University of Zurich UZH

A Test of Lepton Flavor Universality with $B^0 \to K^+\pi^-\ell^+\ell^-$ Decays at LHCb

BACHELOR THESIS

AUTHOR
NATHAN BAUDIS

SUPERVISORS
PROF. NICOLA SERRA
DR. RAFAEL SILVA COUTINHO
DEPARTMENT OF PHYSICS
UNIVERSITY OF ZURICH

21ST OF JUNE, 2021

Abstract

The rare decays $B^0 \to (K^+\pi^-)^*_J \mu^+\mu^-$ and the previously unobserved $B^0 \to (K^+\pi^-)^*_J e^+e^-$ are examined in the context of a Lepton Flavor Universality measurement of the ratio $R_{K\pi}$ between their two branching fractions. The analysis is performed using data collected at LHCb during 2016 at a center-of-mass energy of $13\,\text{TeV}$ with an integrated luminosity of $1.6\,\text{fb}^{-1}$. The dilepton mass-squared range considered is $1.1 < q^2 < 6.0\,\text{GeV}^2/c^4$. Two regions of the invariant mass $m(K^+\pi^-)$ above the resonance $K^*_1(892)^0$ are studied separately for both of the lepton final states: $1000 < m(K^+\pi^-) < 1835\,\text{MeV}/c^2$ and $1895 < m(K^+\pi^-) < 2600\,\text{MeV}/c^2$. Invariant mass fits on the reconstructed B^0-mass are performed to determine the observed decay yields. The electron channel $B^0 \to K^+\pi^-e^+e^-$ is observed in the lower and upper $m(K^+\pi^-)$ regions with a statistical significance of $5.8\,\sigma$ and $0.86\,\sigma$, respectively. Pseudo-experiments are performed to account for biases in the fits and estimate the obtainable sensitivity on $R_{K\pi}$. The statistical sensitivity on $R_{K\pi}$ using the 2011–2018 LHCb dataset is expected to be at the level of $11\,\%$, providing a novel and complementary probe of Lepton Flavor Universality.

Acknowledgements

Contents

Chapter 1

Introduction

The Standard Model of particle physics has served its purpose well by delivering precise explanations for a multitude of phenomena in a remarkably concise framework. However it cannot be the end of the story, as it does not include descriptions of observed phenomena such as dark matter or the amount of matter-antimatter asymmetry in our Universe. Direct searches for New Physics beyond the Standard Model, carried out by looking for signs of new particles in high energy collisions have revealed the last missing piece of the Standard Model, the Higgs boson, but otherwise not delivered any such signs [1, 2].

Broadly speaking, indirect searches are another type of approach to New Physics searches, where new particles are not directly detected through their decay products but by the signatures they impart onto other decays during the decay process. Recently, hints and evidence have been emerging from such indirect searches that one of the features of the Standard Model, so-called Lepton Flavor Universality (LFU), may not in fact be realized in nature. This is part of a bigger set of observed anomalies that appear to be forming a coherent picture of physics beyond the Standard Model [3, 4].

Lepton Flavor Universality states that the charged leptons, i.e. electrons and their heavier cousins, muons and taus, have identical properties and interactions aside from their mass. To increase the sensitivity to New Physics via a test of LFU, very rare decays are studied, where new virtual particles have a chance to compete and measurably change the Standard Model prediction [3, 5].

One of the goals of the LHCb experiment at the Large Hadron Collider is to conduct such indirect searches in the decay of b-hadrons, bound states including a heavy bottom quark which quickly decay [6]. By precisely comparing the rate of rare decays of such hadrons into different leptons, e.g. muons and electrons, their behavior can be compared, giving a test of LFU.

Measurements at LHCb have been providing the most precise tests of this kind [3], with the most recent result constituting the first evidence for the breaking of Lepton Flavor Universality [4]. However, the statistical significance is as of yet too low to conclude that New Physics have indeed been found. To this end, more data will be gathered with upcoming upgrades to the LHC and more years of data-taking. Additionally, more decays are being investigated to offer potential tests of LFU. In this thesis, $B^0 \to K^+\pi^-\ell^+\ell^-$ decays are explored as such a new test using data collected at LHCb, where ℓ denotes muons and electrons.[1] With the same final state particles as the previously studied $B^0 \to K^{*0}\ell^+\ell^-$ decays [7], many criteria in the selection of events can be transferred or adapted in a straight-forward way.

The remainder of this thesis is structured as follows: In Chapter 2, a brief introduction to the Standard Model of particle physics is given, with a contextualization of Lepton Flavor Universality within it. This is followed by a description of previous measurements of LFU, with a focus on the category of electroweak neutral-current decays. It is in such decays that evidence for a

breaking of LFU has been observed at LHCb and which include the decays studied in this thesis. Chapter 3 describes the Large Hadron Collider and the LHCb experiment, giving context for how the data used in this thesis was obtained and its characteristics. In Chapter 4 the data samples used are described, together with the applied selection criteria. Contributions from background channels are identified and filtered out. Chapter 5 details the analysis consisting of fits on the reconstructed B^0 mass performed on the data. A sensitivity study is performed in Chapter 6 by generating pseudo-experiments using the fit results from the previous chapter. Lastly, Chapter 7 provides a conclusion and summarizes further steps to be performed in the analysis of the studied decays.

---

[1]The inclusion of charge-conjugated particles and decays such as $\bar{B}^0 \to K^- \pi^+ \mu^+ \mu^-$ is implied throughout.

Chapter 2

Theory

2.1 The Standard Model

The Standard Model of particle physics (SM) describes the fundamental particles and their interactions in a unified picture, providing an explanation and powerful predictions for virtually all observed phenomena. The SM includes three fundamental forces, excluding gravity [1, 2].

The forces are described in the framework of a local relativistic quantum field theory. The description of these interactions is remarkably based on the common principle of gauge invariance. The SM is based on the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$, which through spontaneous symmetry breaking is reduced to $SU(3)_C \times U(1)_Q$. This is caused by the non-zero vacuum expectation value of the Higgs field, which breaks the symmetry through a vacuum state that is not left invariant under the symmetry transformation anymore. $SU(2)_L \times U(1)_Y$ describes the unified electroweak interaction, and $SU(3)_C$ the strong interaction. Each of the generators of these groups is associated with a gauge boson (a vector boson, i.e. of spin 1), which mediates the interaction. For $SU(2)_L \times U(1)_Y$ there are four gauge bosons γ, W^+, W^- and Z, out of which only the photon γ, the mediator of the electromagnetic interaction, is massless due to the spontaneous symmetry breaking to $U(1)_Q$. For $SU(3)_C$ there are eight generators, associated with eight massless gluons mediating the strong interaction [1, 2].

Matter is made up of fermionic particles (with spin 1/2), divided into quarks and leptons. Both quarks and leptons are grouped into three generations, with equal quantum numbers but different masses:

$$
\begin{array}{cccc}
 & \text{I} & \text{II} & \text{III} \\
\text{Quarks}: & \begin{pmatrix} u \\ d \end{pmatrix}, & \begin{pmatrix} c \\ s \end{pmatrix}, & \begin{pmatrix} t \\ b \end{pmatrix} \\
\text{Leptons}: & \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, & \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, & \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}
\end{array}
$$

The 'up-type' quarks (u, c, t) have an electrical charge of $+2/3$, and the 'down-type' quarks (d, s, b) one of $-1/3$. The quarks carry the color charge of the strong interaction – each of the quarks comes in three colors. As the weak interaction couples to all of the fermions, the quarks participate in all of the interactions. The leptons are divided into charged (e^-, μ^-, τ^-) and uncharged (ν_e, ν_μ, ν_τ) leptons. The former have electrical charges of -1, the latter are fittingly called neutrinos. They do not carry color charge and the neutrinos thus only participate in the weak interaction, making them hard to detect. All of the fermions have corresponding antiparticles with an opposite electric charge (which are also neutral in the case of neutrinos) [1–3].

Despite the success of the SM, there are several open questions to which it does not provide an answer. These include those of the fundamental nature of dark matter and dark energy, accounting for the majority of matter and energy density in our Universe, respectively. Another unresolved mystery is the observed matter-antimatter asymmetry, for which the required amount of CP violation is far larger than that present in the SM. As the SM does not include gravity, which is described separately by the theory of general relativity, the search for a unified theory of quantum gravity is still ongoing. Further open questions include the structure of the SM, for example why there exist three generations of fermions in particular, and an explanation for the values of the fermion masses [1, 2, 5].

2.2   Lepton Flavor Universality

In the Standard Model, the three charged lepton flavors are identical to each other, aside from their mass. Particularly, the couplings of the gauge bosons mediating the weak interaction are identical to the three lepton flavors. Hence, leptonic decays involving the weak interaction should have equal branching fractions for the different families, aside from phase space differences. This is also known as Lepton Flavor Universality (LFU) [2, 8] .

The fermion fields are organized into the three aforementioned generations, with equal gauge charges. This leads to the same coupling in all generations, also after the spontaneous symmetry breaking through the Higgs mechanism. The only difference between the families is due to the interaction between the Higgs field and fermion fields, called the Yukawa interaction, which gives the fermions their masses. Since the Yukawa couplings to the different families differ, this explicitly breaks the flavor symmetry and results in the families differing in mass.[1] These couplings are free parameters of the Standard Model, which does not explain the values of these different masses. Naturally, this difference in mass does cause a difference in the available phase space available in decays involving the leptons, but this is easily taken into account. Also, these differences often become negligible in the decays studied here involving heavy b-mesons due to the high energies involved [2, 3].

The reason Lepton Flavor Universality is of interest is that it is an 'accidental' symmetry of the SM, which does not provide an explanation for it. Any evidence of this universality being violated would be a sign for New Physics (NP) beyond the SM [3]. Indeed, there are many NP models where LFU is violated. One example is the collection of theories involving leptoquarks, hypothetical particles carrying both baryon and lepton numbers, which can directly interact with a quark and lepton [9]. These leptoquark states are part of various NP theories, including grand unification theories based on larger gauge groups [10]. Another group of theories introduces a new neutral gauge boson, denoted by Z', which couples in a non-universal way to the lepton families [11, 12].

---

[1]The Yukawa interaction has another effect, due to the fact that the mass eigenstates and flavor eigenstates are not equal, i.e. the matrix of Yukawa couplings cannot be diagonalized simultaneously in mass and flavor eigenstates. This results in mixing matrices between the eigenstates, the CKM matrix for quarks and the PMNS matrix for neutrinos [3].

2.3   Tests of LFU

There are various ways to conduct precision tests of LFU. The focus will be on processes involving b decays, as these have shown the largest deviation from LFU, and specifically those involving neutral-current decays, as these include the decays examined in this thesis.

2.3.1   Electroweak sector

Precision measurements of the universality of the Z couplings to the charged leptons have been performed at LEP [13]. The ratios of the leptonic branching fractions are measured as the leptonic partial decay widths, with the following results:

$$\begin{aligned}
\frac{\Gamma_{\mu\mu}}{\Gamma_{ee}} &= \frac{\mathcal{B}(Z \to \mu^+\mu^-)}{\mathcal{B}(Z \to e^+e^-)} = 1.0009 \pm 0.0028 \\
\frac{\Gamma_{\tau\tau}}{\Gamma_{ee}} &= \frac{\mathcal{B}(Z \to \tau^+\tau^-)}{\mathcal{B}(Z \to e^+e^-)} = 1.0019 \pm 0.0032 \,.
\end{aligned} \qquad (2.1)$$

A very good agreement with LFU is observed, as the lepton masses provide only small corrections to the value of unity for the relative widths [13].

Similar measurements have been performed for W-boson decays. Summarized, LFU is found to hold between the first two lepton families with a precision of 0.3% in Z-boson decays and 0.8% in W-boson decays. Between the third and the other two families, this value is also 0.3% in Z-boson decays, but at a larger 3% in W-boson decays, as well as slightly in tension with the SM prediction [3].

2.3.2   Charged-current decays

Another way to test LFU is through semi-leptonic $b \to c\ell^-\bar{\nu}_\ell$ decays, where ℓ^- is any of the three charged leptons. In the Standard Model, these can occur at tree-level, shown in Fig. 2.1. Decays of this type involving electrons and muons, such as $B \to D\mu^-\bar{\nu}_\mu$ and $B \to De^-\bar{\nu}_e$, are consistent with LFU as of date within the experimental uncertainty [3]. However, the large τ mass could make semi-leptonic charged-current B decays more sensitive to NP effects.
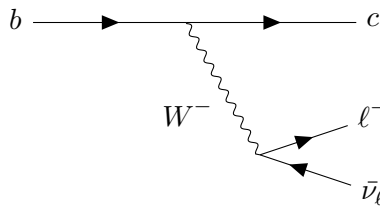
Figure 2.1: The charged current decay $b \to c\ell^-\bar{\nu}_\ell$.

Common observables to probe NP contributions are the ratios of the branching fractions from B decays to D-mesons, with different lepton families in the final states. Specifically, two such ratios that have been studied are

$$R_D = \frac{\mathcal{B}(B^0 \to D^+\tau^-\bar{\nu}_\tau)}{\mathcal{B}(B^0 \to D^+\mu^-\bar{\nu}_\mu)} \quad \text{and} \quad R_{D^*} = \frac{\mathcal{B}(B^0 \to D^{*+}\tau^-\bar{\nu}_\tau)}{\mathcal{B}(B^0 \to D^{*+}\mu^-\bar{\nu}_\mu)} \,. \qquad (2.2)$$

Such quantities are useful, since the quarks involved in both decays are the same, causing the common hadronic form-factors to cancel out. This results in a decrease of the hadronic uncertainties in the theoretical calculation of the value. Experimentally, this also reduces uncertainties from the reconstruction efficiencies [3].

If LFU is violated, then the second and third generation leptons could couple differently to NP processes involved in these decays, which would result in a deviation of such ratios from the SM prediction [8]. Indeed, such deviations from the SM prediction have been observed at the B-factories BaBar and Belle, as well as at LHCb. These results have been averaged by the Heavy Flavor Averaging Group, taking into account the statistical and systematic uncertainties, as well as their correlations [14]. The combined average compared to the SM prediction is depicted in Fig. 2.2.
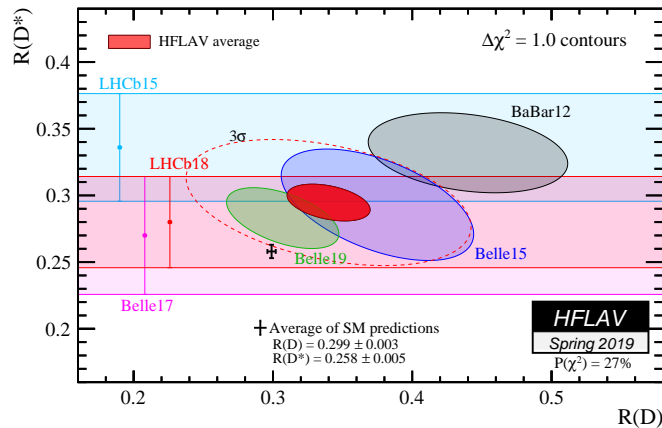

Figure 2.2: The various experimental results on the ratios R_D and R_{D^*}, averaged and compared to the SM prediction. The red ellipse is the average of the values as determined in 2019, and the single data point is the average of the SM predictions. The experimental values and the SM prediction deviate from each other by $3.08\,\sigma$ [14].

The measurement of the decays involving τ leptons is challenging, since these decay quickly before being able to be detected, and always into a state involving a ν_τ which cannot directly be detected. There is also a large background from other partially reconstructed B decays with similar signatures. The measurements at LHCb have used the leptonic $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$ and 3-prong hadronic $\tau^- \to \pi^- \pi^+ \pi^- (\pi^0)\,\nu_\tau$ decays [3, 8]. Due in part to these challenges, charged-current decays have not provided the most sensitive test of NP effects. However, deviations from the SM prediction have still been observed, and further data-taking with the LHC high-luminosity upgrade [15] as well as with Belle II [16], the successor to the Belle experiment, will shed more light on these observations. Furthermore, as mentioned, charged-current decays can serve as a valuable test for NP coupling differently to third-generation leptons.

2.3.3 Neutral-current decays

An ideal way to test LFU is through so called Flavor-Changing Neutral-Current (FCNC) processes such as those involving the transition $b \to s\ell^+\ell^-$, i.e. processes where the flavor of the quark is changed without a change in the electric charge. These FCNC are forbidden at tree-level in the SM, and must proceed through amplitudes involving electroweak loop diagrams, as seen in the left side of Fig. 2.3. This suppression, further increased by the CKM factors involved in transitions between different quark generations, makes them highly sensitive to NP processes.

Such a process involving a scalar leptoquark, occurring at tree-level, is depicted in the right side of Fig. 2.3. In fact, depending on the model, the sensitivity for NP extends up to ranges of $50\,\mathrm{TeV}$, beyond the scale one can directly probe today [3, 11].

