

Measurement of rare leptonic and semi-baryonic B meson decays with the LHCb experiment

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

vorgelegt von

Maik Becker

geboren am 23.10.1993 in Unna

Fakultät Physik

Technische Universität Dortmund

Dortmund, August 2023

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

Gutachter:

Prof. Dr. Johannes Albrecht

Prof. Dr. Kevin Kröniger

Vorsitzender der Prüfungskommission:

Prof. Dr. Christoph Lange

Vertreterin der wissenschaftlichen Mitarbeiter*innen:

Dr. Doris Reiter

Datum des Einreichens der Dissertation:

17. August 2023

Datum der mündlichen Prüfung:

16. Oktober 2023

Abstract

This thesis presents two measurements probing the Standard Model of particle physics with data recorded by the LHCb experiment in Run 1 and Run 2 of the Large Hadron Collider.

The first measurement is a precise branching fraction determination of the purely leptonic $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays. The $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction is determined with a relative uncertainty of 15 %, making it the most precise single-experiment result available at publication time. The $B^0 \rightarrow \mu^+ \mu^-$ decay is not observed. An upper limit on its branching fraction is calculated to be of $\mathcal{O}(10^{-10})$ at a confidence level of 95 %. These published results strongly constrain theoretical models of new physics effects in the flavour sector.

The second analysis is the first search for the semi-baryonic $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays. It is the first time that decays of B_s^0 mesons via flavour-changing neutral $b \rightarrow s\ell^+\ell^-$ transitions into a final state with baryons are studied. The current state of the measurement and preliminary results are presented. Expected upper limits on the branching fractions are determined to be of $\mathcal{O}(10^{-9})$ at a confidence level of 95 %.

Kurzfassung

In dieser Arbeit werden zwei Messungen vorgestellt, die das Standardmodell der Teilchenphysik prüfen. Dazu werden Daten verwendet, die vom LHCb Experiment in Run 1 und Run 2 des Large Hadron Collider aufgezeichnet wurden.

Die erste Messung ist eine präzise Bestimmung der Verzweigungsverhältnisse der rein leptonischen $B_s^0 \rightarrow \mu^+ \mu^-$ und $B^0 \rightarrow \mu^+ \mu^-$ Zerfälle. Das $B_s^0 \rightarrow \mu^+ \mu^-$ Verzweigungsverhältnis wird mit einer relativen Unsicherheit von 15 % bestimmt und ist zum Zeitpunkt der Veröffentlichung das präziseste Ergebnis eines einzelnen Experiments. Der $B^0 \rightarrow \mu^+ \mu^-$ Zerfall wird nicht beobachtet. Die obere Ausschlussgrenze für das Verzweigungsverhältnis des Zerfalls liegt in der Größenordnung von $\mathcal{O}(10^{-10})$ bei einem Konfidenzlevel von 95 %. Diese veröffentlichten Ergebnisse schränken die theoretischen Modelle der Effekte neuer Physik im Flavour-Sektor stark ein.

Die zweite Analyse ist eine Suche nach den semi-baryonischen $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ und $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ Zerfällen. Es ist das erste Mal, dass Zerfälle von B_s^0 Mesonen über flavour-verändernde neutrale $b \rightarrow s\ell^+\ell^-$ Übergänge in einen Endzustand mit Baryonen untersucht werden. Der aktuelle Stand der Messung und vorläufige Ergebnisse werden vorgestellt. Die erwarteten oberen Ausschlussgrenzen für die Verzweigungsverhältnisse liegen in der Größenordnung von $\mathcal{O}(10^{-9})$ bei einem Konfidenzlevel von 95 %.

Contents

1	Introduction	1
I	Theoretical and experimental context	3
2	Theory	5
2.1	The Standard Model of particle physics	5
2.2	Heavy-flavour B physics	8
2.3	Rare $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays	10
2.4	Introducing baryons: $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays	13
3	Experiment	17
3.1	The LHC accelerator and collider complex	17
3.2	The LHCb detector	18
3.2.1	Track and vertex reconstruction	20
3.2.2	Identification of muons and protons	22
3.2.3	Trigger system and B candidate reconstruction	24
3.3	Introduction to LHCb simulation	25
3.4	LHCb terms and variables	25
II	Measurement of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays	29
4	Analysis introduction	31
4.1	Analysis strategy	31
4.2	Selection of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ candidates	32
4.3	Simulation calibration	34
5	Background studies	41
5.1	Hadronic backgrounds	42
5.2	Semileptonic backgrounds	43
5.2.1	$B^0 \rightarrow \pi^- \mu^+ \nu_\mu$	44
5.2.2	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	46
5.2.3	$B^+ \rightarrow \pi^+ \mu^+ \mu^-$ and $B^0 \rightarrow \pi^0 \mu^+ \mu^-$	48
5.2.4	$\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$	49
5.3	Summary of expected background events	51

6	Branching fraction measurement	55
6.1	Normalisation with $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow K^+ \pi^-$ decays	55
6.1.1	Selection efficiencies	56
6.1.2	Tracking efficiency corrections	57
6.1.3	Normalisation factors and cross-checks	60
6.2	Modelling of the invariant mass distribution	63
6.3	Results and interpretation	64
III	First search for $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays	71
7	Strategy and data preparation	73
7.1	Data and simulation samples	74
7.2	Simulation calibration	75
7.2.1	Particle identification	77
7.2.2	Trigger response	80
7.2.3	Event multiplicity and B hadron kinematics	81
7.2.4	Decay model	83
8	Selection of $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ candidates	87
8.1	Trigger requirements	87
8.2	Preselection and multivariate classification	89
8.3	Selection optimisation	96
9	Modelling of invariant mass distributions	99
9.1	Model for $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ decays	99
9.2	Model for $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays	104
10	Expected upper limits	107
10.1	Normalisation with $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ decays	107
10.1.1	Number of normalisation channel events	108
10.1.2	Selection efficiencies	108
10.1.3	Normalisation factors	109
10.2	Results and interpretation	112
11	Conclusion	115
	References	117
	Acknowledgements	131

1 Introduction

The Standard Model of particle physics is a theory that describes the properties of the fundamental particles of nature and their interactions [1–6]. Since its development in the 1960s, it has been utilised to make precise predictions about subatomic processes and the results of particle physics experiments. It has led to major successes, such as the prediction of the Higgs boson in 1964 [7–9], which was finally discovered in 2012 [10, 11] and completed the set of elementary Standard Model particles. So far, the Standard Model is the experimentally most tested and established theory describing the microcosm on a fundamental level.

However, despite its great success, it has been known for several decades that more than the Standard Model is needed to describe the mechanisms of the universe in all its aspects. For example, the Standard Model includes only three of the four fundamental forces of nature: the electromagnetic, the weak, and the strong force. It does not explain the gravitational force due to its special properties as described in Einstein’s general theory of relativity and due to the fact that its impact only becomes noticeable at greater distances and masses. Another problem arises when looking at the cosmological scale. From the rotation of stars around their galactic centre and the observed accelerated expansion of the universe, it can be inferred that the visible matter described by the Standard Model accounts for only 5% of the universe’s energy density. The remaining fraction consists of “dark matter” (26%) and “dark energy” (69%) [12]. The origin and nature of these contributions remain unknown, leaving space for new particles or interactions extending the Standard Model. Furthermore, concerning its internal structure, the Standard Model is introduced with several free parameters. The question arises whether hidden connections between the free parameters exist and whether the Standard Model is only one part of a more complex theory.

To answer the open questions of the Standard Model, particle physicists build experiments, perform measurements, and conduct research on its predictions. They aim to increase the knowledge of the universe’s fundamental building blocks by reducing the uncertainties on the theory’s free parameters and to find physics beyond the Standard Model (BSM), from which it is known that it must exist.

In recent decades, the experimental landscape has partly evolved towards larger experiments that test the Standard Model at high energy scales. The largest particle accelerator and collider in the world is the Large Hadron Collider (LHC), which can produce particles at energies of $\mathcal{O}(1 \text{ TeV})$. These particles can be found in direct searches with large detectors. Since the LHC’s maximum acceleration

power limits the accessible energy range for these direct searches, a second approach motivates the study of particle decays with very low Standard Model rates in indirect searches. It is known that BSM effects must be small (otherwise, they would have been seen already). Hence, the relative enhancement of such effects is expected to be much larger for rare than for more frequent decays. Because the BSM effects are expected to contribute with quantum corrections or new couplings to the fundamental quark and lepton interactions, potential new intermediate particles can contribute off-shell at much higher energies of up to $\mathcal{O}(100 \text{ TeV})$ [13, 14]. By measuring rare decays, these effects and particles can be studied indirectly.

The LHCb experiment at the LHC is specialised in indirect measurements in the beauty and charm quark sector. As part of its rich physics programme, one fundamental class of rare decays contains the transition of a beauty quark into a strange quark and two charged leptons. Recent measurements of such flavour-changing neutral $b \rightarrow s\ell^+\ell^-$ transitions have shown tensions with the Standard Model, suggesting a modified lepton coupling in weak interactions. None of these measurements is statistically significant on its own, but the overall picture suggests more detailed studies on the subject [15].

This thesis presents two analyses involving decays with $b \rightarrow s\ell^+\ell^-$ or related $b \rightarrow d\ell^+\ell^-$ transitions. The first analysis includes a branching fraction measurement of the $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ decays exploiting the entire data set from high-energy proton-proton collisions the LHCb detector recorded between 2011 and 2018. Due to their simple two-body final states, precise Standard Model predictions exist for the decays, making them an excellent probe for BSM effects. The results of this measurement have been published [16, 17]. The second analysis is a search for $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays. This is the first analysis of a decay with a $b \rightarrow s\ell^+\ell^-$ transition having a B meson in the initial state and baryons in the final state. Next to the effects observed in other $b \rightarrow s\ell^+\ell^-$ transitions, this measurement allows to probe for possible new mechanisms affecting the diproton spectrum near its production threshold, as suggested by measurements in similar decay channels [18].

This thesis is divided into three main parts. In the first part, the two analyses are placed in a theoretical and experimental context. Chapter 2 briefly introduces the Standard Model of particle physics and gives a motivation for the performed measurements. Chapter 3 introduces the LHCb detector and essential parts of the data acquisition chain. The second part of the thesis presents the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ branching fraction measurement in Chapters 4 to 6. A focus is set on those aspects of the analysis in which the author of this thesis was significantly involved. The third part presents the current status of the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ analysis and first preliminary results produced for this thesis. This part includes Chapters 7 to 10 and is followed by a short conclusion highlighting the main results of the presented work.

Part I

Theoretical and experimental context

2 Theory

This chapter gives an overview of the theory and the current experimental status of the analysed decay modes. Key aspects of the Standard Model of particle physics are summarised in Section 2.1, with a more detailed look at $b \rightarrow s\ell^+\ell^-$ transitions and b flavour anomalies in Section 2.2. The investigated $B_{(s)}^0 \rightarrow \mu^+\mu^-$ and $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decay modes are discussed in Sections 2.3 and 2.4, respectively, focusing on theoretical predictions and past measurements.

2.1 The Standard Model of particle physics

Particle physics is dedicated to studying the fundamental constituents of the universe, the elementary particles, and the interactions between them. The current best knowledge manifests in the Standard Model (SM) of particle physics, a relativistic gauge quantum field theory described in Refs. [1–6]. Its Lagrange density is derived based on global Poincaré symmetry and local gauge symmetry of the unitary $SU(3)_C \times SU(2)_L \times U(1)_Y$ group. It has 19 independent parameters that must be determined experimentally [19] and describes three of the four fundamental forces of nature: the strong force introduced by the $SU(3)_C$ group, the weak force, and the electromagnetic force. The latter two can be unified into an electroweak interaction represented by the $SU(2)_L \times U(1)_Y$ group. The fourth fundamental force, gravity, is not part of the SM since a quantum-mechanical description of gravitation considering general relativity that can be integrated into the SM has not yet been achieved. Fortunately, due to its weakness on the microscopic scale, the effects of gravitation are negligible for all practical purposes in particle physics.

The SM consists of twelve fermions with spin $\frac{1}{2}$, which form the matter of the universe. A summarising illustration of all elementary particles is given in Fig. 2.1. The fermions can be divided into six quarks (q) and six leptons (ℓ and ν_ℓ) of different flavours. Both quarks and leptons can be organised into three generations of two particles each, which are subject to a hierarchy of increasing particle masses. The six leptons are the electron (e^-), the muon (μ^-), and the tau (τ^-) with an electrical charge of $-1e$, and an associated chargeless lepton neutrino for each of the three lepton flavours. Besides the leptons, there are the three up-type quarks up (u), charm (c), and top (t), and the three down-type quarks down (d), strange (s), and bottom/beauty (b). Each quark generation consists of one up-type and one down-type quark. The electrical charge of up-type quarks is $\frac{2}{3}e$, while it is $-\frac{1}{3}e$ for down-type quarks. All quarks carry a colour charge,

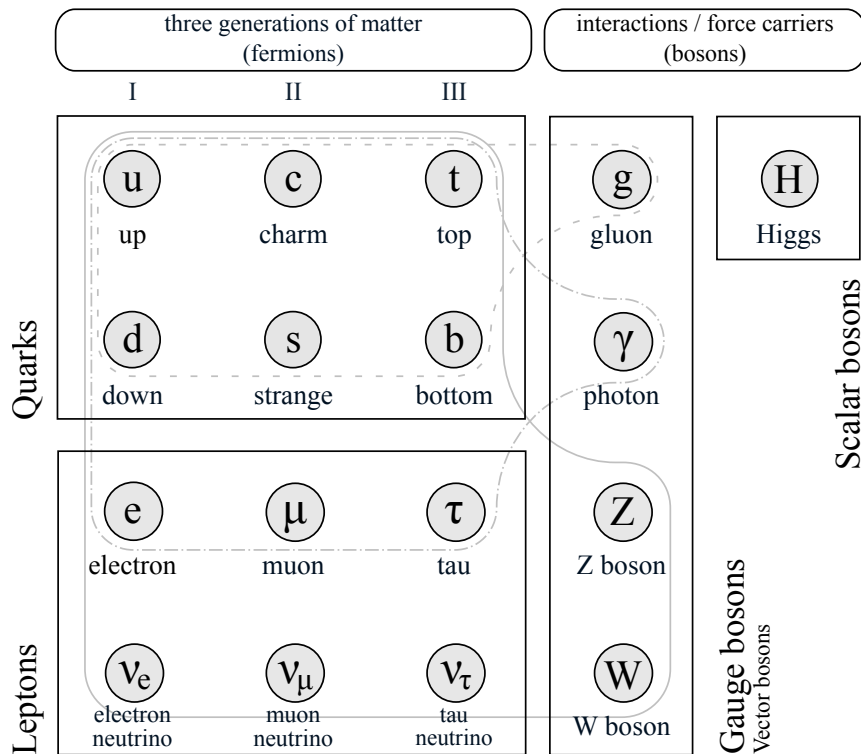


Figure 2.1 – Overview of elementary particles described by the SM. The twelve fermions forming the matter of the universe are shown together with the bosons mediating the fundamental forces. Figure refactored from Ref. [20].

which is the property the strong interaction couples to. The colour charge can be red, green, or blue. Like the electrical charge, also the colour charge is always preserved. Finally, the number of fundamental fermionic particles is doubled by twelve anti-fermions resulting from the multiple solutions of the Dirac equation. These anti-matter particles mirror ordinary-matter particles but with inverted charge-like quantum numbers, such as the electrical charge or the colour. Studying and comparing the properties of matter and anti-matter is an essential part of the research programme of flavour experiments like LHCb.

Next to the properties of the fermions, the SM also describes their interactions by introducing twelve bosonic spin-1 mediator particles that carry the forces. The twelve vector bosons are the massless photon γ carrying the electromagnetic force, the massive W^\pm and Z bosons mediating the weak force, and eight massless gluons for the strong force. Introduced by the theory of quantum chromodynamics (QCD), the gluons can be distinguished by different combinations of colour and anti-colour charges that they carry. Self-coupling is possible among the gluons, which leads to increasing coupling strength for smaller energies and longer distances. Due to this, quarks of high energies behave like free particles (asymptotic freedom) but are forced to colourless bound states at lower energies (confinement). In the lower energy regime, the quarks are bound together to

hadrons. They appear in states of three quarks, the baryons and anti-baryons, or in states of a quark and an anti-quark, the mesons. The hadrons present in the signal modes of this thesis are the B^0 and B_s^0 mesons, consisting of $\bar{b}d$ and $\bar{b}s$ quark pairs, respectively, and the baryonic (anti-)protons that have a uud (or $\bar{u}\bar{u}\bar{d}$) quark content. In recent years, besides mesons and baryons, also the conjunctions of four and five quarks, called tetra- [21] and pentaquarks [22–24], were observed by the LHCb collaboration.

In contrast to the strong force, the weak interaction affects both the leptons and the quarks. The W^\pm and Z mediators couple to the weak isospin, a non-zero property for particles with left-handed chirality. For massive particles, the chirality has a left-handed and a right-handed component, while for neutrinos, the chirality corresponds to the helicity, which is the projection of the particle spin onto its momentum. In the SM, all neutrinos are massless and, therefore, of left-handed chirality. The left-handed fermions form weak isospin doublets, which can be organised as

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L \quad (2.1)$$

and

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

with L denoting the left-handed component. The charged currents of the weak interaction can change the third component of the weak-isospin vector, leading to transitions of left-handed neutrinos into charged leptons and left-handed up-type quarks into down-type quarks and reverse. For quarks, the transitions dominate within a generation but are also possible between them with lower probability. This quark flavour mixing is described by the Cabibbo–Kobayashi–Maskawa matrix V_{CKM} [25, 26] connecting the flavour eigenstates (d, s, b) with the mass eigenstates (d', s', b') via

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.3)$$

The absolute values of the matrix elements representing the square roots of the transition probabilities are quoted without uncertainties from Ref. [27]:

$$|V_{\text{CKM}}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97435 & 0.22500 & 0.00369 \\ 0.22486 & 0.97349 & 0.04182 \\ 0.00857 & 0.04110 & 0.999118 \end{pmatrix}. \quad (2.4)$$

In practice, it follows that processes like decays with unfavourable transitions, *e.g.* from the third to the first generation, are rather unlikely. This important effect is called CKM suppression (or Cabibbo suppression).

The SM is completed by the scalar Higgs boson H^0 as an excitation of the Higgs field. Postulated in 1964, the Higgs mechanism [7–9] explains how the gauge bosons get their mass from the spontaneous symmetry breaking of the electroweak interaction and how the Yukawa coupling connects the fermion masses to the Higgs field.

Since its invention in the 1970s, the SM has been tested with various particle physics experiments at different energy scales. To this day, it has been proven that its concepts are valid. Still, some limitations exist, which are conceptually related or due to observations since its invention. A selection is given below:

- In the original SM formulation, lepton neutrinos must be massless. However, measurements of the neutrino oscillation frequency have shown that they have a non-zero mass [28, 29].
- Gravitation is not considered, because a quantum-mechanical description of its effects on the cosmological scale, as described by general relativity, has not yet been achieved.
- According to astrophysical observations [12], about 5 % of the universe consists of matter described by the SM. The origin and nature of the remaining 95 %, referred to as dark matter and energy, remain unexplained.
- The canonical SM (without neutrino masses and neutrino mixing) is based on 19 free parameters [19], such as coupling constants or particle masses, whose origins and possible connections cannot be derived in the SM.
- The SM formulation allows the violation of symmetry under charge and parity conjugation (CP) in the sector of strong interactions. However, despite many experiments [30], no such violation has ever been observed for QCD. The reason for this coincidence is unknown and is referred to as strong CP problem.

2.2 Heavy-flavour B physics

The decays presented in this thesis proceed via CKM-suppressed $b \rightarrow s\ell^+\ell^-$ or $b \rightarrow d\ell^+\ell^-$ transitions, which require an electroweak current that switches the down-type quark flavour without changing the electrical charge. Such a flavour-changing neutral current (FCNC) does not exist as fundamental interaction in the SM and, therefore, cannot occur in first order of perturbation calculation (tree level). Instead, a combination of multiple charged currents is needed (loop level), leading to strong suppression of the affected decays. However, new interactions are proposed in several new physics (NP) models that allow FCNC processes on tree level, altering the predicted decay rates. Because small new effects might have a relatively strong impact on the observed rates, the measurement of rare B meson decays is a sensitive probe for NP models.

When discussing and comparing the experimental and theoretical findings for different B meson decays, it is beneficial to formulate the theoretical predictions within an effective field theory (EFT) using an operator product expansion. In this model-independent formalism, the transition amplitude \mathcal{A} of a B meson decaying to a final state f is given as

$$\mathcal{A}(B \rightarrow f) = \langle f | \mathcal{H}_{\text{eff}} | B \rangle = \frac{G_{\text{F}}}{\sqrt{2}} V_{tb} V_{t(d,s)}^* \sum_{i=7,9,10} C_i(\lambda) \langle f | \mathcal{O}_i(\lambda) | B \rangle, \quad (2.5)$$

where \mathcal{H}_{eff} is the effective Hamiltonian of the transition, G_{F} is the Fermi constant, and V_{ij} are the CKM matrix elements of the involved flavour transitions. The introduced operators \mathcal{O}_i describe the vertices of the effective theory with the couplings C_i , also known as Wilson coefficients. The operators cover non-perturbative effects at long distances, while the coefficients contain information about perturbative short-distance effects at high energies. Both the operators and coefficients depend on the energy scale λ , chosen as the b quark mass. The Wilson coefficients are independent of the initial and final states of the decay and can be determined, *e.g.* with lattice QCD by matching the amplitude expression in the EFT with that in the full theory using the W^\pm mass as threshold energy. For $b \rightarrow s\ell^+\ell^-$ transitions, only three operators are contributing, which comprise the radiative $b \rightarrow s\gamma(\rightarrow \ell^+\ell^-)$ transition (\mathcal{O}_7), vector couplings (\mathcal{O}_9), and axial-vector couplings (\mathcal{O}_{10}). The associated Wilson coefficients can be examined by analysing different decay channels, and global fits to multiple observables can be constructed to obtain an overall picture of agreement with the SM values as seen in Fig. 2.2.

In recent years, $b \rightarrow s\ell^+\ell^-$ transitions were studied in several measurements covering differential and total branching fractions [33–36], angular distributions [37, 38], and lepton-flavour universality (LFU) ratios [39–41] between electron and corresponding muon modes. So far, no single measurement shows a significant deviation from the SM predictions, but in the combination of all results, a tension is seen in the muon sector concerning the SM value of $C_9^{\mu\mu}$. Depending on the included analyses and the construction of the global fit, slight tensions are also observed for $C_{10}^{\mu\mu}$. Measurements of LFU ratios in charged and uncharged currents further stress these anomalies since some imply different coupling strengths of the weak force to electrons and muons. Most prominent were the first measurements of R_K and R_{K^*} [39, 42], defined as the ratio between the decay rates of the $B^+ \rightarrow K^+\ell^+\ell^-$ (or $B^0 \rightarrow K^{*0}\ell^+\ell^-$) electron and muon modes, which the SM expects to be close to one. Those measurements showed a tension of up to 3σ with respect to the SM. However, the latest measurements of R_K and R_{K^*} [31, 32], that employ a larger LHCb data set, show better agreement with the SM prediction of equal coupling strength. Another class of decays are charged $b \rightarrow c\ell\nu_\ell$ transitions, which have been analysed by the Belle, BaBar, and LHCb collaborations in Refs. [43–46]. A discrepancy from the SM at a level of 2σ is seen between the muon and tau modes, which is still present. Further ten-

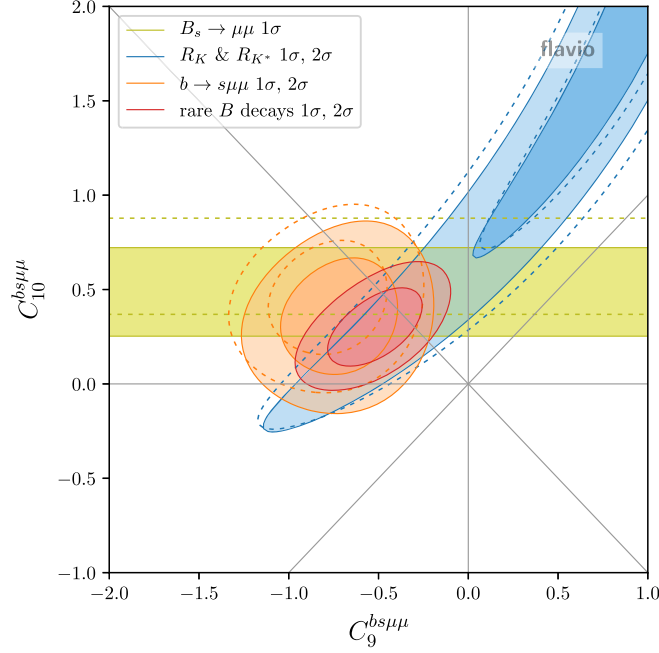


Figure 2.2 – Likelihood contours in the Wilson coefficient plane $C_9^{bs\mu\mu}$ and $C_{10}^{bs\mu\mu}$. The contours are shown for the previous $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction measurement, the lepton-flavour universality ratios R_K and R_{K^*} (not including the latest LHCb results from Refs. [31, 32]), a combination of several $b \rightarrow s\mu^+\mu^-$ observables, and the global fit. The dashed lines indicate the contours before the inclusion of newer updates. Figure taken from Ref. [15].

sions are reported, *e.g.* from the branching fraction measurement of $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays [47] and from the angular distribution of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays [48, 49]. A detailed summary can be found in Ref. [15]. Given these findings, the overall picture of the flavour anomalies remains unclear. To solve the issue, it is necessary to continue measuring $b \rightarrow s\ell^+\ell^-$ transitions on larger data sets and in different decay channels.

2.3 Rare $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays

The rare leptonic $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays are studied in this thesis. Both decays proceed via the weak interaction and have low rates because the needed FCNC transitions are only allowed on loop level and include CKM suppression (see Section 2.2). Additionally, the final-state muons must be in opposite spin states with the same helicity to conserve angular momentum. This condition leads to strong helicity suppression by a factor of m_μ^2/m_b^2 since the currents of the weak interaction can only couple to the left-handed (right-handed) component of particles (anti-particles). The leading-order Feynman diagrams for the decays are shown in Fig. 2.3. The current best SM predictions

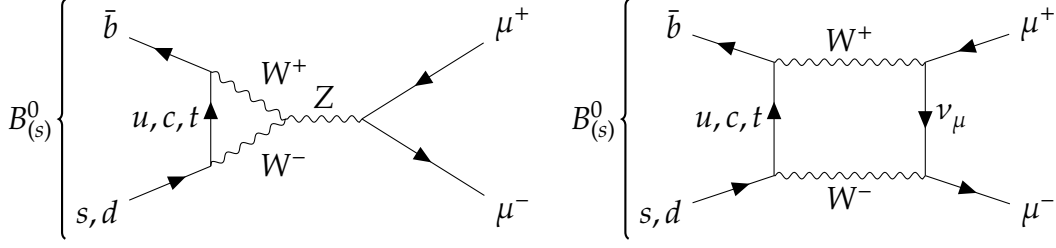


Figure 2.3 – Two leading-order Feynman diagrams for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays. The flavour-changing neutral current can proceed via (left) penguin or (right) box diagram.

are calculated in Ref. [50] and yield

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)_{\text{SM}} = (3.66 \pm 0.14) \times 10^{-9} \quad (2.6)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)_{\text{SM}} = (1.03 \pm 0.05) \times 10^{-10} . \quad (2.7)$$

The branching fractions are determined with quite small uncertainties due to the simplicity of the purely leptonic final state. Calculations profit from recent progress in understanding electroweak effects at next-to-leading order [51], QCD effects at next-to-next-to-leading order [52], and lattice QCD [53]. Also, corrections from virtual photon exchange are now calculated with higher precision than in previous times [50, 54], further reducing the uncertainties on the SM values. The uncertainties on the CKM matrix inputs highly dominate the remaining uncertainties.

Within the EFT formalism, the SM branching fractions can be expressed as

$$\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+ \mu^-)_{\text{SM}} = \tau_{B_{(s)}^0} \frac{G_F^2 \alpha^2}{16\pi^2} f_{B_{(s)}^0}^2 m_\mu^2 m_{B_{(s)}^0} \sqrt{1 - \frac{4m_\mu^2}{m_{B_{(s)}^0}^2} |V_{tb} V_{t(d,s)}^*|^2} |C_{10}^{\text{SM}}|^2, \quad (2.8)$$

with the fine-structure constant α , and the $B_{(s)}^0$ meson lifetime τ and decay constant f . The branching fractions are only sensitive to the Wilson coefficient C_{10} . However, additional NP contributions are still discussed, given the uncertainties of the present measured and calculated values. Therefore, $B_s^0 \rightarrow \mu^+ \mu^-$ decays are a suitable probe for NP effects. Many models introduce new contributions from pseudo-scalar and scalar currents that can be described by the composed coefficients

$$P = \frac{C_{10}^L - C_{10}^R}{C_{10}^{\text{SM}}} + \frac{m_{B_{(s)}^0}^2}{2m_\mu} \frac{m_b}{m_b + m_{d,s}} \left(\frac{C_P^L - C_P^R}{C_{10}^{\text{SM}}} \right) \quad (2.9)$$

and

$$S = \sqrt{1 - \frac{4m_\mu^2}{m_{B_{(s)}^0}^2} \frac{m_{B_{(s)}^0}^2}{2m_\mu} \frac{m_b}{m_b + m_{d,s}} \left(\frac{C_S^L - C_S^R}{C_{10}^{\text{SM}}} \right)}. \quad (2.10)$$

These coefficients contain the quark and lepton masses, Wilson coefficients for additional left- or right-handed scalar and pseudo-scalar contributions, C_S and C_P , and also the coefficient C_{10}^R describing a potential right-handed component of C_{10} . In the SM, it is $P = 1$ and $S = 0$, meaning that with measurements of the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fractions the Wilson coefficient $C_{10}^{\text{SM}} = C_{10}^L$ can be directly accessed.

In NP models, S and P can have different values. No attempt is made in this thesis to review all possible NP contributions. The most discussed models introduce new bosons like leptoquarks [55–57] or additional Higgs bosons within a supersymmetric extension of the SM [58, 59]. As right-handed contributions to C_{10} , additional neutral Z' vector bosons are considered [60, 61].

To lower the number of potential NP effects, most models are constructed to obey the minimal flavour violation hypothesis [62]. According to this hypothesis, the NP effects should be equally strong on both $b \rightarrow s \ell^+ \ell^-$ and $b \rightarrow d \ell^+ \ell^-$ transitions. The ratio of the B_s^0 decay rate and the B^0 decay rate can be compared between calculation and measurement to test the hypothesis. Based on Eq. (2.8), the SM value of the ratio is calculated to be

$$\mathcal{R}_{\mu^+ \mu^-}^{\text{SM}} = \frac{\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)_{\text{SM}}}{\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)_{\text{SM}}} = 0.0281 \pm 0.0016 . \quad (2.11)$$

Another observable is derived from the sizeable relative decay width difference $y_s = \Delta\Gamma_s/(2\Gamma_s)$ of the heavy and light mass eigenstates of the B_s^0 meson. In the SM, the CP asymmetry parameter is given by $\mathcal{A}_{\Delta\Gamma} = +1$ and, therefore, only the heavy mass eigenstate decays into two muons. The effective $B_s^0 \rightarrow \mu^+ \mu^-$ lifetime is the B_s^0 lifetime seen in the $B_s^0 \rightarrow \mu^+ \mu^-$ decay. It is derived as

$$\tau(B_s^0 \rightarrow \mu^+ \mu^-) = \frac{\tau_{B_s^0}}{1 - y_s^2} \left(\frac{1 + 2\mathcal{A}_{\Delta\Gamma} y_s + y_s^2}{1 + \mathcal{A}_{\Delta\Gamma} y_s} \right) , \quad (2.12)$$

and offers a complementary probe of NP in the scalar and pseudo-scalar sectors by giving access to the CP structure of the decay. New effects might be seen, even if the branching fractions match the SM predictions.

At the beginning of the work presented in this thesis, the world's most precise values for the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fractions were provided by a combination [63] of the corresponding LHCb [64], CMS [65], and ATLAS [66] results:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (2.69_{-0.35}^{+0.37}) \times 10^{-9} \quad (2.13)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 1.9 \times 10^{-10} \quad \text{at } 95\% \text{ CL} . \quad (2.14)$$

Given the experimental uncertainties, the measured value for the $B_s^0 \rightarrow \mu^+ \mu^-$ decay is in tension with the calculated SM value at 2.1σ . Although this finding alone might not be concerning, it is worth noting that the observed 25 % deviation from the SM is in good agreement with the general picture of flavour anomalies

seen in $b \rightarrow s\ell^+\ell^-$ transitions [15]. Recently and after the publication of the LHCb result presented in this thesis, the CMS collaboration provided a new measurement for the $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ decays on a larger data set [67]:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) = \left(3.83_{-0.36}^{+0.38} \text{ (stat)}_{-0.21}^{+0.24} \text{ (syst)}\right) \times 10^{-9} \quad (2.15)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-) < 1.9 \times 10^{-10} \quad \text{at 95 \% CL .} \quad (2.16)$$

The result for $B_s^0 \rightarrow \mu^+\mu^-$ reduces the tension with the SM value. A discussion including the results presented in this thesis can be found in Section 6.3. In Ref. [63], the ratio of B_s^0 and B^0 rates was measured to be

$$\mathcal{R}_{\mu^+\mu^-} = 0.021_{-0.025}^{+0.030}, \quad (2.17)$$

which is insignificant and fully compatible with the predicted SM value. The effective $B_s^0 \rightarrow \mu^+\mu^-$ lifetime was determined by the LHCb and CMS collaborations. A combination of both results [63] led to

$$\tau(B_s^0 \rightarrow \mu^+\mu^-) = \left(1.91_{-0.35}^{+0.37}\right) \text{ ps .} \quad (2.18)$$

In the latest CMS analysis [67], the effective lifetime was measured to be

$$\tau(B_s^0 \rightarrow \mu^+\mu^-) = \left(1.83_{-0.20}^{+0.23} \text{ (stat)}_{-0.04}^{+0.04} \text{ (syst)}\right) \text{ ps ,} \quad (2.19)$$

giving a consistent picture of the measurements.

2.4 Introducing baryons: $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays

The second set of decays analysed in this thesis comprises the rare semi-baryonic $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays. It is the first time $b \rightarrow s\ell^+\ell^-$ transitions are studied with a B meson in the initial and baryons in the final state. A first search for the decays is performed to measure their decay rates and to study the effect of threshold enhancement, which describes the peaking of the baryon-antibaryon pair mass spectrum near the kinematic threshold. This effect has been seen in several three- and four-body decay modes [18], but is not fully understood yet. Theoretical approaches range from the involvement of intermediate resonances (e.g. $X(1835)$ or f_2 states) to non-perturbative QCD effects that affect the quark fragmentation process [68]. There are only a few published measurements that employ proton-antiproton pairs in the final state, which are described in the following.

The semileptonic $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ and $B^+ \rightarrow p\bar{p}e^+\nu_e$ decays were analysed by the Belle and LHCb collaborations [69, 70]. It was found that the measured branching fractions are one order of magnitude lower than the theoretical predictions at that time [71]. Following these observations, the theoretical calculations were

revised and now agree with the experimental results [72, 73]. Figure 2.4 shows the differential branching fraction of the muon mode as a function of the invariant diproton mass.

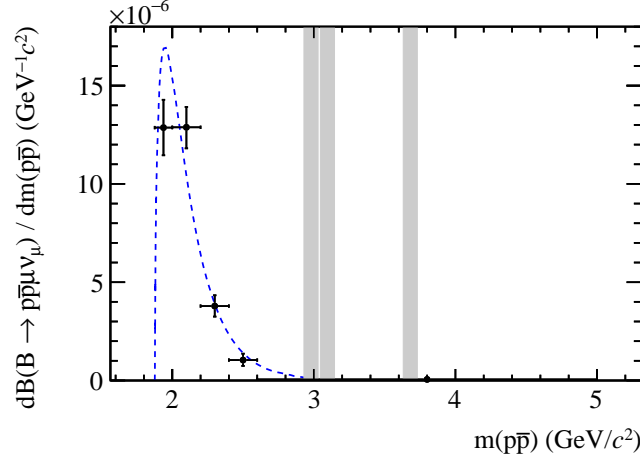


Figure 2.4 – Differential branching fraction of the $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ mode as a function of the invariant diproton mass. The LHCb measurement (black points) is shown together with the theoretical prediction (dotted curve). Grey bands indicate vetoes on charmonium resonances. Figure modified from Ref. [70].

The purely baryonic $B^0 \rightarrow p\bar{p}p\bar{p}$ decay was found by the LHCb collaboration, and evidence for the $B_s^0 \rightarrow p\bar{p}p\bar{p}$ decay is reported in Ref. [74]. The measured rate of the B^0 mode matches with theoretical calculations. For the B_s^0 mode, the rate is surprisingly large. However, this could be explained by statistical fluctuations in the small data set. If the high rate is confirmed in future measurements, this would be of great interest since to reach the $p\bar{p}p\bar{p}$ final state, an s quark must be converted, and, therefore, a CKM-suppressed process would be expected.

The resonant $B_s^0 \rightarrow J/\psi p\bar{p}$ and $B^0 \rightarrow J/\psi p\bar{p}$ decays were observed by the LHCb experiment for the first time [75]. The invariant mass distribution of the $J/\psi p\bar{p}$ system is shown in Fig. 2.5. The branching fractions were measured to be

$$\mathcal{B}(B_s^0 \rightarrow J/\psi p\bar{p}) = (3.58 \pm 0.19 \text{ (stat)} \pm 0.39 \text{ (syst)}) \times 10^{-6} \quad (2.20)$$

$$\mathcal{B}(B^0 \rightarrow J/\psi p\bar{p}) = (4.51 \pm 0.40 \text{ (stat)} \pm 0.44 \text{ (syst)}) \times 10^{-7}. \quad (2.21)$$

While the branching fraction of the B^0 mode agrees with expectations, the decay rate of the B_s^0 mode is much larger than expected. When looking at the quark-level processes, the B^0 meson decays via a CKM-suppressed $b \rightarrow c\bar{c}d$ transition with the diproton system originating directly from the hadronisation of the d quark. For the B_s^0 meson, the underlying process must annihilate the s quark and create the diproton system from gluons alone. Such processes are affected by the Okubo–Zweig–Iizuka (OZI) rule [76] of QCD, which predicts a suppression if the initial states can be separated from the final states by only

removing internal gluons. However, precise QCD calculations are difficult for the low-energy system present. Figure 2.6 shows the leading-order Feynman diagrams for the two processes. In a subsequent LHCb amplitude analysis of the B_s^0 mode, evidence was found for a pentaquark state in the invariant $J/\psi p$ and $J/\psi \bar{p}$ systems [77]. The contribution of this possible state to the decay rate has not yet been conclusively determined.

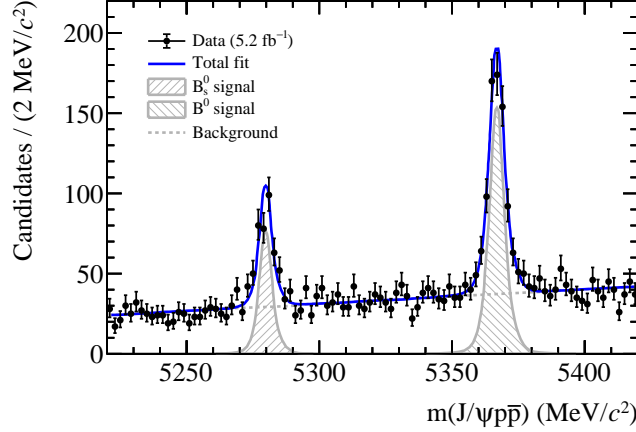


Figure 2.5 – First observation of the $B_s^0 \rightarrow J/\psi p\bar{p}$ and $B^0 \rightarrow J/\psi p\bar{p}$ decays. The data and a fit to the reconstructed $J/\psi p\bar{p}$ mass distribution are shown. Figure modified from Ref. [75].

Given these findings, the study of more decays with proton-antiproton pairs in the final state is advised to understand the observed decay rate of the B_s^0 mesons. The $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays are an obvious choice given that selecting events with a dimuon pair is highly efficient at LHCb. Unfortunately, no predictions for the decay rates of the two modes exist so far. In Fig. 2.6, the leading-order Feynman diagrams are presented, which show a similar structure compared to the respective diagrams of the resonant decays. An additional internal loop gives stronger CKM suppression. The rates of the decays are expected to be very small. The orders of magnitude for the branching fractions are estimated by looking at $b \rightarrow s\mu^+\mu^-$ and $b \rightarrow d\mu^+\mu^-$ processes, where the corresponding $b \rightarrow sJ/\psi$ and $b \rightarrow dJ/\psi$ rates are measured. For the B^0 mode, decays such as $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}J/\psi$, or $\Lambda_b^0 \rightarrow p\pi^-\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow p\pi^-J/\psi$ can be used. It turns out that the rates of the $b \rightarrow d\mu^+\mu^-$ transitions are about three orders of magnitude smaller than the respective $b \rightarrow dJ/\psi$ rates. Following this rule, the $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ branching fraction is expected to be of $\mathcal{O}(10^{-10})$. The $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ decay rate is estimated to be of $\mathcal{O}(10^{-9}) - \mathcal{O}(10^{-8})$, assuming that the large $B_s^0 \rightarrow J/\psi p\bar{p}$ rate is caused by hadronic effects in the diproton system, which also affect the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ mode. In case the enhanced rate of $B_s^0 \rightarrow J/\psi p\bar{p}$ decays is caused by resonances or effects in the $J/\psi p$ system, like pentaquarks or rescattering effects, one would not expect to see it in the non-resonant $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ mode.

