, with variations up to 0.5% due to its dependence on the jet $p_{\mathrm{T}}$ and $\eta$ as well as on the number of primary vertices per event. The correction factor increases at high and low pseudorapidity because the efficiency of the calorimeter for detecting neutral particles decreases at the edge of the acceptance. It also decreases with increasing number of primary vertices since it is more likely to include incorrectly in the particle flow downstream tracks and energy deposit in the calorimeter which belong to a different primary vertex. As described in Section 4.5.1, an uncertainty between 2% and 5% is assigned to the jet energy correction to account for mismodelling of the detector response to jet in simulation.

### Jet energy resolution

The jet $p_{\mathrm{T}}$ resolution is evaluated from simulation, by comparing the true and reconstructed jets, and is found to be between 10% and 15% [90]. It has been observed that there is little dependence on pseudorapidity and on the number of primary vertices. The resolution is optimal for jets with $30 < p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and worsens for higher $p_{\mathrm{T}}$, as shown in Figure 2.12. For low transverse momentum jets, the resolution increases with $p_{\mathrm{T}}$ as typically the number of tracks increases reducing the effects of poorly measured tracks. For high transverse momentum jets, the calorimeter cells start to saturate and therefore part of the energy deposited is not recorded, and neutral recovered energy increases, leading to larger uncertainties on the energy of each cluster.

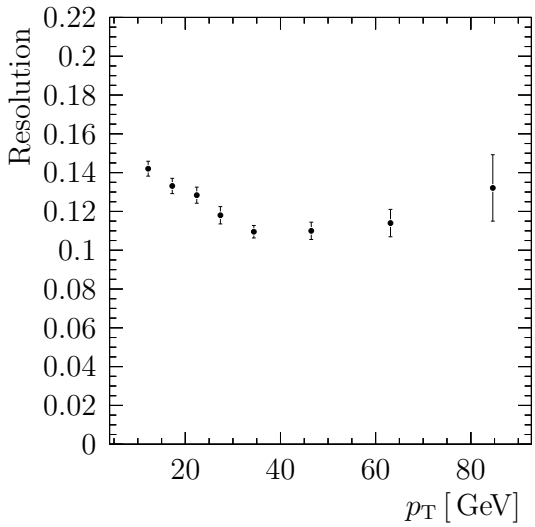



**Figure 2.12:** Jet resolution as a function of jet $p_{\mathrm{T}}$ for jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. From [90].

# Chapter 3

# Analysis Strategy and Data Selection

The search for right-handed neutrinos in the decay of an on-shell $W$ with the data recorded by the LHCb detector in the year 2012 is presented in this dissertation. The first part of the analysis, described in this chapter, consists of optimising the selection for the signal candidates and defining regions of the phase space enhanced in the main backgrounds, called control regions. The analysis strategy is outlined in Section 3.1, followed by the data and simulated samples for the signal and backgrounds in Section 3.2. In Section 3.3 the different steps of the signal selection are illustrated together with the evaluation of the corresponding efficiency. The control regions for the main backgrounds are defined in Section 3.4. The selections adopted for other decay channels, used for normalisation and for the evaluation of efficiency corrections and systematic uncertainties, are introduced in Section 3.5.

## 3.1 Analysis strategy

The aim of the analysis is to search for a heavy right-handed neutrino occurring in the decay of an on-shell $W$. The $W$ decays leptonically to a muon and the heavy neutrino, which itself decays to a muon and a quark-anti-quark pair, as shown in Figure 1.7. From the assumption that the heavy neutrino is a Majorana particle, it follows that in the final state both same- and opposite-sign muon pairs appear.
Dominant backgrounds are charged weak currents, in particular $pp \to W + X$ with $W \to \mu\nu$ and $W \to \tau\nu$, neutral electroweak Drell-Yan processes $pp \to \gamma/Z^{(*)} + X$ with $\gamma/Z^{(*)} \to \mu\mu, \tau\tau$, and heavy flavour $b\bar{b} \to X\mu$ and $c\bar{c} \to X\mu$.
The analysis strategy consists of splitting the dataset in a signal region and orthogonal control regions enhanced in different backgrounds. This is realised by training three different multivariate classifiers. The main backgrounds are normalised in their control

regions and the yields obtained extrapolated to the signal region, where a fit to the muon-jet invariant mass is performed to extract the signal yield for different heavy neutrino mass hypotheses. The decay channel $W \to \mu\nu$ is used to normalise the branching fraction of the decays to heavy neutrinos. In absence of a signal, an upper limit is set for both same- and opposite-sign muons decays on $\mathcal{B}(N \to \mu\text{jet})\,|V_{\mu N}|^2$, where $V_{\mu N}$ is the mixing between the active muon neutrino and a generic sterile neutrino $N$ (Eq. 1.17). The analysis has been developed as blinded, meaning that the data in the signal region have not been inspected until the analysis strategy had been consolidated. The data have been recorded by the LHCb detector in 2012 but data for the full Run I of the experiment will be included in future analyses. In this search the heavy neutrino is assumed to decay promptly, but as shown in Figure 1.9, for low masses it is expected to have a lifetime $\geq 10\,\text{ps}$. Hence, the current analysis has been developed such that it can be easily extended to look for a long-lived neutrino, as described in more details in Section 3.3.

The main analysis steps are the following:

- signal selection: it consists of two parts, first a preselection is performed followed by a tighter selection based on multivariate classifiers (Section 3.3);
- identification and selection of the main Standard Model backgrounds (Section 3.4);
- efficiency estimation from simulation for signal and normalisation channels (Section 3.3.4);
- description of the signal model adopted for the fit to the $\mu$-jet invariant mass: to estimate the contribution of backgrounds, fits are performed in the control regions and the yields are extrapolated to the signal region (Section 4.3);
- normalisation channel yield estimation (Section 4.2);
- systematic uncertainties evaluation (Section 4.5);
- limit setting by fitting the signal model for different signal hypothesis to the data (Section 4.6).

## 3.2 Data Samples

In this section the data samples used for the analysis are presented. Other than the data collected by the LHCb experiment, simulated samples have been generated for the signal and the main expected Standard Model backgrounds. Data and MC samples have been processed through the same processing chain.

### 3.2.1 Data

The analysis uses data collected by the LHCb experiment in the year 2012 at 8 TeV, corresponding to a sample of 1929 $\mathrm{pb}^{-1}$, of which 978 $\mathrm{pb}^{-1}$ magnetic field up (MU) and 951 $\mathrm{pb}^{-1}$ magnetic field down (MD). As described in Section 2.4, offline the data are split according to dedicated selections known as stripping lines. This analysis uses the so called `WMuLine` stripping line, which requires one muon from `StdAllLooseMuons` with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. Muons classified as `StdAllLooseMuons` satisfy a loose binary selection, known as *isMuon* [95], based on the number of muon stations hit within a field of interest around the track extrapolation. Both the required number of stations hit and the area of the field of interest depend on the momentum of the track candidate.
In total 14263610 MU and 14119263 MD single muon events are selected by the stripping.

### 3.2.2 Simulated Signal

Simulated samples for the signal decay $W \to \mu N$, where $N \to \mu q\overline{q}$, are generated with both opposite- and same-sign muons, in equal amount. Since the mass and lifetime of the neutrino are unknown, several samples were produced with different properties. Both MU and MD samples have been generated for six different heavy neutrino masses within the targeted mass range of 5-50 GeV. The details of each sample are summarised in Table 3.1. The signal process is generated using MadGraph 5 [96] and is subsequently processed with Pythia 8 [97] to simulate the parton showering. The model [98] used to generate the signal is the $\nu$MSM, introduced in Section 1.5.2.
Several cuts are applied at generator level to ensure that the two muons are within the acceptance of the detector for each event. The efficiency of this selection is referred to as generator level efficiency $\varepsilon_{gen}$. No requirement is applied at the MadGraph level, to simplify the generator level efficiency estimation. Generator level cuts are applied within the Gauss [99] framework. At least one muon coming from the decay of a

**Table 3.1:** Simulated prompt signal samples: the number of events and generator level efficiency per sample are indicated.

| Mass [ GeV] | Events | | $\varepsilon_{gen}$ | |
|---|---|---|---|---|
| | MU | MD | MU | MD |
| 5 | 49799 | 49800 | 0.2193 ± 0.0009 | 0.2177 ± 0.0009 |
| 10 | 49799 | 49600 | 0.2033 ± 0.0008 | 0.2023 ± 0.0008 |
| 15 | 50000 | 49900 | 0.1945 ± 0.0008 | 0.1941 ± 0.0008 |
| 20 | 49800 | 50000 | 0.1856 ± 0.0008 | 0.1878 ± 0.0008 |
| 30 | 50000 | 49799 | 0.1766 ± 0.0007 | 0.1756 ± 0.0007 |
| 50 | 50000 | 50000 | 0.1575 ± 0.0006 | 0.1568 ± 0.0006 |

$W$ boson with momentum $p > 2\,\text{GeV}$ must be within the acceptance of the detector, requiring $\theta < 300\,\text{mrad}$. Another muon with the same requirements on momentum and acceptance must be in the decay of a heavy neutrino, where the heavy neutrino comes from the decay of the $W$. The generator level efficiency decreases with increasing heavy neutrino mass due to the detector acceptance.

### 3.2.3 Simulated Backgrounds

The main backgrounds for the $W \to \mu\mu$jet decay are described in detail in Section 3.4. A summary of the samples used and their size is given in Table 3.2. All the simulated samples for the backgrounds are generated using Pythia 8.

**Table 3.2:** Simulated background samples: the number of events and generator level efficiency per sample are indicated.

| Decay mode | Events | | $\varepsilon_{gen}$ [%] | |
|---|---|---|---|---|
| | MU | MD | MU | MD |
| $W \to \mu\nu$ | 2492826 | 2495730 | 24.725 ± 0.030 | 24.715 ± 0.030 |
| $W \to \tau\nu$ | 1068652 | 1072163 | 13.474 ± 0.019 | 13.445 ± 0.019 |
| $Z \to \mu^+\mu^-$ | 1994991 | 2035996 | 36.906 ± 0.042 | 36.834 ± 0.042 |
| $Z \to \tau^+\tau^-$ | 2015495 | 2031495 | 36.270 ± 0.300 | 36.867 ± 0.300 |
| $b\bar{b}$ | 514348 | 512052 | 0.090 ± 0.0002 | 0.090 ± 0.0002 |
| $c\bar{c}$ | 511292 | 575029 | 0.017 ± 0.0001 | 0.017 ± 0.0001 |

An additional sample, used for efficiencies determination, is produced for the decay $\Upsilon \to \mu\mu$. The decay is generated using Pythia 8.

The following generator level cuts are applied:

- $W \to \mu\nu$, $W \to \tau\nu$: one lepton with $p_{\text{T}} > 10\,\text{GeV}$ is required to be in the acceptance of the detector ($\theta < 400\,\text{mrad}$) and to be produced by a $W$ or $Z/\gamma^*$ boson;

- $Z \to \mu\mu$, $Z \to \tau\tau$: one lepton with $p_{\text{T}} > 4\,\text{GeV}$ is required to be in the acceptance of the detector ($\theta < 400\,\text{mrad}$) and to be produced by a $W$ or $Z/\gamma^*$ boson. Additionally, the invariant mass of boson must be $> 40\,\text{GeV}$ for the $Z \to \mu\mu$ sample;

- $b\bar{b}$, $c\bar{c}$: one muon with $p_{\text{T}} > 18\,\text{GeV}$ is required to be in the acceptance of the detector ($\theta < 400\,\text{mrad}$);

## 3.3 Signal selection

In the following sections, the reconstruction steps are outlined for the signal decay $W \to \mu\mu$jet with both opposite-sign and same-sign muons as well as the normalisation channel $W \to \mu\nu$. The selection requirements for the decay channels $Z \to \mu\mu$, $\Upsilon \to \mu\mu$, which are used for the efficiencies evaluation, are summarised in Section 3.5.

The $W \to \mu\nu$ channel is used as normalisation channel. In order for systematic uncertainties in the selection of the prompt muon to cancel, the selection requirements for the prompt muon in the signal channel and the normalisation are chosen to be identical. The selection is split into a pre-selection and a final selection based on trained multivariate classifiers, in particular boosted-decision trees (BDTs). Three different classifiers are trained and are used to define three regions in the phase space: the signal region and two orthogonal control regions enhanced in different backgrounds.
The preselection is presented first, then the optimisation of the classifiers is discussed and finally the cuts that identify signal candidates are outlined.

### 3.3.1 Preselection

The preselection is performed in two steps. First a high $p_{\mathrm{T}}$ prompt muon is selected, identical for signal and normalisation channels. Subsequently events with a $W \to \mu\mu$jet candidate are selected.

#### Selection of the prompt muon from $W$

This analysis uses the `WMuLine` stripping line, which requires one muon from `StdAll-LooseMuons` with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. Events are retained if this muon passes all the three (L0, HLT1, HLT2) high-$p_{\mathrm{T}}$ single muon trigger lines, consecutively listed below:

- `L0Muon`: it accepts events with less than 600 hits in the scintillating pad detector (SPD) and with a muon of $p_{\mathrm{T}} > 1.5\,\mathrm{GeV}$.
- `Hlt1SingleMuonHighPt`: it accepts events that fired `L0Muon` line and have a muon with $p_{\mathrm{T}} > 4.8\,\mathrm{GeV}$, $p > 8\,\mathrm{GeV}$ and track $\chi^2/\mathrm{ndf} < 4$.
- `Hlt2SingleMuonHighPt`: it accepts events that have a muon with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$.

Furthermore, a set of additional cuts is applied:

- Since after the muon system, the energy deposits are an important factor in reducing the background, the muon is required to be within the acceptance of both calorimeters;

- To reject pions and kaons punching through the calorimeters and misidentified as muons, a cut is applied on the relative energy deposition in the calorimeters in front of the muon track, $(E_{\mathrm{ECAL}} + E_{\mathrm{HCAL}})/p < 4\%$;

- To remove high-$p_{\mathrm{T}}$ tracks with a poorly determined momentum, the significance of the track curvature of the muon from the W

$$\frac{|q/p|}{\sigma(q/p)}, \tag{3.1}$$

  is required to be greater than 10, where $q$ is the charge, $p$ the momentum of the track and $\sigma$ is the uncertainty on their ratio. This requirement removes almost 25% bad tracks in data;

- To remove spurious contributions from fake tracks, which leads to odd peaks in the $\eta\phi$ distribution of selected candidates, the high $p_{\mathrm{T}}$ muon is required to have hits in the TT detector before the magnet and a track $\chi^2$ probability larger than 0.01.

Muon candidates that satisfy these requirements are called $\mu_W$ in the following.

### $W \to \mu_W N$ candidate reconstruction

Once events with a high-$p_{\mathrm{T}}$ muon have been selected, the reconstruction of the $W \to \mu_W \mu_N$jet candidates proceeds through the following steps. First, cuts which constitute the preselection of the second muon in the event, labelled as $\mu_N$ in the following, are applied. The second muon is selected among the `StdAllLooseMuons` and is required to have $p_{\mathrm{T}} > 3\,\mathrm{GeV}$, $|q/p|/\sigma(q/p) > 10$ and be within the acceptance of both calorimeters. The invariant mass of this muon with a different muon that passes the $\mu_W$ selection is required to be in the range $20-70\,\mathrm{GeV}$. At this stage events are classified in two categories depending on whether the muons are same-sign (SS) or opposite sign (OS).
No preselection cut is applied to the IP of $\mu_N$, since it can be used as an handle to separate prompt and long-lived neutrinos.

Then jets are reconstructed using the anti-$k_t$ [91] algorithm implemented in the FastJet [100] package. To prevent overlap between the jets and the muons, `StdAllLooseMuons` with a $p_{\mathrm{T}}$ larger than $2\,\mathrm{GeV}$ are removed from the particle flow inputs. As this was shown to improve the resolution, neutral recovery contributions are also excluded from the jet search. A cone size of 0.5, the default in LHCb, is adopted [1].
As can be seen from Figure 3.1a, the resolution of the reconstructed mass of the heavy

[1] Other cone sizes, 0.3 and 0.7, have been investigated: a cone size of 0.3 leads to a narrow mass resolution for heavy neutrino masses smaller than 15 GeV, while R = 0.5 or R = 0.7 are more appropriate for higher masses. Eventually, the cone size of 0.5 has been adopted for all mass points because the jet energy correction is not currently available for a 0.3 radius.

neutrino is highly correlated to the resolution of the reconstructed mass of the $W$. This is due to the fact that they are both dominated by the uncertainty on the jet energy. Hence, it has been decided to exploit the known mass of the $W$ to constraint the four-momentum of the $\mu_W\mu_N$jet system. A scale factor $\alpha$ for the jet four-momentum is introduced such that the total four-momentum of the reconstructed $W$ results in:

$$p_W \;=\; p_{\mu_W} \;+\; p_{\mu_N} \;+\; \alpha\, p_{\text{jet}}. \tag{3.2}$$

Requiring that $|p_W| = m_W$, leads to the following expression for $\alpha$:

$$\alpha \;=\; \frac{\sqrt{(m_W^2 - m_{\mu_W\mu_N}^2)\, m_{\text{jet}}^2 + (p_{\mu_W\mu_N} \cdot p_{\text{jet}})^2} \;-\; (p_{\mu_W\mu_N} \cdot p_{\text{jet}})}{m_{\text{jet}}^2}, \tag{3.3}$$

where

- $m_W = 80.385$ GeV [32] is the known $W$ mass;
- $p_{\mu_W\mu_N}$ is the four-vector of the dimuon system;
- $m_{\mu_W\mu_N} \equiv |p_{\mu_W\mu_N}|$ is the mass of the dimuon system;
- $p_{\text{jet}}$ is the four-momentum of the jet;
- $m_{\text{jet}} \equiv |p_{\text{jet}}|$ is the mass of the jet.

Once the scale factor $\alpha$ is applied to the jet four-momentum, the heavy neutrino mass $m(\mu_N\text{jet})$ is recomputed. Figure 3.1b illustrates its effect on the correlation between the reconstructed neutrino mass and the reconstructed $W$ mass. Furthermore, as can be seen from Figure 3.1c where the neutrino mass distribution before and after the $W$ mass constraint are compared, the resolution improves significantly.

Selected jets must have $p_{\text{T}} > 10\,\text{GeV}$ and at least a long track with $p_{\text{T}} > 1.2\,\text{GeV}$. Further cuts are applied on the properties of the reconstructed jet in order to remove poorly reconstructed jets and fake jets. It is therefore required that CPF$> 0.1$ and MPT$> 1.2\,\text{GeV}$, where:

- CPF (charge $p_{\text{T}}$ fraction) is the ratio of the $p_{\text{T}}$ contribution from charged particles and the total $p_{\text{T}}$ of the jet.
- MPT is the minimal value required for the $p_{\text{T}}$ of the maximum $p_{\text{T}}$ particle in the jet. A requirement on the MPT rejects jets composed of soft particles and neutrals.

These cuts have been optimised using $Z \to \mu\mu + \text{jet}$ events [101].

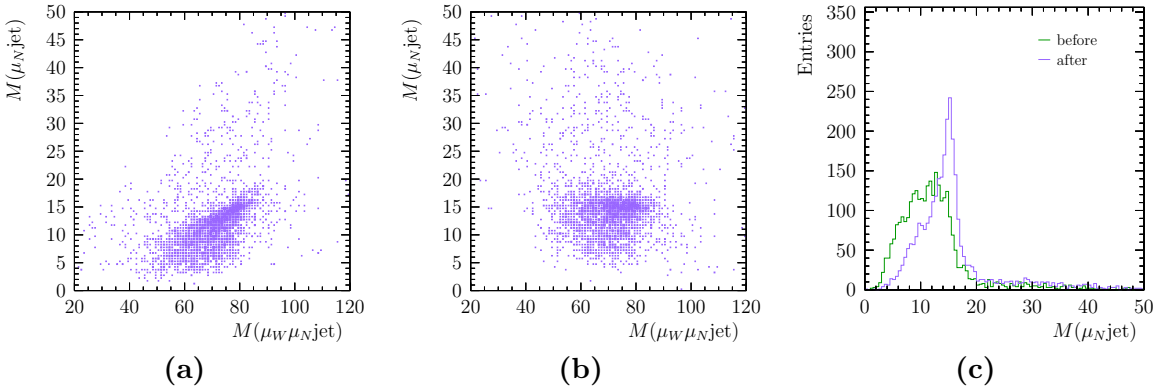



**Figure 3.1:** Effect of the $W$ mass constraint on simulated events with a neutrino mass of 15 GeV. The correlation between the $\mu_N$jet mass and the $\mu_W\mu_N$jet mass is shown in (a) before and in (b) after the $W$ mass constraint is applied to the $\mu_N$jet mass. In (c) the $\mu_N$jet mass distribution before and after the $W$ mass constraint are compared.

Jets are combined with an $\mu_N$ candidate to form $N \to \mu_N$jet candidates, which are required to have a mass smaller than 80 GeV and a $p_{\mathrm{T}} > 10$ GeV. The selected $N$ candidate is combined with a $\mu_W$ candidate to form $W \to \mu_W\mu_N$jet candidates.
The assignment of the two muons is ambiguous if they both pass the $\mu_W$ selection. As shown in Figure 3.2, for every $W$ candidate the jet can be combined with either of the muons giving two combinations for the neutrino. In order to prevent double counting an additional selection is applied, requiring that the mass of the $\mu_N$jet combination is smaller than that of the $\mu_W$jet combination. In this way the ambiguity is removed, as can be seen in Figure 3.2.
Finally, only the candidates within a 20 GeV window from the known $W$ mass value [32] are retained. A summary of the preselection criteria is outlined in Table 3.3.

### 3.3.2 Multivariate Selection

In this analysis three different multivariate classifiers have been trained. Two classifiers are trained for the muon identification. The third classifier uses event kinematics to distinguish the signal from remaining background. Among the many different MVA methods available, the boosted decision tree (BDT) has been adopted in this analysis.

#### Boosted decision trees

A decision tree [102] is a binary tree structural classifier: starting from the root node, a sequence of binary splits is applied. Every split is performed cutting on the discriminating variable which has the highest separating power between signal and background;

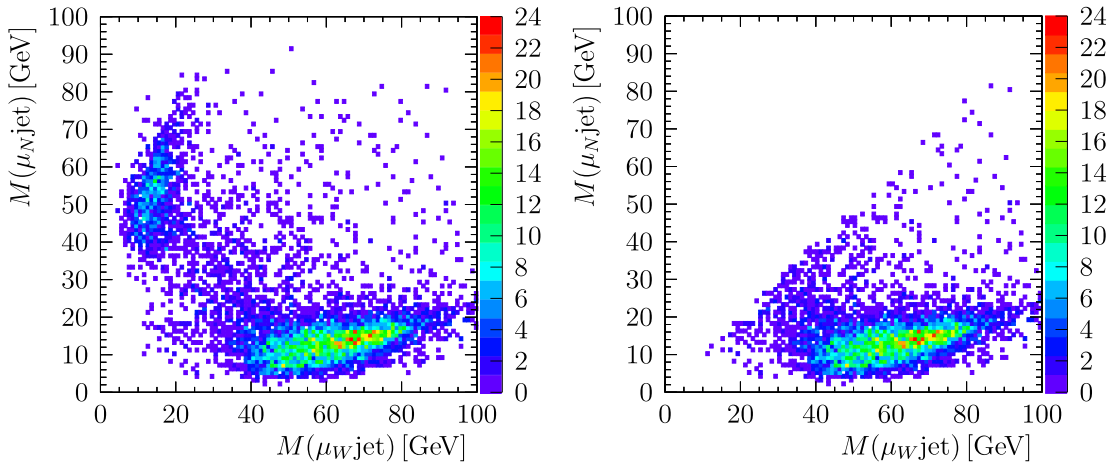



**Figure 3.2:** Invariant mass of $\mu_N$jet as a function of invariant mass of the $\mu_W$jet for candidate events (a) before and (b) after selecting the lightest combination, in the signal sample with a 15 GeV neutrino.

**Table 3.3:** Preselection requirements.

| | |
|---|---|
| $\mu_W$ | `WMuLine` (`IsMuon` and $p_{\mathrm{T}} > 20\,\mathrm{GeV}$)<br>$p_{\mathrm{T}} < 70\,\mathrm{GeV}$<br>in acceptance ECAL and HCAL<br>$(E_{\mathrm{ECAL}} + E_{\mathrm{HCAL}})/p < 4\%$<br>$(q/p)/\sigma(q/p) > 10$<br>num TT hits $> 0$ and $\mathrm{Prob}(\chi^2_{\mathrm{trk}}) > 0.01$<br>$2 < \eta < 4.5$ |
| $\mu_N$ | `IsMuon` and $p_{\mathrm{T}} > 3\,\mathrm{GeV}$<br>in acceptance ECAL and HCAL<br>$(q/p)/\sigma(q/p) > 10$<br>$2 < \eta < 4.5$ |
| jet | $p_{\mathrm{T}} > 10\,\mathrm{GeV}$<br>one long track with $p_{\mathrm{T}} > 1.2\,\mathrm{GeV}$<br>CPF$> 0.1$ and MPT$> 1.2\,\mathrm{GeV}$ |
| $N \to \mu_N$jet | $m_N < 80$ GeV |
| $W \to \mu_W \mu_N$jet | $60 < m_W < 100$ GeV<br>$20 < m_{\mu\mu} < 70\,\mathrm{GeV}$ |

thus, the same variable can be used in several nodes, while it can be that other variables are not used at all. In this way, the feature space is split into many regions and every region can be classified either as signal-like or background-like region, depending on the nature of the major number of events in the final leaf node, as it is called a terminal

node. A drawback of the method is the instability with respect to statistical fluctuation: if two variables have similar discriminating power at a split node, a statistical fluctuation can lead to the choice of one variable instead of the other one, giving rise to an altered classifier response. To overcome this problem, the decision tree is boosted. The boosting consists in generating a forest of decision trees instead of a single tree, and classifying an event according to the major number of times it has been recognised as signal-like or background-like in every tree. This process allows to stabilise the classifier response and improve the separation performance. The trees are trained sequentially on the same sample and the overall performance is improved at each step based on the performance of the previous tree. This is realised by assigning an higher weight to the events misclassified in the previous tree. Eventually the response of all the trees is combined in a single classifier, known as Boosted Decision Tree (BDT).
Since the heavy neutrino has unknown mass, it would be ideal to have a classifier that is uniformly efficient with respect to the mass of the neutrino. This can be achieved in different ways. Mass dependent variables can be excluded from the training features, but this is undesirable if they have high separation power. The other possibility, adopted in this analysis, is to use the uBoost method [103], which consists in training boosted decision trees keeping the signal response efficiency flat in some multivariate space. In uBoost not only the misclassification rate, but also the uniformity of the signal efficiency contribute to the weights.
The three classifiers are referred to as $\mu_W$ uBDT, $\mu_N$ uBDT and kinematic uBDT. All three of them are boosted decision trees, trained requiring the signal efficiency to be flat against the true neutrino mass. The $\mu_W$ uBDT does not need to be unbiased with respect to the neutrino mass, but, since the performances of the uBDT classifier are comparable to other classifiers trained the uBDT has been adopted in order to have the same type of classifiers for all three BDTs.

### Input variables

The input variables for the three BDTs are listed in Table 3.4. The training of the muon identification BDTs is based on particle identification and isolation observables:

- The variables $E_{\mathrm{ECAL}}/p$ and $E_{\mathrm{HCAL}}/p$ represent the relative energy deposited in the electronic (ECAL) and hadronic (HCAL) calorimeters with respect to the momentum of the track.

- The particle identification variable PID$\mu$ [87] provides a measure of how likely is the muon mass hypothesis relative to the pion hypothesis:

$$\Delta \log \mathcal{L}_{\mu\pi} = \log \mathcal{L}(\mu) - \log \mathcal{L}(\pi), \tag{3.4}$$

  as described in Section 2.5.3.

- The variables named IT5, IPF5, IPF3 and IPF1 are isolation variables, which give the fraction of $p_{\mathrm{T}}$ carried by the seed particle with respect to the $p_{\mathrm{T}}$ of all the neutral and charged particles in a cone of a certain size around the particle itself. The cone size is specified in the variable name: 1, 3, 5 stand for 0.1, 0.3, 0.5 respectively. IT indicates that the variable is obtained via the Isolation Tool, which uses as input long tracks (charged particles) and photons (neutral particles), while the IPF variables use as input the Particle Flow, described in Section 2.5.4. Both IT and IPF variables are exploited, since the distributions carry different information, due to the different inputs. Specifically, particle flow isolation variables of different cone sizes are useful for the training of the $\mu_N$ uBDT classifier, while they do not add any more information than IT5 for $\mu_W$ uBDT, as can be seen comparing Figure 3.3 and Figure 3.4.

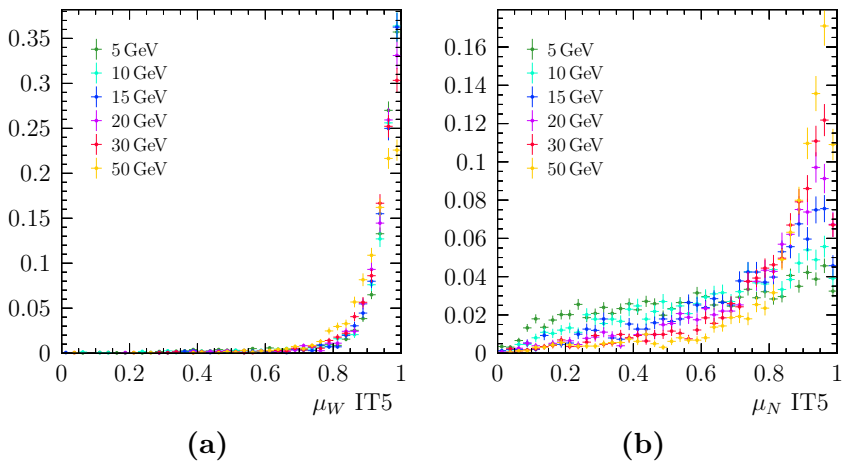



**Figure 3.3:** Isolation tool variables with cone size 0.5 for (a) $\mu_W$ and (b) $\mu_N$.

The kinematics of the candidate is exploited to train the third uBDT. In addition to the dimuon mass $m\mu_W\mu_N$, the $\mu_W\mu_N$jet invariant mass and the jet transverse momentum, the following variables are used:

- $\Delta R$, which is the radial distance between the $\mu_N$ and jet, defined as:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}, \tag{3.5}$$

  where with $\eta$ and $\phi$ the pseudorapidity and azimuthal angle, respectively.

- $\cos\theta$, where the angle $\theta$ is between the two muons in the rest frame of the heavy neutrino $N$.

- Missing $p_{\mathrm{T}}$, which consists of the transverse component of the sum of the four-momentum of all particles used as particle flow input. Since in the lab frame the

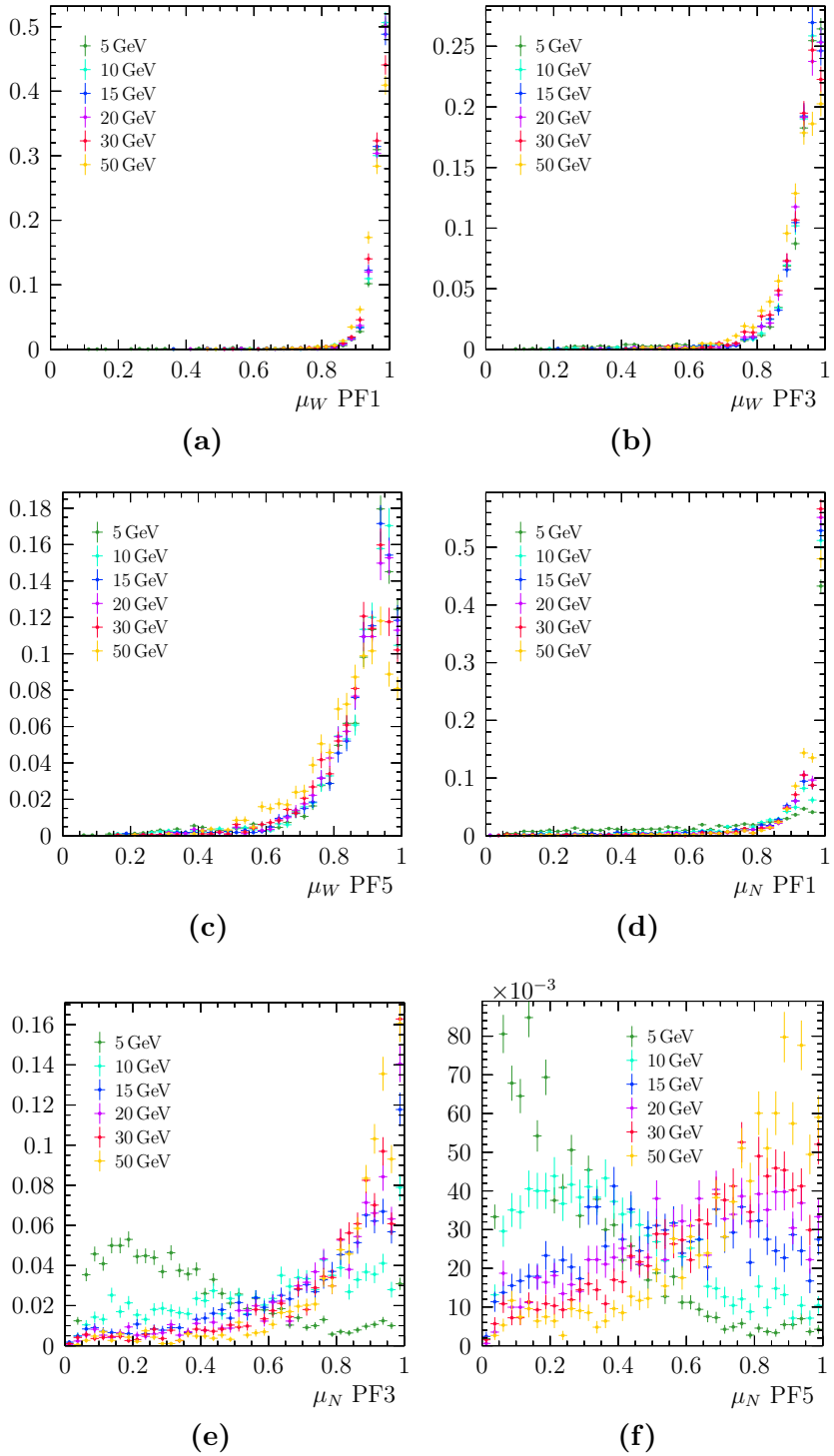



**Figure 3.4:** Particle flow isolation variables for $\mu_W$ with cone size (a) 0.1, (b) 0.3, (c) 0.5 and for $\mu_N$ with cone size (d) 0.1,(e) 0.3, (f) 0.5 for different mass points.

total transverse momentum is expected to be zero, any missing particle would lead to an increase in total transverse momentum, which is therefore denoted as missing $p_{\mathrm{T}}$.

The agreement between data and MC has been checked on the $Z \to \mu\mu$ sample for the input variables of the $\mu_W$ uBDT classifier. The PID$\mu(\mu_W)$ and IT5$(\mu_W)$ variables have been independently re-weighted (since they are uncorrelated) prior to performing the training of the classifiers.

**Table 3.4:** Input variables for the uBDTs.

| $\mu_W$ uBDT | $\mu_N$ uBDT | kinematics uBDT |
|---|---|---|
| IT5$(\mu_W)$ | $p_{\mathrm{T}}$ $(\mu_N)$ | $\cos\theta$ |
| $E_{\mathrm{ECAL}}/p(\mu_W)$ | $E_{\mathrm{ECAL}}/p(\mu_N)$ | $\Delta R$ |
| $E_{\mathrm{HCAL}}/p(\mu_W)$ | $E_{\mathrm{HCAL}}/p(\mu_N)$ | $p_{\mathrm{T}}$(jet) |
| PID$\mu(\mu_W)$ | PID$\mu(\mu_N)$ | $m_{\mu\mu}$ |
| | IPF5$(\mu_N)$ | $m(\mu_W\mu_N$jet) |
| | IPF3$(\mu_N)$ | missing $p_{\mathrm{T}}$ |
| | IPF1$(\mu_N)$ | |

### Training samples

Two statistically independent samples are needed, one to train and one to test the classifier. To exploit the full samples in the training of the classifiers, the $k$-fold cross-validation method [104] is adopted. It consists in randomly splitting the sample in $k$ folds (in this analysis $k = 5$), of which ($k$-1) folds form the training set and the remaining one is used for the validation and application of the classifier. The signal sample used for the training is a mix of the six same-sign muons signal MC samples. The $\mu_W$ uBDT is trained using as background sample a control region of the data (defined by the impact parameter IP$(\mu_W) > 40\,\mu$m and IP$(\mu_N) > 100\,\mu$m). As can be seen from Figure 3.5a, where the data in the control region are overlaid with the expected backgrounds, the background consists mainly of heavy flavour decays $b\bar{b} \to X\mu$. The $\mu_N$ and kinematics uBDTs are trained using as background sample a mix of the six expected MC backgrounds, where each background is weighted such that the ratio of the backgrounds corresponds to the expected one (Figure 3.5b). The background estimation procedure is described in Section 3.4.1.

### MVA output

A few tests are performed on the classifier output to evaluate its performance:

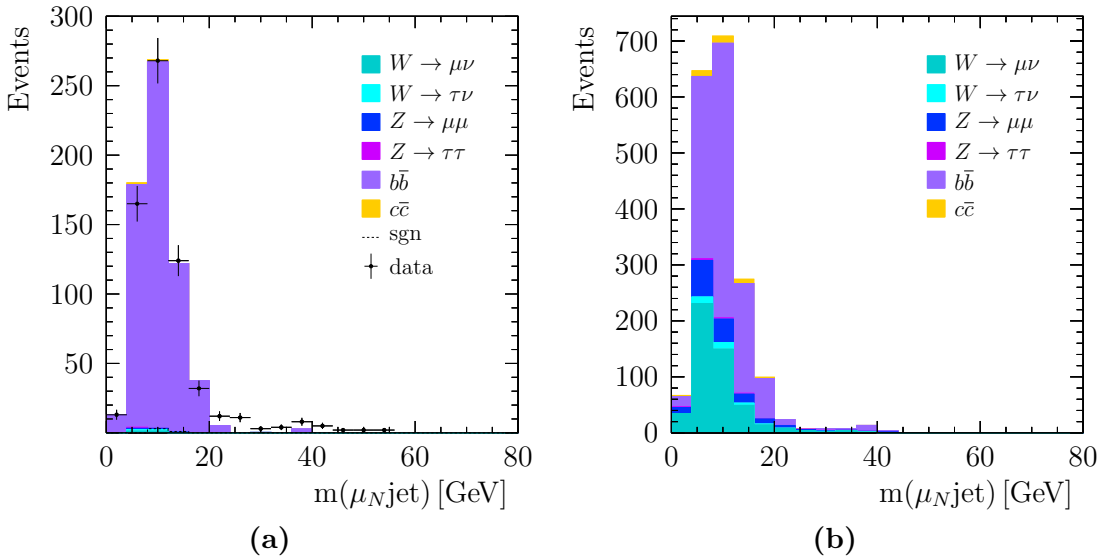



**Figure 3.5:** Background distribution used for the training: (a) data in the control region compared to the expected backgrounds for $\mu_W$ uBDT training and (b) MC background composition used for $\mu_N$ and kinematic uBDT training.

- Overtraining check.
  Overtraining occurs when the feature space has been split too many times with respect to the amount of events and the classifier is subject to statistical fluctuations, such that when the method is tested on a statistically independent sample it gives rise to different results. Hence, to check that no overtraining has occurred, distributions of training and test samples are compared for both signal and background, as shown in Figure 3.6. It can be seen that the BDT response gives a good separation between signal and background and no significant overtraining is observed for any of the classifiers.

- ROC curve.
  The performance of the classifier at various threshold settings can be visualised with the receiver operating characteristic (ROC) curve, which shows background rejection as a function of signal efficiency. The bigger the area under the curve, the more optimal the classifier trained. The roc curves for the three classifiers are shown in Figure 3.7.

- Importance of variables.
  The importance of the variables used in the training of the classifiers has been evaluated by removing one by one each of the training variables and comparing the area under the ROC ($AUC$) curves. A smaller area is an indication of the higher impact of that variable in the training. The variable importance is reported in Figure 3.8 as $1 - AUC$ for each of the training variables. As a reference, also

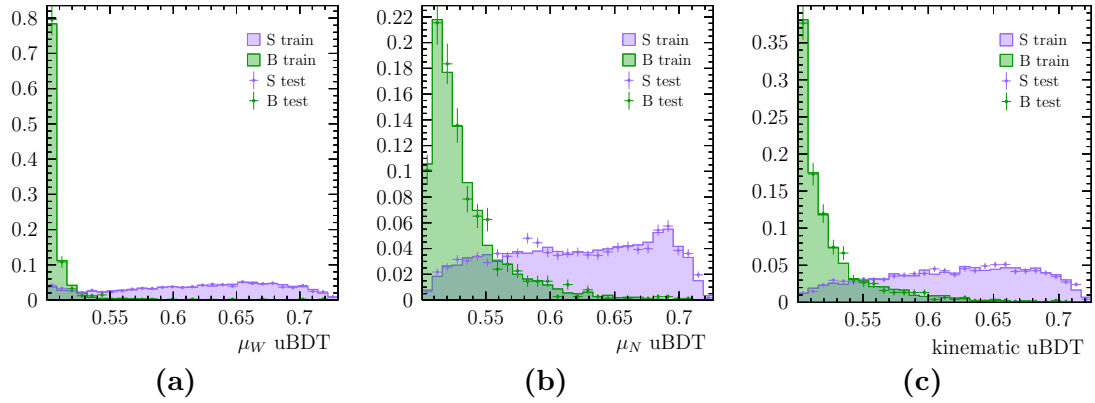



**Figure 3.6:** BDT distribution for signal (S) and background (B) in training and test samples (overtraining check) for (a) $\mu_W$ uBDT , (b) $\mu_N$ and (c) kinematic uBDT.

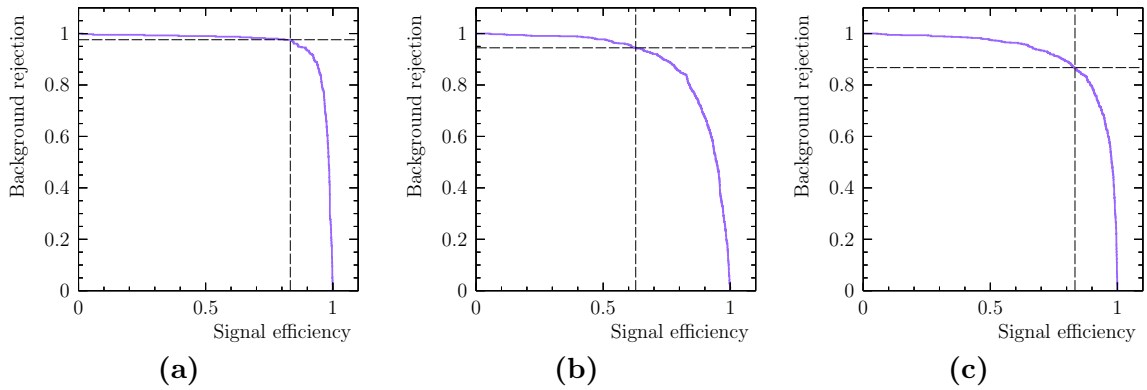



**Figure 3.7:** ROC for (a) $\mu_W$ uBDT, (b) $\mu_N$ uBDT and (c) kinematic uBDT. The signal efficiency is evaluated on the combination of the six mass samples. The operation point for the 15 GeV heavy neutrino mass sample is indicated by dashed lines.

the case where all the variables are used in the training is included and labelled as *all*. It illustrates that for the $\mu_W$ uBDT the isolation variable dominates the training of the classifier. For the other two classifiers there is not such a difference between the input variables.

- Folds validation.
  The ROC curves of different folds are compatible with each other. The overtraining check has been performed also per fold separately showing no significant mismatch between test and training samples.

- Comparison same-sign (SS) and opposite-sign (OS).
  The distributions for OS and SS signal events are compared for each of the three classifiers in Figure 3.9. Although there are small differences, the distributions

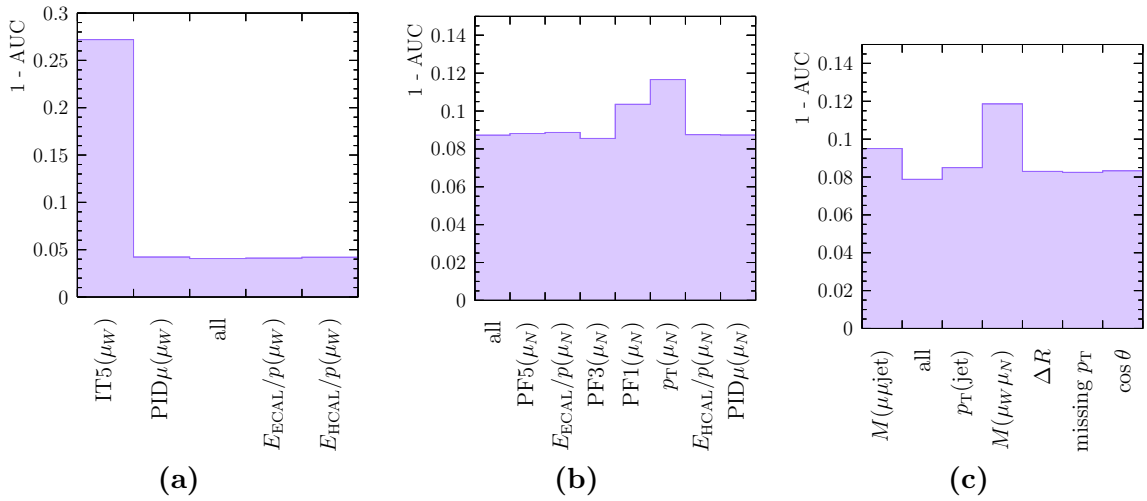



**Figure 3.8:** Variable importance, expressed as (1 - AUC) for (a) the $\mu_W$ uBDT, (b) the $\mu_N$ uBDT and (c) the kinematics uBDT.

for same-sign and opposite-sign muons follow the same trend and therefore the classifiers are straightforwardly applied to the opposite-sign muons samples as well.

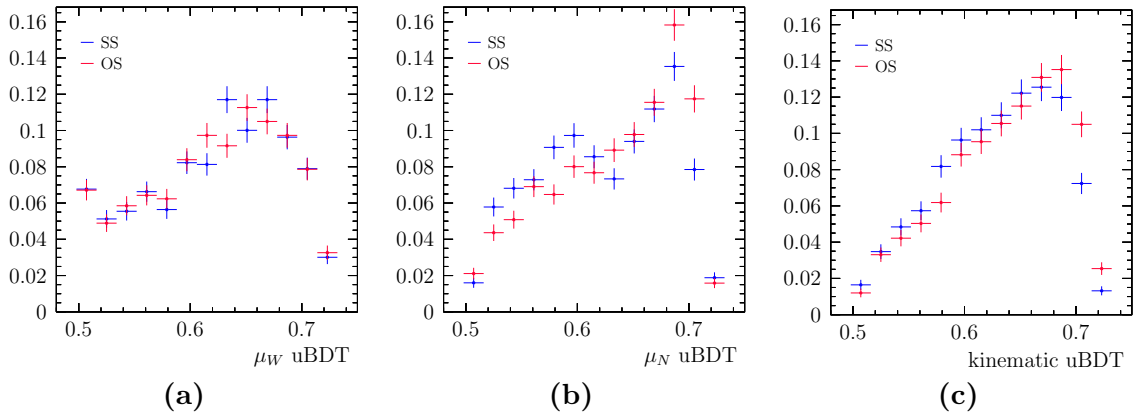



**Figure 3.9:** Distributions of (a) $\mu_W$ uBDT, (b) $\mu_N$ uBDT and (c) kinematic uBDT for same- and opposite-sign muons signal events of 15 GeV mass neutrino. Preselection (Table 3.3) is applied.

The MC background composition per classifier is shown in Figure 3.10. The backgrounds are normalised to the expected yield and compared to the distribution for a simulated 15 GeV neutrino signal. In the case of the $\mu_W$ uBDT classifier the $b\bar{b}$ distribution peaks at low values of the classifier, while the $W \to \mu\nu$ background distribution overlaps with the signal one, as expected. The other two classifiers discriminate between the signal

and all the main backgrounds.

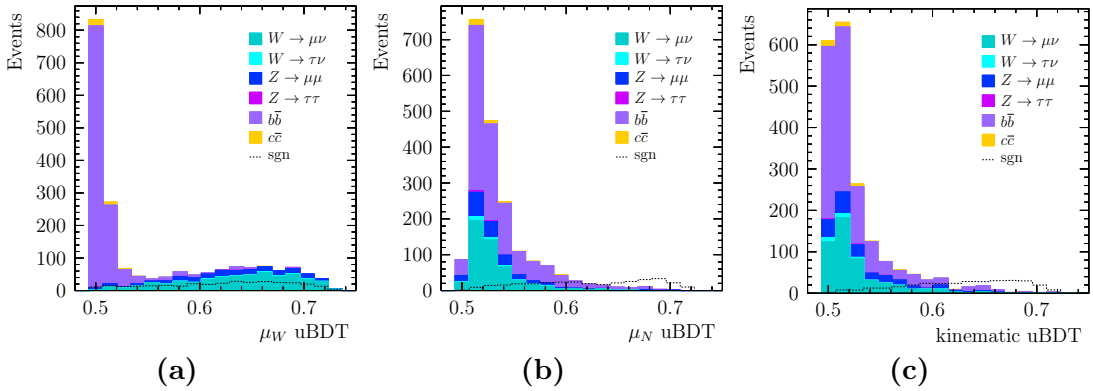



**Figure 3.10:** Background composition (stack) and overlaid signal (dashed line) for (a) $\mu_W$ uBDT, (b) $\mu_N$ uBDT and (c) kinematic uBDT distributions, prior the application of uBDT cuts. The signal is normalised to the expected yield, assuming a coupling $10^3$ times higher than the current limit.

### Cuts optimisation

The optimisation of the cut on each classifier is performed by maximising the Punzi figure-of-merit [105]:

$$\frac{\varepsilon(S)}{\sqrt{B} + \frac{1}{2}\sigma}, \tag{3.6}$$

where $\varepsilon(S)$ is the signal efficiency, $B$ the background yield and $\sigma$ is set to 3. The signal efficiency is computed from the simulated signal samples. The background yield is obtained from the same-sign muons sample:

- for $\mu_W$ uBDT, as the product of the events left in the $b\bar{b}$ control region after the cut on the classifier and a transfer factor, which consists in the fraction of MC events in the signal with respect to the control region, computed from the $b\bar{b}$ MC sample;

- for $\mu_N$ and kinematic uBDTs, by counting the number of MC background events surviving after the cut.

The optimisation is performed independently per mass point. First the Punzi figure of merit is evaluated for the $\mu_W$ uBDT as a function of the cut on the uBDT as shown in Figure 3.11a, where the different colours correspond to different mass points. The optimal working point is quite stable between the different samples. The sharp peak on the right is due to the absence of background events at the corresponding classifier threshold.

After applying the optimised $\mu_W$ uBDT cut, the Punzi figure of merit is evaluated simultaneously for the other two classifiers. An example, using as the heavy neutrino with 15 GeV mass as signal sample, is shown in Figure 3.11b. Since no significant difference in the optimal cut is found for any of the classifiers between the difference mass point, the same cut is adopted for all the signal samples. The efficiency per signal sample for the chosen cuts is reported in Table 3.5. The efficiency for the 5 GeV neutrino mass sample is about 30% lower than for the other mass samples for the $\mu_N$ and kinematic uBDT cuts. This can be attributed to the fact that the mass distribution for the 5 GeV signal sample and backgrounds overlap, therefore the training of the uBDT disfavoured the lowest mass point given that the signal used for the training is a combination of all of simulated signal samples. The efficiency values are also evaluated separately per background component and are are collected in Table 3.6.

**Table 3.5:** Signal efficiency in percent per mass sample for the optimised uBDT cuts. The efficiency of each cut is evaluated over the events surviving the preselection.

| Mass | 5 GeV | 10 GeV | 15 GeV | 20 GeV | 30 GeV | 50 GeV |
|---|---|---|---|---|---|---|
| $\mu_W$ uBDT $> 0.55$ | $84.8 \pm 0.7$ | $86.3 \pm 0.7$ | $83.3 \pm 0.8$ | $83.9 \pm 0.8$ | $80.9 \pm 1.0$ | $75.9 \pm 1.5$ |
| $\mu_N$ uBDT $> 0.60$ | $51.5 \pm 1.0$ | $58.3 \pm 1.0$ | $62.8 \pm 1.0$ | $67.8 \pm 1.1$ | $76.9 \pm 1.1$ | $92.5 \pm 0.9$ |
| kinematic uBDT $> 0.62$ | $55.3 \pm 1.0$ | $56.5 \pm 1.0$ | $58.4 \pm 1.1$ | $60.3 \pm 1.1$ | $66.4 \pm 1.2$ | $68.5 \pm 1.6$ |

**Table 3.6:** Efficiency in percent evaluated separately per background component for the optimised uBDT cuts.

| Background | $W \to \mu\nu$ | $W \to \tau\nu$ | $Z \to \mu\mu$ | $Z \to \tau\tau$ | $b\bar{b}$ | $c\bar{c}$ |
|---|---|---|---|---|---|---|
| $\mu_W$ uBDT $> 0.55$ | $92.6 \pm 1.3$ | $92.0 \pm 2.5$ | $90.5 \pm 1.1$ | $87.5 \pm 4.8$ | $4.9 \pm 1.0$ | $14.1 \pm 3.9$ |
| $\mu_N$ uBDT $> 0.60$ | $2.8 \pm 0.8$ | $1.8 \pm 1.2$ | $4.5 \pm 0.7$ | $4.2 \pm 2.9$ | $5.9 \pm 1.1$ | $6.4 \pm 2.8$ |
| kinematic uBDT $> 0.62$ | $1.4 \pm 0.6$ | $0.0 \pm 0.0$ | $7.2 \pm 0.9$ | $4.2 \pm 2.9$ | $3.1 \pm 0.8$ | $3.8 \pm 2.2$ |

### 3.3.3 Full selection

On top of trigger and preselection, additional requirements are applied to select the signal candidates. These additional requirements, referred to collectively as full selection, are summarised in Table 3.7 and consist of the BDT classifiers cuts described in the previous section and cuts on the IP of both muons to remove background from non-prompt muons.

The $\mu_N$jet mass distributions in the signal region for the six mass points is shown in Figure 3.12. The resolution degrades with increasing mass, since the probability for the heavy neutrino to decay into two jets instead of one increases.

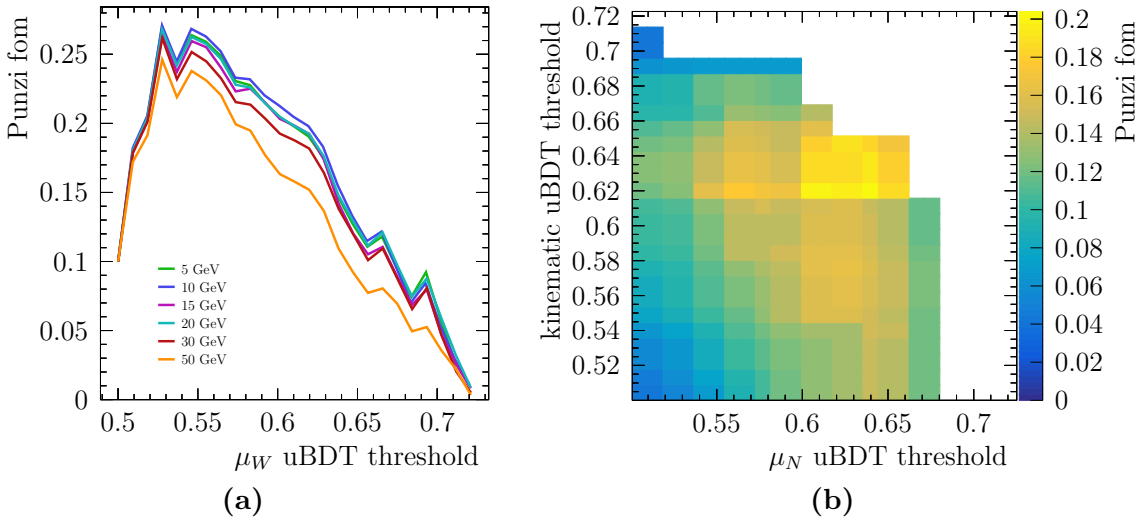



**Figure 3.11:** Punzi figure of merit as a function of (a) the cut on the classifier response for $\mu_W$ uBDT and (b) the simultaneous cuts on $\mu_N$ uBDT and kinematic uBDT for the 15 GeV neutrino mass signal sample.

**Table 3.7:** Full selection requirements.

| |
|---|
| $\mathrm{IP}(\mu_W) < 0.04\,\mathrm{mm}$ |
| $\mu_W$ uBDT $> 0.55$ |
| $\mathrm{IP}(\mu_N) < 0.1\,\mathrm{mm}$ |
| $\mu_N$ uBDT $> 0.60$ |
| kinematics uBDT $> 0.62$ |

### 3.3.4 Signal efficiency

The reconstruction, stripping and selection efficiencies have been evaluated from simulation. Corrections to the reconstruction efficiency of the two muons are described in Section 4.4. A summary of the reconstruction, stripping and selection efficiencies for signal with same-sign and opposite-sign muons samples can be found in Table 3.8 and Table 3.9, respectively, and for normalisation sample in Table 3.10. The generator level cut efficiency differs between normalisation and signal simulated samples. For the signal channel the efficiency is lower due to the requirement of an additional muon in the acceptance, as described in Section 3.2. The relative efficiencies of the normalisation channel with respect to the different signal sample with same- and opposite-sign muons are collected in Table 3.11. It can be seen that the efficiencies for same-sign and opposite-sign muons samples are very similar.

