

Search for the lepton flavour violating decays $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ with LHCb Run 2 data

Présentée le 16 février 2024

Faculté des sciences de base
Laboratoire de physique des hautes énergies OS
Programme doctoral en physique

pour l'obtention du grade de Docteur ès Sciences

par

Sebastian SCHULTE

Acceptée sur proposition du jury

Prof. L. Villard, président du jury
Prof. O. Schneider, directeur de thèse
Prof. J. A. de Vries, rapporteur
Dr H. Cliff, rapporteur
Dr C. G. Lange, rapporteur



To my uncle Manfred, in loving memories ...

It is not the mountain we conquer but ourselves.

—Edmund Hillary

Acknowledgements

Looking back on my journey at LPHE, it has truly been one of the most remarkable experiences of my life. Throughout this time, I've had the privilege of crossing paths with amazing people who have significantly influenced and shaped my personal and professional growth. I would like to express my heartfelt gratitude and acknowledgement to the following people.

Firstly, I thank my supervisor Prof. Olivier Schneider for providing me with the opportunity to do a PhD at this lab. Especially for being a constant support during both the highs and lows, providing valuable input for my analysis and his efforts in maintaining a positive atmosphere at the lab.

I also would like to thank Prof. Frédéric Blanc, whose expert knowledge was a valuable contribution to my work, as well as his unwavering support throughout my journey.

A big thank you belongs to Dr. Renato Quagliani, with whom I worked most of the time. In my view, Renato is an exceptional physicist whose insightful input into my work has significantly enhanced my understanding of high-energy physics and professional growth.

I would also like to thank Prof. Lesya Shchutska, Dr. Guido Haefeli, Esther Hofmann and Corinne Craman for their consistent support and encouragement.

Big thanks go to my beloved colleagues who have become good friends over time. Honourable mentions belong to my office mates Anna, Violaine, Pavol, Veronica, Raphaël, Lino, Guillaume and Matthieu.

Throughout my doctoral journey, I lived in an apartment that I shared with fantastic people from various corners of the world. My flatmates—Vladimir, Michele, Gabriel, Clemens, and Afonso—consistently supported me, and together, we evolved into a close-knit family. I am deeply grateful to them and want to express my heartfelt thanks.

Finally, I want to extend my heartfelt appreciation to my loved ones for their steadfast support and continuous presence during this journey. I owe a debt of gratitude to my partner Kirsten, my brother Andreas and my parents for their unwavering commitment and unconditional love, which made this journey possible.

Lausanne, December 29, 2023

S. Schulte

Abstract

A large variety of new physics models suggest that the rates for lepton flavour violating b -hadron decays may be much higher than predicted in the Standard Model, which leads to a high interest in the search for such decays.

This thesis presents the search for the lepton-flavour violating decays $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ with the data sample collected by the LHCb experiment between 2016 and 2018, which corresponds to an integrated luminosity of 5.4 fb^{-1} of proton-proton collisions. The analysis is performed while blinding the signal mass region, to avoid any bias introduced by the experimenter. The $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ branching fractions are measured with respect to the high statistics $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ decay, with a fit to the invariant $e^\pm \mu^\mp$ mass distribution in six independent data samples characterized by different background levels and signal mass resolutions due to electron bremsstrahlung. A dedicated selection is used to reduce background from random combinations and specific physics decays with a multivariate analysis and particle identification requirements, respectively. Efficiencies are determined from simulation, corrected for mismodeling effects with data-driven methods, and validated by measuring $r_{J/\psi} = \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-) / \mathcal{B}(J/\psi \rightarrow e^+ e^-)$. The main physics backgrounds are identified to arise from B^0 decays to the $K^+ \pi^-$, $\pi^+ \pi^-$, $\pi^+ e^- \bar{\nu}_e$ and $\pi^+ \mu^- \bar{\nu}_\mu$ final states, and taken into account in the mass fit; the yields of the two components peaking in the signal mass window, $B^0 \rightarrow K^+ \pi^-$ (12.2 events) and $B^0 \rightarrow \pi^+ \pi^-$ (7.7 events), are estimated from simulation and validated with two independent data-driven methods. Pseudo-experiments are performed to test the fit stability, study potential biases, and estimate the sensitivity of the analysis.

In absence of signals and without any systematic uncertainty considered, the average expected upper limits at 95%(90%) confidence level are estimated to be $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 7.3(5.9) \cdot 10^{-10}$, and $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 2.3(1.8) \cdot 10^{-9}$ or $< 1.8(1.5) \cdot 10^{-9}$ assuming the B_s^0 decay amplitude is completely dominated by the light (short-lived) or heavy (long-lived) mass eigenstate of the $B_s^0 - \bar{B}_s^0$ system, respectively. This is an improvement in sensitivity by a factor 1.6 for $B^0 \rightarrow e^\pm \mu^\mp$ and 2.8 for $B_s^0 \rightarrow e^\pm \mu^\mp$ (when dominated by the heavy eigenstate), with respect to the previous best expected limits published by LHCb with the 2011 and 2012 data.

In addition, this thesis presents the quality assurance process of silicon photomultiplier detectors used for the Scintillating Fibre (SciFi) tracker of the recent LHCb detector upgrade.

Keywords: particle physics, LHCb, LHC, rare decays, flavour, lepton flavour violation, SciFi tracker, SiPMs

Abstrakt

Eine Vielzahl neuer physikalischer Modelle suggeriert deutlich höhere Zerfallsraten für Lepton-Flavour-verletzende Zerfälle von b -Hadronen als vom Standardmodell vorhergesagt, was zu einem hohen Interesse an solchen Zerfällen führt.

Die vorliegende Dissertation behandelt die Suche nach den Lepton-Flavour-verletzenden Zerfällen $B^0 \rightarrow e^\pm \mu^\mp$ und $B_s^0 \rightarrow e^\pm \mu^\mp$, unter der Verwendung von Daten, die zwischen 2016 und 2018 am LHCb-Experiment aus Proton-Proton-Kollisionen aufgezeichnet wurden und einer integrierten Luminosität von 5.4 fb^{-1} entsprechen. Der signal-sensitive Bereich wird während der Analyse ausgeblendet, um jegliche Verzerrung durch den Experimentator zu vermeiden.

Die Verzweungsverhältnisse $B^0 \rightarrow e^\pm \mu^\mp$ und $B_s^0 \rightarrow e^\pm \mu^\mp$ werden im Verhältnis zu dem Zerfall $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ mit einem Fit der $e^\pm \mu^\mp$ Massenverteilung, in sechs unabhängigen Datensätzen mit unterschiedlichen Untergrundniveaus und Signalsensitivitäten aufgrund der Elektronenbremsstrahlung, gemessen. Zur Reduzierung des kombinatorischen Untergrunds sowie der Beiträge von spezifischen physikalischen Zerfällen, werden eine multivariate Analyse und Identifikationskriterien verwendet. Effizienzen werden anhand von Simulationen ermittelt und datengestützte Methoden werden zur Korrektur von Modellierungsfehlern angewandt und mit der Messung von $r_{J/\psi} = \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-) / \mathcal{B}(J/\psi \rightarrow e^+ e^-)$ validiert. Die wesentlichen Untergründe resultieren aus Zerfällen von B^0 -Mesonen in $K^+ \pi^-$, $\pi^+ \pi^-$, $\pi^+ e^- \bar{\nu}_e$ und $\pi^+ \mu^- \bar{\nu}_\mu$ und werden in die Analyse einbezogen. Der Beitrag der beiden Komponenten, die im Signalmassenfenster auftreten, $B^0 \rightarrow K^+ \pi^-$ (12.2 Ereignisse) und $B^0 \rightarrow \pi^+ \pi^-$ (7.7 Ereignisse), werden mithilfe von Simulationen geschätzt und mit zwei unabhängigen Methoden validiert. Pseudo-Experimente werden durchgeführt, um die Stabilität des Fits und mögliche Verzerrungen zu untersuchen.

In Abwesenheit eines Signals und ohne Berücksichtigung der systematischen Unsicherheiten sind die geschätzten oberen Grenzwerte, für ein Konfidenzniveau von 95% (90%), $\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) < 7.3(5.9) \cdot 10^{-10}$ und $\mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) < 2.3(1.8) \cdot 10^{-9}$ bzw. $< 1.8(1.5) \cdot 10^{-9}$, vorausgesetzt, die Zerfallamplitude des B_s^0 wird vollständig von den leichten (kurzlebigen) oder schweren (langlebigen) Masseneigenzuständen des $B_s^0 - \bar{B}_s^0$ -Systems dominiert. Dies stellt eine Sensitivitätssteigerung um den Faktor 1.6 für $B^0 \rightarrow e^\pm \mu^\mp$ und 2.8 für $B_s^0 \rightarrow e^\pm \mu^\mp$ (falls der schwere Eigenzustand dominiert) dar, im Vergleich zu den bisher besten zu erwarteten Grenzwerten, die von der LHCb-Kollaboration mit den Daten von 2011 und 2012 veröffentlicht wurden.

Des Weiteren präsentiert diese Dissertation den Qualitätskontrollprozess für Silizium-Photomultiplier-Detektoren, die im Scintillating Fibre (SciFi)-Tracker des jüngsten Upgrades des

Abstrakt

LHCb-Detektors eingesetzt werden.

Schlüsselwörter: Teilchenphysik, LHCb, LHC, seltene Zerfälle, Flavour, Lepton-Flavour-Verletzung, SciFi-Tracker, SiPMs

Contents

| | |
|---|------------|
| Acknowledgements | i |
| Abstract (English/Deutsch) | iii |
| 1 Theoretical framework and introduction to lepton flavour violation | 1 |
| 1.1 The fundamental particles | 1 |
| 1.2 Particle interactions | 2 |
| 1.2.1 Quantum Chromodynamics | 3 |
| 1.2.2 The electroweak interaction | 4 |
| 1.2.3 The Higgs mechanism | 6 |
| 1.3 Lepton Flavour Violation in b -hadron decays | 8 |
| 1.4 Limits on lepton flavour violating decays | 10 |
| 2 Experimental setup | 13 |
| 2.1 The Large Hadron Collider | 13 |
| 2.2 The LHCb experiment | 15 |
| 2.2.1 The tracking system | 16 |
| 2.2.2 Particle identification | 21 |
| 2.2.3 The trigger system | 26 |
| 2.2.4 Stripping | 27 |
| 2.2.5 Electron reconstruction | 27 |
| 3 The LHCb Run 3 upgrade | 29 |
| 3.1 Run 3 detector upgrade | 29 |
| 3.2 Run 3 trigger upgrade | 31 |
| 3.3 The SciFi tracker upgrade | 31 |
| 3.3.1 Detector design | 31 |
| 3.3.2 SiPMs for the SciFi tracker | 33 |
| 3.3.3 SiPM quality assurance | 34 |
| 4 Search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ | 37 |
| 4.1 Analysis strategy | 37 |
| 4.2 Data and simulation samples | 39 |
| 4.3 Event selection | 40 |

Contents

| | | |
|-------|---|-----|
| 4.3.1 | Stripping selection | 40 |
| 4.3.2 | Decay-tree fitting | 43 |
| 4.3.3 | Truth-matching | 43 |
| 4.3.4 | Pre-selection | 43 |
| 4.3.5 | Trigger selection | 46 |
| 4.3.6 | Particle identification | 48 |
| 4.3.7 | Multivariate analysis | 50 |
| 4.3.8 | Optimisation and binning in multivariate classifier response | 54 |
| 4.4 | Normalisation mass fits | 56 |
| 4.5 | Efficiencies and corrections | 60 |
| 4.5.1 | Geometrical acceptance | 60 |
| 4.5.2 | Track reconstruction efficiency corrections | 61 |
| 4.5.3 | PID efficiencies | 62 |
| 4.5.4 | L0 trigger efficiency corrections | 69 |
| 4.5.5 | Kinematic corrections | 75 |
| 4.5.6 | Efficiency determination | 79 |
| 4.5.7 | $r_{J/\psi}$ cross-check | 80 |
| 4.5.8 | B_s^0 mass eigenstates and lifetime correction | 82 |
| 4.5.9 | $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-)$ efficiencies | 82 |
| 4.6 | Physics background | 83 |
| 4.6.1 | Expectations from corrected simulation | 83 |
| 4.6.2 | Validation of the expected peaking backgrounds with $B_{(s)}^0 \rightarrow hh'$ decays | 86 |
| 4.6.3 | Pass-fail cross-check validation of peaking backgrounds | 89 |
| 4.7 | Fit to data | 91 |
| 4.7.1 | Fit strategy | 91 |
| 4.7.2 | Mass resolution calibration | 92 |
| 4.7.3 | Signal models | 95 |
| 4.7.4 | Physics background models | 97 |
| 4.7.5 | Pseudo-experiments | 100 |
| 4.8 | Branching fraction limits | 104 |
| 4.8.1 | Expected limits from pseudo-experiments | 105 |
| 4.9 | Systematic uncertainties | 109 |
| 4.9.1 | Efficiency-related systematics | 110 |
| 4.9.2 | Non-efficiency related systematics | 112 |
| 4.9.3 | Impact of systematic uncertainties on the limits | 116 |
| 4.10 | Summary and conclusion | 116 |

A Appendix 121

| | | |
|-------|---|-----|
| A.1 | Multivariate analysis | 121 |
| A.1.1 | Track isolation | 121 |
| A.1.2 | Sculpting of the combinatorial background | 121 |

| | | |
|-------|--|------------|
| A.2 | Normalisation fits | 122 |
| A.3 | Efficiencies and corrections | 122 |
| A.3.1 | Geometrical acceptance | 122 |
| A.3.2 | Track reconstruction efficiency corrections | 125 |
| A.3.3 | PID efficiencyies | 125 |
| A.3.4 | L0 trigger efficiency corrections | 136 |
| A.3.5 | Kinematic corrections | 145 |
| A.3.6 | $r_{J/\psi}$ cross-check | 152 |
| A.4 | Physics background | 152 |
| A.4.1 | Background expectations from corrected simulation | 152 |
| A.4.2 | Validation of the expected peaking backgrounds with $B_{(s)}^0 \rightarrow hh'$ decays | 159 |
| A.5 | Pass-fail cross-check validation of peaking backgrounds | 166 |
| | Bibliography | 171 |
| | Curriculum Vitae | 179 |

1 Theoretical framework and introduction to lepton flavour violation

This chapter provides a short overview of the underlying theoretical concepts in elementary particle physics. Firstly, the fundamental particles and the particle interactions are introduced, followed by a brief explanation of lepton flavour violating b -hadron decays.

1.1 The fundamental particles

The Standard Model of Particle Physics (SM) [1–7] is a theoretical framework that describes the fundamental particles of matter and three of their interactions, the electromagnetic, the weak, and the strong interaction. It is based on a set of relativistic gauge Quantum Field Theories in which particles and antiparticles are described by the wave functions of the four-vector position x , denoted in the following as $\Psi(x)$ and $\bar{\Psi}(x)$. The particles of the SM can be grouped based on their spin into fermions, with half-integer spin, and bosons with integer spin. These are the constituents of all known matter and are considered to be fundamental, *i.e.*, particles without any internal structure.

The fermions are categorized as quarks and leptons into three generations, each comprised of a so-called up-type quark, a down-type quark, a charged lepton, and a light neutral lepton, neutrino, with the same lepton flavour as the charged lepton. The leptons occur in three lepton flavour types, muon, electron and tau. Neutrinos do not carry any electromagnetic charge, while the charged leptons carry exactly one unit of the elementary charge. The electric charge of the quarks is fractional, with values of $-1/3$ and $+2/3$. Furthermore, quarks are the only SM fermions that interact via the strong interaction and therefore carry a so-called colour charge (red R, green G, or blue B). The corresponding antifermions, which carry opposite electric charges, are also described by the Standard Model.

The bosons in the SM arise as a consequence of the particle interactions, where they are introduced as mediators. There are 13 bosons in the SM: 12 vector bosons with spin one, and one scalar Goldstone boson with spin zero. The mediators of the weak and electromagnetic interactions are the charged W^\pm bosons, the neutral vector bosons Z^0 and γ (photon). Strong

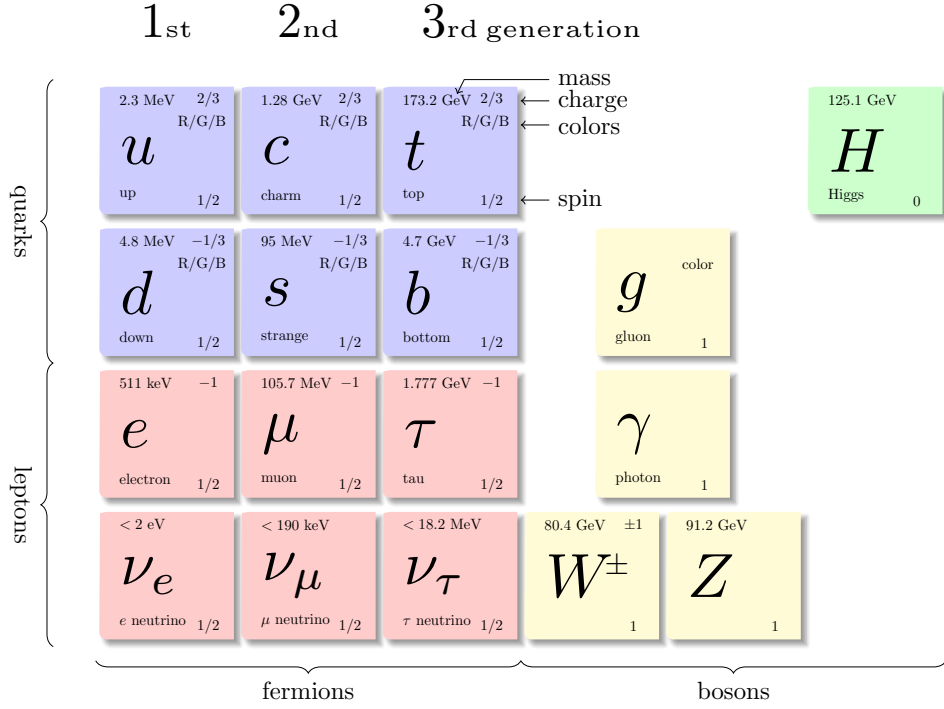


Figure 1.1 – The particle content of the Standard Model, with three quark and lepton generations, 12 vector bosons (yellow), and one scalar boson (green). The graphic is inspired from Ref. [8].

interactions are mediated by eight massless gluon fields, g . The recently discovered scalar Higgs boson plays a crucial role in the Standard Model. It arises from the introduction of the Higgs mechanism, which is responsible for the masses of the elementary particles. The full SM content is displayed in Fig. 1.1.

1.2 Particle interactions

The particle interactions in the Standard Model are based on the principle of local gauge invariance. In this principle, the interaction of a fermion wave function can be expressed as local gauge transformations of the following form:

$$\Psi \longrightarrow \Psi'(x) = e^{i\alpha(x)}\Psi(x), \quad (1.1)$$

where $\alpha(x)$ is a phase factor that depends on the chosen symmetry group. The Lagrangian of the system is required to be invariant under a Lie group of these transformations. The term “local” refers to the fact that these transformations depend on the space-time coordinates. The invariance is achieved by covariant derivatives, resulting in the introduction of vector gauge fields that are the generators of the Lie groups, *i.e.* the gauge bosons mentioned in Sec. 1.1.

In summary, the Standard Model is composed of two Quantum Field Theories: Quantum Chromodynamics (QCD), which describes the interactions of quarks and gluons under gauge transformations of the $SU(3)_C$ group, and the unified electroweak theory, which describes the interactions of charged particles and neutrinos under the $SU(2)_L \times U(1)_Y$ symmetry groups.

The fact that some of the gauge bosons are massive violates the local $SU(2)_L \times U(1)_Y$ gauge invariance. To overcome this issue, the Higgs mechanism is introduced, where the symmetry is spontaneously broken by a scalar field with a non-zero vacuum expectation value. This leads to additional mass terms for the gauge bosons in the Lagrangian. Furthermore, the fermion mass terms are introduced by so-called Yukawa couplings, which arise from interactions between the fermion fields and the Higgs field.

1.2.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) describes the strong interaction of massive spin-1/2 quarks and massless spin-1 gluons. This interaction is based on the non-abelian gauge group $SU(3)_C$, with the strong charge referred to as colour (C). The QCD Lagrangian can be expressed as (adapted from Ref. [9])

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}F_{\mu\nu}^a F_a^{\mu\nu} + \sum_f \bar{\Psi}_f^C (i\gamma^\mu D_\mu - m_f) \Psi_f^C. \quad (1.2)$$

The first term is the field strength tensor (adapted from Ref. [10])

$$F_{\mu\nu}^a = \partial_\nu G_\mu^a - \partial_\mu G_\nu^a + g_s f^{abc} G_\mu^b G_\nu^c, \quad (1.3)$$

which is comprised of eight massless gluon fields denoted as G_μ^a . The interaction strength is quantified by g_s and f^{abc} denotes the structure constants of the strong interaction, where a, b, c run over all eight degrees of freedom of the $SU(3)_C$ group. The eight degrees of freedom refer to the eight generators of the $SU(3)_C$ group, denoted as T . The second term in Eq. (1.2), which is summed over all quark flavours ($f = u, d, c, s, t, b$), includes the covariant derivative (adapted from Ref. [10])

$$D_\mu = \partial_\mu + i g_s T^a G_\mu^a, \quad (1.4)$$

as well as the term m_f representing the fermion masses. In the Lagrangian, γ denotes the Dirac matrices. The quarks form $SU(3)_C$ colour triplets with the notation C = red, green, blue, while anti-quarks form triplets of the conjugated group and carry the corresponding anti-colours. Furthermore, the gluons themselves are carrying colour charges, responsible for gluon-gluon self-coupling interactions.

The coupling strength of the strong interaction strongly depends on the energy scale. It is typically described by $\alpha_s = g_s^2/4\pi$ [10]. At high energies, the coupling strength becomes sufficiently small, *i.e.*, free particle-like behaviour can be observed, which is called **asymptotic freedom**. However, for low energies, the coupling strength increases drastically and can be

expressed as [9]

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2N_f) \ln(\frac{Q^2}{\Lambda_{\text{QCD}}^2})}. \quad (1.5)$$

N_f denotes the number of quark flavours, that have to be taken into account at an energy Q^2 . In perturbation theory, ultraviolet divergences lead to a breakdown of the perturbation at an energy scale Λ_{QCD} , which is of the order of a few hundred MeV [9]. As a consequence of the behaviour of the coupling strength expressed in Eq. (1.5), quarks and gluons cannot be observed as free particles propagating over macroscopic scales, since the coupling strength increases for decreasing energies, resulting in the **confinement** of quarks. Experimentally, only the colour-neutral bound states, the hadrons (baryons or mesons), can be observed, which represent the singlet states of the $SU(3)_C$ group. In this thesis, B mesons play an important role. These are bound states of \bar{b} quarks, such as the B_s^0 meson, which is composed of a \bar{b} and an s quark.

1.2.2 The electroweak interaction

A unified theory of the **electromagnetic force** and the **weak interaction** was introduced by Glashow, Weinberg, and Salam (GSW) [1, 6, 7], based on the $SU(2)_L \times U(1)_Y$ symmetry group. The generators of the $SU(2)_L \times U(1)_Y$ group introduce four gauge vector fields: a boson triplet $\vec{W} = (W_1, W_2, W_3)^T$, and a boson singlet B_0 . These can be associated with the observed physical particles, as explained in the following. Moreover, the $SU(2)$ symmetry implies the conservation of a quantum number known as weak isospin, which is analogous to spin in quantum mechanics. The third component of the weak isospin is denoted as I_3 .

An important consequence of the $SU(2)$ symmetry transformation is the maximal violation of parity conservation, *i.e.*, the weak interaction acts differently on fermions depending on their chirality. Fermions with negative chirality (left-handed fermions) with $I_3 = \pm 1/2$ form weak isospin doublets, while fermions with positive chirality (right-handed) are weak isospin singlets with $I_3 = 0$. Therefore, only left-handed fermions and right-handed antifermions interact via the weak force.

The electromagnetic interaction, described by the $U(1)$ symmetry group, does not distinguish between left- and right-handed particles. The weak hypercharge Y is an additional quantum number connected to the electromagnetic charge Q and the weak isospin component I_3 via the Gell-Mann–Nishijima formula [10]:

$$Q = I_3 + \frac{Y}{2}. \quad (1.6)$$

The fermion isospin eigenstates (d', s', b'), also known as flavour eigenstates, are related to the fermion mass eigenstates (d, s, b) that couple to the gauge fields via the Cabibbo-Kobayashi-

Maskawa (CKM) matrix [11, 12]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.7)$$

The CKM matrix elements are determined by experimental measurements and the magnitudes of its matrix elements are [13]

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.97373 \pm 0.00031 & 0.2243 \pm 0.0005 & (3.82 \pm 0.20) \cdot 10^{-3} \\ 0.221 \pm 0.004 & 0.975 \pm 0.006 & (40.8 \pm 1.4) \cdot 10^{-3} \\ (8.6 \pm 0.2) \cdot 10^{-3} & (41.5 \pm 0.9) \cdot 10^{-3} & 1.014 \pm 0.029 \end{pmatrix}. \quad (1.8)$$

These magnitudes correspond to different transition probabilities between the quark states, where, for instance, a transition from a t quark to a b quark is much more probable than a transition from a t quark to a d quark.

In order to unify the electromagnetic and weak interactions, the complete GSW Lagrangian can be expressed as (modified from Ref. [10])

$$\mathcal{L}_{GSW} = -\frac{1}{4}F_{\mu\nu}^a F_a^{\mu\nu} f_{\mu\nu} f^{\mu\nu} + \sum_f \bar{\Psi}_{f_R} (i\gamma^\mu D_{\mu_R}) \Psi_{f_R} + \sum_f \bar{\chi}_{f_L} (i\gamma^\mu D_{\mu_L}) \chi_{f_L}. \quad (1.9)$$

The first term describes the kinematics of the fields, with the field strength tensors given by (adapted from Ref. [10])

$$F_{\mu\nu}^a = \partial_\nu W_\mu^a - \partial_\mu W_\nu^a + g\epsilon^{abc} W_\mu^b W_\nu^c \quad (1.10)$$

and

$$f_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu. \quad (1.11)$$

ϵ^{abc} is the Levi-Civita symbol with $a, b, c = 1, 2, 3$. The second Lagrangian term describes the interaction of right-handed fermion singlets, Ψ_{f_R} , with the covariant derivative (adapted from Ref. [10])

$$D_{\mu_R} = \partial_\mu + g' \frac{Y_R}{2} B_\mu. \quad (1.12)$$

χ_{f_L} in the third term of Eq. (1.9) denotes the left-handed fermion doublets. The corresponding covariant derivative is (adapted from [10])

$$D_{\mu_L} = \partial_\mu + ig' \frac{Y_L}{2} B_\mu + ig \vec{I} \cdot \vec{W}_\mu. \quad (1.13)$$

g and g' refer to the coupling strength of the interactions. The introduced gauge fields can be related, with simple mathematical transformations, to the experimentally observed particles. In case of weak charged current interactions, the observed mediators are the W^\pm bosons,

which can be constructed from the gauge fields as [10]

$$W_\mu^\pm = \frac{1}{2}(W_\mu^1 \pm iW_\mu^2). \quad (1.14)$$

The mediators of the neutral current interactions are the Z^0 boson and the photon γ (denoted as A_μ), which can be obtained by an orthogonal rotation of the gauge fields with the Weinberg angle Θ_W (adapted from Ref. [10]):

$$\begin{pmatrix} A_\mu \\ Z_\mu^0 \end{pmatrix} = \begin{pmatrix} \cos\Theta_W & \sin\Theta_W \\ -\sin\Theta_W & \cos\Theta_W \end{pmatrix} \cdot \begin{pmatrix} B_\mu \\ W_\mu^0 \end{pmatrix}. \quad (1.15)$$

Θ_W can be obtained from the coupling constants and the hypercharge as (adapted from Ref. [10])

$$\cos\Theta_W = \frac{g}{\sqrt{g^2 + g'^2 Y_L^2}}. \quad (1.16)$$

Finally, the observed neutral vector bosons are given by

$$A_\mu = \frac{gB_\mu - g'Y_L W_\mu^0}{\sqrt{g^2 + g'^2 Y_L^2}} \quad \text{and} \quad Z_\mu^0 = \frac{gW_\mu^0 + g'Y_L B_\mu}{\sqrt{g^2 + g'^2 Y_L^2}}. \quad (1.17)$$

With the GSW theory, the unification of two fundamental interactions is achieved in a beautiful manner just by introducing a set of four gauge fields. However, gauge invariance only holds for massless gauge bosons. In the next section, a mass generation mechanism is introduced that preserves gauge invariance for massive gauge bosons.

1.2.3 The Higgs mechanism

This section explains the mass generation mechanism in the SM and follows Ref. [10], unless specified otherwise.

In the unified electroweak interaction, the local $SU(2)_L \times U(1)_Y$ gauge invariance is broken by the fact that the observed mediators, the gauge bosons, are massive particles. To solve this issue, Brout, Englert, Guralnik, Hagen, Kibble, and Higgs [14–17] introduced a complex scalar field:

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi_0(x) \end{pmatrix}. \quad (1.18)$$

This so-called Higgs field forms a doublet of the weak isospin component I_3 with a hypercharge of one.

The interaction of the Higgs field with the electroweak gauge fields is described by the Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi, \Phi^\dagger), \quad (1.19)$$

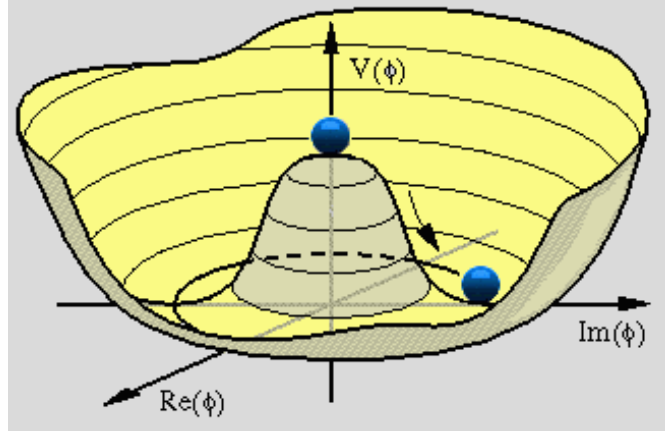


Figure 1.2 – The Higgs potential for $\mu^2 > 0$ and $\lambda > 0$ [18].

where Φ is the Higgs doublet, Φ^\dagger is the complex conjugate, and $V(\Phi, \Phi^\dagger)$ the Higgs potential. D_μ are the covariant derivatives defined in Eq. (1.12), which introduce the coupling to \vec{W}_μ and B_μ . In its most general form, $\mathcal{L}_{\text{Higgs}}$ follows the $SU(2)_L \times U(1)_Y$ symmetry of the electroweak gauge theory. However, in the second term of the Lagrangian the potential

$$V(\Phi, \Phi^\dagger) = -\mu^2 \Phi^\dagger \Phi + \frac{\lambda}{4} (\Phi^\dagger \Phi)^2 \quad (1.20)$$

is introduced. The parameters μ and λ characterize the potential shape, displayed in Fig. 1.2 for $\mu^2 > 0$ and $\lambda > 0$. The potential describes the self-interaction of the Higgs field. Satisfying $\Phi^\dagger \Phi = 2\mu^2/\lambda$, the potential takes its minimum value in the complex plane as

$$\langle \Phi_0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad \text{with } v = \frac{2\mu}{\sqrt{\lambda}}. \quad (1.21)$$

In the ground state $\langle \Phi \rangle_0$ of the system, the symmetry $SU(2)_L \times U(1)_Y$ is broken, *i.e.*, the system keeps only the local $U(1)_Y$ gauge symmetry of the electromagnetic force, which is called **spontaneous symmetry breaking**. By expanding the ground state in the form $\phi_0(x) = v + H(x) + i\chi(x)$, Eq. (1.18) can be expressed as

$$\Phi(x) = \begin{pmatrix} \phi^\dagger(x) \\ (v + H(x) + i\chi(x))/\sqrt{2} \end{pmatrix}, \quad (1.22)$$

where $\chi(x)$ and ϕ^\dagger can be eliminated by choosing a gauge where $\chi(x) = \phi^\dagger = 0$, *i.e.* these components are unphysical degrees of freedom. Then the potential, Eq. (1.20), becomes

$$V(\Phi, \Phi^\dagger) = \mu^2 H^2 + \frac{\mu^2}{v} H^3 + \frac{\mu^2}{4v^2} H^4. \quad (1.23)$$

H can be identified with a scalar particle, the famous Higgs boson with a mass of $m_H = \sqrt{2}\mu$.

Chapter 1. Theoretical framework and introduction to lepton flavour violation

In the Lagrangian of Eq. (1.19), the coupling between the electroweak gauge fields \vec{W}_μ and B_μ and the Higgs are a consequence of the kinetic term. The measured electroweak gauge bosons W^\pm and Z can be derived as

$$m_W = \frac{1}{2}g\nu \quad \text{and} \quad m_Z = \frac{1}{2}\sqrt{g^2 + g'^2}\nu. \quad (1.24)$$

Also massive charged fermions can be explained by the introduction of the so-called Yukawa coupling between the Higgs field and the fermions, described by [10]

$$\mathcal{L}_{\text{Yukawa}} = -\sum_f m_f \bar{\Psi}_f \Psi_f - \sum_f \frac{m_f}{\nu} \bar{\Psi}_f \Psi_f H. \quad (1.25)$$

The wavefunctions Ψ_f are summed up over all massive fermions and couple with individual constants, proportional to the fermion mass

$$m_f = G_f \frac{\nu}{\sqrt{2}}, \quad (1.26)$$

where G_f is the Yukawa coupling constant.

The Higgs mechanism shows in a very elegant way how the particles gain their mass. However, the mass of the Higgs boson, as well as the fermion masses, can only be determined directly via measurements. The first observations of the Higgs boson, announced on the 4th of July 2012 by the ATLAS and CMS collaborations [19, 20], have been a cornerstone for the Standard Model and its accuracy.

1.3 Lepton Flavour Violation in b -hadron decays

This thesis is primarily focused on the search for the lepton flavour violating decays $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$. The motivation behind this search is explained in the following.

The leptons of the Standard Model, introduced in Sec. 1.1, can be grouped into three different lepton flavours, the electron, the muon, and the tau flavour, denoted in the following as e, μ and τ . The lepton flavour is described by the lepton family number, L_i , with $i = e, \mu, \tau$. Each lepton family contains a charged lepton and a corresponding neutrino, which both share the same lepton family number, *e.g.*, the electron and the electron neutrino form a lepton family with $L_e = 1$.

Lepton Flavour Violation (LFV) refers to particle transitions where the lepton family number is not conserved, *e.g.*, the transition from a muon to an electron and a photon $\mu \rightarrow e\gamma$. For this decay, the initial state has the lepton family numbers $L_\mu = 1$ and $L_e = 0$, while the final state has the lepton family numbers $L_\mu = 0$ and $L_e = 1$, indicating a violation of lepton flavour conservation.

With the discovery of neutrino oscillations [21], *e.g.*, $\nu_\mu \rightarrow \nu_\tau$, LFV processes are well es-

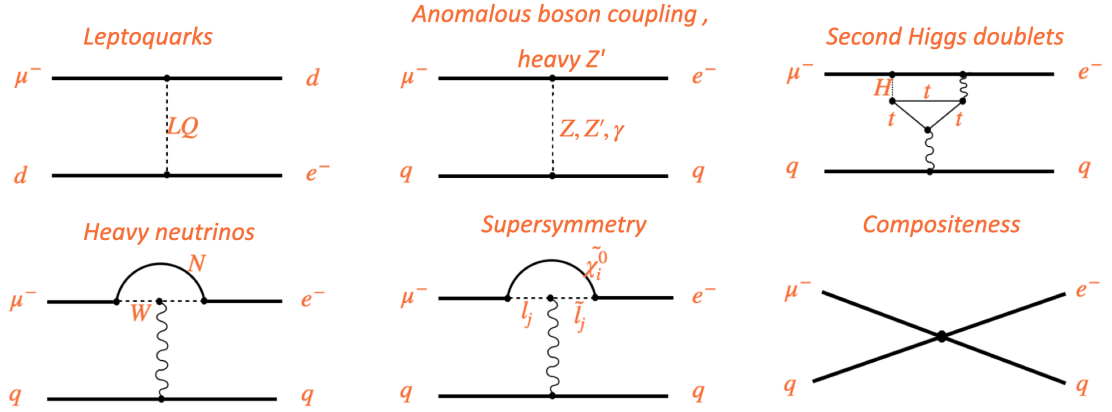


Figure 1.3 – Beyond-Standard-Model processes with cLFV, inspired by Ref. [30].

established in the neutral lepton sector. Furthermore, the CKM matrix (see Eq. (1.7)) shows that the weak interaction mixes different quark flavour eigenstates, hence flavour is also not conserved in the quark sector. Despite the observed flavour violation of quarks and neutral leptons, there are no signs of LFV processes in the charged lepton sector, which is referred to as **charged Lepton Flavour Violation (cLFV)**. The current strongest experimental limit on the branching fraction of a cLFV decay is given by the search for the $\mu \rightarrow e \gamma$ decay, which is $\mathcal{B}(\mu^+ \rightarrow e^+ \gamma) < 4.2 \cdot 10^{-13}$ at 90% confidence level [22]. The corresponding Standard Model prediction is $\mathcal{B}(\mu^+ \rightarrow e^+ \gamma) \sim 10^{-54}$ [23].

Several theoretical models beyond the Standard Model predict the existence of lepton flavour violating processes, for example models with leptoquarks [24, 25], a model with a new gauge Z' boson [26], models with a heavy Dirac neutrino singlet [27], models with a Higgs doublet [28], the Pati-Salam model [29] or supersymmetric and composite lepton-quark models [30]. Figure 1.3 illustrates Feynman diagrams of beyond-Standard-Model processes with lepton flavour violation.

While the Standard Model, including the neutrino oscillations, predicts branching fractions of cLFV processes typically of the order $\mathcal{O}(10^{-44})$ or smaller [31, 32], new physics processes beyond the SM can enhance these branching fractions by several orders of magnitude. Therefore, the observation of cLFV processes would be a clear sign for new physics beyond the Standard Model.

The search for cLFV in b -hadron decays is particularly interesting since it can be related to one of the most extensively researched topics in particle physics: **Lepton Flavour Universality (LFU)**. LFU is a fundamental assumption in the Standard Model that the coupling of the gauge bosons to the leptons is independent of the lepton family. This implies that for a given process, e.g., $b \rightarrow s l^+ l^-$, the decay rates for different lepton flavours ($l = e, \mu, \tau$) should be equal. In the past, measurements of these decay rates have shown tensions compared to the Standard Model predictions [25, 33]. However, recent studies performed by the LHCb collaboration are

consistent with LFU [34, 35]. Nevertheless, the study of LFU remains of great interest and is still intensely studied [34–37].

There is a clear relation between the ratio

$$R_K = \frac{\mathcal{B}(B \rightarrow K \mu^+ \mu^-)}{\mathcal{B}(B \rightarrow K e^+ e^-)} \quad (1.27)$$

and LFV processes, for example in leptoquark models [25]. In particular, the decays of interest in this thesis, $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$, can be directly related to LFU [25]:

$$\frac{\mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)}{\mathcal{B}(B_{(s)}^0 \rightarrow \mu^\pm \mu^\mp)} \sim 0.01 \left(\frac{1 - R_K}{0.23} \right)^2. \quad (1.28)$$

The observed deviations in LFU tests, as well as several new physics models predicting cLFV, have made the search for such processes one of the most interesting topics in modern particle physics.

1.4 Limits on lepton flavour violating decays

The current most stringent upper limits on the branching fractions of charged Lepton Flavour Violation processes have been determined from the MEG and the SINDRUM I and II experiments, at 90% confidence level:

$$\begin{aligned} \mathcal{B}(\mu^+ \rightarrow e^+ \gamma) &< 4.2 \cdot 10^{-13} \quad [22]; \\ \mathcal{B}(\mu^+ \rightarrow e^+ e^+ e^-) &< 1.0 \cdot 10^{-12} \quad [38]; \\ \mathcal{B}(\mu^+ \text{Au} \rightarrow e^- \text{Au}) &< 7.0 \cdot 10^{-13} \quad [39]. \end{aligned}$$

Furthermore, several searches for cLFV processes have been performed in the b quark sector by the Belle, BaBar, and LHCb collaborations. The branching fraction upper limits set by the LHCb collaboration are typically at the level of $\mathcal{O}(10^{-9})$ or larger.

The LHCb collaboration performed two searches for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ decays based on the data taken during the first LHC run period, Run 1, in 2011 and 2012. A first analysis was based on the 2011 dataset and established the first limits on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions at 90%(95%) confidence level [40]:

$$\begin{aligned} \mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) &< 2.8(3.7) \cdot 10^{-9}; \\ \mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) &< 1.1(1.4) \cdot 10^{-8}. \end{aligned}$$

These results have been improved by an updated analysis, based on the full Run 1 data-set (2011 and 2012) with improved analysis techniques. Furthermore, the upper limits on the $B_s^0 \rightarrow e^\pm \mu^\mp$ branching fraction have been determined for the two different mass eigenstates of the B_s^0 meson, the light and heavy mass eigentates which result from flavour mixing [41].

The two mass eigenstates have distinct lifetimes, which impact the selection efficiencies and thus the upper limit on the branching fraction. The so-far best obtained upper limits on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions by the LHCb collaboration at 90%(95%) confidence level are [42]

$$\begin{aligned}\mathcal{B}(B^0 \rightarrow e^\pm \mu^\mp) &< 1.0(1.3) \cdot 10^{-9}; \\ \mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) &< 6.0(7.2) \cdot 10^{-9} \text{ for the light mass eigenstate;} \\ \mathcal{B}(B_s^0 \rightarrow e^\pm \mu^\mp) &< 5.4(6.3) \cdot 10^{-9} \text{ for the heavy mass eigenstate.}\end{aligned}$$

In the second LHC data-taking period from 2015 until 2018 (Run 2), the LHCb experiment operated at a much higher center-of-mass energy (13 TeV instead of 8 TeV) and significantly increased luminosity (see Fig. 2.4), which resulted in a massive increase in the amount of collected data and hence better sensitivity for searches for cLFV decays. Therefore, the aim of this analysis, which uses the data collected between 2016 and 2018, is to further improve the search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ decays, with respect to the current results.

From a pure statistical perspective, a rough estimate of the improvement on the upper limit can be derived, considering the expected number of signal events, defined as

$$\mathcal{N}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) = \sigma(b\bar{b}) \cdot \varepsilon(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) \cdot \mathcal{L} \cdot \mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp), \quad (1.29)$$

where $\sigma(b\bar{b})$ is the $b\bar{b}$ production cross-section, $\varepsilon(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)$ the total efficiency and \mathcal{L} the integrated luminosity. To determine the increase in statistics from Run 1 to Run 2, two factors have to be taken into account. Firstly, the $b\bar{b}$ production cross-section increased by a factor of two. Secondly, the integrated luminosity increased from roughly from 3.2 fb^{-1} to 5.4 fb^{-1} . The increase in statistics can now be estimated, assuming that the selection efficiency remains unchanged, by evaluating the ratio of Eq. (1.29) between Run 2 and Run 1:

$$\begin{aligned}\frac{\mathcal{N}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)_{\text{Run 2}}}{\mathcal{N}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)_{\text{Run 1}}} &= \frac{\sigma(b\bar{b})_{\text{Run 2}}}{\sigma(b\bar{b})_{\text{Run 1}}} \cdot \frac{\mathcal{L}_{\text{Run 2}}}{\mathcal{L}_{\text{Run 1}}} \\ &= 2 \cdot \frac{5.4 \text{ fb}^{-1}}{3.2 \text{ fb}^{-1}} = 3.38.\end{aligned} \quad (1.30)$$

The estimated relative increase in statistics from Run 1 to Run 2 is about 3.38. The relative statistical error is given by taking the square root of the relative statistical increase and is 1.83. Therefore, an improvement on the $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ branching fraction limits of roughly a factor of 2 can be expected, considering only the increase in statistics.

2 Experimental setup

The Large Hadron Collider beauty experiment (LHCb) is based at the Large Hadron Collider (LHC) [43]. It is one of the four main experiments at the LHC, detecting highly energetic particle collisions. LHCb is specifically designed for heavy-flavour physics and the investigation of CP violation.

In the following, a brief overview of the experimental setup is presented. First, the LHC is introduced, followed by a description of the LHCb detector and a short overview of the particle reconstruction at LHCb.

2.1 The Large Hadron Collider

The LHC is located between 45 and 170 m below ground, on the Franco-Swiss border, near Geneva, in a circular tunnel with a circumference of approximately 27 km. Currently the LHC is the most powerful particle accelerator of all times and is specifically designed to perform high-energy hadron collisions, such as proton-proton (pp), lead-lead ($PbPb$) or proton-lead (pPb) collisions.

The particles are accelerated in a chain of smaller accelerators (see Fig. 2.1), before they are injected into the LHC. In Run 2, protons underwent the following acceleration steps:

1. The particles are accelerated in the Linear Accelerator LINAC2 to an energy level of 50 MeV¹;
2. The particles reach an energy of approximately 1.4 GeV in the Proton Synchrotron Booster (PSB);
3. The Proton Synchrotron (PS) accelerates the particles towards 25 GeV;
4. In the large Super Proton Synchrotron (SPS), the particles reach an energy of 450 GeV;
5. The particles exiting the SPS are injected into the LHC, where they are accelerated in two parallel beam pipes. In one beam pipe particles are accelerated clockwise and in

¹Figure 2.1 shows the acceleration chain for the year 2020, where LINAC2 is replaced with LINAC4.

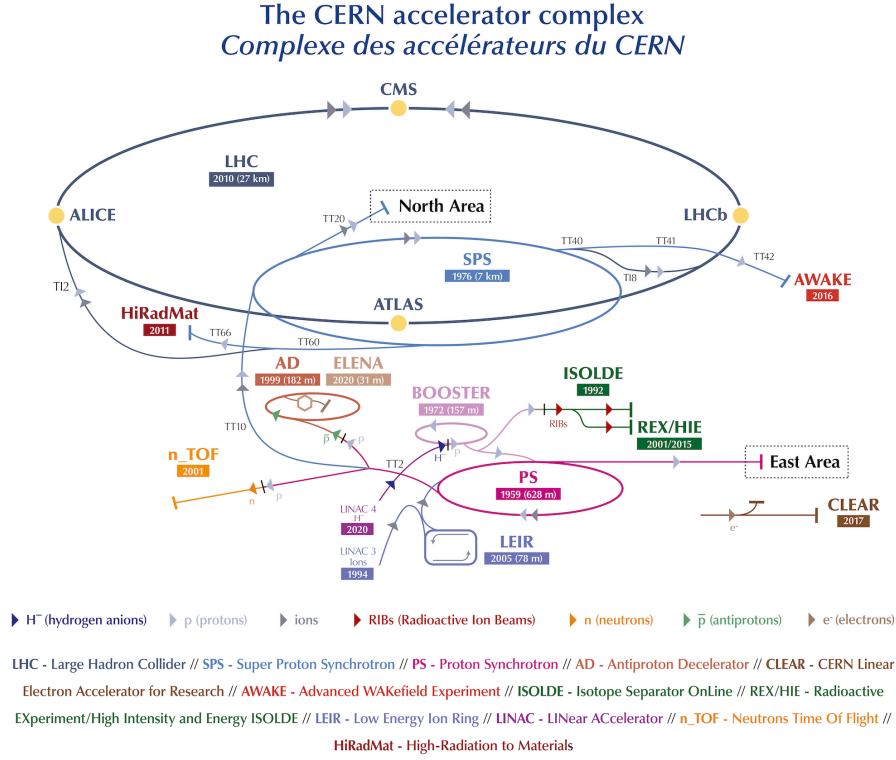


Figure 2.1 – The CERN accelerator complex in 2020 [44].

the other beam pipe counter-clockwise. The particles are accelerated up to 6.5 TeV in Run 2 before they are brought in collision.

In the LHC the particles are grouped in bunches containing 10^{11} particles with a bunch spacing of 25 ns. The LHC beams are accelerated by sixteen superconducting radio-frequency cavities and guided by 1232 main superconducting dipole magnets. 392 superconducting quadrupole magnets are used for beam focusing. The magnets are operating with a maximal field strength above 8 TeV under constant cooling with superfluid helium at 2 K. [43]

There are four interaction points at the LHC, where the particle beams cross and the main experiments **ATLAS**, **CMS**, **ALICE** and **LHCb** are placed. ATLAS² [45] and CMS³ [46] are multi-purpose detectors covering the full 4π solid angle around the collision point and are designed for high precision measurements of the Standard Model parameters and the search for new physics. The ALICE⁴ experiment [47] also covers the full 4π solid angle and is specifically designed to investigate the quark-gluon plasma in heavy ion collisions. Compared to ATLAS, CMS, and ALICE, the design of LHCb is very different and is discussed in Sec. 2.2.

²A Toroidal LHC Apparatus

³Compact-Muon-Solenoid

⁴A Large Ion Collider Experiment

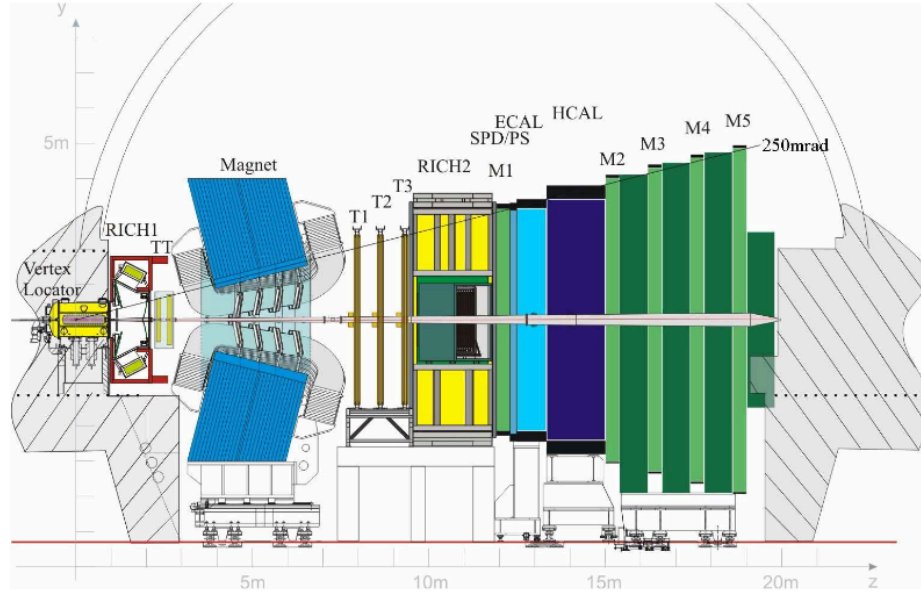


Figure 2.2 – Illustration of the LHCb detector [48].

2.2 The LHCb experiment

The Large Hadron Collider Beauty experiment (LHCb) [48], is single-arm forward spectrometer (see Fig. 2.2), specifically designed for studying physic processes involving b and c quarks. The asymmetric detector design is optimised for the measurement of $b\bar{b}$ pairs, where the corresponding production angle with respect to the beam axis peaks in the forward and backward regions (see Fig. 2.3 left). The $b\bar{b}$ pair production cross-section at LHC is very large (see Fig. 2.3 right).

The angular LHCb acceptance covers the range from 10 to 250 mrad in the vertical plane and 10 to 300 mrad in the horizontal plane, which corresponds to a pseudorapidity⁵ range of $1.6 < \eta < 4.9$. As shown in Fig. 2.2, the z axis is aligned with the beam, pointing from the collision point to the other end of the experiment, while the y axis is vertical and the x axis completes the right-handed coordinate system.

During its operation time from 2010 to 2018 (Run 1 and Run 2), the LHCb experiment recorded a large amount of data of pp collisions at several centre-of-mass energies, as shown in Fig. 2.4. To perform high precision studies, *e.g.*, the search for rare b processes, LHCb consists of different detector layers, each optimized for the measurement of specific quantities. An overview of the sub-detector systems is given in the following.

⁵ $\eta = -\ln(\tan(\frac{\Theta}{2}))$, where Θ is the angle between the particle momentum and the beam axis.

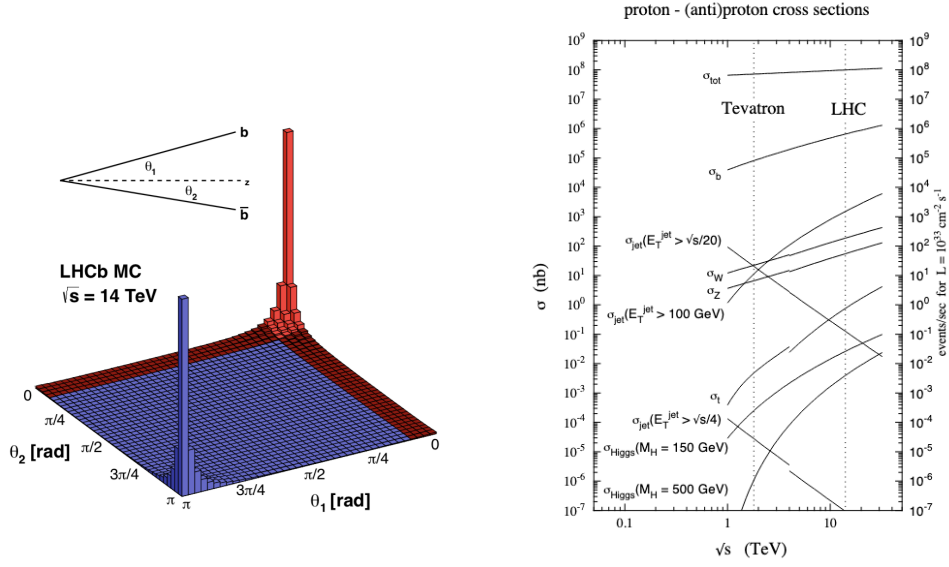


Figure 2.3 – Left: $b\bar{b}$ pair production distribution with respect to the beam axis [49]. The results are obtained from simulation at a centre-of-mass energy $\sqrt{s} = 13$ TeV. The LHCb acceptance is illustrated in red. Right: hard scattering cross-sections as a function of the centre-of-mass energy \sqrt{s} [50].

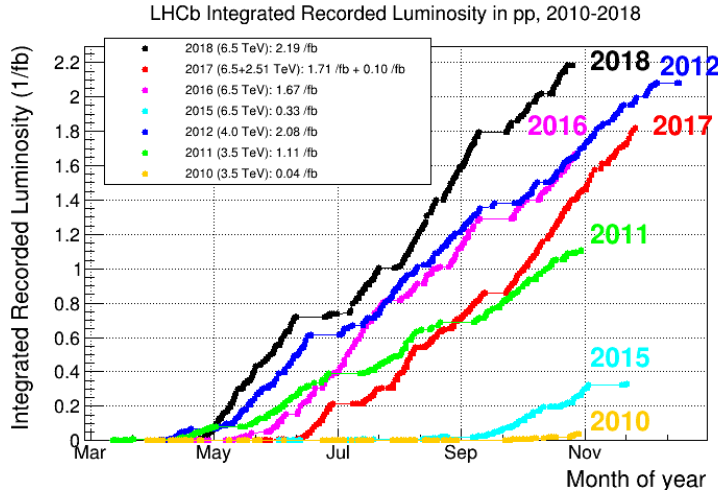


Figure 2.4 – Recorded integrated luminosity in pp collisions by LHCb during Run 1 and Run 2 [51].

2.2.1 The tracking system

The tracking system, which consists of several sub-detectors, is designed to reconstruct the trajectory of a particle from its creation vertex to its decay vertex. This provides important information, for example for the determination of the particle lifetime.

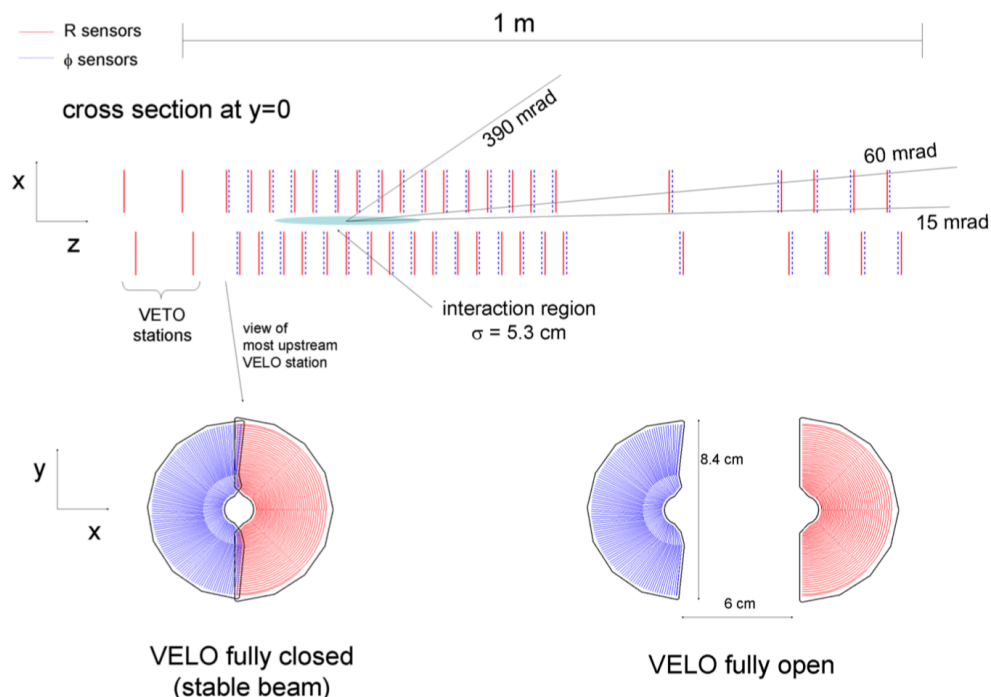


Figure 2.5 – Illustration of the VELO. Top: arrangement of the silicon strip modules along the beam axis. Bottom: cross-section of the detector modules in the closed and open state [52].

The VELO

The VERtEX LOcator (VELO) is the first sub-detector closest to the point where the protons collide. It is based on silicon strip modules, which have a circular shape around the beam direction (see Fig. 2.5). The modules are located around the interaction region covering a length of approximately 1 m along the beam line. A vacuum vessel, disconnected from the LHC primary vacuum, encloses the system with a thin (0.5 mm) aluminum foil, which also shields the modules from radio-frequency noise generated by the LHC beam. Each module consists of two halves with a radial distance of 7 mm to the LHC beam. The modules contain two different sensors. The R-sensors measure the radial distance from the beam axis and the Φ -sensors measure the azimuthal angle. The position of the sensor plane allows the determination of the coordinate along the beam axis. [52]

The VELO achieves outstanding performances, *e.g.*, an impact parameter resolution of less than $35 \mu\text{m}$ for particles with $p_T > 1 \text{ GeV}/c$ [53].

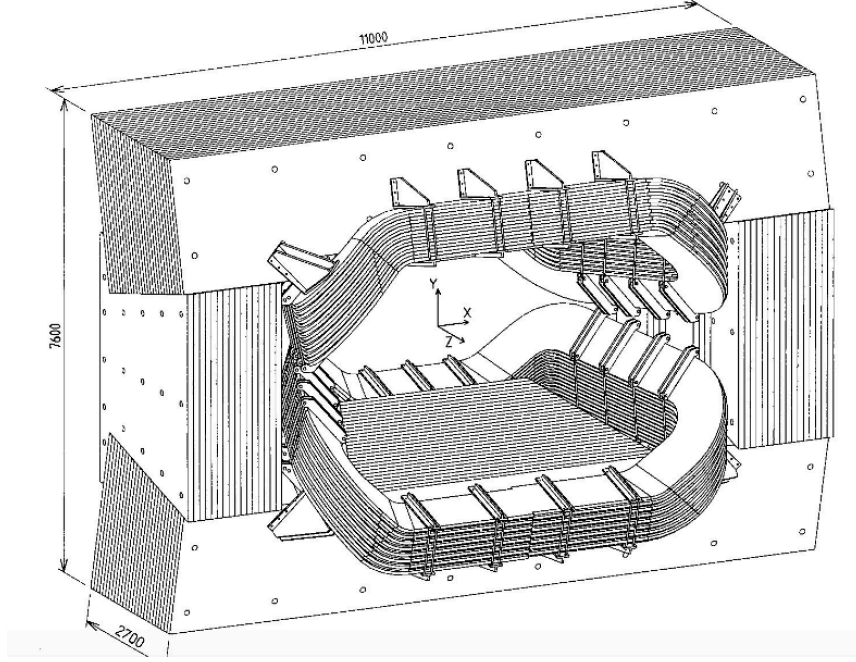


Figure 2.6 – Schematic drawing of the dipole magnet [54].

The magnet

The momentum of a charged particle is measured by bending the particle track with a giant dipole magnet placed between the Tracker Turicensis (upstream) and the first tracking station T1 (downstream).

The magnet, illustrated in Fig. 2.6, is comprised of two trapezoidal coils, each bent at 45° to a saddle-like shape. The coils are mounted in a window-frame yoke. The \vec{B} field is vertical, bending the particles in the horizontal plane. It reaches an integrated magnetic field strength of around 4 Tm, which allows the measurement of the particle momentum with a precision of 0.4% up to 200 GeV/c. [54]

In order to cope with systematic errors induced by asymmetric effects in particle detection, the magnet polarity of the magnet is reversed at regular intervals during data taking.

The Tracker Turicensis

The so-called Tracker Turicensis (TT) is located upstream of the dipole magnet and consists of four layers of silicon strip detectors grouped in two sub-stations TTa and TTb. There are three different layer geometries, as shown in Fig. 2.7. The so-called x layers (first and fourth layers) have vertical strips and are part of TTa and TTb. The strips of the u/v layers are rotated by an angle of $\pm 5^\circ$ with respect to the vertical direction. The u layer is part of TTa and the v layer is

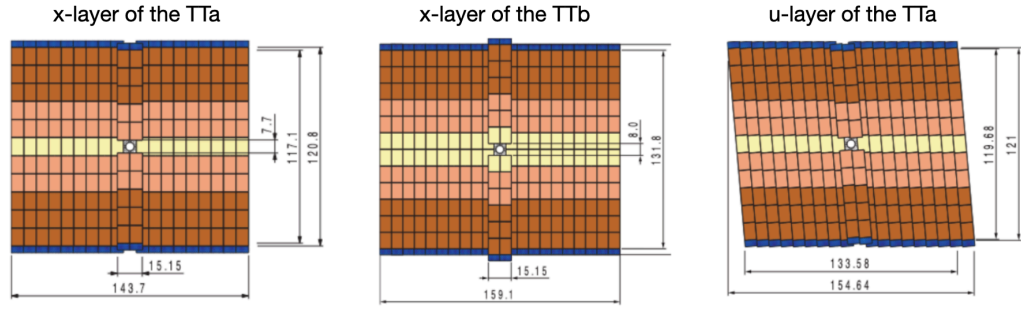


Figure 2.7 – Cross-sections of the different TT layers. The dimensions are given in cm. Taken from [55].

part of TTb. The silicon sensors are mounted in nine ladders, each with a length of 11 sensors in TTa and 12 in TTb, covering the left and right active areas of the station. Furthermore, ladders with a length of five sensors in TTa and six in TTb cover the area above and below the beam pipe. [55]

The Inner Tracker

The Inner Tracker (IT) is placed downstream of the dipole magnet and comprised of three tracking stations, T_1 , T_2 , and T_3 . The IT is located close to the beam pipe, where the hit occupancy is higher compared to the outer regions of the tracking stations. In total the IT covers only a very small fraction ($< 2\%$) of the LHCb acceptance. However, around 20% of all particles produced by the pp collisions of the LHC pass through that area. Due to the high particle density in that area, a high fine spatial resolution is required and achieved by silicon microstrip detectors. [56]

The IT stations are each built of four layers of detectors with the $x - u - v - x$ layout. The detectors are mounted in four thermally and electronically shielded boxes, which are arranged around the beam pipe. In each box, each layer is made of seven detector modules. While the detector modules placed on the top and the bottom of the beam pipe contain one detector,

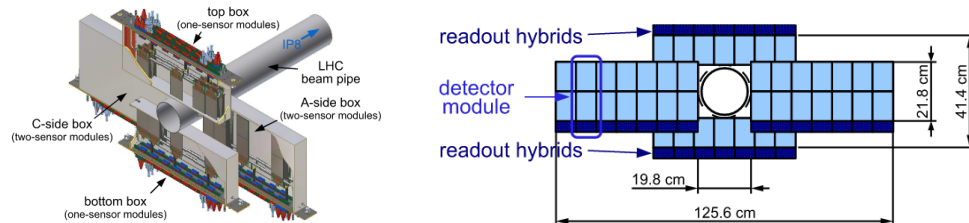


Figure 2.8 – Left: the four detector boxes of an IT station arranged around the beam pipe [48]. Right: x layer of the second IT station [48].

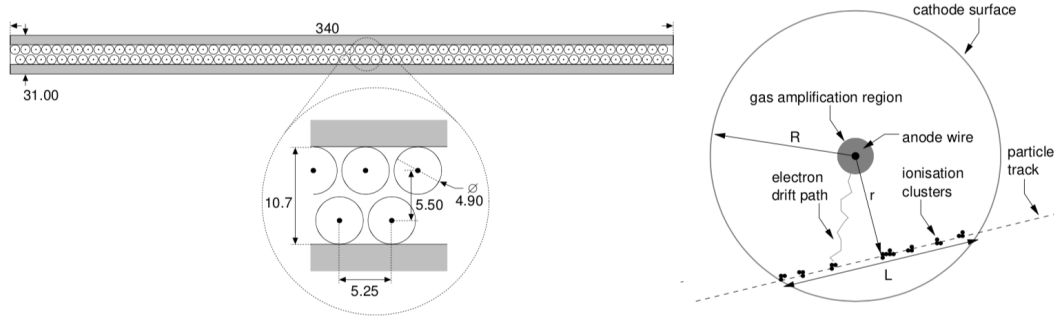


Figure 2.9 – Left: arrangement of the straw tubes in one OT module [58]. Right: Cross-section of a straw tube [58].

the modules on the sides contain two silicon detectors. Additionally, each module contains an electronic readout board. Figure 2.8 shows the four detector boxes of a station and the layout of an x layer.

The IT provides a good single-hit resolution and very accurate measurements in the horizontal plane, where the charged particles are bent by the dipole magnet.

The Outer Tracker

The Outer Tracker (OT) [57] is located in the outer regions of the tracking stations T_1 , T_2 , and T_3 , surrounding the IT. The OT is based on gas detectors measuring the hit position of charged particles with straw tubes (see Fig. 2.9).

Similar to the IT, each OT tracking station is built with the $x - u - v - x$ layout. There are 18 OT modules, placed symmetrically to each other in the $x - z$ plane. A module contains 128 straw tubes arranged in monolayers, which are staged on top of each other. These monolayers are vertically split into two halves. To avoid having dead zones, the splitting of the monolayers is slightly shifted. A support structure, the so-called C-Frame, is used to carry the detector modules. Figure 2.10 shows a sketch of the OT and the straw tube modules.

The straw tubes are filled with a gas mixture of Ar (70%), CO_2 (28.5%), and O_2 (1.5%). In the middle of a tube, a gold-coated tungsten wire works as an anode connected to a high voltage supply. The cylindrical outside of the tube is the cathode, collecting the charge created by the ionization of the gas. Charged particles induce ionization clusters (see Fig. 2.9). The drift time of these ionization clusters is measured to determine the hit positions of the particles. With this approach, a spatial resolution of $200 \mu\text{m}$ in the $z - y$ plane is achieved.

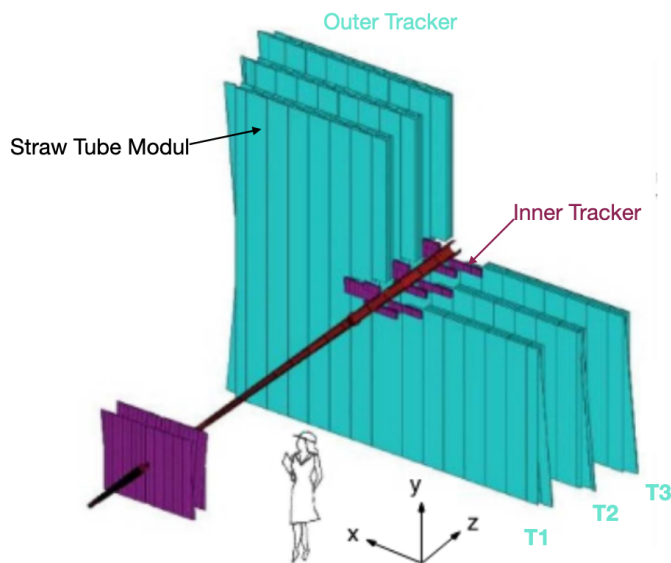


Figure 2.10 – Sketch of the Outer Tracker with the position of the straw tube modules [48].

Track reconstruction

The VELO, the TT, and the tracking stations are used for the reconstruction of the trajectory of a charged particle and the determination of its momentum. However, not all charged particles pass through the full tracking system and can therefore be classified in different categories [55].

Tracks passing through the full tracking system are called **long tracks** and have a very precisely determined momentum. The momentum resolution of tracks passing only through the VELO and the TT is worse. Such tracks exit the detector acceptance after the TT due to the bending by the magnet field and are called **upstream tracks**. There are also tracks not detected by the VELO, the so-called **downstream tracks**, which can be related to particles decaying outside of the VELO, like Λ baryons. Tracks only measured by the VELO, called **VELO tracks**, can often be related to tracks in the backward direction. Additionally, tracks only in the tracking stations T_1 , T_2 , and T_3 , the so-called **T tracks**, can be related to secondary interactions and provide information for particle identification in RICH2 (see Sec. 2.2.2).

2.2.2 Particle identification

The Particle Identification (PID) system, takes the input of several detector components into account, each optimised for the detection of specific particles, which allows for example the distinguishing between kaons and pions. In the following, a short overview of the PID system is given.

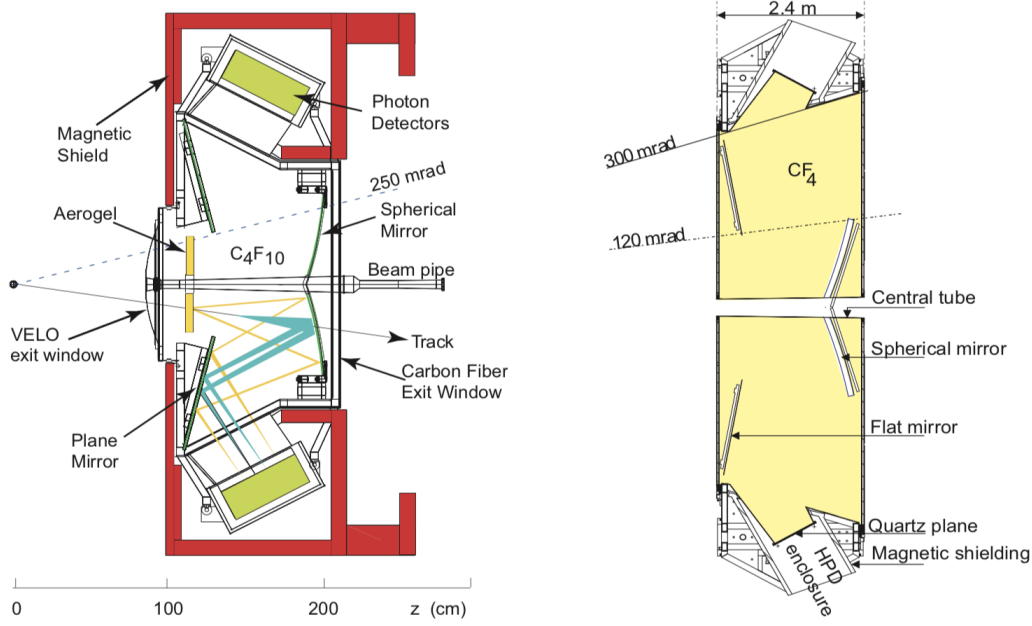


Figure 2.11 – Cross-sections of RICH1 (left) and RICH2 (right) [59].

