

# Lepton Flavour Violation searches and scintillating fibre tracker testing in LHCb

Thèse n. 9251 2018  
présenté le 24 janvier 2019  
à la Faculté des Sciences de Base  
laboratoire LPHE  
programme doctoral en Physique  
École Polytechnique Fédérale de Lausanne  
pour l'obtention du grade de Docteur ès Sciences  
par

Guido Andreassi

acceptée sur proposition du jury:

Prof Paolo Ricci, président du jury  
Prof Aurelio Bay, directeur de thèse  
Prof Frédéric Courbin, rapporteur  
Prof Johannes Albrecht, rapporteur  
Dr Flavio Archilli, rapporteur

Lausanne, EPFL, 2019



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE





# Acknowledgements

These have been four intense years; they went by fast, and yet they have been rich of events, changes and milestones that define by sharp lines the person I am today.

I received in 2015 the valuable opportunity to start a PhD in the LPHE laboratory at EPFL to whose leaders and professors – Aurelio Bay, Olivier Schneider, Tatsuya Nakada, Frédéric Blanc and Guido Haefeli – I will always be grateful: they have welcomed me and formed me as a Physicist. Prof. Aurelio Bay has been the best possible thesis director, as he let me free to set my path, without ever leaving me alone. He supported my ambitions, corrected my mistakes and always provided new input if I was getting stuck.

There are, then, those who followed me more closely, in my everyday work. To Flavio and Luca, I owe most of what I learned about research, its methods, its techniques. We did not just work together; we also became good friends, which I am very happy for.

I had the occasion to work with other colleagues and friends, although for shorter periods: Plamen, Maria Elena, Chitsanu and many others. It has been a wonderful trip together and I learned a lot from each of you! Also all the other colleagues from LPHE, with whom I did not have the occasion to work directly, deserve undoubtedly to be thanked: you are a fantastic group, which I am proud of having been part of. We started friendships that I hope will continue in the future.

But my activity took place, for a large portion, at CERN: there I shared coffees and beers (mostly beers...) with a noisy group of compatriots; some are very old friends, with some others I met there, and I am very glad about it. *Grazie ragazzi!*

Who really made all this experience special, though, is the one beauty that does not decay: Lucrezia. Four years ago we both left our hometown, Rome, for different destinations. Distance did not scare us, and we can now certainly say that we have been an invincible team. You have been my strength and made these years much more meaningful.

The first in line to believe in me has always been my family. They helped me in finding and chasing my dreams. When I was a kid, while tucking me in, my parents were encouraging me to hop onboard the *carro dei sogni* (the wagon of dreams). Today this wagon keeps on driving me through the more complex ambitions of an adult. My mother, my father, my grandmother and my sister Camilla are still on my side in this journey. They are my points of reference.





# Abstract

This thesis presents new results on searches for lepton-flavour-violating particle decays, performed in the LHCb experiment at CERN.

Two analyses are discussed in detail. The first one concerns the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  decay. No evidence of signal is observed and two upper limits are set, respectively for the  $B^0$  and  $B_s^0$  mesons:  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 1.0(1.3) \times 10^{-9}$  and  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 6.0(7.2) \times 10^{-9}$ . The second analysis concerns the baryon sector: the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  decay is searched. This study is currently under review in the LHCb collaboration and the information shown here should be considered as preliminary. The whole procedure is described here in detail.

Furthermore, the present thesis contains a description and the results of the quality assurance process for the production of mats of scintillating fibres for the SciFi, a new tracker for the LHCb detector developed in view of the Run III of LHC.

Key words: particle physics, LHCb, LHC, rare decays, flavour, lepton flavour violation, tracker, scintillating fibres.



# Résumé

Cette thèse présente des nouveaux résultats des recherches de désintégrations de particules qui violent la conservation de la saveur leptonique, effectuées par l'expérience LHCb au CERN.

Deux analyses sont discutées. La première concerne la désintégration  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ . Aucun signal n'est observé, et deux limites supérieures sont établis, respectivement pour le meson  $B_s^0$  et  $B_s^0$  :  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 1.0(1.3) \times 10^{-9}$  et  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 6.0(7.2) \times 10^{-9}$ . La seconde analyse concerne le secteur des baryons : la désintégration  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  est recherchée. Cette étude est actuellement en phase de révision dans la collaboration LHCb et les informations montrées dans cette thèse doivent être considérées comme préliminaires. L'entière procédure est décrite ici en détail.

En outre, cette thèse contient une description et les résultats du processus de assurance qualité pour la production de *mats* de fibres scintillantes pour le SciFi, un nouveau trajectographe pour le détecteur LHCb, développé en vue du Run III de LHC.

Mots clefs : physique des particules, LHCb, LHC, désintégrations rares, saveur, violation des saveurs leptoniques, trajectographe, fibres scintillantes.



# Contents

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| <b>Acknowledgements</b>                                                 | <b>i</b>   |
| <b>Abstract (English/Français)</b>                                      | <b>iii</b> |
| <b>List of figures</b>                                                  | <b>xi</b>  |
| <b>List of tables</b>                                                   | <b>xix</b> |
| <b>1 Introduction</b>                                                   | <b>1</b>   |
| <b>2 Theory</b>                                                         | <b>3</b>   |
| 2.1 The Standard Model of Particle Physics . . . . .                    | 3          |
| 2.1.1 Interactions . . . . .                                            | 4          |
| 2.1.2 Three generations . . . . .                                       | 6          |
| 2.1.3 Lepton Flavour Violation: current theoretical scenario . . . . .  | 10         |
| 2.1.4 Lepton Flavour Violation: current experimental scenario . . . . . | 11         |
| 2.2 Challenges of the Standard Model . . . . .                          | 11         |
| <b>3 The LHCb experiment</b>                                            | <b>13</b>  |
| 3.1 The Large Hadron Collider . . . . .                                 | 13         |
| 3.2 The Large Hadron Collider beauty experiment: LHCb . . . . .         | 15         |
| 3.3 Tracking system . . . . .                                           | 20         |
| 3.3.1 Vertex Locator . . . . .                                          | 20         |
| 3.3.2 The dipole magnet . . . . .                                       | 24         |
| 3.3.3 Tracker Turicensis . . . . .                                      | 24         |
| 3.3.4 Inner Tracker (IT) . . . . .                                      | 25         |
| 3.3.5 Outer Tracker (OT) . . . . .                                      | 26         |
| 3.4 Particle Identification System . . . . .                            | 28         |
| 3.4.1 The RICH detectors . . . . .                                      | 28         |
| 3.4.2 The calorimeter system . . . . .                                  | 29         |
| 3.4.3 The muon system . . . . .                                         | 31         |
| 3.5 Data treatment . . . . .                                            | 32         |
| 3.5.1 Particle Identification . . . . .                                 | 32         |
| 3.5.2 Trigger . . . . .                                                 | 34         |
| 3.5.3 LHCb software framework . . . . .                                 | 37         |

|          |                                                                                                                   |            |
|----------|-------------------------------------------------------------------------------------------------------------------|------------|
| 3.5.4    | Stripping . . . . .                                                                                               | 38         |
| 3.5.5    | Tracks classification . . . . .                                                                                   | 39         |
| 3.5.6    | Reconstruction of electrons . . . . .                                                                             | 40         |
| <b>4</b> | <b>The LHCb upgrade</b>                                                                                           | <b>43</b>  |
| 4.0.1    | Detector upgrade . . . . .                                                                                        | 44         |
| 4.0.2    | Trigger upgrade . . . . .                                                                                         | 44         |
| 4.1      | The Scintillating fibre Tracker - Sci-Fi . . . . .                                                                | 45         |
| 4.1.1    | Detector overview and layout . . . . .                                                                            | 47         |
| 4.1.2    | Fibre mats . . . . .                                                                                              | 50         |
| 4.1.3    | Fibre mats quality assurance . . . . .                                                                            | 53         |
| 4.2      | Light yield tests of fibre mats from EPFL . . . . .                                                               | 57         |
| 4.2.1    | Goal . . . . .                                                                                                    | 57         |
| 4.2.2    | Setup . . . . .                                                                                                   | 58         |
| 4.2.3    | Results . . . . .                                                                                                 | 60         |
| <b>5</b> | <b>Search for the lepton-flavour violating decays <math>B_{(s)}^0 \rightarrow e^\pm \mu^\mp</math></b>            | <b>65</b>  |
| 5.1      | Selection . . . . .                                                                                               | 82         |
| 5.1.1    | Trigger . . . . .                                                                                                 | 82         |
| 5.1.2    | Stripping and pre-selection . . . . .                                                                             | 83         |
| 5.1.3    | The BDTS discriminant . . . . .                                                                                   | 84         |
| 5.1.4    | PID . . . . .                                                                                                     | 86         |
| 5.1.5    | The BDT classifier . . . . .                                                                                      | 87         |
| 5.1.6    | Determination of the BDT PDF for signal . . . . .                                                                 | 91         |
| 5.2      | Calibration of the invariant mass . . . . .                                                                       | 100        |
| 5.2.1    | Data-driven correction to the mass resolution . . . . .                                                           | 102        |
| 5.2.2    | Fraction of HasBremAdded . . . . .                                                                                | 104        |
| 5.3      | Backgrounds . . . . .                                                                                             | 104        |
| 5.3.1    | Peaking backgrounds . . . . .                                                                                     | 105        |
| 5.4      | Normalisation . . . . .                                                                                           | 111        |
| 5.4.1    | Invariant-mass fits . . . . .                                                                                     | 111        |
| 5.4.2    | Normalisation factors . . . . .                                                                                   | 113        |
| 5.5      | Upper limit determination . . . . .                                                                               | 113        |
| 5.6      | Results . . . . .                                                                                                 | 115        |
| 5.7      | Systematic errors . . . . .                                                                                       | 116        |
| 5.8      | Conclusions . . . . .                                                                                             | 117        |
| <b>6</b> | <b>Search for the lepton-flavour violating decay <math>\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp</math></b> | <b>119</b> |
| 6.1      | Data and simulated samples . . . . .                                                                              | 120        |
| 6.1.1    | Data . . . . .                                                                                                    | 120        |
| 6.1.2    | Simulation . . . . .                                                                                              | 120        |
| 6.2      | Selection . . . . .                                                                                               | 121        |
| 6.2.1    | Trigger . . . . .                                                                                                 | 121        |

|                 |                                                                                                                             |            |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------|------------|
| 6.2.2           | Stripping . . . . .                                                                                                         | 122        |
| 6.2.3           | Pre-selection . . . . .                                                                                                     | 122        |
| 6.3             | Re-weighting of the signal simulated sample . . . . .                                                                       | 125        |
| 6.4             | Multivariate Classifier . . . . .                                                                                           | 126        |
| 6.4.1           | Definition . . . . .                                                                                                        | 126        |
| 6.4.2           | Training . . . . .                                                                                                          | 129        |
| 6.4.3           | Performances . . . . .                                                                                                      | 130        |
| 6.4.4           | Response on data . . . . .                                                                                                  | 131        |
| 6.4.5           | Cut optimisation . . . . .                                                                                                  | 135        |
| 6.5             | Backgrounds . . . . .                                                                                                       | 135        |
| 6.6             | Calibration of signal invariant mass . . . . .                                                                              | 138        |
| 6.6.1           | Signal model . . . . .                                                                                                      | 138        |
| 6.6.2           | Fit to $\Lambda_b^0 \rightarrow \Lambda^0 J/\psi (\rightarrow \mu\mu)$ . . . . .                                            | 142        |
| 6.6.3           | Signal model for $\Lambda_b^0 \rightarrow \Lambda^0 e\mu$ . . . . .                                                         | 143        |
| 6.7             | Normalisation . . . . .                                                                                                     | 144        |
| 6.7.1           | Normalisation channel yield . . . . .                                                                                       | 145        |
| 6.8             | Selection efficiencies . . . . .                                                                                            | 146        |
| 6.8.1           | Geometrical acceptance . . . . .                                                                                            | 148        |
| 6.8.2           | Trigger . . . . .                                                                                                           | 148        |
| 6.8.3           | Reconstruction and stripping . . . . .                                                                                      | 148        |
| 6.8.4           | Offline selection . . . . .                                                                                                 | 149        |
| 6.8.5           | Particle identification . . . . .                                                                                           | 149        |
| 6.8.6           | Selection on the BDT response . . . . .                                                                                     | 150        |
| 6.8.7           | Total Efficiency . . . . .                                                                                                  | 152        |
| 6.9             | Results . . . . .                                                                                                           | 152        |
| 6.9.1           | Invariant-mass fit . . . . .                                                                                                | 152        |
| 6.9.2           | CLs limit . . . . .                                                                                                         | 152        |
| 6.9.3           | Treatment of systematic uncertainties . . . . .                                                                             | 154        |
| 6.9.4           | Uncertainty on the branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ . . . . . | 155        |
| 6.10            | Conclusions and prospects . . . . .                                                                                         | 155        |
| <b>7</b>        | <b>Conclusion and outlook</b>                                                                                               | <b>159</b> |
| <b>Appendix</b> |                                                                                                                             | <b>161</b> |
| <b>A</b>        | <b>Mass resolution of <math>e - \mu</math> final states</b>                                                                 | <b>163</b> |
| A.0.1           | Invariant mass resolution in a relativistic 2-body decay . . . . .                                                          | 163        |
| A.0.2           | Extrapolation to the case of $\Lambda_b \rightarrow \Lambda^0 e\mu$ . . . . .                                               | 165        |
| A.0.3           | Energy correlations . . . . .                                                                                               | 165        |
| <b>B</b>        | <b>Invariant mass fits to <math>B^0 \rightarrow K^+ \pi^-</math></b>                                                        | <b>167</b> |
| <b>C</b>        | <b>Training features of the BDT classifier for <math>B_{(s)}^0 \rightarrow e^\pm \mu^\mp</math></b>                         | <b>171</b> |

## Contents

---

|          |                                                                                                                 |            |
|----------|-----------------------------------------------------------------------------------------------------------------|------------|
| <b>D</b> | <b>Selection efficiencies for <math>B_{(s)}^0 \rightarrow e^\pm \mu^\mp</math> backgrounds.</b>                 | <b>173</b> |
| <b>E</b> | <b>Training features of the BDT classifier for <math>\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp</math></b> | <b>175</b> |
| <b>F</b> | <b>HOP variables</b>                                                                                            | <b>179</b> |
|          | <b>Bibliography</b>                                                                                             | <b>188</b> |
|          | <b>Curriculum Vitae</b>                                                                                         | <b>189</b> |



# List of Figures

|     |                                                                                                                                                                                                                                                                                                                                                                                          |    |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1 | Schematic representation of the fundamental particles of the Standard Model.                                                                                                                                                                                                                                                                                                             | 4  |
| 2.2 | A sketch of the experimental setup used by Conversi, Pancini and Piccioni. Figure from [11]. . . . .                                                                                                                                                                                                                                                                                     | 6  |
| 2.3 | The lowest-order Feynman diagram for Michel decay. Figure from [1]. . . . .                                                                                                                                                                                                                                                                                                              | 6  |
| 2.4 | The weak interaction couplings of the d, s, u and c in terms of the Cabibbo angle, $\theta_C$ . Figure from [1]. . . . .                                                                                                                                                                                                                                                                 | 7  |
| 2.5 | The weak interaction couplings of the first generation of leptons, in their allowed helicity combinations, in the relativistic limit. Figure from [1]. . . . .                                                                                                                                                                                                                           | 7  |
| 2.6 | Examples of Feynman diagrams of LFV processes in beyond-standard-model theories. Figure adapted from [27]. . . . .                                                                                                                                                                                                                                                                       | 10 |
| 3.1 | The LHC Accelerator Complex. Figure from [62]. . . . .                                                                                                                                                                                                                                                                                                                                   | 14 |
| 3.2 | Recorded luminosity at LHCb. Figure from [63]. . . . .                                                                                                                                                                                                                                                                                                                                   | 15 |
| 3.3 | The LHCb detector, with its sub-detectors, side view. Figure from [64]. . . . .                                                                                                                                                                                                                                                                                                          | 16 |
| 3.4 | Angular distribution of $b\bar{b}$ quarks couples produced in $pp$ collisions at LHC at a centre-of-mass energy of 7 (left), 8 (centre) and 13 (right) TeV. $\theta_b$ and $\theta_{\bar{b}}$ are the angles of the quarks momentum with respect to the beam axis. Figures from [65].                                                                                                    | 17 |
| 3.5 | Standard Model cross sections as a function of the centre-of-mass energy, $\sqrt{s}$ , of the collider. The dashed lines corresponds to the Tevatron energy of 1.96 TeV and the nominal LHC energy of 14 TeV. The discontinuity in some of the cross sections at 4 TeV is due to the switch from proton-antiproton to proton-proton collisions at that energy. Figure from [66]. . . . . | 18 |
| 3.6 | Development of the instantaneous luminosity for LHCb, ATLAS and CMS during a typical LHC fill. After reaching the design value for LHCb, the luminosity is stabilised with a tolerance of 5% for about 15 hours by adjusting the transversal overlap of the colliding beams. Figure from [68]. . . . .                                                                                   | 19 |
| 3.7 | Average instantaneous luminosity over LHC fill number in 2012 (left) and 2017 (right). . . . .                                                                                                                                                                                                                                                                                           | 19 |
| 3.8 | Average number of visible interactions per bunch crossing $\mu$ over LHC fill number in 2012 (left) and 2017 (right). . . . .                                                                                                                                                                                                                                                            | 19 |

## List of Figures

---

|      |                                                                                                                                                                                                                                                                                                                                                                                                                             |    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.9  | Arrangement of the VELO silicon modules along the direction of the beam. The angles indicated by solid lines are: the crossing angle for minimum-bias events (60 mrad), minimum (15 mrad) and maximum (390 mrad) angle for which three or more stations are crossed. The front face of the first modules is illustrated, in the bottom part of the figure, in both its closed and open positions. Figure from [69]. . . . . | 21 |
| 3.10 | A schematic view of the R- and $\Phi$ - measuring sensors. On the $\Phi$ -sensor, the strips on two adjacent modules are shown, to highlight the stereo angle. Although the physical radius of the R-sensors is slightly bigger due to practical constraints, the sensitive area is the same. Figure from [64]. . . . .                                                                                                     | 22 |
| 3.11 | View of the inside of the secondary vacuum container of the VELO. The corrugations close to the beam axis allow to minimise the material seen by incoming tracks, while those at the side allow the overlap between the two halves of each module. Figure from [69]. . . . .                                                                                                                                                | 23 |
| 3.12 | Left: perspective view of the LHCb dipole magnet. Right: magnetic field intensity along the $z$ -axis. Figures from [70]. . . . .                                                                                                                                                                                                                                                                                           | 24 |
| 3.13 | Left: layout of TT layers. Right: structure of a half module. Figures from [71]. . .                                                                                                                                                                                                                                                                                                                                        | 25 |
| 3.14 | View of the four IT detector boxes arranged around the LHC beampipe. Figure from [64]. . . . .                                                                                                                                                                                                                                                                                                                              | 26 |
| 3.15 | Layout of an $x$ detector layer in the second IT station. The lengths shown are in cm and they refer to the active area of the Inner Tracker. Figure from [64]. . . .                                                                                                                                                                                                                                                       | 26 |
| 3.16 | Arrangement of OT straw-tube modules in layers and stations. . . . .                                                                                                                                                                                                                                                                                                                                                        | 27 |
| 3.17 | Left: section of an OT module, showing the arrangement of the tubes. Right: section of a single tube. Figures from [74]. . . . .                                                                                                                                                                                                                                                                                            | 28 |
| 3.18 | Left: side view of the RICH 1. The aerogel has been removed in Run II. Right: side view of the RICH 2. Figures from [76]. . . . .                                                                                                                                                                                                                                                                                           | 29 |
| 3.19 | Relationship between $\theta_c$ and the momentum for different particles for the three media in the two RICH sub-detectors. Figure from [64]. . . . .                                                                                                                                                                                                                                                                       | 30 |
| 3.20 | Lateral segmentation of the SPD/PS and ECAL (left) and the HCAL (right). One quarter of the detector front face is shown. The cell dimensions reported in the left figure refer to the ECAL. Figures from [78] . . . . .                                                                                                                                                                                                    | 31 |
| 3.21 | Side view of the muon system in the $y$ - $z$ plane. Figure from [79]. . . . .                                                                                                                                                                                                                                                                                                                                              | 32 |
| 3.22 | Left: front view of a quadrant of a muon station. Each rectangle represents one chamber. Right: division into logical pads of four chambers belonging to the different regions of station M1. Stations M2-M3 (M4-M5) have twice (half) the number of pad columns per chamber with respect to M1, in each corresponding region, while the number of pad rows per chamber is the same. Figures from [64].                     | 33 |
| 3.23 | Sketch of the interactions of different types of particles with the components of a <i>traditional</i> particle physics detector. Figure from [80]. . . . .                                                                                                                                                                                                                                                                 | 33 |
| 3.24 | LHCb trigger scheme during the Run I (left) and Run II (right) data taking periods. Figures from [82]. . . . .                                                                                                                                                                                                                                                                                                              | 36 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                         |    |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.25 | Illustration of the track types in the LHCb detector. The main component of the magnetic field $B_y$ is shown above as a function of the $z$ coordinate. Figure from [72]. . . . .                                                                                                                                                                                                                      | 39 |
| 4.1  | The LHC commissioning schedule from year 2015 to 2021. LS stands for Long Shutdown, while EYETS indicates the Extended Year-End Technical Stop. Figure from [94]. . . . .                                                                                                                                                                                                                               | 44 |
| 4.2  | LHCb trigger scheme foreseen for Run III. Figure from [82]. . . . .                                                                                                                                                                                                                                                                                                                                     | 45 |
| 4.3  | The LHCb detector as it will be after the upgrade, with its sub-detectors, side view.                                                                                                                                                                                                                                                                                                                   | 46 |
| 4.4  | A schematic view of one station of the SciFi Tracker. Figure from [95]. . . . .                                                                                                                                                                                                                                                                                                                         | 48 |
| 4.5  | The three stations of the SciFi tracker shown between the dipole magnet, on the left, and the RICH2, on the right. Figure from [95]. . . . .                                                                                                                                                                                                                                                            | 49 |
| 4.6  | SiPM array mounted on flex PCB on the left, zoom into the single channel size and the fibre mat on the right. Figure from [99]. . . . .                                                                                                                                                                                                                                                                 | 50 |
| 4.7  | Top: Package with two 64-channels silicon dies. Bottom left: the gap between two silicon dies shown under the microscope. Bottom right: a pixel with optical trenches. Figure from [95]. . . . .                                                                                                                                                                                                        | 51 |
| 4.8  | Sketch of the longitudinal section of a scintillating fibre. Light is produced in the core material and then trapped and propagated within the fibre through total internal reflection. The claddings refraction indices are indicated, along with the maximum incidence angles at which emitted light is captured, corresponding to the indicated 5.35% of the solid angle. Figure from [100]. . . . . | 52 |
| 4.9  | Part of the cross section of a fibre mat. . . . .                                                                                                                                                                                                                                                                                                                                                       | 52 |
| 4.10 | Segments of the cross section of a fibre mat produced at EPFL, illuminated at the opposite end to enhance contrast, in order to evaluate the cleanness of the cut and the regularity of the geometry. On the right, the fibre cladding is also clearly visible. . . . .                                                                                                                                 | 53 |
| 4.11 | The expected dose in the $x$ - $y$ plane at $z = 783$ cm (T1 position) after an integrated luminosity of $50 \text{ fb}^{-1}$ . Figure from [95]. . . . .                                                                                                                                                                                                                                               | 54 |
| 4.12 | The SiPM-end of a fibre mat produced at EPFL. The kapton foil (black reflective surface) and the endpiece are clearly visible. . . . .                                                                                                                                                                                                                                                                  | 55 |
| 4.13 | Scheme of the production steps and quality assurance measurements for the production of SciFi fibre mats. . . . .                                                                                                                                                                                                                                                                                       | 56 |
| 4.14 | Sketch of the light yield measurement setup. A, B, C and D indicate the positions for the $^{90}\text{Sr}$ source. Only position A is used in nominal tests, while the others have been used for attenuation length measurements. . . . .                                                                                                                                                               | 58 |
| 4.15 | Light yield (left) and cluster size (right) as a function of the horizontal distance between the SiPM array and the fibre mat. These scans have been performed on a test mat without mirror, for this reason, the absolute values on the $y$ axes are not representative of the nominal ones observed in production mats. . . . .                                                                       | 59 |

## List of Figures

---

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.16 | Light yield (left) and cluster size (right) as a function of the vertical alignment between the SiPM array and the fibre mat. These scans have been performed on a test mat without mirror, for this reason, the absolute values on the y axes are not representative of the nominal ones observed in production mats. . . . .                                                                                                                                                                            | 60 |
| 4.17 | Example of light yield test report, as saved in pdf format for each mat. The report shows, in black, the graph of the light yield per channel and per cluster as well as the cluster size. Each plot is compared to the analogous one from a reference mat, in red, to help immediately spotting unforeseen characteristics. In addition, for each of these graphs, the relative distribution is shown on the right. . . . .                                                                              | 61 |
| 4.18 | Attenuation length on three of the early mats produced at EPFL. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                   | 62 |
| 4.19 | Distribution of light yield of the scintillating fibre mats produced at EPFL (left) and scatter plot of fibre mat ID versus light yield (right). Data are reported both from tests performed with and without mirror. . . . .                                                                                                                                                                                                                                                                             | 62 |
| 4.20 | Distribution (left) and scatter plot (right) of the ratio between the light yield with and without the mirror. . . . .                                                                                                                                                                                                                                                                                                                                                                                    | 63 |
| 5.1  | Distribution of the BDTS classifier output after pre-selection for simulated signal and data sideband background. . . . .                                                                                                                                                                                                                                                                                                                                                                                 | 86 |
| 5.2  | Optimisation of PID requirement for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ with respect to $B_{(s)}^0 \rightarrow h^+ h^-$ mis-identification. The two sets of points correspond to the optimisation run with and without the ProbNNk requirement. The dashed lines represent the PID cut used in the 2013 analysis [46]. . . . .                                                                                                                                                                          | 87 |
| 5.3  | (left) ROC curve for the new BDT, in red, compared with the BDT used in Ref. [46], in black. The ROC curve is zoomed in the region of high rejection. (right) Relative importance for the BDT inputs. . . . .                                                                                                                                                                                                                                                                                             | 89 |
| 5.4  | (left) BDT response (before flattening) for the signal (blue) and background (red) samples used for training (points) and testing (filled histograms). The Kolmogorov-Smirnov test's p-value is overlaid. (right) BDT response after flattening compared between simulated signal and OS data sidebands. . . . .                                                                                                                                                                                          | 90 |
| 5.5  | BDT response vs mass distribution for data events in the (left) OS right mass sideband and (right) SS sample. The red points show the mean value of BDT for each mass bin. . . . .                                                                                                                                                                                                                                                                                                                        | 90 |
| 5.6  | Invariant mass distributions of $B^0 \rightarrow K^+ \pi^-$ candidates in Run I data in the most background-like (top) and the most signal-like (bottom) BDT bins with a PID requirement $ \Delta LL_{K-\pi}  < 5$ . The red solid line shows the $B^0$ signal, the green one shows the $B_s^0$ component, the yellow dashed one the one from $\Lambda_b^0 \rightarrow p h$ where the proton is misidentified as a kaon or pion. The combinatorial background is shown by the purple dashed line. . . . . | 94 |
| 5.7  | (Top) fraction of $B^0 \rightarrow K^+ \pi^-$ data in each BDT bin for different $\Delta LL_{K-\pi}$ cuts ( $\kappa = 5 \dots 10$ ). Each BDT bin is fitted with a constant, to check consistency along $\kappa$ . (bottom) For better visualisation the fractions are shifted by $0.25 \cdot (i - 1)$ . . . . .                                                                                                                                                                                          | 96 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.8  | BDT signal PDF for $B^0 \rightarrow K^+ \pi^-$ candidates in Run I data. The statistical uncertainty is included as a red band and total uncertainty as a grey band. The systematic uncertainties are discussed in Section 5.1.6. The distribution for $B^0 \rightarrow K^+ \pi^-$ simulated candidates (blue hatched) is superimposed. . . . .                                                                                                                                                                                                                                        | 97  |
| 5.9  | Invariant mass distributions of $B^0 \rightarrow K^+ \pi^-$ candidates from Run I data in the most background-like (top) and the most signal-like (bottom) BDT bins for $ \Delta LL_{K-\pi}  < \kappa$ cut, with $k = 5$ with the alternative PDF used for the evaluation of the systematic error. The red solid line shows the $B^0$ signal, the green one shows the $B_s^0$ component, the yellow dashed one the one from $\Lambda_b^0 \rightarrow ph$ where the proton is misidentified as a kaon or pion. The combinatorial background is shown by the purple dashed line. . . . . | 99  |
| 5.10 | Trigger efficiency ratio between $B_s^0 \rightarrow e^\pm \mu^\mp$ and $B^0 \rightarrow K^+ \pi^-$ for the different BDT bins. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 100 |
| 5.11 | Shape of the BDT response for candidates with and without bremsstrahlung photons recovered and for the whole sample, for signal. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                               | 101 |
| 5.12 | Invariant mass fits to $B_s^0 \rightarrow e^\pm \mu^\mp$ simulated candidates with HasBremAdded=0 (left) and HasBremAdded=1 (right). The blue, solid line is a Double-Sided Crystal Ball distribution with all parameters free in the fit. . . . .                                                                                                                                                                                                                                                                                                                                     | 101 |
| 5.13 | Dilepton invariant mass $m(J/\psi)$ fits for $B^+ \rightarrow J/\psi(e^+ e^-)K^+$ (left) and $B^+ \rightarrow J/\psi(\mu^+ \mu^-)K^+$ (right) simulated candidates (top) and 2012 data (bottom). The blue line is the total distribution, while the green line is a Double-Sided Crystal Ball distribution and the red line is an exponential component for the background. . . . .                                                                                                                                                                                                    | 103 |
| 5.14 | Invariant-mass distribution of $B^0 \rightarrow \pi e$ candidates with the fit superimposed. $m_{\mu\mu}$ indicates the $B$ invariant mass, calculated in the muon mass hypothesis for both child particles. . . . .                                                                                                                                                                                                                                                                                                                                                                   | 107 |
| 5.15 | Invariant-mass distributions of $\Lambda_b^0 \rightarrow p\mu\nu$ (left) and $B^0 \rightarrow \pi\mu\nu$ (right) candidates passing the $B_s^0 \rightarrow e^\pm \mu^\mp$ full selection. The plots show separately candidates with one bremsstrahlung photon emitted by the particle reconstructed as an electron ( $1\gamma$ ) and no photons emitted ( $0\gamma$ ) together with their sum. . . . .                                                                                                                                                                                 | 109 |
| 5.16 | $\Lambda_b^0 \rightarrow p\mu\nu$ (left) and $B^0 \rightarrow \pi\mu\nu$ (right) simulated candidates passing the signal selection, in the BDT range $[0.4, 0.5]$ for events with HasBremAdded = 0 (top) and HasBremAdded = 1 (bottom). . . . .                                                                                                                                                                                                                                                                                                                                        | 110 |
| 5.17 | Limits on Lepton Flavour Violating decays, updated in August 2017. Figure from the Heavy Flavor Averaging Group Report [119]. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 118 |
| 6.1  | Scatter plot of the $\Lambda_b^0$ flight distance versus $\alpha_{\text{HOP}}$ for signal (blue) and combinatorial background (red). The black line represents the cut applied on signal. . . . .                                                                                                                                                                                                                                                                                                                                                                                      | 125 |
| 6.2  | Distribution of $p(\Lambda_b^0)$ (top), $p_T(\Lambda_b^0)$ (middle) and $\chi_{\text{ORIVX}}^2(\Lambda^0)$ (bottom) on $s$ -weighted data (black points) and on simulation before (red lines) and after (blue boxes) the re-weighting. The left plot describes Run I samples, while the right one describes Run II. . . . .                                                                                                                                                                                                                                                            | 127 |

|      |                                                                                                                                                                                                                                                                                                                                                               |     |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.3  | Distribution of $t(\Lambda_b^0)$ on s-weighted data (black points) and on simulation before (red lines) and after (blue boxes) the re-weighting. The left plot describes Run 1 samples, while the right one describes Run 2. . . . .                                                                                                                          | 128 |
| 6.4  | Receiver Operating Characteristics curves of the four multivariate classifiers considered for the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ analysis. . . . .                                                                                                                                                                                         | 129 |
| 6.5  | Response of the BDT used in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ analysis for the signal (blue) and background (red) samples used for training (points) and testing (filled histograms). The Kolmogorov-Smirnov test's p-value is overlaid. . . . .                                                                                          | 130 |
| 6.6  | Receiver Operating Characteristics curves for the five cross-validation folds of the BDT classifier used in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ analysis. The thickness of the lines represents the measurement uncertainty. The Area Under the Curve (AUC) is also indicated. . . . .                                                      | 131 |
| 6.7  | Response of the BDT classifier used in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ analysis, compared for HasBremAdded categories (top), LHC run (middle) and $\Lambda^0$ track type (bottom). In each plot, the sample is integrated over the other categories. . . . .                                                                            | 132 |
| 6.8  | Receiver Operating Characteristics curves of the BDT classifier used in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ analysis, for the eight categories in which the analysis is split. Each point in the curves is shown with its binomial errors on both axes. . . . .                                                                             | 133 |
| 6.9  | Running cut efficiency on $\text{BDT}^{fix-hop}$ evaluated on $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ data and simulated samples. The left plots correspond to LL $\Lambda^0$ tracks and the right plots to DD $\Lambda^0$ tracks. The top row corresponds to Run I events and the bottom row corresponds to Run II. . . . . | 134 |
| 6.10 | $\alpha_{\text{HOP}}$ distribution on $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow e^\pm e^\mp)$ data and MC for Run I +Run II candidates, in both track categories. . . . .                                                                                                                                                                        | 134 |
| 6.11 | Punzi figure of merit and signal selection efficiencies as a function of the cut on BDT for Run I events. The categories without (with) bremsstrahlung photons added are shown in the top (bottom) row, and those with LL (DD) tracks are shown in the left (right) column. The red dashed lines indicate the optimal cuts. . . . .                           | 136 |
| 6.12 | Punzi figure of merit and signal selection efficiencies as a function of the cut on BDT for Run I events. The categories without (with) bremsstrahlung photons added are shown in the top (bottom) row, and those with LL (DD) tracks are shown in the left (right) column. The red dashed lines indicate the optimal cuts. . . . .                           | 137 |
| 6.13 | Invariant mass of simulated $\Lambda_b^0 \rightarrow \Lambda_c^{0+} (\rightarrow \Lambda^0 \mu \nu) e \nu$ candidates reconstructed as $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . . . . .                                                                                                                                                            | 138 |
| 6.14 | Mass preselection for $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow e^\pm e^\mp)$ channel displayed on 2016 data. The different colours correspond to the different cuts shown in Figure a. . . . .                                                                                                                                                  | 140 |
| 6.15 | Histogram of $m(p\pi ee)$ - $m(ee)$ for 2016 data candidates. . . . .                                                                                                                                                                                                                                                                                         | 141 |
| 6.16 | Fit to $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow e^\pm e^\mp)$ simulation for LL (top) and DD (bottom) candidates and for the $0\gamma$ (left) and $1\gamma$ (right) categories. . . . .                                                                                                                                                         | 142 |
| 6.17 | Fit on $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow e^\pm e^\mp)$ Run I and Run II simulation for LL (top) and DD (bottom) candidates and for the $0\gamma$ (left) and $1\gamma$ (right) categories. . . . .                                                                                                                                        | 143 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.18 Fit to $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ Run I and Run II data LL (left) and DD (right) candidates. . . . .                                                                                                                                                                                                                                                                                                                                                                                              | 144 |
| 6.19 Fit to $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ simulated candidates. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 145 |
| 6.20 Fits to the invariant mass of $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ simulated candidates built from LL (left) and DD (right) tracks. . . . .                                                                                                                                                                                                                                                                                                                                                                 | 146 |
| 6.21 Fits to the invariant mass of $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ data candidates built from LL (left) and DD (right) tracks; Run I (top) and Run II (bottom). . . . .                                                                                                                                                                                                                                                                                                                                     | 147 |
| 6.22 Blind mass fits to the invariant mass of $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ data. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                       | 153 |
| 6.23 CLs limit scan on the branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . This limit is obtained using the average branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ from the PDF [2].                                                                                                                                                                                                                                                                                       | 154 |
| 6.24 Relative uncertainties on nuisance parameters and global observables entering the computation of the branching fraction upper limit on $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . . . . .                                                                                                                                                                                                                                                                                                                                             | 156 |
| 6.25 CLs limit scan on the branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . This limit is obtained using the average branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ from the PDF [2], and removing all the systematic constraints. . . . .                                                                                                                                                                                                                                  | 157 |
| 6.26 CLs limit scan on the ratio, $r$ , of the branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ over the branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ . . . . .                                                                                                                                                                                                                                                                                                             | 157 |
| A.1 Histograms of the energies of the two leptons in $\Lambda_b \rightarrow \Lambda^0 e \mu$ (figures A.1a and A.1b) and $\Lambda_b \rightarrow \Lambda^0 J/\psi(ee)$ (figures A.1c and A.1d). In both cases, on the left the energies are expressed in the frame of the center of mass of the two leptons, while on the right they are in the laboratory rest frame. On top of each histogram, the Pearson correlation factor is shown. . . . .                                                                                                    | 166 |
| B.1 Invariant-mass distributions of $B^0 \rightarrow K^+ \pi^-$ candidates in Run I data in different BDT bins with a PID requirement $ \Delta LL_{K-\pi}  < 5$ . . . . .                                                                                                                                                                                                                                                                                                                                                                           | 168 |
| B.2 Invariant-mass distributions of $B^0 \rightarrow K^+ \pi^-$ from Run I data in different BDT bins for $ \Delta LL_{K-\pi}  < \kappa$ cut value $\kappa = 5$ with the alternative PDF used for the evaluation of the systematic error. . . . .                                                                                                                                                                                                                                                                                                   | 169 |
| C.1 Distributions of 4 out of 12 variables used in the $B_s^0 \rightarrow e^\pm \mu^\mp$ BDT training for simulated signal (blue) and background from opposite-sign data sidebands (red). From left to right and from top to bottom: the proper $B_s^0$ lifetime, $t(B_{(s)}^0)$ ; the square root of the minimum impact parameter $\chi^2$ for the two tracks, $\sqrt{\min(\text{IP}\chi^2)}$ ; the impact parameter of the $B_{(s)}^0$ , $\text{IP}(B_{(s)}^0)$ ; the distance of closest approach between the two daughter tracks, DOCA. . . . . | 171 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| C.2 | Distributions of 8 out of 12 variables used in the $B_s^0 \rightarrow e^\pm \mu^\mp$ BDT training for simulated signal (blue) and background from opposite-sign data sidebands (red). From left to right and from top to bottom: the isolation of the two tracks $I(e\mu)$ ; the cosine of the angle between the muon momentum in the B rest frame and the vector perpendicular to the $B_s^0$ momentum and the beam axis, $\cos \theta_k$ ; the transverse momentum of the $B_s^0$ , $p_T(B_{(s)}^0)$ ; the isolation of the $B_s^0$ , $I(B_{(s)}^0)$ ; the maximum transverse momentum of the two daughter tracks, $p_{T,\max}$ ; the flight distance of the $B_s^0$ with respect to its primary vertex, $FD(B_{(s)}^0)$ ; the $\chi^2$ of the decay vertex of the $B_s^0$ , $\chi_{DV}^2$ ; the difference of pseudo-rapidity between the two tracks, $\Delta\eta$ . . . . . | 172 |
| E.1 | Distributions of 6 out of 11 variables used in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ BDT training for simulated signal events (blue) and background events from data sidebands (red). From left to right and from top to bottom: the distance of closest approach between the two lepton tracks, DOCA, the direction angle between the flight direction of the $\Lambda^0$ and the direction of its momentum, $DIRA(\Lambda^0)$ , the $\chi^2$ of the origin vertex of the $\Lambda^0$ , $\chi_{ORIVX}^2(\Lambda^0)$ , the impact parameter $\chi^2$ of the $\Lambda_b^0$ , $\chi_{IP}^2(\Lambda_b^0)$ , the sum of the impact parameter $\chi^2$ of the two leptons, $\chi_{IP}^2(leptons)$ , the sum of the transverse momentum asymmetries of the four final-state tracks, $\text{sum}(P_T \text{ asymmetry})$ . . . . .                                     | 176 |
| E.2 | Distributions of 5 out of 11 variables used in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ BDT training for simulated signal events (blue) and background events from data sidebands (red). From left to right and from top to bottom: the sum of the cone isolations of the four final-state tracks, $\text{sum}(\text{isolation})$ , the HOP factor, $\alpha_{\text{HOP}}$ , the transverse momentum of the $\Lambda_b^0$ factor, $p_T(\Lambda_b^0)$ , the difference of pseudo-rapidity between the two hadron tracks, $\Delta\eta(hadrons)$ , the $\chi^2$ of the distance between the decay vertex of the $\Lambda_b^0$ and the production vertex of the two leptons. . . . .                                                                                                                                                                                    | 177 |
| E.3 | Linear correlation factors between the BDT training variables in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ analysis in signal (left) and background (right). . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 178 |
| F.1 | Representation of the kinematic of a $B \rightarrow Y_h X_e$ decay, highlighting the quantities relevant for the definition of the variable HOP. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 179 |



# List of Tables

|      |                                                                                                                                                                                                                                                                                                                                   |     |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 2.1  | The forces experienced by different particles. Table adapted from [1]. . . . .                                                                                                                                                                                                                                                    | 5   |
| 4.1  | Light yield grades definition. . . . .                                                                                                                                                                                                                                                                                            | 56  |
| 5.1  | Trigger selection for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ . . . . .                                                                                                                                                                                                                                                             | 83  |
| 5.2  | Trigger selection for $B^+ \rightarrow J/\psi(\mu^+ \mu^-) K^+$ . . . . .                                                                                                                                                                                                                                                         | 83  |
| 5.3  | Selection for $B_s^0 \rightarrow e^\pm \mu^\mp$ . . . . .                                                                                                                                                                                                                                                                         | 84  |
| 5.4  | Selection for $B^0 \rightarrow K^+ \pi^-$ . . . . .                                                                                                                                                                                                                                                                               | 84  |
| 5.5  | Selection for $B^+ \rightarrow J/\psi K^+$ . . . . .                                                                                                                                                                                                                                                                              | 85  |
| 5.6  | PID selection for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ . . . . .                                                                                                                                                                                                                                                                 | 86  |
| 5.7  | Fractions of signal events surviving a requirement on the flattened BDT response, $\text{BDT}_{flat}^{sig} > X$ , for simulated signal and combinatorial background (data) compared between the new and old [46] BDT. . . . .                                                                                                     | 91  |
| 5.8  | Yields from the $B^0 \rightarrow K^+ \pi^-$ fit and PID efficiency corrections per BDT bin at $\kappa = 5$ . The first bin contains the full BDT range, to which the yield of all the others bins is later subtracted. . . . .                                                                                                    | 93  |
| 5.9  | Fractions of $B^0 \rightarrow K^+ \pi^-$ candidates falling into each BDT bin for data and simulated signal. The error quoted for data is the total (stat. and syst.), while for the MC it is just the statistical. . . . .                                                                                                       | 98  |
| 5.10 | Fraction of HasBremAdded, $f_{brem}$ , in data and MC for $B^+ \rightarrow J/\psi(\rightarrow e^+ e^-) K$ . Both reweighted and unweighted MC is in agreement with data. . . . .                                                                                                                                                  | 104 |
| 5.11 | Mis-identification probabilities for candidates from various $B_{(s)}^0 \rightarrow h^+ h^-$ decays to be reconstructed as $B_s^0 \rightarrow e^\pm \mu^\mp$ . The probabilities for the $B_s^0$ decays are assumed to be the same, as the kinematics of particles from 2-body $B^0$ and $B_s^0$ decays are very similar. . . . . | 106 |
| 5.12 | Branching ratio, total efficiency and number of expected background events . .                                                                                                                                                                                                                                                    | 106 |
| 5.13 | $B \rightarrow hh'$ fraction $c_X$ , total efficiency and number of expected peaking background candidates. . . . .                                                                                                                                                                                                               | 108 |
| 5.14 | Branching ratio, total efficiency (including geometric) and number of expected partially reconstructed background events. . . . .                                                                                                                                                                                                 | 109 |
| 5.15 | Yield (N) of the $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+$ normalisation channel and corresponding selection efficiencies, for Run I data. Each efficiency is calculated given all the previous cuts. Uncertainties are statistical only. . . . .                                                                     | 112 |

## List of Tables

---

|      |                                                                                                                                                                                                                                     |     |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.16 | Yield (N) of the $B^0 \rightarrow K^+ \pi^-$ normalisation channel and corresponding selection efficiencies, for Run I data. Each efficiency is calculated given all the previous cuts. Uncertainties are statistical only. . . . . | 112 |
| 5.17 | Total selection efficiencies for signal and normalisation channels. . . . .                                                                                                                                                         | 114 |
| 5.18 | Input for normalisation: branching fractions and fragmentation factor. . . . .                                                                                                                                                      | 114 |
| 5.19 | $B_s^0 \rightarrow e^\pm \mu^\mp$ and $B^0 \rightarrow e^\pm \mu^\mp$ single-event sensitivities. . . . .                                                                                                                           | 114 |
| 5.20 | $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ invariant-mass fit results. . . . .                                                                                                                                                           | 115 |
| 5.21 | Expected and observed upper limits for $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$ at 90%(95%) CL in light and heavy mass-eigenstate hypotheses. . . . .                                                                         | 115 |
| 5.22 | Expected and observed upper limits for $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp)$ at 90%(95%) CL. . . . .                                                                                                                         | 115 |
| 6.1  | Luminosity recorded at the LHCb experiment and centre-of-mass energy of the $pp$ collisions in each year included in the $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ analysis data sample. .                                  | 120 |
| 6.2  | Simulated samples and their number of produced events, per year. . . . .                                                                                                                                                            | 121 |
| 6.3  | Trigger selection for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . . . . .                                                                                                                                                   | 122 |
| 6.4  | Requirements of the Bu2LLK_meLine and Bu2LLK_mmLine. When a cut differs between DD and LL candidates, the cut for DD is reported in parentheses. . . .                                                                              | 123 |
| 6.5  | Offline pre-selection cuts applied on $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ with LL and DD hadron tracks.                                                                                                               | 124 |
| 6.6  | Offline pre-selection cuts applied on $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ with LL and DD hadron tracks. . . . .                                                                                | 124 |
| 6.7  | Double-semileptonic background to $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ and their branching fractions.                                                                                                                  | 135 |
| 6.8  | Results of $\mathcal{C}$ factor evaluation. . . . .                                                                                                                                                                                 | 144 |
| 6.9  | Normalisation channel, $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ , yields in the considered categories.                                                                                              | 146 |
| 6.10 | Geometrical acceptance efficiencies for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ and $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ .                                                            | 148 |
| 6.11 | Trigger efficiencies for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ and $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ . . . . .                                                                   | 148 |
| 6.12 | Reconstruction and Stripping efficiencies. . . . .                                                                                                                                                                                  | 149 |
| 6.13 | Offline selection efficiencies from $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ MC with LL tracks. . . . .                                                                                                                    | 149 |
| 6.14 | Offline selection efficiencies from $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ MC with DD tracks. . . . .                                                                                                                    | 150 |
| 6.15 | Offline selection efficiencies from $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ MC with LL tracks. . .                                                                                                 | 150 |
| 6.16 | Offline selection efficiencies from $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ MC with DD tracks. . .                                                                                                 | 150 |
| 6.17 | PID efficiencies for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ and $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ , divided per year and track type. . . . .                                      | 151 |
| 6.18 | Efficiencies of the cut to the BDT response on signal in the eight analysis categories.                                                                                                                                             | 151 |
| 6.19 | Total selection efficiencies for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ and $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ in each category. . . . .                                           | 152 |
| D.1  | Geometric, PID, trigger and reconstruction plus selection efficiencies. Only decays with MC events passing selection are shown. . . . .                                                                                             | 173 |

# 1 Introduction

Our current understanding of the elementary constituents of matter is condensed in the Standard Model (SM) of particle physics, a remarkably successful theory that accommodates the known particles and forces in a clear scheme and that has proven capable in the past to even predict the existence of yet unobserved particles.

Nevertheless, despite providing an apparently complete picture, the SM appears fine-tuned on experimental observations and does not explain some mis-balances in the current composition of matter.

Some aspects of the interactions and relations between different elementary particles are also a source of yet unanswered questions: quarks and leptons are organised in a similar scheme, with three replicas of a base *doublet*, but the interactions between these replicas seem to obey different conservation laws that do not arise spontaneously from the theory. Furthermore, the existence of these conservation laws is questioned by theories beyond the SM where new hypothetical mediators allow new kinds of processes.

The present thesis tries to investigate this open question by probing the interactions between quarks and leptons at high energy scales, through the analysis of data collected with the LHCb detector from proton-proton collisions at the LHC.

LHCb has been operating from 2011 to 2018 and will resume its activity in 2021 after a substantial upgrade, mostly involving its tracking system. Some details concerning the testing and quality assurance of scintillating fibre mats that will compose the new tracker – the SciFi – are also discussed in this thesis.



## 2 Theory

This section introduces the reader to the theory that constitutes the foundation of the research presented in this thesis. Only a general, non-exhaustive, overview of the most relevant topics is provided.

### 2.1 The Standard Model of Particle Physics

The nature of elementary particles and their interactions, i.e. the forces acting between them, is described by the Standard Model of particle physics, a gauge quantum field theory that unifies the particles and the forces in one single picture, where the latter are represented by the exchange of some specific particles.

In the SM, particles are classified according to their properties:

- fermions have half-integer spin and their wave-functions obey the Dirac equation;
- bosons have integer spin and their wave-functions obey the Klein-Gordon equation.

The building blocks of all matter are the elementary particles, *i.e.* those that do not have an internal structure, as opposed to composite particles.

According to the current experimental evidence, these particles are:

- the **twelve elementary fermions**:
  - the six quarks:  $u, d, c, s, t, b$  plus their respective anti-quarks;
  - the six leptons: electron ( $e$ ), muon ( $\mu$ ), tau ( $\tau$ ) and their respective neutrinos  $\nu_e, \nu_\mu, \nu_\tau$ , plus all their respective anti-leptons;
- the **five bosons**  $W^\pm, Z^0, \gamma$  (photon),  $g$  (gluon),  $H$  (Higgs boson)

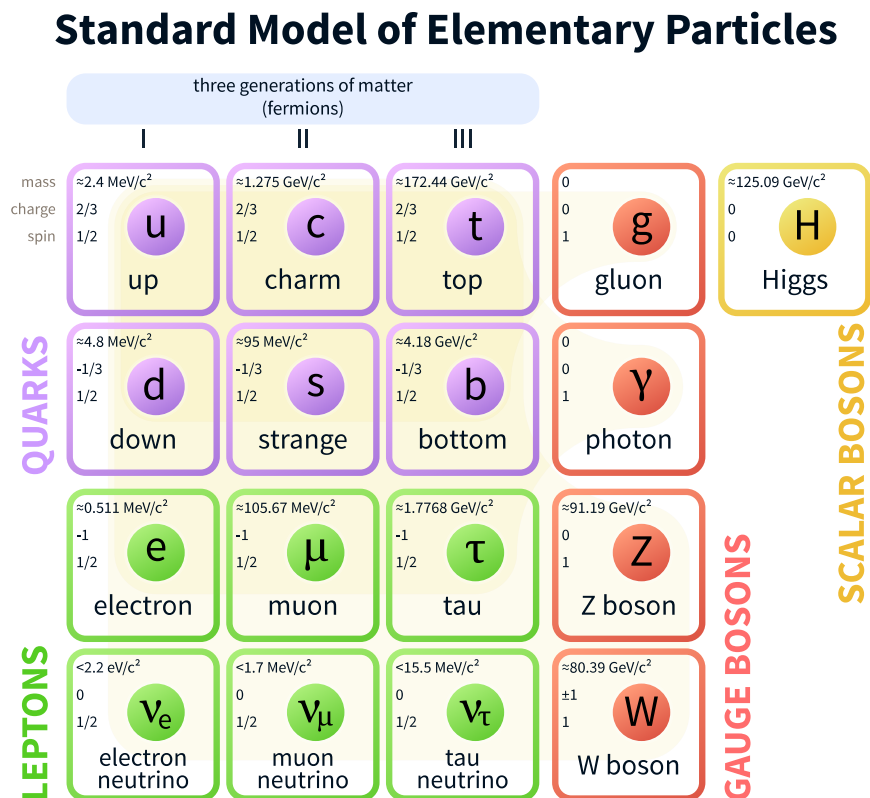


Figure 2.1 – Schematic representation of the fundamental particles of the Standard Model.

Figure 2.1 shows the known elementary particles of the Standard Model, with their quantum numbers.

The six different kinds of quarks and the six different kinds of leptons are referred to as *flavours*. The study of their origin and interactions is called **flavour physics** and it represents the specific domain of study of this thesis.

Unlike the other fermions, neutrinos are predicted to be massless in the SM. One consequence is that they should only be observed in their left-handed helicity state, while anti-neutrinos should be only right-handed. This point is important for the further discussion in Section 2.1.2.

2.1.1 Interactions

Table 2.1 summarises the effect of forces on the different elementary fermions.

The elementary bosons are responsible for the fundamental forces of the SM: these interactions happen indeed through the exchange of the respective mediator bosons, as described in the present section.

**Photons** mediate the electromagnetic interaction. They are massless, electrically neutral and

Table 2.1 – The forces experienced by different particles. Table adapted from [1].

|         |           |         |           |            | strong | e.m. | weak |
|---------|-----------|---------|-----------|------------|--------|------|------|
| Quarks  | up-type   | u       | c         | t          | ✓      | ✓    | ✓    |
|         | down-type | d       | s         | b          |        |      |      |
| Leptons | charged   | $e$     | $\mu$     | $\tau$     |        | ✓    | ✓    |
|         | neutrinos | $\nu_e$ | $\nu_\mu$ | $\nu_\tau$ |        |      |      |

have spin 1. Electrically charged particles interact with each other by the exchange of a virtual photon.

**The  $Z^0$  and  $W^\pm$**  mediate the weak interaction. The three gauge bosons have spin equal to 1. Unlike the photon, the electroweak bosons are massive: the masses of the  $Z^0$  and  $W^\pm$  are, respectively  $(91.1876 \pm 0.0021) \text{ GeV}/c$  and  $(80.358 \pm 0.015) \text{ GeV}/c$  [2]. The  $W^\pm$  is also electrically charged. The electromagnetic and weak forces are unified under the common picture of the electroweak force known as the Glashow-Weinberg-Salam Model [3–5]. The electroweak sector interacts under the Abelian symmetry group  $U(1) \times SU(2)$ .

**Gluons** mediate the strong interaction, responsible of the formation and transformations of hadrons. A total of 8 massless and electrically neutral gluons with spin 1 are present. They carry colour charge and are exchanged by coloured particles. Each flavour of quark exists in three colours (and three anti-colours), while the gluons carry a colour-anticolour charge, allowing them to interact with different quarks and with each other, but not with leptons, these latter being color-less. The theory that describes strong interactions is called Quantum Chromo Dynamic (QCD) and it is related to the non-Abelian symmetry group  $SU(3)$ .

Due to the *color confinement* [6, 7], quarks are not observed directly, but they form color-neutral composite particles called hadrons. These particles, as the elementary ones, can be classified in bosons and fermions with the same criterion, according to their spin. Hadrons composed of one quark and one anti-quark have integer spin and are called **mesons**, while hadrons composed of three quarks (or three anti-quarks) have fractional spin and are called **baryons**. Mesons are thus bosons while baryons are fermions.

Bound states with more than three quarks, referred to as *exotic* states are also possible, and have been observed [8].

The **Higgs boson** acts a different role: through the *Higgs Mechanism*, the  $W$  and  $Z$  bosons acquire mass. Below some very high temperature threshold, spontaneous symmetry breaking happens in the interaction between the Higgs field and the two bosons, causing them to have mass [9]. Furthermore, the Higgs coupling to fermions through a Yukawa term in the SM Lagrangian allows also the fundamental fermions to acquire a mass term.

The picture of forces is completed by gravity, which is not included in the SM as its effects at the sub-atomic scale are negligible, although numerous attempts of integrating it in extended theories are being performed, for example by suggesting the existence of an additional boson,

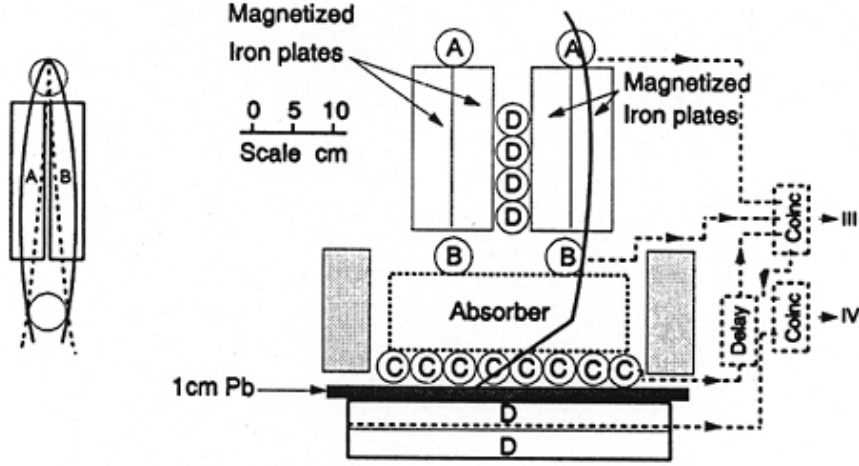


Figure 2.2 – A sketch of the experimental setup used by Conversi, Pancini and Piccioni. Figure from [11].

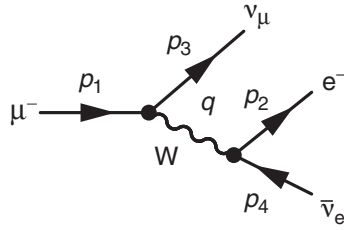


Figure 2.3 – The lowest-order Feynman diagram for Michel decay. Figure from [1].

the *graviton* [10].

### 2.1.2 Three generations

#### A historical introduction

In 1947, following the historical experiment performed by Conversi, Pancini and Piccioni [11] (see Figure 2.2), it appeared clear that the new particle discovered earlier, in 1937, by Neddermeyer and Anderson [12] could not be the mediator of the strong force predicted by Yukawa [13], as it was initially believed. This particle, initially named *mesotron*, and later  $\mu$  *meson* and *muon*, was indeed not captured in nuclei of light materials, but had instead properties very similar to those of the electron.

The idea that the muon could be a sort of replica of the electron appeared soon plausible. It was then necessary to investigate whether and how the muon can decay to an electron.

In the same year, neutral kaons were observed for the first time. As it was clear a few years later, they contain the *s* quark. There was then clearly a second generation of elementary fermions.



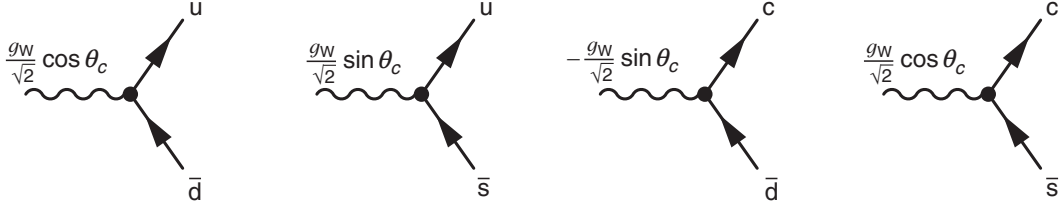


Figure 2.4 – The weak interaction couplings of the d, s, u and c in terms of the Cabibbo angle,  $\theta_C$ . Figure from [1].

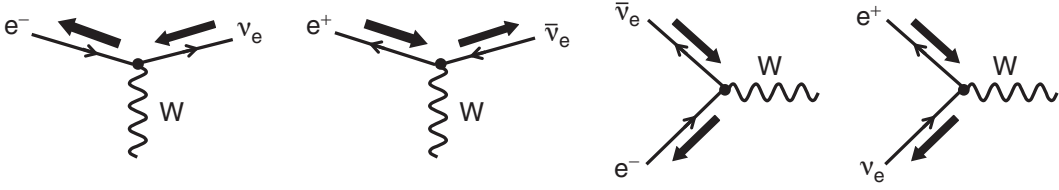


Figure 2.5 – The weak interaction couplings of the first generation of leptons, in their allowed helicity combinations, in the relativistic limit. Figure from [1].

The decay of the muon into an electron plus two neutral particles not compatible with being photons was already known in the 1950s, and five years later it was understood that the two neutral particles had a different nature. This decay is today known as *Michel decay*:  $\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e$ , see Figure 2.3. The observation of this decay, opposed to the non-observation of the long searched  $\mu^- \rightarrow e^- \gamma$  brought to the conclusion of the existence of the conserved lepton family numbers  $L_e$  and  $L_\mu$ , corresponding to the number of leptons of each family minus the number of anti-leptons of the same family. Conversely, in the hadronic sector, transitions between different generations [14] were observed in weak interactions.

The third generation was discovered much later. It took until 1995 to discover the  $t$  quark [15, 16] and until 2000 for the  $\nu_\tau$  [17]. The third conserved lepton family number  $L_\tau$  was then introduced.

### Today's picture

It is nowadays clear that the fundamental fermions in the SM are organised in three *generations*, as indicated in Figure 2.1. Each generation contains two quarks with opposite charge sign ( $+2/3$ ,  $-1/3$ ), and a couple of leptons, one charged ( $-1$ ) and one neutral. The three generations differ by the mass of these particles. The first is what composes the ordinary matter, while particles from the other generations are obtained in high-energy processes.

Despite this classification being in common between quarks and leptons, the way weak interactions behave with the three families is different for the two kinds of fermions.

In the lepton sector of the SM, the three family numbers are conserved. In the above-mentioned  $\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e$  decay, the destruction of the muon is balanced by the creation of a muonic neutrino to conserve  $L_\mu$ , and the creation of the electron is balanced by the creation of the electronic anti-neutrino, to conserve  $L_e$ . This phenomenon is known as *lepton flavour conservation*.

Figure 2.5 shows the possible couplings of the first generation of leptons to the  $W$  boson, in the allowed helicity configurations.

Nevertheless, the conservation of the lepton flavour is accidental in the SM, meaning that it is not related to the gauge structure of the theory, but it is rather a consequence of the absence of a mass term for neutrinos in the Lagrangian, in turn linked to the predicted absence of right-handed neutrinos. More specifically, it is possible to apply the same unitary transformation to the charged leptons and to neutrinos of the three generations, diagonalising the matrix of the Yukawa couplings for charged leptons to the Higgs field without introducing cross-generation terms in the Lagrangian of the interaction with the  $W$  boson, unlike in the case of quarks, where two different transformations are necessary to diagonalise up-type and down-type quarks, giving rise to the CKM matrix introduced later in this chapter.

However, this particular effect vanishes in minimal extensions of the SM, such as those introducing a second Higgs boson or, notably, those with a neutrino mass term. These latter are of particular interest, since the observation of neutrino oscillations [18–20] does imply that neutrinos are actually massive, besides being itself an observation of change of flavour in the lepton sector.

Furthermore, the oscillation of neutrinos indirectly allows the change of flavour in charged leptons to diagrams with loops involving the neutrinos, but this results in processes that are so rare (for example  $\mathcal{O}(10^{-54})$  for  $\mu \rightarrow e\gamma$ ) that they are far from experimental reach. Any clear observation of *Lepton Flavour Violation* (LFV) in charged leptons would thus be a sign of the existence of New Physics (NP) processes. The phenomenon of LFV is sometimes more precisely indicated as *charged lepton flavour violation* (CLFV) in the context of new searches, as in neutral leptons the violation is already observed.

In addition to the conservation of the lepton flavour, for years, experimental evidences such as the observed decay rates of muons and tau leptons or the couplings of the weak bosons to leptons [21] suggested that the strength of the weak interaction is the same for all lepton flavours, i.e. that the weak coupling is the same for the three lepton families. This law is known as *Lepton Flavour Universality* (LFU).

Conversely, in the quark sector, while the  $Z$  boson cannot alter the flavour of the particles it interacts with, thus forbidding flavour-changing neutral currents (FCNC) at the tree level<sup>1</sup>, the  $W$  boson can couple to different quark families in the same vertex. Transitions between the different generations are thus possible. Examples are shown in Figure 2.4.

Furthermore, the flavour universality is also not observed in the quark sector.

---

<sup>1</sup>The expression *tree level* refers to processes that can be described by Feynman diagrams with no loops.

This behaviour of quarks was originally explained by the Cabibbo hypothesis, in which the weak eigenstates of  $d$  and  $s$  quarks differ from the mass eigenstates, with respect to which they are rotated by an angle  $\theta_C$ . At the time of this formulation the charm quark and the third generation of quarks were not yet known, so the relation between the two different bases could be expressed as:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \quad (2.1)$$

The generalisation to the six quarks came with the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.2)$$

The CKM matrix can be parametrised with two rotation angles and a complex phase. Nevertheless, as it is nearly diagonal, it is convenient to represent it as a function of the parameter  $\lambda = \sin\theta_C = 0.225$  and the other three real parameters,  $A$ ,  $\rho$  and  $\eta$ . To  $\mathcal{O}(\Lambda^4)$  the CKM matrix then can be parameterised as

$$\begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}, \quad (2.3)$$

where the presence of the complex terms allows CP violation in the quark sector. The different magnitude of the elements of this matrix implies different probabilities of transitions between quarks. The norm of its elements are approximately:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \simeq \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix}, \quad (2.4)$$

which means that, for example, a  $u \rightarrow d$  transition is about  $(100)^2$  times more likely than  $t \rightarrow d$  and slightly more rare than  $t \rightarrow b$ . As mentioned above, the CKM matrix is indeed nearly diagonal, which implies that coupling between different generations are more rare than within the same, although being possible, unlike in the lepton sector. The same-family couplings are also not all equally likely by construction, explaining the absence of universality in the quark sector.

Hints of non-universality also in the lepton sector have been recently observed in B meson decays, although with too low statistical significances to draw solid conclusions, in measurements performed by the BaBar, Belle and LHCb experiments [22–26]. Updates of these studies are being performed in LHCb with the latest collision data to reduce these uncertainties.

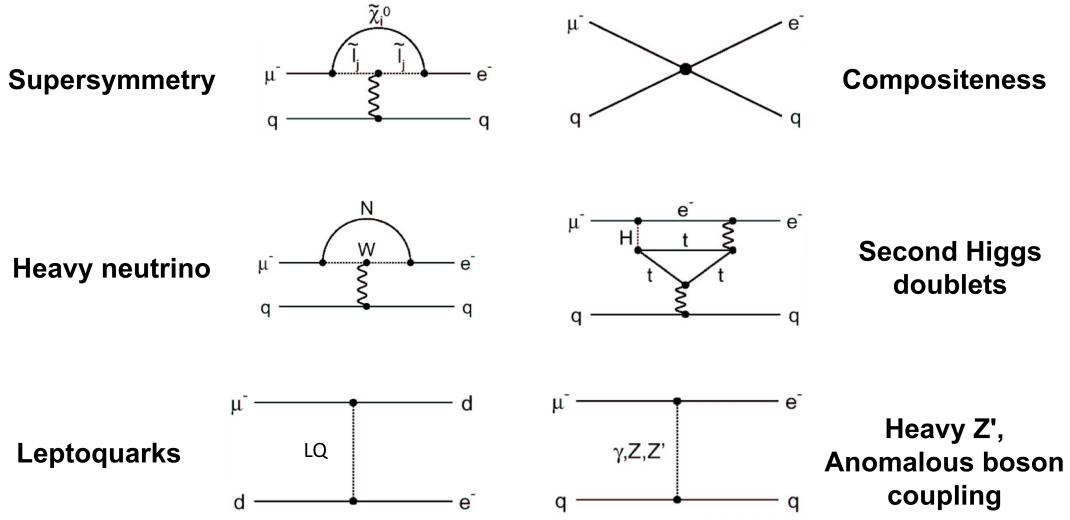


Figure 2.6 – Examples of Feynman diagrams of LFV processes in beyond-standard-model theories. Figure adapted from [27].

### 2.1.3 Lepton Flavour Violation: current theoretical scenario

The existence of lepton-flavour-violating processes is predicted by a large variety of theoretical models alternative to the SM. These include models with a new gauge  $Z'$  boson [28] or leptoquarks [29, 30], or models with heavy singlet Dirac neutrinos [31], supersymmetric models [32], the Pati-Salam model [33], models with composite leptons/quarks and models with additional Higgs doublets. Examples of Feynman diagrams of LFV processes in these beyond-standard-model (BSM) theories are shown in Figure 2.6.

Such models predict in some cases enhancements of multiple orders of magnitude for the branching fractions of LFV processes with respect to what allowed by the minimal extension of the SM with neutrino oscillations. In most cases, the predicted mass of the new mediators is very high, which explains why they have never been observed directly, although they can mediate processes at accessible energy scales off their mass shell.

The recent hints of violation of the lepton flavour universality mentioned in the previous section generated further interest in LFV phenomena, since the two effects could be strongly linked [34]. The existence of new LFU-violating mediators would indeed lead to the potential existence of a new basis in which leptons and quarks appear in these new interactions, different from the mass basis and from the electro-weak one. The transformation that rotates the elementary fermions from their mass base to this new base would generate LFV as a side-effect, as explained in Section 2.1.2.

General reviews about some of these hypothetical new particles can be found in [2].