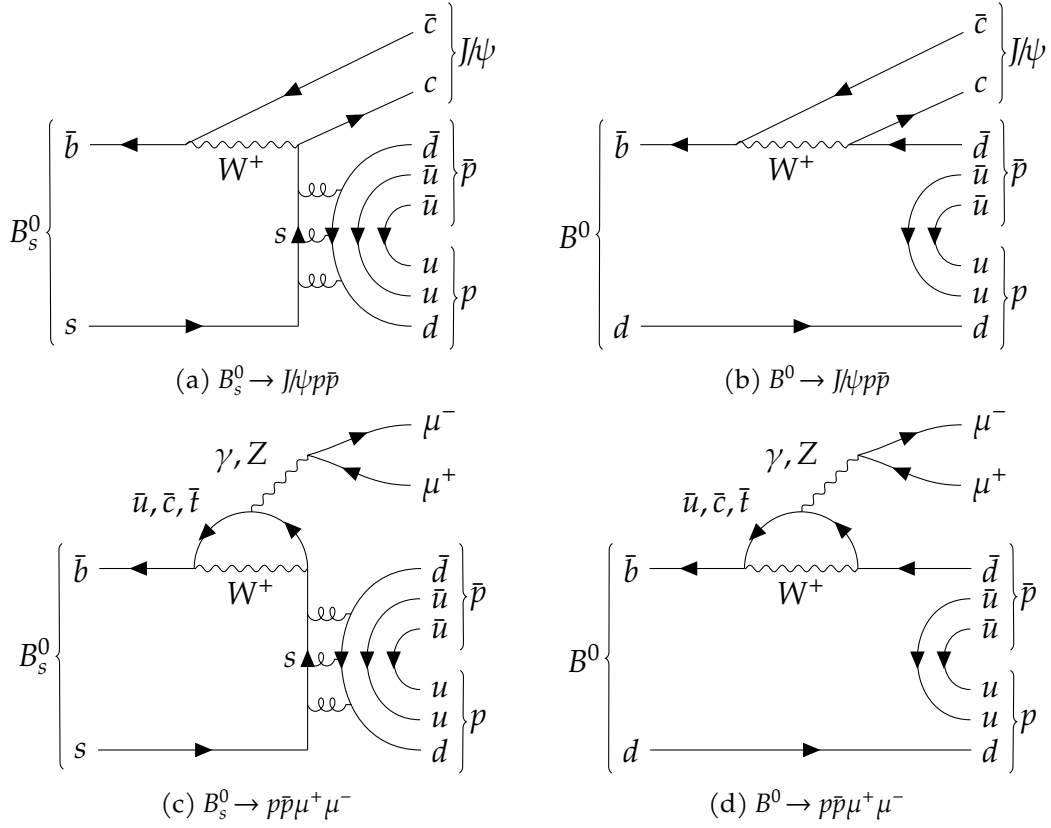


Figure 2.6 – The leading-order Feynman diagrams for the resonant $B_{(s)}^0 \rightarrow J/\psi p \bar{p}$ and non-resonant $B_{(s)}^0 \rightarrow p \bar{p} \mu^+ \mu^-$ decays. In contrast to the B^0 mode, the B_s^0 modes are additionally suppressed by the OZI rule.

A search for $B_{(s)}^0 \rightarrow p \bar{p} \mu^+ \mu^-$ decays and a precision measurement of their branching fractions are performed to give more insight into the described anomaly. It is unclear whether the decays can be observed with the current statistical power of the LHCb data set. Even if the decays are not found, a calculated upper limit on the rates would also be a valuable input for the theoretical discussion of the studied threshold enhancement.

3 Experiment

The experimental setup that was used to record the analysed data is presented. First, the proton-proton accelerator and collider complex LHC is described briefly in Section 3.1, followed by a more detailed description of the LHCb detector and the data-taking environment in Section 3.2. Given its importance for the measurements, a short introduction to LHCb simulation is given in Section 3.3. Finally, in Section 3.4, some LHCb specific terms and variables are defined that will appear in later parts of the thesis.

3.1 The LHC accelerator and collider complex

The LHCb experiment is located at the Large Hadron Collider (LHC) of the European Organisation for Nuclear Research (CERN). The LHC [78, 79] is the world's largest ring-shaped superconducting hadron accelerator and collider with a circumference of about 26.7 km, located below the surface in the former tunnel of the Large Electron-Positron Collider (LEP).

The ring consists of eight straight sections and eight arcs, equipped with acceleration cavities and bending magnets capable of speeding up protons and holding them on a circular path. The bending is done by 1232 superconducting dipole magnets that are cooled down to 1.9 K with liquid nitrogen and helium and reach a magnetic field strength of up to 8.3 T. For acceleration, eight superconducting radio-frequency (RF) cavities are used. In addition, a series of different multipole magnets are installed along the ring to keep the particle beams focused. Surrounded by the magnets and RF cavities, the protons are gathered in up to 2808 bunches of $\mathcal{O}(10^{11})$ particles that circulate in opposite directions in two spatially separated beam pipes. The bunches were lined up with a temporal spacing of 50 ns, which was reduced to 25 ns as of 2015 due to better beam control. The beam pipes are evacuated to an ultra-high vacuum of 10^{-13} to 10^{-14} bar to prevent the highly-energised particles from interacting with air molecules.

While taking proton-proton collision data, the LHC was operating at different beam energies ranging from 3000 (3500) GeV in the year 2011 (2012) to 6500 GeV between 2015 and 2018. Since the LHC cannot accelerate protons from rest, it is linked to preceding accelerator facilities powered by CERN. The protons are obtained by ionising hydrogen atoms and raised to an energy of 50 MeV using the linear accelerator LINAC2. Then, they are transferred into a series of ring accelerators that gradually increase the particles' energy: the Booster,

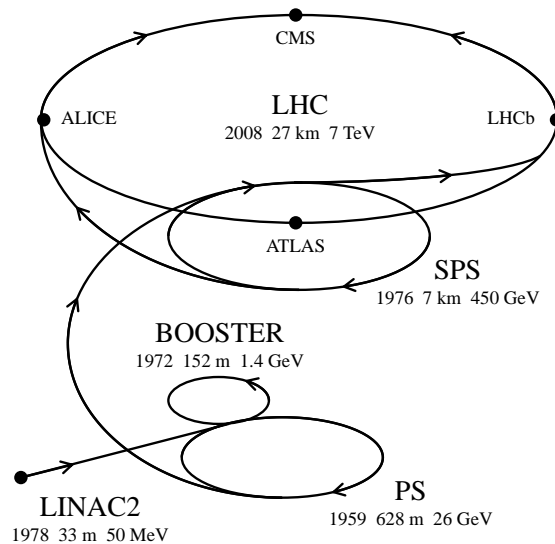


Figure 3.1 – Scheme of the LHC and its preceding accelerators at CERN. The designed output beam energy, length, and year of commissioning are given for each accelerator. Arrows denote the proton path. Figure refactored from Ref. [80].

the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). With an energy of 450 GeV, the protons leave the SPS and are injected clockwise and counterclockwise into the LHC to reach the desired colliding energy. Figure 3.1 shows a scheme of the LHC and its preceding accelerators.

At four distinct locations in the LHC tunnel, the two beams are collimated and crossed, resulting in particle collisions with a centre-of-mass energy of up to 13 TeV and high luminosity. At these points, large underground caverns host particle detectors that surround the interaction points and try to catch the long-living products of the particle interactions. The four large experiments at the LHC are ATLAS [81], CMS [82], ALICE [83] and LHCb [84]. The two multi-purpose detectors ATLAS and CMS are known for their significant contributions in the Higgs and top-quark sector, especially for the discovery of the Higgs boson in 2012 [10, 11]. In contrast, the ALICE experiment makes major contributions in the field of heavy ion physics studying the quark-gluon plasma when ions are filled into the collider. The LHCb experiment used to collect the data for the two measurements presented in this thesis is described in the following sections.

3.2 The LHCb detector

The LHCb experiment comprises the actual detector components in the underground cavern and the subsequent data acquisition and processing systems that are partly above ground. It is operated by an international collaboration of more than 1500 scientists from 96 institutes in 21 countries. By the time of writing this thesis, the author was part of the LHCb collaboration.

The LHCb detector [84] is a single-arm particle spectrometer covering a range from 10 to 300 (250) mrad in the forward angular bending (non-bending) plane. Figure 3.2 shows the arrangement of the different detector components. The detector is optimised to measure processes involving b quarks, such as B_s^0 and B^0 decays. At the LHC, the production of heavy b quark pairs in proton-proton collisions (via the gluon fusion process $gg \rightarrow b\bar{b}$) is dominant for small angles between their flight direction and the beam pipe [85]. Consequently, there are two regions, forward and backwards, where most b hadrons are produced. Due to limited space in the underground cavern and for the sake of an optimised instrumentation, the detector covers only one of the two regions containing around 25% of the produced $b\bar{b}$ pairs.

In the first two data-taking periods of the LHC from 2011 to 2012 (Run 1) and from 2015 to 2018 (Run 2), the LHCb detector collected a substantial sample of b hadrons [86], which is ideally suited for the analysis of rare B meson decays.

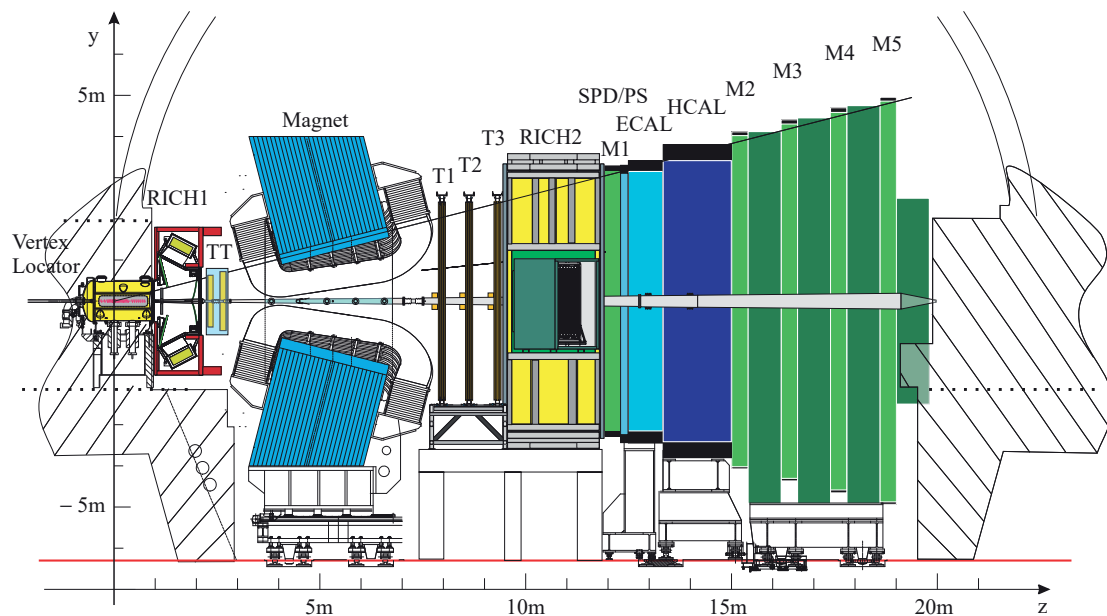


Figure 3.2 – Cross section of the LHCb experiment with named detector components. The coordinate system is right-handed with the x -axis pointing to the centre of the LHC, the y -axis pointing to the surface, and the z -axis pointing along the beam pipe in the forward direction. Figure taken from Ref. [87].

A particle detector records the partial energy deposit of particles in its detector cells. Different detector systems and technologies are needed to develop hypotheses about the final state particles, their kinematics and the general event topology for a complete reconstruction of a collision event. In this context, the detector subsystems serve different tasks like particle identification, tracking, or vertex reconstruction. However, the overall event interpretation relies on information from every subsystem. In the following, the LHCb detector components in their Run 2 settings are described.

3.2.1 Track and vertex reconstruction

An essential part of the event reconstruction is the particle track reconstruction, which aims at finding the trajectories of particles passing through the detector. By using methods like pattern recognition and track fitting, track candidates and the total number of tracks in an event (the event multiplicity) can be derived. In the case of charged particles and with a known magnetic field, a momentum estimate for all three spatial directions can be calculated based on the track curvature. A second essential part is the vertex reconstruction. A vertex is a point in space where at least two particle tracks intersect. Primary vertices (PV) can be found in the interaction region and are defined as the points of proton-proton collisions. In contrast, secondary vertices (SV) can be reconstructed by looking for intersections of particle tracks incompatible with being assigned to a PV. The B_s^0 and B^0 mesons are known to fly several millimetres through the detector before decaying due to their lifetime of around 1.5 ps and the relativistic boost. Hence, resolving the B decay vertices and determining their positions is a powerful feature for distinguishing between particles that result from detached B meson decays and the ones from processes at the PV.

To find the particle tracks and vertices, the LHCb detector is equipped with a dedicated tracking system. It consists of a vertex locator (VELO), a large dipole magnet, and four planar tracking stations, of which one, the Tracker Turicensis (TT), is located upstream, and the three others (T1–T3) are located downstream of the magnet. For the reconstruction of muon tracks, additional information from a dedicated muon system are used (see Section 3.2.2).

The VELO [88, 89] is made of 42 semi-circular silicon strip modules that surround the proton-proton interaction region along the beam direction up to a distance of 8.2 mm. The sensor modules are placed perpendicular to the beam pipe and in two halves inside a vacuum vessel. They provide a measurement of particle hits in the polar coordinates r and ϕ , and achieve a best spatial resolution of 4 μm in a region of high track density. An aluminium foil separates the vacuum of the beam pipe from the VELO. To protect the VELO from hard radiation damage when the LHC is ramping up and the proton beams are unfocused, the detector halves can be pulled out of the interaction area in a fully automated procedure. Due to its high resolution and proximity to the interaction region, the VELO is essential for reconstructing primary and secondary vertices and for differentiating many close tracks near the interaction point. For high-energetic particles, an impact parameter resolution regarding the PVs of less than 35 μm is reached.

The large LHCb dipole magnet [90] deflects charged particles in the xz -plane with an integrated magnetic field of 4 Tm. The magnetic field is measured and serves as input for the reconstruction algorithms. The magnet can be operated with two opposed polarities, called MagUp and MagDown. When taking data, the polarity is changed regularly. In the analyses presented in this thesis, the two resulting data sets are not treated separately but are always combined. This

step is justified since the decay modes under study have symmetric final states, and no differences between both data sets are expected.

The tracking system is completed by the planar tracking stations TT and T1–T3. The three stations downstream of the magnet are composed of an inner part, the Inner Tracker (IT) [91], and an outer part, the Outer Tracker (OT) [92, 93], with the former being closer to the beam pipe. The TT and the IT are built as silicon trackers with four layers of silicon micro-strip sensors in each station. To increase the spatial resolution in y direction, the second and third layers are rotated into a stereo configuration with angles of $+5^\circ$ or -5° . In contrast to the silicon trackers, the OT is a drift-tube gas detector with 200 gas-tight straw-tube modules and a mixture of Argon, CO_2 , and O_2 as counting gas. The OT is significantly larger than the IT but also with smaller hit resolution, which is a design choice justified by the lower particle flux in the outer detector regions.

Tracks can be reconstructed with information from all subdetector systems or only from some of them, resulting in different track types. In Fig. 3.3, a sketch of the track types at LHCb is shown. For the analyses presented in this thesis, only tracks reconstructed with information from all tracking detectors (the long tracks) are considered since these tracks have the highest precision.

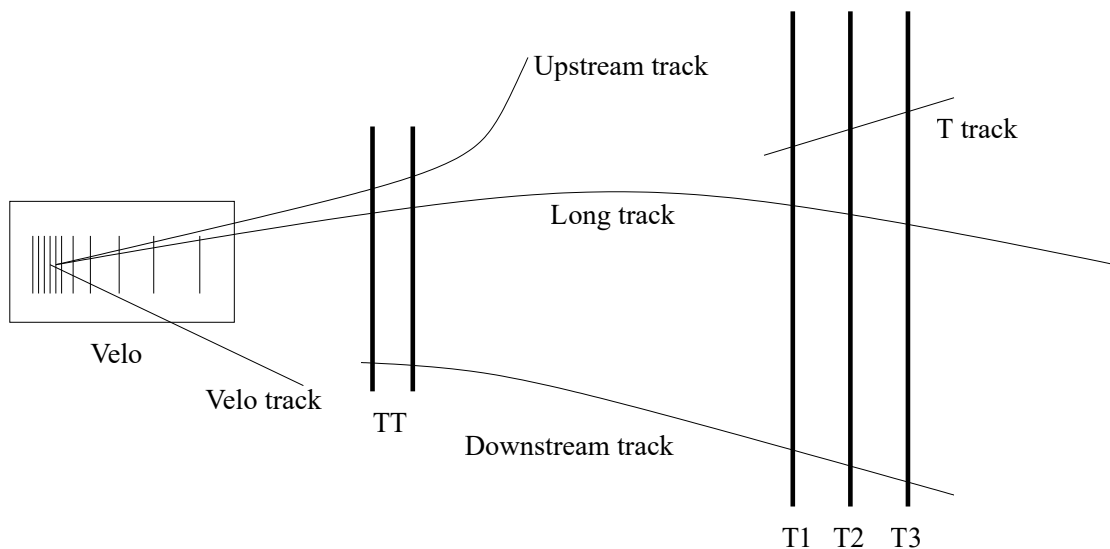


Figure 3.3 – Sketch of the LHCb tracking system. The VELO and the four tracking stations are shown. The different track types are indicated by lines showing which subdetector information is used to reconstruct the track. Figure taken from Ref. [94].

The LHCb tracking system reconstructs charged particle tracks with a track finding efficiency of 96%. The momentum resolution is measured between 0.5 and 0.8% depending on the particle kinematics, while the resulting mass resolution is determined to be around 0.5% [95].

Despite this good performance, the wrong assignment of hits to a track may lead to misreconstruction and must be addressed in an analysis when calculating the reconstruction efficiency. For the presented analyses, the effect is small due to

the observables of interest being ratios of two decays. Still, it has to be considered for the high-precision measurement of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays, while for the search for $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays, this tracking efficiency correction is omitted.

3.2.2 Identification of muons and protons

The LHCb detector is equipped with a particle identification (PID) system to efficiently assign particle hypotheses to reconstructed track candidates. It is composed of two Ring Imaging Cherenkov detectors (RICH1 and RICH2), a scintillating pad detector (SPD), the preshower detector (PS), an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL), and a system of five muon chambers (M1–M5).

The RICH system [96, 97] aims to differentiate between charged hadrons such as pions, kaons, and protons over a wide momentum range. The separation is done by exploiting the Cherenkov radiation that charged particles emit when passing through a radiator medium at a velocity greater than the phase velocity of light in that medium. The emission angle is given by $\theta = \arccos(1/n\beta)$ relative to the flight direction of the particle, with n being the refraction index of the radiator medium and β being the particle velocity divided by the speed of light. To reach a wide angular acceptance, the RICH1 detector is placed upstream of the magnet as close as possible to the VELO. It covers the low and intermediate momentum region of 2 to 60 GeV/c. The RICH2 detector is installed downstream of the magnet behind the tracking stations and covers the higher momentum region of 15 to 100 GeV/c. As radiator gases, fluorocarbon compounds with low dispersion are used, which is C_4F_{10} with $n = 1.0014$ in RICH1 and CF_4 with $n = 1.0005$ in RICH2. In Run 1, an additional layer of aerogel was placed as a radiator medium at the entrance of RICH1 to improve the identification of slow kaons but removed for Run 2 due to low efficiency. Both detectors have similar optical systems for deflecting the Cherenkov cones to hybrid photodetectors outside the acceptance. The resulting ring images on the detector planes can be translated to particle velocities, which allow for setting mass hypotheses when including the momentum information from the assigned particle tracks.

The main calorimeter system [98, 99] is composed of an electromagnetic and a hadronic calorimeter. Both calorimeters follow a classical shashlik-type structure with alternating scintillator and absorber material. For the ECAL, lead layers are used as an absorber with a total thickness of 25 radiation lengths, while for the relatively small HCAL (its size corresponds to only 5.6 nuclear interaction lengths), iron layers were chosen. Both detectors measure the energy deposit and position of electrons, photons, and hadrons. While the ECAL gives essential input for identifying electrons and photons by analysing their characteristic shower shapes, the HCAL is optimised to provide information on charged and neutral hadrons. Together with the upstream SPD and PS detectors, which support distinguishing electron showers from high-energetic π^0 mesons, particles can

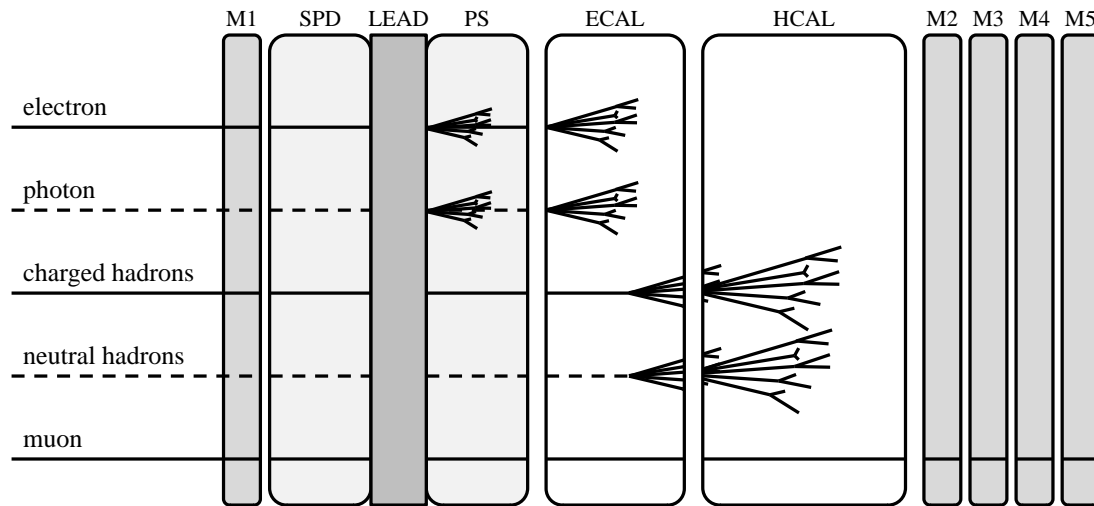


Figure 3.4 – Sketch of the LHCb calorimeter and muon system. For different particle classes, the characteristic shower patterns are shown. Figure refactored from Ref. [100].

be identified based on their different detector signatures, as seen in Fig. 3.4. Moreover, the measurements of the transverse energy E_T in both calorimeters serve as input for the LHCb fast online event selection (see Section 3.2.3).

The PID system is completed by the muon identification system [101, 102], which is of outstanding importance for the two presented analyses. It comprises five stations (M1–M5) arranged along the beam pipe. Station M1 is placed upstream, while station M2–M5 are downstream of the calorimeters. Summing over all stations, the muon system is equipped with 1380 multi-wire proportional chambers. For the innermost region of station M1, gas electron multipliers (GEM) were chosen for their better radiation hardness. As with the calorimeters, the active detector layers are laterally segmented with higher resolution in the inner region and lower resolution in the outer region. They alternate with 80 cm thick iron absorbers to select penetrating muons, resulting in a total absorber thickness of 20 interaction length together with the calorimeter system. Therefore, only muons with a momentum of at least 6 GeV/ c can cross all five stations and leave the detector. The system provides spatial data for the muon track reconstruction and is designed for a muon transverse momentum resolution of 20%. The high muon identification efficiency of around 95% together with a low misidentification rate of around 0.5% makes the detector well suited for analyses with muons in the final state [103]. Background pollution can arise via high-energetic hadrons crossing the calorimeters and reaching the muon stations (punch-through) or via particles decaying within the detector into muons but being unrelated to the physics process of interest (decay-in-flight). Like the calorimeters, the muon system, with its measurement of the transverse momentum of particles, provides important information for the fast online event selection that is described in the following.

3.2.3 Trigger system and B candidate reconstruction

The LHC is able to provide its experiments with a peak instantaneous luminosity of $\mathcal{L} = 1 \times 10^{34} \text{ cm s}^{-1}$ decreasing over the time of a proton fill. For the LHCb interaction point, the luminosity is levelled down to $\mathcal{L} = 4 \times 10^{32} \text{ cm s}^{-1}$ by reducing the overlap of crossing proton bunches. This luminosity levelling is necessary since LHCb is not designed to handle the LHC peak luminosity but profits from more stable data-taking conditions over a longer period, resulting in more stable reconstruction efficiencies. However, with an effective collision rate of 15 MHz in Run 1 and 30 MHz in Run 2 and an average event size of 60 to 70 kB [104], the LHC is still able to provide more data than the detector can read out and store on tape [105–107]. To reduce the event rate and retain only events with valuable content for the LHCb physics programme (like events with B meson decays), the detector is equipped with a system for fast online event selection: the trigger [108, 109].

The system consists of a hardware trigger (L0) and a two-staged high-level software trigger (HLT1 and HLT2). The trigger is configured via trigger lines, which are named sets of requirements the different trigger stages evaluate on the event data. Only events that pass at least one trigger line at every stage are finally stored on tape. The L0 trigger criteria are implemented in the hardware of the calorimeter and muon station read-out systems. At a very early data acquisition stage, the L0 selects events by requiring a minimum transverse momentum p_T of particles in the muon system or a minimum transverse energy E_T of particles in the calorimeter. To further reduce the event rate, the trigger tries to reject events containing multiple proton-proton collisions (pile-up) or events with high track multiplicity for which the event reconstruction would take too long. The former is done by introducing a pile-up veto based on the reconstruction of all PVs in the VELO [110], while the latter is done by setting upper limits on the number of hits in the SPD. These requirements lower the event rate to 1 MHz. The second part of the trigger is implemented in software running on an event filter server farm. The first stage, the HLT1, performs inclusive selections on one- and two-track signatures from a partial event reconstruction with only a subset of the detector information. Also, the PV reconstruction is done at this stage. Decisions are based on properties like track or vertex fit qualities (usually quantified by the corresponding fit χ^2), the p_T of particles or the displacement of tracks from any primary vertex. Especially the latter is an essential feature in selecting tracks from detached heavy flavour B decays. With an input rate of 110 kHz in Run 2 and by using information from the HLT1, the subsequent HLT2 stage finally reconstructs the entire event allowing to set requirements on its topology and other more complex analysis-specific variables. With an event rate of 12.5 kHz in Run 2, the events passing HLT2 go to tape for later analysis.

All software for the trigger configuration, event reconstruction, and creation of the analysis-ready data samples is developed under the Gaudi framework [111]. The trigger configurations for the HLT1 and HLT2 stages are implemented in

the Moore package [112]. The reconstruction sequences for the proto-particle candidates having a track and a particle hypothesis assigned are defined with the Brunel software [113]. These particle candidates are persisted on tape, filtered by a centralised loose preselection, and combined with the DaVinci software [114] into analysis-specific decay candidates. The data of these final candidates are then stored on disk as Root files [115] that can be used for a high-level data analysis.

3.3 Introduction to LHCb simulation

Next to the data, the detector records, Monte Carlo simulated samples of signal and specific background modes are produced to study detection efficiencies and frequency distributions for different variables. The software packages for producing the simulation samples are bundled in the Gauss framework [116]. The actual collisions are simulated with the Pythia8 package [117] using a dedicated LHCb configuration [118]. For the decay of (intermediate) particles, the EvtGen package [119] is used, while the Photos package [120] takes radiative corrections to the final-state particles and resonances into account. The propagation through and interaction with the detector material is simulated by the Geant4 toolkit [121]. In the last step, the Boole software [122] takes on digitising the energy deposition of particles in the detector cells. After this, simulated events are treated exactly as real data with the further processing as described above.

From an analysis perspective, it is important to notice that the LHCb simulation is not a perfect representation of the signal in data. Therefore, as seen in both analyses, a significant part of the analysis procedure is understanding the simulation and calibrating it according to well-known data control channels. The calibrations are done in a way that the selection efficiencies and multivariate classifier outputs can be trusted or by calculating estimates on the systematic uncertainties that arise from the simulation imperfections.

3.4 LHCb terms and variables

Data passing the trigger is stored on tape as explained in Section 3.2.3. Since many different analyses are performed within LHCb, each with different requirements for the reconstructed decay candidates to be examined, the data are sorted using a centralised loose preselection.

Similar to the trigger, this preselection is defined via lines, which are named sets of requirements configured within the DaVinci software. Both analyses presented in this thesis exploit lines that select events with a dimuon pair, good-quality tracks, and a vertex displaced from any PV. Details on the preselection and trigger requirements chosen for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ and $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ analyses are given in Section 4.2 and Chapter 8. For the preselection and subsequent

high-level data analysis, the DaVinci software provides a variety of variables that are calculated for each decay candidate, some of which have a specific definition in the LHCb context:

- After reconstructing candidates for the decay of interest from trigger accepted events, a trigger category is assigned to each candidate and for each trigger line indicating whether the event was triggered only by the reconstructed signal tracks (TOS), independent of the signal tracks (TIS) or explicitly by both (TOB). Figure 3.5 shows the logic behind this scheme. The categories are used to measure the trigger efficiencies on data via the TISTOS method (see Section 7.2.2) or to study the trigger response for different particles. Since the muon identification rate at LHCb is high, the muon and dimuon triggers are very efficient. Therefore, the category TOS is often preferred for decays with muons in the final state. However, as seen in the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ analysis, no specific trigger category is sometimes chosen in the signal candidate selection to analyse the full trigger output and maximise the data set statistics.
- The PID system offers two different approaches to quantify the probability that a particle candidate belongs to a particular species. The first approach uses log-likelihood differences between the hypothesis of a candidate to be a true particle X and the hypothesis of this particle to be a pion. These log-likelihood differences are provided independently by the RICH, calorimeter and muon systems and are combined to get one analytical variable:

$$\text{PID}_X = \sum_i \ln \mathcal{L}_i(X) - \ln \mathcal{L}_i(\pi) \quad (3.1)$$

The second approach combines track reconstruction and PID information of simulated events to train neural networks optimised to distinguish the different particle species. One neural network each is trained for kaons, pions, protons, muons, electrons, and ghost tracks. Ghost tracks result from the incorrect combination of unrelated hits in the detector and are treated as background. The neural networks return as output variables probabilities for a particle candidate to be a true particle X . These output variables are referred to as ProbNN(X).

- Another variable for muon identification is named isMuon [103]. This binary variable is set depending on the number of muon stations that registered a hit within a field of interest around the track extrapolation. The number of stations required for a positive decision depends on the particle momentum, while for the size of the field of interest, the expected muon scattering within the stations is considered. The variable provides a powerful complementary approach to the muon PID evaluation, as it is not used in the neural networks described above.

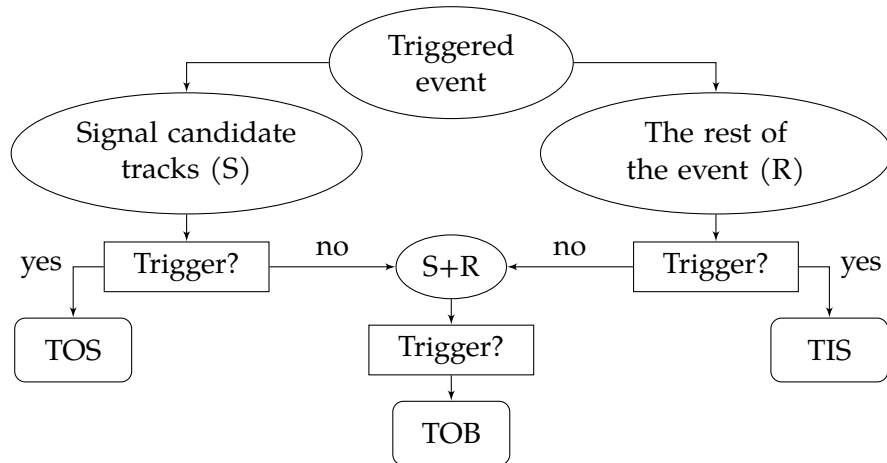


Figure 3.5 – Logic behind the LHCb trigger categories. After the decay candidate reconstruction, a trigger category (TOS, TIS, TOB) is assigned to each candidate and for each trigger line. Figure refactored from Ref. [123].

- For simulated samples, truth-matching information is provided, given that the properties of the truly generated decays are known. After reconstruction, all simulated decay candidates are assigned to a background category that indicates how closely the reconstructed candidate matches the generated decay. A category with a low number means a larger match. The categories relevant for the presented analyses are categories 0 and 10, which contain the correctly reconstructed decays, and category 50, which contains candidates whose reconstructed B mass is too low due to bremsstrahlung losses.

Other variables used in the analyses are listed and defined in Table 3.1.

3 Experiment

Table 3.1 – Glossary of variable definitions used in the selection of signal candidates.

Variable	Explanation
$m(ab)$	Invariant mass of the combination of particle a and b
Q	Electric charge of the particle
p	Absolute momentum of the particle
p_T	Transverse momentum of the particle, which is the component of the total momentum transverse to the beam pipe
η	Pseudo-rapidity, which is defined as $\eta = -\ln \tan(\theta/2)$, where θ is the angle between the momentum of the particle and the beam pipe. The geometrical LHCb acceptance is $2 < \eta < 5$
$\chi_{\text{IP}}^2(\text{PV})$	Difference in the χ^2 of the PV fit when fitting with or without the particle track/candidate
$\chi_{\text{FD}}^2(\text{PV})$	Distance of the decay vertex of the particle with respect to the PV in units of χ^2
χ_{vertex}^2	χ^2 of the vertex fit of the particle
χ_{track}^2	χ^2 of the track fit of the particle
DIRA	Cosine of the flight direction angle, which is defined as angle between the momentum vector of the particle and the vector, which points from the origin vertex to the decay vertex of the particle
DOCA χ^2	Distance of closest approach between two particles in units of χ^2
hasRich/hasMuon	Decision whether the RICH/muon chamber system registers a track in the event
GhostProb	Probability of a track to be a ghost track
isMuon	Decision according to the number of muon stations where a hit is found within a field of interest around the track extrapolation
$a \text{ PID}_b$	Measure for the probability of particle a to be particle b calculated as difference of log-likelihoods
$a \text{ ProbNN}(b)$	Probability of particle a to be particle b calculated by a neural network
nTracks	Number of fitted tracks in the event
dof	Number of degrees of freedom

Part II

Measurement of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays

4 Analysis introduction

The first presented analysis aims to measure the branching fractions of the rare electroweak $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays using the entire proton-proton collision data set recorded by the LHCb detector in Run 1 and Run 2 of the LHC. In addition, a determination of the effective $B_s^0 \rightarrow \mu^+ \mu^-$ lifetime and a first search for the $B_s^0 \rightarrow \mu^+ \mu^- \gamma$ decay without explicit photon reconstruction are performed.

The analysis was carried out by a group of 14 people who actively contributed to the internal analysis documentation. The presented thesis concentrates on the author's individual contributions. It includes studies on differences between recorded and simulated data samples, background contributions to the signal channel, and improvements for the branching fraction normalisation. However, essential findings by the other proponents and results from previous analyses are summarised if needed to ensure comprehensibility and provide context. Emphasis is laid on clarifying those contributions made by others.

Section 4.1 describes the analysis strategy for the branching fraction measurement, followed by a brief discussion of the signal selection in Section 4.2 and the simulation calibration in Section 4.3. Chapter 5 presents work for the background estimation in the signal channel. Finally, in Chapter 6, the branching fraction calculation is outlined, and improvements in the signal efficiency determination are explained in more detail. The chapter concludes with the presentation and interpretation of the analysis results.

According to the focus of this thesis, which is the branching fraction measurement of $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays, both the measurements of the effective $B_s^0 \rightarrow \mu^+ \mu^-$ lifetime and the branching fraction of $B_s^0 \rightarrow \mu^+ \mu^- \gamma$ decays are not discussed. Instead, these aspects are examined in Refs. [86, 124–126].

The analysis is summarised in two journal publications: A short presentation of the main results is given in Ref. [16], while a more detailed and extended discussion can be found in Ref. [17].

4.1 Analysis strategy

The core strategy for the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fraction measurement is adopted from previous LHCb analyses of these decay modes. Compared to the previous measurement [64], the analysed data set size is doubled. The past analysis exploited the complete Run 1 but only 25 % of the current Run 2 data set. However, the analysis now contains the complete set of proton-proton collision data collected at centre-of-mass energies of 7, 8, and 13 TeV. The data

size corresponds to an integrated luminosity of 9 fb^{-1} . Next to the recorded data, simulated events are used to study signal selection efficiencies, distributions of decay candidates in different variables, and prominent exclusive background modes.

The $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fractions are calculated relative to the branching fractions of the two normalisation modes $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow K^+ \pi^-$, which have high yields in the LHCb data. The process $J/\psi \rightarrow \mu^+ \mu^-$ is always assumed. For both normalisation modes, the branching fractions are determined with high precision. The $B^+ \rightarrow J/\psi K^+$ mode is chosen because its two final state muons lead to similar trigger and particle identification system responses as in the signal channel. In contrast, the $B^0 \rightarrow K^+ \pi^-$ mode is chosen for its abundance, yielding a small statistical uncertainty on the normalisation, and its simple two-body decay topology, which it has in common with the signal modes.

4.2 Selection of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ candidates

To separate the signal events from the different background processes, a dedicated signal selection is implemented. The selection strategy for the signal and normalisation channels consists of several parts: a trigger selection, requirements on the topology and kinematics of the event, requirements on particle identification variables, and the selection by a multivariate classifier.

Regarding the trigger, the most inclusive selection approach is chosen for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ channels. No specific set of trigger lines is required for accepting an event, but all events passing the trigger are kept. This strategy was developed for earlier studies of the decay modes and is motivated by the high signal efficiencies for $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ decays of 96 % and 90 %. At the same time, the strong background suppression of the remaining selection supports this strategy. The observed high signal efficiencies are due to the excellent performance of LHCb's muon identification system. The differences between the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ channels are mainly caused by the kinematics of the modes, with the muons from the J/ψ being less energetic on average and, therefore, harder to detect.

For the $B^0 \rightarrow K^+ \pi^-$ channel, where no muons are present in the final state, an inclusive TIS selection at the L0 stage and a selection of exclusive TIS lines at the HLT1 stage are developed. This new trigger strategy, independent of the signal candidate, reduces the differences between the signal and the $B^0 \rightarrow K^+ \pi^-$ trigger lines and yields sufficient statistics for the normalisation. Compared to the previous analysis [64], where a more inclusive approach was chosen, the strategy allows for more precise efficiency estimates in the $B^0 \rightarrow K^+ \pi^-$ channel. For the HLT2 stage, an exclusive TOS trigger selects events with a two-body decay topology. The $B^0 \rightarrow K^+ \pi^-$ trigger efficiencies for all three trigger stages are measured using the data-driven TISTOS method [123] (explained in Section 7.2.2) and yield around 4.3 % on Run 1 and 7.3 % on Run 2 data.

The data that passes the trigger is further reduced with the selection being adopted from the previous analysis without changes. On top of the centralised preselection, additional loose requirements are imposed to select decay candidates with high-quality reconstructed tracks and vertices. Only candidates with SVs that are clearly displaced from any PV are chosen. Additional requirements, *e.g.* on the invariant dimuon mass and the muons' transverse momentum, are applied to shrink the phase-space volume and reduce the data size while keeping the signal efficiency high. Further background is removed by applying a loose requirement on the response of a first multivariate classifier, which has a high $B_s^0 \rightarrow \mu^+ \mu^-$ signal efficiency of 92 % measured on simulation. The classifier mainly exploits topological variables, *e.g.* the angle between the B momentum direction and the direction defined by the vector joining PV and SV, the impact parameter of the B candidate, and the vertex χ^2 of the SV.

All muon tracks are required to satisfy the isMuon condition and the more restrictive requirement

$$\text{ProbNN}_{\text{comb}} = \text{ProbNN}(\mu) \times (1 - \text{ProbNN}(K)) \times (1 - \text{ProbNN}(p)) > q, \quad (4.1)$$

with q being dependent on the data-taking period. The approach aims at efficiently selecting muons while rejecting backgrounds from misidentified kaons and protons using the ProbNN and isMuon particle identification variables presented in Section 3.4. Although it was optimised for the $B^0 \rightarrow \mu^+ \mu^-$ mode in Ref. [127], a selection review is performed as part of the presented analysis but no improvement regarding the choice of q is found. The PID signal selection efficiency is evaluated on data using the LHCb-wide hosted PIDCalib package [128] together with large calibration samples of $J/\psi \rightarrow \mu^+ \mu^-$ decays, in which the J/ψ mesons originate from inclusive b decays. For the Run 1 data, the PID signal efficiency of the above selection is measured to be 81 %, while for the Run 2 data, it is 84 %.

