

Probing the Standard Model of Particle Physics with rare B decays

Dissertation zur Erlangung des akademischen Grades Dr. rer. nat.

vorgelegt von

Tobias Tekampe

geboren am 27.10.1988 in Ahaus

Fakultät Physik
Technische Universität Dortmund

August 25, 2019

Der Fakultät Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften vorgelegte Dissertation.

Gutachter:

Dr. Johannes Albrecht

Prof. Dr. Kevin Kröninger

Vorsitzender der Prüfungskommission:

Prof. Dr. Heinz Hövel

Vertreter der wissenschaftlichen Mitarbeiter:

Dr. Bärbel Siegmann

Datum der mündlichen Prüfung:

26.8.2019

Abstract

This thesis discusses two analyses of rare B meson decays that are sensitive to effects beyond the Standard Model of particles physics. Both measurements are based on the LHCb Run 1 data sample of proton-proton collisions corresponding to an integrated luminosity of 3 fb^{-1} recorded at centre-of-mass energies of 7 and 8 TeV.

The first measurement is a search for light scalar particles in $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays. Additionally to the Standard Model contributions without intermediate resonances, a specific supersymmetric model in which the decays proceed via light scalar and pseudoscalar particles is studied. No signal candidates are observed such that upper limits on the branching fractions of the B^0 and B_s^0 decays of the order of 10^{-9} and 10^{-10} are set at 95 % confidence level. These yield an improvement of about one order of magnitude compared to the previous most stringent limits.

The second measurement is a test of lepton flavour universality with the ratio of branching fractions of the decays $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$. The analysis is still in progress. Hence, no final result can be presented. The ratio of signal yields obtained from the LHCb Run 1 data sample is determined with a statistical uncertainty of 17 %, which can be reduced to 8 % by extending the measurement to the Run 2 data sample.

Zusammenfassung

Diese Arbeit beschreibt die Analyse zweier, auf Effekte jenseits des Standardmodells der Teilchenphysik sensitiver, seltener B -Mesonzerfälle. Der analysierte LHCb Run 1 Datensatz wurde bei Proton-Proton-Kollisionen mit einer Schwerpunktsenergie von 7 und 8 TeV aufgezeichnet und entspricht einer integrierten Luminosität von 3 fb^{-1} .

Zunächst wird eine Suche nach neuen, leichten, skalaren Teilchen in den Zerfallskanälen $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ beschrieben. Neben der Suche nach standardmodellartigen Zerfällen ohne Zwischenresonanzen wird außerdem die Sensitivität auf eine spezielle supersymmetrische Erweiterung des Standardmodells, in der die Zerfälle über leichte, (pseudo)skalare Teilchen ablaufen, bestimmt. Es werden keine Signalkandidaten gefunden, sodass Ausschlussgrenzen der Ordnung 10^{-9} und 10^{-10} auf die Verzweungsverhältnisse der B^0 und B_s^0 -Zerfälle entsprechend einem Konfidenzintervall von 95 % gesetzt werden. Diese stellen eine Verbesserung der vorigen Ausschlussgrenzen um etwa eine Größenordnung dar.

Im zweiten Teil wird durch die Bestimmung des Quotienten aus den Verzweungsverhältnissen der Zerfälle $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ und $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$ Lepton-Flavour-Universalität getestet. Die Messung ist derzeit noch nicht abgeschlossen, sodass kein Endergebnis präsentiert werden kann. Das Verhältnis der Anzahlen an beobachteten Zerfällen basierend auf dem LHCb Run 1 Datensatz wird mit einer statistischen Unsicherheit von 17 % bestimmt. Durch Hinzunahme des Run 2 Datensatzes kann die Unsicherheit auf 8 % reduziert werden.



Contents

1	Introduction	1
2	Status quo in flavour physics	3
2.1	The Standard Model of particle physics	3
2.2	Search for new (pseudo)scalar particles	7
2.3	Lepton flavour universality	10
3	The LHCb experiment at the LHC	17
3.1	Tracking system	19
3.2	Particle identification	21
3.3	Reconstruction of bremsstrahlung	23
3.4	Variable definitions	25
3.5	Trigger system	27
3.6	Data processing	29
4	Common techniques for rare decay analyses	33
4.1	Branching fraction measurements	33
4.2	Hypatia function	34
4.3	Backgrounds and how to deal with them	35
4.4	Background subtraction with the $s\mathcal{P}$ lot method	37
4.5	Simulation correction	38
4.6	Particle identification calibration	40
4.7	Data driven trigger efficiencies	42
5	Search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays	45
5.1	Strategy	45
5.2	Selection	46
5.3	Physical backgrounds	55
5.4	Efficiencies	57
5.5	Systematic uncertainties and cross checks	60
5.6	Fits and limit setting	69
5.7	Results and conclusion	72

6	Test of lepton flavour universality with $R_{K\pi\pi}$	77
6.1	Strategy	77
6.2	Simulation corrections	79
6.2.1	Kinematic corrections	79
6.2.2	Particle identification calibration	87
6.3	Selection	89
6.3.1	Trigger selection	89
6.3.2	Preselection	91
6.3.3	Invariant mass intervals	94
6.3.4	Veto against physical backgrounds	95
6.3.5	Machine learning based selection	100
6.3.6	Selection optimisation	108
6.3.7	Low mass background	110
6.4	Signal yields	113
6.5	Estimation of the achievable sensitivity	117
7	Conclusion	121
A	Appendix	131
A.1	Stripping selection	131
A.1.1	Search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays	131
A.1.2	Test of lepton flavour universality with $R_{K\pi\pi}$	134
A.2	Comparisons between measured and simulated data for $R_{K\pi\pi}$	135
A.2.1	Features that are subject to simulation corrections	135
A.2.2	Features entering the BDT	137
A.2.3	Particle identification quantities	139
A.3	Background shapes for $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$	140

1 Introduction

The Standard Model of Particle Physics (SM), developed in the 1960's [1–3], is a theory of elementary particles and their dynamics. It is capable to precisely describe the outcome of collider experiments and predicted the existence of several elementary particles prior to their discovery. From an experimental point of view, the last missing constituent foreseen by the SM, the Higgs boson, has been discovered in 2012 [4,5]. However, the corresponding theoretical principles were already postulated in 1964 [6–8]. Now the question arises: What is next?

Since many years thousands of physicists are working on enhancing our understanding of elementary particles and their interactions. Good progress has been made in reducing the uncertainties to which the degrees of freedom of the SM such as particle masses or CKM parameters are measured. However, it is evident that the SM cannot be the final answer, as it is unable to describe several observations. To name some examples: We live in a world made of matter, although the asymmetry between matter and antimatter predicted by the SM is far too small to yield an explanation [9]. Cosmologists observe trajectories of stars that are incompatible with the mass of their galaxies derived from their visible matter content. This lead to the introduction of a hypothetical form of matter, which interacts via the gravitational force but does not participate in the electromagnetic interaction and is hence invisible [10]. None such dark matter particle is foreseen in the SM. Naively one would assume that the expansion of the universe is slowing down due to the gravitational attraction of all matter. Measurements however show that the universe is expanding at an accelerating rate. The energy that drives this acceleration is called dark energy as analogously to dark matter it seems to be invisible for experiments [11]. Recent measurements show that dark energy and dark matter are responsible for around 69 % and 26 % of the energy content of the universe [12]. This shows how incomplete the SM is, as it covers only 5 % of the energy density of the universe.

From the open questions we know that Beyond Standard Model (BSM) effects must exist. And based on the fact that we did not find them yet, we can assume that they occur at extremely high energy scales that are not directly accessible for today's experiments or have faint couplings.

Both assumptions motivate the study of rare decays. Firstly, if the coupling to BSM effects is small, studying rates of abundant decays is not very promising, as the BSM contributions will hide behind the large signal caused by the SM decay. Secondly, searches for specific decays of potential new particles require these to be produced on shell and are hence limited by the available energy, which for the data analysed in this thesis is at the order of 1 TeV. Contrary, studies of rare decays are indirect

searches for BSM effects. These involve tests of agreement between SM predictions and experimental observations. Rare decays are typically mediated by higher order quantum corrections where mediators exist only virtually. Therefore, masses of contributing intermediate particles are less constrained by energy conservation. That way, scales up to 100 TeV can be probed [13].

One class of thoroughly studied rare decays are $b \rightarrow s\ell^+\ell^-$ transitions, where a beauty quark converts into a strange quark while emitting two charged leptons. Interestingly, recent measurements of these type of decays point towards a consistent pattern of effects that cannot be described by the SM. In several individual analyses of $b \rightarrow s\ell^+\ell^-$ quark transitions anomalies have been reported by the Large Hadron Collider beauty (LHCb) collaboration [14–19]. These results regarded on their own represent only insignificant deviations from the SM. In order to combine the information Wilson coefficients are fit to these measurements and good agreement with the SM values is observed, such that the SM is not yet excluded. However, in comparison to models with BSM contributions the latter are favoured by more than 5 standard deviations [20–22]. Further studies of $b \rightarrow s\ell^+\ell^-$ processes are needed to clarify the picture and maybe finally find the so long sought after physics beyond the Standard Model.

This thesis begins with a brief introduction to the SM together with a motivation for the measurements that are discussed in chapters 5 and 6. In chapter 3 the LHCb detector, which recorded the data that is analysed in this thesis is introduced and the treatment of the data before the actual analyses is described. After that, chapter 4 details experimental methods that are commonly used for measurements of rare B decays performed by the LHCb collaboration. In chapter 5 a search for BSM scalar and pseudoscalar particles in non-resonant and BSM $B_{(s)}^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ decays¹ is discussed. In chapter 6 the current status of a test of lepton flavour universality based on $B^+ \rightarrow K^+\pi^+\pi^-e^+e^-$ and $B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-$ decays is presented.

¹Charge conjugation is implied throughout this thesis if not stated differently.

2 Status quo in flavour physics

In this chapter, first a brief introduction into the SM is given. Then, the search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays and after that the measurement of $R_{K\pi\pi}$ are motivated. Hereby, current anomalies along with possible explanations involving extensions to the SM are discussed.

2.1 The Standard Model of particle physics

The currently best description of elementary particles and their interactions is given by the SM. Despite omitting one known fundamental force, gravity, the SM successfully describes and predicts experimental results. An example is the discovery of the last missing elementary particle predicted by the SM, the Higgs boson. It arises as a consequence of implementing the masses of the Z^0 and W^\pm bosons in the SM and generates the masses of the elementary fermions. The Higgs boson was discovered in 2012 by the ATLAS [4] and CMS [5] collaborations, but was predicted already in 1964 [6–8]. In the following, only a brief introduction into the SM is given. If not stated differently, the information is taken from Refs. [23,24], where a more extensive discussion of the material can be found.

In the SM, matter consists of twelve fermions, subdivided into six quarks and six leptons. Both are further split into three generations containing two particles each. This structure can be visualised for quarks as

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}, \quad (2.1)$$

and for leptons as

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}_L, \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}_L, \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}_L, \quad \text{and} \quad \begin{pmatrix} e^- \\ \mu^- \\ \tau^- \end{pmatrix}_R. \quad (2.2)$$

Here, left and right handed leptons denoted by a corresponding subscript L or R are distinguished. The reason for distinguishing leptons depending on their chirality will emerge in the section discussing the weak interaction.

For each quark and each lepton a corresponding antiparticle exists, carrying opposite charge-type quantum numbers, denoted by either the different sign in the superscript for charged leptons or by an over-line for all other fermions. Quarks can be categorised in up-type, which are up (u), charm (c) and top (t) quark and down-type quarks, namely down (d), strange (s) and bottom or beauty (b) quark. The first

category carries an electric charge of $+2/3$ and the latter $-1/3$. Among the leptons the three generations are named electron (e), muon (μ) and tauon (τ). Each generation consists of a charged, massive lepton and a neutral, massless neutrino.

In the SM, interactions are described by the exchange of twelve different gauge bosons. The strong force is mediated by eight massless gluons and couples to a charge referred to as colour, which is only carried by quarks and gluons themselves. Three different colours exist, namely red, green and blue. This choice is motivated by the optical fact that light beams of these three primary colours combined result in white. Consequently, bound states of quarks, called hadrons, are colour singlets, realised by combining all three different kinds of colours. In contrast to optics, in the SM also anticolours exist, such that a colour neutral state can also be realised by pairing a colour with the corresponding anticolour. The first kind of bound states are called baryons and consist of three quarks, while the latter are mesons and consist of a quark-antiquark pair. In this thesis mostly mesons such as kaons, pions or various kinds of B mesons occur. Examples of the latter, that are studied in this thesis, are B^0 , B_s^0 , and B^+ mesons, which are made up of a \bar{b} quark bound to a d , s or u quark, respectively. The most prominent baryons are the neutron (udd) and the proton (uud). For a long time it was unclear if also larger bound states for example consisting of four quarks and an antiquark, so called pentaquarks, exist or not. Indeed these are also colour singlets and perfectly allowed in the SM and in fact resonances consistent with pentaquark states were discovered by the LHCb collaboration in 2015 [25]. This discovery lead to further extensive searches, resulting in the recent discovery of two more pentaquark states [26].

The coupling constant of the strong force increases with distance. As a consequence creating a quark-antiquark pair out of the vacuum at a certain separation of two quarks is energetically favoured. Therefore, quarks are only observed in bound states. This principle is referred to as confinement. Contrary, the strong coupling decreases asymptotically for high energies (small distances), which results in asymptotic freedom. Another consequence is that at low energies (large distances) perturbation theory is not applicable to calculate processes mediated by the strong force.

The gauge boson associated with the electromagnetic force is the photon. This interaction acts on all charged particles. In contrast to the strong and the electromagnetic force, the weak force is carried by massive gauge bosons. The weak charged current is described by W^\pm boson exchange, which on interaction with a quark alters the quark flavour from up-type to down-type or vice versa. On top of that, there is a neutral weak current, which is mediated by the Z^0 boson and conserves quark flavour. All SM particles except for photons and gluons participate in the weak interaction. The coupling of the weak gauge bosons depend on the chirality of the respective fermion. In particular, the weak force acts on left-handed particles and right-handed antiparticles. For massless particles, such as neutrinos in the SM, chirality is equal to the projection of the spin along the momentum called helicity. Keeping in mind that the only interaction neutrinos participate in is the weak force, these elementary fermions can occur only in one chirality configuration. Massive particles always carry

a fraction of each chirality, such that certain decays can be strongly suppressed due to helicity requirements, but not entirely excluded. An overview of all particles that are described by the SM and the interactions they participate in is shown in Figure 2.1.

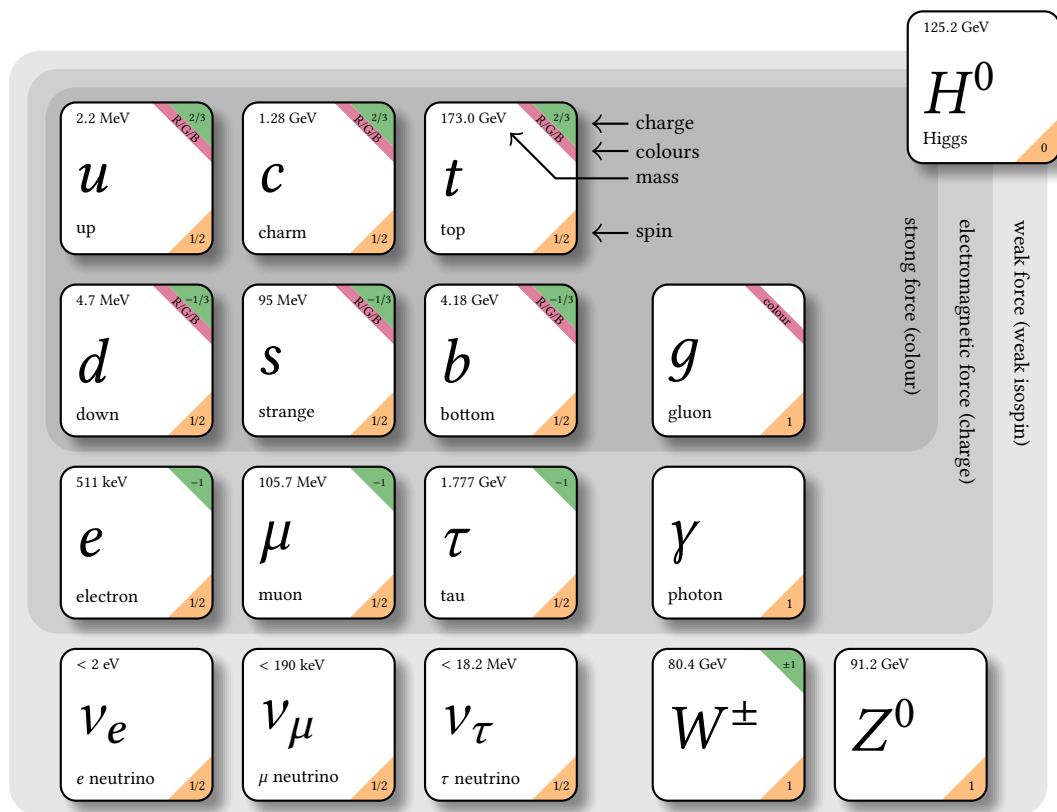


Figure 2.1: Overview of all elementary particles particles and forces described by the SM. Additionally, the charge, mass and spin of each particle is shown. Figure modified from Ref. [27] with updated numerical values from Ref. [28].

The mathematical expression of the SM can be derived from the Dirac Lagrangian by demanding invariance under local transformations corresponding to the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. Here, the group $SU(3)_C$ is responsible for generating gluon fields, $SU(2)_L$ represents the weak interaction and $U(1)_Y$ creates the photon field.

The rare decays that are studied in this thesis are Flavour Changing Neutral Current (FCNC) processes, which means that the quark flavour, but not the electrical charge is changed. In the SM, only charged weak currents mediated by the W^\pm bosons change quark flavour. Therefore, in Feynman diagrams of FCNC processes a W^\pm boson needs to interact twice with the same fermion line in order for the overall current to be neutral. This results in loop level processes as presented in Figure 2.2 making the described decays rare. In the SM, flavour changes were first explained by the GIM mechanism [29], which with introducing the third generation of quarks was extended to the CKM mechanism [30,31]. Here, eigenstates of quarks under the weak interaction are related to their mass eigenstates in terms of a complex 3×3 matrix, referred to as CKM matrix

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \quad \text{with} \quad |(V_{\text{CKM}})| = \begin{pmatrix} 0.97446 & 0.22452 & 0.00365 \\ 0.22438 & 0.97359 & 0.04214 \\ 0.00896 & 0.04133 & 0.999105 \end{pmatrix}. \quad (2.3)$$

The numerical values are taken from [28] and uncertainties were omitted in favour of readability. The probability for a transition of the kind $q \rightarrow q'$ is proportional to $|V_{qq'}|^2$. The CKM matrix is unitary, resulting in relations of the type

$$\sum_{k=1}^3 V_{ki}^* V_{kj} = \delta_{ij}, \quad (2.4)$$

with the Kronecker delta δ , which is one for $i = j$ and zero otherwise.

The dynamics of FCNC processes can be described by an effective field theory [32, 33] separating different contributions by a particular mass scale. In heavy quark effective theory a typical choice is the b quark mass. The theory is expressed in terms of an effective Hamiltonian, which for $b \rightarrow s \ell^+ \ell^-$ decays reads

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i C_i(\mu) Q_i(\mu), \quad (2.5)$$

where G_F denotes the Fermi constant, μ the energy scale, C_i are Wilson coefficients and Q_i local operators. The amplitude of a B meson decaying into a final state f is calculated as

$$A(B \rightarrow f) = \langle f | \mathcal{H}_{\text{eff}} | B \rangle = \frac{G_F}{\sqrt{2}} \sum_i V_{tb} V_{ts}^* C_i(\mu) \langle f | Q_i(\mu) | B \rangle. \quad (2.6)$$

This ansatz is called operator product expansion. It separates low energy contributions expressed as operators from high energy contributions in the form of Wilson coefficients. The latter describe effects from heavy particles such as the weak gauge bosons, the Higgs particle, the top quark, or any heavy not yet discovered particle. While Wilson coefficients are calculable from perturbation theory, non perturbative QCD effects are captured in the operators, which can be interpreted as effective vertices and the Wilson coefficients as corresponding couplings. The dominant processes for the FCNC transitions that are studied in this thesis are described by the operators Q_7 , Q_9 , Q_{10} . Corresponding Feynman diagrams are presented in Figure 2.2. In the effective Hamiltonian, the unitarity of the CKM matrix, more precisely the relation

$$V_{cb}V_{cs}^* = -V_{tb}V_{ts}^* - V_{ub}V_{us}^* \quad (2.7)$$

is exploited to eliminate contributions proportional to the factor $V_{cb}V_{cs}^*$. Furthermore, terms that are proportional to $V_{ub}V_{us}^*$ are neglected. This can safely be done, as the CKM factor is 50 times smaller than the other two and an additional suppression of the order of 10^{-4} relative to the c quark contribution arises from the loop function due to the very small mass of the u quark compared to that of the c quark [33]. For $b \rightarrow d\ell^+\ell^-$ decays the effective Hamiltonian is similar, however the contribution of the u quark cannot be neglected.

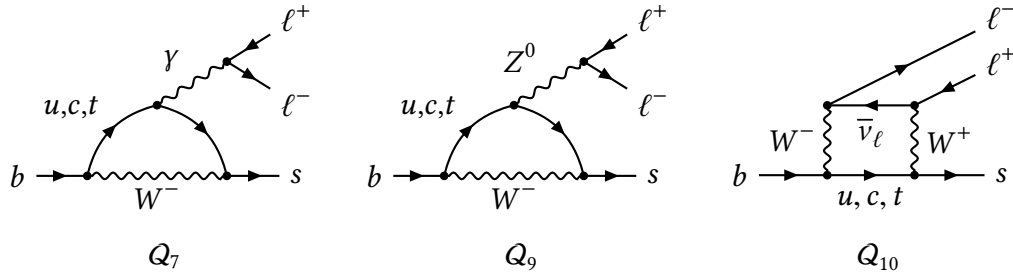


Figure 2.2: Operators and Feynman diagrams of the most dominant contributions to $b \rightarrow s\ell^+\ell^-$ decays.

2.2 Search for new (pseudo)scalar particles

A possible solution to various shortcomings of the SM is to introduce new particles such as light pseudoscalars. These new light particles must interact very weakly with SM matter to explain that they are still undiscovered. Extensions to the SM with pseudoscalar couplings to SM quarks can induce new flavour-changing processes at the one-loop level [34]. Hence, quark transitions of the kind $b \rightarrow sP$ are introduced and consequently contribute to $b \rightarrow s\ell^+\ell^-$ processes if the new pseudoscalar P decays into two charged leptons. These contributions can be searched for by measuring the branching fractions of B decays that proceed via the mentioned quark transition.

In the year 2005 the HyperCP collaboration reported an anomaly found in a search for the decay $\Sigma^+ \rightarrow p\mu^+\mu^-$ [35]. In total three decay candidates were found all of

which situated within an interval with a width of $1 \text{ MeV}/c^2$. This could be interpreted as indication for a decay of an unknown pseudoscalar particle P with a mass of $m(P) = (214.0 \pm 0.5) \text{ MeV}/c^2$ into a pair of muons. Figure 2.3 shows the measured di-muon mass distribution in comparison with two different simulation models of the SM decay and with simulated $\Sigma^+ \rightarrow pP$ decays.

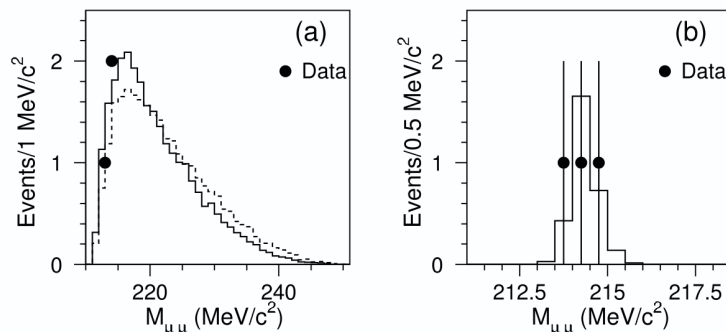


Figure 2.3: Di-muon invariant mass distribution of the three $\Sigma^+ \rightarrow p\mu^+\mu^-$ candidates observed by the HyperCP collaboration, left in comparison with two different types of SM simulation and right in comparison with $\Sigma^+ \rightarrow pP$ simulation [35].

The HyperCP anomaly has been studied in the context of a Minimal Supersymmetric Standard Model (MSSM) [36]. In supersymmetric models a new kind of symmetry transformation that translates fermions into bosons and vice versa is introduced. By requiring the SM Lagrangian to be invariant under this supersymmetry transformation new particle fields emerge, such that each fermion is assigned a bosonic partner and each boson a fermionic one. The mass difference between these two is set by the supersymmetric breaking scale. One of the new particles introduced by this model is the *goldstino*. This Goldstone fermion emerges from spontaneous supersymmetry breaking. It has a scalar superpartner S and a pseudoscalar one P both referred to as *sgoldstinos* and both neutral under all gauge interactions. The HyperCP measurement cannot be explained by an intermediate scalar *sgoldstino* S , as this would violate constraints coming from the branching fractions of rare kaon decays. Hence, the observed anomaly is interpreted to be caused by the pseudoscalar *sgoldstino* P . Applying this model in the context of $B_{(s)}^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ decays leads to a constraint of the branching fraction of $\mathcal{B}(B_s^0 \rightarrow SP) \lesssim 10^{-4}$ [37], which depending on the branching fractions of the sub-decays $S \rightarrow \mu^+\mu^-$ and $P \rightarrow \mu^+\mu^-$ might be within the sensitivity of the LHCb experiment. The predicted branching fraction of the decay in the MSSM scheme depends on the mass of the *sgoldstinos* as can be seen in Figure 2.4.

The dominant SM decay process of a B_s^0 meson into a final state composed of four muons proceeds via a J/ψ and a ϕ resonance. Multiplying the measured branching fractions of the underlying decays results in $\mathcal{B}(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi(\rightarrow \mu^+\mu^-)) = (1.83 \pm 0.18) \times 10^{-8}$ [38], which is already at the order a typical rare decay and within the reach of the LHCb experiment with the Run 1 data set. Depending on the flavour of the neutral B meson, the non-resonant

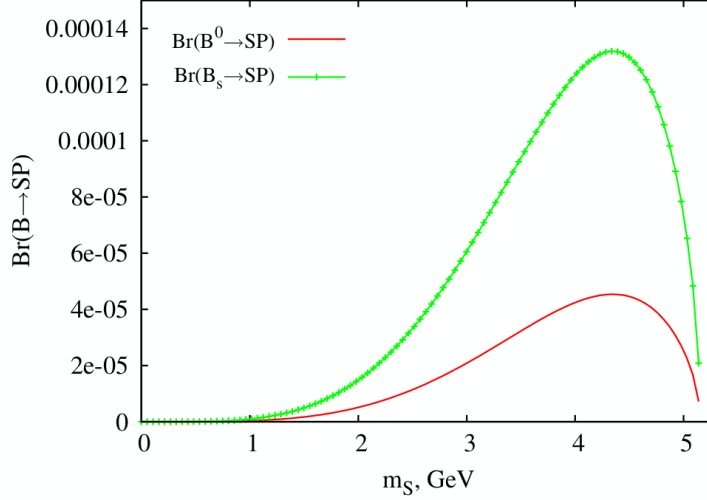


Figure 2.4: Branching fraction of the decays $B_{(s)}^0 \rightarrow SP$ in a supersymmetric extension to the SM as a function of the scalar *sgoldstino* mass. The mass of the pseudoscalar *sgoldstino* is fixed to $214 \text{ MeV}/c^2$ [37].

$B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decay proceeds through a $b \rightarrow s$ or a $b \rightarrow d$ FCNC process. In the dominant non-resonant SM amplitude, one muon pair is produced by an electroweak loop, while the other originates from a virtual photon. The branching fraction of the non-resonant $B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decay has been estimated based on the *double conversion ratio* $\rho_{\ell\ell'\ell'e'} = \frac{\Gamma(B_s^0 \rightarrow \ell\ell'\ell'e')}{\Gamma(B_s^0 \rightarrow \gamma\gamma)}$ by multiplication with the branching fraction of the $B_s^0 \rightarrow \gamma\gamma$ decay. The result is $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) \approx 3.5 \times 10^{-11}$ [39]. In a more recent publication [40], which was carried out after the measurement presented in this thesis was published estimates on both B decays are presented. The numerical values are

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) \approx (0.9 - 1.0) \times 10^{-10}, \quad \text{and} \quad (2.8)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) \approx (0.4 - 4.0) \times 10^{-12}. \quad (2.9)$$

The estimated branching fractions are below the sensitivity of the LHCb experiment making this search a null test for the SM. Therefore, observing a significant signal would be a clear sign of BSM physics. Feynman diagrams describing all mentioned processes are illustrated in Figure 2.5. The only preceding experimental study of the non-resonant decays was performed by the LHCb collaboration based on data corresponding to an integrated luminosity of 1 fb^{-1} recorded in 2011 [41]. The limits at 95% confidence level on the branching fractions were determined to be

$$\begin{aligned} \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) &< 1.6 \times 10^{-8}, \\ \mathcal{B}(B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) &< 6.6 \times 10^{-9}, \\ \mathcal{B}(B_s^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) &< 1.6 \times 10^{-8}, \\ \mathcal{B}(B^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) &< 6.3 \times 10^{-9}. \end{aligned}$$

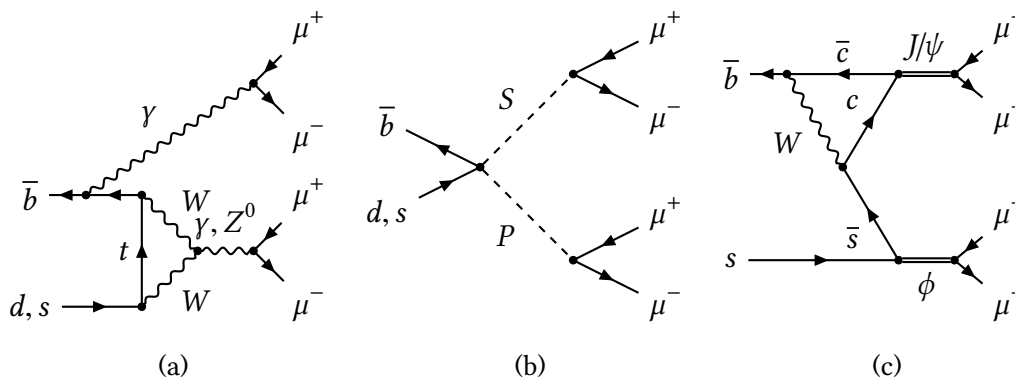


Figure 2.5: Feynman diagrams for (a) the non-resonant $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decay, (b) a supersymmetric $B_{(s)}^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)$ decay and (c) the resonant $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ decay.

The analysis discussed in chapter 5 supersedes this measurement by employing an improved analysis strategy and extending the search to the data sample recorded in 2012, which is a factor of two larger than that recorded in 2011.

2.3 Lepton flavour universality

In the SM, the gauge bosons of the weak interaction couple equally to all lepton flavours. This concept is referred to as lepton flavour universality (LFU) and can be exploited to define clean observables where large theory uncertainties cancel out. One kind of these observables that relies on $b \rightarrow s \ell^+ \ell^-$ FCNC decays is given by LFU ratios, defined as

$$R_H = \frac{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(B \rightarrow H \mu^+ \mu^-)}{dq^2} dq^2}{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(B \rightarrow H e^+ e^-)}{dq^2} dq^2}. \quad (2.10)$$

Here, q^2 is the square of the invariant mass of the di-lepton system and Γ the decay width. Due to the equal hadronic content H in the numerator and denominator uncertainties originating from hadronic form factors cancel in the ratio. Therefore, extremely precise SM predictions of these observables are possible. At a significant distance from the di-lepton production thresholds these ratios are predicted to be

$$R_H = 1 + \mathcal{O}\left(\frac{m_\mu^2}{m_b^2}\right), \quad (2.11)$$

in the SM with an uncertainty of around 1% [42, 43].

Experimentally, this ratio has been evaluated for two kinds of $b \rightarrow s \ell^+ \ell^-$ decays so far. More precisely, $B^+ \rightarrow K^+ \ell^+ \ell^-$ and $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays have been studied to calculate R_K and $R_{K^{*0}}$. The first mentioned has been determined by the LHCb

collaboration to be

$$R_K = 0.846_{-0.054}^{+0.060} (\text{stat})_{-0.014}^{+0.016} (\text{syst}) \quad (2.12)$$

in the range $1 < q^2 < 6 \text{ GeV}^2/c^4$ [16]. A comparison to the theory prediction and to corresponding measurements performed by other experiments is presented in Figure 2.6. Good compatibility between the LHCb measurement and the values reported by the Belle [44] and the BaBar [45] collaborations can be seen. Due to the significantly smaller uncertainty achieved by the LHCb collaboration with respect to the other experiments, a tension with the SM arises. The deviation is quantified as 2.5 standard deviations, which is clearly below the threshold for a *first evidence* of three or even an *observation* of five standard deviations.

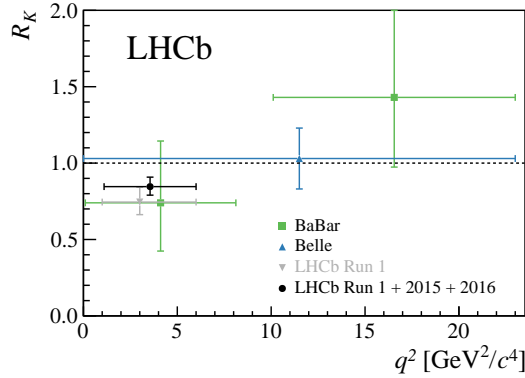


Figure 2.6: Values of R_K measured previously [15] and recently [16] by the LHCb, the BaBar [45] and the Belle [44] collaboration. Additionally, the SM expectation [46, 47] is highlighted. The figure is taken from Ref. [16].

The second $b \rightarrow s \ell^+ \ell^-$ process that has been studied in the context of LFU is given by $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays and the corresponding ratio $R_{K^{*0}}$. The LHCb collaboration reported a measurement of this observable in two q^2 regions with the results [17]

$$R_{K^{*0}} = 0.66_{-0.07}^{+0.11} (\text{stat}) \pm 0.03 (\text{syst}) \quad \text{for } 0.045 < q^2 < 1.1 \text{ GeV}^2/c^4, \text{ and} \quad (2.13)$$

$$R_{K^{*0}} = 0.69_{-0.07}^{+0.11} (\text{stat}) \pm 0.05 (\text{syst}) \quad \text{for } 1.1 < q^2 < 6.0 \text{ GeV}^2/c^4. \quad (2.14)$$

In Figure 2.7 comparisons to corresponding results from BaBar [45] and Belle [48] and to various SM predictions made by different theory groups are visualised. Again, the LHCb measurement achieves the smallest uncertainty and is perfectly compatible with the results reported by the B factories. Regarding the SM predictions a similar pattern as for R_K emerges. In both q^2 regions smaller values than predicted are observed. The deviation between the LHCb measurement and the SM predictions are reported to be 2.1–2.3 and 2.4–2.5 standard deviations for the two bins, respectively.

These LFU measurements point into a consistent direction of BSM effects in $b \rightarrow s \mu^+ \mu^-$ and/or $b \rightarrow s e^+ e^-$ processes. In fact, adding more experimental insights to the picture the trend goes towards muon decay rates being smaller than predicted.

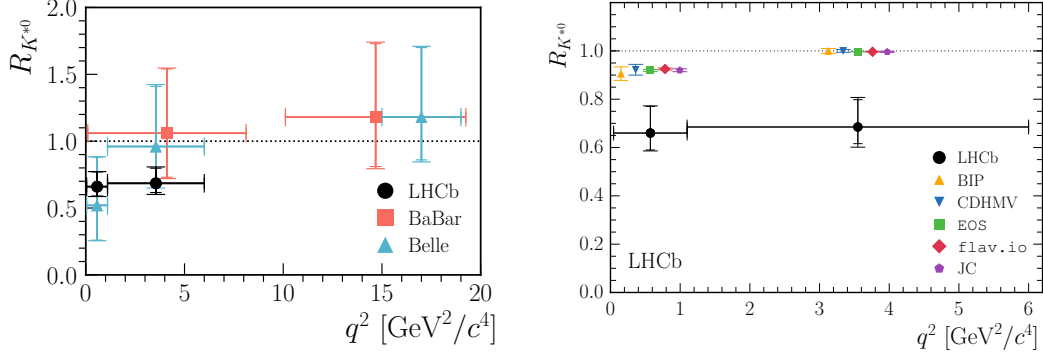


Figure 2.7: Values of $R_{K^{*0}}$ measured by the LHCb collaboration in two bins of di-lepton invariant mass, q^2 [17]. The left figure shows a comparison to the corresponding measurements performed by the BaBar [45] and Belle [48] collaborations. In the right figure the LHCb result is compared to SM predictions calculated by different groups of theorists. Both figures are taken from Ref. [17]. The left plot is reproduced to include the most recent Belle measurement.

Differential branching fractions of $b \rightarrow s\mu^+\mu^-$ processes are systematically measured smaller than expected by the SM. This is illustrated in Figure 2.8, which shows comparisons between theoretical and experimental values for $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B^+ \rightarrow K^+\mu^+\mu^-$ decays. In both a deficiency in the observed differential branching fraction in the q^2 region below the J/ψ resonance can be seen. Interestingly, this is the region in which the previously discussed LFU ratios were determined.

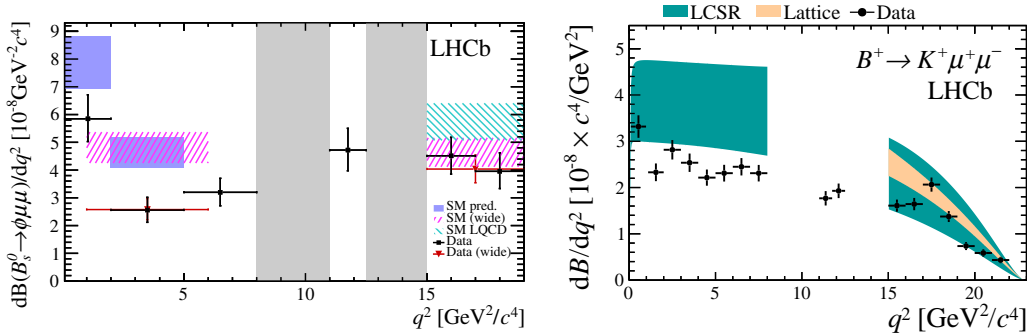


Figure 2.8: Differential branching fractions of two $b \rightarrow s\mu^+\mu^-$ transitions showing consistent downwards deviations in the q^2 region below the J/ψ resonance. The left figure is taken from Ref. [18] and shows values measured by the LHCb collaboration for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays along with the corresponding SM predictions [49–51]. The right figure is taken from Ref. [19] and draws the same comparison for $B^+ \rightarrow K^+\mu^+\mu^-$ decays. Again the measured values were determined by the LHCb collaboration and the SM predictions originate from Refs. [52, 53].

None of the mentioned individual measurements reports a significant deviation. However, the number of different measurements of $b \rightarrow s\ell^+\ell^-$ processes, which are all parametrised by the same effective Hamiltonian given in equation (2.5), can be used to over constrain the system and determine the Wilson coefficients by letting them float in a fit. This has been done by several different groups, e.g. [20–22], using different baselines for example in terms of matrix elements, but converging on a similar conclusions: a negative shift in C_9 , as can be seen in Figure 2.9. Without inclusion of the LFU ratios an explanation due to hadronic effects is possible. However, these largely cancel in those ratios, such that a BSM interpretation is favoured. To explain the unequal effects for di-muon and di-electron decays, the BSM effect must be strongly suppressed or absent in the latter. This is achieved by distinguishing between C_9^e and C_9^μ . Among the favoured models that are capable of accounting for the discussed tensions are leptoquarks [54], which directly couple leptons to quarks. Also models with new heavy gauge bosons, such as a Z' [55], which can mediate a FCNC at tree level are possible.

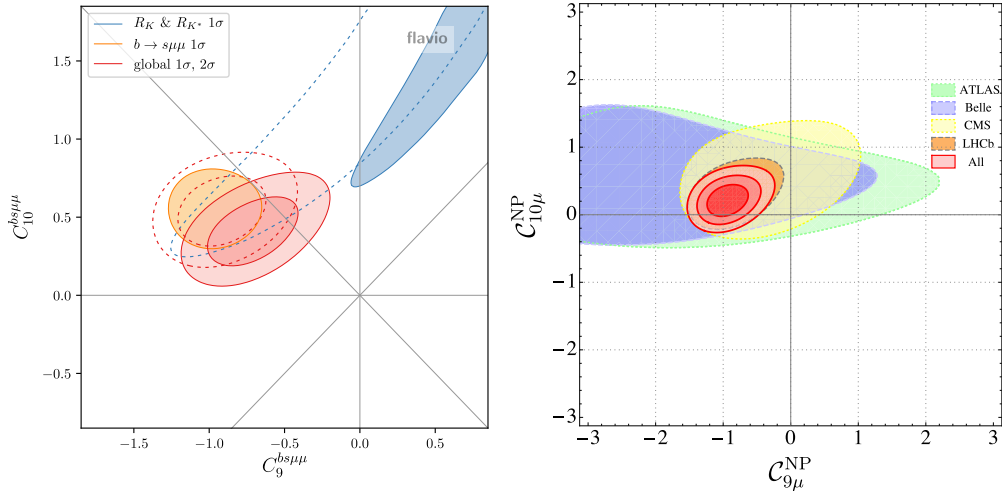


Figure 2.9: Fit results of the shifts in C_9 and C_{10} obtained by different global analyses of $b \rightarrow s\ell^+\ell^-$ processes using different matrix elements, but finding a consistent tension with the SM [20, 21].

In the SM, all FCNC processes are mediated by the weak interaction, which couples only to left handed particles and right handed antiparticles. Therefore, the right handed counterparts to the dominant left handed processes, often denoted by a prime ($C'_i Q'_i$) were omitted so far. However, BSM interactions that potentially violate LFU might give rise to these [56]. Regarding only linear BSM contributions, which given the current experimental precision is sufficient, LFU ratios can be expressed as

$$R_K \approx 1 + \Delta_+, \quad (2.15)$$

$$R_{K^{*0}} \approx 1 + p(\Delta_- - \Delta_+) + \Delta_+, \quad (2.16)$$

$$R_{K_1} \approx 1 + p'(\Delta_+ - \Delta_-) + \Delta_-, \quad (2.17)$$

with

$$\Delta_{\pm} = 2\text{Re} \left[\frac{C_{LL}^{\text{NP},\mu} \pm C_{RL}^{\text{NP},\mu}}{C_{LL}^{\text{SM}}} - \frac{C_{LL}^{\text{NP},e} \pm C_{RL}^{\text{NP},e}}{C_{LL}^{\text{SM}}} \right] \quad (2.18)$$

Here, p and p' are polarisation fractions of the respective hadron as defined in [56]. In both cases discussed here, the polarisation fractions are close to 1, making the corresponding LFU ratios highly sensitive to right handed currents. In equation (2.18), the Wilson coefficients C_{XY} are defined in a chirality projected basis. The corresponding operators are

$$Q_{XY}^{\ell} = \bar{s}\gamma_{\mu}P_X b\bar{\ell}\gamma^{\mu}P_Y\ell \quad \text{with} \quad X, Y \in \{L,R\}, \quad (2.19)$$

where γ_{μ} are Dirac matrices and $P_{L,R}$ the chiral projectors. The Wilson coefficients relate to the standard basis as

$$C_{LL}^{\ell} = C_9^{\ell} - C_{10}^{\ell}, \quad C_{LR}^{\ell} = C_9^{\ell} + C_{10}^{\ell}, \quad (2.20)$$

$$C_{RL}^{\ell} = C_9^{\prime,\ell} - C_{10}^{\prime,\ell}, \quad C_{RR}^{\ell} = C_9^{\prime,\ell} + C_{10}^{\prime,\ell}. \quad (2.21)$$

Equation (2.18) exploits that in the SM $C_{LL}^{\mu} = C_{LL}^e$ and $C_{LR}^{\ell} = C_{RL}^{\ell} = C_{RR}^{\ell} = 0$. In equation (2.17) only the K_1 family is included, while the corresponding measurement presented in chapter 6 targets a determination of $R_{K\pi\pi}$ integrated over all processes that mediate non-resonant $B^+ \rightarrow K^+ \pi^+ \pi^- \ell^+ \ell^-$ decays. However, the overall rate is clearly dominated by processes proceeding via the K_1 family as shown by the Belle collaboration [57]. The authors of Ref. [56] suggest to not only measure LFU ratios, but also double ratios such as $R_{K^{*0}}/R_K$. Here, left handed BSM contributions cancel, which allows for the determination of the handedness of a potential BSM current.

The precisions of global fits depend on the precision and number of the available individual measurements of $b \rightarrow s\ell^+\ell^-$ decays and also on the theoretical uncertainties of their predictions. Therefore, it is beneficial to add more individual measurements of LFU ratios. One such ratio is $R_{K\pi\pi}$, where not a distinct resonance but the entire $K^+ \pi^+ \pi^-$ spectrum, dominated by the $K_1^+(1270)$ and $K_1^+(1400)$ is included. Analogously to R_K and $R_{K^{*0}}$ the ratio,

$$R_{K\pi\pi} = \frac{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-)}{dq^2} dq^2}{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-)}{dq^2} dq^2} \quad (2.22)$$

is defined where the di-lepton invariant mass range considered here ranges from $q^2 = 1.1$ to $7 \text{ GeV}^2/c^4$. While the $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$ decay has not yet been observed the $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ decay has been discovered by the LHCb collaboration [58]. Its branching fraction was measured to be

$$\mathcal{B}(B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-) = (4.36_{-0.27}^{+0.29} \text{ (stat)} \pm 0.21 \text{ (syst)} \pm 0.18 \text{ (norm)}) \times 10^{-7}, \quad (2.23)$$

where the individual uncertainties are statistical, systematic and due to the uncertainty of the branching fraction of the normalisation mode. The branching ratio of the $B^+ \rightarrow K_1^+(1270)\mu^+\mu^-$ decay predicted by the SM is reported to be $(2.3_{-1.0}^{+1.3} \text{ }_{-0.3}^{+0.0}) \times 10^{-6}$ [59]. Accounting for the branching fraction of the underlying decay, $\mathcal{B}(K_1^+(1270) \rightarrow K^+\pi^+\pi^-) = (35.7 \pm 3.7)\%$ [38], the measured and predicted values agree within uncertainties. However, this neglects contributions from amplitudes that do not proceed via an intermediate $K_1^+(1270)$ resonance.

The Belle collaboration performed an amplitude analysis of the decays $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$ and $B^+ \rightarrow \psi(2S)K^+\pi^+\pi^-$ [57]. The $m(K^+\pi^-)$, $m(\pi^+\pi^-)$ and $m(K^+\pi^+\pi^-)$ spectra were studied and parametrised in a three-dimensional fit. In both decay channels the $K_1^+(1270)$ resonance is reported to yield the dominant contribution to the decay followed by the $K_1^+(1400)$ meson. In this thesis an integrated measurement over the full spectrum is performed. Especially the signal yield of the di-electron mode is expected to be too small to draw significant conclusions concerning contributions from different resonances. However, regarding the previously mentioned test for right handed currents, it could be beneficial to account for different hadronic contributions in the future with more data available. The current state of the measurement of $R_{K\pi\pi}$ is presented in chapter 6.