**Table 3.8:** Efficiency for selecting signal events with same-sign muons, evaluated from simulation.

| Efficiency | Signal 5 GeV | Signal 10 GeV | Signal 15 GeV | Signal 20 GeV | Signal 30 GeV | Signal 50 GeV |
|---|---|---|---|---|---|---|
| | | | Selection Sequential | | | |
| generator level | $21.85 \pm 0.06$ | $20.28 \pm 0.06$ | $19.43 \pm 0.06$ | $18.67 \pm 0.05$ | $17.61 \pm 0.05$ | $15.71 \pm 0.05$ |
| $\mu_W$ reco+strip | $47.56 \pm 0.16$ | $47.96 \pm 0.16$ | $48.57 \pm 0.16$ | $48.57 \pm 0.16$ | $46.39 \pm 0.16$ | $44.81 \pm 0.16$ |
| trigger | $69.28 \pm 0.58$ | $68.35 \pm 0.61$ | $68.95 \pm 0.62$ | $69.66 \pm 0.62$ | $70.72 \pm 0.61$ | $71.61 \pm 0.57$ |
| $\mu_W$ presel | $82.88 \pm 0.57$ | $83.06 \pm 0.59$ | $84.41 \pm 0.59$ | $83.55 \pm 0.60$ | $81.15 \pm 0.63$ | $63.34 \pm 0.72$ |
| $\mu_W$ fullsel | $81.16 \pm 0.65$ | $82.20 \pm 0.66$ | $80.38 \pm 0.70$ | $80.65 \pm 0.70$ | $78.56 \pm 0.73$ | $73.78 \pm 0.83$ |
| $\mu$jet reco+strip | $26.31 \pm 0.16$ | $24.71 \pm 0.15$ | $22.59 \pm 0.14$ | $22.71 \pm 0.14$ | $23.92 \pm 0.15$ | $27.58 \pm 0.16$ |
| $\mu$jet presel | $68.90 \pm 0.86$ | $70.52 \pm 0.87$ | $66.85 \pm 0.93$ | $60.68 \pm 0.96$ | $50.60 \pm 1.00$ | $36.03 \pm 1.06$ |
| $\mu$jet fullsel | $36.98 \pm 1.08$ | $42.98 \pm 1.12$ | $47.84 \pm 1.21$ | $50.57 \pm 1.26$ | $58.27 \pm 1.39$ | $65.50 \pm 1.74$ |
| | | | Selection Cumulative | | | |
| generator level | $21.85 \pm 0.06$ | $20.28 \pm 0.06$ | $19.43 \pm 0.06$ | $18.67 \pm 0.05$ | $17.61 \pm 0.05$ | $15.71 \pm 0.05$ |
| $\mu_W$ reco+strip | $10.39 \pm 0.05$ | $9.73 \pm 0.04$ | $9.44 \pm 0.04$ | $9.07 \pm 0.04$ | $8.17 \pm 0.04$ | $7.04 \pm 0.03$ |
| trigger | $7.20 \pm 0.07$ | $6.65 \pm 0.07$ | $6.51 \pm 0.07$ | $6.32 \pm 0.06$ | $5.78 \pm 0.06$ | $5.04 \pm 0.05$ |
| $\mu_W$ presel | $5.97 \pm 0.07$ | $5.52 \pm 0.07$ | $5.49 \pm 0.07$ | $5.28 \pm 0.06$ | $4.69 \pm 0.06$ | $3.19 \pm 0.05$ |
| $\mu_W$ fullsel | $4.84 \pm 0.07$ | $4.54 \pm 0.07$ | $4.42 \pm 0.07$ | $4.26 \pm 0.06$ | $3.68 \pm 0.06$ | $2.36 \pm 0.04$ |
| $\mu$jet reco+strip | $1.27 \pm 0.02$ | $1.12 \pm 0.02$ | $1.00 \pm 0.02$ | $0.97 \pm 0.02$ | $0.88 \pm 0.02$ | $0.65 \pm 0.01$ |
| $\mu$jet presel | $0.88 \pm 0.02$ | $0.79 \pm 0.02$ | $0.67 \pm 0.02$ | $0.59 \pm 0.01$ | $0.45 \pm 0.01$ | $0.23 \pm 0.01$ |
| $\mu$jet fullsel | $0.32 \pm 0.01$ | $0.34 \pm 0.01$ | $0.32 \pm 0.01$ | $0.30 \pm 0.01$ | $0.26 \pm 0.01$ | $0.15 \pm 0.01$ |

**Table 3.9:** Efficiency for selecting signal events with opposite-sign muons, evaluated from simulation.

| Efficiency | Signal 5 GeV | Signal 10 GeV | Signal 15 GeV | Signal 20 GeV | Signal 30 GeV | Signal 50 GeV |
|---|---|---|---|---|---|---|
| | | | Selection Sequential | | | |
| generator level | $21.85 \pm 0.06$ | $20.28 \pm 0.06$ | $19.43 \pm 0.06$ | $18.67 \pm 0.05$ | $17.61 \pm 0.05$ | $15.71 \pm 0.05$ |
| $\mu_W$ reco+strip | $47.56 \pm 0.16$ | $47.96 \pm 0.16$ | $48.57 \pm 0.16$ | $48.57 \pm 0.16$ | $46.39 \pm 0.16$ | $44.81 \pm 0.16$ |
| trigger | $69.80 \pm 0.58$ | $69.20 \pm 0.62$ | $70.00 \pm 0.62$ | $68.07 \pm 0.65$ | $69.80 \pm 0.63$ | $71.76 \pm 0.57$ |
| $\mu_W$ presel | $82.94 \pm 0.57$ | $84.10 \pm 0.59$ | $82.66 \pm 0.62$ | $83.64 \pm 0.62$ | $81.21 \pm 0.64$ | $63.73 \pm 0.72$ |
| $\mu_W$ fullsel | $80.98 \pm 0.65$ | $79.71 \pm 0.70$ | $79.63 \pm 0.72$ | $82.53 \pm 0.70$ | $79.06 \pm 0.74$ | $73.20 \pm 0.83$ |
| $\mu$jet reco+strip | $26.45 \pm 0.16$ | $23.52 \pm 0.15$ | $22.31 \pm 0.14$ | $21.45 \pm 0.14$ | $23.29 \pm 0.15$ | $27.68 \pm 0.16$ |
| $\mu$jet presel | $72.18 \pm 0.83$ | $69.74 \pm 0.90$ | $67.00 \pm 0.94$ | $60.99 \pm 0.99$ | $54.47 \pm 1.01$ | $41.33 \pm 1.08$ |
| $\mu$jet fullsel | $40.42 \pm 1.07$ | $55.90 \pm 1.17$ | $55.30 \pm 1.22$ | $59.60 \pm 1.27$ | $60.49 \pm 1.35$ | $72.73 \pm 1.52$ |
| | | | Selection Cumulative | | | |
| generator level | $21.85 \pm 0.06$ | $20.28 \pm 0.06$ | $19.43 \pm 0.06$ | $18.67 \pm 0.05$ | $17.61 \pm 0.05$ | $15.71 \pm 0.05$ |
| $\mu_W$ reco+strip | $10.39 \pm 0.05$ | $9.73 \pm 0.04$ | $9.44 \pm 0.04$ | $9.07 \pm 0.04$ | $8.17 \pm 0.04$ | $7.04 \pm 0.03$ |
| trigger | $7.25 \pm 0.07$ | $6.73 \pm 0.07$ | $6.61 \pm 0.07$ | $6.17 \pm 0.06$ | $5.70 \pm 0.06$ | $5.05 \pm 0.05$ |
| $\mu_W$ presel | $6.02 \pm 0.07$ | $5.66 \pm 0.07$ | $5.46 \pm 0.07$ | $5.16 \pm 0.07$ | $4.63 \pm 0.06$ | $3.22 \pm 0.05$ |
| $\mu_W$ fullsel | $4.87 \pm 0.07$ | $4.51 \pm 0.07$ | $4.35 \pm 0.07$ | $4.26 \pm 0.07$ | $3.66 \pm 0.06$ | $2.36 \pm 0.04$ |
| $\mu$jet reco+strip | $1.29 \pm 0.02$ | $1.06 \pm 0.02$ | $0.97 \pm 0.02$ | $0.91 \pm 0.02$ | $0.85 \pm 0.02$ | $0.65 \pm 0.01$ |
| $\mu$jet presel | $0.93 \pm 0.02$ | $0.74 \pm 0.02$ | $0.65 \pm 0.02$ | $0.56 \pm 0.01$ | $0.46 \pm 0.01$ | $0.27 \pm 0.01$ |
| $\mu$jet fullsel | $0.38 \pm 0.01$ | $0.41 \pm 0.01$ | $0.36 \pm 0.01$ | $0.33 \pm 0.01$ | $0.28 \pm 0.01$ | $0.20 \pm 0.01$ |

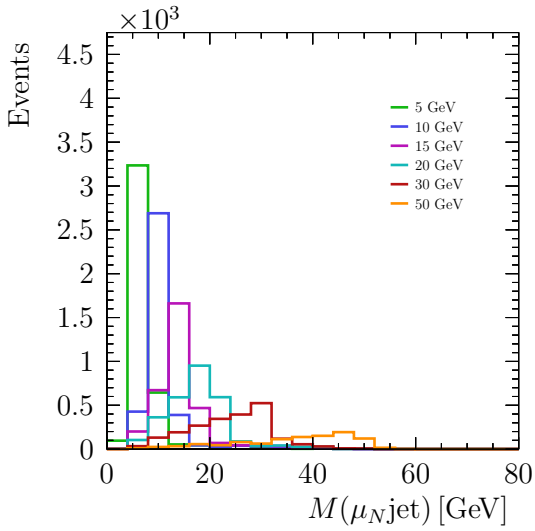



**Figure 3.12:** Comparison of $N$ mass distributions of different signal hypothesis for same-sign muons sample. The signal is normalised to the expected yield, assuming a coupling $10^3$ times higher than the current limit.

**Table 3.10:** Efficiency for selecting $W \to \mu\nu$ events evaluated from simulation.

| Efficiency | Sequential | Cumulative |
|---|---|---|
| generator level | $24.72 \pm 0.03$ | $24.72 \pm 0.03$ |
| reco+strip | $60.78 \pm 0.02$ | $15.03 \pm 0.02$ |
| trigger | $73.90 \pm 0.03$ | $11.10 \pm 0.01$ |
| presel | $82.77 \pm 0.03$ | $9.19 \pm 0.01$ |
| fullsel | $82.76 \pm 0.03$ | $7.61 \pm 0.01$ |

## 3.4 Backgrounds selection

The main backgrounds for this analysis are the following:

**Table 3.11:** Relative efficiency of the normalisation channel with respect to signal. The uncertainty is statistical.

| $\varepsilon_{norm}/\varepsilon_{sgn}$ | same sign | opposite sign |
|---|---|---|
| 5 GeV | $23.4 \pm 0.9$ | $20.2 \pm 0.7$ |
| 10 GeV | $22.4 \pm 0.8$ | $18.4 \pm 0.6$ |
| 15 GeV | $23.8 \pm 0.8$ | $21.2 \pm 0.7$ |
| 20 GeV | $25.6 \pm 0.9$ | $22.9 \pm 0.8$ |
| 30 GeV | $29.3 \pm 1.1$ | $27.1 \pm 1.0$ |
| 50 GeV | $49.6 \pm 2.2$ | $38.8 \pm 1.6$ |

- charged weak currents, in particular $pp \to W + X$ with $W \to \mu\nu$ and $W \to \tau\nu$. From truth level studies on the $W \to \mu\nu$ MC sample, it was found that that the $\mu_W$ is a real muon, while $\mu_N$ is a real muon $\sim 20\%$ of the times, coming mainly from pions ($\sim 48\%$), kaons ($\sim 16\%$) and $D^0$ ($\sim 12\%$). About 60% of the times it is a pion, and about 19% a kaon.

- neutral electroweak Drell-Yan processes, $pp \to \gamma/Z^{(*)} + X$ with $\gamma/Z^{(*)} \to \mu\mu, \tau\tau$. To simulate this background samples are used of DY production with mass $>$ 40 GeV (referred to as $Z \to \mu\mu$ in this analysis) and also with mass $>$ 10 GeV (referred to as low-mass Drell-Yan in this analysis).

- heavy flavour: $b\bar{b} \to X\mu$ and $c\bar{c} \to X\mu$. From truth level studies on the $b\bar{b}$ sample, it was found that the $\mu_W$ is a real muon 97% of the time and the $\mu_N$ 65% of the times. The $\mu_W$ mainly comes from the decay of B mesons, while the $\mu_N$ comes mainly from the decay of D mesons.

The contribution of light QCD processes $(u, d, s) \to X\mu$ is evaluated from data.
In the same sign muons channel the Drell-Yan type backgrounds are highly suppressed; in the opposite sign muon channel the contribution from $Z \to \mu\mu$ and $Z \to \tau\tau$ are reduced by the preselection cut on the invariant mass of the two muons, but the low-mass Drell-Yan remains a prominent background.
Three control regions have been selected, one enhanced in electroweak $W$ backgrounds, indicated as $W$ region, one in heavy flavour backgrounds, indicated as $b\bar{b}$ region, and one in light QCD backgrounds, indicated as QCD region. The cuts that define the control regions, applied on top of trigger and preselection, are summarised in Table 3.12.

**Table 3.12:** Cuts applied to define the three different phase space regions.

| | IP($\mu_W$) [mm] | $\mu_W$ uBDT | $\mu_N$ uBDT | global uBDT | IP($\mu_N$) [mm] |
|---|---|---|---|---|---|
| signal | $< 0.04$ | $> 0.55$ | $> 0.60$ | $> 0.62$ | $< 0.1$ |
| $W$ region | $< 0.04$ | $> 0.55$ | $< 0.60$ | $< 0.62$ | $< 0.1$ |
| $b\bar{b}$ region | $> 0.04$ | $< 0.55$ | $< 0.60$ | $< 0.62$ | $> 0.1$ |
| QCD region | $< 0.04$ | $< 0.55$ | $> 0.60$ | $> 0.62$ | $< 0.1$ |

In each region the predicted background composition and yield are compared to the data to confirm that no other contribution has been neglected.

An additional region, defined by the requirements in Table 3.14, is used to normalise the $Z \to \mu\mu$ background.

### 3.4.1 Yields prediction

The expected yields for the background samples are estimated for each set of cuts according to

$$N_{exp} = \mathcal{L} \cdot \sigma \cdot \mathcal{B} \cdot \varepsilon \ , \tag{3.7}$$

where

- $\mathcal{L}$ is the integrated luminosity;
- $\sigma$ is the theoretical cross section. The theoretical cross section for $Z \to ll$ and $W \to l\nu$ are the cross section at NNLO from [106] scaled by the branching ratio from [32] ;
- $\mathcal{B}$, where applicable, is the branching ratio of the process, taken from [32];
- $\varepsilon$ is the selection efficiency given by the product of the generator cut efficiency, the stripping and reconstruction efficiency and, after selection, the selection efficiency:

$$\varepsilon = \varepsilon_{gen} \cdot \varepsilon_{strip+reco} \cdot \varepsilon_{sel} \tag{3.8}$$

The selection efficiency is evaluated from MC as the ratio of MC events in a given region and the total number of events. The uncertainty on the expected number of events is obtained by propagating the uncertainty of 1.16% on the luminosity determination and the uncertainties on the efficiencies. The uncertainties on branching ratio and cross-section are considered negligible. The values of the parameters used for each background sample are reported in Table 3.13, with the exception of generator level efficiencies, which can be found in Table 3.2. Since no production cross-section measurement is available for the low-mass Drell-Yan sample, the shape is scaled such that the $Z \to \mu\mu$ peak is normalised to the corresponding expected yield.

**Table 3.13:** Cross section, branching ratio and selection efficiencies for the background samples studied. The selection efficiencies are evaluated using the simulated samples.

| Decay Mode | $\sigma$[nb] | $\mathcal{B}$ | $\varepsilon_{strip+reco}$ (MD) [%] | $\varepsilon_{strip+reco}$ (MU) [%] |
|---|---|---|---|---|
| $W \to \mu\nu$ | 112.778 | 0.106 | $0.126 \pm 0.002$ | $0.126 \pm 0.002$ |
| $W \to \tau\nu$ | 112.778 | 0.020 | $0.089 \pm 0.003$ | $0.088 \pm 0.003$ |
| $Z \to \mu\mu$ | 33.573 | 0.034 | $0.255 \pm 0.004$ | $0.237 \pm 0.003$ |
| $Z \to \tau\tau$ | 33.573 | 0.034 | $0.020 \pm 0.001$ | $0.021 \pm 0.001$ |
| $b\bar{b}$ | 1646.890 | 1.000 | $0.898 \pm 0.013$ | $1.040 \pm 0.014$ |
| $c\bar{c}$ | 1754.700 | 1.000 | $0.388 \pm 0.008$ | $0.447 \pm 0.009$ |

### 3.4.2 Background Composition per Region

Figure 3.13 shows the candidate neutrino mass distribution in the control regions for both the OS and SS sample. The data are overlaid with sum of all background contributions, where each background is normalised to the yield expected from the simulation.

As can be observed, the $W \to \mu\nu$ and heavy flavour backgrounds are well isolated in the respective control regions. For both the SS and OS sample the predicted shape of the background distribution is in good agreement with the data. The distributions for other observables are in agreement as well: as an example, the di-muon mass, $\cos\theta$ and $p_{\mathrm{T}}$ distributions for the two muons are shown in Appendix B. Regarding the light QCD region, it can be seen that for the SS muons sample the expected backgrounds account only for part of the events observed in data. Hence, the data events in excess are attributed to light QCD decays. For the OS muons sample the excess of events over the expected backgrounds is not as evident due to the dominant Drell-Yan contribution. The light QCD background in the QCD control region is then obtained by subtracting from the total number of events in the region the estimated yields of the simulated backgrounds. This procedure is described in more detail in Section 4.3.2.

The backgrounds prediction in the signal region is shown in Figure 3.14. Only few backgrounds events are expected for the SS channel, while the Drell-Yan component is significant in the OS channel.

The MC sample for the process $Z \to \mu\mu$ is used as a cross check for the yield estimation method. In Figure 3.15 the $Z$ mass distributions obtained from data and MC samples applying the selection in Table 3.14 are compared. In terms of shape, the peak is larger in data than in simulation due to imperfect detector calibration. The number of estimated and measured events are in good agreement: the $Z$ yield is expected to be $122671 \pm 1442$, while on data it amounts to 124816. Hence the $Z$ expectation is about 2% lower than the data yield.

## 3.5 Selection of other decay modes

#### $W \to \mu\nu$

The $W \to \mu\nu$ decay is the normalisation channel. The selection adopted for the high-$p_{\mathrm{T}}$ muon is the same as for the selection of the signal, such that systematic uncertainties related to the selection of the normalisation channel cancel in the calculation of the branching ratio. The requirements are summarised in Table 3.14. The transverse momentum of the muon, after the application of the selection requirements, is shown in Figure 3.16c.

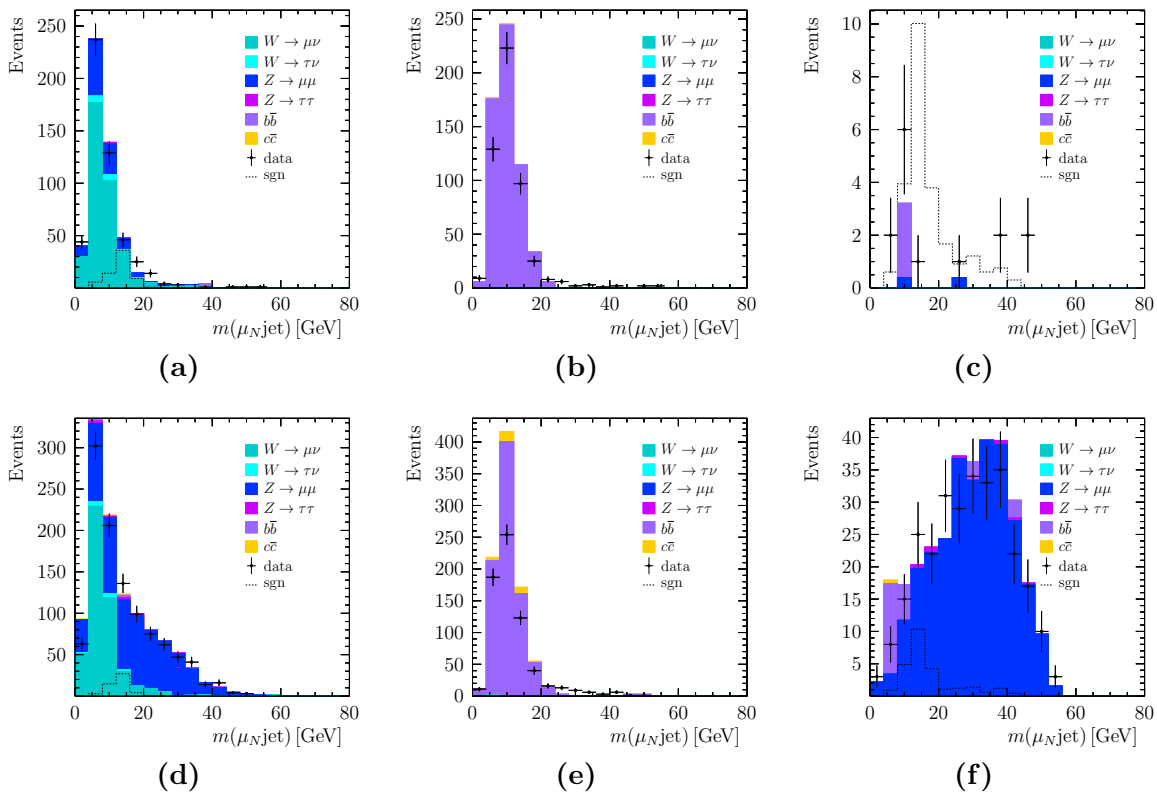



**Figure 3.13:** Background composition predicted by the simulation (stack), signal (dashed line) and data (filled circles) in (a)(c) the $W$ region, (b)(d) $b\bar{b}$ region and (c)(f) light QCD region, on top for same-sign muons channel and at the bottom for opposite-sign muons channel. The signal is normalised to the expected yield, assuming a coupling $10^3$ times higher than the current limit.

### $Z \to \mu\mu$

The $Z \to \mu\mu$ channel is used for efficiencies and systematics evaluation. The Z candidate is reconstructed combining two hard muons, each with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. The Z mass is required to be $\mathrm{m}_Z > 60\,\mathrm{GeV}$. At least one of the muons is required to be TOS with respect to the same lines as the signal channel. Further selection requirements are listed in Table 3.14. The invariant mass distribution of the two muons in data, after the application of the selection requirements, is shown in Figure 3.16a.

### $\Upsilon \to \mu\mu$

The $\Upsilon \to \mu\mu$ is used for the evaluation of efficiency corrections, reported in section 4.4.4 The $\Upsilon$ candidate is reconstructed combining two medium hard muons, each with $p_{\mathrm{T}} > 2\,\mathrm{GeV}$. The $\Upsilon$ mass is required to be in the mass range $9.3 < m_\Upsilon < 9.7\,\mathrm{GeV}$ and the vertex $\chi^2/\mathrm{ndf} < 16$. Further selection requirements are listed in Table 3.14. The

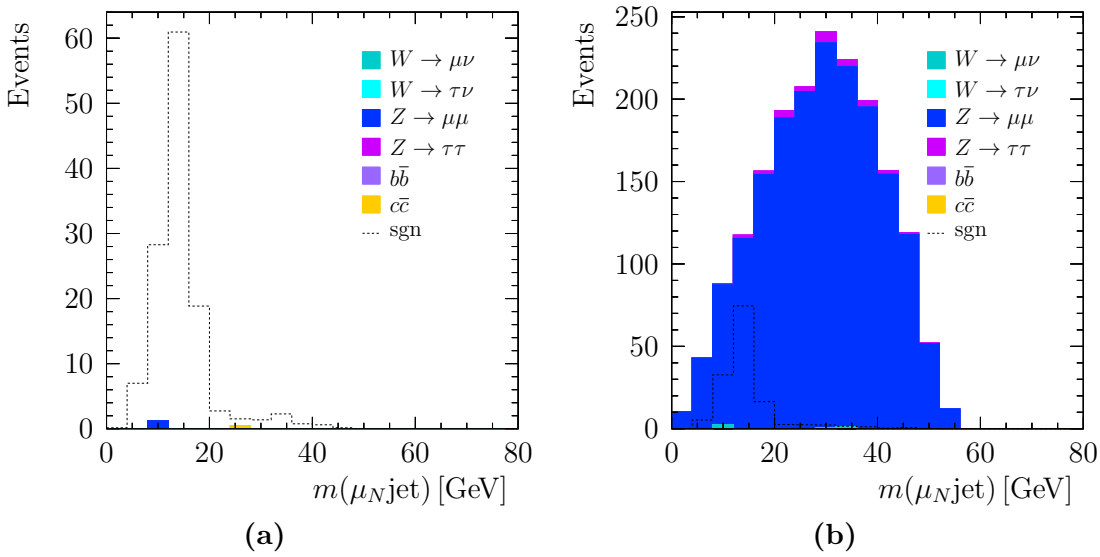



**Figure 3.14:** Predicted background composition in the signal region (stack) for same-sign muons (a) and opposite sign-muons (b). The signal is normalised to the expected yield, assuming a coupling $10^3$ times higher than the current limit.

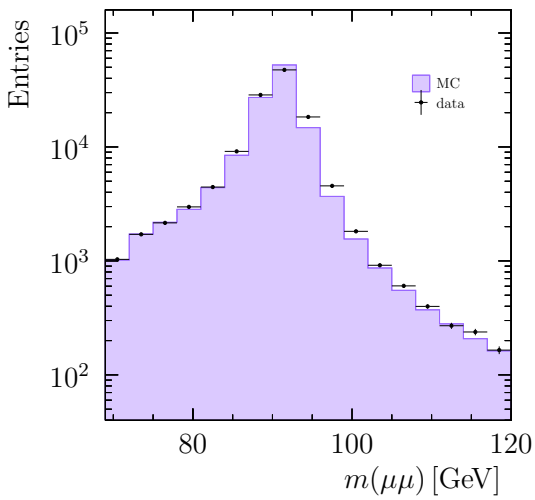
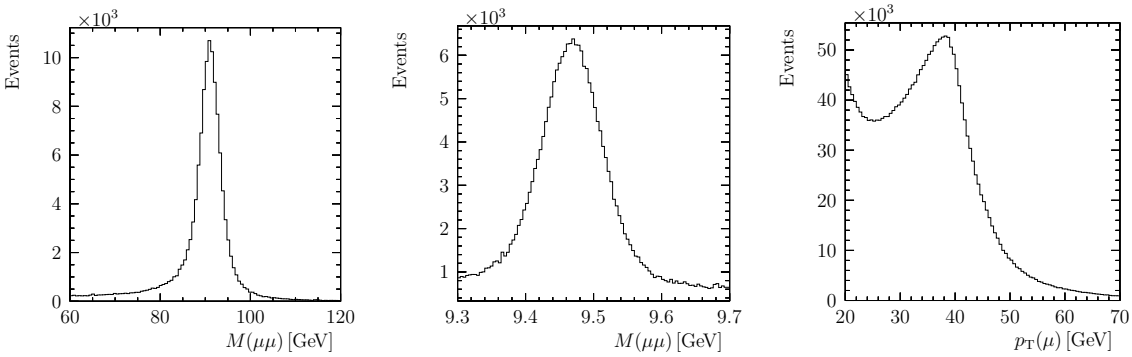



**Figure 3.15:** Comparison of $Z$ mass distribution between data and simulation.

invariant mass distribution of the two muons in data, after the application of the selection requirements, is shown in Figure 3.16b. In data events are taken from the `WMuLine`, which implies effectively that one of the two muons always has $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. For the simulation a standard $\Upsilon \to \mu\mu$ sample is used, that does not have this $p_{\mathrm{T}}$ cut. Since the kinematics of the two samples is rather different, for the evaluation of the muon identification efficiency muon candidates are binned in $p_{\mathrm{T}}$ and $\eta$.

**Table 3.14:** Selection requirements on $W \to \mu\nu$, $Z \to \mu\mu$, $\Upsilon \to \mu\mu$ decay channels.

| Channel | Selection |
|---|---|
| $W \to \mu\nu$ | `WMuLine` (`IsMuon` and $p_{\mathrm{T}} > 20\,\mathrm{GeV}$)<br>$p_{\mathrm{T}} < 70\,\mathrm{GeV}$<br>in acceptance ECAL and HCAL<br>$(E_{\mathrm{ECAL}} + E_{\mathrm{HCAL}})/p < 4\%$<br>$(q/p)/\sigma(q/p) > 10$<br>num TT hits $> 0$ and $\mathrm{Prob}(\chi^2_{\mathrm{trk}}) > 0.01$<br>$\mathrm{IP}_{\mu} < 40\,\mu\mathrm{m}$<br>$2 < \eta < 4.5$<br>$\mu_W$ uBDT $> 0.55$ |
| $Z \to \mu\mu$ | $p_{\mathrm{T}}(\mu) > 20\,\mathrm{GeV}$<br>$60 < \mathrm{m}(\mu\mu) < 120\,\mathrm{GeV}$<br>$\mathrm{IP}(\mu) < 0.04$ |
| $\Upsilon \to \mu\mu$ | both muons $p_{\mathrm{T}} > 3\,\mathrm{GeV}$<br>$9.3 < \mathrm{m}(\mu\mu) < 9.7\,\mathrm{GeV}$ |




**(a)** **(b)** **(c)**

**Figure 3.16:** Invariant mass distribution of (a) $Z$ and (b) $\Upsilon(1S)$and (c) transverse momentum of the muon in data.

# Chapter 4

# Results

The aim of the analysis is to search for sterile neutrinos in a particular final state. This is done by measuring the production of signal events relative to ordinary $W \to \mu\nu$ events. The analysis measures the branching fraction times mixing amplitude, extracted from the data as:

$$\mathcal{B}(N \to \mu\text{jet})\,|V_{\mu N}|^2 = \frac{N_{\text{sig}}}{N_{\text{norm}}}\frac{\varepsilon_{\text{norm}}}{\varepsilon_{\text{sig}}}\left(1 - \frac{m_N^2}{m_W^2}\right)^{-2}\left(1 + \frac{m_N^2}{2m_W^2}\right)^{-1}, \tag{4.1}$$

The mixing of the active muon neutrino with a heavy right-handed neutrino $|V_{\mu N}|^2$ is also estimated by computing $\mathcal{B}(N \to \mu\text{jet})$ as described in Section 1.7.2, assuming $|V_{eN}|^2 = |V_{\tau N}|^2 = 0$.
Several elements are needed to perform the measurement:

- $N_{sgn}$ is the signal yield determined by a fit to the data surviving the full signal selection. The signal model is described in Section 4.3.
- $N_{norm}$ is the yield for the normalisation channel $W \to \mu\nu$ determined by a fit to the data surviving the normalisation selection. The fit is described in Section 4.2.
- $\varepsilon_{\text{norm}}/\varepsilon_{\text{sgn}}$ is the ratio $\varepsilon_{ratio}$ between the normalisation channel selection efficiency and the signal selection efficiency, collected in Table 3.11. Corrections to $\varepsilon_{ratio}$ are evaluated in Section 4.4 to account for differences between data and simulations.
- $\left(1 - \frac{m_N^2}{m_W^2}\right)^2\left(1 + \frac{m_N^2}{2m_W^2}\right)$ is a phase suppression factor with $m_N$ the heavy neutrino mass under consideration and $m_W$ the $W$ mass [32].
- $\mathcal{B}(N \to \mu\text{jet})$ is determined from a theoretical prediction.

Furthermore, systematic uncertainties on the measurement are estimated, as described in Section 4.5, and included in determination of the upper limit. The statistical procedure adopted for the limit calculation and the results obtained are illustrated in Section 4.6.

## 4.1 Binned maximum likelihood fit

An extended binned maximum likelihood method is adopted for all the fits using the HistFactory [107] package.
In general, a fit model consists of a normalised probability density function (PDF) $pdf(\vec{x};\vec{\theta})$ of the variables describing the data points $\vec{x}$ and the unknown parameters $\vec{\theta}$, which determine the shape of the PDF. The extended maximum likelihood method consists in finding the values of the parameters $\vec{\theta}$ that maximise the likelihood:

$$\begin{aligned}\mathcal{L} &= \frac{(\mu_{sgn}+\mu_{bkg})^{N_{obs}} \cdot e^{-(\mu_{sgn}+\mu_{bkg})}}{N_{obs}!} \times \prod_{i=1}^{N_{obs}} pdf(x_i;\vec{\theta}) \\ &= \frac{e^{-(\mu_{sgn}+\mu_{bkg})}}{N_{obs}!} \times \prod_{i=1}^{N_{obs}} \Big(\mu_{sgn}\mathcal{S}(x_i;\vec{\theta}) + \mu_{bkg}\mathcal{B}(x_i;\vec{\theta})\Big)\end{aligned} \tag{4.2}$$

The first factor is the Poisson distribution $P(N_{obs}|\mu_{sgn}+\mu_{bkg})$ which describes the probability of observing the actual number of events given the expected number of events for signal $\mu_{sgn}$ and for background $\mu_{bkg}$. The *pdf* in the second factor can be expressed in terms of its signal and background components as $pdf(x_i;\vec{\theta}) = f_{sgn}\mathcal{S}(x_i;\vec{\theta}) + f_{bkg}\mathcal{B}(x_i;\vec{\theta})$ with $f_{sgn} = \frac{\mu_{sgn}}{\mu_{sgn}+\mu_{bkg}}$ and $f_{bkg} = \frac{\mu_{bkg}}{\mu_{sgn}+\mu_{bkg}}$.
In the case of a binned fit, the likelihood is the product of Poisson distributions for each bin of the templates:

$$\mathcal{L} = \prod_{b\in bins} P(n_b|\nu_b^{sgn} + \nu_b^{bkg}), \tag{4.3}$$

where $n_b$ is the number of events observed in the bin $b$ and $\nu_b^{sgn}$ and $\nu_b^{sgn}$ are the expected number of signal and background events in $b$, respectively.
Constraints on the nuisance parameters[1] need to be included in this expression for the likelihood. Depending on their nature, they can be modelled by Gaussian or Poisson distributions. Typically, the Gaussian distribution describes systematic uncertainties on the nuisance parameters:

$$G(\tilde{\theta}|\theta,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(\tilde{\theta}-\theta)^2}{2\sigma^2}}, \tag{4.4}$$

where $\tilde{\theta}$ is the maximum likelihood estimate of the observable $\theta$. Usually $\theta$ is scaled such that the distribution has unit variance $G(\tilde{\theta}|\theta,1)$.
If the template is statistically limited, bin-by-bin statistical fluctuation are modelled in the likelihood. A nuisance parameter $\theta$ is added per bin to account for the fact that the true rate might differ from the MC expectation. The nuisance parameter is shared

[1]Nuisance refer to parameters whose values are not of interest, but are also not known a priori and need to be obtained from the measurement. An example is an overall data/MC yield mismatch factor.

between the different components of the fit model and is associated to the total Monte Carlo estimate and the total statistical uncertainty in that bin. The MC estimate is then treated as an external measurement described by a Poisson distribution:

$$P(\tilde{\theta}|\theta\lambda) = \frac{(\theta\lambda)^{\tilde{\theta}} e^{-\theta\lambda}}{\tilde{\theta}!} \tag{4.5}$$

where $\lambda$ is the nominal value of $\tilde{\theta}$, which corresponds to the squared ratio between the total statistical uncertainty in the bin and MC estimate in the bin, and the nominal value of $\theta$ is 1, such that $\tilde{\theta}$ fluctuates around $\lambda\theta$.
In this analysis Gaussian constraints are used to describe both the systematic uncertainties and the number of background events estimated in the control regions and then extrapolated to the signal region. Regarding the templates shape, only statistical uncertainties on the simulated templates are included. Hence, both the normalisation and signal channels distributions are modelled by the likelihood function:

$$\mathcal{L} = \prod_{b \in bins} P\left(n_b | \left(\nu_b^{sgn}(\mu, \vec{\theta}) + \sum_{n \in Nbkgs} \nu_{bn}^{bkg}(\vec{\theta})\right)\right) \times \prod_{i=1}^{N_\alpha} G(\tilde{\alpha}_i | \alpha_i, \sigma_i) \times \prod_{i=1}^{N_\gamma} P(\tilde{\gamma}_i | \gamma_i \lambda_i), \tag{4.6}$$

where $N_{bkg}$ is the number of background components in the fit, $N_\alpha$ indicates the number of nuisance parameters which are constrained by a Gaussian distribution and $N_\gamma$ the ones constrained by a Poisson distribution, with $\vec{\theta} = \vec{\alpha} + \vec{\gamma}$. In Eq. 4.6, the dependence on the parameter of interest (PoI) $\mu$ is also explicitly indicated. This represents the yield of $W \to \mu\nu$ component in the normalisation channel fit and $\mathcal{B}(N \to \mu\text{jet})\,|V_{\mu N}|^2$ in the signal model. The PoI is set to zero when testing the background-only hypothesis, as described in Section 4.6.1.

## 4.2 Determination of $W \to \mu\nu$ yield for normalisation

A fit is performed of the $p_{\mathrm{T}}(\mu_W)$ distribution to extract the yield for the normalisation channel $W \to \mu\nu$. The fit is performed independently for positively and negatively charged muons, to account for the difference in production rate at LHCb. The $W^+$ is mainly produced by $u\overline{d}$ collisions and the $W^-$ by $d\overline{u}$ collisions. Hence, given the quark content in proton-proton collisions, the production rate is higher for $W^+$ than for $W^-$. Moreover, since $u$ quarks carry on average higher proton momentum fraction than $d$ quarks, the kinematic distributions of $W^+$ and $W^-$, as well as of the lepton the $W$ decays into, are rather different [108].
The fit is performed in eight bins of muon pseudorapidity and the following contributions

are considered:

- $W \to \mu\nu$: signal shape from MC after the normalisation selection (Table 3.14) is applied. Positively and negatively charged muons have different $p_{\mathrm{T}}$ spectrum, therefore the shapes are obtained separately. The template obtained with MadGraph has been compared to the one of Pythia 8 showing no significant difference at generator level. However, it is known that neither template describes the data properly due to higher order corrections. The $W$ cross-section depends on the angular distribution $(\theta, \phi)$ of the outgoing charged lepton in the $W$ rest frame and on the rapidity $Y$ and transverse momentum $p_{\mathrm{T}}$ of the boson itself in the laboratory frame [109, 110]. The angular dependence of the differential cross-section can be described as a sum of nine harmonic polynomials of the second order whose coefficients $A(p_{\mathrm{T}}, Y)$ implicitly describe the hadronic process. The dependence of these coefficients on the $p_{\mathrm{T}}$ and $Y$ of the $W$ is corrected for NLO contributions to the predictions of DYTurbo [111].

- $W \to \tau\nu$, $Z \to \mu\mu$, $Z \to \tau\tau$: the shape is taken from simulation after applying the normalisation channel selection. Since no difference in the $p_{\mathrm{T}}$ spectrum is observed between positively and negatively charged muons, the template is obtained with no requirement on the charge.

- $b\bar{b}$, $c\bar{c}$: since the shape does not significantly change between the different selection stages, a high statistics template is obtained from the simulation sample after applying the preselection cut. No requirement on the charge is applied and one template is used to fit the yield for both $b\bar{b}$ and $c\bar{c}$ samples since these have compatible shapes.

- QCD (hadron misidentification): template extracted from a sideband of the data selected by applying the normalisation selection with the $\mu_W$ uBDT cut reversed and the requirement on energy deposits reversed and tightened to $(E_{\mathrm{ECAL}} + E_{\mathrm{HCAL}})/p > 0.3$. No charge separation is required.

The templates are obtained per $\eta$ bin. Figure 4.1 shows the templates for the components listed normalised to unity and integrated over the full $\eta$ range. Since the $W \to \mu\nu$ and the $Z \to \mu\mu$ shapes are rather similar, the $Z \to \mu\mu$ yield needs to be constrained. Therefore, first a fit of the $Z$ mass is performed on the Z data sample (Table 3.14) with a template extracted from the $Z \to \mu\mu$ MC sample. The expected yield in the normalisation sample is then obtained as:

$$\frac{N_{norm}(\mathrm{MC})}{N_Z(\mathrm{MC})} N_Z(\mathrm{fit}), \tag{4.7}$$

where the expected number of events is evaluated on the $Z \to \mu\mu$ MC applying the normalisation selection for $N_{norm}(\mathrm{exp})$ and the $Z$ selection for $N_Z(\mathrm{exp})$. The expected

yield and its uncertainty are used to set a Gaussian constraint on the $Z \to \mu\mu$ yield. Since the $Z \to \tau\tau$ and $b\bar{b}$ contributions are expected to be minor, the yields are fixed to the expected values. The $W \to \tau\nu$ background is normalised to a fraction of the $W \to \mu\nu$ yield, where the fraction is fixed to its value from simulation. The total yield for the normalisation channel $W \to \mu\nu$ results in $1762519 \pm 45171$ , of which 57% $W^+$ and 43% $W^-$.
The fit results per bin are collected in Appendix C. A fit example for one $\eta$ bin is shown in Figure 4.2. The data to fit ratio shows an imperfect description of the data by the adopted templates, particularly for high $p_{\mathrm{T}}$ muons at high $\eta$.
The ratio of the measured yields for positively and negatively charged muons as a function of $\eta$ is compared to the expected ratio and the one measured in the $W$ cross-section analysis, performed on data collected in 2012 by the LHCb detector [112]. Also in this analysis the $W$ yield has been obtained from a fit to the transverse momentum spectrum of the muon in eight bins of pseudorapidity. The data selection is the same with two exceptions. First, a veto on a second muon was used, which could not be applied in this analysis as it would remove the signal. Second, to improve the background rejection a muon isolation variable is exploited, which in the current analysis has been replaced by the three uBDTs. Hence, the yields obtained for positively and negatively charged muons can not be directly compared between the two analyses, but the ratio between $W^+$ and $W^-$ yields as a function of pseudorapidity must show the same trend. As can be seen from Figure 4.3a, the three trends are indeed in agreement. In Figure 4.3b the measured ratio is shown separately for the different components. As expected the ratio for $Z \to \mu\mu$, $Z \to \tau\tau$ and $b\bar{b}$ components is constant as a function of $\eta$ and in agreement with one within the uncertainty. For the QCD background it is observed a dependence

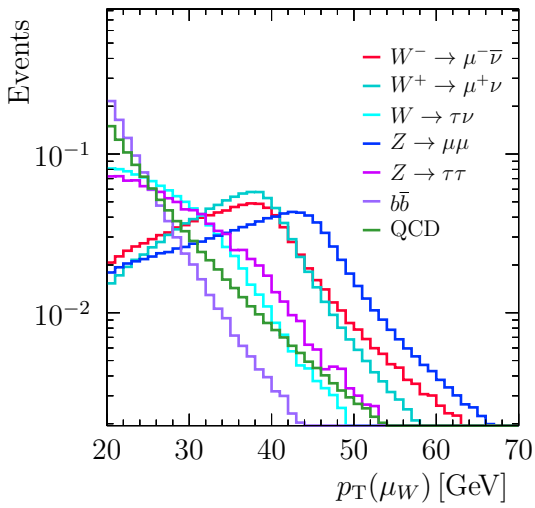



**Figure 4.1:** Templates, normalised to unity, for the signal and background components used in the normalisation channel fit.

on $\eta$. A systematic uncertainty is assigned to the normalisation yield to account for this dependence.

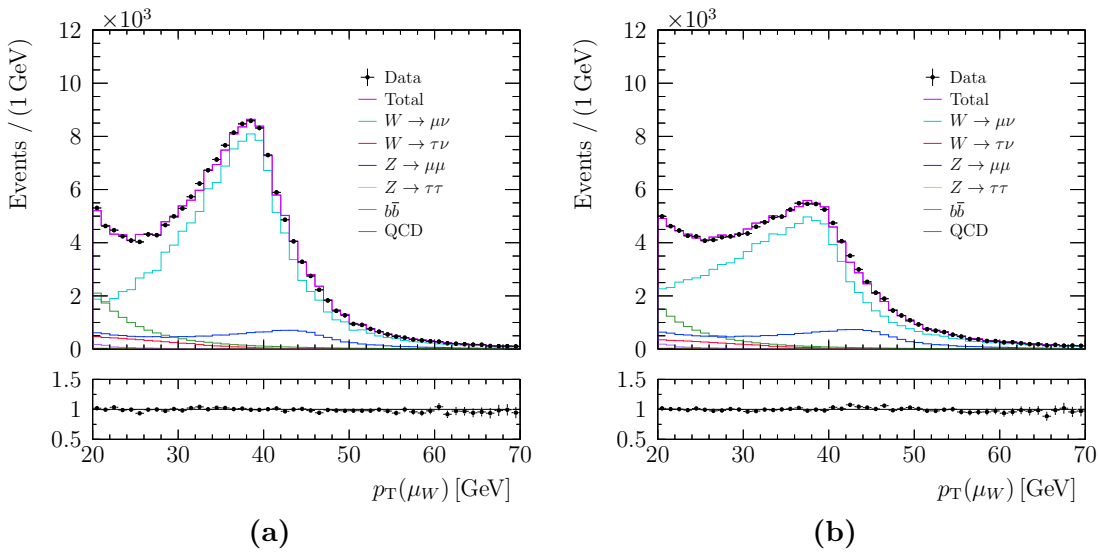



**Figure 4.2:** Fit to the (a) positive and (b) negative muon $p_{\mathrm{T}}$ spectra in the bin $3.00 < \eta < 3.25$ for to normalisation channel. Data (points) are compared to the fit result (magenta line). The fitted contributions are shown separately: $W \to \mu\nu$ (light blue), $Z \to \mu\mu$ (blue), qcd (dark green), $W \to \tau\nu$ (red), $Z \to \tau\tau$ (orange) and $b\bar{b}$ (violet).

## 4.3 Determination of the neutrino mass fit model

The signal yield is determined from a fit to the $\mu_N$ mass. First the yield for each of the three main background categories is obtained from a fit to the $W$ and $b\bar{b}$ control regions. The background prediction is then extrapolated to the signal region via transfer factors determined from simulation and used to constrain the background yields. The expected yield for the light QCD background in the signal region is obtained with a data-driven method. The neutrino mass fit model is then built and validated in regions of the phase space that lie between the signal and the control regions, called validation regions.

### 4.3.1 Determination of background yields in control regions

The fit to the control regions are performed separately for positively and negatively charged muons, as for the normalisation channel. The templates are obtained from the fully simulated samples. The normalisation is left free to vary for each component, except for the $Z \to \mu\mu$ which is constrained to the yield obtained with its dedicated

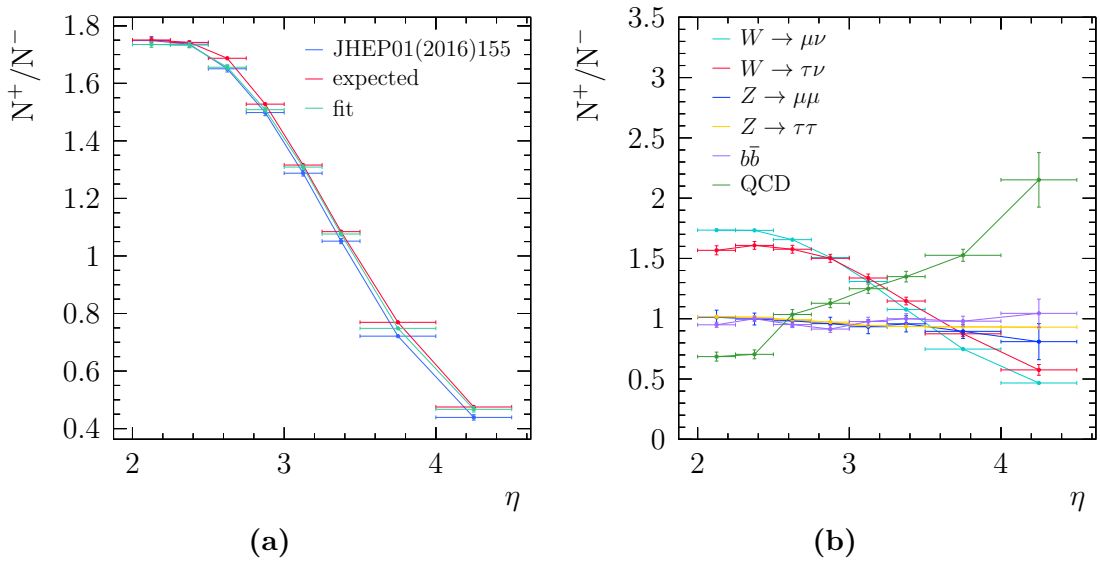



**Figure 4.3:** Ratio of measured yields for positively and negatively charged muons as a function of muon pseudorapidity. In (a) the measured ratio is compared to the expected one and the one measured in the $W$ cross-section analysis [112]. In (b) the measured ratio is shown split by components.

selection. The fitting method is the same as used for the normalisation channel.
The fit has been performed independently for the $W$ and $b\bar{b}$ control regions. In the case of the $b\bar{b}$ control region a fit is not really necessary, since the contributions from other components is expected to be negligible. Nevertheless, the fit projections in Figure 4.4 show that the MC template fits well the data. The fits have also been performed with TFractionFitter Root [113], a tool to perform template fitting implemented within ROOT, as a cross check and the results are in agreement. The fit results are collected in Table 4.1 and Table 4.2. It can be seen that the agreement between the fitted and the expected yield obtained as described in Section 3.4 is good but not perfect, motivating the approach to obtain a normalisation of the background from these control regions.
The fitted background yields are then extrapolated to the signal region via transfer factors, given by the ratio of the number of MC events in the signal region and the number of MC events in the background CR. If the number of events in the signal region is 0, then the number of events is assumed to be $1 \pm 1$, in order to avoid to introduce a bias due to statistical fluctuations. The transfer factor has a statistical uncertainty which is added in quadrature to the statistical uncertainty on the yield from the fit. The number of background events expected in the signal region by extrapolating the yields in the control regions are collected in Table 4.3 for both same-sign and opposite-sign muons events. A few background events are expected in the signal region for same-sign muons samples, while in the case of opposite-sign muons about 1600 events are expected

due to the irreducible Drell-Yan like background.

**Table 4.1:** Fit result for the $W$ control region. The yield obtained from the fit and predicted are reported and compared to the observed number of events on data.

| Charge | $W \to \mu\nu$ fit | $W \to \mu\nu$ prediction | $Z \to \mu\mu$ fit | $Z \to \mu\mu$ prediction | Observed |
|---|---|---|---|---|---|
| plus | $259 \pm 24$ | $204 \pm 15$ | $52 \pm 6$ | $49 \pm 3$ | 312 |
| minus | $148 \pm 47$ | $161 \pm 14$ | $47 \pm 3$ | $51 \pm 3$ | 194 |

**Table 4.2:** Fit result for the $b\bar{b}$ control region. The yield obtained from the fit and predicted are reported and compared to the observed number of events on data.

| Charge | $b\bar{b}$ fit | $b\bar{b}$ prediction | Observed |
|---|---|---|---|
| plus | $253 \pm 16$ | $245 \pm 26$ | 253 |
| minus | $250 \pm 16$ | $334 \pm 31$ | 250 |

**Table 4.3:** Extrapolated events to the signal region for same-sign and opposite-sign muons events.

| Background | Same sign | Opposite sign |
|---|---|---|
| $W \to \mu\nu$ | $1.3 \pm 1.3$ | $3.8 \pm 2.3$ |
| $b\bar{b}$ | $2.4 \pm 2.4$ | $2.4 \pm 2.4$ |
| $Z \to \mu\mu$ | $0.7 \pm 0.5$ | $1672.9 \pm 154.7$ |

### 4.3.2 Determination of light QCD yield in the signal region

The light QCD contribution in the signal region is estimated from data. Assuming that the $\mu_W$ uBDT selection factorises from the other selections that suppress the QCD background, its efficiency is estimated on the single muon sample. Then it is applied to the light QCD region, that is identical to the signal region apart from the $\mu_W$ uBDT requirement as defined in Table 3.12.
The efficiency can be calculated as:

$$\varepsilon^{\text{QCD}}_{\mu_W \text{uBDT}} = \frac{N^{\text{QCD}}_{\text{fit}}}{N^{\text{QCD}}_{\text{before}}}, \tag{4.8}$$

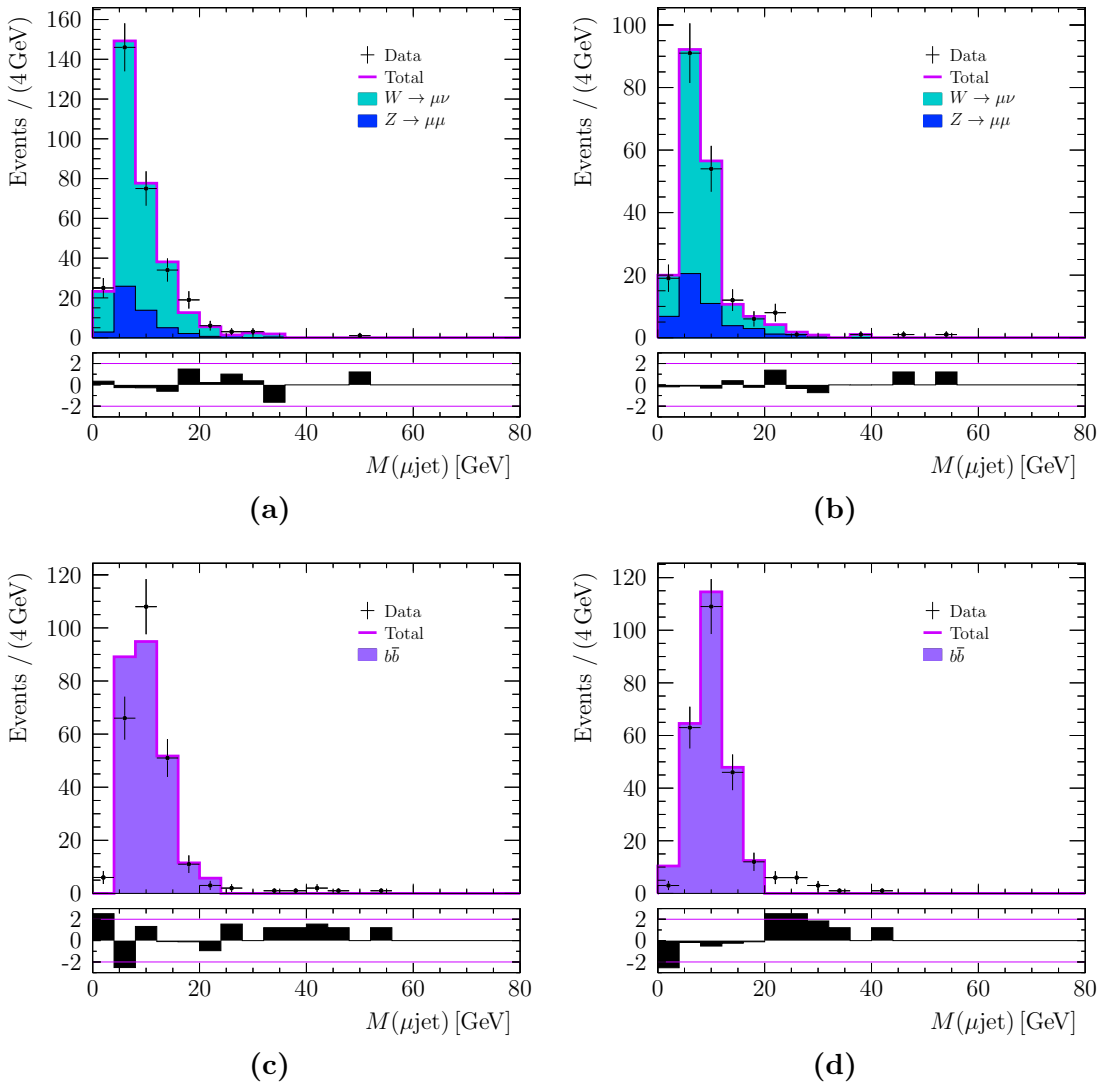



**Figure 4.4:** Fit to (a) positively and (b) negatively charged muons in the $W$ control region and to (c) positively and (d) negatively charged muons in the $b\bar{b}$ control region.

where $N_{\text{fit}}^{\text{QCD}}$ is the QCD fitted yield and $N_{\text{before}}^{\text{QCD}}$ is the number of QCD events before applying the cut. The latter is obtained as:

$$N_{\text{before}}^{\text{QCD}} = N_{\text{before}} - \sum_{\text{backgrounds}} \frac{N_{\text{fit}}^{b}}{\varepsilon_{\text{MC}}^{b}}, \tag{4.9}$$

with $N_{\text{before}}$ the total number of events before applying the cut, $N_{\text{fit}}^{b}$ the fitted yield for the background $b$ and $\varepsilon_{\text{MC}}^{b}$ the $\mu_W$ uBDT cut efficiency for the background $b$ from simulation. The efficiency of the $\mu_W$ uBDT requirement results in $\varepsilon_{\mu_W\text{uBDT}}^{\text{QCD}} = 17 \pm 3\%$. The number of light QCD events in the QCD region $N_{\text{QCDR}}^{QCD}$ is obtained by subtracting

to the total number of events the expected yields for the simulated backgrounds. The number of events per component in the light QCD region is collected in Table 4.4 for both same-sign and opposite-sign muons. Finally, the number of expected light QCD events in the signal region is given by the product

$$N_{\mathrm{SR}}^{\mathrm{QCD}} = \frac{\varepsilon_{\mu W\,\mathrm{uBDT}}^{\mathrm{QCD}}}{1 - \varepsilon_{\mu W\,\mathrm{uBDT}}^{\mathrm{QCD}}} \cdot N_{\mathrm{QCDR}}^{\mathrm{QCD}}, \tag{4.10}$$

and amounts to $0.8 \pm 1.1$ and $0.6 \pm 4.0$ for same-sign and opposite-sign muons, respectively.

**Table 4.4:** Total number of events in data and estimated number of background events in the light QCD control region.

| events | same sign | opposite sign |
|---|---|---|
| total | 14.0 | 287.0 |
| $W \to \mu\nu$ | $2.3 \pm 1.6$ | $2.3 \pm 1.6$ |
| $W \to \tau\nu$ | $0.5 \pm 0.4$ | $0.5 \pm 0.4$ |
| $Z \to \mu\mu$ | $0.4 \pm 0.3$ | $251.5 \pm 7.4$ |
| $Z \to \tau\tau$ | $0.4 \pm 0.3$ | $3.4 \pm 0.8$ |
| $b\bar{b}$ | $5.6 \pm 3.9$ | $25.0 \pm 8.3$ |
| $c\bar{c}$ | $1.1 \pm 0.7$ | $1.1 \pm 0.7$ |
| light QCD | $3.8 \pm 5.7$ | $3.3 \pm 20.4$ |

### 4.3.3 Neutrino mass fit model

To build the neutrino mass fit model the following components are needed:

- background shapes;
- transfer factors to extrapolate the background yields in the control regions to the signal region;
- signal shapes for the six mass hypotheses.

In the simulation, only a few events survive in the signal region. Therefore, the shapes for the different backgrounds are obtained from elsewhere. This procedure relies on the assumption that the shape of the neutrino invariant mass distribution is the same for the final selection and for the selection used to make the template. First, a shape for each background is obtained from the respective MC sample and control region selection. For $Z \to \mu\mu$ it would be the $W$ region. This shape is then compared to the one given

by reversing one by one the cuts that make the control region orthogonal to the signal one. To increase the statistics of the template, the control region and the reversed cut distributions are also compared to the shape obtained when only preselection and trigger requirements are applied. For each of the considered background components, it is found that the neutrino invariant mass distribution is in good agreement for these different selections. The high statistics template is eventually used to build the mass fit model. The only exception is given by the opposite-sign muons $Z \to \mu\mu$ component, for which there is a sufficient number of events in the signal region and there is no need to obtain the shape from somewhere else. Hence, for this component only the shape is obtained by applying the full signal selection. The fit is performed to both positive and negative $W$ combined, hence no charge requirement is applied to generate the templates.

The background yields fitted in the control regions allow to constrain the normalisation of the respective backgrounds in the signal region. The yield is extrapolated via a transfer factor, which is computed from the MC as the ratio between the number of selected events in the signal region and in the control region. The extrapolated mean and statistical error are used to set the mean and sigma of a gaussian constraint applied to the normalisation parameter. In the case of the $b\bar{b}$ component, the mean of the gaussian is set to be the sum of the extrapolated $b\bar{b}$ yield and expected light QCD yield, while the width is set to the propagated uncertainty, under the assumption that the light QCD background has the same shape as the $b\bar{b}$ background.
As for the background shapes, the templates for the six signal mass points are obtained from simulation, applying trigger and preselection requirements.

### 4.3.4 Validation of the neutrino mass fit model

The neutrino mass fit model is tested in validation regions that are in between the control regions and the signal region. The selection cuts defining the validation regions are obtained starting from the signal region selection and reversing a specific cut. This procedure has been applied to each of the cuts that make the signal region orthogonal to the control regions (Table 3.12). An example of validation region is shown in Figure 4.5, where the data are compared to the sum of the backgrounds, each of which is normalised to the extrapolated yield. The shape of each backgrounds is the one used in the signal model for all the validation regions but the two opposite-sign muons regions obtained by reversing the $\mu_N$ uBDT and kinematic uBDT requirements. These cuts modify the shape of the $Z \to \mu\mu$ background. Hence, in these two validation regions the shape used for the $Z \to \mu\mu$ component is obtained by reversing the corresponding cut. The comparison for each of the validation regions can be found in Appendix D showing good agreement. The events expected, extrapolated from the fit and counted in the data sample are compared in Table 4.5 and in Table 4.6 for the same- and opposite-sign muons samples, respectively.

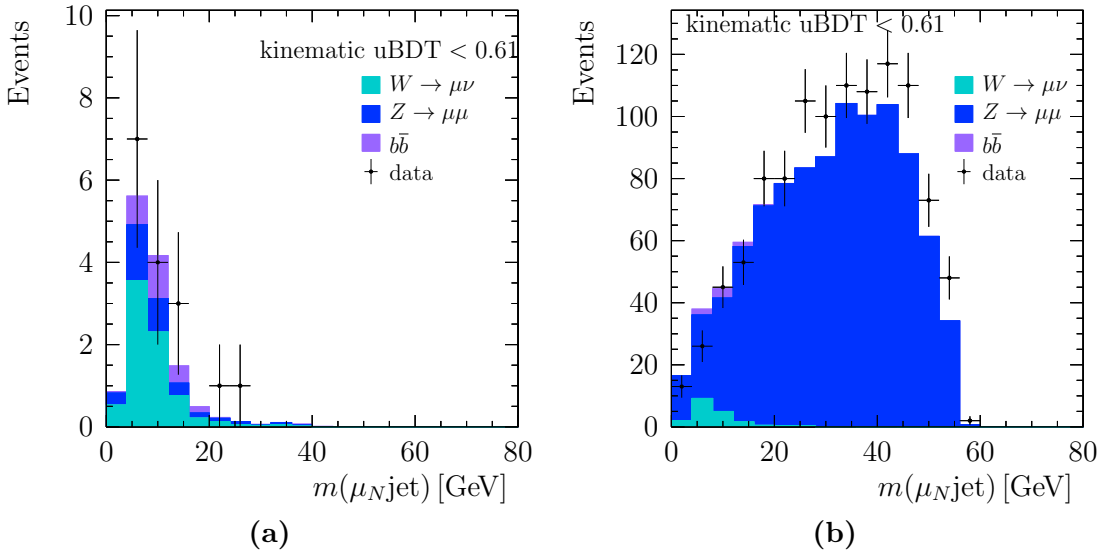



**Figure 4.5:** Comparison between data (points) and expected backgrounds (stack) in the validation regions obtained by reversing one cut in the signal selections (a) same-sign muons and (b) opposite-sign muons. The cut on kinematic uBDT is reversed.