A veto on the dimuon mass is introduced to remove combinations of two muons with an invariant mass that falls into the J/ψ mass region. The veto suppresses background contributions like $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$ decays. Figure 4.1 shows the distribution of candidates rejected by the J/ψ veto in Run 2 data in the plane of the dimuon mass and the main multivariate classifier (described below) after applying kinematic and acceptance cuts. The veto efficiency on $B_s^0 \rightarrow \mu^+ \mu^-$ signal is about 99.8 % in Run 1 and 99.7 % in Run 2. At the same time, the number of rejected signal candidates is negligible in all but the lowest region of the main multivariate classifier.

The last part of the selection is a multivariate classifier with high separation power, implemented as a Boosted Decision Tree (BDT) [129]. This BDT is trained with $B_s^0 \rightarrow \mu^+ \mu^-$ signal simulation against $b \rightarrow \mu \mu X$ simulation and exploits the decay topology. Key input variables are two track isolation variables based on individual BDTs that classify the separation of the signal muon candidates from other tracks in the event and yield a high separation between signal and back-

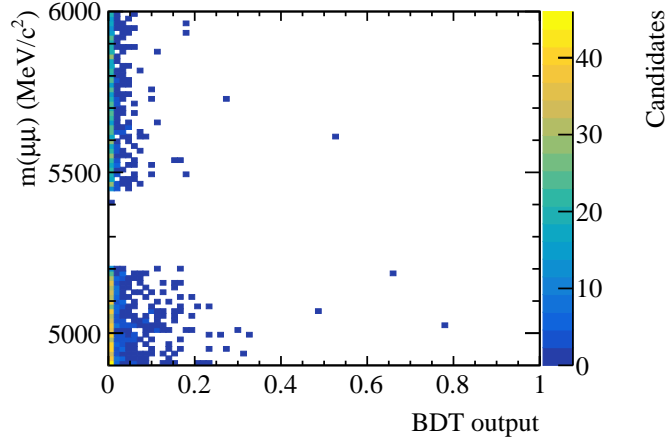


Figure 4.1 – Invariant dimuon mass as a function of the signal BDT output for candidates fulfilling the kinematic and acceptance requirements but being rejected by the J/ψ veto. The signal region is excluded. The distribution is shown for the Run 2 data set.

ground candidates. Instead of selecting the data based on the classification, the BDT output is used to bin the data. By transforming the BDT output distribution to be flat over the range from zero to one in simulation, the signal efficiency is kept stable between all bins and the most background is gathered in the lowest BDT bin. Then, the branching fractions are obtained by extracting the number of signal candidates from simultaneous fits to the invariant dimuon mass in bins of the BDT output. All bins except for the lowest one are used in the branching fraction fit. This strategy maximises the sensitivity to the rare modes with respect to a strategy that applies a requirement on the signal BDT output rejecting true signal decays.

4.3 Simulation calibration

It is known that the distribution of signal events in several selection variables does not match perfectly between data and simulation. The differences must be corrected if reliable information such as signal efficiencies, BDT fractions, or signal mass distributions are to be obtained from simulation.

In the course of this thesis, differences in kinematic and topological variables are examined. To calibrate these variables, dedicated per-event weights are calculated for the simulated samples using gradient-boosted reweighters (GBR) [130] trained for each data-taking year individually. A GBR is based on an ensemble of decision trees that are successively applied to the difference between data and simulation. In this process, weights based on the gradient of data-simulation similarity are determined at each iteration to improve the reweighting. For the reweighting, a method has to be applied to access the signal distributions in data.

To achieve this, the sPlot technique [131] is used. It can be applied to control channel data in which the distributions of signal and background in at least one variable is known. The technique uses a fit to these known distributions to extract weights (called sWeights) that can be used to project the signal distribution out of any other variable that is not correlated with the fit variable. Since the hadronisation processes for B_s^0 and B^0 mesons are different, the reweighting for the two signal channels is done with different control modes. For $B_s^0 \rightarrow \mu^+ \mu^-$, the $B_s^0 \rightarrow J/\psi \phi$ channel is chosen, while for $B^0 \rightarrow \mu^+ \mu^-$, the $B^+ \rightarrow J/\psi K^+$ channel is used assuming the b quark hadronisation fractions f_u and f_d to be the same. The fit is performed on the reconstructed B mass distribution, with the invariant dilepton mass constrained to be the J/ψ mass increasing the mass peak resolution.

The kFolding method [132] with $k = 5$ folds is used to prevent training on statistical fluctuations of the input data. This method divides the input data into k parts, training the reweighter with $k - 1$ parts and validating the result using the remaining part. A rotation of the assignment to the training and test data set results in k trained versions of the same reweighter. For the weight predictions on independent data sets, the mean prediction of all reweighters is used. The reweighters take the transverse momentum p_T , the pseudorapidity η , and the χ_{IP}^2 of the B candidate as training inputs. These variables have been chosen since they are largely uncorrelated, and a disagreement between the signal in data and simulation is prominent. The use of the track multiplicity variable nTracks, which is also insufficiently described in simulation, is investigated. However, a strong correlation with the signal BDT is found, leading to a separate correction for the track multiplicity, as described below. The distributions from the simulation are reweighted to match the distributions of the signal in data. The input distributions used in the reweighting are shown in Fig. 4.2, where sWeighted data is compared to unweighted and reweighted simulation. After the reweighting, a better agreement of the distributions can be seen.

Since fewer simulated events are available for the $B_s^0 \rightarrow J/\psi \phi$ mode than for the $B^+ \rightarrow J/\psi K^+$ mode, the B_s^0 reweighter is not trained for each year individually but for the two-year periods 2011/2012, 2015/2016 and 2017/2018. This procedure is done to get higher stability in the reweighter training. As the algorithm does not provide an uncertainty on the GBR weights taking the limited statistics of the training samples into account, additional reweighters with half and a quarter of the training statistics are trained to estimate the variation. These variations are later considered in the uncertainty calculations of the measured branching fraction.

The trigger selection efficiency can be calculated using simulated samples. However, the accuracy of this method is limited as the trigger response is not perfectly described in simulation. Therefore, the efficiency is determined using the already mentioned TISTOS method [123] on control channel data. The method itself is explained in Section 7.2, where the trigger calibration for the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decay is discussed. For the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ and $B^+ \rightarrow J/\psi K^+$ modes in this analysis, the calibration is performed with data of reconstructed and selected $B^+ \rightarrow J/\psi K^+$

4 Analysis introduction

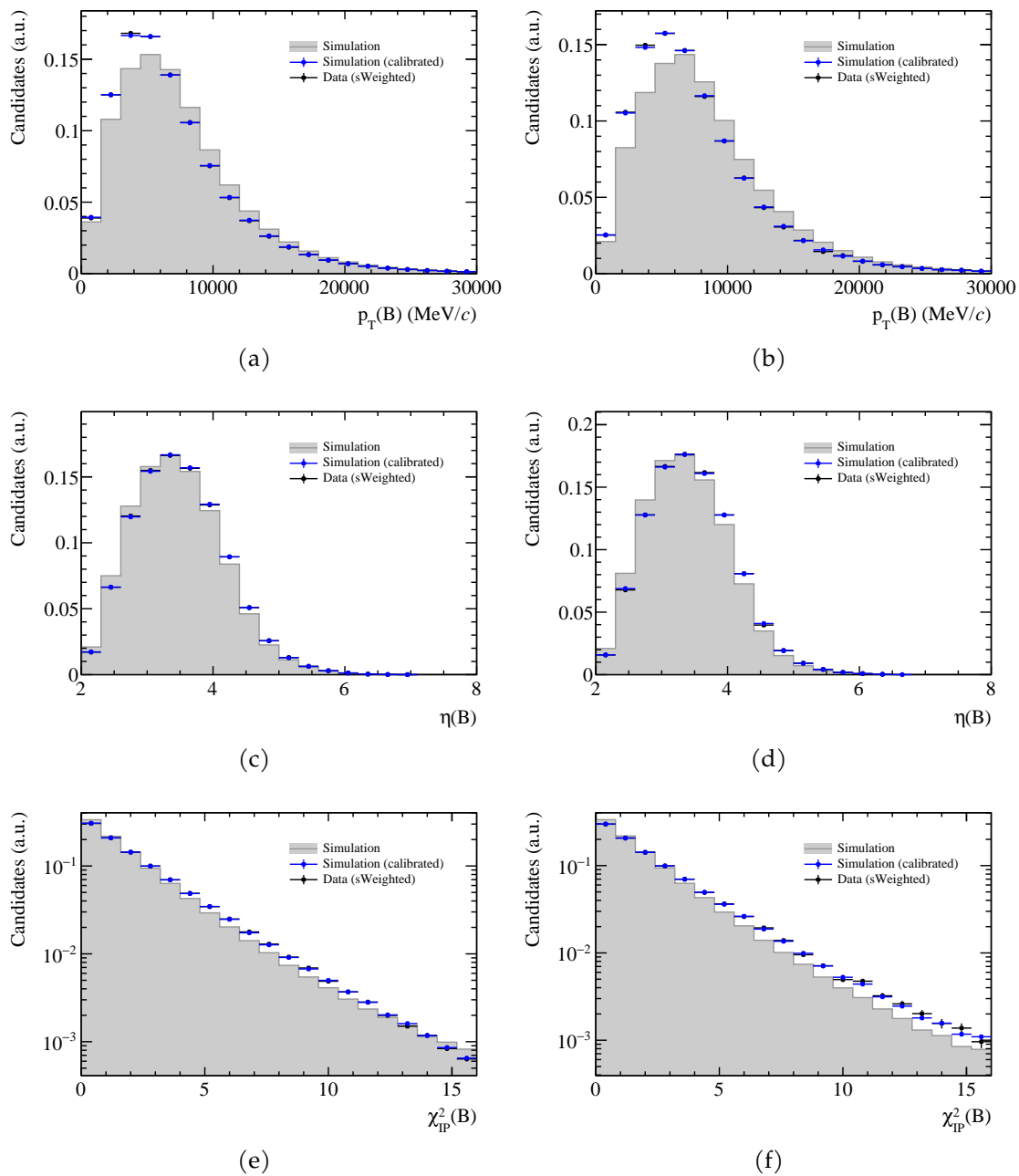


Figure 4.2 – Input distributions for the gradient-boosted reweighters, which are trained in the (left) $B^+ \rightarrow J/\psi K^+$ and (right) $B_s^0 \rightarrow J/\psi \phi$ decay channel. The sWeighted data is compared to plain and calibrated simulation. The 2018 data set is shown.

events. The efficiencies are calculated as a function of the maximum muon transverse momentum and the product of the two muon transverse momenta as these are the variables the L0 muon trigger selects on. Then, the resulting efficiency maps are convoluted with the kinematics of the simulated samples to get an integrated trigger efficiency per mode. The procedure assumes the signal and the rest of the event are kinematically independent. However, this is not entirely given, leading to a separate calibration in different kinematical bins. For the $B^0 \rightarrow K^+\pi^-$ mode, the candidates are triggered independently of the signal using a TIS selection at the L0 and HLT1 stage, as mentioned in Section 4.2. The approach to compute the calibrated trigger efficiency follows the one for the muon modes. Instead of using the muon kinematics, the efficiency maps are calculated with the TISTOS method in bins of the momentum and the pseudorapidity of the parent B meson. In contrast, the HLT2 efficiency is taken directly from simulation as it is found to be sufficiently well described.

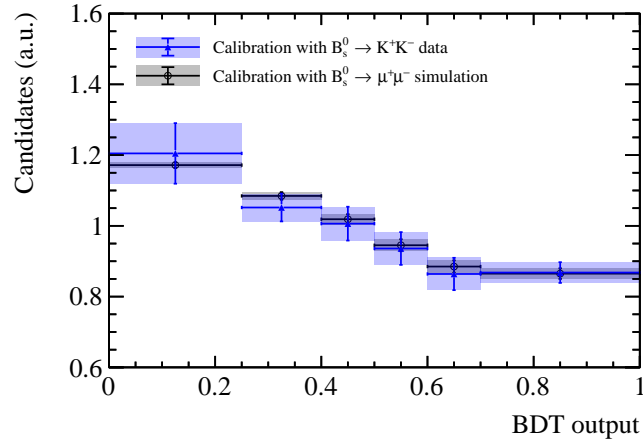
Besides the kinematic and trigger corrections, further calibrations are applied: Corrections to the PID variables and misidentification rates are discussed in Section 5.1. A dedicated method to calibrate the track reconstruction efficiency is presented in Section 6.1.2.

For the $B_s^0 \rightarrow \mu^+\mu^-$ channel, an additional correction is applied to account for the relative decay-width asymmetry measured for the heavy and light B_s^0 meson mass eigenstates [133, 134]. Both states have different lifetimes. However, in the SM, only the heavy mass eigenstate decays into two muons. In simulation, an averaged lifetime is used for the B_s^0 decay. The reconstruction and selection depend slightly on the lifetime. Therefore, correction factors for the decay time selection efficiency are derived from the known decay-width asymmetry in bins of the signal BDT output. In the B^0 system, the decay-width asymmetry is close to zero, which leads to a negligible lifetime difference and no correction.

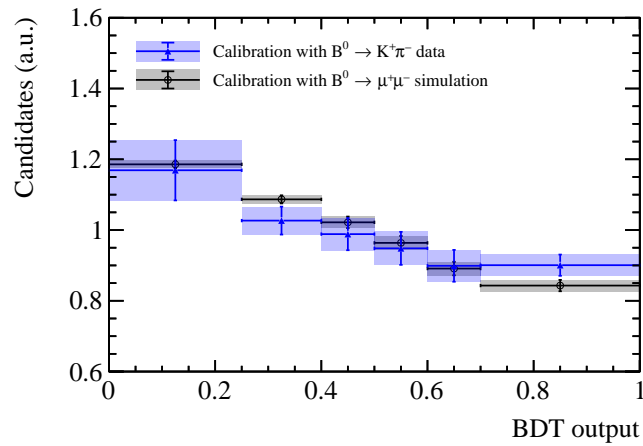
For the branching fraction fit, the $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ mass distributions are modelled by double-sided Crystal Ball functions [135] (see Section 6.2). The mean and width of these models are calibrated on control channel data. The mean is estimated from $B_s^0 \rightarrow K^+K^-$ and $B^0 \rightarrow K^+\pi^-$ decays, respectively, while the width is derived from the interpolation of the mass resolution between different charmonium and bottomonium resonances. The tail parameters of the Crystal Ball functions are determined from $B_{(s)}^0 \rightarrow \mu^+\mu^-$ simulation taking the calibrated mass resolutions into account.

A final correction is applied to model the BDT fractions correctly. The BDT fractions describe the distribution of candidates over the full BDT range and are calculated in bins of the signal BDT output. They can be cross-checked with data by comparing to the $B^0 \rightarrow K^+\pi^-$ and $B_s^0 \rightarrow K^+K^-$ signal in control channel data. After applying all corrections described above, the BDT fractions are still insufficiently well described in simulation. The remaining discrepancies result from the improper description of the event multiplicity and the correlation of the BDT output distribution with the non-signal part of the event, which is introduced by the isolation variables in the BDT training. The differences

are addressed by performing a further two-dimensional reweighting, using the nTracks distribution and the BDT output distribution for binning. The correction factors are derived by comparing the distributions in $B^+ \rightarrow J/\psi K^+$ data and simulation. Then, the factors are applied to the two-body signal and control channels. Figure 4.3 shows the BDT output distribution in the signal channel data. The calibration is obtained in two ways, either with $B_s^0 \rightarrow K^+ K^-$ and $B^0 \rightarrow K^+ \pi^-$ data or with $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ simulation. In both cases, all described corrections are applied. Both methods have a high agreement, whereas the calibration with the simulated samples shows significantly smaller uncertainties. Indeed, it turns out that the uncertainty on the BDT calibration can be reduced by a factor of ten down to 1% using the simulation-based method performed in this analysis for the first time.



(a)



(b)

Figure 4.3 – BDT output from calibrated (a) $B_s^0 \rightarrow K^+K^-$ and (b) $B^0 \rightarrow K^+\pi^-$ channel data compared to either $B^0 \rightarrow \mu^+\mu^-$ or $B_s^0 \rightarrow \mu^+\mu^-$ calibrated simulation. Within the uncertainties, the distributions are consistent. The calibration based on the simulated samples shows ten times smaller uncertainties and is used for the analysis. The plots for the Run 2 data set are shown.

5 Background studies

This chapter presents the evaluation of background contributions to the signal channel. Despite an optimised $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ signal selection with strong background suppression, several irreducible background decays are left in the data set and must be estimated and modelled for the branching fraction fit. The prominent background modes considered in the analysis are:

- $B_{(s)}^0 \rightarrow h^+ h^{(\prime)-}$ decays, with $h^{(\prime)}$ being real kaons or pions wrongly identified as muons, which contaminate the B^0 signal region;
- $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays, which pollute the left mass sideband due to a pion misidentified as a muon and an undetected neutrino;
- $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decays, which pollute the left mass sideband due to a kaon misidentified as a muon and an undetected neutrino;
- $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ and $B^0 \rightarrow \pi^0 \mu^+ \mu^-$ decays, with two real muons in the final state, which contribute to the left mass sideband;
- $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ decays, with a proton misidentified as a muon, which mainly populate the left sideband, but also contribute to the signal mass region;
- $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$ decays, with one muon from the J/ψ and one from the B_c^+ decay forming a good vertex. The background mainly populates the left sideband but also contributes to the signal mass region. Studies on this decay can be found in Ref. [86];
- $B_s^0 \rightarrow \mu^+ \mu^- \gamma$ decays, with two real muons in the final state and an escaped photon contaminating the left mass sideband. The decay rate study is beyond this thesis's scope but is covered in the presented LHCb publication as an additional observable. Details can be found in Ref. [125].

In addition, the fit contains an exponential component per BDT bin to cover combinatorial background that results from random combinations of unrelated muons in the detector passing the selection. The fit directly determines the contribution of the combinatorial background via these free components.

The work presented in this thesis comprises investigations for the above semileptonic decays resulting from a single hadron misidentification or missing particles in the final state. However, for completeness, also findings for the fully hadronic background contributions are briefly summarised, together with a description of a new data-driven strategy for obtaining the hadron-muon misidentification rate.

5.1 Hadronic backgrounds

The $B_{(s)}^0 \rightarrow h^+ h'^{-}$ backgrounds pollute the signal region when both hadrons are wrongly identified as muons. They are challenging since the backgrounds are peaking near the B^0 mass, limiting the sensitivity on the $B^0 \rightarrow \mu^+ \mu^-$ mode. Therefore, a reliable determination of the hadron misidentification rates is essential to disentangle the signal and background components in the fit correctly.

The PIDCalib package [128] was used to calibrate the misidentification rates in the previous analyses. The LHCb-wide hosted package exploits clean calibration samples of $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+) \pi^+$ decays to obtain $\pi \rightarrow \mu$ and $K \rightarrow \mu$ misidentification maps in bins of the hadron's momentum and transverse momentum. By applying the maps to the kinematics of the background channels, data-driven estimates of the misidentification rates can be made, and systematic uncertainties can be attached, with the latter being considerably large in previous analyses.

However, this update analysis found that the PIDCalib approach constantly underestimates the kaon and pion misidentification rates when applying a strict muon PID requirement to the hadrons. The problem originates mainly from low-energetic kaons and pions decaying to muons within the detector according to $\mathcal{B}(\pi \rightarrow \mu\nu) \approx 99\%$ and $\mathcal{B}(K \rightarrow \mu\nu) \approx 64\%$ [27]. These decay-in-flight hadrons cause a momentum loss in the reconstruction due to the missing neutrinos and create a sizeable tail in the invariant $K\pi$ distribution, which becomes visible when applying the muon selection. For pions, the tail is shown in Fig. 5.1a. At the same time, the PIDCalib approach relies on the sPlot method to extract information about the pure signal kinematics from the calibration channel. It assumes that

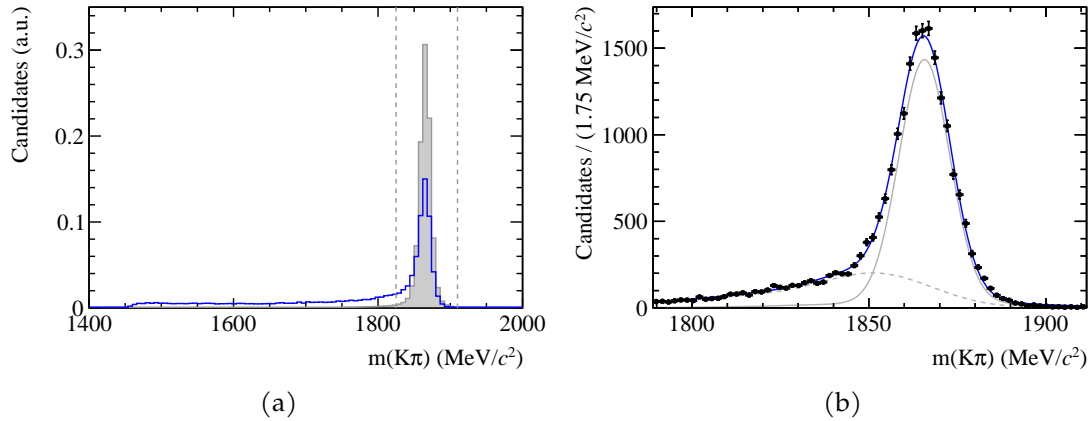


Figure 5.1 – (a) Invariant $K\pi$ mass from $D^0 \rightarrow K^- \pi^+$ simulation with no PID selection applied (grey) and with the muon selection applied to the pion (blue). The dashed lines mark the PIDCalib mass window. (b) Fit to the $D^0 \rightarrow K^- \pi^+$ calibration sample with the muon selection applied to the pion. The solid grey component corresponds to the insufficient model PIDCalib is using. The dashed grey component describes the contribution from decay-in-flight pions.

the invariant $K\pi$ mass shape and the muon selection are independent and that all calibration events are gathered in the fit mass range. Both conditions are no longer met, causing the sPlot method to fail.

A novel procedure is developed to account for the problem: The misidentification rates are now determined from fits to the calibration data sets before and after applying the muon selection following a fit-and-count approach. The fits are done in kinematic bins of momentum and transverse momentum and consider the decay-in-flight hadrons by modelling the distribution's tail in the selected $K\pi$ sample as seen in Fig. 5.1b. The events outside the PIDCalib fit mass range are estimated from large samples of $D^0 \rightarrow K^- \pi^+$ decays, additionally studied for this analysis. Compared to the previous PIDCalib approach, the final effect measured on $B_{(s)}^0 \rightarrow h^+ h^{(\prime)-}$ decays is a more than 50 % increase of the misidentification rate per track. Independent and data-driven cross-checks with $B^0 \rightarrow K^+ \pi^-$ decays are performed, validating the new procedure. The double misidentification probabilities for Run 1 and Run 2 data are calculated separately for the exclusive $B^0 \rightarrow K^+ \pi^-$, $B^0 \rightarrow \pi^+ \pi^-$, $B_s^0 \rightarrow \pi^+ K^-$, and $B_s^0 \rightarrow K^+ K^-$ modes. Based on these numbers, the amount of expected $B_{(s)}^0 \rightarrow h^+ h^{(\prime)-}$ events in bins of the signal BDT output is determined and stated in Table 5.1.

5.2 Semileptonic backgrounds

Next to purely hadronic $B_{(s)}^0 \rightarrow h^+ h^{(\prime)-}$ backgrounds, several semileptonic backgrounds contaminate the signal mass region, especially the lower mass sideband. This thesis investigates the exclusive decays $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$, $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$, $B^+ \rightarrow \pi^+ \mu^+ \mu^-$, $B^0 \rightarrow \pi^0 \mu^+ \mu^-$, and $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$. A discussion of the important background contribution from $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$ decays can be found in Ref. [86].

All studied backgrounds have one or two real muons in the final state. Contamination of the signal channel arises from the partial reconstruction of the decays or misidentification of the hadrons. For all semileptonic background modes, the number of expected candidates is estimated by normalising to the $B^+ \rightarrow J/\psi K^+$ channel as it is done for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ channel (see Section 6.1):

$$N^{\text{expected}} = \frac{f_x}{f_u} \times \underbrace{\frac{N(B^+ \rightarrow J/\psi K^+)}{\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \varepsilon(B^+ \rightarrow J/\psi K^+)}}_{\beta_x} \times \mathcal{B} \times \varepsilon. \quad (5.1)$$

Here, x denotes one of the initial state hadrons (or its relevant quark flavour), making the β factor dependent on the involved hadronisation fractions. The total detection efficiencies ε are estimated from simulation, except for the PID efficiency, calculated with PIDCalib for muons and protons and with the new data-driven approach for pions and kaons. Since it is needed for the branching fraction fit, the background expectations are computed in bins of the signal BDT output. For the B^+ , B^0 , and B_s^0 background contributions, the kinematic

weights derived in Section 4.3 are considered as they substantially affect the signal in high BDT bins. The same procedure as for $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ is applied to the background channels to obtain the weights. As the sum of weights for a channel is not normalised to one, the weights are not used in the absolute detection efficiency calculations. Nevertheless, they are used to correct the different BDT fractions, *i.e.* the distribution of signal events over the full BDT range.

5.2.1 $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$

The semileptonic $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decay has a branching fraction [136] of

$$B(B^0 \rightarrow \pi^- \mu^+ \nu_\mu) = (1.50 \pm 0.06) \times 10^{-4} \quad (5.2)$$

and enriches the background in the signal channel when the pion is wrongly identified as a muon. Due to the missing energy of the escaping neutrino, the decay spectrum is shifted toward lower masses. Still, contributions dominant in the left mass sideband affect the sensitivity to the rare modes by tails into the signal region or at least by correlations with other background components. Consequently, the decay's contribution and invariant mass shape per BDT bin must be estimated and modelled carefully.

The background is examined using simulated $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ samples generated with the latest revision of the simulation software for this analysis. To speed up the generation process, the invariant $\pi\mu$ mass is forced to be above $4500 \text{ MeV}/c^2$. This requirement allows producing more events in a reasonable amount of time, resulting in a data sample corresponding to six million signal events for Run 1 and 16 million events for Run 2.

The simulated samples are produced with the ISGW2 form factor model developed in Refs. [137, 138]. As can be seen in Fig. 5.2a, the generated q^2 spectrum

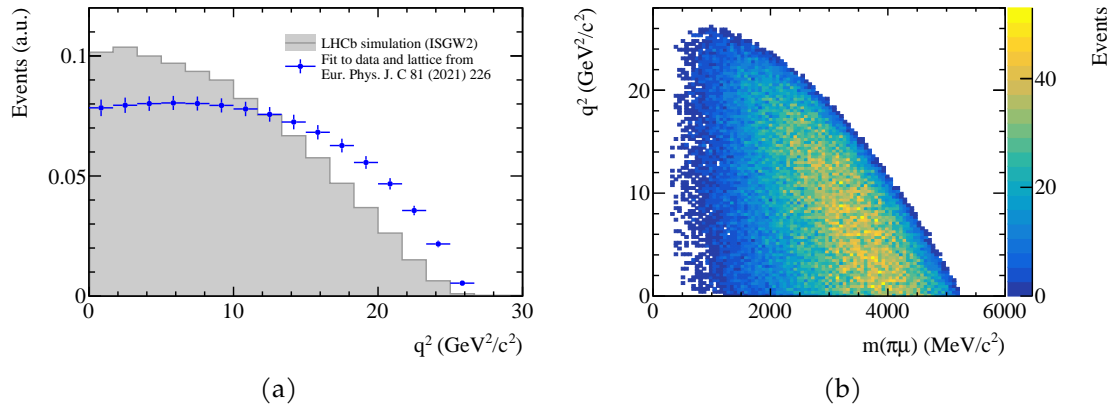


Figure 5.2 – (a) Generated q^2 distribution of $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays using the ISGW2 form factor model and a fit to data and recent lattice QCD results. (b) Generated $\pi\mu$ mass versus generated q^2 distribution in the simulated $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ sample.

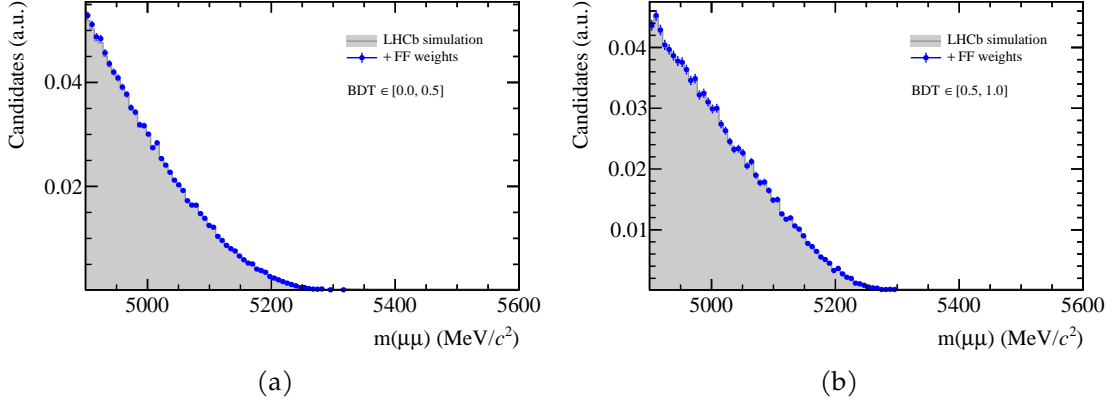


Figure 5.3 – Invariant mass distributions of $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays in the reconstructed dimuon channel with and without the form factor weights. The distributions are shown for (a) $\text{BDT} \in [0.0, 0.5]$ and (b) $\text{BDT} \in [0.5, 1.0]$.

produced with this model disagrees with the spectrum obtained from a fit to recent data and lattice QCD calculations, using the BCL parametrisation presented in Refs. [139, 140]. From bin-wise comparisons, weights are computed to correct the simulated events:

$$w_1 = 0.772 \pm 0.035 \quad (q^2 \in [0.00, 1.67] \text{ GeV}^2/c^2) \quad (5.3)$$

$$w_2 = 0.767 \pm 0.032 \quad (q^2 \in [1.67, 3.33] \text{ GeV}^2/c^2) \quad (5.4)$$

$$w_3 = 0.802 \pm 0.032 \quad (q^2 \in [3.33, 5.00] \text{ GeV}^2/c^2). \quad (5.5)$$

The uncertainty determination on the weights is fully revisited for this analysis. Now, it relies on the limited statistics of the used simulation sample and the uncertainty of the BCL model, which is derived by the following procedure: Different form factors are sampled according to the coefficients and the covariance matrix, both given in Ref. [140]. For each set of form factors, q^2 values are generated from the function of the differential decay width. The number of generated q^2 values is chosen high so that statistical uncertainties become negligible. Then, the q^2 values are binned, and the means and standard deviations of the bin contents are determined across the different form factor samples. It is assumed that the weight uncertainties are fully correlated.

As shown in Fig. 5.2b, only the region with $q^2 < 5 \text{ GeV}^2/c^2$ is relevant for correcting the simulated samples. This region is due to the invariant mass cut on the dimuon candidate at $4900 \text{ MeV}/c^2$, which is applied before the signal fit and strictly constrains the q^2 spectrum.

Following the correction, a study is performed to show that the form factor reweighting has no effect on the $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ mass shape in the dimuon channel. This can be seen in Fig. 5.3, where the invariant mass distributions with and without the form factor weights are compared in two different BDT regions. The

5 Background studies

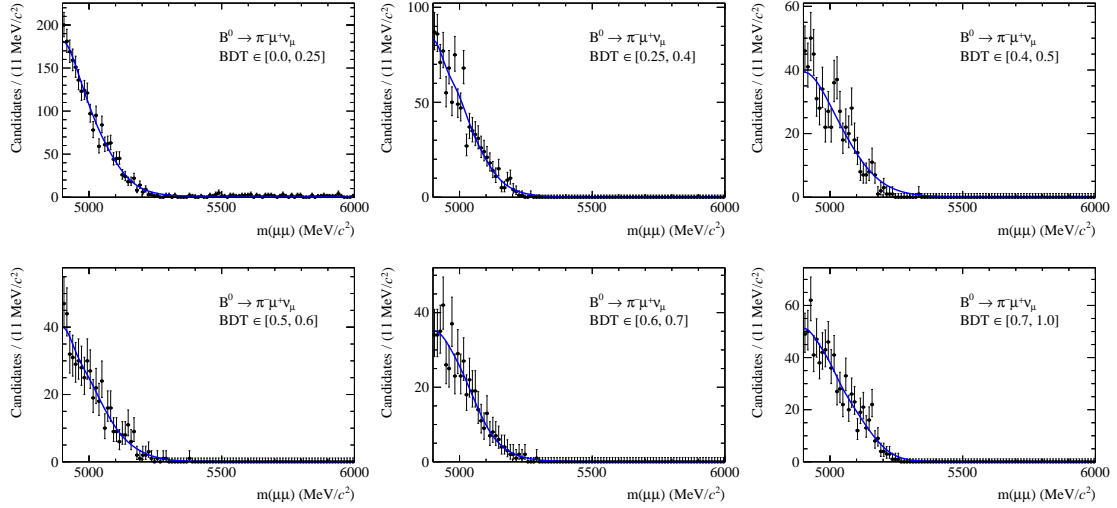


Figure 5.4 – Invariant mass distribution of $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays in the reconstructed dimuon channel and shape fitted to simulation. The distributions are shown separately for the six BDT bins.

invariant mass distributions in bins of the BDT are fitted on simulation with one-dimensional kernel density estimators using Gaussian kernels. In Fig. 5.4, the fit results are shown.

The muon selection efficiency is calculated from data with PIDCalib in bins of the momentum, transverse momentum, and the number of tracks in the event. The same procedure is applied to the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ signals. For the pion, the same approach as for the $B_{(s)}^0 \rightarrow h^+ h^{(\prime)-}$ backgrounds is used.

In Table 5.1, the amount of expected $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ events per BDT bin is listed. The quoted uncertainties include the uncertainties on the branching fraction, the normalisation with $B^+ \rightarrow J/\psi K^+$, and the total efficiencies from simulation and calibration samples; this applies to all background estimates given in the rest of this chapter.

5.2.2 $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$

Similar to $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays, the semileptonic $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decay can represent a significant peaking background for the analysis when the kaon is misidentified as a muon. The branching fraction of the decay has recently been measured by LHCb [141] to be

$$B(B_s^0 \rightarrow K^- \mu^+ \nu_\mu) = (1.06 \pm 0.10) \times 10^{-4}, \quad (5.6)$$

which is 25% lower than the value used in the last LHCb analysis, where the branching fraction was estimated from only theory inputs. The rate is of the same order of magnitude as for $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays. However, the contribution

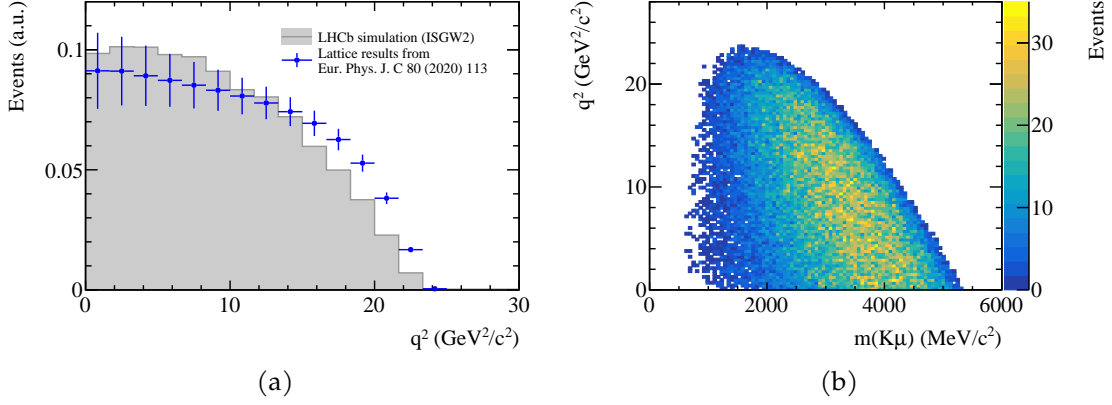


Figure 5.5 – (a) Generated q^2 distribution of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decays using the ISGW2 form factor model and recent lattice QCD results. (b) Generated $K\mu$ mass versus generated q^2 distribution in the simulated $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ sample.

from $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decays is expected to be smaller due to the larger mass shift, explained by the mass difference of kaons and pions and the lower fragmentation fraction of the B_s^0 meson.

Again, large simulated signal samples are produced to study the background contribution. Similar to the $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ case, a form factor reweighting is applied to the q^2 spectrum. A difference is that only lattice QCD calculations and no fits to data are available for the form factors. In Fig. 5.5a, the q^2 spectrum obtained from the already mentioned ISGW2 model in LHCb simulation is compared to the spectrum generated according to lattice QCD predictions based on the recent FLAG average [53]. Both distributions agree better than in the case of $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays. Still, the simulation consistently overestimates the background in the lower q^2 region, requiring a correction to avoid biasing the branching fraction fit.

With the same procedure as for $B^0 \rightarrow \pi^- \mu^+ \nu_\mu$ decays, weights are derived to scale the simulation in the different q^2 regions:

$$w_1 = 0.93 \pm 0.16 \quad (q^2 \in [0.00, 1.67]) \quad (5.7)$$

$$w_2 = 0.90 \pm 0.14 \quad (q^2 \in [1.67, 3.33]) \quad (5.8)$$

$$w_3 = 0.88 \pm 0.12 \quad (q^2 \in [3.33, 5.00]) \quad (5.9)$$

The uncertainties on the weights in the relevant region with $q^2 < 5 \text{ GeV}^2/c^2$ are much larger than for the pionic background, which is caused by the significant uncertainties of the form factors at low q^2 .

In Table 5.1, the expected numbers of events in bins of the BDT are listed. The invariant mass distributions in bins of the BDT are fitted on simulation with one-dimensional kernel density estimators using Gaussian kernels. The fit results are shown in Fig. 5.6.

5 Background studies

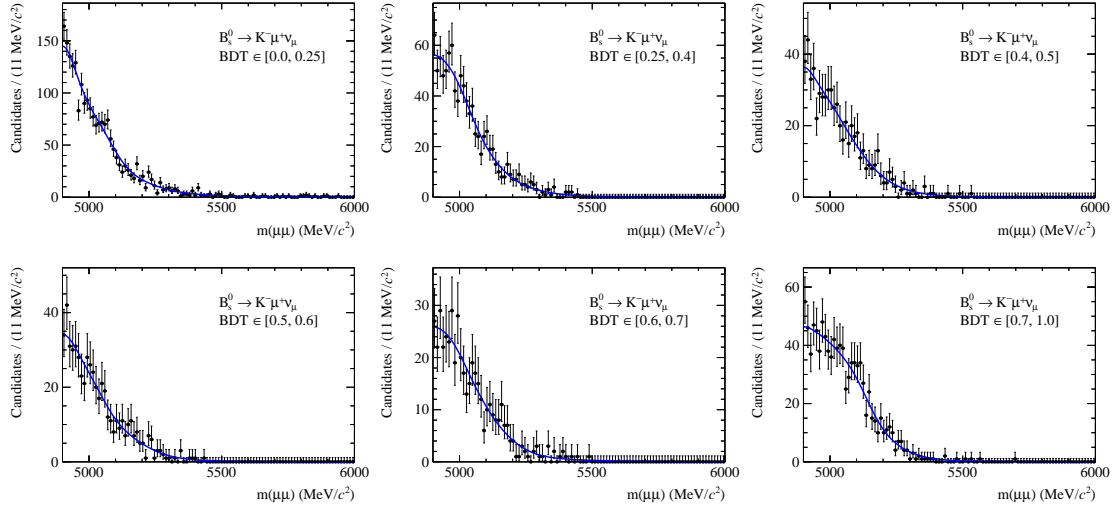


Figure 5.6 – Invariant mass distribution of $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ decays in the reconstructed dimuon channel and shape fitted to simulation. The distributions are shown separately for the six BDT bins.