The RICH detectors

Two Ring Imaging Cherenkov (RICH) detectors are used to identify charged particles. RICH1 is located upstream of the magnet, while RICH2 is placed downstream, behind the tracking stations.

To distinguish between particles, the RICH system uses Cherenkov radiation. When a charged particle traverses a medium with a velocity v , larger than the speed of light c in that medium, so-called Cherenkov light is emitted in a cone with a half-opening angle Θ_c , given by

$$\cos(\Theta_c) = \frac{1}{n \cdot \beta}, \quad (2.1)$$

with $\beta = \frac{v}{c}$ and n the refraction index of the radiator medium. The measurement of Θ_c allows the determination of the particle speed. Together with the information on the particle momentum, provided by the tracking system, the particle mass can be determined and the particle can be identified.

For the detection of the Cherenkov light Hybrid Photon Detectors (HPDs) are used, in both RICH detectors. The emitted Cherenkov photons are focused and guided via a mirror system to the HPDs, which are placed outside of the angular acceptance of the RICH system. RICH1 is optimized to identify particles in a momentum range from 1 to 60 GeV/ c and is designed with two radiators, one which uses silica aerogel⁶ and one which uses C4F10 gas. On the other

⁶The silica aerogel radiator has been removed since the early operation days of the experiment.

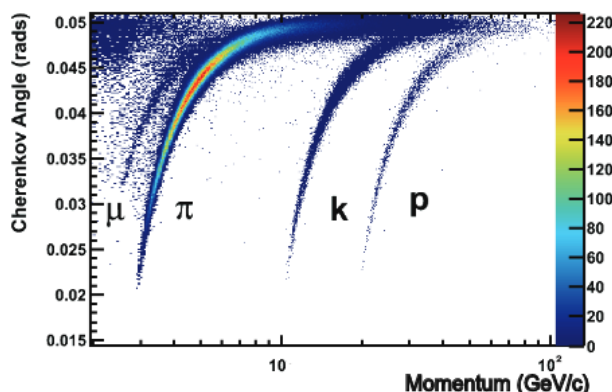


Figure 2.12 – Cherenkov angle for muon (μ), pion (π), kaon (k) and proton (p) tracks as a function of particle momentum [60].

hand, RICH2 uses CF_4 and focuses on a momentum range of 15–100 GeV/c . [60]

The detector layout is displayed in Fig. 2.11. This setup provides an excellent particle identification, as illustrated in Fig. 2.12, where the Cherenkov angle for different particle tracks as a function of particle momentum is displayed. A detailed description of the RICH system can be found in Ref. [60].

The calorimeters

The calorimeter system is used for the identification of electrons, photons and hadrons, and for the measurement of the corresponding particle energies and positions. Furthermore, hadron, electron and photon candidates are selected, based on their transverse energy for the first trigger level (see Sec. 2.2.3).

The calorimeter system is comprised of an electromagnetic calorimeter (ECAL), detecting electrons and photons, and a hadron calorimeter (HCAL). The identification of electrons is especially challenging, due to large background contributions occurring from charged pions. Therefore, the electromagnetic shower is detected in longitudinal segments, using a pre-shower detector (PS), placed in front of the ECAL. Furthermore, the scintillator pad detector (SPD), placed in front of the PS, is used to reject photons, by selecting charged particles. In between the SPD and PS a thin lead converter is placed. [48]

All four components of the calorimeter are based on the same principle, using a structure of alternating planes with scintillators and absorbers. Incident particles interact with the absorber, triggering a shower of secondary particles, which interact with the scintillators. The scintillating photons are guided with wavelength shifting fibres to photomultiplier tubes.

A lateral segmentation is used for the SPD/PS, the ECAL, and the HCAL to cope with the

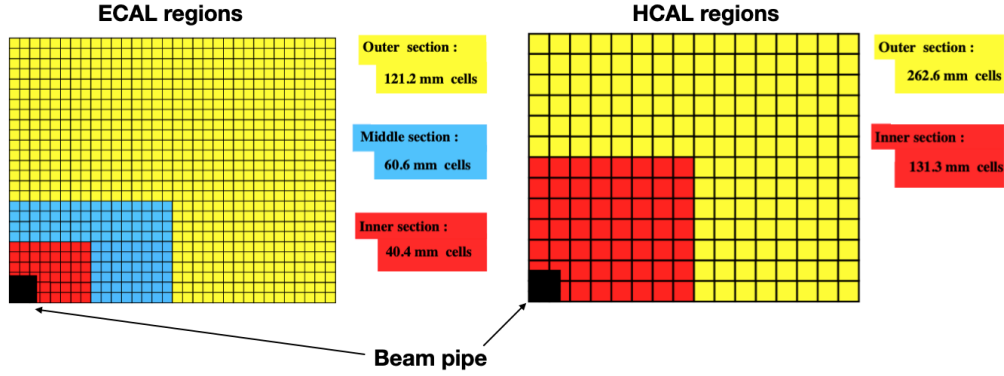


Figure 2.13 – One quarter of the ECAL and HCAL front faces [61]. Left: three ECAL regions with different cell sizes. Right: two HCAL regions with different cell sizes.

variation in the hit density by two orders of magnitude. The ECAL is segmented into three different cell size zones. The cell size of the innermost region is chosen to be close to the Molière radius⁷, ensuring that the majority of the deposited energy is contained in a group of four cells [61]. The HCAL cells are chosen to be larger than the ECAL cells, given the larger dimensions of the hadronic showers. Furthermore, a lateral segmentation into two HCAL zones is chosen. The ECAL and HCAL calorimeter regions are illustrated in Fig. 2.13.

The ECAL provides a relative resolution of $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\%$ and the HCAL $\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\%$ [61], where the energy E is in GeV.

The muon detectors

The muon detection system provides accurate identification of muons, as well as a precise measurement of the muon momentum at a very high particle rate. Furthermore, it supplies information for the hardware-level trigger.

The muon system is comprised of five muon stations M1–M5, where M1 is located downstream before the PS, while M2–M5 are placed behind the calorimeters. M1–M3 provide a very high spatial resolution in the bending plane, which allows an accurate determination of the particle track direction and the measurement of the transverse momentum. M4 and M5 are optimised for the detection of penetrating particles.

The muon stations M2–M5 are based on Multi-Wire Proportional Chambers (MWPC), operating with a gas mixture of 45% Ar, 15% CO₂ and 40% CF₄. The M1 station is built with Gas Electron Multiplier detectors (GEMs) to deal with a large occupancy due to the high particle flux. Three muon filters (iron shields) are placed in between M2–M4. In addition, a fourth iron shield is placed directly behind the last muon station. The muon stations are sub-divided into four regions with different dimensions of the logical pad, to account for the variation in the

⁷The Molière radius is a constant, which describes the transversal dimension of an electromagnetic shower.

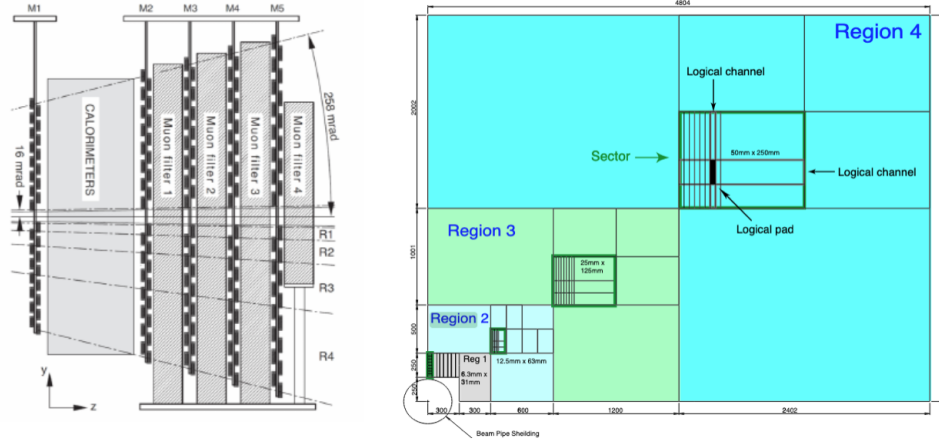


Figure 2.14 – Left: cross-section of the muon system. Right: illustration of one quarter of the front face of a muon station, with the different detection regions and logical pad sizes [62].

particle flux observed between the inner and outer regions of the detector. The dimensions of the detection regions (R1–R4) and the logical pad size scale by a factor of two, from one region to the next. [62]

The design of the muon system is shown in Fig. 2.14 (left) and the muon station layout is illustrated in Fig. 2.14 (right).

The muon system provides an angular acceptance of 20–306 mrad in the bending plane and 16–258 mrad in the non-bending plane. 20% of all muons from semi-leptonic b -hadron decays are detected by the system.

Particle identification variables

Specific algorithms are designed to identify the nature of each particle in the detector acceptance, by taking into account several parameters, *e.g.*, particle momentum or pseudorapidity. The main PID variables are:

- **ProbNN**: response of a Neural Network (NN) trained to identify specific particles [63]. For example, ProbNNe is close to one for a track produced by an electron a close to zero in the opposite case.
- **Delta-Log-Likelihood (ΔLL)**: variable computed as the natural logarithm of the ratio of likelihoods for a given track, to be compatible with two particle hypotheses, h and h'

$$\Delta LL = \ln \frac{L(t|h)}{L(t|h')}, \quad (2.2)$$

where t denotes the set of variables describing the particle track [63]. h' is usually set to the pion hypothesis, while for each possible h a specific ΔLL variable is defined [63]. In

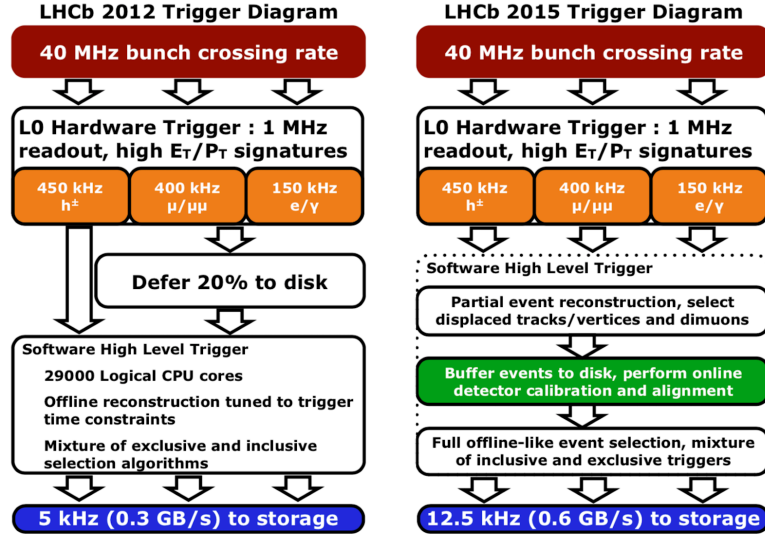


Figure 2.15 – Schematic drawing of the trigger flow [65]. Left: Run 1. Right: Run 2.

the following, the ΔLL is denoted as PID, *e.g.*, PIDE for $h = e$ and $h' = \pi$.

- **IsMuon**: boolean variable indicating whether or not hits are detected in the muon stations that can be matched to hits detected in the tracking stations [63].

Requirements placed on the different PID variables are usually specifically chosen for a data analysis. In this search for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$, requirements are imposed on the ProbnNe, ProbnNmu and PIDE variables.

2.2.3 The trigger system

The trigger system [64] is crucial for the separation of interesting physics events (signal) from background events. Signal candidates are events that fulfill a specific selection, the so-called trigger lines. The trigger selection is analysis specific and takes into account decay-specific features, *e.g.*, transverse momentum p_T , or the flight distance. The system is composed of a hardware-level trigger and two software-level triggers.

The hardware-based **Level-0 (L0)** trigger is the first step in the trigger selection chain operating at the LHC bunch-crossing rate of 40 MHz. It reduces the event rate significantly toward 1 MHz before the full detector data is read out. For the L0 selection, information from the calorimeter and the muon system is used. In addition, the pile-up system of the VELO, which uses the R-sensors is taken into account. The L0 trigger selects the electron and hadron candidates with the highest transverse energy E_T among all recorded clusters of an event in the calorimeter system. Furthermore, the muon candidate with the highest transverse momentum p_T relative to all other reconstructed muon candidates in an event is chosen. [64]

The second step of the trigger cascade is the **High-Level Trigger (HLT)**, which is completely

software-based and runs on a computer farm. The HLT system reduces the event rate to a few kHz and can be divided into two sub-systems. **HLT1** takes the information provided by all detector components into account. Events are partially reconstructed in the VELO and tracks that can be associated with a muon or a high impact parameter with respect to the primary vertex (IP) are extrapolated further into the tracking system. **HLT2** selects tracks with a $p_T > 300 \text{ MeV}/c$ and fully reconstructs these events. It uses a Boosted Decision Tree for the event reconstruction and is optimized for specific physics processes of interest. Only events passing the full trigger selection are saved, while the remaining events are discarded. [64]

Events selected by the trigger and containing a signal candidate of interest can be assigned to three different categories:

- **Trigger On Signal (TOS)** are events that still pass the trigger selection if everything is removed, except the signal candidate.
- **Triggered Independently of Signal (TIS)** are events that still pass the trigger selection if the signal candidate is removed.
- **Trigger On Both (TOB)** are events where the presence of the signal alone or the rest of the event alone is not enough for passing the trigger selection, instead both components are necessary for a positive trigger decision.

Figure 2.15 shows the full trigger flow for Run 1 and Run 2. In the case of Run 2 new real-time alignment and calibration are applied, which have improved the trigger selection, compared to Run 1.

2.2.4 Stripping

After the full event reconstruction, a first central offline selection is applied, called the stripping. The stripping contains analysis-specific selection requirements, which are organised in so-called stripping lines. The stripping output is stored on disk and accessible for further target-specific analyses.

Generally, there are 12 stripping streams, which group the stripping selections into physics categories, *e.g.*, the dimuon stripping stream is used to select two muon candidates.

2.2.5 Electron reconstruction

The reconstruction of electrons plays an important role in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ search. Compared with other charged particles, electrons are very light, which leads to strong emission of bremsstrahlung. In the present analysis, the energy loss due to bremsstrahlung affects significantly important quantities like mass shapes and selection efficiencies. Therefore, it is important to take these effects into account in the electron reconstruction.

Considering an electron candidate, there are in general two bremsstrahlung scenarios one

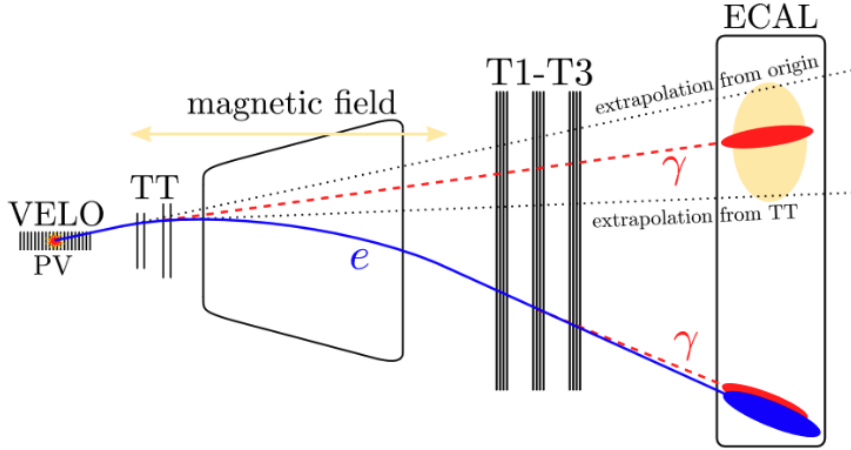


Figure 2.16 – The emission of bremsstrahlung photons (red dashed lines) before and after the magnet is illustrated [66].

has to consider, illustrated in Fig. 2.16. Firstly, an electron emits a photon after the magnet. In that case, the trajectory of the photon will be close to the trajectory of the electron. Both particles will hit the calorimeter in the same region so that the photon and the electron will be reconstructed as one single cluster. In the second case, the electron emits a photon before the magnet. While the electron will be bent by the magnetic field, the photon trajectory is not affected and the two particles will hit the ECAL in different regions. Therefore, the electron and the corresponding bremsstrahlung photon can not be treated as a single object. The particle reconstruction needs to take care of this by searching for photon clusters in ECAL which can be associated with the electron candidate. However, this can lead to the case that no photon cluster in the ECAL could be found while the electron actually emitted bremsstrahlung. Vice versa, a wrongly identified photon cluster could be assigned to an electron candidate which did not emit a bremsstrahlung photon.

To deal with these issues, the analysis can be performed in two categories, depending on whether or not a bremsstrahlung photon can be associated with the electron candidate.

3 The LHCb Run 3 upgrade

The amount of collected data by LHCb during Run 1 and Run 2 was limited mainly by the readout rate of the detector. For the third LHC data-taking period, Run 3 from 2022 to 2025, the LHCb detector has undergone a major upgrade in view of operation with an increased instantaneous luminosity up to $2 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ and a centre-of-mass energy of 13.6 TeV, while improving the readout rate. The increase in luminosity results, together with expected higher cross-sections, in a drastic rise of interactions per bunch crossing, which adds difficulty to isolate interesting events from the rest of the events, the so-called pile-up. Furthermore, the higher occupancy and the increasing damage caused by radiation needs to be taken into account as well. The LHCb upgrade is designed to collect in the next 10 years up to 50 fb^{-1} of data, which is five times as much as in the past [67].

To improve the data-taking performance and cope with the challenging Run 3 conditions, the hardware is upgraded to allow the readout of all detector components at 40 MHz. Consequently, several sub-detectors were fully replaced, using different technologies compared to the previous setup. The tracking stations for example have been replaced by a new detector using scintillating fibres, the so-called Scintillating Fibre (SciFi) tracker. The LHCb group at the École polytechnique fédérale de Lausanne (EPFL) has made a major contribution to the SciFi tracker project, which includes work performed within this thesis. Furthermore, the current L0 hardware trigger system has been removed and replaced by a fully software-based trigger system.

This chapter provides an overview of the upgrade, with emphasis on the SciFi tracker, as well as a summary of the contribution of this thesis to the SciFi tracker construction.

3.1 Run 3 detector upgrade

The layout of the upgraded detector is displayed in Fig. 3.1. The silicon micro-strip detectors of the VELO are replaced with silicon pixel detectors, which offer a high granularity and an improvement in the pattern recognition. Furthermore, the innermost acceptance region of

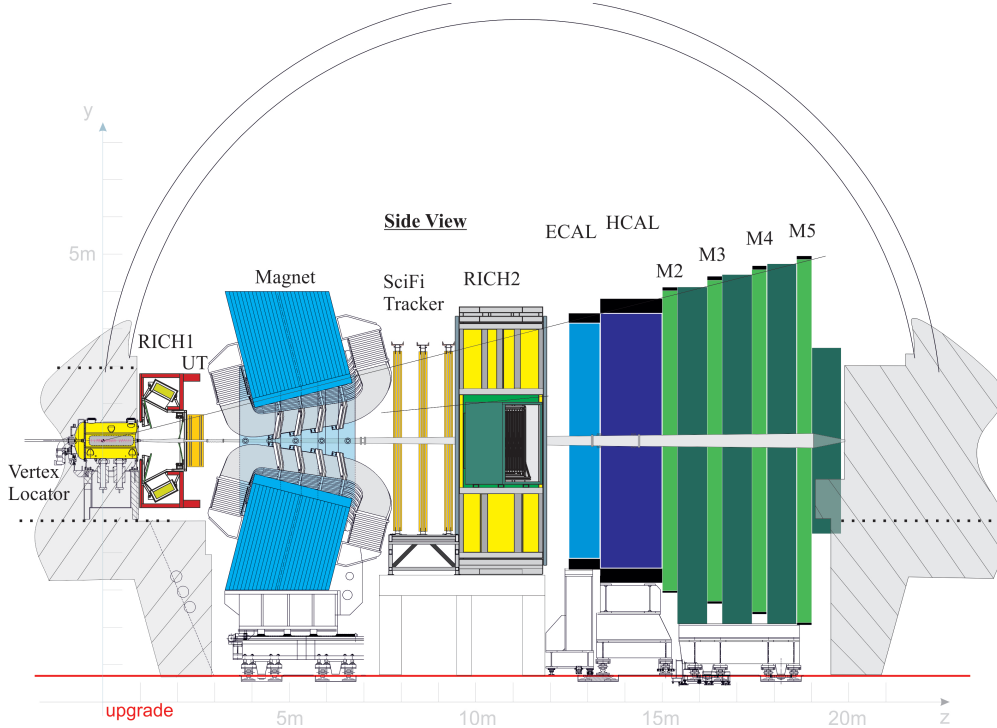


Figure 3.1 – LHCb detector layout for Run 3 [69].

the VELO is improved, placing detectors closer to the beam pipe (5.1 mm, instead of 8.1 mm previously). A detailed description of the VELO upgrade can be found in Ref. [68].

As explained in Ref. [70], the TT was completely replaced by the Upstream Tracker (UT), which is based on four layers of silicon micro-strip detectors. The new detectors provide a higher granularity compared to the previous system. The inner detector sensors have a very fine strip pitch of $95\ \mu\text{m}$, which is half the size of the strip pitch of outer sensors.

The IT and OT are both replaced with the SciFi tracker. The new detector does not have any passive matter like cooling pipes in the acceptance of the detector. A further advantage of this new setup is the reduction of the processing time, since instead of evaluating the IT and OT, only a single detector needs to be read out. A more detailed description of the SciFi detector is given in Sec. 3.3.

The particle identification system plays a crucial role for the software trigger and also went through several changes. While the structure of the RICH system remained mainly the same, the RICH1 mirrors have been replaced by mirrors with a larger focal length, providing a better reflectivity, which helps handling the rise in occupancy. For a faster readout, the HPDs have been replaced by multi-anode photomultiplier tubes. The calorimeter system stayed mostly unchanged. However, to cope with the 40 MHz readout, the front-end and back-end electronics had to be upgraded. 238 new front-end boards are used, each designed to read out 32 photomultiplier tubes [71]. The SPD and the PS are completely removed, as well as the

first muon station M1. These components provided mainly information for the L0 hardware trigger, which will no longer exist in Run 3. The full upgrade of the PID system is described in Ref. [71].

3.2 Run 3 trigger upgrade

With the Run 3 upgrade new challenges arise, which have to be taken into account to prevent the experiment from high efficiency losses. The previous experimental setup was strongly limited by the readout rate of the L0 hardware trigger (approximately 1 MHz). Therefore, the main goal of the trigger upgrade was the replacement of the current trigger system with a full software based trigger system, running on an Event Filter Farm (EFF). [64]

By taking into account the hardware upgrades explained in Sec. 3.1, the tracking sequence can be reduced dramatically. For example, the full HLT tracking sequence, taking the SciFi upgrade into account, takes less than 10 ms, which does match perfectly with the maximum event processing time of the EFF of 13 ms [64]. This allows to perform the full tracking sequence within the HLT system.

3.3 The SciFi tracker upgrade

3.3.1 Detector design

The SciFi tracker is based on scintillating fibres acting as active detection material. Charged particles crossing the fibres deposit energy in form of light pulses, which are transported to the photodetectors at the end of the fibres.

The SciFi tracker is located between the magnet and the RICH2 (see Fig. 3.1) and built of three tracking stations, where each station is composed of two movable halves (left and right). Each tracking station has four detection layers of vertical fibres, oriented in the $x - u - v - x$ geometry, where the two midlayers are rotated by a $\pm 5^\circ$ stereo angle. Each detection layer contains 12 modules. The 5 m high and 0.52 m wide modules are made from eight mats of scintillating fibres. Figure 3.2 shows the three tracking stations and the arrangement of the detection layers. The scintillating fibres have a round cross section with a diameter of $250 \mu\text{m}$ and are based on a polystyrene core, doped with p-terphenyl (PT) and tetraphenyl-butadiene (TPB). They provide a very high light yield and a fast decay constant of 2.8 ns, which makes them very suitable for the application in the LHCb SciFi tracker. The reflection indices are decreasing from the core towards the outside, which optimises the guidance of the scintillating light by total reflection in the fibre. The 242.4 cm long and 13.5 cm wide fibre mats contain six layers of scintillating fibres glued together. Schematic drawings of a module and of a fibre are displayed in Fig. 3.3.

A mirror is glued at the inner end of each module, reflecting inward travelling photons towards

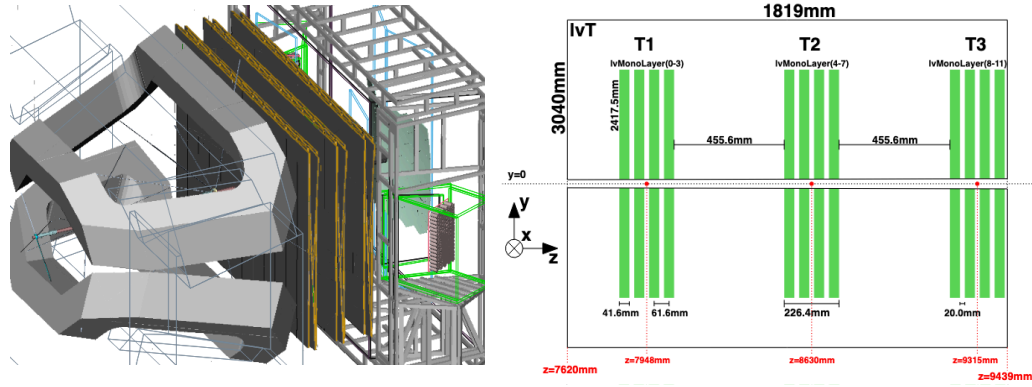


Figure 3.2 – Left: the SciFi tracking stations, placed between the magnet and the RICH2 [70]. Right: layout of the detection layers [70].

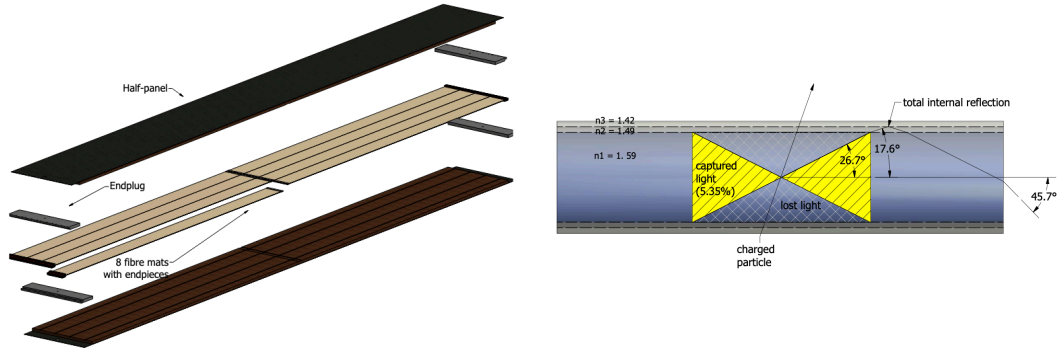


Figure 3.3 – Left: full SciFi detector module with 8 fibre mats [72]. Right: side view of a single fibre, illustrated with a crossing particle [72].

the outer end of the module, where four multichannel Silicon Photo Multipliers (SiPMs) with a channel pitch of $250\ \mu\text{m}$ are mounted. The light pulses in the fibres, generated by crossing particles, are detected by the SiPMs, which are operating in cold boxes at -40°C to mitigate damage by radiation. More details of the SciFi tracker design can be found in Ref. [70].

3.3.2 SiPMs for the SciFi tracker

The SiPMs used for the SciFi tracker are multichannel photo-detectors based on matrices of avalanche photo diodes (APDs). The SiPMs are composed of two silicon dies each with 64 channels. Each channel has a width of $250\ \mu\text{m}$ and 104 pixels of APDs. The total active area is $(32.54 \times 1.625)\ \text{mm}^2$ and does not include the $220\ \mu\text{m}$ gap between the dies. A $105\ \mu\text{m}$ thick epoxy protection layer is applied on top of the SiPM surface. A stiffener plate, as well as a temperature sensor are mounted on the backside of the device, which is connected with a Kapton flex cable to the readout electronics. A detailed view of an SiPM assembly and of the silicon dies is given in Fig. 3.4. More details about the SiPM design can be found in Ref. [73].

Each pixel, with a surface of $(57.5 \times 62.5)\ \mu\text{m}^2$, is comprised of a quenching resistor R_Q connected in series with a diode working in the Geiger-Müller regime (GM-APDs). The pixels are connected in parallel (see Fig. 3.5 left). Incoming photons, from the scintillating fibres, produce electron-hole pairs, which can trigger electron avalanches and make the diodes conductive. If the voltage applied to the diode is above the so-called breakdown voltage V_{BD} , the triggered avalanche is self-sustaining. In the initial state, the diodes are operated at a voltage $V_{bias} > V_{BD}$ and are not conductive. They behave basically like a charged capacitor (see Fig. 3.5 right). The triggered electron avalanche makes the diode conductive and the capacitor discharges. The resulting electric current decreases exponentially and is limited by R_Q . The voltage on the diode drops towards V_{BD} . Once the current drops below $10\text{--}30\ \mu\text{A}$ the avalanche is terminated and the diode is not conductive anymore. The time it takes until the diode is recharged at a voltage of V_{bias} is known as the recovery time τ_{rec} [74]. The recovery time of a pixel is approximately 70 ns and the breakdown voltage is approximately 50 V at 25°C . The breakdown voltage decreases non-linearly with decreasing temperature. The difference $\Delta V = V_{bias} - V_{BD}$ is called overvoltage and plays an important role for several properties, which were studied intensively at EPFL [73]. An example is the Dark Count Rate (DCR),

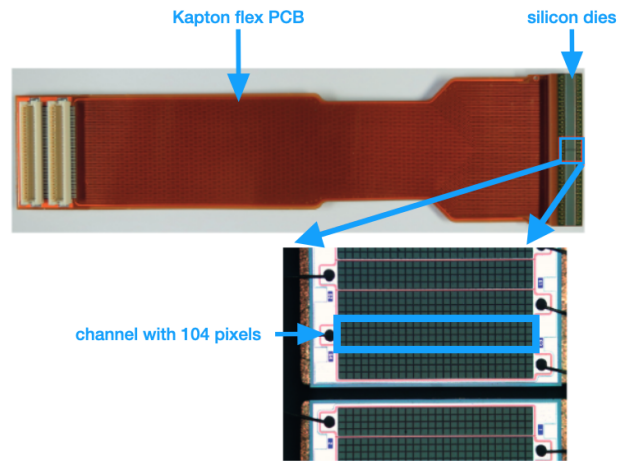


Figure 3.4 – Picture of an SiPM assembly and zoom into the silicon dies [73].

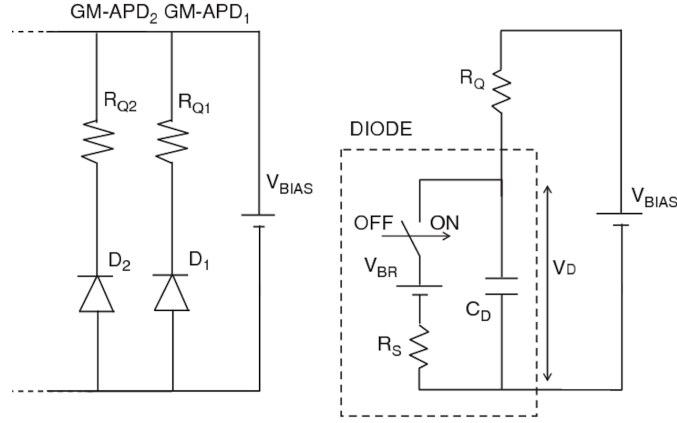


Figure 3.5 – Left: circuit diagram of pixels connected in parallel [74]. Right: circuit diagram of a pixel [74].

i.e. the rate of random thermally generated electron-hole pairs that trigger an avalanche in the pixel. The DCR increases with the overvoltage and the temperature. Another example is pixel-to-pixel cross-talk, where photons generated by the avalanche of a triggered pixel fire a direct or delayed avalanche in a neighbouring pixel, causing correlated noise. Furthermore, trapped charged carriers from a generated avalanche can trigger a secondary avalanche in the same pixel, known as After Pulse (AP). A further challenge is the damage caused by radiation. Intense studies were done at EPFL [73] to understand these effects, which are mitigated by operating the devices in the final detector at a temperature of -40°C .

The SiPMs used for the LHCb SciFi tracker are produced by Hamamatsu. In total 5500 detectors were produced, while 3840 are needed for the experiment. The remaining good detectors are preserved as spare parts.

3.3.3 SiPM quality assurance

Each fibre mat of the final detector is attached to four SiPMs, which are mounted in a cold box with a common bias voltage supply. For the optimal operation of the devices it is therefore crucial that the detectors in each cold box are selected to share common properties. On one hand it is important that the entire active surface of the four SiPMs is as smooth as possible to minimise light losses in the interface region. On the other hand the common voltage supply must be taken into account. Only devices with a similar breakdown voltage are grouped together. Differences up to 500 mV can be compensated. Furthermore, the functionality of the devices need to be checked, *e.g.*, SiPMs with dead channels, which give no signal, need to be identified. In order to ensure the quality of all SiPM detectors used for the SciFi tracker, each device went through several tests at EPFL.

As a first step, the SiPMs went through a full electrical characterisation with a low light amplitude spectrum analysis to measure V_{BD} . Eight SiPMs are measured simultaneously in an

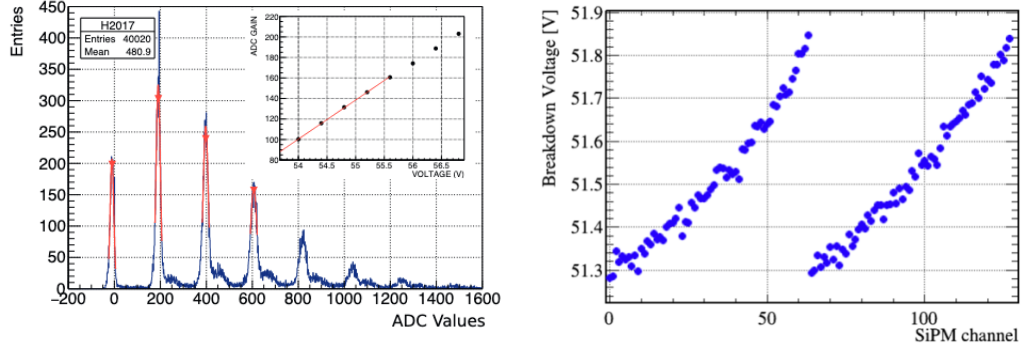


Figure 3.6 – Left: illustration of the determination of the breakdown voltage. Recorded light pulses are fitted with Gaussian distributions and the resulting gain for different overvoltages is linearly approximated [73]. Right: Measured breakdown voltages for all channels of one SiPM.

electrically and thermally insulated box. A heat spreader is installed in the box, which cools the devices to -25°C and a bar is used to inject fast light pulses, which are detected by the devices while scanning over a range of different overvoltages. The readout electronics contains an analogue-to-digital converter (ADC). The recorded ADC pulses are fitted with Gaussian distributions (see Fig. 3.6 left). The ADC gain, which is proportional to ΔV is determined by the distance of the mean of the Gaussian distributions. The obtained gain is then linearly approximated and V_{BD} is determined as the voltage when the gain is zero. This method allows a very fast and quite accurate measurement of V_{BD} for each channel of the SiPM (see Fig. 3.6 right). Furthermore, it reveals non-working channels. In addition, the temperature sensor on the backside of the SiPMs is monitored during the light injection, to ensure that the devices are fully functional.

The second step of the SiPM quality assurance procedure is a thickness measurement of the epoxy surface via optical focusing under a microscope. The thickness is measured for all devices at both outer edges and in the centre of the device. The central region is observed to be a bit thicker than the outer regions, resulting in a banana-like shape (see Fig. 3.7 left). The mean deviation in thickness is evaluated and devices with a deviation larger than $100\text{ }\mu\text{m}$ are discarded.

Finally, the active surface of the SiPMs is inspected optically under a microscope. Devices with large impurities, scratches and damages (see Fig. 3.7 right) are discarded. After the full inspection and characterisation, groups of four SiPMs with similar V_{BD} and similar deviation in thickness are defined.

Almost the full first year of this PhD was dedicated to the SiPM testing. All devices were characterized and tested. Only 1.5 % were discarded due to optical imperfections and 0.6 % were discarded due to electrical issues.

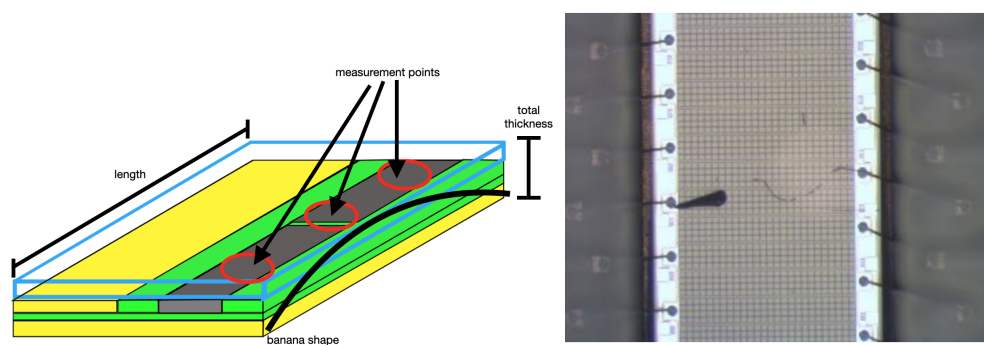


Figure 3.7 – Left: illustration of the SiPM surface measurement. The blue rectangle represents the epoxy layer. Right: SiPM surface with a large impurity.

4 Search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$

4.1 Analysis strategy

The goal of the analysis is the search for the $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ decays, by measuring (or setting upper limits on) the corresponding branching fractions. In the following, the notation $B_{(s)}^0$ is used to either refer to the B^0 or the B_s^0 mesons, charge-conjugated modes are implied throughout.

Experimentally, the $B_{(s)}^0 \rightarrow e^+ \mu^-$ and $B_{(s)}^0 \rightarrow e^- \mu^+$ decay modes cannot be distinguished in the untagged and time-integrated analysis presented here. The experimental efficiencies are assumed to be the same for $B_{(s)}^0 \rightarrow e^+ \mu^-$ and $B_{(s)}^0 \rightarrow e^- \mu^+$ and the final branching fractions of interest are defined as

$$\mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) = \mathcal{B}(B_{(s)}^0 \rightarrow e^+ \mu^-) + \mathcal{B}(B_{(s)}^0 \rightarrow e^- \mu^+). \quad (4.1)$$

The search for the $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ decays is performed, using the high-statistics $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ decay mode as normalisation channel, by evaluating the corresponding branching fractions as

$$\begin{aligned} \mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) &= \mathcal{B}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+) \cdot \frac{f_u}{f_{d(s)}} \\ &\cdot \frac{\varepsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)}{\varepsilon(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)} \cdot \frac{\mathcal{N}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)}{\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)}. \end{aligned} \quad (4.2)$$

The $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ normalisation branching fraction is equal to

$$\mathcal{B}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+) = \mathcal{B}(B^+ \rightarrow J/\psi K^+) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-), \quad (4.3)$$

where the values are taken from the Particle Data Group (PDG) [13]. The fragmentation fractions, $f_{u,d,s}$, describe the probability that a \bar{b} quark forms a weakly-decaying bound

state with an up, down, or strange quark. The ratio f_s/f_u has been measured by the LHCb collaboration at 13 TeV [75], while f_u/f_d is assumed to be one.

The signal yields, $\mathcal{N}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)$, as well as the normalisation yield, $\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)$, are determined with invariant mass fits to the corresponding data samples, collected with the LHCb experiment between 2016 and 2018. A dedicated selection, designed to reduce the background pollution in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data samples, is applied to the real data, as well as to Monte Carlo (MC) simulation:

- firstly, loose selection requirements are applied to the samples in the stripping-, trigger-, and pre-selection (see Sec. 4.3);
- secondly, particle identification (PID) requirements are applied to reduce the pollution from specific physics decays (see Sec. 4.3.6);
- thirdly, a Boosted Decision Tree (BDT) is trained and applied to the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data samples, to reduce the pollution from combinatorial background (see Sec. 4.3.7).

The selection efficiencies ε entering Eq.(4.2) are determined from simulation, taking data-driven corrections into account, which are introduced in Sec. 4.5 to take care of effects arising from the imperfect modelling in simulation.

In order to demonstrate that the corrected efficiencies are properly estimated, a cross-check is performed. Since there is no suitable decay mode with an electron and muon in the final state, the cross-check is performed using the well-known $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ decays. This approach allows to test the obtained selection efficiency corrections for electrons and muons, by measuring the branching fraction ratio:

$$\begin{aligned} r_{J/\psi} &= \frac{\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)}{\mathcal{B}(J/\psi \rightarrow e^+ e^-)} = \frac{\mathcal{B}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)}{\mathcal{B}(B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+)} \\ &= \frac{\varepsilon(B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+)}{\varepsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)} \cdot \frac{\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)}{\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+)}. \end{aligned} \quad (4.4)$$

The yields $\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+)$ and $\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)$ are obtained with a fit to the invariant B^+ mass distributions. The selection efficiencies ε are corrected, using the same methods as for the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal channels. With all corrections taken into account, the measured result for $r_{J/\psi}$ is expected to agree with the PDG value of 0.9983 ± 0.0078 [13].

One of the biggest challenges of this analysis is the bremsstrahlung recovery of the electron in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ final state. The final state electron emits bremsstrahlung as illustrated in Sec. 2.2.5. A dedicated algorithm is applied to search for clusters in the electromagnetic calorimeter from bremsstrahlung photons, which can be associated to the emitting electron track. Thus the total deposited energy in the electromagnetic calorimeter by the electron is given by the sum of the deposited energies associated with the electron track and the bremsstrahlung photons, $E_{\text{total}}(e) = E_{\text{ECAL}}(e) + \sum_{j=1}^{N_\gamma} E_{\text{ECAL}}^j(\gamma)$. However, the bremsstrahlung recovery is imperfect, meaning bremsstrahlung clusters may not be found for an electron

track or that wrong clusters may be associated to the track. This not only affects the shapes of the invariant mass distributions, but also the selection efficiencies. Therefore, the analysis is performed in bremsstrahlung categories, considering different bremsstrahlung recovery cases for the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal and the $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ control channel:

- **0 γ** : no bremsstrahlung is recovered ($B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$);
- **1 γ** : bremsstrahlung is recovered for one electron track ($B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$);
- **2 γ** : bremsstrahlung is recovered for two electron tracks (only $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$).

Two main background sources are considered: random electron-muon track combinations, so-called combinatorial background, and specific physics background. The later is estimated from corrected simulation. Two data-driven methods are used to test these background predictions using for example $B_{(s)}^0 \rightarrow hh'$ decays, with $h = \pi^+, K^+, p$ and $h' = \pi^-, K^-, \bar{p}$.

In the final steps of this analysis, the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions are determined with a simultaneous fit to the invariant $B_{(s)}^0$ mass distribution in several fit samples, defined by the bremsstrahlung categories, and regions of the Boosted Decision Tree (BDT) response (see Sec. 4.7). Furthermore, an upper limit for each branching fraction will be determined with the CL_s method [76]. Until all steps and methods of this analysis are finalised and validated, the analysis is performed with a blinded signal window, *i.e.* $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal events in the $B_{(s)}^0$ mass region between 5.1 and 5.6 GeV/ c^2 are removed, to avoid bias introduced by the experimenter.

4.2 Data and simulation samples

The data samples were collected during Run 2 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV in 2016, 2017 and 2018 and correspond to integrated luminosities of 1.6, 1.7 and 2.1 fb⁻¹, respectively, for a total of 5.4 fb⁻¹. The data sample collected in 2015 is excluded in this analysis, due to the low statistics which corresponds to an integrated luminosity of only 0.3 fb⁻¹. During each data-taking year, the polarity of dipole magnetic field was alternated.

Monte Carlo (MC) simulated events are used for several stages of this analysis, *e.g.* the training of a multivariate classifier, the determination of selection efficiencies, or the study of possible background sources. Therefore, simulation samples for several decay modes are processed for each data-taking year configuration (see Table 4.1). The MC samples are generated for both magnet polarities with approximately equal proportions. Furthermore, the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal MC samples are generated with an equal probability for the two final states $B_{(s)}^0 \rightarrow e^+ \mu^-$ and $B_{(s)}^0 \rightarrow e^- \mu^+$.

Table 4.1 – List of simulated samples. The number of generated events is the sum over the data-taking periods 2016, 2017 and 2018.

| Decay | Generated events [10^6] |
|--|-----------------------------|
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 3 |
| $B^0 \rightarrow e^\pm \mu^\mp$ | 6 |
| $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ | 13 |
| $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ | 25 |
| $B^0 \rightarrow K^+ \pi^-$ | 12 |
| $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ | 15 |
| $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ | 3 |
| $B^0 \rightarrow \pi^+ \pi^-$ | 12 |
| $B^0 \rightarrow p \bar{p}$ | 6 |
| $B_s^0 \rightarrow \pi^+ K^-$ | 12 |
| $B_s^0 \rightarrow K^+ K^-$ | 12 |
| $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \pi^+$ | 4 |
| $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) \pi^+$ | 5 |
| $\Lambda_b^0 \rightarrow p K^-$ | 6 |
| $\Lambda_b^0 \rightarrow p \pi^-$ | 6 |
| $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ | 12 |
| $\Lambda_b^0 \rightarrow p e^- \bar{\nu}_e$ | 3 |

4.3 Event selection

Selection requirements are designed to reduce the background pollution as much as possible. The selection proceeds in the following five stages:

- firstly, decays of interest are selected by applying a loose selection in the stripping;
- secondly, loose offline requirements are applied to ensure that the analysis working point is aligned to the calibration samples used for data-driven corrections;
- dedicated trigger requirements are chosen to refine the selection;
- particle identification requirements are chosen to reduce the pollution from specific physics backgrounds;
- a multivariate analysis (MVA) is performed to remove combinatorial background.

In the following, each selection step is explained in detail.

4.3.1 Stripping selection

Dedicated stripping algorithms (see Sec. 2.2.4) are used to select the signal and control channel candidates. The following stripping lines are used:

1. $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$: StrippingLFVB2eMuLine
2. $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$: StrippingBs2MuMuLinesBu2JPsiKLine

3. $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$: StrippingLFVBu2KJPseeLine
4. $B_{(s)}^0 \rightarrow hh'$: StrippingBs2MuMuLinesNoMuIDLine

and the corresponding requirements are listed in the Tables 4.2–4.5.

The requirements are applied to the following variables:

- **track χ^2/ndf** : χ^2 of the track fit divided by the number of degrees of freedom;
- **IP χ^2** : χ^2 difference of the primary vertex fit with and without the B -meson candidate;
- **DOCA**: distance of closest approach between two tracks;
- **TrackGhostProb**: probability that a track is reconstructed from random hits (fake track);
- **IsMuon**: track can be associated to a muon candidate (see Sec. 2.2.2);
- **InMuonAcc**: particle track in muon acceptance;
- **PIDe**: Delta-Log-Likelihood ratio under the electron hypothesis (see Sec. 2.2.2);
- **vertex χ^2** : χ^2 of the vertex fit;
- **PV distance χ^2** : χ^2 of the distance to the primary vertex;
- **m** : invariant mass;
- **DIRA**: cosine of the angle between the momentum of the B candidate and the flight direction defined from the production and decay vertices;
- **p_T** : transverse momentum;
- **IsLong**: requirement to select particles from long tracks (see Sec. 2.2.1); distinguishes particles with respect by their lifetime and ensures here that a track can be associated to a long-lived particle (kaon);
- **τ** : proper decay time.

Table 4.2 – Stripping requirements applied to select $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ decays in StrippingLFVB2eMuLine.

| Variable | Applied to | Requirement |
|-----------------------------|-------------|-------------------------|
| track χ^2/ndf | μ, e | < 3 |
| IP χ^2 | μ, e | > 25 |
| DOCA | μ, e | $< 0.3 \text{ mm}$ |
| p_T | μ, e | $> 250 \text{ MeV}/c$ |
| TrackGhostProb | μ | < 0.3 |
| IsMuon | μ | TRUE |
| PIDe | e | > -2 |
| $ m(e\mu) - m_{B_{(s)}^0} $ | $B_{(s)}^0$ | $< 1.2 \text{ GeV}/c^2$ |
| vertex χ^2 | $B_{(s)}^0$ | < 9 |
| PV distance χ^2 | $B_{(s)}^0$ | > 225 |
| IP χ^2 | $B_{(s)}^0$ | < 25 |
| DIRA | $B_{(s)}^0$ | > 0 |

Chapter 4. Search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$

Table 4.3 – Stripping requirements applied to select $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ decays in Strip-pingBs2MuMuLinesBu2JPsiKLine.

| Variable | Applied to | Requirement |
|----------------------------------|---------------------|-------------------------|
| track χ^2/ndf | μ^+, μ^-, K^+ | < 4 |
| IP χ^2 | μ^+, μ^-, K^+ | > 25 |
| DOCA | μ^+, μ^- | $< 0.3 \text{ mm}$ |
| p_T | μ^+, μ^-, K^+ | $> 250 \text{ MeV}/c$ |
| TrackGhostProb | μ^+, μ^-, K^+ | < 0.4 |
| IsMuon | μ^+, μ^- | TRUE |
| IsLong | K^+ | TRUE |
| $ m(\mu^+ \mu^-) - m_{J/\psi} $ | J/ψ | $< 100 \text{ MeV}/c^2$ |
| vertex χ^2 | J/ψ | < 9 |
| PV distance χ^2 | J/ψ | > 225 |
| DIRA | J/ψ | > 0 |
| $ m(\mu^+ \mu^- K^+) - m_{B^+} $ | B^+ | $< 500 \text{ MeV}/c^2$ |
| vertex χ^2 | B^+ | < 45 |
| IP χ^2 | B^+ | < 25 |

Table 4.4 – Stripping requirements applied to select $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ decays in Strip-pingLFVBu2KJPsieeLine.

| Variable | Applied to | Requirement |
|------------------------------|-----------------|--------------------------|
| track χ^2/ndf | e^+, e^-, K^+ | < 3 |
| IP χ^2 | e^+, e^-, K^+ | > 25 |
| DOCA | e^+, e^- | $< 0.3 \text{ mm}$ |
| p_T | e^+, e^-, K^+ | $> 250 \text{ MeV}/c$ |
| TrackGhostProb | K^+ | < 0.3 |
| PIDe | e^+, e^- | > 2 |
| IsLong | K^+ | TRUE |
| $ m(e^+ e^-) - m_{J/\psi} $ | J/ψ | $< 1000 \text{ MeV}/c^2$ |
| vertex χ^2 | J/ψ | < 9 |
| PV distance χ^2 | J/ψ | > 169 |
| DIRA | J/ψ | > 0 |
| $ m(e^+ e^- K^+) - m_{B^+} $ | B^+ | $< 600 \text{ MeV}/c^2$ |
| vertex χ^2 | B^+ | < 45 |
| IP χ^2 | B^+ | < 25 |

Table 4.5 – $B_{(s)}^0 \rightarrow hh'$ stripping requirements in `StrippingBs2MuMuLinesNoMuIDLine`, used to select $B^0 \rightarrow K^+\pi^-$ decays. The line builds candidates under the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ hypothesis.

| Variable | Applied to | Requirement |
|----------------------------|-------------|--------------------------------|
| track χ^2/ndf | h, h' | < 4 |
| IP χ^2 | h, h' | > 25 |
| DOCA | h, h' | $< 0.3 \text{ mm}$ |
| p_T | h, h' | $\in [0.25, 40] \text{ GeV}/c$ |
| p | h, h' | $< 500 \text{ GeV}/c$ |
| TrackGhostProb | h, h' | < 0.4 |
| InMuonAcc | h, h' | TRUE |
| $ m(hh') - m_{B_{(s)}^0} $ | $B_{(s)}^0$ | $< 500 \text{ MeV}/c^2$ |
| vertex χ^2 | B_0^s | < 9 |
| PV distance χ^2 | $B_{(s)}^0$ | > 225 |
| IP χ^2 | $B_{(s)}^0$ | < 25 |
| DIRA | $B_{(s)}^0$ | > 0 |
| τ | $B_{(s)}^0$ | $< 13.248 \text{ ps}$ |
| p_T | $B_{(s)}^0$ | $> 350 \text{ MeV}/c$ |

4.3.2 Decay-tree fitting

The resolution of the reconstructed invariant masses is improved by using the Decay-Tree Fitter (DTF) package [77], fitting the decay tree of a given decay mode under some constraints. For the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal candidates, the final-state particles are constrained to originate from the same point in space (vertex constraint), while for the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ candidates, the dilepton mass is constrained to the known J/ψ mass and a three-prong vertex constraint is applied. The resulting B -candidate mass is denoted with m_{DTF} in the following.

4.3.3 Truth-matching

To ensure that the particles in simulation are correctly identified, the truth-matching selection is applied in simulation. The reconstructed particles of a decay in simulation are selected only if they can be associated with the corresponding true particles of the true decay chain.

4.3.4 Pre-selection

Pre-selection requirements are applied on top of the stripping. On one hand, they are chosen to ensure that the decay kinematics of the signal and normalisation channels are comparable. Therefore, selection requirements on the following variables are applied:

- τ : proper decay time of the B meson candidate;

Chapter 4. Search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$

Table 4.6 – $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ pre-selection requirements.

| Variable | Applied to | Requirement |
|---------------------------------|-------------|---|
| nSPDHits | | < 450 |
| $m_{\text{DTF}}(e^\pm \mu^\mp)$ | $B_{(s)}^0$ | $\in [4.9, 6.1] \text{ GeV}/c^2$ |
| τ | $B_{(s)}^0$ | < 13.248 ps ($\sim 9 \cdot \tau_{B_s^0}$) |
| IP χ^2 | $B_{(s)}^0$ | < 9 |
| p_T | $B_{(s)}^0$ | > 0.5 GeV/c |
| TrackGhostProb | e | < 0.4 |
| p_T | e | $\in [0.5, 15] \text{ GeV}/c$ |
| p | e | $\in [3, 200] \text{ GeV}/c$ |
| InAccECAL | e | TRUE |
| CALO_region | e | ≥ 0 |
| HasCALO | e | TRUE |
| HasRICH | e | TRUE |
| abs(CALO_ECAL_xyProjection) | e | !(x < 363.6 mm & y < 282.6 mm) |
| p_T | μ | $\in [0.8, 40] \text{ GeV}/c$ |
| p | μ | $\in [3, 200] \text{ GeV}/c$ |
| InMuonAcc | μ | TRUE |
| IsMuon | μ | TRUE |
| HasRICH | μ | TRUE |

Table 4.7 – $B_{(s)}^0 \rightarrow hh'$ pre-selection requirements. $B_{(s)}^0 \rightarrow hh'$ decays are used to build $B^0 \rightarrow K^+ \pi^-$ candidates.

| Variable | Applied to | Requirement |
|----------------|-------------|------------------------------------|
| nSPDHits | | < 450 |
| IP χ^2 | $B_{(s)}^0$ | < 9 |
| log(1-DIRA) | $B_{(s)}^0$ | < -12 |
| DeltaR | $B_{(s)}^0$ | < 2.0 |
| HasCALO | h, h' | TRUE |
| HasRICH | h, h' | TRUE |
| IP χ^2 | h, h' | > 25 |
| $m(K^+ \pi^-)$ | $B_{(s)}^0$ | $\in [5.20, 5.45] \text{ GeV}/c^2$ |

- $m(\ell^+ \ell^-)$: reconstructed dilepton mass (only for $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$);
- $p_{J/\psi}/p_{K^+}$: ratio of the J/ψ and kaon momenta (only for $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$).

On the other hand, it is essential to ensure that the working point of the analysis is in agreement with the calibration samples used for the data-driven corrections of the selection efficiencies (see Sec. 4.5), *i.e.*, the selection should match the selection of the calibration samples. Hence,

Table 4.8 – $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ pre-selection requirements.

| Variable | Applied to | Requirement |
|------------------------------|----------------|---|
| nSPDHits | | < 450 |
| $m_{\text{DTF}}(J/\psi K^+)$ | B^+ | $\in [5.2, 5.6] \text{ GeV}/c^2$ |
| τ | B^+ | < 13.248 ps ($\sim 9 \cdot \tau_{B_s^0}$) |
| $m(\mu^+ \mu^-)$ | J/ψ | $\in [2996.9, 3196.9] \text{ MeV}/c^2$ |
| $p_{J/\psi}/p_{K^+}$ | $J/\psi, K^+$ | $\in [0.4, 19]$ |
| p_T | μ^+, μ^- | $\in [0.8, 40] \text{ GeV}/c$ |
| p | μ^+, μ^- | $\in [3, 200] \text{ GeV}/c$ |
| InMuonAcc | μ^+, μ^- | TRUE |
| IsMuon | μ^+, μ^- | TRUE |
| HasRICH | μ^+, μ^- | TRUE |
| p_T | K^+ | > 0.4 GeV/c |
| p | K^+ | $\in [2, 200] \text{ GeV}/c$ |
| InMuonAcc | K^+ | TRUE |
| HasRICH | K^+ | TRUE |
| HasCALO | K^+ | TRUE |
| TrackGhostProb | K^+ | < 0.4 |
| track χ^2/ndf | K^+ | < 3 |

requirements on the particle momentum and transverse momentum are chosen accordingly, in addition to cuts on the following variables:

- **InAccECAL**: track is in the ECAL acceptance region;
- **CALO_region**: track point to on of the regions in the calorimeter system;
- **HasCALO**: track left hits in the calorimeter;
- **HasRICH**: track left hits in the RICH;
- **CALO_ECALS_xProjection**: x -coordinate of the hits deposited in the ECAL;
- **CALO_ECALS_yProjection**: y -coordinate of the hits deposited in the ECAL;
- **InMuonAcc**: track is in the Muon acceptance region;
- **nSPDHits**: number of hits in the Scintillating Pad Detector.

The pre-selection applied to $B_{(s)}^0 \rightarrow hh'$ data and simulation (see Table 4.7), contains additional requirements to reduce the combinatorial background on:

- **log₁₀(1-DIRA)**: where DIRA is defined in Sec. 4.3.1;
- **DeltaR**: quadratic sum of the difference in pseudorapidity ($\Delta\eta = \eta_e - \eta_\mu$) and azimuthal angle ($\Delta\phi = \phi_e - \phi_\mu$) between the two daughter tracks (e, μ),

$$\text{DeltaR} = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (4.5)$$

In addition, fiducial cuts are introduced to consider for example the ranges of the subsequent

Table 4.9 – $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ pre-selection requirements.

| Variable | Applied to | Requirement |
|------------------------------|---------------|---|
| nSPDHits | | < 450 |
| $m_{\text{DTF}}(J/\psi K^+)$ | B^+ | $\in [5.2, 5.6] \text{ GeV}/c^2$ |
| τ | B^+ | $< 13.248 \text{ ps } (\sim 9 \cdot \tau_{B_s^0})$ |
| $m(e^+ e^-)$ | J/ψ | $\in [2.6, 3.6] \text{ GeV}/c^2$ |
| $p_{J/\psi}/p_{K^+}$ | $J/\psi, K^+$ | $\in [0.4, 19]$ |
| TrackGhostProb | e^+, e^- | < 0.4 |
| p_T | e^+, e^- | $\in [0.5, 15] \text{ GeV}/c$ |
| p | e^+, e^- | $\in [3, 200] \text{ GeV}/c$ |
| InAccECAL | e^+, e^- | TRUE |
| CALO_region | e^+, e^- | ≥ 0 |
| HasCALO | e^+, e^- | TRUE |
| abs(CALO_ECAL_xyProjection) | e^+, e^- | $!(x < 363.6 \text{ mm} \ \& \ y < 282.6 \text{ mm})$ |
| p_T | K^+ | $> 0.4 \text{ GeV}/c$ |
| p | K^+ | $\in [2, 200] \text{ GeV}/c$ |
| InMuonAcc | K^+ | TRUE |
| HasRICH | K^+ | TRUE |
| HasCALO | K^+ | TRUE |
| TrackGhostProb | K^+ | < 0.4 |
| track χ^2/ndf | K^+ | < 3 |

invariant mass fits. The full pre-selection requirements are listed in Tables 4.6–4.9.

4.3.5 Trigger selection

The L0 trigger is chosen to select signal candidates, triggering on a single muon (L0Muon) or a single electron (L0Electron) candidate. A cut on the transverse energy of the electron candidate in the ECAL ensures that the working point is in a region where the efficiency of the L0Electron trigger approaches a plateau (see Sec. 4.5.4). The HLT1 trigger requirements are applied on top of the L0 lines. Tracks are required to trigger either the Hlt1TrackMVADecision line or the Hlt1TrackMuonMVADecision line. Events with one or two good tracks of interest are reconstructed, based on the response of a Multivariate Analysis (MVA). On top of the HLT1 trigger lines, candidates of interest are required to trigger one of two HLT2 trigger lines. The chosen HLT2 trigger lines, Hlt2Topo2BodyDecision and Hlt2TopoMu2BodyDecision, fire for two-body decays. The Hlt2B2HHB2HHDecision line is applied in addition, to select displaced two-body B hadron decays. Only candidates of the type “Trigger On Signal” (TOS) are retained, as explained in Sec. 2.2.3. The full trigger selection for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ is shown in Table 4.10.

The trigger selection for $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ (see Table 4.11) and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$

Table 4.10 – $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ trigger selection.

| Trigger level | Trigger selection | Applied to |
|---------------|---|-------------|
| L0 | L0MuonDecision | μ |
| | OR | |
| | L0ElectronDecision AND L0Calo_ECAL_realET > 3000 MeV | e |
| HLT1 | Hlt1TrackMVADecision | $B_{(s)}^0$ |
| | OR | |
| | Hlt1TrackMuonMVADecision | μ |
| HLT2 | Hlt2TopoMu2BodyDecision | $B_{(s)}^0$ |
| | OR | |
| | Hlt2Topo2BodyDecision | $B_{(s)}^0$ |
| | OR | |
| | Hlt2B2HHB2HHDecision | $B_{(s)}^0$ |

Table 4.11 – $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ trigger selection.

| Trigger level | Trigger selection | Applied to |
|---------------|--------------------------|--------------------|
| L0 | L0MuonDecision | μ^+ or μ^- |
| HLT1 | Hlt1TrackMVADecision | B^+ |
| | OR | |
| | Hlt1TrackMuonMVADecision | μ^+ or μ^- |
| HLT2 | Hlt2TopoMu2BodyDecision | B^+ |
| | OR | |
| | Hlt2Topo2BodyDecision | B^+ |
| | OR | |
| | Hlt2Topo3BodyDecision or | B^+ |
| | OR | |
| | Hlt2TopoMu3BodyDecision | B^+ |

(see Table 4.12) follows the same strategy. However, due to the $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ decay topology, the Hlt2TopoMu3BodyDecision and Hlt2Topo-3BodyDecision are included, triggering on three-body decays. The muon-specific trigger lines are only used for $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and the electron-specific trigger lines are only used for $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$. The Hlt2B2HHB2HHDecision line is excluded for the control channels. The trigger strategy applied to $B_{(s)}^0 \rightarrow hh'$ is chosen specifically for the method used to validate the corrected misidentification rates in Sec. 4.6.2.

Table 4.12 – $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ trigger selection.

| Trigger level | Trigger selection | Applied to |
|---------------|---|----------------|
| L0 | L0ElectronDecision AND L0Calo_ECAL_realET > 3000 MeV | e^+ or e^- |
| HLT1 | Hlt1TrackMVADecision | B^+ |
| HLT2 | Hlt2Topo2BodyDecision OR Hlt2Topo3BodyDecision or | B^+ |

4.3.6 Particle identification

Particle Identification (PID, see Sec. 2.2.2) requirements are needed to reduce the background from specific physics decays. Several requirements are applied at different stages of the selection chain:

- loose PID selections are applied at the stripping level to both electrons and muons;
- offline PID selections are applied to reduce physics background in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ selections;
- an additional offline PID selection is applied to the kaon track in $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ decays to suppress the contamination from $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) \pi^+$.

Since the PID is not well modeled in the simulation, a corrected PID selection efficiency is determined with calibration data samples (see Sec. 4.5.3). The specific PID requirements, applied in the selection chain, are elaborated in the following.

Stripping PID selection

In the stripping lines StrippingLFVBu2KJPsiLine, StrippingLFVB2eMuLine and StrippingBs2MuMuLinesBu2JPsiKLine, the following requirements are applied:

- **PIDe > 2**: applied to the electron track in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$;
- **PIDe > -2**: applied to the electron tracks in $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$;
- **isMuon**: applied to the muon tracks in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$.

To correct the overall PID selection efficiency and misidentification rate, all the simulation samples are re-stripped removing the PIDe requirements. However, the IsMuon requirement is not removed, except for the $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ low-statistics sample.

Electron and muon PID requirements

The most dangerous background decay modes are hadronic two-body B meson decays, $B_{(s)}^0 \rightarrow hh'$. Such decays can peak in the signal mass window and significantly impact the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$

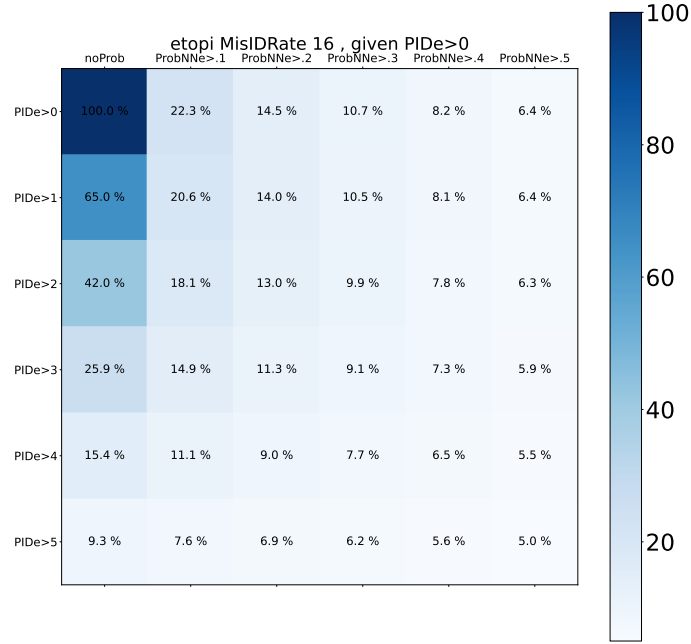


Figure 4.1 – Efficiencies of various requirements $PIDe$ and $ProbNNe$ on the pion from $D^0 \rightarrow K^+ \pi^-$ background in $B^\pm \rightarrow K^\pm e^+ e^-$ data candidates for the data-taking year 2016. The efficiencies are quoted with respect to the $PIDe > 0$ selection requirement. These studies have been performed for Refs. [34, 35].

branching fraction measurement. These peaking backgrounds arise from the misidentification of the reconstructed final-state particles, *e.g.* a true pion can be misidentified as a muon.

To remove as much of this background as possible, while maintaining the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiency as high as possible, additional requirements are applied to the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ electron and muon candidates:

- **$PIDe > 3$ AND $ProbNNe > 0.4$:** applied to the electron track;
- **$ProbNN\mu > 0.4$:** applied to the muon track.

These PID requirements are also applied to the final-state particles of the $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$, in the context of the $r_{J/\psi}$ measurement (see Sec. 4.5.7).

More stringent PID selections have been studied, leading to the conclusion that they would still result in the presence of irreducible misidentification background. However, a tighter PID selection would reduce significantly the signal sensitivity and is therefore not considered.

This decision is supported by studies of the misidentification efficiency performed for Refs. [34, 35]. The $D^0 \rightarrow K^+ \pi^-$ background, where the pion is misidentified as an electron, has been analysed in a sample of $B^\pm \rightarrow K^\pm e^+ e^-$ data candidates with a two-dimensional scan over the PID_e and the ProbNN_e requirements. The $\pi \rightarrow e$ misidentification efficiencies are determined with respect to $\text{PID}_e > 0$ and are presented for the data-taking year 2016 in Fig. 4.1.⁸ The results of this study reveal that the benefit of applying tighter requirements scales only linearly for the misidentification rate, while the signal efficiency significantly decreases.

Kaon PID requirements

To reduce the impact of $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) \pi^+$ decays, a PID requirement is applied to the kaon track in the $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ selection:

- **$\text{PID}_k > 0$ AND $\text{ProbNN}_k \cdot (1 - \text{ProbNN}_p) > 0.05$.**

These choices, as well as the corresponding PID corrections (see Sec. 4.5.3) are taken from Refs. [34, 35].

4.3.7 Multivariate analysis

On top of all selections described previously, a Multivariate Analysis (MVA) is performed to isolate random electron-muon track combination, so-called combinatorial background, from signal decays.

A Boosted Decision Tree (BDT) is trained, using chosen physical quantities to discriminate between the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal and combinatorial background. This machine-learning technique, has been developed specifically for data mining and pattern recognition, as described in Ref. [78], and is widely used in high energy physics. The signal is represented in the BDT training by $B^0 \rightarrow e^\pm \mu^\mp$ simulation, while the background is represented by data from the upper and lower mass sidebands ($m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.5, 5.1] \cup [5.6, 6.5] \text{ GeV}/c^2$). Data and simulation samples are merged over all data-taking years. Furthermore, the efficiency corrections described in Sec. 4.5 are taken into account in simulation, by reweighing the signal sample accordingly. Overall, the training samples contain 947'970 signal events (simulation) and 100'797 background events (data). To account for this difference, a signal sub-sample is randomly drawn for the classifier training to match the size of the data sample. This procedure ensures balanced numbers of signal and background events in the training samples, avoiding a bias towards the larger sample. In addition, the background sample is normalised according to the reweighted effective sample size of the simulation. Furthermore, the signal and background samples are split into five folds (k-folding [79]). While four of the folds are used for the BDT training, the BDT response is evaluated on the remaining fold. This procedure is repeated, alternating the training and testing folds, which results in five trained classifiers.

⁸The misidentification efficiencies determined for the years 2017 and 2018 show very similar trends and are compatible with 2016.

Various alternative multivariate algorithms and sets of discriminating variables have been studied. The best performance is found for the 14 following training features, which are chosen based on their discriminating power:

1. **B^0 DOCA**: distance of closest approach between the two daughter tracks of the B -meson;
2. **min lepton IP**: minimum impact parameter of the daughter tracks;
3. **Long Track Isolation**: response of a classifier, trained for the isolation of long tracks (see App. A.1.1);
4. **Velo Track Isolation**: response of a classifier, trained for the isolation of VELO tracks (see App. A.1.1);
5. **μ cosnk**: the cosine of the angle between the muon momentum in the B rest frame and the vector perpendicular to the B momentum and the z axis;
6. **B^0 ENDVERTEX χ^2** : χ^2 of the B -meson decay vertex;
7. **B^0 PV IP**: B -meson impact parameter with respect to the primary vertex (PV);
8. **$B^0 p_T$** : B -meson transverse momentum;
9. **min lepton p_T** : minimum transverse momentum of the daughter tracks;
10. **min lepton IP χ^2** : minimum impact parameter χ^2 of the daughter tracks (e, μ);
11. **log(1-DIRA)**: with DIRA defined in Sec. 4.3.1;
12. **DeltaR**: defined in Sec. 4.3.1;
13. **B^0 IP χ^2** : B -meson impact parameter χ^2 ;
14. **B^0 FD**: B -meson flight distance.