3 The LHCb experiment at the LHC

In this thesis decays of B mesons are studied, which are produced in proton-proton (pp) collisions provided by the Large Hadron Collider (LHC) [60] during the years 2011 and 2012, referred to as Run 1. Therefore, the description of the LHCb experiment in this chapter represents the state and performance during this period. The LHC is the world's largest and most powerful synchrotron operated by the European Organization for Nuclear Research (CERN) at the French-Swiss border close to Geneva. The centre-of-mass energy (\sqrt{s}) of the pp collisions delivered by the LHC was increased stepwise. During Run 1 collisions at $\sqrt{s} = 7$ and 8 TeV were provided. After the Long Shutdown 1 (LS1), that was mainly used to consolidate splices between the LHC cryomagnets and ended in late 2015, the energy was increased to 13 TeV [61]. The period after LS1 ending in December 2018 is referred to as LHC Run 2. The proton beams of the LHC consist of up to 2808 bunches, each containing around 10^{11} protons.

The data analysed for the measurements presented in this thesis are recorded with the LHCb detector [62], which is one of the four major experiments at the LHC. The remaining three main experiments are ALICE [63], ATLAS [64] and CMS [65]. The first intends to study the strong interaction at extremely high energy densities at which quarks can be asymptotically free and form a quark-gluon plasma. For these studies the LHC specifically delivers lead-lead or proton-lead collisions in dedicated runs. The latter two experiments are general purpose detectors that are constructed to measure decays of heavy particles and therefore are built around the pp -interaction point, such that tracks propagating perpendicular to the beam axis can be reconstructed. In contrast to the LHCb experiment ATLAS and CMS favour large instantaneous luminosities (up to 40 pp -interactions per bunch crossing) in order to collect as much data as possible. The LHCb experiment is build to operate at an average instantaneous luminosity of $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ [62], which is far lower than the design luminosity of the LHC of $1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. However, during LHC Run 1 the detector was able to take data at twice its design luminosity due to the outstanding performance of the hard- and software. The remaining difference of more than one order of magnitude between LHC and LHCb design luminosities is compensated by a lower focusing of the beams. Additionally, the transversal beam overlap is adjusted in real-time such that the instantaneous luminosity delivered to the LHCb detector remains constant. An advantage of running at lower luminosities is that events dominantly consist of single pp -interactions, resulting in a lower occupancy in the detector, which allows for more precise measurements and causes lower radiation damage. Despite the reduced luminosity, the data gathered by the LHCb collaboration is the largest $b\bar{b}$ -data sample in the world making it the ideal basis for the study of rare

B decays. As an illustration the increase in integrated luminosity and correspondingly the number of produced b hadrons is shown in Figure 3.1.

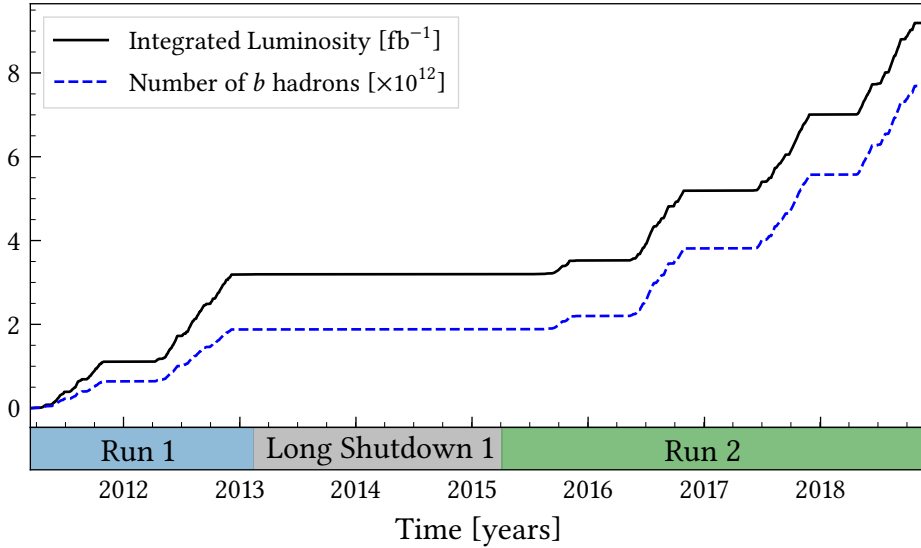


Figure 3.1: Amount of integrated luminosity and corresponding number of produced b hadrons collected by the LHCb collaboration as a function of time.

The LHCb detector is designed for precision measurements of the beauty and charm system. In particular rare decays and CP violating processes are studied. In LHC proton-proton collisions, bottom quarks are produced dominantly by the fusion of two gluons and a subsequent decay into a $b\bar{b}$ pair. Since the gluons carry a statistically distributed portion of the proton impulse, high asymmetries between the gluon impulses are likely. This results in a high momentum of the $b\bar{b}$ system and hence in a large boost along the beam axis. Exploiting this boost the LHCb detector is realised as a single arm forward spectrometer covering a pseudo rapidity range of $2 < \eta < 5$.

As can be seen in Figure 3.2 the detector consists of multiple subsystems of which all parts that are crucial for the presented measurements are detailed in the following. Hereby the geometry is described in a Cartesian coordinate system, where the x -axis is horizontal, the y -axis vertical and the z -axis along the beam direction. The coordinate system is centred at the proton-proton interaction point called primary vertex (PV).

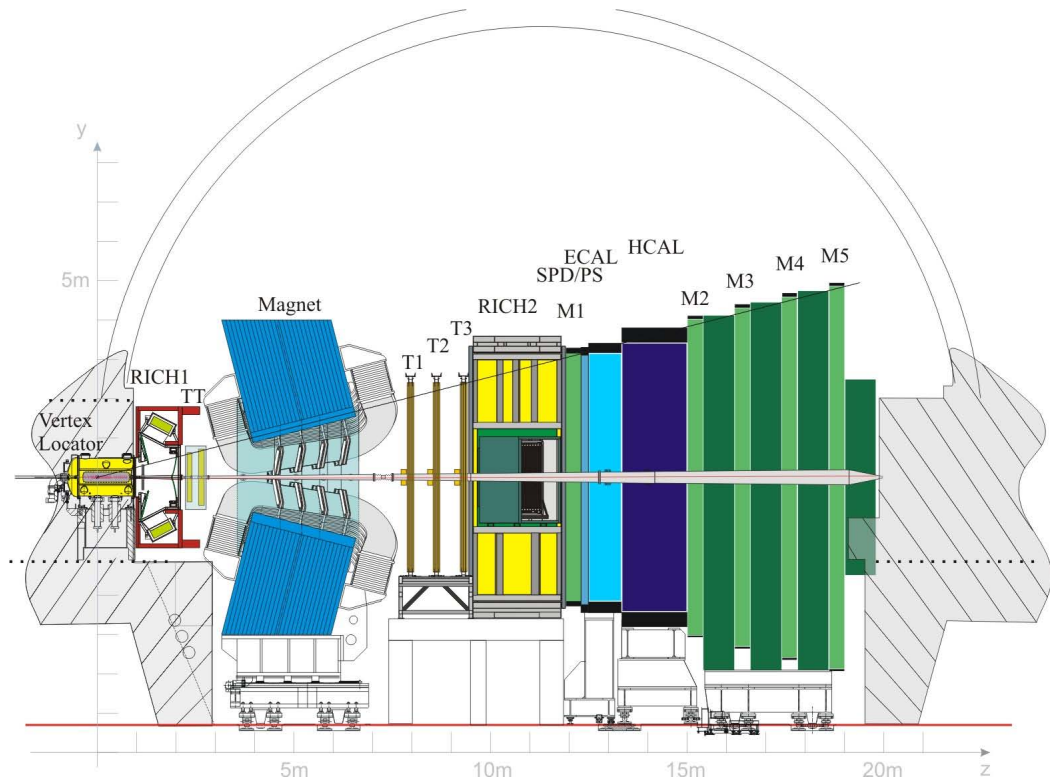


Figure 3.2: Schematic view of the LHCb detector [62]. The interaction point is at the centre of the coordinate system. A hypothetical unstoppable particle produced in a proton-proton collision inside the vertex locator (VELO) propagating into the detector acceptance traverses in order, a ring-imaging Cherenkov detector (RICH), the tracker Turicensis (TT), the dipole magnet, tracking stations (T1–T3), a second RICH, the muon station (M1), the calorimeter system consisting of a scintillator pad detector (SPD), a preshower detector (PS), an electromagnetic calorimeter (ECAL) and a hadron calorimeter (HCAL), and finally four more muon stations (M2–M5).

3.1 Tracking system

In order to reconstruct the trajectories of charged particles information from multiple detector systems is combined. These systems are based on two different technologies: For high occupancy regions semiconductor detectors in form of silicon microstrip sensors are employed, while in the outer regions of the experiment drift-tubes are sufficient. Around the pp -collision point the VELO [66] is situated. For optimal precision of the measurements of the PV and the displaced secondary vertex (SV), for this work most importantly the B decay vertices, the sensors must be positioned as close to the beam as possible. Therefore, the VELO is constructed in two halves, each installed on a movable support-structure, such that the delicate measuring device can be moved out of the LHC beam aperture during injection. While the proton beams are proclaimed stable by the LHC control centre, the VELO can be moved as close as 7 mm from the beam. Each half consists of 21 modules of which there are two types. One

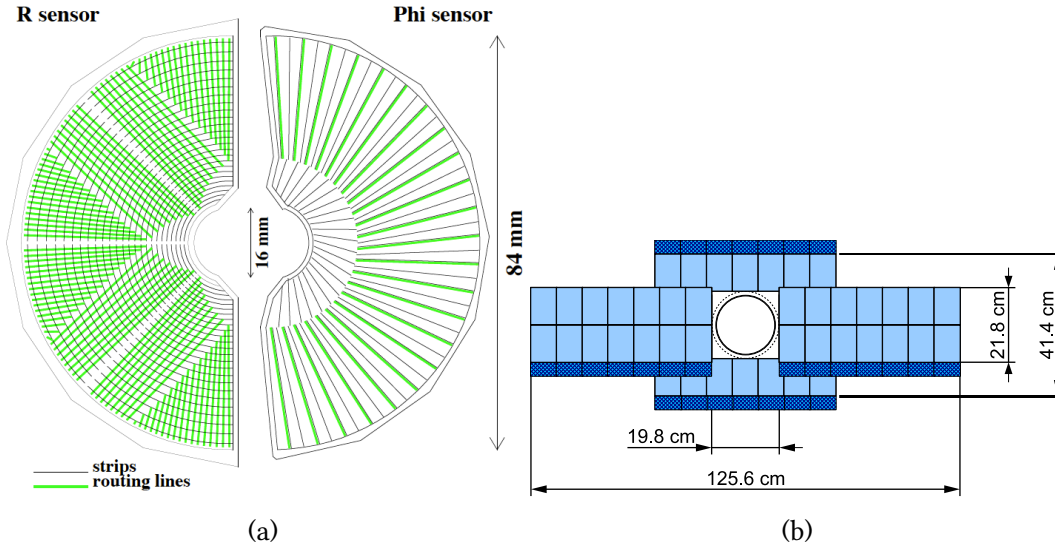


Figure 3.3: Figure (a) shows a sketch of R and Phi sensors of the VELO [66] and figure (b) a sketch of the IT layout [62]. In both sketches the circular hole in the centre accounts for the beam-pipe.

for measuring the distance of a hit from the z -axis (R sensor) and one for measuring the azimuthal angle (Phi sensor). As can be seen in Figure 3.3a the difference is in the orientation of the silicon strips. For R sensors they are arranged tangentially, while for Phi sensor the orientation is radial. The modules in the two halves are arranged such that in closed state there is a small overlap. That way all tracks inside the LHCb acceptance pass through at least three modules. The PV resolution depends strongly on the number of measured tracks originating from the PV. For 25 tracks the resolution in x and y coordinates amounts to $13\ \mu\text{m}$ while in z direction a resolution of $71\ \mu\text{m}$ is achieved. An important criterion to distinguish between particles produced in the PV and those originating from a SV is the impact parameter (IP). The IP resolution achieved by the VELO is $12\ \mu\text{m}$ in the xy -plane for tracks with high momentum and drops to $35\ \mu\text{m}$ for particles with transverse momentum of $1\ \text{GeV}/c^2$.

Downstream of the VELO, the next part of the tracking system is the TT, situated between the first RICH and the magnet. The TT covers the full acceptance of the LHCb detector. Its main purpose is to detect low energy particles whose trajectories are bent out of the detector by the magnet thus not reaching tracking stations T1-T3. Additionally, the reconstruction of particles that decay outside of the VELO relies on the TT. Together with the IT, which is situated in the inner and therefore highly occupied regions of the tracking stations T1-T3, the TT forms the LHCb silicon tracker project. Both are organised in four layers arranged in a x - u - v - x pattern, where x layers are vertically orientation and u and v layers are rotated by $\pm 5^\circ$ around the z -axis to enable a hit resolution also in y -direction. As indicated in Figure 3.3b, the IT instruments cross shaped areas of 120 by 40 cm around the beam pipe in the tracking stations T1-T3. The outer tracker (OT), a drift-tube detector, covers the region outside of the IT out to 250 cm in the vertical and 300 cm in the horizontal

plane. Also this device consists of four layers organised in a $x-u-v-x$ pattern in every station. The OT has a fast read-out time of less than 50 ns and a drift coordinate resolution of 210 μm [67]. Another important component of the tracking system is given by the dipole magnet. With its integrated field strength of 4 T m for tracks with a length of 10 m it bends the trajectories of charged particles and thereby allows for momentum measurement based on the curvature of reconstructed tracks. The momentum resolution of the tracking system depends on the track momentum and ranges from $\delta p/p = 0.5\%$ for particles with below 20 GeV/ c to 0.8% for particles with 100 GeV/ c [68].

3.2 Particle identification

This section details the detector components that are dominantly involved in calculating particle identification (PID) quantities. The final states that are studied in this thesis are composed of four different particles, namely kaons, pions, muons and electrons, which need to be distinguished from one another and from further particle species such as protons. Charged hadrons, like pions and kaons, are mainly identified by the RICH system. Electrons and the photons they emit as bremsstrahlung are measured by the ECAL. For muons, the muon stations play an important role. Also signatures measured in the HCAL contribute to particle identification. In the following, these three detector components are detailed.

The LHCb experiment is equipped with two RICH detectors [69]. RICH1 is situated after the VELO and RICH2 between the last tracking station T3 and the first muon station M1. They differ in terms of gas filling. While RICH1 is filled with C_4F_{10} ($n = 1.0014$), RICH2 contains CF_4 ($n = 1.0005$). The refractive indices n correspond to a temperature of 0° and pressure of 1013 mbar. The reason for using different gases is that RICH1 is designed to measure particles with a momentum between 1 and 60 GeV/ c , while RICH2 targets those with 15 to 100 GeV/ c . The detection principle exploits the Cherenkov effect, which occurs when a charged particle propagates in a medium at a speed greater than the phase velocity of light in the corresponding substance. The Cherenkov effect manifests in the emission of photons in a forward cone with an opening angle θ_c given by

$$\cos \theta_c = \frac{c}{nv}, \quad (3.1)$$

where c is the speed of light in vacuum, n the refractive index of the medium and v the velocity of the particle. As indicated in Figure 3.4a the emitted light is guided by mirrors onto hybrid photon detectors (sHPDs), where it is detected in a circular pattern. Based on the radius of the circle the opening angle of the cone can be calculated. Together with the information from the tracking system, the mass of the underlying particle can be determined. The figure shows Aerogel as an additional radiator. This material was found to absorb the surrounding C_4F_{10} and thereby compromise the

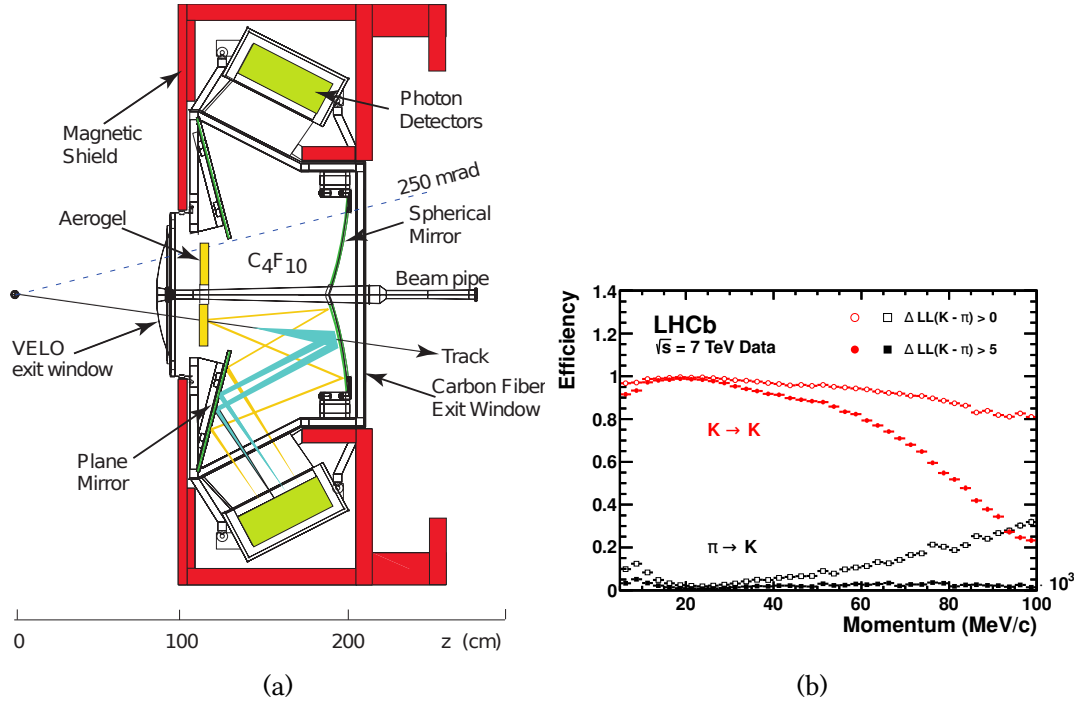


Figure 3.4: Figure (a) shows a side view of RICH1 [62] and figure (b) the efficiencies of pion identification and kaon misidentification for different $DLL_{K\pi}$ requirements as a function of the momentum [69].

Cherenkov angle resolution. Therefore, the Aerogel has been removed at the end of the data taking period in 2011.

The calorimeter system [70] is made up by the SPD, the PS, the ECAL and the HCAL and is situated between the first and the second muon station. The purpose of the SPD is to differentiate between charged and neutral particles, and thereby avoid that neutral pions and photons cause positive decisions in the electron hardware trigger. Downstream of the SPD a 15 mm thick layer of lead, corresponding to 2.5 radiation lengths followed by the PS is located. This component serves to distinguish between electrons and charged hadrons, which can be achieved due to the fact that electrons in contrast to e.g. pions are likely to produce a shower due to interactions with the lead. Following the PS the ECAL and after that the HCAL are installed. Both consists of alternating layers of shower material and scintillator pads. For the ECAL the shower material is lead, while for the HCAL iron is employed. The first mentioned corresponds to 25 radiation lengths, the second to 5.6 hadronic interaction lengths. The response of the calorimeter system is not only used for particle identification, but is also a crucial input to the first trigger stage, which is further discussed in 3.5.

The muon system consists of five stations. The first situated before the calorimeter system for better transverse momentum resolution and the remaining four at the very end of the detector. Each station except for M1 is followed by a muon shield consisting of a 80 cm thick iron layer in order to stop remaining hadrons. The active part of the detector consists of multi-wire proportional chambers (sMWPCs) except for the inner

part of M1, which is instrumented with Triplegas electron multipliers (sGEMs). The individual muon stations can be divided into four different regions concerning the granularity of the readout cells, which is higher in the centre region and decreases with the distance from the beam axis. That way the particle flux is approximately the same in the four regions of a given station. As a consequence, the spacial resolution in x and y is lower at large distances from the beam pipe. However, in these regions the resolution is anyway limited due to multiple scattering.

In order to separate muons from any other type of particles the boolean quantity $IsMu\text{on}$ is introduced [71]. It is assigned the value `True` if hits in the muon stations inside a field of interest around the extrapolation of a reconstructed track are detected. Depending on the momentum of the track hits in different combinations of individual Muon stations are required. The efficiency of the $IsMu\text{on}$ flag is measured on a $J/\psi \rightarrow \mu^+ \mu^-$ control sample using a tag and probe method. The average efficiency is around 98 %, while the misidentification efficiencies for protons, pions and kaons depending on their transverse momentum lie between 0.1 and 1.3 %, 0.2 and 5.6 % and 0.6 and 4.5 %, respectively.

A better purity of muons and additionally the separation between electrons, pions, kaons and protons is achieved based on two different types of quantities [68]. One is given by the difference in log-likelihood (DLL) of one of the mentioned particle types x with regard to the pion mass, $DLL_{x\pi}$. Here, the individual $DLL_{x\pi}$ of the detected signals in the RICH, the calorimeter and the muon system are added linearly. The performance for the kaon-pion separation as a function of the momentum can be seen in Figure 3.4b. In contrast to the boolean $IsMu\text{on}$ flag the $DLL_{x\pi}$ are continuous quantities such that the efficiencies and misidentification rates depend on the required threshold. For $DLL_{K\pi}$ for example requiring the value to be greater than zero at average keeps 95 % of true kaons while rejecting 90 % of all pions. The second approach also accounts for the correlations between the different detector systems by employing an artificial neural network (ANN). The predicted probability for a track to be of the species x made by the ANN is referred to as $ProbNNx$.

3.3 Reconstruction of bremsstrahlung

Bremsstrahlung is emitted by charged particles when their direction of flight is altered or their velocity is reduced due to electromagnetic interaction. This happens for example in material interactions or in magnetic fields. The resulting photons are emitted in direction of flight and the energy a particle loses due to bremsstrahlung is inversely proportional to its mass to the fourth power. Therefore, the effect is significantly more distinct for electrons than it is for muons. In fact at the LHCb experiment energy losses due the emission of bremsstrahlung are negligible for muons.

Photons are detected by the ECAL, whereby a transverse momentum of larger than 75 MeV/ c is required. For technical reasons, such as limited disk space, photons in simulated data are only stored if their momentum is larger than 100 MeV/ c . The

performance of the photon reconstruction is studied based on a simulation of electrons traversing the LHCb detector [72]. Here, three cases that differ in the region where the electron emits the photon are distinguished. Firstly, the photon is emitted after the particle traversed the magnet region, e.g. the z -coordinate of the origin of the photon is at $z > 7$ m. These photons are mostly detected in the same calorimeter cell as the corresponding electron making the two undistinguishable. However, this does not compromise the measurement of the electron momentum. The reason for this is that the momentum is measured from the curvature of the track, which is not altered by photon emission after the magnet region as here the track is not bent. The second case consists of photons that are emitted within the magnet region, e.g. $3 < z < 7$ m. This happens so rarely that it is negligible. The most important scenario is that the electron emits bremsstrahlung before being bent by the magnet, e.g. at $z < 3$ m, because here energy is lost before the momentum is measured. About 50 % of the photons in simulated data (e.g. that pass the momentum threshold) satisfy the transverse momentum requirement, such that they are reconstructable. Of these another 50 % is reconstructed in the ECAL if emitted before the magnet region. For successful reconstruction of a photon, it must be within the acceptance of the ECAL and must not share a calorimeter cell with another particle.

The procedure to add the energy of reconstructed photons to the corresponding electron track is illustrated in Figure 3.5. The sketch shows a top view of the involved detector components, which are the VELO, the TT and the ECAL. A search window in the latter, indicated by the hatched region, is defined by extrapolating the track of the considered electron linearly once from its origin vertex, which is the B vertex for decays studied in this thesis, and once from the intersection point with the TT. A line in the xy -plane is constructed by connecting the points where the two extrapolations intersect the ECAL. All photon clusters within a distance of 2σ from this line are considered as bremsstrahlung that was emitted by the electron track. Here, σ is defined as the combined uncertainty of the track extrapolation and the location of the cluster barycentre. As the ECAL only measures deposited energy, the direction of the momentum is deduced by assuming that the photon originates from the PV. The resulting four-momentum is then added to the electron.

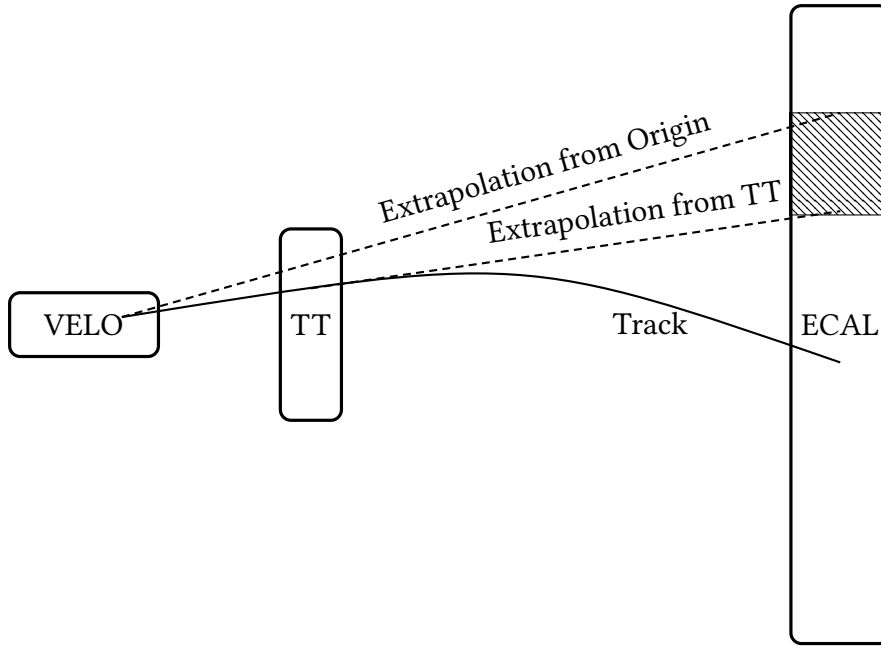


Figure 3.5: Illustration of the bremsstrahlung recovery at LHCb reproduced from Ref. [72].

3.4 Variable definitions

In this section variables that are used to distinguish signal from background candidates throughout this thesis are explained and defined. A very frequently mentioned quantity is the transverse momentum, p_T , which is defined as the component of a track's momentum perpendicular to the z -axis. Another frequently used quantity related to the trajectories of reconstructed particles is the pseudo rapidity, which is defined as

$$\eta = -\ln [\tan (\theta / 2)] \quad (3.2)$$

with θ being the angle between the particle momentum and the beam axis.

Figure 3.6 shows a sketch of a $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ signal event. The two proton beams approach from the left and right, colliding at the PV. Additional to the $B_{(s)}^0$ meson from the example many more particles are produced in the collisions, which were omitted in the sketch for the sake of readability. Based on the tracks of these particles that are reconstructed in the VELO the location of the PV is measured. The precise knowledge of the PV location is the foundation to calculate many properties of the B meson such as the lifetime, the flight distance or the impact parameter significance (χ_{IP}^2) with respect to the PV. The quality of reconstructed tracks and vertices is measured in terms of χ^2 (χ_{trk}^2 and χ_{vtx}^2 respectively) that are defined following the eponymous statistical hypothesis test. Often these are divided by the number of degrees of freedom (ndf). The χ_{IP}^2 is defined as the difference between the χ_{vtx}^2 of a vertex fit with and without the considered track. Taking the event in the sketch as an example the reconstructed momentum of the $B_{(s)}^0$ meson ($p_{\text{reco}}(B_{(s)}^0)$) is hardly compatible

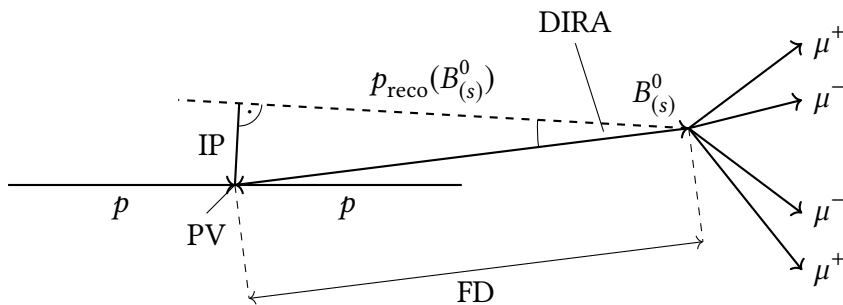


Figure 3.6: This sketch shows the signal part of a $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ event. Two protons collide at the PV, produce a $B_{(s)}^0$ meson, which travels a bit and then decays into four muons. The solid bold lines indicate the true trajectories of the individual particles. The bold dashed line represents the direction of the momentum of the reconstructed $B_{(s)}^0$ meson, $p_{\text{reco}}(B_{(s)}^0)$. Furthermore, the impact parameter (IP), flight distance (FD) and direction angle (DIRA) of the $B_{(s)}^0$ candidate with regard to the PV are indicated.

with originating from the PV. Therefore, the χ^2 of the PV fit is significantly different in fits with and without the $B_{(s)}^0$ candidate. Hence, low χ_{IP}^2 means good compatibility for a given track with originating from the regarded vertex. Another measure for the compatibility of a track to originate from a given vertex is the direction angle (DIRA). The DIRA is defined as the angle between the momentum of a track and a straight line connecting the regarded vertex to the decay vertex. The smaller the DIRA the more compatible a track is with originating from the regarded vertex. The flight distance (FD) is calculated as the distance between the regarded and the decay vertex of a particle. It is a valuable measure to reject tracks that were created in the proton-proton collision or to distinguish between long lived and short lived intermediate particles. Based in the FD the flight distance significance χ_{FD}^2 is defined. This quantity additionally incorporates the uncertainties originating from the underlying vertex fits, which is done by dividing the FD by a χ^2 calculated from the square of the distance between the two vertices divided by the sum of their variances. Yet another way to quantify if two tracks were created in the same interaction is given by the distance of closest approach (DOCA). In the case of more than two tracks typically the minimum or maximum DOCA of all combinations is regarded depending on whether to veto or to require the particles to originate from the respective vertex. The DOCA significance χ_{DOCA}^2 is defined analogously to the χ_{FD}^2 .

For the measurement of $R_{K\pi\pi}$ detailed in chapter 6 two isolation variables ISO_1^{vtx} and ISO_2^{vtx} are employed. These are defined as the minimum change in the χ_{vtx}^2 when adding one or two more tracks to the decay vertex of the B meson. Small values indicate that there are at least one or two more tracks in the event that are compatible with being produced by the decay of the B meson. This can be the case if the B meson was reconstructed only partially.

3.5 Trigger system

The bunch crossing rate at the LHCb experiment is 40 MHz, which corresponds to a rate of visible inelastic pp collisions of about 11 MHz. Not all collisions include processes that are of interest to the physics program of the LHCb experiment. On top of that saving events with an average size of 60 kB [73] to hard drives at this rate is not feasible as it would result in a data rate of more than 600 TB s^{-1} . Therefore, the event rate is reduced by a trigger system [74, 75], which consists of a hardware trigger (L0) and a high level trigger (HLT). The L0, implemented in custom made electronics operated synchronously with the 40 MHz bunch crossing frequency, decides based on signals in the calorimeter and the muon system if an event is worth saving or not. Thereby, the event rate is reduced to 1 MHz. At this rate the full detector can be read out such that the HLT is enabled to perform a full event reconstruction. The HLT is a software application that is run on the event filter farm consisting of around 30 000 CPU cores. In LHC Run 1 the output event rate of the HLT was between 2 and 5 kHz depending on the operating conditions.

The muon L0 selects events based on the (product of the) transverse momenta of muon candidates measured solely based on the muon stations. The demanded requirements vary between the years 2011 and 2012 and are detailed in Table 3.1. Additionally, requirements on the track multiplicity are imposed. Depending on whether the single muon or the di-muon criterion was satisfied the number of hits in the SPD is required to be less than 600 or 900.

The L0 electron trigger selects events based on the deposited energy in the ECAL. The highest cluster found in the ECAL is required to correspond to a transversal energy of 2.5 GeV (3 GeV) during 2011 (2012) data taking. Additionally, one or two hits in the PS and at least one hit in the SPD respectively in the cells lying before the regarded calorimeter cluster are demanded. The number of hits in the SPD is required to be smaller than 600 in order to avoid too long processing times in the HLT.

Table 3.1: Requirements applied to the muon candidates transverse momenta by the L0 muon triggers. [75]

Case	Requirement 2011	Requirement 2012
$p_T(\mu)$ [GeV/c]	> 1.48	> 1.76
$p_T(\mu_1) \times p_T(\mu_2)$ [GeV ² /c ²]	$> (1.296)^2$	$> (1.6)^2$

The HLT is subdivided into two stages, HLT1 and HLT2. For candidates that pass the L0 the HLT is evaluated and similarly events that satisfy the HLT1 are processed by the HLT2. At HLT1 level a partial event reconstruction is performed and requirements are made on the track quality, (transverse) momentum, displacement from the PV and two track combinations. The exact requirements that are imposed by the different HLT1 trigger lines that select the data, which is analysed in this thesis are detailed in Table 3.2. The muon part of HLT1 is only evaluated for events that were selected by

the L0 muon triggers. Contrary `TrackAllL0` is calculated for every event passing L0. Besides requirement on kinematic quantities, which were introduced in section 3.4, demands are made on the number of hits the respective track (did not) caused in different subdetectors of the tracking system.

Table 3.2: Requirements necessary for positive decisions of the considered HLT1 trigger lines. [74]

HLT1 Trigger Line	TrackAllL0	TrackMuon	DiMuonLowMass
IP [mm]	> 0.1	> 0.1	-
χ_{IP}^2	> 16	> 16	> 3
p_{T} [GeV/c]	> 1.7	> 1	> 0.5
p [GeV/c]	> 10	> 8	> 6
$\chi_{\text{trk}}^2/\text{ndf}$	< 2.5	< 2	< 4
Number VELO hits/track	> 9	-	-
Number missed VELO hits/track	< 3	-	-
Number OT+IT \times 2 hits/track	> 16	-	-
DOCA [mm]	-	-	< 0.2
χ_{vtx}^2	-	-	< 25
$m(\mu^+ \mu^-)$ [GeV/c ²]	-	-	> 1

The data analysed for the measurements presented in this thesis are selected by two different kinds of HLT2 trigger lines. The `DimuonDetached` trigger lines impose manually defined criteria that are typical for di-muon vertices displaced from the PV. These are summarised in Table 3.3. The second category is given by topological inclusive B trigger lines [76]. Here, combinations consisting of two, three or four tracks, which is reflected in the name of the respective line following the pattern `Topo[2, 3, 4]BodyBDT`, are built. Additionally, special cases for events containing electrons or muons (`Topo[2, 3, 4](Mu, E)BodyBDT`) exist exploiting the particle identification quantities introduced in section 3.2. In case of muons, at least one of the particles in the combination is required to fulfil the `IsMuon` criterion, while for the electron lines at least one particle must have `DLLe π > -2`.

The core of the topological trigger lines consists of a boosted decision tree (BDT) that identifies interesting events based on properties of the combination of two, three or four high quality tracks. A special BDT flavour called `Bonsai BDT` is employed that transforms the extensive `if-else` instruction, which make up an ordinary BDT into a look up table by discretising the input features. This is necessary to achieve the evaluation speed needed to run in real time during data taking.

Combinations of multiple tracks are build iteratively, requiring the DOCA between the first and the second track, or the already existing two or three track combination and the newly added track to be smaller than 0.2 mm. In order to enable partially

Table 3.3: Requirements imposed by HLTDimuonDetached HLT2 trigger lines. [74]

HLT2 Trigger DiMuon Line	Detached	DetachedHeavy
$\chi_{\text{trk}}^2/\text{ndf}$	< 5	< 5
χ_{IP}^2	> 9	-
$m(\mu^+\mu^-)$ [GeV/ c^2]	> 1	> 2.95
χ_{FD}^2	> 49	> 25
χ_{vtx}^2	> 25	< 25
$p_{\text{T}}(\mu^+\mu^-)$ [GeV/ c]	> 1.5	-

reconstructed decays causing a positive trigger decision the training features of the BDT include the corrected mass of the combination, calculated as

$$m_{\text{cor}} = \sqrt{m^2 + |p_{\text{T}}^{\text{miss}}|^2} + |p_{\text{T}}^{\text{miss}}|, \quad (3.3)$$

where m is the invariant mass of the combination and $p_{\text{T}}^{\text{miss}}$ is the missing momentum transverse to the direction of flight of the combination assuming it originates from the PV. Particles originating directly from the PV are rejected by demands on the IP or by requiring the mass of all underlying $(n - 1)$ -body objects to be larger than $2.5 \text{ GeV}/c^2$. Besides the corrected mass, the BDT is trained on the sum of the individual track transverse momenta, the minimum transverse momentum among the tracks in a combination, the mass of the combination, the DOCA, the χ_{IP}^2 and the χ_{FD}^2 .

In order to record as many data as possible, the trigger system is subject to permanent optimisation. Therefore, the requirements were not constant during the whole data taking period. The demands summarised throughout this section should be understood as references. A very detailed description of the different trigger configuration employed during LHCb Run 1 is given in Ref. [77].

3.6 Data processing

Data passing the criteria imposed by the trigger system are written to disk. When a significant amount of data has been collected or an important software update is released a stripping campaign is initiated. The stripping can be considered as a preparation of the data for efficient physics analysis by sorting and selecting the recorded events. The input are reconstructed tracks and information on these tracks, such as the response of the PID system. In particular for all data samples studied in this thesis only long tracks are considered. These are defined as tracks that were recorded in all subdetectors of the tracking system. In a first step, these tracks are organised in different containers assigning an identity to the underlying particle. An example for such a container is StdAllLooseMuons illustrated in Figure 3.7. In order for a track to enter this container, the IsMuon tag (see section 3.2) is demanded.

The stripping system follows a modular layout, such that one container is the input for the next. This allows for efficient data pre-processing and eases further extensions, for example if additional decays shall be studied. Staying with `StdAllLooseMuons` and the stripping selection of $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ candidates as an example, the next tighter container is `StdLooseDiMuon`. Here, two oppositely charged muons measured in the same event are combined and required to form a common vertex. For the definitions of the employed quantities, as well as others used in the following, see section section 3.4. This kind of successive container chain exists for all commonly used particles, such as kaons, pions, electrons and muons. Important to note here is that there is a significant overlap between the tracks in different containers. Regarding `StdAllLooseMuons` and `StdAllLooseKaons` that are illustrated in Figure 3.7 as an example the only difference between these two containers is the `IsMuon` requirement. As a result at this stage all muon candidates are also kaon candidates. The differentiation between different kinds of particles is mainly applied manually in later stages of the analyses as most stripping containers impose only soft PID criteria, if at all.

In the last container of each of the stripping lines that yield the data for the measurements presented in this thesis the desired final state is built. For the $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ stripping selection illustrated in Figure 3.7 requirements on the track quality and the ghost track classifier response [78] for the hadron candidates are applied. For muon candidates the selection is based on the transverse momentum and the χ_{IP}^2 with regard to the PV. Additionally, the invariant masses of the di-muon and the $K^+ \pi^+ \pi^-$ system are loosely demanded to be within the kinematically allowed ranges. Furthermore, the di-muon and $K^+ \pi^+ \pi^-$ combinations are each required to form a common vertex in terms of vertex fit quality. Candidates originating from the PV are rejected by applying demands on the χ_{IP}^2 and χ_{FD}^2 .

The reconstructed B^\pm candidate is demanded to from a good quality vertex, have a cosine of the direction angle to be larger than 0.9995, be compatible with originating from the PV in terms of χ_{IP}^2 and have a significant distance between origin and decay vertex (χ_{FD}^2). On top of that, a requirement is made on the invariant mass of the reconstructed B^\pm candidate, which allows for statistical modelling of the distribution by including sufficient data at both sides of the mass peak.

Stripping lines and their associated data are organised in streams depending on the nature of the final state particles. The data including the final states analysed for the search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ presented on chapter 5 is located in the `DIMUON` stream, while those analysed for the measurement of $R_{K\pi\pi}$ are located in the `LEPTONIC` stream. Therefore, only the output files of the respective stream need to be processed in order to acquire the data for the respective measurement. This brings huge advantages in terms of CPU and bandwidth usage.

In the search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays that is presented in chapter 5 three final states are regarded. Besides the B decays into four muons, $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ and $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ decays are studied. The respective requirements applied in the stripping are detailed in Appendix A.1.1.

The data analysed for determining the $R_{K\pi\pi}$ ratio originates from two different stripping lines, depending on the flavour of the final state leptons. The selection of the data containing the di-muon final state is sketched in Figure 3.7 and that for the di-electron mode in Appendix A.1.2. A particularity of the stripping selection for these modes in comparison to those employed for the search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays is that the di-lepton combination is already build in intermediate stripping containers, namely `StdLooseDiElectron` and `StdLooseDiMuon`. In the di-electron case the recovery of bremsstrahlung photons as detailed in section 3.3 is applied to all tracks that enter this container.

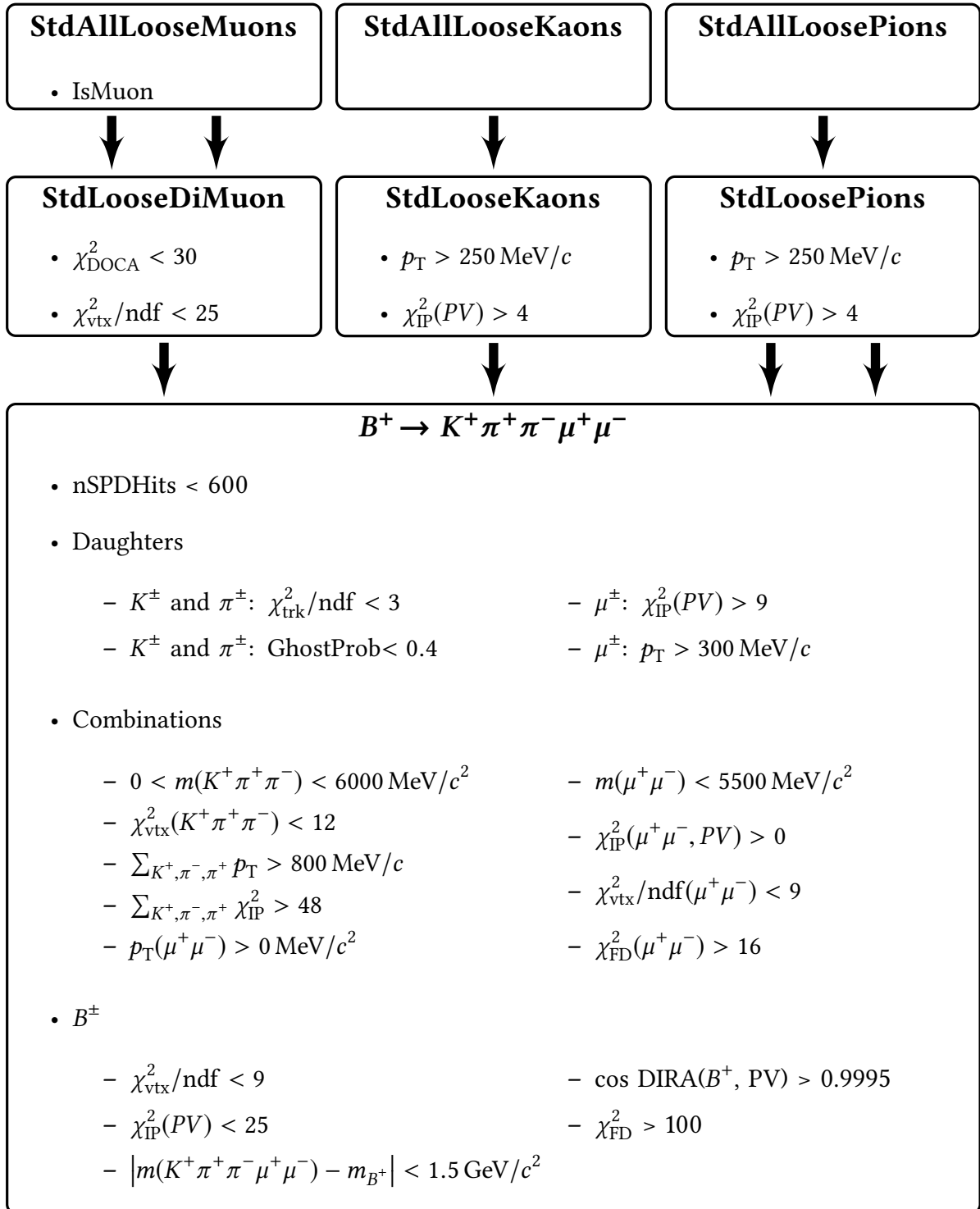


Figure 3.7: Sketch of the connection between individual stripping containers that are employed in selecting B^+ decays into the $K^+ \pi^+ \pi^- \mu^+ \mu^-$ final state.

4 Common techniques for rare decay analyses

In this chapter techniques are introduced that are commonly used in measurements of rare decays performed by the LHCb collaboration. Here, the basic principles are explained, while details such as the configuration and the outcome of the application are detailed in the corresponding analysis chapters.

4.1 Branching fraction measurements

The branching fraction of a decay of particle B into final state Y is defined as

$$\mathcal{B}(B \rightarrow Y) = \frac{N(B \rightarrow Y)}{N(B)}, \quad (4.1)$$

which means, that in order to measure a branching fraction not only the amount of decays of the type $B \rightarrow Y$ needs to be known, but also the amount of produced B particles. The latter in the case of B mesons can be calculated from the integrated luminosity \mathcal{L}_{int} , the $b\bar{b}$ production cross-section $\sigma_{b\bar{b}}$ and the hadronisation probability f . The resulting equation is

$$\mathcal{B}(B \rightarrow Y) = \frac{N(B \rightarrow Y)}{\mathcal{L}_{\text{int}} \times \sigma_{b\bar{b}} \times 2 \times f}, \quad (4.2)$$

which introduces large systematic uncertainties Especially $\sigma_{b\bar{b}}$ is known with an uncertainty of 17 % and 12 % for data taken in the years 2011 [79] and 2012 [80]. These systematic uncertainties can be avoided by measuring the $B \rightarrow Y$ branching fraction relative to a well known, experimentally easily accessible decay, such as $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$. Hereby, the integrated luminosity and the $b\bar{b}$ cross-section in equation (4.2) cancel out at the cost of the necessity to study an additional decay. The newly introduced uncertainty originating from the branching fraction of the normalisation mode, $\mathcal{B}(\text{norm})$, is typically at the percent level. The statistical uncertainty on the signal yield of the normalisation mode, $N(\text{norm})$ is negligible compared to the respective quantity of the rare mode. By introducing such a normalisation, equation (4.2) becomes

$$\mathcal{B}(B \rightarrow Y) = N(B \rightarrow Y) \times \frac{\mathcal{B}(\text{norm}) \times f_{q'}}{N(\text{norm}) \times f_q}, \quad (4.3)$$

where f_q and $f_{q'}$ are the probability for a b quark to hadronise with a u , d or s quark (f_u, f_d, f_s) depending on which flavour the B meson carries. An additional complication is introduced by imperfections of the detector, the reconstruction software and the selection applied to the data. All measured quantities need to be corrected by the efficiencies that arise from the geometric acceptance of the detector, imperfect track reconstruction and selection criteria that are applied to suppress background contamination. Taking these effects into account, equation (4.3) becomes

$$\mathcal{B}(B \rightarrow Y) = N(B \rightarrow Y) \times \eta, \quad (4.4)$$

where everything except for the number of signal candidates is encapsulated in the normalisation factor η defined as

$$\eta = \frac{\mathcal{B}(\text{norm})}{N(\text{norm})} \times \frac{f_u}{f_q} \times \frac{\varepsilon(\text{norm})}{\varepsilon(B \rightarrow Y)}. \quad (4.5)$$

The quantities that need to be measured are the amount of signal and normalisation decays and the corresponding efficiencies.