Figure 2.3: Flavor-Changing Neutral Current process $b \to s\ell^+\ell^-$ in the SM (left) and a NP model involving a scalar leptoquark (right) [9]. In the SM, this necessitates a suppressed loop diagram, while the leptoquark would enable such a decay at tree-level.

Analogous to the R_D ratio in charged-current decays, a convenient observable to study is the ratio of branching fractions with different lepton families in the final state:

$$R_{H_s} = \frac{\mathcal{B}(H_b \to H_s \mu^+ \mu^-)}{\mathcal{B}(H_b \to H_s e^+ e^-)} \, , \tag{2.3}$$

where H_b and H_s denote hadrons containing b- and s-quarks, respectively [3]. Most previous measurements of such ratios, including the ones showing the largest deviation from the SM prediction, have looked at B-mesons decaying to kaons, where the kaon can be in an excited state ($H_b = B$ and $H_s = K^{(*)}$). The SM predicts $R_{K^{(*)}} = 1$ with theoretical uncertainties on the order of $1\,\%$ [17]. Again, hadronic uncertainties are essentially canceled out [11, 17]. A difficulty with measuring this ratio is the different behavior of electrons and muons in the detector, especially in terms of detection efficiency. These differences in the context of the LHCb detector are detailed further in Sect. 3.2.2. Such experimental systematic uncertainties can be reduced by measuring $R_{K^{(*)}}$ as a double ratio, where the branching ratios are normalized to those of the resonant J/ψ mode:

$$R_{K^{(*)}} = \frac{\mathcal{B}(B \to K^{(*)}\mu^+\mu^-)}{\mathcal{B}(B \to K^{(*)}J/\psi(\to \mu^+\mu^-))} \, / \, \frac{\mathcal{B}(B \to K^{(*)}e^+e^-)}{\mathcal{B}(B \to K^{(*)}J/\psi(\to e^+e^-))} \, . \tag{2.4}$$

This double ratio can be analyzed instead of the single ratio since the resonant $J/\psi \to \ell^+\ell^-$ decays are found to be consistent with LFU, i.e. the ratio $\mathcal{B}(J/\psi \to e^+e^-) \, / \, \mathcal{B}(J/\psi \to \mu^+\mu^-)$ is found to be consistent with unity [18, 19]. Since the experimental signatures of the resonant and non-resonant decays are similar, systematic effects are reduced so that in the end, the uncertainty on the ratio is dominated by the statistical uncertainty [18]. The ratio of this resonant channel also gives a good cross-check, verifying that the difference in experimental signatures for muons and electrons are well understood and accounted for [5].

The ratio denoted by R_{K^*} of $B^0 \to K^{*0}\ell^+\ell^-$ decays has been analyzed by the LHCb collaboration, where K^{*0} is the $K_1^*(892)^0$ resonance, detected in the $K^{*0} \to K^+\pi^-$ decay mode. The most recently reported result from 2017 shows a tension of up to $2.5\,\sigma$ (standard deviations) with the SM prediction [7]. Recently, the LHCb collaboration presented results (in preprint) on R_K using the full data collected up to 2018, analyzing branching fractions of the decays $B^+ \to K^+\ell^+\ell^-$. The ratio of branching fractions was measured to be $R_K = 0.846^{+0.042+0.013}_{-0.039-0.012}$, where the first uncertainty is statistical and the second systematic, deviating from the SM prediction by $3.1\,\sigma$

[4]. This value is the same as previously published in Ref. [18] with a smaller dataset, but with reduced uncertainties, crossing the 3σ threshold for 'evidence'. An overview of results on R_K and R_{K^*} by LHCb and the B-factories BaBar and Belle is shown in Fig. 2.4. Furthermore, a test of LFU has been performed with $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays, with a ratio between the electron and muon channels of $R_{pK}^{-1} = 1.17_{-0.16}^{+0.18} \pm 0.07$. While this ratio is still consistent with the SM prediction, this is the first test of LFU performed with b-baryons [20].
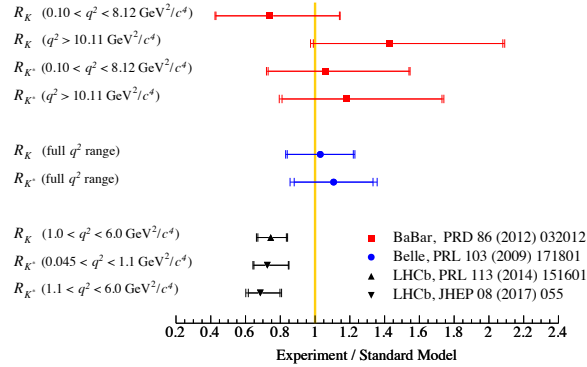

Figure 2.4: An overview of the different results on $R_{K^{(*)}}$ as measured by BaBar, Belle and LHCb. The LHCb value for R_K has been updated to the one mentioned above. The LHCb measurements are the most precise and show the largest discrepancies to date [3].

In addition to modifying the branching fractions, contributions from NP can also change the angular distribution of the detected particles in the final state. Such relevant angular observables have been studied in $b \to s\mu^+\mu^-$ decays and, similarly to the branching fraction ratios, show a tension of around 3σ with respect to the SM prediction [21].

2.3.4 Outlook

As has been seen, there are tensions with the LFU prediction of the SM observed in multiple decays. A natural explanation for the fact that such anomalies have been observed only in semi-leptonic B decays and not in electroweak precision tests or semi-leptonic K and π decays is that the NP responsible for breaking LFU is mainly coupled to the third generation of quarks and leptons. Indeed, there are leptoquark theories that are quite simple and effective at explaining the observed anomalies to date [22].

While the observations look promising, the observed tensions with LFU are not sufficient as to claim a discovery yet, set at a deviation of 5σ. For one, additional statistics are required, which will be obtained by further analysis of the LHC Run 2 data, the upcoming high-luminosity upgrade as well as Belle II, the successor to the Belle experiment. Secondly, if this effect really is a sign of New Physics, it must be visible in certain other decays not yet examined. In this thesis, the decays $B^0 \to K^+\pi^-\ell^+\ell^-$ with $\ell = e, \mu$ are studied. This is part of the attempt to study and measure the ratio

$$R_{K\pi} = \frac{\mathcal{B}(B^0 \to K^+\pi^-\mu^+\mu^-)}{\mathcal{B}(B^0 \to K^+\pi^- J/\psi(\to \mu^+\mu^-))} \Bigg/ \frac{\mathcal{B}(B^0 \to K^+\pi^- e^+ e^-)}{\mathcal{B}(B^0 \to K^+\pi^- J/\psi(\to e^+ e^-))} \qquad (2.5)$$

analogous to R_{K^*} but above the $K_1^*(892)$ resonance, which is again measured as a double ratio.

Chapter 3

The LHCb experiment

3.1   The LHC

The Large Hadron Collider (LHC) is a hadron accelerator and collider, with a length of 26.7 km. It is installed in the tunnel that was previously used for the LEP experiment at CERN, the European Organization for Nuclear Research. The aim of the LHC is to search for physics beyond the Standard Model, through various direct and indirect searches. Another design goal was to find direct evidence for the last missing piece of the Standard Model, the Higgs boson, which was successfully discovered by experiments at the LHC in 2012 [23, 24].
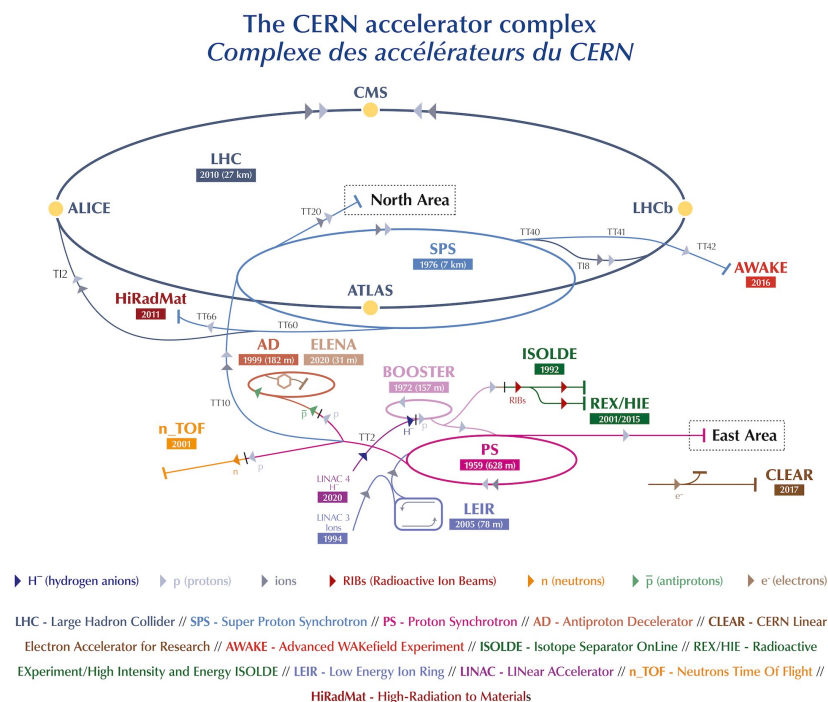
Figure 3.1: The CERN accelerator complex. At the top, the large LHC ring can be seen, with the four large experiments at the collision points [25].

The LHC is a particle-particle collider, designed to, in its main operation mode, collide protons with each other at center-of-mass energies up to 14 TeV. Its requirement for a high luminosity necessitates the use of p-p collisions as opposed to p-$\bar{p}$ collisions.[1] This calls for a a different accelerator design compared to particle-antiparticle colliders, where a common pipe and magnetic field can be used for both beams which have opposite charges and directions, while the two proton

beams at the LHC have separate magnetic fields and vacuum chambers. To save space and costs however, the magnets use a complex 'two-in-one' design, with a shared cryostat. This magnet system is one of the largest challenges and limiting factors at the LHC, using superconducting magnets cooled to below $2\,\mathrm{K}$ with liquid helium, at fields above $8\,\mathrm{T}$ [23].

The protons are accelerated and brought to collision at the interaction points in bunches, with a nominal bunch spacing of $25\,\mathrm{ns}$, corresponding to a frequency of $40\,\mathrm{MHz}$. Recording all of the interactions at the LHC would produce a data amount of one petabyte every second, an unfeasible amount to store or analyze. However, the interesting events are rare, and the experiments have trigger systems filtering these out to be kept for analysis. Even this reduction of many orders of magnitude corresponds to a data amount of $15\,\mathrm{PB}$ per year, needing a large computing grid across many countries to process [23, 24].

There are four interaction points, where the experiments ATLAS, CMS, ALICE and LHCb are installed. The ATLAS and CMS experiments are both high-luminosity, general purpose experiments with a wide physics program, designed for a peak luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ during the nominal proton operation [23].

The coverages of detectors are usually described with the azimuthal angle ϕ measured around the beam axis, and the polar angle θ relative to the beam axis. One rescales the polar angle by defining the pseudorapidity as $\eta = -\ln\tan(\theta/2)$. The ATLAS and CMS experiments have a toroidal structure around the interaction point, designed to have an almost hermetic coverage, covering essentially the full range of the azimuthal angle ϕ and pseudorapidities of around $|\eta| < 3.0$ [26, 27].[2]

The LHCb experiment, detailed in Sect. 3.2, is a dedicated B-physics experiment operating at a lower peak luminosity of $10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The LHC is at times also operated with heavy ion beams (such as lead), with the dedicated ion experiment ALICE studying strongly interacting matter and the quark-gluon plasma, operating at a peak luminosity of $10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [28].

In the next few years, the LHC will be upgraded to the High-Luminosity LHC, allowing an increased peak luminosity of $5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and a tenfold increase of the integrated luminosity compared to the LHC [15].

3.2   The LHCb detector

The LHCb experiment is a dedicated heavy flavor physics experiment at the LHC, schematically shown in Fig. 3.2. It is a single-arm forward spectrometer, whose shape is motivated by the angular distribution of the b- and $\bar{b}$-hadrons produced at the interaction point. This distribution is shown in Fig. 3.3 in terms of the polar angle relative to the beam axis. It is highly non-isotropic, with the hadrons mostly produced in small cones in the forwards and backwards direction. The LHCb detector is then built to cover one of these directions, in the pseudorapidity range $2 < \eta < 5$ [29].

---

[1]The luminosity $\mathcal{L}$ describes the event rates, where for a given process, the number of of interactions is $N = \sigma \int \mathcal{L}\,\mathrm{d}t = \sigma\,L$, where σ denotes the cross section for that process given by the dynamics, and L the integrated luminosity. They are often specified in *barns* (b) and inverse barns, respectively, where $1\,\mathrm{b} = 1 \times 10^{-28}\,\mathrm{m}^2$ [1].

[2]The pseudorapidity coverages vary between the various tracking and caliometry systems, which are not covered in detail here.

The LHCb experiment's main goal is the search for indirect evidence of New Physics, in processes involving CP violation and rare decays of b- and c-hadrons. The single source of CP violation in the Standard Model cannot explain the observed matter-antimatter asymmetry in the universe. Hence, new sources of CP violation are needed, evidence of which might be seen in heavy flavor physics. As discussed in the previous chapter, studying rare decays provides a sensitive way to probe NP contributions, which can change the branching fractions, angular observables or introduce new sources of CP violation [6].
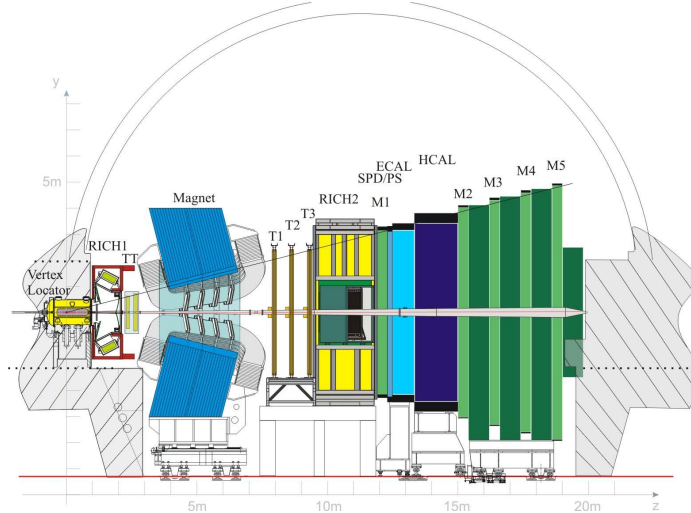

Figure 3.2: A schematic view of the LHCb detector [30].

Compared to e^+e^- B factories, which collide electrons and positrons at the $\Upsilon(4S)$ resonance, decaying into a $B\bar{B}$ pair, the LHCb experiment has various advantages. These include a higher cross section, a larger boost of the hadrons as well as a higher number of b-hadron species present in the decays. However, this comes at the cost of higher background levels, for example worsening the ability to reconstruct final states with missing or neutral particles [6]. Since the colliding particles (protons) are composite particles, the center-of-mass (c.o.m.) energy is not well known and the $b\bar{b}$ pairs are produced with a broad spectrum of energies. Thus, the invariant mass of the undetected particles in an event is not known. However, the direction of flight can be exploited instead to estimate the momentum, requiring a good resolution of the vertex detector [3].

3.2.1 Particle tracking and reconstruction

The momentum of charged particles can be measured by their bending in a magnetic field due to the Lorentz force acting perpendicularly to their direction of travel. To this end, the LHCb uses a warm dipole magnet with an integrated field of about $4\,\mathrm{T\,m}$. This magnet can be inverted to reduce systematic effects on CP-violating observables [29]. The tracking system consists of different sub-detectors, measuring the positions of particles as they pass through in the forwards direction. This allows a reconstruction of the track and the corresponding bending radius.

Vertex Locator

The Vertex Locator (VELO) is closest to the interaction point at a distance of $8\,\mathrm{mm}$ from the beam axis, and located inside a vacuum tank. It consists of silicon modules measuring the distance r and the azimuthal angle ϕ of charged particles as they excite charges in the sensors when passing through.
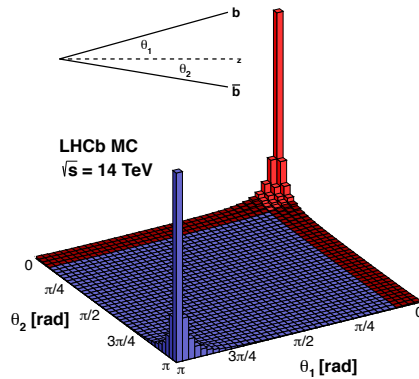
Figure 3.3: The polar angles of b- and $\bar{b}$-hadrons produced at a c.o.m. energy of $14\,\text{TeV}$ relative to the beam axis [31].