The distributions of these variables are displayed in Figs. 4.2 and 4.3.

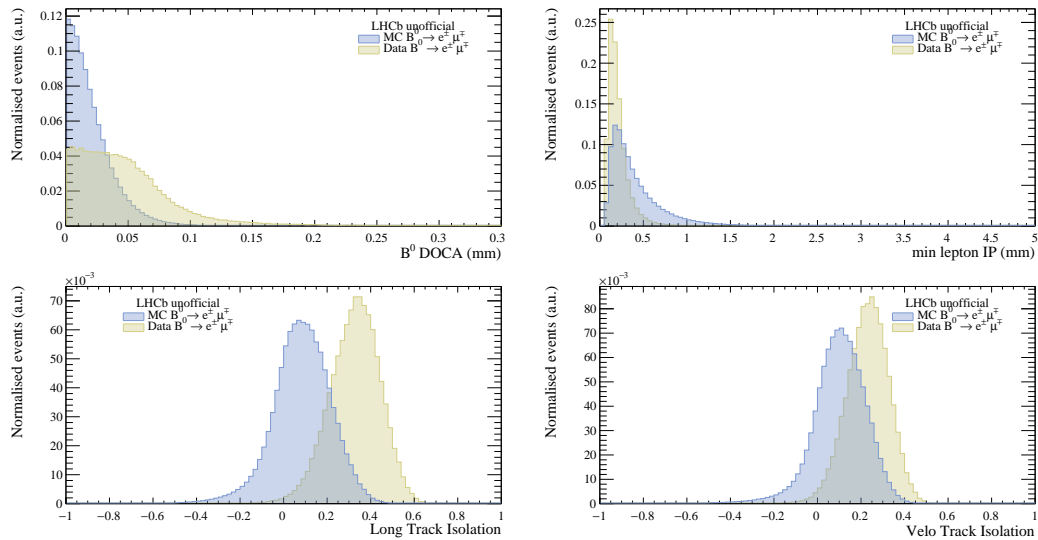


Figure 4.2 – Distributions of discriminating variables used for the BDT training. Blue: Simulation. Yellow: Data sidebands.

Chapter 4. Search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$

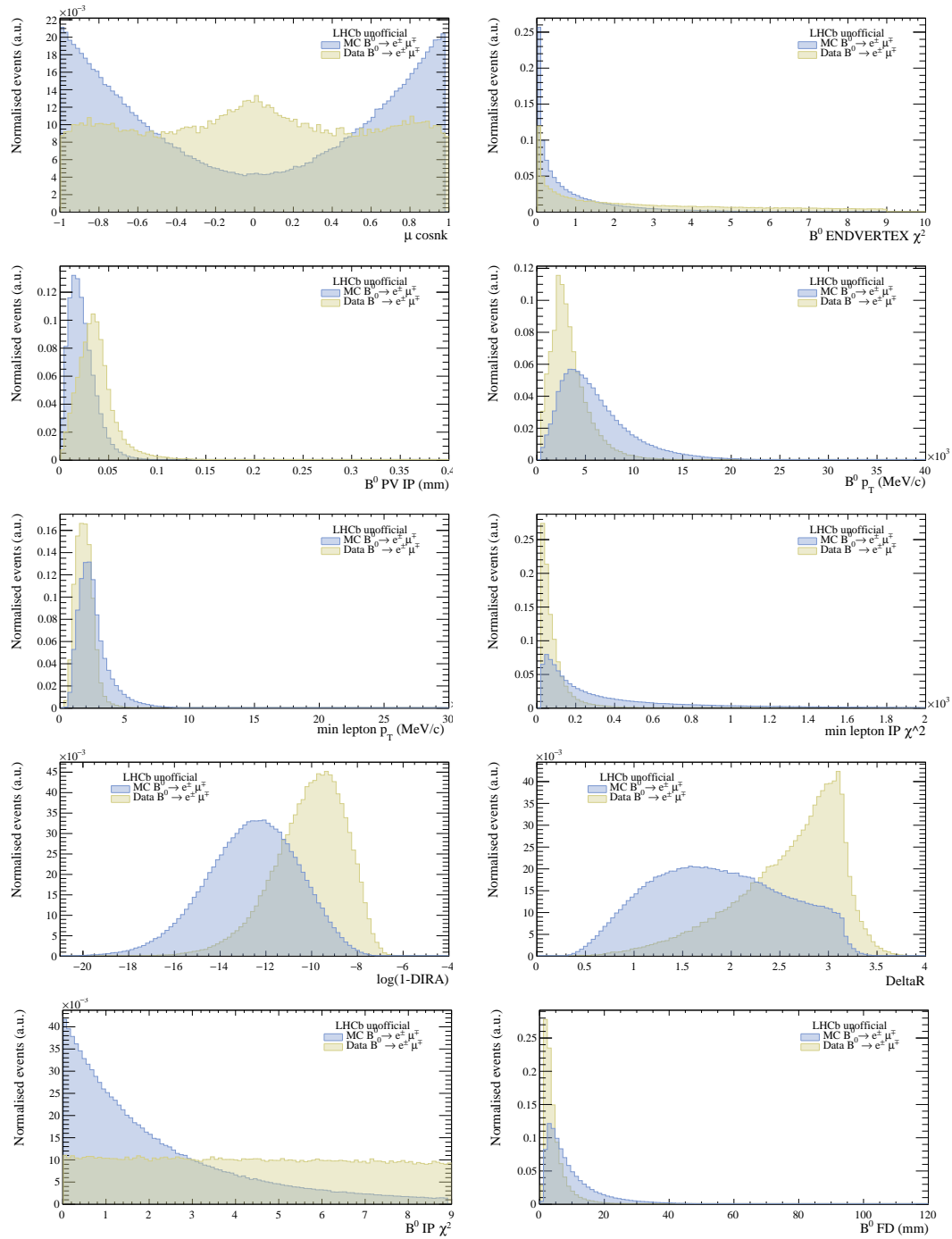


Figure 4.3 – Distributions of discriminating variables used for the BDT training. Blue: Simulation. Yellow: Data sidebands.

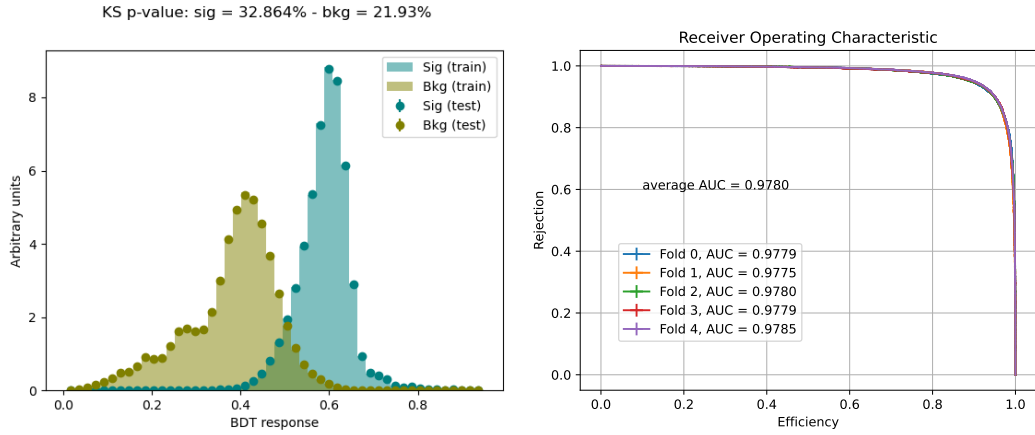


Figure 4.4 – Left: Distribution of the BDT response, for the signal and background test and training samples, summed over all five folds. Right: Receiver Operating Characteristic-Curve, *i.e.* signal efficiency as a function of the background rejection, for each fold.

Among the different algorithms trained, the Adaboost algorithm from the `scikit-learn` machine learning library [80] is found to have the best performance and is therefore chosen as baseline.

Figure 4.4 shows that the test and training samples are compatible for signal and background, which is confirmed by the p-values obtained with the Kolmogorov-Smirnov test [81]. Therefore, over-training can be excluded. The trade-off between the signal efficiency and the background rejection is illustrated by the Receiver Operating Characteristic (ROC) curves displayed in Fig. 4.4 (right). The performance of the five BDTs is found to be compatible. Furthermore,

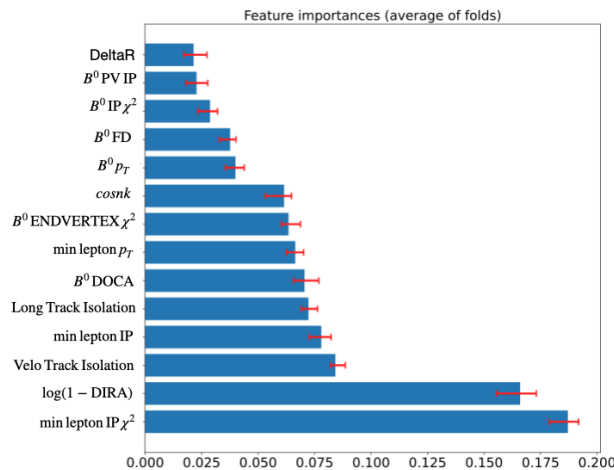


Figure 4.5 – Importance of the discriminating variables in the BDT training. The red error bars cover the variations observed in the five folds.

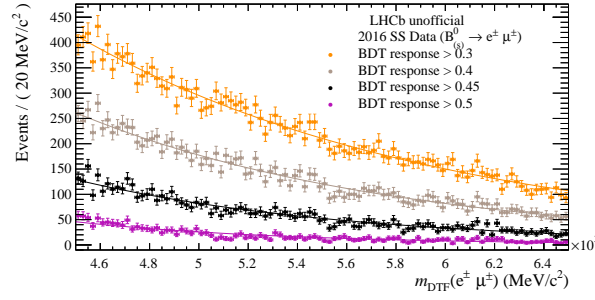


Figure 4.6 – Invariant mass distribution of same-sign $B_{(s)}^0 \rightarrow e^\pm \mu^\pm$ candidates in the 2016 data, for various requirements on the BDT response.

the Area Under the Curve (AUC) is close to 98%. The importance of each discriminating variable in the classifier training, combining all folds, is shown in Fig. 4.5.

The absence of sculpting effects on the combinatorial background, *i.e.* of an influence of a BDT requirement on the shape of the combinatorial background, is tested using Same Sign (SS) $B_{(s)}^0 \rightarrow e^\pm \mu^\pm$ data candidates, which are ideal for the representation of the combinatorial background over the full mass range. An exponential function is fitted to their invariant mass distribution, while applying different cuts on the BDT response. The fits performed in the data-taking year 2016 are shown in Fig. 4.6, while the corresponding plots for the data-taking years 2017 and 2018 can be found in Fig. A.1 in App. A.1.2. The results of these studies, which are compatible among all data-taking years, show that the mass distribution is well modeled by an exponential function, independently of the applied BDT requirement. Hence, no sculpting effects are observed.

4.3.8 Optimisation and binning in multivariate classifier response

In the final step of this analysis, the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal branching fractions are fitted simultaneously in three statistically independent BDT regions, divided into two bremsstrahlung categories, 0γ and 1γ , while merging all data-taking years (see Sec. 4.7.1). These three BDT regions provide different signal sensitivities, due to their different signal-to-background ratios. The requirement on the BDT response is only optimised to remove the region with the highest expected combinatorial background pollution.

The expected combinatorial background in the signal mass region ($m_{DTF}(e^\pm \mu^\mp) \in [5.1, 5.6]$ GeV/c^2) and the $B^0 \rightarrow e^\pm \mu^\mp$ selection efficiency are studied, as a function of the BDT requirement. For each scan point of the BDT requirement, the expected combinatorial background in the signal mass region is determined from an exponential fit of the upper and lower mass sidebands ($m_{DTF}(e^\pm \mu^\mp) \in [4.9, 5.1] \cup [5.6, 6.1]$ GeV/c^2) in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data, merged over all data-taking years and bremsstrahlung categories. The expected combinatorial yield, shown in Fig. 4.7 (left) for each scan point, is then determined by extrapolating the fitted exponential into the signal mass region.

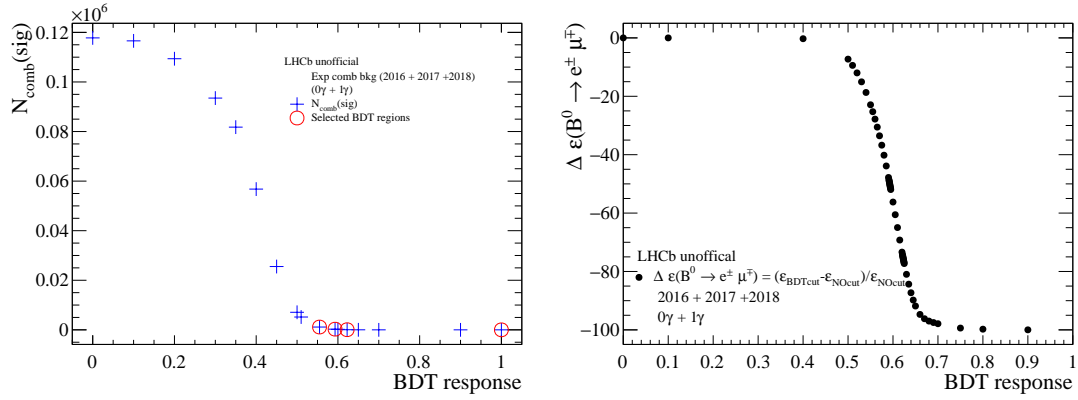


Figure 4.7 – Left: expected combinatorial background in the signal mass window as a function of the requirement on the BDT response. The red circles show the boundaries chosen for the partitioning. Right: relative $B^0 \rightarrow e^\pm \mu^\mp$ selection efficiency decrease (in %) as function of the requirement on the BDT response.

For each scan point, the $B^0 \rightarrow e^\pm \mu^\mp$ selection efficiency in the signal mass range is evaluated from full selected simulation, taking the corrections of Sec. 4.5 into account. The efficiencies are averaged over all bremsstrahlung categories and data-taking years, weighted by the integrated luminosity of each year, and displayed in Fig. 4.7 (right) relative to the efficiency when no BDT requirement is applied. The studies show that a drastic reduction in the expected combinatorial background is achieved for cuts on the BDT response above 0.5, while the signal selection efficiency starts to decrease significantly for tighter BDT requirements.

It is therefore decided to discard the region with BDT response smaller than 0.555 where expected pollution by combinatorial background candidates is very high. The remaining range in BDT response, which contains around 80% of the signal candidates is partitioned in three regions, labeled 0, 1 and 2, corresponding to equal signal populations. The bounds of the selected three BDT regions are shown in Fig. 4.7 (left) and listed in Table 4.13 where also the total signal efficiency for each BDT region is presented.

Table 4.13 – Definition of the three BDT regions, and their signal efficiencies averaged over all data-taking years and bremsstrahlung categories, weighted by the integrated luminosity of each year. Only the binomial errors are displayed here.

| BDT region | Ranges of BDT response | Signal efficiency [%] |
|------------|------------------------|-----------------------|
| 0 | 0.5550 – 0.5930 | 0.593 ± 0.002 |
| 1 | 0.5930 – 0.6225 | 0.595 ± 0.002 |
| 2 | 0.6225 – 1.0000 | 0.584 ± 0.002 |

4.4 Normalisation mass fits

A fit to the invariant B^+ mass distribution, $m_{\text{DTF}}(J/\psi K^+)$, is performed to determine the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ yields in data, which are needed for the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction measurement (see Eq. (4.2)) and the $r_{J/\psi}$ cross-check (see Eq. (4.4)). All selections introduced in Sec. 4.3 are applied to the data and the simulation samples. The fit is performed for each data-taking period simultaneously on the following four data samples:

- μ : $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ candidates;
- $e_{0\gamma}$: $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ candidates with zero bremsstrahlung photons recovered;
- $e_{1\gamma}$: $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ candidates with one bremsstrahlung photon recovered;
- $e_{2\gamma}$: $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ candidates with two bremsstrahlung photons recovered.

The invariant mass distribution is fitted in the range of $[5.2, 5.6]$ GeV/c^2 , where the lower limit is chosen to reduce the pollution from partially reconstructed events. For each of the above four samples, the following components are considered: $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ signal, $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) \pi^+$ misidentification backgrounds and combinatorial background.

The $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) \pi^+$ mass shapes are determined from simulated events, using the following models:

- $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$: sum of two double-sided Crystal Ball functions with common mean;
- $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$: single double-sided Crystal Ball function, for each bremsstrahlung category;
- $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \pi^+$: single double-sided Crystal Ball function;
- $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) \pi^+$: single double-sided Crystal Ball function, for each bremsstrahlung category.

The simulation samples are not corrected. The fits to the $m_{\text{DTF}}(J/\psi K^+)$ distributions for the

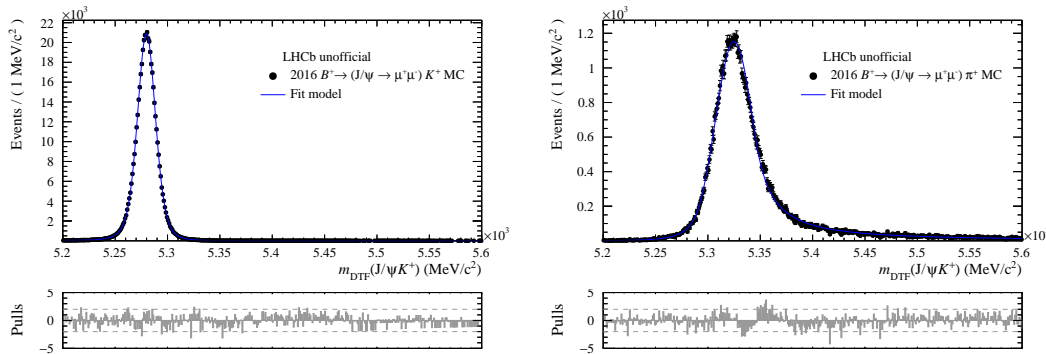


Figure 4.8 – Fitted $m_{\text{DTF}}(J/\psi K^+)$ distribution of $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ (left) and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \pi^+$ (right) simulation for the data-taking year 2016.

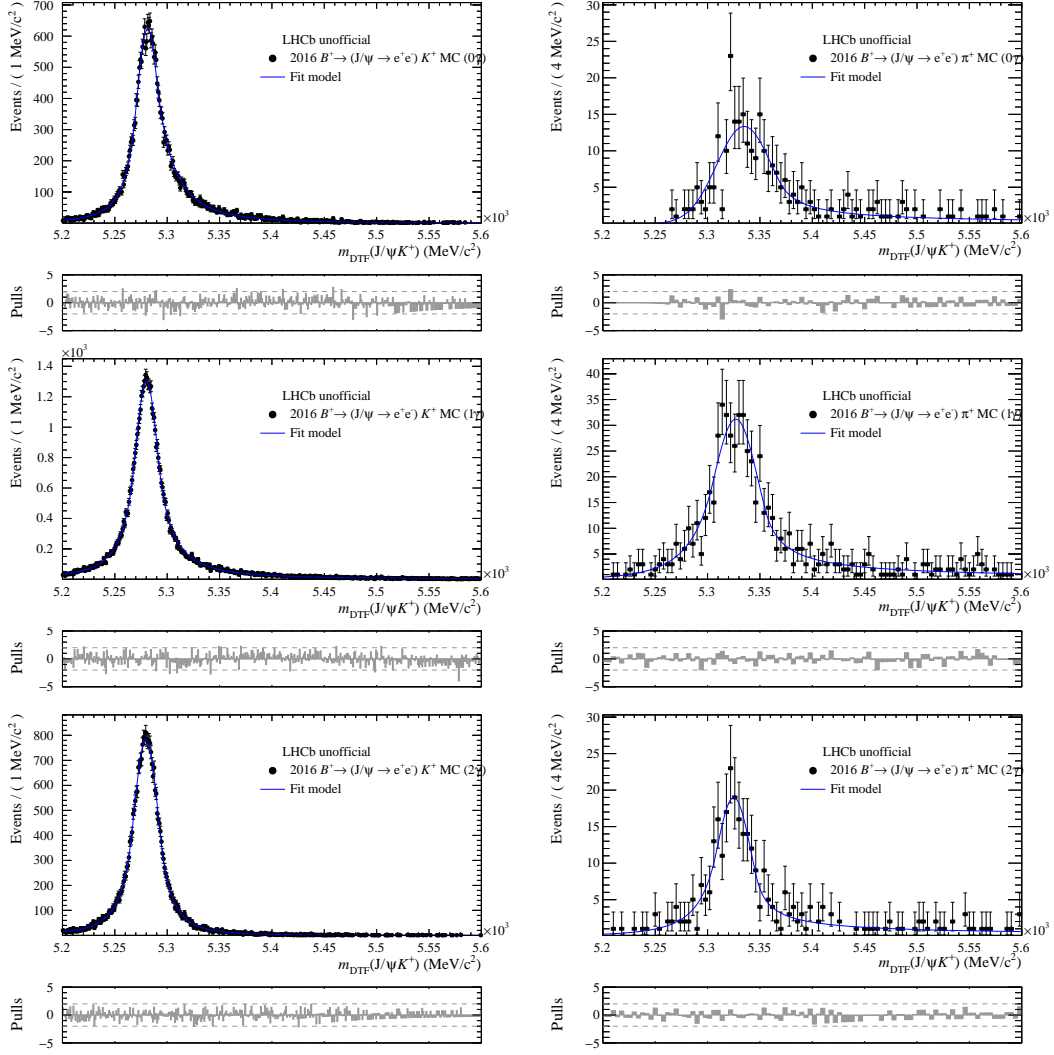


Figure 4.9 – Fitted $m_{\text{DTF}}(J/\psi K^+)$ distribution of $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ (left) and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) \pi^+$ (right) simulation for the data-taking year 2016. Top: 0γ . Middle: 1γ . Bottom: 2γ .

data-taking year 2016 are presented in Figs. 4.8 and 4.9. The fit results for the data-taking years 2017 and 2018 are consistent with the results presented here for 2016 and can be found in App. A.2.

The analytical mass shapes determined in simulation are used in the fits to the data with all parameters fixed. However, for each Crystal Ball function k , two additional free parameters, $\Delta\mu^k$ and s^k , are introduced to allow for modified central values and widths, respectively: $\mu_{\text{data}}^k = \mu_{\text{MC}}^k + \Delta\mu^k$, and $\sigma_{\text{data}}^k = \sigma_{\text{MC}}^k \cdot s^k$. These new free parameters are shared between the $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) \pi^+$ components.

In each of the four fitted samples j , the yield of the misidentified background $B^+ \rightarrow (J/\psi \rightarrow$

Table 4.14 – Fitted yields of $B^+ \rightarrow J/\psi K^+$ decays in data.

| Year | $\mathcal{N}_{B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+}$ | $\mathcal{N}_{B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+}$ (0 γ) | $\mathcal{N}_{B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+}$ (1 γ) | $\mathcal{N}_{B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+}$ (2 γ) |
|------|--|---|---|---|
| 2016 | 528'151 \pm 742 | 21'810 \pm 191 | 40'865 \pm 279 | 20'944 \pm 170 |
| 2017 | 536'205 \pm 747 | 21'552 \pm 186 | 41'400 \pm 280 | 21'144 \pm 174 |
| 2018 | 658'111 \pm 828 | 26'374 \pm 206 | 50'622 \pm 309 | 25'857 \pm 190 |
| Sum | 1'722'467 \pm 1'340 | 69'736 \pm 338 | 132'887 \pm 502 | 67'945 \pm 308 |

$\ell^+ \ell^- \pi^+$ ($\mathcal{N}_{\text{misID}}^j$) is parametrised as a function of the corresponding signal yield ($\mathcal{N}_{\text{sig}}^j$) as

$$\mathcal{N}_{\text{misID}}^j = \frac{\mathcal{B}(B^+ \rightarrow J/\psi \pi^+)}{\mathcal{B}(B^+ \rightarrow J/\psi K^+)} \cdot \mathcal{N}_{\text{sig}}^j \cdot f_{\text{scale}}, \quad (4.6)$$

with $\mathcal{B}(B^+ \rightarrow J/\psi \pi^+) = (3.92 \pm 0.08) \cdot 10^{-5}$ and $\mathcal{B}(B^+ \rightarrow J/\psi K^+) = (1.020 \pm 0.019) \cdot 10^{-3}$ [13]. The ratio of PID efficiencies, denoted as f_{scale} , is shared among all samples and left floating. The combinatorial background is modeled in each sample with an exponential function and the corresponding yields and slope parameters are left floating. The obtained $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ yields are presented in Table 4.14. The resulting fits are shown in Figs. 4.10–4.12.

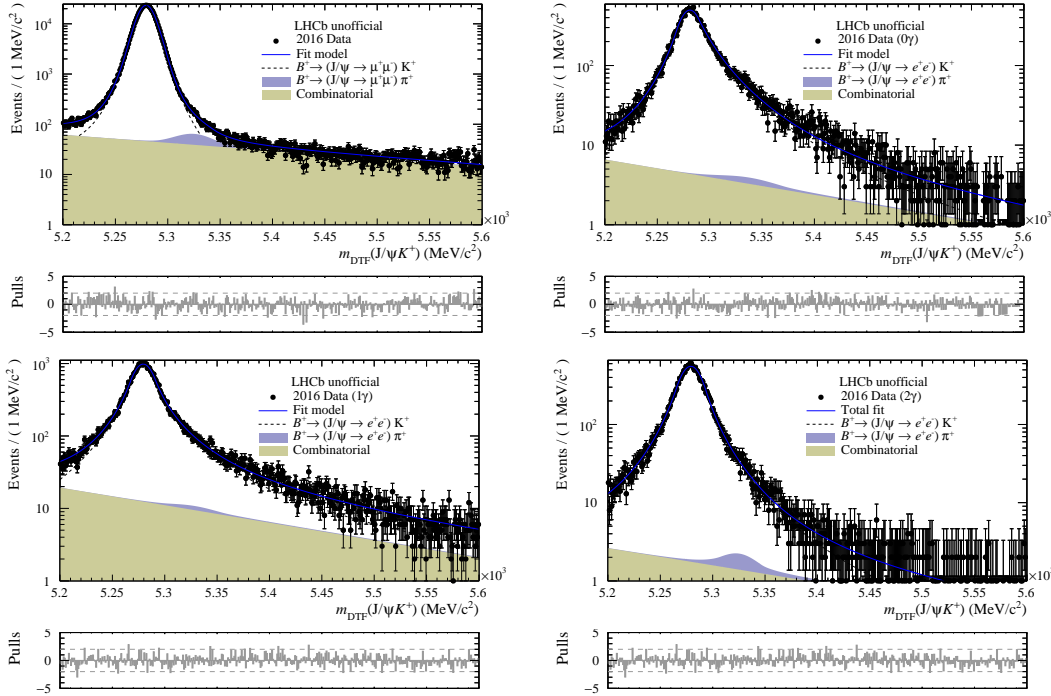


Figure 4.10 – 2016 normalisation fit. The $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data are fitted simultaneously with $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ data, split into three bremsstrahlung categories (0 γ , 1 γ and 2 γ).

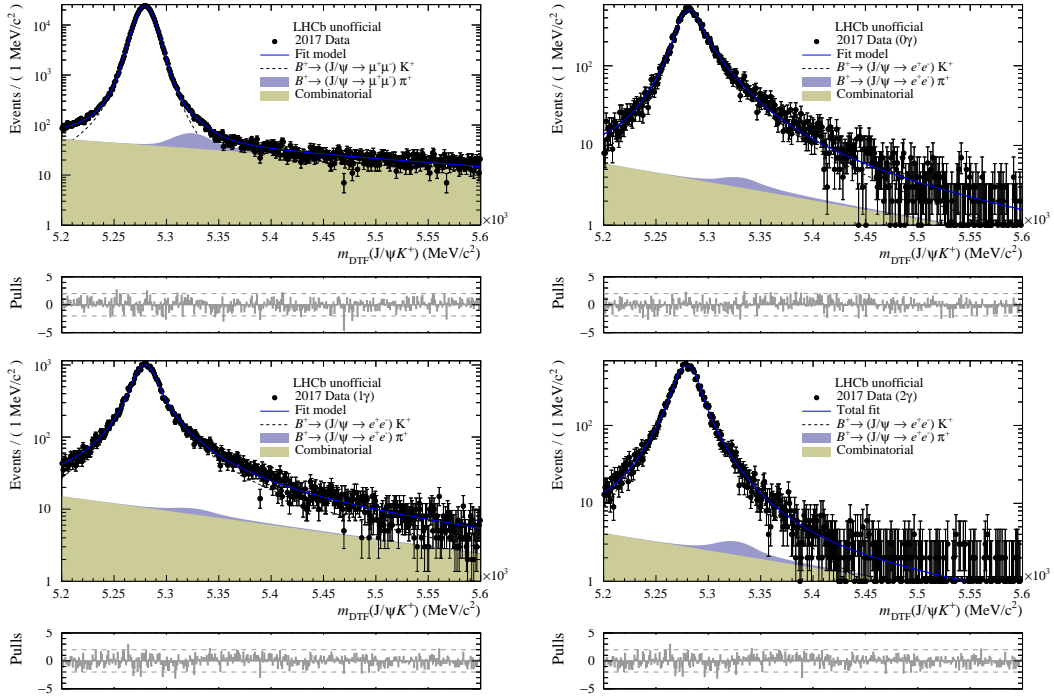


Figure 4.11 – 2017 normalisation fit. The $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data are fitted simultaneously with $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ data, split into three bremsstrahlung categories (0γ , 1γ and 2γ).

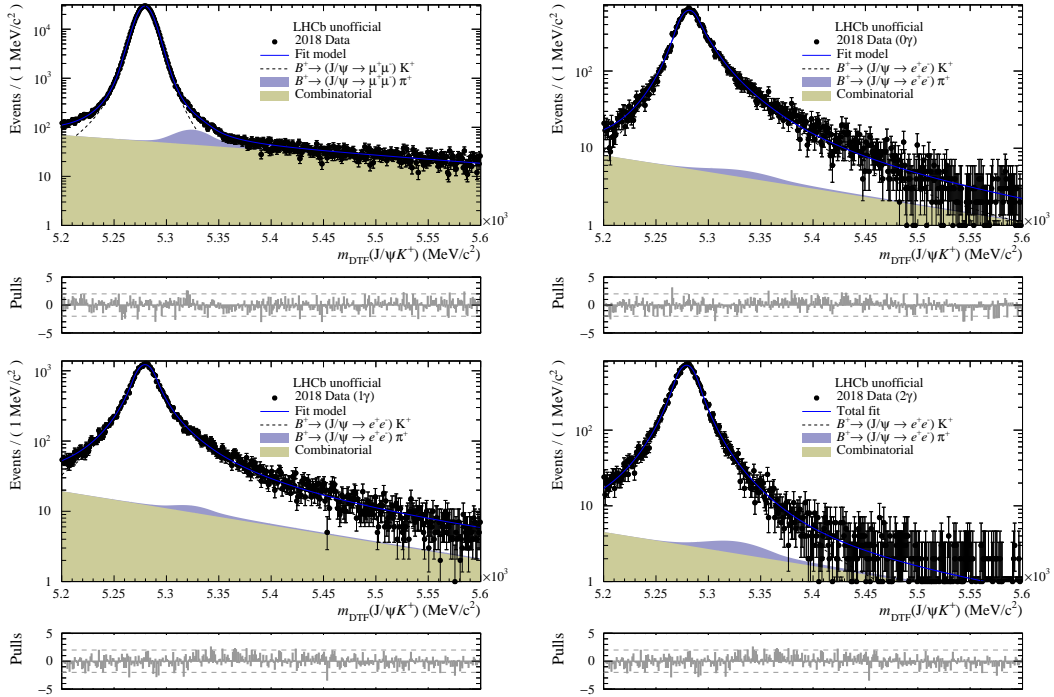


Figure 4.12 – 2018 normalisation fit. The $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data are fitted simultaneously with $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ data, split into three bremsstrahlung categories (0γ , 1γ and 2γ).

4.5 Efficiencies and corrections

The $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ selection efficiencies play a crucial role in the determination of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions (see Eq.(4.2)). The total selection efficiency is determined considering the full selection chain, which takes into account the geometrical acceptance, the stripping, the track reconstruction, the pre-selection, the trigger selection, and the PID requirements. For the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$, the requirements on the BDT are also considered (see Sec. 4.3.8).

Simulation is known to describe poorly the efficiencies, in particular those related to track reconstruction, PID selection and L0 trigger. Furthermore, the kinematic variables of the B mesons are not well modeled in simulation. Therefore, data-driven corrections are introduced to account for these discrepancies.

In the following, all ingredients entering the total selection efficiency and the corresponding corrections are introduced. Firstly, the geometrical acceptance is discussed, before the correction strategies for the track reconstruction, the particle identification, the L0 trigger response and the B kinematics are presented. Finally, the determination of the total selection efficiency is explained.

4.5.1 Geometrical acceptance

The geometrical acceptance is calculated as

$$\varepsilon_{\text{geo}} = \frac{\mathcal{N}_{\text{gen}}^{\text{acc}}}{\mathcal{N}_{\text{gen}}^{4\pi}}, \quad (4.7)$$

where $\mathcal{N}_{\text{gen}}^{\text{acc}}$ is the number of generated B decays of interest with all daughter tracks falling into the LHCb acceptance, while $\mathcal{N}_{\text{gen}}^{4\pi}$ is the total number of generated B decays in the full 4π solid angle. The values for ε_{geo} , taken from the LHCb production webpage, are presented in Table 4.15 for the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ decays, and in Table A.1 of App. A.3.1 for background decays.

Table 4.15 – Geometrical acceptance ε_{geo} (in %) for simulated $B^0 \rightarrow e^\pm \mu^\mp$, $B_s^0 \rightarrow e^\pm \mu^\mp$, $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ decays.

| Year | $B^0 \rightarrow e^\pm \mu^\mp$ | $B_s^0 \rightarrow e^\pm \mu^\mp$ | $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ | $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ |
|------|---------------------------------|-----------------------------------|--|--|
| 2016 | 0.1930 ± 0.0003 | 0.1937 ± 0.0005 | 0.1730 ± 0.0004 | 0.1730 ± 0.0004 |
| 2017 | 0.1931 ± 0.0003 | 0.1937 ± 0.0005 | 0.1731 ± 0.0003 | 0.1733 ± 0.0004 |
| 2018 | 0.1929 ± 0.0003 | 0.1939 ± 0.0005 | 0.1737 ± 0.0003 | 0.1731 ± 0.0004 |

4.5.2 Track reconstruction efficiency corrections

The electron and muon track reconstruction efficiencies are corrected with tag-and-probe methods, using $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ decays in data and simulation. Track weights are determined, as a function of the electron or muon kinematic variables, they are computed as the data/simulation ratio of the electron [82] or muon [83] track reconstruction efficiencies, respectively. The weights, ω_{TRK} , are applied per electron or muon track. The final event weight is given by the product of the two lepton weights, $\omega_{\text{TRK}} = \omega_{\text{TRK}}(\ell_1) \cdot \omega_{\text{TRK}}(\ell_2)$.

The electron tracking weights are obtained as

$$\omega_{\text{TRK}}(e) = \frac{\mathcal{P}(\text{long}|\text{VELO})^{\text{data}}}{\mathcal{P}(\text{long}|\text{VELO})^{\text{MC}}} . \quad (4.8)$$

Here, $\mathcal{P}(\text{long}|\text{VELO})$ denotes the probability of reconstructing an electron track through the full spectrometer (so-called long tracks, see Sec. 2.2.1), given that the track is already reconstructed in the VELO, as determined with the tag-and-probe method of Ref. [82]. These weights are determined for each data-taking year, in four different regions of the electron pseudorapidity η . Furthermore, each pseudorapidity region is binned in the transverse momentum, p_T , and the azimuthal angle, ϕ , of the electron candidate. The obtained weights for the data-taking year 2016 are shown in Fig. 4.13. The corresponding weights for the data-taking years 2017

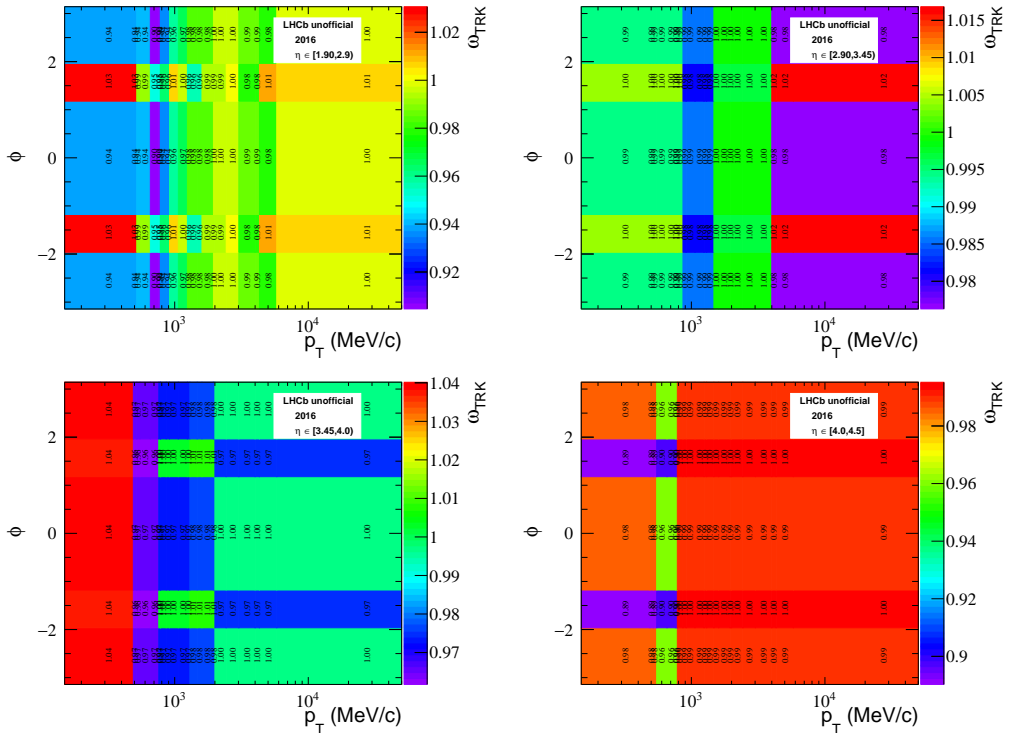


Figure 4.13 – 2016 electron tracking efficiency weight, ω_{TRK} , as a function of the electron p_T and ϕ in four regions of electron η .

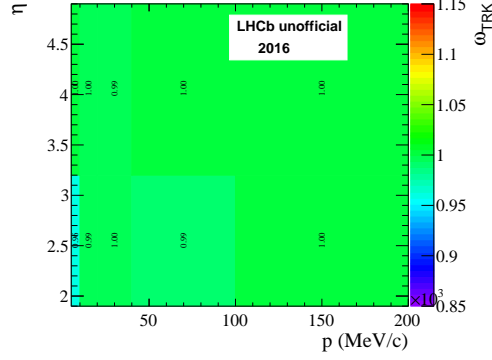


Figure 4.14 – 2016 muon tracking efficiency weight, ω_{TRK} , as a function of the muon momentum and pseudorapidity.

and 2018 are displayed in Fig. A.5 of App. A.3.2.

The muon tracking weights are determined for each data-taking year, in bins of the muon momentum p and pseudorapidity η , as

$$\omega_{\text{TRK}}(\mu) = \frac{\varepsilon_{\text{TRK}}(\mu)^{\text{data}}}{\varepsilon_{\text{TRK}}(\mu)^{\text{MC}}}, \quad (4.9)$$

where $\varepsilon_{\text{TRK}}(\mu)$ is the track reconstruction efficiency determined with the tag-and-probe method of Ref. [83]. The obtained results for the data-year 2016 are shown in Fig. 4.14, while those for the data-taking year 2017 and 2018 can be found in Fig. A.4 of App. A.3.2.

4.5.3 PID efficiencies

The PID efficiencies are determined as a function of kinematic observables of the electron, muon or kaon, from data samples, denoted as calibration datasets, produced centrally with a dedicated selection. To ensure the validity of the obtained PID efficiencies for this analysis, several requirements are applied in the pre-selection (see Sec. 4.3.4) to align the working point of this analysis with the calibration datasets. The final PID efficiency of an event is given as the product of the PID efficiencies determined for all daughter tracks, *i.e.* the PID efficiency for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ is computed as $\varepsilon_{\text{PID}}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) = \varepsilon_{\text{PID}}(e) \cdot \varepsilon_{\text{PID}}(\mu)$ and for $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ as $\varepsilon_{\text{PID}}(B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+) = \varepsilon_{\text{PID}}(\ell^-) \cdot \varepsilon_{\text{PID}}(\ell^+) \cdot \varepsilon_{\text{PID}}(K^+)$.

In the following, the framework used to determine the efficiencies for the electron, muon and kaon PID requirements is briefly introduced, followed by a detailed presentation of the results for the different cases.

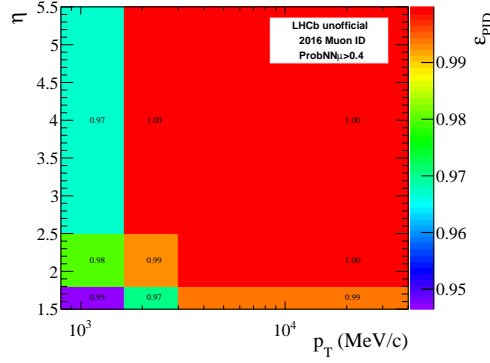


Figure 4.15 – Muon PID efficiencies as a function of the muon transverse momentum p_T and pseudorapidity η , determined for the data-taking year 2016.

Muon PID efficiency

The PID requirement applied to the muon candidates in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ selection is

- **ProbNN μ > 0.4.**

Its efficiency is determined with the `PIDCalib2` package, an upgraded version of the `PIDCalib` package [84], as a function of the muon p_T and η . In `PIDCalib2`, the efficiency for a given PID requirement is determined using background-subtracted data obtained with the `sPlot` technique [85] from a Maximum Likelihood fit to the invariant mass distribution of a $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ calibration dataset. The binning schemes are obtained by studying the one-dimensional behaviours of the efficiency as a function of muon p_T and η , and chosen to capture the observed trends. A binning in the track multiplicity, is not considered, since no noticeable trend is observed as a function of track multiplicity.

The muon PID efficiencies are displayed in Fig. 4.15 for the data-taking year 2016, and in Fig. A.6 of App. A.3.3 for the years 2017 and 2018.

Electron PID efficiency

The PID requirement applied to the electron candidates in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ selection is

- **PID e > 3 AND ProbNNe > 0.4.**

Its efficiency is determined separately for each data-taking year, for the bremsstrahlung categories 0γ and 1γ , and for two regions of the track multiplicity (`nTracks`), binned in the electron transverse momentum p_T and pseudorapidity η . A tag-and-probe technique is used, where the PID cut is applied to one of the two electron candidates of $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$, the probe electron, while the second electron, denoted as tag electron, is selected

with a tight PID requirement $\text{PID}_e > 5$. This approach is chosen over PIDCalib2 due to the substantial impact of the PID requirement on the resolution of the B mass relying on background-subtracted data as used by PIDCalib2 is not recommended when the fit model varies across the phase space. The PID efficiency is computed as

$$\epsilon_{\text{PID}}^k = \frac{\mathcal{N}_{\text{pass}}^k}{\mathcal{N}_{\text{pass}}^k + \mathcal{N}_{\text{fail}}^k}, \quad (4.10)$$

where $\mathcal{N}_{\text{pass,fail}}^k$ denotes the number of $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ decays of the calibration dataset where the probe electron candidate in bin k passes or fails the PID requirement. In each bin k of transverse electron momentum p_T and the electron pseudorapidity η , the $m_{\text{DTF}}(J/\psi K^+)$ distributions of the pass and fail samples are fitted simultaneously with ϵ_{PID}^k as shared parameter $\epsilon_{\text{PID}}(e)^k$, relating the two yields:

$$\mathcal{N}_{\text{fail}}^k = \frac{1 - \epsilon_{\text{PID}}^k}{\epsilon_{\text{PID}}^k} \cdot \mathcal{N}_{\text{pass}}^k. \quad (4.11)$$

In the fit, the $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ signal component is modeled with a Crystal Ball function, while an exponential function is used to model the combinatorial background. Furthermore,

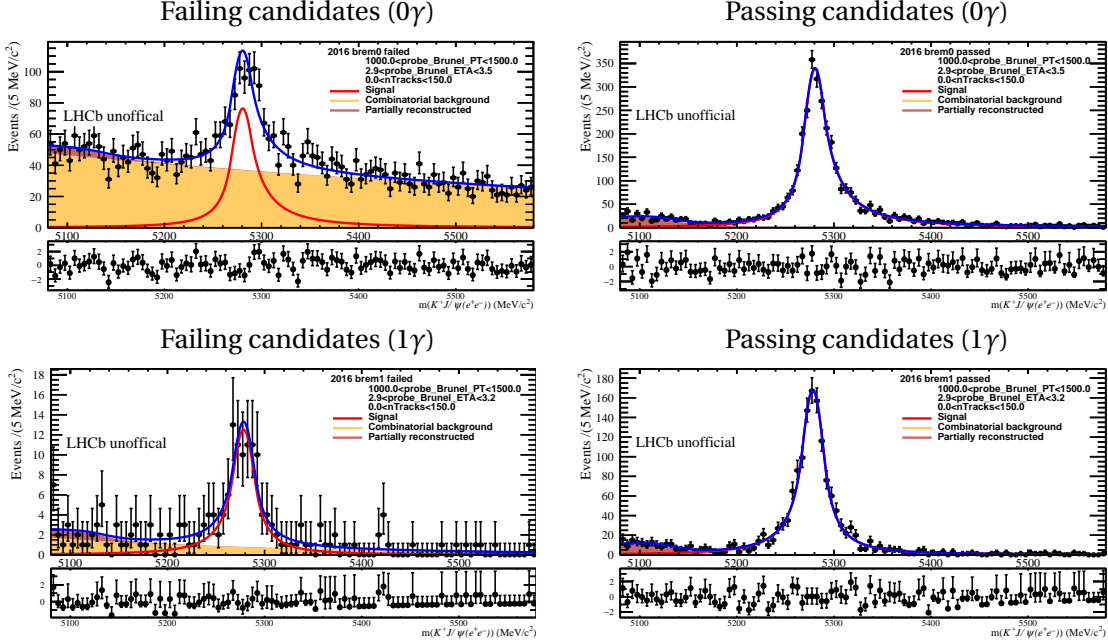


Figure 4.16 – $m_{\text{DTF}}(J/\psi K^+)$ mass distributions, in 2016 data, for $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ candidates failing (left) and passing (right) the electron PID requirement in one (p_T , η , nTracks) bin. The fits are described in the text. The top (bottom) row shows the results for the bremsstrahlung category 0γ (1γ).

partially reconstructed decays are modeled, in each bin of the two fit categories, with a one-dimensional kernel-density estimator, consisting of a superposition of Gaussian functions [86], extracted from simulated $B^0 \rightarrow K^{*0}(\rightarrow K^+\pi^-)J/\psi$ decays passing or failing the electron PID requirement. The partially reconstructed background yield is related in each bin to the $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ signal yields as

$$\mathcal{N}_{\text{fail,pass}}^{k,\text{PartReco}} = f^{k,\text{PartReco}} \cdot \mathcal{N}_{\text{fail,pass}}^k, \quad (4.12)$$

where $f^{k,\text{PartReco}}$ is a shared parameter between the pass and fail subsamples. Figure 4.16 shows an example of the $m_{\text{DTF}}(J/\psi K^+)$ distributions for one $(p_T, \eta, n\text{Tracks})$ bin and the two bremsstrahlung categories.

The binning schemes are obtained by studying the electron PID efficiency separately as a function of the electron p_T and η with the fit-and-count method just described above and PIDCalib2. These one-dimensional studies also allow the comparison of the results obtained with the two different methods. The results are presented in Fig. 4.17 for the data-taking year 2016 and separately for both bremsstrahlung categories, using binning schemes different from the final one. The plots also show the normalised p_T and η distributions of the $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ calibration dataset. The results for the data-taking years 2017 and 2018 are found to be compatible and are shown in Figs. A.7 and A.8 of App. A.3.3. The trends of the electron PID efficiency for both methods are compatible. Based on these one-dimensional studies, a binning scheme is chosen in p_T and η , for the final determination of the electron PID efficiencies with the fit-and-count method, such that the observed efficiency trends are captured. Similar studies of the PID efficiency as function of $n\text{Tracks}$ show no dramatic

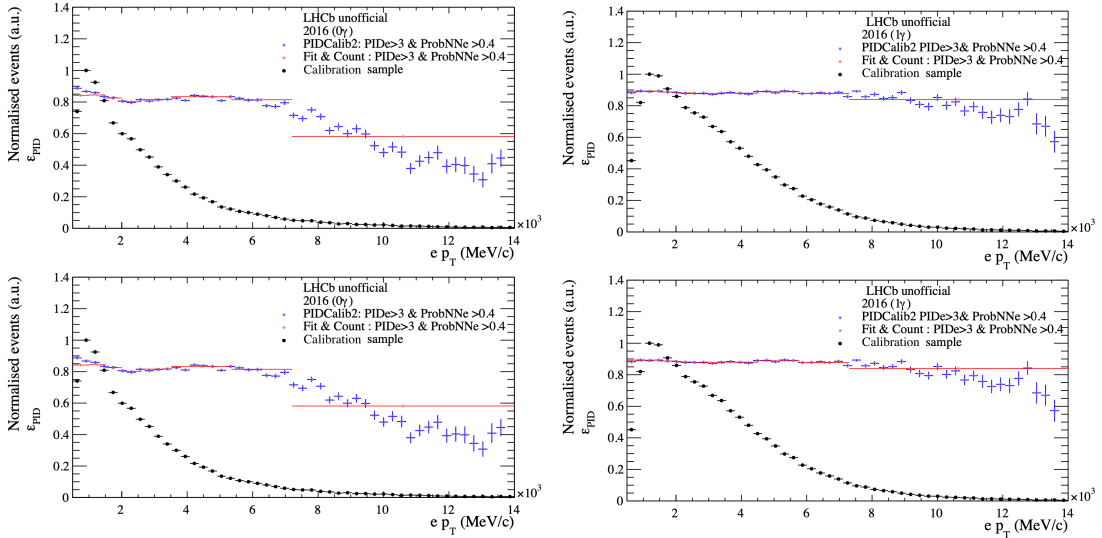


Figure 4.17 – 2016 electron PID efficiencies, obtained with the fit-and-count method (red) and PIDCalib2 (blue) as a function of the electron transverse momentum (top) and the pseudorapidity (bottom) for the bremsstrahlung categories 0γ (left) and 1γ (right). The electron p_T and η distributions of the calibration sample are also shown.

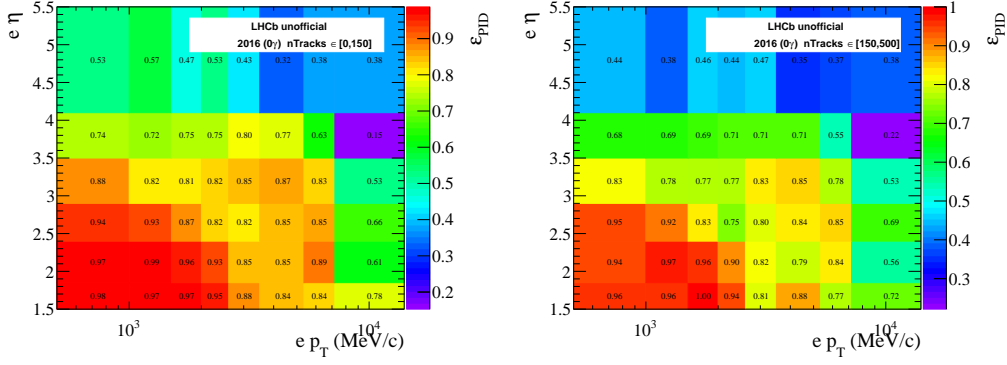


Figure 4.18 – Electron PID efficiencies determined for the data-taking year 2016 and bremsstrahlung category 0γ as a function of the electron transverse momentum p_T and pseudorapidity η , for the low (left) and high (right) nTracks regions.

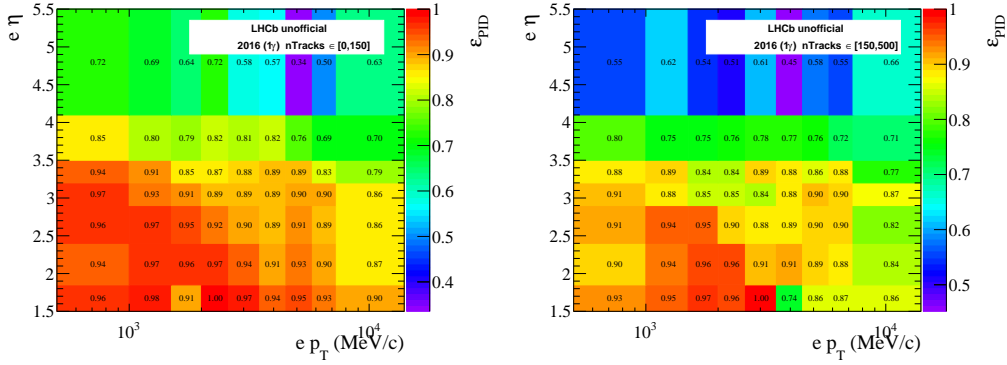


Figure 4.19 – Electron PID efficiencies determined for the data-taking year 2016 and bremsstrahlung category 1γ as a function of the electron transverse momentum p_T and pseudorapidity η , for the low (left) and high (right) nTracks regions.

trends. Therefore, the choice of the two nTracks regions is fiducial to avoid low statistics regions. The final PID efficiencies obtained with the fit-and-count method are presented for the data-taking year 2016 in Figs. 4.18 and 4.19. The corresponding efficiencies for the data-taking years 2017 and 2018 are presented in Figs. A.9 and A.10 of App. A.3.3.

Kaon PID efficiency

The PID requirement applied to the kaon candidates in the $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ selection is

- $PIDk > 0$ AND $\text{ProbNN}k \cdot (1 - \text{ProbNN}p) > 0.05$.

Its efficiency is determined with the PIDCalib2 package in five regions of the track multiplicity, nTracks, as a function of the kaon momentum and the kaon pseudorapidity [34, 35] and shown in Fig. 4.20 for the data-taking year 2016, and in Figs. A.11 and A.12 of App. A.3.3 for the

data-taking years 2017 and 2018.

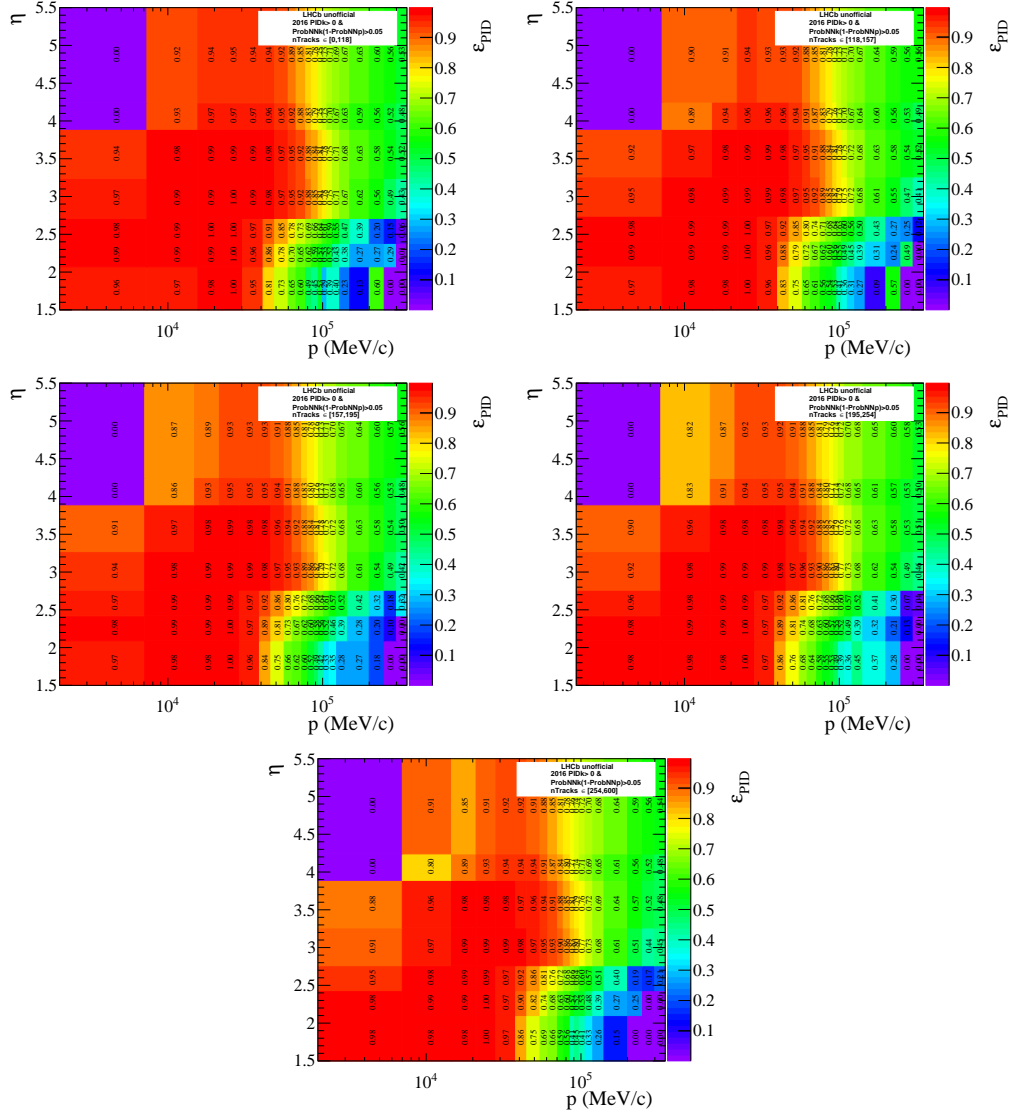


Figure 4.20 – Kaon PID efficiencies for the data-taking year 2016 as a function of the kaon momentum and pseudorapidity for the five regions in increasing track multiplicity (from top to bottom and left to right) [34, 35].

MisID efficiencies

In addition to the PID efficiencies mentioned so far, misidentification (MisID) efficiencies are also determined for the cases when a kaon, a pion or a proton is misidentified as a muon or an electron, as well as for the cases when an electron or muon is misidentified as a muon or an electron, respectively.

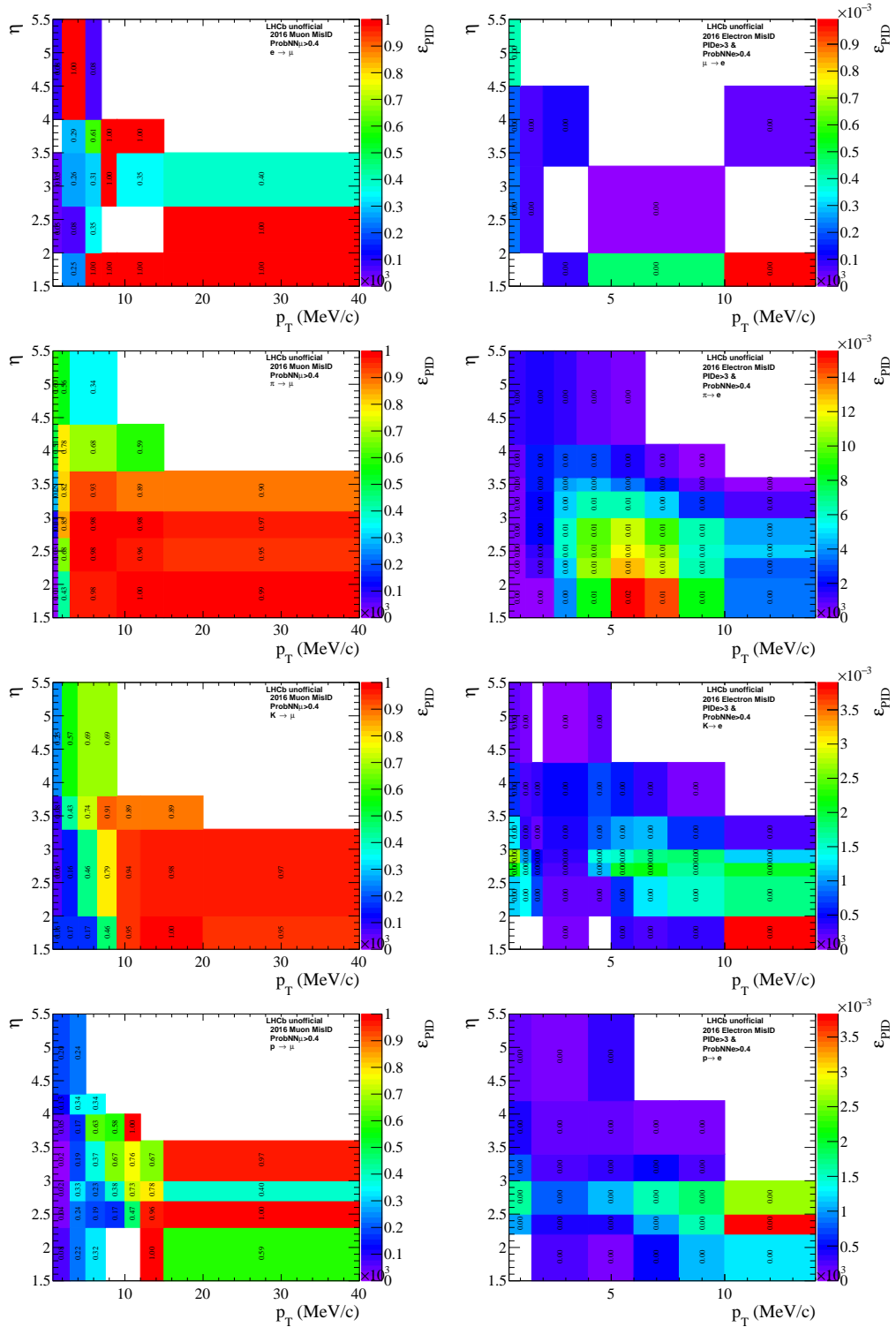


Figure 4.21 – Left: 2016 efficiencies for the misidentification of an electron, a pion, a kaon or a proton (from top to bottom) as a muon. Right: 2016 efficiencies for the misidentification a muon, a pion, a kaon or a proton (from top to bottom) as an electron. The empty bins (white) are caused by the low statistics regions in the calibration datasets.

The MisID efficiencies, which are needed to estimate the level of physics background in the selected sample of $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ candidates (see Sec. 4.6.1), are determined with the PIDCalib2 package, using dedicated calibration datasets [84], as a function of the particle transverse momentum p_T and the pseudorapidity η . The binning schemes are chosen to capture the efficiency trends observed when studying the efficiency as a function of the particle p_T or η only.

The obtained efficiencies for the misidentification of an electron, a pion, a kaon or a proton as a muon and the misidentification of a muon, a pion, a kaon or a proton as an electron are presented in Fig. 4.21 for the data-taking year 2016, and in Figs. A.13 and A.14 of App. A.3.3. The empty bins (white) are caused by the low statistics regions in the calibration datasets and are treated by assuming an efficiency of zero. Furthermore, this effects are taken into account in the systematic uncertainties (see Sec. 4.9).

isMuon efficiency

The isMuon PID requirement is applied to the muon candidate in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ reconstruction (see Sec. 4.3.1).

This requirement is suspected to be modeled well enough in simulation and, therefore, its efficiency does not need to be estimated with a data-driven method. However, this is nevertheless done for the $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ background estimation of Sec. 4.6.1. The reason is that the simulated samples run very low in statistics when applying the full selection chain, which makes the determination of the selection efficiency and of the shape of the invariant mass distribution very difficult. Therefore, the isMuon requirement is removed in the stripping selection and the isMuon efficiency is taken into account instead when determining the $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ selection efficiency and the shape of the $m_{\text{DTF}}(e^\pm \mu^\mp)$ distribution as shown in Sec. 4.7.4.

The isMuon efficiencies for a true muon, pion, kaon, proton or electron are determined, using PIDCalib2, as a function of the muon transverse momentum p_T and pseudorapidity η . The results are presented in Figs. A.15–A.17 of App. A.3.3 for all three data-taking years.

4.5.4 L0 trigger efficiency corrections

The L0 trigger response is known to be badly modeled in simulation and a data-driven approach, called the TISTOS method [87], is used to correct the L0 trigger efficiencies.

The efficiencies are determined for the L0ElectronDecision_TOS and L0MuonDecision_TOS trigger lines, by applying Trigger Independent of Signal (TIS) and Trigger On Signal (TOS) (see Sec. 2.2.3) selections to the lepton tracks in $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data and simulation. For each lepton track the L0 efficiency is calculated in bins of

the lepton kinematics as

$$\epsilon_i^{\text{TOS}} = \frac{\mathcal{N}_i^{\text{TISTOS}}}{\mathcal{N}_i^{\text{TIS}}}, \quad (4.13)$$

where $\mathcal{N}_i^{\text{TISTOS}}$ denotes the number of events passing the TIS and TOS selections in bin i , and $\mathcal{N}_i^{\text{TIS}}$ denotes the number of events selected only with the TIS requirement. The following TIS selection is applied to the B^+ candidates:

- L0MuonDecision_TIS OR L0HadronDecision_TIS.

The L0 trigger correction is then determined as an event weight, ω_{L0} , by calculating the data/MC efficiency ratio as

$$\omega_{\text{L0}}(B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-)) = \frac{1 - (1 - \epsilon_{\text{data}}^{\text{TOS}}(\ell^+)) \cdot (1 - \epsilon_{\text{data}}^{\text{TOS}}(\ell^-))}{1 - (1 - \epsilon_{\text{MC}}^{\text{TOS}}(\ell^+)) \cdot (1 - \epsilon_{\text{MC}}^{\text{TOS}}(\ell^-))}, \quad (4.14)$$

where $\epsilon(\ell)$ refers to the efficiency determined from data or simulation for a specific lepton track $\ell = e, \mu$.

For the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal, the L0 trigger correction weight is computed by adapting Eq. (4.14) as

$$\omega_{\text{L0}}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) = \frac{1 - (1 - \epsilon_{\text{data}}^{\text{TOS}}(e)) \cdot (1 - \epsilon_{\text{data}}^{\text{TOS}}(\mu))}{1 - (1 - \epsilon_{\text{MC}}^{\text{TOS}}(e)) \cdot (1 - \epsilon_{\text{MC}}^{\text{TOS}}(\mu))}. \quad (4.15)$$

In the following, a more detailed description of the determination of the L0 muon and electron trigger efficiencies is given.

L0 muon efficiency

The L0 trigger efficiency for the L0MuonDecision_TOS trigger line is obtained from $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data and simulation, for both muon tracks. In data, the full selection chain, except the L0MuonDecision_TOS selection, is taken into account, while in simulation the above-introduced tracking corrections and PID efficiencies are taken into account, as well as the kinematic corrections introduced in Sec. 4.5.5.

The L0 trigger efficiency ϵ^{TOS} is determined as a function of the muon p_T and pseudorapidity η , with a binning scheme chosen to capture the observed efficiency trends. The results for the data-taking year 2016 are presented in Fig. 4.22. The results for the data-taking years 2017 and 2018 are found to be compatible with those for 2016 and can be found in Figs. A.18 and A.19 of App. A.3.4.

L0 electron efficiency

The L0 trigger efficiency for the L0ElectronDecision_TOS trigger line is obtained from $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ data and simulation, for both electron tracks and separately for

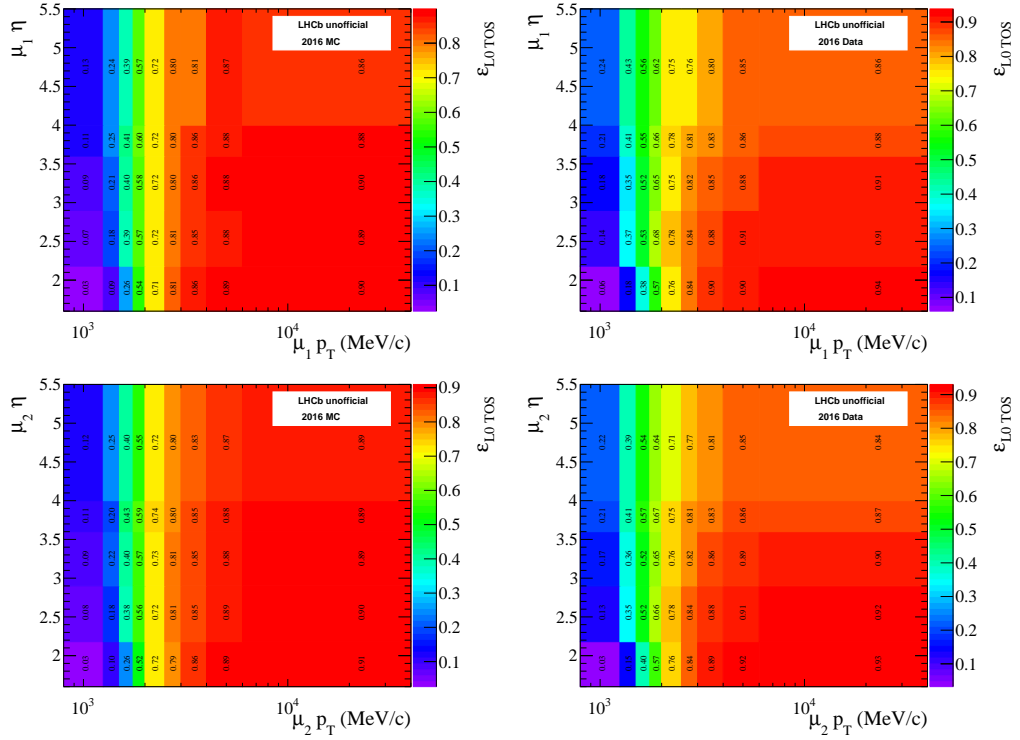


Figure 4.22 – L0 trigger efficiency ϵ^{TOS} determined from 2016 $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ simulation (left) and data (right) for both the μ^+ (top) and μ^- (bottom) tracks, as a function of the muon p_T and η .

the two bremsstrahlung categories 0γ and 1γ . In data the full selection chain, except the LOElectronDecision_TOS selection, is taken into account, while in simulation the above-introduced tracking corrections and PID efficiencies are taken into account, as well as the kinematic corrections introduced in Sec. 4.5.5. The efficiency is also determined in three regions of the electromagnetic calorimeter, denoted as Calo region 0, 1, and 2, because of the non-constant granularity of the ECAL system (Sec. 2.2.2).

In addition, each of these three calorimeter regions is binned in the transverse energy deposited in the ECAL, LOCalo_ECAL_realET. The binning scheme is chosen to capture the observed behaviour of the efficiency. The obtained efficiencies are fitted with the following error function:

$$f(x) = a_0 \cdot a_5^{\frac{-(x-a_1)}{\sqrt{2} \cdot a_2}} + a_3 \cdot \left(1 + \operatorname{erfc} \left(\frac{x - a_1}{\sqrt{2} \cdot a_4} \right) \right) + a_6, \quad (4.16)$$

where x is the central value of LOCalo_ECAL_realET and a_i are the fit parameters.

