Fakultät für
Mathematik, Informatik
und Naturwissenschaften

RWTH AACHEN UNIVERSITY

Test of Lepton Universality with $b \to s\,\ell^+\ell^-$ Decays at LHCb

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Stephan Escher M. Sc.
aus
Hagen

Berichter: Univ.-Prof. Dr. rer. nat. Stefan Schael
Dr. rer. nat. Christoph Michael Langenbruch

Tag der mündlichen Prüfung: 24.11.2022

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

Abstract

Tests of lepton universality are currently among the most discussed observables in flavour physics and show tensions with Standard Model of Particle Physics predictions in recent measurements. The analysis presented in this thesis constitutes the first simultaneous measurement of lepton universality in the modes $B^+ \to K^+\ell^+\ell^-$ (R_K) and $B^0 \to K^{*0}\ell^+\ell^-$ ($R_{K^{*0}}$). Two regions of the squared invariant mass of the dilepton system (q^2) are evaluated, where the low q^2 region covers 0.1–1.1 GeV^2/c^4 and the central q^2 region 1.1–6.0 GeV^2/c^4. The analysis exploits the 2011 to 2018 datasets of LHCb corresponding to an integrated luminosity of 9 fb^{-1}. To avoid bias in the results, the complete analysis was performed in a blinded manner. The results of the four lepton universality ratios are

$$\begin{aligned} R_K \quad \text{low} \quad &= \quad 0.875 \pm 0.070 \pm 0.026^{+0.095}_{-0.000} \\ R_K \quad \text{central} \quad &= \quad 0.904 \pm 0.039 \pm 0.008^{+0.022}_{-0.000} \\ R_{K^{*0}} \quad \text{low} \quad &= \quad 0.856 \pm 0.075 \pm 0.032^{+0.071}_{-0.000} \\ R_{K^{*0}} \quad \text{central} \quad &= \quad 0.954 \pm 0.062 \pm 0.016^{+0.085}_{-0.000}. \end{aligned}$$

The first two uncertainties quoted above are the statistical and systematic uncertainties from a simultaneous fit of the invariant B mass distributions. The third uncertainty is assigned as a conservative estimate of the instabilities of the R_K and $R_{K^{*0}}$ ratios when tightening the electron identification criteria. This effect was observed post-unblinding and is under investigation by the LHCb collaboration. This problem has to be resolved before these results can be published.

R_K is measured in the low q^2 region for the first time. The calibration of the signal efficiencies was successfully validated using the corresponding $B^+ \to K^+ J/\psi\,(\to \ell^+\ell^-)$ and $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ ratios. The invariant B mass fits were checked for bias with pseudo-experiments. The presented analysis is a product of a group effort from several members of the Rare Decays working group of LHCb to which the author of this thesis has made significant contributions.

Zusammenfassung

Untersuchungen der Leptonen Universalität zählen momentan zu den meist diskutierten Größen der Flavour Physik und weisen in vergangenen Messungen Spannungen mit Standard Model of Particle Physics Vorhersagen auf. Die in dieser Dissertation präsentierte Analyse stellt die erste simultane Messung von Leptonen Universalität in den Kanälen $B^+ \to K^+\ell^+\ell^-$ (R_K) und $B^0 \to K^{*0}\ell^+\ell^-$ ($R_{K^{*0}}$) dar. Es wurden zwei Regionen der quadrierten invarianten Masse des Zwei-Leptonen Systems (q^2) untersucht, wobei die tiefe q^2 Region 0.1–1.1 GeV^2/c^4 abdeckt und die zentrale q^2 Region 1.1–6.0 GeV^2/c^4. Die Analyse nutzt die 2011 bis 2018 Datensätze von LHCb, was einer integrierten Luminosität von $9\,\mathrm{fb}^{-1}$ entspricht. Um einen Bias der Ergebnisse zu verhindern, wurde die gesamte Analyse blind durchgeführt. Die Ergebnisse der vier Leptonen Universalitäts-Verhältnisse sind

$$\begin{aligned}
R_K \quad \text{tief} \quad &= \quad 0.875 \pm 0.070 \pm 0.026^{+0.095}_{-0.000} \\
R_K \quad \text{zentral} \quad &= \quad 0.904 \pm 0.039 \pm 0.008^{+0.022}_{-0.000} \\
R_{K^{*0}} \ \text{tief} \quad &= \quad 0.856 \pm 0.075 \pm 0.032^{+0.071}_{-0.000} \\
R_{K^{*0}} \ \text{zentral} \quad &= \quad 0.954 \pm 0.062 \pm 0.016^{+0.085}_{-0.000}.
\end{aligned}$$

Die ersten beiden Unsicherheiten, die oben genannt werden, sind die statistische und systematische Unsicherheit aus dem simultanen Fit an die Verteilungen der invarianten B Masse. Die dritte Unsicherheit ist als eine konservative Abschätzung der Instabilitäten der R_K und $R_{K^{*0}}$ Verhältnisse unter Anziehen der Elektronen PID Selektionskriterien angefügt. Dieser Effekt wurde Post-Unblinding beobachtet und wird von der LHCb Kollaboration untersucht. Dieses Problem muss vor der Publikation der Ergebnisse gelöst werden.

Zum ersten Mal ist R_K im tiefen q^2 Bereich gemessen. Die Kalibrierung der Signaleffizienzen wurde erfolgreich mit den zugehörigen $B^+ \to K^+ J/\psi (\to \ell^+\ell^-)$ und $B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$ Verhältnissen überprüft. Die Fits an die invariante B Masse wurden mit Pseudo-Experimenten nach möglichen Bias untersucht. Die vorgestellte Analyse ist das Ergebnis einer Gruppenarbeit von mehreren Mitgliedern der Rare Decays Arbeitsgruppe von LHCb, wobei der Autor dieser Doktorarbeit signifikante Beiträge geleistet hat.

Contents

1 Introduction

Today, the Standard Model of Particle Physics (SM) [1, 2, 3] represents the most successful theory to describe nature in terms of fundamental particles and their interactions. Since the discovery of the Higgs boson in 2012 by the ATLAS [4] and CMS [5] collaborations at CERN, every predicted fundamental particle of this theory has been confirmed by experiment. However, it can only be considered as an approximation at low energies of a more global theory, since there are several observations in nature, such as the matter-antimatter asymmetry observed in the universe, that are not explained by the SM.

The Large Hadron Collider (LHC) was constructed to test the established theory and find signs of physics beyond the SM. So far, evidence for New Physics (NP) has eluded detection in direct searches, which is limited by the available energy of the proton beams at LHC. It is possible, that NP particles are too heavy for direct production. The measurements performed by ATLAS and CMS allow to set limits on their masses and couplings to SM processes, see for example Refs. [6] and [7]. Another of the pursued avenues to test the SM is the precision measurement of loop suppressed decays in which NP particles could enter virtually which allows to probe much higher energy scales up to and beyond 100 TeV [8]. These processes are highly suppressed in the SM ($\mathcal{B} \approx \mathcal{O}(10^{-7})$) and referred to as rare decays. Here, NP effects can make significant contributions leading to deviations from the decay rates or angular distribution predicted in the SM. The Large Hadron Collider beauty (LHCb) experiment, one of the four main experiments situated at the LHC, was designed to perform precision measurements of *CP* violation and rare decays by collecting and analysing large datasets of beauty and charm hadron decays.

In the SM, the couplings to leptons of all described fundamental forces are independent from the three lepton generations; a property known as Lepton Flavour Universality (LFU). Over the last years, an increasing interest in LFU tests in $b \to s\,\ell^+\ell^-$ transistions has arisen, where $\ell^+\ell^-$ indicates either a dimuon or a dielectron final state. These processes constitute Flavor Changing Neutral Currents (FCNC) which are forbidden at tree-level in the SM. Thus, they only occur at loop level in so-called penguin or box diagrams (see Fig. 2.4). The ratios of branching fractions of the muon over electron final states are predicted to be unity, if lepton masses can be neglected. The SM predictions are very precise with total uncertainties at the level of 1 % [9], since hadronic uncertainties cancel. Consequently, a measured deviation between theory and experiment in LFU observables would be a clear indicator of NP.

The first test of LFU in $b \to s\,\ell^+\ell^-$ decays at LHCb was performed in 2014, using the mode $B^+ \to K^+\ell^+\ell^-$ [10]. The measured result deviated from SM expectations at the level of 2.5 standard deviations (σ). Subsequent analyses of different decay modes

such as $B^0 \to K^{*0}\ell^+\ell^-$ [11] and $\Lambda_b^0 \to pK\ell^+\ell^-$ [12] as well as of $B^+ \to K^+\ell^+\ell^-$ on a larger data sample [13] have shown similar tensions with the SM expectation. Inspired by these results, various extensions of the SM are explored today, featuring interactions that violate LFU, *e.g.* leptoquark [14] or Z' model [15]. Still, the question remains whether these tensions are caused by actual NP effects or are statistical fluctuations. It is therefore vital to perform further measurements to improve the experimental precision and explore other decay modes.

This thesis presents the first simultaneous measurement of LFU in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays based on the full 9 fb^{-1} Run 1 and Run 2 LHCb dataset, where the branching fraction ratios of the dimuon over the dielectron mode ratios are called R_K and $R_{K^{*0}}$, respectively. The simultaneous approach entails various advantages such as the constraint of the contamination from partially reconstructed $B^0 \to K^{*0}e^+e^-$ decays in the $B^+ \to K^+e^+e^-$ mode. The ratios R_K and $R_{K^{*0}}$ are determined in two regions of the transferred four momentum squared (q^2) of the $b \to s\,\ell^+\ell^-$ process. A key aspect of the analysis are the experimental challenges of electrons at LHCb. Due to their comparably small mass, electrons produce a large amount of bremsstrahlung in contrast to muons. This leads to a significantly reduced resolution of the reconstructed invariant masses of both the dilepton system and the B meson. Further, the trigger system at LHCb relies heavily on the Electromagnetic Calorimeter (ECAL) subdetector for decays to electron final states, whose occupancy is much higher compared to the muon stations that are used to trigger the muon modes. In order to balance the rates, higher thresholds are needed for the electron trigger reducing its efficiency. This reduced yield and poorer resolution are the reason why all analysis aspects are designed to optimise the performance of the electron modes and subsequently are translated to the muon mode in order to align their treatment and improve cancellation of biases in the ratios. For the presented test of LFU, novel approaches to the selection, the fits to the invariant B mass as well as the efficiency calibration have been developed resulting in an analysis with an improved sensitivity compared to its predecessors.

Whilst this measurement is the product of a team of several analysts, the work performed in the course of this thesis made significant contributions to all aspects of the analysis. One of the main contributions was made to the applied selection to the electron modes, which, compared to similar LFU analyses, applies comparatively tight criteria yielding a high signal purity in the datasets. This provides a good control over residual backgrounds in the invariant mass fit allowing to perform the fits in a wide range of the reconstructed B mass in order to extract the signal yields. Another main field of work was the calibration and calculation of efficiencies that, together with the signal yields, are an essential input to calculate the desired ratio observables (for further details see Chap. 4).

The structure of this thesis is as follows: First, an overview of the theoretical context and the experimental status of $b \to s\,\ell^+\ell^-$ transitions focusing on LFU measurements is provided in Chap. 2. Next, the LHCb detector and software environment used to gather the analysed dataset are described in Chap. 3. Chapter 4 describes the analysis of the simultaneous measurement of R_K and $R_{K^{*0}}$ in both regions of q^2. Lastly, a summary of the results together with an outlook to future LFU measurements is given in Chap. 5.

2 Theoretical and Experimental Overview

This chapter first gives a brief overview of the theoretical background based on Refs. [16, 17, 18, 19]. Secondly, the current experimental status and motivation for the analysis are presented.

2.1 The Standard Model of Particle Physics

Since millennia humans try to unravel the mystery that is the universe they live in. A long journey of progress that continues to this day and may never fully end. Today, the area of science that tries to understand the laws of nature at the smallest scales is called Particle Physics. The current state of knowledge about the basic constituents of matter and their interactions is contained in the Standard Model of Particle Physics (SM) [1, 2, 3]. This model is designed to describe elementary particles at high energies and is implemented as a relativistic quantum field theory. The SM encompasses three of the four fundamental forces: electromagnetism, weak interaction and strong interaction; which are all derived by requiring gauge invariance. Only gravity is not contained, but is known to have a negligible impact on the probed processes at the energy scales used in this thesis.

2.1.1 Elementary Particles

In the SM all known fundamental particles are divided into two groups, bosons and fermions. The bosons carry integer spin and are the mediators of the fundamental forces which are topic of the subsequent section. The fermions have spin $1/2$, thus obeying Pauli's exclusion principle [20], which forbids two fermions that are equal in all quantum numbers to occupy the same energy state. They are further divided into the sub-groups of quarks and leptons, which both have three generations of particles as listed in Tab. 2.1. Each generation of quarks contains an up-type quark (u, c, t) with an elementary electric charge (Q) $Q = +2/3$ and a down-type quark (d, s, b) with $Q = -1/3$. Quarks are the only fermions that carry colour charge and thus participate in the strong interaction. The lepton generations each contain a charged lepton (e^-, μ^-, τ^-) carrying $Q = -1$ and a corresponding electrically neutral neutrino (ν_e, ν_μ, ν_τ).

All hadrons are compositions of quarks and antiquarks, most notably the first generation of quarks (u and d) form the nucleons (proton and neutron). Together with the electron from the first generation of leptons, the nucleons form all stable atoms that populate the universe. A direct consequence of the quantum field theory is the existence of a corresponding antiparticle for each particle, so that for example the electron has the

positron as an antimatter partner. Indeed such antiparticles are known to exist ever since the discovery of the positron in 1932 [21]. Particles and their antimatter partner share the same mass and spin, however they carry oppositely signed physical charges, so that the positron carries $Q = +1$.

Table 2.1.: SM fermions of all three generations. The quoted masses are compiled from Ref. [22].

	Quarks			Leptons		
Generation	Particle	Q	Mass	Particle	Q	Mass
1^{st}	u	$+2/3$	$2.16^{+0.49}_{-0.26}\,\text{MeV}/c^2$	e^-	-1	$511.0\ \text{keV}/c^2$
	d	$-1/3$	$4.67^{+0.48}_{-0.17}\,\text{MeV}/c^2$	ν_e	0	$< 1.1\ \ \text{eV}/c^2$ (90 % CL)
2^{nd}	c	$+2/3$	$1.27 \pm 0.02\,\text{GeV}/c^2$	μ^-	-1	$105.7\,\text{MeV}/c^2$
	s	$-1/3$	$93^{+11}_{-5}\,\text{MeV}/c^2$	ν_μ	0	$< 0.19\,\text{MeV}/c^2$ (90 % CL)
3^{rd}	t	$+2/3$	$172.76 \pm 0.30\,\text{GeV}/c^2$	τ^-	-1	$1777\,\text{MeV}/c^2$
	b	$-1/3$	$4.18^{+0.03}_{-0.02}\,\text{GeV}/c^2$	ν_τ	0	$< 18.2\,\text{MeV}/c^2$ (95 % CL)

2.1.2 Interactions

The SM combines Quantum Chromodynamics (QCD), which describes strong interactions, with the Glashow-Weinberg-Salam Electroweak theory (EW), which unites the weak force and electromagnetism (Quantum Electrodynamics (QED)). These theories all share the foundation of invariance under local phase transformation, *i.e.* gauge invariance, and are combined under the symmetry group:

$$\text{SU}(3)_C \otimes \text{SU}(2)_L \otimes \text{U}(1)_Y \ .$$

Demanding gauge invariance of Lagrangians in relativistic field theories leads to the introduction of new vector fields through the covariant derivative. These vector fields relate to the spin 1 gauge bosons of the SM that are listed in Tab. 2.2.

Table 2.2.: SM bosons for all three interactions. The quoted masses are compiled from Ref. [22].

Interaction	Boson	Mass	Couples to
strong	$8g$	0	colour charge
weak	$W^\pm$	$80.38\,\text{GeV}/c^2$	weak charge
	Z	$91.19\,\text{GeV}/c^2$	weak & el. charge
electromagnetic	γ	0	el. charge

Quantum Chromodynamics

QCD is a gauge theory based on a non-abelian symmetry group $\mathrm{SU}(3)_C$, where C denotes colour charge. In the SM there are three different states of colour (`red`, `green`, `blue`) that are conserved in all interactions. The only fermions participating in the strong interaction are the quarks. The masses of the various quark flavours are extremely different as shown in Tab. 2.1. The gauge bosons of QCD are eight gluons which carry a positive and a negative unit of colour and anti-colour. Since they carry colour charge themselves, gluons can interact which each other. Thus, adding to the fundamental process of $q \to q + g$ also gluon-gluon vertices with three- and four-gluon interactions. These primitive vertices are shown in Fig. 2.1[1].

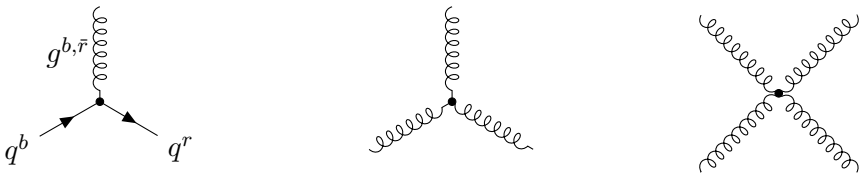

Figure 2.1.: Primitive vertices in QCD. On the left a $q \to q + g$ interaction is shown, where an incoming quark with colour charge `blue` changes to `red`. The emitted gluon carries `blue` and anti-`red` colours. The other diagrams show three- and four-gluon interactions.

The coupling constant of the strong interaction (α_s) is, despite its name, a function of the separation distance of the interacting particles or in other words the transferred four momentum squared (q^2). A phenomenon that is called 'running' coupling. At increasing values of q^2 (small distances) α_s drops until the particles are barely interacting with each other and are called 'asymptotically free'. This property enables to perform reliable perturbative calculations with methods such as the Feynman calculus in QCD at high energies. On the other side, at increasing distances (small q^2) α_s rises and the energy needed to further separate two interacting particles becomes so large, that a new quark-anitquark pair is created. For this reason quarks never occur as free particles but only in colourless bound states (hadrons), a property known as 'confinement'. At low q^2 QCD is in its non-perturbative regime, where methods such as Light-Cone Sum Rules (LCSR) [24] and lattice QCD [25] are employed for theoretical calculation.

Electroweak Theory

The EW theory is a chiral theory classifying fermion fields as left-handed doublets (weak isospin $I = 1/2$) and right-handed singlets ($I = 0$). The $\mathrm{SU}(2)_L$ vector field triplet W^i_μ ($i = 1, 2, 3$) couples only to the left-handed doublets, as indicated by the L, and are associated with the massless W_i bosons. The $\mathrm{U}(1)_Y$ vector field singlet B_μ couples to both chiralities and is associated with the massless B boson that couples to weak

[1] All Feynman diagrams in this thesis are drawn with the `tikz-feynman` tool [23].

hypercharge (Y). The Gell-Mann−Nishijima formula relates the electric charge Q with the third component of weak isospin (I_3) and Y:

$$Q = I_3 + \frac{1}{2}Y\,.$$

The charged $W^{\pm}$ bosons of weak interaction are linear combinations of W^1_μ and W^2_μ:

$$W^{\pm}_\mu = (1/\sqrt{2})(W^1_\mu \mp iW^2_\mu)\,,$$

while the neutral Z and the photon γ are linear combinations of W^3_μ and B_μ:

$$\begin{pmatrix}\gamma\\ Z\end{pmatrix} = \begin{pmatrix}\cos\theta_{\mathrm{W}} & \sin\theta_{\mathrm{W}}\\ -\sin\theta_{\mathrm{W}} & \cos\theta_{\mathrm{W}}\end{pmatrix}\begin{pmatrix}B_\mu\\ W^3_\mu\end{pmatrix}\,,$$

with the weak mixing angle θ_{W} (Weinberg angle). Similar to QCD the EW theory has a non-abelian symmetry group causing direct couplings between $W^{\pm}$ and Z as well as between $W^{\pm}$ and γ. These processes add to the basic primitive vertices of EW theory displayed in Fig. 2.2.

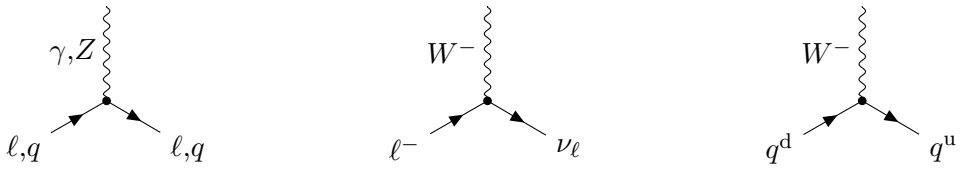

Figure 2.2.: Primitive vertices in EW theory. The ℓ denotes any charged lepton, while $q^{\mathrm{u}}, q^{\mathrm{d}}$ indicate up- and down-type quarks, respectively. Additionally, there are direct couplings between $W^{\pm}$ and Z and between $W^{\pm}$ and γ.

The coupling strength of the EW interactions are defined by the vertex factors entering for each vertex in the Feynman calculus:

$$\text{EM: } ig_e\gamma^\mu, \quad \text{Charged weak: } \frac{-ig_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5), \quad \text{Neutral weak: } \frac{-ig_Z}{2}\gamma^\mu(c_V - c_A\gamma^5)\,,$$

with the gamma matrices γ^μ, the coupling constants $g_e = \sqrt{4\pi\alpha_{EM}}$, $g_W = \frac{g_e}{\sin\theta_{\mathrm{W}}}$ and $g_Z = \frac{g_e}{\sin\theta_{\mathrm{W}}\cos\theta_{\mathrm{W}}}$ as well as the coefficients $c_V = \frac{1}{2}I_3 - Q\sin^2\theta_{\mathrm{W}}$ and $c_A = \frac{1}{2}I_3$. Notice that these factors are independent of the mass and generation of the lepton involved, a property known as Lepton Flavour Universality (LFU) and of great importance for the work presented in this thesis.

The weak fine structure constant $\alpha_W \approx \frac{1}{29.5}$ is about 5 times larger compared to the electromagnetic $\alpha_{EM} \approx \frac{1}{137}$. Clearly the attribute 'weak' is not referring to the strength of the weak interaction but to the fact, that in most processes the transferred energy is small with respect to the mass of the weak bosons ($q^2 \ll (mc)^2$, with $m = m_{W,Z}$). This results in small values of the weak propagator

$$\frac{-i(g_{\mu\nu} - q_\mu q_\nu/m^2\mathrm{c}^2)}{q^2 - m^2\mathrm{c}^2}\,,$$

where $g_{\mu\nu}$ is the metric, compared to other interactions (if allowed). The weak vertex factor contains both a vector (γ^μ) and an axial vector ($\gamma^\mu\gamma^5$) coupling term. Consequently, parity (P) is violated in weak interactions. Electromagnetic and strong interactions on the other hand are pure vector couplings and thus conserve P.

2.1.3 Spontaneous Symmetry Breaking and Higgs Mechanism

The reason why the massless bosons of EW theory combine to form the massive bosons of the weak interaction ($W^\pm$, Z) and the massless boson of electromagnetism (γ) is, that the underlying $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry is broken down to the $\mathrm{U}(1)_Q$ electric charge symmetry. The underlying mechanism is called spontaneous symmetry breaking and is described in the following: The Lagrangian ($\mathcal{L}$) of the EW theory possesses a continuous global symmetry under rotations $\phi \rightarrow e^{i\theta}\phi$, where ϕ is a scalar field (Higgs field). The potential $\mathcal{U}(\phi)$ of the Lagrangian has an infinite set of minima (ground states) for non-zero values of ϕ. The sub-theories of the SM are perturbative theories that consider the various vector fields as fluctuations around the ground state, *i.e.* the vacuum expectation value (VEV). When the Lagrangian is re-expressed as a function of a deviation (η) around the VEV, the transformed Lagrangian is no longer invariant under rotations, since the vacuum does not share the symmetry of the inital Lagrangian. Thus, the symmetry is broken by spontaneously selecting an arbitrary ground state. According to Goldstone's theorem [26], the breaking of a continuous global symmetry creates one (or more) massless scalar field(s), so-called Goldstone bosons. However, there is no candidate for such bosons amongst the known elementary particles.

To solve this dilemma the Higgs mechanism comes into play [27, 28, 29]: By choosing a specific gauge it is possible to transform the Lagrangian in a way, that the massless scalar field disappears. The resulting Lagrangian has a massive gauge field A^μ and a massive scalar η called Higgs boson. This particle was considered the last missing experimental confirmation of the SM until its discovery in 2012 by the ATLAS [4] and CMS [5] collaborations at CERN.

In the minimal form of the SM, the Higgs boson is a SU(2) doublet that couples to the weak bosons, the (massive) fermions via the Yukawa couplings and to itself. Thus, the Higgs mechanism and Yukawa couplings are responsible for generating the masses of all massive particles in the SM. However, it fails to predict the value of these masses, which need to be determined experimentally.

2.1.4 Flavour Sector

In the SM there is no process changing the flavour of leptons, *i.e.* the lepton familiy number is always conserved[2]. The same is true for quarks in all processes except for charged weak interactions. The coupling of $W^\pm$ bosons to quarks can change the quark flavour. The reason for this is, that the $W^\pm$ do not couple to the physical quarks d,

[2]The observation of neutrino oscillations [30] have shown that such processes, while forbidden in the SM, do occur in nature.

s, b which form the basis of diagonal mass states. Instead the charged weak bosons couple to the 'Cabibbo-rotated' states d', s', b' which form the basis where the $W^{\pm}$ interactions are diagonal. These interaction states where first introduced in the Glashow, Iliopoulos and Maiani (GIM) mechanism [31] for the first and second quark generation and later extended to all three generations in the Cabibbo, Kobayashi and Maskawa (CKM) formalism [32]. Here, the unitary matrix V_{CKM} transforms between mass and interaction basis:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} =: V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} .$$

The elements of V_{CKM} are labelled with the quarks involved in the interaction, *e.g.* V_{ud} gives the coupling strength of the $W^{\pm}$ in the u-d vertex. They can be re-parameterised as combinations of $\cos\theta_{ij}$, $\sin\theta_{ij}$ and $e^{i\delta}$, where θ_{ij} are the three real mixing angles and δ is the Kobayashi-Maskawa phase. The complex phase δ is the only source of *CP* violation in the SM (assuming the scalar sector with a single Higgs doublet).

The SM offers no predictions for the values of the V_{CKM} matrix elements, these need to be determined by experiment. It has been found, that the CKM matrix is close to the unit matrix with small off-diagonal elements. Hence, transitions between quark generations can occur through the charged current but are suppressed in the SM, this effect is known as Cabbibo suppression.

Flavour Changing Neutral Currents

Flavor Changing Neutral Currents (FCNC) are processes that change the flavour of a quark while preserving its electric charge. They are forbidden in the SM at one-vertex level (tree-level) for all neutral gauge bosons. In the case of weak interactions they can however occur at loop-level with internal $W^{\pm}$. For instance, FCNCs occur in neutral meson mixing diagrams that change the flavour of both initial quarks as shown in Fig. 2.3 for the $B^0 - \overline{B}^0$ process.

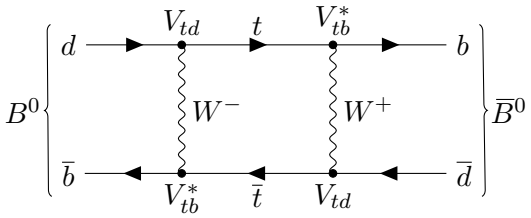

Figure 2.3.: Box diagram of $B^0 - \overline{B}^0$ mixing.

In the diagram, the intermediate quark can also be a u- or c-quark, though their transitions are suppressed by the GIM mechanism, which explains the amplitude suppression of a decay by the mass squared of the internal quark m_i^2/m_W^2 with $i = u, c, t$, so that the t contribution clearly dominates.

Another type of FCNC processes are rare decays that change only one quark flavour of the initial meson (or baryon) while the other quarks are just 'spectators', *i.e.* they do not take part in the interaction. An example are semileptonic $b \to s\,\ell^+\ell^-$ transitions that are represented by so-called 'penguin' and box diagrams as shown in Fig. 2.4. An analogous argument about u and c contributions as for the mixing diagram holds for $b \to s\,\ell^+\ell^-$ transitions.

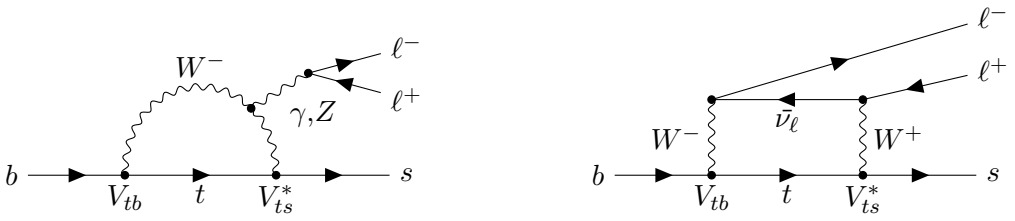

Figure 2.4.: Example penguin (left) and box (right) diagram of a $b \to s\,\ell^+\ell^-$ transition.

Since they only occur at loop-level, involve an off-diagonal CKM element and experience GIM suppression, FCNCs are highly suppressed in the SM. These processes are therefore an exciting field of study because even small effects Beyond Standard Model (BSM) or in other words New Physics (NP) contributions can have a significant impact on several characteristics of $b \to s\,\ell^+\ell^-$ decays, such as the measured branching ratios or angular observables.

2.2 Motivation

The SM is an extremely successful theory which so far has survived every experimental test (with neutrino oscillations being the only exception) and provided numerous predictions which later were confirmed by experiment. An example is the prediction of the c-quark by the GIM mechanism [31] four years before its discovery in 1974 at Brookhaven [33]. However, the SM can not be a complete theory as indicated by various sources of experimental evidence:

- The SM provides no candidate for dark matter nor an explanation of dark energy.
- Neutrino oscillations, while forbidden in the SM, have been measured by the Super-Kamiokande collaboration [30].
- The observed matter anit-matter asymmetry of the universe cannot be explained by the SM.
- The SM does not encompass gravity.

Additionally, the SM contains many free parameters such as quark and lepton masses or the parameters of the CKM matrix, which currently can only be determined by experiment. A complete theory would provide such information on its own and ideally would depend on fewer parameters. A key goal of today's particle physics programme is to

find extensions to the established theory to account for the aforementioned deficits. To this end, experiments perform high precision tests of SM predictions to find deviations between theory and nature. These searches aim for NP processes that involve new particles or interactions. Two strategies are pursued: on the one hand, direct searches try to produce NP particles in collision experiments and detect their characteristic decay signatures. Whilst this method is able to provide indisputable evidence of a NP particle, it is limited by the centre-of-mass energy provided by the collider. Since so far NP particles have escaped direct detection, they are assumed to be very heavy ($\mathcal{O}(\,\text{TeV})$ or higher) and lower limits on their masses have been set. Indirect searches on the other hand perform precision measurements of processes, where the SM contribution is already well understood. Here, possible NP particles can appear virtually in loop processes which allows to probe energy scales much higher than those present in the studied process. Alternatively, new interactions, that couple only weakly to SM particles, can even enter the process at tree-level. Indirect searches are limited by the precision of the measurement. Today, indirect searches for NP effects in rare decays receive major attention from many experimenters and theorists, because of recent experimental results that are discussed in the following section.

2.3 The $b \to s\,\ell^+\ell^-$ Transition

The $b \to s\,\ell^+\ell^-$ transition (see example Feynman diagram in Fig. 2.4) is subject of high interest in todays particle physics programme. As mentioned in Sec. 2.1.4 it is an excellent probe for NP effects because of the small amplitudes of the SM processes. New quantum fields which introduce heavy BSM particles could affect the decay rates or angular distributions of these processes leading to deviations from SM predictions. Quark confinement forces the initial as well as the final state quark to be arranged in a meson (or baryon) bound state. When probing the $b \to s\,\ell^+\ell^-$ transition it is therefore neccessary to study a process such as $B^+ \to K^+\ell^+\ell^-$ or $B^0 \to K^*(892)^0\ell^+\ell^-$. Currently, there are interesting tensions between SM predictions and measurements of differential branching fraction measurements. For example the LHCb Collaboration has measured the branching fractions of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$ as a function of the transferred momentum (q^2) [34, 35]. In the case of $b \to s\,\ell^+\ell^-$ transitions, q^2 is equal to the invariant mass squared of the dilepton system. Figure 2.5 shows the results of these measurements compared to the SM prediction. Particularly at low q^2 tensions between data and theory are visible (at a 1–3 σ level). Similar results recently emerged from measurements of the $B_s^0 \to \phi\mu^+\mu^-$ differential branching fraction [36].

Calculating a SM prediction of the decay rates requires detailed knowledge of the hadronic sector. In practice, calculating the charm loops within the QCD computations of these decays introduces the largest source of theoretical uncertainty. The accuracy of these uncertainties is still being discussed amongst theorists and is an active field of research. Thus, it is not guaranteed whether the tensions of the differential branching fraction measurements stem from actual NP effects or from underestimated hadronic uncertainties.

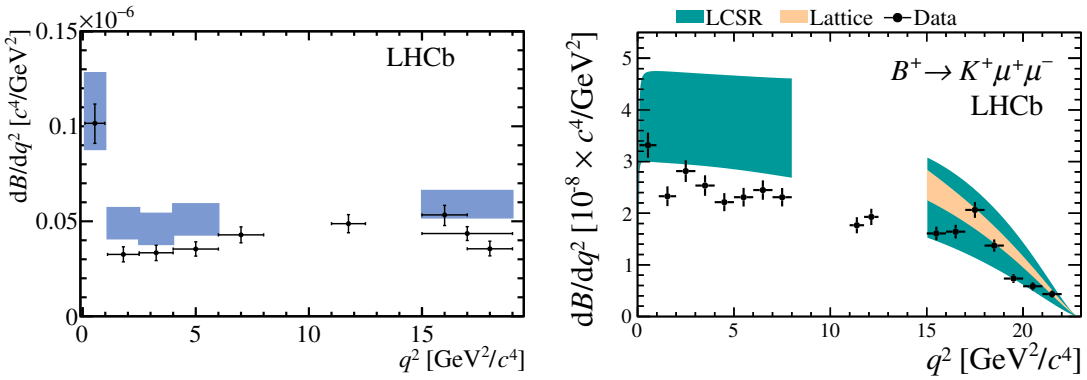

Figure 2.5.: Measurements of $B^0 \to K^{*0}\mu^+\mu^-$ (left) and $B^+ \to K^+\mu^+\mu^-$ (right) performed by LHCb Colaboration [34, 35]. Especially at low q^2 the results show tension with theory.

A way to overcome this obstacle is to construct observables that are (nearly) independent of the hadronic sector, such as ratio measurements: By calculating the ratio of branching fractions of decays sharing the hadronic system but differ in the leptonic part, it is possible to cancel the hadronic uncertainties at first order. These so-called LFU ratios [37] are defined as

$$R_h := \frac{\int_{q^2_{min}}^{q^2_{max}} \frac{\mathrm{d}\Gamma(B \to h\mu^+\mu^-)}{\mathrm{d}q^2}\mathrm{d}q^2}{\int_{q^2_{min}}^{q^2_{max}} \frac{\mathrm{d}\Gamma(B \to he^+e^-)}{\mathrm{d}q^2}\mathrm{d}q^2}\,, \tag{2.1}$$

where h is the final state hadron of the transition ($K^+, K^{*0}, \phi, ...$). The integration runs over a q^2 intervall that is equally defined for muon and electron decays. Because all SM interactions couple with the same strength to all leptons (except Yukawa couplings), as discussed in Sec. 2.1.2, the SM predictions for these ratios are close to unity. Deviations from unity occur at low q^2 ($< 1\,\mathrm{GeV}^2/c^4$) due to the difference of electron and muon masses resulting in different kinematic phase-space as well as from effects induced by light meson resonances such as η or π^0 [38]. In the q^2 region ($0.1\,\mathrm{GeV}^2/c^4 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$) probed in the analysis presented in this thesis, the theoretical uncertainties are at the 1 % level and are dominated by QED corrections [9]. Because of the clean theory predictions, a significant deviation from unity of an experimental measurement would be a clear sign of NP.

The most precise LFU ratio measurements (prior to the analysis presented in this thesis) are the determinations by the LHCb Collaboration of R_K [13] using the full LHCb Run 1 and Run 2 dataset (integrated luminosity of $9\,\mathrm{fb}^{-1}$) and $R_{K^{*0}}$ [11] using the Run 1 dataset only ($3\,\mathrm{fb}^{-1}$):

$$R_K = 0.846\ ^{+\ 0.042}_{-\ 0.039}\,(\mathrm{stat})^{+\ 0.013}_{-\ 0.012}\,(\mathrm{syst}) \quad \text{for } 1.1 \quad < q^2 < 6.0\ \ \mathrm{GeV}^2/c^4 \tag{2.2}$$

$$R_{K^{*0}} = \begin{cases} 0.66\ ^{+\ 0.11}_{-\ 0.07}\,(\mathrm{stat}) \pm 0.03\,(\mathrm{syst}) & \text{for } 0.045 < q^2 < 1.1\ \ \mathrm{GeV}^2/c^4 \\ 0.69\ ^{+\ 0.11}_{-\ 0.07}\,(\mathrm{stat}) \pm 0.05\,(\mathrm{syst}) & \text{for } 1.1 \quad < q^2 < 6.0\ \ \mathrm{GeV}^2/c^4\,. \end{cases} \tag{2.3}$$

The R_K ratio shows a 3.1 σ tension with the SM prediction [39], while for $R_{K^{*0}}$ a 2.4–2.5 σ tension with the prediction [39, 40] is reported. The measurements are summarised and compared to results from other experiments in Fig. 2.6. The analysis presented in this thesis performs a simultaneous measurement of R_K and $R_{K^{*0}}$ in two bins of q^2 using the full Run 1 and Run 2 dataset, which represents the highest statistics LFU test to date and the first measurement of R_K in the $0.1\,\text{GeV}^2/c^4 < q^2 < 1.1\,\text{GeV}^2/c^4$ region.

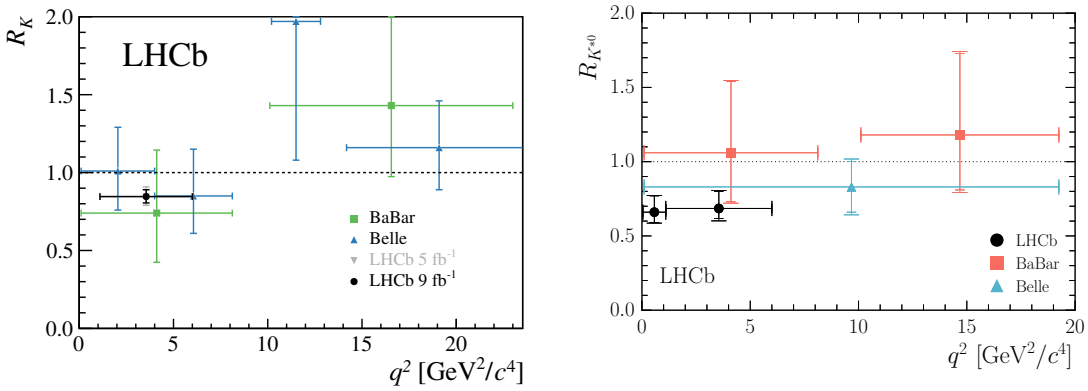

Figure 2.6.: Measurements of R_K [13] (left) and $R_{K^{*0}}$ [11] (right) performed by LHCb Collaboration, which show tension with the SM, compared to the measurements from Belle [41] and BaBar [42] which are compatible with theory but have much larger statistical uncertainties.

In recent years, several measurements of differential branching fraction and angular measurements of muon decay modes as well as LFU tests show tensions with theoretical predictions, they form the so-called 'flavour anomalies' [35, 43, 44, 45, 46, 47, 48, 13, 11, 12, 49]. Currently the most popular extensions of the SM that can explain these tensions between experiment and theory are leptoquark models and Z' models (see *e.g.* Refs. [14, 15]), an exemplary Feynman diagram for each is shown in Fig. 2.7.

Figure 2.7.: Examples for BSM Feynman diagrams contributing to $b \to s\,\ell^+\ell^-$: Left (right) a leptoquark (Z') transition.

2.4 Effective Field Theory

The electroweak $b \to s\,\ell^+\ell^-$ transitions probed in this thesis involve hadrons, consequently QCD effects must be taken into account. Two very different energy scales are involved: The scale of the weak interaction is given by the mass of the $W^\pm$ boson of $m_W \approx$

$80\,\mathrm{GeV}/c^2$, while the QCD scale for hadronic dynamics is given by $\Lambda_{\mathrm{QCD}} \approx 0.2\,\mathrm{GeV}/c^2$, which marks the beginning of the non-perturbative regime of QCD (see Sec. 2.1.2). To simultaneously handle the high energy (or 'short-distance') electroweak and low-energy (or 'long-distance') QCD effects, the framework of an Effective Field Theory (EFT) [50, 51] is employed for theory calculations. The EFT is a generalised representation of the Fermi EW theory, where the electroweak interactions are considered point-like by integrating out the heavy gauge bosons and top quark. In EFT, the resulting effective couplings are given by the so-called Wilson coefficients $\mathcal{C}_i$ together with corresponding local operators $\mathcal{O}_i$ that describe the remaining light SM fields. The operators are of mass dimension six since they describe four-fermion vertices, where each fermion field is of dimension $^3/_2$. While the Wilson coefficients cover the short-distance physics above the defined scale μ, the operators contain the long-distance part below. Both $\mathcal{C}_i$ and $\mathcal{O}_i$ are dependent on the scale used for renormalisation. Typically, the Wilson coefficients are evaluated at the electroweak scale (m_W) to exploit the perturbative QCD regime and subsequently propagated down to μ, which for $b\to s\,\ell^+\ell^-$ transitions is given by the b mass.

Each flavour transition can be expressed via an effective Hamiltonian $\mathcal{H}_{\mathrm{eff}}$ summing over all contributing Wilson coefficients and operators

$$\mathcal{H}_{\mathrm{eff}} = \frac{G_F}{\sqrt{2}} \sum_i \lambda \mathcal{C}_i(\mu)\mathcal{O}_i(\mu) + \mathrm{h.c.}\,, \tag{2.4}$$

where G_F is the Fermi constant and λ contains the CKM factor(s). Based on this, it is possible to write the decay amplitudes for an $i \to f$ transition as

$$A_f = \langle f|\mathcal{H}_{\mathrm{eff}}|i\rangle = \frac{G_F}{\sqrt{2}} \sum_i \lambda \mathcal{C}_i(\mu)\langle f|\mathcal{O}_i(\mu)|i\rangle\,. \tag{2.5}$$

In case of the $b\to s\,\ell^+\ell^-$ transitions studied in this thesis, the employed effective Hamiltonian [52] is given by

$$\mathcal{H}_{\mathrm{eff}}^{b\to s\,\ell^+\ell^-} = -\frac{4G_F V_{tb}V_{ts}^*}{\sqrt{2}} \sum_i \mathcal{C}_i^\ell \mathcal{O}_i^\ell + \mathrm{h.c.}\,, \tag{2.6}$$

where $\mathcal{C}_i$ and $\mathcal{O}_i$ are renormalised at $\mu \approx m_b$ and a lepton specific index $\ell = e,\,\mu,\,\tau$ is added to allow for lepton non-universality. The effective four-fermion coupling of the $b\to s\,\ell^+\ell^-$ transition is sketched in Fig. 2.8.

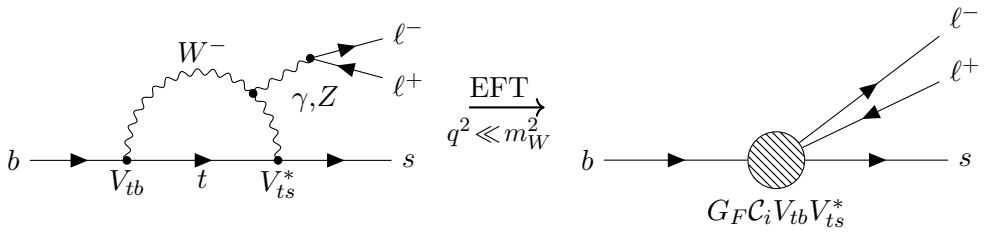

Figure 2.8.: $b\to s\,\ell^+\ell^-$ transition in full theory (left) and effective theory (right).

The most relevant operators contributing to the studied processes are the semileptonic vector- ($\mathcal{O}_9$) and axial vector- ($\mathcal{O}_{10}$) operators defined as

$$\begin{aligned} \mathcal{O}_9^\ell &= (\bar{s}\gamma^\mu P_L b)(\bar{\ell}\gamma_\mu \ell)\,, \qquad & \mathcal{O}_9'^\ell &= (\bar{s}\gamma^\mu P_R b)(\bar{\ell}\gamma_\mu \ell)\,, \\ \mathcal{O}_{10}^\ell &= (\bar{s}\gamma^\mu P_L b)(\bar{\ell}\gamma_\mu \gamma_5 \ell)\,, \qquad & \mathcal{O}_{10}'^\ell &= (\bar{s}\gamma^\mu P_R b)(\bar{\ell}\gamma_\mu \gamma_5 \ell)\,, \end{aligned} \tag{2.7}$$

with the left- and right-handed chiral projectors $P_{L,\,R} = {}^1\!/_2(1 \mp \gamma_5)$. Here, the prime indicates the operators with flipped lepton chirality. At very low q^2 the electromagnetic dipole operator $\mathcal{O}_7$ becomes increasingly relevant for the $b \to s\,\ell^+\ell^-$ transition, where the process is mediated via a nearly-on-shell photon ($b \to s\,\gamma^{(*)}$, $\gamma^{(*)} \to \ell^+\ell^-$)

$$\mathcal{O}_7^\ell = \frac{m_b}{e}(\bar{s}\sigma^{\mu\nu} P_L b) F^{\mu\nu}\,, \qquad \mathcal{O}_7'^\ell = \frac{m_b}{e}(\bar{s}\sigma^{\mu\nu} P_R b) F^{\mu\nu}\,. \tag{2.8}$$

However, this region is largely excluded in the presented analysis to reduce the photon pole contribution (see Sec. 4.1).

In the SM $\mathcal{O}'_{9,\,10} = 0$ since the weak interaction is left-handed, but this does not have to be the case for BSM processes which might have a right-handed component. Additionally, scalar $\mathcal{O}_S$, pseudo-scalar $\mathcal{O}_P$ and tensor $\mathcal{O}_T$ operators which are suppressed in the SM could play an important role for NP effects.

A great feature of Wilson coefficients is that they are universal for all $b \to q$ (where $q = d,\, s$) processes allowing to probe them using measurements of various decay modes and observables. Thus, they represent a tool to combine experimental input from several sources to test the SM with higher precision.

2.5 *b* Hadron Production at LHCb

The heavy quark production at LHCb is governed by the strong interaction via gluon or quark fusion resulting in the creation of quark-antiquark pairs. The Feynman diagrams of the leading order processes are shown in Fig. 2.9.

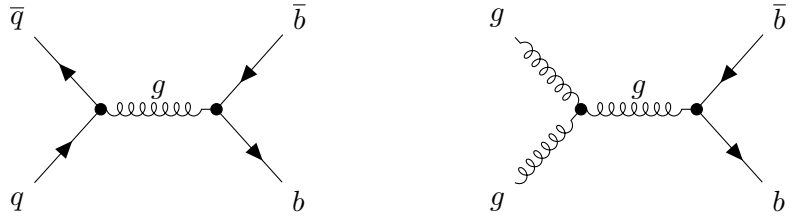

Figure 2.9.: Example leading order Feynman diagrams of $b\bar{b}$ production processes at LHCb: quark-antiquark annihilation (left), gluon-gluon fusion (right).

Because of confinement, the b-quarks immediately combine with lighter quarks to form mesons or baryons after creation. The probabilities to combine to a mesons with an u-, d-, s- or c quark are called fragmentation fractions f_u, f_d, f_s and f_c. In this thesis,

only the relative fraction of f_s over the sum of f_u and f_d is used for background studies. Assuming $f_u = f_d$, this ratio is measured to be

$$\frac{f_s}{f_u + f_d} = \begin{cases} 0.124 \pm 0.010 \quad (7\,\text{TeV}) \;\; [53] \\ 0.122 \pm 0.006 \;\; (13\,\text{TeV}) \;\; [54]\,. \end{cases}$$

Additionally, the ratio with the Λ_b^0-baryon fragmentation fraction, which is the most commonly produced b-baryon at LHCb, is used

$$\frac{f_{\Lambda_b^0}}{f_u + f_d} = \begin{cases} 0.223 \pm 0.036 \quad (7\,\text{TeV}) \;\; [55] \\ 0.259 \pm 0.018 \;\; (13\,\text{TeV}) \;\; [54]\,. \end{cases}$$

The quarks and gluons of the colliding protons have a large range of momenta, since the energy of each proton is distributed over all of its partons. Thus, there is typically a large momentum asymmetry between the quarks or gluons involved in the $b\bar{b}$ creation process resulting in a high boost in either forward or backward direction. Combined with the comparably high lifetime of the $B_{(s)}$-meson, this leads to a displaced decay vertex of the b-hadron which is an extremely useful property to separate b-decays from other particles created in the collision. The distribution of polar angles of the generated $b\bar{b}$ pairs can be seen in Fig. 2.10.

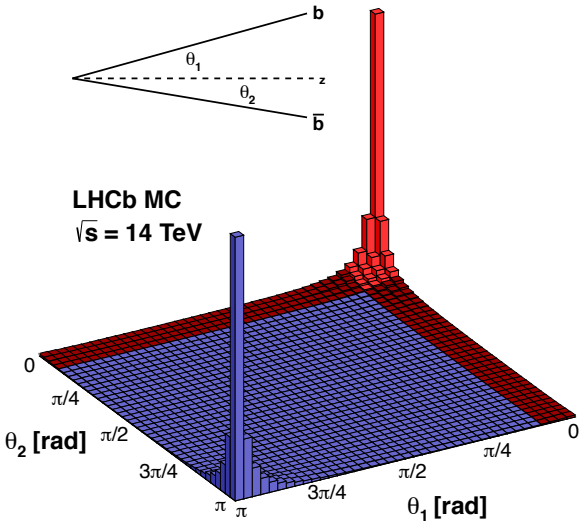

Figure 2.10.: Production rates of the $b\bar{b}$ pairs as a function of the polar angles. The region covered by the LHCb acceptnace is shown in red. Figure taken from [56].

3 The LHCb Experiment

The LHCb experiment [57] is designed to perform a rich flavour physics programme at the LHC at CERN. With precision measurements of *CP* violation and rare decays of beauty and charm hadrons it has excellent tools to test the SM and perform indirect searches for BSM particles.

The analysis presented in this thesis is based on LHCb data taken during 2011–2018. This chapter will first give some key information on the LHC and secondly introduce the LHCb detector and software environment used to gather the analysed dataset, highlighting the aspects that are most important to study $b \to s\,\ell^+\ell^-$ transitions. More detailed reviews of the detector can be found in Refs. [58, 59].

3.1 The LHC at CERN

Currently, the particle collider providing the largest centre-of-mass energy is the Large Hadron Collider (LHC) [60] located at the European Organisation for Nuclear Research (CERN) in Geneva. It is built inside the 26.7 km long tunnel that earlier accommodated the Large Electron-Positron Collider (LEP). Targeting an instantaneous luminosity of $\mathcal{L} = 1 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, the two proton beams of the LHC must have a high intensity. This excludes the usage of anti-protons, since it is not feasible to produce large amounts of anti-matter, and explains the particle-particle collider configuration. Therefore, the counter-rotating beams must traverse opposite magnetic dipole fields in order to stay on their circular tracks. To fit into the spacial limitations set by the LEP tunnel, the LHC uses twin bore magnets allowing the two beams to share the same ring with two separate vacuum chambers. The design energy of 7 TeV per particle beam (resulting in a centre-of-mass energy of 14 TeV) implies a peak dipole field of 8.3 T which requires superconducting magnets. So far the LHC has operated at 7–8 TeV in Run 1 (2011–2012 data-taking periods) and at 13 TeV in Run 2 (2015–2018). Inside the beams, the protons are arranged in bunches of around 10^{11} particles with a spacing of 50 ns in Run 1 and 25 ns in Run 2 corresponding to a bunch crossing frequency of 20 MHz and 40 MHz, respectively.

The two proton beams are brought to collision at four points alongside the ring. These interaction points are the locations of the four largest experiments at CERN: ATLAS [61] and CMS [62] are general-purpose experiments surrounding their beam collision points almost hermetically. The ALICE [63] experiment studies heavy ion physics in lead-lead collisions. The LHCb experiment specialises on beauty and charm physics and will be subject of the subsequent sections.

3.2 The LHCb Detector

The Large Hadron Collider beauty (LHCb) experiment is specialised for heavy flavour physics with the main goal of searching for indirect evidence of New Physics (NP) in *CP* violation and rare decays of hadrons containing a b or c quark. To this end, precision measurements of various processes having a precise SM prediction are performed. LHCb has collected the presently largest dataset of heavy flavour decays by exploiting the large b and c production cross-section and high luminosity at the LHC [60]. However, compared to B factories at e^+e^- colliders these advantages come at the price of the unknown initial state and a high multiplicity hadronic environment, originating from the compound structure of the protons. Typically several hundreds of long-lived particles per pp collision traverse the LHCb detector. In order to control the pile-up[1] and thus the detector occupancy, the instantaneous luminosity is kept at a constant value of about $\mathcal{L} = 4 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ by adjusting the beta function of the colliding beams. To date, the cumulative integrated luminosity collected by LHCb is about $9\,\mathrm{fb}^{-1}$ (see Fig. 3.1).

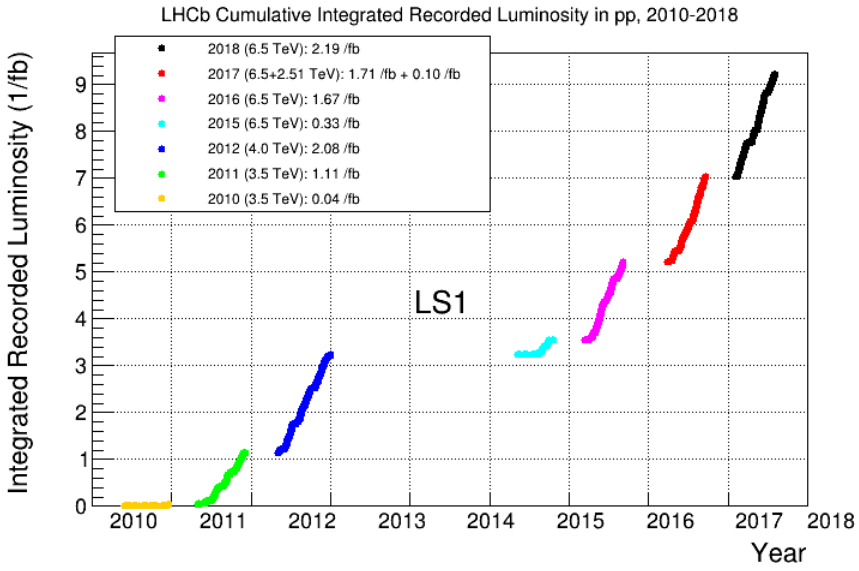

Figure 3.1.: Integrated luminosity collected by LHCb during Run 1 and Run 2. Figure taken from [64].

The LHCb detector is built as a single-arm forward spectrometer with an angular coverage in forward direction from 15 mrad to 300 (250) mrad in the bending (non-bending) plane which corresponds to about 4 % of the solid angle. However, because of the high correlation and boost of the b- and $\bar{b}$-hadron production (see Sec. 2.5), about 25 % of all produced $\bar{b}$-pairs fall into this acceptance. The detector is composed of several subdetectors that are each specialised to provide information on position, identity or energy of the particles traversing them. The subdetectors are briefly described in the following sections. The LHCb detector layout is shown in Fig. 3.2. The figure also shows how the coordinate frame of LHCb is defined: the z-axis coincides with the beam axis and points

[1] Number of pp collisions in a single bunch crossing.

toward the muon stations. Perpendicular to the z-axis, the horizontal x- and vertical y-axis form a right-handed coordinate system.

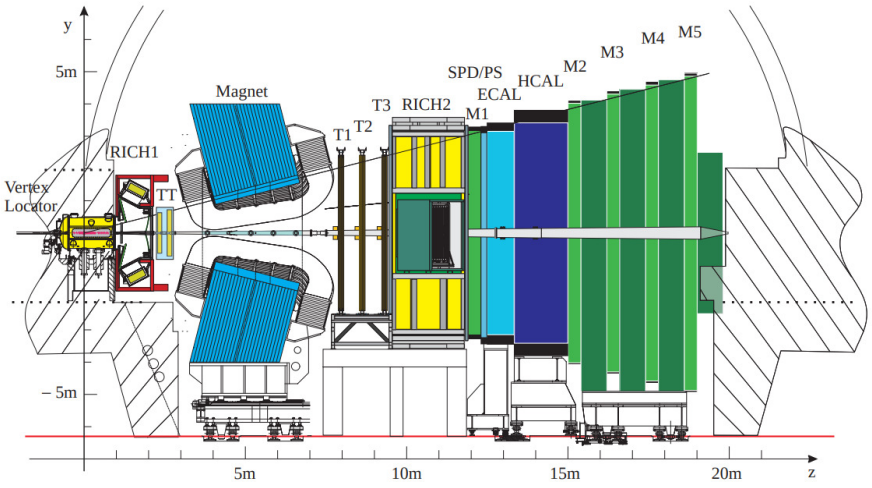

Figure 3.2.: Layout of the LHCb detector. Figure taken from [58].

3.2.1 Tracking System

Accurate momentum resolution and vertex reconstruction are crucial requirements for data analysis at LHCb. These quantities are provided by a tracking system, which determines the trajectories and momenta of charged particles. The measurement of momenta is realised by bending the trajectories of charged particles using a dipole magnet. The reconstruction of trajectories is based on information from three tracking subdetectors that measure the location of traversing particles: the VELO surrounding the interaction region of the colliding proton beams, the TT placed upstream of the magnet as well as the downstream tracking stations T1-T3. The following subsections will provide further details on the magnet and tracking subdetectors as well as into the reconstruction of tracks.

The LHCb Magnet

The magnet used in the LHCb spectrometer is a warm dipol magnet providing an integrated field over 10 m of about 4 Tm. Charged particles traversing the magnetic field are predominantly bent in the x-z plane. Reconstructing the bent trajectories allows to determine the sign of the particles electric charge as well as to estimate its momentum. The magnet polarity is regularly changed during each data-taking period to reduce the systematic uncertainty stemming from detector asymmetries.

The VELO

The Vertex Locator (VELO) is a silicon vertex tracker, which is the subdetector located closest to the beam interaction point. The active region of the detector is placed inside a vacuum vessel which is separated from the LHC vacuum by a corrugated aluminium foil (also referred to as RF-foil). While this design is chosen to minimise the amount of material traversed by the particles, the foil still contributes significantly to the total material budged. The VELO is composed of two halves both with 21 silicon modules arranged along the beam axis (see Fig. 3.3). Both sides of each module are equipped with a sensor, the one side measuring the radial distance r the other the azimuthal angle ϕ. Because of the larger beam aperture during injection, the modules can be retracted to prohibit damage to the sensitive regions. During data-taking, the active region of the modules approaches the beam to a distance of 8 mm. The small distance to the interaction point and high resolution provided by the VELO allows for precise measurement of displaced secondary vertices, which are a key signature of b- and c-hadron decays.

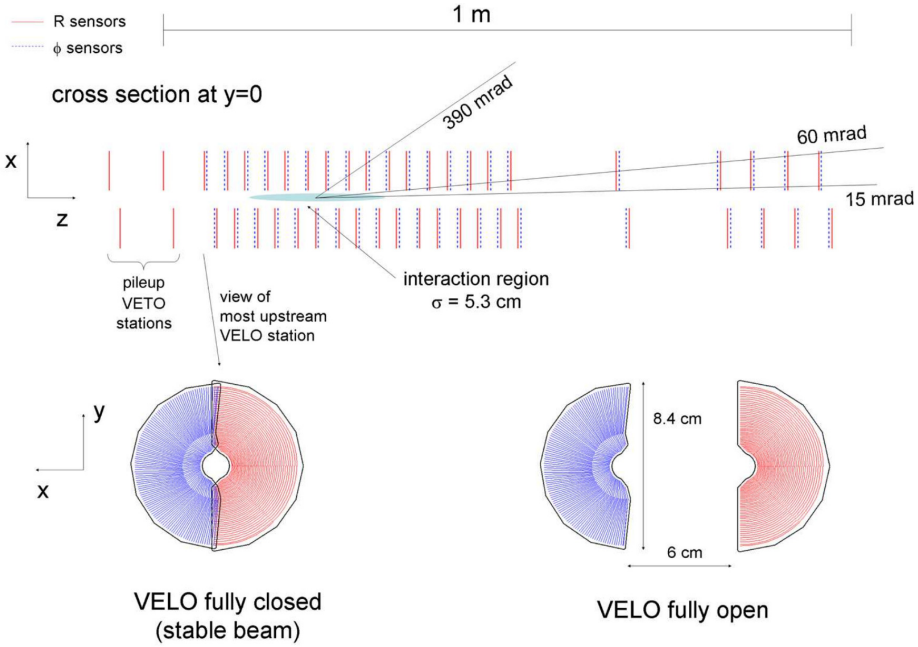

Figure 3.3.: The top figure shows a cross section in the x-z plane of the VELO silicon modules in closed position. The bottom figure shows a front view of a single silicon module in fully closed (left) and fully opened position (right). Figure taken from [57].

Silicon Tracker

There are two Silicon Tracker (ST) consisting of silicon strips installed at LHCb: the Tracker Turicensis (TT) located upstream of the magnet and the Inner Tracker (IT) that

forms the inner part of the downstream tracking stations T1-T3. Each of the four ST stations consists of four detector layers, where the first and last layers are vertical while the second and third are rotated by $-5\,^\circ$ and $+5\,^\circ$, respectively. This so-called x-u-v-x arrangement allows for a three-dimensional track reconstruction with an enhanced track resolution in the x-z bending plane, which is most important for the momentum resolution. The intrinsic hit efficiency of both STs is determined to be greater than 99.7 % while the single hit resolution is found to be about 50 µm; the average occupancy in the sectors closest to the beam pipe which are exposed to the highest particle density does not exceed a few percent [59].

Each of the TT layers is composed of several half modules together covering the full LHCb detector acceptance. A single module covers half the height of the acceptance and is equipped with seven silicon sensors that are divided into two or three different readout groups. The readout electronics are mounted outside at the top and bottom of the layers and fall outside the acceptance, thus minimising the material budged and resolution penalties due to Multiple Scattering (MS). The layout of a full TT layer is displayed in Fig. 3.4.

The three stations of the IT are each embedded in four detector boxes as shown in Fig. 3.4. The boxes contain the four detection layers arranged in a x-u-v-x layout, where each layer is divided into seven detector modules.

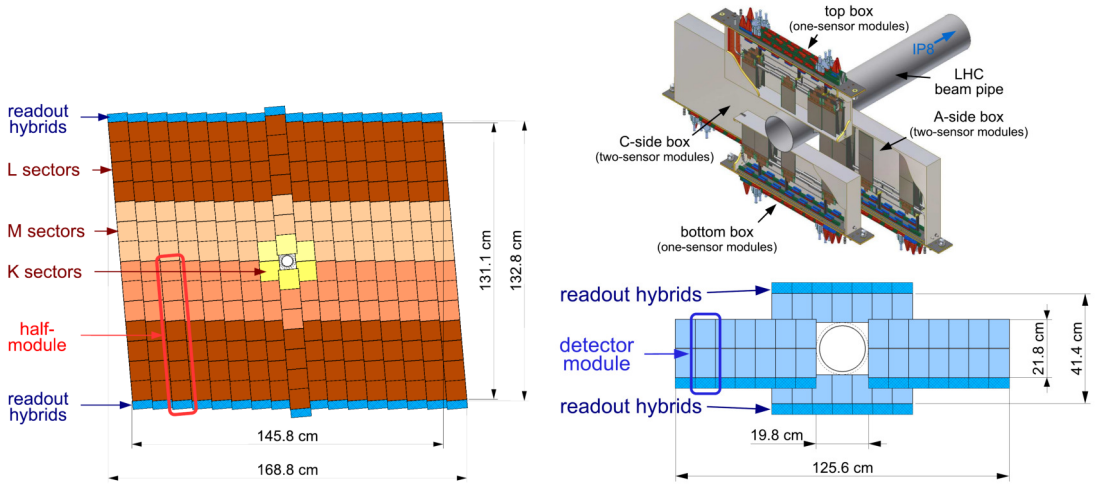

Figure 3.4.: The left figure displays the layout of the v layer of the TT showing the different readout sectors in different colours. The right figure shows the arrangement of the four detector boxes around the beam pipe making up one station of the IT (top) and the layout of the modules in an x layer of the IT (bottom). Figures taken from [57].

Outer Tracker

The Outer Tracker (OT) covers the large outer area of the downstream tracking stations. It is a straw tube drift-time detector built as an array of gas-tight modules, that each contain two layers of drift-tubes of 4.9 mm diameter. The tubes are filled with a gas mixture of Argon (70 %) and CO_2 (30 %). Charged particles traversing the gas-tubes ionise gas-molecules which are electrically attracted by the anode. Because of the synchronisation to the collision time, it is possible to determine the position of the particle hit by measuring the drift time of the ionised clusters. The OT reaches a spatial resolution of about 200 µm and a drift time of less than 50 ns.

Track Reconstruction

Various types of particle tracks are defined at LHCb depending on the subdetectors contributing to their reconstruction (see Fig. 3.5):

- **Long tracks** have hits both upstream (VELO and optionally TT) and downstream (T1-T3) of the magnet. Therefore they have the most precise momentum estimate and are most commonly used by physics analyses.
- **Upstream tracks** in general originate from low momentum particles that traverse the VELO and TT after which they are deflected outside the LHCb acceptance by the magnetic field.
- **Downstream tracks** are important for the reconstruction of long-lived particles such as K_S^0 that decay downstream of the VELO and are built from hits in the TT and T1-T3.
- **VELO tracks** only have hits in the VELO modules. They are useful to reconstruct the primary vertices.
- **T-tracks** are reconstructed only from hits in the T1-T3 stations.

The analysis presented in this thesis only considers long tracks. These are reconstructed starting with the search for a straight line from at least three hits in both r and ϕ sensors of the VELO. Afterwards, information from the T1-T3 stations is added to the VELO tracks following two strategies:

- **Forward tracking:** First, a trajectory is formed from the VELO track and a single hit in the T stations. This trajectory is extrapolated to the other T-stations to find additional hits in a defined search window.
- **Track matching:** A standalone algorithm is used to form track segments in the T1-T3 tracking stations. These segments are combined with the VELO track by extrapolating both to the bending plane of the magnet.

After a set of quality requirements are applied, the long track candidates of both algorithms are combined. This yields the final set of long tracks used for offline analysis. In the last step, hits in the TT are added if they are consistent with the trajectory of a long track.

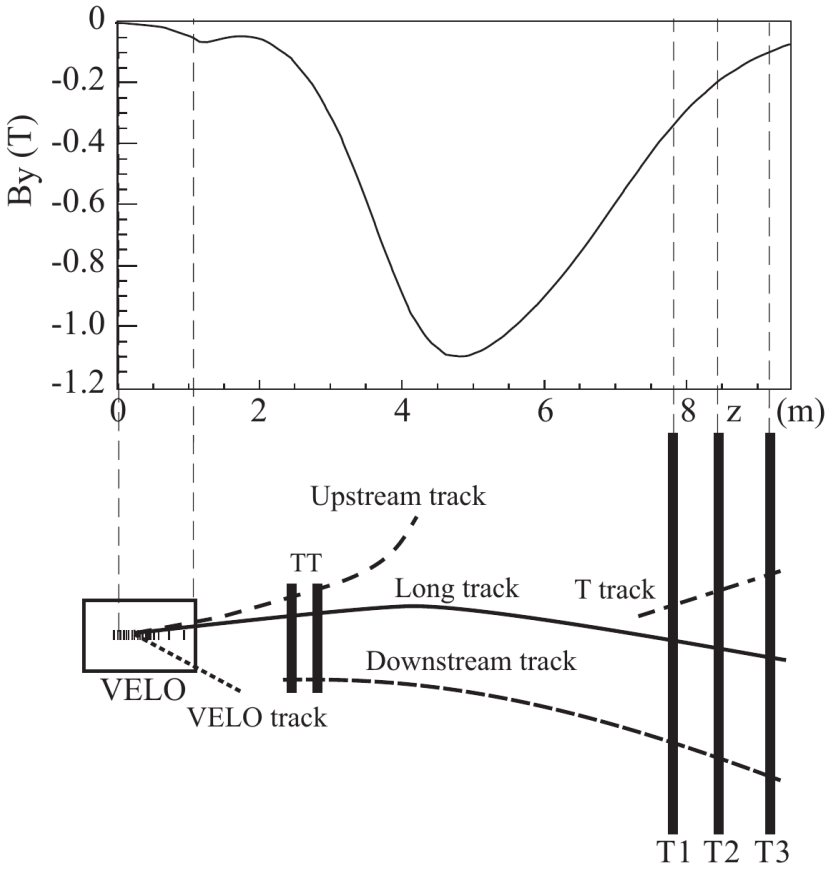

Figure 3.5.: Illustration of subdetectors contributing to the various track types defined at LHCb. The main component of the magnetic field (B_y) is plotted above. Figure taken from [59].