**Table 4.5:** Number of events observed in data, expected from simulation (labelled as *expected*) and extrapolated to the validation region from the fitted yield in the control regions (labelled as *fit*), in different validation regions for same-sign muons samples. Each region is defined starting from the signal selection and reversing one specific cut.

| Region | Label | data | fit | expected |
|---|---|---|---|---|
| sgn | $\text{IP}(\mu_W) > 0.04\,\text{mm}$ | 1 | $4 \pm 3$ | $8 \pm 6$ |
| sgn | $\mu_W$ uBDT $< 0.55$ | 14 | $4 \pm 3$ | $8 \pm 6$ |
| sgn | $\mu_N$ uBDT $< 0.60$ | 13 | $13 \pm 4$ | $15 \pm 7$ |
| sgn | kinematic uBDT $< 0.62$ | 16 | $13 \pm 4$ | $15 \pm 7$ |
| sgn | $\text{IP}(\mu_N) > 0.1\,\text{mm}$ | 2 | $4 \pm 3$ | $8 \pm 6$ |

**Table 4.6:** Number of events observed in data, expected from simulation (labelled as *expected*) and extrapolated to the validation region from the fitted yield in the control regions (labelled as *fit*), in different validation regions for opposite-sign muons samples. Each region is defined starting from the signal selection and reversing one specific cut.

| Region | Label | data | fit | expected |
|---|---|---|---|---|
| sgn | $\mathrm{IP}(\mu_W) > 0.04\,\mathrm{mm}$ | 74 | $43 \pm 5$ | $42 \pm 8$ |
| sgn | $\mu_W$ uBDT $< 0.55$ | 287 | $320 \pm 28$ | $285 \pm 17$ |
| sgn | $\mu_N$ uBDT $< 0.60$ | 183 | $242 \pm 22$ | $214 \pm 13$ |
| sgn | kinematic uBDT $< 0.62$ | 1070 | $972 \pm 82$ | $846 \pm 24$ |
| sgn | $\mathrm{IP}(\mu_N) > 0.1\,\mathrm{mm}$ | 22 | $17 \pm 3$ | $20 \pm 7$ |

## 4.4 Efficiency corrections

The efficiency of the selection requirements for both normalisation and signal samples has been obtained from simulation and needs to be corrected for MC-data differences. This is achieved by computing corrections from control samples, such as $Z \to \mu\mu$, which are then applied to the efficiency of the sample under consideration. Corrections are applied to the reconstruction efficiency of the two muons, to the efficiency of the global events cut within the L0 muon trigger line and to the efficiency of the $\mu_W$ uBDT and $\mu_N$ uBDT cuts. All the corrections, with the exception of the one related to the global event cut, are evaluated in bins of kinematic variables, that is, the pseudorapidity and the momentum or transverse momentum of the muon. The efficiency correction $\rho_i$ for the bin $i$ is defined as:

$$\rho_i \equiv \frac{\varepsilon(\text{data bin } i)}{\varepsilon(\text{simulation bin } i)}, \tag{4.11}$$

where *data* and *simulation* refer to the control sample. The per bin corrections are then applied to the signal or normalisation sample to get the corrected efficiency as:

$$\varepsilon(\text{corrected}) = \frac{\sum_i \rho_i M_i}{\sum_i N_i}, \tag{4.12}$$

with $N_i$ the number of events in the bin $i$ before the cut under evaluation is applied and $M_i$ the number of events after the cut is applied. The average efficiency correction over the whole kinematic range is then obtained as the ratio of corrected and uncorrected efficiency, leading to:

$$\bar{\rho} = \frac{\varepsilon(\text{corrected})}{\varepsilon(\text{uncorrected})} = \frac{\sum_i \rho_i M_i}{\sum_i M_i}. \tag{4.13}$$

Hence, it is computed by integrating over the $M_i$ distribution after the correction is applied. A systematic uncertainty per correction factor is assigned by propagating the uncertainty per kinematic bin to the correction factors.
The correction factors obtained are collected in Table 4.7 and Table 4.8 for the same- and opposite-sign muons signal samples, respectively. The dominant correction is due to the global event cut efficiency.

### 4.4.1 Muon reconstruction

#### High-$p_{\mathrm{T}}$ Muon

The reconstruction efficiency for high-$p_{\mathrm{T}}$ muons has been measured at LHCb from data collected in 2012 [114]. A tag-and-probe method is applied on a $Z \to \mu\mu$ sample, which is a pure sample of dimuon pairs. The method consists of reconstructing, triggering and identifying a muon, which is used as tag, and evaluating the efficiency of a specific requirement on another reconstructible object, called the probe. The definition of

reconstructible object depends on the particular selection requirement to be evaluated and is described for each case in the following.
In events with high-$p_{\mathrm{T}}$ jets the occupancies tend to be higher, but it has been verified [115] that the efficiencies evaluated on data with the tag-and-probe method do not significantly depend on the jet $p_{\mathrm{T}}$ or the jet multiplicity. Hence, the scale factors from [114] are used in the analysis presented here. The scale factors are applied as a function of muon transverse momentum and pseudorapidity.
The overall muon reconstruction efficiency can be factorised in three distinct components:

$$\varepsilon^{rec} = \varepsilon^{trk} \times \varepsilon^{id} \times \varepsilon^{trg} \tag{4.14}$$

where

- $\varepsilon^{trk}$ is the efficiency for a muon track in the event to be reconstructed as a long track, given that the muon track is reconstructed by combining hits in the muon stations and the TT in a so-called MuonTT track. The uncertainty on the per bin correction is the sum in quadrature of statistical and systematic uncertainty. The systematic uncertainty accounts for background contamination. The total uncertainty varies between 0.1-0.9%;

- $\varepsilon^{id}$ is the efficiency for identifying a long track as a muon according to the *isMuon* requirement, given that the muon is reconstructed as long track. The per bin uncertainty is due to corrections applied to the data and varies between 0.3-1.2%.

- $\varepsilon^{trg}$ is the efficiency for a muon to pass the single muon trigger lines at L0, HLT1 and HLT2, given that the muon is reconstructed as long track and identified according to the *isMuon* requirement. The uncertainty on the per bin efficiency correction is statistical and varies between 0.3-1.3%;

The three components have been evaluated separately. The ratio of efficiencies evaluated in data and simulation are used to correct the $\mu_W$ reconstruction efficiency.

### Low-$p_{\mathrm{T}}$ Muon

The low-$p_{\mathrm{T}}$ muon is not required to pass the trigger lines. Therefore the overall muon reconstruction efficiency is given by the track reconstruction efficiency and the identification efficiency. The track reconstruction efficiency for muon $p_{\mathrm{T}} > 3\,\mathrm{GeV}$ is evaluated using the tag-and-probe method in data and simulation. Corrections are obtained from two different control samples, $J/\psi \to \mu\mu$ and a $Z \to \mu\mu$, since neither of them allows to fully cover the kinematic range of the $\mu_N$. For the $Z \to \mu\mu$ sample, the same corrections as for the $\mu_W$ are applied. The ratio of track reconstruction efficiency between data and simulation from the $J/\psi \to \mu\mu$ sample has been evaluated in [116] in bins of $\eta$ and $p$. The per bin ratio is compatible with unity and measured with an uncertainty of 0.4%,

including both statistical and systematic uncertainty, in 2012. The average corrections per signal sample obtained with the two control samples are in agreement within the uncertainty. The scale factors obtained from the $J/\psi \to \mu\mu$ sample are used eventually to correct the low-$p_{\mathrm{T}}$ muon tracking efficiency. Regarding the muon identification efficiency, the corrections from the $Z \to \mu\mu$ sample are used to weight the kinematic distribution, as for the $\mu_W$.

### 4.4.2 Global event cut

The L0 trigger line required in this analysis contains a global event cut (GEC): the hit multiplicity in the SPD, labelled in the following as nSPDhits, must be less than 600. This requirement removes events with high occupancy which would require a long processing time in the HLT.
The nSPDhits distributions in data and simulation are significantly different, as shown in Figure 4.6a for the $Z \to \mu\mu$ channel. The comparison between simulated $Z \to \mu\mu$, $W \to \mu\nu$ and signal distributions can be seen in Figure 4.6b. In the following it is assumed that the difference between data and simulation can essentially be described as a continuous transformation. Then using real and simulated $Z \to \mu\mu$ events it is derived the nSPDhits threshold that needs to be applied to simulated events in order to obtain the same efficiency as the actual cut applied to the data.
First the efficiency for nSPDhits $< 600$ is evaluated on $Z \to \mu\mu$ data events that have been selected by the L0 dimuon trigger. This trigger line has a cut nSPDhits $< 900$, rather than 600, whose efficiency is almost 100%. Second, the cut corresponding to this efficiency on the $Z \to \mu\mu$ simulation sample is found. The procedure was performed separately for events with different number of primary vertices as well as to all events together. A dependence on the number of primary vertices is observed, but since the distribution of the number of primary vertices is well reproduced in the simulation the correction to the GEC efficiency is evaluated with no requirement on the number of primary vertices. Statistical uncertainties on the efficiencies are calculated as binomial confidential interval and propagated to the cut uncertainty. The new cut, found to be $401 \pm 1$, is then applied to both the normalisation channel and to the signal channels to evaluate the efficiencies. The efficiency correction factor per signal and normalisation samples is obtained as the ratio of the efficiency corresponding to the new cut and the efficiency of the nSPDhits $> 600$ requirement.

### 4.4.3 $\mu_W$ uBDT cut

The kinematic coverage of $\mu_W$ largely overlaps with the coverage of muons from a $Z$ decay. The efficiency for the $\mu_W$ uBDT cut is therefore evaluated on a sample of $Z \to \mu\mu$ events as a function of pseudorapidity and transverse momentum of both muons. The

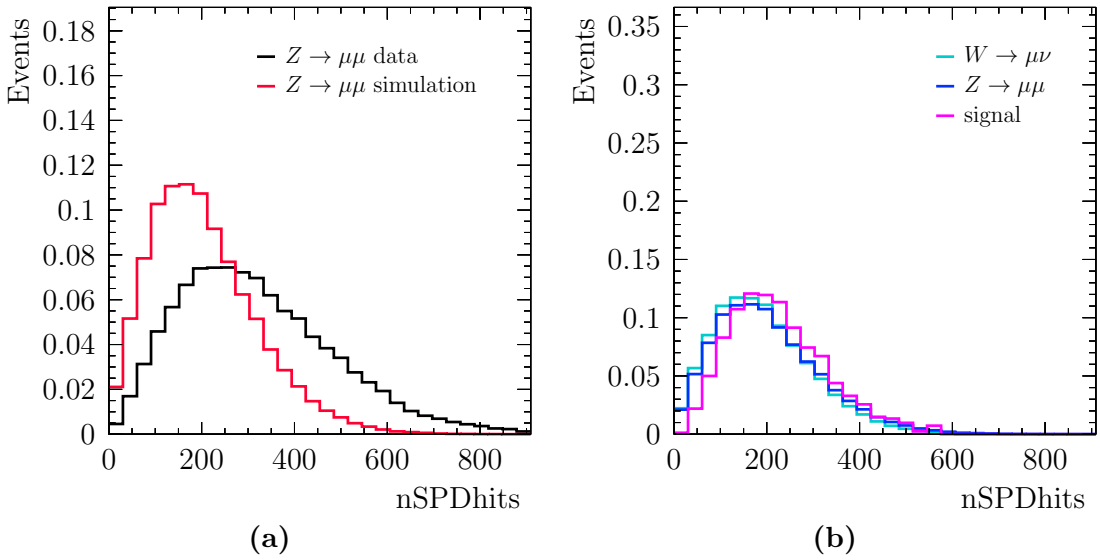



**Figure 4.6:** Distributions of nSPDhits for (a) $Z \to \mu\mu$ events in data and simulation and (b) for simulated $Z \to \mu\mu$, $W \to \mu\nu$ and signal events.

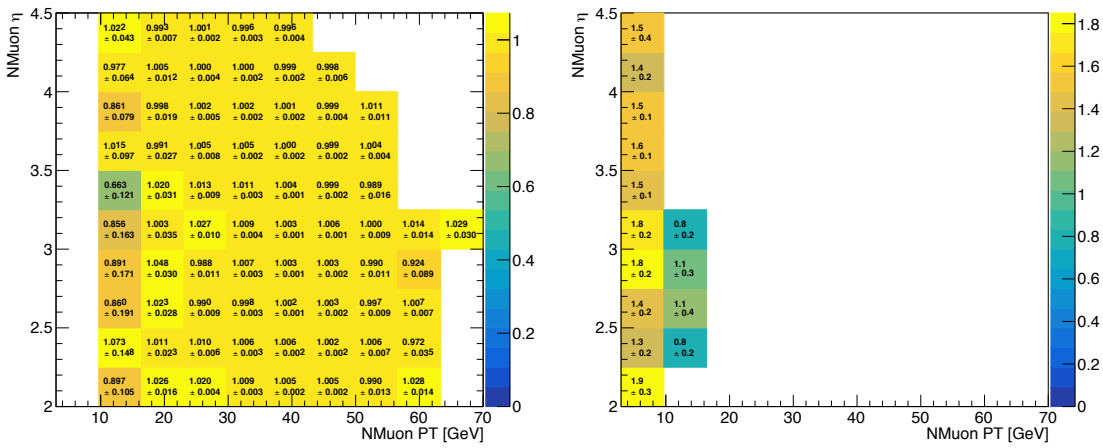
$Z$ events are selected according to the requirements in Table 3.14. The ratio between the muon efficiency in $Z \to \mu\mu$ simulation and data represents the correction factor for the signal efficiency obtained from simulation.

### 4.4.4 $\mu_N$ uBDT cut

The $p_{\mathrm{T}}$ spectrum of the $\mu_N$ extends to lower values with respect to the muons from a $Z$ decay. Therefore an additional sample of $\Upsilon(1S) \to \mu\mu$ events is used to evaluate the efficiency correction for muons with transverse momentum lower than 10 GeV.
The selection for the $Z \to \mu\mu$ sample is modified in the following way: the trigger requirements are applied to the high-$p_{\mathrm{T}}$ muon and the cut on the $p_{\mathrm{T}}$ of the other muon is softened from 20 to 3 GeV. No additional background is expected in the $Z \to \mu\mu$ sample due to the looser cut on the $\mu_N$ $p_{\mathrm{T}}$. The correction to the efficiency is then obtained in bins of transverse momentum and pseudorapidity for the low-$p_{\mathrm{T}}$ muon as the ratio between the efficiency evaluated on the data sample and on the simulated sample. As can be seen from Figure 4.7a, the lowest $p_{\mathrm{T}}$ bin, for $p_{\mathrm{T}} < 10$ GeV, is not covered. For $p_{\mathrm{T}} < 10$ GeV, the $\Upsilon(1S) \to \mu\mu$ sample is used. The $\Upsilon(1S)$events are selected according to the requirements in Table 3.14. Since this sample is not as pure as the $Z \to \mu\mu$, the number of signal events before and after the $\mu_N$ uBDT cut is obtained by fitting the data with a Hypatia [117] function[2]. In Figure 4.8 is shown a fit example

[2]The Hypatia function is a generalisation of the Crystal Ball function, which consists of a Gaussian describing the detector resolution and a tail on the left-hand side parametrising the final state radiation. Differently from the Crystal Ball, the Hypatia function has a generalised hyperbolic core to take into

for $3.25 < \eta < 3.50$ and $3.0 < p_{\mathrm{T}} < 9.7\,\mathrm{GeV}$ muons and it can be seen that the efficiency of the $\mu_N$ uBDT cut is low for muons coming from $\Upsilon(1S)$. This is due to the fact that the $\mu_N$ uBDT classifier exploits the transverse momentum observable to reject the low-$p_{\mathrm{T}}$ background. Since after the application of the full selection the $\mu_N$ spectrum is harder, the lowest $p_{\mathrm{T}}$ bins do not affect significantly the correction factor.
Comparing Figure 4.7a and 4.7b, obtained using $Z \to \mu\mu$ and $\Upsilon(1S) \to \mu\mu$ events, respectively, it can be seen that the two samples are complementary in terms of kinematic coverage and therefore both are used to evaluate the correction to the efficiency of the $\mu_N$ uBDT cut.




**(a)** **(b)**

**Figure 4.7:** Efficiency corrections in bins of $(\eta, p_{\mathrm{T}})$ for the $\mu_N$ uBDT cut efficiency evaluated (a) on the $Z \to \mu\mu$ sample and (b) on the $\Upsilon(1S) \to \mu\mu$. The white areas correspond to empty bins.

## 4.5 Systematic uncertainties

In this section systematic uncertainties on the computation of the signal strength due to the selection criteria and to the normalisation channel fit are presented.

### 4.5.1 Efficiency ratio uncertainty

#### Efficiency Corrections

The systematic uncertainties on the efficiency correction factors for signal and normalisation channels are reported in Table 4.7 and Table 4.8, for same- and opposite-sign

account per-event mass uncertainties that would distort the Gaussian distribution.

**Table 4.7:** Efficiency correction factors for the same-sign signal samples. The uncertainty is systematic.

| Correction | Signal 5 GeV | Signal 10 GeV | Signal 15 GeV | Signal 20 GeV | Signal 30 GeV | Signal 50 GeV | Normalisation |
|---|---|---|---|---|---|---|---|
| $\mu_W$ id | $0.979 \pm 0.002$ | $0.979 \pm 0.002$ | $0.978 \pm 0.002$ | $0.977 \pm 0.002$ | $0.973 \pm 0.002$ | $0.960 \pm 0.004$ | $0.984 \pm 0.001$ |
| $\mu_W$ tracking | $0.995 \pm 0.002$ | $0.995 \pm 0.002$ | $0.995 \pm 0.002$ | $0.995 \pm 0.002$ | $0.995 \pm 0.003$ | $0.994 \pm 0.004$ | $0.995 \pm 0.002$ |
| $\mu_W$ trigger | $1.064 \pm 0.002$ | $1.063 \pm 0.002$ | $1.064 \pm 0.002$ | $1.064 \pm 0.002$ | $1.065 \pm 0.003$ | $1.081 \pm 0.007$ | $1.063 \pm 0.003$ |
| $\mu_N$ id | $0.984 \pm 0.002$ | $0.984 \pm 0.002$ | $0.984 \pm 0.002$ | $0.984 \pm 0.002$ | $0.983 \pm 0.002$ | $0.978 \pm 0.003$ | - |
| $\mu_N$ tracking | $1.010 \pm 0.001$ | $1.011 \pm 0.001$ | $1.010 \pm 0.001$ | $1.010 \pm 0.001$ | $1.009 \pm 0.001$ | $1.008 \pm 0.001$ | - |
| $\mu_W$ BDT | $0.980 \pm 0.002$ | $0.980 \pm 0.003$ | $0.979 \pm 0.003$ | $0.977 \pm 0.003$ | $0.974 \pm 0.005$ | $0.964 \pm 0.023$ | $0.977 \pm 0.005$ |
| $\mu_N$ BDT | $1.012 \pm 0.001$ | $1.012 \pm 0.001$ | $1.011 \pm 0.001$ | $1.010 \pm 0.001$ | $1.009 \pm 0.001$ | $1.006 \pm 0.003$ | - |
| nSPD hits | $0.888 \pm 0.004$ | $0.885 \pm 0.004$ | $0.873 \pm 0.005$ | $0.870 \pm 0.005$ | $0.861 \pm 0.005$ | $0.868 \pm 0.005$ | $0.939 \pm 0.001$ |

**Table 4.8:** Efficiency correction factors for the opposite-sign signal samples. The uncertainty is systematic.

| Correction | Signal 5 GeV | Signal 10 GeV | Signal 15 GeV | Signal 20 GeV | Signal 30 GeV | Signal 50 GeV | Normalisation |
|---|---|---|---|---|---|---|---|
| $\mu_W$ id | $0.979 \pm 0.002$ | $0.979 \pm 0.002$ | $0.978 \pm 0.002$ | $0.977 \pm 0.002$ | $0.973 \pm 0.002$ | $0.960 \pm 0.004$ | $0.984 \pm 0.001$ |
| $\mu_W$ tracking | $0.995 \pm 0.002$ | $0.995 \pm 0.002$ | $0.995 \pm 0.002$ | $0.995 \pm 0.002$ | $0.995 \pm 0.003$ | $0.994 \pm 0.004$ | $0.995 \pm 0.002$ |
| $\mu_W$ trigger | $1.063 \pm 0.002$ | $1.064 \pm 0.002$ | $1.064 \pm 0.002$ | $1.064 \pm 0.002$ | $1.065 \pm 0.003$ | $1.080 \pm 0.006$ | $1.063 \pm 0.003$ |
| $\mu_N$ id | $0.982 \pm 0.002$ | $0.982 \pm 0.002$ | $0.982 \pm 0.002$ | $0.982 \pm 0.002$ | $0.981 \pm 0.002$ | $0.977 \pm 0.003$ | - |
| $\mu_N$ tracking | $1.010 \pm 0.001$ | $1.010 \pm 0.001$ | $1.010 \pm 0.001$ | $1.010 \pm 0.001$ | $1.010 \pm 0.001$ | $1.008 \pm 0.001$ | - |
| $\mu_W$ BDT | $0.981 \pm 0.002$ | $0.980 \pm 0.003$ | $0.980 \pm 0.003$ | $0.979 \pm 0.003$ | $0.974 \pm 0.005$ | $0.963 \pm 0.022$ | $0.977 \pm 0.005$ |
| $\mu_N$ BDT | $1.011 \pm 0.002$ | $1.012 \pm 0.001$ | $1.012 \pm 0.001$ | $1.010 \pm 0.001$ | $1.009 \pm 0.001$ | $1.008 \pm 0.003$ | - |
| nSPD hits | $0.888 \pm 0.004$ | $0.885 \pm 0.004$ | $0.873 \pm 0.005$ | $0.870 \pm 0.005$ | $0.861 \pm 0.005$ | $0.868 \pm 0.005$ | $0.939 \pm 0.000$ |

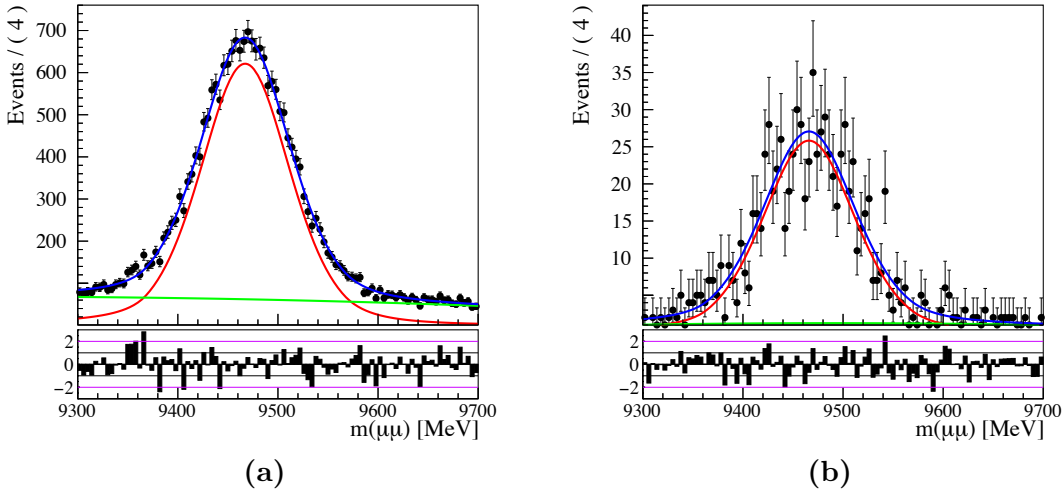



**Figure 4.8:** Fit to the $\Upsilon(1S) \to \mu\mu$ data sample for $3.25 < \eta < 3.50$ and $3.0 < p_{\mathrm{T}} < 9.7\,\mathrm{GeV}$ (a) before and (b) after the requirement on $\mu_N$ uBDT. The background component is in green, the signal component in red and the total in blue.

muons, respectively. Eventually a systematic uncertainty needs to be assigned to the efficiency ratio. The uncertainties are expected to largely cancel when signal and normalisation channels selections are identical. A conservative approach has been adopted, which consists of dropping the smallest uncertainty when computing a ratio of highly positively correlated efficiencies. The uncertainties on the different correction factors are then added in quadrature. The ratio of normalisation and signal correction factors are collected in Table 4.9. The relative uncertainty on the correction factors is of the order of 2%.

**Table 4.9:** Ratio of efficiency correction factors between normalisation and signal samples. The uncertainty is systematic.

| $\varepsilon_{norm}/\varepsilon_{sgn}$ | same sign | opposite sign |
|---|---|---|
| 5 GeV | $1.05 \pm 0.01$ | $1.06 \pm 0.01$ |
| 10 GeV | $1.06 \pm 0.01$ | $1.06 \pm 0.01$ |
| 15 GeV | $1.07 \pm 0.01$ | $1.07 \pm 0.01$ |
| 20 GeV | $1.08 \pm 0.01$ | $1.08 \pm 0.01$ |
| 30 GeV | $1.10 \pm 0.01$ | $1.10 \pm 0.01$ |
| 50 GeV | $1.12 \pm 0.03$ | $1.12 \pm 0.03$ |

### Jet Energy Scale and Resolution

Uncertainties are assigned to account for the possibility that the detector response to jets is different between data and simulation.
The uncertainty on the jet energy scale determination has three distinct sources: charged particles, neutral particles and the fraction of jet $p_{\mathrm{T}}$ carried by ghosts. The dominant contribution arises from the disagreement between data and simulation for neutral particles, which yields an uncertainty of about 5% [118]. The effect of this uncertainty on the efficiency ratio is determined by reducing the jet energy by 5% in simulation. The jet energy resolution uncertainty is obtained by smearing the jet $p_{\mathrm{T}}$. This is realised by taking a random number from a Gaussian of mean 1 and width 0.1 for each jet. Both energy scale and resolution uncertainties affect the reconstruction and stripping efficiency as well the selection efficiency, since the jet transverse momentum is among the variables used in the training of the kinematic uBDT. Hence, the efficiency is recalculated and the relative difference with respect to the efficiency previously obtained is assigned as systematic uncertainty. Both uncertainties due to the jet energy scale and jet energy resolution are between 5% and 9%. A fluctuation of $\sim 5\%$ between the different mass points is expected due to the low statistics available.

### Jet Identification

The uncertainty due to the jet identification requirements is taken from the production cross-section analysis of $W$ and $Z$ in association with jets at the centre of mass energy of 8 TeV [118], since the same cuts on jet MPT and CPF are applied. The uncertainty is obtained by tightening each cut both in data and simulation and comparing the fraction of events rejected in data with respect to the simulation. The MPT requirement is tightened from 1.2 GeV to 2.4 GeV and the CPF requirement from 10% to 20%. The variable which shows the worst agreement between data and simulation is used to set the uncertainty. The overall uncertainty results to be 1.7%.

### Missing $p_{\mathrm{T}}$ observable

An uncertainty arises due to the mismatch between data and simulation of the missing $p_{\mathrm{T}}$ observable, which is used for the training of the kinematic uBDT. The uncertainty is estimated by comparing data and simulation for the $Z \to \mu\mu$ channel. A new value of missing $p_{\mathrm{T}}$ is extracted per event by requiring that the cumulative distribution for MC matches the data one. Then the total efficiency is recalculated and the relative difference with respect to the efficiency previously obtained is assigned as systematic uncertainty. The systematic uncertainty results to be a few percents.

**Table 4.10:** Systematic uncertainties on the efficiency ratio in percent for the same-sign muons samples. Fluctuations of about five percent-points on the uncertainties due to energy scale and energy resolution are expected due to the low statistics of the samples.

| Signal | 5 GeV | 10 GeV | 15 GeV | 20 GeV | 30 GeV | 50 GeV |
|---|---|---|---|---|---|---|
| efficiency corrections | 0.85 | 0.86 | 0.89 | 0.90 | 0.98 | 2.60 |
| jet energy scale | 8.65 | 7.44 | 8.66 | 9.08 | 10.73 | 4.92 |
| jet energy resolution | 7.70 | 4.20 | 5.73 | 5.42 | 7.74 | 2.46 |
| jet identification | 1.70 | 1.70 | 1.70 | 1.70 | 1.70 | 1.70 |
| missing pt | 0.95 | 1.92 | 0.98 | 1.13 | 1.36 | 2.46 |
| total | 11.77 | 8.96 | 10.60 | 10.81 | 13.45 | 6.78 |

**Table 4.11:** Systematic uncertainties on the efficiency ratio in percent for the opposite-sign muons samples. Fluctuations of about five percent-points on the uncertainties due to energy scale and energy resolution are expected due to the low statistics of the samples.

| Signal | 5 GeV | 10 GeV | 15 GeV | 20 GeV | 30 GeV | 50 GeV |
|---|---|---|---|---|---|---|
| efficiency corrections | 0.85 | 0.86 | 0.89 | 0.91 | 0.97 | 2.49 |
| jet energy scale | 5.60 | 6.71 | 8.23 | 11.26 | 8.29 | 8.65 |
| jet energy resolution | 3.62 | 5.62 | 6.28 | 7.32 | 6.41 | 6.57 |
| jet identification | 1.70 | 1.70 | 1.70 | 1.70 | 1.70 | 1.70 |
| missing pt | 0.12 | 0.59 | 0.54 | 0.90 | 2.39 | 2.56 |
| total | 6.93 | 8.98 | 10.54 | 13.60 | 10.92 | 11.56 |

#### Total uncertainty

The uncertainty on the energy scale represents the dominant systematic uncertainty on the signal efficiency. The ratios of efficiencies between the normalisation and each signal channel are collected in Table 4.12. The correction factors are applied. It can be seen that statistical and systematic uncertainties are of the same order.

### 4.5.2 Background yields uncertainties

The systematic uncertainties on transfer factors are neglected. Due to the few background events expected in the signal region, the statistical uncertainties dominate.

**Table 4.12:** Ratio of efficiencies between normalisation and signal channels. Corrections are applied. The first uncertainty is statistical, the second uncertainty is systematic.

| $\varepsilon_{norm}/\varepsilon_{sgn}$ | same sign | opposite sign |
|---|---|---|
| 5 GeV | $25 \pm 1 \pm 3$ | $21 \pm 1 \pm 1$ |
| 10 GeV | $24 \pm 1 \pm 2$ | $19 \pm 1 \pm 2$ |
| 15 GeV | $26 \pm 1 \pm 3$ | $23 \pm 1 \pm 2$ |
| 20 GeV | $28 \pm 1 \pm 3$ | $25 \pm 1 \pm 3$ |
| 30 GeV | $32 \pm 1 \pm 4$ | $30 \pm 1 \pm 3$ |
| 50 GeV | $55 \pm 2 \pm 3$ | $43 \pm 2 \pm 4$ |

### 4.5.3 Normalisation yield uncertainty

Several sources of systematic uncertainties are considered in the determination of the normalisation channel yield. To estimate the uncertainty both the shape and normalisation of the templates are varied. For each variation the fit is repeated and the deviations from the $W \to \mu\nu$ yield for each source are added in quadrature. A contribution has been evaluated for each of the following variations:

- The QCD template is replaced by an exponential.
- The weights applied to the $W \to \mu\nu$ template are reduced of 10% towards unity.
- To account for the different QCD yield between the fit to positively and negatively charged muons, the normalisation of the QCD component is fixed bin by bin to the average of the yields.
- The yields for $Z \to \mu\mu$, $Z \to \tau\tau$, $b\bar{b}$ components, that are fixed in the fit, are varied individually by $\pm 1\sigma$: between the two, the largest variation is assigned as the systematic uncertainty.

The different uncertainties and their combination are collected in Table 4.13.

**Table 4.13:** Systematic uncertainties on the normalisation channel yield in percentage.

| | |
|---|---|
| QCD shape | 1.053 |
| $W \to \mu\nu$ shape | 2.495 |
| QCD yield | 0.762 |
| $Z \to \mu\mu$ yield | 0.073 |
| $b\bar{b}$ yield | 0.008 |
| $Z \to \tau\tau$ yield | 0.009 |
| total | 2.831 |

## 4.6 Upper limit determination

In order to assess the compatibility of the data with a given hypothesis a statistical test is performed. First, the statistical hypothesis test method used is described, then the results are presented.

### 4.6.1 Statistical hypothesis test

The first step in the frequentist statistical hypothesis test is to define the so called null hypothesis $H_0$. In the case of limit setting $H_0$ consists of a description of the data including background and signal, which in this analysis is a heavy neutrino with a given mass, to be tested against the background-only hypothesis $H_1$. To evaluate the validity of the hypothesis, a unique function of the observables describing the data, named test statistic, is defined. The test statistic adopted in this analysis is the profile likelihood ratio:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\vec{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\hat{\vec{\theta}}})}, \tag{4.15}$$

where

- $\mu$ is the parameter of interest, that is $\mathcal{B}(N \to \mu\text{jet})\,|V_{\mu N}|^2$;
- $\hat{\hat{\vec{\theta}}}$ is the value of the nuisance parameters $\vec{\theta}$ that maximises the likelihood for a specific $\mu$ given the data, called conditional likelihood;
- $\hat{\theta}$ and $\hat{\mu}$ are the maximum likelihood estimators, that is the best estimate of $\theta$ and $\mu$ from the fit to the observed data and therefore maximise the unconditional likelihood.
- $\mathcal{L}$ is the likelihood defined in Eq. 4.6.

The profile likelihood test statistics has the advantage of being independent on the nuisance parameters. In order to estimate the compatibility to the data, the knowledge of the probability density functions for the test statistic for both the null and alternate hypotheses is required. These distributions are obtained by generating pseudo-experiment, in this analysis 5000, for each hypothesis from the likelihood function. The value of $\lambda$ varies between 0 and 1. An high value is an indication of compatibility of the data with the estimation of $\mu$. For convenience, the test statistic is reformulated as:

$$t_\mu = -2 \ln \lambda(\mu), \tag{4.16}$$

with high values of $t_\mu$ indicating low compatibility with the data. This is quantified by measuring the $p$-value:

$$p = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|H)dt_\mu, \tag{4.17}$$

where $f(t_\mu|H)$ is the probability density of the test statistics. Hence, the $p$-value represents the probability assuming an hypothesis $H$ of finding data which are equal or less compatible with $H$ than the observed data. An hypothesis is excluded if the $p$-value is below a certain threshold, which for the exclusion of the signal-plus-background hypothesis is usually set to 0.05. This value corresponds to 95% confidence level ($2\sigma$), when converted into a one-sided Gaussian significance. In this analysis the requirement is placed not on the $p$-value itself, but on the $CL_s$, defined as:

$$CL_s = \frac{CL_{s+b}}{CL_b} \equiv \frac{p_{s+b}}{1 - p_b}, \tag{4.18}$$

where $p_{s+b}$ is the $p$-value under the signal-plus-background hypothesis and $p_b$ under the background-only hypothesis. The use of the $CL_s$ prevents the exclusion of models where the experiment is not sensitive, hence when the two test statistics distributions are close to each other. The $CL_s$ method is illustrated in Figure 4.9a: the shaded area of the distribution for the test statistics under signal-plus-background hypothesis (purple) represents $CL_{s+b}$ and the shaded area of the distribution under background-only hypothesis (green) is $CL_b$.
To obtain the upper limit, the parameter of interest (POI) is scanned and the $CL_s$ is evaluated for each value from the data. The measured upper limit is the value of the POI for which $CL_s = 0.05$. This can be compared to the expected upper limit, which is obtained as the median of the test statistics distribution under the background-only hypothesis. An example of the expected upper limit is shown in Figure 4.9b as a function of the POI value. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ regions of the distribution.
This procedure is performed for every signal mass point in order to study the upper limit variation as a function of the mass of the heavy neutrino.

### 4.6.2 Unblinded result

The expected limit is first evaluated under the assumption of no signal. In order to keep the result blinded, the limit setting procedure is performed on a pseudoexperiment dataset, generated according to the neutrino mass fit model with no signal injected. After the analysis procedure had been established, the data in the signal region have been uncovered. The number of events observed in data in the signal region amounts to 4 and 1496 for same-sign and opposite-sign muons, respectively, and in both cases is in agreement with the background expectation within the assigned uncertainties.

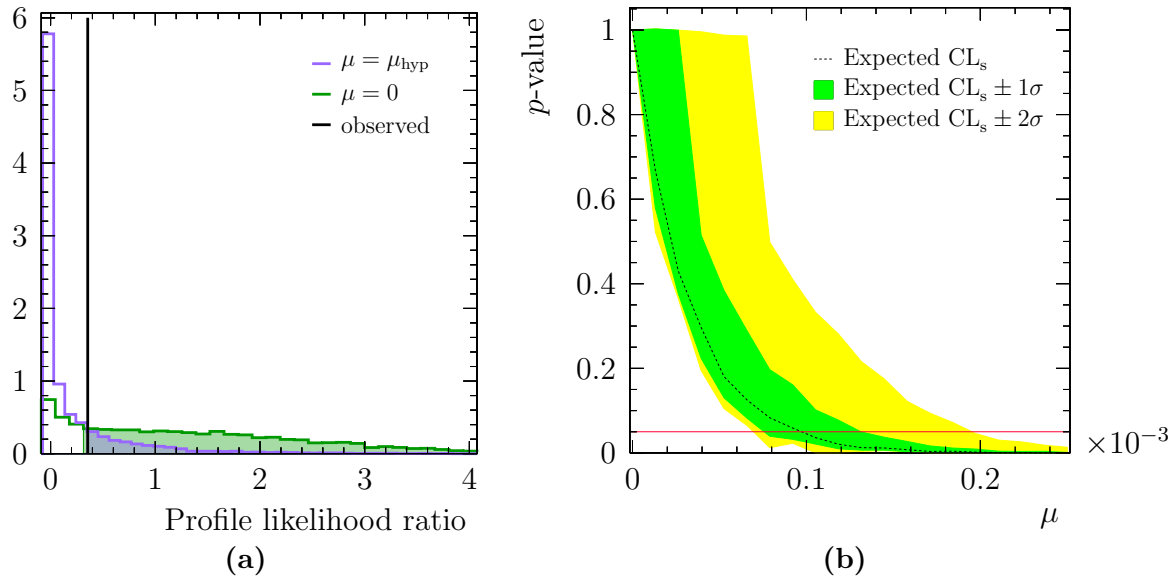



**Figure 4.9:** Illustration of the $CL_s$ method. In (a) distribution of the test statistics under background-only hypothesis ($\mu = 0$) and signal-plus-background hypothesis for a specific value of the parameter of interest $\mu_{\text{hyp}}$. The black line represent the observed value, the green shaded area $CL_b$ and the purple shaded area $CL_{s+b}$. In (b) the expected $CL_s$ is shown as a function of the POI value.

Fits are performed for the six mass hypotheses. The fits to the $\mu$jet mass in the signal region under a 15 GeV neutrino signal hypothesis are shown in Figure 4.10. No signal is observed and hence an upper limit is set on $\mathcal{B}(N \to \mu\text{jet})\,|V_{eN}|^2$ as a function of heavy neutrino mass. The upper limits are then scaled by $\mathcal{B}(N \to \mu\text{jet}) = 0.51$, computed from [77] assuming $|V_{eN}|^2 = |V_{\tau N}|^2 = 0$, leading to the results of Figure 4.11. Both the expected upper limit according to the background only hypothesis and measured upper limit are reported. For the same-sign muons sample, since the number of events is small, the expected limit trend follows closely the variation of the efficiency ratio with mass. The measured limits start to diverge from the expected one at 30 GeV neutrino mass and fall out the two standard deviation band for a 50 GeV neutrino mass. The worse limit measured with respect to the expectation can be attributed to the two data events at $M(\mu\text{jet}) = 31.5$ GeV and 39.1 GeV. The probability of two or more events in the mass range 28-60 GeV is 0.4% given an expected number of background events of 0.1. Nevertheless, since they are at high heavy neutrino mass, while all the backgrounds peak at low neutrino mass, a non-zero signal yield is obtained under the hypothesis of a 30 GeV or 50 GeV neutrino, as shown in Figure 4.12. For the opposite-sign muons samples the expected limit is a factor 5 to 10 worse due to the irreducible background from Drell-Yan process.

In conclusion, the heavy neutrino hypothesis is excluded at 95% CL in the mass range

5-50 GeV for coupling of the order of few times $10^{-4}$ and $10^{-3}$-$10^{-2}$ for same-sign and opposite-sign muons, respectively.

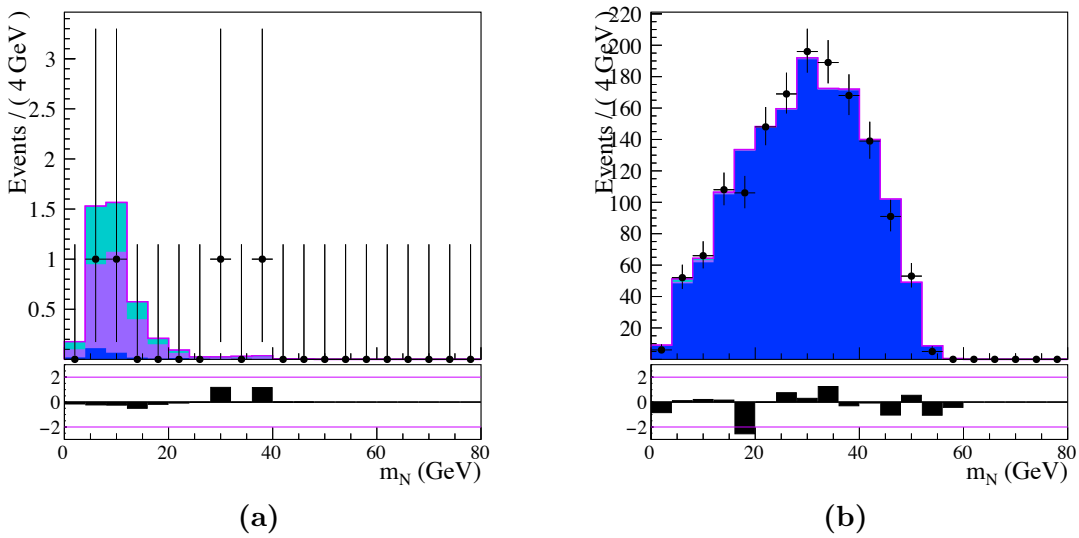



**Figure 4.10:** Fit to the $\mu$jet mass distribution for (a) same-sign and (b) opposite-sign muons. The signal component corresponds to a 15 GeV neutrino. Data (points) are compared to the fit result (magenta line). The fitted contributions are shown separately: $W \to \mu\nu$ (light blue), $Z \to \mu\mu$ (blue), qcd (purple) and signal (pink).

## 4.7 Outlook

The expected upper limit is extrapolated to the Run 1 and Run 2 dataset, corresponding to 9 fb$^{-1}$, and Runs 1 to 4 dataset (Upgrade I), corresponding to 50 fb$^{-1}$. For the same-sign muons case, two different hypotheses are considered:

- the observed background scales according to the luminosity $\mathcal{L}$ and hence the limit scales as $1/\sqrt{\mathcal{L}}$;
- the selection is optimised in order to reduce the background to a negligible level and hence the limit scales as $1/\mathcal{L}$.

Realistically, the reachable sensitivity would be in between and hence the extrapolated sensitivity is shown as a band. For the opposite-sign muons case, a significant amount of background is present and therefore the extrapolated limit is obtained according to the first hypothesis. The limit measured in this analysis and the extrapolated sensitivities can be observed in Figure 4.13 for both same-sign and opposite-sign muons. With an

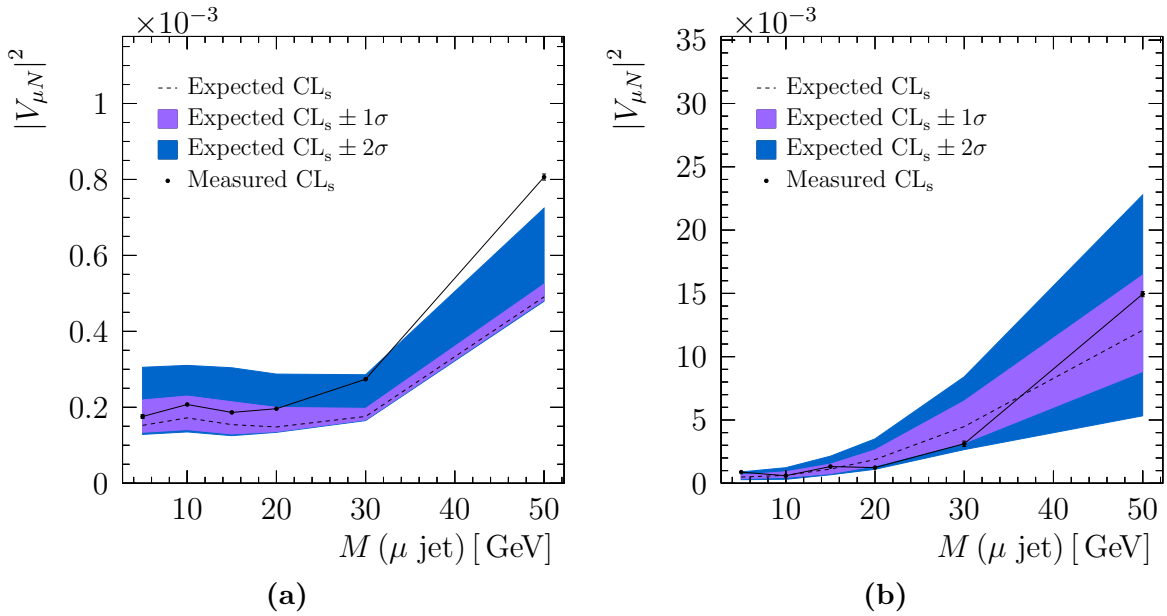
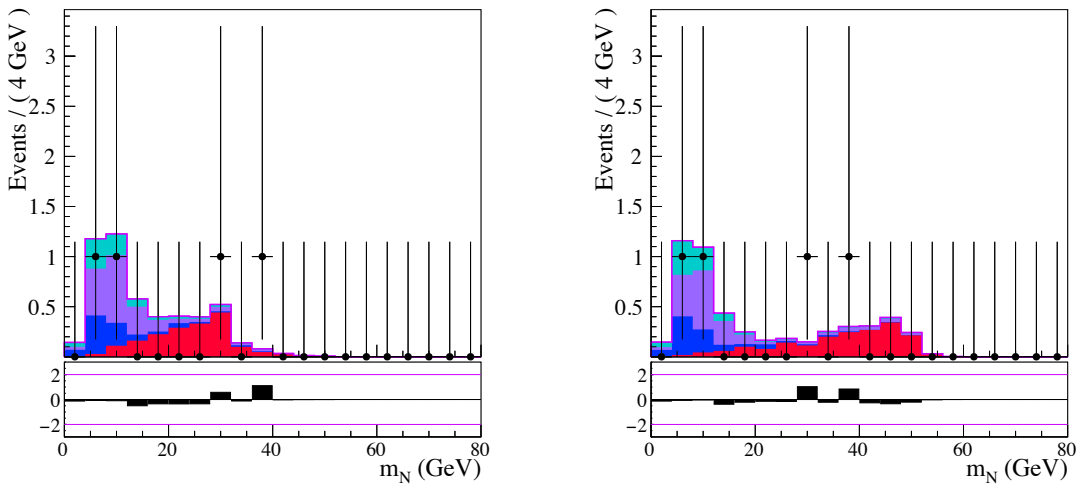


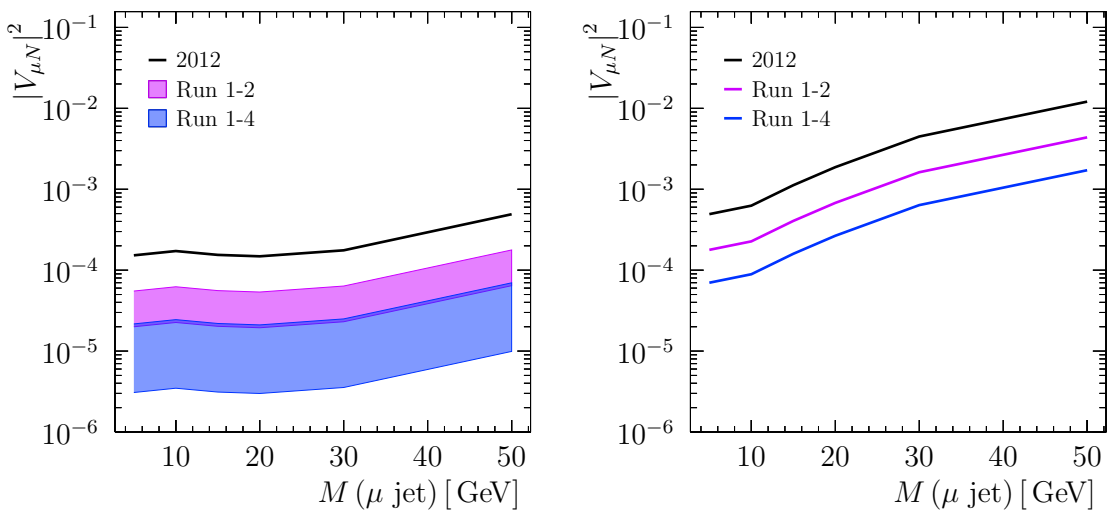

**Figure 4.11:** Expected (dashed line) and measured (filled line) limits on $|V_{\mu N}|^2$ at 95% confidence level for (a) the same-sign muons sample and (b) the opposite-sign muons sample.

integrated luminosity of $50\,\mathrm{fb}^{-1}$, a better sensitivity than the current most stringent limit could be reached in the mass range 5-50 GeV for the same-sign muons channel.




(a) (b)

**Figure 4.12:** Fit to the $\mu$jet mass distribution for same-sign muons. The signal component corresponds to (a) a 30 GeV and (b) a 50 GeV neutrino. Data (points) are compared to the fit result (magenta line). The fitted contributions are shown separately: $W \to \mu\nu$ (light blue), $Z \to \mu\mu$ (blue), qcd (purple) and signal (pink).




(a) (b)

**Figure 4.13:** Extrapolated limits on $|V_{\mu N}|^2$ at 95% confidence level for (a) the same-sign and (b) opposite-sign muons sample to Run 1 and 2 dataset (purple) and Run 1 to 4 dataset (blue). The upper limit obtained in this analysis using the data acquired during 2012 is also shown (black).

# Part II

# Silicon Sensors for the VELO Upgrade

# Chapter 5

# The VELO Upgrade

The VErtex LOcator (VELO) surrounds the proton-proton collision point and is dedicated to the reconstruction of primary and secondary vertices as well as part of the tracking system. In this Chapter the upgrade of the VELO is introduced. Firstly, the motivation for the upgrade and the main changes to the LHCb detector are summarised in Section 5.1. Secondly, the specific constraints on the VELO detector are outlined as well as their implications on the detector design in Section 5.2. Thirdly, the VELO layout, with focus on the main components, is described in Section 5.3. Finally, the estimated performance of the upgraded detector is evaluated and compared to the performance of the previous detector in Section 5.4.

## 5.1 LHCb Upgrade

The LHCb detector is currently being upgraded [119] during the long shutdown of the LHC in 2019-2020 in order to increase the precision on key observables that are now statistically limited, and to extend the reach of the detector to a new range of physics signatures. The instantaneous luminosity will increase to $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, a factor five higher with respect to Run I and II luminosity, and the experiment is expected to collect an integrated luminosity of 50 $\mathrm{fb}^{-1}$ by the end of Run IV (2030). The increase in instantaneous luminosity can be obtained with the current LHC machine, which has been designed to deliver an instantaneous luminosity up to $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at the general purpose detectors and up to 50% of that at the LHCb interaction point [120]. The LHC will only upgrade during LS3, which is foreseen for 2024-2026, to High Luminosity LHC. The operational conditions at the LHCb interaction point in Run I-II and at the upgrade are compared in Table 5.1.
With the Run I-II detector the yield recorded for hadronic channels would saturate around a luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, as shown in Figure 5.1. This is due to the hardware trigger, which reduces the data rate from 30 MHz, the event rate of the LHC, to a readout rate of 1.1 MHz. The benefits of the increased instantaneous luminos-

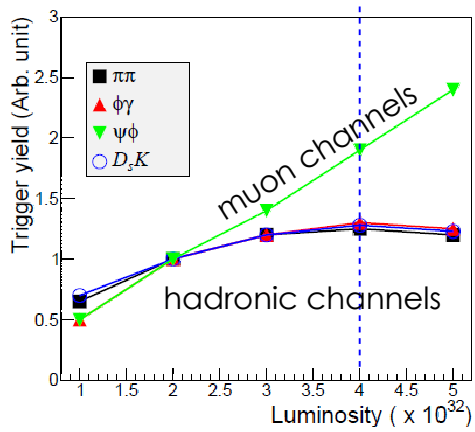



**Figure 5.1:** Trigger yield as a function of instantaneous luminosity for different decays of $B$ meson. The dashed line represents the instantaneous luminosity of Run I-II. From [121].

ity are limited for the L0 hadron trigger, in contrast to the L0 muon trigger, due to additional cuts on the number of interactions per event and on track multiplicity. It has therefore been decided to remove the hardware trigger and adopt a fully software-based trigger. The upgrade trigger strategy, summarised in Figure 5.2, is based on the real time reconstruction, alignment and calibration algorithms developed during Run II. Signal classification is performed at the trigger level employing the information from all subdetectors and limited or no further offline reconstruction is foreseen.

The main challenge brought by the removal of the hardware trigger, is that the data from each subdetector must be read out and sent directly to the CPU farm for each collision at a rate of 30 MHz. In addition to the higher rate, the increase in pile-up, track multiplicity and radiation damage represent challenging aspects for the detector upgrade. In order to comply with these conditions, the majority of the readout electronics and data acquisition system is upgraded and some subdetectors are fully replaced. The particle identification subdetectors largely retain the existing layout [123]. The RICH system, which provides particle identification of charged hadrons, maintains the

**Table 5.1:** Comparison between the operational conditions at the LHCb interaction point of Run I-II and the upgrade: $\mu$ is the average number of visible interaction per bunch crossing, $\sqrt{s}$ is the centre of mass energy, $\mathcal{L}$ is the instantaneous luminosity and $\int \mathcal{L}$ is the integrated luminosity.

| | $\mu$ | $\sqrt{s}$ [ TeV ] | $\mathcal{L}$ [$\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$] | $\int \mathcal{L}dt$ [$\mathrm{fb}^{-1}$ ] |
|---|---|---|---|---|
| Run I-II | 1.1 | 7,8,13 | $4 \times 10^{32}$ | 9 |
| Run III-IV | 5.5 | 14 | $2 \times 10^{33}$ | 50 |

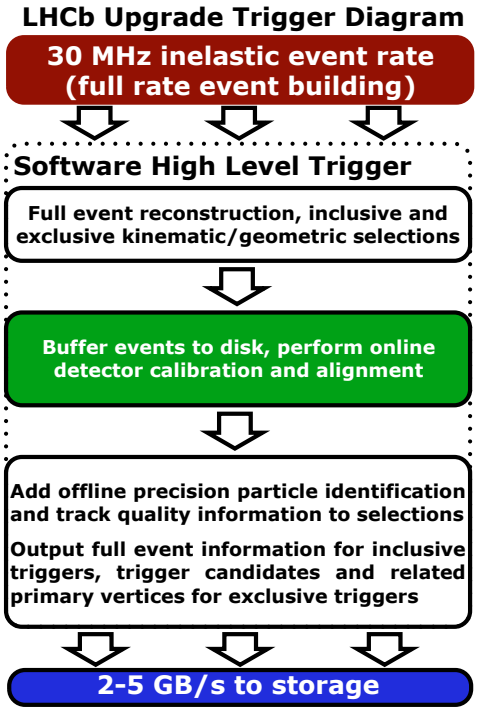



**Figure 5.2:** Trigger strategy for the LHCb upgrade from [122].

same overall structure. The main modification applied to RICH 1 consists of the removal of the silica areogel, leaving only $C_4F_{10}$ gas as radiator. With the resulting free space, the optical system can be replaced to increase the image area of the Cherenkov rings and therefore cope with the higher occupancy. The second major change involves the HPD photodetectors, which are replaced with multianode photomultipliers with external readout electronics.

The upgraded calorimeter system consists of two subdetectors, ECAL and HCAL. The scintillating pad detector and the preshower are removed, since their main use was within the hardware trigger. The modules of ECAL and HCAL remain unchanged, although eventually some of the inner cells of the ECAL might need to be replaced during LS3, due to radiation damage. The photomultipliers are retained with a reduction of gain to ensure longer lifetime under high-luminosity operation, which is compensated by a gain increase in the upgraded readout electronics. Among the five stations of the muon system, the first one, M1, which is placed in front of the calorimeters, is removed. Due to the high hit occupancy at upgrade luminosity, the association of M1 hits to the track segments from the muon stations would no longer be efficient. The innermost part of station M2 (the first muon station downstream the calorimeters) is expected to experience a too high flux of particles. Therefore an additional shielding is installed around the beam-pipe behind the HCAL. Also for the muon stations the readout electronics is replaced to comply with a 40 MHz readout.

In contrast to the particle identification system, all the tracking subdetectors are replaced [124]. The VELO silicon strip detector is replaced by a pixel detector and is described in detail in the following sections.
The tracking station before the magnet (TT) is replaced by the Upstream Tracker (UT). The UT employs the same technology as the TT, silicon strips, but with finer segmentation and larger coverage. The new station consists of four planes in the nominal LHCb stereo configuration x-u-v-x. The main element of the plane is the stave, a support structure which contains the cooling tube and hosts on the surface strip sensors and ASICs[1] alternated with flex cables carrying data lines and HV. In order to provide full coverage the sensors on the two sides of the stave are overlapping and the staves are staggered. Three different sensors geometries are used, with increasing granularity at shorter radial distance from the beam. The innermost sensors have a semicircular cutout to maximise the acceptance near the beam line. The UT is expected to accumulate a maximum integrated fluence of $5 \times 10^{14} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$. To mitigate the impact of radiation damage the sensors should be kept at a temperature of $-5°\mathrm{C}$.
The Outer Tracker (OT) and Inner Tracker (IT) are replaced by a single detector using scintillating fibres, the Scintillating Fibre tracker (SciFi). The SciFi consists of three stations of four planes each, in the same stereo configuration, each plane composed by densely packed scintillating fibres. The scintillation light is recorded by silicon photomultipliers (SiPM) at the end of the fibre and the signal is read out by a custom developed chip. The maximum integrated fluence at the position of the SiPMs, that are 2.5 m far from the beam axis, is estimated to be $10^{12} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$. To reduce the thermal noise of the SiPM after the exposure to such fluence, the SiPMs are cooled down to $-40°\mathrm{C}$ by 3D printed titanium cold bars.

## 5.2 VELO Constraints

The increase in luminosity, the 40 MHz readout and the necessity of maintaining or even improving the physics performance drive the design choices for the upgraded VELO. The main constraints, and their implications for the detector, are outlined below.

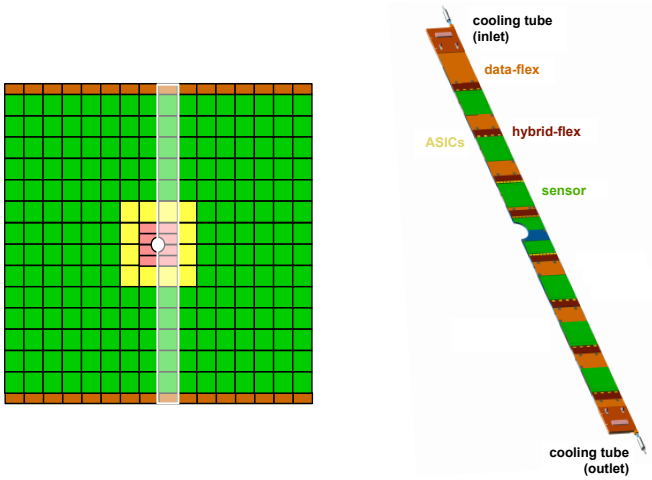
**Data Rate**

The increase in luminosity leads to an increase in particle flux and occupancy. The number of particle hits-per-event at the $2 \times 10^{33}\,\mathrm{cm^{-2}\,s^{-1}}$ luminosity varies as

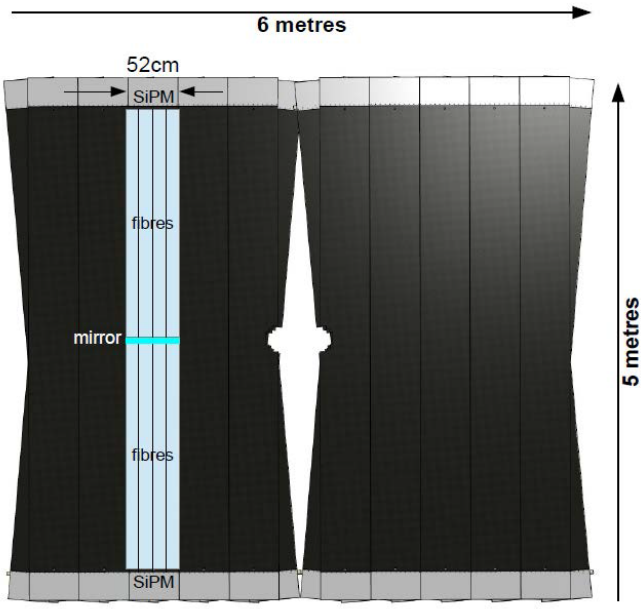
$$N \sim 5.2 \times r^{-1.9}, \tag{5.1}$$

where $r$ is the radius in cm [125]. The event repetition rate is expected to vary

[1] Application Specific Integrated Circuit




**(a)**




**(b)**

**Figure 5.3:** In (a) a UT plane consisting of vertical staves, where different colours indicate different sensor geometries, and illustration of a stave. In (b) SciFi front view of a module. From [124].

from an average of 27 MHz to a peak of 40 MHz, depending on the bunch filling scheme. The output data rate is 20 Gbytes/s for the ASICs located closest to the beam.

### Radiation

The upgraded VELO has to withstand an integrated luminosity of 50 $\mathrm{fb}^{-1}$. The radiation exposure is expected to be nonuniform; as shown in Figure 5.4 the fluence

depends on the distance from the interaction point along the $z$ direction and on the radial distance from the beam. At the end of lifetime, the innermost part of the sensor in a module close to the interaction point has been exposed to a fluence of $8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$.