### 2.1.4 Lepton Flavour Violation: current experimental scenario

The most stringent bounds on LFV processes to date are found in the muon sector. This is because the large muon lifetime ( $\sim 2.2 \times 10^{-6} \text{ s}$  [2]) allows for high-intensity muon beams and thus high statistical precision. Both the high precision and the relatively low muon mass, result in a low multiplicity of decay channels, and therefore are favourable for the sensitivity. The MEG experiment, the SINDRUM experiment and the SINDRUM II experiment have found the current best upper limits of  $5.7 \times 10^{-13}$ ,  $1.0 \times 10^{-12}$  and  $7 \times 10^{-13}$  on the branching fractions ( $\mathcal{B}$ ) of  $\mu^+ \rightarrow e^+ \gamma$ ,  $\mu^+ \rightarrow e^+ e^- e^+$  and  $\mu^- \text{Au} \rightarrow e^- \text{Au}$  [35–37] respectively. These limits will be further reduced in the near future by the follow-up experiments MEG II, Mu3e and Mu2e with expected sensitivities up to the order of  $10^{-17}$  [38–40].

LFV searches were also performed at LEP at the Z pole [41, 42] and at higher masses [43].

Another significant contribution to LFV searches is provided at the *B factories* by Belle and BaBar and at LHC by the LHCb experiment, which enable amongst other types of channels the study of  $\tau$  leptons and mesons decaying into asymmetrically flavoured lepton pairs. Apart from a few exceptions, typical limits on  $\tau$  and meson decays from both the B factories and the LHCb are of the order  $10^{-6} - 10^{-8}$ . Until now, LHCb has set the following upper limits:

- $\mathcal{B}(\tau^- \rightarrow \mu^- \mu^+ \mu^-) < 4.6 \times 10^{-8}$  [44],
- $\mathcal{B}(\tau^- \rightarrow \bar{p} \mu^+ \mu^-) < 3.3 \times 10^{-7}$  and  $\mathcal{B}(\tau^- \rightarrow p \mu^+ \mu^-) < 4.4 \times 10^{-7}$  [44],
- $\mathcal{B}(D^0 \rightarrow e^\pm \mu^\mp) < 1.3 \times 10^{-8}$  [45],
- $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 1.1 \times 10^{-8}$  and  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 2.8 \times 10^{-9}$  [46],
- $\mathcal{B}(B^+ \rightarrow K^- \mu^+ \mu^+) < 5.4 \times 10^{-8}$  and  $\mathcal{B}(B^+ \rightarrow \pi^- \mu^+ \mu^+) < 5.8 \times 10^{-8}$  [47],
- $\mathcal{B}(D^+ \rightarrow \pi^+ \mu^+ \mu^-) < 7.3 \times 10^{-8}$ ,  $\mathcal{B}(D_s^+ \rightarrow \pi^+ \mu^+ \mu^-) < 4.1 \times 10^{-7}$ ,  $\mathcal{B}(D^+ \rightarrow \pi^- \mu^+ \mu^+) < 2.2 \times 10^{-8}$  and  $\mathcal{B}(D_s^+ \rightarrow \pi^- \mu^+ \mu^+) < 1.2 \times 10^{-7}$  [48]

where some of these decays also violate other conservation laws such as those for the baryon number and lepton number. Future prospects in the LFV field are the Belle II upgrade for the Belle experiment [49], an upgraded accelerator complex at Fermilab giving opportunities for rare kaon decay searches [50] and further searches at the LHCb.

A more detailed introduction to the topic of lepton flavour violation, both from the theoretical and experimental point of view, can be found in [51].

## 2.2 Challenges of the Standard Model

The questions and the tensions introduced in the previous section are not the only indication of a potential incompleteness of the SM.

## Chapter 2. Theory

---

With the discovery of the Higgs boson at the LHC in 2012 [52, 53], the last fundamental particle predicted by the SM has been identified, completing the puzzle and providing a strong confirmation of the theory.

However, some questions remain totally or partially unanswered. Some of the most well-known critical points are:

1. as introduced in Section 2.1.2, the observation of neutrino oscillations suggests that neutrinos have a non-zero mass, contrarily to what assumed in the SM, although it is possible to extend the theory to accommodate these masses;
2. experimental evidences suggest the existence of dark matter and dark energy, which are not explained in the SM;
3. beyond-standard-model theories (supersymmetry, for example) are needed to explain the origin of *hierarchy problem*, i.e. the large difference between the strength of the electroweak force and of gravity.
4. the CP violation observed in the SM is insufficient to explain the observed imbalance of matter and anti-matter in the universe;
5. the SM contains a total of 18 free parameters which are tuned to experimental results: the 9 masses of the fermions, the 3 gauge couplings, the 4 parameters of the CKM matrix (3 angles and one phase), the mass of the  $Z^0$  and the one of the Higgs boson; resulting in the so-called *naturalness problem*, i.e. the idea that too many parameters are fine-tuned to data, with no existing explanation for their specific values.

These and many other observations hint for the existence of yet unobserved particles and forces, and strongly motivate research in this field.

## 3 The LHCb experiment

LHCb is one of the four major experiments at the Large Hadron Collider (LHC) [54]. It is specifically designed for performing precise measurements in the heavy-flavour sector, aiming to investigate new physics phenomena in CP violation and rare decays of hadrons containing beauty and charm quarks.

The present chapter contains a brief description of the experimental apparatus used to collect the data analysed in this thesis. After a quick general overview of the accelerator and the detector, the main sub-detectors and their performance are discussed.

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest and most powerful particle accelerator ever built. It is located at the European Organization for Nuclear Research (Conseil Européen pour la Recherche Nucléaire - CERN), on the border between France and Switzerland.

The accelerator consists of a 27-km-long double-ring synchrotron built inside the old LEP [55] (Large Electron-Positron Collider) tunnel, about 100 meters underground.

In the two rings, protons travel in opposite directions and collide in eight interaction points, four of which correspond to the positions of the four major particle detectors - ATLAS [56], CMS [57], ALICE [58] and LHCb. The latter will be presented more in detail in the next sections. The other three experiments at LHC - LHCf [59], TOTEM [60] and MoEDAL [61] - are located respectively near ATLAS, CMS and LHCb.

Before entering the LHC rings, protons are produced and accelerated in 4 steps: they are obtained ionizing hydrogen atoms and injected in bunches into the initial linear accelerator (LINAC2), which boosts them up to an energy of 50 MeV. They then enter the Proton Synchrotron Booster (BOOSTER) where they reach an energy of 26 GeV and finally, are accelerated in the Super Proton Synchrotron (SPS) to an energy of 450 GeV. A scheme is shown in Figure 3.1

By design, the number of protons in one bunch is  $1.15 \times 10^{11}$  and each beam contains up to

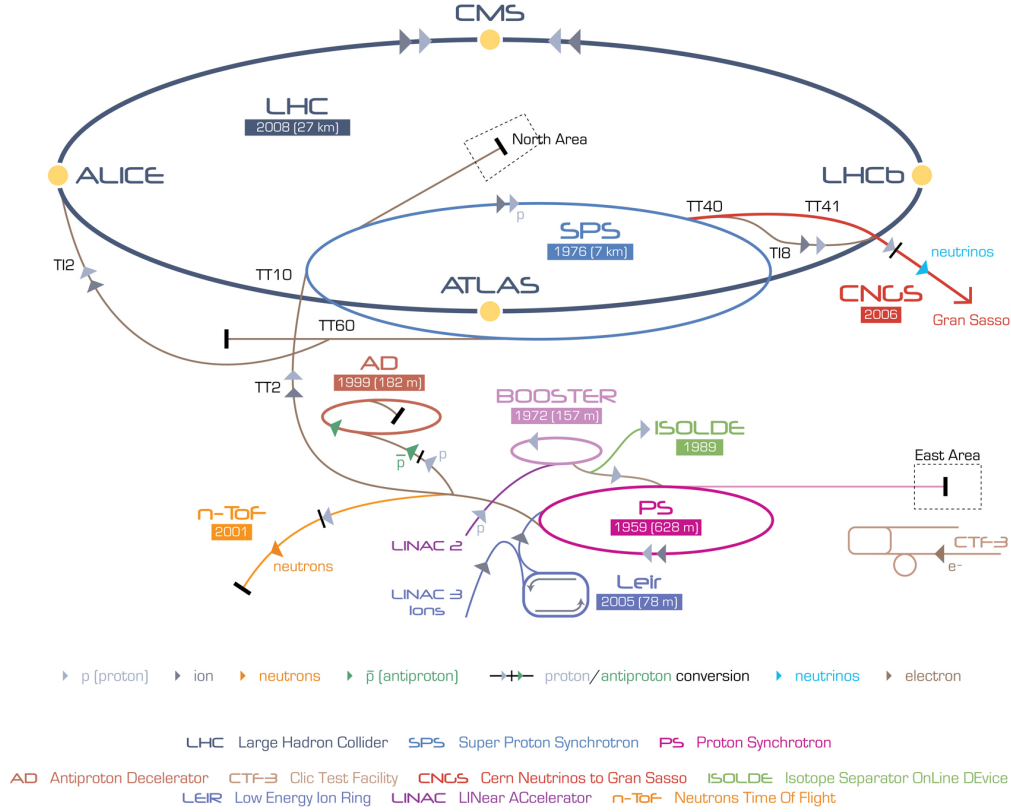


Figure 3.1 – The LHC Accelerator Complex. Figure from [62].

2808 bunches. The minimum time separation between proton bunches in the beams is 25 ns, corresponding to a bunch-crossing frequency of 40 MHz.

During its first period of activity, known as *Run I*, the LHC has been operating at a center-of-mass energy of 7 TeV in 2010 and 2011, and 8 TeV in 2012, while in the second period, *Run II*, from 2015 to 2018 the record energy of 13 TeV has been reached. The accelerator is also designed to collide heavy ions, mainly Pb and Ar: dedicated runs are performed each year for about one month to study ion-ion or ion-proton collisions. Only  $p-p$  collisions are analysed in this thesis.

An important parameter of the LHC is the luminosity,  $\mathcal{L}$ , which links the cross-section,  $\sigma$ , of a process to the rate of events produced for such process in a particle collider through the relation

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma. \quad (3.1)$$

The luminosity can be computed from the beam parameters as

$$\mathcal{L} = \frac{n_B N_p^2 f_{\text{rev}}}{4\pi\sigma_T^2} F \quad (3.2)$$



### 3.2. The Large Hadron Collider beauty experiment: LHCb

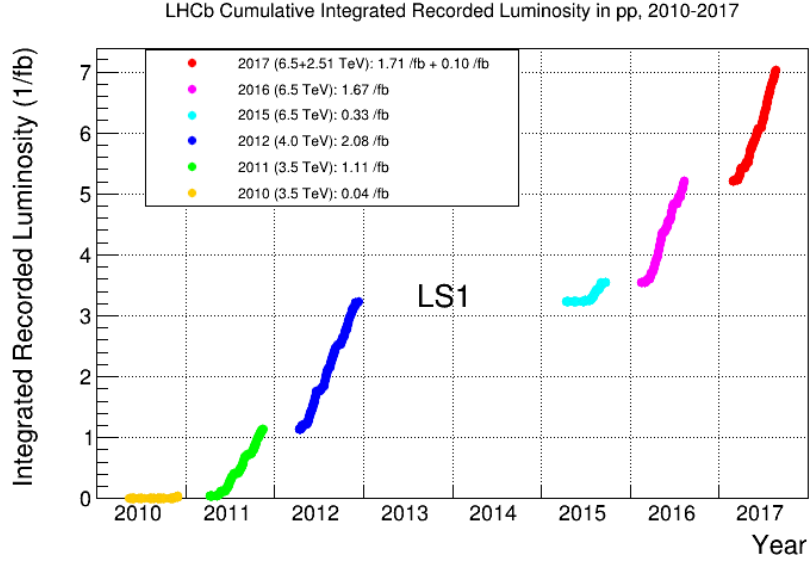


Figure 3.2 – Recorded luminosity at LHCb. Figure from [63].

where  $n_B$  is the number of bunches per beam,  $N_p$  is the number of particles per bunch,  $f_{rev}$  is the revolution frequency and  $\sigma_T$  is the transverse beam size at the interaction point.  $F$  is a parameter that quantifies the effect of the non-zero crossing angle at the interaction point.

Integrating Equation 3.1 over time defines the integrated luminosity  $\mathcal{L}_{int}$ , measured in units of inverse barn,  $b^{-1} = 10^{24} \text{cm}^{-2}$

$$\mathcal{L}_{int} = \int \mathcal{L} dt = \frac{N}{\sigma} \quad (3.3)$$

As explained further in Section 3.2, the instantaneous luminosity is reduced in the LHCb collision point to match specific requirements. Figure 3.2 shows the recorded luminosity in LHCb from 2010 to 2017.

### 3.2 The Large Hadron Collider beauty experiment: LHCb

LHCb, acronym of Large Hadron Collider beauty, is a single-arm forward spectrometer (see Figure 3.3).

This design is suggested by the production angle distribution of  $b\bar{b}$  pairs, which peaks in the forward and backward region as shown in Figure 3.4.

The geometrical acceptance of LHCb has an angular coverage of  $[10 - 250]$  mrad vertically (bending plane) and  $[10 - 300]$  mrad horizontally (non-bending plane), corresponding to the

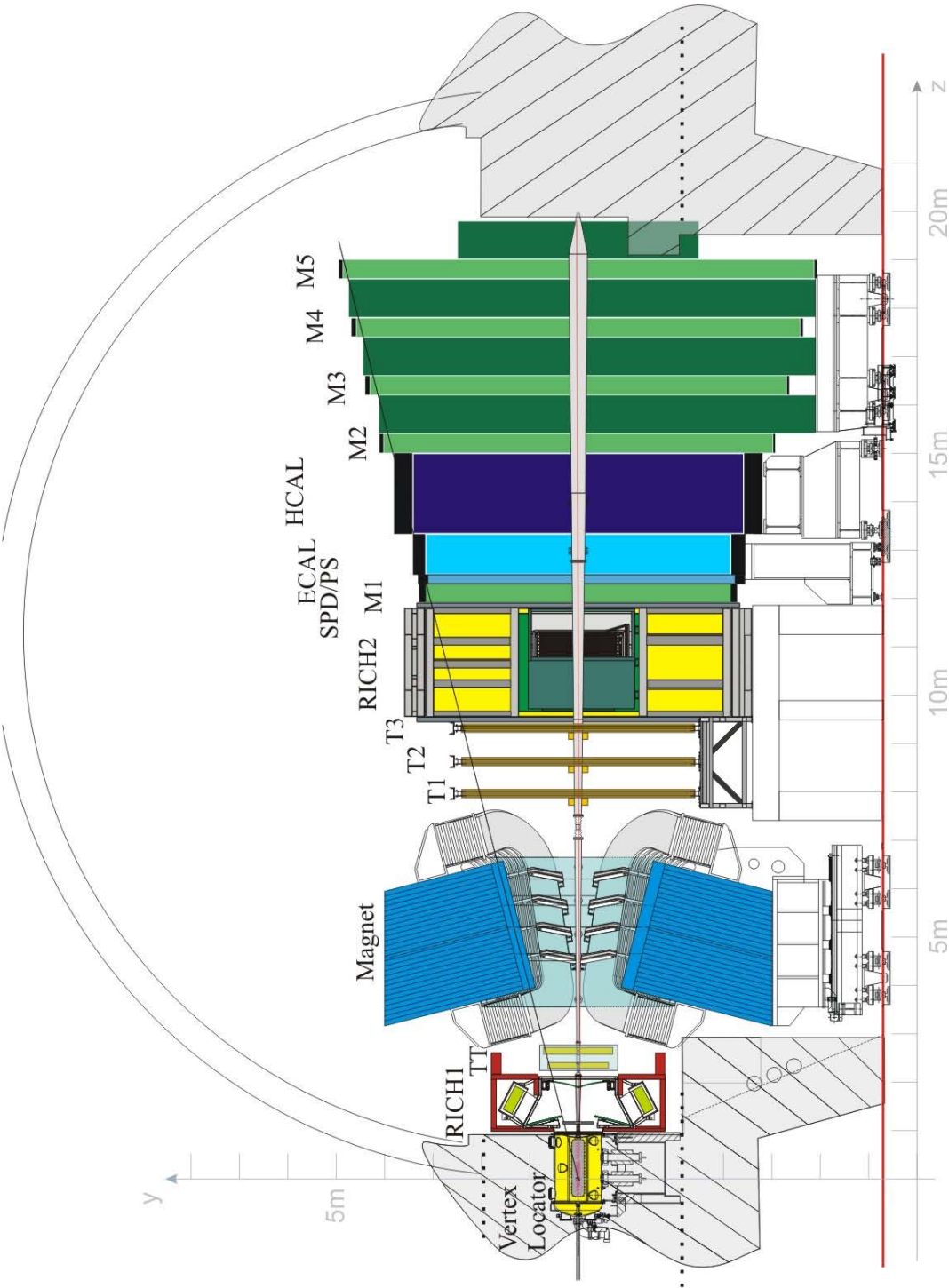


Figure 3.3 – The LHCb detector, with its sub-detectors, side view. Figure from [64].

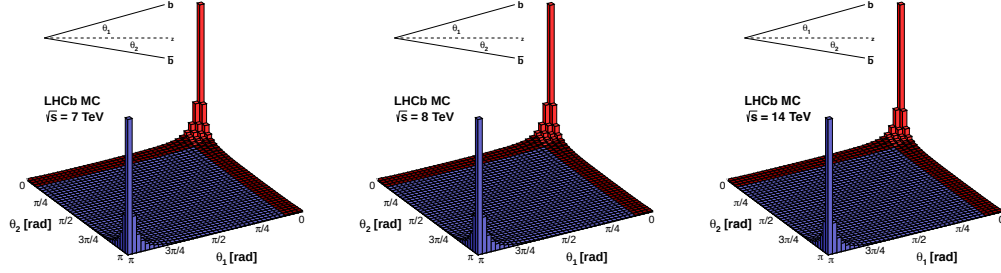


Figure 3.4 – Angular distribution of  $b\bar{b}$  quarks couples produced in  $pp$  collisions at LHC at a centre-of-mass energy of 7 (left), 8 (centre) and 13 (right) TeV.  $\theta_b$  and  $\theta_{\bar{b}}$  are the angles of the quarks momentum with respect to the beam axis. Figures from [65].

pseudo-rapidity region  $1.6 < \eta < 4.9$ .<sup>1</sup>

The production of  $b\bar{b}$  couples in  $pp$  collisions has a large cross-section (see Figure 3.5), and this makes LHCb a suitable detector for analyses of rare phenomena and precision studies. Furthermore, the high average momentum of the produced  $b$ - and  $c$ -quark mesons – about 80 GeV/ $c$  – implies that such particles travel on average for 1 cm before decaying, providing a clean and specific signature, observable at LHCb thanks to the precise vertex reconstruction.

The cross-section of production of  $b$  hadrons is in fact so large that a leveling of the instantaneous luminosity delivered by the LHC in the LHCb interaction point is made necessary in order to avoid too large event multiplicities and substantial radiation damage. The instantaneous luminosity is lowered by more than one order of magnitude with respect to the design value of  $10^{34} \text{cm}^{-2} \text{s}^{-1}$ , delivered to ATLAS and CMS. This is achieved by introducing an offset between the sections of the two colliding beams at the interaction point. This procedure is explained, together with other luminosity leveling techniques in Ref. [67]. This allows to obtain a sufficiently low average number of visible interactions per bunch crossing,  $\mu$ , and thus a low pileup<sup>2</sup>, around  $1 \div 2$ .

An example of the development of the LHCb instantaneous luminosity compared to ATLAS and CMS can be observed in Figure 3.6.

Figures 3.7 and 3.8 show respectively the average instantaneous luminosity and the average  $\mu$  over the LHC fill number in years 2012 and 2017, representing respectively Run I and Run II. Further general details on the performances of the LHCb detector are given in [64] and [68].

<sup>1</sup>Pseudo-rapidity is defined by  $\eta = -\ln(\tan(\theta/2))$ ,  $\theta$  being the angle between the particle momentum and the beam axis.

<sup>2</sup>The average number of  $pp$  interactions in visible events.

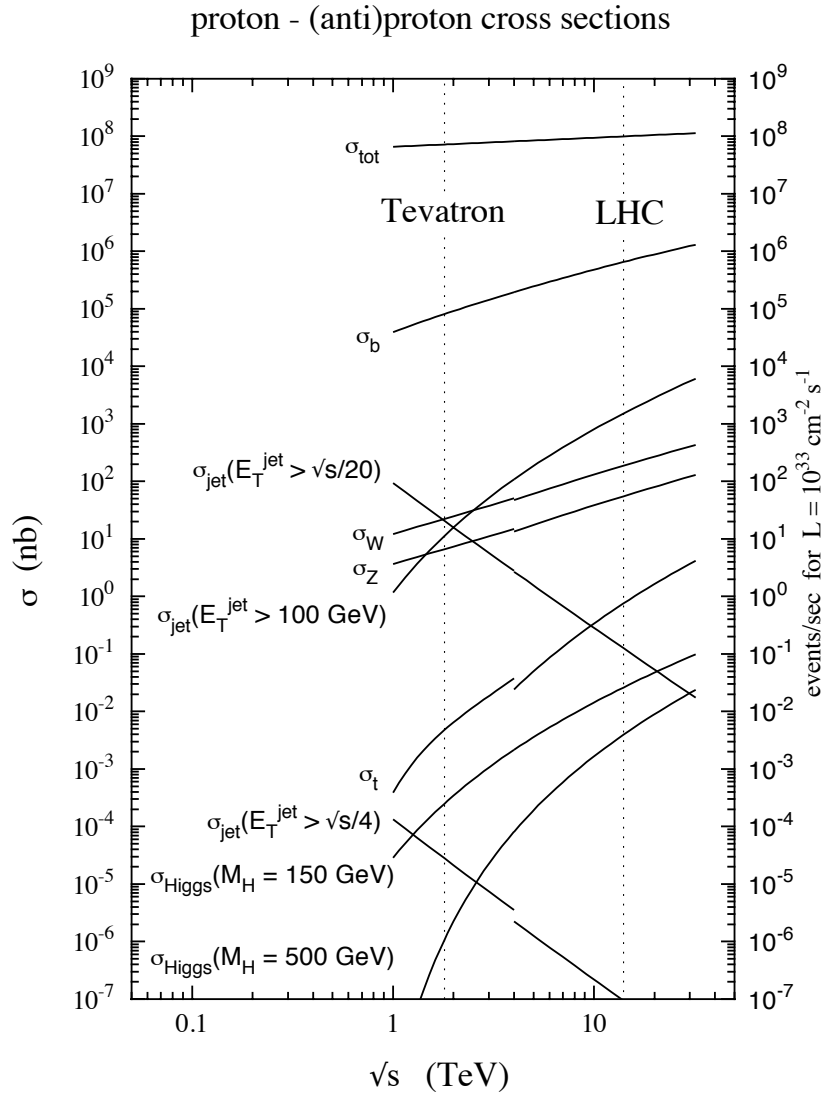


Figure 3.5 – Standard Model cross sections as a function of the centre-of-mass energy,  $\sqrt{s}$ , of the collider. The dashed lines corresponds to the Tevatron energy of 1.96 TeV and the nominal LHC energy of 14 TeV. The discontinuity in some of the cross sections at 4 TeV is due to the switch from proton-antiproton to proton-proton collisions at that energy. Figure from [66].

### 3.2. The Large Hadron Collider beauty experiment: LHCb

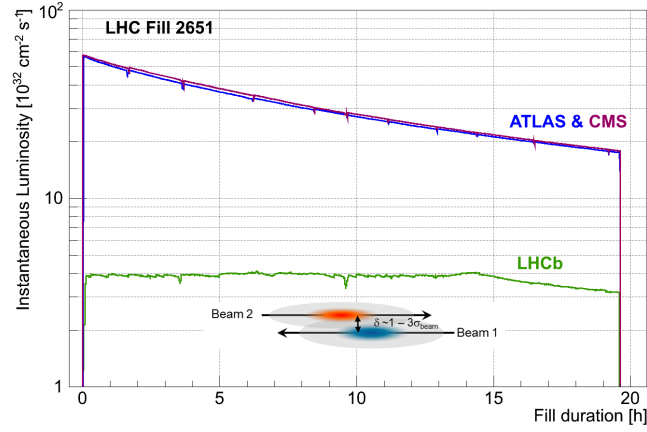


Figure 3.6 – Development of the instantaneous luminosity for LHCb, ATLAS and CMS during a typical LHC fill. After reaching the design value for LHCb, the luminosity is stabilised with a tolerance of 5% for about 15 hours by adjusting the transversal overlap of the colliding beams. Figure from [68].

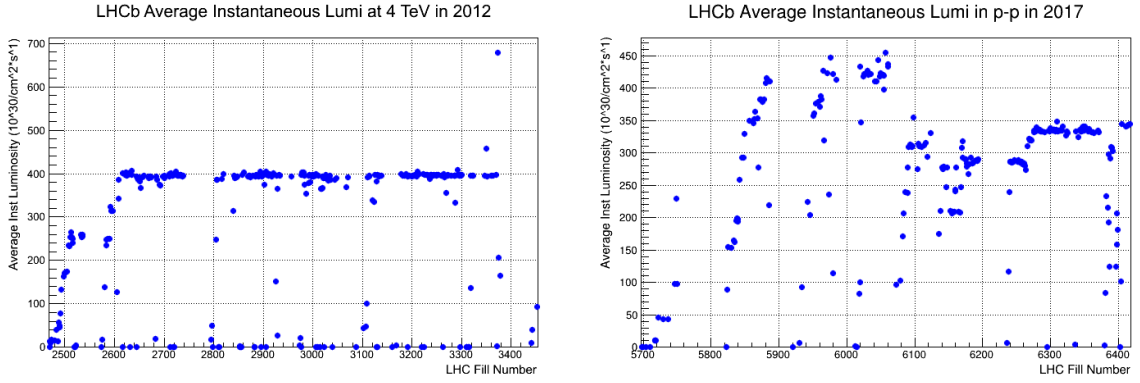


Figure 3.7 – Average instantaneous luminosity over LHC fill number in 2012 (left) and 2017 (right).

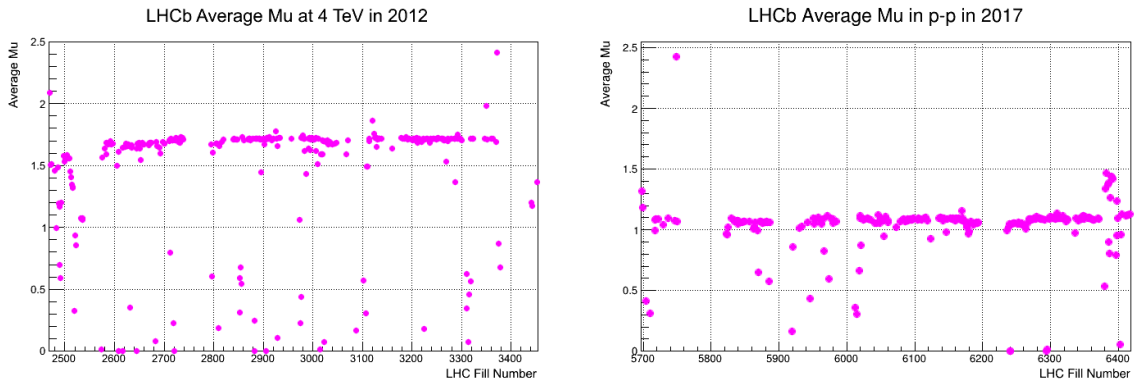


Figure 3.8 – Average number of visible interactions per bunch crossing  $\mu$  over LHC fill number in 2012 (left) and 2017 (right).

### 3.3 Tracking system

As already mentioned in the previous section, tracks from hadrons containing  $b$  and  $c$  quarks can be well identified by reconstructing their primary and secondary vertex (PV, SV), representing respectively the point in space where the hadron was created and where it decayed. This task is achieved by the VErteX LOcator (VELO).

Five types of particles are considered stable within the LHCb detector: electrons, muons, proton, kaons and pions. Their tracks are reconstructed in the tracking stations, that, in combination with the dipole magnet, also allows to evaluate their momentum, by measuring the bending experienced in the magnetic field, downstream of the VELO.

Track segments upstream of the dipole magnet are reconstructed in the Tracker Turicensis (TT), which also provides a preliminary momentum estimation and predicts the track position downstream of the magnet. Upstream track segments are then matched to downstream ones, provided by the three tracking stations ( $T$  stations) named T1, T2, T3, allowing for a precise measurement of track momenta with a resolution of  $\Delta p/p = 0.4\%$  at  $p = 5 \text{ GeV}/c$  to  $\Delta p/p = 0.6\%$  at  $p = 100 \text{ GeV}/c$  and a reconstruction efficiency up to 96% for tracks traversing the whole spectrometer.

The following sections provide further details on each section of the tracking system.

#### 3.3.1 VErteX LOcator

To the purpose of locating primary vertices (PV) and identifying the displaced secondary vertices (SV), the Vertex Locator (VELO) provides precise measurements of track coordinates close to the interaction region. It consists of a series of circular silicon strip modules arranged along the beam direction as shown in Figure 3.9.

The radial distance from the beam at which modules are placed is smaller than the minimum aperture required for the injection of proton bunches in the LHC, and therefore the VELO is designed to be retractable: each module is composed by two separate halves, that are only closed to form a circle (with a small overlap allowing relative alignment) during the *stable beams* phases of LHC.

Each module is composed of two parts:

- R-sensors, segmented in concentric semi-circles, thanks to which it is possible to measure the radial distance from the beam axis;
- $\Phi$ -sensors, segmented radially, for measuring the azimuthal angle.

The third coordinate, along the beam axis, is provided by the knowledge of the position of each sensor plane within the experiment.

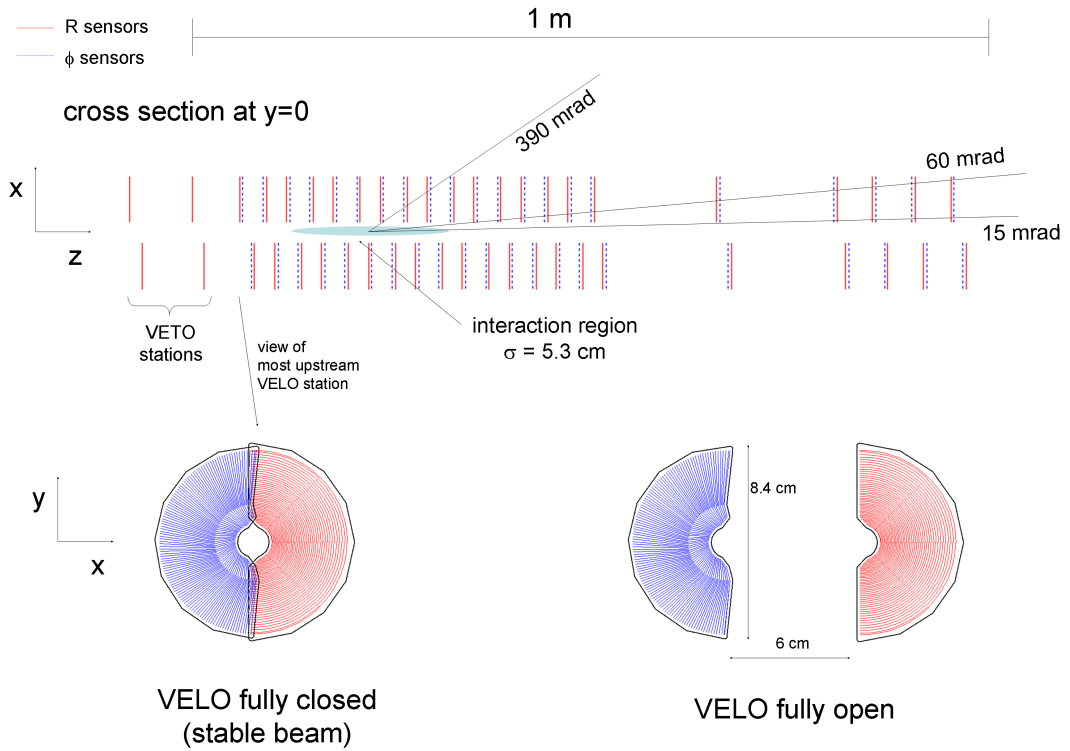


Figure 3.9 – Arrangement of the VELO silicon modules along the direction of the beam. The angles indicated by solid lines are: the crossing angle for minimum-bias events (60 mrad), minimum (15 mrad) and maximum (390 mrad) angle for which three or more stations are crossed. The front face of the first modules is illustrated, in the bottom part of the figure, in both its closed and open positions. Figure from [69].

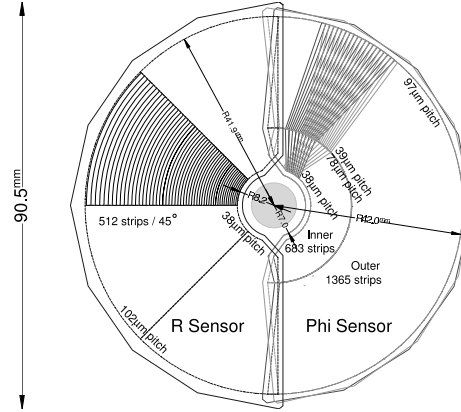


Figure 3.10 – A schematic view of the R- and  $\Phi$ - measuring sensors. On the  $\Phi$ -sensor, the strips on two adjacent modules are shown, to highlight the stereo angle. Although the physical radius of the R-sensors is slightly bigger due to practical constraints, the sensitive area is the same. Figure from [64].

Besides covering the entire forward acceptance of LHCb, the VELO does also partially cover the back hemisphere, in order to improve the identification of the primary vertex.

The sensors are  $300\mu\text{m}$ -thick, radiation tolerant, built using the n-implants in n-bulk technology. To reduce the strip occupancy, the strip pitch is not constant: in the R-modules it increases linearly from the inner to the outer edge, while the  $\Phi$ -sensors are divided in two parts, the outer of which has approximately twice the number of strips as the inner region. The strips in the inner and outer regions of the  $\Phi$ -sensors are skewed in opposite directions, to improve pattern recognition. Furthermore, adjacent  $\Phi$ -sensors have reversed skew with respect to each other, to obtain a stereo view that allows to better reject ghost tracks. A sketch of two opposite halves of the two types of modules is shown in Figure 3.10.

The VELO sensors are enclosed in a *secondary* vacuum vessel (as opposed to the *primary* (LHC) vacuum) made of a thin aluminum foil that also has the purpose of shielding the modules from the radio-frequency (RF) field of the LHC beams. This foil is thin ( $0.5\text{ mm}$ ) and it has a corrugated shape, in order to minimise the amount of material transversed by incoming particles and allow the superposition of the two halves of each module. Figure 3.11 illustrates this arrangement. More details of the mechanical design can be found in Ref. [69].

The VELO reaches excellent performances, with a signal-to-noise ratio better than 18 and a hit resolution that can reach  $4\mu\text{m}$ , depending on the angle of the track. For what concerns the resolution of the PV position measurement, it is mainly due to the number of tracks produced in a pp collision. For an average event it is  $42\mu\text{m}$  in the  $z$  direction and  $10\mu\text{m}$  perpendicular to the beam.



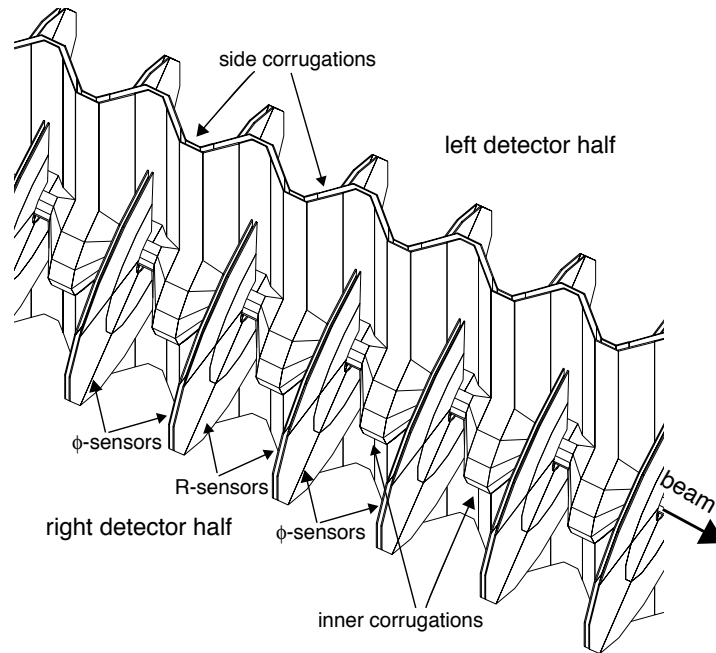


Figure 3.11 – View of the inside of the secondary vacuum container of the VELO. The corrugations close to the beam axis allow to minimise the material seen by incoming tracks, while those at the side allow the overlap between the two halves of each module. Figure from [69].

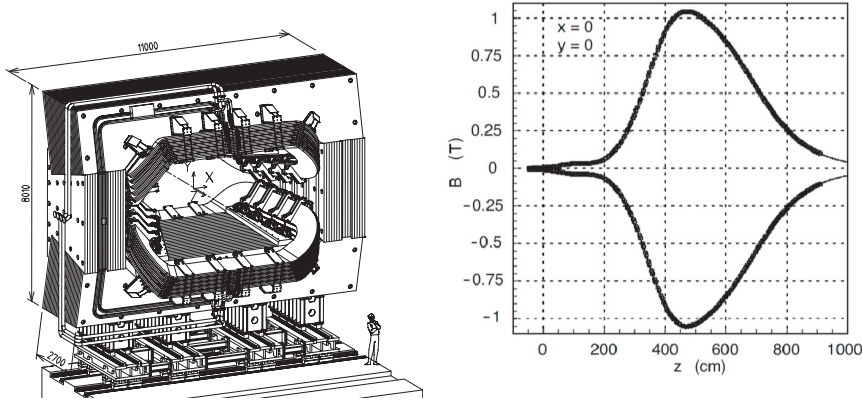


Figure 3.12 – Left: perspective view of the LHCb dipole magnet. Right: magnetic field intensity along the  $z$ -axis. Figures from [70].

### 3.3.2 The dipole magnet

A warm dipole magnet is used in LHCb to bend the tracks of charged particles, in order to be able to measure their momentum. It is located between the TT and the first tracking station T1, composed of two saddle-shaped coils placed mirror-symmetrically to each other in a window-frame yoke with sloping poles, to match the detector acceptance.

The magnet provides an integrated magnetic field of about 4 Tm for tracks of 10 m length (*i.e.* tracks passing through the entire tracking system). The main component of the  $\vec{B}$  field is along the  $y$ -axis and thus it bends charged tracks in the  $x$ - $z$  plane. The very precise (at the order of  $10^{-4}$ ) knowledge of this field along the detector acceptance allows for the tracking detectors to perform momentum measurements on charged particles with a precision of about 0.4% for low momenta and 0.6% up to 200 GeV/c. A sketch of the LHCb dipole magnet and the magnetic field intensity in the  $z$  direction is shown in Figure 3.12.

The magnet polarity is reversed frequently during data-taking in order to keep under control systematics due to left-right effects in the detector, which might bias precision measurements, potentially introducing an asymmetry in the detection and reconstruction efficiency of particle of opposite charge.

Further details on the design and performances of the magnet are given in [70].

### 3.3.3 Tracker Turicensis

Placed upstream of the magnet, the Tracker Turicensis (TT) consists of four layers of silicon strip sensors, subdivided in two stations (TTa, TTb) separated by 27 cm along the  $z$  direction, as shown in Figure 3.13. Being close to the magnet, in the TT a bending magnetic field of about 0.15 T is present; this allows to improve the momentum estimation for the charged particles. The layers are arranged in a  $x$ - $u$ - $v$ - $x$  configuration: the first and last one are perpendicular to

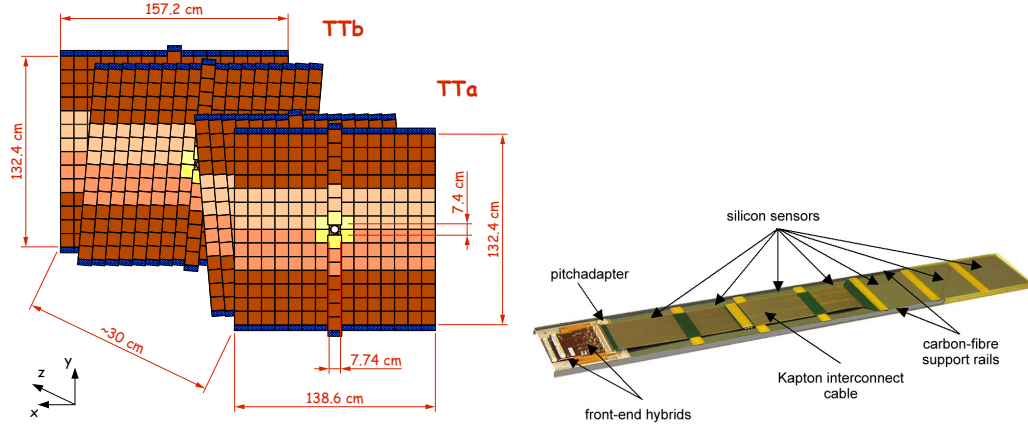


Figure 3.13 – Left: layout of TT layers. Right: structure of a half module. Figures from [71].

the  $x$ - $z$  plane, while the  $u$  and  $v$  layers are tilted by an angle of  $5^\circ$  in opposite directions around the  $z$  axis. This allows to resolve ambiguities that would occur with a simpler  $x$ - $y$  arrangement in case of multiple hits.

Each layer is composed of *half modules*, that cover half the height of the LHCb acceptance. They consist of a row of seven silicon sensors, made of 512 strips with a pitch size of  $183 \mu\text{m}$ , providing a hit position resolution of about  $50 \mu\text{m}$  in the bending plane.

The readout chips, the cooling system and the module supports are located above and below the active area of the detector, outside of its acceptance.

Further details on the TT can be found in [72].

#### 3.3.4 Inner Tracker (IT)

The inner tracker is placed in the central (closer to the beam) part of the three tracking stations T1, T2 and T3, downstream the magnet. The hit occupancy is higher in this area, with respect to the outer region of the stations, occupied by the OT: the IT covers less than 2% of the LHCb acceptance, but it is intercepted by about 20% of the tracks produced in pp collisions. A finer detector granularity is therefore required. Silicon microstrip sensors are employed to address this requirement.

Similarly to the TT, each of the three stations is composed by four layers, arranged in a  $x$ - $u$ - $v$ - $x$  configuration. The stations are also sub-divided in four electrically and thermally insulated detector boxes, arranged around the LHC beam pipe as shown in Figure 3.14, so that each layer is in fact split in four parts, one in each box together with the corresponding part of the other three layers. Each of these parts contains seven modules. The modules located to the sides of the beam pipe consist of two silicon sensors and a readout board, while those at the

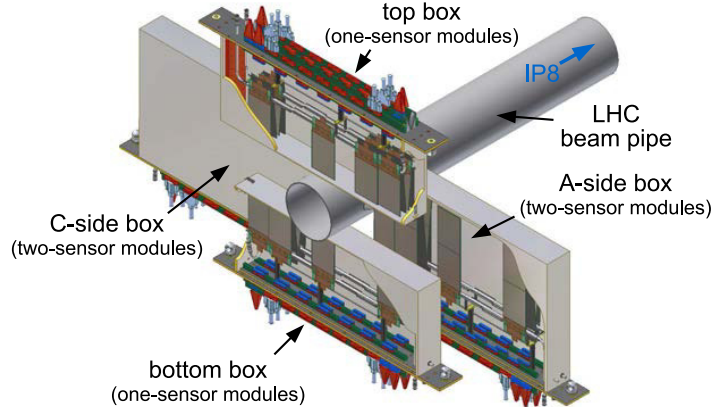


Figure 3.14 – View of the four IT detector boxes arranged around the LHC beampipe. Figure from [64].

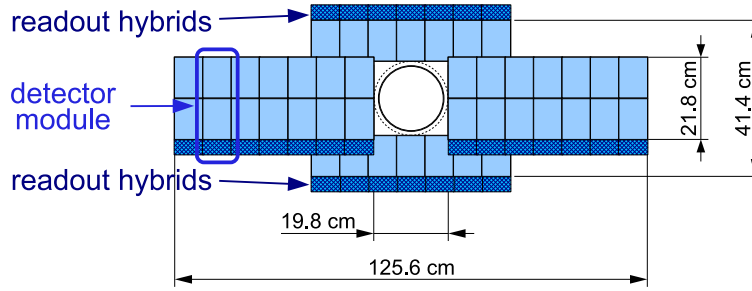


Figure 3.15 – Layout of an  $x$  detector layer in the second IT station. The lengths shown are in cm and they refer to the active area of the Inner Tracker. Figure from [64].

top and at the bottom only contain one detector each, plus a readout board. This allows to achieve precise measurements in the bending (horizontal) plane and a sufficient resolution for tracks reconstruction in the vertical plane. The arrangement of the silicon sensors in an  $x$  layer is sketched in Figure 3.15. Globally, the single-hit resolution is about  $50\text{ }\mu\text{m}$ . More details about the Inner Tracker design can be found in [73].

#### 3.3.5 Outer Tracker (OT)

The Outer Tracker is a drift time detector situated in the outer region of the three tracking stations T1-T3, surrounding the Inner Tracker and covering the remaining acceptance. It follows the same arrangement as the IT, having four layers in  $x$ - $u$ - $v$ - $x$  configuration per station. It is made of straw gas drift tubes with an inner diameter of 5.0 mm. Each layer of the Outer Tracker is made of 18 modules, arranged vertically and symmetrically around the beam pipe, each containing 128 tubes in turn arranged in two staggered *monolayers*. The single layers are

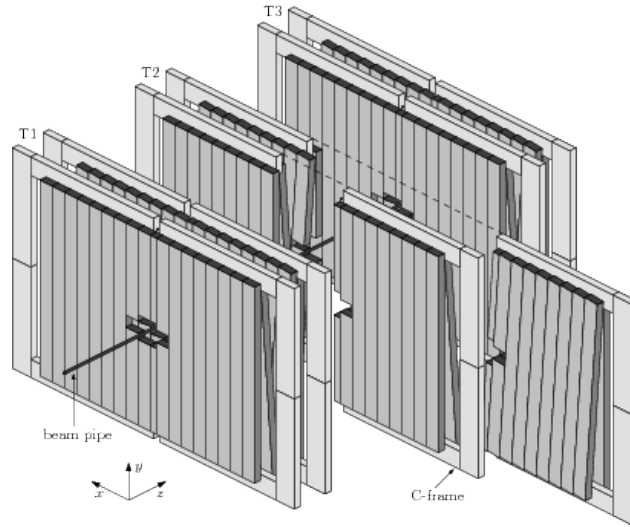


Figure 3.16 – Arrangement of OT straw-tube modules in layers and stations.

split longitudinally in an upper and a lower halves with respect to the plane  $y=0$ . The central modules are shorter to leave space for the IT in the middle. The position of the splitting in two sections is shifted between the two monolayers, in order to avoid insensitive regions in the centre of the module.

Each layer is mounted on two (left-right) supporting structures called *C-frames*. These structures are retractable, so that maintenance operations and easier access to the IT are made possible. Figure 3.16 shows a sketch of the OT modules and their structural elements.

The drift tubes are composed by an anode wire supplied by a high voltage potential surrounded by a cylindrical wall made of conductive material that collects the charge produced by the ionization of the gas induced by the transversing charged particles. A schematic section of an OT module and of a single tube can be seen in Figure 3.17. The filling gas is a combination of Argon (70%),  $\text{CO}_2$  (28.5%) and  $\text{O}_2$  (1.5%). The front-end electronics measures the drift time of the ionization clusters produced by charged particles transversing the straw tubes and ionizing the gas.

Thanks to a precise measurement of the drift time of the ionisation clusters produced by charged particles with respect to the bunch-crossing time, a position resolution of  $200\text{ }\mu\text{m}$  in the bending plane is achieved.

More detailed information about the Outer Tracker is available in [75].

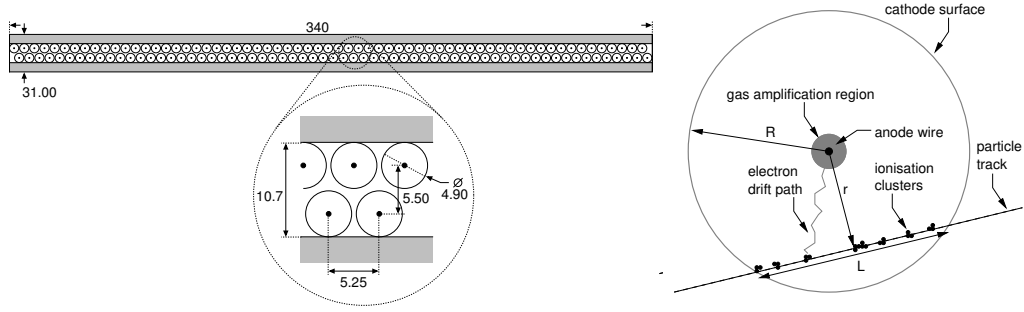


Figure 3.17 – Left: section of an OT module, showing the arrangement of the tubes. Right: section of a single tube. Figures from [74].

### 3.4 Particle Identification System

As mentioned in the previous sections, different types of charged particles ( $e$ ,  $\mu$ ,  $\pi$ ,  $K$ ,  $p$ ) transverse the detector, leaving distinctive signatures in different subsystems. An efficient and clean identification of such particles is a key requirement for distinguishing signal from background in physics analyses. The information collected from the two Ring Imaging CHerenkov (RICH) detectors, the calorimeters and the muon system collectively help to identify these charged particles, while neutral ones, such as photons and neutral pions, are identified using mainly information from the electromagnetic calorimeter.

#### 3.4.1 The RICH detectors

LHCb has two RICH detectors placed respectively upstream of the TT and downstream the three tracking stations T1-T3.

Charged particles are identified in the RICH using the Cherenkov light, that they emit when travelling through a dielectric medium (called *radiator*) at a speed faster than the speed of light in such medium. The Cherenkov effect is indeed the electromagnetic analogous of the *sonic boom* in acoustics: the charged particle would cause photon emission from the medium, only in a cone of half-opening angle  $\theta_c$ , which has its vertex in the particle itself and extends backwards. The angle  $\theta_c$  is linked to the refractive index of the medium,  $n$ , and to the ratio between the speed  $v$  of the particle and the speed of light,  $\beta = v/c$ , by the following formula:

$$\cos \theta_c = \frac{1}{n \cdot \beta}. \quad (3.4)$$

Knowing a particle's momentum (from the tracking detectors) and being able to obtain its speed through a measurement of  $\theta_c$ , it is then possible to determine its mass and thus the type of particle.

The choice of using two different RICH detectors arises from the need to be sensitive over the wide momentum spectrum for the particles in the LHCb acceptance. The upstream

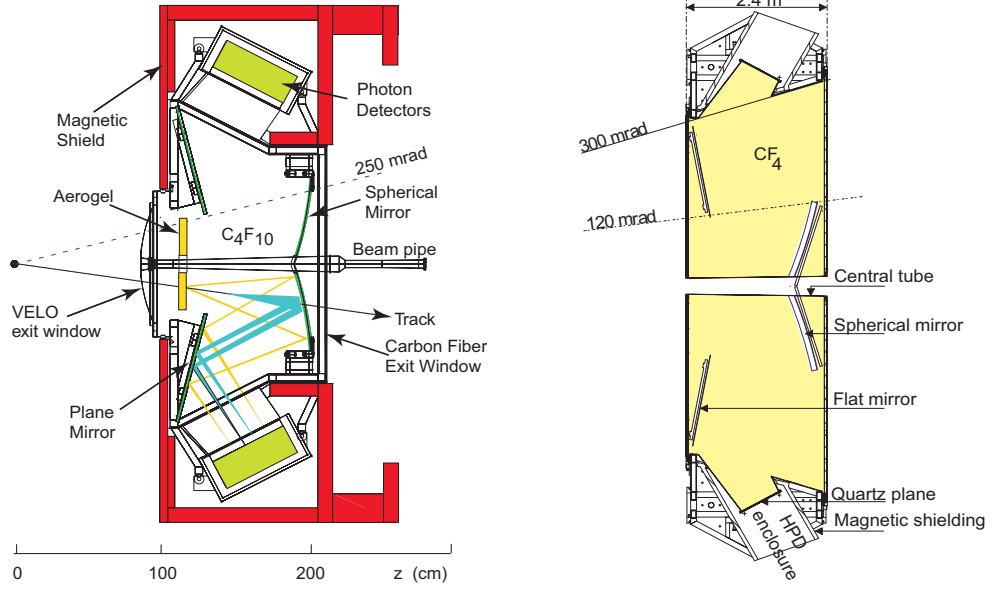


Figure 3.18 – Left: side view of the RICH 1. The aerogel has been removed in Run II. Right: side view of the RICH 2. Figures from [76].

sub-detector, RICH 1, performs best on the low-momentum ( $1\text{--}60\text{ GeV}/c$ ) particles, while the downstream sub-detector, RICH 2, covers the high momentum range, from  $\sim 15\text{ GeV}/c$  up to and beyond  $100\text{ GeV}/c$ . The acceptance of RICH 2 is smaller than the one of RICH 1; this is possible because the high-momentum tracks are concentrated at low polar angles.

A sketch of the two RICH detectors is shown in Figure 3.18.

In both RICH detectors the Cherenkov light is focused and steered out of acceptance by a system of mirrors which guide the photons (in the wavelength range  $200\text{--}600\text{ nm}$ ) towards a matrix of Hybrid Photon Detectors (HPDs), shielded against the magnetic field.

Figure 3.19 allows to appreciate the excellent separation power between different types of particles given by the two RICH detectors. The separation between kaons and pions, in particular, crucially relies on these detectors, given that the two particle's signature in other sub-systems is very similar.

Further details on the RICH can be found in [77].

#### 3.4.2 The calorimeter system

The calorimeter system fulfills multiple tasks. It measures the energy of hadrons, electrons and photons, it contributes to the identification of such particles and it provides very fast measurements of the transverse energy  $E_T$  to the hardware-level trigger. In order to keep a steady trigger rate, the calorimeter is self-calibrated using an embedded radioactive source during data-taking periods.

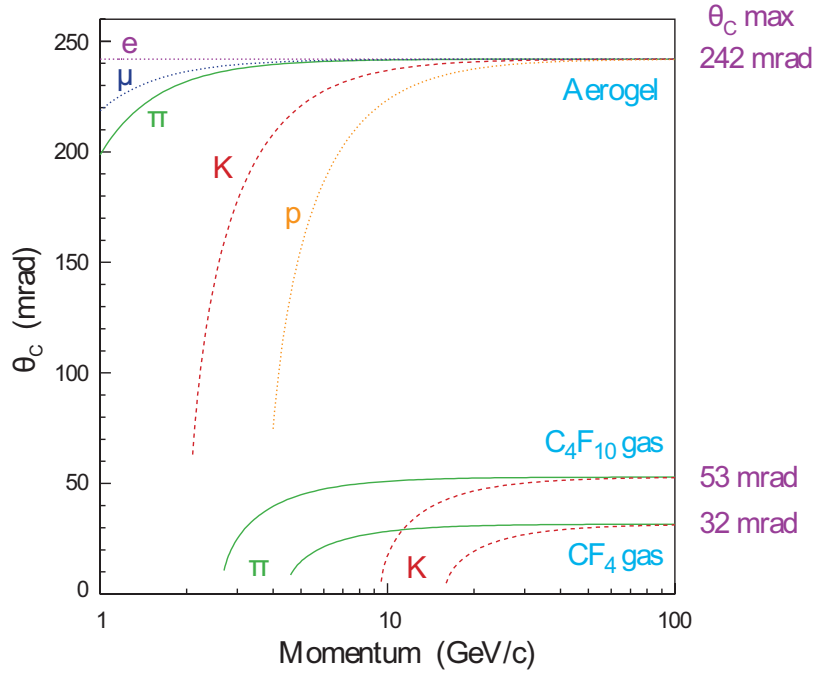


Figure 3.19 – Relationship between  $\theta_c$  and the momentum for different particles for the three media in the two RICH sub-detectors. Figure from [64].

The system is composed by four elements, in order of position along  $z$ :

- a Scintillating Pad Detector (SPD);
- a Preshower (PS);
- an electromagnetic calorimeter (ECAL);
- a hadronic calorimeter (HCAL).

The ECAL and the HCAL consist of alternating planes of scintillating and absorbing material (respectively lead and iron). The incident particles interact with the latter, creating a cascade of secondary particles that hit the scintillators, causing the emission of photons proportional in number to the energy of the incident particle. Photons are then transported through wavelength shifting fibres to photomultiplier tubes attached to the rear mechanical structure of each module. The thickness of ECAL was chosen to be 25 radiation lengths, while the one of the HCAL corresponds to 5.6 interaction lengths and it is limited by space constraints.

The PS and SPD are separated from each other by a 15 mm-thick lead layer, corresponding to 2.5 electromagnetic interaction lengths and to  $\sim 0.06$  hadronic interaction lengths. They are mainly used for initiating the electromagnetic shower from electrons and photons. The presence of the PS before the ECAL provides a longitudinal segmentation, necessary for



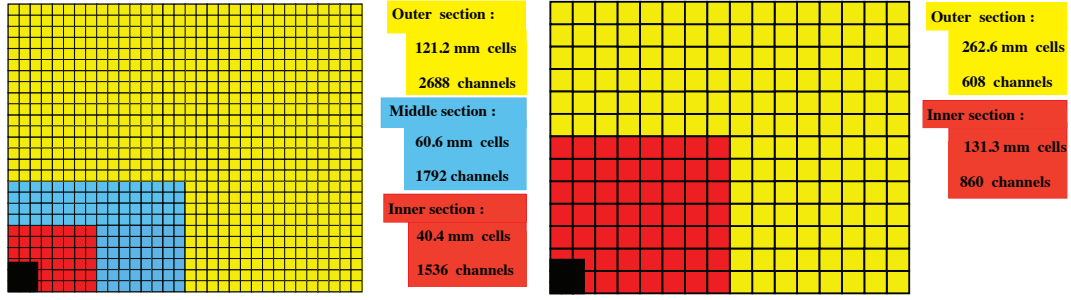


Figure 3.20 – Lateral segmentation of the SPD/PS and ECAL (left) and the HCAL (right). One quarter of the detector front face is shown. The cell dimensions reported in the left figure refer to the ECAL. Figures from [78]

discriminating charged pions, while the SPD allows to reject the background from high- $E_T$ <sup>3</sup> neutral pions.

The PS, SPD, ECAL and HCAL adopt a variable lateral segmentation, as shown in Figure 3.20. This choice is dictated by the fact that the hit density varies by two orders of magnitude over the calorimeter surface, depending on the radial distance from the beam. The PS, SPD and ECAL are segmented into three different regions, while the HCAL is divided in two zones with larger cell sizes, given the dimensions of the hadronic showers.

The energy resolution obtained from the ECAL is  $\frac{\sigma_E}{E} = 1\% + \frac{10\%}{\sqrt{E}}$ , while for the HCAL it is  $\frac{\sigma_E}{E} = 9\% + \frac{65\%}{\sqrt{E}}$ , where  $E$ , the energy of the particle, is expressed in GeV.

More information on the calorimeter system is available in [78].

#### 3.4.3 The muon system

The muon system of LHCb consists of five rectangular tracking stations. The first station (M1) is placed upstream of the PS, while the remaining four stations (M2, M3, M4 and M5) are located downstream of the calorimeter system, interleaved with iron absorbers to select penetrating muons.

Stations M1–M3 have a high spatial resolution in the bending plane, and are mainly used to find the direction of the tracks and to calculate the transverse momentum of the candidate muon, while stations M4 and M5 have a limited spatial resolution and are used for the identification of penetrating particles. The total thickness of the muon system and of the calorimeters correspond to 20 interaction lengths.

With an angular acceptance in the interval  $[20, 306]$  mrad in the bending plane and  $[16, 258]$  in the non-bendin plane, the muon system intercepts about 20% of the muons from inclusive  $b$  semileptonic decays.

<sup>3</sup>The transverse energy of an object is defined as  $E_T = \sqrt{m^2 + p_T^2}$ , where  $p_T$  is its momentum transverse to the beam, and  $m$  is its mass.

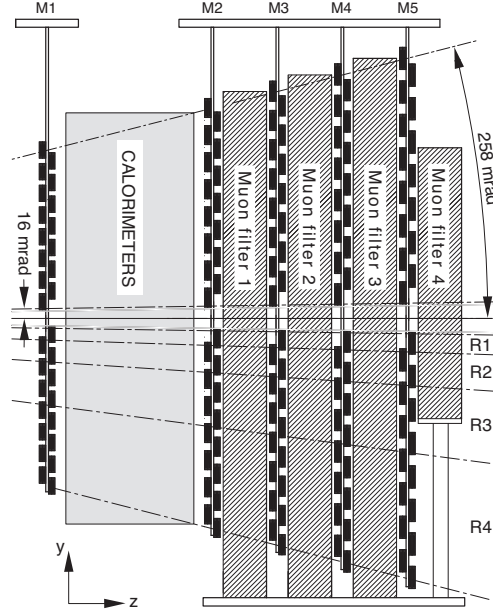


Figure 3.21 – Side view of the muon system in the  $y$ - $z$  plane. Figure from [79].

Each station is subdivided in four regions with dimensions and segmentation scaling as 1:2:4:8, in order to level the channel occupancy with respect to the distance from the beam. Figures 3.22 and 3.21 show this layout. In the muon stations multi wire proportional chambers are used, operating with a gas mixture of Ar (45%), CO<sub>2</sub> (15%) and CF<sub>4</sub> (40%), with the exception of the inner part of the most upstream station where GEM (Gas Electron Multiplier) chambers are employed, due to their higher resistance to the sizeable flux of particles.

Besides measuring the momenta of the muons, the muon system is used, like the calorimeter, to provide fast information on the transverse momentum,  $p_T$ , of the tracks to the hardware-lever trigger.

The muon system is described in detail in [79].

## 3.5 Data treatment

### 3.5.1 Particle Identification

Each type of particle leaves a characteristic *signature* in a detector, depending on the kind of interaction(s) with the materials that it transverses. A representation of how different kinds of particles behave in different subsystems of a generic HEP detector is provided in Figure 3.23.

Dedicated algorithms are used in LHCb to best exploit the information from the different sub-detectors for distinguishing between muons, protons, kaons, pions and electrons. They

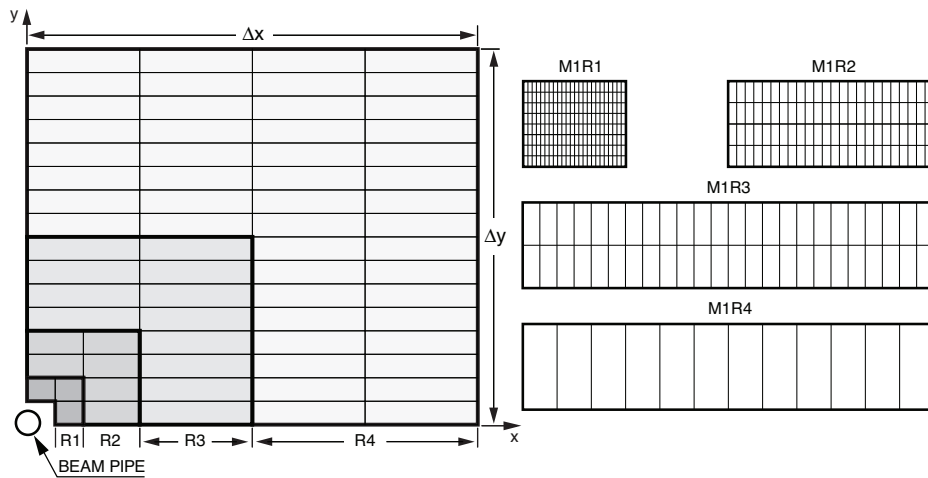


Figure 3.22 – Left: front view of a quadrant of a muon station. Each rectangle represents one chamber. Right: division into logical pads of four chambers belonging to the different regions of station M1. Stations M2-M3 (M4-M5) have twice (half) the number of pad columns per chamber with respect to M1, in each corresponding region, while the number of pad rows per chamber is the same. Figures from [64].

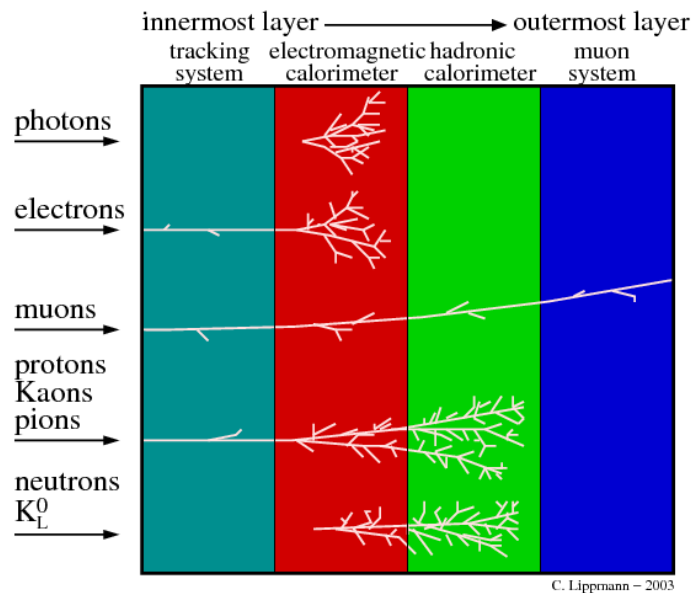


Figure 3.23 – Sketch of the interactions of different types of particles with the components of a *traditional* particle physics detector. Figure from [80].

can be briefly summarised by splitting them in three categories:

- **i sMuon**. This is a boolean variable indicating whether or not there are hits in the muon stations that can be associated to those from the tracking system, within a pre-defined geometric tolerance.
- **Delta-Log-Likelihood** or  $\Delta LL$ . These algorithms compute the difference between the natural logarithm of the likelihoods of a given track being compatible with two different particle hypothesis:

$$\Delta LL = \ln \mathcal{L}(t|h) - \ln \mathcal{L}(t|h') = \ln \left( \frac{\mathcal{L}(t|h)}{\mathcal{L}(t|h')} \right), \quad (3.5)$$

where  $t$  represents the set of values of the variables describing the observed track, coming from the different sub-detectors, and  $h$  and  $h'$  are the two different hypotheses.  $h'$  is conventionally always set to the pion hypothesis, while a different  $\Delta LL$  variable is defined for each possible  $h$ .

- **ProbNN**. These are the responses of Neural Network classifiers trained to recognise the different kinds of particles using as features the data provided by the sub-detectors, including tracking information, and taking into account correlations between detector systems.

The separation performance of the algorithms described above depends on the momentum and pseudo-rapidity ( $\eta$ ) of the tracks. For example, on Run I data, a cut on  $\Delta LL_{K-\pi} > 0$  gives on average a kaon identification efficiency around 95%, with  $< 10\%$  misidentification probability from pions.

Further details on the performances and strategy of particle identification in LHCb can be found in [77] and [81].

### 3.5.2 Trigger

The LHCb trigger is organised in three different levels, applied in cascade:

- **L0**, or *Level-0* is a hardware trigger, operating synchronously with the bunch-crossing frequency. It lowers the event rate to a maximum of  $\sim 1$  MHz. At this rate, the whole detector can be read out.

The L0 trigger uses information provided by three sub-systems: the calorimeter system, the muon system and the VELO pile-up system (composed of the two r-sensors placed upstream the interaction region) and it identifies the hadron, electron and photon clusters with the highest transverse energy in the calorimeters and the two tracks with the highest transverse <sup>4</sup> momentum ( $p_T$ ) in the muon chambers.

---

<sup>4</sup>The expression *transverse momentum*, denoted as  $p_T$ , is commonly used to indicate the component of the

- HLT1 is the first stage of the High Level Trigger (HLT), which consists of a C++ application running on a computer farm made of  $\sim 29,000$  CPU cores and uses information for all the sub-detectors. HLT, being a software application, is flexible and can be tuned to meet the experimental needs, adjusting the selection criteria according to physics priorities. The HLT1 provides a fast confirmation of the L0 decision using more refined information: a partial event reconstruction is performed at this stage on high- $p_T$  and high- $p$  tracks. Their impact parameter with respect to the primary vertex ( $IP$ ) is computed and used as a discriminating variable as well as the presence in the event of secondary vertices produced by the decay of flying particles. HLT1 reduces the event rate by a factor 30 with respect to the input received from L0.
- HLT2 is the second stage of the HLT. It uses a full event reconstruction and it consists of a series of selection lines run in parallel, corresponding each to a specific physics decay topology matching events of interest for the LHCb physics program.

The output of HLT2 is then written to disk to be further analysed, while the events that are not selected by the trigger are irretrievably lost.

The first stage of the trigger is emulated on Monte Carlo events to reproduce its effect on simulated data. The software stages are applied on simulation as they are on data. Figure 3.24 shows the flows of the trigger selection in place during the LHC Run I and Run II at LHCb. The improvement between the two runs is mainly due to the introduction of a new procedure for real-time alignment and calibration of the detector, as a first step towards the new trigger configuration foreseen for the LHCb upgrade before Run III (see Section 4). A detailed description of this procedure is provided in [83].

A physics decay candidate selected by the LHCb trigger is classified according to the following four definitions:

- TOS: the positive decision of the trigger is or could have been determined exclusively by particles belonging to the signal candidate itself, without depending on other objects (tracks, energy deposits...) belonging to the rest of the event.
- TIS: on the contrary, the trigger decision is or could have been determined by objects not belonging to the signal decay. TIS signal candidates are therefore accidentally unbiased with respect to the trigger.
- TOB: neither the signal candidate nor the rest of the event alone can generate a positive trigger decision, but tracks from both are necessary.
- Dec: any candidate that passed the trigger. TOS, TIS or TOB.

---

space momentum of a particle perpendicular to the beam axis. Similarly, the transverse energy  $E_T$  is  $\sqrt{m^2 + p_T^2}$  for a particle of mass  $m$ .