The tracks are fitted with a Kalman filter [65] taking into account effects from material interactions. The quality of the reconstructed track is measured by the $\chi^2/_{\text{ndof}}$ of the fit. If two (or more) tracks share more than 70 % of the hits, the track of the best quality is chosen.

The efficiency of reconstructing a charged particle traversing the full tracking system is determined via the `tag&probe` method based on $J/\psi(\mu^+\mu^-)$ decays, where one of the muons (`tag`) is fully reconstructed while the other (`probe`) is only partially reconstructed using a defined subset of the tracking system allowing to probe the remaining tracking subdetectors. If a fully reconstructed long track is successfully matched to the partially reconstructed `probe` it is defined as efficient. Both on data and simulated events, the efficiency is measured to be greater than 95 % for particles with a momentum of $5\,\text{GeV}/c < p < 200\,\text{GeV}/c$. The momentum resolution $\frac{\delta p}{p}$ probed with the same decay mode is found to be 0.5 % and 0.8 % for particles with a momentum of $20\,\text{GeV}/c$ and $100\,\text{GeV}/c$, respectively. The rate of ghost tracks, which are mis-reconstructed tracks that do not correspond to a real particle, is found to be about 7 % [59].

Primary Vertex and Impact Parameter

Because of the relatively high lifetime and large boost of the B-meson, the displacement of the decay vertex of the b- or c-hadrons (Secondary Vertex (SV)) from the location of the pp interaction (Primary Vertex (PV)) is a crucial feature in order to separate signal tracks from the high background level originating from the underlying event. Accurate tracking information is required to determine both PV and the Impact Parameter (IP)[2] of the reconstructed tracks. The reconstruction of primary vertices begins with a search for seeds, *i.e.* locations where a high number of tracks are close to each other. Subsequently, an iterative fitting procedure to every found seed is performed during which every track is assigend a weight according to the change in χ^2 of the PV fit if the track is added (χ^2_{IP}) [66].

The IP resolution is driven by the track quality, material interactions such as MS and the distance between the PV and the first hits used to form the track. The latter is minimised by the close distance of the VELO modules and the interacting beams.

3.2.2 Particle Identification System

Besides a good track reconstruction, LHCb is optimised to provide a high separation power of the various particle species. The Particle Identification (PID) system of LHCb consists of the calorimeter system, muon stations and most prominently two Ring Imaging Cherenkov Detectors (RICH) detectors, which are a unique feature amongst the four large experiments at the LHC. The RICH detectors are most important to identify charged hadrons, while for muons and electrons the key discrimination power is provided by the muon stations and the calorimeter system, respectively. The following subsections will introduce the detector components contributing to the PID system.

RICH Detectors

LHCb is equipped with two RICH detectors: RICH 1 is located between the VELO and TT while RICH 2 is located downstream of the T1-T3 stations. To cover the full momentum spectrum, the two RICH detectors are designed for different momentum ranges: RICH 1 is filled with C_4F_{10} gas, covering charged particles in the low momentum range of 1–60 GeV/c; RICH 2 on the other hand uses a CF_4 radiator and is designed to cover high momentum particles from 15 GeV/c to 100 GeV/c and above. In contrast to RICH 1, the RICH 2 detector does not cover the full LHCb acceptance but only the region closer to the beam pipe ($\pm$15 mrad to $\pm$120 mrad(horizontal) and $\pm$100 mrad(vertical)), where most particles of high momentum traverse.

The RICH detectors exploit the property of charged particles travelling through a dielectric medium with a velocity faster than the speed of light in that medium to produce Cherenkov radiation [67]. Cherenkov radiation is created in a cone around the particle track with an opening angle of $\cos\theta = 1/n\beta$, where n is the refractive index of the medium

[2] The IP is defined as the distance of closest approach of a track from the PV.

and $\beta = v/c$ is the velocity of the particle. Thus, measuring the Cherenkov angle θ allows to estimate the velocity which together with the momentum estimate from the tracking system is used to calculate the mass of the particle, which allows to identify the particle species. Figure 3.6 shows the Cherenkov angle as a function of the momentum for different particles species. This figure also shows the kinematic limitations of the RICH detectors: While at low momenta the velocity of kaons and protons does not suffice to create Cherenkov radiation, the θ vs. p distributions overlap at high momenta.

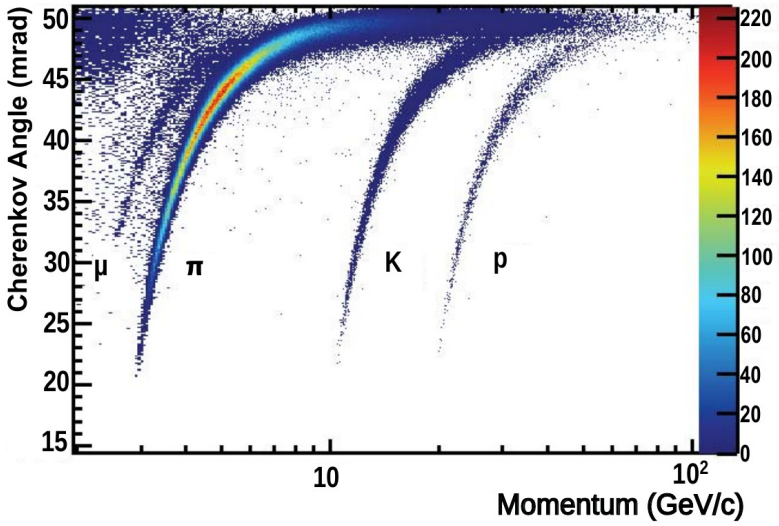

Figure 3.6.: Cherenkov angle θ as a function of the momentum for different particle species in the C_4F_{10} radiator (RICH 1). Electrons form the cluster at the low momentum and high Cherenkov angle region (top left). Figure taken from [68].

To measure the angle θ, the emitted photons are focused and guided by a combination of spherical and planar mirrors onto photo multipliers that are placed outside LHCb's acceptance. This design enables to reduce the material budget and to protect the photo multipliers from radiation damage.

The known momentum of a track is used to compute the Cherenkov angle for all possible mass hypotheses (electron, muon, pion, kaon and proton). For each hypothesis a likelihood $\mathcal{L}$ is calculated based on the probability that the detected photons were emitted with the respective Cherenkov angle. An important variable for analyses is the difference of the logarithmic likelihoods (`DLL`) to separate two particle species such as kaons from pions:

$$\texttt{DLL}_{K,\pi} := \log \mathcal{L}(K) - \log \mathcal{L}(\pi) \ .$$

The PID performance of the RICH system is determined on control samples where the particle species are already known from tracking information. Exclusive decay modes with large branching fractions such as $J/\psi \to \ell^+\ell^-$ with $\ell = e, \mu$, $K^0_S \to \pi^+\pi^-$, $\Lambda \to p\pi^-$ and $D^{*+} \to D^0\ (\to K^-\pi^+)\pi^+$ are used to gather these samples. The kaon identification and pion misidentification efficiencies as a function of the particle momentum are shown

in Fig. 3.7 for different selections on $\texttt{DLL}_{K,\pi}$. Averaged over the momentum range $2-100\,\text{GeV}/c$ the kaon efficiency is measured to be about 95 % at a pion misidentification rate of 10 % for $\texttt{DLL}_{K,\pi} > 0$ [68]. The high PID efficiency is crucial to suppress the large background from pions. Also commonly used for PID selection is the so-called `ProbNN` variable, which is calculated from a neural network trained on information from all subdetectors [69]. Similarly to `DLL`, a separate training is performed for each particle hypothesis. Because of the additional information the `ProbNN` variables typically have a higher selection power, see Fig. 3.7.

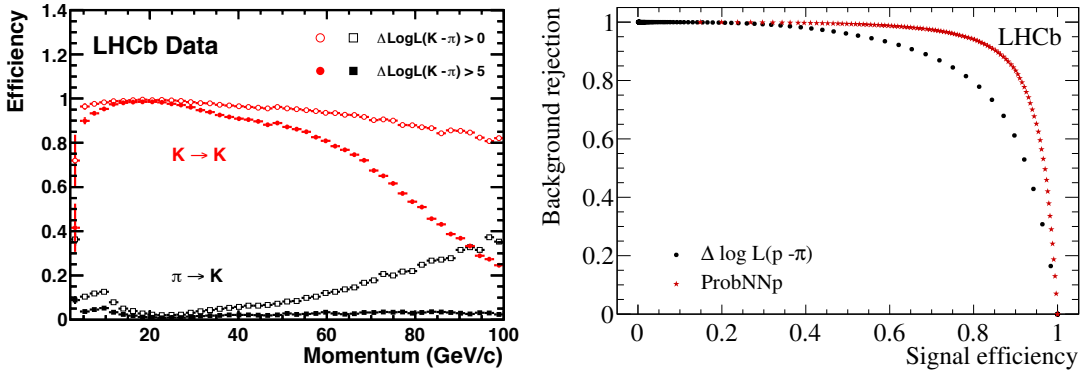

Figure 3.7.: The left figure shows the kaon identification (red) and pion misidentification (black) efficiencies as a function of the particle momentum. On the right, the comparison of signal and background rejection efficiency between `DLL` and `ProbNN` variables for protons is displayed. Figures taken from [70].

3.2.3 Calorimeter System

The calorimeter system at LHCb is located downstream of the RICH 2 detector and the first muon station (M1) and consists of the Scintillating Pad (SPD), Pre-Shower (PS), Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL) subdetectors. Its main functions are selecting high transverse energy (E_T) deposits for the Level-0 Hardware Trigger (L0), which is the first trigger stage of LHCb, as well as providing information on electron, photon and hadron identity. This section focusses on the desgin and PID performance of the calorimeter system based on [71, 72], while the trigger aspects are detailed further in Sec. 3.3.1.

The SPD and PS are identical detectors consisting of scintillator pads. They are designed to separate electrons, photons and pions by exploiting the characteristic differences of the induced electromagnetic showers, see Fig. 3.8. The thickness of the lead plate separating the two detectors is chosen such that electrons and photons begin to shower much more often than hadrons (2.5 radiation lengths (X_0) vs. 6 % hadronic shower length). Photons, that in contrast to charged particles traverse undetected by the tracking system and the SPD, begin to shower in the lead plate and produce a signal in both PS and ECAL. Consequently, ECAL clusters with neither track nor SPD cluster associated are identified

as photons. Summarised, the SPD separates neutral from charged particles, while the PS separates electrons and photons from hadrons.

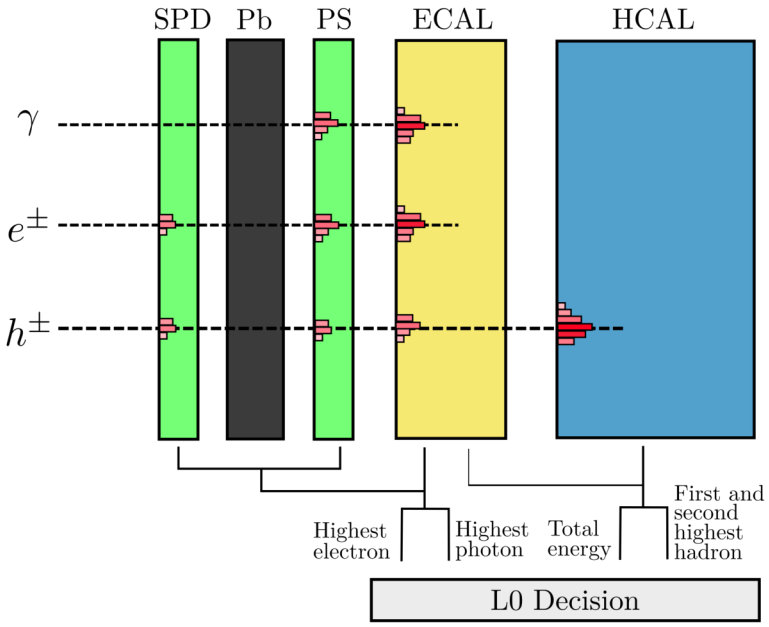

Figure 3.8.: Scheme of the subdetectors of the calorimeter system and their input about measured cluster energy for the L0 trigger. An illustration of photon, electron and hadron tracks and their energy deposits is shown.

Conventionally, the calorimeters are built as a succession of absorber material and active regions. Traversing particles induce showers in the dense absorber material, which is lead and iron for the ECAL and HCAL, respectively. The shower particles create light in active regions made of scintillating material. To cope with the higher track density close to the beam axis, the scintillator plates possess a higher granularity, *i.e.* smaller sizes of active cells, in the inner area. The light is guided by wavelength-shifting fibres towards photomultipliers for detection and readout.

The ECAL is designed as a so-called 'shashlik' calorimeter, *i.e.* a sampling structure of alternating layers of scintillator tiles (4 mm) and lead absorber (2 mm). The total length of 84 mm corresponds to 25 X_0, ensuring the full containment of electromagnetic showers. The resolution of the ECAL for electrons is found to be

$$\frac{\sigma(E)}{E} = \frac{(9.0 \pm 0.5)\,\%}{\sqrt{E}} + 1\,\% \;,$$

where E is the energy of the particle in GeV [72]. Hadrons on the other hand only deposit a small fraction of their energy in the ECAL. Their main shower occurs in the HCAL having a thickness of 5.6 hadronic interaction lengths. Because of its moderate

resolution, the main purpose of the HCAL is to provide hadronic high energy clusters for the L0 trigger. The energy resolution of the HCAL for pions is [72]

$$\frac{\sigma(E)}{E} = \frac{(67 \pm 5)\,\%}{\sqrt{E}} + 10\,\% \ .$$

3.2.4 Muon System

Efficient triggering and identification of muons is crucial for the rich analysis programme of decays to muon final states. To this end, a muon system consisting of five muon chambers is installed (see Fig. 3.9 left). The muon stations M2-M5 are located downstream of the HCAL and form the outermost part of the spectrometer. Since these are separated by 80 cm thick lead absorbers, muons are the only SM particles (except neutrinos) that can reach the last stations. Combined with the calorimeter material this yields a total of 20 interaction lengths, which requires a muon of at least 6 GeV to reach the last station. M1 is located downstream of RICH 2 and is used to improve the p_{T} estimate for the muon L0 trigger. Each station is divided into four regions providing a finer segmentation in areas closer to the beam axis as shown in Fig. 3.9 right. To reach the designed goal of a 95 % efficient muon trigger, the single hit efficiency of each muon chamber needs to be above 99 % [73]. This requirement is met by using Multiwire Proportional Chambers (MWPCs) in all muon stations over the full acceptance of the spectrometer. Only the inner region of M1 is equipped with Gas Electron Multiplier (GEM) detectors instead, to cope with the higher particle density.

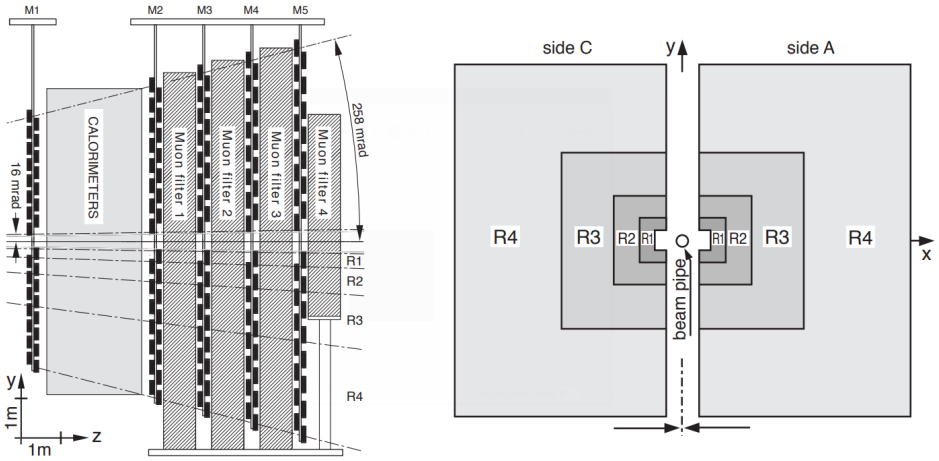

Figure 3.9.: Cross-section of the muon system (left). Layout of a single muon station displaying the four regions of granularity (right). Figure taken from [73].

Three steps are applied to a track associated with a muon candidate to ensure a high muon identification efficiency [74]:

1. A search for hits in the M2-M5 stations in a window around the extrapolated track is performed. Since it is necessary to penetrate both calorimeter and absorbers

between muon stations, this requirement reduces the hadron misidentification to percent level while maintaining a high muon efficiency.

2. A likelihood is calculated based on the distance between M2-M5 hits and the extrapolated track under both muon and non-muon hypothesis.
3. The combined likelihood including calorimeter and RICH information is calculated for each mass hypothesis. The resulting $\texttt{DLL}_{\mu,\pi}$ is used as discriminating variable in offline analyses.

3.3 Data Flow at LHCb

To process and store the data of every proton-proton collision occurring at the 40 MHz bunch crossing rate of the LHC would result in huge CPU and storage costs. In fact, this is an impossible task for LHCb in its current configuration because of limitations by subdetector response times and read-out electronics. Therefore, a data flow is in place that reduces the data rate by orders of magnitude to a level that is economical for storing data to disk. Since most of the events contain QCD interactions that are of no interest to the LHCb physics programme, the reduction of data does not entail an immense loss for offline[3] analyses because efficient selections are in place that exploit signatures typical for b- or c-hadron decays. The following sections introduce the stages of the data flow encompassing a trigger system, including a hardware and two online software trigger stages, and several offline applications. In addition, the LHCb simulation is introduced that uses a processing chain similar to real data. The full data flow is visualised in Fig. 3.10.

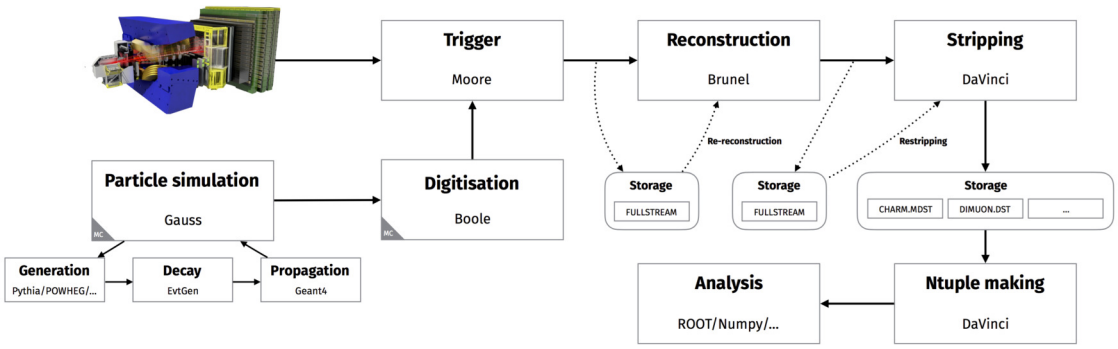

Figure 3.10.: The data flow with the assiciated applications of both real data and simulation. Figure taken from [75].

3.3.1 Trigger System

A multi-level trigger system is used at LHCb in order to reduce the memory usage and to filter out events that are not relevant for offline analysis while retaining events containing

[3]Online algorithms are run in parallel to the data-taking, while offline algorithms are run afterwards.

b- or c-hadron decays. The trigger system successively reduces the data rate from the 40 MHz bunch crossing rate of the LHC down to about 5 kHz and 13 kHz for Run 1 and Run 2, respectively. Various trigger classification algorithms, so-called trigger lines, are implemented each introducing a different set of requirements. To gain flexibility, the configurations and rates of the lines are adjustable. These settings are stored in a Trigger Configuration Key (TCK) that is unique for every configuration. The trigger stages are described in the following sections based on Refs. [76, 77, 78].

L0 Hardware Trigger

The first stage is the L0 trigger that is fully implemented in hardware and designed to operate synchronously with the interaction rate. Based on information from the calorimeter and muon systems, which are the only parts of the detector that can be read out at the bunch crossing rate, it searches for events containing a high transverse energy cluster or a high momentum muon track, since these events are more likely to include a heavy flavour decay. The L0 trigger reduces the data rate to about 1 MHz at which the full detector can be read out so that later trigger stages have more information and computation time at their disposal.

The L0 calorimeter trigger searches for high E_{T} deposits in clusters of 2×2 cells in the ECAL or HCAL. Using information from SPD, PS, ECAL or HCAL it classifies clusters as photons, electrons or hadrons and checks whether the E_{T} deposit exceeds the defined threshold of the `L0Photon`, `L0Electron` or `L0Hadron` trigger lines, respectively, see Fig. 3.8.

The L0 muon trigger is very efficient because of the clean signals from the muon stations. It searches for straight tracks in the five muon stations. The transverse momentum p_{T} of the candidate is estimated by reconstructing the track direction under the assumption that the particle originates from the interaction point. The two muon candidates having the highest p_{T} are used to form the trigger decision based on the highest muon only (`L0Muon`) or on the transverse momentum product of both (`L0DiMuon`).

To reduce the computation time in the subsequent software trigger stages, a requirement on the event multiplicity is applied by limiting the number of hits in the SPD detector to $\texttt{nSPDHits} < 600$ in Run 1 and $\texttt{nSPDHits} < 450$ in Run 2. Within most years the used trigger thresholds are changed between data-taking periods to optimise data throughput. The thresholds used to acquire most of the data in a given year are shown in Tab. 3.1 for the L0 triggers used in this thesis.

High Level Trigger

Events passing the L0 trigger are further processed by the two High Level Trigger (HLT) software trigger stages. Here, more complex decisions are performed to further select events that contain decays of interest for offline analyses. The HLT trigger lines are composed of a set of reconstruction and selection steps. An event is selected by the HLT trigger, if at least one trigger line from both HLT stages returns a positive decision. In

Table 3.1.: Thresholds of the main L0 trigger lines used for the majority of the data-taking in each year. The values are extracted from the TCKs using the `Moore` application [79] of LHCb.

Year	`L0Muon` [GeV/c]	`L0Electron` [GeV]	`L0Hadron` [GeV]
2011	$p_{\mathrm{T}} > 1.48$	$E_{\mathrm{T}} > 2.50$	$E_{\mathrm{T}} > 3.50$
2012	$p_{\mathrm{T}} > 1.76$	$E_{\mathrm{T}} > 3.00$	$E_{\mathrm{T}} > 3.70$
2015	$p_{\mathrm{T}} > 2.80$	$E_{\mathrm{T}} > 2.70$	$E_{\mathrm{T}} > 3.60$
2016	$p_{\mathrm{T}} > 1.80$	$E_{\mathrm{T}} > 2.40$	$E_{\mathrm{T}} > 3.70$
2017	$p_{\mathrm{T}} > 1.35$	$E_{\mathrm{T}} > 2.11$	$E_{\mathrm{T}} > 3.46$
2018	$p_{\mathrm{T}} > 1.75$	$E_{\mathrm{T}} > 2.38$	$E_{\mathrm{T}} > 3.75$

order to deal with the high computational effort, the processing is done by the Event Filter Farm (EFF) consisting of 900 nodes. Because of the higher trigger rates, 800 additional nodes were added for Run 2. The software responsible for the High Level Trigger (HLT) is the `Moore` application [79].

The first stage, called HLT1, performs a partial reconstruction of the event. The main goal of HLT1 is to find displaced long tracks with a high p_{T}, as they are a signature of B- and D-meson decays. To this end, a full three-dimensional pattern recognition is performed to reconstruct VELO tracks. Candidates for PV's are formed from vertices that have more than five VELO tracks associated. Subsequently, displaced VELO tracks are selected by requiring a minimum χ^2_{IP} to all PV's. Selected VELO tracks are extrapolated to the T1-T3 stations to form long tracks, where the applied tracking algorithm uses a simplified description of the material budget and fewer iterations with respect to the offline tracking described in Sec. 3.2.1. This leads to a slightly worse resolution but reduces the computational effort, which is always a key factor for the HLT trigger. Additionally, HLT1 performs a fast muon identification for events selected by the L0 muon trigger by searching for hits in the muon stations that can be matched to the track of the muon candidate.

The HLT2 performs a full event reconstruction based on information from all subdetectors. For Run 2 the reconstruction performed by HLT2 is identical to what is used in offline analyses. This is made possible by running the calibration of the detector, which needs to be re-done for every fill of the LHC, in parallel of the data-taking. On the other hand, in Run 1 the calibration was done after the data-taking which required to reprocess the events. The HLT2 includes a set of inclusive and exclusive lines. The presented analysis fully relies on inclusive HLT2 lines, where inclusive means that even a partial reconstruction of the B candidate is sufficient, which makes the lines efficient for a large array of B decay topologies [80]. These topological lines require at least two charged tracks that have significant IP separation and are associated to the same PV[4]. Further, the tracks must form a vertex downstream of the PV with a high χ^2_{FD} and PID criteria

[4] A track is associated to the PV to which it has the smallest IP.

for muon and electron candidates are applied. Additional tracks are added to form a 2 to 4 body candidate, if these tracks pass similar requirements. In RUN 2 the HLT2 trigger lines rely on multivariate classifiers, such as a Boosted Decision Tree (BDT) [81], while in RUN 1 also cut-based selections where implemented. The cut-based HLT2 lines were dropped in RUN 2, because the BDT lines enhanced both signal efficiency and robustness against detector instabilities that can occur during online running.

3.3.2 Offline Processing

Because of the complexity of the events and the rich variety of analyses performed at LHCb a high flexibility is demanded on the offline processing. Triggered events are further processed by the `BRUNEL` application [82]. Here, a full event reconstruction is performed using information from all subdetectors to transform detector hits into physical quantities such as tracks and clusters. Since this reconstruction step is very computationally intensive, the `BRUNEL` reconstruction is run centrally rather than by individual analysis groups. The reconstructed event information is stored in Data Summary Tape (DST) files which is the format used by all physics analyses.

However, the output DST files from `BRUNEL` are not directly used by analyses. To reduce computational effort, a prior set of selections is applied to filter events that are of interest for a specific kind analysis. This so-called 'stripping' is controlled by the `DaVinci` application [83]. In this software, signal decays are reconstructed by adding the mass hypothesis to tracks and combining the decay chain. Similar final states and stripping selections are group in 'streams' to save disk space and to simplify the access. The output DST files from stripping are used by individual analyses to extract their decay channels of interest.

3.3.3 Simulation

Most analyses performed at LHCb use simulation of the studied physical processes and their detector response. Among the main applications for simulated samples are the determination of efficiencies, background studies and extraction of Probability Density Functions (PDF's) for fits. The simulation uses a Monte Carlo (MC) approach to model the creation of particles in proton-proton collisions and their physical processes as well as their interaction with the spectrometer.

The simulation of initial particle generation and their propagation through the detector is performed by the `Gauss` [84] software package. `Gauss` relies on the `Pythia` progamme [85] to handle the production of particles in proton-proton collisions. Both decay and time evolution of hadronic particles are described by the `EvtGen` package [86] which is specialised on b-hadron decays. The simulation of QED radiative corrections, *i.e.* Final-State Radiation (FSR) and bremsstrahlung, is performed by the `PHOTOS` alogrithm [87]. The `GEANT4` toolkit simulates the interaction of particles with material and creates hits in the sensitive detector regions. Finally, the `BOOLE` algorithm [88] converts the `GEANT4` hits to the digitised response of the readout electronics and the L0 trigger. The output

data from BOOLE has the same format as real data enabling to run HLT and offline reconstruction in the same way (see Fig. 3.10).

At LHCb, the simulation shows deviations from real data in some aspects. In particular, the generated kinematics, responses of PID and trigger systems and event multiplicity are among the quantities that exhibit differences. For instance, the used TCKs for the L0 trigger can vary between different runs of the LHC in a data-taking year (as discussed in Sec. 3.3.1), while for simulated samples only one TCK is used per year. Consequently, each analysis needs to assess possible biases caused by these differences or calibrate the simulation with data-driven methods.

3.4 Electron Reconstruction and Identification

At LHCb, electrons are experimentally much more challenging than muons for the following reasons:

- **L0 trigger:** Due to the worse resolution and higher occupancy of the ECAL, the thresholds used for L0Electron are higher compared to L0Muon (see Sec. 3.3.1) resulting in lower trigger efficiencies.
- **Resolution:** Electrons produce a large amount of bremsstrahlung because of their low mass and are more affected by MS resulting in a worse momentum resolution.
- **PID:** The good momentum resolution and additional information from the muon stations leads to a more accurate PID estimate for muons.

However, the analysis presented in this thesis performs a LFU measurement and must therefore reliably control the different behaviour of the two lepton species. This section provides insight into the identification and reconstruction of electrons.

3.4.1 Calorimeter Response

Due to their low mass, electrons create Cherenkov radiation under the maximal angle $\cos\theta_{\max} = 1/n$ already at comparatively small momenta leading to large overlaps with muon and pion distributions, see Fig. 3.6. Thus, the RICH system only provides separation for very low energy electrons. The main driver for electron PID is the calorimeter system. As mentioned in Sec. 3.2.3, electrons create a signature in the calorimeter system that differs from photons and hadrons, see Fig. 3.8. In short, electrons create hits in the SPD detector which separates them from photons. Further, electrons deposit most of their energy in the ECAL while hadrons create their main shower in the HCAL. To distinguish electrons from charged hadrons, reference histograms are created for each subdetector of the calorimeter system plotting the ratio of measured energy deposit and momentum of the associated particle track. The ECAL histogram of E/p shown in Fig. 3.11 is most important for electron-hadron separation.

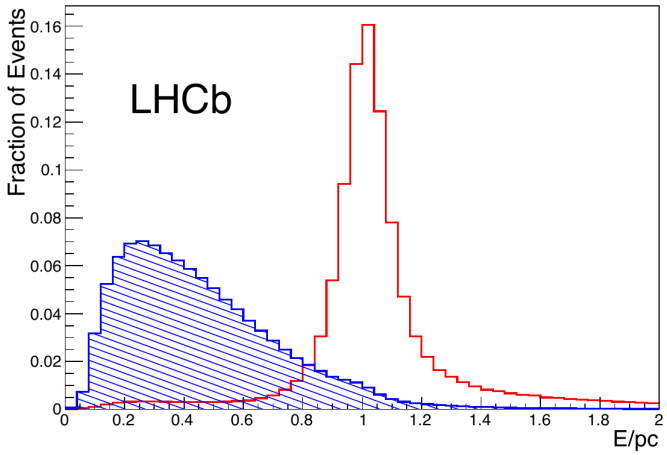

Figure 3.11.: E/p distribution in the ECAL for electrons (red) and hadrons (blue) measured on a part of the 2011 data. Figure taken from [59].

3.4.2 Reconstruction and Bremsstrahlung

In contrast to muons, the resolution of electrons suffers a large penalty due to bremsstrahlung. Because of their considerably lower masses ($m_e = 511\,\mathrm{keV}/c^2$ vs. $m_\mu = 105.7\,\mathrm{MeV}/c^2$) electrons can loose a significant part of their energy due to radiation losses in material interactions while this represents a negligible effect for muons. At LHCb, the momentum estimate of charged particles is based on the curvature of the track caused by the magnetic field. Consequently, bremsstrahlung created in material interactions upstream of the magnet introduces a bias to the momentum calculation. The energy that is carried away by the photon reduces the reconstructed momentum of the electron resulting in a lower reconstructed mass of the mother particle. Photons created behind the magnet do not affect the curvature of the track and therefore do not bias momentum nor energy estimations. Figure 3.12 (left) compares the fully reconstructed dilepton mass systems of a dielectron to a dimuon final-state.

A dedicated recovery algorithm is in place to assign the energy of reconstructed bremsstrahlung photons created before the magnet to the electron track, see Fig. 3.13. Because of the high boost the photons are emitted colinearly to the electron track. Since photons traverse unaffected by the magnetic field, they hit the ECAL in a different position than the electron. The algorithm uses a search window for bremsstrahlung photons, defined by extrapolations of the electron track at the VELO as well as the TT station, that represent the first and last source of material upstream the magnet, to the ECAL. Any photon clusters found in this window are assigned to the electron track and their energies are added to the momentum estimate. This strategy significantly improves the resolution of electron resolution as shown in Fig. 3.12 (right) even though there are some constraints: First, the photon energy estimate is limited by the resolution of the ECAL. Second, not all photons can be successfully reconstructed and assigned to the correct track. Last, in some cases photons from different sources are falsely reconstructed as bremsstrahlung leading to an overestimation of the original electron energy.

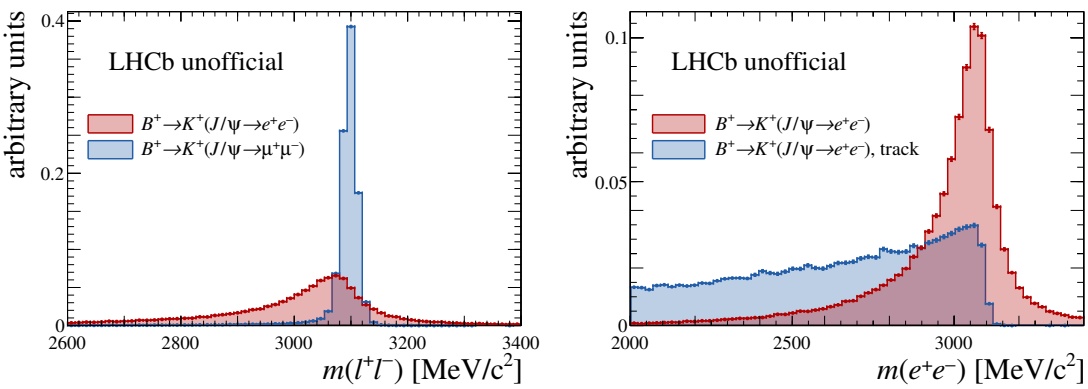

Figure 3.12.: Distribution of $m(\ell^+\ell^-)$ in 2018 $B^+ \to K^+J/\psi$ MC. The left plot shows J/ψ decaying into two electrons (red) and two muons (black). The right plot shows the dielectron mass with full reconstructed momenta (red) and momenta from tracking system only (black).

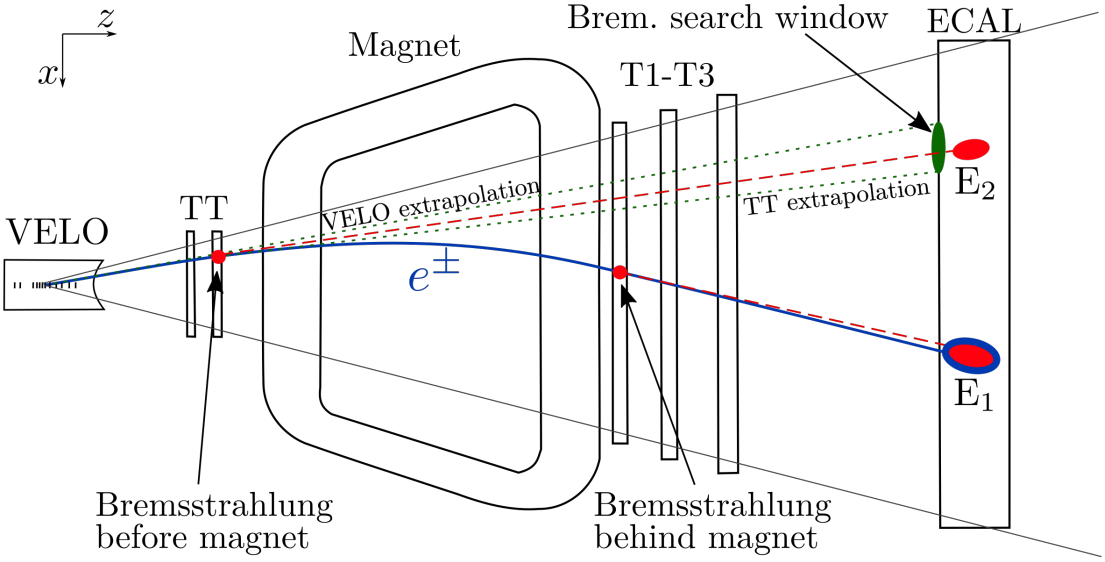

Figure 3.13.: Bremsstrahlung reconstruction at LHCb.

The assignment of bremsstrahlung photons to a track is a typical signature for electrons leading to an improved identification compared to electron candidates without added photons. This is demonstrated in Fig. 3.14, which compares the electron PID variables used in the presented analysis on fully selected 2018 data showing a much more reliable identification for electrons with associated photons.

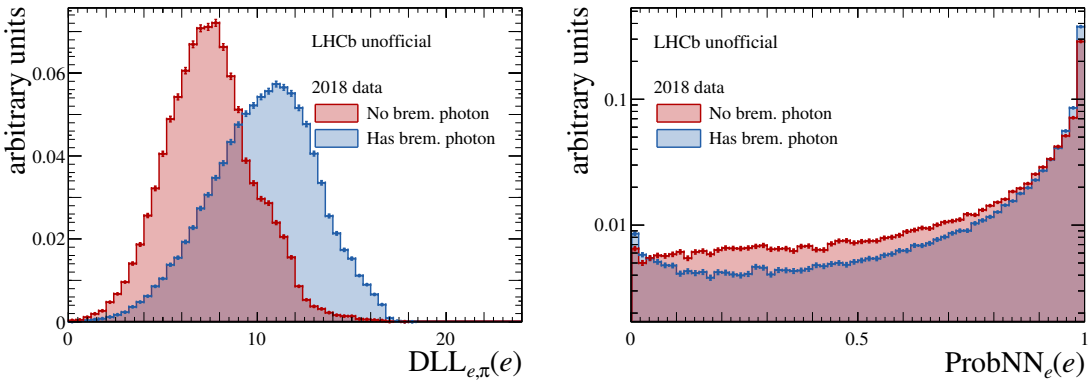

Figure 3.14.: Distribution of $\texttt{DLL}_{e,\pi}(e)$ (left) and $\texttt{ProbNN}_e(e)$ (right) in fully selected 2018 data of the $B^+ \to K^+ J/\psi (\to e^+e^-)$ channel. The red (blue) distribution shows electron candidates having no (at least one) bremsstrahlung photon assigned.

4 Test of Lepton Universality with $b \to s\,\ell^+\ell^-$ Transitions

This chapter describes the simultaneous measurement of the ratios R_K and $R_{K^{*0}}$ exploiting the full Run 1 and Run 2 datasets of LHCb, which represents the most precise test of LFU in $b \to s\,\ell^+\ell^-$ transistions to date. The work performed in the context of this thesis has made significant contributions to all aspects of the presented analysis, while the main focus lay on the selection of physical backgrounds (Sec. 4.3.4), the calibration and calculation of efficiencies (Sec. 4.4), the invariant mass fits to the J/ψ and $\psi(2S)$ resonant modes (Sec. 4.6 and App. D) as well as the calculation of systematic uncertainties and the $r^{J/\psi}$ and $R^{\psi(2S)}$ cross-checks (Secs. 4.7 and 4.8). In total, the analysis is a product of a group effort from several people of the Rare Decays (RD) working group of LHCb. The author of this thesis is listed collaboration internally as one of the three contact authors together with Renato Quagliani (PhD) and Alex Seuthe. Additionally, substantial contributions have been made by Simone Bifani (PhD), Sebastian Schmitt as well as by Ryan Calladine (PhD) [89], Fabrice Desse (PhD) [90] and Da Yu Tou (PhD) [91] in context of their referenced PhD theses. The analysis group is currently preparing a paper [92] to publish the results and has documented the entire analysis in great detail in the analysis note given in Ref. [93]. This chapter aims to provide a complete account of the analysis that is less focused on technical aspects. Throughout this thesis, a variety of variables are used which are detailed in App. B.

The chapter is structured as follows: First, the analysis strategy is explained and the relevant data samples are introduced in Sec. 4.1 and 4.2, respectively. Subsequently, the criteria applied to select the signal events is described in Sec. 4.3. Sections 4.4 and 4.5 report the calibration and calculation of signal efficiencies. The simultaneous fit to the reconstructed B mass used to extract the signal yields is shown in Sec. 4.6. Lastly, in Secs. 4.7 to 4.10 systematic uncertainties, cross-checks and the final results are presented.

4.1 Strategy

The presented analysis is the first simultaneous test of LFU in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays[1], where the K^{*0} meson is reconstructed in the $K^+\pi^-$ final state.

A good understanding of the q^2 differential rate of the studied decays is essential for the presented measurement since it strongly varies as a function of q^2. Figure 4.1 illustrates the distribution of the decay rate and highlights the dominant Wilson coefficients in the

[1] In this thesis the charged conjugated decays are always implied and treated analogously.

respective q^2 regions. On the far left side, the figure shows the peak created by the $b \to s\,\gamma^{(*)}$ process ($\mathcal{C}_7'$) which is the dominant contribution at very low q^2. This so-called photon pole has a stronger impact on electron modes, since the very low q^2 region is kinematically forbidden for the much heavier muon and tau leptons. Furthermore, the photon pole is only present for the $K^{*0}\ell^+\ell^-$ final state, because the $b \to s\,\gamma^{(*)}$ transistion violates angular momentum conservation for pseudoscalar mesons such as K^+. The plateau area on the right side of the photon pole is populated by some light-meson resonances such as $\rho(770)$ and $\phi(1020)$. These modes possess rather small branching ratios ($B(\rho \to \ell^+\ell^-) \approx 4.6 \cdot 10^{-5}$, $\mathcal{B}(\phi \to \ell^+\ell^-) \approx 2.9 \cdot 10^{-4}$ [22]) and the measurements of electron and muon final states are consistent with lepton universality. Otherwise this area is governed by the $b \to s\,\ell^+\ell^-$ transition via the decay modes of interest: $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$, which are referred to as 'signal' or 'rare' modes in this thesis. Here, an interference of the coefficients $\mathcal{C}_7'$, $\mathcal{C}_9'$ and $\mathcal{C}_{10}'$ is present. Above this plateau region, the decay rate is clearly dominated by decays of a B-meson to charmonia (J/ψ, $\psi(2S)$) via the $b \to s\,c\bar{c}$ transition, with a subsequent decay of the charmonia to a dilepton pair. The resonant J/ψ tree-level decay is also referred to as the 'control' or 'normalisation' channel for reasons given below. Figure 4.2 shows the Feynman diagrams for both rare and resonant modes. Finally, at very high values of q^2 the rare decay mode again dominates the decay rate and the region is characterised by an interference of $\mathcal{C}_9'$ and $\mathcal{C}_{10}'$, with small contributions from high mass $c\bar{c}$ resonances that are above the $D\overline{D}$ threshold.

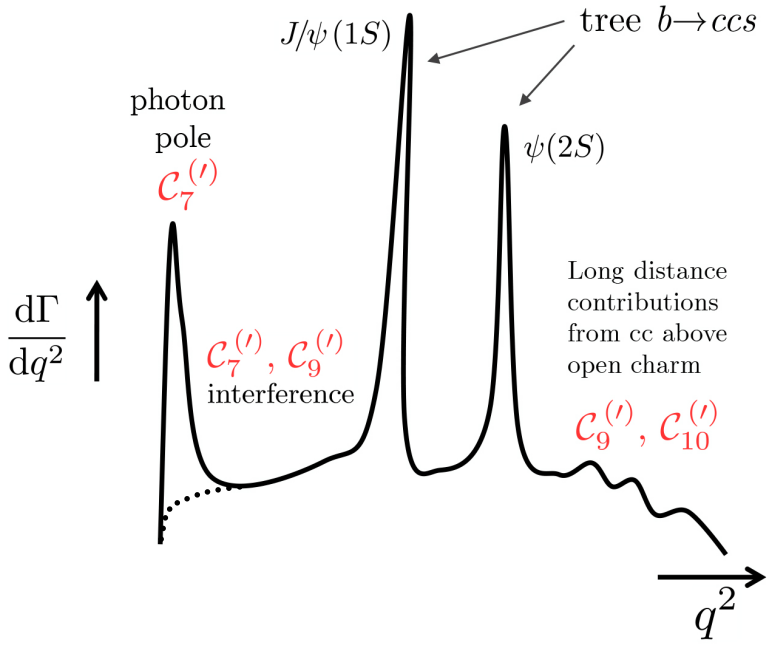

Figure 4.1.: Sketch of the $B^0 \to K^{*0}\ell^+\ell^-$ decay rate as a function of q^2. The dotted line indicates the absence of the photon pole for the $B^+ \to K^+\ell^+\ell^-$ decay, that otherwise has an analogous distribution. Figure taken from [93].

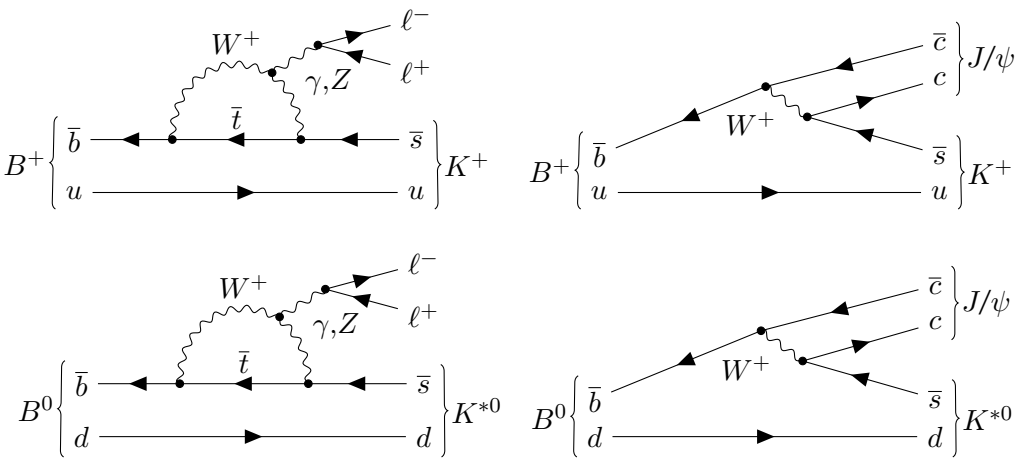

Figure 4.2.: Rare mode (left) and resonant (right) Feynman diagram for the B^+ mode (top) and B^0 mode (bottom). The rare mode also has additional contributions like the box diagram as discussed in Sec. 2.1.4. The $\psi(2S)$ resonant diagram is identical to the J/ψ resonant.

To further illustrate the kinematic phase-space of the involved rare and resonant processes, the q^2 system is plotted against the reconstructed invariant mass of the B meson for both muon and electron modes in Fig. 4.3. The two prominent B mass ($m_B = 5279.5\,\mathrm{MeV}/c^2$) peaks at $q^2 = 9.59\,\mathrm{GeV}^2/c^4$ and $q^2 = 13.59\,\mathrm{GeV}^2/c^4$ correspond to the J/ψ and $\psi(2S)$ resonant modes, respectively. The diagonal bands around these peaks stem from mass resolution and bremsstrahlung. In the muon plot, the rare mode is clearly visible as well as a vertical band at the B mass. The horizontal bands left and right to the resonant modes originate from background sources such as partially reconstructed or combinatorial backgrounds, which are further discussed in due course. These types of background populate the entire phase-space, thus they need to be considered when exploring the mass system of both rare and resonant modes. Comparing the muon and electron plot, it becomes apparent that electron channels have worse resolution and lower yields due to bremsstrahlung and the high occupancy of the electron L0 hardware trigger as already mentioned in Sec. 3.4. Therefore, all analysis aspects are designed to optimise the performance of the electron channel and ported as coherently as possible to the muon mode.

In order to get a clean sample of the rare decays, the analysis is performed in the q^2 region between the photon pole and the J/ψ resonance. The lower boundary is set to $0.1\,\mathrm{GeV}^2/c^4$ to reduce the photon pole contribution to a negligible level. Additionally, below this threshold the R_K and $R_{K^{*0}}$ theory predictions depart from unity and have increased uncertainties [39]. The upper threshold is set to $6.0\,\mathrm{GeV}^2/c^4$ to control the J/ψ resonance, which has a large tail to lower q^2 values due to the poor electron resolution (see Sec. 3.4.2). This region is split into two bins of q^2: The so-called 'low' q^2 bin goes from 0.1–1.1 GeV^2/c^4 and the 'central' q^2 bin from 1.1–6.0 GeV^2/c^4. This is done for two reasons:

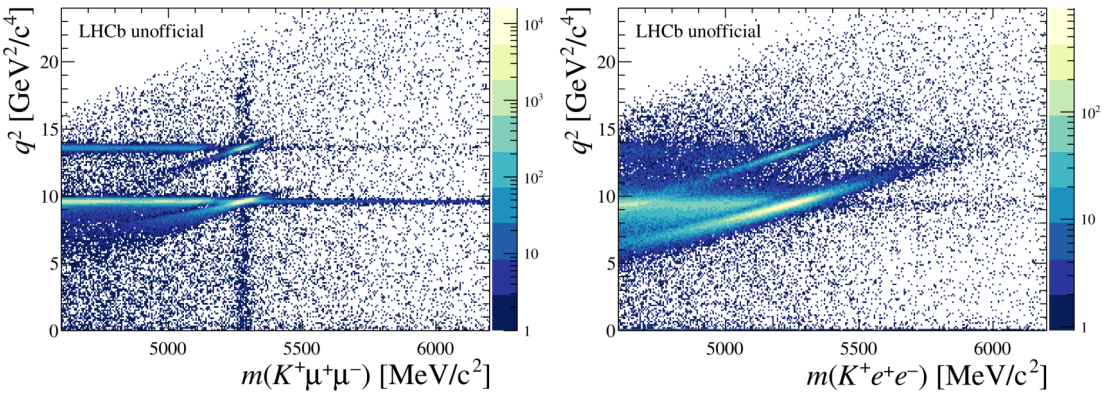

Figure 4.3.: The q^2 system plotted against the invariant mass of the B meson for 2018 data of $K^+\ell^+\ell^-$ candidates in the dimuon final state (left) and dielectron (right). The plots for $K^+\pi^-\ell^+\ell^-$ candidates show qualitatively the same behaviour and are therefore omitted.

1. The low q^2 bin contains all light-meson resonances leaving the central q^2 bin, which is the bin of highest statistics and thus greatest importance, clean. In this analysis, the impact of the light-meson resonances on the low q^2 measurement is ignored, since the current sample size and electron resolution does not allow these contributions to be removed. As a result, the low q^2 measurement is slightly biased towards the SM prediction due to the lepton universality of these modes. Therefore, this strategy can be considered as conservative.

2. The finer the q^2 binning the easier it becomes to discriminate between different types of BSM scenarios because the contributions of various Wilson coefficients changes with q^2. Figure 4.4 gives further insight into this motivation, which shows the q^2 dependence of R_K and $R_{K^{*0}}$ for the SM hypothesis as well as for different NP scenarios for certain Wilson coefficients calculated with an EFT framework. The distribution of R_K is flat above the dimuon threshold for both SM and NP cases, while BSM contributions can lead to significant q^2 dependencies for $R_{K^{*0}}$.

Even though not all previously mentioned arguments hold for both R_K and $R_{K^{*0}}$ (*e.g.* photon pole), the same q^2 binning is chosen since a simultaneous measurement of the $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays is performed.

Experimentally, the ratios R_K and $R_{K^{*0}}$ are determined in a given q^2 bin by determining the efficiency ϵ and signal yield $\mathcal{N}$ for both muon and electron rare modes. The efficiencies are calculated using simulated samples (see Sec. 4.5), while the yields are determined from

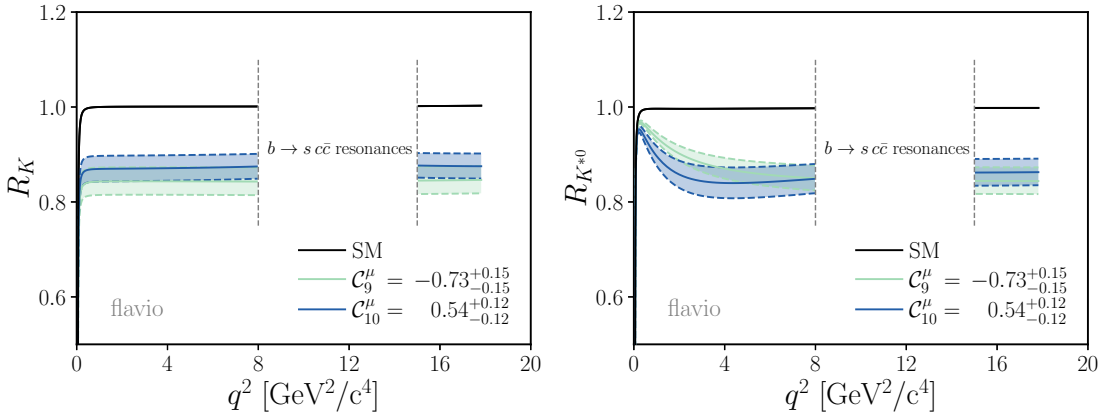

Figure 4.4.: R_K (left) and $R_{K^{*0}}$ (right) as a function of q^2 for both SM and NP scenarios (see legend) that are calculated using the `flavio` software package [94]. Contributions from $c\bar{c}$ resonances are removed. The NP values are taken from global fits [95] based on experimental input on rare B decays (prior to the results presented in this thesis).

one-dimensional maximum likelihood fits to the invariant B-meson mass (see Sec. 4.6). The ratios are thus defined as [2]

$$r_K := \frac{\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \to K^+e^+e^-)} = \frac{\mathcal{N}_{B^+\to K^+\mu^+\mu^-}}{\mathcal{N}_{B^+\to K^+e^+e^-}} \cdot \frac{\varepsilon_{B^+\to K^+e^+e^-}}{\varepsilon_{B^+\to K^+\mu^+\mu^-}}\,. \tag{4.1}$$

$$r_{K^{*0}} := \frac{\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{\mathcal{B}(B^0 \to K^{*0}e^+e^-)} = \frac{\mathcal{N}_{B^0\to K^{*0}\mu^+\mu^-}}{\mathcal{N}_{B^0\to K^{*0}e^+e^-}} \cdot \frac{\varepsilon_{B^0\to K^{*0}e^+e^-}}{\varepsilon_{B^0\to K^{*0}\mu^+\mu^-}}\,, \tag{4.2}$$

The tree-level charmonia modes also play an important role for this analysis. In particular, the J/ψ resonant mode has a branching ratio that is about two orders of magnitude larger than the rare mode [22]. Due to its abundance, this mode is exploited extensively for various exercises like efficiency calibration and signal shape extraction, which are introduced in later sections. Furthermore, the single ratio of the $J/\psi \to \ell^+\ell^-$ modes, that is defined in an analogous way to the rare mode

$$r_K^{J/\psi} := \frac{\mathcal{N}_{B^+\to K^+J/\psi(\to\mu^+\mu^-)}}{\mathcal{N}_{B^+\to K^+J/\psi(\to e^+e^-)}} \cdot \frac{\varepsilon_{B^+\to K^+J/\psi(\to e^+e^-)}}{\varepsilon_{B^+\to K^+J/\psi(\to\mu^+\mu^-)}}\,, \tag{4.3}$$

$$r_{K^{*0}}^{J/\psi} := \frac{\mathcal{N}_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}}{\mathcal{N}_{B^0\to K^{*0}J/\psi(\to e^+e^-)}} \cdot \frac{\varepsilon_{B^0\to K^{*0}J/\psi(\to e^+e^-)}}{\varepsilon_{B^0\to K^{*0}J/\psi(\to\mu^+\mu^-)}}\,, \tag{4.4}$$

represents one of the most important tests of the analysis framework: Since lepton universality is experimentally well established for the J/ψ mode [22]

$$\frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(J/\psi \to e^+e^-)} = 0.9983 \pm 0.0077\,, \tag{4.5}$$

[2] In this thesis, a small r denotes a single ratio of branching fractions while a capital R stands for a double ratio.

probing the agreement with unity of $r^{J/\psi}$ is a powerful cross-check. Moreover, because $r^{J/\psi}$ is unity, it is possible to use the J/ψ mode as normalisation channel and define the final ratios for R_K and $R_{K^{*0}}$ as

$$R_K := \frac{r_K}{r_K^{J/\psi}}\,, \quad R_{K^{*0}} := \frac{r_{K^{*0}}}{r_{K^{*0}}^{J/\psi}}\,. \tag{4.6}$$

This strategy greatly increases the robustness of the measurement, since many systematic uncertainties cancel at leading order, which is especially important for the experimentally challenging electron mode. The double ratio approach is tested in another cross-check, where the double ratio of the two charmonia modes is calculated

$$R_K^{\psi(2S)} := \frac{r_K^{\psi(2S)}}{r_K^{J/\psi}}\,, \quad R_{K^{*0}}^{\psi(2S)} := \frac{r_{K^{*0}}^{\psi(2S)}}{r_{K^{*0}}^{J/\psi}}\,, \tag{4.7}$$

where $r_K^{\psi(2S)}$ and $r_{K^{*0}}^{\psi(2S)}$ are the single ratios of the $\psi(2S)$ resonant modes. Both the single ratio $r^{J/\psi}$ and double ratio $R^{\psi(2S)}$ cross-checks are reported in Sec. 4.8.

The strategy of this analysis is to perform a simultaneous measurement of R_K and $R_{K^{*0}}$ offers several advantages: First, it gives the ability to constrain the partially reconstructed background from the B^0 decay in the B^+ sample. This way, the experimental sensitivity is improved and possible correlations between the measured LFU observables are treated correctly. However, this source of partially reconstructed background is only relevant for the electron mode, since the good resolution of muons renders it negligible. Secondly, the efficiency calibration is performed coherently for B^+ and B^0 modes and calculated using data and simulated samples of both $B^+ \to K^+ J/\psi$ and $B^0 \to K^{*0} J/\psi$ decays. The extracted corrections are shown to be interchangeable, thus providing a method of cross-check as well as a handle on correlations that are unique to this analysis. Lastly, it enables to report the full statistical and systematic covariance matrices of the four LFU measurements allowing a correct combination of results in global fits (see Sec. 4.10.3).

To avoid potential biases, the analysis was performed in a blinded way, meaning that the results for the LFU observables were not evaluated before the analysis procedure was finalised. Since all measured signal modes are known to exist, it was chosen to blind the signal efficiencies rather than yields. This way, it was easier to ensure a good fit stability and correct background modelling prior to unblinding. The final unblinding was performed after all cross-checks and the unblinding checklist described in Sec. 4.9 were passed successfully.

4.2 Data and Simulation Samples

4.2.1 Data Samples

The measurements presented in this thesis are based on pp collision data taken at LHCb. The centre of mass energy of the LHC and the integrated luminosity for each data taking

year are shown in Tab. 4.1. Throughout this thesis, the data is separated into three run periods: 2011 and 2012 data samples are referred to as Run 1, 2015 and 2016 as Run 2p1 and 2017 and 2018 as Run 2p2.

Table 4.1.: Summary of the data samples used for the presented analysis.

Year	$\mathcal{L}$ [fb^{-1}]	$\sqrt{s}$ [TeV]
2011	1.1	7
2012	2.1	8
2015	0.3	13
2016	1.7	13
2017	1.7	13
2018	2.2	13

4.2.2 Trigger

As introduced in Sec. 3.3.1, the trigger system at LHCb consists of a L0 hardware trigger and a two-staged software trigger (HLT1 and HLT2). Since electrons and muons leave different signatures in the detector, it is not possible to fully align the trigger selection between these modes without a large loss of signal. However, the choice of trigger paths is designed to treat the different modes probed in this analysis as coherently as possible. In fact, the trigger selection is completely identical between B^+ and B^0 modes.

L0 Hardware Trigger

In this analysis two L0 trigger categories are used for both electrons and muons:

1. The primary category uses events triggered independently of the signal (`TIS`) final state particles, meaning that the L0 trigger was fired on a particle from the underlying event. Here, the L0 muon, electron and hadron trigger lines are considered so that the trigger category, which is referred to as L0I, is defined as

$$\text{L0I} := \texttt{L0Muon\_TIS} \,\|\, \texttt{L0Electron\_TIS} \,\|\, \texttt{L0Hadron\_TIS}\,. \tag{4.8}$$

2. The secondary trigger category consists of events that are triggered on the signal (`TOS`) final state leptons. For the dimuon (dielectron) channel, at least one of the two leptons must fire the `L0Muon` (`L0Electron`) trigger. Therefore, the individual trigger decisions of the two final state leptons are linked with the logical 'OR' condition. This trigger category is referred to as L0L, or as L0M and L0E when talking about the muon and electron mode, separately. Since it is the secondary

category, the L0I term needs to be inverted in order to avoid double counting. Consequently, L0L is an exclusive category and defined as

$$\text{L0M} := \texttt{L0Muon\_TOS}(\mu^+,\,\mu^-)\ \&\&\ !\text{L0I}\,, \tag{4.9}$$

$$\text{L0E} := \texttt{L0Electron\_TOS}(e^+,\,e^-)\ \&\&\ !\text{L0I}\,. \tag{4.10}$$

This choice of L0 categories is novel for LFU searches in rare $b \to s\,\ell^+\ell^-$ transistions at LHCb. Previous measurements (*e.g.* [11, 13]) were using the `L0Electron` as inclusive, `L0Hadron` as secondary and L0I as tertiary category for the electron channel, while for the muon channel `L0Muon` was the only category used. The motivation for this change is that the L0 trigger response is much more similar between electron and muon modes for the L0I category compared to L0L. Hence, a primary category is formed which minimises the L0 induced differences between the two types of leptons. Furthermore, the inverted term used to form the exclusive secondary category is aligned between electron and muon modes. Discarding the `L0Hadron` category comes only with a small loss of signal yield, since most events are recovered by the L0I category. In addition, the trigger efficiency of this category is very challenging to calibrate and usually comes with a large systematic uncertainty.

HLT Software Trigger

Table 4.2 shows the full HLT selection used in this analysis. The HLT1 selection is fully aligned between muon and electron modes. The used HLT1 lines introduce cut-based requirements on the events such as the existance of a displaced long track with p_{T} larger than $500\,\mathrm{MeV}/c$ (as described in Sec. 3.3.1). Only information from the tracking system is used here. At the HLT2 stage, so-called topological (`Topo`) trigger lines are exploited. These lines try to reconstruct B meson decays from two-, three- and four-track vertices, where only tracks of charged final state particles are considered. In this analysis, the four-track lines are discarded for the B^0 mode in order to align the selection with the three-body final state of the B^+ decays. The signal loss for the B^0 electron mode induced by dropping the four-track lines is vary small ($< 1\%$). The HLT2 lines apply requirements such as a significant displacement of the B vertex from the PV on the events. Here, a response from multivariate classifiers, such as bonsai BDTs (BBDT) [81], is calculated based on properties of both the vertex and its associated final state tracks. Additionally, topological lines dedicated for decays to final states including at least one or two muon(s) (`TopoMu(Mu)`) or electron(s) (`TopoE(E)`) are used. These lines apply a softer requirement on the multivariate classifier response, if at least one of the tracks associated to the vertex passes muon or electron PID criteria.

4.2.3 Stripping

The stripping lines used to select the dimuon and dielectron modes are called `Bu2LLKmmLine` and `Bu2LLKeeLine2`, respectively. These lines are specifically designed for LFU measurements in various $b \to s\,\ell^+\ell^-$ decay modes such as $B^0 \to K^{*0}\ell^+\ell^-$ and $B^+ \to K^+\ell^+\ell^-$ and

Table 4.2.: Summary of the HLT1 and HLT2 lines used in this analysis.

Stage	Year	Muon modes	Electron modes
HLT1	Run 1	`Hlt1TrackAllL0`	
	Run 2	`Hlt1TrackMVA`	
HLT2	Run 1	`Hlt2Topo[2,3]BodyBBDT`	
		`Hlt2TopoMu[2,3]BodyBBDT`	`Hlt2TopoE[2,3]BodyBBDT`
	2015	`Hlt2Topo[2,3]Body`	
		`Hlt2TopoMu[2,3]Body`	
	2016, Run 2p2	`Hlt2Topo[2,3]Body`	
		`Hlt2TopoMu[2,3]Body`	`Hlt2TopoE[2,3]Body`
		`Hlt2TopoMuMu[2,3]Body`	`Hlt2TopoEE[2,3]Body`

make up the first set of offline selection criteria. The full set of requirements introduced by these lines is summarised in Tab. 4.3, where the used variables are defined in App. B.

Table 4.3.: Requirements from the `Bu2LLKmmLine` and `Bu2LLKeeLine2` stripping lines which are used in all years of data-taking. Cuts that are unique to B^+ or B^0 modes are labelled accordingly.

Applied to	**Requirement**
Event	`nSPDHits` $< 600(450)$ Run 1 (Run 2)
B	$\lvert m - m_B^{\mathrm{PDG}}\rvert < 1500\,\mathrm{MeV}/c^2$
	`DIRA` > 0.9995
	$\chi^2_{\mathrm{IP}}(\mathrm{PV}) < 25$
	$\chi^2_{\mathrm{vtx}}(\mathrm{SV})/\mathrm{ndf} < 9$
	$\chi^2_{\mathrm{PV,\,SV}} > 100$
K^{*0} (B^0 only)	$\lvert m - m_{K^{*0}}^{PDG}\rvert < 300\,\mathrm{MeV}/c^2$
	$p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$
	$\chi^2_{\mathrm{vtx}}(\mathrm{SV})/\mathrm{ndf} < 25$
K	$\mathtt{DLL}_{K,\pi} > -5$
	$\chi^2_{\mathrm{IP}}(\mathrm{PV}) > 9$
	$p_{\mathrm{T}} > 400\,\mathrm{MeV}/c$ (B^+ only)
π (B^0 only)	$\chi^2_{\mathrm{IP}}(\mathrm{PV}) > 9$
$\ell\ell$	$m < 5500\,\mathrm{MeV}/c^2$
	$\chi^2_{\mathrm{vtx}}(\mathrm{SV})/\mathrm{ndf} < 9$
	$\chi^2_{\mathrm{PV,\,SV}} > 16$
μ	`isMuon`, `hasMuon`
	$p_{\mathrm{T}} > 300\,\mathrm{MeV}/c$
	$\chi^2_{\mathrm{IP}}(\mathrm{PV}) > 9$
e	$\mathtt{DLL}_{e,\pi} > 0$
	$p_{\mathrm{T}} > 300\,\mathrm{MeV}/c$
	$\chi^2_{\mathrm{IP}}(\mathrm{PV}) > 9$

4.2.4 Simulated Samples

In this thesis, the used simulated samples, also referred to as MC samples, are usually produced for a specific decay mode. The generation with `Pythia` is repeated until an event is produced that features the requested decay mother particle[3]. Subsequently, this particle is forced to decay into the requested final state, where also intermediate resonances can be specified. Two different types of simulated samples are used: On the one hand, information on kinematic variables directly after the event production with `Pythia` is stored in `generator-level` simulation samples. The only selection criterion applied to these samples is that all charged final state particles of the simulated decay mode must fall into the acceptance of LHCb. On the other hand, `detector-level` samples contain the full response of interactions from the generated particles with the detector and are designed to mimic the data samples. In addition to the reconstructed information that is available in both data and MC, the simulated samples also provide the 'true' event information which is unbiased from any detector effects. In the following, this information is referred to as `MCTruth`. Most notably, the `MCTruth` information is used to perform the so-called 'truth-matching': all final state particles must have the correct mass hypothesis assigned and are matched to the correct mother particle to form the requested signal decay. Furthermore, reconstructed quantities of the signal candidate, such as tracks and clusters, must relate to the true generated signal particles. The truth-matching is always applied to simulated samples of both rare and resonant modes in order to extract results from correctly identified signal decays. However, this method is obviously not applicable on data samples. Here, contributions from misreconstructed events, such as the misidentification of a final state track, are treated as sources of background (see Sec 4.3).

Simulated samples are used throughout the analysis for the following purposes:

- The calculation of signal efficiencies for rare and resonant modes;
- The extraction of mass lineshapes to model both signal and background components in invariant mass fits;
- The study of various sources of background from physical processes;
- The training of classifiers that are used to filter out backgrounds.

Consequently, a large number of simulated samples for both signal and background decay modes is required to achive these versatile tasks. It is important that the simulated samples accurately represent data to extract reliable results. In this analysis, a set of data-driven corrections is applied to calibrate the simulation (see Sec. 4.4). To fully correct the simulated PID response, it is necessary to remove the PID cuts from the stripping requirements when processing the samples. Modified versions of the `Bu2LLK` stripping lines are used for MC samples, where the cuts on $\texttt{DLL}_{K,\pi}(K) > -5$, $\texttt{isMuon}(\mu), \texttt{hasMuon}(\mu)$ and $\texttt{DLL}_{e,\pi}(e) > 0$ are excluded.

[3]This particle is always some type of b-hadron for the samples used in this analysis.

Physical Model

The simulated samples for both $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ modes rely on the `BTOSLLBALL` 6 model from the `EvtGen` package [86], where the theoretical calculation of the form factors are taken from Ref. [96] for B^+ and Ref. [97] for B^0 decays. Any bias from the assumptions in these calculations is expected to cancel at leading order in the double ratio measurement. Nevertheless, a systematic uncertainty on the choice of the model is assigned as discussed in Sec. 4.7.1.

However, there are some decays, $\Lambda_b^0 \to pK\ell^+\ell^-$ and $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$ for instance, where no reliable theoretical model exists. The hadronic part of these decays often proceeds via intermediate resonances such as Λ^0 or $K_1^+(1270)$ for the $\Lambda_b^0 \to pK\ell^+\ell^-$ and $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$ decay, respectively. Because of the lack of a physical model, the simulated samples use the 'phase-space' model from `EvtGen` which produces the hadronic final states without these intermediate resonances. Both these decays are important sources of background for the B^0 mode and therefore extracting their correct lineshape in the reconstructed B^0 mass system as well as their efficiency is crucial. Thus, data-driven corrections are applied to these MC samples to correct the hadronic mass system.