These precise coordinate measurements close to the interaction region are important for c- and b-hadron decays, which have secondary decay vertices close to the primary vertex. Identifying these vertices with a high precision is necessary to achieve a good resolution of the decay lifetime and direction, allowing the measurement of flavor oscillations and the momentum of the hadron, as well as the rejection of various background sources.

Trackers

Next to the VELO, the tracking system consists of four further planar tracking stations. The Tracker Turicensis (TT) is placed upstream of the dipole magnet, and the tracking stations T1-T3 downstream of it. The downstream tracking stations are divided into two parts, the Inner Trackers (IT) and Outer Trackers (OT).

The TT and IT both use silicon microstrip sensors, like the VELO. The TT is $150\,\text{cm}$ wide and $130\,\text{cm}$ high with a total active area of $8.4\,\text{m}^2$ and covers the full acceptance of the experiment. It is important for the reconstruction of so-called upstream tracks of low-momentum particles with a momentum below $1.5\,\text{GeV}/c$, which are bent out of the acceptance of the detector by the magnet before reaching the other tracking stations. The IT is $120\,\text{cm}$ wide and $40\,\text{cm}$ wide with an active area of $4.0\,\text{m}^2$. It covers the central area of the downstream tracking stations with a higher particle flux. Both the TT and IT have a single-hit resolution of $50\,\mu\text{m}$.

The OT is a drift-tube gas detector with a total of about $55\,000$ straw-tube channels, each with an inner diameter of $4.9\,\text{mm}$. It covers the outer regions with a lower particle flux, with occupancies not exceeding $10\,\%$ at a luminosity of $2 \times 10^{32}\,\text{cm}^{-2}\,\text{s}^{-1}$. As charged particles traverse the tubes they will ionize the gas, which is measured by the drift of the electrons to the central wire. The tubes are filled with a mixture of Argon and CO_2, giving a fast drift time of below $50\,\text{ns}$. The drift-coordinate resolution of $200\,\mu\text{m}$ is larger than the resolution of the silicon sensors, but sufficient. An event with the recorded hits in all of the trackers and the reconstructed tracks is shown in Fig. 3.4 [6, 29].

3.2.2 Particle identification

Particle identification (PID) is essential for the goals of LHCb. For instance, the reconstruction and separation of π^0 and photons is essential for the study B-meson decays, as well as the
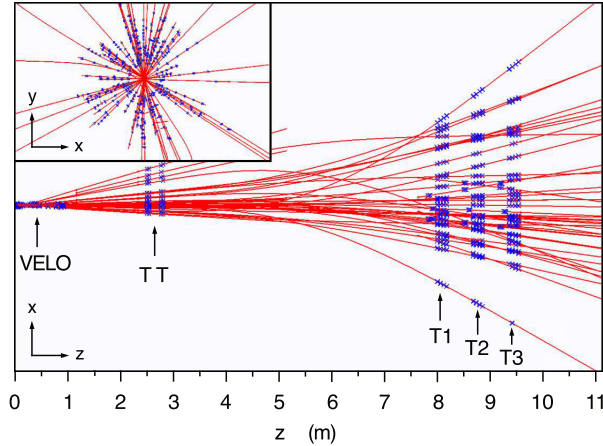
Figure 3.4: An event at the LHCb, with the assigned hits in the different tracking sub-detectors in blue and reconstructed tracks in red. The insert in the top left shows a magnified region at the VELO [29].

separation of pions and kaons. There are multiple sub-detectors that contribute to the PID of specific particles, whose information can in certain cases be combined for an optimal performance [29].

RICH detectors

Two Ring Imaging Cherenkov detectors (RICH) are used to identify charged particles, in particular hadrons (π, K, p), covering the full momentum range. They can also complement the other identification systems in the case of electrons and muons. The detector upstream of the dipole magnet (RICH1) provides identification for particles in the momentum range 1–60 GeV/c, while the detector downstream (RICH2) covers the higher momentum range 15–100 GeV/c.

The RICH detectors make use of the fact that in a medium with a refractive index n, the velocity of light is reduced to c/n, where c denotes the velocity of light in a vacuum. Particles traveling faster than this velocity emit a cone of electromagnetic radiation with a characteristic angle of $\cos\theta_{\mathrm{c}} = (n\beta)^{-1}$, where $\beta = v/c$. By measuring this angle for a given refractive index, one can determine a particle's velocity. If its momentum has been measured, this gives the particle's mass through $p = \gamma m v$, and therefore its species [1].

The RICH detectors focus the Cherenkov light with a combination of spherical and flat mirrors, after which it is detected with Hybrid Photon Detectors, allowing a determination of the emittance angle. The material in the RICH detectors must be chosen so that the particles of interest are fast enough to emit Cherenkov light. However, as the refractive index is increased, distinguishing particles of different masses becomes more challenging. The Cherenkov angle θ_{c} as a function of momentum is displayed in Fig. 3.5 for the materials used in the LHCb detector. RICH1 contains aerogel and fluorobutane (C_4F_{10}) and covers the lower momentum range, while RICH2 contains CF_4 gas, covering the high-momentum range detailed above [29].

Calorimeters

The calorimeter system measures the energies deposited by electrons, photons and hadrons together with their location, while also providing particle identification. In particular, it provides the main way of identifying neutral particles. It consists of an electromagnetic calorimeter (ECAL) followed by a hadronic calorimeter (HCAL).
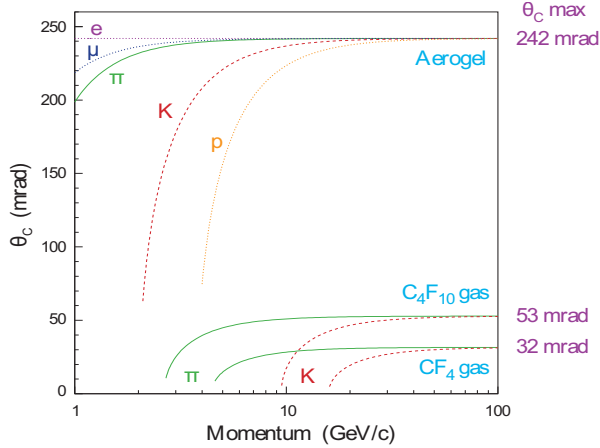
Figure 3.5: The Cherenkov angle as a function of momentum for different particle species and materials used in the LHCb RICH detectors. One sees that aerogel provides a good distinction between the hadrons for low momenta, while CF_4 does so for high momenta, motivating their uses in RICH1 and RICH2, respectively [29].

The calorimeters detect particles using the same principle: the traversing particles excite scintillation light out of the active calorimeter material, which is transmitted by fibers and recorded by photomultiplier tubes. To measure the energy of the incoming particles, calorimeters are designed to completely absorb them. The layers of scintillating material are interleaved with a passive absorber material, causing the particles to interact and produce secondary particle showers. The ECAL and HCAL use lead and iron as their passive absorber material, respectively. The particles from these showers are detected in the active layers, and can produce further showers until the secondary particle energy is low enough to be absorbed. The size of the shower allows for a determination of the initial energy through the amount of scintillation light recorded, while the shape of the shower can provide PID information.

There are two additional sub-detectors, the Preshower Detector (PS) and the Scintillator Pad Detector (SPD), both consisting of scintillator and lead layers. They provide additional information for particle identification, placed before the ECAL. The PS helps to reduce the high background of charged pions, while the SPD assists in the separation of electrons and photons.

Photons and neutral pions are detected in the ECAL, where a distinction is made via the shape of the cluster; the $\pi^0 \to \gamma\gamma$ decay is detected either as two separated clusters or one merged cluster. Electrons, which are important for the measurement of semi-leptonic decays, pose the most demanding particle identification. They are mainly identified based on the balance of momentum associated with the track and energy deposited in the ECAL. Additionally, their energy resolution is worsened by the emission of bremsstrahlung. There are algorithms in place to correct for this, by associating photon clusters in the ECAL with electron tracks. If a photon is emitted upstream of the magnet, this results in a separate cluster in the ECAL, which can be extrapolated to the electron track. An emission downstream of the magnet will not lead to a separate cluster. This correction is illustrated in Fig. 3.6 [6, 29].

Muon system

The muon system provides identification for muons and information for the trigger system for the selection of high-p_T muons. An efficient offline muon identification is also important for the tagging and reconstruction of muonic B decays, many of which are CP-sensitive.
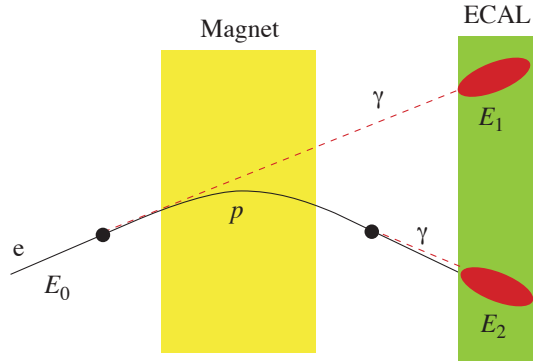
14

Figure 3.6: An illustration of the bremsstrahlung correction for electrons, where photons emitted upstream of the magnet resulting in a separate cluster can be associated with the electron track [29].

The muon system consists of five stations (M1–M5) along the beam axis, with M1 placed upstream and M2 to M5 downstream of the calorimeters. The total detector area is $435\,\mathrm{m}^2$. M1's placement serves to improve the p_T resolution, which is important for the trigger system. It also helps to reduce the background from electromagnetic showers due to late interactions in the calorimeters. M2 to M5 are interleaved with iron absorbers to select high-energy muons; a muon must have a minimum momentum of $6\,\mathrm{GeV}/c$ to traverse all five stations [29, 32].

3.2.3 Trigger

As mentioned previously, the proton bunches cross at a rate of $40\,\mathrm{MHz}$ at the interaction point. Due to the constraints on offline computing power and storage, only a fraction of these interactions can be kept for analysis, which is decided by the LHCb trigger system. It consists of two stages, the Level-0 trigger (L0) and the High Level Trigger (HLT).

The L0 trigger reduces the initial bunch crossing rate of $40\,\mathrm{MHz}$ to a rate of $1\,\mathrm{MHz}$ at which the entire detector can actually be read out. It is fully synchronous with this bunch crossing rate, and is implemented in custom hardware (various types of Field Programmable Gate Arrays) to allow for sufficiently fast processing – for each crossing, the data must be processed in $2\,\mathrm{\mu s}$. The L0 trigger receives partial detector information from three different sources which can be read out sufficiently fast. The pile-up system uses silicon sensors to provide information on the primary vertices and track multiplicities, allowing a determination of the luminosity. The calorimeter trigger PID information from the calorimeters and selects those with high E_T. The L0 muon trigger takes data from the muon chambers, determining p_T with a resolution of about $20\,\%$, and selects the two muons with the highest p_T in each quadrant. The L0 decision unit then collects the information from these sources, and uses simple logic to make a boolean decision per bunch crossing.

The HLT is a software application which is run on the event filter farm in $29\,500$ separate instances. It further reduces the event rate from $1\,\mathrm{MHz}$ down to $2\,\mathrm{kHz}$, and can make use of the full event data from those events selected by the L0 trigger. The HLT is itself divided into two stages: in the first one (HLT1), a partial event reconstruction is performed, which reduces the rate down to $30\,\mathrm{kHz}$, allowing the second stage (HLT2) to perform a full pattern recognition on the remaining events. The reconstruction algorithms are similar to those employed offline, with some simplifications to be less CPU intensive. The resulting output rate of $2\,\mathrm{kHz}$ is stored for offline analysis. Since the HLT is implemented in software, it can readily be adapted to changes in conditions and physics goals, as well as improvements in event reconstruction algorithms.

15

Chapter 4

Dataset

The data studied was collected at LHCb during 2016 at a center-of-mass energy of $13\,\mathrm{TeV}$ with an integrated luminosity of $1.6\,\mathrm{fb}^{-1}$. The relevant decays are

- $B^0 \to (K^+\pi^-)^*_J\, \mu^+\mu^-$

- $B^0 \to (K^+\pi^-)^*_J\, e^+e^-$

- $B^0 \to K^{*0}e^+e^-$

- $B^0 \to K^{*0}\mu^+\mu^-$

where the first two are the signal decays studied in this thesis, and the other two are used as a reference. Additionally, Monte-Carlo samples of the following decays are used in the analysis:

- $B^0 \to (K^+\pi^-)^*_J\, \mu^+\mu^-$

- $B^0 \to K^{*0}e^+e^-$

These are generated without any dynamic effects of the different kaon resonances, but simply with kinematic phase space considerations. The latter is used to obtain the $B^0 \to (K^+\pi^-)^*_J\, e^+e^-$ signal shape, as no simulation was available for this decay.

As the resonance $K^*_1(892)^0$ (denoted by K^{*0} as done previously) is detected and reconstructed through its decay products $K^{*0} \to K^+\pi^-$, the final state particles are the same in the $B^0 \to K^{*0}\ell^+\ell^-$ and $B^0 \to (K^+\pi^-)^*_J\ell^+\ell^-$ decays, where $(K^+\pi^-)^*_J$ denotes that the kaon and pion include contributions from various excited states K^*_J with spin J (for simplicity, the decays are often labeled by $B^0 \to K^+\pi^-\ell^+\ell^-$ in the following). The difference between the decays lies in the invariant mass $m(K^+\pi^-)$, which is required to be within $100\,\mathrm{MeV}/c^2$ of the $K^*_1(892)^0$ mass for said resonance, i.e. in the range 792–992 MeV/c^2.[1]

For the $B^0 \to (K^+\pi^-)^*_J\ell^+\ell^-$ decays studied here, two regions of the invariant mass $m(K^+\pi^-)$ above the K^{*0} resonance are considered separately, for both of the lepton final states: 1000–1835 MeV/c^2 and 1895–2600 MeV/c^2. They will also be denoted by the 'lower' and 'upper' region, respectively. These are located below and above the excluded resonance $D^0 \to K^+\pi^-$ ($m_{D^0} = 1865\,\mathrm{MeV}/c^2$ [10]). This separation is done to study any systematic differences between the regions, as well as a comparison of the number of observed signal events and sensitivities.

---

[1]The invariant mass m_0 of a system of particles is the Lorentz-invariant quantity which can be computed through the relativistic energy-momentum relation

$$\left(m_0 c^2\right)^2 = \left(\sum_i E_i\right)^2 - \left\|\sum_i \boldsymbol{p}_i c\right\|^2 , \tag{4.1}$$

where the sums go over the energies and momenta of the individual particles. If the particles are decay products, the invariant mass gives the mass of the decaying particle [1].

The spectrum of the various K_J^* states is depicted in Fig. 4.1. One sees that the $m(K^+\pi^-)$ regions considered here contain a multitude of these states. As these are the same for the two lepton final states which enter into the $R_{K\pi}$ ratio, such an inclusive measurement can be performed without needing to separate these states .
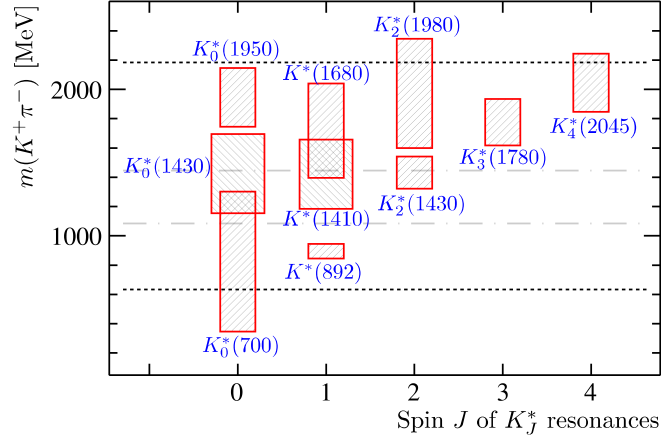

Figure 4.1: The spectrum of K_J^* states, arranged by their spin J and the invariant mass of their decay products $K^+\pi^-$, where the height of the boxes indicates the width of the resonance [33].

The lower region includes the $K_{0,2}^*(1430)$ resonances, which have been studied for the muon channel $B^0 \to K^+\pi^-\mu^+\mu^-$ in Ref. [34]. The observed distribution of events as a function of $m(K^+\pi^-)$ is shown in Fig. 4.2. The sharp peak of the K^{*0} resonance around $892\,\mathrm{MeV}/c^2$ is visible, with a flattened spectrum at higher masses and a smaller peak at the studied region. The electron channel has not been observed so far.
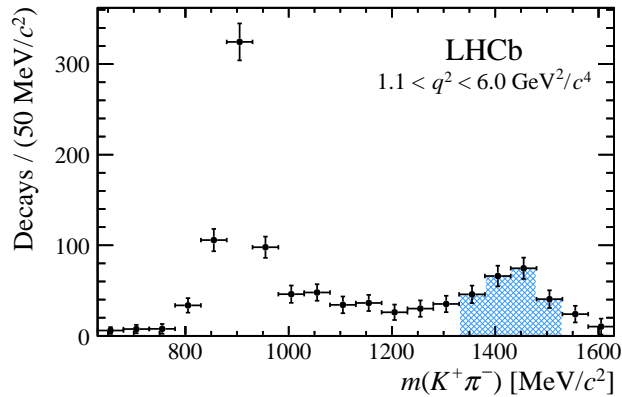
Figure 4.2: The invariant mass $m(K^+\pi^-)$ in $B^0 \to K^+\pi^-\mu^+\mu^-$ events previously measured at LHCb. The highlighted region shows the analyzed events at the $K_2^*(1430)$ resonance [34].