The distributions of LOCalo_ECAL_realET for the TIS and TISTOS electrons from $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ decays, as well as the corresponding L0 trigger efficiency fitted with Eq. (4.16), are shown in Fig. 4.23 (0γ) and Fig. 4.24 (1γ) for 2016 simulation and in Fig. 4.25 (0γ) and Fig. 4.26 (1γ) for 2016 data. The working point of this analysis at LOCalo_ECAL_realET > 3000

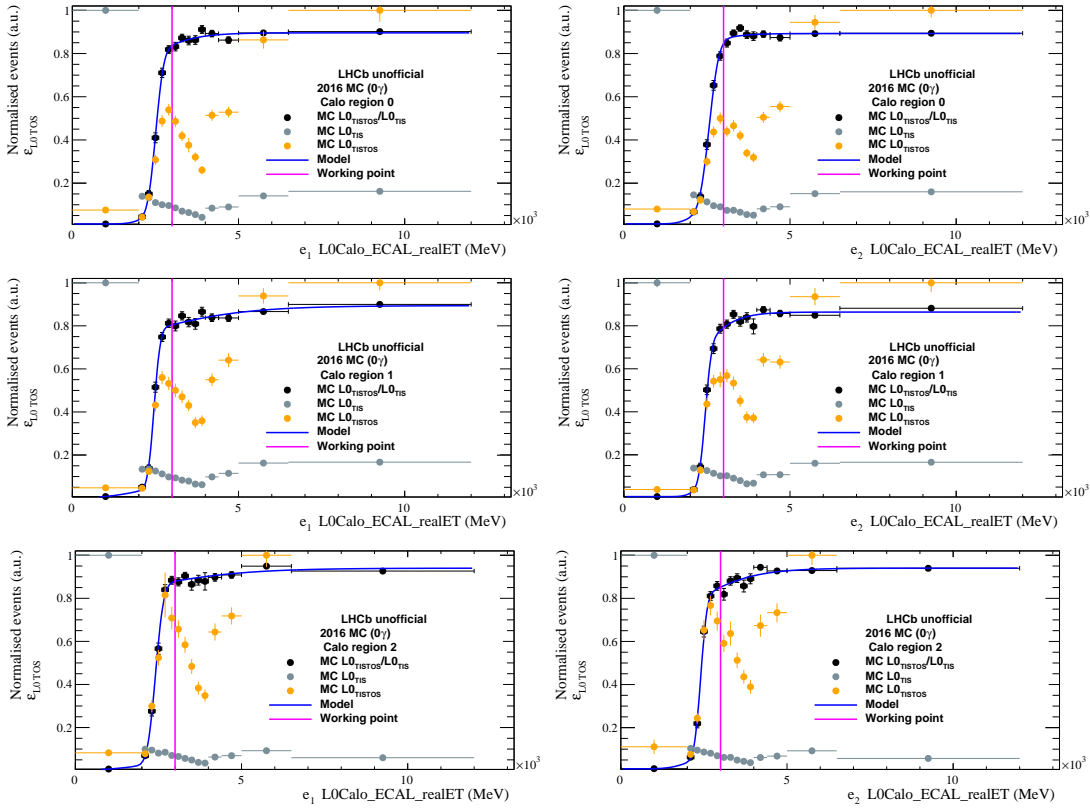


Figure 4.23 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 0γ from $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ simulation for the data-taking year 2016, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

MeV (see Sec. 4.3.5), also shown on the figures, is close to the region where the efficiency curves flatten. The results for the data-taking years 2017 and 2018 are presented in App. A.3.4.

In order to determine the final electron L0 trigger weights, ω_{L0} , the fitted efficiency curve $f(x)$ for each event in simulation is used rather than the binned results.

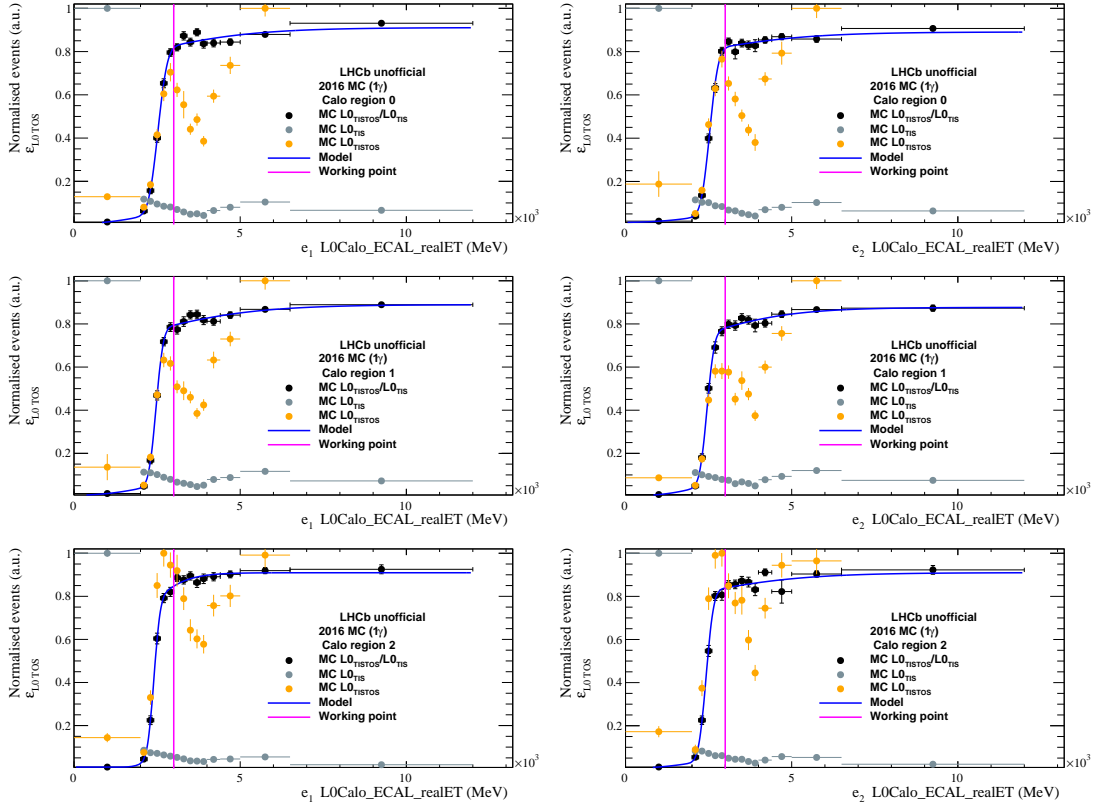


Figure 4.24 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 1γ from $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ simulation for the data-taking year 2016, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

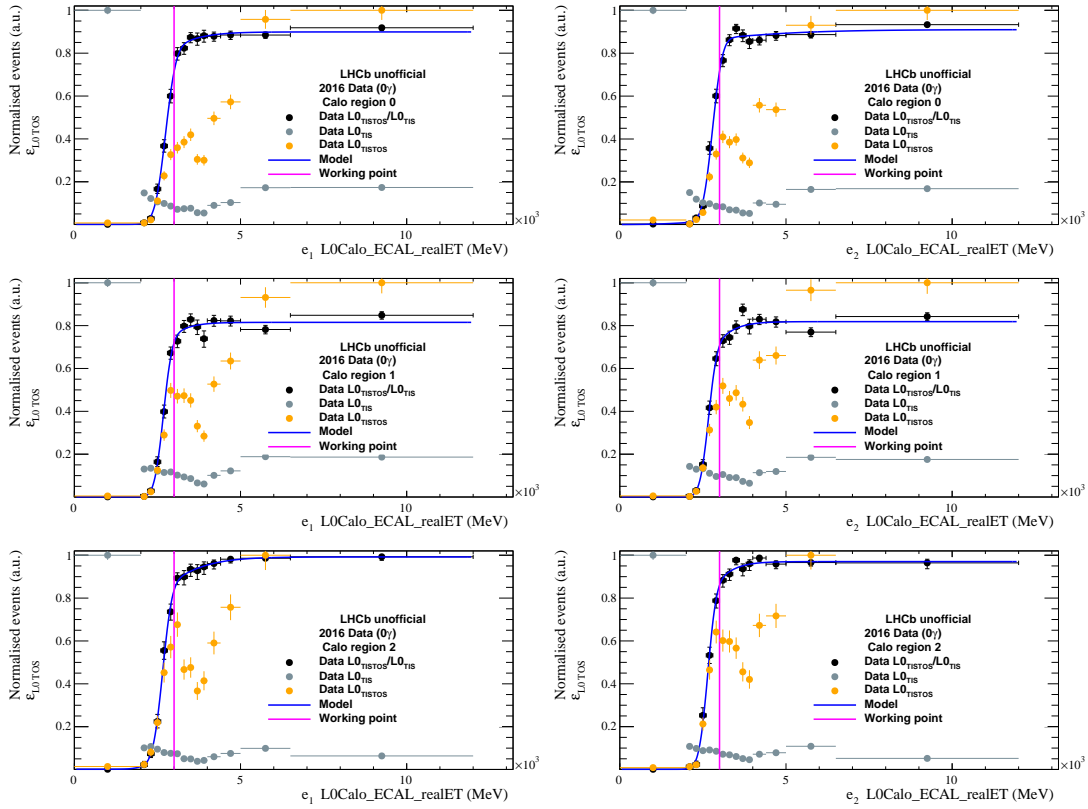


Figure 4.25 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 0γ from $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ data for the data-taking year 2016, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

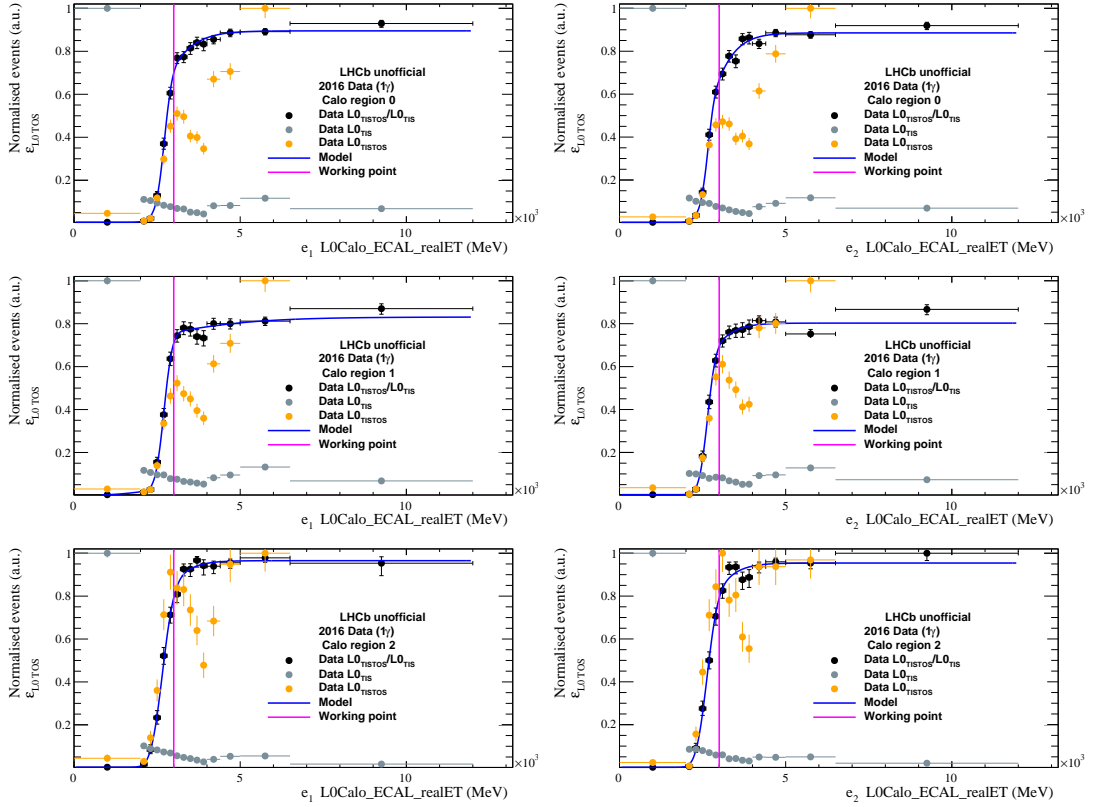


Figure 4.26 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 1γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ data for the data-taking year 2016, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

4.5.5 Kinematic corrections

Significant discrepancies are observed when comparing the distributions of B kinematic variables, such as the B momentum, in data and simulation. This indicates modeling deficiencies in simulation. A kinematic reweighting of the simulated events is therefore performed using event weights, ω_{BKIN} , accounting for the observed deviations. However, since $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ has not been discovered yet the discrepancies between data and simulation cannot be investigated directly with the signal decay modes and an alternative approach is used, deriving the kinematic weights from $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-)K^+$ data and simulation.

A comparison using uncorrected $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-)K^+$ simulation and background-subtracted $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-)K^+$ data obtained with the sPlot technique [85] reveals large discrepancies in the distributions of the B^+ momentum, transverse momentum, pseudorapidity, as well as the track multiplicity nTracks (see Fig. 4.27 (left) for the year 2016, and App. A.3.5 for the years 2017 and 2018).

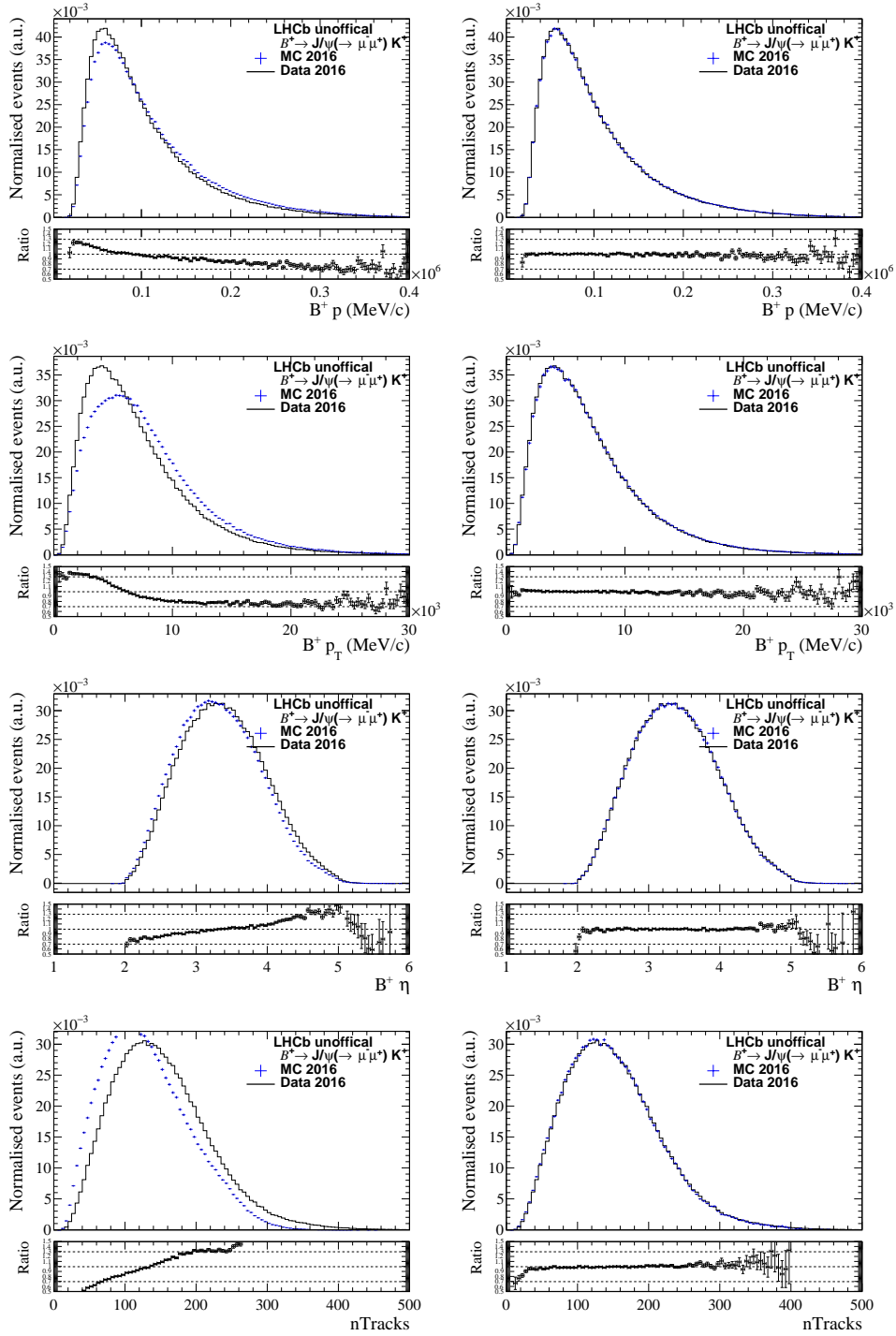


Figure 4.27 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ background-subtracted data (black histogram) and simulation (blue points) for the data-taking year 2016. Left: before any correction. Right: after all corrections.

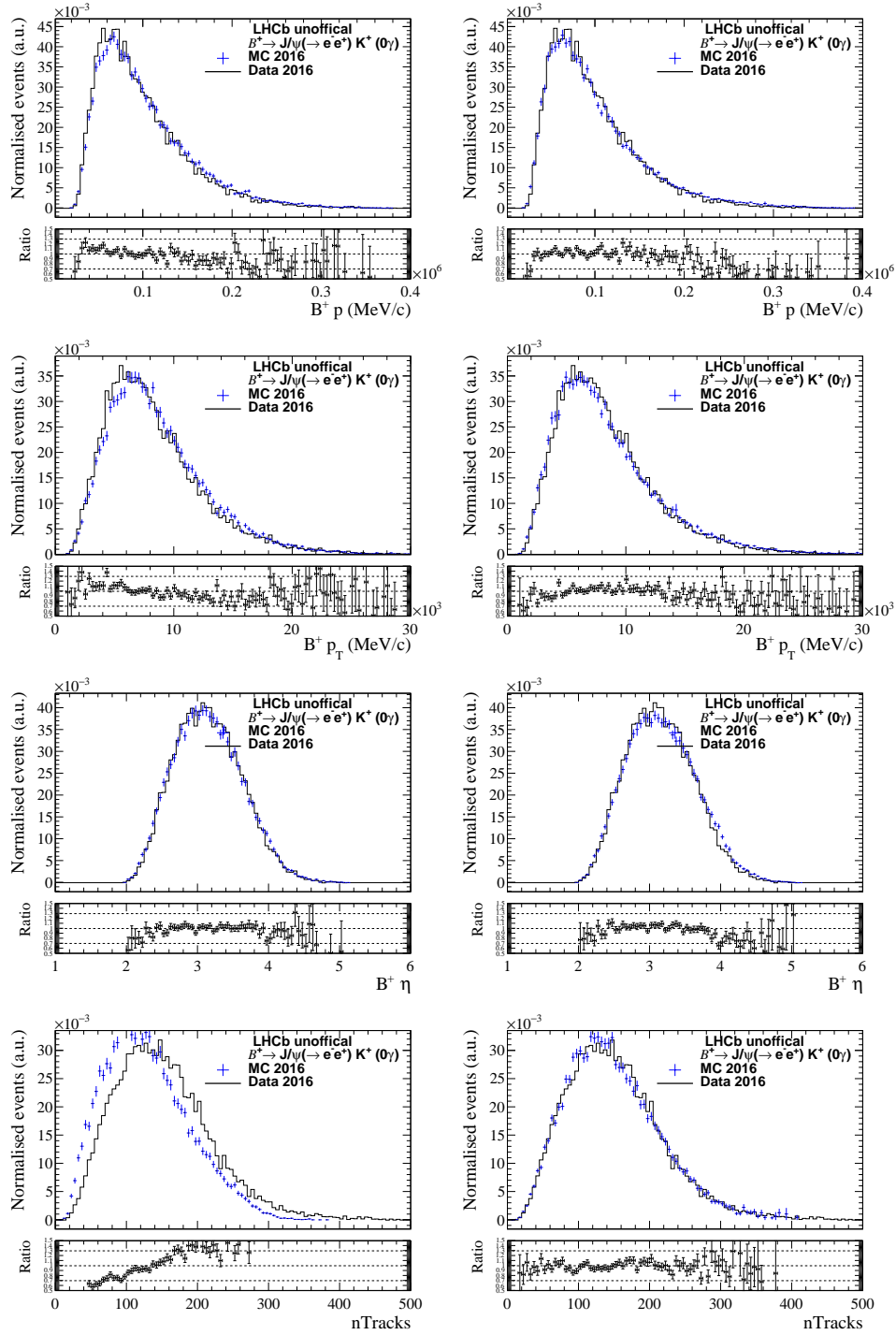


Figure 4.28 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ background-subtracted data (black histogram) and simulation (blue points) for the data-taking year 2016 and bremsstrahlung category 0γ . Left: before any correction. Right: after all correction.

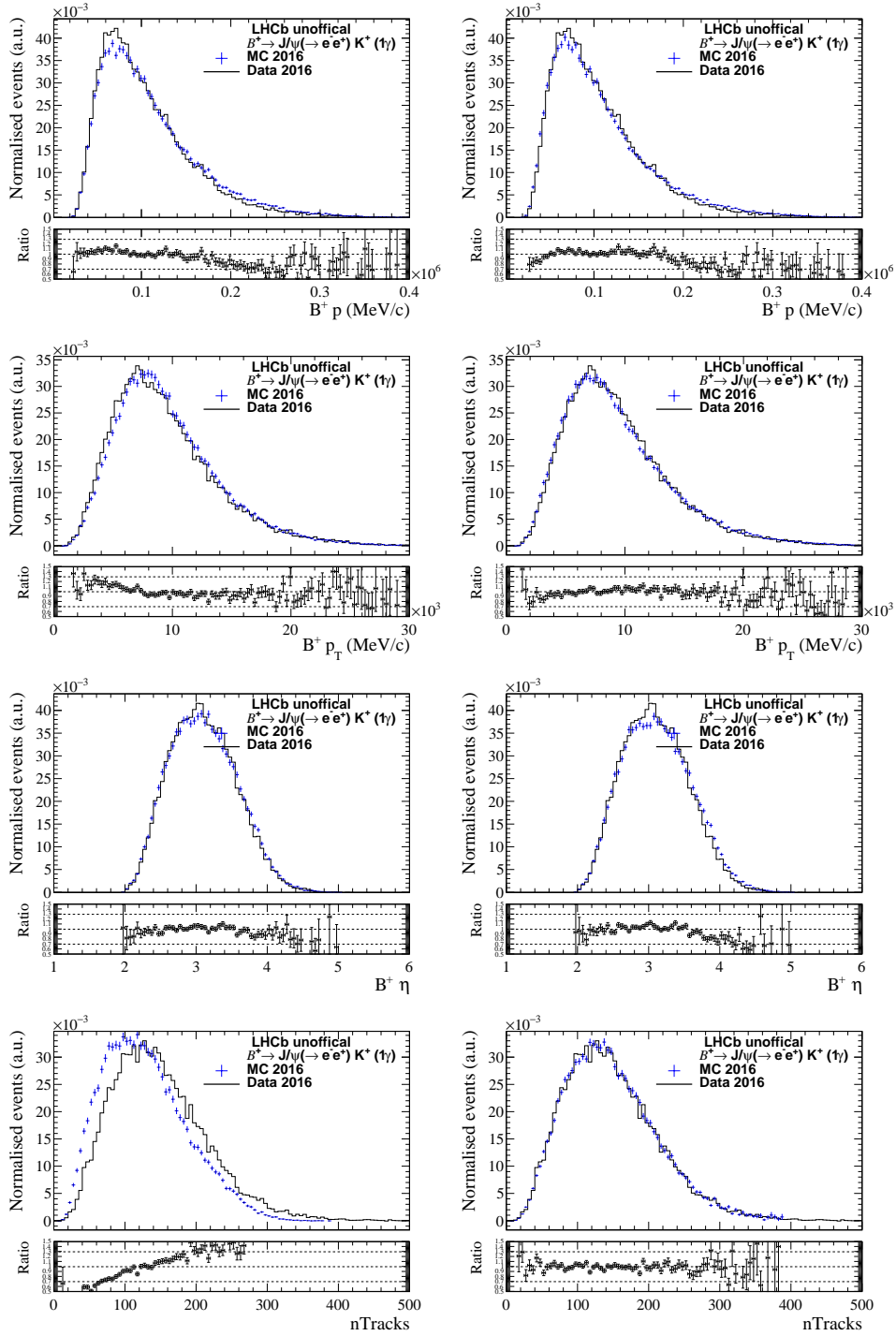


Figure 4.29 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ background-subtracted data (black histogram) and simulation (blue points) for the data-taking year 2016 and bremsstrahlung category 1γ . Left: before any correction. Right: after all corrections.

To account for these discrepancies, event weights are determined with the GBReweighter package [88], where a classifier is trained with background-subtracted $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data and simulation. The B^+ momentum, transverse momentum, pseudorapidity, and nTracks are used as reweighting variables. In simulation, the tracking (ω_{TRK}) and L0 trigger (ω_{L0}) efficiency corrections are taken into account, as well as the data-driven PID efficiencies. A classifier is trained for each data-taking year. The kinematic weights ω_{BKIN} to simulation are determined at generator level, *i.e.* before the stripping selection and event reconstruction. Therefore, the weights are obtained by evaluating the response of the classifier on the true B^+ mesons in the LHCb acceptance before any reconstruction and selection. Then the weights are ported to the candidates after reconstruction and stripping selection, by matching the event and run number at generator level and at reconstruction level.

To study the impact of the obtained corrections, several observables are compared, using fully corrected $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ simulation and background-subtracted $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ data ($\ell = e, \mu$). In the simulation, the tracking, PID, L0 trigger and B kinematic corrections are taken into account, by assigning to each event a combined correction weight, $\omega_{\text{TRK}} \cdot \varepsilon_{\text{PID}} \cdot \omega_{\text{L0}} \cdot \omega_{\text{BKIN}}$, and not applying the PID requirements.

The B^+ momentum, transverse momentum, pseudorapidity, and nTracks distributions in fully corrected $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ simulation are found to be in very good agreement with the corresponding distributions in background-subtracted data, as shown in Fig. 4.27 (right) for the data-taking year 2016. Using the same kinematic corrections weights, the agreement between simulation and data is also satisfactorily improved for the $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ samples, separately in each of the two bremsstrahlung categories 0γ and 1γ , as displayed in Fig. 4.28 (0γ) and Fig. 4.29 (1γ), respectively. The results for the data-taking years 2017 and 2018 are found to be compatible (see App. A.3.5). A similar trend is observed for all data-taking years, with an improvement in the $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ corrected simulation that is not as perfect as for $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$, which can be expected given that the ω_{BKIN} corrections are determined from $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ samples.

4.5.6 Efficiency determination

The total selection efficiency for a given decay channel factorises as

$$\varepsilon = \varepsilon_{\text{geo}} \cdot \varepsilon_{\text{off}}, \quad (4.17)$$

where ε_{geo} is the geometrical acceptance (see Sec. 4.5.1) and ε_{off} the efficiency calculated from simulated events, considering the stripping selection, the pre-selection, the trigger selection and the PID selection. Without any corrections, ε_{off} can be simply calculated by dividing the number of events passing the full selection ($N_{\text{sel,trig,PID}}$) by the total number of events before

any selection and reconstruction, *i.e.* at generator level (N_{gen}):

$$\varepsilon_{\text{off}} = \frac{N_{\text{sel,trig,PID}}}{N_{\text{gen}}}. \quad (4.18)$$

In order to take into account the event-by-event efficiency corrections for tracking, PID and L0 trigger, the number of events $N_{\text{sel,trig,PID}}$ is replaced by the sum of weights $\omega_{\text{TRK}} \cdot \varepsilon_{\text{PID}} \cdot \omega_{\text{L0}}$ over the $N_{\text{sel,trig}}$ simulated events passing the full selection except the PID requirements. Hence, Eq. (4.18) becomes:

$$\varepsilon_{\text{off}} = \frac{\sum_{i=1}^{N_{\text{sel,trig}}} \omega_{\text{TRK}}^i \cdot \varepsilon_{\text{PID}}^i \cdot \omega_{\text{L0}}^i}{N_{\text{gen}}}. \quad (4.19)$$

In order to take the kinematic corrections into account, which are defined at generator level, the event weight ω_{BKIN} must be incorporated in both the numerator and the denominator of Eq. (4.19). The kinematic weights are determined under global event cut on the number of hits in the Scintillating Pad Detector, $\text{nSPDhits} < 450$. To factorize effects related to this requirement out, ω_{BKIN} is summed in the denominator only over events passing this requirement. Hence the final corrected selection efficiency at the reconstruction level is calculated as follows:

$$\varepsilon_{\text{off}} = \frac{\sum_{i=1}^{N_{\text{sel,trig}}} \omega_{\text{TRK}}^i \cdot \varepsilon_{\text{PID}}^i \cdot \omega_{\text{L0}}^i \cdot \omega_{\text{BKIN}}^i}{\sum_{j=1}^{N_{\text{gen,nSPDhits}<450}} \omega_{\text{BKIN}}^j}. \quad (4.20)$$

For $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$, the weights are defined as

- $\omega_{\text{TRK}} = \omega_{\text{TRK}}(e) \cdot \omega_{\text{TRK}}(\mu)$,
- $\varepsilon_{\text{PID}} = \varepsilon_{\text{PID}}(e) \cdot \varepsilon_{\text{PID}}(\mu)$,
- $\omega_{\text{L0}} = \frac{1 - (1 - \varepsilon_{\text{data}}^{\text{TOS}}(e)) \cdot (1 - \varepsilon_{\text{data}}^{\text{TOS}}(\mu))}{1 - (1 - \varepsilon_{\text{MC}}^{\text{TOS}}(e)) \cdot (1 - \varepsilon_{\text{MC}}^{\text{TOS}}(\mu))}$,

while for $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$, with $\ell = e, \mu$:

- $\omega_{\text{TRK}} = \omega_{\text{TRK}}(\ell^+) \cdot \omega_{\text{TRK}}(\ell^-)$,
- $\varepsilon_{\text{PID}} = \varepsilon_{\text{PID}}(\ell^+) \cdot \varepsilon_{\text{PID}}(\ell^-) \cdot \varepsilon_{\text{PID}}(K^+)$,
- $\omega_{\text{L0}} = \frac{1 - (1 - \varepsilon_{\text{data}}^{\text{TOS}}(\ell^+)) \cdot (1 - \varepsilon_{\text{data}}^{\text{TOS}}(\ell^-))}{1 - (1 - \varepsilon_{\text{MC}}^{\text{TOS}}(\ell^+)) \cdot (1 - \varepsilon_{\text{MC}}^{\text{TOS}}(\ell^-))}$.

4.5.7 $r_{J/\psi}$ cross-check

The data-driven efficiency corrections are validated by measuring the branching fraction ratio $r_{J/\psi}$ of Eq. (4.4), which is measured experimentally to be 0.9983 ± 0.0078 [13].

The $r_{J/\psi}$ ratio is computed for each data-taking year in three different bremsstrahlung cate-

4.5. Efficiencies and corrections

Table 4.16 – Measured $r_{J/\psi}$ ratios for different bremsstrahlung categories and corrections. The quoted uncertainties are from the statistical uncertainties on the fitted $B^+ \rightarrow J/\psi K^+$ yields.

| Year | Correction | 0γ | 1γ | 2γ | $0\gamma+1\gamma+2\gamma$ |
|------|--|-------------------|-------------------|-------------------|---------------------------|
| 2016 | uncorrected | 1.298 ± 0.015 | 1.442 ± 0.013 | 1.454 ± 0.016 | 1.422 ± 0.010 |
| 2016 | ω_{TRK} | 1.286 ± 0.015 | 1.432 ± 0.013 | 1.447 ± 0.016 | 1.412 ± 0.010 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}^{\text{no kaon}}$ | 1.209 ± 0.014 | 1.356 ± 0.012 | 1.378 ± 0.016 | 1.337 ± 0.009 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}$ | 1.211 ± 0.014 | 1.358 ± 0.012 | 1.378 ± 0.016 | 1.339 ± 0.009 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}}$ | 1.142 ± 0.014 | 1.231 ± 0.011 | 1.197 ± 0.014 | 1.212 ± 0.008 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}} \cdot \omega_{L0}$ | 1.075 ± 0.013 | 1.156 ± 0.011 | 1.122 ± 0.013 | 1.138 ± 0.008 |
| 2017 | uncorrected | 1.260 ± 0.013 | 1.359 ± 0.011 | 1.405 ± 0.014 | 1.352 ± 0.008 |
| 2017 | ω_{TRK} | 1.261 ± 0.013 | 1.362 ± 0.011 | 1.410 ± 0.014 | 1.355 ± 0.008 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}^{\text{no kaon}}$ | 1.138 ± 0.012 | 1.241 ± 0.010 | 1.297 ± 0.013 | 1.235 ± 0.008 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}$ | 1.141 ± 0.012 | 1.244 ± 0.010 | 1.299 ± 0.013 | 1.238 ± 0.008 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}}$ | 1.071 ± 0.011 | 1.125 ± 0.009 | 1.134 ± 0.012 | 1.119 ± 0.007 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}} \cdot \omega_{L0}$ | 1.045 ± 0.011 | 1.094 ± 0.009 | 1.098 ± 0.012 | 1.088 ± 0.007 |
| 2018 | uncorrected | 1.318 ± 0.013 | 1.421 ± 0.011 | 1.458 ± 0.014 | 1.411 ± 0.008 |
| 2018 | ω_{TRK} | 1.317 ± 0.013 | 1.423 ± 0.011 | 1.463 ± 0.014 | 1.413 ± 0.008 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}^{\text{no kaon}}$ | 1.155 ± 0.011 | 1.267 ± 0.010 | 1.320 ± 0.013 | 1.258 ± 0.007 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}$ | 1.156 ± 0.011 | 1.268 ± 0.010 | 1.321 ± 0.013 | 1.259 ± 0.007 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}}$ | 1.089 ± 0.011 | 1.146 ± 0.009 | 1.146 ± 0.011 | 1.137 ± 0.007 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}} \cdot \omega_{L0}$ | 1.053 ± 0.011 | 1.101 ± 0.009 | 1.092 ± 0.011 | 1.092 ± 0.007 |

gories (0γ , 1γ and 2γ), as wells as for the combined case, using the measured $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ yields of Table 4.14 and the selection efficiencies determined from simulation. The various efficiency corrections are applied in sequence in order to study their impact on $r_{J/\psi}$. The obtained corrected efficiencies can be found in App. A.3.6 and the $r_{J/\psi}$ results, taking only the statistical uncertainties into account, are shown in Table 4.16.

The efficiency corrections to the simulation, especially the L0 trigger and B kinematic corrections are found to be the largest correction factors. Nevertheless, the fully corrected $r_{J/\psi}$ values still show a residual deviation from unity of approximately 10%. One should keep in mind that, the applied selections of $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ decays differ, which results in non identical coverages of the phase-space. Furthermore, imperfections in the corrections to simulation, which may occur from phase-space regions not relevant for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$, can have a magnified effect on the $r_{J/\psi}$ measurement. Indeed, a bias in the electron PID efficiency and L0 trigger electron corrections, for example, would affect $r_{J/\psi}$ twice, if the two electron candidates of $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ are strongly correlated in the biased phase-space region, while the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ efficiency would be affected only once. Consequently, the power of the $r_{J/\psi}$ cross-check performed here is limited, although still useful to gain confidence in the validity of the applied corrections to simulation. The systematic uncertainties due to the miscalibration of the selection efficiencies can be determined from this $r_{J/\psi}$ test, assuming that the residual discrepancy occurs from calibration mismodeling and are propagated to the

$B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ results, as described in Sec. 4.9.

4.5.8 B_s^0 mass eigenstates and lifetime correction

Flavour mixing [41] of the B^0 and B_s^0 mesons, results in two different $B_{(s)}^0$ mass eigenstates, denoted as $B_{(s),L}^0$ (light) and $B_{(s),H}^0$ (heavy). The mass eigenstates have different decay lifetimes, affecting the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiencies and thus the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction measurements and upper limits. However, since no significant decay-width difference has been observed in the B^0 system, flavour mixing in this analysis is only considered for the B_s^0 meson.

In general, B_s^0 mesons decaying to a given final state consist of a mixture of $B_{s,L}^0$ and $B_{s,H}^0$ eigenstates, with lifetimes $\tau_L = (1.429 \pm 0.007)$ ps and $\tau_H = (1.624 \pm 0.009)$ ps [13], respectively. In the case of $B_s^0 \rightarrow e^\pm \mu^\mp$, this mixture is unknown, especially if New Physics would be at play.⁹ Therefore, the effective lifetime of $B_s^0 \rightarrow e^\pm \mu^\mp$ decays can, a priori, be anywhere between τ_L and τ_H . Hence the $B_s^0 \rightarrow e^\pm \mu^\mp$ search is repeated under these two extreme assumptions. For this, the $B_s^0 \rightarrow e^\pm \mu^\mp$ simulated decays, which are generated with a B_s^0 lifetime of $\tau_{\text{gen}} = (1.512 \pm 0.007)$ ps [89], are assigned new event weights

$$\omega_{\tau_{L,H}} = \frac{e^{-\frac{\tau}{\tau_{L,H}}}}{e^{-\frac{\tau}{\tau_{\text{gen}}}}}, \quad (4.21)$$

where τ is the proper B_s^0 lifetime of an event in the LHCb acceptance at generator level. These weights are then used to compute the $B_s^0 \rightarrow e^\pm \mu^\mp$ efficiencies using a modified version of Eq. (4.20):

$$\varepsilon_{\text{off,L,H}}(B_s^0 \rightarrow e^\pm \mu^\mp) = \frac{\sum_{i=1}^{N_{\text{sel,trig}}} \omega_{\tau_{L,H}}^i \cdot \omega_{\text{TRK}}^i \cdot \varepsilon_{\text{PID}}^i \cdot \omega_{\text{L0}}^i \cdot \omega_{\text{BKIN}}^i}{\sum_{j=1}^{N_{\text{gen,nSPDhits}<450}} \omega_{\tau_{L,H}}^j \cdot \omega_{\text{BKIN}}^j}. \quad (4.22)$$

4.5.9 $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-)$ efficiencies

The $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal and $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ selection efficiencies, shown in Table 4.17, are determined from fully corrected simulation, following the strategy presented in Sec. 4.5.6, and averaged over all data-taking years, using the integrated luminosity as weights.

The efficiencies of $B^0 \rightarrow e^\pm \mu^\mp$, and $B_s^0 \rightarrow e^\pm \mu^\mp$, for the three lifetimes τ_{gen} , τ_L and τ_H , are computed separately for the two bremsstrahlung categories and the three BDT regions introduced in Sec. 4.3.8, which correspond to the six samples used in the simultaneous invariant mass fit presented in Sec. 4.7.1.

⁹In the Standard Model, the $B_s^0 \rightarrow e^\pm \mu^\mp$ decay is expected to be dominated by the heavy eigenstate, like the $B_s^0 \rightarrow \mu^+ \mu^-$ decay.

Table 4.17 – Corrected $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiencies, averaged over all data-taking years, using the integrated luminosity as weights. Only statistical uncertainties are shown.

| Decay mode | | ε [%] | | | |
|--|-------------------|-------------------|------------------------------------|------------------------------------|------------------------------------|
| $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ | | 1.699 ± 0.003 | | | |
| | Lifetime [ps] | Brem | $\varepsilon_{\text{BDTbin0}}$ [%] | $\varepsilon_{\text{BDTbin1}}$ [%] | $\varepsilon_{\text{BDTbin2}}$ [%] |
| $B^0 \rightarrow e^\pm \mu^\mp$ | | 0γ | 0.2612 ± 0.0010 | 0.2488 ± 0.0009 | 0.2425 ± 0.0009 |
| $B^0 \rightarrow e^\pm \mu^\mp$ | | 1γ | 0.3312 ± 0.0011 | 0.3458 ± 0.0011 | 0.3416 ± 0.0011 |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.512 ± 0.007 | 0γ | 0.2684 ± 0.0014 | 0.2576 ± 0.0013 | 0.2566 ± 0.0013 |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.512 ± 0.007 | 1γ | 0.3394 ± 0.0016 | 0.3529 ± 0.0016 | 0.3659 ± 0.0016 |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.429 ± 0.007 | 0γ | 0.2661 ± 0.0014 | 0.2509 ± 0.0014 | 0.2397 ± 0.0013 |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.429 ± 0.007 | 1γ | 0.3372 ± 0.0016 | 0.3445 ± 0.0016 | 0.3426 ± 0.0016 |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.624 ± 0.009 | 0γ | 0.3081 ± 0.0015 | 0.3025 ± 0.0015 | 0.3174 ± 0.0015 |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.624 ± 0.009 | 1γ | 0.3887 ± 0.0017 | 0.4130 ± 0.0018 | 0.4516 ± 0.0019 |

4.6 Physics background

4.6.1 Expectations from corrected simulation

The background from specific physics decays needs to be studied, since these decay modes can lead to a noticeable reduction in the signal sensitivity. The most concerning physics background sources when reconstructing $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ candidates can occur from hadronic two-body b -hadron decays, such as $B_{(s)}^0, \Lambda_b^0 \rightarrow h^+ h^-$, where $h = \pi, K, p$. Such decay modes can have a distinctive peaking behaviour within the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal mass range. Additionally, semi-leptonic decays, such as $B^0 \rightarrow \pi^+ \ell^- \bar{\nu}_\ell$ and $\Lambda_b^0 \rightarrow p \ell^- \bar{\nu}_\ell$ ($\ell = e, \mu$), need to be taken into account, since the corresponding branching fractions are relatively large. In order to suppress these backgrounds, the PID requirements introduced in Sec. 4.3.6 are applied.

The expected pollution for each background decay j surviving the event selection is estimated for each fit sample k (corresponding to one of the 2 bremsstrahlung categories and one of the 3 BDT regions introduced in Sec. 4.3.8) using simulated events, where the backgrounds are reconstructed as $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ candidates. The estimation is performed with respect to the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ decay mode as

$$\mathcal{N}_{j,\text{exp}}^k = \frac{f_j}{f_u} \cdot \frac{\mathcal{B}_j}{\mathcal{B}(B^+ \rightarrow J/\psi K^+) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)} \cdot \frac{\varepsilon_j^k}{\varepsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)} \cdot \mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+), \quad (4.23)$$

where ε_j^k denotes the selection efficiency for a specific background mode j in category k in the full mass range, and $\varepsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+) = (1.699 \pm 0.003)\%$ the selection efficiency

Table 4.18 – Branching fractions and ratio of fragmentation fractions, used in the estimation of the physics background levels, as well as expected number of physics background events in the 0γ and 1γ bremsstrahlung categories, summed over the three BDT regions in the full mass region. A dash indicates that no event of the simulated sample passed the selection.

| Background mode j | \mathcal{B}_j [13] | f_j/f_u [75, 90] | $\mathcal{N}_{j,\text{exp}}^{0\gamma}$ | $\mathcal{N}_{j,\text{exp}}^{1\gamma}$ |
|---|---------------------------------|---------------------|--|--|
| $B^0 \rightarrow K^+ \pi^-$ | $(1.96 \pm 0.05) \cdot 10^{-5}$ | 1 | 12.2 ± 2.4 | – |
| $B^0 \rightarrow K^+ K^-$ | $(7.8 \pm 1.5) \cdot 10^{-8}$ | | 0.011 ± 0.004 | – |
| $B^0 \rightarrow \pi^+ \pi^-$ | $(5.12 \pm 0.19) \cdot 10^{-6}$ | | 7.7 ± 1.0 | – |
| $B^0 \rightarrow p \bar{p}$ | $(1.25 \pm 0.32) \cdot 10^{-8}$ | | $(1.8 \pm 0.8) \cdot 10^{-5}$ | – |
| $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ | $(1.50 \pm 0.06) \cdot 10^{-4}$ | | 11 ± 4 | 26 ± 9 |
| $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ | $(1.50 \pm 0.06) \cdot 10^{-4}$ | | 94 ± 4 | – |
| $B_s^0 \rightarrow \pi^+ K^-$ | $(5.8 \pm 0.7) \cdot 10^{-6}$ | 0.2539 ± 0.0079 | 0.96 ± 0.18 | – |
| $B_s^0 \rightarrow K^+ K^-$ | $(2.66 \pm 0.22) \cdot 10^{-5}$ | | 0.93 ± 0.30 | – |
| $B_s^0 \rightarrow \pi^+ \pi^-$ | $(7.0 \pm 1.0) \cdot 10^{-7}$ | | 0.27 ± 0.04 | – |
| $B_s^0 \rightarrow p \bar{p}$ | $< 1.5 \cdot 10^{-8}$ (90% CL) | | $(6 \pm 4) \cdot 10^{-6}$ | – |
| $\Lambda_b^0 \rightarrow p \pi^-$ | $(4.5 \pm 0.8) \cdot 10^{-6}$ | 0.518 ± 0.036 | 0.23 ± 0.08 | – |
| $\Lambda_b^0 \rightarrow p K^-$ | $(5.4 \pm 1.0) \cdot 10^{-6}$ | | 0.20 ± 0.07 | – |
| $\Lambda_b^0 \rightarrow p e^- \bar{\nu}_e$ | $(4.1 \pm 1.0) \cdot 10^{-4}$ | | 0.0015 ± 0.0015 | 1.2 ± 1.0 |
| $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ | $(4.1 \pm 1.0) \cdot 10^{-4}$ | | 19 ± 4 | – |

determined for the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ normalisation mode. All selection efficiencies are determined from simulation, taking into account the full selection, as well as all event-by-event correction weights obtained from data-driven methods, as explained in Sec. 4.5. Furthermore, the selection efficiencies are weighted by the integrated luminosity of each data-taking year. The yield of the normalisation mode, merged over the three data-taking years, is $\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+) = 1\,722\,467 \pm 1\,340$, as determined in Sec. 4.4 (see Table 4.14). The normalisation branching fractions $\mathcal{B}(B^+ \rightarrow J/\psi K^+) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-) = (1.020 \pm 0.019) \cdot 10^{-3} \cdot (5.961 \pm 0.033) \cdot 10^{-2}$ are taken from Ref. [13]. The assumed fragmentation fraction ratios f_j/f_u and background branching fractions are listed in Table 4.18.

The focus of this background study is on one hand on the identification of the peaking backgrounds in the signal mass window, $m_{\text{DTF}}(e^\pm \mu^\mp) \in [5.1, 5.9] \text{ GeV}/c^2$, on the other hand on the determination of the expected backgrounds in the mass sidebands, $m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.9, 5.1] \cup [5.6, 6.1] \text{ GeV}/c^2$, which might be absorbed by the modeling of the combinatorial background with an exponential. Thus, the expected physics background is estimated in these different mass regions, as well as in the full mass range, $m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.9, 61] \text{ GeV}/c^2$. The background selection efficiencies ε_j^k are determined for each mass range as reported in App. A.4.1.¹⁰

The expected numbers of events for each background decay are computed for all three mass

¹⁰No samples of $B^0 \rightarrow K^+ K^-$, $B_s^0 \rightarrow p \bar{p}$ and $B_s^0 \rightarrow \pi^+ \pi^-$ decays have been simulated; the efficiencies of these modes are assumed to be the same as those of $B_s^0 \rightarrow K^+ K^-$, $B^0 \rightarrow p \bar{p}$ and $B^0 \rightarrow \pi^+ \pi^-$, respectively.

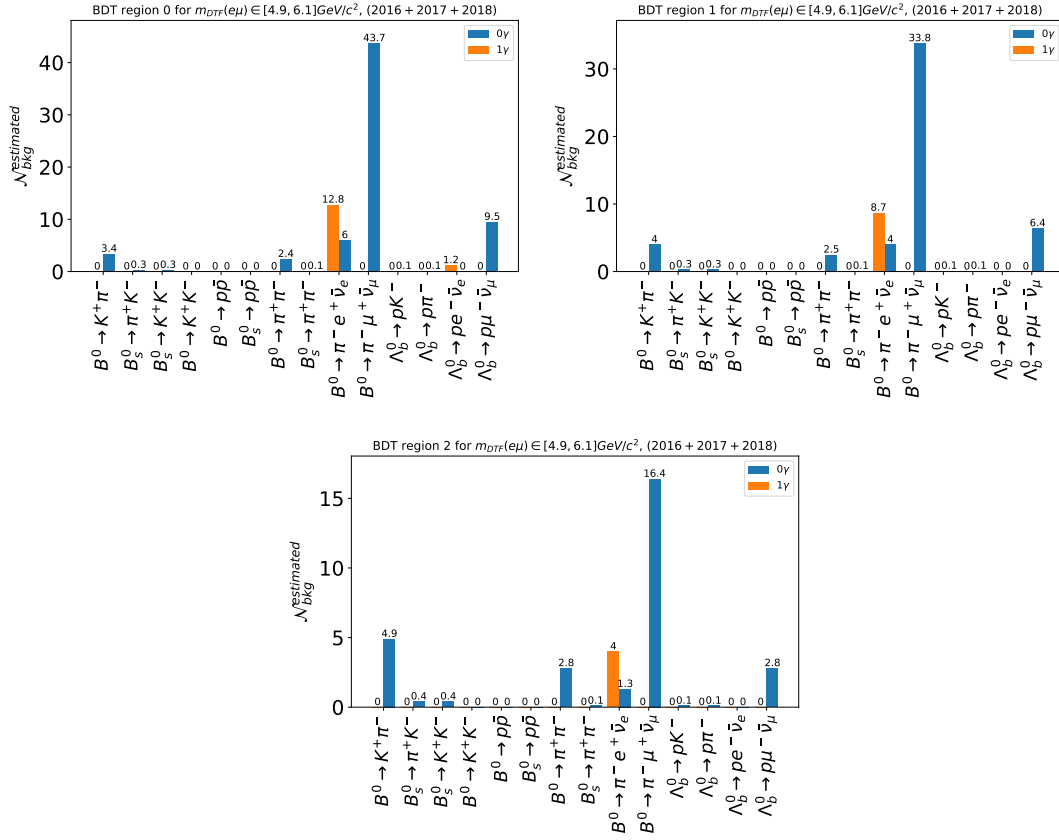


Figure 4.30 – Expected number of physics background events in each sample (0 γ and 1 γ in the 3 BDT regions), as determined from corrected simulation in the full mass range ($m_{DTR}(e^\pm\mu^\mp) \in [4.9, 6.1] \text{ GeV}/c^2$).

ranges and shown in App. A.4.1. In Fig. 4.30, The expected background events $\mathcal{N}_{j,\text{exp}}^k$ for the full mass range are shown in Fig. 4.30 and summed over all BDT regions in Table 4.18.

The most dangerous backgrounds occur from $B^0 \rightarrow K^+\pi^-$ and $B^0 \rightarrow \pi^+\pi^-$ candidates in bremsstrahlung category 0 γ , while no peaking $B_{(s)}^0, \Lambda_b^0 \rightarrow h^+h^-$ backgrounds are expected in bremsstrahlung category 1 γ . This can be explained by the recovery of the bremsstrahlung photon in the 1 γ category, which almost certainly refers to a true electron candidate. However, the estimated background contribution arising from the semi-leptonic decays $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$, $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$, and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ is significantly larger, but populate mainly the sideband mass regions (see Table A.8 in App. A.4.1). The $B^0 \rightarrow K^+\pi^-$, $B^0 \rightarrow \pi^+\pi^-$, $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ components will be modeled in the final fit of the $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$ branching fractions and their yields will be constrained to their expectations (see Sec. 4.7). The $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ is not considered, since it is assumed that this component can be absorbed by the model of the combinatorial background.

To validate these background estimations, two independent methods are used, as described

in the following sections.

Despite the here studied backgrounds, other decay modes such as $B_c^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-)$ or $B_s^0 \rightarrow K^+ \ell^- \bar{\nu}_\ell$ ($\ell = \mu, e$) might additional pollute the samples. The study of the impact of these decay modes is discussed in Sec. 4.9.

4.6.2 Validation of the expected peaking backgrounds with $B_{(s)}^0 \rightarrow hh'$ decays

To cross-check the rate of true pions or kaons misidentified as muons or electrons and therefore to test the accuracy of the predicted physics background of Sec. 4.7.4, a data-driven approach is used, based on $B_{(s)}^0 \rightarrow hh'$ data and simulation.

Data samples of $B^0 \rightarrow K^+ \pi^-$ candidates are built by selecting one of the two hadrons, called Tag, as a kaon (or pion) with a tight PID requirement, $\text{ProbNN}k > 0.8$ (or $\text{ProbNN}\pi > 0.8$), and assigning the kaon (or pion) mass to it. The misidentification rate is tested with the second hadron, denoted as Probe. The Probe is assigned to the pion (or kaon) mass and must pass the electron or muon PID requirement introduced in Sec. 4.3.6 for the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection.

Fits are then performed to the $m(K^+ \pi^-)$ mass distributions, to measure the yields of several $B_{(s)}^0 \rightarrow hh'$ components. In addition, the expected yield of each fit component is estimated from $B_{(s)}^0 \rightarrow hh'$ simulation, taking the data-driven PID efficiencies of Sec. 4.5.3 into account, following the same approach used to estimate the physics background in Sec. 4.6.1. The selection of the $B^0 \rightarrow K^+ \pi^-$ samples, the $K^+ \pi^-$ mass fits, and the yield estimation from simulation is described in App. A.4.2.

Table 4.19 (4.20) compares the estimated $B_{(s)}^0 \rightarrow hh'$ yields from the corrected simulation with the yields observed in data, when tagging a kaon and the Probe is subject to the electron (muon) PID requirement, while Table 4.21 (4.22) corresponds to the case when the Tag is a pion.

The analysis of the results obtained for the data-taking year 2016 reveals that the expected number of events matches almost perfectly the predictions. For the data-taking years 2017 and 2018 the results are still compatible. However, the results for the three data-taking year show partially very different trends. This might occur from several sources. Firstly, for different years, different trigger configurations are used, which can affect the behaviour of the L0 hadron trigger efficiency. Secondly, the `L0HadronDecision_TOS` trigger response is not corrected in this study and is assumed to be correctly described in simulation, when the transverse energy in the hadronic calorimeter is above 4000 MeV. Furthermore, studies presented here are performed with very low statistics samples, which results in large statistical uncertainties.

Table 4.19 – Yields of various $B_{(s)}^0 \rightarrow hh'$ decays expected from the corrected simulation (\mathcal{N}_{MC}) and observed in data ($\mathcal{N}_{\text{data}}$) when one of the two tracks (Tag) is selected as a kaon and the other (Probe) as an electron.

| Year | Decay | \mathcal{N}_{MC} | $\mathcal{N}_{\text{data}}$ | $\frac{\mathcal{N}_{\text{data}}}{\mathcal{N}_{\text{MC}}}$ |
|------|-------------------------------|---------------------------|-----------------------------|---|
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 136.1 ± 9.4 | 116.0 ± 12.5 | 0.85 ± 0.11 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 12.0 ± 2.4 | 9.6 ± 3.0 | 0.80 ± 0.29 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 10.9 ± 1.5 | 8.8 ± 2.1 | 0.81 ± 0.22 |
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 3.4 ± 0.7 | 2.8 ± 0.90 | 0.84 ± 0.32 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 136.6 ± 9.2 | 128 ± 12 | 0.94 ± 0.11 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 11.9 ± 1.8 | 10.6 ± 2.6 | 0.90 ± 0.26 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 10.9 ± 1.5 | 10.2 ± 1.7 | 0.93 ± 0.19 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 3.1 ± 0.6 | 2.9 ± 0.9 | 0.92 ± 0.35 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 145 ± 14 | 144 ± 14 | 0.99 ± 0.13 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 12.4 ± 2.5 | 10.9 ± 3.7 | 0.90 ± 0.36 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 11.6 ± 1.8 | 11.0 ± 2.0 | 0.95 ± 0.22 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 3.6 ± 0.8 | 3.5 ± 1.2 | 0.96 ± 0.40 |

Table 4.20 – Yields of various $B_{(s)}^0 \rightarrow hh'$ decays expected from the corrected simulation (\mathcal{N}_{MC}) and observed in data ($\mathcal{N}_{\text{data}}$) when one of the two tracks (Tag) is selected as a kaon and the other (Probe) as a muon.

| Year | Decay | \mathcal{N}_{MC} | $\mathcal{N}_{\text{data}}$ | $\frac{\mathcal{N}_{\text{data}}}{\mathcal{N}_{\text{MC}}}$ |
|------|-------------------------------|---------------------------|-----------------------------|---|
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 97.1 ± 7.7 | 95.1 ± 12.8 | 0.98 ± 0.15 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 63.5 ± 7.4 | 56.9 ± 12.7 | 0.90 ± 0.23 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 8.6 ± 1.2 | 8.3 ± 1.6 | 0.96 ± 0.23 |
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 2.5 ± 0.6 | 2.46 ± 0.92 | 0.97 ± 0.43 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 96.4 ± 7.5 | 115 ± 15 | 1.2 ± 0.18 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 62.0 ± 6.5 | 66 ± 14 | 1.07 ± 0.25 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 8.1 ± 1.2 | 9.7 ± 1.8 | 1.21 ± 0.28 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 2.50 ± 0.58 | 3.0 ± 1.1 | 1.18 ± 0.51 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 111.01 ± 12.05 | 127 ± 15 | 1.14 ± 0.18 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 66.13 ± 7.97 | 79.0 ± 13.2 | 1.19 ± 0.25 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 8.55 ± 1.4 | 9.3 ± 2.8 | 1.08 ± 0.37 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 2.85 ± 0.7 | 3.3 ± 1.3 | 1.15 ± 0.54 |

Table 4.21 – Yields of various $B_{(s)}^0 \rightarrow hh'$ decays expected from the corrected simulation (\mathcal{N}_{MC}) and observed in data ($\mathcal{N}_{\text{data}}$) when one of the two tracks (Tag) is selected as a pion and the other (Probe) as an electron.

| Year | Decay | \mathcal{N}_{MC} | $\mathcal{N}_{\text{data}}$ | $\frac{\mathcal{N}_{\text{data}}}{\mathcal{N}_{\text{MC}}}$ |
|------|-------------------------------|---------------------------|-----------------------------|---|
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 54.9 ± 3.5 | 53 ± 16 | 0.97 ± 0.29 |
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 14.5 ± 2.8 | 13.6 ± 3.4 | 0.94 ± 0.30 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.2 ± 0.3 | 1.1 ± 0.4 | 0.94 ± 0.39 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 0.63 ± 0.59 | 0.59 ± 0.50 | 0.95 ± 1.19 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 57.1 ± 3.6 | 63.2 ± 22.1 | 1.11 ± 0.39 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 14.7 ± 2.7 | 18.8 ± 4.8 | 1.28 ± 0.40 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.20 ± 0.26 | 1.55 ± 0.53 | 1.29 ± 0.53 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 0.54 ± 0.35 | 0.68 ± 0.68 | 1.26 ± 1.48 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 61.0 ± 4.0 | 75 ± 24 | 1.23 ± 0.41 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 15.2 ± 4.4 | 17.0 ± 4.3 | 1.12 ± 0.43 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.24 ± 0.38 | 1.44 ± 0.92 | 1.16 ± 0.82 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 0.55 ± 0.64 | 0.63 ± 0.64 | 1.13 ± 1.74 |

Table 4.22 – Yields of various $B_{(s)}^0 \rightarrow hh'$ decays expected from the corrected simulation (\mathcal{N}_{MC}) and observed in data ($\mathcal{N}_{\text{data}}$) when one of the two tracks (Tag) is selected as a pion and the other (Probe) as a muon.

| Year | Decay | \mathcal{N}_{MC} | $\mathcal{N}_{\text{data}}$ | $\frac{\mathcal{N}_{\text{data}}}{\mathcal{N}_{\text{MC}}}$ |
|------|-------------------------------|---------------------------|-----------------------------|---|
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 81.8 ± 6.9 | 83 ± 13 | 1.01 ± 0.19 |
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 47.9 ± 3.2 | 49.4 ± 9.1 | 1.03 ± 0.20 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 6.11 ± 0.91 | 6.0 ± 1.7 | 0.98 ± 0.32 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 3.9 ± 1.3 | 4.0 ± 1.3 | 1.03 ± 0.47 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 87.1 ± 7.1 | 70 ± 14 | 0.81 ± 0.18 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 51.3 ± 3.3 | 41.5 ± 9.3 | 0.81 ± 0.19 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 7.2 ± 1.0 | 5.8 ± 1.5 | 0.81 ± 0.23 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 3.39 ± 0.83 | 2.74 ± 0.82 | 0.81 ± 0.31 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 97 ± 11 | 111 ± 17 | 1.14 ± 0.22 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 57.9 ± 3.9 | 67 ± 12 | 1.16 ± 0.23 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 7.2 ± 1.2 | 8.3 ± 1.9 | 1.15 ± 0.33 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 3.6 ± 1.3 | 4.1 ± 1.5 | 1.15 ± 0.59 |

Considering all these points, it is concluded that for all Tag and Probe configurations, the predictions from simulation and the observed number of events in data are compatible within an accuracy of 10–30%, which can be directly related to the validation of the misidentification rates, obtained with the corrected PID efficiencies. In general, only minor effects of over/under estimation are observed, which however are at a level that is not concerning, considering the statistical uncertainties of the mass fits they refer to.

4.6.3 Pass-fail cross-check validation of peaking backgrounds

The size of the peaking backgrounds is further validated using $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data candidates passing the stripping, pre-selection and trigger selection introduced in Sec. 4.3. For this cross-check data in the different BDT regions and from the different data-taking years are merged. Furthermore, since the results presented in Sec. 4.6.1 show that peaking backgrounds are only expected in the bremsstrahlung category 0γ , only the corresponding events are selected. The pass-fail method is applied using two different datasets of $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ candidates, defined as follows:

1. $\mu_{\text{fail}} e_{\text{fail}}$: the muon track is selected with the `isMuon` requirement applied and the inverted signal muon PID requirement $\text{ProbNN}\mu < 0.4$, while the electron track is selected with PID requirement applied in the stripping, $\text{PID}e > -2$, and the inverted electron PID requirement, $\text{NOT}(\text{ProbNN}e > 0.4 \text{ AND } \text{PID}e > 3)$.
2. $\mu_{\text{pass}} e_{\text{fail}}$: the muon track is selected with the nominal muon PID requirements (`isMuon` and $\text{ProbNN}\mu > 0.4$) and the electron track is selected with the stripping PID requirement ($\text{PID}e > -2$) and inverted muon PID requirement, $\text{NOT}(\text{ProbNN}e > 0.4 \text{ AND } \text{PID}e > 3)$. This dataset is used since with the application of the `isMuon` requirement on the muon final state, most of the misidentified pions or kaons are expected to populate the high values regime of $\text{ProbNN}\mu$, which makes the $\text{ProbNN}\mu$ variable less powerful in the classification of muons with respect to pions, respectively, kaons.

The principle of the pass-fail method is to reweight these datasets to make them representative of the $\mu_{\text{pass}} e_{\text{pass}}$ dataset in the signal mass region. The point is that the $\mu_{\text{pass}} e_{\text{pass}}$ dataset needs to remain blinded, while the $\mu_{\text{fail}} e_{\text{fail}}$ and $\mu_{\text{pass}} e_{\text{fail}}$ datasets can be freely examined. In these samples, the tracks failing the lepton PID requirements (*i.e.* passing the inverted lepton PID requirements) are assumed to be either pions or kaons and are given a weight defined as

$$\omega_{\text{pass/fail}}(h \rightarrow \ell) = \frac{\varepsilon_{\text{pass}}(h \rightarrow \ell)}{\varepsilon_{\text{fail}}(h \rightarrow \ell)}, \quad (4.24)$$

with $h = K, \pi$ and $\ell = e, \mu$, and where $\varepsilon_{\text{pass}}(h \rightarrow \ell)$ and $\varepsilon_{\text{fail}}(h \rightarrow \ell) = 1 - \varepsilon_{\text{pass}}(h \rightarrow \ell)$ are the probabilities that a true hadron h passes or fails the PID requirement applied after the stripping for lepton ℓ . These efficiencies are determined using `PIDCalib2` in bins of the hadron transverse momentum and pseudorapidity, following the approach of Sec. 4.5.3. They are extracted from $D^{*+} \rightarrow D^0(K\pi)\pi^+$ calibration data, where the applied selection to the final state particles is aligned with the selection applied to $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$. Furthermore, to account for correlations of the misidentification rate and the L0 trigger response, the $\omega_{\text{pass/fail}}(h \rightarrow \ell)$ weights are evaluated in two different trigger response categories, once requiring that the `LOMuon_TOS` or `LOElectron_TOS` trigger line has fired, and once that one of the trigger lines has not fired. The resulting weights $\omega_{\text{pass/fail}}(K \rightarrow \mu)$, $\omega_{\text{pass/fail}}(K \rightarrow e)$, $\omega_{\text{pass/fail}}(\pi \rightarrow \mu)$ and $\omega_{\text{pass/fail}}(\pi \rightarrow e)$ are presented in Figs. A.38–A.41 of App. A.5.

In the $\mu_{\text{fail}} e_{\text{fail}}$ dataset, each $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ candidate is assigned a weight is equal to the product

of the weights applied to the muon and electron candidates, $\omega_{\text{pass/fail}} = \omega_{\text{pass/fail}}(h \rightarrow e) \cdot \omega_{\text{pass/fail}}(h \rightarrow \mu)$, while in the $\mu_{\text{pass}} e_{\text{fail}}$ dataset, only the weight of the electron candidate is applied, $\omega_{\text{pass/fail}} = \omega_{\text{pass/fail}}(h \rightarrow e)$. The weights are assigned, depending on the trigger category in which the event falls. The true nature of the selected lepton candidates is not known, *i.e.* one can not distinguish perfectly if a kaon or a pion is misidentified as a lepton. Therefore, if the value of ProbNNk for a lepton candidate exceeds 0.2, the weight $\omega_{\text{pass/fail}}(K \rightarrow \ell)$ is applied, otherwise the weight $\omega_{\text{pass/fail}}(\pi \rightarrow \ell)$ is applied.

For the $\mu_{\text{fail}} e_{\text{fail}}$ and $\mu_{\text{pass}} e_{\text{fail}}$ datasets, an extended unbinned maximum likelihood fit to the reconstructed $m(\pi^+ \pi^-)$ mass distribution is performed, assigning the pion mass hypothesis to both tracks. The fit model includes the following components:

- $B^0 \rightarrow K^+ \pi^-$: modeled with a double-sided Crystal Ball function with shape parameters extracted from simulation;
- $B_s^0 \rightarrow \pi^+ K^-$: modeled with the same shape as used for $B^0 \rightarrow K^+ \pi^-$, except that the mean of the distribution is shifted by the B_s^0 - B^0 mass difference of 87.42 MeV [13];
- $B_s^0 \rightarrow K^+ K^-$: modeled with a double-sided Crystal Ball function with shape parameters extracted from simulation;
- $B^0 \rightarrow \pi^+ \pi^-$: modeled with a double-sided Crystal Ball function with shape parameters extracted from simulation;
- Combinatorial background: modeled with an exponential function.

All shape parameters are fixed in the fit to the values obtained from simulation, except the exponential slope of the combinatorial background. While the yield of the combinatorial background is left floating, the yields of the four $B_{(s)}^0 \rightarrow hh'$ components are fixed to the expectations presented in Table 4.18.

The fitted $m(\pi^+ \pi^-)$ distributions is shown Fig. 4.31 for the reweighed $\mu_{\text{pass}} e_{\text{fail}}$ and $\mu_{\text{fail}} e_{\text{fail}}$ datasets. In both cases, the fit quality is satisfying, indicating that the predicted $B_{(s)}^0 \rightarrow hh'$ yields, estimated from simulation, are in agreement with the observed mass distribution results in data.

This cross-check relies on the reweighing of the same data as used for the final $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction fit, however selected in a control region by applying inverted PID requirements, and thus bypassing the use of simulation. Therefore, the obtained results provide a strong validation, together the cross-check performed in Sec. 4.6.2, that the predicted $B_{(s)}^0 \rightarrow hh'$ backgrounds in Sec. 4.6.1 are valid and the estimates of the peaking backgrounds entering the final $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction fit are unbiased.

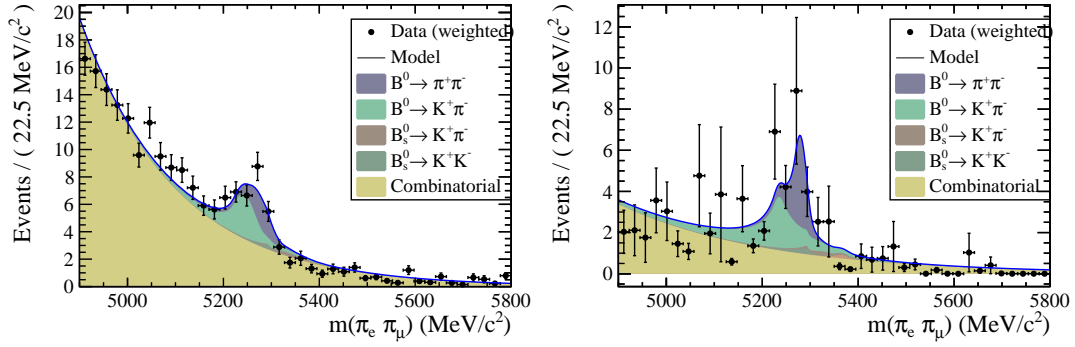


Figure 4.31 – Reweighted $m(\pi^+ \pi^-)$ mass distribution (black data points) of $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ candidates in the bremsstrahlung category 0γ of the $\mu_{\text{pass}} e_{\text{fail}}$ (right) and $\mu_{\text{fail}} e_{\text{fail}}$ (left) datasets, merged over the three data-taking years. In the superimposed fit, the yields of each $B_{(s)}^0 \rightarrow hh'$ component is fixed to expectation from simulation.

4.7 Fit to data

In the following, the extended unbinned maximum likelihood fit of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions and its essential components are presented together with studies, based on pseudo-experiments of the fit stability and pulls. Firstly, the chosen fit strategy is discussed, followed by the introduction of the mass shapes employed in the fit model.

4.7.1 Fit strategy

The two signal branching fractions, $\mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)$, are fitted relative to the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ normalisation decay mode, simultaneously on six data samples, each corresponding to one of the two bremsstrahlung categories (0γ and 1γ) and to one of three regions of the BDT response (see Sec. 4.3.8). The simultaneous fit is performed as an extended unbinned maximum likelihood fit to the invariant $B_{(s)}^0$ mass distribution in the range $m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.9, 6.1] \text{ GeV}/c^2$. To ensure fit stability, the data samples are merged over all data-taking years.

From Eq.(4.2) the signal yield in each sample ($k = 1, \dots, 6$) can be expressed as

$$\mathcal{N}^k(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) = \frac{f_{d(s)}}{f_u} \cdot \frac{\varepsilon^k(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)}{\varepsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)} \cdot \frac{\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+) \cdot \mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)}{\mathcal{B}(B^+ \rightarrow J/\psi K^+) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)}, \quad (4.25)$$

where the total selection efficiencies, ε , for the signal and normalisation decay modes, shown in Table 4.17 in Sec. 4.5.9, are determined from simulation, taking into account all the data-driven corrections of Sec. 4.5 and then averaged over all data-taking years using the integrated luminosity as weights. The ratio of the fragmentation fractions, $\frac{f_{d(s)}}{f_u}$, the normalisation branching fractions, $\mathcal{B}(B^+ \rightarrow J/\psi K^+) \cdot \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$, as well as the normalisation yield

Table 4.23 – Normalisation yield, normalisation branching fractions, and fragmentation function ratios at 13 TeV entering the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ fit.

| Parameter | Value |
|--|--|
| $\mathcal{N}(B^+ \rightarrow (J\psi \rightarrow \mu^+ \mu^-) K^+)$ | 1722467 ± 1340 (Table 4.14) |
| $\mathcal{B}(B^+ \rightarrow J\psi K^+)$ | $(1.020 \pm 0.019) \cdot 10^{-3}$ [13] |
| $\mathcal{B}(J\psi \rightarrow \mu^+ \mu^-)$ | $(5.961 \pm 0.033) \cdot 10^{-2}$ [13] |
| f_d/f_u (at 13 TeV) | 1 |
| f_s/f_u (at 13 TeV) | 0.254 ± 0.008 [75] |

$\mathcal{N}(B^+ \rightarrow (J\psi \rightarrow \mu^+ \mu^-) K^+)$ are fixed in the fit to the values presented in Table 4.23. The two signal branching fractions, $\mathcal{B}(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)$, are shared parameters among all six samples and are left floating between -1 and $+1$. The $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal mass shapes are described in each sample by fixed bifurcated double-sided Crystal Ball functions and are discussed in Sec. 4.7.3.