For the $\Lambda_b^0 \to pK\ell^+\ell^-$ decay the used correction of the hadronic system was developed in the amplitude analysis of the $\Lambda_b^0 \to pKJ/\psi$ resonant decay reported in Ref. [98]. Here, weights are applied to simulation as a function of the p (in our case π), K and J/ψ four-momenta based on `MCTruth` information to correct for the differences between data and simulation in the hadronic system (see Fig. 4.5).

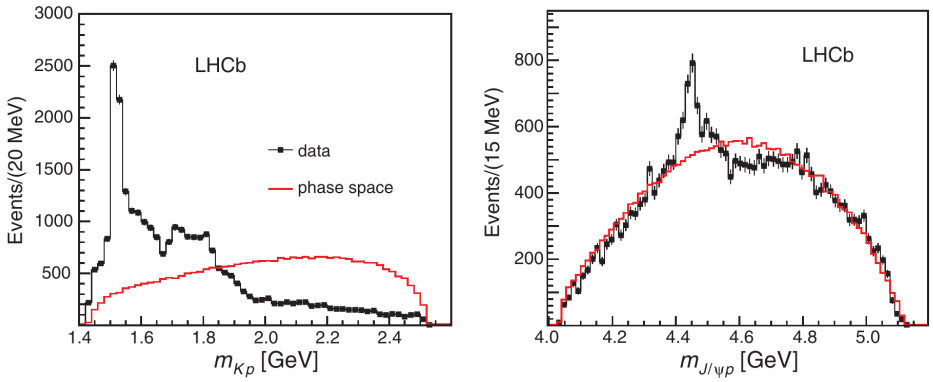

Figure 4.5.: Invariant mass system of pK^- (left) and pK^-J/ψ (right) for phase-space MC (red) and real data (black). Figure taken from Ref. [98].

The correction for the $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$ decay was developed by the analysis team of the $R_{K\pi\pi}$ measurement, which is currently on-going at LHCb. Similarly to the $\Lambda_b^0 \to pK\ell^+\ell^-$ correction, a reweighting of the simulated $m(K^+\pi^+\pi^-)$, $m(K^+\pi^-)$ and $m(\pi^+\pi^-)$ mass systems is performed to match the distributions seen in $B^+ \to K^+\pi^+\pi^-J/\psi$ data (see Fig. 4.6).

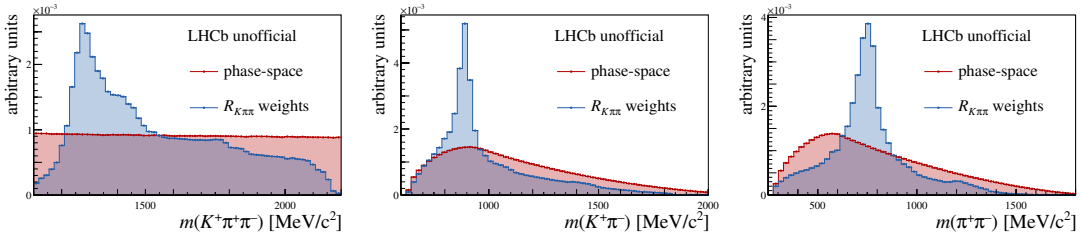

Figure 4.6.: The $m(K^+\pi^+\pi^-)$, $m(K^+\pi^-)$ and $m(\pi^+\pi^-)$ (from left to right) invariant mass system of $B^+ \to K^+\pi^+\pi^- e^+ e^-$ MC with the phase-space model (red) and after the hadronic mass reweighting (black).

4.3 Offline Selection

The pp collisions provided by the LHC result in a very busy environment in the detector and leave the initial state of the decay undefined. Thus, the task of reconstructing decays is rather challenging at LHCb and significant levels of background need to be rejected. There are several sources of background that need to be considered in this analysis, which are divided into three main categories:

1. **Misidentified:** As discussed in Sec. 3.2.2, the PID system of the detector allows to assign a mass hypothesis to the final state tracks. Decay modes to a different final state can be misreconstructed as signal decays if one or more final state particles have assigned a wrong mass hypothesis. Especially for low and high momentum tracks, it is not possible to obtain a reliable separation between the different particle species resulting in misidentification probabilities of up to several percent.
2. **Partially reconstructed:** Decay modes to a higher number of final state particles can be misreconstructed as signal decays, if some final state particles are not reconstructed.
3. **Combinatorial:** Random tracks from the event can be combined to form a decay vertex which can be interpreted as a signal decay. Usually, some of the tracks did originate from a common source, *e.g.* a true $J/\psi \to \ell^+\ell^-$ can be combined with a random kaon track to form a $B^+ \to K^+ J/\psi\,(\to \ell^+\ell^-)$ candidate. This is the reason why the amount of combinatorial background is much higher in the J/ψ region with respect to the low and central q^2 bins.

An offline selection is applied to the samples that are selected by the stripping lines, in order to increase the signal purity. This is particularly of importance for the rare modes due to the low signal yield. This analysis uses a tight selection compared to similar LFU measurements at LHCb in order to get a clean working point, which improves the control over residual backgrounds and allows to perform the fits in a wider range of the invariant B mass.

The selection consists of several stages: First, the mass windows for the used q^2 regions and the reconstructed B mass are defined for rare and resonant modes in Sec. 4.3.1. Second, a set of requirements is introduced to ensure a good reconstruction quality of

the signal candidates. This selection step is described in 4.3.2. Third, the cuts on the PID response are described in Sec. 4.3.3. Fourth, the vetoes against physical background modes from misidentified and partially reconstructed decays are explained in Sec. 4.3.4. Lastly, the Multivariate Analysis (MVA) based selections against combinatorial and partially reconstructed backgrounds are subject of Sec. 4.3.5. There is some residual level of background left in the data samples even after the full selection is applied. These backgrounds are included in the invariant mass fits as described in Sec. 4.6.

4.3.1 Mass Windows

The q^2 windows used for the rare decays were already motivated in Sec. 4.1 and are identical for muon and electron modes. For the J/ψ resonant channel, the q^2 window for the electron mode is larger with respect to the muon mode to account for the poorer resolution. Similarly, the window of the reconstructed B meson mass is wider in the electron case. The mass window for the muon rare mode starts at higher values compared to the resonant mode to fully exclude partially reconstructed backgrounds in the rare fits, while the same window is used between the respective electron channels. For B^0 modes, the $K\pi$ system is selected in a 100 MeV window around the $K^*(892)^0$ mass so that the sample is dominated by the resonant P-wave contribution. Indeed, in this region the $K\pi$ S-wave contribution is found to be relativley small [43] and similar between electron and muon modes [11].

Except for the K^{*0} mass window cut, the identical mass selection is applied to B^+ and B^0 modes. All mass windows used to measure R_K and $R_{K^{*0}}$ are defined in Tab. 4.4. Different mass windows, which are defined in App. D, are used for the $r^{J/\psi}$ and $R^{\psi(2S)}$ cross-checks, since here the reconstructed B mass systems are calculated using a J/ψ ($\psi(2S)$) mass constraint on the dilepton system. Throughout the thesis, the J/ψ and $\psi(2S)$ mass constraints are calculated via the `DecayTreeFitter` (DTF) alogrithm [99], which constrains the invariant mass of the dilepton system to the known J/ψ ($\psi(2S)$) mass. This way, the resolution of the resonant channels is enhanced, especially for the electron modes.

Definition of q^2

Differences between the true squared dilepton mass system (q^2_{true}) and the reconstructed q^2 distributions can arise from the imperfect reconstruction of the lepton kinematics by the LHCb detector. For muons only small variations from q^2_{true} occur and are mainly driven by the single track momentum resolution. Whereas electrons can have a substantial deviation between q^2_{true} and q^2 due to bremsstrahlung, leading to a significant fraction of events which migrate into a different region of q^2. To take bin-migration correctly into account in simulation, two definitions of the squared invariant dilepton mass are used in this analysis:

- The nominal distribution calculated from the reconstructed lepton four-momenta $q^2 = |p_{\ell^+} + p_{\ell^-}|^2$.

Table 4.4.: Summary of the mass windows used for the R_K and $R_{K^{*0}}$ measurements.

Type	Applied to	Requirement
q^2	low q^2	0.1–1.1 GeV^2/c^4
	central q^2	1.1–6.0 GeV^2/c^4
	$J/\psi\,(\mu^+\mu^-)$	$\lvert m(\mu^+\mu^-) - m^{\mathrm{PDG}}_{J/\psi}\rvert < 100\,\mathrm{MeV}/c^2$
	$J/\psi\,(e^+e^-)$	6.0–11.0 GeV^2/c^4
$m(B)$	$\mu^+\mu^-$ rare	5150–5850 MeV/c^2
	$J/\psi\,(\mu^+\mu^-)$	5100–6100 MeV/c^2
	e^+e^- modes	4600–6200 MeV/c^2
$m(K^{*0})$	B^0 modes	$\lvert m(K\pi) - m^{\mathrm{PDG}}_{K^{*0}}\rvert < 100\,\mathrm{MeV}/c^2$

- The so-called 'true-q^2' distribution calculated as $q^2_{\mathrm{true}} = |p_B - p_h|^2$ with h being a K^+ and $K^+\pi^-$ for B^+ and B^0 modes, respectively.

The q^2_{true} distribution is used to define q^2 windows on generator-level MC enabling a consistent definition between muon and electron channel before FSR and reconstruction effects. On fully reconstructed data and simulation, the nominal q^2 definition is used to define the mass window. This way, the impact of bin-migration on the q^2 selection efficiency is properly accounted for, if the decay model and detector resolution are correctly simulated. For the choice of the decay model (`BTOSLLBALL`) a systematic uncertainty is assigned by studying the efficiency difference when using different models (see Sec. 4.7.1), while the simulated detector resolution is calibrated as described in Sec. 4.4.6.

4.3.2 Reconstruction and Acceptance

There are several requirements placed on the final state particles that are used to form the signal candidate. Cuts placed on the χ^2 of the track and the `GhostProb` variable [100] ensure good quality tracks. A minimum threshold on the momentum and transverse momentum is required for all final state particles, to get a reliable response from the RICH detectors. The threshold values are chosen to align the selection with the calibration samples used to correct the PID efficiency in simulated samples (see Sec. 4.4.1). Further, there are cuts on acceptance variables (`hasRICH`, `hasCalo` and `inAccMuon`) to ensure that each final state track is provided with reliable information from the subdetectors, that are important to the respective type of the particle. There are ECAL specific requirements for electron candidates to veto tracks that fall outside the ECAL acceptance or into the region closest to the beam pipe. Additionally, a minimum distance between the two electron tracks extrapolated to the ECAL plane is required. The reason for this cut is,

that the PID corrections applied to simulation assume an uncorrelated PID response for the two electrons. However, a study performed by our analysis group [101] has shown, that for tracks that are close in the ECAL plane, the clusters of the electrons can overlap and bias the PID efficiency. The ECAL distance cut significantly reduces this effect and the residual bias is treated as a systematic uncertainty (see Sec. 4.7.1). Lastly, there are cuts against so-called 'clone' tracks applied: Two tracks are considered clones when they share at least 70 % of their hits in the tracking system. Clone tracks occur for instance when the same VELO track is matched to two different track segments in the T1-T3 stations. Since the angle between clone tracks is close to zero, this type of background is very efficiently suppressed by requiring a minimum angle of 0.5 mrad between all final state particles. The full list of quality selection criteria is given in Tab. 4.5.

Table 4.5.: Summary of the reconstruction quality and acceptance requirements.

Type	**Applied to**	**Requirement**
Quality & Acceptance	all tracks	$\chi^2_{track}/ndf < 3$
		`GhostProb` < 0.4
		$p < 200\,\text{GeV}/c$
		`hasRICH`
	K, π	$p_\text{T} > 250\,\text{MeV}/c$
		$p > 2\,\text{GeV}/c$
		`inAccMuon`
	μ	$p_\text{T} > 800\,\text{MeV}/c$
		$p > 3\,\text{GeV}/c$
		`inAccMuon`
	e	$p_\text{T} > 500\,\text{MeV}/c$
		$p > 3\,\text{GeV}/c$
		`hasCalo`
ECAL	e	$!(x_{ECAL} < 363.6\,\text{mm}$ && $y_{ECAL} < 282.6\,\text{mm})$
		$\text{region}_{ECAL} \geq 0$
		$d_{ECAL}(e, e) > 100$ mm
Clones	all tracks	$\theta(\ell_1, \ell_2) > 0.5\,\text{mrad}$
		$\theta(\ell_{1,2}, h) > 0.5\,\text{mrad}$

4.3.3 PID Response

Selection criteria on the PID response of all final state particles are applied to reduce the pollution from misidentified backgrounds. For each particle sepcies, a requirement on `ProbNN` variables is introduced. Additionally, the cuts on `DLL` variables from the stripping lines have been tightened for both kaons and electrons. The cut thresholds are largely inspired by the RUN 1 $R_{K^{*0}}$ analysis [11]. The applied PID selection criteria are summarised in Tab. 4.6. The effect of the PID selection on misidentified backgrounds is discussed in the following section.

Table 4.6.: Summary of the PID selection.

Type	Applied to	Requirement
PID	K	$\texttt{DLL}_{K,\pi} > 0$ $\texttt{ProbNN}_K \cdot (1 - \texttt{ProbNN}_p) > 0.05$
	π	$\texttt{ProbNN}_\pi \cdot (1 - \texttt{ProbNN}_K) \cdot (1 - \texttt{ProbNN}_p) > 0.1$
	μ	$\texttt{ProbNN}_\mu > 0.2$
	e	$\texttt{DLL}_{e,\pi} > 2$ $\texttt{ProbNN}_e > 0.2$

4.3.4 Physical Backgrounds

Several decay modes constitute sources of background for the signal samples due to errors such as partial reconstruction and misidentification. Additionally, decays to a final state featuring fewer particles than the signal decay can be combined with random tracks from the underlying event to form a signal candidate. These so-called over-reconstructed decays fall into the category of combinatorial background.

When studying backgrounds from misidentified final states particles, it is always useful to probe the mass system under the reversed mass hypothesis: For example the decay $B_s^0 \to \phi e^+ e^-$, with ϕ decaying further to K^+K^-, can be reconstructed as a signal $B^0 \to K^{*0} e^+ e^-$ decay, if the K^- is misidentified as a π^-. By assigning the kaon mass hypothesis to the pion candidate, it is possible to reconstruct the ϕ resonance for this background mode when probing the $m(K^+\pi^-_{\to K})$ mass system. This mass system is calculated using the reconstructed momentum information of the pion candidate and the known kaon mass (and the four momentum of the K^+). For misidentifications of hadrons as electron candidates it is necessary to use the reconstructed momentum that is purely calculated from tracking information, since for electrons the full reconstructed momenta are modified by the bremsstrahlung recovery algorithm (see Sec. 3.4.2). It is usually possible to introduce an efficient veto on intermediate resonances that are unique to the

background modes. The reconstructed mass of the b-hadron for these backgrounds is biased by the misidentification, since the final state mass enters in its calculation. The direction and amount of the bias are resulting from the difference between the assigned mass hypothesis of the misidentied particle and its true mass, so that the $B_s^0 \to \phi e^+ e^-$ no longer peaks at the B_s^0 mass but at lower values close to the B^0 mass.

This section introduces the physical modes that are relevant sources of background for the studied data samples and describes strategies used to reduce their contribution. To this end, studies on data samples as well as on dedicated simulated samples of the respective background modes are presented. It is important to note that background modes are always reconstructed as signal decays when processing the MC samples, in order to extract the efficiency of the background with respect to the studied signal process. The studies focus on the experimentally more challenging electron modes and will use Run 2p2 samples as example.

Determining Residual Background Levels

To quantify the residual background pollution of the studied modes, their expected yields in the signal region after full selection are calculated. The expected yield of a background mode N_{bkg} can be calculated by normalising to a well known channel, *i.e.* a normalisation channel

$$N_{\text{bkg}} = \epsilon_{\text{bkg}} \cdot \frac{\mathcal{B}_{\text{bkg}}}{\mathcal{B}_{\text{norm}}} \cdot \frac{\mathcal{N}_{\text{norm}}}{\epsilon_{\text{norm}}} \cdot \frac{f_{\text{bkg}}}{f_{\text{norm}}} =: \epsilon_{\text{bkg}} \cdot \alpha_{\text{norm}} \,, \tag{4.11}$$

where the $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ mode is used as normalisation channel for background studies relevant for B^+ samples, while $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ is used for B^0 samples. The branching fractions of both background ($\mathcal{B}_{\text{bkg}}$) and normalisation ($\mathcal{B}_{\text{norm}}$) channel are taken from the PDG [22]. The yield of the normalisation channel $\mathcal{N}_{\text{norm}}$ is determined with mass fits described in Sec. 4.6. Potential differences between background and normalisation mode in the hadronisation process involved to create the b-hadron are accounted for by the fragmentation fractions f_{bkg} and f_{norm}, which are defined in Sec. 2.5. The efficiencies of background (ϵ_{bkg}) and normalisation mode ϵ_{norm} are calculated on simulated samples as described in Sec. 4.5.

However, for many background modes the efficiency cannot be determined reliably via the frequentist approach, since no (or very few) simulated events survive the full selection, because of the limited size of the samples and the large suppression factor. In these cases, an upper limit on the efficiency and thus on the expected background yield is calculated: A Bayesian approach is chosen where the posterior PDF $p(\epsilon_{\text{bkg}}|n_{\text{sel}},\, n_{\text{gen}})$ for the selection efficiency ϵ_{bkg} with respect to the number of selected (n_{sel}) and generated (n_{gen}) events of the studied simulated sample is defined via Jeffrey's prior as detailed in Ref. [102]. With this ansatz, the posterior PDF reads as

$$p(\epsilon_{\text{bkg}}|n_{\text{sel}},\, n_{\text{gen}}) = \text{Be}\left(\epsilon_{\text{bkg}}|n_{\text{sel}} + \frac{1}{2}, n_{\text{gen}} - n_{\text{sel}} + \frac{1}{2}\right) \,, \tag{4.12}$$

with the Euler Beta function $\text{Be}(x|a,b)$. The upper limits are derived with an inverse transform sampling method by extracting the efficiency 10^6 times from the distribution

given by the PDF in Eq. 4.12. Additionally, the full normalisation factor α_{norm} is sampled from a Gaussian function, where the mean is given by its nominal value and the standard deviation by the propagated uncertainties. The product of ϵ_{bkg} and α_{norm} yields the distribution of the expected background yield from which an upper limit is extracted at a chosen confidence level of 90 %. The sampled distributions are shown in Fig. 4.7 for the $B^+ \to (\overline{D}^0 \to K^+ e^- \bar{\nu_e})\pi^+$ decay, which is a background for B^+ modes if the pion is misreconstructed as an electron.

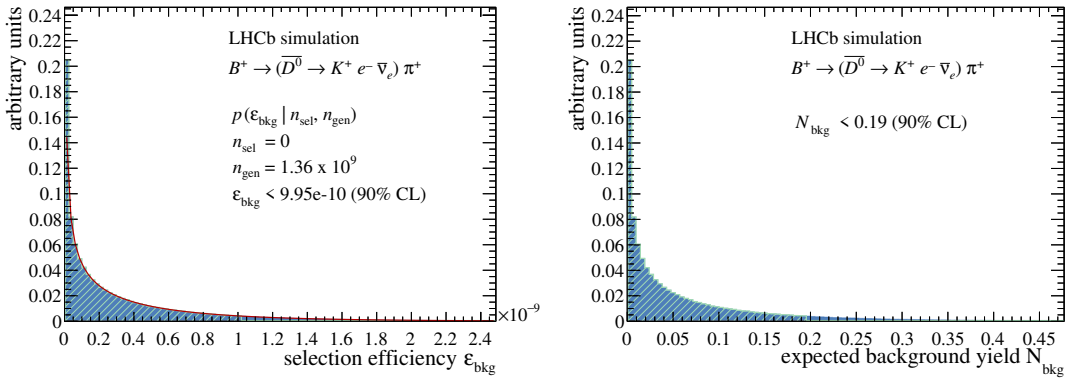

Figure 4.7.: Sampled distributions for the selection efficiency (left) and expected yield of the $B^+ \to (\overline{D}^0 \to K^+ e^- \bar{\nu_e})\pi^+$ background mode in central q^2 from 2018 MC. The red curve in the left plot displays the PDF used to sample the efficiency. The 90 % confidence region is above the green area. The expected background yield relates to the full B^+ RUN 2P2 central q^2 sample.

In the following, the Bayesian method is used whenever less than 20 simulated events remain after full selection. The full list of expected yields for all background modes in every year of data taking is given in the analysis note [93]. Some key values are quoted in the next sections to demonstrate the residual level of contamination of the studied processes after full selection. It is important to note that the calculated background yields do not enter the final R_K and $R_{K^{*0}}$ measurements, instead the remaining background pollution is covered by the systematic uncertainties on the modelled backgrounds in the mass fits (see Se. 4.7.2).

Misidentified Backgrounds in B^0 Modes

First, relevant sources of misidentified backgrounds are examined. As mentioned above, these usually originate from decays to final states different to the studied signal. However, in some cases the signal decay cause its own source of misidentified background, if two of the final state particles have assigned the mass hypothesis associated with the other particle's type. This background category is referred to as 'swap'. All relevant sources of misidentified backgrounds for B^0 modes are separated according to their type of particle

misidentification and are listed in Tab. 4.7. In the following, the various sources of misidentification are studied individually.

Table 4.7.: Summary of relevant sources of misidentified backgrounds for B^0 modes. The $N_{\text{bkg}}/N_{\text{sig}}$ column shows the ratio of the expected background yield calculated with the Bayesian method and the rare yield taken from Tab. 4.25 for the electron mode in central q^2 (if the component is not modelled in the mass fit).

Type	**Decay mode**	q^2 **region**	$N_{\text{bkg}}/N_{\text{sig}}$
$K \to \pi$	$B_s^0 \to (\phi \to K^+K^-)(J/\psi \to \ell^+\ell^-)$	J/ψ	Modelled in fit
	$B_s^0 \to (\phi \to K^+K^-)\ell^+\ell^-$	rare	$< 0.3\,\%$
$p \to \pi$	$\Lambda_b^0 \to pKJ/\psi(\to \ell^+\ell^-)$	J/ψ	Modelled in fit
	$\Lambda_b^0 \to pK\ell^+\ell^-$	rare	$< 0.2\,\%$
$\pi \to \ell$	$B^0 \to (D^{*-} \to (\overline{D}^0 \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$	all	$< 0.4\,\%$
	$B^0 \to (\overline{D}^0 \to K^+\pi^-)\pi^-\ell^+\nu_\ell$	all	$< 0.1\,\%$
	$B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$	all	$< 0.5\,\%$
$K \leftrightarrow \pi$ swap	$B^0 \to (K^{*0} \to K^+\pi^-)(J/\psi \to \ell^+\ell^-)$	J/ψ	Modelled in fit
	$B^0 \to (K^{*0} \to K^+\pi^-)\ell^+\ell^-$	rare	$< 0.4\,\%$
$h \leftrightarrow \ell$ swap	$B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$	all	$< 0.4\,\%$
	$B^0 \to K^{*0}\psi(2S)(\to \ell^+\ell^-)$	all	$< 0.1\,\%$

***K* → *π* misidentification:** $B_s^0 \to (\phi \to K^+K^-)\ell^+\ell^-$ decays constitute the only relevant source for $K \to \pi$ misidentification. This background peaks close to the signal in the reconstructed B mass, because of the mass difference between the kaon and pion, making it particularly dangerous. Due to the incorrectly reconstructed mass of the hadronic system, a lot of these events fall outside the K^{*0} mass window. The relative efficiency for all applied offline selection criteria (except for the background vetoes introduced in this section) of this background mode is shown in Fig. 4.8, where the last step refers to the MVA selection that is introduced in the next section. A tightened PID requirement of $\texttt{ProbNN}_\pi(\pi) > 0.8$ is applied at the ϕ resonance in the $m(K^+\pi^-_{\to K})$ mass system to further suppress this background, while simultaneously maintaining a high signal efficiency. Both background rejection and signal efficiency for this background veto are shown in Fig. 4.8. With the tools described above, the expected yield of this background for the full Run 2p2 B^0 sample is found to be 0.32 and 0.79 events for low and central q^2, respectively. This corresponds to a background over signal yield ratio of $N_{\text{bkg}}/N_{\text{sig}} < 0.3\,\%$ for both low and central q^2, where the signal yields are taken from Tab. 4.25. Thus, the residual background level after full selection is considered to be negligible for the rare mode, while for the J/ψ resonant mode the remaining background is modelled in the mass fits.

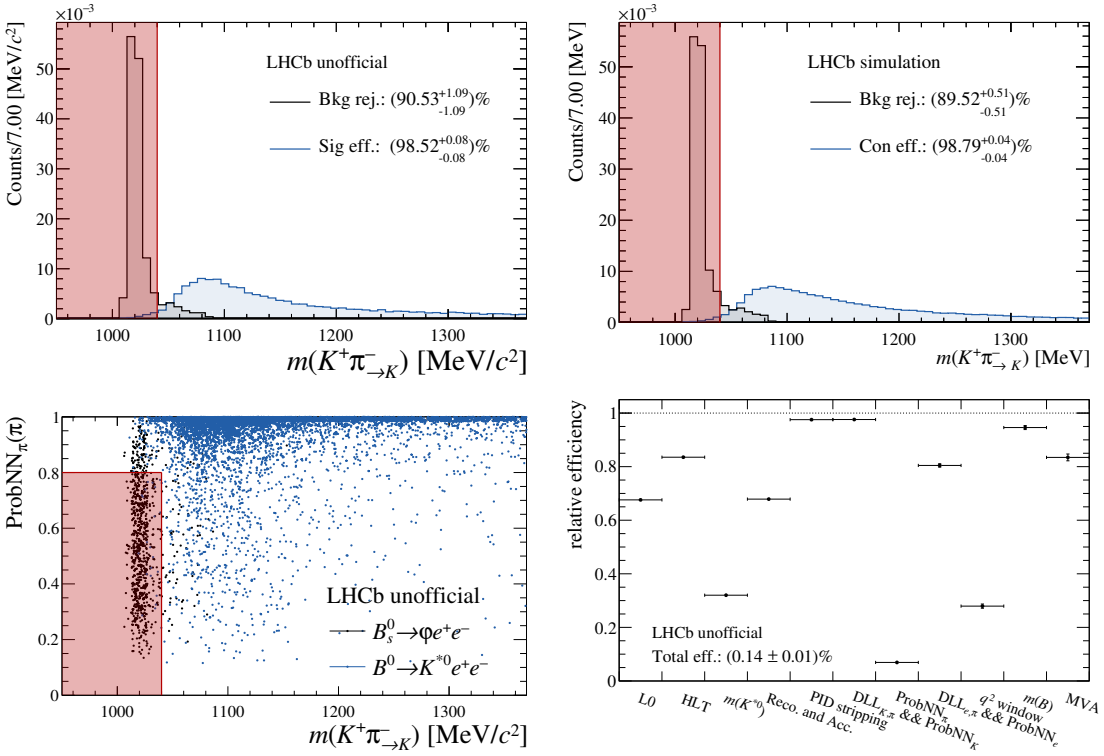

Figure 4.8.: The top plots show the efficiency of the veto against $\phi \to KK$ decays for central (J/ψ) q^2 region on the left (right). For central (J/ψ) q^2 $B^0_s \to (\phi \to K^+K^-)e^+e^-$ $(B^0_s \to (\phi \to K^+K^-)(J/\psi \to \ell^+\ell^-))$ MC is used as background proxy. The red area is rejected by the veto. On the bottom left is a two-dimensional plot of the veto in central q^2. The bottom right plot shows the relative efficiency of all offline selection steps for $B^0_s \to (\phi \to K^+K^-)e^+e^-$ MC in central q^2.

$p \to \pi$ misidentification: Because of baryon number conservation in the SM, $pX\ell^+\ell^-$ (with $X = K, \pi$) final states must originate from a baryonic mother particle. The Λ^0_b particle is the most frequently produced b-baryon at LHCb and upon $p \to X$ misidentification it falls into the signal region. Thus, the study focuses on $\Lambda^0_b \to pK\ell^+\ell^-$ decays, since $\Lambda^0_b \to p\pi\ell^+\ell^-$ is Cabibbo-suppressed. The efficiency scan in Fig. 4.9 shows, that this background is strongly suppressed by the K^{*0} mass window and the PID requirements. However, no dedicated veto is placed against this background since in neither the $m(K^+\pi^-_{\to p})$ nor the $m(K^+\pi^-_{\to p}e^+e^-)$ mass system good separation between signal and background distribution is found, which is a crucial requirement for an efficient cut. For low and central q^2 the remaining background level is found to be $N_{\text{bkg}}/N_{\text{sig}} < 0.2\,\%$. Similarly to the $B^0_s \to \phi\ell^+\ell^-$ background, the $\Lambda^0_b \to pK\ell^+\ell^-$ background is modelled in the J/ψ resonant fits, while it is considered negligible for the rare mode.

$\pi \to \ell$ misidentification: The only relevant source of $\pi \to \ell$ misidentification stems from semileptonic cascade decays $B^0 \to (D^{*-} \to (\overline{D}^0 \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$, $B^0 \to (\overline{D}^0 \to$

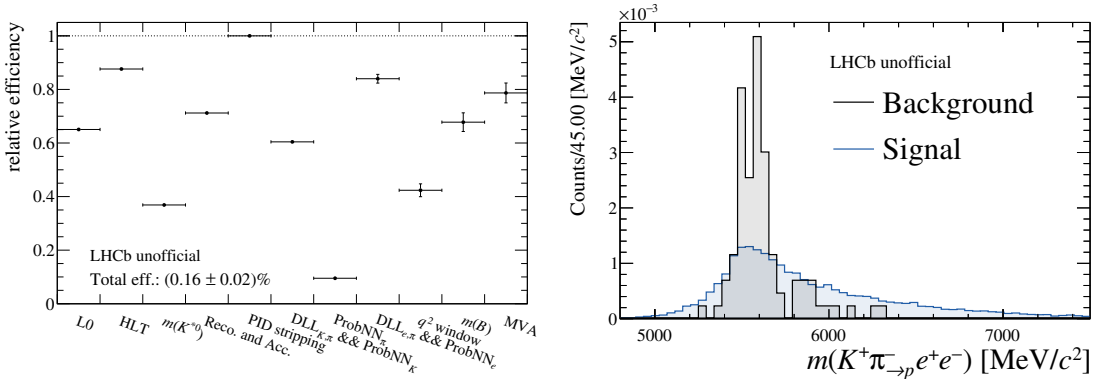

Figure 4.9.: The left plot shows the relative efficiency of all offline selection steps for $\Lambda_b^0 \to pKe^+e^-$ MC in central q^2. The right plot shows the $m(K^+\pi^-_{\to p}ee)$ distribution for both background and signal MC.

$K^+\pi^-)\pi^-\ell^+\nu_\ell$ and $B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$. These modes also fall into the category of partially reconstructed backgrounds, because of the missing neutrino. This is why these backgrounds populate the lower mass region below the signal in the reconstructed B^0 mass, making them less dangerous. The branching ratios of these modes can be several orders of magnitude higher compared to the signal mode as can be seen in the PDG [22]. Thus, a good rejection is nevertheless needed to control this background. Because of the incomplete reconstruction of the final state, a good separation is provided by the $\texttt{DIRA}(B^+)$ variable, which enters in the stripping requirements and the training of the MVA (see Sec. 4.3.5). Except for the last mentioned semileptonic decay, these modes are strongly suppressed by the K^{*0} mass selection. Further, dedicated vetoes on the intermediate $\overline{D}^0$ and D^- mesons are placed, which introduce a strong PID requirement in a $30\,\mathrm{MeV}/c^2$ window of the $m(K^+\ell^-_{\to\pi})$ and $m(K^+\pi^-\ell^-_{\to\pi})$ mass system, respectively. The efficiencies of both signal selection and background rejection of these vetoes are shown in Fig. 4.10. Moreover, the effect of these open charm vetoes are demonstrated on data in the central q^2 region in Fig. 4.11. The data plots show, that the dimuon mode exhibits stronger peaks at the open charm resonances. This can be explained by decays in flight of pions and kaons to muons ($\pi \to \mu\nu_\mu$, $K \to \mu\nu_\mu$), which are strongly suppressed for electrons and thus increase the relative misidentification rate for muons.

After the vetoes, all studied semileptonic backgrounds are beneath the 0.5 % level in the $N_{\mathrm{bkg}}/N_{\mathrm{sig}}$ ratio for the signal and resonant q^2 regions. Because of the typical mass shape for a partially reconstructed background, the residual pollution is expected to be absorbed by the combinatorial background component modelled in the mass fits.

***K* ↔ *π* swaps:** The main source of $K \leftrightarrow \pi$ swaps is the $B^0 \to K^{*0}J/\psi$ resonant channel, however also heavier $K\pi$ resonances can contribute. These events are strongly suppressed by the PID selection, because of the double misidentification. After the full selection, the background level is found to be $N_{\mathrm{bkg}}/N_{\mathrm{sig}} < 0.4\,\%$ in the J/ψ mode. Since the leptonic system is correctly reconstructed, these events only have a small leakage

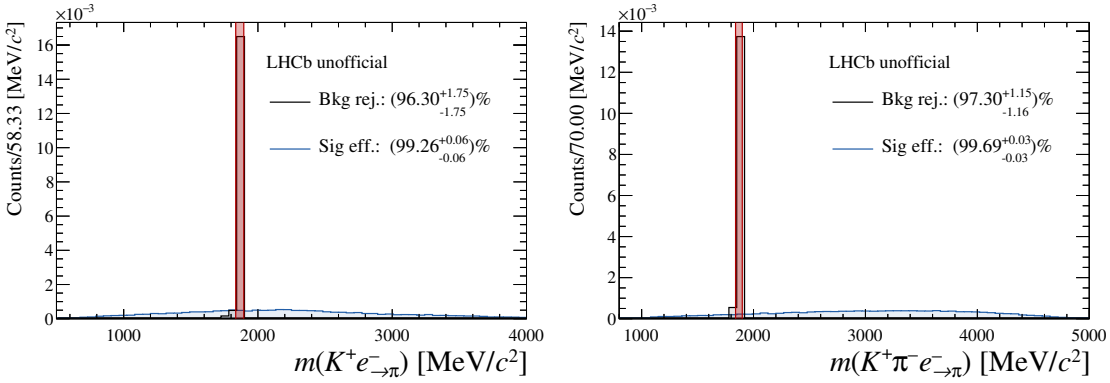

Figure 4.10.: Distribution of the $m(K^+e^-_{\to\pi})$ mass system from $B^0 \to (D^{*-} \to (\overline{D}^0 \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$ MC (left) and $m(K^+\pi^-e^-_{\to\pi})$ from $B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\pi^-)\ell^+\nu_\ell$ (right) in central q^2 together with the efficiency of the open charm vetoes.

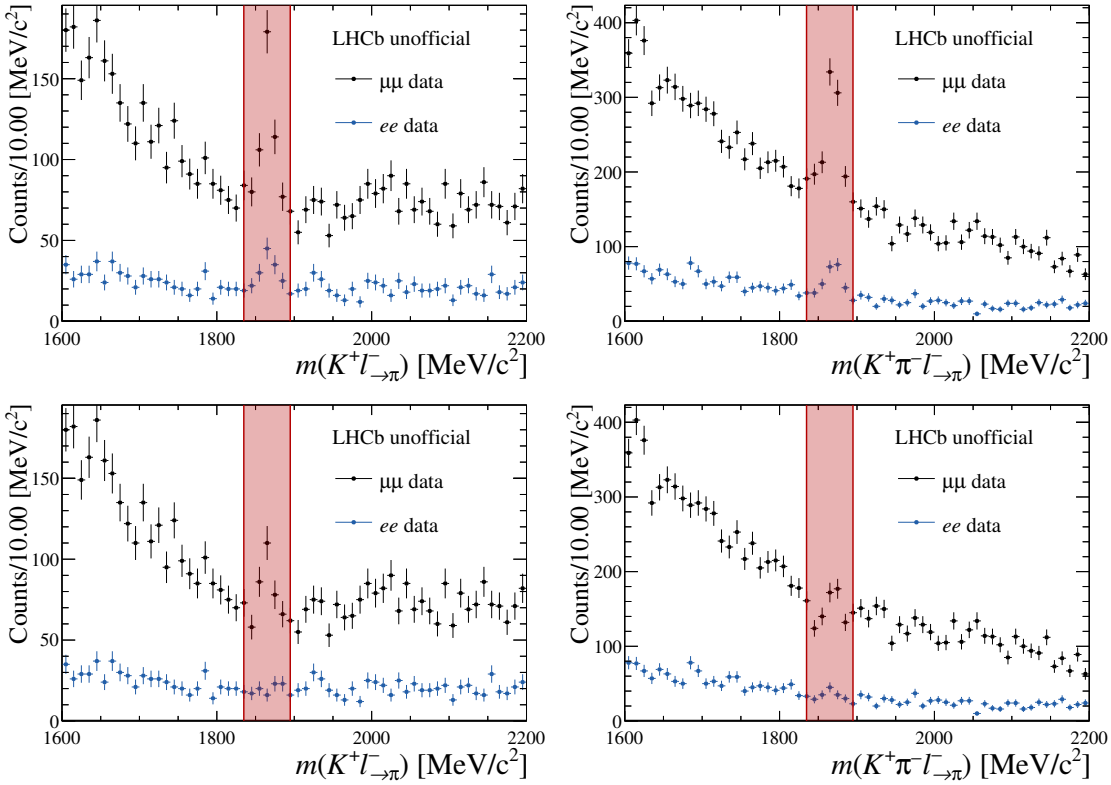

Figure 4.11.: Distribution of the $m(K^+\ell^-_{\to\pi})$ (left) and $m(K^+\pi^-\ell^-_{\to\pi})$ (right) mass systems on 2012 central q^2 data of the dielectron (blue) and dimuon channel (black). The top (bottom) plots show the distribution before (after) the open charm vetoes.

into the central q^2 region just like the nominal J/ψ resonant mode. The pollution from hadronic swaps from the rare mode itself is negligible ($N_{\rm bkg}/N_{\rm sig} < 0.4\,\%$).

$h \leftrightarrow \ell$ swaps: Swaps from J/ψ (or $\psi(2S)$) resonant decays, where a final state hadron h is swapped with a final state lepton, are much more dangerous since they no longer peak at the mass of the $c\bar{c}$ resonance. Thus, they can easily enter the signal q^2 regions. The charges of the swapped particles need to have the same sign, so that the cases of $K^+ \leftrightarrow \ell^+$ and $\pi^- \leftrightarrow \ell^-$ need to be considered. These events are strongly suppressed by the PID cuts, however due to the abundance of $B \to c\bar{c}X$ decays they still remain relevant which is why dedicated vetoes are applied. In the muon case, it is sufficient to apply a cut around the J/ψ and $\psi(2S)$ resonances in the $m(\mu^+_{\to K}\mu^-)$ and $m(\mu^+\mu^-_{\to\pi})$ mass systems to further suppress this background (see Fig. 4.12). For electrons, a similar veto would be very inefficient due to the worse resolution of the leptonic system. Therefore, a J/ψ (and $\psi(2S)$) mass constrained fit of the misidentified $m(K^+_{\to e}\pi^- e^+_{\to K} e^-)$ and $m(K^+\pi^-_{\to e} e^+ e^-_{\to\pi})$ mass systems is performed. The constraint greatly improves the resolution, allowing to introduce an efficient veto (see Fig. 4.13). The residual background level is at $N_{\text{bkg}}/N_{\text{sig}} < 0.3\,\%$ ($< 0.4\,\%$) in both low and central q^2 for the muon (electron) mode.

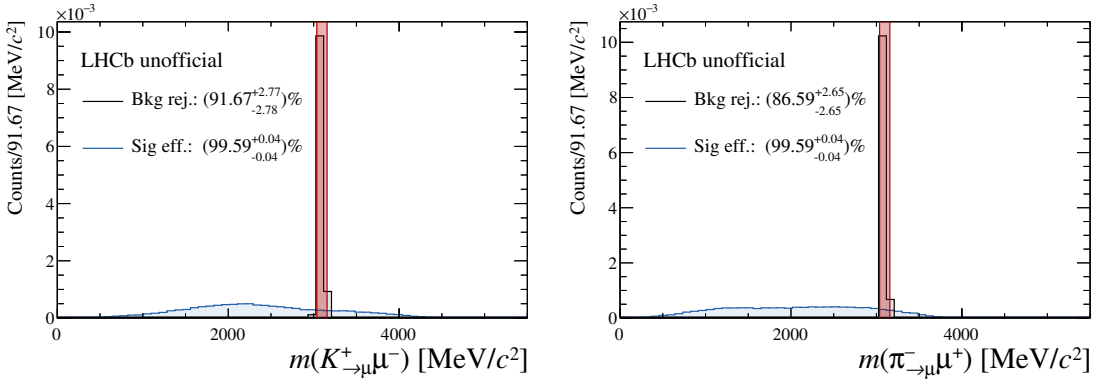

Figure 4.12.: Distribution of the $m(\mu^+_{\to K}\mu^-)$ (left) and $m(\mu^+\mu^-_{\to\pi})$ (right) mass systems in central q^2 together with the efficiency of the muon swap vetoes.

Partially Reconstructed Backgrounds in B^0 Modes

Partially reconstructed backgrounds pollute the lower mass sideband, because of the lost momentum carried by the particles, which are not reconstructed. However, especially for electron modes this type of background still remains important, since the poor resolution leads to an overlap with the left tail of the signal shape in the mass fits. All relevant sources of partially reconstructed backgrounds for B^0 modes are listed in Tab. 4.8.

Double semileptonic: The double semileptonic cascade decay $B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\ell^-\bar{\nu}_\ell)\ell^+\nu_\ell$ results in the same final state (also via a K^{*0} resonance) as the signal mode, which is why both PID and the K^{*0} mass window selection can not suppress this background. Only the two neutrinos are unique to the semileptonic mode making it a partially reconstructed background. A study based on a `generator-level` MC of the

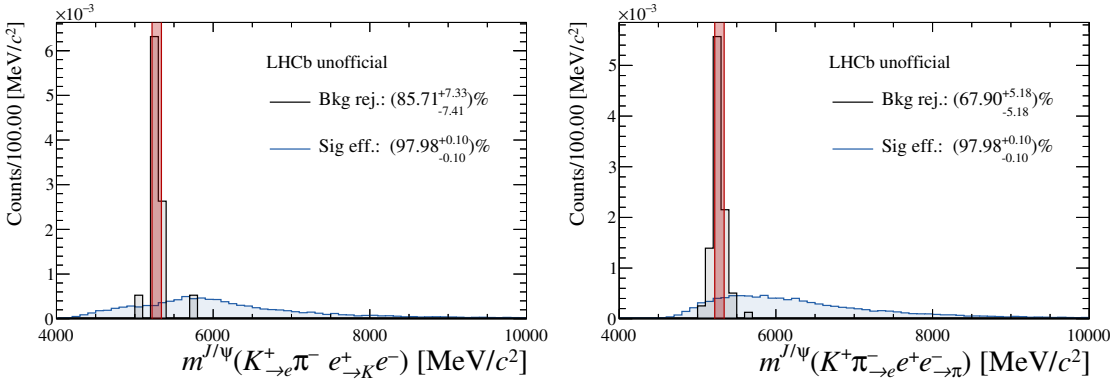

Figure 4.13.: Distribution of the $m^{J/\psi}(K^+_{\to e}\pi^- e^+_{\to K}e^-)$ (left) and $m^{J/\psi}(K^+\pi^-_{\to e}e^+e^-_{\to\pi})$ (right) mass systems from a J/ψ mass constrained fit in central q^2 together with the efficiency of the electron swap vetoes.

Table 4.8.: Summary of relevant sources of partially reconstructed backgrounds for B^0 modes. The $N_{\text{bkg}}/N_{\text{sig}}$ column shows the ratio of the expected background yield calculated with the Bayesian method and the rare yield taken from Tab. 4.25 for the electron mode in central q^2 (if the component is not modelled in the mass fit).

Decay mode	**q^2 region**	**$N_{\text{bkg}}/N_{\text{sig}}$**
$B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)\ell^-\bar{\nu}_\ell)\ell^+\nu_\ell$	all	$< 0.1\,\%$
$B^+ \to K^+\pi^+\pi^- J/\psi(\to \ell^+\ell^-)$	J/ψ	Modelled in fit
$B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$	rare	Modelled in fit

background mode has shown that due to the missing momentum of the neutrinos less than 1 % of these decays fall into the invariant mass region probed in the electron fits (even less in the muon case). Even so, this decay gives a significant contribution to the lower mass side band of the rare mode because of its large branching ratio (see PDG [22]). A dedicated veto is placed on the $m(K^+\pi^-\ell^-)$ mass system, which in case of this background mode must have smaller values than the mass of the D^- meson. The signal efficiency and background rejection of the veto are shown in Fig. 4.14. The residual background level is at $N_{\text{bkg}}/N_{\text{sig}} < 0.1\,\%$ in both low and central q^2.

Hadronic and charmonium: The main source of partially reconstructed background, where a final state hadronic is not reconstructed, is coming from $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$ decays. The $K^+\pi^+\pi^-$ system can be produced directly or stem from intermediate resonances such as $K_1^+(1270)$ mesons. These events can be reconstructed as signal candidates, if the π^+ is lost. In the J/ψ resonant mode, an additional source of background arises from higher $c\bar{c}$ resonances such as $B^0 \to K^{*0}(\psi(2S) \to J/\psi\,\pi^+\pi^-)$, where the

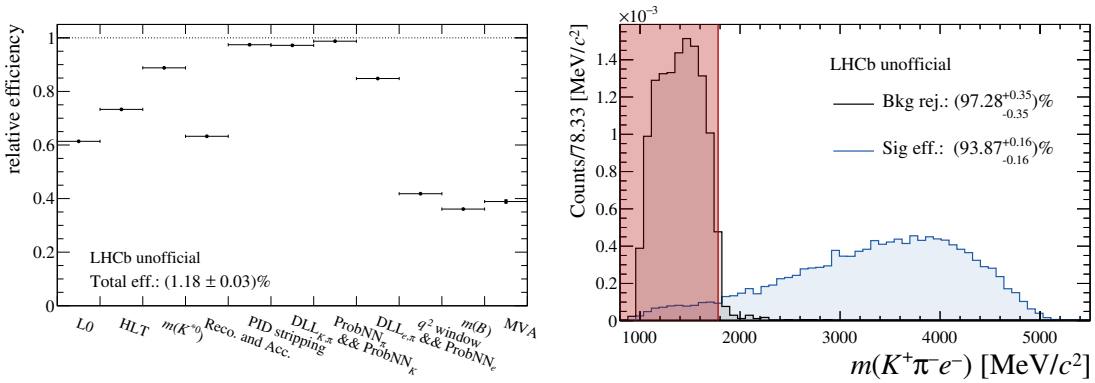

Figure 4.14.: The left plot shows the relative efficiency of all offline selection steps for $B^0 \to (D^- \to (K^{*0} \to K^+\pi^-)e^-\bar{\nu}_e)e^+\nu_e$ MC in central q^2. The right plot shows the $m(K^+\pi^-e^-)$ distribution for both background and signal MC.

particles accompanying the J/ψ are not reconstructed. These modes are referred to as 'charmonium' partially reconstructed backgrounds. For both signal and J/ψ resonant channel partially reconstructed modes are modelled in the mass fits. However, for the signal q^2 region a dedicated MVA based selection (see Sec. 4.3.5) is applied to reduce this background component.

Over-Reconstructed Backgrounds in B^0 Modes

Over-reconstructed backgrounds stem from $B^+ \to K^+\ell^+\ell^-$ decays, when a pion from the underlying event is added to the candidate. As mentioned before, this type of background is falling into the category of combinatorial backgrounds and is thus suppressed by the combinatorial MVA. Additionally, a dedicated requirement on the $m(K^+\ell^+\ell^-)$ mass system is applied cutting all events above 5100 MeV/c^2, which roughly represents the mass difference of the B^0 and π^- mesons. The signal efficiency and background rejection of the veto are shown in Fig. 4.15. The residual background level is at $N_{\text{bkg}}/N_{\text{sig}} < 0.1\,\%$ in both low and central q^2.

Misidentified Backgrounds in B^+ Modes

An analogous strategy as used for the B^0 modes is applied to the misidentified backgrounds of the B^+ modes. All relevant sources of misidentified backgrounds for B^+ modes are separated according to their type of particle misidentification and are listed in Tab. 4.9.

$\pi \to K$ misidentification: The only relevant source of $\pi \to K$ misidentification originates from the Cabibbo-suppressed $B^+ \to \pi^+\ell^+\ell^-$ decay. Since no intermediate resonance occurs in the hadronic system, it is not possible to introduce a dedicated veto

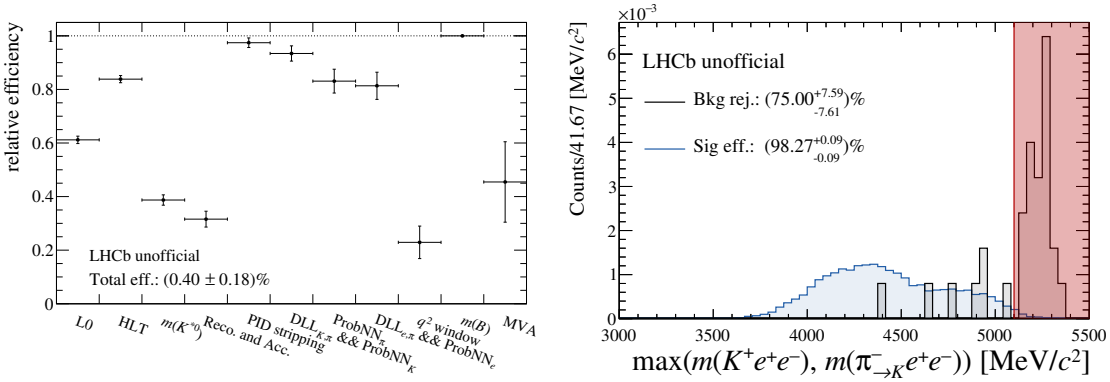

Figure 4.15.: The left plot shows the relative efficiency of all offline selection steps for $B^+ \to K^+e^+e^-$ MC in central q^2. The right plot shows the $m(K^+e^+e^-)$ distribution for both background and signal MC.

Table 4.9.: Summary of relevant sources of misidentified backgrounds for B^+ modes. The $N_{\rm bkg}/N_{\rm sig}$ column shows the ratio of the expected background yield calculated with the Bayesian method and the rare yield taken from Tab. 4.25 for the electron mode in central q^2 (if the component is not modelled in the mass fit).

Type	**Decay mode**	q^2 **region**	$N_{\rm bkg}/N_{\rm sig}$
$\pi \to K$	$B^+ \to \pi^+ J/\psi(\to \ell^+\ell^-)$	J/ψ	Modelled in fit
	$B^+ \to \pi^+\ell^+\ell^-$	rare	$< 0.1\,\%$
$\pi \to \ell$	$B^+ \to (\overline{D}^0 \to K^+\pi^-)\ell^+\nu_\ell$	all	$< 0.1\,\%$
	$B^+ \to (\overline{D}^0 \to K^+\ell^-\bar{\nu}_\ell)\pi^+$	all	$< 0.1\,\%$
$h \leftrightarrow \ell$ swap	$B^+ \to K^+ J/\psi(\to \ell^+\ell^-)$	all	$< 0.5\,\%$
	$B^+ \to K^+\psi(2S)(\to \ell^+\ell^-)$	all	$< 0.1\,\%$

against this mode. For the J/ψ resonant mode, this background is modelled in the invariant mass fits. For the rare mode, the residual background level is found to be $N_{\rm bkg}/N_{\rm sig} < 0.1\,\%$ and considered negligible.

$\pi \to \ell$ misidentification: Similarly as for B^0, the relevant sources for $\pi \to \ell$ misidentification stem from semileptonic cascade decays such as $B^+ \to (\overline{D}^0 \to K^+\pi^-)\ell^+\nu_\ell$ and $B^+ \to (\overline{D}^0 \to K^+\ell^-\bar{\nu}_\ell)\pi^+$. A strong PID requirement is applied to the reconstructed $m(K^+\pi^-_{\to\ell})$ mass system around the D^0 mass in order suppress the $B^+ \to (\overline{D}^0 \to K^+\pi^-)\ell^+\nu_\ell$ mode. For the $B^+ \to (\overline{D}^0 \to K^+\ell^-\bar{\nu}_\ell)\pi^+$ mode, a cut is applied on $m(K^+\ell^-)$ to reject events below the D^0 mass. Both vetoes and their efficiencies are shown in Fig. 4.16. After full selection the residual background level is $N_{\rm bkg}/N_{\rm sig} < 0.1\,\%$ for both background modes in low and central q^2.

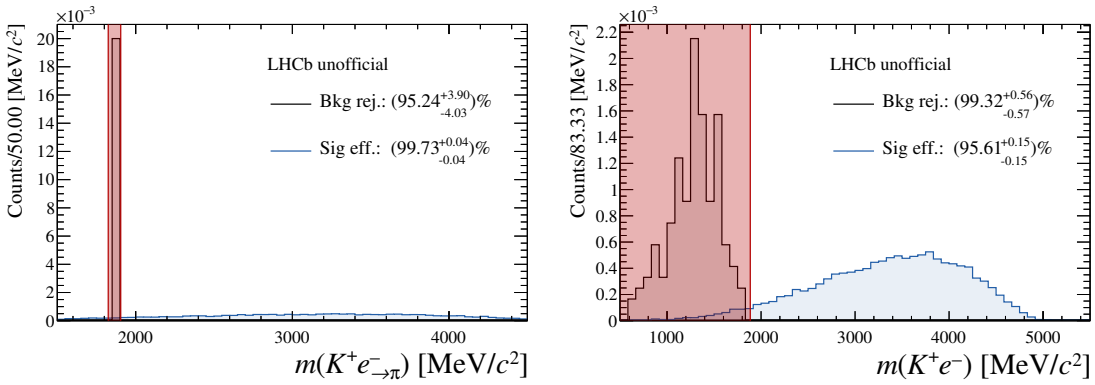

Figure 4.16.: The left plot shows the $m(K^+e^-_{\to\pi})$ distribution with the veto on the D^0 resonance for both $B^+ \to (\overline{D}{}^0 \to K^+\pi^-)e^+\nu_e$ background and signal MC. On the right, the $m(K^+e^-)$ mass system is shown for $B^+ \to (\overline{D}{}^0 \to K^+e^-\bar{\nu}_e)\pi^+$ background and signal MC.

***h* ↔ *ℓ* swaps:** For B^+ modes only $K \leftrightarrow \ell^+$ swaps need to be considered. The placed vetoes are designed in an analogous way to the respective B^0 cuts and are shown in Fig. 4.17. The residual background level is at $N_{\text{bkg}}/N_{\text{sig}} < 0.2\,\%$ ($< 0.5\,\%$) in both low and central q^2 for the muon (electron) mode.

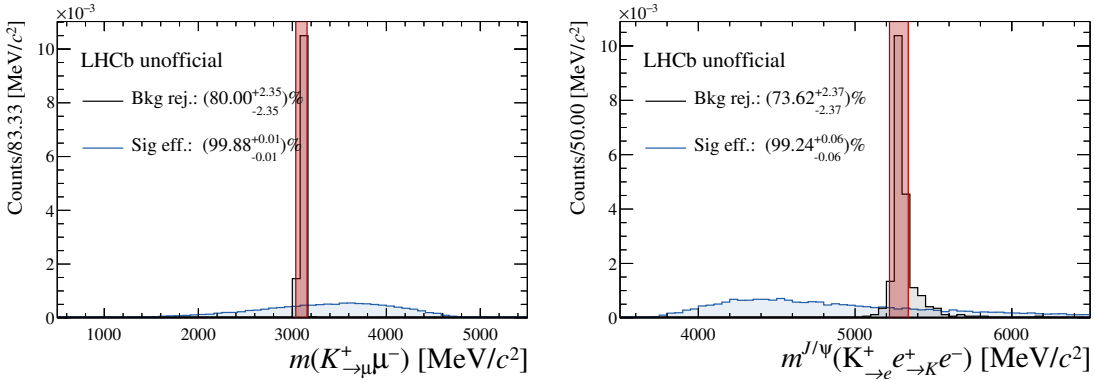

Figure 4.17.: Distribution of the $m(\mu^+_{\to K}\mu^-)$ (left) and $m^{J/\psi}(K^+_{\to e}e^+_{\to K}e^-)$ (right) mass systems in central q^2 together with the efficiency of the swap vetoes.

Partially Reconstructed Backgrounds in B^+ Modes

All relevant sources of partially reconstructed backgrounds for B^+ modes are listed in Tab. 4.10.

Double semileptonic: A similar behaviour to its B^0 counterpart is shown by the double semileptonic cascade decay $B^+ \to (\overline{D}{}^0 \to K^+\ell^-\bar{\nu}_\ell)\ell^+\nu_\ell$. Again, the probability of falling into the B^+ mass region used in the fits is below 1 %. The veto that was

Table 4.10.: Summary of relevant sources of partially reconstructed backgrounds for B^+ modes. The $N_{\text{bkg}}/N_{\text{sig}}$ column shows the ratio of the expected background yield calculated with the Bayesian method and the rare yield taken from Tab. 4.25 for the electron mode in central q^2 (if the component is not modelled in the mass fit).

Decay mode	**q^2 region**	**$N_{\text{bkg}}/N_{\text{sig}}$**
$B^+ \to (\overline{D}^0 \to K^+\ell^-\bar{\nu}_\ell)\ell^+\nu_\ell$	all	$< 0.1\,\%$
$B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$	J/ψ	Modelled in fit
$B^0 \to K^{*0}\ell^+\ell^-$	rare	Modelled in fit
$B^+ \to K^+\eta' \to (e^+e^-\gamma)$	low	Modelled in fit

introduced for the $B^+ \to (\overline{D}^0 \to K^+\ell^-\bar{\nu}_\ell)\pi^-$ decay is also efficient to reject the double semileptonic cascade decay, which is demonstrated in Fig. 4.18. The residual background level is $N_{\text{bkg}}/N_{\text{sig}} < 0.1\,\%$ in low and central q^2.

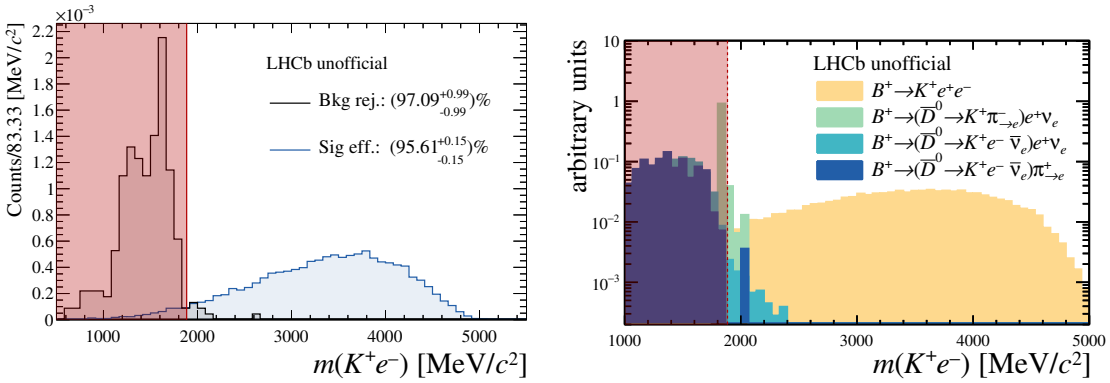

Figure 4.18.: On the left, the $m(K^+e^-)$ distribution is shown for both $B^+ \to (\overline{D}^0 \to K^+e^-\bar{\nu}_e)e^+\nu_e$ background and signal MC. The right plot shows the same distribution adding also the other cascade modes considered for the B^+ modes.

Hadronic and charmonium: These types of background are treated in the same way as introduced for the B^0 modes. The main source of hadronic partially reconstructed background originates from $B^0 \to K^{*0}\ell^+\ell^-$ decays, where also the non-resonant $K^+\pi^-$ mode and the isospin partners via the K^{*+} meson need to be considered, when modelling the residual background in the invariant mass fits (see Sec.4.6). For the dielectron mode, also the $B^+ \to K^+\eta' \to (e^+e^-\gamma)$ mode needs to be considered, which decays into the same final state as the signal except for the lost photon. Since the electrons from the η' are very soft, this mode is only relevant for the low q^2 region and is modelled in the fit.

Summary of Physical Backgrounds

The full list of requirements applied to reduce physical backgrounds for the B^0 modes can be found in Tab. 4.11. Many of these are updated or added with respect to the RUN 1 $R_{K^{*0}}$ analysis [11]. These modifications are a result of detailed studies based on new MC samples and provide an increase in signal efficiency of approximately 10 %, while also providing improved background rejection especially against semi leptonic cascades and $h \leftrightarrow \ell$ swaps.

Table 4.11.: Summary of the physical background selection for B^0 modes.

Decay mode	**Requirement**	q^2 **region**
$B_s^0 \to \phi\ell^+\ell^-$	$!(m(K^+K^+_{\to\pi}) < 1040\,\text{MeV}/c^2$ && $\texttt{ProbNN}_\pi(\pi) < 0.8)$	all
$B^0 \to (\overline{D}^0 \to K^+\pi^-)\pi^-\ell^+\nu_\ell$	$!(\lvert m(K^+\ell^-_{\to\pi}) - m^{\text{PDG}}_{D^0}\rvert < 30\,\text{MeV}/c^2$ && $\texttt{ProbNN}_\ell(\ell) < 0.8)$	all
$B^0 \to (D^- \to K^{*0}\pi^-)\ell^+\nu_\ell$	$!(\lvert m(K^+\pi^-\ell^-_{\to\pi}) - m^{\text{PDG}}_{D^-}\rvert < 30\,\text{MeV}/c^2$ && $\texttt{ProbNN}_\ell(\ell) < 0.8)$	all
$h \leftrightarrow \ell$ swap	$!(\lvert m(\mu_{\to h}\mu) - m^{\text{PDG}}_{J/\psi}\rvert < 60\,\text{MeV}/c^2$ && $\texttt{ProbNN}_\mu(\mu) < 0.8)$	all $\mu\mu$
	$!(\lvert m^{J/\psi}(h_{\to e}he_{\to h}e) - m^{\text{PDG}}_{B^0}\rvert < 60\,\text{MeV}/c^2$ && $\texttt{ProbNN}_e(e) < 0.8)$	all ee
$B^0 \to (D^- \to K^{*0}\ell^-\bar{\nu_\ell})\ell^+\nu_\ell$	$m(K^+\pi^-\ell^-) > 1780\,\text{MeV}/c^2$	low, central
$B^+ \to K^+\ell^+\ell^-$	$max(m(K\ell^+\ell^-), m(K^+_{\to\pi}\ell^+\ell^-)) < 5100\,\text{MeV}/c^2$	all

The list of requirements applied to suppress physical backgrounds on the B^+ mode is given in Tab. 4.12. This set of selection criteria has similar powerful background rejection as for the B^0 modes and an even higher signal efficiency, since fewer background modes need to be vetoed. The signal efficiencies and background rejections of all applied vetoes are summarised in Fig. 4.19.

Table 4.12.: Summary of the physical background selection for B^+ modes.

Decay mode	**Requirement**	q^2 **region**
$B^+ \to (\overline{D}^0 \to K^+\pi^-)\ell^+\nu_\ell$	$!(\lvert m(K^+\ell^-_{\to\pi}) - m^{\text{PDG}}_{D^0}\rvert < 40\,\text{MeV}/c^2$ && $\texttt{ProbNN}_\ell(\ell) < 0.8)$	all
$K \leftrightarrow \ell$ swap	$!(\lvert m(\mu^+_{\to K}\mu^-) - m^{\text{PDG}}_{J/\psi}\rvert < 60\,\text{MeV}/c^2$ && $\texttt{ProbNN}_\mu(\mu) < 0.8)$	all $\mu\mu$
	$!(\lvert m^{J/\psi}(K^+_{\to e}e^+_{\to K}e^-) - m^{\text{PDG}}_{B^+}\rvert < 60\,\text{MeV}/c^2$ && $\texttt{ProbNN}_e(e) < 0.8)$	all ee
$B^+ \to (\overline{D}^0 \to K^+\ell^-\bar{\nu_\ell})\ell^+\nu_\ell$	$m(K^+\ell^-) > 1885\,\text{MeV}/c^2$	low, central

All background modes that have a significant contamination in the data samples even after the full selection is applied are modelled in the mass fits as described in Sec. 4.6. These modes are listed in Tab. 4.23.

Additional studies on backgrounds involving double hadronic to electron misidentifications from decays such as $B^+ \to K^+\pi^+\pi^-$ for B^+ and $B^0 \to (K^{*0} \to K^+\pi^-)\pi^+\pi^-$ for B^0 modes have been investigated. The residual contamination after full selection is estimated in a data-driven way and covered by a systematic uncertainty. These studies can be reviewed in Sec. 4.7.2.

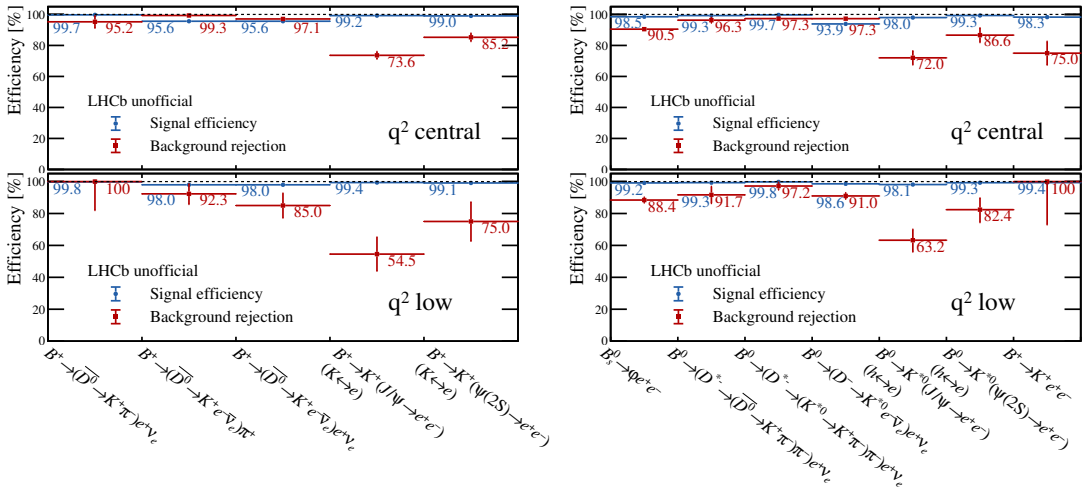

Figure 4.19.: Summary of the signal efficiency and background rejection of all vetoes against physical backgrounds for B^+ (left) and B^0 (right) calculated on 2018 MC samples.

Table 4.13.: Backgrounds that cannot be vetoed that must be modelled in the mass fits, together with the q^2 regions they are important for

Decay mode	**Relevant for q^2 region**	**Importance**
$B^+ \to \pi^+ J/\psi (\to \ell^+\ell^-)$	J/ψ	yes
$B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$	J/ψ	yes
$B^0 \to K^{*0}\psi(2S) (\to \ell^+\ell^-)$	$\psi(2S)$	yes
$B^+ \to \pi^+\ell^+\ell^-$	signal	negligible
$B^0 \to K^{*0}\ell^+\ell^-$	signal	in ee mode
$B^+ \to K^+\eta' (\to e^+e^-\gamma)$	low	in ee mode

4.3.5 MVA Selection and HOP

After applying all previous selection stages, the signal purity is already at an acceptable level for the J/ψ resonant mode. However, the data samples of the rare electron modes still remain dominated by combinatorial and partially reconstructed backgrounds, which is why additional requirements on MVA responses and on the so-called HOP mass [103] are applied. In particular, a multivariate classifier dedicated to suppress combinatorial background is used for both muon and electron modes in all probed q^2 regions, while for the rare electron mode an additional MVA as well as the HOP requirement are applied to reduce partially reconstructed backgrounds. These selection steps are described in this section.

Combinatorial MVA

The distributions of a large set of kinematic variables is different for combinatorial background events with respect to signal. However, none of these variables alone provide separation between signal and background distributions that can be exploited with a simple selection criterion that gives strong background rejection with high signal efficiency. Instead, a MVA approach is chosen that takes multiple input variables into account to increase the separation power. In this analysis, the `CatBoost` algorithm [104] is used, because it provides the best performance of all tested MVA algorithms.

The first step when using a multivariate classifier is the so-called 'training' phase in which the algorithm learns how to separate signal from background. Separate classifiers are trained for the muon and electron modes, B^+ and B^0 channels as well as for the three run periods resulting in a total of 12 combinatorial MVAs. Each classifier is used for all studied regions of q^2 and both L0 trigger categories. The training requires proxy samples for both signal and background. Since the selection is designed to increase the sensitivity to the FCNC decays, simulated samples of the rare mode ($B^+ \to K^+\ell^+\ell^-$ for B^+ and $B^0 \to K^{*0}\ell^+\ell^-$ for B^0 modes) are used as signal proxy. Taking real data would not provide a sufficient sample size for the training and would also increase the risk of introducing a bias. The background proxy is extracted from real data in the upper mass sideband defined as $m(B) > 5400\,\text{MeV}/c^2$ and $m(B) > 5600\,\text{MeV}/c^2$ for muon and electron modes, respectively, with the upper limit being set by the stripping lines (see Tab. 4.3). The used phase-space of the background proxy is further selected to fall into the q^2 region of the rare modes ($0.1\,\text{GeV}^2/c^4 < q^2 < 6.0\,\text{GeV}^2/c^4$) and can be viewed in Fig. 4.20. This particular phase-space region is selected, since unlike the lower mass sideband, it is free of partially reconstructed backgrounds and thus provides a clean proxy of combinatorial events.