5.2.3 $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ and $B^0 \rightarrow \pi^0 \mu^+ \mu^-$

The $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ and $B^0 \rightarrow \pi^0 \mu^+ \mu^-$ decays can mimic the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decay due to two final state muons forming a high-quality vertex and an escaping pion. While the dimuon mass distribution will not reach the signal region, it could affect the left sideband. The $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ decay has a branching fraction [136] of

$$\mathcal{B}(B^+ \rightarrow \pi^+ \mu^+ \mu^-) = (1.75 \pm 0.22) \times 10^{-8}. \quad (5.10)$$

The $B^0 \rightarrow \pi^0 \mu^+ \mu^-$ decay has not been observed so far, but a theoretical estimate [142] of its rate relative to $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ can be used:

$$\frac{\mathcal{B}(B^0 \rightarrow \pi^0 \mu^+ \mu^-)}{\mathcal{B}(B^+ \rightarrow \pi^+ \mu^+ \mu^-)} = 0.47^{+0.22}_{-0.18}. \quad (5.11)$$

From this, an estimate for the branching fraction is obtained:

$$\mathcal{B}(B^0 \rightarrow \pi^0 \mu^+ \mu^-) = (0.82 \pm 0.36) \times 10^{-8}. \quad (5.12)$$

The strategy to estimate the background contribution remains unchanged compared to the previous $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ analysis. However, larger simulation samples are used for the two decays, with both muons required to be in the detector acceptance. The expected numbers of events in bins of the BDT are listed in Table 5.1. In Fig. 5.7, the fit results are shown. The same mass shape is used for both $B^+ \rightarrow \pi^+ \mu^+ \mu^-$, and $B^0 \rightarrow \pi^0 \mu^+ \mu^-$ decays, as the dimuon mass distributions agree well in the simulated samples.

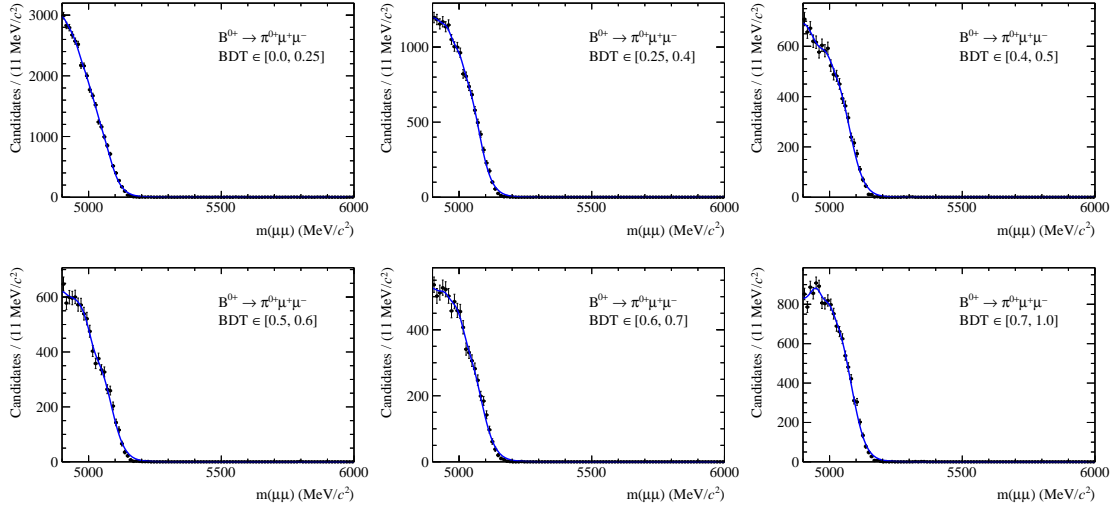


Figure 5.7 – Invariant mass distribution of $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ and $B^0 \rightarrow \pi^0 \mu^+ \mu^-$ decays in the reconstructed dimuon channel and shape fitted to simulation. The distributions are shown separately for the six BDT bins.

5.2.4 $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$

The baryonic $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ decay can imitate the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decay when the proton is wrongly reconstructed as a muon. It can happen due to noise or punch-through muons from proton showers in the calorimeters (as explained in Section 3.2.2). In Ref. [143], the LHCb collaboration measured the branching fraction of the decay to be

$$B(\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu) = (4.1 \pm 1.0) \times 10^{-4}, \quad (5.13)$$

with the uncertainty being dominated by the knowledge of the $\Lambda_b^0 \rightarrow p$ form factors at low q^2 . The hadronisation fraction of Λ_b^0 baryons is taken from the LHCb measurement [144] of

$$r_\Lambda(p_T) \equiv \frac{f_{\Lambda_b^0}}{f_u + f_d}(p_T) = A [p_1 + \exp(p_2 + p_3 \times p_T)], \quad (5.14)$$

where

$$\begin{aligned} A &= 1.000 \pm 0.061 \\ p_1 &= (7.93 \pm 1.41) \times 10^{-2} \\ p_2 &= -1.022 \pm 0.047 \\ p_3 &= (-0.107 \pm 0.002) \text{ GeV}^{-1}. \end{aligned}$$

The relation $\beta_\Lambda = 2r_\Lambda(p_T) \times \beta_u$ is obtained by assuming the hadronisation fractions f_u and f_d to be equal. Since the ratio r_Λ has a p_T dependence, the hadronisation fraction of the Λ_b^0 baryon is incorporated in the background efficiency by

5 Background studies

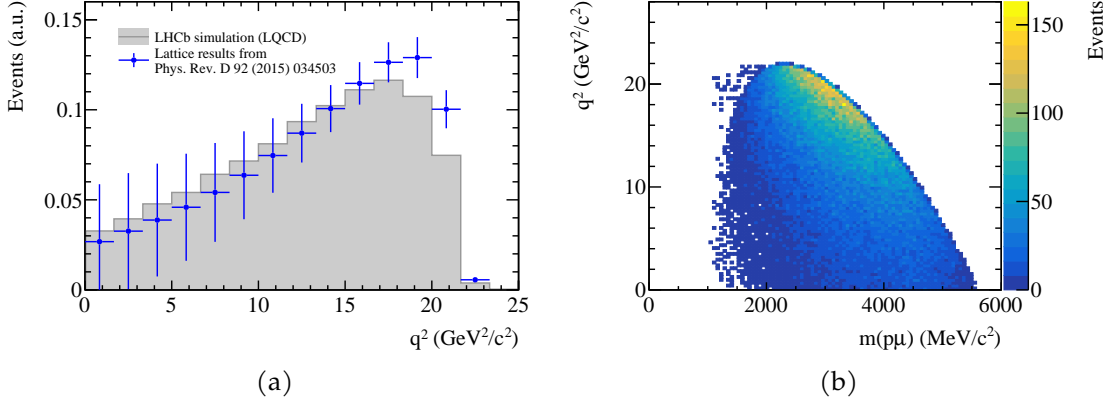


Figure 5.8 – (a) Generated q^2 distribution of $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ decays using a lattice QCD form factor model in simulation and using the model with updated values. (b) Generated $p\mu$ mass versus generated q^2 distributions in the simulated $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ sample.

assigning per-event-weights of $2r_\Lambda(p_T)$ and then using β_μ to normalise. In the above equation, the transverse momentum of the combination formed by the muon and the proton is used, as it corresponds with the one from the $\Lambda_c^+\mu^-$ pair in $\Lambda_b^0 \rightarrow \Lambda_c^+\mu^-\bar{\nu}_\mu X$ decays used in the LHCb measurement.

Simulated samples of about two million signal events for Run 1 and 14 million events for Run 2 are used to study the background. The events were generated requiring $m(p\mu) > 4500 \text{ MeV}/c^2$ and using the lattice QCD form factor model given in Ref. [145]. Similar to $B^0 \rightarrow \pi^-\mu^+\nu_\mu$ and $B_s^0 \rightarrow K^-\mu^+\nu_\mu$ decays, a correction is applied to the q^2 spectrum considering the theoretical uncertainties on the form factor description. In Fig. 5.8a, the q^2 spectrum in simulation is compared to the spectrum generated according to the same lattice QCD model but with updated form factor values from Ref. [146]. From the comparison, weights are obtained to scale the simulated events:

$$w_1 = 0.82 \pm 0.97 \quad (q^2 \in [0.00, 1.67]) \quad (5.15)$$

$$w_2 = 0.83 \pm 0.82 \quad (q^2 \in [1.67, 3.33]) \quad (5.16)$$

$$w_3 = 0.81 \pm 0.66 \quad (q^2 \in [3.33, 5.00]) \quad (5.17)$$

The given uncertainties are due to the limited statistics of the simulated samples and the uncertainties in the lattice QCD model, which are derived by varying the form factors within the corresponding parameter uncertainties. The uncertainties on the weights are assumed to be fully correlated.

As shown in Fig. 5.8b, only the region with $q^2 < 5 \text{ GeV}^2/c^2$ is relevant for the analysis. The expected number of $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ background events in bins of the BDT are listed in Table 5.1.

The invariant mass distributions in bins of the BDT are fitted on the simulated samples with one-dimensional kernel density estimators using Gaussian kernels.

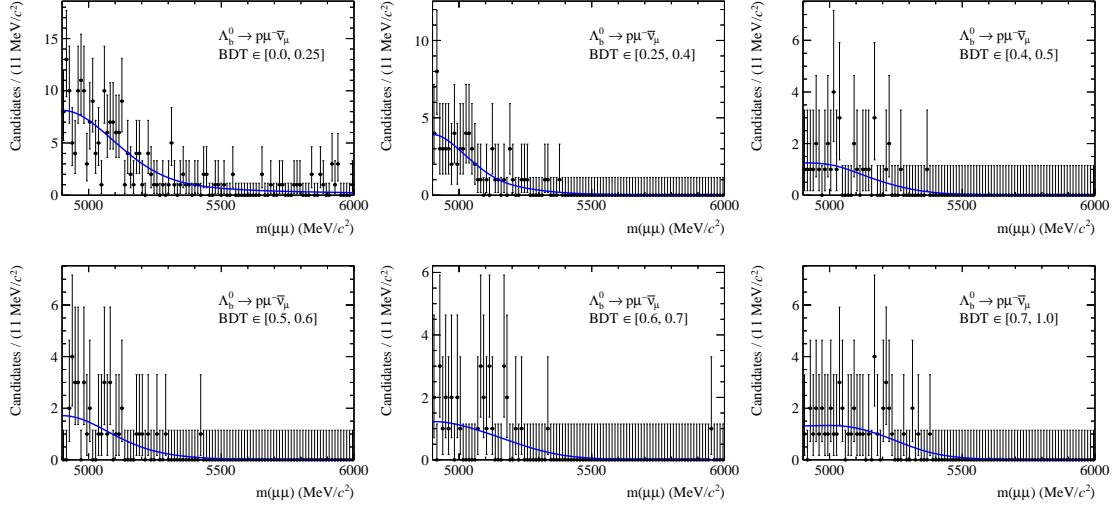


Figure 5.9 – Invariant mass distribution of $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ decays in the reconstructed dimuon channel and shape fitted to simulation. The distributions are shown separately for the six BDT bins.

In Fig. 5.9, the fit results are shown. Potential small shape variations due to form factor uncertainties are not considered in the analysis, as the actual contribution of $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ decays is found to be very small. This step is justified as follows:

The large form factor uncertainties of about 100 % in the relevant q^2 regions are considered by simultaneously moving all form factor weights up and down. It is found that the mass shape only varies if the form factor weights come close to the physical boundary at zero. The variation can be seen in Figs. 5.10a and 5.10b. However, this is an unphysical scenario, as a comparison of the lattice QCD model to a light-cone sum rules (LCSR) model [147] (physically well motivated at low q^2) in Fig. 5.10c shows. Contrary, both the lattice QCD and LCSR models come to very similar central values for the form factors around $q^2 < 5 \text{ GeV}^2/c^2$. Additionally, smaller form factors in this region would reduce the $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ yield in the LHCb acceptance. It has been tested that the large form factor uncertainties propagated to the $\Lambda_b^0 \rightarrow p\mu^-\bar{\nu}_\mu$ yield have a negligible impact on any of the signal branching fractions and the effective lifetime measurement compared to a fixed yield. Therefore, it is concluded that reasonably small mass shape variations cannot affect the intended branching fraction and effective lifetime measurements either.

5.3 Summary of expected background events

The expected numbers of background events in bins of the signal BDT output and in the dimuon mass range from 4900 to 6000 MeV/c^2 are summarised in Table 5.1 for Run 1 and Run 2, respectively. All estimates are used to constrain the background components in the branching fraction fit described in Section 6.2.

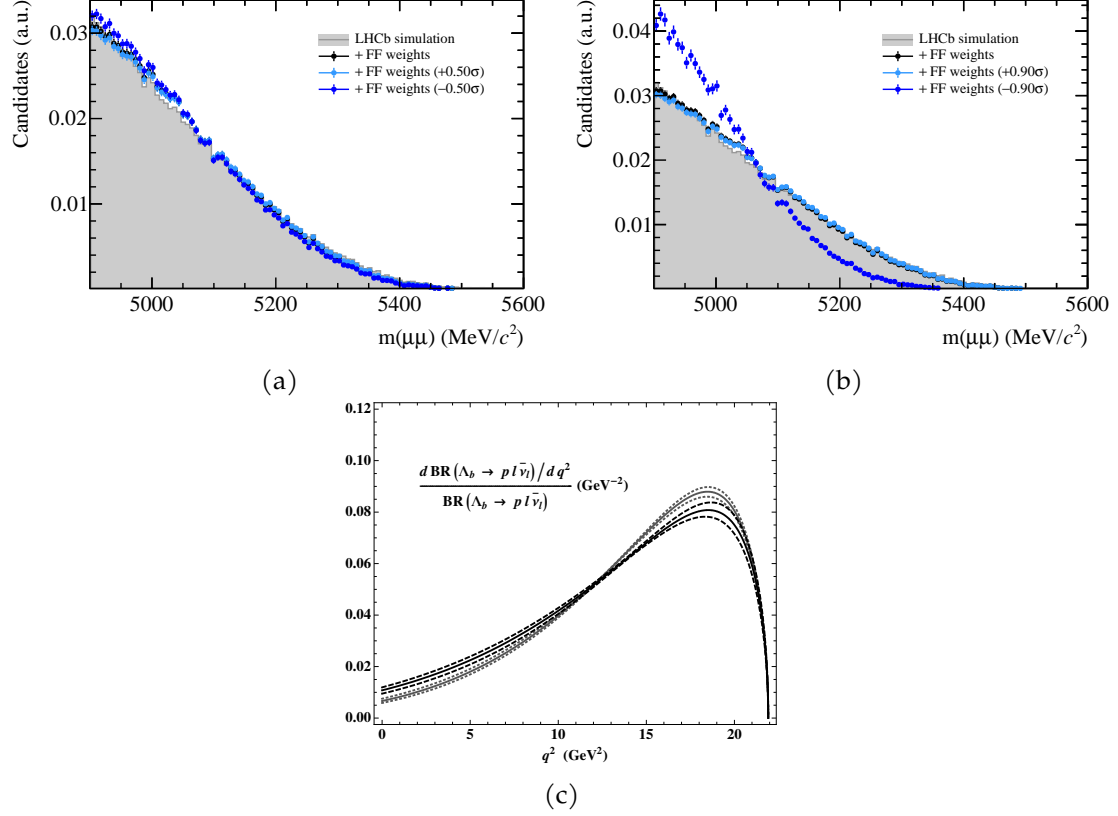


Figure 5.10 – Invariant mass distributions of $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ decays with and without form factor weights applied. The form factor weights are additionally moved up and down by (a) 0.5σ and (b) 0.9σ . (c) Normalised q^2 distribution of $\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$ decays according to an LCSR form factor model [147]. The black (grey) lines indicate the axial-vector (pseudoscalar) contributions and the dashed lines indicate their uncertainties. In the low q^2 region relevant for this analysis, both the lattice QCD and LCSR model indicate similar, clearly non-zero, central values for the form factors.

Table 5.1 – Exclusive background yields per BDT bin with total estimated uncertainties calculated for (a) Run 1 and (b) Run 2 data. The quoted uncertainties include the uncertainties on the branching fraction, the normalisation with $B^+ \rightarrow J/\psi K^+$, and the total efficiencies from simulation and calibration samples.

(a)

BDT region	$B_{(s)}^0 \rightarrow h^+ h^{(\prime)-}$	$B^0 \rightarrow \pi^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B^+ \rightarrow \pi^+ \mu^+ \mu^-$	$B^0 \rightarrow \pi^0 \mu^+ \mu^-$	$\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$	$B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$
[0.0-0.25]	6.1 ± 1.4	14.7 ± 1.4	2.0 ± 0.5	6.1 ± 0.8	1.9 ± 0.8	0.6 ± 0.7	24.4 ± 1.3
[0.25-0.4]	3.8 ± 0.9	7.4 ± 0.9	1.30 ± 0.33	2.7 ± 0.4	1.0 ± 0.4	0.29 ± 0.32	3.27 ± 0.23
[0.4-0.5]	1.57 ± 0.20	5.6 ± 0.7	0.73 ± 0.23	1.45 ± 0.19	0.66 ± 0.29	0.18 ± 0.20	1.54 ± 0.14
[0.5-0.6]	1.61 ± 0.21	3.9 ± 0.6	1.6 ± 0.4	1.17 ± 0.16	0.63 ± 0.28	0.20 ± 0.22	0.86 ± 0.10
[0.6-0.7]	1.65 ± 0.21	4.8 ± 0.7	1.40 ± 0.35	0.95 ± 0.13	0.62 ± 0.27	0.19 ± 0.21	0.61 ± 0.08
[0.7-1.0]	5.3 ± 0.7	9.1 ± 1.1	4.0 ± 0.9	1.36 ± 0.18	1.3 ± 0.6	0.32 ± 0.35	0.43 ± 0.07
[0.0-1.0]	20.1 ± 2.6	46 ± 4	11.0 ± 2.4	13.8 ± 1.8	6.0 ± 2.7	1.8 ± 2.0	31.1 ± 1.6

(b)

BDT region	$B_{(s)}^0 \rightarrow h^+ h^{(\prime)-}$	$B^0 \rightarrow \pi^- \mu^+ \nu_\mu$	$B_s^0 \rightarrow K^- \mu^+ \nu_\mu$	$B^+ \rightarrow \pi^+ \mu^+ \mu^-$	$B^0 \rightarrow \pi^0 \mu^+ \mu^-$	$\Lambda_b^0 \rightarrow p \mu^- \bar{\nu}_\mu$	$B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$
[0.0-0.25]	10.7 ± 1.7	63 ± 5	3.7 ± 1.0	22.9 ± 2.9	6.7 ± 3.0	1.8 ± 1.9	79 ± 4
[0.25-0.4]	6.2 ± 1.0	34.4 ± 2.8	3.1 ± 0.8	9.9 ± 1.3	3.6 ± 1.6	1.1 ± 1.2	11.7 ± 0.6
[0.4-0.5]	2.90 ± 0.26	22.3 ± 2.0	2.5 ± 0.7	5.0 ± 0.6	2.3 ± 1.0	0.8 ± 0.8	4.59 ± 0.27
[0.5-0.6]	2.82 ± 0.26	18.9 ± 1.8	3.6 ± 0.9	4.0 ± 0.5	2.2 ± 1.0	0.9 ± 0.9	2.85 ± 0.19
[0.6-0.7]	2.68 ± 0.24	18.1 ± 1.8	3.7 ± 0.9	3.1 ± 0.4	2.2 ± 1.0	0.9 ± 1.0	1.56 ± 0.12
[0.7-1.0]	8.1 ± 0.7	36.0 ± 3.1	8.6 ± 1.9	4.5 ± 0.6	4.4 ± 1.9	1.6 ± 1.8	0.85 ± 0.08
[0.0-1.0]	33.4 ± 2.9	193 ± 13	25 ± 5	49 ± 6	21 ± 9	7 ± 8	101 ± 5

6 Branching fraction measurement

In this chapter, the results of the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fraction measurements are presented. The normalisation procedure is outlined in Section 6.1. It comprises the strategy, the efficiency calculation, and a presentation of the normalisation constants, including validation studies. A description of the mass model used in the branching fraction fit is given in Section 6.2. Lastly, the branching fraction fit results are shown and interpreted in Section 6.3.

6.1 Normalisation with $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow K^+ \pi^-$ decays

The efficiency-corrected number of observed $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ events can be translated into branching fractions using the integrated luminosity of the data set, the production cross-section of $b\bar{b}$ pairs at LHCb, and the hadronisation fraction for B_s^0 and B^0 mesons. However, the integrated luminosity comes with significant uncertainty. Therefore, the signal branching fractions are calculated with another approach: The number of observed signal events is normalised to the $B^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow J/\psi K^+$ decay modes. The normalisation equation is

$$\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+ \mu^-) = \frac{1}{\underbrace{\sum_y \langle \beta^{(y)} \rangle \varepsilon^{(y)}(B_{(s)}^0 \rightarrow \mu^+ \mu^-)}_{\alpha_{(s)}}} \left(\times \frac{f_d}{f_s} \right) \times N(B_{(s)}^0 \rightarrow \mu^+ \mu^-), \quad (6.1)$$

where α and α_s are the single-event sensitivities for the signal modes. The factor $\langle \beta^{(y)} \rangle$ is calculated as sum over all data-taking years y to be the uncertainty-weighted average of

$$\beta_u = \frac{N(B^+ \rightarrow J/\psi K^+)}{\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \varepsilon(B^+ \rightarrow J/\psi K^+)} \quad (6.2)$$

and

$$\beta_d = \frac{N(B^0 \rightarrow K^+ \pi^-)}{\mathcal{B}(B^0 \rightarrow K^+ \pi^-) \times \varepsilon(B^0 \rightarrow K^+ \pi^-)}. \quad (6.3)$$

Here, N is the observed number of signal events, and ε denotes the measured efficiency. The process $J/\psi \rightarrow \mu^+ \mu^-$ is always assumed. While the number of signal events is determined from simultaneous fits to selected signal candidates

in the Run 1 and Run 2 data sets, the efficiencies are obtained using a mixture of simulation-based and data-driven techniques as explained in Section 4.3 and in the following sections. The branching fractions of the normalisation modes [27]

$$\mathcal{B}(B^+ \rightarrow J/\psi K^+) = (6.00 \pm 0.16) \times 10^{-5} \quad (6.4)$$

$$\mathcal{B}(B^0 \rightarrow K^+ \pi^-) = (1.96 \pm 0.05) \times 10^{-5} \quad (6.5)$$

are used as external inputs, with the $J/\psi \rightarrow \mu^+ \mu^-$ branching fraction being included in the above value. The hadronisation fraction ratio f_s/f_d enters into the calculation of α_s and contributes significantly to the uncertainty. It is found to depend on the b hadron momentum and the centre-of-mass energy of the collision. In previous analyses, it was tried to avoid this ratio by measuring $B_s^0 \rightarrow \mu^+ \mu^-$ relative to the decay mode $B_s^0 \rightarrow J/\psi \phi$ with $\phi \rightarrow K^+ K^-$. However, no improvement was found due to the uncertainty on the introduced $B_s^0 \rightarrow J/\psi \phi$ branching fraction. The value of f_s/f_d has been measured by LHCb in different modes and at different collision energies, like in semileptonic $B \rightarrow D\mu X$ decays at 7 and 13 TeV [144, 148], hadronic $B \rightarrow Dh$ decays [149, 150], and $B \rightarrow J/\psi X$ decays [151], both at 7, 8, and 13 TeV. In this analysis, new phase-space integrated combination values from Ref. [152] with reduced uncertainties are used:

$$f_s/f_d (7 \text{ TeV}) = 0.2390 \pm 0.0076 \quad (6.6)$$

$$f_s/f_d (8 \text{ TeV}) = 0.2385 \pm 0.0075 \quad (6.7)$$

$$f_s/f_d (13 \text{ TeV}) = 0.2539 \pm 0.0079 . \quad (6.8)$$

The correlations among the three values are considered by only using the 13 TeV value and combining it with a measurement of

$$\frac{f_s/f_d (\text{Run 2})}{f_s/f_d (\text{Run 1})} = 1.064 \pm 0.007 , \quad (6.9)$$

provided in the same publication. In addition, it is assumed that isospin symmetry holds, allowing f_s/f_d to be used in connection with the $B^+ \rightarrow J/\psi K^+$ normalisation channel.

6.1.1 Selection efficiencies

In the following the signal selection efficiencies are presented, which can be divided into three partial efficiencies: the geometrical detector acceptances, the reconstruction and selection efficiencies, and the trigger efficiencies, such that

$$\varepsilon = \varepsilon_{\text{Acc}} \times \varepsilon_{\text{RecSel|Acc}} \times \varepsilon_{\text{Trig|AccRecSel}} . \quad (6.10)$$

Each efficiency is evaluated with the candidates passing the previous stage.

The detector geometry causes the first efficiency since not all signal decays occur within the detector acceptance. The fraction of signal decays having all

decay products within the acceptance compared to all signal decays produced defines the efficiency. It is computed using simulated samples for the signal and normalisation modes. The physical LHCb detector acceptance in the forward angular bending plane is given by a polar angle ranging from approximately 10 to 300 mrad. Nevertheless, all events where the decay products fly under an angle smaller than 400 mrad are accepted in simulation. This looser requirement is chosen to allow for the recovery of particles, bent by the magnetic field. The geometrical detector acceptance efficiencies are listed in Table 6.1. The efficiencies are similar for the signal and the $B^0 \rightarrow K^+ \pi^-$ decays but lower for the $B^+ \rightarrow J/\psi K^+$ decays, which is expected given the additional third final state track.

The reconstruction efficiency is the ratio of reconstructed decay candidates within the detector acceptance. In contrast, the selection efficiency is the fraction of all reconstructed candidates selected by the requirements presented in Section 4.2. The selection and reconstruction efficiency calculation relies on simulation and collected data. The efficiencies depend on the decay's characteristics, like the number of particles in the final state, their kinematics, the quality of the track reconstruction, or the quality of the particle identification. Also, effects caused by the magnetic field and by interactions of the particles with the detector material are included. The simulation-based efficiencies for the different channels are shown in Table 6.1. Since only the ratio of signal and normalisation efficiencies matters in the branching fraction calculation, variations due to different simulation software versions and slightly different detector performances between the data-taking years cancel out.

The calculation of the trigger efficiencies is not part of this thesis but can be found in Ref. [126].

6.1.2 Tracking efficiency corrections

The simulated samples are corrected to describe the signal in real data more accurate. In the course of this thesis, especially the track reconstruction efficiency is studied. The procedure to account for the corresponding simulation imperfections is described in the following:

The efficiency to reconstruct a particle track is measured on $J/\psi \rightarrow \mu^+ \mu^-$ data using the tag-and-probe method [153]. The measured tracking efficiencies are included as correction factors to the reconstruction and selection efficiencies, which are determined on the simulated samples and include the simulated tracking efficiencies. Because the efficiencies depend on the track kinematics, the corrections are determined separately for all data-taking years in bins of the track momentum and pseudorapidity. The tracking efficiency correction factor for a given decay channel is calculated from these tracking maps as

$$C^{\text{ch}} = \prod_{tr} C_{tr}^{\text{ch}} = \prod_{tr} \left\{ \sum_i (f_i^{\text{tr}} C_i) \beta_{\text{RW}} \beta_{\text{mat}}^{\text{tr}} \right\}, \quad (6.11)$$

6 Branching fraction measurement

Table 6.1 – Geometrical detector acceptance efficiencies, estimated as the fraction of decays contained in the polar angle region from 10 to 300 mrad, and reconstruction and selection efficiencies for the signal and normalisation channels. Next to the efficiencies, the number of candidates at the two selection stages is also given. All numbers are evaluated on simulated samples.

Channel	ε^{Acc}	N^{Acc}	$\varepsilon^{\text{RecSel Acc}}$	$N^{\text{AccRecSel}}$
$B_s^0 \rightarrow \mu^+ \mu^-$				
2011	$(18.377 \pm 0.039)\%$	534 499	$(33.78 \pm 0.06)\%$	180 540
2012	$(18.668 \pm 0.023)\%$	2 080 642	$(31.487 \pm 0.032)\%$	655 129
2015	$(19.296 \pm 0.037)\%$	1 011 344	$(32.96 \pm 0.05)\%$	333 341
2016	$(19.420 \pm 0.047)\%$	2 011 123	$(33.176 \pm 0.033)\%$	667 217
2017	$(19.273 \pm 0.043)\%$	2 012 366	$(33.198 \pm 0.033)\%$	668 064
2018	$(19.373 \pm 0.048)\%$	1 008 582	$(33.23 \pm 0.05)\%$	335 158
$B^0 \rightarrow \mu^+ \mu^-$				
2011	$(18.329 \pm 0.038)\%$	508 999	$(33.62 \pm 0.07)\%$	171 101
2012	$(18.691 \pm 0.032)\%$	498 027	$(30.95 \pm 0.07)\%$	154 132
2015	$(19.282 \pm 0.041)\%$	1 072 675	$(32.40 \pm 0.05)\%$	347 595
2016	$(19.382 \pm 0.050)\%$	2 006 533	$(32.616 \pm 0.033)\%$	654 455
2017	$(19.364 \pm 0.049)\%$	2 071 197	$(32.600 \pm 0.033)\%$	675 209
2018	$(19.423 \pm 0.048)\%$	1 047 670	$(33.40 \pm 0.05)\%$	349 967
$B^+ \rightarrow J/\psi K^+$				
2011	$(16.410 \pm 0.033)\%$	762 312	$(18.37 \pm 0.04)\%$	140 048
2012	$(16.702 \pm 0.042)\%$	5 047 318	$(16.648 \pm 0.017)\%$	840 275
2015	$(17.373 \pm 0.033)\%$	2 072 461	$(17.370 \pm 0.026)\%$	359 980
2016	$(17.342 \pm 0.029)\%$	4 182 716	$(17.757 \pm 0.019)\%$	742 742
2017	$(17.394 \pm 0.040)\%$	10 006 407	$(17.759 \pm 0.012)\%$	1 777 045
2018	$(17.354 \pm 0.042)\%$	9 912 257	$(17.745 \pm 0.012)\%$	1 758 921
$B^0 \rightarrow K^+ \pi^-$				
2011	$(17.735 \pm 0.024)\%$	775 505	$(25.93 \pm 0.05)\%$	201 069
2012	$(18.978 \pm 0.051)\%$	8 581 113	$(24.103 \pm 0.015)\%$	2 068 318
2015	$(19.589 \pm 0.037)\%$	2 009 237	$(25.537 \pm 0.031)\%$	513 096
2016	$(19.660 \pm 0.049)\%$	4 006 284	$(25.543 \pm 0.022)\%$	1 023 331
2017	$(19.680 \pm 0.046)\%$	4 008 116	$(25.572 \pm 0.022)\%$	1 024 970
2018	$(19.655 \pm 0.048)\%$	4 109 859	$(25.511 \pm 0.022)\%$	1 048 447

6.1 Normalisation with $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow K^+ \pi^-$ decays

where the index tr runs over the final state tracks, and i runs over the bins of the tracking map. The parameter C_i is the value of the tracking map in bin i , and the factor denoting the corresponding fraction of reconstructed and selected tracks in this bin is given as

$$f_i^{tr} = N_i^{tr, \text{AccRecSel}} / N^{\text{AccRecSel}} . \quad (6.12)$$

The factor β_{RW} provides an uncertainty per track on the correction resulting from different reweighting techniques in Ref. [153] that adjust the simulation according to data:

$$\beta_{\text{RW}} = 1 \pm 0.004 \quad (\text{for } 2011/2012) \quad (6.13)$$

$$\beta_{\text{RW}} = 1 \pm 0.008 \quad (\text{for } 2015/2016) \quad (6.14)$$

$$\beta_{\text{RW}} = 1 \pm 0.004 \quad (\text{for } 2017/2018) . \quad (6.15)$$

The factor β_{mat}^{tr} takes into account the increase of the tracking uncertainty for kaon and pion tracks compared to muons due to the hadronic interactions of the former with the detector material:

$$\beta_{\text{mat}}^K = 1 \pm 0.011 \quad (6.16)$$

$$\beta_{\text{mat}}^\pi = 1 \pm 0.014 \quad (6.17)$$

$$\beta_{\text{mat}}^\mu = 1 . \quad (6.18)$$

The resulting tracking efficiency corrections for all channels are listed in Table 6.2.

Regarding the tracking map coverage, the tag-and-probe method used in Ref. [153] allows for avoiding most selection requirements on the probe tracks, apart from the L0 trigger and loose transverse momentum requirements needed to remove poorly reconstructed tracks. The maps are defined in the track momentum range from 5 to 200 GeV and pseudorapidity range from 1.9 to 4.9. However, the signal and normalisation channels used in this analysis will produce final state tracks outside the tracking maps. The percentages of these number of tracks are between 5 and 15 % across the 2, 3, and 4 body final states.

The track reconstruction efficiencies in the simulation agree with those determined on data [153]. Not applying a correction to the outlier tracks is not expected to have a significant effect. Nevertheless, the effect of the outlier tracks is measured by calculating the average correction factors for the channels in two ways: firstly, using no correction factors for the outlier tracks and secondly, using the correction factors from the adjacent tracking map bins. The difference is assigned to the track correction as a systematic β_{out} . The uncertainty arising from outlier tracks is found to be small compared to other uncertainties that are not related to the tracking corrections.

Finally, the correction uncertainties are reduced in the normalisation ratios because (i) the tracking efficiency correction maps are used for all the track types, (ii) the reweighting uncertainty β_{RW} cancels for all the track ratios and,

most importantly, (iii) the hadron uncertainty β_{mat} cancels (reduces) in the ratio of same (different) hadron types. What remains are the uncertainties due to outlier tracks and the limited statistics of the tracking maps and the simulation samples. The correction ratio per track is then expressed in the following way:

$$R(C_{tr}^{\text{ch}}) = \frac{\sum_i f'_i C_i}{\sum_j f_j C_j} \times \frac{\beta'_{\text{out}}}{\beta_{\text{out}}} \times \frac{\beta'_{\text{mat}}}{\beta_{\text{mat}}}. \quad (6.19)$$

The uncertainty from the simulation statistics available after reconstruction and selection and the simulation and data statistics used to calculate the bin-by-bin corrections is derived as:

$$\begin{aligned} \sigma_{\text{stat}}^2(R) &= \sum_i \left(\frac{\partial R}{\partial f'_i} \right)^2 \sigma(f'_i)^2 + \sum_i \left(\frac{\partial R}{\partial f_i} \right)^2 \sigma(f_i)^2 + \sum_i \left(\frac{\partial R}{\partial C_i} \right)^2 \sigma(C_i)^2 \quad (6.20) \\ &= \left(\frac{1}{C_{tr}^{\text{ch}}} \right)^4 \left\{ (C_{tr}^{\text{ch}})^2 \sum_i (C_i \sigma(f'_i))^2 + (C_{tr}^{\text{ch}\prime})^2 \sum_i (C_i \sigma(f_i))^2 \right. \\ &\quad \left. + \sum_i (f'_i C_{tr}^{\text{ch}} - f_i C_{tr}^{\text{ch}\prime})^2 \sigma(C_i)^2 \right\}. \quad (6.21) \end{aligned}$$

Table 6.2 – Tracking efficiency corrections to the simulation-based reconstruction and selection efficiencies. The correction factors C^{ch} are given year-wise for the signal and normalisation channels.

Year	$C^{\text{ch}}(B_s^0 \rightarrow \mu^+ \mu^-)$	$C^{\text{ch}}(B^0 \rightarrow \mu^+ \mu^-)$	$C^{\text{ch}}(B^+ \rightarrow J/\psi K^+)$	$C^{\text{ch}}(B^0 \rightarrow K^+ \pi^-)$
2011	1.007 ± 0.007	1.007 ± 0.007	1.021 ± 0.014	1.007 ± 0.019
2012	1.006 ± 0.006	1.006 ± 0.007	1.020 ± 0.014	1.006 ± 0.019
2015	0.999 ± 0.014	0.999 ± 0.014	0.994 ± 0.019	0.999 ± 0.023
2016	0.991 ± 0.012	0.991 ± 0.012	0.988 ± 0.018	0.991 ± 0.021
2017	0.999 ± 0.011	0.999 ± 0.011	0.997 ± 0.018	0.999 ± 0.021
2018	0.998 ± 0.012	0.998 ± 0.012	0.991 ± 0.018	0.998 ± 0.021

6.1.3 Normalisation factors and cross-checks

The single-event sensitivities α and α_s are calculated according to Eq. (6.1) from the data-corrected efficiencies and the measured numbers of signal events in the normalisation channels. These normalisation constants are summarised in Table 6.3 together with the expected numbers of rare $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ events, calculated using the branching fraction predictions [50]. The numbers show that both normalisation channels, $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow K^+ \pi^-$, have similar statistical sizes. Combining both leads to normalisation constants with reduced uncertainties lowering the uncertainties of the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fractions.

Table 6.3 – Summary of the normalisation constants and numbers of expected signal events for the Run 1, Run 2, and the combined data set. The number of expected signal events is derived using the branching fraction predictions in Ref. [50].

	$B^0 \rightarrow \mu^+ \mu^-$	$B_s^0 \rightarrow \mu^+ \mu^-$
Run 1		
$\alpha(B^+ \rightarrow J/\psi K^+)$	$(2.85 \pm 0.09 \pm 0.06) \times 10^{-11}$	$(1.130 \pm 0.050 \pm 0.023) \times 10^{-10}$
$\alpha(B^0 \rightarrow K^+ \pi^-)$	$(2.73 \pm 0.10 \pm 0.12) \times 10^{-11}$	$(1.08 \pm 0.05 \pm 0.05) \times 10^{-10}$
$\alpha(\text{combined})$	$(2.83 \pm 0.07 \pm 0.06) \times 10^{-11}$	$(1.120 \pm 0.044 \pm 0.020) \times 10^{-10}$
N_{expected}	$3.64 \pm 0.20 \pm 0.07$	$32.7 \pm 1.8 \pm 0.6$
Run 2		
$\alpha(B^+ \rightarrow J/\psi K^+)$	$(8.54 \pm 0.26 \pm 0.13) \times 10^{-12}$	$(3.19 \pm 0.14 \pm 0.05) \times 10^{-11}$
$\alpha(B^0 \rightarrow K^+ \pi^-)$	$(8.27 \pm 0.25 \pm 0.26) \times 10^{-12}$	$(3.09 \pm 0.13 \pm 0.10) \times 10^{-11}$
$\alpha(\text{combined})$	$(8.47 \pm 0.16 \pm 0.10) \times 10^{-12}$	$(3.168 \pm 0.115 \pm 0.033) \times 10^{-11}$
N_{expected}	$12.16 \pm 0.63 \pm 0.14$	$115.5 \pm 6.1 \pm 1.2$
Run 1+2		
$\alpha(B^+ \rightarrow J/\psi K^+)$	$(6.57 \pm 0.19 \pm 0.08) \times 10^{-12}$	$(2.505 \pm 0.108 \pm 0.030) \times 10^{-11}$
$\alpha(B^0 \rightarrow K^+ \pi^-)$	$(6.35 \pm 0.18 \pm 0.17) \times 10^{-12}$	$(2.42 \pm 0.10 \pm 0.06) \times 10^{-11}$
$\alpha(\text{combined})$	$(6.52 \pm 0.10 \pm 0.06) \times 10^{-12}$	$(2.485 \pm 0.086 \pm 0.023) \times 10^{-11}$
N_{expected}	$15.80 \pm 0.80 \pm 0.16$	$147.3 \pm 7.6 \pm 1.3$

Several cross-checks of the normalisation procedure validate the efficiency computation and yield extraction. The first data-driven check compares the ratio of the $B^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow J/\psi K^+$ branching fractions with the value provided in Ref. [27]. The check comprises the evaluation of

$$\frac{\mathcal{B}(B^0 \rightarrow K^+ \pi^-)}{\mathcal{B}(B^+ \rightarrow J/\psi K^+)} = \frac{N(B^0 \rightarrow K^+ \pi^-)}{N(B^+ \rightarrow J/\psi K^+)} \times \frac{\varepsilon(B^+ \rightarrow J/\psi K^+)}{\varepsilon(B^0 \rightarrow K^+ \pi^-)} \times \frac{f_u}{f_d} \quad (6.22)$$

for the different data-taking years. In Table 6.4, the results are given assuming isospin symmetry, *i.e.* $f_u = f_d$. Excellent stability can be seen between the different years. Compared to the world average value, a small tension is visible, which could be explained by taking the actual value of f_u/f_d into account. No LHC measurement is available for this hadronisation ratio. The average value from the B factories for the production of two charged over two neutral B mesons is $f^{+/-}/f^{00} = 1.059 \pm 0.027$ [140]. Adding this value to the computation would make the branching fraction ratio perfectly agree with the reference value. However, it was decided to not include the ratio in the normalisation calculation given the differences between LHCb and the B factories. A second cross-check is performed by probing the relative B_s^0 and B^+ production ratio

$$R \equiv \frac{f_s}{f_u} \times \frac{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)}{\mathcal{B}(B^+ \rightarrow J/\psi K^+)} = \frac{N(B_s^0 \rightarrow J/\psi \phi)}{N(B^+ \rightarrow J/\psi K^+)} \times \frac{\varepsilon(B^+ \rightarrow J/\psi K^+)}{\varepsilon(B_s^0 \rightarrow J/\psi \phi)}. \quad (6.23)$$

6 Branching fraction measurement

Table 6.4 – Ratio of the branching fractions of $B^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow J/\psi K^+$ decays obtained from data using the numbers of observed events and the corresponding selection efficiencies. The decay $J/\psi \rightarrow \mu^+ \mu^-$ is implied. The ratios show good stability between the different data subsets and are compatible with the PDG value provided in Ref. [27].