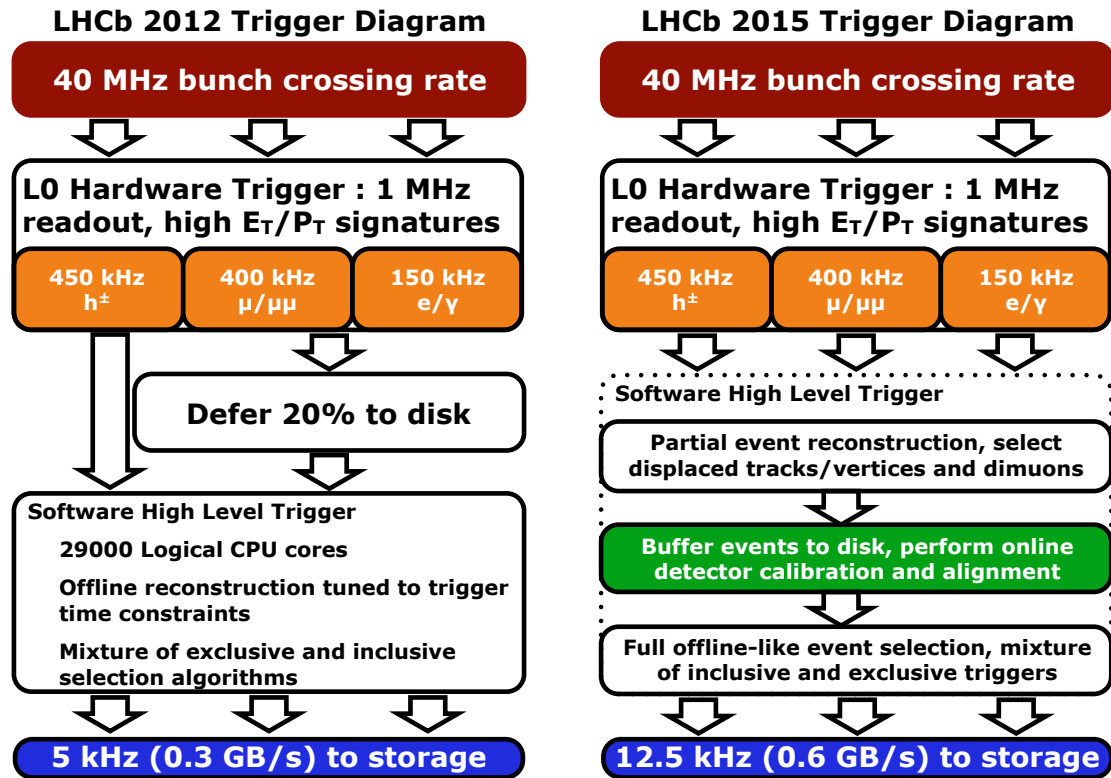


Figure 3.24 – LHCb trigger scheme during the Run I (left) and Run II (right) data taking periods. Figures from [82].

For each candidate and for each trigger line, three boolean variables are stored, corresponding to the three above-mentioned categories.

The fact that one candidate can be at the same time TOS and TIS, combined with the absence of trigger biases on TIS events provides a very useful way of computing the trigger efficiency as

$$\varepsilon = \frac{N^{\text{TIS\&TOS}}}{N^{\text{TIS}}} \quad (3.6)$$

where  $N^{\text{TIS\&TOS}}$  are the number of events classified as TIS and TOS at the same time, and  $N^{\text{TIS}}$  is the number of events classified as TIS. This procedure of determining trigger efficiencies is known in LHCb as TIS-TOS method; further details about it can be found in [84]. More details on the LHCb trigger in general are provided in [85].

### 3.5.3 LHCb software framework

The LHCb software is based on the Gaudi framework. The core software is mainly written in C++ and it is interfaced through Python configuration scripts.

The main components of the framework are described in the following sub-sections.

#### Gauss

A key ingredient of data analysis in particle physics is the study of data simulated with the Monte Carlo (MC) method.

This allows to improve the knowledge of what one should expect from an experiment. In particular it is often used for studying the background of a given signal, for calculating properties of the selection chain such as efficiency and purity and for *training* some multivariate analysis tools that are later used on data. More details are available in Ref. [86].

In LHCb, MC simulations are generated by the application Gauss, which consists of a collection of libraries for physics simulation. It generates the physical process of interest through the PYTHIA [87] generator package, that simulates the physics inherent the p-p interaction and the hadronization process. PYTHIA has been specifically tuned to reproduce the correct track multiplicities in the LHCb acceptance. The physics of  $b$  decays is handled by a specific package, called EvtGen [88], originally developed for the experiments BaBar and CLEO. The final-state radiation is generated using PHOTOS [89]. The detector response is simulated in a second stage with the Geant4 [90, 91] package, taking into account a very precise description of the detector geometry and the details of the physics processes behind the operation of each subdetector.

#### Boole

Boole simulates the response of the detector, translating charges and currents to what would be the output of the readout electronics as well as the decisions of the L0 hardware trigger.

This process is called *digitisation*. A detailed description of the response of each sub-system is obtained thanks to a careful calibration with test beams and with known processes in real data.

### Brunel

The digital signals coming from the readout electronics of the different sub-systems are then used, in real data as well as in simulation, to reconstruct the event. It is at this point that single hits are combined to extrapolate tracks and particle identification algorithms are run. Physical objects are at this point stored in *.dst* (data storage tape) files for *offline* analysis.

### Moore

Moore is the application that runs the HLT1 and HLT2 triggers on both real data and simulated events. The configuration of the triggers is provided with a *Trigger Configuration Key* (TCK), that defines the sequence of algorithms and cuts.

### DaVinci

The DaVinci software is used for offline data analysis. It combines the final-state particles in each event to search for candidates corresponding to the physics decay of interest. In a first stage, DaVinci is used to run a first series of simple cuts aiming at performing a preliminary selection of the wanted candidates. This process is called *stripping*, and it is explained in deeper detail in Section 5.1.2. DaVinci is also used for more user-end purposes, such as translating the *.dst* files in ROOT [92] ntuples, simultaneously computing and storing additional variables that will be used later in the analysis.

### 3.5.4 Stripping

The *stripping* is the first offline selection of physics events applied after the full reconstruction. It is executed centrally on each available dataset: each ongoing analysis has one or more *stripping lines*, i.e. python scripts containing dedicated selection algorithms. The stripping is executed once on every set of data (the year of collection usually defines the set) and re-run later – still collectively – in case important improvements that can affect the performance of the analyses are achieved in the selection or in the offline reconstruction algorithms.

The output of all the lines is written to disk and made accessible to the collaboration, so that the analysts are able to retrieve the candidates selected by their stripping line for further analysis. The stored events are organised in twelve different *streams*, according to the general kind of physics of the line by which they were selected. For example, lines selecting events containing two muons will store their output to the *dimuon* stream, and lines selecting charm physics will write in the *charm* stream. The streams are in turn organised in two groups, based



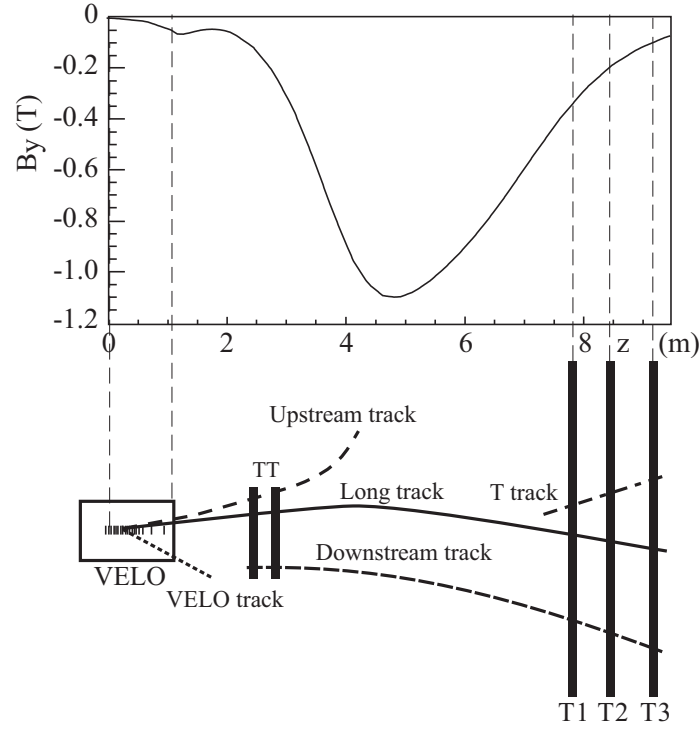


Figure 3.25 – Illustration of the track types in the LHCb detector. The main component of the magnetic field  $B_y$  is shown above as a function of the  $z$  coordinate. Figure from [72].

on what kind of information is retained: *dst* streams contain all of the event's information, while *micro-dst* (or *mdst*) streams only contain a limited set of information relative to the signal candidate. This allows to significantly reduce the size of the files stored to disk (trading it for a higher allowed rate) for studies not requiring full information on the event and detector hits.

An alternative approach to the stripping for analysis data selection is represented by the *turbo* stream, which takes advantage of the on-line reconstruction performed in the HLT to store to disk directly the candidates identified by the trigger. For further details see [93].

### 3.5.5 Tracks classification

The tracks of charged particles that leave sufficient hits in the LHCb detector are reconstructed combining hits from the different components of the tracking system to form particle trajectories.

These tracks are classified in five different types, depending on their path in the detector, as illustrated in Figure 3.25:

- **Long tracks:** these are the tracks that go through the entire tracking system from the VELO to the T stations. This gives the most precise momentum determination achievable in LHCb.
- **Upstream tracks:** these tracks traverse only the VELO and the TT stations. They have in general lower momentum than the long tracks and are therefore bent out of the detector acceptance by the magnetic field before entering the tracking stations T1-T3.
- **Downstream tracks:** transversing only the TT and T stations, these tracks are not seen in the VELO. They are mostly originated from  $K_S^0$  mesons or  $\Lambda$  baryons decaying outside the VELO acceptance;
- **VELO tracks:** as opposed to the others, these tracks are only seen in the VELO. They are usually tracks having a large angle with respect to the beam, or backward tracks, and are used for improving the reconstruction of primary vertices. No momentum measurement is available for VELO tracks;
- **T tracks:** these tracks only produce hits in the T stations; they are typically produced in secondary interactions, and are mainly used for the global pattern recognition in RICH2.

More details on the reconstruction of tracks in LHCb can be found in [72] and [64].

### 3.5.6 Reconstruction of electrons

Of particular interest for the work presented in the following of this thesis is the treatment of electron final-states in LHCb.

Given their low mass, the interaction of electrons with matter is often accompanied by the emission of bremsstrahlung radiation, as opposed to other final-state particles, in particular to muons, which leave on the contrary a very clean signature in the detector.

This loss of energy by radiation entails the need of dedicated reconstruction algorithms and specific precautions in analysis strategies. If an electron emits a photon after having passed through the magnet, the two particles will likely proceed on very close trajectories and hit the same region of the calorimeter, being thus reconstructed as one single object. If, on the other hand, a photon is emitted upstream or inside the high magnetic field region, it will not follow the same trajectory as the electron, as it is not charged, and it will thus hit a different region of the calorimeter.

Dedicated reconstruction algorithms look for clusters in the ECAL that can be compatible with bremsstrahlung photons and assign their energy to the electron track that emitted them, to compensate mismeasurements of momenta that would in turn imply errors in the determination of the quality of the track and its origin vertex. The resulting momentum resolution, though, is degraded by the finite resolution on the photon detection, and this reflects on a degradation of the spectra of momenta, invariant masses and chi-squares of tracks and

vertices.

Furthermore, two opposite kinds of errors can happen in this procedure:

1. a fake or a background photon could be mistakenly identified as originating from the electron, and added to it;
2. a real bremsstrahlung photon might not be detected or not be identified correctly, being in fact lost.

For the above-mentioned reasons, in the two physics analyses presented further in this thesis, each final-state electron is classified in two categories depending on whether or not a bremsstrahlung photon was added to it in the reconstruction. This information is stored in a boolean variable called `HasBremAdded`. This allows to study the selection efficiencies and the spectra separately for the two categories, combining then appropriately the results according to the expected relative fractions.



## 4 The LHCb upgrade

The capability to record a large number of physics events is a key feature for an experiment that studies rare phenomena and performs precision measurements. LHCb has obtained important results during the Run I and Run II phases of the LHC, and the collection of new data in Run III (2020-2025) will reduce statistical uncertainties on many existing analyses and allow to perform new ones.

Currently, though, the amount of recordable interesting data is limited by the maximum readout rate. The instantaneous luminosity delivered by the LHC is indeed levelled in order to meet this limitation, as previously mentioned in Section 3.2.

For this reason, the LHCb detector will undergo a major upgrade during the shutdown period (LS2) that follows the end of Run II in 2018. This upgrade will allow to run at higher instantaneous luminosities ( $2 \times 20^{33} \text{cm}^{-2} \text{s}^{-1}$ ) while keeping the same or higher efficiency. Figure 4.1 shows the schedule of the LHC commissioning from year 2015 to 2021.

Two main sets of modifications will be applied:

- Detector upgrade, aiming at being able to read out all the subsystems at 40 MHz and to cope with the significantly increased pile-up and track multiplicity while at the same time increasing the measurements precision.
- L0 upgrade: the hardware trigger will be replaced by a software one, with a much higher flexibility. For example, the possibility to cut on calculated, potentially complex, quantities will allow to lower the transverse momentum thresholds significantly increasing the efficiency on many decays and opening to kinematic regions that are currently inaccessible.

A brief overview of these two topics is provided in the next subsections, and a more detailed description of the downstream tracker upgrade – relevant to this thesis – can be found in Section 4.1.

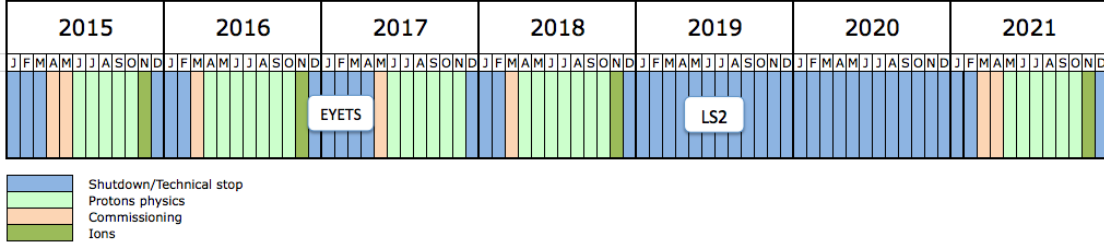


Figure 4.1 – The LHC commissioning schedule from year 2015 to 2021. LS stands for Long Shutdown, while EYETS indicates the Extended Year-End Technical Stop. Figure from [94].

### 4.0.1 Detector upgrade

Major modifications will be applied to different subdetectors of LHCb.

The whole tracking system will be replaced. In particular, new hybrid pixel sensors will replace the VELO, covering a region closer to the beam (5.1 mm compared to the current 8.1 mm); the TT will be replaced by the UT, composed of four layers of silicon detectors with higher granularity and covering a larger acceptance in the central region; the IT and OT will be removed and a new Scintillating Fibre Tracker (SciFi) will be installed. The new tracking system will be read out at 40 MHz and will allow a more resource-efficient track reconstruction.

More details on the tracker upgrade can be found in [95], and a more extensive description of the SciFi is provided in this thesis in Section 4.1.

Partial upgrades of the particle identification subsystems are also scheduled. The RICH will see its mirrors replaced with new ones having higher reflectivity and larger focal length, to cope with the expected increase in occupancy, and its HPDs substituted by Multi-anode PMTs (MaPMT), allowing faster readout. The readout electronics of the ECAL and HCAL will be upgraded to meet the new 40 MHz requirement.

Further details can be found in [96].

As for the muon system, minor changes are foreseen: the M1 station will be removed, as its purpose was mainly to provide input for the hardware L0 trigger; the region around the beam pipe in correspondence to the M2 station will be better shielded to counterbalance the expected higher occupancy from showers originating from material interaction and, like in the other subdetectors, the readout electronics will be substituted to allow a 40 MHz readout. A more detailed description of the muon system upgrade is provided in [96].

### 4.0.2 Trigger upgrade

A pile-up above 5 is expected at Run III running conditions, which should be compared to the current 1.6. This requires the trigger to be extremely efficient. The limited information presently available at L0 level would lead to prohibitive efficiency losses.

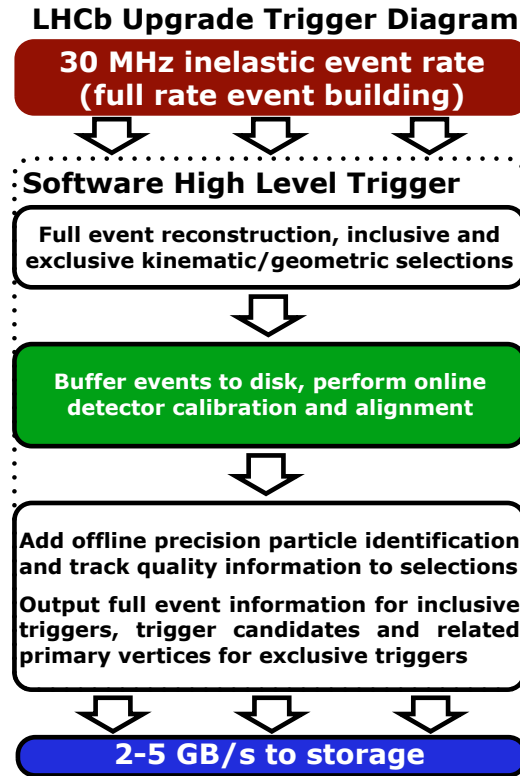


Figure 4.2 – LHCb trigger scheme foreseen for Run III. Figure from [82].

The new L0 software will allow a much faster readout and will be able to process more complex information: a fast and simplified track reconstruction algorithm will indeed run online, followed by the online detector alignment procedure that was already introduced in Run II, allowing to perform physics analysis directly on the data coming from the trigger. Figure 4.2 shows the new trigger flow scheme, to be compared with Figure 3.24. For Run I and Run II diagrams, the shown input rate is 40 MHz, while it is 30 MHz in the new scheme. The reason for this is that the current detector can only be read at 1 MHz, so the figure of interest is the overall bunch crossing rate, which has to be filtered by the hardware trigger, whereas, for the upgrade detector, every bunch crossing can be actually read out, thus what matters is the non-empty bunch crossing rate (30 MHz), as this is what has to be reduced by the HLT. A complete description of the LHCb trigger upgrade for Run III can be found in [97].

Figure 4.3 shows a side view of the LHCb detector as it will be after the upgrade.

## 4.1 The Scintillating fibre Tracker - Sci-Fi

In the upgrade phase of LHCb, both the IT, composed of silicon microstrips and the OT, composed of straw drift tubes, will be replaced by a single new detector: the scintillating fibre tracker, SciFi.

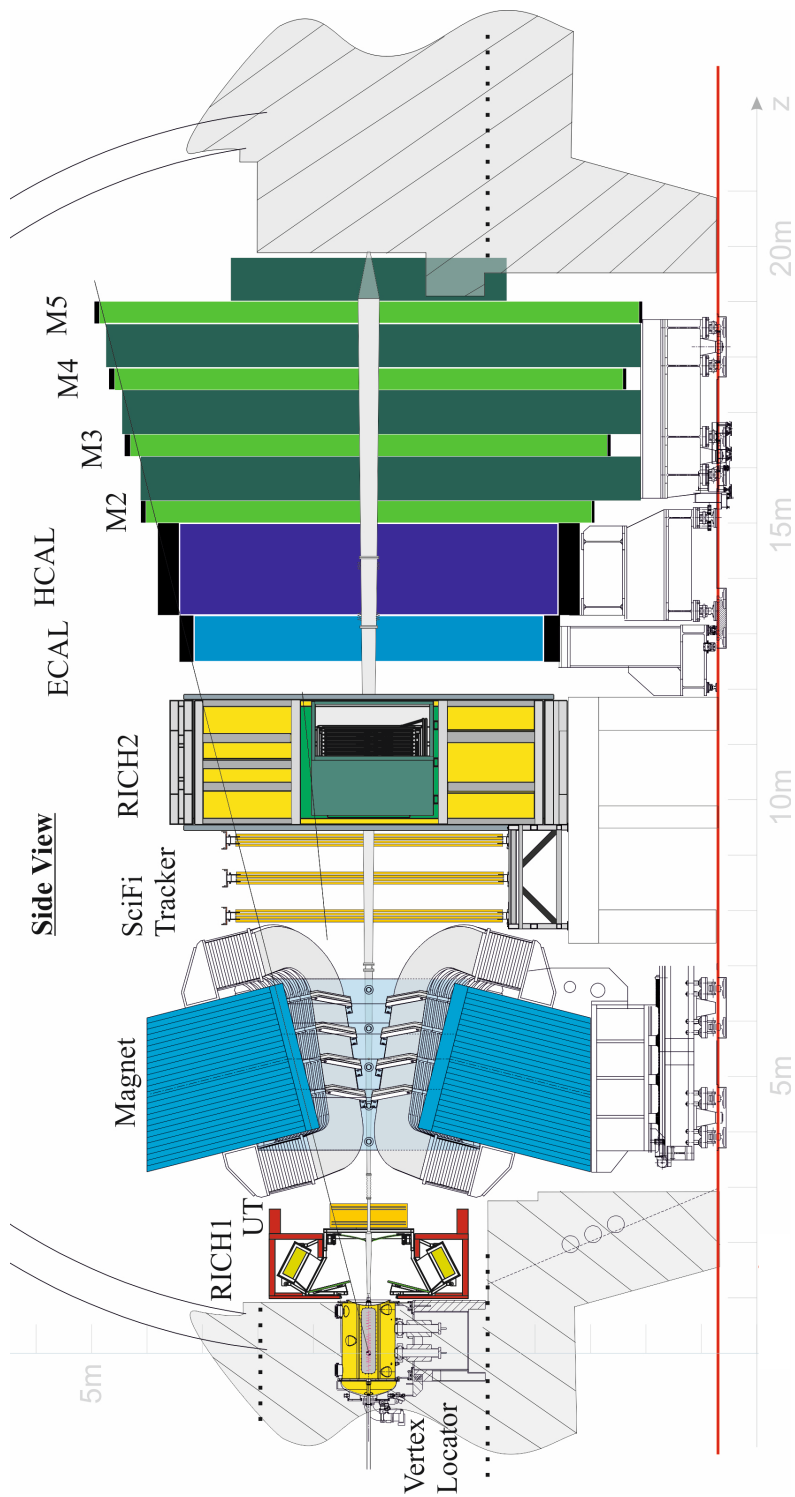


Figure 4.3 – The LHCb detector as it will be after the upgrade, with its sub-detectors, side view.



Charged particles are detected in the SciFi by collecting the photons emitted when these particles transverse a series of layers of fibres made of scintillating material.

The details of the SciFi design and construction procedures are summarised in this section and in Ref. [98].

As mentioned in the previous sections, the LHCb detector in the upgrade will run at the increased instantaneous luminosity of  $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  with a bunch crossing time of 25 ns and an average pile-up of about 5. The current tracking system was designed to operate with lower occupancies than the ones expected at these conditions, and will therefore be replaced by the SciFi. The new tracker will need to provide high spacial resolution, high hit efficiency, clean and fast signal, and it will also have to age slowly enough to keep good performances during the whole Run III, allowing LHCb to reach the planned recorded integrated luminosity of  $50 \text{ fb}^{-1}$ .

The most important requirements mentioned in the technical design report [95] are indeed:

- hit detection efficiency as high as possible ( $\sim 99\%$ ), with a very low ( $< 10\%$ ) noise-to-signal at any point along the detector;
- single-hit spatial resolution in the bending plane of the magnet  $\leq 100 \mu\text{m}$  (this figure originates from the lower bound obtained from the limitations due to multiple scattering effects in the extrapolation of tracks from the VELO);
- amount of material in the acceptance region as low as possible, (less than 1% of the radiation length for each layer) to keep under control the effects of multiple scattering;
- read-out electronics able to run at a frequency of 40 MHz;
- capability to operate for the full lifetime of the LHCb upgrade, reaching at least the planned integrated luminosity of  $50 \text{ fb}^{-1}$ .

### 4.1.1 Detector overview and layout

This section contains an overview of the main elements of the SciFi detector and a description of its layout. Its fibre mats, being more closely a subject of this thesis, are described in deeper detail in Section 4.1.2.

Like the current IT and OT, the SciFi Tracker will be arranged in three stations – T1, T2, T3 – each composed of four detection layers arranged in a  $x$ - $u$ - $v$ - $x$  configuration as shown in Figures 4.4 and 4.5.

A rectangular hole in the centre of each layer will accommodate the beam pipe. Each layer will cover  $6 \text{ m} \times 5 \text{ m}$  in the  $x - y$  plane. The plane is sub-divided into five-metres-high modules, with a width of 0.52 m, resulting in 12 modules per plane. Each module is in turn composed of eight 2.5 metres-long, 130.45 millimetres-wide mats of scintillating fibres; four placed side-by-side on the top half and four similarly on the bottom half along the  $x$  direction. At the top

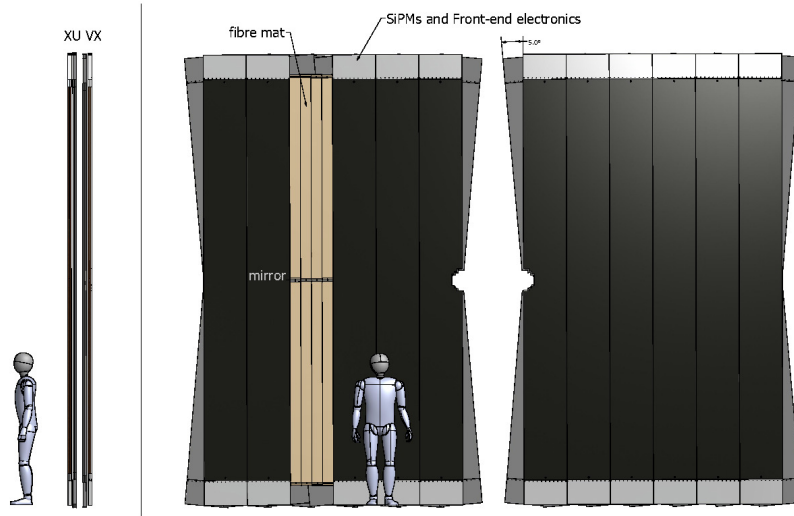


Figure 4.4 – A schematic view of one station of the SciFi Tracker. Figure from [95].

and at the bottom of each module are located 16 Silicon Photomultiplier (SiPM) arrays – 4 per mat – that allow to detect the shining light produced in the fibres. At the opposite end of each mat, *i.e.* at the half of each module, reflective mirrors are installed. More details on the fibre mats are provided in Section 4.1.2, preceded by a description of the scintillating fibres in Subsection 4.1.1.

The above-mentioned SiPMs are solid-state photomultipliers of small size and with high granularity. Specifically, the sensors employed for the SciFi are produced by the company Hamamatsu Photonics K.K.

SiPMs are composed of Avalanche Photodiodes (APD), each constituting a pixel. The APDs are operated in Geiger-Mode, *i.e.* a reverse bias voltage slightly above the breakdown voltage is applied. In this way, a single electron-hole pair generated by the absorption of a photon can trigger an avalanche current.

The SciFi SiPM detectors are composed of multichannel arrays of 128 channels, with a pitch of  $250\text{ }\mu\text{m}$  and an height of  $1.625\text{ mm}$  corresponding to that of a fibre mat. The 128 channels are in turn arranged on two adjacent 64-channels silicon dies. Each channel is composed of 104 pixels with a pixel size of about  $60 \times 60\text{ }\mu\text{m}^2$ , connected in parallel. This rather large pixel size allows to reach a high photon detection efficiency (above 50% at peak). Further details on the engineering of the SiPMs can be found in [99].

Figure 4.6 shows a SiPM array mounted on the cable (flex PCB) that connects it to the front-end electronics, while Figure 4.7 shows three photos at increasing zoom levels, where the structure of arrays, channels and pixels can be observed.

The SciFi readout system is composed of front-end electronics and back-end electronics. Front-end electronics boards are attached directly to the detector modules. They include a custom integrated circuit (ASIC) called the low-Power Asic for the sCIntillating Fibres traCker

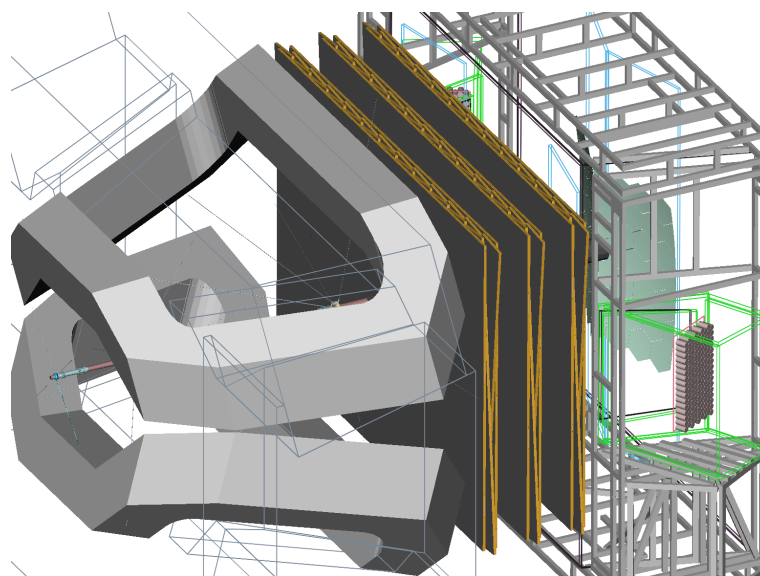


Figure 4.5 – The three stations of the SciFi tracker shown between the dipole magnet, on the left, and the RICH2, on the right. Figure from [95].

(quite imaginatively acronymised into PACIFIC) along with a Field Programmable Gate Array (FPGA) responsible for the clustering algorithm. The PACIFIC boards amplify, shape and integrate the SiPM signal over the LHC bunch crossing period, converting it to a digital signal with adjustable thresholds. The signals from the single channels are then merged in clusters by the FPGA, according to a specific algorithm based on identifying single channels or groups of neighbouring channels exceeding specific charge thresholds. The back-end electronics, on the other hand, will be placed outside the LHCb cavern. They elaborate the information from the front-end modules and recognises patterns to reconstruct tracks.

The production of all the components of the SciFi requires to be closely followed and monitored. A crucial operation is the continuous evaluation of the characteristics of the produced elements, through accurate measurements of specific parameters. This procedure is indicated as *quality assurance* (QA). Section 4.1.3 contains a description of the quality assurance procedure for the SciFi fibre mats, and an overview of some of the results obtained in the context of this thesis is provided in Section 4.2.

### Scintillating fibres design and construction

The scintillating fibres used in the SciFi are produced by the company Kuraray. They have been selected for their fast decay constant of 2.8 ns and their high light yield. The fibres have a round cross section with a diameter of 250  $\mu\text{m}$ . The core material is doped polystyrene, surrounded by a double cladding with decreasing refractive indices allowing to guide the scintillation light along the fibre via total internal reflection. Figure 4.8 shows a sketch of the longitudinal section of a fibre. As discussed further in this section, an important parameter

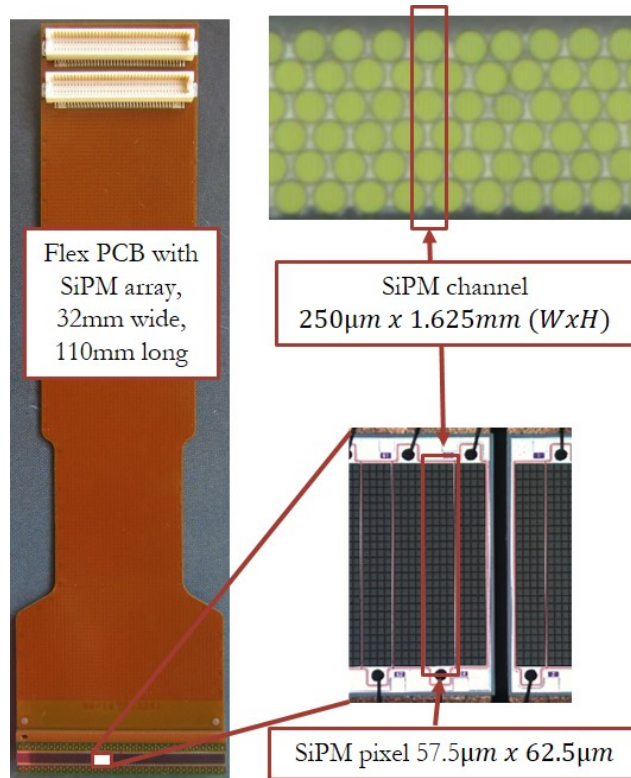


Figure 4.6 – SiPM array mounted on flex PCB on the left, zoom into the single channel size and the fibre mat on the right. Figure from [99].

of the fibres is the attenuation length, which is expected to decrease in time because of the radiation to which the detector material will be exposed in the LHCb cavern.

Before winding the fibres into mats, their quality is analysed at CERN, and possible bumps in their diameter are removed, either by shrinking or by cutting them [101]. Further details on the scintillating fibres can be found in [100].

### 4.1.2 Fibre mats

The scintillating fibres are wound in 6 layers to form rectangular mats with a length of 2.4 metres and a width of 13.5 cm.

To reach the design spatial resolution better than 100 µm the position of the single fibre along the mat needs to be known precisely, which implies that the fibre needs to be parallel along the mat. For this reason, the fibre is wound on a threaded wheel having a pitch of 275 µm. This distance is intentionally slightly larger than the fibre diameter, in order to accommodate fluctuations of the latter without altering the path of the neighbouring fibres. The positioning of the first layer follows these threads on the wheel surface, while in every successive layer the fibres are guided by the underlying ones; their position is therefore shifted by half the

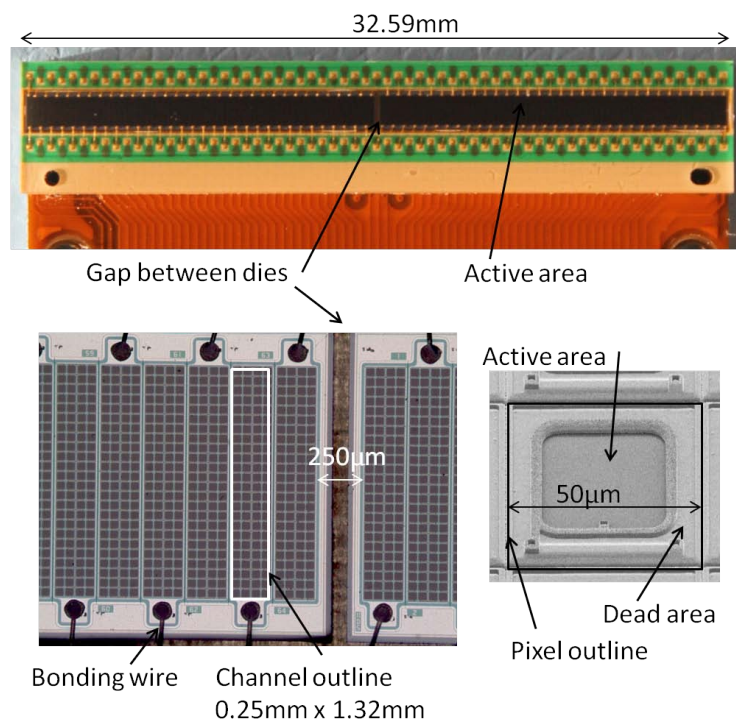


Figure 4.7 – Top: Package with two 64-channels silicon dies. Bottom left: the gap between two silicon dies shown under the microscope. Bottom right: a pixel with optical trenches. Figure from [95].

pitch size with respect to the previous and next layers. The cross section of a mat is shown in Figures 4.9 and 4.10.

Between consecutive layers, liquid glue is applied to preserve the geometry and to keep the mat together. The fibre tension is kept constant during the winding to allow a precise positioning. At the end of the winding process, after the curing of the glue, the mat is cut transversally to the fibre direction, removed from the wheel and flattened.

In order to help the correct alignment during the assembly of the modules, *alignment pins* are positioned on the first layer of each mat. These are semicircle-shaped bumps formed by the glue entering dedicated holes on the surface of the wheel.

The transparency of the fibre is negatively affected by the radiation present around the beam pipe. Figure 4.11 shows the integrated ionising dose predicted after the Run III operations. As expected, it peaks in the region closer to the beam. Irradiation measurements performed by the LHCb SciFi group have shown that the total loss of signal in this region is expected to be around 35% [102]. In order to preventively counterbalance this loss, the fibre mats that are found to have higher light yield will be placed closer to the beam pipe.

A 25-µm-thick kapton foil is glued on each side of the mat in order to protect it from mechani-

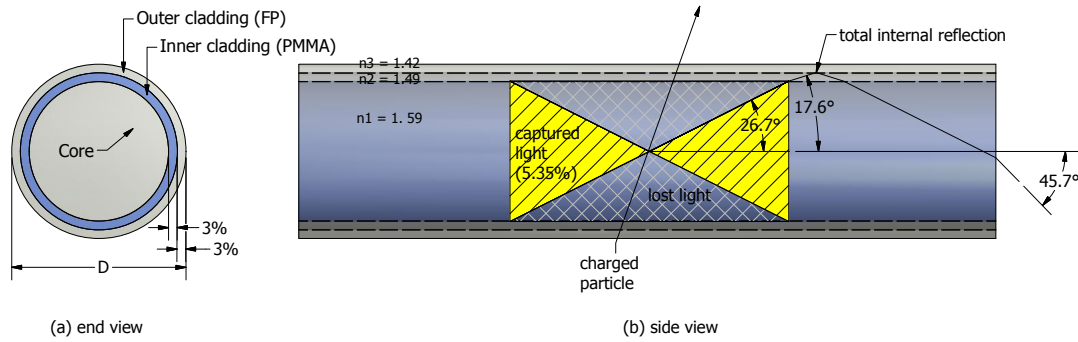


Figure 4.8 – Sketch of the longitudinal section of a scintillating fibre. Light is produced in the core material and then trapped and propagated within the fibre through total internal reflection. The claddings refraction indices are indicated, along with the maximum incidence angles at which emitted light is captured, corresponding to the indicated 5.35% of the solid angle. Figure from [100].

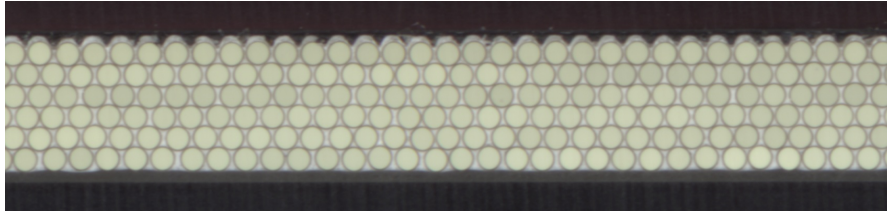


Figure 4.9 – Part of the cross section of a fibre mat.

cal damages and from light. At the ends of each mat two different plastic endpieces are glued, to support its structure and allow the alignment to the SiPMs on one end and the gluing of a mirror on the other. Both endpieces have holes that allow a precise alignment when mounted in the jig. Figure 4.12 shows the SiPM-end of a mat, coated with the kapton foil and with the endpiece glued.

In order to provide a smooth and flat surface against which the SiPMs will be placed, a precise diamond milling of the end of the fibre mats and endpieces is performed. This allows to have a uniform contact with the photomultipliers, without gaps. The mats are also cut longitudinally to match the required width within a precision of  $150\ \mu\text{m}$ .

On the opposite side of where the SiPMs will be placed, a mirror foil is glued, to reflect the light travelling in this direction, redirecting it to the photomultipliers with minimum loss (mostly due to attenuation).

Before and after the mirror gluing, a light yield test is performed on each mat. This procedure and its results are described in detail in Section 4.2.

Technical details about the realisation of the SciFi scintillating fibre mats are provided in Ref. [100], while Ref. [103] provides a didactic introduction.



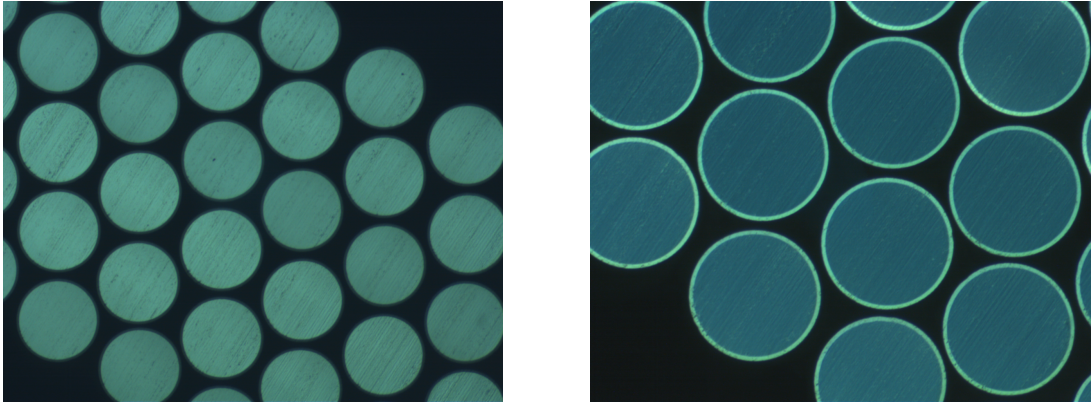


Figure 4.10 – Segments of the cross section of a fibre mat produced at EPFL, illuminated at the opposite end to enhance contrast, in order to evaluate the cleanness of the cut and the regularity of the geometry. On the right, the fibre cladding is also clearly visible.

#### 4.1.3 Fibre mats quality assurance

The production of the scintillating fibre mats for the SciFi is carried out in parallel in four different production centres located in Lausanne, Aachen, Dortmund and Moscow. Common procedures have been established and carefully followed in order to obtain homogeneous results, and quality measurements are performed after some of the production steps. These tests not only are needed to reject defective mats, but are also used to grade the quality of the accepted mats in order to classify them with the purpose of placing similar ones in regions of the detector with similar hit occupancies, with the best ones in the busiest regions, as anticipated in the previous section.

Figure 4.13 shows a scheme of the production steps and quality assurance measurements for the SciFi fibre mats.

#### Online monitoring

A first frequent issue that potentially compromises the quality of the mats is encountered during the winding process: imprecisions in the fibre positioning might cause the jump of one groove or the overlap of two fibres on the same position. Such imprecisions can be due to bumps in the fibres or to a transversal wobbling of the winding wheel (order of 100  $\mu\text{m}$ ), caused by microscopic imperfections in its geometry.

In order to prevent this, an industrial camera and a dedicated multivariate-analysis-based software are used to automatically detect such errors, that are then corrected manually.

#### Shrinking of the fibres

As mentioned in the previous section, a moderate tension is applied to the fibre during the winding process. This helps to position the fibre with precision, but it causes an extension

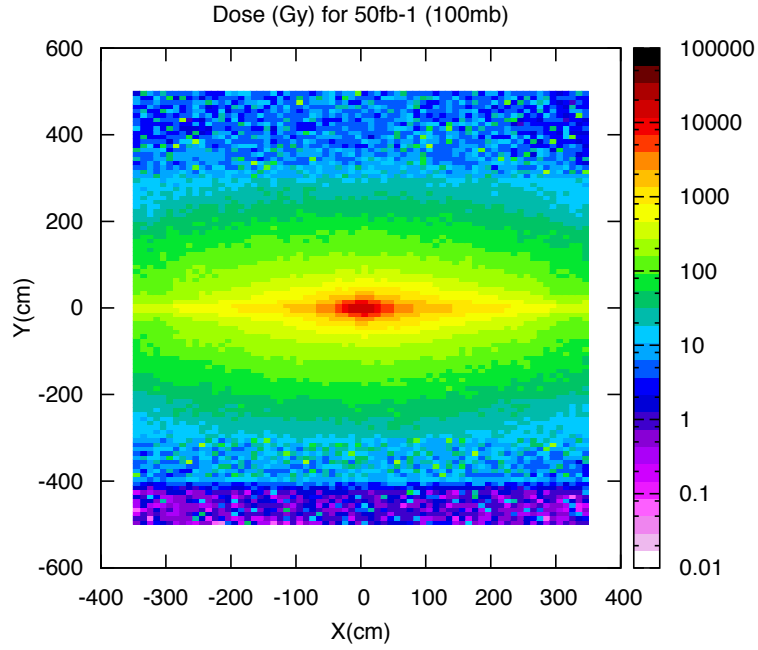


Figure 4.11 – The expected dose in the  $x$ - $y$  plane at  $z = 783$  cm (T1 position) after an integrated luminosity of  $50 \text{ fb}^{-1}$ . Figure from [95].

of about 1 cm on the full length of the mat. When the mat is cut on the winding wheel, this tension is instantly released and the fibres shrink. In addition to this *fast shrinking component*, it has been observed that the length of the mat keeps reducing slowly for approximately the following 30 days (*slow shrinking component*) by a length varying from a fraction of a millimetre up to about 2 mm. This is problematic since the mat has to fit precisely in the supporting frame. The relative length change has been measured on a set of mats during a few weeks after the winding; it has been observed that it follows a double-exponential PDF. A tolerance of 0.2 mm has been established, and the endpieces are glued to the mats only when the expected future length change does not exceed this threshold.

### Optical scan

After the optical cut, quality assurance measurements are performed *a posteriori*.

The first one is the optical scan of the mat section: the quality of the cut, the good shape (fibres section, geometrical arrangement) and the transparency of the fibres are checked by taking high-resolution images of the SiPM end of the mats. Two scans are performed: for the first one, only the light source internal to the scanner is used, while the second picture is obtained with an additional light placed at the opposite end. The transparency of the fibres is evaluated by comparing the two images. A pattern recognition software is used to identify the section of the single fibres and find deviations from the expected circular shape and the relative position, that can result from mistakes in the winding process or from a bad cutting.



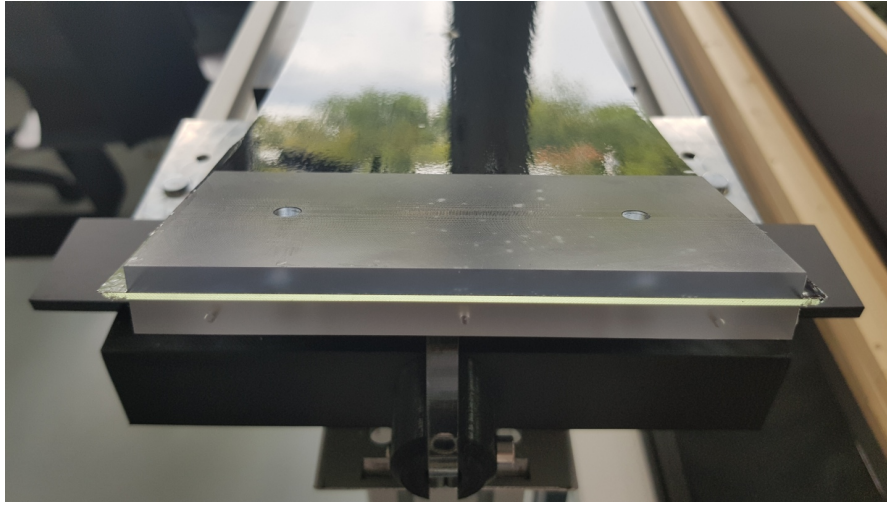


Figure 4.12 – The SiPM-end of a fibre mat produced at EPFL. The kapton foil (black reflective surface) and the endpiece are clearly visible.

A grade from 1 (best) to 4 (worst) is then attributed to each mat. The evaluation is based on the number of fibres lying out of the SiPM bounds, which in turn is defined to be slightly larger (maximum spread in the vertical plane:  $\pm 75 \mu\text{m}$ ) than the mat cross section to allow imperfections in the relative alignment between the photodetector and the mat.

#### Light yield test

The *light yield* (LY) of a fibre mat is the average amount of photons detected by the SiPMs when a charged particle transverses the scintillating fibres, measured in photoelectrons (p.e.). This figure is essential for characterising the mats, *i.e.* spotting possible defects such as non-uniform sensitivity, bad mirror gluing or damages to the fibres.

In order to measure the light yield, a  $^{90}\text{Sr}$  radioactive source is placed above the mat, close to the mirror-end, to illuminate it with electrons, produced by  $\beta$  decay.

The electrons produce scintillation light when crossing the fibres, as it would happen in operating conditions. This light is detected with an array of SiPMs with 512 channels, whose response is processed by a dedicated software running on a commercial PC.

The light yield measurement is performed twice on each mat: before and after gluing the mirror, and the two results are compared. The procedure and the results of the light yield tests of the fibre mats produced at EPFL are discussed in detail in Section 4.2.

Each mat receives a grade from 1 (best) to 4 (worst) given by the number of channels reading a light yield below a pre-determined threshold, computed as 2 RMS below the mean light yield of a set of several analogous light yield measurement. Table 4.1 shows the grades definition.

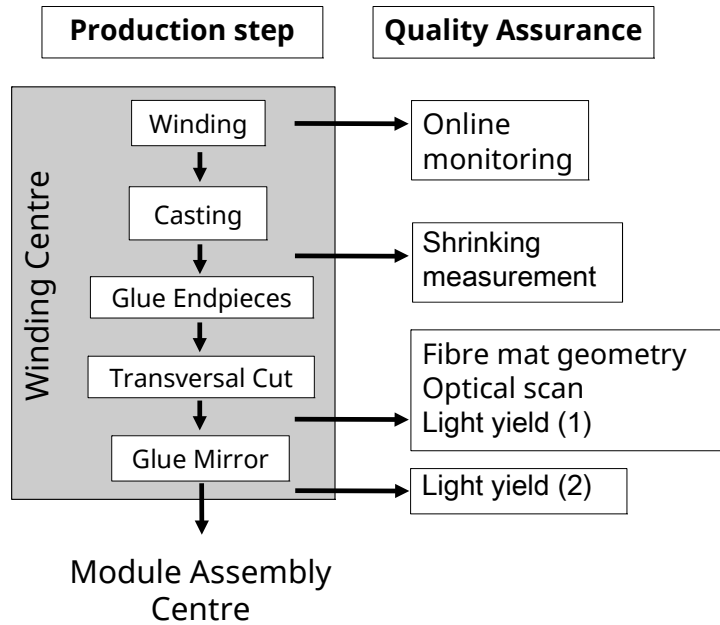


Figure 4.13 – Scheme of the production steps and quality assurance measurements for the production of SciFi fibre mats.

Table 4.1 – Light yield grades definition.

| Grade | # of low-LY channels |
|-------|----------------------|
| 1     | $\leq 20$            |
| 2     | $\in [21, 50]$       |
| 3     | $\in [51, 200]$      |
| 4     | $> 200$              |

### Geometry measurements

As a last check before shipping the mats to the module assembly centres, the geometrical properties are evaluated. Four parameters are considered:

- the length;
- the amount of excess residual glue;
- the presence of cracks;
- the integrity of the alignment pins.

Each of these receives a score from A (best) to D (worst) according to precise criteria. For example, the grade on the length is defined in intervals, from  $[2424.0, 2424.5]$  mm, which is the

nominal length and corresponds to grade A, to the extreme case where the mat has shortened so much that it does not fit anymore in the jig, corresponding to grade D.

### Overall grade

The scores assigned with the optical scan and with the light yield tests contribute to a single *performance mark*, defined as the worst of the two. Similarly, the *geometrical mark* is the worst of the marks assigned to the four geometrical measurements.

The final grade is then formed by a letter and a number, representing these two grades. Furthermore, an integer approximation of the light yield (with mirror) in photoelectrons (p.e.) is indicated.

For example, an excellent-quality mat will have grade A1-21.

Further details on the SciFi quality assurance procedures can be found in [95] and [100].

## 4.2 Light yield tests of fibre mats from EPFL

### 4.2.1 Goal

An introduction to the light yield measurement is provided in Section 4.1.3. Its purpose is to characterise the mats, grading them according to the amount of light that they can produce when transversed by charged particles and to spot possible defects and damages.

This measurement is performed twice on each mat: a first time right after the optical scan, and a second time after gluing the mirror. The first measurement allows to identify immediately defects like possible cracks in the mat or bad quality of the fibre, while the second, compared with the first, allows to assess the gain due to the mirror reflection, and to spot misaligned or detached mirrors. Each measurement, before and after the mirror gluing, is in turn repeated twice: once with and once without a 3-mm-thick plastic absorber placed between the mat and the scintillating trigger (described in Section 4.2.2). This absorber allows to get rid of very-low-energy electrons, that loose more energy than a minimum ionising particle (higher kinetic energy), thus obtaining conditions more similar to the operations ones.

In addition to serving as a quality assurance test, the LY test has driven the mat production since its beginning, providing an immediate feedback. This allowed to define the manufacturing procedures by spotting those actions that tend to damage the product.

The **light yield** is measured **per channel**, across the 512 channels of the SiPM array used in the setup. A light yield **per cluster** is also defined, summing on the channel that form each cluster, as well as a light yield **per mat**, given by the average of the per-cluster light yield across all the channels to which it is assigned. The latter defines the final LY figure.

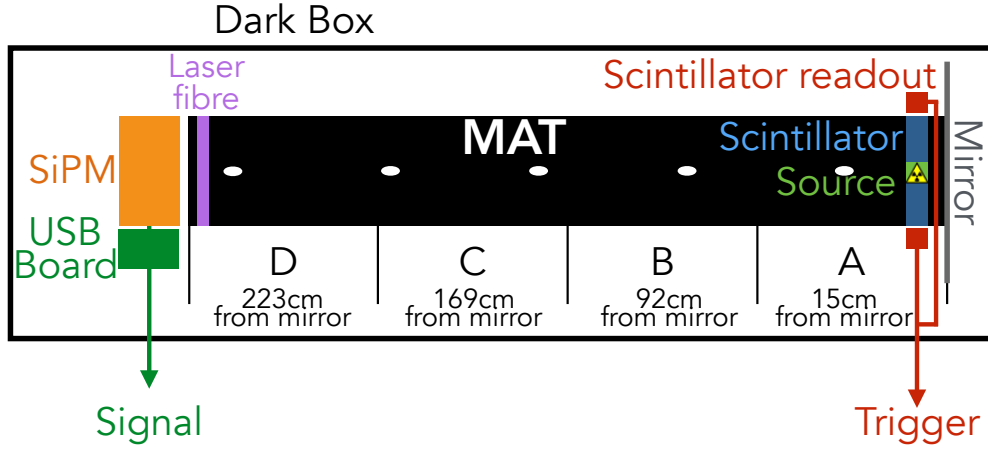


Figure 4.14 – Sketch of the light yield measurement setup. A, B, C and D indicate the positions for the  $^{90}\text{Sr}$  source. Only position A is used in nominal tests, while the others have been used for attenuation length measurements.

#### 4.2.2 Setup

The experimental setup used to perform the LY test is sketched in Figure 4.14 and described in this section.

As mentioned in Section 4.1.3, the light yield is measured by irradiating the mats with electrons produced in the  $\beta^-$  decay



followed by



The half-life of  $^{90}\text{Sr}$  is about 29 years, which makes it a durable source, suited for this purpose. The energy of the emitted electron is 0.546 MeV. The intermediate yttrium atom has instead a half-life of about 60 hours and produces electrons with an energy of 2.28 MeV.

The readout is triggered by a scintillating bar detector placed below the mat, and the light is collected on the opposite end by an array of SiPMs, mounted on a sliding cart that allows to get the detectors as close as possible to the mat during the scan, and to retract them while the mat is brought into position, in order to protect them from accidental scratches. Furthermore, the alignment of the SiPMs with respect to the mat can be finely adjusted horizontally and vertically using three micrometer screw gauges, in order to obtain the highest and most uniform achievable result.

Figure 4.15 shows the impact of the positioning between mat and SiPMs on the LY: after a 100  $\mu\text{m}$  plateau, the number of detected photoelectrons drops drastically, while the cluster size results at first artificially enlarged, as expected from simple geometrical reasons, and

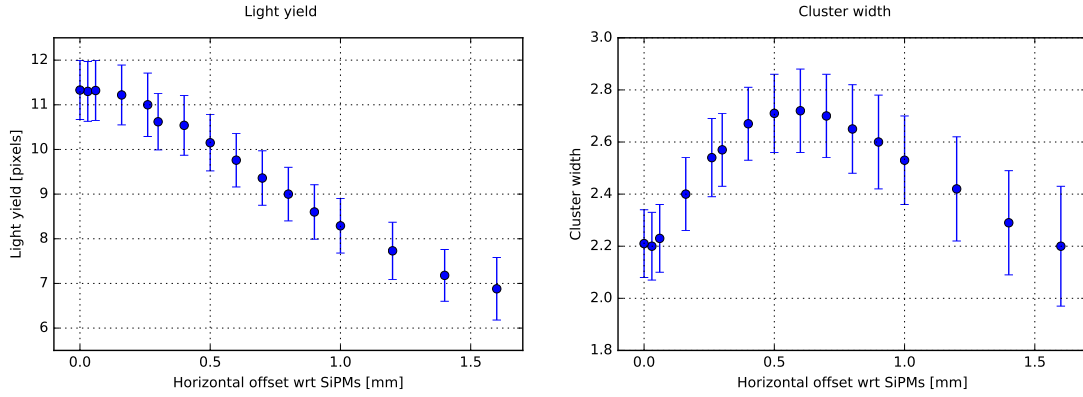


Figure 4.15 – Light yield (left) and cluster size (right) as a function of the horizontal distance between the SiPM array and the fibre mat. These scans have been performed on a test mat without mirror, for this reason, the absolute values on the y axes are not representative of the nominal ones observed in production mats.

subsequently drops due to attenuation. Figure 4.16 shows instead the effect of the vertical misalignment. It is evident that the rather smooth central part of the distributions allows a tolerance larger than  $\pm 5 \mu\text{m}$  in the alignment.

Underneath the mat, on the SiPMs side is placed a scratched fibre connected to a laser, used to inject photons into the mat (the capton foil is, on purpose, not present at this specific site) for calibration purposes. The scratches allow the light to escape the fibre and to be scattered in multiple directions.

The amount of light produced in the scintillating fibres is negligible compared to a normal ambient light, thus the measurement needs to be performed in the dark. For this reason the whole setup is placed inside a closed box.

The SiPM array is connected via a USB board to a commercial PC, where a dedicated software reads and controls the various components and processes the information into readable plots and figures. A calibration is performed before each measurement. The following steps describe the measurement procedure:

- the ADC pedestal of each channel is identified, using data collected with a random trigger and no light source;
- the one-photon peak of each channel is identified using the light injected by means of the scratched fibre;
- from the pedestal and the one-photon peak, the gain is computed;
- the actual measurement is performed, using the scintillator trigger, and histograms of the raw ADC value, cluster size and LY per channel and per cluster are filled;

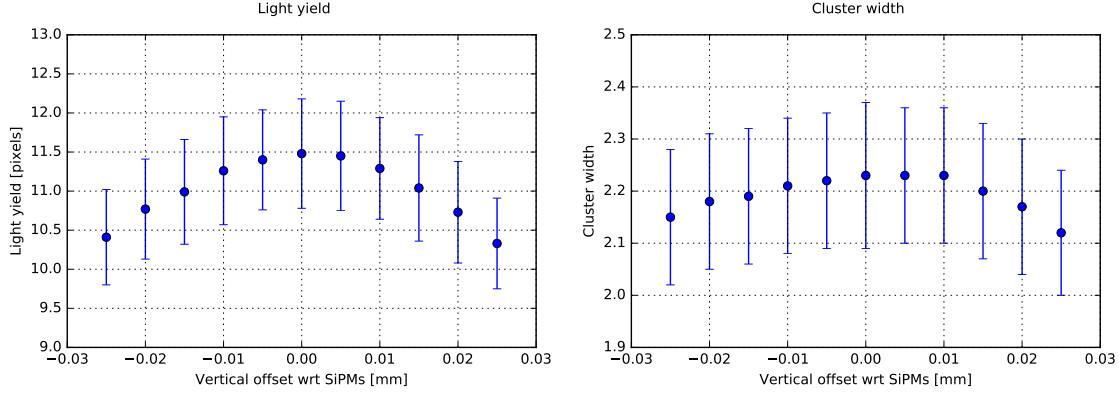


Figure 4.16 – Light yield (left) and cluster size (right) as a function of the vertical alignment between the SiPM array and the fibre mat. These scans have been performed on a test mat without mirror, for this reason, the absolute values on the y axes are not representative of the nominal ones observed in production mats.

- when a sufficiently-high data sample has been collected, the output is stored as a pdf summary sheet and as raw data.

Figure 4.17 shows part of the summary sheet for one of the mats produced at EPFL, containing the graph of the light yield per channel and per cluster as well as the cluster size. In addition to this, *housekeeping data* such as the temperature and bias voltages of the SiPMs are stored. It can be observed that some channels, on a periodic pattern, show a significantly lower-than-average yield. This is due to the 250- $\mu\text{m}$  gap present between the two 64-channels silicon dies form an SiPM and between neighbouring SiPMs.

In addition to the routine LY measurements, other checks have been performed using this setup. Moving the radioactive source along the positions A, B and C indicated in Figure 4.14 (D is not used as it is too much affected by the contribution from light produced in the fibre cladding), for example, allowed to study the attenuation length of the fibres, see Figure 4.18, which was found to be around 3 meters. Similarly, the straightness of the alignment pins has been checked on some mats by placing the source on a plastic support fitting a collimator. This rectangle-shaped support can be aligned with the pins simply by placing its long edge at contact with them; doing so, it was possible to check that the same channels are illuminated by the collimated source when the support is shifted along the length of the mat.

### 4.2.3 Results

Between September 2016 and July 2018, 500 scintillating fibre mats have been produced and tested at EPFL, meeting the scheduled goal. A rate of 8 wound mats per week has been reached in the last months of work. Along with the other production centres, around 10'000 km of fibre have been wound into  $\sim 1200$  mats, to which a number of spare units has to be added.

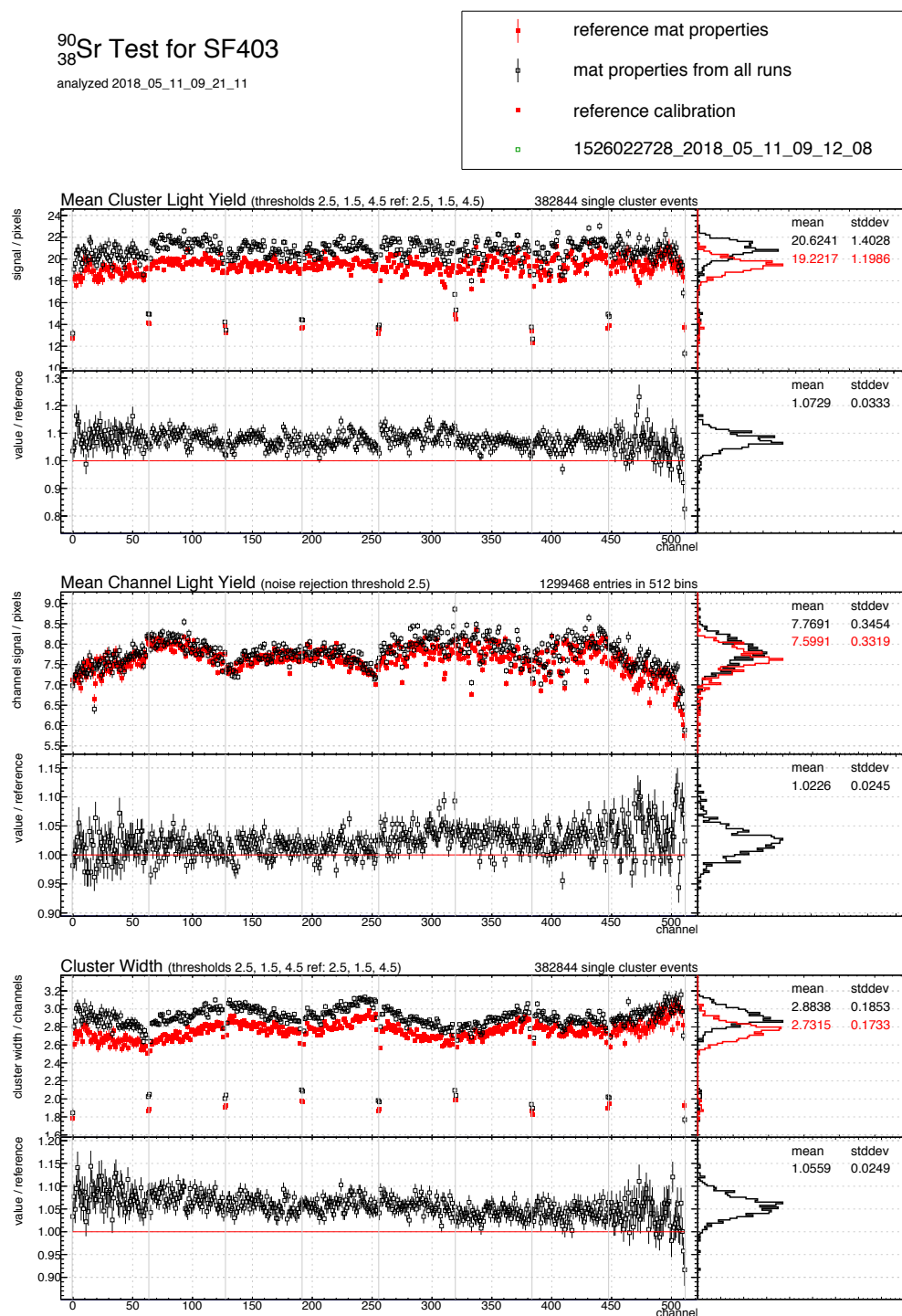


Figure 4.17 – Example of light yield test report, as saved in pdf format for each mat. The report shows, in black, the graph of the light yield per channel and per cluster as well as the cluster size. Each plot is compared to the analogous one from a reference mat, in red, to help immediately spotting unforeseen characteristics. In addition, for each of these graphs, the relative distribution is shown on the right.

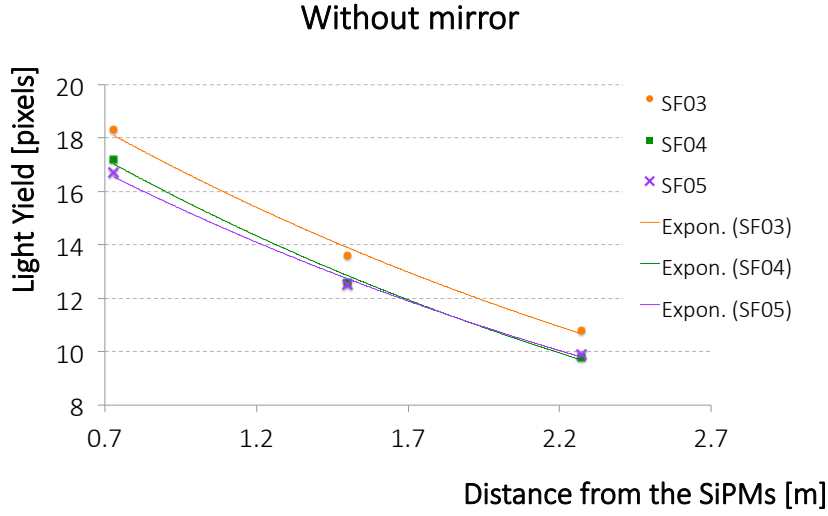


Figure 4.18 – Attenuation length on three of the early mats produced at EPFL.

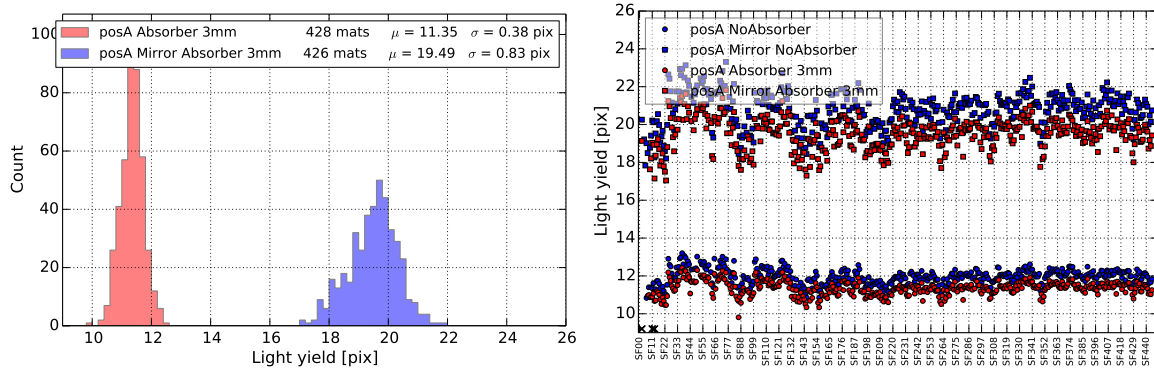


Figure 4.19 – Distribution of light yield of the scintillating fibre mats produced at EPFL (left) and scatter plot of fibre mat ID versus light yield (right). Data are reported both from tests performed with and without mirror.

The light yield test procedure has proven solid and effective. It has allowed to constantly monitor the quality of the production steps and of the material.

The mats produced at EPFL show an average LY of about 19.5 p.e. with mirror and 11.0 p.e. without mirror, in line with what observed in the other centres.

Figure 4.19 shows the distribution and the graph of the LY of most of the mats (defective test mats excluded), while Figure 4.20 shows the ratio between LY with and without mirror, which exhibits a rather narrow distribution, as expected from good-quality mats and clean mirror gluing. The LY values used in these plots are corrected for temperature fluctuations. The temperature of the SiPMs has indeed been observed to affect the measurement: the LY decreases linearly with the increasing temperature, with an angular coefficient of about  $-0.15 \text{ p.e./}^{\circ}\text{C}$ . It is interesting to notice how both graphs stabilise over time, due to the increasing acquired experience of the team involved in the production.