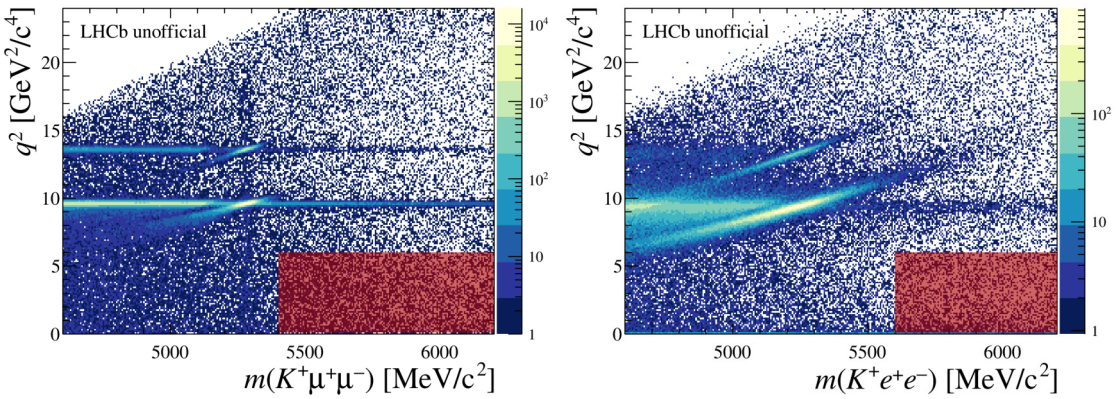

Figure 4.20.: The q^2 system plotted against the invariant mass of the B meson for 2018 data after full selection (except MVA and HOP) of $K^+\ell^+\ell^-$ candidates in the dimuon final state (left) and dielectron (right). The red area shows the phase-space region used to extract the background proxy for the combinatorial MVA.

Since the same signal MC samples are also used to extract the signal efficiencies and the background proxy is overlapping with the mass window used in the rare mode fits, it is crucial to protect against biases from this double usage. Thus, the `k-folding` method [105] with 10 folds is used: The signal and background proxy samples are divided randomly into 10 parts. Next, 10 separate classifiers are trained iteratively each using $^9/_{10}$ of the proxy samples for training and the remaining $^1/_{10}$ to apply the response of the MVA. The $^1/_{10}$ fraction used to apply the response is unique to each of the 10 classifiers. This way, nearly the full sample is available to train the MVAs, while no MVA is applied to events that where also used for its training.

All selection steps from the previous sections are applied to the proxy samples. To increase the available statistics in the background sample for B^0 modes, the K^{*0} mass window selection is loosened to $|m(K\pi) - m^{\mathrm{PDG}}_{K^{*0}}| < 200\,\mathrm{MeV}/c^2$. Still the statistics of the background sample is in each case smaller than the signal proxy. In order to reduce computational effort, the same number of signal and background events are used meaning that events from the larger sample are randomly removed until the sample sizes are equal. The available statistics of both signal and background samples are given in Tab. 4.14.

Table 4.14.: Statistics of signal and background proxy samples used to train the combinatorial MVA. The number in brackets denotes the full size of the sample before aligning to the smaller proxy.

Classifier		**Signal events**	**Background events**
$B^+ \to K^+\mu^+\mu^-$	Run 1	23217 (102991)	23217 (limiting)
	Run 2p1	40417 (65970)	40417 (limiting)
	Run 2p2	68222 (122538)	68222 (limiting)
$B^+ \to K^+e^+e^-$	Run 1	4715 (18498)	4715 (limiting)
	Run 2p1	10489 (26291)	10489 (limiting)
	Run 2p2	16286 (63019)	16286 (limiting)
$B^0 \to K^{*0}\mu^+\mu^-$	Run 1	3884 (43132)	3884 (limiting)
	Run 2p1	5331 (66247)	5331 (limiting)
	Run 2p2	9382 (118276)	9382 (limiting)
$B^0 \to K^{*0}e^+e^-$	Run 1	1051 (42800)	1051 (limiting)
	Run 2p1	2095 (62512)	2095 (limiting)
	Run 2p2	3869 (66997)	3869 (limiting)

The combinatorial MVA is trained on variables that describe the kinematics of the decay as well as the quality of tracks and vertices, since these are expected to differ between signal and combinatorial events. The same variables are used for electron and muon

modes as well as between the three run periods, while they differ between B^+ and B^0 modes. Preliminary MVAs are trained on a large set of input variables, that are subsequently sorted by their discrimination power. In an iterative procedure, the lowest ranked variables are removed until a significant change ($> 1\,\%$) of the area underneath the Receiver Operating Characteristic (ROC) curve occurs. This way, 16 variables remain for B^+ and 23 variables for B^0 classifiers which are listed in Tab. 4.15 and Tab. 4.16, respectively.

Table 4.15.: Summary of the input variables used to train the combinatorial MVA for B^+ modes.

Particle	**Variables**
B^+	p_T, χ^2_IP, χ^2_FD, $\chi^2_\mathrm{vtx}/\mathrm{ndf}$, $\chi^2_\mathrm{DTF}/\mathrm{ndf}$, `DIRA`
$\ell^+\ell^-$	p_T, χ^2_IP, χ^2_FD, $\chi^2_\mathrm{vtx}/\mathrm{ndf}$, `DIRA`
K^+	p_T, χ^2_IP
ℓ	$min(p_\mathrm{T}(\ell^+), p_\mathrm{T}(\ell^-))$, $min, max(\chi^2_\mathrm{IP}(\ell^+), \chi^2_\mathrm{IP}(\ell^-))$

Table 4.16.: Summary of the input variables used to train the combinatorial MVA for B^0 modes.

Particle	**Variables**
B^0	p_T, χ^2_IP, χ^2_FD, $\chi^2_\mathrm{vtx}/\mathrm{ndf}$, $\chi^2_\mathrm{DTF}/\mathrm{ndf}$, `DIRA`
K^{*0}	p_T, χ^2_IP, χ^2_FD, $\chi^2_\mathrm{vtx}/\mathrm{ndf}$, `DIRA`
$\ell^+\ell^-$	p_T, χ^2_IP, χ^2_FD, $\chi^2_\mathrm{vtx}/\mathrm{ndf}$, `DIRA`
K^+, π^-	$min, max(p_\mathrm{T}(K^+), p_\mathrm{T}(\pi^-))$, $min, max\,(\chi^2_\mathrm{IP}(K^+), \chi^2_\mathrm{IP}(\pi^-))$
ℓ	$min(p_\mathrm{T}(\ell^+), p_\mathrm{T}(\ell^-))$, $min, max(\chi^2_\mathrm{IP}(\ell^+), \chi^2_\mathrm{IP}(\ell^-))$

After the training, the classifiers provide a variable (response) which is used to introduce a selection requirement against combinatorial background. However, before a cut threshold is determined, several cross-checks are performed to ensure a successful training. In the following, these checks are demonstrated using the Run 2p2 B^+ classifiers, which are representative for all trainings. First, the ROC curves of all 10 folds are compared to test the stability in the individual trainings. Figure 4.21 shows, that the trend of the curves are very compatible and the averaged area under the curve (AUC) is 0.993 ± 0.001 and 0.9928 ± 0.0004 for the electron and muon classifier, respectively, which demonstrates the high separation power. Moreover, a comparison of the MVA response on the dataset used for training against the remaining $1/10$ part of the data is performed. A significant difference in these distributions would be a sign of overtraining, *i.e.* the MVA is sensitive to statistical fluctuations in the training samples. This check is shown in Fig. 4.22 where no sign of overtraining is visible.

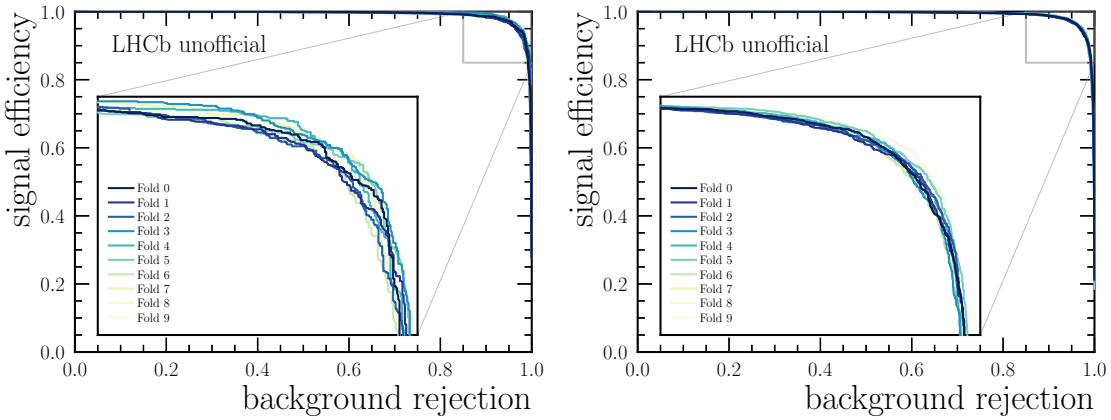

Figure 4.21.: ROC curves of all 10 folds for the Run 2p2 electron (left) and muon (right) combinatorial MVA.

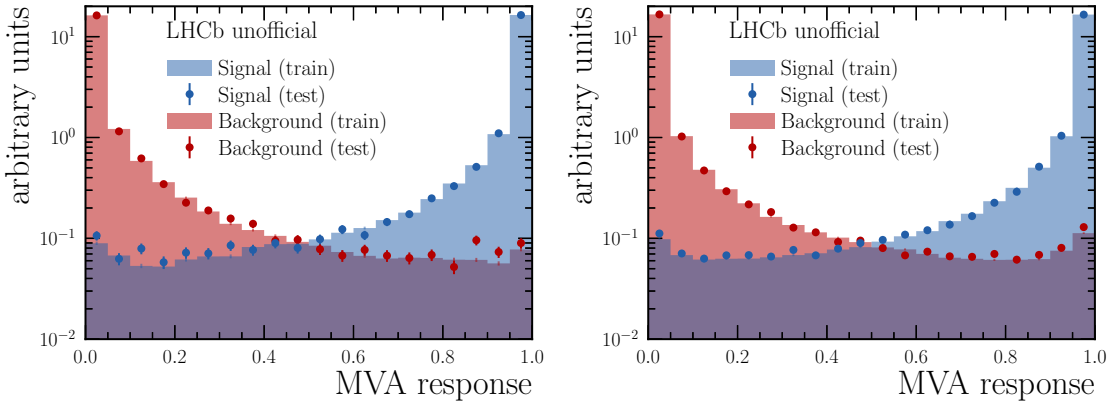

Figure 4.22.: Overtraining plot averaged over all folds for the Run 2p2 electron (left) and muon (right) combinatorial MVA. The blue and red areas show the MVA response on the signal and background data used for training, while the blue and red data points show the same on the remaining 1/10 part of the data.

The MC samples used in the training do not have the corrections to simulation discussed in Sec. 4.4 applied. The reason for this is, that classifiers trained on corrected simulation showed no significant difference in their performance. To ensure that no bias arises from the training on uncorrected MC, the response of the MVA is compared on the J/ψ resonant channel between simulation and data. In order to increase the signal purity, the data used in this check is selected in a tight window around the B^+ mass in the reconstructed $K^+\ell^+\ell^-$ mass system with a J/ψ mass constraint applied ($|m^{J/\psi}(K^+\ell^+\ell^-) - m^{\mathrm{PDG}}_{B^+}| < 60\,\mathrm{MeV}/c^2$).

The simulated samples used in this cross-check have the full MC corrections applied, in order to test the agreement on the distributions used to extract the final signal efficiencies. The comparison is shown in Fig. 4.23 again for the Run 2p2 B^+ classifiers. A

good agreement between data and corrected MC is visible enabling a reliable efficiency estimation of the MVA selection. Residual differences are covered by the systematic uncertainty assigned to the non-flatness of the double ratio as detailed in Sec. 4.7.1. Similar plots for variables most important to the analysis (including all MVA training variables) are given in App. C.

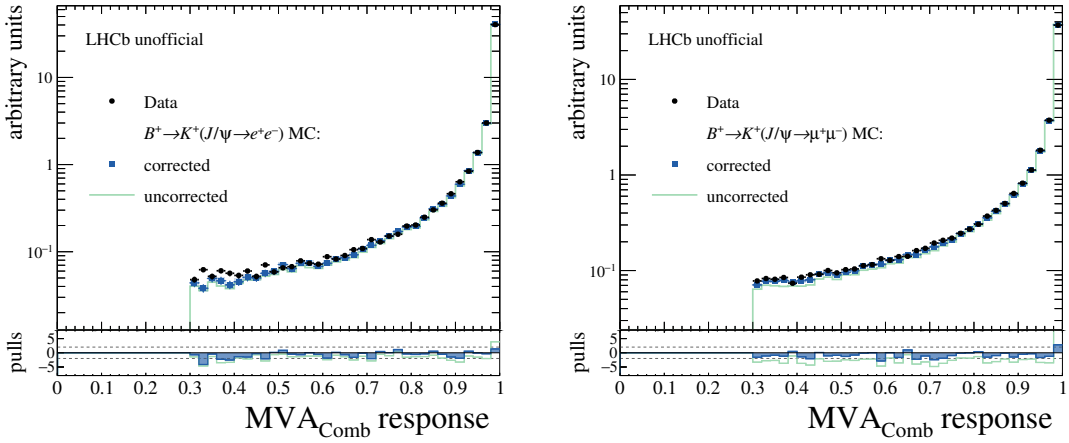

Figure 4.23.: Comparison of data (black) vs. MC with (blue) and without (green) corrections applied on Run 2p2 electron (left) and muon (right) samples for the combinatorial MVA response. The full selection is applied with loosened MVA cuts.

Partially Reconstructed MVA

The strategy used for the training of the partially reconstructed MVA is analogous to the combinatorial MVA. As mentioned above, this multivariate selection is only trained and applied for the electron modes. Simulated samples of the rare mode are used as signal proxy. In contrast to the combinatorial MVA, simulated samples are used as background proxy as well. For B^+ modes $B^0 \to K^{*0}e^+e^-$ MC is used as background proxy, while B^0 classifiers are trained with $B^+ \to K^+\pi^+\pi^-e^+e^-$ MC. The available statistics of the training samples are detailed in Tab. 4.17.

The variables used to train the MVAs are unchanged between the three run periods and different for B^+ and B^0 modes. In both cases, a set of 14 variables are determined in an analogous way as used for the combinatorial MVA. As before, the variables describe the decay kinematic as well as track and vertex quality, however the training of these classifiers also features 'isolation' variables which are defined in App. B. The used variables are listed in Tab. 4.18 for B^+ and Tab. 4.19 for B^0 classifiers.

The same cross-checks are performed to test the quality of the training of the partially reconstructed MVAs, which can be seen for the Run 2p2 B^+ classifier in Fig. 4.24. The AUC score averaged over all folds is 0.831 ± 0.006, which is a lot lower compared to the combinatorial MVA, but still represents good discrimination power. The lower AUC score reflects, that partially reconstructed backgrounds are much more similar to signal

Table 4.17.: Statistics of signal and background proxy samples used to train the partially reconstructed MVA. The number in brackets denotes the full size of the sample before aligning to the smaller proxy.

Classifier		Signal events	Background events
$B^+ \to K^+e^+e^-$	Run 1	18498 (limiting)	18498 (35568)
	Run 2p1	26291 (limiting)	26291 (96348)
	Run 2p2	61975 (63019)	61975 (limiting)
$B^0 \to K^{*0}e^+e^-$	Run 1	26824 (42800)	26824 (limiting)
	Run 2p1	62512 (limiting)	62512 (73913)
	Run 2p2	27027 (66997)	27027 (limiting)

Table 4.18.: Summary of the input variables used to train the partially reconstructed MVA for B^+ modes.

Particle	Variables
B^+	p_{T}, χ^2_{IP}, DIRA, $\chi^2_{\mathrm{vtx,iso}}$, $m_{\mathrm{vtx,iso}}$
$\ell^+\ell^-$	χ^2_{IP}, χ^2_{FD}, DIRA
K^+	p_{T}
e	$min(\chi^2_{\mathrm{IP}}(\ell^+), \chi^2_{\mathrm{IP}}(\ell^-))$, $max(\mathrm{MULT}_{\mathrm{cone,iso}}(\ell^+), \mathrm{MULT}_{\mathrm{cone,iso}}(\ell^-))$, $min, max(\mathrm{PT}_{\mathrm{cone,iso}}(\ell^+), \mathrm{PT}_{\mathrm{cone,iso}}(\ell^-))$, $min(\mathrm{IT}_{\mathrm{cone,iso}}(\ell^+), \mathrm{IT}_{\mathrm{cone,iso}}(\ell^-))$

Table 4.19.: Summary of the input variables used to train the partially reconstructed MVA for B^0 modes.

Particle	Variables
B^0	χ^2_{IP}, χ^2_{vtx}/ndf, χ^2_{DTF}/ndf, DIRA, $\chi^2_{\mathrm{vtx,iso}}$
K^{*0}	χ^2_{IP}, χ^2_{vtx}/ndf, DIRA
$\ell^+\ell^-$	χ^2_{IP}, χ^2_{FD}, DIRA
K^+, π^-	$min(p_{\mathrm{T}}(K^+), p_{\mathrm{T}}(\pi^-))$
e	$min(\mathrm{MULT}_{\mathrm{cone,iso}}(e^+), \mathrm{MULT}_{\mathrm{cone,iso}}(e^-))$

events. No indicators for overtraining can be seen in the distributions of training and testing samples. The data vs. MC comparison for the response of the classifier is shown in Fig. 4.25.

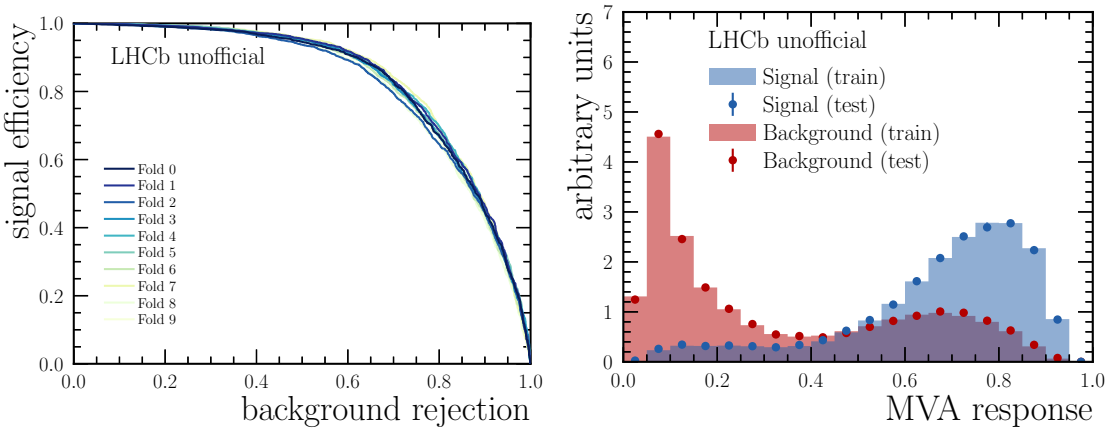

Figure 4.24.: Cross-checks for the Run 2p2 electron partially reconstructed MVA: ROC curves of all 10 folds (left). Overtraining plot averaged over all folds (right).

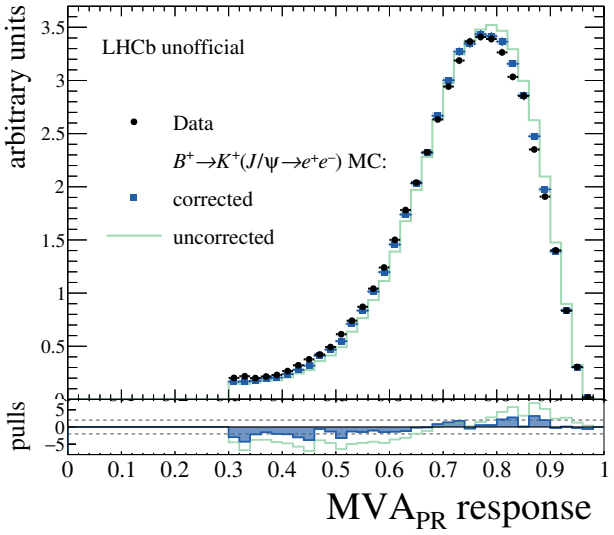

Figure 4.25.: Comparison of data (black) vs. MC with (blue) and without (green) corrections applied on Run 2p2 electron samples for the partially reconstructed MVA response. The full selection is applied with loosened MVA cuts.

Cut Optimisation

In order to increase the sensitivity to the rare modes, the cut values on the response of the MVAs are optimised. Similarly to the MVA training, the optimisation is performed separately for B^+ and B^0 channel, muon and electron modes and in the three run periods. Additionally, separate cut threshold are determined for the low and central q^2 regions. For muon modes, a one-dimensional optimisation of the combinatorial MVA is done, while for electron modes a two-dimensional optimisation together with the partially

reconstructed MVA is performed, which yields a total of 24 separately determined cut thresholds. Technically, this is done by maximising the Figure Of Merit (FOM) defined as $P = {}^{N_{\text{sig}}}/\sqrt{N_{\text{sig}} + N_{\text{bkg}}}$, where N_{sig} and N_{bkg} are the expected number of signal and background events at a given MVA cut value.

The value of N_{sig} is not taken directly from fits to data to avoid the introduction of a bias, instead it is calculated using MC samples: First, the number of events (N_{MC}) after full selection is calculated in a signal window of $m(B) \pm 50\,\text{MeV}/c^2$ for muon and 5150–5350 MeV/c^2 for electron channels. Next, the number of expected signal events N_{exp} is determined by normalising to the J/ψ resonant mode with a similar approach as used in Eq. 4.11:

$$N_{\text{exp}} = \epsilon_{\text{sig}} \cdot \frac{\mathcal{B}_{\text{sig}}}{\mathcal{B}_{\text{norm}}} \cdot \frac{\mathcal{N}_{\text{norm}}}{\epsilon_{\text{norm}}}\,, \tag{4.13}$$

where $\mathcal{B}_{\text{sig}}$ and ϵ_{sig} are the branching ratio and efficiency of the rare mode of interest, while $\mathcal{B}_{\text{norm}}$, ϵ_{norm} and $\mathcal{N}_{\text{norm}}$ are the branching ratio, efficiency and yield of the respective J/ψ mode. The branching ratios are taken from PDG [22] and the efficiencies and yields are determined on simulated samples (see Sec. 4.5) and with invariant mass fits (see Sec. 4.6), respectively. Using this, a factor $\alpha_{\text{sig}} := {}^{N_{\text{exp}}}/_{N_{\text{MC}}}$ is determined by which the number of selected events with a given MVA cut is scaled

$$N_{\text{sig}} = \alpha_{\text{sig}} \cdot N_{\text{MC}|\text{cut(MVA)}}\,. \tag{4.14}$$

The value of N_{bkg} is determined by fitting to the invariant B mass in the upper and lower sideband regions. The signal windows (defined above) are blinded in these fits. N_{bkg} is calculated by interpolating the fitted background components to the signal window.

Finally, the MVA response is scanned in equidistant steps and for each N_{sig} and N_{bkg} are calculated. Subsequently, the maximum of the FOM yields the cut threshold that are listed in Tab. 4.20. The cut values for the electron modes were found to be compatible across the three run periods, which is why they were aligned. For the J/ψ and $\psi(2S)$ resonant channels looser cut thresholds are chosen in order to account for the much lower background to signal ratio in these modes.

HOP Selection

The last selection step applies a requirement on the so-called 'HOP' mass [103], which further suppresses the partially reconstructed backgrounds in electron modes by exploiting the kinematics of the decay: Under an ideal reconstruction, the vectorial sum of the momenta of all final state particles is parallel to the direction between the PV and the decay vertex of the B meson. However, bremsstrahlung (or not reconstructed particles) can introduce an orthogonal component for signal events, so that the ratio

$$\alpha_{\text{HOP}} := \frac{p_{\text{T}}(K^+)}{p_{\text{T}}(e^+e^-)}\,,$$

Table 4.20.: Summary of the combinatorial ($\mathrm{MVA_{Comb}}$) and partially reconstructed ($\mathrm{MVA_{PR}}$) MVA cut values.

Channel	q^2 region	Run 1	Run 2p1	Run 2p2
$B^+ \to K^+\mu^+\mu^-$	low	$\mathrm{MVA_{Comb}} > 0.70$	$\mathrm{MVA_{Comb}} > 0.85$	$\mathrm{MVA_{Comb}} > 0.85$
	central	$\mathrm{MVA_{Comb}} > 0.70$	$\mathrm{MVA_{Comb}} > 0.80$	$\mathrm{MVA_{Comb}} > 0.80$
$B^+ \to K^+e^+e^-$	low	$\mathrm{MVA_{Comb}} > 0.90$ && $\mathrm{MVA_{PR}} > 0.40$		
	central	$\mathrm{MVA_{Comb}} > 0.90$ && $\mathrm{MVA_{PR}} > 0.40$		
$B^0 \to K^{*0}\mu^+\mu^-$	low	$\mathrm{MVA_{Comb}} > 0.29$	$\mathrm{MVA_{Comb}} > 0.54$	$\mathrm{MVA_{Comb}} > 0.55$
	central	$\mathrm{MVA_{Comb}} > 0.63$	$\mathrm{MVA_{Comb}} > 0.77$	$\mathrm{MVA_{Comb}} > 0.64$
$B^0 \to K^{*0}e^+e^-$	low	$\mathrm{MVA_{Comb}} > 0.50$ && $\mathrm{MVA_{PR}} > 0.50$		
	central	$\mathrm{MVA_{Comb}} > 0.90$ && $\mathrm{MVA_{PR}} > 0.40$		

where the $B^+ \to K^+e^+e^-$ mode is taken as an example, deviates from unity. Since, bremsstrahlung photons are emitted in flight direction of the electron, it is possible to correct the dielectron momentum with

$$\vec{p}_{\mathrm{corr}}(e^+e^-) = \alpha_{\mathrm{HOP}} \cdot \vec{p}(e^+e^-)\,.$$

For partially reconstructed backgrounds the missing particle(s) most likely are not emitted in the same direction as the electrons, so that this correction is biased. The corrected momentum is used to recalculate the invariant mass of the B meson $m_{\mathrm{HOP}}(B)$, which is shown for signal and partially reconstructed background MC in Fig. 4.26.

The cut value on the $m_{\mathrm{HOP}}(B)$ variable is determined with an analogous optimisation procedure as used for the MVA selections. The values are found to be compatible across the three run periods. For both B^+ and B^0 electron modes, the requirement applied on the HOP mass is $m_{\mathrm{HOP}}(B) > 4800\,\mathrm{MeV}/c^2$ for low q^2 and $m_{\mathrm{HOP}}(B) > 4700\,\mathrm{MeV}/c^2$ for central q^2.

4.4 Calibration of Simulated Samples

As mentioned before, simulated samples are used for various exercises throughout the analysis, most notably the efficiency calculation for the LFU measurements. To calculate reliable results, the simulated samples must model the real data well. However, as detailed in Sec. 3.3.3 the LHCb simulation is known to deviate from data in several aspects, consequently a number of data-driven corrections are applied. The following list gives all simulated quantities that are corrected for:

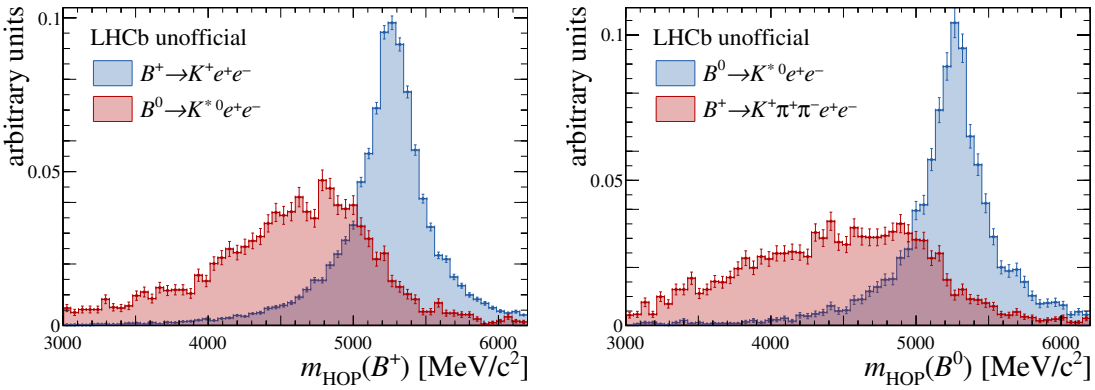

Figure 4.26.: Distribution of $m_{\mathrm{HOP}}(B)$ on 2018 simulated signal and partially reconstructed background samples. For the B^+ mode (left) $B^0 \to K^{*0} e^+ e^-$ MC is used as background while for B^0 (right) $B^+ \to K^+ \pi^+ \pi^- e^+ e^-$ is taken.

1. The efficiency of the PID selection;
2. The electron track reconstruction efficiency;
3. The kinematic of the B meson at generator-level and event multiplicity;
4. The response of both L0 and HLT trigger;
5. The reconstruction of the signal candidate;
6. The efficiency of the q^2 selection.

In each case, the correction is determined using calibration samples from data and applied to MC with per-event weights. This is done separately for all years of data-taking. In order to account for possible correlations between the individual corrections, they are calculated sequentially with all previous corrections applied.

There are dedicated calibration samples for the PID and electron track reconstruction corrections, which are described in the next sections. For all other correction steps, data from the J/ψ resonant channels is used. Each of these corrections is calculated on both B^+ and B^0 modes yielding two sets of weights, that are referred as $w(B^+)$ and $w(B^0)$, respectively. Subsequently, the full corrections from both chains are applied to all rare and resonant mode MC samples, since they proved to be portable between B^+ and B^0 modes (which will be shown in Sec. 4.8). Thus, it is possible to switch between the two correction chains, to cross-check the results. In fact, the final LFU measurements of B^+ modes use efficiencies based on $w(B^0)$ corrections and vice versa, in order to minimise correlations of the efficiency corrections and the fitted J/ψ resonant mode.

In the following, this section will describe the various MC corrections in the same order as listed above.

4.4.1 Particle Identification Response

For simulated samples, the PID criteria of both stripping (see Tab. 4.3) and offline selection (see Tab. 4.6) are calibrated to determine the correct selection efficiency. The strategy for this calibration is to substitute the PID selection on MC with weights, that contain the efficiency of these cuts measured on calibration data samples. These calibration samples are recorded for each particle species during the general data-taking and are described in Ref. [106]. In general, these samples are taken from decay processes that allow the particle type of interest to be identified without directly using PID information. This way, the sample is unbiased from the PID response of the studied particle allowing the efficiency of a given PID requirement to be measured. Various selection criteria ensure a good signal purity of the calibration samples and the residual background contamination is removed statistically using the *s$\mathcal{P}$lot* method [107], which is based on an invariant mass fit to the full sample.

The calibration is performed using the `PIDCalib` package [108], which uses a `tag`&`probe` method similar to the tracking reconstruction (see Sec. 3.2.1) to determine the efficiencies. The PID efficiency is defined as the ratio of the number of events in the calibration sample after and before applying the PID requirements of interest. Since the efficiency is known to depend on both kinematics of the particle and multiplicity of the event, this ratio is evaluated in bins of the momentum and pseudorapidity of the particle as well as of `nTracks`. Thus, efficiency maps are obtained, which are used to apply per-event weights to the simulated samples. This approach is used in the presented analysis for kaon, pion and muon candidates for which representative efficiency maps ($\epsilon_{\text{PID}}(K,\,\pi,\,\mu)$) are shown in Fig. 4.27.

This strategy is used in a large number of analyses at LHCb and known to successfully calibrate the PID response for kaons, pions and muons. However, for electrons this *sPlot*-based approach is not applicable for the following reasons: In contrast to muons and hadrons, the calibration samples of electrons have comparably lower statistics and larger background contamination, which can lead to non-physical values for the *s$\mathcal{W}$eight*s in low populated areas in the kinematic phase-space. Furthermore, the *s$\mathcal{P}$lot* method requires the variable used for fitting (invariant mass of the B-meson) to be uncorrelated to the variables which are studied applying the *s$\mathcal{W}$eight*s. However, for electrons there are non-negligible correlations between the B-meson mass and the track momentum of the electron for tracks having assigned bremsstrahlung photons.

Thus, a different approach, called `fit`&`count`, is chosen to calibrate the electron PID response. Here, for each bin of p, η and `nTracks` a simultaneous fit to the events in the calibration sample which passed and the events which failed the PID requirements of interest is performed. The mass fits follow a similar strategy as used for the nominal electron fits that are introduced in Sec. 4.6. This allows the efficiency of the applied PID selection to be calculated as

$$\epsilon_{\text{PID}}(e) = \frac{N_{\text{passed}}}{N_{\text{passed}} + N_{\text{failed}}}\,. \tag{4.15}$$

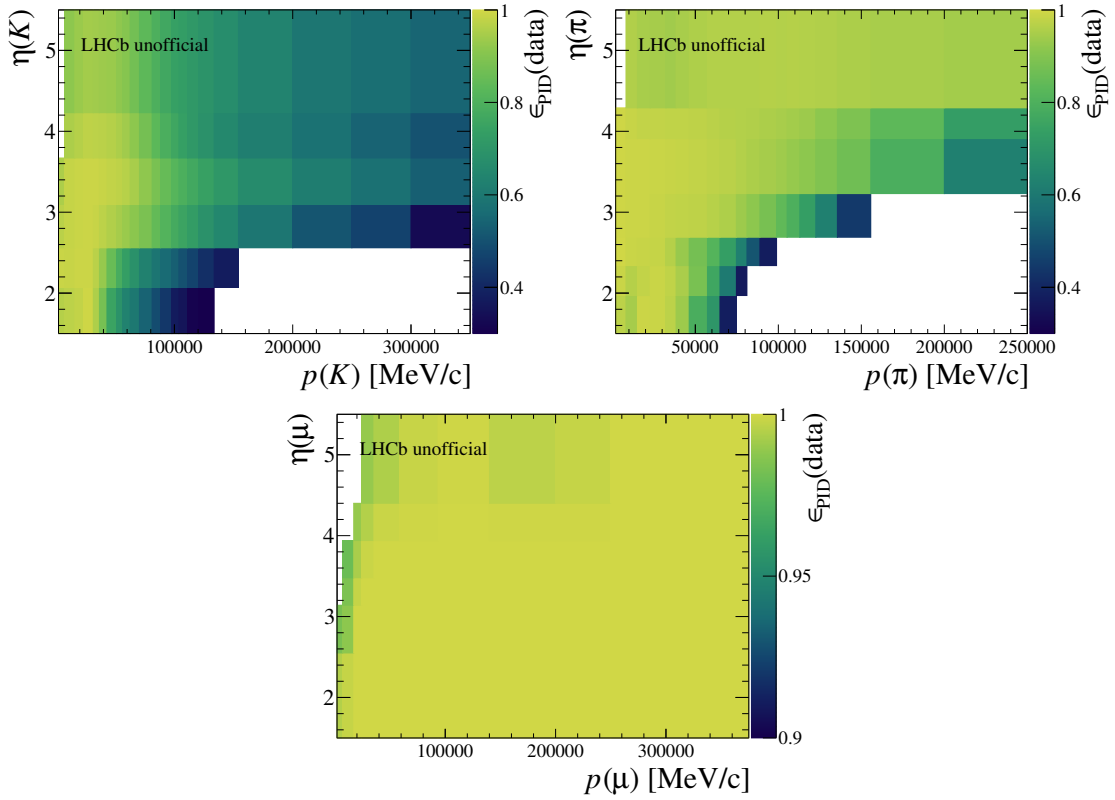

Figure 4.27.: PID efficiency maps from 2018 calibration samples for kaons (left), pions (right) and muons (bottom). The maps for the first `nTracks` bin are shown. Phase-space regions that do not contain any events are shown in white.

Another difference between electrons and other particle species lies in the correlation of the PID efficiency between two signal tracks in an event. The PID response of the two signal muons or the hadrons are uncorrelated, which justifies the separate determination of efficiencies as described above. In contrast, studies based on simulated samples have shown, that for two electrons, having a small relative distance in the ECAL, the PID efficiencies can correlate due to overlapping clusters [101]. This was already mentioned in the selection strategy for electrons in Sec. 4.3.2 and was reason to require a minimal distance in the ECAL between the electron tracks. To further reduce the impact of this potential source of bias, the PID correction follows a different strategy for electrons. Instead of fully substituting the PID response on MC with the efficiency weights, the electron PID selection is applied on the simulated PID response, which enables to account for possible correlations between the tracks. In order to correct the MC response, the single electron PID efficiency is determined with Eq. 4.15 on both data and MC samples and the data over MC ratio is applied as a per-event weight

$$w_{\mathrm{PID}}(e) = \frac{\epsilon^{\mathrm{data}}_{\mathrm{PID}}(e)}{\epsilon^{\mathrm{MC}}_{\mathrm{PID}}(e)}\,. \tag{4.16}$$

To ensure reliable results, the used data and MC samples describe the same decay channel and are aligned in selection. This so-called 'reweighting' approach is also used for most of the remaining MC correction steps in this analysis.

As discussed in Sec. 3.4.2, the PID efficiency for electrons depends on whether the candidate has bremsstrahlung assigned, which is the reason why the calibration sample is further divided (prior to the fitting) into candidates that have zero or at least one photon assigned. For each bremsstrahlung category, one electron PID correction map is shown in Fig. 4.28.

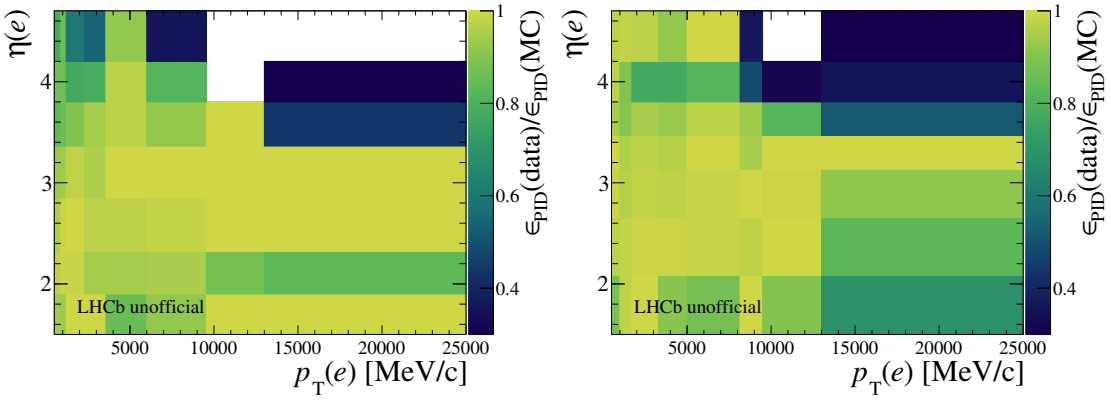

Figure 4.28.: PID efficiency maps from 2018 calibration samples for electrons for candidates with zero (left) and at least one (right) bremsstrahlung photon. The maps for the first `nTracks` bin are shown. Phase-space regions that do not contain any events are shown in white.

In summary, the total PID efficiency on MC for electron modes is calculated by applying the full electron PID selection (Sel(PIDe)) and multiplying by the product of the two data over MC single electron weights and the single data efficiency weights of the hadrons. The following formulas show the full PID efficiency computation for each signal final state:

$$
\begin{aligned}
K^+\mu^+\mu^- &: \quad \epsilon_{\text{PID}} = \epsilon_{\text{PID}}(K^+) \cdot \epsilon_{\text{PID}}(\mu^+) \cdot \epsilon_{\text{PID}}(\mu^-)\,, \\
K^{*0}\mu^+\mu^- &: \quad \epsilon_{\text{PID}} = \epsilon_{\text{PID}}(K^+) \cdot \epsilon_{\text{PID}}(\pi^-) \cdot \epsilon_{\text{PID}}(\mu^+) \cdot \epsilon_{\text{PID}}(\mu^-)\,, \\
K^+e^+e^- &: \quad \epsilon_{\text{PID}} = \epsilon_{\text{PID}}(K^+) \cdot w_{\text{PID}}(e^+) \cdot w_{\text{PID}}(e^-) \mid \text{Sel(PID}e\text{)}\,, \\
K^{*0}e^+e^- &: \quad \epsilon_{\text{PID}} = \epsilon_{\text{PID}}(K^+) \cdot \epsilon_{\text{PID}}(\pi^-) \cdot w_{\text{PID}}(e^+) \cdot w_{\text{PID}}(e^-) \mid \text{Sel(PID}e\text{)}\,.
\end{aligned}
$$

4.4.2 Electron Track Reconstruction

The track reconstruction efficiency of electron candidates is calibrated by the LHCb collaboration via the `tag&probe` method and is documented in Ref. [109]. In this measurement, $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decays are exploited, where the final state kaon and one electron are used as `tag` and have a rather stringent selection applied (good track quality and PID response). For the other electron (`probe`), only track information coming from the VELO detector is required. Subsequently, the efficiency for the `probe` to

form a long track by combining the VELO segment with information from the remaining tracking system is measured on both data and simulated samples. The per-event weights used to correct the simulation are calculated as

$$w_{\mathrm{TRK}} = \frac{\epsilon_{\mathrm{TRK}}(\mathrm{Data})}{\epsilon_{\mathrm{TRK}}(\mathrm{MC})}\,, \tag{4.17}$$

where the phase-space is parameterised in bins of p_{T}, η and the azimuthal angle ϕ of the `probe` track. Figure 4.29 shows correction maps for 2017 samples that contain the data over MC track reconstruction efficiency ratios. The final tracking correction is computed as the product of the two single electron track correction weights.

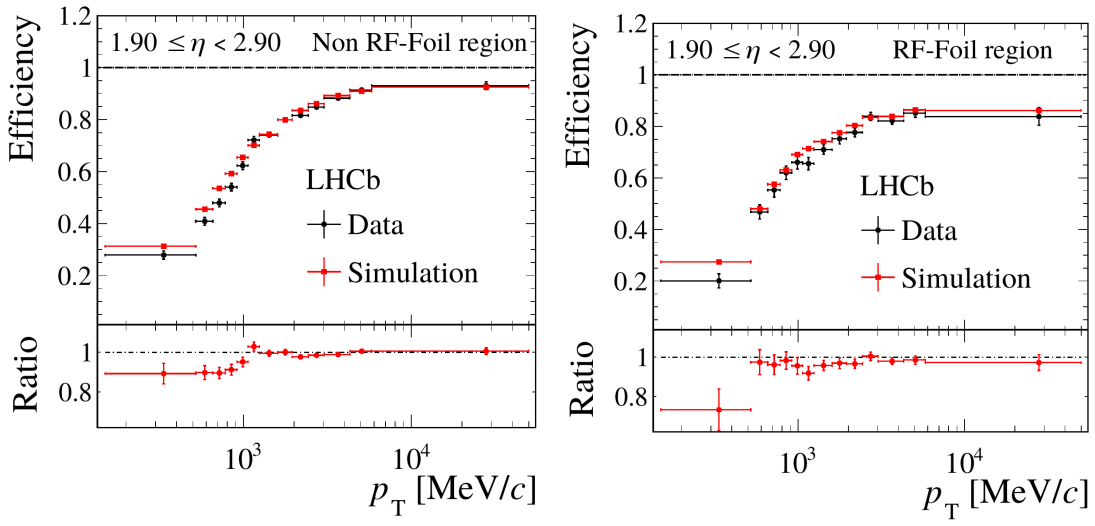

Figure 4.29.: Electron track reconstruction efficiency measured on data (black) and simulation (red). The maps are from 2017 samples in the low η region for two regions of ϕ: tracks parallel (left) and not parallel (right) to the RF-foil. Figures taken from [109].

For muons, the tracking efficiency is known to be well simulated, while for hadrons any data and MC differences cancel even in single ratios such as $r^{J/\psi}$. Thus, this kind of correction is not needed for other particles species than electrons.

4.4.3 Kinematic and Multiplicity

The simulated production kinematics of the B-meson as well as the multiplicity of the underlying event are known to show deviations from data. The reason for the kinematic discrepancies are an imperfect description of the pp collisions in `Pythia`, while the multiplicity suffers from limitations in the simulated material budget (especially of the first muon station (M1)) and in the production of low momentum particles in interactions with detector material.

To address these shortcomings, weights are assigned to the simulated samples, so that the reweighted distributions match those on data. The kinematic weights are parameterised in p, p_{T} and η of the B-meson while `nTracks` is chosen as proxy for multiplicity.

The weights are calculated from fully selected data and simulated samples from the J/ψ resonant muon modes triggered by the inclusive `L0Muon` line. This mode is used to benefit from the high purity and statistics provided by the muon data samples. Residual backgrounds are statistically removed from the data samples using the *sPlot* method. The correction extracted from the `L0Muon` sample is applied to simulated samples of the signal and resonant modes in both trigger categories. Therefore, the correction needs to be portable between muon and electron modes as well as between trigger categories. In principle, there are no expected differences between the generated kinematics and multiplicity between muon and electron decays, which is demonstrated by Fig. 4.30 showing the proxies used for the reweighting on generator-level samples. However, since fully reconstructed samples are used to calculate the weights, because there is no data proxy at generator-level, it is necessary to ensure that no other differences between data and MC bias the weight calculation. The selected `L0Muon` sample benefits from well modelled PID and trigger responses on simulation. To further ensure that only kinematic and multiplicity differences are corrected for, a first set of L0 and HLT trigger corrections (see following sections) is applied to the used MC samples. Additionally, the data sample is selected by the TCK that was used for MC production to align the trigger thresholds.

The actual calculation of the weights $w_{\text{Kin\&Mult}}$ is performed with the `GBReweighter` from the `HEP-ml` software library [110]. Here, a BDT based approach is used to simultaneously reweight all kinematic and multiplicity variables. This offers a highly tunable tool, that fully accounts for correlations between the variables. Again the `k-folding` method is applied to train and apply the BDT in order to avoid biases. To validate that the training of the BDT was successful, the agreement between data and simulated samples is checked after applying the weights. This is shown for the `L0Muon` mode which is used for the training, in Fig. 4.31 and for the L0I electron mode in Fig. 4.32[4]. The excellent data and MC agreement in the L0I electron mode demonstrates the good portability of the correction, since at no point information on kinematics and multiplicity from the electron modes or the L0I trigger category enters into the calculation of the weights. The good agreement in other variables, such as the kinematics of final state or intermediate particles, which is demonstrated in App. C, further underlines the high quality of this correction.

It is worth to mention that `nTracks` is chosen to correct the event multiplicity, because this variable is also used to parameterise the correction of the PID response. However, the reweighting of `nTracks` does not improve the agreement with data on all other simulated multiplicity variables. Especially the number of hits in the SPD (`nSPDHits`), which is much more strongly affected by the mismodelled material budget of the M1, still shows large discrepancies (see App. C). Since the stripping lines apply a requirement on `nSPDHits` (see Tab. 4.3), the data and MC difference of this variable poses a potential problem. However, as shown in Fig. 4.30 the generated event multiplicity is nearly identical between muon and electron modes, so that any bias on the efficiency of the `nSPDHits` cut cancels in the double (and even single) ratios. To further ensure

[4]The distributions of corrected MC also have the trigger corrections from the following sections applied.

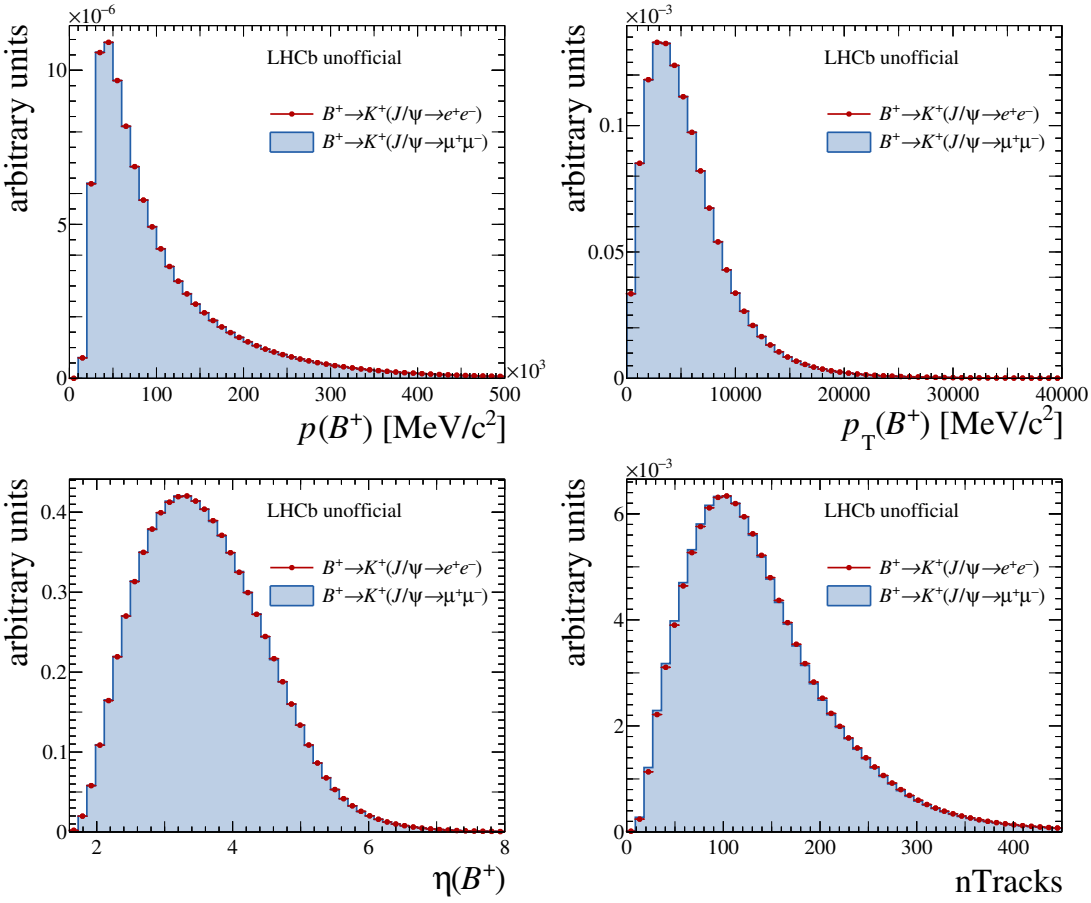

Figure 4.30.: Distributions of the variables used for the kinematic and multiplicity reweighting on generator-level MC of $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ (blue) and $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ (red) in Run 1.

the robustness of the efficiency estimation, a systematic uncertainty is assigned for the correction of the event multiplicity (see Sec. 4.7.1).

4.4.4 L0 Trigger

The simulated trigger response is known to deviate from data, because of misalignments in the TCKs between data and simulation (see Sec. 3.3.3) and other imperfectly modelled quantities like detector occupancy. In order to correct the efficiencies of the required L0 trigger lines on simulation, they need to be extracted from data. To this end, the J/ψ resonant modes $B^+ \to K^+ J/\psi$ and $B^0 \to K^{*0} J/\psi$ selected on data are used.he difference in trigger selection for muons and electrons requires to exploit both $J/\psi \to \mu^+\mu^-$ and $J/\psi \to e^+e^-$ channels. The samples have the full selection applied except for the L0 trigger requirements. To further suppress backgrounds on the data sample, the J/ψ mass constrained B mass system is selected in a 45 MeV wide window around the B-meson mass.

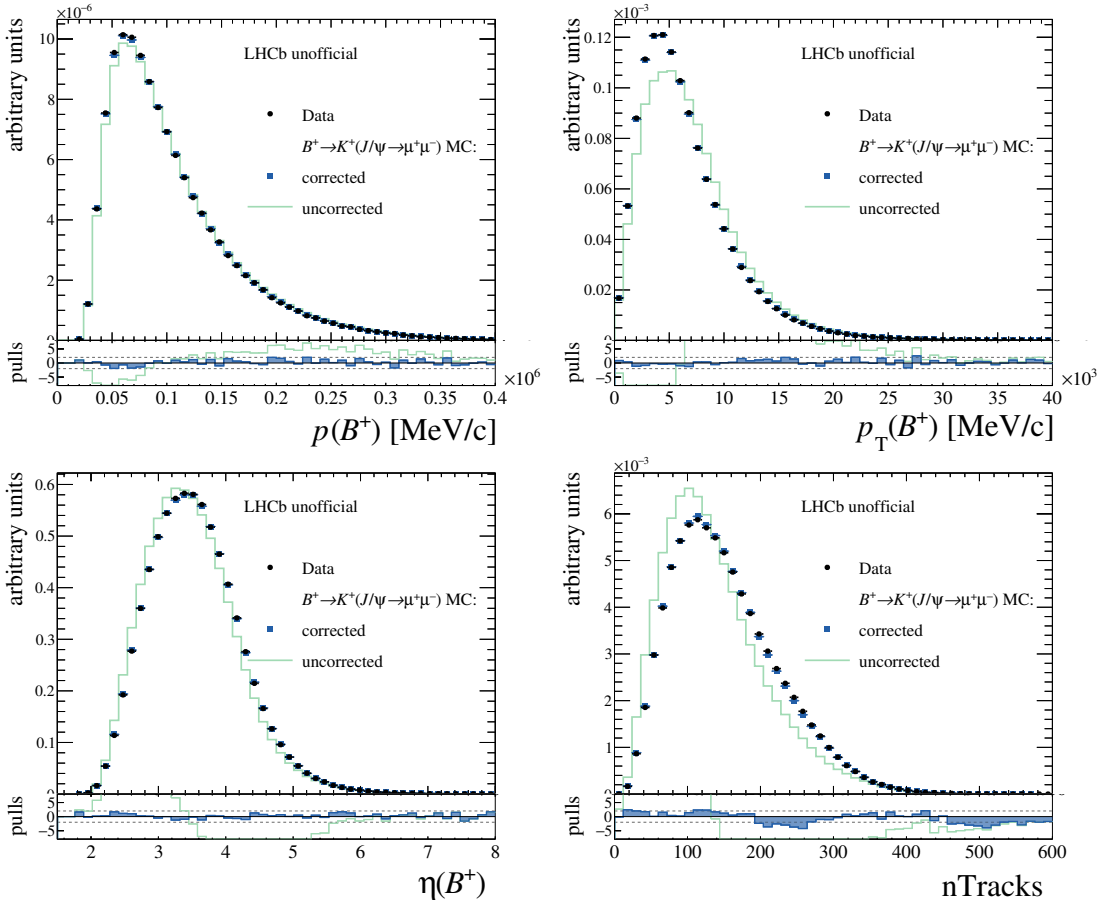

Figure 4.31.: Distributions of the variables used for the kinematic and multiplicity reweighting on data (black) and on MC with (blue) and without (green) applying the corrections. The samples from the $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ channel in Run 1 `L0Muon` (inclusive).

A difficulty arises from the fact that the trigger efficiencies are not directly accessible in data, since all of the events that did not pass the trigger requirements were discarded. Thus, the so-called `TISTOS` method [111] is used to calibrate the trigger response: A `tag`&`probe` approach is used to calculate the efficiency of a trigger line on a `tag` sample that is gathered by another, independent, trigger line. The efficiency of the studied L0 trigger line ($\epsilon_{\texttt{L0}}$) is then given by the fraction of all events in the `tag` sample ($N_{\texttt{tag}}$) that also have fired the `probe` trigger line ($N_{\texttt{tag\&probe}}$)

$$\epsilon_{\texttt{L0}} = \frac{N_{\texttt{tag\&probe}}}{N_{\texttt{tag}}}\,. \tag{4.18}$$

This exercise is performed on both data and simulated samples having applied the same selection. The corrections to simulation are calculated via the data over MC ratio approach

$$w_{\text{L0}} = \frac{\epsilon_{\text{L0}}^{\text{data}}}{\epsilon_{\text{L0}}^{\text{MC}}}\,. \tag{4.19}$$

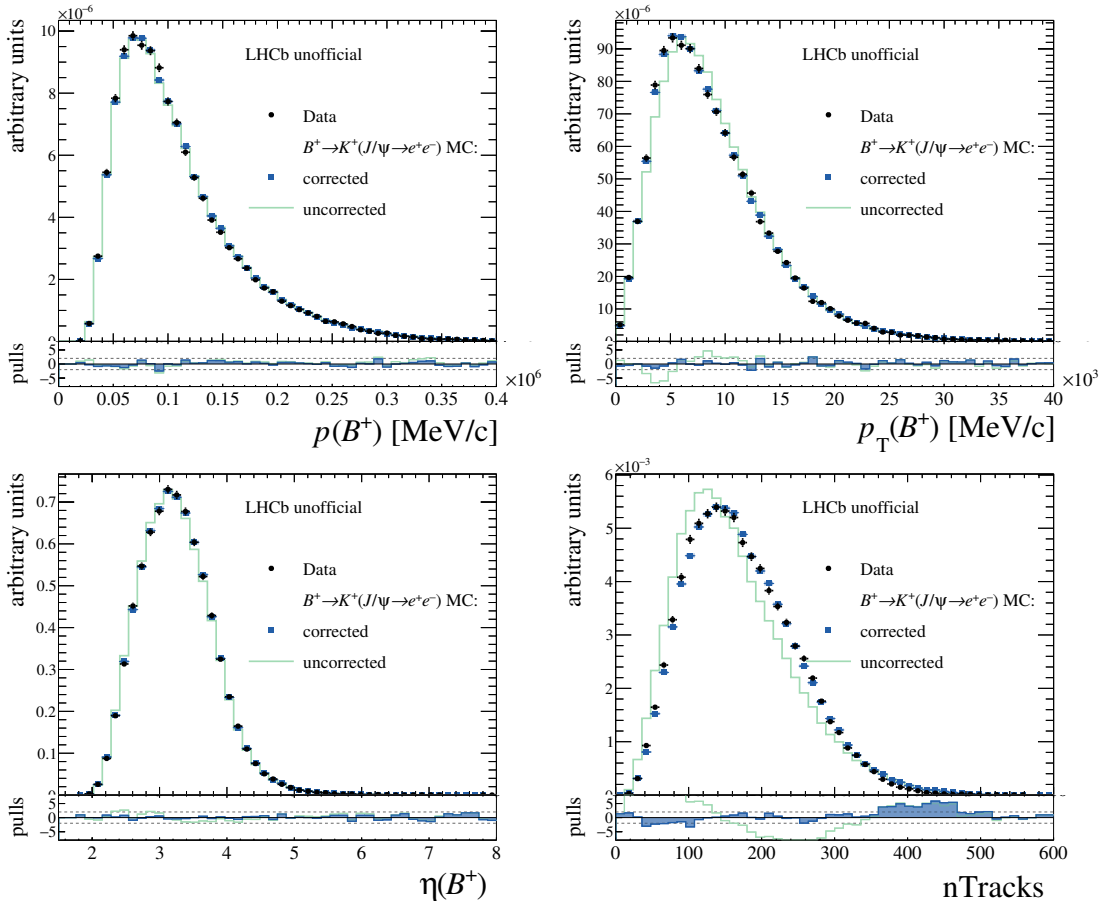

Figure 4.32.: Distributions of the variables used for the kinematic and multiplicity reweighting on data (black) and on MC with (blue) and without (green) applying the corrections. The samples from the $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ channel in Run 1 L0I.

Similar to previous MC corrections, the L0 efficiency exhibits a significant dependence on certain kinematic variables, especially those which are used to form the trigger decision. These variables vary between the different L0 lines exploited in this analysis and are used to parameterise the trigger correction. In the following subsections, the trigger corrections for each used L0 line are described.

Correction of `L0Muon`

The `L0Muon` line is fired by muon candidates that exceed a certain threshold on the transverse momentum of the track $p_T(\mu)$ (see Sec. 3.3.1). Thus, the calibration of this trigger line is perfomed in bins of $p_T(\mu)$ and also in three different regions of the pseudorapidity $\eta(\mu)$, which also shows variations of the efficiency. Events from the L0 electron and hadron lines, that were triggered independent of the signal (`TIS`), are used as the `tag` sample. To verify that this `tag` sample is indeed independent of the probed `L0Muon`

line and yields unbiased results, two other `tag` samples from different L0 lines are also used to extract the efficiencies and have shown compatible results. This cross-check is performed for the calibration of all used L0 lines and is documented in Ref. [93].

The dependency of the efficiency on the variable used to form the trigger decision, known as the 'turn-on' curve, can be seen in Fig. 4.33, which shows the trigger efficiencies calculated with the `TISTOS` method for both data and simulated samples as well as the ratio of the two.

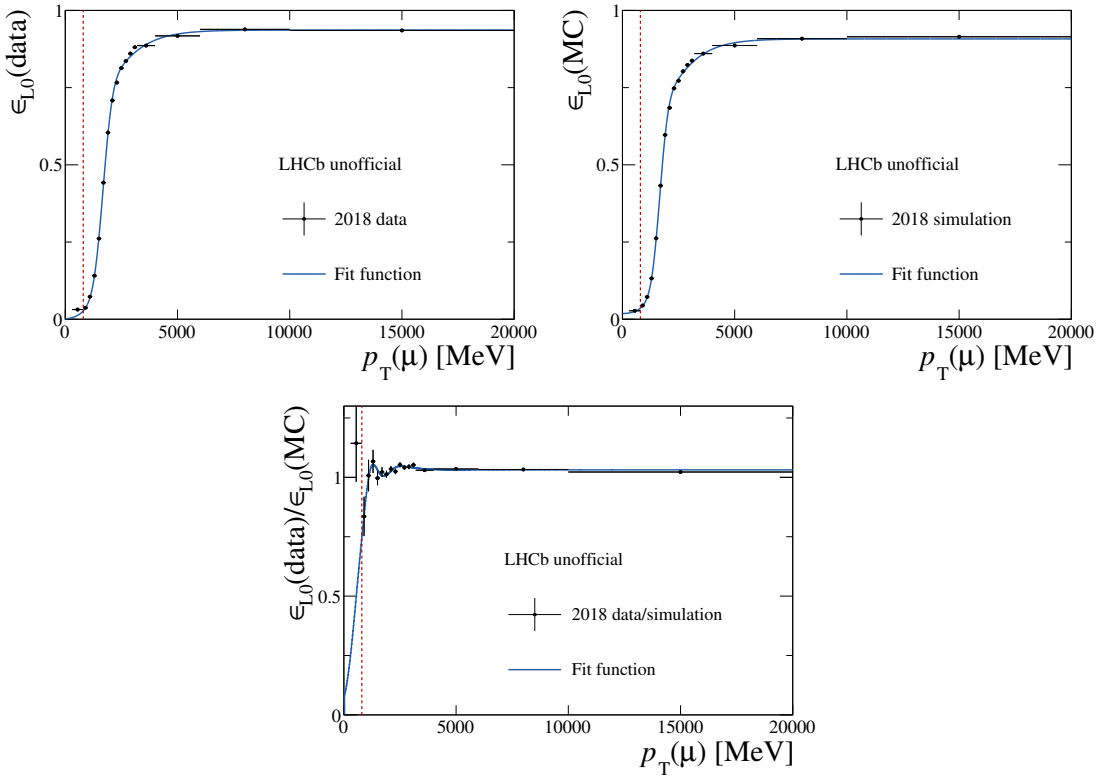

Figure 4.33.: Trigger turn-on curve plots showing the efficiency of the `L0Muon` trigger on data (left) and MC (right). The bottom plot shows the data over MC efficiency ratio. The vertical line indicates the trigger threshold value. Representatively, efficiencies in the low $\eta(\mu)$ bin from 2018 data and simulated samples of the B^+ mode are used for these plots.

To reduce the bias from the chosen binning scheme, the turn-on curve is parameterised with a function, which is used to extract the measured efficiencies. A combination of an error function and a Gompertz function is used to perform this task

$$f(x) = f_1 \cdot \left(1 + \mathrm{erf}\left(\frac{x - t_0}{\sqrt{2}\cdot\sigma_1}\right)\right) + f_2 \cdot s^{\exp\left(\frac{-(x-t_0)}{\sqrt{2}\cdot\sigma_2}\right)} + a\,, \tag{4.20}$$

where $f_{1,2}$ are the relative fractions between the two functions, $\sigma_{1,2}$ the resolution parameters, t_0 the trigger turn-on point, s the skewness of the Gompertz function and a

the global offset. The motivation for the error function is in its ability to model the typical shape of turn-on curves: A plateau region on the left where the $p_{\mathrm{T}}(\mu)$ becomes so small that the efficiency goes to zero and another plateau at high values of $p_{\mathrm{T}}(\mu)$, where the saturation efficiency is reached; both linked with a steep ascent that is dominated by the resolution. Since the error function possesses point symmetry around t_0, it is not perfectly suited to describe the slightly softer entry into the high plateau region compared to the left plateau as can be observed in Fig. 4.33. Thus, the Gompertz function is added, which can model this asymmetric behaviour with its skewness parameter s.

Correction of `L0Electron`

The strategy for the calibration of `L0Electron` is very similar to `L0Muon`. Here, the amount of deposited transverse energy in the ECAL by an electron candidate $E_T^{\mathrm{ECAL}}(e)$ is used to form the trigger decision. The trigger efficiency is also dependent on the position of the cluster in the ECAL, since there are different granularities in the three regions of the calorimeter. Thus, the parameterisation of the calibration is done in bins of $E_T^{\mathrm{ECAL}}(e)$ separately for the different ECAL regions. The `tag` sample is collected with the `L0Muon` and `L0Hadron` trigger lines from `TIS` events. The trigger turn-on curves for both data and simulated samples are fitted again with the function in Eq. 4.20 and are displayed in Fig. 4.34 together with the data over MC ratio distribution. Especially for the `L0Electron` calibration, the fit to the turn-on curve can not describe the binned efficiencies perfectly, because of large fluctuations in the high plateau area. Any bias arising from this effect is covered by a systematic uncertainty on the binning and parameterisation of the L0 trigger calibration (see Sec. 4.7.1).

Correction of L0I

Lastly, the calibration of the L0I trigger line is discussed. This category consists in large parts of events that are triggered by final state particles originating from the opposite b-quark of the signal $b\bar{b}$ creation process. Because of the correlation between the b and $\bar{b}$ quarks, the transverse momentum of the signal B-meson is used to parameterise this efficiency. In addition, the multiplicity of the event has a significant impact on the L0I line, which is why the parameterisation is performed in bins of $p_{\mathrm{T}}(B)$ and `nTracks`. Both L0L and `L0Hadron` lines fired on signal final state particles are used to gather the `tag` sample. Since the L0I line is expected to be independent of the signal decay, the efficiency should be equal for the dimuon and dielectron channels. This has been verified by extracting efficiencies with the `TISTOS` method from J/ψ resonant muon and electron decays and proving their compatibility. This study can be reviewed in Ref. [93]. Thus, the calibration of the L0I line is extracted on the dimuon channel to benefit from its higher statistics and purity, and subsequently used to correct simulated samples of both muon and electron channels.

Figure 4.35 shows the turn-on curves of the L0I category, which are fitted with a single Gompertz function, and the resulting data over simulation ratio.

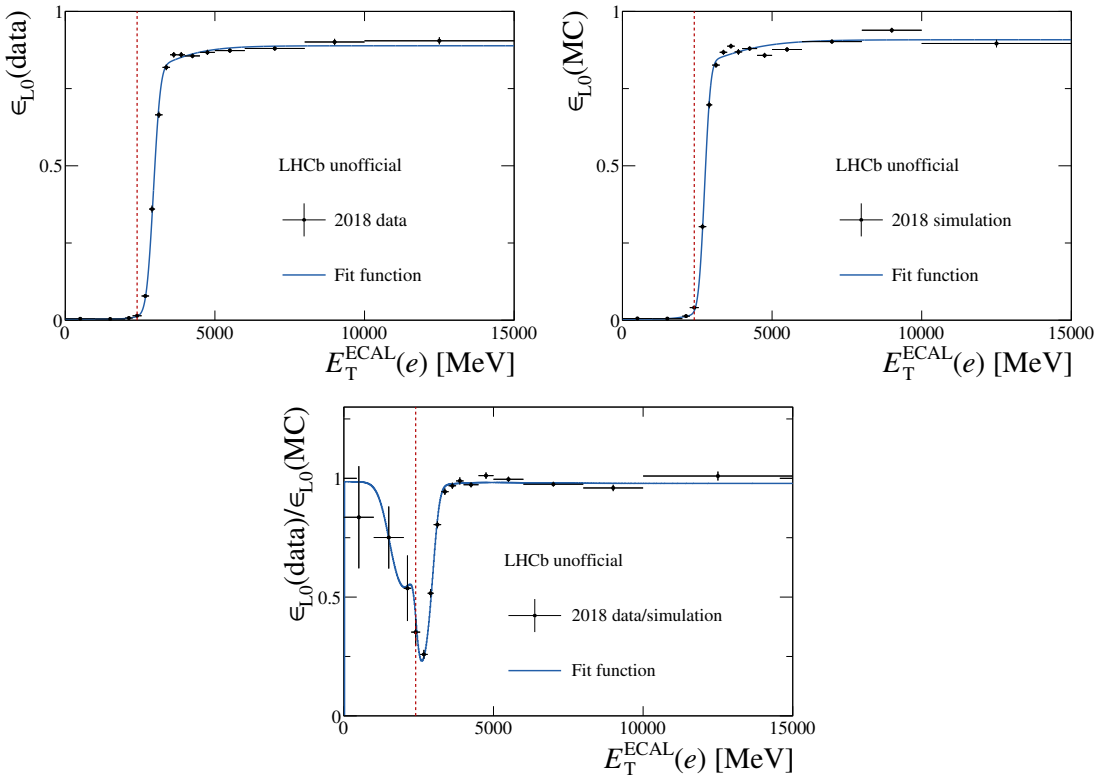

Figure 4.34.: Trigger turn-on curve plots showing the efficiency of the `L0Electron` trigger on data (left) and MC (right). The bottom plot shows the data over MC efficiency ratio. The vertical line indicates the trigger threshold value. Representatively, efficiencies in the inner ECAL region from 2018 data and simulated samples of the B^+ mode are used for these plots.