The studies of Sec. 4.6.1, conducted to estimate the expected physics background, show that the most significant background pollution in the signal window is expected to occur from $B^0 \rightarrow K^+ \pi^-$ (12.2 events) and $B^0 \rightarrow \pi^+ \pi^-$ (7.7 events). Furthermore, when considering the full mass range, significant pollution is expected to arise from the semi-leptonic decays $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ (37 events) and $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ (94 events). Consequently, these background decay modes are considered in the fit and their yield in each sample k is expressed as

$$\mathcal{N}_j^k = \frac{\mathcal{N}_{j,\text{exp}}^k}{\sum_{k'} \mathcal{N}_{j,\text{exp}}^{k'}} \cdot \mathcal{N}_j^{\text{tot}}. \quad (4.26)$$

The index $j = 1, \dots, 4$ represents the background decay modes $B^0 \rightarrow K^+ \pi^-$, $B^0 \rightarrow \pi^+ \pi^-$, $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ and $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$. $\mathcal{N}_{j,\text{exp}}^k$ represents the expected yield of background j , in fit sample k in the full fit range as shown in Fig. 4.30. The total background yields, $\mathcal{N}_j^{\text{tot}}$, are four fit parameters shared among all samples and Gaussian-constrained to their expectations, $\sum_{k'} \mathcal{N}_{j,\text{exp}}^{k'}$, shown in Table 4.18.

The mass shapes of the physics backgrounds are fixed from simulation as described in Sec. 4.7.4.

Finally, the fit model includes a combinatorial background component for each sample k with a free-floating yield $\mathcal{N}_{\text{comb}}^k$ and an exponential shape with free slope parameter τ^k .

4.7.2 Mass resolution calibration

The shape of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal mass distributions are obtained from fits to $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ simulation. However, data/simulation discrepancies require a correction of the mass line shape. Therefore, the widths of the invariant $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass distributions are multiplied

with correction factors C which are obtained using a data-driven approach. Since the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass distributions are unknown in the data, the width correction factors are derived from invariant J/ψ mass fits of $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ data and simulation samples, following the approach of the Run 1 analysis [42].

The correction factor, C , is defined as the ratio of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass resolution in data and simulation:

$$C = \frac{\sigma_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}^{\text{data}}}{\sigma_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp}^{\text{MC}}} . \quad (4.27)$$

Consider a decay with two oppositely-charged leptons in the final state, $X \rightarrow \ell^+ \ell^-$, where the energy of the leptons are E_{ℓ^+} and E_{ℓ^-} and the lepton momenta form an angle α . If the masses of the two leptons can be neglected with respect to the mass of the decaying particle, m_X , the following relation holds:

$$m_X = \sqrt{2E_{\ell^+} E_{\ell^-} (1 - \cos \alpha)} . \quad (4.28)$$

Considering now the limit case where the resolution on the angle α is perfect and the energy resolutions, σ_E , are independent, the mass resolution on m_X can be expressed as

$$\sigma_{X \rightarrow \ell^+ \ell^-} = \frac{m_X}{2} \sqrt{\left(\frac{\sigma_E}{E}\right)_{\ell^+}^2 + \left(\frac{\sigma_E}{E}\right)_{\ell^-}^2} . \quad (4.29)$$

Applying this to a $J/\psi \rightarrow \ell^+ \ell^-$ decay, the relative energy resolution, assumed to be the same for both leptons, can then be expressed as

$$\left(\frac{\sigma_E}{E}\right)_\ell = \sqrt{2} \frac{\sigma_{J/\psi \rightarrow \ell^+ \ell^-}}{m_{J/\psi}} . \quad (4.30)$$

If Eq. (4.30) is now inserted for $J/\psi \rightarrow \mu^+ \mu^-$ and $J/\psi \rightarrow e^+ e^-$ in Eq. (4.29) evaluated for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$, under the assumption that the final state leptons of $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ decays have the same relative energy resolutions as those from J/ψ decays, Eq. (4.29) becomes

$$\sigma_{B_{(s)}^0 \rightarrow e^\pm \mu^\mp} = \frac{m_{B_{(s)}^0}}{m_{J/\psi}} \sqrt{\frac{\sigma_{J/\psi \rightarrow e^+ e^-}^2 + \sigma_{J/\psi \rightarrow \mu^+ \mu^-}^2}{2}} . \quad (4.31)$$

The correction factor defined in Eq. (4.27) can therefore be expressed as

$$C = \sqrt{\frac{\left(\sigma_{J/\psi \rightarrow e^+ e^-}^{\text{data}}\right)^2 + \left(\sigma_{J/\psi \rightarrow \mu^+ \mu^-}^{\text{data}}\right)^2}{\left(\sigma_{J/\psi \rightarrow e^+ e^-}^{\text{MC}}\right)^2 + \left(\sigma_{J/\psi \rightarrow \mu^+ \mu^-}^{\text{MC}}\right)^2}} . \quad (4.32)$$

The core of the $J/\psi \rightarrow \mu^+ \mu^-$ mass distribution can be described with a symmetric function with a single width $\sigma_{J/\psi \rightarrow \mu\mu}$. However, the $J/\psi \rightarrow e^+ e^-$ and $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass distributions are very asymmetric due to bremsstrahlung. Therefore, it makes sense to model these distributions

with a bifurcated function, introducing different mass widths on the left and the right of the mass peak (σ_L and σ_R). In addition σ_L and σ_R may show dependencies on the bremsstrahlung categories 0γ and 1γ and are therefore considered separately for each case. Taking all these assumptions into account the final expression for the correction factor becomes:

$$C_{L/R,0\gamma/1\gamma} = \sqrt{\frac{(\sigma_{J/\psi \rightarrow e^+ e^-, L/R, 0\gamma/1\gamma}^{\text{data}})^2 + (\sigma_{J/\psi \rightarrow \mu^+ \mu^-}^{\text{data}})^2}{(\sigma_{J/\psi \rightarrow e^+ e^-, L/R, 0\gamma/1\gamma}^{\text{MC}})^2 + (\sigma_{J/\psi \rightarrow \mu^+ \mu^-}^{\text{MC}})^2}}. \quad (4.33)$$

The correction factors $C_{L/R,0\gamma/1\gamma}$ are determined by fitting the $m_{\text{DTF}}(J/\psi)$ distribution in $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ data and simulation, merged over the three data-taking years. The single component of the $m_{\text{DTF}}(J/\psi)$ distribution in $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data and simulation is modeled with the sum of two Crystal Ball Functions sharing their mean and width. In the fit to data, an additional exponential function is used to model the combinatorial background. The fit results are shown in Fig. 4.32. The single component of the $m_{\text{DTF}}(J/\psi)$ distribution in $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ data and simulation is modeled with a bifurcated double-sided Crystal Ball function. In the fit to data, an additional exponential function is used to model the combinatorial background. The fit results are shown in Fig. 4.33. The final correction factors of Eq. (4.33) are reported together with the obtained Gaussian mass resolutions in Table 4.24.

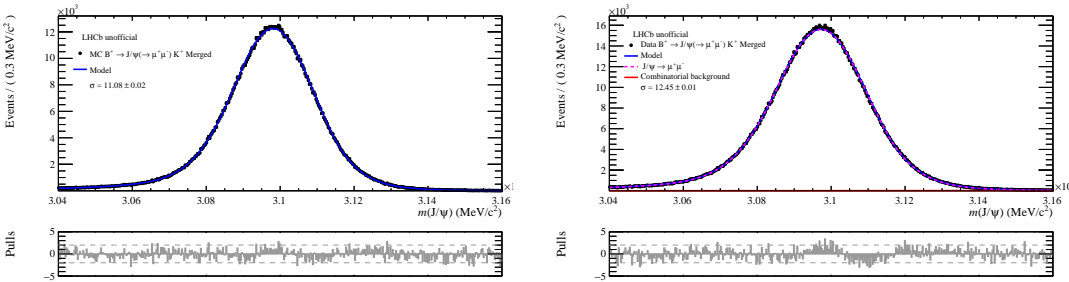


Figure 4.32 – Fit of the $m_{\text{DTF}}(J/\psi)$ distribution in $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ simulation (left) and data (right), merged over the data-taking years 2016, 2017, and 2018.

Table 4.24 – $J/\psi \rightarrow \ell^+ \ell^-$ ($\ell = \mu, e$) core mass resolutions in simulation and data (all years merged), in MeV/c^2 , and correction factors C to apply on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass resolutions from simulation.

| | | L, 0γ | R, 0γ | L, 1γ | R, 1γ |
|---|------------------|-----------------|-----------------|-----------------|-----------------|
| $\sigma_{J/\psi \rightarrow \mu^+ \mu^-}^{\text{MC}}$ | 11.08 ± 0.02 | | | | |
| $\sigma_{J/\psi \rightarrow \mu^+ \mu^-}^{\text{data}}$ | 12.45 ± 0.01 | | | | |
| $\sigma_{J/\psi \rightarrow e^+ e^-}^{\text{MC}}$ | | 71.2 ± 2.9 | 16.5 ± 0.3 | 52.0 ± 1.1 | 33.6 ± 0.6 |
| $\sigma_{J/\psi \rightarrow e^+ e^-}^{\text{data}}$ | | 81.3 ± 0.6 | 18.8 ± 0.3 | 60.3 ± 0.3 | 30.1 ± 0.2 |
| C | | 1.14 ± 0.05 | 1.14 ± 0.02 | 1.16 ± 0.02 | 0.92 ± 0.02 |

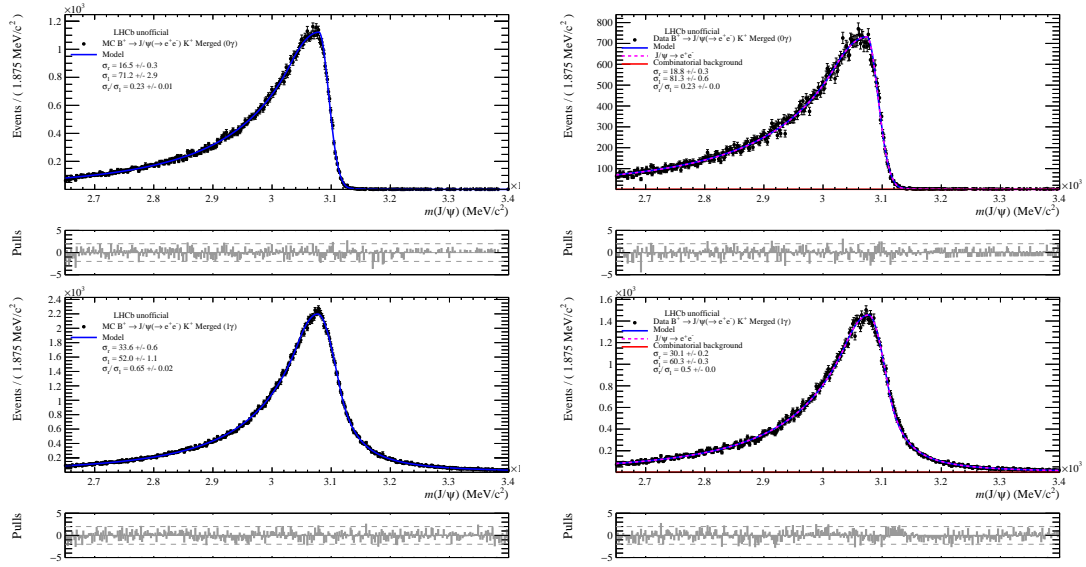


Figure 4.33 – Fit of the $m_{\text{DTF}}(J/\psi)$ distribution in $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ simulation (left) and data (right) for bremsstrahlung category 0 γ (top row) and 1 γ (bottom row).

4.7.3 Signal models

In the fit to the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass distributions, the B^0 and B_s^0 signals are each modeled with a bifurcated double-sided Crystal Ball function, with shape parameters fixed to values determined in fully corrected simulation, separately for each of the six samples. However, the left and right widths of the bifurcated function are multiplied by the correction factors determined in Sec. 4.7.2:

$$\sigma_{\text{L/R},0\gamma/1\gamma}^{\text{data}} = C_{\text{L/R},0\gamma/1\gamma} \cdot \sigma_{\text{L/R},0\gamma/1\gamma}^{\text{MC}}. \quad (4.34)$$

The fitted mass distributions in simulated $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ samples are shown in Figs. 4.34 and 4.35, respectively. The figures also show the model with the width correction applied.

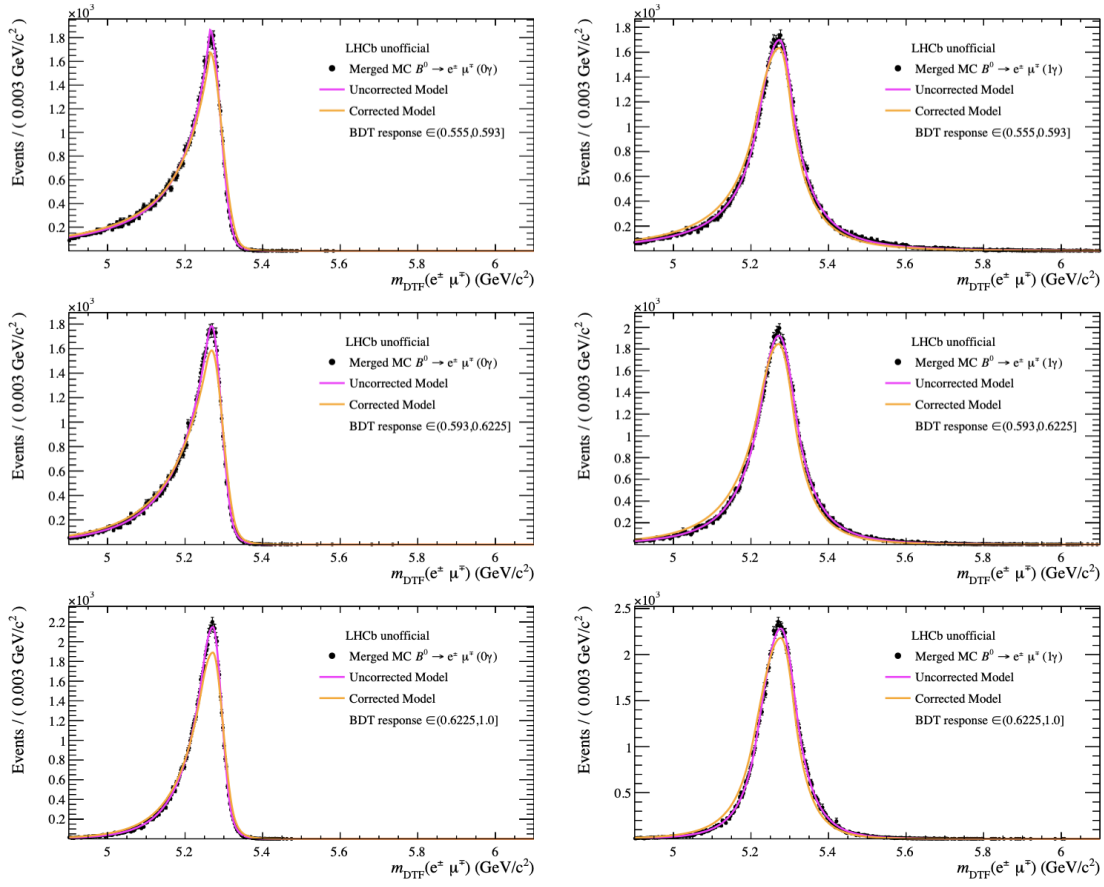


Figure 4.34 – Distribution of $m_{\text{DTF}}(e^\pm \mu^\mp)$ for $B^0 \rightarrow e^\pm \mu^\mp$ candidates in fully corrected simulation, passing the full reconstruction and selection chain and falling in the 0γ (left) or 1γ (right) bremsstrahlung category, and in BDT region 0 (top), 1 (middle) or 2 (bottom). The pink (orange) curves represent the fitted model without (with) the width correction.

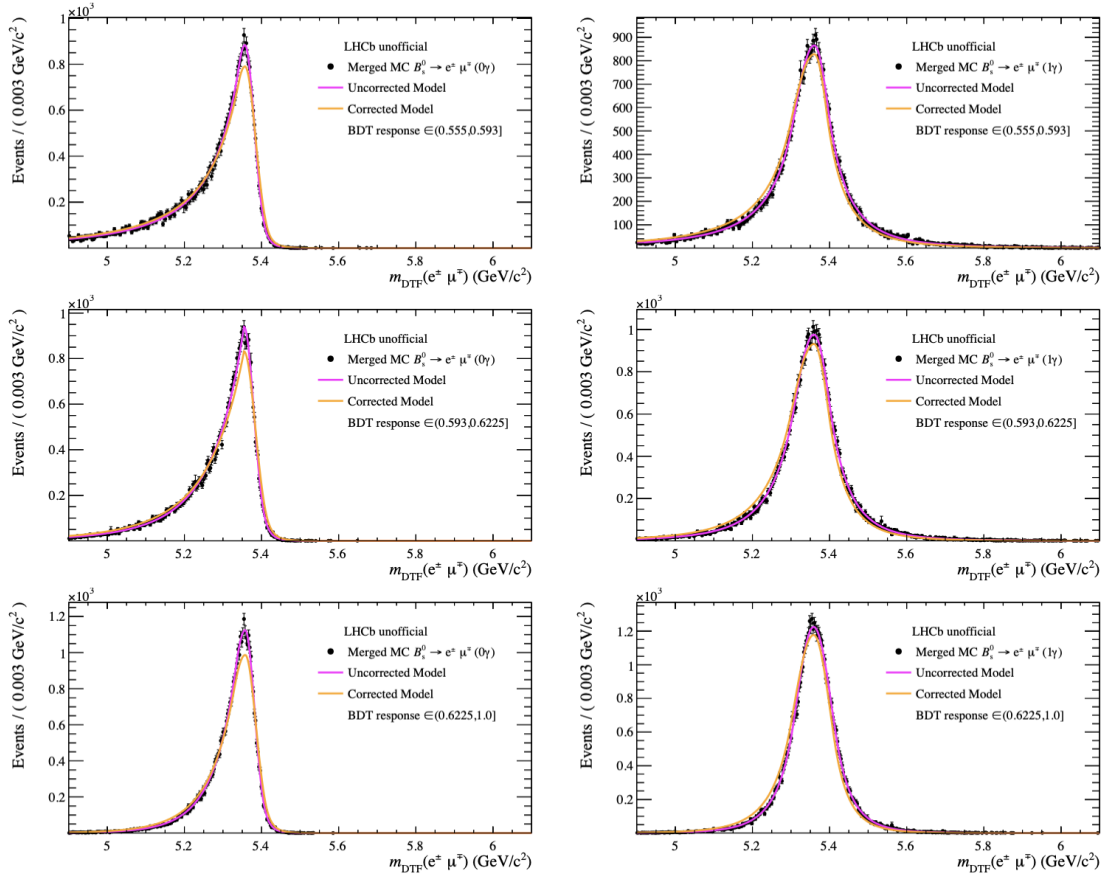


Figure 4.35 – Distribution of $m_{\text{DTF}}(e^\pm\mu^\mp)$ for $B_s^0 \rightarrow e^\pm\mu^\mp$ candidates in fully corrected simulation, passing the full reconstruction and selection chain and falling in the 0γ (left) or 1γ (right) bremsstrahlung category, and in BDT region 0 (top), 1 (middle) or 2 (bottom). The pink (orange) curves represent the fitted model without (with) the width correction.

4.7.4 Physics background models

Based on the study of Sec. 4.6.1, the following physics backgrounds are included in the $m_{\text{DTF}}(e^\pm\mu^\mp)$ fit model: $B^0 \rightarrow K^+\pi^-$, $B^0 \rightarrow \pi^+\pi^-$, $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ for the 0γ bremsstrahlung category, and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ only for the 1γ bremsstrahlung category.

Each of the peaking backgrounds, $B^0 \rightarrow K^+\pi^-$ and $B^0 \rightarrow \pi^+\pi^-$, is modeled with the sum of two Crystal Ball functions with common mean and width. Each of the two semi-leptonic backgrounds, $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ and $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$, is modeled with a superposition of Gaussian functions using a kernel-density estimator. The shape of each physics background component is determined, in each BDT region and each relevant bremsstrahlung category, using fully corrected simulation merged over all data-taking years. The fits to the $B^0 \rightarrow K^+\pi^-$ and $B^0 \rightarrow \pi^+\pi^-$ simulated samples are shown in Fig. 4.36, while those to the $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ and $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ are displayed in Fig. 4.37 and Fig. 4.38, respectively.

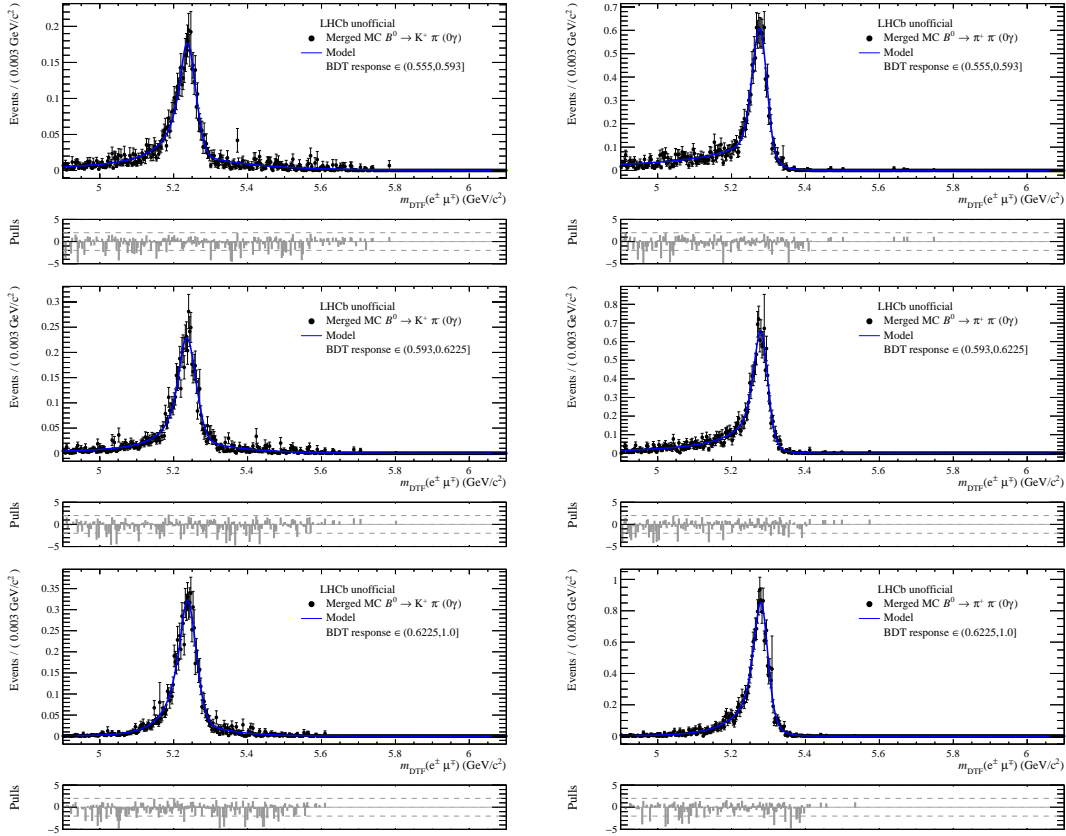


Figure 4.36 – Distribution of $m_{\text{DTF}}(e^\pm \mu^\mp)$ for $B^0 \rightarrow K^+ \pi^-$ (left) and $B^0 \rightarrow \pi^+ \pi^-$ (right) simulated events, falling in the 0γ bremsstrahlung category and in BDT region 0 (top row), 1 (middle row) or 2 (bottom row). The blue curves represent the fitted model.

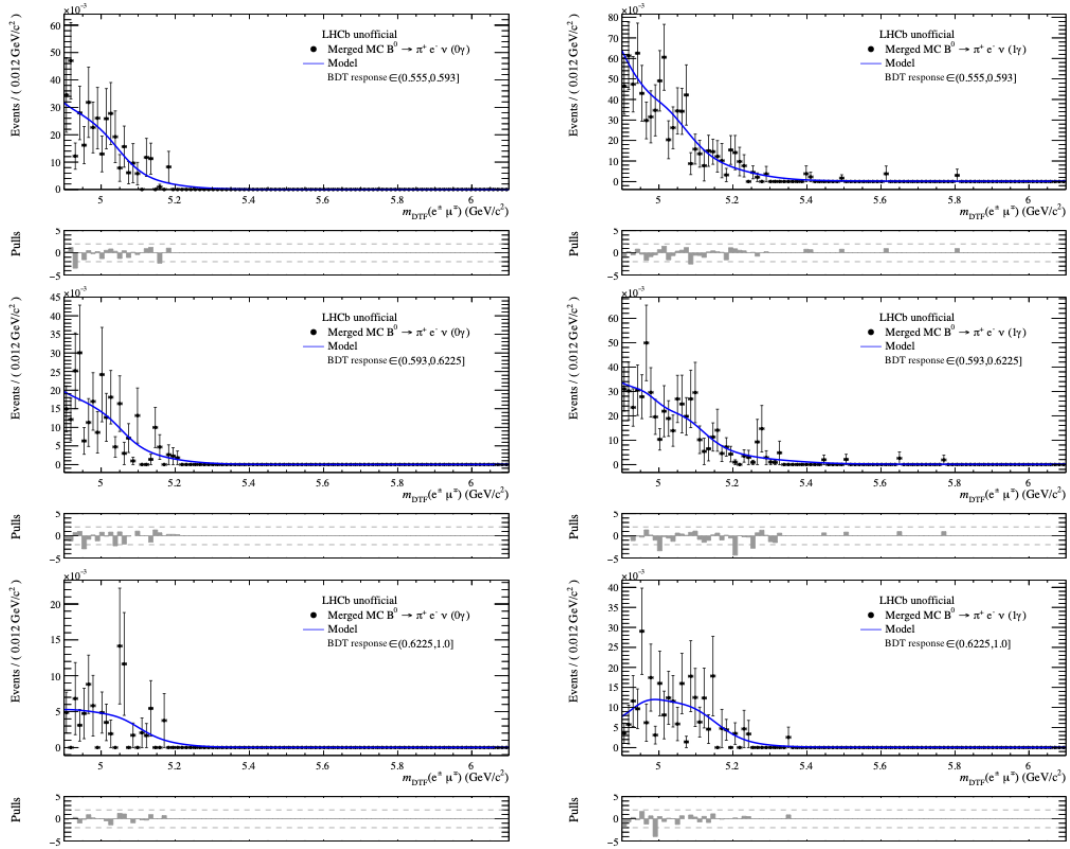


Figure 4.37 – Distribution of $m_{\text{DTF}}(e^\pm\mu^\mp)$ for $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ simulated events, falling in the 0γ (left) and 1γ (right) bremsstrahlung category and in BDT region 0 (top row), 1 (middle row) or 2 (bottom row). The blue curves represent the fitted model.

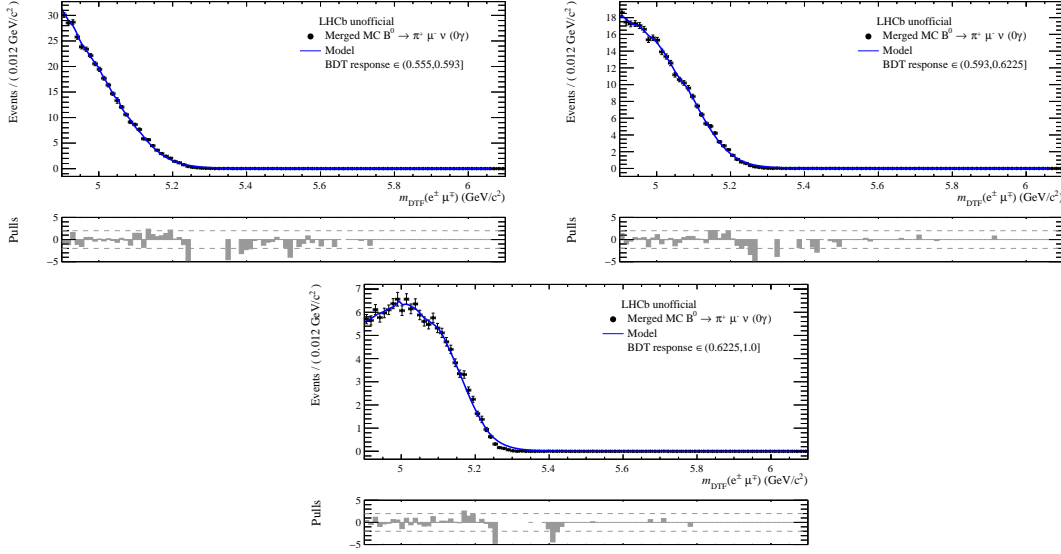


Figure 4.38 – Distribution of $m_{\text{DTF}}(e^\pm \mu^\mp)$ for $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ simulated events, falling in the 0γ bremsstrahlung category and in BDT region 0 (top right), 1 (top left) or 2 (bottom). The blue curves represent the fitted model.

4.7.5 Pseudo-experiments

Pseudo-experiments are generated to assess the stability and robustness of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass fit described in Sec. 4.7.1, as well as to study potential biases in the fitted parameters and their estimated uncertainties, and to obtain the expected limits on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal branching fractions.

A total of 500 pseudo-experiments with the same background level as observed in data are generated under the null hypothesis, *i.e.* under the hypothesis that no signal is present following several steps:

1. **Fit to data sidebands:** a simultaneous fit of the six samples, following the strategy presented in Sec. 4.7.1, is performed to the $m_{\text{DTF}}(e^\pm \mu^\mp)$ mass distribution of $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data, considering only the sideband regions, *i.e.* $m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.9.5.1] \cup [5.6, 6.1]$ GeV/c^2 . The $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal yields as well as the $B^0 \rightarrow K^+ \pi^-$ and $B^0 \rightarrow \pi^+ \pi^-$ yields are fixed to zero. The only remaining components in this so-called blind fit are the combinatorial background and the semi-leptonic backgrounds ($B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$), since these are the only relevant components in the sidebands. The $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ yields are Gauss-constrained, as explained in Sec. 4.7.1. The fit is performed applying the full selection chain and the corresponding result is displayed in Fig. 4.39. In each sample, $k = 1, \dots, 6$, the combinatorial background yields, $\mathcal{N}_{\text{comb}}^k$, and the semi-leptonic background yields, $\mathcal{N}_{\pi\mu\nu}^k$ and $\mathcal{N}_{\pi e\nu}^k$, are then estimated in the full mass range $m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.9.6.1]$ GeV/c^2 from the result of the sideband fit.
2. **Generation of combinatorial background:** for each pseudo-experiment, the number of

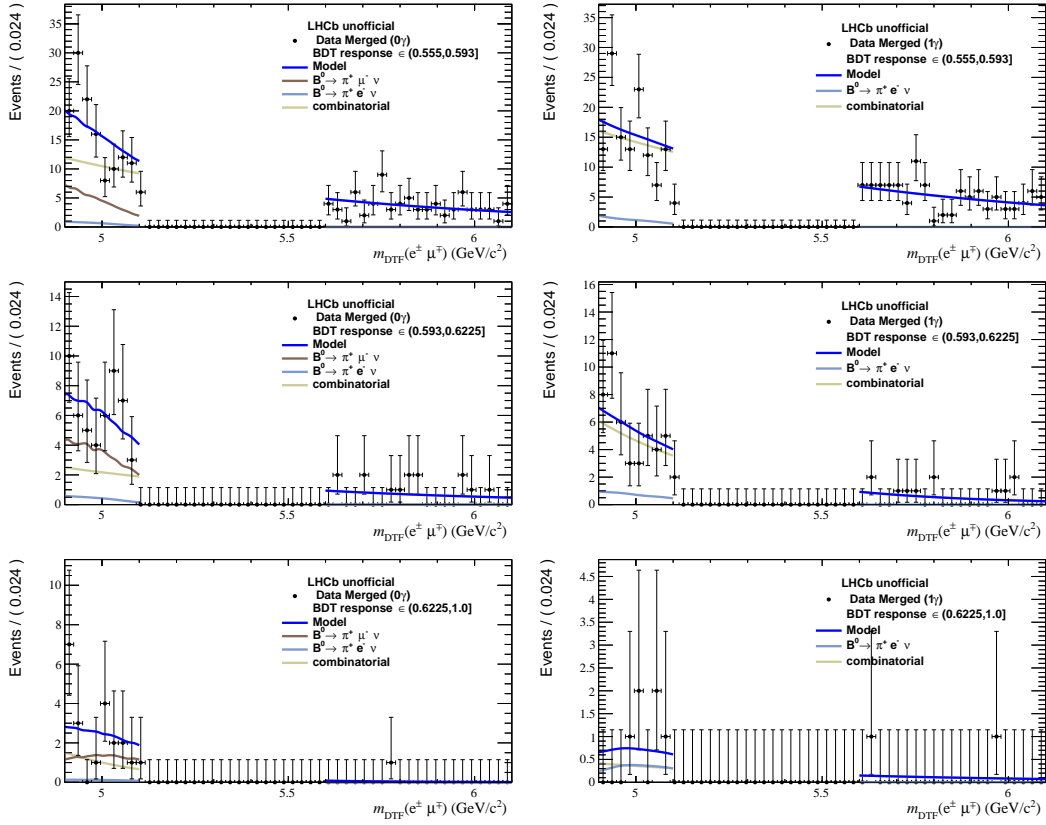


Figure 4.39 – Sidebands of the $m_{\text{DTF}}(e^\pm\mu^\mp)$ mass distributions in the blinded $B^0 \rightarrow e^\pm\mu^\mp$ data samples, with $m_{\text{DTF}}(e^\pm\mu^\mp) \in [4.9, 5.1] \cup [5.6, 6.1] \text{ GeV}/c^2$. The samples correspond to the 0γ (left) and 1γ (right) bremsstrahlung category in BDT region 0 (top), 1 (middle), or 2 (bottom). The fit is performed simultaneously in all six samples. The $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ yields are Gauss-constrained.

combinatorial background events in each sample k is obtained by drawing from a Poisson distribution of mean $\mathcal{N}_{\text{comb}}^k$. The combinatorial background events are generated, over the full mass range, with the exponential function fitted on the data sidebands.

3. **Generation of semi-leptonic backgrounds:** in analogy to the combinatorial background, the $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ yields are obtained by drawing from Poisson distribution of means $\mathcal{N}_{\pi\mu\nu}^k$ and $\mathcal{N}_{\pi e\nu}^k$, respectively. The $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ events are then generated over the full mass range using the fixed shapes from simulation, presented in Sec. 4.7.4.
4. **Generation of peaking backgrounds:** the $B^0 \rightarrow K^+\pi^-$ and $B^0 \rightarrow \pi^+\pi^-$ backgrounds are generated for each pseudo-experiment with the functions determined in Sec. 4.7.4. The number of generated events in each sample is drawn from a Poisson distribution of mean equal to the corresponding number of expected events obtained in Sec. 4.6.1.

The generated pseudo-experiments are fitted over the full mass range as described in Sec. 4.7.1, with the exception that the total $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ yields, defined in Eq. (4.26),

Chapter 4. Search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$

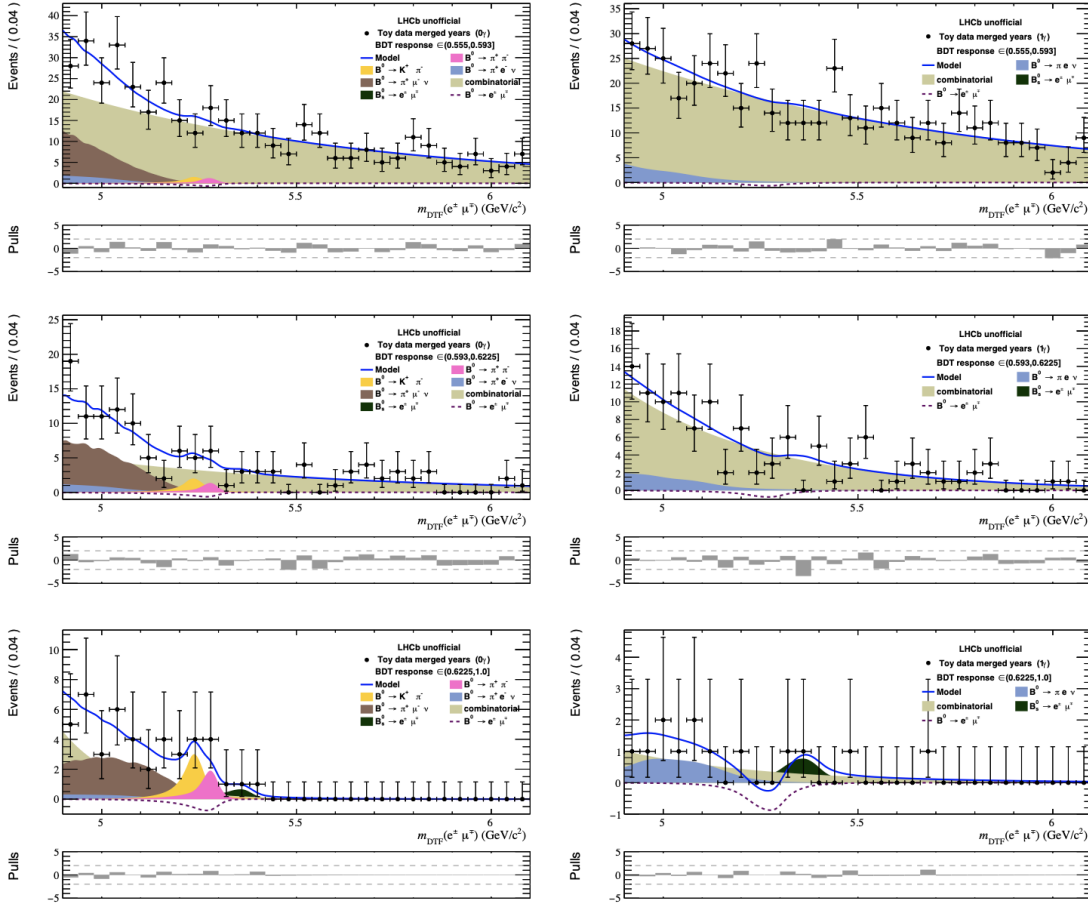


Figure 4.40 – $m_{\text{DTF}}(e^\pm \mu^\mp)$ distributions for one pseudo-experiment with six samples (left = 0γ , right = 1γ , top/middle/bottom = BDT region 0/1/2) generated without $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signals. The blue curves represent the result of a simultaneous fit of the six samples, with a model accounting for exponential combinatorial backgrounds (brown), as well as $B^0 \rightarrow K^+ \pi^-$ (orange), $B^0 \rightarrow \pi^+ \pi^-$ (pink), $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ (brown) and $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ (light blue) backgrounds with total yields Gaussian-constrained to expectations. The yields of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal (green) and (violet) are free in the fit.

are Gaussian-constrained to the total yields determined in the blind fit to data, $\sum_{k=1}^6 \mathcal{N}_{\pi\mu\nu}^k$ and $\sum_{k=1}^6 \mathcal{N}_{\pi e\nu}^k$, respectively. The number of successful (*i.e.* converged) fits among the 500 pseudo-experiments is 470, showing a stability above 95%. An example for one pseudo-experiment is shown in Fig. 4.40.

To assess potential biases in the 18 fitted parameters, $\hat{\Theta}^i \pm \Delta\Theta^i$ ($i = 1, \dots, 18$), the pull values are computed as

$$\text{Pull}^i = \frac{\hat{\Theta}^i - \Theta_0^i}{\Delta\Theta^i}, \quad (4.35)$$

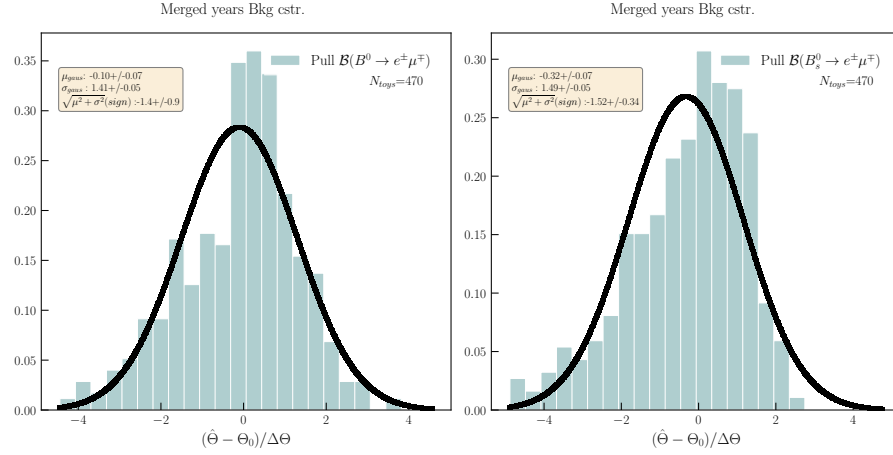


Figure 4.41 – Pull distributions of the $B^0 \rightarrow e^\pm \mu^\mp$ (left) and $B_s^0 \rightarrow e^\pm \mu^\mp$ (right) branching fractions fitted in 470 successful pseudo-experiments generated without $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal components. The curves represent the results of Gaussian fits.

where Θ_0^i is the true value of parameter i , used for the generation of the pseudo-experiments. For the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal branching fractions, Θ_0 is zero, while for the combinatorial background yield in sample k $\Theta_0^i = \mathcal{N}_{\text{comb}}^k$ determined from the data sideband fit. For the total semi-leptonic background yields, Θ_0^i is set to the values determined in the blind fit $\sum_{k=1}^6 \mathcal{N}_{\pi\mu\nu}^k$ and $\sum_{k=1}^6 \mathcal{N}_{\pi e\nu}^k$, while for the total peaking background yields Θ_0^i is set to the expected values determined in Sec. 4.6.1.

The pulls are calculated for each successful toy, and their distributions are fitted with a Gaussian function. The pull distributions for the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal branching fractions are displayed in Fig. 4.41, and a summary of the fitted mean and sigma, determined from the Gaussian fits, is shown in Fig. 4.42 for all fit parameters.

The obtained sigma for both signal branching fractions is significantly larger than one, indicating that the fit is underestimating the uncertainties. Furthermore, the obtained mean for the $B_s^0 \rightarrow e^\pm \mu^\mp$ branching fraction shows the largest deviation from zero among the obtained means. This is a result of the asymmetric tails of the distributions, which is suspected to be related to the low statistics. The likelihood method is unbiased in the limit of an infinite statistics, which is clearly not the case here. However, despite these issues, the results presented in Fig. 4.42 do not show any further large biases and the observed effects are planned to be studied in detail as described in Sec. 4.9.

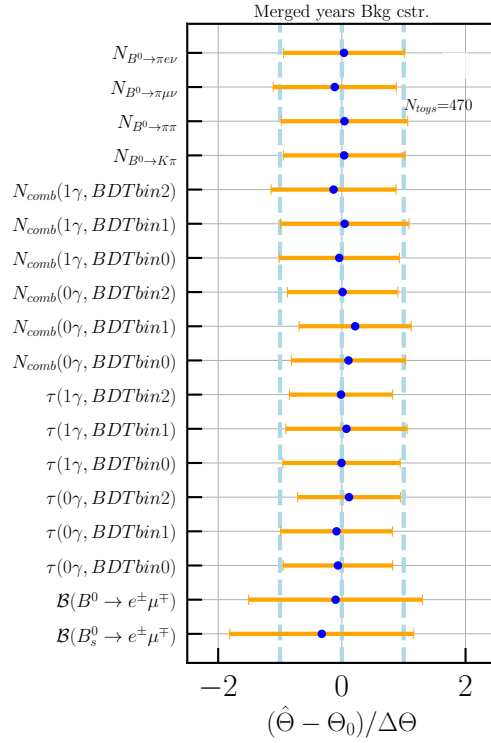


Figure 4.42 – Mean (blue markers) and sigma (orange bars) of the Gaussians fits to the pull distributions of the 18 fitted parameters in 470 successful pseudo-experiments, generated without $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal components.

4.8 Branching fraction limits

If the fitted $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions are compatible with zero, upper limits will be determined with the CL_s method [76].

The upper limit is evaluated by testing the signal plus background hypothesis (H_{s+b}) against the background-only (H_b) hypothesis, where the corresponding models are defined as the product of the signal (background) yield and the probability density functions of the signal (background) components and used to build the ratio of likelihoods L , defined as [91]

$$\lambda(\mu) = \frac{L(\mu|\hat{\hat{\theta}})}{L(\hat{\mu}|\hat{\hat{\theta}})} \quad (4.36)$$

where μ represents a hypothesised value for the parameter of interest and $\hat{\hat{\theta}}$ the value of the nuisance parameters θ which maximise the likelihood for that specific hypothesis. $\hat{\mu}$ and $\hat{\hat{\theta}}$ are the maximum likelihood estimators, obtained by fitting the data sample with the model of H_{s+b} or H_b , *i.e.* the values belong to the global maximum. As shown in Ref. [91] a test statistic q_μ can be defined, taking the likelihood ratio into account, and the p -value for a hypothesised is computed, integrating the probability density function (PDF) of test statistic with respect to

the observed value in the dataset as

$$p_\mu = \int_{q_{\text{obs}}}^{\infty} f(q_\mu|\mu) dq_\mu. \quad (4.37)$$

The PDF of the test statistic, $f(q_\mu|\mu)$ is unknown and can be determined with MC simulation, however, since this approach is very time and computing intensive, this analysis uses the asymptotic approach for likelihood-based tests of new physics, presented in Ref. [91], where it is shown that the distribution of the likelihood ratio can be approximated by a χ^2 distribution in the limit of a large sample size.

From Eq. (4.37) an upper limit at 95% (90%) confidence level, for a single hypothesis, can be simply determined by finding a value of μ , scanning over a pre-defined range, for which $p_\mu = 0.05$ (0.10). However, if the PDF of the test static for H_{s+b} and H_b are very similar the test becomes insensitive for testing μ values in the high sidebands of the PDFs. Therefore, the upper limit with the CL_s is determined with the ratio:

$$CL_s = \frac{CL_{s+b}}{1 - CL_b} = \frac{p_{s+b}}{1 - p_b}, \quad (4.38)$$

where p_{s+b} is the p -value determined for μ under H_{s+b} and p_b under H_b .

The sensitivity of measurement is evaluated by determining the expected upper limit under the H_b assumption and computing the median of q_μ with the Asimov data, which replaces the ensemble of distributions generated for each hypothesised μ with a single representative one [91].

To determine the upper limit for one of the two signal branching fractions, a model is constructed under the signal plus background hypothesis and the background only hypothesis, based on the result of the simultaneous $m_{\text{DTF}}(e^\pm\mu^\mp)$ fit, presented in Sec. 4.7.1. The expected upper limits are determined for each pseudo-experiment introduced in Sec. 4.7.5. The observed upper limits will be determined in a similar way from the results of the fit to the unblinded data.

The limit for each signal branching fraction is determined while fixing the other one to zero. The remaining fit parameters are considered as nuisance parameters. The one-sided single branching fraction limits are determined at 90% and 95% confidence level (CL).

4.8.1 Expected limits from pseudo-experiments

The expected upper limits on the $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$ branching fractions are calculated as described above, using the fit results of the pseudo experiments from Sec. 4.7.5. Furthermore, the expected upper limit on the $B_s^0 \rightarrow e^\pm\mu^\mp$ branching fraction is also determined in case one of the two mass eigenstates is completely dominating the decay amplitude. These expected upper

Table 4.25 – Mean of the distributions of the expected upper limits on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions evaluated for a single branching fraction. For $B_s^0 \rightarrow e^\pm \mu^\mp$ the expected limits are also determined assuming the lifetimes of the light and heavy mass eigenstates. Systematic uncertainties are not included. The results are compared to those of the Run 1 analysis [42].

| Decay | Assumed lifetime [ps] | Mean exp. limit 95% (90%) CL | Exp. limit Run 1 [42] 95% (90%) CL | Obs. limit Run 1 [42] 95% (90%) CL |
|-----------------------------------|--------------------------|---------------------------------|---------------------------------------|---------------------------------------|
| $B^0 \rightarrow e^\pm \mu^\mp$ | | $7.3 (5.9) \cdot 10^{-10}$ | $1.2 (0.9) \cdot 10^{-9}$ | $1.3 (1.0) \cdot 10^{-9}$ |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.512 ± 0.007 | $2.2 (1.8) \cdot 10^{-9}$ | – | – |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.429 ± 0.007 | $2.3 (1.9) \cdot 10^{-9}$ | – | $7.2 (6.0) \cdot 10^{-9}$ |
| $B_s^0 \rightarrow e^\pm \mu^\mp$ | 1.624 ± 0.009 | $1.8 (1.5) \cdot 10^{-9}$ | $5.0 (3.9) \cdot 10^{-9}$ | $6.3 (5.4) \cdot 10^{-9}$ |

limits are obtained for each successful pseudo-experiment. As an example, the expected upper limits for the pseudo-experiment presented in Fig. 4.40 is shown in Fig. 4.43. The distributions of the expected upper limits are presented in Fig. 4.44. The mean values of these distributions are reported in Table 4.25, where also the observed and expected limits obtained in the Run 1 analysis [42] are presented.

When comparing the mean values expected upper limit distributions of this study with the expected upper limits of the Run 1 analysis at 95% (90%) CL, the improvement factors are found to be 1.6 (1.5) for $B^0 \rightarrow e^\pm \mu^\mp$ and 2.8 (2.9) for $B_s^0 \rightarrow e^\pm \mu^\mp$ assuming the heavy eigenstate dominates the decay amplitude. The improvement factors are roughly compatible with the very naive expectation of two, given in Sec. 1.4. Of course, one has to consider that the analysis presented here, deviates significantly from the Run 1 analysis, *e.g.* the estimated peaking backgrounds in the Run 1 analysis are close to zero and are therefore not taken into account.

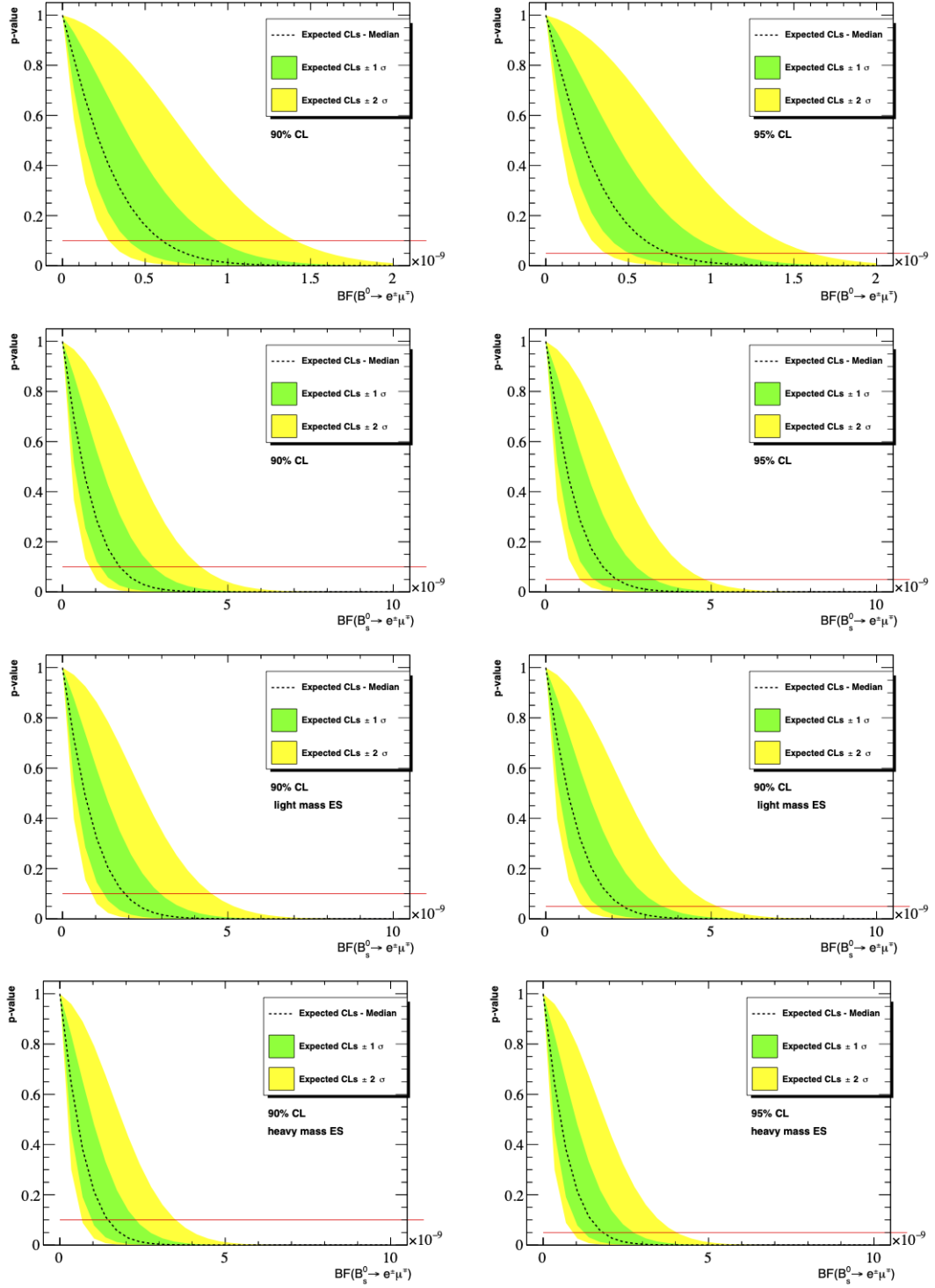


Figure 4.43 – CL_s scan of a single pseudo-experiment to determine the expected upper limits at 90% (left) and 95% (right) CL on the $B^0 \rightarrow e^\pm \mu^\mp$ (first row), $B_s^0 \rightarrow e^\pm \mu^\mp$ (second row), and $B_s^0 \rightarrow e^\pm \mu^\mp$ assuming the decay amplitude is dominated by the light or heavy mass eigenstate (third or fourth row, respectively).

Chapter 4. Search for $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$

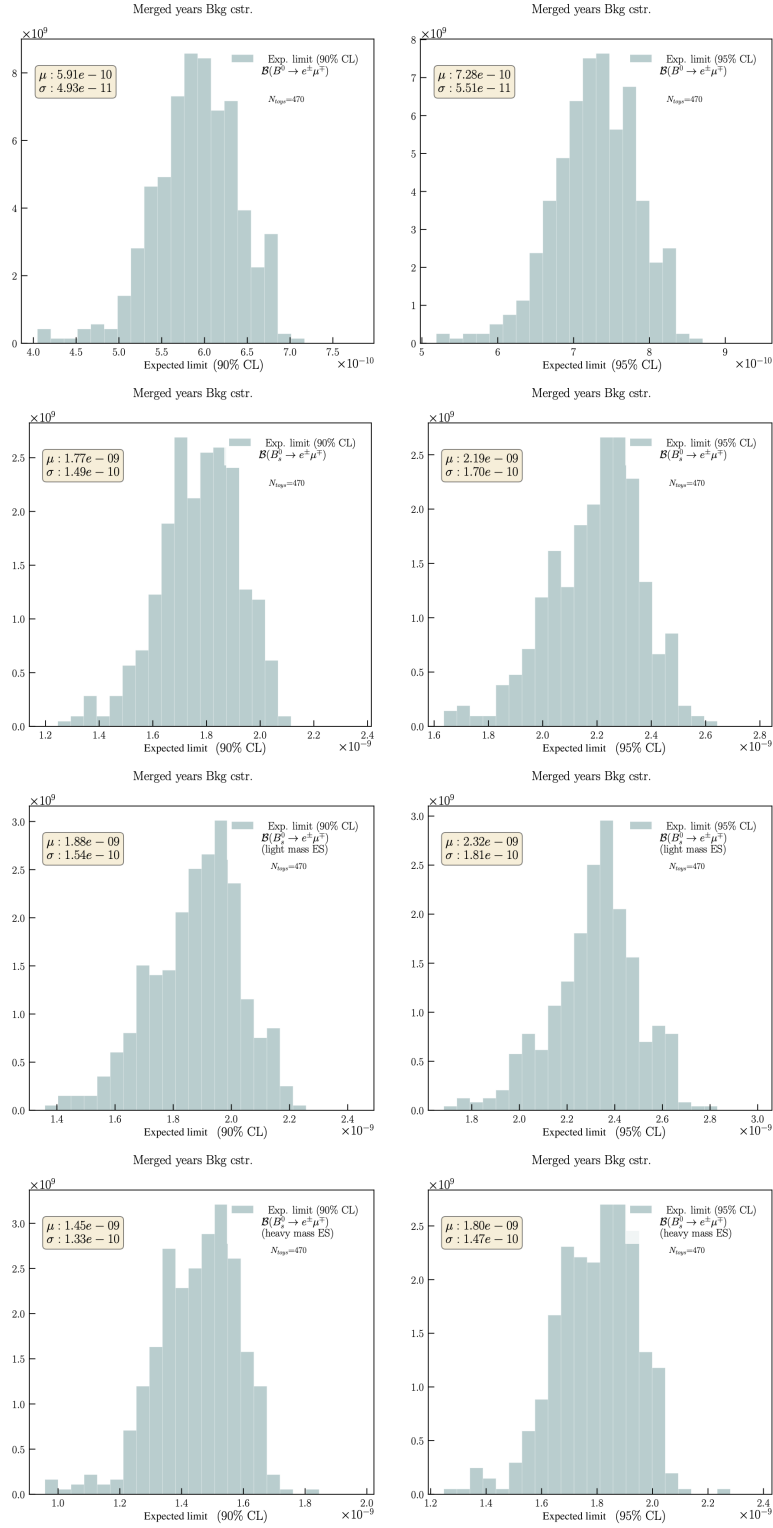


Figure 4.44 – Distributions of the expected upper limits at 90% (left) and 95% (right) CL on the $B^0 \rightarrow e^\pm \mu^\mp$ (first row), $B_s^0 \rightarrow e^\pm \mu^\mp$ (second row), and $B_s^0 \rightarrow e^\pm \mu^\mp$ assuming the decay amplitude is dominated by the light or heavy mass eigenstate (third or fourth row, respectively).

4.9 Systematic uncertainties

Relative systematic uncertainties in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction analysis are suspected to be mainly driven by the efficiencies and are taken into account, under the assumption that they depend only on the bremsstrahlung category, by introducing two new fit parameters $c_{\text{sys}}^{\text{brem}}$ (brem=0 γ , 1 γ) multiplying the signal yields as

$$\mathcal{N}_{\text{sys}}^k(B_{(s)}^0 \rightarrow e^\pm \mu^\mp) = c_{\text{sys}}^{\text{brem}} \cdot \mathcal{N}^k(B_{(s)}^0 \rightarrow e^\pm \mu^\mp), \quad (4.39)$$

where $\mathcal{N}^k(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)$ represents the signal yields for each of the six samples k defined in Eq.(4.25) and $\mathcal{N}_{\text{sys}}^k(B_{(s)}^0 \rightarrow e^\pm \mu^\mp)$ the yields including systematic uncertainties. The multiplicative factors $c_{\text{sys}}^{\text{brem}}$ are shared between the $B_s^0 \rightarrow e^\pm \mu^\mp$ and $B^0 \rightarrow e^\pm \mu^\mp$ signal yields, in all BDT regions. Furthermore, they are constrained with a Gaussian of mean one and of width equal to the sum in quadrature of all the evaluated relative systematic uncertainties.

Two classes of systematics effects are considered: those related to the evaluation of the selection efficiencies and all other additional systematic effects, for example related to the evaluation of the signal yields.

The efficiency-related systematic sources are:

- effects arising from the electron miscalibration in simulation;
- uncertainties arising from the truth-matching criteria applied to simulation;
- effects arising from the statistics of the simulation samples used to determine the selection efficiencies;
- the coverage PID, L0 trigger and tracking corrections, applied to simulation;
- the impact of the track multiplicity, nTracks, in the kinematic reweighting;
- the portability of the L0 trigger corrections to $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ simulation;
- BDT response mismodeling.

The considered non-efficiency-related systematic effects are:

- uncertainties on the fixed parameters in the mass fit;
- determination of the normalisation yields;
- $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal model;
- impact from physics background constrains;
- fit bias and uncertainties underestimation, discovered in the pseudo-experiment studies presented in Sec. 4.7.5;
- additional background sources not included in the model.

It has not been feasible to finalise the evaluation of the systemic uncertainties within the timescale of this thesis. Nevertheless, the current results are discussed, as well as the methods used for the determination of all listed systematic uncertainties. Furthermore, studies with pseudo-experiments are performed to investigate the impact of the systematic uncertainties on the expected limits of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions.

4.9.1 Efficiency-related systematics

Electron miscalibration in simulation

The method presented here has been developed as part of a Master thesis [66], which was dedicated to the determination of the systematic effects due to the electron calibration for this analysis. Therefore, only a brief overview of the applied method is presented in the following.

The $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiencies entering the branching fraction measurements are determined from simulation. However, the cross-check in Sec. 4.5.7 shows that $r_{J/\psi}$ deviates from unity by approximately 10%, which is assumed to be due to residual mismodeling effects of the electron reconstruction and selection, since those are tested in Ref. [66] to be well modeled for the muons by measuring the $B^+ \rightarrow J/\psi K^+$ branching fraction with $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ corrected simulation and background-subtracted data, which is in agreement with the PDG value of $(1.020 \pm 0.019) \cdot 10^{-3}$ [13]. Therefore, weights are computed to correct for the electron miscalibration in simulation, which is assumed to be responsible for the $r_{J/\psi}$ deviations. Efficiency correction weights are obtained for each electron track of the $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ final state and used to derive a systematic uncertainty for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ efficiencies.

The efficiency correction weights are determined to bring $r_{J/\psi}$ to unity, for each bremsstrahlung category, in two-dimensional bins of a kinematic observable for the two leptons of the $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ decays. The exercise is repeated for three different kinematic observables $\log_{10}(p_T)$, $\log_{10}(p)$ and η of the two lepton tracks. The electron correction weights, ω_{calib} , are determined under the following constraints:

$$r_{J/\psi \text{ brem}}^{i,j} \cdot \omega_{\text{calib}}(e^+)_\text{brem}^i \cdot \omega_{\text{calib}}(e^-)_\text{brem}^j = 1, \quad (4.40)$$

where i and j are the bin indices of the two electrons. Then a relative variation of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiency is derived as

$$\frac{\Delta \epsilon_{\text{sel}}^{\text{tot}}}{\epsilon_{\text{sel}}^{\text{tot}}} = \frac{\sum_{i=1}^{N_{\text{sel, trig}}} \omega_{\text{TRK}}^i \cdot \epsilon_{\text{PID}}^i \cdot \omega_{\text{L0}}^i \cdot \omega_{\text{BKIN}}^i \cdot \omega_{\text{calib}}^i(e) - \sum_{i=1}^{N_{\text{sel, trig}}} \omega_{\text{TRK}}^i \cdot \epsilon_{\text{PID}}^i \cdot \omega_{\text{L0}}^i \cdot \omega_{\text{BKIN}}^i}{\sum_{i=1}^{N_{\text{sel, trig}}} \omega_{\text{TRK}}^i \cdot \epsilon_{\text{PID}}^i \cdot \omega_{\text{L0}}^i \cdot \omega_{\text{BKIN}}^i}, \quad (4.41)$$

following from Eqs. (4.17) and (4.20). This variation is computed for each data-taking year, as well as for the luminosity average over all three data-taking years, in all three BDT regions for the bremsstrahlung categories 0γ and 1γ .

A global relative systematic uncertainty on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ efficiency is then defined as the largest observed value of the relative variations determined for the average of the three data-taking years. This maximum is found to happen for the $\log_{10}(p)$ observable in BDT region 2 and bremsstrahlung category 1γ and is equal to 8.49% for $B^0 \rightarrow e^\pm \mu^\mp$ and 8.52% for $B_s^0 \rightarrow e^\pm \mu^\mp$ [66].

Truth-matching criteria

The truth-matching (see Sec. 4.3.3) applied in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ simulation is not fully efficient, and may therefore bias the efficiency estimation. The efficiency of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ truth-matching requirement is estimated from simulation by applying the requirement after the stripping selection without any further offline selection as

$$\varepsilon_{\text{truth}} = \frac{N_{\text{strip,truth}}}{N_{\text{strip}}}, \quad (4.42)$$

where N_{strip} is the number of events passing the stripping selection and $N_{\text{strip,truth}}$ the correctly truth matched events. The efficiency is found to be approximately 97%.

An alternative way to determine the efficiency from simulation is to relax the truth-matching criteria and allow for tracks to be a ghost, *i.e.* a track not matched to any true particle. A systematic uncertainty can then be assigned as the difference between the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiency determined while including the ghost candidates and the nominal efficiency, calculated without the ghost candidates.

Simulation statistics

Systematic effects arising from the limited size of the simulation samples used to determine the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiencies are evaluated by making use of the bootstrapping method [92], with the Poisson-bootstrap technique [93]. Each simulated signal event is weighted with a weight drawn from a Poisson distribution with an expectation value of one. The weights are used to recompute the selection efficiency, taking all corrections into account. The procedure is performed 100 times, resulting in an efficiency distribution. The systematic uncertainty is determined as the spread of this distribution.

Coverage of the efficiency corrections

A systematic uncertainty is evaluated by taking the statistics and the uncertainties from the calibration samples, used for the determination of the corrected PID efficiencies, the L0 trigger corrections, and the tracking corrections, into account.

The systematic effect is determined with the bootstrapping method [92], evaluating the selection efficiency, 100 times, each time using a variation of the corrections. For each variation, the selection efficiencies are determined by smearing the tracking and L0 trigger corrections, as well as the data-driven PID efficiencies, in each bin. The corresponding bin content is randomly varied within its statistical uncertainty using a Gaussian function. Wherever this approach is performed to corrections that represent efficiencies, the gaussian smearing is forced to be between zero and one. A set of varied histograms is generated by repeating this procedure multiple times and used to determine new efficiencies for each variation. The new efficiency models are then used to compute the ratio with respect to the nominal value and

build a distribution. An absolute systematic uncertainty is determined as the shift of the mean value of the distribution.

Kinematic reweighting

To evaluate systematic effects arising from kinematic reweighting presented in Sec. 4.5.5 and using the track multiplicity, nTracks, are evaluated with a second set of kinematic corrections is obtained without making use of the nTracks variable as reweighting proxy. The systematic uncertainty is computed as the difference between the selection efficiencies with the new set of kinematic weights and the result with the nominal one.

Portability of the L0 trigger corrections

The L0 trigger corrections are determined in Sec. 4.5.4 with the TISTOS method from $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ data and simulation. The portability of these corrections can be verified by applying the TISTOS method in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ simulation and comparing the results per electron/muon track with the ones obtained from $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ simulation. In case of a disagreement between the nominal corrections and the corrections obtained from $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ simulation, the observed deviation is considered as a systematic uncertainty.

Mismodeling of the BDT response

To assess the systematic effects from the BDT requirements applied to the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ samples, the $J/\psi \rightarrow \mu^+ \mu^-$ decays are used, applying the BDT to $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data and corrected simulation. The J/ψ is treated as a B meson and observables are studied, comparing background-subtracted data and corrected simulation. The observed discrepancies are taken into account to derive correction weights as a function of the highest ranked discriminating variables of the BDT. The weights are used to compute the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiency from fully corrected simulation. The difference of the thus obtained efficiency and the nominal $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ selection efficiency can then, if need, be further propagated as a systematic uncertainty.

4.9.2 Non-efficiency related systematics

Fixed parameters in the mass fit

The statistical uncertainty of the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ normalisation yield entering Eq.(4.25), as well as the known uncertainties of the external branching ratios, $\mathcal{B}(B^+ \rightarrow J/\psi K^+)$ and $\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$, and the relative uncertainty of the ratio f_s/f_u can be introduced as a single systematic uncertainty on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal yields. The relative uncertainties are presented in Table 4.26 for each parameter. Alternatively, it is considered, to constrain each fixed

Table 4.26 – Relative uncertainties of the normalisation yield, normalisation branching fractions, and fragmentation function ratios at 13 TeV entering the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ fit.

| Parameter | Relative uncertainty |
|---|-----------------------------------|
| $\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)$ | $7.78 \cdot 10^{-4}$ (Table 4.14) |
| $\mathcal{B}(B^+ \rightarrow J/\psi K^+)$ | 1.9% [13] |
| $\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$ | 0.55% [13] |
| f_s/f_u (at 13 TeV) | 3.1% [75] |

component in the fit model with a Gaussian distribution with a mean value of one and a width of the known uncertainties.

Evaluation of the normalisation yields

The $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ normalisation yield is a fixed parameter in the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ mass fit. While systematic effects related to the corresponding statistical uncertainty are described above, further effects arising from the method used to extract the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ normalisation yields, as described in Sec. 4.4, need to be taken into account by considering different variations of the method:

- $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ mass fit with an alternative model, *e.g.* using an Hypatia function for the signal model;
- $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ mass fit, using the $m(K + \mu^+ \mu^-)$ invariant mass (without the J/ψ mass constraint) instead of $m_{\text{DTF}}(J/\psi K^+)$;
- excluding the kaon PID requirement in the $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ selection, which affects the invariant mass fit and the normalisation efficiency $\varepsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)$.

The maximum variation observed among the three approaches can be taken as systematic uncertainty.

Signal model

In order to take systematic effects related to the line shape of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ models into account, the width scaling factor C , introduced in Sec. 4.7.2, can be determined with an alternative modelling of the $m_{\text{DTF}}(J/\psi)$ distributions. For the $J/\psi \rightarrow \mu^+ \mu^-$ model the sum of two Crystal Ball functions can be replaced with a double-sided Crystal Ball function, while for the $J/\psi \rightarrow e^+ e^-$ model a sum of two Crystal Ball functions, sharing their mean and width, can be used instead of the bifurcated Crystal Ball function.

Instead of determining a left and right scaling factor, this choice results in a single scaling factor. To propagate this scaling factor into the final fit model of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions, the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal shapes need to be adapted as well, by replacing the bifurcated

functions with the sum of two Crystal Ball functions, sharing the mean and the width. Pseudo-experiments with the alternative model and scaling factor can then be performed and the variation of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions can then be considered as a systematic uncertainty.

Physics background yields constraints

Systematic effects related to the Gaussian constraints applied to the physics background ($B^0 \rightarrow K^+ \pi^-$, $B^0 \rightarrow \pi^+ \pi^-$, $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ and $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$) yields, are evaluated by changing central values and the widths of the constraints.

Therefore, an alternative size of the corresponding uncertainties can be taken into account by considering the results of the $B_{(s)}^0 \rightarrow hh'$ cross-check, presented in Sec. 4.6.2, taking the maximum deviations between estimated and observed events into account, which is of the size of approximately 20%. Pseudo-experiments can then be performed with the alternative constraints and the impact of related systematic effects can be studied and further ported to the final results, if needed.

Fit bias and uncertainties underestimation

The $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction pull distributions, determined in Sec. 4.7.5, show that the fit is underestimating the uncertainties and that the $B_s^0 \rightarrow e^\pm \mu^\mp$ branching fraction pull is biased.