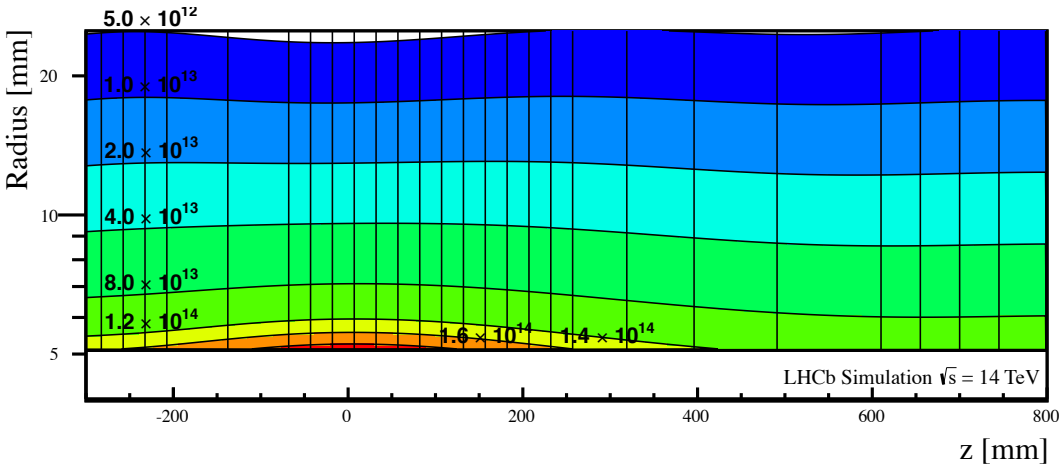



**Figure 5.4:** Estimated integrated fluence per $\mathrm{fb}^{-1}$ in upgraded conditions as a function of radial distance from the beam and position along the beam direction ($z$ axis). The fluence is in units of $1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$. From [125], based on a GEANT4 simulation.

### Physics Performance

The pattern recognition and track reconstruction must be fast and efficient. This requires hit efficiencies higher than 99% and an optimised layout of the detector which ensures that more than 99% of the tracks within the LHCb acceptance traverse at least four modules.

The main resolution figure-of-merit to evaluate the performance of the detector is the impact parameter (IP) resolution, that is the precision with which the secondary vertex particles can be separated from the primary vertex. The upgraded VELO is required to maintain and if possible improve the performance with respect to the previous VELO, with $\sigma(\mathrm{IP}_x) \sim 20\,\mu\mathrm{m}$ for tracks with high transverse momentum [126]. Due to the forward geometry of LHCb, the impact parameter resolution is determined by the resolution in the $xy$ plane, which can be approximated by the following expression [127]:

$$\begin{aligned}\sigma_{\mathrm{IP}}^2 &= \frac{r_1^2}{{p_{\mathrm{T}}}^2}\left(0.0136\ \mathrm{GeV}/c\sqrt{\frac{x}{X_0}}\left(1+0.038\ln\left(\frac{x}{X_0}\right)\right)\right)^2 + \frac{\Delta_{0n}^2\sigma_1^2+\Delta_{01}^2\sigma_n^2}{\Delta_{1n}^2}\\ &= \sigma_{\mathrm{MS}}^2+\sigma_{\mathrm{res}}^2,\end{aligned} \tag{5.2}$$

where $r_1$ is the radius of the first measured point on the track; $p_\mathrm{T}$ is the transverse momentum of the track; $x/X_0$ is the fraction of radiation length before the second measured point on the track; $\sigma_i$ the measurement errors on the measurement point $i$ and $\Delta_{ij}$ is the distance between the points $i$ and $j$, with 0 referring to the vertex, 1 to the first measured point and $n$ to the point measured at the $n$th plane downstream, with the assumption that only two planes contribute to the determination of the track parameters.
The IP resolution therefore consists of two terms: the multiple scattering term $\sigma_\mathrm{MS}$, which accounts for multiple scattering in the detector material and the RF foil, and the resolution term $\sigma_\mathrm{res}$, which accounts for the hit resolution. The multiple scattering term depends on the position of the RF foil, the radius of the first measured point and the amount of material before the second measured point. The resolution term depends on the single hit resolution and the distance between measured points. The second measurement plane depends on the multiple scattering in the detector: for low momentum tracks the average scattering angle is larger and mostly the first two measurement planes contribute to the IP resolution, while for high momentum tracks the scattering is smaller and more planes downstream contribute to the track measurement. Hence both the multiple scattering and resolution terms depend on the momentum of the particle. Both terms contribute for low $p_\mathrm{T}$ tracks, while the resolution term is dominant for high $p_\mathrm{T}$ tracks.

The increase in particle multiplicity and pattern recognition requirements lead to the decision to replace micro-strips with hybrid pixel detectors, to increase granularity and thereby reduce the occupancy per sensitive element. The term hybrid refers to the fact that sensor and electronics are fabricated separately and then bonded. In this way sensor and ASIC can be optimised independently. The change of technology from strips to hybrid pixel detectors has several advantages. Pixels have higher granularity and therefore cope better with the particle occupancy of the upgraded detector, while strips would not be able to assign the hit positions unambiguously. The pixel technology simplifies the track reconstruction algorithms, as the three dimensional position of the hit is immediately provided. Moreover, pixel detectors are more robust in a high radiation environment due to the smaller capacitance with respect to the strips and hence higher signal-to-noise ratio.
The high radiation dose is critical for both the sensor, due to the risk of thermal runaway at the sensor tip, and the ASIC, which is required to be radiation hard and single event upset [2] (SEU) tolerant. Moreover, the non uniformity of the radiation exposure across the sensor represents the most critical aspect for the sensor R&D, as discussed in Section 7.1. The ASIC, called VeloPix, has been custom developed for the VELO upgrade, designed to be radiation hard and to cope with high rate and occupancy. A

[2] Bit flip in the electronics induced by large deposit from ionising particles in a small volume of silicon.

new cooling system has been developed using a two phase $CO_2$ cooling and directly integrating microchannels to the silicon substrate of the module. The cooling must remove the power dissipated by the ASICs and keep the tiles at a temperature $< -20°C$ in order to minimise the effects of radiation damage and to protect the tip of the sensor from thermal runaway effect[3]. Important constraints on the detector geometry arise from Eq. 5.2. Since the impact parameter resolution is proportional to the radial distance to the first measure point of the track, the active area of the detector is placed at a closer distance to the beam with respect to the previous VELO. The sensor closest point to the beam is moved from the 8.2 mm of Run I-II to 5.1 mm, with an average distance of 5.9 mm. Another crucial aspect is the amount of material crossed by the particles before being detected. The major contribution to the material budget is due to the RF foil, since particles may pass through the foil multiple times and at large angles. Therefore, the RF foil is required to be thin, needs to be adapted to the new shape of the modules and is placed closer to the beam. Despite a thinner RF foil, the material budget of the upgraded VELO increases, with a total radiation length about 30% higher with respect to the previous detector. In terms of material each module of the upgraded VELO consists of 200 μm thick sensor, 200 μm thick ASIC, 20 μm thick spherical Tin-Lead (SnPb) bump bonds and a 500 μm silicon substrate, while the module of the previous VELO of 300 μm thick $R$ and 300 μm $\phi$ sensors.

## 5.3 VELO Layout

The layout of the detector planes of the upgraded detector is similar to the Run I-II layout. The upgraded VELO consists of two movable halves that can be retracted during beam injection. Each half consists of 26 modules, that are arranged perpendicularly to the beam. The $z$ position of the modules has been optimised to provide the best performance in terms of IP resolution and track reconstruction, based on a Geant4 detector simulation. The modules are separated from the primary vacuum by the RF foil. As shown in Figure 5.5, each module has two sensor tiles placed in a L shape on the front and two on the back, overlapping by 110 μm to provide full coverage over the LHCb acceptance. Each tile consists of a sensor bump bonded to three VeloPix ASICs. The ASICs are connected to low voltage (LV), control and data cables via a printed circuit board known as hybrid. The high voltage (HV) is delivered to the backplane of the sensor. In the module cross-section of Figure 5.6, it can be seen that tiles and hybrids are mounted on both sides of a substrate where the cooling microchannels are embedded. In the following, the main components of the upgraded detector are described.

[3]Thermal runaway is a consequence of bulk damage of the sensor due to high irradiation. The increase in leakage current leads to an increased power dissipation and heating of the sensor, which causes a further increase in leakage current and dissipated power. The consequence is a positive feedback if the sensor is not properly cooled.

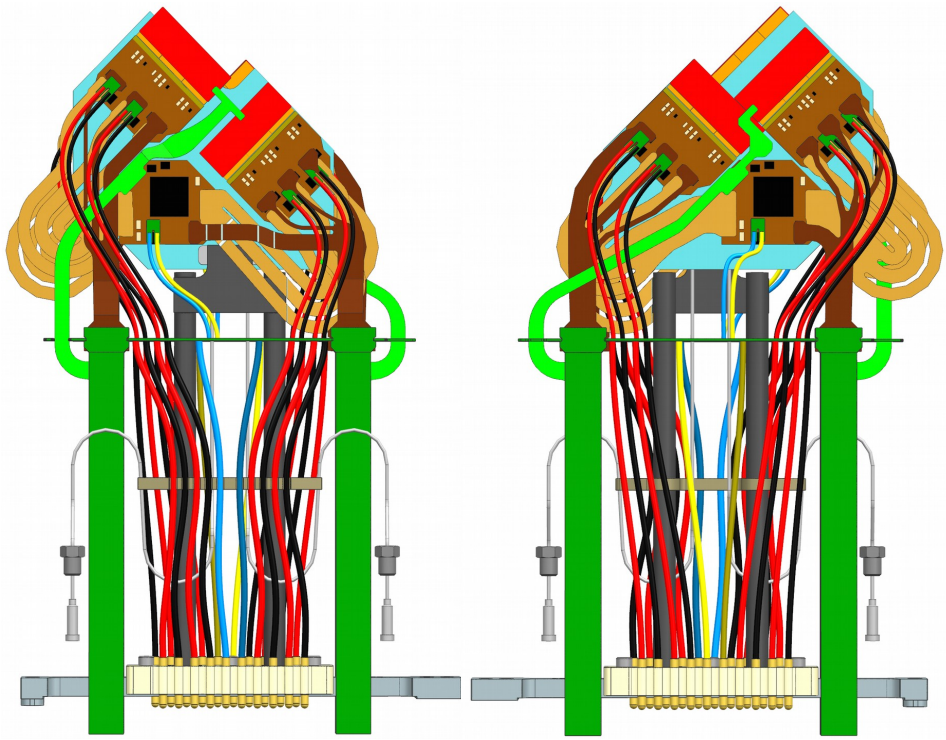


**Figure 5.5:** Front (left) and back (right) schematic view of a module. The red rectangles represent the tiles.

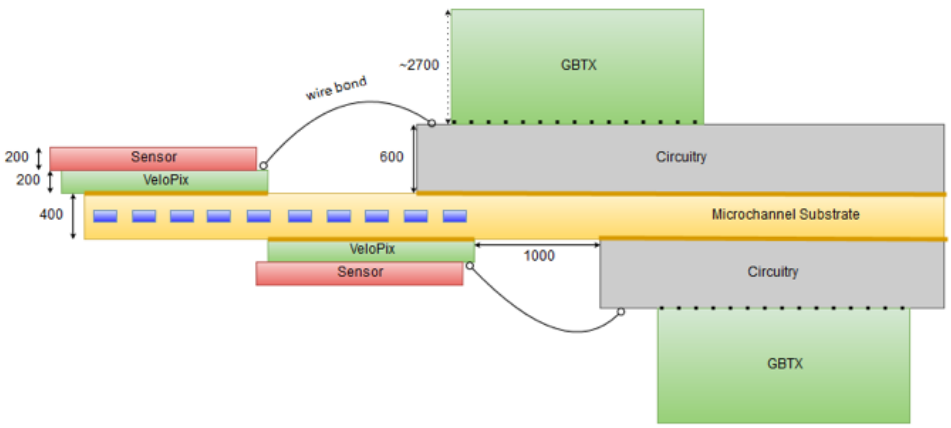



**Figure 5.6:** Module cross-section.

### 5.3.1 Sensors

The sensor consists of $768 \times 256$ pixels and is bump bonded to three ASICs. The pixel pitch is 55 µm, with the exception of the pixels covering the regions between ASICs that are elongated by half of the interchip distance, from 55 µm to 137.5 µm. Various sensors with different geometry and of different bulk type have been considered. The prototypes have been produced by two vendors, Hamamatsu (HPK) and Micron, and tested in a beam of minimum ionising particles in order to choose the best candidate for the VELO upgrade. Since the VeloPix ASIC has been developed in parallel to the sensor, the prototypes have been tested in combination with the Timepix3 ASIC [128]. Similar to the VeloPix, this chip belongs to the Medipix/Timepix family. The Timepix3 ASIC has also been used to build a dedicated beam telescope, called Timepix3 telescope, to test the prototype sensors for the VELO upgrade. The Timepix3 telescope, described in Section 7.3, is a high rate beam telescope with data-driven readout which provides a fast and robust pattern recognition and track reconstruction. The telescope has been operated in a 180 GeV mixed hadron beam (protons and pions) at the CERN SPS for several beam tests. The different prototypes tested, the requirements that they have to fulfil and the results of the testbeam program are presented in Chapter 7.

### 5.3.2 VeloPix

The VeloPix ASIC [129] has been custom developed for the LHCb upgrade, to be able to cope with the harsh radiation environment and the high data rate. The chip consists of $256 \times 256$ pixels of 55 µm pitch. The sensor is thinned down to a thickness of 200 µm to reduce the material traversed by the particles.
The chip must withstand up to 400 MRad over its lifetime. It is designed in 130 nm technology which proved to be radiation hard [130] and it is protected against single event upsets using a triple modular redundancy scheme [131], based on triplicated logic which gives as result the vote of the three outputs.
According to Eq. 5.1, the particle flux decreases quickly with the radial distance from the beam. Figure 5.7 shows the average number of tracks crossing each VeloPix per bunch crossing, which corresponds to more than 900 Mhits/ s/chip for the highest occupancy chips.
The chip is required to provide the time of arrival and the position ($xy$) of the hits. Since it is read out at the bunch crossing frequency of the LHC, the time of arrival needs to be measured with a precision of 25 ns, in order to assign the hits to the correct bunch crossing. To reduce the output bandwidth the readout is binary, data driven and zero suppressed: as soon as the hit is recorded by the pixel, it is timestamped and labelled with the pixel address and sent off chip. The bunch crossing identification number gives the particle timestamp. Another feature to reduce the output bandwidth consists of grouping $2 \times 4$ pixels in a so-called superpixel. This exploits the fact that the average

number of hits created by a track, referred to as cluster, is about 2.2 hits. Pixels hit within the same superpixel share the address of the superpixel and the bunch crossing identification number. Therefore, hits arriving at the same time in adjacent pixels are combined in the same data packet. The $xy$ position of the hit is provided combining the address of the superpixel with the pixel hit within the superpixel. Clusters crossing the boundary between superpixels are transmitted in different packets and are grouped together at a later stage in the processing chain.
The VeloPix is readout column-wise, with the column direction always directed away from the interaction, as indicated by the arrow in Figure 5.7. Each ASIC has four high-speed output links. All four links will be active for the innermost chips due to the high data-rate, while less links are needed for the outermost chips and hence less links are activated, reducing the power consumption. The VeloPix consumes up to 3 W per ASIC, which is a substantial amount of power that must be removed by the cooling system.

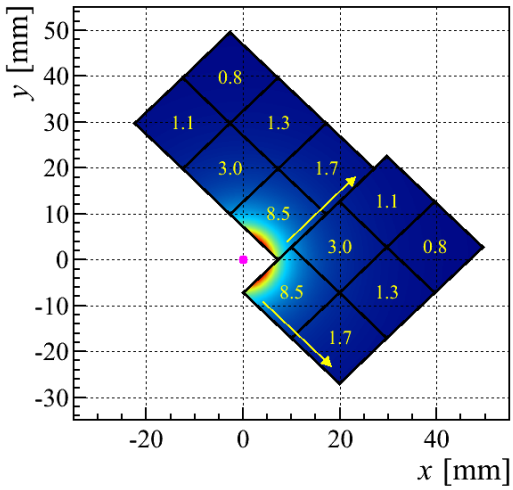



**Figure 5.7:** Mean number of particles crossing an ASIC per bunch crossing from [125]. The yellow arrows indicate the readout directions. The nominal position of the beam is indicated by the red dot.

### 5.3.3 Microchannel Cooling

The cooling for the upgraded VELO must satisfy the following requirements:

- The innermost region of the sensor is exposed to the highest fluence and therefore has the highest leakage current due to the radiation damage (see Section 6.2). To avoid thermal runaway, the innermost region of the sensor must be kept at a maximum temperature of $-20°$C.

- The cooling is required to dissipate a total power of 43 W per module. The 12 ASICs contribute most of the power. Additional 5 W are generated by the hybrid and up to 2 W can be produced by the sensors at the end of lifetime.
- The contribution to the material budget should be minimal.
- The cooling system must be leak tight since the modules are placed in vacuum.

The cooling solution that is adopted consists of bi-phase $CO_2$ circulating in microchannels etched into the silicon substrate [132]. The evaporative cooling [81], already employed in the previous VELO, is a process which exploits the phase transition of the coolant from liquid to vapour. The $CO_2$ is brought at the boiling point in the detector, such that it evaporates while absorbing the heat from the modules. Therefore the coolant is in two different phases, liquid and vapour. Since the process is isothermal, a low variation of the temperature across the module is expected. The choice of $CO_2$ as a coolant is made taking into account its radiation hardness, low viscosity which makes it easily circulate, and high latent heat. The coolant circulates in the substrate underneath the ASICs, minimising the distance between the heat source and the coolant. In this way the material in the acceptance is also reduced, since the substrate serves also as mechanical support for ASICs, sensors and hybrids. Moreover, mechanical distortions of the module, due to mismatch between the expansion coefficients of the different components, are avoided by choosing silicon as material for the substrate as well as detection medium.

### 5.3.4 RF Box

The geometry of the RF foil needs to be adapted to fit the L-shaped modules and to account for the different detector layout along the $z$ axis. In order to reduce the material budget, the thickness of the box is required to be $\lesssim 250\,\mu\text{m}$ of aluminium. In addition, the box must be leak tight. The technique adopted to shape the box is different with respect to that of the previous VELO, which was pressed into shape. The upgraded box is obtained by milling a single homogeneous block of aluminium. The maximum thickness of the part of the foil facing the beam is $250\,\mu\text{m}$. The possibility of thinning the foil further via chemical etching is currently evaluated.

## 5.4 Simulated Performance

A few key features are compared to evaluate the performance of the upgraded detector with respect to that of the previous VELO. The results are obtained using samples of minimum bias events, simulated and reconstructed with the LHCb software [133], at an average number of interactions per bunch crossing of $\nu = 7.6$. Two fundamental figures

of merit are the impact parameter resolution and the track reconstruction efficiency. The IP resolution for the upgraded and previous VELO in upgraded conditions are compared in Figure 5.8a. Several factors affect the difference in IP resolution in comparison to the previous VELO:

- a shorter distance between the beam and the first active element;
- a higher radiation length of a factor 1.3;
- a smaller distance between the modules;
- a coarser pitch at small radius: pixels have a constant pitch of 55 μm, while at low radius the strips have a smaller pitch of 40 μm.

Despite the larger material budget and the coarser pitch, the IP resolution of the upgraded VELO improves significantly. This is mainly due to the smaller distance from the beam of the first measured point, which reduces both the multiple scattering and resolution terms according to Eq. 5.2. For high momentum tracks, where the error on the IP is dominated by the resolution term, which is expected to degrade due to the coarser pitch, the two factors approximately cancel resulting in similar performances between the two detectors.
The tracking efficiency is evaluated by comparing the number of correctly reconstructed tracks to the number of reconstructible tracks according to the simulation. In Figure 5.8b the tracking efficiency is shown as a function of the azimuthal angle $\phi$. The loss in efficiency for the previous VELO at $\phi = \pm 90°$ can be attributed to different effects. Firstly, the inner radius of the sensors is not constant, but it increases at $\phi = \pm 90°$, leading to a loss in efficiency for tracks at high pseudorapidity. Secondly, the overlap of $R$ and $\phi$ sensors in some regions is not perfect. This is compounded by the fact that the VELO has been operated with a distance between the two halves 200 μm larger than designed. The efficiency for the upgraded VELO improves significantly, while reducing the ghost rate by more than 50% and 90% with respect to the previous VELO in Run I-II and upgraded conditions, respectively. The tracking efficiency is also more uniform, despite the square geometry of the tiles. Furthermore, the layout of the VELO has been optimised to mitigate the efficiency loss if the gap between the modules is larger than expected.

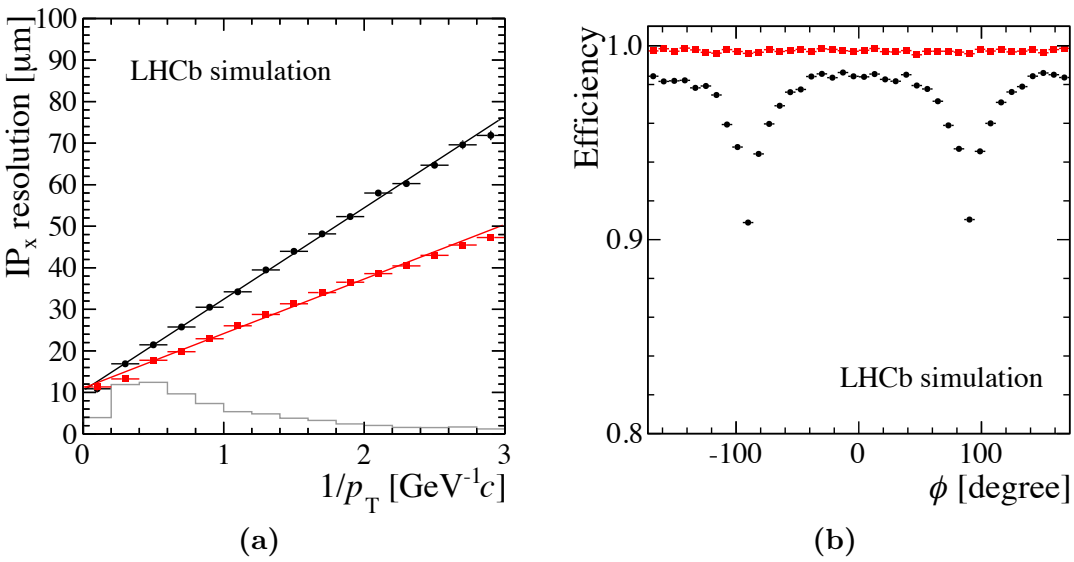



**Figure 5.8:** Impact parameter $x$ resolution as a function of inverse $p_{\mathrm{T}}$ (a) and track reconstruction efficiency as a function of azimuthal angle $\phi$ (b). The black circles represent the Run I-II VELO and the red circles the upgraded VELO. The performance of both upgraded and previous VELO are evaluated at $\nu = 7.6$ and $\sqrt{s} = 14\,\mathrm{TeV}$. The light grey histogram represents the $1/p_{\mathrm{T}}$ spectrum of $b$-hadron decay products. From [125].

# Chapter 6

# Silicon Sensors

In this thesis silicon pixel detectors have been studied in view of the upgrade of the LHCb VELO detector. The main working principles of silicon detectors are introduced in Section 6.1. The effect of radiation damage on silicon sensors, which is crucial for the design and operation of the detector in an environment such as the LHC, is described in Section 6.2. The different sensor types and their main characteristics are presented in Section 6.3.

## 6.1 Sensor concepts

Particle detection exploits the energy deposited by the particle due to the interaction with the material, which partly leads to ionisation. Silicon pixel devices are designed such that the electron-hole pairs, generated by ionisation, drift under the effect of an electric field and induce an electric signal which is subsequently recorded by a dedicated readout chip.

### 6.1.1 Interaction of particles with matter

Highly energetic particles traversing a medium maintain an almost constant velocity across and deposit energy mainly via single electromagnetic interactions with the electrons of the material, leading to ionisation along their path.
The distribution of the energy deposited, known as straggling function, is shown in Figure 6.1 for 500 MeV pions in silicon of different thicknesses. The straggling function is asymmetric, with a tail towards higher energies due to highly energetic single collisions. These are mainly due to highly energetic electrons, known as $\delta$ or knock-on electrons, which themselves contribute to the ionisation. Consequently, the mean energy loss rate for thin sensors is almost double the most probable value $\Delta_p/x$ of the distribution, where $\Delta/x$ is the energy loss per unit thickness. The mean energy loss rate is predicted by the Bethe-Bloch [32] formula, which estimates the average loss per unit length of

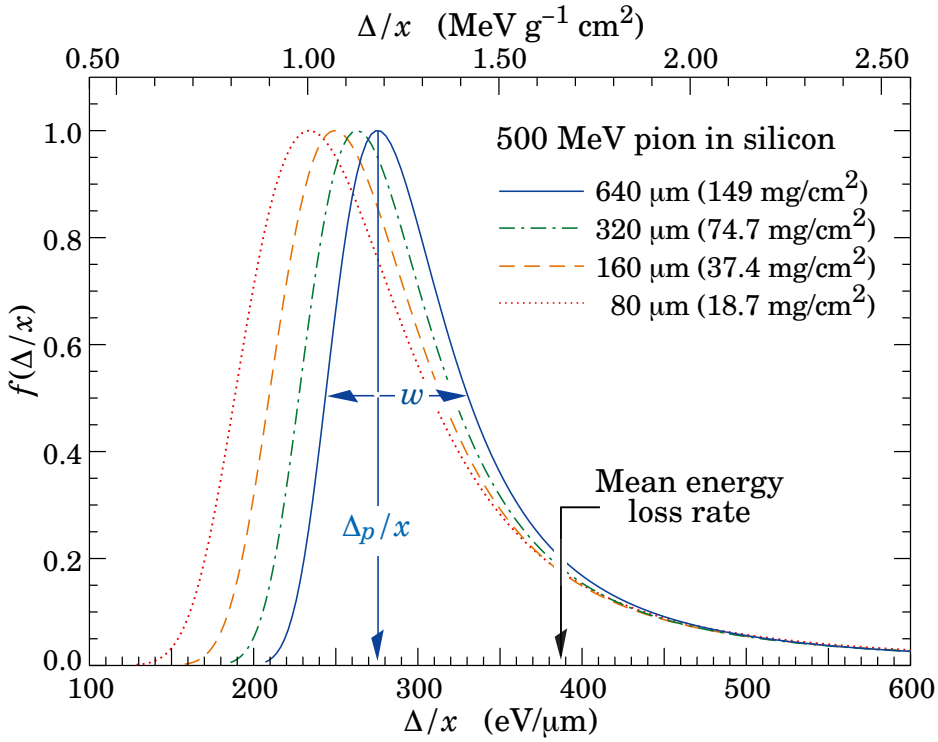



**Figure 6.1:** Energy loss in silicon per unit thickness for 500 MeV pions. $\Delta_p/x$ scales logarithmically with the silicon thickness. From [32].

highly energetic particles within a few percent accuracy. Both the most probable value (MPV) and the width of the distribution can be estimated for a thin detector[1] with the Landau-Vavilov function[2] [134]:

$$f_L(x, \Delta) = \frac{\varphi(\lambda)}{\xi}, \tag{6.1}$$

with

$$\varphi(\lambda) = \frac{1}{\pi} \int_0^\infty e^{-\pi y/2} \cos(y \ln y + \lambda y) dy, \tag{6.2}$$

and

$$\lambda = \frac{\Delta - \langle \Delta \rangle}{\xi} - \beta^2 - \ln(k) - 1 + C_E. \tag{6.3}$$

The function has been obtained by solving an integral transport equation with the Laplace transform method and in the approximation of free electrons [135]. The parameter $k$ is defined as $k = \langle \Delta \rangle / E_{max}$, where $\Delta$ is the energy loss, $E_{max}$ the maximum

[1] A detector is considered thin if the fraction of energy loss in the detector is negligible.

[2] The difference between the Landau and Vavilov formulations lies in the region of applicability, defined in terms of $k$.

energy transferable in a single collision and $C_E = 0.5772$ the Euler constant. The term $\xi = (K/2)(Z/A)z^2(x/\beta^2)\rho$ depends on the charge $z$ of the incident particle in units of electron charge, the atomic number $Z$ and the atomic mass $A$ of the material, the density $\rho$, the velocity in units of speed of light $\beta$ and $K = 0.307075\,\mathrm{MeV\,cm^2}$. The most probable value of the distribution results in [32]:

$$\Delta_p = \xi \left[ \ln \frac{2m_e c^2 \beta^2 \gamma^2}{I} + \ln \frac{\xi}{I} + 0.2 - \beta^2 - \delta \right], \tag{6.4}$$

where $I$ is the mean excitation energy in eV and the $\delta$ contribution, included at a later stage by Bichsel [136], accounts for density effect corrections to the ionising energy loss. While the mean energy loss per unit thickness described by the Bethe-Bloch formula does not depend on the detector thickness $x$, the most probable value of the Landau-Vavilov distribution scales logarithmically with the thickness. The distributions for thinner detectors in Figure 6.1 have a significantly larger width with respect to the one predicted by Eq. 6.1. In order to explain this, atomic electron binding energy effects, not taken into account in the Landau-Vavilov formulation, need to be introduced. This results in a convolution of the Landau-Vavilov function with a Gaussian:

$$f(x, \Delta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} f_L(x, \Delta - \tau) e^{-\frac{\tau^2}{2\sigma^2}} d\tau, \tag{6.5}$$

where $\sigma = \sqrt{\sigma_{res}^2 + \sigma_{noise}^2 + x\delta_2}$ accounts for both the detector resolution $\sigma_{res}$ and the electronic noise $\sigma_{noise}$ as well as the electron binding energy $\delta_2$ [135]:

$$\delta_2 = \frac{8}{3}\frac{\xi}{x} \sum_i I f_i \ln \frac{2m_e c^2 \beta^2}{I_i}, \tag{6.6}$$

where $f_i$ is the fraction of electrons in shell $i$. The convolution leads to a larger width and a small shift of the MPV towards higher energies. With decreasing material thickness the number of interactions decreases and $\delta_2$ becomes the dominant term, driving the width of the charge distribution.

The MPV of the convolution of the Gaussian and Landau distributions is used to describe the energy deposited in thin silicon detectors such as the sensors investigated in this thesis.

An average energy of 3.6 eV is needed to liberate an electron-hole pair in silicon [137]. The subsequent transport of charge carriers proceeds via diffusion and drift. The former consists of the random movement of the free charge carriers with a kinetic energy proportional to the temperature. The latter occurs in the presence of an electric field, which moves the carriers along the field lines.

### 6.1.2 The p-n junction

The basis of silicon sensors is the p-n junction. A p-n junction is formed between silicon doped with donor atoms that have one valence electron more than silicon atoms (n-material), and silicon doped with acceptor atoms that have one valence electron less than silicon atoms (p-material), when the n- and p-material are brought into contact. Due to the difference in doping concentration at the junction, holes diffuse towards the n-side and electrons diffuse towards the p-side, where they recombine. This leads to an area around the junction interface empty of free charge carriers, known as depleted region. This region is also referred to as space charge region, since a region in the p-material is negatively charged and a region in the n-material is positively charged. A so called built-in potential arises and consequently an electric field counteracting the diffusion. Once equilibrium between the electric field and diffusion is reached, the diffusion stops. In order to further increase the depleted volume, the junction is externally biased with a reverse bias voltage, i.e. the voltage is applied with the same polarity as the built-in voltage. The total depleted depth, which is given by the sum of the depleted depth in the n-side and p-side, depends on the voltage $V$ and doping concentration according to [137]:

$$W = w_n + w_p = \sqrt{\frac{2\epsilon_0\epsilon_{Si}(V_{bi}+V)}{e}\left(\frac{1}{N_a}+\frac{1}{N_d}\right)} \tag{6.7}$$

where $\epsilon_0$ is the permittivity of free space, $\epsilon_{Si}$ the relative permittivity of silicon, $V_{bi}$ the built-in voltage of about 0.6 V, $V$ the externally applied voltage and $N_a$ and $N_d$ are the concentrations of acceptors and donors, respectively. Silicon pixel sensors consist of a lightly doped bulk with highly doped segmented implants at the chip side and a highly doped implant at the backside. Given that the built-in voltage is negligible compared to the typical operation voltage $V$ and that the implants have typical doping levels of the order of $10^{17}$-$10^{18}\,\mathrm{cm}^{-3}$, Eq. 6.7 can be approximated as:

$$W \approx \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}V}{eN_a}}, \tag{6.8}$$

for a sensor with highly doped $n$-type implants on a lightly doped $p$-type bulk ($N_d \gg N_a$, n-on-p). In silicon pixel detectors one of the implants is segmented, as illustrated in Figure 6.2. The sketch shows the case of a partially depleted sensor. The depleted volume starts from the junction at the $n^+$ electrodes and extends into the bulk of the sensor. Charges liberated in the depleted bulk drift under the effect of the electric field, while charges liberated in the non-depleted volume diffuse. Ideally, the sensor would always be operated fully depleted, such that all the charges liberated in the bulk of the detector drift under the effect of the electric field and contribute to the signal, as described in Section 6.1.3. The voltage at which the sensor is fully depleted, with $W$ equal to the full thickness of the device, is called depletion voltage. For a uniformly

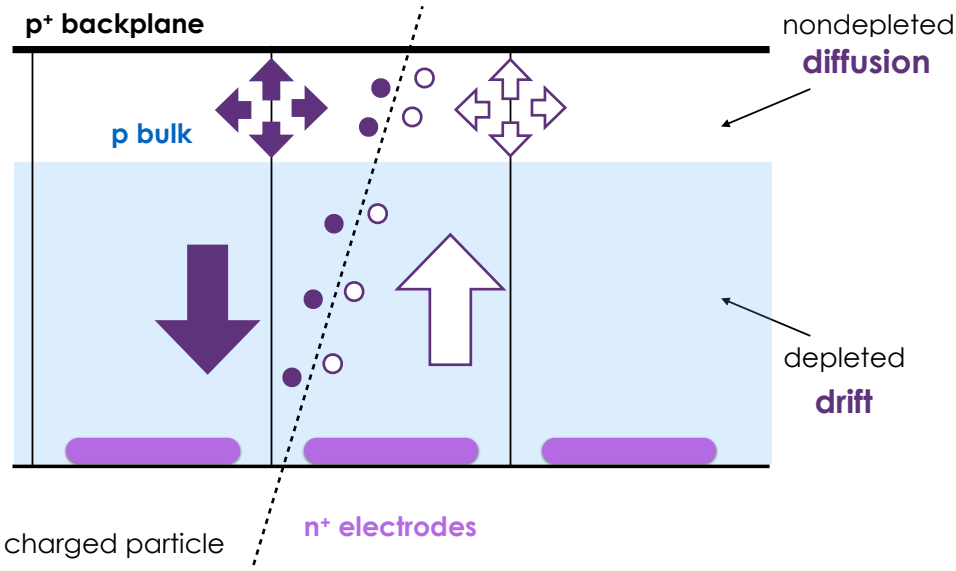



**Figure 6.2:** Sketch of a partially depleted n-on-p sensor. A charged particle traversing the detector is represented by a dashed line. The black and white circles represent electron-hole pairs liberated along the trajectory of the particle.

doped bulk the electric field due to the space charge has a linear dependence on depth: if the sensor is partially depleted, the electric field is maximum at the junction and zero at the border of the depleted volume, while if the sensor is overdepleted there is an additional constant field. Until saturation, the velocity $\vec{v}$ of the carriers is proportional to the electric field $\vec{E}$ according to:

$$\vec{v} = \mu \vec{E}, \tag{6.9}$$

where $\mu$ the mobility of the carriers, which differs between electrons and holes. Electrons have an almost three times higher mobility than holes. Electrons and holes move in opposite directions, but induce current of the same sign in the electrodes because they have opposite charges.

Overdepleting the sensor brings the advantage of a faster charge collection time, but degrades the spatial resolution because of the smaller lateral diffusion [138]. The limit on the operating voltage is imposed by electrical breakdown. Charge carriers are not only liberated by ionising particles, but are also generated thermally. The current induced by the latter, the so called leakage current or dark current, is therefore present even when there are no impinging particles. The leakage current depends exponentially on temperature [139]:

$$I \propto T^2 e^{\left(-\frac{E_g}{2kT}\right)}, \tag{6.10}$$

where $E_g$ is the bandgap energy and $k$ the Boltzmann constant. Figure 6.3 shows the typical trend of current as a function of voltage (IV curve). Initially the current increases

with voltage due to the increase of the depleted volume, reaching a plateau when the sensor is fully depleted. The current does not further increase until the breakdown voltage is reached. Then the high field at the junction results in an avalanche, that is, the electrons are sufficiently energetic to liberate secondary electrons in cascade, causing a rapid increase in current which could permanently damage the sensor. Avalanche breakdown can also occur at lower voltages in specific regions of the sensor, like at the edge or near the implants, unless specific structures are introduced to prevent this as described in Section 6.3.2.

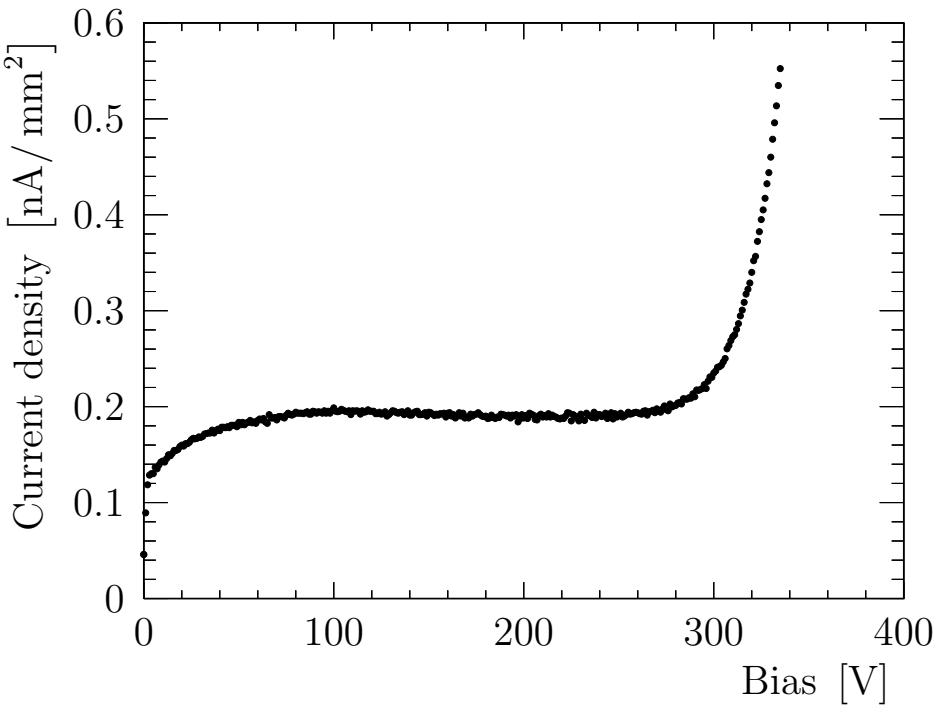



**Figure 6.3:** Typical current trend as a function of voltage.

### 6.1.3 Signal formation

Electrons and holes drift due to the electric field, inducing a current at the electrodes as soon as the carriers start to move in the sensor. The induction in a segmented detector such as a pixel sensor is described by Ramo's theorem [140]:

$$i = e\vec{v} \cdot \vec{E_w}, \tag{6.11}$$

where $\vec{v}$ is the velocity of the electron or hole under the electric field and $\vec{E_w}$ is the weighting field. While the electric fields determines the trajectory of charge carriers, the weighting field accounts for the geometry of the detector, describing the coupling of

the charge at a given position to a specific electrode. This weighting field is obtained by applying to this electrode unit potential and to the neighbouring electrodes zero potential [139]. The shape of the weighting potential can be observed in Figure 6.4: most of the signal is induced in the last part of the drift path towards, or away from, the pixel electrodes. If the integration time is larger than the collection time, the integral of the induced current corresponds to the charge liberated by the particle, assuming no trapping or recombination.

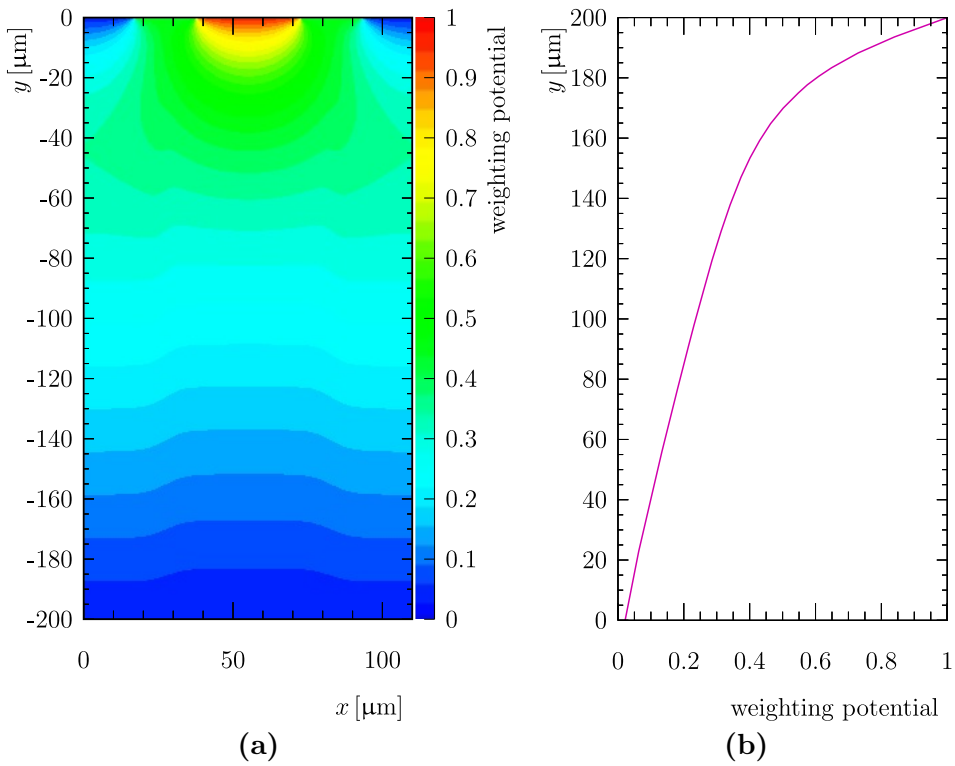



**Figure 6.4:** Weighting potential for a 200 µm thick sensor with 35 µm implant width simulated with TCAD (Technology Computer Assisted Design [141]): in (a) as a function of position in the pixel and in (b) as a function of depth in the middle of the pixel.

## 6.2 Radiation damage

Both the surface and the bulk of the sensor are sensitive to radiation damage. Surface damage consists of defects at the silicon surface and in the insulating layers on the surface of the sensor introduced for protection purposes. While surface damage does not have direct consequences for the operation of the detector, the bulk damage represents a limiting factor and is therefore described in more detail in the following.

### 6.2.1 Microscopic level

Most of the energy released by the particle is due to the interaction with atomic electrons and results in ionisation, but a small fraction is lost due to the interaction with the nuclei and determines the displacement of silicon atoms from their lattice sites. The fraction of this non-ionising energy loss (NIEL) depends on the energy of the particle. In order to compare damage induced by different particles of different energy, 1 MeV neutrons are used as a reference. According to the NIEL scaling hypothesis [142] the damage is quantified in terms of the cross-section for atomic displacement, with $1\,\mathrm{MeV}\,\mathrm{n_{eq}} = 95\,\mathrm{MeV\,mb}$ as reference value.
The NIEL does not account for how the damage is distributed. Primary point-like defects are vacancies and interstitials, schematically shown in Figure 6.5. Depending on the energy of the impinging particle, the recoil silicon atoms can be sufficiently energetic to induce defects themselves, generating clusters of defects usually in a localised area. The energy threshold to displace a silicon atom and produce a vacancy-interstitial pair is about 25 eV, while at least about 5 keV are needed to generate a cluster of defects [143]. The energy required for protons and neutrons to produce the same damage as electrons is smaller due to their higher mass. Furthermore, electrically charged particles interact mainly via the electromagnetic force with atoms screened by electron clouds generating mainly point-like defects and less clusters than neutral particles, that only interact via nuclear interactions [137]. Vacancies and interstitial can also combine among themselves

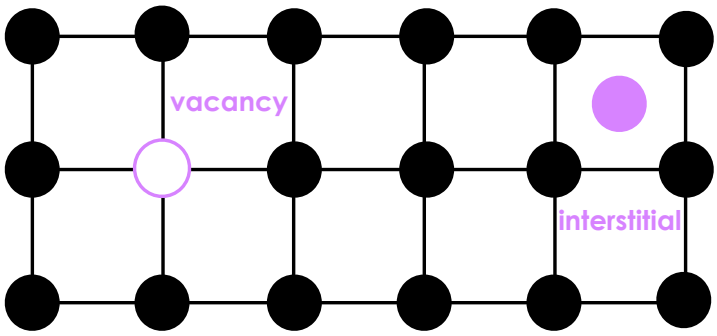



**Figure 6.5:** Sketch of radiation induced defects in silicon: vacancy on the left and interstitial on the right.

or with impurity atoms already present in the crystal. While vacancies and interstitial defects can move through the lattice at room temperature, clusters and complex combinations of them are usually stable. The mobility of the defects strongly depends on temperature: the higher the temperature, the higher the migration velocity.
Clusters and point like defects contribute differently to sensor damage. Hence, not all the macroscopic effects observed can be described with the NIEL scaling hypothesis.

### 6.2.2 Macroscopic level

Defects affect the properties of silicon. The main effects observed at the macroscopic level, the increase of leakage current, change in the doping concentration and trapping of the charge carriers, have a direct impact on the operation of the detectors. Furthermore, the changes in the space charge distribution introduced by radiation affect the electric field in the bulk.

**Change in doping concentration**

Before irradiation the concentration of donors and acceptors is dominant at the respective sides of the junction. After irradiation electrically charged defects can act as donors or acceptors modifying the effective doping concentration and the distribution of space charge. Most of the radiation defects in silicon are acceptor-like, leading to an increase of the negative space charge. The effective doping concentration is defined as the difference between the donor and acceptor like states. The full depletion voltage $V_{dep}$ changes accordingly [137]:

$$V_{dep} = \frac{e\,|N_{eff}|\,d^2}{2\epsilon_{Si}\epsilon_0}, \tag{6.12}$$

where $N_{eff}$ is the effective doping concentration and a uniform space charge is assumed. Figure 6.6 shows the fluence dependence of the effective doping concentration $N_{eff}$ and of the full depletion voltage $U_{\mathrm{dep}}$ for a sensor with an n-type bulk. The effective doping concentration decreases due to donor removal and the increase of acceptors up to a fluence $\Phi_{\mathrm{eq}}$ of $2 \times 10^{12} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ where it reaches a minimum. The minimum corresponds to the same number of donors and acceptors and therefore null space charge. For higher fluences the number of acceptors is dominant and increases with fluence. Therefore the silicon undergoes the so called type or space charge inversion, from an n-type bulk to a p-type bulk. Since the full depletion voltage is proportional to the effective doping concentration, it increases with fluence for both inverted n-type silicon and p-type silicon. The change in doping concentration depends on material and particle type and hence does not scale according to the NIEL hypothesis.

**Leakage current**

The presence of defects acting as carrier generation-recombination centres causes an increase of leakage current. The increase of leakage current $I_{\mathrm{vol}}$ is proportional to the fluence and independent of sensor type or impurity concentration of the silicon [144]:

$$\frac{I_{\mathrm{vol}}}{V} = \frac{I_{\mathrm{vol},\Phi=0}}{V} + \alpha\Phi, \tag{6.13}$$

where $I_{\mathrm{vol},\Phi=0}$ is the current before irradiation, $V$ is the volume and $\alpha$ is the current

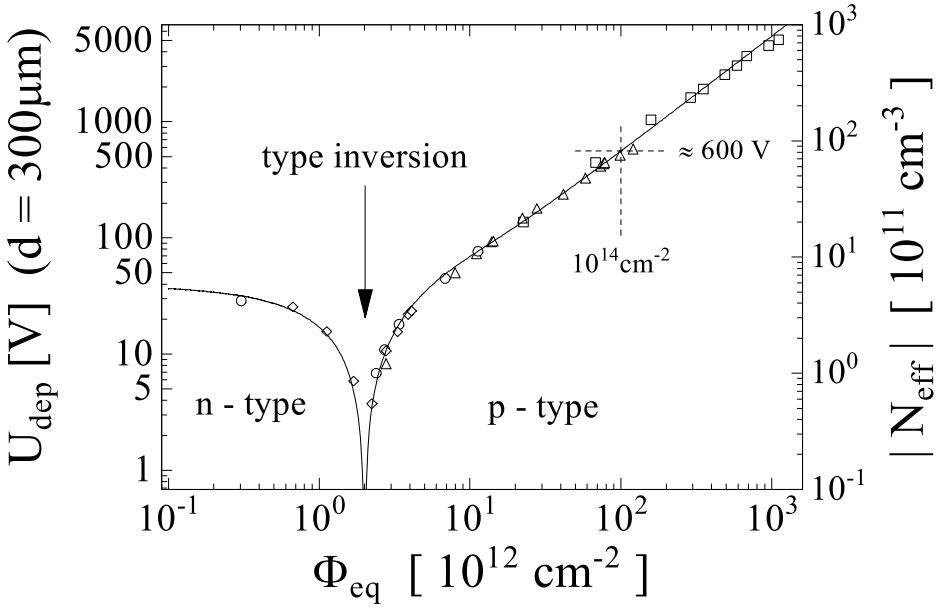



**Figure 6.6:** Change in effective concentration and depletion voltage as a function of fluence $\Phi_{\mathrm{eq}}$ for an n-type bulk 300 μm thick silicon sensor [144].

related damage rate, which depends on particle type and fluence $\Phi$. The leakage current scales according to the NIEL hypothesis.

**Trapping**

Defects in silicon can act as trapping centres and capture signal charges. Eventually charges will be released due to thermal excitation or an external field, but if the charge is held for a time longer than the integration time of the front-end electronics signal it does not contribute to the measured charged, leading to a degradation of the charge collection efficiency. The trapping time $\tau_t$ is inversely proportional to the concentration of trapping centres and therefore to fluence:

$$\frac{1}{\tau_t(\Phi)} = \frac{1}{\tau_{t,\Phi=0}} + \gamma\Phi \tag{6.14}$$

where $\gamma$ depends on particle type, temperature and time after irradiation. A shorter charge collection time reduces the probability to get trapped: hence electrons are less likely to be trapped than holes given their higher mobility. The inverse trapping time is shown in Figure 6.7 separately for electrons and holes. The trapping probability does not scale according to the NIEL hypothesis: it has been observed that charged particles induce about 30-40% more trapping centres than neutrons with the same NIEL [145].

**Change in electric field**

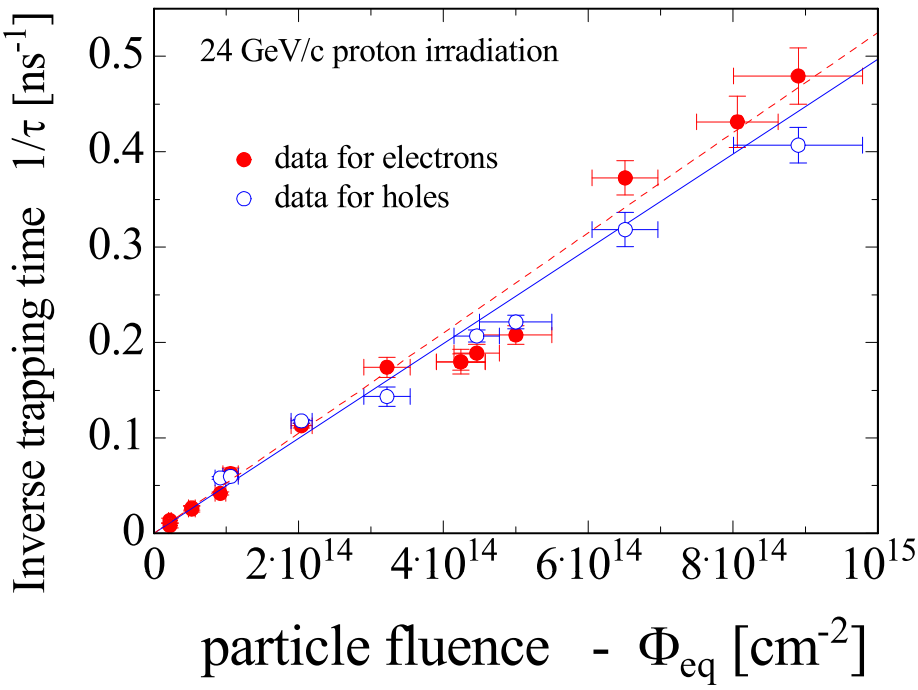



**Figure 6.7:** Inverse trapping time as a function of fluence for electrons and holes after irradiation with 24 GeV protons. From [144].

After high fluence irradiation the electric field has a complex structure. While for a nonirradiated device the space charge is homogeneous over the bulk and hence the electric field has a linear dependence on depth, for a highly irradiated sensor the electric field profile has a double peak structure with maxima at the implant side and at the backplane. This is a consequence of radiation induced defects in silicon: free carriers are generated and while the electrons drift towards the p-type implants, holes drift towards n-type implants. This produces an high density of negative charge, which corresponds to n-type doping, near the p-type implants and of positive charge, which corresponds to p-type doping, near the n-type implants, respectively. Part of the charge is trapped by the radiation induced defects, leading to two space charge regions separated by an electrically neutral region [146]. Hence, a p-n junction is formed at both sides of the sensor. The electric field is present in the whole sensor bulk, with a peak at the pixel implant side, as intended, and a smaller peak at the back side. This can be seen from Figure 6.8, where the sum of electron and hole velocity is measured as a function of sensor depth for different operation voltages [147].

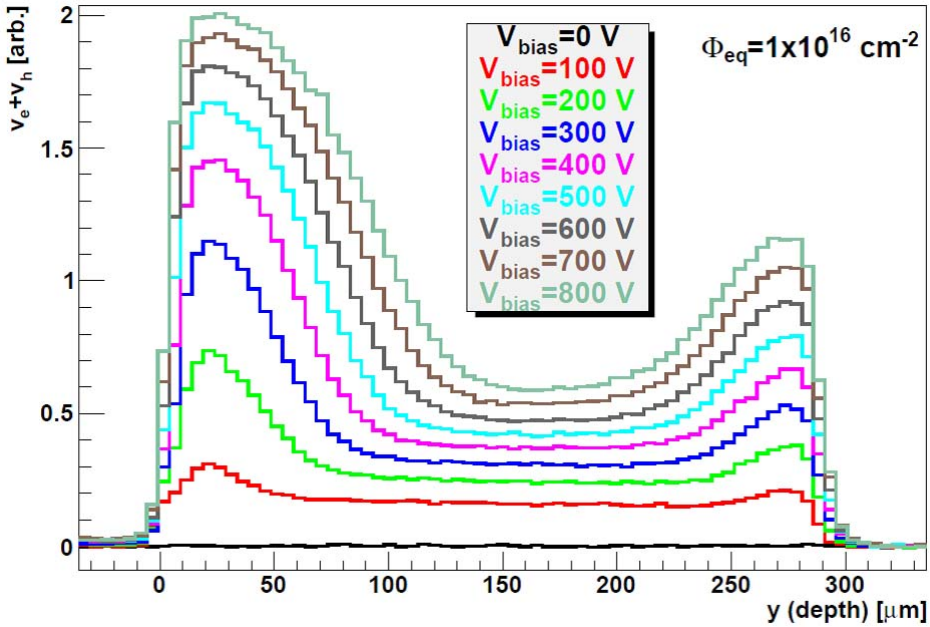



**Figure 6.8:** Drift velocity profiles for a neutron irradiated p-type strip sensor operated at different voltages. From [147].

### 6.2.3 Annealing

The damage due to non-ionising radiation can be partially recovered with a process known as annealing, which consists of exposing the sensor to a certain temperature for a given amount of time. Annealing is possible owing to the mobility of some of the defects (vacancies and interstitials), which is strongly temperature dependent. However, the damage can only be partially recovered. The annealing can be accelerated or decelerated by increasing or decreasing the temperature.
The effective doping concentration $N_{\text{eff}}$ changes with annealing time and temperature, as analytically described by the Hamburg model [148]:

$$\Delta N_{\text{eff}} = N_C(\phi) + N_a(\phi, T_a, t) + N_Y(\phi, T_a, t), \tag{6.15}$$

where

- $N_C$ represents the stable damage. The stable damage can not be recovered, since the removal of donors and the increase of acceptors caused by irradiation depends solely on fluence and not on temperature.

- $N_a$ represents short term annealing. Short term annealing depends not only on fluence, but also on the temperature $T_a$ and the time after irradiation $t$. It is observed as a decrease of the depletion voltage for a short time after irradiation (at

room temperature it lasts a few days). It is therefore beneficial for the operations of the detector and is also referred to as beneficial annealing.

- $N_Y$ represents long term annealing. Like the beneficial annealing it depends on temperature and time after irradiation. Long term annealing follows in time the short term annealing and consists of an increase of depletion voltage due to the increase of the negative space charge region. Hence, it is not beneficial for the operations of the detector and is also referred to as reverse annealing.

Annealing affects the trapping of charge carriers as well. It has been observed that the trapping-time damage constant decreases by about 35% for electrons, while at the same time it increases by 30% for holes [149]. In terms of current related damage annealing, the damage related constant $\alpha$ of Eq. 6.13 is observed to decrease with annealing time [148]. In order to avoid reverse annealing effects, detectors need to be cooled down also during shutdown periods. This has been the case for the VELO detector during the first long shutdown and the winter shutdowns. Only towards the end of Run 2 of data taking the VELO sensors needed to be annealed for about a day at about 20° to mitigate the effect of radiation damage.

### 6.2.4 Impact on detector operations

The different radiation related damage effects affect the operation of the device. The leakage current depends exponentially on temperature, according to Eq. 6.10. Hence, since the increase in leakage current caused by irradiation leads to an increase in temperature, this causes a positive feedback, known as thermal runaway, unless the device is cooled down. Furthermore, cooling is crucial not only during operations of the detector, but also in shutdown periods such that reverse annealing effects are avoided. The operation voltage needs to be adjusted with the change in effective doping concentration. With increasing radiation damage, the bias voltage needs to be increased to operate the sensors fully depleted. It is also possible, after a certain fluence, that the full depletion voltage becomes too high to be applied and sensors are operated partially depleted. Therefore the charge collection efficiency degrades and the amount of charge collected is reduced. At high fluence trapping becomes the limiting factor with a significant signal reduction due to the high concentration of trapping centres.

## 6.3 Pixel sensors

The main features of the sensor designs are discussed, with particular focus on the characteristics chosen for the sensors under consideration for the VELO upgrade.

### 6.3.1 Sensor types

Depending on the doping of the pixel electrodes and the bulk of the sensors, four different sensor types can be distinguished:

- n-type electrodes into p-type bulk (n-on-p);
- n-type electrodes into n-type bulk (n-on-n);
- p-type electrodes into p-type bulk (p-on-p);
- p-type electrodes into n-type bulk (p-on-n).

If the collection electrodes are of different type than the bulk the depletion region starts from the electrode side and the sensor can be operated partially depleted. If the collection electrodes have instead the same type as the bulk then the depletion start from the backplane: since the backplane is not segmented, the sensor needs to be operated fully depleted to get precise position information. Moreover, double-sided processing is then needed to place guard rings, described in the next section, on the back-side. Sensors with n-type bulk are subject to type inversion after fluences of the order of $10^{12} \cdot 1\,\mathrm{MeV}\,\mathrm{n_{eq}}\,\mathrm{cm}^{-2}$. A p-on-n sensor therefore undergoes type inversion and effectively becomes a p-on-p sensor, with the depletion layer starting from the backside and need to be operated fully depleted. Hence, n-on-p and n-on-n types are used when radiation hardness is required and are these types considered for the VELO upgrade. These two types of sensors are illustrated in Figure 6.9. The depletion region is indicated: it always starts from the p-n junction and extends in the bulk of the sensors. After type inversion, the n-on-n sensor behaves as an n-on-p sensor with the junction at the pixel electrode side.

### 6.3.2 Guard rings

Mechanical damage from the dicing makes the sensor edge conductive and it will therefore be at the backside potential, while the pixel electrodes are at ground potential. This can lead to two potential risks:

- the depletion volume extends laterally reaching the cutting edge, where the mechanical damage acts as carrier generation centre with a consequently high increase of leakage current;
- the surface gets the same potential as the back side, and as a consequence an electron layer accumulates extending up to the lateral end of the junction. Hence, there is a high potential difference between the electron layer and the pixel electrode, consequently a high electric field and low breakdown voltage.