Trigger Weight Calculation

So far, all L0 trigger lines are considered as inclusive during the calibration, because the `TISTOS` method is not directly applicable to the exclusive L0L categories, which veto the `TIS` events that are required as `tag` sample in Eq. 4.18. For the L0I category, which is the inclusive category in this analysis, the trigger correction directly follows from Eq. 4.19

$$w_{\mathrm{L0}}^{\mathrm{L0I}} = \frac{\epsilon_{\mathrm{L0I}}^{\mathrm{data}}}{\epsilon_{\mathrm{L0I}}^{\mathrm{MC}}}\,, \tag{4.21}$$

while for the exclusive L0L categories the final weight is given by the product of the inclusive L0L weight and the ratio of the inverted L0I efficiencies

$$w_{\mathrm{L0}}^{\mathrm{L0L}} = \frac{\epsilon_{\mathrm{L0L}}^{\mathrm{data}}}{\epsilon_{\mathrm{L0L}}^{\mathrm{MC}}} \cdot \frac{1-\epsilon_{\mathrm{L0I}}^{\mathrm{data}}}{1-\epsilon_{\mathrm{L0I}}^{\mathrm{MC}}}\,. \tag{4.22}$$

Another complication arises in the L0L category from the possibility of both final state leptons to have fired the trigger line. As discussed in Sec. 4.2.2, the individual trigger

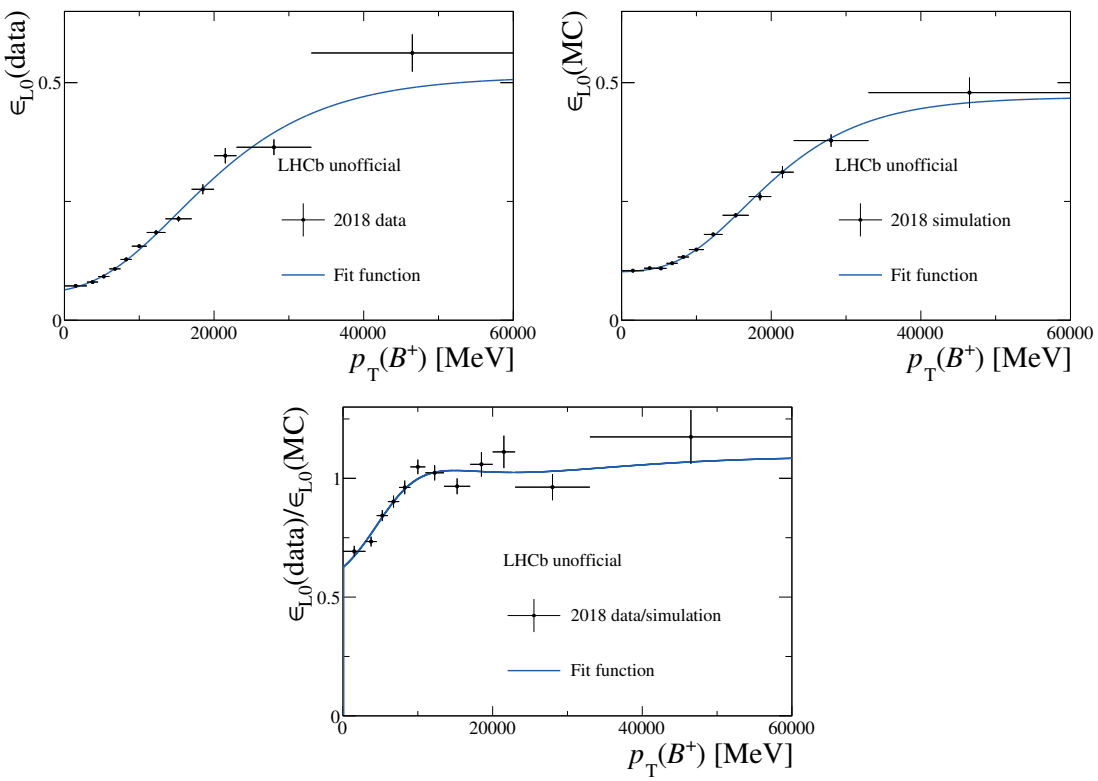

Figure 4.35.: Trigger turn-on curve plots showing the efficiency of the L0I trigger on data (left) and MC (right). The bottom plot shows the data over MC efficiency ratio. Representatively, efficiencies in the low `nTracks` bin from 2018 data and simulated samples of the $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ channel are used for these plots.

decisions of the two final state leptons are linked with the logical 'OR' condition so that their single efficiencies need to be combined using set theory

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\,.$$

Thus, equation 4.22 can be written as a function of the single lepton efficiencies

$$w_{\text{L0}}^{\text{L0L}} = \frac{(1 - (1 - \epsilon_{\text{L0L}}^{\text{data}}(\ell^+)) \cdot (1 - \epsilon_{\text{L0L}}^{\text{data}}(\ell^-)))}{(1 - (1 - \epsilon_{\text{L0L}}^{\text{MC}}(\ell^+)) \cdot (1 - \epsilon_{\text{L0L}}^{\text{MC}}(\ell^-)))} \cdot \frac{1 - \epsilon_{\text{L0I}}^{\text{data}}}{1 - \epsilon_{\text{L0I}}^{\text{MC}}}\,. \tag{4.23}$$

In conclusion, the formulas given in Eq. 4.21 and Eq. 4.23 are used to correct the selection efficiency on MC for the inclusive L0I and exclusive L0L categories, respectively.

4.4.5 HLT Trigger

The calibration of the HLT software trigger is done with the `TISTOS` method in an analogous way as for the L0 trigger. The two stages (HLT1 and HLT2) of the software trigger, which are linked with the logical 'AND' condition, are considered simultaneously.

However, because of correlations to the L0 response, individual sets of HLT corrections are calculated for each L0 category. Data and simulation samples of the J/ψ resonant mode, which have the full selection applied except for the HLT requirements, are used to calculate the efficiencies. Applying the data over MC ratio approach, the weights to correct the HLT response on simulation are calculated as

$$w_{\text{HLT}} = \frac{\epsilon^{\text{data}}_{\text{HLT}}}{\epsilon^{\text{MC}}_{\text{HLT}}} = \frac{\left(\frac{N_{\texttt{tag\&probe}}}{N_{\texttt{tag}}}\right)^{\text{data}}}{\left(\frac{N_{\texttt{tag\&probe}}}{N_{\texttt{tag}}}\right)^{\text{MC}}} \,. \tag{4.24}$$

A set of different HLT categories, which are independent to the signal `TOS` lines used in this analysis, are used as `tag` sample and are defined in Ref. [93]. In general the HLT efficiency shows a comparatively flat behaviour in all variables. The strongest dependence is found to be on the `nTracks` variable, which is therefore used to parameterise the calibration maps. Similar to the L0 correction, a systematic uncertainty is assigned on the choice of binning and parameterisation.

In each bin, the number of events in the `tag` ($N_{\texttt{tag}}$) and probe ($N_{\texttt{tag\&probe}}$) subsamples are determined with a `fit&count` approach similar to the one used for electron PID calibration. A simultaneous fit to the `tag` and `probe` samples is performed to extract these yields and calculate the HLT efficiency. The HLT efficiencies extracted with the `TISTOS` method for both data and simulated samples as well as the ratio between the two are displayed in Fig. 4.36 and Fig. 4.37 for the L0I and L0L categories, respectively.

4.4.6 Reconstruction of Signal Candidates

Differences between data and simulation originating from the reconstruction process are corrected for with the same method as used for the kinematic and multiplicity correction described in Sec. 4.4.3. The χ^2 of the impact parameter of both the B-meson ($\chi^2_{\text{IP}}(B^+)$) and J/ψ-meson ($\chi^2_{\text{IP}}(J/\psi)$) are used to correct the simulation, since reweighting these quantities is improving the agreement with data in many other variables of the reconstruction process as well (see App. C). The weights calculated to correct the reconstruction process are referred to as w_{Rec}. This correction is particularly important to calibrate the efficiency of the MVA requirements, which rely on the reconstruction variables in their training. In contrast to the kinematic and multiplicity correction, the reconstruction reweighting is not portable between electron and muons, because of differences in the reconstruction between the lepton species. Thus, the calculation of this correction is performed separately on muon and electron J/ψ resonant channels as well as for L0 trigger categories. The validation of this correction step is again achieved by studying the agreement of the reweighted simulated distributions with data. These plots are displayed for a large number of variables in App. C.

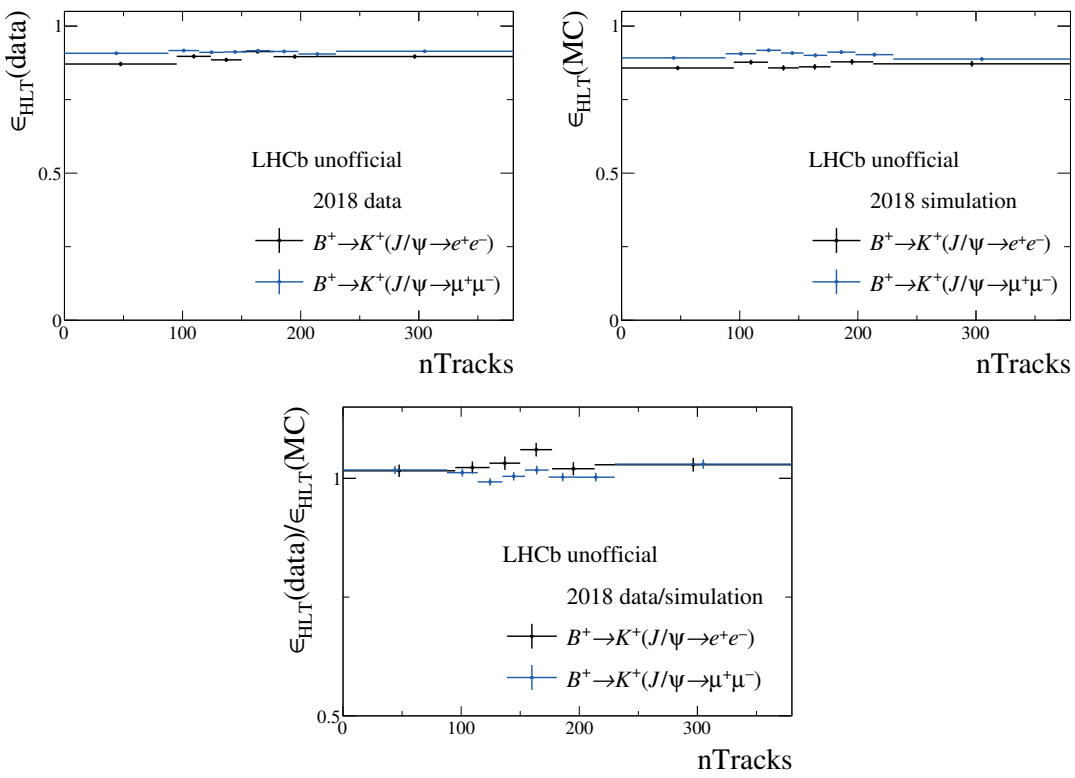

Figure 4.36.: HLT efficiencies for the L0I trigger category on data (left) and MC (right). The bottom plot shows the data over MC efficiency ratio. Representatively, efficiencies from 2018 data and simulated samples of the $B^+ \to K^+ J/\psi (\to \mu^+\mu^-)$ (blue) and $B^+ \to K^+ J/\psi (\to e^+e^-)$ (black) channel are used for these plots.

4.4.7 q^2 Selection Efficiency

The last step of the MC correction chain is the calibration of the selection efficiency of the q^2 windows used in the fits (see Tab. 4.4). The mass resolution of both the reconstructed B-meson and the q^2 system are known to be overestimated in simulation of electron channels with respect to data, while they are accurately described for muons. The main reasons for the mismodelled resolution on electron decays are on the one hand caused by the imperfectly simulated material budget resulting in a different behaviour of bremsstrahlung between data and simulation. On the other hand, the simulated detector occupancy, especially of the ECAL, differs from data, leading to a variation in the efficiency of the bremsstrahlungs recovery process. The resolution difference between data and simulation of the reconstructed B-meson mass causes a negligible effect on the signal efficiency estimation, since the efficiency of this selection is close to 100 % due to the wide range of the invariant B mass system used in the fits (see Sec. 4.6) . In contrast, the q^2 selection for the signal region is a stringent cut and its efficiency does not cancel in the ratio with the resonant channel. Moreover, the bin-migration effect described in Sec. 4.3.1 is dependent on the resolution of q^2.

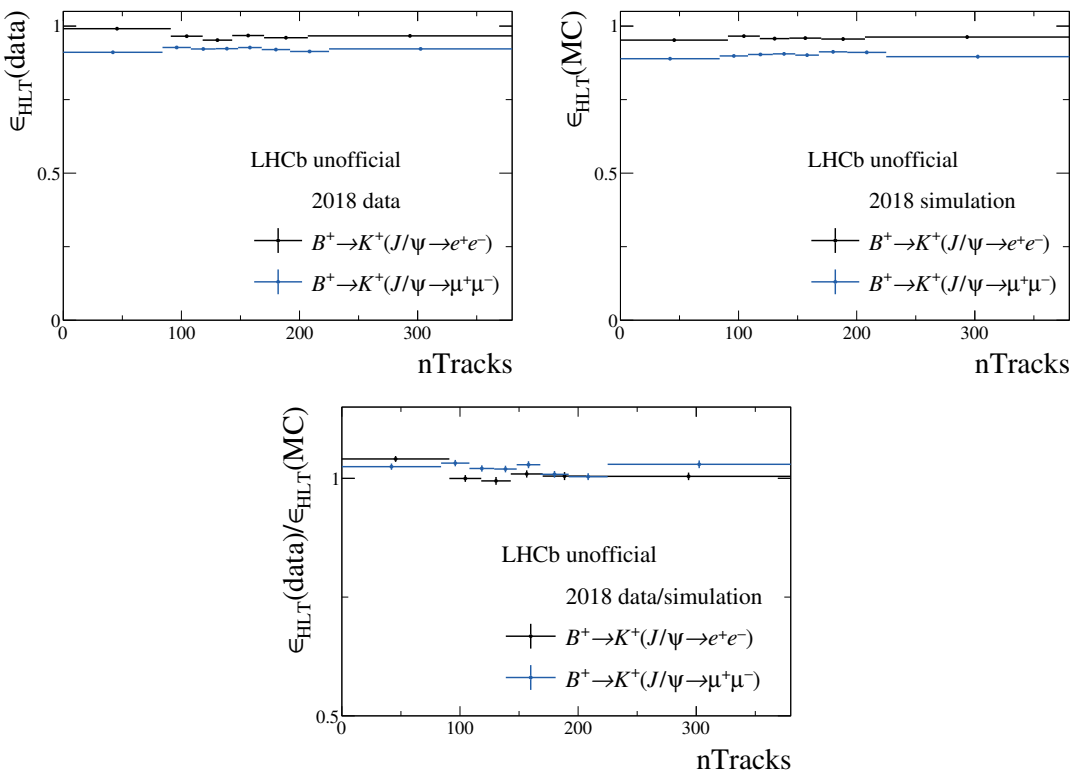

Figure 4.37.: HLT efficiencies for the L0L trigger category on data (left) and MC (right). The bottom plot shows the data over MC efficiency ratio. Representatively, efficiencies from 2018 data and simulated samples of the $B^+ \to K^+ J/\psi (\to \mu^+\mu^-)$ (blue) and $B^+ \to K^+ J/\psi (\to e^+e^-)$ (black) channel are used for these plots.

To correct the resolution and thus the efficiency of the mass window cuts, a 'smearing' of the q^2 spectrum is employed. A fit to the $m(e^+e^-)$ mass system on simulated samples of the J/ψ resonant modes is performed using a Double-Sided Crystal Ball (DSCB) function (as defined in Sec. 4.6.1) to describe the mass shape. This allows to extract the width σ_{MC} and mean value μ_{MC} of the distribution. Subsequently, a fit to $m(e^+e^-)$ on real data of the J/ψ mode is performed (see Fig. 4.38), where a scaling factor of the width $s_\sigma = {}^{\sigma_{\text{data}}}/_{\sigma_{\text{MC}}}$ and a shift parameter of the mean value $\Delta(\mu) = \mu_{\text{data}} - \mu_{\text{MC}}$ are allowed to float, while keeping the other shape parameters fixed to the converged values from the fit to simulation. Since the resolution difference between data and MC is strongly dependent on the bremsstrahlung recovery mechanism, the fits are performed for the three bremsstrahlungs categories defined in Sec. 4.6.1, separately.

With the assumption that the $m(e^+e^-)$ mass system on data is accurately described by simulation when using a distribution of width $\sigma_{\text{MC}} \cdot s_\sigma$ and with a mean value of $\mu_{\text{MC}} + \Delta(\mu)$, a smeared q^2 mass value is calculated as:

$$m_{\text{smeared}} = s_\sigma \cdot m_{\text{reco}} + \Delta(\mu) + (1 - s_\sigma) \cdot (\mu_{\text{MC}} + m_{\text{true}} - m_{J/\psi}^{\text{PDG}})\,, \tag{4.25}$$

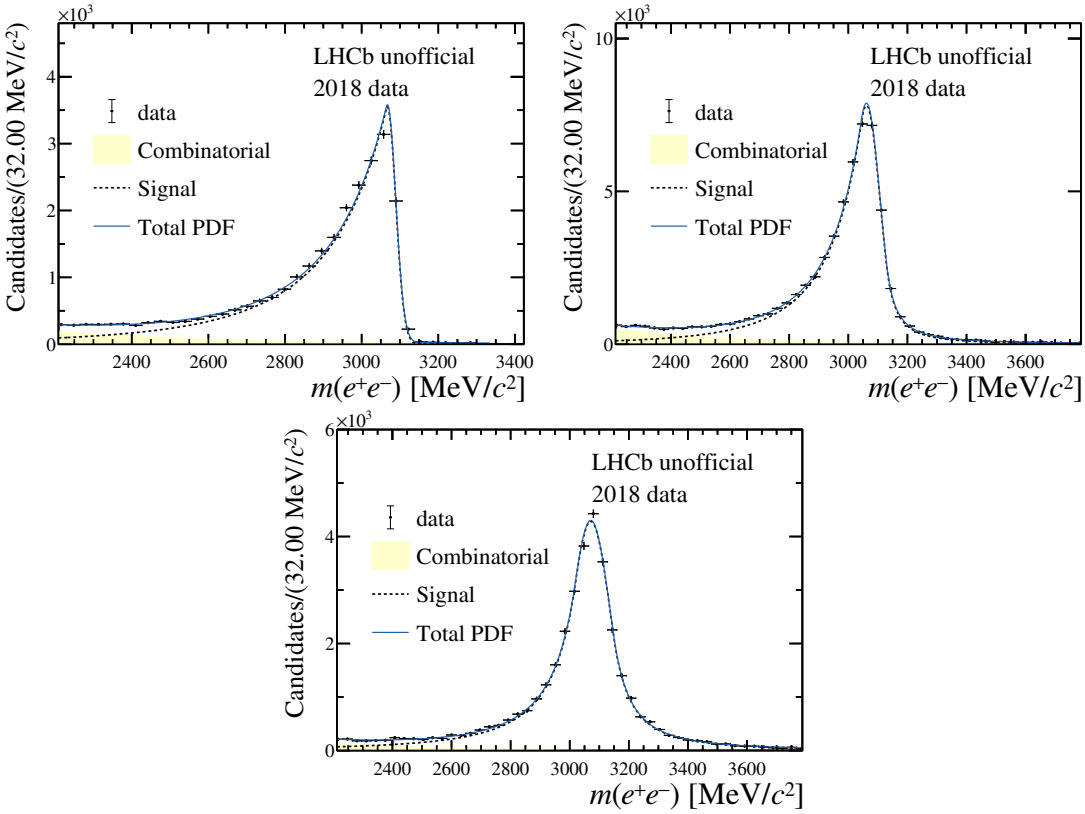

Figure 4.38.: Fit to the $m(e^+e^-)$ mass system of the $B^+ \to K^+ J/\psi (\to e^+e^-)$ channel for 2018 data in the L0I trigger category. The top left (right) plot shows the fit to the Brem0 (Brem1) category, while the bottom plot shows the Brem2 category. The various components used in the fit are detailed in the legends.

where m_{reco} is the reconstructed mass on simulation, m_{true} the generated dielectron mass before FSR and $m_{J/\psi}^{\text{PDG}}$ the measured J/ψ mass from PDG [22]. The formula effectively takes the generated mass (m_{true}) as a start value for each event. Two effects that shift the mean of the mass distribution are taken into account: Firstly, FSR and the detector resolution cause a shift of the reconstructed events and enter into m_{reco}. Secondly, differences between data and MC lead to another shift which is accounted for by adding $\Delta(\mu)$. Additionally, the scaling factor s_σ is applied to the distance between the reconstructed and generated mass on simulation to match the width observed on data.

To validate the method, a second fit to data is performed, where the initial mass shape is extracted on the smeared simulated $m(e^+e^-)$ distribution. These fits result in a good agreement of s_σ with unity and of $\Delta(\mu)$ with zero, which indicates a successful correction.

4.5 Efficiencies

The full efficiency ϵ_{Full} of a mode to be successfully reconstructed and selected by all applied requirements is given by the formula :

$$\epsilon_{\text{Full}} = \epsilon_{\text{Acc}} \cdot \epsilon_{\text{Trg}} \cdot \epsilon_{\text{Rec}} \cdot \epsilon_{\text{Sel}} \,, \tag{4.26}$$

with the individual efficiencies of the acceptance cut ϵ_{Acc}, the trigger lines ϵ_{Trg}, the reconstruction ϵ_{Rec} and the full offline selection ϵ_{Sel}. In the following, the strategies used to obtain these individual efficiencies are detailed.

4.5.1 Detector Acceptance

All final state particles of the probed decay modes need to have a polar angle of 10 to 400 mrad to fall into the geometric acceptance of the LHCb detector. The efficiency of this requirement, ϵ_{Acc}, is extracted on simulated samples of both rare and J/ψ resonant modes, where events are generated with `Pythia`. For the rare mode, separate values of ϵ_{Acc} are evaluated for low and central q^2, since this efficiency is known to depend on q^2. Also the centre-of-mass energy of the proton beams has an impact on this efficiency, which is why they are calculated for 7 TeV (2011), 8 TeV (2012) and 13 TeV (2015–2018), separately. The values of ϵ_{Acc} are reported in Tab. 4.21. The simulated samples used to calculate ϵ_{Acc} do not have the corrections from Sec. 4.4 applied. The only correction step with potential relevance for these generator-level samples is the reweighting of the B-meson kinematic, which however proved to have a negligible impact on the calculated efficiencies.

4.5.2 Reconstruction and Selection

The remaining efficiencies required in Eq. 4.26 are determined from fully reconstructed simulation, which have the complete set of MC corrections applied. This includes the full product of per-event weights

$$w_{\text{Full}} = w_{\text{PID}} \cdot w_{\text{TRK}} \cdot w_{\text{Kin\&Mult}} \cdot w_{\text{L0}} \cdot w_{\text{HLT}} \cdot w_{\text{Rec}}$$

and the usage of the smeared q^2 distribution.

All MC corrections that reweight distributions on simulation to match those on data and which do not describe direct efficiencies must be normalised in order to correctly determine the efficiencies. Specifically, this is necessary for the kinematic and multiplicity correction as well as for the reconstruction reweighting. To correct for as many biases as possible, the normalisation of the weights is performed at the earliest accessible stage. For the kinematic and multiplicity correction the weights are normalised directly at generator-level, while for the reconstruction reweighting the normalisation is done after stripping selection since the used variables describe the detector response and are therefore not accessible at earlier stages. Additionally, the L0 selection is applied for the reconstruction

Table 4.21.: Summary of the detector acceptance cut efficiencies of rare mode MC in low and central q^2 as well as of J/ψ resonant mode MC. The values are reported for all used configurations of the proton beam energy and are given in percent.

		ϵ_{Acc} [%]			
$\sqrt{s}$	q^2 Region	$K^+\mu^+\mu^-$	$K^+e^+e^-$	$K^{*0}\mu^+\mu^-$	$K^{*0}e^+e^-$
7 TeV	low	17.06 ± 0.05	16.79 ± 0.05	16.45 ± 0.03	16.02 ± 0.05
	central	16.01 ± 0.02	15.86 ± 0.02	15.43 ± 0.02	15.24 ± 0.04
	J/ψ	15.96 ± 0.03	15.85 ± 0.03	15.41 ± 0.03	15.29 ± 0.03
8 TeV	low	17.38 ± 0.03	17.03 ± 0.04	16.76 ± 0.02	16.36 ± 0.03
	central	16.36 ± 0.02	16.20 ± 0.02	15.74 ± 0.02	15.57 ± 0.02
	J/ψ	16.30 ± 0.03	16.23 ± 0.02	15.73 ± 0.02	15.65 ± 0.02
13 TeV	low	18.38 ± 0.04	18.09 ± 0.04	17.81 ± 0.03	17.29 ± 0.04
	central	17.42 ± 0.02	17.25 ± 0.02	16.73 ± 0.02	16.60 ± 0.03
	J/ψ	17.37 ± 0.02	17.23 ± 0.02	16.73 ± 0.02	16.62 ± 0.02

weight normalisation, because this correction step is not portable between the trigger categories (as already discussed in Sec. 4.4.6). Consequently, the efficiency of the trigger, full event reconstruction and selection is given as:

$$\begin{aligned}\epsilon_{\text{Trg}} \cdot \epsilon_{\text{Rec}} \cdot \epsilon_{\text{Sel}} = & \frac{\sum_{\text{Det}} w_{\text{Full}} \mid \text{Sel(Full)}}{\sum_{\text{Gen}} w_{\text{Kin\&Mult}} \mid \text{Sel}(\text{nSPD}, q^2_{\text{true}})} \\ & \cdot \frac{\sum_{\text{Det}} w_{\text{PID}} \cdot w_{\text{TRK}} \cdot w_{\text{Kin\&Mult}} \cdot w_{\text{L0}} \cdot w_{\text{HLT}} \mid \text{Sel(Norm)}}{\sum_{\text{Det}} w_{\text{Full}} \mid \text{Sel(Norm)}} ,\end{aligned} \tag{4.27}$$

where the sum of weights is calculated over events from detector- ($\sum_{\text{Det}}$) or generator-level ($\sum_{\text{Gen}}$) MC samples at a given selection stage Sel(X). Sel(Full) indicates the full selection, while Sel(Norm) refers to the selection stage used to normalise the reconstruction reweighting as described above. There are only two selection requirements applied to the generator-level MC Sel(nSPD, q^2_{true}): First, the `nSPDHits` cut from the stripping lines is applied in order to exclude it from the efficiency estimation. The reason for this is the poor modelling of the `nSPDHits` distribution in simulation as discussed in Sec. 4.4.3. Due to the independence of the event multiplicity from the lepton type, the bias of removing the `nSPDHits` cut in the calculation cancels in the efficiency ratios. Second, the q^2 window selection of the given bin is applied on the q^2_{true} distribution to account for bin migration as discussed in Sec. 4.3.1. Consequently, the efficiency for the q^2_{true} value of a simulated candidate to fall into the window of the studied q^2 region is not accounted for but cancels in the muon over electron ratios.

The efficiencies for the LFU measurements are calculated by applying the $w(B^0)$ corrections to all B^+ modes and $w(B^+)$ corrections to all B^0 modes. The portability of the MC calibration has been verified individually for each correction step and the corresponding studies are reported in Ref. [93]. Additionally, the cross-checks in Sec. 4.8 of single $r^{J/\psi}$ and double $R^{\psi(2S)}$ ratios are performed with both correction chains, to further demonstrate their portability. The final efficiencies for both rare and J/ψ resonant modes are reported in Tab. 4.22 for each year of data taking.

4.5.3 Correlations

Correlations from various sources enter the efficiency calculation and need to be taken into account. One source of correlations stems from the usage of J/ψ resonant data and simulated samples to calculate various MC corrections. However, the size of these correlations is reduced by swapping the correction chains between B^+ and B^0 modes. Another source originates from the porting of some corrections between modes, *e.g.* the $w_{\text{Kin\&Mult}}$ corrections are calculated on the J/ψ resonant inclusive `L0Muon` samples and are applied to all modes and trigger categories. Further, there are correlations between efficiencies of different run periods, trigger categories and decay modes which stem from using the same methods to correct the simulation, such as the chosen parameterisation of L0 corrections, across all cases.

The correlations are accounted for by including covariance matrices in the final fit for the LFU observables. To this end, the correlations are evaluated with the `bootstrapping` technique: A set of $n = 100$ weights from a Poisson distribution with a mean value of unity is assigned to each event in the data and simulated samples. For all n sets of weights, the full MC calibration and all efficiencies are recalculated. This results in n efficiency values which follow a multidimensional Normal distribution as shown in Ref. [93].

The sample covariance for efficiencies ϵ_1 and ϵ_2, *e.g.* of two different run periods, is calculated as

$$\text{cov}_{1,2} = \frac{1}{n-1}\sum_{i=1}^{n}(\epsilon_1^i - \bar{\epsilon}_1)\cdot(\epsilon_2^i - \bar{\epsilon}_2)\,,$$

where ϵ^i is the efficiency from the i-th `bootstrap` and $\bar{\epsilon}$ the average efficiency. The `bootstrapping` also accounts for the statistical uncertainty σ on the efficiencies coming from the limited size of the calibration samples, which is defined as

$$\sigma_1 = \sqrt{\text{cov}_{1,1}} = \left(\frac{1}{n-1}\sum_{i=1}^{n}(\epsilon_1^i - \bar{\epsilon}_1)^2\right)^{1/2}.$$

The definition of the correlation $\text{cor}_{1,2}$ between two efficiencies follows as

$$\text{cor}_{1,2} = \frac{\text{cov}_{1,2}}{\sigma_1\cdot\sigma_2}\,. \tag{4.28}$$

In practice, the covariance matrix is of dimension $m = 24$ covering the four LFU bins in the three run periods as well as in both L0 trigger categories. In the final fits for the LFU

Table 4.22.: Summary of total efficiencies of the rare modes in low and central q^2 as well as of J/ψ resonant mode MC for all run periods and trigger categories. The uncertainties follow from the `bootstrapping` technique. All values are given in permille. The efficiency of the q^2_{true} selection is not included in these efficiencies, only bin migration effects are accounted for.

			$\epsilon_{\text{Full}}\,[10^{-3}]$			
Run Period	q^2 Region	Trigger	$K^+\mu^+\mu^-$	$K^+e^+e^-$	$K^{*0}\mu^+\mu^-$	$K^{*0}e^+e^-$
Run 1	low	L0I	3.93 ± 0.06	1.52 ± 0.05	1.46 ± 0.02	0.54 ± 0.01
		L0L	10.07 ± 0.13	1.26 ± 0.04	3.48 ± 0.04	0.44 ± 0.01
	central	L0I	3.73 ± 0.05	1.38 ± 0.04	1.30 ± 0.01	0.51 ± 0.01
		L0L	9.39 ± 0.10	1.27 ± 0.03	3.06 ± 0.02	0.44 ± 0.01
	J/ψ	L0I	4.64 ± 0.05	1.68 ± 0.04	1.69 ± 0.01	0.65 ± 0.01
		L0L	12.76 ± 0.12	1.81 ± 0.04	4.39 ± 0.03	0.70 ± 0.01
Run 2p1	low	L0I	5.25 ± 0.09	2.50 ± 0.06	2.06 ± 0.03	0.84 ± 0.02
		L0L	11.61 ± 0.17	2.39 ± 0.06	4.31 ± 0.04	0.74 ± 0.02
	central	L0I	5.05 ± 0.06	2.30 ± 0.04	1.86 ± 0.03	0.87 ± 0.02
		L0L	11.30 ± 0.12	2.28 ± 0.04	3.82 ± 0.04	0.81 ± 0.02
	J/ψ	L0I	6.13 ± 0.07	2.71 ± 0.05	2.36 ± 0.02	1.03 ± 0.01
		L0L	14.75 ± 0.14	3.15 ± 0.05	5.30 ± 0.03	1.18 ± 0.01
Run 2p2	low	L0I	5.34 ± 0.06	2.56 ± 0.05	2.07 ± 0.02	0.86 ± 0.02
		L0L	13.14 ± 0.13	2.28 ± 0.03	4.71 ± 0.04	0.79 ± 0.02
	central	L0I	5.12 ± 0.05	2.32 ± 0.03	1.89 ± 0.02	0.84 ± 0.01
		L0L	12.56 ± 0.08	2.26 ± 0.03	4.18 ± 0.03	0.78 ± 0.01
	J/ψ	L0I	6.13 ± 0.05	2.71 ± 0.04	2.34 ± 0.01	1.02 ± 0.01
		L0L	16.41 ± 0.09	3.09 ± 0.03	5.82 ± 0.02	1.17 ± 0.01

observables, the efficiencies are used in the mass fits to extract the desired ratio values directly. However, the efficiencies are not fixed, but are Gaussian-constrained around their nominal values using the covariance matrix which follows from the `bootstrapping` approach described above.

Contributions to the covariance matrix from other sources, such as systematic uncertainties, will also enter into the final fit. Computation and size of these contributions are detailed in Sec. 4.7.1.

4.6 Invariant Mass Fits

In order to determine the yields of both rare and resonant modes used to measure the LFU observables, one-dimensional maximum likelihood fits to the invariant mass of the reconstructed B-meson candidates, which survive the full selection, are performed. The fits are performed simultaneously across all modes, run periods and L0 trigger categories, allowing constraints between different modes to be introduced, *e.g.* the yield of the cross-feed background in the $B^+ \to K^+e^+e^-$ fit is constrained by the signal yield of the $B^0 \to K^{*0}e^+e^-$ channel. The fits to the rare modes are unbinned, while fits to the high statistics resonant modes are binned in the invariant mass of the B-meson to reduce the computational effort. To improve the resolution of the invariant mass, the DTF algorithm is exploited to constrain all final state particles originating from the same PV. Further, the DTF algorithm is used to constrain the dilepton system to the J/ψ and $\psi(2S)$ masses when studying the $r^{J/\psi}$ and $R^{\psi(2S)}$ ratios. The fitter implementation is based on the `RooFit` software package [112]. Alongside the signal, various sources of background need to be considered in the invariant mass fits as discussed in Sec. 4.3. Fully selected and calibrated simulated samples of both signal and background modes are used to determine PDFs describing the mass lineshapes for each component, with combinatorial background being the one exception (see below). Individual PDFs are obtained for each run period and trigger category in order to account for kinematic differences.

The strategy used to model the individual components is discussed in the next sections. Afterwards, the setup of the simultaneous fit and its results are presented. Lastly, the fit result is validated using pseudo-experiments.

4.6.1 Signal Shape

The PDFs used to describe the signal component in the mass fits are extracted from simulated samples of the relevant decay modes. Analytical functions such as Gaussian and Crystal Ball (CB) [113] functions, the latter being a combination of a Gaussian core with an power-law tail, are used to describe the signal shape in the mass fits. Different functions are used for electron and muon modes as well as for J/ψ and $\psi(2S)$ resonant modes, when fitting with or without the dilepton mass constraint. However, the type of signal PDF is the same for B^+ and B^0 decays, therefore only B^+ modes are shown for demonstration in the plots below.

For fit to the unconstrained muon J/ψ resonant channel, a modified CB function is used to model the signal shape, which has power-law tails both above and below the Gaussian core and is referred to as Double-Sided Crystal Ball (DSCB) function. Especially due to the q^2 window selection a departure from the expected behaviour in the tails is visible

motivating the addition of two Gauss functions to the DSCB to further improve the modelling in the lower and upper mass region. The same function is used to fit the simulation of the rare muon channel. Example fits to Run 2p2 $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ MC in J/ψ q^2 as well as to $B^+ \to K^+\mu^+\mu^-$ MC in low and central q^2 are shown in Fig. 4.39.

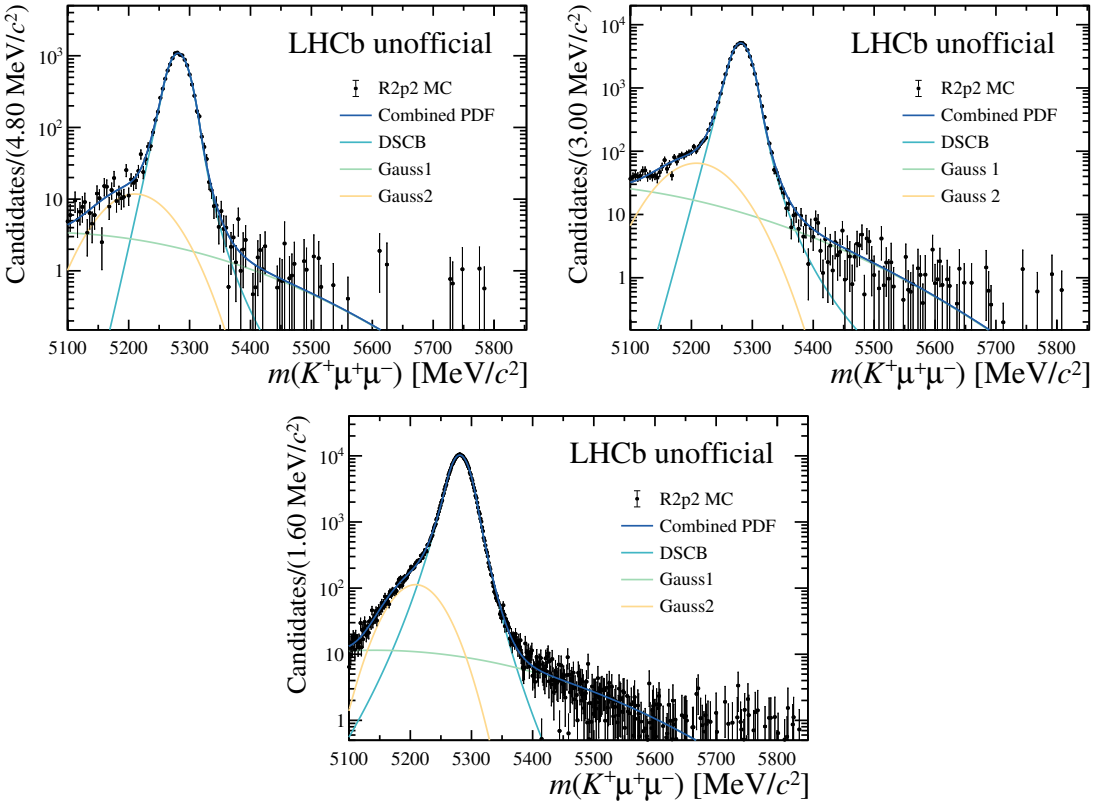

Figure 4.39.: B mass distribution of Run 2p2 simulated $B^+ \to K^+\mu^+\mu^-$ candidates in low (left) and central (right) q^2 as well as $B^+ \to K^+ J/\psi\,(\to \mu^+\mu^-)$ candidates in J/ψ q^2. The shapes resulting from the fit of the separate analytical functions together with their combination are shown in different colors (see legend).

For electron modes the B mass shape of the signal strongly depends on the response of the bremsstrahlung recovery algorithm. Consequently, the candidates are divided into three bremsstrahlung categories:

Brem0: no photon is assigned to the dielectron pair;

Brem1: one photon is assigned to the dielectron pair;

Brem2: at least two photons are assigned to the dielectron pair.

Subsequently, individual fits to each bremsstrahlung category are performed, again using a DSCB function to describe the signal peak. In case of the Brem0 category, a single Gauss function is added to improve the modelling of the lower mass tail, while for Brem1

and Brem2 a second Gauss function is added for the higher mass region. Figure 4.40 shows the PDFs resulting from fits to the individual bremsstrahlung categories together with their combination, which is used for the data fit.

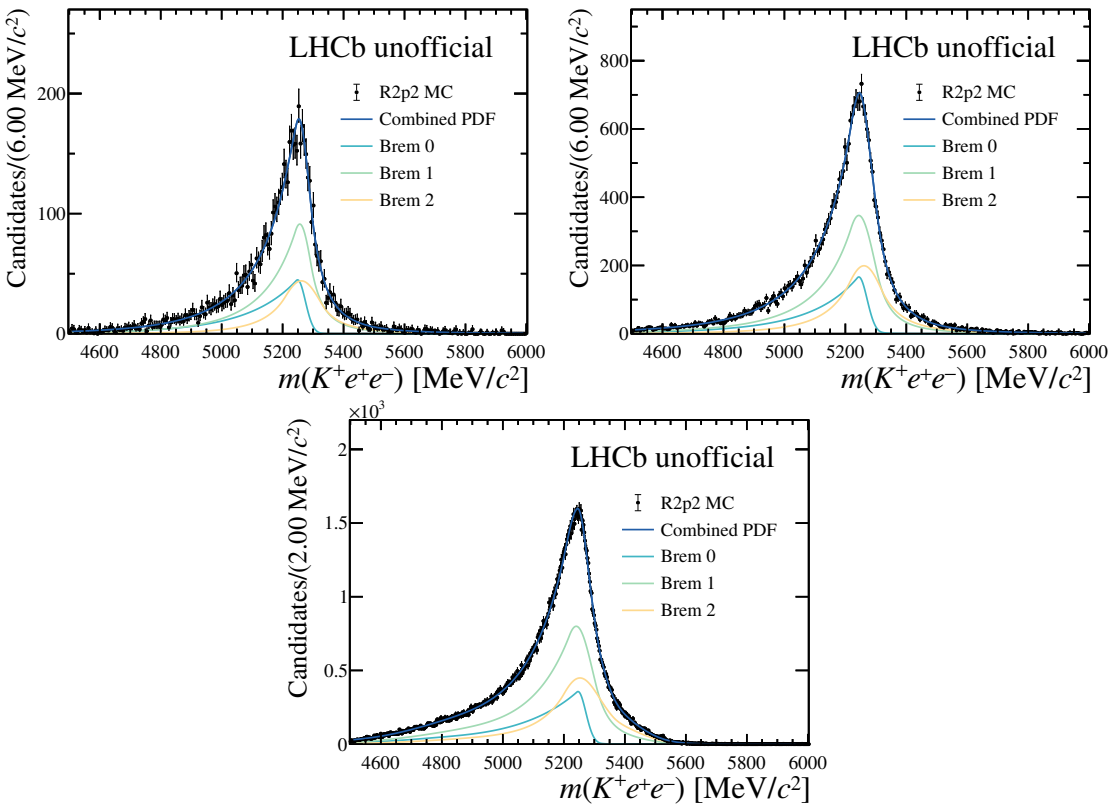

Figure 4.40.: B mass distribution of Run 2p2 simulated $B^+ \to K^+e^+e^-$ candidates in low (left) and central (right) q^2 as well as $B^+ \to K^+J/\psi\,(\to e^+e^-)$ candidates in J/ψ q^2. The shapes resulting from the fit to the three bremsstrahlung categories together with their combination are shown in different colors (see legend).

4.6.2 Combinatorial Background

The combination of random tracks in the event is one of the key backgrounds for all modes and regions of q^2. This background is distributed continuously above the kinematic threshold, *e.g.* $m(K^+) + m(J/\psi)$ in case of combinatorial background with a true J/ψ for the $B^+ \to K^+J/\psi$ mode. Therefore, the background is modelled with a simple exponential function in the mass fits, since the kinematic threshold falls outside the fit range. However, it is possible that the applied selection introduces a departure of the expected exponential behaviour. A study based on Same Sign (SS) data is performed to validate the shape of the combinatorial background after full selection. SS data refers to samples of real data events gathered from the reconstruction of $K^\pm\ell^+\ell^+$ and $K^\pm\ell^-\ell^-$ candidates for B^+ modes and $K^\pm\pi^\mp\ell^+\ell^+$ and $K^\mp\pi^\pm\ell^-\ell^-$ candidates for B^0 modes.

Thus, these samples are populated purely by combinatorial events (assuming lepton number conservation). The distribution of fully selected SS data is well described by an exponential function for all resonant modes as well as for muon rare modes. However, when looking at the rare electron modes in an extended B mass range down to 4000 MeV, a sculpting of the mass distribution is visible that is caused by the additional HOP and partially reconstructed MVA selection criteria applied in these q^2 bins. Figure 4.41 shows the mass distribution of SS data for all rare electron modes, where the full Run 1 and Run 2 datasets are used, in order to improve the statistics after the full selection. Especially for the low q^2 bins, a clear sculpting is visible that also enters into the nominal mass range used in the fits. To describe this effect, the exponential function is multiplied by a turn-on curve

$$f(m) = \frac{1}{1 + \exp\left(-a \cdot (m - m_0)\right)} \cdot e^{b \cdot m},$$

where a controls the slope and m_0 the position of the turn-on part. The two parameters of the turn-on function are extracted from the fit to SS data and are fixed in the final data fit. This way, the sculpting of the combinatorial background in the rare electron fits is accounted for while maintaining the fit stability. Additionally, the slope of the exponential function is the only degree of freedom for the shape of the combinatorial background for all modes in the simultaneous fit. A systematic uncertainty is assigned to the uncertainties of the parameters describing the turn-on curve (see Sec. 4.7.2).

4.6.3 Physical Backgrounds

In the following, all backgrounds considered in the mass fits (see Tab. 4.23), that stem from physical processes, are introduced. Further, the strategy used to describe their shape in the reconstructed B mass is described. Most backgrounds are modelled with the `RooKeysPDF` class from `RooFit`, which approximates the simulated mass shape with a superposition of one-dimensional Gaussian kernels also referred to as Kernel Density Estimator (KDE). An example plot of a KDE modelling is shown in Fig. 4.42 for Run 2p2 simulation of $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ reconstructed as B^0 candidates. Other backgrounds are modelled with analytical functions similar to the signal mass shapes. In general, the mass shape of each background is fixed from the modelling on simulation allowing only the yield of to float in the final fit to data.

B^+ Resonant Modes

The only relevant source of misidentification backgrounds for the $B^+ \to K^+J/\psi$ modes is coming from the Cabibbo-suppressed $B^+ \to \pi^+J/\psi\,(\to \ell^+\ell^-)$ decay, which is modelled with a DSCB (see Fig. 4.43). Further, two types of partially reconstructed backgrounds need to be considered in mass fits to these modes: the 'hadronic' partially reconstructed modes ($B^+ \to XJ/\psi\,(\to \ell^+\ell^-)$) from decays such as $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$; and the 'charmonium' partially reconstructed modes ($B^+ \to K^+\psi(2S)(\to XJ/\psi\,)$) from decays

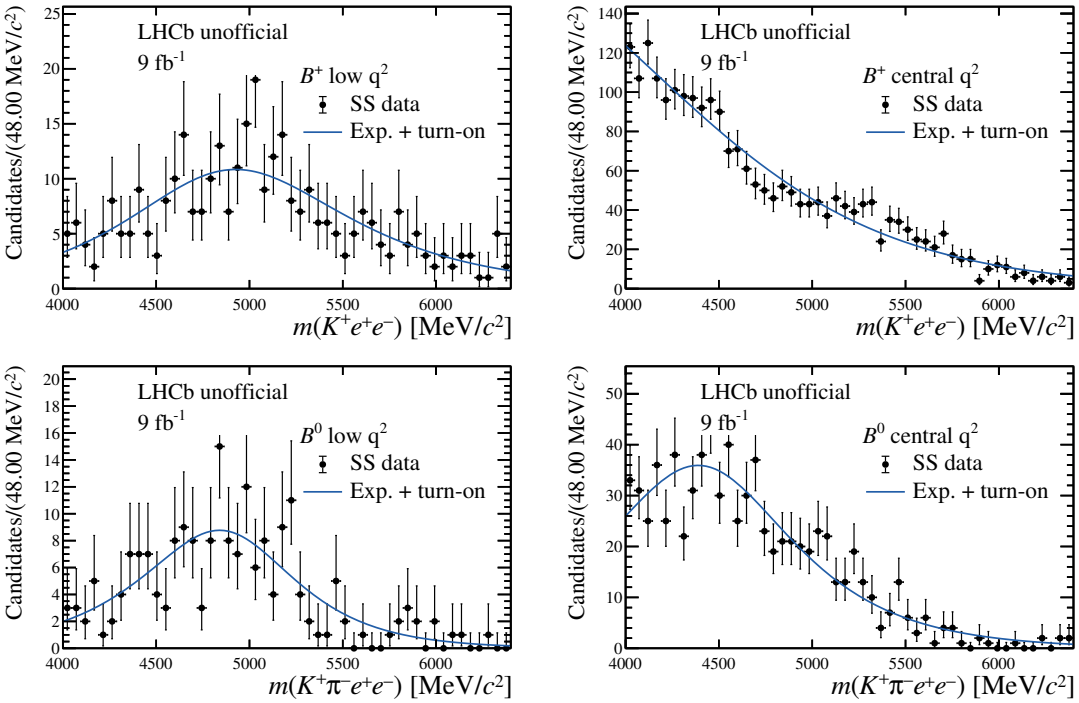

Figure 4.41.: B mass distribution of fully selected SS data merged over the Run 1 and Run 2 datasets for all rare electron modes. The fit result of the exponential function with an additional turn-on term is shown in blue.

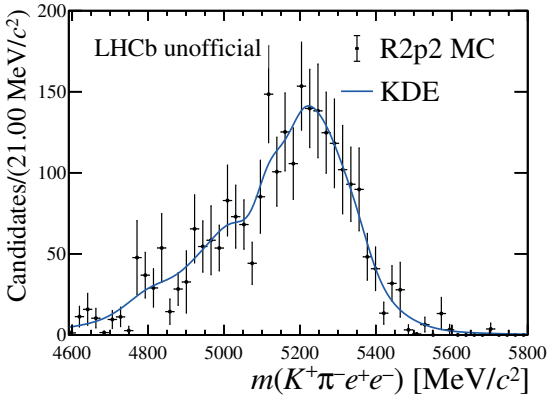

Figure 4.42.: Mass distribution of Run 2p2 $\Lambda_b^0 \to pKJ/\psi(\to e^+e^-)$ simulation reconstructed as B^0 candidates in J/ψ q^2 region together with the mass shape of the KDE.

of higher $c\bar{c}$ resonances such as $\psi(2S) \to \pi\pi J/\psi$. A large number of channels can contribute to both of these background types, thus inclusive simulated samples containing a multitude of decay modes of B^+, B^0 and B_s^0 mesons are used to extract the mass shape via a `RooKeysPDF`. The relative fractions between the B meson species is fixed

according to their relative production fractions and decay rates. Another partially reconstructed background from $B_s^0 \to \overline{K}^{*0} J/\psi\,(\to \ell^+\ell^-)$ decays is modelled with a separate `RooKeysPDF`, since this mode is not included in the inclusive samples.

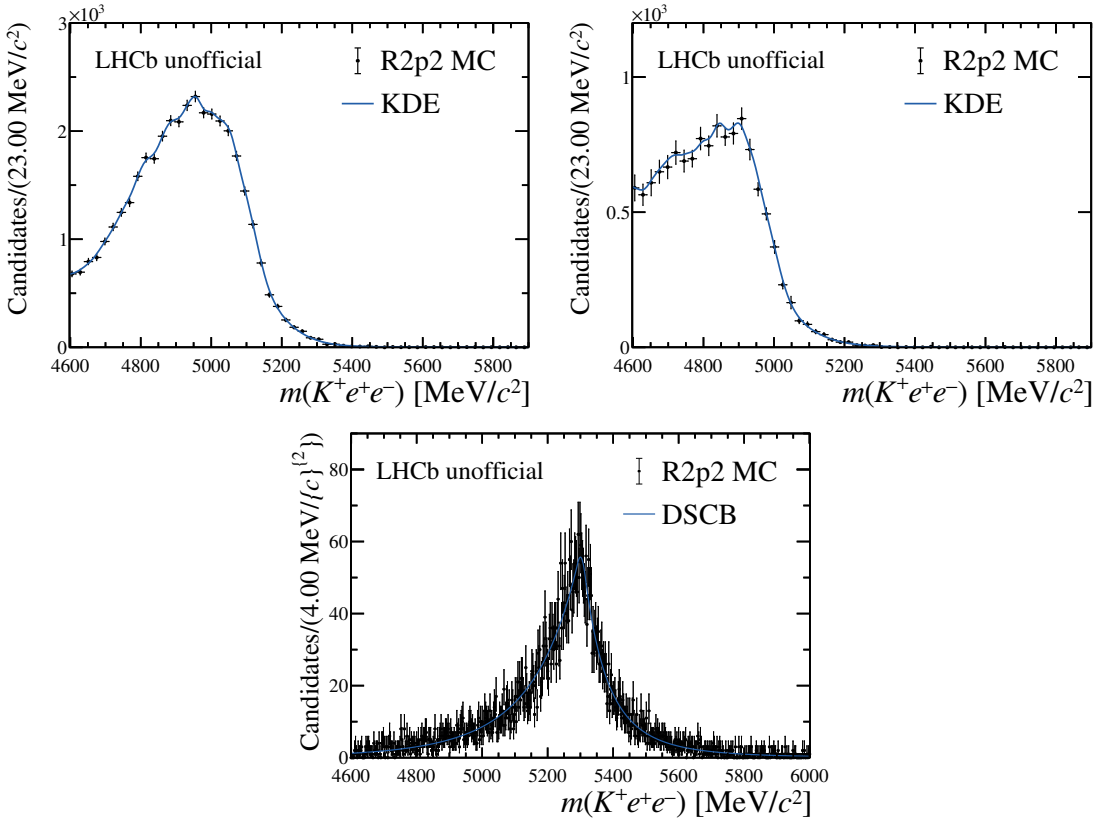

Figure 4.43.: Mass shapes for backgrounds considered in B^+ resonant modes shown for Run 2p2 MC in L0I. The top left (right) plot shows hadronic (charmonium) partially reconstructed background models, while the $B^+ \to \pi^+ J/\psi\,(\to \ell^+\ell^-)$ misidentification background is shown at the bottom.

B^0 Resonant Modes

The relevant sources of misidentification that need to be considered when fitting to the $B^0 \to K^{*0} J/\psi$ modes are coming from $\Lambda_b^0 \to pK J/\psi\,(\to \ell^+\ell^-)$ and $B_s^0 \to \phi J/\psi\,(\to \ell^+\ell^-)$ decays as well as from $B^0 \to K^{*0} J/\psi$ decays themselves with a hadronic swap ($K \leftrightarrow \pi$). All of these backgrounds are modelled using `RooKeysPDF`s. The hadronic and charmonium partially reconstructed backgrounds are treated in an analogous way to the B^+ resonant fits. Additionally, another background arises from $B_s^0 \to \overline{K}^{*0} J/\psi\,(\to \ell^+\ell^-)$ decays. Since they share the same final state, the same PDF as used for the J/ψ resonant mode is used to model this background, shifted by the B^0 and B_s^0 mass difference.

Rare Modes

Misidentification backgrounds are found to be negligible for all rare modes and are thus not considered in the mass fits. Further, in case of rare muon modes, there is no need to account for any partially reconstructed backgrounds due to the small mass range. Thus, only the combinatorial background is modelled besides the signal in low and central q^2 muon fits for both B^+ and B^0 modes.

On the other hand, for rare electron modes it is necessary to include partially reconstructed components in the mass fits for both low and central q^2 regions, because of the extended mass region and worse resolution compared to muon modes. In mass fits to the electron B^0 mode, this background type is modelled with simulated samples of $B^+ \to K^+\pi^+\pi^- e^+e^-$ decays, while two individual partially reconstructed decay modes are included in B^+ modes: one from $B^0 \to K^{*0}e^+e^-$ decays and another from the isospin partner decay $B^+ \to K^{*+}e^+e^-$, where the K^{*+} decays into $K^+\pi^0$. Additionally, the $B^+ \to K^+\eta'(\to e^+e^-\gamma)$ decay is considered for $B^+ \to K^+e^+e^-$ decays in the low q^2 region. All these modes are modelled with `RooKeysPDF` extracted from simulated samples of the respective decays modes.

Lastly, another important component is needed to describe the leakage background from the J/ψ resonant mode in fits to the electron modes in central q^2. Since only events with a significant energy loss in the reconstructed J/ψ can enter in this q^2 window, the reconstructed B mass is shifted to lower values as well so that these events populate the low invariant mass region. This background is modelled with simulated samples of the J/ψ resonant mode selected in central q^2 using a `RooKeysPDF` to describe the B mass shape.

4.6.4 The Simultaneous Mass Fit

This section introduces the setup of the simultaneous fit in terms of its floating parameters, the implemented constraints and the usage of efficiencies as well as the result of the final data fit. A summary of all applied constraints in the fit is given in Tab. 4.24.

Mass Shapes

As described in Sec. 4.6.1, the mass shapes of both rare and resonant modes are extracted from fits to simulated samples of the respective decay mode. When fitting to real data, the parameters of the signal PDFs are fixed to the values obtained from simulation. However, to account for residual data and MC differences, a scaling factor of the width $s_\sigma = \sigma_{\text{data}}/\sigma_{\text{MC}}$ and a shift parameter of the mean value $\Delta(\mu) = \mu_{\text{data}} - \mu_{\text{MC}}$ of the PDF are in the data fits. These parameters are independent between electron and muon channels as well as between different run periods and trigger categories. However, they are shared between rare and resonant modes in order to improve the fit stability in the low statistics FCNC fits.

Table 4.23.: Sources of physical background which must be taken into account in the simultaneous mass fits.

Channel	**Type**	**Background mode**
$B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$	Misidentification	$B_s^0 \to (\phi \to K^+K^-)(J/\psi \to \ell^+\ell^-)$
		$\Lambda_b^0 \to pKJ/\psi\,(\to \ell^+\ell^-)$
		$B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ $(K \leftrightarrow \pi)$
	Part. Reco.	$B \to XJ/\psi\,(\to \ell^+\ell^-)$
		$B^0 \to K^{*0}\psi(2S)(\to XJ/\psi\,)$
	Same Final State	$B_s^0 \to \overline{K}^{*0} J/\psi\,(\to \ell^+\ell^-)$
$B^+ \to K^+ J/\psi\,(\to \ell^+\ell^-)$	Misidentification	$B^+ \to \pi^+ J/\psi\,(\to \ell^+\ell^-)$
	Part. Reco.	$B \to XJ/\psi\,(\to \ell^+\ell^-)$
		$B^+ \to K^+\psi(2S)(\to XJ/\psi\,)$
		$B_s^0 \to \overline{K}^{*0} J/\psi\,(\to \ell^+\ell^-)$
$B^0 \to K^{*0} e^+e^-$	Part. Reco.	$B^+ \to K^+\pi^+\pi^-e^+e^-$
	Leakage	$B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$
$B^+ \to K^+ e^+e^-$	Part. Reco.	$B^0 \to K^{*0} e^+e^-$
		$B^+ \to K^{*+} e^+e^-$
		$B^+ \to K^+\eta'(\to e^+e^-\gamma)$
	Leakage	$B^+ \to K^+ J/\psi\,(\to e^+e^-)$

The slope of the exponential function used to describe the combinatorial background is free to float in the data fits, while the parameters of the turn-on function in the rare electron modes are fixed from the SS data fits. The mass shapes of all other background components are fully fixed from the MC fits.

Background Yield Constraints

For many background components the yield is not free to float, but is constrained based on efficiency ratios and the signal yields in the simultaneous fit. This way, both the fit stability and the sensitivity to the signal yields are improved. The idea of these background yield constraints is to perform a normalisation according to Eq. 4.11 using efficiencies extracted from fully selected and calibrated simulated samples and a normalisation channel which is dependent on the type of background. The constraint is implemented as a penalty term in the likelihood calculation using a Gaussian function, where the mean is given by the expectation value of the normalisation and the width by the propagated uncertainties.

All yields of misidentification backgrounds as well as of the $B_s^0 \to \overline{K}^{*0} J/\psi\,(\to \ell^+\ell^-)$ component used in fits to the J/ψ region are constrained to the respective resonant mode. For example, the yield of the $\Lambda_b^0 \to pKJ/\psi\,(\to e^+e^-)$ component is normalised to the $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ yield taking into account the ratio of their efficiencies and branching fractions as well as the different fragmentation fractions.

The yield of the leakage backgrounds from $J/\psi\,(e^+e^-)$ resonant decays into the central q^2 window is calculated in an analogous way using the yield of the resonant mode in the J/ψ q^2 region as normalisation channel. A conservative uncertainty of 20 % is assigned to this constraint to reflect the large uncertainty on the simulation based efficiency estimate for the leakage component as well as to account for any residual differences between data and MC.

The partially reconstructed background yield from $B^0 \to K^{*0} e^+e^-$ and $B^+ \to K^{*+} e^+e^-$ decays in the $B^+ \to K^+e^+e^-$ fit is normalised to the signal yield of the fit to the $B^0 \to K^{*0}e^+e^-$ channel. This feature is one of the advantages of the simultaneous measurement of R_K and $R_{K^{*0}}$, since it improves not only sensitivity, but also enables a coherent measurement and the full extraction of the correlation matrix. The calculation of the constraint takes the absence of the K^{*0} mass window selection in B^+ modes into account. The relative efficiencies of $B^0 \to K^{*0}e^+e^-$ between the signal reconstruction and the partial reconstruction as $B^+ \to K^+e^+e^-$ enter in the normalisation. For the isospin conjugate $B^+ \to K^{*+}e^+e^-$ decay an additional factor, f_{iso}, is calculated, accounting for the relative branching fractions of $B^0 \to K^{*0}\ell^+\ell^-$ and $B^+ \to K^{*+}\ell^+\ell^-$, which are taken from measurements of the muon channels (see Refs. [34, 35]). Additional contributions from non-resonant $B^0 \to K^+\pi^-e^+e^-$ (S-wave) decays are considered in this constraint as well. The reconstruction and selection efficiency is obtained from $B^0 \to K^+\pi^-e^+e^-$ simulated samples which rely on the phase-space model for the hadronic system. The branching fraction of the S-wave mode is taken from Ref. [34] and the isospin conjugate S-wave modes are accounted for by including f_{iso} in the calculation.

Another source of partially reconstructed backgrounds considered in the fit to the $B^+ \to K^+e^+e^-$ mode comes from $B^+ \to K^+\eta'(\to e^+e^-\gamma)$ decays, which is relevant for low q^2 only. The yield of this component is normalised to the yield of the $B^+ \to K^+J/\psi(\to e^+e^-)$ channel.

In all other modes, the yield of the partially reconstructed component is allowed to float freely. The yield of the combinatorial background is unconstrained in all modes.

Table 4.24.: List of all constraints applied in the simultaneous fit.

Category	Constrained parameter	Constrained to
Signal shape	s_σ of the rare modes	s_σ of the J/ψ mode
	$\Delta(\mu)$ of the rare modes	$\Delta(\mu)$ of the J/ψ mode
Background Yields	All misidentified backgrounds in J/ψ modes	J/ψ mode yield
	$B_s^0 \to K^{*0}J/\psi$ background in $B^0 \to K^{*0}J/\psi$ modes	J/ψ mode yield
	Leakage from J/ψ mode into central q^2	J/ψ mode yield
	Partially reconstructed backgrounds in $B^+ \to K^+e^+e^-$	$B^0 \to K^{*0}e^+e^-$ yield
	$B^+ \to K^+\eta'(\to e^+e^-\gamma)$ in $B^+ \to K^+e^+e^-$ low q^2	J/ψ mode yield

Data Fit Results

The final simultaneous fit is designed such that the LFU ratios R_K and $R_{K^{*0}}$ are extracted directly. To this end, the efficiencies of rare and resonant modes are injected into the data fit as discussed in Sec. 4.5.3. The plots of the mass fits are given in Fig. 4.44 and 4.45 for muon and electron modes, respectively. The yields of the rare channels resulting from the simultaneous fit are given in Tab. 4.25, while the values of the LFU observables are given in Sec. 4.10. The yields of the J/ψ resonant mode are provided in App. D. One-dimensional scans of the logarithmic likelihood of the fit are performed for all four LFU bins as well as for the two single $r^{J/\psi}$ parameters (one for B^+ and one for B^0 modes) that are floating in the fit. They are calculated by forcing the studied parameter to a value different from the best-fit value of the nominal fit and re-minimising the likelihood with respect to the nuisance parameters. Afterwards, the difference of the resulting logarithmic likelihood from the likelihood obtained by the baseline fit is evaluated. This procedure is done for multiple steps of the parameter inside the scanned window. The scans are given in Fig. 4.46, where the central values of the LFU ratios are scaled with a blinding factor. Further, the one, two and three σ confidence intervals are plotted, that are calculated assuming Wilks theorem [114], which relates the likelihood difference with the χ^2 distribution as $-2\Delta\log(\mathcal{L}) = \chi^2$ when scanning with one degree of freedom. The scans of the R_K and $R_{K^{*0}}$ ratios show a slightly asymmetric behaviour, which is caused by the small statistics in the rare electron modes. In the next section, the strategy used to properly account for the effect of these asymmetries on the final result is discussed.

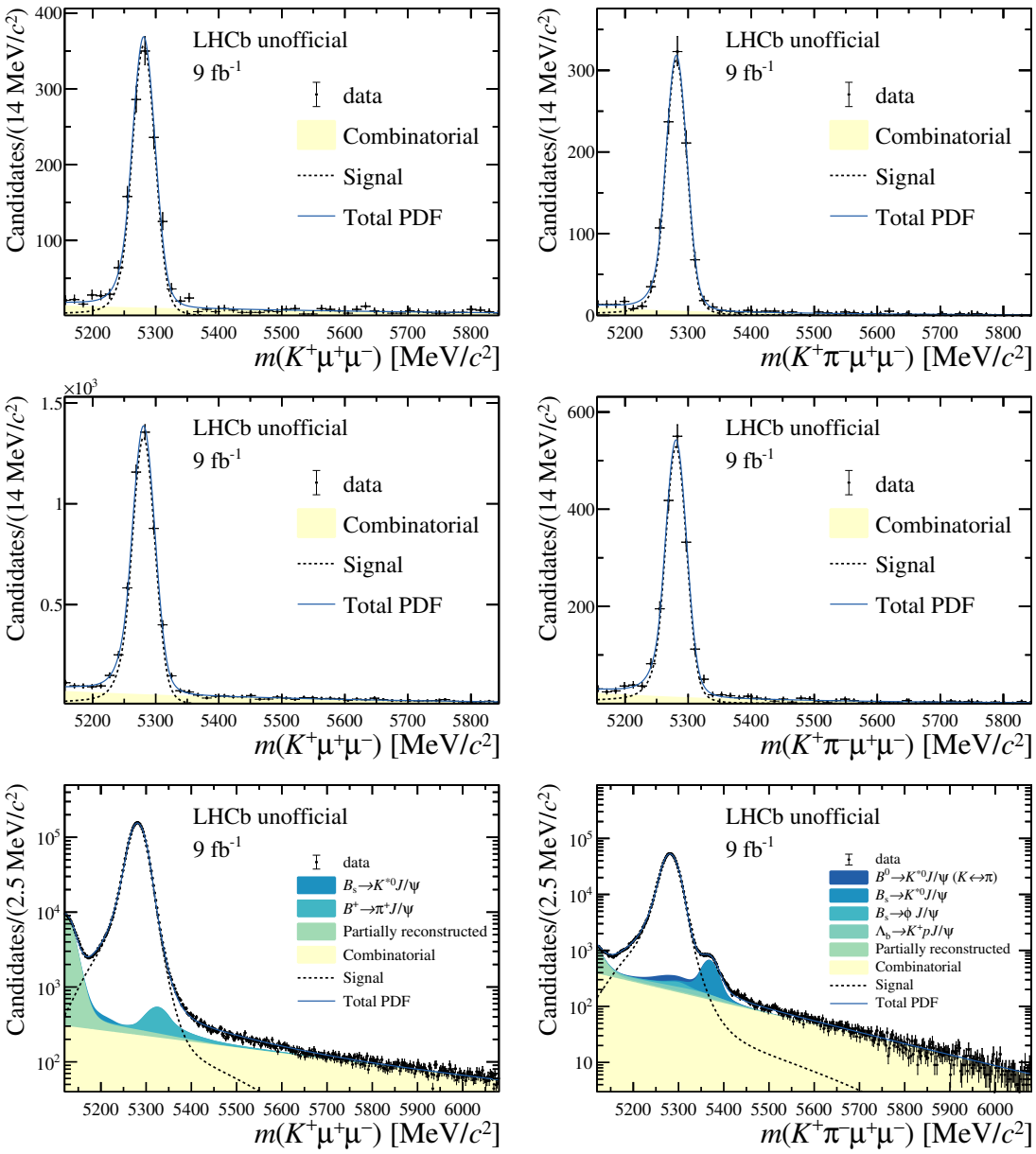

Figure 4.44.: Result from the simultaneous fit of the muon channels for R_K and $R_{K^{*0}}$ in the full Run 1 and Run 2 dataset. The plots result from the sum of data samples and PDFs over the individual run periods and trigger categories. The low, central and J/ψ q^2 regions are shown from top to bottom and the left (right) plots show the B^+ (B^0) modes. The various components used in the fit are detailed in the legends. Notice that the J/ψ resonant plots are shown with a logarithmic scale on the y-axis.