Since the particles in the final state are the same as those in the $B^0 \to K^{*0}\ell^+\ell^-$ decays studied in the analysis of $R_{K^{*0}}$, many aspects in terms of the selection criteria applied will be shared here. However, the difference in the invariant mass $m(K^+\pi^-)$ has an effect on certain exclusive backgrounds, as will be seen below.

4.1 Preselection

4.1.1 Combinatorial background

A large source of background is the combinatorial background, arising from the random association of particles produced in different decays, matching the signature of the signal decay. As a function of the invariant mass of the final state particles (i.e. the reconstructed B^0-mass), the number of combinatorial background events has an exponential shape, and will be modeled as such when performing the invariant mass fits.

This combinatorial background is reduced through the use of a multivariate analysis, using boosted decision trees (BDTs). These map each event to a continuous response, scoring it higher if it is more probable to be signal. Typically, a cut is placed on this BDT response, removing all events below a certain value. This value is chosen by maximizing a figure of merit, e.g. the statistical significance $S/\sqrt{S+B}$ where S and B denote the number of signal and background events, respectively [35].

The classifier used here originates from the R_{K^*} analysis and is trained on simulated $B^0 \to K^{*0}\ell^+\ell^-$ events (corrected for differences between data and simulation) as well as data from the right-side band with a reconstructed mass above $5450\,\mathrm{MeV}/c^2$ for the background. Other network architectures such as Bayesian neural networks have been explored by another student and were found to offer a similar performance, however at the expense of a more difficult training procedure. It is possible that the performance may slightly improve by training on $B^0 \to K^+\pi^-\ell^+\ell^-$ events, however the most important features used by the classifier are not directly related to the kinematics, but to the fits to the tracks and primary vertex of the B^0.

The classifier uses *XGBoost* as the framework for the BDTs, and assigns a response denoted by `XGBUnbiasedLabelScore`. Previously, a cut of `XGBUnbiasedLabelScore` > 0.99 has been used on $B^0 \to K^{*0}\ell^+\ell^-$ events. For $B^0 \to K^+\pi^-\ell^+\ell^-$ events, the background levels were found to be substantially higher for the same cut. While the efficiency on signal and background events may be similar, this can be explained by the fact that in the two $m(K^+\pi^-)$ regions considered here, the available phase space for combinatorial background events is larger, increasing the number of background events compared to signal events. Thus, a higher, reasonable but arbitrary cut of `XGBUnbiasedLabelScore` > 0.999 was chosen for this analysis. In the future analysis of $R_{K\pi}$, the optimal value for the cut will have to be determined, optimizing between the background rejection and signal efficiency for a maximized significance.

4.1.2 Particle identification

The particle identification is accomplished with a neural network approach. The information from the various sub-detectors providing particle identification (as detailed in Sect. 3.2.2) is taken as input, and `ProbNN` parameters are assigned to the detected particles. These lie between 0 and 1 and can be viewed as probabilities, e.g. `ProbNN`$_K$ is the probability that the particle associated with a certain track is a kaon. Simple cuts can be applied on these variables, but cuts on the combinations of the form `ProbNN`$_K \cdot (1 - $`ProbNN`$_\pi)$ have found to be effective, in this case requiring a certain threshold for the track being associated with a kaon and not a pion.

The following requirements are applied:

- For kaon candidates: `ProbNN`$_K \cdot (1 - $`ProbNN`$_p) > 0.05$

- For pion candidates: `ProbNN`$_\pi \cdot (1 - $`ProbNN`$_K) \cdot (1 - $`ProbNN`$_p) > 0.1$

- For electron candidates: `ProbNN`$_e$ > 0.2

- For all final state particles: `GhostProb` < 0.4

The last requirement attempts to remove 'ghost' tracks, which are misidentified tracks from random combinations of hits. They are characterized by a bad quality of the track fit and missing hits [36].

Furthermore, there are requirements on the momentum and its transverse component:

- For kaon and pion candidates: $p_{\mathrm{T}} > 250 \,\mathrm{MeV}/c$

- For electron candidates: $p_{\mathrm{T}} > 500 \,\mathrm{MeV}/c$ & $p > 3000 \,\mathrm{MeV}/c$

This improves the particle identification as this is the range covered by the data used for calibration of the PID system.

4.1.3   Signal selection

A histogram of the candidate events in terms of the invariant mass of the final state particles and the dilepton mass-squared $m(\ell^+\ell^-)^2 \equiv q^2$ is shown in Fig. 4.3. For both the muon and electron channels, the horizontal charmonium resonances at a fixed mass-squared q^2 are visible, originating from J/ψ and $\psi(2S)$ decays at masses of $3.097 \,\mathrm{GeV}/c^2$ and $3.686 \,\mathrm{GeV}/c^2$, respectively [10].

For the electrons, these structures are smeared out due to their bremsstrahlung emission. While the previously mentioned bremsstrahlung correction is applied, this effect is limited due to the resolution of the ECAL [7]. Visually, these distributions are quite similar to those observed for $B^0 \to K^{*0}\ell^+\ell^-$ decays in Ref. [7].
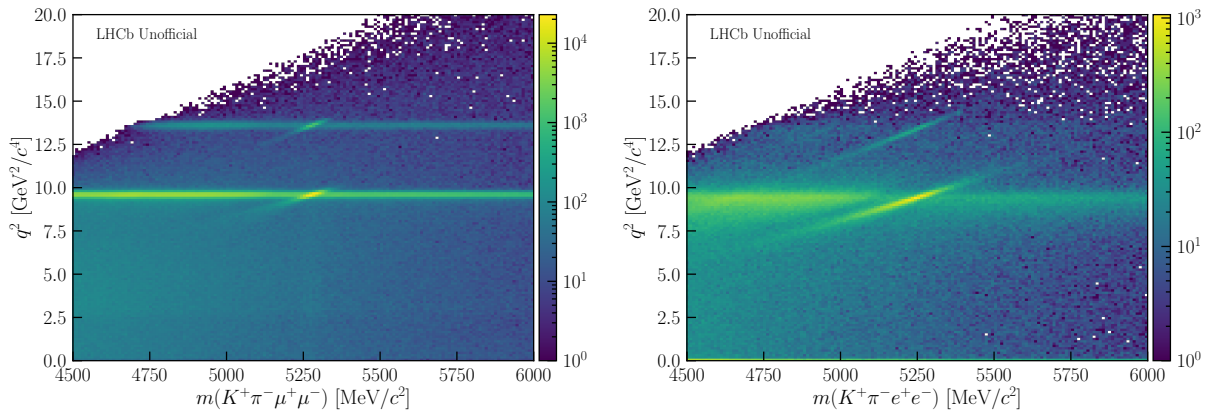

Figure 4.3: The number of $B^0 \to K^+\pi^-\ell^+\ell^-$ candidate events, for muon (left) and electron (right) final states, in the combined lower and upper regions. They are displayed as a function of the dilepton mass-squared q^2 and the invariant mass of the final state particles, associated with a B^0. The particle identification criteria have been applied, but not the BDT cut for the combinatorial background.

To select the rare decay mode, directly decaying to the lepton pair via the electroweak loop diagram (as seen in Fig. 2.3), the following cut on the dilepton mass-squared range is applied: $1.1 < q^2 < 6.0 \,\mathrm{GeV}^2/c^4$. For the electron final states, this quantity is computed with constraints to the B^0-mass and primary vertex, which improves the resolution. The resonant J/ψ channel

will need to be analyzed for the determination of the double ratio $R_{K\pi}$, but will not be studied here.

The studied q^2 region is the same 'central' region which has been used in the analyses of R_K and R_{K^*}. The ratio R_{K^*} has separately been measured in the lower q^2 region $0.045 < q^2 < 1.1\,\mathrm{GeV}^2/c^4$ as well, where the lower limit corresponds to the kinematic threshold of the two muons [7]. This region could also be considered in the analysis of $R_{K\pi}$, but this was not explored here. Furthermore, with the mentioned constraints on q^2 one could explore extending the upper limit to $7.0\,\mathrm{GeV}^2/c^4$ while still avoiding leakage from the resonant J/ψ channel, but this was also not done for this analysis.

4.1.4 Cloning

Pairs of tracks are defined as 'clones' if more than $70\,\%$ of hits and parts of the T-stations are shared. As the tracking pattern recognition contains a series of algorithms which 'build' different types of tracks, such cloned tracks may be produced at the end of the sequence. The Clone Killer algorithm run at the end of the track reconstruction attempts to reduce these, but some of these cloned tracks survive [37]. This is particularly the case for electrons, whose track quality is worsened by the emission of bremsstrahlung.

Cloned tracks are characterized by very small opening angles between them. A commonly used and simple method to remove such events is to require a minimum angle between the various tracks. An efficient cut that has been used, retaining more than $99\,\%$ of simulated signal events, is a minimum angle of $0.5\,\mathrm{mrad}$.

The opening angles are calculated with the momentum vectors associated to the tracks (before the bending by the dipole magnet):

$$\cos\theta_{12} = \frac{\boldsymbol{p}_1 \cdot \boldsymbol{p}_2}{|\boldsymbol{p}_1||\boldsymbol{p}_2|}\,. \tag{4.2}$$

The opening angles between the tracks in both the candidate events of both decays are shown in Figs. 4.4 and 4.5 (note the logarithmic scale). For the electron final states, a peak can be seen for very small angles in all cases, except the angles between the two hadrons and the two leptons. This is indeed removed everywhere by requiring $\theta > 0.5\,\mathrm{mrad}$, cutting more than half of the candidate events. For the muons, with much cleaner tracks, one sees that the cloning is less of an issue. The same cuts on the opening angles are applied, removing a few percent of the events.
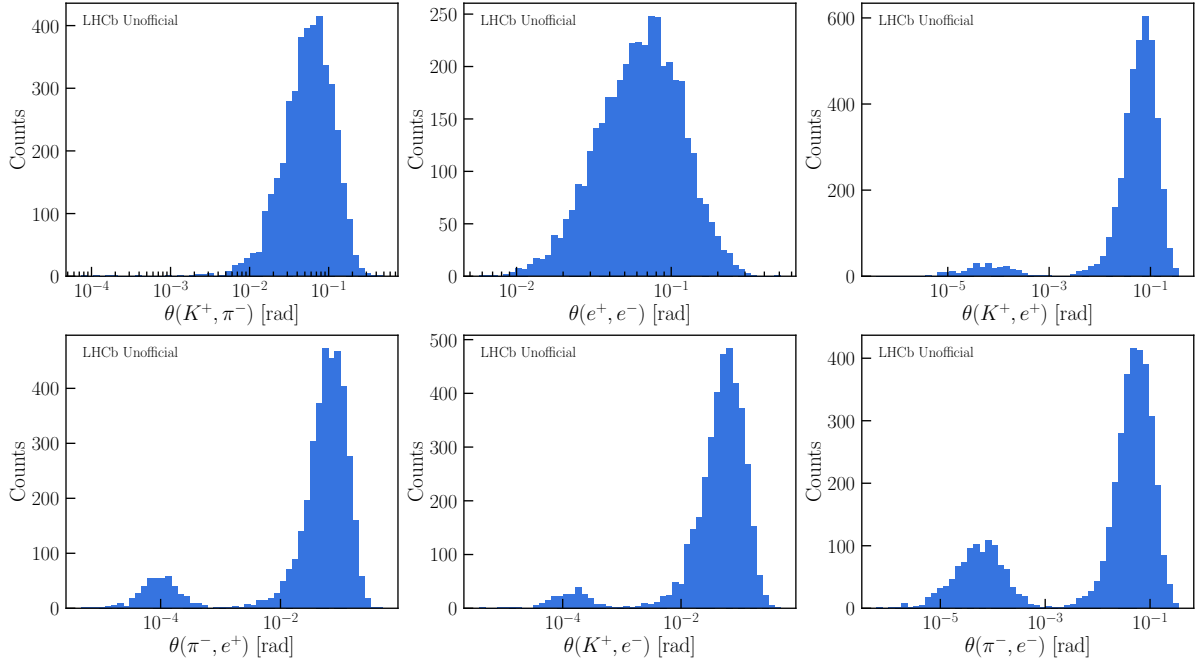
Figure 4.4: The opening angles between the tracks in $B^0 \to K^+\pi^- e^+ e^-$ candidate events.

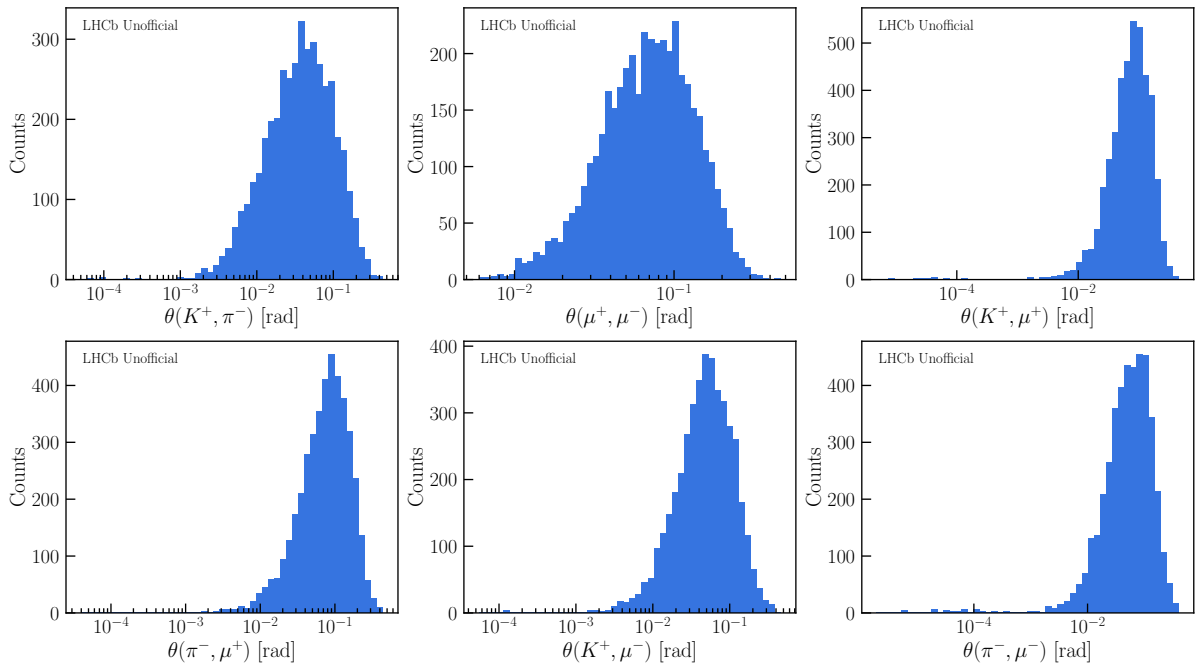Figure 4.5: The opening angles between the tracks in $B^0 \to K^+\pi^- \mu^+ \mu^-$ candidate events.

4.2 Exclusive backgrounds

In addition to the selection criteria mentioned so far for the particle identification and combinatorial background, targeted cuts are made to reduce the contribution from specific decays. These can peak in the signal region, typically due to particle misidentification.

The cuts applied to the $R_{K^{*0}}$ analysis are applied here, including cuts to reduce background from the decays $B^+ \to K^+ \ell^+ \ell^-$, $B_s^0 \to \phi \ell^+ \ell^-$, $\Lambda_b^0 \to pK J/\psi(\to \ell^+ \ell^-)$, $B^0 \to \bar{D}^0(\to K^+ \pi^-)\pi^- \ell^+ \nu_\ell$ as well as the resonant channel $B^0 \to K^+ \pi^- J/\psi(\to \ell^+ \ell^-)$ with particle misidentification.

For instance, for the last one, the requirements are such that all events with

$$|m(K^+_{\to l^+} l^-) - m_{J/\psi}| < 60 \,\text{MeV}/c^2 \ \ \& \ \ l^+ \texttt{ProbNN}_l < 0.8 \tag{4.3}$$

as well as those satisfying

$$|m(\pi^-_{\to l^-} l^+) - m_{J/\psi}| < 60 \,\text{MeV}/c^2 \ \ \& \ \ l^- \texttt{ProbNN}_l < 0.8 \tag{4.4}$$

are removed. The same requirements are applied to filter contributions from the $\psi(2S)$ resonance, analogous but replacing $m_{J/\psi}$ by $m_{\psi(2S)}$. Here, terms such as $m(K^+_{\to l^+} l^-)$ denote the invariant mass of the opposite-sign kaon and lepton under a changed mass hypothesis, where it is assumed that the kaon is a misidentified lepton. Hence, the first requirement filters events where a kaon and lepton have been swapped in the identification, and the second one likewise for a pion and lepton swap.[2] Contributions from the $B^0 \to K^+ \pi^- J/\psi(\to \ell^+ \ell^-)$ decays are further studied in Sect. 4.2.2.