Year	$N(B^+ \rightarrow J/\psi K^+)$	$N(B^0 \rightarrow K^+ \pi^-)$	$\frac{\varepsilon(B^+ \rightarrow J/\psi K^+)}{\varepsilon(B^0 \rightarrow K^+ \pi^-)}$	$\frac{B(B^0 \rightarrow K^+ \pi^-)}{B(B^+ \rightarrow J/\psi K^+)}$
PDG				0.326 ± 0.012
2011	$(3.479 \pm 0.008) \times 10^5$	$(3.73 \pm 0.13) \times 10^3$	32.4 ± 2.9	0.347 ± 0.034
2012	$(7.780 \pm 0.012) \times 10^5$	$(1.032 \pm 0.023) \times 10^4$	25.4 ± 1.3	0.338 ± 0.019
2015	$(1.676 \pm 0.005) \times 10^5$	$(4.43 \pm 0.14) \times 10^3$	12.7 ± 0.7	0.335 ± 0.021
2016	$(1.0369 \pm 0.0015) \times 10^6$	$(2.37 \pm 0.06) \times 10^4$	15.0 ± 1.0	0.342 ± 0.025
2017	$(1.0820 \pm 0.0014) \times 10^6$	$(2.43 \pm 0.06) \times 10^4$	15.0 ± 0.7	0.337 ± 0.019
2018	$(1.3208 \pm 0.0015) \times 10^6$	$(2.75 \pm 0.06) \times 10^4$	15.9 ± 1.0	0.331 ± 0.022
All				0.337 ± 0.010

The ratio is corrected for the branching fractions of the decay modes used for the check. Since the branching fraction of the $B_s^0 \rightarrow J/\psi \phi$ mode comes with significant uncertainty, the check probes the ratio trend for different centre-of-mass energies. In Fig. 6.1, the resulting values are visualised for different energies and a linear function is fitted yielding a slope of $(5.88 \pm 5.42) \times 10^{-4} \text{ TeV}^{-1}$. This slope is compatible with the ratio being a constant and with the slope stated in Ref. [151], which is $(1.27 \pm 0.27) \times 10^{-3} \text{ TeV}^{-1}$. Given the relative stability between the different data-taking years and the consistency with the reference value, the cross-check is considered as passed.

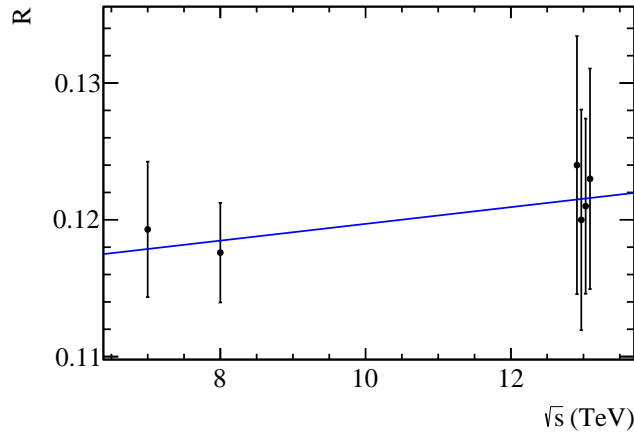


Figure 6.1 – Ratio of the hadronisation fractions f_s and f_u corrected by the branching fractions of the $B_s^0 \rightarrow J/\psi \phi$ and $B^+ \rightarrow J/\psi K^+$ decays and obtained for different centre-of-mass energies. The data points at 13 TeV are slightly shifted for presentation only.

6.2 Modelling of the invariant mass distribution

An unbinned extended maximum likelihood fit is performed to extract the branching fractions of $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays from selected data. The fit is performed simultaneously in bins of the signal BDT output and for the Run 1 and Run 2 data sets. The BDT bin boundaries are $[0.0, 0.25, 0.4, 0.5, 0.6, 0.7, 1.0]$, resulting in six bins. The five highest bins are used in the simultaneous fit. In contrast, the lowest bin is discarded given the domination of combinatorial background and the negligible amount of signal contained in it. Still, the fraction of the lowest bin is considered when calculating the total normalisation of the BDT shape. With this strategy, the analysis differs from the previous one, which used only five bins in total (with the highest bin going from 0.6 to 1.0).

The $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ signal mass distributions are modelled by double-sided Crystal Ball functions [135], defined as

$$f(m|\mu, \sigma, \alpha_l, n_l, \alpha_r, n_r) = N \begin{cases} \left(\frac{n_l}{\alpha_l}\right)^{n_l} \exp\left(-\frac{\alpha_l^2}{2}\right) \left(-\frac{m-\mu}{\sigma} + \frac{n_l}{\alpha_l} - \alpha_l\right)^{-n_l}, & \text{if } \frac{m-\mu}{\sigma} < -\alpha_l \\ \left(\frac{n_r}{\alpha_r}\right)^{n_r} \exp\left(-\frac{\alpha_r^2}{2}\right) \left(\frac{m-\mu}{\sigma} + \frac{n_r}{\alpha_r} - \alpha_r\right)^{-n_r}, & \text{if } \frac{m-\mu}{\sigma} > \alpha_r \\ \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right), & \text{otherwise.} \end{cases} \quad (6.24)$$

The function consists of a Gaussian core with mean μ and width σ and has power-law tails on both sides with slopes n_l and n_r . The corresponding starting points are α_l and α_r in units of σ . The fit is done in a dimuon mass window going from 4900 to 6000 MeV/ c^2 and the means and widths are constrained by Gaussian functions as explained in Section 4.3. Physical background contributions are described with Gaussian kernel density estimators [154] for each BDT bin individually, as shown in Chapter 5. The contribution from the combinatorial background, *i.e.* random combinations of muons passing the selection, is modelled by an unconstrained exponential function per run period with a common slope across the different BDT regions. Within the total fit model, parameters are shared wherever they are in common between the different fit components. This procedure accounts for correlations between the fits on different data samples and comprises, *e.g.* the values of f_s/f_d and $B(B^+ \rightarrow J/\psi K^+)$. The fitted branching fractions are free parameters linked to the number of observed signal events via the normalisation equation. Equation (6.1) is modified with the estimated fraction of signal candidates falling into the considered BDT bin:

$$N_{\text{sig}}^{r,i} = \alpha_{(s)}^r \times \frac{f_{\text{BDT,sig}}^{r,i} c_{\text{PID,sig}}^{r,i} c_{\text{time,sig}}^{r,i}}{\sum_j f_{\text{BDT,sig}}^{r,j} c_{\text{PID,sig}}^{r,j} c_{\text{time,sig}}^{r,j}} \times B_{\text{sig}}. \quad (6.25)$$

Here, f_{BDT} is the estimated and Gaussian-constrained fraction of signal candidates in BDT bin i for a given data-taking period r . The correction factors c_{PID} and c_{time} account for the dependencies of the BDT on the PID selection and B de-

cay time acceptance, which are incorrectly modelled in the simulation. These factors are chosen to be fixed in the fit because of their negligible uncertainties.

Similar equations describe the amount of $B_{(s)}^0 \rightarrow h^+ h'^-$ and semileptonic background events in each BDT region. For the semileptonic backgrounds discussed in Section 5.2, Equation (5.1) gets modified for the fit as

$$N^{\text{expected}} = \frac{f_x}{f_u} \times \frac{N(B^+ \rightarrow J/\psi K^+) \times \mathcal{B} \times \varepsilon}{\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \varepsilon(B^+ \rightarrow J/\psi K^+)} \times \frac{f_{\text{BDT,bkg}}^{r,i}}{\sum_j f_{\text{BDT,bkg}}^{r,j}}. \quad (6.26)$$

All parameters related to the normalisation channel are shared in the fit, while Gaussian functions constrain all numbers used to calculate the background contributions.

6.3 Results and interpretation

The result of the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fraction fit is visualised in Fig. 6.2, showing projections of the dimuon mass distribution in the five fitted BDT bins for Run 1 and Run 2 data. The extracted branching fraction values are:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.09_{-0.45}^{+0.48}) \times 10^{-9} \quad (6.27)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (1.20_{-0.75}^{+0.85}) \times 10^{-10}. \quad (6.28)$$

In Fig. 6.3, the result for the most sensitive region with $\text{BDT} > 0.5$ is presented. Statistical significance is calculated with Wilks' theorem [155] using the difference in the log-likelihood of fits with and without the signal component. The $B_s^0 \rightarrow \mu^+ \mu^-$ signal is clearly visible and measured with a significance of 10.3σ . In contrast, the $B^0 \rightarrow \mu^+ \mu^-$ signal is seen with only 1.7σ above the background-only hypothesis, mainly due to pollution from remaining $B_{(s)}^0 \rightarrow h^+ h'^-$ background decays and limited statistics. Using frequentist statistics, an upper limit for the branching fraction is calculated referring to the maximum value of the confidence interval that encompasses the true value of the parameter at a specific confidence level (CL). This means that if the same experiment is repeated an infinite number of times, the percentage of confidence intervals containing the true parameter value equals the CL. Consequently, the upper limit is established at a specific CL to ensure that the percentage of limits exceeding the true parameter value is equivalent to the CL. The upper limit is calculated using the CL_s method [156] with pseudo-experiments and a profile likelihood ratio as a one-sided test statistic as it is implemented in Ref. [157, 158] and explained in more detail in Section 10.2. The CL_s value distribution for the signal parameter $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)$ is shown in Fig. 6.4. The upper limit is found to be

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 2.6 \times 10^{-10} \quad \text{at } 95\% \text{ CL}. \quad (6.29)$$

The compatibility of the result with the SM prediction is checked by plotting the two-dimensional profile likelihood contours for the two branching fractions. In Fig. 6.5, the contours are calculated from the ratio of likelihoods having the signal branching fractions either fixed or left free in the fit.

Another free parameter is the branching fraction of the $B_s^0 \rightarrow \mu^+ \mu^- \gamma$ decay. This decay is a signal mode in the presented analysis (and discussed in Ref. [125]) but treated as a background component in this thesis. As can be seen in Fig. 6.3, the fit suggests a negative value of $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^- \gamma) = (-2.53_{-0.44}^{+0.46}) \times 10^{-9}$ for the branching fraction with a significance of 1.6σ . Since negative branching fractions are unphysical, the measured value can be interpreted as a result of an under-fluctuation of the data in the left sideband. Its impact on the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fraction measurement is tested by repeating the fit and setting the $B_s^0 \rightarrow \mu^+ \mu^- \gamma$ component to zero. This approach shows compatible results, with the mean value of the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction being increased by 2 %, while the upper limit on the $B^0 \rightarrow \mu^+ \mu^-$ branching fraction gets decreased by 12 %. No significant improvement or deterioration in the fit-to-data agreement in the left data sideband can be seen, indicating a correct estimation and modelling of the semileptonic backgrounds.

All systematic uncertainties to the fitted branching fractions are implemented in the fit as Gaussian-constrained nuisance parameters. Therefore, the uncertainties in Eqs. (6.27) and (6.28) comprise the statistical and systematic component. Both components are separated by repeating the fit with all nuisance parameters being fixed to the final values in the nominal fit giving only the statistical uncertainty. Then, the systematic uncertainty is obtained by subtracting the statistical component from the total uncertainty in quadrature:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.09_{-0.44}^{+0.46} (\text{stat.})_{-0.10}^{+0.14} (\text{syst.})) \times 10^{-9} \quad (6.30)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (1.20_{-0.74}^{+0.83} (\text{stat.})_{-0.15}^{+0.14} (\text{syst.})) \times 10^{-10} . \quad (6.31)$$

Statistical limitations dominate the analysis. For the $B^0 \rightarrow \mu^+ \mu^-$ branching fraction, the statistical uncertainty is more than five times the systematic uncertainty. The leading systematics are caused by the limited precision of f_s/f_d , which is 3 % for the $B_s^0 \rightarrow \mu^+ \mu^-$ mode, and the knowledge of the background estimates, which is 9 % for the $B^0 \rightarrow \mu^+ \mu^-$ mode.

The ratio of the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ branching fractions is measured as it gives insights into the minimal flavour violation hypothesis. By taking the correlations between both branching fraction measurements into account, the ratio is found to be

$$\mathcal{R}_{\mu^+ \mu^-} = 0.039_{-0.024}^{+0.030} (\text{stat.})_{-0.004}^{+0.006} (\text{syst.}) . \quad (6.32)$$

Since the B^0 decay is not yet observed, the CL_s method is used to calculate an upper limit on the ratio:

$$\mathcal{R}_{\mu^+ \mu^-} < 0.095 \quad \text{at } 95 \% \text{ CL} . \quad (6.33)$$

6 Branching fraction measurement

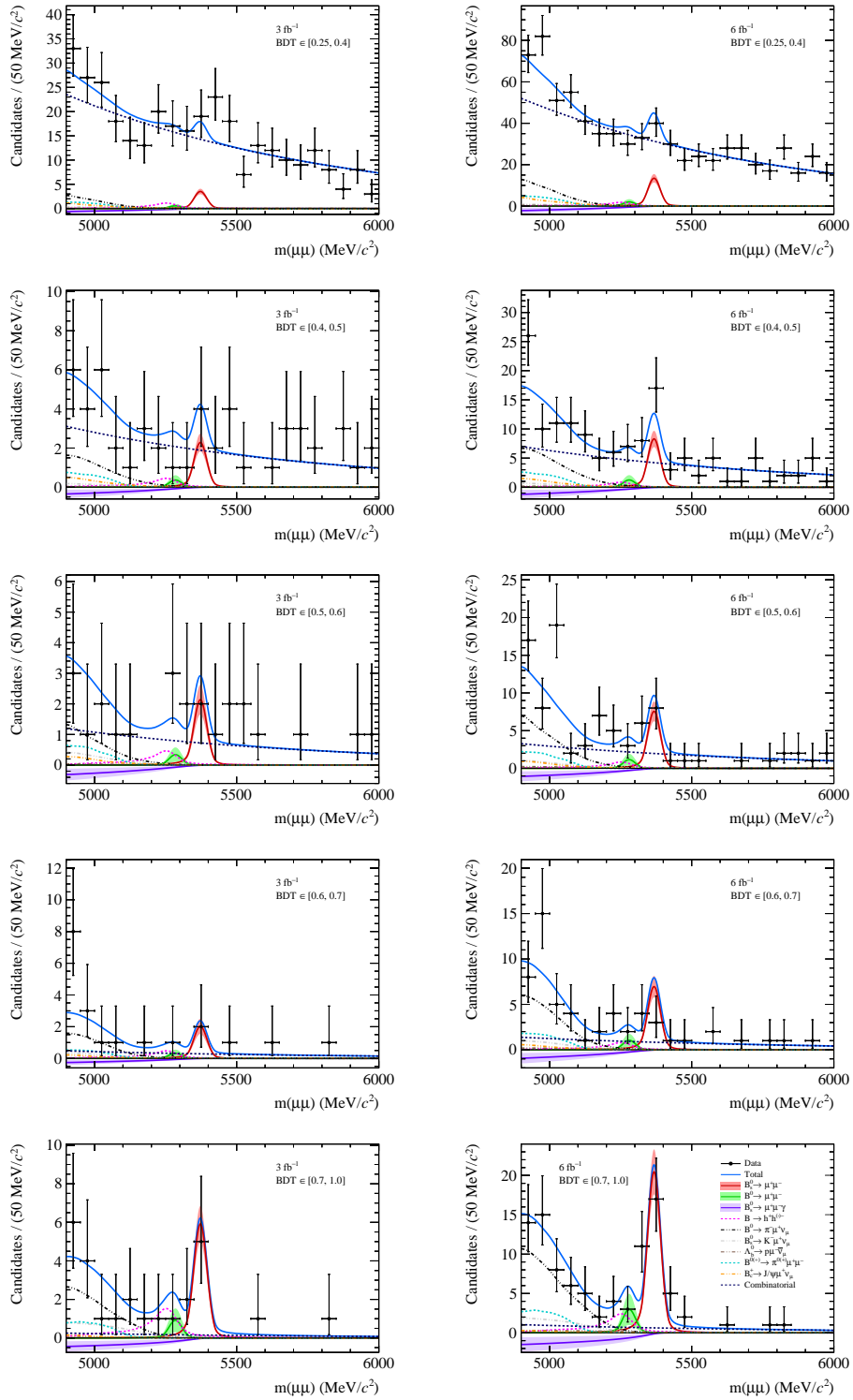


Figure 6.2 – Invariant mass distribution of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ candidates for the (left) Run 1 and (right) Run 2 sample in different bins of the signal BDT output. The fit result with the different fit components is overlaid. For signal, coloured bands indicate the fit uncertainty on the component.

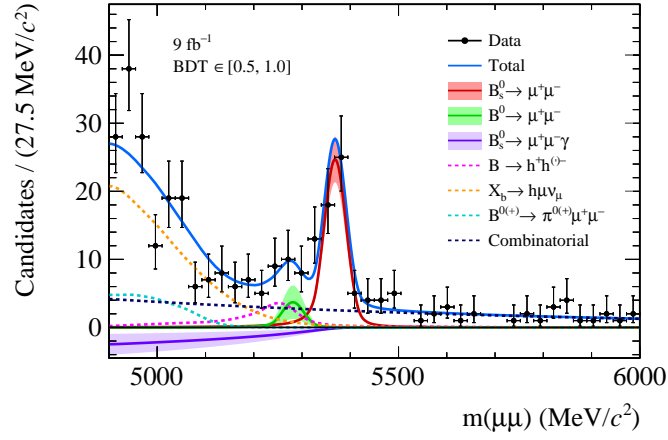


Figure 6.3 – Invariant mass distribution of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ candidates for $\text{BDT} > 0.5$. The fit result with the different fit components is overlaid. For signal, coloured bands indicate the fit uncertainty on the component.

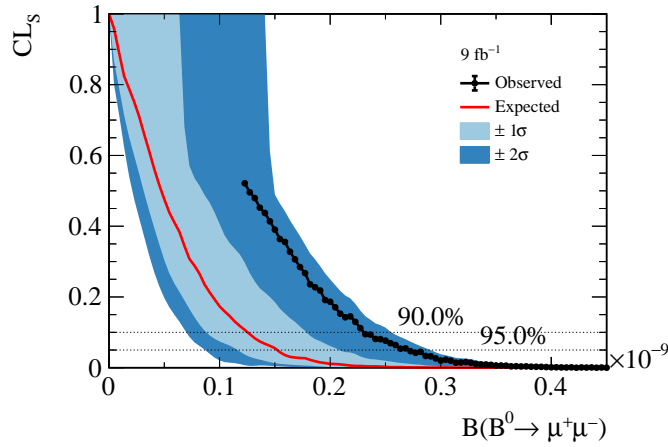


Figure 6.4 – Result of the CL_s scan performed to extract an upper limit on the $B^0 \rightarrow \mu^+ \mu^-$ branching fraction. The observed CL_s value distribution is shown and compared to the expected distribution under the background-only hypothesis derived from pseudo-experiments. Colour bands indicate the 1σ and 2σ regions around the expected values. The 95% confidence level is marked as a dotted line.

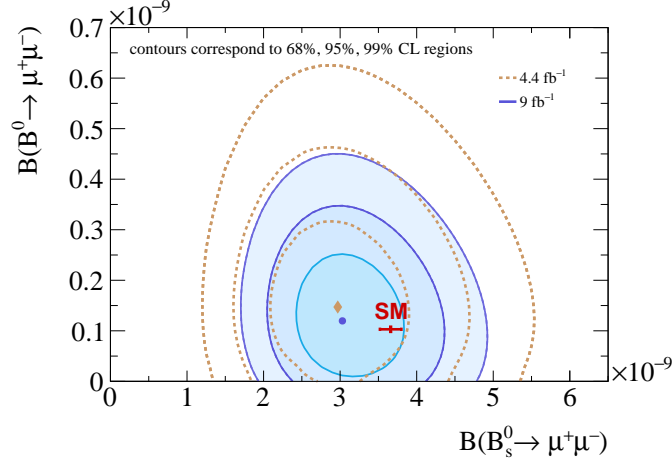


Figure 6.5 – Likelihood contours for the $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ branching fractions. The measured values are marked with a blue dot. The blue contours correspond to the 68 %, 95 %, and 99 % CL regions. The dashed brown contours indicate the results from the previous LHCb measurement [64]. The SM values [50] are shown in red.

The presented results were the most precise measurements by a single experiment by the time of publication. Recently, the CMS collaboration published an even more precise measurement of the branching fractions [67]. A combination of the LHCb with this result is still pending. Both results will bring the LHC average of the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction closer to the theoretical prediction. Nevertheless, more investigations on larger data samples are needed to clarify anomalies seen in $b \rightarrow s\ell^+\ell^-$ decays and answer the question of NP contributions in this sector. Additionally, more data is needed for an observation of the $B^0 \rightarrow \mu^+\mu^-$ mode and a more precise determination of $\mathcal{R}_{\mu^+\mu^-}$, which will allow testing the minimal flavour hypothesis with more certainty.

Since 2022, the LHCb detector is recording data again. For this purpose, it was upgraded with a new particle tracker based on scintillating fibers (the SciFi tracker), a newly built VELO, and a completely redesigned trigger system. It is planned that this Upgrade I detector [159] and the foreseen Upgrade II detector (running under the conditions of the high-luminosity LHC) collect a data sample of around 300 fb^{-1} [160]. With this data sample, it is expected to observe the $B^0 \rightarrow \mu^+\mu^-$ decay and reduce the uncertainty on $\mathcal{R}_{\mu^+\mu^-}$ to 10%. The statistical uncertainty on the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction is projected to be reduced to 1.8%, which is smaller than the expected systematic uncertainty of 4%. The total relative uncertainty is expected to be around 5%, which would be competitive to the uncertainties of the current theoretical predictions. The corresponding projections of the CMS collaboration for the same time period yield a relative uncertainty of the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction of about 11% percent at 3000 fb^{-1} . Here, the measurement of CMS is significantly limited by its muon trigger under the high-luminosity LHC conditions. This indicates that

also in future LHC runs the LHCb experiment will take a leading role in the measurement of $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays.

Next to the improved branching fraction measurements, the larger data sample will serve for a more precise determination of the effective $B_s^0 \rightarrow \mu^+ \mu^-$ lifetime, which was also measured in the presented analysis. Details on this study can be found in Ref. [124]. Future and more precise measurements of this observable will lead to a better understanding of the time-dependent CP asymmetry of $B_s^0 \rightarrow \mu^+ \mu^-$ decays and to constraints on the parameter $\mathcal{A}_{\Delta\Gamma}$ providing a better sensitivity to possible NP contributions in the CP sector.

Part III

First search for $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays

7 Strategy and data preparation

The second analysis presented in this thesis aims to measure the branching fractions of the semi-baryonic $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays for the first time. The final state with two muons and two protons was already studied by LHCb in resonant $B_s^0 \rightarrow J/\psi p\bar{p}$ and $B^0 \rightarrow J/\psi p\bar{p}$ processes, where the J/ψ meson decays muonically [75, 77]. The now presented search is performed in the non-resonant decay channel, with the J/ψ resonance being vetoed in the reconstructed dimuon mass distribution. Processes in the resonant channel are used to normalise the branching fractions following a procedure similar to the one described in Section 6.1 for the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ analysis.

The measured observables are the ratios of branching fractions

$$r_s = \frac{\mathcal{B}(B_s^0 \rightarrow p\bar{p}\mu^+\mu^-)}{\mathcal{B}(B_s^0 \rightarrow J/\psi p\bar{p}) \times \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)} \quad (7.1)$$

and

$$r_d = \frac{\mathcal{B}(B^0 \rightarrow p\bar{p}\mu^+\mu^-)}{\mathcal{B}(B^0 \rightarrow J/\psi p\bar{p}) \times \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}. \quad (7.2)$$

Upper limits on the branching fractions are calculated for the modes for which the measured signal significance does not exceed 5σ above the background-only hypothesis.

The measurement is performed blindly, with the non-resonant signal region discarded for the time the analysis is developed. Only when the analysis strategy has been reviewed by the LHCb collaboration the blinded region is examined. This strategy prevents the measurement from being biased by an intentional analysis design. Figure 7.1 shows the different phase-space regions considered in the analysis. The phase-space of the signal decay is restricted due to the mass of the two protons. Since it is $m(B_s^0) - 2m(p) \approx 3490 \text{ MeV}/c^2$, charmonium resonances heavier than the J/ψ meson are kinematically not allowed. Additionally, the protons in the resonant decay channel are less energetic on average than in the non-resonant channel, which is due to the J/ψ resonance further restricting the available phase space.

The analysis is conducted by a small group of two people, with the author of this thesis being the main contributor. The exploited data and simulated samples are described in Section 7.1, followed by the simulation calibration in Section 7.2. In Chapter 8, the selection requirements of the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ signal candidates are presented. Chapter 9 shows how the invariant mass distributions of the signal are modelled for the branching fraction determination. Finally,

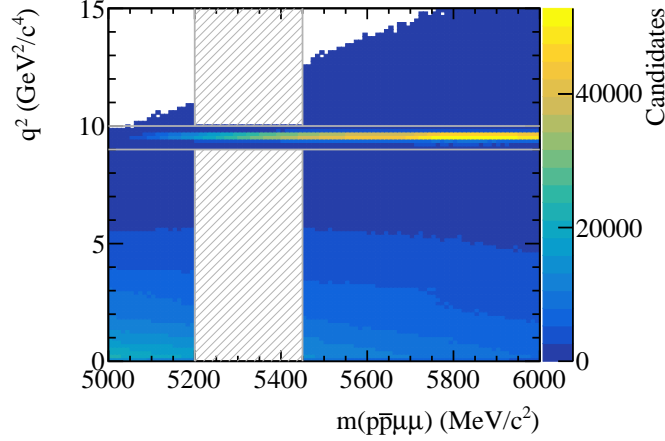


Figure 7.1 – Reconstructed $p\bar{p}\mu^+\mu^-$ mass versus $q^2 = m(\mu^+\mu^-)^2$ distribution in the 2018 data set. Lines separate the different phase-space regions considered in the analysis. The region with $q^2 \in [9, 10] \text{ GeV}^2/c^4$ contains resonant $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ decays and is called the resonant channel. The signal region with $m(p\bar{p}\mu^+\mu^-) \in [5200, 5450] \text{ MeV}/c^2$ is blinded and will not be examined until the analysis strategy is established and reviewed. The remaining regions are referred to as background sidebands used to develop and validate the selection.

the normalisation procedure and preliminary results prepared for this thesis are shown in Chapter 10. By the time of writing, the analysis is still ongoing, and the signal region in the non-resonant channel is blinded. Nonetheless, this thesis calculates and discusses expected upper limits on the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ branching fractions.

7.1 Data and simulation samples

The measurement uses the LHCb proton-proton collision data set collected in Run 1 and Run 2 of the LHC at centre-of-mass energies of 7, 8, and 13 TeV. The combined data set size corresponds to an integrated luminosity of 9 fb^{-1} .

Next to the recorded data, simulated samples are used in the analysis to study the properties of the invariant mass distributions of the signal and normalisation modes, train multivariate classifiers against the background pollution, and measure selection efficiencies. The $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ signal decays are simulated with a generic phase-space decay model, having all matrix elements of the transition amplitude constant in the computation. This model does not describe the decays properly because it leads to only unpolarised decay products with distorted mass and angular distributions. Nevertheless, the model is used because a more appropriate SM calculation of the decays is unavailable. A possible approach to compensate for this mismodelling is discussed in Section 7.2.4 and comprises a calibration of the simulated diproton mass distribution.

To save disk space, the $B_s^0 \rightarrow J/\psi p \bar{p}$ and $B^0 \rightarrow J/\psi p \bar{p}$ samples are simulated as part of a filtered production, which means that the centralised preselection (presented in Section 8.2) is directly applied after the simulation step. The corresponding filtering efficiencies in the order of 40 % are consistently included in the reconstruction and selection efficiencies stated in Section 10.1.

A truth-matching selection is developed to obtain the correct and fully reconstructed signal in the simulation. The selection is based on the background category tool introduced in Section 3.4. The tool assigns each candidate to a category depending on how similar the reconstructed candidate is to the generated signal decay. Depending on the B hadron flavour and the parent particle of the lepton pair, the correctly reconstructed signal decays fall into different categories. In addition to the candidates in these categories, candidates in the category with low reconstructed B mass are accepted. This is done to account for true signal particles that suffer from energy losses due to soft final-state radiation. The truth-matching criteria are consistently applied when the simulation is considered. Some signal candidates might not pass the truth-matching criteria due to inefficiencies in the truth-matching selection. It is planned to treat these candidates in the analysis as one possible source of systematic uncertainty.

Because of the low amount of expected signal decays in data, the branching fractions are measured using the combined Run 1 and Run 2 data samples. When merging the data sets, the corresponding simulation samples of the different years are combined accordingly and weighted to match the luminosity ratio measured in the data.

7.2 Simulation calibration

It is known that the LHCb simulation is not a perfect representation of the signal in data. Because the precision of the analysis depends on reliable signal efficiency estimations and well-described signal proxies in the training of multivariate classifiers, several variables used in the selection must be calibrated in the simulation. The calibration is performed by applying a collection of single-event weights to the simulated samples that change the relative importance of each decay candidate. The following sections describe how these weights are determined and which calibration steps have already been performed.

To compare data with simulation and validate the calibration procedure, a background-free decay sample must be obtained in data similar to a sample of the actual decays under study. A natural choice would be a sample of the already observed resonant $B_s^0 \rightarrow J/\psi p \bar{p}$ decay. However, the amount of signal candidates for this decay is quite low. So far, the $B_s^0 \rightarrow J/\psi p \bar{p}$ mode is only used for loose checks on the invariant mass of the diproton system. Complementary, the abundant decays $B_s^0 \rightarrow J/\psi \phi$ and $B^0 \rightarrow J/\psi K^{*0}$ with $\phi \rightarrow K^+ K^-$ and $K^{*0} \rightarrow K^+ \pi^-$ are exploited, which have a similar topology and trigger response due to the presence of two hadrons and two muons in their final states. These modes are

referred to as further control modes used to calibrate the muon identification response, the event multiplicity distribution, the B meson kinematics, and the trigger response.

The $B_s^0 \rightarrow J/\psi\phi$ and $B^0 \rightarrow J/\psi K^{*0}$ decays are selected in the dimuon region around the J/ψ resonance, with identical trigger requirements and the same centralised preselection used for the signal channel. A summary of the trigger lines and explanations on the preselection are given in Chapter 8. Further selection requirements are applied, as shown in Table 7.1 to select the ϕ and K^{*0} resonances, reduce the number of contributing background processes, and align the samples with calibration data sets.

Table 7.1 – Selection requirements for the $B_s^0 \rightarrow J/\psi\phi$ and $B^0 \rightarrow J/\psi K^{*0}$ decays.

Particle(s)	Variable	Requirement
ϕ	$ m(K^+K^-) - 1020 \text{ MeV}/c^2 $	$< 10 \text{ MeV}/c^2$
K^{*0}	$ m(K^+\pi^-) - 895 \text{ MeV}/c^2 $	$< 50 \text{ MeV}/c^2$
K^\pm	PID_K	> -20
π^\pm	PID_K	< 10
h^\pm/μ^\pm	$\chi_{\text{track}}^2/\text{ndf}$	< 4
	GhostProb	< 0.4
	p	$\in [3, 500] \text{ GeV}/c$
	p_T	$\left\{ \begin{array}{ll} \in [0.80, 40] \text{ GeV}/c & \text{for Run 1} \\ \in [0.25, 40] \text{ GeV}/c & \text{for Run 2} \end{array} \right.$

After selecting the control channel decays, fits to the reconstructed $J/\psi K^+ K^-$ and $J/\psi K^+ \pi^-$ mass distributions are performed with the dimuon mass being constrained to the known J/ψ mass. The fit model consists of a Hypatia function [161] describing the signal and an exponential function describing the background component. The Hypatia function generalises the Crystal Ball function [135], consisting of a hyperbolic core and power law tails. For better convergence behaviour, the complexity of the Hypatia function is reduced by fixing the parameters $\zeta = \beta = 0$ in the function definition, where ζ is a shape parameter of the hyperbolic core and β denotes the asymmetry of the function. In addition, Gaussian constraints from fits to the simulated samples are added to the parameters a_L , n_L , a_R , and n_R , which describe the starting points and slopes of the tails left and right of the hyperbolic core. These constraints are necessary because there are not enough events in the data to describe the tails overlaid by background processes properly. The fits are performed individually for all data-taking years. The results for the 2018 data set are shown as an example in Fig. 7.2, while the results for the other years look similar.

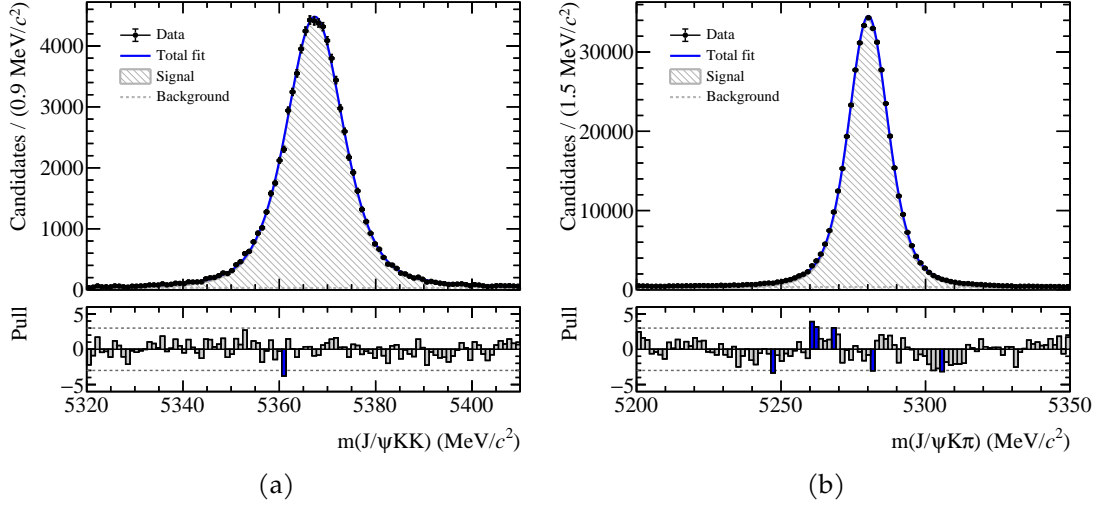


Figure 7.2 – Fit to the reconstructed (a) $J/\psi K^+ K^-$ and (b) $J/\psi K^+ \pi^-$ mass distributions in the $B_s^0 \rightarrow J/\psi \phi$ and $B^0 \rightarrow J/\psi K^{*0}$ control channels. The dimuon mass is constrained to the J/ψ mass to increase the signal resolution. The fit is used for statistical background subtraction with the sPlot method to perform comparison studies between data and simulation. As an example, the 2018 data set is shown.

After fitting the mass models to data, all free fit parameters except for the yields are fixed to their fitted values and the sPlot method [131] is used to subtract the background component. This procedure is identical to the procedure described in Section 4.3 for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ analysis. The weighted distributions are used to compare the signal distributions in control channel data with simulation. The sPlot method is only valid without correlations between the discriminating variable, which is the reconstructed $J/\psi h^+ h^{(\prime)-}$ mass, and the probed variables. Essentially, this requirement is satisfied thanks to the excellent resolution of the invariant B mass, obtained with the help of the J/ψ mass constraint applied to the dimuon system.

The selection of $B_s^0 \rightarrow J/\psi p \bar{p}$ decays follows the selection of $B_s^0 \rightarrow p \bar{p} \mu^+ \mu^-$ decays described in Chapter 8, which is possible due to the same final state considered. Only the selected q^2 region and the selection requirements on the multivariate classifiers are different. A fit to the invariant $J/\psi p \bar{p}$ mass distribution on the combined Run 1 and Run 2 data set is presented in Chapter 9. After fitting, the sPlot method is again used to subtract the data sample's background component.

7.2.1 Particle identification

The first group of variables to calibrate are the particle identification variables. As described in Chapter 8, requirements on the PID variables are imposed on both the final state muons and hadrons. A data-driven approach is chosen for the calibration, exploiting a resampling method that replaces the PID response

for each track with a new one generated from parameterised PID distributions in data. These PID distributions on data are described using clean data sets for particular decays provided by the LHCb PIDCalib group [128]. The calibration data sets contain $J/\psi \rightarrow \mu^+ \mu^-$ decays from inclusive b decays and $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ decays, with subsequent $\Lambda_c^+ \rightarrow p K^- \pi^+$ decays. From these samples, the muon and proton PID distributions can be accessed. In the calibration samples, the signal is selected without PID requirements, and the remaining background is subtracted by isolating the signal component with the sPlot method.

The distributions in the analysis samples are calibrated by fitting the calibration samples with kernel density estimators (KDE) in multiple phase-space regions and randomly drawing new values for the PID variables from the resulting probability density functions. This method provides good results when considering only one variable per particle track. However, it is limited in preserving all information by the loss of correlations between the variables of the same track. An alternative approach aims to transform the existing PID variables in simulation instead of resampling them, preserving the correlations. For this approach, additional KDE fits to large simulated samples of the calibration modes are required. The decision of whether to use this transformation or the above resampling approach is based on the question of whether the correlations for a single track must be conserved for further analysis steps. With both methods, the new distributions in the simulation are expected to show higher agreement with the signal in data.

Selection requirements on the various PID variables are applied at different stages of the analysis. For muons, the isMuon requirement is imposed together with a cut of $\text{PID}_\mu > -3$ in the centralised preselection. In the subsequent selection, the PID_μ and $\text{ProbNN}(\mu)$ distributions of both muons enter in a multivariate classifier training (see Chapter 8). To align the selection of muons in the calibration and analysis samples, the hasRich flag and $3 < p < 150 \text{ GeV}/c$ are required on each muon track. For the Run 1 samples, an additional requirement on the muon's transverse momentum of $p_T > 800 \text{ MeV}/c$ is applied due to strict requirements in the corresponding calibration sample.

The resampling method is used to calibrate the muon PID variables since no other PID responses from the muon tracks are used in the signal selection. The distributions of the resampled PID variables are parameterised using three variables that are correlated with the PID response: the transverse momentum p_T and pseudorapidity η of the track, and the number of tracks in the event (nTracks). The binning in three variables is performed to approximately preserve the kinematic dependence of the PID response and the correlations between the different tracks of the event. While the first two variables are well-modelled in simulation, the event multiplicity is known to be mismodelled since it relies on the underlying event, which is not perfectly described. There are two possible ways out: The first approach is to assume that only the PID efficiency depends on the multiplicity and run the resampling method with a calibrated nTracks distribution while using the uncalibrated distribution for the rest of the analysis. The second

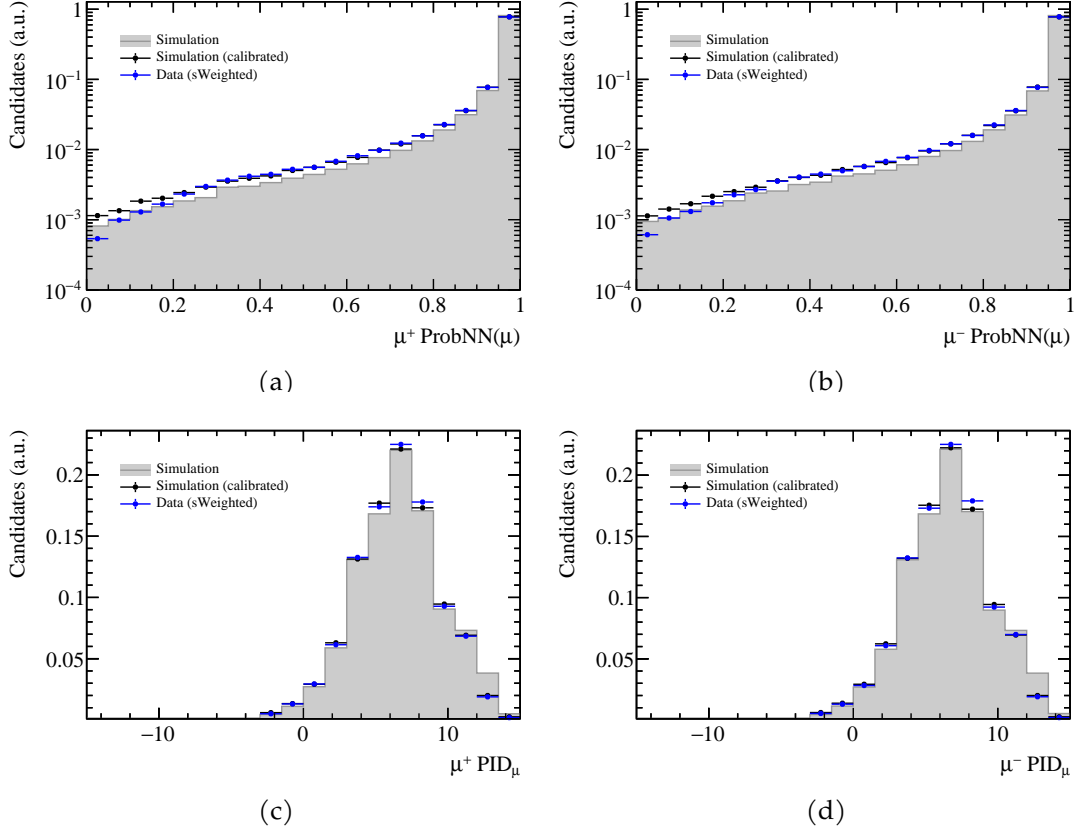


Figure 7.3 – Comparison of muon PID distributions in 2018 $B^0 \rightarrow J/\psi K^{*0}$ data with calibrated and plain simulation. All corrections are applied to the simulation, improving the compatibility with the signal seen in sWeighted data.

approach explicitly calibrates the nTracks distribution as part of the correction chain. Since both approaches are correct to the same extent, it was decided to follow the second approach and calibrate the nTracks distribution in conjunction with the B meson kinematics. This calibration is detailed in Section 7.2.3.