## 4.2. Light yield tests of fibre mats from EPFL

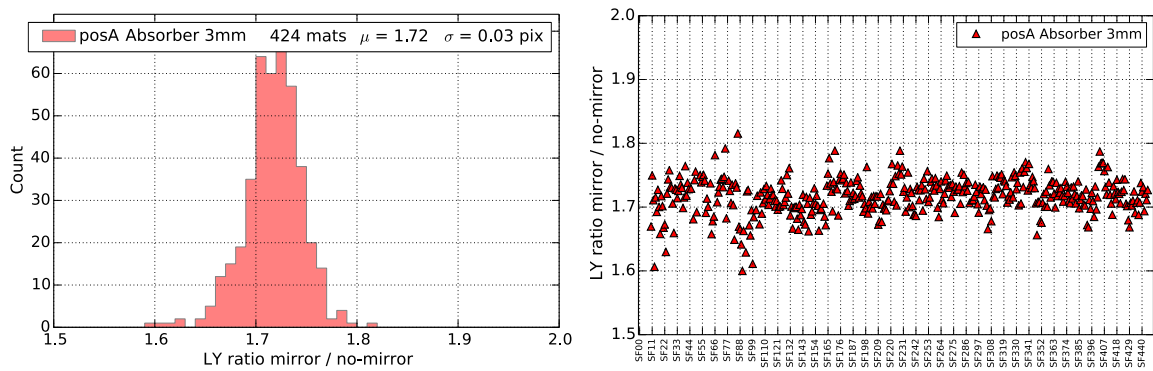


Figure 4.20 – Distribution (left) and scatter plot (right) of the ratio between the light yield with and without the mirror.



## 5 Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

As introduced in Chapter 2 of this thesis, the study of lepton-flavour-violating processes is a very promising way of testing the SM and probing for the existence of new mediators. Furthermore, the current experimental scenario (see review of flavour anomalies in Refs [34, 104]) demands to deeply investigate possible anomalies in the flavour sector.

The LHCb detector is optimised for the reconstruction of  $b$ -hadron decays. The search for the decay of the  $B$  meson into a couple of oppositely-charged, different-flavour leptons constitutes an optimal probe for LFV. Muons provide a very clean signature in the detector. The reconstruction of electrons is also possible but it must be corrected for bremsstrahlung radiation losses.  $\tau$  leptons are more difficult because only a partial reconstruction of their decay products is possible. This makes the  $B^0 \rightarrow e^\pm \mu^\mp$  and  $B_s^0 \rightarrow e^\pm \mu^\mp$  decays<sup>1</sup> almost ideal channels to be studied.

A search for these decays had already been performed in LHCb on  $1 \text{ fb}^{-1}$  of integrated luminosity, and published in 2013 [46]. The following limits at 95% confidence level were obtained on the signals branching fractions:

$$\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 1.4 \times 10^{-8} \quad (5.1)$$

$$\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 3.7 \times 10^{-9}, \quad (5.2)$$

which established the most stringent bounds to these decays at that time.

A new analysis has been performed in LHCb as a subject of this thesis, using the whole Run I data sample, amounting to  $3 \text{ fb}^{-1}$  and profiting from an improved reconstruction and selection strategy. A dedicated Boosted Decision Tree (BDT) has been developed to reject combinatorial background, optimised selection criteria are applied to reduce the contribution from mis-identified particles, and a more accurate approach has been followed for extracting the limit.

The present chapter contains a description of this new analysis, published in 2018 on the

---

<sup>1</sup>Inclusion of charge conjugate processes is implied throughout the text.

Journal of High Energy Physics [105].

The following pages contain the published article (the full list of authors is omitted and can be found in the reference), while deeper details on the most relevant aspects work on for this thesis are discussed in the next sections.

RECEIVED: October 12, 2017

REVISED: January 25, 2018

ACCEPTED: February 28, 2018

PUBLISHED: March 13, 2018

# Search for the lepton-flavour violating decays

$$B_{(s)}^0 \rightarrow e^\pm \mu^\mp$$



## The LHCb collaboration

*E-mail:* [flavio.archilli@cern.ch](mailto:flavio.archilli@cern.ch)

**ABSTRACT:** A search for the lepton-flavour violating decays  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  is performed based on a sample of proton-proton collision data corresponding to an integrated luminosity of  $3 \text{ fb}^{-1}$ , collected with the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. The observed yields are consistent with the background-only hypothesis. Upper limits on the branching fraction of the  $B_s^0 \rightarrow e^\pm \mu^\mp$  decays are evaluated both in the hypotheses of an amplitude completely dominated by the heavy eigenstate and by the light eigenstate. The results are  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 6.3 (5.4) \times 10^{-9}$  and  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 7.2 (6.0) \times 10^{-9}$  at 95% (90%) confidence level, respectively. The upper limit on the branching fraction of the  $B^0 \rightarrow e^\pm \mu^\mp$  decay is also evaluated, obtaining  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 1.3 (1.0) \times 10^{-9}$  at 95% (90%) confidence level. These are the strongest limits on these decays to date.

**KEYWORDS:** B physics, Branching fraction, Flavor physics, Hadron-Hadron scattering (experiments), Rare decay

**ARXIV EPRINT:** [1710.04111](https://arxiv.org/abs/1710.04111)

## Contents

|   |                              |    |
|---|------------------------------|----|
| 1 | Introduction                 | 1  |
| 2 | Detector and simulation      | 2  |
| 3 | Selection                    | 3  |
| 4 | BDT training and calibration | 4  |
| 5 | Normalisation                | 5  |
| 6 | Backgrounds                  | 7  |
| 7 | Mass calibration             | 8  |
| 8 | Results                      | 8  |
| 9 | Summary                      | 11 |
|   | The LHCb collaboration       | 15 |

---

## 1 Introduction

Processes that are suppressed or forbidden in the Standard Model (SM) are sensitive to potential contributions from new mediators, even if their masses are inaccessible to direct searches. Despite the fact that lepton-flavour violating (LFV) decays are forbidden within the SM, neutrino oscillation phenomena are proof that lepton flavour is not conserved in the neutral sector. However, LFV decays have not yet been observed, and their observation would be clear evidence of physics beyond the SM.

The study of LFV decays is particularly interesting in light of hints of lepton non-universality (LNU) effects in semileptonic decays [1] and  $b \rightarrow s\ell\ell$  transitions [2, 3], which could be associated with LFV processes [4]. Possible explanations of these hints can be found in various scenarios beyond the SM, e.g. models with a new gauge  $Z'$  boson [5] or leptoquarks [6, 7]. In these models, the branching fractions of the  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  decays<sup>1</sup> can be enhanced up to  $10^{-11}$ . Other models also predict possible enhancement for  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  decays, e.g. heavy singlet Dirac neutrinos [8], supersymmetric models [9] and the Pati-Salam model [10]. The most stringent published limits on the branching fractions of these decays are currently  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 1.4 \times 10^{-8}$  and  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 3.7 \times 10^{-9}$  at 95% confidence level (CL) from the LHCb collaboration using data corresponding to  $1 \text{ fb}^{-1}$  of integrated luminosity [11].

---

<sup>1</sup>Inclusion of charge conjugate processes is implied throughout the text.

This article presents an analysis performed on a larger data sample, corresponding to an integrated luminosity of  $3\text{ fb}^{-1}$  of  $pp$  collisions collected at centre-of-mass energies of 7 and 8 TeV by the LHCb experiment in 2011 and 2012. In addition to a larger data sample, this analysis benefits from an improved selection and in particular a better performing multivariate classifier for signal and background separation. It supersedes the previous LHCb search for  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  decays [11].

Two normalisation channels are used: the  $B^0 \rightarrow K^+ \pi^-$  decay which has a similar topology to that of the signal, and the  $B^+ \rightarrow J/\psi K^+$  decay, with  $J/\psi \rightarrow \mu^+ \mu^-$ , which has an abundant yield and a similar purity and trigger selection. To avoid potential biases,  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates in the signal region,  $m_{e^\pm \mu^\mp} \in [5100, 5500] \text{ MeV}/c^2$ , where  $m_{e^\pm \mu^\mp}$  is the invariant mass of the  $e^\pm \mu^\mp$  pair, were not examined until the selection and fitting procedure were finalised.

## 2 Detector and simulation

The LHCb detector [12, 13] is a single-arm forward spectrometer covering the pseudorapidity range  $2 < \eta < 5$ , designed for the study of particles containing  $b$  or  $c$  quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the  $pp$  interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of momentum,  $p$ , of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/ $c$ . The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of  $(15 + 29/p_T) \mu\text{m}$ , where  $p_T$  is the component of the momentum transverse to the beam, in GeV/ $c$ . Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the muon and calorimeter systems, followed by a software stage that applies a full reconstruction of the event. The  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates must fulfill the requirements of the electron or muon triggers. At the hardware stage, the electron trigger requires the presence of a cluster in the electromagnetic calorimeter with a transverse energy deposit,  $E_T$ , of at least 2.5 (3.0) GeV for 2011 (2012) data. The muon trigger selects muon candidates with  $p_T$  higher than 1.5 (1.8) GeV/ $c$  for 2011 (2012) data. The software stage requires a two-track secondary vertex identified by a multivariate algorithm [14] to be consistent with the decay of a  $b$  hadron with at least two charged tracks, and at least one track with high  $p_T$  and large IP with respect to any PV.

Simulated samples are used to evaluate geometrical, reconstruction and selection efficiencies for both signal and backgrounds, to train multivariate classifiers and to determine the shapes of invariant mass distributions of both signal and backgrounds. In the simula-

## Chapter 5. Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

tion,  $pp$  collisions are generated using PYTHIA [15] with a specific LHCb configuration [16]. Decays of hadronic particles are described by EVTGEN [17], in which final-state radiation is generated using PHOTOS [18]. The interaction of the generated particles with the detector, and its response, are simulated using the GEANT4 toolkit [19, 20] as described in ref. [21].

### 3 Selection

The  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates in the events passing the trigger selection are constructed by combining pairs of tracks producing good quality secondary vertices that are separated from any PV in the downstream direction by a flight distance greater than 15 times its uncertainty. Only  $B_{(s)}^0$  candidates with  $p_T > 0.5 \text{ GeV}/c$  and a small impact parameter  $\chi^2$ ,  $\chi_{\text{IP}}^2$ , are considered, where the  $\chi_{\text{IP}}^2$  of a  $B_{(s)}^0$  candidate is defined as the difference between the  $\chi^2$  of the PV reconstructed with and without the considered candidate. The PV with the smallest  $\chi_{\text{IP}}^2$  is associated to the  $B_{(s)}^0$  candidate. The measured momentum of electron candidates is corrected for the loss of momentum due to bremsstrahlung. This correction is made by adding to the electron the momentum of photons consistent with being emitted from the electron before the magnet [22]. Since bremsstrahlung can affect the kinematic distribution of  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates, the sample is split into two categories: candidates in which no photon is associated with the electron and candidates for which one or more photons are recovered. The fraction of electrons with recovered bremsstrahlung photons is about 60% for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  decays. Only  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates with  $m_{e^\pm \mu^\mp} \in [4900, 5850] \text{ MeV}/c^2$  are retained to be further analysed.

Particles forming the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates are required to be well identified as an electron and a muon [23], using information from the Cherenkov detectors, the calorimeters and the muon stations. These identification criteria are optimised to keep high signal efficiency while maximising the rejection power for the two-body hadronic  $B$  decays,  $B \rightarrow h^+ h'^-$ , which are the major peaking backgrounds.

In order to reduce combinatorial background — combinations of two random tracks that can be associated to a common vertex — a loose requirement on the response of a multivariate classifier trained on simulated events is applied to the signal candidates. This classifier takes the following geometrical variables as input: the direction of the  $B_{(s)}^0$  meson candidate; its impact parameter with respect to the assigned PV, defined as the PV with which it forms the smallest  $\chi_{\text{IP}}^2$ ; the separation between the two outgoing leptonic tracks at their point of closest approach; and the minimum IP of each lepton particle with respect to any PV. In total 22 020  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates are selected, which are mainly comprised of combinatorial background that is made up of true electrons and muons.

The normalisation channels are selected with requirements as similar as possible to those used for the signal. The selection for  $B^0 \rightarrow K^+ \pi^-$  candidates is the same as for the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  channel, except for the particle identification criteria which are changed into hadronic particle identification requirements. Similarly, the  $B^+ \rightarrow J/\psi K^+$  candidate selection is also kept as similar as possible, applying the same selection used for the signal to the dimuon pair from the  $J/\psi$ , except for the particle identification requirements. Additionally, loose quality requirements are applied on the  $B^+$  vertex and particle identification is



required on both muons. Finally, a  $60 \text{ MeV}/c^2$  mass window around the nominal  $J/\psi$  mass and the requirement  $1.4 < 1 + p_{J/\psi}/p_K < 20.0$  is used. The latter removes backgrounds that have a least one track that is misidentified and another that is not reconstructed, mainly  $B \rightarrow J/\psi \pi^+ X$ , where  $X$  can be one or more particles.

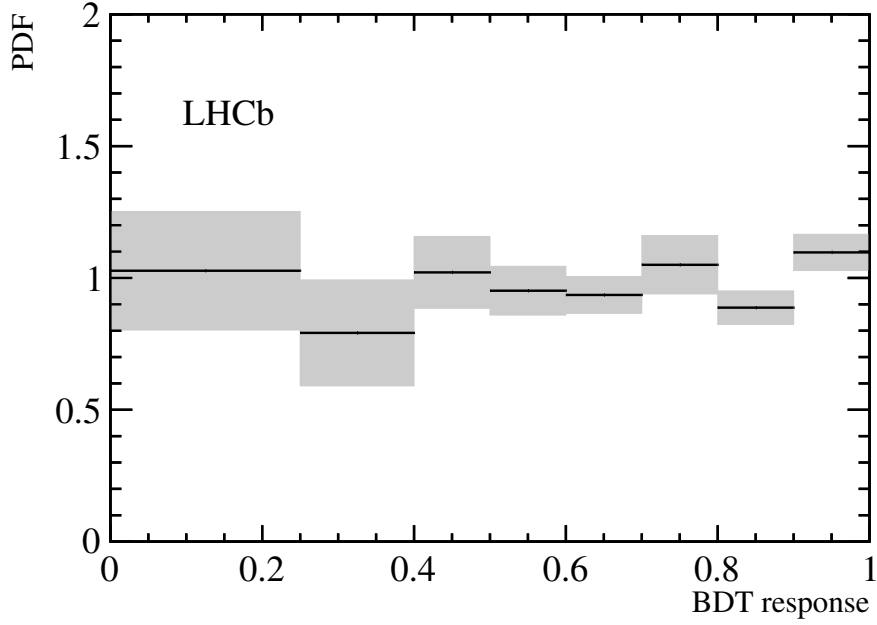
## 4 BDT training and calibration

A Boosted Decision Tree (BDT) classifier is used to separate the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  signal from the combinatorial background. The BDT is trained using a simulated sample of  $B_s^0 \rightarrow e^\pm \mu^\mp$  events to describe the signal and a data sample of same-sign  $e^\pm \mu^\pm$  candidates to describe the combinatorial background. The following input variables are used: the proper decay time of the  $B_{(s)}^0$  candidate; the minimum  $\chi_{\text{IP}}^2$  of the two leptons with respect to the assigned PV; the IP of the  $B_{(s)}^0$  candidate with respect to its PV; the distance of closest approach between the two lepton tracks; the degree of isolation of the two tracks with respect to the other tracks in the same event [24]; the transverse momentum of the  $B_{(s)}^0$  candidate; the cosine of the angle between the muon momentum in the  $B_{(s)}^0$  candidate rest frame and the vector perpendicular to the  $B_{(s)}^0$  candidate momentum and the beam axis; the flight distance of the  $B_{(s)}^0$  candidate with respect to its PV; the  $\chi^2$  of the  $B_{(s)}^0$  candidate decay vertex; the maximum transverse momentum of the two decay products and their difference in pseudorapidity.

The BDT response is transformed such that it is uniformly distributed in the range  $[0,1]$  for the signal, while peaking at zero for the background. The linear correlation between the BDT response and the dilepton invariant mass is found to be around 4%.

Since the BDT is trained using only kinematic information of a two-body  $B_{(s)}^0$  decay, its response is calibrated using  $B^0 \rightarrow K^+ \pi^-$  decays as a proxy. To avoid biases,  $B^0 \rightarrow K^+ \pi^-$  candidates are selected from candidates where the trigger decision did not depend on the presence of the  $B^0$  decay products. Furthermore, the candidates are weighted to emulate the effect of the lepton triggers and the particle identification requirements. The number of  $B^0 \rightarrow K^+ \pi^-$  candidates in bins of BDT response is determined by fitting the  $K^+ \pi^-$  invariant mass distribution. As expected, the BDT response is found to be consistent with a uniform distribution across the range  $[0,1]$ . The distribution of the BDT response is also checked on a  $B^0 \rightarrow K^+ \pi^-$  simulated sample and a uniform distribution is obtained. Candidates with a value smaller than 0.25 are then excluded, as this region is highly contaminated by background, leaving a total of 476 signal candidates. The signal candidates are classified in a binned two-dimensional space formed by the BDT response and the two bremsstrahlung categories. The expected probability density function (PDF) of the BDT response for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  decays with recovered bremsstrahlung photons is shown in figure 1.

Unrecovered bremsstrahlung photons emitted by signal electrons can affect the BDT response and are not accounted for in the calibration procedure since hadrons do not emit significant bremsstrahlung. The impact of bremsstrahlung on the BDT response distribution is evaluated using simulation and a correction is applied where no bremsstrahlung is recovered.



**Figure 1.** Expected distribution of the BDT response for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  decays with recovered bremsstrahlung photons obtained from the  $B^0 \rightarrow K^+ \pi^-$  control channel. The total uncertainty is shown as a light grey band. Each bin is normalised to its width.

## 5 Normalisation

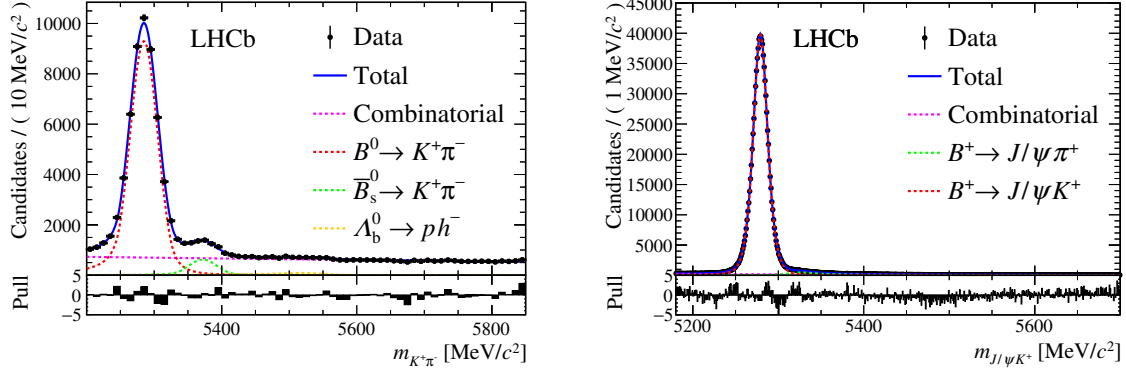
The  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  yields are obtained from a fit to the lepton-pair invariant mass distribution and translated into branching fractions according to

$$\begin{aligned} \mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) &= \sum_i w^i \frac{\mathcal{B}_{\text{norm}}^i}{N_{\text{norm}}^i} \frac{\varepsilon_{\text{norm}}^i}{\varepsilon_{\text{sig}}} \frac{f_q}{f_{d(s)}} \frac{\mathcal{L}_{\text{norm}}^i}{\mathcal{L}_{\text{sig}}} \times N_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp} \\ &= \alpha_{B_{(s)}^0} \times N_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}, \end{aligned} \quad (5.1)$$

where the index  $i$  identifies the normalisation channel and  $N_{\text{norm}}^i$  and  $\mathcal{B}_{\text{norm}}^i$  are its number of candidates and its branching fraction. The signal yields are denoted by  $N_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}$  and the factors  $f_q$  indicate the probabilities that a  $b$  quark fragments into a  $B^0$  or  $B_s^0$  meson. Assuming  $f_d = f_u$ , the fragmentation probability for the  $B^0$  and  $B^+$  channels is set to  $f_d$ . The value of  $f_s/f_d$  used is measured in  $pp$  collision data at  $\sqrt{s} = 7$  TeV by the LHCb collaboration and is evaluated to be  $0.259 \pm 0.015$  [25]. The two normalisation channels are averaged with weights  $w^i$  proportional to the square of the inverse of the uncertainty related to their branching fractions and yields. A correction has also been applied for the marginal difference in luminosity,  $\mathcal{L}$ , between the channels. The branching fractions of the signal decays include both charge configurations of the final-state particles,  $e^+ \mu^-$  and  $e^- \mu^+$ , so that  $\mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) \equiv \mathcal{B}(B_{(s)}^0 \rightarrow e^+ \mu^-) + \mathcal{B}(B_{(s)}^0 \rightarrow e^- \mu^+)$ . The results of the two fits are shown in figure 2 and the measured yields are reported in table 1.

|                              | Yield               |
|------------------------------|---------------------|
| $B^0 \rightarrow K^+\pi^-$   | $49\,907 \pm 277$   |
| $B^+ \rightarrow J/\psi K^+$ | $913\,074 \pm 1106$ |

**Table 1.** Yields of normalisation channels obtained from fits to data.



**Figure 2.** Invariant mass distributions of the two normalisation channels with fit functions superimposed: (left)  $B^0 \rightarrow K^+\pi^-$  and (right)  $B^+ \rightarrow J/\psi K^+$ . Pull distributions are shown below each plot.

The efficiency  $\varepsilon_{\text{sig(norm)}}$  for the signal (normalisation) channels depends on several factors: the geometric acceptance of the detector, the probability for particles to produce hits in the detector which can be reconstructed as tracks, and the efficiency of the selection requirements that are applied both in the trigger and selection stages, which includes the particle identification requirements. The ratios of acceptance, reconstruction and selection efficiencies are evaluated using simulation with the exception of the trigger and particle identification efficiencies, which are not well reproduced by simulation, and are calibrated using data [26, 27]. Calibration samples where the trigger decision was independent of the candidate decay products are used to study the trigger efficiency. From these samples,  $B^+ \rightarrow J/\psi K^+$  candidates, with  $J/\psi \rightarrow e^+e^-$  and  $J/\psi \rightarrow \mu^+\mu^-$ , are used to study the requirements for the electrons and muons, respectively. The efficiencies are determined as a function of the  $p_T$  and IP for the muon and  $E_T$  and IP for the electron. The single-track efficiencies are then combined with a weighted average over the properties of the electron and muon tracks of a  $B_s^0 \rightarrow e^\pm \mu^\mp$  simulated sample.

Particle identification efficiencies are evaluated using calibration samples where the identity of one of the particles can be inferred by means uncorrelated to particle identification requirements. A tag-and-probe method is applied on  $J/\psi \rightarrow \mu^+\mu^-$  and  $J/\psi \rightarrow e^+e^-$  decay samples, where only one lepton, the tag, is required to be well identified and the identity of the other lepton is deduced. The single-track efficiencies, calculated as a function of kinematic variables, are then combined and averaged using the momentum distributions of the leptons in a  $B_s^0 \rightarrow e^\pm \mu^\mp$  simulated sample.

The two normalisation factors  $\alpha_{B_s^0}$  and  $\alpha_{B^0}$  are determined to be  $(2.48 \pm 0.17) \times 10^{-10}$  and  $(6.16 \pm 0.23) \times 10^{-11}$ . The total efficiencies for the  $B^0 \rightarrow e^\pm \mu^\mp$ ,  $B_s^0 \rightarrow e^\pm \mu^\mp$ ,

## Chapter 5. Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

$B^+ \rightarrow J/\psi K^+$  and  $B^0 \rightarrow K^+ \pi^-$  decays are respectively  $(2.22 \pm 0.05)\%$ ,  $(2.29 \pm 0.05)\%$ ,  $(2.215 \pm 0.035)\%$  and  $(0.360 \pm 0.021)\%$ , where the efficiencies for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  are for the full BDT and bremsstrahlung category range.

To validate the normalisation procedure, the ratio between the measured branching fractions of  $B^0 \rightarrow K^+ \pi^-$  and  $B^+ \rightarrow J/\psi K^+$  is determined as

$$R_{\text{norm}} = \frac{N_{B^0 \rightarrow K^+ \pi^-} \times \varepsilon_{B^+ \rightarrow J/\psi K^+}}{N_{B^+ \rightarrow J/\psi K^+} \times \varepsilon_{B^0 \rightarrow K^+ \pi^-}} = 0.332 \pm 0.002 (\text{stat}) \pm 0.020 (\text{syst}), \quad (5.2)$$

where  $\varepsilon_{B^+ \rightarrow J/\psi K^+}$  and  $\varepsilon_{B^0 \rightarrow K^+ \pi^-}$  are the selection efficiencies for the  $B^0 \rightarrow K^+ \pi^-$  and  $B^+ \rightarrow J/\psi K^+$  decays respectively. A correction of about 1% is applied in order to take into account the difference in luminosity between the two channels. The value obtained for  $R_{\text{norm}}$  is in excellent agreement with the measured value of  $0.321 \pm 0.013$  [28].

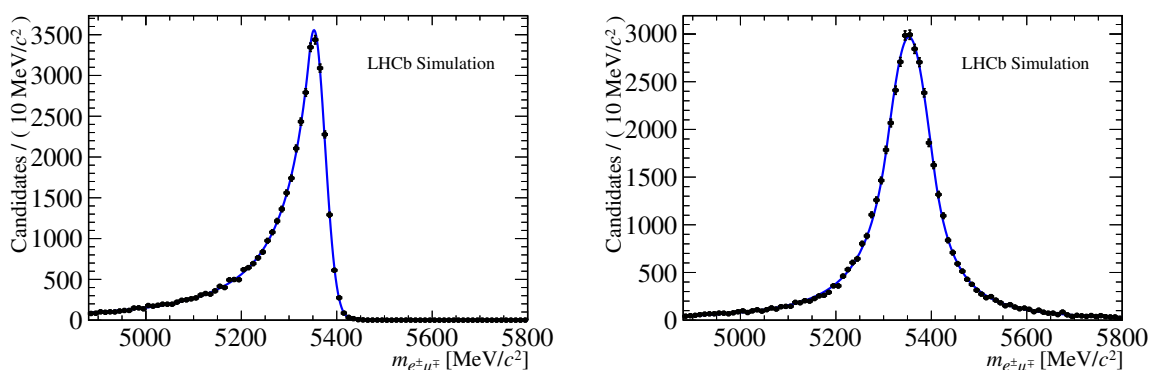
## 6 Backgrounds

In addition to the combinatorial background, the signal region is also potentially polluted by backgrounds from exclusive decays where one or more of the final-state particles are misidentified or not reconstructed. The potentially most dangerous of these backgrounds are hadronic  $B \rightarrow h^+ h'^-$  decays where both hadrons are misidentified as an electron-muon pair, resulting in peaking structures near the  $B_s^0 \rightarrow e^\pm \mu^\mp$  signal mass. Other decays which could contribute, especially at low invariant masses, are  $B_c^+ \rightarrow J/\psi \ell'^+ \nu_{\ell'}$  with  $J/\psi \rightarrow \ell^+ \ell^-$ ,  $B^0 \rightarrow \pi^- \ell^+ \nu_{\ell}$ ,  $\Lambda_b^0 \rightarrow p \ell^- \bar{\nu}_{\ell}$  and  $B^+ \rightarrow \pi^+ J/\psi$  with  $J/\psi \rightarrow \ell^+ \ell^-$ , where  $\ell/\ell'^\pm = e^\pm$  or  $\mu^\pm$ . These decays do not peak under the signal but are potentially abundant. The expected number of candidates from each possible background decay that pass the signal selection is evaluated using simulation. The candidates are normalised to the number of  $B^+ \rightarrow J/\psi K^+$  decays found in data as

$$N_X = N_{B^+ \rightarrow J/\psi K^+} \frac{f_q}{f_u} \frac{\mathcal{B}(X)}{\mathcal{B}(B^+ \rightarrow J/\psi K^+) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)} \frac{\varepsilon(X)}{\varepsilon(B^+ \rightarrow J/\psi K^+)}, \quad (6.1)$$

where  $N_X$  is the expected number of candidates from the  $X$  decay that fall into the  $B_s^0 \rightarrow e^\pm \mu^\mp$  signal mass window;  $f_q$  is the fragmentation fraction;  $\mathcal{B}(X)$ ,  $\mathcal{B}(B^+ \rightarrow J/\psi K^+)$  and  $\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$  are respectively the branching fractions of the decay under study,  $B^+ \rightarrow J/\psi K^+$  and  $J/\psi \rightarrow \mu^+ \mu^-$  [28];  $\varepsilon(X)$  is the efficiency for each considered decay to pass the  $B_s^0 \rightarrow e^\pm \mu^\mp$  selection; and  $\varepsilon(B^+ \rightarrow J/\psi K^+)$  is the efficiency for  $B^+ \rightarrow J/\psi K^+$  candidates to pass the respective selection.

The mass and BDT distributions of these background modes are evaluated using simulated samples, while the probabilities of misidentifying kaons, pions and protons as muons or electrons are determined from  $D^{*+} \rightarrow D^0 \pi^+$  with  $D^0 \rightarrow K^- \pi^+$  and  $\Lambda \rightarrow p \pi^-$  decays selected from data. The expected total number of  $B \rightarrow h^+ h'^-$  candidates is  $0.11 \pm 0.02$  in the full BDT range, which is negligible. This yield estimation is cross-checked using data. A sample of  $B \rightarrow h^+ h'^-$  decays is selected by applying only a partial  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  selection: only the signal electron PID requirements are applied while the second particle is required to be identified as a pion. The application of these criteria still leaves a sizeable peak to



**Figure 3.** Distribution of the  $m_{e^\pm\mu^\mp}$  invariant mass of simulated  $B_s^0$  candidates with no (left) and one or more (right) recovered bremsstrahlung photons. The overlaid fit function is a modified Crystal Ball function with two tails on opposite sides.

be fit in data. The yield of decays identified as  $B_{(s)}^0 \rightarrow e^\pm\pi^\mp$  is then modified to take into account the probability of a pion to be misidentified as a muon. After this correction the expected yield is compatible with the yield obtained using the simulation.

The expected yields of most of the other backgrounds are also found to be negligible. The only backgrounds which are relevant are  $B^0 \rightarrow \pi^-\mu^+\nu_\ell$  and  $\Lambda_b^0 \rightarrow p\ell^-\bar{\nu}_\ell$  for which  $55 \pm 3$  and  $82 \pm 39$  candidates, respectively, are expected in the full BDT range. The contributions from these two decays are included in the fit model.

## 7 Mass calibration

The invariant-mass distribution of  $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$  candidates is modelled by a modified Crystal Ball function [29] with two tails on opposite sides defined by two parameters each. The signal shape parameters are obtained from simulation, with data-driven scale factors applied to the core resolution to correct for possible data-simulation discrepancies. For this purpose, since there is no appropriate control channel with an electron and a muon in the final state,  $J/\psi \rightarrow e^+e^-$  and  $J/\psi \rightarrow \mu^+\mu^-$  decays are analysed comparing the mass resolution in data and simulation. The results are then combined to reproduce the effect on an  $e^\pm\mu^\mp$  final state. Corrections to the widths of the mass are of the order of 10%. Since bremsstrahlung can significantly alter the mass shape by enhancing the tails, the fit model for  $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$  candidates is obtained separately for the two bremsstrahlung categories (see figure 3). The mass shape parameters are found to be independent of the particular BDT bin chosen and a single model for each bremsstrahlung category is therefore used.

## 8 Results

The data sample is split into two bremsstrahlung categories, which are further divided into seven subsets each depending on the BDT response covering the range from 0.25 to 1.0, with boundaries 0.25, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0. The region with BDT response

## Chapter 5. Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

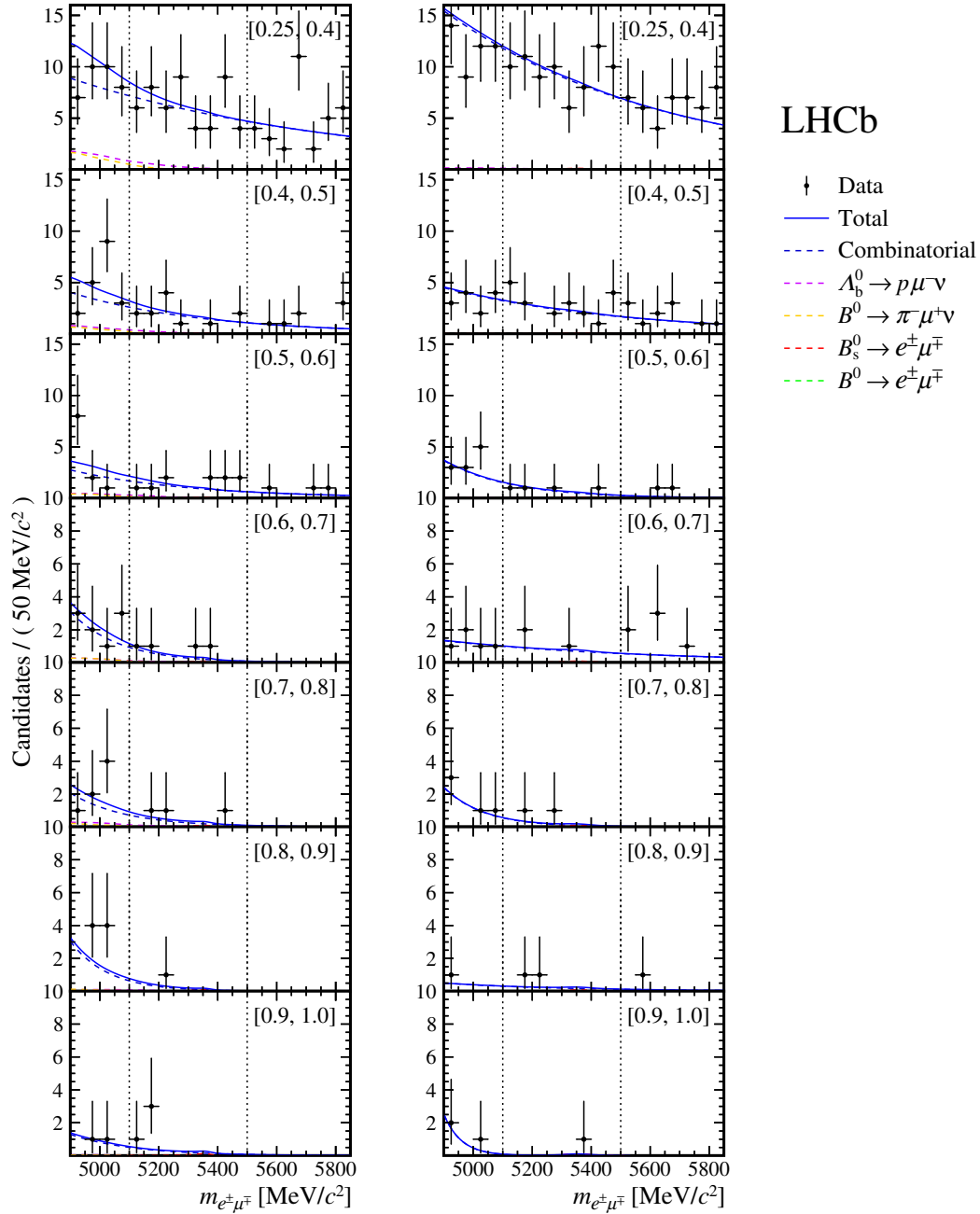
| channel                                        | expected                   | observed                   |
|------------------------------------------------|----------------------------|----------------------------|
| $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$ | $5.0 (3.9) \times 10^{-9}$ | $6.3 (5.4) \times 10^{-9}$ |
| $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp)$   | $1.2 (0.9) \times 10^{-9}$ | $1.3 (1.0) \times 10^{-9}$ |

**Table 2.** Expected (assuming no signal) and observed upper limits for  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$  and  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp)$  at 95% (90%) CL. The upper limit on the  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$  is evaluated under the assumption of pure heavy eigenstate contribution on the decay amplitude.

lower than 0.25, which is mostly populated by combinatorial background, is excluded from the fit. The  $B^0 \rightarrow e^\pm \mu^\mp$  and  $B_s^0 \rightarrow e^\pm \mu^\mp$  yields are obtained from a single unbinned extended maximum likelihood fit performed simultaneously to the  $m_{e^\pm \mu^\mp}$  distributions in each subset. The  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  fractional yields and the mass shape parameters in each category are Gaussian-constrained according to their expected values and uncertainties. The combinatorial background is modelled with an exponential function with independent yield and shape parameters in each subset. The exclusive backgrounds are included as separate components in the fit. Their mass shapes are modelled using nonparametric functions determined from simulation for each bremsstrahlung category. The overall yields and fractions of these backgrounds are Gaussian-constrained to their expected values. The result of this fit is shown in figure 4.

No significant excess of  $B^0 \rightarrow e^\pm \mu^\mp$  or  $B_s^0 \rightarrow e^\pm \mu^\mp$  decays is observed and upper limits on the branching fractions are set using the CL<sub>s</sub> method [30]. The ratio between the likelihoods in two hypotheses, signal plus background and background only, is used as the test statistic. The likelihoods are computed with nuisance parameters fixed to their nominal values. Pseudoexperiments, in which the nuisance parameters are varied according to their statistical and systematic uncertainties, are used for the evaluation of the test statistic. The resulting CL<sub>s</sub> scans are shown in figure 5 and upper limits at 95% and 90% confidence level are reported in table 2.

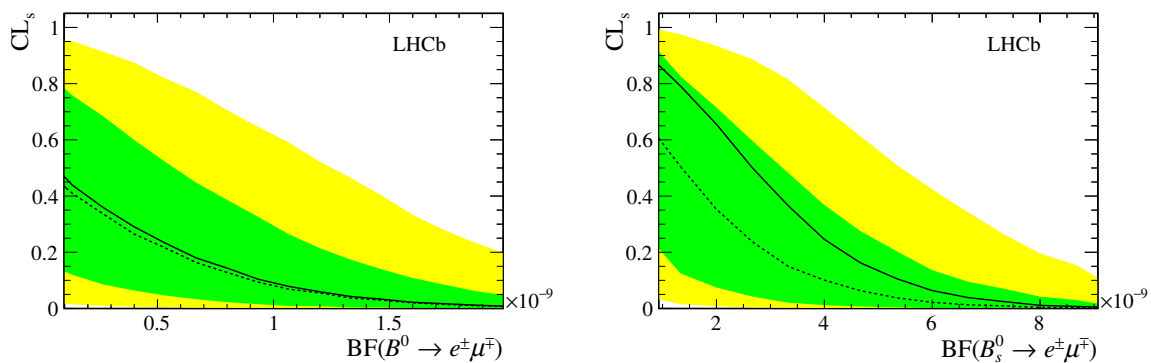
Several systematic uncertainties can affect the evaluation of the limit on the  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  branching fractions through the normalisation formula in eq. (5.1) and the fit model used to evaluate the signal yields. The systematic uncertainties are taken into account for the limit computation by constraining the respective nuisance parameters in the likelihood fit with a Gaussian distribution having the central value of the parameter as the mean and its uncertainty as the width. The nuisance parameters for the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  yields are related to the calibration of the BDT response, the parameters of the signal shape, the estimated yields of the  $B^0 \rightarrow \pi^- \mu^+ \nu_\ell$  and  $\Lambda_b^0 \rightarrow p \ell^- \bar{\nu}_\ell$  backgrounds and the fractional yield per bremsstrahlung category. For the limit on the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  branching fractions, the nuisance parameters are in addition related to the signal efficiency, whose uncertainty is dominated by the systematic uncertainty on the trigger efficiencies, and the uncertainties on the efficiencies, branching fractions and yields of the normalisation channels. For the  $B_s^0 \rightarrow e^\pm \mu^\mp$  branching fraction estimation, eq. (5.1) also includes the hadronisation fraction  $f_s/f_d$ , which dominates the systematic uncertainty for the normalisation. The overall impact on the limits is evaluated to be below 5%.



**Figure 4.** Distributions of the invariant mass of the  $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$  candidates,  $m_{e^\pm\mu^\mp}$ , divided into bins of BDT response and two bremsstrahlung categories (left) without and (right) with bremsstrahlung photons recovered. The result of the fit is overlaid and the different components are detailed. The edges of the range that was examined only after finalising the selection and fit procedure are delimited by gray dashed vertical lines. This region includes 90% of the potential signal candidates. Given the result obtained from the fit, the  $B^0 \rightarrow e^\pm\mu^\mp$  component is not visible in the plots.



## Chapter 5. Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$



**Figure 5.** Results of the  $CL_s$  scan used to obtain the limit on (left)  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp)$  and (right)  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$ . The background-only expectation is shown by the dashed line and the  $1\sigma$  and  $2\sigma$  bands are shown as dark (green) and light (yellow) bands respectively. The observed limit is shown as the solid black line.

The two  $B_s^0$  mass eigenstates are characterised by a large lifetime difference. Depending on their contribution to the decay amplitude, the selection efficiency and the BDT shape can be affected. Given the negligible difference in lifetime for the  $B^0$  system, this effect is not taken into account for the  $B^0 \rightarrow e^\pm \mu^\mp$  limit evaluation. Two extreme cases can be distinguished: when only the heavy or the light eigenstate contributes to the total decay amplitude. For example, if the only contribution to the LFV  $B_s^0 \rightarrow e^\pm \mu^\mp$  decay is due to neutrino oscillations, it is expected that the amplitude is dominated by the heavy eigenstate as for the  $B_s^0 \rightarrow \mu^+ \mu^-$  decay [24]. As the contribution to the total amplitude from the heavy and light eigenstate can have an effect on the acceptance, the limit on  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$  is evaluated in the two extreme cases. The one reported in table 2 and obtained from the  $CL_s$  scan in figure 5, is evaluated assuming only a contribution from the heavy eigenstate. For the light eigenstate case the limit is found to be  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 7.2 (6.0) \times 10^{-9}$  at 95% (90%) CL.

## 9 Summary

In summary, a search for the LFV decays  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  is performed using  $pp$  collision data collected at centre-of-mass energies of 7 and 8 TeV, corresponding to a total integrated luminosity of  $3 \text{ fb}^{-1}$ . No excesses are observed for these two modes and upper limits on the branching fractions are set to  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 6.3 (5.4) \times 10^{-9}$  and  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 1.3 (1.0) \times 10^{-9}$  at 95% (90%) CL, where only a contribution from the heavy eigenstate is assumed for the  $B_s^0$  meson. If the  $B_s^0$  amplitude is completely dominated by the light eigenstate, the upper limit on the branching fraction becomes  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 7.2 (6.0) \times 10^{-9}$  at 95% (90%) CL. These results represent the best upper limits to date and are a factor 2 to 3 better than the previous results from LHCb [11].

## Acknowledgments

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff



at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (U.S.A.). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (U.S.A.). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), ANR, Labex P2IO, ENIGMASS and OCEVU, and Région Auvergne-Rhône-Alpes (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, the Royal Society, the English-Speaking Union and the Leverhulme Trust (United Kingdom).

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] LHCb collaboration, *Measurement of the ratio of branching fractions  $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$* , *Phys. Rev. Lett.* **115** (2015) 111803 [[arXiv:1506.08614](https://arxiv.org/abs/1506.08614)] [[INSPIRE](#)].
- [2] LHCb collaboration, *Test of lepton universality with  $B^0 \rightarrow K^{*0}\ell^+\ell^-$  decays*, *JHEP* **08** (2017) 055 [[arXiv:1705.05802](https://arxiv.org/abs/1705.05802)] [[INSPIRE](#)].
- [3] LHCb collaboration, *Test of lepton universality using  $B^+ \rightarrow K^+\ell^+\ell^-$  decays*, *Phys. Rev. Lett.* **113** (2014) 151601 [[arXiv:1406.6482](https://arxiv.org/abs/1406.6482)] [[INSPIRE](#)].
- [4] D. Guadagnoli, *Flavor anomalies on the eve of the run-2 verdict*, *Mod. Phys. Lett. A* **32** (2017) 1730006 [[arXiv:1703.02804](https://arxiv.org/abs/1703.02804)] [[INSPIRE](#)].
- [5] A. Crivellin, L. Hofer, J. Matias, U. Nierste, S. Pokorski and J. Rosiek, *Lepton-flavour violating  $B$  decays in generic  $Z'$  models*, *Phys. Rev. D* **92** (2015) 054013 [[arXiv:1504.07928](https://arxiv.org/abs/1504.07928)] [[INSPIRE](#)].
- [6] D. Bečirević, S. Fajfer, N. Košnik and O. Sumensari, *Leptoquark model to explain the  $B$ -physics anomalies,  $R_K$  and  $R_D$* , *Phys. Rev. D* **94** (2016) 115021 [[arXiv:1608.08501](https://arxiv.org/abs/1608.08501)] [[INSPIRE](#)].
- [7] I. de Medeiros Varzielas and G. Hiller, *Clues for flavor from rare lepton and quark decays*, *JHEP* **06** (2015) 072 [[arXiv:1503.01084](https://arxiv.org/abs/1503.01084)] [[INSPIRE](#)].
- [8] A. Ilakovac, *Lepton flavor violation in the Standard Model extended by heavy singlet Dirac neutrinos*, *Phys. Rev. D* **62** (2000) 036010 [[hep-ph/9910213](https://arxiv.org/abs/hep-ph/9910213)] [[INSPIRE](#)].

## Chapter 5. Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

- [9] R.A. Diaz, R. Martinez and C.E. Sandoval, *Improving bounds on flavor changing vertices in the two Higgs doublet model from  $B^0$ - $\bar{B}^0$  mixing*, *Eur. Phys. J. C* **46** (2006) 403 [[hep-ph/0509194](#)] [[INSPIRE](#)].
- [10] J.C. Pati and A. Salam, *Lepton number as the fourth color*, *Phys. Rev. D* **10** (1974) 275 [*Erratum ibid.* **D 11** (1975) 703] [[INSPIRE](#)].
- [11] LHCb collaboration, *Search for the lepton-flavor violating decays  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$* , *Phys. Rev. Lett.* **111** (2013) 141801 [[arXiv:1307.4889](#)] [[INSPIRE](#)].
- [12] LHCb collaboration, *The LHCb detector at the LHC*, 2008 *JINST* **3** S08005 [[INSPIRE](#)].
- [13] LHCb collaboration, *LHCb detector performance*, *Int. J. Mod. Phys. A* **30** (2015) 1530022 [[arXiv:1412.6352](#)] [[INSPIRE](#)].
- [14] V.V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, 2013 *JINST* **8** P02013 [[arXiv:1210.6861](#)] [[INSPIRE](#)].
- [15] T. Sjöstrand, S. Mrenna and P.Z. Skands, *A brief introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [[arXiv:0710.3820](#)] [[INSPIRE](#)].
- [16] LHCb collaboration, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, *J. Phys. Conf. Ser.* **331** (2011) 032047 [[INSPIRE](#)].
- [17] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 [[INSPIRE](#)].
- [18] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, *Eur. Phys. J. C* **45** (2006) 97 [[hep-ph/0506026](#)] [[INSPIRE](#)].
- [19] GEANT4 collaboration, J. Allison et al., *GEANT4 developments and applications*, *IEEE Trans. Nucl. Sci.* **53** (2006) 270 [[INSPIRE](#)].
- [20] GEANT4 collaboration, S. Agostinelli et al., *GEANT4: a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250 [[INSPIRE](#)].
- [21] LHCb collaboration, *The LHCb simulation application, Gauss: design, evolution and experience*, *J. Phys. Conf. Ser.* **331** (2011) 032023 [[INSPIRE](#)].
- [22] LHCb collaboration, *Measurement of the  $B^0 \rightarrow K^{*0} e^+ e^-$  branching fraction at low dilepton mass*, *JHEP* **05** (2013) 159 [[arXiv:1304.3035](#)] [[INSPIRE](#)].
- [23] F. Archilli et al., *Performance of the muon identification at LHCb*, 2013 *JINST* **8** P10020 [[arXiv:1306.0249](#)] [[INSPIRE](#)].
- [24] LHCb collaboration, *Measurement of the  $B_s^0 \rightarrow \mu^+ \mu^-$  branching fraction and effective lifetime and search for  $B^0 \rightarrow \mu^+ \mu^-$  decays*, *Phys. Rev. Lett.* **118** (2017) 191801 [[arXiv:1703.05747](#)] [[INSPIRE](#)].
- [25] LHCb collaboration, *Measurement of the fragmentation fraction ratio  $f_s/f_d$  and its dependence on B meson kinematics*, *JHEP* **04** (2013) 001 [[arXiv:1301.5286](#)] [[LHCb-CONF-2013-011](#)] [[INSPIRE](#)].
- [26] S. Tolk, J. Albrecht, F. Dettori and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, [LHCb-PUB-2014-039](#), CERN, Geneva Switzerland, (2014).
- [27] L. Anderlini et al., *The PIDCalib package*, [LHCb-PUB-2016-021](#), CERN, Geneva Switzerland, (2016).

- 
- [28] PARTICLE DATA GROUP collaboration, C. Patrignani et al., *Review of particle physics*, *Chin. Phys. C* **40** (2016) 100001 [[INSPIRE](#)].
- [29] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-prime and Upsilon resonances*, Ph.D. thesis, Institute of Nuclear Physics, Krakow Poland, (1986) [[INSPIRE](#)].
- [30] A.L. Read, *Presentation of search results: the  $CL_s$  technique*, *J. Phys. G* **28** (2002) 2693 [[INSPIRE](#)].

## 5.1 Selection

Selecting candidates efficiently while keeping a high rejection power on all possible backgrounds is a delicate task, which becomes even more challenging in an analysis where the searched signal has never been observed before and has different characteristics from any SM process. Relying completely on simulated events is not a viable option since some aspects of the detector response cannot be perfectly reproduced. This section provides details on the strategies that have been adopted.

A loose selection is applied in the stripping, and it is described in Section 5.1.2. This is followed by a first multivariate classifier, described in Section 5.1.3, that reduces the data sample to a manageable level while maintaining a high signal efficiency. After this, a second BDT is applied and used to bin the data. Furthermore, trigger and PID cuts are applied to enhance the signal efficiency while removing background from other decays such as  $B_{(s)}^0 \rightarrow h^+ h^-$ .

### 5.1.1 Trigger

$$B_{(s)}^0 \rightarrow e^\pm \mu^\mp: \text{L0} \otimes \text{HLT1}$$

At L0 level (defined in Section 3.5), events are selected with the requirement of having passed the L0Muon or L0Electron lines. The former requires a transverse momentum  $p_T$  larger than 1480 MeV/ $c$  and nSPDHits < 600. The latter consists mainly of two requirements: the transverse energy  $E_T$  of the track must be larger than 2.5 GeV and the number of hits in the SPD, nSPDHits, must not exceed 600.

For HLT1 the aim is to trigger on displaced muon tracks. For this purpose Hlt1TrackMuon is used on the muon. Of top of that Hlt1TrackAllL0 is also used both for the muon and the electron. These lines consist mainly of requirements on the momentum and transverse momentum of the tracks:

- $p > 8.0 \text{ GeV}/c$  and  $p_T > 1.0 \text{ GeV}/c$  in Hlt1TrackMuon;
- $p > 10.0 \text{ GeV}/c$  and  $p_T > 1.7 \text{ GeV}/c$  in Hlt1TrackAllL0;

along with fiducial cuts on the number of detector hits and on the impact parameter  $\chi^2$  of a particle with respect to its origin vertex,  $\chi_{\text{IP}}^2$ .

$$B_{(s)}^0 \rightarrow e^\pm \mu^\mp: \text{HLT2}$$

The HLT2 strategy is dictated by the need to use the normalisation channel  $B^0 \rightarrow K^+ \pi^-$ . To guarantee an overlap between this channel and  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ , a topological trigger is used, selecting displaced two-body decays with the kinematics of a  $B$ . For reference, the names of the lines that have been used are: Hlt2Topo2BodyBBDTDecision\_TOS, Hlt2TopoMu2BodyBBDTDecision\_TOS

and Hlt2B2HHDecision\_TOS.

$$B_{(s)}^0 \rightarrow e^\pm \mu^\mp: \text{L0} \otimes \text{HLT1} \otimes \text{HLT2}$$

The full list of trigger requirements used to select  $B_s^0 \rightarrow e^\pm \mu^\mp$  signal candidates is reported in Table 5.1.

Table 5.1 – Trigger selection for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ .

|                               | <b>Muon</b>                                             | <b>Electron</b> |
|-------------------------------|---------------------------------------------------------|-----------------|
| L0                            | L0Muon                                                  | L0Electron      |
| HLT1                          | Hlt1TrackAllL0<br>or Hlt1TrackMuon                      | Hlt1TrackAllL0  |
| <b><math>B_{(s)}^0</math></b> |                                                         |                 |
| L0 $\otimes$ HLT1             | Muon or ( Electron and not Muon )                       |                 |
| HLT2                          | Hlt2Topo2BodyBBDT or Hlt2TopoMu2BodyBBDT<br>or Hlt2B2HH |                 |

### Trigger for normalisation channels

The normalisation channels are selected with criteria as similar as possible to those applied on signal, in order to reduce possible systematic effects introduced by different cuts. For reference, the trigger selection that has been used for  $B^0 \rightarrow K^+ \pi^-$  is

L0HadronDecisionTOS & Hlt1TrackAllL0DecisionTOS & Hlt2B2HHDecisionTOS,

while the one for  $B^+ \rightarrow J/\psi(\mu^+ \mu^-) K^+$  is shown in Table 5.2.

Table 5.2 – Trigger selection for  $B^+ \rightarrow J/\psi(\mu^+ \mu^-) K^+$ .

|                         | <b>Muons</b>                                                                            | <b>Kaons</b>   |
|-------------------------|-----------------------------------------------------------------------------------------|----------------|
| L0                      | L0Muon                                                                                  |                |
| HLT1                    | Hlt1TrackAllL0<br>or Hlt1TrackMuon                                                      | Hlt1TrackAllL0 |
| <b><math>B^+</math></b> |                                                                                         |                |
| HLT2                    | Hlt2Topo2BodyBBDT or Hlt2TopoMu2BodyBBDT<br>or Hlt2Topo3BodyBBDT or Hlt2TopoMu3BodyBBDT |                |

### 5.1.2 Stripping and pre-selection

Separate algorithms are used to select  $B_s^0 \rightarrow e^\pm \mu^\mp$ , same-sign (SS)  $B_s^0 \rightarrow e^\pm \mu^\pm$ ,  $B^0 \rightarrow K^+ \pi^-$  and  $B^+ \rightarrow J/\psi K^+$  where the requirements are kept as similar as possible. For  $B^0 \rightarrow K^+ \pi^-$ ,

in order to ensure that the selection is similar to  $B_s^0 \rightarrow e^\pm \mu^\mp$ , the hadrons are required to be within the muon acceptance. Additional cuts are then applied on top of the stripping to further reduce the size of the samples. The full list of stripping requirements is reported in Tables 5.3, 5.4 and 5.5, where DOCA is the distance of closest approach between the two tracks, VDS is the B meson flight distance significance, ghost prob is the probability of a track not being associated to any real charged particle. The BDTS discriminant is described in Section 5.1.3.

Table 5.3 – Selection for  $B_s^0 \rightarrow e^\pm \mu^\mp$ .

| Cut                       | applied to  | value                                    |
|---------------------------|-------------|------------------------------------------|
| track $\chi^2/\text{ndf}$ | $\mu / e$   | $< 3$                                    |
| ghost prob                | $\mu / e$   | $< 0.3$                                  |
| DOCA                      |             | $< 0.3 \text{ mm}$                       |
| $\chi_{\text{IP}}^2$      | $\mu / e$   | $> 25$                                   |
| $p_{\text{T}}$            | $\mu / e$   | $> 0.25 \text{ and } < 40 \text{ GeV}/c$ |
| $p$                       | $\mu / e$   | $< 500 \text{ GeV}/c$                    |
| IsMuon                    | $\mu$ only  | true                                     |
| $\chi_{\text{vtx}}^2$     | $B_{(s)}^0$ | $< 9$                                    |
| VDS                       |             | $> 15$                                   |
| $\chi_{\text{IP}}^2$      | $B_{(s)}^0$ | $< 25$                                   |
| $t$                       |             | $< 9 \cdot \tau_{PDG}(B_s^0)$            |
| BDTS                      |             | $> 0.05$                                 |
| $p_{\text{T}}(B_s^0)$     | $B_{(s)}^0$ | $> 0.5 \text{ GeV}/c$                    |

Table 5.4 – Selection for  $B^0 \rightarrow K^+ \pi^-$ .

| Cut                       | applied to  | value                                    |
|---------------------------|-------------|------------------------------------------|
| track $\chi^2/\text{ndf}$ | $h^+ / h^-$ | $< 3 (< 4)$                              |
| ghost prob                | $h^+ / h^-$ | $< 0.3$                                  |
| DOCA                      |             | $< 0.3 \text{ mm}$                       |
| $\chi_{\text{IP}}^2$      | $h^+ / h^-$ | $> 25$                                   |
| $p_{\text{T}}$            | $h^+ / h^-$ | $> 0.25 \text{ and } < 40 \text{ GeV}/c$ |
| $p$                       | $h^+ / h^-$ | $< 500 \text{ GeV}/c$                    |
| InMuonAcc                 | $h^+ / h^-$ | true                                     |
| $\chi_{\text{vtx}}^2$     | $B_{(s)}^0$ | $< 9$                                    |
| VDS                       |             | $> 15$                                   |
| $\chi_{\text{IP}}^2$      | $B_{(s)}^0$ | $< 25$                                   |
| $t$                       | $B_{(s)}^0$ | $< 9 \cdot \tau_{PDG}(B_s^0)$            |
| BDTS                      |             | $> 0.05$                                 |
| $p_{\text{T}}(B_s^0)$     | $B_{(s)}^0$ | $> 0.5 \text{ GeV}/c$                    |

### 5.1.3 The BDTS discriminant

A requirement on the response of a BDT multivariate classifier called BDTS is used to further reduce the size of the background sample before the final BDT classifier is trained. The variables entering the BDTS are:

Table 5.5 – Selection for  $B^+ \rightarrow J/\psi K^+$ .

| Cut                       | applied to | value                                           |
|---------------------------|------------|-------------------------------------------------|
| track $\chi^2/\text{ndf}$ | $\mu / h$  | $< 3$                                           |
| ghost prob                | $\mu$ only | $< 0.3$                                         |
| DOCA                      |            | $< 0.3\text{mm}$                                |
| $\chi_{\text{IP}}^2$      | $\mu$ only | $> 25$                                          |
| $p_T$                     | $\mu$ only | $> 0.25$ and $< 40 \text{ GeV}/c$               |
| $p$                       | $\mu$ only | $< 500 \text{ GeV}/c$                           |
| IsMuon                    | $\mu$ only | true                                            |
| $\chi_{\text{vtx}}^2$     | $J/\psi$   | $< 9$                                           |
| VDS                       |            | $> 15$                                          |
| $\Delta M$                |            | $ M(\mu\mu) - m_{J/\psi}  < 60 \text{ MeV}/c^2$ |
| $\chi_{\text{IP}}^2$      | $B^+$      | $< 25$                                          |
| $t$                       | $B^+$      | $< 9 \cdot \tau_{PDG}(B_s^0)$                   |
| BDTS                      |            | $> 0.05$                                        |
| $m_{J/\psi K}$            |            | $> 5180 \text{ MeV}$ and $< 5700 \text{ MeV}$   |
| $1 + p_{J/\psi} / p_K$    |            | $> 1.4$ and $< 20$                              |

- the impact parameter (IP(B)) and impact parameter  $\chi^2$  (IP $\chi^2$ (B)) of the  $B$  candidate;
- the  $\chi^2$  of the B vertex (VCHI2);
- the angle between the direction of the momentum of the  $B$  candidate and the direction defined by the secondary and the primary vertices (DIRA);
- the minimum distance between the two daughter tracks (DOCA);
- the minimum impact parameter of each child particle with respect to any primary vertex (minIP( $\mu$ )).

In order to minimise the systematic uncertainty in the normalisation factors, the same BDTS cut is also applied to the normalisation channels. For the  $B^+ \rightarrow J/\psi K^+$  mode, the  $\chi^2$  of the secondary vertex is substituted by the  $\chi^2$  of the  $J/\psi$  vertex, the flight distance is computed between the  $J/\psi$  vertex and the primary vertex and the DOCA is computed between the two muons from the  $J/\psi$  decay.

This way, the distributions of all the variables but minIP and DOCA, are very similar for  $B_s^0 \rightarrow e^\pm \mu^\mp$ ,  $B^0 \rightarrow K^+ \pi^-$  and  $B^+ \rightarrow J/\psi K^+$ , resulting in a similar efficiency for signal and normalisation channels, as in Ref. [106].

The cut BDTS  $> 0.05$  has an efficiency of  $\sim 92\%$  on the signal. The distribution of the BDTS response for simulated signal and data sideband background is shown in Figure 5.1 with the cut already applied.

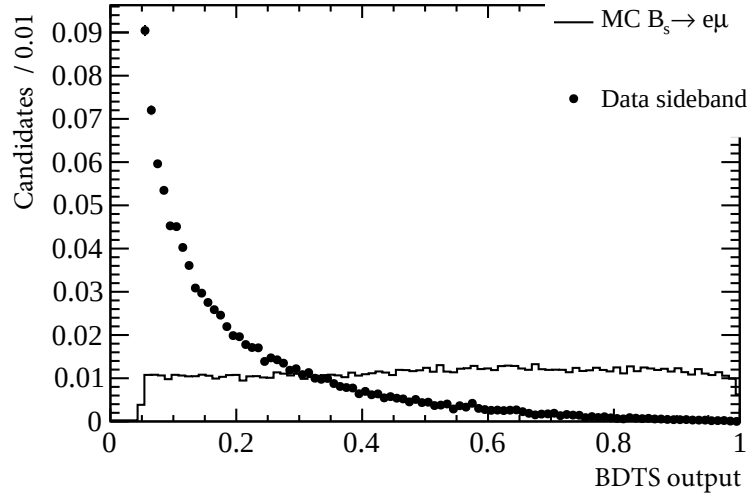


Figure 5.1 – Distribution of the BDTs classifier output after pre-selection for simulated signal and data sideband background.

#### 5.1.4 PID

Table 5.6 summarises the particle identification requirements applied on the signal candidate.

The requirement on the electron, being particularly delicate, has been optimised by maximising its rejection with respect to a cocktail of  $B_{(s)}^0 \rightarrow h^+ h^-$  ( $h^+ h^- = KK, \pi\pi, K\pi$ ) simulated mis-identified events, which are the most dangerous background, while keeping a similar signal efficiency. The rejection rate is calculated as:

$$\text{FOM} = \sum_{B_{(s)}^0 \rightarrow hh} \frac{f_{d,s}}{f_d} \mathcal{B}(B_{(s)}^0 \rightarrow h^+ h^-) \epsilon_{hh \rightarrow e\mu}^{\text{PID}},$$

where  $\frac{f_{d,s}}{f_d}$  is the relative hadron production fraction,  $\mathcal{B}(B_{(s)}^0 \rightarrow h^+ h^-)$  the branching fraction of each  $B_{(s)}^0 \rightarrow h^+ h^-$  decay and  $\epsilon_{hh \rightarrow e\mu}^{\text{PID}}$  is the mis-identification probability of the specific hadrons to be identified as electrons and muons.

Figure 5.2 shows the signal efficiency and background rejection for different cuts that have been analysed, involving different types and versions of PID variables.

Table 5.6 – PID selection for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ .

| Particle | Variable                                                                | Value    |
|----------|-------------------------------------------------------------------------|----------|
| Muon     | $\text{ProbNNmu} \cdot (1 - \text{ProbNNk}) \cdot (1 - \text{ProbNNp})$ | $> 0.4$  |
| Electron | DLLe                                                                    | $> 5.5$  |
| Electron | ProbNNk                                                                 | $< 0.95$ |



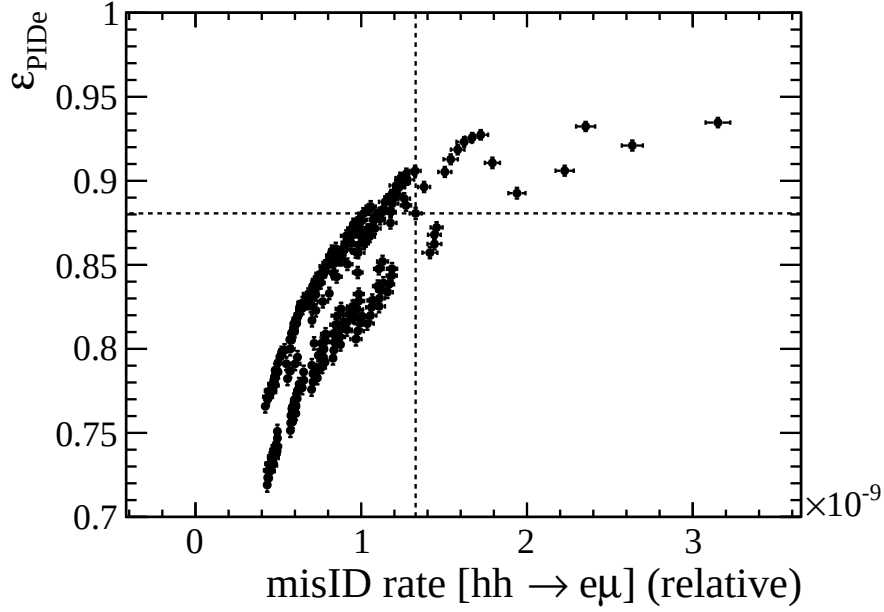


Figure 5.2 – Optimisation of PID requirement for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  with respect to  $B_{(s)}^0 \rightarrow h^+ h^-$  mis-identification. The two sets of points correspond to the optimisation run with and without the ProbNNk requirement. The dashed lines represent the PID cut used in the 2013 analysis [46].

### 5.1.5 The BDT classifier

A Boosted Decision Tree (BDT), designed using the TMVA package [107] is employed to separate the  $B_s^0 \rightarrow e^\pm \mu^\mp$  signal from the combinatorial background. The BDT is trained using a sample of  $B_s^0 \rightarrow e^\pm \mu^\mp$  simulated events ( $\sim 120000$ ) for the signal, and same-sign data events ( $\sim 60000$  events) representative of combinatorial background. Two independent halves of the samples are used for training and test. The choice to use the SS data is due to the need of training the BDT on an independent sample with respect to the one used at a later stage to fit the combinatorial background in bins of BDT.

The BDT response distribution on signal,  $\text{BDT}^{sig}$ , is subsequently *flattened* in the range  $[0, 1]$  such that the simulated signal is uniformly distributed while the background peaks at zero. This is achieved through the following transformation:

$$\text{BDT}_{flat}^{sig}(x) = \frac{\int_{-\infty}^x \text{BDT}^{sig}(y) dy}{\int_{-\infty}^{+\infty} \text{BDT}^{sig}(y) dy} \quad (5.3)$$

which does not change the signal/background separation but is convenient to split the BDT range in subregions of known signal efficiency.

The shape of the BDT variables is then validated on  $B^0 \rightarrow K^+ \pi^-$  data (see Section 5.1.6) in order to exclude any potential bias due to mis-modeling of the BDT input variables used in the training.

The range of the BDT variable is split in 8 bins:

[0, 0.25], [0.25, 0.4], [0.4, 0.5], [0.5, 0.6], [0.6, 0.7], [0.7, 0.8], [0.8, 0.9], [0.9, 1.0];

this is the binning that is used for the rest of the analysis in the study of backgrounds, efficiencies, and in the final mass fit. The first two bins are wider because, in data, they correspond to the region with the highest amount of background, and increasing their width improves the expected signal-to-noise ratio.

### BDT definition and performance

The BDT operator used in the current analysis is built with twelve features describing the signal kinematics through its geometry, cleanliness and reconstruction quality:

- the  $B$  proper time ( $t$ );
- the minimum impact parameter  $\chi^2$  for the two tracks IP  $\chi^2$ ;
- the impact parameter of the  $B$ , IP( $B$ );
- the distance of closest approach between the two daughter tracks (DOCA);
- the isolation of the two tracks with respect to any other track in the event  $I(\mu)$  [108], defined as the number of tracks in a cone of angle 0.27 rad around the reconstructed track (electron or muon) ;
- the transverse momentum of the  $B$ , ( $p_T(B)$ );
- the cosine of the angle between the muon momentum in the  $B$  rest frame and the vector perpendicular to the  $B$  momentum and the beam axis ( $\cos \theta_k$ );
- the  $B$  isolation  $I(B)$  (CDF definition [109]):

$$I(B) = \frac{|\vec{p}_T^{e\mu}|}{\sum_i p_T^i - |\vec{p}_T^{e\mu}|} \quad (5.4)$$

where the sum is over all tracks with  $\sqrt{\Delta\eta^2 + \Delta\Phi^2} \leq 1$  ( $\eta$ =pseudo-rapidity,  $\Phi$ =azimuthal angle) wrt  $\vec{p}^{e\mu}$ ;

- the flight distance of the  $B$  meson with respect to its PV, FD;
- the maximum transverse momentum  $p_{T,max}$  of the two daughter tracks;
- the  $\chi^2$  of the decay vertex of the  $B$  meson,  $\chi_{DV}^2$ ;
- the difference of pseudo-rapidity between the two daughter tracks,  $\Delta\eta$ .

The distributions of the twelve variables for  $B_s^0 \rightarrow e^\pm \mu^\mp$  signal MC,  $B_s^0 \rightarrow e^\pm \mu^\mp$  data sidebands and  $B_s^0 \rightarrow e^\pm \mu^\pm$  data are compared in Appendix C.