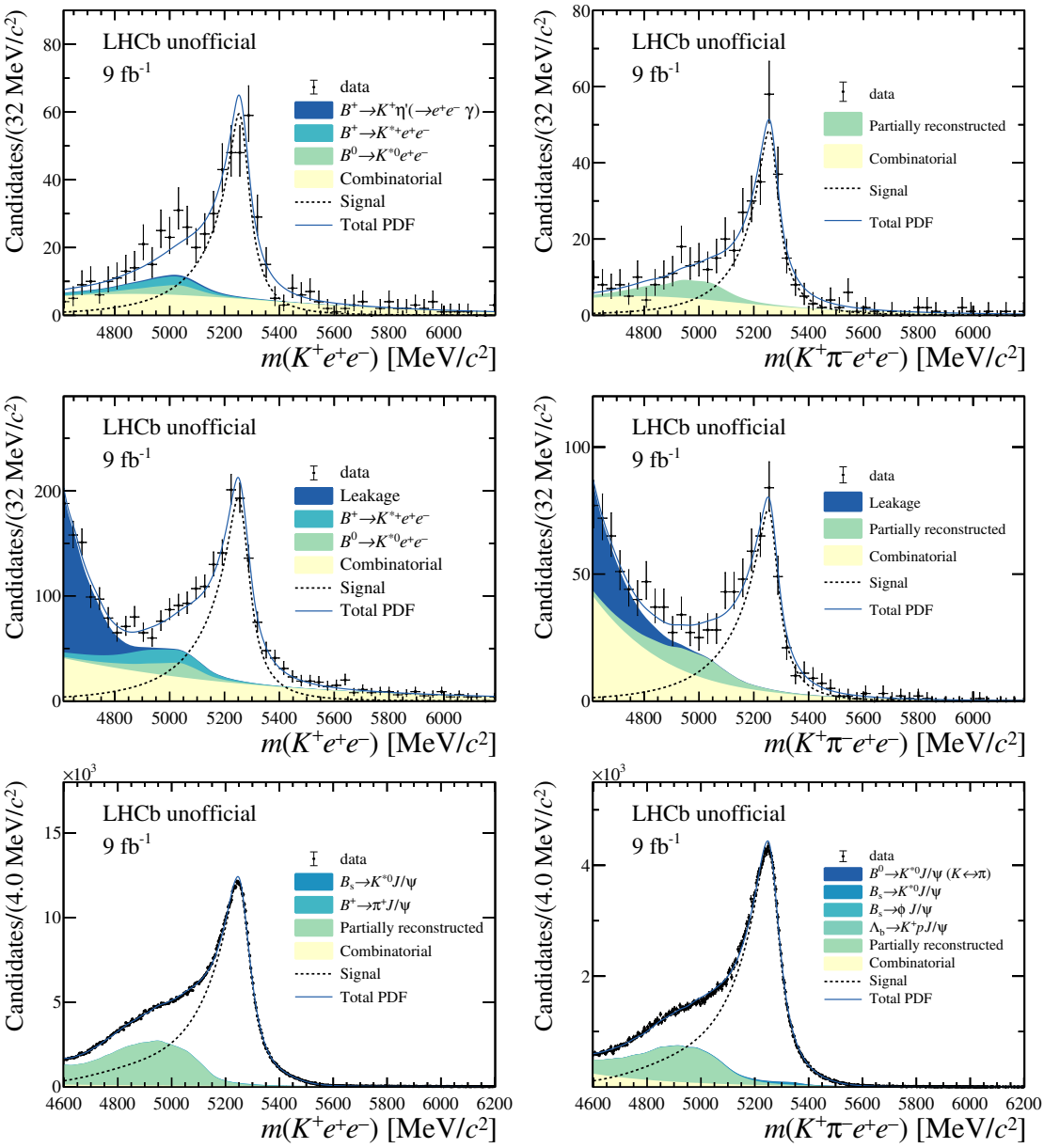

Figure 4.45.: Result from the simultaneous fit of the electron channels for R_K and $R_{K^{*0}}$ in the full RUN 1 and RUN 2 dataset. The plots result from the sum of data samples and PDFs over the individual run periods and trigger categories. The low, central and J/ψ q^2 regions are shown from top to bottom and the left (right) plots show the B^+ (B^0) modes. The various components used in the fit are detailed in the legends.

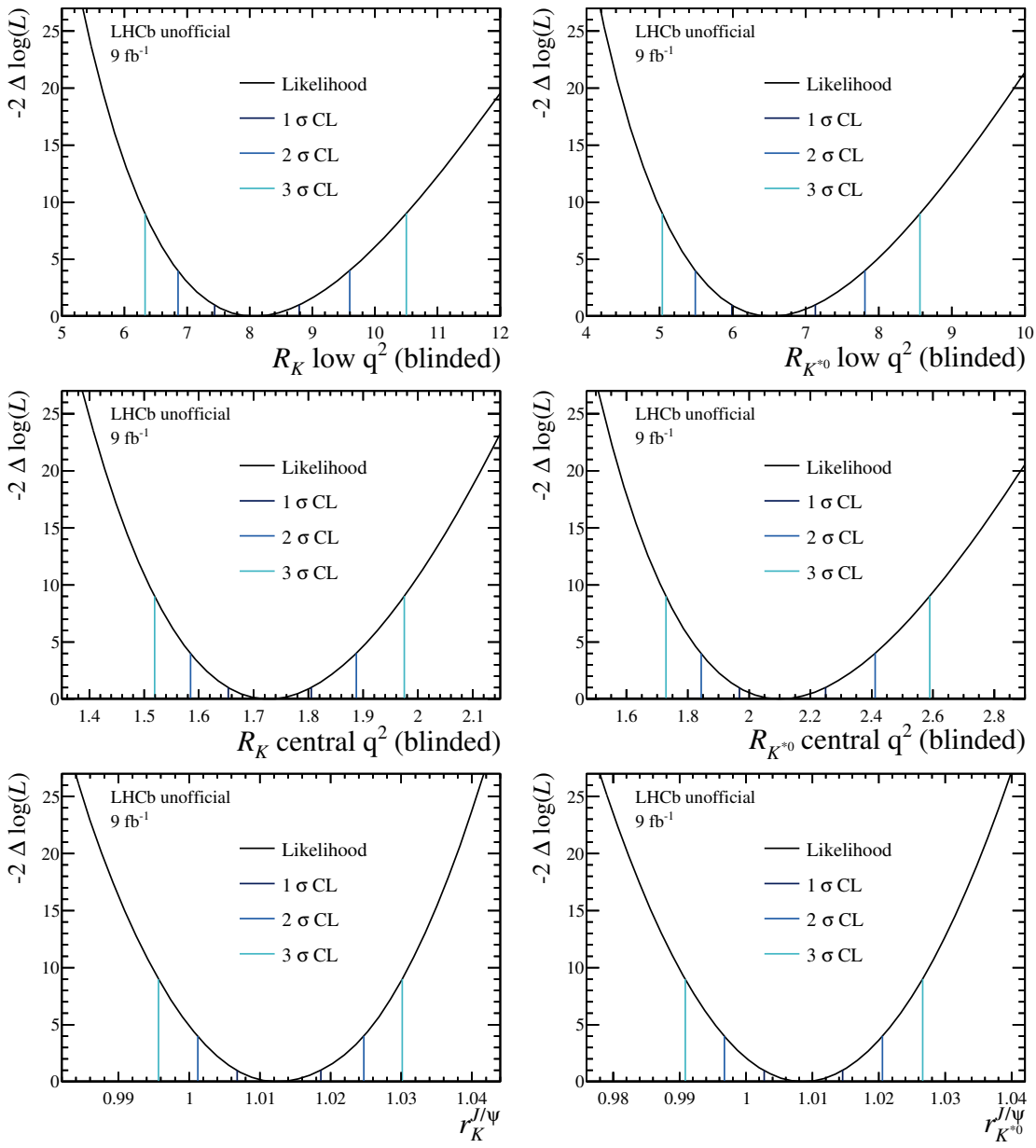

Figure 4.46.: One dimensional likelihood scans from the simultaneous fit for R_K (left) and $R_{K^{*0}}$ (right) in low (top) and central (middle) as well as for the $r^{J/\psi}$ single ratios (bottom).

Table 4.25.: Mass yields of the rare modes from the simultaneous fit in all q^2 regions, run periods and trigger categories.

Mode	q^2 region	Run period	$\mathcal{N}(\mu^+\mu^-)$ L0I	$\mathcal{N}(\mu^+\mu^-)$ L0L	$\mathcal{N}(e^+e^-)$ L0I	$\mathcal{N}(e^+e^-)$ L0L
$B^+ \to K^+\ell^+\ell^-$	low	Run 1	85.43 ± 10.70	237.88 ± 17.00	34.45 ± 7.81	45.21 ± 8.14
		Run 2p1	88.30 ± 10.30	229.98 ± 16.80	45.76 ± 9.33	41.18 ± 8.29
		Run 2p2	168.60 ± 14.20	439.78 ± 22.90	94.26 ± 13.90	85.48 ± 12.00
	central	Run 1	307.09 ± 20.20	899.23 ± 33.40	154.88 ± 17.90	150.96 ± 15.70
		Run 2p1	339.53 ± 21.10	831.12 ± 31.90	162.76 ± 18.30	175.88 ± 17.20
		Run 2p2	680.68 ± 29.70	1630.40 ± 44.70	283.63 ± 25.60	318.55 ± 23.50
$B^0 \to K^{*0}\ell^+\ell^-$	low	Run 1	83.90 ± 9.96	171.37 ± 14.00	25.82 ± 6.71	32.09 ± 6.87
		Run 2p1	99.97 ± 10.30	152.02 ± 13.10	29.43 ± 8.41	34.15 ± 7.61
		Run 2p2	153.70 ± 13.20	340.27 ± 19.50	72.41 ± 11.20	74.75 ± 10.60
	central	Run 1	114.42 ± 11.70	266.49 ± 18.10	51.93 ± 10.30	42.84 ± 7.97
		Run 2p1	154.98 ± 13.60	293.66 ± 18.60	56.27 ± 10.50	62.49 ± 10.10
		Run 2p2	296.63 ± 18.90	616.62 ± 27.00	139.34 ± 14.80	126.60 ± 13.90

4.6.5 Validation with Pseudo-experiments

In order to validate the stability of the simultaneous fit for the LFU observables and check for possible biases, a large number of pseudo-experiments are performed. The result of the converged fit to the nominal dataset is used to generate pseudo-datasets, where the yields of the separate fit components are varied according to a Poisson distribution. Subsequently, these datasets are fitted with the nominal fit model and minimisation strategy. A fit to a pseudo-dataset is classified as 'failed', if the minimisation did not converge allowing to calculate the failure rate, which is a good indicator of the fit stability. The failure rate of the fit is found to be at about 7 %, which is at an acceptable level considering the low statistics in the rare electron modes and the high complexity of the fit. Results from failed pseudo-experiments are excluded in the following studies.

The so-called 'pull' of all fit components is calculated, which is defined as the difference between the converged value of the pseudo-experiment from the generated value devided by the uncertainty of the fitted parameter. Combining the pulls of all pseudo-experiments allows to study the pull distribution. For an unbiased fit with correctly estimated uncertainties, this distributions is described by a Gaussian function centered at zero and with a width of unity. The pull distributions of the LFU ratio values are given in Fig. 4.47, while the results for other fit components such as yields of the various backgrounds are given in Ref. [93].

The results of the pull distributions show a good agreement of the pull width with unity, which confirms the coverage of the fit uncertainties. However, especially in low q^2 a small bias ($\mathcal{O}(0.2\,\sigma(\text{stat.}))$) on the LFU central values is present, which originates from the correlation of signal yields and combinatorial background in rare electron modes.

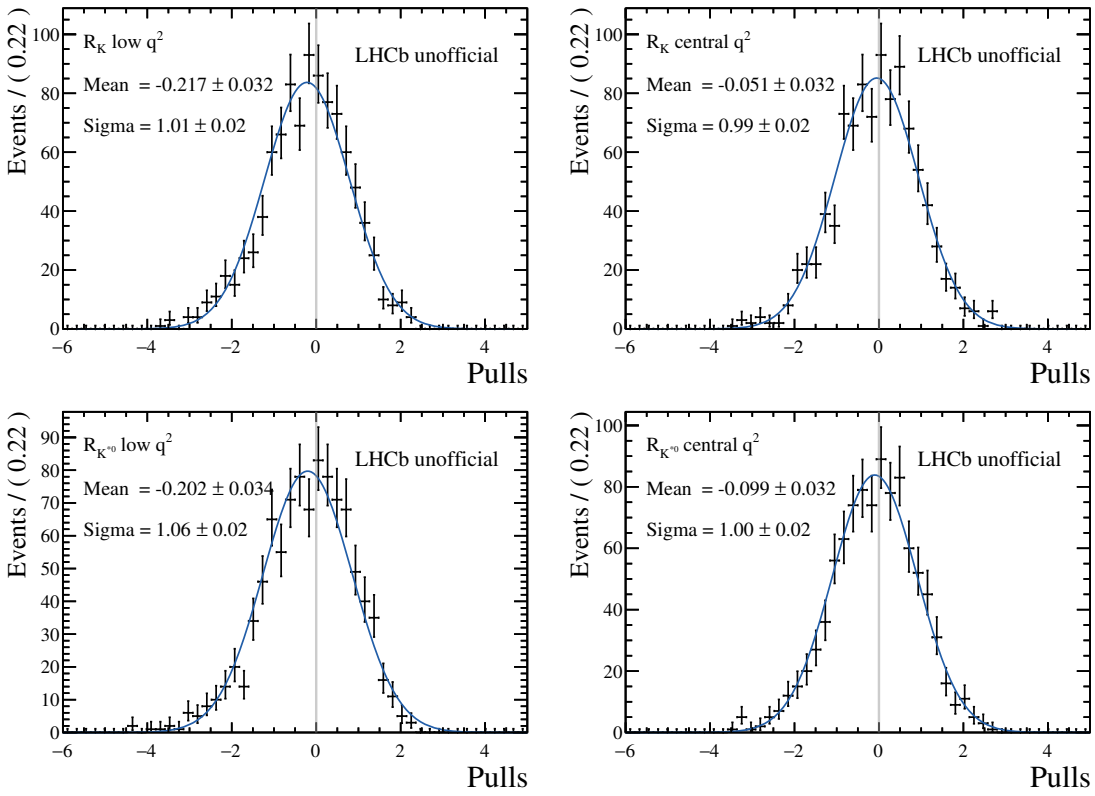

Figure 4.47.: Pull distributions of R_K (top) and $R_{K^{*0}}$ (bottom) in low (left) and central (right) q^2 calculated from 1000 pseudo-experiments. A Gaussian function is fitted to the distribution and the resulting values of mean and width are given in the plots.

After the post-unblinding studies discussed in Sec. 4.10.1 are concluded, the individual results of the four LFU ratios will be corrected for the observed bias on the central value. The final compatibility with the SM expectation is evaluated with pseudo-experiments (see Sec. 4.10.2) which correctly account for the observed biases and the non-Gaussian behaviour of the likelihood (see Sec. 4.10).

4.7 Systematic Uncertainties

Several sources of systematic uncertainties are introduced by the chosen methods and strategies used throughout the analysis. They can be divided into systematic uncertainties related to the efficiency calculation of rare and resonant modes and those associated with the simultaneous fit to the mass distributions.

4.7.1 Efficiency Calculation

The systematic uncertainties from the efficiency determination can be further divided into a statistical component from the limited size of the simulated samples used to determine the efficiencies and a component associated to chosen methods used during the generation and calibration of the simulation. The statistical component is calculated with the `bootstrapping` technique described in Sec. 4.5.3 and is thus already covered by the Gaussian-constraints on the efficiencies that are used in the direct fits for the LFU ratios. The methodological component is calculated by varying the assumptions made in the efficiency calculation and measuring the impact on the double ratios. However, due to the chosen approach of using the MC calibration calculated on B^0 samples to correct the efficiencies of B^+ modes (and vice versa), a statistical independence of the efficiencies from the fitted yield of the control mode is achieved. Thus, only the pure efficiency ratios need to be considered when calculating the systematic uncertainties. Separate uncertainties are calculated for the four LFU bins, the three run periods as well as for both L0 trigger categories. The covariances between these individual values are calculated with Eq. 4.5.3 yielding a 24×24 matrix for each systematic uncertainty. In the end, the sum over all covariance matrices yields the final matrix that is used in the mass fit.

In the following, the considered sources of systematic uncertainties for the efficiency determination are introduced and the strategies used to calculate their sizes is outlined; the more detail on the technical aspects of the procedures are given in Ref. [93]. The absolute size of each systematic as well as the total uncertainty on each LFU observable is given in Tab. 4.26.

Decay Model

The default model for both simulated $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays is the `BTOSLLBALL` 6 model [86], which uses the form factor calculations from Ref. [96] for B^+ decays and from Ref. [97] for B^0. Alternative form factors from Ref. [115] and Ref. [116] are used to evaluate the impact of the used decay model on the efficiencies. The choice of form factors impacts the distribution of both q^2 and the angles between final state particles. The nominal simulated samples are reweighted in these quantities to match the distribution of the alternative models. The efficiencies are recalculated using these weights and the systematic uncertainty is derived from the difference to the nominal efficiencies. For B^+ modes this source of uncertainty is insignificant, while it represents one of the largest systematics for B^0 modes and is at the level of 12 % of the statistical uncertainty.

PID Calibration

Several uncertainties arise from the PID calibration procedure, which is described in Sec. 4.4.1:

- **Statistics:** The uncertainty from the limited size of the data calibration samples is determined via the `bootstrapping` technique.
- **Binning:** The maps used to correct the PID response are calculated using a certain binning scheme in bins of the particle's momentum and pseudorapidity as well as of `nTracks`. Alternative schemes are used to reevaluate the PID calibration and efficiencies to extract an uncertainty.
- **Factorisation:** In the calibration process, the efficiencies of the PID cuts are evaluated separately for each particle species and the total efficiency is given by the product. This assumes that the efficiencies are uncorrelated between the individual final state particles. However, as mentioned in Sec. 4.4.1, there is a non negligible correlation between two final state electrons that are close in the ECAL. As mentioned before, two strategies are applied to reduce this effect, nonetheless a systematic uncertainty is assigned to cover residual biases. To this end, the 'true' PID defined as

  $$\epsilon_{\text{PID}}^{\text{true}} = \frac{N(\text{All PID cuts})}{N(\text{total})}$$

  and the factorised PID as used in the nominal approach

  $$\epsilon_{\text{PID}} = \prod_p \frac{N(p \text{ PID cut})}{N(\text{total})}$$

  are calculated on uncorrected simulated samples, where in the second formula the product of all final state particles p is used. The relative difference between these approaches is used to estimate the factorisation bias. This strategy slightly overestimates the factorisation bias by design, because it also calculates a bias from kinematic correlations between the final state particles that are actually properly accounted for by the applied parameterisation in p_{T} and η. Still, this uncertainty is small compared to the total systematic, as can be seen in Tab. 4.26. The combined PID related systematic is less than 6 % of the statistical uncertainty in all cases.

Kinematic and Multiplicity

The data and MC differences of the B kinematics and event multiplicity are corrected with a reweighting method based on samples of the muon resonant channels selected with the inclusive `L0Muon` trigger line (see Sec. 4.4.3). To evaluate the systematic uncertainty coming from the choice of this calibration sample, the inclusive L0I trigger category is used instead and new efficiencies based on the alternative set of weights are calculated. Moreover, a systematic uncertainty on the chosen multiplicity variable (`nTracks`) used to perform the reweighting is calculated by redoing the reweighting with the number of reconstructed tracks in the VELO (`nVeloTracks`). The resulting systematic is about 5–8 % relative to the statistical uncertainty, where the largest part originates from the multiplicity related uncertainty.

Trigger Calibration

Several uncertainties arise from the L0 and HLT trigger calibration procedure, which are described in Secs. 4.4.4 and 4.4.5, respectively:

- **Binning:** Similarly to the PID corrections, also the trigger calibration depends on the binning used for the maps and a systematic uncertainty is calculated in a analogous way.
- **Factorisation:** Another similarity to the PID calibration arises from the assumption that the trigger response of the final state particles is uncorrelated. Since the L0L category is defined via the logical 'OR' condition between the trigger responses of the two signal leptons, it is possible that a correlation between the individual trigger responses is introduced. A systematic uncertainty is calculated by reparameterising the correction maps of the L0E category with $\max(E_T^{\mathrm{ECAL}}(e^+), E_T^{\mathrm{ECAL}}(e^-))$ to extract the efficiency of the dielectron system directly. This effect is found to be negligible for the `L0Muon` category so that no uncertainty is assigned for dimuon modes.
- **L0I:** The correction maps between the muon and electron L0I category are found to be compatible motivating to port this correction from the muon to the electron mode in order to benefit from the higher statistics. A systematic uncertainty is assigned, by using the electron maps for all electron channels and reevaluating the efficiencies.
- **HLT parameterisation:** The baseline variable used to parameterise the HLT correction maps is `nTracks`. To assess a systematic uncertainty, the corrections are reevaluated using $p_{\mathrm{T}}(B)$ instead.

The combined systematic of the trigger calibration is of subordinate size compared to other systematic uncertainties and at the level of 2–4 % of the statistical uncertainty.

Smearing of the q^2 System

Differences between data and MC in the resolution of the electron q^2 distribution are corrected with a smearing method based on fits to the $m(e^+e^-)$ system as discussed in Sec. 4.4.7. Several systematic uncertainties are calculated on the used procedure:

- **Fit model:** An alternative fit model is used to describe the $m(e^+e^-)$ distribution on both data and simulation.
- **Selection:** A tight cut on the reconstructed B mass with a J/ψ mass constraint is used to reduce the background level on data. To assess a systematic uncertainty, the chosen cut window is varied in order to study the dependence on the background level.
- **L0 categories:** In the nominal approach, the resulting smearing parameters are combined between the two L0 trigger categories. An uncertainty is assigned by

using the individual values to extract the efficiency correction separately for L0I and L0L.

These variations yield new values for the smearing parameters σ_{MC} and μ_{MC}, which are used to reevaluate the efficiencies. This systematic is rather small with a size of 3–5 % of the statistical uncertainty.

HOP Selection

The HOP mass veto is only applied to the rare electron mode, which is why a bias from potential data and MC differences does not cancel in the ratios. Thus, a systematic uncertainty is assigned by calculating the difference on the double ratio when the HOP cut is applied to the electron J/ψ mode as well. The resulting uncertainty has a size of 4–6 % of the statistical uncertainty.

Stability of LFU Ratios

Residual differences between data and MC in both rare and resonant modes can lead to a non-flatness of the observed ratio, *i.e.* a trend of the ratio when calculated as a function of a variable such as the transverse momentum of the B meson. For the J/ψ modes, the flatness of $r^{J/\psi}$ is tested by evaluating the ratio in equi-populated bins of many key variables. For example, this is shown for Run 2p2 samples in App. E before and after applying the calibration to simulation. The result of this check shows that the flatness is greatly improved after applying the correction chain, however residual differences, especially in some multiplicity observables, remain. If the rare mode shows a similar trend to the resonant mode, the bias from this non-flatness will cancel in the double ratios. In order to assign a systematic uncertainty, it is therefore necessary to determine the bias from the non-flatness on the double ratio itself. Due to their low statistics, it is not feasible to study the rare mode yields split into bins, which is why the so-called flatness parameter d_f is calculated:

$$d_f = \frac{\sum_i^n \epsilon^i_{signal,\mu} \cdot \mathcal{Y}^i_\mu}{\sum_i^n \epsilon^i_{signal,\mu}} \cdot \frac{\sum_i^n \epsilon^i_{J/\psi,\mu}}{\sum_i^n N^i_\mu} \Bigg/ \frac{\sum_i^n \epsilon^i_{signal,e} \cdot \mathcal{Y}^i_e}{\sum_i^n \epsilon^i_{signal,e}} \cdot \frac{\sum_i^n \epsilon^i_{J/\psi,e}}{\sum_i^n N^i_e} - 1\,, \tag{4.29}$$

where n is the number of bins, $\epsilon^i_{signal,J/\psi}$ are the efficiencies of rare and resonant mode in bin i, N^i_ℓ are the resonant mode yields in bin i and $\mathcal{Y}^i_\ell$ refers to the efficiency corrected resonant yield in bin i ($\mathcal{Y}^i_\ell = {}^{N^i_\ell}/_{\epsilon^i_{J/\psi,\ell}}$). Essentially, the d_f parameter describes an artificial double ratio, where the rare mode yields are substituted by the resonant yields scaled by the efficiency ratio of rare and resonant mode. This parameter is calculated for each studied variable and the one yielding the largest values is used as a systematic uncertainty. Appendix F shows the values of d_f for several key variables. For most LFU ratios this uncertainty represents the largest systematic related to efficiencies. The size corresponds to about 24 % of the statistical uncertainty for the $R_{K^{*0}}$ measurement in low q^2, while it is at the order of 10 % for the other ratios.

4.7.2 Simultaneous Fit

In the case of systematic uncertainties related the fit model of the rare modes, an approach based on pseudo-experiments (as introduced in Sec. 4.6.5) is used. To this end, each pseudo-experiment is fitted twice, once with the nominal fit model and once with an alternative model used to evaluate the given systematic uncertainty. The difference of the studied LFU ratio between the alternative (R_X^{alt}) and the nominal (R_X^{nom}) fit model is fitted with a Gaussian function as shown in Fig. 4.48. The systematic uncertainty is derived from the values of mean and width of the Gaussian.

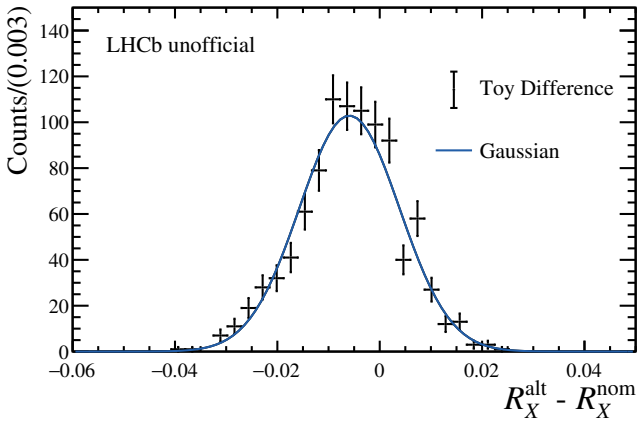

Figure 4.48.: Example plot of the LFU ratio difference from pseudo-experiments fitted with the nominal and an alternative model.

Similarly to the efficiency related systematic uncertainties, a separate evaluation of the uncertainty for each LFU bin, run period and L0 category is performed. However, since the uncertainties are derived from the same simultaneous fit, possible correlations between them are properly accounted for. In the following the considered sources of systematic uncertainties related to the simultaneous fit are introduced.

Fixed Parameters

Various parameters are fixed, rather than constrained, in the simultaneous fit. In order to compute an uncertainty, the pseudo-experiments are fitted with a fit model, where these parameters are shifted by their uncertainties in both directions ($\pm 1\,\sigma$). The largest contribution to this systematic is due to the parameters of the turn-on function used to describe the combinatorial background in rare electron modes. Because of the low statistics in the SS data samples used to fit the turn-on shape (see Sec. 4.6.2), the uncertainty of these parameters is rather large. For the low q^2 LFU ratios the limited knowledge of these parameters and the resulting differences of the combinatorial shape represent the dominant source of systematic uncertainty and is 27 % (25 %) of the statistical uncertainty for R_K ($R_{K^{*0}}$). This contribution is less important for central q^2 with 5 % for R_K and 9 % for $R_{K^{*0}}$.

Next, the fractions of the relative yields between the three separate bremsstrahlung categories used to model the signal shape of electron modes are investigated. These fractions are taken from simulation and are fixed in the data fit. A systematic uncertainty is assigned, by varying the relative fractions by $\pm 1\,\%$ in the final signal model for all electron modes, which covers well the observed differences of these parameters between data and simulated samples of the J/ψ resonant modes. The size of this contribution is below $4\,\%$ of the statistical uncertainty for all LFU ratios.

Lastly, the yields of the partially reconstructed backgrounds in $B^+ \to K^+ e^+ e^-$ modes are constrained to the signal yield of the $B^0 \to K^{*0} e^+ e^-$ mode using the strategy described in Sec. 4.6.4. Various parameters enter into the calculation for these constraints such as the isospin branching fraction factor f_{iso} and the branching fractions of the S-wave component. The parameters are varied by a conservative uncertainty (up to $\pm 50\,\%$) and the constraint is recalculated and used in pseudo-experiments. This source of systematic uncertainty makes a rather small contribution of about $1--2\,\%$ of the statistical uncertainty for all LFU ratios.

In total, the resulting systematic due to fixed parameters is at the level of 25–27 % and 6–9 % of the statistical uncertainty for the ratios in low and central q^2, respectively.

Physical Backgrounds

Another uncertainty arises from the modelling and constraining of the partially reconstructed backgrounds in fits to the rare electron channels. For B^0 modes, this background type is modelled with simulated samples of $B^+ \to K^+\pi^+\pi^- e^+ e^-$ decays, where a phase-space model is used to generate the hadronic system. These samples are reweighted to match the resonant structures observed in data (see Sec. 4.2.4). A systematic uncertainty is derived, by modelling this background component in pseudo-experiments with a mixture of simulated samples from K_1^+ and K_2^{*+} resonant decays that are generated with a full amplitude model.

The parameters that enter the constraint of the partially reconstructed backgrounds in $B^+ \to K^+ e^+ e^-$ modes (that were already considered for the systematic uncertainty due to fixed fit parameters) can also affect the mass shape of this background component. The parameters are varied by $\pm 1\,\sigma$ of their associated uncertainties yielding an alternative fit model which is used in pseudo-experiments.

Since the yield of the partially reconstructed background is floating freely in $B^0 \to K^{*0} e^+ e^-$ modes, a change of the mass shape of the background component has a bigger impact on the fit with respect to B^+ modes. Consequently, the systematic uncertainty is larger for $R_{K^{*0}}$ observables with a size of 16 % for low and 8 % for central q^2, while for R_K ratios this source of systematic is less important and does not exceed 6 %.

Fit to the J/ψ Modes

Also in the J/ψ mode, the signal shape as well as the composition and modelling of backgrounds is not described perfectly. Here, the largest uncertainty is introduced by 'charmonium' partially reconstructed backgrounds (as defined in Sec. 4.6.3) in the electron modes. Because of the high statistics of the J/ψ mode, even systematic uncertainties below the percent level can make a significant contribution to the total uncertainty budget.

The systematic uncertainty related to the fit of the J/ψ resonant modes is not computed via pseudo-experiments. Instead, the following four different fit setups are studied and evaluated:

1. The nominal J/ψ mode fits used to normalise the LFU ratios as described in Sec. 4.6.
2. The fits to the invariant B mass with a mass constraint of the dilepton system to the J/ψ meson mass that are used for the $r^{J/\psi}$ and $R^{\psi(2S)}$ cross-checks. These fits are given in App. D.
3. Another fit to the unconstrained B mass is performed, where the partially reconstructed backgrounds are suppressed by requiring an additional cut on the J/ψ constrained B mass ($m^{J/\psi}(B) > 5200\,\text{MeV}$).
4. The J/ψ mass-constrained B system is fitted in a wider mass range in the electron modes (4650–6200 MeV) in order to further test the modelling of the partially reconstructed backgrounds.

For the last two setups, the yields resulting from the fits are given in App. D. The difference of the $r^{J/\psi}$ values between the nominal unconstrained (setup 1) and constrained (setup 2) fits is calculated. Both the modelling of the signal shapes and the correlation of the signal component to the various backgrounds that are modelled in the fits behave very differently in the two setups, since the mass constraint applied in setup 2 greatly improves the mass resolution, especially for electron modes. Thus, the difference between setups 1 and 2 is a good measure for the uncertainty on the signal shape and the interplay with background components. Further, the difference between the two alternative setups (3 and 4) is computed to account for the mismodelling of partially reconstructed backgrounds. These two differences are added in quadrature yielding the final systematic uncertainty assigned to the fit of the J/ψ modes. The resulting systematic is at the level of 5–9 % of the statistical uncertainty for all LFU ratios.

Residual Backgrounds

Lastly, a systematic uncertainty is assigned to residual backgrounds in fits to the rare electron modes originating from decays such as $B^+ \to K^+\pi^+\pi^-$ and $B^0 \to K^{*0}\pi^+\pi^-$ for B^+ and B^0 channels, respectively. Since a double misidentification ($\pi \to e$) is required to reconstruct these modes as signal candidates, this source of background is strongly suppressed by the PID requirements. Nonetheless, a significant contamination from these

modes can occur due to their large branching ratios compared to the signal channels. The level of contamination is determined in a data-driven way: The mass hypothesis of a pion is assigned to both electrons of the signal candidates, which yields a mass system in which these background modes peak at the B mass. Further, the PID selection criteria on the electrons are inverted to increase the statistics of this background source. The resulting dataset is fitted in order to compute the yield of this background. Finally, these yields are corrected for the misidentification rate given the full nominal PID selection, which is extracted from calibration samples. This systematic is one of the dominant sources for the R_K measurement in low q^2 with a size of 21 % of the statistical uncertainty. For the other ratios it is of moderate size with 6–11 %.

However, tests performed post-unblinding have shown that these modes might have been underestimated and may lead to noticeable contamination in the data samples after the full selection. This cross-check as well as the studies that are currently performed are detailed in Sec. 4.10.1.

4.7.3 Results

The resulting central values for all sources of systematic uncertainty are quoted in Tab. 4.26, while the full correlation matrices are given in Ref. [93].

In order to calculate the impact of the systematic uncertainties on the final measurements, the combined covariance matrix, resulting from the sum over all individual sources, is used in the simultaneous mass fit. This is achieved by adding another factor to the double ratio defined in Eq. 4.6, which is implemented as a 24×24 dimensional Gaussian-constraint of unity mean and the width is given by the covariance matrix of the total systematic uncertainty.

Table 4.26.: Results for the individual sources of systematic uncertainties on the R_K and $R_{K^{*0}}$ ratios in the low and central q^2. The values are derived from the weighted averages of central values from all run periods and L0 trigger categories (ignoring correlations). The total uncertainty follows from the quadratic sum over all sources in a given q^2 bin. For comparison, the statistical uncertainties from the invariant mass fit is given in the last row. All values are given in percent relative to the measured central value.

	Relative systematic uncertainty [%]			
Source	R_K low q^2	R_K central q^2	$R_{K^{*0}}$ low q^2	$R_{K^{*0}}$ central q^2
Decay model	0.09	0.08	0.83	0.76
Particle identification	0.17	0.22	0.10	0.12
Kinematics and multiplicity	0.35	0.26	0.57	0.52
Trigger	0.27	0.16	0.26	0.13
q^2 smearing	0.30	0.19	0.28	0.31
HOP selection	0.25	0.24	0.33	0.33
Stability of double ratios	0.78	0.38	1.79	0.47
Fixed fit parameters	1.87	0.24	1.88	0.60
Physical backgrounds	0.24	0.20	1.24	0.51
J/ψ fit model	0.35	0.35	0.40	0.40
Residual backgrounds	1.50	0.30	0.80	0.40
Total syst. uncertainty	2.63	0.83	3.22	1.57
Total stat. uncertainty	6.99	3.88	7.53	6.22

4.8 Cross-checks

This section describes the various cross-checks that are performed to test both efficiencies and their calibration as well as the invariant mass fits used to extract the signal yields. First, the integrated $r^{J/\psi}$ and $R^{\psi(2S)}$ ratios are discussed. Next, scans of the logarithmic likelihood are performed for both single and double ratio cross-checks including the full set of systematic uncertainties from Sec. 4.7. Additionally, the $r^{J/\psi}$ ratios are calculated differentially in bins of various key variables in App. E. Lastly, in App. G the $R^{\psi(2S)}$ ratio is fitted using a wider range for the q^2 window in the $\psi(2S)(e^+e^-)$ channel.

4.8.1 $r^{J/\psi}$ and $R^{\psi(2S)}$ Ratios

Two of the most important cross-checks of the analysis are the $r^{J/\psi}$ single and the $R^{\psi(2S)}$ double ratios defined in Eq. 4.3 and Eq. 4.7, respectively. Both are determined from a simultaneous fit to the J/ψ and $\psi(2S)$ resonant modes that is described in App. D. Since the $r^{J/\psi}$ ratios are defined as single ratios, there is no cancellation of systematic uncertainties from the analysis procedure related to a specific lepton species. For this reason, systematic uncertainties will have a larger impact on these ratios. Further, possible biases that cancel in double ratios persist in the $r^{J/\psi}$ test and can lead to a deviation from unity making this a very stringent cross-check for both efficiencies and their calibration as well as for the J/ψ mass fits. The $R^{\psi(2S)}$ ratios are a very useful tool to test the double ratio approach and to verify the cancellation of systematic uncertainties. Additionally, this cross-check is used to test the portability of the MC corrections, that are extracted from calibration samples of the J/ψ resonant modes, to other regions of q^2.

The systematic uncertainties, that are discussed in Sec. 4.7, are not included in this section but will enter in the next section. Here, only the statistical uncertainties on the yields from the invariant mass fits and those which follow from the `bootstrapping` method are considered. The values for both $r^{J/\psi}$ and $R^{\psi(2S)}$ and their uncertainties are given in Tab. 4.27. Even without including the systematic uncertainties, a good agreement with unity is visible throughout all run periods and trigger categories. Comparing the three types of uncertainties, the expected behaviour is present: The $r^{J/\psi}$ single ratio is dominated by the `bootstrapping` uncertainty from the simulation calibration, while this uncertainty nicely cancels in the double ratio. Because of the much smaller branching ratio of the $\psi(2S)$ resonant channels with respect to J/ψ, the uncertainty on the $R^{\psi(2S)}$ ratio is dominated by the yields of the $\psi(2S)$ modes.

In Figs. 4.49 and 4.50, the ratios are shown for each consecutive step of the simulation calibration procedure to demonstrate the impact of the efficiency corrections. The clear trend of the $r^{J/\psi}$ single ratios is driven by the large impact of the calibration chain on electron efficiencies. Since such inaccuracies do not cancel in the single ratio, the good agreement of the final value with the theory prediction of unity is a very powerful indicator of the control over the efficiencies in this analysis. In contrast, the $R^{\psi(2S)}$ ratios show a nearly constant behaviour, which demonstrates the cancellation of biases as well as the robustness of the double ratios. Further, the good portability between the MC corrections obtained from B^+ and B^0 calibration samples is underlined by the compatibility of the results when interchanging the $w(B^+)$ and $w(B^0)$ correction chains as shown in the plots.

4.8.2 $r^{J/\psi}$ and $R^{\psi(2S)}$ with Systematic Uncertainties

In this section, the full set of systematic uncertainties are used in the simultaneous fit (as described in Sec. 4.7) for the ratios of the resonant modes. In this section, the results of $r^{J/\psi}$ are extracted from a fit to the J/ψ modes only. This allows to use a dedicated set of systematic uncertainties, that are evaluated on the single ratio only. Consequently, these

Table 4.27.: Values of the $r^{J/\psi}$ and $R^{\psi(2S)}$ ratios after applying the full MC calibration. The baseline correction chain is applied *i.e.* $w(B^0)$ for B^+ modes and $w(B^+)$ for B^0 modes. The three uncertainties (from left to right) are defined as: statistical on the yields from the mass fits; statistical on efficiencies from the finite MC sample sizes; the `bootstrapping` uncertainty from the simulation calibrations.

Category	$r_K^{J/\psi}$	$r_{K^{*0}}^{J/\psi}$
Run 1 L0I	$1.043 \pm 0.005 \pm 0.003 \pm 0.028$	$1.027 \pm 0.010 \pm 0.004 \pm 0.016$
Run 1 L0L	$1.054 \pm 0.004 \pm 0.002 \pm 0.025$	$1.033 \pm 0.008 \pm 0.003 \pm 0.018$
Run 2p1 L0I	$1.025 \pm 0.004 \pm 0.002 \pm 0.019$	$0.986 \pm 0.008 \pm 0.004 \pm 0.010$
Run 2p1 L0L	$1.050 \pm 0.004 \pm 0.002 \pm 0.012$	$1.022 \pm 0.007 \pm 0.004 \pm 0.010$
Run 2p2 L0I	$1.002 \pm 0.003 \pm 0.003 \pm 0.012$	$0.991 \pm 0.006 \pm 0.004 \pm 0.007$
Run 2p2 L0L	$0.992 \pm 0.003 \pm 0.002 \pm 0.008$	$1.008 \pm 0.005 \pm 0.004 \pm 0.006$
	$R_K^{\psi(2S)}$	$R_{K^{*0}}^{\psi(2S)}$
Run 1 L0I	$0.997 \pm 0.021 \pm 0.005 \pm 0.002$	$1.051 \pm 0.045 \pm 0.009 \pm 0.001$
Run 1 L0L	$0.981 \pm 0.016 \pm 0.004 \pm 0.003$	$0.987 \pm 0.033 \pm 0.007 \pm 0.003$
Run 2p1 L0I	$0.949 \pm 0.017 \pm 0.004 \pm 0.001$	$1.032 \pm 0.039 \pm 0.008 \pm 0.003$
Run 2p1 L0L	$0.987 \pm 0.014 \pm 0.003 \pm 0.003$	$0.990 \pm 0.029 \pm 0.006 \pm 0.004$
Run 2p2 L0I	$0.996 \pm 0.013 \pm 0.004 \pm 0.001$	$0.957 \pm 0.025 \pm 0.006 \pm 0.001$
Run 2p2 L0L	$1.000 \pm 0.010 \pm 0.003 \pm 0.002$	$1.057 \pm 0.023 \pm 0.006 \pm 0.002$

uncertainties are expected to be significantly larger compared to those used in the fit for $R^{\psi(2S)}$, since there is no cancellation in a double ratio. The full covariance matrices from systematic uncertainties used in these fits are given in Ref. [93]. A simultaneous fit to the full dataset of both B^+ and B^0 channels is performed and one- and two-dimensional profiles of the likelihood are calculated. The result for both $r^{J/\psi}$ and $R^{\psi(2S)}$ are shown in Figs. 4.51 and 4.52, respectively, and are both compatible with unity well within the $2\,\sigma$ area. The final values for the single ratios are $r_K^{J/\psi} = 1.042 \pm 0.025$ and $r_{K^{*0}}^{J/\psi} = 1.022 \pm 0.024$ and for the double ratios $R_K^{\psi(2S)} = 0.989 \pm 0.008$ and $R_{K^{*0}}^{\psi(2S)} = 1.014 \pm 0.014$, where the uncertainties include both statistical and systematic uncertainties.

The asymmetric behaviour of the two-dimensional likelihood contours shows that the $r^{J/\psi}$ values between B^+ and B^0 modes are clearly correlated. The correlations originate from the systematic uncertainties that dominate the single ratios. As expected, the $R^{\psi(2S)}$ scans are very symmetrical, since these ratios are driven by statistical uncertainties and the main part of systematic uncertainties cancels in the double ratio.

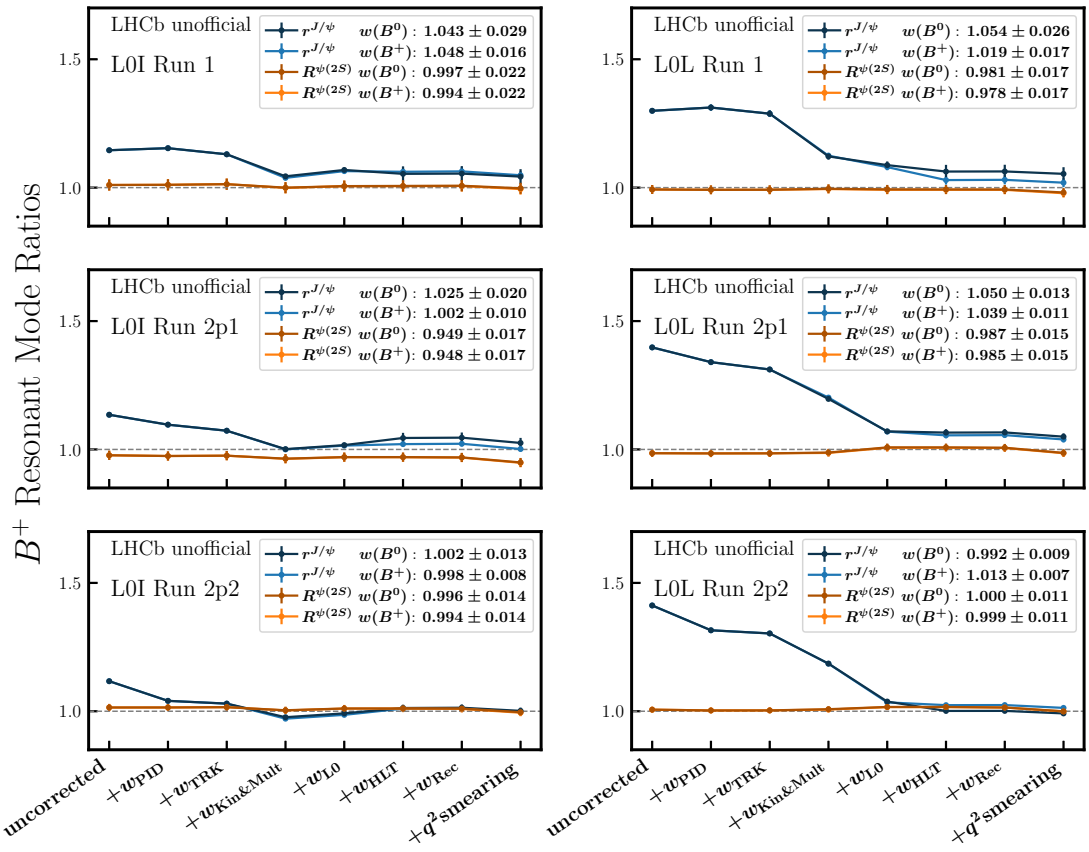

Figure 4.49.: Evolution of the $r^{J/\psi}$ single and the $R^{\psi(2S)}$ double ratios of B^+ modes with each step of the MC calibration chain as indicated by the x-axis. The left (right) plots show the L0I (L0L) trigger category and the three run periods are shown from top to bottom.

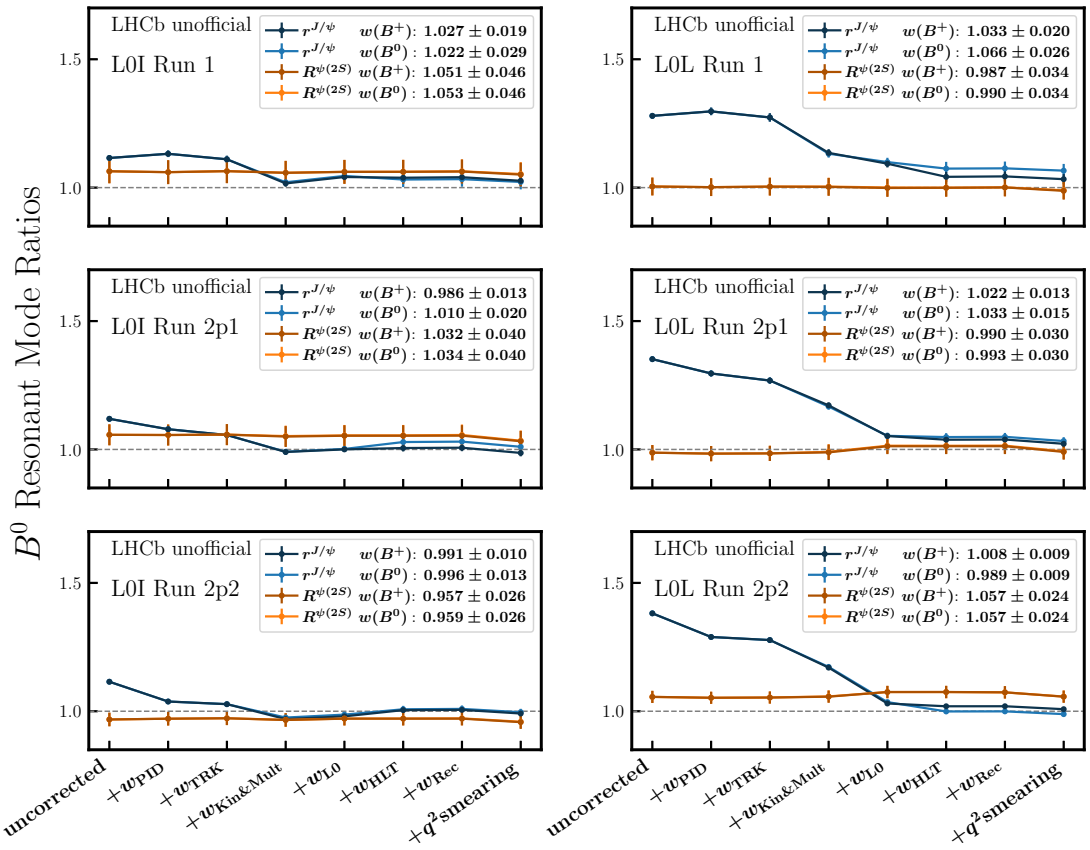

Figure 4.50.: Evolution of the $r^{J/\psi}$ single and the $R^{\psi(2S)}$ double ratios of B^0 modes with each step of the MC calibration chain as indicated by the x-axis. The left (right) plots show the L0I (L0L) trigger category and the three run periods are shown from top to bottom.

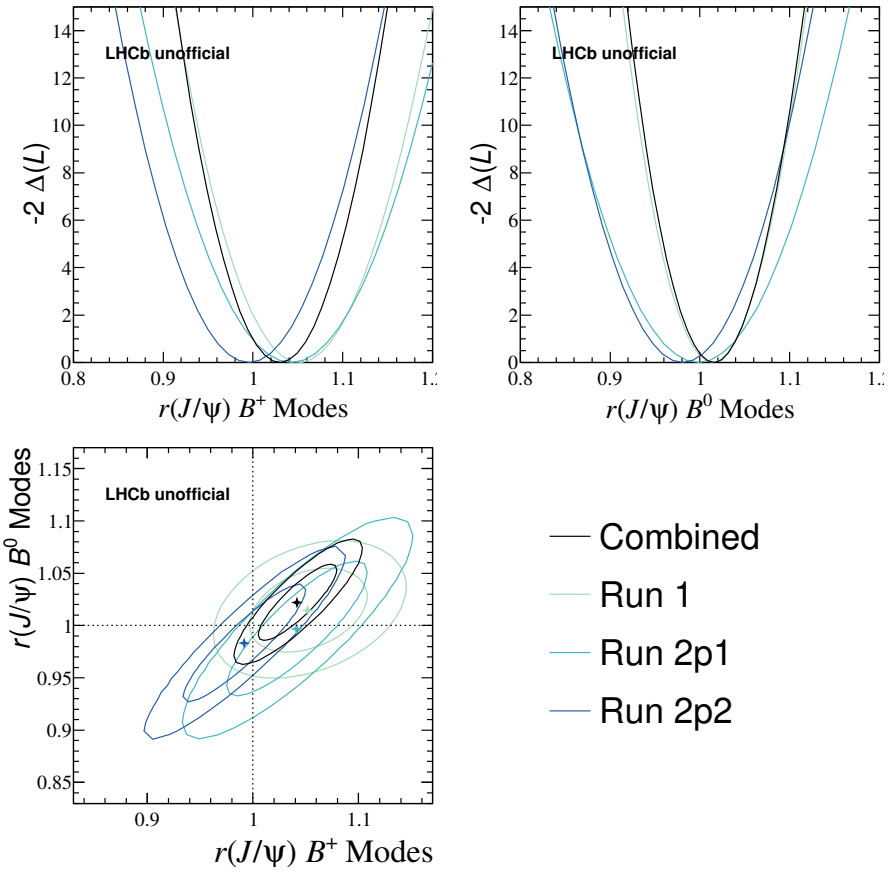

Figure 4.51.: $r^{J/\psi}$ results from a simultaneous fit of the B^+ and B^0 resonant channels to the full $9\,\text{fb}^{-1}$ dataset including the complete set of systematic uncertainties. The left (right) plot shows a one-dimensional scan over the logarithmic likelihood for the B^+ (B^0) modes. The bottom plot shows the two-dimensional scan with the 1 and $2\,\sigma$ contours.

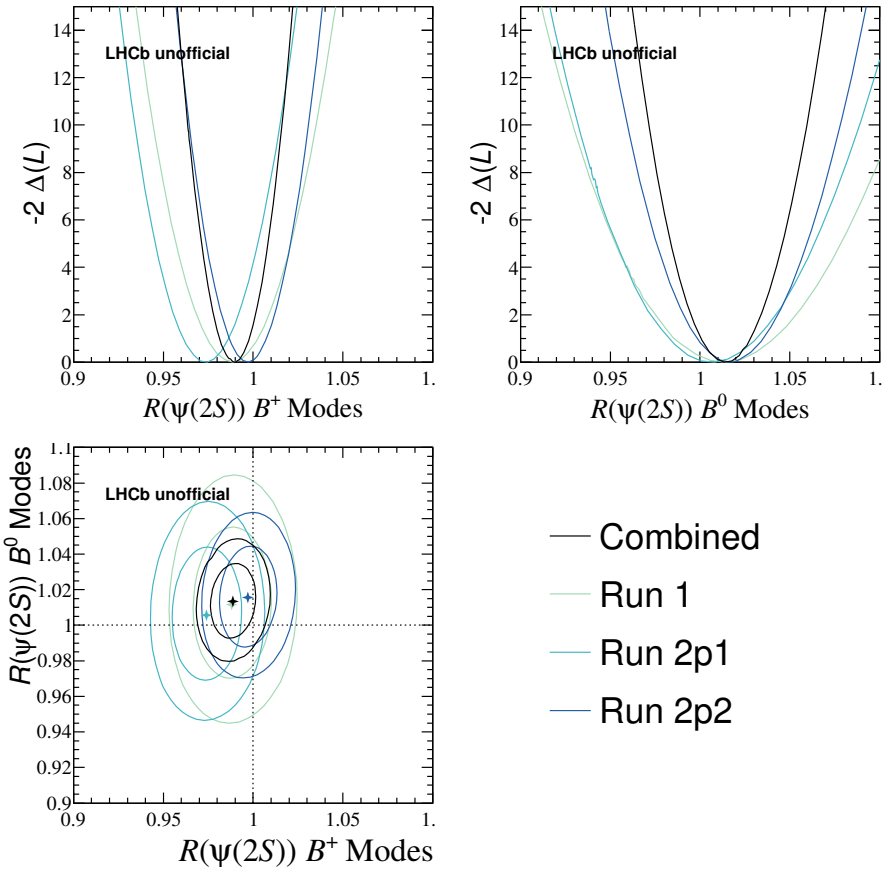

Figure 4.52.: $R^{\psi(2S)}$ results from a simultaneous fit of the B^+ and B^0 resonant channels to the full $9\,\text{fb}^{-1}$ dataset including the complete set of systematic uncertainties. The left (right) plot shows a one-dimensional scan over the logarithmic likelihood for the B^+ (B^0) modes. The bottom plot shows the two-dimensional scan with the 1 and $2\,\sigma$ contours.

4.9 Unblinding

The final unblinding of the LFU observables was dependent on successfully passing several checks. These pre-unblinding checks were established together with the review committee of this analysis and are as follows:

1. The agreement of each LFU observable across the individual run periods and trigger categories is evaluated in a blinded manner in order to test the self-consistency of the analysis.
2. Measurement of the branching ratios of the rare muon channels in RUN 1 and RUN 2 and comparison with the published values.
3. Measurement of R_K (for central q^2 only) and comparison with the value published in Ref. [13].

When comparing to other analyses, the presented analysis is referred to as 'R_X'. In the following, the passing criteria and results of these checks are documented.

4.9.1 Self-Consistency

The consistency of the LFU ratios across the three run periods and two L0 trigger categories is evaluated by comparing the difference of the logarithmic likelihood ($\Delta \log(\mathcal{L})$) of the simultaneous fit to the full $9\,\text{fb}^{-1}$ dataset between the following two setups:

1. **Nominal:** The nominal setup of the simultaneous fit is used to fit a single R_K and $R_{K^{*0}}$ value in both low and central q^2. Together with the two values for $r^{J/\psi}$ (one for B^+ and one for B^0 modes) this yields a total amount of 6 ratio parameters in the fit.
2. **Splitted:** A similar setup is used allowing to fit a separate ratio value for each run period and trigger category. Thus, a total of 24 double ratio (4 LFU bins fitted on 6 separate datasets) and 12 single ratio (2 $r^{J/\psi}$ values fitted on 6 separate datasets) parameters are minimised in the fit.

The result of the fit to data with the nominal setup is used to generate pseudo-datasets, which are subsequently fitted with both the nominal and splitted setup. The likelihood difference $|\log(\mathcal{L}_{\text{Splitted}}) - \log(\mathcal{L}_{\text{Nominal}})|$ is calculated for each pseudo-experiment. The resulting distribution of this test statistic represents a good proxy for the expected deviation due to statistical effects. The self-consistency test is classified as passed, if the likelihood difference observed from the fits to data is within the $2.5\,\sigma$ area of the distribution from pseudo-experiments. The result is given in Fig. 4.53 showing that about 13 % of pseudo-experiments yield a larger deviation than the data fits so that the $2.5\,\sigma$ criterion is met comfortably. The distribution of the pseudo-experiments is also rather well described by a χ^2 distribution for ndf = 30, which equals the difference of degrees of freedom between the two fit setups. The observable departure from the expected distribution is manly driven by the low statistics in the rare electron modes.

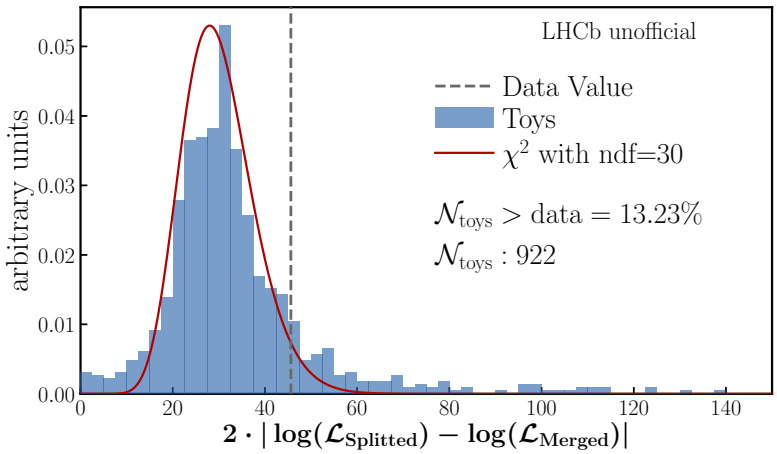

Figure 4.53.: Logarithmic likelihood difference between the two fit setups used to evaluate the self-consistency of the measured ratios between the various run periods and trigger categories. The distribution from pseudo-experiments is shown in blue, while the result from the fits to nominal data is indicated in gray.

4.9.2 Muon Branching Ratios

The branching ratios of the rare muon channels $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ are compared to the published values from LHCb analyses on Run 1 data given in Refs. [35] and [34], respectively. The comparison is performed in both the low and central q^2 region, where the low q^2 range is changed to 0.1–0.98 GeV^2/c^4 in order to align with the published analyses. Following the same strategy that is used in these analyses, the rare mode differential branching ratios ($d\mathcal{B}_{\text{rare}}/dq^2$) averaged over a q^2 bin of width $q^2_{\text{max}} - q^2_{\text{min}}$ are calculated by normalising to the respective J/ψ resonant channel, so that the final values are computed as

$$\frac{d\mathcal{B}_{\text{rare}}}{dq^2} = \frac{\mathcal{N}_{\text{rare}}}{\mathcal{N}_{J/\psi}} \cdot \frac{\varepsilon_{J/\psi}}{\varepsilon_{\text{rare}}} \cdot \frac{\mathcal{B}_{J/\psi}}{q^2_{\text{max}} - q^2_{\text{min}}},$$

where the same values for the branching ratios of the resonant modes ($\mathcal{B}_{J/\psi}$) are used as in the quoted analyses. For the B^0 mode another factor enters into this equation, namely $1 - F_s^{\text{rare}}/1 - F_s^{J/\psi}$, which subtracts the S-wave ($K\pi$ non-resonant) component F_s to measure the branching ratio of the P-wave (K^{*0} resonant) component only. The values for F_s are taken from Ref. [34].

The muon branching ratio check is performed in four stages that were executed subsequently:

1. Comparison of the measured muon branching ratios from the Run 1 and Run 2 results of this analysis to validate the self-consistency.
2. Comparison of the Run 2 result of this analysis with the published values.
3. Comparison of the Run 1 result of this analysis with the published values.

4. Comparison of the combined result of this analysis with the published values.

Each check is considered passed, if no measured deviation exceeds the $3\,\sigma$ threshold in all four studied q^2 bins. The systematic uncertainties discussed in Sec. 4.7 are not considered in these comparisons. For the third stage, the overlap of the datasets used in this and the published analyses need to be considered, since both are performed on the LHCb RUN 1 dataset. Due to the different selection strategy between the analyses, the datasets are not identical. The amount of events selected in this analysis that are shared with the published analyses is found to be at the level of 60 % for both B^+ and B^0 modes.

The result of all four comparisons in all four studied q^2 regions are given in Fig. 4.54. The figure shows that no result exceeds the $3\,\sigma$ criteria so that the muon branching ratio test is considered as successful. However, a relatively high tension is found between the RUN 1 results of $B^+ \to K^+\mu^+\mu^-$ in central q^2. Further investigations were performed to study this difference in more detail and no pathological problems were found. The tension is attributed partly to a fluctuation and to systematic differences of the analysis strategies, which are not accounted for in this cross-check. Detailed studies on the systematic differences have shown a reduced tension when aligning the approaches between the two analyses (see Ref. [93]).

4.9.3 R_K in Central q^2

The last check performed prior to the full unblinding of the analysis is the comparison of R_K in central q^2 with the published result from the LHCb analysis of the full $9\,\text{fb}^{-1}$ dataset given in Ref. [13]. Deviations from the published result can originate from differences in the calibration and calculation of efficiencies, the setup used for the invariant mass fits as well as different selections leading to a finite overlap of the fitted data samples. The chief distinction between the analyses is the much tighter selection applied to the rare electron mode in the analysis presented in this thesis. This leads to a significantly higher signal purity and control over the residual backgrounds allowing to fit the $B^+ \to K^+e^+e^-$ channel in a wider range of the reconstructed B mass (4600–6200 MeV in R_X vs. 4880–6200 MeV).

First, a rough calculation of the compatibility for the R_K values in central q^2 between the R_X and the published analysis is presented. Here, only statistical deviations from the rare electron mode are considered, which is the dominant source of uncertainty for the R_K measurement. Since both analyses exploit the same LHCb data sample, the overlap of the electron data sample after full selection must be taken into account. After analysing the datasets of both analyses it is found that about 75 % of the events in the R_X sample are shared with the published analysis, while about 60 % of the events from the published analysis are also included in the R_X sample. The allowed statistical deviation between the two analyses can be estimated with

$$\sigma_{\Delta(x,\,y)} = \left[(1 - f_y)\cdot\sigma_x^2 + (1 - f_x)\cdot\sigma_y^2\right]^{1/2}\,, \tag{4.30}$$

where $f_i = N_{\text{shared}}/N_{\text{total}}$ is the ratio of shared over total events with i referring to either the R_X or the published analysis. Taking the overlap values from above and the statistical

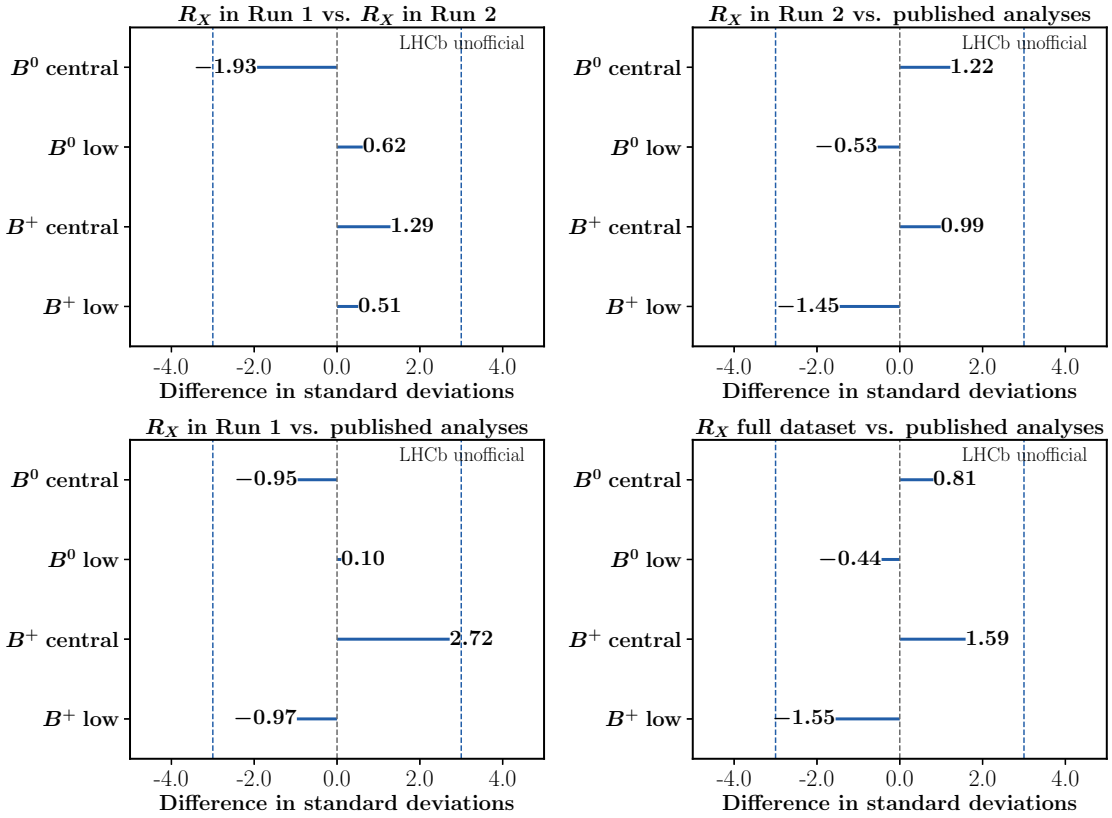

Figure 4.54.: Results of the muon branching ratio check in the four separate stages from left to right and top to bottom. The muon branching ratios of $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ in low and central q^2 measured in this analysis are compared with the results of the published Run 1 LHCb analyses.

uncertainties of both analyses (σ_i), the estimated $1\,\sigma$ deviation between them is 3.8 % (relative to the published value). With the observed difference of the central values the compatibility between the two analyses is found to be $1.81\,\sigma$.

Now, a more detailed computation of the $1\,\sigma$ deviation between the analyses is performed. Again, the allowed statistical deviation of the rare electron mode is considered by calculating the overlap of the data samples between the two analyses and using Eq. 4.30, where the statistical uncertainties on the yields of the rare electron modes are used as input for σ_i. The same approach is used for the statistical deviation of the rare muon mode. Additionally, the statistical deviation coming from the different levels of signal purity in the rare electron sample is considered, which strongly affects the sensitivity to the signal. This component is estimated based on pseudo-experiments, where the residual background levels of both analyses are emulated to study the impact on the signal yield uncertainty. Lastly, the systematic uncertainties are added in quadrature, by conservatively assuming them to be uncorrelated. Combining all sources, the overall expected $1\,\sigma$ deviation is found to be 4.1 % relative to the published value.

After unblinding the R_K central q^2 bin in R_X, the compatibility with the published result is found to be at $1.67\,\sigma$, which exceeds the passing threshold of $1.5\,\sigma$. Hence, more detailed studies are performed. Further investigations have shown, that the difference is mainly driven by the tighter electron PID selection used in the presented analysis. For clarity, the different electron PID selections used in the two analyses as well as the tight PID setup mentioned below are detailed in Tab. 4.28. When aligning the PID selection to the one used in the published analysis, a nearly perfect agreement between the R_K central values is found. A possible explanation for the change of R_K when relaxing the PID criteria is an underestimation of misidentification backgrounds of the type discussed in Sec. 4.7.2 in the published analysis.

Table 4.28.: Summary of the different electron PID selections used in the published R_K central analysis, the R_X analysis and the tight PID setup. The listed criteria are applied to both electrons of the signal candidate.

Setup	PID selection
R_K central analysis	$\texttt{DLL}_{e,\pi} > 3$
R_X analysis	$\texttt{DLL}_{e,\pi} > 2$ && $\texttt{ProbNN}_e > 0.2$
Tight PID	$\texttt{DLL}_{e,\pi} > 3$ && $\texttt{ProbNN}_e > 0.4$

The result of R_K in central q^2 of the R_X analysis (with the nominal PID selection) was further cross-checked: The control over efficiency variations between both PID setups is tested by studying the single $r^{J/\psi}$ ratio, which is found to be stable in all tested setups. Additionally, the yields obtained in the fits to the rare electron mode are checked by fitting in an even wider mass range of 4050–6200 MeV, which further tests the control over residual backgrounds in the mass fit. Finally, the stability of the PID working point of the R_X analysis was tested by using an even tighter electron PID selection (see Tab. 4.28). Both efficiencies and mass yields were reevaluated and a compatible value of the R_K ratio in central q^2 was found with the baseline result from R_X. After obtaining satisfying results on all these checks, the decision to proceed to the full unblinding was formed together with the review committee. More details on the comparison of R_K in central q^2 and the perfomed cross-checks are given in Ref. [93].