4.2 Hypatia function

In branching fraction measurements performed by the LHCb collaboration signal yields are usually determined by modelling the reconstructed B mass distribution. This is done utilizing the extended maximum likelihood method. Under the assumption that the measured mass of each B candidate has the same uncertainty it follows a Gaussian distribution. In reality the uncertainty of the measured mass depends for example on the momentum resolution which again depends on the momentum itself. As a consequence the mass uncertainties follow a continuous distribution. This is accounted for by the Hypatia function [81] by marginalising over the per-candidate mass uncertainty resulting in a generalised hyperbolic distribution as the core of the distribution. On top of that effects such as missed or wrongly assigned bremsstrahlung photons and also multiple scattering result in tails to both sides of the measured mass peak. These tails are accounted for by parametrising the left and right side of the distribution with power-laws of degree n_1 and n_2 using the transition parameters α_1 and α_2 as can be seen in equation (4.6). The presented parametrisation is the consequence of fixing the parameter λ to one and β to zero as it is done throughout this thesis. The consequence of this parameter choice is that the core of the distribution corresponds to a symmetric hyperbolic function as specified in Table 1 in Ref. [81]. The full definition of the Hypatia function used to parametrise the B mass distributions in both measurements presented in this thesis is

$$I(m|\mu, \sigma, \zeta, \alpha_1, n_1, \alpha_2, n_2) \quad (4.6)$$

$$\propto \begin{cases} G(m, \mu, \sigma, \lambda, \zeta) & \text{for } -\alpha_1 < \frac{m-\mu}{\sigma} < \alpha_2, \\ \frac{G(\mu+\alpha_1\sigma, \mu, \sigma, \zeta)}{(1-m/(n_1 H(\mu+\alpha_1\sigma, \mu, \sigma, \zeta) - \alpha_1\sigma))^{n_1}} & \text{for } -\alpha_1 > \frac{m-\mu}{\sigma}, \\ \frac{G(\mu-\alpha_2\sigma, \mu, \sigma, \zeta)}{(1-m/(n_2 H(\mu+\alpha_2\sigma, \mu, \sigma, \zeta) - \alpha_2\sigma))^{n_2}} & \text{for } \alpha_2 < \frac{m-\mu}{\sigma}, \end{cases}$$

with m being the modelled observable, e.g. the B mass and

$$G(m, \mu, \sigma, \lambda, \zeta) = ((m - \mu)^2 + A^2(\zeta)\sigma^2)^{\frac{1}{4}} K_{\frac{1}{2}} \left(\zeta \sqrt{1 + \left(\frac{m - \mu}{A(\zeta)\sigma} \right)^2} \right), \quad (4.7)$$

$$H(\mu + \alpha_2\sigma, \mu, \sigma, \zeta) = \frac{G(m, \mu, \sigma, \lambda, \zeta)}{G'(m, \mu, \sigma, \lambda, \zeta)}, \quad (4.8)$$

$$A^2(\zeta) = \frac{\zeta K_1(\zeta)}{K_2(\zeta)}. \quad (4.9)$$

Hereby $K_\lambda(\zeta)$ are modified Bessel functions of the second kind. The parameters μ and σ can be interpreted as mean and width of the core distribution while ζ changes the curvature of the flanks of the distribution as illustrated in Figure 4 in Ref. [81].

4.3 Backgrounds and how to deal with them

Before the experiment started operation, LHCb data were expected to suffer significantly more from background contamination than data recorded at B factories. However, as can be seen in Figure 4.1, LHCb data samples yield extremely clean signals after the right treatment, which is detailed in the following. Backgrounds can be split into several categories, regarding their origin and also the measures to reduce them. Figure 4.1 shows data corresponding to 3 fb^{-1} of proton-proton collisions mea-

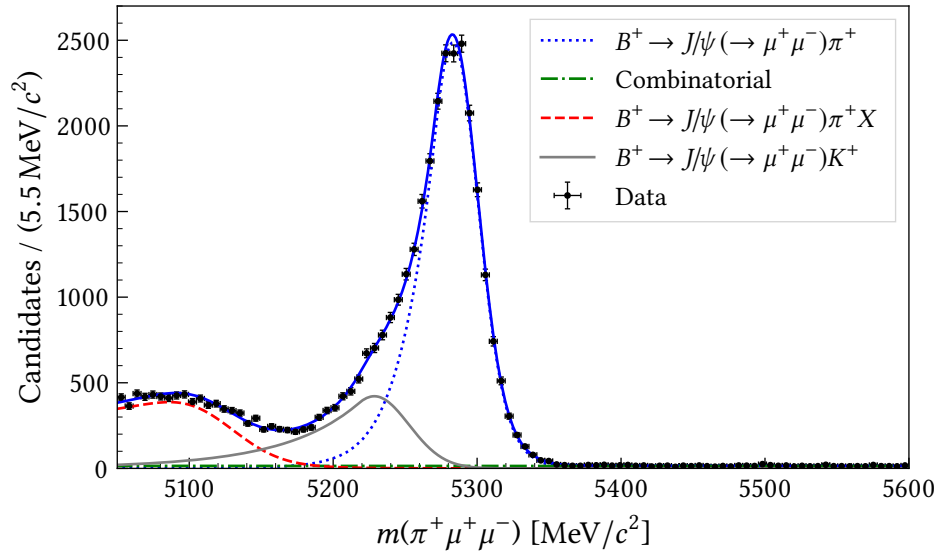


Figure 4.1: Distribution of $\pi^+ \mu^+ \mu^-$ invariant mass in the LHCb Run 1 data sample after applying a selection to measure $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \pi^+$ decays. Several components are identified and modelled by different PDFs as explained in the text.

sured by LHCb during Run 1. The data are selected in order to extract the amount of $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \pi^+$ signal candidates and was used for a measurement that is not

presented in this thesis [82]. Nevertheless, this distribution is well suited to discuss the composition of LHCb data. Besides the large signal component that is described by the blue dotted line and peaks at the known B mass three different kinds of background are considered in the model that is fitted to data.

One category is made up by partially reconstructed decays, where one or more decay products are not reconstructed, and hence results typically in a lower invariant mass of the reconstructed B meson. For muonic final states the mass resolution of the LHCb experiment is sufficient to exclude partially reconstructed decays by rejecting candidates with an invariant B mass lower than $5175 \text{ MeV}/c^2$. In the case of electron final states, the emission of bremsstrahlung compromises the mass resolution, such that a shape for partially reconstructed background needs to be included in the fit model. In the example, partially reconstructed decays are represented by the red, dashed line. They mainly originate from $B^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^{*0} (\rightarrow K^+ \pi^-)$ decays, where the kaon is not reconstructed.

Another kind of background arises from particle misidentification, where one or more particles in the final state are wrongly identified. Depending on the mass difference between the true particle and the assigned identity, the measured B mass is altered. The LHCb experiment's particle identification system, most dominantly the RICH detectors and the muon system, allow the calculation of particle identification variables as detailed in section 3.2. By placing requirements on these variables, candidates with misidentified final state particles can be efficiently suppressed. Placing selection requirements on data is always a trade off between background reduction and signal efficiency. Therefore, in some cases it is beneficial to incorporate this type of background in the fit model as done in the example by the grey solid line, instead of reducing it to a negligible level. Here, $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ decays, where the kaon is falsely identified as a pion, cause a structure peaking at $5229 \text{ MeV}/c^2$ with a large tail towards lower masses. The shape is fixed in the presented fit and is extracted from selected $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ data, where the kaon mass is swapped for the pion mass in the reconstruction.

The most frequent kind of background in raw data is combinatorial, which arises when the sought-after final state is made up by a random combination of particles originating from different decays. The fact that the final state particles do not originate from the same B decay compromises the decay tree fit quality and affect also the kinematic distributions of the reconstructed B meson. The fraction of combinatorial background that pass cut based selection steps are typically suppressed further by applying a machine learning algorithm that is trained to perform a binary classification between signal and combinatorial background. In the example the combinatorial background is modelled by an exponential function represented by the green dash-dotted line. The data presented in Figure 4.1 already passed a machine learning based selection, such that the combinatorial background has already been reduced significantly.

4.4 Background subtraction with the $s\mathcal{P}$ lot method

The data studied throughout this thesis consist of different components as discussed in section 4.3. One main task of data analyses in high energy physics is to isolate the sought-after signal component from background. This can for example be achieved by *sideband subtraction*, which means that tight cuts around a peak in the invariant mass of a reconstructed particle are applied. This approach brings the disadvantages that only a subset of the available data are used and that distributions of other variables in the same data sample might be biased. Another method widely used in particle physics to deconvolute a mixed sample and to determine the signal yield is the unbinned extended maximum likelihood fit. For studies of rare decays at LHCb typically the invariant mass distribution of the reconstructed B meson is modelled like in the example in section 4.3. The fit results in one PDF for each individual component that is represented in the fit model describing its distribution. Often not just the invariant B mass distribution of the signal component is of interest, but also for example the transverse momentum distribution. Therefore, the $s\mathcal{P}$ lot technique [83] has been developed. It introduces a formalism to calculate weights based on the mentioned fit result. These $sWeights$ can be used to project out the signal (or any other) component of any dimension that is uncorrelated with the modelled dimension.

Assuming a fit model consisting of N_S different components, the total PDF is given by the sum of the sub-PDFs f_i describing the individual classes of candidates,

$$\mathcal{F} = \sum_{i=1}^{N_S} N_i f_i. \quad (4.10)$$

This model is fitted to data employing the extended maximum likelihood method, which means that N_i are not normalisation factors, but Poissonian terms. Applying the maximum likelihood method not only estimates the shape parameters, but also the number of candidates that are described by each individual sub-PDF or with other words belong to the class of candidates that is described by the regarded sub-PDF. Based on the model, $sWeight$ for the n th class are calculated using the equation

$${}_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N_S} V_{nj} f_j(y_e)}{\sum_{k=1}^{N_S} N_k f_k(y_e)}, \quad (4.11)$$

where V_{nj} is the correlation between the yields of the classes n and j , f_j and f_k are the sub-PDFs describing the respective classes and y_e is the discriminative variable, in our case the reconstructed B mass. These weights can then be used to create histograms that show the distribution of the corresponding category in an arbitrary feature that must be uncorrelated to the dimensions used for discrimination. Additionally, efficiencies can be measured for one of the modelled classes by counting the sum of the respective weights before and after applying a selection.

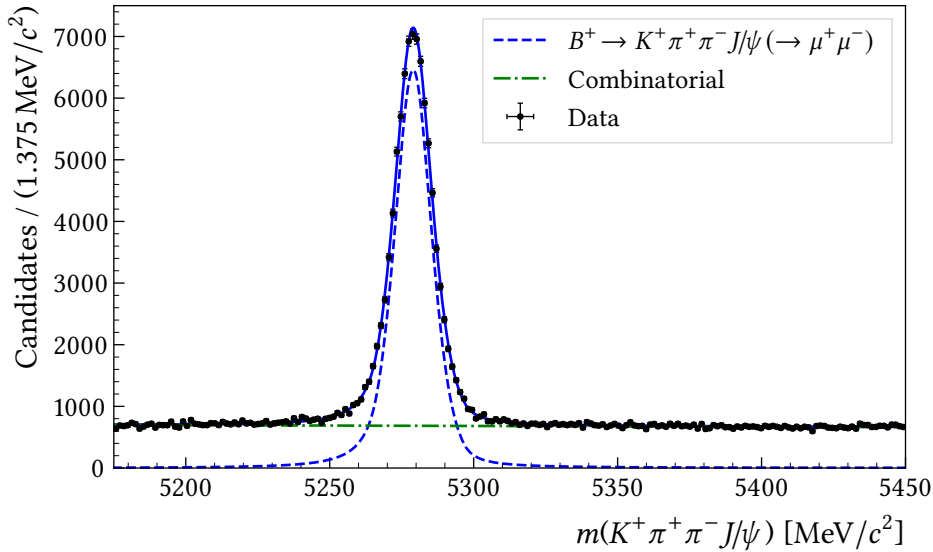


Figure 4.2: Invariant mass distribution of the $K^+ \pi^+ \pi^- \mu^+ \mu^-$ system. Here, the invariant di-muon mass is constrained to that of the J/ψ meson. The distribution is modelled by a PDF accounting for $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \mu^+ \mu^-)$ decays and combinatorial background.

As an example $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \mu^+ \mu^-)$ data before applying the final part of the selection (see chapter 6), which reduces backgrounds to a negligible level, is used. As can be seen in Figure 4.2, the invariant mass of the reconstructed B meson with the di-muon system constrained to the J/ψ mass is parametrised by a model consisting of one component for the signal and one for the combinatorial background candidates. Based on the presented fit *sWeights* are calculated and used in the following sections to compare simulated and measured data.

4.5 Simulation correction

Generally speaking measured data are well reproduced by LHCb simulation. There are only a small number of features that are known to be not well reproduced and must be corrected in order to be able to extract accurate efficiencies and to train well performing machine learning classifiers. Besides the PID response, which will be discussed in section 4.6, the occupancy and the momentum distribution of B mesons are among the quantities that need to be corrected. One major challenge arises from correlations, which can result in degrading one feature when correcting another. Therefore, it is advisable to correct correlated features simultaneously. However, this quickly leads to problems originating from the curse of dimensionality. In the following two methods to improve the agreement between simulated and measured data are discussed.

Histogram based approach

The most common approach is to create normalised multi-dimensional histograms, where the dimensions are given by the features that need to be corrected in simulated data. One histogram is filled from background subtracted measured data, the other from simulated data. Dividing the histogram filled from measured data by that filled from simulated data in a bin-wise manner results in a histogram of weights that map the considered distributions in simulation to those in data. This approach has multiple disadvantages. One is the choice of the binning scheme, which needs to be a compromise of sufficient population in each bin and good sensitivity to the shape of the distributions. The main problem though is that the number of bins and also the number of dimension are strongly limited by the curse of dimensionality. The population density is inverse proportional to the volume, which increases exponentially with the number of dimensions. Therefore, at maximum three dimensions can be corrected simultaneously with the histogram approach given the size of the LHCb Run 1 data samples. Adding more dimensions makes the procedure unreliable by introducing extremely large weights originating from small denominators when dividing the two histograms.

Machine learning approach

A novel way to avoid the arbitrariness in the choice of a binning scheme and also use the available amount of data in an optimal way is given by the GB-reweighting algorithm [84]. The method is inspired by the gradient boosting technique, hence the name. It is trained on simulated and background-subtracted measured data. The core of the algorithm consists of decision trees. The cut optimiser of the individual trees splits real data (RD) and simulated data (MC) in a way that maximises

$$\chi^2 = \sum_{l=1}^{l=N_l} \frac{(w_l(\text{MC}) - w_l(\text{RD}))^2}{w_l(\text{MC}) + w_l(\text{RD})}, \quad (4.12)$$

where N_l is the number of leaves and $w_l(\text{RD})$ ($w_l(\text{MC})$) is the sum of weights made up by the RD (MC) part in the l th leaf. After each iteration, weights corresponding to

$$w = \begin{cases} w, & \text{for RD,} \\ w \times \frac{w(\text{RD})}{w(\text{MC})}, & \text{for MC,} \end{cases} \quad (4.13)$$

are assigned to the training data, where $w(\text{MC})$ and $w(\text{RD})$ are the sum of weights made up by RD or MC data in the respective leaf.

This technique is very similar to the previously described histogram approach. Each individual decision tree of the GB-reweighting can be seen as a histogram with the dimensions and the binning scheme configured in a way that reveals the largest differences between measured and simulated data. The formula used to calculate the individual weights is the same for both methods. The GB-reweighting is an iterative

algorithm, that applies this optimised histogram approach many times, each time taking into account the weights from the previous step.

The maximum number of bins (leafs) of each decision tree is defined by the maximal depth (d) and amounts to 2^d . This technique ensures reliability also in higher dimensions, as the number of features used in a split and also the minimum amount of data required in a leaf are configurable, such that large weights can be avoided.

4.6 Particle identification calibration

With every requirement on the PID criteria not only is the background from particle misidentification reduced, but signal is also lost. As discussed in section 4.1 the exact knowledge of how much signal is lost is a crucial input for branching ratio measurements. The efficiencies of the PID requirements cannot be easily measured on simulated data as the response of the particle identification system is not well modelled. Figure 4.3 shows a PID quantity before and after calibration in comparison to measured data. For this particular case it can be seen that the performance of the PID system is overestimated in simulated data, but nicely recovered by the calibration. Two different kinds of PID calibration techniques are used at the LHCb experiment [85]. One is to calculate efficiencies of PID requirements directly from calibration data and the other to correct PID distributions in simulated data based on calibration data. In order to avoid any bias from background contamination the signal part of the calibration data is projected out by applying the \mathcal{P} lot technique (see section 4.4).

Calibration data for kaons and pions are taken from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays. From the charge of the pion created in the decay of the D^{*+} meson, the flavour of the D^0 meson is deduced. Based on this information, the kaon and the pion from the D^0 decay can be identified unambiguously through their charge.

For electrons and muons, $B^+ \rightarrow J/\psi(\rightarrow \ell^+\ell^-)K^+$ decays where $\ell^+\ell^-$ is either e^+e^- or $\mu^+\mu^-$ are used. Here, the data are selected by a tag and probe method, which means that a stringent PID requirement is applied to one of the two leptons while the other is used for the calibration.

For protons the calibration data are taken from $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow pK^-\pi^+)\pi^-$ decays. Protons from Λ decays typically have a significantly lower momentum than those created in heavy-flavour decays. Therefore, also protons from Λ_b^0 decays are included.

An intuitive way to measure the efficiency of an arbitrary requirement on those calibration data sets is to take the ratio of their sizes before and after placing the requirement. Unfortunately, this will not yield accurate efficiencies as for example the track kinematics and the occupancy of the detector have a large influence on the PID response of the LHCb detector. Two different approaches exist to account for kinematic and occupancy effects in the efficiency calculation. One very intuitive method is to create an efficiency map, which comes with the downside of the necessity to recreate the map every time a new requirement is evaluated. A more practical

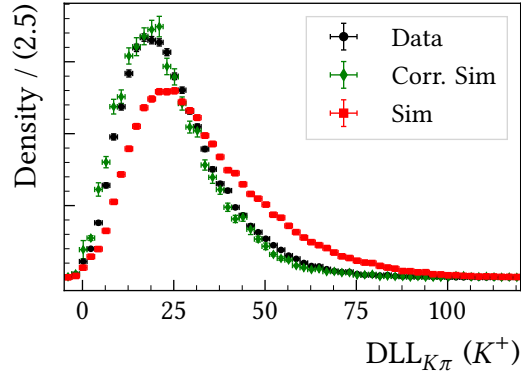


Figure 4.3: Log-likelihood based PID classifier for a kaon candidate to be a kaon in $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \mu^+ \mu^-)$ simulation before and after applying the PID calibration and in respective measured data employing *sWeights* to avoid contributions from background.

approach implies replacing the PID response for every track in the data set, which is referred to as resampling. Here, again different techniques are available. All methods that are used throughout this thesis are briefly explained in the following.

PID efficiency maps

Efficiency maps address the previously mentioned dependence of the efficiency by creating a histogram of efficiencies in dimensions reflecting track kinematics and detector occupancy [85]. The maps are created by filling two histograms with the *sWeight* of each candidate before and after applying a certain PID criterion and then dividing these two bin by bin. Typically this is done in three dimensions, which are the track multiplicity in the event, the momentum of the track and the pseudo rapidity of the track. One main problem of this method is the choice of a binning scheme. A compromise between the sensitivity to the underlying probability density function and the statistical uncertainty given by the population in each bin is to be made. The efficiency of the underlying PID requirement is measured based on a simulated sample of the studied decay. For each simulated candidate the efficiency from the corresponding bin in the efficiency map is looked up. The efficiency of the applied selection is computed by calculating the average.

Classical resampling approach

In order to be able to easily probe different PID selections for a given data sample the per track value of the PID response is redrawn [85]. In the classical resampling approach, the calibration data are split in regions comparable to the bins of the efficiency map approach. This results in one distribution of the regarded PID variable for each bin. For a given track the corresponding bin is looked up and a new PID response

is drawn. Hereby, the probability to draw an individual value is proportional to the *sWeight* assigned to each entry in the calibration data. This method again involves the choice of a binning scheme, which again comes with the problems originating from the finite size of the calibration data sample.

Resampling with Meerkat

An approach that avoids binning the data is given by the Meerkat resampling technique [86]. Here, a multi-dimensional kernel density estimation (KDE) is performed on the calibration data, where the dimensions are the regarded PID response and its dependencies, e.g. kinematic and detector occupancy variables. Instead of looking up the distribution to sample from in the corresponding bin, the KDE is evaluated and for the probability to draw an individual value not the *sWeight* of each individual value from calibration data, but the amplitude of the KDE is used. This technique is referred to as *PIDGen* and gives good agreement for the individual PID quantities. However, if aggregated PID quantities are employed this method is not sufficient because here the PID responses are redrawn, which breaks the correlation with any other variable than those used in the parametrisation. This issue is resolved by the *PIDCorr* method. Here, the PID response in simulation is modelled in the same way as in data. Based on the two KDEs a transformation is calculated that translates the simulated PID response into a new distribution that matches the calibration PID response far better. As a consequence, variables that are correlated to the simulated PID response are also correlated to the corrected PID response. Especially for the neural network based ProbNN variables, which depend on many more features than just the kinematics and occupancy, this is beneficial. Furthermore, when placing requirements on combinations of PID variables or training a machine learning algorithm on them, the correct reproduction of the correlation is crucial. The downside of this method is that a large amount of simulated data is required for the KDE to be reliable throughout the whole phase space.

4.7 Data driven trigger efficiencies

The response of the trigger system is not perfectly reproduced in the LHCb Monte Carlo simulation. In order to determine accurate efficiencies, the TisTos method [87] has been developed. It allows to measure trigger efficiencies on measured data by exploiting the fact that different parts of an event can cause a positive trigger decision. Throughout this thesis mostly particles that belong to the studied signal decay are regarded. For this method however it is important to keep in mind that there are far more particles in an event than the three to five that make up the studied final states. The signature of any particle can yield a positive trigger decision, which will cause the whole event to be saved. And a peculiarity of the LHCb trigger system is that it saves the information which particle is responsible for a positive trigger decision.

Events that pass the trigger are categorised into three types, which are Triggered On Signal (TOS), where the part of the event that belongs to the signal satisfies the trigger requirements, Triggered Independent of Signal (TIS), where the part of the event that does not belong to the signal meets the trigger criteria and Triggered On Both (TOB) where both parts were needed to yield a positive trigger decision. Important to note is, that events can be TIS and TOS at the same time.

Generally the trigger efficiency $\varepsilon_{\text{Trigger}}$ can be written as

$$\varepsilon_{\text{Trigger}} = \frac{N_{\text{Trigger}}}{N} \quad (4.14)$$

$$= \frac{N_{\text{Trigger}}}{N_{\text{TIS}}} \frac{N_{\text{TIS}}}{N} \quad (4.15)$$

$$= \frac{N_{\text{Trigger}}}{N_{\text{TIS}}} \varepsilon_{\text{TIS}}, \quad (4.16)$$

where N is the number of candidates before the trigger selection, N_{Trigger} is the number of candidates that pass the trigger selection and N_{TIS} is the number of candidates belonging to the TIS category. If the signal part of an event is completely uncorrelated to the rest of the event, ε_{TIS} can simply be calculated as

$$\varepsilon_{\text{TIS}} = \frac{N_{\text{TIS\&TOS}}}{N_{\text{TOS}}}. \quad (4.17)$$

Inserting equation (4.17) into equation (4.16) yields

$$\varepsilon_{\text{Trigger}} = \frac{N_{\text{Trigger}}}{N_{\text{TIS}}} \frac{N_{\text{TIS\&TOS}}}{N_{\text{TOS}}}. \quad (4.18)$$

From this step it is straight forward to calculate the amount of signal candidates before the trigger selection by

$$N = \frac{N_{\text{Trigger}}}{\varepsilon_{\text{Trigger}}} \quad (4.19)$$

$$= \frac{N_{\text{TIS}} \times N_{\text{TOS}}}{N_{\text{TIS\&TOS}}}. \quad (4.20)$$

The part of the trigger that is relevant for the measurements presented in this thesis is designed to select decays of B mesons mainly based on transverse momentum (p_{T}) and impact parameter (IP) requirements. Due to the fact that a B meson contains a b quark, which is produced in pairs, it is very likely that the part of the event that is responsible for firing the TIS decision originates from the decay of a b hadron that contains the respective other b quark. This b hadron is correlated to the B meson as far as the momentum spectrum is concerned, which breaks the assumption. This problem is addressed by applying the TisTos method in small bins of the B meson phase space, for example in momentum and pseudo rapidity, such that the number of candidates before the trigger selection becomes

$$N = \sum_{i=1}^{N_{\text{bins}}} \frac{N_{\text{TIS}}(i) \times N_{\text{TOS}}(i)}{N_{\text{TIS\&TOS}}(i)}, \quad (4.21)$$

with $N(i)$ being the number of candidates belonging to the respective category in the i th bin. In Ref. [87] the trigger efficiency measured on simulation using the TisTos method is compared to its true value. It is reported that not accounting for the phase space dependence by using only one bin results in an efficiency that is overestimated by 5.5 %. This shows that a carefully chosen binning scheme is crucial to determine an accurate efficiency and as a consequence reduce the resulting systematic uncertainty. The statistical uncertainty of the TisTos method is driven by the size of the population in each bin. As the least represented category is TisTos, a good approach for constructing a binning scheme is choosing the bin borders in a way such that equal amounts of TisTos candidates fall into each bin. The optimal binning scheme is a compromise between agreeing central values of the efficiencies measured on simulation by means of Monte Carlo truth information and by the TisTos method while achieving a low statistical uncertainty when applying the TisTos method to measured data.

5 Search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays

This chapter describes the search for non-resonant $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays carried out by the LHCb collaboration based on a data sample of proton-proton-collisions corresponding to an integrated luminosity of $\mathcal{L}_{\text{int}} = 3 \text{ fb}^{-1}$. The measurement was performed by a team of three people and was published in March 2017 in JHEP¹ [88]. My main contributions are the normalisation, the estimation of systematic uncertainties and the validation of the analysis strategy. The chapter is structured as follows: First, the strategy is explained, then the signal selection and studies concerning potential signal contamination from physical background processes are detailed. After that the signal efficiencies and systematic uncertainties are evaluated and several studies regarding the analysis validation are performed. Finally, the extraction of the results is described followed by a conclusion.

5.1 Strategy

The main goal of this measurement is the search for non-resonant $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays. To accomplish this a very stringent signal selection needs to be developed that efficiently rejects almost all background candidates. In order to minimise the systematic uncertainty, the search is performed normalised to the decay $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ as described in section 4.1. This measurement is performed blindly, which means that the signal region in reconstructed $B_{(s)}^0$ mass is only investigated after the complete analysis strategy is finalised and the developed workflow is validated through several cross-checks. This procedure reduces chances to bias the measurement by fine tuning the analysis based on the observed signal. The quantities that are measured before *unblinding* are the amount of signal candidates in the normalisation mode and the efficiencies of signal and normalisation mode. The SM branching fraction estimates discussed in section 2.2 are beyond the sensitivity of the analysed data sample. Therefore, a machinery based on the CL_s method is prepared to calculate (expected) upper limits on the branching fractions. Additionally to the SM decay, the sensitivity to a process mediated by new intermediate particles predicted by the MSSM as introduced in section 2.2 is probed. This is done based on simulated data of the decays $B_{(s)}^0 \rightarrow S (\rightarrow \mu^+ \mu^-) P (\rightarrow \mu^+ \mu^-)$ that proceed via a new pseudoscalar and a new scalar particle denoted as P and S . Their masses are set to $m(P) = 214 \text{ MeV}/c^2$ and $m(S) = 2.5 \text{ GeV}/c^2$. The first choice is motivated by the HyperCP result [35] and the latter maximises the theoretical branching fraction [37]. Two different lifetimes

¹Within the LHCb collaboration I am one of two contact authors of this publication.

of both particles are simulated. One compatible with a prompt decay, the other allowing for a displaced vertex. Upper limits are only evaluated for the first case as this measurement is optimised for the SM process and hence loses sensitivity if intermediate particles with non-negligible lifetimes cause the muons to originate from more than one distinguishable vertex. In order to validate and where needed to correct the simulated data $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ decays are examined.

5.2 Selection

The selection developed in the following is specifically optimised for the SM signal. For this purpose the corresponding signal simulation is generated with a uniform probability across the decay phase space. As the LHCb detector does not instrument the full solid angle, events where the decay products do not propagate into the geometric acceptance of the experiment are discarded as early as possible in order to save computation time. The resulting efficiency is measured as the ratio of accepted to generated events. Moreover, the simulated sample is processed by the tracking and stripping software in the same step, such that the respective efficiencies are measured simultaneously.

The first selection criteria are applied already during data taking by the trigger system. For this measurement data are analysed that pass at least one of the L0, one of the HLT1 and one of the HLT2 lines given in Table 5.1. Only candidates where the signal part of the event caused a positive trigger decision are considered. The underlying selection criteria are detailed in section 3.5.

Table 5.1: Trigger lines that select the data analysed for this measurement. For a detailed discussion about the underlying criteria see section 3.5.

L0	HLT1	HLT2
Muon	TrackAllL0	Topo2BodyBBDT
DiMuon	TrackMuon	Topo3BodyBBDT
	DiMuonLowMass	Topo4BodyBBDT
		TopoMu2BodyBBDT
		TopoMu3BodyBBDT
		TopoMu4BodyBBDT
		DiMuonDetached
		DiMuonDetachedHeavy

The first step of the offline selection is applied centrally in the stripping as detailed in section 3.6. The three studied final states are selected by the stripping lines `B24MuLine`, `Bu2JPsiKLine` and `Bs2JPsiPhiLine`. The precise selection criteria are illustrated in Figures A.1 to A.3.

Table 5.2: Selection windows of resonances.

Particle	Mass range
ϕ	950–1090 MeV/ c^2
J/ψ	3000–3200 MeV/ c^2
$\psi(2S)$	3600–3800 MeV/ c^2

In the preselection the number of hits in the SPD is required to be fewer than 600 to restrict the analysis to events with a reasonable detector occupancy. Additionally, the amount of fake tracks is reduced by requiring the ghost track classifier response [78] to be less than 0.3, which is a general recommendation made by the LHCb tracking group. The preselection is common for all decays that are studied during the search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays.

Data that proceeds via intermediate states as well as as non-resonant data are selected by requirements on the invariant mass of two-particle combinations. For non-resonant data the three mass windows specified in Table 5.2 are removed from the invariant mass distributions of all di-muon combinations with neutral charge. In order to validate the extensive use of simulated data, a large sample of real data that is as similar as possible to the signal decay is needed. Here, $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ is chosen, where the reconstructed masses of the J/ψ and the ϕ meson are required to be within the corresponding intervals defined in Table 5.2. Similar demands are made on the J/ψ mass for the normalisation mode $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$. For a cross-check (see section 5.5) the branching ratio of the decay $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow \mu^+ \mu^-)$ is measured. Here, signal candidates are selected by requiring one pair of oppositely charged muons to be in the mass interval corresponding to the ϕ meson, while the other pair must be in the J/ψ range.

Due to the large branching fraction and the clean signal from the $J/\psi \rightarrow \mu^+ \mu^-$ decay, there is no complicated strategy needed to extract the $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ signal. Additionally to the preselection, only cuts to log-likelihood based PID variables are applied. In detail, the muon candidates are demanded to be more likely a muon than a pion ($DLL_{\mu\pi} > 0$) and more likely a pion than a kaon ($DLL_{K\pi} < 0$). On top of that, the difference in log-likelihood between the kaon-candidate being a kaon or a pion ($DLL_{K\pi}$) is required to be larger than five. In Figure 5.1 the $DLL_{K\pi}$ distribution from kaon and pion PID calibration data is shown. The mentioned criterion is indicated by the dashed line. It can be seen that the vast majority of pion candidates fail the applied requirement. Additionally, the branching fraction of the $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ decay is about 26 times larger than that of the $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \pi^+$ decay [28], such that contributions due to misidentification are negligible.

Before developing the final part of the selection, the agreement between measured and simulated data is tested by comparing the corresponding samples of $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ decays. The signal component in measured data is

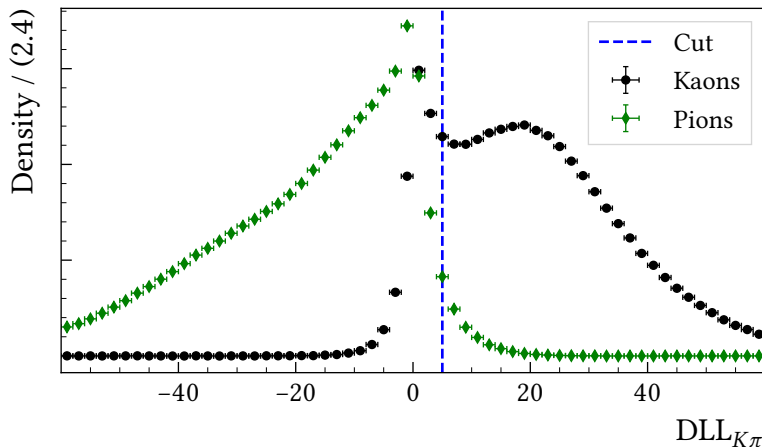


Figure 5.1: The difference in log-likelihood between kaon and pion hypothesis for kaon and pion PID calibration data. The line indicates the requirement applied to select $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$ candidates.

isolated from backgrounds through the \mathcal{P} lot technique as described in section 4.4. Therefore, an extended maximum likelihood fit to the invariant mass of the $\mu^+ \mu^- K^+ K^-$ system is performed. The signal component of the distribution is modelled by a Hypatia function [81] (for more details see section 4.2) and combinatorial background by an exponential function. Figure 5.2 shows the distribution and the fitted model for data passing different PID criteria. In Figure 5.2a a tail towards the right hand side of the signal peak can be seen. This feature is created by $B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^{*0}(\rightarrow K^+ \pi^-)$ decays, where the pion originating from the decay of the K^* meson is wrongly identified as a kaon. Placing requirements on the PID of the two kaon candidates reduces the tail as can be seen in Figures 5.2b and 5.2c. These requirements also reduce the level of combinatorial background, such that a small peak at the B^0 mass becomes visible. The influence of the different fit results of the three presented scenarios on the final result are tested and no significant alterations are observed.

Some quantities that are important for distinguishing signal from background are known to be simulated inaccurately. Therefore, the gradient boosting reweighting technique (see section 4.5) is employed to correct the imperfectly simulated track multiplicity, B transverse momentum, the χ^2 of the decay vertex fit and the smallest χ_{IP}^2 with respect to any PV. The first two quantities are correlated with the particle identification variables and the latter two are the most important features of the multivariate classifier that is trained in the next step of the selection. The weighting algorithm is trained on simulated and measured $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow K^+ K^-)$ decays. The background in measured data is suppressed through the \mathcal{P} lot technique as explained in section 4.4. The trained algorithm is applied to every simulated sample used for this measurement. Distributions of these four features in measured and in simulated data before and after applying the weights are presented in Figure 5.3. Good agreement between corrected simulation and measured data is observed.

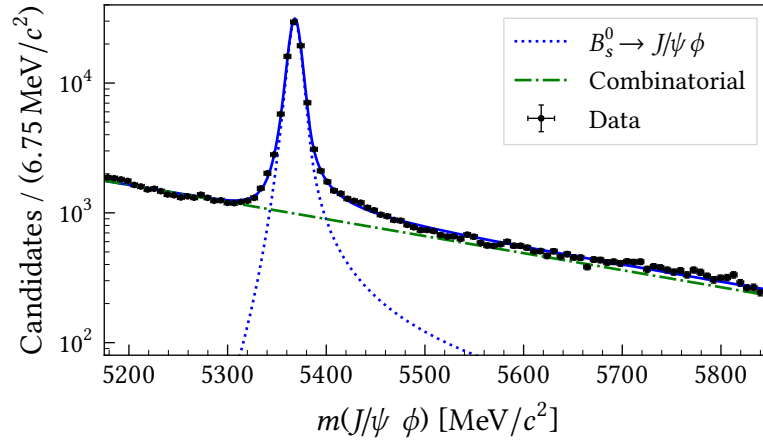
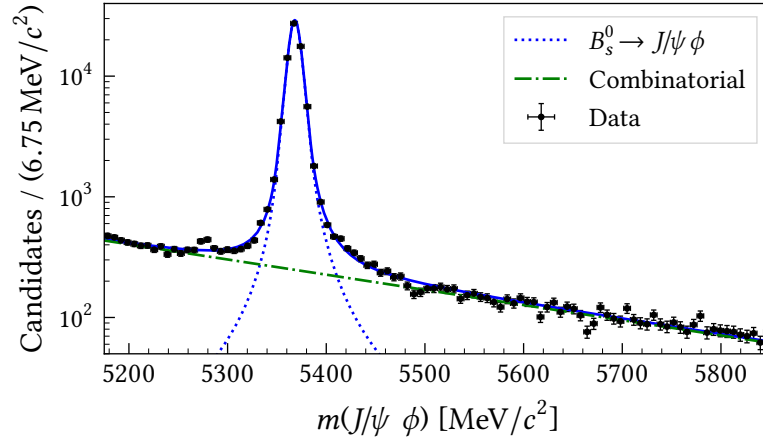
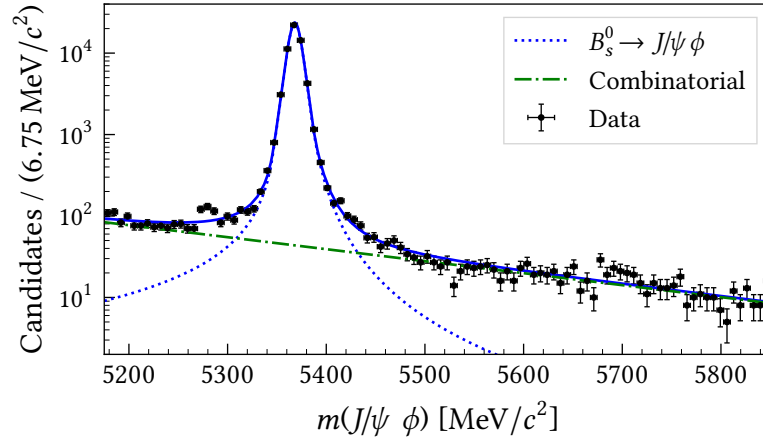
(a) No ProbNN_K(b) ProbNN_K > 0.02(c) ProbNN_K > 0.2

Figure 5.2: Invariant mass distribution of the $\mu^+ \mu^- K^+ K^-$ system with J/ψ and ϕ mass constraints in data corresponding to an integrated luminosity of 3 fb^{-1} passing the trigger and preselection. From top to bottom different thresholds on ProbNN_K for both kaon candidates are applied.

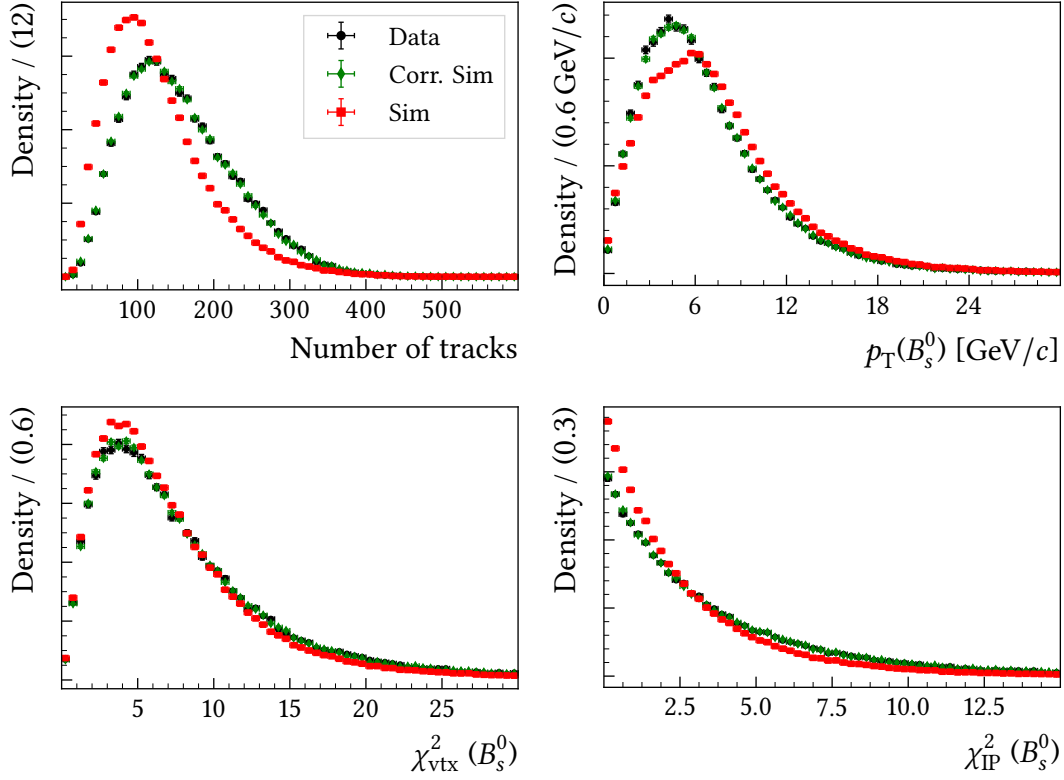


Figure 5.3: Distributions of the set of features used as input into the weighting algorithm, namely the number of tracks in the event, the transverse momentum of the reconstructed B_s^0 meson, the χ^2/ndf of the B_s^0 vertex fit and the χ_{IP}^2 of the B_s^0 meson. Black circles represent $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ decays in measured data isolated from background through the $s\mathcal{P}$ lot technique (see section 4.4), red squares raw simulation and green diamonds corrected simulation. The shown variables are explained in section 3.4.

Table 5.3: Definitions of intervals in the B^0 and B_s^0 reconstructed invariant mass distributions.

	Mass interval (MeV/ c^2)
Near sidebands	[5020, 5220] and [5426, 5626]
Far sidebands	[4360, 5020] and [5626, 6360]
Signal region	$[m(B^0) - 60, m(B_s^0) + 60]$
B_s^0 search region	$[m(B_s^0) - 40, m(B_s^0) + 40]$
B^0 search region	$[m(B^0) - 40, m(B^0) + 40]$

Another set of inaccurately simulated quantities consists of the responses of the PID system. The efficiencies used to calculate the final result are measured based on efficiency maps. However, in order to find the optimal selection many different PID requirements are tested. Creating a new map for every regarded threshold is very inefficient. Therefore, also the classical resampling approach is employed. As discussed in section 4.6 both techniques rely on a binning scheme, which here is composed of ten bins in the track momentum, five bins in the track pseudo rapidity and five bins in the track multiplicity dimension. The bin borders are chosen in a way such that each bin is equally populated. This ensures reasonable statistical uncertainties on all bins of an efficiency map and a sufficient sample size to draw the PID response from in case of resampling.

For the following steps the blinded data are split into different regions all of which are detailed in Table 5.3. First the upper and lower mass sidebands consisting of data below and above the removed mass region are defined. These are assumed to contain only background and are further subdivided into near and far sidebands. The first pair is used to optimise the selection, the latter to determine the expected background yield in the signal area in order to calculate the expected upper limit on the branching fractions later on. This procedure avoids bias from optimising and measuring based on the same data. Furthermore, search regions are defined individually for the B_s^0 and B^0 meson, corresponding approximately to two standard deviations of the mass resolution around the respective mass. For the calculation of the final result only candidates inside these regions are considered.

After the data are sufficiently cleaned by the preselection and the simulation is ensured to be accurate, a multivariate classifier is trained in order to reduce the remaining combinatorial background while losing as few signal as possible. Implementations of decision trees with gradient boosting from the Scikit-Learn [89] and the TMVA [90] frameworks are compared to the MatrixNet (MN) [91], which is a proprietary algorithm developed by Yandex. As a metric on how well a given algorithm can distinguish between signal and background candidates the area under the Receiver Operating Characteristic (ROC AUC) is examined. Here, the MN performs slightly better than the two boosted decision trees and is therefore applied for this

measurement. The classifier is trained on 16000 background candidates from the near and the far sideband as defined in Table 5.3 and 100000 weighted simulated signal candidates of respectively $B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ and $B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays. A ten-fold cross validation is performed, which means that the training data are split into ten sub-samples of equal size and ten independent classifiers are trained. Each classifier is trained on merged data from nine samples and applied to the remaining one. This is done in every possible permutation, such that each candidate is predicted by an unbiased classifier. For predicting candidates that were not present in the training data one of the ten classifiers is chosen randomly. The MN is trained on the following properties of the $B_{(s)}^0$ -candidate:

- The decay time: Because of the large Lorentz boost B mesons are expected to decay at an average distance of 7 mm from the PV. Via the decay time signal candidates that are formed of tracks that for example originate from the PV can be rejected.
- The vertex quality: For signal candidates all four muons are produced in the same B decay, such that the combination of the four muon tracks results in a good quality vertex. Combinatorial background candidates however are formed of tracks that originate from different decays, such that the vertex quality is compromised.
- The momentum and transverse momentum: Due to the production mechanism of $b\bar{b}$ quark pairs in proton-proton collisions B mesons are created with a characteristic (transverse) momentum distribution, which is different from that resulting from a random track combination.
- The χ_{IP}^2 quantifies the impact of the $B_{(s)}^0$ candidate on the PV reconstruction. If the $B_{(s)}^0$ meson is not compatible with originating from the PV, respecting it in the fit will shift the reconstructed location significantly, resulting in a large χ_{IP}^2 .
- The cosine of the DIRA is a measure for the difference between the direction of the momentum of the reconstructed B meson and the path the B meson must have taken if it originates from the PV.

The features are explained in more detail in section 3.4. Only properties of the B candidate are chosen, such that the classifier can be applied to other B decays for validation purposes and to estimate a systematic uncertainty as discussed in section 5.5. The effect of the weighting algorithm on the distributions of the features the MN is based on are given in Figure 5.4. After applying the weights only small differences remain, which are accounted for by a systematic uncertainty discussed in section 5.5.