The Receiver Operating Characteristic (ROC) curve<sup>2</sup> and the importance of the input variables are shown in Figure 5.3. The performances are evaluated by training the BDT on half of the MC sample and applying it to the other half and to the opposite-sign (OS) data sidebands, defined by the cut  $m_{e\mu} \in [4900, 5100] \cup [5500, 5850] \text{ MeV}/c^2$ . The results are compared with the BDT that was used in Ref. [46]

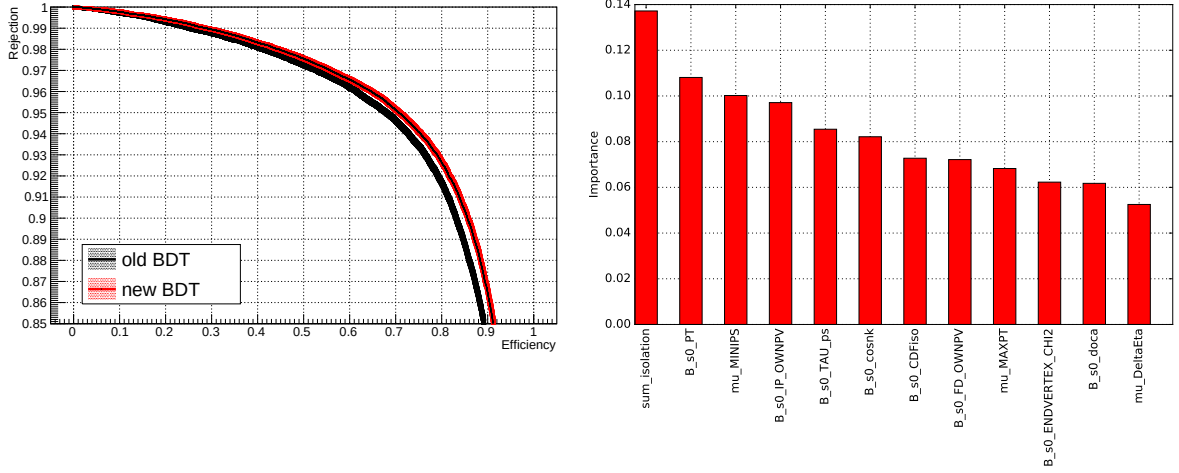


Figure 5.3 – (left) ROC curve for the new BDT, in red, compared with the BDT used in Ref. [46], in black. The ROC curve is zoomed in the region of high rejection. (right) Relative importance for the BDT inputs.

Table 5.7 shows the fractions of signal events surviving different cuts on the flat BDT response for signal (MC) and combinatorial background (data) compared between the new and old BDT.

Figure 5.4 shows the BDT response distribution before the flattening for signal and background, comparing the training and test samples, which are found to be in good agreement, showing no sizable sign of over-training.

The correlation of the BDT response with the mass of the combinatorial background has also been checked, as it could generate fictitious excesses in the signal region. Figure 5.5 shows the two-dimensional histogram of BDT response versus the mass of the candidates from the right data sideband and same-sign data. The overlaid red points show the mean value of BDT for each mass bin. The Pearson's correlation factor is found to be  $\sim -0.2\%$  on the OS data sidebands, and  $\sim -3.8\%$  on the SS sample.

<sup>2</sup>The ROC curve describes the diagnostic ability of a binary multivariate classifier by showing the efficiency and rejection power corresponding to different cuts on its response.

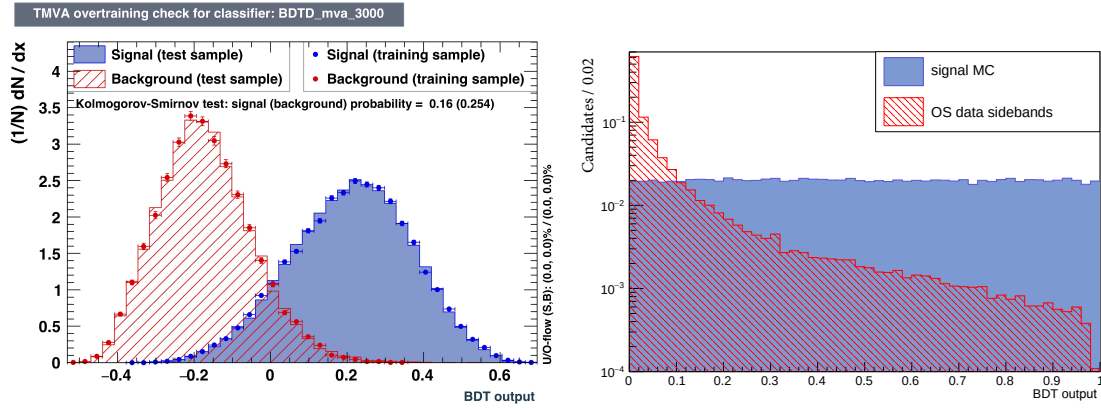


Figure 5.4 – (left) BDT response (before flattening) for the signal (blue) and background (red) samples used for training (points) and testing (filled histograms). The Kolmogorov-Smirnov test's p-value is overlaid. (right) BDT response after flattening compared between simulated signal and OS data sidebands.

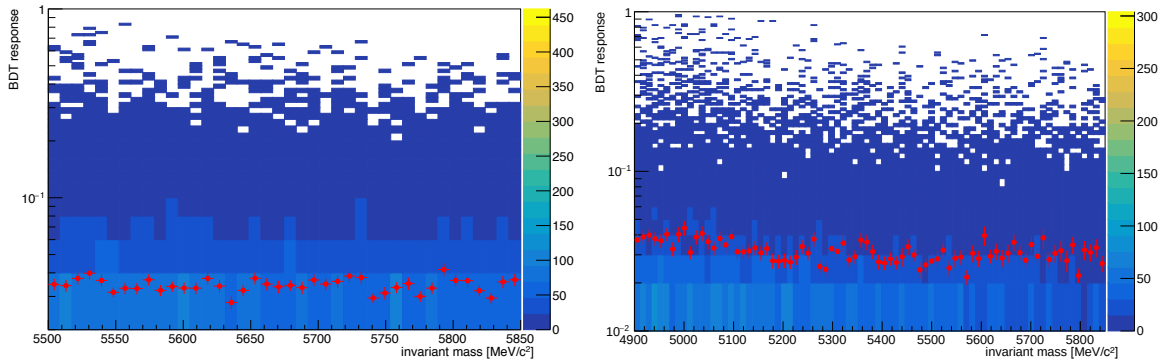


Figure 5.5 – BDT response vs mass distribution for data events in the (left) OS right mass sideband and (right) SS sample. The red points show the mean value of BDT for each mass bin.

Table 5.7 – Fractions of signal events surviving a requirement on the flattened BDT response,  $\text{BDT}_{flat}^{sig} > X$ , for simulated signal and combinatorial background (data) compared between the new and old [46] BDT.

| cut | $f_{comb}^{new}$ | $f_{sig}^{new}$ | $f_{comb}^{old}$ | $f_{sig}^{old}$ | $\frac{(S/B)^{new}}{(S/B)^{old}}$ |
|-----|------------------|-----------------|------------------|-----------------|-----------------------------------|
| 0.1 | 0.1366           | 0.9024          | 0.2138           | 0.9311          | 1.5                               |
| 0.2 | 0.0740           | 0.8020          | 0.1110           | 0.8484          | 1.4                               |
| 0.3 | 0.0483           | 0.6997          | 0.0688           | 0.7609          | 1.3                               |
| 0.4 | 0.0340           | 0.5973          | 0.0478           | 0.6658          | 1.3                               |
| 0.5 | 0.0239           | 0.4955          | 0.0342           | 0.5677          | 1.3                               |
| 0.6 | 0.0165           | 0.3973          | 0.0241           | 0.4625          | 1.3                               |
| 0.7 | 0.0104           | 0.2978          | 0.0158           | 0.3525          | 1.3                               |
| 0.8 | 0.0057           | 0.1997          | 0.0090           | 0.2334          | 1.4                               |
| 0.9 | 0.0022           | 0.0989          | 0.0033           | 0.1104          | 1.3                               |

### 5.1.6 Determination of the BDT PDF for signal

As explained in Section 5.1.5, the BDT classifier is trained using  $B_s^0 \rightarrow e^\pm \mu^\mp$  simulated events to model the signal, and same-sign data events for background and it is flattened using  $B_s^0 \rightarrow e^\pm \mu^\mp$  Monte Carlo as well. Nevertheless, in order to check the absence of dangerous inconsistencies between data and simulation, potentially resulting in a wrong estimation of the expected number of signal in each BDT bin, its PDF for signal is subsequently compared with data.

Since the BDT is designed to only use kinematic information of a two-body neutral  $B$  meson decay,  $B_{(s)}^0 \rightarrow h^+ h'^{-}$  (where  $h$  and  $h'$  are charged kaons or pions) is a suitable channel to calibrate the classifier, acting as a proxy for the  $B_s^0 \rightarrow e^\pm \mu^\mp$  decay. However, given the branching fractions of the neutral  $B$  mesons into two hadrons, only the most frequent  $B^0 \rightarrow K^+ \pi^-$  process is considered in the following. In order to distinguish this channel from the others and therefore to identify the correct mass hypothesis of the final state particles, a requirement on the  $\Delta LL_{K-\pi}$  variable of the final state particles is applied. The effect of misidentified components in the  $B^0 \rightarrow K^+ \pi^-$  yield per BDT bin is evaluated varying this requirement. The cut on the  $\Delta LL_{K-\pi}$  variable could however induce a bias as its separation power is highly correlated with the kinematics of the final state particle. To reduce this effect, the yields are corrected by the efficiency of the  $\Delta LL_{K-\pi}$  cut, evaluated as a function of particle momenta and pseudo-rapidity.

#### $\Delta LL_{K-\pi}$ cut efficiency determination

The efficiency is estimated using the PIDCalib [110] tool. This is done separately per polarity and year of data taking, to avoid bias due to detector conditions and occupancy.

PIDCalib provides generalised calibration samples and uses tag-and-probe and  $sWeights$

methods to extract particle-identification efficiencies.  $D^{*\pm} \rightarrow (D^0 \rightarrow K^\pm \pi^\mp) \pi^\pm$  events are used to calculate the efficiency as a function of the momentum of the final state hadron,  $p$  and its pseudo rapidity,  $\eta$ . The binning scheme used throughout this analysis to determine the efficiency of the  $\Delta LL_{K-\pi}$  cuts is:

- $p$ : 2 bins for  $0 < p < 10 \text{ GeV}/c$ , 45 bins for  $10 \text{ GeV}/c < p < 100 \text{ GeV}/c$ , 20 bins for  $100 \text{ GeV}/c < p < 150 \text{ GeV}/c$  4 bins for  $150 \text{ GeV}/c < p < 500 \text{ GeV}/c$ ;
- $\eta$ : 10 bins for  $1 < \eta < 6$ .

As the calibration samples used by PIDCalib have different kinematics with respect to  $B^0 \rightarrow K^+ \pi^-$ , in order to determine the PID efficiency for the latter, the above-mentioned efficiencies are folded with the kinematics  $(p, \eta)$  from  $B^0 \rightarrow K^+ \pi^-$  simulated events.

For each event of the MC, the PID selection efficiency is smeared 1000 times according to its uncertainty; each value is then used to obtain a different global PID efficiency, as the average of the  $i$ -th smeared values of all events. The mean of the 1000 averages is then taken as the final efficiency and its root-mean-square error is used as the uncertainty. The PID efficiencies are obtained per BDT bin and  $\Delta LL_{K-\pi}$  cut, and are used to correct the obtained yields before calculating the fraction of events per BDT bin.

### BDT PDF determination

The sample used to calibrate the BDT is the  $B^0 \rightarrow K^+ \pi^-$  sample defined in Section 5.1.2 with the following additional requirements applied:

- both hadrons in the muon acceptance;
- L0\_Electron (TIS) || L0\_Muon (TIS) || L0\_Photon (TIS);
- Hlt1\_TrackAllL0 (TIS) || Hlt1\_SingleMuonHighPT (TIS) ||  
Hlt1\_TrackMuon (TIS) || Hlt1\_TrackPhoton (TIS) ||  
Hlt1\_DiMuonLowMass (TIS) || Hlt1\_DiMuonHighMass (TIS);
- Hlt2\_B2HH;

where the meaning of TIS is explained in Section 3.5.2.

The total sample is divided into the eight BDT bins indicated in Section 5.1.5. The goal of the BDT calibration is to determine the fraction of  $B^0 \rightarrow K^+ \pi^-$  events for each BDT bin through a fit of the invariant mass of the two hadrons,  $m(K\pi)$ . The correct mass hypothesis for each hadron is chosen by cutting on the  $\Delta LL_{K-\pi}$  value. If  $\Delta LL_{K-\pi} > \kappa$ , with  $\kappa$  positively defined, the hadron is identified as a kaon, otherwise if  $\Delta LL_{K-\pi} < -\kappa$  it is identified as a pion. All the events with  $|\Delta LL_{K-\pi}| < \kappa$  are rejected. In order to exclude the partially reconstructed background, the mass window is chosen to run from 5200 to 5850  $\text{MeV}/c^2$ . The upper bound is chosen

taking into account the invariant mass cut in the stripping selection. The stability of the result under different  $\Delta LL_{K-\pi}$  cuts is investigated by scanning from  $\kappa = 5$  to  $\kappa = 10$  in steps of 0.5. To determine the number of events per bin, a maximum-likelihood binned fit is performed. The number of events in the most background-like BDT response bin is obtained by fitting the full BDT range and subtracting the number of events falling in all the other bins. This allows to have always very clean peaks in the mass fit.

The resulting invariant-mass distributions are fitted with a Double-Sided Crystal Ball distribution [111] for the  $B^0$  signal peak, with the tail parameters constrained from  $B^0 \rightarrow K^+ \pi^-$  MC. A  $B_s^0 \rightarrow K^+ \pi^-$  component is also present, and modelled in the same way. As the  $B_s^0$  yield is small compared to fluctuations of the combinatorial background, the  $B_s^0$  mean is fixed relative to the  $B^0$  mean adding the known mass difference between the  $B^0$  and  $B_s^0$  from the PDG, while its width is obtained by correcting the width of the  $B^0$  by a factor that corresponds to the ratio between the  $B^0$  and  $B_s^0$  widths obtained from simulation. For the combinatorial background, an exponential function is used. A component accounting for a small background from  $\Lambda_b^0 \rightarrow ph$  is included, as a Double Sided Crystal Ball distribution with all the parameters constrained from simulation. The  $B^0 \rightarrow \pi^+ \pi^-$  and  $B^0 \rightarrow K^+ K^-$  components are found to be negligible for  $\kappa = 5$ , therefore excluded from the fit model. Figure 5.6 shows the fits for  $\kappa = 5$  in the most background-like and most signal-like BDT bins as examples. All the bins are shown in Appendix B.

From the fit, the number of  $B^0$  candidates in each bin is obtained, which is then corrected for the cut on  $\Delta LL_{K-\pi}$  to obtain a PID-independent quantity. The number of  $B^0$  candidates along with the PID efficiency are evaluated as described before and the fit yields are reported in Table 5.8, along with the PID efficiency correction for  $\kappa = 5$ .

Table 5.8 – Yields from the  $B^0 \rightarrow K^+ \pi^-$  fit and PID efficiency corrections per BDT bin at  $\kappa = 5$ . The first bin contains the full BDT range, to which the yield of all the others bins is later subtracted.

| BDT bin      | data yield | data yield<br>stat. error | PID<br>efficiency correction | PID<br>efficiency correction error |
|--------------|------------|---------------------------|------------------------------|------------------------------------|
| [0.0, 1.0]   | 8993.74    | 192.26                    | 1.964                        | $3 \times 10^{-6}$                 |
| [0.25, 0.40] | 1209.40    | 41.79                     | 1.962                        | $6 \times 10^{-6}$                 |
| [0.40, 0.50] | 926.54     | 37.02                     | 2.029                        | $1 \times 10^{-5}$                 |
| [0.50, 0.60] | 864.60     | 35.28                     | 2.086                        | $9 \times 10^{-6}$                 |
| [0.60, 0.70] | 864.39     | 30.04                     | 2.097                        | $1 \times 10^{-5}$                 |
| [0.70, 0.80] | 841.13     | 33.91                     | 2.082                        | $1 \times 10^{-5}$                 |
| [0.80, 0.90] | 860.21     | 30.08                     | 2.062                        | $9 \times 10^{-6}$                 |
| [0.90, 1.00] | 920.06     | 30.97                     | 2.046                        | $2 \times 10^{-5}$                 |

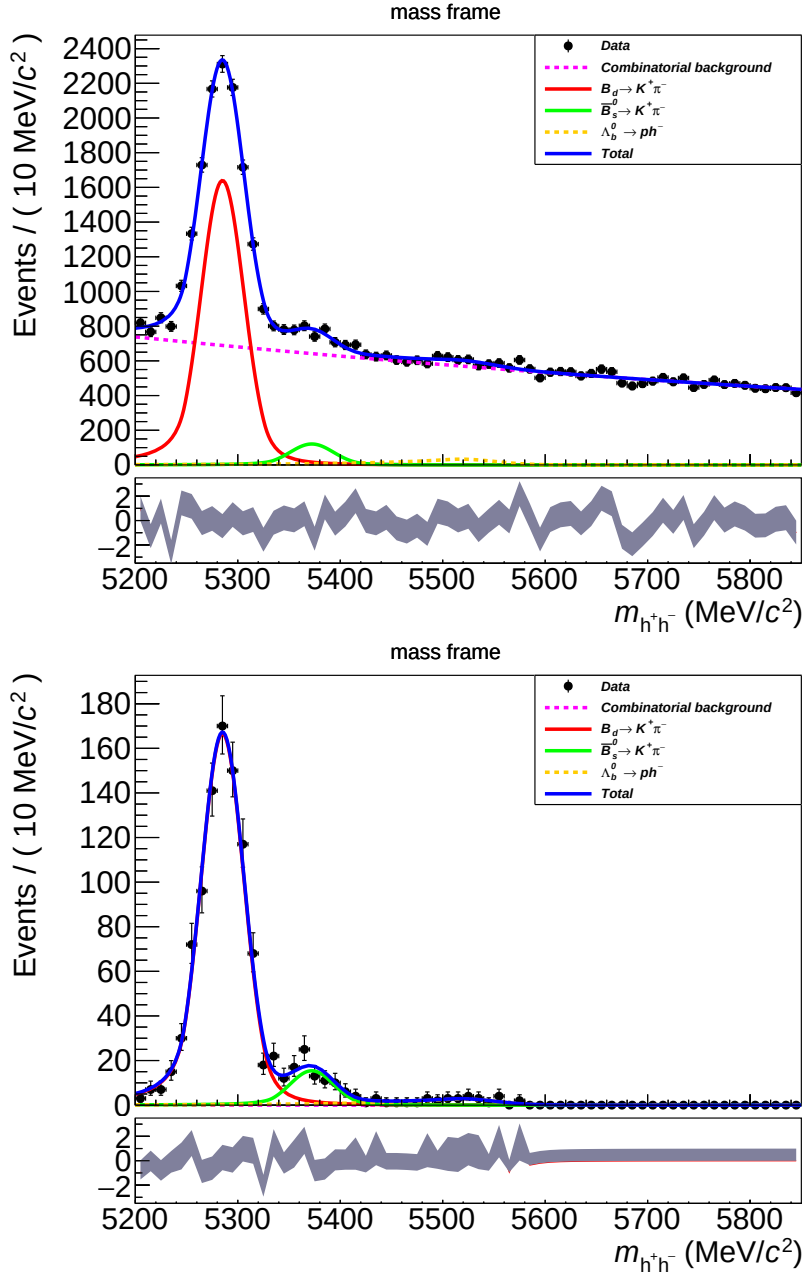


Figure 5.6 – Invariant mass distributions of  $B^0 \rightarrow K^+ \pi^-$  candidates in Run I data in the most background-like (top) and the most signal-like (bottom) BDT bins with a PID requirement  $|\Delta LL_{K-\pi}| < 5$ . The red solid line shows the  $B^0$  signal, the green one shows the  $B_s^0$  component, the yellow dashed one the one from  $\Lambda_b^0 \rightarrow p h^-$  where the proton is misidentified as a kaon or pion. The combinatorial background is shown by the purple dashed line.

The fraction of events for the  $i$ -th BDT bin and PID cut  $\kappa$  is then determined as:

$$r_i(\kappa) = N_{B^0,i}(\kappa) / \sum_j (N_{B^0,j}(\kappa)) \quad (5.5)$$



where  $N_{B_d,i}(\kappa)$  is the number of events after dividing by the PID efficiency for  $i$ -th BDT bin and PID cut  $\kappa$ . Figure 5.7 shows the fraction of events per BDT bin as a function of the PID requirement. The fraction of events is found to be very stable (within 1%) over the range of  $\Delta LL_{K-\pi}$  cuts considered.

To take into account all of the information from the fits at different PID cuts, a weighted average  $\bar{r}_i$  is determined over the whole PID cut range, taking into account the correlations between  $r_i$  for different values of  $\kappa$ :

$$\begin{aligned}\bar{r}_i &= \sigma_{\bar{r}_i}^2 \sum_a c_{i;a,b}^{-1} r_{i,a} \\ \sigma_{\bar{r}_i}^2 &= \frac{1}{\sum_{a,b} c_{i;a,b}^{-1}} \\ \text{with } c_{i;a,b}^{-1} &: \text{element of the inverted covariance matrix} \\ a, b &: \text{indices of the different data points for the different } \Delta LL_{K-\pi} \text{ cuts,}\end{aligned}$$

$\sigma_{\bar{r}_i}$  is also the estimated statistical error on  $\bar{r}_i$ .

The covariance matrix for the  $i$ -th BDT bin is defined as

$$\begin{aligned}c_{i;a,b} &= \sigma_{r_{i,a}} \cdot \sigma_{r_{i,b}} \cdot \rho_{i,a,b} \\ \text{with } \sigma_{r_{i,a}} &: \text{statistical error on } r_{i,a} \text{ propagated from Equation (5.5).}\end{aligned}$$

$\rho_{i,a,b}$  is the correlation estimated using the number of events in the  $i$ -th BDT surviving  $\Delta LL_{K-\pi}$  cuts of  $\kappa$  ( $n_{i;a}$ ) and those surviving cuts of  $\kappa' > \kappa$  ( $n_{i;b}$ ):

$$\rho_{i,a,b} = \sqrt{\frac{n_{i;b}}{n_{i;a}}}.$$

The formula for the correlation is derived from comparing the binomial uncertainty on  $\varepsilon = n_{i;b}/n_{i;a}$  which is  $\sigma_\varepsilon^2 = \varepsilon(1-\varepsilon)/n_{i;a}$  with the result of gaussian error propagation:

$$\sigma_\varepsilon^2 = \left( \frac{\partial \varepsilon}{\partial n_{i;b}} \sigma_{n_{i;b}} \right)^2 + \left( \frac{\partial \varepsilon}{\partial n_{i;a}} \sigma_{n_{i;a}} \right)^2 + 2\rho \left( \frac{\partial \varepsilon}{\partial n_{i;b}} \sigma_{n_{i;b}} \right) \left( \frac{\partial \varepsilon}{\partial n_{i;a}} \sigma_{n_{i;a}} \right)$$

using  $\sigma_{n_{i;b}} = \sqrt{n_{i;b}}$  and  $\sigma_{n_{i;a}} = \sqrt{n_{i;a}}$ .

Figure 5.8 shows the BDT signal PDE. This distribution has also been checked for the  $B_s^0 \rightarrow \pi^+ K^-$  candidates, which can be obtained from the same fit, and it was found to be flat within the errors.

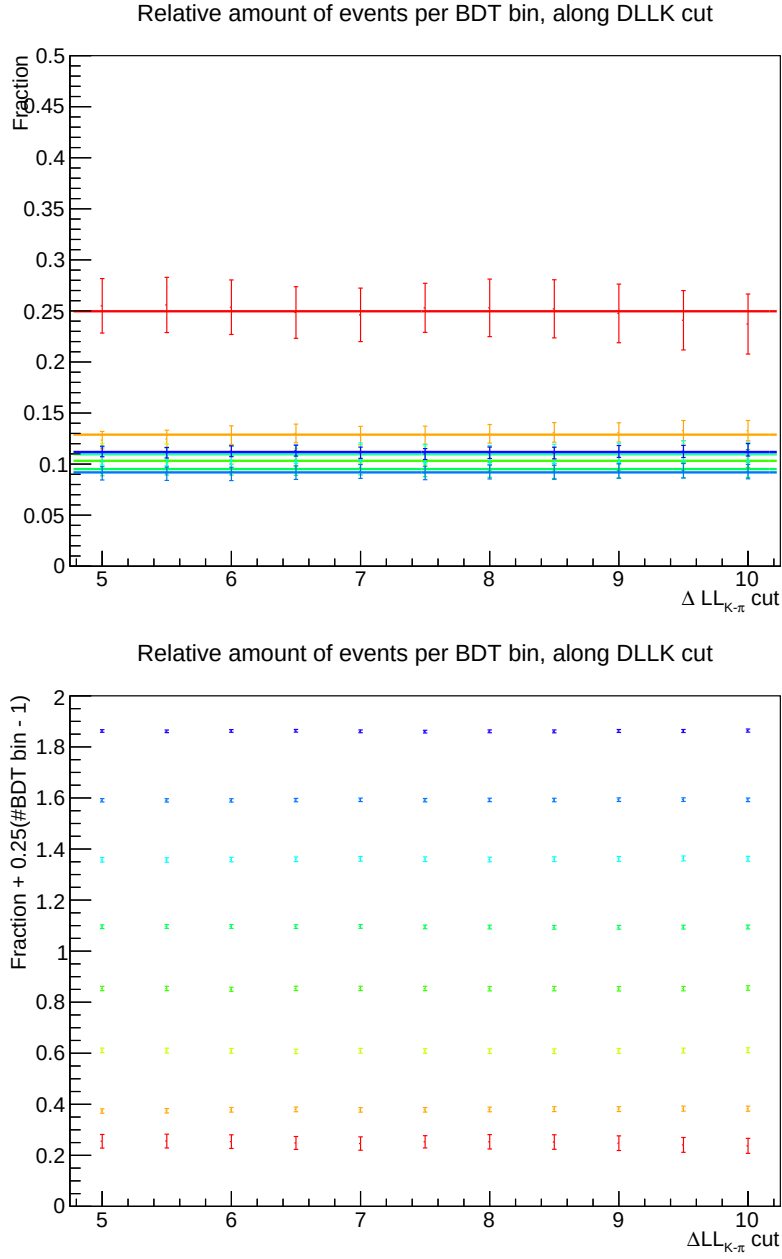


Figure 5.7 – (Top) fraction of  $B^0 \rightarrow K^+ \pi^-$  data in each BDT bin for different  $\Delta LL_{K-\pi}$  cuts ( $\kappa = 5 \dots 10$ ). Each BDT bin is fitted with a constant, to check consistency along  $\kappa$ . (bottom) For better visualisation the fractions are shifted by  $0.25 \cdot (i - 1)$ .

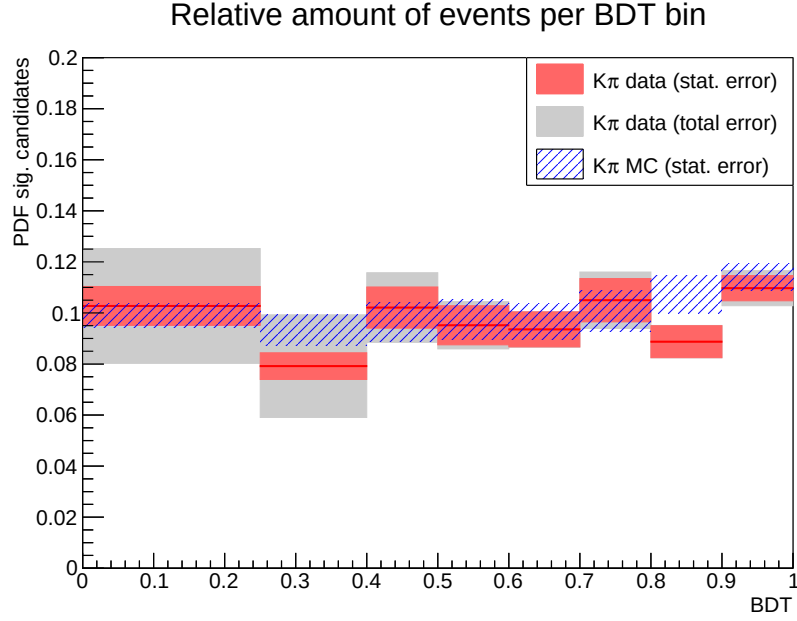


Figure 5.8 – BDT signal PDF for  $B^0 \rightarrow K^+ \pi^-$  candidates in Run I data. The statistical uncertainty is included as a red band and total uncertainty as a grey band. The systematic uncertainties are discussed in Section 5.1.6. The distribution for  $B^0 \rightarrow K^+ \pi^-$  simulated candidates (blue hatched) is superimposed.

The PDFs shown in these plots have been corrected for the trigger requirements placed on  $B_s^0 \rightarrow e^\pm \mu^\mp$  relative to  $B^0 \rightarrow K^+ \pi^-$ , as explained in Section 5.1.6.

The statistical uncertainty is included as a red band and the total uncertainty as a grey band. The results are summarised in Table 5.9. The agreement between data and MC allows to exclude the existence of sizeable biases in the BDT PDF estimation.

### Systematics

The following contributions are considered for the evaluation of systematic uncertainties on the calibration of the BDT:

1. **PID consistency:** a linear  $\chi^2$  fit to the data points in the PID cut range for each BDT bin is performed. If  $r_i$  at  $\kappa = 5$  is found to be significantly different from the baseline value  $\bar{r}_i$ , the difference is taken as a systematic uncertainty. Significant is defined as  $|r_{\text{inter}} - \bar{r}| > \sqrt{\sigma_{\text{inter}}^2 + \sigma_{\bar{r}}^2}$  where  $\sigma_{\text{inter}}$  is the uncertainty extracted from the fit on the extrapolated value at  $\kappa = 5$ . It is found that all BDT bins are consistent over their PID range; therefore, this systematic is neglected.
2. **PID efficiency determination:** To check the PID efficiency, this latter is determined using a binning with twice as many bins as described at the beginning of this section.

Table 5.9 – Fractions of  $B^0 \rightarrow K^+ \pi^-$  candidates falling into each BDT bin for data and simulated signal. The error quoted for data is the total (stat. and syst.), while for the MC it is just the statistical.

| BDT bin      | data<br>fraction | data<br>fraction<br>error | MC<br>fraction | MC<br>fraction<br>error |
|--------------|------------------|---------------------------|----------------|-------------------------|
| [0.00, 0.25] | 0.257            | 0.056                     | 0.247          | 0.005                   |
| [0.25, 0.40] | 0.119            | 0.030                     | 0.140          | 0.006                   |
| [0.40, 0.50] | 0.102            | 0.014                     | 0.096          | 0.008                   |
| [0.50, 0.60] | 0.095            | 0.009                     | 0.097          | 0.008                   |
| [0.60, 0.70] | 0.094            | 0.007                     | 0.096          | 0.007                   |
| [0.70, 0.80] | 0.105            | 0.011                     | 0.101          | 0.008                   |
| [0.80, 0.90] | 0.089            | 0.006                     | 0.107          | 0.007                   |
| [0.90, 1.00] | 0.110            | 0.007                     | 0.114          | 0.005                   |

The difference with the efficiency determined using the nominal binning is taken as a systematic per BDT bin.

3. **Fraction of events outside the mass window:** The amount of  $B^0$  events outside the mass window, due to the cutoff at  $5200 \text{ MeV}/c^2$ , is about 0.8%. This number is stable within  $\sim 0.3\%$ . The BDT calibration is performed with a correction factor which is the reciprocal of the number of events in the mass window, and the result per bin is compared with the nominal result. The difference is assigned as a systematic.
4. **Fit model:** The analysis is repeated with a different fit model. Instead of the nominal model, the signal shapes are described by Johnson distributions [112] with all the parameters free, except the mass difference between the  $B^0$  and  $B_s^0$  peaks, which is fixed to the PDG value as in the nominal model. The mass distributions for  $\kappa = 5$  in the most background-like and in the most signal-like BDT response bins are shown in Figure 5.9. All the distributions can be observed in Appendix B. The difference between the two models is assigned as a systematic.

### BDT PDF correction for trigger efficiency

In order to account for the difference of trigger efficiency between the requirements that are applied to select the  $B^0 \rightarrow K^+ \pi^-$  sample and those that are applied on the signal, a correction is applied to the shape of the BDT on  $B^0 \rightarrow K^+ \pi^-$ . Each bin of the BDT response is multiplied by the following factor:

$$t_i = \frac{\epsilon_{i,e\mu}^{Trig|Reco\&Sel}}{\epsilon_{i,K\pi}^{Trig|Reco\&Sel}}, \quad (5.6)$$

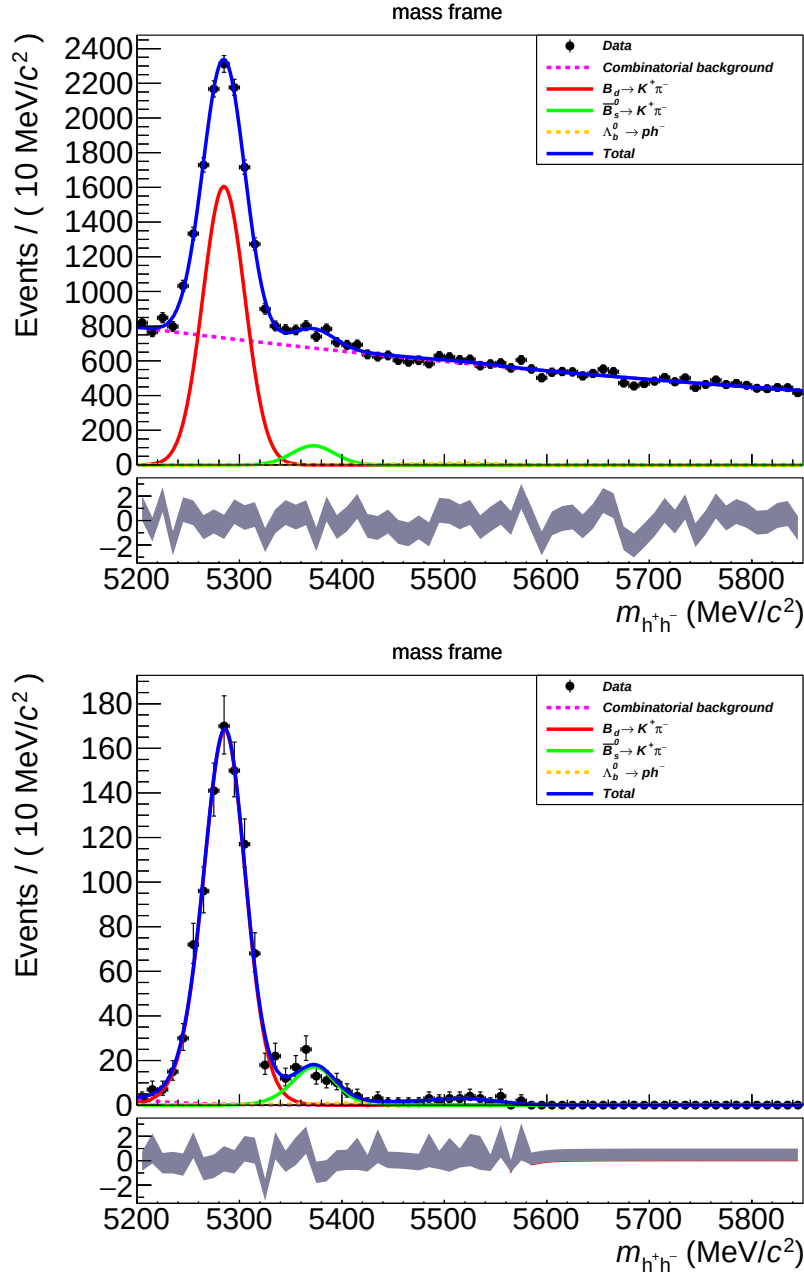


Figure 5.9 – Invariant mass distributions of  $B^0 \rightarrow K^+ \pi^-$  candidates from Run I data in the most background-like (top) and the most signal-like (bottom) BDT bins for  $|\Delta LL_{K-\pi}| < \kappa$  cut, with  $k = 5$  with the alternative PDF used for the evaluation of the systematic error. The red solid line shows the  $B^0$  signal, the green one shows the  $B_s^0$  component, the yellow dashed one the one from  $\Lambda_b^0 \rightarrow p h^-$  where the proton is misidentified as a kaon or pion. The combinatorial background is shown by the purple dashed line.

where  $\varepsilon_{i,K\pi}^{Trig|Reco\&Sel}$  is the efficiency of the  $B^0 \rightarrow K^+ \pi^-$  trigger cut evaluated on  $B^0 \rightarrow K^+ \pi^-$  simulation and  $\varepsilon_{i,e\mu}^{Trig|Reco\&Sel}$  is the efficiency of the signal trigger cut, in the  $i$ -th BDT bin. The

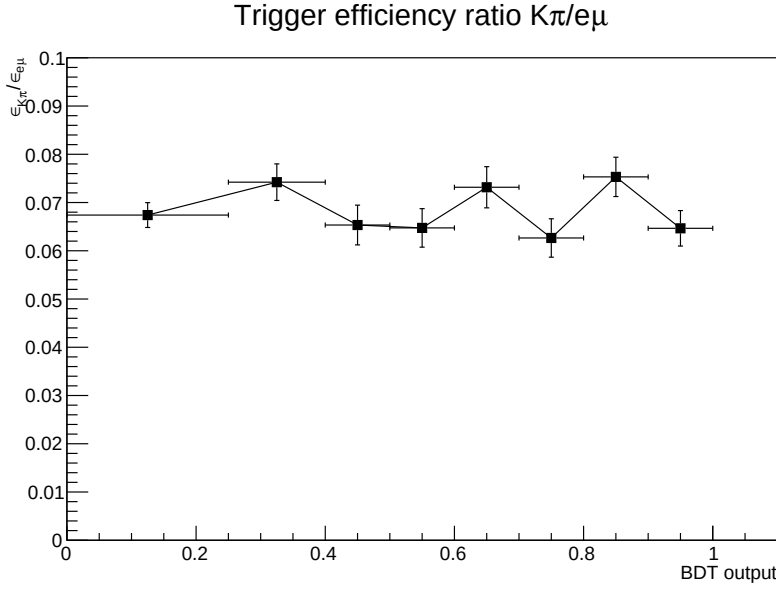


Figure 5.10 – Trigger efficiency ratio between  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow K^+ \pi^-$  for the different BDT bins.

plot of  $t_i$  in bins of BDT is shown in Figure 5.10, which shows a flat distribution, within the errors, thanks to the choice of TIS trigger requirements on  $B^0 \rightarrow K^+ \pi^-$ .

### BDT shape for HasBremAdded categories

The BDT response is flattened on signal MC, where both categories HasBremAdded=1 and HasBremAdded=0 are considered. Figure 5.11 shows the shape of the BDT response on an independent signal MC sample for the two categories and for the whole sample.

The  $BDT_{flat}^{sig}$  shape obtained from the  $B^0 \rightarrow K^+ \pi^-$  sample is representative of the distribution that would be obtained on a  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  sample with HasBremAdded=1, as evidenced by the analysis of the BDT features.

The ratio between these shapes is used at a later stage as a correction to the shape obtained from the BDT calibration, to extrapolate the expected BDT response for candidates with HasBremAdded=1 and HasBremAdded=0.

## 5.2 Calibration of the invariant mass

The knowledge of the signal mass shape is necessary to perform the fit to data. This section describes the procedure followed for the determination of such shape.

The mass PDF for the signal is obtained separately for the two bremsstrahlung categories

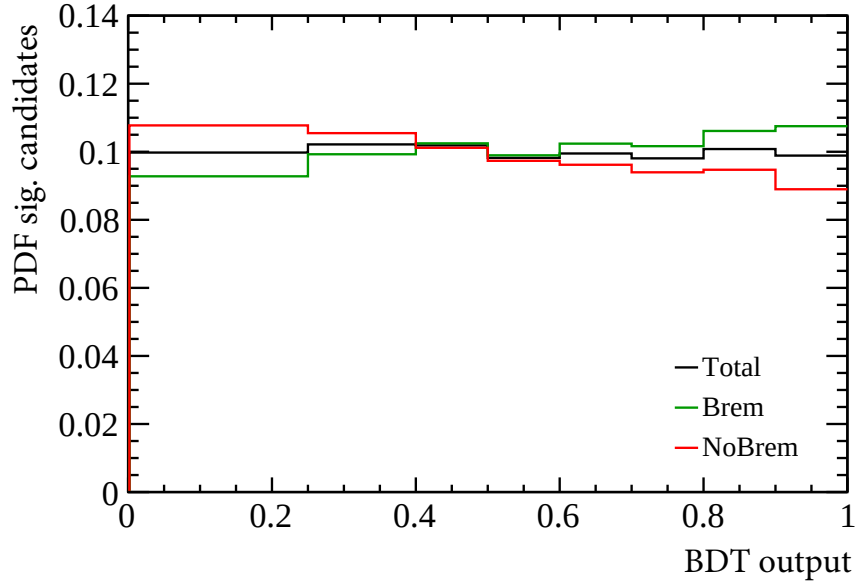


Figure 5.11 – Shape of the BDT response for candidates with and without bremsstrahlung photons recovered and for the whole sample, for signal.

(HasBremAdded). This is important because Bremsstrahlung changes significantly the mass shape. The BDT bin, on the other hand, has been observed to have no influence on it.

The models are obtained by fitting the 2-body invariant mass of simulated  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates with a Double-Sided Crystal Ball function with all parameters free. The mass fits for the most signal-like BDT bin are shown in Figure 5.12.

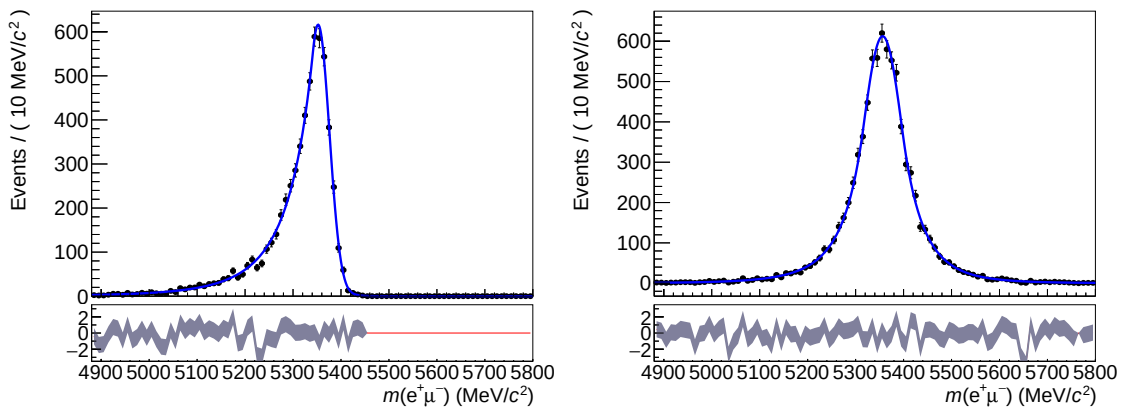


Figure 5.12 – Invariant mass fits to  $B_s^0 \rightarrow e^\pm \mu^\mp$  simulated candidates with HasBremAdded=0 (left) and HasBremAdded=1 (right). The blue, solid line is a Double-Sided Crystal Ball distribution with all parameters free in the fit.

### 5.2.1 Data-driven correction to the mass resolution

In order to take into account data-simulation differences in the resolution of the core of the mass distributions (i.e. the  $\sigma$  of the Double-Sided Crystal Ball), the PDFs obtained from MC are modified by multiplying their widths by a factor  $\mathcal{C}$  defined as:

$$\mathcal{C} = \frac{\left(\frac{\sigma_M}{M}\right)_{e\mu, DATA}}{\left(\frac{\sigma_M}{M}\right)_{e\mu, MC}}, \quad (5.7)$$

where the first subscript indicates that these quantities refer to the decay  $B_s^0 \rightarrow e^\pm \mu^\mp$ , the labels DATA and MC indicate which kind of events each quantity refers to, and

$$\left(\frac{\sigma_M}{M}\right)_i \simeq \frac{1}{2} \cdot \sqrt{\left[\left(\frac{\sigma_{E_1}}{E_1}\right)^2 + \left(\frac{\sigma_{E_2}}{E_2}\right)^2\right]} \quad (5.8)$$

with the subscripts  $1$  and  $2$  referring to the two daughter particles of a generic two-body decay. The opening angle resolution is here neglected with respect to the energy resolution.

Since there is no suitable proxy channel with only an electron and a muon in the final state, a measurement is performed using  $B^+ \rightarrow J/\psi(e^+e^-)K^+$  and  $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$  decays, and combining the results.  $\mathcal{C}$  can be obtained by measuring  $\left(\frac{\sigma_M}{M}\right)_{ee, DATA}$ ,  $\left(\frac{\sigma_M}{M}\right)_{ee, MC}$ ,  $\left(\frac{\sigma_M}{M}\right)_{\mu\mu, DATA}$  and  $\left(\frac{\sigma_M}{M}\right)_{\mu\mu, MC}$ , where the subscripts  $ee$  and  $\mu\mu$  refer to the dilepton invariant mass for  $B^+ \rightarrow J/\psi(e^+e^-)K^+$  and  $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ , respectively:

$$\mathcal{C} = \frac{\sqrt{\left[\left(\frac{\sigma_M}{M}\right)_{ee, DATA}^2 + \left(\frac{\sigma_M}{M}\right)_{\mu\mu, DATA}^2\right]}}{\sqrt{\left[\left(\frac{\sigma_M}{M}\right)_{ee, MC}^2 + \left(\frac{\sigma_M}{M}\right)_{\mu\mu, MC}^2\right]}}. \quad (5.9)$$

The factor  $\mathcal{C}$  has to be applied to the width obtained from  $B_s^0 \rightarrow e^\pm \mu^\mp$  simulation to reproduce the one expected in  $B_s^0 \rightarrow e^\pm \mu^\mp$  data. Further details on this procedure can be found in Appendix A.

The  $J/\psi$  mass is therefore fitted for  $B^+ \rightarrow J/\psi(e^+e^-)K^+$  and  $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$  candidates to obtain the needed resolutions. In both cases the fit is performed initially on simulation, using a Double-Sided Crystal Ball distribution with all the parameters free, and subsequently on data, adding a free exponential component to account for the background, and bounding the tail parameters of the signal to the values obtained on simulation using gaussian constraints. These fits are shown in Figure 5.13. The value of  $\mathcal{C}$  is found to be  $\mathcal{C} = 1.12 \pm 0.04$  on 2012 data and  $\mathcal{C} = 1.06 \pm 0.03$  on 2011 data.



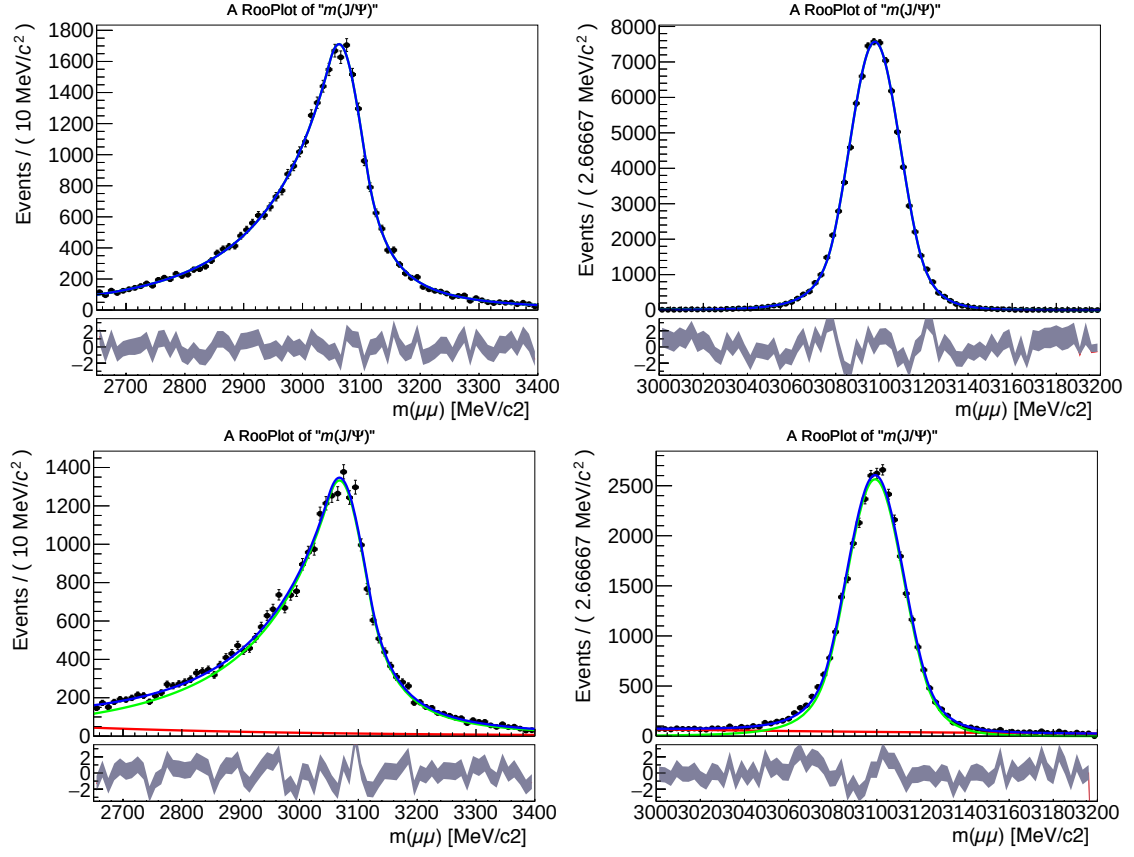


Figure 5.13 – Dilepton invariant mass  $m(J/\psi)$  fits for  $B^+ \rightarrow J/\psi(e^+e^-)K^+$  (left) and  $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$  (right) simulated candidates (top) and 2012 data (bottom). The blue line is the total distribution, while the green line is a Double-Sided Crystal Ball distribution and the red line is an exponential component for the background.

### 5.2.2 Fraction of HasBremAdded

The fraction of events with bremsstrahlung photon added, after the whole selection, is determined from simulated signal, and it is found to be:

$$f_{brem} = 0.5796 \pm 0.0017. \quad (5.10)$$

The uncertainty is included as a systematic in the mass fit for the upper limit determination.

As an additional cross-check,  $f_{brem}$  is compared between data and simulation using samples of  $B^+ \rightarrow J/\psi (\rightarrow e^+ e^-) K$  decays. The result is given in Table 5.10. As the ECAL performance is dependent on nSPDHits, a reweighting to data nSPDHits distribution is applied to MC to account for the large difference in occupancy in MC and data. Data and simulation provide compatible results.

Table 5.10 – Fraction of HasBremAdded,  $f_{brem}$ , in data and MC for  $B^+ \rightarrow J/\psi (\rightarrow e^+ e^-) K$ . Both reweighted and unweighted MC is in agreement with data.

|                     | <b>Data</b>          | <b>MC (unweighted)</b> | <b>MC (reweighted (nSPDHits))</b> |
|---------------------|----------------------|------------------------|-----------------------------------|
| $e^+$               | $(48.00 \pm 0.20)\%$ | $(46.9 \pm 0.5)\%$     | $(48.0 \pm 0.5)\%$                |
| $e^-$               | $(47.89 \pm 0.19)\%$ | $(47.3 \pm 0.5)\%$     | $(47.6 \pm 0.5)\%$                |
| $\langle e \rangle$ | $(47.94 \pm 0.17)\%$ | $(47.1 \pm 0.4)\%$     | $(47.8 \pm 0.4)\%$                |

## 5.3 Backgrounds

The study of possible backgrounds is of central importance as these may alter in different possible ways the signal yield.

Three main sources of background are considered for this analysis:

- combinatorial background: candidates formed by random combination of tracks;
- candidates formed by tracks from other decays where the final particles are mis-identified;
- candidates formed by tracks from other decays where some of the final particles are not reconstructed and the remaining ones are the same as for the signal.

The combinatorial background is rejected by the multivariate classifier described in Section 5.1.5 and the remaining candidates are modeled in the fit by an exponential function. As the BDT is trained to distinguish combinatorial background from signal, the amount and shape of the combinatorial are different depending on the BDT bin. Therefore, independent parameters are used for the exponential for each bin.

The amount of background candidates from other decays that is left after full selection is studied in the following sections. The backgrounds that peak in the 2-body invariant mass are

found to be negligible, while a non-negligible amount of events remains, in particular from semileptonic decays where a neutrino is not reconstructed. These backgrounds do not peak in the invariant mass but they have a decaying shape the search window. They are parametrised in the fit model using non-parametric models: RooKeysPdfs [113].

### 5.3.1 Peaking backgrounds

Peaking backgrounds are candidates from other decays than the signal where one or more particles are lost or mis-identified. A particularly dangerous category of decays that is considered is  $B_{(s)}^0 \rightarrow h^+ h^-$  decays, as they would peak in invariant mass under the signal. Furthermore, semileptonic decays of  $B$  mesons and  $\Lambda_b^0$  are considered. In this case a neutrino is not reconstructed and therefore these candidates are not peaking in invariant mass and they have a broad shape, but they can still pollute the sample. Finally,  $B_c^+$  and  $B^+$  decays to  $J/\psi$  where one or more particles are not reconstructed or misidentified are considered.

The expected yields of exclusive decays are shown in the following sections. In particular, for the  $B_{(s)}^0 \rightarrow h^+ h^-$  decays they are estimated using two different methods: first using simulation re-normalised to  $B^+ \rightarrow J/\psi K^+$  data and, secondly, rescaling the yield of  $B_{(s)}^0 \rightarrow h^+ h^-$  candidates found in data with a dedicated selection.

#### Expected $B_{(s)}^0 \rightarrow h^+ h^-$ yields normalising with respect to $B^+ \rightarrow J/\psi K^+$

Simulated samples can be used to estimate the selection efficiencies but, as an arbitrary number of events can be generated, it is needed to re-normalise the results using a channel which is well understood in data. The  $B^+ \rightarrow J/\psi K^+$  decay is a clean and abundant channel, ideal to be used as normalisation.

The mis-identification probabilities for the considered decays are reported in Table 5.11. All wrong ID combinations are considered when more than one is possible. For example, the  $B^0 \rightarrow K^+ \pi^-$  decay can be mis-identified as  $B_s^0 \rightarrow e^\pm \mu^\mp$  if the following two combinations occur:  $K \rightarrow e$  and  $\pi \rightarrow \mu$  or  $K \rightarrow \mu$  and  $\pi \rightarrow e$ . In this example the total mis-identification probability is calculated as  $\varepsilon^{tot} = \varepsilon(K \rightarrow e)\varepsilon(\pi \rightarrow \mu) + \varepsilon(K \rightarrow \mu)\varepsilon(\pi \rightarrow e)$ . Note that no PID requirement is applied to select the  $B^+ \rightarrow J/\psi K^+$  sample (except for the isMuon flag).

The expected numbers of candidates after the full  $B_s^0 \rightarrow e^\pm \mu^\mp$  selection are reported in Table 5.12 together with the parameters used for the estimation: the branching ratios of the decays, the fragmentation fractions and the total efficiencies.

The total estimated amount of  $B_{(s)}^0 \rightarrow h^+ h^-$  candidates passing the full selection is  $0.07 \pm 0.02$ .

Table 5.11 – Mis-identification probabilities for candidates from various  $B_{(s)}^0 \rightarrow h^+ h^-$  decays to be reconstructed as  $B_s^0 \rightarrow e^\pm \mu^\mp$ . The probabilities for the  $B_s^0$  decays are assumed to be the same, as the kinematics of particles from 2-body  $B^0$  and  $B_s^0$  decays are very similar.

| Decay                            | Mis-identification probability ( $\times 10^{-4}$ ) |
|----------------------------------|-----------------------------------------------------|
| $B_{(s)}^0 \rightarrow K\pi$     | $0.41 \pm 0.09$                                     |
| $B_{(s)}^0 \rightarrow \pi\pi$   | $1.60 \pm 0.26$                                     |
| $B_{(s)}^0 \rightarrow KK$       | $0.16 \pm 0.04$                                     |
| $B_{(s)}^0 \rightarrow p\bar{p}$ | $0.05 \pm 0.49$                                     |

### Expected $B_{(s)}^0 \rightarrow h^+ h^-$ yields obtained from data

In order to cross-check the results from the previous section the estimation is also performed using as a starting point a  $B_{(s)}^0 \rightarrow h^+ h^-$  data sample where the  $B_s^0 \rightarrow e^\pm \mu^\mp$  selection is only partially applied. This method allows to have a sizeable peak to fit which can be then corrected for the difference between the partial and the full selection.

On top of the stripping, a  $\text{BDT}_{flat}^{sig} > 0.5$  cut is applied. The same selection as for the signal is then applied, except for the particle identification criteria: one of the two child particles is required to pass the same electron PID requirement used for the signal, while the second one must fulfil a pion requirement. This means that we are effectively building  $B^0 \rightarrow \pi e$  candidates. The PID requirements used are:

$$p_1 \rightarrow (\text{DLLe} > 5.5 \ \&\& \ \text{ProbNNK} < 0.95) \text{ and} \quad (5.11)$$

$$p_2 \rightarrow \text{ProbNNpi} \cdot (1 - \text{ProbNNK}) \cdot (1 - \text{ProbNNp}) > 0.5. \quad (5.12)$$

Table 5.12 – Branching ratio, total efficiency and number of expected background events

| Decay                        | BR ( $\times 10^{-5}$ ) | $f_q/f_u$       | Efficiency ( $\times 10^{-8}$ ) | Expected events       |
|------------------------------|-------------------------|-----------------|---------------------------------|-----------------------|
| $B^0 \rightarrow K\pi$       | $1.96 \pm 0.05$         | 1.00            | $0.190 \pm 0.044$               | $0.032 \pm 0.007$     |
| $B^0 \rightarrow \pi\pi$     | $0.51 \pm 0.02$         | 1.00            | $0.655 \pm 0.121$               | $0.029 \pm 0.005$     |
| $B^0 \rightarrow KK$         | $0.008 \pm 0.002$       | 1.00            | $0.066 \pm 0.017$               | $0.00007 \pm 0.00003$ |
| $B^0 \rightarrow p\bar{p}$   | $0.0015 \pm 0.0007$     | 1.00            | $0.016 \pm 0.165$               | $0.0000 \pm 0.0002$   |
| $B_s^0 \rightarrow K\pi$     | $0.56 \pm 0.06$         | $0.27 \pm 0.02$ | $0.190 \pm 0.044$               | $0.0023 \pm 0.0006$   |
| $B_s^0 \rightarrow \pi\pi$   | $0.07 \pm 0.01$         | $0.27 \pm 0.02$ | $0.655 \pm 0.121$               | $0.0011 \pm 0.0004$   |
| $B_s^0 \rightarrow KK$       | $2.54 \pm 0.16$         | $0.27 \pm 0.02$ | $0.066 \pm 0.017$               | $0.004 \pm 0.001$     |
| $B_s^0 \rightarrow p\bar{p}$ | $0.003 \pm 0.002$       | $0.27 \pm 0.02$ | $0.016 \pm 0.165$               | $0.00000 \pm 0.00001$ |

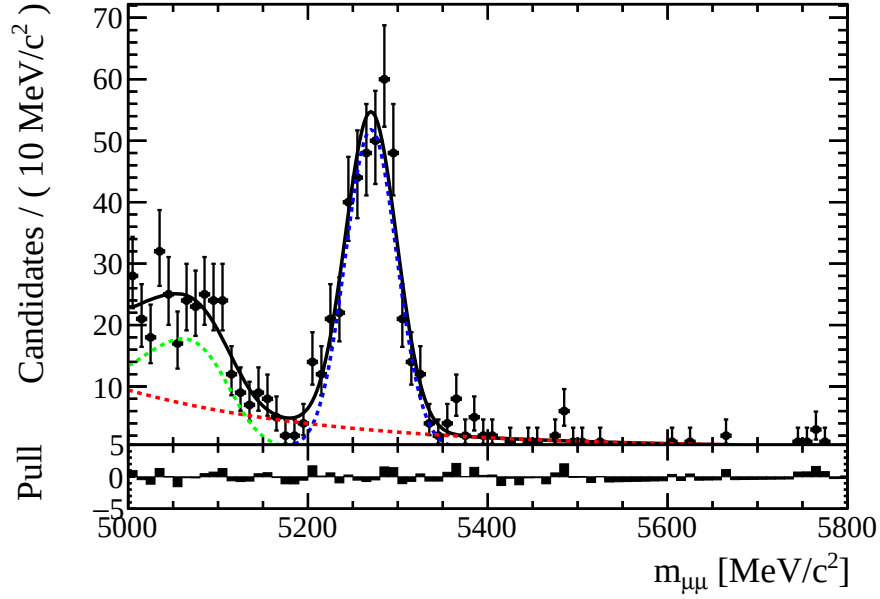


Figure 5.14 – Invariant-mass distribution of  $B^0 \rightarrow \pi e$  candidates with the fit superimposed.  $m_{\mu\mu}$  indicates the  $B$  invariant mass, calculated in the muon mass hypothesis for both child particles.

The expected number of candidates for each channel is then calculated by multiplying the obtained yield by the probability of misidentifying also the second hadron. The following formula is used:

$$N_X = N_{B^0 \rightarrow \pi e} \cdot c_X \cdot \frac{\varepsilon^{strip}(X)}{\varepsilon^{strip}(B^0 \rightarrow \pi e)} \cdot \frac{\varepsilon^{PID}(h \rightarrow \mu)}{\varepsilon^{PID}(h \rightarrow \pi)}, \quad (5.13)$$

where  $\varepsilon^{strip}$  is the stripping efficiency and  $c_X$  is a factor that takes into account that the fitted sample contains various channels in different proportions. This factor is defined as

$$c_X = \frac{(f_q)_X \cdot \mathcal{B}_X \cdot \varepsilon_X^{PID}(h_1 h_2 \rightarrow \pi e)}{\sum (f_q)_k \cdot \mathcal{B}_k \cdot \varepsilon_k^{PID}(h_1 h_2 \rightarrow \pi e)}. \quad (5.14)$$

The number of  $B^0 \rightarrow \pi e$  candidates is obtained fitting the invariant-mass distribution of the sample which is reported in Figure 5.14. The PDF used for the fit includes a Gaussian for the  $B^0 \rightarrow \pi e$  yield, an Argus [114] function convoluted with a Gaussian for the partially-reconstructed background and an exponential for the combinatorial background. All parameters are floating in the fit.

The expected numbers of candidates for each decay considered are reported in Table 5.13 and the total number of  $B_{(s)}^0 \rightarrow h^+ h^-$  decays into the  $B_s^0 \rightarrow e^\pm \mu^\mp$  sample is expected to be  $0.11 \pm 0.02$ .

Table 5.13 –  $B \rightarrow hh'$  fraction  $c_X$ , total efficiency and number of expected peaking background candidates.

| Decay                        | $c_X$           | Eff. ratio ( $\times 10^{-4}$ ) | Expected events       |
|------------------------------|-----------------|---------------------------------|-----------------------|
| $B^0 \rightarrow K\pi$       | $0.34 \pm 0.03$ | $1.814 \pm 0.230$               | $0.044 \pm 0.008$     |
| $B^0 \rightarrow \pi\pi$     | $0.60 \pm 0.06$ | $1.201 \pm 0.147$               | $0.052 \pm 0.009$     |
| $B^0 \rightarrow KK$         | $< 0.01$        | $10.768 \pm 1.464$              | $0.00015 \pm 0.00007$ |
| $B^0 \rightarrow p\bar{p}$   | $< 0.01$        | $1.706 \pm 2.999$               | $0.0003 \pm 0.0006$   |
| $B_s^0 \rightarrow K\pi$     | $0.02 \pm 0.00$ | $1.814 \pm 0.230$               | $0.0032 \pm 0.0007$   |
| $B_s^0 \rightarrow \pi\pi$   | $0.02 \pm 0.01$ | $1.201 \pm 0.147$               | $0.0020 \pm 0.0006$   |
| $B_s^0 \rightarrow KK$       | $0.01 \pm 0.01$ | $10.768 \pm 1.464$              | $0.008 \pm 0.002$     |
| $B_s^0 \rightarrow p\bar{p}$ | $< 0.01$        | $1.706 \pm 2.999$               | $0.00002 \pm 0.00003$ |

### Expected $B_{(s)}^0 \rightarrow h^+ h^-$ yields: conclusion

The two methods described are found to be in agreement and predict that only  $\sim 0.1$  background candidates will fall into the  $B_s^0 \rightarrow e^\pm \mu^\mp$  sample. As this number is significantly lower than 1 it is concluded that the contribution from these decays is negligible and therefore there is no need to add a component to model them in the fit.

### Partially-reconstructed backgrounds

The number of candidates from semileptonic decays and decays including a  $J/\psi$  which fall into the  $B_s^0 \rightarrow e^\pm \mu^\mp$  sample are estimated using the same method described in Section 5.3.1; namely, from simulation re-normalised using the yield of  $B^+ \rightarrow J/\psi K^+$  candidates observed in data. This study is performed in the invariant mass window  $[4900 - 6000] \text{ GeV}/c^2$  which is the one used for the fit.

Expected numbers of candidates are reported in Table 5.14 together with the parameters used for the estimations: the branching ratios of the decays, the fragmentation fractions and the total efficiencies. The efficiencies split in their PID, trigger and reconstruction components are reported in Appendix D.

Table 5.14 only reports channels for which at least one simulated event passes the full selection, which is not the case for  $B^+ \rightarrow J/\psi (ee)K$  and  $\Lambda_b^0 \rightarrow p e \nu$ . For these decays an evaluation is done of how many events are expected after stripping before BDTS, trigger and PID requirements. For  $B^+ \rightarrow J/\psi (ee)K \sim 80$  events are expected, all of which fall below  $4.4 \text{ GeV}/c^2$  in  $m(e\mu)$ , far from the mass window used for the fit. This background would therefore not significantly pollute the analysis sample. For  $\Lambda_b^0 \rightarrow p e \nu$ ,  $22 \pm 11$  events are expected to fall into the analysed mass window after stripping. This amount would be reduced especially by the fact that the proton must be misidentified as a muon. The maximum probability for this mis-identification is 14%, only in a narrow region of the phase space and  $< 1\%$  in most of the phase-space and the

trigger efficiency for similar decays is  $\sim 35\%$ . Therefore, even considering that the maximum mis-ID rate would hold everywhere, this brings the number of expected events to  $\sim 1$  which would be further reduced by the BDTs requirements. Therefore the contribution due this channel is negligible, especially when compared to the amounts of other backgrounds, and any possible contribution can be safely included in the combinatorial exponential.

In summary, the only relevant contributions are from  $\Lambda_b^0 \rightarrow p\mu\nu$  and  $B^0 \rightarrow \pi\mu\nu$ . Figure 5.15 shows the invariant-mass distributions of the candidates passing the selection, which are peaking at low masses, far from the signal peak.

These two channels are taken into account as an additional contributions in the mass fit for the limit determination modeled using RooKeysPdfs fitted on simulation. Figure 5.16 shows these distributions for the third BDT bin, as an example.

Table 5.14 – Branching ratio, total efficiency (including geometric) and number of expected partially reconstructed background events.

| Decay                                  | BR ( $\times 10^{-5}$ ) | $f_q/f_u$ | Efficiency ( $\times 10^{-8}$ ) | Expected events     |
|----------------------------------------|-------------------------|-----------|---------------------------------|---------------------|
| $B^0 \rightarrow \pi\mu\nu$            | $14.50 \pm 0.50$        | 1.00      | $44.587 \pm 0.898$              | $55.2 \pm 2.9$      |
| $B^0 \rightarrow \pi e\nu$             | $14.50 \pm 0.50$        | 1.00      | $0.1064 \pm 0.0245$             | $0.13 \pm 0.03$     |
| $B_c^+ \rightarrow J/\psi(\mu\mu)e\nu$ | $0.31 \pm 0.14$         | 0.01      | $11.170 \pm 3.044$              | $0.0015 \pm 0.0008$ |
| $B_c^+ \rightarrow J/\psi(ee)\mu\nu$   | $0.31 \pm 0.14$         | 0.01      | $10.625 \pm 2.949$              | $0.0014 \pm 0.0008$ |
| $\Lambda_b^0 \rightarrow p\mu\nu$      | $41.00 \pm 10.00$       | 0.81      | $29.141 \pm 11.081$             | $82.4 \pm 38.7$     |
| $B^+ \rightarrow J/\psi(\mu\mu)K$      | $6.115 \pm 0.18$        | 1.00      | $0.207 \pm 0.120$               | $0.11 \pm 0.06$     |

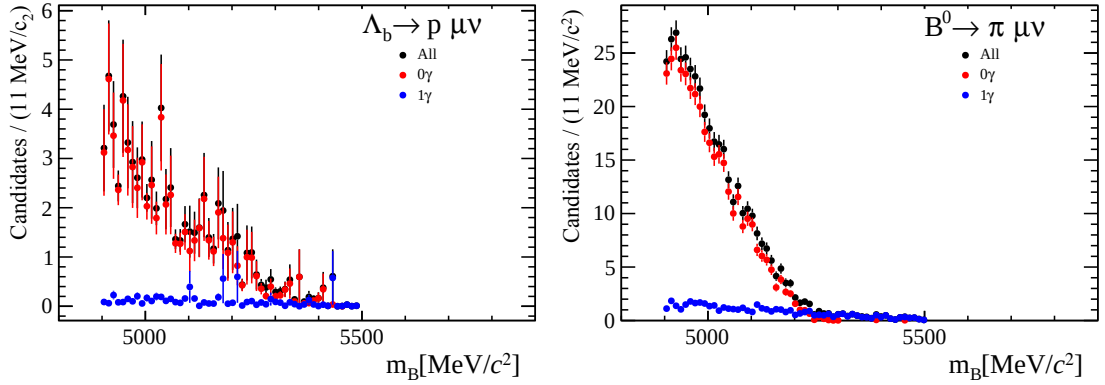


Figure 5.15 – Invariant-mass distributions of  $\Lambda_b^0 \rightarrow p\mu\nu$  (left) and  $B^0 \rightarrow \pi\mu\nu$  (right) candidates passing the  $B_s^0 \rightarrow e^\pm \mu^\mp$  full selection. The plots show separately candidates with one bremsstrahlung photon emitted by the particle reconstructed as an electron ( $1\gamma$ ) and no photons emitted ( $0\gamma$ ) together with their sum.