4.10 Results

The unblinded results for all four LFU observables from the simultaneous fit to the full Run 1 and Run 2 $9\,\text{fb}^{-1}$ data sample of LHCb are given in Fig. 4.55. The combined compatibility with the theory prediction will be evaluated using pseudo-experiments. Further, an interpretation of the measured ratios in terms of Wilson coefficients will be presented. Sections 4.10.2 and 4.10.3 describe the strategies that will be used to obtain

these results, however the outcome lies beyond the scope of this thesis because of the on-going studies performed post-unblinding that are subject of the subsequent section.

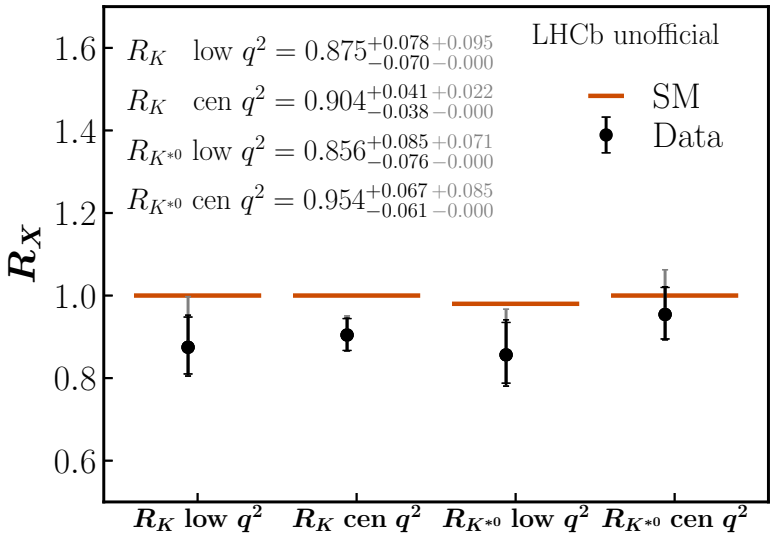

Figure 4.55.: The unblinded results of the four LFU ratios from the simultaneous fit to the full Run 1 and Run 2 9 fb^{-1} data sample of LHCb. The experimental uncertainties resulting from the fit including the systematic uncertainties discussed in Sec. 4.7 are displayed in black (the pure statistical uncertainty is displayed by the additional horizontal bar), while the uncertainty including the additional systematic due to electron PID instabilities (see Sec. 4.10.1) is shown in grey. The theory predictions are taken from Refs. [39, 40].

4.10.1 Post-Unblinding Checks

After the full unblinding, the cross-check using the tight PID setup, that is discussed during the comparison of the R_K central measurement with the published LHCb result in Sec. 4.9.3, was performed for the other three LFU ratios as well. Unlike for R_K central, the results of the other ratios are not stable when tightening the PID criteria but move significantly towards the SM expectation. The value of the results presented in this thesis is that they are unbiased as they are obtained in a blinded manner. Therefore, the central values can not be changed after unblinding but systematic uncertainties can be reevaluated.

To get a first understanding of the origin of these variations with the electron PID selections, contamination from residual misidentification backgrounds have been studied for analogue reasons as discussed for R_K central in Sec. 4.9.3. For this, all LFU observables are recalculated using the nominal PID selection but excluding events from the Brem0 category (as defined in Sec. 4.6.1). The reason behind this test is, that misidentifications involving the final state electrons mostly occur in the Brem0 category, since

the addition of bremsstrahlung photons strongly improves the identification of electrons (see Sec. 3.4.1). The result of all for LFU ratios obtained from the Brem1 and Brem2 categories only is found to be compatible with the tight PID scenario reinforcing the hypothesis of residual backgrounds from misidentification. Thus, extensive studies have been launched to investigate the source and amount of such backgrounds.

So far, decay channels involving a double hadronic to electron misidentification are considered to explain at least a part of the observed deviations. One of these decays is the aforementioned $B^+ \to K^+\pi^+\pi^-$ mode with a double $\pi \to e$ misidentification, but also $B^+ \to K^+K^+K^-$ and $B^+ \to (D^0 \to K^+\pi^-)K^+$ decays featuring $K \to e$ misidentifications are found to play a role (similar channels are studied for the $R_{K^{*0}}$ modes). Additionally, modes such as $B^0 \to K^+\pi^-\pi^0$ are investigated, which can be reconstructed as a $B^+ \to K^+e^+e^-$ candidate under a single $\pi \to e$ misidentification and when the other electron stems from a photon conversion in interactions with detector material of one of the photons from the π^0 decay. First studies have shown, that these modes have a peaking structure below the B mass and can enter all regions of q^2 considered in this analysis.

Currently, various approaches are discussed to estimate the residual contamination from these hadron to electron misidentification backgrounds after the PID selection of the R_X analysis. One pursued strategy studies the behaviour of the results when successively tightening the cut value on the $\texttt{ProbNN}_e$ variable starting at the nominal working point $\texttt{ProbNN}_e > 0.2$ and going up to $\texttt{ProbNN}_e > 0.6$. This scan over $\texttt{ProbNN}_e$ cuts shows, that the R_K ratio in central q^2 is very robust, while the other three measured LFU ratios show a clear trend towards higher values between the $\texttt{ProbNN}_e$ cut values 0.2 and 0.4. All four ratios show a stable behaviour above $\texttt{ProbNN}_e > 0.4$. In the context of this thesis, a systematic uncertainty on the stability of the ratios against changes of the electron PID working point is estimated by taking the difference of the nominal result and the saturation value observed in the scan over $\texttt{ProbNN}_e$ cut thresholds at values between 0.4 and 0.6. The resulting values are 0.095 and 0.022 for R_K in low and central q^2 as well as 0.071 and 0.085 for $R_{K^{*0}}$, which are taken as an asymmetric systematic uncertainty towards higher values. This additional systematic is assigned in this thesis as a very conservative and preliminary estimate of the observed effect. Currently, a large effort is made by the LHCb collaboration to improve the stability of the ratios and reduce the associated systematic uncertainty. At this point, it is unclear whether changes to the central values of the unblinded result shown in Fig. 4.55 will be introduced following the final outcome of these studies. Also the systematic uncertainty assigned to the residual background level discussed in Sec. 4.7.2 might be part of future changes. While the author of this thesis is part of the team performing these studies, their outcome is beyond the scope of this thesis.

4.10.2 SM Compatibility

In order to evaluate the overall compatibility of the presented measurement with the SM expectation, the likelihood ratio between the nominal fit and the SM hypothesis is calculated. From the nominal fit a likelihood function is constructed

$$\mathcal{L} = \mathcal{L}(\vec{R}, \vec{\lambda}|\text{data})\,, \tag{4.31}$$

with the LFU observables $\vec{R} = \{R_K^{\text{low}}, R_{K^*}^{\text{low}}, R_K^{\text{central}}, R_{K^*}^{\text{central}}\}$ and nuisance parameters $\vec{\lambda}$ yielding the best estimates for these parameters given the used data. A second maximisation of the likelihood is performed setting the $\vec{R}$ parameters to the SM values. To compare these two likelihoods their ratio is calculated logarithmically

$$-2\Delta\ln\mathcal{L}(\vec{R}|\text{data}) = -2\Big[\ln\mathcal{L}(\vec{R}^{\text{SM}}, \vec{\lambda}'|\text{data}) - \ln\mathcal{L}(\vec{R}, \vec{\lambda}|\text{data})\Big]\,. \tag{4.32}$$

To account for the small uncertainty of the SM prediction, the difference to the SM values, $\Delta\vec{R} = \vec{R} - \vec{R}^{\text{sm}}$, is fitted for, which allows to add the theory uncertainties as multivariate normal constraints.

Now, the calculation of the p-value for the SM hypothesis is dependent on the $-2\Delta\ln\mathcal{L}$ distribution. Since it is not guaranteed that the experimental uncertainties are Gaussian distributed, the p-value is not simply calculated assuming a χ^2 distribution and Wilks theorem [114]. Indeed, the likelihood scans performed in Sec. 4.6.4 have shown some deviations from Gaussian behaviour of the LFU ratios caused by the small statistics in the rare electron modes. Instead, the test statistics of $-2\Delta\ln\mathcal{L}$ is obtained from pseudo-experiments. These pseudo-experiments are designed to mimic the SM scenario, which is achieved by generating the rare mode yields as expected from theory. For each pseudo-experiment, the rare mode yields are obtained from the known decay rates of the J/ψ mode and the relative efficiencies of signal and resonant mode. Next, the value of $-2\Delta\ln\mathcal{L}$ is determined for all pseudo-experiments and the p-value of the SM hypothesis follows as the fraction of pseudo-experiments producing a larger value than the one obtained on data. This p-value will be considered as the key result of this analysis, however an additional interpretation of the measurement in terms of effective couplings is described in the next section.

4.10.3 Interpretation in Terms of Effective Couplings

The measured LFU observables can be interpreted by employing the framework of an effective field theory in terms of the $b \to s\,\mu^+\mu^-$ effective couplings, *i.e.* Wilson coefficients as described in Sec. 2.4. Since Wilson coefficients are universal for all $b \to q$ processes, they offer a powerful tool to combine the result of the four LFU ratios measured in this analysis, taking into account the full experimental correlations. Several scenarios involving one- and two-dimensional variations of the coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ are considered that are currently widely discussed in literature on global fits, *e.g.* Ref. [95].

A similar approach based on a likelihood ratio test as used for the SM p-value is exploited. Instead of comparing to the SM values, theory predictions and their uncertainties are

calculated using the `flavio` software package [94] depending on the Wilson coefficients $\vec{R}^{\mathrm{NP}}(\mathcal{C}_i)$ in various NP scenarios. Scans over the Wilson coefficient space are performed, where for each point $-2\Delta\ln\mathcal{L}$ is determined using the fitting framework of the R_X analysis. This way, the full experimental covariance matrix is properly accounted for in the scans. Figure 4.56 shows some example scans based on a SM like pseudo-experiment. The scans allow to extract the best fit value of the coefficients as well as the p-value of the SM hypothesis in each scenario.

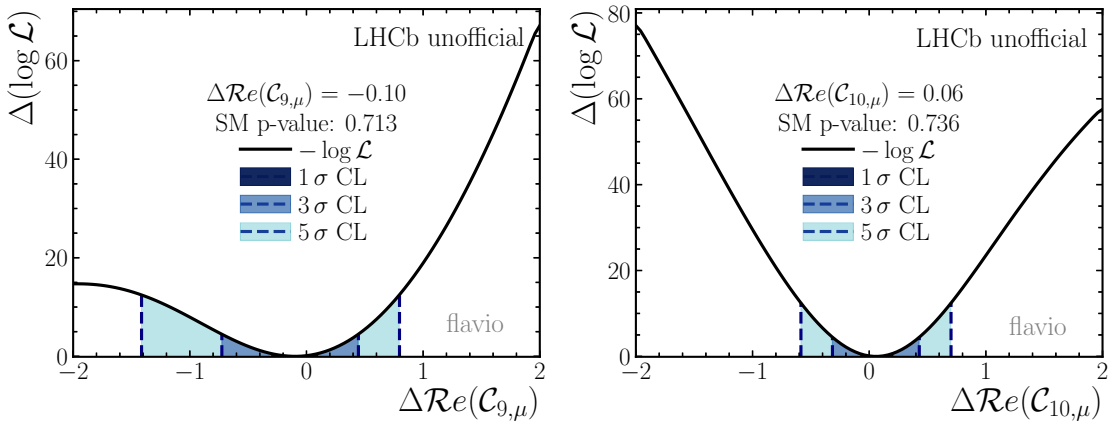

Figure 4.56.: Example scans over the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ based on a SM like pseudo-experiment. The figure is only included to illustrate the method and does not include information on the final LFU observables. The confidence level intervals are calculated using Wilks theorem [114].

5 Conclusions and Outlook

This thesis presents the first simultaneous measurement of lepton universality in the observables R_K and $R_{K^{*0}}$ for both the low and central region of the mass of the dilepton system squared, q^2. The full RUN 1 and RUN 2 dataset of LHCb is exploited, which corresponds to an integrated luminosity of $9\,\text{fb}^{-1}$ in total. The simultaneous approach allows to constrain the background from $B^0 \to K^{*0}e^+e^-$ decays in the $B^+ \to K^+e^+e^-$ mode, which improves the signal sensitivity. Furthermore, it allows to extract the full experimental covariance matrix of the four LFU ratios. The latter is vital input when combining the four measurements in global fits, *e.g.* interpreting them in an effective field theory as discussed in Sec. 4.10.3. The presented analysis is a product of a group effort from several members of the Rare Decays working group of LHCb to which the author of this thesis has made significant contributions (for details see Chap. 4).

The used selection is optimised for the experimentally challenging electron modes and applies rather tight criteria in comparison to previous LFU analyses at LHCb, yielding a high signal purity. In particular, the selection suppresses misidentified backgrounds using the PID response and mass vetoes and employs criteria against partially reconstructed and combinatorial backgrounds based on dedicated MVA selections. The high signal purity allows to perform the simultaneous fit to the reconstructed B mass in a wide range (4600–6200 MeV), providing a good handle over the residual background contributions.

The efficiencies are determined with dedicated fully simulated Monte Carlo (MC) samples of the signal and normalisation modes. A novel strategy to calibrate simulation is employed and verified by comparing the distribution in various key quantities (*e.g.* $p_\text{T}(B)$) between data and corrected MC. Moreover, the $r^{J/\psi}$ single ratio provides a very stringent cross-check of the calibration method. Also the $R^{\psi(2S)}$ double ratio is a powerful cross-check designed to test the portability of the efficiency corrections to different regions of q^2 and to validate the double ratio strategy. To calculate these ratios, a simultaneous fit to the J/ψ and $\psi(2S)$ resonant modes is performed, where the efficiencies are included to directly extract the desired ratios. In addition, the systematic uncertainties discussed in Sec. 4.7 are included in the fit. The final values of both single and double ratio cross-checks from the fit to the full dataset are:

$$\begin{aligned} B^+ \text{ modes: } r^{J/\psi} &= 1.042 \pm 0.025 \quad \text{and} \quad R^{\psi(2S)} = 0.989 \pm 0.008 \\ B^0 \text{ modes: } r^{J/\psi} &= 1.022 \pm 0.024 \quad \text{and} \quad R^{\psi(2S)} = 1.014 \pm 0.014\,. \end{aligned}$$

The good agreement with unity of these ratios demonstrates the great control over efficiencies in this analysis.

The determination of the four LFU observables follows an analogous approach as used for $r^{J/\psi}$ and $R^{\psi(2S)}$ and is based on a simultaneous fit to the rare and J/ψ resonant modes.

Using pseudo-experiments, the stability of the simultaneous fit is successfully tested and the presence of biases is checked for. Table 5.1 gives the values of R_K and $R_{K^{*0}}$ in both q^2 regions and the uncertainties from the simultaneous fit including the systematic uncertainties discussed in Sec. 4.7. The second uncertainty represents the systematic due to electron PID instabilities that are introduced in Sec. 4.10.1. Currently, a large effort is made by the LHCb collaboration to understand the origin of these instabilities and reduce the preliminary systematic uncertainty that is assigned in this thesis. Without this additional systematic, this result constitutes the most precise measurement of $R_{K^{*0}}$ in low and central q^2 as well as of R_K in central q^2, while it represents the first measurement of R_K in low q^2. The final compatibility with the SM predictions will be evaluated with the methods described in Sec.4.10.2 after the studies of electron PID instabilities are completed.

Table 5.1.: Values of the LFU ratios R_K and $R_{K^{*0}}$ in both low and central q^2 measured with a simultaneous fit to the full RUN 1 and RUN 2 LHCb dataset. The first uncertainty is coming from the simultaneous fit including the systematics from Sec. 4.7, while the second uncertainty represents the systematic due to electron PID instabilities discussed in Sec. 4.10.1. For comparison, SM predictions are quoted that are taken from Refs. [39, 40].

LFU bin	Simultaneous fit result	SM prediction
R_K low	$0.875^{+0.078+0.095}_{-0.070-0.000}$	1.00 ± 0.01
R_K central	$0.904^{+0.041+0.022}_{-0.038-0.000}$	1.00 ± 0.01
$R_{K^{*0}}$ low	$0.856^{+0.085+0.071}_{-0.076-0.000}$	0.98 ± 0.01
$R_{K^{*0}}$ central	$0.954^{+0.067+0.085}_{-0.061-0.000}$	1.00 ± 0.01

The measured value of R_K in central q^2 deviates by $1.67\,\sigma$ from the previously published value by LHCb [13]. Details on the calculation and origin of the observed deviation are given in Sec. 4.9.3. A graphic comparison of the results from both analyses is given in Fig. 5.1. The figure also includes the results from the RUN 1 $R_{K^{*0}}$ analysis [11]. It can be seen, that the resulting values from the presented analysis are closer to the SM expectation compared to previous LHCb measurements, which can be partly explained by the cleaner selection working point used in the presented analysis making it less affected by the electron PID instabilities.

As shown in Tab. 4.26, the systematic due to electron PID instabilities even dominates over the statistical uncertainty in some LFU bins, which is why the current investigations are vital to further improve the sensitivity. However, without this systematic the presented measurement is dominated by the statistical uncertainties, which are driven by the signal yield of the rare electron modes. Thus, it is crucial to update the analysis with LHCb data from future run periods. For example, RUN 3 is scheduled to start in 2022 and is expected to collect a dataset corresponding to an estimated integrated luminosity

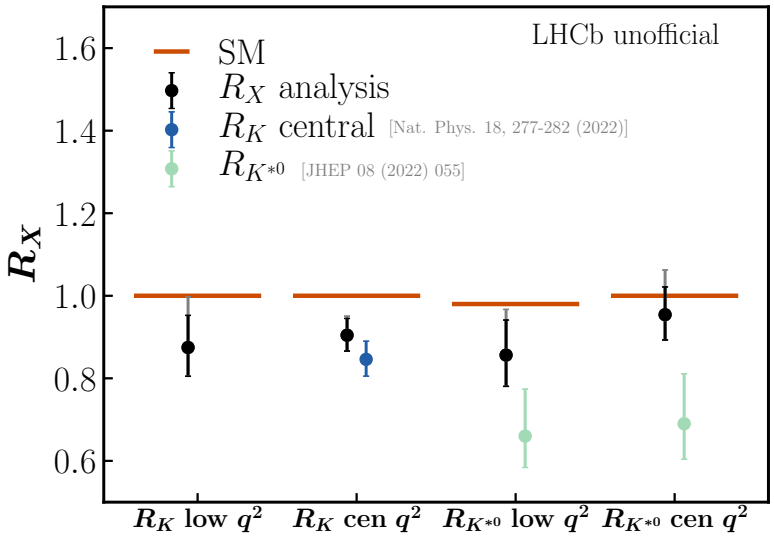

Figure 5.1.: The results of R_K and $R_{K^{*0}}$ in low and central q^2 regions from the analysis presented in this thesis (black) compared to previous results from LHCb (green and blue) and the SM expectation (orange). The black lines show the quadratic sum of the statistical and systematic uncertainties, the grey lines indicate the post-unblinding uncertainties from the variations of the electron PID selection. Note that these systematic uncertainties were not investigated for the previously published LHCb results.

of $23\,\text{fb}^{-1}$. Further, the upgrades to the LHCb detector [117] that are currently installed as well as new analysis methods will likely also allow to reduce the systematic budget. Most importantly, starting in Run 3 the LHCb detector will operate without the L0 trigger and fully rely on software triggers. This will enable a much more coherent treatment of electron and muon final states as well as a more efficient and flexible data-taking.

Also other experiments such as Belle II [118] and CMS [62] are currently collecting and analysing datasets to provide independent measurements of R_K and $R_{K^{*0}}$. Since they use largely orthogonal setups, these results will be a vital cross-check of the LHCb measurements. In addition, analyses of different decay modes such as $B_s^0 \to \phi\ell^+\ell^-$ (R_ϕ) and $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$ ($R_{K\pi\pi}$), which are currently performed at LHCb, will provide further avenues to test lepton universality.

A Acronyms

ALICE A Large Ion Collider Experiment

ATLAS A Toroidal LHC ApparatuS

$\mathcal{B}$ Branching Ratio

BDT Boosted Decision Tree

BSM Beyond Standard Model

CB Crystal Ball

CERN European Organisation for Nuclear Research

CMS Compact Muon Solenoid

CKM Cabibbo, Kobayashi and Maskawa

DSCB Double-Sided Crystal Ball

DTF `DecayTreeFitter`

ECAL Electromagnetic Calorimeter

EFT Effective Field Theory

EW Electroweak theory

FCNC Flavor Changing Neutral Currents

FOM Figure Of Merit

FSR Final-State Radiation

GIM Glashow, Iliopoulos and Maiani

HCAL Hadronic Calorimeter

HLT High Level Trigger

IP Impact Parameter

IT Inner Tracker

KDE Kernel Density Estimator

LCSR Light-Cone Sum Rules

LEP Large Electron-Positron Collider

LFU Lepton Flavour Universality

LHC Large Hadron Collider

LHCb Large Hadron Collider beauty

L0 Level-0 Hardware Trigger

MC Monte Carlo

MIP Minimum Ionising Particle

MS Multiple Scattering

MVA Multivariate Analysis

NP New Physics

OT Outer Tracker

PDF Probability Density Function

PID Particle Identification

PV Primary Vertex

QED Quantum Electrodynamics

QCD Quantum Chromodynamics

RD Rare Decays

RICH Ring Imaging Cherenkov Detectors

ROC Receiver Operating Characteristic

SM Standard Model of Particle Physics

SS Same Sign

ST Silicon Tracker

SV Secondary Vertex

TCK Trigger Configuration Key

TT Tracker Turicensis

VELO Vertex Locator

B Glossary of Variables

A summary of variables used throughout the analysis is presented. The variables are introduced using the $B^+ \to K^+\mu^+\mu^-$ decay as example.

Tracking:

$p_{\mathrm{T}}(K^+)$	Kaon's transverse momentum with respect to the beam axis
$\eta(K^+)$	Angle between the track of the kaon and the beam axis
`DIRA`(B^+)	Angle between a line drawn from the PV to the decay vertex of the B^+ meson and its vector of momentum
$\theta(K^+, \mu^+)$	Angle between the kaon and μ^+ candidate
$\mathrm{IP}(B^+)$	Shortest distance between the reconstructed B^+ track and the associated PV
$\mathrm{FD}(B^+)$	Distance between the decay vertex of the B^+ and the associated PV
$\chi^2_{\mathrm{IP}}(B^+)$	χ^2 difference of a PV when reconstructed with and without the B^+ candidate (A measure of how likely the track originated from the vertex)
$\chi^2_{\mathrm{FD}}(B^+)$	χ^2 of the FD (A measure of how well the SV is separated from the PV)
$\chi^2_{\mathrm{PV,\,SV}}(B^+)$	χ^2 difference between a fit assuming all tracks stem from the PV and a fit separating the B^+ candidate to a SV
$\chi^2_{\mathrm{vtx}}(B^+)$	χ^2 of the decay vertex fit of the B^+ candidate
$\chi^2_{\mathrm{DTF}}(B^+)$	χ^2 of the `DecayTreeFitter` [99] fit of the B^+ decay chain
`GhostProb`(K^+)	Probability for the kaon track of being a fake track from combinations of random hits in the tracking system
$x_{ECAL}(K^+)$	x coordinate of the kaon track extrapolated to the ECAL plane
$y_{ECAL}(K^+)$	y coordinate of the kaon track extrapolated to the ECAL plane
$d_{ECAL}(\mu^+, \mu^-)$	Distance between the two muon tracks extrapolated to the ECAL plane
$\mathrm{region}_{ECAL}(K^+)$	ECAL region hit by the kaon. Assigned value: Outer=0, middle=1, inner=2, out of acceptance<0

PID:

$\texttt{DLL}_{x,\pi}(\mu^+)$	Log-likelihood difference between the mass hypotheses $x = K$, p, μ, e and the pion hypothesis for the muon candidate based on inputs from the RICH and calorimeter subdetectors
$\texttt{ProbNNx}(\mu^+)$	Probability for the muon candidate of being a particle of type $x = \pi$, K, p, μ, e based on information from the full LHCb detector
$\texttt{isMuon}(\mu^+)$	Candidate is compatible with being a muon based muon station and calorimeter information
$\texttt{inAccMuon}(\mu^+)$	Muon candidate track falls into the acceptance of the muon stations
$\texttt{hasMuon}(\mu^+)$	Muon candidate has information from the muon stations assigned
$\texttt{hasCalo}(\mu^+)$	Muon candidate has information from the calorimeter system assigned
$\texttt{hasRICH}(\mu^+)$	Muon candidate has information from the RICH assigned

Global event:

`nTracks`	Number of reconstructed tracks
`nVeloTracks`	Number of VELO tracks
`nSPDHits`	Number of hits recorded in the SPD
`nPVs`	Number of primary vertices

Isolation:

$m_{\text{vtx,iso}}(B^+)$	Mass of the candidate with the smallest χ^2 difference, when sequentially adding other tracks of the event to the B vertex to form a new vertex
$\chi^2_{\text{vtx,iso}}(B^+)$	χ^2 of the candidate with the smallest χ^2 difference, when sequentially adding other tracks of the event to the B vertex to form a new vertex
$\text{MULT}_{\text{cone,iso}}(K^+)$	Number of other tracks within a 0.5 mrad wide cone around the track of the kaon
$\text{PT}_{\text{cone,iso}}(K^+)$	Summarised p_T of other tracks within a 0.5 mrad wide cone around the track of the kaon
$\text{IT}_{\text{cone,iso}}(K^+)$	p_T ratio of kaon and all other tracks within a 0.5 mrad wide cone around the track of the kaon

C Comparison of Data and Simulation

The data and simulated samples used for this cross-check are from the J/ψ resonant mode and have the full selection applied (MVA cuts are loosened). Additionally, a 60 MeV mass window around the B mass on the J/ψ constrained B mass system is applied. As an example Run 2p2 samples are shown in the L0I trigger category which are representative for all run periods. The full calibrations to simulation from Sec. 4.4 are applied for the 'corrected' MC distribution.

C.1 B^+ Muon Mode

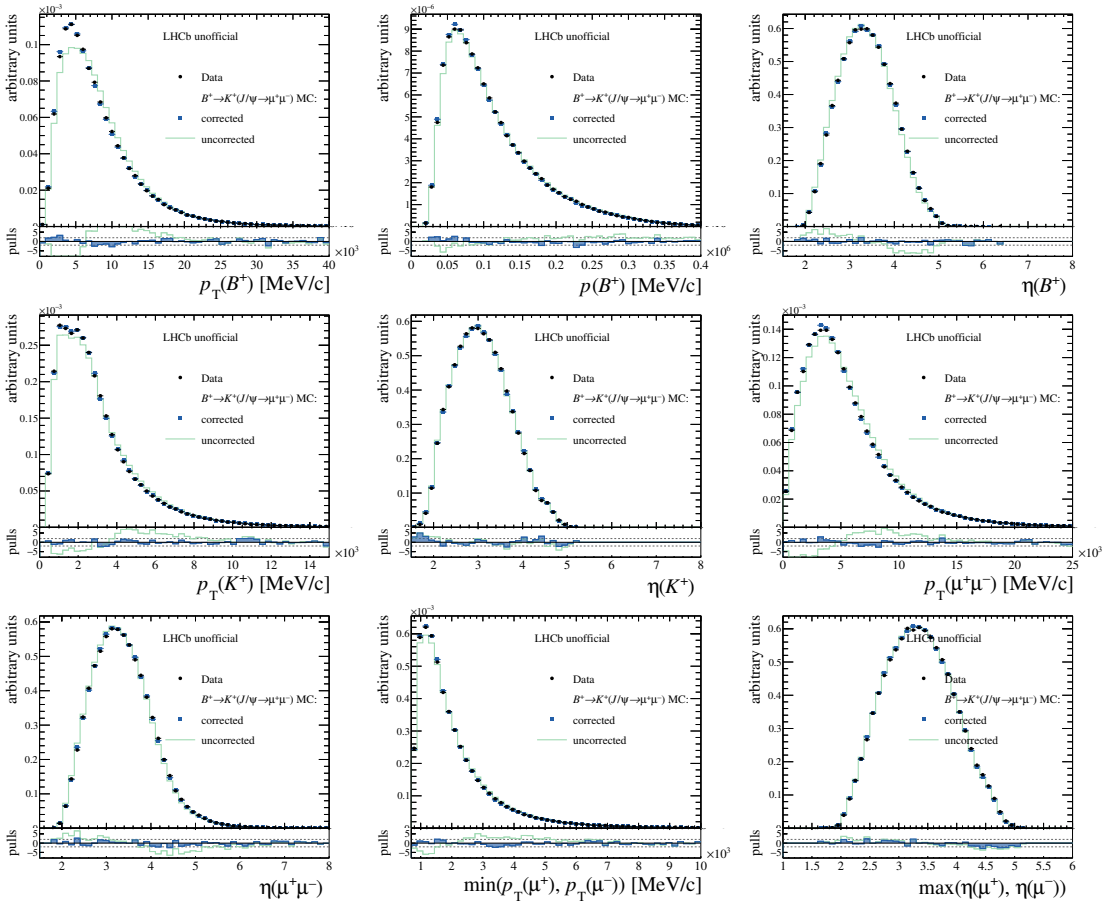

Figure C.1.: Kinematic variables.

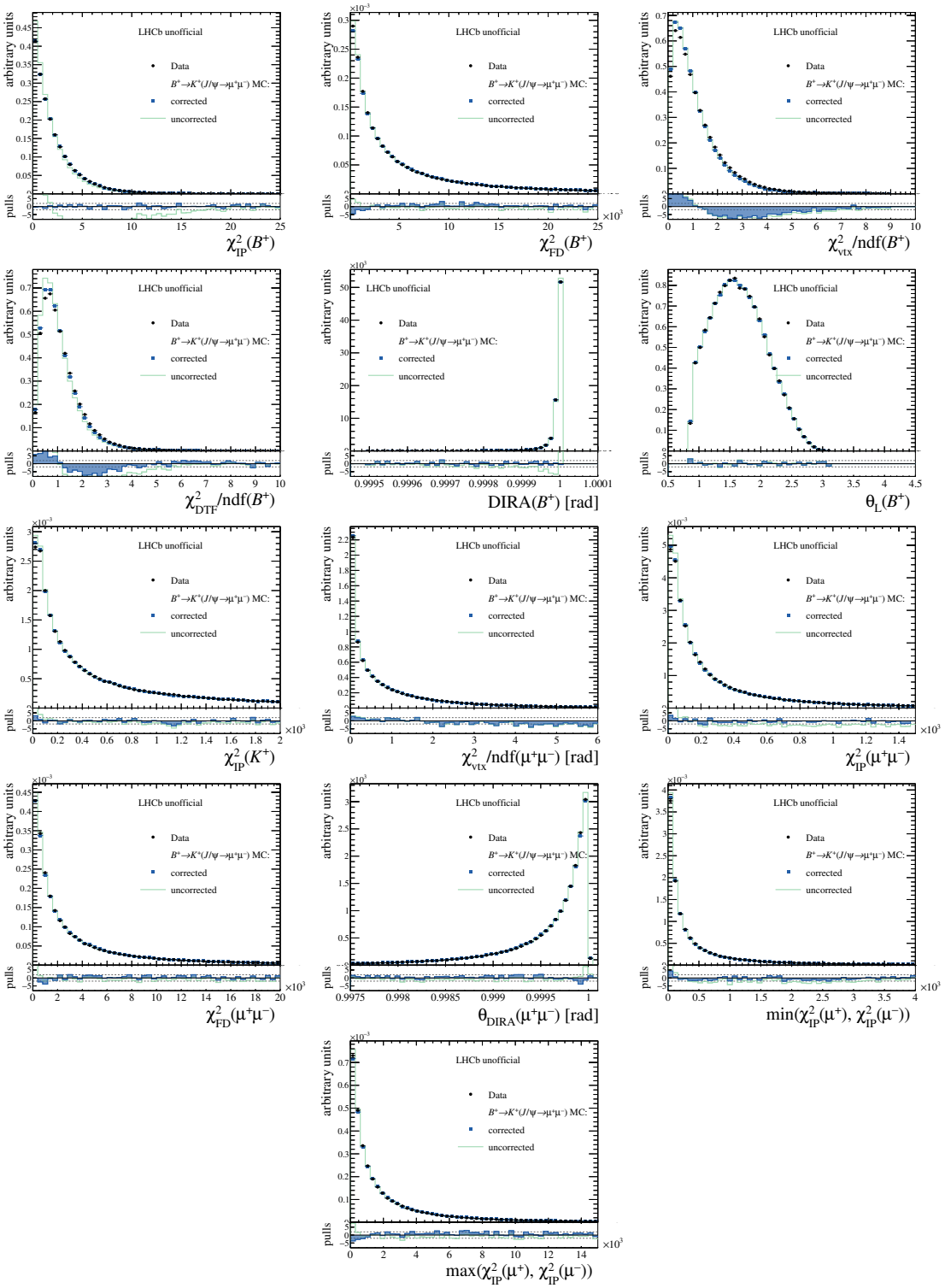

Figure C.2.: Reconstruction variables.

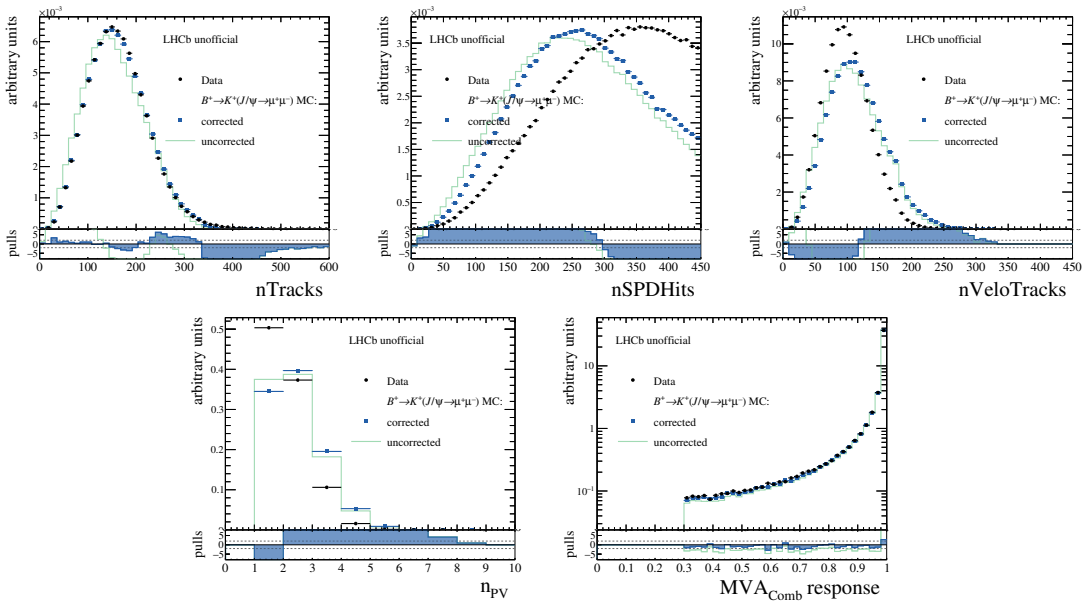

Figure C.3.: Multiplicity and MVA variables.

C.2 B^+ Electron Mode

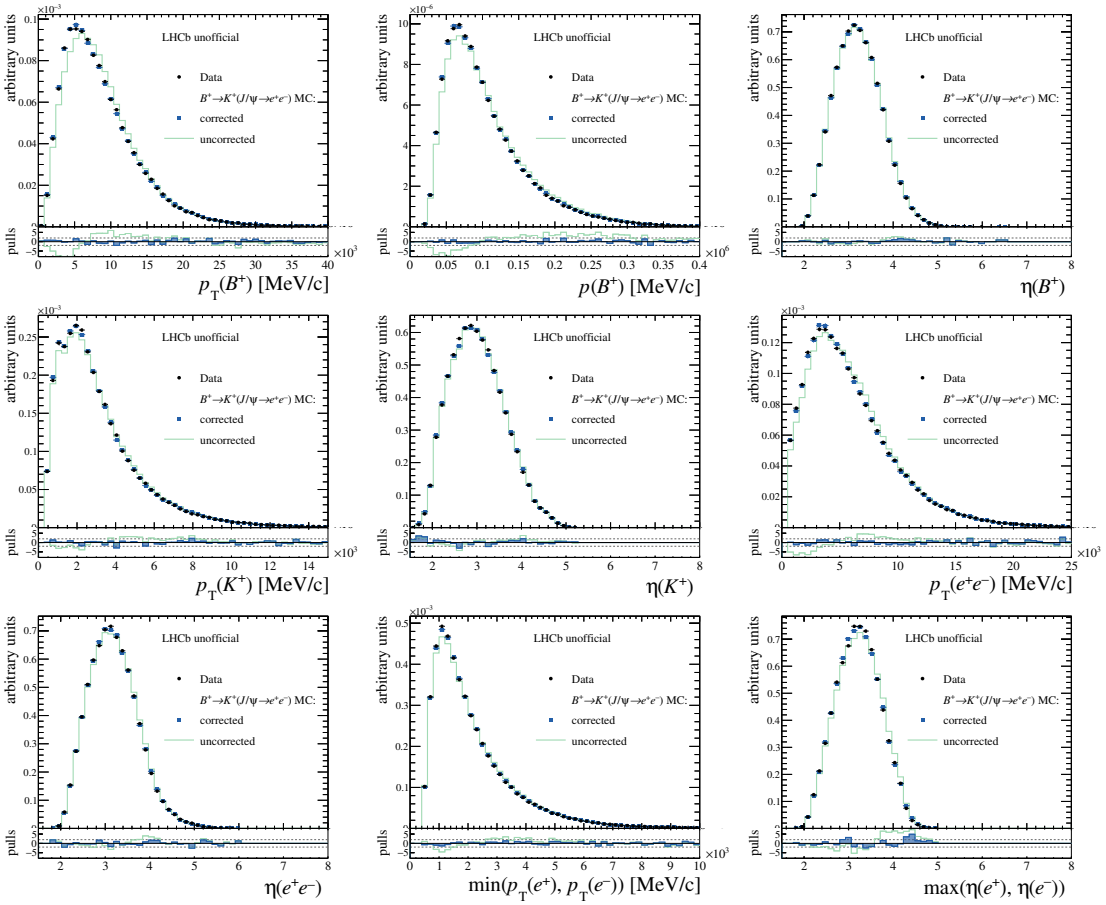

Figure C.4.: Kinematic variables.

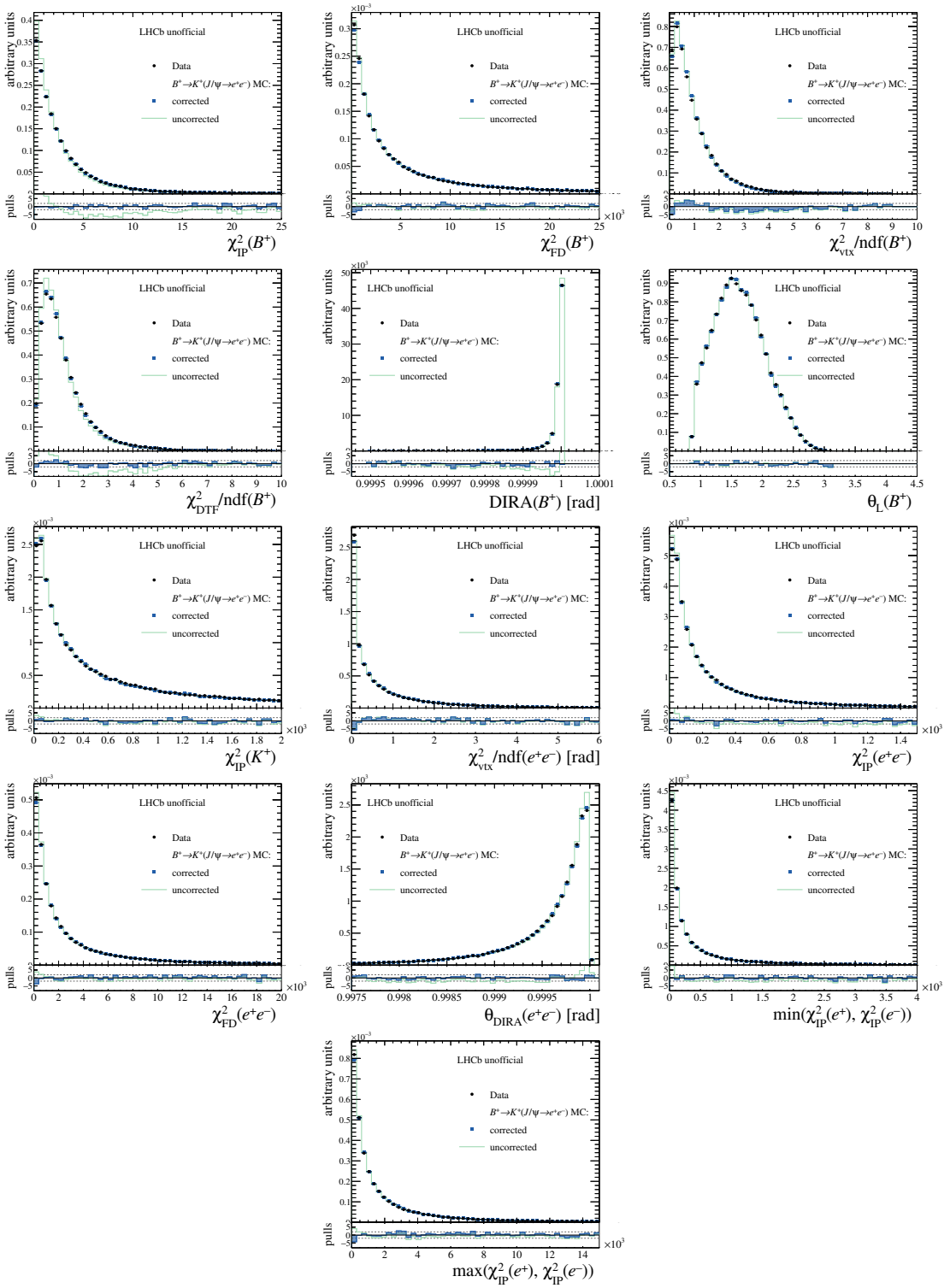

Figure C.5.: Reconstruction variables.

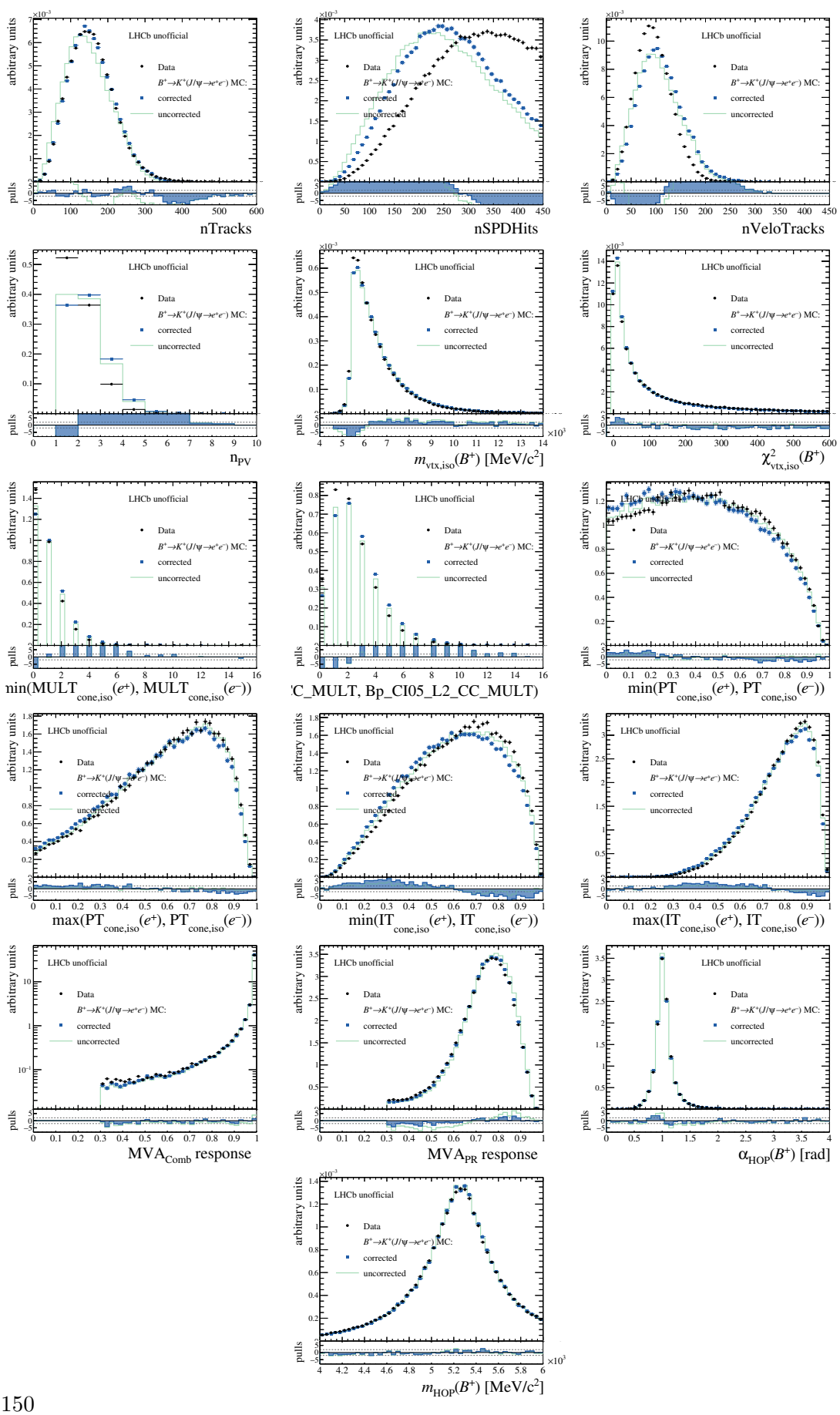

Figure C.6.: Multiplicity, isolation, MVA and HOP variables.

C.3 B^0 Muon Mode

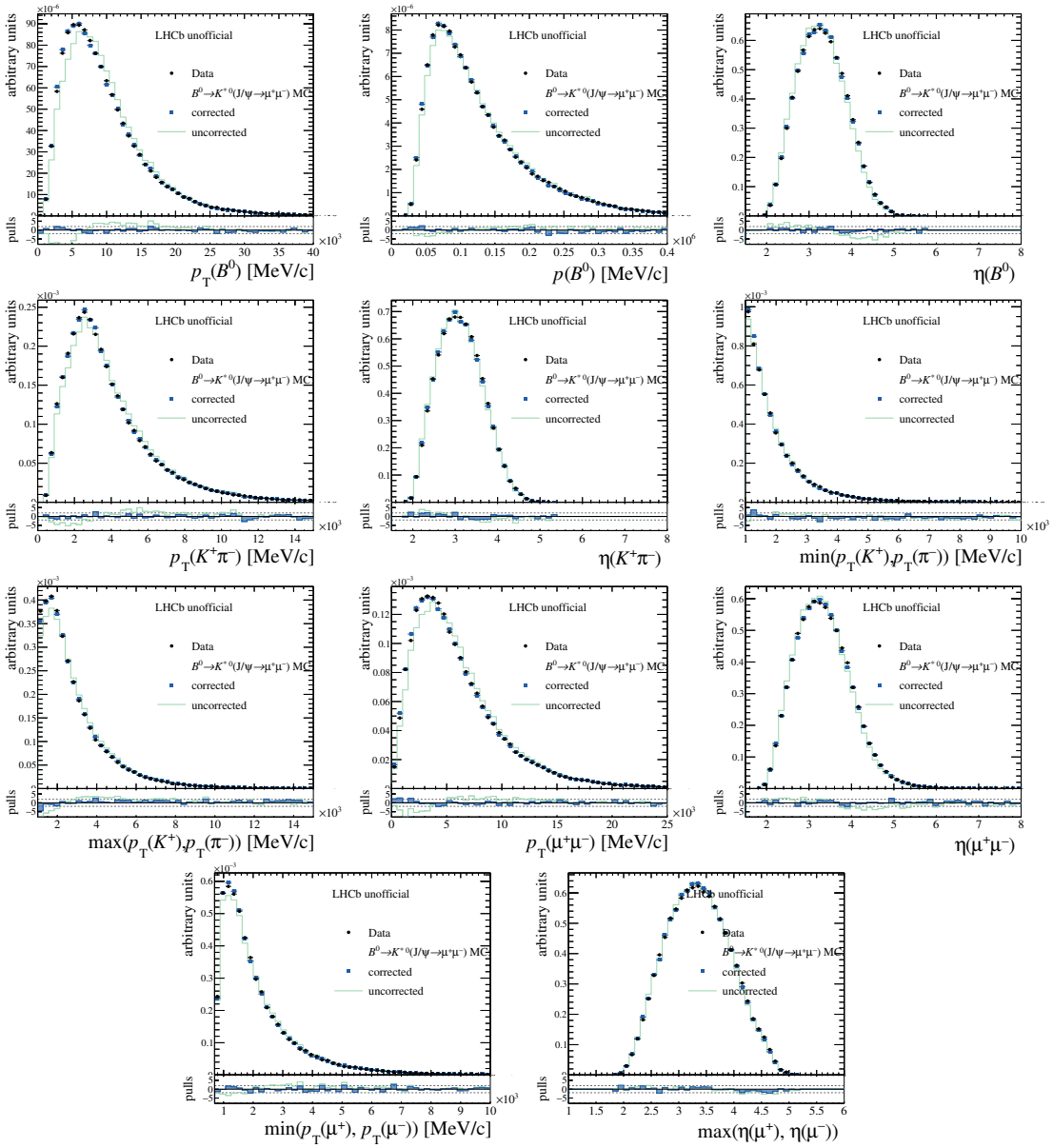

Figure C.7.: Kinematic variables.

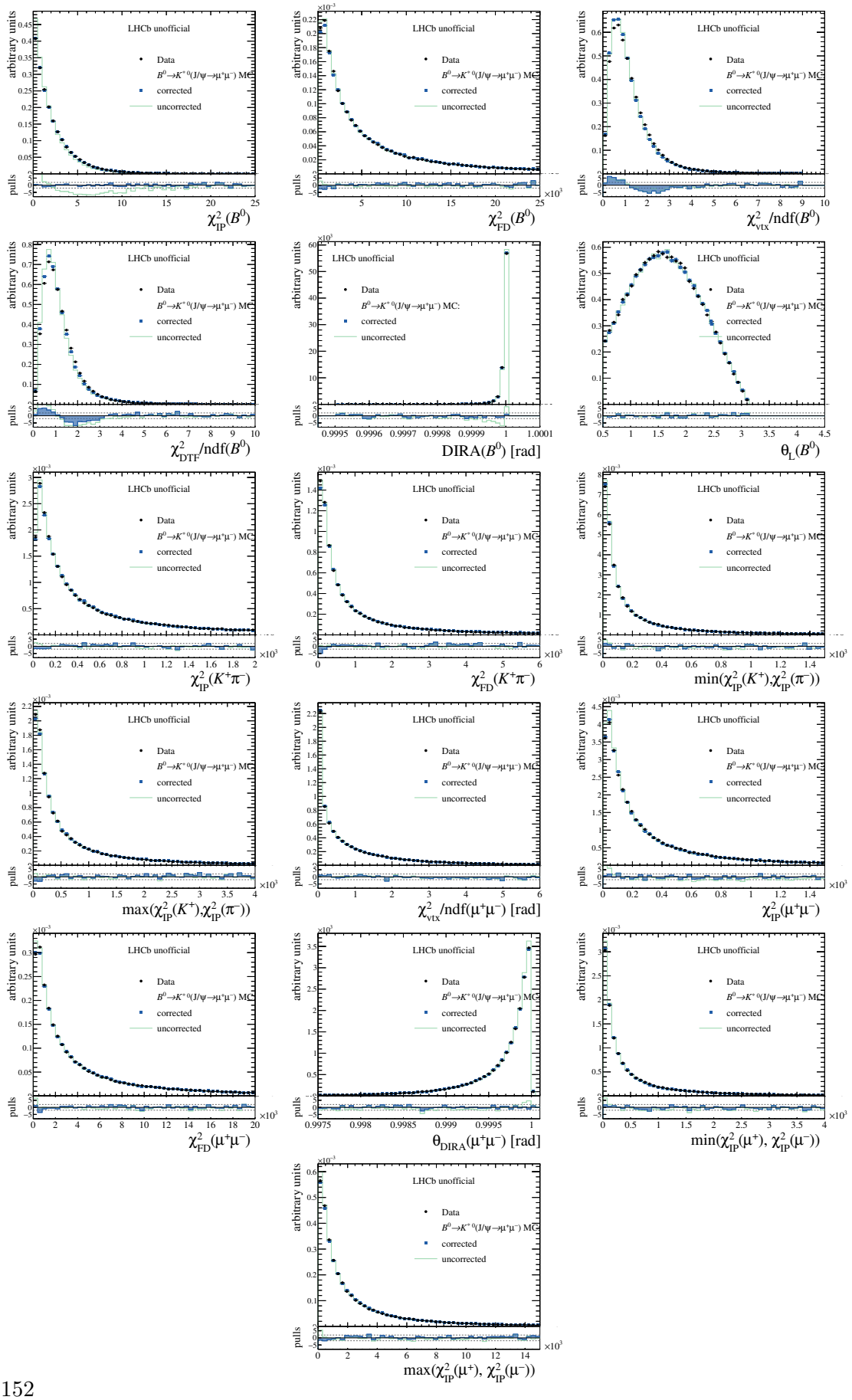

Figure C.8.: Reconstruction variables.

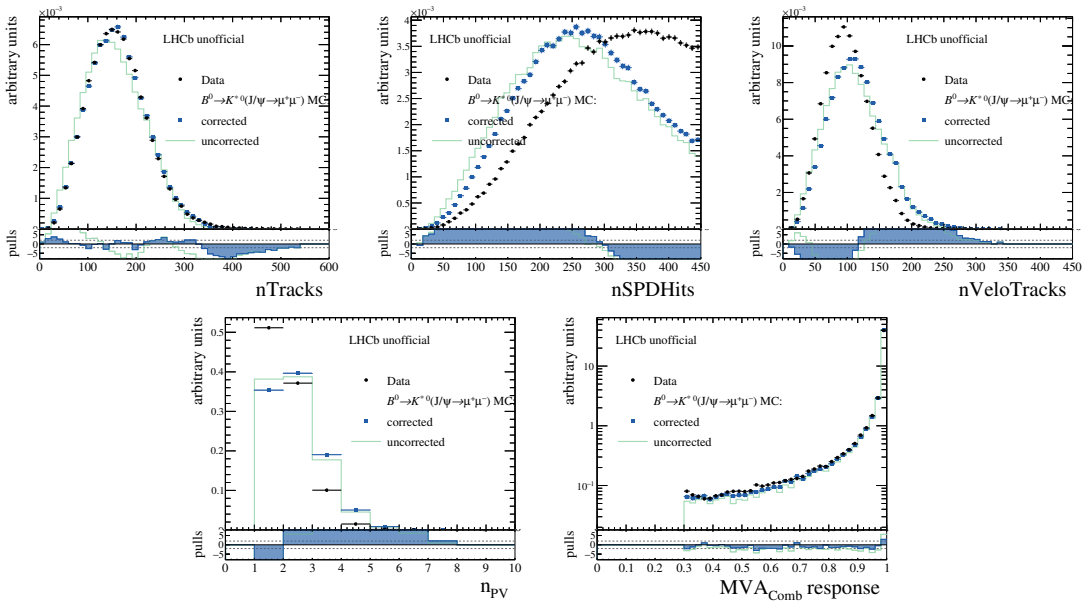

Figure C.9.: Multiplicity and MVA variables.

C.4 B^0 Electron Mode

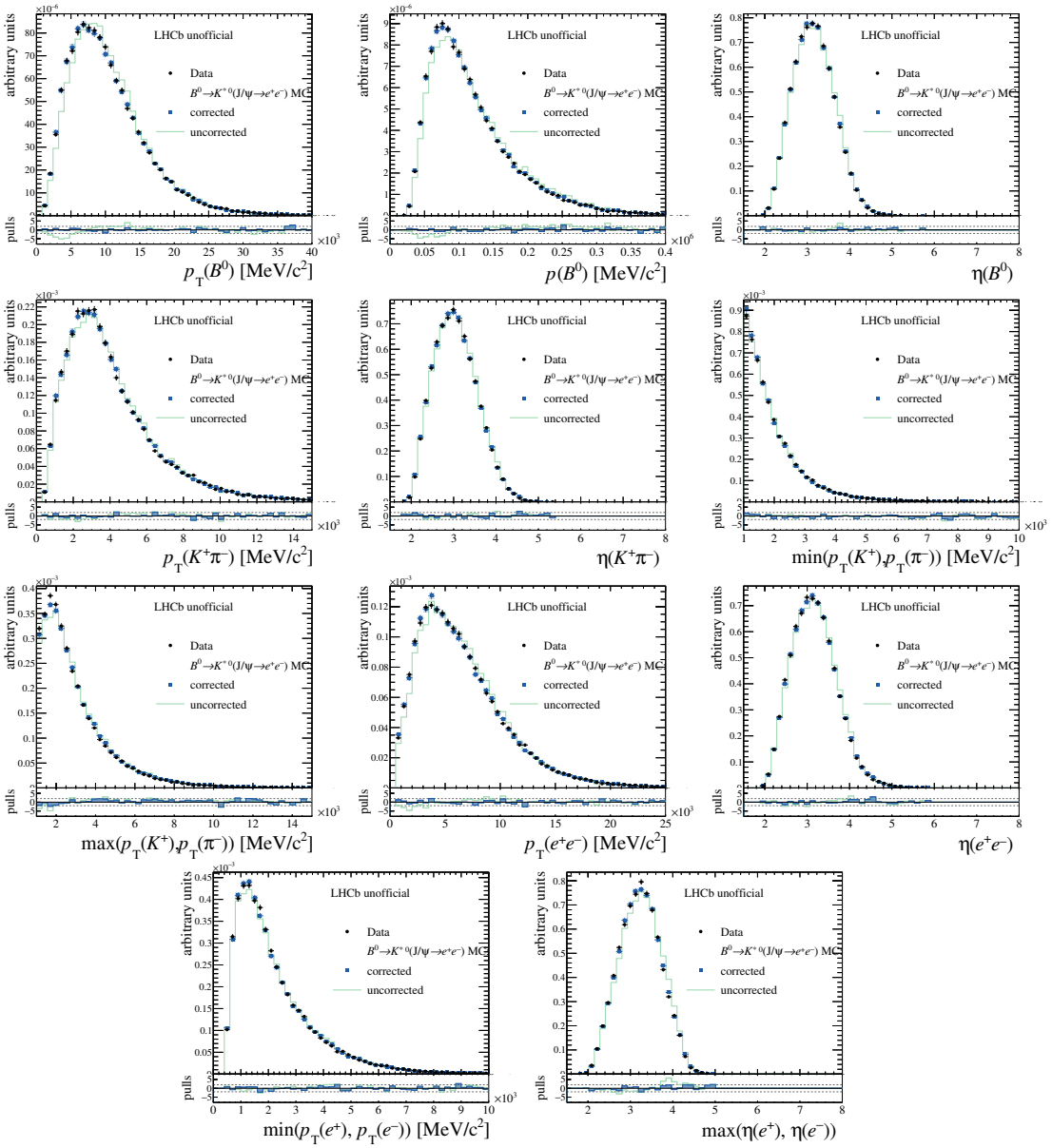

Figure C.10.: Kinematic variables.

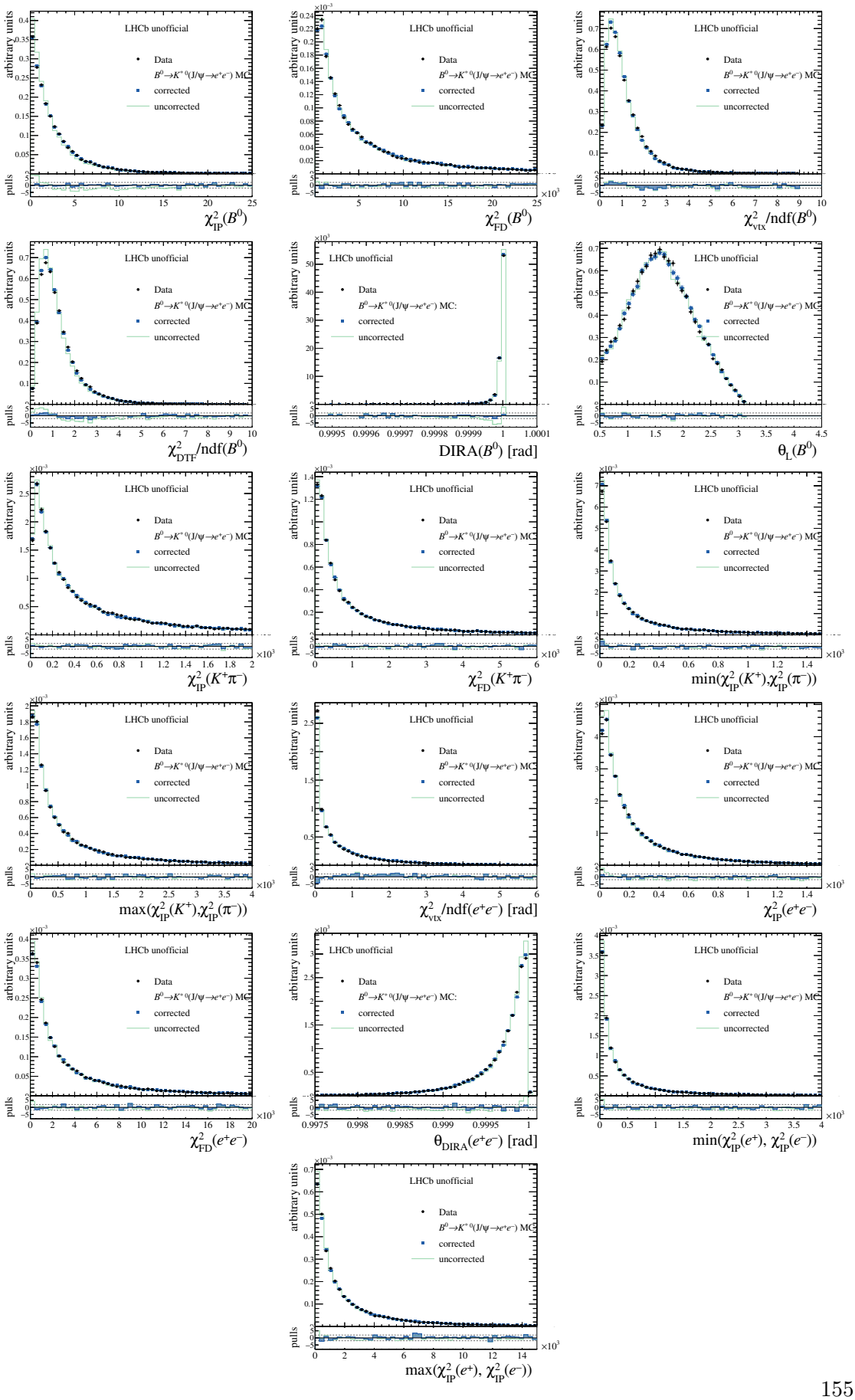

Figure C.11.: Reconstruction variables.

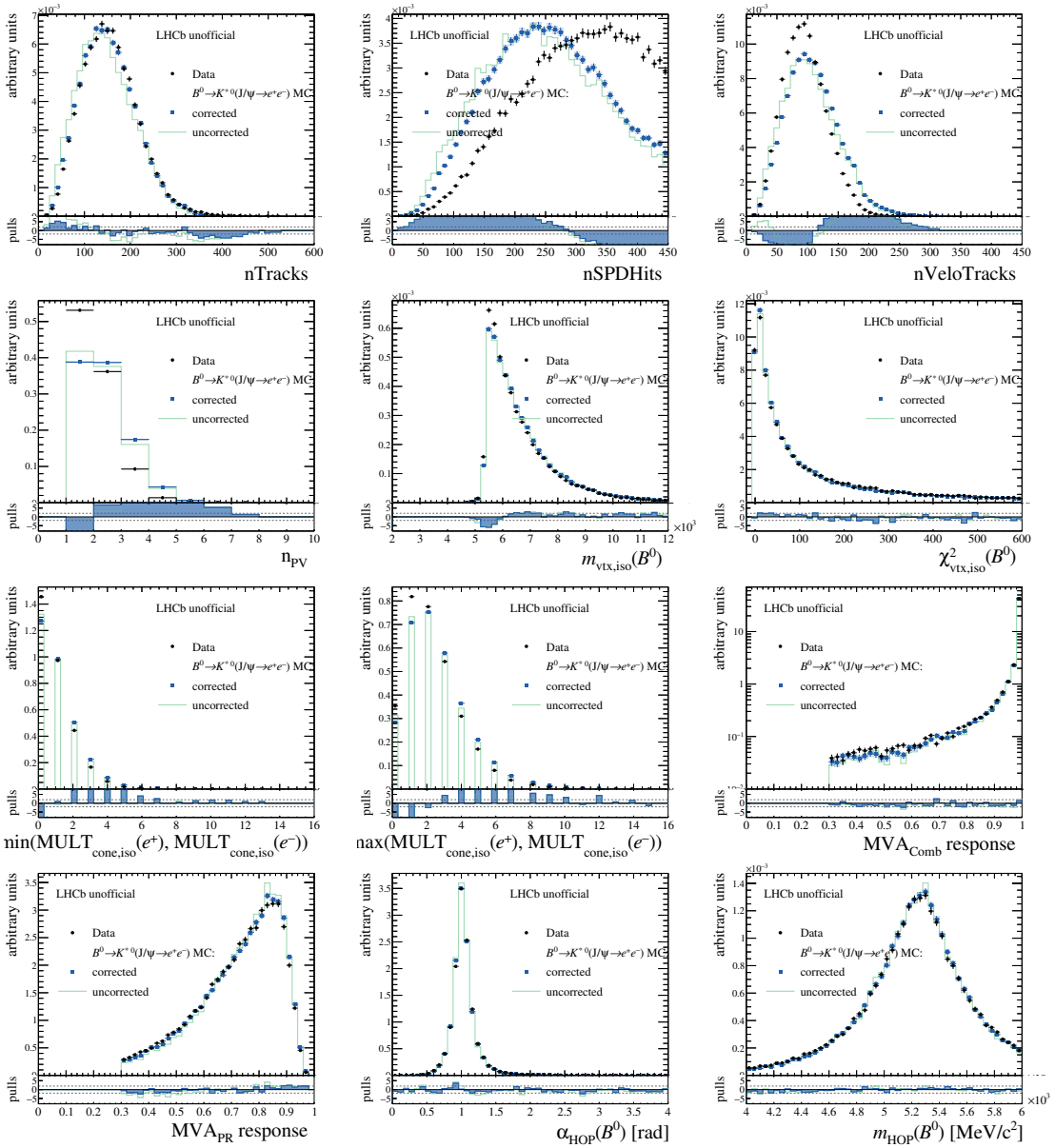

Figure C.12.: Multiplicity, isolation, MVA and HOP variables.

D Fit to the Resonant Channels

The $r^{J/\psi}$ and $R^{\psi(2S)}$ cross-checks, which are reported in Sec. 4.8, are based on a simultaneous fit to the J/ψ and $\psi(2S)$ resonant modes. A mass constraint of the lepton system to the known values of $m(J/\psi)$ and $m(\psi(2S))$ is used to improve the resolution of the reconstructed B mass. The sizes of both q^2 regions and B mass windows used in the fit are reported in Tab. D.1. As discussed in Sec. G, the $R^{\psi(2S)}$ cross-check is performed with two different widths of the electron $\psi(2S)$ q^2 window to test the control over backgrounds in this channel.

Table D.1.: Summary of the mass windows used for the $r^{J/\psi}$ and $R^{\psi(2S)}$ cross-checks.

Type	**Applied to**	**Requirement**
q^2	$J/\psi\,(\mu^+\mu^-)$	$\lvert m(\mu^+\mu^-) - m^{\text{PDG}}_{J/\psi}\rvert < 100\,\text{MeV}/c^2$
	$J/\psi\,(e^+e^-)$	6.0–11.0 GeV^2/c^4
	$\psi(2S)(\mu^+\mu^-)$	$\lvert m(\mu^+\mu^-) - m^{\text{PDG}}_{\psi(2S)}\rvert < 100\,\text{MeV}/c^2$
	$\psi(2S)(e^+e^-)$	11.0–15.0 GeV^2/c^4
	$\psi(2S)(e^+e^-)$ (wide)	9.9–16.4 GeV^2/c^4
$m(B)$	$J/\psi\,(\mu^+\mu^-)$	5100–6100 MeV/c^2
	$J/\psi\,(e^+e^-)$	4900–6200 MeV/c^2
	$\psi(2S)(\mu^+\mu^-)$	5100–5750 MeV/c^2
	$\psi(2S)(e^+e^-)$	5100–5750 MeV/c^2
$m(K^{*0})$	B^0 modes	$\lvert m(K\pi) - m^{\text{PDG}}_{K^{*0}}\rvert < 100\,\text{MeV}/c^2$

The strategy is analogous to the fit of the J/ψ modes in the simultanoues fit for the R_K and $R_{K^{*0}}$ observables. Again, Gaussian and CB functions are used to model the signal mass shape for the electron modes from simulated samples, while a Hypatia [119] function is used to describe the shape of the muon modes. The same backgrounds, that are considered in the fits to the J/ψ mode, are modelled in the $\psi(2S)$ fits. However, two additional backgrounds are described in the $\psi(2S)(e^+e^-)$ fits: First, the leakage from the prominent J/ψ resonance into the $\psi(2S)$ q^2 window is described with a `RooKeysPDF`. Second, the charmonium partially reconstructed backgrounds are modelled individually, since the lost energy of the hadrons from the decay of $c\bar{c}$ resonances (*e.g.* $\psi(2S) \to$

$\pi\pi J/\psi$) is recovered by the $\psi(2S)$ mass constraint. Thus, these modes peak under the signal region making them particularly dangerous. Also this background component is described with a `RooKeysPDF` fixed from a fit to simulated samples of these background modes.