4.2.1 Partially reconstructed backgrounds

Decays of the type $B^0 \to K^+ \pi^- \ell^+ \ell^- X$ can contribute to the background, specifically as a partially reconstructed background, if the decay products denoted by X are not reconstructed. The reconstructed particles are the same as for the signal decay, and will contribute to the background if the invariant mass lies within the same region.

Since the products X are not detected, the invariant mass $m(K^+ \pi^- \ell^+ \ell^-)$ will mostly peak and fall below the B^0 mass where the actual signal peaks. Thus, the partially reconstructed background can be reduced by kinematic requirements, for example by requiring the reconstructed B^0 mass to lie above $5150 \,\text{MeV}/c^2$, as done for the muon channel studied in Ref. [7]. However, this is less effective for electrons with the decreased mass resolution. Nonetheless, partially reconstructed backgrounds are not studied or modeled in the mass fits in this analysis. As detailed in Chapter 5, no significant contributions from such backgrounds are observed in the data.

4.2.2 Charmonium particle swap

With its large branching fraction, the resonant $B^0 \to K^+ \pi^- J/\psi(\to \ell^+ \ell^-)$ decay can easily contribute to the background when one of the leptons is misidentified as a hadron, as the

---

[2]The alternative mass hypothesis is considered by recomputing the invariant mass in Eq. (4.1), where energies of the particles with a changed mass m_i under the new hypothesis are recomputed with the energy-momentum relation $E_i^2 = \|\boldsymbol{p}_i c\|^2 + (m_i c^2)^2$.

aforementioned cut on q^2 is no longer effective. Contributions from this decay were identified in the data sample.

The various invariant mass combinations, such as $m(K^+\pi^-\ell^+)$, $m(\pi^-\ell^-)$ etc. were observed to look for any signs of resonances from peaking backgrounds. Peaks were observed in the opposite-sign pion and lepton combination, i.e. in $m(\pi^-\ell^+)$. This invariant mass was recomputed under a different hypothesis where the pion is a misidentified lepton, denoted by $m(\pi^-_{\to\ell^-}\ell^+)$. These distributions are displayed in Figs. 4.6 and 4.7 for the muon and electron candidate events, respectively. Note that all of the subsequent plots also have a cut applied on the invariant mass $m(K^+\pi^-\ell^+\ell^-)$, where the same region is studied as in the mass fits performed in Chapter 5. For the muons, the large peak in both the lower and upper regions above the K^{*0} resonance matches the invariant mass of the J/ψ meson. Additionally, a smaller peak is visible from the $\psi(2S)$ meson. While this latter peak is not discernible for the electron events with reduced statistics and resolution, the J/ψ peak is still apparent.

This clearly shows that the altered kinematics of the larger $m(K^+\pi^-)$ regions increase the leakage from the charmonium channels with pion-lepton swaps. The cuts applied in Eqs. 4.3 and 4.4 thus have to be adapted for the $m(K^+\pi^-)$ regions studied here.

Figure 4.6: The invariant mass $m(\pi^-\mu^+)$ for $B^0 \to K^{*0}\mu^+\mu^-$ (left) and $B^0 \to K^+\pi^-\mu^+\mu^-$ candidate events in the lower (middle) and upper regions (right). The masses of the charmonium mesons J/ψ and $\psi(2S)$ are displayed.

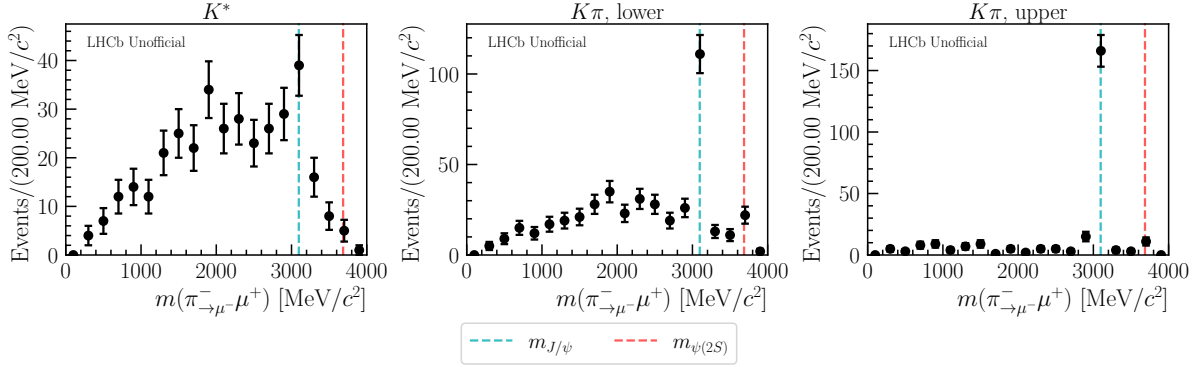Figure 4.7: The invariant mass $m(\pi^-e^+)$ for $B^0 \to K^{*0}e^+e^-$ (left) and $B^0 \to K^+\pi^-e^+e^-$ candidate events in the lower (middle) and upper regions (right). The masses of the charmonium mesons J/ψ and $\psi(2S)$ are displayed.
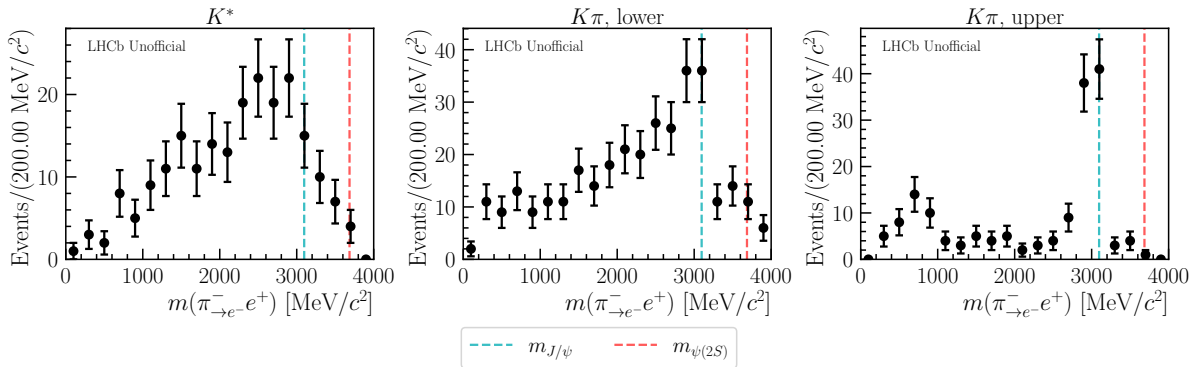
To remove these charmonium contributions through a suitable condition and proceed with the analysis, the particle identification variables were examined to look for any systematic differences between the candidate events in the charmonium peaks and those outside. A suitable variable was found to be the difference of the $\mathtt{ProbNN}_\pi$ variables assigned to the same-sign pion and lepton, i.e. $\pi^-\mathtt{ProbNN}_\pi - l^-\mathtt{ProbNN}_\pi$. The distribution of candidate events in terms of this variable and the invariant mass $m(\pi^-_{\to\ell^-}\ell^+)$ is displayed in Fig. 4.8. For the muon final states, the large J/ψ peak is visible, concentrated around values of the $\mathtt{ProbNN}$ difference slightly above zero, while the bulk of the other events lies at higher values of said difference. This supports the premise that the peaks seen in the mass distribution stem from the resonant charmonium decays, where the same-sign lepton and pion have been swapped in the particle identification.

For the electron states a similar pattern is observed, but less clearly due to the mentioned lower statistics as well as higher background levels. Additionally, the invariant mass peak of the J/ψ meson appears to be shifted downwards. A possible explanation is that the electrons misidentified as pions were not corrected for any bremsstrahlung effects.

The charmonium contributions were removed 'by eye', where sensible conditions on the two variables used in Fig. 4.8 were chosen to cut away a rectangular window. For the muon final states, the events with

$$|m(\pi^-_{\to\mu^-}\mu^+) - m_{J/\psi}| < 100\,\mathrm{MeV}/c^2 \ \ \& \ \ (\pi^-\mathtt{ProbNN}_\pi - \mu^-\mathtt{ProbNN}_\pi) < 0.8 \qquad (4.5)$$

are removed, and for the electron final states those with

$$|m(\pi^-_{\to e^-}e^+) - (m_{J/\psi} - 100\,\mathrm{MeV}/c^2)| < 100\,\mathrm{MeV}/c^2 \ \ \& \ \ (\pi^-\mathtt{ProbNN}_\pi - e^-\mathtt{ProbNN}_\pi) < 0.8\,. \tag{4.6}$$

In the latter case, the charmonium masses are shifted downwards by $100\,\mathrm{MeV}/c^2$ due to the observed shift of the peak. As before, in both cases the analogous cuts are applied replacing $m_{J/\psi}$ by $m_{\psi(2S)}$. The resulting invariant mass distributions after these cuts can be found in Appendix B. Note that the chosen cut performs better on the lower region, whereas some background events remain in the upper region.
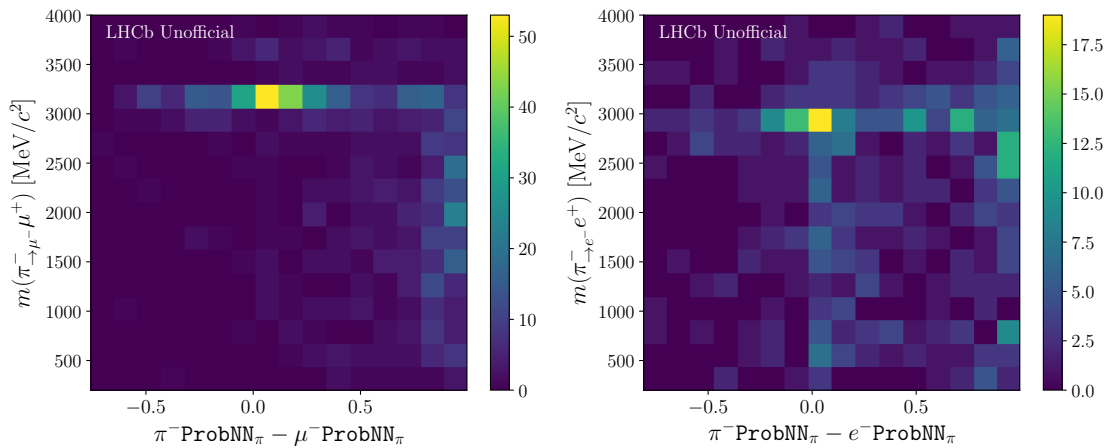

Figure 4.8: The number of $B^0 \to K^+\pi^-\mu^+\mu^-$ (left) and $B^0 \to K^+\pi^-e^+e^-$ (right) candidate events as a function of the invariant mass $m(\pi^-_{\to\ell^-}\ell^+)$ and the difference of the $\mathtt{ProbNN}_\pi$ parameters assigned to the same-sign pion and lepton.

Naturally, a proper background study, beyond the scope of this thesis, would use simulated events to determine such a cut, with the optimum balance of signal efficiency and background rejection. It has nonetheless been shown that certain background cuts such as for swaps with the resonant J/ψ channel will have to be re-optimized. While this was done by directly looking at the data, the backgrounds were filtered in a reasonable way to receive an estimate of the signal yield and statistical uncertainty in the succeeding chapter.

Chapter 5

Analysis

The signal yields are obtained by fits to the invariant mass $m(K^+\pi^-\ell^+\ell^-)$ of the final state decay products. The shape of the signal is obtained from fits to Monte-Carlo simulations. The fitting is performed with *zfit*, a fitting library built on top of *TensorFlow* [38].

5.1   Fitting method

The shape used to model the distribution of signal events is a so-called double Crystal Ball (DCB), defined as [38, 39]

$$f_{\text{sig}}(x; \mu, \sigma, \alpha_{\text{L}}, n_{\text{L}}, \alpha_{\text{R}}, n_{\text{R}}) = \begin{cases} A_{\text{L}} \left(B_{\text{L}} - \frac{x-\mu}{\sigma} \right)^{-n_{\text{L}}}, & \frac{x-\mu}{\sigma} < -\alpha_{\text{L}} \\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right), & -\alpha_{\text{L}} \leq \frac{x-\mu}{\sigma} \leq \alpha_{\text{R}} \\ A_{\text{R}} \left(B_{\text{R}} - \frac{x-\mu}{\sigma} \right)^{-n_{\text{R}}}, & \frac{x-\mu}{\sigma} > \alpha_{\text{R}} \end{cases} \quad (5.1)$$

with

$$A_i = \left(\frac{n_i}{|\alpha_i|} \right)^{n_i} \exp\left(-\frac{|\alpha_i|^2}{2} \right), \quad i = \text{L, R}.$$
$$B_i = \frac{n_i}{|\alpha_i|} - |\alpha_i| \quad (5.2)$$

It consists of a Gaussian core, modeling the resolution of the detector, with power-law tails on both sides, modeling the broadening of the signal shape due to bremsstrahlung emission. While this effect is much more prominent for the electrons, the shape was used for both the electron and muon decays. The parameters μ and σ define the center and width of the Gaussian, respectively. The parameters α_{L} and α_{R} determine the transition point to the power-laws with the exponents n_{L} and n_{R}.

For the background, the combinatorial contribution is modeled with an exponential shape:

$$f_{\text{bkg}}(x; \lambda) = \lambda e^{-\lambda x}. \quad (5.3)$$

It will be seen that this component alone describes the observed background shape well.

The fits are performed using the maximum likelihood principle. Given a probability density function (PDF) $P(x; a)$ which models the data x, parameterized by some parameters a, the

likelihood L is defined as the probability that the model assigns to the data sample $\{x_1, \ldots, x_N\}$:

$$L(x_1, \ldots, x_N; a) = \prod_i P(x_i; a).$$ (5.4)

The optimal parameters $\hat{a}$ are determined as to maximize the likelihood, i.e. the probability that the data sample is observed under the model $P(x; a)$:

$$\hat{a} = \underset{a}{\mathrm{argmax}}\, L(x_1, \ldots, x_N; a).$$ (5.5)

In practice, instead of maximizing the likelihood, the negative logarithm of the likelihood is minimized. This likelihood maximization is equivalent to a minimization of the cross-entropy between the data and model probability distributions, where the data distribution is estimated from the data sample, as the true distribution is unknown.

The fits are performed with extended maximum likelihood fits, where the normalization of $P(x; a)$ is not fixed to unity anymore, i.e. the PDF is replaced by another function with an arbitrary normalization λ (also called an 'extended PDF'). This is appropriate as the number of signal and background events are unknown – in fact, this normalization giving the signal yield is the key value to be obtained from the fits. Specifically, the PDFs defined in Eqs. (5.1) and (5.3) are multiplied by such factors λ which correspond to the signal and background yields.

With this modification, the likelihood assigned to the given data sample with N events is multiplied by the Poisson probability of obtaining N events with a mean of λ:

$$e^{-\lambda}\frac{\lambda^N}{N!}.$$

This term ties the normalization λ to the number of events in the data sample.

The symmetric error is determined using the covariance matrix, obtained from the inversion of the Hessian matrix of the log-likelihood:

$$\mathrm{cov}^{-1}(a_i, a_j) = -\left\langle \frac{\partial^2 \ln L}{\partial a_i \partial a_j} \right\rangle.$$ (5.6)

The errors on the parameters a_i are then the square roots of the diagonal terms in this covariance matrix [40].

To summarize, unbinned extended maximum likelihood fits are performed. The total function is the sum of two PDFs, each weighted by the respective yield and Poisson term:

$$f_{\mathrm{tot}}(x; a_j, \lambda_j) = \lambda_{\mathrm{sig}} f_{\mathrm{sig}}(x; a_{\mathrm{sig}}) + \lambda_{\mathrm{bkg}} f_{\mathrm{bkg}}(x; a_{\mathrm{bkg}}),$$ (5.7)

where a_j denotes the parameters a_{sig} and a_{bkg} that the signal and background PDFs depend on as previously defined, and likewise for λ_j. Which of these parameters are fixed and varied during the fits is detailed in the succeeding sections. The likelihood in the fits is given by

$$L(x_1, \ldots, x_N; a_j, \lambda_j) = e^{-\lambda_{\mathrm{tot}}}\frac{\lambda_{\mathrm{tot}}^N}{N!} \prod_i f_{\mathrm{tot}}(x_i; a_j, \lambda_j),$$ (5.8)

where $\lambda_{\mathrm{tot}} = \lambda_{\mathrm{sig}} + \lambda_{\mathrm{bkg}}$.