To investigate if the origin of these issues is related to the low statistics, a study is performed with 500 pseudo-datasets, generated with five times the statistics of the nominal pseudo-datasets presented in Sec. 4.7.5. Each of these pseudo-datasets is produced under the background only hypothesis following the procedure described in Sec. 4.7.5, while taking into account five times the amount of combinatorial and physics backgrounds in each fit sample k , by multiplying the nominal background yields ($\mathcal{N}_{\text{comb}}^k$, $\mathcal{N}_{\pi\pi}^k$, $\mathcal{N}_{K\pi}^k$, $\mathcal{N}_{\pi\mu\nu}^k$ and $\mathcal{N}_{\pi e\nu}^k$) by five.

Pseudo-experiments are performed with the pseudo-datasets with five times more statistics, to study the fit stability and compute the pull distributions. Each pseudo-dataset is fitted with the nominal fit model presented in Sec. 4.7.1. While the efficiency ratios in Eq. (4.25) are kept the same, the normalisation yield $\mathcal{N}(B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+)$ is multiplied by five. Furthermore, the constraints on the background yields (see Eq. (4.26)) are adjusted to take the increase in statistics into account.

The increase in statistics leads to a success rate of 100%, *i.e.* the fit for all 500 pseudo-experiments converges. The pull values of all fit parameters are computed with Eq. (4.35). The pull distributions of the $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ branching fractions, fitted with a Gaussian function, are presented in Fig. 4.45. The obtained means of the Gaussian fits are in good agreement with zero, indicating that there are no biases present anymore. Furthermore,

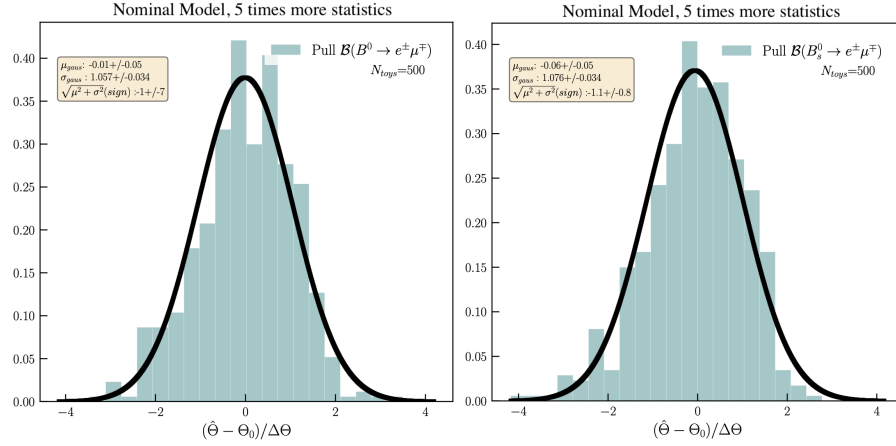


Figure 4.45 – Pull distributions of the $B^0 \rightarrow e^+ \mu^-$ (left) and $B_s^0 \rightarrow e^+ \mu^-$ (right) branching fractions fitted in 500 successful pseudo-experiments generated without $B_{(s)}^0 \rightarrow e^+ \mu^-$ signal components and five times more statistics. The curves represent the results of Gaussian fits.

the widths of the fitted Gaussian functions are very close to one, which shows that the fit is not underestimating the uncertainties. These results state clearly that the $B_s^0 \rightarrow e^+ \mu^-$ branching fraction bias and the large underestimation of the uncertainties observed in Sec. 4.7.5 are related to the low statistics in the nominal data sample.

While the $B_s^0 \rightarrow e^+ \mu^-$ branching fraction bias observed in pseudo-experiments can easily be taken into account as a systematic uncertainty, it is less clear how to deal with the large underestimation of the uncertainties. These issues are currently under discussion, and alternative approaches to incorporate these effects into the determination of the upper limits are considered, *e.g.* setting the limits with the Feldman-Cousins method [94].

Additional physics background

The most dangerous background sources of the analysis are studied in Sec. 4.6.1. However, additional background decays such as $B_c^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-)$ and $B_s^0 \rightarrow K^+ \ell^- \bar{\nu}_\ell$ ($\ell = \mu, e$), as well as B^+ decays with a D meson, might also be present in the signal sample.

The contribution of these decays is suspected to be negligible, since the lower boundary of the chosen fit range of $4.9 \text{ GeV}/c^2$ ensures that part of the backgrounds, which assume on top of missing energy also the loss of an extra particle is not a problem. Furthermore, the kaon misidentification rate with an electron or muon is smaller than the corresponding pion misidentification rate (see Fig. 4.21), which further supports this statement, *e.g.* for $B_s^0 \rightarrow K^+ \ell^- \bar{\nu}_\ell$ decays.

To investigate a possible background contribution from these decay modes, simulated samples are generated with an increased mass range and will be studied.

Table 4.27 – Mean values of the distributions of the expected upper limits for the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions for various assumptions for the total relative systematic uncertainty.

| Total relative systematic uncertainty | Mean exp. limit $B^0 \rightarrow e^\pm \mu^\mp$ (95% CL) | Mean exp. limit $B_s^0 \rightarrow e^\pm \mu^\mp$ (95% CL) |
|--|---|---|
| 0% | $7.3 \cdot 10^{-10}$ | $2.2 \cdot 10^{-9}$ |
| 10% | $7.3 \cdot 10^{-10}$ | $2.2 \cdot 10^{-9}$ |
| 20% | $7.6 \cdot 10^{-10}$ | $2.3 \cdot 10^{-9}$ |
| 30% | $8.1 \cdot 10^{-10}$ | $2.4 \cdot 10^{-9}$ |
| 40% | $9.4 \cdot 10^{-10}$ | $2.7 \cdot 10^{-9}$ |
| 50% | $14.5 \cdot 10^{-10}$ | $3.6 \cdot 10^{-9}$ |

4.9.3 Impact of systematic uncertainties on the limits

The impact of systematic uncertainties on the expected $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction limits is studied with 500 pseudo-experiments, introducing different total relative systematic uncertainties of 10%, 20%, 30%, 40%, and 50%, as shown in Eq. (4.39). The two multiplicative factors are constrained using a Gaussian distribution, where the mean is set to one, while the width of the Gaussian corresponds to the size of the total relative systematic uncertainty.

The evaluation of the expected $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fraction limits at 95% CL are repeated for each tested relative systematic uncertainty value. For the $B_s^0 \rightarrow e^\pm \mu^\mp$ case, the expected upper limit is only re-evaluated without reweighting the B_s^0 lifetime, *i.e.* the light and heavy mass eigenstates are not considered.

The results of this study, shown in Table 4.27, demonstrate that a systematic uncertainty of up to 20% does not impact significantly the expected limits on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions. However, the results also reveal that a relative systematic uncertainty above 30% significantly degrades the expected branching fraction limits. This effect becomes especially visible for a relative systematic uncertainty of 50%, where the mean of the expected limits on the $B^0 \rightarrow e^\pm \mu^\mp$ branching fraction is significantly larger than the results obtained in the Run 1 analysis. Nevertheless, such high relative systematic uncertainties are unrealistic, considering for example that the overall impact of the systematic uncertainties on the branching fraction limits in the Run 1 analysis is found to be below 5% [42], and would lead to the conclusion that there is something problematic in the design of the analysis. The dominant systematic effects in this analysis are expected to arise from the mismodeling of the electron reconstruction and selection which is estimated to be around 10%.

4.10 Summary and conclusion

This thesis describes an analysis for the search of the lepton flavour violating decays $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ with the LHCb Run 2 data collected between 2016 and 2018.

A dedicated selection chain is carefully designed, to minimise the background pollution. To reduce background arising from misidentified decays, *e.g.* $B^0 \rightarrow K^+ \pi^-$ or $B^0 \rightarrow \pi^+ \pi^-$, particle identification requirements are applied, while a BDT is trained to remove as much combinatorial background as possible. The latest isolation variables, which have been developed especially for the search of $B_s^0 \rightarrow \mu^+ \mu^-$ and provide an excellent discrimination power, are included in the BDT training. A loose BDT requirement is determined to remove most of the combinatorial background. Furthermore, intervals of the BDT response are defined, which are taken into account in the final fit of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions.

The normalisation yield is extracted with a simultaneous fit to $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ data, which ensures that the misidentification rate arising from $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) \pi^+$ is well under control.

Selection efficiencies are determined from simulated samples while taking data-driven corrections into account. Corrections for the electron and muon track reconstruction are determined with tag-and-probe methods. For the L0 trigger efficiency, correction weights are obtained for the L0MuonDecision_TOS and the L0ElectronDecision_TOS trigger lines with the TISTOS method. Furthermore, data-driven PID efficiencies are determined for all PID requirements applied in the analysis. While the muon and kaon PID efficiencies are determined with the PIDCalib2 framework, the electron PID efficiencies are determined with the fit-and-count method. The electron tracking and L0ElectronDecision_TOS trigger corrections, as well as the electron PID efficiencies, are evaluated and applied separately for the bremsstrahlung categories 0γ and 1γ . Furthermore, the tracking and L0 trigger corrections, as well as the PID efficiencies, are applied on a per-track basis, which is a state-of-the-art method and used in high profile analysis like the lepton universality tests at LHCb. Finally, weights are determined to correct for the observed discrepancies in $B_{(s)}^0$ kinematic variables and the track multiplicity between background subtracted data and simulation. The kinematic corrections are determined with $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ data and simulation and then ported to the simulated $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ and $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ samples.

To test the validity of the obtained corrections, the ratio $r_{J/\psi}$ of the $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ branching fractions for $\ell = \mu, e$ is computed, which shows a significant improvement when applying the corrections, while still a residual deviation from unity of approximately 10% is observed. This discrepancy is suspected to arise from phase-space regions enhanced by the different selections applied to $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ and $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ and may not be relevant for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$. Consequently, the validation power of $r_{J/\psi}$ is limited. Nevertheless, the observed improvement of $r_{J/\psi}$, when taking all corrections into account, is significant, which increases the confidence in the validity of the corrections applied to simulation. Furthermore, the observed offset from unity is used to evaluate the systematic uncertainties arising from mismodeling effects.

The background pollution, arising from misidentified particles and surviving the selection, is estimated. The results show that a significant amount of background is expected from $B^0 \rightarrow$

$K^+ \pi^-$, $B^0 \rightarrow \pi^+ \pi^-$, $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ and $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ decays, which are therefore included in the fit model of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions.

The validity of the predicted background is tested with two independent methods. Firstly, a cross-check is performed with $B_{(s)}^0 \rightarrow hh'$ data and simulation, which confirms the validity of the pion and kaon misidentification within an accuracy of 10–20%. Only minor over/under estimating effects are observed, which are at a level that is not concerning. Secondly, the pass-fail method is used to validate the kaon and pion misidentifications by taking directly the predicted level of $B^0 \rightarrow hh'$ events into account. The pass-fail method uses blinded $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data and does not rely on simulation. Its result confirms the validity of the background predictions.

The $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions are evaluated with a simultaneous fit to the invariant B mass distribution in six data samples, which are defined by three BDT regions and the two bremsstrahlung categories 0γ and 1γ . Pseudo-experiments are drawn to test the fit stability and the absence of biases, as well as to derive the expected limits on the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions.

The fit stability and robustness are found to be very high since more than 95% of the pseudo-experiments lead to a converged fit. A small bias is observed in the pulls of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ branching fractions, as well as a large underestimation of the uncertainties. A further study with pseudo-experiments shows clearly that these effects are related to the low statistics of the data sample and an adequate response to take the observed issues into account will be further investigated.

The 95% (90%) CL upper limits expected in absence of signals and without considering any systematic uncertainty, show an improvement, with respect to Run 1, by a factor 1.6 (1.5) for the $B^0 \rightarrow e^\pm \mu^\mp$ branching fraction and a factor of 2.8 (2.9) for the $B_s^0 \rightarrow e^\pm \mu^\mp$ branching fraction (if the decay amplitude is dominated by the heavy mass eigenstate). These improvements are roughly compatible with the naive expectations considering only the increase in statistics.

The strategy to derive the systematic uncertainties is presented. Furthermore, a study is performed to investigate the impact of the relative systematic uncertainties on the branching fraction expected limits, revealing that the expected upper limits show only minor changes for a relative systematic uncertainty up to 20%. The final relative systematic uncertainty is expected to be at most of this size, since the dominant source is suspected to be the mismodeling of the electron reconstruction which is approximately 10%.

Despite the significant improvement of the expected upper branching fraction limits, the sensitivity of this search with Run 2 data is still not high enough to test the $\mathcal{O}(10^{-12})$ regime where new physics models predict $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$. While the presented search is statistically limited, it holds great promise for further development and enhancement with the inclusion of the larger sample that LHCb collaboration plans to collect during Run 3. The results of this analysis are still blinded. Therefore, this conclusion is limited to the expected results and

needs to be re-assessed, once the data is unblinded. A statistically significant determination of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ yields in the fit to the unblinded data would be very exciting and a clear sign of new physics.

A Appendix

A.1 Multivariate analysis

A.1.1 Track isolation

In Sec. 4.3.7, the long and VELO track isolation variables (see also Sec. 2.2.3) are used, since they provide a significant discrimination between signal and combinatorial background. These track isolation variables were originally designed for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ analysis [95] and are applied in this analysis without any modifications and are used to separate signal from combinatorial background, under the assumption that signal tracks are isolated from other tracks of the event compared to background tracks.

A Boosted Decision Tree is trained to separate the isolated from the non-isolated tracks. In the classifier training, $B_s^0 \rightarrow \mu^+ \mu^-$ simulation is used as a proxy for isolated tracks, while for the non-isolated tracks inclusive B hadron simulation is used ($b\bar{b} \rightarrow \mu^+ \mu^- X$).

The long and VELO track isolation variables are determined by evaluating the BDT response for each long, respectively, VELO track with respect to the two final state particles, which are in the present case the electron and muon tracks of the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ signal.

A.1.2 Sculpting of the combinatorial background

Figure A.1 shows the results of the combinatorial background study of Sec. 4.3.7 for the data-taking years 2017 and 2018.

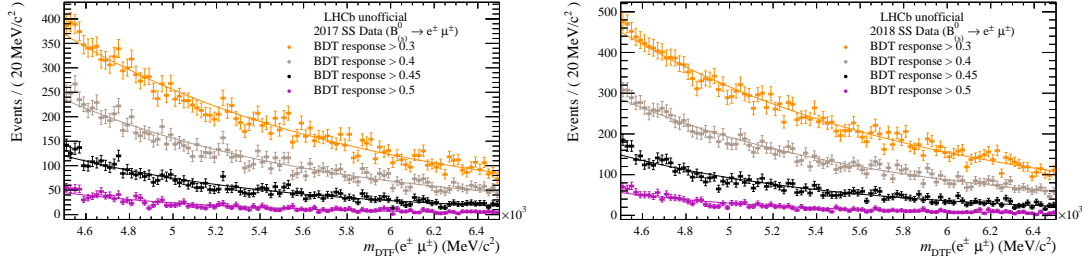


Figure A.1 – Invariant mass distribution of same-sign $B_{(s)}^0 \rightarrow e^\pm \mu^\pm$ candidates in the 2017 (left) and 2018 (right) data, for various requirements on the BDT response.

A.2 Normalisation fits

The fit to $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) \pi^+$ simulation, to extract the mass shapes as explained in Sec. 4.4, are presented in Figs. A.2 and A.3 for the data-taking years 2017 and 2018.

A.3 Efficiencies and corrections

A.3.1 Geometrical acceptance

Table A.1 lists the geometrical acceptance of the various background decay modes.

Table A.1 – Geometrical acceptance ε_{geo} in % for the simulated background decays.

| Decay | 2016 | 2017 | 2018 |
|---|-----------------------|-----------------------|-----------------------|
| $B^0 \rightarrow K^+ \pi^-$ | 0.1964 ± 0.0004 | 0.1966 ± 0.0004 | 0.1965 ± 0.0005 |
| $B^0 \rightarrow \pi^+ \pi^-$ | 0.1931 ± 0.0005 | 0.1937 ± 0.0005 | 0.1937 ± 0.0005 |
| $B^0 \rightarrow p \bar{p}$ | 0.21334 ± 0.00006 | 0.21400 ± 0.00005 | 0.21393 ± 0.00006 |
| $B_s^0 \rightarrow \pi^+ K^-$ | 0.1963 ± 0.0005 | 0.1969 ± 0.0005 | 0.1956 ± 0.0005 |
| $B_s^0 \rightarrow K^+ K^-$ | 0.1994 ± 0.0005 | 0.1995 ± 0.0005 | 0.1984 ± 0.0005 |
| $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ | 0.1808 ± 0.0009 | 0.1799 ± 0.0009 | 0.1813 ± 0.0009 |
| $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ | 0.00701 ± 0.00002 | 0.00699 ± 0.00002 | 0.00699 ± 0.00003 |
| $\Lambda_b^0 \rightarrow p K^-$ | 0.2029 ± 0.0006 | 0.2033 ± 0.0005 | 0.2043 ± 0.0011 |
| $\Lambda_b^0 \rightarrow p \pi^-$ | 0.2009 ± 0.0006 | 0.2006 ± 0.0005 | 0.2022 ± 0.0011 |
| $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ | 0.01322 ± 0.00005 | 0.01321 ± 0.00005 | 0.01325 ± 0.00005 |
| $\Lambda_b^0 \rightarrow p e^- \bar{\nu}_e$ | 0.2152 ± 0.0006 | 0.2152 ± 0.0005 | 0.2146 ± 0.0006 |

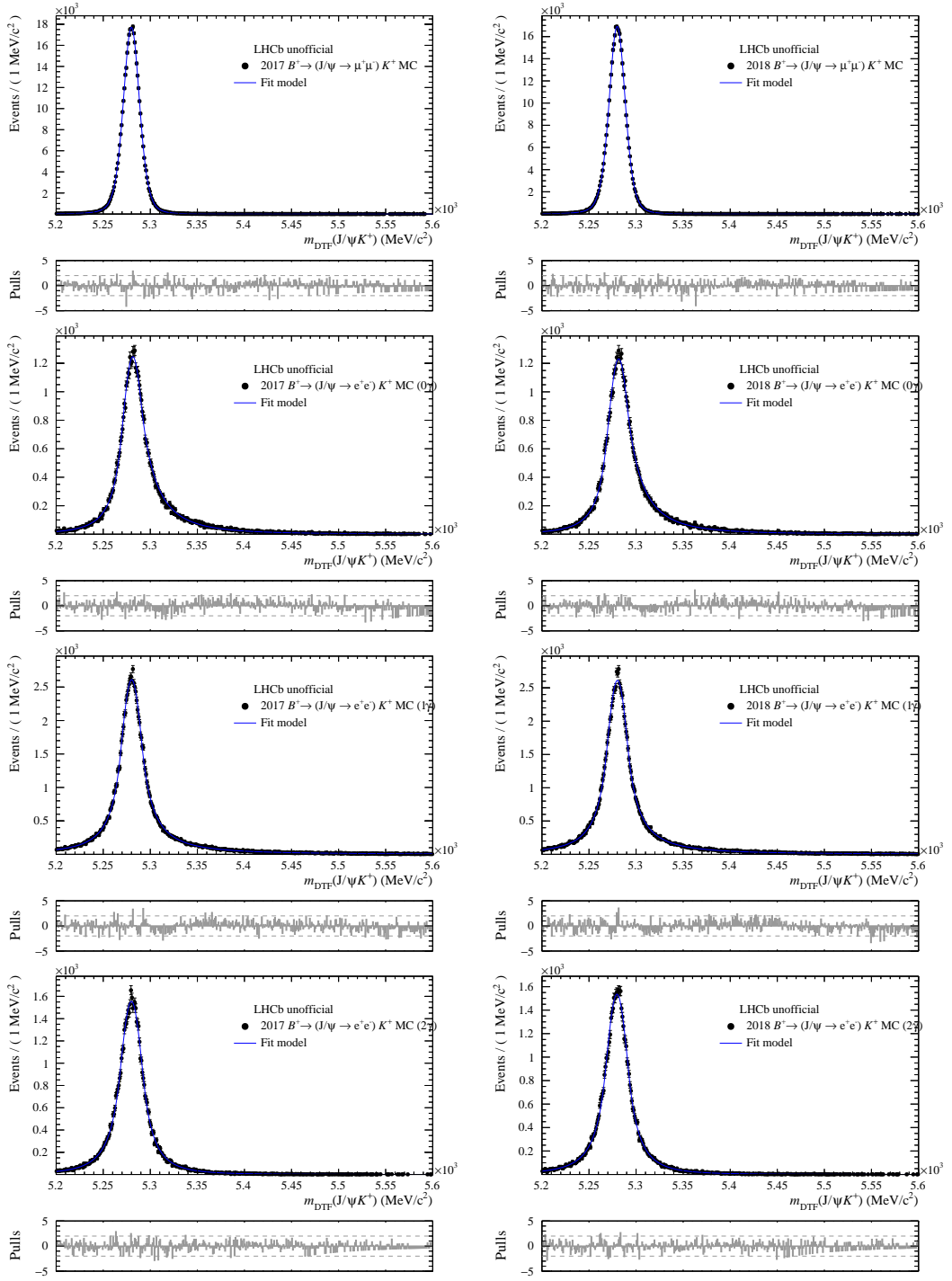


Figure A.2 – Fitted $m_{\text{DTF}}(J/\psi K^+)$ distribution of $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ (0γ second row, 1γ third row, 2γ fourth row) for the data-taking year 2017 (left) and 2018 (right).

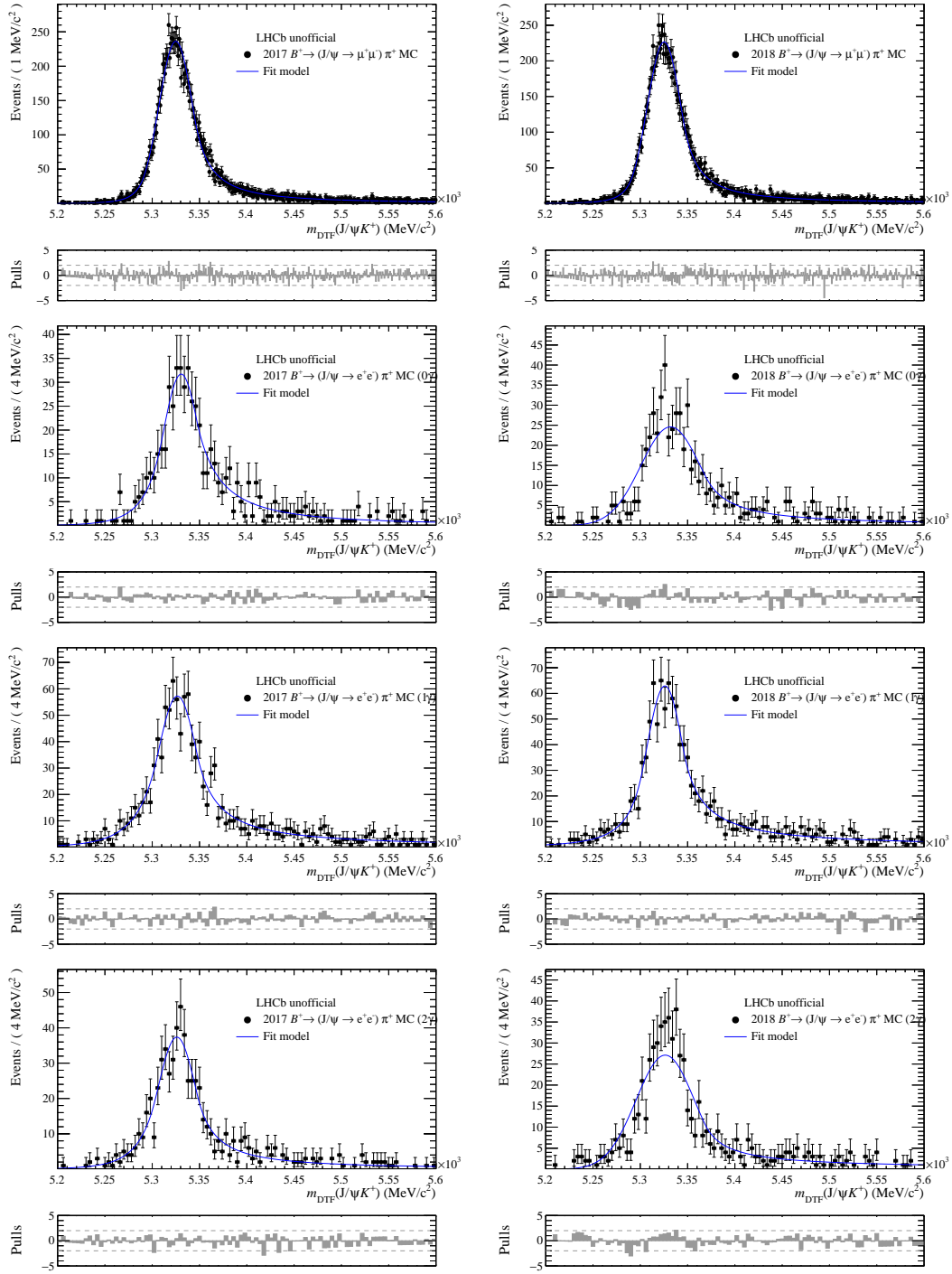


Figure A.3 – Fitted $m_{\text{DTF}}(J/\psi K^+)$ distribution of $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \pi^+$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) \pi^+$ (0γ second row, 1γ third row, 2γ fourth row) for the data-taking year 2017 (left) and 2018 (right).

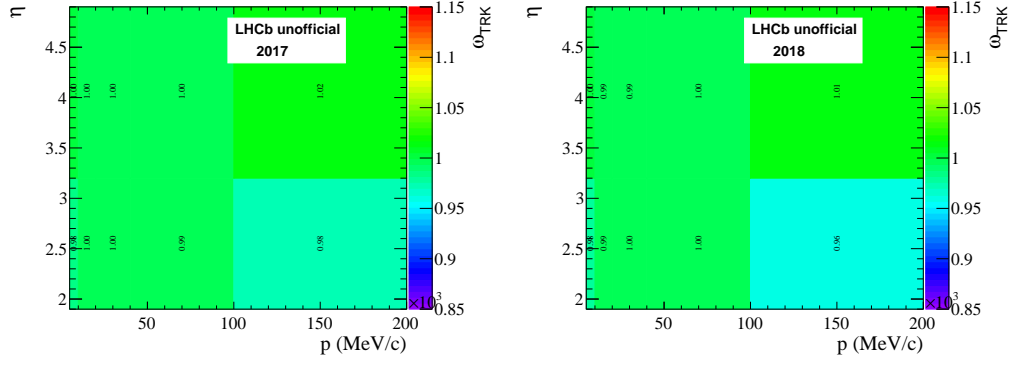


Figure A.4 – 2017 (left) and 2018 (right) muon tracking efficiency weight, ω_{TRK} , as a function of the muon momentum and pseudorapidity.

A.3.2 Track reconstruction efficiency corrections

The muon and electron tracking efficiency weights are shown in Figs. A.4 and A.5 for the data-taking years 2017 and 2018, respectively.

A.3.3 PID efficiencies

The muon PID efficiencies for the data-taking years 2017 and 2018 are shown in Fig. A.6 as a function of the muon p_T and η .

Figures A.7 Fig. A.8 show the results of the one-dimensional electron PID efficiency studies are for the data-taking years 2017 and 2018. The final electron PID efficiencies determined for the data-taking years 2017 and 2018 are shown in Figs. A.9 and A.10, for the bremsstrahlung categories 0γ and 1γ , respectively.

The kaon PID efficiencies determined in five nTracks regions as a function of the kaon momentum and pseudorapidity are shown in Figs. A.11 and A.12 for the data-taking years 2017 and 2018, respectively.

The PID efficiencies for true pions, kaons, protons and electrons misidentified as muons, as well as for true pions, kaons, protons and muons misidentified as electrons, are shown in Figs. A.13 and A.14 for the data-taking years 2017 and 2018, respectively.

Efficiencies of the `isMuon` requirement for true muons, electrons, protons, pions and kaons are shown as a function of the transverse momentum and pseudorapidity in Figs. A.15–A.17 for the data-taking years 2016, 2017 and 2018, respectively.

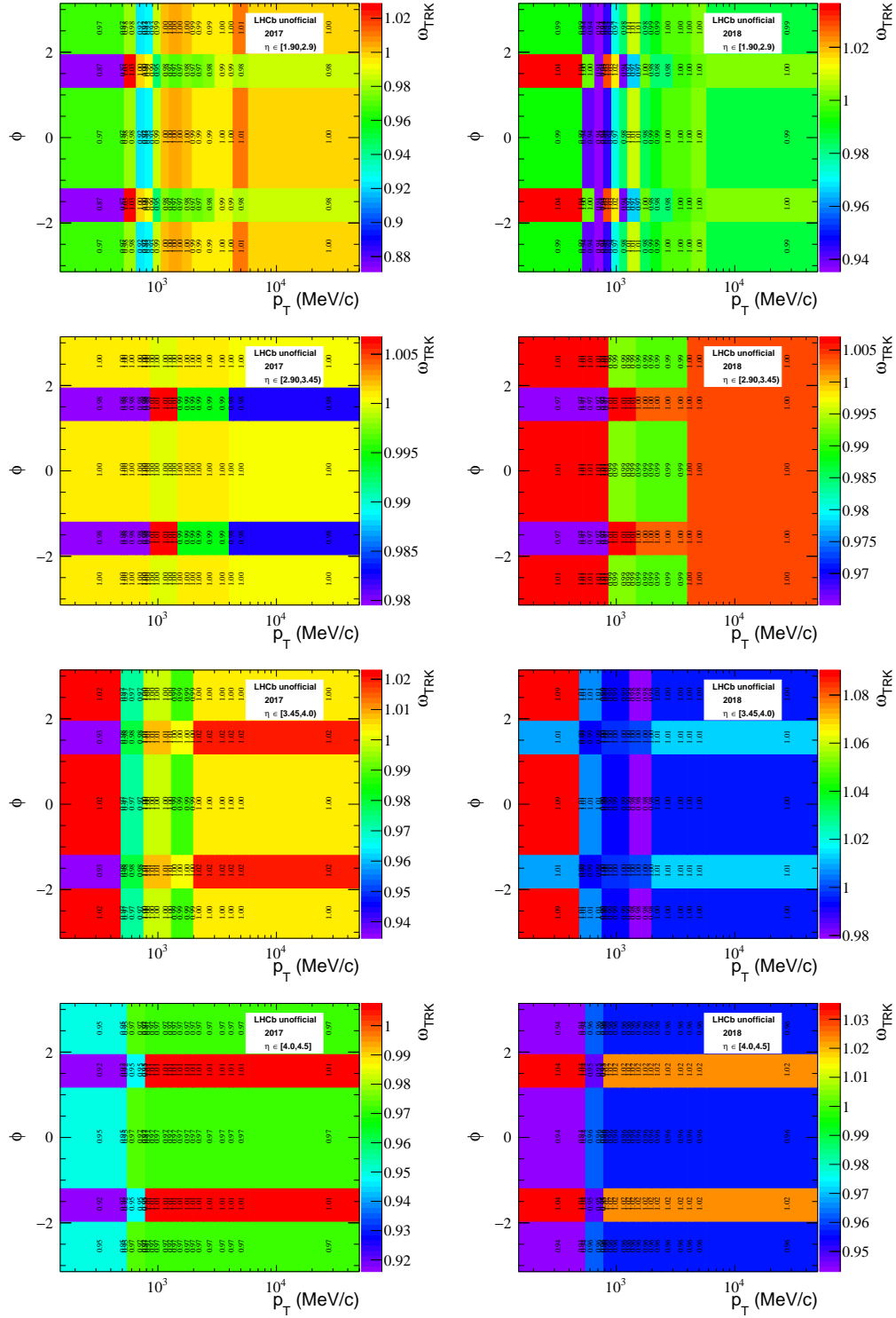


Figure A.5 – 2017 (left) and 2018 (right) electron tracking efficiency weight, ω_{TRK} , as a function of the electron p_T and ϕ in four regions of electron η .

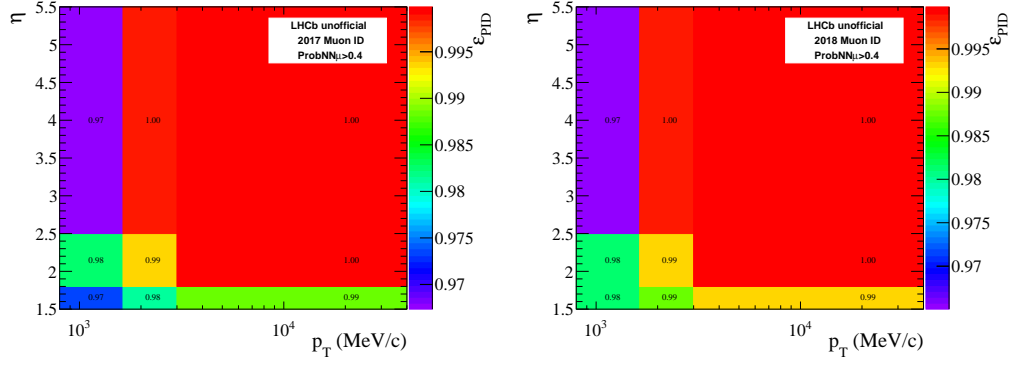


Figure A.6 – Muon PID efficiencies as a function of the muon transverse momentum p_T and pseudorapidity η , determined for the data-taking years 2017 (left) and 2018 (right).

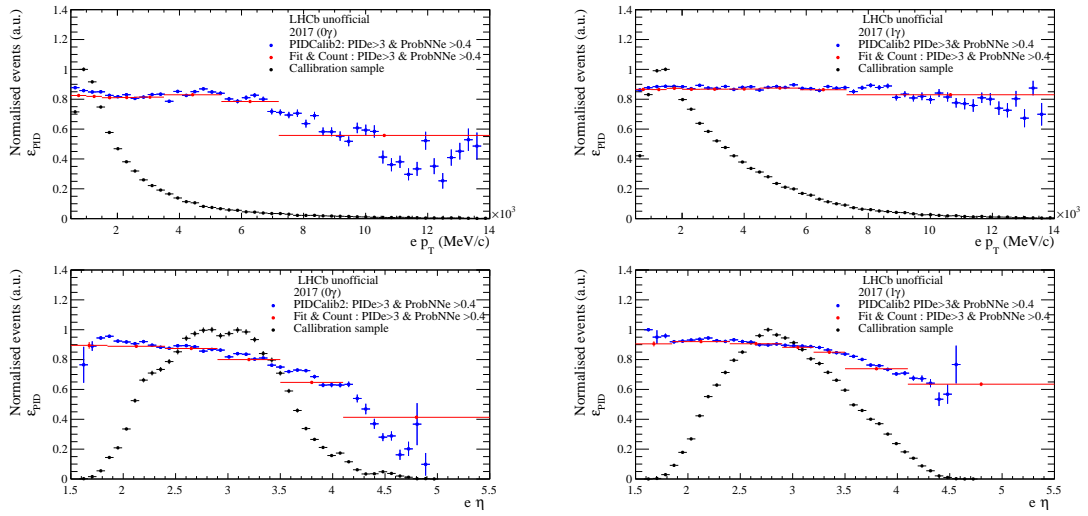


Figure A.7 – 2017 electron PID efficiencies, obtained with the fit-and-count method (red) and PIDCalib2 (blue) as a function of the electron transverse momentum (top) and the pseudorapidity (bottom) for the bremsstrahlung categories 0γ (left) and 1γ (right). The electron p_T and η distributions of the calibration sample are also shown.

Appendix A. Appendix

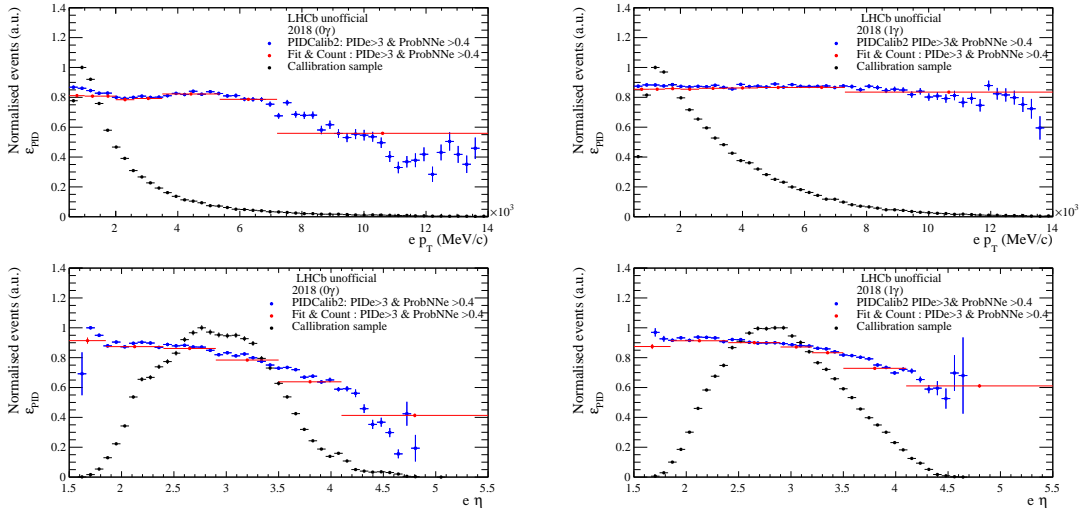


Figure A.8 – 2018 electron PID efficiencies, obtained with the fit-and-count method (red) and PIDCalib2 (blue) as a function of the electron transverse momentum (top) and the pseudorapidity (bottom) for the bremsstrahlung categories 0γ (left) and 1γ (right). The electron p_T and η distributions of the calibration sample are also shown.

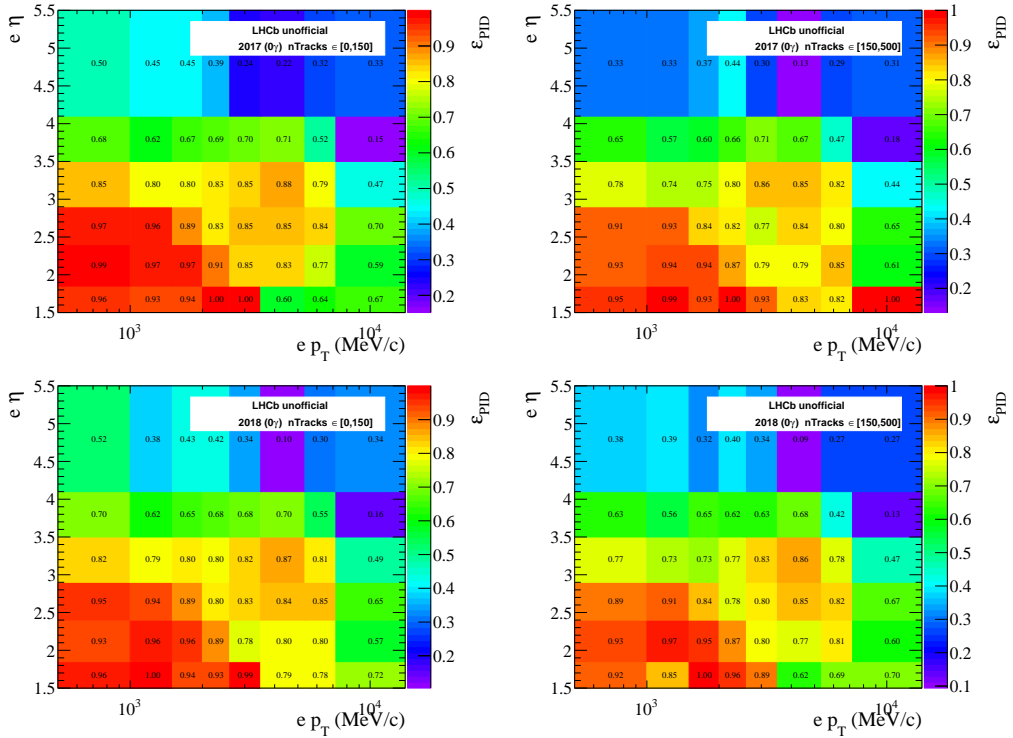


Figure A.9 – Electron PID efficiencies determined for the data-taking years 2017 (top) and 2018 (bottom) for the bremsstrahlung category 0γ as a function of the electron transverse momentum p_T and pseudorapidity η , for the low (left) and high (right) $nTracks$ regions.

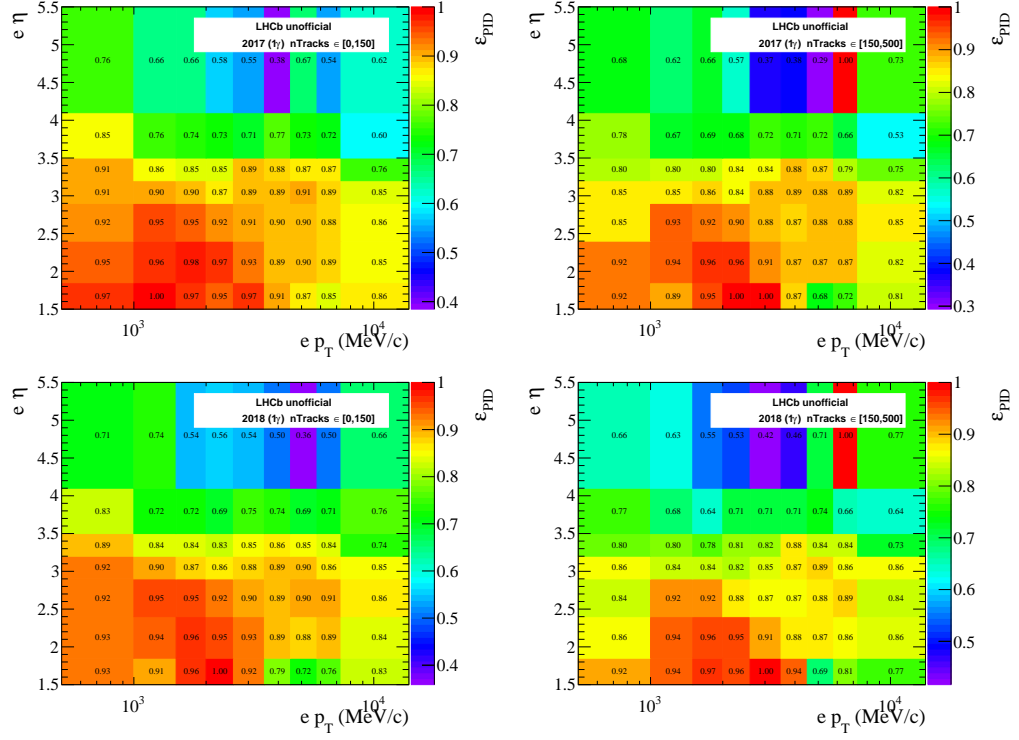


Figure A.10 – Electron PID efficiencies determined for the data-taking years 2017 (top) and 2018 (bottom) for the bremsstrahlung category 1γ as a function of the electron transverse momentum p_T and pseudorapidity η , for the low (left) and high (right) nTracks regions.

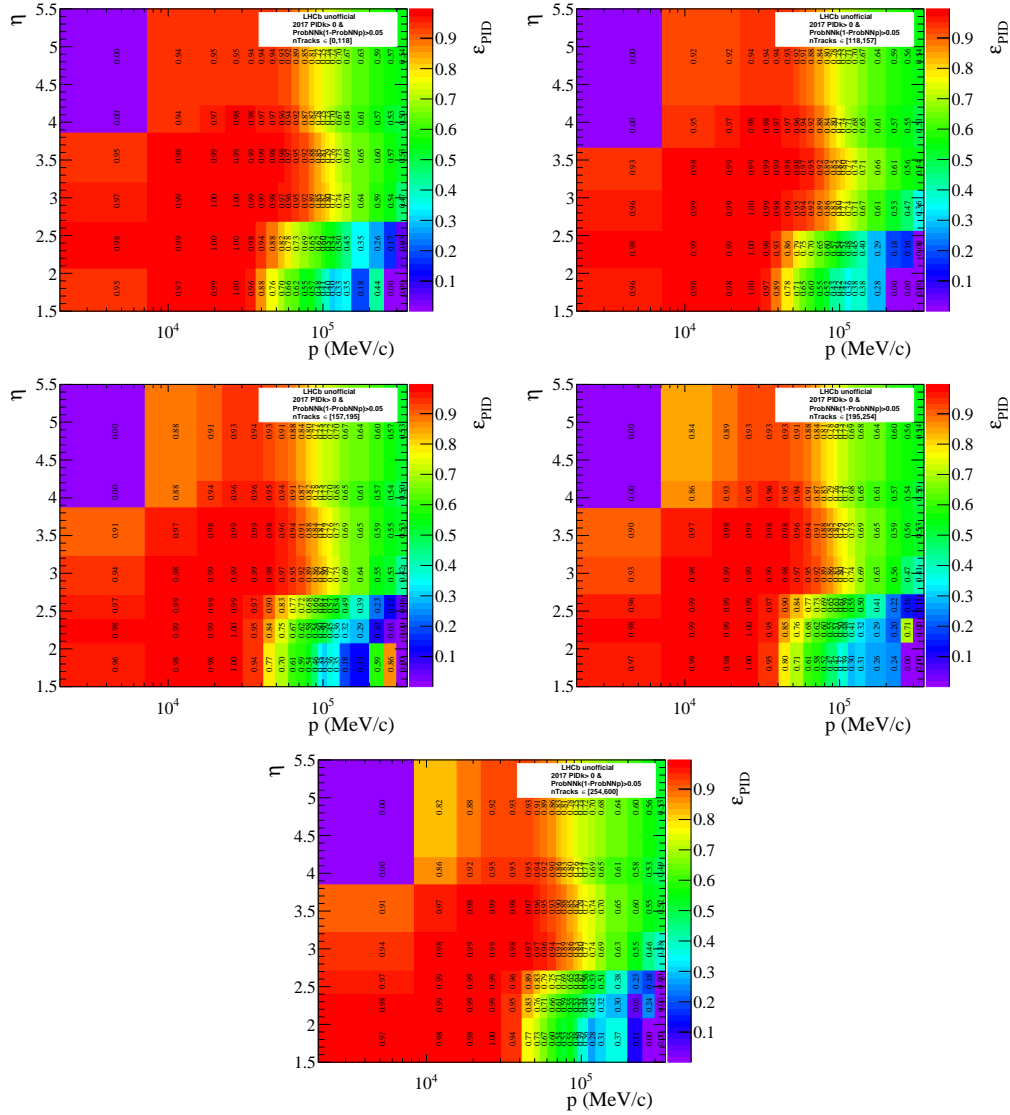


Figure A.11 – Kaon PID efficiencies for the data-taking year 2017 as a function of the kaon momentum and pseudorapidity for the five regions in increasing track multiplicity (from top to bottom and left to right) [34, 35].

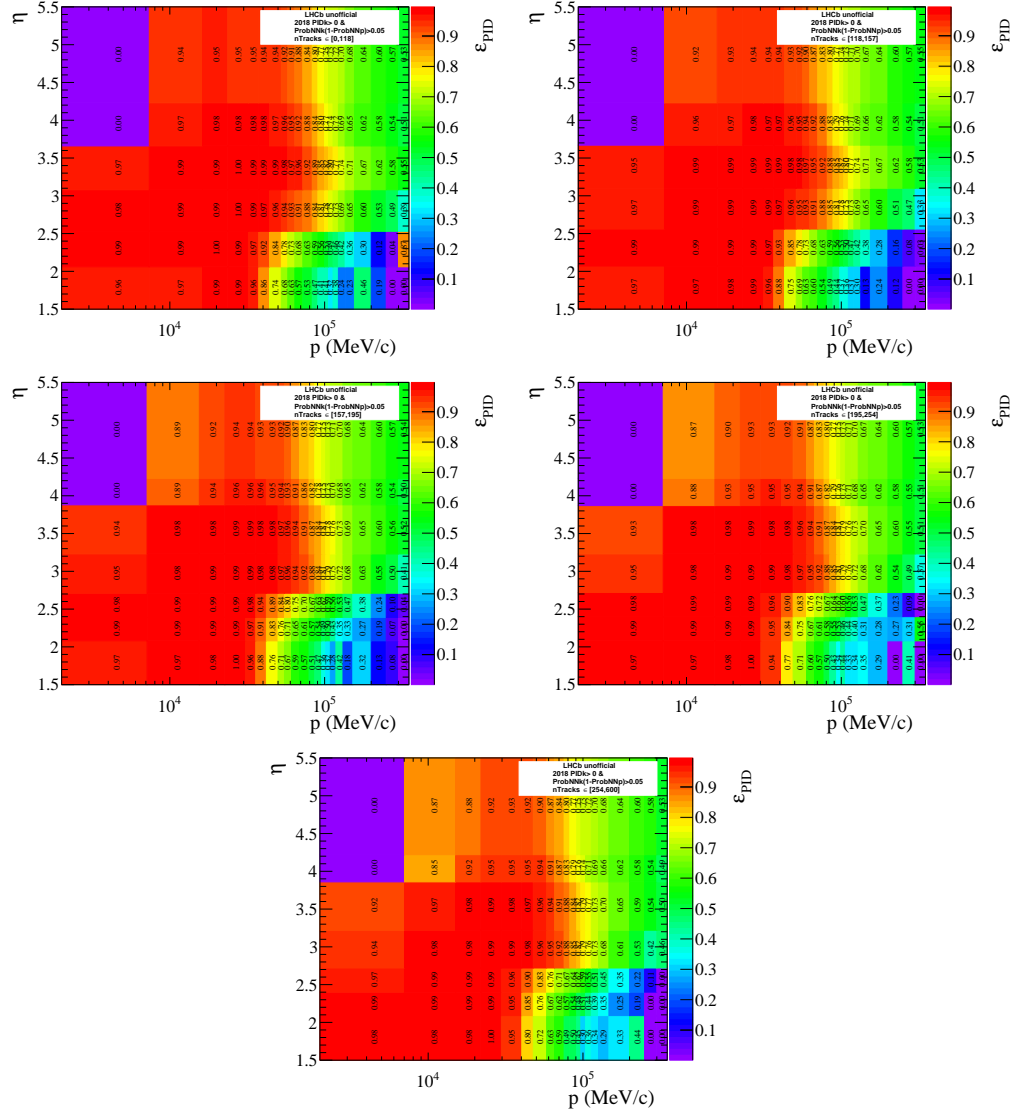


Figure A.12 – Kaon PID efficiencies for the data-taking year 2018 as a function of the kaon momentum and pseudorapidity for the five regions in increasing track multiplicity (from top to bottom and left to right) [34, 35].

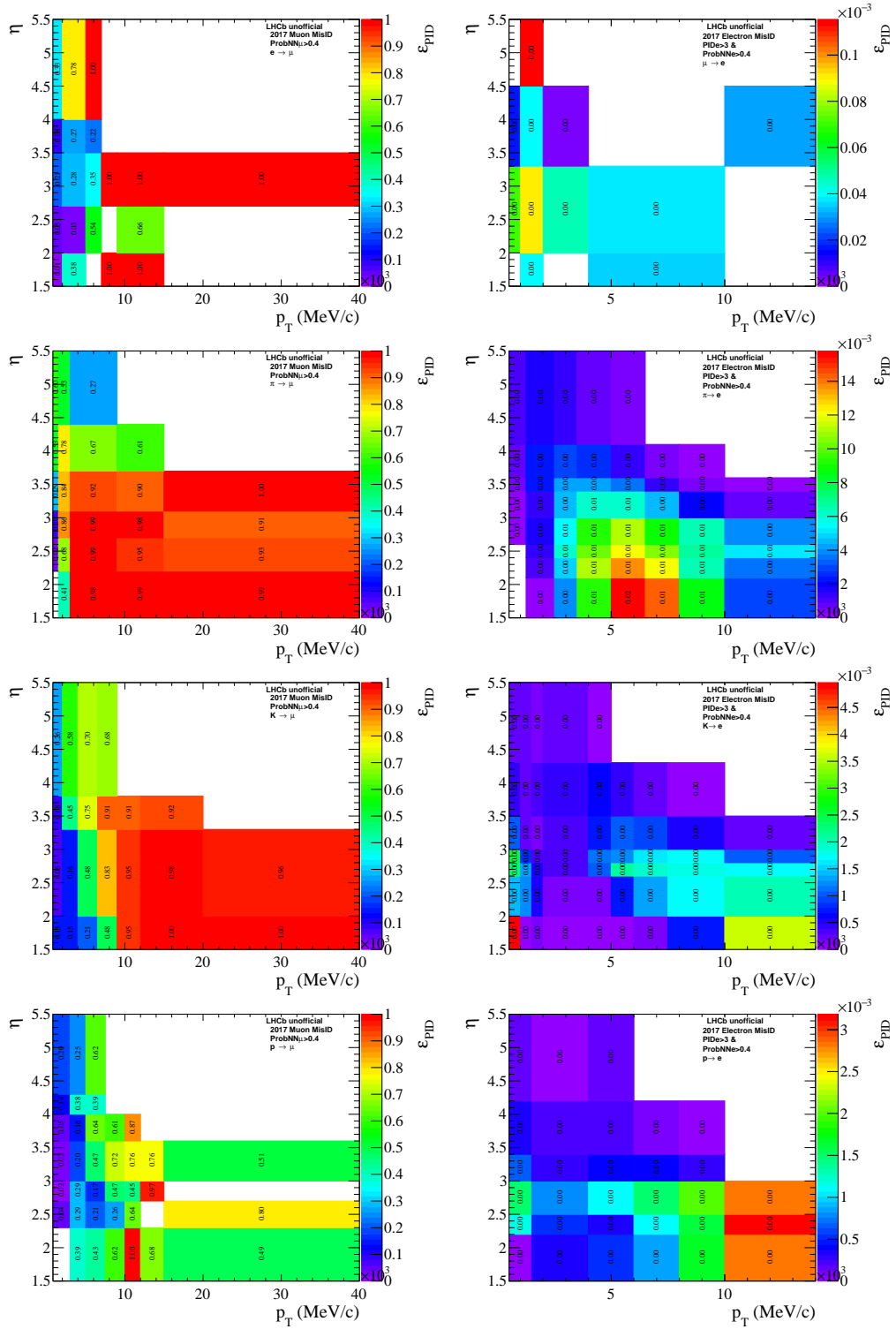


Figure A.13 – Left: 2017 efficiencies for the misidentification of an electron, a pion, a kaon or a proton (from top to bottom) as a muon. Right: 2017 efficiencies for the misidentification a muon, a pion, a kaon or a proton (from top to bottom) as an electron. The empty bins (white) are caused by the low statistics regions in the calibration datasets.

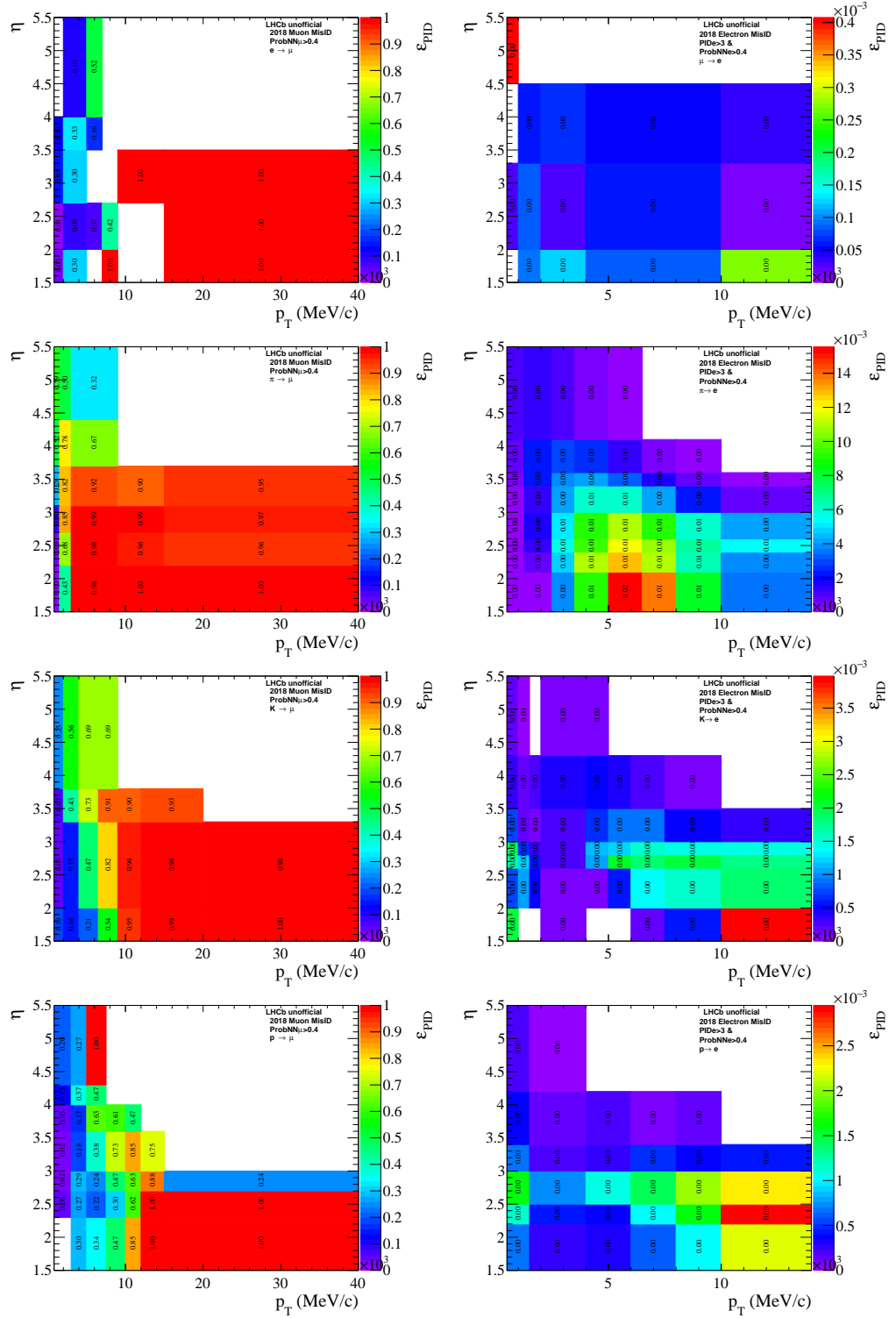


Figure A.14 – Left: 2018 efficiencies for the misidentification of an electron, a pion, a kaon or a proton (from top to bottom) as a muon. Right: 2018 efficiencies for the misidentification a muon, a pion, a kaon or a proton (from top to bottom) as an electron. The empty bins (white) are caused by the low statistics regions in the calibration datasets.

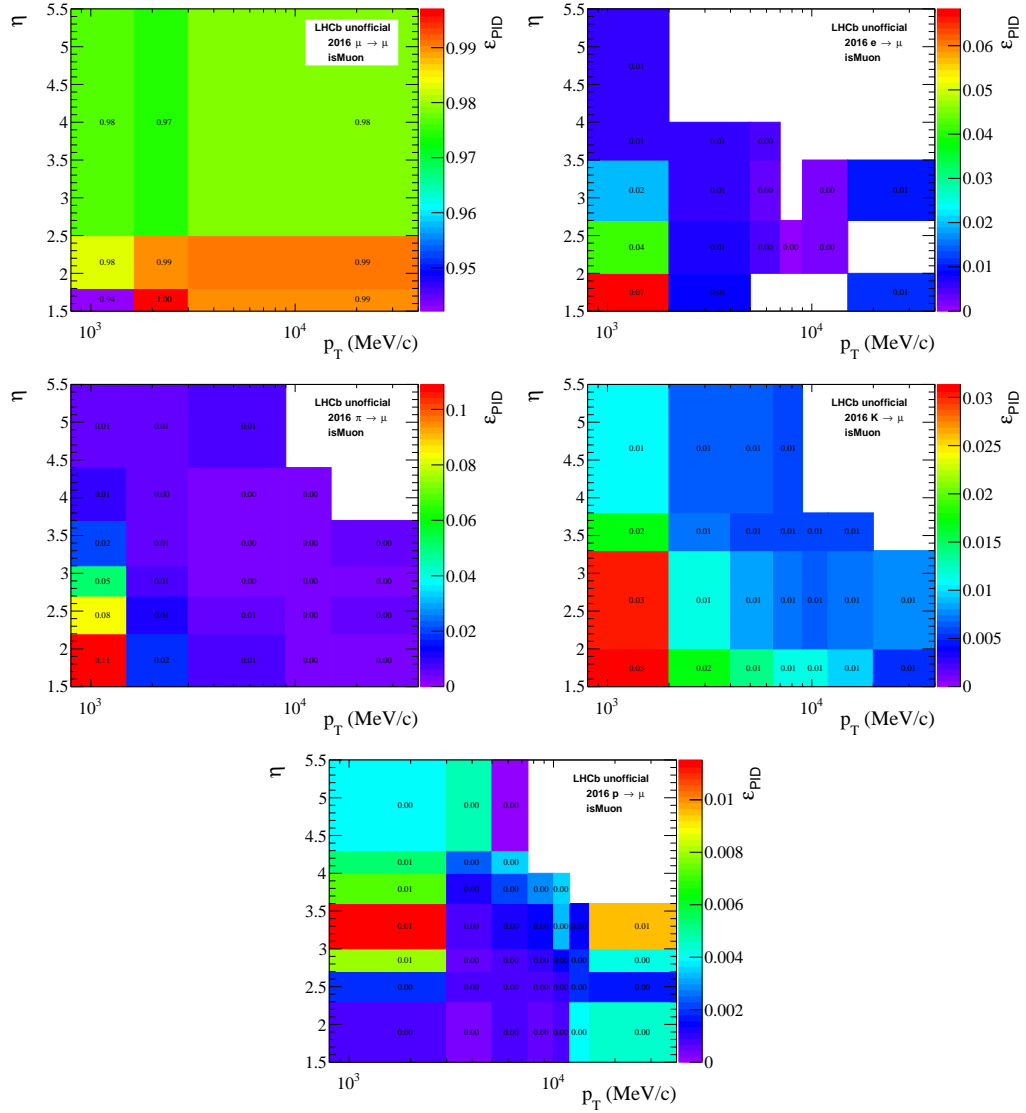


Figure A.15 – Efficiencies of the isMuon requirement, determined for the data-taking year 2016 as a function of p_T and η for true muons (top left), electrons (top right), pions (middle left), kaons (middle right) and protons (bottom).

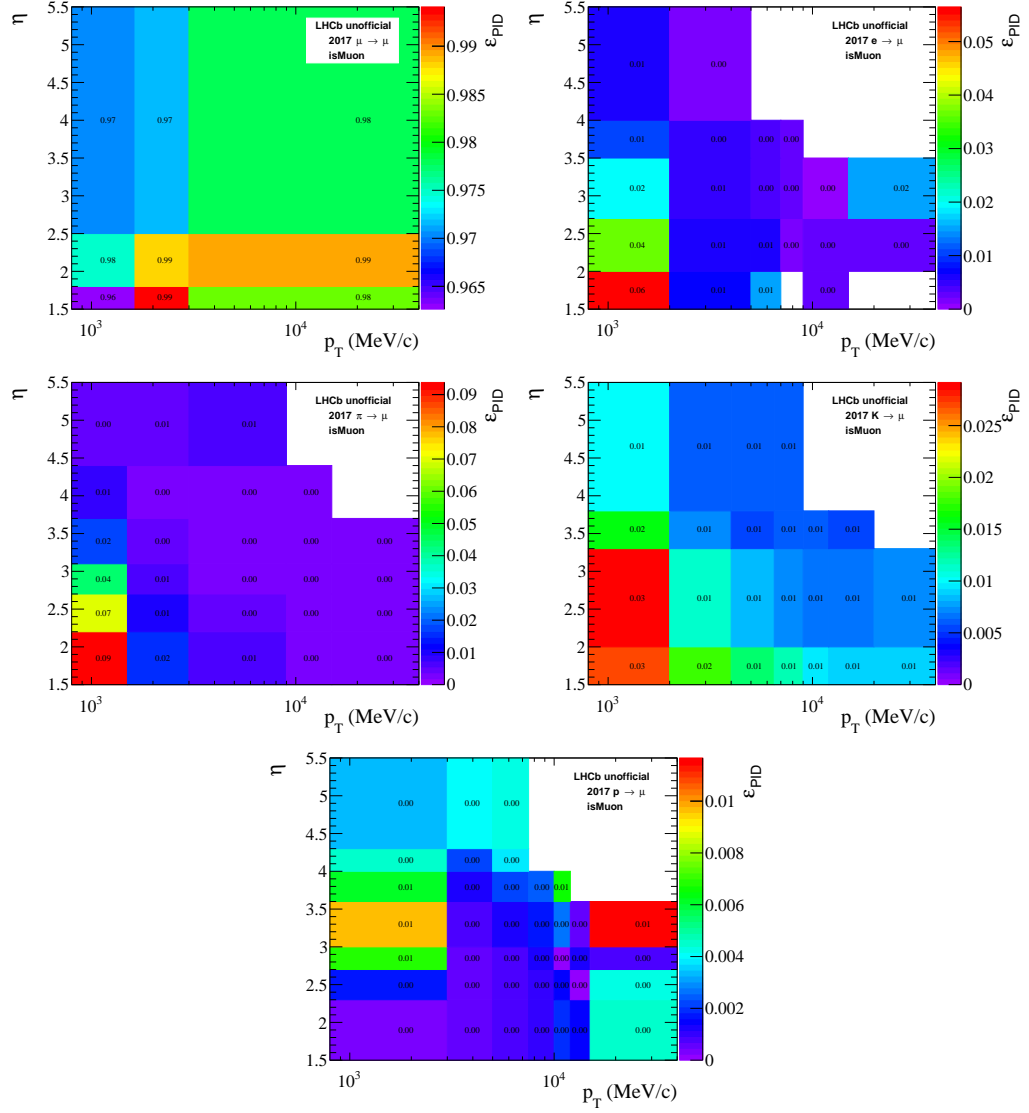


Figure A.16 – Efficiencies of the `isMuon` requirement, determined for the data-taking year 2017 as a function of p_T and η for true muons (top left), electrons (top right), pions (middle left), kaons (middle right) and protons (bottom).

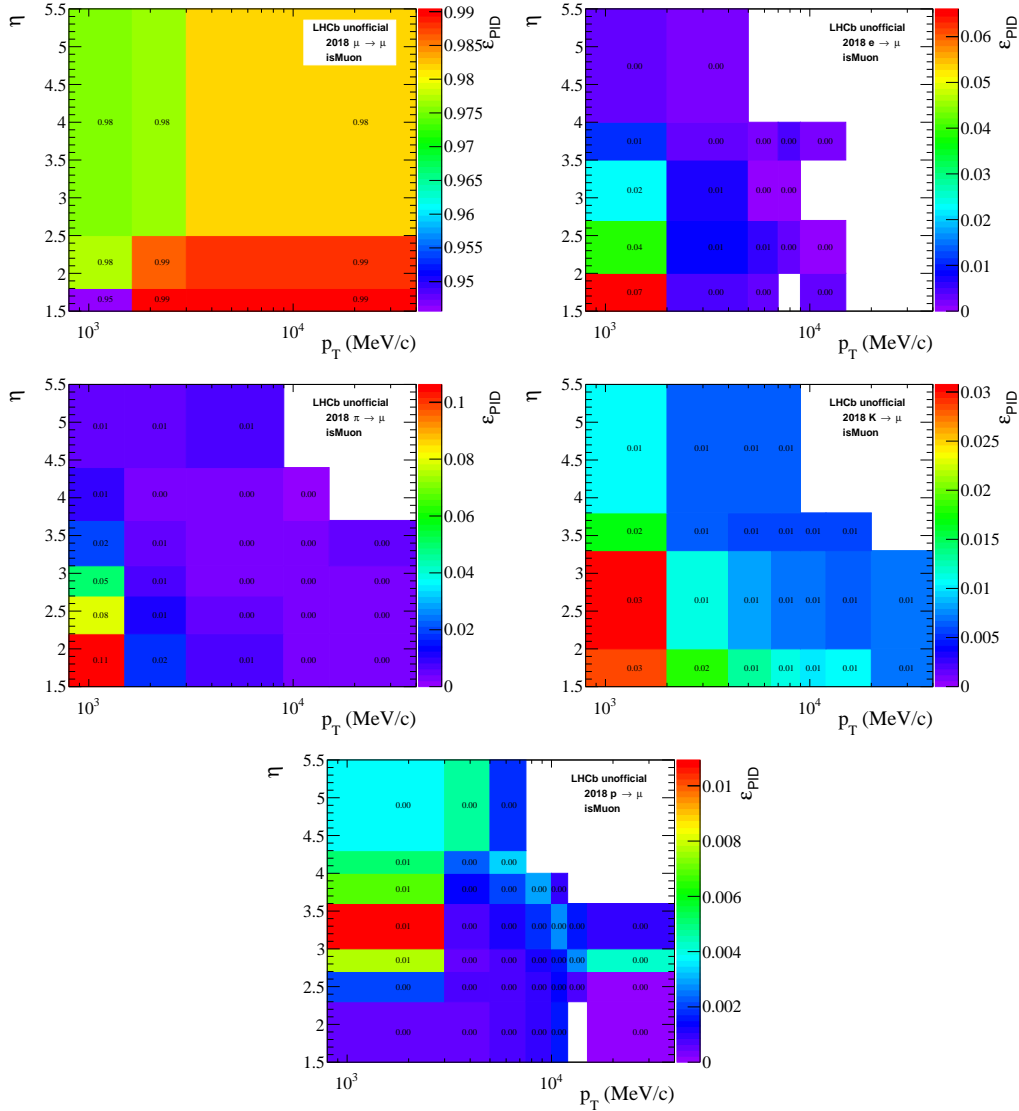


Figure A.17 – Efficiencies of the `isMuon` requirement, determined for the data-taking year 2018 as a function of p_T and η for true muons (top left), electrons (top right), pions (middle left), kaons (middle right) and protons (bottom).

A.3.4 L0 trigger efficiency corrections

The L0 muon trigger efficiencies as a function of the muon p_T and η are shown in Figs. A.18 and A.19 for the data-taking year 2017 and 2018, respectively.

The L0 electron trigger efficiencies as a function of the transversal energy deposited in the different ECAL regions are shown in Fig. A.20 (0γ , 2017 simulation), Fig. A.21 (1γ , 2017 simulation), Fig. A.22 (0γ , 2017 data), Fig. A.23 (1γ , 2017 data), Fig. A.24 (0γ , 2018 simulation), Fig. A.25 (1γ , 2018 simulation), Fig. A.26 (0γ , 2018 data) and Fig. A.27 (1γ , 2018 data).