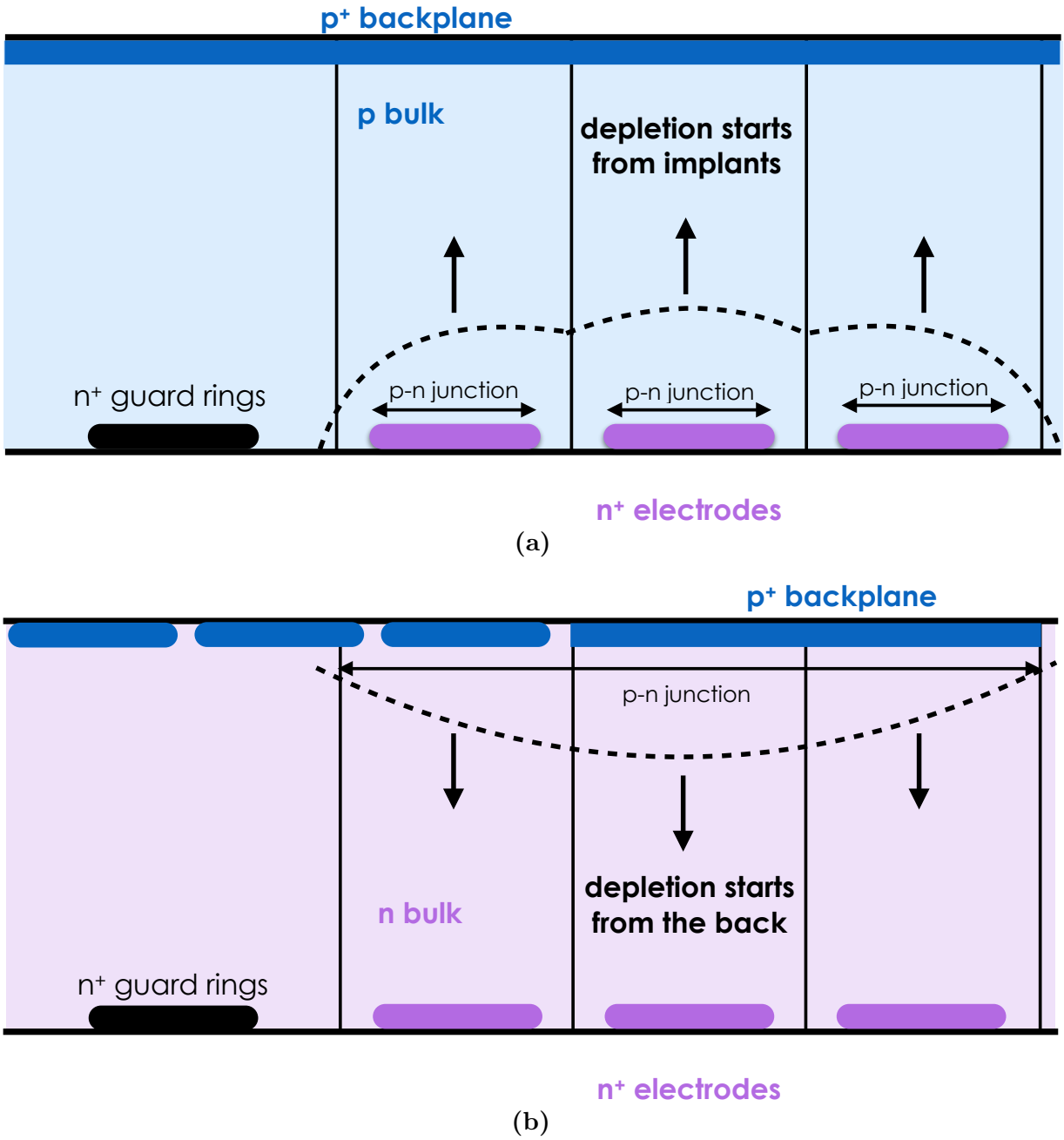



**Figure 6.9:** Sketch of a sensor with n-type electrodes on (a) a p-type bulk and (b) an n-type bulk (before irradiation). In the n-on-n sensor there are also backplane guard rings, implanted underneath the edge pixels.

Guard rings are structures that have been introduced to provide a gradual drop of potential towards the edge and isolate the pixel electrodes from the cutting edge for n-on-p sensors. The guard rings are placed around the pixel matrix, extending the physical edge of the sensor. The area between the physical edge and edge of the pixel matrix is

inactive, but contributes to the material budget and increase the distance between the interaction region and the first active pixels. Hence, the size of the guard ring is a trade between the amount of insensitive material introduced and the bias voltage that can be applied. As described in Section 5.2, the fluence is expected to be highly nonuniform. This implies that the guard ring size should account for lateral depletion of low irradiated sensor at the high voltage required for the highly irradiated part. Therefore, in view of the VELO upgrade different guard ring designs have been considered, as described in Section 7.4.

# Chapter 7

# Testbeam

A wide range of sensor characteristics for various prototypes have been tested in a beam of minimum ionising particles in view of the LHCb VELO upgrade. In particular this chapter focuses on the performance of sensors at the edge of the pixel matrix. The requirements the sensors must fulfil are outlined in Section 7.1. The prototype sensors, coupled to Timepix3 ASICs, are discussed in Section 7.2; the specifications and the different sensor designs are presented, followed by a description of the Timepix3 ASIC and its calibration. Then in Section 7.3 the Timepix3 telescope, built to evaluate the performance of the prototype assemblies for the VELO upgrade, is introduced. The offline reconstruction of the telescope data and the main performance figures that are essential for the sensor characterisation are described. Finally, the testbeam results regarding the edge behaviour for all tested prototype sensors are presented in Section 7.4.

## 7.1 Sensors Requirements

The upgrade of the LHCb detector brings challenges in terms of data rate and occupancy, but the most challenging aspect from the perspective of sensor R&D is the nonuniform radiation exposure. As shown in Figure 5.7, the innermost region of the sensor will be highly irradiated, while the outermost region is irradiated by a factor 40 to 120 less, depending on the distance along the beam axis with respect to the collision point. The irradiation itself causes a degradation of the charge collection efficiency due to changes in the doping concentration and trapping of the charge carriers. This effect can be partially compensated by increasing the bias voltage applied. Due to the change in doping concentration and therefore reduced lateral depletion, the HV tolerance is expected to improve with radiation. Increasing the voltage applied would be enough to recover a good signal, but the innermost part of the sensor will be highly irradiated and hence requires high bias, while the outermost part of the same sensor is less irradiated and must sustain the same bias without suffering from avalanche breakdown effects. At the end of lifetime the hottest part of the sensor is exposed to a fluence of $8 \times$

| | |
|---|---|
| Thickness | $\leq 200\,\mu\text{m}$ |
| Type | n-on-n or n-on-p |
| Full depletion voltage ($V_{FD}$) before irradiation | $< 300$ V |
| Maximum guard ring width | $450\,\mu\text{m}$ |
| Pre-irradiation breakdown voltage | $3 \times V_{FD}$ |
| Breakdown voltage after non-uniform irradiation | $> 1000$ V |
| Collected charge after full irradiation | $> 6000\ e^-$ |
| Efficiency after full irradiation | $> 99\%$ |

**Table 7.1:** Sensor requirements.

$10^{15} \cdot 1\,\text{MeV}\,\text{n}_{\text{eq}}\,\text{cm}^{-2}$. Such high radiation requirements led to the choice of an n-side readout, which benefits from the high electric field [150] at the readout side, allows for shorter collection times and it is less sensitive to trapping with respect to a p-type readout [151, 152]. With $\text{p}^+$-pixel implants the current is mainly induced by the motion of holes, while with $\text{n}^+$-pixel implants mainly by the motion of electrons. Since the charge collected depends on the collection time and electrons have a mobility about three times higher than the holes while the trapping rate is approximately the same, electrons collecting electrodes are preferred for highly irradiated sensors. The n-type electrodes can be implemented both on an n-type and p-type silicon bulk.
Regarding the thickness of the sensor, the choice was made to construct thinner sensors than the $300\,\mu\text{m}$ strips of the previous VELO. In this way the amount of material is reduced in the acceptance. Furthermore, the edge can be narrower, since the distance from the active volume to the edge needs to be larger than the sensor thickness to avoid that the depletion volume extends laterally up to the sensor edge. After irradiation, the effect of charge trapping in thinner sensors is mitigated because the drift path of the signal charge is reduced with respect to a thicker sensor. Therefore the choice has been made for a maximum thickness of $200\,\mu\text{m}$.
The sensors are required to withstand a bias voltage of 1000 V in order to provide sufficient signal up to the full radiation fluence. The charge collected must be higher than 6000 electrons, to mitigate the effect of timewalk and be sufficiently above threshold to ensure particle detection with high efficiency, since the charge can be shared up to four pixels and the typical threshold is $1000\ e^-$. The cluster finding efficiency must be higher than 99% in order to reduce the risk of missing the first measured point on a track, which dominates the impact parameter resolution. The requirements for the sensors are summarised in Table 7.1. The sensor R&D program includes the evaluation of HV tolerance, charge collection and cluster finding efficiency, as well as spatial resolution for the different sensor types presented in section Section 7.2. While these characteristics are evaluated in the centre of the sensor, it is also crucial to monitor the behaviour at the edge of the sensor.

### 7.1.1 Behaviour at the Edge

During operations of the detector the sensor edge is the part closest to the interaction region and hence where the occupancy is the highest. At the same time, the impact parameter resolution depends on the radial distance of the first measured point on a track and the detector hit resolution. Therefore the sensor must be fully efficient up to the edge and distortions of the electric field in the sensor must be avoided as that would lead to a degradation of the spatial resolution. In order to protect the active area from edge effects, all sensors have a guard ring structure to gradually reduce the electric field towards the edge. The guard rings have different designs depending on the sensor type. The study presented in Section 7.4 is devoted to evaluate the different prototype assemblies focusing on the performance at the edge of the sensor.

## 7.2 Assemblies

More than 30 different assemblies have been tested in order to assess the best candidate for the VELO upgrade. Included are both single sensors, $256 \times 256$ pixels with $55 \times 55$ $\mu m^2$ pixel pitch, and so-called triples which are $3 \times 1$ tiles coupled to three readout ASICs. In the R&D phase the sensors have been read out by the Timepix3 chip. The cross-section of a n-on-p assembly is sketched in Figure 7.1. As shown in the sketch, the inner guard ring is grounded, while the outer guard ring is left floating. In the case of n-on-n sensors, as described in Section 6.3, there are backside guard rings. The assembly is glued to a 635 µm thick AlN ceramic board and wirebonded to a custom made kapton-copper hybrid, also glued to the ceramic substrate.
Since the radiation tolerance is a key factor in the sensor R&D, all the assemblies have been irradiated up to the maximum fluence expected after $50\,\mathrm{fb}^{-1}$ of integrated luminosity, with uniform and nonuniform profiles and different types of irradiation with protons and neutrons at several facilities.

### 7.2.1 Sensor description

The prototype sensors have been produced by two different manufacturers, Hamamatsu (HPK) and Micron Semiconductor Ltd. The baseline design, according to the physics and detector requirements listed in the previous section, consists of 200 µm thick sensors, n-on-p type, with 450 µm guard ring (GR) size. Also other options have been considered, with different implant width, thickness, guard ring size and bulk type.
Hamamatsu proposed two different implant widths of 35 and 39 µm and also sensors with a more conservative guard ring width of 600 µm, in case of HV tolerance issues with the narrower guard ring design. Micron provided both 200 µm thick n-on-p sensors and 150 µm thick n-on-n sensors. Regarding the guard ring design, for Micron n-on-n sensors

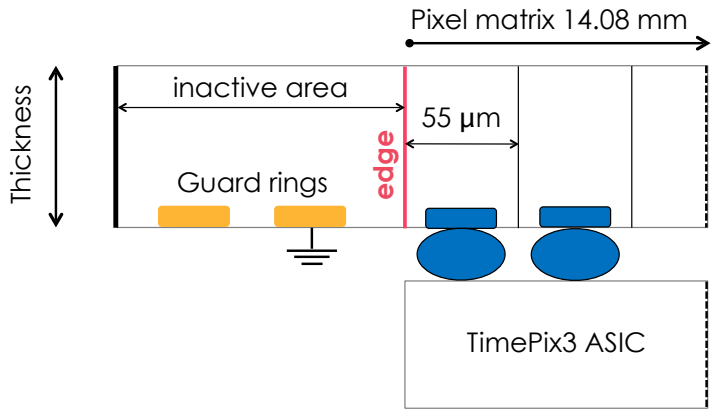



**Figure 7.1:** Cross-section of an n-on-p assembly.

it is possible to have also a reduced edge distance of 250 µm, because the backside guard rings can be implanted underneath edge pixels such that the guard ring area slightly overlaps with the active region. The assemblies can be divided in three categories, whose main characteristics are summarised in Table 7.2. In terms of edge performance, the Micron n-on-n family has been further classified in sensors with 450 µm guard ring width and 250 µm guard ring width.

For a larger clearance between the sensor and the RF foil and to maintain a constant

**Table 7.2:** Prototype assemblies.

| Vendor | Type | Thickness | GR size | Implant width |
|---|---|---|---|---|
| HPK | n-on-p | 200 µm | 450,600 µm | 39,35 µm |
| Micron | n-on-p | 200 µm | 450 µm | 36 µm |
| Micron | n-on-n | 150 µm | 450,250 µm | 36 µm |

distance between the active region and edge, it has been decided to round the sensor corners. Sensors with rounded corners have been provided by HPK. The prototypes used for the study presented in Section 7.4 do not have this feature, since it has been introduced at a later stage of prototyping. It has been verified afterwards that no difference is introduced by rounded corners in terms of edge performance.

### 7.2.2 Timepix3 ASIC

The Timepix3 ASIC [128], which is derived from the Medipix family of chips, is used as a proxy of VeloPix for prototype testing. The chip, designed in 130 nm CMOS technology, consists of a matrix of 256 × 256 square pixels of 55 µm pixel pitch. It supports a zero-suppression scheme, such that only pixels hit are read out. The readout is continuous and triggerless with the possibility of a data-driven mode, meaning that as soon as a

hit is recorded it is sent out of the chip. In Table 7.3 the specifications of Timepix3 and VeloPix are compared. The major difference concerns the data rate the chip can handle: the VeloPix maximum hit rate per pixel is more than a factor ten than that of the Timepix3. The Timepix3 ASIC has several advantages when compared to VeloPix with respect to sensor characterisation. In one of the three different data acquisition modes provided, the Timepix3 ASIC can measure simultaneously the threshold crossing time, denoted as Time-of-Arrival (ToA), and the time the signal is above threshold, denoted as Time-over-Threshold (ToT). The former provides the hit timestamp, with a resolution of 1.56 ns, which corresponds to the bin size of the per pixel Time-to-Digital converter. The latter is proportional to the energy deposited and can be converted into equivalent units of collected electrons via a charge calibration process. For the VeloPix the ToT value can only be read out at a slow rate for calibration purposes only and the timestamp has a resolution of 25 ns.

A simplified diagram of the Timepix3 logic is presented in Figure 7.2. Each pixel has

**Table 7.3:** Specifications of the Timepix3 and VeloPix ASICs.

| Specification | Timepix3 | VeloPix |
|---|---|---|
| pixel size | $55 \times 55\,\mu m^2$ | $55 \times 55\,\mu m^2$ |
| matrix size | $256 \times 256$ | $256 \times 256$ |
| timestamp resolution | 1.56 ns | 25 ns |
| Time over Threshold | yes | low rate only |
| peak pixel hit rate | 80 MHit/s | 900 MHit/s |
| power consumption | $< 2$ W/ASIC | $< 3$ W/ASIC |
| radiation hardness | no spec. | $> 400$ MRad |
| single event upset robust | no | yes |

an analog and digital front-end and $2 \times 4$ pixels are grouped in a super pixel structure with shared readout logic. Differently from the VeloPix, the super pixel does not have the purpose of reducing the data rate by grouping several hits in one packet. The analog front-end receives the signal from the sensor through the input pad, converts it into a voltage and compares it with a pre-set threshold. If the voltage is above threshold, then a logic 1 is generated and sent to the digital front-end. The main elements of the analog front-end are a charge sensitive pre-amplifier with a leakage current compensation feedback circuitry known as Krummenacher scheme [153], a single threshold discriminator and circuitry for testpulse injection. In the Krummenacher scheme a current, denoted as $I_{krum}$, is provided: half of it is used to discharge the signal and the other half for leakage current compensation. When the sensor has a large leakage current, due for example to high radiation damage, the $I_{krum}$ must be increased, but this has the drawback of worsening the ToT resolution. This is due to the fact that increasing the $I_{krum}$ leads to a faster discharge, hence a shorter Time-over-Threshold,

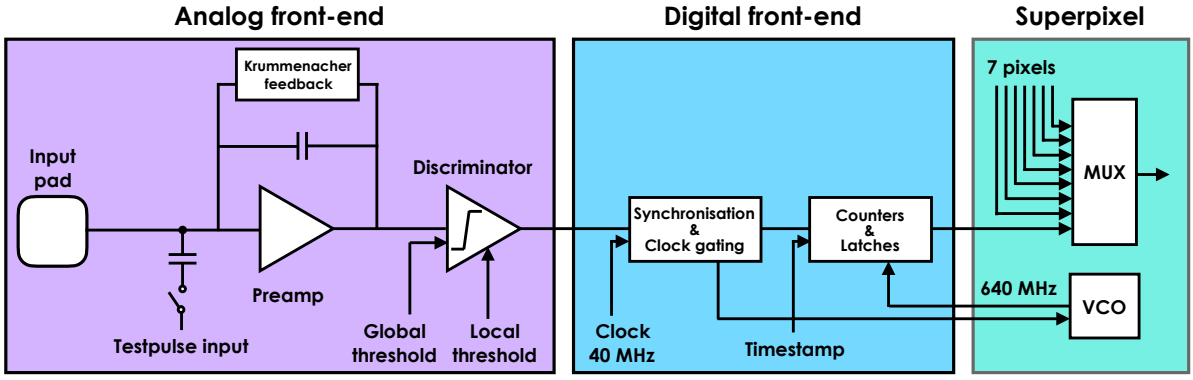



**Figure 7.2:** Simplified diagram of Timepix3 logic, reproduced from [128].

while the clock frequency does not change. Therefore the resolution is worse and the uncertainty due to digitisation larger, as described in Section 8.2. For nonirradiated sensors the $I_{krum}$ is set to 10, giving a ToT value for the MPV of around 120 counts[1] for a 200 µm thick sensor, while for irradiated sensors it is usually 20, giving a ToT value for the MPV of around 25 counts. The discriminator compares the output voltage from the pre-amplifier to a global threshold. It is also provided with a 4-bit digital-to-analog converter, known as trimdac, for local threshold tuning in order to compensate for pixel to pixel variations. The trimdac is indicated as *local threshold* in Figure 7.2. The trimdac settings are obtained via a procedure called equalisation.
The signal at the output of the discriminator is processed in the digital front-end. An external 40 MHz clock is distributed to the whole pixel matrix for synchronisation and ToT counting. When the discriminator output rises, it activates the voltage-controlled oscillator (VCO) in the super pixel logic which provides a 640 MHz clock to all eight pixels in the superpixel. A pixel that is hit counts the number of 640 MHz clock cycles up to the first rising edge of the 40 MHz clock. The two clocks are combined to get more precise values for ToA.
The global threshold must be chosen to remove noise hits. Hence the threshold value is usually set to be at least six times above the sum in quadrature of the noise level and the threshold variation. For Timepix3 the average noise for the whole pixel matrix is measured to be about 60 $e^-$ [154] and the threshold dispersion after equalisation is about 30 $e^-$, therefore the threshold could be set at around 500 $e^-$. In practice, the default threshold during the data taking is 1000 $e^-$ to ensure a measurement virtually free of noise.

[1] ToT counts are the number of cycles of a 40 MHz clock for which the signal is over threshold.

### Threshold equalisation

Pixel-to-pixel variations lead to slightly different baseline voltage of the pre-amplifier and, since the Timpix3 ASIC has a global threshold for the whole pixel matrix, this could lead to significant threshold dispersion across the pixel matrix. A threshold equalisation procedure is needed to tune the per-pixel trimdac settings such that this effect is mitigated. The first step consists of finding the baseline voltage of the pre-amplifier per pixel. This is obtained by scanning the global threshold and measuring the number of noise counts at each step: the centroid of the noise distribution obtained represents the baseline voltage in units of threshold. This scan is performed twice per pixel: first setting the trimdac to the minimum (0) and then to the maximum (15). The two distributions of the baseline voltage for all the pixels can be seen in Figure 7.3 in blue and red, respectively. It can be seen that changing the trimdac value shifts the effective threshold. Hence, a target threshold is chosen as the midpoint between the two distribution and the optimal setting per pixel is identified by interpolation (assuming a linear relation between the two endpoints) as the trimdac value which gives an effective threshold as close as possible to the midpoint. The black distribution is the result of the equalisation. Some dispersion is still present due to the binning of the trimdac, but it is significantly smaller, of the order of 30 $e^-$. If it is not possible to find a trim value that brings the effective threshold at most one trimdac step away from the target threshold, the pixel is masked. Usually the number of masked pixels is $\sim 1‰$. The equalisation procedure is performed before and after irradiation of the sensor to account for possible differences due to radiation damage in the ASIC.

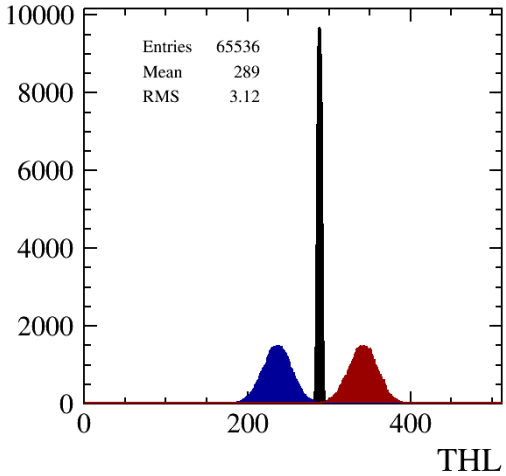



**Figure 7.3:** Example of a threshold equalisation. The blue distribution represents the threshold distribution for all the pixels with trimdacs set to 0, the red distribution for trimdacs set to 15 and the black distribution for optimised trimdacs settings after equalisation. The legend refers to the black distribution. From [155].

### Charge calibration

A charge calibration is needed to convert the ToT value recorded into the corresponding number of electrons. The calibration is performed by injecting testpulses into the analog front-end via a capacitor of design value 3 fF. Per voltage step 100 testpulses are injected and the average of the measured ToT values is studied as a function of the known charge (Figure 7.4), obtained as $Q = C \times V$ where $C$ is the capacitance and $V$ the injected voltage. The variance on the 100 testpulses is assigned as uncertainty to the ToT value. A calibration curve is obtained by fitting the data with the so-called surrogate function:

$$\mathrm{ToT}(q) = p_0 + p_1 q - \frac{c}{q - t}, \tag{7.1}$$

where $p_0$ and $p_1$ represent the gain and offset of the linear part of the function and $c$ and $t$ denote the curvature and asymptote of the non-linear part. The non-linearity close to threshold is due to the charge sensitivity of the discriminator. The measured ToT value for a given injected charge varies from pixel-to-pixel due to variations in the discharge current, therefore a calibration curve is determined for every pixel individually. The fitted parameter values are then used to convert the recorded ToT counts to unit of electrons via the inverse surrogate function:

$$q(\mathrm{ToT}) = \frac{p_1 t + \mathrm{ToT} - p_0 + \sqrt{(p_0 + p_1 t - \mathrm{ToT})^2 + 4 p_1 c}}{2 p_1}. \tag{7.2}$$

A typical fit can be seen in Figure 7.4.

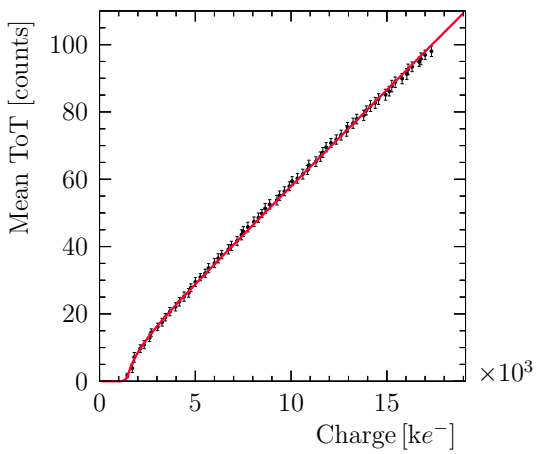



**Figure 7.4:** Example fit of the mean ToT response over 100 tespulses as a function of the injected charge.

### Timing effects

An effect that occurs in the chip, known as timewalk, affects timing measurements and needs to be corrected for if the arrival time of a signal must be determined with high precision. Timewalk refers to the time difference in recording signals of different amplitude that occurred simultaneously: as shown in the sketch of Figure 7.5, the signal rise time is independent of the signal amplitude and a signal with smaller amplitude crosses the threshold later than a signal with larger amplitude. This is due to the combined effect of a finite rise time and a single constant threshold. An additional contribution to timewalk is due to the response time of the discriminator, which is charge dependent. An offline correction can be obtained by exploiting the simultaneous measurement of ToA and ToT provided by the Timepix3 chip, as described in Section 8.3.

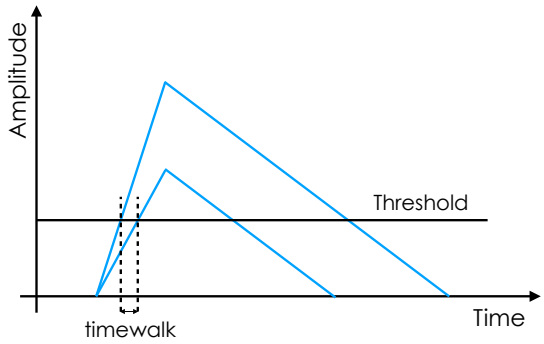



**Figure 7.5:** Illustration of the timewalk effect.

## 7.3 Timepix3 Telescope

An extensive test beam program has been carried out at the SPS H8 beamline at CERN to characterise the sensors at high rate. The beam is a mixed charged hadron beam ($\sim 67\%$ protons, $\sim 30\%$ pions) at 180 GeV. The beam at the SPS has a spill structure, where each spill lasts about 4.5 s and contains roughly $10^6$ particles. The advantage of a minimum ionising beam of particles to test the prototype assemblies is that beam particles have a well defined energy and trajectory through the sensor because of small multiple scattering. This allows accurate measurements and to probe the sensor within the dimension of one pixel of 55 µm pitch. In view of the sensor characterisation for the LHCb VELO upgrade, a telescope based on Timepix3 has been built. The simultaneous measurement of ToA and ToT by the Timepix3 chip makes it ideal for fast and robust pattern recognition and track reconstruction. First the telescope is described, then the main steps of the offline reconstruction are outlined and finally the main performance figures are presented.

### 7.3.1 Hardware

The Timepix3 telescope [156], shown in Figure 7.6, is a high rate, data-driven beam telescope. The telescope comprises two arms of four planes each, where each plane is rotated by about 9° around the $x$ and $y$ axes to optimise the single-hit resolution. A global right-handed coordinate system is adopted where the $z$ direction is along the beam axis and the $y$ directions points upwards. Each plane consists of a 300 µm thick p-on-n sensor bump-bonded to a Timepix3 ASIC, which is read out by the Speedy PIxel Detector Readout (SPIDR) [157] board, specifically designed for the readout of Medipix3 and Timepix3 chips at maximum rate. Each SPIDR board reads out two planes in parallel for a total of four boards. An additional board is dedicated to the Device Under Test (DUT). The eight telescope planes are synchronised by a central logic unit (TLU) which provides a clock, as well as a signal to synchronise the time counters and a shutter signal to synchronise the start and stop of the data flow. The DUT is placed in the middle of telescope, on a remotely controlled motion stage which can translate in $x$ and $y$ directions and rotates about the $y$ axis. The central stage is provided with a vacuum box in order to test irradiated devices at high voltage. The DUT cooling system consists of a combination of a glycol chiller typically operated at a temperature of −15°C and a Peltier element. In standard operation conditions the telescope bias is 150 V while the full depletion voltage of the sensor is about 50 V.

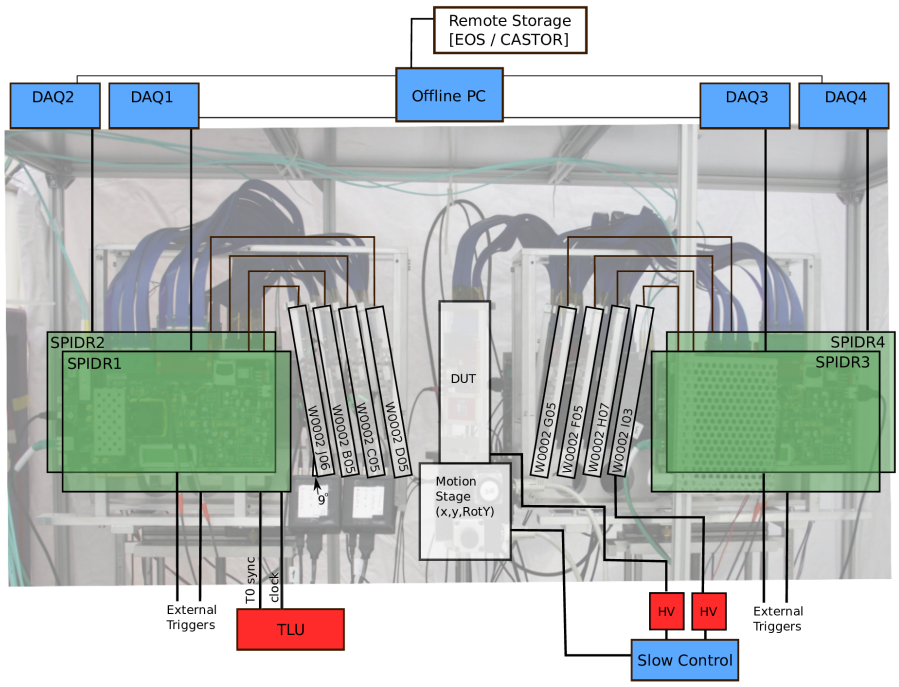



**Figure 7.6:** Photograph of the Timepix3 telescope with overlaid the labels for the main components, described in the text.

### 7.3.2 Offline reconstruction

The offline processing is performed with the KEPLER application, based on the GAUDI event-processing framework [158]. The track reconstruction is performed in three stages: clustering, pattern recognition and track fitting. The DUT is excluded from this procedure in order not to bias the result. Once the track is obtained, the cluster on the DUT is associated to the track. Before the full track reconstruction is performed, the telescope planes are aligned in space and time. Each of these steps is described in the following sections.

#### Tracking

Firstly the hits are time ordered, then neighbouring pixels recorded within a time window of 10 ns are grouped into a cluster. In the default configuration a maximum of one pixel gap is allowed within a cluster, to account for masked or dead pixels; the number of allowed missing hits can be increased in analyses where the tracks are at a shallow angle with respect to the sensor, such as in the grazing angles analysis described in Chapter 8. The cluster position in local coordinates is defined as the centre of the pixel for one-pixel clusters, while for a multi-pixels cluster it is computed as the charge weighted average of the pixels positions, known as Centre-of-Gravity method. The cluster time is assigned to be the timestamp of the earliest hit in cluster, such that the time spread due to timewalk of low charge hits is reduced. Clusters with a width larger than three columns or rows are rejected in order to remove clusters due to $\delta$-rays.
The pattern recognition is time based: a track seed is generated from a pair of clusters, one in the most upstream plane and one in the second plane, that are within a time window of 10 ns. The seed track is then extrapolated downstream by requiring a hit within a time window of 10 ns and an opening angle of 0.01 rad, under the assumption of straight tracks. After a plane is added, the track seed is updated. This is iterated for all the clusters, excluding from the pattern recognition the clusters that are already associated to a track. Eight clusters are required per track. The track is then fitted with a straight line. To reject tracks with bad fit quality, the fit $\chi^2$/ndf is required to be less than 10. The track time is the average time of the clusters forming the track. After the telescope tracks are reconstructed, the DUT cluster to track association is performed. The clusters on the DUT are associated to the tracks if they fall within a time window of 100 ns and a square spatial window of 150 μm from the track intercept.

#### Time alignment

The time alignment compensates for small time offsets between the planes due to difference in the delays introduced by cables and electronics and difference in time of flight of the particles. The offsets are given by the mean of the biased track time residuals

distribution per plane, where the time residual is the difference between the track and cluster time and biased refers to the fact that the plane is included in the tracking. The offsets are measured for each run, but they are found to be the same within 15 ps. The time alignment is crucial especially in runs where the DUT is at grazing angles (described in Chapter 8), where the spatial association of DUT clusters to tracks is not possible.

### Spatial alignment

The spatial alignment provides the relation between the local coordinates of each sensor plane and the global reference frame of the telescope. A good alignment is needed to be able to exploit the pointing resolution of the telescope and probe the DUT with a precision that is a fraction of the pixel pitch. The alignment procedure is performed first for the telescope planes, followed by the alignment of the DUT with respect to the telescope. The alignment procedure is track based and is iterated multiple times with more restrictive selection requirements at each stage.
The algorithm adopted for the telescope alignment is Millepede [159], which simultaneously fits the tracks states and six alignment parameters per sensor and minimises the sum of the track residuals. The track states are $(x_0, y_0, t_x, t_y)$, where $x_0$ and $y_0$ are the initial global positions along the $x$ and $y$ axes, while $t_x$ and $t_y$ are the track slopes in $x$ and $y$. The alignment parameters are $(X, Y, Z, \theta_X, \theta_Y, \theta_Z)$, where $X, Y, Z$ are the translations and $\theta_{X,Y,Z}$ the rotations around the $X, Y, Z$ axes, respectively. A sample of about $10^4$ tracks is reconstructed using a preliminary alignment, the algorithm is run and the alignment parameters updated. At the first iteration only $X$, $Y$ and $\theta_Z$ are optimised. The procedure is repeated and gradually the other alignment parameters are included, while the cuts on spatial association and track fit quality are tightened. At the last iteration all the geometrical parameters are optimised. The unbiased resolution per plane, obtained by excluding the plane from the pattern recognition, is about 5 µm both in $x$ and $y$ directions after the track based alignment procedure is performed. The variation in the width of the residual distribution is less than 0.2 µm and the width is very stable, proving the high quality of the alignment.
Once the telescope alignment is optimised, the DUT is aligned. The DUT alignment is performed using the Minuit [160] algorithm. During the DUT alignment procedure the telescope alignment is fixed and $10^4$ tracks are reconstructed by the telescope. If a cluster on the DUT is within a spatial and temporal window from the track intercept then the cluster is associated to the track. The alignment of the DUT relies on the minimisation of the residuals as for the telescope, but the $\chi^2$ depends only on one set of geometrical parameters. The process is iterated, reducing the spatial and time window cuts for the cluster association. The alignment procedure is performed per run. The residual distribution for one-pixel clusters is expected to be flat and centred at 0. An

indication of good alignment is that the mean value of the residual distribution is smaller than the pointing resolution of the telescope and no correlation is observed between the residuals and the track intercept coordinates. An example is shown in Figure 7.7 for a Micron n-on-p sensor placed perpendicularly to the beam: the horizontal band indicates that no misalignment occurs.

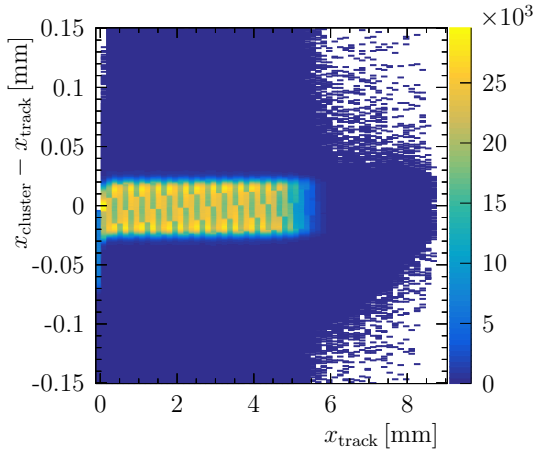



**Figure 7.7:** Unbiased residuals in the $x$ direction as a function of the track intercept in $x$ for a Micron n-on-p assembly after spatial alignment is performed.

### 7.3.3 Performance

The most relevant telescope performance figures for the DUT performance evaluation are the track time resolution and pointing resolution. After the telescope planes are time aligned, the measured intrinsic resolution of each plane is approximately 1.1 ns in the standard operation condition where the telescope bias is 150 V. Assuming the eight planes provide uncorrelated measurements, the track time resolution yields 370 ps. The pointing resolution is the precision of the track position extrapolated to the DUT position and is evaluated using a simulation. The resolution as function of $z$ is shown in Figure 7.8: the blue band represents the predicted pointing resolution, which is $1.69 \pm 0.16\,\mu\text{m}$ and $1.55 \pm 0.16\,\mu\text{m}$ [156] in the $x$ and $y$ direction, respectively, at the DUT position. The simulation is validated by comparing the residuals of the telescope planes from data (red markers) and simulation (green markers).

## 7.4 Prototype sensor edge analysis

The sensor charge collection properties are investigated at the sensor edge by positioning the DUT such that the beam illuminates the edge of the sensor. Unless stated

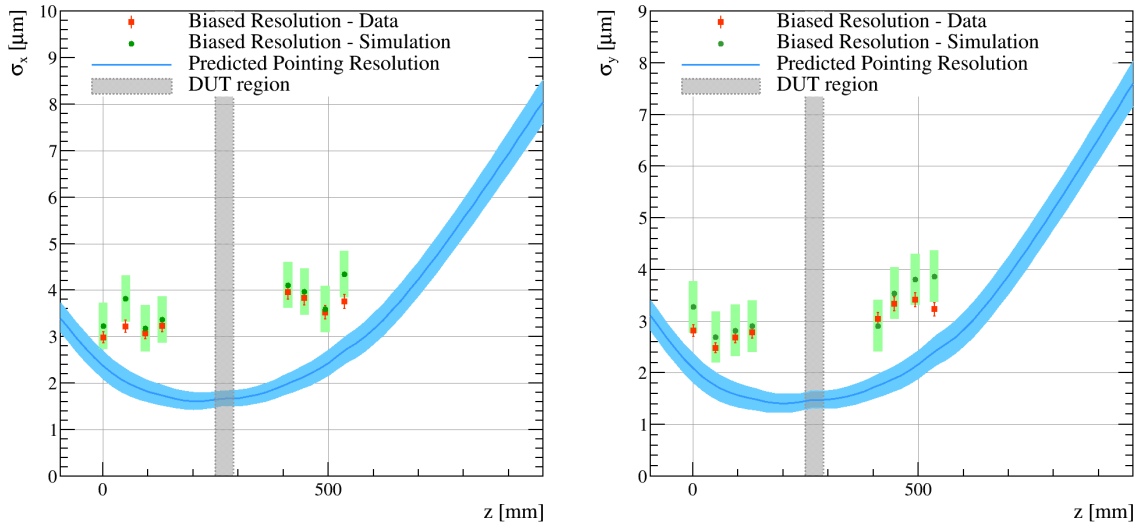



**Figure 7.8:** Predicted pointing resolution (blue solid line) in the $x$ (left) and $y$ (right) direction as a function of the position along the beam axis. The blue band represents the uncertainty on the predicted pointing resolution. The grey region indicates the DUT position. The biased resolution from simulation and data are indicated by green and red markers, respectively. The green band represents the systematic uncertainty on the simulate biased resolution. Figure taken from [156].

otherwise, the dataset has been acquired placing the sensor perpendicular to the beam. The following measurements are reported for one corner of the sensor, but it has been verified that the other corners show the same behaviour. The sensors are operated at a bias voltage where they are fully depleted and the threshold is set at 1000 $e^-$. The results reported are for single chip sensors, nonirradiated as well as uniformly neutron irradiated, but all prototypes tested show no difference regarding the behaviour at the edge between single and triple sensors or between different types of irradiation.
Figure 7.9 shows the hitmap for a Micron n-on-p sensor with the beam illuminating the bottom left corner, which corresponds to column and row zero, and its projection on the columns of the pixel matrix. It can be seen that there is an excess of hits in the first column and first row of the pixel matrix, which amounts to about 30% more hits with respect to the average number of hits for pixels away from the edge. In order to investigate the excess of hits observed for Micron n-on-p assemblies in the first column, the cluster charge distributions for different columns are compared integrating over all rows, where the cluster charge is assigned to the column with the highest charge. In Figure 7.10 the four different types of prototype assemblies are compared and it can be seen that the first pixel columns have a different distribution: the peaks of the Landau distributions for different columns overlap, but hits accumulate at low charges for the first column. Some low charge hits are expected because for some charge from hits in the guard ring region diffuses to the first column. The HPK sensor with 450 μm guard

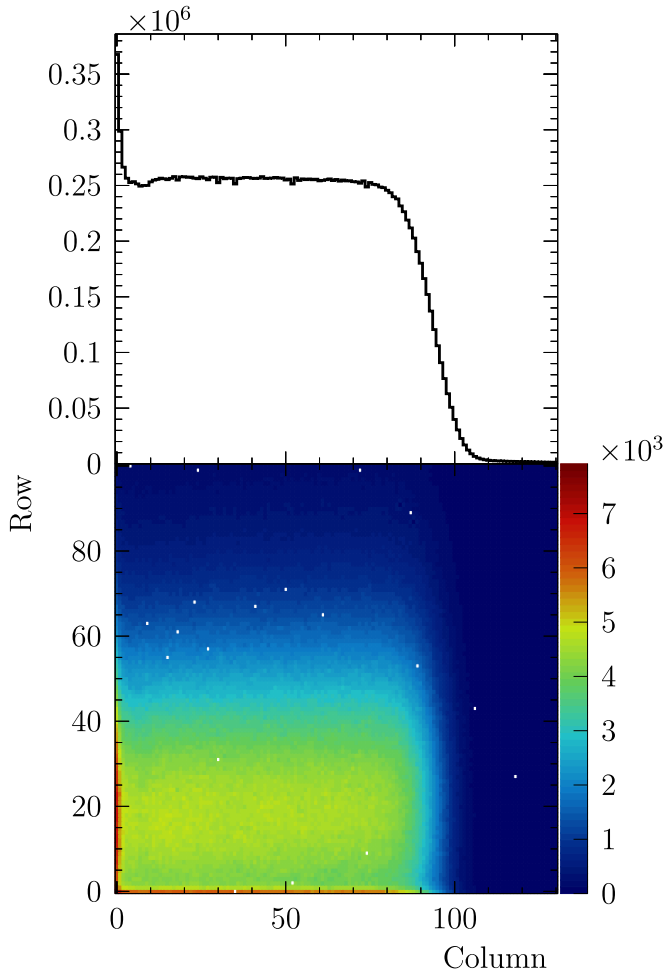



**Figure 7.9:** Hitmap (bottom) and its projection on the columns (top) of a 200 µm thick Micron n-on-p sensor with 450 µm guard rings operated at 200 V.

rings (GR) represents the best scenario. The number of hits is slightly higher for the Micron n-on-n sensor with 450 µm guard rings and grows substantially for the Micron n-on-p sensor with 450 µm guard rings. This suggests that a larger amount of charge is collected from the guard ring region. For the Micron n-on-n sensor with a 250 µm guard rings instead the Landau distribution has a different shape: it is broader with respect to the other columns. It is important to notice that for each kind of sensor such effect on the charge distribution involves only the first column of the pixel matrix. The one but first column has the same distribution as a column far from the edge, without any shoulder.

Using the track intercept position provided by the telescope, it is possible to investigate the excess of hits as a function of the track position on the DUT. Figure 7.11 shows the unbiased residuals in the $x$ direction as a function of the track intercept with the DUT along the $x$ direction. In the plot the solid line represents the edge of the pixel ma-

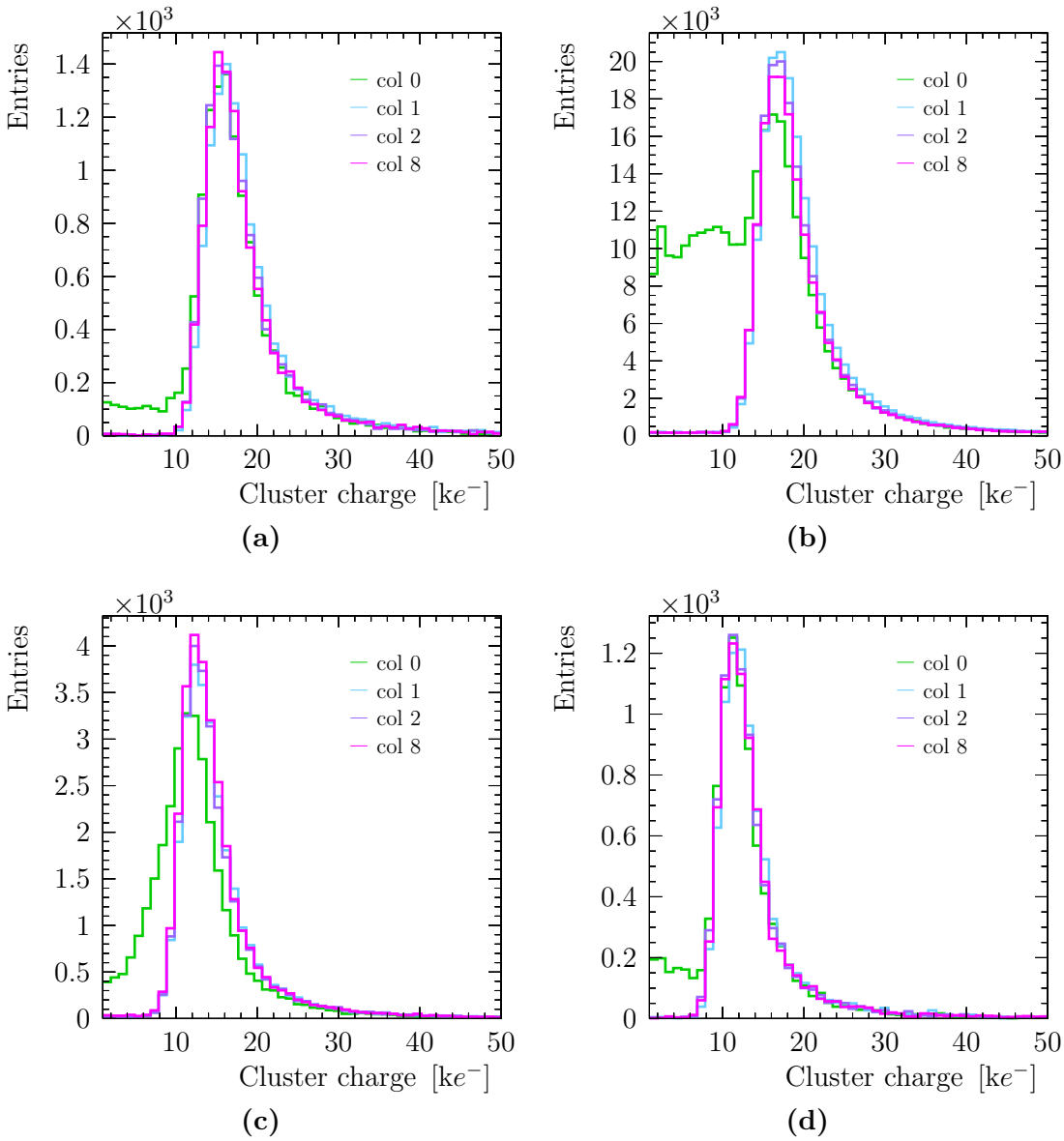



**Figure 7.10:** Cluster charge distribution of hits in different columns for (a) a 200 µm thick HPK n-on-p sensor with 450 µm guard rings, (b) a 200 µm thick Micron n-on-p sensor with 450 µm guard rings, (c) a 150 µm thick Micron n-on-n sensor with 250 µm guard rings, (d) a 150 µm thick Micron n-on-n sensor with 450 µm guard rings.

trix, which corresponds to the ref line in Figure 7.1, while the dashed lines the borders between pixels. Here only the columns closest to the edge are shown, but the pattern is reproduced in the whole sensor. As expected, two pixels clusters are located at the border between pixels and one pixel cluster in the middle of the pixel. It can be seen that there are clusters associated to tracks beyond the edge of the pixel matrix. Among

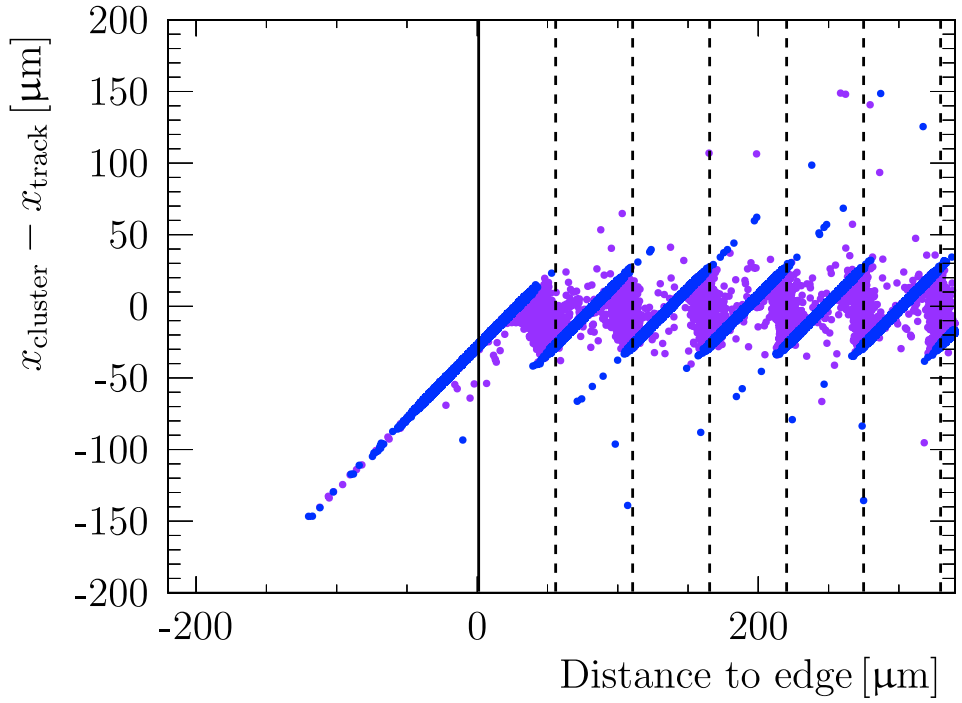



**Figure 7.11:** Unbiased residuals in the $x$ direction as a function of the distance of the track intercept from the edge. The sensor is a 200 µm thick Micron n-on-p sensor with 450 µm guard rings operated at 380 V. Purple markers indicate two pixel clusters and blue markers one pixel clusters. The filled line represents the edge of the pixel matrix and the dashed lines the borders between pixels.

these clusters, 80% are one pixel clusters and 20% are two pixels clusters in one column, due to charge shared between two pixels in the row direction. Therefore the shoulder of the green distribution in Figure 7.10 can be due to tracks passing through the guard ring region, whose charge is collected in the first column of the pixel matrix.

This is verified by studying the cluster charge distribution as a function of the distance of the track intercept from the edge of the pixel matrix. The expected behaviour is to collect charge from tracks traversing the pixel matrix, but not from tracks passing through the guard ring area apart from some diffused charge. Figure 7.12 show that every type of sensor presents a low charge deposit at the edge. The HPK and Micron n-on-n sensors with 450 µm guard rings collect the full charge up to the end of the pixel matrix, where a small deposit at low charges appears due to charges diffusing from the guard ring area. The charge excess for the Micron n-on-p sensor has a linear behaviour and extends beyond the edge, involving an order of magnitude more tracks than the others, while the shape of the charge distribution for Micron n-on-n sensor with 250 µm backside guard rings indicates a loss of charge in the last pixel near the edge. It is important to notice that releasing the cuts on spatial association of clusters with tracks does not lead to a further extension of the charge distribution beyond the edge

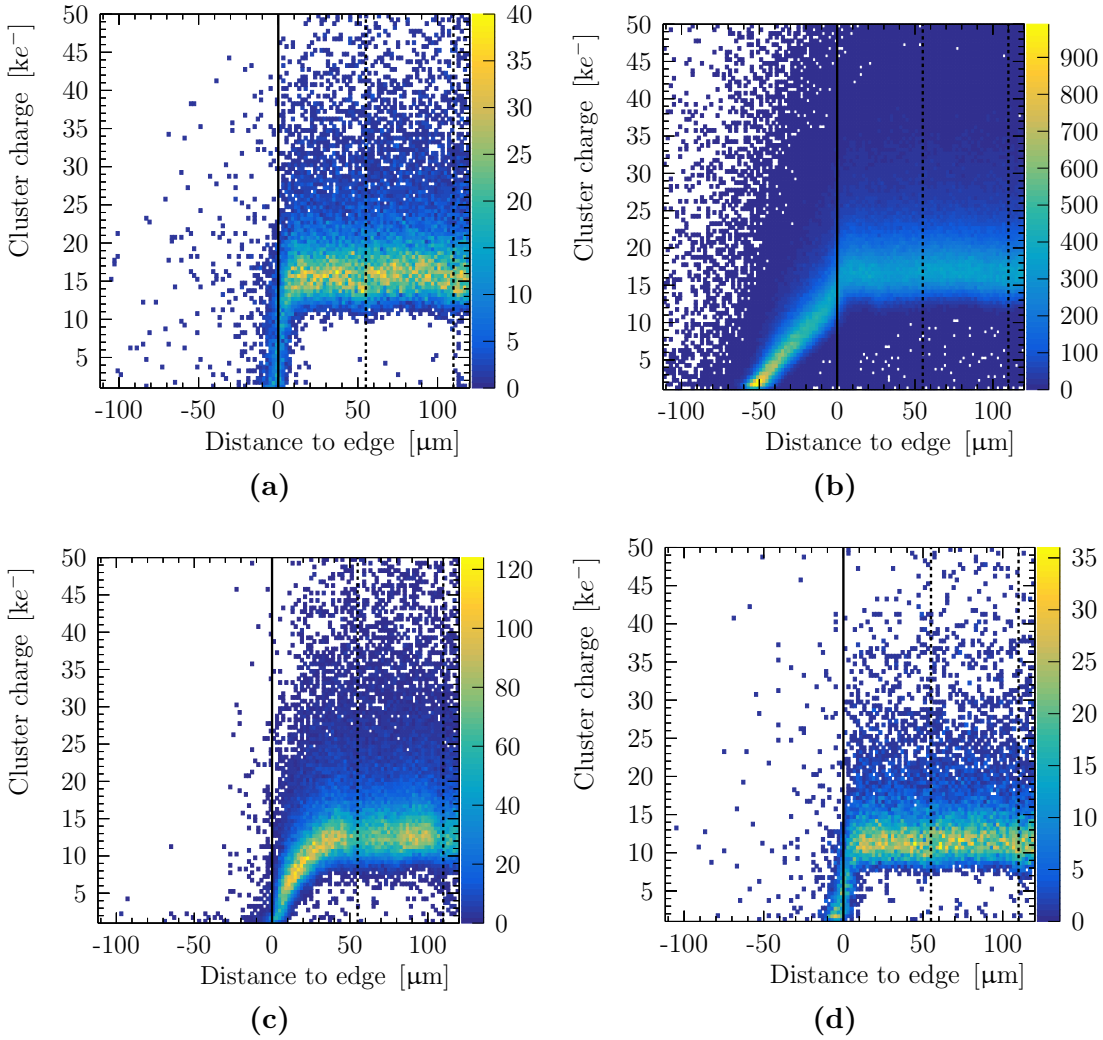



**Figure 7.12:** Cluster charge distribution as a function of the distance to the edge of the associated track intercept for (a) a 200 μm thick HPK n-on-p sensor with 450 μm guard rings, (b) a 200 μm thick Micron n-on-p sensor with 450 μm guard rings, (c) a 150 μm thick Micron n-on-n sensor with 250 μm guard rings, (d) a 150 μm thick Micron n-on-n sensor with 450 μm guard rings. All the sensors are operated at 200 V.

of the pixel matrix. In order to quantitatively compare different prototype sensors the two-dimensional distribution is sliced in fine bins of distance from the edge. The cluster charge distribution per bin is then fitted with a Gaussian convoluted with a Landau, as described in Section 6.1.1. The Most Probable Value (MPV) of the Landau distribution as a function of the track distance from the edge in the $x$ direction can be seen in Figure 7.13 for all the tested nonirradiated assemblies. The uncertainty assigned to the MPV is statistical only, while no uncertainty is assigned to the distance for clarity, but corresponds to the slice width of 5.5 μm. A colour code is employed to distinguish

between the three families of sensors: shades of green for HPK n-on-p, blue for Micron n-on-p and purple for Micron n-on-n. The solid line represents the edge of the pixel matrix and the dashed lines the borders between pixels. All the Micron n-on-p sensors, without exception, collect charge from tracks crossing the sensor beyond the edge. None of the HPK sensors shows this behaviour apart from some charge due to diffusion. Regarding the Micron n-on-n, a difference is observed between the two different guard ring sizes for the two sensors available.

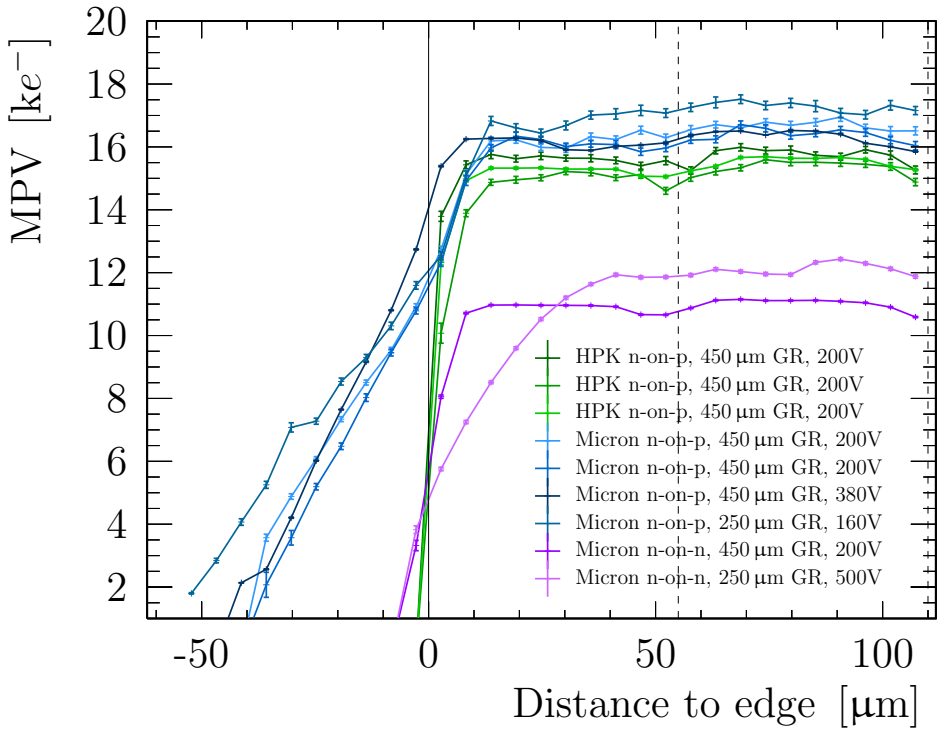



**Figure 7.13:** Cluster charge MPV as a function of the distance to the edge of the associated track intercept for all the nonirradiated assemblies tested. The solid line represents the border of the pixel matrix and the dashed lines the borders between pixels. The label GR states for *guard rings*.

### 7.4.1 Charge Collection Region

In order to further investigate the edge effect of the prototype assemblies, in particular the unexpected behaviour of the Micron n-on-p and Micron n-on-n with 250 μm guard rings sensors, the testbeam programme includes angle scans, focusing on the range between -12° and 12°, where 0° corresponds to perpendicular track incidence.

### Micron n-on-p 450 µm guard rings

Figure 7.14 shows the charge distribution as a function the $x$ position of the intercept of the track with the sensor at three different angles: -12°, 0° and +12°. It can be seen that the shape of the charge distribution changes and that the point of maximum charge collection which is nearest to the physical edge, indicated with a red cross on the plot, shifts towards the guard ring area when increasing the angle.
The simplest geometrical model that can explain this effect is a tilted border between the charge collection region of the guard rings and that of the pixels, as indicated in the sketches on the right of the same figure. In the schemes the black arrow represents the point nearest to the physical edge where the charge deposited by the particle is fully collected, while the shaded area indicates the charge collection region of the guard rings. At -12° the slope of the charge deposit is still present and extends far beyond the edge of the matrix: the charge is fully collected by the sensor up to the point indicated by the arrow, then the sensor starts to pick up charge from the guard ring region. At +12° the slope in the charge deposit almost disappears: the charge is either fully collected by the sensor or fully collected by the guard rings.
The shift of the point of maximum charge collection nearest to the physical edge can also be explained looking at the sketches. The blue line is the 0 of the local $x$ coordinate on the sensor, and the red cross is the position of the intercept between the track and the sensor where the charge deposited by the particle is fully collected by the sensor. Figure 7.16a shows that the shift of the cross from the line is proportional to the tangent of the angle, according to:

$$s = d \times \tan(\text{angle}), \tag{7.3}$$

where $s$ is the shift and $d$ refers to the intercept point of the track with the sensor, which is located in the middle of the active depth; since the sensors are operated at their full depletion voltage or higher, this corresponds to half the thickness. The sign of the slope indicates that the excess of charge occurs at the pixel side, thus we gain charge from the front part of the sensor. The fitted value for $d$ is $96 \pm 7$ µm, which is in agreement with the expectation of half thickness.
The tilted border hypothesis could be verified by a TCAD[2] simulation, but this was not pursued because the details of the guard ring designs is proprietary information.