5.2 Muon channel

Invariant mass fits are performed to the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidate events with the aforementioned selection cuts applied and the functional form described in the previous section.

The optimal DCB parameters parametrizing the signal are determined from fits to simulated events, performed for both $m(K^+\pi^-)$ regions and shown in Fig. 5.1. For the muons with a clean experimental signature, the signal is a distinct peak at the B-mass $m_{B^0} = 5280\,\text{MeV}/c^2$ [10]. The $m(K^+\pi^-)$ region does not influence the signal shape in a substantial way (as can be verified in Appendix A, where the fit parameters are listed).
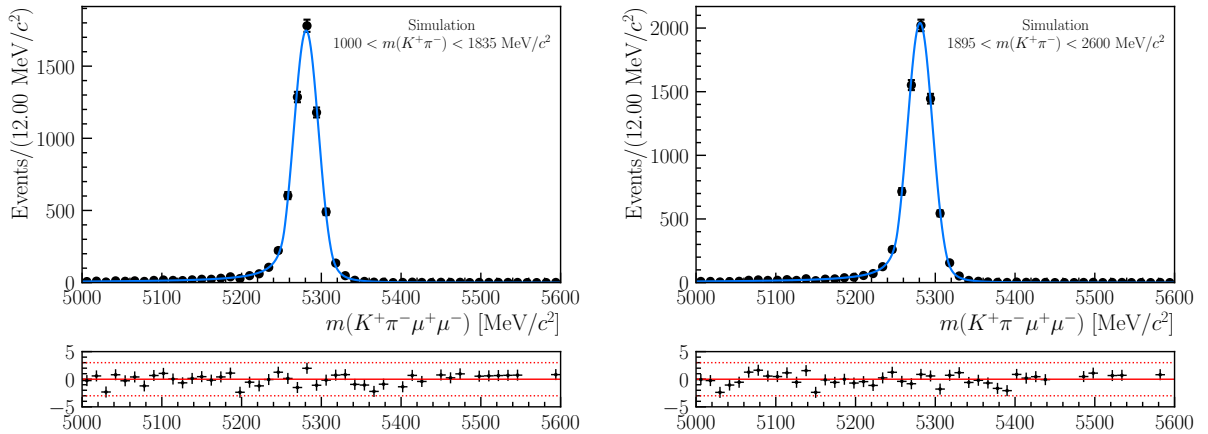

Figure 5.1: Invariant mass fits to Monte-Carlo samples of $B^0 \to K^+\pi^-\mu^+\mu^-$ events, in the lower (left) and upper regions (right). Below the mass distribution and fit, the residuals are shown, which are the differences of the data points and values of the fitted function divided by the uncertainty on the data points. Such residuals will also be shown for all subsequent mass fits.

The region of the invariant mass considered (in simulation and data) is 5000–5600 MeV/c^2. The lower limit could be chosen higher to limit partially reconstructed background, but with this choice it can be examined if there are any substantial contributions from such background.

As mentioned, the signal and background distributions are weighted by the signal and background yields, respectively, which are varied during the fit. In the fits to the data, the other parameters that are allowed to be varied are the slope of the background and the μ and σ of the signal PDF. This allows for a shifting of the Gaussian center and scaling of the width compared to the simulation. The other DCB parameters are kept fixed as determined from the simulation. Additional background sources could be modeled in the fit, such as $\bar{B}_s^0 \to K^+\pi^-\mu^+\mu^-$, but such contributions are found to be negligible.

The data fits are shown in Fig. 5.2, with the obtained signal yields listed in Tab. 5.1. The remaining parameters are listed in Appendix A. In the lower $m(K^+\pi^-)$ region, a clean signal peak is visible, with minimal background that appears to be well described by the exponential combinatorial background, and no characteristic shapes from partially reconstructed background. In the upper region, a peak in the signal region is still visible, but the number of events is reduced and the relative uncertainty on the signal yield is larger. This reduction could be expected firstly due to the more limited phase space as well as the dynamics of the higher K^* resonances, seen to be more suppressed [10].
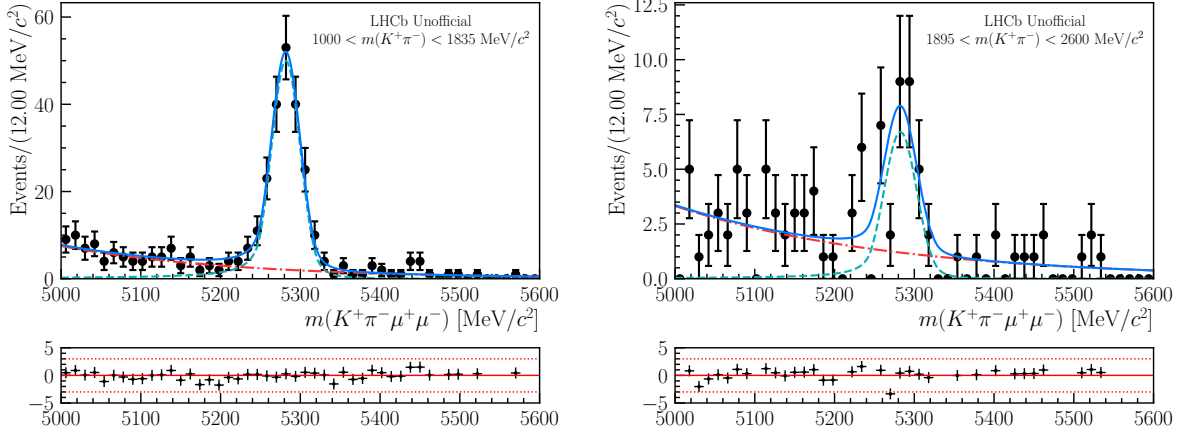
Figure 5.2: Invariant mass fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ data in the lower (left) and upper (right) regions. The total fitted function is plotted in blue, while the signal and background components are displayed separately in green and red, respectively.

Table 5.1: Signal yields obtained from the $B^0 \to K^+\pi^-\mu^+\mu^-$ fits, with the statistical uncertainty.

$m(K^+\pi^-)$	Yield
1000–1835 MeV/c^2	208 ± 16
1895–2600 MeV/c^2	30 ± 7

To compare the statistics with the resonant K^{*0} region, the distribution of candidate events in terms of the invariant mass $m(K^+\pi^-)$ is displayed in Fig. 5.3. A background-subtraction is performed using the *sPlot* technique, implemented in the *hepstats* library [41]. This procedure works by weighting each of the events according to the signal distribution that was obtained in the mass fit (here performed on the entire data, with the signal shape obtained from the $B^0 \to K^+\pi^-\mu^+\mu^-$ simulation in the combined $m(K^+\pi^-)$ region) [42].
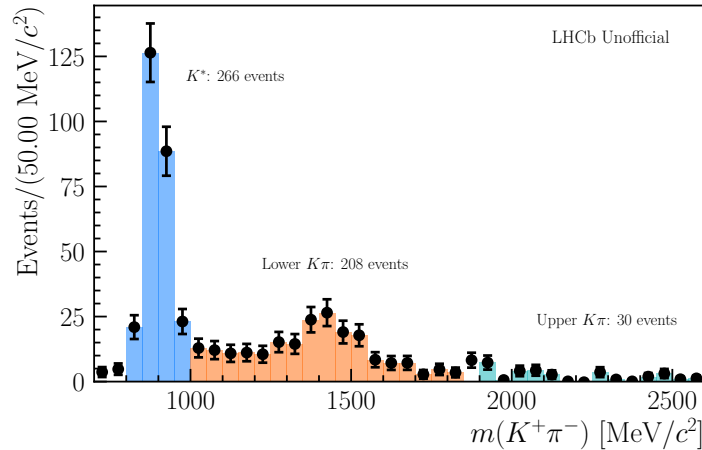

Figure 5.3: Background-subtracted $B^0 \to K^+\pi^-\mu^+\mu^-$ candidate events as a function of the invariant mass $m(K^+\pi^-)$. The resonant K^{*0}, lower and upper $K^+\pi^-$ regions are highlighted with the number of respective events obtained from mass fits. A similarity to the previously observed spectrum in Fig. 4.2 is observed.

While the upper region has comparatively few events as seen in the mass fit, the lower region contains slightly fewer, but a comparable amount to the resonant K^{*0} region with $(78 \pm 8)\,\%$ of the events. Note however that this is with the selection applied that was chosen for the $B^0 \rightarrow K^+\pi^-\ell^+\ell^-$ events. This ratio is slightly below but compatible with the ratio of $(84 \pm 7)\,\%$ observed in Ref. [34] (as seen in Fig. 4.2) between events in the $m(K^+\pi^-)$ region of 1030–1630 MeV/c^2 and the K^{*0} region. It should be noted that the former region is smaller than the lower region considered here, however it does include the bulk of the events observed here as seen in Fig. 5.3. A similar distribution will not be displayed for the electron channel due to the lower statistics, but the relative statistics are similar.

5.3 Electron channel

The invariant mass fits to $B^0 \rightarrow K^+\pi^-e^+e^-$ candidate events are performed in a similar fashion as for the muon final states. As previously mentioned, no simulated samples were available for $B^0 \rightarrow K^+\pi^-e^+e^-$ events, hence the signal shape was obtained from simulated $B^0 \rightarrow K^{*0}e^+e^-$ samples. The fit to the samples is displayed in Fig. 5.4, where one sees that the broad DCB shape describes the (simulated) signal well. Again, the fit parameters are listed in Appendix A.
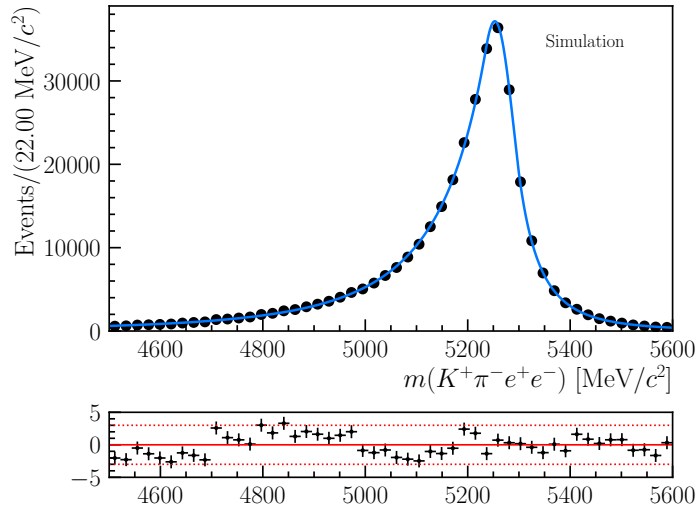

Figure 5.4: Invariant mass fit to a Monte-Carlo sample of $B^0 \rightarrow K^{*0}e^+e^-$ events, used to model the signal shape in $B^0 \rightarrow K^+\pi^-e^+e^-$ data fits.

Due to the worsened mass resolution, as is also visible in the simulation in Fig. 5.4, the invariant mass region considered in the fit is 4500–5800 MeV/c^2. As with the dilepton mass-squared q^2, for the electrons the invariant mass $m(K^+\pi^-e^+e^-)$ of the final state particles is computed with a constraint on the primary vertex, improving the resolution.

For the electron channel, the shape of the signal is fixed, i.e. only the signal and background yields as well as the shape of the background are varied during the fit. Allowing the Gaussian parameters μ and σ to vary as for the muons did not result in a good convergence in the fit. For this reason, they are fixed, but shifted and scaled, respectively, with respect to the values obtained from the simulation, where the shifting and scaling is taken to be the same as for the muon fits. This shifting and scaling is usually fixed from the J/ψ control channels, i.e. in this

case $B^0 \to K^+\pi^- J/\psi(\to \ell^+ e^-)$. This was beyond the scope of this thesis but will have to be done in the future analysis of these decays.

As mentioned, the signal shape for the electrons is broader than for the muons due to the bremsstrahlung emitted. In fact, the signal shape is dependent on the number of bremsstrahlung clusters that were detected and used for the bremsstrahlung correction. Hence, such bremsstrahlung categories have been separated and modeled with separate values obtained from simulation in previous analyses such as the measurement of $R_{K^{*0}}$ in Ref. [7]. Additionally, the fits have been separated into different trigger categories separating for instance the events where one of the electrons or one of the hadrons satisfies the hardware trigger requirements. This is done since the again the resolution is influenced, as well as the background contributions. These categories are then fit simultaneously [7]. However, these bremsstrahlung and trigger categories were combined into one fit for this analysis.

The data fits are shown in Fig. 5.5 and the obtained signal yields are listed in Tab. 5.2. The remaining parameters are again listed in Appendix A. In the lower region, a signal peak is observed, albeit with a relatively large relative uncertainty on the signal yield. As for the muon channel, the upper region suffers from even lower statistics, where the fit does return a positive signal yield of a few events, but it is not clear that any signal is actually present. To judge this quantitatively, hypothesis tests were performed, evaluating the significance with which the null hypothesis (i.e. the hypothesis under which no signal is observed) is rejected.[1] The obtained significances are listed in Tab. 5.2 as well. The lower region with a significance of $5.8\,\sigma$ shows a clear observation of this previously unobserved decay, while the upper region with $0.86\,\sigma$ is not conclusive in this regard. Note that this is a purely statistical significance, further systematic studies are required to validate this claim.
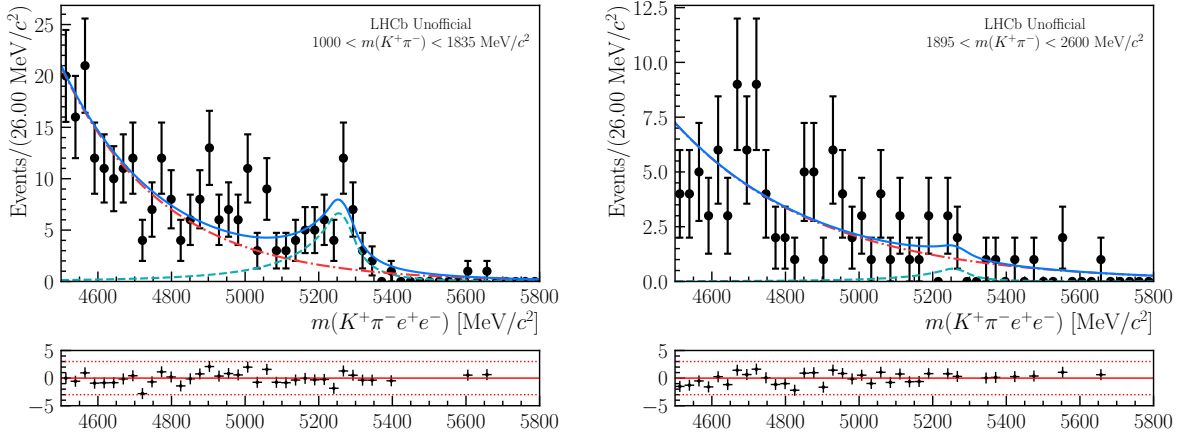

Figure 5.5: Invariant mass fits to the $B^0 \to K^+\pi^- e^+ e^-$ data in the lower (left) and upper (right) regions. The total fitted function is plotted in blue, while the signal and background components are displayed separately in green and red, respectively.

---

[1]The p-value, giving the probability that the data is compatible with the null hypothesis, is computed using the test statistic q_0, given by the likelihood ratio:

$$q_0 = \begin{cases} -2\ln \dfrac{L(\lambda = 0, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\lambda}, \hat{\boldsymbol{\theta}})}, & \hat{\lambda} \geq 0 \\ 0 & , \quad \hat{\lambda} < 0\,, \end{cases} \qquad (5.9)$$

Table 5.2: Signal yields obtained from the $B^0 \to K^+\pi^- e^+ e^-$ fits, with the statistical uncertainty and significance.

$m(K^+\pi^-)$	Yield	Significance
1000–1835 MeV/c^2	56 ± 12	$5.8\,\sigma$
1895–2600 MeV/c^2	5 ± 6	$0.86\,\sigma$

Again, sources of partially reconstructed background were not considered in this fit. As in the muon channel, when looking at the data distribution, the combinatorial background appears to model the background relatively well, with no distinct shape from a partially reconstructed background. Such background is kinematically limited compared to the K^{*0} decay due to the higher requirement on the invariant mass $m(K^+\pi^-)$. However, to reach a more thorough conclusion on such partially reconstructed backgrounds, they will have to be studied with simulation.

where $\hat{\hat{\boldsymbol{\theta}}}$ and $\hat{\boldsymbol{\theta}}$ denote the values that maximize the likelihood for a signal yield λ fixed to zero and a freely floating one with fit value $\hat{\lambda}$, respectively. Wilks' theorem states that this statistic asymptotically approaches the χ^2-distribution. The p-value is obtained by integrating the PDF of the statistic q_0 for all values above $q_{0,\text{obs}}$. The p-value is converted into the equivalent statistical significance

$$Z = \phi^{-1}(1-p) = \sqrt{q_0}\,, \tag{5.10}$$

where ϕ is the cumulative distribution of the standard Gaussian and the second equality follows from Wilks' theorem [43].