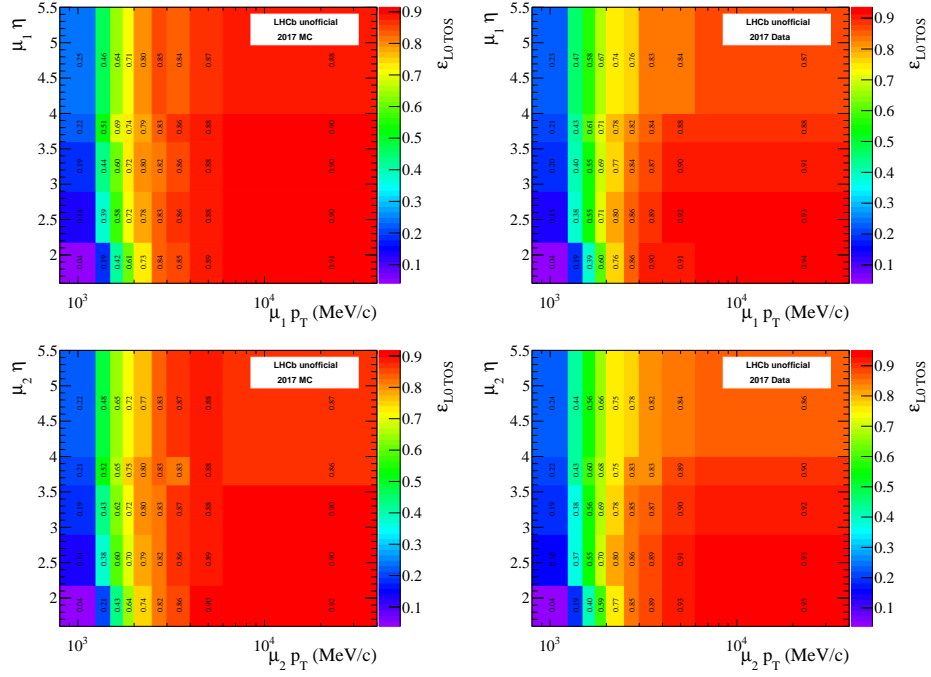


Figure A.18 – L0 trigger efficiency ϵ^{TOS} determined from 2017 $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ simulation (left) and data (right) for both the μ^+ (top) and μ^- (bottom) tracks, as a function of the muon p_T and η .

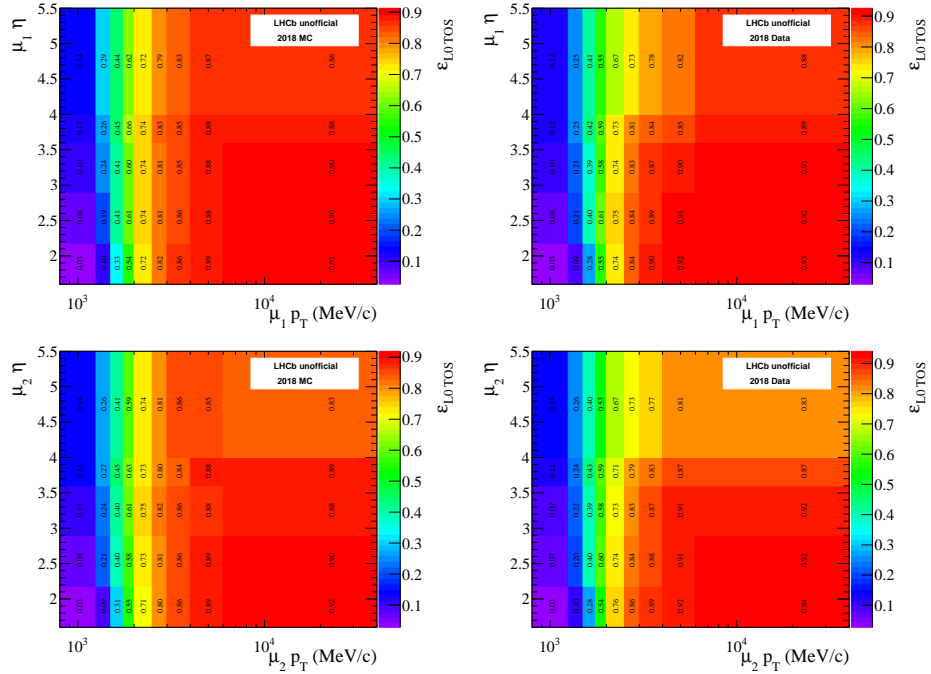


Figure A.19 – L0 trigger efficiency ϵ^{TOS} determined from 2018 $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ simulation (left) and data (right) for both the μ^+ (top) and μ^- (bottom) tracks, as a function of the muon p_T and η .

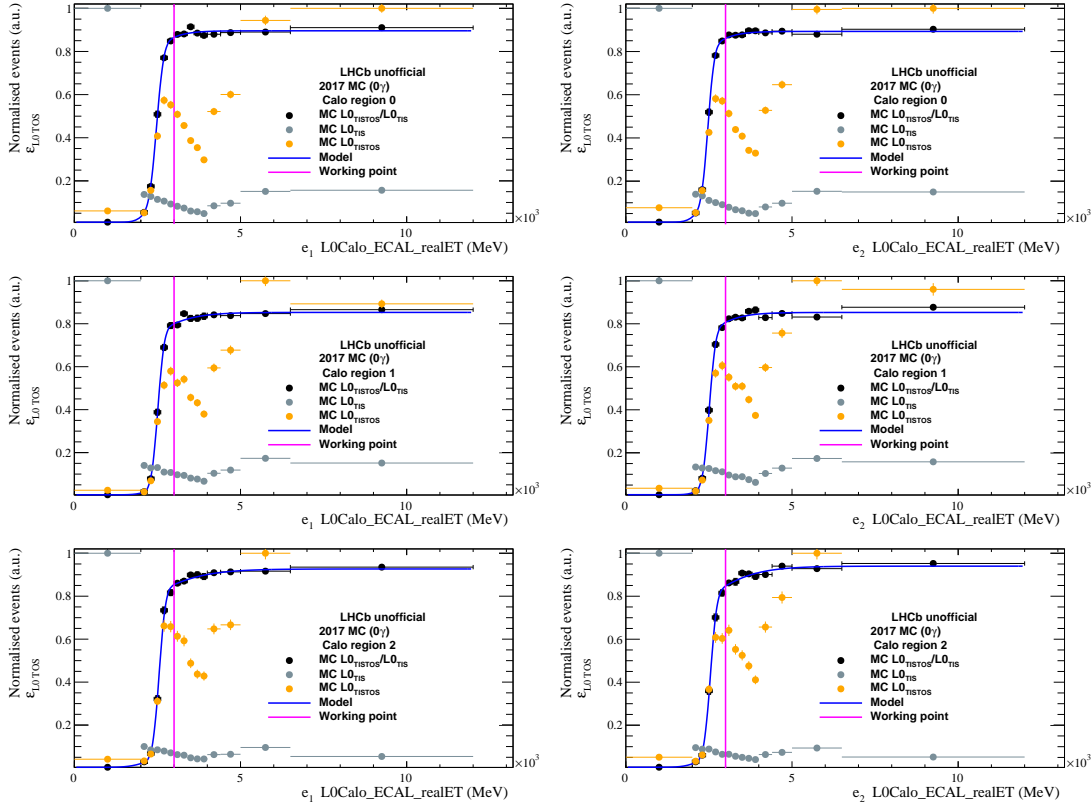


Figure A.20 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 0γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ simulation for the data-taking year 2017, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

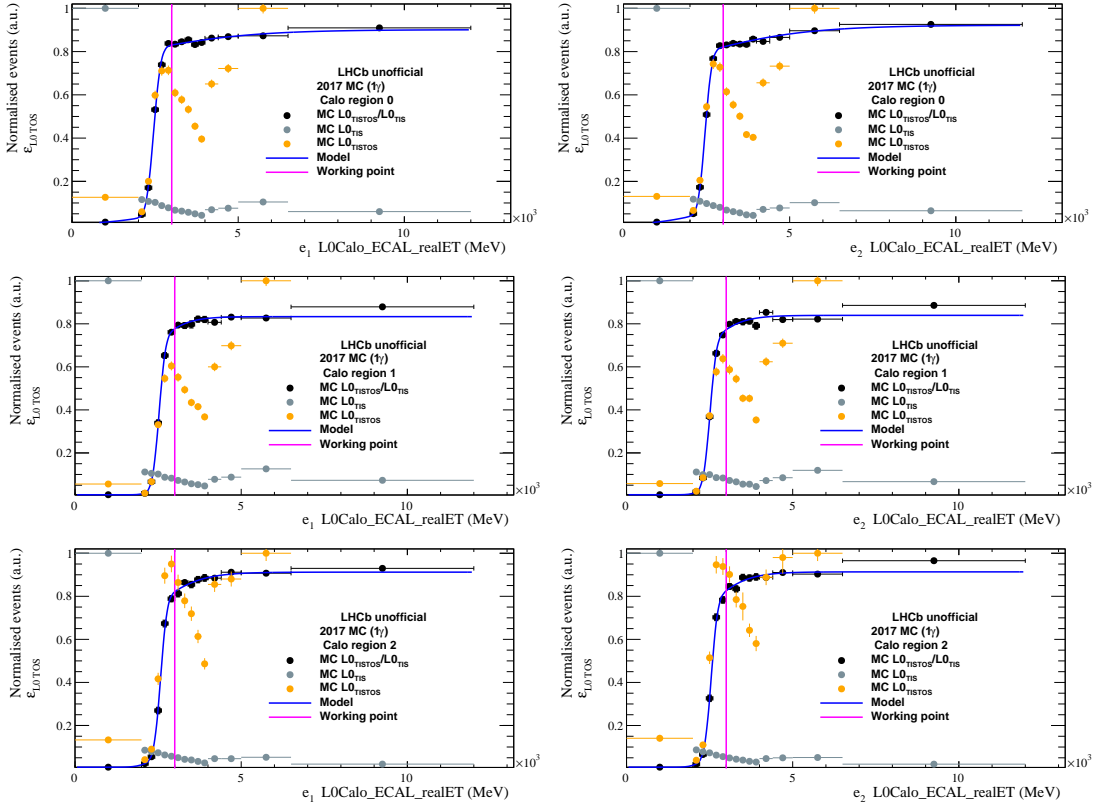


Figure A.21 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 1γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ simulation for the data-taking year 2017, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAL_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

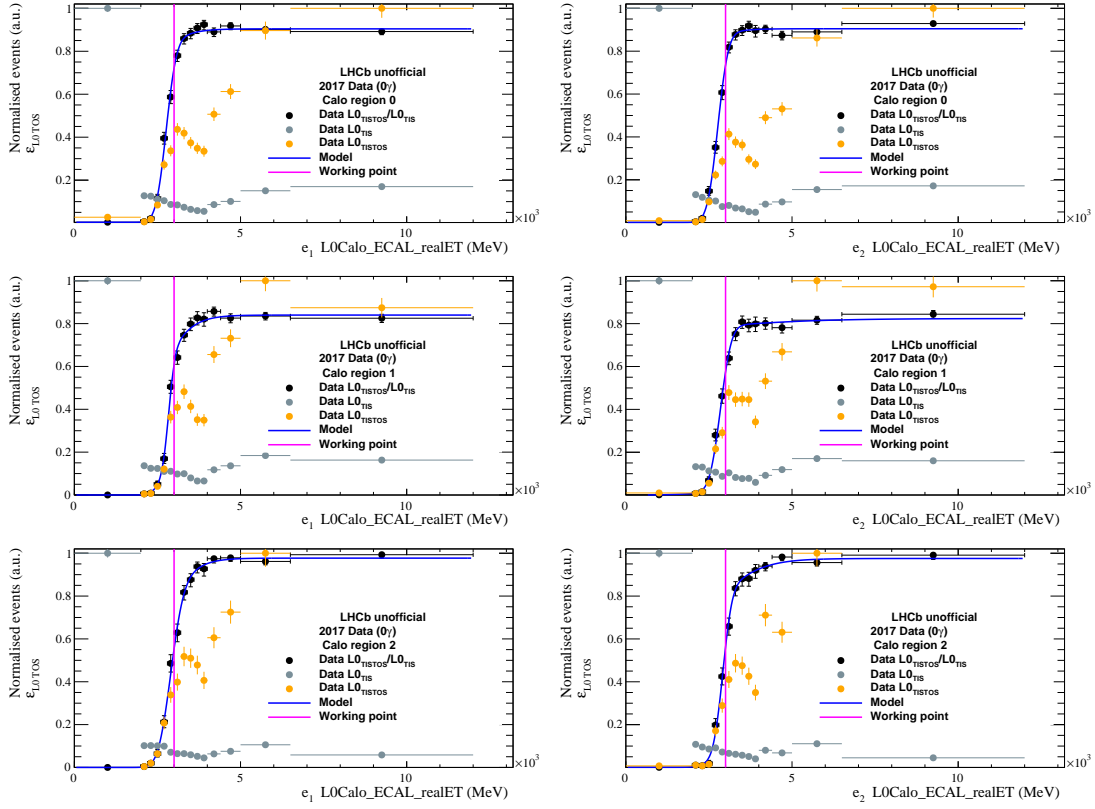


Figure A.22 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 0γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ data for the data-taking year 2017, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

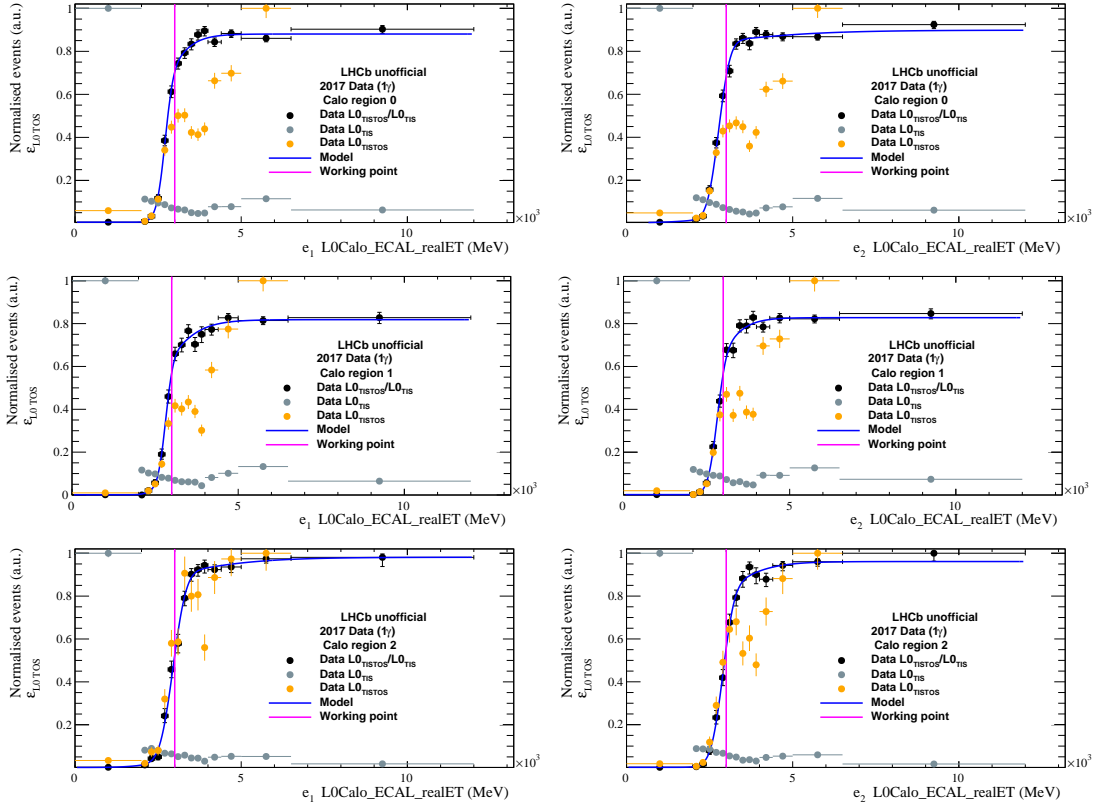


Figure A.23 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 1γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ data for the data-taking year 2017, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

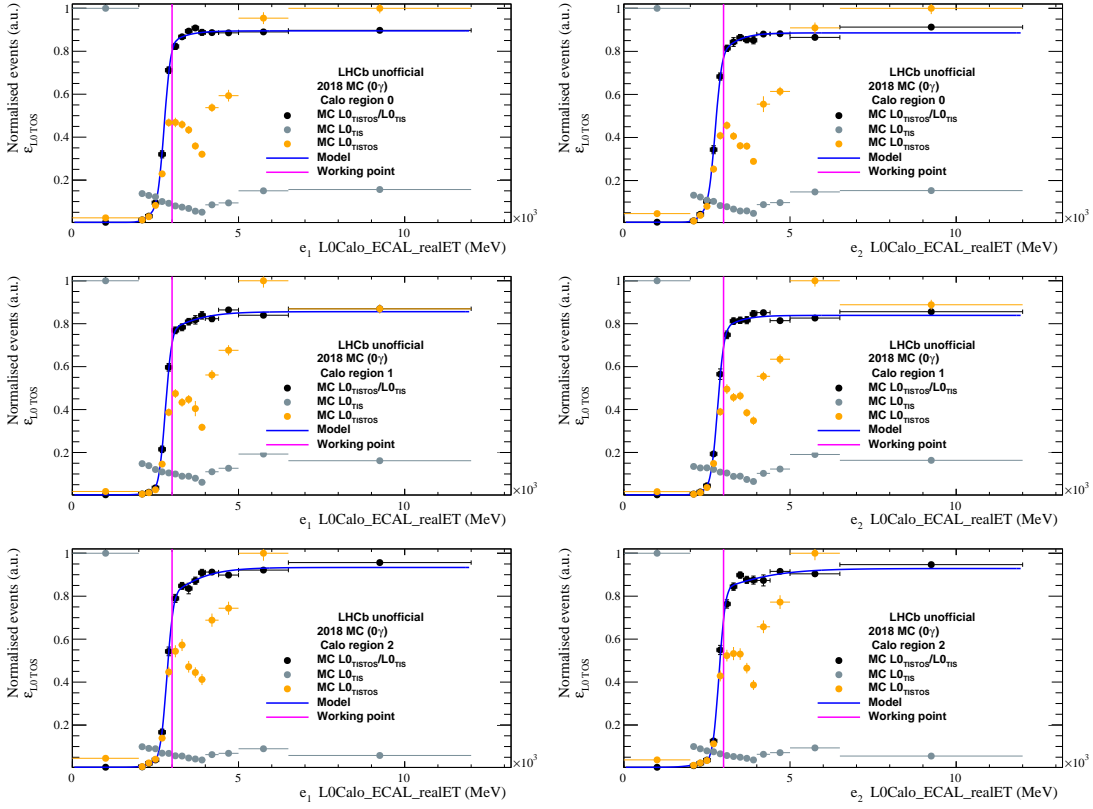


Figure A.24 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 0γ from $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ simulation for the data-taking year 2018, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

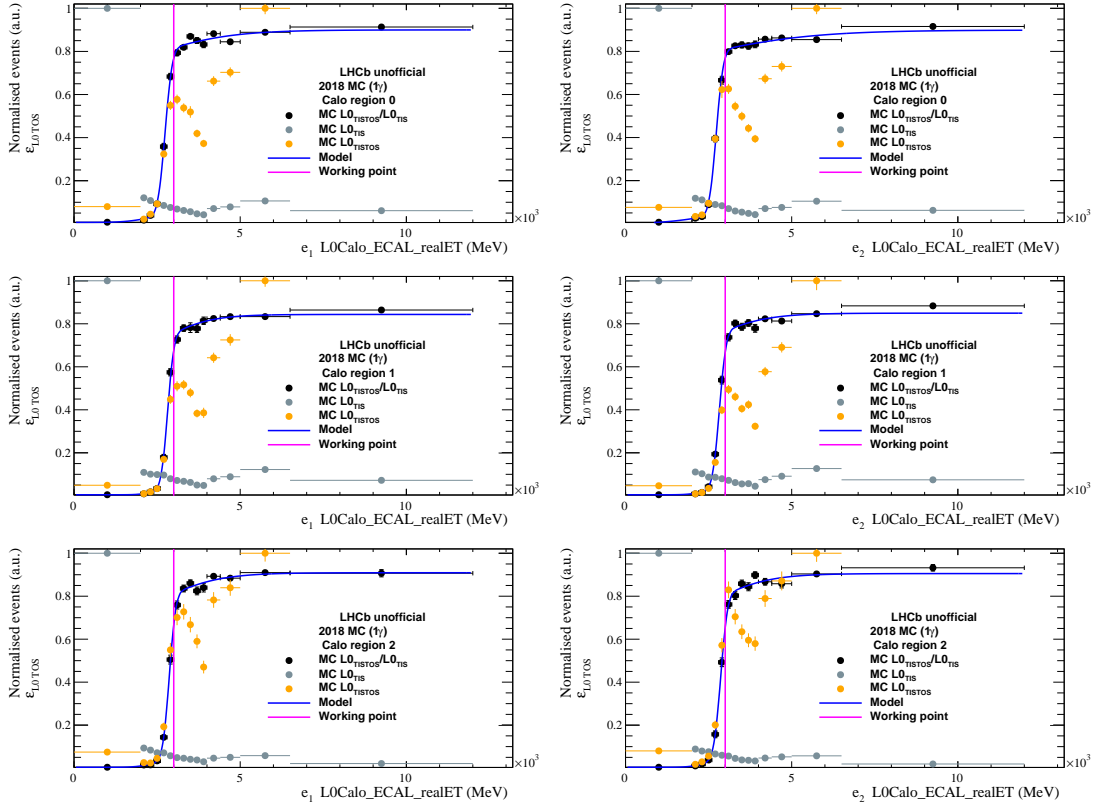


Figure A.25 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 1γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ simulation for the data-taking year 2018, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAL_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

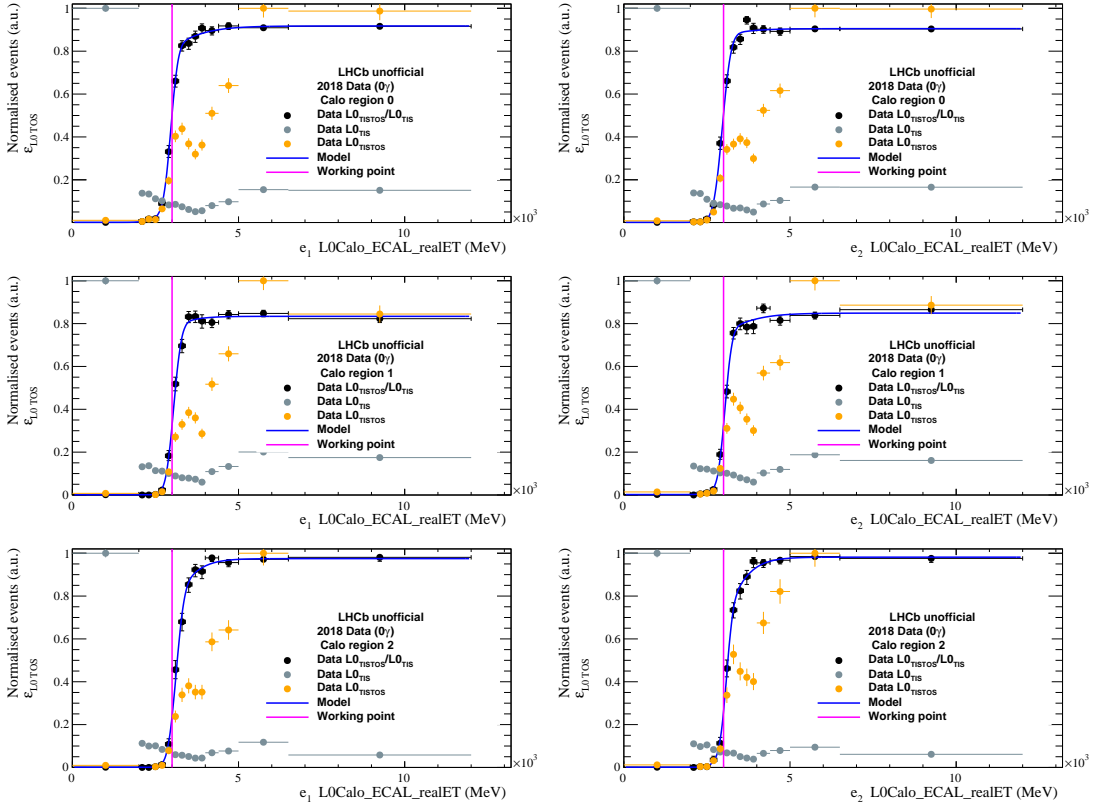


Figure A.26 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 0γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ data for the data-taking year 2018, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

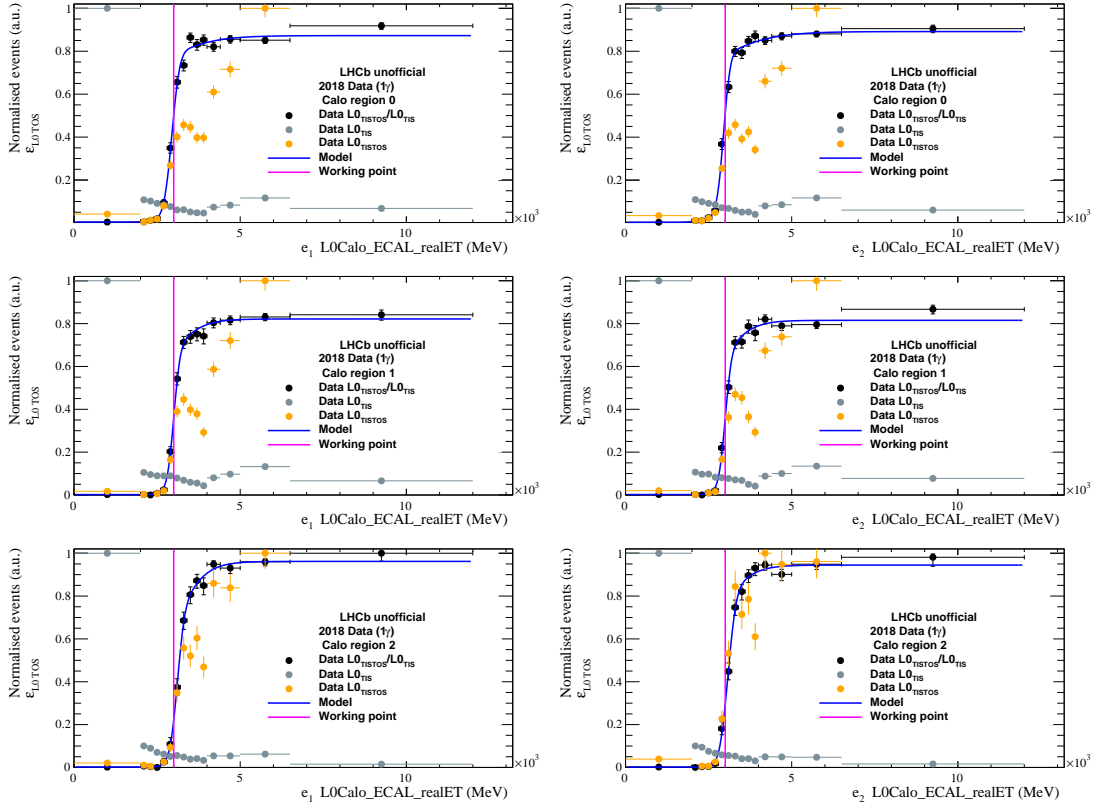


Figure A.27 – Normalised distribution of the transverse energy deposited in the ECAL by TIS (grey points) and TISTOS (orange points) electrons in bremsstrahlung category 1γ from $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ data for the data-taking year 2018, and corresponding L0 trigger efficiency (black points) fitted with an error function (blue curve). The analysis working point, $L0Calo_ECAl_realET > 3000$ MeV, is marked (pink line). The rows correspond to the three ECAL regions and the columns to two electrons e^+ and e^- .

A.3.5 Kinematic corrections

The distributions of the B^+ momentum, transverse momentum, pseudorapidity and nTracks in background subtracted data, uncorrected simulation and corrected simulation are shown in Figs. A.28 and A.29 for $B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+$ for the data-taking years 2017 and 2018. The distributions for $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ are shown in Figs. A.30–A.33 for the data-taking years 2017 and 2018 and the bremsstrahlung categories 0γ and 1γ .

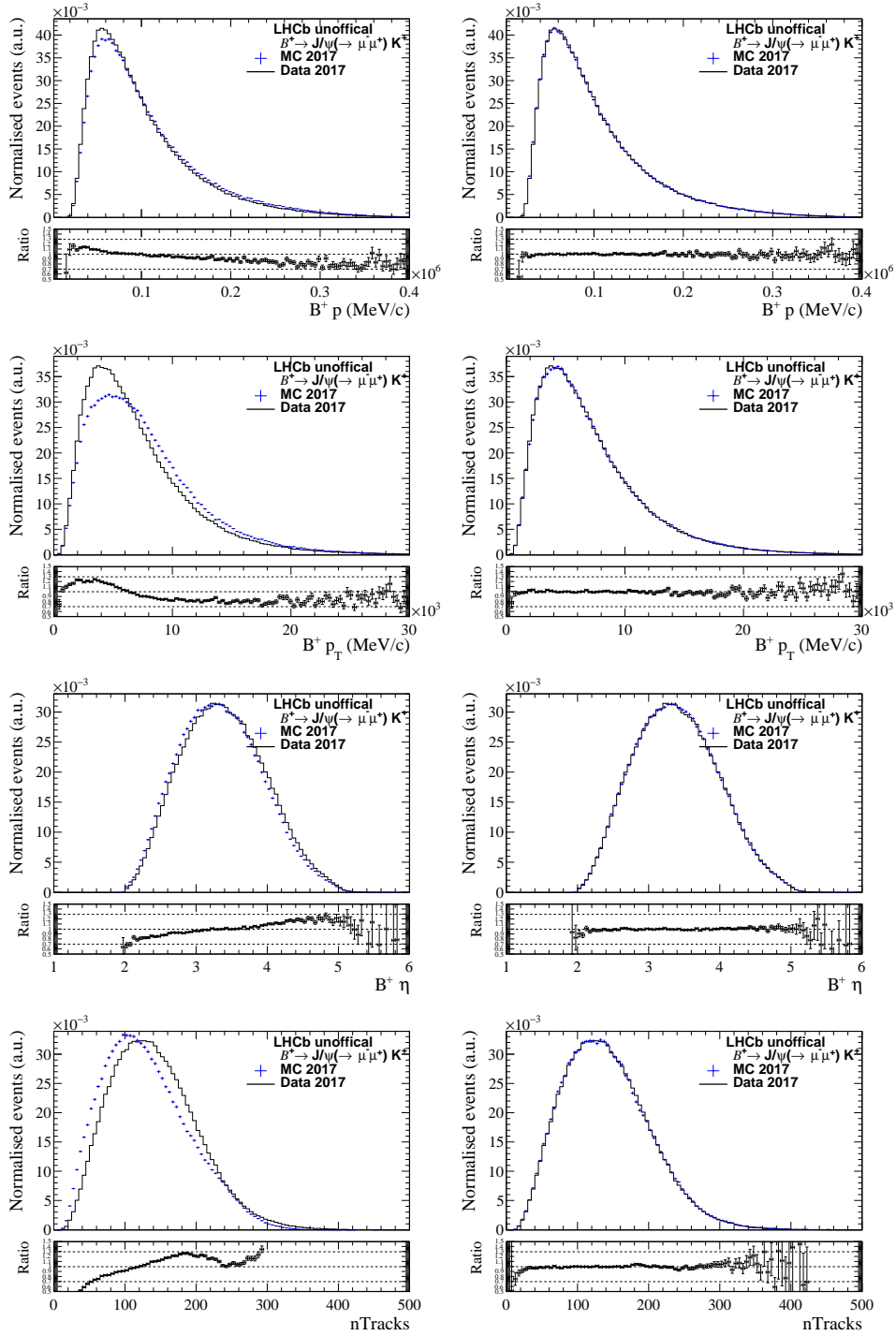


Figure A.28 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ background-subtracted data (histogram) and simulation (points) for the data-taking year 2017. Left: before any correction. Right: after all corrections.

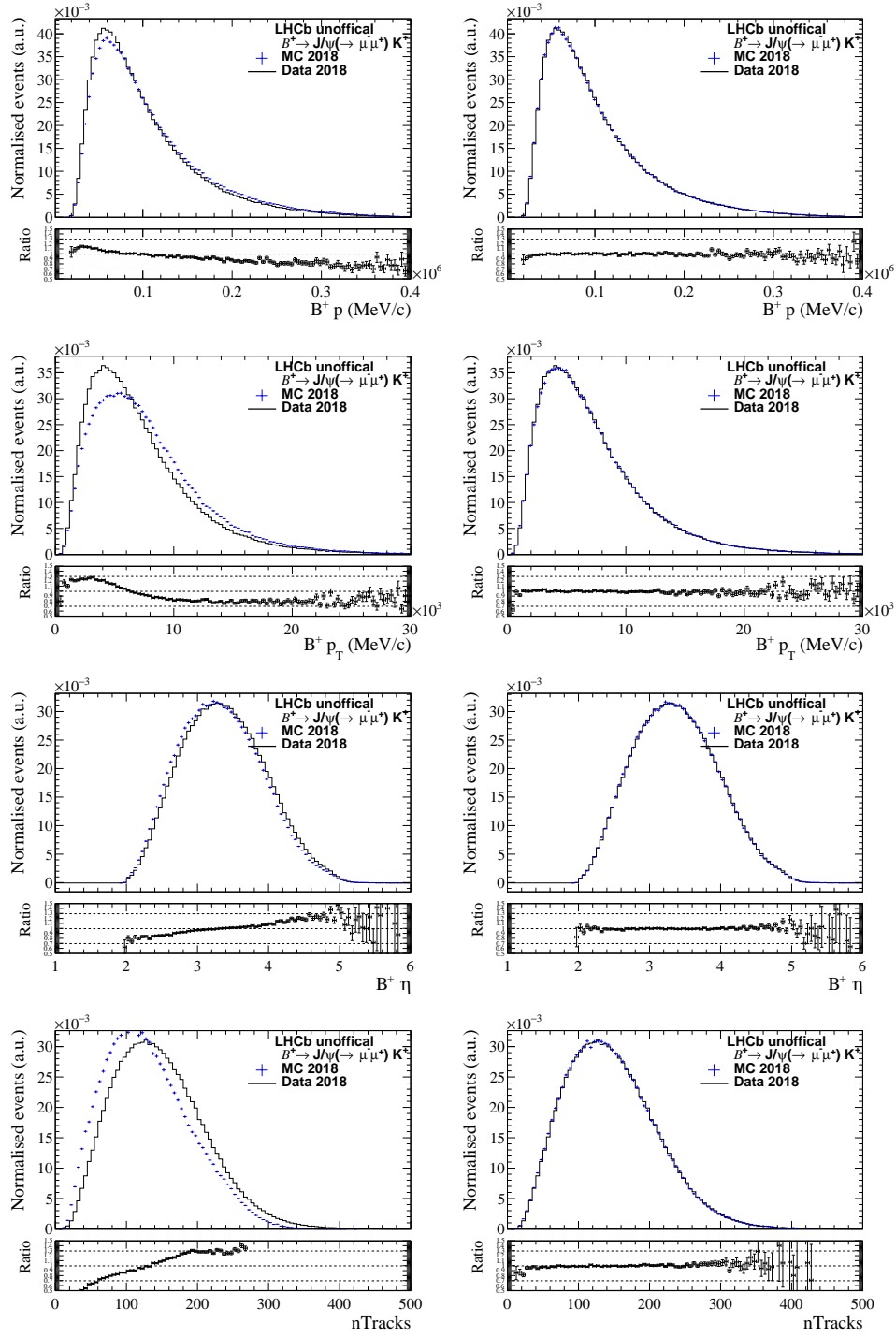


Figure A.29 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ background-subtracted data (histogram) and simulation (points) for the data-taking year 2018. Left: before any correction. Right: after all corrections.

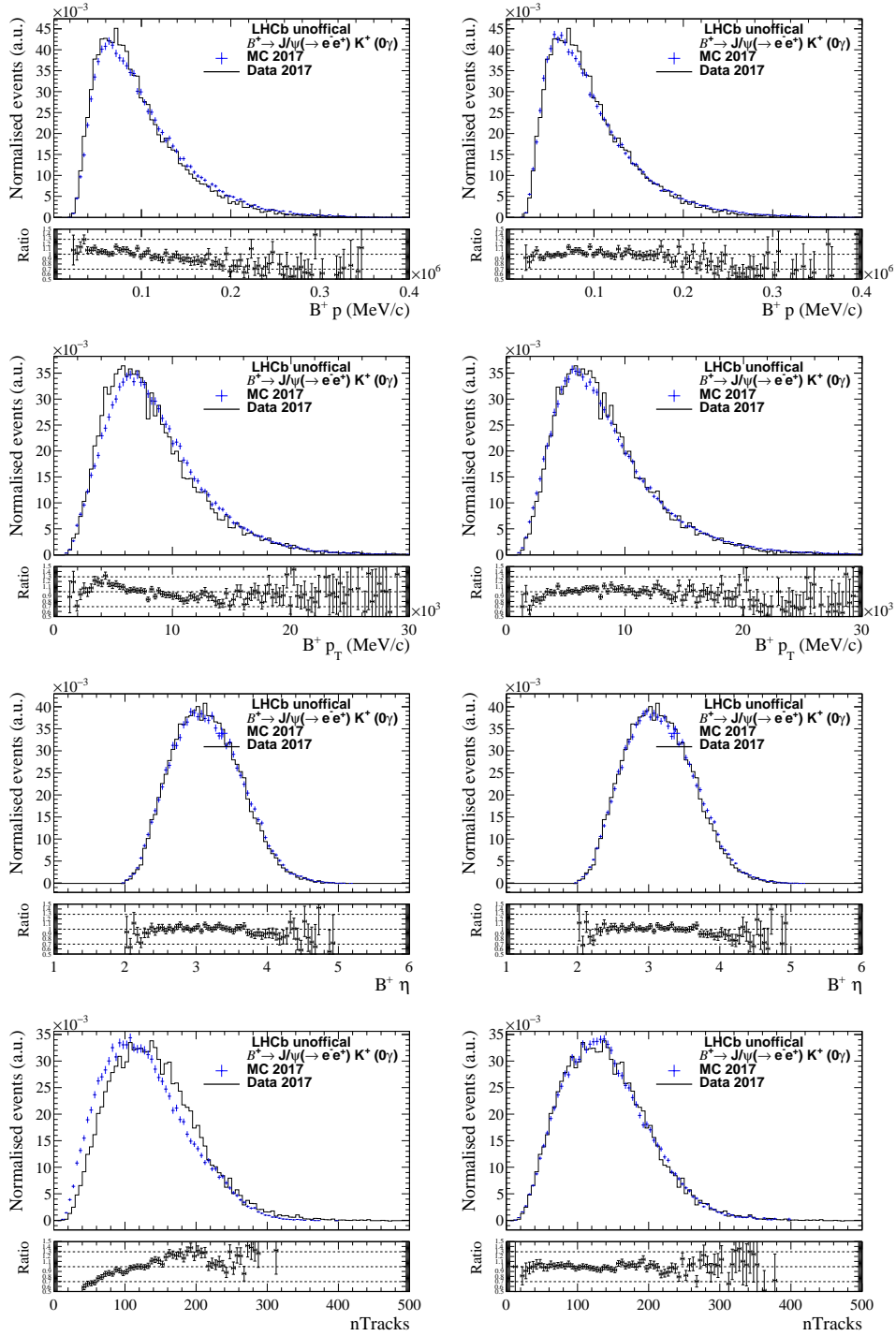


Figure A.30 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ background-subtracted data (histogram) and simulation (points) for the data-taking year 2017 and bremsstrahlung category 0γ . Left: before any correction. Right: after all corrections.

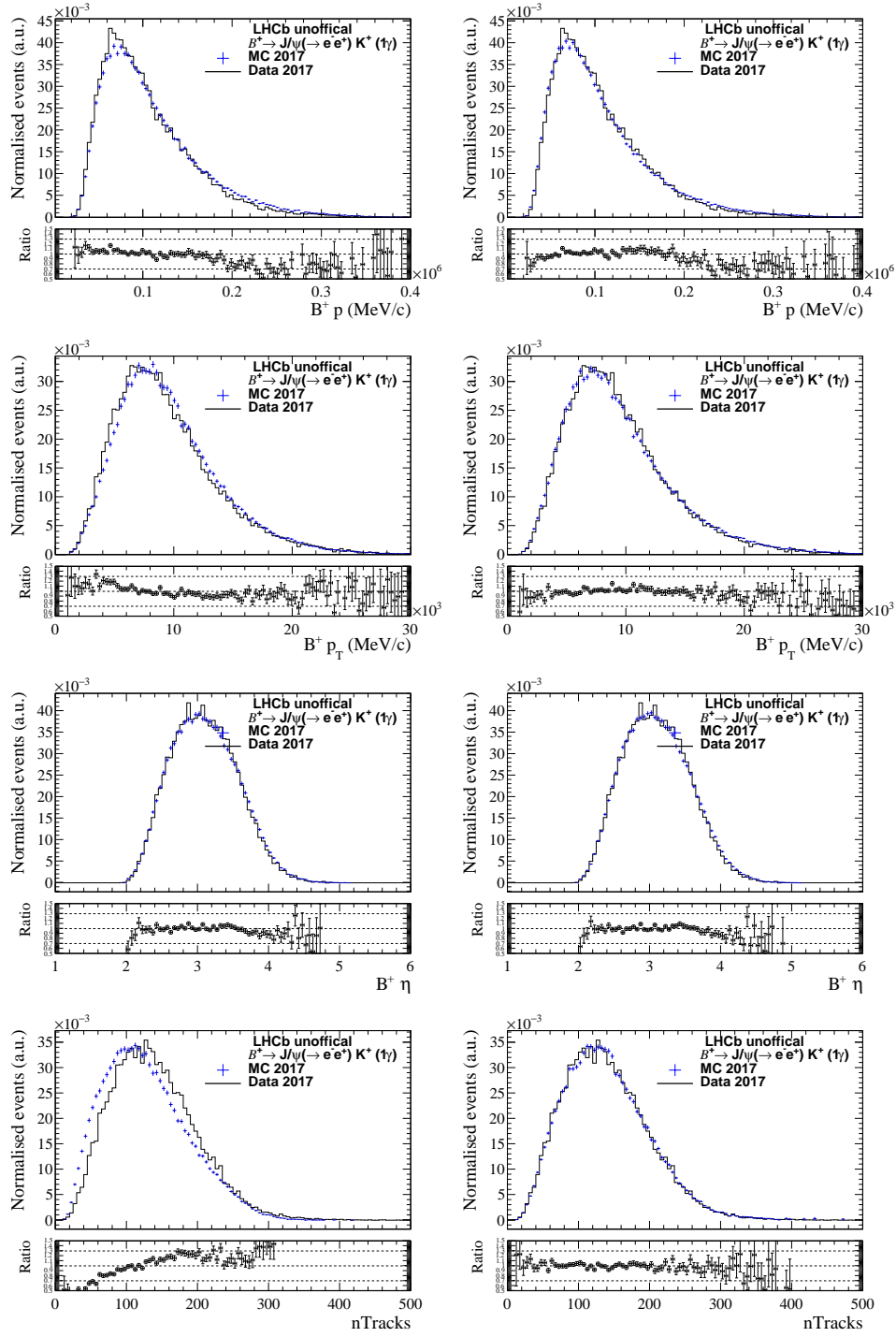


Figure A.31 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ background-subtracted data (histogram) and simulation (points) for the data-taking year 2017 and bremsstrahlung category 1γ . Left: before any correction. Right: after all corrections.

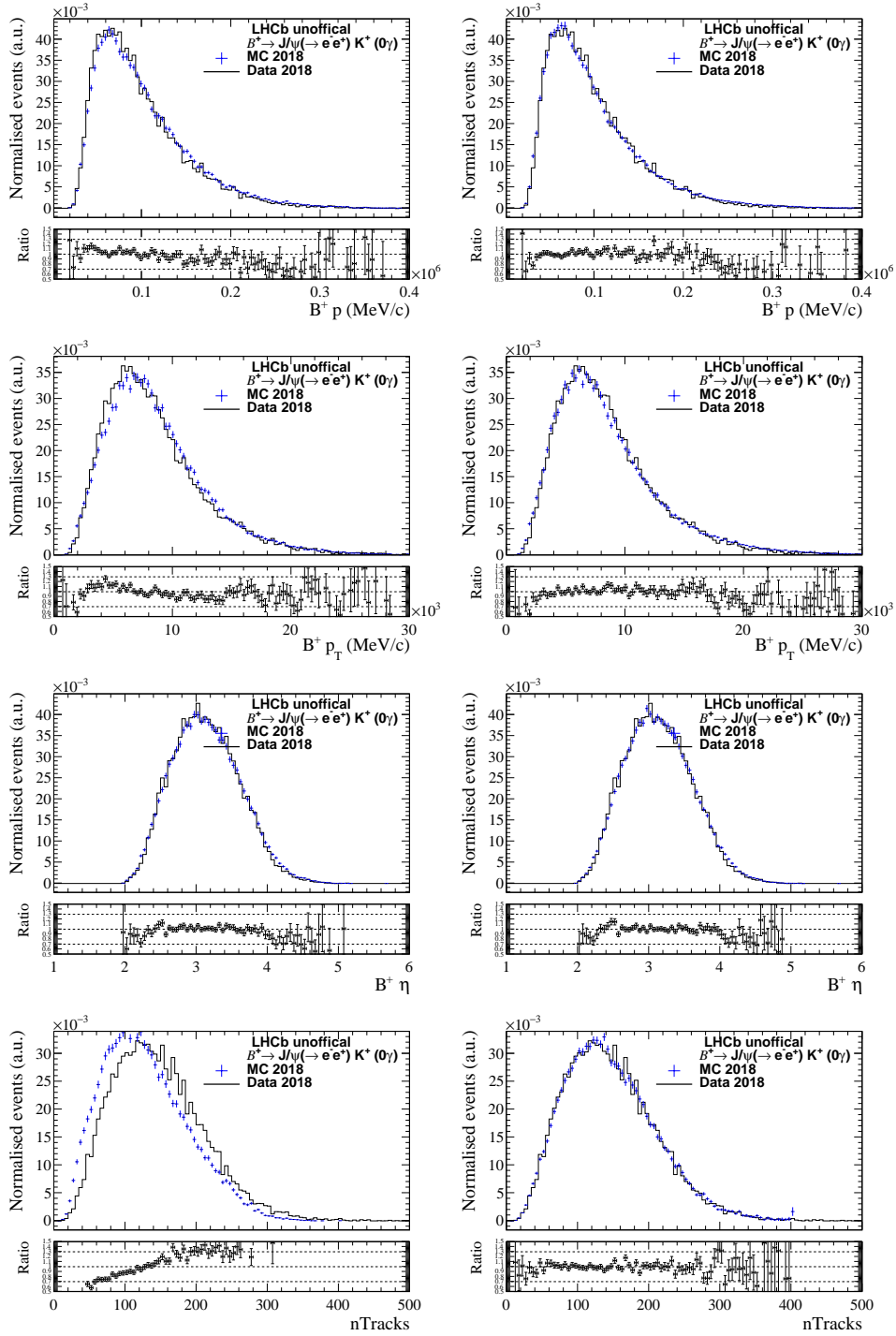


Figure A.32 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow e^+e^-)K^+$ background-subtracted data (histogram) and simulation (points) for the data-taking year 2018 and bremsstrahlung category 0γ . Left: before any correction. Right: after all corrections.

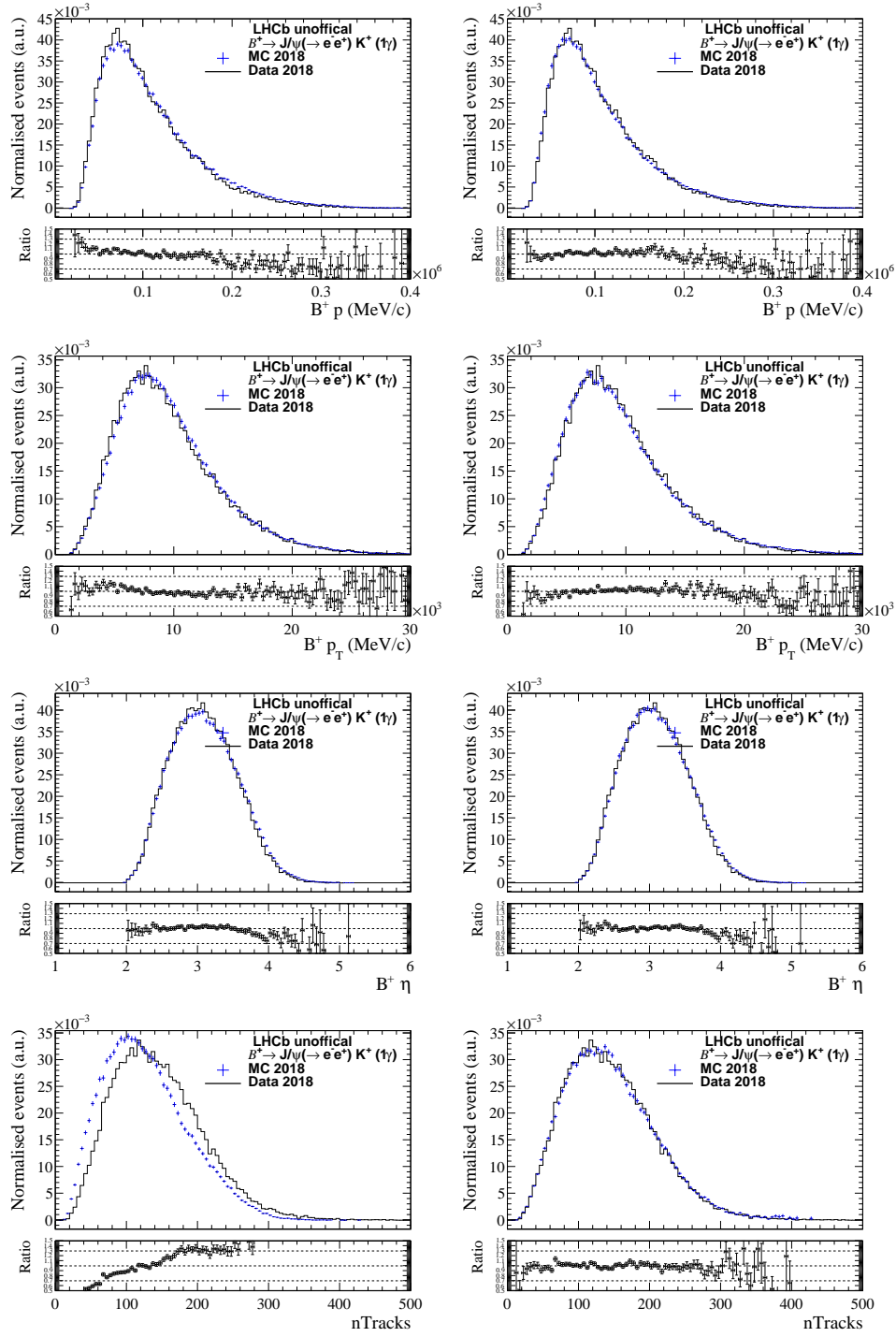


Figure A.33 – Distributions of the B^+ momentum, transverse momentum, pseudorapidity and track multiplicity in $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ background-subtracted data (histogram) and simulation (points) for the data-taking year 2018 and bremsstrahlung category 1γ . Left: before any correction. Right: after all corrections.

Appendix A. Appendix

Table A.2 – $B^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) K^+$ and $B^+ \rightarrow (J/\psi \rightarrow e^+ e^-) K^+$ selection efficiencies in simulation (in %) for different bremsstrahlung categories, data-taking years and corrections, applied in sequence, starting from the uncorrected case. Only the statistical uncertainties are presented.

| Year | Correction | $\mu^+ \mu^-$ | $e^+ e^-, 0\gamma$ | $e^+ e^-, 1\gamma$ | $e^+ e^-, 2\gamma$ |
|------|---|-------------------|---------------------|---------------------|---------------------|
| 2016 | uncorrected | 1.684 ± 0.005 | 0.0903 ± 0.0006 | 0.1878 ± 0.0009 | 0.1011 ± 0.0006 |
| 2016 | ω_{TRK} | 1.657 ± 0.005 | 0.0880 ± 0.0006 | 0.1835 ± 0.0009 | 0.0990 ± 0.0006 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}_{\text{nokaon}}}$ | 1.647 ± 0.005 | 0.0822 ± 0.0006 | 0.1727 ± 0.0009 | 0.0937 ± 0.0006 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}$ | 1.627 ± 0.005 | 0.0814 ± 0.0006 | 0.1710 ± 0.0009 | 0.0926 ± 0.0006 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}}$ | 1.618 ± 0.005 | 0.0763 ± 0.0006 | 0.1541 ± 0.0008 | 0.0800 ± 0.0006 |
| 2016 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}} \cdot \omega_{\text{L0}}$ | 1.691 ± 0.005 | 0.0751 ± 0.0006 | 0.1513 ± 0.0008 | 0.0783 ± 0.0006 |
| 2017 | uncorrected | 1.787 ± 0.004 | 0.0905 ± 0.0005 | 0.1875 ± 0.0007 | 0.1009 ± 0.0005 |
| 2017 | ω_{TRK} | 1.771 ± 0.004 | 0.0897 ± 0.0004 | 0.1862 ± 0.0007 | 0.1003 ± 0.0005 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}_{\text{nokaon}}}$ | 1.759 ± 0.004 | 0.0805 ± 0.0004 | 0.1686 ± 0.0007 | 0.0917 ± 0.0005 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}$ | 1.744 ± 0.004 | 0.0800 ± 0.0004 | 0.1675 ± 0.0007 | 0.0911 ± 0.0005 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}}$ | 1.765 ± 0.004 | 0.0760 ± 0.0004 | 0.1533 ± 0.0006 | 0.0804 ± 0.0004 |
| 2017 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}} \cdot \omega_{\text{L0}}$ | 1.761 ± 0.004 | 0.0740 ± 0.0004 | 0.1487 ± 0.0006 | 0.0778 ± 0.0004 |
| 2018 | uncorrected | 1.704 ± 0.004 | 0.0900 ± 0.0004 | 0.1863 ± 0.0007 | 0.0996 ± 0.0005 |
| 2018 | ω_{TRK} | 1.685 ± 0.004 | 0.0889 ± 0.0004 | 0.1844 ± 0.0007 | 0.0988 ± 0.0005 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}_{\text{nokaon}}}$ | 1.674 ± 0.004 | 0.0774 ± 0.0004 | 0.1631 ± 0.0007 | 0.0886 ± 0.0004 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}}$ | 1.662 ± 0.004 | 0.0770 ± 0.0004 | 0.1621 ± 0.0007 | 0.0880 ± 0.0004 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}}$ | 1.653 ± 0.004 | 0.0722 ± 0.0004 | 0.1457 ± 0.0006 | 0.0759 ± 0.0004 |
| 2018 | $\omega_{\text{TRK}} \cdot \epsilon_{\text{PID}} \cdot \omega_{\text{BKIN}} \cdot \omega_{\text{L0}}$ | 1.655 ± 0.004 | 0.0698 ± 0.0004 | 0.1401 ± 0.0006 | 0.0724 ± 0.0004 |

A.3.6 $r_{J/\psi}$ cross-check

In Table A.2, the $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ selection efficiencies used for the computation of $r_{J/\psi}$ (see Sec. 4.5.7) are presented. Only the statistical uncertainties are presented. The efficiencies are determined from corrected $B^+ \rightarrow (J/\psi \rightarrow \ell^+ \ell^-) K^+$ simulation, taking into account the tracking, L0 trigger and B kinematic corrections, as well as the data-driven PID efficiencies, introduced in Sec. 4.5. The selection efficiencies are obtained in a sequence for each correction step, as well as without any correction.

A.4 Physics background

A.4.1 Background expectations from corrected simulation

The expected physics backgrounds (see Sec. 4.6.1), are estimated for three different mass ranges: the full mass range used for the final fit, $m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.9, 61] \text{ GeV}/c^2$, the blinded signal mass range, $m_{\text{DTF}}(e^\pm \mu^\mp) \in [5.1, 5.6] \text{ GeV}/c^2$, and the sideband mass range, $m_{\text{DTF}}(e^\pm \mu^\mp) \in [4.9, 5.1] \cup [5.6, 6.1] \text{ GeV}/c^2$. To perform this estimation, the selection efficiencies for all physics backgrounds are determined from simulation for all three mass ranges, in the six samples

Table A.3 – Selection efficiency for several background decays, determined for each of the final six fit samples, merged for all data-taking years in the full mass range, $m_{\text{DTF}}(e^\pm\mu^\mp) \in [4.9, 6.1]\text{GeV}/c^2$.

| Decay | Brem | $\varepsilon(\text{BDT region 0})$ | $\varepsilon(\text{BDT region 1})$ | $\varepsilon(\text{BDT region 2})$ |
|---|-----------|------------------------------------|------------------------------------|------------------------------------|
| $B^0 \rightarrow K^+\pi^-$ | 0γ | $(1.0 \pm 0.4) \cdot 10^{-7}$ | $(1.2 \pm 0.4) \cdot 10^{-7}$ | $(1.5 \pm 0.5) \cdot 10^{-7}$ |
| $B_s^0 \rightarrow \pi^+K^-$ | 0γ | $(1.1 \pm 0.4) \cdot 10^{-7}$ | $(1.3 \pm 0.4) \cdot 10^{-7}$ | $(1.5 \pm 0.4) \cdot 10^{-7}$ |
| $B_s^0 \rightarrow K^+K^-$ | 0γ | $(2.3 \pm 1.2) \cdot 10^{-8}$ | $(2.8 \pm 1.5) \cdot 10^{-8}$ | $(3.2 \pm 1.7) \cdot 10^{-8}$ |
| $B^0 \rightarrow p\bar{p}$ | 0γ | $(2.2 \pm 1.5) \cdot 10^{-10}$ | $(1.7 \pm 1.3) \cdot 10^{-10}$ | $(4.7 \pm 2.9) \cdot 10^{-10}$ |
| $B^0 \rightarrow \pi^+\pi^-$ | 0γ | $(2.8 \pm 0.7) \cdot 10^{-7}$ | $(2.9 \pm 0.7) \cdot 10^{-7}$ | $(3.3 \pm 0.7) \cdot 10^{-7}$ |
| $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ | 0γ | $(2.4 \pm 1.4) \cdot 10^{-8}$ | $(1.6 \pm 0.9) \cdot 10^{-8}$ | $(5.2 \pm 3.5) \cdot 10^{-9}$ |
| $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ | 0γ | $(1.75 \pm 0.09) \cdot 10^{-7}$ | $(1.35 \pm 0.08) \cdot 10^{-7}$ | $(6.5 \pm 0.6) \cdot 10^{-8}$ |
| $\Lambda_b^0 \rightarrow pK^-$ | 0γ | $(1.2 \pm 0.7) \cdot 10^{-8}$ | $(1.3 \pm 0.8) \cdot 10^{-8}$ | $(1.7 \pm 1.0) \cdot 10^{-8}$ |
| $\Lambda_b^0 \rightarrow p\pi^-$ | 0γ | $(1.7 \pm 1.0) \cdot 10^{-8}$ | $(1.9 \pm 1.1) \cdot 10^{-8}$ | $(2.2 \pm 1.3) \cdot 10^{-8}$ |
| $\Lambda_b^0 \rightarrow pe^-\bar{\nu}_e$ | 0γ | $(4 \pm 4) \cdot 10^{-12}$ | – | – |
| $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ | 0γ | $(2.7 \pm 0.5) \cdot 10^{-8}$ | $(1.8 \pm 0.4) \cdot 10^{-8}$ | $(7.9 \pm 2.8) \cdot 10^{-9}$ |
| $B^0 \rightarrow K^+\pi^-$ | 1γ | – | – | – |
| $B_s^0 \rightarrow \pi^+K^-$ | 1γ | – | – | – |
| $B_s^0 \rightarrow K^+K^-$ | 1γ | – | – | – |
| $B^0 \rightarrow p\bar{p}$ | 1γ | – | – | – |
| $B^0 \rightarrow \pi^+\pi^-$ | 1γ | – | – | – |
| $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ | 1γ | $(5.1 \pm 3.1) \cdot 10^{-8}$ | $(3.5 \pm 2.0) \cdot 10^{-8}$ | $(1.6 \pm 0.9) \cdot 10^{-8}$ |
| $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow pK^-$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow p\pi^-$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow pe^-\bar{\nu}_e$ | 1γ | $(3.5 \pm 2.6) \cdot 10^{-9}$ | – | – |
| $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ | 1γ | – | – | – |

used in the final fit (see Sec. 4.1) and corresponding to the three BDT regions and the two bremsstrahlung categories 0γ and 1γ . Furthermore, the efficiencies are determined merged over all data-taking years, while weighting the simulated samples with the corresponding integrated luminosity of each data-taking year. The obtained selection efficiencies are presented in Table A.3 for the full signal region, in Table A.4 for the blinded signal region and in Table A.5 for the sideband region. The expected physics backgrounds are shown in Table A.6 for the full signal region, in Table A.7 for the blinded signal region and in Table A.8 for the sideband region.

Appendix A. Appendix

Table A.4 – Selection efficiency for several background decays, determined for each of the final six fit samples, merged for all data-taking years, in the signal mass window, $m_{\text{DTF}}(e^\pm\mu^\mp) \in [5.1, 5.6]\text{GeV}/c^2$.

| Decay | Brem | $\varepsilon(\text{BDT region 0})$ | $\varepsilon(\text{BDT region 1})$ | $\varepsilon(\text{BDT region 2})$ |
|---|-----------|------------------------------------|------------------------------------|------------------------------------|
| $B^0 \rightarrow K^+\pi^-$ | 0γ | $(9.0\pm 3.5)\cdot 10^{-8}$ | $(1.1\pm 0.4)\cdot 10^{-7}$ | $(1.4\pm 0.5)\cdot 10^{-7}$ |
| $B_s^0 \rightarrow \pi^+K^-$ | 0γ | $(1.03\pm 0.34)\cdot 10^{-7}$ | $(1.2\pm 0.4)\cdot 10^{-7}$ | $(1.5\pm 0.4)\cdot 10^{-7}$ |
| $B_s^0 \rightarrow K^+K^-$ | 0γ | $(2.0\pm 1.1)\cdot 10^{-8}$ | $(2.6\pm 1.4)\cdot 10^{-8}$ | $(3.1\pm 1.6)\cdot 10^{-8}$ |
| $B^0 \rightarrow p\bar{p}$ | 0γ | – | – | – |
| $B^0 \rightarrow \pi^+\pi^-$ | 0γ | $(2.4\pm 0.6)\cdot 10^{-7}$ | $(2.6\pm 0.6)\cdot 10^{-7}$ | $(3.2\pm 0.7)\cdot 10^{-7}$ |
| $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ | 0γ | $(2.0\pm 1.3)\cdot 10^{-9}$ | $(1.6\pm 1.2)\cdot 10^{-9}$ | $(8\pm 5)\cdot 10^{-10}$ |
| $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ | 0γ | $(2.02\pm 0.32)\cdot 10^{-8}$ | $(2.12\pm 0.33)\cdot 10^{-8}$ | $(1.77\pm 0.30)\cdot 10^{-8}$ |
| $\Lambda_b^0 \rightarrow pK^-$ | 0γ | $(1.0\pm 0.6)\cdot 10^{-8}$ | $(1.2\pm 0.7)\cdot 10^{-8}$ | $(1.6\pm 1.0)\cdot 10^{-8}$ |
| $\Lambda_b^0 \rightarrow p\pi^-$ | 0γ | $(1.6\pm 0.9)\cdot 10^{-8}$ | $(1.8\pm 1.1)\cdot 10^{-8}$ | $(2.2\pm 1.3)\cdot 10^{-8}$ |
| $\Lambda_b^0 \rightarrow pe^-\bar{\nu}_e$ | 0γ | $(4\pm 4)\text{e-}10$ | – | – |
| $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ | 0γ | $(8.5\pm 2.9)\cdot 10^{-9}$ | $(7.6\pm 2.7)\cdot 10^{-9}$ | $(4.3\pm 1.9)\cdot 10^{-9}$ |
| $B^0 \rightarrow K^+\pi^-$ | 1γ | – | – | – |
| $B_s^0 \rightarrow \pi^+K^-$ | 1γ | – | – | – |
| $B_s^0 \rightarrow K^+K^-$ | 1γ | – | – | – |
| $B^0 \rightarrow p\bar{p}$ | 1γ | – | – | – |
| $B^0 \rightarrow \pi^+\pi^-$ | 1γ | – | – | – |
| $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ | 1γ | $(1.0\pm 0.6)\cdot 10^{-8}$ | $(7\pm 4)\cdot 10^{-9}$ | $(4.7\pm 2.9)\cdot 10^{-9}$ |
| $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow pK^-$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow p\pi^-$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow pe^-\bar{\nu}_e$ | 1γ | $(2.3\pm 2.3)\cdot 10^{-9}$ | – | – |
| $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ | 1γ | – | – | – |

Table A.5 – Selection efficiency for several background decays, determined for each of the final six fit samples, merged for all data-taking years, in the mass sideband region, $m_{\text{DTF}}(e^\pm\mu^\mp) \in [4.9, 5.1] \cup [5.6, 6.1] \text{ GeV}/c^2$.

| Decay | Brem | $\varepsilon(\text{BDT region 0})$ | $\varepsilon(\text{BDT region 1})$ | $\varepsilon(\text{BDT region 2})$ |
|---|-----------|------------------------------------|------------------------------------|------------------------------------|
| $B^0 \rightarrow K^+\pi^-$ | 0γ | $(1.3 \pm 0.7) \cdot 10^{-8}$ | $(1.1 \pm 0.6) \cdot 10^{-8}$ | $(4.4 \pm 2.5) \cdot 10^{-9}$ |
| $B_s^0 \rightarrow \pi^+K^-$ | 0γ | $(1.2 \pm 0.7) \cdot 10^{-8}$ | $(9 \pm 5) \cdot 10^{-9}$ | $(2.6 \pm 1.6) \cdot 10^{-9}$ |
| $B_s^0 \rightarrow K^+K^-$ | 0γ | $(2.9 \pm 1.7) \cdot 10^{-9}$ | $(2.9 \pm 1.7) \cdot 10^{-9}$ | $(1.2 \pm 0.7) \cdot 10^{-9}$ |
| $B^0 \rightarrow p\bar{p}$ | 0γ | $(2.2 \pm 1.5) \cdot 10^{-10}$ | $(1.7 \pm 1.3) \cdot 10^{-10}$ | $(4.7 \pm 2.9) \cdot 10^{-10}$ |
| $B^0 \rightarrow \pi^+\pi^-$ | 0γ | $(4.2 \pm 2.1) \cdot 10^{-8}$ | $(2.9 \pm 1.6) \cdot 10^{-8}$ | $(1.0 \pm 0.6) \cdot 10^{-8}$ |
| $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ | 0γ | $(2.2 \pm 1.3) \cdot 10^{-8}$ | $(1.4 \pm 0.9) \cdot 10^{-8}$ | $(4.4 \pm 3.1) \cdot 10^{-9}$ |
| $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ | 0γ | $(1.55 \pm 0.09) \cdot 10^{-7}$ | $(1.14 \pm 0.08) \cdot 10^{-7}$ | $(4.8 \pm 0.5) \cdot 10^{-8}$ |
| $\Lambda_b^0 \rightarrow pK^-$ | 0γ | $(2.2 \pm 1.3) \cdot 10^{-9}$ | $(1.6 \pm 1.0) \cdot 10^{-9}$ | $(1.0 \pm 0.6) \cdot 10^{-9}$ |
| $\Lambda_b^0 \rightarrow p\pi^-$ | 0γ | $(1.4 \pm 0.9) \cdot 10^{-9}$ | $(9 \pm 5) \cdot 10^{-10}$ | $(4.0 \pm 2.7) \cdot 10^{-10}$ |
| $\Lambda_b^0 \rightarrow pe^-\bar{\nu}_e$ | 0γ | – | – | – |
| $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ | 0γ | $(1.8 \pm 0.4) \cdot 10^{-8}$ | $(1.06 \pm 0.33) \cdot 10^{-8}$ | $(3.5 \pm 1.7) \cdot 10^{-9}$ |
| $B^0 \rightarrow K^+\pi^-$ | 1γ | – | – | – |
| $B_s^0 \rightarrow \pi^+K^-$ | 1γ | – | – | – |
| $B_s^0 \rightarrow K^+K^-$ | 1γ | – | – | – |
| $B^0 \rightarrow p\bar{p}$ | 1γ | – | – | – |
| $B^0 \rightarrow \pi^+\pi^-$ | 1γ | – | – | – |
| $B^0 \rightarrow \pi^+e^-\bar{\nu}_e$ | 1γ | $(4.2 \pm 2.5) \cdot 10^{-8}$ | $(2.8 \pm 1.6) \cdot 10^{-8}$ | $(1.1 \pm 0.7) \cdot 10^{-8}$ |
| $B^0 \rightarrow \pi^+\mu^-\bar{\nu}_\mu$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow pK^-$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow p\pi^-$ | 1γ | – | – | – |
| $\Lambda_b^0 \rightarrow pe^-\bar{\nu}_e$ | 1γ | $(1.2 \pm 1.2) \cdot 10^{-9}$ | – | – |
| $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ | 1γ | – | – | – |

Table A.6 – Estimated physics backgrounds in the full mass range.

| Decay | BDTregion0 | | BDTregion1 | | BDTregion2 | | Total |
|---|---------------------------------|---------------|---------------------------------|------------|---------------------------------|---------------|---------------------------------|
| | 0 γ | 1 γ | 0 γ | 1 γ | 0 γ | 1 γ | |
| $B^0 \rightarrow K^+ \pi^-$ | 3.4 \pm 1.3 | – | 4.0 \pm 1.4 | – | 4.9 \pm 1.6 | – | 12.2 \pm 2.4 |
| $B_s^0 \rightarrow \pi^+ K^-$ | 0.28 \pm 0.10 | – | 0.31 \pm 0.10 | – | 0.36 \pm 0.11 | – | 0.96 \pm 0.18 |
| $B_s^0 \rightarrow K^+ K^-$ | 0.26 \pm 0.14 | – | 0.32 \pm 0.17 | – | 0.36 \pm 0.19 | – | 0.93 \pm 0.30 |
| $B^0 \rightarrow K^+ K^-$ | 0.0030 \pm 0.0017 | – | 0.0037 \pm 0.0021 | – | 0.0041 \pm 0.0023 | – | 0.011 \pm 0.004 |
| $B^0 \rightarrow p \bar{p}$ | (4.6 \pm 3.4) $\cdot 10^{-6}$ | – | (3.6 \pm 2.8) $\cdot 10^{-6}$ | – | (1.0 \pm 0.7) $\cdot 10^{-5}$ | – | (1.8 \pm 0.8) $\cdot 10^{-5}$ |
| $B_s^0 \rightarrow p \bar{p}$ | (1.4 \pm 1.7) $\cdot 10^{-6}$ | – | (1.1 \pm 1.4) $\cdot 10^{-6}$ | – | (3.0 \pm 3.5) $\cdot 10^{-6}$ | – | (6 \pm 4) $\cdot 10^{-6}$ |
| $B^0 \rightarrow \pi^+ \pi^-$ | 2.4 \pm 0.6 | – | 2.5 \pm 0.6 | – | 2.8 \pm 0.6 | – | 7.7 \pm 1.0 |
| $B_s^0 \rightarrow \pi^+ \pi^-$ | 0.084 \pm 0.023 | – | 0.086 \pm 0.024 | – | 0.098 \pm 0.026 | – | 0.27 \pm 0.04 |
| $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ | 6.0 \pm 3.5 | 13 \pm 8 | 4.0 \pm 2.4 | 9 \pm 5 | 1.3 \pm 0.9 | 4.0 \pm 2.3 | 37 \pm 10 |
| $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ | 43.7 \pm 3.1 | – | 33.8 \pm 2.6 | – | 16.4 \pm 1.6 | – | 94 \pm 4 |
| $\Lambda_b^0 \rightarrow p K^-$ | 0.056 \pm 0.034 | – | 0.06 \pm 0.04 | – | 0.08 \pm 0.05 | – | 0.20 \pm 0.07 |
| $\Lambda_b^0 \rightarrow p \pi^-$ | 0.07 \pm 0.04 | – | 0.07 \pm 0.05 | – | 0.09 \pm 0.05 | – | 0.23 \pm 0.08 |
| $\Lambda_b^0 \rightarrow p e^- \bar{\nu}_e$ | 0.0015 \pm 0.0015 | 1.2 \pm 1.0 | – | – | – | – | 1.2 \pm 1.0 |
| $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ | 9.5 \pm 3.1 | – | 6.4 \pm 2.3 | – | 2.8 \pm 1.2 | – | 19 \pm 4 |

Table A.7 – Estimated physics backgrounds in the blinded signal region.

| Decay | BDTregion0 | | BDTregion1 | | BDTregion2 | | Total |
|---|---------------------|---------------|---------------------|---------------|---------------------|---------------|---------------------|
| | 0 γ | 1 γ | 0 γ | 1 γ | 0 γ | 1 γ | |
| $B^0 \rightarrow K^+ \pi^-$ | 3.0 \pm 1.2 | – | 3.7 \pm 1.3 | – | 4.7 \pm 1.5 | – | 11.3 \pm 2.3 |
| $B_s^0 \rightarrow \pi^+ K^-$ | 0.25 \pm 0.09 | – | 0.29 \pm 0.10 | – | 0.36 \pm 0.11 | – | 0.90 \pm 0.17 |
| $B_s^0 \rightarrow K^+ K^-$ | 0.22 \pm 0.13 | – | 0.29 \pm 0.16 | – | 0.34 \pm 0.19 | – | 0.86 \pm 0.27 |
| $B^0 \rightarrow K^+ K^-$ | 0.0026 \pm 0.0015 | – | 0.0033 \pm 0.0019 | – | 0.0040 \pm 0.0023 | – | 0.0099 \pm 0.0033 |
| $B^0 \rightarrow p \bar{p}$ | – | – | – | – | – | – | – |
| $B_s^0 \rightarrow p \bar{p}$ | – | – | – | – | – | – | – |
| $B_s^0 \rightarrow \pi^+ \pi^-$ | 2.1 \pm 0.5 | – | 2.2 \pm 0.6 | – | 2.7 \pm 0.6 | – | 7.0 \pm 1.0 |
| $B_s^0 \rightarrow \pi^+ \pi^-$ | 0.071 \pm 0.021 | – | 0.077 \pm 0.022 | – | 0.095 \pm 0.025 | – | 0.24 \pm 0.04 |
| $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ | 0.50 \pm 0.32 | 2.4 \pm 1.5 | 0.40 \pm 0.29 | 1.7 \pm 1.0 | 0.20 \pm 0.12 | 1.2 \pm 0.7 | 6.4 \pm 2.0 |
| $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ | 5.0 \pm 0.8 | – | 5.3 \pm 0.8 | – | 4.4 \pm 0.8 | – | 14.8 \pm 1.4 |
| $\Lambda_b^0 \rightarrow p K^-$ | 0.046 \pm 0.028 | – | 0.054 \pm 0.033 | – | 0.08 \pm 0.05 | – | 0.18 \pm 0.06 |
| $\Lambda_b^0 \rightarrow p \pi^-$ | 0.06 \pm 0.04 | – | 0.07 \pm 0.04 | – | 0.09 \pm 0.05 | – | 0.22 \pm 0.08 |
| $\Lambda_b^0 \rightarrow p e^- \bar{\nu}_e$ | 0.0015 \pm 0.0015 | 0.8 \pm 0.8 | – | – | – | – | 0.8 \pm 0.8 |
| $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ | 3.0 \pm 1.3 | – | 2.7 \pm 1.2 | – | 1.5 \pm 0.8 | – | 7.2 \pm 1.9 |

Table A.8 – Estimated physics backgrounds in the sideband region.

| Decay | BDTregion0 0 γ | BDTregion0 1 γ | BDTregion1 0 γ | BDTregion1 1 γ | BDTregion2 0 γ | BDTregion2 1 γ | Total |
|---|---------------------------------|--------------------------|---------------------------------|--------------------------|---------------------------------|--------------------------|---------------------------------|
| $B^0 \rightarrow K^+ \pi^-$ | 0.41 \pm 0.24 | – | 0.35 \pm 0.20 | – | 0.14 \pm 0.08 | – | 0.91 \pm 0.33 |
| $B_s^0 \rightarrow \pi^+ K^-$ | 0.029 \pm 0.017 | – | 0.021 \pm 0.013 | – | 0.006 \pm 0.004 | – | 0.057 \pm 0.021 |
| $B_s^0 \rightarrow K^+ K^-$ | 0.033 \pm 0.019 | – | 0.032 \pm 0.019 | – | 0.014 \pm 0.008 | – | 0.079 \pm 0.028 |
| $B^0 \rightarrow K^+ K^-$ | 0.00038 \pm 0.00023 | – | 0.00037 \pm 0.00023 | – | 0.00016 \pm 0.00010 | – | 0.00091 \pm 0.00034 |
| $B^0 \rightarrow p \bar{p}$ | (4.6 \pm 3.4) $\cdot 10^{-6}$ | – | (3.6 \pm 2.8) $\cdot 10^{-6}$ | – | (1.0 \pm 0.7) $\cdot 10^{-5}$ | – | (1.8 \pm 0.8) $\cdot 10^{-5}$ |
| $B_s^0 \rightarrow p \bar{p}$ | (1.4 \pm 1.7) $\cdot 10^{-6}$ | – | (1.1 \pm 1.4) $\cdot 10^{-6}$ | – | (3.0 \pm 3.5) $\cdot 10^{-6}$ | – | (6 \pm 4) $\cdot 10^{-6}$ |
| $B^0 \rightarrow \pi^+ \pi^-$ | 0.36 \pm 0.18 | – | 0.25 \pm 0.14 | – | 0.09 \pm 0.05 | – | 0.70 \pm 0.23 |
| $B_s^0 \rightarrow \pi^+ \pi^-$ | 0.013 \pm 0.007 | – | 0.009 \pm 0.005 | – | 0.0031 \pm 0.0018 | – | 0.024 \pm 0.008 |
| $B^0 \rightarrow \pi^+ e^- \bar{\nu}_e$ | 5.5 \pm 3.2 | 10 \pm 6 | 3.6 \pm 2.2 | 7 \pm 4 | 1.1 \pm 0.8 | 2.8 \pm 1.6 | 30 \pm 9 |
| $B^0 \rightarrow \pi^+ \mu^- \bar{\nu}_\mu$ | 38.7 \pm 2.8 | – | 28.5 \pm 2.3 | – | 12.0 \pm 1.3 | – | 79 \pm 4 |
| $\Lambda_b^0 \rightarrow p K^-$ | 0.010 \pm 0.006 | – | 0.007 \pm 0.005 | – | 0.0047 \pm 0.0029 | – | 0.023 \pm 0.008 |
| $\Lambda_b^0 \rightarrow p \pi^-$ | 0.0055 \pm 0.0035 | – | 0.0036 \pm 0.0022 | – | 0.0016 \pm 0.0011 | – | 0.011 \pm 0.004 |
| $\Lambda_b^0 \rightarrow p e^- \bar{\nu}_e$ | – | 0.4 \pm 0.4 | – | – | – | – | 0.4 \pm 0.4 |
| $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ | 6.5 \pm 2.3 | – | 3.7 \pm 1.5 | – | 1.3 \pm 0.7 | – | 11.5 \pm 2.8 |

A.4.2 Validation of the expected peaking background with $B_{(s)}^0 \rightarrow hh'$ decays

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

The $B^0 \rightarrow K^+\pi^-$ samples are built in $B_{(s)}^0 \rightarrow hh'$ data and simulation, starting from the stripping output of the StrippingBs2MuMuLinesNoMuIDLine line and selecting a Tag hadron (pion or kaon) with one of the following PID requirements:

- **Tag π :** ProbNN $\pi > 0.8$;
- **Tag K :** ProbNN $k > 0.8$,

while on the second hadron, the Probe hadron, the $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$ muon or electron PID requirement, introduced in Sec. 4.3.6, is applied.