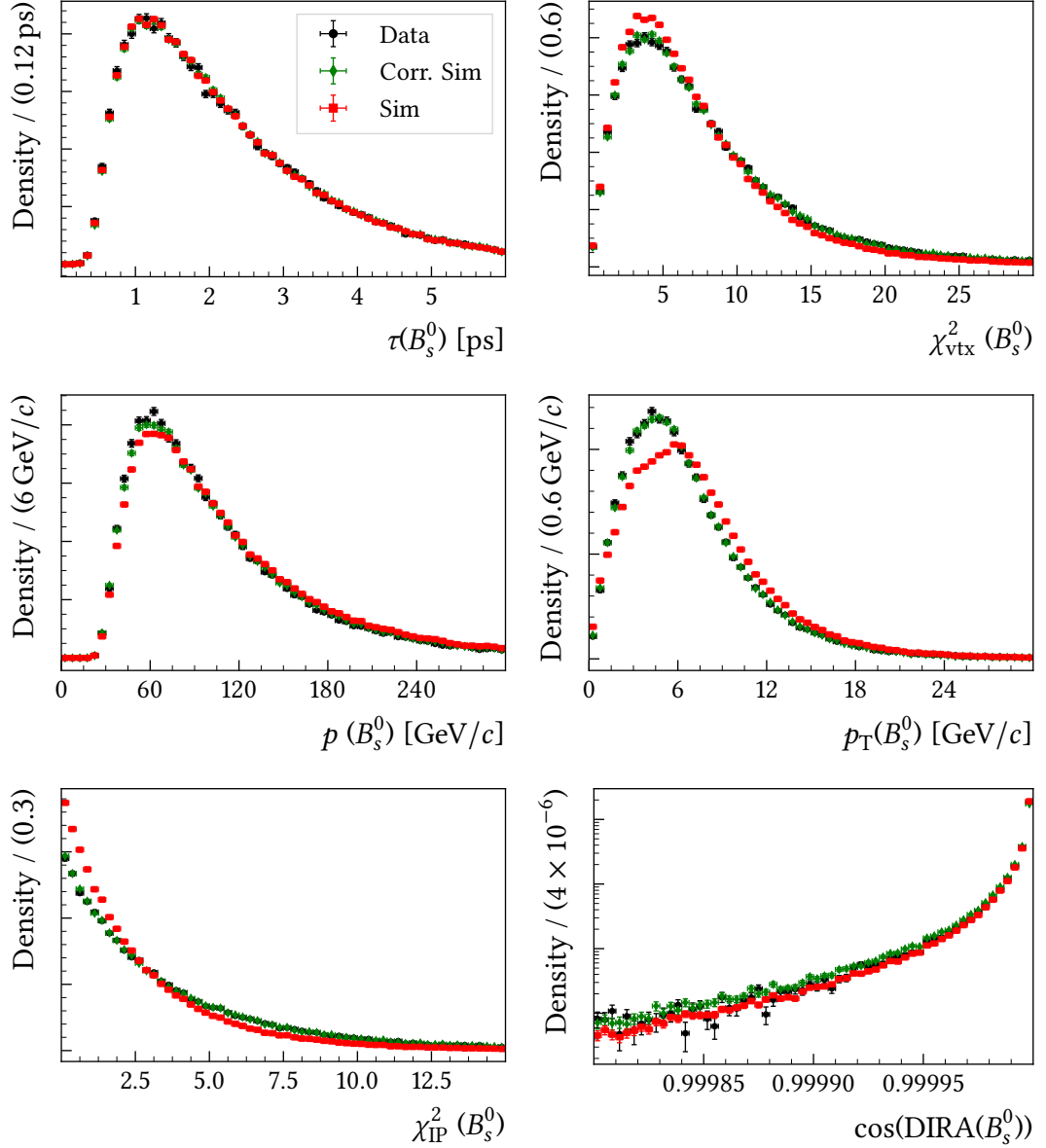


Figure 5.4: Distributions of features the MN is trained on: B meson lifetime, B meson vertex quality, B meson momentum and transverse momentum, B meson χ_{IP}^2 and the DIRA. Black circles represent measured $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ decays isolated from data by *sWeights*, red squares raw simulation and green diamonds corrected simulation.

The final step of the selection consists of a cut on particle identification variables and the MN response. These two cuts are correlated and therefore must be optimised simultaneously. As PID variable the difference in log-likelihood between the muon candidate being a muon or a pion ($DLL_{\mu\pi}$) is regarded. To ease the search for the optimal cut point, the minimum $DLL_{\mu\pi}$ between all four muon candidates is calculated, which reduces the dimensionality of the optimisation problem from five to two. The space spanned by the minimum muon PID and the MN response (optimisation space) is scanned for the maximum of the figure of merit [92]

$$\text{FoM} = \frac{\epsilon_{\text{signal}}}{\sigma/2 + \sqrt{N_{\text{bkg}}^{\text{expected}} \times \epsilon_{\text{bkg}}}}. \quad (5.1)$$

Here, the intended significance in terms of standard deviations, σ , is set to 3, but very similar selection criteria are found when using five. The expected amount of background candidates within the signal region before applying cuts on the MN response or the PID, $N_{\text{bkg}}^{\text{expected}}$, is determined by fitting an exponential function to the near sidebands and calculating the integral over the signal area. The invariant mass distribution of the four-muon system together with the fitted exponential function and the mentioned regions is presented in Figure 5.5. The expected background yield in the signal region is determined to be $N_{\text{bkg}}^{\text{expected}} = 1734 \pm 30$.

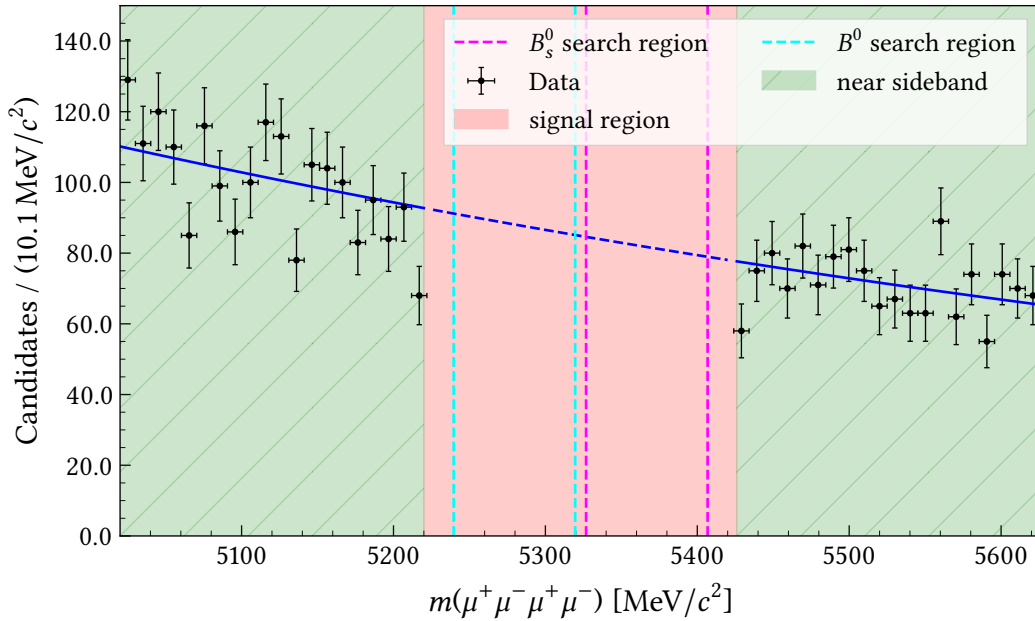


Figure 5.5: Distribution for the four-muon invariant mass before applying the final part of the selection. The data are parametrised by an exponential function. Additionally, various regions as defined in Table 5.3 are indicated.

Then a grid is spanned over the optimisation space and for each grid point, that is each unique combination of $\min(\text{DLL}_{\mu\pi})$ and MN response, the FoM is evaluated. Hereby the signal efficiency $\varepsilon_{\text{signal}}$ is measured based on corrected simulation and the background efficiency ε_{bkg} is determined on the near sidebands of the measured data. The result of the optimisation is presented in Figure 5.6, which shows the course of the FoM in the optimisation space. On the left the full range is shown and on the right a closer view to the region of highest FoM is illustrated. In this region only a small number of background candidates remains, while still a significant amount of simulated signal is present. Therefore, the FoM undergoes jumps at thresholds, where background candidates are situated while further removal of signal candidates causes a continuous alteration. Consequently, the plot shows square shaped artefacts in which the FoM is continuous. At the maximal FoM only one background candidate remains in the optimisation part of the data, such that the background is reduced by more than 99.9% with a signal efficiency of around 55%.

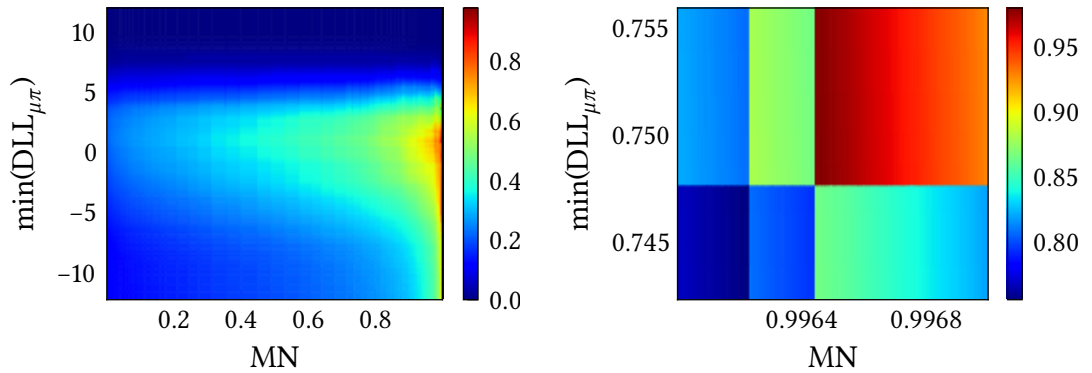


Figure 5.6: Figure of merit as a function of the optimisation space. On the left the full space is shown and on the right the region around the optimal selection.

5.3 Physical backgrounds

Several processes that can possibly fake the $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ signal and thereby bias the measured signal yield are studied. Contributions involving particle misidentification are quantified by measuring the frequency of misidentifying kaons and pions with kinematic similar to those of the muons originating from $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays. The misidentification rates are provided in form of PID performance maps as explained in section 4.6. The regarded kinematic distributions originate from

$B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ simulation that passes the full selection. The results are

$$p(2\pi \rightarrow 2\mu) = 8.24 \times 10^{-5}, \quad (5.2)$$

$$p(4\pi \rightarrow 4\mu) = 7.48 \times 10^{-9}, \quad (5.3)$$

$$p(2K \rightarrow 2\mu) = 3.49 \times 10^{-4}, \quad (5.4)$$

$$p(4K \rightarrow 4\mu) = 1.37 \times 10^{-7}. \quad (5.5)$$

$$(5.6)$$

Misidentified pions are stronger suppressed by the selection than kaons due to a requirement on $\text{DLL}_{K\pi}$. Decays that were explicitly tested regarding their possible contribution to fully selected $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ data are listed in Table 5.4. For

Table 5.4: Possible sources of peaking backgrounds, with their branching fraction and their effective branching fraction calculated by multiplying the misidentification probability. The considered decay modes for intermediate particles are $\mu^+ \mu^-$ for J/ψ and $\psi(2S)$, $K^+ \pi^-$ for K^{*0} , $\pi^+ \pi^-$ for ρ^0 and $K^+ K^-$ for the ϕ meson. Branching ratios are taken from [93] if not stated differently.

Decay mode	\mathcal{B}	$\mathcal{B} \times \text{misID rate}$	Expectation
$B_s^0 \rightarrow \pi^+ \pi^- \mu^+ \mu^-$	8.6×10^{-8} [94]	7.1×10^{-12}	0
$B^0 \rightarrow \pi^+ \pi^- \mu^+ \mu^-$	2.1×10^{-8} [94]	1.7×10^{-12}	0
$B^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$	$< 2.3 \times 10^{-4}$	$< 6.6 \times 10^{-12}$	0
$B^0 \rightarrow K^{*0} \mu^+ \mu^-$	1.1×10^{-6}	3.8×10^{-10}	0
$B_s^0 \rightarrow D_s^+ (\rightarrow K^+ K^- \pi^+) \pi^-$	1.6×10^{-4}	4.6×10^{-12}	0
$B_s^0 \rightarrow \phi \mu^+ \mu^-$	7.6×10^{-7}	2.7×10^{-10}	0 (ϕ veto)
$B^0 \rightarrow \psi(2S) K^*$	3.7×10^{-6}	1.3×10^{-9}	0 ($\psi(2S)$ veto)
$B^0 \rightarrow J/\psi K^{*0}$	5.3×10^{-5}	1.9×10^{-8}	0 (J/ψ veto)
$B_s^0 \rightarrow J/\psi K^{*0}$	3.2×10^{-6}	1.1×10^{-9}	0 (J/ψ veto)
$B_s^0 \rightarrow J/\psi \rho^0$	7.6×10^{-6}	6.3×10^{-10}	0 (J/ψ veto)

example $B_{(s)}^0 \rightarrow \pi^+ \pi^- \mu^+ \mu^-$ decays are probed by multiplying the reported misidentification rate $p(2\pi \rightarrow 2\mu)$ to the measured branching ratio. The result is of the order of 1×10^{-12} . Keeping in mind that the branching ratio of $B_s^0 \rightarrow \mu^+ \mu^-$, which is of the order of 1×10^{-9} , was at the edge of being measurable given the size of the LHCb Run 1 data sample, this process can be neglected. The same is valid for $B^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ decays. On the contrary the branching fraction of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay is of the order of 1×10^{-6} [38] and thereby sizeable enough to make further studies necessary. Therefore, the selection efficiency is measured on full $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ simulation using the same methods as for the signal efficiency. Just one out of 402 000 simulated decays passes the full selection, such that no contribution is expected. Full simulation is also studied for $B_s^0 \rightarrow D_s^+ (\rightarrow K^+ K^- \pi^+) \pi^-$ decays. All but two out of 800 000 simulated

candidates are found to be removed by only applying the `isMuon` requirement for all daughters. Hence, also this decay is neglected. Any decay proceeding via one or more of the intermediate resonances ϕ , J/ψ or $\psi(2S)$ is explicitly vetoed by applying corresponding cuts to the invariant mass of all four possible neutral combinations of two muons. The performance of this requirement is studied on 300 000 simulated $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ decays and the J/ψ (ϕ) veto alone is found to remove 99.99 % (100 %) of all candidates. In summary, all regarded sources of physical background were found to contribute only on a negligible level, if at all.

5.4 Efficiencies

The combined efficiency of all mentioned selection steps is measured on simulated data and amounts to $(0.580 \pm 0.003) \%$ and $(0.568 \pm 0.003) \%$ for the $B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ and the $B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decay, respectively. The sensitivity to the MSSM model described in section 2.2 is studied by measuring the efficiency on corresponding simulation. Both the P and S particles are simulated with a decay width of $\Gamma = 0.1 \text{ MeV}/c^2$, which corresponds to an immediate decay, such that the four muons in the final state are compatible with originating from a common vertex. The efficiency to select this kind of MSSM processes is found to be the same for B_s^0 and B^0 and amounts to $(0.648 \pm 0.003) \%$. The developed selection is specifically optimised to search for the SM decay. For example as a consequence of the MN being trained on features such as the vertex quality, the efficiency will drop for processes with intermediate particles whose lifetime is non-negligible. This can be seen when comparing the individual efficiencies stated in Tables 5.5 to 5.7. Here, Table 5.5 shows the efficiencies measured for the SM process, Table 5.6 for the decay via new (pseudo) scalar particles with negligible lifetime and Table 5.7 for the same beyond the standard model scenario with lifetimes of 5 ps for both new particles. The geometric acceptance is slightly higher for decays proceeding through intermediate resonance, especially the reconstruction, but also the MN efficiency is significantly lower though if the intermediate state has a non-negligible lifetime. The efficiencies for the BSM scenario with long-lived intermediate states are only presented for illustration. They could be significantly increased by developing a respective selection. This measurement focusses on the search for the SM process, though. Therefore, no explicit limit on the branching ratios of the BSM decays with long-lived intermediate states is calculated.

The simulated samples studied here are generated with fixed masses of S and P . To test the dependence of the final result on these masses, the selection efficiency is measured in bins of di-muon invariant mass. An efficiency variation of $\mathcal{O}(20 \%)$ is observed. Differences between the selection of the signal decays and the normalisation mode $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ lie in the PID cuts and in the fact that no multivariate selection is applied. The total efficiency for $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ is $(1.495 \pm 0.006) \%$. For all efficiencies the stated uncertainties originate from the limited sample size of the simulation and are treated as systematic uncertainties.

Table 5.5: The selection efficiencies in percent for the $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$ and $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decay. The values are measured step wise, such that the total efficiency is defined by their product.

Efficiency, %	$B^+ \rightarrow J/\psi K^+$	$B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$	$B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$
Geometric acceptance	15.780 ± 0.062	15.290 ± 0.030	15.247 ± 0.042
Reco and Stripping	18.834 ± 0.022	13.451 ± 0.027	13.716 ± 0.028
Preselection	96.813 ± 0.026	97.287 ± 0.034	97.347 ± 0.034
Dimuon mass vetoes	-	65.069 ± 0.105	64.389 ± 0.105
Trigger	84.987 ± 0.047	89.744 ± 0.083	90.351 ± 0.081
Muon DLL $_{\mu\pi}$	93.130 ± 0.001	83.734 ± 0.023	83.905 ± 0.025
Muon DLL $_{K\pi}$	76.024 ± 0.000	-	-
Kaon DLL $_{K\pi}$	86.309 ± 0.013	-	-
MN	-	65.974 ± 0.216	66.353 ± 0.216
40 MeV/ c^2 mass window	-	87.983 ± 0.183	87.974 ± 0.183
Total	1.4950 ± 0.0062	0.5675 ± 0.0029	0.5800 ± 0.0032

Table 5.6: The selection efficiencies in percent for the $B_{(s)}^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)$ decays, where S and P decay instantaneously.

Efficiency, %	$B_s^0 \rightarrow SP$	$B^0 \rightarrow SP$
Geometric acceptance	16.940 ± 0.030	17.040 ± 0.030
Reco and Stripping	13.720 ± 0.024	13.620 ± 0.024
Preselection	96.960 ± 0.037	96.790 ± 0.038
Dimuon mass vetoes	65.350 ± 0.105	65.700 ± 0.105
Trigger	90.150 ± 0.082	89.750 ± 0.083
Muon DLL $_{\mu\pi}$	83.840 ± 0.017	83.690 ± 0.021
MN	66.000 ± 0.153	66.260 ± 0.151
40 MeV/ c^2 mass window	88.200 ± 0.128	88.210 ± 0.127
Total	0.650 ± 0.003	0.650 ± 0.003

Table 5.7: The selection efficiencies in percent for the $B_{(s)}^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)$ decays where S and P have a lifetime of 5 ps.

Efficiency, %	$B_s^0 \rightarrow SP$	$B^0 \rightarrow SP$
Geometric acceptance	16.810 ± 0.016	16.890 ± 0.016
Reco and Stripping	4.790 ± 0.016	4.400 ± 0.015
Preselection	97.360 ± 0.076	96.970 ± 0.083
Dimuon mass vetoes	86.160 ± 0.163	84.350 ± 0.176
Trigger	91.600 ± 0.139	91.650 ± 0.144
Muon $DLL_{\mu\pi}$	83.600 ± 0.034	83.450 ± 0.036
MN	42.190 ± 0.272	44.400 ± 0.278
40 MeV/ c^2 mass window	86.980 ± 0.285	87.250 ± 0.280
Total	0.190 ± 0.002	0.180 ± 0.002

5.5 Systematic uncertainties and cross checks

In general two categories of systematic uncertainties arise. One is due to external inputs that come with an uncertainty. The second category is made up by differences between simulated and measured data and is accessed by comparing values determined on both data sources.

In the following all studied uncertainties are briefly described in individual paragraphs. Their effects are summarised in Table 5.8. At the end of this section cross checks are presented that validate the analysis strategy and show that all measured quantities yield consistent results within the assigned uncertainties.

External inputs

The most dominant source of systematic uncertainty, entering only the limit on the branching ratio of the B_s^0 decay though, originates from the ratio of hadronisation probabilities

$$\frac{f_s}{f_d} = 0.259 \pm 0.015 \quad (5.7)$$

that has been measured by LHCb [95]. At the time the measurement described in this chapter was published a dependence of this quantity on the B transverse momentum was observed with a significance corresponding to three standard deviations. In a recent update [96] the significance increased to 3.6 standard deviations. However, for the analysis described in this thesis this dependency does not need to be incorporated as the quoted value represents the average $\frac{f_s}{f_d}$ calculated based on B decays measured at LHCb, and is hence assumed to be precise for the analysed data.

There are two further input parameters for the final result that are taken from external sources and come with an uncertainty, which are the branching fractions of the normalisation decay. For both the world averages [38],

$$\mathcal{B}(B^+ \rightarrow J/\psi K^+) = (1.027 \pm 0.031) \times 10^{-3} \quad \text{and} \quad (5.8)$$

$$\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-) = (5.961 \pm 0.033) \times 10^{-2}, \quad (5.9)$$

are taken.

Particle identification efficiency

One example of imperfectly modelled distributions are the responses of the PID system. In order to be still able to extract accurate efficiencies of the applied particle identification criteria, performance histograms are employed as explained in section 4.6. Different binning schemes are tested and the observed deviation of the resulting efficiency, which amounts to 1% for $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ and 0.5% for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$, is treated as systematic uncertainty. For muons the calibration

samples are based on $J/\psi \rightarrow \mu^+ \mu^-$ decays, where the muons are required to have a transverse momentum of larger than 800 MeV/c. Unfortunately around 50 % of all simulated $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays contain at least one muon that falls below this threshold. The solution to this pitfall is to expand the calibration data to muons from $D_s^+ \rightarrow \phi(\rightarrow \mu^+ \mu^-) \pi^+$ decays. This decay offers a very clean signature such that no PID criteria need to be applied in order to extract its signal. The muons from $D_s^+ \rightarrow \phi(\rightarrow \mu^+ \mu^-) \pi^+$ are required to have a transverse momentum of larger than 300 MeV/c at stripping level, such that 3 % of all simulated $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays still contain at least one muon that lies in a phase space region that is not covered. Therefore, a systematic uncertainty of 3 % is assigned.

Data simulation agreement

Deviations between simulated and real data in distributions other than the PID response are corrected by a weighting algorithm based on boosted decision trees as discussed in section 5.2. The algorithm is trained on measured and simulated data of $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \phi(\rightarrow K^+ K^-)$ decays. The trained reweighter is then applied to every simulated data sample used for this measurement. To account for differences between $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \phi(\rightarrow K^+ K^-)$ and $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays, that can compromise the correctness of this adaptation, the MN efficiency is compared between the three modes and the relative difference of 3.6 % is assigned as systematic uncertainty on the reweighting technique for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ simulation. One major difference between $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \phi(\rightarrow K^+ K^-)$ decays originates from the number of daughter particles. This difference is reflected also in the features that are subject to the reweighting. Therefore, the difference between the $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+$ selection efficiencies calculated with and without weights, which amounts to 2.3 %, is assigned as systematic uncertainty.

Table 5.8: Summary of systematic uncertainties affecting the single event sensitivities along with the total systematic uncertainty calculated by adding up the individual components in quadrature. The dominating uncertainty arising from f_u/f_s only contributes to in case of the B_s^0 decay. The uncertainty of the stated selection efficiencies arising from the limited number of simulated events is 0.5 % for $B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ and 0.4 % for all other considered decay modes.

Source	Value [%]
f_u/f_s	5.8
$\mathcal{B}(B^+ \rightarrow J/\psi K^+)$	3.0
$\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$	0.1
PID data phase space coverage	3.0
PID binning $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$	1.0
PID binning $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$	0.5
Weighting $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$	3.6
Weighting $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$	2.3
MN efficiency	0.3
Trigger efficiency	3.0
Mass resolution	0.5
Track finding efficiency	1.7
Photon radiation $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$	0.9
Photon radiation $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$	0.5
Track mismatching	0.6
Size of simulated samples	0.4 – 0.5
Normalisation fit	0.3
Combined B_s^0 SM	9.2
Combined B^0 SM	7.2
Combined B_s^0 MSSM	9.2
Combined B^0 MSSM	7.2

Multivariate Classifier efficiency

To quantify remaining discrepancies between measured and simulated data that are not corrected for by applying the weights, the MN efficiency for $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ decays is regarded. On simulated data the efficiency is given by the ratio of the sum of correction weights before and after applying the requirement on the MN response. On measured data the \mathcal{P} lot method (see section 4.4) is employed to produce weights, which allow for a signal efficiency measurement without background contribution. The resulting efficiencies are $(70.2 \pm 0.2) \%$ and $(70.4 \pm 0.2) \%$ and the relative difference of 0.3 % is assigned as systematic uncertainty. The performance of the MN is expected to be similar between the different $B_{(s)}^0$ decays, because it is trained only on properties of the $B_{(s)}^0$ meson, that are similar if not equal for different final states. The efficiencies mentioned here are significantly different from those stated in Table 5.5, as the selection that is applied previous to measuring these values is different from that defined in section 5.2. More precisely the selection that is applied to $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ candidates at stripping level is also applied to $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ candidates to avoid artificial differences.

Trigger efficiency

The trigger efficiency measured on simulation by requiring the trigger criteria is compared to values obtained from applying the TisTos method (see section 4.7) to simulation and data. This technique requires a large, clean sample, which is found in $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ data. The remaining background is suppressed by applying the \mathcal{P} lot method (see section 4.4). Hereby a single fit to a merged sample of TIS, TOS and TISTOS candidates is performed instead of individual fits to all three categories. To validate this decision, the mass distributions of these three categories are compared in Figure 5.7 and good agreement is observed.

The TisTos method is applied in bins of the B meson phase space. Here, seven bins in p_T and five bins in p_z are chosen, which is a compromise between good sensitivity to the shape of the distributions and a sufficient population of each bin. The measured efficiencies extracted from simulation by applying the trigger criteria ($\epsilon_{\text{Cut}}^{\text{MC}}$) and through the TisTos method on simulated and real data ($\epsilon_{\text{TisTos}}^{\text{MC}}$ and $\epsilon_{\text{TisTos}}^{\text{Data}}$) are

$$\epsilon_{\text{Cut}}^{\text{MC}} = (84.99 \pm 0.05) \%, \quad (5.10)$$

$$\epsilon_{\text{TisTos}}^{\text{MC}} = (84.7 \pm 0.7) \%, \quad (5.11)$$

$$\epsilon_{\text{TisTos}}^{\text{Data}} = (82.7 \pm 0.8) \%. \quad (5.12)$$

The efficiency measured by requiring the trigger criteria on simulation is the base line

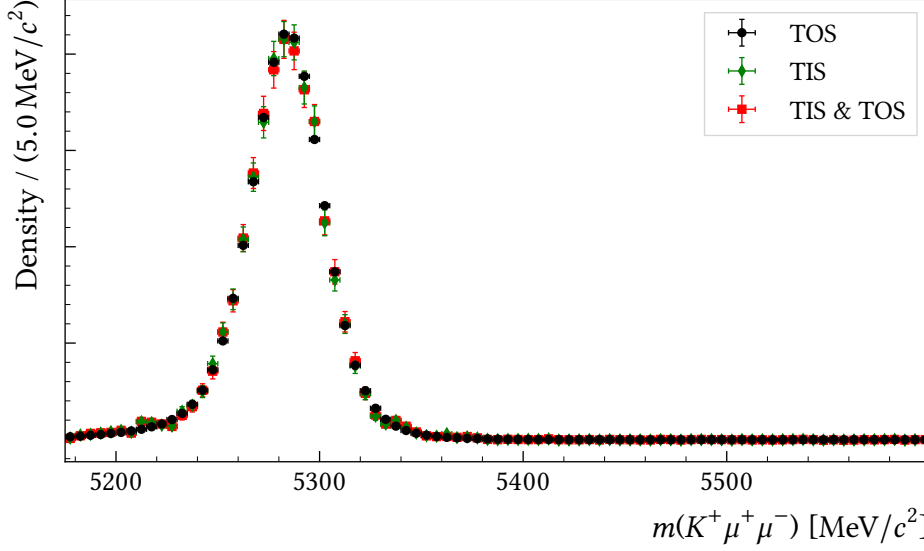


Figure 5.7: Distribution of the invariant $K^+ \mu^+ \mu^-$ mass of $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ decays that fall into the TIS, TOS and TISTOS category. The histograms are weighted by the $sWeights$ obtained from a single fit to a merged sample of all three categories.

that is used to compute the final result. The systematic uncertainty is given by

$$\Delta_{\text{Trigger}}^{\text{Syst.}} = \frac{\sqrt{(\epsilon_{\text{TisTos}}^{\text{Data}} - \epsilon_{\text{Cut}}^{\text{MC}})^2 + (\epsilon_{\text{TisTos}}^{\text{MC}} - \epsilon_{\text{Cut}}^{\text{MC}})^2 + \Delta_{\text{stat.}}(\epsilon_{\text{TisTos}}^{\text{Data}})^2 + \Delta_{\text{stat.}}(\epsilon_{\text{TisTos}}^{\text{MC}})^2}}{\epsilon_{\text{Cut}}^{\text{MC}}} \quad (5.13)$$

$$= 3 \%$$

The first term accounts for the difference between the base line trigger efficiency and the value measured on data, the second for the difference between two methods used to measure the efficiency, and the third and fourth for the statistical uncertainties on the TisTos efficiencies measured on data and simulation ($\Delta_{\text{stat.}}(\epsilon_{\text{TisTos}}^{\text{Data}})$ and $\Delta_{\text{stat.}}(\epsilon_{\text{TisTos}}^{\text{MC}})$).

All regarded trigger lines are based on (di)-muon signatures, of which the $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays contain twice as many as $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$. For this reason the systematic is expected to cancel partially in the ratio of efficiencies between $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$, such that 3% is assigned as systematic uncertainty for the final result.

Mass resolution

For the final result, every candidate that falls inside a region of $\pm 40 \text{ MeV}/c^2$ around the known $B_{(s)}^0$ mass is assumed to be signal. This interval corresponds approximately to two standard deviations of the mass resolution. In simulated data that passed all selection steps, this region contains 93.276% of all candidates. For measured $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi (\rightarrow K^+ K^-)$ data the ${}_s\mathcal{P}$ lot technique (see section 4.4) is applied

in order to avoid any bias due to the presence of background candidates. The fraction of the sum of $sWeights$ within $\pm 40 \text{ MeV}/c^2$ around the B_s^0 mass is 92.831 %. Both, simulated and measured data samples are so large, that the statistical uncertainties on both numbers are below one-tenth of a percent and are therefore omitted. The relative difference between the two values of 0.5 % is treated as systematic uncertainty.

Track finding efficiency

The track finding efficiency cannot be extracted from simulation with a sufficient precision. Especially due to the difference of the number of tracks in the sought-after and the normalisation decay the effect is not expected to cancel in the normalisation. The LHCb tracking group provides histograms containing the ratio of the track finding efficiencies on data and simulation in bins of momentum and pseudorapidity [97]. Taking the average of the corresponding value from the histogram for the $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$ and the $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ simulated data samples, and then calculating the ratio between these yields

$$\frac{\frac{\epsilon_{3\text{-tracks}}^{\text{MC}}}{\epsilon_{3\text{-tracks}}^{\text{Data}}}}{\frac{\epsilon_{4\text{-tracks}}^{\text{MC}}}{\epsilon_{4\text{-tracks}}^{\text{Data}}}} = 0.978 \pm 0.011. \quad (5.14)$$

The nominal value is treated as a correction factor to the efficiency, while the uncertainty is assigned as systematic uncertainty. The method that is applied to create the efficiency map is unreliable for low momentum tracks. Therefore, the map starts at a minimal momentum of $5 \text{ GeV}/c$, while in simulated data the momentum distributions of the tracks start at $3 \text{ GeV}/c$. Following an advice of the LHCb tracking group, tracks with momenta less than $5 \text{ GeV}/c$, are treated as having $5 \text{ GeV}/c$ and an additional 5 % systematic uncertainty is added for the individual track. This results in a per candidate systematic uncertainty of $i \times 5 \%$ where i is the number of tracks with momentum less than $5 \text{ GeV}/c$ in the corresponding candidate. The total contribution is calculated by averaging the per candidate value over the simulated data sample and is already included in the uncertainty stated in Equation (5.14). An additional 1.3 % of systematic uncertainty is assigned due to material interactions. The value corresponds to the average recommended systematic uncertainty stated in [97]. The uncertainties are added in quadrature and result in a total systematic uncertainty of 1.7 %.

Reconstruction effects

The reconstruction and stripping efficiency is measured from the ratio of generated events and correctly reconstructed candidates in the simulated data sample after applying the stripping criteria. A systematic uncertainty arises from the definition of *correctly reconstructed*. The LHCb software arranges simulated candidate according to different categories regarding their reconstruction. One category is made up by

correctly reconstructed *signal*. Another is referred to as *quasi signal*, which occurs for example if a decay via an intermediate resonance is reconstructed although the final state particles have been produced directly. The *low mass background* category is assigned if a decay product of the simulated B meson is not used to form the B candidate. For example a photon can be produced, but not reconstructed. The *fake* category is assigned if one or more tracks that are used to construct the candidate are considered to be *fake* tracks, referred to as ghosts. During the track reconstruction, hits are collected and combined into tracklets, that are at a later stage used to build tracks [98]. Hits that were caused by different particles can accidentally be matched to the same tracklet. If 30 % of all hits of a track are wrongly matched it is considered a ghost and any candidate build from that track is assigned *fake* category.

Depending on the regarded final state, 0.9 % or 0.5 % of the reconstructed candidates in signal simulation are assigned the *fake* category and 0.6 % the *low mass background* category. More thorough studies could be carried out in order to implement a correct treatment of these misclassified signal candidates. Given the small number of affected candidates the impact of said studies on the final result would be negligible. Therefore, these fractions are assigned as systematic uncertainty.

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

As a cross check regarding the treatment of the normalisation mode, the extracted number of signal candidates is compared to the expectation. The determination of the the first mentioned quantity is discussed in section 5.6. The latter is calculated based on equation (4.2). The different centre-of-mass energies of the collisions the LHC delivered in 2011 and 2012 result in different $b\bar{b}$ production cross-sections for the two years, such that equation (4.2) becomes

$$\begin{aligned}
 N^{\text{expected}} &= \{ \mathcal{L}_{\text{int}}(7 \text{ TeV}) \times \sigma_{b\bar{b}}(7 \text{ TeV}) + \mathcal{L}_{\text{int}}(8 \text{ TeV}) \times \sigma_{b\bar{b}}(8 \text{ TeV}) \} \times 2 \times f_u \quad (5.15) \\
 &\quad \times \mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-) \times \varepsilon_{B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+} \\
 &= 635\,000 \pm 67\,000.
 \end{aligned}$$

All quantities needed for the calculation are summarised in Table 5.9. The expected and measured numbers of signal candidates agree within their uncertainties, which are, especially due to $\sigma_{b\bar{b}}$, large. Nevertheless, this cross check shows that the measured normalisation yield and the efficiency for $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ are consistent with previously measured results.

Table 5.9: The expected $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$ signal yield and the numerical input for the computation.

Quantity	Value
$\sigma_{b\bar{b}}(7 \text{ TeV})$	$(288 \pm 48) \text{ mb}$ [79]
$\sigma_{b\bar{b}}(8 \text{ TeV})$	$(298 \pm 36) \text{ mb}$ [80]
f_u	$(40.1 \pm 0.8) \%$ [93]
$\mathcal{L}_{\text{int}}(7 \text{ TeV})$	$(0.9858 \pm 0.0017) \text{ fb}^{-1}$
$\mathcal{L}_{\text{int}}(8 \text{ TeV})$	$(1.9897 \pm 0.0017) \text{ fb}^{-1}$
$\mathcal{B}(B^+ \rightarrow J/\psi K^+)$	$(1.027 \pm 0.031) \times 10^{-3}$ [99]
$\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$	$(5.9610 \pm 0.0033) \times 10^{-2}$ [99]
Selection efficiency	$(1.4758 \pm 0.0081) \%$

$B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ branching fraction

A more profound cross check is performed by measuring the ratio of branching fractions of the $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$ and the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ decays

$$\begin{aligned}
 R &= \frac{\mathcal{B}(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-))}{\mathcal{B}(B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+)} & (5.16) \\
 &= \frac{\epsilon_{B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+}}{\epsilon_{B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)}} \times \frac{N_{B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)}}{N_{B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+}} \times \frac{f_u}{f_s}. & (5.17)
 \end{aligned}$$

The $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ decay itself proceeds via a FCNC and thus has a low branching fraction. Observing a significant signal of this decay validates that the developed selection is capable of rejecting enough background while keeping sufficient signal to measure rare decays. Comparing the measured value to the precisely known world average validates the accuracy of the measured efficiencies and thereby the measures taken to correct for imperfectly simulated features. In order to extract the corresponding signal the selection concerning the regarded intervals in the different di-muon mass distributions has to be altered. In contrast to the non-resonant decays, here one pair of oppositely charged muons is required to be situated inside the ϕ interval, while the other must fall into the J/ψ window. The rest of the selection is kept unchanged. A breakdown of the underlying efficiencies together with the total efficiency, calculated by their product is given in Table 5.10. Multiplying the world averages of the measured branching ratios of the underlying decays summarised in Table 5.11 gives

$$\mathcal{B}(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)) = (1.83 \pm 0.18) \times 10^{-8}. \quad (5.18)$$

Table 5.10: Efficiency breakdown for $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ candidates.

Selection	Efficiency [%]
Geometric acceptance	15.150 ± 0.040
Reco and Stripping	13.085 ± 0.044
Preselection	97.334 ± 0.066
Dimuon mass vetoes	94.609 ± 0.082
Trigger	92.180 ± 0.100
Muon DLL $_{\mu\pi}$	84.192 ± 0.025
MN	68.590 ± 0.210
Total	0.971829 ± 0.000050

 Table 5.11: World averages of the underlying branching fractions of the $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ decays.

Branching Fraction	World Average
$\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)$	$(1.07 \pm 0.09) \times 10^{-3}$ [99]
$\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$	$(5.961 \pm 0.033) \times 10^{-2}$ [99]
$\mathcal{B}(\phi \rightarrow \mu^+ \mu^-)$	$(2.87 \pm 0.19) \times 10^{-4}$ [99]
$\mathcal{B}(B^+ \rightarrow J/\psi K^+)$	$(1.027 \pm 0.031) \times 10^{-3}$ [99]

An extended unbinned maximum likelihood fit to the four-muon invariant mass distribution results in a signal yield of $N(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)) = 38.9 \pm 6.7$. The modelled fully selected data are shown in Figure 5.8. In the distribution a clean signal peak can be seen. The amount of combinatorial background that passes the full selection is almost negligible. The fit model consists of a Gaussian function to describe the signal part and an exponential function for combinatorial background. The efficiency to detect, reconstruct and select $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ decays is measured on simulated data to be

$$\varepsilon(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)) = (0.6006 \pm 0.0028) \%, \quad (5.19)$$

using the same correction techniques as developed for the $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ simulation. The measured ratio and that calculated from the world averages are

$$R_{\text{measured}} = (3.54 \pm 0.70) \times 10^{-4} \quad \text{and} \quad (5.20)$$

$$R_{\text{world avg}} = (3.00 \pm 0.34) \times 10^{-4}. \quad (5.21)$$

For this decay no studies regarding systematic uncertainties are carried out. The stated values show agreement within their statistical uncertainties and thereby validate the whole analysis chain.

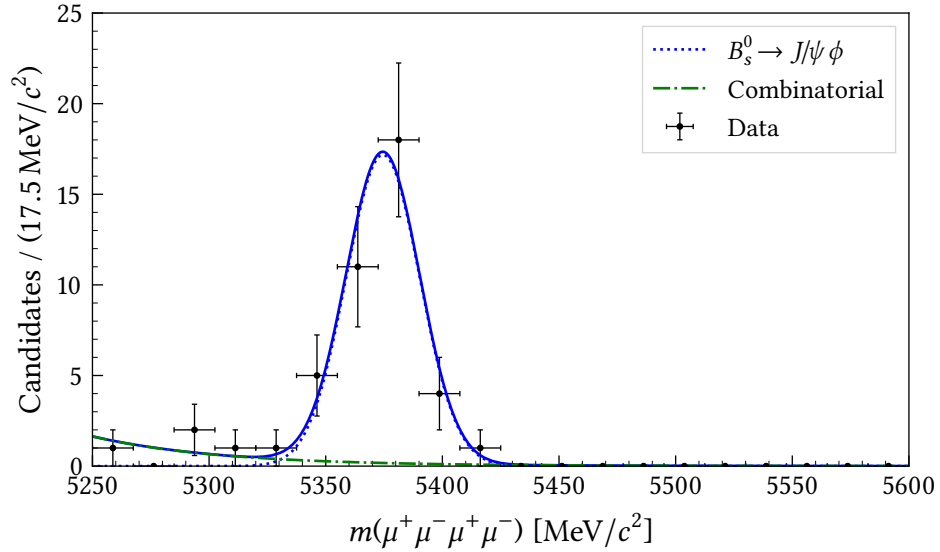


Figure 5.8: Four-muon invariant mass distribution in data passing the full selection. Here, the requirements on the individual di-muon invariant masses are adjusted in order to select $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)\phi(\rightarrow \mu^+ \mu^-)$ decays.

5.6 Fits and limit setting

The amount of signal candidates in the normalisation mode $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$ is extracted by performing an unbinned extended maximum likelihood fit to the reconstructed B^+ mass distribution. The signal component is modelled by a Hypatia function [81], which is detailed in section 4.2. Combinatorial background is described by an exponential function and partially reconstructed background by a Gaussian function that at a given point transitions into an exponential function. This model is chosen because partially reconstructed background that falls into the regarded B^+ mass range dominantly originates from $B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^{*0}(\rightarrow K^+ \pi^-)$ decays, where the K^{*0} decays further into a kaon and a pion and the pion is not reconstructed. This leads to reconstructed B^+ mass of $m_{B^0} - m_{\pi^+} \approx 5140 \text{ MeV}/c^2$ and produces a large tail towards lower masses caused by the momentum of the missed π^+ meson. The mass distribution together with the fitted functions is illustrated in Figure 5.9. The measured yield of normalisation mode signal candidates that pass the full selection is

$$N_{B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+} = 687\,890 \pm 920. \quad (5.22)$$

A systematic uncertainty due to the parametrisation of the signal invariant mass distribution is determined by probing different models, fit ranges and fit methods. Results from unbinned and binned maximum likelihood fits are compared. Additionally, the impact of setting the lower edge of the fit range to the transition point of the model for partially reconstructed background and thereby replacing the model with a Gaussian function is tested. A variation of the signal yield of 0.3 % is observed and assigned

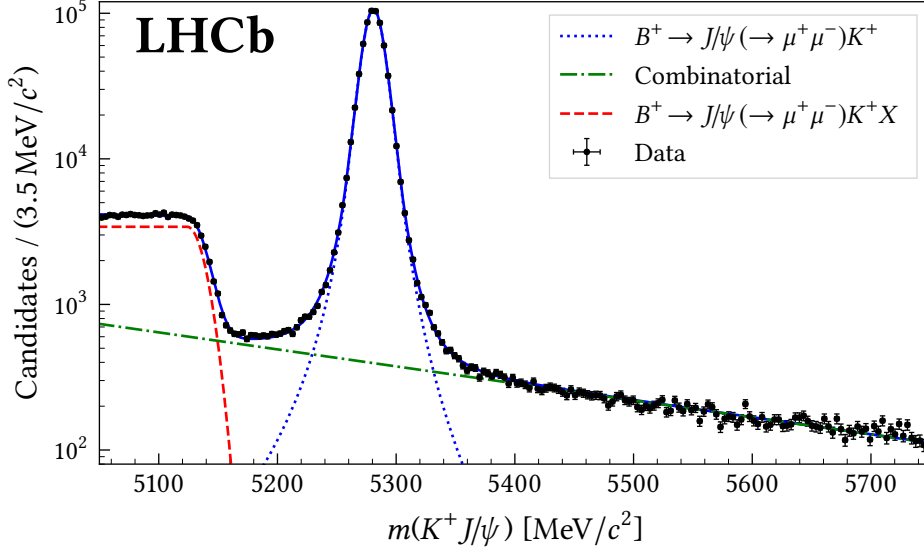


Figure 5.9: Invariant mass distribution of the $K^+ \mu^+ \mu^-$ system, where the di-muon invariant mass is constrained to that of the J/ψ meson. The distribution is modelled by a PDF as described in the text.

as systematic uncertainty. The normalisation constants as defined in equation (4.5) in section 4.1 are calculated to be

$$\eta_s^{\text{SM}} = (8.65 \pm 0.80) \times 10^{-10}, \quad (5.23)$$

$$\eta_d^{\text{SM}} = (2.29 \pm 0.16) \times 10^{-10}, \quad (5.24)$$

$$\eta_s^{\text{MSSM}} = (7.75 \pm 0.72) \times 10^{-10}, \quad (5.25)$$

$$\eta_d^{\text{MSSM}} = (2.01 \pm 0.14) \times 10^{-10} \quad (5.26)$$

for the B_s^0 and the B^0 decay modes in the SM and the MSSM scenarios. The reported uncertainties are the sum in quadrature over the respective statistical and systematic uncertainty. Here, the ratio of the hadronisation probabilities f_u/f_d is assumed to be unity. As a consequence f_s/f_u is equal to f_s/f_d , such that a previously measured value (see Equation (5.7)) can be used for computing the normalisation constants. This leads to a significantly larger uncertainty on η_s compared to η_d . The branching fractions predicted by the SM are at least three orders of magnitude below the current sensitivity of the LHCb experiment. Therefore, it is assumed that only upper limits on the branching fractions can be set. Before investigating the blinded signal region, the expected amount of background candidates for both search regions are estimated by parametrising the far sidebands (see Table 5.3) and integrating the model over the B^0 and B_s^0 search intervals. Two models, namely an exponential function and a second order polynomial, are probed, where the exponential is used as baseline model. The difference between the expected background yields is assigned as systematic

uncertainty. The found expected background yields are

$$N_{\text{bkg}}^{\text{expected}}(B^0) = 0.55_{-0.19}^{+0.24} (\text{stat}) \pm 0.20 (\text{syst}) \quad (5.27)$$

$$N_{\text{bkg}}^{\text{expected}}(B_s^0) = 0.47_{-0.18}^{+0.23} (\text{stat}) \pm 0.18 (\text{syst}) \quad (5.28)$$

in the B^0 and the B_s^0 search region. The Hybrid- CL_s method [100–102] is employed to calculate upper limits. For this method a range of assumed branching fractions is scanned and for each the agreement with the expected or observed data, depending on whether determining the expected or observed limit is calculated. This is done by computing the CL_s , which is the ratio of p -values of the hypotheses that the data contains background and signal C_{s+b} and that only background is present CL_b . Hereby, a counting experiment is assumed. Hence, the p -values are calculated assuming a Poissonian distribution for the respective amount of candidates. In order to account for the uncertainties of the expected background yields and the normalisation constants, the CL_s of each regarded branching fractions is calculated multiple times fluctuating the mentioned quantities by sampling them from a prior and taking the average. As priors log-normal and gamma distributions were tested with compatible outcome.