Chapter 6

Sensitivity study

Given the results of the invariant mass fits, a sensitivity study is performed. The quality of the fits is studied through simulated pseudo-experiments giving an improved estimate of the value and uncertainty of the parameter of interest, which is the signal yield for the various decays. Subsequently, the statistical error on $R_{K\pi}$ stemming from the contributions of the rare decay mode yields is estimated. This estimation is also performed for an extrapolation onto the entire LHCb dataset.

6.1 Pseudo-experiments

The quality of the fits is studied by resampling from the sample probability distribution which was constructed in the data fit. The parameters obtained in the data fit are then corrected for biases which become apparent in this sampling.

2000 pseudo-experiments (or 'toy' experiments) are simulated for each of the decays. This is done by sampling from the total (combined signal and background) event distribution obtained from the mass fits in the previous chapter. The number of events in each pseudo-experiment is sampled from a Poisson distribution with a mean given by the total yield obtained in the mass fits.

Each of these pseudo-experiments is fit with the same method as for the data fit, resulting in a distribution of signal yields and their statistical uncertainties. From these, the pull distribution is computed, with the pull defined by

$$p = \frac{\lambda_{\mathrm{p}} - \lambda}{\sigma_{\lambda_{\mathrm{p}}}}, \tag{6.1}$$

where λ is the 'true' value for the signal yield obtained from the data fit and λ_{p} the result of a particular pseudo-experiment with uncertainty $\sigma_{\lambda_{\mathrm{p}}}$.

For a proper and unbiased fit, the pull distribution should follow a standard Gaussian distribution with a mean of zero and width of one. If this is not the case, then the deviation of the pull distribution from this standard Gaussian can be used to correct the values from the original data fit in order to obtain an unbiased pull distribution [44]. If the pull distribution has a mean and width of μ_p and σ_p, then the data fit values can be corrected as [45]

$$\begin{aligned} \lambda^{\mathrm{c}} &= \lambda - \mu_p \sigma_\lambda \\ \sigma_\lambda^{\mathrm{c}} &= \sigma_p \sigma_\lambda, \end{aligned} \tag{6.2}$$

where λ is the signal yield with uncertainty σ_λ.

The pull distributions for the pseudo-experiments are shown in Figs. 6.1 and 6.2, together with a fitted Gaussian (using an unbinned maximum likelihood fit) and the obtained values for the mean μ and width σ. The distributions all closely match a standard Gaussian, with the exception of the lower region for the electron channel, where $\mu = 0.24 \pm 0.02$, showing that this fit is slightly biased to higher values of the signal yield. Using Eq. (6.2), the values obtained in the data fits are corrected for biases and listed in Tab. 6.1. As expected from the pull distributions, the largest change is in the lower region for the electron channel, where the yield is reduced by a few events. Additionally, the uncertainty in the upper region of the electron channel is slightly increased. The reason for the bias in the lower electron channel is unresolved – low yields can cause such biases, but it is then unclear why the same behavior isn't observed in the upper electron channel. However, by correcting the yields, this bias was taken into account for the estimation of the sensitivity on $R_{K\pi}$ performed in the next section.
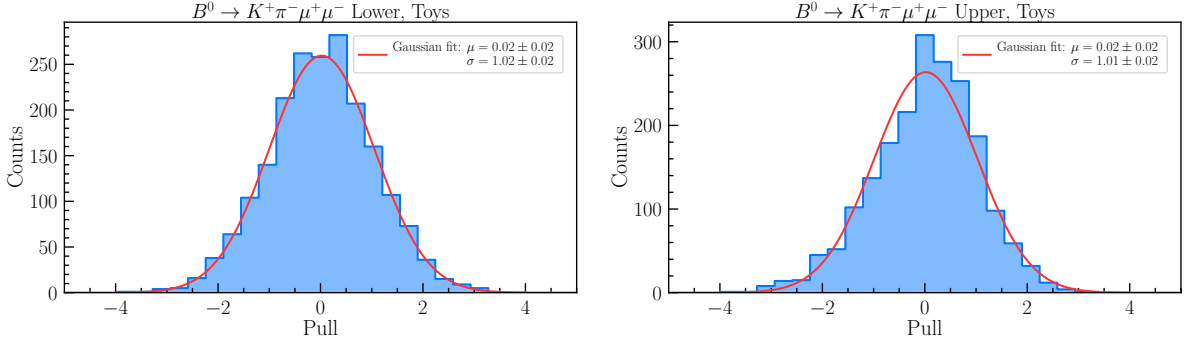

Figure 6.1: Pull distribution in the $B^0 \to K^+\pi^-\mu^+\mu^-$ pseudo-experiments in the lower (left) and upper (right) regions, with a Gaussian fit and the obtained values with their statistical uncertainties.

Figure 6.2: Pull distribution in the $B^0 \to K^+\pi^-e^+e^-$ pseudo-experiments in the lower (left) and upper (right) regions, with a Gaussian fit and the obtained values with their statistical uncertainties.
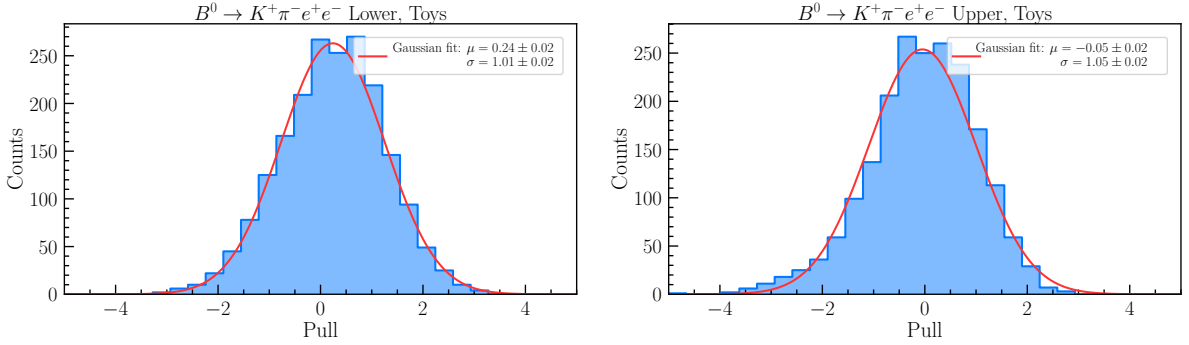
Table 6.1: Yields from the data fits corrected for biases found in the pull distribution of pseudo-experiments.

Decay	Yield
$B^0 \to K^+\pi^-\mu^+\mu^-$, lower	208 ± 16
$B^0 \to K^+\pi^-\mu^+\mu^-$, upper	30 ± 7
$B^0 \to K^+\pi^-e^+e^-$, lower	53 ± 12
$B^0 \to K^+\pi^-e^+e^-$, upper	5 ± 7

6.2 Extrapolation to the entire LHCb dataset

To estimate the relative uncertainty that could be obtained on the yields as well as the ratio $R_{K\pi}$ with the entire LHCb dataset (up to 2018), the corrected signal yields are scaled up by the integrated luminosity, listed in Tab. 6.2 for each year together with the center-of-mass energy. As the b-meson production cross section has been seen to approximately scale with the center-of-mass energy [46], the integrated luminosities from lower energies are scaled by the factors 7/13 and 8/13 for 2011 and 2012, respectively. This gives a factor of 4.6 for the expected events from the 2016 data to the full dataset. The statistical uncertainties, assuming to be uncorrelated between the different years, are scaled by the square root of said factor. The resulting yields are listed in Tab. 6.3.

Table 6.2: Integrated luminosity at LHCb and center-of-mass energy by year [47].

Year	L [fb^{-1}]	E [TeV]
2011	1.0	7
2012	2.0	8
2015	0.30	13
2016	1.6	13
2017	1.7	13
2018	2.1	13

Table 6.3: Projected yields onto the entire LHCb dataset.

Decay mode	Yield
$B^0 \to K^+\pi^-\mu^+\mu^-$, lower	954 ± 35
$B^0 \to K^+\pi^-\mu^+\mu^-$, upper	138 ± 15
$B^0 \to K^+\pi^-e^+e^-$, lower	243 ± 25
$B^0 \to K^+\pi^-e^+e^-$, upper	24 ± 14

The contributions to the ratio $R_{K\pi}$ in Eq. (2.5) from the rare decays studied is:

$$R_{K\pi} \propto \frac{\mathcal{B}(B^0 \to K^+\pi^-\mu^+\mu^-)}{\mathcal{B}(B^0 \to K^+\pi^-e^+e^-)} = \frac{N_\mu}{N_e}\frac{\varepsilon_e}{\varepsilon_\mu}, \qquad (6.3)$$

where N_μ denotes the muon channel yield, ε_μ its selection efficiency and likewise for the electron channel. The relative uncertainty on $R_{K\pi}$ consists of the relative uncertainties of the factors in the right-hand side of Eq. (6.3), such as σ_{N_μ}/N_μ, added in quadrature. While the efficiencies, which are not studied here, carry uncertainties, the dominant contributions to the uncertainty will stem from the signal yields, and the electron channel yield in particular. To estimate the sensitivity on $R_{K\pi}$, the relative uncertainties on the (corrected) data fit values and said values scaled to the entire dataset are taken from Tabs. 6.1 and 6.3, respectively, and added in quadrature. The highest sensitivity, i.e. lowest relative uncertainty, was found to be obtained in the lower $m(K^+\pi^-)$ region. The resulting values are listed in Tab. 6.4. The same analysis including the mass fits was also performed for the lower and upper $m(K^+\pi^-)$ regions combined, but the sensitivity was found to be worsened as compared to the lower region. The fits and sensitivity values are found in Appendix C.

Table 6.4: The relative uncertainty on $R_{K\pi}$, in the lower $m(K^+\pi^-)$ region, stemming from the signal yield uncertainties, for the values obtained from the fits to the 2016 data as well as extrapolated to the entire LHCb dataset.

	$R_{K\pi}$ rel. uncertainty
2016 data	23.8 %
2011–2018	11.1 %

As a comparison, the recently reported value of $R_K = 0.846^{+0.042+0.013}_{-0.039-0.012}$ in Ref. [4], using the entire 2011–2018 dataset, has an (upper) relative statistical uncertainty of 5.0 %. Said value for the latest reported value of $R_{K^*} = 0.69^{+0.11}_{-0.07} \pm 0.05$ (in the q^2 region of 1.1–6 GeV^2/c^4), using an integrated luminosity of 3 fb^{-1}, is 16 %. Scaling this uncertainty to the entire dataset with the same method as was done for the estimate on $R_{K\pi}$ would give a relative uncertainty on R_{K^*} of 7.8 %. Compared to these two values, the estimated value on $R_{K\pi}$ is higher. However, this value should be improved with an optimized preselection and selection, quite feasibly to below 10 %.

Chapter 7

Conclusion

Using data collected at LHCb in 2016, the decays $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-e^+e^-$ have been observed and studied, for the first time for the latter. The results were used to estimate the obtainable sensitivity on the ratio $R_{K\pi}$ as another way to test Lepton Flavor Universality.

The difference in kinematics compared to the previously studied K^{*0} decays were found to influence the levels of combinatorial, exclusive as well as partially reconstructed backgrounds. In particular, exclusive backgrounds from the resonant charmonium channels were seen to increase due to the larger kinematic overlap. Similarly, the level of combinatorial background increased due to the larger $m(K^+\pi^-)$ region, prompting a higher (arbitrary) cut on the BDT response. On the other hand, the altered kinematics suggest a reduced partially reconstructed background, and no such contributions were observed.

The study of all the aforementioned backgrounds as well as numerous other aspects will have to be improved for a proper analysis of the examined decays and the ratio $R_{K\pi}$. The cuts on exclusive backgrounds will be re-optimized by observing the effect of such cuts on simulated signal and exclusive background events, determining the number of expected events in the data and maximizing the sensitivity. The cut on the combinatorial background in the multivariate analysis is optimized in a similar way.

The q^2 range of 1.1–6 GeV^2/c^4 considered in this thesis could be extended on the lower and upper end. Furthermore, to avoid potential biases, the signal regions, particularly the mass distribution in the electron channel, would be blinded during the analysis. Once it has been ensured that every step has been properly carried out, they are revealed at the end together with the obtained signal yield and efficiency.

While the numerical values obtained for the sensitivity on $R_{K\pi}$ depend on the various preliminary selections made, the overall statistics were seen to be comparable but slightly reduced compared to the K^{*0} decays. Note that while it has been concluded that the lower $m(K^+\pi^-)$ region provides the best sensitivity, with an improved selection, it is possible that an increased sensitivity could be obtained by looking at the combined lower and upper $m(K^+\pi^-)$ regions or even some other range of the invariant mass.

In conclusion, this analysis has shown that while a measurement of $R_{K\pi}$ may not produce the leading sensitivity for New Physics in the realm of Lepton Flavor Universality, it can nonetheless provide another cross-check of such effects and strengthen their predictive power in combination with other measurements.

Appendix

A   Mass fit parameters

Table A1: The DCB parameters obtained from fits to $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ simulation.

Channel	μ [MeV/c^2]	σ [MeV/c^2]	α_{L}	n_{L}	α_{R}	n_{R}
$\mu^+\mu^-$, lower	5280.98 ± 0.24	15.48 ± 0.24	1.76 ± 0.06	1.22 ± 0.09	2.07 ± 0.10	2.61 ± 0.35
$\mu^+\mu^-$, upper	5280.84 ± 0.22	15.67 ± 0.20	1.80 ± 0.06	1.18 ± 0.09	2.19 ± 0.09	2.51 ± 0.34
e^+e^-	5252.91 ± 0.34	40.2 ± 0.5	0.426 ± 0.007	3.33 ± 0.05	1.037 ± 0.016	3.24 ± 0.08

Table A2: The parameters obtained from fits to $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-e^+e^-$ candidate events (the signal yields are listed in the main text). λ_{bkg} denotes the background yield and λ the slope of the combinatorial background as defined in Eq. (5.3), while Δ_μ and a_σ denote the shifting and scaling of μ and σ, respectively. Note that these last two values are fixed for the electron channels to the corresponding values from the muon channels.

Channel	λ_{bkg}	λ [(MeV/c^2)$^{-1}$]	Δ_μ [MeV/c^2]	a_σ
$\mu^+\mu^-$, lower	128 ± 13	0.0046 ± 0.0006	0.8 ± 1.5	1.16 ± 0.09
$\mu^+\mu^-$, upper	69 ± 9	0.0036 ± 0.0008	3 ± 4	1.08 ± 0.19
e^+e^-, lower	218 ± 17	0.0036 ± 0.0004		
e^+e^-, upper	105 ± 12	0.0026 ± 0.0004		

B   Invariant mass distributions

The invariant mass distributions $m(\pi^-_{\to l^-}\ell^+)$ and q^2 are displayed for the candidate events with the final selection used in the analysis, and background-subtracted using the *sPlot* technique.
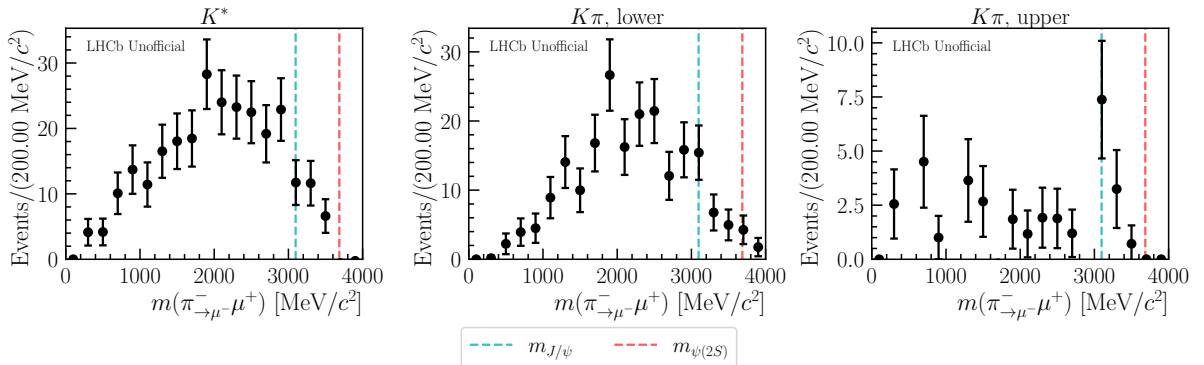

Figure B.1: The background-subtracted invariant mass $m(\pi^-_{\to\mu^-}\mu^+)$ for $B^0 \to K^{*0}\mu^+\mu^-$ (left) and $B^0 \to K^+\pi^-\mu^+\mu^-$ candidate events in the lower (middle) and upper regions (right) with all of the selection cuts applied.
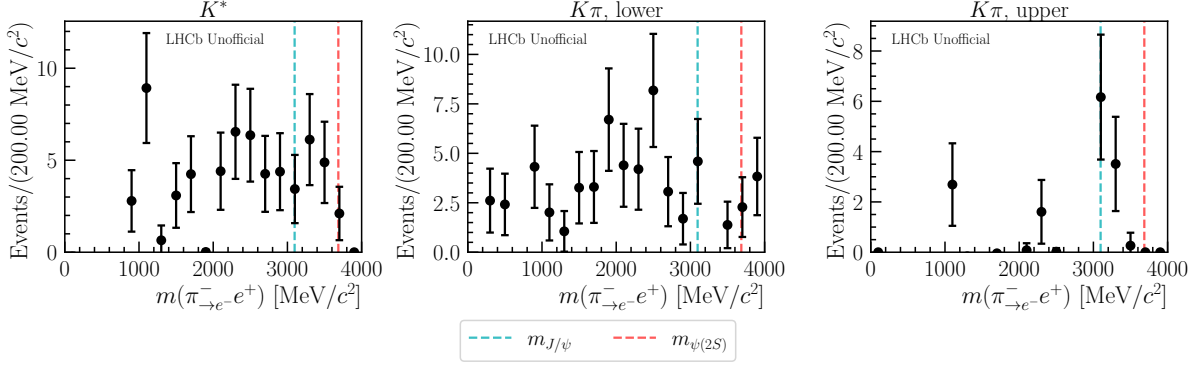
Figure B.2: The background-subtracted invariant mass $m(\pi^-_{\to e^-} e^+)$ for $B^0 \to K^{*0} e^+ e^-$ (left) and $B^0 \to K^+ \pi^- e^+ e^-$ candidate events in the lower (middle) and upper regions (right) with all of the selection cuts applied.