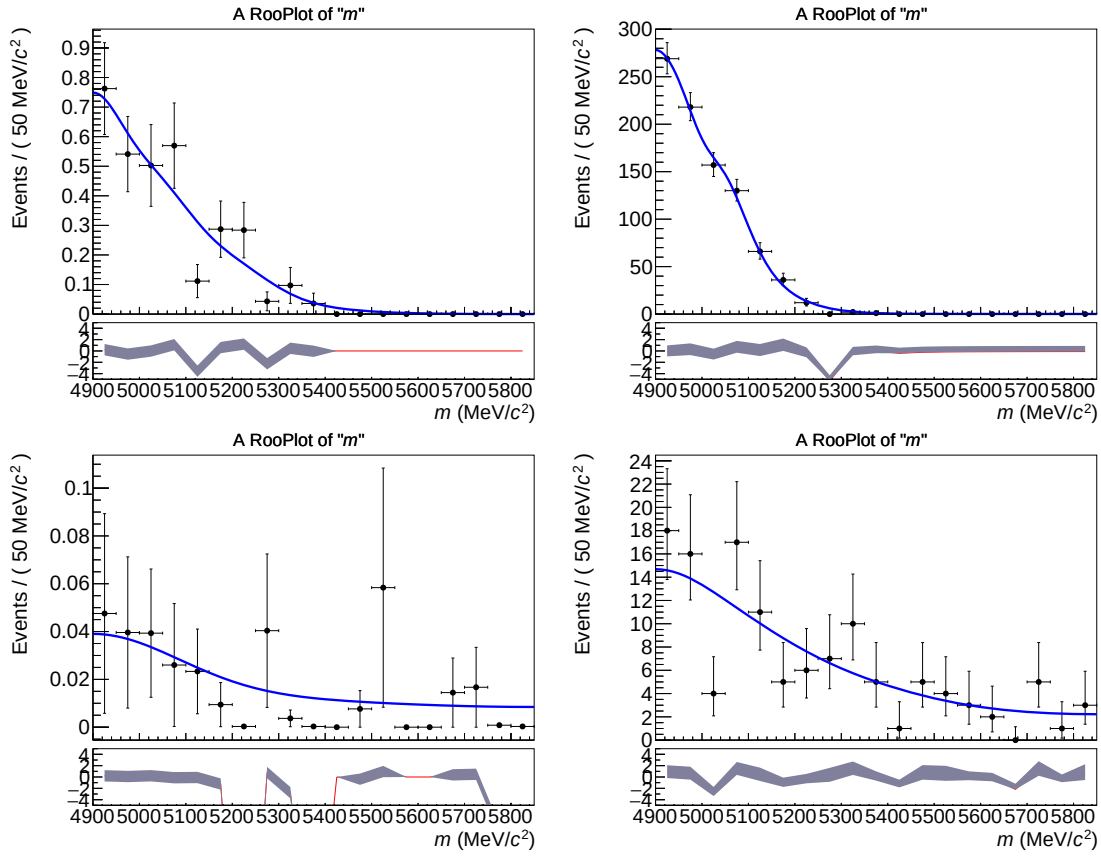


Figure 5.16 –  $\Lambda_b^0 \rightarrow p\mu\nu$  (left) and  $B^0 \rightarrow \pi\mu\nu$  (right) simulated candidates passing the signal selection, in the BDT range  $[0.4, 0.5]$  for events with  $\text{HasBremAdded} = 0$  (top) and  $\text{HasBremAdded} = 1$  (bottom).



## 5.4 Normalisation

In order to extract the branching fraction of the signal, the number of observed  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates,  $N_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}$ , has to be compared to the number of candidates of a *calibration channel*, with a well-known branching fraction,  $N_{\text{cal}}$ . The decays  $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$  and  $B^0 \rightarrow K^+ \pi^-$  are chosen as calibration channels respectively for the clean signature and for the topological similarity to the signal. The branching fraction is calculated as

$$\mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) = \frac{f_{\text{cal}}}{f_{d,s}} \cdot \frac{\epsilon_{\text{cal}}}{\epsilon_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}} \cdot \frac{N_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}}{N_{\text{cal}}} \mathcal{B}_{\text{cal}}, \quad (5.15)$$

where  $f_x$  is the hadronisation fraction of a  $b$  quark with  $x \in \{u, d, s, c\}$ ,  $\epsilon_y$  is the efficiency for the detection of decay channel  $y \in \{\text{cal}, B^0 \rightarrow e^\pm \mu^\mp, B_s^0 \rightarrow e^\pm \mu^\mp\}$  and  $\mathcal{B}_{\text{cal}}$  is the branching fraction of the normalisation channel. With this formula the one-event sensitivity is defined as,

$$\alpha_{\text{cal}} \equiv \frac{f_{\text{cal}}}{f_{d,s}} \cdot \frac{\epsilon_{\text{cal}}}{\epsilon_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}} \cdot \frac{\mathcal{B}_{\text{cal}}}{N_{\text{cal}}}. \quad (5.16)$$

The  $\alpha_{\text{cal}}$  from the two normalisation channels are then combined into a single factor, averaging them with weights proportional to the inverse of their uncertainties. The determination of the efficiencies is discussed in the next section. The yields of the normalisation channels,  $N_{\text{cal}}$ , are determined using maximum-likelihood fits to their respective invariant-mass spectra, given in Section 5.4.1. Finally, the normalisation factors, *i.e.* the single-event sensitivities, are given in Section 5.4.2.

### 5.4.1 Invariant-mass fits

This section describes the maximum-likelihood fits to the invariant-mass used to determine the yields of the normalisation channels. The fits can be observed in Figure 2 of the published paper, at the beginning of this chapter.

#### $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$

Due to its abundance, cleanliness and to its similar trigger selection to the signal, the  $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$  is an ideal normalisation channel. The invariant mass  $m_{J/\psi K}$  is modeled by an Hypatia distribution [115], with tail parameters fixed from simulation and mean, width and  $\lambda$  floating in the fit. Combinatorial background is described by an exponential PDF, while partially-reconstructed decays are cut away by requiring  $m_{J/\psi K} > 5180 \text{ MeV}$ .

An additional mis-ID background component from  $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \pi^+$  ( $\mathcal{B}$  about 3% of the signal) is included in the fit.

## Chapter 5. Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

Table 5.15 reports the yield from the fit, along with the PID and selection corrections.

Table 5.15 – Yield (N) of the  $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$  normalisation channel and corresponding selection efficiencies, for Run I data. Each efficiency is calculated given all the previous cuts. Uncertainties are statistical only.

|                            |                    |
|----------------------------|--------------------|
| N                          | $913074 \pm 1106$  |
| $\epsilon_{gen}$           | $16.63 \pm 0.07\%$ |
| $\epsilon_{Reco\&Sel gen}$ | $18.01 \pm 0.01\%$ |
| $\epsilon_{PID Reco\&Sel}$ | $97.81 \pm 0.02\%$ |
| $\epsilon_{trigger PID}$   | $75.8 \pm 1.0\%$   |

### $B^0 \rightarrow K^+ \pi^-$

The  $B^0 \rightarrow K^+ \pi^-$  decay is an optimal normalisation channel thanks to its topology similar to the signal.

The yield and its uncertainty are obtained with the same strategy used for the BDT calibration, described in Section 5.1.6, with a mass fit performed on the full BDT range [0, 1] and for  $|\Delta LL_{K-\pi}| < 10$ . The trigger selection, however, differs from the one used for the calibration of the BDT shape, as indicated in Section 5.1.1.

Table 5.16 reports the yield from the fit, along with the PID and selection corrections.

Table 5.16 – Yield (N) of the  $B^0 \rightarrow K^+ \pi^-$  normalisation channel and corresponding selection efficiencies, for Run I data. Each efficiency is calculated given all the previous cuts. Uncertainties are statistical only.

|                            |                        |
|----------------------------|------------------------|
| N                          | $49907 \pm 277$        |
| $\epsilon_{gen}$           | $18.97 \pm 0.05\%$     |
| $\epsilon_{Reco\&Sel gen}$ | $23.15 \pm 0.06\%$     |
| $\epsilon_{PID Reco\&Sel}$ | $38.5046 \pm 0.0001\%$ |
| $\epsilon_{L0 PID}$        | $31.0 \pm 0.2\%$       |
| $\epsilon_{HLT1 L0}$       | $83.5 \pm 0.4\%$       |
| $\epsilon_{HLT2 HLT1}$     | $82.0 \pm 0.3\%$       |

The negligible uncertainty on the PID efficiency is due to the high statistics of the PIDCalib sample. The systematic error, computed as in Section 5.1.6 is included in the systematic uncertainty.

For the computation of the trigger efficiencies, the following approach is used:

- L0 and HLT1 efficiencies are obtained from  $B^0 \rightarrow K^+ \pi^-$  data using the TIS-TOS method (see Section 3.5.2). For what concerns L0, the TIS L0MuonDecision line is used. The result has been cross-checked using L0HadronDecision, and no large difference was

observed. For HLT1 the Hlt1TrackAllL0Decision\_TIS line is used.

- Hlt2 efficiency is obtained from simulation, after re-weighting for the PID efficiency obtained with PIDCalib, to reproduce the correct effect of the particle identification requirements.

Systematic uncertainties are treated in the same way as in Section 5.1.6. In this case, the systematic uncertainty dominates over the statistical one.

### Cross-check between the two normalisation channel

As a cross-check for possible mis-estimations of the yields and efficiencies of the normalisation channels, the ratio between the measured branching ratios of the two normalisation channels,  $R_{norm}$ , is compared with the one obtained from the branching fractions in the PDG [2].

The measured value, combining the results from Sections 5.4.1 and 5.4.1 is:

$$R_{norm} = \frac{N_{B^0 \rightarrow K\pi} \times \epsilon_{B^+ \rightarrow J/\psi K}^{sel} \times \epsilon_{B^+ \rightarrow J/\psi K}^{gen} \times \mathcal{L}_{B^+ \rightarrow J/\psi K}}{N_{B^+ \rightarrow J/\psi K} \times \epsilon_{B^0 \rightarrow K\pi}^{sel} \times \epsilon_{B^0 \rightarrow K\pi}^{gen} \times \mathcal{L}_{B^0 \rightarrow K\pi}} = 0.332 \pm 0.020 \quad (5.17)$$

where  $\mathcal{L}$  is the luminosity of the data used in the analysis, which differ slightly between the two samples.

This result is in excellent agreement with the fraction of branching ratios obtained from the PDG:  $R_{norm}^{PDG} = 0.321 \pm 0.013$ .

As a further crosscheck, the  $B^0 \rightarrow K^+ \pi^-$  yield is measured for different DLL cuts, similarly to what is done in Section 5.1.6, and the result, corrected for the PID efficiency, is found to be stable within less than 1%.

### 5.4.2 Normalisation factors

Table 5.17 summarises the global selection efficiencies, while Table 5.18 shows the branching fractions and the fragmentation fraction entering the computation of the normalisation factor. The single-events sensitivity values ( $\alpha$ ), which are the weighted average between the two normalisation channels, are given in Table 5.19.

## 5.5 Upper limit determination

The CLs method [117] is used to compute the branching ratio upper limits for  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  decays.

The method consists in generating multiple replicas of two types of simulated datasets: one

Table 5.17 – Total selection efficiencies for signal and normalisation channels.

| Channel                                                |                       | efficiency              |
|--------------------------------------------------------|-----------------------|-------------------------|
| $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$                  | with brem recovery    | $(1.354 \pm 0.034)\%$   |
| $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$                  | without brem recovery | $(0.935 \pm 0.021)\%$   |
| $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ |                       | $(2.30 \pm 0.04)\%$     |
| $B^0 \rightarrow K^+ \pi^-$                            |                       | $(0.3598 \pm 0.0034)\%$ |

Table 5.18 – Input for normalisation: branching fractions and fragmentation factor.

| Parameter                                                           | value                                |
|---------------------------------------------------------------------|--------------------------------------|
| $\mathcal{B}(B^0 \rightarrow K^+ \pi^-)$                            | $(1.96 \pm 0.05) \times 10^{-5}$ [2] |
| $\mathcal{B}(B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+)$ | $(6.10 \pm 0.20) \times 10^{-5}$ [2] |
| $\frac{f_s}{f_d}$                                                   | $0.256 \pm 0.015$ [116]              |

Table 5.19 –  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  single-event sensitivities.

| Channel                           | $\alpha$                          |
|-----------------------------------|-----------------------------------|
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | $(2.45 \pm 0.17) \times 10^{-10}$ |
| $B^0 \rightarrow e^\pm \mu^\mp$   | $(6.16 \pm 0.23) \times 10^{-11}$ |

containing only background, and one containing background plus signal at a given branching fraction. The datasets are used to build the distributions of the test statistics  $-2\ln Q$  respectively for the two hypotheses: *Background-only* and *Signal + Background*.

The test-statistic  $Q$  at a given branching fraction  $\mathcal{B} = F$  is defined as:

$$Q = \frac{\mathcal{L}_{S+B}(\mathcal{B} = F)}{\mathcal{L}_B(\mathcal{B} = 0)}, \quad (5.18)$$

where  $\mathcal{L}_{S+B}$  and  $\mathcal{L}_B$  are the Likelihood functions respectively using the *Signal + Background* and the *Background-only* model.

The two distributions obtained for  $Q$ ,  $\mathcal{Q}_{S+B}$  and  $\mathcal{Q}_B$ , are then compared with the value of  $Q$  measured on the real data,  $Q_{obs}$ , by computing the two quantities

$$CL_{s+b} = \int_{-\infty}^{Q_{obs}} \mathcal{Q}_{S+B} dQ \quad (5.19)$$

and

$$CL_b = \int_{-\infty}^{Q_{obs}} \mathcal{Q}_B dQ. \quad (5.20)$$

$\text{CL}_s$  is then obtained as

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b}. \quad (5.21)$$

The expected limit is obtained similarly, by replacing  $Q_{obs}$  with the expected value from  $\mathcal{Q}_B$ . This procedure is repeated for different hypothetical values of the signal branching fraction, obtaining the  $\text{CL}_s$  scan.

The implementation of the CLs method provided by the RooStats [118] package is used to compute the limit.

## 5.6 Results

The measured branching fractions obtained from the best fit to the data are reported in Table 5.20. For the  $B_s^0$ , both the light and heavy mass-eigenstate hypotheses are considered, as the non-negligible lifetime difference between the two eigenstates affects the selection efficiencies.

Table 5.20 –  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  invariant-mass fit results.

| Channel                                        | light mass eigenstate          | heavy mass eigenstate          |
|------------------------------------------------|--------------------------------|--------------------------------|
| $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$ | $(1.4 \pm 1.3) \times 10^{-9}$ | $(1.3 \pm 1.2) \times 10^{-9}$ |
| $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp)$   | $(0 \pm 1.7) \times 10^{-9}$   |                                |

A mild excess is observed in the  $B_s^0 \rightarrow e^\pm \mu^\mp$  yield, with a significance of about  $1\sigma$ .

The upper limits for the branching fractions are reported in Table 5.21 and 5.22, where for the  $B_s^0$ , both light and heavy mass-eigenstate hypotheses are considered.

Table 5.21 – Expected and observed upper limits for  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$  at 90%(95%) CL in light and heavy mass-eigenstate hypotheses.

| channel                                        | light mass eigenstate     |                           | heavy mass eigenstate     |                           |
|------------------------------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
|                                                | expected                  | observed                  | expected                  | observed                  |
| $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp)$ | $4.5(5.6) \times 10^{-9}$ | $6.0(7.2) \times 10^{-9}$ | $3.9(5.0) \times 10^{-9}$ | $5.4(6.3) \times 10^{-9}$ |

Table 5.22 – Expected and observed upper limits for  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp)$  at 90%(95%) CL.

| channel                                      | expected                    | observed                  |
|----------------------------------------------|-----------------------------|---------------------------|
| $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp)$ | $0.91(1.24) \times 10^{-9}$ | $1.0(1.3) \times 10^{-9}$ |

The CLs scans are shown in the published paper, reported at the beginning of this chapter.

## 5.7 Systematic errors

The systematic uncertainties are accounted for by adding Gaussian constraints in the final fit, allowing each parameter to fluctuate with a standard deviation corresponding to its uncertainty. The components are summarised here and organised according to their origin. Each contribution is referred to the parameter itself and not on its effect on the estimated limit.

### Due to MC and calibration samples statistics

- Signal efficiency for each bremsstrahlung category, the maximum systematic error amounts to  $\sim 0.5\%$ ;
- Mass window cut,  $0.2\%$ ;
- BDT shape correction for each bremsstrahlung category (see Section 5.1.6), order of  $3\%$ ;
- Bremsstrahlung fraction,  $0.5\%$ ;
- Exclusive background yields, order of  $40\%$ ;

### Inputs from outside

- $\frac{f_s}{f_d}$ ,  $6\%$ ;
- Branching fraction of the normalisation channels,  $2.6\%$  for  $B^0 \rightarrow K^+ \pi^-$  and  $3.3\%$  for  $B^+ \rightarrow J/\psi K^+$ ;

### Fits to data

- Signal shape parameters for each bremsstrahlung category, where the dominant one is on the power of the tail parameter which is of the order of  $20\%$ ;
- Normalisation yield, on  $B^0 \rightarrow K^+ \pi^-$  is  $6\%$  and on  $B^+ \rightarrow J/\psi K^+$  is  $\sim 0.8\%$ ;
- BDT fractions as determined in Section 5.1.6, order of  $25\%$ ;
- PID efficiencies,  $\sim 0.6\%$ ;
- Trigger efficiencies, order of  $3\%$ ;
- Correction factor to the mass resolution described in Section 5.2.1, order of  $5\%$ ;

The impact of the two largest systematics, *i.e.* exclusive background yield and the BDT shape uncertainties ( $40\%$  and  $25\%$  respectively) on the upper limits are evaluated by re-computing them fixing these parameters and comparing with the nominal limit. The effects are found to be  $4\%$  and  $5\%$  respectively for the two sources of systematic errors.

## 5.8 Conclusions

In summary, a search for the LFV decays  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  is performed in  $pp$  collision data corresponding to a total integrated luminosity of  $3 \text{ fb}^{-1}$ . No excesses are observed for both  $B_d^0$  and  $B_s^0$ , and upper limits are evaluated.

For  $B_s^0$  they were found to be  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 6.0(7.2) \times 10^{-9}$  at 90(95)% CL for the light mass eigenstate hypothesis, and  $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 5.4(6.3) \times 10^{-9}$  at 90(95)% CL for the heavy mass eigenstate hypothesis. For  $B^0$  the limit is  $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 1.0(1.3) \times 10^{-9}$ .

These results represent the world's best upper limits on the branching fractions of LFV decays from  $B$  mesons to date, as shown in Figure 5.17.

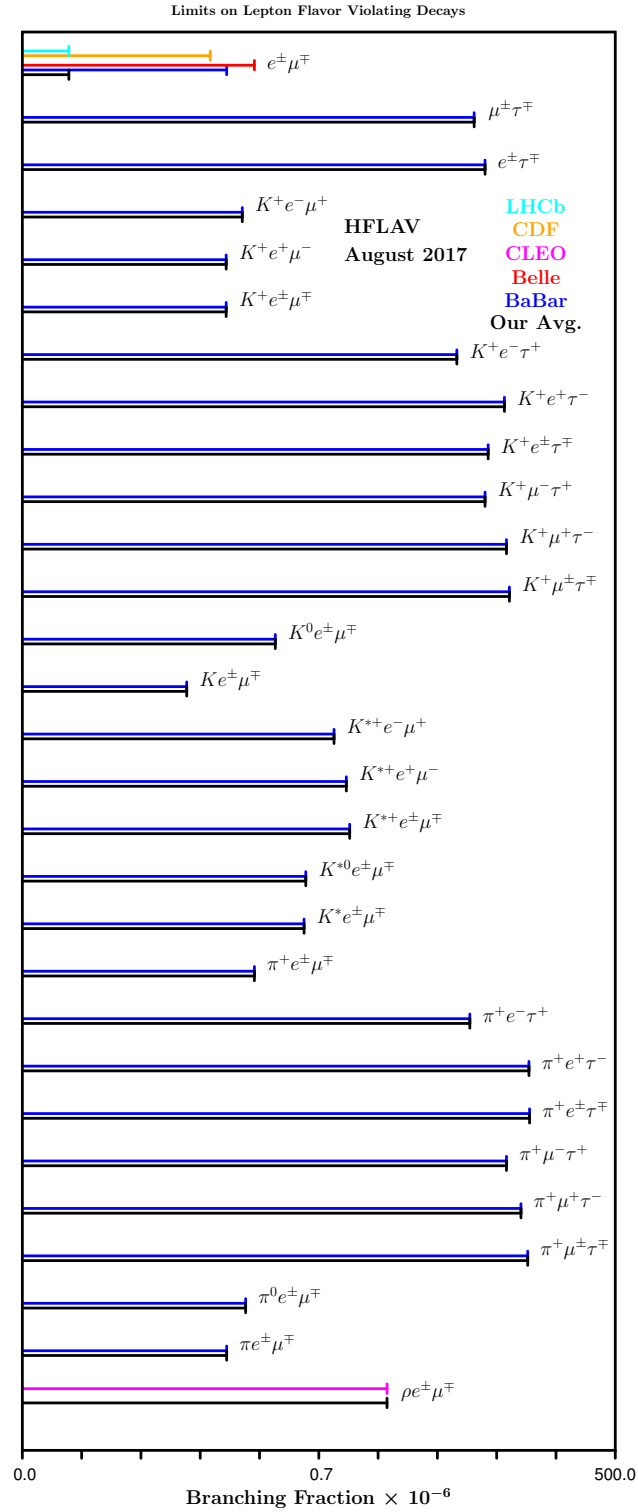


Figure 5.17 – Limits on Lepton Flavour Violating decays, updated in August 2017. Figure from the Heavy Flavor Averaging Group Report [119].



## 6 Search for the lepton-flavour violating decay $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$

Lepton Flavour Violation is not restricted to the meson sector: baryons could participate to new physics processes as well. Being fermions, the baryons obey to different Lagrangian interaction terms and have a different (half-integer) spin, which generates decays with different angular structures. They therefore can provide complementary and independent information on new physics phenomena.

The  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  process, with  $\Lambda^0 \rightarrow p\pi^-$ , is an ideal candidate for studying LFV in the baryon sector. Such decay would indeed provide a clean signature in the detector, with four charged tracks and a detached vertex from the  $\Lambda_b^0$  decay ( $\tau \simeq 1.5 \times 10^{-12}$  s [2]). Moreover, only very few SM processes can mimic the signal, generating backgrounds.

While being practically (except for high-order diagrams with neutrino oscillations) zero in the SM, the branching fraction of the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  would be significantly enhanced in alternative models in which the existence of new mediators contribute to the process. In some of these models the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  branching fraction reach  $\mathcal{O}(10^{-9})$  [120], becoming potentially accessible to experimental searches.

At LHC, the production of  $\Lambda_b^0$  baryons is abundant [121], and LHCb provides a unique opportunity to perform such study, which has never been explored before.

The present chapter contains a description of the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis in LHCb, performed in the context this thesis. The study is still ongoing, and potentially subject to minor changes; for this reason, the signal mass region is kept blind until the analysis is finalised and approved in the LHCb collaboration.

A description of the analysis strategy and a projection of the expected results is provided here, without repeating in detail aspects that are shared with the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis, described in the previous chapter.

Table 6.1 – Luminosity recorded at the LHCb experiment and centre-of-mass energy of the  $pp$  collisions in each year included in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis data sample.

| Run    | Year | Energy (TeV) | Luminosity ( $\text{fb}^{-1}$ ) |
|--------|------|--------------|---------------------------------|
| Run I  | 2011 | 7            | $\sim 1.11$                     |
|        | 2012 | 8            | $\sim 2.08$                     |
| Run II | 2015 | 13           | $\sim 0.33$                     |
|        | 2016 | 13           | $\sim 1.67$                     |
|        | 2017 | 13           | $\sim 1.61$                     |

## 6.1 Data and simulated samples

The measurement is performed using  $pp$  collision data collected by the LHCb experiment during the Run 1 and Run 2 of LHC, up to year 2017. Monte Carlo samples reproducing the data-taking conditions are used to simulated the signal and the exclusive backgrounds.

### 6.1.1 Data

Table 6.1 summarises the recorded luminosity and centre-of-mass energy for each of the years included in the present analysis' data sample.

The data are selected by multi-purpose stripping algorithms (described in Section 6.2.2) reconstructing decays of the type  $H_b \rightarrow h \ell \ell'$ , where  $H_b$  indicates a b-hadron, and  $\ell$  and  $\ell'$  indicate the two leptons.  $h$  represents different types of hadrons including kaons, pions and  $\Lambda^0$ s. Only the case where  $H_b$  is a  $\Lambda_b^0$  and  $h$  is a  $\Lambda^0$  is considered for selecting candidates in this analysis.

### 6.1.2 Simulation

Simulated samples are used to obtain the efficiencies, study the backgrounds and extract the invariant mass shapes as explained in the following sections. Table 6.2 lists the simulated samples used in this analysis, together with the number of events generated per each data-taking year. The  $B_d^0 \rightarrow K_S^0 \ell^+ \ell^-$  samples use the BTOSLLBALL [88] decay model. All other MC samples have been generated with a phase-space decay model.

Simulated samples for 2017 running conditions are not currently available in LHCb; therefore, 2016 samples are used in the study of 2017 data. The data-taking conditions being the same between the two years, this choice introduces no bias.

Table 6.2 – Simulated samples and their number of produced events, per year.

| Decay                                                                                | 2011 | 2012 | 2015 | 2016       |
|--------------------------------------------------------------------------------------|------|------|------|------------|
| $B_d^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K_S^0$                            | 2M   | 2M   | 10M  | 14M (S26)  |
| $B_d^0 \rightarrow J/\psi(\rightarrow e^+ e^-) K_S^0$                                |      | 500k | 2M   | 4M (S26)   |
| $B_d^0 \rightarrow K_S^0 \mu^+ \mu^-$                                                | 1M   |      |      |            |
| $B_d^0 \rightarrow K_S^0 e^+ e^-$                                                    | 2M   |      |      | 2M         |
| $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$              | 500k | 1M   | 1M   | 4M         |
| $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$                  | 50k  | 100k | 30k  | 150k(S28)  |
| $\Lambda_b^0 \rightarrow \Lambda^0 \mu^\pm \mu^\mp$                                  | 2M   |      |      |            |
| $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$                                    | 50k  | 100k | 30k  | 150k (S28) |
| $\Xi_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \Lambda^0$                      | 2M   | 2M   |      |            |
| $\Xi_b^{(0/+)} \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \Xi^{0/+}$                | 2M   | 2M   | 600k | 3M         |
| $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow \Lambda^0 \mu^+ \nu_\mu) e^- \nu_e$ |      | 500k |      | 500k       |
| $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow \Lambda^0 e^+ \nu_e) \mu^- \nu_\mu$ |      | 50k  |      | 50k        |
| $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow \Lambda^0 \pi^+) \mu^- \nu_\mu$     |      | 50k  |      | 50k        |
| $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow \Lambda^0 \pi^+) e^- \nu_e$         |      | 50k  |      | 50k        |

## 6.2 Selection

The trigger, stripping and cut-based selection criteria used on the signal and normalisation channels are presented in this section. After this first sample cleaning, a multivariate classifier is employed to further reduce the combinatorial background. Its details are discussed in Section 6.4.

### 6.2.1 Trigger

Signal candidates are selected at L0 level using the muon track, which provides a clean and efficient trigger response. The L0Muon requirement is then followed by the logic OR response from the HLT1 algorithms HLT1TrackMuon and HLT1TrackAllL0, applied again on the muon track. A description of these requirements is provided in Section 5.1.1.

Topological lines, based on the geometrical arrangement of tracks and vertices are applied at HLT2 level.

The full list of trigger requirements used to select  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  signal candidates is reported in Table 6.3. The same trigger requirements – applied on either muon for what concerns L0 – are used on  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  candidates to select the normalisation sample (see Section 6.7).

Table 6.3 – Trigger selection for  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ .

|      | <b>Run1</b>   | <b>Run2</b> |
|------|---------------|-------------|
| L0   | L0Muon        |             |
|      | TrackAllL0    | TrackMVA    |
| HLT1 | or            | or          |
|      | TrackMuon     | TrackMuon   |
|      | Topo2BodyBBDT | Topo2Body   |
| HLT2 | or            | or          |
|      | Topo3BodyBBDT | Topo3Body   |

### 6.2.2 Stripping

A dedicated stripping algorithm, called Bu2LLK\_meLine, is used to select the signal,  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ , while a different one, called Bu2LLK\_mmLine is used to select  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  candidates.

Table 6.4 reports the cuts applied in both algorithms.  $\Lambda^0$  candidates are reconstructed under two categories, depending on the  $p$  and  $\pi$  tracks reconstruction. If the  $\Lambda^0$  decay takes place inside the VELO (see Section 3.5.5), the two hadrons will be reconstructed as Downstream tracks, while in the opposite case, they will yield two Long tracks. The two cases are indicated respectively as DD and LL candidates. Most of the stripping cuts are the same for the two categories.

### 6.2.3 Pre-selection

A series of loose pre-selection cuts are applied on top of the stripping requirements, in order to further clean the data. The cuts applied on the signal sample  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  are reported in Table 6.5, while those used on the normalisation channel  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  are in Table 6.6. The efficiencies of these requirements are discussed in Section 6.8.

Fiducial cuts are meant to reject unphysical events, resulting from reconstruction errors. A lower cut on the  $\Lambda_b^0$  invariant mass allows to remove combinatorial background in a region which is not needed for its estimation, being too far from the signal peak. The  $J/\psi$  invariant mass is used to veto resonant backgrounds on signal, while, on the opposite, it allows to isolate the resonant  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  channel for the normalisation. Similarly, the  $\Lambda^0$  mass cut allows to better isolate decays with a resonant hadronic structure. HasDet requirements consist in the check that a specific track left hits in a specific sub-detector, which is necessary to compute some quantities such as the PID variables. For the nature of Down and Long tracks, these requirements differ between the DD and LL cases. Lastly, the HOP-FD cut exploits the correlation between the flight distance (FD) of the  $\Lambda_b^0$  and  $\alpha_{\text{HOP}}$  (see Appendix F) to isolate the

Table 6.4 – Requirements of the Bu2LLK\_meLine and Bu2LLK\_mmLine. When a cut differs between DD and LL candidates, the cut for DD is reported in parentheses.

| Particle      | Bu2LLK_meLine                                                    | Bu2LLK_mmLine         |
|---------------|------------------------------------------------------------------|-----------------------|
| $e$           | $p_T > 300 \text{ MeV}/c$                                        | -                     |
|               | $\chi_{IP}^2 > 9$                                                | -                     |
|               | $DLL_e > 0$                                                      | -                     |
| $\mu$         | $p_T > 300 \text{ MeV}/c$                                        |                       |
|               | $\chi_{IP}^2 > 9$                                                |                       |
|               | IsMuon                                                           |                       |
| $e\mu$        | $0.1 < m < 5.5 \text{ GeV}^2/c^4$                                | -                     |
|               | $\chi_{ORIVX}^2 < 9$                                             | -                     |
|               | $\chi_{FD}^2 > 16$                                               | -                     |
| $\mu\mu$      | -                                                                | $\chi_{DOCA}^2 < 30$  |
|               | -                                                                | $\chi_{ORIVX}^2 < 25$ |
| $\pi$         | $p > 2 \text{ GeV}/c$                                            |                       |
|               | $p_T > 300 \text{ MeV}/c$                                        |                       |
|               | $\chi_{IP}^2 > 9(4)$                                             |                       |
| $p$           | $p > 2(4) \text{ GeV}/c$                                         |                       |
|               | $p_T > 300 \text{ MeV}/c$                                        |                       |
|               | $\chi_{IP}^2 > 9$                                                |                       |
|               | $DLL_p > -5.0$                                                   |                       |
| $\Lambda^0$   | $\text{abs}(m - m_{\Lambda^0}) < 0.035(0.064) \text{ GeV}^2/c^4$ |                       |
|               | $\chi_{ORIVX}^2 < 30(25)$                                        |                       |
| $\Lambda_b^0$ | $\chi_{ORIVX}^2 < 9$                                             |                       |
|               | $\chi_{IP}^2 > 25$                                               |                       |
|               | $\text{DIRA} > 0.9995$                                           |                       |
|               | $\chi_{FD}^2 > 25$                                               |                       |

## Chapter 6. Search for the lepton-flavour violating decay $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$

Table 6.5 – Offline pre-selection cuts applied on  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  with LL and DD hadron tracks.

| Description      | Cut (LL)                                                           | Cut (DD)                     |
|------------------|--------------------------------------------------------------------|------------------------------|
| Fiducial         |                                                                    | $\chi_{FD}^2(\Lambda^0) > 0$ |
|                  |                                                                    | $Z_{DV}(\Lambda^0) > 0$      |
|                  |                                                                    | $Z_{DV}(\Lambda^0) < 2330$   |
|                  |                                                                    | $t(\Lambda^0) * 1000 > 0.5$  |
|                  |                                                                    | $t(\Lambda^0) * 1000 < 2000$ |
|                  |                                                                    | $\text{DIRA}(\Lambda^0) > 0$ |
| Mass             |                                                                    | $M(\Lambda_b^0) > 4300$      |
| $J/\psi$ veto    | $M(e^+ \mu^-) \notin [\sqrt{9.0} * 1000, \sqrt{10.1} * 1000]$      |                              |
| $\Lambda^0$ mass | $\text{abs}(M(p\pi) - 1115.68) < 15$                               |                              |
| HasDet           | $\mu_{\text{HasMuon}}$                                             | $\mu_{\text{HasMuon}}$       |
|                  | $e_{\text{HasCalo}}$                                               | $e_{\text{HasCalo}}$         |
|                  | $p_{\text{HasRich}}$                                               |                              |
| HOP-FD           | $M_{HOP}(\Lambda_b) > (2900 + 170 * \log(\chi_{FD}^2(\Lambda_b)))$ |                              |

Table 6.6 – Offline pre-selection cuts applied on  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  with LL and DD hadron tracks.

| Description      | Cut (LL)                                               | Cut (DD)                     |
|------------------|--------------------------------------------------------|------------------------------|
| Fiducial         |                                                        | $\chi_{FD}^2(\Lambda^0) > 0$ |
|                  |                                                        | $Z_{DV}(\Lambda^0) > 0$      |
|                  |                                                        | $Z_{DV}(\Lambda^0) < 2330$   |
|                  |                                                        | $t(\Lambda^0) * 1000 > 0.5$  |
|                  |                                                        | $t(\Lambda^0) * 1000 < 2000$ |
|                  |                                                        | $\text{DIRA}(\Lambda^0) > 0$ |
| Mass             |                                                        | $M(\Lambda_b^0) > 4800$      |
| HasDet           | $\mu_{\text{HasMuon}}$                                 | $\mu_{\text{HasMuon}}$       |
|                  | $e_{\text{HasMuon}}$                                   | $e_{\text{HasCalo}}$         |
|                  | $p_{\text{HasRich}}$                                   |                              |
| $\Lambda^0$ mass | $\text{abs}(M(p\pi) - 1115.68) < 15$                   |                              |
| $J/\psi$ mass    | $\text{abs}(M(\Lambda_b)^{PV, \Lambda^0} - 3096) < 50$ |                              |

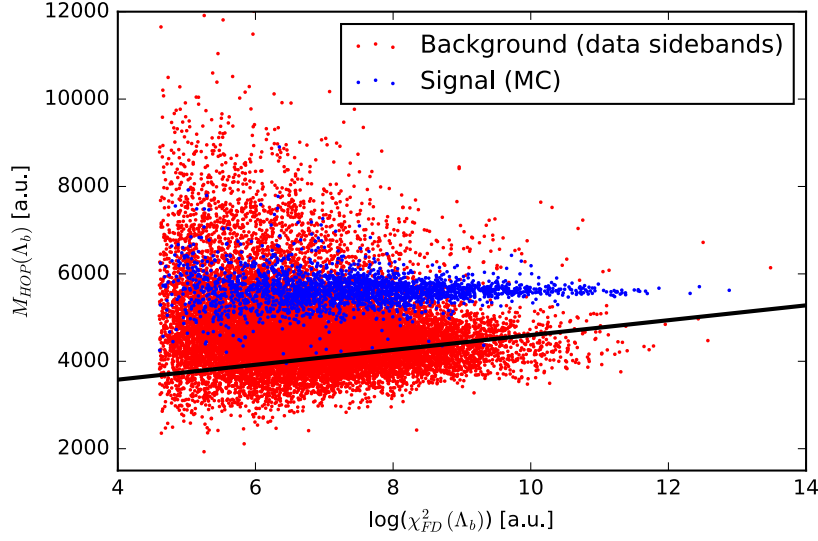


Figure 6.1 – Scatter plot of the  $\Lambda_b^0$  flight distance versus  $\alpha_{HOP}$  for signal (blue) and combinatorial background (red). The black line represents the cut applied on signal.

signal very efficiently from the combinatorial background. Figure 6.1 shows the scatter plot of these two variables on simulated signal and data combinatorial background, along with the linear cut, chosen with the aim of keeping  $\sim 100\%$  signal efficiency while removing as much background as possible.

### 6.3 Re-weighting of the signal simulated sample

A re-weighting procedure is applied on the simulated samples in order to correct them for kinematic mis-modelling.

Weights are extracted by comparing  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  simulation against s-weighted [122] data. On both samples, the same pre-selection and trigger requirements as for the normalisation fit, are applied. The weights are extracted by comparing the distributions of  $p(\Lambda_b^0)$ ,  $p_T(\Lambda_b^0)$  and the origin vertex  $\chi^2$  of the  $\Lambda^0$ ,  $\chi^2_{ORIVX}(\Lambda^0)$ . An additional weight is included to re-model the  $\Lambda_b^0$  lifetime, using the updated value of  $1.470 ps$ , from the PDG [2], compared to the outdated  $1.451 ps$  used in the generation of the MC sample. This weight is obtained independently from the others, and calculated analitically as:

$$w(t) = \exp \left[ t \cdot \frac{\tau_{wa} - \tau_{gen}}{\tau_{wa} \tau_{gen}} \right], \quad (6.1)$$

where  $\tau_{wa}$  and  $\tau_{gen}$  are respectively the world average mean lifetime and the generated one.

Figure 6.2 shows the distributions of  $p(\Lambda_b^0)$ ,  $p_T(\Lambda_b^0)$  and  $\chi^2_{ORIVX}(\Lambda^0)$  on s-weighted data and on simulation before and after the application of the weights. The imperfect agreement after

the re-weighting is due to differences between the binnings used in the plots and the ones used for the re-weighting. These latter are chosen with the aim of maximising the granularity while avoiding to have bins which are too little populated for one of the samples, which would lead to a large uncertainty in the weights.

Figure 6.3 shows the distribution of the decay time of the  $\Lambda_b^0$ ,  $t(\Lambda_b^0)$ , on s-weighted data and on simulation before and after the application of the kinematic weights.

The effect of the re-weighting has been checked on the other variables relevant for this analysis such as those used in the multivariate classifier: an overall improvement in the agreement with s-weighted data has been observed. The remaining differences are accounted for by comparing the response of the classifier between data and MC as described in Section 6.4.4.

## 6.4 Multivariate Classifier

### 6.4.1 Definition

A multivariate classifier is used in this analysis to suppress the combinatorial background. The classifier makes use of eleven topological features that describe the kinematics of the signal:

- the distance of closest approach of the two leptons,  $\text{DOCA}(\text{leptons})$ ;
- the direction angle between the flight direction of the  $\Lambda^0$  and the direction of its momentum,  $\text{DIRA}(\Lambda^0)$ ;
- the  $\chi^2$  of the origin vertex of the  $\Lambda^0$ ,  $\chi_{ORIVX}^2(\Lambda^0)$ ;
- the impact parameter  $\chi^2$  of the  $\Lambda_b^0$ ,  $\chi_{IP}^2(\Lambda_b^0)$ ;
- the sum of the impact parameter  $\chi^2$  of the two leptons,  $\chi_{IP}^2(\text{leptons})$ ;
- the sum of the transverse momentum asymmetries of the four final-state tracks,  $\text{sum}(p_T \text{ asymmetry})$ ;
- the sum of the cone isolations<sup>1</sup> of the four final-state tracks,  $\text{sum}(\text{isolation})$ ;
- the HOP factor,  $\alpha_{\text{HOP}}$ ;
- the transverse momentum of the  $\Lambda_b^0$  factor,  $p_T(\Lambda_b^0)$ ;
- the difference of pseudo-rapidity between the two hadron tracks,  $\Delta\eta(\text{hadrons})$ ;
- the  $\chi^2$  of the distance between the decay vertex of the  $\Lambda_b^0$  and the production vertex of the two leptons,  $\text{leptons\_vertex\_detachment}$ .

---

<sup>1</sup>The isolation variable is defined as in the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis, see Section 5.1.5



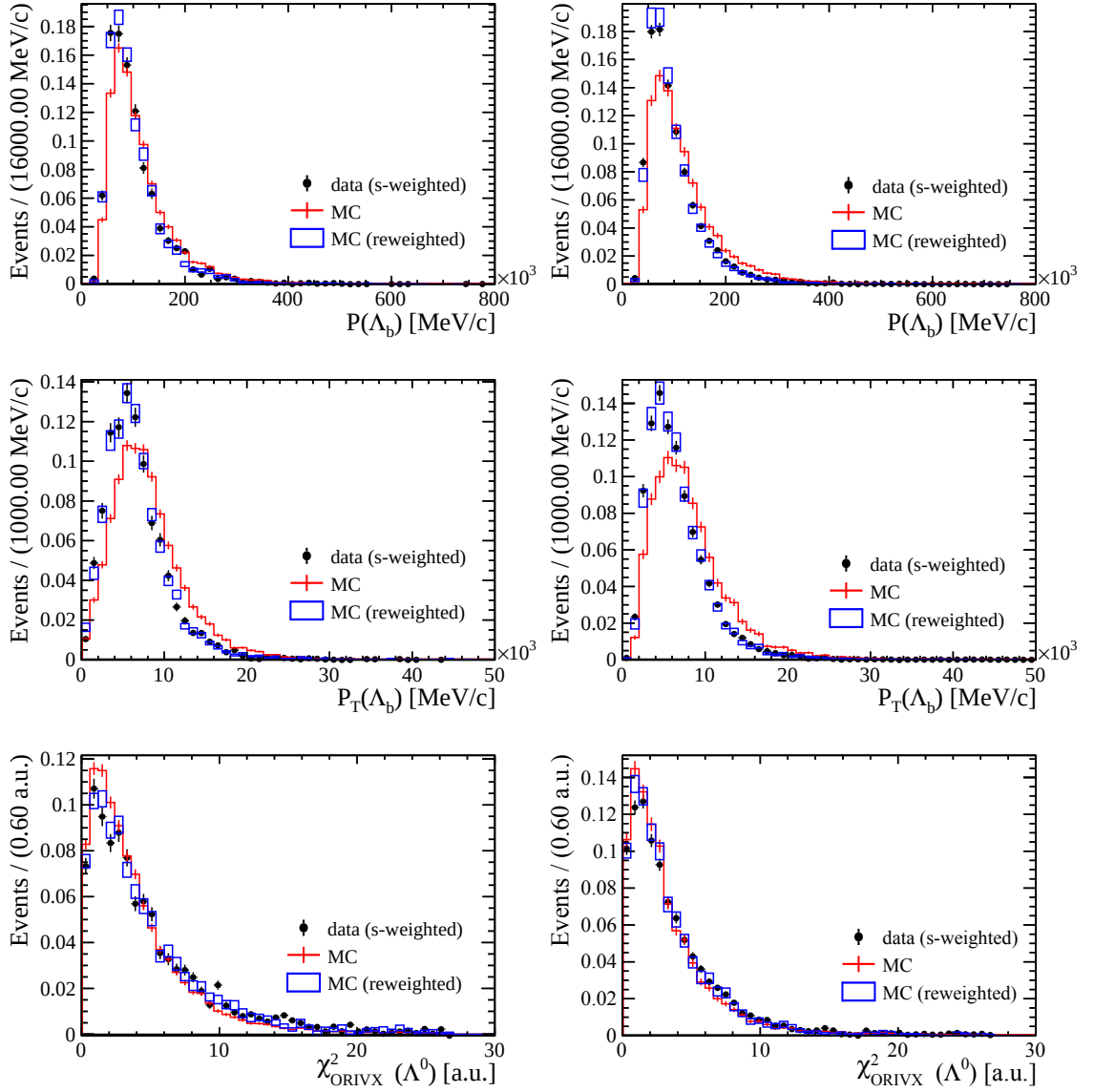


Figure 6.2 – Distribution of  $p(\Lambda_b^0)$  (top),  $p_T(\Lambda_b^0)$  (middle) and  $\chi^2_{ORIVX}(\Lambda^0)$  (bottom) on s-weighted data (black points) and on simulation before (red lines) and after (blue boxes) the re-weighting. The left plot describes Run I samples, while the right one describes Run II.

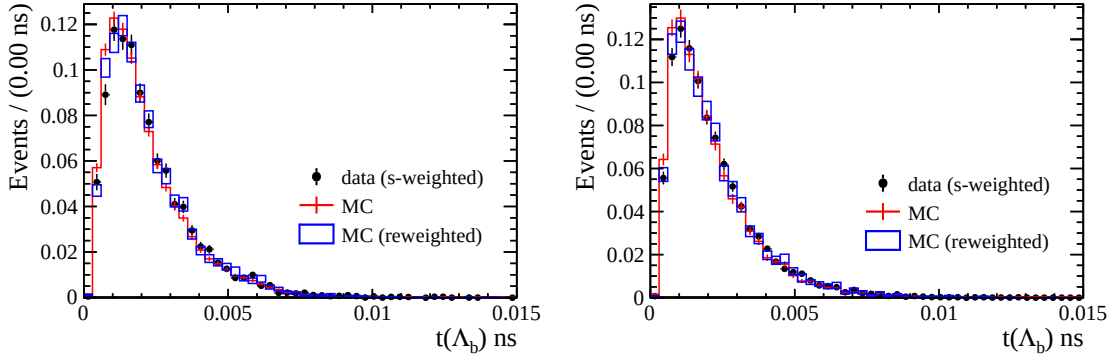


Figure 6.3 – Distribution of  $t(\Lambda_b^0)$  on s-weighted data (black points) and on simulation before (red lines) and after (blue boxes) the re-weighting. The left plot describes Run 1 samples, while the right one describes Run 2.

These variables are chosen according to their discriminating power, the accuracy of their simulation (evaluated comparing  $\Lambda_b^0 \rightarrow \Lambda^0 J/\psi$  s-weighted data against simulated samples), the low correlation with the invariant mass of the background and the moderate differences between their PDFs in the 8 signal categories. The latter requirement assures a homogeneous response of the classifier across the categories.

Different types and implementations of classifiers are compared in order to find the best-performing algorithm. The integral under the ROC curve (AUC) is used as a figure of merit to compare the following four classifiers:

1. a Gradient Boosting Decision Tree (BDT) from XGboost [123];
2. a Gradient Boosting Decision Tree from Scikit-Learn [124];
3. a Gradient Boosting Decision Tree from TMVA [107];
4. an Artificial Neural Network (Multi-layer Perceptron) from Scikit-Learn.

A randomised scan of the most important training parameters is performed for each of these classifiers in order to find the configuration giving the highest performance. The four classifiers, trained using their optimal parameter configurations, are then compared. Further details on the training can be found in Section 6.4.2.

The four algorithms obtain similar results, and their ROC curves can be seen in Figure 6.4. Nevertheless, the time needed to execute the training differs largely, with number 1 and 2 being significantly (roughly 10 times) faster than the other two.

Classifier 1, hereinafter referred to simply as BDT, is chosen for this analysis. It is configured with 900 trees having a maximum depth of 3, and the learning rate parameter is set to 0.02.

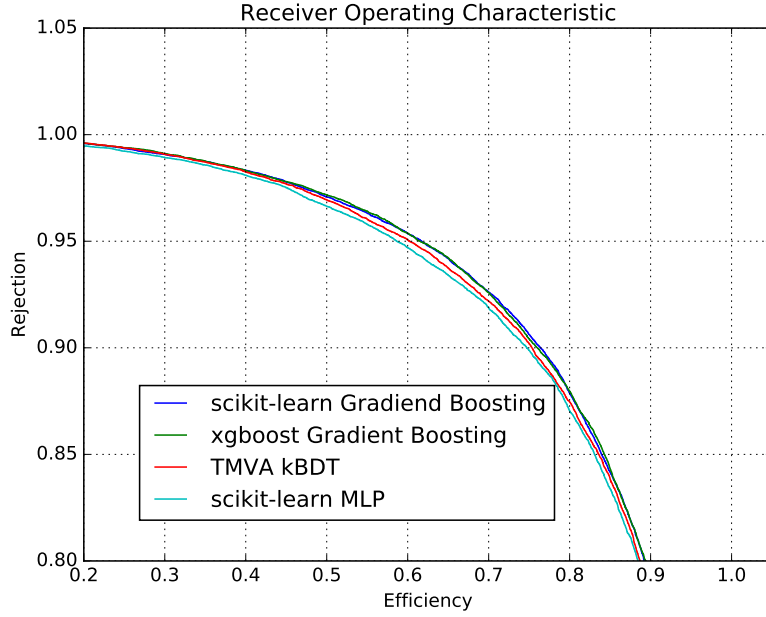


Figure 6.4 – Receiver Operating Characteristics curves of the four multivariate classifiers considered for the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis.

This configuration is determined by maximising the integral of the ROC curve on an unbiased sample.

### 6.4.2 Training

The BDT is trained on a signal simulated sample against combinatorial background data taken from the mass sidebands around the expected signal in the range

$$M(\Lambda e \mu) < 5000 \text{ MeV} \parallel M(\Lambda e \mu) > 5800 \text{ MeV}.$$

The sample consist of about 60000 candidates in total, half of signal and half of background. This size is dictated by the available MC statistics: the data candidates, being more abundant, are drawn randomly from the full samples to match the size of the MC. The ratio of the number of Run I over Run II events is scaled in MC to match the one observed in the data sample, in order to avoid training bias. The simulated samples are weighted as described in Section 6.3. The distributions of the training variables for signal and background are shown in Appendix E, along with their mutual correlations. In order to exploit all of the available simulated sample and avoid performance estimation bias, the *k-folding* cross-validation technique [125] is used, with  $k = 5$ .

The absence of over-training is assessed by comparing the BDT response (hereinafter referred

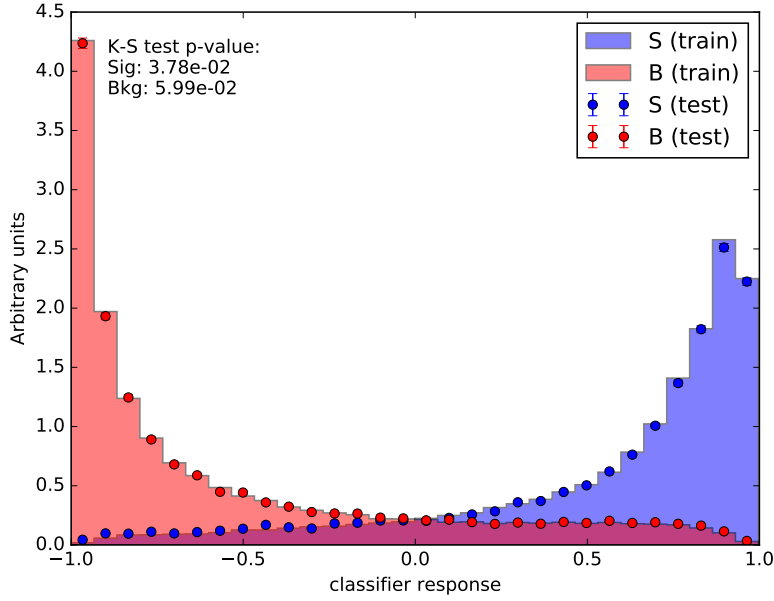


Figure 6.5 – Response of the BDT used in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis for the signal (blue) and background (red) samples used for training (points) and testing (filled histograms). The Kolmogorov-Smirnov test’s p-value is overlaid.

to as BDT) distributions on the testing and training samples, summing on the folds. The two distributions, shown in Figure 6.5 along with the p-values of the Kolmogorov-Smirnov test, are found to be compatible, as expected in case of no over-training.

### 6.4.3 Performances

Figure 6.6 shows the ROC curves for the five training folds, each evaluated on the respective testing fold. The five classifiers show compatible performances, indicating the homogeneity between the training results across the folds. The Area Under the Curve (AUC) is also indicated, as a figure of merit for quantifying the performance.

The full testing sample, consisting of the union of the five test folds, each with its corresponding BDT response, is then split once for HasBremAdded categories, once for LHC run and once for  $\Lambda^0$  track type in order to compare the distribution of the response along these categories. The results are shown in Figure 6.7.

Finally, the ROC curves corresponding to the eight categories resulting from the full split of the sample are compared in Figure 6.8. The curves show similar but not identical performances. This is fully expected given the different size of the samples in the training dataset and, most importantly, because of the small differences in the distribution of the training variables. The bremsstrahlung radiation, for example, smears the distribution of momenta, while the two

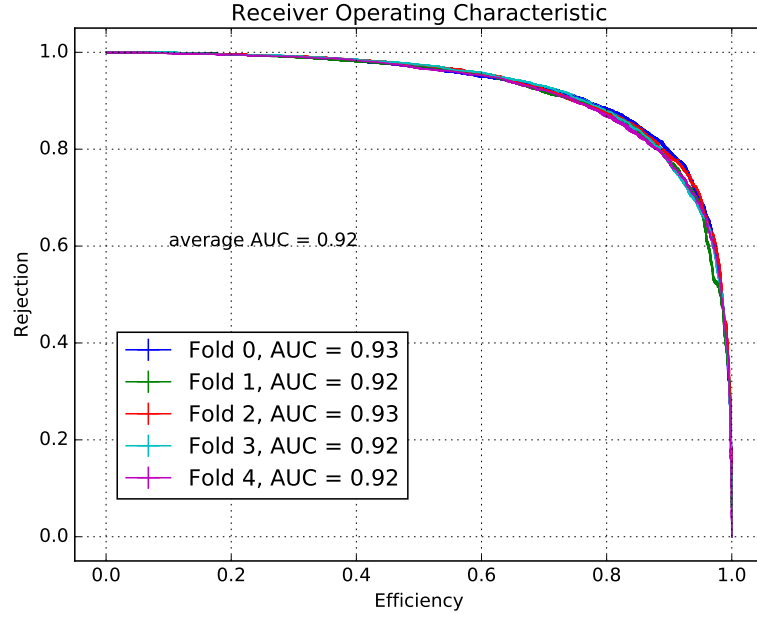


Figure 6.6 – Receiver Operating Characteristics curves for the five cross-validation folds of the BDT classifier used in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis. The thickness of the lines represents the measurement uncertainty. The Area Under the Curve (AUC) is also indicated.

$\Lambda^0$  track types have different reconstruction accuracies because of the different amount of information available on the hadron tracks. The small differences in performance between the Run I and Run II samples are also due to differences in the reconstruction and selection of the candidates.

The eight categories are therefore treated separately for what concerns the determination of the BDT efficiencies, described in the next sections.

#### 6.4.4 Response on data

In order to avoid biases in the estimation of the BDT performance due to inaccuracies in the simulation, the response of the classifier must be cross-checked on data. Nevertheless, given the unique signature of the searched signal, no SM process can be used as its proxy, to obtain a reliable estimate of the BDT response. As a viable alternative, the latter is compared between data and simulation on a same known SM process. The  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  decay is used for this purpose: an s-weighted [122] data sample is compared to signal MC. This channel has been chosen for its clean and abundant yield.

The  $\alpha_{\text{HOP}}$  variable, present as one of the features on the BDT, cannot be reproduced on this sample, as it is not defined in absence of electrons. For this reason,  $\alpha_{\text{HOP}}$  is assigned the default value of 1 (corresponding to the ideal signal case) for all the candidates, at the moment of computing the BDT response on  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ , and it is then checked separately on

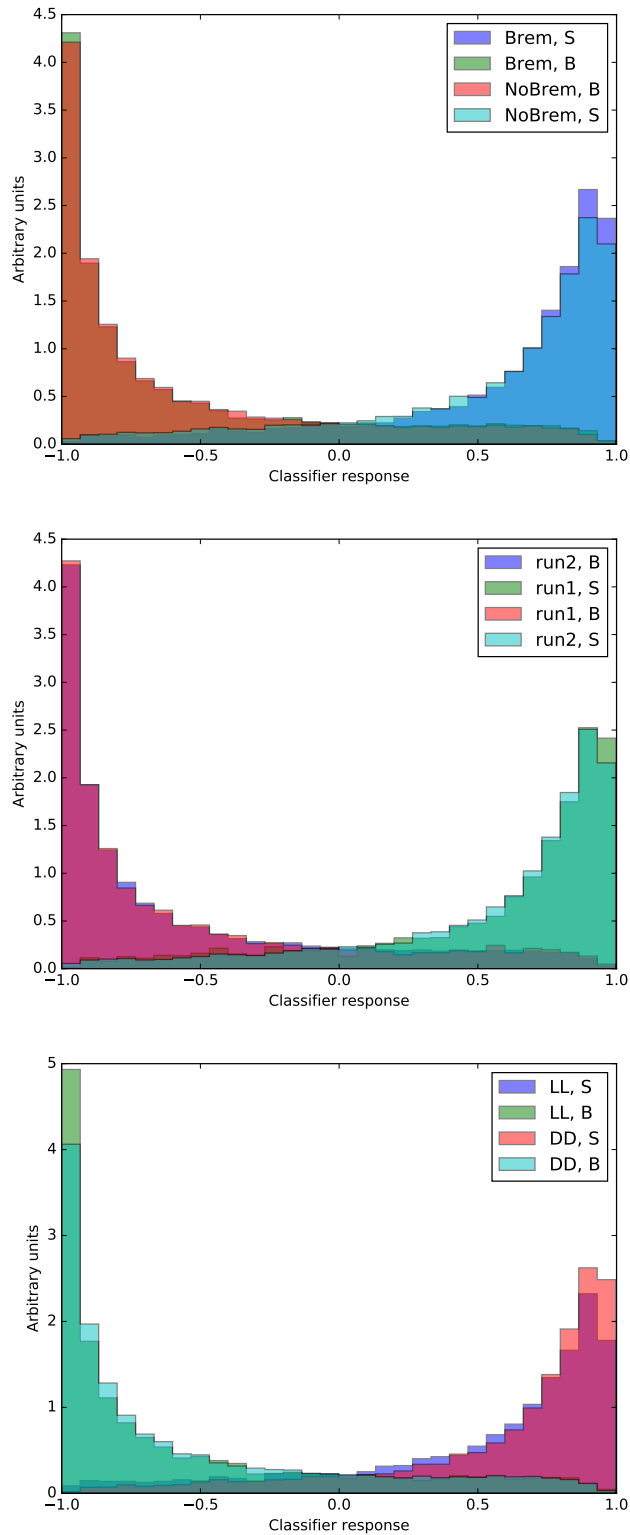


Figure 6.7 – Response of the BDT classifier used in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis, compared for HasBremAdded categories (top), LHC run (middle) and  $\Lambda^0$  track type (bottom). In each plot, the sample is integrated over the other categories.

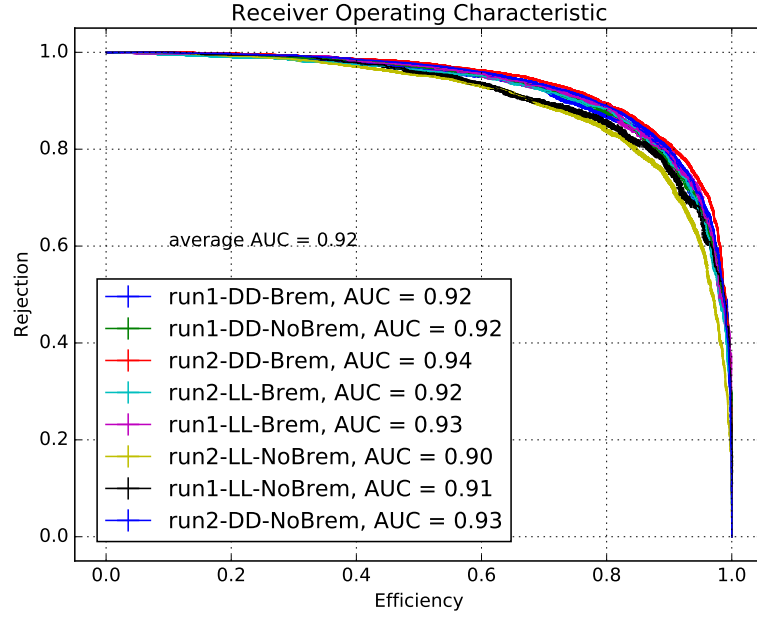


Figure 6.8 – Receiver Operating Characteristics curves of the BDT classifier used in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis, for the eight categories in which the analysis is split. Each point in the curves is shown with its binomial errors on both axes.

samples of  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$  data and MC. The BDT response with  $\alpha_{\text{HOP}}$  fixed to 1 is going to be referred to as  $\text{BDT}^{\text{fix-hop}}$ , hereinafter.

Figure 6.9 shows the efficiency of a running cut on the  $\text{BDT}^{\text{fix-hop}}$  distribution on  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  in Run I and Run II, for LL and DD  $\Lambda^0$  tracks. The response of the BDT appears well reproduced, especially in Run II events. The relative difference between the response on data and simulation in each category is assigned as a systematic uncertainty on the nominal response from signal MC, and propagated to the final result. This uncertainty is always below 1.5% of the respective value.

Figure 6.10 shows the comparison of the  $\alpha_{\text{HOP}}$  variable's distribution between data and simulation, using Run I + Run II s-weighted data and the corresponding simulated sample, in order to maximise the available statistics from data. No split is performed in the hadron track types, since  $\alpha_{\text{HOP}}$  mainly depends on the lepton tracks. The agreement is satisfactory, and no additional systematic error is assigned.

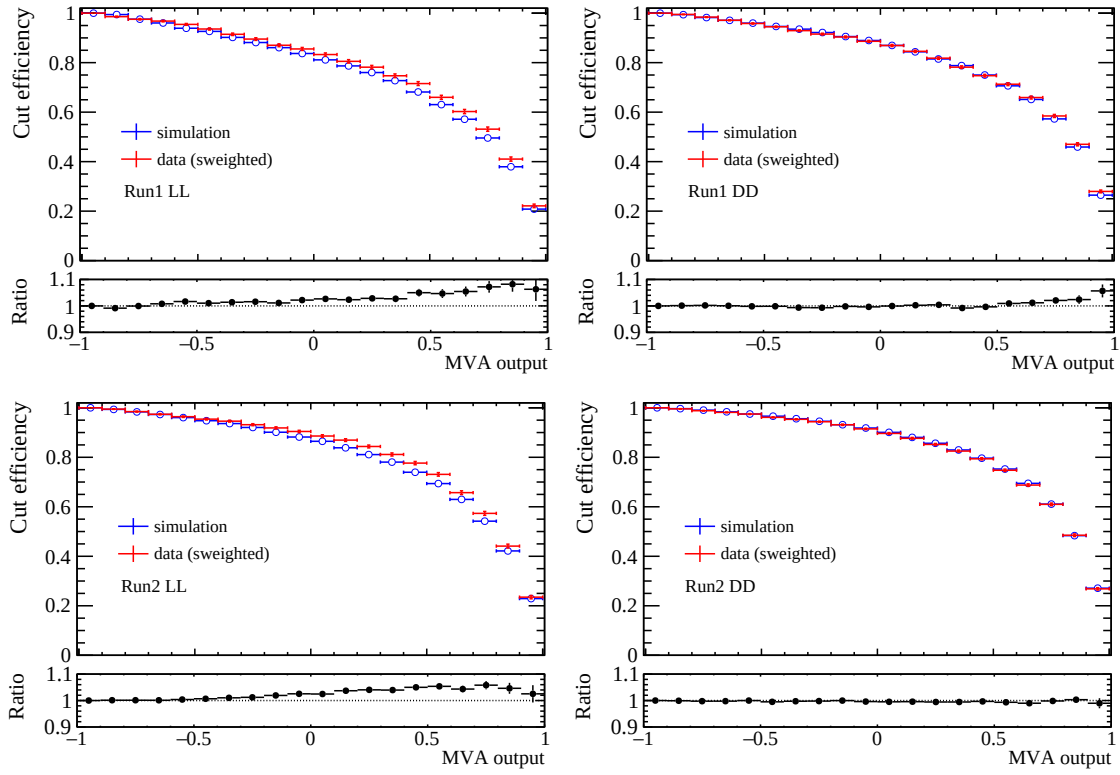


Figure 6.9 – Running cut efficiency on  $\text{BDT}^{fix-hop}$  evaluated on  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  data and simulated samples. The left plots correspond to LL  $\Lambda^0$  tracks and the right plots to DD  $\Lambda^0$  tracks. The top row corresponds to Run I events and the bottom row corresponds to Run II.

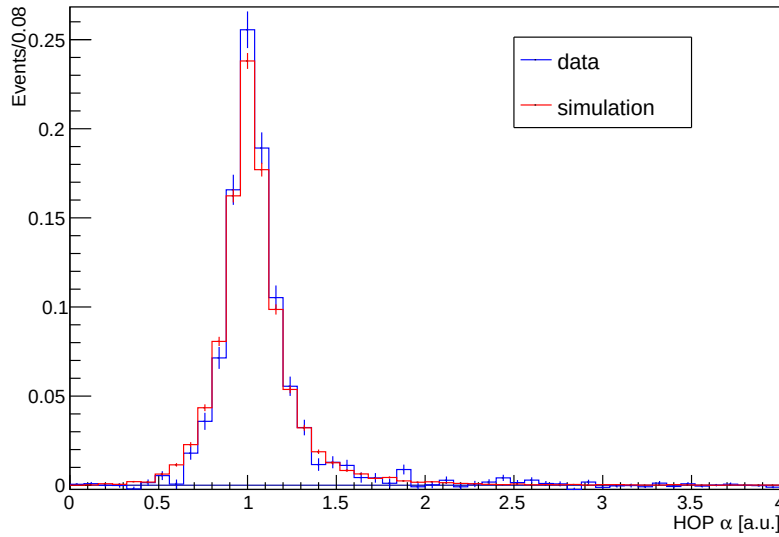


Figure 6.10 –  $\alpha_{\text{HOP}}$  distribution on  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$  data and MC for Run I + Run II candidates, in both track categories.



Table 6.7 – Double-semileptonic background to  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  and their branching fractions.

| Decay                                                                              | Branching fraction   |
|------------------------------------------------------------------------------------|----------------------|
| $\Lambda_b^0 \rightarrow \Lambda_c^{0+} (\rightarrow \Lambda^0 \ell \nu) \ell \nu$ | $2.3 \times 10^{-3}$ |
| $\Lambda_b^0 \rightarrow \Lambda_c^{0+} (\rightarrow \Lambda^0 \pi) \ell \nu$      | $8.0 \times 10^{-4}$ |
| $\Lambda_b^0 \rightarrow \Lambda_c^{0+} (\rightarrow \Lambda^0 \ell \nu) \pi$      | $1.8 \times 10^{-4}$ |
| $\Lambda_b^0 \rightarrow \Lambda_c^{0+} (\rightarrow \Lambda^0 \pi) \pi$           | $6.3 \times 10^{-5}$ |

#### 6.4.5 Cut optimisation

The optimal cut on the BDT response is optimised independently in each of the 8 analysis categories, by maximising the *Punzi figure of merit* [126], defined as

$$P = \frac{\varepsilon(t)}{a/2 + \sqrt{B(t)}}, \quad (6.2)$$

where  $t$  is the cut value,  $\varepsilon$  is the signal efficiency,  $B$  is the expected number of background events and  $a$  is the number of standard deviations of significance at which the analysis aims to observe the signal. Specifically,  $a$  is fixed to 3.  $B$  is obtained by fitting the signal sidebands with an exponential function and extrapolating in the signal region, while  $\varepsilon$  is computed from signal simulated events, weighted as described in Section 6.3.

The cut  $t$  is scanned in the domain of definition of the BDT response. The resulting  $\varepsilon$  and  $P$  plots are shown in Figures 6.11 and 6.12. The best cut, corresponding to the maximum  $P$ , is shown by the red dashed line. In order to reduce the dependency on statistical fluctuations, the  $\varepsilon$  and  $P$  curves are smoothened by taking their moving average over 5 points. The best cut obtained this way deviates only slightly from the one obtained from the raw data.

### 6.5 Backgrounds

The absence of SM decays producing a peaking structure with an electron, a muon and a  $\Lambda^0$  baryon makes the present search rather clean in terms of backgrounds, besides the combinatorial contribution. Furthermore, the veto on the dilepton invariant mass around the  $J/\psi$  resonance, described in Section 6.2.3 allows to suppress otherwise abundant resonant misidentified backgrounds such as  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  and  $B^0 \rightarrow (K_s^0 \rightarrow \pi^+ \pi^-) (J/\psi \rightarrow \mu^+ \mu^-)$ .