### Micron n-on-n 250 µm guard rings

The procedure followed for the Micron n-on-n sensor is analogous to the one described in the previous section.
Looking at Figure 7.15, the different shape of the charge distribution and the shift of the point of maximum charge collection nearest to the edge can again be explained to

[2]Technology Computer Aided Design

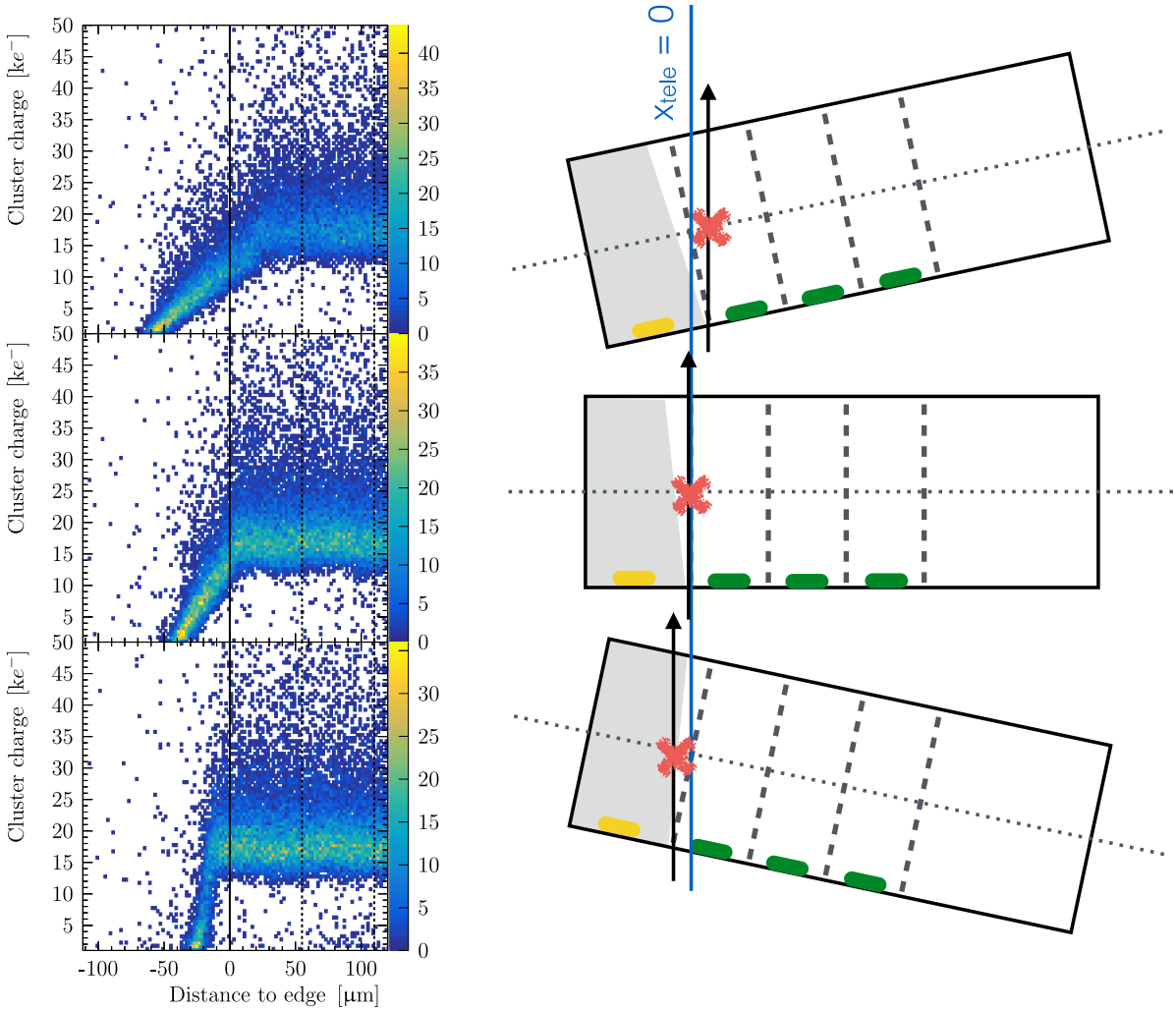



**Figure 7.14:** Cluster charge as a function of the distance to the edge of the track intercept (left) and related sketch (right) for a 200 μm thick Micron n-on-p sensor with 450 μm guard rings operated at 380 V. The sensor is rotated with respect to the beam by −12° (top), 0° (middle) and +12° (bottom).

first order by a tilted border between the collection region of the sensor and the one of the guard ring. Differently from the Micron n-on-p case, the border is tilted in the opposite direction, such that instead of gaining charge from the guard ring area there is a loss of charge in the last pixel. At -12°, even though the slope is very steep, the point of maximum charge collection is not at the edge of the pixel but within the pixel boundary: looking at the related sketch, this is due to the fact that the charge is either fully collected by the last pixel or fully collected by the guard rings. At +12° the charge loss expands a bit more instead, because the charge is fully collected by the sensor up to the arrow, then the sensor starts to lose charge because it is collected by the guard rings.

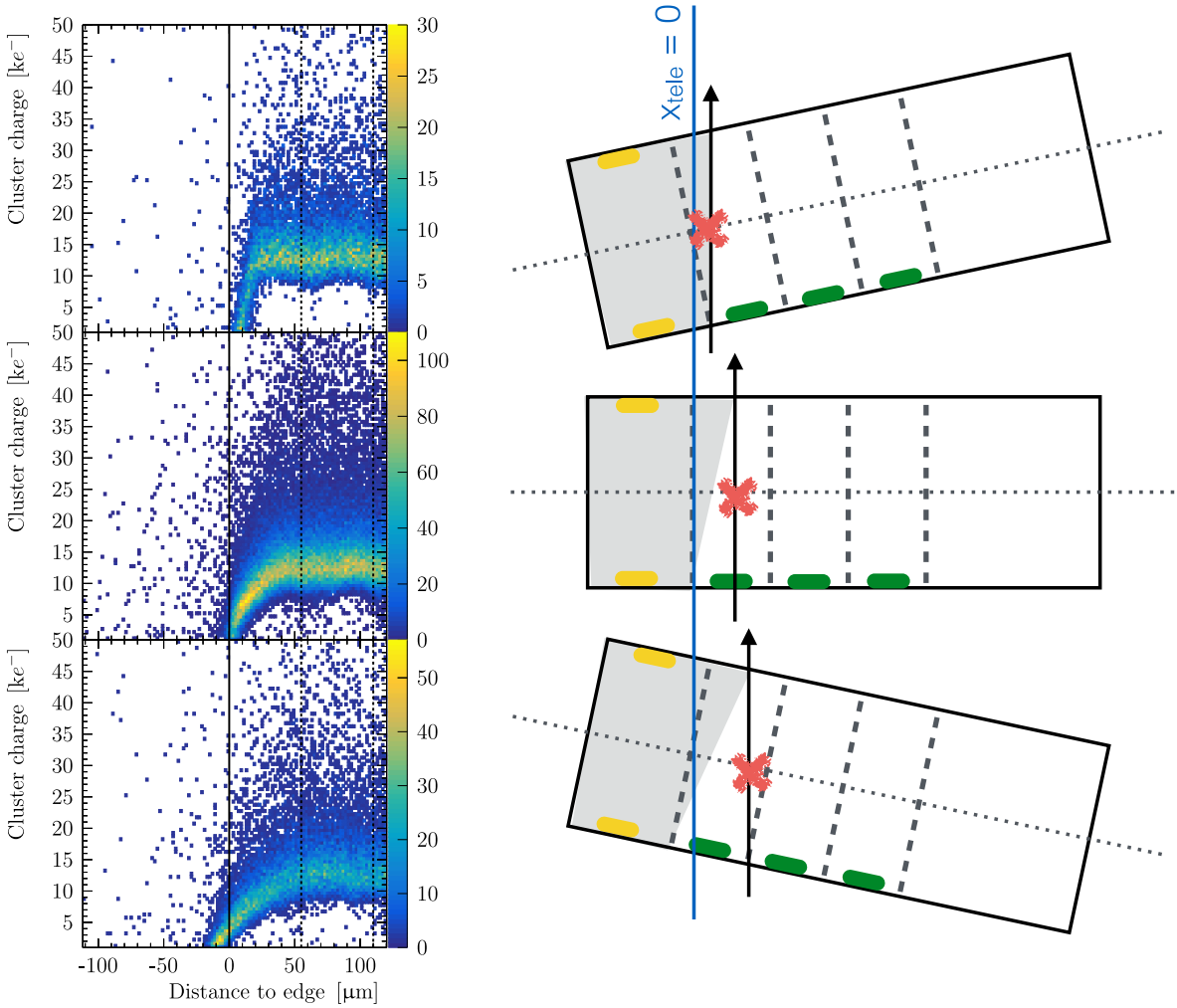



**Figure 7.15:** Cluster charge as a function of the distance to the edge of the track intercept (left) and related sketch (right) for a 150 µm thick Micron n-on-n sensor with 250 µm guard rings operated at 200 V. The sensor is rotated with respect to the beam by −12° (top), 0° (middle) and +12° (bottom).

The point of maximum charge collection nearest to the physical edge shifts in the opposite direction with respect to the n-on-p sensor. Figure 7.16b shows that the slope has opposite sign, indicating that the loss of charge occurs at the backplane side. The fitted value for $d$ is $70 \pm 5$ µm, which is in agreement with the expectation of half thickness.

### 7.4.2 Excess hits investigation

The excess of hits at the edge, as shown in Figure 7.9, has been investigated as a function of the bias voltage applied, track incidence angle and position within the pixel in the

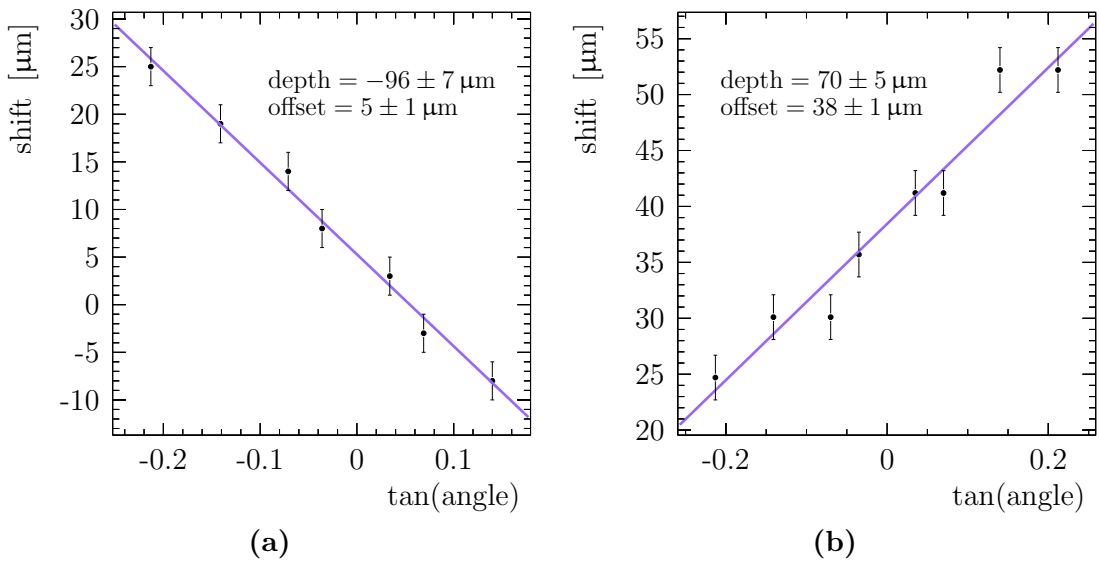



**Figure 7.16:** Position of the point nearest to the physical edge where the charge deposited by the particle is fully collected by the last pixel column/row as a function of the angle between the sensor and the beam. In (a) for a 200 μm thick Micron n-on-p sensor with 450 μm guard rings operated at 380 V and in (b) for a 150 μm thick Micron n-on-n sensor with 250 μm guard rings operated at 200 V.

$y$ direction. Furthermore, the probability for the second column of the pixel matrix to be affected is examined. The excess of hits in the first column of the pixel matrix is computed with respect to the average number of hits away from the edge. These studies, described in the following sections, have been performed using a Micron n-on-p sensor with a 450 μm guard rings.

**Track intercept position in $y$ direction**

The shape of the charge distribution does not depend on the track intercept within a pixel along the $y$ direction, but the excess of hits decreases by about 15% towards the border between pixels, as can be seen in Figure 7.17. This can be attributed to the effect of threshold and charge sharing.

**Bias dependence**

The assembly has been tested at a bias voltage that was much higher than the depletion voltage, in order to check if the excess of hits is related to it. The effect is present at every voltage tested, but the slope of the linear charge deposit beyond the edge becomes smaller when increasing the voltage applied. At 100 V operation voltage (full depletion voltage is around 40 V for Micron n-on-p) the linear region extends up to $\sim$ 70 μm from the edge, while at 380 V it extends up to $\sim$ 50 μm. In the same way the number of excess hits decreases of about 20% while increasing the voltage applied, since the higher the voltage, the stronger the electric field.

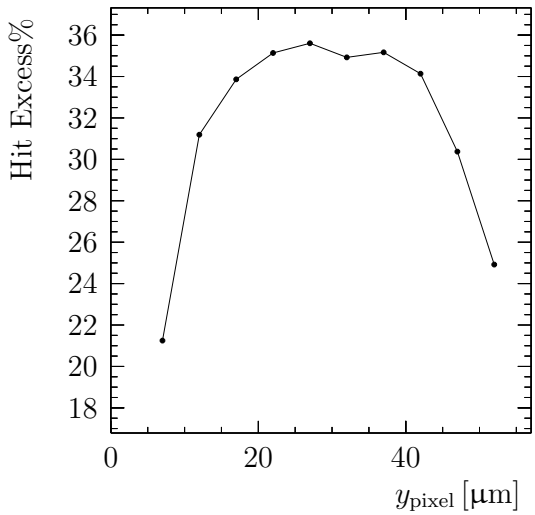



**Figure 7.17:** Excess in the number of hits in the first column of the pixel matrix as a function of the track intercept within a pixel in the $y$ direction. The sensor is a nonirradiated 200 μm thick Micron n-on-p sensor with 450 μm guard rings operated at 200 V.

The higher field leads to faster charge collection, less diffusion and hence smaller excess. The variation of the hit excess as a function of the applied voltage can be seen in Figure 7.18a.

**Angle dependence**

Rotating the sensor with respect to the beam, such that the track incidence angle varies between $-12°$ and $+12°$, leads to a lower excess of hits. The decrease in the excess of hits has different trends when rotating the sensor to positive or negative values with respect to the beam. This can be understood from the sketches in Figure 7.14.

**Excess in the second column**

As can be seen from Figure 7.9, also the second last column shows an excess of hits. Exploiting the telescope information, these hits can be attributed to tracks going through the last column of the sensor. Hence, the distortion of the electric field is such that the second last column collects charge from tracks going through the last column.

**Source cross-check**

As a cross-check, the edge behaviour observed in testbeam data has been investigated using a $^{90}$Sr radioactive source. An excess of hits in the first column and row of the pixel matrix has been observed, as well as the low charge deposit. Data have been collected operating the sensor at different voltages and the hit excess

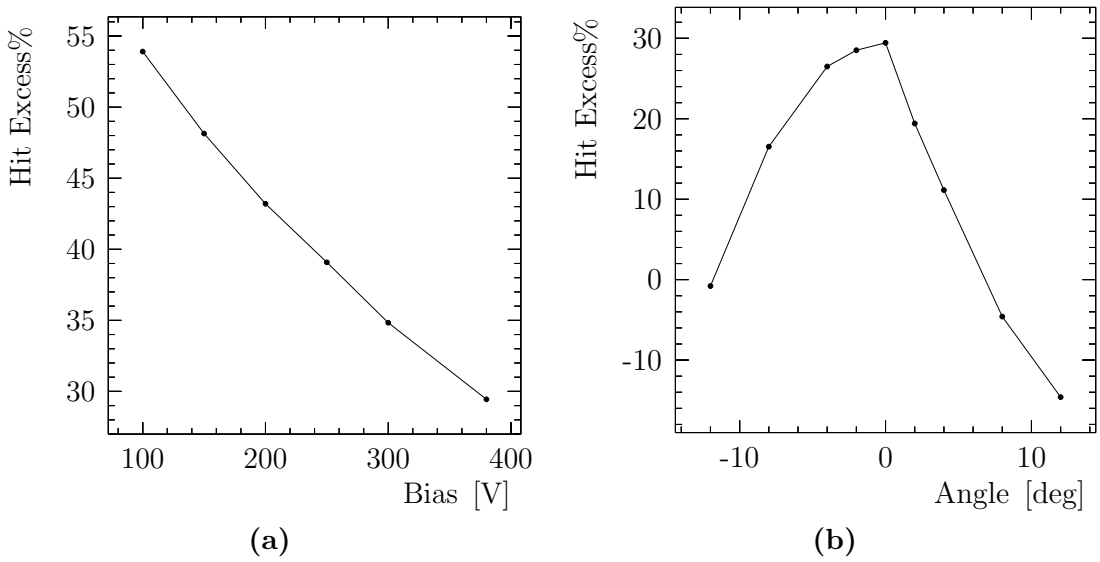



**Figure 7.18:** Excess in the number of hits in the first column of the pixel matrix as a function of (a) the bias voltage applied and (b) the track angle. The sensor is a nonirradiated 200 μm thick Micron n-on-p sensor with 450 μm guard rings.

dependence observed has the same trend as the one obtained from testbeam data. The hit excess is on average lower, most likely due to the fact that tracks from a source do not have a specific direction of incidence. The hit excess decreases when the track is not perpendicular to the sensor.

### 7.4.3 Effect of irradiation

To evaluate the effect of radiation on the edge performance several irradiated sensors have been tested with different fluence profiles. Part of the assemblies have been neutron irradiated at the JSI facility in Ljubljana. The radiation fluence is $8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$, which is the expected maximum fluence at the hottest tip of the sensor after 50 $\mathrm{fb^{-1}}$ of data has been collected, and is uniform along the sensor. The fluence uncertainty is estimated to be of the order of 10%. Part of the assemblies have been nonuniformly proton irradiated in the CERN-PS IRRAD facility. The irradiation profile is such that the irradiation at the edge is $2 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$. The irradiation is expected to reduce the signal because of the change of doping concentration and trapping of the charge carriers, even if this is partially recovered applying high bias voltage. In Figure 7.19 the irradiated Micron n-on-p sensors tested are compared to a nonirradiated sensor of the same type. There is a degradation of the collected cluster charge and the linear charge deposit extending beyond the edge has now strongly reduced for both uniformly and nonuniformly irradiated sensors. This can be linked to the fact that, due to the high irradiation, the sensors are not fully depleted even at very high voltage. The

dips at the border between pixels, which are observed also for the other types of sensors are due to charge sharing in combination with non-zero threshold.
The MPV profiles as a function of the distance to the edge along the $x$ direction for the different types of sensors after irradiation are collected in Figure 7.20. After irradiation by both protons and neutrons with fluences ranging from 2 to $8 \times 10^{15} \cdot 1\,\mathrm{MeV}\,\mathrm{n_{eq}}\,\mathrm{cm}^{-2}$, the behaviour at the edge is comparable among the different prototype sensors.

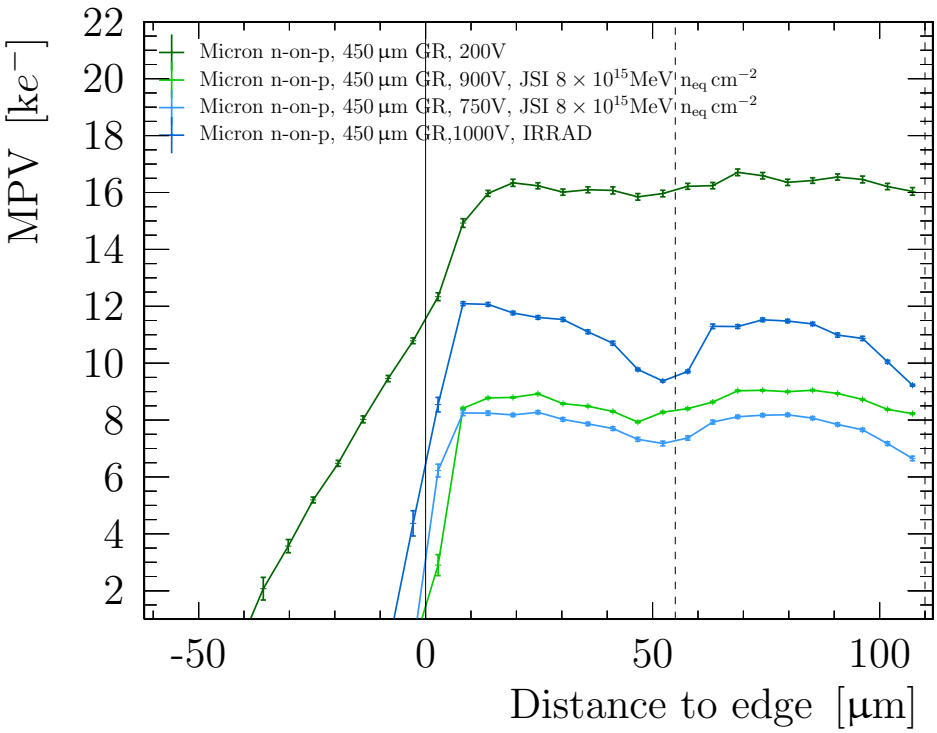



**Figure 7.19:** MPV of the cluster charge as a function of the distance to the edge of the associated track intercept for all the irradiated 200 µm Micron n-on-p sensors with 450 µm guard rings. Two sensors have been uniformly neutron irradiated at JSI and one sensor nonuniformly proton irradiated at IRRAD. A nonirradiated sensor of the same type is shown for comparison. The solid line represents the border of the pixel matrix and the dashed lines the borders between pixels.

### 7.4.4 Summary

A wide range of silicon pixel sensor prototypes has been tested with a charged particle beam using the Timepix3 telescope in order to select the best candidate for the LHCb VELO upgrade. Owing to the precise pointing resolution and timestamp provided by the telescope, a variety of studies has been performed to compare the different sensors. Both HPK and Micron devices respect the specifications in terms of HV-tolerance [161, 162], charge collection and cluster finding efficiency [163], before and after irradiation.

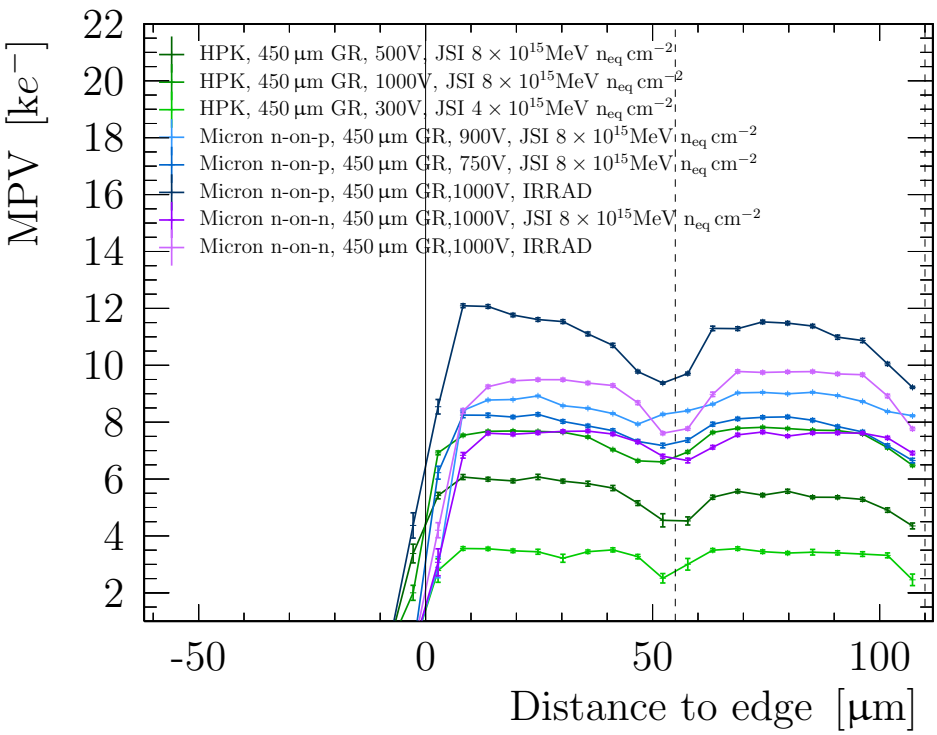



**Figure 7.20:** MPV of the cluster charge as a function of the distance to the edge of the associated track intercept for all the assemblies uniformly irradiated at JSI and nonuniformly irradiated at IRRAD. The solid line represents the border of the pixel matrix and the dashed lines the borders between pixels.

The spatial resolution has been also evaluated and found to be comparable among the different families of prototypes [138]. Nevertheless, a difference has been found regarding the performances at the edge of the pixel matrix. It has been observed that, among all the nonirradiated prototypes tested, two types exhibit an unexpected behaviour. The Micron n-on-p with 450 µm guard rings sensors accumulate charge in the outermost pixel from tracks traversing the sensor beyond the edge of the pixel matrix, while Micron n-on-n with 250 µm guard rings sensors do not collect the full charge in the last pixel. Both effects can be attributed to different guard ring designs which lead to a tilted border between the collection region of the guard ring and the collection region of the pixel matrix. In the case of the Micron n-on-p sensor charge is gained from the guard ring region at the pixel implant side, while in the case of the Micron n-on-n charge from the outermost pixels is lost to the guard ring area. After irradiation the effect is strongly reduced, both at full and half fluence. In view of the VELO upgrade the observed edge effect is critical, since it would increase the occupancy in the part of the sensor where the occupancy is already the highest. Moreover, it leads to a loss of spatial resolution for the first measured point. Hence, it played a crucial role in the decision for the sensor technology. Eventually the studies performed contributed to the decision of adopting

the 200 μm thick HPK n-on-p sensors with 39 μm implant width and 450 μm guard ring size for the VELO upgrade.

# Chapter 8

# Grazing Angles Analysis

The high statistics acquired and the wide range of different kind of prototypes tested during the several testbeam campaigns (Chapter 7) allowed to use and further develop innovative methods for bulk studies of the sensor. The method described in this chapter is known as the *grazing angle* or *edge on* technique [164, 165]. It consists of placing the device under test almost parallel to a beam of minimum ionising particles, such that particle tracks traverse it longitudinally. In this way charge is liberated at a known depth in the bulk material. The advantage of this method is that a large area of the sensor can be probed. The grazing angle technique has many applications. It is one of the few methods known to determine the depletion voltage of a sensor [166]. It has also been applied for charge diffusion studies in the silicon [167] and to perform intrinsic resolution studies [168]. An alternative technique consists of illuminating the sensor from the side with infra-red radiation [169]; the penetration depth of the infra-red light in the energy range near the silicon band-gap ranges from a few μm to a few mm [170], hence centre pixels of large matrices such as in Timepix3 are difficult to probe.
In this thesis the grazing angle method is used to investigate the evolution of the charge collection profile in the bulk of a pixel sensor. Owing to the possibility of simultaneously measuring the signal time-of-arrival (ToA) and Time-over-Threshold (ToT) with the Timepix3 chip, both the charge collection and time required to cross the threshold (time-to-threshold) have been measured as a function of depth. The three families of sensors described in Section 7.2 are compared throughout the chapter, before and after different types and levels of irradiation. Firstly, the selection of the data is outlined in Section 8.1, followed by the description of how depth, charge collection and time-to-threshold and the associated uncertainties have been measured in Section 8.2. A correction is applied to the time-to-threshold measurement to disentangle sensor and chip effects and is described in Section 8.3. The charge collection and time-to-threshold profiles are studied for nonirradiated sensors in Section 8.4 and for sensors irradiated to full fluence in Section 8.5. The impact of radiation damage is further investigated at different fluences by studying devices that were nonuniformly irradiated to full fluence

in Section 8.6.

## 8.1 Data Selection

In the grazing angle setup the sensor is placed almost parallel to the beam such that the particle forms long tracks across multiple adjacent pixels (Figure 8.1). The path length of the track in the sensor, referred to as cluster length, is defined as the number of adjacent columns.

Firstly, the clusters on the Device Under Test (DUT) are associated to telescope tracks.

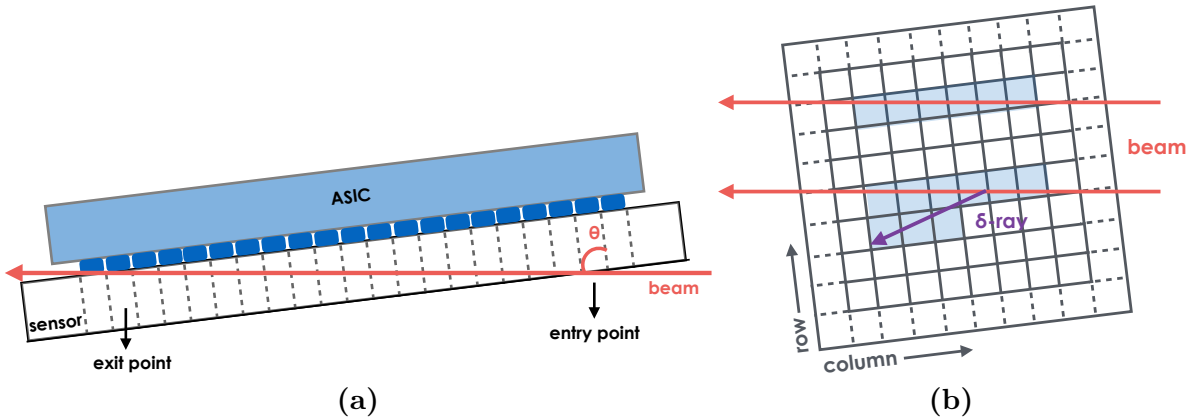



**Figure 8.1:** Illustration of the grazing angle setup: top (a) and front (b) view of the sensor. In (a) the entry and exit point of the track are indicated, as well as the angle $\theta$ of the track with respect to the sensor. In (b) two different types of track are represented, with and without the emission of a $\delta$-ray.

The telescope time resolution is $\sim 0.4$ ns after the time alignment procedure, as described in Section 7.3.3. An excellent time resolution is crucial since good spatial alignment of the DUT with the telescope is not feasible at the large angles used in the grazing setup. The time of a cluster is defined as the earliest timestamp of the pixel hits forming the cluster and the time of a track is as the average of the times of the clusters forming the track. A cluster is considered associated to a track if it is found within a time window of 10 ns around the track. Given the average rate at the CERN SPS of 2M tracks for a spill of 4.5 s, the time cut is sufficient to perform the cluster-track association, expecting a background rate of about 0.5%. Only associated clusters are retained.

Furthermore, among the associated clusters, multiple clusters with the same timestamp and therefore associated to the same track, referred to as broken clusters, are removed. Further selection is then performed according to the following criteria:

- The cluster is allowed to span only one pixel row in order to remove particle tracks with delta-ray emissions or secondary particle production. Figure 8.1b illustrates

two different types of tracks, with and without the emission of a $\delta$-ray. The first kind of track spans only one row and it is retained for the analysis, the latter extends over two rows or more due to the emission of a $\delta$-ray and therefore is removed. A $\delta$-ray can also be released within the same row. In this case the energy deposition is larger, but no cut is applied on the charge to avoid the introduction of a bias in the charge collection profile. This cut rejects $\sim 70\%$ and $< 50\%$ of the tracks for a nonirradiated and irradiated sensor, respectively.

- A gap of up to three empty pixels is allowed within the cluster to account for dead and masked pixels and missing hits because of a signal lower than the threshold.
- The edges of the sensor are excluded: the entry and exit point of the track can not be within three pixels from the edge of the pixel matrix to ensure that the full cluster is within the detector area.
- A cut on the cluster length is applied exploiting the relation between the incident angle $\theta$ and the cluster length:

$$N(\theta) = \frac{\tan\theta \times t}{pitch}, \tag{8.1}$$

where $N$ is the number of pixels forming the cluster, $t$ is the active depth of the sensor and *pitch* is the pixel pitch. A fit to the cluster length distribution for a given angle is performed with a Gaussian distribution, ignoring the tail on the left. Clusters with length outside a one standard deviation window from the fitted mean are rejected.

Data is taken for four sets of angles: 83, 85, 87, 89 degrees. The fitted mean cluster length is then studied as a function of the track angle and fitted with Eq. 8.1, where the pitch is fixed to be 55 µm and $\theta = \alpha + \epsilon$ with $\alpha$ fixed to the known angle and $\epsilon$ allowed to vary to account for a possible offset. The angle offset and the effective depletion depth are obtained as the result of the fit in case of nonirradiated or uniformly irradiated sensors. Th angle offset is found to be of the order of 0.05 degrees.

## 8.2 Fit and Systematics

From Figure 8.1a can be seen that the position of each pixel within the cluster can be associated to a certain depth. Exploiting the relation between the hit pixel position and the depth traversed by the track, it is possible to investigate the charge collection and the time needed for the charge to cross threshold as a function of depth. A 0 µm depth corresponds to the chip side of the sensor and hence to the pixel electrode side, while the full depth (150 µm or 200 µm depending on the device) to the backplane. The closer

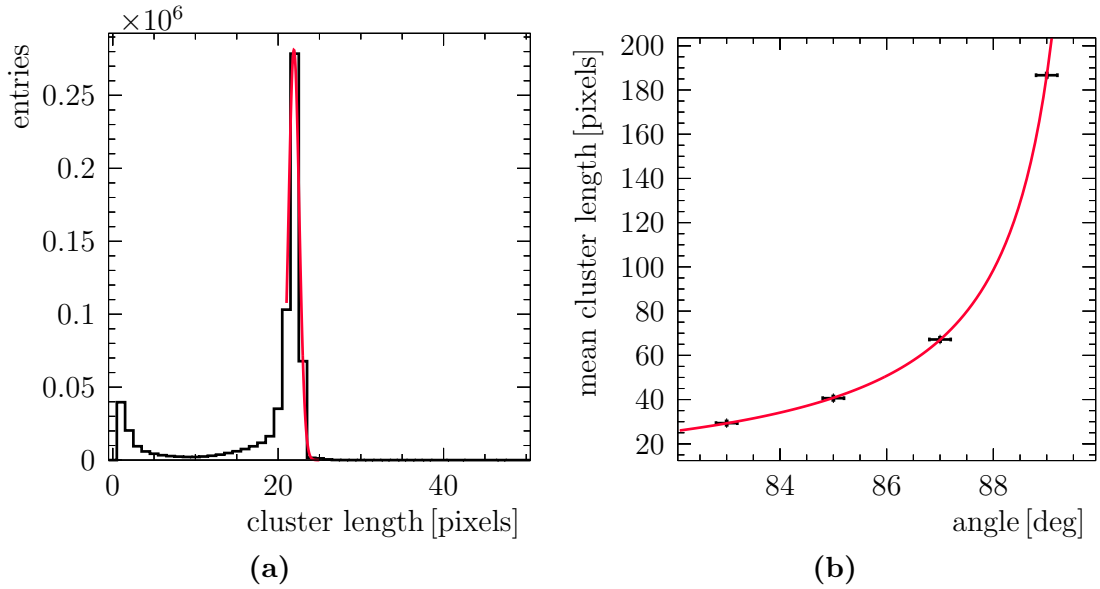



**Figure 8.2:** Example of cluster length fit for a 150 µm thick nonirradiated sensor at 83 degrees with respect to the beam (a) and typical fit to cluster length as a function of angle (b).

is the angle to 90 degrees, the bigger is the cluster length, giving a finer sampling and therefore a better determination of the depth. The plots presented here are obtained with the sensor placed at a 85 degrees angle with respect to the beam, giving a depth step of about 5 µm. In the following, the procedure to determine depth, charge collected and time-to-threshold is described.

**Depth Measurement**

The depth $d(\mathrm{i})$ corresponding to the pixel i can be obtained by inverting Eq. 8.1:

$$d(\mathrm{i}) = \frac{pitch \times N(\mathrm{i})}{\tan\theta}, \tag{8.2}$$

where $N(\mathrm{i})$ is the position of pixel i within the cluster starting from the track entry point. The depth step is assigned as systematic uncertainty to the measured depth and it accounts for the propagation of the uncertainty on the angle offset and a possible missing hit at the beginning or end of the cluster. The effect of the depth uncertainty on the charge and time-to-threshold distributions has been checked and found to be negligible.

**Charge Collection Measurement**

The charge collected is measured by performing a fit with a Landau curve convoluted with a Gaussian, as described in Section 6.1, to the hit charge distribution at a given depth. The value quoted as collected charge is the Most Probable Value

(MPV) of the Landau component. Three sources of systematic uncertainty are considered:

- Charge calibration
  The systematic uncertainty due to the charge calibration is obtained by generating 300 pseudo-experiments to evaluate how the correlated uncertainties of the surrogate function fit parameters affect the MPV. For each pseudo-experiment, a new set of surrogate parameters is generated per pixel based on the initial surrogate function fit parameters and their covariance matrix. The charge of each hit is then computed from the ToT with the new set of parameters and the hit charge distribution per depth is refitted. A pull distribution is obtained per depth by computing the difference between each of the pseudo-experiment MPVs and the initial MPV. The pull distributions are centred at 0 and the RMS of the distribution, removing outliers, is assigned as systematic uncertainty to the MPV. The charge calibration systematic uncertainty $\sigma_{cal}$ is found to be 30 $e^-$ for a nonirradiated sensor and 50 $e^-$ for a sensor irradiated at full fluence.

- Digitisation
  A systematic uncertainty is assigned due to the discrete values of ToT. Per pixel and ToT value a digitisation uncertainty is obtained as:

$$\sigma_{\text{dig}} = \frac{\text{charge/ToT}}{\sqrt{12}}. \tag{8.3}$$

  Figure 8.3 shows the digitisation uncertainty as a function of charge for a pixel in a nonirradiated sensor. The error is constant and of the order of 40 $e^-$ for hit charges higher than $\sim 2500$ $e^-$. This is expected, since for this range of charges the surrogate function (Eq. 7.1) is dominated by the linear term. For low charges, the non-linear term dominates and the uncertainty rapidly increases. Hence the digitisation uncertainty is expected to be significantly higher for irradiated sensors due to the degradation of the charge collection.
  To estimate the effect of such uncertainty on the MPV of the charge distribution, 300 pseudo-experiments have been generated. For each pseudo-experiment, 100 charge values for each pixel are randomly generated from a Gaussian distribution with the initial charge as mean and the corresponding digitisation error as width. The hit charge distribution per depth is then refitted. A pull distribution is obtained by computing the difference between each of the pseudo-experiment MPVs and the initial MPV. A significant shift from zero is observed in the pull distributions. Since the width of the pull distributions is negligible compared to the shift, the shift of the MPV is as-

signed as systematic uncertainty to the MPV. The digitisation systematic uncertainty $\sigma_{dig}$ is assigned to be 20 $e^-$ and 50 $e^-$ for a nonirradiated sensor and for a sensor irradiated at full fluence, respectively.

- Testpulse Amplitude
  To evaluate the uncertainty due to the testpulse amplitude, an alternative technique has been employed to calibrate the detector [155]. The DUT has been exposed to an Americium ($^{241}$Am) source, which features several X-ray lines. The response of the detector has been measured and compared to the calibration obtained from testpulses. The two methods agree within 3%, hence the systematic uncertainty is assigned to be 3% of the measured charge.

The overall uncertainty on the MPV of the charge distribution is given by the sum in quadrature of the three independent contributions and results in 36 $e^-$ for a nonirradiated sensor and 71 $e^-$ for an irradiated sensor. The statistical uncertainty from a Landau convoluted with a Gaussian fit, which is of the order of 5 $e^-$, is negligible.

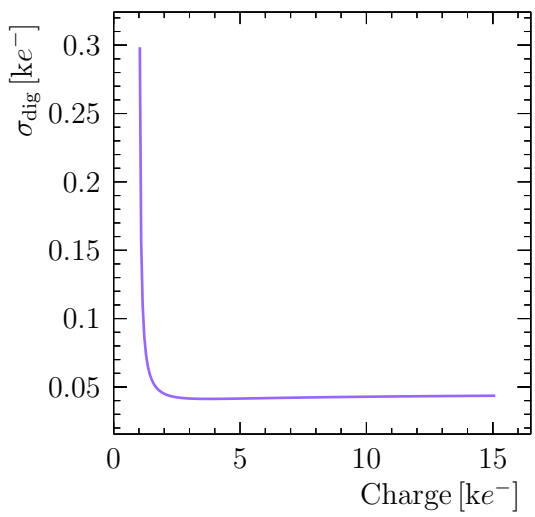



**Figure 8.3:** Example of digitisation error as a function of charge deposited for a single pixel.

### Time-to-Threshold Measurement

The time-to-threshold of a hit is obtained by subtracting the track time provided by the telescope from the hit time. The average time-to-threshold per depth is obtained by fitting the time-to-threshold distribution with a Cruijff [171] function, in order to take into account the asymmetry in the distribution introduced by timewalk. The Cruijff function is a Gaussian with different left-right widths and

non-Gaussian tails:

$$f(x; x_0, \sigma_L, \sigma_R, \alpha_L, \alpha_R) = \begin{cases} \exp\left(-\frac{(x-x_0)^2}{2(\sigma_L^2+\alpha_L(x-x_0)^2)}\right), & \text{if } x < x_0, \\ \exp\left(-\frac{(x-x_0)^2}{2(\sigma_R^2+\alpha_R(x-x_0)^2)}\right), & \text{if } x > x_0, \end{cases} \tag{8.4}$$

where $x_0$ is the mean, $\sigma_{L,R}$ is the left-right width and $\alpha_{L,R}$ parametrises the left-right tail. The left and right resolutions represent the accuracy with which the track is timestamped. The measured resolution $\sigma_{\text{meas}}$ is a combination of the intrinsic resolution $\sigma_{\text{intr}}$ of the sensor and the track time resolution of the telescope $\sigma_{\text{tele}}$, measured to be $\sim 400\,\text{ps}$ [156]:

$$\sigma_{\text{meas}}^2 = \sigma_{\text{intr}}^2 + \sigma_{\text{tele}}^2. \tag{8.5}$$

The intrinsic resolutions are assigned as uncertainties to the fitted time-to-threshold value, ranging from a few ns for nonirradiated sensors to about 10 ns for irradiated sensors.

## 8.3 Timewalk Correction

To understand the properties of the DUT in terms of charge collection time, sensor effects must be disentangled from ASIC effects. The total time taken from when the charge is liberated and the moment the signal crosses the threshold is the combination of multiple factors:

- induction time of the charges drifting in the depleted region;
- diffusion time if the charge is liberated in the nondepleted region;
- time for the integrated charge to overcome the threshold;
- timewalk if the charge collected is small, as described in Section 7.2.2.

The timewalk contribution becomes significant after irradiation, due to the degraded charge collection. Since the process of timewalk is a pure electronics effect, it is necessary to correct for it.
In order to minimise sensor effects and extract the timewalk curve, only charges liberated at small depth, up to $\sim 25\,\mu\text{m}$ from the electrodes, are selected. In Figure 8.4a the timewalk curves for charges liberated at two different depths are compared. In both cases lower charge signals cross threshold later by effects of timewalk and imposed threshold, but it can be seen that charges liberated close to the electrodes and to the backplane give significantly different contributions to the measured timewalk curve. When charges are liberated close to the backplane, the charge collected is in general smaller as a

consequence of the charge trapping due to radiation damage in the bulk of the sensor and the collection time is longer due to the effect of drift and weighting field, as discussed in Section 8.4.2.
The timewalk curve obtained from charges liberated close to the electrodes is validated for some assemblies by comparing it to the timewalk curve determined by injecting a testpulse with known charge in the pixel front-end (Figure 8.4b). The shape variation of the timewalk curve is negligible, leading to the conclusion that the profile obtained is sufficiently representative of a pure electronics effect. The horizontal bar on the hit charge of the testpulse curve is due to the binning and it is not representative of the charge uncertainty.
The curve is fitted for each sensor with the following expression:

$$\mathrm{t}(q) = \frac{A}{q - q_0} + C, \tag{8.6}$$

where $t = t_{\mathrm{hit}} - t_{\mathrm{track}}$, $q$ is the charge, $q_0$ the charge corresponding to the asymptote, $A$ the slope and $C$ the offset. The inverse function is then used to correct the measured time-to-threshold of each hit.

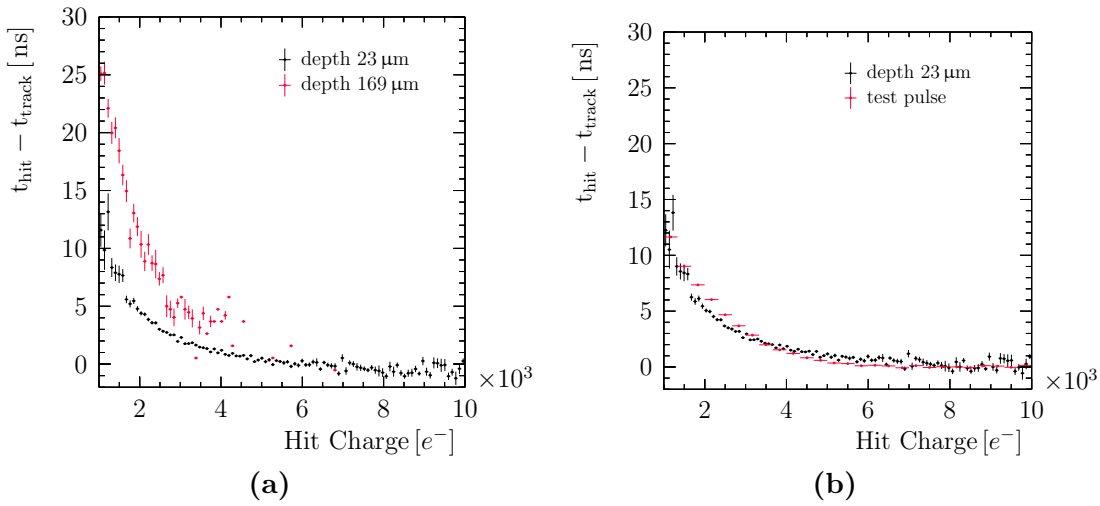



**Figure 8.4:** In (a) timewalk profiles for charges liberated at a depth close to the electrode (black) and close to the backplane (red) and in (b) comparison between timewalk curve from testpulse data (red) and from test beam data (black) with charges liberated close to the electrode. The sensor is a 200 µm thick nonuniformly irradiated HPK n-on-p sensor at 1000 V bias voltage.

## 8.4 Nonirradiated Results

An important check for the validity of the grazing angle technique is to measure the depletion depth for a nonirradiated device operated at a bias voltage above full depletion and compare it to the nominal thickness. The active depth of the sensor is obtained from the fit to the mean cluster length as a function of angle (Figure 8.2b). This procedure can be applied to nonirradiated and uniformly irradiated sensors since the thickness of the depleted region is uniform across the pixel matrix. In Figure 8.5 the measured active depth is studied as a function of bias applied. The depth increases as a function of bias, according to Eq. 6.7, and saturates when the sensor is fully depleted. HPK sensors reach full depletion at $100 - 120$ V and Micron sensors at less than 40 V. The active depth at a voltage above depletion agrees within the error with the nominal thickness and below depletion it is in agreement with the expectation.

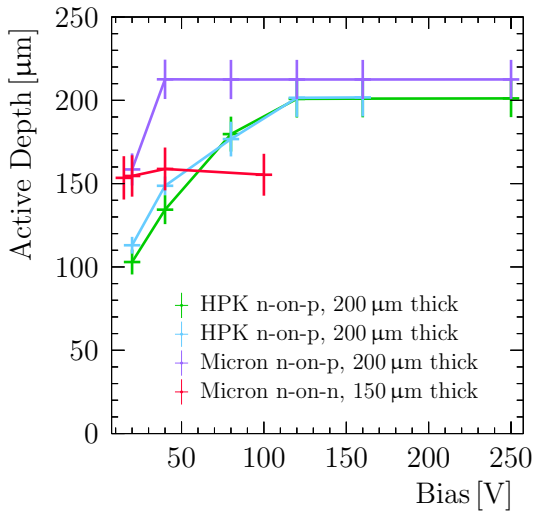



**Figure 8.5:** Active depth as a function of bias for nonirradiated Hamamatsu n-on-p, Micron n-on-p and Micron n-on-n sensors.

### 8.4.1 Charge Collection

The charge collected as a function of depth for a 200 µm thick HPK sensor at different bias voltages is shown in Figure 8.6a. For a nonirradiated sensor and bias voltage above depletion the charge collected is constant and equal to the charge expected for the full thickness of the sensor. At a track angle of 85° the path traversed by the particle is 55.2 µm in each pixel, with the exception of the first and last pixel where it can be shorter. Since the number of electron-hole pairs generated by a minimum ionising particle in silicon is $70 - 80$ per µm [32], the charge deposited per pixel is expected to be between $\sim 3800\ e^-$ and $\sim 4400\ e^-$, which is compatible with the measured value.

At bias voltages lower than 120 V the sensor is partially depleted and the charge linearly drops starting at the border between depleted and nondepleted volume. As can be seen from Figures 8.6a and 8.6c, the charge is collected from the nondepleted volume up to a depth of about 20 µm from the border. This is an effect known as charge migration [172, 173]. If charge is liberated in the nondepleted volume close enough to the border, it can reach the depleted volume by diffusion and then drift to the electrodes.
The linear charge drop suddenly ends at around 2 k$e^-$: no charge is collected below that despite being still significantly far from the threshold of 1 k$e^-$. This is due to the longer collection time needed for the charges to diffuse from the nondepleted region, as discussed in Section 8.4.2.
The Micron n-on-p and Micron n-on-n devices show the same trend, as can be seen in Figure 8.6c and Figure 8.6e, respectively. From the profiles of the three different prototypes it can be seen that the depletion in an n-on-p sensor starts from the electrodes, while in an n-on-n sensor it starts from the backplane.

### 8.4.2 Time-to-Threshold

The time-to-threshold is studied as a function of depth, providing complementary information to the charge collection profile. Figures 8.6b, 8.6d and 8.6f show the result for HPK n-on-p, Micron n-on-p and Micron n-on-n, respectively.
For a fully depleted sensor, the time needed to cross the threshold is less than 5 ns. The time-to-threshold profile exhibits an increase with depth, that can be attributed to a combination of different effects. According to Eq. 6.11, the induced current depends on the drift velocity of the charge carriers, assumed to be constant, the number of charge carriers and the weighting field. Since the weighting field is not uniform, but increases towards the pixel electrodes, most of the signal is induced while drifting near the electrodes. The contribution from holes and electrons depends on the position in the sensor where the charges are generated. If charges are liberated close to the electrodes, electrons drift for a short distance while holes drift for almost the full thickness of the detector. Hence the main contribution is due to the motion of the holes, which have lower mobility and hence longer induction time. On the contrary, if charges are liberated close to the backplane, the main contribution to the induced current is from the motion of electrons. Finally, the time for the integrated charge to overcome the 1000 $e^-$ threshold needs to be taken into account. For a partially depleted sensor, there is an additional contribution to the time-to-threshold from charges migrating from the nondepleted region due to diffusion. The lifetime of electrons and holes in the nondepleted region is expected to be of the order of µs given the doping concentration. The diffusion time extends the time-to-threshold up to 100 ns. Beyond that no hit is registered, due to the time window of 100 ns imposed in the clustering process.
Since the sensor is nonirradiated and the collected charge is higher than 3000 $e^-$, the

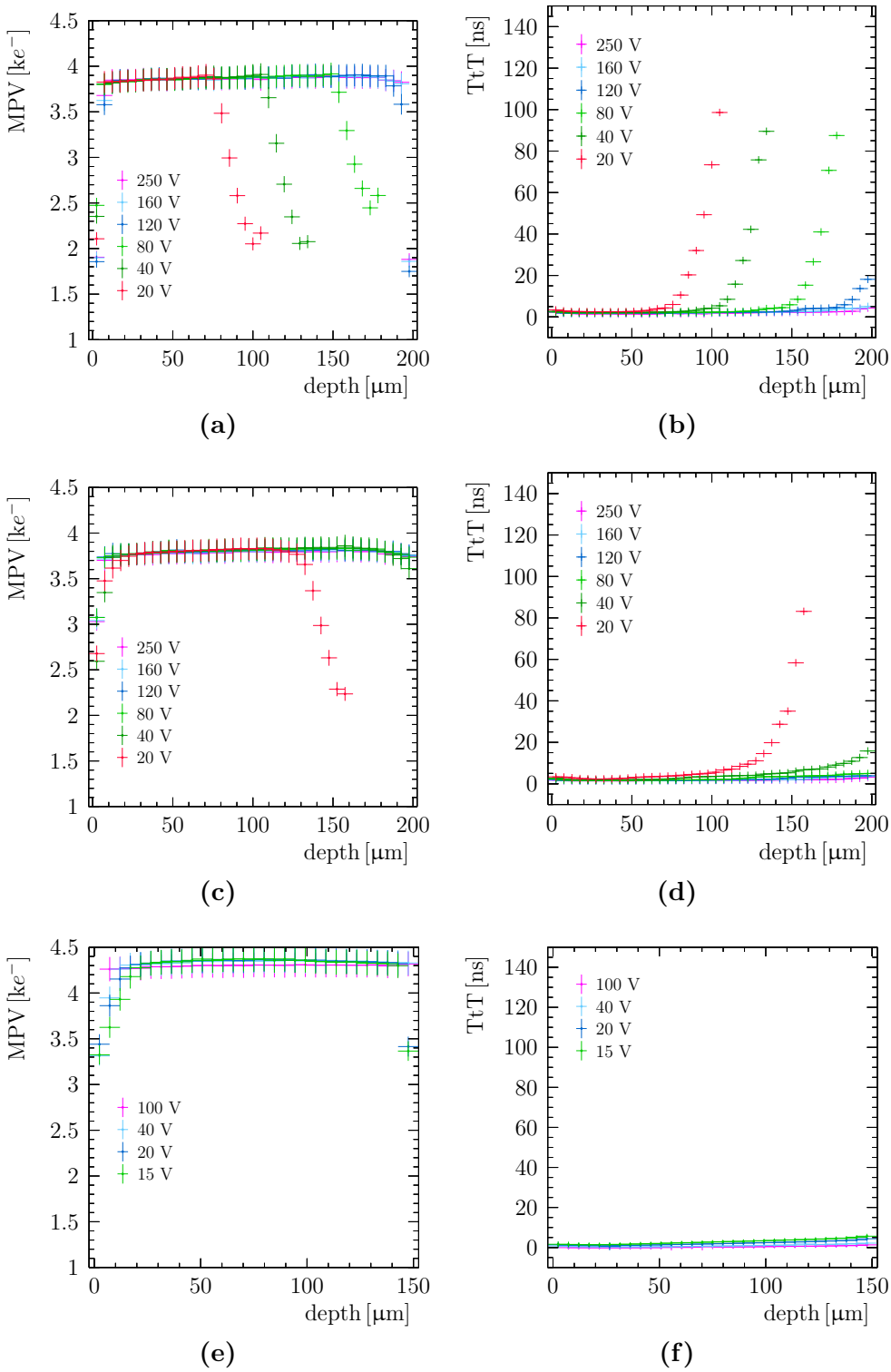



**Figure 8.6:** Charge collected (left) and time-to-threshold (right) as a function of depth for a 200 µm thick nonirradiated HPK n-on-p sensor (a)(b), a 200 µm thick nonirradiated Micron n-on-p sensor (c)(d) and a 150 µm thick nonirradiated Micron n-on-n sensor (e)(f). The 0 µm depth corresponds to the pixel electrodes side and 200 (top and middle) or 150 (bottom) µm to the backplane.

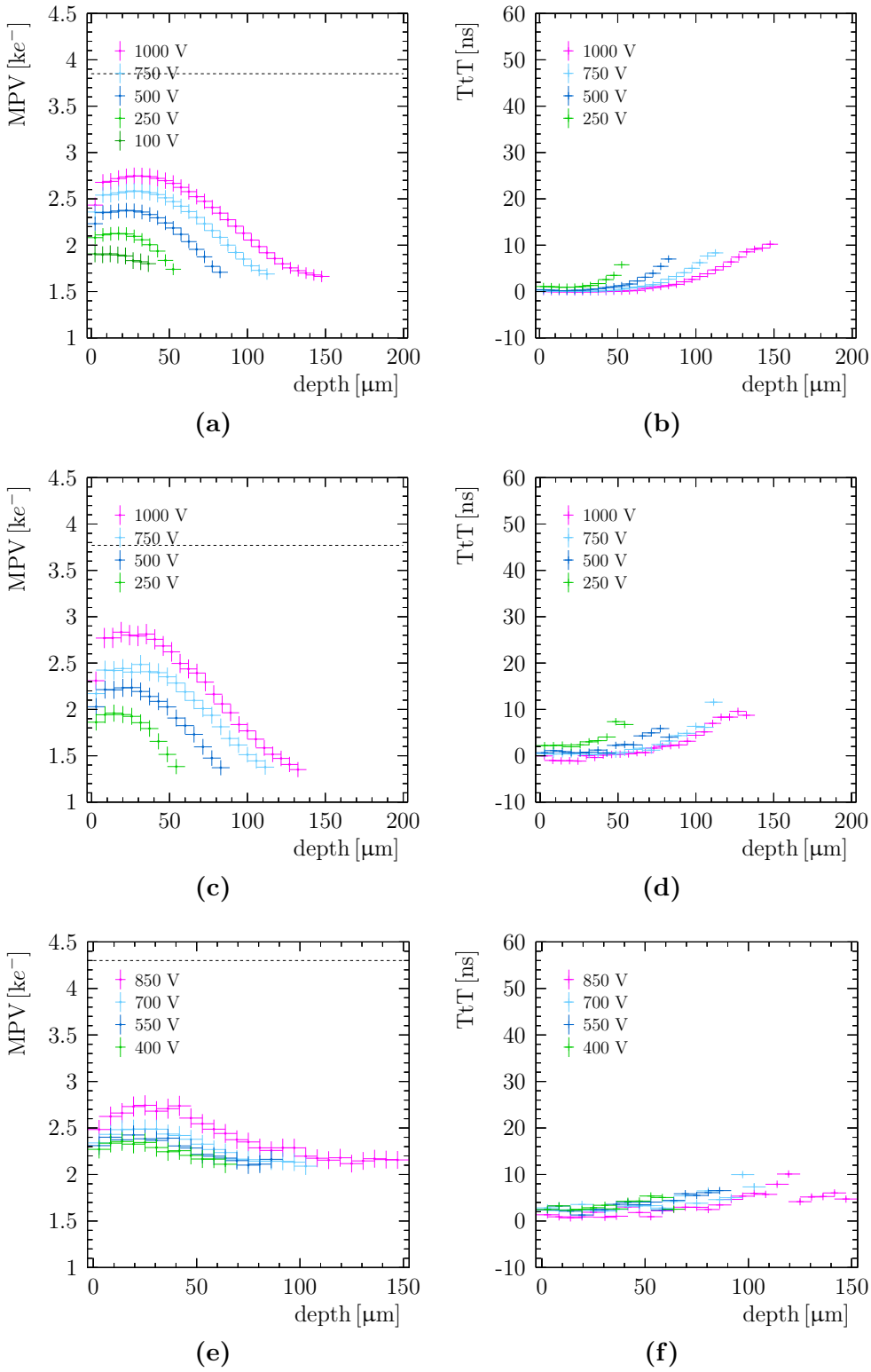



**Figure 8.7:** Charge collected (left) and time-to-threshold (right) as a function of depth for a 200 µm thick HPK n-on-p sensor (a)(b), a 200 µm thick Micron n-on-p sensor (c)(d) and a 150 µm thick Micron n-on-n sensor (e)(f). The 0 µm depth corresponds to the pixel electrodes side and 200 (top and middle) or 150 (bottom) µm to the backplane. The dashed line indicates the charge collected by a nonirradiated sensor of the same type. The sensors are uniformly irradiated to $8 \times 10^{15} \cdot 1\,\mathrm{MeV}\,\mathrm{n_{eq}}\,\mathrm{cm}^{-2}$.

timewalk correction has a negligible effect. It affects only slightly the low charges diffusing from the nondepleted volume. An evidence of the negligible effect of timewalk, is represented by the fact that left and right widths of the time-to-threshold distribution are comparable, as can be seen from Figure 8.8a.

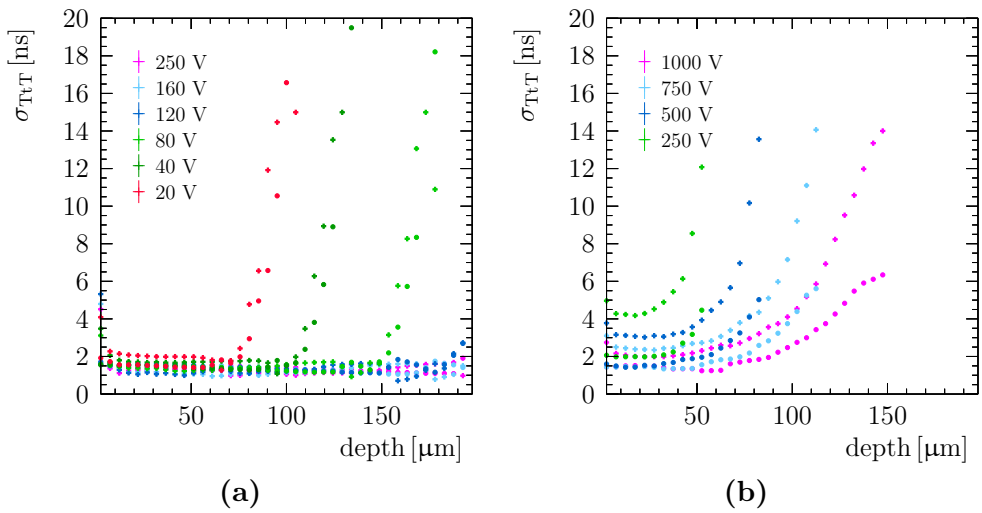



**Figure 8.8:** Time-to-threshold uncertainty on the single measurement as a function of depth for a nonirradiated sensor (a) and a sensor uniformly irradiated to $8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ (b). The dot represents the right width and the cross the left width of the time-to-threshold distribution. The sensor is a 200 μm thick HPK n-on-p sensor.