The algorithm to calculate one individual CL_s value for a given branching fraction \mathcal{B} is as follows:

1. Start with the expected background yield (see equations (5.27) and (5.28)) and the expected signal yield based on the assumed branching fraction \mathcal{B} and the normalisation constant η

$$S' = \frac{\mathcal{B}}{\eta}, \quad (5.29)$$

$$B' = N_{\text{bkg}}^{\text{expected}}. \quad (5.30)$$

2. Sample B and S from a prior that accounts for the respective uncertainty on S' and B' .
3. Calculate C_{s+b} and CL_b as

$$CL_{s+b} = \sum_{n=0}^{n_{\text{obs}}} \mathcal{P}_{S+B}(n), \quad (5.31)$$

$$CL_b = \sum_{n=0}^{n_{\text{obs}}} \mathcal{P}_B(n), \quad (5.32)$$

$$CL_s = \frac{CL_{s+b}}{CL_b}, \quad (5.33)$$

where $\mathcal{P}_\lambda(k)$ denotes a Poisson distribution with expectation value λ evaluated for k observations. The observed number of candidates n_{obs} is hereby sampled from a Poissonian distribution with expectation value set to B when determining the expected limit.

4. Repeat steps 2 and 3 many times and calculate the average CL_s

The upper limit is defined as the branching fraction where

$$CL_s \leq 1 - \alpha \quad (5.34)$$

with α being the desired confidence level (CL), which for this measurement is 95 %.

The expected upper limits are the best comparison to the previously performed measurement [88] as they are less prone to fluctuations in the search regions. The expected upper limits at 95 % CL are

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) < 6.9 \times 10^{-10}, \quad (5.35)$$

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) < 2.6 \times 10^{-9}, \quad (5.36)$$

$$\mathcal{B}(B^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) < 6.0 \times 10^{-10}, \quad (5.37)$$

$$\mathcal{B}(B_s^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) < 2.3 \times 10^{-9}, \quad (5.38)$$

for B^0 and B_s^0 in the SM and the MSSM case. In comparison to the measurement based on 1 fb^{-1} of integrated luminosity [41] these values are smaller by a factor of ten for the B^0 limits and a factor of six for the B_s^0 limits.

5.7 Results and conclusion

Unblinding the signal region revealed no additional B candidates, in either the B^0 or the B_s^0 search region, such that the observed signal yield is zero for both. The four-muon invariant mass distribution of the complete data sample passing the full selection is presented in Figure 5.10. Overall sixteen candidates pass the full selection of which fifteen lie in the far sidebands and one in the near sidebands. Of the sixteen surviving candidates fourteen are situated at lower and two at higher four-muon invariant mass with regard to the signal region. This is due to the exponentially distributed combinatorial background. Regarding the invariant masses of all possible neutral di-muon combinations in Figure 5.11 no pileup of candidates at a specific invariant mass is observed. Hence, no hint of a particle, SM like or BSM like, decaying into a pair of muons is found.

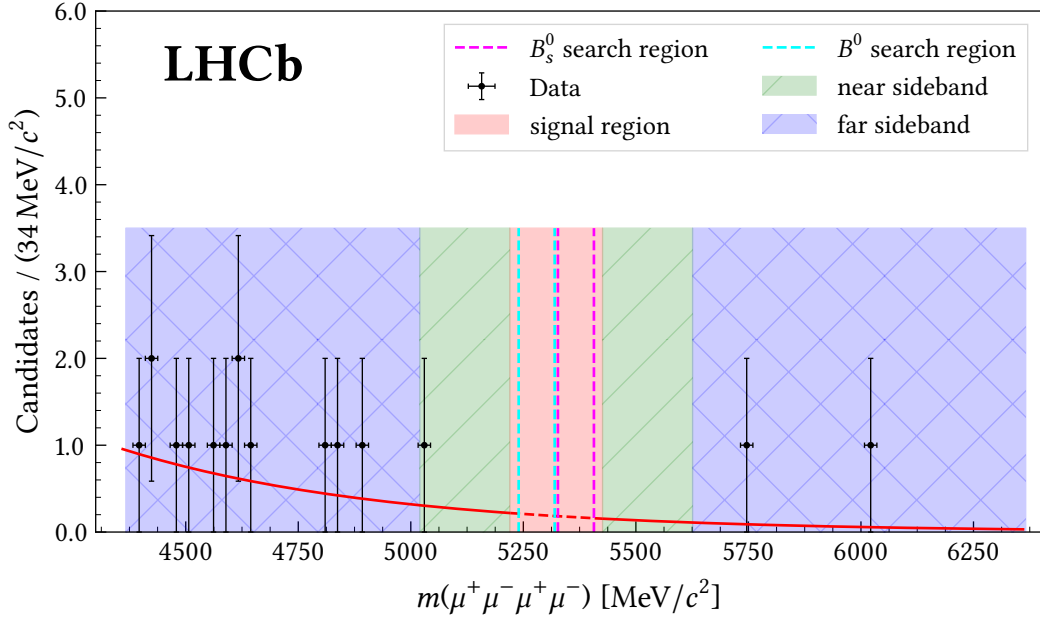


Figure 5.10: Distribution for the four-muon invariant mass after applying the full selection. The data are parametrised by an exponential function. Additionally, various regions as defined in Table 5.3 are indicated.

As the expected invariant mass distributions of the two sought-after decays slightly overlap, the limits for the B^0 and B_s^0 mode are anti-correlated due to cross-feed between the two search regions. This is accounted for in the limit calculation as presented in Figure 5.12. As a consequence of observing zero signal candidates in both search regions, the observed limits lie at the lower error bound of the expected limits, which are defined by the constraint that it is not possible to observe fewer than zero candidates. Neglecting contributions from the B^0 to the B_s^0 decay and vice versa the observed upper limits at 95 % CL are

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) < 6.9 \times 10^{-10}, \quad (5.39)$$

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) < 2.5 \times 10^{-9}, \quad (5.40)$$

$$\mathcal{B}(B^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) < 6.0 \times 10^{-10}, \quad (5.41)$$

$$\mathcal{B}(B_s^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) < 2.2 \times 10^{-9}. \quad (5.42)$$

Previous to this measurement the only experimental study of $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays was carried out by LHCb based on 1 fb^{-1} of integrated luminosity recorded in 2011. These previous upper limits have been reduced by a factor of 6.4 (7.3) for the SM (MSSM) mode in the case of the B_s^0 decay and by a factor of 9.5 (10.5) for the B^0 decay. An improvement of a factor of $\sqrt{3} \approx 1.7$ is expected due to the increased sample size when assuming a Poissonian distribution. The main improvements however originate from a machine learning based selection and a more optimal choice of the normalisation mode. In the previous measurement, $B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^{*0}(\rightarrow K^+ \pi^-)$

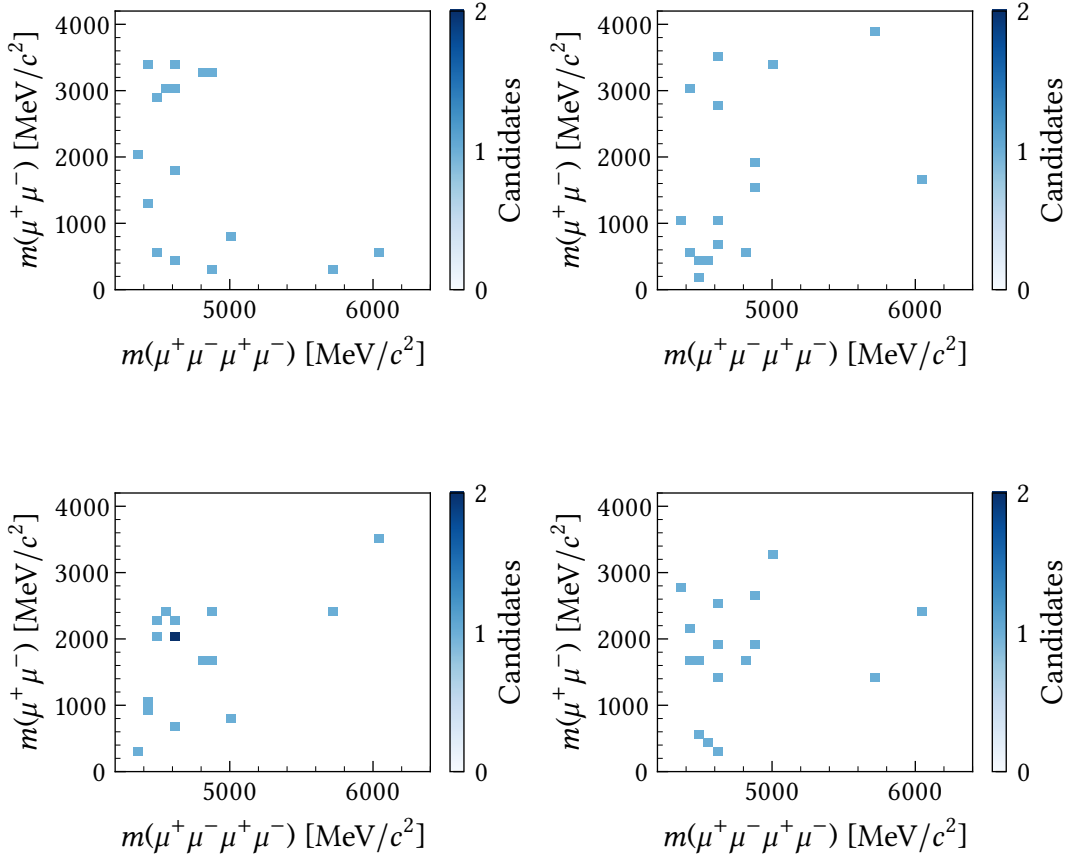


Figure 5.11: Two dimensional histograms of fully selected $B_{(s)}^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ data showing the four-muon invariant mass on the x and the di-muon invariant mass on the y -axis. The four plots correspond to the four different combinations of oppositely charged muons.

decays were used for normalisation, which implied a large systematic uncertainty originating from the S-wave fraction and the less precisely measured branching fraction. However, the measurement documented in this thesis is specifically tuned for the SM decays. More precisely, the selection is optimised for finding four muons that originate from a common vertex. This results in a low sensitivity for any potential BSM contribution via intermediate particles with non-negligible lifetimes, e.g. displaced vertices. The influence of the ad hoc chosen masses of the particles S and P is studied by measuring the selection efficiencies in different di-muon invariant mass ranges. A variation of the order of 20 % is observed. Recently the LHCb collaboration found evidence for the $\Sigma^+ \rightarrow p\mu^+\mu^-$ decay and measured the branching fraction to be compatible with the SM prediction and the HyperCP result. However, the di-muon invariant mass distributions did not display any feature that would hint towards an intermediate resonance [103].

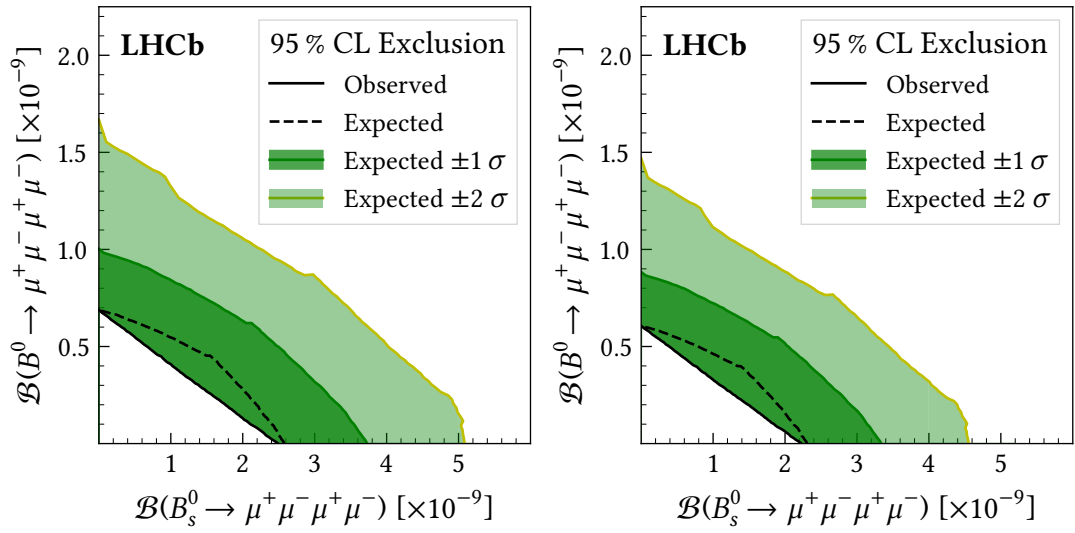


Figure 5.12: Expected and observed 95 % CL exclusion for the SM and the MSSM scenario for the B^0 and the B_s^0 decay as a function of each other. Both figures are refactored from [41].

6 Test of lepton flavour universality with $R_{K\pi\pi}$

This chapter details the test of lepton flavour universality based on $B^+ \rightarrow K^+ \pi^+ \pi^- \ell^+ \ell^-$ decays, where ℓ is either a muon or an electron. The measurement is still work in progress. More precisely, the efficiencies are not yet fully determined, such that this thesis does not contain a measured value of $R_{K\pi\pi}$. Furthermore, studies concerning systematic uncertainties are still ongoing. However, based on similar measurements previously published by the LHCb collaboration [15–17] it is assumed that the uncertainty on $R_{K\pi\pi}$ will be dominated by the limited sample size. Therefore, the ratio of signal yields is determined and based on that the precision of the full measurement is estimated.

The analysis is carried out by a small group. My main contributions are the development and optimisation of the signal selection including studies on physical backgrounds and the determination of the signal yields. In the following, first the strategy of the measurement is described. After that the corrections applied to simulated data to better reproduce measured data are discussed. Next, the signal selection is detailed. Finally, the ratio of signal yields is determined on a subsample of the available data and extrapolated to the full sample size in order to estimate the achievable sensitivity.

6.1 Strategy

For this measurement two final states that differ only in the flavour of charged leptons need to be analysed. Decays with electrons in the final state are more challenging to measure for the LHCb experiment than those with muons due to the imperfect reconstruction of bremsstrahlung. Therefore, the main focus is laid on optimising the analysis strategy for the di-electron final state and, where possible, treat the di-muon final state in the exact same way. This allows arising systematic uncertainties to cancel out when calculating the ratio of yields and efficiencies. Regarding the normalisation explained in section 4.1, quantities that introduce sizeable systematic uncertainties do not enter the ratio of branching fractions, even if only regarding the two rare decays. Nevertheless, the resonant modes are employed to calculate a double ratio. Thus, systematic uncertainties that are different between the two lepton flavours still cancel. In particular the PID selection and the trigger system are based on different detector

components for muons and electrons. Therefore, the double ratio

$$R_{K\pi\pi} = \underbrace{\frac{N_{K^+\pi^+\pi^-\mu^+\mu^-}^{\text{rare}}}{N_{K^+\pi^+\pi^-e^+e^-}^{\text{rare}}} \times \frac{N_{K^+\pi^+\pi^-e^+e^-}^{\text{resonant}}}{N_{K^+\pi^+\pi^-\mu^+\mu^-}^{\text{resonant}}}}_{R_{K\pi\pi}^{\text{yields}}} \times \underbrace{\frac{\varepsilon_{K^+\pi^+\pi^-e^+e^-}^{\text{rare}}}{\varepsilon_{K^+\pi^+\pi^-\mu^+\mu^-}^{\text{rare}}} \times \frac{\varepsilon_{K^+\pi^+\pi^-\mu^+\mu^-}^{\text{resonant}}}{\varepsilon_{K^+\pi^+\pi^-e^+e^-}^{\text{resonant}}}}_{R_{K\pi\pi}^{\text{efficiencies}}}, \quad (6.1)$$

including the resonant $B^+ \rightarrow K^+\pi^+\pi^-J/\psi (\rightarrow \ell^+\ell^-)$ decays that proceed via an intermediate J/ψ meson is computed. Here, ε represents the efficiency to detect, reconstruct and select the respective signal candidates and N stands for the signal yield of the corresponding decay mode.

The resonant and non-resonant decays are distinguished through the invariant mass of the di-lepton system. The composition of the hadronic systems was studied by the Belle collaboration [57] and a rich spectrum of intermediate particles is reported to contribute to the decay as can be seen in Figure 6.1. However, for the measurement presented in this thesis no difference between the individual processes is made. Instead, an integrated measurement over all contributing decays is performed. Nevertheless, measures are taken to reproduce the hadronic spectra in simulated data as detailed in section 6.2.1.

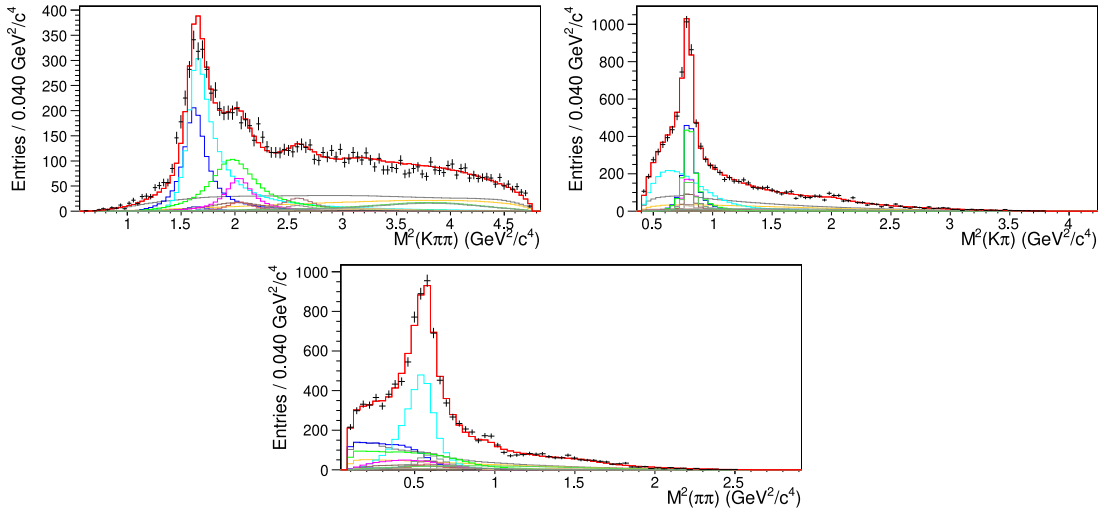


Figure 6.1: Invariant mass distributions of the $K^+\pi^+\pi^-$, $K^+\pi^-$ and $\pi^+\pi^-$ system reported by the Belle collaboration in an amplitude analysis of resonant $B^+ \rightarrow K^+\pi^+\pi^-\ell^+\ell^-$ decays [57]. Measured data are represented by black crosses, while the step functions correspond to various fit components that model background and contributions that proceed non-resonantly as well as via different intermediate states.

In order to avoid a possible bias, a blinding strategy is employed. The parametrisation and the knowledge of the exact composition of the B mass distribution of the di-electron mode is one of the main challenges of this measurement. Therefore, in contrast to the previously discussed search for $B_{(s)}^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ decays, not the mass

distribution, but the efficiencies are blinded. As this measurement is yet unfinished the efficiencies are still blind. Hence, the goal of this thesis is the determination of $R_{K\pi\pi}^{\text{yields}}$. This quantity is defined as the part of the $R_{K\pi\pi}$ ratio that consists of the signal yields as indicated in equation (6.1).

In the following, at first the agreement between simulated and measured data is probed and corrected where required. For this comparison, it is necessary to apply parts of the signal selection that is detailed in section 6.3. In the first selection step, the data sample is reduced to a manageable size to ease further analysis by removing obvious backgrounds. After that physical processes that can potentially fake the signal are studied and explicitly vetoed if necessary. At this stage the agreement between simulated and measured data in terms of the $K^+\pi^+\pi^-$ system and kinematic quantities is probed and corrected as needed. Next, a machine learning algorithm is optimised and trained to distinguish between signal candidates and combinatorial background. Then, requirements on the machine learning classifier and PID responses are optimised and applied. In the final step of the signal selection, the remaining low mass background is suppressed. Then the signal yields are determined by parametrising the measured $K^+\pi^+\pi^-\ell^+\ell^-$ invariant mass distributions via simultaneous unbinned extended maximum likelihood fits.

6.2 Simulation corrections

Many parts of the described measurement rely on Monte Carlo simulation, which does not perfectly reproduce measured distributions and must therefore be corrected. On top of that the signal simulation is generated with a simplified model, which needs to be accounted for. This is discussed in the first part of this section together with corrections of kinematic and track multiplicity variables. The second part deals with the calibration of the PID variables.

6.2.1 Kinematic corrections

The $K^+\pi^+\pi^-$ system in the studied final states can be produced in many different ways involving a rich spectrum of intermediate states. In this analysis an inclusive measurement of the full spectrum is performed. Hence, no difference is made between individual processes contributing to the creation of the $K^+\pi^+\pi^-$ system. However, due to the unavailability of a reliable model of the exact composition of the $K^+\pi^+\pi^-$ spectrum, signal samples are generated with an equal probability throughout the invariant mass of the $K^+\pi^+\pi^-$ system between 1100 and 2400 MeV/ c^2 . The individual momenta of kaons and pions are simulated with an equal probability through the available phase space. The same is true for the lepton pair in non-resonant signal simulation whereas in the resonant case the VLL model of the EvtGen package [104] is employed, which describes the decay of a vector meson into a pair of charged leptons.

In order to increase the agreement between simulated and measured data $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \ell^+ \ell^-)$ decays are regarded. Since at LHCb final states with muons are measured with a higher efficiency compared to those with electrons, $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \mu^+ \mu^-)$ decays are used as a proxy for measured data in the simulation correction procedure, which benefits from larger sample sizes. The $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow e^+ e^-)$ decay mode is employed to check whether the correction procedure developed on the di-muon mode can be adapted for the di-electron mode.

To avoid bias due to background contamination in measured data, *sWeights* (see section 4.4) are employed. These are obtained from an extended maximum likelihood fit to the invariant mass distribution of the $K^+ \pi^+ \pi^- J/\psi$ system. The di-lepton invariant mass is constrained to the J/ψ mass to improve the mass resolution and thereby increase the distance between the signal peak and partially reconstructed B decays. The latter are situated at lower invariant mass and can be avoided by starting the parametrised mass range at $5175 \text{ MeV}/c^2$. The signal part of the di-muon final state is modelled by a Hypatia function [81] and combinatorial background by an exponential function. For further explanations see section 4.3 for a discussion of different backgrounds, section 4.2 for the Hypatia function and section 4.4 for *sWeights*. For the di-electron final state the shape of the signal part is different depending on whether no, one or both electrons were assigned corrections for bremsstrahlung during the reconstruction. Therefore, each of the three categories is modelled by an individual Hypatia function. Due to the presence of combinatorial background, it is not possible to determine the parameters defining the tails of the Hypatia function solely from measured data. Therefore, the degrees of both power-law tails n_1 and n_2 and for the di-electron mode also the transition parameters α_1 and α_2 are fixed to values obtained from fits to the corresponding simulated data. The modelled $K^+ \pi^+ \pi^- J/\psi$ invariant mass distributions are presented in Figures 6.2 and 6.3. The signal yields amount to $137\,320 \pm 490$ for the di-muon final state and $15\,240 \pm 150$ for the di-electron mode.

These studies are based on data that already passed several selections. To be more precise, all decay products fell into the geometric acceptance of the LHCb detector, the tracks were successfully reconstructed, and the trigger, stripping and preselection criteria were fulfilled. It must be assumed that the efficiencies introduced by these requirements are biased by the inadequately simulated features, such that an unbiased comparison between measured and simulated data is not feasible. Therefore, the corrections to simulated data need to be applied at *generator level*, that is before the detector response is simulated. Hence, data are needed that reflect the measured signal unfolded from any efficiency. This is achieved by dividing the *sWeight* that is assigned to each candidate by the efficiency to pass the afore mentioned criteria. Here, the efficiency is measured as a function of the invariant masses of the $\pi^+ \pi^-$ and the $K^+ \pi^-$ system based on signal simulation such that a bias originating from the simplified simulation model is avoided. Simulated data generated under the 2011 and 2012 data taking conditions are treated separately to account for different centre-of-mass energies and different stripping and reconstruction efficiencies.

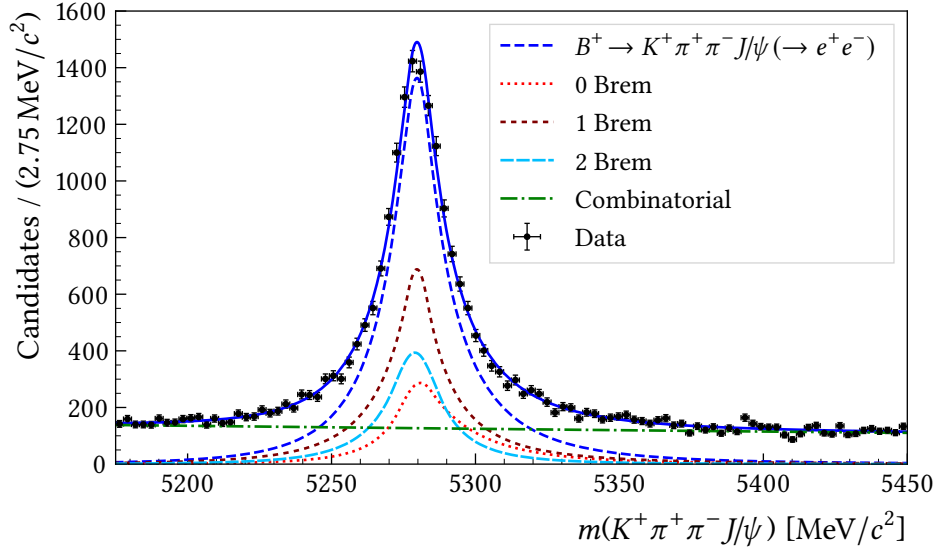


Figure 6.2: Invariant mass distributions of the $K^+ \pi^+ \pi^- e^+ e^-$ system. Here, the invariant di-electron mass is constrained to that of the J/ψ meson. Shown is the full Run 1 data sample that passed the stripping, trigger and preselection requirements. The distribution is modelled by a PDF as detailed in the text.

This results in a sample that in combination with the resulting weights is unfolded from any effect that introduces an efficiency such that the weighted data represents the signal as a 100 % efficient LHCb detector would see it. In the following, these data are used as the reference that simulated data needs to match.

Simulated data are corrected in three steps: In the first step a GB-reweighter (see section 4.5) is trained to enhance the agreement of the invariant mass distributions of the $\pi^+ \pi^-$, the $K^+ \pi^-$ and the $K^+ \pi^+ \pi^-$ system. In the second step another GB-reweighter is trained to correct differences in the pseudo rapidity and transverse momentum of the B meson. These two variables define the direction in which the B meson travels through the detector and thereby also the detector regions that are traversed. Therefore, this correction is essential for determining accurate efficiencies for the geometric acceptance and the track reconstruction.

In the third step the track multiplicity quantified by the number of hits in the SPD and the number of tracks reconstructed in the VELO is corrected. The track multiplicity is an important measure for the detector occupancy, which has a significant influence on for example the particle identification performance for hadrons. These are dominantly identified in the RICH detectors, that look for rings of Cherenkov light, which are much harder to reconstruct if many overlapping rings are present. On top of that, when multiple particles hit the same calorimeter cell, only one can be reconstructed, such that also the PID and trigger response for electrons is expected to depend on the track multiplicity.

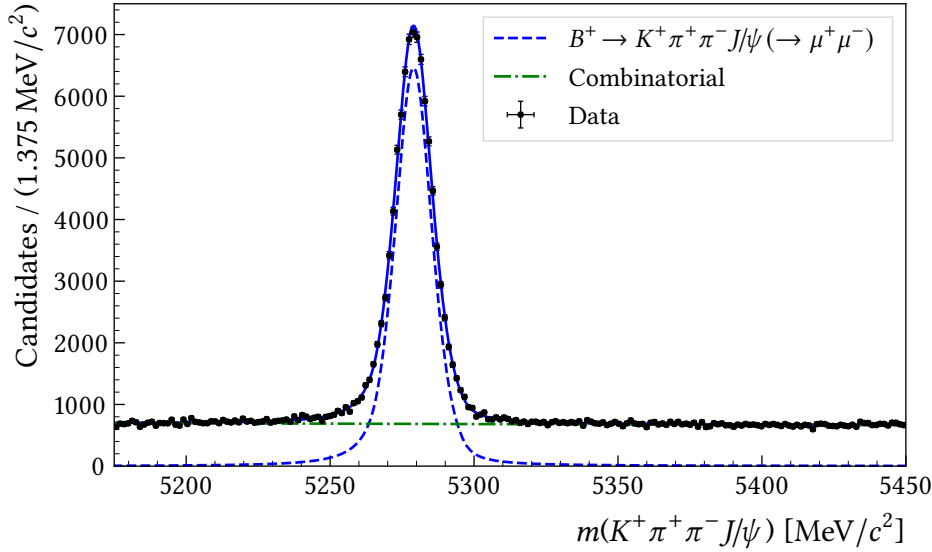


Figure 6.3: Invariant mass distributions of the $K^+ \pi^+ \pi^- \mu^+ \mu^-$ system. Here, the invariant di-muon mass is constrained to that of the J/ψ meson. Shown is the full Run 1 data sample that passed the stripping, trigger and preselection requirements. The distribution is modelled by a PDF as detailed in the text.

A comparison between measured and simulated data before and after the correction procedure is presented in Figure 6.4. Shown are the invariant mass distributions of the different hadron combinations. It can be seen that the simplified simulation model is successfully corrected such that it reproduces the measured invariant mass distributions of the three regarded particle combinations. The features that are corrected in the second and third step are presented in Appendix A.2.1. In general the agreement between simulated and measured data is significantly enhanced not only in the di-muon case, but also for di-electron decays. Another important aspect is that the successive correction steps are well compatible, such that one step does not decrease the agreement that was achieved by the previous step.

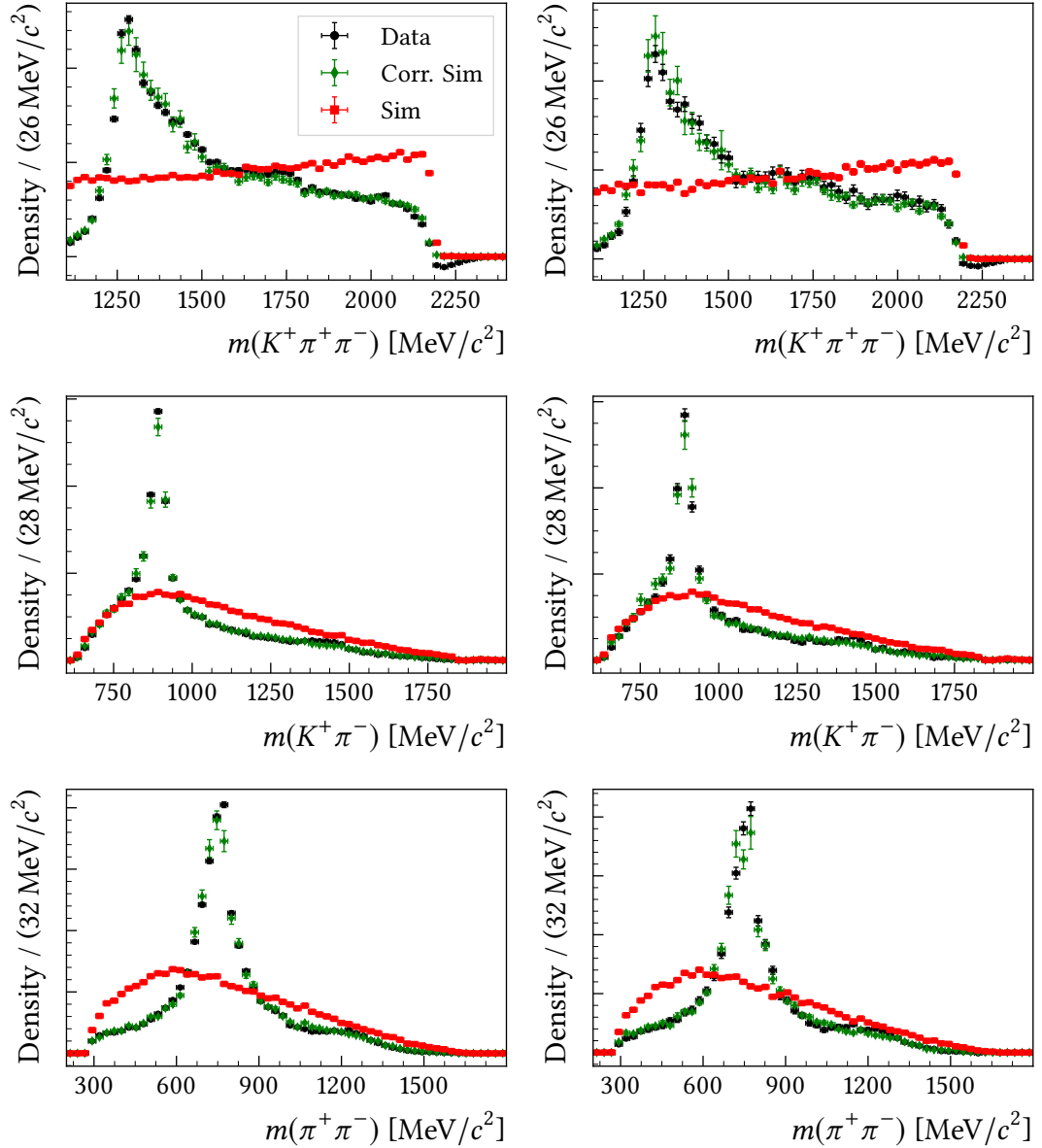


Figure 6.4: Invariant mass spectra of the $K^+ \pi^+ \pi^-$, $K^+ \pi^-$, and $\pi^+ \pi^-$ systems. Shown are simulated data before and after applying all three correction steps that are detailed in the text and measured data. For the latter $sWeights$ are employed in order to isolate the signal component. On the left resonant di-muon and on the right resonant di-electron decays are presented.

In an amplitude analysis of the $K^+\pi^+\pi^-$ final state performed by the Belle collaboration [57] the distribution of the squared $K^+\pi^+\pi^-$ mass is presented. The program WebPlotDigitizer [105] is used to extract the measured distribution from the plotted histogram. Thus, it is possible to compare the distributions measured from LHCb data and from the corrected simulation to the Belle result as done in Figure 6.5. For both lepton flavours the spectrum is dominated by the peak of the $K_1^+(1270)$ meson. However a second peak above the $K_1^+(1270)$ mass and another peak corresponding presumably to the $K^+(1600)$ meson that are evident in Belle data are not clearly visible in LHCb data. The former can be caused by multiple resonances such as the $K_1^+(1400)$, $K^*(1410)$ or $K_2^{*+}(1430)$ mesons and causes the $K_1^+(1270)$ peak in LHCb data to be wider towards higher masses. Due to a smaller boost of the produced B mesons and the resulting larger opening angle between their decay products the mass resolution of the Belle detector exceeds that of the LHCb detector. Therefore, the spectrum measured by the B factory shows multiple distinct peaks that are not clearly visible in data recorded with the LHCb experiment. In general, the presented histograms of corrected simulation show larger uncertainties in more densely populated regions. This is a consequence of the simulation model, which generates a uniformly distributed $K^+\pi^+\pi^-$ invariant mass. Therefore, the correction weights are largest in the $K_1^+(1270)$ peak. The analogous effect is observed for the K^{*0} peak in the $K^+\pi^-$ spectra. The hadrons are generated with an equal probability throughout the available phase space and therefore do not reproduce any resonant structure. Consequently, large weights need to be assigned in order to form the K^{*0} peak. Comparing the measured LHCb data between the two lepton flavours, the uncertainties in the di-electron case are slightly larger compared to the di-muon mode. This is a consequence of the smaller signal yield caused by lower efficiencies in electron reconstruction compared to muon reconstruction.

The final part of the signal selection detailed in section 6.3.5 is based on a machine learning classifier. This algorithm is trained on measured background and simulated signal decays to identify the sought after decays based on eleven features. Therefore, it is crucial that these features are accurately simulated. In Figure 6.6 the measured distributions of a subset of the used features are compared to those observed in simulated data before and after applying the correction. The remaining six features can be found in Appendix A.2.2. For a detailed explanation of the mentioned quantities see section 3.4. In most features good consistency is observed even without correction. Especially for the transverse momentum and vertex quality of the B^+ meson the agreement is further improved. The most significant difference between raw simulated and measured data is observed in the distributions of the DIRA of the $K^+\pi^+\pi^-$ system with respect to the PV. The peak around one is narrower in simulated than in measured data, which is successfully corrected by the described procedure. Remaining differences between simulated and measured data will be accounted for by a systematic uncertainty. Corresponding studies are still being finalised.

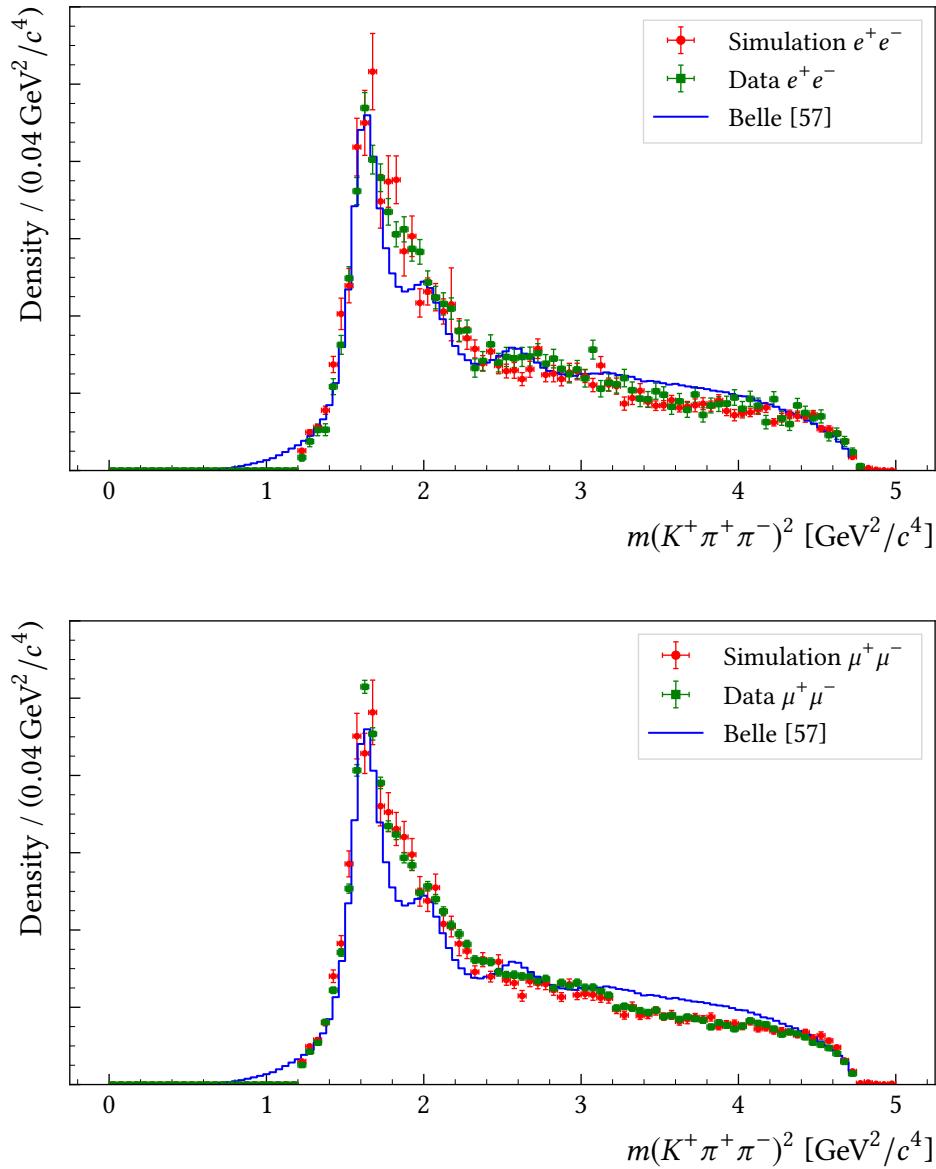


Figure 6.5: Comparison of the $m(K^+\pi^+\pi^-)^2$ spectrum in measured and corrected simulated data overlaid by the distribution reported by the Belle collaboration [57]. In the upper plot $B^+ \rightarrow K^+\pi^+\pi^- J/\psi (\rightarrow e^+e^-)$ decays are presented and in the lower $B^+ \rightarrow K^+\pi^+\pi^- J/\psi (\rightarrow \mu^+\mu^-)$ decays.

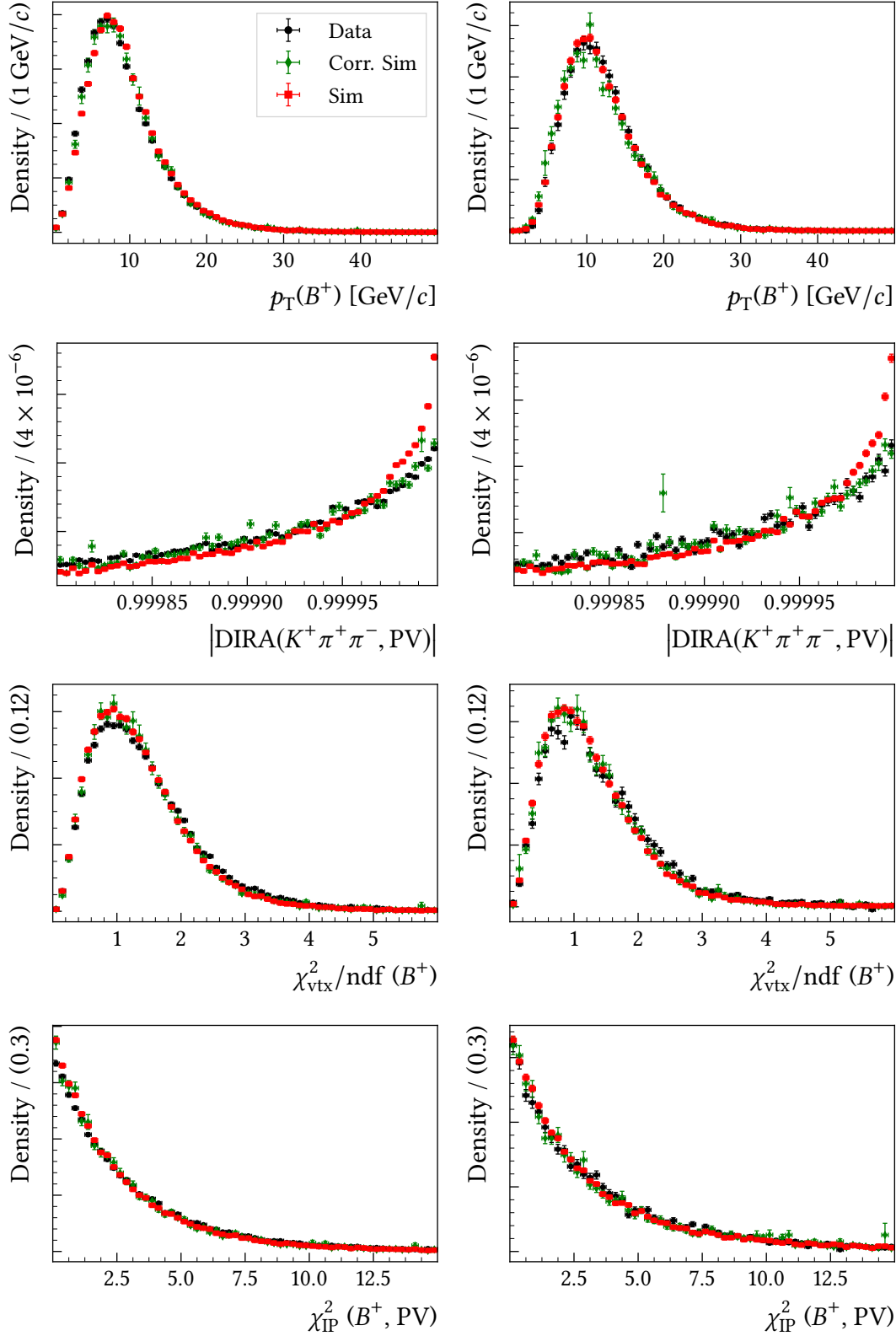


Figure 6.6: Comparison between measured data, raw simulation and corrected simulation for the first four out of eleven features used for the machine learning based selection. Shown are resonant di-muon decays (left) and resonant di-electron decays (right). For measured data *sWeights* are employed in order to isolate the signal component. For the definitions of the mentioned quantities see section 3.4.

6.2.2 Particle identification calibration

In the signal selection several different PID quantities are employed. Requirements on the likelihood based PID classifiers $DLL_{e\pi}$ and $DLL_{K\pi}$ are already applied at stripping level, such that the stripping selection excluding the PID criteria needs to be rerun on all simulated samples in order to avoid bias. This step is possible only for simulated data, such that comparisons between simulated and measured data throughout this section already include the PID requirements illustrated in Figures 3.7 and A.4. In an early stage of the selection presented in Section 6.3.2 requirements on neural network based PID quantities are made. For further details on the different particle identification quantities see Section 3.2. Applying the same PID requirement to both oppositely charged particles of the same type in the final state is equivalent to applying the requirement to the minimum PID response of the two. Therefore, in the following the minima between respectively both leptons and both pions are considered, such that

$$\min DLL_{e\pi} = \min(DLL_{e\pi}(e^+), DLL_{e\pi}(e^-)), \quad (6.2)$$

$$\min \text{ProbNN}_\pi = \min(\text{ProbNN}_\pi(\pi^+), \text{ProbNN}_\pi(\pi^-)), \quad (6.3)$$

$$\min \text{ProbNN}_\ell = \min(\text{ProbNN}_\pi(\ell^+), \text{ProbNN}_\pi(\ell^-)), \quad (6.4)$$

where ℓ is either μ or e . Furthermore, for the final part of the selection discussed in section 6.3.6 the composite PID quantities

$$\text{comp} \mathcal{PID}_\pi = \text{ProbNN}_\pi \times (1 - \text{ProbNN}_K) \times (1 - \text{ProbNN}_p), \quad (6.5)$$

$$\text{comp} \mathcal{PID}_K = \text{ProbNN}_K \times (1 - \text{ProbNN}_p), \quad (6.6)$$

are defined and for the pions again the minimum of the two oppositely charged candidates is calculated as

$$\min \text{comp} \mathcal{PID}_\pi = \min(\text{comp} \mathcal{PID}_\pi(\pi^+), \text{comp} \mathcal{PID}_\pi(\pi^-)). \quad (6.7)$$

The composite PID quantities bring the advantage of not just representing good compatibility with the desired particle species, but also penalise compatibility with undesired species.

The response of the PID system in simulated data is corrected by the *PIDCorr* method, which is detailed in section 4.6. The distributions of the different PID responses used throughout this measurement are illustrated in Figure 6.7 for the $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow e^+ e^-)$ decay mode and in Appendix A.2.3 for the respective di-muon decay mode. The three different line shapes represent the measured signal component isolated from background through *sWeights* and simulated data before and after calibrating the respective PID response. It can be seen that the raw simulated PID responses are significantly different from those observed in measured data. After calibration though these differences are significantly reduced, such that good agreement between calibrated and measured PID response can be seen. Due to the smaller sample size the uncertainties in case of measured $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow e^+ e^-)$ decays are considerably larger than those for the corresponding di-muon data.