Although not peaking, double-semileptonic decays like those reported in Table 6.7 could however constitute a considerable source of background. The two modes with a pion in the final state have a very low branching fraction. Furthermore, the misidentification probability of a pion to a muon or an electron is negligible ( $\sim 1\%$  and  $\sim 3\%$ , respectively). For this reason,

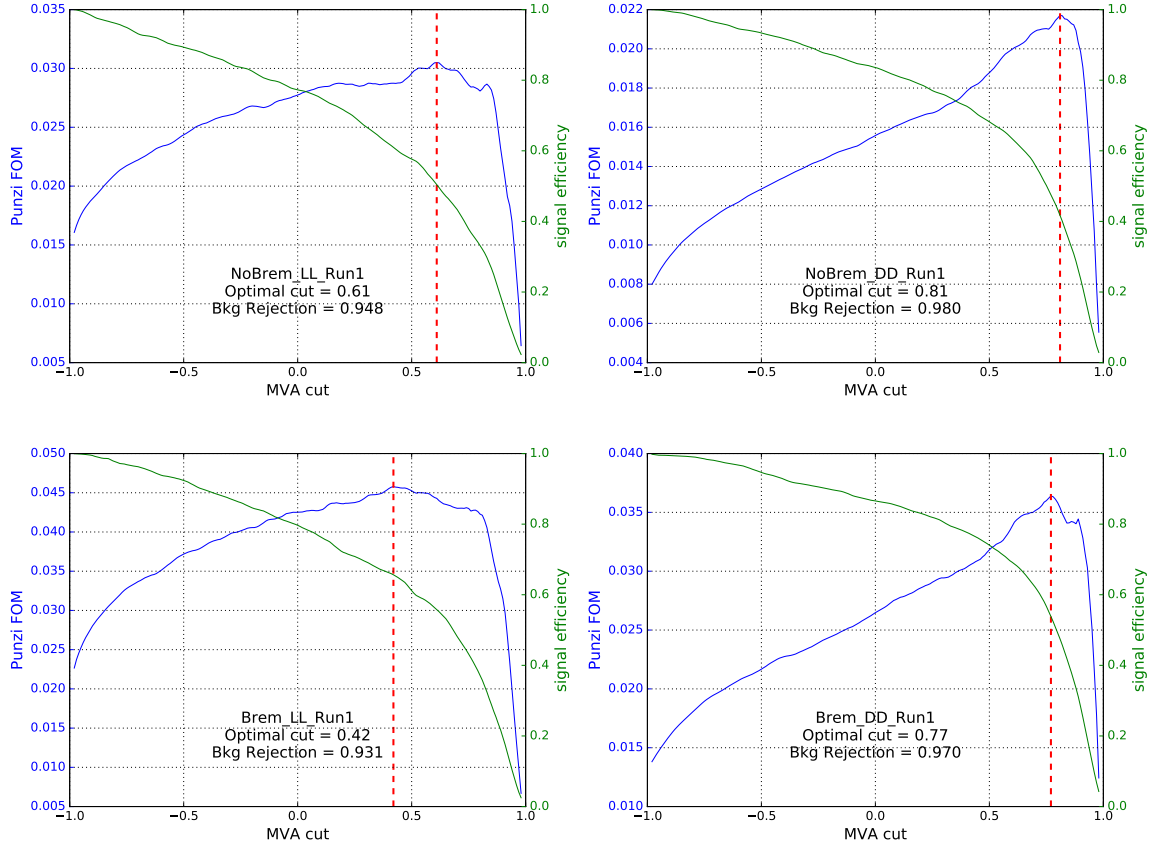


Figure 6.11 – Punzi figure of merit and signal selection efficiencies as a function of the cut on BDT for Run I events. The categories without (with) bremsstrahlung photons added are shown in the top (bottom) row, and those with LL (DD) tracks are shown in the left (right) column. The red dashed lines indicate the optimal cuts.

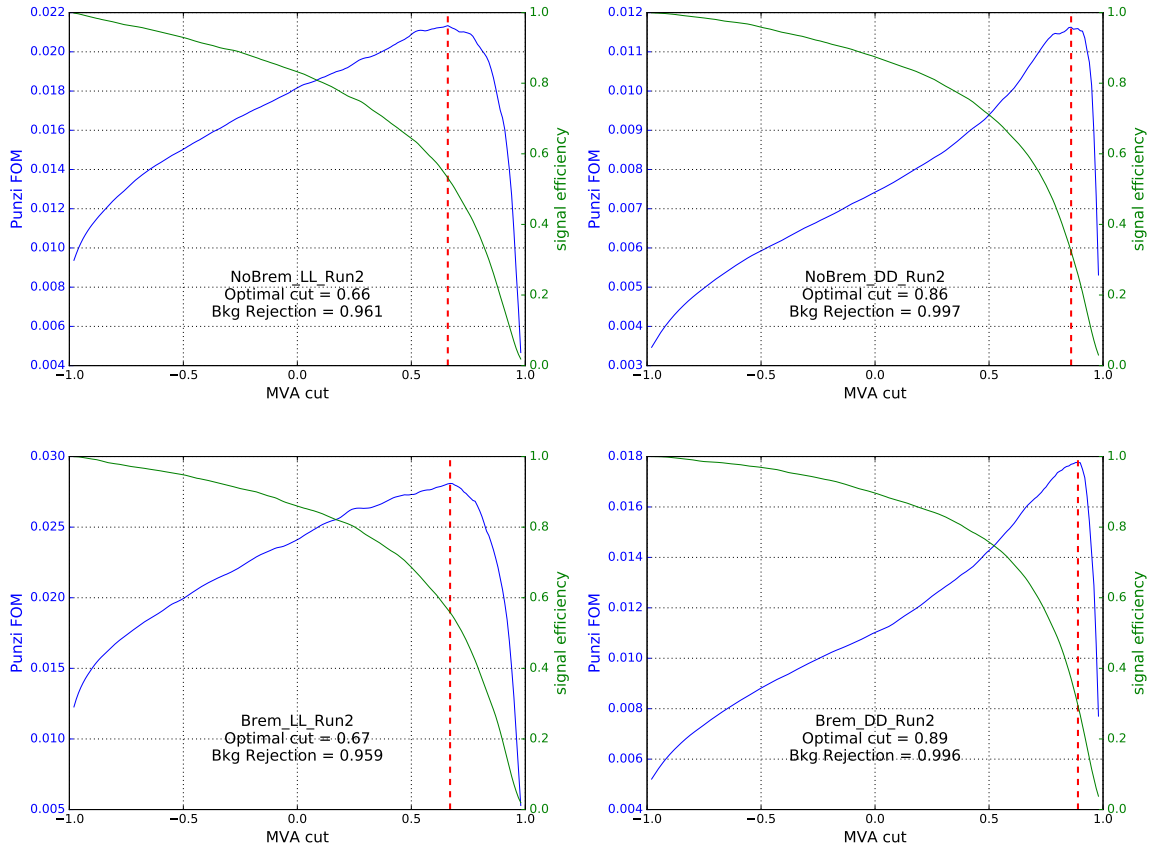


Figure 6.12 – Punzi figure of merit and signal selection efficiencies as a function of the cut on BDT for Run I events. The categories without (with) bremsstrahlung photons added are shown in the top (bottom) row, and those with LL (DD) tracks are shown in the left (right) column. The red dashed lines indicate the optimal cuts.

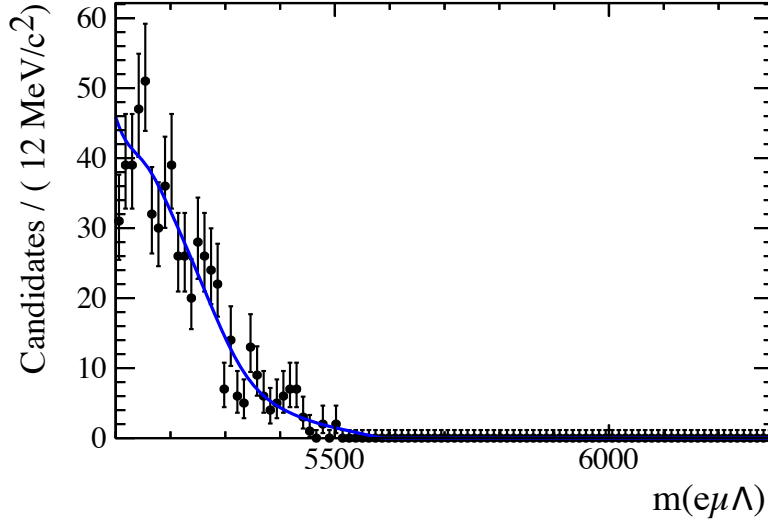


Figure 6.13 – Invariant mass of simulated  $\Lambda_b^0 \rightarrow \Lambda^{0+}(\rightarrow \Lambda^0 \mu \nu) e \nu$  candidates reconstructed as  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ .

only the first decay in the table is analysed.

As can be observed in Figure 6.13, the invariant mass shape of the  $\Lambda_b^0 \rightarrow \Lambda^{0+}(\rightarrow \Lambda^0 \mu \nu) e \nu$  reconstructed as  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  has a decaying shape in the analysis mass region considered for the analysis. Depending on its normalisation, this shape could strongly contribute to the the observed data.

The normalisation of this background is under study. Depending on the result, a non-parametric shape describing its contribution will be added to the data mass fit described in Section 6.9.1.

## 6.6 Calibration of signal invariant mass

### 6.6.1 Signal model

The invariant-mass resolution of the signal can be wrongly estimated in simulation. For this reason, the width of the signal mass shape obtained from MC must be corrected by a factor,  $\mathcal{C}$ , that takes into account the differences between data and simulation.

As in the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis, because of the absence of a calibration channel with an electron and a muon in the final state, the information on the resolution must be extracted from the combination of a muonic and an electronic decay channel.  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$  and  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  are used in this analysis, for their evident similarity with the signal.

While on two-body decays the  $\mathcal{C}$  factor can be obtained from simple kinematic considerations

as explained in Appendix A, the four-body structure of  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  requires a different, empirical treatment. The chosen approach consists in finding the value for  $\sigma_{e\mu}^{\text{data}}$  by performing an expansion around  $\sigma_{\mu\mu}^{\text{data}}$  (or equivalently around  $\sigma_{ee}^{\text{data}}$ ):

$$\sigma_{e\mu}^{\text{data}} = \sigma_{\mu\mu}^{\text{data}} + (\sigma_{e\mu}^{\text{MC}} - \sigma_{\mu\mu}^{\text{MC}}) \cdot \frac{\sigma_{ee}^{\text{data}} - \sigma_{\mu\mu}^{\text{data}}}{\sigma_{ee}^{\text{MC}} - \sigma_{\mu\mu}^{\text{MC}}}. \quad (6.3)$$

The  $\mathcal{C}$  factor can be then calculated as

$$\mathcal{C} = \frac{\sigma_{e\mu}^{\text{data}}}{\sigma_{e\mu}^{\text{MC}}}. \quad (6.4)$$

Note that in this case the corrected parameters can be calculated directly and  $\mathcal{C}$  is only reported to express the relative difference between data and simulation.

The quantities in Formula 6.3 are obtained via invariant-mass fits to data and simulation, described in the next sections.

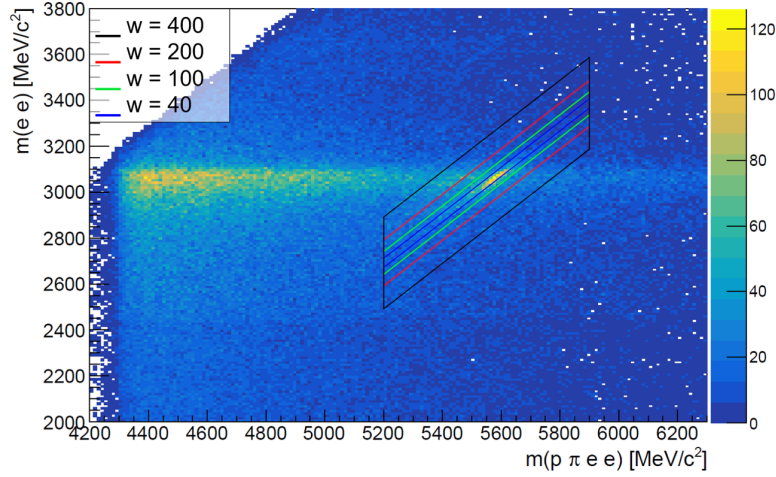
#### Fit to $\Lambda_b^0 \rightarrow \Lambda^0 J/\psi (\rightarrow ee)$

The fit to  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow e^\pm e^\mp)$  requires some prior treatment of the data. The following initial considerations are made:

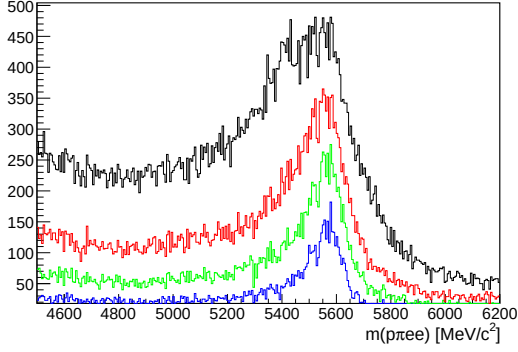
1. the decay cannot be cleanly fit without any further selection after the stripping, specific cuts are thus applied;
2. bremsstrahlung emission has a sizeable impact on the shape of the mass spectrum;
3.  $B^0 \rightarrow (J/\psi \rightarrow ee)(K_s^0 \rightarrow \pi\pi)$  decays where a pion is mis-identified as a proton cannot be cleanly fit as they have a broad shape under the signal peak, thus they have to be suppressed using PID information.

**Mass preselection** Figure 6.14a shows a two-dimensional histogram of the invariant masses  $m(p\pi ee)$  vs  $m(ee)$ . The invariant masses are recalculated constraining  $m(p\pi)$  to the known  $\Lambda^0$  mass value [2] and the  $\Lambda_b^0$  to originate from primary vertex. Note that in this case a  $J/\psi$  mass constraint cannot be applied, as we are interested in transferring the resolution from the calibration to the signal channel, which is non-resonant.

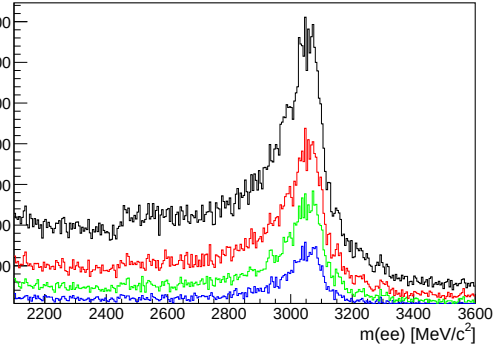
The oval peak corresponds to  $\Lambda_b^0$  decaying into  $J/\psi$ . The diagonal bounds on raw data come from kinematic limits. The horizontal yellow band is populated by partially-reconstructed background candidates, where the approximately constant  $m(ee)$  indicates that a true  $J/\psi$  is combined with other particles.



(a) Histogram of  $m(\Lambda_b)$  vs  $m(J/\Psi)$ . At low values of the  $x$ -axis, a partially reconstructed component is present and must be cut away.



(b)  $\Lambda_b^0$  invariant-mass distribution.



(c)  $J/\psi$  invariant-mass distribution.

Figure 6.14 – Mass preselection for  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$  channel displayed on 2016 data. The different colours correspond to the different cuts shown in Figure a.

$\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$  candidates are isolated by applying a geometrical cut in the  $(m(p\pi ee), m(ee))$  plane: a band with a constant vertical width  $w$  for any  $x = \text{const}$  slice, as shown in Figure 6.14a. This constant width allows to avoid edge effects on  $m(ee)$  when projected on  $m(p\pi ee)$ . The width has been set to  $w = 200 \text{ MeV}/c^2$  (red) which is a good compromise between high efficiency on  $J/\psi$  candidates and the exclusion of background.

Figures 6.14b and 6.14c show the projections of the cut events on  $x$ -axis and  $y$ -axis respectively, namely  $m(p\pi ee)$  and  $m(ee)$ .

**Proton identification** The decay  $B^0 \rightarrow (K_s^0 \rightarrow \pi^+ \pi^-)(J/\psi \rightarrow e^+ e^-)$  is an important source of background due to mis-identifications of pions as protons. A cut on the proton identification probability (ProbNN) is sufficient to suppress it. In Figure 6.15, the blue histogram shows the invariant mass difference  $\Delta m = m(p\pi ee) - m(ee)$ . A clear peak appears at  $\sim 2500 \text{ MeV}/c^2$

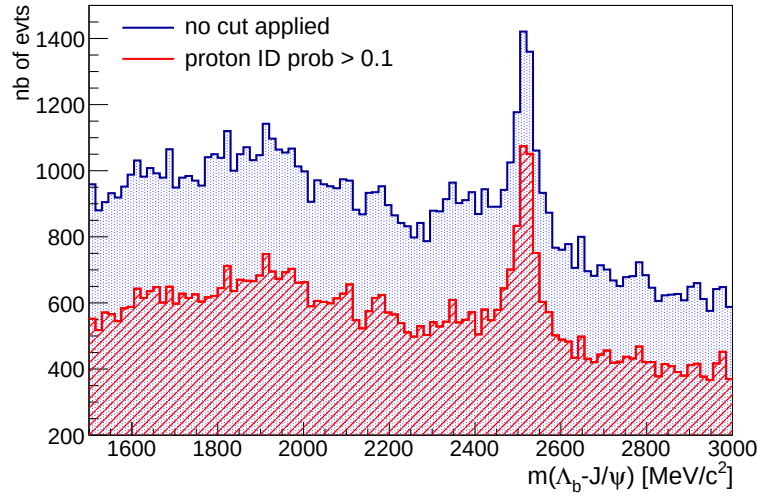


Figure 6.15 – Histogram of  $m(p\pi ee)-m(ee)$  for 2016 data candidates.

corresponding to  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^+e^-)$  decays and a smaller structure on the left is due to  $K_S^0$  decays. After applying a proton identification cut,  $\text{ProbNN} > 0.1$ , this background is reduced to a negligible level, as can be seen in the red histogram.

**Invariant-mass shapes** Bremsstrahlung emission affects the invariant-mass shape of the candidates. In the category with no bremsstrahlung emission, the energy distribution of electrons is given by a Crystal Ball [111] (CB) shape.

On the other hand, in the category with bremsstrahlung, the right tail of the distribution is populated by events for which the correction is over-estimated, as explained in the previous chapter for the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis. The invariant mass distribution is then described by a Double-sided Crystal Ball (DCB) shape, namely a Gaussian core with power tails on both sides.

In both cases, the core width,  $\sigma$ , can be used as a measurement of the invariant-mass resolution.

In  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ , bremsstrahlung emission happens from a single lepton, unlike in this calibration channel, where two electrons contribute to the radiation. Thus, for the purpose of correcting the mass resolution events with bremsstrahlung photons recovered for both electrons are removed, and, to avoid confusion, the candidates with bremsstrahlung not emitted by any of the electrons are indicated as  $0\gamma$ , and those with one radiating electron are indicated as  $1\gamma$ .

**Fits to data** As the analysis is performed in categories of track type and bremsstrahlung, the calibration sample is fit in the same categories. It was instead decided not to split the sample for Run I and Run II and perform a unique calibration for both runs. The resolutions indeed are not expected to strongly depend on the run, and keeping them together reduced the statistical uncertainty of the calibration.

The signal PDF used to fit data is obtained from the MC beforehand. The categories are fit independently and all parameters are allowed to vary. Figure 6.16 shows the fits on simulation for the considered track types and bremsstrahlung categories.

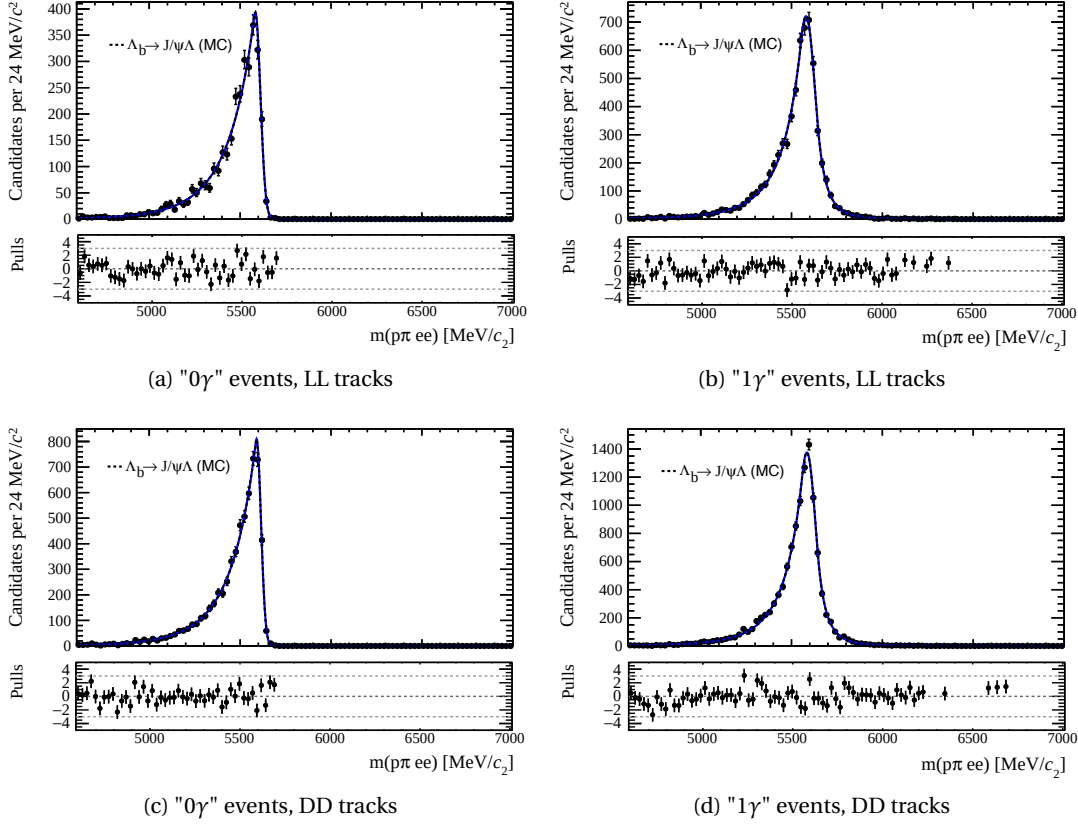


Figure 6.16 – Fit to  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$  simulation for LL (top) and DD (bottom) candidates and for the  $0\gamma$  (left) and  $1\gamma$  (right) categories.

The tail parameters obtained from the fit to simulation are subsequently fixed in the fit to data. The mean and width parameters of the shape are instead allowed to vary freely and the value obtained from simulation is only used as initial value. The small combinatorial background component in data that survives the selection is modelled in each category with an exponential PDF. The fits on the four considered categories are performed independently. Figure 6.17 show the results.

### 6.6.2 Fit to $\Lambda_b^0 \rightarrow \Lambda^0 J/\psi(\rightarrow \mu\mu)$

In the case of  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ , no preselection in mass is necessary as the peak has a better resolution and stands clearly above the background. On the other hand, to simplify the fit and for consistency with the electron mode, a PID cut is performed to remove the  $B^0 \rightarrow (K_S^0 \rightarrow \pi^+ \pi^-)(J/\psi \rightarrow \mu^+ \mu^-)$  background. The same cut is used as for the electron case



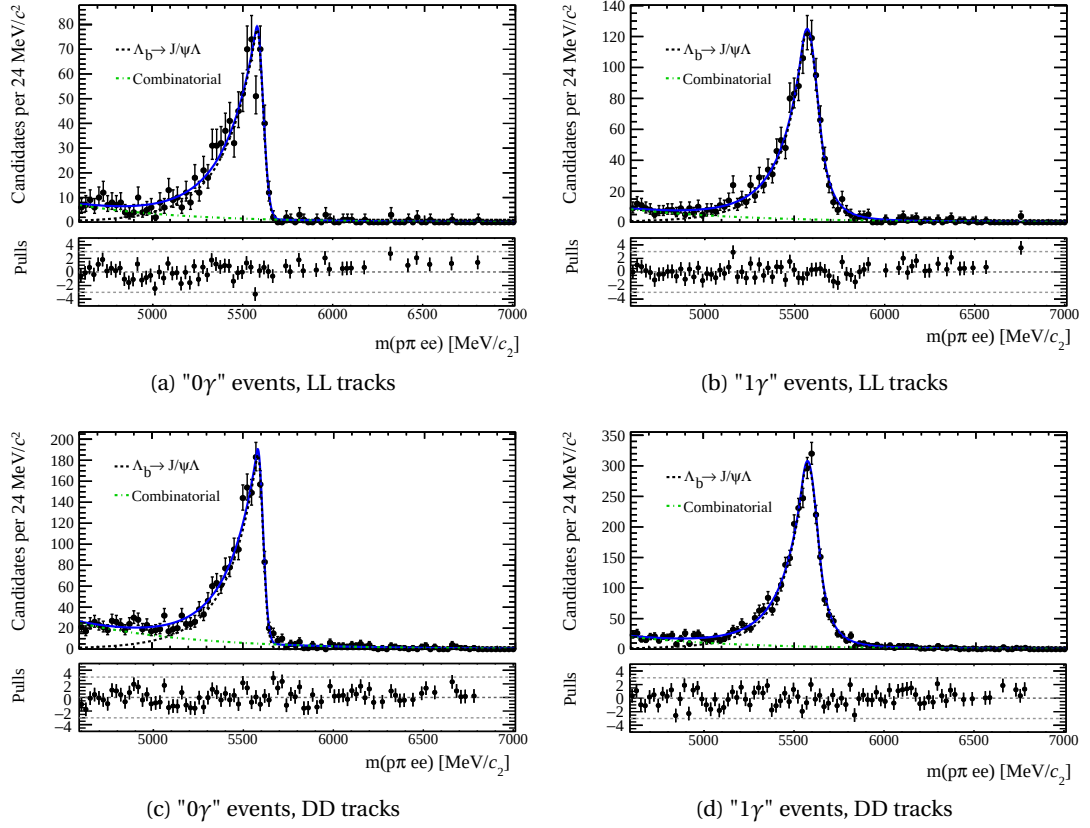


Figure 6.17 – Fit on  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow e^\pm e^\mp)$  Run I and Run II simulation for LL (top) and DD (bottom) candidates and for the 0 $\gamma$  (left) and 1 $\gamma$  (right) categories.

$\text{ProbNN} > 0.1$ . It was checked that cutting on PID does not significantly change the resolution of the  $\Lambda_b^0$  peak which is the quantity of interest for this calibration.

Following the same procedure as for the electron channel,  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  simulated candidates are fitted beforehand in order to obtain shape parameters. A DCB function is used to model the signal in all categories. The combinatorial background is described by an exponential shape with independent slope for each category.

Figure 6.18 shows fits to data invariant-mass distributions including both runs for long and downstream tracks.

### 6.6.3 Signal model for $\Lambda_b^0 \rightarrow \Lambda^0 e \mu$

Figure 6.19 shows fits to  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  simulated candidates separated in two bremsstrahlung categories. These shapes are used to fit data as explained in Section 6.9.1.

The signal PDF widths are then corrected using the factors calculated as described in Equation 6.4, shown in Table 6.8.

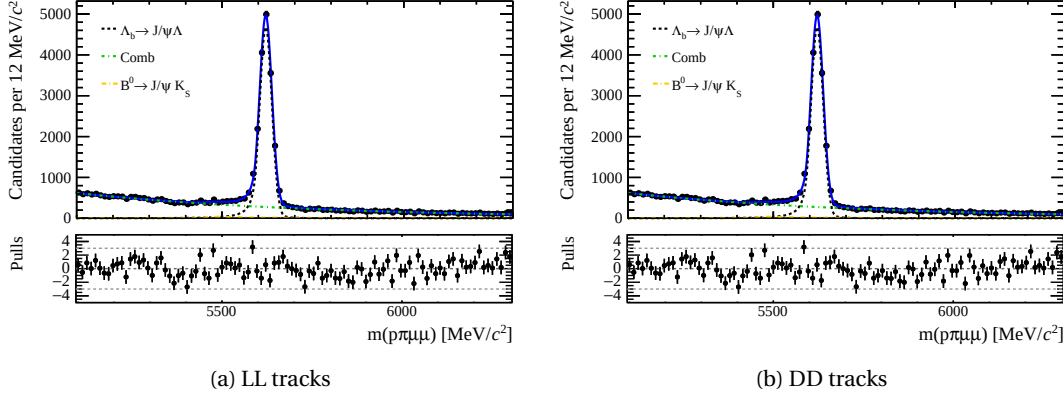


Figure 6.18 – Fit to  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  Run I and Run II data LL (left) and DD (right) candidates.

Table 6.8 – Results of  $\mathcal{C}$  factor evaluation.

|                                | $\mathcal{C}$     |
|--------------------------------|-------------------|
| <b>LL <math>0\gamma</math></b> | $1.156 \pm 0.071$ |
| <b>DD <math>0\gamma</math></b> | $1.129 \pm 0.060$ |
| <b>LL <math>1\gamma</math></b> | $1.140 \pm 0.051$ |
| <b>DD <math>1\gamma</math></b> | $1.124 \pm 0.038$ |

## 6.7 Normalisation

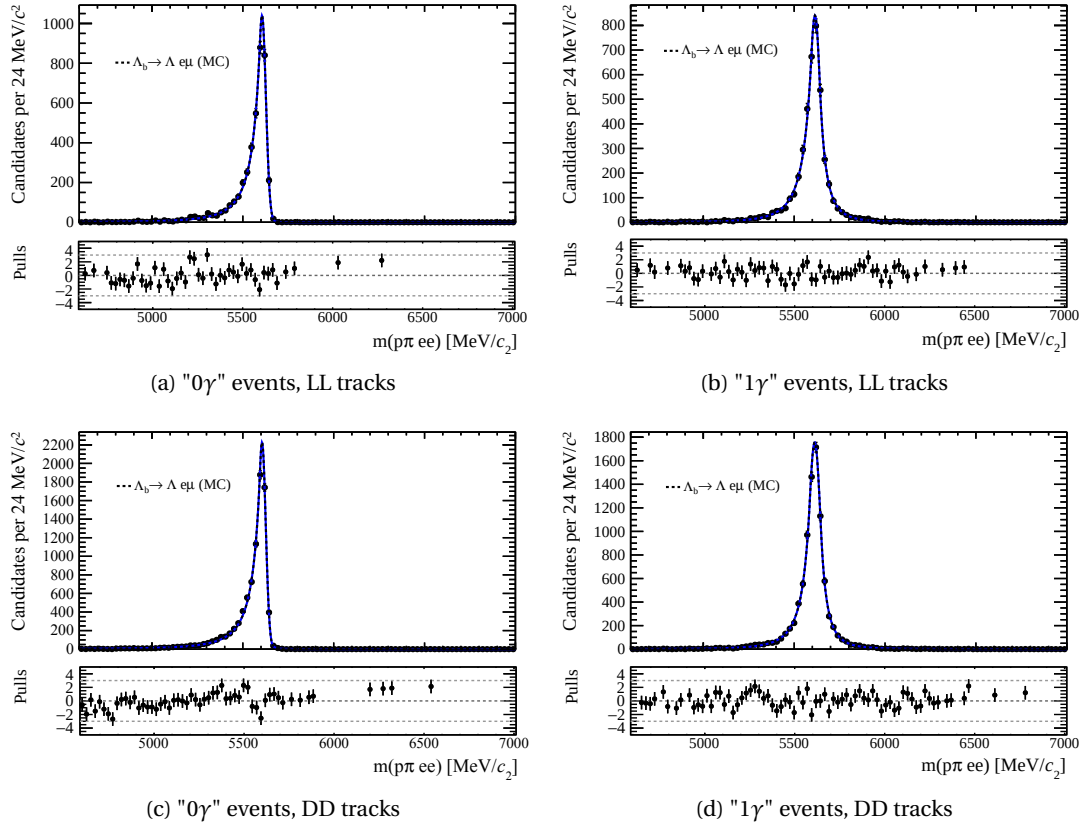
In order to translate an observed (or excluded) number of signal candidates to a branching fraction, the  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  decay is used as normalisation channel. The signal branching fraction is obtained as

$$\mathcal{B}_{\text{sig}} = \frac{N_{\text{sig}} \cdot \varepsilon_{\text{norm}}}{N_{\text{norm}} \cdot \varepsilon_{\text{sig}}} \times \mathcal{B}_{\text{norm}}, \quad (6.5)$$

where  $N$  indicates the number of observed candidates, obtained from the mass fits,  $\varepsilon$  indicates the total selection and reconstruction efficiency, and the subscript indicates whether these quantities are referred to the signal or to the normalisation channel.

The  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  events are selected as described in Section 6.2, and the same multivariate classifier applied on signal is used against combinatorial background. This allows to reduce systematic uncertainties from the classifier's efficiency, in the ratio with the signal yield. Only a single modification is applied to the classifier: as described previously in Section 6.4.4, the  $\alpha_{\text{HOP}}$  variable, not being defined for the  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  sample, is fixed to 1.

A cut on the response of the classifier,  $\text{BDT}^{\text{fix-hop}} > 0.6$  is applied.

Figure 6.19 – Fit to  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  simulated candidates.

The yield is extracted through a maximum-likelihood fit explained further in this section.

The LL and DD categories that are used to split the signal candidates, as well as the Run I and Run II categories, are also used as a criterion to perform independent fits on  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ . On the contrary, bremsstrahlung categories, not being defined on this channel, are not present in the normalisation fits.

### 6.7.1 Normalisation channel yield

The invariant-mass distribution of  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  decays is fitted to extract the yield to be used for normalisation. In this case, contrarily to what described in Section 6.6, we are not interested in preserving the resolution of the signal. We are instead interested in obtaining the signal yield in the most clean possible way reducing both statistical and systematic uncertainties. For this reason the fits are performed on the invariant mass variable calculated using a constraint on the invariant mass of the two muons to match the known  $J/\psi$  mass. This allows to improve the resolution on the  $\Lambda_b^0$  and at the same time to push partially-reconstructed backgrounds further into the sidebands, outside of the fitting mass window.

This method allows to only consider two background components:

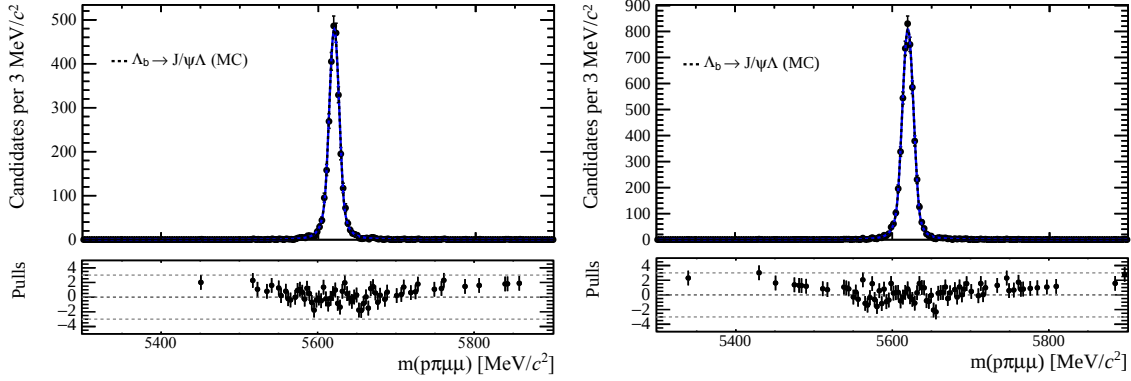


Figure 6.20 – Fits to the invariant mass of  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  simulated candidates built from LL (left) and DD (right) tracks.

|               | LL            | DD              |
|---------------|---------------|-----------------|
| <b>Run I</b>  | $1880 \pm 45$ | $4044 \pm 69$   |
| <b>Run II</b> | $4979 \pm 73$ | $12154 \pm 118$ |

Table 6.9 – Normalisation channel,  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ , yields in the considered categories.

- Combinatorial, described with an exponential PDF with separate slope parameters for each category.
- $B^0 \rightarrow (K_S^0 \rightarrow \pi^+ \pi^-)(J/\psi \rightarrow \mu^+ \mu^-)$ : this background falls below the signal peak and needs to be modelled. It is important especially for DD candidates. In the LL category, a stripping cut on the PID reduces its contribution, but the same cut cannot be applied for DD tracks as its efficiency cannot be studied due to the absence of an abundant calibration channel. This component is modelled with a non-parametric PDF obtained from MC.

The signal is also modelled using a DCB shape. Simulated  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  candidates are fit first. The results of these fits are shown in Figure 6.20. All parameters are allowed to float independently for each category. The shape obtained fitting the MC is then used to fit data with all tail parameters fixed. The mean and width parameters are instead free to float.

Finally, the fit performed on data is shown in Figure 6.21 and the obtained yields, which will be used to calculate the normalisation factors, are listed in Table 6.9.

## 6.8 Selection efficiencies

The efficiencies of geometrical acceptance, reconstruction and selection of the data are discussed in this section, both for what concerns signal ( $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ ) and normalisation ( $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$ ) candidates.

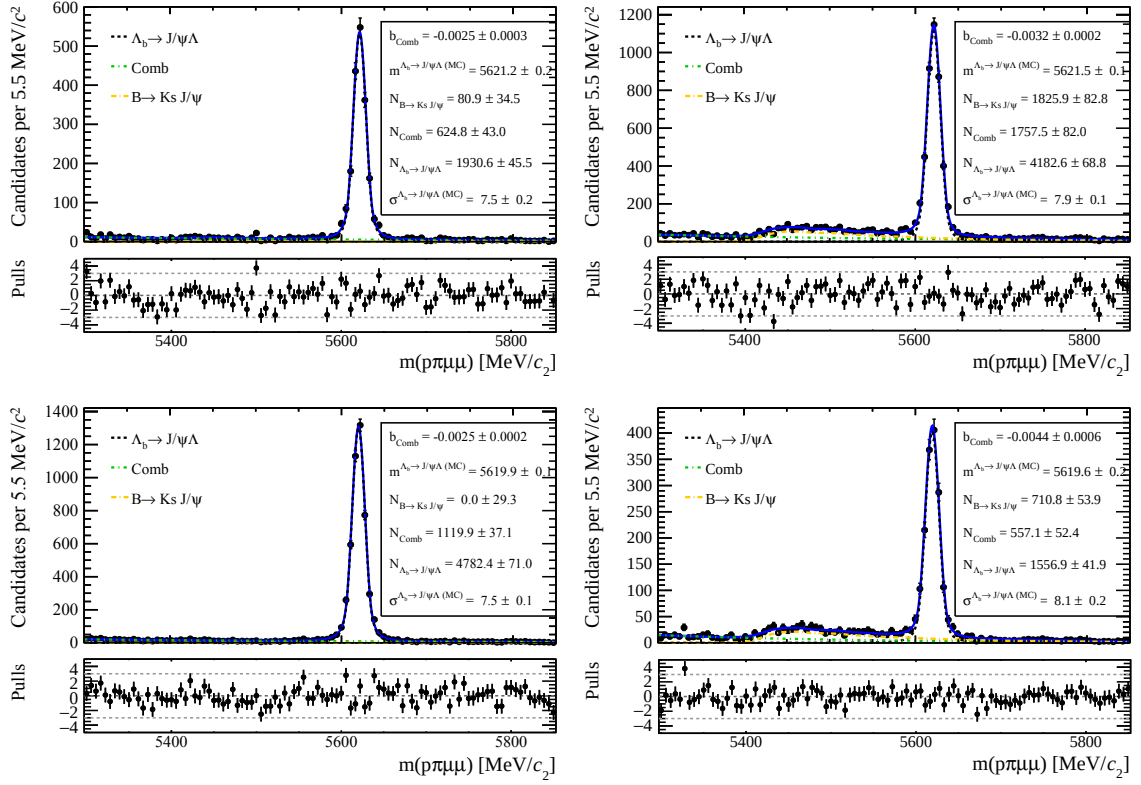


Figure 6.21 – Fits to the invariant mass of  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^+\mu^-)$  data candidates built from LL (left) and DD (right) tracks; Run I (top) and Run II (bottom).

### 6.8.1 Geometrical acceptance

Geometrical acceptance efficiencies are computed during the generation of the MC, where precise information about the detector shape and active area is used. The efficiency is extracted as the fraction between the number of events falling in the acceptance and the total number of generated events.

Table 6.10 contains the geometrical acceptance efficiencies for signal and normalisation channels, split by magnet polarity.

Table 6.10 – Geometrical acceptance efficiencies for  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  and  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ .

| Year | $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ |                       | $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ |                       |
|------|---------------------------------------------------|-----------------------|--------------------------------------------------------------------------|-----------------------|
|      | Mag Up                                            | Mag Down              | Mag Up                                                                   | Mag Down              |
| 2011 | $0.19123 \pm 0.00049$                             | $0.18998 \pm 0.00048$ | $0.1881 \pm 0.000727$                                                    | $0.1897 \pm 0.000719$ |
| 2012 | $0.19414 \pm 0.0005$                              | $0.19321 \pm 0.00051$ | $0.1881 \pm 0.000727$                                                    | $0.1897 \pm 0.000719$ |
| 2015 | $0.20389 \pm 0.00077$                             | $0.20568 \pm 0.00078$ | $0.19838 \pm 0.0005$                                                     | $0.19832 \pm 0.00049$ |
| 2016 | $0.20437 \pm 0.00073$                             | $0.20474 \pm 0.00073$ | $0.19838 \pm 0.0005$                                                     | $0.19832 \pm 0.00049$ |

### 6.8.2 Trigger

Trigger efficiencies for L0 and HLT1 are computed using a tag-and-probe technique on  $B^+ \rightarrow (J/\psi \rightarrow \mu^\pm \mu^\mp) K^+$  data: candidates are selected by triggering on one of the two muons, and the efficiency is extracted for the other one, thus from an un-biased and clean sample of muon tracks. This allows to extract efficiencies directly from data, without relying on the correct modelling of the trigger response on MC. Conversely, HLT2 efficiencies are obtained directly from simulation, as the response of this trigger is well reproduced. Table 6.11 contains the total trigger efficiencies for the signal and normalisation channels.

Table 6.11 – Trigger efficiencies for  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  and  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ .

| Year                                                                     | Run I               | Run II              |
|--------------------------------------------------------------------------|---------------------|---------------------|
| $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$                        | $0.1955 \pm 0.0013$ | $0.3466 \pm 0.0012$ |
| $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ | $0.4396 \pm 0.0020$ | $0.6373 \pm 0.0014$ |

### 6.8.3 Reconstruction and stripping

Reconstruction and stripping efficiencies are computed together by measuring the ratio, on simulation, between the number of candidates stored after the stripping and the number of signal events generated in the detector acceptance. Candidates with LL and DD tracks are considered separately. The track type is a property that arises in the reconstruction, thus it is not defined before. The efficiencies shown in this section are therefore normalised to the total number of generated events in acceptance, *i.e.* they include the fraction of events recon-

structured as LL and DD, respectively. These fraction are roughly 30% for LL and (consequently) 70% for DD.

The full list of stripping requirements is provided in Section 6.2.2.

All reconstruction and stripping efficiencies, for  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  and  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  in both LHC runs and track types, are reported in Table 6.12.

Table 6.12 – Reconstruction and Stripping efficiencies.

| Sample                                            | RunI-LL               | RunI-DD               | RunII-LL              | RunII-DD              |
|---------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ | $0.01253 \pm 0.00035$ | $0.01905 \pm 0.00032$ | $0.01497 \pm 0.00038$ | $0.01907 \pm 0.00045$ |
| $\Lambda_b^0 \rightarrow \Lambda^0 J/\psi$        | $0.01253 \pm 0.00035$ | $0.02896 \pm 0.00053$ | $0.01497 \pm 0.00038$ | $0.02700 \pm 0.00051$ |

#### 6.8.4 Offline selection

The efficiency of offline selection cuts is obtained from simulation. The ratio of the number of correctly-reconstructed signal events before and after applying the cuts is computed, re-weighting the samples as described in Section 6.3.

The efficiencies are computed for each cut forming the selection, always conditionally to the previous cut. The total efficiency is also reported.

The names used to identify the cuts are those explained in Section 6.2.3.

Tables 6.13 and 6.14 show the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  offline selection efficiencies for the LL and DD categories respectively.

Table 6.13 – Offline selection efficiencies from  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  MC with LL tracks.

| Selection                          | 2011                                  | 2012                                  | 2015                                  | 2016                                  |
|------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| $\hookrightarrow$ HasDet           | $0.9027 \pm 0.0030$                   | $0.9019 \pm 0.0025$                   | $0.8932 \pm 0.0041$                   | $0.8980 \pm 0.0019$                   |
| $\hookrightarrow$ Mass             | 1.0                                   | 1.0                                   | 1.0                                   | 1.0                                   |
| $\hookrightarrow$ $J/\psi$ veto    | $0.9368 \pm 0.0025$                   | $0.9344 \pm 0.0021$                   | $0.9335 \pm 0.0033$                   | $0.9356 \pm 0.0015$                   |
| $\hookrightarrow$ $\Lambda^0$ mass | $0.99739 \pm 0.00054$                 | $0.99832 \pm 0.00036$                 | $0.99754 \pm 0.00068$                 | $0.99730 \pm 0.00033$                 |
| $\hookrightarrow$ Fiducial         | $0.9361 \pm 0.0025$                   | $0.9585 \pm 0.0017$                   | $0.9561 \pm 0.0027$                   | $0.9510 \pm 0.0013$                   |
| $\hookrightarrow$ HOP-FD           | $0.99858 \pm 0.00041$                 | $0.99818 \pm 0.00038$                 | $0.99881 \pm 0.00048$                 | $0.99870 \pm 0.00024$                 |
| <b>Total</b>                       | <b><math>0.7943 \pm 0.0035</math></b> | <b><math>0.8123 \pm 0.0028</math></b> | <b><math>0.8078 \pm 0.0047</math></b> | <b><math>0.8128 \pm 0.0022</math></b> |

Similarly, Tables 6.15 and 6.16 show the  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  offline selection efficiencies for the LL and DD categories respectively.

#### 6.8.5 Particle identification

Particle identification efficiencies are obtained using the PIDCalib tool, already discussed in the previous chapter for the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis. The description of the procedure is not

## Chapter 6. Search for the lepton-flavour violating decay $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$

Table 6.14 – Offline selection efficiencies from  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  MC with DD tracks.

| Selection                        | 2011                                  | 2012                                  | 2015                                  | 2016                                  |
|----------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| $\hookrightarrow$ HasDet         | $0.9129 \pm 0.0020$                   | $0.9190 \pm 0.0016$                   | $0.9060 \pm 0.0027$                   | $0.9119 \pm 0.0013$                   |
| $\hookrightarrow$ Mass           | $0.999944 \pm 0.000056$               | 1.0                                   | $0.99981 \pm 0.00014$                 | $0.999933 \pm 0.000038$               |
| $\hookrightarrow J/\psi$ veto    | $0.9361 \pm 0.0017$                   | $0.9334 \pm 0.0014$                   | $0.9319 \pm 0.0023$                   | $0.9343 \pm 0.0011$                   |
| $\hookrightarrow \Lambda^0$ mass | $0.99659 \pm 0.00043$                 | $0.99614 \pm 0.00037$                 | $0.99616 \pm 0.00060$                 | $0.99596 \pm 0.00029$                 |
| $\hookrightarrow$ Fiducial       | $0.9773 \pm 0.0011$                   | $0.97931 \pm 0.00081$                 | $0.9771 \pm 0.0014$                   | $0.97689 \pm 0.00067$                 |
| $\hookrightarrow$ HOP-FD         | $0.99726 \pm 0.00038$                 | $0.99747 \pm 0.00029$                 | $0.99804 \pm 0.00042$                 | $0.99715 \pm 0.00024$                 |
| <b>Total</b>                     | <b><math>0.8255 \pm 0.0024</math></b> | <b><math>0.8301 \pm 0.0019</math></b> | <b><math>0.8360 \pm 0.0031</math></b> | <b><math>0.8403 \pm 0.0015</math></b> |

Table 6.15 – Offline selection efficiencies from  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  MC with LL tracks.

| Selection                        | 2011                                  | 2012                                  | 2015                                | 2016                                  |
|----------------------------------|---------------------------------------|---------------------------------------|-------------------------------------|---------------------------------------|
| $\hookrightarrow$ Mass           | $0.99748 \pm 0.00076$                 | $0.99565 \pm 0.00076$                 | $0.9951 \pm 0.0028$                 | $0.99605 \pm 0.00082$                 |
| $\hookrightarrow$ HasDet-p       | 1.0                                   | 1.0                                   | 1.0                                 | $0.99483 \pm 0.00094$                 |
| $\hookrightarrow$ HasDet- $\mu$  | 1.0                                   | 1.0                                   | 1.0                                 | $0.8822 \pm 0.0042$                   |
| $\hookrightarrow \Lambda^0$ mass | $0.99840 \pm 0.00060$                 | $0.99796 \pm 0.00052$                 | $0.9984 \pm 0.0016$                 | $0.99609 \pm 0.00087$                 |
| $\hookrightarrow$ Fiducial       | $0.9265 \pm 0.0039$                   | $0.9532 \pm 0.0025$                   | $0.9346 \pm 0.0099$                 | $0.9516 \pm 0.0028$                   |
| $\hookrightarrow J/\psi$ mass    | $0.9389 \pm 0.0037$                   | $0.9376 \pm 0.0029$                   | $0.930 \pm 0.011$                   | $0.9349 \pm 0.0033$                   |
| <b>Total</b>                     | <b><math>0.8591 \pm 0.0044</math></b> | <b><math>0.8923 \pm 0.0031</math></b> | <b><math>0.860 \pm 0.012</math></b> | <b><math>0.7775 \pm 0.0051</math></b> |

Table 6.16 – Offline selection efficiencies from  $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$  MC with DD tracks.

| Selection                        | 2011                                  | 2012                                  | 2015                                  | 2016                                  |
|----------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| $\hookrightarrow$ HasDet- $\mu$  | 1.0                                   | 1.0                                   | 1.0                                   | $0.8836 \pm 0.0030$                   |
| $\hookrightarrow J/\psi$ mass    | $0.9387 \pm 0.0020$                   | $0.9384 \pm 0.0015$                   | $0.9453 \pm 0.0055$                   | $0.9374 \pm 0.0024$                   |
| $\hookrightarrow$ Fiducial       | $0.9796 \pm 0.0012$                   | $0.97849 \pm 0.00094$                 | $0.9747 \pm 0.0039$                   | $0.9757 \pm 0.0015$                   |
| $\hookrightarrow$ Mass           | $0.99749 \pm 0.00043$                 | $0.99591 \pm 0.00041$                 | $0.9982 \pm 0.0010$                   | $0.99545 \pm 0.00064$                 |
| $\hookrightarrow \Lambda^0$ mass | $0.99717 \pm 0.00045$                 | $0.99735 \pm 0.00032$                 | $0.9959 \pm 0.0015$                   | $0.99732 \pm 0.00049$                 |
| <b>Total</b>                     | <b><math>0.9238 \pm 0.0020</math></b> | <b><math>0.9141 \pm 0.0016</math></b> | <b><math>0.9208 \pm 0.0058</math></b> | <b><math>0.8096 \pm 0.0035</math></b> |

repeated here, and the efficiencies can be found in Table 6.17, split per year of data taking and track type. The efficiencies for 2017 cannot yet be directly determined due to the unavailability of calibration samples, currently in preparation. Therefore, 2016 efficiencies are used for this year.

### 6.8.6 Selection on the BDT response

BDT efficiencies for signal and normalisation channels are reported in Table 6.18. They are determined as described in the following paragraphs.



Table 6.17 – PID efficiencies for  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  and  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ , divided per year and track type.

| Sample                                                                   | Year | Track Type | Efficiency            |
|--------------------------------------------------------------------------|------|------------|-----------------------|
| $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$                        | 2011 | LL         | $0.9535 \pm 0.0015$   |
|                                                                          |      | DD         | $0.9744 \pm 0.0014$   |
|                                                                          | 2012 | LL         | $0.9466 \pm 0.0012$   |
|                                                                          |      | DD         | $0.9721 \pm 0.0012$   |
|                                                                          | 2015 | LL         | $0.9121 \pm 0.0018$   |
|                                                                          |      | DD         | $0.94073 \pm 0.0018$  |
|                                                                          | 2016 | LL         | $0.9393 \pm 0.0007$   |
|                                                                          |      | DD         | $0.9660 \pm 0.0007$   |
| $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ | 2011 | LL         | $0.9427 \pm 0.0007$   |
|                                                                          |      | DD         | $0.9592 \pm 0.0006$   |
|                                                                          | 2012 | LL         | $0.9367 \pm 0.0005$   |
|                                                                          |      | DD         | $0.95716 \pm 0.00045$ |
|                                                                          | 2015 | LL         | $0.9016 \pm 0.0006$   |
|                                                                          |      | DD         | $0.9225 \pm 0.0006$   |
|                                                                          | 2016 | LL         | $0.92377 \pm 0.00020$ |
|                                                                          |      | DD         | $0.95086 \pm 0.00018$ |

$$\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$$

The efficiencies of the different cuts on BDT applied in the eight analysis categories are determined from simulation, using the test set for each training fold in order to avoid biases. The procedure described in Section 6.4.4 subsequently allows to obtain an estimate on the systematic error of this efficiency.

Table 6.18 – Efficiencies of the cut to the BDT response on signal in the eight analysis categories.

| Sample                                                                   | Category  | Efficiency (Run I)  | Efficiency (Run II) |
|--------------------------------------------------------------------------|-----------|---------------------|---------------------|
| $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$                        | Brem-LL   | $0.657 \pm 0.011$   | $0.561 \pm 0.018$   |
|                                                                          | NoBrem-LL | $0.505 \pm 0.008$   | $0.531 \pm 0.018$   |
|                                                                          | Brem-DD   | $0.539 \pm 0.006$   | $0.295 \pm 0.017$   |
|                                                                          | NoBrem-DD | $0.420 \pm 0.006$   | $0.324 \pm 0.017$   |
| $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ | LL-RunI   | $0.5929 \pm 0.0047$ | $0.639 \pm 0.008$   |
|                                                                          | DD-RunI   | $0.6674 \pm 0.0042$ | $0.704 \pm 0.008$   |

$$\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$$

The cut applied on the response of the *modified* BDT with fixed  $\alpha_{\text{HOP}}$  ( $\text{BDT}^{\text{fix-hop}}$ , see Section 6.4.4) is  $\text{BDT}^{\text{fix-hop}} > 0.6$ , common between all the categories. Its efficiency is estimated directly from  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  MC, and its systematic uncertainty obtained from the difference with the same efficiency computed on s-weighted data.

### 6.8.7 Total Efficiency

Table 6.19 contains the total efficiency per category on signal and normalisation channels. These efficiencies do not contain the fraction of events for the relative category, *i.e.* they are normalised independently, in order to allow an easier comparison between the two channels.

Table 6.19 – Total selection efficiencies for  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  and  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  in each category.

| Category       | $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ | $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ |
|----------------|---------------------------------------------------|--------------------------------------------------------------------------|
| Brem-LL-Run1   | $(8.43 \pm 0.19) \times 10^{-4}$                  | $(1.558 \pm 0.030) \times 10^{-3}$                                       |
| NoBrem-LL-Run1 | $(6.48 \pm 0.15) \times 10^{-4}$                  |                                                                          |
| Brem-DD-Run1   | $(7.34 \pm 0.15) \times 10^{-4}$                  | $(1.880 \pm 0.033) \times 10^{-3}$                                       |
| NoBrem-DD-Run1 | $(5.72 \pm 0.13) \times 10^{-4}$                  |                                                                          |
| Brem-LL-Run2   | $(1.202 \pm 0.044) \times 10^{-3}$                | $(2.672 \pm 0.056) \times 10^{-3}$                                       |
| NoBrem-LL-Run2 | $(1.139 \pm 0.043) \times 10^{-3}$                |                                                                          |
| Brem-DD-Run2   | $(6.63 \pm 0.41) \times 10^{-4}$                  | $(3.078 \pm 0.061) \times 10^{-3}$                                       |
| NoBrem-DD-Run2 | $(7.29 \pm 0.41) \times 10^{-4}$                  |                                                                          |

## 6.9 Results

### 6.9.1 Invariant-mass fit

As anticipated in the introduction of this chapter, the present analysis is currently being finalised, and therefore the signal mass region of the data will be kept blind until the strategy is fully approved by the LHCb collaboration.

A preliminary maximum-likelihood fit to the invariant mass sidebands has been performed and can be observed in Figure 6.22, divided into the eight categories in which the analysis is performed. The background is currently assumed to be dominated by combinatorial events, and it will be possibly updated when the study introduced in Section 6.5 is finalised.

The signal component, not present in the blind fit, is described by a Double-sided Crystal Ball PDE, as described in Section 6.6, with parameters constrained to their respective values on simulation.

### 6.9.2 CLs limit

In case of no evidence of signal, the  $\text{CL}_s$  [117] technique will be used to compute an upper limit on the branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . The procedure is the same used in the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis, described in Section 5.5. Also in this case, the implementation provided in the RooStats [118] package is used. In particular, the frequentist method with toys datasets is adopted.

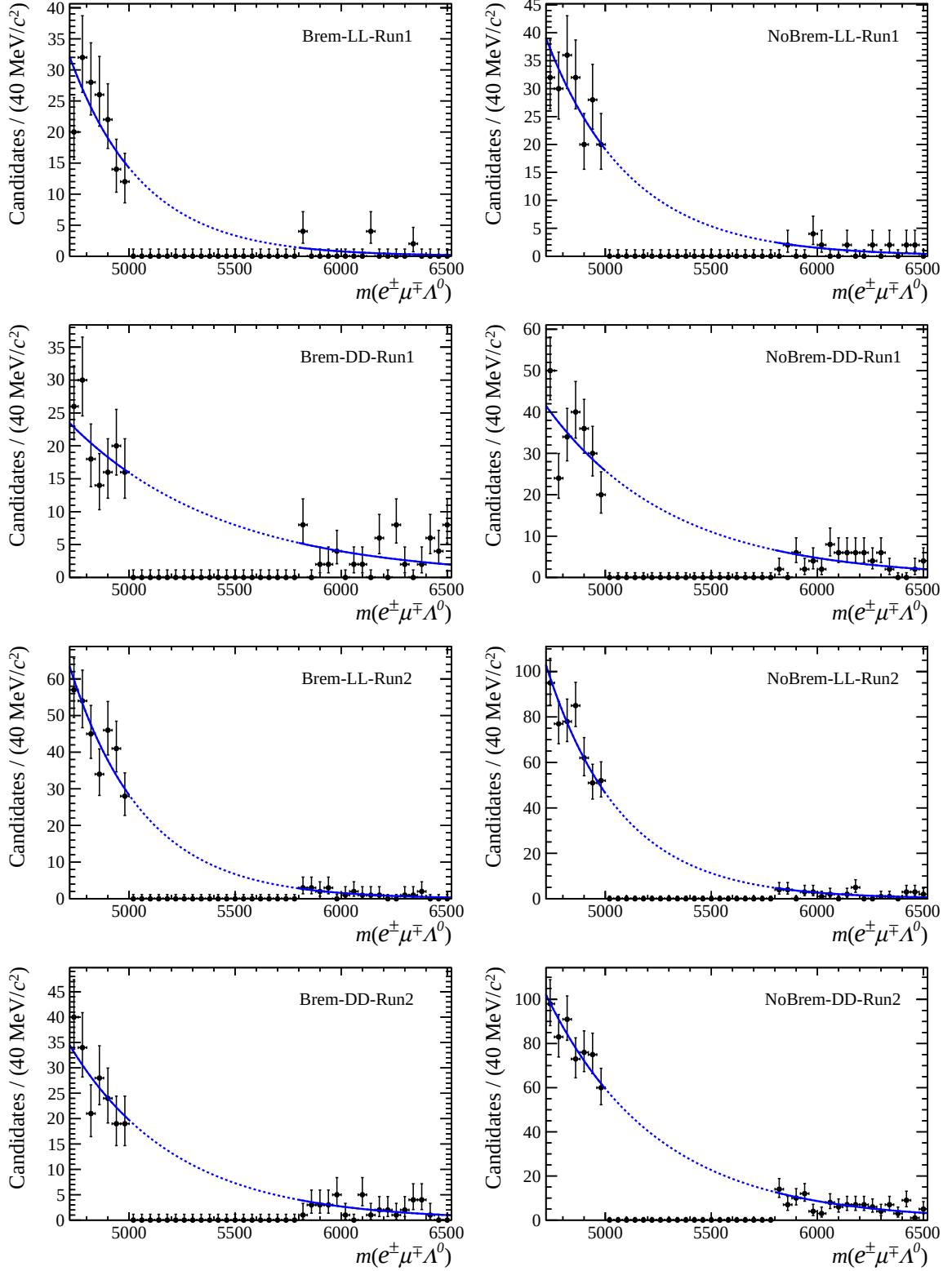


Figure 6.22 – Blind mass fits to the invariant mass of  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  data.

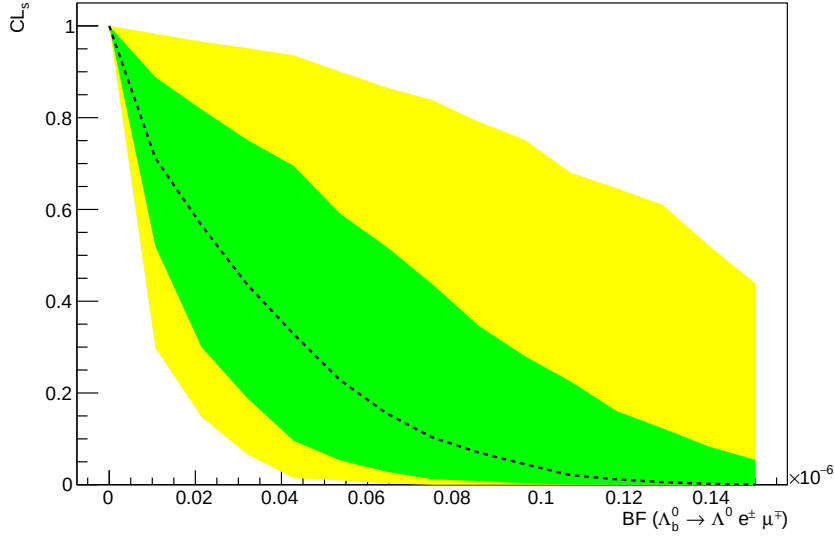


Figure 6.23 – CLs limit scan on the branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . This limit is obtained using the average branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  from the PDF [2].

The single-event sensitivity,  $\alpha$ , defined by Equation 6.5 with  $N_{\text{sig}} = 1$  is found to be

$$\alpha = 3.1 \times 10^{-9}. \quad (6.6)$$

A preliminary expected limit scan, shown in Figure 6.23, provides an upper boundary of

$$9.43 \times 10^{-8} \text{ at } 95\% \text{CL}.$$

The reason for the very large uncertainty in this plot is addressed in the following pages.

### 6.9.3 Treatment of systematic uncertainties

As in the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis, systematic uncertainties on parameters such as selection efficiencies, signal and background shapes and fractions of each category are propagated to the CLs limit by allowing each parameter to fluctuate in the toys and in the fit PDFs, by multiplying the likelihood function with a gaussian distribution centered in the nominal value, and having the parameter's systematic uncertainty as the standard deviation. Figure 6.24 shows the relative uncertainty on each parameter. These parameters include:

- the branching fraction and the number of observed candidates of the normalisation channel, respectively `BFnorm` and `Nnorm`;
- the fraction, `frac`, of signal events expected in each of the eight categories;

- the selection efficiency in each category for the signal, `eff_sig`, and the normalisation channel, `eff_norm`;
- the parameters of the Crystal Ball signal PDF for each category, obtained as described in Section 6.6, `m` (mean), `s` ( $\sigma$ ), `a` and `a2` (tail parameters).

Fixing all the constrained parameters to their central value yields a significantly narrower limit band, shown in Figure 6.25, although the effect on the expected limit is tiny, reducing its value (at 95% CL) to  $9.32 \times 10^{-8}$  *i.e.* by about 1%.

#### 6.9.4 Uncertainty on the branching fraction of $\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)$

The uncertainty on the branching fraction of the normalisation channel largely dominates between the systematic uncertainties, with its relative magnitude of about 54%:

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0 J/\psi) = (5.7 \pm 3.1) \times 10^{-4}. \quad (6.7)$$

This reflects in the very large error band on the branching fraction upper limit shown in Figure 6.23. For this reason, a limit is evaluated also directly on the ratio of branching fractions of signal and normalisation:

$$r = \frac{\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp)}{\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp))}. \quad (6.8)$$

This quantity can indeed be calculated without explicit knowledge on  $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp))$ , since Equation 6.5 would become:

$$r = \frac{N_{\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp} \cdot \mathcal{E}_{\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp)}}{N_{\Lambda_b^0 \rightarrow \Lambda^0 J/\psi} \cdot \mathcal{E}_{\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp}}. \quad (6.9)$$

A result obtained on  $r$  could be reinterpreted at a later time, with possible new measurements of  $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0(J/\psi \rightarrow \mu^\pm \mu^\mp))$ , to obtain a more precise estimate of  $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp)$ .

Figure 6.26 shows the  $\text{CL}_s$  scan on  $r$ , which results in an upper limit of

$$1.67 \times 10^{-4} \text{ at } 95\% \text{CL}.$$

## 6.10 Conclusions and prospects

An overview of the search for the lepton flavour violating decay  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ , performed in the context of this thesis, was provided in this chapter.

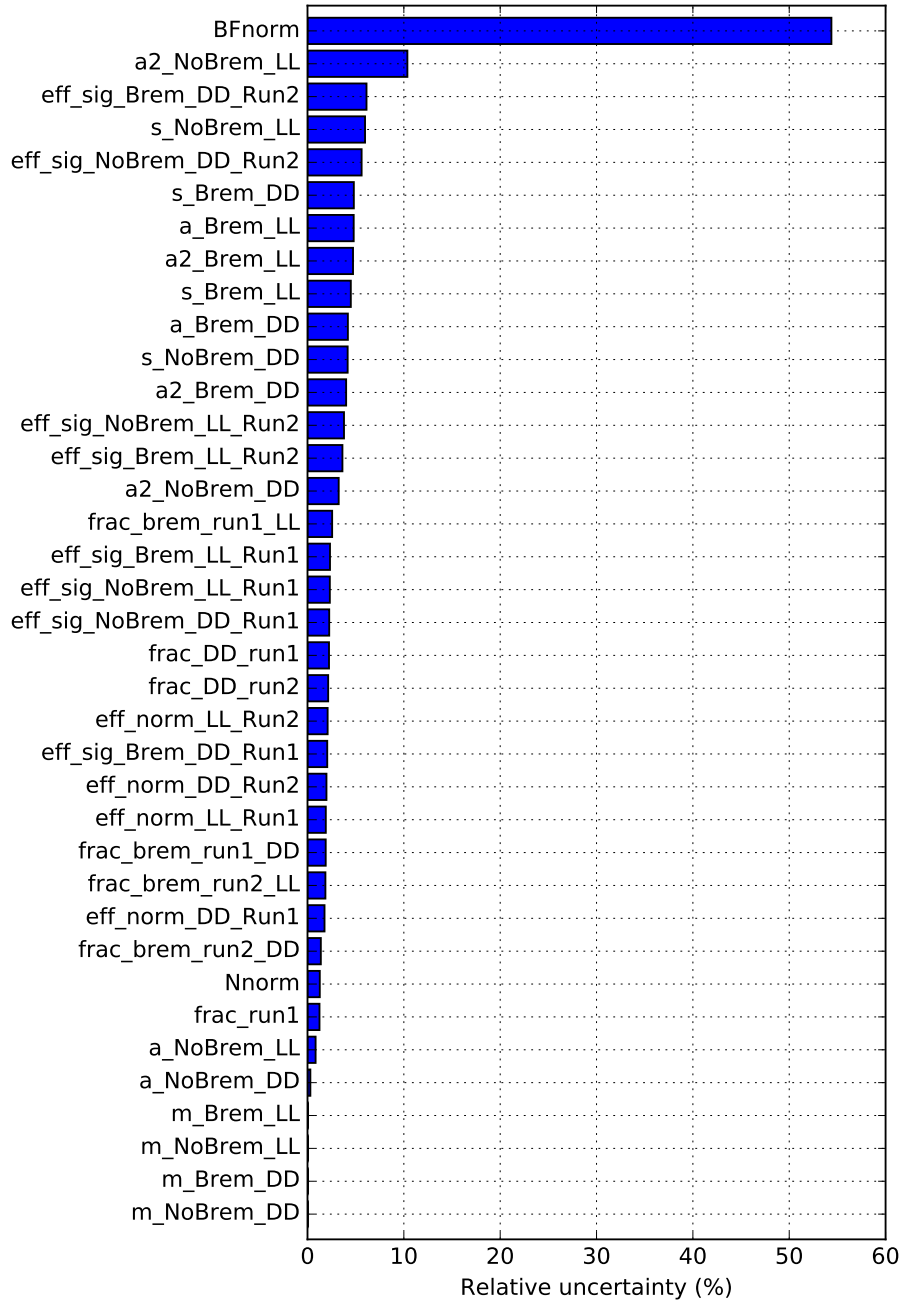


Figure 6.24 – Relative uncertainties on nuisance parameters and global observables entering the computation of the branching fraction upper limit on  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ .

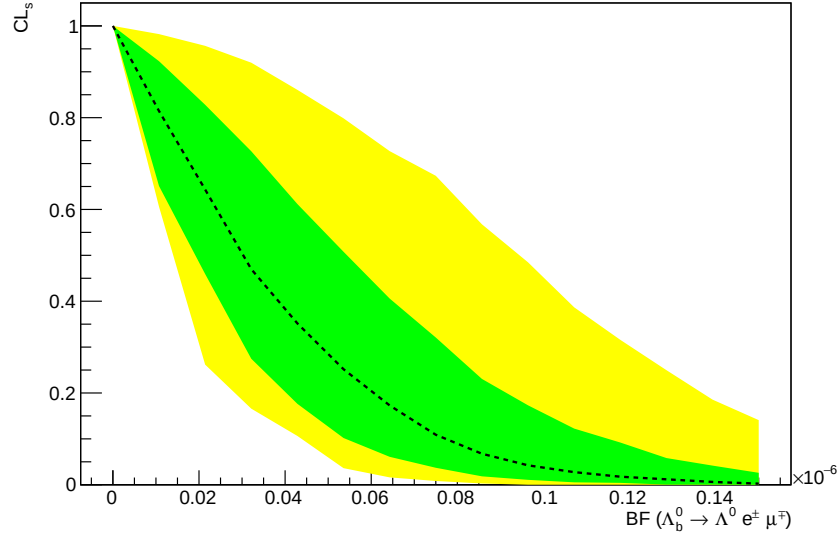


Figure 6.25 – CL<sub>s</sub> limit scan on the branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$ . This limit is obtained using the average branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  from the PDF [2], and removing all the systematic constraints.