## 8.5 Uniformly Irradiated Results

The sensors analysed in this section have been neutron irradiated to the full fluence of $8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ at the TRIGA Mark-II reactor at JSI Ljubljana [174]. Since the irradiation is uniform across the sensor, it is possible to extract the depletion depth as a function of voltage per device. Figure 8.9 shows the measured active depth, defined as the depth from which a charge larger than the threshold of 1000 $e^-$ is obtained, as a function of the operating voltage for different prototypes and it can be directly compared with Figure 8.5 to understand the effects of irradiation on the silicon. A smaller active region despite the higher bias voltage applied is macroscopic evidence of the change in effective doping concentration. The effective active depth increases linearly with the bias, but none of the n-on-p sensors reaches full depletion. The n-on-n sensor instead is type inverted at this fluence and is therefore more depleted then the other prototypes at the same voltage, getting to the full thickness at 1000 V.

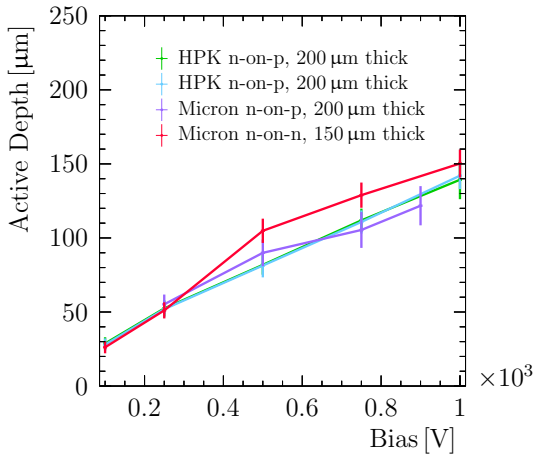



**Figure 8.9:** Active depth as a function of bias voltage for uniformly neutron irradiated HPK n-on-p, Micron n-on-p and Micron n-on-n sensors.

### 8.5.1 Charge Collection

The trend of the charge collected as a function of depth for uniformly irradiated sensors, as illustrated in Figures 8.7a, 8.7c and 8.7e for HPK n-on-p, Micron n-on-p and Micron n-on-n, respectively, is quite different from the behaviour observed for nonirradiated devices.

Firstly, there is an overall signal reduction due to charge trapping, which is not recovered by the increase in bias voltage. Not all the charge liberated at a given depth is collected and this decreases with distance from the electrodes because charge needs to travel over a longer distance and hence has a larger probability to be trapped. The charge profile at each bias voltage shows a slight increase close to the electrodes. This is due to two different effects: the lower field between the neighbouring implants and hole trapping, since for charges liberated close to the electrodes the current is mainly induced by the motion of holes. Secondly, no charge is collected from the nondepleted volume. This can be attributed to charge trapping and slow drift in combination with the integration time of the front-end. The time for the integration of the signal is limited, hence the discharge can start while still in the process of integrating; this is especially relevant for a small amount of charge. Figure 8.10 shows the number of hits recorded as a function of depth for the HPK n-on-p sensor, confirming what is observed in the charge collection profile. The number of hits decreases with depth and quickly drops approaching the nondepleted region. No effect of a doubly peaked electric field according to the double junction model for highly irradiated sensors, described in Section 6.2, is observed. This can be attributed to a combination of the small amount of charge and the low weighting field at the backside of the sensor.

The percentage of charge loss due to irradiation per depth is illustrated in Figure 8.11 for the two different types of n-on-p sensors. The charge loss is obtained from the ratio

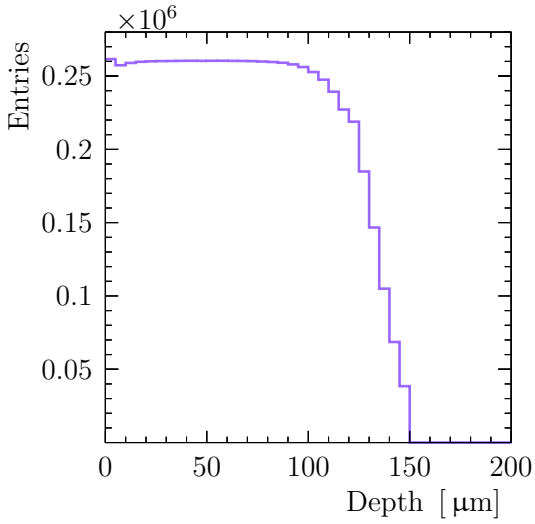



**Figure 8.10:** Number of hits as a function of depth for a uniformly irradiated 200 µm thick HPK n-on-p sensor.

of the charge collected between a device irradiated to $8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ operated at 1000 V and a nonirradiated one of the same type fully depleted (250 V). The percentage of charge loss after full irradiation varies with depth between about 25% close to the electrodes and 60% at the border of the active region. These values are compatible to what reported in literature for similar fluences [175]. The HPK n-on-p and Micron n-on-p devices show a similar trend: the charge loss increases with depth up to the active volume of the sensor after irradiation, while nothing is collected from the nondepleted volume. The fraction of charge collected ($1 - \mathrm{charge}_{loss}$) integrated over the full thickness has been compared to the fraction of charge collected by a sensor placed perpendicular to the beam and they are found to be in agreement within 5%.
Since charge calibration data for JSI irradiated Micron n-on-p and Micron n-on-n are not available[1], the results presented for these two sensors in this section and Section 8.5.2 have been obtained by using the most heavily irradiated region of nonuniformly irradiated Micron n-on-p and n-on-n sensors, respectively, irradiated at IRRAD as described in Section 8.6. For the Micron n-on-n sensor the threshold settings differ between data taking (700 $e^-$) and testpulse calibration (1000 $e^-$), hence to compare 300 $e^-$ have been added to the measured charge.

### 8.5.2 Time to Threshold

The time-to-threshold profiles for sensors irradiated to full fluence are shown in Figures 8.7b, 8.7d and 8.7f for HPK n-on-p, Micron n-on-p and Micron n-on-n, respectively. The trend is similar for the different prototypes: the time to cross the threshold

[1]The devices are no longer operational.

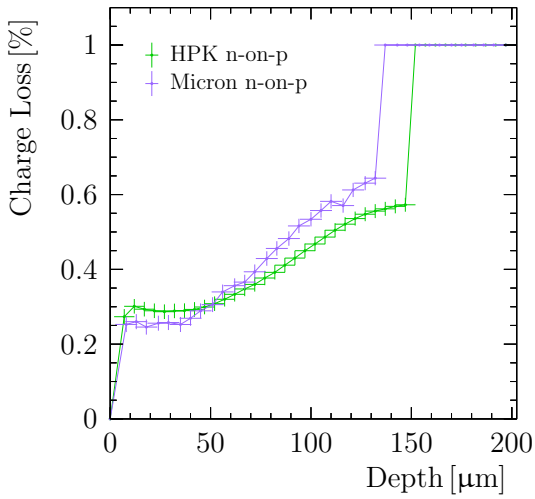



**Figure 8.11:** Percentage of charge loss as a function of depth for a HPK n-on-p and a Micron n-on-p sensors.

increases with depth, up to 10 ns, independent of the voltage applied. Differently from the nonirradiated case, there is no additional contribution due to charges diffusing from the nondepleted region, since these charges are not collected as described in Section 8.4. However, the contribution of timewalk becomes significant, since the charge collected is lower than 3000 $e^-$. The timewalk effect also broadens the time-to-threshold distribution, leading to an asymmetric uncertainty on the single measurement that varies from 3 ns up to 15 ns depending on the depth, as can be seen from Figure 8.8b.
In Figure 8.12a the time-to-threshold profile for a HPK n-on-p sensor at 1000 V is compared to the profile obtained by applying the timewalk correction described in Section 8.3. The timewalk curve is fitted and a correction to the time-to-threshold of each hit is applied as a function of charge. The correction increases as the charge decreases, hence with depth, leading to a smaller mean value and narrower distribution compared to the not corrected case. The corrected time-to-threshold spread results in less than 3 ns along the whole sensor depth and the uncertainty spans from $\sim$ 1.5 ns at small depth up to $\sim$ 4 ns at the border of the active region.
In Figure 8.12b the corrected profiles at the highest (1000 V) and lowest (250 V) voltage available are compared to the time-to-threshold profile for the same type of sensor before irradiation and fully depleted. The profiles agree within the uncertainties. The main difference with respect to the nonirradiated case is represented by the larger uncertainty on the single measurement: close to the electrodes the uncertainty is comparable, while at the border of the active region is almost double, due to possible residual timewalk.

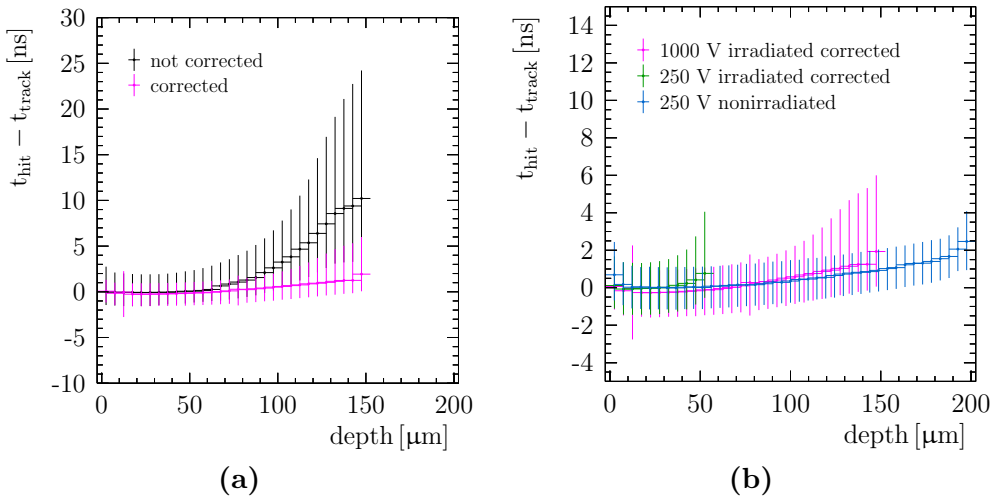



**Figure 8.12:** In (a) comparison between time-to-threshold profiles before and after timewalk correction. In (b) comparison between time-to-threshold profiles for a nonirradiated sensor operated at 250 V and for a uniformly irradiated sensor operated at 250 V and 1000 V. The sensors are 200 μm thick HPK n-on-p sensors. The uncertainty is assigned as uncertainty on the single measurement, as described in Section 8.2.

## 8.6 Non Uniformly Irradiated Results

The sensors analysed in this section have been irradiated with 24 GeV protons in the CERN-PS IRRAD facility [176]. The irradiation of these assemblies is nonuniform, following the shape of the illumination by the proton beam. The nonuniform irradiation profile is representative for the expected condition of the upgraded VELO and allows to perform accurate studies of the charge collection as a function of fluence. All the assemblies have been tested without additional controlled annealing, with the exception of the Micron n-on-n sensor that underwent controlled annealing for 80 minutes at 60°C.

### 8.6.1 Fluence Profile

The beam profile of the IRRAD proton beam was reconstructed combining the activity map of the assembly itself after irradiation and the dosimetry measurement provided by the facility. The most reliable map of the sensor activity is obtained from the out of spill hitmap. Background hits from the activated assembly are always present, while beam tracks only during the spill. Figure 8.13(a) shows the number of hits as a function of time: the peaks (in purple) correspond to the beam spills, while the flat background (in blue) is visible in between spills. Only the latter hits are selected to produce the hitmap. The dosimetry measurement (Figure 8.13(b)) has been performed attaching an aluminium foil to the back of one of the ceramic boards. The foil was cut in six

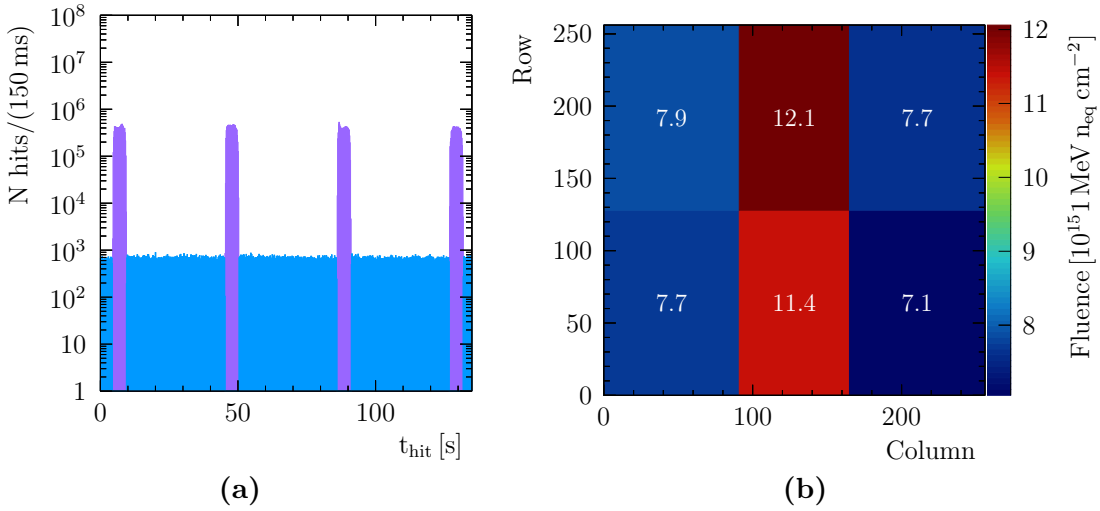



**Figure 8.13:** Number of hits as a function of time (a) and dosimetry results for the sensors irradiated at IRRAD (b).

pieces and for each one of these there is a measurement of the average fluence with 10% accuracy. The conversion from proton to neutron equivalents is done using a hardness factor[2] of 0.62 [148]. The three assemblies under study have been irradiated at the same time and were oriented perpendicularly to the beam, but the profile centre can be shifted with respect to the sensors. The beam profile is then reconstructed through the following steps. A cluster size smaller than five and a hit ToT less than 15 counts are required to select only photon and electron clusters coming from the decay of activated elements in the chip. The out of spill map is then fitted with a 2D Gaussian to get the centre of the profile per sensor and the width along columns and rows. Since the beam should have the same profile on all sensors, the width is fixed to the average width among all sensors. The overall normalisation is obtained minimising the $\chi^2$ with respect to the foil (including the 10% uncertainty from the dosimetry measurement).
The reconstructed profile for the HPK n-on-p sensor is shown in Figure 8.14a. After irradiation the activated sensors were subject to the grazing angle studies at the CERN SPS H8 beamline.

### 8.6.2 Charge Collection and Time To Threshold

The fluence varies by about a factor four between the most and the least irradiated areas of the sensor. Therefore charge collection and time-to-threshold have been investigated splitting the pixel matrix in 16 elliptical rings of constant width that follow the shape of the beam, as shown in Figure 8.14b. The clusters are assigned to the region

[2]Energy dependent scale factor to convert the radiation damage from a certain type of particle to the equivalent damage for 1 MeV neutrons.

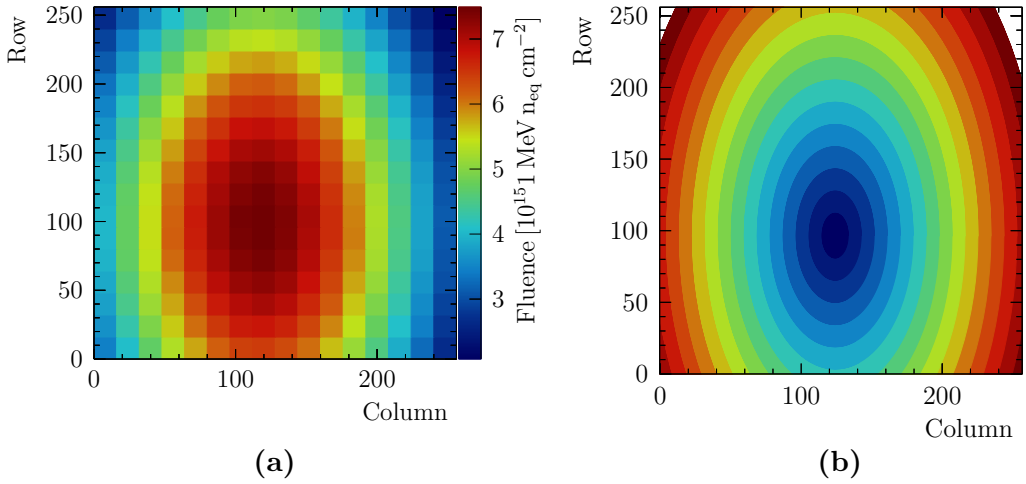



**Figure 8.14:** Reconstructed fluence profile in (a) and adopted fluence binning in (b).

depending on the exit point of the track. Figures 8.16a and 8.16b show how the charge and time-to-threshold profiles change as a function of depth from the most internal ring (dark blue) to the most external one (dark red) for a HPK n-on-p sensor operated at 1000 V. The depletion depth increases when the fluence decreases, as expected. The shape of both profiles changes: charges collected in the most internal region follow the typical shape observed for uniformly irradiated sensors (decreasing charge collection and increasing time-to-threshold due to lower average charge), which then tends to flatten out moving towards the edges of the device, where the low fluence corresponds to a lower trapping rate and larger depletion depth. The saturation of the curves at large depths is due to the fact that the charge distribution approaches the threshold and part of the distribution is cut away. This leads to a enhancement of the tail of the Landau distribution and consequently the most probable value of the distribution is pushed to higher values. Wiggles in the profile are due to the low statistics available. In this plot the error bars are omitted for clarity.

The Micron n-on-p sensor has a similar behaviour to the HPK n-on-p. For the Micron n-on-n sensor instead the span between charge liberated close to the backplane and close to the electrodes is smaller leading to a flatter profile compared to the other type of sensors. The same difference is observed in the time-to-threshold profile.

The timewalk curve used to correct the time-to-threshold profile is obtained using hits collected from the whole pixel matrix after verifying that there is no difference in the shape of the curve between hits from the most and least irradiated region. In Figure 8.15 the time-to-threshold profiles from different fluences after timewalk correction are compared and no significant difference is observed.

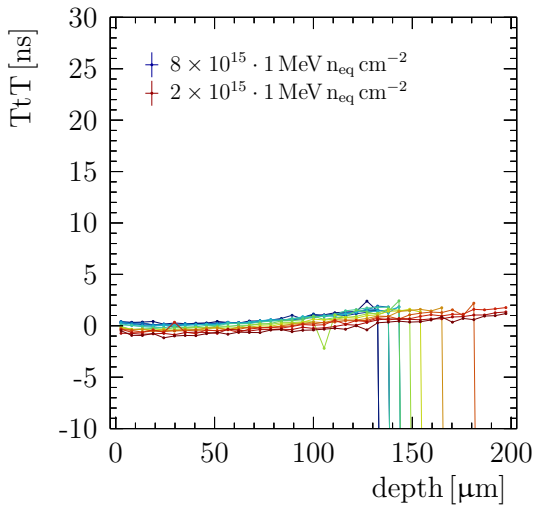



**Figure 8.15:** Timewalk corrected time-to-threshold as a function of depth for different fluences. The fluence increases from red to blue and the lowest and highest values are indicated. The sensor is a nonuniformly irradiated 200 μm thick Micron n-on-p sensor. Error bars are omitted.

### 8.6.3 Fluence Dependence

Given the fluence profile, it is possible to investigate the charge and collection time as a function of fluence. Figure 8.17a shows an example of charge profile at 83 μm depth, obtained by splitting the sensor in 16 × 16 bins and extracting the MPV of the charge distribution per bin. The profile resembles the irradiation profile of Figure 8.14: the most irradiated regions are less depleted due to the change in effective doping concentration, hence the charge collected is about 50% smaller. The charge collection for charges liberated at different depths in the bulk of the sensors is studied as a function of fluence at different operation voltages in Figure 8.18 for a HPK n-on-p sensor. It can be seen that at low voltage the charge is collected only up to ∼ 90 μm depth from the electrodes and it decreases as a function of fluence. Increasing the bias voltage, the charge collected at a certain depth is higher and charge is collected from deeper in the sensor, up to ∼ 170 μm depth at 1000 V.

At depths smaller than 80 μm, there is a sizeable deviation from the trend in the highest fluence region: the charge collected exhibits an increase of ∼ 15%. This effect, which happens only at 1000 V bias, is recognised as the so called charge multiplication or avalanche effect [177–179]. This mechanism occurs in highly irradiated silicon sensors, which can be biased at high voltage without breakdown, creating locally a high enough field to give avalanche breakdown. The charge multiplication is localised close to the electrodes where the electric field is the highest. If charges are liberated in the high field volume (close to the electrodes) they will give rise to the multiplication effect, while most of the charges liberated deeper in the bulk will experience charge trapping

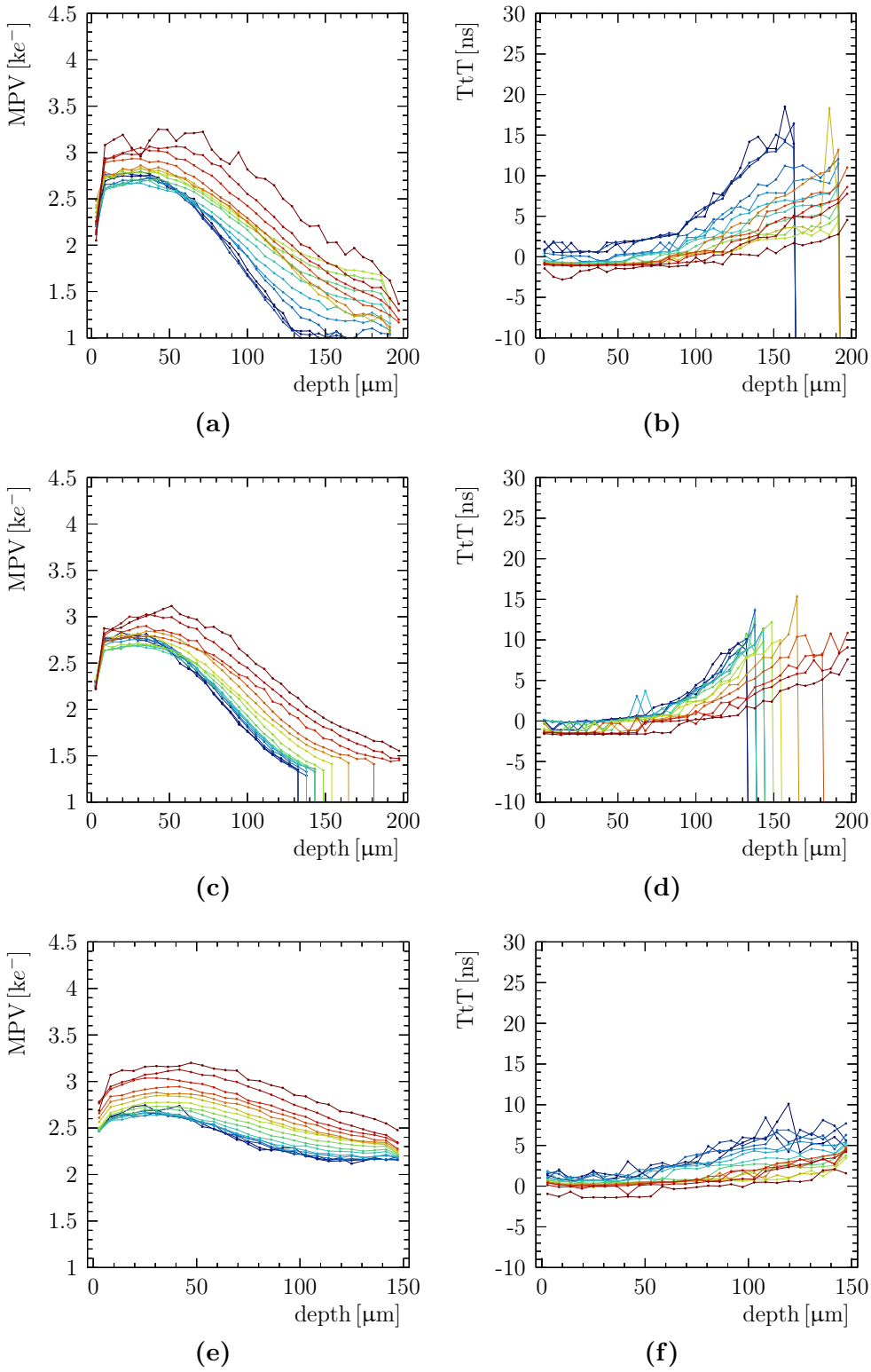



**Figure 8.16:** Charge collected (left) and time-to-threshold (right) as a function of depth for a 200 µm thick HPK n-on-p sensor (a)(b), a 200 µm thick Micron n-on-p sensor (c)(d) and a 150 µm thick Micron n-on-n sensor (e)(f). The sensors are non uniformly irradiated up to full fluence, with increasing fluence from red ($2 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$) to blue ($8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$). The error bars are omitted. Colour coding according to the rings in Figure 8.14b.

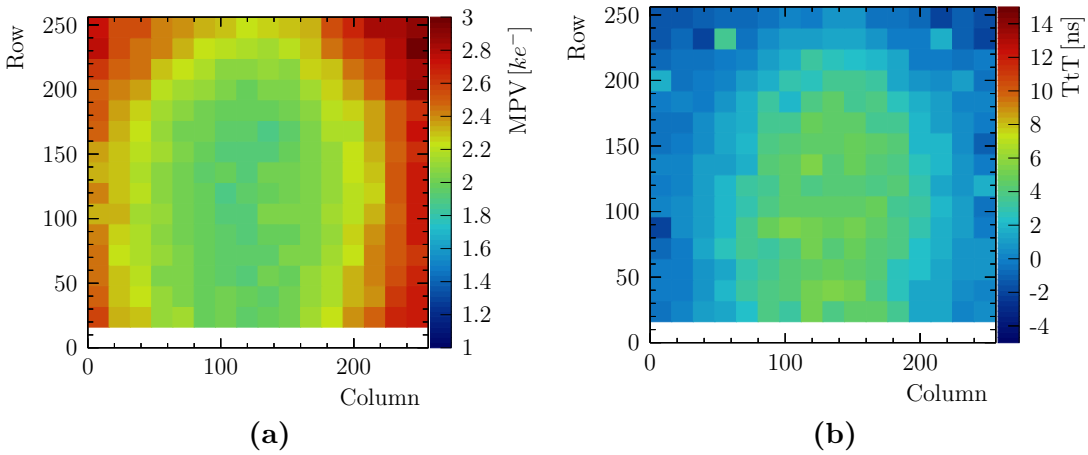



**Figure 8.17:** Example of (a) charge collection profile and (b) time-to-threshold profile. The sensor is a HPK sensor nonuniformly irradiated up to full fluence.

before reaching the high field region. The remaining charges still cross the high field region, but the effect is probably too small to be seen as clearly as for charges liberated close to the electrodes. Avalanche multiplication occurs in both Hamamatsu n-on-p and Micron n-on-p sensors at 1000 V, while the Micron n-on-n sensor does not present any effect even at the highest voltage tested. The charge multiplication effect has not been observed for uniformly neutron irradiated sensors, even when operated at 1000 V. This can be attributed to the different nature of irradiation leading to different damage in the silicon [180].

Another difference between n-on-p and n-on-n sensors is that the n-on-n sensor depends less on fluence, as can be seen from Figure 8.19a. The charge collected by the last pixel, at 150 µm depth, is examined as a function of fluence for different operation voltages in Figure 8.19b. It can be seen that the charge collected never reaches the threshold value of 1000 $e^-$ and charge is collected up to the full thickness even at 400V, confirming that Micron n-on-n has a larger active depth compared to the other types of sensors.

## 8.7 Summary

The grazing angle technique proves to be a powerful method to study charge collection efficiency properties in the bulk of the sensors, like depletion depth and charge collection. Moreover, exploiting the simultaneous measurement of ToA and ToT provided by the Timepix3 chip, it was possible to investigate the time-to-threshold profile and the different factors contributing to it.

For nonirradiated sensors the most probable value of collected charge is constant as a

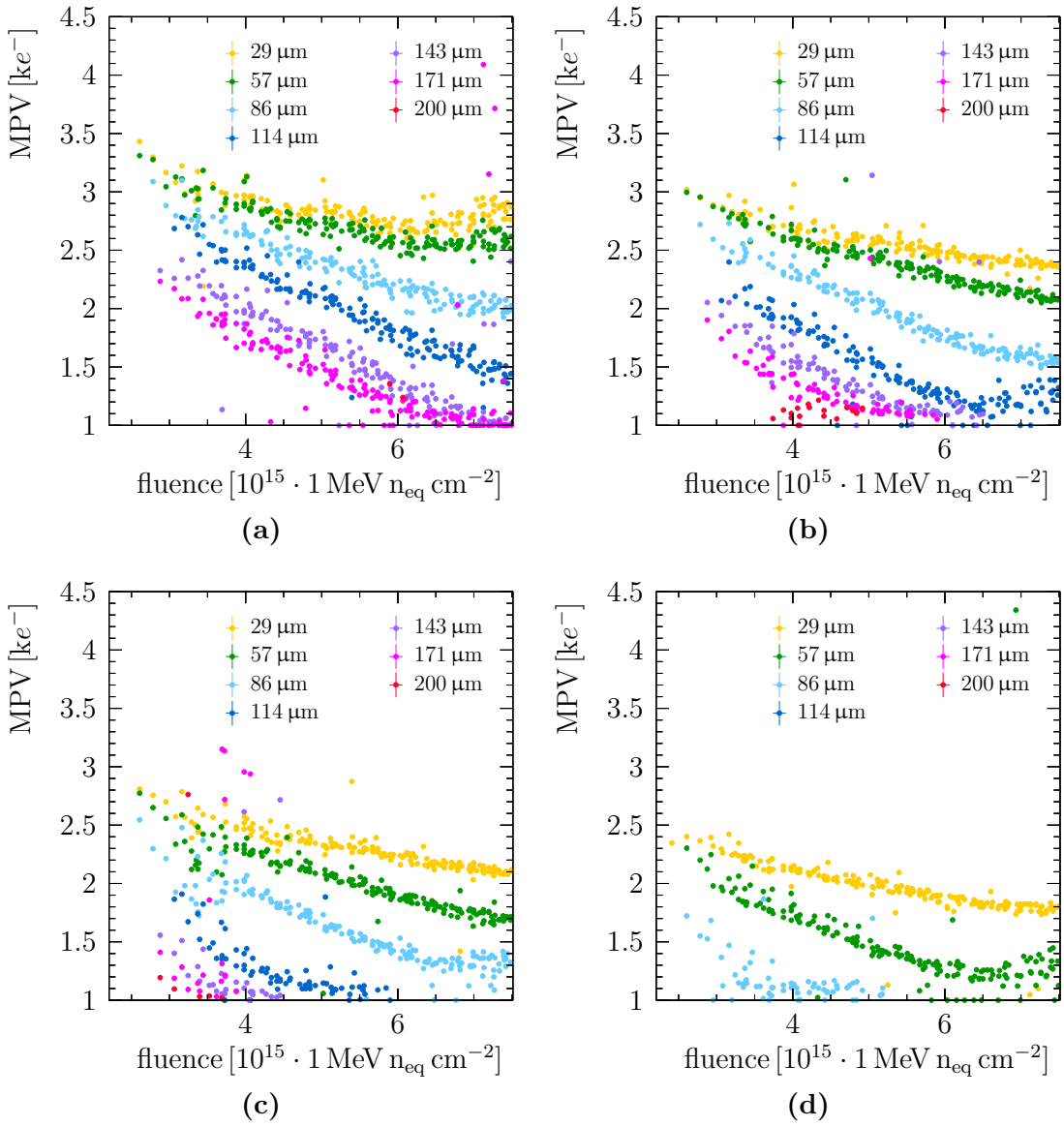



**Figure 8.18:** Charge collected as a function of fluence from different depths for a 200 µm HPK n-on-p sensor operated at (a) at 1000 V, (b) at 750 V, (c) at 500 V and (d) at 250 V.

function of depth once full depletion is reached and higher than 3500 $e^-$. The time-to-threshold is therefore barely affected by timewalk, but it is extended by the time needed for the charge to migrate from the nondepleted volume (up to about 20 µm from the border to the depleted volume). For sensors uniformly irradiated to the full fluence it is observed that the charge is collected mostly from depths close to the pixel electrode. Due to radiation damage there is in general a reduction of the charge collected caused by trapping leading to a contribution to the time-to-threshold. Being able to disentangle

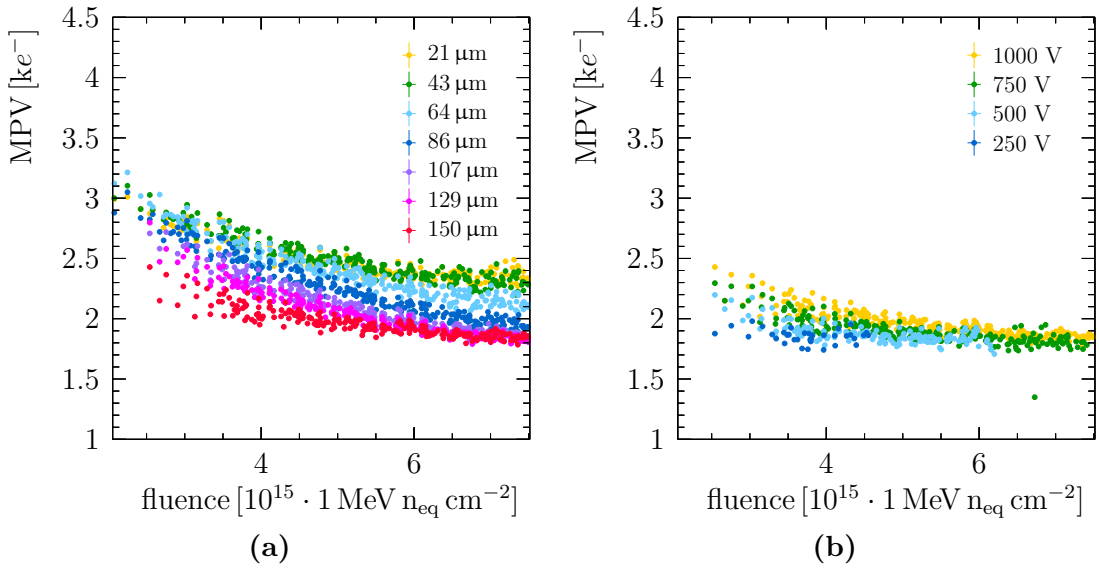



**Figure 8.19:** In (a) charge collected as a function of fluence from different depths for a sensor operated at 1000 V and in (b) charge collected as a function of fluence for charges liberated at 150 µm depths at different operation voltages. The sensor is a 150 µm Micron n-on-n sensor.

sensor and chip effects, the time-to-threshold has been corrected for timewalk, leading to a flatter profile which resembles the nonirradiated one in terms of shape. Moreover, it is possible to see that the time required for the charge to be collected is shorter due to the higher operating voltage. Nonuniformly irradiated sensors allowed the study of charge collection and time-to-threshold variations as a function of fluence. In particular there is a clear enhancement of charge collected from depths up to 80 µm at an operating voltage of 1000 V due to charge multiplication. The charge multiplication effect has been observed for proton but not neutron irradiated sensors.
Regarding the different families of prototypes tested, there is a clear difference between Micron n-on-p and HPK on one side and Micron n-on-n on the other. It is not clear yet if this difference is due to the different bulk type of the sensor or to the different thickness; more devices are needed to further investigate this.

## 8.8 Outlook

The grazing angle technique has been proven very useful for detailed timing and charge collection efficiency studies. Possible improvements to the grazing angle technique and additional studies are presented here. The role that the grazing angle technique could play in the R&D of future detectors is also outlined.

### 8.8.1 Possible extra studies

Further studies can be performed using the grazing angle technique:

- As mentioned in Section 8.1, spatial alignment of the sensor could not be performed due to the large angle at which the sensors are placed with respect to the beam. However, clusters could be associated to tracks using the timing information. Hence one can select a sample of tracks with definite cluster length, as described in Section 8.1, and use it to align the sensors exploiting the knowledge of entry and exit points of the track. Even though this alignment might not be as good as for devices placed perpendicularly to the beam, it would open up a series of possible studies, for example regarding broken clusters or the entry point of the track, which exploit the knowledge of the position of track intercept with the sensor.

- Broken clusters, as defined in Section 8.1, have been removed from the data sample for the grazing angle analysis. They could be studied in more detail and possibly help in understanding why the two peaks expected in the double junction model are not observed.

- The cluster length distribution could be investigated as a function of the entry point of the track, after alignment, for nonuniformly irradiated sensors.

- It would be interesting to compare the measured charge profile as a function of fluence for different depths to simulation, in order to better understand the charge multiplication effect.

### 8.8.2 Future sensor R&D

A second upgrade of the LHCb detector, known as Upgrade II [181], is foreseen during the Long Shutdown 4 of the LHC (2030). With a tenfold increase with respect to the Upgrade I detector in terms of pile-up, radiation exposure and occupancy, the timing information becomes essential to suppress vertex misassociation and reduce the probability of combining hits belonging to different tracks. For the VELO a time resolution of better than 50 ps per hit would be extremely helpful. As can be seen from Figure 8.12, the time to threshold for the sensor type chosen for Upgrade I is more than 2 ns for nonirradiated sensors and irradiated sensors after timewalk correction, far from the goal of Upgrade II. Several possibilities are being considered in terms of sensor technology to provide fast timing, from thin planar sensors to Low Gain Avalanche Detectors (LGAD) and 3D sensors. In order to characterise fast silicon sensors, a high time resolution telescope is needed. The successor of the Timepix3 chip, Timepix4, which has recently been submitted, will have 55 µm spatial and 195 ps time granularity and would be the

ideal candidate to build the next-generation beam telescope. Assuming eight planes as the current telescope, a 20-30 ps track time resolution can be achieved which, combined with the excellent pointing resolution, will allow the evaluation the most suitable sensor technology: in this context, the grazing angle technique is an ideal technique to study the time capabilities of such devices.

# Appendix A

# Neutrino Mass Matrix Diagonalisation

The block diagonalisation of the mass matrix Eq. 1.10 is presented here for the one generation case. The mass matrix $M^{D+M}$ for one generation in the basis $(\nu_L, \nu_R^c)$ results in the following $2 \times 2$ symmetric matrix:

$$M^{D+M} = \begin{pmatrix} 0 & M^D \\ M^D & M_R^M \end{pmatrix}. \tag{A.1}$$

The eigenvalues of the matrix, obtained by solving the characteristic polynomial and assuming $\theta \equiv M^D/M_R^M << 1$, are:

$$\begin{aligned} m_\nu &= -M_R^M \frac{(M_R^M)^2}{(M^D)^2} = -M_R^M \theta^2, \\ m_N &= M_R^M + M_R^M \frac{(M_R^M)^2}{(M^D)^2} = M_R^M + M_R^M \theta^2, \end{aligned} \tag{A.2}$$

with eigenstates, respectively:

$$\begin{aligned} n &= \left(1 - \frac{1}{2}\theta^2\right) \begin{pmatrix} 1 \\ -\theta \end{pmatrix} (\nu_L, \nu_R^c) = \left(1 - \frac{1}{2}\theta^2\right)\nu_L - \theta\nu_R^c, \\ N &= \left(1 - \frac{1}{2}\theta^2\right) \begin{pmatrix} \theta \\ 1 \end{pmatrix} (\nu_L, \nu_R^c) = \left(1 - \frac{1}{2}\theta^2\right)\nu_R^c + \theta\nu_L. \end{aligned} \tag{A.3}$$

# Appendix B

# Background composition per region

The data and expected backgrounds distributions for different observables are compared in the $W$ and $b\bar{b}$ regions are shown for both same-sign and opposite-sign muons.

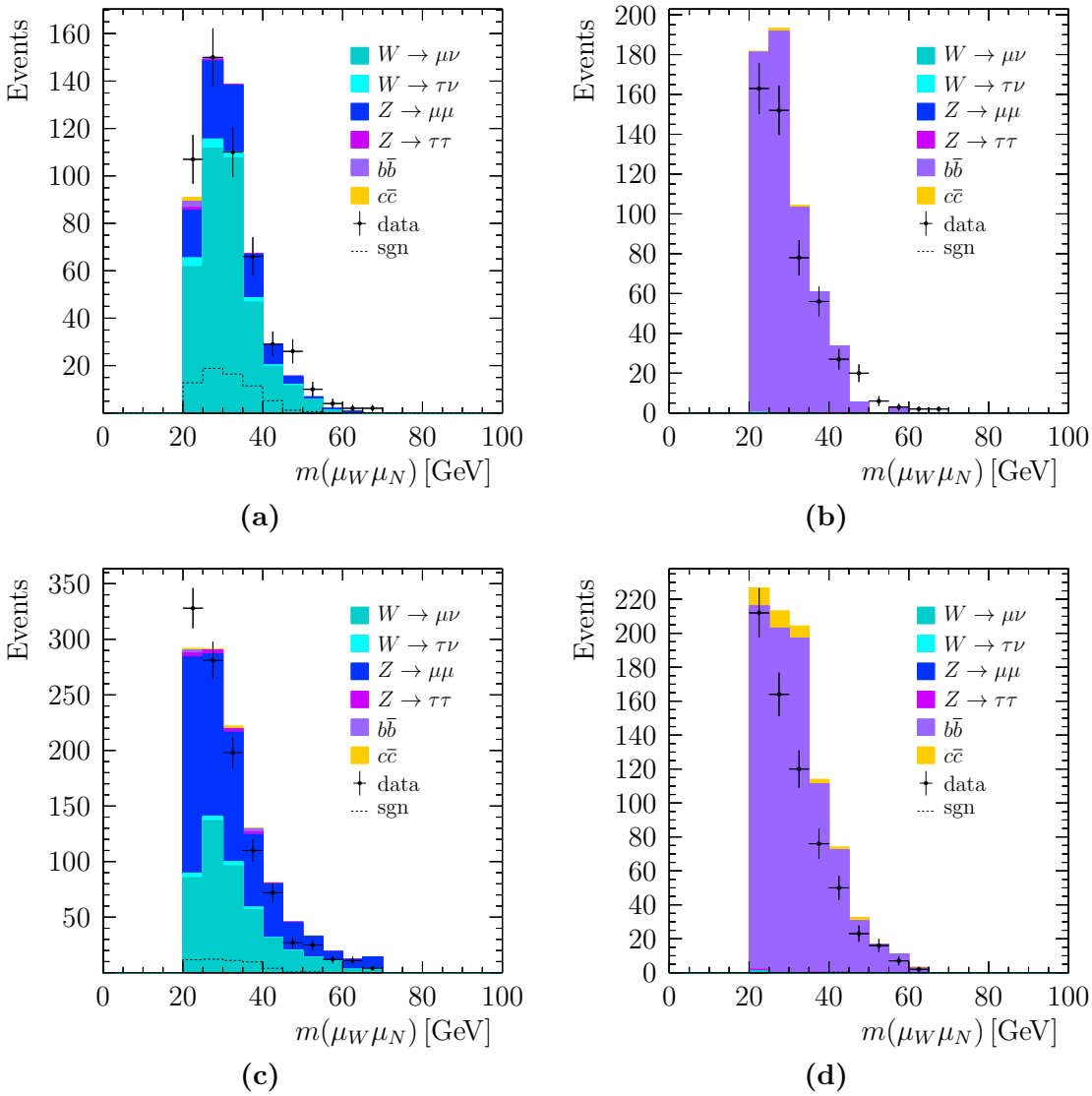



**Figure B.1:** Background composition predicted by the simulation (stack), signal (dashed line) and data (filled circles) in (a) for same-sign muons in the $W$ region, (c) opposite-sign muons in the $W$ region, (b) same-sign muons in the $b\bar{b}$ region and (d) opposite-sign muons in the $b\bar{b}$ region. The signal is normalised to the expected yield, assuming a coupling $10^3$ times higher than the current limit.

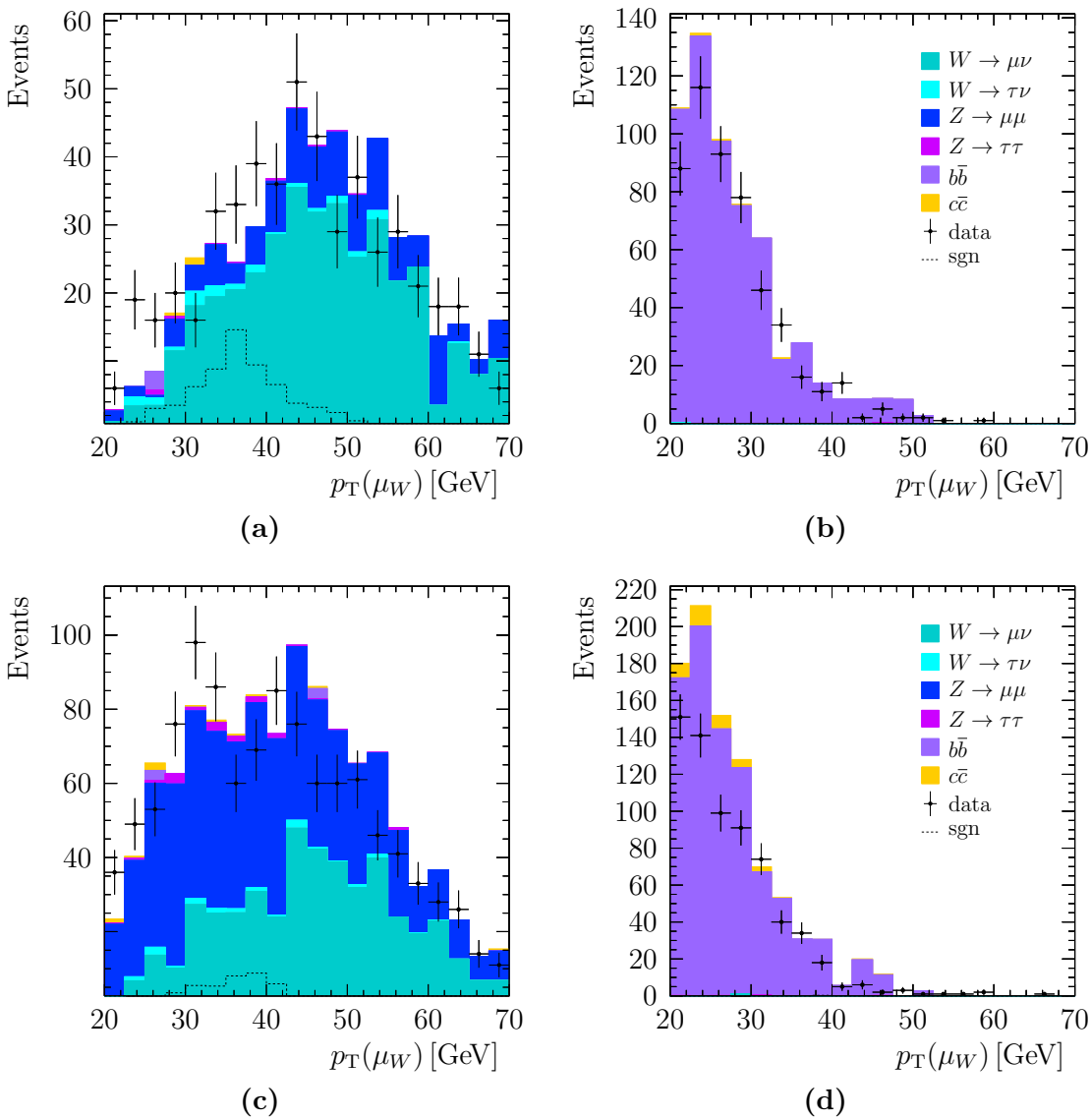



**Figure B.2:** Background composition predicted by the simulation (stack), signal (dashed line) and data (filled circles) in (a) for same-sign muons in the $W$ region, (c) opposite-sign muons in the $W$ region, (b) same-sign muons in the $b\bar{b}$ region and (d) opposite-sign muons in the $b\bar{b}$ region. The signal is normalised to the expected yield, assuming a coupling $10^3$ times higher than the current limit.

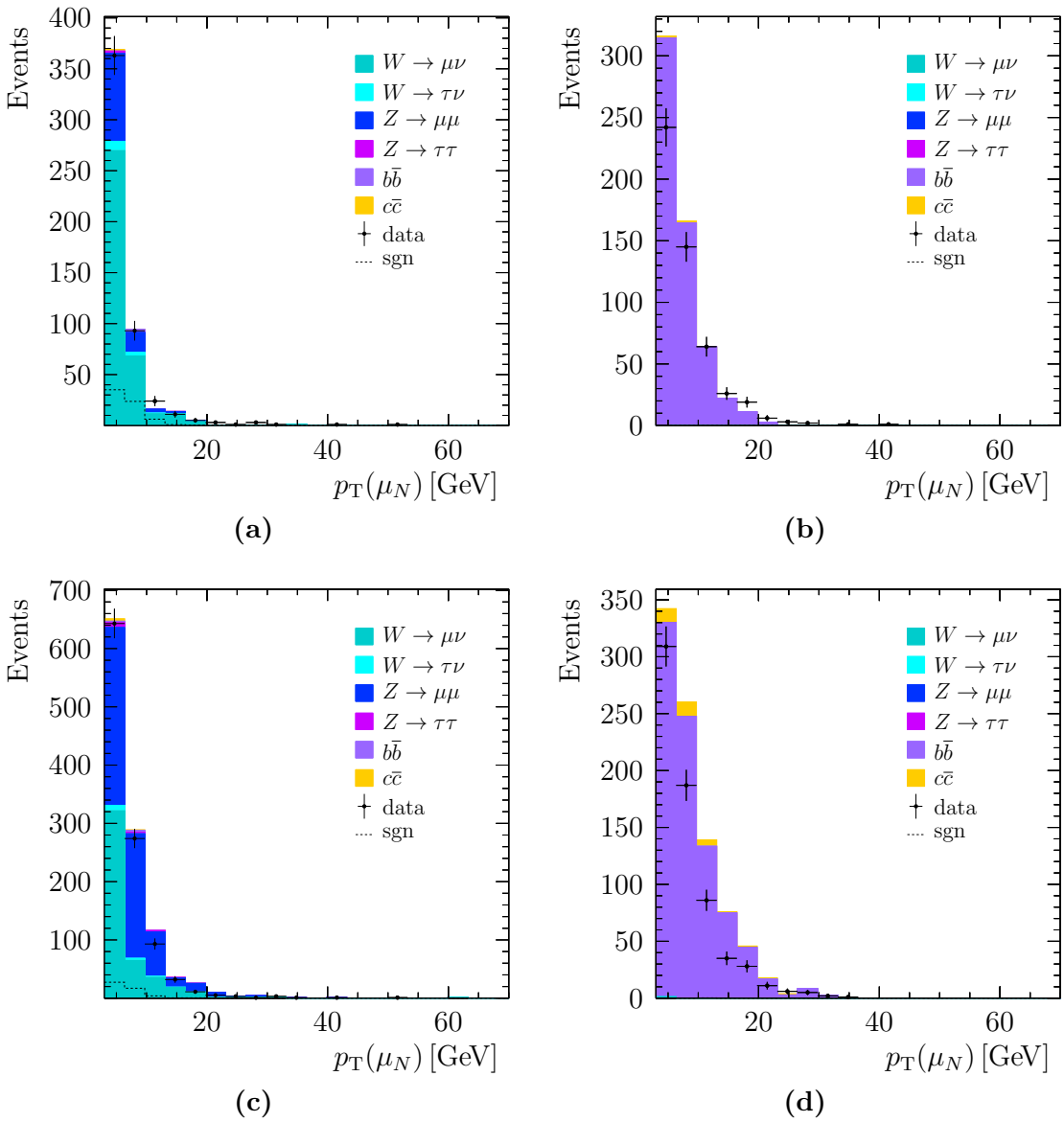



**Figure B.3:** Background composition predicted by the simulation (stack), signal (dashed line) and data (filled circles) in (a) for same-sign muons in the $W$ region, (c) opposite-sign muons in the $W$ region, (b) same-sign muons in the $b\bar{b}$ region and (d) opposite-sign muons in the $b\bar{b}$ region. The signal is normalised to the expected yield, assuming a coupling $10^3$ times higher than the current limit.

# Appendix C

# Normalisation channel yields

The $W \to \mu\nu$ yields per pseudorapidity bin, obtained from the fit to the transverse momentum spectrum of the muon are collected in Tables C.1 and C.2 for positively and negatively charged muons, respectively.

**Table C.1:** $W^+ \to \mu^+\nu$ channel fit result in bins of muon pseudorapidity.

| Eta | $W \to \mu\nu$ | $W \to \tau\nu$ | $Z \to \mu\mu$ | $Z \to \tau\tau$ | $b\bar{b}$ | qcd |
|---|---|---|---|---|---|---|
| 2.00 - 2.25 | $152\,236 \pm 523$ | $4492 \pm 67$ | $8239 \pm 26$ | $585 \pm 24$ | $3885 \pm 62$ | $8750 \pm 355$ |
| 2.25 - 2.50 | $191\,477 \pm 572$ | $6508 \pm 81$ | $11\,807 \pm 37$ | $810 \pm 28$ | $4009 \pm 63$ | $9528 \pm 374$ |
| 2.50 - 2.75 | $177\,692 \pm 536$ | $6040 \pm 78$ | $10\,806 \pm 34$ | $744 \pm 27$ | $2839 \pm 53$ | $12\,984 \pm 345$ |
| 2.75 - 3.00 | $151\,897 \pm 485$ | $5108 \pm 71$ | $9258 \pm 29$ | $629 \pm 25$ | $2081 \pm 46$ | $15\,309 \pm 312$ |
| 3.00 - 3.25 | $121\,568 \pm 427$ | $3894 \pm 62$ | $7366 \pm 23$ | $498 \pm 22$ | $1468 \pm 38$ | $15\,979 \pm 276$ |
| 3.25 - 3.50 | $93\,235 \pm 376$ | $2778 \pm 53$ | $5987 \pm 19$ | $391 \pm 20$ | $1117 \pm 33$ | $15\,397 \pm 253$ |
| 3.50 - 4.00 | $91\,894 \pm 387$ | $2600 \pm 51$ | $7600 \pm 24$ | $428 \pm 21$ | $1126 \pm 34$ | $20\,460 \pm 285$ |
| 4.00 - 4.50 | $10\,475 \pm 202$ | $263 \pm 16$ | $1257 \pm 4$ | $52 \pm 7$ | $160 \pm 13$ | $4092 \pm 156$ |

**Table C.2:** $W^- \to \mu^-\nu$ channel fit result in bins of muon pseudorapidity.

| Eta | $W \to \mu\nu$ | $W \to \tau\nu$ | $Z \to \mu\mu$ | $Z \to \tau\tau$ | $b\bar{b}$ | qcd |
|---|---|---|---|---|---|---|
| 2.00 - 2.25 | $86\,881 \pm 398$ | $2862 \pm 53$ | $8607 \pm 27$ | $578 \pm 24$ | $4085 \pm 64$ | $11\,981 \pm 276$ |
| 2.25 - 2.50 | $109\,605 \pm 443$ | $4045 \pm 64$ | $12\,458 \pm 39$ | $814 \pm 29$ | $4002 \pm 63$ | $12\,759 \pm 302$ |
| 2.50 - 2.75 | $105\,733 \pm 437$ | $3832 \pm 62$ | $11\,553 \pm 36$ | $759 \pm 28$ | $2982 \pm 55$ | $12\,680 \pm 302$ |
| 2.75 - 3.00 | $99\,067 \pm 430$ | $3401 \pm 58$ | $10\,013 \pm 31$ | $656 \pm 26$ | $2274 \pm 48$ | $14\,038 \pm 306$ |
| 3.00 - 3.25 | $91\,547 \pm 417$ | $2910 \pm 54$ | $8094 \pm 25$ | $537 \pm 23$ | $1504 \pm 39$ | $13\,390 \pm 300$ |
| 3.25 - 3.50 | $85\,310 \pm 409$ | $2421 \pm 49$ | $6473 \pm 20$ | $409 \pm 20$ | $1119 \pm 33$ | $12\,269 \pm 299$ |
| 3.50 - 4.00 | $121\,883 \pm 503$ | $2972 \pm 55$ | $7931 \pm 25$ | $477 \pm 22$ | $1150 \pm 34$ | $15\,087 \pm 377$ |
| 4.00 - 4.50 | $22\,498 \pm 239$ | $456 \pm 21$ | $1261 \pm 4$ | $64 \pm 8$ | $153 \pm 12$ | $2299 \pm 180$ |

# Appendix D

# Validation regions

The comparisons between data and expected backgrounds in each of the validation regions are shown in Figure D.1 and Figure D.2 for same-sign and opposite-sign muons, respectively.

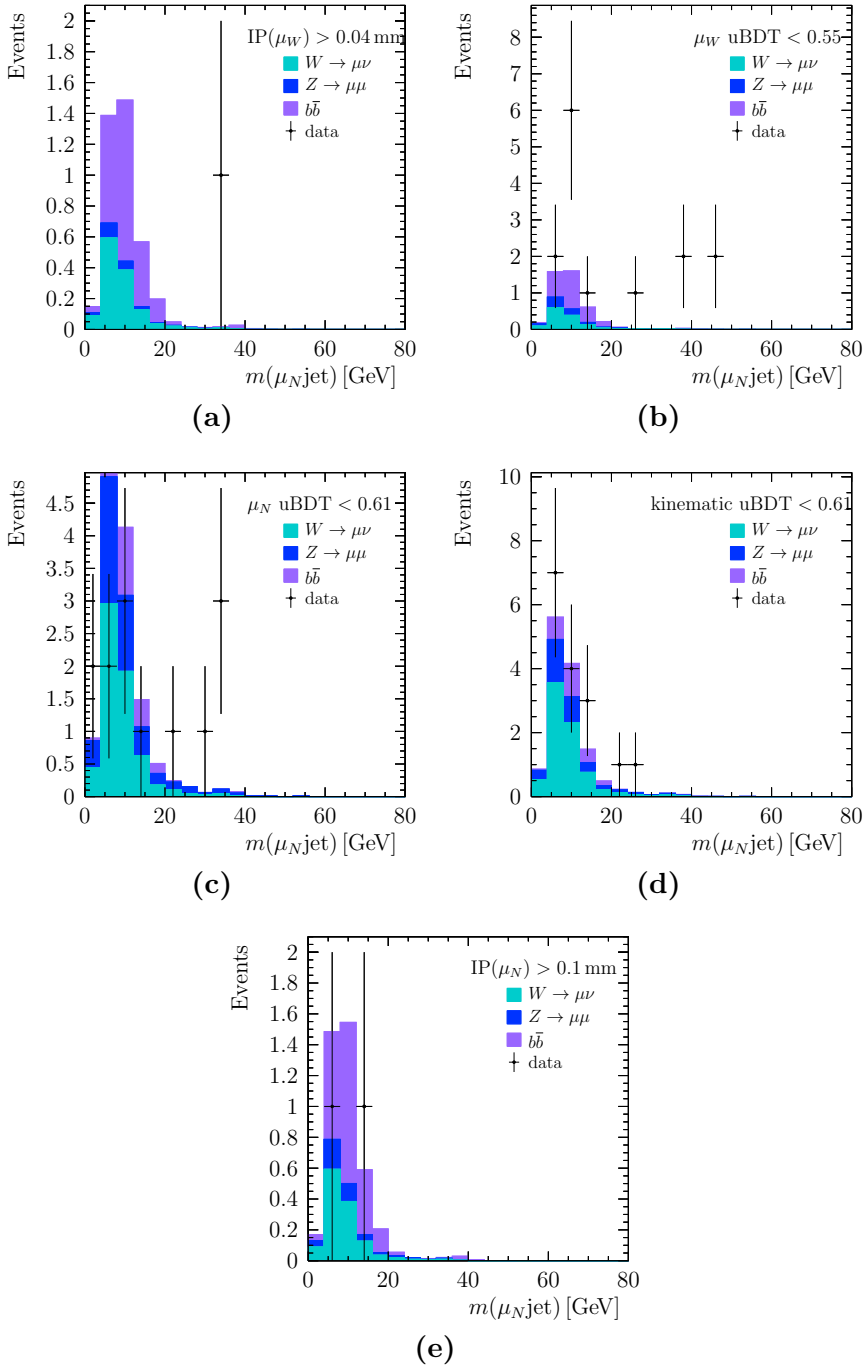



**Figure D.1:** Comparison between data (points) and expected backgrounds (stack) in the validation regions obtained by reversing one cut in the signal selections for same-sign muons. The reversed cut is indicated in each figure.