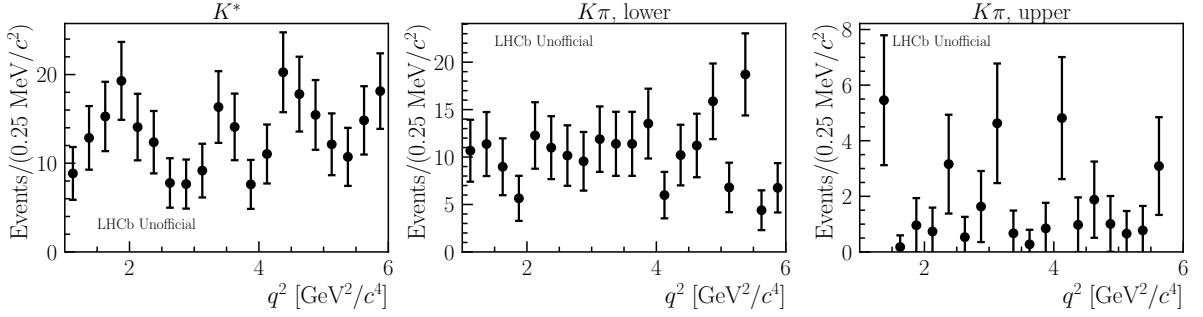Figure B.3: The background-subtracted dilepton mass-squared for $B^0 \to K^{*0} \mu^+ \mu^-$ (left) and $B^0 \to K^+ \pi^- \mu^+ \mu^-$ candidate events in the lower (middle) and upper regions (right) with all of the selection cuts applied.

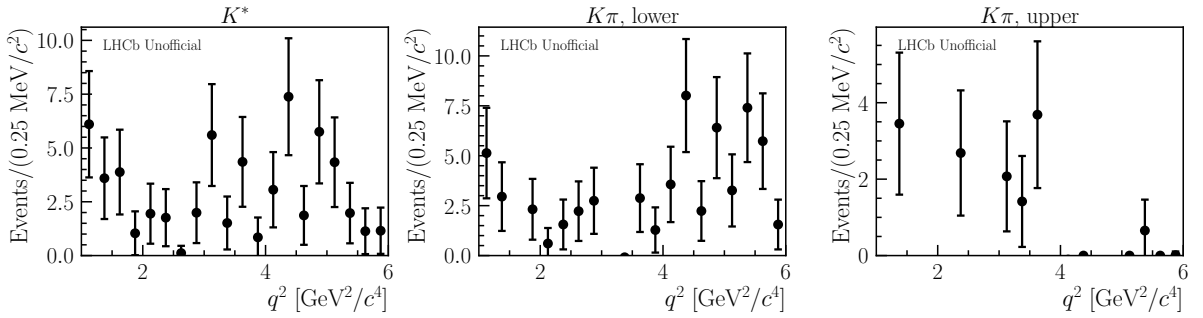Figure B.4: The background-subtracted dilepton mass-squared for $B^0 \to K^{*0} e^+ e^-$ (left) and $B^0 \to K^+ \pi^- e^+ e^-$ candidate events in the lower (middle) and upper regions (right) with all of the selection cuts applied.

39

C Analysis in the combined lower and upper regions

Shown are the invariant mass fits and sensitivity study result in the combined lower and upper $m(K^+\pi^-)$ regions. Note how a lower sensitivity is obtained compared to the lower region due to the higher uncertainty in the upper region.
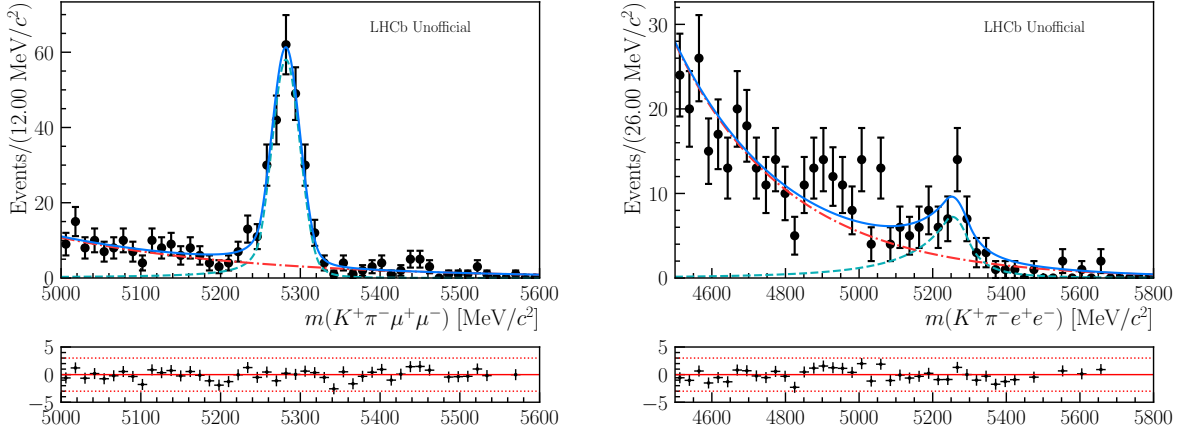

Figure C.1: Invariant mass fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ (left) and $B^0 \to K^+\pi^-e^+e^-$ (right) data in the lower and upper regions combined. The total fitted function is plotted in blue, while the signal and background components are displayed separately in green and red, respectively.
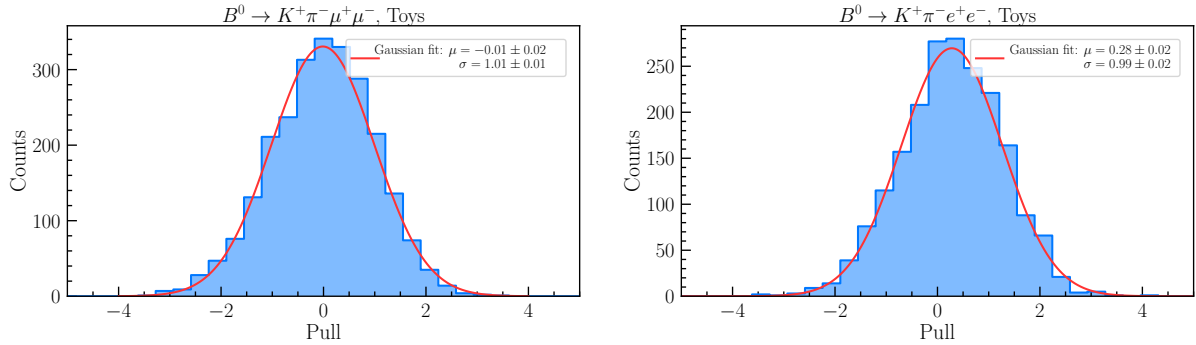

Figure C.2: Pull distribution in the $B^0 \to K^+\pi^-\mu^+\mu^-$ (left) and $B^0 \to K^+\pi^-e^+e^-$ (right) pseudo-experiments the combined lower and upper regions, with a Gaussian fit and the obtained values with their statistical uncertainties.

Table C1: The relative uncertainty on $R_{K\pi}$, in the combined upper and lower $m(K^+\pi^-)$ region, stemming from the signal yield uncertainties, for the values obtained from the fits to the 2016 data as well as extrapolated to the entire LHCb dataset.

	$R_{K\pi}$ rel. uncertainty
2016 data	24.6 %
2011–2018	11.5 %

Bibliography

1. Thomson, M. *Modern Particle Physics* (Cambridge University Press, 2013).

2. Altarelli, G. & Wells, J. *Collider physics within the standard model: a primer* (Springer, Cham, 2017).

3. Bifani, S., Descotes-Genon, S., Vidal, A. R. & Schune, M.-H. Review of lepton universality tests in B decays. *Journal of Physics G: Nuclear and Particle Physics* **46,** 023001 (2018).

4. LHCb collaboration *et al. Test of lepton universality in beauty-quark decays* 2021. arXiv: `2103.11769 [hep-ex]`.

5. Archilli, F., Bettler, M. O., Owen, P. & Petridis, K. A. Flavour-changing neutral currents making and breaking the standard model. *Nature* **546,** 221–226 (2017).

6. LHCb detector performance. *International Journal of Modern Physics A* **30,** 1530022. eprint: `https://doi.org/10.1142/S0217751X15300227` (2015).

7. Aaij, R. *et al.* Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays. *Journal of High Energy Physics* **2017,** 55 (2017).

8. Müller, K. & on behalf of the LHCb Collaboration. Tests of Lepton Flavour Universality at LHCb. *Journal of Physics: Conference Series* **1271,** 012009 (2019).

9. Crivellin, A., Müller, D. & Ota, T. Simultaneous explanation of $R(D^{(*)})$ and $b \to s\mu^+\mu^-$: the last scalar leptoquarks standing. *Journal of High Energy Physics* **2017,** 40 (2017).

10. Group, P. D. *et al.* Review of Particle Physics. *Progress of Theoretical and Experimental Physics* **2020.** 083C01. eprint: `https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf` (Aug. 2020).

11. Alda, J., Guasch, J. & Peñaranda, S. Some results on lepton flavour universality violation. *The European Physical Journal C* **79,** 588 (2019).

12. Celis, A., Fuentes-Martín, J., Jung, M. & Serôdio, H. Family nonuniversal $Z^{'}$ models with protected flavor-changing interactions. *Phys. Rev. D* **92,** 015007 (1 2015).

13. Precision electroweak measurements on the Z resonance. *Physics Reports* **427,** 257–454 (2006).

14. Heavy Flavor Averaging Group *et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018* 2019. arXiv: `1909.12524 [hep-ex]`.

15. Apollinari, G., Brüning, O., Nakamoto, T. & Rossi, L. Chapter 1: High Luminosity Large Hadron Collider HL-LHC. High Luminosity Large Hadron Collider HL-LHC. *CERN Yellow Report.* 21 pages, chapter in High-Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report, 1–19. 21 p (2017).

16. Kou, E. *et al.* The Belle II Physics Book. *Prog. Theor. Exp. Phys.* **2019** (ed Kou, E.) 123C01. 654 p. arXiv: `1808.10567` (2018).

17. Bordone, M., Isidori, G. & Pattori, A. On the standard model predictions for R_K and R_{K^*}. *The European Physical Journal C* **76,** 440 (2016).

18. Aaij, R. *et al.* Search for Lepton-Universality Violation in $B^+ \to K^+\ell^+\ell^-$ Decays. *Phys. Rev. Lett.* **122,** 191801 (19 2019).

19.	Ablikim, M. *et al.* Precision measurements of $\mathcal{B}[\psi(3686) \to \pi^+\pi^- J/\psi]$ and $\mathcal{B}[J/\psi \to l^+l^-]$. *Phys. Rev. D* **88,** 032007 (3 2013).

20.	Aaij, R. *et al.* Test of lepton universality with $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays. *Journal of High Energy Physics* **2020,** 40 (2020).

21.	LHCb collaboration *et al. Angular analysis of the* $B^+ \to K^{*+}\mu^+\mu^-$ *decay* 2020. arXiv: `2012.13241 [hep-ex]`.

22.	Buttazzo, D., Greljo, A., Isidori, G. & Marzocca, D. B-physics anomalies: a guide to combined explanations. *Journal of High Energy Physics* **2017,** 44 (2017).

23.	Evans, L. & Bryant, P. LHC Machine. *Journal of Instrumentation* **3,** S08001–S08001 (2008).

24.	Evans, L. & Evans, L. *The Large Hadron Collider: a marvel of technology; 2nd ed.* (EPFL Press, Lausanne, 2018).

25.	Mobs, E. The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019. General Photo (2019).

26.	The ATLAS Collaboration *et al.* The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation* **3,** S08003–S08003 (2008).

27.	The CMS Collaboration *et al.* The CMS experiment at the CERN LHC. *Journal of Instrumentation* **3,** S08004–S08004 (2008).

28.	The ALICE Collaboration *et al.* The ALICE experiment at the CERN LHC. *Journal of Instrumentation* **3,** S08002–S08002 (2008).

29.	The LHCb Collaboration *et al.* The LHCb Detector at the LHC. *Journal of Instrumentation* **3,** S08005–S08005 (2008).

30.	Lindner, R. *LHCb layout_2. LHCb schema_2* LHCb Collection. 2008.

31.	Elsässer, C. $\bar{b}b$ *production angle plots* `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html` (2021-02-25).

32.	Barbosa-Marinho, P. R. *et al. LHCb muon system: Technical Design Report* (CERN, Geneva, 2001).

33.	Aaij, R. *et al.* Model-Independent Observation of Exotic Contributions to $B^0 \to J/\psi K^+\pi^-$ Decays. *Phys. Rev. Lett.* **122,** 152002 (15 2019).

34.	Aaij, R. *et al.* Differential branching fraction and angular moments analysis of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region. *Journal of High Energy Physics* **2016,** 65 (2016).

35.	Hoecker, A. *et al.* TMVA - Toolkit for Multivariate Data Analysis. arXiv: `physics/0703039 [physics.data-an]` (Mar. 2007).

36.	Needham, M. *Identification of Ghost Tracks using a Likelihood Method* tech. rep. LHCb-2008-026. CERN-LHCb-2008-026. LPHE-2008-004 (CERN, Geneva, 2008).

37.	Rodrigues, E. *Dealing with clones in the tracking* tech. rep. LHCb-2006-057. CERN-LHCb-2006-057 (CERN, Geneva, 2006).

38.	Eschle, J., Puig Navarro, A., Silva Coutinho, R. & Serra, N. zfit: Scalable pythonic fitting. *SoftwareX* **11,** 100508 (2020).

39.	Skwarnicki, T. *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances* PhD thesis (Cracow, INP, 1986).

40.	Barlow, R. J. *Statistics: a guide to the use of statistical methods in the physical sciences* (Wiley, Chichester, 1989).

41. Marinangeli, M., Eschle, J., Rodrigues, E. & Pollack, B. *scikit-hep/hepstats: 0.4.0* version v0.4.0. Apr. 2021.

42. Pivk, M. & Le Diberder, F. sPlot: A statistical tool to unfold data distributions. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **555,** 356–369 (2005).

43. Cowan, G., Cranmer, K., Gross, E. & Vitells, O. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C* **71,** 1554 (2011).

44. Demortier, L. & Lyons, L. Everything you always wanted to know about pulls, CDF Public Note 5776 (2002).

45. Karbach, T. M. & Schlupp, M. *Constraints on Yield Parameters in Extended Maximum Likelihood Fits* tech. rep. Comments: 6 pages, 7 figures (2012). arXiv: `1210.7141`.

46. Aaij, R. *et al.* Measurement of the b-Quark Production Cross Section in 7 and 13 TeV pp Collisions. *Phys. Rev. Lett.* **118,** 052002 (5 2017).

47. The LHCb Collaboration. *Standard set of performance numbers* `https://lhcb.web.cern.ch/speakersbureau/html/PerformanceNumbers.html` (2021-01-19).