The resulting comparison of the muon PID distributions in 2018 data and calibrated simulation is presented in Fig. 7.3. A much better agreement of the muon PID distributions between the calibrated simulation and the sWeighted data is seen.

Next, the proton PID variables must be calibrated. No requirements on these variables are set in the centralised preselection. However, in the further selection, several proton PID and ProbNN variables serve as inputs for a multivariate signal classifier, which is only based on PID information. Because both, the kaon and proton PID responses of particles, are exploited, the transformation method is used to preserve the correlations between the variables belonging to the same track. As for the muon PID calibration, the variable transformation is performed in bins of the transverse momentum and pseudorapidity of the particle track, as well as the event multiplicity.

The validation for this calibration is still pending. Since the $B_s^0 \rightarrow J/\psi p\bar{p}$ sample is too small to draw any meaningful conclusion from the sWeighted distributions, a larger and clean sample of protons must be obtained to compare the simulation and the signal in data properly. One possible channel could be $\Lambda_b^0 \rightarrow pK^-\mu^+\mu^-$, which can be selected with high efficiency. However, the proton kinematics need to be checked beforehand and compared to the ones in the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ sample as they can differ due to lifetime differences of the Λ_b^0 baryon and the B mesons.

7.2.2 Trigger response

A calibration of the trigger response is needed since the exact trigger configuration was changed multiple times during data-taking, while the simulated samples were produced with only one specific configuration. In addition, the L0 trigger response is not well-modelled in simulation, mainly due to the insufficient description of the underlying events. As detailed in Section 8, at the L0 stage, a positive decision of the single-muon or dimuon trigger is required. The single-muon trigger uses configuration-dependent requirements based on the muon's transverse momentum and the event multiplicity. The dimuon trigger applies a requirement based on the product of the transverse momenta of the two muons and also the event multiplicity. Given these selections, using the uncalibrated trigger response from simulation would lead to wrong efficiency estimates. Hence, the L0 trigger calibration is computed using the data-driven TISTOS method [123]:

On data, the total number of events before the trigger selection is unknown. The TISTOS method assumes that the fraction of TIS-triggered and TOS-triggered events are independent. With this, the number of events before the trigger selection can be stated as

$$N_{\text{tot}} = \frac{N_{\text{TOS}}}{\epsilon_{\text{TOS}}} = \frac{N_{\text{TIS}}N_{\text{TOS}}}{N_{\text{TISTOS}}}, \quad (7.3)$$

with the TOS efficiency being calculated as the fraction of TISTOS triggered events on the TIS subsample:

$$\epsilon_{\text{TOS}} = \frac{N_{\text{TISTOS}}}{N_{\text{TIS}}}. \quad (7.4)$$

The correction weights to simulation are then extracted for each year individually by calculating the ratio of the TOS trigger efficiencies measured on data and simulation:

$$w = \frac{\epsilon_{\text{TOS}}^{\text{Data}}}{\epsilon_{\text{TOS}}^{\text{MC}}}. \quad (7.5)$$

Because the kinematics of the parent B mesons are correlated with the underlying event and the L0 trigger decisions are based on the transverse momenta of their daughter particles, the assumption of independent TIS and TOS samples

only holds for small phase-space regions. Therefore, the resulting correction weights are measured as a function of the transverse momenta of both muons. The two-dimensional binning scheme is determined using a k -dimensional tree binning algorithm, creating bins with variable sizes that are uniformly filled. This approach obtains weights with the same relative uncertainties independent of the phase-space region. The binning scheme is calculated for each year individually based on the smallest subsample, which is the TISTOS selected data sample by construction. The two-dimensional correction maps are then used to assign per-event weights to all simulated samples.

Because the signal channel is blinded and the $B_s^0 \rightarrow J/\psi p\bar{p}$ signal yield is too small, the $B^0 \rightarrow J/\psi K^{*0}$ control channel is used to calculate the L0 trigger correction. The corresponding data and simulation samples are selected as described in Section 7.2, but instead of requiring the TOS trigger decisions, only signal-independent events with at least one positive TIS decision are considered. From these samples, the number of TIS and TISTOS events are determined. Figure 7.4 shows the trigger response calibration maps for the different data-taking years. While the majority of weights are close to one, the weights in the region of low transverse momentum are substantially larger for the 2015 and 2016 data. Throughout these years of data taking, many trigger configurations with different muon trigger criteria were used. The simulated samples, however, were generated with only a single configuration that was used to collect the majority of the data. The effect of this in the weights is seen at low transverse momenta in the region where data was recorded with lower thresholds than the corresponding simulation.

In contrast to the L0 trigger, the HLT trigger response is assumed to be well-modelled in simulation as it is known to be the case from other analyses with dimuon final states. The reason for this is the excellent muon track identification and reconstruction. No attempt is made in this analysis to check the assumption due to the low amount of data in the $B_s^0 \rightarrow J/\psi p\bar{p}$ channel and the resulting large statistical uncertainties.

7.2.3 Event multiplicity and B hadron kinematics

The next calibration concerns the simulated event kinematics. In the LHCb simulation, the transverse momentum dependence of produced $B_{(s)}^0$ mesons is not correctly modelled. Also, the event multiplicity is usually underestimated because of the mismodelling of the underlying event.

To correct these variables, the simulation samples are weighted using gradient-boosted reweighters (GBR) [130] trained for each data-taking year individually. To prevent training on statistical fluctuations of the input data, the k -folding method [132] with $k = 5$ is performed, while for the weight predictions on independent data sets, the mean prediction of all five reweighters is used. The procedure is the same as documented for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ analysis in Section 4.3. The trained reweighters take as inputs the transverse momentum p_T and pseu-

7 Strategy and data preparation

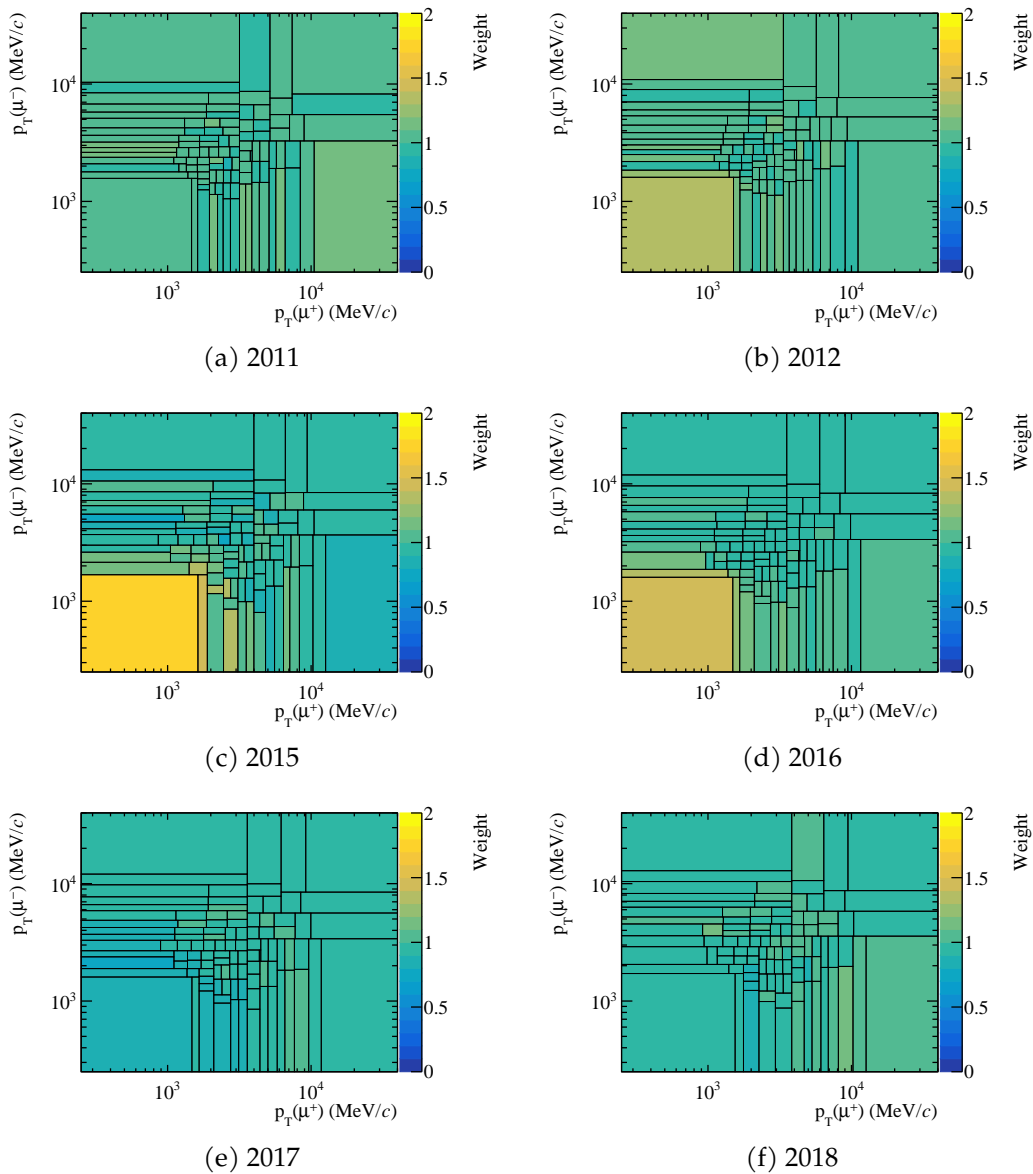


Figure 7.4 – L0 trigger correction maps derived from the $B^0 \rightarrow J/\psi K^{*0}$ data sample. The weights are given as a function of both muon transverse momenta. The binning schemes are determined on the TISTOS selected sample with a k -dimensional tree binning algorithm to get weights with stable relative uncertainties.

rapidity η of the $B_{(s)}^0$ candidate and the number of tracks in the event (nTracks) as a proxy for the event multiplicity. Since the hadronisation processes of the decaying B mesons are different with respect to the kinematics of the resulting particles, the reweighting for the B_s^0 and B^0 modes is performed with different control channels. For B^0 decays, the $B^0 \rightarrow J/\psi K^{*0}$ channel is chosen, while for B_s^0 decays, the $B_s^0 \rightarrow J/\psi \phi$ channel is used. As an example, the calibrated input distributions are shown in Fig. 7.5 for the 2018 data set. The calibration weights are attached as per-event weights to all simulated samples.

7.2.4 Decay model

A further calibration concerns the decay model used for the simulated samples. The rate of a given decay is proportional to the square of its transition amplitude. The transition amplitudes depend on the QCD form factor parametrisations for the decays, which theoretical calculations must provide. These calculations are not existing for the $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ and $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decay modes. Therefore, the simulated samples were produced by setting the matrix elements of the transition amplitudes as constant, making it a phase-space-only description. Such a description is insufficient for modelling the physical processes because it leads to distorted mass and angular distributions. In the scope of this thesis, the reconstructed diproton mass spectrum is looked at since it can be enhanced by resonances or other effects that affect the quark fragmentation process as discussed in Section 2.4.

Such enhancing contributions are not considered in the phase-space-only simulation. For the resonant $B_s^0 \rightarrow J/\psi p\bar{p}$ and $B^0 \rightarrow J/\psi p\bar{p}$ decays, this can be shown by comparing the reconstructed diproton mass spectrum between sWeighted data and the corresponding simulation. The comparisons are shown in Figs. 7.6a and 7.6b. Despite significant statistical uncertainties in the data sample especially for the B_s^0 mode, a deviation from the simulated sample is visible. Before calibrating the diproton spectra according to the spectra in sWeighted data, further distributions related to the decay model and the selection should be checked and compared. As indicated in Ref. [77], the $J/\psi p$ and $J/\psi \bar{p}$ mass spectra also show resonant structures that may need to be considered. So far, no correction of the decay model in $B_s^0 \rightarrow J/\psi p\bar{p}$ and $B^0 \rightarrow J/\psi p\bar{p}$ simulation has been applied and no systematic uncertainty estimated.

Regarding the non-resonant $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ modes, the diproton spectra in data are unknown and also, no simulation based on SM calculations exists. For a correction of these samples, it is assumed that the diproton spectra are similar to the spectrum in $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays, for which a decay model based on perturbative QCD (pQCD) calculations exist [71]. Figure 7.6c shows the invariant diproton mass distributions of $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays from phase-space-only and pQCD simulation. Both distributions refer to the true diproton properties before any reconstruction. From the comparison, correction weights

7 Strategy and data preparation

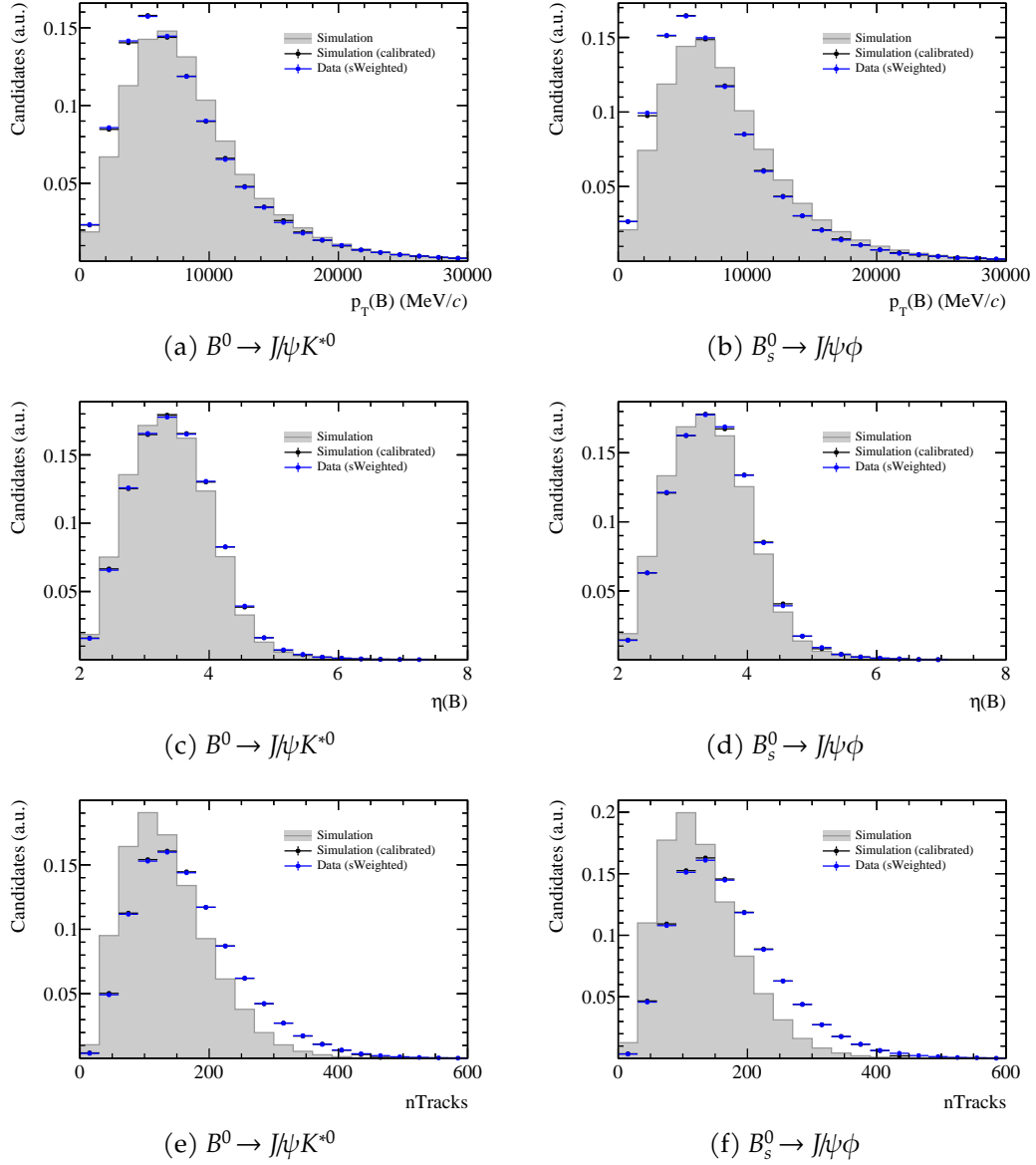


Figure 7.5 – Input distributions for the gradient-boosted reweighters, which are trained in the (left) $B^0 \rightarrow J/\psi K^{*0}$ and (right) $B_s^0 \rightarrow J/\psi \phi$ decay channel. The sWeighted data is compared to plain and calibrated simulation. The 2018 data set is shown.

can be extracted in fine bins of the diproton mass that can be applied as per-event weights to the non-resonant $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ samples. With these weights applied to the signal simulation, the diproton spectra are not expected to agree ideally with the unknown spectra in the data. The kinematics of $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays differ, and the different flavours of the B mesons contribute to a difference in the calculation of the transition amplitude. Nevertheless, the structure of the involved hadronic form factors is expected to be the same. That means that the corrected distributions should describe the decays more accurately than the raw distributions of the phase-space-only simulation. So far, the derived weights are included in the efficiency calculation of the non-resonant signal modes and implemented in the training of the multivariate signal classifiers. In a simple test, the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ signal efficiency is calculated with and without the weights, resulting in a relative difference of around 20%. It is planned to perform further studies on this calibration since insufficient knowledge of the correct decay model is expected to contribute significantly to the systematic uncertainty of the measurement.

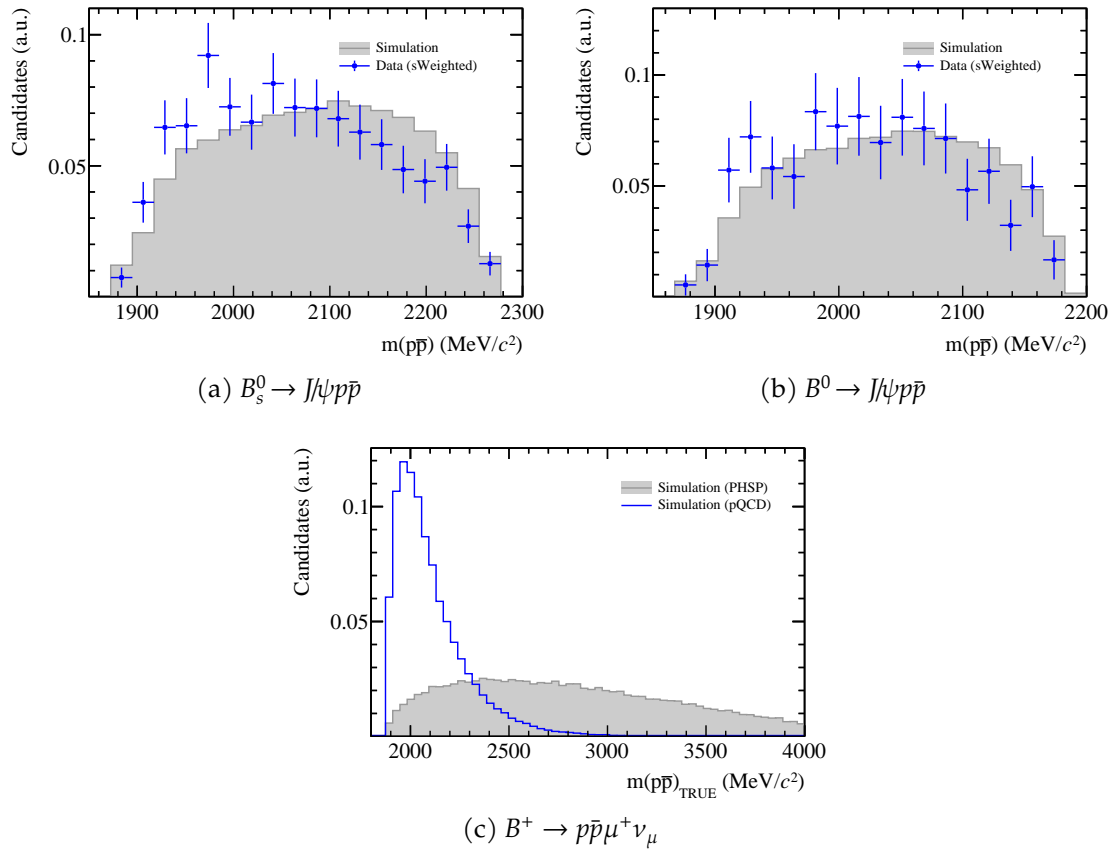


Figure 7.6 – Reconstructed diproton mass distribution of resonant (a) $B_s^0 \rightarrow J/\psi p\bar{p}$ and (b) $B^0 \rightarrow J/\psi p\bar{p}$ decays in phase-space simulation and sWeighted data. (c) Generated diproton mass distribution of $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ decays in phase-space simulation and in simulation produced with the perturbative QCD model developed in Ref. [71].

8 Selection of $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ candidates

This chapter describes the selection of $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ signal candidates. The entire selection consists of the trigger selection, a centralised preselection, additional loose requirements on the track kinematics and track fit qualities, and the selection based on two multivariate signal classifiers. Also, the corresponding resonant decay modes are selected as similarly as possible due to the same final state. The only differences between the resonant and non-resonant decay channel selections are the q^2 regions considered and the requirements on the multivariate classifier outputs.

8.1 Trigger requirements

As explained in Section 3.2.3, the LHCb trigger system includes a hardware trigger and a two-staged high-level software trigger. The trigger stages are organised in different sets of trigger requirements, the trigger lines, each optimised for different physics use cases. Trigger lines chosen for the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ analysis are given in Table 8.1. Their choice is motivated by other LHCb searches of rare decays with muons in the final state. On the L0 stage, the muon and dimuon lines exploit the measurement of the tracks' transverse momenta in the muon chambers to efficiently select events containing high-energetic muons. While the single-muon trigger sets a lower threshold on the transverse momentum for a single muon, the dimuon trigger decides based on the product of the two largest muon transverse momenta. In addition, both L0 lines impose a requirement on the event multiplicity by setting an upper threshold on the number of recognised hits in the SPD detector (nSPDHits). This threshold suppresses events with many particle tracks for which the event reconstruction would take too long. The L0 trigger threshold settings were changed multiple times during data-taking to stabilise the detector read-out dependent on the experimental conditions. In Table 8.2, the threshold values for the transverse momentum and nSPDHits employed to record the majority of data for each year are given [108, 109, 162]. For the events that pass the L0 trigger, on the HLT1 stage, a reconstruction of the particle tracks and primary vertices is performed. At this stage, the TrackMuon and TrackAllL0 trigger lines accept all events that passed the L0 muon trigger or the L0 stage in general. In contrast, the TrackMVA and TwoTrackMVA trigger lines are based on multivariate classifiers. They filter the events based on

properties like the minimum transverse momentum and the impact parameter significance (with respect to the PV) of the tracks. Finally, in the HLT2 stage, the events are fully reconstructed. The DiMuonDetached trigger line filters on properties typical for dimuon vertices, such as the distance of PV and SV. Additionally, a series of inclusive B trigger lines is included, which are based on boosted decision trees and select decays with a two-, three-, or four-body topology. Some of these lines explicitly require identified muons to be part of the decay.

For all L0 and HLT lines, the corresponding TOS requirements are implied, avoiding a strong dependence on the underlying event, which is not well-modelled in simulation, while keeping the signal efficiency high.

Table 8.1 – Summary of trigger lines used to select $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays.

	Run 1	Run 2
L0	Muon	Muon
	DiMuon	DiMuon
HLT1	TrackMuon	TrackMuon
	TrackAllL0	TrackMVA
		TwoTrackMVA
HLT2	DiMuonDetached	DiMuonDetached
	Topo2BodyBBDT	Topo2Body
	Topo3BodyBBDT	Topo3Body
	Topo4BodyBBDT	Topo4Body
	TopoMu2BodyBBDT	TopoMuMu2Body
	TopoMu3BodyBBDT	TopoMuMu3Body
	TopoMu4BodyBBDT	TopoMuMu4Body
		TopoMu3Body

Table 8.2 – Thresholds of the L0 muon and dimuon trigger lines [108, 109, 162].

		2011	2012	2015	2016	2017	2018
		L0Muon					
$p_T(\mu)$ (GeV/c)	>	1.48	1.76	2.8	1.8	1.35	1.35
nSPDHits	<	600	600	450	450	450	450
		L0DiMuon					
$p_T(\mu_1) \times p_T(\mu_2)$ (GeV ² /c ²)	>	1.69	2.56	1.69	2.25	1.69	1.69
nSPDHits	<	900	900	600	600	600	600

8.2 Preselection and multivariate classification

The candidates in the signal and control channels are constructed from events that passed the centralised preselection summarised in Table 8.3. This centralised preselection serves the purpose of harmonising parts of the selection requirements for similar decay modes across the experiment's different analysis groups. Explanations of the different variables are given in Table 3.1. This selection step aims to remove large parts of the background while keeping the signal efficiency high. Requirements on the dimuon, diproton, and $B_{(s)}^0$ meson vertex quality $\chi_{\text{vtx}}^2/\text{ndf}$ and on the DOCA between the tracks reduce combinatorial background contributions. This background reduction is also achieved by cutting on the cosine of the angle between the flight direction of the B meson and the vector joining the PV with the SV (DIRA). In the case of a correct and fully reconstructed decay, this variable is expected to be close to one since no momentum was carried away by one or more undetected or wrongly attributed particles. The displaced B decay vertex is selected by requirements on the flight distance significance χ_{FD}^2 and the χ_{IP}^2 of both the B meson and its decay products. Particle identification requirements are imposed on the muon tracks using the PID_μ and isMuon variables introduced in Section 3.4.

Further loose selection requirements are applied on top of the centralised preselection throughout the analysis to select the relevant $B_{(s)}^0$ mass range and align the signal and control channel data with calibration samples. These additional requirements are summarised in Table 8.4.

As the last part of the selection, the separation power of multivariate classifiers is exploited to suppress background contributions efficiently. Two extreme-gradient-boosting decision tree classifiers (BDT) [163] are trained. This algorithm is chosen for its high efficiency, flexibility and speed of training.

The strategy is to have two complementary classifiers: The first classifier is trained with kinematic and topological variables, while the second classifier is trained with PID information only. Although using a multivariate signal classifier with PID input variables is not a common practice in searches for rare decays at LHCb, the strategy in this analysis serves the crucial purpose of efficiently separating protons from more abundant pions and kaons. Because of the high separation power of one single classifier and the limited data available for training, both classifiers are trained on the same input data, meaning that the training processes are performed in parallel and not consecutively. A study using only one classifier with kinematic and PID variables as inputs is planned and might serve as a cross-check for the actual selection strategy.

Two data samples are prepared for the training to represent the signal and background contributions. The simulated $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ candidates are used as a signal proxy with trigger and preselection applied. All corrections described in Section 7.2 are considered. As a proxy for the background processes, the upper and lower sideband data is used. The lower sideband is defined by the reconstructed $p\bar{p}\mu^+\mu^-$ mass in the range from 5000 to 5200 MeV/c^2 , while

Table 8.3 – Requirements of the centralised preselection.

Particle(s) or Event	Variable	Requirement
$B_{(s)}^0$	$m(hh\mu\mu)$	$> 4700 \text{ MeV}/c^2$
	$m(hh\mu\mu)$	$< 7000 \text{ MeV}/c^2$
	$ \sum_{hh\mu\mu} Q $	< 3
	$\chi_{\text{vertex}}^2/\text{dof}$	< 8
	$\chi_{\text{FD}}^2(\text{PV})$	> 64
	$\chi_{\text{IP}}^2(\text{PV})$	< 16
	$\max(\chi_{\text{IP}}^2(\text{daughters, PV}))$	> 9
	DIRA	> 0.9999
h^+h^-	$m(hh)$	$< 6200 \text{ MeV}/c^2$
	DOCA χ^2	< 20
	$\chi_{\text{vertex}}^2/\text{dof}$	< 12
	$\chi_{\text{FD}}^2(\text{PV})$	> 16
	$\max(\chi_{\text{IP}}^2(\text{daughters, PV}))$	> 6
	DIRA	> -0.9
h^\pm	hasRich	True
	$\chi_{\text{IP}}^2(\text{PV})$	> 6
	GhostProb	< 0.5
$\mu^+\mu^-$	$m(\mu\mu)$	$< 7100 \text{ MeV}/c^2$
	$\chi_{\text{vertex}}^2/\text{dof}$	< 12
	$\chi_{\text{FD}}^2(\text{PV})$	> 9
	$\max(\chi_{\text{IP}}^2(\text{daughters, PV}))$	> 6
	DIRA	> -0.9
μ^\pm	hasMuon	True
	isMuon	True
	PID_μ	> -3
	$\chi_{\text{IP}}^2(\text{PV})$	> 6
	GhostProb	< 0.5
Event	nTracks	< 600

Table 8.4 – Additional selection requirements to select $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays.

Particle(s)	Variable	Requirement
$B_{(s)}^0$	$m(hh\mu\mu)$	$> 5000 \text{ MeV}/c^2$
	$m(hh\mu\mu)$	$< 6000 \text{ MeV}/c^2$
h^\pm/μ^\pm	$\chi_{\text{track}}^2/\text{ndf}$	< 4
	GhostProb	< 0.4
	p	$\in [3, 500] \text{ GeV}/c$
	p_T	$\begin{cases} \in [0.80, 40] \text{ GeV}/c & \text{for Run 1} \\ \in [0.25, 40] \text{ GeV}/c & \text{for Run 2} \end{cases}$

the upper sideband is defined from 5450 to 6000 MeV/c^2 . The selection of the background proxy is made consistently with the selection of the signal proxy. The same classifiers are used for the resonant and non-resonant q^2 regions, but the requirements on the classifier outputs are optimised separately for the different ranges. This strategy ensures that the systematic uncertainties induced by the classifier selection will partly cancel in the normalisation ratio. However, due to the resonant channel's different background composition, the classifiers are trained only in the non-resonant q^2 region.

The classifier training is performed independently for Run 1 and Run 2 data to consider the different data-taking conditions. In total, around 83 000 simulated signal candidates and 1 200 000 background candidates from data are used for training the Run 1 classifiers, while around 195 000 signal and 2 340 000 background candidates are used for the Run 2 classifiers.

The classifier algorithm has several hyperparameters which modify the classification model and its complexity. The parameters that differ from the default value of the algorithm are given in Table 8.5. To lower the probability of an over-trained model that corresponds too closely to the training data set, the maximum depth of the decision trees and the learning rate are decreased. One common way to compensate for the slower learning rate would be to increase the number of estimators (which is equivalent to the number of boosting iterations). In this analysis, however, another approach is chosen by not fixing the number of estimators but introducing an early-stopping routine that stops the learning procedure when no classification improvement was achieved over the last ten boosting iterations. Additionally, a scaling factor is set automatically within the training process to account for the high signal and background candidates imbalance in the training data. This factor is calculated as the ratio of available signal and background candidates.

Regarding the choice of input variables, a set of 13 variables is used for training the kinematic classifier, while for the PID classifier, 11 variables are chosen. A list

Table 8.5 – Hyperparameters used for the classifier training.

Parameter	BDT (kinematic)	BDT (PID)
$N_{\text{estimators}}$	chosen by early-stopping	
max(tree depth)	4	4
learning rate	0.05	0.05
w_{scale}	automatically set	

of the used variables ranked by their importance for the classifiers can be found in Fig. 8.1. The corresponding distributions are shown in Figs. 8.2 and 8.3. Each plot compares the data sideband representing background with the simulation representing signal. The sets of discriminating input variables are determined systematically by a recursive feature elimination (RFE) to avoid redundancies in the information by different variables. The RFE procedure starts with larger sets of standard input variables. First, the classifier is trained with all these variables. Then, the least important variables, as determined by the algorithm, are dropped in an iterative procedure until a termination criterion defined by a metric is reached. This metric is chosen as the area under the receiver operating characteristics (ROC) curve. The ROC curve results from plotting the rate of true positive classifications against the rate of false positive classifications. For this metric, a hypothetical all-knowing classifier shows an area under the curve (AUC) of 1.0, while a randomly guessing classifier has an AUC of 0.5. Based on this metric, the termination criterion for the RFE is chosen to be reached by a design choice when the relative difference to the classifier trained with all variables is more than 0.1 %.

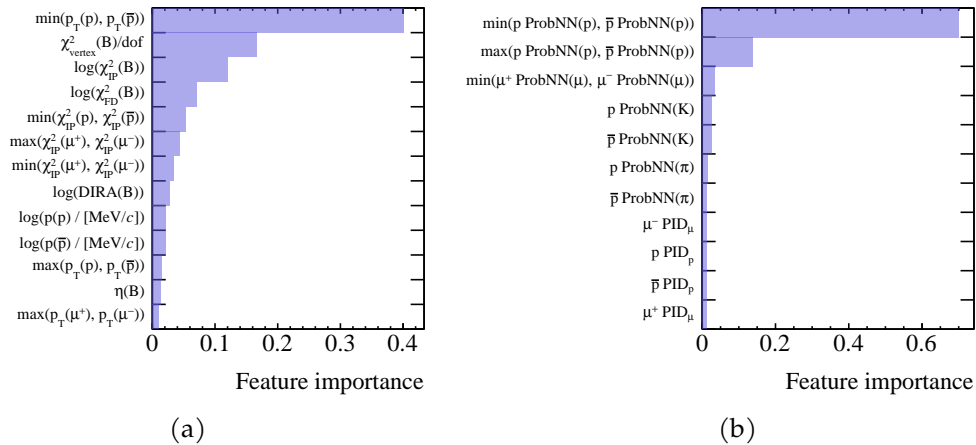


Figure 8.1 – Variables used for the training of the (a) kinematic and (b) PID classifier. The variables are ranked by their importance as determined by the classification algorithm.

8.2 Preselection and multivariate classification

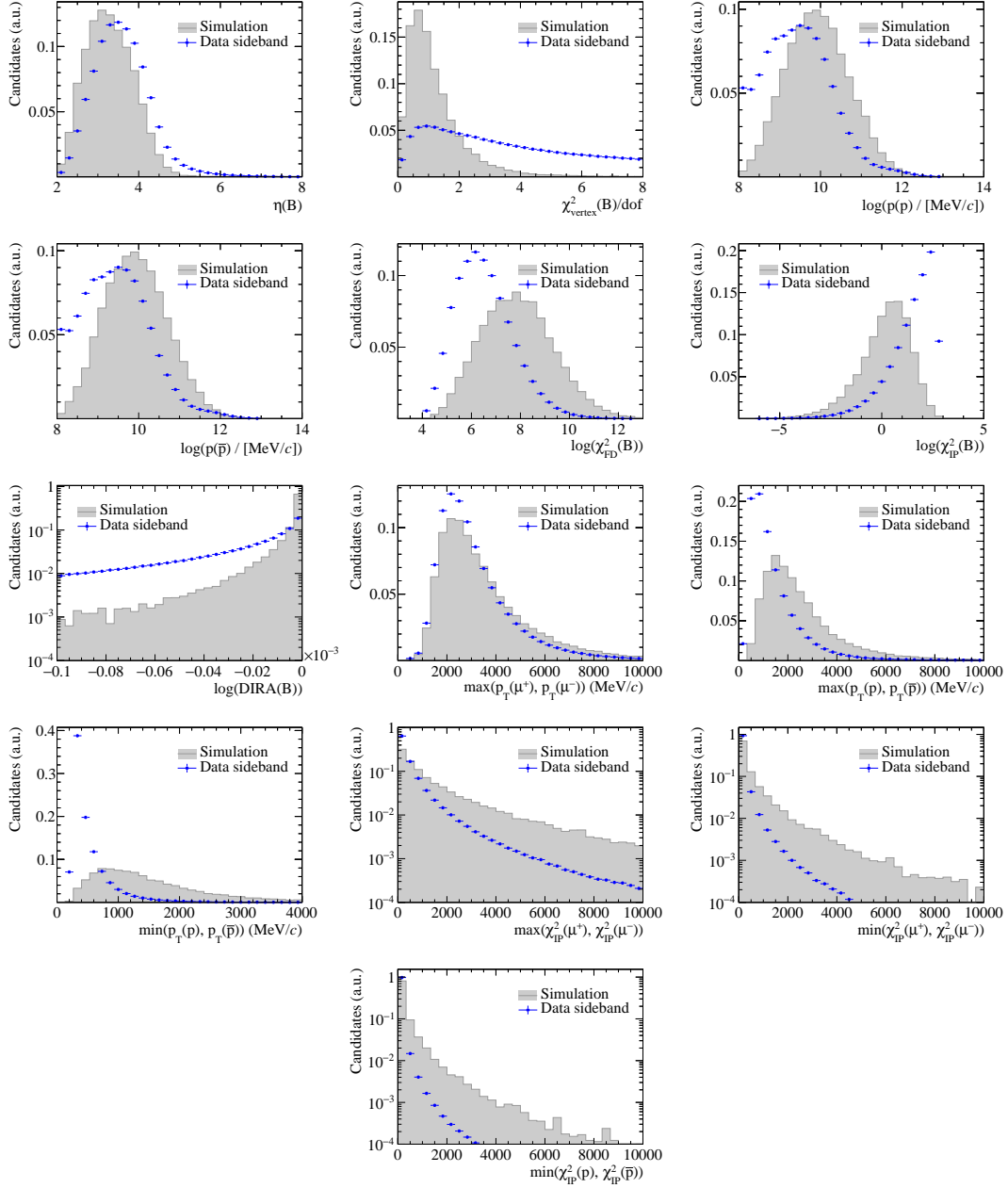


Figure 8.2 – Distributions of the variables used as training inputs for the kinematic classifier. The Run 2 sideband data is compared with signal simulation.

8 Selection of $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ candidates

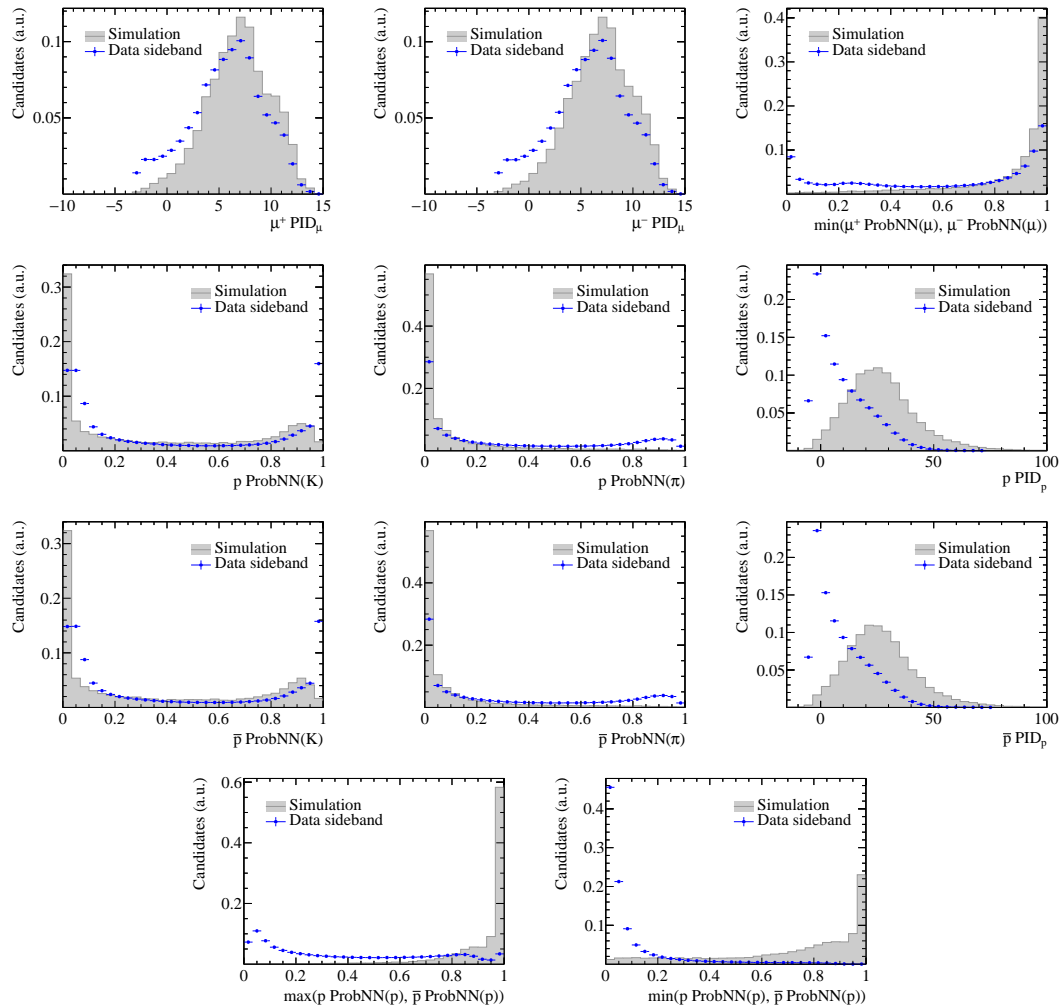


Figure 8.3 – Distributions of the variables used as training inputs for the PID classifier. The Run 2 sideband data is compared with signal simulation.