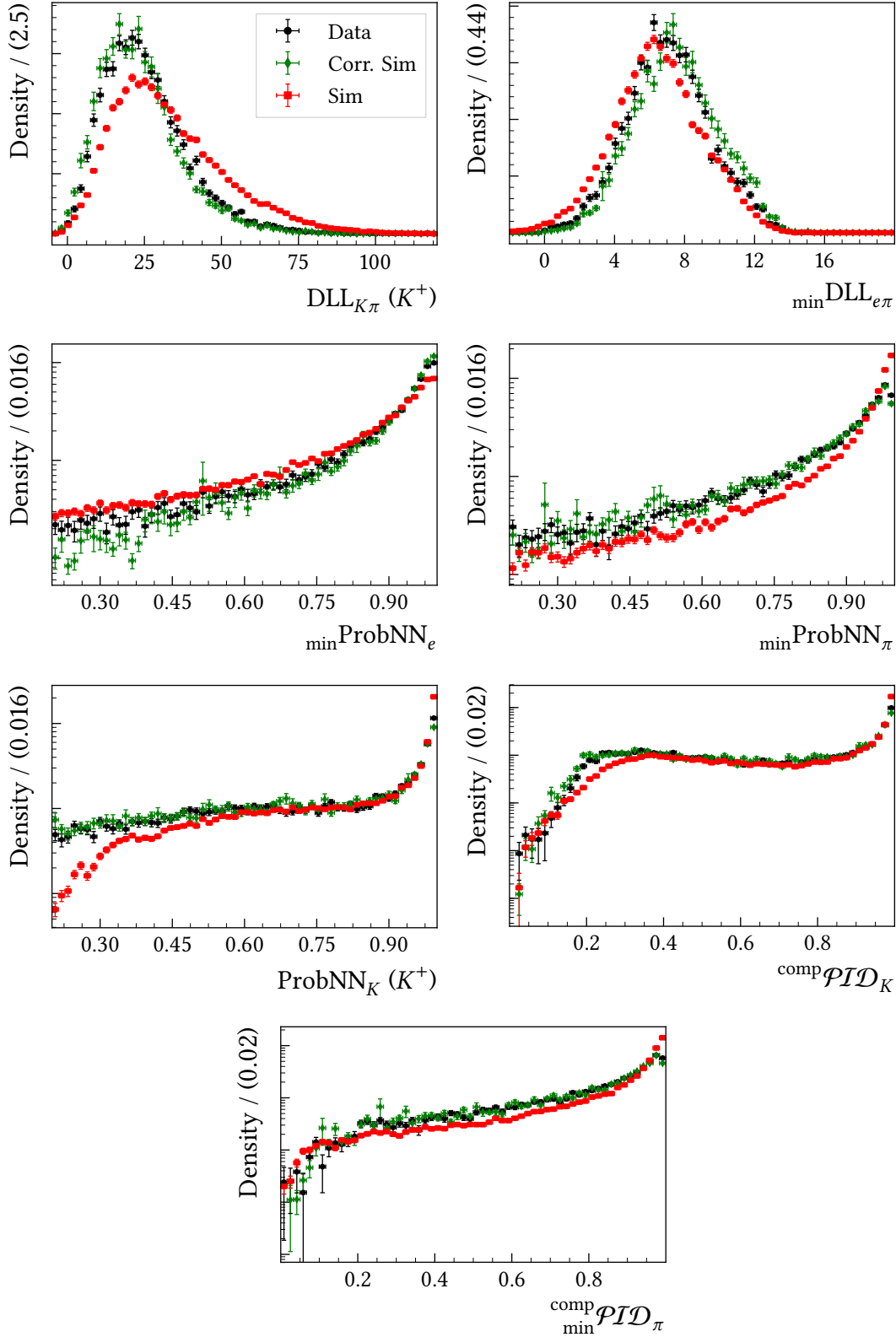


Figure 6.7: Comparison between measured, raw simulated and corrected simulated data for the PID quantities employed to select the di-electron final state. The comparison is based on simulated and measured $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow e^+ e^-)$ decays. For the latter $sWeights$ are employed to avoid contributions from background.

6.3 Selection

In this section the individual steps of the signal selection are detailed. First, the considered trigger lines are discussed. Then the data sample is cleaned from obvious backgrounds and selected sources of physical processes that can potentially fake the $B^+ \rightarrow K^+ \pi^+ \pi^- \ell^+ \ell^-$ signal. After that, the split between resonant and non-resonant data is explained. Finally, a machine learning classifier is trained and selection criteria on its response and different PID quantities are optimised.

6.3.1 Trigger selection

For this measurement candidates are considered that satisfy the MTOS requirement in case of the di-muon decays or either the TIS or ETOS criteria in case of the di-electron modes. The mentioned categories are defined in Table 6.1. The di-electron data are subdivided at lowest trigger level into two categories, which are based on different selection criteria and hence require different approaches to correct for disagreements between data and simulation. These corrections are still under development and are hence not presented in this thesis. The first category (ETOS) consists of events that were recorded due to at least one of the two signal electrons satisfying the requirements made by the L0Electron trigger. Here, trigger decisions are based on the energy deposited in the ECAL by the electron candidate. Hence, corrections will be determined as a function of this quantity. The second category (TIS) includes candidates that do not belong to the first category and achieved a positive trigger decision through a particle that is not associated with the B^+ signal decay. As most trigger lines select events based on the transverse momentum of tracks, corrections to TIS and HLT efficiencies measured in simulated data will be determined as a function of the transverse momentum of the B candidate. In contrast to the measurement of R_{K^*0} performed by the LHCb collaboration no third category based on the hadron trigger is considered. It was shown that the estimation of the hadron trigger efficiency is challenging and hence results in a large systematic uncertainty [17]. Part of the signal candidates lost due to this decision is recovered in the TIS category. The remaining signal loss is determined to be 2.7%. The criteria demanded by the trigger system are subject to permanent optimisation in order to always record as much relevant data as possible. To obtain a consistent data sample the tightest selection applied by the trigger during the considered data taking periods is reproduced and demanded for all studied data. This leads to requirements that are applied to the respective lepton that caused a positive trigger decision. For muons the trigger decisions are based on the transverse momentum and for electrons on the transverse energy deposited in the ECAL.

Additionally to the L0 criteria, candidates must fulfil at least one HLT1 and at least one HLT2 requirement specified in Table 6.2 separately for the two considered lepton flavours. The individual requirements made by all mentioned trigger lines are detailed in 3.5.

Table 6.1: Definition of the trigger categories together with the considered L0 trigger lines and additional criteria that represent the most stringent trigger selection during the considered data taking period. The underlying criteria of each mentioned trigger line are detailed in section 3.5.

μ^\pm : MTOS	μ^\pm L0MuonDecision TOS
	$p_T(\mu^\pm) > 1.76 \text{ GeV}/c$
e^\pm : ETOS	e^\pm L0ElectronDecision TOS
	$E_T(e^\pm) > 2.96 \text{ GeV}$
e^\pm : TIS	B L0Global TIS
	not ETOS

Table 6.2: Data is required to pass at least one of the HLT1 and one of the HLT2 trigger lines presented in this table individually for the di-electron and di-muon decays. The underlying criteria of each mentioned trigger line are detailed in section 3.5.

electron	HLT1	B TrackAllL0Decision TOS
	HLT2	B Topo[2,3,4](E)BodyBBDTDecision TOS
muon	HLT1	B TrackAllL0Decision TOS
		\parallel B TrackMuonDecision TOS
	HLT2	B Topo[2,3,4](Mu)BodyBBDTDecision TOS
		\parallel B DiMuonDetachedDecision TOS

6.3.2 Preselection

Due to the very soft PID requirements demanded in the stripping the data contain a significant amount of misidentified particles as can be seen in Figure 6.8. The presented histograms are based on data that already passed the stripping, trigger and preselection except for the requirements made on the neural network based PID quantities. The upper two histograms illustrate the distribution of the PID variable quantifying the agreement for pion and muon candidates to match their assigned identity. Both show peaks at zero that indicate contributions from misidentified particles. The bottom two figures show two dimensional distributions of different PID responses for respectively the same candidates. The response matching the assigned identity is plotted on the x -axis and that of a chosen undesired identity on the y -axis. Especially in the bottom left figure a large cluster of kaon candidates with low ProbNN_K and large ProbNN_π can be seen. This feature is to be interpreted as kaon candidates that are misidentified pions. A similar effect, less pronounced though, can be seen in the bottom right figure, which shows ProbNN_e versus ProbNN_π for electron candidates.

In order to remove this clearly visible background, requirements are placed on the neural network based PID quantity that matches the assigned identity of each final state particle to be larger than 0.2. Different versions of the neural networks that predict the PID variables exist. Relevant to the presented measurement are MC12TuneV3 and MC12TuneV2. Due to better performance the former is employed to quantify compatibility for a given particle to be an electron or muon and the latter for compatibility with being a pion, kaon or proton.

Besides the physically motivated requirements, the preselection also incorporates technical aspects. The PID calibration data that are used in section 6.2.2 to correct inadequately simulated PID responses undergoes a selection applied by the LHCb PID group. In order to avoid a potential bias, these requirements are also applied to the analysed simulated and measured data samples as part of the preselection. Electron candidates traversing the very central region of the ECAL, which is not read out, are vetoed. As the ECAL signatures are crucial inputs to PID and trigger decisions for electrons the efficiencies to select these candidates is very low, such that the loss of signal efficiency due to the introduced veto is small. Furthermore, the transverse energy deposited by the electron candidates in the ECAL will be used to parametrise simulation corrections to account for mismodelling of the trigger system, which would fail for the vetoed candidates. Following a general recommendation made by the LHCb tracking group the fake track classifier (GhostProb) [78] of each final state track is required to be less than 0.3. The mass of the $K^+\pi^+\pi^-$ system is required to be between 1100 and 2400 MeV/c^2 , which is the interval that is populated by the signal simulation.

All mentioned requirements are summarised in Table 6.3.

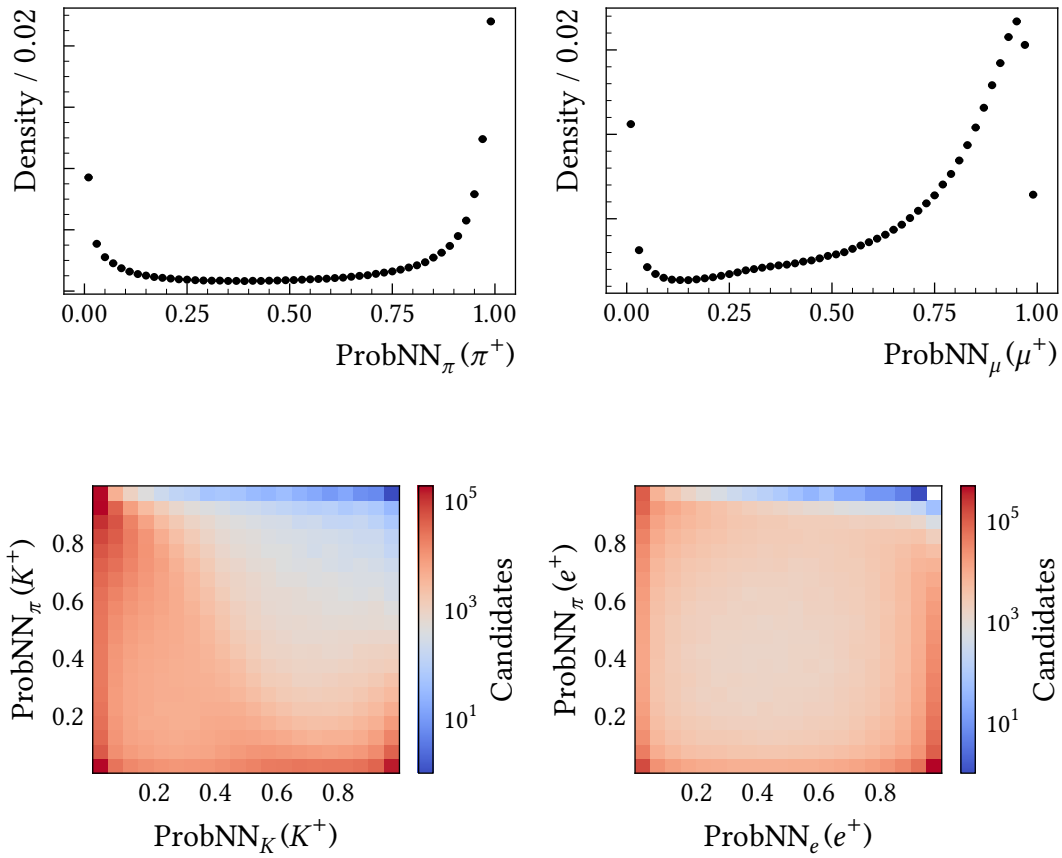


Figure 6.8: Histograms of particle identification classifier responses for all four different particle species present in the final states regarded for this measurement. The shown data correspond to the full Run 1 data sample passing the stripping, trigger and preselection except for the introduced PID criteria.

Table 6.3: Preselection requirements applied to simulated and measured data samples to remove obvious background and ease further processing.

Adapted from PID Calibration	
π^\pm, K^\pm	within the acceptance of the muon system
e^\pm, π^\pm, K^\pm	have associated Cherenkov rings in the RICH system
e^\pm	have associated clusters in the calorimeter system
μ^\pm	IsMuon
K^\pm, e^\pm, μ^\pm	$p > 3 \text{ GeV}/c$
π^\pm	$p > 2 \text{ GeV}/c$
K^\pm	$p_T > 200 \text{ MeV}/c$
π^\pm	$p_T > 250 \text{ MeV}/c$
μ^\pm	$p_T > 800 \text{ MeV}/c$
e^\pm	$p_T > 500 \text{ MeV}/c$
Veto against misidentified particles	
K^\pm	$\text{ProbNN}_K \text{ MC12TuneV2} > 0.2$
π^\pm	$\text{ProbNN}_\pi \text{ MC12TuneV2} > 0.2$
e^\pm	$\text{ProbNN}_e \text{ MC12TuneV3} > 0.2$
μ^\pm	$\text{ProbNN}_\mu \text{ MC12TuneV3} > 0.2$
Veto deactivated ECAL region	
e^\pm	ECAL cluster x position $> 363.6 \text{ mm}$
e^\pm	ECAL cluster y position $> 282.6 \text{ mm}$
Recommendation by Tracking Group	
$K^\pm, \pi^\pm, \mu^\pm, e^\pm$	Track GhostProb < 0.3
Mass range of $K^+ \pi^+ \pi^-$ system	
$K^+ \pi^+ \pi^-$	$1100 < m < 2400 \text{ MeV}/c^2$

6.3.3 Invariant mass intervals

The analysed data are divided into two samples by the stripping selection. One includes the decays with electrons in the final state, the other those with muons. In the following, an additional split is introduced in order to separate the decays proceeding via the rare FCNC processes from those where the lepton pair is produced via an intermediate J/ψ resonance. These resonant decays are tree level processes and have branching fractions that are two orders of magnitude larger than those of the rare decays [38]. For this reason and because of the equal composition of the final states, the resonant modes are chosen for the normalisation and to validate the analysis strategy. Resonant and non-resonant data are distinguished from one another through the di-lepton invariant mass squared, q^2 . The lower boundary for non-resonant decays is set to $q^2 = 1.1 \text{ GeV}^2/c^4$. This guarantees a sufficiently large distance from the di-lepton production threshold, such that differences between the branching ratios of the two studied lepton flavours that originate from phase space effects are avoided. Many previous measurements of rare B decays set the border between the non-resonant and the J/ψ region at $q^2 = 6 \text{ GeV}^2/c^4$ [15,17], as especially for the $J/\psi \rightarrow e^+e^-$ decay the J/ψ peak is smeared out towards lower masses due to not reconstructed bremsstrahlung. However, to increase the signal yield of the non-resonant mode for this measurement $q^2 = 7 \text{ GeV}^2/c^4$ is chosen. Remaining resonant signal is removed by vetoing an interval around the B^+ mass in the $K^+\pi^+\pi^-J/\psi$ invariant mass calculated constraining the invariant mass of the di-lepton system to the known J/ψ mass. The vetoed interval includes 95 % of all simulated $B^+ \rightarrow K^+\pi^+\pi^-J/\psi (\rightarrow e^+e^-)$ candidates. Throughout this thesis only non-resonant data in the region 1.1 to $7 \text{ GeV}^2/c^4$ are regarded. Once these are successfully analysed, the measurement can be extended into additional q^2 regions. Another mass region, namely the upper side band is defined, which is assumed to contain only combinatorial background. The data falling into this region are used later on to calculate background efficiencies and to train machine learning classifiers. The exact definitions of all mentioned regions are summarised in Table 6.4.

Table 6.4: Definitions of intervals in the invariant mass of different particle combinations.

Name	Dimension	Range
Resonant signal	q^2	7–11 GeV^2/c^4
Non-resonant signal	q^2	1.1–7 GeV^2/c^4
	$m(K^+\pi^+\pi^-J/\psi)$	not in 5139.3–5476.2 MeV/c^2
Upper side band	$m(K^+\pi^+\pi^-\ell^+\ell^-)$	> 5650 MeV/c^2

6.3.4 Vetoes against physical backgrounds

Several decays are studied that can potentially fake the $B^+ \rightarrow K^+ \pi^+ \pi^- \ell^+ \ell^-$ signals. In a first step the branching fractions of potential contributions are regarded and every process with a branching fraction of the order of 10^{-9} or less is discarded. The branching fraction of the non-resonant $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ decay has been measured to be $(4.3 \pm 0.4) \times 10^{-7}$ and around 150 signal candidates were observed in the q^2 region that roughly corresponds to the interval studied in this thesis [58]. Assuming the same efficiency for background and signal candidates, processes with branching fractions of the order of 10^{-9} contribute at the order of one candidate. However, due to particle misidentification and/or only partial reconstruction of the original decay the efficiencies are significantly smaller making these processes negligible.

In the next step the RapidSim [106] package is used to generate simulated samples of all relevant processes. RapidSim performs a *phase space simulation*, which means that decay products equally populate the available phase space. The major difference to the LHCb simulation framework is that the detector response and all following steps, such as the reconstruction and the trigger system are omitted. Instead, effective resolution models are applied to smear the simulated momenta. Consequently, only very basic quantities, as for example the momenta and therefore also invariant mass distributions are available, while high level features such as PID responses and vertex or track qualities are not. The advantage of RapidSim however is that the required computing time is negligibly small compared to that of the nominal simulation software. With RapidSim, it is feasible to generate decent sample sizes within seconds, whereas the full detector simulation already needs several minutes for a single event making the use of computing clusters inevitable.

Because of the absence of high level features it is not possible to apply the machine learning classifier introduced in section 6.3.5 to samples generated with RapidSim. However, the PID performance can at least be estimated. As described in section 4.6, the PID performance is parametrised in terms of momentum, pseudo rapidity and detector occupancy quantified by the track multiplicity in the corresponding event. The latter quantity is not simulated by RapidSim. Following the most conservative approach, two dimensional projections of the PID performance maps are created by taking the bin showing the largest misidentification rate along the track multiplicity dimension. Based on these histograms rates for every regarded final state to pass the PID criteria demanded in the preselection are calculated.

The decays $B^+ \rightarrow J/\psi (\rightarrow \pi^+ \pi^- \pi^+ \pi^- \pi^0) K^+$ and $B^+ \rightarrow J/\psi (\rightarrow \pi^+ \pi^- \pi^0) K^+ \pi^+ \pi^-$ have branching fractions of the order 10^{-5} . When misidentifying two oppositely charged pions as leptons these decays can fake the signal if the neutral pion is not reconstructed. In the di-muon case the partial reconstruction results in a sufficient shift of the reconstructed B^+ mass to separate these decays from the signal. However, due to the lower resolution caused by missed bremsstrahlung for the di-electron mode this is not the case. Therefore, the misidentification rate of two pions as two electrons needs to be studied. Based on the two-dimensional efficiency maps the probability to

misidentify one pion as an electron is estimated to be 1.2% resulting in a suppression of the mentioned decays of the order of 10^{-4} . Therefore, these backgrounds are reduced to a negligible level.

Decays via D mesons

Contributions that are still not ruled out by the above mentioned methods are either vetoed explicitly or regarded in full detector simulation, such that the efficiency of the complete selection can be determined. Explicit vetoes are applied against hadronically decaying D^0 and D^\pm mesons by removing a mass window from 1838 to 1898 MeV/c^2 in the $K^\pm\pi^\mp$, $K^\pm\ell^\mp_{\rightarrow\pi^\mp}$, $K^\pm\pi^\mp\ell^\mp_{\rightarrow\pi^\mp}$ and $K^\pm\pi^\pm\pi^\mp\ell^\mp_{\rightarrow\pi^\mp}$ invariant mass distributions. The subscripts indicate that the corresponding invariant mass has been calculated taking into account misidentification of a lepton as a pion. The vetoed region is the same for the D^0 and D^\pm mesons, ignoring the small mass difference in favour of simplicity, and covers more than 99% of the simulated D mesons.

Full detector simulation is regarded for the double semileptonic decay $B^+ \rightarrow D^-(\rightarrow K^+\pi^-\ell^-\bar{\nu}_\ell)\pi^+\ell^+\nu_\ell$ and the two semileptonic decays $B^+ \rightarrow D^0(\rightarrow K^+e^-\bar{\nu}_e)\pi^+\pi^+\pi^-$ and $B^+ \rightarrow D^-(\rightarrow K^+\pi^-e^-\bar{\nu}_e)\pi^+\pi^+$. Neutrinos are invisible to the LHCb detector such that D mesons that undergo a semileptonic decay do not cause a sharp mass peak. The shape of the peak is significantly smeared towards lower masses caused by the missing energy of the neutrino. Therefore, these decays can escape the introduced vetoes against contributions from D meson decays. The branching fractions of these decays are too large to produce sufficient amounts of simulated data to exclude them by directly applying the full selection to measure the efficiency. However, under the assumption that the efficiencies of the PID and the BDT requirements factorise, the efficiency of the full signal selection can be estimated based on the available small simulated samples. In order to validate this assumption the track multiplicity, transverse momenta and pseudo rapidities of all final state particles are compared before and after applying the BDT requirement based on non-resonant di-electron signal simulation. The agreement is found to be equally well for all final state particles. As two out of the three potential background processes involve the misidentification of a pion as an electron, Figure 6.9 illustrates this comparison for one of the two electron candidates. Only small differences, mostly at lower transverse momenta are observed. Calculating the expected contributions based on the measured efficiencies, branching fractions, the integrated luminosity and the $b\bar{b}$ production cross section shows, that additional measures need to be taken in order to exclude these background sources. In order to suppress the misidentification of pions as electrons the $\text{DLL}_{e\pi}$, representing the difference in log-likelihood between electron and pion hypothesis, is required to be larger than 3. The threshold is chosen ad hoc and is the same as applied for the measurement of $R_{K^{*0}}$, where it is reported to reduce the pion-electron misidentification to a negligible level [17]. The expected background yields of the three considered decays are between 0.02 and 0.16. Hence, it is concluded that contributions from the three semileptonic decays are reduced to a

negligible level.

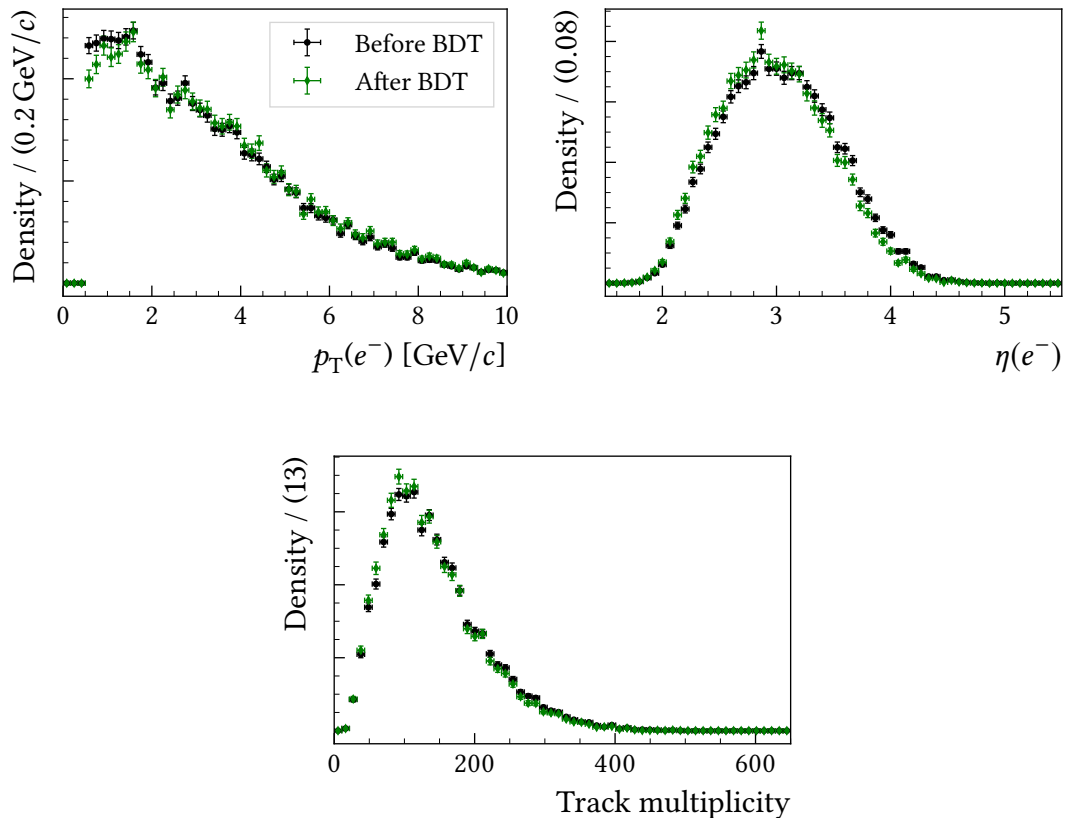


Figure 6.9: Comparison of variables correlated with the PID efficiencies. Shown are simulated data of non-resonant di-electron decays before and after applying the BDT selection.

Resonant decays

When swapping one of the final state mesons with one of the final state leptons $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \ell^+ \ell^-)$ decays can escape the respective vetoes and contribute to the non-resonant signal. This issue is addressed by removing a $100 \text{ MeV}/c^2$ wide area around the known J/ψ mass in the invariant mass distributions of all neutral combinations of a final state lepton with a final state meson. Hereby, the mass of the meson is replaced by that of the respective lepton when calculating the invariant mass of the two particle system. On top of that, these decays can contribute in the di-electron case if a large amount of energy was emitted as bremsstrahlung and not recovered by the detector. As this background cannot be removed efficiently, it is accounted for in the fit model as detailed in section 6.4.

Higher charmonia

Contributions from $B^+ \rightarrow \psi(2S)(\rightarrow J/\psi(\rightarrow \ell^+ \ell^-)\pi^+ \pi^-)K^+$ do not fall into the non-resonant data samples due to the veto against resonant decays, but can significantly enhance the $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi(\rightarrow \ell^+ \ell^-)$ signal yields. Therefore, a requirement is placed on the invariant mass of the $J/\psi \pi^+ \pi^-$ system. Here, the invariant mass of the di-lepton system is constrained to the J/ψ mass. In order to determine the necessary size of the vetoed region the distribution is modelled by a Gaussian function accounting for $B^+ \rightarrow \psi(2S)(\rightarrow J/\psi(\rightarrow \ell^+ \ell^-)\pi^+ \pi^-)K^+$ decays and an exponential function to parametrise combinatorial background. The respective mass peak is wider for the di-electron final state because of the energy loss due to bremsstrahlung. Therefore, the study is performed based on $B^+ \rightarrow \psi(2S)(\rightarrow J/\psi(\rightarrow e^+ e^-)\pi^+ \pi^-)K^+$ data and adapted for the respective di-muon mode. Figure 6.10 shows the invariant mass of the $J/\psi \pi^+ \pi^-$ system modelled by the mentioned PDF. Although the Gaussian function does not perfectly describe the observed shape, it is sufficient to assign a veto corresponding to five standard deviations estimated by the fit. The interval ranges from 3663 to 3709 MeV/c^2 and is illustrated by the shaded region.

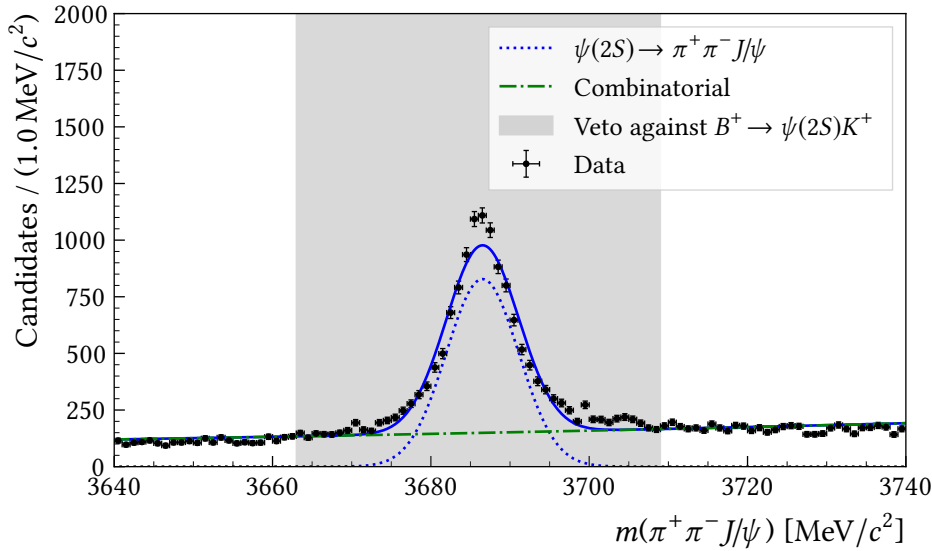


Figure 6.10: Invariant mass distribution of the $e^+ e^- \pi^+ \pi^-$ system in resonant data modelled by a PDF as detailed in the text. Here, the di-electron mass is constrained to that of the J/ψ -meson.

An additional charmonium resonance that needs to be considered as background is the χ_{c1} meson. Similar to the decay via the $\psi(2S)$ particle, $B^+ \rightarrow \chi_{c1}(3872)(\rightarrow J/\psi \pi^+ \pi^-)K^+$ decays give rise to the resonant signal yields. As before, the peak in the $m(J/\psi \pi^+ \pi^-)$ distribution is modelled by a Gaussian PDF and an exponential function. A mass window from 3848 to 3894 MeV/c^2 corresponding to three standard deviations of the Gaussian function is removed. Here, the veto is chosen to be narrower than that against the similar process proceeding via a $\psi(2S)$ meson because the branching fraction is smaller by two orders of magnitude [28].

As can be seen in Figure 6.11 both vetoes fall into a sparsely populated region of the corresponding distribution in corrected $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow e^+ e^-)$ signal simulation. Hence, the loss of resonant signal yield due to the assigned vetoes is small.

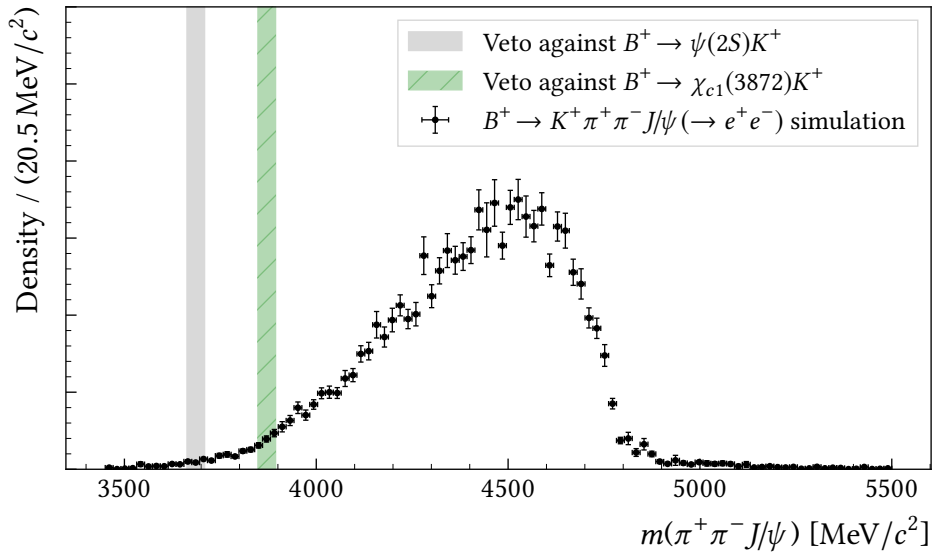


Figure 6.11: Invariant mass distribution of the $e^+ e^- \pi^+ \pi^-$ system in simulated resonant di-electron decays. Here, the di-electron mass is constrained to that of the J/ψ -meson.

$B^0 \rightarrow K^+ \pi^- \ell^+ \ell^-$ decays

Another type of processes that can give rise to the searched signals are $B^0 \rightarrow K^+ \pi^- \ell^+ \ell^-$ decays. The meson pair is produced dominantly via an intermediate K^{*0} meson. Here, again resonant and non-resonant decays exist. In order to fake the $B^+ \rightarrow K^+ \pi^+ \pi^- \ell^+ \ell^-$ signals a pion needs to be randomly added in the reconstruction of the decay. Data obtained from full detector simulation of the four decay modes $B^0 \rightarrow K^{*0} J/\psi (\rightarrow \ell^+ \ell^-)$ and $B^0 \rightarrow K^{*0} \ell^+ \ell^-$, respectively for the two studied lepton flavours, are analysed. For resonant $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \ell^+ \ell^-)$ decays, the non-resonant type of background can be neglected as the branching fractions differ by three orders of magnitude. Resonant decays however must be vetoed, which is efficiently done by requiring $m(K^+ \pi^- J/\psi) < 5175 \text{ MeV}/c^2$. Thanks to the excellent

mass resolution due to the constrained di-lepton mass more than 99 % of the regarded background process is vetoed while almost no resonant signal is lost.

Regarding non-resonant signal, contamination from $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decays in the di-muon signal can be excluded by regarding the $m(K^+ \pi^- \mu^+ \mu^-)$ and $m(K^+ \pi^- J/\psi)$ distributions in data that passes the complete selection. Not a single event falls into a region corresponding to two standard deviations around the B mass simultaneously in $m(K^+ \pi^+ \pi^- \mu^+ \mu^-)$ and $m(K^+ \pi^- \mu^+ \mu^-)$ and no hint of candidates clustering around the B mass neither in $m(K^+ \pi^- \mu^+ \mu^-)$ nor in $m(K^+ \pi^- J/\psi)$ is found.

For the di-electron mode however, resonant decays can leak into the non-resonant q^2 interval due to not reconstructed bremsstrahlung. In fact a peak at the B mass is visible in the $m(K^+ \pi^- J/\psi)$ distribution non-resonant di-electron data after applying the full selection. Therefore, the interval from 5175 to 5350 MeV/c^2 is vetoed, which reduces this background to a negligible level.

The rate of non-resonant $B^0 \rightarrow K^{*0} e^+ e^-$ decays contaminating the non-resonant di-electron signal is estimated based on the efficiency, which is determined by applying the complete non-resonant selection to the corresponding simulation. The expected contribution is calculated based on the measured efficiency, the branching fraction, the integrated luminosity and the $b\bar{b}$ production cross section and amounts to 2.5 ± 0.4 . As simulated $B^0 \rightarrow K^{*0} e^+ e^-$ decays show a broad distribution between 4500 and 5800 MeV/c^2 in the $m(K^+ \pi^+ \pi^- e^+ e^-)$ dimension, contributions in the signal area are negligible.

All introduced vetoes are summarised in Table 6.5. In conclusion, the majority of the studied background sources is excluded, partially by adding corresponding vetoes. Only non-resonant di-electron decays need to be respected in the fit model as these leak into the q^2 region below the J/ψ mass due to the loss of bremsstrahlung.

6.3.5 Machine learning based selection

In order to reduce contamination from combinatorial background a BDT [107] with gradient boosting [108] from the *LightGBM* implementation [109] is employed. Training a supervised machine learning algorithm technically means minimizing a function referred to as loss function that quantifies the agreement between the predictions made by the algorithm and the truth labels on the training data. Smaller function value typically means better agreement. Here, binary log loss, which is a special case of cross-entropy, is employed as it is the standard choice for binary classification in the *LightGBM* framework and known to yield good convergence.

One feature of *LightGBM* with regard to most other implementations is, that decision trees are not grown level-wise, but leaf-wise, such that not a maximum depth, but a maximum number of leafs is set as a hyper parameter. During training, the split that brings the largest loss reduction is chosen such that individual nodes are added to the decision tree instead of splitting all nodes in the lowest level and adding a complete new level of leafs.

Table 6.5: Summary of selection requirements applied to suppress physical backgrounds. The data samples the respective requirements are applied to are denoted by 1: resonant di-muon, 2: rare di-muon, 3: resonant di-electron and 4: rare di-electron.

Background source	Dimensions	Requirements	applied to
hadronic D decays	$m(K^\pm \pi^\mp)$ $m(K^\pm \ell^\mp \rightarrow \pi^\mp)$ $m(K^\pm \pi^\mp \ell^\mp \rightarrow \pi^\mp)$ $m(K^\pm \pi^\pm \pi^\mp \ell^\mp \rightarrow \pi^\mp)$	veto 1838–1898 MeV/ c^2	1,2,3,4
semileptonic D decays	$DLL_{e\pi}(e^+)$ $DLL_{e\pi}(e^-)$	> 3	3,4
lepton - hadron swap	$m(\ell^- K^+ \rightarrow \ell^+)$ $m(\ell^- \pi^+ \rightarrow \ell^+)$ $m(\ell^+ \pi^- \rightarrow \ell^-)$	veto 3047–3147 MeV/ c^2	2,4
Higher charmonia	$m(\pi^+ \pi^- J/\psi)$	veto 3848–3894 MeV/ c^2 veto 3663–3709 MeV/ c^2	1,3
$B^0 \rightarrow K^{*0} \ell^+ \ell^-$	$m(K^+ \pi^- J/\psi)$	veto 5175–5350 MeV/ c^2	1,2,3,4

To prevent any bias from training the classifier on the same events it is meant to predict, 10-fold cross validation is performed. This means that the training data are split into ten parts of equal size and each classifier is trained on nine parts and used to predict the remaining part. This is done in all ten possible permutations, such that an unbiased classifier for each part of the data is provided. The classifiers are trained on background from the upper B mass sideband and signal simulation. For the di-electron final state the signal proxy consists of 20421 and the background proxy of 7617 candidates. The corresponding sizes of the di-muon samples amount to 31355 and 36658. Before starting the training, the background and signal samples are balanced in terms of the sum of sample weights, by assigning a corresponding weight to each background candidate, while using the weight that corrects for differences between data and simulation (see section 6.2.1) for the signal sample.

The feature selection is optimised for $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$ decays by *recursive feature elimination*, which means that a classifier is trained on all possible feature candidates. After training a BDT the least important feature is dropped and a new classifier is trained. This procedure is repeated until only one feature remains. The whole procedure is performed ten times varying the splits among the ten folds and the random seed used for subsampling, which together with other hyper parameters is explained later in this section. That way the influence of statistical fluctuations on the feature selection is reduced. The separation power of each trained classifier is quantified by the ROC AUC. A BDT assigns a continuous value between zero (most likely background) and one (most likely signal) to each candidate. The ROC curve is obtained by plotting the false positive rate against the true positive rate obtained

when requiring various thresholds on the BDT response. In particle physics the ROC curve is often defined as true positive rate against one minus false positive rate as done in Figure 6.16. These two dimensions represent the signal efficiency and background rejection. The ROC AUC is defined as the area under the ROC curve and is equivalent for both types of ROC curves. A perfect separation between signal and background yields a ROC AUC of one, while a perfect random split results in a value of 0.5.

During the *recursive feature elimination* the ROC AUC of every trained classifier is stored and later the average between the ten iterations is calculated. The mean ROC AUC as a function of the number of training features is illustrated in Figure 6.12. The number of training features for the final classifier is chosen ad hoc to be the point where the average ROC AUC starts to decrease significantly, which here is 11. The feature ranking shown in Table 6.6 is created by calculating the mean rank of each feature candidate. The first eleven features are chosen for training the classifier for both the di-muon and the di-electron mode.

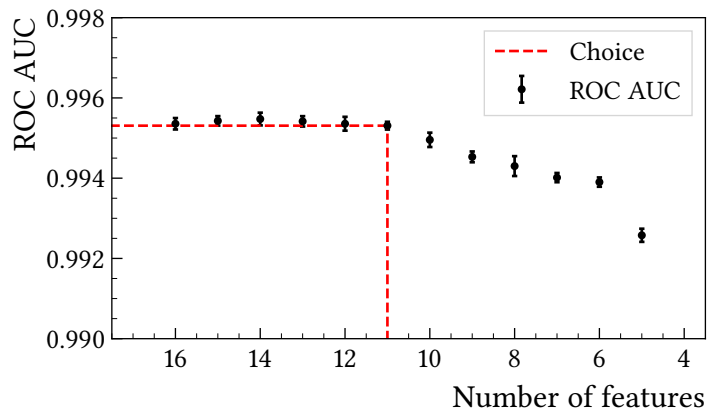


Figure 6.12: The mean ROC AUC as a function of the number of features the BDT is trained on. The number of features for the final BDT is chosen to be eleven as indicated by the red dashed line.

Table 6.6: Result of the recursive feature elimination. The first eleven features are used in the classifiers. For the definitions of the stated quantities see section 3.4.

Rank	Variable	Mean	Std Dev
1	$p_T(B^+)$	0.0	0.0
2	$\chi_{\text{FD}}^2(J/\psi, \text{PV})$	1.0	0.0
3	$\chi_{\text{vtx}}^2/\text{ndf}(B^+)$	2.0	0.0
4	$\chi_{\text{IP}}^2(B^+, \text{PV})$	3.0	0.0
5	$\chi_{\text{FD}}^2(K^+\pi^+\pi^-, \text{PV})$	4.0	0.0
6	ISO_2^{vtx}	5.0	0.0
7	$ \text{DIRA}(K^+\pi^+\pi^-, \text{PV}) $	6.3	0.5
8	$\max_{\text{DOCA}}(K^+, \pi^+, \pi^-, \mu^+, \mu^-)$	6.8	0.6
9	$\min_{\text{IP(PV)}}(K^+, \pi^+, \pi^-, \mu^+, \mu^-)$	8.1	0.3
10	$p(B^+)$	8.8	0.6
11	$\sum_{K^+, \pi^+, \pi^-, \ell^+, \ell^-} p_T$	10.1	0.3
12	ISO_1^{vtx}	11.4	0.9
13	$\chi_{\text{IP}}^2(\ell^+, \text{PV}) + \chi_{\text{IP}}^2(\ell^-, \text{PV})$	11.8	0.6
14	$\chi_{\text{vtx}}^2/\text{ndf}(K^+\pi^+\pi^-)$	13.0	0.6
15	$\text{FD}(J/\psi)$	13.7	0.6
16	$\chi_{\text{vtx}}^2/\text{ndf}(J/\psi)$	15.0	0.0

After choosing the final set of features, the classifier used to discriminate signal from background is trained. A pitfall with machine learning algorithms is that they can become sensitive to statistical fluctuations in the training data sample. This scenario is called overfitting and can be avoided for example by early stopping. This means that more trees are trained (more boosting rounds performed) until no further reduction of the loss function calculated on the validation sample is observed for the last 50 iterations. Hereby, the learning rate is fixed to 0.01. For the final prediction the iteration with the smallest loss value on the validation sample is used for each fold individually. As an additional measure against overfitting the complexity of the individual decision trees is limited by setting the maximum number of leaves to 31. This number would mean a maximum depth of four for a depth-wise tree. However, given the leaf-wise decision tree implemented in *LightGBM* a depth of 15 is possible. All other hyper parameters are optimised in multiple grid searches individually for each of the studied lepton flavours. The resulting optimal configurations are given in Table 6.7. The meanings of the considered parameters are:

Table 6.7: The hyper parameters that yield the optimal classification for the di-electron and di-muon sample.

Parameter	Value for e^+e^-	Value for $\mu^+\mu^-$
Data subsampling	0.95	1
Data subsampling freq.	6	-
Feature subsampling	0.95	1
Min child samples	500	160
Min child weight	4	9
Min split gain	0	0
l1	0.001	0
l2	0.001	0.01

- **Data subsampling:** Train each individual decision tree on a different subsample of the available training data. The subsample is redrawn in a certain frequency specified by an additional hyper parameter. This measure reduces the sensitivity to statistical fluctuation in an equivalent way as stochastic gradient decent. Here, only the di-electron selection has been found to benefit from data subsampling.
- **Feature subsampling:** Train each individual decision tree on a randomly chosen subsample of the available features. The numerical value represents the fraction of features that is present in each subsample. Again, only the di-electron classification benefits.

- **Min child samples and weight:** Required amount of candidates or sum of weights to form a leaf.
- **Min split gain:** Minimum reduction of the loss function required to split a leaf.
- **l1 and l2:** Parameters for L1 and L2 regularisation terms that are added to the loss function.

Measures against overfitting are found to have a larger impact on the classifier performance for the di-electron mode than for the di-muon mode. This is to be expected, due to the larger sizes of the training samples for the latter. In particular the background sample used for the di-muon classification is larger by almost a factor five.

The importance of each feature averaged among the ten folds is presented in Figure 6.13. The shown uncertainties correspond to the standard deviation. It can be seen that the importances are equal for the di-electron and di-muon sample. Although the central values suggest a different order of feature importances, these differences are insignificant due to the overlapping uncertainties. A comparison of the BDT response on training and validation data can be found in Figure 6.14. Especially for the di-electron mode the classifier performs better on the data it is trained on than on the validation sample. This is to be expected due to the smaller amount of data available for the training of the di-electron classifier. The same effect can be seen in the learning curves showing the ROC AUC as a function of the number of boosting rounds presented in Figure 6.15. However, based on the learning curves the presence of critical overfitting can be excluded. More precisely, the performance on the validation data increases continuously when performing more boosting rounds and asymptotically reaches a maximum, which triggers the early stopping. The overfitting does not reach a level where the performance measured on the validation sample starts decreasing with additional boosting rounds.

The capability of the final classifiers to separate signal and combinatorial background is quantified by the ROC AUC which amounts to 0.9931 for the di-electron and 0.9960 for the di-muon decay mode. The ROC curves are shown in Figure 6.16, where the area under the curve is highlighted by the shaded region. Both results represent excellent performance in solving the required classification task.