The plots of the mass fits are given in Fig. D.1 and D.2 for muon and electron modes, respectively. The yields resulting from the simultaneous fit are given in Tab. D.2, while the values of $r^{J/\psi}$ and $R^{\psi(2S)}$ are given in Sec. 4.8.

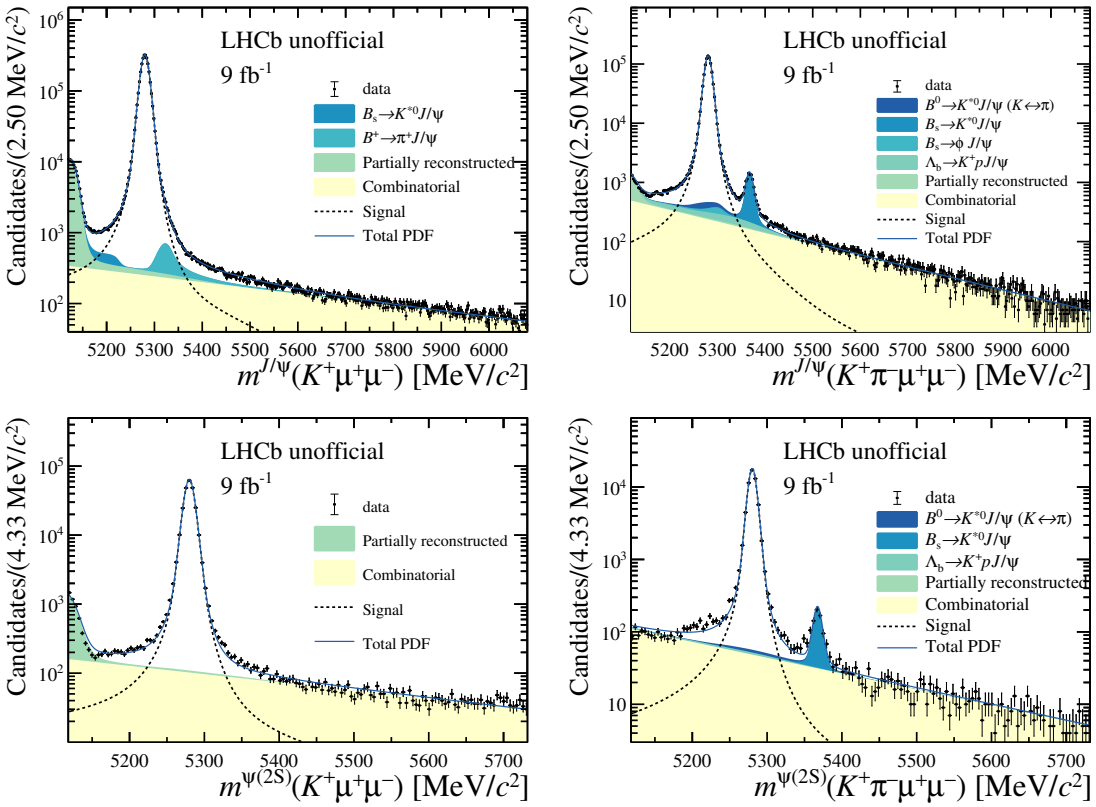

Figure D.1.: Result of the muon channels from the simultaneous fit to the resonant modes in the full Run 1 and Run 2 dataset. The plots result from the sum of data samples and PDFs over the individual run periods and trigger categories. The J/ψ and $\psi(2S)$ q^2 regions are shown from top to bottom and the left (right) plots show the B^+ (B^0) modes. The various components used in the fit are detailed in the legends. Notice that the plots are shown with a logarithmic scale on the y-axis.

To evaluate the systematic uncertainty for the fit model for the J/ψ mode fits, the fits are performed in four setups which differ in selection, mass range or whether a mass constraint to the dilepton system is applied. The setups are defined in Sec. 4.7.2, where the yields from setups 3 and 4 are given in Tab. D.3.

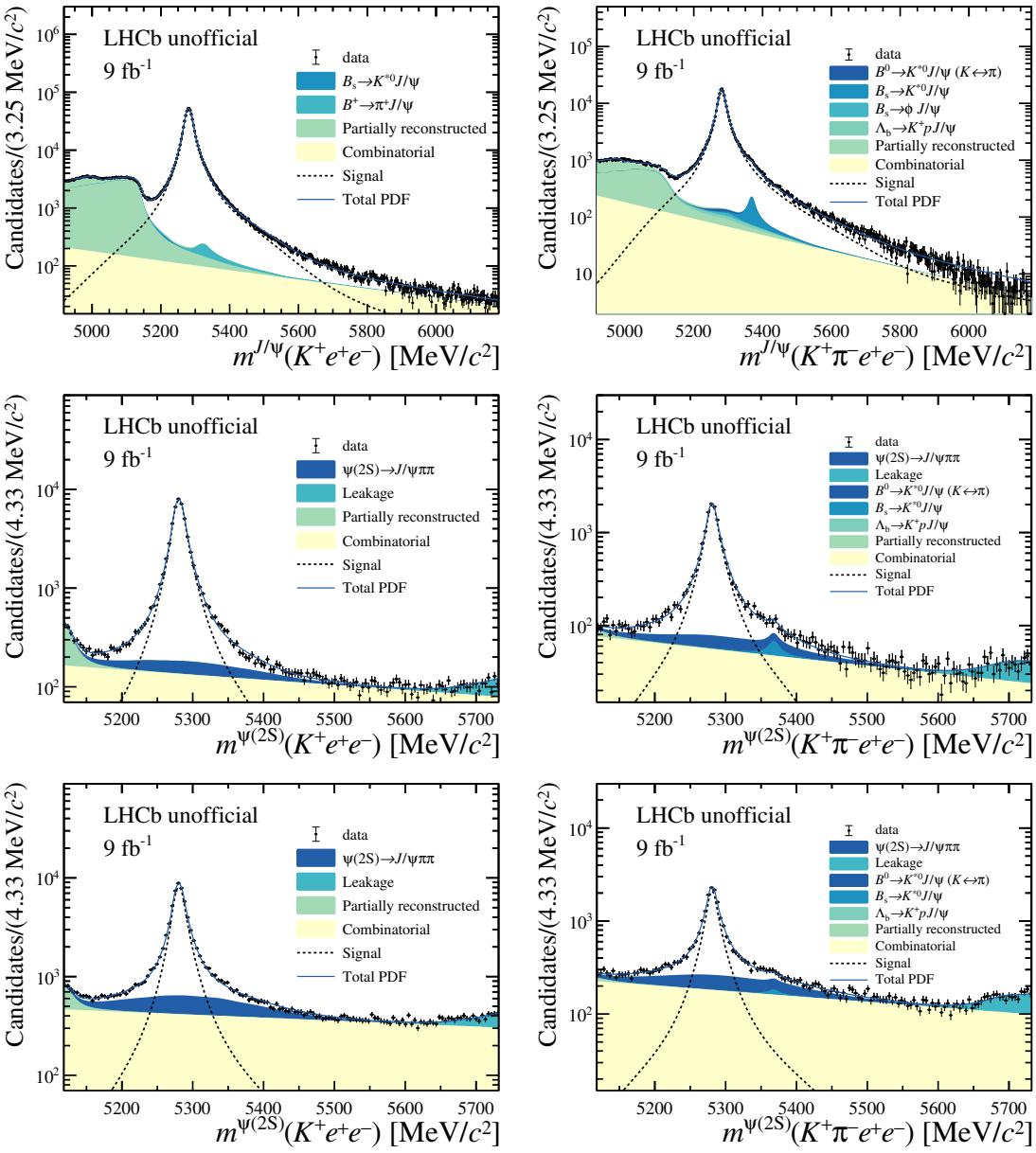

Figure D.2.: Result of the electron channels from the simultaneous fit to the resonant modes in the full Run 1 and Run 2 dataset. The plots result from the sum of data samples and PDFs over the individual run periods and trigger categories. The J/ψ, $\psi(2S)$ and wide $\psi(2S)$ q^2 regions are shown from top to bottom and the left (right) plots show the B^+ (B^0) modes. The various components used in the fit are detailed in the legends. Notice that the plots are shown with a logarithmic scale on the y-axis.

Table D.2.: Mass yields from the simultaneous fit to the resonant modes in all q^2 regions, run periods and trigger categories.

Mode	Run period	$\mathcal{N}(\mu^+\mu^-)$ L0I	$\mathcal{N}(\mu^+\mu^-)$ L0L	$\mathcal{N}(e^+e^-)$ L0I	$\mathcal{N}(e^+e^-)$ L0L
$B^+ \to K^+ J/\psi (\to \ell^+\ell^-)$	Run 1	198420 ± 451	556610 ± 755	68718 ± 287	74922 ± 301
	Run 2p1	208460 ± 463	521840 ± 732	90405 ± 326	107420 ± 362
	Run 2p2	396570 ± 638	1074800 ± 1050	174290 ± 457	205130 ± 504
$B^+ \to K^+ \psi(2S) (\to \ell^+\ell^-)$	Run 1	15996 ± 128	47185 ± 220	4558 ± 86	6140 ± 96
	Run 2p1	16714 ± 131	44249 ± 213	6405 ± 99	8628 ± 111
	Run 2p2	31833 ± 181	90511 ± 305	11773 ± 135	16195 ± 150
$B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$	Run 1	65946 ± 261	172640 ± 425	24301 ± 207	26792 ± 201
	Run 2p1	72463 ± 273	169040 ± 417	31810 ± 229	37226 ± 239
	Run 2p2	138640 ± 378	349510 ± 601	60803 ± 319	70305 ± 328
$B^0 \to K^{*0} \psi(2S) (\to \ell^+\ell^-)$	Run 1	4018 ± 64	10532 ± 104	1238 ± 48	1712 ± 54
	Run 2p1	4340 ± 67	10319 ± 103	1623 ± 54	2235 ± 63
	Run 2p2	8007 ± 91	21209 ± 148	3217 ± 75	4080 ± 81

Table D.3.: Mass yields from the simultaneous fit to the J/ψ resonant modes in all run periods and trigger categories. The yields are given for the setups 3 and 4 of the systematic uncertainty evaluation of the J/ψ mode fit model (defined in Sec. 4.7.2).

Mode	Run period	$\mathcal{N}(\mu^+\mu^-)$ L0I	$\mathcal{N}(\mu^+\mu^-)$ L0L	$\mathcal{N}(e^+e^-)$ L0I	$\mathcal{N}(e^+e^-)$ L0L
Setup 3					
$B^+ \to K^+ J/\psi (\to \ell^+\ell^-)$	Run 1	199060 ± 458	557680 ± 767	66481 ± 265	72074 ± 288
	Run 2p1	209030 ± 469	523490 ± 740	86485 ± 304	103490 ± 333
	Run 2p2	397560 ± 646	1077600 ± 1060	168080 ± 424	199160 ± 463
$B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$	Run 1	66280 ± 271	173550 ± 441	23360 ± 168	26189 ± 175
	Run 2p1	72921 ± 281	169980 ± 429	30898 ± 185	35767 ± 201
	Run 2p2	139500 ± 390	351160 ± 617	58742 ± 257	68402 ± 275
Setup 4					
$B^+ \to K^+ J/\psi (\to \ell^+\ell^-)$	Run 1	198470 ± 451	556780 ± 756	68456 ± 291	74626 ± 305
	Run 2p1	208490 ± 408	521870 ± 733	88897 ± 286	106580 ± 356
	Run 2p2	396680 ± 639	1074800 ± 1050	172540 ± 484	204470 ± 493
$B^0 \to K^{*0} J/\psi (\to \ell^+\ell^-)$	Run 1	65989 ± 262	172770 ± 426	24465 ± 183	27194 ± 187
	Run 2p1	72483 ± 274	169140 ± 418	31931 ± 209	37339 ± 220
	Run 2p2	138770 ± 379	349500 ± 602	60780 ± 322	70918 ± 302

E Flatness of $r^{J/\psi}$

In order to test the stability of the ratio measurements, the $r^{J/\psi}$ cross-check reported in Sec. 4.8.1 is performed differentially in bins of various variables that are of importance to the analysis. Therefore, both efficiencies and mass fits are calculated separately for each bin. Figures E.1 to E.4 show the result of this test using the Run 2p2 dataset as an example. The plots show that the flatness is greatly improved when the full MC correction chain is applied. Again, the good portability of the corrections based on B^+ and B^0 calibration samples is visible when comparing the results from $w(B^+)$ with $w(B^0)$ corrections.

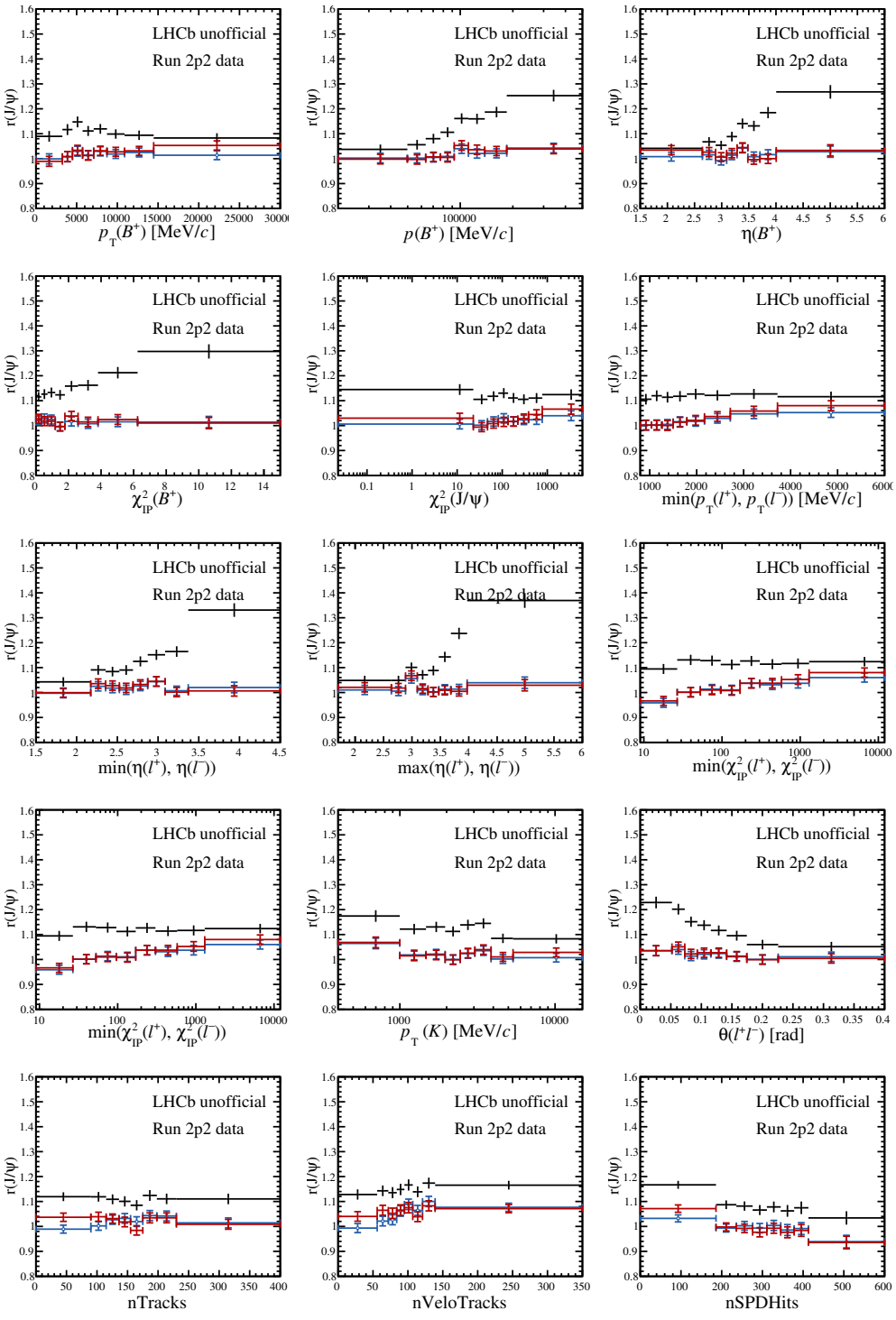

Figure E.1.: $r^{J/\psi}$ flatness check in various key variables fo the B^+ modes selected by the L0I category. The black distributions shows the results based on uncorrected MC samples. The results from fully calibrated efficiencies are shown in blue for $w(B^+)$ and in red for $w(B^0)$.

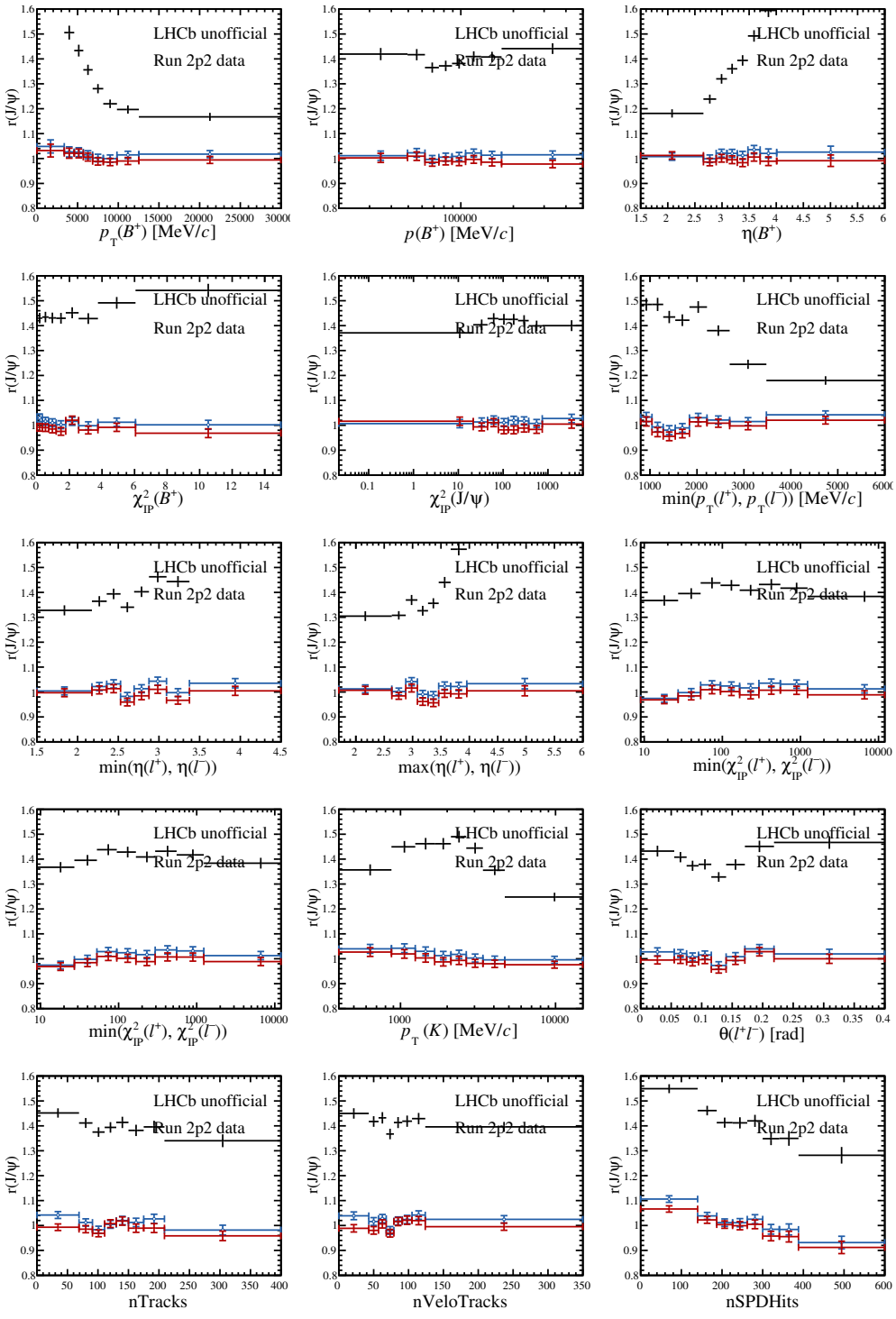

Figure E.2.: $r^{J/\psi}$ flatness check in various key variables fo the B^+ modes selected by the L0L category. The black distributions shows the results based on uncorrected MC samples. The results from fully calibrated efficiencies are shown in blue for $w(B^+)$ and in red for $w(B^0)$.

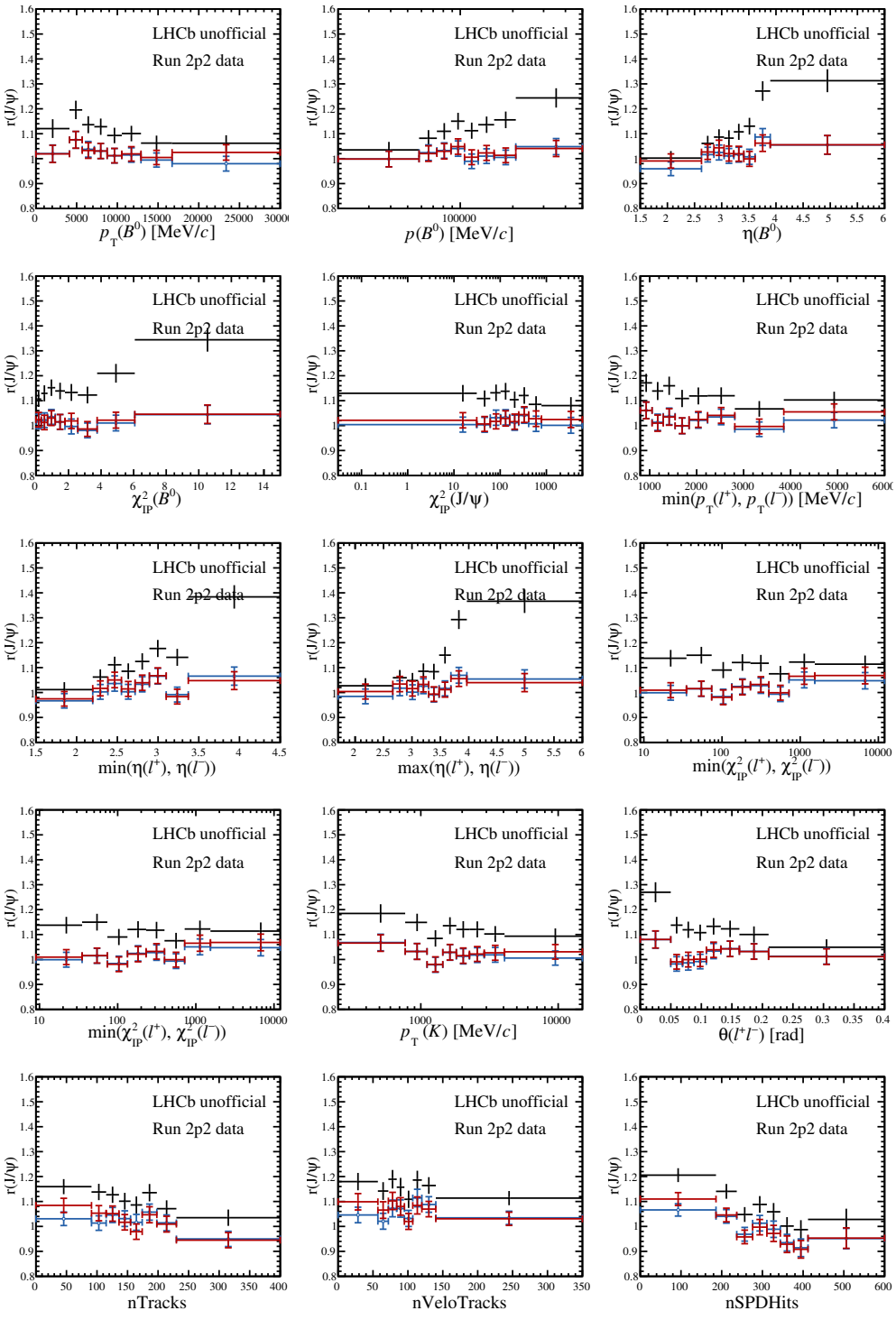

Figure E.3.: $r^{J/\psi}$ flatness check in various key variables fo the B^0 modes selected by the L0I category. The black distributions shows the results based on uncorrected MC samples. The results from fully calibrated efficiencies are shown in blue for $w(B^+)$ and in red for $w(B^0)$.

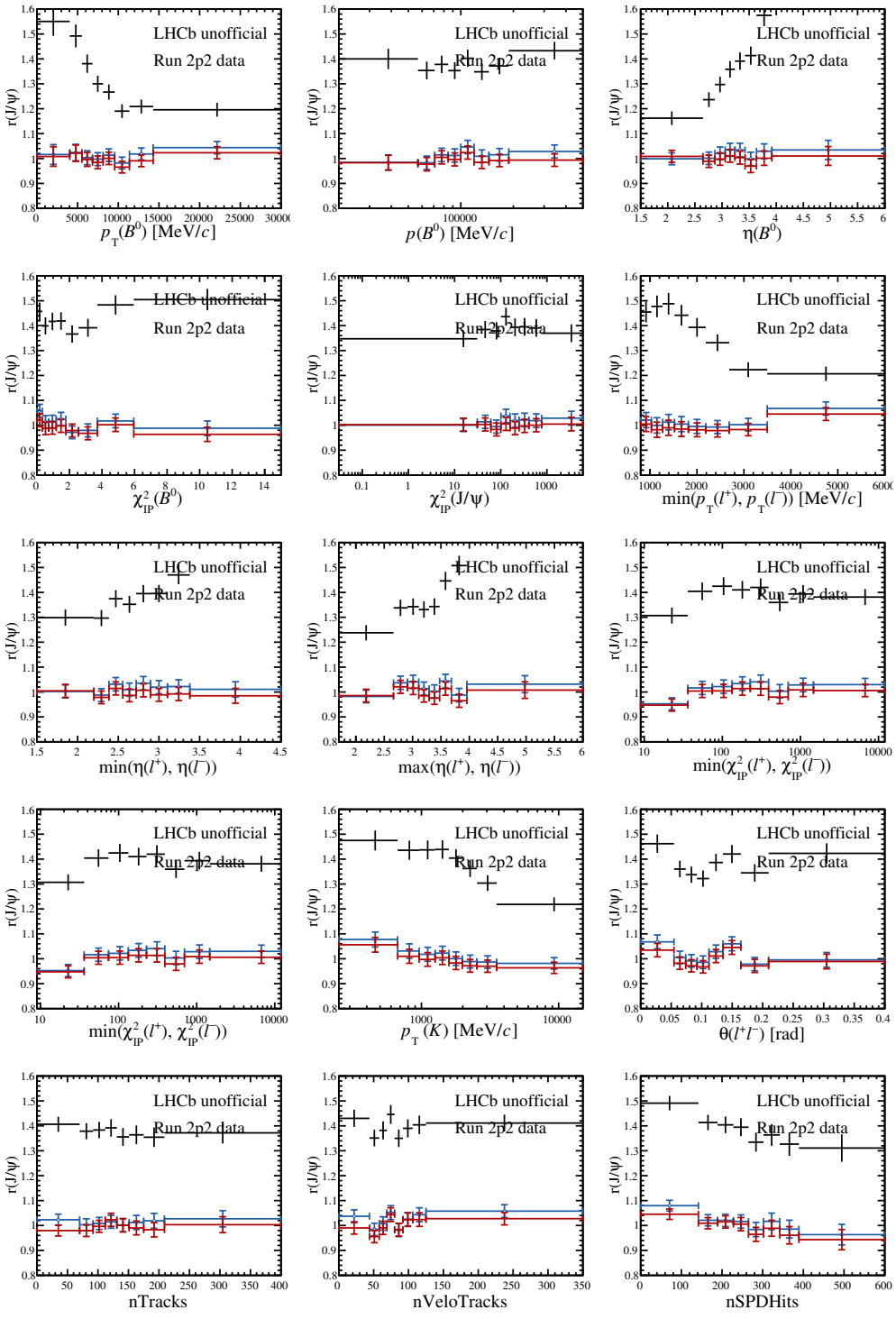

Figure E.4.: $r^{J/\psi}$ flatness check in various key variables fo the B^0 modes selected by the L0L category. The black distributions shows the results based on uncorrected MC samples. The results from fully calibrated efficiencies are shown in blue for $w(B^+)$ and in red for $w(B^0)$.

F Flatness Systematic

The values for the flatness parameter d_f, which is introduced in Sec. 4.7.1, is shown for several key variables in Figs. F.1 to F.4. The flatness parameter is calculated with the full MC calibration applied based on both the $w(B^+)$ and $w(B^0)$ correction chains. The figures show the d_f values for all four studied LFU bins and in both L0 trigger categories. In general, the size of the flatness parameter is much reduced after the full MC correction chain is applied.

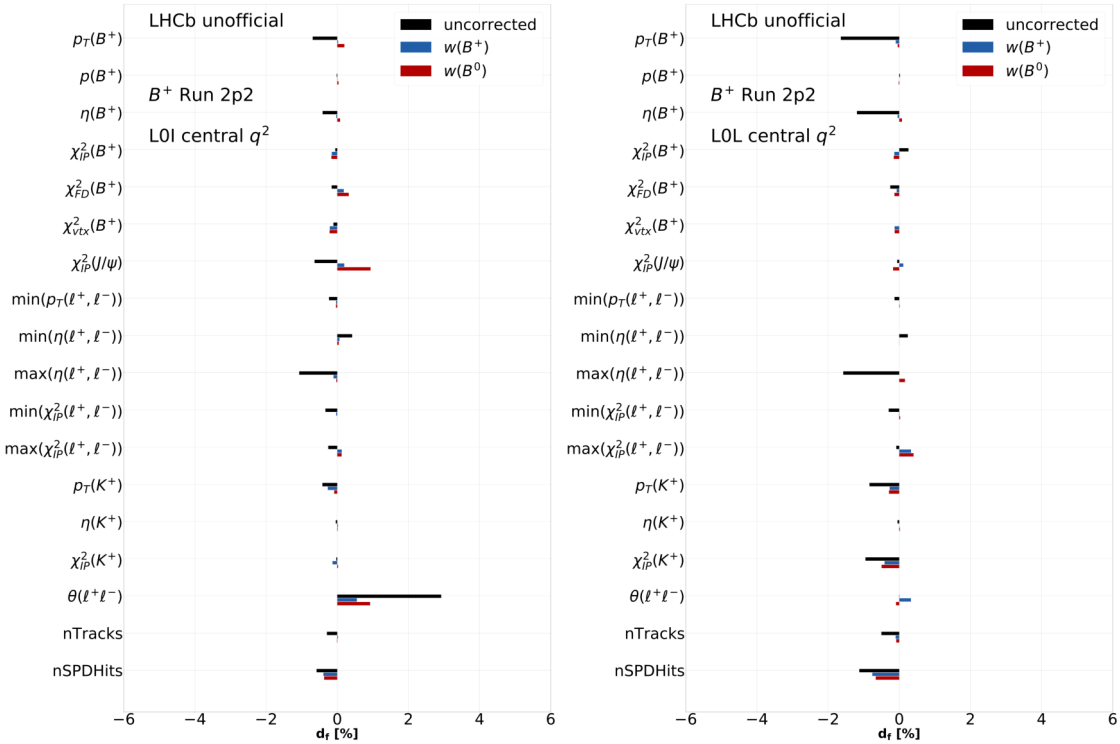

Figure F.1.: The values of the flatness parameter d_f for several key variables evaluated for the R_K double ratio in central q^2 based on Run 2p2 samples. The left (right) plot shows the L0I (L0L) trigger category. The values are given in percent relative to the measured central value of the LFU ratio.

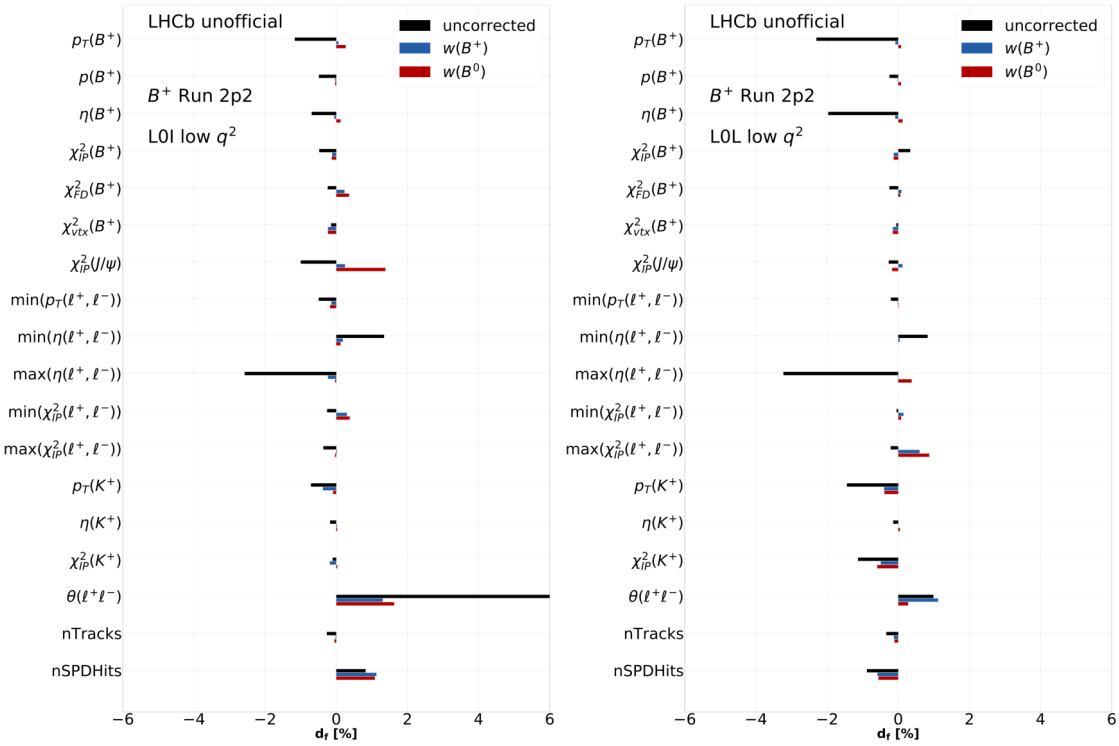

Figure F.2.: The values of the flatness parameter d_f for several key variables evaluated for the R_K double ratio in low q^2 based on Run 2p2 samples. The left (right) plot shows the L0I (L0L) trigger category. The values are given in percent relative to the measured central value of the LFU ratio.

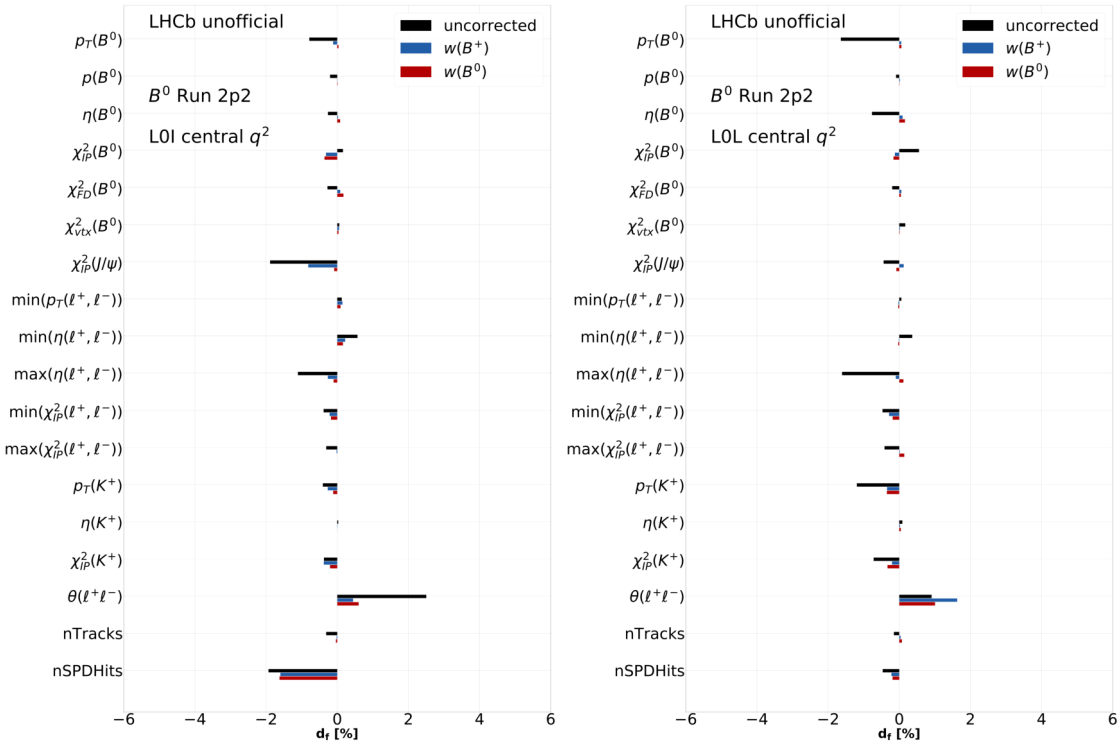

Figure F.3.: The values of the flatness parameter d_f for several key variables evaluated for the $R_{K^{*0}}$ double ratio in central q^2 based on Run 2p2 samples. The left (right) plot shows the L0I (L0L) trigger category. The values are given in percent relative to the measured central value of the LFU ratio.

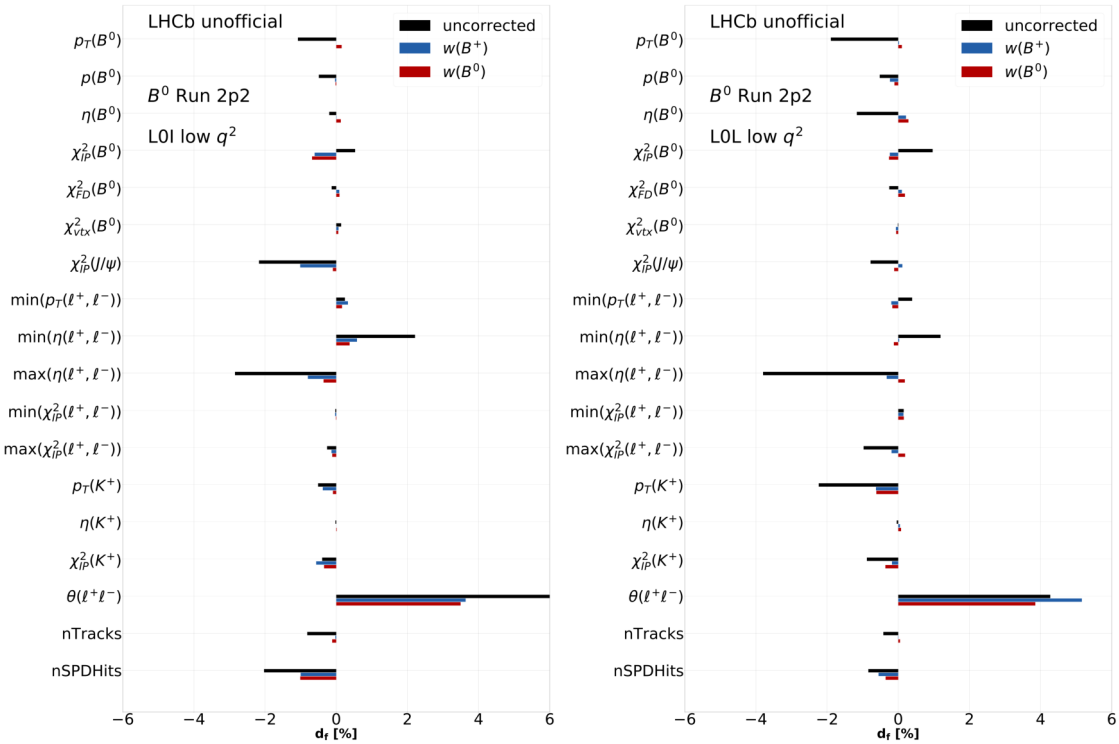

Figure F.4.: The values of the flatness parameter d_f for several key variables evaluated for the $R_{K^{*0}}$ double ratio in low q^2 based on Run 2p2 samples. The left (right) plot shows the L0I (L0L) trigger category. The values are given in percent relative to the measured central value of the LFU ratio.

G $R^{\psi(2S)}$ in a Wide q^2 Range

As an additional cross-check the $R^{\psi(2S)}$ double ratio is calculated using a larger q^2 window for the $\psi(2S)(e^+e^-)$ mode (9.9–16.4 GeV^2/c^4). The wider q^2 range significantly increases the background contamination, especially from $\psi(2S) \to J/\psi X$ decays. Thus, the control over the modelling and composition of the backgrounds considered in the mass fit is tested. The comparison of $R^{\psi(2S)}$ between the nominal and wide q^2 results can be seen in Tab. G.1. The rows called 'Combination' show the result from a simultaneous fit to the full run period or even the full dataset, instead of fitting the individual categories.

Table G.1.: Values of the $R^{\psi(2S)}$ ratios after applying the full MC calibration for fits to the nominal and wide q^2 windows of the $\psi(2S)(e^+e^-)$ mode.

Mode	Run Period	L0 Category	Nominal q^2	Wide q^2
B^+	Run 1	L0I	0.997 ± 0.021	0.997 ± 0.023
		L0L	0.982 ± 0.017	0.983 ± 0.018
		Combination	0.988 ± 0.013	0.989 ± 0.014
	Run 2p1	L0I	0.947 ± 0.017	0.973 ± 0.019
		L0L	0.987 ± 0.015	1.000 ± 0.015
		Combination	0.971 ± 0.011	0.990 ± 0.012
	Run 2p2	L0I	0.994 ± 0.014	0.993 ± 0.014
		L0L	1.000 ± 0.011	1.011 ± 0.011
		Combination	0.998 ± 0.008	1.004 ± 0.009
	All Runs	Combination	0.987 ± 0.006	0.996 ± 0.006
B^0	Run 1	L0I	1.052 ± 0.045	1.037 ± 0.048
		L0L	0.987 ± 0.034	0.999 ± 0.037
		Combination	1.013 ± 0.027	1.015 ± 0.029
	Run 2p1	L0I	1.030 ± 0.040	1.052 ± 0.045
		L0L	0.991 ± 0.030	1.003 ± 0.032
		Combination	1.006 ± 0.024	1.022 ± 0.026
	Run 2p2	L0I	0.955 ± 0.026	0.940 ± 0.028
		L0L	1.059 ± 0.024	1.045 ± 0.025
		Combination	1.017 ± 0.018	1.004 ± 0.018
	All Runs	Combination	1.015 ± 0.013	1.012 ± 0.013

Bibliography

[1] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961. 10.1016/0029-5582(61)90469-2. *cited on pages 1, 3.*

[2] Abdus Salam and John Clive Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13:168–171, 1964. 10.1016/0031-9163(64)90711-5. *cited on pages 1, 3.*

[3] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967. 10.1103/PhysRevLett.19.1264. *cited on pages 1, 3.*

[4] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):129, Sep 2012. 10.1016/J.PhysLetB.2012.08.020. *cited on pages 1, 7.*

[5] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):3061, Sep 2012. 10.1016/J.PhysLetB.2012.08.021. *cited on pages 1, 7.*

[6] The ATLAS Collaboration. Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector. *Phys. Rev. D*, 104(11), Dec 2021. 10.1103/physrevd.104.112005. *cited on page 1.*

[7] The CMS Collaboration. Search for new particles in events with energetic jets and large missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 2021(11), Nov 2021. 10.1007/jhep11(2021)153. *cited on page 1.*

[8] Andrzej J. Buras and Jennifer Girrbach. Towards the identification of new physics through quark flavour violating processes. *Reports on Progress in Physics*, 77(8):086201, Aug 2014. 10.1088/0034-4885/77/8/086201. *cited on page 1.*

[9] Marzia Bordone, Gino Isidori, and Andrea Pattori. On the standard model predictions for R_K and R_{K^*}. *The European Physical Journal C*, 76(8), Aug 2016. 10.1140/epjc/s10052-016-4274-7. *cited on pages 1, 11.*

[10] LHCb collaboration. Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays. *Phys. Rev. Lett.*, 113(15), Oct 2014. 10.1103/physrevlett.113.151601. *cited on page 1.*

[11] LHCb collaboration. Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays. *JHEP*, 2017(8), Aug 2017. 10.1007/jhep08(2017)055. *cited on pages 2, 11, 12, 12, 44, 50, 53, 66, 138.*

[12] LHCb collaboration. Test of lepton universality with $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays. *JHEP*, 2020(5), May 2020. 10.1007/jhep05(2020)040. *cited on pages 2, 12.*

[13] LHCb collaboration. Test of lepton universality in beauty-quark decays. *Nature*

Physics, Mar 2022. 10.1038/s41567-021-01478-8. *cited on pages 2, 11, 12, 12, 44, 128, 130, 138.*

[14] A. Angelescu, D. Beirevi, D. A. Faroughy, and O. Sumensari. Closing the window on single leptoquark solutions to the B-physics anomalies. *JHEP*, 2018(10), Oct 2018. 10.1007/JHEP10(2018)183. *cited on pages 2, 12.*

[15] Andrzej J. Buras and Jennifer Girrbach. Left-handed Z' and Z FCNC quark couplings facing new $b \to s\mu^+\mu^-$ data. *JHEP*, 2013(12), Dec 2013. 10.1007/jhep12(2013)009. *cited on pages 2, 12.*

[16] Guido Altarelli. The Standard Model of Particle Physics, 2015. arXiv:hep-ph/0510281. *cited on page 3.*

[17] Yuval Grossman. Introduction to flavor physics, 2010. arXiv:1006.3534. *cited on page 3.*

[18] David Griffiths. *Introduction to elementary particles*. WILEY-VCH, 2008. *cited on page 3.*

[19] Monika Blanke. Introduction to Flavour Physics and CP Violation, 2017. arXiv:1704.03753. *cited on page 3.*

[20] Wolfgang Pauli. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Zeitschrift für Physik*, 31:765–783, 1925. 10.1007/BF02980631. *cited on page 3.*

[21] Carl D. Anderson. The Apparent Existence of Easily Deflectable Positives. *Science*, 76(1967):238–239, 1932. 10.1126/science.76.1967.238. *cited on page 4.*

[22] R. M. Barnett et al. Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, Aug 2020. 10.1093/ptep/ptaa104. *cited on pages 4, 4, 38, 41, 41, 54, 58, 61, 75, 93.*

[23] J. P. Ellis. TikZ-Feynman: Feynman diagrams with TikZ. *Computer Physics Communications*, 210:103123, Jan 2017. 10.1016/j.cpc.2016.08.019. *cited on page 5.*

[24] Aoife Bharucha, David M. Straub, and Roman Zwicky. $B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules. *JHEP*, 2016(8), Aug 2016. 10.1007/jhep08(2016)098. *cited on page 5.*

[25] C. Bouchard at al. Rare decay $B \to K\ell^+\ell^-$ form factors from lattice QCD. *Phys. Rev. D*, 88(5), Sep 2013. 10.1103/physrevd.88.054509. *cited on page 5.*

[26] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys. Rev.*, 127:965–970, 1962. 10.1103/PhysRev.127.965. *cited on page 7.*

[27] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. 10.1103/PhysRevLett.13.508. *cited on page 7.*

[28] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964. 10.1103/PhysRevLett.13.585. *cited on page 7.*

[29] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. 10.1103/PhysRevLett.13.321. *cited on page 7.*

[30] Y. Ashie et al. Evidence for an Oscillatory Signature in Atmospheric Neutrino Oscillations. *Phys. Rev. Lett.*, 93:101801, Sep 2004. 10.1103/PhysRevLett.93.101801. *cited on pages 7, 9.*

[31] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D*, 2:1285–1292, Oct 1970. 10.1103/PhysRevD.2.1285. *cited on pages 8, 9.*

[32] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, Feb 1973. 10.1143/PTP.49.652. *cited on page 8.*

[33] J. J. Aubert et al. Experimental Observation of a Heavy Particle J. *Phys. Rev. Lett.*, 33:1404–1406, Dec 1974. 10.1103/PhysRevLett.33.1404. *cited on page 9.*

[34] LHCb collaboration. Measurement of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction. *JHEP*, Nov 2016. 10.1007/JHEP11(2016)047. *cited on pages 10, 11, 106, 106, 129, 129.*

[35] LHCb collaboration. Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays. *JHEP*, Jun 2014. 10.1007/JHEP06(2014)133. *cited on pages 10, 11, 12, 106, 129.*

[36] LHCb collaboration. Branching fraction measurements of the rare $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to f_2'(1525)\mu^+\mu^-$ decays, 2021. arXiv:2105.14007. *cited on page 10.*

[37] S. Bifani, S. Descotes-Genon, A. R. Vidal, and M.-H. Schune. Review of lepton universality tests in B decays. *Journal of Physics G: Nuclear and Particle Physics*, 46(2):023001, Dec 2018. 10.1088/1361-6471/aaf5de. *cited on page 11.*

[38] Tobias Huber, Enrico Lunghi, Mikoaj Misiak, and Daniel Wyler. Electromagnetic logarithms in $B \to X_s\ell^+\ell^-$. *Nuclear Physics B*, 740(1-2):105137, Apr 2006. 10.1016/j.nuclphysb.2006.01.037. *cited on page 11.*

[39] Marzia Bordone, Gino Isidori, and Andrea Pattori. On the standard model predictions for R_K and R_{K^*}. *The European Physical Journal C*, 76(8), Aug 2016. 10.1140/epjc/s10052-016-4274-7. *cited on pages 12, 12, 39, 133, 138.*

[40] B. Capdevila, S. Descotes-Genon, L. Hofer, and J. Matias. Hadronic uncertainties in $B \to K^*\mu^+\mu^-$: a state-of-the-art analysis. *JHEP*, 2017(4), Apr 2017. 10.1007/jhep04(2017)016. *cited on pages 12, 133, 138.*

[41] Belle Collaboration. Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for $B \to K^{(0)}\ell^+\ell^-$. *Phys. Rev. Lett.*, 103(17), Oct 2009. 10.1103/PhysRevLett.103.171801. *cited on page 12.*

[42] BaBar Collaboration. Measurement of branching fractions and rate asymmetries in the rare decays $B \to K^{(0)}\ell^+\ell^-$. *Phys. Rev. D*, 86(3), Aug 2012. 10.1103/Phys-

RevD.86.032012. *cited on page 12.*

[43] LHCb collaboration. Measurements of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\pi^-\mu^+\mu^-$ differential branching fraction. *JHEP*, 11(11), Jun 2016. 10.1007/JHEP11(2016)047. *cited on pages 12, 50.*

[44] LHCb Collaboration. Branching Fraction Measurements of the Rare $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to f_2'(1525)\mu^+\mu^-$ Decays. *Phys. Rev. Lett.*, 127:151801, Oct 2021. 10.1103/PhysRevLett.127.151801. *cited on page 12.*

[45] LHCb Collaboration. Measurement of CP-Averaged Observables in the $B^0 \to K^{*0}\mu^+\mu^-$ Decay. *Phys. Rev. Lett.*, 125:011802, Jul 2020. 10.1103/PhysRevLett.125.011802. *cited on page 12.*

[46] LHCb Collaboration. Angular Analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ Decay. *Phys. Rev. Lett.*, 126:161802, Apr 2021. 10.1103/PhysRevLett.126.161802. *cited on page 12.*

[47] LHCb Collaboration. Angular analysis of the rare decay $B_s^0 \to \phi\mu^+\mu^-$. *JHEP*, 2021(11), Nov 2021. 10.1007/jhep11(2021)043. *cited on page 12.*

[48] LHCb Collaboration. Analysis of Neutral B-Meson Decays into Two Muons. *Phys. Rev. Lett.*, 128:041801, Jan 2022. 10.1103/PhysRevLett.128.041801. *cited on page 12.*

[49] LHCb Collaboration. Tests of Lepton Universality Using $B^0 \to K_s^0\ell^+\ell^-$ and $B^+ \to K^{*+}\ell^+\ell^-$ decays. *Phys. Rev. Lett.*, 128(19), May 2022. 10.1103/physrevlett.128.191802. *cited on page 12.*

[50] Gerhard Buchalla, Andrzej J. Buras, and Markus E. Lautenbacher. Weak decays beyond leading logarithms. *Reviews of Modern Physics*, 68(4):11251244, Oct 1996. 10.1103/revmodphys.68.1125. *cited on page 13.*

[51] T. Mannel. *Effective Field Theories in Flavour Physics*. Springer, 2004. *cited on page 13.*

[52] Gudrun Hiller and Ivan Nisandzic. R_K and R_{K^*} beyond the standard model. *Phys. Rev. D*, 96(3):035003, 2017. *cited on page 13.*

[53] LHCb collaboration. Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics. *JHEP*, 2013(4), Apr 2013. 10.1007/jhep04(2013)001. *cited on page 15.*

[54] LHCb collaboration. Measurement of b hadron fractions in 13 TeV pp collisions. *Phys. Rev. D*, 100(3), Aug 2019. 10.1103/physrevd.100.031102. *cited on page 15, 15.*

[55] LHCb collaboration. Measurement of b hadron production fractions in 7 TeV pp collisions. *Phys. Rev. D*, 85(3), Feb 2012. 10.1103/physrevd.85.032008. *cited on page 15.*

[56] LHCb collaboration. Christian Elsässer. $\bar{b}b$ production angle plots. lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html, visited at 04.10.21. *cited on page 15.*

[57] LHCb collaboration. The LHCb Detector at the LHC. *JINST*, 3, 2008. 10.1088/1748-0221/3/08/S08005. *cited on pages 17, 20, 21.*

[58] LHCb collaboration. *LHCb reoptimized detector design and performance: Technical Design Report*. Technical design report. LHCb. CERN, Geneva, 2003. 10.1088/1748-0221/3/03/630827. *cited on pages 17, 19.*

[59] LHCb collaboration. LHCb detector performance. *International Journal of Modern Physics A*, 30(07):1530022, Mar 2015. 10.1142/s0217751x15300227. *cited on pages 17, 21, 23, 23, 34.*

[60] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08), Aug 2008. 10.1088/1748-0221/3/08/s08001. *cited on pages 17, 18.*

[61] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003–S08003, Aug 2008. 10.1088/1748-0221/3/08/s08003. *cited on page 17.*

[62] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, Aug 2008. 10.1088/1748-0221/3/08/s08004. *cited on pages 17, 139.*

[63] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002–S08002, Aug 2008. 10.1088/1748-0221/3/08/s08002. *cited on page 17.*

[64] LHCb collaboration. LHCb performance plots. lbgroups.cern.ch/public/OperationsPlots, visited at 04.10.21. *cited on page 18.*

[65] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, 262(2):444–450, 1987. 10.1016/0168-9002(87)90887-4. *cited on page 23.*

[66] Marcin Kucharczyk, Piotr Morawski, and Mariusz Witek. Primary Vertex Reconstruction at LHCb. Technical report, CERN, Geneva, Sep 2014. CERN-LHCb-PUB-2014-044. *cited on page 24.*

[67] P. A. Cherenkov. Visible Radiation Produced by Electrons Moving in a Medium with Velocities Exceeding that of Light. *Phys. Rev.*, 52:378–379, Aug 1937. 10.1103/PhysRev.52.378. *cited on page 24.*

[68] LHCb RICH group. Performance of the LHCb RICH detector at the LHC. *Eur. Phys. J. C*, 73:2431. 25 p, Nov 2012. 10.1140/epjc/s10052-013-2431-9. *cited on pages 25, 26.*

[69] LHCb. D. Derkach et al. Machine-Learning-based global particle-identification algorithms at the LHCb experiment. *Journal of Physics: Conference Series*, 1085:042038, Sep 2018. 10.1088/1742-6596/1085/4/042038. *cited on page 26.*

[70] LHCb collaboration. PID Plots for Conference. twiki.cern.ch/twiki/bin/view/LHCb/PIDConferencePlots, visited at 04.10.21. *cited on page 26.*

[71] LHCb Collaboration. *LHCb calorimeters: Technical Design Report*. Technical design report. LHCb. CERN, Geneva, 2000. cds.cern.ch/record/494264. *cited on page 26.*

[72] C. Abellán Beteta et al. Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC, 2020. arXiv:2008.11556. *cited on pages 26, 27, 28.*

[73] A. A. Alves et al. Performance of the LHCb muon system. *Journal of Instrumentation*, 8(02):P02022P02022, Feb 2013. 10.1088/1748-0221/8/02/p02022. *cited on page 28, 28.*

[74] F. Archilli et al. Performance of the Muon Identification at LHCb. *Journal of Instrumentation*, 8(10):P10020–P10020, Oct 2013. 10.1088/1748-0221/8/10/p10020. *cited on page 28.*

[75] LHCb Starterkit. The LHCb data flow. lhcb.github.io/starterkit-lessons/first-analysis-steps/dataflow.html, visited at 13.10.21. *cited on page 29.*

[76] LHCb Collaboration. *LHCb trigger system: Technical Design Report*. Technical design report. LHCb. CERN, Geneva, 2003. CERN-LHCC-2003-031. *cited on page 30.*

[77] R. Aaij et al. The lhcb trigger and its performance in 2011. *Journal of Instrumentation*, 8(04):P04022P04022, Apr 2013. 10.1088/1748-0221/8/04/p04022. *cited on page 30.*

[78] R. Aaij et al. Design and performance of the LHCb trigger and full real-time reconstruction in run 2 of the LHC. *Journal of Instrumentation*, 14(04):P04013–P04013, Apr 2019. 10.1088/1748-0221/14/04/p04013. *cited on page 30.*

[79] LHCb Collaboration. The Moore project. lhcbdoc.web.cern.ch/lhcbdoc/moore/master/index.html, visited at 08.10.21. *cited on page 31, 31.*

[80] Vladimir V Gligorov, Christopher Thomas, and Michael Williams. The HLT inclusive B triggers. Technical report, CERN, Geneva, Sep 2011. CERN-LHCb-PUB-2011-016. *cited on page 31.*

[81] V. V. Gligorov and M. Williams. Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree. *Journal of Instrumentation*, 8(02), Feb 2013. 10.1088/1748-0221/8/02/p02013. *cited on pages 32, 44.*

[82] LHCb Collaboration. The Brunel project. lhcbdoc.web.cern.ch/lhcbdoc/brunel/, visited at 13.10.21. *cited on page 32.*

[83] LHCb Collaboration. The DaVinci project. lhcbdoc.web.cern.ch/lhcbdoc/davinci/, visited at 13.10.21. *cited on page 32.*

[84] M. Clemencic et al. The LHCb Simulation Application, Gauss: Design, Evolution and Experience. *Journal of Physics: Conference Series*, 331(3):032023, Dec 2011. 10.1088/1742-6596/331/3/032023. *cited on page 32.*

[85] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 2006(05):026–026, May 2006. 10.1088/1126-6708/2006/05/026.

cited on page 32.

[86] David J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, 462(1):152–155, 2001. 10.1016/S0168-9002(01)00089-4. *cited on pages 32, 48, 113.*

[87] P. Golonka and Z. Was. PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays. *The European Physical Journal C*, 45(1):97107, Jan 2006. 10.1140/epjc/s2005-02396-4. *cited on page 32.*

[88] LHCb Collaboration. The BOOLE project. lhcbdoc.web.cern.ch/lhcbdoc/boole/, visited at 22.10.21. *cited on page 32.*

[89] Ryan B. Calladine. *Determination of the $r^K_{J/\psi}$ and $r^{K^*}_{J/\psi}$ Ratios for Lepton Universality.* PhD thesis, University of Birmingham, 2021. etheses.bham.ac.uk/id/eprint/11577/. *cited on page 37.*

[90] Fabrice Desse. *Angular analysis of $B^0 \to K^{*0}e^+e^-$ decays with the LHCb detector and upgrade of the electronics of the calorimeters.* PhD thesis, University Paris-Saclay, 2020. https://cds.cern.ch/record/2730230. *cited on page 37.*

[91] Da Y. Tou. *Test of lepton flavour universality in $b \to s\ell\ell$ decays.* PhD thesis, Sorbonne University, 2020. http://www.theses.fr/2020SORUS409. *cited on page 37.*

[92] LHCb Collaboration. Test of lepton universality in $b \to s\ell^+\ell$ decays, 2022. in preparation (title preliminary). *cited on page 37.*

[93] Johannes Albrecht, Simone Bifani, Julian Alexander Boelhauve, Ryan Bernard Calladine, Fabrice Desse, Stephan Escher, Vladimir Gligorov, Christoph Michael Langenbruch, Vitalii Lisovskyi, Simon Nieswand, Francesco Polci, Renato Quagliani, Sebastian Schmitt, Marie-Helene Schune, Alex Seuthe, Eluned Anne Smith, Da Yu Tou, and Nigel Watson. Test of lepton flavour universality with $b \to s\ell^+\ell$ decays. *CERN Docment Server*, Nov 2020. cds.cern.ch/record/2745141. *cited on pages 37, 38, 55, 86, 87, 90, 96, 96, 111, 113, 120, 123, 130, 132.*

[94] David M. Straub. flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond, 2018. arXiv:1810.08132. *cited on pages 41, 136.*

[95] Wolfgang Altmannshofer and Peter Stangl. New Physics in Rare B Decays after Moriond 2021. *The European Physical Journal C*, 81(10), Oct 2021. 10.1140/epjc/s10052-021-09725-1. *cited on pages 41, 135.*

[96] Patricia Ball and Roman Zwicky. New results on $B \to \pi$, K, η decay form factors from light-cone sum rules. *Phys. Rev. D*, 71(1), Jan 2005. 10.1103/physrevd.71.014015. *cited on pages 48, 113.*

[97] Patricia Ball and Roman Zwicky. $B_{s,d} \to \rho$, ω, K^*, ϕ decay form factors from light-cone sum rules revisited. *Phys. Rev. D*, 71(1), Jan 2005. 10.1103/physrevd.71.014029. *cited on pages 48, 113.*

[98] LHCb Collaboration. Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays. *Phys. Rev. Lett.*, 115:072001, Aug 2015. 10.1103/PhysRevLett.115.072001. *cited on page 48, 48.*

[99] Wouter D. Hulsbergen. Decay chain fitting with a Kalman filter. *Nucl. Instrum. Meth.*, 552(3):566575, Nov 2005. 10.1016/j.nima.2005.06.078. *cited on pages 50, 143.*

[100] Michel De Cian, Stephen Farry, Paul Seyfert, and Sascha Stahl. Fast neural-net based fake track rejection in the LHCb reconstruction. Technical report, CERN, Geneva, Mar 2017. CERN-LHCb-PUB-2017-011. *cited on page 51.*

[101] Eluned Anne Smith, Sebastian Schmitt, and Christoph Michael Langenbruch. On the PID-response dependency for dielectron pairs. Technical report, CERN, Geneva, Jun 2021. CERN-LHCb-INT-2021-008. *cited on pages 52, 79.*

[102] D Casadei. Estimating the selection efficiency. *Journal of Instrumentation*, 7(08):P08021P08021, Aug 2012. 10.1088/1748-0221/7/08/p08021. *cited on page 54.*

[103] Marie-Helene Schune, Francesco Polci, and Martino Borsato. HOP an additional tool for decays involving electrons. Technical report, CERN, Geneva, Nov 2015. CERN-LHCb-INT-2015-037. *cited on pages 67, 75.*

[104] L. Prokhorenkova et al. CatBoost: unbiased boosting with categorical features, Jan 2019. arXiv:1706.09516. *cited on page 68.*

[105] Avrim Blum, Adam Kalai, and John Langford. Beating the Hold-out: Bounds for K-Fold and Progressive Cross-Validation. In *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, page 203208, New York, NY, USA, 1999. Association for Computing Machinery. 10.1145/307400.307439. *cited on page 69.*

[106] R. Aaij et al. Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2. *EPJ Techniques and Instrumentation*, 6(1), Feb 2019. 10.1140/epjti/s40485-019-0050-z. *cited on page 78.*

[107] M. Pivk and F.R. Le Diberder. sPlot: A statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, 555(1-2):356369, Dec 2005. 10.1016/j.nima.2005.08.106. *cited on page 78.*

[108] L. Anderlini et al. The PIDCalib package. Technical report, CERN, Geneva, Jul 2016. CERN-LHCb-PUB-2016-021. *cited on page 78.*

[109] LHCb Collaboration. Measurement of the electron reconstruction efficiency at LHCb. *JINST*, 14:P11023. 21 p, Sep 2019. 10.1088/1748-0221/14/11/P11023. *cited on pages 80, 81.*

[110] Alex Rogozhnikov. Reweighting with Boosted Decision Trees. *Journal of Physics: Conference Series*, 762:012036, Oct 2016. 10.1088/1742-6596/762/1/012036. *cited on page 82.*

[111] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino. Data driven trigger efficiency determination at LHCb. Technical report, CERN, Geneva, May 2014. CERN-LHCb-PUB-2014-039. *cited on page 84.*

[112] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003. inspirehep:physics/0306116. *cited on page 98.*

[113] T. Skwarnicki. *A study of the radiative cascade transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986. DESY-F31-86-02. *cited on page 98.*

[114] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):60 – 62, 1938. 10.1214/aoms/1177732360. *cited on pages 107, 135, 136.*

[115] Aoife Bharucha, David M. Straub, and Roman Zwicky. $B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules. *JHEP*, 2016(8), Aug 2016. 10.1007/jhep08(2016)098. *cited on page 113.*

[116] N. Gubernari, A. Kokulu, and D. van Dyk. $B \to P$ and $B \to V$ form factors from B-meson light-cone sum rules beyond leading twist. *JHEP*, 2019(1), Jan 2019. 10.1007/jhep01(2019)150. *cited on page 113.*

[117] LHCb Collaboration. Framework TDR for the LHCb Upgrade: Technical Design Report. Technical report, CERN, Apr 2012. cds.cern.ch/record/1443882. *cited on page 139.*

[118] Belle II Collaboration. Belle II Technical Design Report, 2010. 10.48550/ARXIV.1011.0352. *cited on page 139.*

[119] D. Martínez Santos and F. Dupertuis. Mass distributions marginalized over per-event errors. *Nucl. Instrum. Meth.*, 764:150155, Nov 2014. 10.1016/j.nima.2014.06.081. *cited on page 157.*

Acknowledgments

First of all, i want to thank Prof. Schael for giving me the great opportunity to do my PhD at the Aachen LHCb group. Also thanks for all the guidance and fruitful conversations we had.

Special thanks to Christoph Langenbruch for being my supervisor and second examiner of my thesis. I am grateful to be part of the amazing working group you have founded and to have witnessed how each member evolved over the last years under your guidance. Thanks for all the time and effort you put in, to help us become better scientists.

A huge thanks to the whole RWTH Rare Decays working group for creating such a nice working environment. It says a lot about a working group when so many Bachelor students stay for their Masters and PhDs as well. A special thanks goes to Eluned who, as our first post-doc, taught us a great deal and was always there when things got difficult. Thanks to Hendrik and Leon who shared an office with me when i started my PhD. It was great to learn the first steps in Rare Decays together with you. Thanks to Marcel and Sophie who were, together with myself, the first permanent students in our working group and shared an awesome time with me in 'The Office'. Further, i want to thank my fellow RX people from Aachen Sebastian and Simon for the great collaboration on the project and for sharing the unique experience that was RX. Last but not least, i want to thank Jan-Marc, Johannes[1] and Thomas who also shared an office with me from time to time, helped a lot with many questions that i had and for all the fun we had both on and off work.
I truly had a great time working as a part of this group and many valuable friendships were made.

Next, a huge thanks goes out to the RX analysis group at LHCb, who helped performing the analysis presented in this thesis. First, i want to mention Simone and Renato, who collaborated with me on this project from day 1. Thanks for all the great conversations, stuff you taught me about physics and analysis as well as for all the shared success and struggles we had. Also thanks to Alex, Da Yu, Fabrice, Ryan, Sebastian and Simon who all made invaluable contributions in course of their PhD theses. Additionally, i want to thank Christoph, Eluned, Marie-Helene, Nigel, Vava and Vitali who all did a great job in supervising this analysis. I have to highlight the effort from Vava here, who in the past two years helped a lot to converge this project, by coordinating and guiding our work on a daily basis.

[1] Shout-out to the JHeuel Support, the entire working group is in your dept.

I want to thank Christoph, Eluned, Jan-Marc, Mark and Sebastian who reviewed this manuscript. All your suggestions and comments helped immensely and are very appreciated.

Another thanks goes out to Roman, Jan-Marc and Simon who met with me at CERN for a SciFi testbeam campaign during my masters thesis and had a great impact on my decision to transfer to RWTH for my PhD.

Another huge thanks goes out to my family and friends who helped me a lot and provided a good balance. Especially i want to thank my fiancé Nadine who was very supportive and patient during all this time.

Eidesstattliche Erklärung

Ich, Stephan Escher

erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;

2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;

3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;

4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;

5. Alle wesentlichen Quellen von Unterstützung wurden benannt;

6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;

7. Kein Teil dieser Arbeit wurde vor deren Einreichung veröffentlicht.

01.12.2022