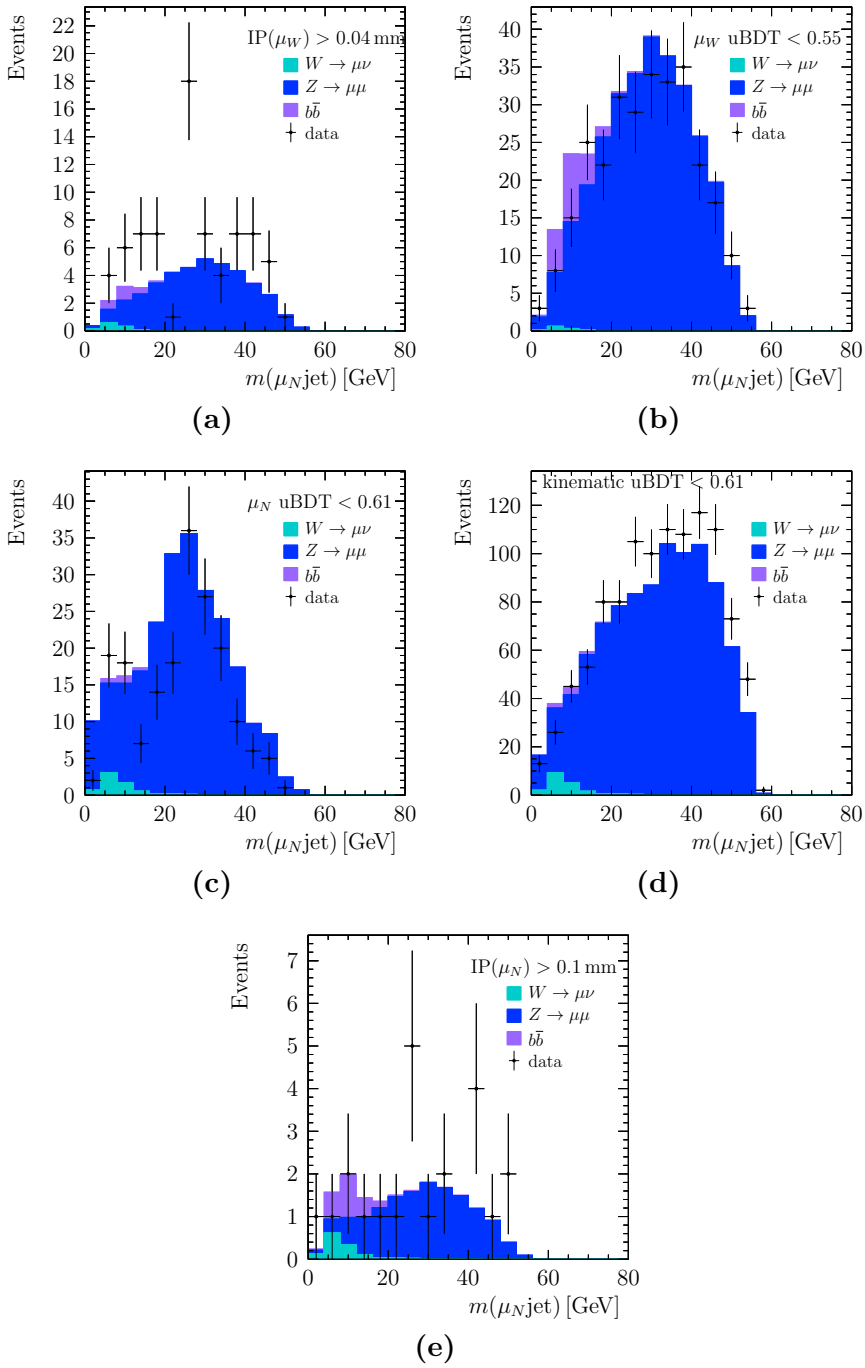



**Figure D.2:** Comparison between data (points) and expected backgrounds (stack) in the validation regions obtained by reversing one cut in the signal selections for opposite-sign muons. The reversed cut is indicated in each figure.

# References


[1] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, [160(1964)].

[2] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, [157(1964)].

[3] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[4] J. Schwinger, *The theory of quantized fields*, Phys. Rev. **82** (1951) 914.

[5] W. Bonivento *et al.*, *Proposal to Search for Heavy Neutral Leptons at the SPS*, `arXiv:1310.1762`.

[6] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [Usp. Fiz. Nauk161,no.5,61(1991)].

[7] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. **D51** (1995) 379, `arXiv:hep-ph/9404302`.

[8] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, Helv. Phys. Acta **6** (1933) 110, [Gen. Rel. Grav.41,207(2009)].

[9] A. Bosma, *21-cm line studies of spiral galaxies. II. The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types*, Astron. J. **86** (1981) 1825.

[10] V. C. Rubin, D. Burstein, W. K. Ford, Jr. and N. Thonnard, *Rotation velocities of 16 SA galaxies and a comparison of Sa, Sb, and SC rotation properties*, Astrophysical Journal **289** (1985) 81.

[11] Planck, P. A. R. Ade *et al.*, *Planck 2015 results. XIII. Cosmological parameters*, Astron. Astrophys. **594** (2016) A13, `arXiv:1502.01589`.

[12] J. Frieman, M. Turner, and D. Huterer, *Dark Energy and the Accelerating Universe*, Ann. Rev. Astron. Astrophys. **46** (2008) 385, `arXiv:0803.0982`.

[13] B. T. Cleveland *et al.*, *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, Astrophys. J. **496** (1998) 505.

[14] SAGE, J. N. Abdurashitov *et al.*, *Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period*, Phys. Rev. **C80** (2009) 015807, `arXiv:0901.2200`.

[15] GALLEX, W. Hampel *et al.*, *GALLEX solar neutrino observations: Results for GALLEX IV*, Phys. Lett. **B447** (1999) 127.

[16] M. Altmann *et al.*, *Complete results for five years of gno solar neutrino observations*, Physics Letters B **616** (2005), no. 3 174 .

[17] Kamiokande, Y. Fukuda *et al.*, *Solar neutrino data covering solar cycle 22*, Phys. Rev. Lett. **77** (1996) 1683.

[18] Super-Kamiokande Collaboration, K. Abe *et al.*, *Solar neutrino measurements in super-kamiokande-iv*, Phys. Rev. D **94** (2016) 052010.

[19] SNO Collaboration, B. Aharmim *et al.*, *Combined analysis of all three phases of solar neutrino data from the sudbury neutrino observatory*, Phys. Rev. C **88** (2013) 025501.

[20] KamLAND, A. Gando *et al.*, *Reactor On-Off Antineutrino Measurement with KamLAND*, Phys. Rev. **D88** (2013), no. 3 033001, `arXiv:1303.4667`.

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

[22] M. G. Aartsen *et al.*, *Neutrino oscillation studies with icecube-deepcore*, Nuclear Physics B **908** (2016) .

[23] K2K Collaboration, M. H. Ahn *et al.*, *Measurement of neutrino oscillation by the K2K experiment*, Phys. Rev. D **74** (2006) 072003.

[24] P. Adamson *et al.*, *Combined analysis of $\nu_\mu$ disappearance and $\nu_\mu \to \nu_e$ appearance in minos using accelerator and atmospheric neutrinos*, Physical Review Letters **112** (2014) .

[25] NOvA Collaboration, P. Adamson *et al.*, *Measurement of the neutrino mixing angle $\theta_{23}$ in nova*, Phys. Rev. Lett. **118** (2017) 151802.

[26] OPERA Collaboration, N. Agafonova *et al.* , *Discovery of $\tau$ neutrino appearance in the cngs neutrino beam with the opera experiment*, Phys. Rev. Lett. **115** (2015) 121802.

[27] Daya Bay Collaboration, F. P. An *et al.*, *New Measurement of Antineutrino Oscillation with the Full Detector Configuration at Daya Bay*, Phys. Rev. Lett. **115** (2015) 111802.

[28] RENO Collaboration, J. H. Choi *et al.*, *Observation of Energy and Baseline Dependent Reactor Antineutrino Disappearance in the RENO Experiment*, Phys. Rev. Lett. **116** (2016) 211801.

[29] Double Chooz collaboration, *Measurement of $\theta_{13}$ in double chooz using neutron captures on hydrogen with novel background rejection techniques*, Journal of High Energy Physics **2016** (2016) 163.

[30] A. Strumia and F. Vissani, *Neutrino masses and mixings and...*, `arXiv:hep-ph/0606054`.

[31] C. Giganti, S. Lavignac, and M. Zito, *Neutrino oscillations: the rise of the PMNS paradigm*, Prog. Part. Nucl. Phys. **98** (2018) 1, `arXiv:1710.00715`.

[32] Particle Data Group, M. Tanabashi *et al.* , *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001.

[33] S. Bilenky, *Introduction to the physics of massive and mixed neutrinos*, Lect. Notes Phys. **817** (2010) 1.

[34] A. Boyarsky, O. Ruchayskiy, and M. Shaposhnikov, *The Role of sterile neutrinos in cosmology and astrophysics*, Ann. Rev. Nucl. Part. Sci. **59** (2009) 191, `arXiv:0901.0011`.

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

[36] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*, Phys. Lett. **B174** (1986) 45.

[37] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, *On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe*, Phys. Lett. **155B** (1985) 36.

[38] F. R. Klinkhamer and N. S. Manton, *A saddle-point solution in the weinberg-salam theory*, Phys. Rev. D **30** (1984) 2212.

[39] A. Pilaftsis and T. E. J. Underwood, *Resonant leptogenesis*, Nucl. Phys. **B692** (2004) 303, `arXiv:hep-ph/0309342`.

[40] A. Pilaftsis and T. E. J. Underwood, *Electroweak-scale resonant leptogenesis*, Phys. Rev. **D72** (2005) 113001, `arXiv:hep-ph/0506107`.

[41] V. V. Khoze and G. Ro, *Leptogenesis and Neutrino Oscillations in the Classically Conformal Standard Model with the Higgs Portal*, JHEP **10** (2013) 075, `arXiv:1307.3764`.

[42] M. Shaposhnikov, *A Possible symmetry of the nuMSM*, Nucl. Phys. **B763** (2007) 49, `arXiv:hep-ph/0605047`.

[43] LSND, A. Aguilar-Arevalo *et al.*, *Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam*, Phys. Rev. **D64** (2001) 112007, `arXiv:hep-ex/0104049`.

[44] MiniBooNE, A. A. Aguilar-Arevalo *et al.*, *A Combined $\nu_\mu \to \nu_e$ and $\bar{\nu}_\mu \to \bar{\nu}_e$ Oscillation Analysis of the MiniBooNE Excesses*, 2012. `arXiv:1207.4809`.

[45] C. Giunti and M. Laveder, *Statistical Significance of the Gallium Anomaly*, Phys. Rev. **C83** (2011) 065504, `arXiv:1006.3244`.

[46] G. Mention *et al.*, *The Reactor Antineutrino Anomaly*, Phys. Rev. **D83** (2011) 073006, `arXiv:1101.2755`.

[47] Intensity Frontier Neutrino Working Group, A. de Gouvea *et al.*, *Working Group Report: Neutrinos*, in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013. `arXiv:1310.4340`.

[48] T. Asaka, S. Blanchet, and M. Shaposhnikov, *The nuMSM, dark matter and neutrino masses*, Phys. Lett. **B631** (2005) 151, `arXiv:hep-ph/0503065`.

[49] M. Shaposhnikov, *Is there a new physics between electroweak and Planck scales?*, in *Astroparticle Physics: Current Issues, 2007 (APCI07) Budapest, Hungary, June 21-23, 2007*, 2007. `arXiv:0708.3550`.

[50] D. Gorbunov and M. Shaposhnikov, *How to find neutral leptons of the $\nu$MSM?*, JHEP **10** (2007) 015, `arXiv:0705.1729`, [Erratum: JHEP11,101(2013)].

[51] L. Canetti, M. Drewes, T. Frossard, and M. Shaposhnikov, *Dark Matter, Baryogenesis and Neutrino Oscillations from Right Handed Neutrinos*, Phys. Rev. **D87** (2013) 093006, `arXiv:1208.4607`.

[52] M. Drewes and B. Garbrecht, *Combining experimental and cosmological constraints on heavy neutrinos*, Nucl. Phys. **B921** (2017) 250, `arXiv:1502.00477`.

[53] E. Nardi, E. Roulet, and D. Tommasini, *Limits on neutrino mixing with new heavy particles*, Phys. Lett. **B327** (1994) 319, `arXiv:hep-ph/9402224`.

[54] L. Lello and D. Boyanovsky, *Searching for sterile neutrinos from $\pi$ and $K$ decays*, Phys. Rev. **D87** (2013) 073017, `arXiv:1208.5559`.

[55] G. Azuelos *et al.*, *Constraints on massive neutrinos in $\pi \to e\nu$ decay*, Phys. Rev. Lett. **56** (1986) 2241.

[56] PIENU, M. Aoki *et al.*, *Search for Massive Neutrinos in the Decay $\pi \to e\nu$*, Phys. Rev. **D84** (2011) 052002, `arXiv:1106.4055`.

[57] K. Assamagan *et al.*, *Search for a heavy neutrino state in the decay $\pi^+ \to \mu^+ + \nu_\mu$*, Phys. Lett. **B434** (1998) 158.

[58] T. Yamazaki *et al.*, *Search for Heavy Neutrinos in Kaon Decay*, , [Conf. Proc.C840719,262(1984)].

[59] E949, A. V. Artamonov *et al.*, *Search for heavy neutrinos in $K^+ \to \mu^+\nu_H$ decays*, Phys. Rev. **D91** (2015), no. 5 052001, `arXiv:1411.3963`, [Erratum: Phys. Rev.D91,no.5,059903(2015)].

[60] G. Bernardi *et al.*, *Further limits on heavy neutrino couplings*, Phys. Lett. B **203** (1987) 332.

[61] NA3 Collaboration *et al.*, *Direct photon production from pions and protons at* 200 GeV/$c$, Zeitschrift für Physik C Particles and Fields **31** (1986) 341.

[62] CHARM, F. Bergsma *et al.*, *A Search for Decays of Heavy Neutrinos in the Mass Range 0.5-GeV to 2.8-GeV*, Phys. Lett. **166B** (1986) 473.

[63] S. A. Baranov *et al.*, *Search for heavy neutrinos at the IHEP-JINR neutrino detector*, Phys. Lett. **B302** (1993) 336.

[64] WA66, A. M. Cooper-Sarkar *et al.*, *Search for Heavy Neutrino Decays in the BEBC Beam Dump Experiment*, Phys. Lett. **160B** (1985) 207.

[65] FMMF, E. Gallas *et al.*, *Search for neutral weakly interacting massive particles in the Fermilab Tevatron wide band neutrino beam*, Phys. Rev. **D52** (1995) 6.

[66] NuTeV, E815, A. Vaitaitis *et al.*, *Search for neutral heavy leptons in a high-energy neutrino beam*, Phys. Rev. Lett. **83** (1999) 4943, `arXiv:hep-ex/9908011`.

[67] NOMAD, P. Astier *et al.*, *Search for heavy neutrinos mixing with tau neutrinos*, Phys. Lett. **B506** (2001) 27, `arXiv:hep-ex/0101041`.

[68] Belle, D. Liventsev *et al.*, *Search for heavy neutrinos at Belle*, Phys. Rev. **D87** (2013), no. 7 071102, `arXiv:1301.1105`, [Erratum: Phys. Rev.D95,no.9,099903(2017)].

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

[70] B. Shuve and M. E. Peskin, *Revision of the LHCb Limit on Majorana Neutrinos*, Phys. Rev. **D94** (2016), no. 11 113007, `arXiv:1607.04258`.

[71] A. Atre, T. Han, S. Pascoli, and B. Zhang, *The Search for Heavy Majorana Neutrinos*, JHEP **05** (2009) 030, `arXiv:0901.3589`.

[72] L3, O. Adriani *et al.*, *Search for isosinglet neutral heavy leptons in Z decays*, Phys. Lett. **B295** (1992) 371.

[73] DELPHI, P. Abreu *et al.*, *Search for neutral heavy leptons produced in Z decays*, Z. Phys. **C74** (1997) 57, [Erratum: Z. Phys.C75,580(1997)].

[74] CMS, A. M. Sirunyan *et al.*, *Search for heavy neutral leptons in events with three charged leptons in proton-proton collisions at $\sqrt{s} = 13\ TeV$*, Phys. Rev. Lett. **120** (2018), no. 22 221801, `arXiv:1802.02965`.

[75] G. Cvetic, A. Das, and J. Zamora-Saa, *Probing heavy neutrino oscillations in rare W boson decays*, `arXiv:1805.00070`.

[76] J. C. Helo, S. Kovalenko, and I. Schmidt, *On sterile neutrino mixing with $\nu_\tau$*, Phys. Rev. **D84** (2011) 053008, `arXiv:1105.3019`.

[77] J. C. Helo, S. Kovalenko, and I. Schmidt, *Sterile neutrinos in lepton number and lepton flavor violating decays*, Nucl. Phys. **B853** (2011) 80, `arXiv:1005.1607`.

[78] LHCb, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[79] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[80] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[81] LHCb VELO, E. Jans, *Operational aspects of the VELO cooling system of LHCb*, PoS **Vertex2013** (2013) 038.

[82] P. David, *Search for exotic long-lived particles with the LHCb detector*, PhD thesis, Feb, 2016, Presented 07 Jul 2016.

[83] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[84] S. Tolk, *Discovery Of Rare B decays. First observation of the $B_s \to \mu^+\mu^-$ decays, first evidence of the $B^0 \to \mu^+\mu^-$ decays.*, PhD thesis, Dec, 2015, Presented 08 Apr 2016.

[85] LHCb, P. Perret and X. Vilasís-Cardona, *Performance of the LHCb calorimeters during the period 2010-2012*, J. Phys. Conf. Ser. **587** (2015), no. 1 012012.

[86] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[87] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015), no. 07 1530022, `arXiv:1412.6352`.

[88] W. Barter *et al.*, *Jets reconstruction and performances at LHCb*, Tech. Rep. LHCb-INT-2012-015. CERN-LHCb-INT-2012-015, CERN, Geneva, May, 2012.

[89] LHCb, R. Aaij *et al.*, *Study of forward $Z$ + jet production in pp collisions at $\sqrt{s} = 7\ TeV$*, JHEP **01** (2014) 033, `arXiv:1310.8197`.

[90] W. Barter, *Z boson and associated jet production at the LHCb experiment*, Mar, 2014. Presented 27 May 2014.

[91] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-$k_t$ jet clustering algorithm*, JHEP **04** (2008) 063, `arXiv:0802.1189`.

[92] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally invariant $K_t$ clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406** (1993) 187.

[93] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, JHEP **08** (1997) 001, `arXiv:hep-ph/9707323`.

[94] M. Cacciari and G. P. Salam, *Dispelling the n3 myth for the kt jet-finder*, Physics Letters B **641** (2006), no. 1 57 .

[95] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[96] J. Alwall *et al.*, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079. 158 p.

[97] T. Sjöstrand, S. Mrenna, and P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2007) 852.

[98] E. Izaguirre and B. Shuve, *Multilepton and Lepton Jet Probes of Sub-Weak-Scale Right-Handed Neutrinos*, Phys. Rev. **D91** (2015), no. 9 093010, `arXiv:1504.02470`.

[99] LHCb, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[100] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, Eur. Phys. J. **C72** (2012) 1896, `arXiv:1111.6097`.

[101] O. A. De Aguiar Francisco, M. Potterat, and M. Rangel, *LHCb jet identification*, Tech. Rep. LHCb-INT-2014-021. CERN-LHCb-INT-2014-021, CERN, Geneva, Mar, 2014.

[102] A. Hocker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, `arXiv:physics/0703039`.

[103] J. Stevens and M. Williams, *uBoost: A boosting method for producing uniform selection efficiencies from multivariate classifiers*, JINST **8** (2013) P12013, `arXiv:1305.7248`.

[104] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: Bounds for k-fold and progressive cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, (New York, NY, USA), pp. 203–208, ACM, 1999. doi: 10.1145/307400.307439.

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

[106] CMS, S. Chatrchyan *et al.*, *Measurement of inclusive W and Z boson production cross sections in pp collisions at* $\sqrt{s}$ *= 8 TeV*, Phys. Rev. Lett. **112** (2014) 191802, `arXiv:1402.0923`.

[107] ROOT Collaboration, K. Cranmer *et al.*, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016, New York U., New York, Jan, 2012.

[108] S. Catani, G. Ferrera, and M. Grazzini, *W Boson Production at Hadron Colliders: The Lepton Charge Asymmetry in NNLO QCD*, JHEP **05** (2010) 006, `arXiv:1002.3115`.

[109] E. Mirkes and J. Ohnemus, *W and Z polarization effects in hadronic collisions*, Phys. Rev. **D50** (1994) 5692, `arXiv:hep-ph/9406381`.

[110] E. Richter-Was and Z. Was, *W production at LHC: lepton angular distributions and reference frames for probing hard QCD*, Eur. Phys. J. **C77** (2017), no. 2 111, `arXiv:1609.02536`.

[111] S. Camarda *et al.*, *DYTurbo: Fast predictions for Drell-Yan processes*, `arXiv:1910.07049`.

[112] LHCb collaboration, R. Aaij *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** (2016) 131, `arXiv:1605.00951`.

[113] R. Brun and F. Rademakers, *Root - an object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **389** (1997), no. 1 81 , New Computing Techniques in Physics Research V.

[114] S. Farry and N. Chiapolini, *A measurement of high-pT muon reconstruction efficiencies in 2011 and 2012 data*, Tech. Rep. LHCb-INT-2014-030. CERN-LHCb-INT-2014-030, CERN, Geneva, Jul, 2014.

[115] W. J. Barter, *Muon reconstruction efficiencies in events containing jets*, Tech. Rep. LHCb-INT-2014-036. CERN-LHCb-INT-2014-036, CERN, Geneva, Sep, 2014.

[116] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, `arXiv:1408.1251`.

[117] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, `arXiv:1312.5000`.

[118] W. J. Barter, S. Bifani, and S. Farry, *Measurement of forward W and Z boson production in association with jets in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, Tech. Rep. LHCb-ANA-2015-067, CERN, Geneva, May, 2016.

[119] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007.

[120] LHCb collaboration, *Expression of interest for an LHCb upgrade*, CERN-LHCC-2008-007.

[121] LHCb collaboration, *Letter of Intent for the LHCb Upgrade*, CERN-LHCC-2011-001.

[122] P. Seyfert, *Tracking, Vertexing and data handling strategy for the LHCb upgrade*, p. 039, 07, 2018. doi: 10.22323/1.309.0039.

[123] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022.

[124] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001.

[125] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021.

[126] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[127] A. Papadelis, M. Merk, and E. Jans, *Characterisation and commissioning of the LHCb VELO detector*, PhD thesis, 2009, Presented on 17 Jun 2009.

[128] T. Poikela *et al.*, *Timepix3: A 65k channel hybrid pixel readout chip with simultaneous toa/tot and sparse readout*, Journal of Instrumentation **9** (2014) C05013.

[129] T. Poikela *et al.*, *The VeloPix ASIC*, JINST **12** (2017), no. 01 C01070.

[130] R. Plackett *et al.*, *Measurement of Radiation Damage to 130nm Hybrid Pixel Detector Readout Chips*, in *Proceedings, Topical Workshop on Electronics for Particle Physics (TWEPP09)*, CERN, CERN, 2009. doi: 10.5170/CERN-2009-006.157.

[131] R. E. Lyons and W. Vanderkulk, *The use of triple-modular redundancy to improve computer reliability*, IBM Journal of Research and Development **6** (1962) 200.

[132] VELO Group, CERN PH-DT, O. A. de Aguiar Francisco *et al.*, *Evaporative $CO_2$ microchannel cooling for the LHCb VELO pixel upgrade*, JINST **10** (2015), no. 05 C05014.

[133] T. Bird *et al.*, *VP Simulation and Track Reconstruction*, Tech. Rep. LHCb-PUB-2013-018. CERN-LHCb-PUB-2013-018, CERN, Geneva, Oct, 2013.

[134] S. Meroli, D. Passeri, and L. Servoli, *Energy loss measurement for charged particles in very thin silicon layers*, JINST **6** (2011) P06013.

[135] G. Hall, *Ionisation energy losses of highly relativistic charged particles in thin silicon layers*, Nuclear Instruments and Methods in Physics Research **220** (1984), no. 2 356 .

[136] H. Bichsel, *Straggling in thin silicon detectors*, Rev. Mod. Phys. **60** (1988) 663.

[137] N. Wermes, L. Rossi, P. Fischer, and T. Rohe, *Pixel Detectors, From Fundamentals to Applications*, Springer-Verlag, 2006. doi: 10.1007/3-540-28333-1.

[138] E. Buchanan, *Spatial Resolution Studies for the LHCb VELO Upgrade*, PhD thesis, Oct, 2018, Presented 31 Oct 2018.

[139] H. Spieler, *Semiconductor Detector Systems*, Series on Semiconductor Science and Technology, OUP Oxford, 2005.

[140] S. Ramo, *Currents induced by electron motion*, Proceedings of the IRE **27** (1939) 584.

[141] *Sentaurus Device User Guide.*, Synopsys **j-2014.09** (2014).

[142] G. Lindstrom, *Radiation damage in silicon detectors*, Nucl. Instrum. Meth. **A512** (2003) 30.

[143] M. Moll, *Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties*, PhD thesis, Hamburg U., 1999.

[144] M. Moll, *Displacement damage in silicon detectors for high energy physics. Displacement Damage in Silicon Detectors for High Energy Physics*, IEEE Trans. Nucl. Sci. **65** (2018), no. 8 1561.

[145] G. Kramberger *et al.*, *Effective trapping time of electrons and holes in different silicon materials irradiated with neutrons, protons and pions*, Nucl. Instrum. Meth. **A481** (2002) 297.

[146] V. Eremin, E. Verbitskaya, and Z. Li, *The origin of double peak electric field distribution in heavily irradiated silicon detectors*, Nucl. Instrum. Meth. **A476** (2002) 556.

[147] G. Kramberger *et al.*, *Modeling of electric field in silicon micro-strip detectors irradiated with neutrons and pions*, JINST **9** (2014), no. 10 P10016.

[148] M. Moll, E. Fretwurst, M. Kuhnke, and G. Lindstrom, *Relation between microscopic defects and macroscopic changes in silicon detector properties after hadron irradiation*, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms **186** (2002), no. 1 100 .

[149] O. Krasel *et al.*, *Measurement of trapping time constants in proton-irradiated silicon pad detectors*, IEEE Trans. Nucl. Sci. **51** (2004), no. 6 3055.

[150] S. Marti i Garcia, P. P. Allport, G. Casse, and A. Greenall, *A model of charge collection for irradiated p+n detectors*, Nucl. Instrum. Meth. **A473** (2001) 128.

[151] G. Casse *et al.*, *First results on the charge collection properties of segmented detectors made with p-type bulk silicon*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **487** (2002), no. 3 465 .

[152] G. Kramberger *et al.*, *Determination of effective trapping times for electrons and holes in irradiated silicon*, Nucl. Instrum. Meth. **A476** (2002) 645.

[153] F. Krummenacher, *Pixel detectors with local intelligence: an ic designer point of view*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **305** (1991), no. 3 527 .

[154] M. De Gaspari *et al.*, *Design of the analog front-end for the Timepix3 and Smallpix hybrid pixel detectors in 130 nm CMOS technology*, JINST **9** (2014), no. 01 C01037.

[155] M. Vicente Barreto Pinto, K. Carvalho Akiba, and E. Ribeiro Polycarpo Macedo, *Caracterização do TimePix3 e de sensores resistentes à radiação para upgrade do VELO*, Feb, 2016. Presented 04 Sep 2015.

[156] K. Akiba *et al.*, *LHCb VELO Timepix3 Telescope*, `arXiv:1902.09755`.

[157] J. Visser *et al.*, *SPIDR: a read-out system for Medipix3 & Timepix3*, JINST **10** (2015), no. 12 C12028.

[158] M. Clemencic *et al.*, *Recent developments in the LHCb software framework Gaudi*, J. Phys. Conf. Ser. **219** (2010) 042006.

[159] V. Blobel, *Software alignment for tracking detectors*, Nucl. Instrum. Meth. **A566** (2006) 5.

[160] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, .

[161] V. Franco Lima, *LHCb Vertex Locator Upgrade Development and Rare b-quark Decays in LHCb*, PhD thesis, 2019, Presented 03 Sep 2019.

[162] E. Dall'Occo, *The Timepix3 Telescope and Sensor R&D for the LHCb VELO Upgrade*, in *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, pp. 1–8, Nov, 2018. doi: 10.1109/NSSMIC.2018.8824535.

[163] S. E. Richards, *Characterisation of silicon detectors for the LHCb Vertex Locator Upgrade*, PhD thesis, Nov, 2017, Presented 15 Dec 2017.

[164] V. Chiochia *et al.*, *Simulation of the cms prototype silicon pixel sensors and comparison with test beam measurements*, .

[165] S. Meroli, D. Passeri, and L. Servoli, *Measurement of charge collection efficiency profiles of cmos active pixel sensors*, Journal of Instrumentation **7** (2012), no. 09 P09011.

[166] B. Henrich, W. Bertl, K. Gabathuler, and R. Horisberger, *Depth Profile of Signal Charge Collected in Heavily Irradiated Silicon Pixels*, Tech. Rep. CMS-NOTE-1997-021, CERN, Geneva, Apr, 1997.

[167] E. J. Schioppa *et al.*, *Study of charge diffusion in a silicon detector using an energy sensitive pixel readout chip*, IEEE Transactions on Nuclear Science **62** (2015) 2349.

[168] L. Servoli, S. Meroli, D. Passeri, and P. Tucceri, *Measurement of submicrometric intrinsic spatial resolution for active pixel sensors*, Journal of Instrumentation **8** (2013), no. 11 P11007.

[169] G. Kramberger *et al.*, *Investigation of Irradiated Silicon Detectors by Edge-TCT*, IEEE Trans. Nucl. Sci. **57** (2010), no. 4 2294.

[170] I. Abt *et al.*, *Characterization of silicon microstrip detectors using an infrared laser system*, Nucl. Instrum. Meth. **A423** (1999) 303.

[171] BaBar, P. del Amo Sanchez *et al.*, *Study of $B \to X\gamma$ Decays and Determination of $|V_{td}/V_{ts}|$*, Phys. Rev. **D82** (2010) 051101, `arXiv:1005.4087`.

[172] N. Croitoru, P. G. Rancoita, and A. Seidman, *Charge migration contribution to the sensitive layer of a silicon detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **234** (1985), no. 3 443 .

[173] P. Tsopelas, *A silicon pixel detector for LHCb*, PhD thesis, 2016, presented 21 Nov 2016.

[174] L. Snoj, G. Žerovnik, and A. Trkov, *Computational analysis of irradiation facilities at the jsi triga reactor*, Applied Radiation and Isotopes **70** (2012), no. 3 483 .

[175] A. Ducourthial *et al.*, *Performance of Thin Planar n-on-p silicon pixels after HL-LHC radiation fluences*, `arXiv:1810.07279`.

[176] M. Glaser *et al.*, *Radiation test facilities in the new PS East Hall at CERN*, in *1999 Fifth European Conference on Radiation and Its Effects on Components and Systems. RADECS 99 (Cat. No.99TH8471)*, pp. 136–141, 1999.

[177] M. Mikuz *et al.*, *Study of anomalous charge collection efficiency in heavily irradiated silicon strip detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **636** (2011), no. 1 S50 , 7th International Hiroshima Symposium on the Development and Application of Semiconductor Tracking Detectors.

[178] G. Casse *et al.*, *Evidence of enhanced signal response at high bias voltages in planar silicon detectors irradiated up to* $2.2 \times 10^{16} n_{eq} cm^{-2}$, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **636** (2011), no. 1 S56 , 7th International Hiroshima Symposium on the Development and Application of Semiconductor Tracking Detectors.

[179] J. Lange *et al.*, *Charge multiplication properties in highly irradiated epitaxial silicon detectors*, PoS **VERTEX2010** (2010) 025.

[180] A. Junkes, *Status of defect investigations*, Proceedings of Science (2011).

[181] LHCb collaboration, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, `arXiv:1808.08865`.

# Summary

The Standard Model of particle physics provides an accurate description of the fundamental constituents of matter and their interactions, but several experimental observations, such as the excess of matter over anti-matter, the existence of dark matter and oscillations of neutrinos between different flavours, remain unexplained. At the Large Hadron Collider (LHC) these puzzles are investigated by colliding proton beams at high energies and studying the products of these collisions. Detectors are placed at the four collision points around the LHC and collected data at the centre of mass energies of 7 TeV (2011), 8 TeV (2012) and 13 TeV (2015-2018). The design and construction of such detectors and their upgrades represent a technological challenge. A detector is composed of several different subdetectors, each of which plays a specific role in particle identification or track reconstruction and is often built with custom made components after a long R&D process. The detector used for the studies of this thesis is the LHCb detector. In the first part of the thesis a search for heavy neutrinos using data collected by the LHCb detector during 2012 is presented. The second part of the thesis is dedicated to the characterisation of novel silicon pixel sensors for the upgrade of the experiment.

### Search for heavy neutrinos

In the Standard Model (SM) neutrinos are introduced as massless particles. Many different experiments observed that neutrinos oscillate between flavours, leading to the conclusion that at least two neutrino types have a mass. This posed the question of how to account for the neutrino mass. In the simplest scenario neutrinos become massive via the same mechanism as the other leptons and quarks, but there is no explanation for the smallness of their mass. Another possibility is that neutrinos are Majorana particles, that is they are their own antiparticles. In this case, a mechanism known as *see-saw* could explain the smallness of the neutrino masses by introducing new, heavy, particles. The see-saw type I (there are three different mechanism types) adds a number of heavy right-handed neutrinos equal to the number of massive neutrinos. These heavy neutrinos are known as sterile because they can interact only through mixing with the active standard neutrinos. From an experimental point of view this means that heavy neutrinos can be searched for in any decay where active neutrinos are involved. The

two observables that are investigated are the mass of the neutrino and the strength of the mixing with the active neutrino. Depending on the type of search, different areas of the parameter space of these two observables have been probed, but large areas are still unexplored for the mixing of the heavy neutrino with every flavour of active neutrinos.

In this thesis a heavy neutrino in the mass range 5-50 GeV is searched using the data collected with the LHCb detector at the centre of mass energy of 8 TeV. The decay process $W \to \mu N$ with the heavy neutrino decaying promptly as $N \to \mu\ jet$ is investigated exploiting the large branching ratio to $q\overline{q}$ pairs, the high muon reconstruction efficiency of LHCb, and the full reconstruction of the $W$ from its decay products. Both final states with opposite-sign and same-sign muon pairs are considered, because both processes are allowed if the neutrino is a Majorana particle. The main backgrounds are charged weak currents, neutral electroweak processes, heavy and light flavour decays with a muon in the final state.

First, the backgrounds are suppressed by applying requirements on the kinematics and topology of the event and on three custom trained multivariate classifiers. Then, in order to estimate the amount of the remaining backgrounds the dataset has been split in a signal region and two orthogonal regions, denoted control regions, enhanced in the two main backgrounds. The amount of each background is measured in the respective control region and extrapolated to the signal region, yielding 5 and 1600 for same-sign and opposite-sign muons, respectively. The latter is much higher due to the irreducible background from neutral electroweak processes. The number of observed events in data is in agreement with the expected number of events and no signal is observed. Hence, an upper bound is set in the parameter space of the mixing of the heavy neutrino with the muon neutrino for both final state, as can be seen in Figure S.1.

### Silicon sensors for the VELO upgrade

Silicon detectors are widely exploited in high energy physics for vertexing and precision tracking detectors. In the LHCb detector the VErtex LOcator (VELO), the tracking station before the magnet and the inner part of the three tracking stations after the magnet consist of silicon strip detector modules. With the Upgrade I of the detector, currently being constructed and installed (2019-2021), the tracking detectors are being replaced in order to be able to cope with a fivefold increase in terms of pile-up, radiation exposure and track multiplicity. For the VELO, which plays a crucial role as part of the tracking system and in the reconstruction of primary and secondary vertices, strip detectors are replaced by pixel sensors.

The main challenge from the sensor point of view is the high and non uniform ra-

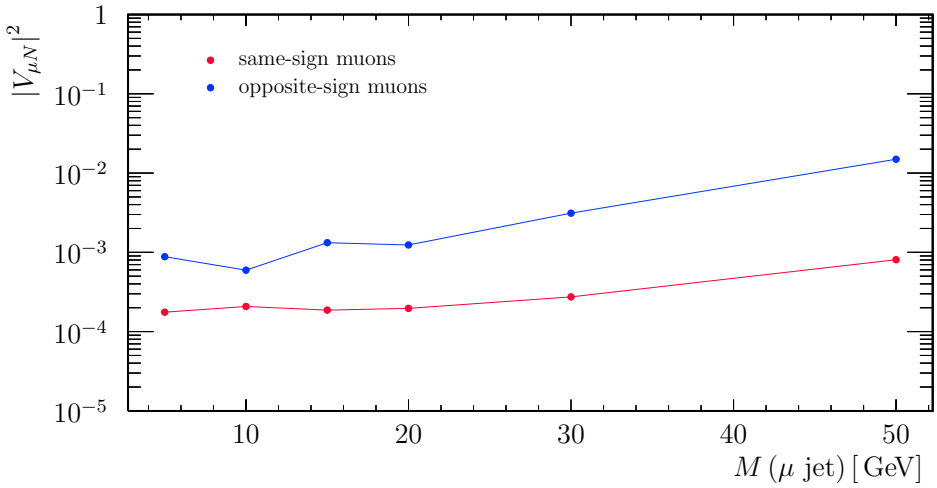



**Figure S.1:** Measured upper limit on the mixing between a heavy neutrino and the muon neutrino in the mass range 5-50 GeV for same-sign and opposite-sign muons in the final states.

diation exposure, with a maximum fluence of $8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ expected in the closest point to the proton-proton collision at the end of the lifetime of the upgraded VELO. In order to assess the best candidate for the VELO upgrade, several types of sensors have been studied, which differ by bulk type, thickness, implant size and guard ring size. A mixed hadron beam of minimum ionising particles at the Super Proton Synchrotron (SPS) has been used to characterise different prototype sensors. The Timepix3 telescope, which consists of eight silicon pixel planes, has been employed to reconstruct particle tracks with a spacial precision better than 2 µm at the position of the device under test and a time resolution of 370 ps per track to evaluate the response of the prototype sensors. The Timepix3 chip, which names the telescope, is used for the readout of the telescope planes as well as the prototype sensors.

Several aspects of the sensors have been investigated before and after irradiation up to the maximum fluence: the tolerance to high voltage, the charge collection and cluster finding efficiency, the spatial resolution and the performance close to the edge of pixel matrix. The latter being studied in detail in this thesis. Around the pixel matrix guard ring structures are placed, which provide a gradual drop of potential towards the edge of the sensor. The inactive area where the guard rings are located extends the physical edge of the sensor and can vary between 250 µm and 600 µm depending on the sensor type. Among all the prototypes tested, a family of sensors, coloured in blue in Figure S.2, exhibited a different behaviour with respect to the other sensors. Additional charge is collected from the guard ring region in the first column of the pixel matrix. Since at the edge the occupancy is already the highest, being the closest part to the

beam, this would lead to a degradation of resolution for the first measured point of the track. This played an important role in the decision of the sensor type, which resulted in 200 µm thick n-on-p sensors by Hamamatsu with 39 µm implant width and 450 µm guard ring size.

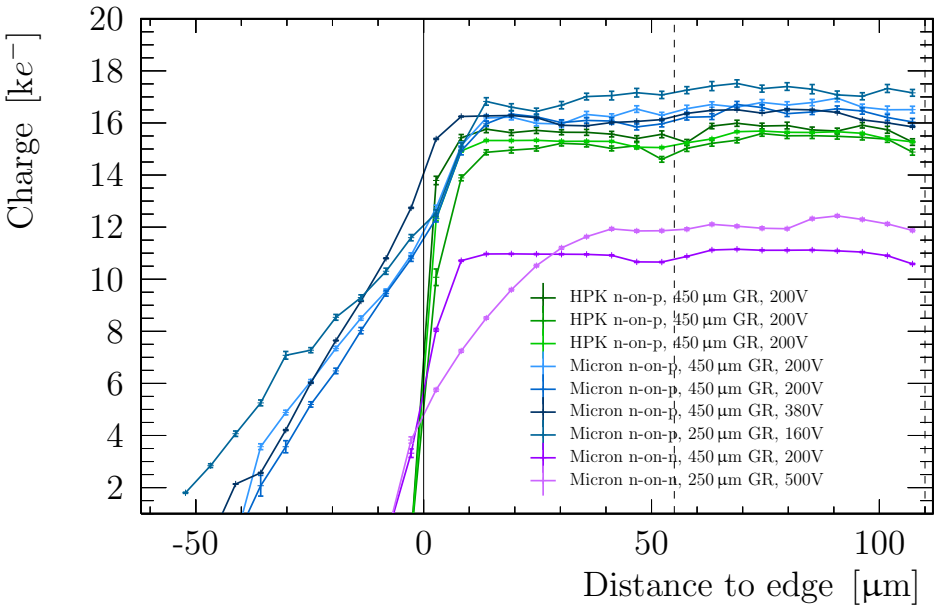



**Figure S.2:** Charge as a function of the distance to the edge of the track intercept for all the nonirradiated assemblies tested. The solid line represents the border of the pixel matrix and the dashed lines the borders between pixels. The label GR states for *guard rings*.

The radiation damage is foreseen to worsen with the increase in luminosity of the future upgrades. Hence, it is crucial to understand the effects of irradiation on silicon. With the grazing angle technique, studied and further developed in this thesis, the evolution of charge collection in the bulk of the sensor can be investigated. The grazing angle technique consists of placing the sensor almost parallel to the beam such that particles cross it longitudinally depositing energy in long clusters, as shown in Figure S.3. This makes it possible to study both of the charge collection and the time needed to cross the threshold, and hence give a signal, as a function of depth.

Charge collection and time to threshold profiles have been studied before and after full uniform irradiation fluence, leading to the observation of several radiation induced defects. In general a degradation of charge collection is observed, since most of the sensor types can no longer reach full depletion despite increasing the applied voltage up to 1000 V, meaning that the electron-hole pairs liberated in part of the volume of the detector do not contribute to the electric signal. Furthermore, the charge collected decreases with distance from the electrodes due to the trapping of charge carriers by radiation

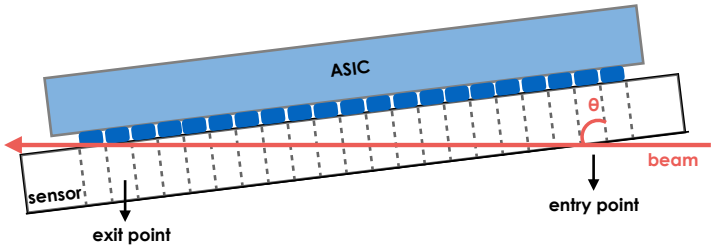



**Figure S.3:** Illustration of the grazing angle setup from the top. The entry and exit point of the track are indicated, as well as the angle $\theta$ of the track with respect to the sensor typically between 83 and 85 degrees.

induced defects in silicon. The smaller amount of charge collected is reflected in the increase of time to overcome the threshold, known as the timewalk effect. The evolution of both charge collection and time to threshold profiles with fluence has been studied with nonuniformly irradiated sensors. The grazing angle is shown to be a powerful technique that will play a crucial role in future development of silicon pixel detectors for high energy physics applications, especially in view of the fact that fast sensor and high precision timing information will become essential.

# Samenvatting

Het standaardmodel van de deeltjesfysica biedt een nauwkeurige beschrijving van de fundamentele bouwstenen der materie, tezamen met hun interacties. Echter, dit model kan verschillende experimentele resultaten, waaronder de overvloed van materie ten opzichte van anti-materie, het bestaan van donkere materie, en neutrino-oscillaties niet verklaren. Door het bestuderen van de restproducten van botsingen van hoogenergetische protonenbundels worden deze puzzels onderzocht met de Large Hadron Collider (LHC). Op vier verschillende punten langs de ring van de LHC komen de bundels samen met een totale botsingsenergie van 7 (2011), 8 (2012) of 13 TeV (2015-2018), en verzamelen detectoren data over de vrijgekomen deeltjes. Het ontwerpen, construeren en verbeteren van dergelijke meetinstrumenten brengen grote technologische uitdagingen met zich mee. Een detector bestaat uit verschillende subdetectoren, welke ieder een specifieke rol heeft in het identificeren van deeltjes of spoorreconstructie. Deze toepassing zorgt voor unieke eisen, en de componenten van de subdetectoren worden vaak, na een lang R&D proces, op maat gemaakt. De detector die gebruikt is voor de studies in dit proefschrift is de LHCb detector. In het eerste deel wordt een zoektocht naar zware neutrinodeeltjes (in de data verzameld in 2012) gepresenteerd. Het tweede deel van deze dissertatie is geweid aan de kenmerken van nieuwe silicium pixelsensoren voor een geplande, grote verbeteringsronde van het experiment.

### Zoektocht naar zware neutrino's

In het Standaard Model (SM) worden neutrino's geïntroduceerd als massaloze deeltjes. Veel verschillende experimenten hebben intussen oscillaties van neutrino's tussen de verschillende smaken (e, mu, tau) waargenomen. Dit heeft geleid tot de conclusie dat ten minste twee typen neutrino's massa hebben. Ook heeft dit geleid tot de vraag waar de massa van de neutrino's vandaan komt. Op deze vraag zijn er verschillende mogelijke antwoorden. In het meest simpele scenario krijgen neutrino's hun massa via hetzelfde mechanisme als de andere leptonen en quarks in het SM. Dit scenario verklaart helaas niet waarom de massa's van de neutrino's zo klein zijn. Een ander scenario is dat de neutrino's Majorana deeltjes zijn, en dus hun eigen anti-deeltje. In dat geval kan een mechanisme, dat bekend staat als *See-Saw*, de geringe massa's verklaren door hulp in te roepen van nieuwe, zware deeltjes. Het See-Saw mechanisme van het Type I (er zijn drie

verschillende types) voegt een aantal zware rechts-handige neutrino's toe, evenveel als het aantal neutrino's met massa. Deze zware neutrino's worden steriel genoemd, omdat hun wisselwerking met de bekende, actieve, neutrino's beperkt is tot een mixingsproces. Dit betekent dat men kan zoeken naar zware neutrino's door de gewone, actieve neutrino's te bestuderen in het lab. Twee belangrijke grootheden die komen kijken bij deze zoektocht zijn de massa van de zware neutrino en de sterkte van het mixen met de actieve neutrino. Alle mogelijke combinaties van waarden voor deze grootheden spannen samen de parameterruimte van dit model. Verschillende gebieden van de parameterruimte van deze twee grootheden zijn reeds onderzocht en uitgesloten. Echter, voor het model waarin het zware neutrino kan mixen met elke smaak van de actieve neutrino's is er nog veel van de parameterruimte open voor een ontdekking.

Dit proefschrift presenteert een zoektocht naar een zwaar neutrino met een massa tussen de 5 en 50 GeV in data van 8 TeV proton-proton botsingen verzameld met de LHCb detector. Er is gezocht naar een signaal van het zware neutrino via het verval $W \to \mu N$, waarbij het zware neutrino, N, direct zelf vervalt als $N \to \mu$jet. Dit maakt goed gebruik van de zeer efficiente reconstructie van muonen met de LHCb detector, de (verwachte) hoge vertakkingsratio van het verval van het zware neutrino en, als laatste, de volledige reconstructie van het W boson via de vervalsproducten. Combinaties waarin de uiteindelijke twee muonen dezelfde lading hebben, alsmede combinaties met een tegenovergestelde lading worden beschouwd, aangezien beide processen mogelijk zijn als het neutrino een Majorana deeltje is. De achtergronden zijn hoofdzakelijk afkomstig van van elektrozwakke processen en vervallen van zware of lichte hadronen waarin een muon wordt geproduceerd.

Om uit de grote hoeveelheid data een signaal te kunnen selecteren van het zware neutrino volgen een aantal stappen. Ten eerste worden de achtergronden onderdrukt met strikte vereisten op de kinematica en topologie van de kandidaten. Vervolgens wordt de dataset gesplitst in een signaalregio en twee controleregio's, die vooral bestaan uit de twee (bekende) achtergronden. De controleregio's zijn nuttig om een een goede inschatting te maken van de hoeveelheid achtergrond dat niet bij de eerste stap verwijderd is. Door middel van extrapolatie is de achtergrond in de signaalregio ingeschat op 5 (1600) kandidaten voor de dataset met de muonen van gelijke (tegenovergestelde) lading. De achtergrond in de laatste dataset is hoger door de grote achtergrond van neutrale elektrozwakke processen. Het aantal signaalkandidaten in de uiteindelijke dataset komt overeen met de verwachting van het SM, namelijk dat geen signaal van een zwaar neutrino wordt waargenomen. Daarom kan een deel van de parameterruimte van de zware neutrino worden uitgesloten, zoals weergegeven in Figuur S.4. De speurtocht naar zware neutrino's wordt voortgezet in Run-3 van de LHC versneller, waarvoor de LHCb detector een grote upgrade ondergaat. Het tweede deel van het proefschrift beschrijft het

R&D onderzoek naar een nieuwe vertex detector sensor voor de LHCb Velo.

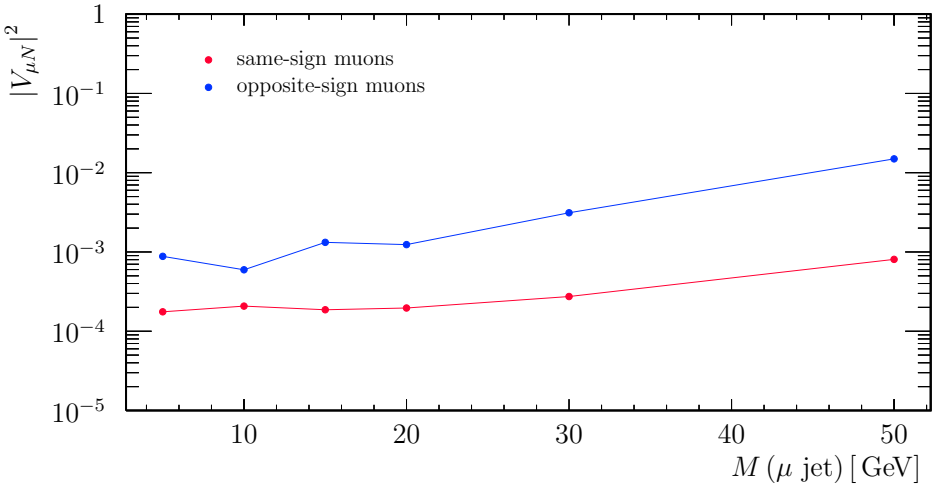



**Figure S.4:** Gezette limieten voor het mixingsproces van zware neutrino's met muon neutrino's voor massa's tussen de 5 en 50 GeV voor de metingen van muonen van gelijke en tegenovergestelde lading.

### Siliciumsensoren voor de VELO upgrade

Siliciumdetectoren worden veel gebruikt in de hoge-energiefysica voor vertexing en deeltjesdetectoren met een hoge precisie. In de LHCb detector bestaan de VErtex LOcator (VELO), de deeltjesdetectoren voor de magneet en het centrale deel van de drie deeltjesdetectoren na de magneet uit modules van silicium strips. Met de eerste upgrade van de LHCb detector, die op dit moment wordt gebouwd en geinstalleerd (2019-2021), worden de deeltjesdetectoren vervangen om om te kunnen gaan met de vervijfvoudiging van de pile-up, wat een verhoogde stralingdosis en deeltjesmultipliciteit met zich mee brengt. Voor de VELO, die een cruciale rol speelt in de reconstructie van sporen en vertices, zullen de stripsensoren worden vervangen door pixels.

De grote uitdaging voor de sensoren is de hoge, niet-uniforme blootstelling aan straling. Aan het eind van de levensduur van de upgrade VELO wordt een een maximale fluence verwacht van $8 \times 10^{15} \cdot 1\,\mathrm{MeV\,n_{eq}\,cm^{-2}}$ voor het dichtstbijzijnde punt van de proton-proton botsingen. Om het beste type sensor voor de VELO upgrade te kiezen zijn verschillende sensoren bestudeerd, met verschillende *bulk types*, diktes, de grootte van het implantaat en de grootte van de *guard ring*. Een gemengde hadronenbundel van minimum ioniserende deeltjes bij de Super Proton Synchrotron (SPS) is gebruikt om de karakteristieke eigenschappen van de verschillende prototype sensoren te onderzoeken. Om precieze studies uit te voeren is er gebruik gemaakt van de Timepix3 telescoop, die zelf bestaat uit acht meetvlakken met silicium pixels. Deeltjes die door de telescoop

reizen kunnen zeer precies worden gereconstrueerd, waardoor de positie van deeltjes in de prototypesensor kan worden bepaald met een resolutie van minder dan 2 µm. Ook meet deze opstelling de tijd met een resolutie van 370 ps, waarmee de respons in de tijd is onderzocht. De Timepix3 chip, waar de telescoop naar vernoemd is, is gebruikt voor zowel het uitlezen van de siliciumsensoren van de telscoop als de prototypesensor.

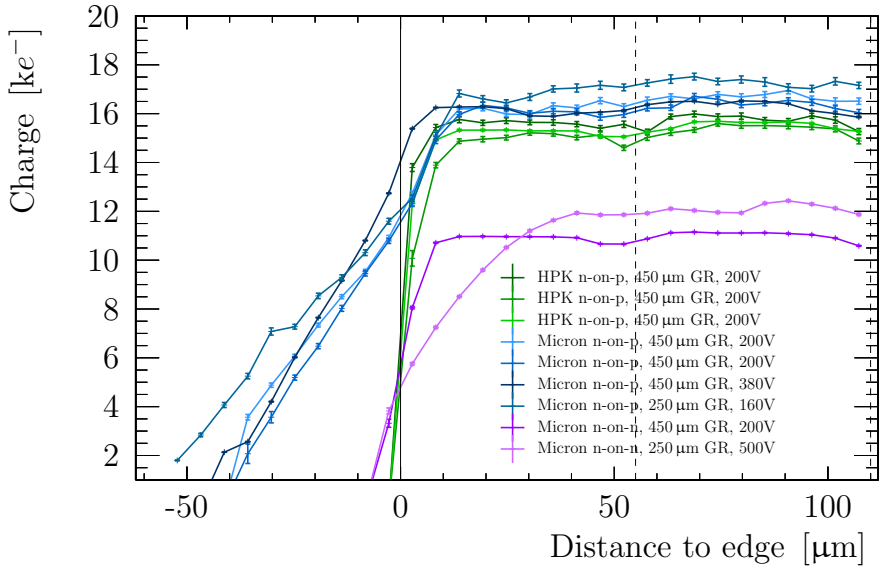



**Figure S.5:** Lading als een functie van de afstand tussen het spoor en de rand van de pixelmatrix voor de verschillende opties, waarbij de sensor vrij is van stralingsschade. De doorlopende lijn weerspiegelt de rand van de pixelmatrix, en de gestreepte lijnen de grenzen van de individuele pixels. De afkorting "GR", gebruikt in de legenda, staat voor *guard rings*.

Verschillende aspecten van de sensoren zijn onderzocht vóór, en na de bestraling tot de maximale fluence: de tolerantie voor hoogspanning, de ladingsopbrengst, de clusteringsefficiëntie, de hitresolutie en, in meer detail, de prestaties van de sensor dichtbij de rand van de pixelmatrix. Rond de pixelmatrix zijn guard rings geplaatst die voor een geleidelijke afname van de potentiaal zorgen aan de rand van de sensor. Door het ongevoelige stuk van de sensoren waar deze guard rings zitten zijn de sensoren groter: afhankelijk van het type sensor, tussen de 250 µm en 600 µm. Bij de verschillende geteste prototypes zat ook een klasse van sensoren, die in blauw zijn gekenmerkt in Figuur S.5. Deze laten een ander gedrag zien in de buurt van de randen: extra lading wordt verzameld in de eerste kolom van de pixel matrix, afkomstig van de regio van de guard ring. Omdat bij de rand de verwachte bezettingsgraad al de hoogste is van de hele sensor (dit zit immers dichtbij de bundel), zou dit leiden tot een vermindering van de resolutie voor de eerste hit op een spoor. Uiteindelijk is er gekozen voor de 200 µm dikke n-on-p sensors van Hamamatsu, met een implantaatdikte van 39 µm en een guard ring

van 450 μm. Het eerder genoemde effect heeft een belangrijke rol gespeeld bij deze keuze.

De stralingsschade zal drastisch toenemen met de verhoging van de luminositeit in de volgende upgrades. Het is dus cruciaal om de effecten van straling op silicium goed te begrijpen. Om de verandering van de ladingsopbrengst in de bulk van de sensor te onderzoeken wordt er gebruik gemaakt van metingen waarbij de sensor onder een grote hoek ten opzichte van de bundel geplaatst was, waardoor de sensor bijna parallel aan de bundel staat. In dit geval schrapen de deeltjes langs de pixels, waarbij het energie achterlaat in de sensor in de vorm van lange clusters. Op deze manier kunnen zowel de ladingsopbrengst als de tijd die nodig is om over de drempelwaarde te komen worden onderzocht als een functie van de diepte. De ladingsopbrengst en de tijd-tot-drempelwaarde zijn al eerder bestudeerd voor sensoren die niet, of volledig (uniform) aan straling zijn blootgesteld, waarbij stralingsschade werd waargenomen.

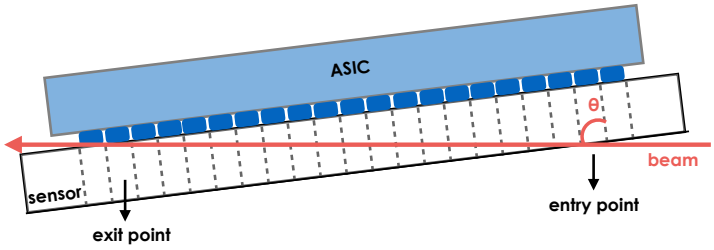



**Figure S.6:** Schematische weergave van een sensor die met een grote hoek ten opzichte van de bundel geplaatst is. Het begin- en eindpunt van het spoor door de sensor zijn aangegeven met pijlen, samen met de hoek $\theta$ van het spoor ten opzichte van de sensor. Deze hoek is typisch tussen de 83 en 85 graden.

Over het algemeen wordt er een verlies van de ladingsopbrengst gezien nadat de sensors een tijd zijn blootgesteld aan straling. De meeste sensoren kunnen niet langer terugvallen naar hun rusttoestand, zelfs als de spanning wordt verhoogd tot 1000V. Dat betekent dat een deel van de vrijgekomen paren van elektronen en elektrongaten niet meer bijdragen aan een elektrisch signaal. Bovendien neemt de ladingsopbrengst af naarmate de afstand van de electrodes toeneemt, doordat ladingsdragers worden ingevangen als een gevolg van de stralingsschade. De afname van de verzamelde lading is merkbaar door een toename van de benodigde tijd om boven de drempelwaarde te komen, een effect dat *timewalk* wordt genoemd. Dit proefschrift presenteert een studie van de effecten van straling op de ladingsopbrengst en de timewalk voor sensoren die op een niet-uniforme wijze zijn bestraald. In het bijzonder is de techniek waarbij de sensor onder hoek ten opzichte van de bundel is geplaatst zeer behulpzaam. Zij zal ook in de toekomst een belangrijke rol spelen in de ontwikkeling van silicium pixel detectoren voor hoge-energie fysica, inachtnemend dat snelle sensoren en precieze tijdsinformatie van essentieel belang zullen zijn.

# Acknowledgments

When you start a PhD, four years seems like a very long time; in reality, it goes so fast that one day you are the "new one" and the one after you have to start writing. Here I want to thank all the people that, directly or indirectly, contributed to this thesis.

First, a giant thanks goes to my supervisors, Martin and Wouter, and my promotor, Marcel. Since day one you all have been a well of knowledge and expertise and your guidance has been fundamental for my personal and professional development. Martin, thank you for your patience in teaching me about sensors and for your way of guiding me in understanding weird plots or in finding answers to my questions. Wouter, it was great to do a new analysis from scratch. Your positive attitude and directness always pushed me to do better. Thank you for that. Marcel, you are the glue of the group. Not only you are a reference point from the physics point of view, but also take care of the well being of every person in the group. I admire your curiosity in any physics topic and your constant willingness to help in any field. Thanks to all of you for giving me the opportunity to work at Nikhef, I could have not asked for a better place and people to carry out my PhD with.

I am grateful to have been part of the INFIERI project. During the schools and workshops I learned a lot, but also met some great people. Thanks to Davide, Maria Teresa, Daniele, André and Luis for the moments we shared around the world.

The testbeam times are unforgettable. Not only I had the chance to learn and work with fantastic experts, but I also had a lot of fun. Thanks to the VELO crew Paula, Kazu, Heinrich, Daniel, Mateus. Emma and Vinny, we met in the shiny lights of the north area and you have been at my side since then. Panos, you have been a great first teacher (I owe you also my driving skills of the CERN van) and one of the first friends at Nikhef. Thank you.

To the VELO operations team, Karol, Kurt, Sasha, a big thanks for the many things you taught me and for always answering my questions, even at 4 in the morning. I could always rely on you.

One of the things I really like at Nikhef is that it is such a big group of enthusiastic physicists. Thanks to Antonio, Niels, Hella, Gerhard, Patrick, Eddie, Jeroen and Tjieerd, backbone of the group. Throughout a few generations of PhD and postdocs I want to thank Laurent, Jacco, Lennart, Roel, Greg, Flaffio, Krista, Vasilis, Stergios, Ivan, Carlos, Mauricio, Roman, Michele, Veronica, Maxime, Sevda, Brian, Igor, Jordi, Hilbrand, and last but not least the new recruits Lera, Sese, Aleksandra, Marjolein, Robbert and Kevin. It has been a pleasure to share coffee breaks, borrels and night outs with you. Silvia, Cristina and Miriam, I am happy we overlapped for at least a short period of time. You brought southern energy into the group. To my co-PhDs, Maarten, Katya and Mick, it was great to share these years with you. Pieter, thanks for welcoming me, answering my questions about physics and helping me starting up in a new city. Thanks Sean for your constant support, our coffees in the early morning were the best way to start the day.

A special thanks to Afroditi, Lucrezia, Maria, Becky, Dimitra and Nat. I really enjoyed our girls times in Amsterdam/Geneva. Dimitra you have been a pillar during these four years, from Duwo to Saint-Genis-Pouilly.

To my friends of always, Baru, Civi, Cerru, Meg, Marco Armani, Pally, Vale, thank you for being there despite the physical distance. A special thanks to Cateri, for designing the beautiful cover of this thesis.

To my family, Fri, zia and especially mami and papi and nonni, you are always there through ups and downs, with innumerable trips through Europe to bring decent prosciutto and biscotti, help with movings and simply to see each other. Thanks for your constant support. Tim, we have been through three different countries, uncountable number of flights, two PhD theses and in all of this mess we never stopped having fun. Thanks for being there.