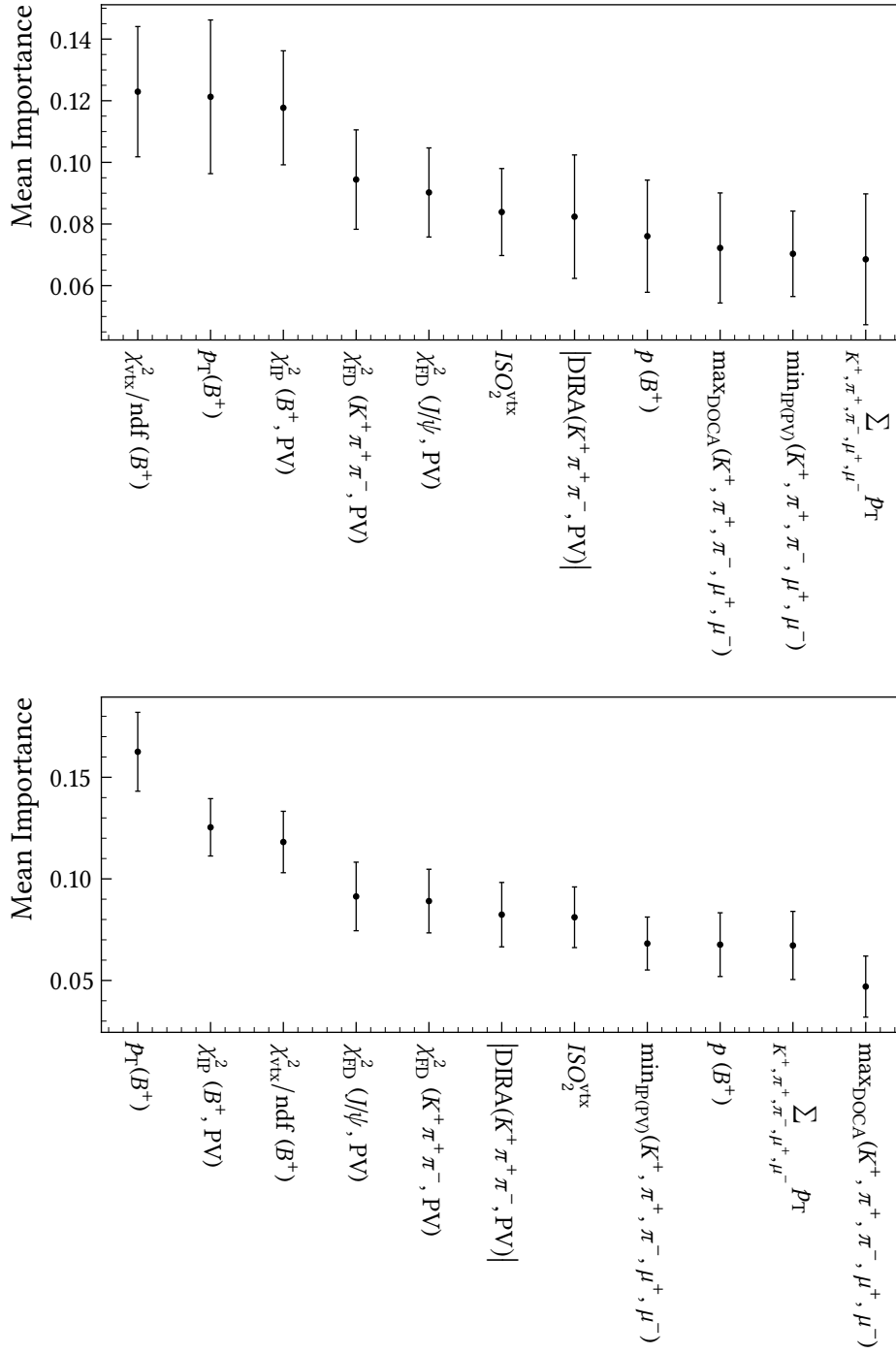


Figure 6.13: The mean of the feature importances of all ten classifiers for (top) the di-electron and (bottom) the di-muon decay. The importance is defined as the fraction of splits that exploit the regarding feature.

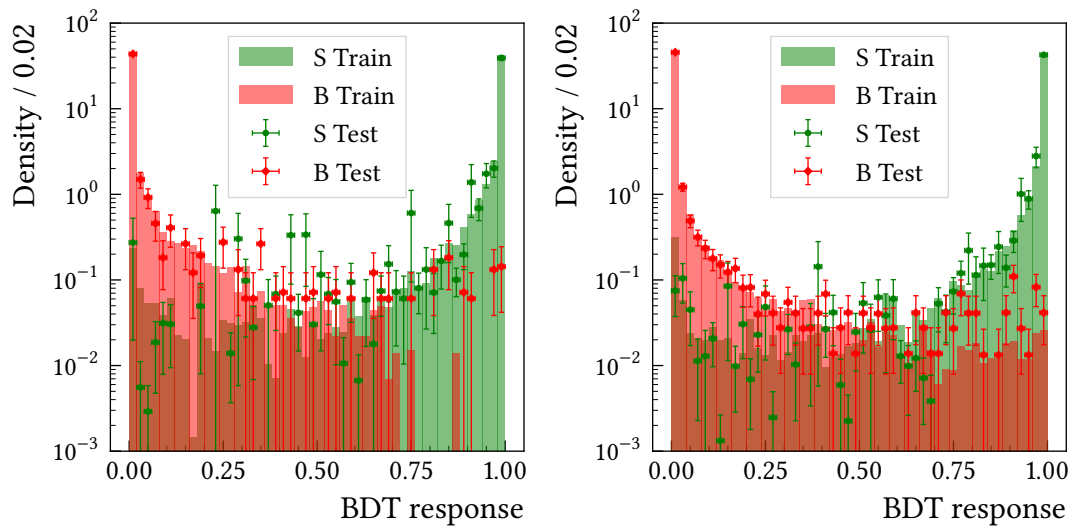


Figure 6.14: Comparison of the BDT response on the signal and background sample between the training and the validation data sample for one fold for (left) the di-electron and (right) the di-muon decay.

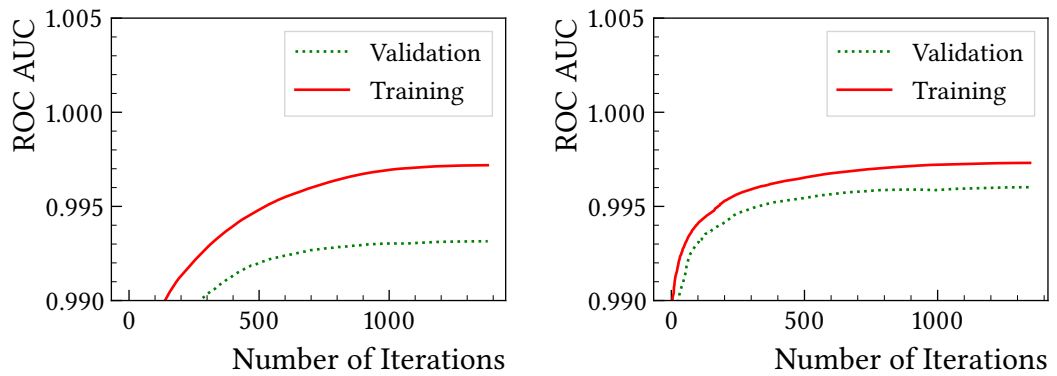


Figure 6.15: The ROC AUC the classifiers achieve on the training and validation sample as function of the number of trained decision trees for (left) the di-electron and (right) the di-muon decay.

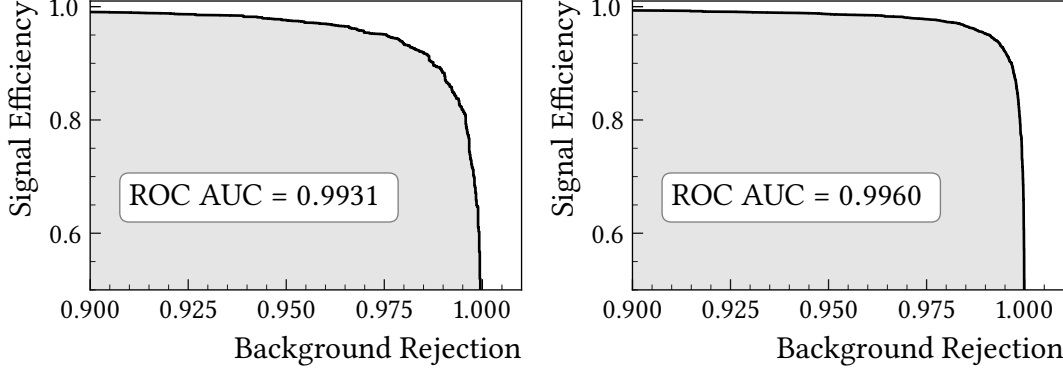


Figure 6.16: The ROC curve of all ten folds for (left) the di-electron and (right) the di-muon decay.

6.3.6 Selection optimisation

The final step of the selection consists of requirements on the BDT response and the PID quantities of the three different particles species that make up the respective final state. As introduced in section 6.2.2 for pions and leptons the minimum PID responses of the two oppositely charged candidates are considered, which is equivalent to applying the same requirement to both, but reduces the dimensionality of the optimisation problem that is detailed in the following. On top of that, composed PID quantities for the two different hadrons in the final state are employed.

The resulting four dimensional space, spanned by the BDT response, $\text{comp}_{\min} \text{PID}_{\pi}$, $\text{comp}_{\min} \text{PID}_K$ and min ProbNN_{ℓ} (in the following referred to as optimisation space), is scanned for the maximum of the figure of merit

$$\text{FoM} = \frac{S}{\sqrt{S+B}}, \quad (6.8)$$

where S and B are the expected signal and background yields. The former is calculated based on the integrated luminosity \mathcal{L}_{int} and the $b\bar{b}$ quark pair production cross section $\sigma_{b\bar{b}}$, which are different for the two centre-of-mass energies of the collisions delivered by the LHC during 2011 and 2012. Additionally, the hadronisation probability f_u , the branching ratio \mathcal{B} and the total efficiency to detect, reconstruct and select the regarding decay ε enter the equation resulting in

$$S = 2 \times \left(\sigma_{b\bar{b}}(7 \text{ TeV}) \times \mathcal{L}_{\text{int}}(7 \text{ TeV}) + \sigma_{b\bar{b}}(8 \text{ TeV}) \times \mathcal{L}_{\text{int}}(8 \text{ TeV}) \right) \times f_u \times \mathcal{B} \times \varepsilon. \quad (6.9)$$

Hereby, the measured branching ratio of the $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ decay is used also for the optimisation of the $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$ signal selection. Both, the b and the \bar{b} quark can hadronise into a B^{\pm} meson such that the factor of two is a consequence of the implication of charge conjugation. The numerical inputs are summarised in Table 6.8. The expected background yield B for every regarded point is measured by modelling the upper mass side band with an exponential function and calculating

Table 6.8: Numerical inputs used for the simultaneous optimisation of the BDT and PID selection.

Quantity	Value
$\sigma_{b\bar{b}}(7 \text{ TeV})$	$(288 \pm 48) \text{ mb}$ [79]
$\sigma_{b\bar{b}}(8 \text{ TeV})$	$(298 \pm 36) \text{ mb}$ [80]
f_u	$(40.1 \pm 0.8) \%$ [93]
$\mathcal{L}_{\text{int}}(7 \text{ TeV})$	$(0.9858 \pm 0.0017) \text{ fb}^{-1}$
$\mathcal{L}_{\text{int}}(8 \text{ TeV})$	$(1.9897 \pm 0.0017) \text{ fb}^{-1}$
$\mathcal{B}(B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-)$	$(4.36 \pm 0.40) \times 10^{-7}$ [58]

the integral over the signal range. This range is defined individually for the di-muon and the di-electron decay and corresponds to the smallest interval containing 95 % of the respective signal simulation. In order to be able to utilise existing numerical methods not the maximum of the FoM, but the minimum of the negative FoM is searched throughout the optimisation space. The simplex method [110] is used as it does not rely on the derivative of the target function, which for the figure of merit especially in the case of tight selections where the background sample size is small, is far from being continuous and therefore not easily calculable numerically. The same is true for the FoM itself. Therefore, the simplex method is ran with 50000 different randomly chosen starting points and at the end the result yielding the smallest negative FoM is chosen. Hereby the requirement on the BDT response is constrained to be between 0.2 and 1. Thereby already more than 95 % of the combinatorial background is excluded with a signal efficiency of 99 %, which makes the search for the optimum much more efficient.

The optimised selections for both final states are presented in Table 6.9 and yield a signal efficiency of 70.1 % (91.5 %) while removing 99.9 % (99.7 %) of the combinatorial background for the di-electron (di-muon) final state. During the optimisation the total efficiency, which according to the analysis strategy is not be investigated, needs to be calculated. This is done internally and only the optimal selection requirements are reported here.

Table 6.9: Simultaneously optimised BDT and PID requirements and expected signal and background yields in a mass interval corresponding to two standard deviations for $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$ and $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ decays.

Decay	BDT	\mathcal{PID}_π	\mathcal{PID}_K	\mathcal{PID}_ℓ
$B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$	0.9933	0.1596	0.0485	0.3929
$B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$	0.9523	0.1104	0.0136	0.2541

6.3.7 Low mass background

At the LHCb experiment one main challenge in studies of rare decays with electrons in the final state is to deal with low mass background. The signal distribution is significantly smeared out towards low mass because of the imperfect reconstruction of bremsstrahlung. Therefore, the estimation of the signal yield based on a fit to the invariant mass distribution of the reconstructed B meson suffers from the presence of background at the lower mass side of the signal peak.

Currently, work is ongoing to understand the exact composition of this background in order to derive efficient ways to suppress it and to develop an accurate fit model. However, the completion of these studies is beyond the scope of this thesis. Therefore, for the results presented in this document selection requirements are chosen ad hoc and the remaining low mass background is assumed to follow the same distribution in resonant and non-resonant di-electron data.

The low mass background is dominantly caused by partially reconstructed decays, which can be efficiently suppressed by the recently developed HOP method [111]. In order to illustrate this technique Figure 6.17 shows the topology of a decay and the variables relevant for the calculations that are detailed in the following. The final state is split into the electron part, denoted as X_e , and the hadronic part referred to as Y_h . For correctly reconstructed B decays the components of the momenta of X_e and Y_h perpendicular to the direction of flight (DoF) of the B meson are balanced. These components are denoted as p_{\perp} . The DoF is measured based on the position of the PV and the SV. Assuming that electrons emit bremsstrahlung in direction of flight not the total electron momentum, but at least its direction is measured correctly. Given this assumption the factor

$$\alpha_{\text{HOP}} = \frac{p_{\perp}(Y_h)}{p_{\perp}(X_e)}, \quad (6.10)$$

is multiplied to the di-electron momentum in order to correct for energy losses due to bremsstrahlung emission.

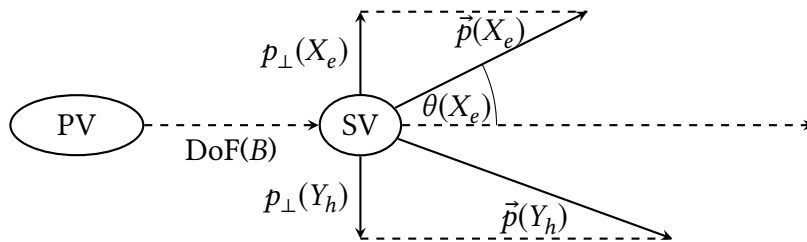


Figure 6.17: Illustration of the kinematics in a decay of the type $B \rightarrow X_e Y_h$. Sketch reproduced from [111]

Compared to the nominally measured B mass the corrected mass, m_{HOP} , shows a degraded resolution. This is mainly due to uncertainties in the measurement of the DoF propagating into the calculation of the angle between the electron system and

the DoF denoted as $\theta(X_e)$. These uncertainties originate from the measurement of the vertex position and depend on the distance between PV and SV. A quantity that accounts for both is the χ_{FD}^2 that was introduced in 3.4. Consequently, a selection requirement in the two dimensional plane spanned by m_{HOP} and $\chi_{\text{FD}}^2(B^+)$ is employed. Similarly to the measurement of R_{K^*0} [17], a cut of the type

$$m_{\text{HOP}} > c + b \times \ln \chi_{\text{FD}}^2(B^+) \quad (6.11)$$

is applied. The selection is optimised using a background sample taken from resonant di-electron data. This sample consists of candidates that pass the complete selection detailed so far and satisfy $m(K^+\pi^+\pi^-J/\psi) < 5175 \text{ MeV}/c^2$. The separation power is probed by creating a ROC curve, presented in Figure 6.18, that shows the background rejection as a function of the signal efficiency. The latter is determined on the corresponding simulated sample. When creating the ROC curve, c is defined based on the interval containing the desired fraction of simulated signal candidates while b is determined by scanning all integer values between zero and 200 for the largest rejection of low mass background. The dependence of the background rejection on the parameter b for the chosen requirement as well the ROC curve are illustrated in Figure 6.18.

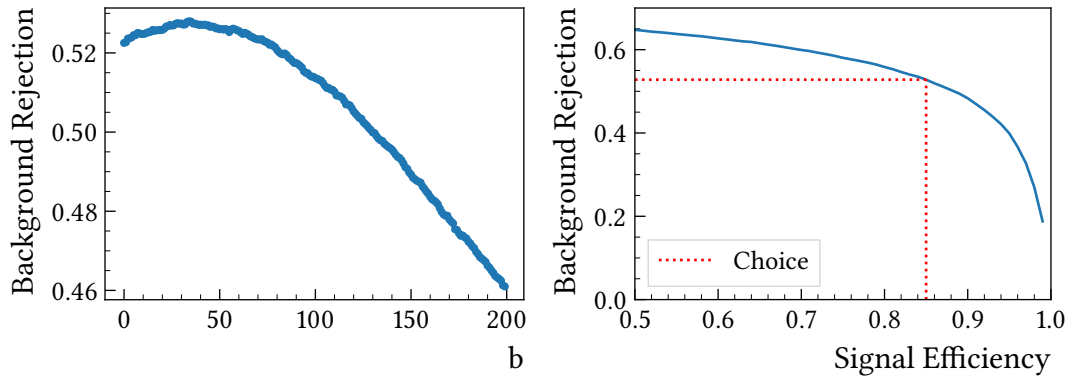


Figure 6.18: Background rejection as a function of b for 85 % signal efficiency (left) and the ROC curve of the HOP requirement along with the chosen selection (right).

The selection requirement

$$m_{\text{HOP}} > 4660 + 34 \times \ln \chi_{\text{FD}}^2(B^+), \quad (6.12)$$

is chosen ad hoc based on the slope of the ROC curve. It corresponds to a signal efficiency of 85 % and rejects about 53 % of the low mass background. The requirement and the plane spanned by m_{HOP} and $\ln \chi_{\text{FD}}^2(B^+)$ for simulated signal decays and for the background sample are illustrated in Figure 6.19.

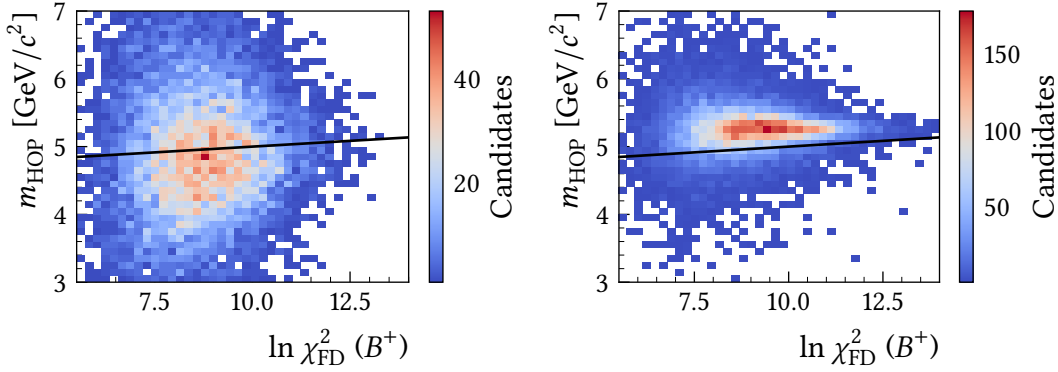


Figure 6.19: Plane spanned by the HOP mass and the logarithm of the flight distance significance in low mass background (left) and in simulated non-resonant decays (right). The straight line represents the applied selection requirement.

Additionally, candidates that pass the non-resonant selection and lie below $5050 \text{ MeV}/c^2$ in $m(K^+ \pi^+ \pi^- e^+ e^-)$ are probed for contributions from misidentified particles. The threshold is chosen in order to be sensitive to the low mass background close to the non-resonant signal. The value corresponds to the lower border of the 68 % interval of the mass distribution, which is determined from simulated data. Two dimensional histograms featuring the neural network based PID variable that quantifies the agreement for the given track to match the desired identity on the x -axis and that for several undesired identities on the y -axis are regarded. Clear signs of kaon and pion candidates that are misidentified electrons are found as can be seen in Figure 6.20. These are vetoed by requiring $\text{ProbNN}_e < 0.2$ for all kaon and pion candidates. The cut is chosen ad hoc and gives a good suppression of misidentified electrons while retaining 95 % of the signal when applied to all three final state hadron candidates.

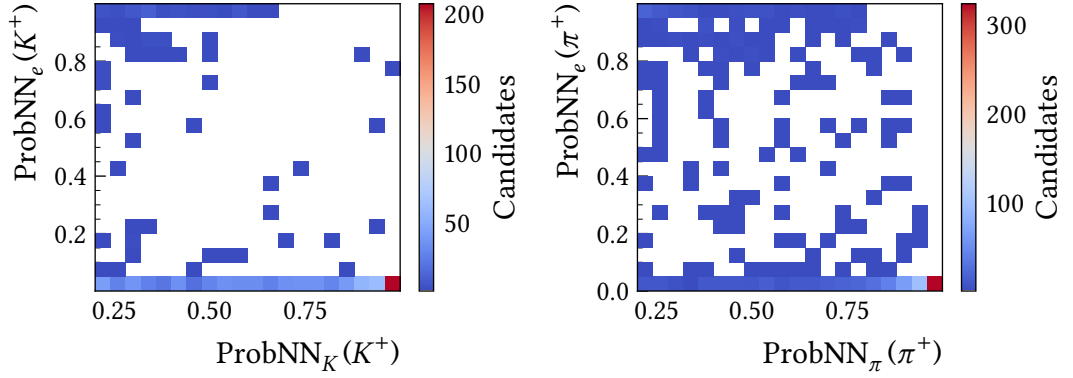


Figure 6.20: Histograms of particle identification classifier responses for kaon and pion candidates. On the x -axis the classifier for the desired identity is plotted, while the y -axis features the compatibility with being a misidentified electron. The shown data correspond to the full Run 1 data sample passing the non-resonant di-electron selection discussed so far.

6.4 Signal yields

The signal yields are determined by simultaneous maximum likelihood fits to the invariant mass distribution of the reconstructed B^+ mesons. Hereby, the shape parameters are partially shared between resonant and rare mode respectively for the two lepton flavours. In contrast to the previously described fits, here the B mass calculated without J/ψ mass constraint is parametrised, as the rare decays do not proceed via an intermediate J/ψ meson. For the di-muon data the mass resolution is sufficient to exclude low mass background by setting the lower bound of the parametrised mass range to $5175 \text{ MeV}/c^2$. Therefore, the fit model consists of only two shapes: An exponential function to account for combinatorial background and a Hypatia function that models the signal component. At first, the model is fitted simultaneously to simulated resonant and non-resonant data to determine the tail parameters $\alpha_1, \alpha_2, n_1, n_2$ and the shape parameter ζ . These are fixed in a subsequent simultaneous fit to measured data. In both fits the mean, the width and ζ are shared between resonant and non-resonant mode. The first two are floating in the fit to measured data, as they are known to be only imperfectly simulated. The signal yields are determined to be

$$N_{K^+\pi^+\pi^-\mu^+\mu^-}^{\text{rare}} = 185 \pm 20 \quad (6.13)$$

$$N_{K^+\pi^+\pi^-\mu^+\mu^-}^{\text{resonant}} = 72\,060 \pm 310. \quad (6.14)$$

The fit to both dimuon modes along with the data is visualised in Figure 6.21.

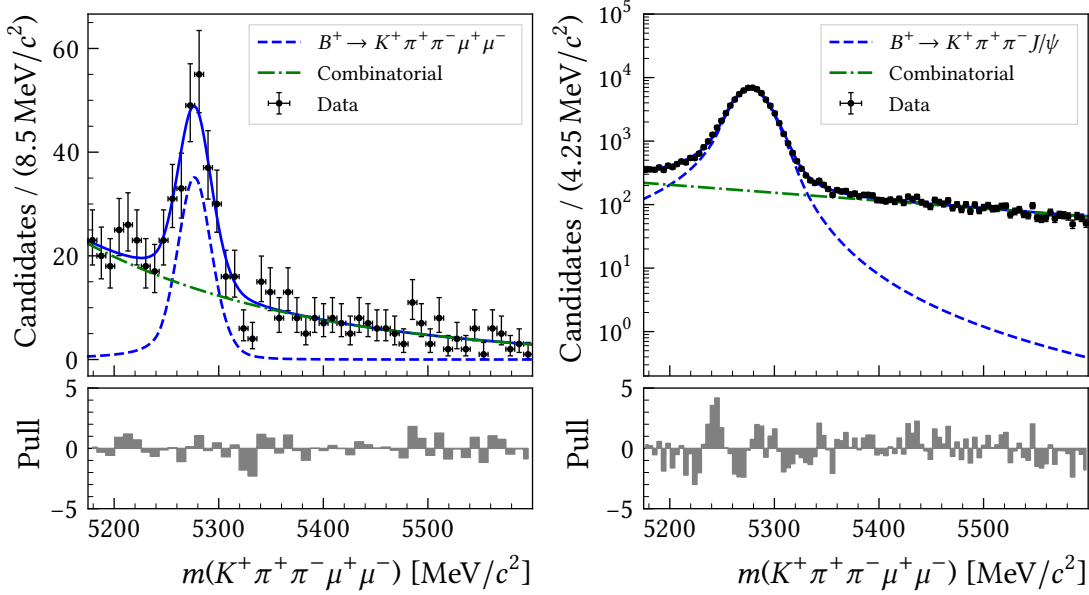


Figure 6.21: The full LHCb Run 1 data sample after applying the (left) non-resonant and (right) resonant selection. Shown is the invariant mass of the $K^+ \pi^+ \pi^- \mu^+ \mu^-$ system parametrised by PDFs as detailed in the text.

For the di-electron modes imperfect reconstruction of bremsstrahlung causes the signal distribution to have a significant tail towards low mass. Consequently, it is not feasible to exclude low mass background by a corresponding choice of the fit range. Hence, an additional shape to account for these background candidates needs to be added to the fit model. In order to determine this shape, candidates that pass the full selection for the resonant di-electron mode and are situated at $m(K^+ \pi^+ \pi^- J/\psi) < 5175 \text{ MeV}/c^2$ are selected and regarded in the $m(K^+ \pi^+ \pi^- e^+ e^-)$ dimension. The distribution is modelled by a Gaussian function that at a given point transitions into an exponential function. The resulting shape is fixed and used to parametrise partially reconstructed background in the non-resonant mode. For the resonant mode, only data with $m(K^+ \pi^+ \pi^- J/\psi) > 5175 \text{ MeV}/c^2$ is considered, such that partially reconstructed decays are excluded.

Another source of background that is accounted for in the fit model for non-resonant di-electron data originates from resonant decays that pass the respective vetoes because of emitting large amounts of energy in the form of bremsstrahlung that is not completely recovered by the detector. Therefore, in the B^+ mass distribution, these candidates interfere with the left handed tail of the signal distribution. The shape is modelled by a KDE applied to simulated resonant di-electron decays that pass the non-resonant selection. The yield of this background is constrained by a Gaussian function, with mean calculated as $f \times N_{K^+ \pi^+ \pi^- e^+ e^-}^{\text{resonant}}$. Here, f is determined based on simulated data as

$$f = \frac{w_{\text{resonant} \rightarrow \text{non-resonant}}}{w_{\text{resonant} \rightarrow \text{resonant}}}, \quad (6.15)$$

with $w_{\text{resonant} \rightarrow \text{non-resonant}}$ being the sum of simulation correction weights (see section 6.2.1) of the simulated resonant di-electron decays that pass the full non-resonant selection, and $w_{\text{resonant} \rightarrow \text{resonant}}$ the analogous quantity after applying the resonant selection. The standard deviation of the Gaussian function is set to the uncertainty originating from the finite size of the simulated data sample. This approach is adapted from the recent measurement of R_K [16].

The parametrisations of both background sources for the two individual trigger categories are presented in Appendix A.3.

Combinatorial background is parametrised by an exponential function individually in resonant and the non-resonant data. For each of the three bremsstrahlung categories a different signal shape is employed. If none of the two electron candidates was assigned a bremsstrahlung correction, a Crystal Ball function [112] with tail towards lower mass is used. In the other two categories wrongly assigned bremsstrahlung corrections cause an additional tail towards higher mass. Therefore, these signal components are each parametrised by one Hypatia function. At first, a simultaneous fit to resonant and non-resonant simulation is performed to determine the parameters that describe the tails of the signal distributions of all three categories, which are not shared between the resonant and rare mode. Shared are the mean, the width and for the Hypatia functions also the parameter ζ . In the fit to measured data, all tail parameters and additionally the parameter ζ are fixed to the values determined from simulated data. Moreover, the difference between the means of the Crystal Ball and the two Hypatia functions are fixed to the distances observed in simulated data. The fraction of signal candidates described by each of the three shapes is feature to a Gaussian constraint. Hereby, the central value is set to the relative number of simulated candidates that make up the respective category. The width is fixed to the uncertainty originating from the limited size of the simulated sample.

The $m(K^+ \pi^+ \pi^- e^+ e^-)$ distributions modelled by the described PDF are illustrated in Figure 6.22 for the ETOS trigger category and in Figure 6.23 for the TIS category. The measured signal yields in ETOS amount to

$$N_{K^+ \pi^+ \pi^- e^+ e^-}^{\text{rare,ETOS}} = 66 \pm 11 \quad (6.16)$$

$$N_{K^+ \pi^+ \pi^- e^+ e^-}^{\text{resonant,ETOS}} = 6939 \pm 83 \quad (6.17)$$

and in TIS to

$$N_{K^+ \pi^+ \pi^- e^+ e^-}^{\text{rare,TIS}} = 41.0 \pm 8.5 \quad (6.18)$$

$$N_{K^+ \pi^+ \pi^- e^+ e^-}^{\text{resonant,TIS}} = 1883 \pm 43. \quad (6.19)$$

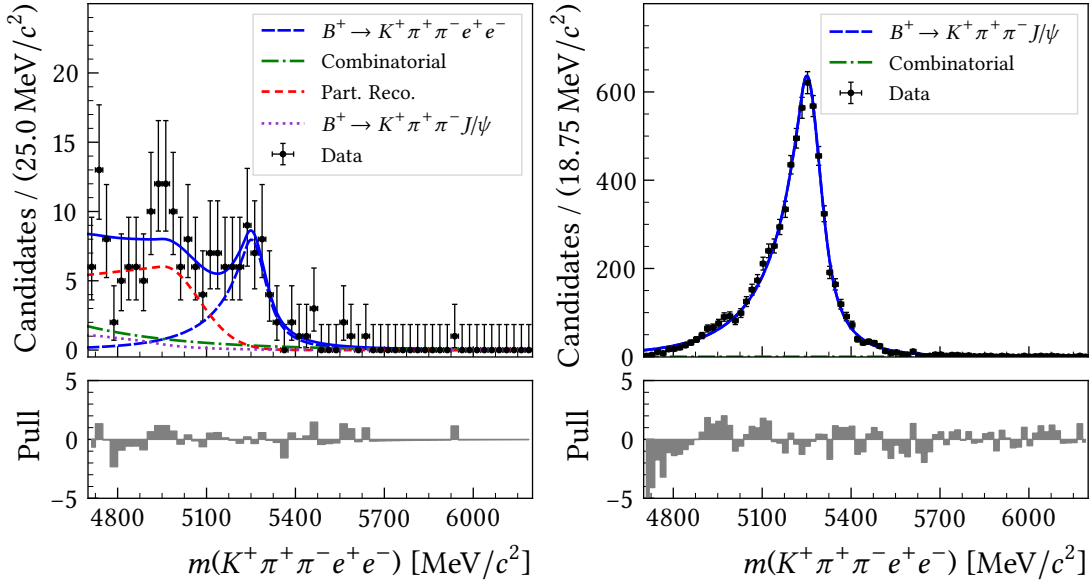


Figure 6.22: The full LHCb Run 1 data sample after applying the (left) non-resonant and (right) resonant selection. Shown is the invariant mass of the $K^+ \pi^+ \pi^- e^+ e^-$ system of candidates that belong to the ETOS category parametrised by PDFs as detailed in the text.

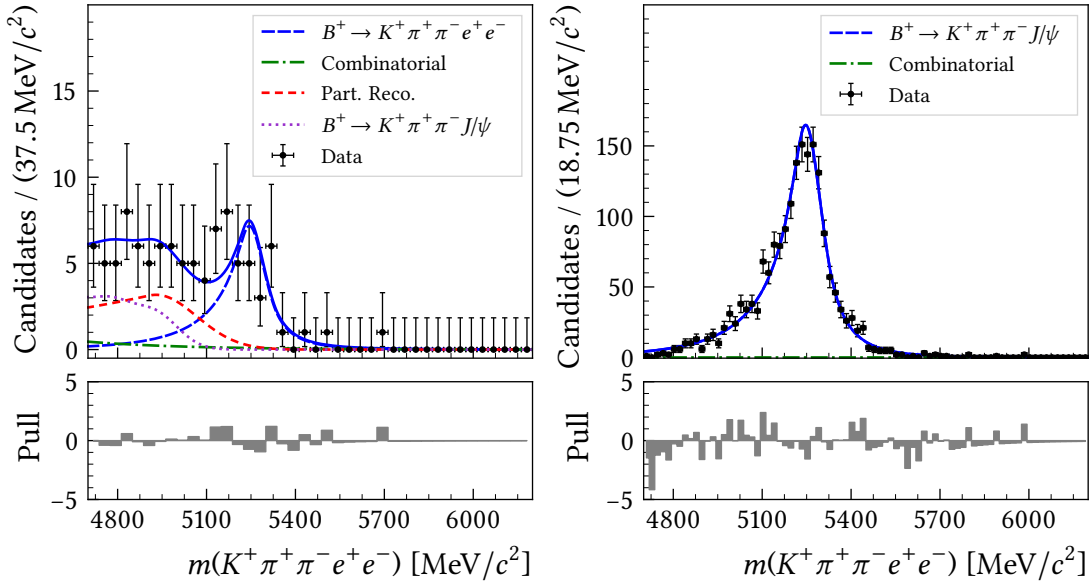


Figure 6.23: The full LHCb Run 1 data sample after applying the (left) non-resonant and (right) resonant selection. Shown is the invariant mass of the $K^+ \pi^+ \pi^- e^+ e^-$ system of candidates that belong to the TIS category parametrised by PDFs as detailed in the text.

The fit strategy is not yet finalised such that all presented results are preliminary. The ratio between the signal yields of resonant and non-resonant di-electron decays differ by three standard deviations between the two trigger categories. This indicates that further investigations concerning the determination of the signal yields is needed. The region above the signal peak in the $K^+\pi^+\pi^-e^+e^-$ invariant mass distribution is approximately flat and contains only a small fraction of the sample, which is comparable to the outcome of similar measurements performed by the LHCb collaboration [15–17]. However, the region below the signal peak is densely populated and needs to be parametrised accurately. Therefore, studies that go beyond the ad hoc treatment discussed in this thesis are currently developed.

Based on the presented fit results the part of the LFU ratio that consists of the signal yields is determined as

$$R_{K\pi\pi}^{\text{yields}} = 0.212 \pm 0.036. \quad (6.20)$$

So far no systematic uncertainties are evaluated, such that the uncertainty of 17 % is purely statistical. For the final goal of this measurement, the determination of $R_{K\pi\pi}$, the ratio of efficiencies still needs to be determined. This quantity is still blind and will not be calculated before the analysis passed the quality assurance measures of the LHCb collaboration. However, based on the signal yields an estimate on the achievable sensitivity of this measurement can be calculated. In comparison to the statistical uncertainties of R_K and R_{K^*0} reported for Run 1 data in the q^2 region 1.1 to 6 GeV^2/c^4 , which are 12 % [16] and 16 % [17], the measurement of $R_{K\pi\pi}$ can be a valuable addition to the general picture of lepton flavour universality.

6.5 Estimation of the achievable sensitivity

With the end of LHC Run 2 the LHCb collaboration has collected a data sample corresponding to an integrated luminosity of 9 fb^{-1} . As the production cross section of $b\bar{b}$ quark pairs increases with the centre-of-mass energy of the proton-proton collisions, the number of recorded B meson decays during Run 2 is not two, but almost four times larger than that in the Run 1 data sample on which the previously discussed $R_{K\pi\pi}^{\text{yields}}$ is determined.

Table 6.10: Data taking periods, centre-of-mass energy of the proton-proton collisions (\sqrt{S}), recorded integrated luminosity (\mathcal{L}_{int}), $b\bar{b}$ quark pair production cross section ($\sigma_{b\bar{b}}$) and corresponding number of $b\bar{b}$ quark pairs ($N(b\bar{b})$).

Year	\sqrt{S} [TeV]	\mathcal{L}_{int} [fb^{-1}]	$\sigma_{b\bar{b}}$ [mb]	$N(b\bar{b})$ [10^9]
2011	7	1	288 [79]	288
2012	8	2	298 [80]	596
2015 - 2018	13	6	560 [113]	3360

In order to estimate the sensitivity that can be achieved when extending the measurement to all available data pseudo experiments are carried out. Therefore, pseudo data samples are generated based on the PDFs that are fitted to the Run 1 data sample. The measured signal and background yields are scaled according to the integrated luminosities and $b\bar{b}$ production cross sections presented in Table 6.10 by the factor

$$f = 1 + \frac{\mathcal{L}_{\text{int}}(13 \text{ TeV}) \times \sigma(13 \text{ TeV})}{\mathcal{L}_{\text{int}}(7 \text{ TeV}) \times \sigma(7 \text{ TeV}) + \mathcal{L}_{\text{int}}(8 \text{ TeV}) \times \sigma(8 \text{ TeV})} \quad (6.21)$$

$$\approx 4.8$$

to represent the size of the total available data sample that corresponds to 9 fb^{-1} .

The fit strategy is left unchanged with respect to that discussed in section 6.4. In total 15 000 pseudo experiments are conducted, 5000 for the di-muon mode and the same number for each of the two trigger categories in di-electron mode.

In order to probe the correctness of the fit procedure, pulls of the signal yields are calculated. The pull of a variable x is defined as

$$\text{Pull}(x) = \frac{x_{\text{fit}} - x_{\text{generated}}}{\sigma(x_{\text{fit}})}, \quad (6.22)$$

where x_{fit} and $\sigma(x_{\text{fit}})$ are the estimate of parameter x and its uncertainty determined by the fit and $x_{\text{generated}}$ is the value the pseudo experiment is generated with. A Pull distribution is expected to follow a standard normal distribution if the parameter estimation is unbiased and the uncertainty is correctly estimated. For a biased measurement, the mean of the pull distribution deviates from zero. If the uncertainties are (underestimated) overestimated by the fit, the width of the pull distribution is (larger) smaller than one. The respective results obtained from the pseudo experiments are presented in Figure 6.24. A Gaussian function is fitted to each of the three pull distributions and the parameters are presented in the respective figure. For the rare di-muon mode the signal yield estimated by the extended maximum likelihood fit is unbiased and the corresponding uncertainty is correctly estimated. However, in the case of the rare di-electron modes a bias in the TIS category is observed. In both categories the uncertainty is overestimated. These results indicate once more, that further studies concerning the parametrisation of the $K^+ \pi^+ \pi^- e^+ e^-$ mass distributions are needed.

The average relative uncertainty on the parameter $R_{K\pi\pi}^{\text{yields}}$ obtained by the pseudo experiments that correspond to the joined Run 1 and Run 2 data samples is 7.8 %. The same precision is expected under the assumption that the statistical uncertainty is of Poissonian nature. The sample generated for the pseudo experiments is 4.8 times as large as that used to determine $R_{K\pi\pi}^{\text{yields}}$ with a precision of 17 %. Scaling the uncertainty corresponding to Poissonian statistics results in

$$\frac{17 \%}{\sqrt{4.8}} \approx 7.8 \%. \quad (6.23)$$

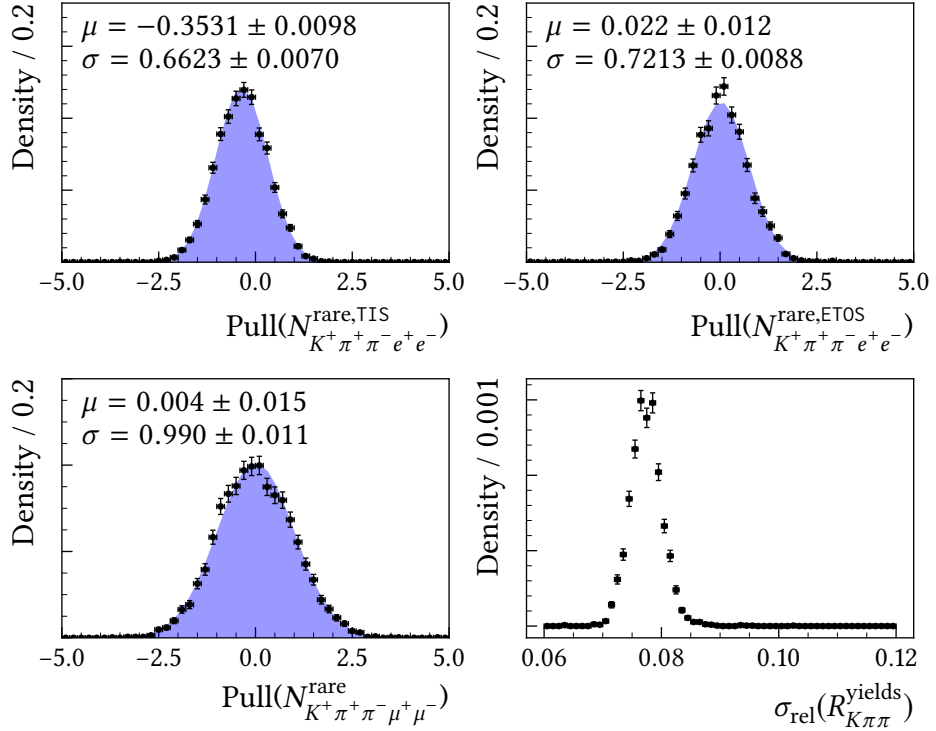


Figure 6.24: Distributions obtained from pseudo experiments. Pull distributions of the signal yields of the non-resonant di-electron mode in the TIS category (top left), the ETOS category (top right) as well as the rare di-muon mode (bottom left). The bottom right distribution shows the relative uncertainty of the parameter $R_{K\pi\pi}^{\text{yields}}$.

The systematic uncertainty on $R_{K^{*0}}$ measured by the LHCb collaboration amounts to 7.2% [17] and is close to the expected statistical uncertainty of this measurement. The main contribution with 5% is due to resonant di-electron decays that pass the corresponding veto due to a swap of a hadron and an electron candidate. An analogous background is studied in section 6.3.4 and is shown to populate an $K^+\pi^+\pi^-e^+e^-$ invariant mass region below the signal. This again underlines the importance of understanding the composition of the lower mass background to reduce the systematic uncertainty such that the full statistical potential of the data sample can be exploited.

7 Conclusion

In this thesis two searches for effects beyond the SM based on proton-proton collision data corresponding to an integrated luminosity of 3 fb^{-1} recorded by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV are presented.

In a search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays no signal candidates are observed, such that upper limits on the branching fractions are determined. These improve the results of the only previous experimental study of these decays by factors of 6.4 and 9.5 for the decay of the B_s^0 and B^0 meson, respectively. Additionally, the sensitivity to these decays proceeding through BSM scalar and pseudoscalar particles is probed and upper limits on the corresponding branching fractions are set. Compared to the previous analysis, these are smaller by a factor of 7.3 and 10.5. An improvement of a factor of $\sqrt{3} \approx 1.7$ is expected due to the increased sample size when assuming a Poissonian distribution. The remaining reduction is due to switching to a machine learning based selection and a more optimal choice of the normalisation mode. The measured upper limits at 95 % confidence level are

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) < 6.9 \times 10^{-10}, \quad (7.1)$$

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-) < 2.5 \times 10^{-9}, \quad (7.2)$$

$$\mathcal{B}(B^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) < 6.0 \times 10^{-10}, \quad (7.3)$$

$$\mathcal{B}(B_s^0 \rightarrow S(\rightarrow \mu^+ \mu^-)P(\rightarrow \mu^+ \mu^-)) < 2.2 \times 10^{-9}. \quad (7.4)$$

The anomaly reported by the HyperCP collaboration [35], which was one of the main motivations for this study, could not be confirmed by the LHCb collaboration. After the measurement presented in this thesis was published, the LHCb collaboration reported first evidence for the $\Sigma^+ \rightarrow p \mu^+ \mu^-$ decay and measured the branching fraction to be compatible with the SM prediction and the HyperCP result. However, the di-muon invariant mass distributions did not display any feature that would hint towards an intermediate resonance [103].

The determined limits are dominantly driven by the sample size, which on inclusion of the data recorded during LHC Run 2 increases by a factor of 4.8. Assuming the same efficiencies for Run 1 and Run 2 data the upper limit of the branching fractions can be assumed to decrease by the factor $\sqrt{4.8} \approx 2.2$ if the measurement is repeated based on all currently available data. Additionally, the analysis strategy leaves room for improvements. In particular vetoing certain intervals in the invariant mass of di-muon systems limits the sensitivity to BSM particles at these particular masses. Therefore, alternative ways to reject contributions from resonant decays have been proposed [114].

The presented measurement of the LFU ratio $R_{K\pi\pi}$ is in an advanced state. A selection based on a machine learning classifier and several PID quantities is optimised. Furthermore, extensive studies concerning contributions from different background processes are performed. The requirements applied to suppress low mass background as well as the parametrisation of the invariant $K^+\pi^+\pi^-e^+e^-$ mass distribution are still under development. In particular, the composition of the low mass background needs to be understood more thoroughly. Therefore, the presented results must be considered preliminary.

The statistical uncertainty of $R_{K\pi\pi}$ measured from the Run 1 data sample is determined to 17 % and is comparable to that of the R_{K^*0} measurement [17]. By extending the analysis to the full Run 2 data sample this uncertainty can be further reduced to 8 %. This shows that in the future $R_{K\pi\pi}$ can be a valuable observable to further clarify the exciting situation in LFU tests. On top of that, on a longer timescale and with even larger samples studies of the composition of the $K^+\pi^+\pi^-$ spectrum can help to determine the chiral structure of a potential BSM effect. For the finalisation of the measurement the determination of the efficiencies and studies concerning systematic uncertainties are still ongoing. Additionally, the whole developed analysis strategy needs to be applied to and validated for Run 2 data.

All in all, exciting times lie ahead in flavour physics. For many of the measurements of $b \rightarrow s\ell^+\ell^-$ processes that show anomalies as discussed in section 2.3 updates with LHCb Run 2 data are in preparation. In the meantime, the Belle II experiment started data taking [115] raising the anticipation to perform complementary studies soon.

Bibliography

- [1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.
- [2] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.
- [3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [4] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1.
- [5] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30.
- [6] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [7] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.
- [8] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [9] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and Antimatter in the Universe*, New J. Phys. **14** (2012) 095012.
- [10] V. Trimble, *Existence and Nature of Dark Matter in the Universe*, Ann. Rev. Astron. Astrophys. **25** (1987) 425.
- [11] M. S. Turner and D. Huterer, *Cosmic Acceleration, Dark Energy and Fundamental Physics*, J. Phys. Soc. Jap. **76** (2007) 111015.
- [12] Planck, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, arXiv:1807.06209.
- [13] A. Buras, *Flavour Expedition to the Zeptouniverse*, PoS **FWNP** (2015) 003.
- [14] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104.