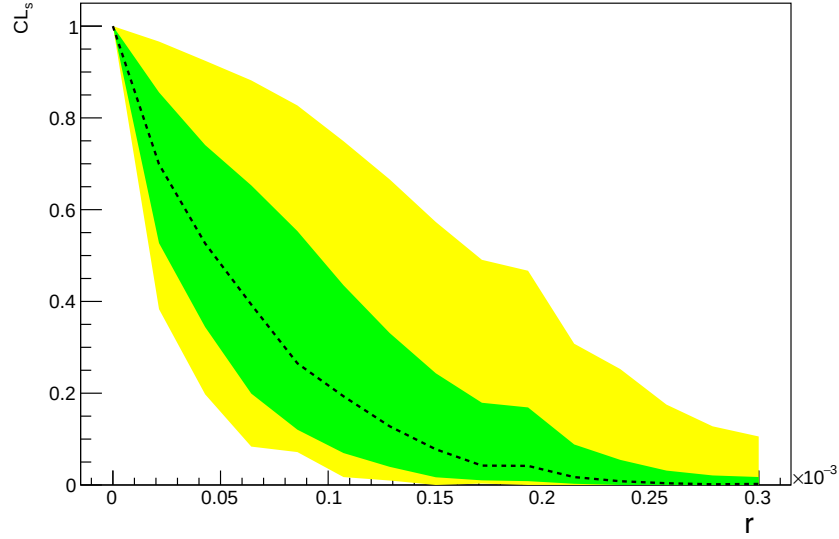


Figure 6.26 – CL<sub>s</sub> limit scan on the ratio,  $r$ , of the branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  over the branching fraction of  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$ .

The analysis is almost finalised, and it is expected to be able to probe a very interesting interval of signal branching fraction, potentially constraining or confirming new physics with LFV. A study of the exclusive background sources is ongoing and will allow to achieve a more precise modelling of the invariant mass distribution below the expected signal.

Further plans include a precise measurement of the branching fraction of the  $\Lambda_b^0 \rightarrow \Lambda^0 (J/\psi \rightarrow \mu^\pm \mu^\mp)$  decay in LHCb, as this is currently poorly known and constitutes the main source of systematic uncertainty in the calculation of the signal branching fraction or its upper limit.



## 7 Conclusion and outlook

In summary, two analyses have been presented in this thesis, searching for lepton flavour violating phenomena, excluded in the Standard Model but predicted by alternative theories. In the first analysis, the  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$  decays are excluded up to branching fractions of the order  $10^{-9}$  using data from the Run 1 of LHC. The second analysis, performed on data collected up to year 2017, will be able to probe for the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  decay down to branching fractions of order  $10^{-8}$  and, in case of no evidence of signal, it will add a tight constraint to new physics models. In the opposite scenario, a hypothetical significant observation would represent an unprecedented revolution.

The processes investigated here are part of a wide range of very interesting phenomena, in a largely unexplored territory. The current situation, which sees mild tensions in the lepton flavour universality assumption and yet no observation of lepton flavour violation demands the continuation of such searches, to find an answer to the open questions mentioned in the introduction of this thesis. LFV studies are indeed currently receiving sizeable attention in the particle physics world, as they can potentially provide explanations to such questions, and a portal to a new sector of particle physics, opening to a new era of research.

The large amount of data expected in the Run III of LHC will help to clarify the current situation, significantly reducing the dominant statistical uncertainties, thus allowing more precise results, both in the case of observation and exclusion of LFV processes.

Thanks to the substantial planned upgrade – part of which was also discussed in this thesis – the LHCb detector will be able to collect and record clean data at an unprecedented rate, overcoming the difficulties due to the increased pile-up. Such data are precious for the above-mentioned studies and for a large variety of other physics analyses, including both completely new topics and updates of previously published papers using the larger available amount of data.



---

## **Appendix**



## A Mass resolution of $e - \mu$ final states

In order to correct the invariant mass resolution observed on Monte Carlo candidates for inaccuracies in the simulation, a data-driven correction factor is applied.

The absence of known decays into an electron-muon couple makes it impossible to derive this figure directly from data. For this reason, the signal final state resolution has to be inferred from two observable channels, one with two electrons and one with two muons, so that the contributions of the two different leptons can be appropriately combined.

### A.0.1 Invariant mass resolution in a relativistic 2-body decay

The invariant mass  $M$  of any two particles in the relativistic approximation ( $E \gg m$ ) can be computed as:

$$\begin{aligned}
 M &= (E_{TOT}^2 - \vec{p}_{TOT} \cdot \vec{p}_{TOT})^{\frac{1}{2}} \\
 &\simeq [(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2) \cdot (\vec{p}_1 + \vec{p}_2)]^{\frac{1}{2}} \\
 &= \sqrt{2E_1E_2(1 - \cos\theta)},
 \end{aligned} \tag{A.1}$$

where the subscripts 1 and 2 refer to the two particles and  $\theta$  is the angle between their space momenta. This allows to easily express the invariant mass resolution in terms of the energy resolutions on the two particles:

$$\begin{aligned}
 \frac{\sigma_M}{M} &\simeq \frac{1}{2} \sqrt{\frac{[2E_2(1 - \cos\theta)\sigma_{E_1}]^2 + [2E_1(1 - \cos\theta)\sigma_{E_2}]^2}{2E_1E_2(1 - \cos\theta)}} \\
 &= \frac{1}{2} \sqrt{\left[\left(\frac{\sigma_{E_1}}{E_1}\right)^2 + \left(\frac{\sigma_{E_2}}{E_2}\right)^2\right]}
 \end{aligned} \tag{A.2}$$

## Appendix A. Mass resolution of $e - \mu$ final states

---

where the dependency on  $\theta$  is neglected. The correlation between the two daughter's energy is also considered negligible. This latter assumption holds well when the parent particle has a sizable boost in the laboratory frame (which is the case for the two analyses presented in this thesis, see Section A.0.3); in the frame of the center-of-mass of the parent particle, on the other hand, the correlation between the kinematics of the two leptons is very high, due to the energy and momentum conservation laws.

In case of two equal particles, for example two electrons or two muons as in  $J/\psi \rightarrow ee$  or  $J/\psi \rightarrow \mu\mu$ :

$$\frac{\sigma_{E_1}}{E_1} = \frac{\sigma_{E_2}}{E_2} \doteq \frac{\sigma_E}{E} \quad (\text{A.3})$$

are expected to be the same, and Equation A.2 can be further simplified:

$$\frac{\sigma_M}{M} \simeq \frac{\sigma_E}{E}. \quad (\text{A.4})$$

This result allows to approximate the energy resolution of an electron (muon) with the mass resolution of a particle decaying into a  $ee$  ( $\mu\mu$ ) final state. Therefore, using Equations A.2 and A.4, the mass resolution for an  $e\mu$  final state can be expressed as

$$\left(\frac{\sigma_M}{M}\right)_{e\mu} \simeq \frac{1}{2} \sqrt{\left[\left(\frac{\sigma_M}{M}\right)_{ee}^2 + \left(\frac{\sigma_M}{M}\right)_{\mu\mu}^2\right]} \quad (\text{A.5})$$

This result has been used to correct the mass resolution of simulated  $B \rightarrow e\mu$  by fitting the mass spectra of  $J/\psi \rightarrow ee$  and  $J/\psi \rightarrow \mu\mu$  decays from  $B^+ \rightarrow J/\psi K^+$ . A correction factor  $C$  to be applied to the MC can be computed as

$$\begin{aligned} C &= \frac{\sqrt{\left(\frac{\sigma_E}{E}\right)_{e,DATA}^2 + \left(\frac{\sigma_E}{E}\right)_{\mu,DATA}^2}}{\sqrt{\left(\frac{\sigma_E}{E}\right)_{e,MC}^2 + \left(\frac{\sigma_E}{E}\right)_{\mu,MC}^2}} \\ &\simeq \frac{\sqrt{\left(\frac{\sigma_M}{M}\right)_{ee,DATA}^2 + \left(\frac{\sigma_M}{M}\right)_{\mu\mu,DATA}^2}}{\sqrt{\left(\frac{\sigma_M}{M}\right)_{ee,MC}^2 + \left(\frac{\sigma_M}{M}\right)_{\mu\mu,MC}^2}} \quad (\text{A.6}) \end{aligned}$$

This correction factor, obtained from a  $J/\psi$  decay can be applied to  $B$  decays despite the larger mass in the assumption that it is constant. Experimental observations show that  $\frac{\sigma_M}{M}$  itself in LHCb is constant in a mass range between those of the  $J/\psi$  and the  $\Upsilon(1S)$  resonances.

### A.0.2 Extrapolation to the case of $\Lambda_b \rightarrow \Lambda^0 e \mu$

In the case of  $\Lambda_b \rightarrow \Lambda^0 e \mu$ , with  $\Lambda^0$  decaying to  $p \pi^-$ , the final state is composed by 4 particles, and Equation A.2 would become more complicated. Nevertheless, the decay can be seen as a two-body decay  $\Lambda_b \rightarrow \Lambda^0 X$ , where X is a fictitious resonance describing the two leptons. These are not necessarily expected to come from a resonance and thus to show a peaking mass structure, but this doesn't compromise the following calculations.

At this point, the mass resolution of the  $\Lambda_b$  can be expressed as

$$\left(\frac{\sigma_M}{M}\right)_{\Lambda_b} \simeq \frac{1}{2} \sqrt{\left[\left(\frac{\sigma_E}{E}\right)_{\Lambda^0}^2 + \left(\frac{\sigma_E}{E}\right)_{e\mu}^2\right]} \quad (\text{A.7})$$

The term  $\left(\frac{\sigma_M}{M}\right)_{e\mu}$  can be obtained from  $ee$  and  $\mu\mu$  resonant decays like in Equation A.5, while  $\left(\frac{\sigma_M}{M}\right)_{\Lambda}$  can be directly measured.

Similarly to Equation A.6, a correction factor can be then computed for the invariant mass of  $\Lambda_b \rightarrow \Lambda^0 e \mu$  as:

$$C \simeq \frac{\sqrt{\left(\frac{\sigma_E}{E}\right)_{e\mu,DATA}^2 + \left(\frac{\sigma_E}{E}\right)_{\Lambda^0,DATA}^2}}{\sqrt{\left(\frac{\sigma_E}{E}\right)_{e\mu,MC}^2 + \left(\frac{\sigma_E}{E}\right)_{\Lambda^0,MC}^2}} \quad (\text{A.8})$$

Nevertheless, in this factor, the term relative to the  $e\mu$  couple still needs to be factorised, as it cannot be directly extracted from data of SM processes. This operation leads to a more complicated formula than Equation A.6, which is not used in the  $\Lambda_b \rightarrow \Lambda^0 e \mu$  analysis.

### A.0.3 Energy correlations

A crucial assumption for the validity of Equation A.2 is the absence of a large correlation between the energies of the two daughter particles. As expected from the conservation of energy in a two-body decays, this would not be the case in the center-of-mass reference system of the mother particle. Nevertheless, due to the large boost that the latter has in LHCb, the 2-dimensional distribution of the two energies is very different in the laboratory frame, and it shows a much lower correlation, as shown in Figure A.1.

## Appendix A. Mass resolution of $e - \mu$ final states

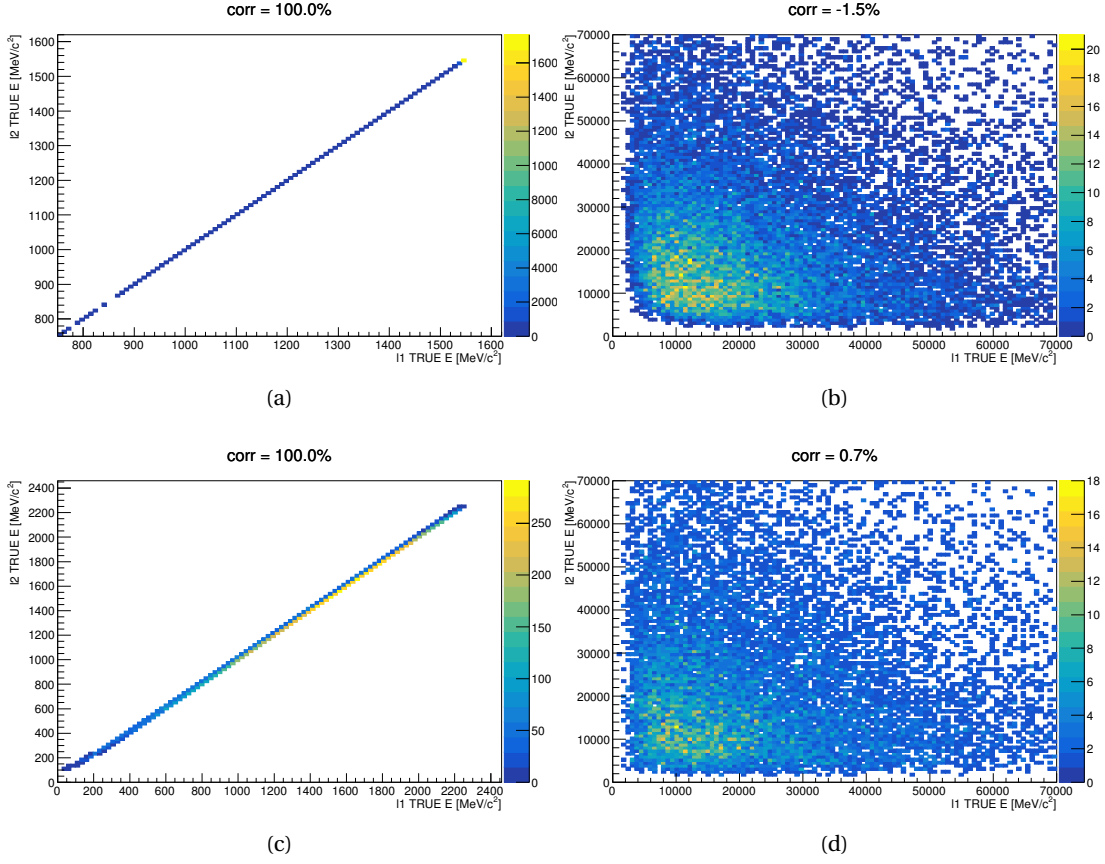


Figure A.1 – Histograms of the energies of the two leptons in  $\Lambda_b \rightarrow \Lambda^0 e \mu$  (figures A.1a and A.1b) and  $\Lambda_b \rightarrow \Lambda^0 J/\psi(ee)$  (figures A.1c and A.1d). In both cases, on the left the energies are expressed in the frame of the center of mass of the two leptons, while on the right they are in the laboratory rest frame. On top of each histogram, the Pearson correlation factor is shown.



## B Invariant mass fits to $B^0 \rightarrow K^+ \pi^-$

Figure B.1 shows the invariant-mass fit to  $B^0 \rightarrow K^+ \pi^-$  performed in the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis for the calibration of the BDT response.

Figure B.2 shows the fits to the same distributions, obtained using the alternative PDF described in Section 5.1.6 for evaluating a systematic error.

The red solid line shows the  $B^0$  signal, the green one shows the  $B_s^0$  component, the yellow dashed one the one from  $\Lambda_b^0 \rightarrow ph$  where the proton is misidentified as a kaon or pion. The combinatorial background is shown by the purple dashed line.

## Appendix B. Invariant mass fits to $B^0 \rightarrow K^+ \pi^-$

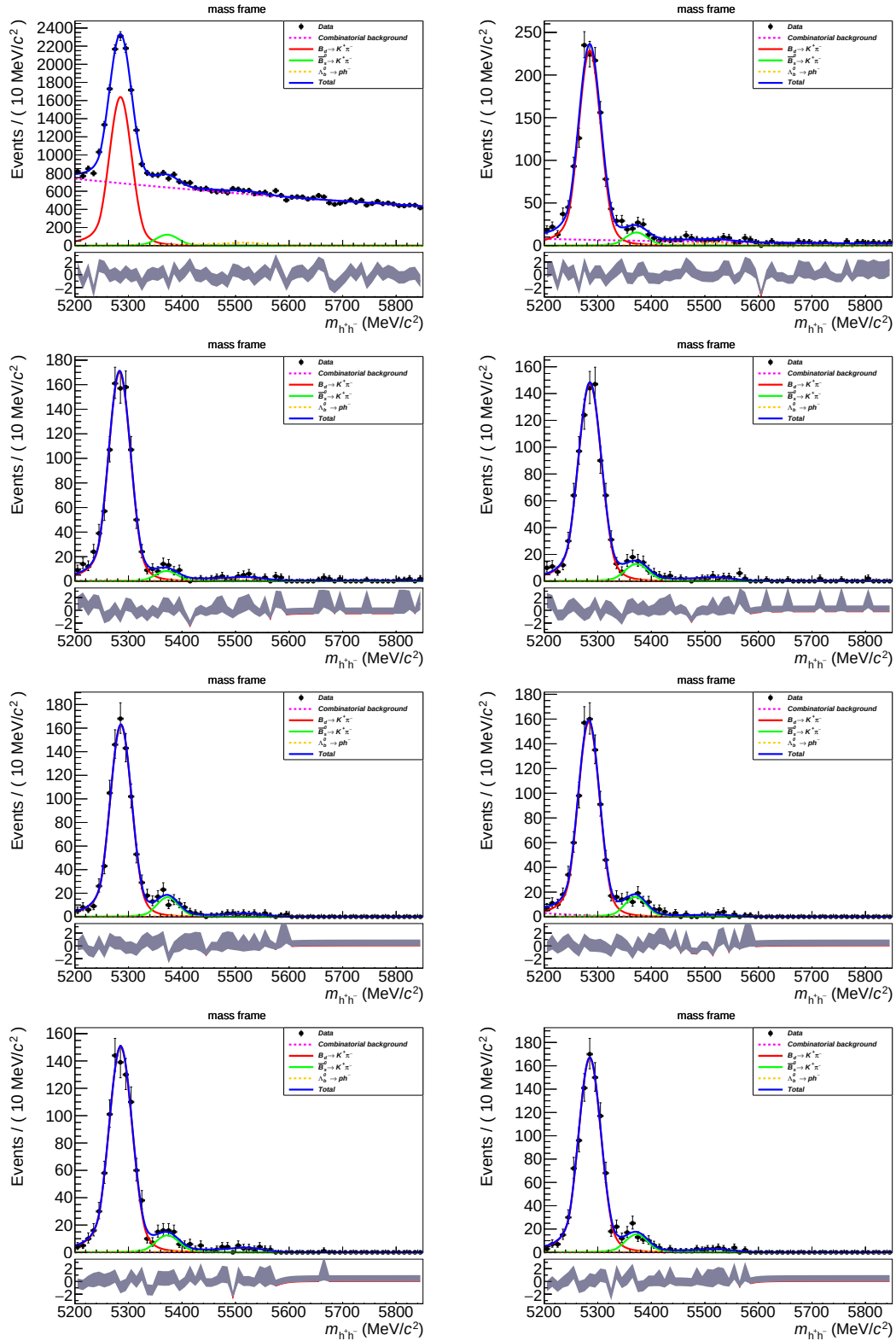


Figure B.1 – Invariant-mass distributions of  $B^0 \rightarrow K^+ \pi^-$  candidates in Run I data in different BDT bins with a PID requirement  $|\Delta LL_{K-\pi}| < 5$ .

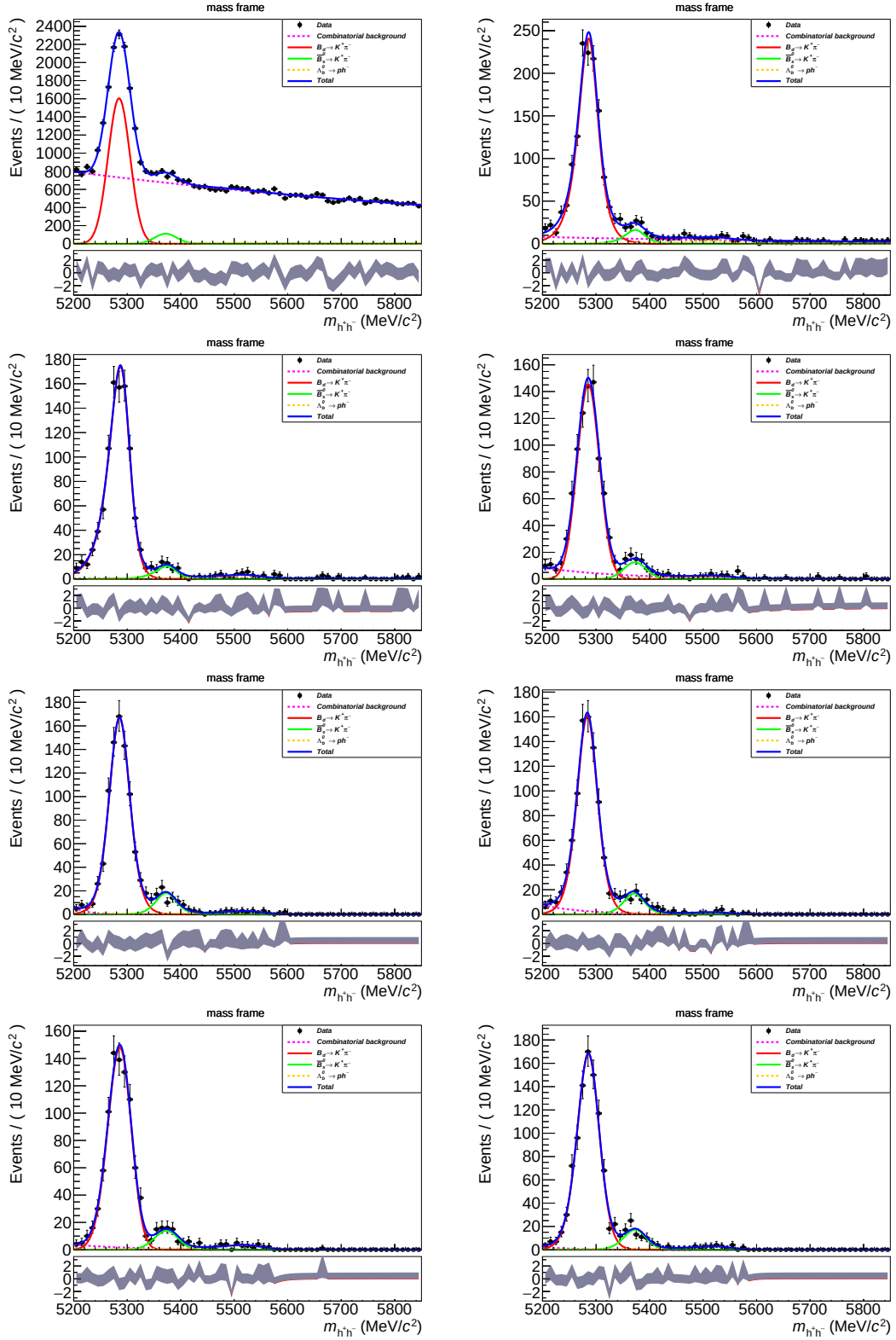


Figure B.2 – Invariant-mass distributions of  $B^0 \rightarrow K^+ \pi^-$  from Run I data in different BDT bins for  $|\Delta LL_{K-\pi}| < \kappa$  cut value  $k = 5$  with the alternative PDF used for the evaluation of the systematic error.



## C Training features of the BDT classifier for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

Figures C.1 and C.2 show the distributions of the twelve training features of the BDT classifier employed in the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis against combinatorial background. The distributions are shown for  $B_s^0 \rightarrow e^\pm \mu^\mp$  signal MC,  $B_s^0 \rightarrow e^\pm \mu^\mp$  data sidebands and  $B_s^0 \rightarrow e^\pm \mu^\pm$  data.

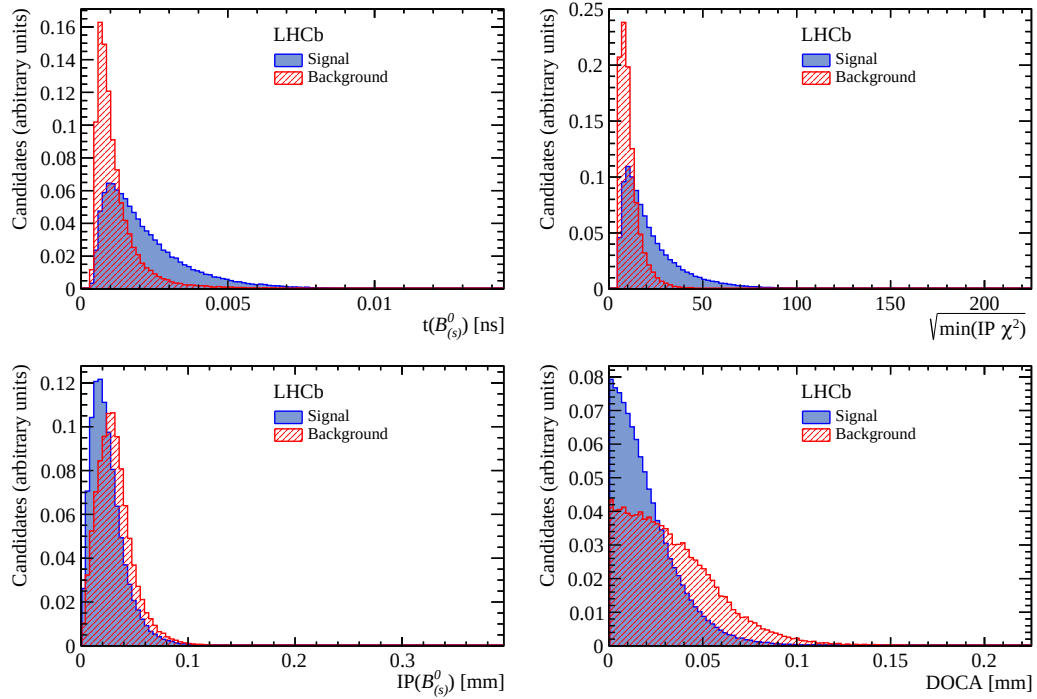


Figure C.1 – Distributions of 4 out of 12 variables used in the  $B_s^0 \rightarrow e^\pm \mu^\mp$  BDT training for simulated signal (blue) and background from opposite-sign data sidebands (red). From left to right and from top to bottom: the proper  $B_s^0$  lifetime,  $t(B_{(s)}^0)$ ; the square root of the minimum impact parameter  $\chi^2$  for the two tracks,  $\sqrt{\min(\text{IP} \chi^2)}$ ; the impact parameter of the  $B_s^0$ ,  $\text{IP}(B_{(s)}^0)$ ; the distance of closest approach between the two daughter tracks, DOCA.

### Appendix C. Training features of the BDT classifier for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

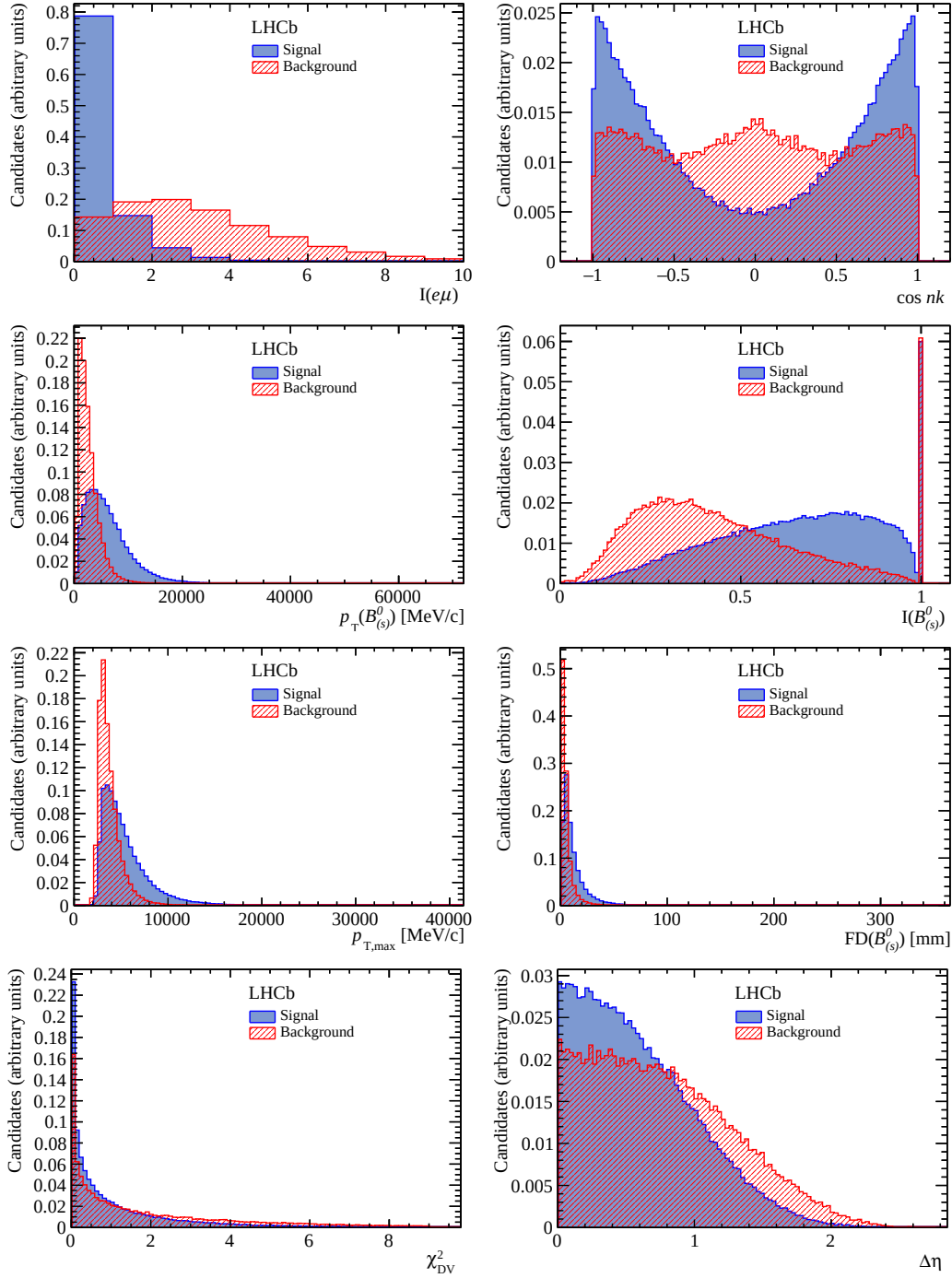


Figure C.2 – Distributions of 8 out of 12 variables used in the  $B_s^0 \rightarrow e^\pm \mu^\mp$  BDT training for simulated signal (blue) and background from opposite-sign data sidebands (red). From left to right and from top to bottom: the isolation of the two tracks  $I(e\mu)$ ; the cosine of the angle between the muon momentum in the B rest frame and the vector perpendicular to the  $B_s^0$  momentum and the beam axis,  $\cos nk$ ; the transverse momentum of the  $B_s^0$ ,  $p_T(B_{(s)}^0)$ ; the isolation of the  $B_s^0$ ,  $I(B_{(s)}^0)$ ; the maximum transverse momentum of the two daughter tracks,  $p_{T,\max}$ ; the flight distance of the  $B_s^0$  with respect to its primary vertex,  $FD(B_{(s)}^0)$ ; the  $\chi^2$  of the decay vertex of the  $B_s^0$ ,  $\chi_{DV}^2$ ; the difference of pseudo-rapidity between the two tracks,  $\Delta\eta$ .

## D Selection efficiencies for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ backgrounds.

This section contains the efficiency of the  $B_s^0 \rightarrow e^\pm \mu^\mp$  selection for decays that can be a background for the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  analysis. The efficiencies are shown in Table D.1, broken down into their PID, trigger and reconstruction and stripping component. The latter also includes the efficiencies due to the mass window: [4800,6000] MeV/ $c^2$ . For the  $B_{(s)}^0 \rightarrow h^+ h^-$  decays the efficiencies are assumed to be the same whether the mother particle is a  $B^0$  or a  $B_s^0$ . The stripping efficiencies here quoted include the DLE and IsMuon cuts whose efficiency is accounted for when calculating the PID efficiency value.

Table D.1 – Geometric, PID, trigger and reconstruction plus selection efficiencies. Only decays with MC events passing selection are shown.

| Decay                                      | Geom                  | PID ( $\times 10^{-4}$ ) | Reco + Sel ( $\times 10^{-4}$ ) | Trigger           |
|--------------------------------------------|-----------------------|--------------------------|---------------------------------|-------------------|
| $B^0 \rightarrow \pi \mu \nu$              | $0.0067 \pm 0.0001$   | $316.29 \pm 5.71$        | $37.48 \pm 0.25$                | $0.558 \pm 0.003$ |
| $B^0 \rightarrow \pi e \nu$                | $0.0067 \pm 0.0001$   | $29.00 \pm 5.51$         | $1.469 \pm 0.115$               | $0.370 \pm 0.038$ |
| $\Lambda_b^0 \rightarrow p \mu \nu$        | $0.0134 \pm 0.0001$   | $119.71 \pm 45.47$       | $34.33 \pm 0.41$                | $0.529 \pm 0.006$ |
| $B^+ \rightarrow J/\psi (\mu \mu) K$       | $0.1666 \pm 0.0005$   | $83.28 \pm 34.04$        | $0.05 \pm 0.01$                 | $0.286 \pm 0.099$ |
| $B_c^+ \rightarrow J/\psi (\mu \mu) e \nu$ | $0.00287 \pm 0.00001$ | $7565.98 \pm 1459.79$    | $1.54 \pm 0.17$                 | $0.333 \pm 0.052$ |
| $B_c^+ \rightarrow J/\psi (e e) \mu \nu$   | $0.00287 \pm 0.00001$ | $7844.04 \pm 1540.71$    | $1.42 \pm 0.16$                 | $0.333 \pm 0.053$ |
| $B^0 \rightarrow K \pi$                    | $0.1898 \pm 0.0005$   | $0.45 \pm 0.10$          | $5.11 \pm 0.31$                 | $0.434 \pm 0.030$ |
| $B^0 \rightarrow \pi \pi$                  | $0.1898 \pm 0.0005$   | $1.56 \pm 0.25$          | $5.11 \pm 0.31$                 | $0.434 \pm 0.030$ |
| $B^0 \rightarrow K K$                      | $0.1898 \pm 0.0005$   | $0.16 \pm 0.04$          | $5.11 \pm 0.31$                 | $0.434 \pm 0.030$ |
| $B^0 \rightarrow p \bar{p}$                | $0.1898 \pm 0.0005$   | $0.04 \pm 0.39$          | $5.11 \pm 0.31$                 | $0.434 \pm 0.030$ |





## **E Training features of the BDT classifier for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$**

Figures E.1 and E.2 show the distributions of the eleven training features of the BDT classifier employed in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis. The distributions are shown for  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  signal MC and the combinatorial background from the data sidebands. Figure E.3 shows the linear correlations between these variables on signal and background.

## Appendix E. Training features of the BDT classifier for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$

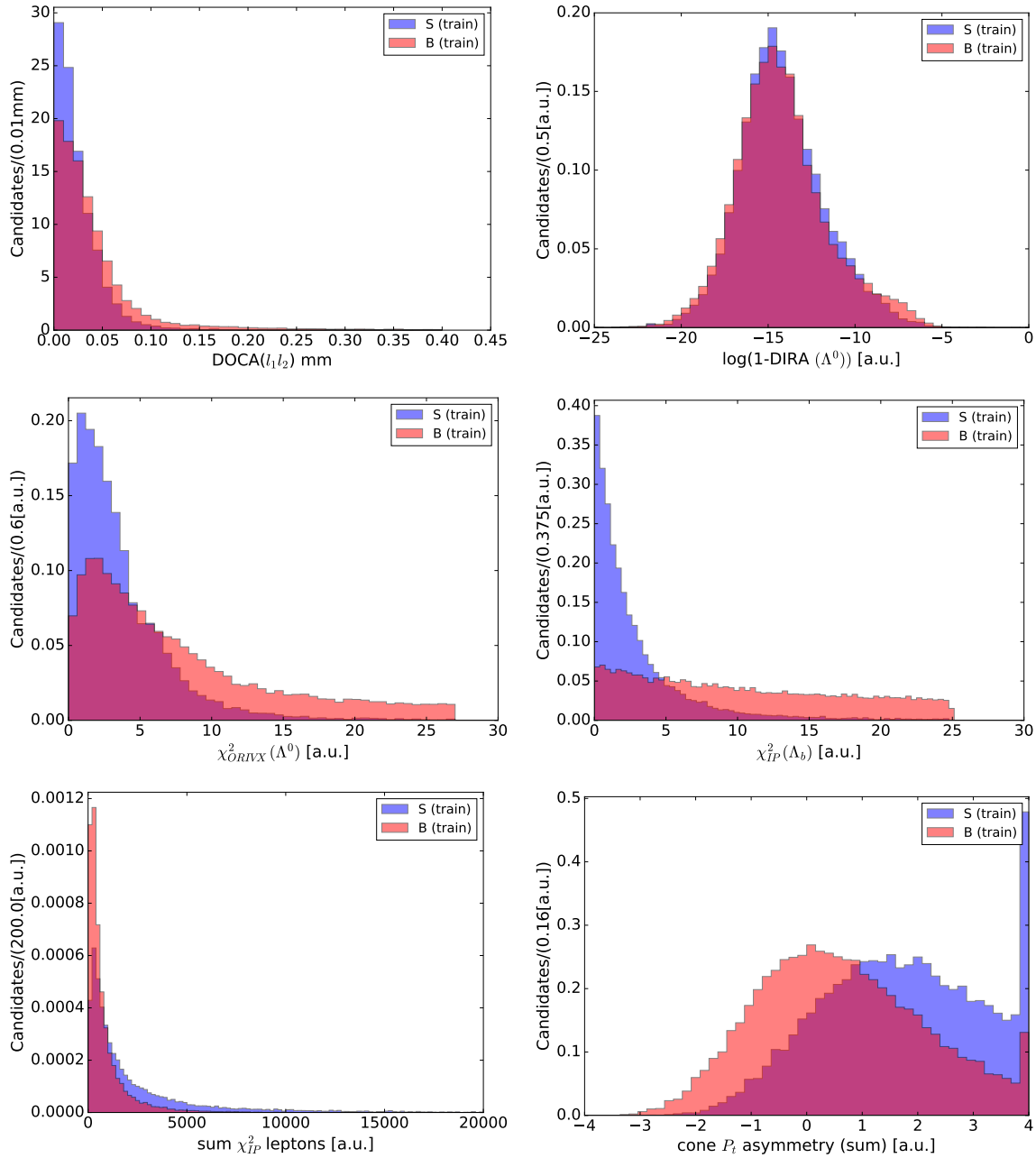


Figure E.1 – Distributions of 6 out of 11 variables used in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  BDT training for simulated signal events (blue) and background events from data sidebands (red). From left to right and from top to bottom: the distance of closest approach between the two lepton tracks, DOCA, the direction angle between the flight direction of the  $\Lambda^0$  and the direction of its momentum, DIRA( $\Lambda^0$ ), the  $\chi^2$  of the origin vertex of the  $\Lambda^0$ ,  $\chi^2_{\text{ORIX}}(\Lambda^0)$ , the impact parameter  $\chi^2$  of the  $\Lambda_b^0$ ,  $\chi^2_{\text{IP}}(\Lambda_b^0)$ , the sum of the impact parameter  $\chi^2$  of the two leptons,  $\chi^2_{\text{IP}}(\text{leptons})$ , the sum of the transverse momentum asymmetries of the four final-state tracks,  $\text{sum}(P_T \text{ asymmetry})$ .

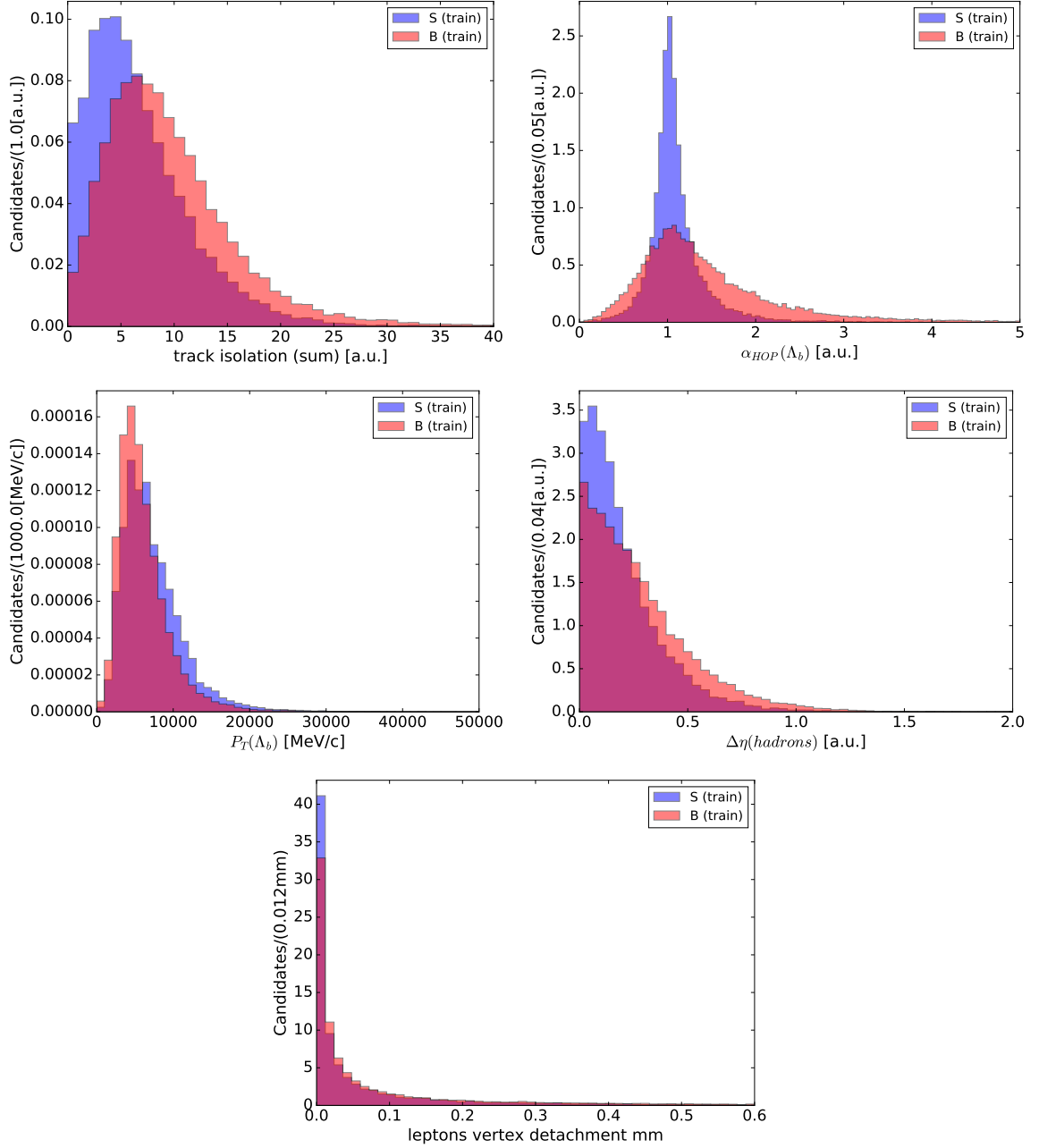


Figure E.2 – Distributions of 5 out of 11 variables used in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  BDT training for simulated signal events (blue) and background events from data sidebands (red). From left to right and from top to bottom: the sum of the cone isolations of the four final-state tracks,  $\text{sum}(\text{isolation})$ , the HOP factor,  $\alpha_{HOP}$ , the transverse momentum of the  $\Lambda_b^0$  factor,  $p_T(\Lambda_b^0)$ , the difference of pseudo-rapidity between the two hadron tracks,  $\Delta\eta(hadrons)$ , the  $\chi^2$  of the distance between the decay vertex of the  $\Lambda_b^0$  and the production vertex of the two leptons.

## Appendix E. Training features of the BDT classifier for $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$

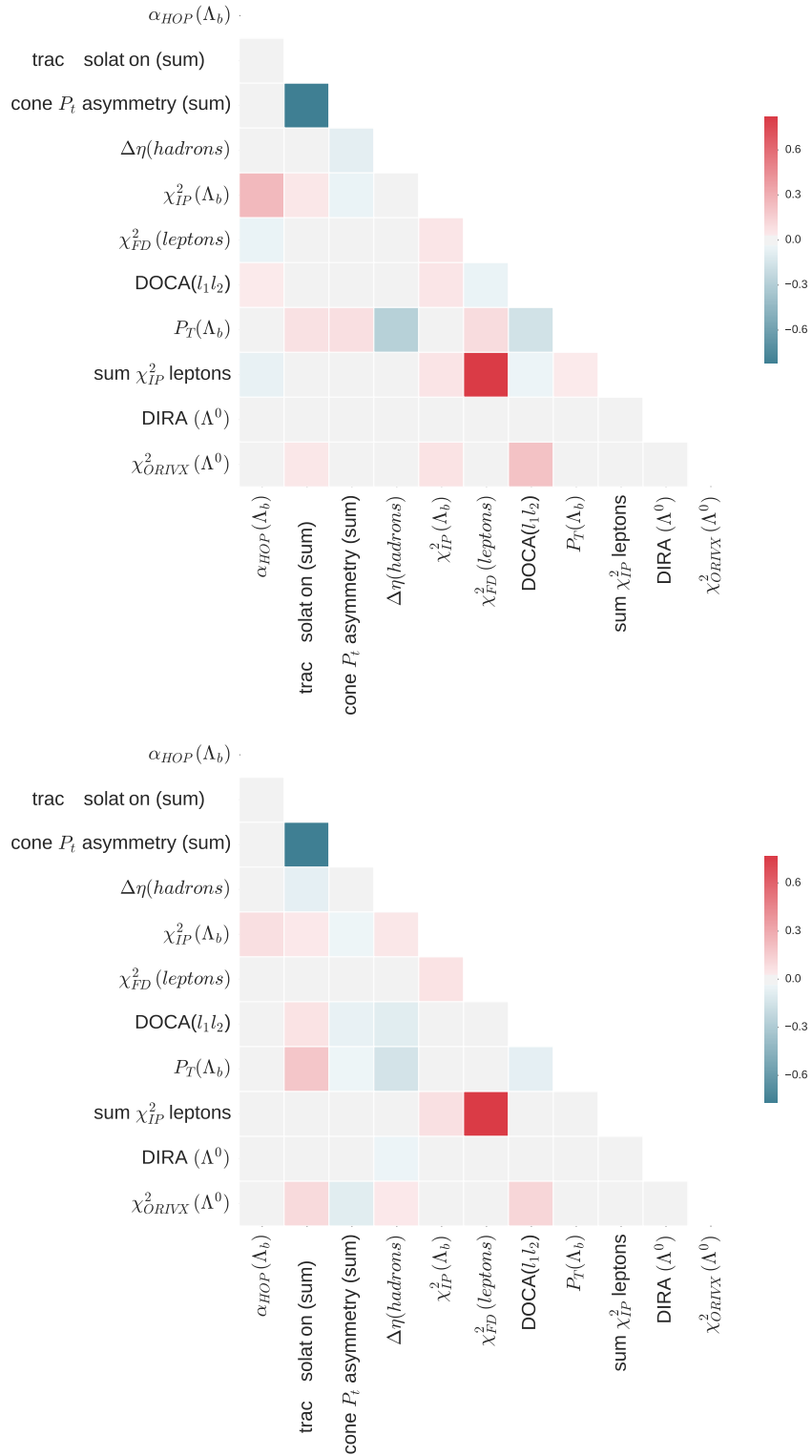


Figure E.3 – Linear correlation factors between the BDT training variables in the  $\Lambda_b^0 \rightarrow \Lambda^0 e^\pm \mu^\mp$  analysis in signal (left) and background (right).

## F HOP variables

Decays with electrons in the final state tend to have a poor mass resolution due to the emission of bremsstrahlung radiation, as discussed in multiple occasions in this thesis. Bremsstrahlung emission can nevertheless be a helpful characteristic of the signal, as it allows to distinguish it from partially-reconstructed backgrounds.

In a generic decay of a flying b-hadron (indicated here with  $B$ ) to a final state containing both electrons (grouped under the name  $X_e$ ) and other, different, particles ( $X_h$ ), one can build the quantity

$$\alpha_{\text{HOP}} = \frac{p_T(Y_h)}{p_T(X_e)}, \quad (\text{E.1})$$

where  $p_T(Y_h)$  and  $p_T(X_e)$  are the combined transverse momenta, with respect to the flight direction of the parent b-hadron, of the non-electronic and the electronic part of the decay, respectively. These quantities are illustrated in Figure F.1. Because of imperfect reconstruction,  $\alpha_{\text{HOP}}$  won't be always exactly 1 as expected from momentum conservation: in some cases too much energy is attributed to the electrons, resulting in a lower  $\alpha_{\text{HOP}}$ , and *vice versa*. For signal candidates, anyway,  $\alpha_{\text{HOP}}$  will be peaked at 1, with short tails on either side. This does not happen for partially-reconstructed backgrounds, where no kinematic constraint

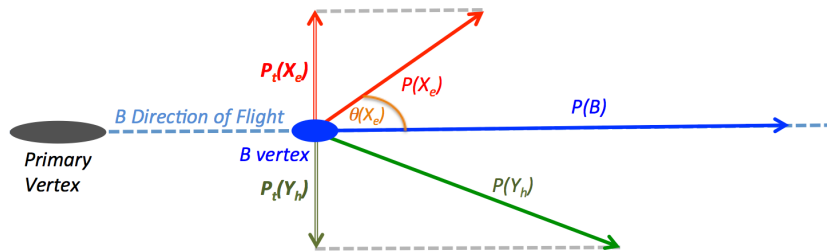


Figure F.1 – Representation of the kinematics of a  $B \rightarrow Y_h X_e$  decay, highlighting the quantities relevant for the definition of the variable HOP.

## Appendix F. HOP variables

---

forces  $p_T(Y_h) \simeq p_T(X_e)$ .

Furthermore,  $\alpha_{\text{HOP}}$  can be used to correct the three-momentum ( $p$ ) of the electrons, to counterbalance the mismeasurement:

$$p^{\text{corr}}(X_e) = \alpha_{\text{HOP}} \times p(X_e), \quad (\text{F.2})$$

forcing the ratio in Equation F.1 to be exactly 1, and allowing to build a *corrected mass* of the parent b-hadron from the four-momentum ( $P^{\text{corr}}(X_e)$ ), obtained using  $p^{\text{corr}}(X_e)$ :

$$(M_{\text{HOP}}^B)^2 = \|P(Y_h) + P^{\text{corr}}(X_e)\|^2. \quad (\text{F.3})$$

The resulting mass has a degraded resolution due to the implicit dependence on the resolution on the relative angle between the momentum of the electron system and the one of the other particles, but, as  $\alpha_{\text{HOP}}$ , it constitutes an excellent tool for background rejection.

The HOP variables were introduced in LHCb in 2016, as part of a study for the  $R_{K^*}$  analysis [23], and they have subsequently been implemented in the LHCb software as a complementary activity in the context of this thesis.

# Bibliography

- [1] M. Thomson, *Modern particle physics*, Cambridge University Press, New York, 2013.
- [2] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.
- [3] S. L. Glashow, *The renormalizability of vector meson interactions*, Nucl. Phys. **10** (1959) 107.
- [4] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [5] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, Il Nuovo Cimento (1955-1965) **11** (1959) 568.
- [6] V. Barger and R. Phillips, *Collider Physics*, Addison-Wesley, Boston, 1997.
- [7] J. Greensite, *An introduction to the confinement problem*, Lect. Notes Phys. **821** (2011) 1.
- [8] A. Ali, J. S. Lange, and S. Stone, *Exotics: Heavy Pentaquarks and Tetraquarks*, Prog. Part. Nucl. Phys. **97** (2017) 123, arXiv:1706.00610.
- [9] G. Bernardi, M. Carena, and T. Junk, *Higgs bosons: theory and searches*, Particle Data Group (2007).
- [10] C. Rovelli, *Notes for a brief history of quantum gravity*, in *Recent developments in theoretical and experimental general relativity, gravitation and relativistic field theories. Proceedings, 9th Marcel Grossmann Meeting, MG'9, Rome, Italy, July 2-8, 2000. Pts. A-C*, pp. 742–768, 2000. arXiv:gr-qc/0006061.
- [11] M. Conversi, E. Pancini, and O. Piccioni, *On the disintegration of negative mesons*, Phys. Rev. **71** (1947) 209.
- [12] S. H. Neddermeyer and C. D. Anderson, *Note on the nature of cosmic-ray particles*, Phys. Rev. **51** (1937) 884.
- [13] H. Yukawa, *On the Interaction of Elementary Particles I*, Proc. Phys. Math. Soc. Jap. **17** (1935) 48, [Prog. Theor. Phys. Suppl.1,1(1935)].
- [14] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.
- [15] CDF, F. Abe *et al.*, *Observation of top quark production in  $\bar{p}p$  collisions*, Phys. Rev. Lett. **74** (1995) 2626, arXiv:hep-ex/9503002.

## Bibliography

---

- [16] D0, S. Abachi *et al.*, *Search for high mass top quark production in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  TeV*, Phys. Rev. Lett. **74** (1995) 2422, arXiv:hep-ex/9411001.
- [17] DONUT, K. Kodama *et al.*, *Observation of tau neutrino interactions*, Phys. Lett. **B504** (2001) 218, arXiv:hep-ex/0012035.
- [18] Super-Kamiokande, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.
- [19] SNO, Q. R. Ahmad *et al.*, *Measurement of the rate of  $\nu_e + d \rightarrow p + p + e^-$  interactions produced by  $^8\text{B}$  solar neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301, arXiv:nucl-ex/0106015.
- [20] SNO, Q. R. Ahmad *et al.*, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv:nucl-ex/0204008.
- [21] SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257, arXiv:hep-ex/0509008.
- [22] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using  $B^+ \rightarrow K^+ \ell^+ \ell^-$  decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.
- [23] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with  $B^0 \rightarrow K^{*0} \ell^+ \ell^-$  decays*, JHEP **08** (2017) 055, arXiv:1705.05802.
- [24] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays  $B \rightarrow K^{(*)} l^+ l^-$* , Phys. Rev. **D86** (2012) 032012, arXiv:1204.3933.
- [25] Belle, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for  $B \rightarrow K^{(*)} \ell^+ \ell^-$* , Phys. Rev. Lett. **103** (2009) 171801, arXiv:0904.0770.
- [26] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions  $\mathcal{B}(B_c^+ \rightarrow J/\psi K^+)/\mathcal{B}(B_c^+ \rightarrow J/\psi \pi^+)$* , JHEP **09** (2016) 153, arXiv:1607.06823.
- [27] R. Bernstein, *Charged Lepton Flavour Violation: An Overview*, Fermilab, Presented at the XIIth Rencontres du Vietnam, ICISE, Quy Nhon, Vietnam, 07, 2016.
- [28] A. Crivellin *et al.*, *Lepton-flavour violating B decays in generic  $Z'$  models*, Phys. Rev. **D92** (2015) 054013, arXiv:1504.07928.
- [29] D. Bečirević, S. Fajfer, N. Košnik, and O. Sumensari, *Leptoquark model to explain the B-physics anomalies,  $R_K$  and  $R_D$* , Phys. Rev. **D94** (2016) 115021, arXiv:1608.08501.
- [30] I. De Medeiros Varzielas and G. Hiller, *Clues for flavor from rare lepton and quark decays*, JHEP **06** (2015) 072, arXiv:1503.01084.
- [31] A. Ilakovac, *Lepton flavor violation in the standard model extended by heavy singlet Dirac neutrinos*, Phys. Rev. **D62** (2000) 036010, arXiv:hep-ph/9910213.



- 
- [32] R. A. Diaz, R. Martinez, and C. E. Sandoval, *Improving bounds on flavor changing vertices in the two Higgs doublet model from  $B^0-\bar{B}^0$  mixing*, Eur. Phys. J. **C46** (2006) 403, arXiv:hep-ph/0509194.
  - [33] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*, Phys. Rev. **D10** (1974) 275, Erratum-ibid. **D11** (1975) 703.
  - [34] D. Guadagnoli, *Flavor anomalies on the eve of the Run-2 verdict*, Mod. Phys. Lett. **A32** (2017) 1730006, arXiv:1703.02804.
  - [35] J. Adam *et al.*, *New Constraint on the Existence of the  $\mu^+ \rightarrow e^+ \gamma$  Decay*, Phys. Rev. Lett. **110** (2013).
  - [36] SINDRUM, U. Bellgardt *et al.*, *Search for the decay  $\mu^+ \rightarrow e^+ e^+ e^-$* , Nucl. Phys. **B299** (1988) 1.
  - [37] W. Bertl *et al.*, *A search for  $\mu$ -e conversion in muonic gold*, The European Physical Journal C (2006) .
  - [38] A. M. Baldini *et al.*, *MEG Upgrade Proposal*, arXiv:1301.7225v2 (2013).
  - [39] A. Blondel *et al.*, *Research Proposal for an Experiment to Search for the Decay  $\mu \rightarrow eee$* , arXiv:1301.6113v1 (2013) arXiv:1301.6113.
  - [40] D. Brown, *Mu2e: a Muon to Electron Conversion Experiment at Fermilab*, Nuclear Physics B - Proceedings Supplements (2014) .
  - [41] DELPHI, P. Abreu *et al.*, *A Search for lepton flavor violation in  $Z^0$  decays*, Phys. Lett. **B298** (1993) 247.
  - [42] R. Akers *et al.*, *A search for Lepton Flavour Violating  $Z^0$  Decays*, Zeitschrift für Physik C **67** (1995) 555.
  - [43] OPAL, G. Abbiendi *et al.*, *Search for lepton flavor violation in  $e^+e^-$  collisions at  $s^{**}(1/2) = 189\text{-GeV} - 209\text{-GeV}$* , Phys. Lett. **B519** (2001) 23, arXiv:hep-ex/0109011.
  - [44] LHCb collaboration, R. Aaij *et al.*, *Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb*, Phys. Lett. **B724** (2013) 36, arXiv:1304.4518.
  - [45] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decay  $D^0 \rightarrow e^\pm \mu^\mp$* , Phys. Lett. **B754** (2016) 167, arXiv:1512.00322.
  - [46] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays  $B_s^0 \rightarrow e^\pm \mu^\mp$  and  $B^0 \rightarrow e^\pm \mu^\mp$* , Phys. Rev. Lett. **111** (2013) 141801, arXiv:1307.4889.
  - [47] LHCb collaboration, R. Aaij *et al.*, *Search for lepton number violating decays  $B^+ \rightarrow \pi^- \mu^+ \mu^+$  and  $B^+ \rightarrow K^- \mu^+ \mu^+$* , Phys. Rev. Lett. **108** (2012) 101601, arXiv:1110.0730.
  - [48] LHCb collaboration, R. Aaij *et al.*, *Search for  $D_{(s)}^+ \rightarrow \pi^+ \mu^+ \mu^-$  and  $D_{(s)}^+ \rightarrow \pi^- \mu^+ \mu^+$  decays*, Phys. Lett. **B724** (2013) 203, arXiv:1304.6365.
  - [49] B. Wang, *The Belle II Experiment and SuperKEKB Upgrade*, arXiv:1511.09434.

## Bibliography

---

- [50] A. S. Kronfeld *et al.*, *Project X: Physics Opportunities*, arXiv:1306.5009.
- [51] L. Calibbi and G. Signorelli, *Charged Lepton Flavour Violation: An Experimental and Theoretical Introduction*, Riv. Nuovo Cim. **41** (2018), no. 2 1, arXiv:1709.00294.
- [52] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.
- [53] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.
- [54] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- [55] S. Myers, *The LEP Collider, from design to approval and commissioning*, John Adams' Lecture, CERN, Geneva, 1991. Delivered at CERN, 26 Nov 1990.
- [56] ATLAS Collaboration, *The atlas experiment at the cern large hadron collider*, Journal of Instrumentation **3** (2008), no. 08 S08003.
- [57] CMS Collaboration, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [58] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [59] LHCf collaboration, O. Adriani *et al.*, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.
- [60] TOTEM collaboration, G. Anelli *et al.*, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.
- [61] MoEDAL Collaboration, J. L. Pinfold, *The MoEDAL Experiment at the LHC – a New Light on the Terascale Frontier*, J. Phys. Conf. Ser. **631** (2015), no. 1 012014.
- [62] C. Lefèvre, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, Dec, 2008.
- [63] LHCb Collaboration, *Large hadron collider beauty experiment public results website*, <http://lhcb-public.web.cern.ch/lhcb-public/>.
- [64] LHCb collaboration, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [65] LHCb Collaboration,  *$b\bar{b}$  production angle plots*, [https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb\\_ProductionAngles.html](https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html).
- [66] CERN, *CERN Workshop on Standard Model Physics (and more) at the LHC.*, (Geneva), CERN, 2000.

- 
- [67] B. Muratori and T. Pieloni, *Luminosity levelling techniques for the LHC*, in *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, pp. 177–181, 2014. arXiv:1410.5646.[,177(2014)], doi: 10.5170/CERN-2014-004.177.
- [68] LHCb collaboration, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015), no. 07 1530022, arXiv:1412.6352.
- [69] LHCb collaboration, *LHCb VELO (Vertex LOCator): Technical Design Report*, CERN-LHCC-2001-011. LHCb-TDR-005.
- [70] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007. LHCb-TDR-001.
- [71] LHCb Collaboration, *LHCb Silicon Tracker - Material for Publications*, <http://www.physik.unizh.ch/groups/lhcb/public/material/>.
- [72] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. LHCb-TDR-009.
- [73] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029. LHCb-TDR-008.
- [74] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.
- [75] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024. LHCb-TDR-006.
- [76] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037. LHCb-TDR-003.
- [77] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.
- [78] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036. LHCb-TDR-002.
- [79] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010. LHCb-TDR-004.
- [80] C. Lippmann, *Particle identification*, Nucl. Instrum. Meth. **A666** (2012) 148, arXiv:1101.3276.
- [81] A. Powell, *Particle Identification at LHCb*, No. LHCb-PROC-2011-008. CERN-LHCb-PROC-2011-008, p. 5, Jan, 2011.
- [82] LHCb Collaboration, *LHCb Trigger Schemes*, <https://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html>.
- [83] G. Dujany and B. Storaci, *Real-time alignment and calibration of the LHCb Detector in Run II*, J. Phys. : Conf. Ser. **664** (2015) 082010. 8 p.

## Bibliography

---

- [84] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, May, 2014.
- [85] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016. LHCb-TDR-016.
- [86] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- [87] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.
- [88] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [89] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.
- [90] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.
- [91] GEANT4, S. Agostinelli *et al.*, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [92] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.
- [93] S. Benson, V. Gligorov, M. A. Vesterinen, and J. M. Williams, *The lhcb turbo stream*, Journal of Physics: Conference Series **664** (2015), no. 8 082004.
- [94] *Lhc commissioning website*, <https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm>.
- [95] LHCb Collaboration, *LHCb Tracker Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-001. LHCb-TDR-015, Feb, 2014.
- [96] LHCb Collaboration, *LHCb PID Upgrade Technical Design Report*, tech. rep., 2013.
- [97] LHCb Collaboration, *LHCb Trigger and Online Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-016. LHCb-TDR-016, May, 2014.
- [98] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001. LHCb-TDR-015.
- [99] A. K. Kuonen *et al.*, *LHCb Scintillating Fibre Tracker Engineering Design Review: Silicon Photomultipliers*, Tech. Rep. LHCb-INT-2016-019. CERN-LHCb-INT-2016-019, CERN, Geneva, Apr, 2016.
- [100] C. Joram *et al.*, *LHCb Scintillating Fibre Tracker Engineering Design Review Report: Fibres, Mats and Modules*, Tech. Rep. LHCb-PUB-2015-008. CERN-LHCb-PUB-2015-008, CERN, Geneva, Mar, 2015.

- 
- [101] A. B. Rodrigues Cavalcante *et al.*, *Shrinking of bumps by drawing scintillating fibres through a hot conical tool*, Tech. Rep. LHCb-PUB-2016-010. CERN-LHCb-PUB-2016-010, CERN, Geneva, Apr, 2016.
  - [102] M. Demmer *et al.*, *Simulation of Light Yield Attenuation Maps for the LHCb SciFi Tracker Upgrade*, Tech. Rep. LHCb-INT-2016-015. CERN-LHCb-INT-2016-015, CERN, Geneva, Mar, 2016.
  - [103] LPHE- Laboratoire de Physique des Hautes Énergies, *The SciFi Mats*, <https://lphe-web.epfl.ch/en/lhcb/les-mats-scifi/>.
  - [104] G. Andreassi, *Tests of Lepton Flavour Universality with  $b \rightarrow s \ell \ell$  transitions at LHCb*, No. LHCb-PROC-2017-024, 2017. arXiv:1708.02515.
  - [105] LHCb, R. Aaij *et al.*, *Search for the lepton-flavour violating decays  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$* , JHEP **03** (2018) 078, arXiv:1710.04111.
  - [106] LHCb collaboration, R. Aaij *et al.*, *Measurement of the  $B_s^0 \rightarrow \mu^+ \mu^-$  branching fraction and effective lifetime and search for  $B^0 \rightarrow \mu^+ \mu^-$  decays*, Phys. Rev. Lett. **118** (2017) 191801, arXiv:1703.05747.
  - [107] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.
  - [108] G. Mancinelli and J. Serrano, *Study of Muon Isolation in the  $B_s^0 \rightarrow \mu^+ \mu^-$  Channel*, Tech. Rep. LHCb-INT-2010-011. CERN-LHCb-INT-2010-011, CERN, Geneva, Mar, 2010.
  - [109] CDF Collaboration, A. Abulencia and others. *Search for  $B_s^0 \rightarrow \mu^+ \mu^-$  and  $B_d^0 \rightarrow \mu^+ \mu^-$  decays in  $p\bar{p}$  collisions with  $cdfii$* , Phys. Rev. Lett. **95** (2005) 221805.
  - [110] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.
  - [111] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
  - [112] N. L. Johnson, *Systems of frequency curves derived from the first law of laplace.*, Trabajos de Estadística **5** (1954) 283.
  - [113] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116, [186(2003)].
  - [114] H. Albrecht *et al.*, *Search for hadronic  $b \rightarrow u$  decays*, Physics Letters B **241** (1990), no. 2 278.
  - [115] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.
  - [116] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio  $f_s/f_d$  and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286,  $f_s/f_d$  value updated in LHCb-CONF-2013-011.

## Bibliography

---

- [117] A. L. Read, *Presentation of search results: the  $cl s$  technique*, Journal of Physics G: Nuclear and Particle Physics **28** (2002), no. 10 2693.
- [118] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, *The RooStats project*, arXiv:1009.1003.
- [119] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of  $b$ -hadron,  $c$ -hadron, and  $\tau$ -lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017) 895, arXiv:1612.07233, updated results and plots available at <https://hflav.web.cern.ch>.
- [120] S. Sahoo and R. Mohanta, *Effects of scalar leptoquark on semileptonic  $\Lambda_b$  decays*, New J. Phys. **18** (2016), no. 9 093051, arXiv:1607.04449.
- [121] LHCb collaboration, R. Aaij *et al.*, *Measurement of  $b$  hadron production fractions in 7 TeV  $pp$  collisions*, Phys. Rev. **D85** (2012) 032008, arXiv:1111.2357.
- [122] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- [123] T. Chen and C. Guestrin, *Xgboost: A scalable tree boosting system*, in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, (New York, NY, USA), pp. 785–794, ACM, 2016. doi: 10.1145/2939672.2939785.
- [124] F. Pedregosa *et al.*, *Scikit-learn: Machine learning in Python*, Journal of Machine Learning Research **12** (2011) 2825.
- [125] M. W. Browne, *Cross-Validation Method*, Journal of Mathematical Psychology **44** (2000), no. 1 108.
- [126] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.

# Guido Andreassi

Birth date: 2<sup>nd</sup> of January 1990

Citizenship: Italian

Address: Avenue de Morges 72b • 1004 Lausanne, Switzerland

E-Mail: guido.andreassi@epfl.ch

## Education

---

|                                                                                                                                       |                  |
|---------------------------------------------------------------------------------------------------------------------------------------|------------------|
| <b>Doctor of Philosophy (PhD) in experimental particle Physics at EPFL</b><br>(École Polytechnique Fédérale de Lausanne), Switzerland | <b>2015-2019</b> |
| Thesis defended in January 2019.                                                                                                      |                  |
| <b>Master's degree in Physics at Università La Sapienza, Rome, Italy</b>                                                              | <b>2012-2014</b> |
| Grade 110/110 cum laude                                                                                                               |                  |
| <b>Bachelor's degree in Physics at Università La Sapienza, Rome, Italy</b>                                                            | <b>2008-2012</b> |

## Scientific experience and activities

---

|                                                                                                                             |                         |
|-----------------------------------------------------------------------------------------------------------------------------|-------------------------|
| <b>Convener of the Very Rare Decays sub-working group in LHCb</b>                                                           | <b>Jan 2019-present</b> |
| <b>Scientific collaborator at CERN (Conseil Européen pour la Recherche Nucléaire)</b>                                       | <b>2014-present</b>     |
| Member of the LHCb (Large Hadron Collider beauty) collaboration                                                             |                         |
| <b>Main author of multiple analyses in LHCb</b>                                                                             | <b>2014-present</b>     |
| Searching new physics phenomena and rare particle decays.<br>Participated as a speaker to several international conferences |                         |
| <b>Teaching assistant</b>                                                                                                   | <b>2015-2018</b>        |
| For master and bachelor courses at EPFL                                                                                     |                         |
| <b>R&amp;D quality assurance activities</b>                                                                                 | <b>2016-2017</b>        |
| For the construction of a scintillating fibre tracker (SciFi) for the LHCb detector                                         |                         |
| <b>Working group liaison in LHCb</b>                                                                                        | <b>2015-2017</b>        |
| For the Rare Decays working group, for two years.                                                                           |                         |
| <b>Internship at PSI - Paul Scherrer Institute</b>                                                                          | <b>Jun-Aug 2013</b>     |
| Worked on the measurement of the positron reconstruction efficiency in the MEG experiment                                   |                         |

## Computing skills

---

### Programming

Excellent in C++, Python, bash; Familiar with SQL, HTML

### Data Analysis

Scikit-learn, XGBoost, TMVA, ROOT, Pandas, NumPy, Matplotlib, Pyplot

### Other

LaTeX, Docker, Jupyter, Git, Subversion, Keynote, Ipython, Anaconda, Microsoft Office, OpenOffice

### Operating systems

Unix, macOS, Microsoft Windows

## Languages

---

**Italian:** Native; **English:** Fluent; **French:** Proficient