To check the classifiers for potential overtraining and increase the data statistics, the training is performed with a k -folding approach where $k = 5$. This procedure is identical to the k -folding procedure described in Section 7.2.3 for the kinematic reweighting. While a classifier is trained with 80 % of the available data set, the remaining 20 % is used to check the classifier prediction. By rotating the fold assignment of the data, five versions of the classifier are trained and combined. When predicting unknown data, the arithmetic mean of the five single classifier predictions is taken as result. In contrast, when predicting data used for training, it is always ensured that only the independent classifier is used for the prediction. Comparisons of the classifier responses on the training and test data set are shown in Fig. 8.4 for the Run 2 classifiers, together with the corresponding ROC curves.

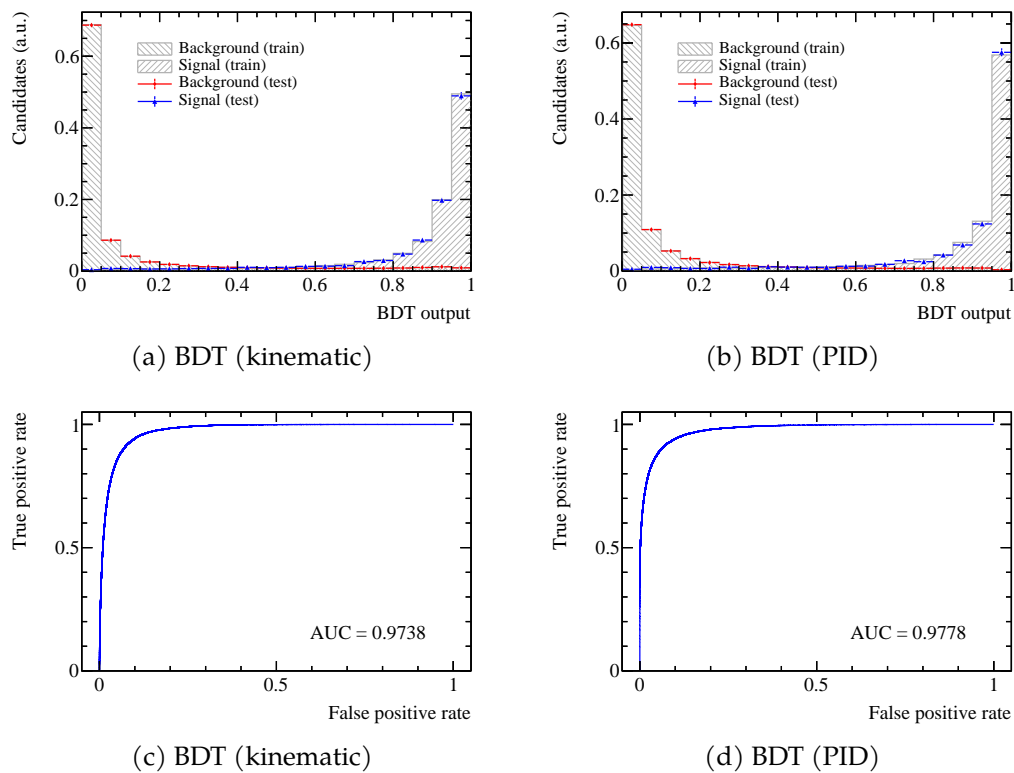


Figure 8.4 – Predictions of the Run 2 (a) kinematic and (b) PID classifiers on the training and test data sets. The corresponding ROC curves are shown below.

8.3 Selection optimisation

The selection requirements on both BDT classifier outputs are optimised to achieve a high sensitivity to the studied decay modes. They are optimised independently for the resonant and non-resonant modes. For the $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ mode, the optimisation is done by fitting the data set with different BDT output requirements and then extracting the number of signal and background candidates. With these numbers, a figure of merit (FOM) is evaluated, which is the significance of the signal,

$$\text{FOM}_{\text{significance}} = \frac{S}{\sqrt{S+B}}, \quad (8.1)$$

where S is the number of signal and B the number of background events measured in a $20 \text{ MeV}/c^2$ region around the B_s^0 mass going from 5346 to 5386 MeV. The fits are performed with a simplified signal model that is not accurate enough for the final mass fit but sufficient for the BDT cut optimisation studies. The model consists of a double-sided asymmetric Crystal Ball function [135] that can be brought to convergence more quickly than the later used and more complex Hypatia function (see Chapter 9). A second-order Chebychev polynomial describes the background. For fits on data, the signal's tail parameters are Gaussian constrained from corresponding fits to simulation. The two-dimensional distributions of the FOM for the Run 1 and Run 2 classifiers are shown in Fig. 8.5. To prove the consistency of the distribution, it is checked that all fits converge for all points on the search grid and that the number of observed background and signal candidates decreases steadily for stricter BDT output requirements. As a result of the $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ selection optimisation, the distributions' maxima are found at

$$\text{BDT (kinematic)} < 0.9 \quad \text{BDT (PID)} < 0.974 \quad \text{for Run 1} \quad (8.2)$$

$$\text{BDT (kinematic)} < 0.425 \quad \text{BDT (PID)} < 0.97 \quad \text{for Run 2} . \quad (8.3)$$

The optimal selection requirements on the BDT classifier outputs for the non-resonant $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays are determined by searching for the maximum of the Punzi FOM [164]. It is defined as

$$\text{FOM}_{\text{Punzi}} = \frac{\epsilon_{\text{sig}}}{\frac{5}{2} + \sqrt{B}}, \quad (8.4)$$

where ϵ_{sig} is the signal efficiency, and B is the expected number of background events at a given selection requirement. The Punzi FOM has the advantage that no assumptions on the decay rate of the rare modes have to be imposed. To avoid convergence problems with the mass fit, the metric is modified with an additional condition requiring at least two candidates left in each sideband

after the selection. The two-dimensional distributions of the Punzi FOM for requirements on the Run 1 and Run 2 classifiers are shown in Fig. 8.6. As a result of the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ selection optimisation, the distributions' maxima are found at

$$\text{BDT (kinematic)} < 0.94 \quad \text{BDT (PID)} < 0.995 \quad \text{for Run 1} \quad (8.5)$$

$$\text{BDT (kinematic)} < 0.94 \quad \text{BDT (PID)} < 0.992 \quad \text{for Run 2} . \quad (8.6)$$

These maxima are not affected by the constraint of at least two candidates per sideband. A comparison of the selection criteria for the resonant and non-resonant channels shows that the optimal kinematic BDT requirement for resonant Run 2 data is significantly lower than for corresponding Run 1 data and corresponding non-resonant data. However, this is not problematic since Fig. 8.5b indicates a broad maximum for different kinematic BDT requirements and consequently the significance depending on this requirement does not change by much.

After applying the entire selection, the data and simulation samples are checked for candidates originating from the same event. Since for rare decays it is highly unlikely that more than one true signal decay occurs in an event, multiple candidates per event mainly result from imperfections in the reconstruction. These multiple candidates are treated by randomly retaining only one candidate and removing the others. In the simulation, if multiple candidates are found, only the candidate in the lowest background category (see Section 3.4) is kept as lower categories contain candidates which are more similar to the truly generated decay. For multiple candidates in the same background category, one candidate is selected randomly, as in data. It turns out that this whole procedure has no significant effect on the signal efficiencies. Only a single multiple candidate in the resonant data set is found and sorted out.

8 Selection of $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ candidates

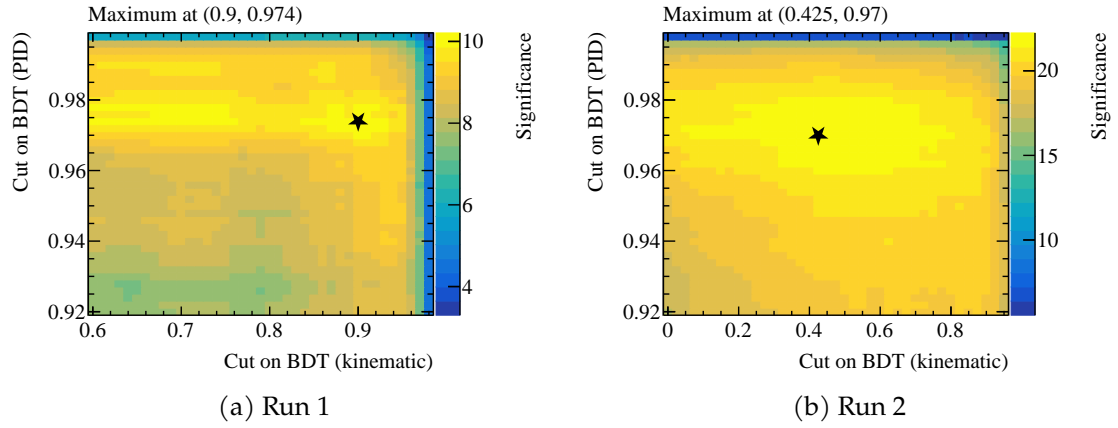


Figure 8.5 – Significance of the $B_s^0 \rightarrow J/\psi p\bar{p}$ signal in a $20 \text{ MeV}/c^2$ region around the B_s^0 mass for different cuts on the kinematic and PID classifier. The distributions' maxima are marked with a star.

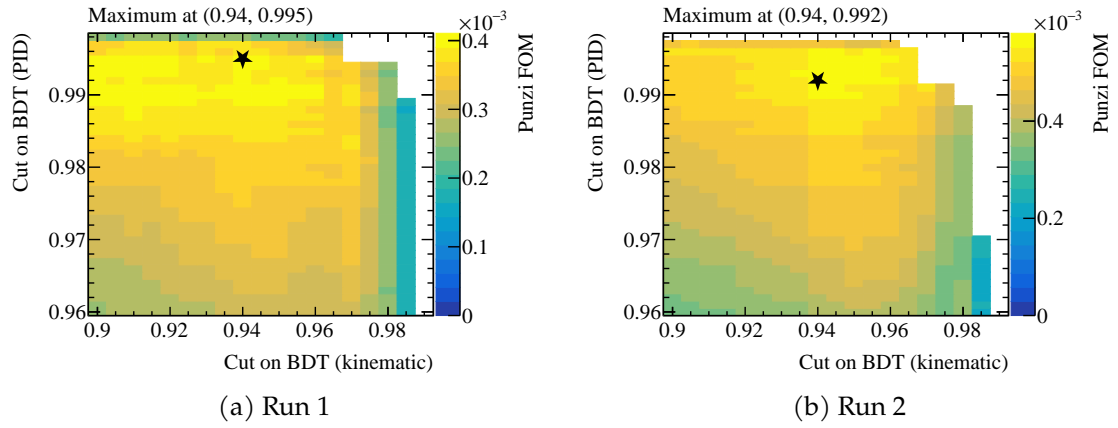


Figure 8.6 – Value of the Punzi FOM for the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ for different kinematic and PID classifier requirements. The distributions' maxima are marked with a star. Empty bins are given if the additional condition of at least two candidates in the upper and lower sideband is not fulfilled.

9 Modelling of invariant mass distributions

This chapter presents the modelling of the invariant mass distributions of the resonant $B_{(s)}^0 \rightarrow J/\psi p \bar{p}$ and non-resonant $B_{(s)}^0 \rightarrow p \bar{p} \mu^+ \mu^-$ decay channels. The numbers of signal candidates in the resonant channel are extracted from a fit to data and used in Section 10.1 to calculate the normalisation constants. Consequently, the fits to the non-resonant simulation and blinded non-resonant data are used to calculate expected upper limits on the $B_s^0 \rightarrow p \bar{p} \mu^+ \mu^-$ and $B^0 \rightarrow p \bar{p} \mu^+ \mu^-$ branching fractions in Section 10.2. All fits are performed as unbinned maximum-likelihood fits with the data sets having all selection requirements of Chapter 8 applied.

9.1 Model for $B_{(s)}^0 \rightarrow J/\psi p \bar{p}$ decays

The fit to the resonant $B_{(s)}^0 \rightarrow J/\psi p \bar{p}$ channel used to determine the number of normalisation channel candidates is performed with a mass constraint on the J/ψ meson. By adding the constraint, the kinematics of the final state particles are more restricted, improving the resolution of the reconstructed $J/\psi p \bar{p}$ mass. In the case of $B_s^0 \rightarrow J/\psi p \bar{p}$ and $B^0 \rightarrow J/\psi p \bar{p}$ decays, this eases the separation of signal and background, especially the clean separation of the B_s^0 and B^0 decay contributions in data. In Fig. 9.1, the $J/\psi p \bar{p}$ mass with and without the constraint is shown for $B_s^0 \rightarrow J/\psi p \bar{p}$ simulation, demonstrating the higher resolution.

Simulated samples of $B_s^0 \rightarrow J/\psi p \bar{p}$ and $B^0 \rightarrow J/\psi p \bar{p}$ decays are used to study the shapes of the signal modes. As for the control channels in Section 7.2, each signal distribution is modelled by a two-sided Hypatia function. The complexity of the Hypatia function is reduced by fixing $\zeta = \beta = 0$ in the function definition. A second-order Chebychev polynomial with coefficients a_0 and a_1 describes the background. The signal model is fitted to the $B_s^0 \rightarrow J/\psi p \bar{p}$ and $B^0 \rightarrow J/\psi p \bar{p}$ simulation samples to constrain the shapes of the distributions for the fit to data. In Fig. 9.2a and Fig. 9.2b, the fit results are shown. For the fit to data, the signal shape tail parameters α_L , α_R , n_L , n_R , and the parameter λ of the Hypatia function are constrained to their fitted values on simulation. The fit to the combined Run 1 and Run 2 data sets is shown in Fig. 9.2c, and the resulting parameter values are listed in Table 9.1.

The stability of the fit procedure is tested with 10 000 pseudo-experiments by repeatedly generating and fitting data sets according to the given fit model. This procedure is done with the fit model described above and the fit values

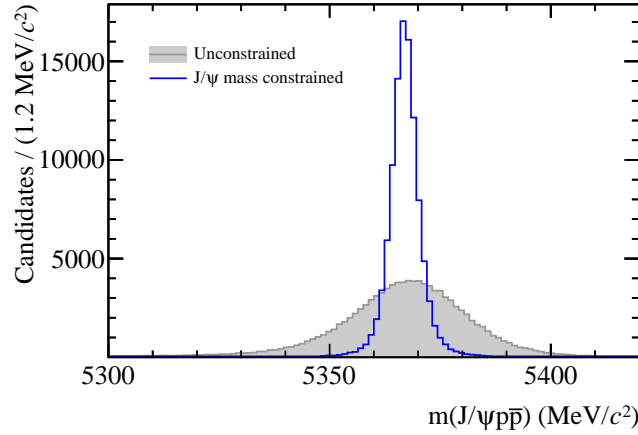


Figure 9.1 – Reconstructed $J/\psi p\bar{p}$ mass in $B_s^0 \rightarrow J/\psi p\bar{p}$ simulation with and without a mass constraint on the J/ψ meson. The 2018 simulation sample is shown. The higher resolution is evident and allows disentangling the B_s^0 and B^0 components in the fit to data.

in Table 9.1 as inputs. The pseudo-experiments are evaluated by examining the uncertainty-weighted residuals, the pulls, between the generated and fitted parameter values. In the case of an unbiased fit with a correct estimation of the fit uncertainties, the pulls are expected to be Gaussian distributed with mean zero and unit width. In Fig. 9.3, the corresponding distributions for the signal and background yield parameters are shown. No significant bias is observed, demonstrating the robustness of the fit.

When evaluating the resulting fit parameters, it is found that the signal resolution in data is about 10 % larger than in simulation. This result is important when discussing the fits to the non-resonant channel where the number of candidates is too small to leave the width parameter freely floating. As a cross-check, the fit to data is repeated with the signal widths fixed to be 10 % larger than in the corresponding fit to simulation. No difference in the number of extracted signal candidates is found, which confirms the result.

As an additional test of the fit model, it is checked that the fit results are compatible when no J/ψ mass constraint is applied. The check is performed to anticipate the behaviour of the fit model when applied to the still blinded non-resonant $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ channel, where no J/ψ meson is present. The fit model is the same one used for the fit with the constrained J/ψ mass, but the ranges of the fit parameters are adjusted to account for the broader widths of the unconstrained mass peaks. Also, the fit procedure is the same as for the constrained fit. First, the signal component of the model is fitted to simulation to constrain the tail parameters and the parameter λ by Gaussian functions. Then, the distribution in data is fitted with the full model, which is shown in Fig. 9.4. In Table 9.2, the fitted B_s^0 and B^0 yields are compared between the J/ψ mass constrained fit and the unconstrained fit. No significant deviation is seen,

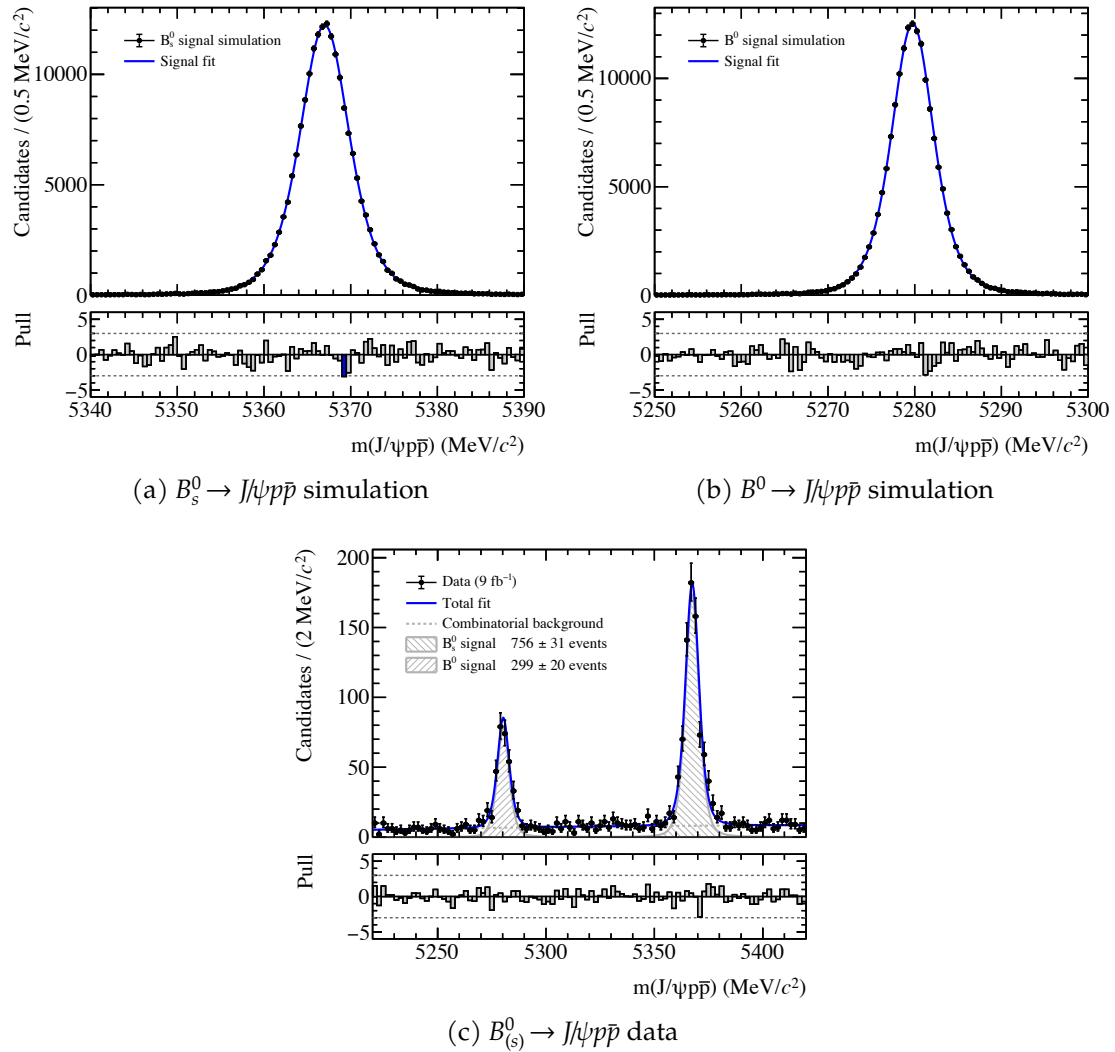


Figure 9.2 – Reconstructed $J/\psi p \bar{p}$ mass distributions of the resonant normalisation channel with a mass constraint on the J/ψ meson. Unbinned maximum-likelihood fits to the (a) $B_s^0 \rightarrow J/\psi p \bar{p}$ simulation, (b) $B^0 \rightarrow J/\psi p \bar{p}$ simulation, and (c) combined Run 1 and Run 2 data are shown.

Table 9.1 – Results of the unbinned maximum-likelihood fits to $B_{(s)}^0 \rightarrow J/\psi p \bar{p}$ simulation and data. In the fit to data, the tail parameters ($\alpha_L, n_L, \alpha_R, n_R$) and the parameter λ are not freely floating but constrained from the fits to simulation. The full 9 fb^{-1} data set is fitted.

		Fit to simulation		Fit to data	
		$B_s^0 \rightarrow J/\psi p \bar{p}$	$B^0 \rightarrow J/\psi p \bar{p}$	$B_s^0 \rightarrow J/\psi p \bar{p}$	$B^0 \rightarrow J/\psi p \bar{p}$
Signal	x_0	5366.99 ± 0.01	5279.77 ± 0.01	5367.47 ± 0.15	5280.30 ± 0.22
	σ_0	6.39 ± 0.09	5.39 ± 0.07	7.02 ± 0.34	6.02 ± 0.45
	λ	-2.33 ± 0.05	-2.26 ± 0.05	-2.32 ± 0.05	-2.26 ± 0.05
	α_L	1.68 ± 0.07	1.63 ± 0.05	1.68 ± 0.07	1.63 ± 0.05
	α_R	1.35 ± 0.05	1.34 ± 0.04	1.35 ± 0.05	1.34 ± 0.04
	n_L	1.93 ± 0.06	2.19 ± 0.08	1.93 ± 0.06	2.19 ± 0.08
	n_R	2.52 ± 0.09	2.57 ± 0.06	2.52 ± 0.09	2.57 ± 0.06
	ζ	0.0	0.0	0.0	0.0
	β	0.0	0.0	0.0	0.0
	N_{sig}			756 ± 31	299 ± 20
	Background	a_0			0.23 ± 0.07
a_1				-0.04 ± 0.07	
N_{bkg}				724 ± 33	

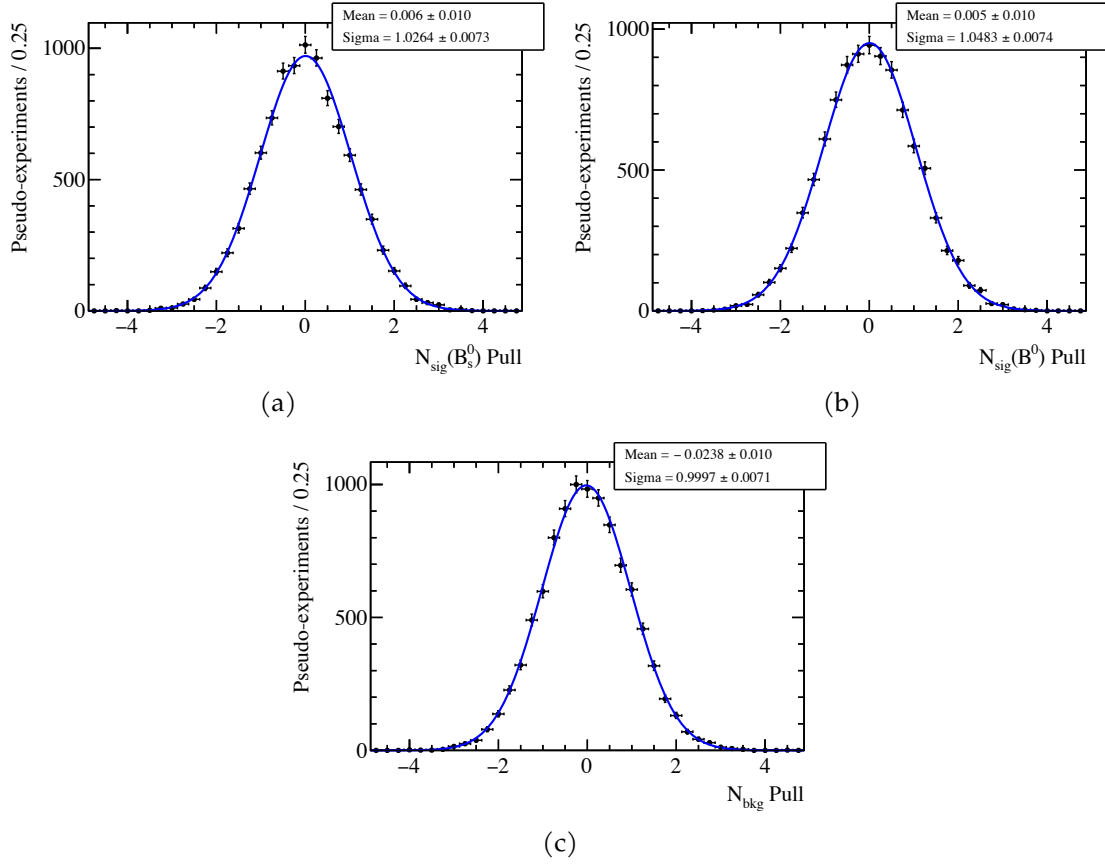


Figure 9.3 – Pull distributions of the signal and background yield parameters extracted from repeated fits to the $B_{(s)}^0 \rightarrow J/\psi p \bar{p}$ channel.

demonstrating the consistency of both fits. Again, the signal resolutions are compared between the simulation and the data. It is found that the B^0 signal resolution in data is about 18 % larger than in simulation, while for the B_s^0 signal resolution, it is 20 %. Repeating the fit with the signal widths fixed to the up-scaled values gives a compatible amount of signal events. It can be assumed that also in the non-resonant channel, the signal resolution is larger in data than in simulation. Therefore, it is planned to consider the relative deviation seen in the resonant channel for the fit in the non-resonant channel.

Table 9.2 – Numbers of signal yields from fits with and without the J/ψ mass constraint. For both modes, the numbers are in agreement within their statistical uncertainty.

	N_{sig} (constrained)	N_{sig} (unconstrained)
$B_s^0 \rightarrow J/\psi p \bar{p}$	756 ± 31	755 ± 40
$B^0 \rightarrow J/\psi p \bar{p}$	299 ± 20	300 ± 29

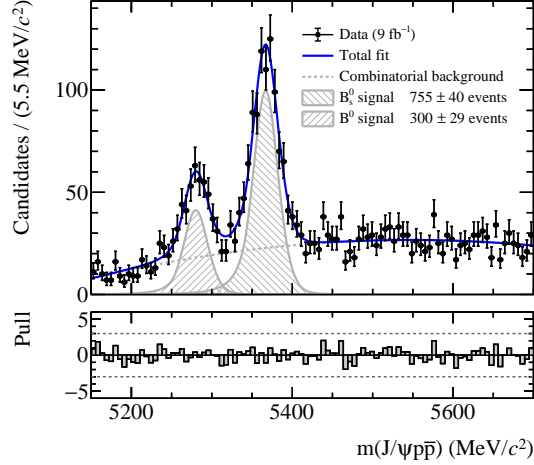


Figure 9.4 – Reconstructed $J/\psi p\bar{p}$ invariant mass distribution in the resonant $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ normalisation channel. A fit to $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ data is shown. No mass constraint on the reconstructed dimuon pair mass is applied.

9.2 Model for $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays

The signal region of the non-resonant $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ channel is still blinded on data. So, in this section, only the signal models are described, together with a background-only fit to the sideband data.

As for the resonant modes, the B_s^0 and B^0 signals are modelled with a two-sided Hypatia function each. The background component is described by an exponential function with parameter c_0 . No kinematic threshold near the B masses exists in the $p\bar{p}\mu^+\mu^-$ mass spectrum as seen in the $J/\psi p\bar{p}$ spectrum, where the minimum mass is given by the rest masses of the J/ψ meson and the two protons. In Fig. 9.5a and Fig. 9.5b, the fits to the simulated samples are shown. For the fit to blinded data, only the exponential background component is used. This fit can be seen in Fig. 9.5c, while the corresponding fit parameters are summarised in Table 9.3. The number of background events in the entire fit window N_{bkg} is given, together with the expected number of events in the blinded signal region N_{bkg}^* calculated by integrating the fitted model in the corresponding region. Next, these fit results are used to calculate expected upper limits on the branching fractions of the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decay modes.

Table 9.3 – Results of the unbinned maximum-likelihood fits to $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ simulation and blinded data. The number of background events in the full fit window N_{bkg} is given. Also, the expected number of events in the blinded signal region N_{bkg}^* is calculated by integrating the fitted model.

		Fit to simulation		Fit to data
		$B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$	$B^0 \rightarrow p\bar{p}\mu^+\mu^-$	
Signal	x_0	5367.94 ± 0.06	5280.71 ± 0.07	
	σ_0	42.94 ± 2.69	41.27 ± 2.90	
	λ	-5.77 ± 0.67	-5.63 ± 0.75	
	α_L	0.62 ± 0.04	0.63 ± 0.06	
	α_R	0.89 ± 0.09	0.86 ± 0.10	
	n_L	1.48 ± 0.04	1.49 ± 0.06	
	n_R	2.37 ± 0.17	2.73 ± 0.21	
	ζ	0.0	0.0	
	β	0.0	0.0	
	Background	c_0		
N_{bkg}^*				11
N_{bkg}				35 ± 7

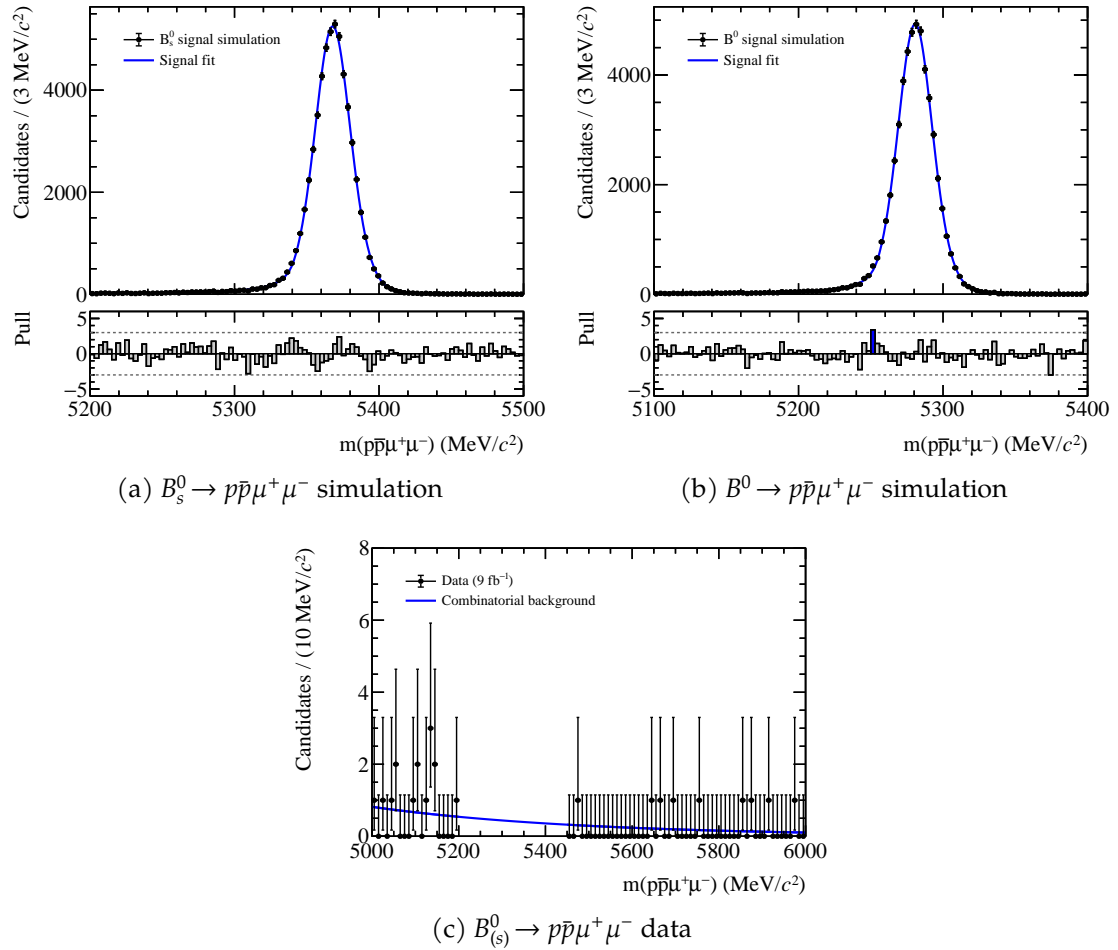


Figure 9.5 – Reconstructed $p\bar{p}\mu^+\mu^-$ invariant mass distributions of the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ signal channel. Unbinned maximum-likelihood fits to (a) $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ simulation, (b) $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ simulation, and (c) combined Run 1 and Run 2 blinded data are shown.

10 Expected upper limits

The presented measurement is ongoing and the signal region of the non-resonant modes is still blinded. Therefore, in this chapter, preliminary expected upper limits on the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ branching fractions are computed. The normalisation procedure for the branching fractions is derived in Section 10.1 followed by the presentation and discussion of the resulting expected upper limits in Section 10.2.

10.1 Normalisation with $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ decays

This section presents the normalisation of the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ signal decay modes with the corresponding resonant decay modes. The signal branching fractions are calculated according to

$$\mathcal{B}(B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-) = \underbrace{\frac{\varepsilon(B_{(s)}^0 \rightarrow J/\psi p\bar{p})}{\varepsilon(B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-)} \times \frac{\mathcal{B}(B_{(s)}^0 \rightarrow J/\psi p\bar{p})}{N(B_{(s)}^0 \rightarrow J/\psi p\bar{p})}}_{\alpha_{(s)}} \times N(B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-). \quad (10.1)$$

In this equation, α_s and α are the normalisation factors for the B_s^0 and B^0 signal decays, respectively. The input N is the measured number of candidates of the corresponding mode, and ε is the total detection efficiency. For the normalisation channel, the process $J/\psi \rightarrow \mu^+\mu^-$ is always considered.

Equation (10.1) is used because the parts of the detection efficiencies and their corresponding uncertainties that are common to the signal and normalisation channel cancel in the ratio. This cancellation motivates the choice of a normalisation channel as similar as possible to the signal channel, with common trigger, reconstruction, and selection procedures. The decays $B_s^0 \rightarrow J/\psi p\bar{p}$ and $B^0 \rightarrow J/\psi p\bar{p}$ are chosen because of their identical final state and the expected similar trigger and particle identification system response. Also, the reconstruction and other selection efficiencies are expected to behave similarly to the signal due to the same final state. In addition, the normalisation channel B hadrons already have the same flavour as the corresponding signal channel B hadrons avoiding the external input f_s/f_d as used in the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ normalisation (see Section 6.1). The total detection efficiencies are split according to the detection stages: geometrical detector acceptance, reconstruction and centralised preselection, and further selection requirements (including the trigger selection and the requirements on

the two multivariate classifier outputs), such that

$$\frac{\varepsilon(B_{(s)}^0 \rightarrow J/\psi p \bar{p})}{\varepsilon(B_{(s)}^0 \rightarrow p \bar{p} \mu^+ \mu^-)} = \frac{\varepsilon_{\text{norm}}^{\text{Total}}}{\varepsilon_{\text{sig}}^{\text{Total}}} = \frac{\varepsilon_{\text{norm}}^{\text{Acc}}}{\varepsilon_{\text{sig}}^{\text{Acc}}} \times \frac{\varepsilon_{\text{norm}}^{\text{RecPre|Acc}}}{\varepsilon_{\text{sig}}^{\text{RecPre|Acc}}} \times \frac{\varepsilon_{\text{norm}}^{\text{Sel|AccRecPre}}}{\varepsilon_{\text{sig}}^{\text{Sel|AccRecPre}}}. \quad (10.2)$$

Here, the efficiency for each stage is computed from candidates that pass the preceding selection stages. The branching fractions of the resonant decays that enter into the normalisation computation are taken from the published LHCb measurement [75] and considered as external inputs:

$$\mathcal{B}(B_s^0 \rightarrow J/\psi p \bar{p}) = (3.58 \pm 0.19 \text{ (stat)} \pm 0.39 \text{ (syst)}) \times 10^{-6} \quad (10.3)$$

$$\mathcal{B}(B^0 \rightarrow J/\psi p \bar{p}) = (4.51 \pm 0.40 \text{ (stat)} \pm 0.44 \text{ (syst)}) \times 10^{-7}. \quad (10.4)$$

The calculation and study of the normalisation factor components are described in the following sections.

10.1.1 Number of normalisation channel events

The normalisation channel candidates have similar selection requirements as the signal channel candidates (see Chapter 8). Differences are the selected q^2 regions and the requirements on the parallel multivariate classifier outputs. In Chapter 9, the extracted number of signal events can be found in Table 9.1.

10.1.2 Selection efficiencies

As part of the detection efficiency, the detector acceptance is defined, in first approximation, as the fraction of generated decays having all the decay products in the geometrical detector acceptance. The total detector acceptance, however, will also be affected by the magnetic field of the LHCb magnet and by the particle interactions with the detector material. These effects are evaluated as part of the reconstruction efficiency. In Table 10.1, the geometrical detector acceptances for the signal and normalisation modes are listed, which are estimated with simulated samples. The decay products of the studied decays are required to move in the direction of the LHCb detector acceptance, defined as the polar angle in the range of [10, 400] mrad. This acceptance is chosen to be larger than the physical LHCb detector acceptance to allow for the recovery of particles deflected into the acceptance by the magnetic field. As expected from the same four-body topology of the decays, the geometrical detector acceptances per year are similar for the signal and normalisation modes.

The reconstruction efficiency is the fraction of decay candidates in the detector acceptance successfully reconstructed by the experiment's software. The selection efficiency is the fraction of reconstructed decay candidates that pass the selection. Both efficiencies depend on the characteristics of the decay channel, like *e.g.* the number of particles in the final state, their kinematic distributions,

the tracking efficiency, and the particle identification efficiency. The combined efficiencies of the reconstruction and centralised preselection are evaluated on truth-matched simulation samples and given in Table 10.1. For each channel, the efficiencies are in overall good agreement between the years. Slight deviations, especially for the years 2011, 2012, and 2015, are due to different detector performances and trigger configurations between Run 1 and Run 2. Because only the ratio of signal and normalisation channel efficiencies enters the final calculation of the normalisation constants, variations in the efficiencies cancel out.

The efficiency of the subsequent selection steps, containing the trigger and BDT selection, is defined on simulated samples as the ratio between the number of reconstructed and fully-selected candidates and the number of generated, reconstructed, preselected, and truth-matched events. Simulation corrections are applied when computing the selection efficiencies (see Section 7.2). The uncertainties are derived exclusively from the statistical size of the simulation sample. Systematic uncertainties resulting from the various simulation corrections are not yet included. The calculated efficiencies are listed in Table 10.1, while a collection of the different partial efficiencies is shown in Table 10.2. As expected, the efficiencies differ between the resonant and non-resonant modes due to the different requirements on the BDT classifier outputs and the q^2 region. No significant deviations are observed when comparing the efficiencies of the B_s^0 modes with the ones of the corresponding B^0 modes.

10.1.3 Normalisation factors

The efficiencies used in the computation of the normalisation factors are the total detection efficiencies, calculated as the product of the geometrical detector acceptance, the reconstruction efficiency, and the efficiency of all selection requirements. In Table 10.1, the total efficiencies are summarised per data-taking year and for the combined Run 1+2 data set. The combined efficiencies are derived by weighting and averaging the single-year efficiencies according to the luminosity fractions in data. Finally, the normalisation factors α_s and α are calculated from Eq. (10.1) and yield

$$\alpha_s = (8.4 \pm 1.1) \times 10^{-10} \quad (10.5)$$

$$\alpha = (2.6 \pm 0.4) \times 10^{-10} , \quad (10.6)$$

where the reported uncertainty is the sum in quadrature over the statistical uncertainty and the uncertainty introduced by the external input of the normalisation channel branching fractions.

10 Expected upper limits

Table 10.1 – Total efficiencies for the $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ signal and $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ normalisation modes calculated on simulated samples. The simulation corrections derived in Section 7.2 are applied. So far, the uncertainties are derived only from the statistical size of the simulation samples. In addition to the numbers for the combined Run 1+2 data set, the efficiencies for the different years are stated.