- [15] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **113** 151601.
- [16] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, arXiv:1903.09252.
- [17] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, JHEP **08** (2017) 055.
- [18] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , JHEP **09** (2015) 179.
- [19] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays*, JHEP **06** (2014) 133.
- [20] J. Aebischer *et al.*, *B-decay discrepancies after Moriond 2019*, arXiv:1903.10434.
- [21] M. Algueró *et al.*, *Addendum: "Patterns of New Physics in $b \rightarrow s \ell^+ \ell^-$ transitions in the light of recent data" and "Are we overlooking Lepton Flavour Universal New Physics in $b \rightarrow s \ell \ell$?"*, arXiv:1903.09578.
- [22] A. Arbey *et al.*, *Update on the $b \rightarrow s$ anomalies*, arXiv:1904.08399.
- [23] D. Griffiths, *Introduction to elementary particles*, Wiley-VCH Verlag GmbH und Co. KGaA, Weinheim, Germany, 7 ed., 2008.
- [24] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons, New York, USA, 1984.
- [25] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) .
- [26] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$* , arXiv:1904.03947.
- [27] C. Burgard, *Sketch of elementary particles*, <http://www.texample.net/tikz/examples/model-physics/>. Accessed: 2019-03-05.
- [28] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. D **98** (2018) 030001.
- [29] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285.
- [30] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.
- [31] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

- [32] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125.
- [33] T. Blake, G. Lanfranchi, and D. M. Straub, *Rare B Decays as Tests of the Standard Model*, Prog. Part. Nucl. Phys. **92** (2017) 50.
- [34] B. Döbrich, F. Ertas, F. Kahlhoefer, and T. Spadaro, *Model-independent bounds on light pseudoscalars from rare B-meson decays*, Phys. Lett. **B790** (2019) 537.
- [35] HyperCP collaboration, H. Park *et al.*, *Evidence for the decay $\Sigma^+ \rightarrow p\mu^+\mu^-$* , Phys. Rev. Lett. **94** (2005) 021801.
- [36] S. V. Demidov and D. S. Gorbunov, *More about sgoldstino interpretation of HyperCP events*, JETP Lett. **84** (2007) 479.
- [37] S. V. Demidov and D. S. Gorbunov, *Flavor violating processes with sgoldstino pair production*, Phys. Rev. **D85** (2012) 077701.
- [38] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.
- [39] Y. Dincer and L. M. Sehgal, *Electroweak effects in the double Dalitz decay $B_s \rightarrow l^+l^-l'^+l'^-$* , Phys. Lett. **B556** (2003) 169.
- [40] A. V. Danilina and N. V. Nikitin, *Four-Leptonic Decays of Charged and Neutral B Mesons within the Standard Model*, Phys. Atom. Nucl. **81** (2018), no. 3 347, [Yad. Fiz.81,no.3,331(2018)].
- [41] LHCb collaboration, R. Aaij *et al.*, *Search for rare $B_{(s)}^0 \rightarrow \mu^+\mu^-\mu^+\mu^-$ decays*, Phys. Rev. Lett. **110** (2013) 211801.
- [42] G. Hiller and F. Kruger, *More model-independent analysis of $b \rightarrow s$ processes*, Phys. Rev. **D69** (2004) 074020.
- [43] M. Bordone, G. Isidori, and A. Pattori, *On the Standard Model predictions for R_K and R_{K^*}* , Eur. Phys. J. **C76** (2016), no. 8 440.
- [44] Belle, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for $B \rightarrow K^{(*)}l^+l^-$* , Phys. Rev. Lett. **103** (2009) 171801.
- [45] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \rightarrow K^{(*)}l^+l^-$* , Phys. Rev. **D86** (2012) 032012.
- [46] HPQCD, C. Bouchard *et al.*, *Standard Model Predictions for $B \rightarrow Kl^+l^-$ with Form Factors from Lattice QCD*, Phys. Rev. Lett. **111** (2013), no. 16 162002, [Erratum: Phys. Rev. Lett.112,no.14,149902(2014)].

- [47] C. Bobeth, G. Hiller, and G. Piranishvili, *Angular distributions of $\bar{B} \rightarrow \bar{K} \ell^+ \ell^-$ decays*, JHEP **12** (2007) 040.
- [48] Belle, A. Abdesselam *et al.*, *Test of lepton flavor universality in $B \rightarrow K^* \ell^+ \ell^-$ decays at Belle*, arXiv:1904.02440.
- [49] W. Altmannshofer and D. M. Straub, *New physics in $b \rightarrow s$ transitions after LHC run 1*, Eur. Phys. J. **C75** (2015), no. 8 382.
- [50] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \rightarrow V \ell^+ \ell^-$ in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098.
- [51] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Calculation of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi \mu^+ \mu^-$ observables using form factors from lattice QCD*, Phys. Rev. Lett. **112** (2014) 212003.
- [52] C. Bobeth, G. Hiller, and D. van Dyk, *More Benefits of Semileptonic Rare B Decays at Low Recoil: CP Violation*, JHEP **07** (2011) 067.
- [53] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, *The Decay $B \rightarrow K \ell^+ \ell^-$ at Low Hadronic Recoil and Model-Independent $\Delta B = 1$ Constraints*, JHEP **01** (2012) 107.
- [54] G. Hiller and M. Schmaltz, *R_K and future $b \rightarrow s \ell \ell$ physics beyond the standard model opportunities*, Phys. Rev. **D90** (2014) 054014.
- [55] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, Phys. Rev. **D89** (2014) 095033.
- [56] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \rightarrow s \ell \ell$* , JHEP **02** (2015) 055.
- [57] Belle, H. Guler *et al.*, *Study of the $K^+ \pi^+ \pi^-$ Final State in $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$ and $B^+ \rightarrow \psi - \text{prime} K^+ \pi^+ \pi^-$* , Phys. Rev. **D83** (2011) 032005.
- [58] LHCb collaboration, R. Aaij *et al.*, *First observations of the rare decays $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$* , JHEP **10** (2014) 064.
- [59] H. Hatanaka and K.-C. Yang, *$K(1)(1270)$ - $K(1)(1400)$ Mixing Angle and New-Physics Effects in $B \rightarrow K(1) l^+ l^-$ Decays*, Phys. Rev. **D78** (2008) 074007.
- [60] L. Evans and P. Bryant, *Lhc machine*, Journal of Instrumentation **3** (2008), no. 08 S08001.
- [61] F. Bordry *et al.*, *The First Long Shutdown (LS1) for the LHC*, in *Proceedings, 4th International Particle Accelerator Conference (IPAC 2013): Shanghai, China, May 12-17, 2013*, p. MOZB202, 2013.
- [62] LHCb Collaboration, A. A. Alves Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008), no. LHCb-DP-2008-001. CERN-LHCb-DP-2008-001 S08005.

- [63] ALICE, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [64] ATLAS, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [65] CMS, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [66] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007.
- [67] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002.
- [68] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022.
- [69] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431.
- [70] LHCb collaboration, P. Perret and X. Vilasis-Cardona, *Performance of the LHCb calorimeters during the period 2010-2012*, J. Phys. Conf. Ser. **587** (2015), no. 1 012012.
- [71] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020.
- [72] D. A. Berninghoff, J. Albrecht, and V. Gligorov, *Bremsstrahlung Recovery of Electrons using Multivariate Methods*, Tech. Rep. LHCb-INT-2016-018. CERN-LHCb-INT-2016-018, CERN, Geneva, Apr, 2016.
- [73] F. Alessio *et al.*, *The lhcb data acquisition during lhc run 1*, Journal of Physics: Conference Series **513** (2014) 012033.
- [74] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022.
- [75] J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. : Conf. Ser. **513** (2014) 012001. 8 p, arXiv:1310.8544.
- [76] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, Tech. Rep. LHCb-PUB-2011-016. CERN-LHCb-PUB-2011-016. LHCb-INT-2011-030, CERN, Geneva, Sep, 2011.
- [77] A. Puig, *The LHCb trigger in 2011 and 2012*, Tech. Rep. LHCb-PUB-2014-046. CERN-LHCb-PUB-2014-046, CERN, Geneva, Nov, 2014.
- [78] J. Brehmer, J. Albrecht, and P. Seyfert, *Ghost probability: an efficient tool to remove background tracks*, Tech. Rep. LHCb-INT-2012-025. CERN-LHCb-INT-2012-025, CERN, Geneva, Sep, 2012.

-
- [79] LHCb collaboration, R. Aaij *et al.*, *Measurement of J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C71** (2011) 1645.
- [80] LHCb collaboration, R. Aaij *et al.*, *Production of J/ψ and Υ mesons in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **06** (2013) 064.
- [81] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150.
- [82] T. Tekampe, *Analysis of the rare decay $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ at LHCb*, Master's thesis, TU Dortmund, Germany, 2014.
- [83] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356.
- [84] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, arXiv:1608.05806.
- [85] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.
- [86] A. Poluektov, *Correction of simulated particle identification response in LHCb using transformation of variables*, Tech. Rep. LHCb-INT-2017-007. CERN-LHCb-INT-2017-007, CERN, Geneva, Apr, 2017.
- [87] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, May, 2014.
- [88] LHCb collaboration, R. Aaij *et al.*, *Search for decays of neutral beauty mesons into four muons*, JHEP **03** (2017) 001.
- [89] F. Pedregosa *et al.*, *Scikit-learn: Machine learning in Python*, Journal of Machine Learning Research **12** (2011) 2825.
- [90] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.
- [91] A. Gulin, I. Kuralenok, and D. Pavlov, *Winning the transfer learning track of Yahoo!'s learning to rank challenge with YetiRank*, JMLR Proceedings **14** (2011) 63.
- [92] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.
- [93] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001, and 2013 partial update for the 2014 edition.

- [94] LHCb collaboration, R. Aaij *et al.*, *Study of the rare B_s^0 and B^0 decays into the $\pi^+\pi^-\mu^+\mu^-$ final state*, Phys. Lett. **B743** (2015) 46.
- [95] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001 CERN-PH-EP-2013-006, LHCb-PAPER-2012-037, f_s/f_d value updated in LHCb-CONF-2013-011.
- [96] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13 TeV pp collisions*, arXiv:1902.06794.
- [97] M. De Cian *et al.*, *Measurement of the track finding efficiency*, Tech. Rep. LHCb-PUB-2011-025. CERN-LHCb-PUB-2011-025, CERN, Geneva, Apr, 2012.
- [98] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, arXiv:1812.10790.
- [99] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.
- [100] R. D. Cousins and V. L. Highland, *Incorporating systematic uncertainties into an upper limit*, Nucl. Instrum. Meth. **A320** (1992) 331.
- [101] A. L. Read, *Presentation of search results: The CL(s) technique*, Journal of Physics G: Nuclear and Particle Physics **28** (2002) 2693.
- [102] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. **A434** (1999) 435.
- [103] LHCb collaboration, R. Aaij *et al.*, *Evidence for the rare decay $\Sigma^+ \rightarrow p\mu^+\mu^-$* , Phys. Rev. Lett. **120** (2018) 221803 LHCb-PAPER-2017-049, CERN-EP-2017-319.
- [104] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [105] A. Rohatgi, S. Rehberg, and Z. Stanojevic, *Webplotdigitizer: Version 4.1*, Jan, 2018. doi: 10.5281/zenodo.1137880.
- [106] G. A. Cowan, D. C. Craik, and M. D. Needham, *RapidSim: an application for the fast simulation of heavy-quark hadron decays*, Comput. Phys. Commun. **214** (2017) 239, The LHCb PID response was implemented in RapidSim after the studies detailed in this thesis were performed.
- [107] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984. ISBN: 978-0412048418.
- [108] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, Annals of Statistics **29** (2000) 1189.

-
- [109] G. Ke *et al.*, *Lightgbm: A highly efficient gradient boosting decision tree*, in *Advances in Neural Information Processing Systems 30* (I. Guyon *et al.*, eds.), pp. 3146–3154. Curran Associates, Inc., 2017.
- [110] J. A. Nelder and R. Mead, *A simplex method for function minimization*, *The Computer Journal* **7** (1965), no. 4 308.
- [111] M.-H. Schune, F. Polci, and M. Borsato, *HOP an additional tool for decays involving electrons*, Tech. Rep. LHCb-INT-2015-037. CERN-LHCb-INT-2015-037, CERN, Geneva, Nov, 2015.
- [112] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
- [113] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b -quark production cross-section in 7 and 13 TeV pp collisions*, *Phys. Rev. Lett.* **118** (2017), no. 5 052002, [Erratum: *Phys. Rev. Lett.*119,no.16,169901(2017)].
- [114] M. Chala, U. Egede, and M. Spannowsky, *Searching new physics in rare B -meson decays into multiple muons*, arXiv:1902.10156.
- [115] M. Prim, *Semileptonic and Leptonic B Decay Results from early Belle II Data*, URL: <https://meetings.triumf.ca/indico/event/48/session/8/contribution/24/material/slides/0.pdf>, May, 2019. Presentation at the Conference on Flavor Physics and CP Violation.

A Appendix

A.1 Stripping selection

A.1.1 Search for $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ decays

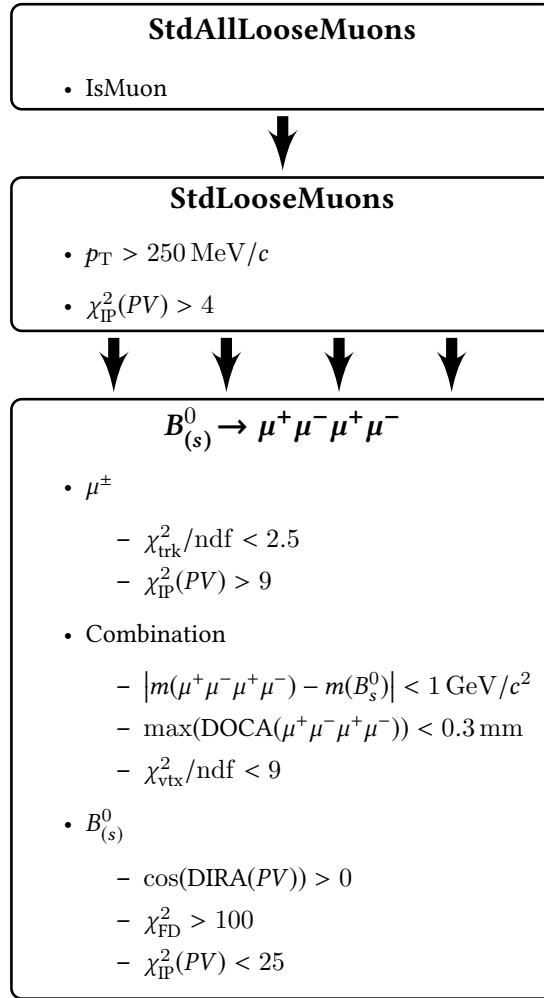


Figure A.1: Sketch of the connection between individual stripping containers that are employed in selecting $B_{(s)}^0 \rightarrow \mu^+ \mu^- \mu^+ \mu^-$ candidates.

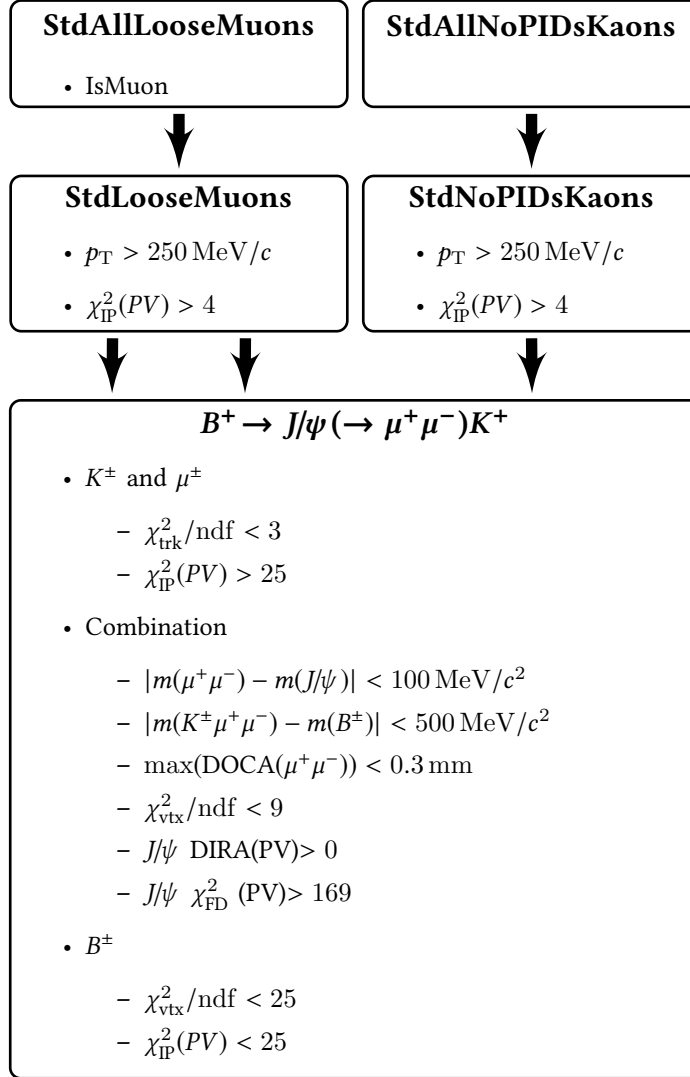


Figure A.2: Sketch of the connection between individual stripping containers that are employed in selecting $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) K^+$ candidates.

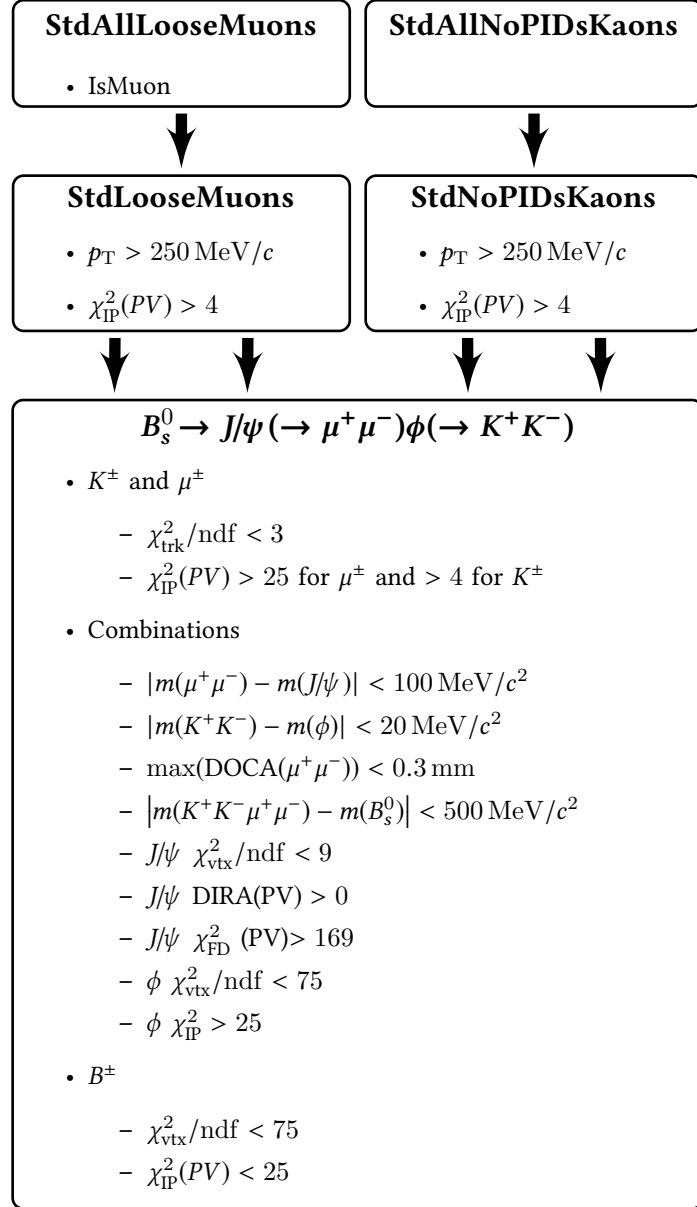


Figure A.3: Sketch of the connection between individual stripping containers that are employed in selecting $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \phi(\rightarrow K^+ K^-)$ candidates.

A.1.2 Test of lepton flavour universality with $R_{K\pi\pi}$

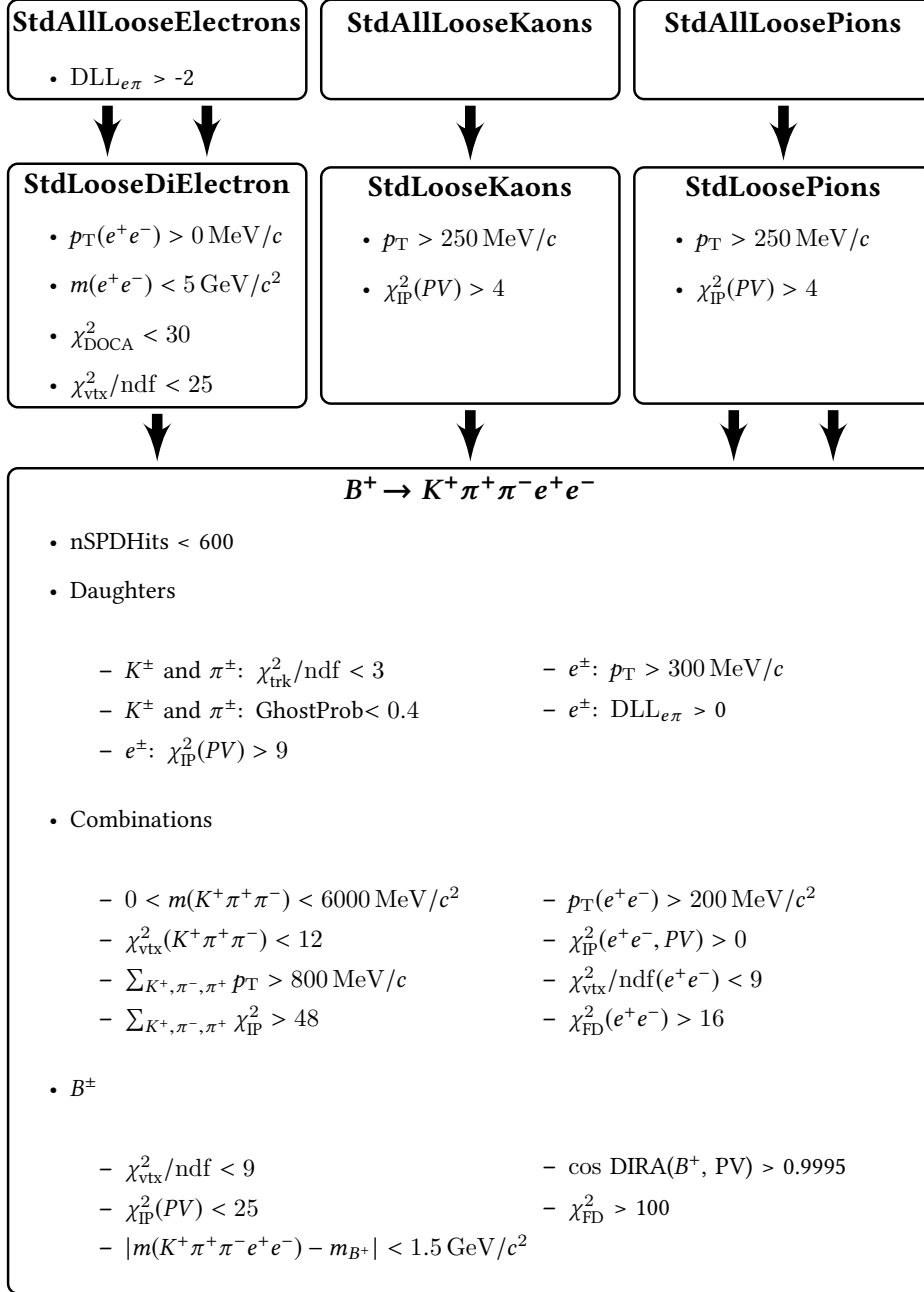


Figure A.4: Sketch of the connection between individual stripping containers that are employed in selecting the $K^+ \pi^+ \pi^- e^+ e^-$ final state.

A.2 Comparisons between measured and simulated data for $R_{K\pi\pi}$

A.2.1 Features that are subject to simulation corrections

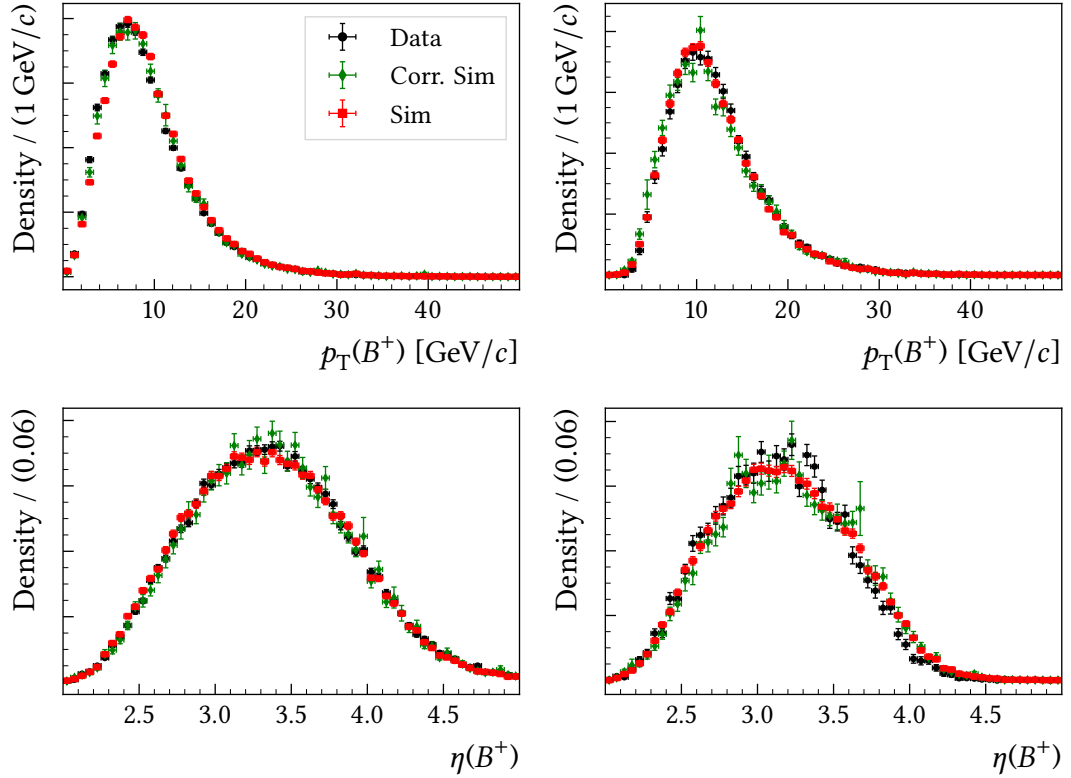


Figure A.5: Comparison between measured data, raw simulation and corrected simulation for the first two out of four features that are corrected in the second and third step as described in section 6.2.1. Shown are resonant di-muon decays (left) and resonant di-electron decays (right). For measured data *sWeights* are employed in order to isolate the signal component.

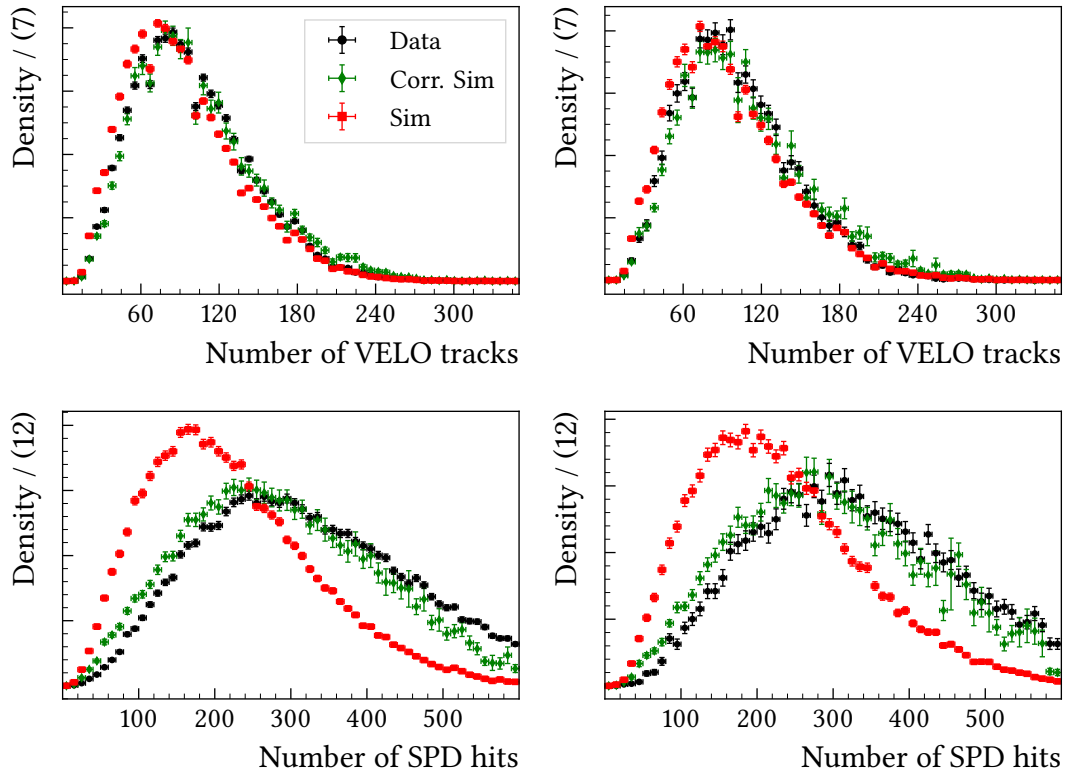


Figure A.6: Comparison between measured data, raw simulation and corrected simulation for the second two out of four features that are corrected in the second and third step as described in section 6.2.1. Shown are resonant di-muon decays (left) and resonant di-electron decays (right). For measured data $sWeights$ are employed in order to isolate the signal component.

A.2.2 Features entering the BDT

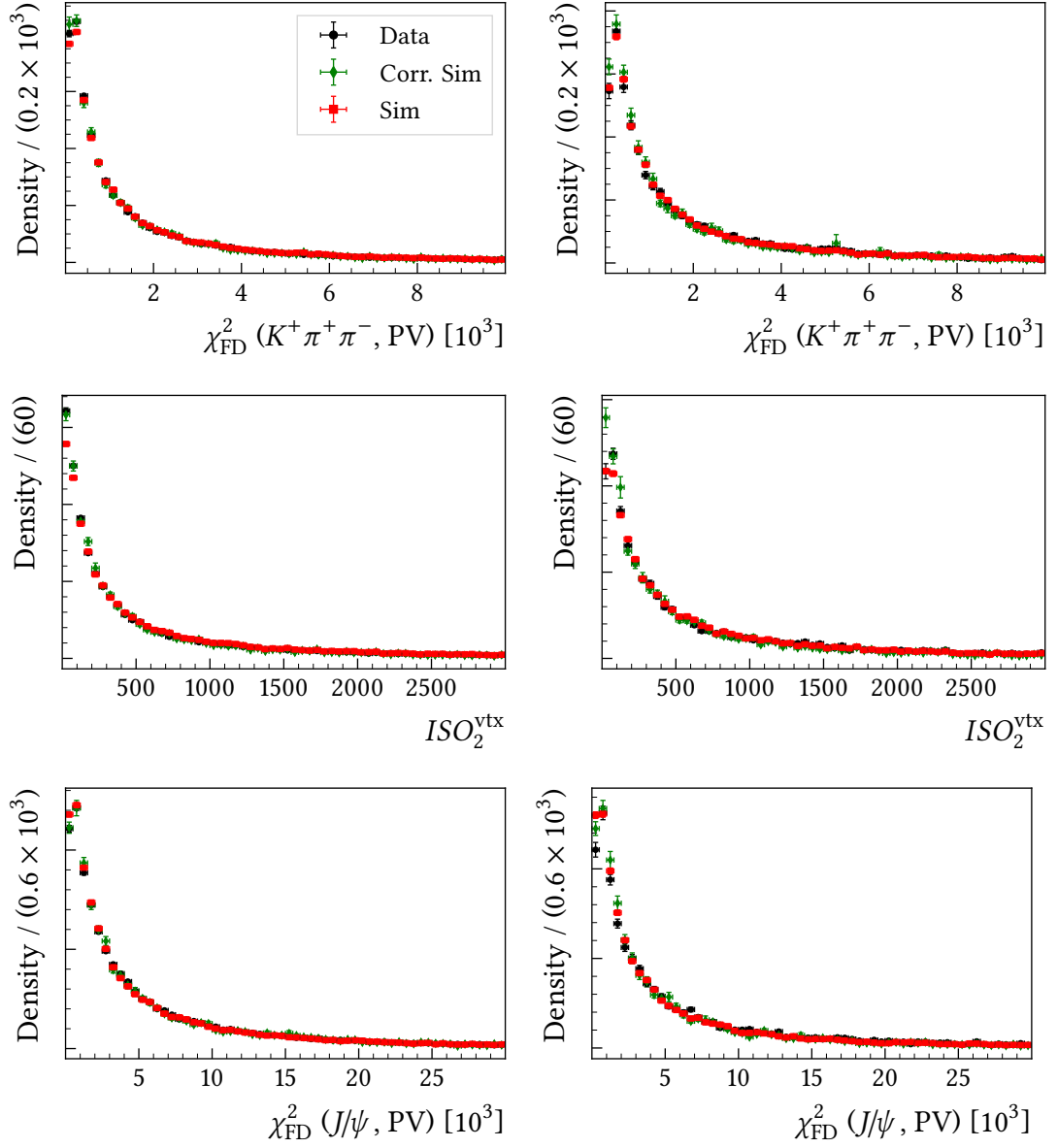


Figure A.7: Feature 5, 6 and 7 used to train a classifier to discriminate signal from combinatorial background. Addendum to Figure 6.6.

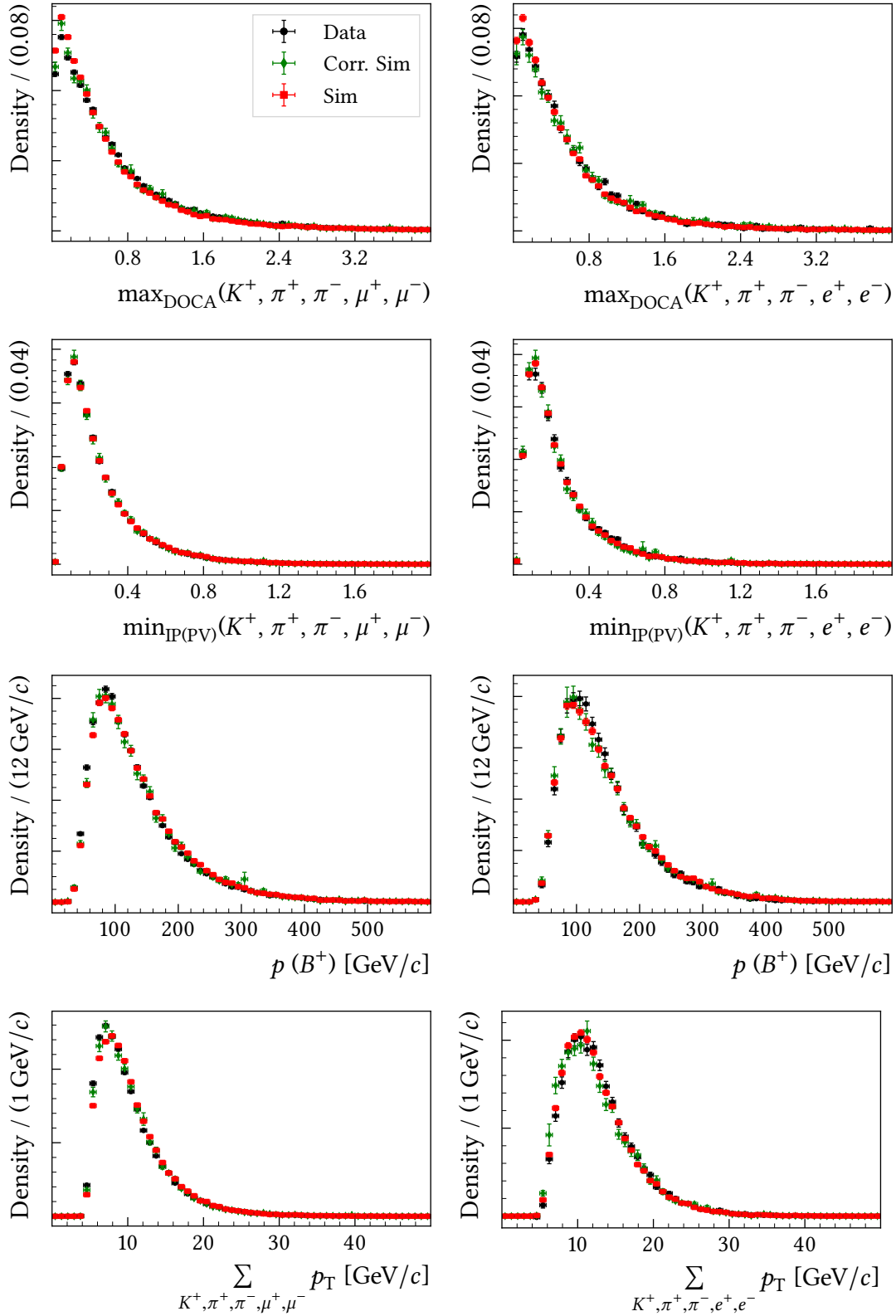


Figure A.8: Feature 8 to 11 used to train a classifier to discriminate signal from combinatorial background. Addendum to Figure 6.6.

A.2.3 Particle identification quantities

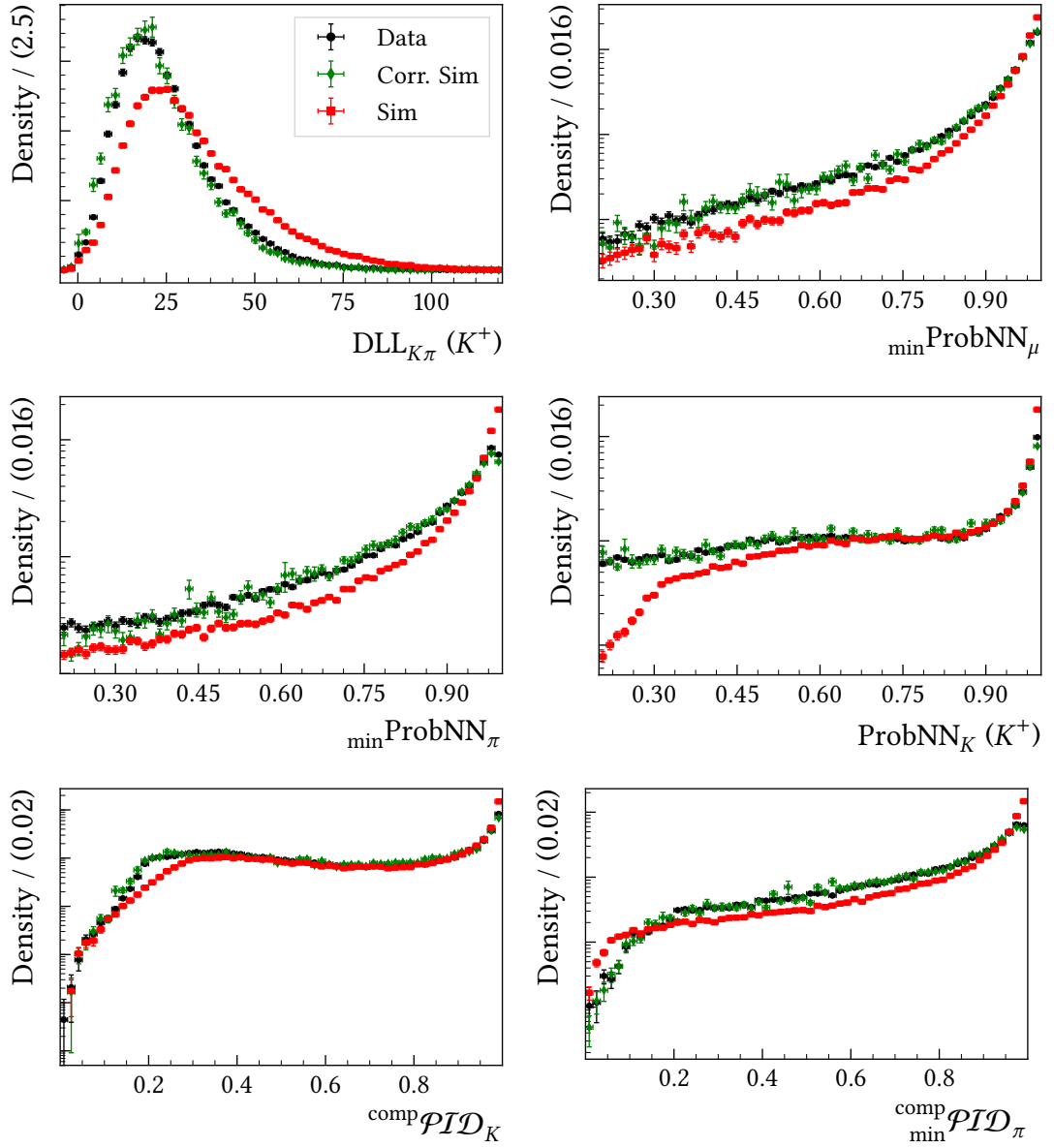


Figure A.9: The PID quantities in the $B^+ \rightarrow K^+ \pi^+ \pi^- J/\psi (\rightarrow \mu^+ \mu^-)$ decays used in the measurement of $R_{K\pi\pi}$ to selected the di-muon final state. Shown are the simulated distributions before and after applying the PID calibration and measured data. For the latter $sWeights$ are employed in order to isolate the signal component.

A.3 Background shapes for $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$

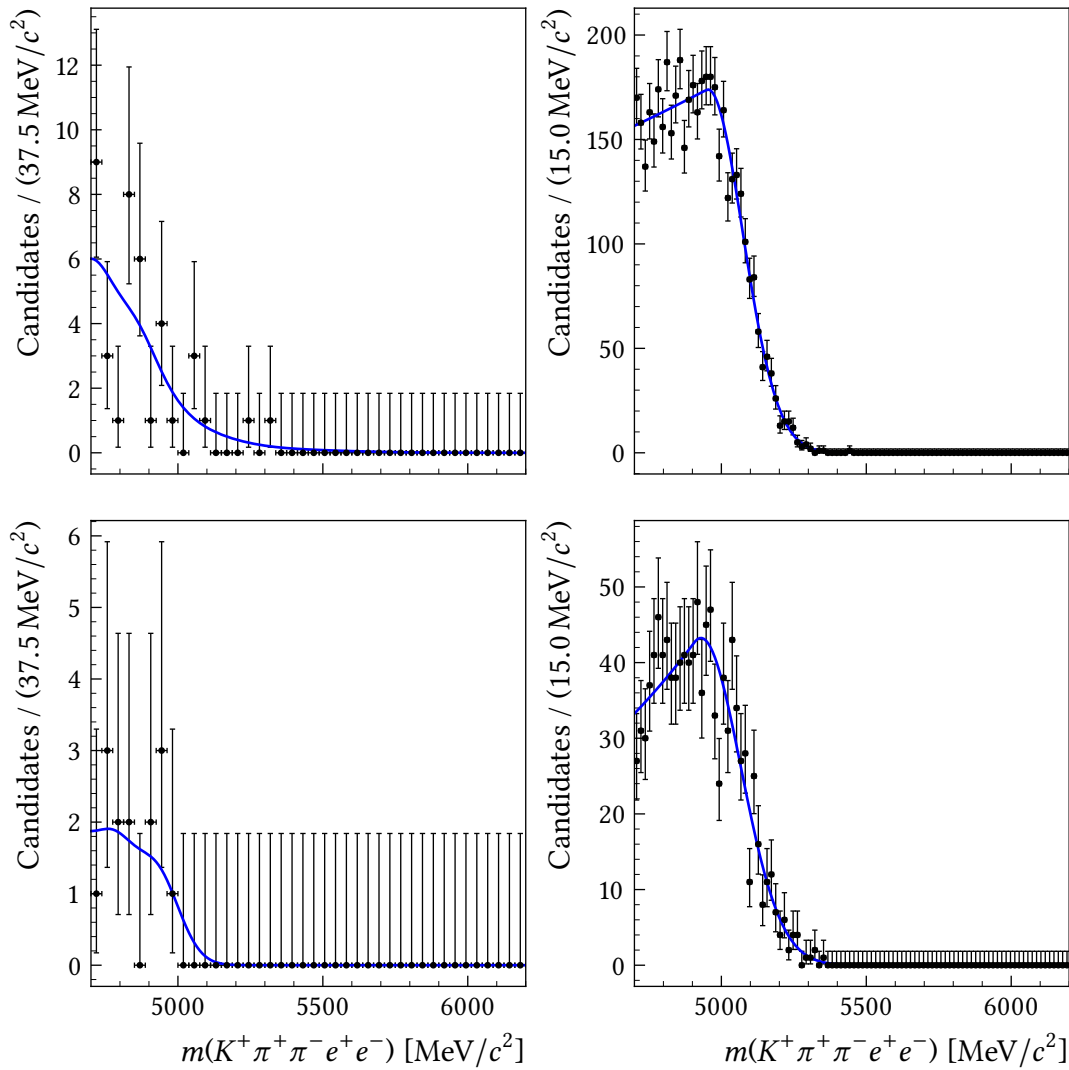


Figure A.10: The left plots show simulated resonant data that pass the non-resonant selection parametrised by a KDE. On the right, partially reconstructed decays parametrised by a Gaussian function that transitions into an exponential function in direction of lower mass are shown. In the top the ETOS and in the bottom the TIS trigger category is presented.

Acknowledgements

An dieser Stelle möchte ich mich zunächst bei meinem Doktorvater Dr. Johannes Albrecht bedanken, der mich auch schon während meiner Masterarbeit betreute. Vielen Dank, dass Du mir stets mit Rat und Tat zur Seite standest und mir neben der Promotion auch die Teilnahme an diversen nationalen wie internationalen Konferenzen und Workshops ermöglicht hast.

Außerdem bedanke ich mich bei der Prüfungskommission bestehend aus Prof. Dr. Kevin Kröniger, Prof. Dr. Heinz Hövel und Dr. Bärbel Siegmann, insbesondere auch für die Realisierung eines Prüfungstermins außerhalb der Vorlesungszeit.

Der gesamten Arbeitsgruppe E5a danke ich für die angenehme gemeinsame Zeit und die vielen gemeinsamen Aktivitäten, wie Rad- oder Kanutouren. Insbesondere der Belegschaft von CP-O3-107, Timon, Moritz, Maik und Alex, möchte ich für die stets heitere Stimmung in unsrem Büro danken. Danke auch an Steffi, die auch nach ihrem Abschied aus der Arbeitsgruppe noch Teile dieser Arbeit Korrektur gelesen hat.

Ganz besonderer Dank gilt auch meiner Familie, die mich nicht erst seit meinem Physikstudium stets beim Erreichen meiner Ziele unterstützt hat. Ohne Euch wäre das nicht möglich gewesen!

Last but not least danke ich meiner Freundin Laura für viel Geduld und Verständnis, gerade in der Endphase meiner Promotion.