The pre-selection presented in Table 4.7 is applied to the B candidate and both hadron tracks, to reduce the combinatorial background, as well as to align the $B^0 \rightarrow K^+\pi^-$ samples with the $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$ selection and with the acceptance requirements applied for the calibration of PID efficiencies. Requirements placed on the $\log(1-\text{DIRA})$ and the ΔR variables are introduced to reduce the combinatorial background. For the alignment with the PID calibration samples the requirements on the HasCALO and HasRICH variables are applied. Furthermore, the fit range of the reconstructed mass distribution, $m(K\pi) \in [5.20, 5.45] \text{ GeV}/c^2$, is chosen to minimize the pollution from partially reconstructed background decays in the lower mass range and from $\Lambda_b^0 \rightarrow p\pi^-$ and $\Lambda_b^0 \rightarrow pK^-$ decays in the upper mass range.

In addition, a dedicated selection presented in Table A.9 is applied to the Tag and Probe hadrons to align the samples with the phase-space region where the data-driven PID efficiencies are valid. Depending on whether the electron PID requirement or the muon PID

Table A.9 – $B_{(s)}^0 \rightarrow hh'$ pre-selection requirements applied to the Tag and Probe hadrons.

| Variable | Applied to | Requirement |
|-------------------------------|--------------------|-------------------------------|
| p_T | Tag h | $\in [0.5, 40] \text{ GeV}/c$ |
| p | Tag h | $\in [3, 200] \text{ GeV}/c$ |
| p_T | Probe h if e | $\in [0.5, 15] \text{ GeV}/c$ |
| p | Probe h if e | $\in [3, 200] \text{ GeV}/c$ |
| p_T | Probe h if μ | $\in [0.8, 40] \text{ GeV}/c$ |
| p | Probe h if μ | $\in [3, 200] \text{ GeV}/c$ |
| InAccEcal | h if e | TRUE |
| LOCalo_ECAl_region | Probe h if e | ≥ 0 |
| abs(LO_CALO_ECAl_xProjection) | Probe h if e | $> 363.6 \text{ mm}$ |
| abs(LO_CALO_ECAl_yProjection) | Probe h if e | $> 282.6 \text{ mm}$ |
| InMuonAcc | Probe h if μ | TRUE |
| isMuon | Probe h if μ | TRUE |

Table A.10 – $B_{(s)}^0 \rightarrow hh'$ trigger selection.

| Trigger level | Trigger Selection | Applied to |
|---------------|---|------------|
| L0 | L0HadronDecision_TOS AND L0Calo_ECAl_realET > 4 GeV | Tag |
| HLT1 | Hlt1TrackMVADecision_TOS | B |
| HLT2 | Hlt2Topo2BodyDecision_TOS | B |

requirement is applied on the Probe hadron, the corresponding PID acceptance criteria are applied, *e.g.*, the InAccEcal, the L0Calo_ECAl_region and the L0_CALO_ECAl requirements for the electron PID cut or the InMuonAcc and isMuon requirements for the muon PID cut.

The trigger selection applied to the $B^0 \rightarrow K^+\pi^-$ data and simulation samples is presented in Table A.10. In general, the applied trigger lines are chosen to select two-body hadron candidates. To avoid any trigger bias on the Probe hadron, the L0 trigger is only applied on the Tag hadron. The requirement L0Calo_ECAl_realET > 4 GeV, which is applied in addition to the L0HadronDecision_TOS line, is required to align the trigger configurations used in data and simulation.

PID efficiencies are determined for the requirement ProbNN k > 0.8 (ProbNN π > 0.8) for the case that a kaon (pion) is a true kaon (pion) or misidentified as a pion (kaon) or proton with the PIDCalib2 package in bins of the kaon (pion) transverse momentum p_T and pseudorapidity η . The PID efficiencies are presented in Figs. A.34 and A.35.

$B_{(s)}^0 \rightarrow hh'$ yields estimated in simulation

The expected number of events in the invariant mass fit of the reconstructed $m(K^+\pi^-)$ distributions to data are estimated with respect to the $B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+$, following the approach presented in Sec. 4.6.1. In the following, the event yields are estimated for $B^0 \rightarrow K^+\pi^-$, $B_s^0 \rightarrow \pi^+K^-$, $B^0 \rightarrow \pi^+\pi^-$, and $B_s^0 \rightarrow K^+K^-$ decays for each Tag and Probe combination, for all three data-taking years.

The final expected $B^0 \rightarrow K^+\pi^-$, $B_s^0 \rightarrow \pi^+K^-$, $B^0 \rightarrow \pi^+\pi^-$ and $B_s^0 \rightarrow K^+K^-$ events are evaluated with Sec. 4.6.1, taking the yields presented in Table 4.14 into account. The corresponding $B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+$ selection efficiencies ($\varepsilon(B^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)K^+)$), are determined from simulation for every data taking year. The $B^0 \rightarrow K^+\pi^-$, $B_s^0 \rightarrow \pi^+K^-$, $B^0 \rightarrow \pi^+\pi^-$ and $B_s^0 \rightarrow K^+K^-$ selection efficiencies, $\varepsilon(B_{(s)}^0 \rightarrow hh')$, are determined from $B_{(s)}^0 \rightarrow h^+h^-$ simulation. The efficiencies are determined considering the full selection chain into account. However, only the corrected PID efficiencies are taken into account, *i.e.* the tracking, L0 trigger and B kinematic corrections are not applied. For the $B_{(s)}^0 \rightarrow h^+h^-$ samples, the corrected efficiency is defined as the product of the Tag and Probe PID efficiencies, $\varepsilon_{\text{PID}} = \varepsilon_{\text{PID}}(\text{Tag}) \cdot \varepsilon_{\text{PID}}(\text{Probe})$.

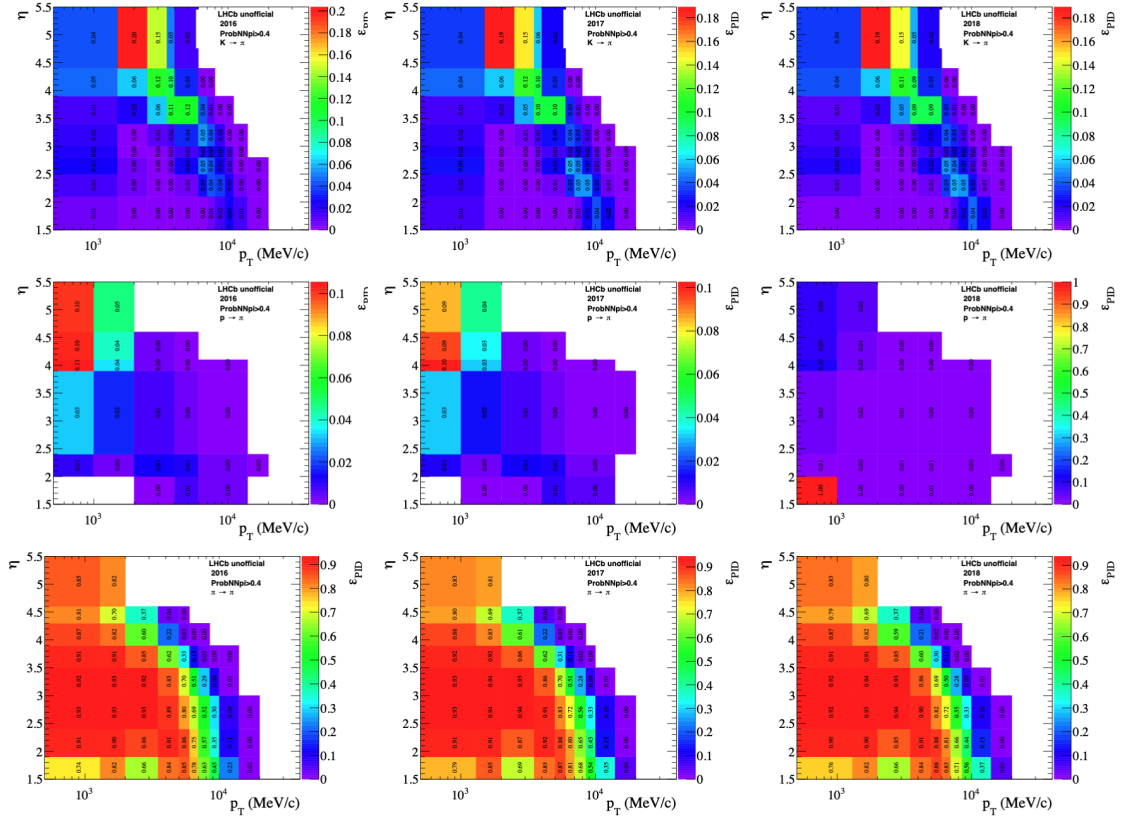


Figure A.34 – PID efficiencies for the misidentification of a kaon (top) or a proton (middle) as a pion and for the pion identification (bottom), determined for the data taking years 2016 (left), 2017 (middle) and 2018 (right).

The corrected PID efficiency of the Probe hadron is taken from Sec. 4.5.3, while the corrected PID efficiency of the Tag hadron is determined with PIDCa1b2, similarly as described in Sec. 4.5.3, for the ProbNN π and ProbNN k cuts (see App. A.4.2).

The expectations for $B^0 \rightarrow K^+ \pi^-$, $B_s^0 \rightarrow \pi^+ K^-$, $B^0 \rightarrow \pi^+ \pi^-$ and $B_s^0 \rightarrow K^+ K^-$, as well as the corresponding selection efficiencies, are presented in Tables A.11–A.14.

Appendix A. Appendix

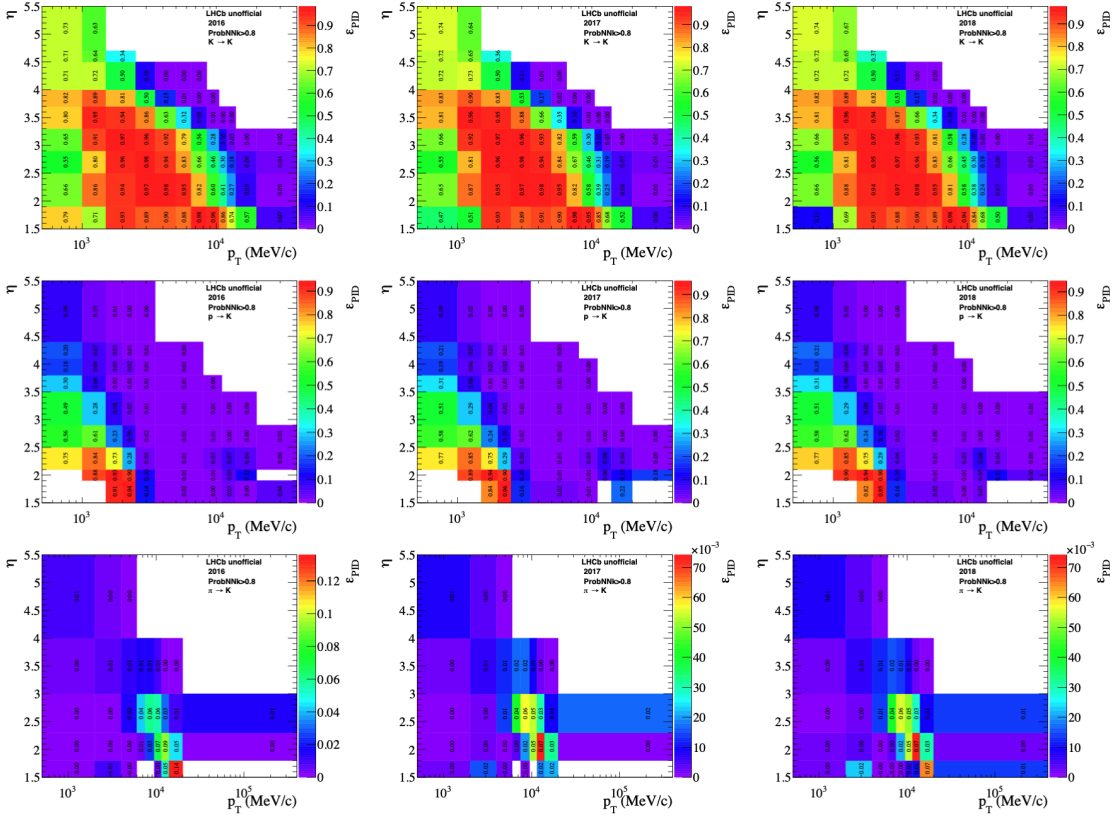


Figure A.35 – PID efficiencies for the misidentification of a pion (top) or a proton (middle) as a kaon and for the kaon identification (bottom), determined for the data taking years 2016 (left), 2017 (middle) and 2018 (right).

Table A.11 – Expected number of events (\mathcal{N}), determined from corrected $B \rightarrow hh'$ simulation, for tagging one hadron as a kaon, while testing the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ electron PID requirement on the second hadron.

| Year | Decay | $\varepsilon(B^+ \rightarrow J/\psi K^+)$ [%] | $\varepsilon(B \rightarrow hh')$ [%] | $\mathcal{N}(B \rightarrow hh')$ |
|------|-------------------------------|---|--------------------------------------|----------------------------------|
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 1.668 ± 0.005 | 0.00133 ± 0.00008 | 136 ± 9 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.668 ± 0.005 | 0.00142 ± 0.00008 | 10.9 ± 1.5 |
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.668 ± 0.005 | 0.000127 ± 0.000025 | 3.4 ± 0.7 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 1.668 ± 0.005 | 0.00034 ± 0.00006 | 12.0 ± 2.3 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 1.776 ± 0.004 | 0.00140 ± 0.00008 | 137 ± 9 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.776 ± 0.004 | 0.00149 ± 0.00009 | 10.9 ± 1.5 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.776 ± 0.004 | 0.000123 ± 0.000025 | 3.1 ± 0.6 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 1.776 ± 0.004 | 0.00035 ± 0.00004 | 11.8 ± 1.8 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 1.696 ± 0.004 | 0.00116 ± 0.00011 | 145 ± 14 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.696 ± 0.004 | 0.00124 ± 0.00011 | 11.6 ± 1.8 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.696 ± 0.004 | 0.000110 ± 0.000023 | 3.6 ± 0.8 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 1.696 ± 0.004 | 0.00028 ± 0.00005 | 12.1 ± 2.5 |

Table A.12 – Expected number of events (\mathcal{N}), determined from corrected $B \rightarrow hh'$ simulation, for tagging one hadron as a kaon, while testing the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ muon PID requirement on the second hadron.

| Year | Decay | $\varepsilon(B^+ \rightarrow J/\psi K^+) [\%]$ | $\varepsilon(B \rightarrow hh') [\%]$ | $\mathcal{N}(B \rightarrow hh')$ |
|------|-------------------------------|--|---------------------------------------|----------------------------------|
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 1.668 ± 0.005 | 0.00095 ± 0.00007 | 97 ± 8 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.668 ± 0.005 | 0.00113 ± 0.00007 | 8.6 ± 1.2 |
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.668 ± 0.005 | $(9.5 \pm 2.2) \cdot 10^{-5}$ | 2.5 ± 0.6 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 1.668 ± 0.005 | 0.00181 ± 0.00013 | 64 ± 7 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 1.776 ± 0.004 | 0.00099 ± 0.00007 | 96 ± 7 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.776 ± 0.004 | 0.00110 ± 0.00007 | 8.0 ± 1.1 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.776 ± 0.004 | 0.000100 ± 0.000022 | 2.5 ± 0.6 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 1.776 ± 0.004 | 0.00185 ± 0.00010 | 62 ± 6 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 1.696 ± 0.004 | 0.00089 ± 0.00009 | 111 ± 12 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.696 ± 0.004 | 0.00091 ± 0.00009 | 8.6 ± 1.4 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.696 ± 0.004 | $(8.7 \pm 2.1) \cdot 10^{-5}$ | 2.9 ± 0.7 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 1.696 ± 0.004 | 0.00153 ± 0.00012 | 66 ± 8 |

Table A.13 – Expected number of events ($B \rightarrow hh'$), determined from corrected $B_{(s)}^0 \rightarrow h^+ h^-$ simulation, for tagging one hadron as a pion, while testing the $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ electron PID requirement on the second hadron.

| Year | Decay | $\varepsilon(B^+ \rightarrow J/\psi K^+) [\%]$ | $\varepsilon(B \rightarrow hh') [\%]$ | $\mathcal{N}(B \rightarrow hh')$ |
|------|-------------------------------|--|---------------------------------------|----------------------------------|
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 1.668 ± 0.005 | 0.000142 ± 0.000027 | 14.5 ± 2.8 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.668 ± 0.005 | 0.000158 ± 0.000028 | 1.21 ± 0.26 |
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.668 ± 0.005 | 0.00206 ± 0.00010 | 54.9 ± 3.5 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 1.668 ± 0.005 | $(1.8 \pm 1.7) \cdot 10^{-5}$ | 0.6 ± 0.6 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 1.776 ± 0.004 | 0.000151 ± 0.000028 | 14.7 ± 2.7 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.776 ± 0.004 | 0.000164 ± 0.000029 | 1.20 ± 0.26 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.776 ± 0.004 | 0.00225 ± 0.00010 | 57 ± 4 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 1.776 ± 0.004 | $(1.6 \pm 1.0) \cdot 10^{-5}$ | 0.54 ± 0.35 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 1.696 ± 0.004 | 0.000121 ± 0.000035 | 15 ± 4 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.696 ± 0.004 | 0.00013 ± 0.00004 | 1.2 ± 0.4 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.696 ± 0.004 | 0.00187 ± 0.00009 | 61 ± 4 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 1.696 ± 0.004 | $(1.3 \pm 1.5) \cdot 10^{-5}$ | 0.6 ± 0.6 |

Appendix A. Appendix

Table A.14 – Expected number of events (\mathcal{N}), determined from corrected $B \rightarrow hh'$ simulation, for tagging one hadron as a pion, while testing the $B \rightarrow hh'$ muon PID requirement on the second hadron.

| Year | Decay | $\varepsilon(B^+ \rightarrow J/\psi K^+) [\%]$ | $\varepsilon(B \rightarrow hh') [\%]$ | $\mathcal{N}(B \rightarrow hh')$ |
|------|-------------------------------|--|---------------------------------------|----------------------------------|
| 2016 | $B^0 \rightarrow K^+ \pi^-$ | 1.668 ± 0.005 | 0.00080 ± 0.00006 | 82 ± 7 |
| 2016 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.668 ± 0.005 | 0.00080 ± 0.00006 | 6.1 ± 0.9 |
| 2016 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.668 ± 0.005 | 0.00180 ± 0.00009 | 47.9 ± 3.2 |
| 2016 | $B_s^0 \rightarrow K^+ K^-$ | 1.668 ± 0.005 | 0.000110 ± 0.000035 | 3.9 ± 1.3 |
| 2017 | $B^0 \rightarrow K^+ \pi^-$ | 1.776 ± 0.004 | 0.00089 ± 0.00007 | 87 ± 7 |
| 2017 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.776 ± 0.004 | 0.00099 ± 0.00007 | 7.2 ± 1.0 |
| 2017 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.776 ± 0.004 | 0.00202 ± 0.00010 | 51.3 ± 3.3 |
| 2017 | $B_s^0 \rightarrow K^+ K^-$ | 1.776 ± 0.004 | 0.000101 ± 0.000023 | 3.4 ± 0.8 |
| 2018 | $B^0 \rightarrow K^+ \pi^-$ | 1.696 ± 0.004 | 0.00078 ± 0.00009 | 98 ± 11 |
| 2018 | $B_s^0 \rightarrow \pi^+ K^-$ | 1.696 ± 0.004 | 0.00077 ± 0.00009 | 7.2 ± 1.2 |
| 2018 | $B^0 \rightarrow \pi^+ \pi^-$ | 1.696 ± 0.004 | 0.00177 ± 0.00009 | 58 ± 4 |
| 2018 | $B_s^0 \rightarrow K^+ K^-$ | 1.696 ± 0.004 | $(8.3 \pm 3.0) \cdot 10^{-5}$ | 3.6 ± 1.3 |

Mass fits of $B^0 \rightarrow K^+ \pi^-$ data samples

The reconstructed $m(K^+ \pi^-)$ distribution is fitted in $B_{(s)}^0 \rightarrow hh'$ data to determine the $B^0 \rightarrow K^+ \pi^-$, $B_s^0 \rightarrow \pi^+ K^-$, $B^0 \rightarrow \pi^+ \pi^-$ and $B_s^0 \rightarrow K^+ K^-$ yields. The fit is performed for each Tag and Probe combination, for all three data-taking years. The obtained yields are then compared to the expected number of events evaluated from simulation.

The $B^0 \rightarrow K^+ \pi^-$, $B_s^0 \rightarrow \pi^+ K^-$, $B^0 \rightarrow \pi^+ \pi^-$ and $B_s^0 \rightarrow K^+ K^-$ components are modeled with the sum of two Crystal Ball functions, sharing the same mean. The tail parameters are determined in $B_{(s)}^0 \rightarrow hh'$ simulation and fixed in the fit to data. All fit components have a shared parameter, which allows shifts on the mean of the distributions, with respect to the parameters obtained in simulation. Similarly, another parameter shared between all components is introduced, allowing the scaling of the width of the distributions in the fit to data.

The $B_s^0 \rightarrow \pi^+ K^-$, $B^0 \rightarrow \pi^+ \pi^-$ and $B_s^0 \rightarrow K^+ K^-$ yields, X , are constrained with respect to the $B^0 \rightarrow K^+ \pi^-$ yield, as

$$\mathcal{N}_X = f_X \cdot \mathcal{N}_{B^0 \rightarrow K^+ \pi^-}, \quad (\text{A.1})$$

where f_X is constrained in the fit and defined as

$$f_X = \frac{f_{s,(d)}}{f_d} \cdot \frac{\varepsilon(X)}{\varepsilon(B^0 \rightarrow K^+ \pi^-)} \cdot \frac{\mathcal{B}(X)}{\mathcal{B}(B^0 \rightarrow K^+ \pi^-)}. \quad (\text{A.2})$$

The values taken for the branching fraction, \mathcal{B} are presented Table 4.18. The selection efficiencies, $\varepsilon(B^0 \rightarrow K^+ \pi^-)$ and $\varepsilon(X)$, are taken from the $B_{(s)}^0 \rightarrow hh'$ simulation estimation, introduced above. The choice of the here introduced constraint is motivated by the low-statistics, observed

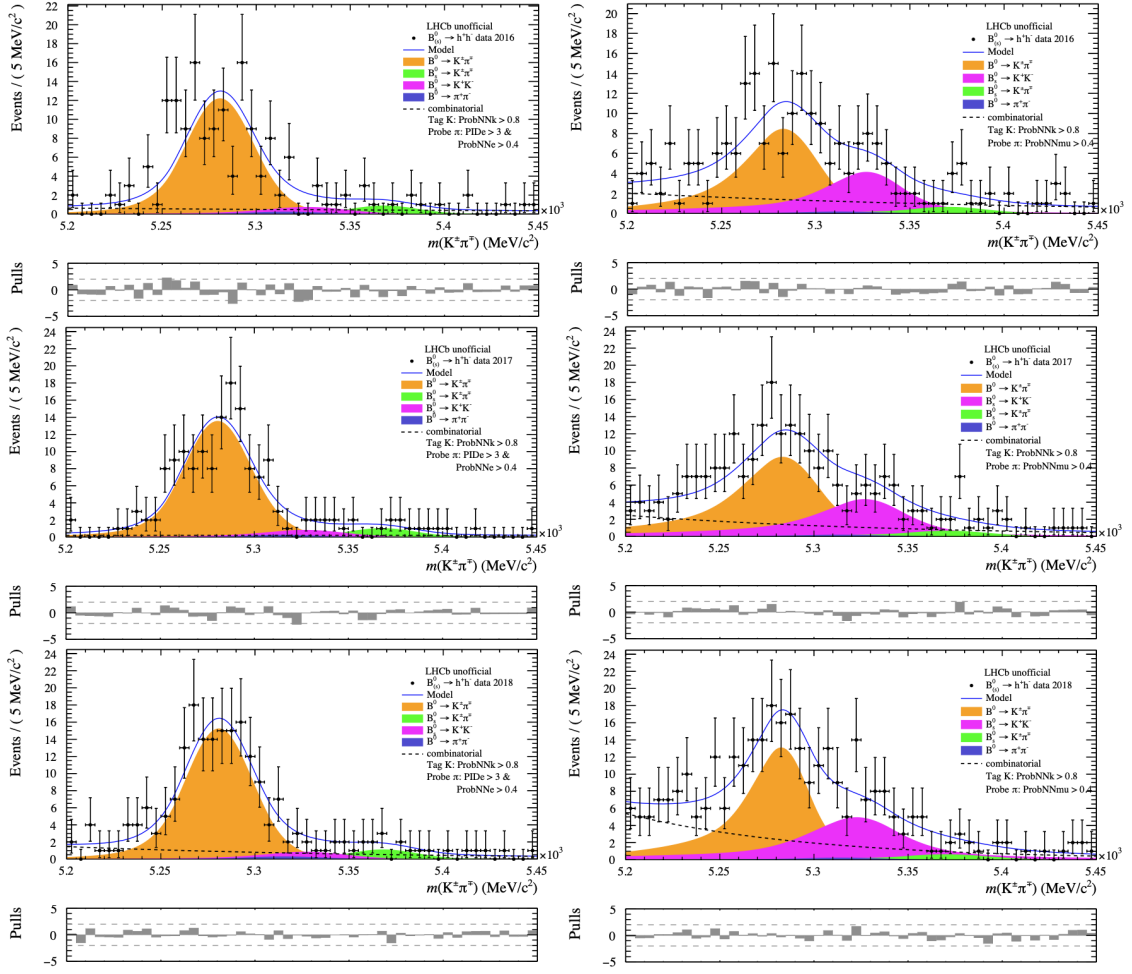


Figure A.36 – Fit to the $m(K\pi)$ distribution reconstructed in $B_{(s)}^0 \rightarrow hh'$ data for the data-taking years 2016 (top), 2017 (middle) and 2018 (bottom). Left: the kaon is tagged and the electron PID cut is probed on the second hadron. Right: the kaon is tagged and the muon PID cut is probed on the second hadron.

in the $B_{(s)}^0 \rightarrow hh'$ data samples and assures the stability and robustness of the fit.

In addition to the peaking components, the combinatorial background is modeled with an exponential for each data-taking year and Tag and Probe combination. The corresponding yield and slope parameters are left floating in the fit.

The fits are performed as extended unbinned maximum likelihood fits to the reconstructed $m(K^+\pi^-)$ distribution. and shown in Figs. A.36 and A.37 for the four tag-and -probe configurations.

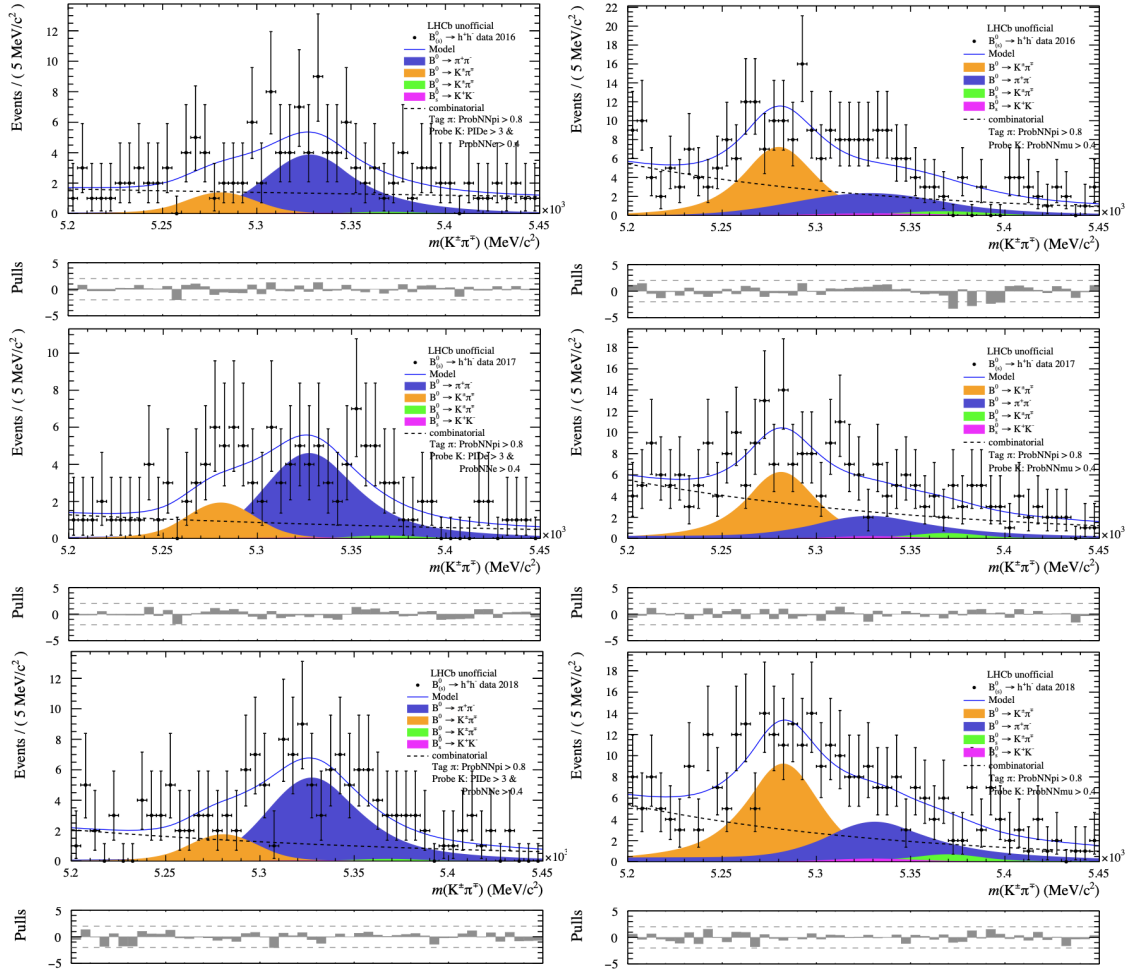


Figure A.37 – Fit to the $m(K\pi)$ distribution reconstructed in $B_{(s)}^0 \rightarrow hh'$ data for the data-taking years 2016 (top), 2017 (middle) and 2018 (bottom). Left: the pion is tagged and the electron PID cut is probed on the second hadron. Right: the pion is tagged and the muon PID cut is probed on the second hadron.

A.5 Pass-fail cross-check validation of peaking backgrounds

The weights $\omega_{\text{pass/fail}}(K \rightarrow e)$, $\omega_{\text{pass/fail}}(K \rightarrow \mu)$, $\omega_{\text{pass/fail}}(\pi \rightarrow e)$ and $\omega_{\text{pass/fail}}(\pi \rightarrow \mu)$, introduced in Sec. 4.6.3, are displayed in Figs. A.38–A.41.

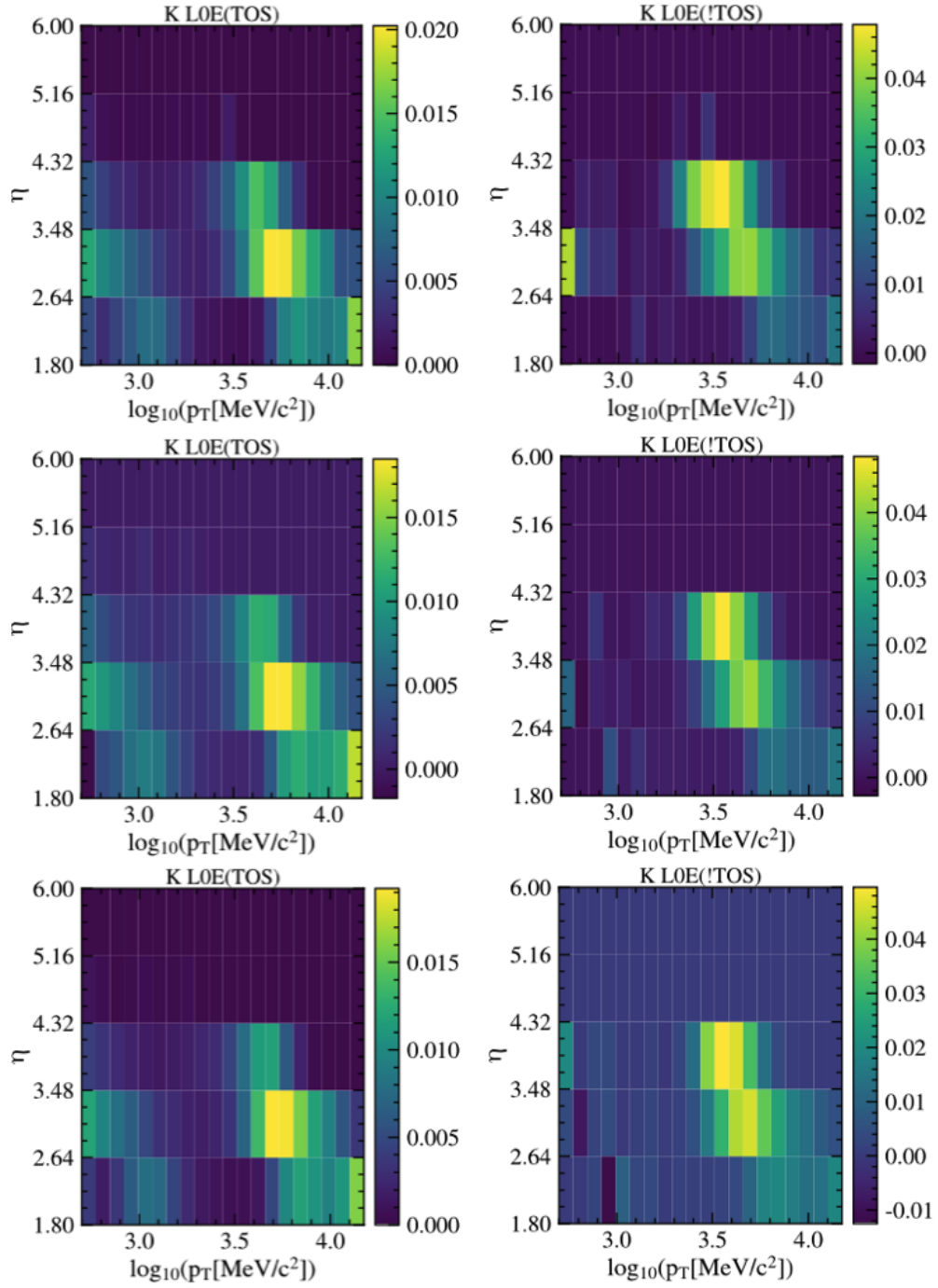


Figure A.38 – Pass-fail weights for the misidentification of a true kaon as an electron, $\omega_{\text{pass/fail}}(K \rightarrow e)$, applied to the electron candidate in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data selected with inverted nominal PID cuts. The weights are obtained for the data taking year 2016 (top), 2017 (middle) and 2018 (bottom), for events firing the L0Electron_TOS trigger line (left) and for events not firing that (right).

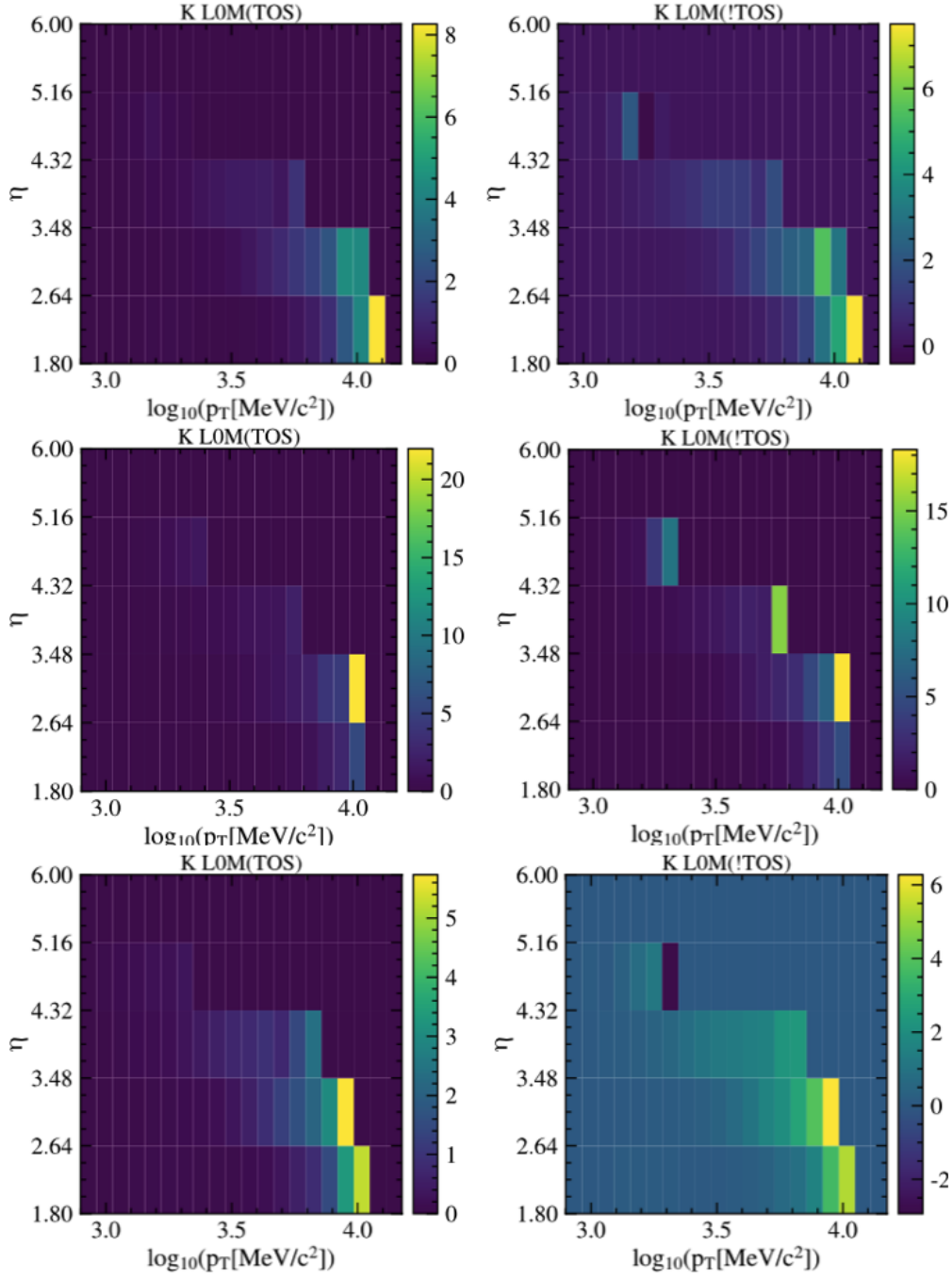


Figure A.39 – Weights for the misidentification of a true kaon as a muon, $\omega_{\text{pass/fail}}(K \rightarrow \mu)$, applied to the muon candidate in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data selected with inverted nominal PID cuts. The weights are obtained for the data taking year 2016 (top), 2017 (middle) and 2018 (bottom), for events firing the LOMuon_TOS trigger line (left) and for events not firing that (right).

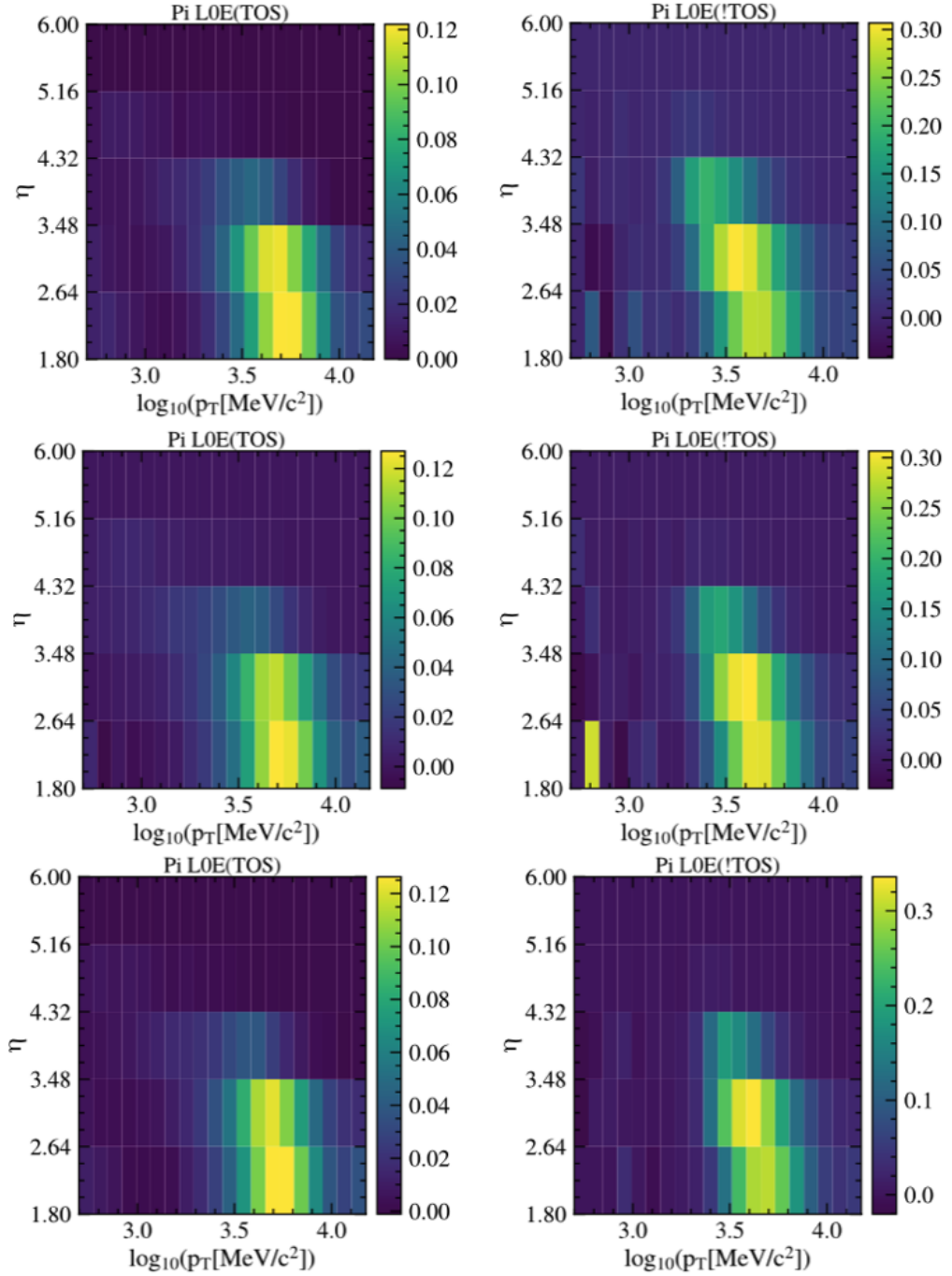


Figure A.40 – Weights for the misidentification of a true pion as an electron, $\omega_{\text{pass/fail}}(\pi \rightarrow e)$, applied to the electron candidate in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data selected with inverted nominal PID cuts. The weights are obtained for the data taking year 2016 (top), 2017 (middle) and 2018 (bottom), for events firing the L0Electron_TOS trigger line (left) and for events not firing that (right).

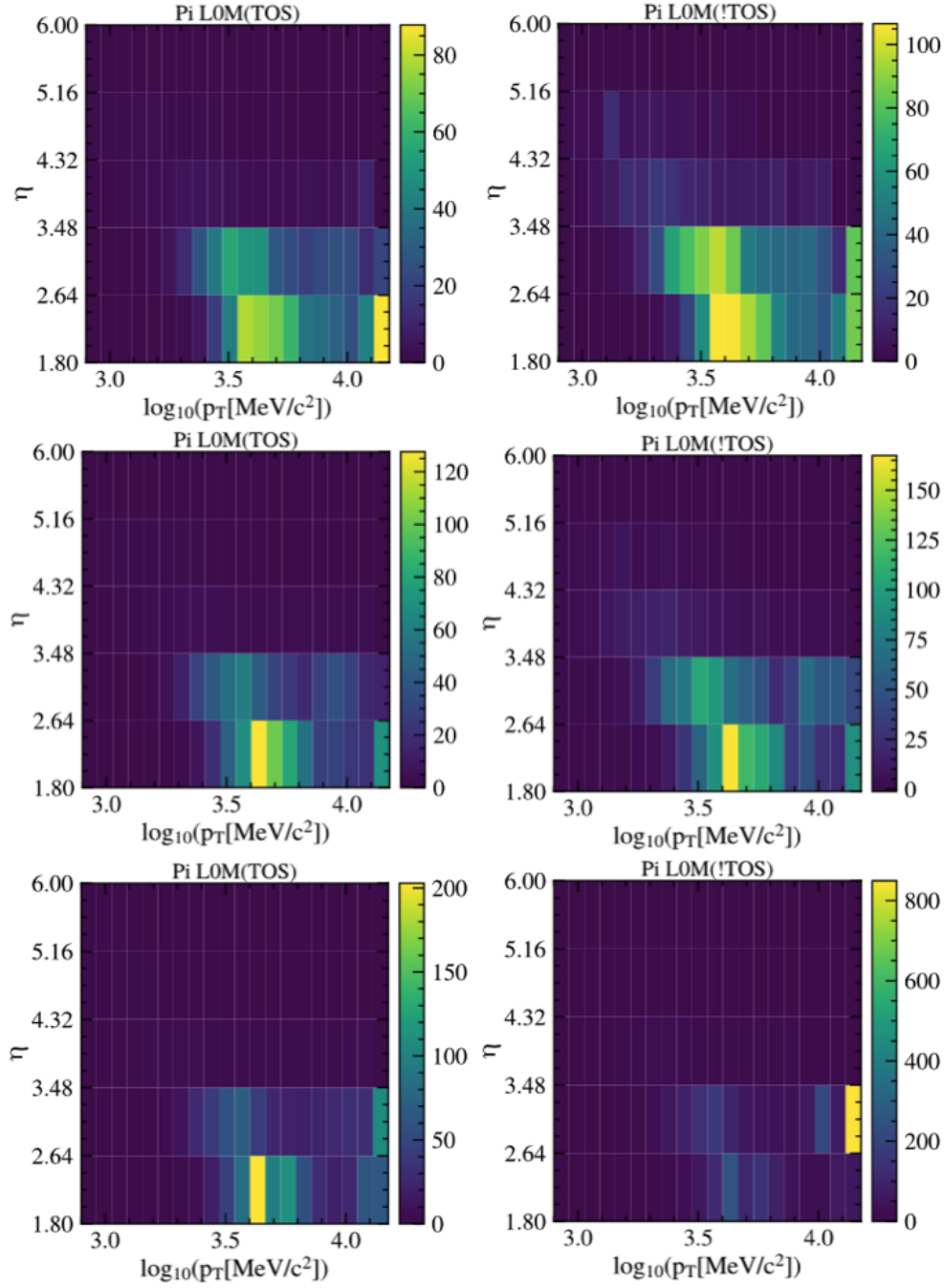


Figure A.41 – Weights for the misidentification of a true pion as a muon, $\omega_{\text{pass/fail}}(\pi \rightarrow \mu)$, applied to the muon candidate in $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ data selected with inverted nominal PID cuts. The weights are obtained for the data taking year 2016 (top), 2017 (middle) and 2018 (bottom), for events firing the LOMuon_TOS trigger line (left) and for events not firing that (right).

Bibliography

- [1] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.
- [2] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. **D2** (1970) 1285.
- [3] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. 1*, Phys. Rev. **D8** (1973) 3633.
- [4] H. D. Politzer, *Reliable perturbative results for strong interactions?*, Phys. Rev. Lett. **30** (1973) 1346.
- [5] H. D. Politzer, *Asymptotic freedom: an approach to strong interactions*, Phys. Rept. **14** (1974) 129.
- [6] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.
- [7] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [8] C. Burgard, *Standard model of physics*, <https://texample.net/tikz/examples/model-physics/>. Accessed: 2021-07-15.
- [9] G. Altarelli, *A QCD primer*, AIP Conf. Proc. **631** (2002) 70, arXiv:hep-ph/0204179.
- [10] W. Hollik, *Quantum field theory and the Standard Model*, in *2009 European School of High-Energy Physics*, 2010, arXiv:1012.3883.
- [11] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.
- [12] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [13] Particle Data Group, R. L. Workman *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2022** (2022) 083C01.
- [14] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.
- [15] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

Bibliography

- [16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.
- [17] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [18] CMS collaboration, D. Horvath for the CMS collaboration, *Search for the Higgs boson: a statistical adventure of exclusion and discovery*, J. Phys. Conf. Ser. **510** (2014) 012001, arXiv:1310.6839.
- [19] ATLAS collaboration, 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.
- [20] CMS collaboration, 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.
- [21] Super-Kamiokande collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.
- [22] MEG collaboration, A. M. Baldini *et al.*, *Search for the lepton flavour violating decay $\mu^+ \rightarrow e^+ \gamma$ with the full dataset of the MEG experiment*, Eur. Phys. J. **C76** (2016) 434, arXiv:1605.05081.
- [23] W. J. Marciano, T. Mori, and J. M. Roney, *Charged lepton flavor violation experiments*, Ann. Rev. Nucl. Part. Sci. **58** (2008) 315.
- [24] 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.
- [25] I. de Medeiros Varzielas and G. Hiller, *Clues for flavor from rare lepton and quark decays*, JHEP **06** (2015) 072, arXiv:1503.01084.
- [26] A. Crivellin, L. Hofer, J. Matias, U. Nierste, S. Pokorski and J. Rosiek, *Lepton-flavour violating B decays in generic Z' models*, Phys. Rev. **D92** (2015) 054013, arXiv:1504.07928.
- [27] A. Ilakovac, *Lepton flavor violation in the standard model extended by heavy singlet Dirac neutrinos*, Phys. Rev. **D62** (2000) 036010, arXiv:hep-ph/9910213.
- [28] 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.
- [29] J. C. Pati and A. Salam, *Lepton number as the fourth color*, Phys. Rev. **D10** (1974) 275, erratum Phys. Rev. **D11** (1975) 703.
- [30] R. Bernstein, *Charged lepton flavour violation: an overview*, PASCOS 2016: 22nd International Symposium on Particles, Strings and Cosmology. Quy Nhon, Vietnam, July, 2016.

-
- [31] W. J. Marciano and A. I. Sanda, *Exotic decays of the muon and heavy leptons in gauge theories*, Phys. Lett. **B67** (1977) 303.
 - [32] S. T. Petcov, *The processes $\mu \rightarrow e + \gamma, \mu \rightarrow e + \bar{e}, \nu' \rightarrow \nu + \gamma$ in the Weinberg-Salam model with neutrino mixing*, Sov. J. Nucl. Phys. **25** (1977) 340, erratum Sov. J. Nucl. Phys. **25** (1977) 698, erratum Yad. Fiz. **25** (1977) 1336.
 - [33] 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.
 - [34] LHCb collaboration, R. Aaij *et al.*, *Measurement of lepton universality parameters in $B^+ \rightarrow K^+ \ell^+ \ell^-$ and $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, Phys. Rev. **D108** (2023) 032002, arXiv:2212.09153.
 - [35] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in $b \rightarrow s \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **131** (2023) 051803, arXiv:2212.09152.
 - [36] LHCb collaboration, *Measurement of the ratios of branching fractions $\mathcal{R}(D^*)$ and $\mathcal{R}(D^0)$* , Phys. Rev. Lett. **131** (2023) 111802, arXiv:2302.02886.
 - [37] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavour universality using $B^0 \rightarrow D^{*-} \tau^+ \nu_\tau$ decays with hadronic τ channels*, Phys. Rev. **D108** (2023) 012018, arXiv:2305.01463.
 - [38] SINDRUM collaboration, U. Bellgardt *et al.*, *Search for the decay $\mu^+ \rightarrow e^+ e^+ e^+$* , Nucl. Phys. **B299** (1988) 1.
 - [39] SINDRUM II collaboration, W. H. Bertl *et al.*, *A search for muon to electron conversion in muonic gold*, Eur. Phys. J. **C47** (2006) 337.
 - [40] LHCb collaboration, R. Aaij *et al.*, *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.
 - [41] M. Blanke, *Introduction to flavour physics and CP violation*, CERN Yellow Rep. School Proc. **1705** (2017) 71, arXiv:1704.03753.
 - [42] LHCb collaboration, 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.
 - [43] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.
 - [44] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*, <https://cds.cern.ch/record/2684277>, 2019. Accessed: 2023-09-19.
 - [45] ATLAS collaboration, G. Aad *et al.*, *The ATLAS experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
 - [46] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

Bibliography

- [47] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [48] LHCb collaboration, A. A. Alves *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [49] C. Elsässer for the LHCb collaboration, $\bar{b}b$ production angle plot, https://lhcb.web.cern.ch/speakersbureau/html/bb_productionangles.html. Accessed: 2021-07-27.
- [50] M. L. Mangano and G. Altarelli, *CERN Workshop on Standard Model Physics (and more) at the LHC*, (Geneva), CERN, 2000, <http://cds.cern.ch/record/425440>. Accessed: 2023-09-20.
- [51] LHCb collaboration, *LHCb integrated luminosity*, <https://http://lhcb-public.web.cern.ch/Images2015/Images2018/IntRecLumiR12.png>. Accessed: 2021-09-06.
- [52] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb VELO TDR: Vertex locator. Technical design report*, Technical design report. LHCb, CERN, Geneva, 2001.
- [53] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, [arXiv:1405.7808](https://arxiv.org/abs/1405.7808).
- [54] LHCb collaboration, S. Amato *et al.*, *LHCb magnet: Technical design report*, Technical design report. LHCb, CERN, Geneva, 2000.
- [55] LHCb collaboration, R. Antunes-Nobrega *et al.*, *LHCb reoptimized detector design and performance: Technical design report*, Technical design report. LHCb, CERN, Geneva, 2003.
- [56] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb inner tracker: Technical design report*, Technical design report. LHCb, CERN, Geneva, 2002.
- [57] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb outer tracker: Technical design report*, Technical design report. LHCb, CERN, Geneva, 2001.
- [58] LHCb Outer Tracker group, R. Arink *et al.*, *Performance of the LHCb outer tracker*, JINST **9** (2014) P01002, [arXiv:1311.3893](https://arxiv.org/abs/1311.3893).
- [59] LHCb collaboration, S. Amato *et al.*, *LHCb RICH: Technical design report*, Technical design report. LHCb, CERN, Geneva, 2000.
- [60] LHCb RICH group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759](https://arxiv.org/abs/1211.6759).
- [61] LHCb collaboration, S. Amato *et al.*, *LHCb calorimeters: Technical design report*, Technical design report. LHCb, CERN, Geneva, 2000.
- [62] LHCb collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb muon system: Technical design report*, Technical design report. LHCb, CERN, Geneva, 2001.

-
- [63] G. Andreassi, *Lepton flavour violation searches and scintillating fibre tracker testing in LHCb*, PhD thesis #9251, EPFL, Lausanne, 2019.
- [64] LHCb collaboration, I. Bediaga *et al.*, *LHCb trigger and online upgrade technical design report*, Technical design report. LHCb, CERN, Geneva, 2014.
- [65] LHCb collaboration, *Trigger schemes*, <http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html>. Accessed: 2021-09-06.
- [66] T. Monnard, *Evaluation of the systematic uncertainties due to the simulation mismodeling in the search for $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$ decays at LHCb*, Master thesis, EPFL, Lausanne, 2023.
- [67] LHCb collaboration, R. Aaij *et al.*, *Letter of intent for the LHCb upgrad*, Letter of intent, CERN, Geneva, 2011.
- [68] LHCb collaboration, I. Bediaga *et al.*, *LHCb VELO upgrade technical design report*, Technical design report. LHCb, CERN, Geneva, 2013.
- [69] LHCb collaboration, *LHCb detector upgrade*, <http://lhcb-public.web.cern.ch/Images2015/Images2018/LHCbUpgrade.jpg>. Accessed: 2021-11-08.
- [70] LHCb collaboration, A. A. Alves Jr *et al.*, *LHCb tracker upgrade technical design report*, Technical design report. LHCb, CERN, Geneva, 2014.
- [71] LHCb collaboration, I. Bediaga *et al.*, *LHCb PID upgrade technical design report*, Technical design report. LHCb, CERN, Geneva, 2013.
- [72] LHCb collaboration, C. Joram, U. Uwer, B. D. Leverington, T. Kirn, S. Bachmann, R. J. Ekelhof and J. Müller, *LHCb Scintillating Fibre tracker engineering design review report: fibres, mats and modules*, Technical design report. LHCb, CERN, Geneva, 2015.
- [73] A. Kuonen, *Development and characterisation of Silicon Photomultiplier Multichannel arrays for the readout of a large scale Scintillating Fibre tracker*, PhD thesis #8842, EPFL, Lausanne, 2018.
- [74] C. Piemonte, *A new silicon photomultiplier structure for blue light detection*, Nuc. Instrum. Meth. **A568** (2006) 224.
- [75] LHCb collaboration, R. Aaij *et al.*, *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B_s^0 decay branching fractions*, Phys. Rev. **D104** (2021) 032005, arXiv:2103.06810.
- [76] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693.
- [77] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.
- [78] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*, Taylor & Francis, 1984.

Bibliography

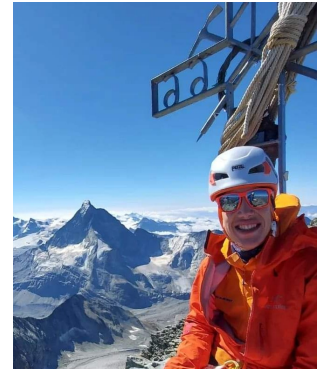
- [79] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: bounds for k-fold and progressive cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, (New York, NY, USA), 203–208, Association for Computing Machinery, 1999.
- [80] F Pedregosa *et al.*, *Scikit-learn: Machine learning in python*, J. Mach. Learn. Res. **12** (2011) 2825, arXiv:1201.0490.
- [81] F. J. Massey, *The Kolmogorov-Smirnov test for goodness of fit*, J. Am. Stat. Assoc. **46** (1951) 68.
- [82] LHCb collaboration, R. Aaij *et al.*, *Measurement of the electron reconstruction efficiency at LHCb*, JINST **14** (2019) P11023, arXiv:1909.02957.
- [83] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.
- [84] LHCb collaboration, L. Anderlini *et al.*, *The PIDCalib package*, Technical design report. LHCb, CERN, Geneva, 2016.
- [85] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- [86] Y.-C. Chen, *A tutorial on kernel density estimation and recent advances*, Biostatistics & Epidemiology **1** (2017) 161, arXiv:1704.03924.
- [87] LHCb collaboration, S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Technical design report. LHCb, CERN, Geneva, 2014.
- [88] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016) 012036, arXiv:1608.05806, https://github.com/arogozhnikov/hep_ml.
- [89] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.
- [90] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13 TeV pp collisions*, Phys. Rev. **D100** (2019) 031102(R), arXiv:1902.06794.
- [91] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727, erratum Eur. Phys. J. **C73** (2013) 2501.
- [92] B. Efron, *Bootstrap methods: Another look at the Jackknife*, Ann. Statist. **7** (1979) 1.
- [93] J. A. Hanley and B. MacGibbon, *Creating non-parametric bootstrap samples using poisson frequencies*, Comp. Meth. Prog. Bio. **83** (2006) 57.
- [94] G. J. Feldman and R. D. Cousins, *A unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873, arXiv:physics/9711021.

- [95] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay properties and search for the $B^0 \rightarrow \mu^+ \mu^-$ and $B_s^0 \rightarrow \mu^+ \mu^- \gamma$ decays*, Phys. Rev. **D105** (2022) 012010, arXiv:2108.09283.



Curriculum Vitae

Sebastian Schulte



Skilled in data analysis and statistics, experienced in collaborative teamwork, certified in project management and a strong interest in analytical problem-solving.

Experience and Projects

- 2018–today** Doctoral Researcher, *École Polytechnique Fédérale de Lausanne* and *European Organization for Nuclear Research (CERN)*, Switzerland.
- **Search for new phenomena in particle physics, analysing large datasets** using statistical and computational methods to assess signal significance, estimate background and establish confidence levels for theoretical predictions.
 - ▷ Managing and developing the data analysis while consistently reporting progress and results of the project to an expert working group at *CERN*.
 - ▷ Data mining with boosted decision trees and neural networks.
 - ▷ Development of algorithms for modeling physical processes using numerical simulation.
 - **Responsible for the quality assurance testing of Silicon Photomultiplier Detectors** for the LHCb experiment at *CERN*.
 - ▷ Co-management of the quality testing, supervising a team of technicians.
 - ▷ Analysing status, progress and issues, while reporting the results to the project executives.
- 2015–2018** Master Student and Scientific Collaborator at CERN, *Technical University of Munich* and *Max Planck Institute for Physics*, Germany.
- **Precision data analysis**, determining theoretical physics models with numerical simulation and developing algorithms for a multi-dimensional statistical minimisation.
- 01–04/2016** Student Trainee, *Max Planck Institute for Physics*, Germany.
- **R&D of innovative muon detectors**, performing electronic measurements and analysing results.
- 2012–2013** Student Trainee, *Heinz Maier-Leibnitz Research Neutron Source*, Germany.
- **Development of a security system** for a new neutron scattering experiment based on a Programmable Logic Controller (PLC).
- 06–07/2009** Electronics Technician for Automation Technology, *Deckel Maho Pfronten GmbH*, Germany.
- **Electromechanical assembly of milling machines**, installing electronic components and programming PLCs.

Education and Certifications

- 2018–today** Ph.D. in Physics, *École Polytechnique Fédérale de Lausanne*, Switzerland.
- **Thesis:** Search for the lepton flavour-violating decays $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ with LHCb Run 2 data. (Defended on the 12.12.2023)
- 11/2023** Certification in Project Management, *PRINCE2 Foundation*, *AXELOS*, United Kingdom.

- 07/2020** Summer School on Machine Learning in High Energy Physics, *École Polytechnique Fédérale de Lausanne*.
- **Topic:** Application of modern machine learning techniques in data mining and data analytics.
- 08/2019** SLAC Summer Institute, *Stanford University*, USA.
- **Topic:** Theoretical concepts and modern experimental measurements in flavour physics.
- 2015–2018** Master of Science in Physics, *Technical University of Munich and Max Planck Institute for Physics*, Germany.
- **Thesis:** Studies for the measurement of the top-quark mass in the $t\bar{t} \rightarrow \text{lepton} + \text{jets}$ channel, with $\sqrt{s} = 13$ TeV proton-proton collision data from the ATLAS experiment.
 - ▷ Passed with distinction.
- 2012–2015** Bachelor of Science in Physics, *Technical University of Munich*, Germany.
- 2011–2012** Student in Electrical Engineering, *Technical University of Munich*, Germany.
- Transition to physics due to a strong interest in more fundamental concepts, yet successfully completed all examinations.
- 2009–2011** Abitur (School Graduation), *College of higher education*, Germany.
- 2006–2009** Apprenticeship Electronics Technician, *Deckel Maho Pfronten GmbH*, Germany.
- Trainee programme with a focus on Automation Technologies and PLC programming.
 - ▷ Diploma from the German Chamber of Industry and Commerce.

Software and Programming Skills

- **Advanced:** Python, Pandas, Microsoft Office, L^AT_EX, Root Data Analysis Framework, Algorithm design
- **Intermediate:** C/C++, Matlab, Siemens Simatic S7, Siemens LOGO, Linux distributions
- **Basic:** Fortran, Tensorflow, Beckhoff Twincat, PILZ PNOZmulti

Languages

- German (native)
- English (fluent, C2)
- French (pre-intermediate, A1–A2)

Interests

- Algorithm design
- Project coordination and management
- Data analysis
- Interdisciplinary applications
- Analytical and strategical thinking
- Statistics

Alpine Climbs

- 2019-today** Highlights: Weissmies (4013m), Matterhorn (4478m), Alalinhorn (4027m), Strahlhorn (4190m), Bis-horn (4151m), Lagginhorn (4010m), Mont Blanc (4809m), Eiger (3967m), Nadelgrat (Dürrenhorn 4035m, Hohberghorn 4218m and Stechnadelhorn 4240m), Dent Blanche (4358m) and Gross Gühorn (4043m).