Channel	ϵ^{Acc}	$\epsilon^{\text{RecPre Acc}}$	$\epsilon^{\text{Sel AccRecPre}}$	ϵ^{Total}
$B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$				
2011	$(16.76 \pm 0.03)\%$	$(10.96 \pm 0.04)\%$	$(7.17 \pm 0.11)\%$	$(0.132 \pm 0.002)\%$
2012	$(17.04 \pm 0.02)\%$	$(9.98 \pm 0.03)\%$	$(6.39 \pm 0.07)\%$	$(0.109 \pm 0.001)\%$
2015	$(18.03 \pm 0.04)\%$	$(11.44 \pm 0.05)\%$	$(6.16 \pm 0.10)\%$	$(0.127 \pm 0.002)\%$
2016	$(18.07 \pm 0.04)\%$	$(11.75 \pm 0.04)\%$	$(8.00 \pm 0.08)\%$	$(0.170 \pm 0.002)\%$
2017	$(18.00 \pm 0.04)\%$	$(11.75 \pm 0.04)\%$	$(11.01 \pm 0.11)\%$	$(0.233 \pm 0.002)\%$
2018	$(17.94 \pm 0.04)\%$	$(11.71 \pm 0.03)\%$	$(9.52 \pm 0.09)\%$	$(0.200 \pm 0.002)\%$
Run 1+2	$(17.63 \pm 0.02)\%$	$(11.23 \pm 0.01)\%$	$(8.39 \pm 0.04)\%$	$(0.169 \pm 0.001)\%$
$B^0 \rightarrow p\bar{p}\mu^+\mu^-$				
2011	$(16.83 \pm 0.03)\%$	$(10.59 \pm 0.04)\%$	$(7.29 \pm 0.11)\%$	$(0.130 \pm 0.002)\%$
2012	$(17.14 \pm 0.02)\%$	$(9.63 \pm 0.03)\%$	$(6.54 \pm 0.07)\%$	$(0.108 \pm 0.001)\%$
2015	$(17.98 \pm 0.04)\%$	$(10.93 \pm 0.05)\%$	$(6.08 \pm 0.10)\%$	$(0.120 \pm 0.002)\%$
2016	$(18.02 \pm 0.04)\%$	$(11.29 \pm 0.03)\%$	$(8.02 \pm 0.08)\%$	$(0.163 \pm 0.002)\%$
2017	$(17.99 \pm 0.04)\%$	$(11.28 \pm 0.04)\%$	$(11.15 \pm 0.11)\%$	$(0.226 \pm 0.002)\%$
2018	$(18.01 \pm 0.04)\%$	$(11.23 \pm 0.03)\%$	$(9.45 \pm 0.09)\%$	$(0.191 \pm 0.002)\%$
Run 1+2	$(17.66 \pm 0.02)\%$	$(10.80 \pm 0.01)\%$	$(8.45 \pm 0.04)\%$	$(0.163 \pm 0.001)\%$
$B_s^0 \rightarrow J/\psi p\bar{p}$				
2011	$(16.72 \pm 0.03)\%$	$(12.15 \pm 0.04)\%$	$(17.43 \pm 0.14)\%$	$(0.354 \pm 0.003)\%$
2012	$(17.04 \pm 0.02)\%$	$(11.11 \pm 0.03)\%$	$(17.59 \pm 0.11)\%$	$(0.333 \pm 0.002)\%$
2015	$(17.96 \pm 0.04)\%$	$(11.62 \pm 0.05)\%$	$(22.43 \pm 0.17)\%$	$(0.468 \pm 0.004)\%$
2016	$(17.93 \pm 0.04)\%$	$(11.98 \pm 0.03)\%$	$(25.29 \pm 0.13)\%$	$(0.544 \pm 0.003)\%$
2017	$(17.89 \pm 0.04)\%$	$(11.97 \pm 0.03)\%$	$(30.00 \pm 0.15)\%$	$(0.643 \pm 0.004)\%$
2018	$(17.94 \pm 0.04)\%$	$(11.89 \pm 0.03)\%$	$(28.01 \pm 0.12)\%$	$(0.597 \pm 0.003)\%$
Run 1+2	$(17.58 \pm 0.02)\%$	$(11.76 \pm 0.01)\%$	$(24.01 \pm 0.06)\%$	$(0.501 \pm 0.001)\%$
$B^0 \rightarrow J/\psi p\bar{p}$				
2011	$(16.72 \pm 0.03)\%$	$(11.31 \pm 0.04)\%$	$(17.54 \pm 0.14)\%$	$(0.332 \pm 0.003)\%$
2012	$(17.04 \pm 0.02)\%$	$(10.22 \pm 0.03)\%$	$(18.07 \pm 0.12)\%$	$(0.315 \pm 0.002)\%$
2015	$(17.93 \pm 0.04)\%$	$(10.62 \pm 0.04)\%$	$(22.91 \pm 0.16)\%$	$(0.436 \pm 0.004)\%$
2016	$(17.85 \pm 0.04)\%$	$(10.91 \pm 0.03)\%$	$(25.94 \pm 0.13)\%$	$(0.505 \pm 0.003)\%$
2017	$(17.93 \pm 0.04)\%$	$(10.90 \pm 0.03)\%$	$(30.67 \pm 0.16)\%$	$(0.599 \pm 0.004)\%$
2018	$(17.99 \pm 0.04)\%$	$(10.94 \pm 0.03)\%$	$(28.99 \pm 0.13)\%$	$(0.570 \pm 0.003)\%$
Run 1+2	$(17.58 \pm 0.02)\%$	$(10.79 \pm 0.01)\%$	$(24.63 \pm 0.06)\%$	$(0.471 \pm 0.001)\%$

Table 10.2 – Partial efficiencies for the signal and normalisation channel candidate selection, evaluated on simulated samples. So far, the uncertainties are derived only from the statistical size of the simulation samples. The different selection steps are described in Section 8. The numbers are provided only for information and are not used in any computations.

Efficiency	$B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$	$B^0 \rightarrow p\bar{p}\mu^+\mu^-$	$B_s^0 \rightarrow J/\psi p\bar{p}$	$B^0 \rightarrow J/\psi p\bar{p}$	
Common	2011	$(58.04 \pm 0.21)\%$	$(57.33 \pm 0.21)\%$	$(79.57 \pm 0.15)\%$	$(78.27 \pm 0.15)\%$
	2012	$(56.58 \pm 0.15)\%$	$(56.23 \pm 0.15)\%$	$(79.92 \pm 0.12)\%$	$(78.56 \pm 0.13)\%$
	2015	$(95.35 \pm 0.09)\%$	$(95.14 \pm 0.09)\%$	$(92.10 \pm 0.11)\%$	$(90.92 \pm 0.11)\%$
	2016	$(95.13 \pm 0.07)\%$	$(95.06 \pm 0.07)\%$	$(92.01 \pm 0.08)\%$	$(90.89 \pm 0.09)\%$
	2017	$(95.39 \pm 0.07)\%$	$(95.22 \pm 0.08)\%$	$(92.83 \pm 0.08)\%$	$(91.65 \pm 0.10)\%$
	2018	$(95.31 \pm 0.06)\%$	$(94.96 \pm 0.07)\%$	$(92.71 \pm 0.07)\%$	$(91.40 \pm 0.08)\%$
q^2 range	2011	$(98.96 \pm 0.06)\%$	$(99.26 \pm 0.05)\%$	$(97.24 \pm 0.07)\%$	$(97.54 \pm 0.06)\%$
	2012	$(98.99 \pm 0.04)\%$	$(99.23 \pm 0.04)\%$	$(97.14 \pm 0.05)\%$	$(97.65 \pm 0.05)\%$
	2015	$(99.45 \pm 0.03)\%$	$(99.56 \pm 0.03)\%$	$(97.17 \pm 0.07)\%$	$(97.49 \pm 0.06)\%$
	2016	$(99.39 \pm 0.02)\%$	$(99.58 \pm 0.02)\%$	$(97.07 \pm 0.05)\%$	$(97.50 \pm 0.05)\%$
	2017	$(99.23 \pm 0.03)\%$	$(99.51 \pm 0.03)\%$	$(97.28 \pm 0.05)\%$	$(97.56 \pm 0.05)\%$
	2018	$(99.29 \pm 0.03)\%$	$(99.55 \pm 0.02)\%$	$(97.07 \pm 0.05)\%$	$(97.58 \pm 0.05)\%$
Trigger	2011	$(66.84 \pm 0.26)\%$	$(66.87 \pm 0.27)\%$	$(75.45 \pm 0.18)\%$	$(76.53 \pm 0.18)\%$
	2012	$(61.56 \pm 0.20)\%$	$(61.89 \pm 0.20)\%$	$(77.52 \pm 0.14)\%$	$(79.00 \pm 0.14)\%$
	2015	$(35.41 \pm 0.21)\%$	$(35.12 \pm 0.21)\%$	$(69.18 \pm 0.20)\%$	$(70.47 \pm 0.19)\%$
	2016	$(46.43 \pm 0.16)\%$	$(46.22 \pm 0.16)\%$	$(78.02 \pm 0.13)\%$	$(78.72 \pm 0.13)\%$
	2017	$(61.84 \pm 0.18)\%$	$(61.34 \pm 0.18)\%$	$(82.55 \pm 0.13)\%$	$(83.18 \pm 0.14)\%$
	2018	$(53.74 \pm 0.16)\%$	$(53.73 \pm 0.16)\%$	$(78.04 \pm 0.12)\%$	$(78.99 \pm 0.13)\%$
BDT (kinematic)	2011	$(51.44 \pm 0.34)\%$	$(52.63 \pm 0.35)\%$	$(64.29 \pm 0.23)\%$	$(65.53 \pm 0.23)\%$
	2012	$(49.88 \pm 0.26)\%$	$(50.83 \pm 0.26)\%$	$(60.45 \pm 0.18)\%$	$(62.22 \pm 0.19)\%$
	2015	$(50.48 \pm 0.37)\%$	$(51.00 \pm 0.37)\%$	$(91.27 \pm 0.15)\%$	$(91.91 \pm 0.13)\%$
	2016	$(49.02 \pm 0.23)\%$	$(48.85 \pm 0.23)\%$	$(90.94 \pm 0.10)\%$	$(91.60 \pm 0.10)\%$
	2017	$(48.67 \pm 0.23)\%$	$(48.84 \pm 0.23)\%$	$(92.40 \pm 0.10)\%$	$(93.07 \pm 0.10)\%$
	2018	$(48.33 \pm 0.21)\%$	$(48.25 \pm 0.21)\%$	$(92.33 \pm 0.09)\%$	$(92.91 \pm 0.09)\%$
BDT (PID)	2011	$(36.31 \pm 0.46)\%$	$(36.41 \pm 0.46)\%$	$(46.43 \pm 0.30)\%$	$(45.80 \pm 0.30)\%$
	2012	$(37.13 \pm 0.35)\%$	$(37.25 \pm 0.35)\%$	$(48.36 \pm 0.24)\%$	$(47.92 \pm 0.25)\%$
	2015	$(36.35 \pm 0.50)\%$	$(35.86 \pm 0.49)\%$	$(39.69 \pm 0.27)\%$	$(39.91 \pm 0.25)\%$
	2016	$(37.16 \pm 0.32)\%$	$(37.51 \pm 0.32)\%$	$(39.91 \pm 0.18)\%$	$(40.60 \pm 0.18)\%$
	2017	$(38.65 \pm 0.32)\%$	$(39.28 \pm 0.33)\%$	$(43.56 \pm 0.20)\%$	$(44.31 \pm 0.20)\%$
	2018	$(38.71 \pm 0.30)\%$	$(38.55 \pm 0.30)\%$	$(43.20 \pm 0.17)\%$	$(44.28 \pm 0.18)\%$

10.2 Results and interpretation

The analysis is still ongoing and the signal mass window of the non-resonant channel is not yet examined. For this thesis, expected upper limits on the branching fractions are calculated based on the normalisation constants and the modelling of the signal and background contributions described in Chapter 9. The upper limits are calculated using the CL_s method [156] as implemented in Ref. [157]. With the same approach, the upper limit on the $B^0 \rightarrow \mu^+ \mu^-$ mode is calculated (see Section 6.3). It is based on a profile likelihood ratio as a one-sided test statistic, defined as

$$q_\mu = \begin{cases} -2 \log \left(\frac{\mathcal{L}(\mu|x)}{\mathcal{L}(\tilde{\mu}|x)} \right) & \text{if } \mu \geq \tilde{\mu} \\ 0 & \text{if } \mu < \tilde{\mu} . \end{cases} \quad (10.7)$$

In this function, \mathcal{L} is the likelihood function of the fit, μ is the scanned parameter of interest (the branching fraction) with its best-fitted value $\tilde{\mu}$, and x is the observable (the reconstructed $p\bar{p}\mu^+\mu^-$ mass). With the test statistic, the p value can be expressed in the Wald approximation [165] representing the compatibility of a tested hypothesis with the data:

$$p_\mu = \int_{q_{\text{obs}}}^{\infty} \left(\Phi \left(\frac{\mu' - \mu}{\sigma} \right) \delta(q_\mu) + \frac{1}{\sqrt{8\pi}q_\mu} \exp \left[-\frac{1}{2} \left(\sqrt{q_\mu} - \frac{\mu - \mu'}{\sigma} \right)^2 \right] \right) dq_\mu . \quad (10.8)$$

Here, μ denotes the parameter of interest of the tested hypothesis with standard deviation σ and its value for a reference hypothesis μ' .

Finally, the p value for the signal-plus-background fit model is normalised to the p value for the background-only model to measure the limit based on the signal rate alone:

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b} . \quad (10.9)$$

For the signal plus background hypothesis, $\mu' = \mu$ is assumed, while it is $\mu' = 0$ for the background-only hypothesis. The confidence intervals are measured by performing 50 000 pseudo-experiments, lowering the dependence on the Wald approximation. When evaluating the expected upper limit on the B_s^0 branching fraction, the B^0 branching fraction is assumed to be zero and vice versa. The effect of Gaussian-constrained parameters in the fit, like *e.g.* the normalisation constant, is considered by randomly generating values for these parameters within their constraints and plugging them in at their best-fitted value at each scanned value of the examined branching fraction.

The CL_s value distributions for the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ branching fractions are shown in Fig. 10.1. The expected upper limits are found to be

$$\mathcal{B}(B_s^0 \rightarrow p\bar{p}\mu^+\mu^-) < 4.3 \times 10^{-9} \quad \text{at 95 \% CL} \quad (10.10)$$

$$\mathcal{B}(B^0 \rightarrow p\bar{p}\mu^+\mu^-) < 1.4 \times 10^{-9} \quad \text{at 95 \% CL} . \quad (10.11)$$

These limits are the first ones calculated based on the complete Run 1 and Run 2 data set of the LHCb experiment. Given the ongoing analysis, these expected upper limits are preliminary since not all systematic uncertainties of the analysis are considered. In particular, this includes an uncertainty caused by the inadequate decay model in simulation. On the other hand, given the few candidates observed in the resonant channel, it is expected that the statistical uncertainty, which is associated with the decay rates, will predominate in the measurement's uncertainty. If this is true, it can be assumed that the expected upper limits calculated in this thesis are of the correct order of magnitude.

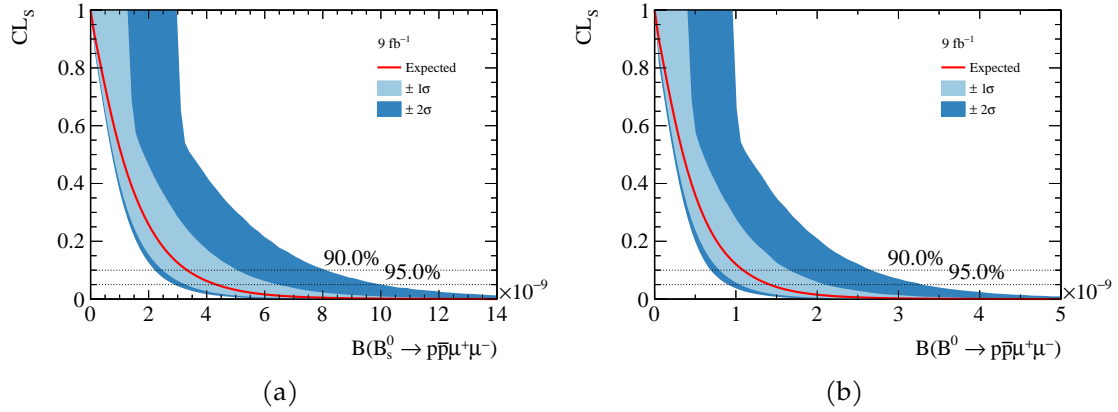


Figure 10.1 – Result of the CL_s scan performed to extract an expected upper limit on the (a) $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and (b) $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ branching fraction under the background-only hypothesis from pseudo-experiments. Colour bands indicate the 1σ and 2σ regions around the expected values. The 95% confidence level is marked as a dotted line.

11 Conclusion

The Standard Model of particle physics is the most comprehensive and precise theory of fundamental particles and their interactions. Nevertheless, despite its great success, it leaves important questions unanswered that motivate the search for new physics contributions in indirect measurements.

This thesis presents two measurements probing the Standard Model with decay modes that include $b \rightarrow s\ell^+\ell^-$ quark transitions. Both measurements exploit large proton-proton collision samples recorded by the LHCb experiment at centre-of-mass energies of 7, 8, and 13 TeV. The data size corresponds to an integrated luminosity of 9 fb^{-1} .

In the first measurement, the branching fractions of $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ decays are determined, which have precise theoretical estimates and can be accessed due to LHCb's optimised detector design. The branching fractions are measured to be

$$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) = \left(3.09_{-0.44}^{+0.46}(\text{stat.})_{-0.10}^{+0.14}(\text{syst.})\right) \times 10^{-9} \quad (11.1)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-) = \left(1.20_{-0.74}^{+0.83}(\text{stat.})_{-0.15}^{+0.14}(\text{syst.})\right) \times 10^{-10} . \quad (11.2)$$

The B^0 mode is detected with a significance of 1.7σ above the background-only hypothesis. Because the decay is not observed, an upper limit on the branching fraction is determined to be

$$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-) < 2.6 \times 10^{-10} \quad \text{at 95\% CL} . \quad (11.3)$$

An upper limit on the ratio of the B_s^0 and B^0 modes is evaluated to be

$$\mathcal{R}_{\mu^+\mu^-} < 0.095 \quad \text{at 95\% CL} . \quad (11.4)$$

The results are presented in two journal articles [16, 17] and have been the most precise single-experiment measurements of these modes by the time of publication.

The second measurement presented in this thesis is a first search for the semi-baryonic $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays. In addition to studying $b \rightarrow s\ell^+\ell^-$ transitions, they give access to possible new mechanisms contributing to the diproton spectrum near its production threshold as indicated by other LHCb analyses. The measurement is in an advanced state. The signal selection has been developed and optimised based on two multivariate classifiers. Simulated samples used to model the signal distributions and estimate

the selection efficiencies were calibrated with respect to their trigger and PID responses, and also the distributions of the B kinematics and event multiplicity were corrected. A calibration of the underlying decay model is still under development. Also, studies on the systematic uncertainties for the measurement are pending. Under these conditions, expected upper limits on the branching fractions are derived from fits to the background in the signal channel to be

$$\mathcal{B}(B_s^0 \rightarrow p\bar{p}\mu^+\mu^-) < 4.3 \times 10^{-9} \quad \text{at 95 \% CL} \quad (11.5)$$

$$\mathcal{B}(B^0 \rightarrow p\bar{p}\mu^+\mu^-) < 1.4 \times 10^{-9} \quad \text{at 95 \% CL} . \quad (11.6)$$

The sensitivities of both measurements presented in this thesis are currently limited by the available size of the data samples, leaving room for new physics contributions such as scalar and pseudo-scalar particles in $b \rightarrow s\ell^+\ell^-$ transitions.

In 2022, the LHCb experiment resumed data-taking with an upgraded detector in Run 3 of the LHC. Equipped with a new tracking system based on scintillating fibers, a new VELO detector, and a completely redesigned trigger system, this Upgrade I detector and the foreseen Upgrade II detector aim to collect a data sample with a size corresponding to 300 fb^{-1} . With such a large data sample, it will be possible to significantly reduce the uncertainties on the $B_s^0 \rightarrow \mu^+\mu^-$ decay close to the precision of the current theoretical predictions and observe the $B^0 \rightarrow \mu^+\mu^-$ decay for the first time. In case they are not already discovered in the currently ongoing measurement, it is also expected to observe the $B_s^0 \rightarrow p\bar{p}\mu^+\mu^-$ and $B^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays in this data. These prospects for both measurements promise more insights and a deeper understanding of the possible influences of new physics on rare processes in the Standard Model. While the $b \rightarrow s\ell^+\ell^-$ anomalies seen in the current data are still puzzling, it can be expected to solve the issue with the data recorded at the LHC in Run 3 and beyond. Both, the measurement of $B_{(s)}^0 \rightarrow \mu^+\mu^-$ and $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ decays, will be important parts to solve this puzzle.

References

- [1] S. L. Glashow. "Partial-symmetries of weak interactions". Nucl. Phys. 22 (1961) 579.
doi: 10.1016/0029-5582(61)90469-2
- [2] A. Salam and J. C. Ward. "Electromagnetic and weak interactions". Phys. Lett. 13 (1964) 168.
doi: 10.1016/0031-9163(64)90711-5
- [3] S. Weinberg. "A model of leptons". Phys. Rev. Lett. 19 (1967) 1264.
doi: 10.1103/PhysRevLett.19.1264
- [4] F. Halzen and A. D. Martin. "Quarks & Leptons: An Introductory Course in Modern Particle Physics". John Wiley & Sons, 1984.
ISBN: 978-0471887416
- [5] D. Griffiths. "Introduction to elementary particles". Wiley-VCH, 2008.
ISBN: 978-3527406012
- [6] A. J. Buras. "Gauge Theory of Weak Decays: The Standard Model and the Expedition to New Physics Summits". Cambridge University Press, 2020.
doi: 10.1017/9781139524100
- [7] P. W. Higgs. "Broken symmetries, massless particles and gauge fields". Phys. Lett. 12 (1964) 132.
doi: 10.1016/0031-9163(64)91136-9
- [8] P. W. Higgs. "Broken symmetries and the masses of gauge bosons". Phys. Rev. Lett. 13 (1964) 508.
doi: 10.1103/PhysRevLett.13.508
- [9] F. Englert and R. Brout. "Broken symmetry and the mass of gauge vector mesons". Phys. Rev. Lett. 13 (1964) 321.
doi: 10.1103/PhysRevLett.13.321
- [10] G. Aad et al. (ATLAS collaboration). "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.
doi: 10.1016/j.physletb.2012.08.020. ARXIV: 1207.7214
- [11] S. Chatrchyan et al. (CMS collaboration). "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". Phys. Lett. B716 (2012) 30.
doi: 10.1016/j.physletb.2012.08.021. ARXIV: 1207.7235

References

- [12] N. Aghanim et al. (Planck collaboration). “Planck 2018 results: Cosmological parameters”. *Astron. Astrophys.* 641 (2020) A6.
DOI: 10.1051/0004-6361/201833910. ARXIV: 1807.06209
- [13] A. J. Buras. “Flavour expedition to the zeptouniverse”. *Proceedings of Science* 220 (2015) 3 (Flavorful Ways to New Physics).
DOI: 10.22323/1.220.0003. ARXIV: 1505.00618
- [14] A. J. Buras and J. Girrbach. “Towards the identification of new physics through quark flavour violating processes”. *Rep. Prog. Phys.* 77 (2014) 086201.
DOI: 10.1088/0034-4885/77/8/086201. ARXIV: 1306.3775
- [15] W. Altmannshofer and P. Stangl. “New physics in rare B decays after Moriond 2021”. *Eur. Phys. J. C* 81 (2021) 952.
DOI: 10.1140/epjc/s10052-021-09725-1. ARXIV: 2103.13370
- [16] R. Aaij et al. (LHCb collaboration). “Analysis of neutral B -meson decays into two muons”. *Phys. Rev. Lett.* 128 (2022) 041801.
DOI: 10.1103/PhysRevLett.128.041801. ARXIV: 2108.09284
- [17] R. Aaij et al. (LHCb collaboration). “Measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay properties and search for the $B^0 \rightarrow \mu^+ \mu^-$ and $B_s^0 \rightarrow \mu^+ \mu^- \gamma$ decays”. *Phys. Rev. D* 105 (2022) 012010.
DOI: 10.1103/PhysRevD.105.012010. ARXIV: 2108.09283
- [18] M.-Z. Wang et al. (Belle collaboration). “Study of the baryon-antibaryon low-mass enhancements in charmless three-body baryonic B decays”. *Phys. Lett. B* 617 (2005) 141.
DOI: 10.1016/j.physletb.2005.05.008. ARXIV: hep-ex/0503047
- [19] M. Thomson. “Modern Particle Physics”. Cambridge University Press, 2013.
ISBN: 978-1107034266
- [20] Wikimedia Commons. “Standard Model of Elementary Particles”. 2019.
URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg
- [21] R. Aaij et al. (LHCb collaboration). “Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^+ \rightarrow J/\psi\phi K^+$ decays”. *Phys. Rev. Lett.* 118 (2017) 022003.
DOI: 10.1103/PhysRevLett.118.022003. ARXIV: 1606.07895
- [22] R. Aaij et al. (LHCb collaboration). “Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi p K^-$ decays”. *Phys. Rev. Lett.* 115 (2015) 072001.
DOI: 10.1103/PhysRevLett.115.072001. ARXIV: 1507.03414

- [23] R. Aaij et al. (LHCb collaboration). “Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$ ”. *Phys. Rev. Lett.* 122 (2019) 222001.
doi: 10.1103/PhysRevLett.122.222001. ARXIV: 1904.03947
- [24] R. Aaij et al. (LHCb collaboration). “Observation of a $J/\psi\Lambda$ resonance consistent with a strange pentaquark candidate in $B^- \rightarrow J/\psi\Lambda\bar{p}$ decays”. 2022.
ARXIV: 2210.10346. Note: to be published in *Phys. Rev. Lett.*
- [25] N. Cabibbo. “Unitary symmetry and leptonic decays”. *Phys. Rev. Lett.* 10 (1963) 531.
doi: 10.1103/PhysRevLett.10.531
- [26] M. Kobayashi and T. Maskawa. “ CP -violation in the renormalizable theory of weak interaction”. *Progr. Theor. Phys.* 49 (1973) 652.
doi: 10.1143/PTP.49.652
- [27] R. L. Workman et al. (Particle Data group). “Review of Particle Physics”. *Prog. Theor. Exp. Phys.* 2022 (2022) 083C01.
doi: 10.1093/ptep/ptac097
- [28] Y. Fukuda et al. (Super-Kamiokande collaboration). “Evidence for oscillation of atmospheric neutrinos”. *Phys. Rev. Lett.* 81 (1998) 1562.
doi: 10.1103/PhysRevLett.81.1562. ARXIV: hep-ex/9807003
- [29] Q. R. Ahmad et al. (SNO collaboration). “Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by 8B solar neutrinos at the Sudbury Neutrino Observatory”. *Phys. Rev. Lett.* 87 (2001) 071301.
doi: 10.1103/PhysRevLett.87.071301. ARXIV: nucl-ex/0106015
- [30] T. Mannel. “Theory and Phenomenology of CP Violation”. *Nucl. Phys. B Proceedings Supplements* 167 (2007) 170.
doi: 10.1016/j.nuclphysbps.2006.12.083
- [31] R. Aaij et al. (LHCb collaboration). “Measurement of lepton universality parameters in $B^+ \rightarrow K^+\ell^+\ell^-$ and $B^0 \rightarrow K^{*0}\ell^+\ell^-$ decays”. 2022.
ARXIV: 2212.09153. Note: to appear in *Phys. Rev. D*
- [32] R. Aaij et al. (LHCb collaboration). “Test of lepton universality in $b \rightarrow s\ell^+\ell^-$ decays”. 2022.
ARXIV: 2212.09152. Note: to appear in *Phys. Rev. Lett.*
- [33] R. Aaij et al. (LHCb collaboration). “Differential branching fractions and isospin asymmetries of $\rightarrow K^{(*)}\mu^+\mu^-$ decays”. *J. High Energy Phys.* 06 (2014) 133.
doi: 10.1007/JHEP06(2014)133. ARXIV: 1403.8044

- [34] R. Aaij et al. (LHCb collaboration). “Differential branching fraction and angular analysis of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays”. *J. High Energy Phys.* 06 (2015) 115.
DOI: 10.1007/JHEP06(2015)115. ARXIV: 1503.07138
- [35] R. Aaij et al. (LHCb collaboration). “Measurements of the S-wave fraction in $B^0 \rightarrow K^+ \pi^- \mu^+ \mu^-$ decays and the $B^0 \rightarrow K^*(892)^0 \mu^+ \mu^-$ differential branching fraction”. *J. High Energy Phys.* 11 (2016) 047.
DOI: 10.1007/JHEP11(2016)047. ARXIV: 1606.04731
- [36] R. Aaij et al. (LHCb collaboration). “Branching fraction measurements of the rare $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow f_2'(1525) \mu^+ \mu^-$ decays”. *Phys. Rev. Lett.* 127 (2021) 151801.
DOI: 10.1103/PhysRevLett.127.151801. ARXIV: 2105.14007
- [37] R. Aaij et al. (LHCb collaboration). “Angular analysis of the rare decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ ”. *J. High Energy Phys.* 11 (2021) 043.
DOI: 10.1007/JHEP11(2021)043. ARXIV: 2107.13428
- [38] R. Aaij et al. (LHCb collaboration). “Amplitude analysis of the $\Lambda_c^+ \rightarrow p K^- \pi^+$ decay and Λ_c^+ baryon polarization measurement in semileptonic beauty hadron decays”. 2022.
ARXIV: 2208.03262. Note: to be published in *Phys. Rev. D*
- [39] R. Aaij et al. (LHCb collaboration). “Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays”. *J. High Energy Phys.* 08 (2017) 055.
DOI: 10.1007/JHEP08(2017)055. ARXIV: 1705.05802
- [40] R. Aaij et al. (LHCb collaboration). “Search for lepton-universality violation in $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays”. *Phys. Rev. Lett.* 122 (2019) 191801.
DOI: 10.1103/PhysRevLett.122.191801. ARXIV: 1903.09252
- [41] R. Aaij et al. (LHCb collaboration). “Test of lepton universality using $\Lambda_b^0 \rightarrow p K^- \ell^+ \ell^-$ decays”. *J. High Energy Phys.* 05 (2020) 040.
DOI: 10.1007/JHEP05(2020)040. ARXIV: 1912.08139
- [42] R. Aaij et al. (LHCb collaboration). “Test of lepton universality in beauty-quark decays”. *Nature Physics* 18 (2022) 277.
DOI: 10.1038/s41567-021-01478-8. ARXIV: 2103.11769
- [43] Y. Sato et al. (Belle collaboration). “Measurement of the branching ratio of $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$ relative to $\bar{B}^0 \rightarrow D^{*+} \ell^- \bar{\nu}_\ell$ decays with a semileptonic tagging method”. *Phys. Rev. D* 94 (2016) 072007.
DOI: 10.1103/PhysRevD.94.072007. ARXIV: 1607.07923
- [44] J.P. Lees et al. (BaBar collaboration). “Evidence for an Excess of $\bar{B} \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ Decays”. *Phys. Rev. Lett.* 109 (2012) 101802.
DOI: 10.1103/PhysRevLett.109.101802. ARXIV: 1205.5442

- [45] R. Aaij et al. (LHCb collaboration). “Measurement of the ratio of the $\mathcal{B}(B^0 \rightarrow D^{*-} \tau^+ \nu_\tau)$ and $\mathcal{B}(B^0 \rightarrow D^{*-} \mu^+ \nu_\mu)$ branching fractions using three-prong τ -lepton decays”. *Phys. Rev. Lett.* 120 (2018) 171802.
DOI: 10.1103/PhysRevLett.120.171802. ARXIV: 1708.08856
- [46] R. Aaij et al. (LHCb collaboration). “Measurement of the ratio of branching fractions $\mathcal{R}(D^*)$ and $\mathcal{R}(D^0)$ ”. 2023.
ARXIV: 2302.02886. Note: submitted to *Phys. Rev. Lett.*
- [47] R. Aaij et al. (LHCb collaboration). “Angular analysis and differential branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ ”. *J. High Energy Phys.* 09 (2015) 179.
DOI: 10.1007/JHEP09(2015)179. ARXIV: 1506.08777
- [48] R. Aaij et al. (LHCb collaboration). “Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity”. *J. High Energy Phys.* 02 (2016) 104.
DOI: 10.1007/JHEP02(2016)104. ARXIV: 1512.04442
- [49] R. Aaij et al. (LHCb collaboration). “Measurement of CP -averaged observables in the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay”. *Phys. Rev. Lett.* 125 (2020) 011802.
DOI: 10.1103/PhysRevLett.125.011802. ARXIV: 2003.04831
- [50] M. Beneke, C. Bobeth, and R. Szafron. “Power-enhanced leading-logarithmic QED corrections to $B_q \rightarrow \mu^+ \mu^-$ ”. *J. High Energy Phys.* 10 (2019) 232.
DOI: 10.1007/JHEP10(2019)232. ARXIV: 1908.07011
- [51] C. Bobeth, M. Gorbahn, and E. Stamou. “Electroweak corrections to $B_{s,d} \rightarrow \ell^+ \ell^-$ ”. *Phys. Rev. D* 89 (2014) 034023.
DOI: 10.1103/PhysRevD.89.034023. ARXIV: 1311.1348
- [52] T. Hermann, M. Misiak, and M. Steinhauser. “Three-loop QCD corrections to $B_s \rightarrow \mu^+ \mu^-$ ”. *J. High Energy Phys.* 12 (2013) 097.
DOI: 10.1007/JHEP12(2013)097. ARXIV: 1311.1347
- [53] S. Aoki et al. (Flavour Lattice Averaging group). “FLAG Review 2019”. *Eur. Phys. J. C* 80 (2020) 113.
DOI: 10.1140/epjc/s10052-019-7354-7. ARXIV: 1902.08191
- [54] M. Beneke, C. Bobeth, and R. Szafron. “Enhanced Electromagnetic Corrections to the Rare Decay $B_{s,d} \rightarrow \mu^+ \mu^-$ ”. *Phys. Rev. Lett.* 120 (2018) 011801.
DOI: 10.1103/PhysRevLett.120.011801. ARXIV: 1708.09152
- [55] G. Hiller and M. Schmaltz. “ R_K and future $b \rightarrow s \ell \ell$ physics beyond the standard model opportunities”. *Phys. Rev. D* 90 (2014) 054014.
DOI: 10.1103/PhysRevD.90.054014. ARXIV: 1408.1627

- [56] S. Sahoo and R. Mohanta. “Scalar leptoquarks and the rare B meson decays”. *Phys. Rev. D* 91 (2015) 094019.
DOI: 10.1103/PhysRevD.91.094019. ARXIV: 1501.05193
- [57] M. Bauer and M. Neubert. “Minimal leptoquark explanation for the $R_{D^{(*)}}$, R_K , and $(g-2)_g$ anomalies”. *Phys. Rev. Lett.* 116 (2016) 141802.
DOI: 10.1103/PhysRevLett.116.141802. ARXIV: 1511.01900
- [58] C. Hamzaoui, M. Pospelov, and M. Toharia. “Higgs-boson-mediated FCNC in supersymmetric models with large $\tan\beta$ ”. *Phys. Rev. D* 59 (1999) 095005.
DOI: 10.1103/PhysRevD.59.095005. ARXIV: hep-ph/9807350
- [59] K. S. Babu and C. Kolda. “Higgs-Mediated $B^0 \rightarrow \mu^+ \mu^-$ in Minimal Supersymmetry”. *Phys. Rev. Lett.* 84 (2000) 228.
DOI: 10.1103/PhysRevLett.84.228. ARXIV: hep-ph/9909476
- [60] A. J. Buras, F. De Fazio, and J. Girrbach. “The anatomy of Z' and Z with flavour changing neutral currents in the flavour precision era”. *J. High Energy Phys.* 02 (2013) 116.
DOI: 10.1007/JHEP02(2013)116. ARXIV: 1211.1896
- [61] A. J. Buras, F. De Fazio, and J. Girrbach. “331 models facing new $b \rightarrow s \mu^+ \mu^-$ data”. *J. High Energy Phys.* 02 (2014) 112.
DOI: 10.1007/JHEP02(2014)112. ARXIV: 1311.6729
- [62] G. D’Ambrosio et al. “Minimal flavour violation: an effective field theory approach”. *Nucl. Phys. B* 645 (2002) 155.
DOI: 10.1016/S0550-3213(02)00836-2. ARXIV: hep-ph/0207036
- [63] CMS The ATLAS and LHCb collaborations. “Combination of the ATLAS, CMS and LHCb results on the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays”. LHCb-CONF-2020-002. 2020.
URL: <https://cds.cern.ch/record/2727207>
- [64] R. Aaij et al. (LHCb collaboration). “Measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction and effective lifetime and search for $B^0 \rightarrow \mu^+ \mu^-$ decays”. *Phys. Rev. Lett.* 118 (2017) 191801.
DOI: 10.1103/PhysRevLett.118.191801. ARXIV: 1703.05747
- [65] A. M. Sirunyan et al. (CMS collaboration). “Measurement of properties of $B_s^0 \rightarrow \mu^+ \mu^-$ decays and search for $B^0 \rightarrow \mu^+ \mu^-$ with the CMS experiment”. *J. High Energy Phys.* 04 (2020) 188.
DOI: 10.1007/JHEP04(2020)188. ARXIV: 1910.12127
- [66] M. Aaboud et al. (ATLAS collaboration). “Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector”. *J. High Energy Phys.* 04 (2019) 098.
DOI: 10.1007/JHEP04(2019)098. ARXIV: 1812.03017

- [67] CMS collaboration. “Measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay properties and search for the $B^0 \rightarrow \mu^+ \mu^-$ decay in proton-proton collisions at $\sqrt{s} = 13$ TeV”. *Phys. Lett. B* 842 (2022) 137955.
doi: 10.1016/j.physletb.2023.137955. ARXIV: 2212.10311
- [68] J. L. Rosner. “Low-mass baryon-antibaryon enhancements in B decays”. *Phys. Rev. D* 68 (2003) 014004.
doi: 10.1103/PhysRevD.68.014004. ARXIV: hep-ph/0303079
- [69] K.-J. Tien et al. (Belle collaboration). “Evidence for semileptonic $B^- \rightarrow p\bar{p}\ell^-\bar{\nu}_\ell$ decays”. *Phys. Rev. D* 89 (2014) 011101.
doi: 10.1103/PhysRevD.89.011101. ARXIV: 1306.3353
- [70] R. Aaij et al. (LHCb collaboration). “Observation of the semileptonic decay $B^+ \rightarrow p\bar{p}\mu^+\nu_\mu$ ”. *J. High Energy Phys.* 03 (2020) 146.
doi: 10.1007/JHEP03(2020)146. ARXIV: 1911.08187
- [71] C. Q. Geng and Y.-K. Hsiao. “Semileptonic $B^- \rightarrow p\bar{p}\ell^-\bar{\nu}_\ell$ decays”. *Phys. Lett. B* 704 (2011) 495.
doi: 10.1016/j.physletb.2011.09.065. ARXIV: 1107.0801
- [72] C. Q. Geng, C.-W. Liu, and T.-H. Tsai. “Revisiting semileptonic $B^- \rightarrow p\bar{p}\ell^-\bar{\nu}_\ell$ decays”. *Phys. Rev. D* 104 (2021) 113002.
doi: 10.1103/PhysRevD.104.113002. ARXIV: 2111.07738
- [73] Y.-K. Hsiao. “Semileptonic baryonic B decays”. *Eur. Phys. J. C* 83 (2023) 300.
doi: 10.1140/epjc/s10052-023-11489-9. ARXIV: 2208.09917
- [74] R. Aaij et al. (LHCb collaboration). “Searches for the rare hadronic decays $B^0 \rightarrow p\bar{p}p\bar{p}$ and $B_s^0 \rightarrow p\bar{p}p\bar{p}$ ”. 2022.
ARXIV: 2211.08847. Note: to appear in *Phys. Rev. Lett.*
- [75] R. Aaij et al. (LHCb collaboration). “Observation of $B_{(s)}^0 \rightarrow J/\psi p\bar{p}$ decays and precision measurements of the $B_{(s)}^0$ masses”. *Phys. Rev. Lett.* 122 (2019) 191804.
doi: 10.1103/PhysRevLett.122.191804. ARXIV: 1902.05588
- [76] S. Okubo. “ φ -meson and unitary symmetry model”. *Phys. Lett.* 5 (1963) 165.
doi: 10.1016/S0375-9601(63)92548-9
- [77] R. Aaij et al. (LHCb collaboration). “Evidence for a new structure in the $J/\psi p$ and $J/\psi \bar{p}$ systems in $B_s^0 \rightarrow J/\psi p\bar{p}$ decays”. *Phys. Rev. Lett.* 128 (2022) 062001.
doi: 10.1103/PhysRevLett.128.062001. ARXIV: 2108.04720
- [78] O. S. Brüning et al. “LHC Design Report: The LHC Main Ring”. CERN Yellow Reports: Monographs. 2004.
doi: 10.5170/CERN-2004-003-V-1

References

- [79] L. Evans and P. Bryant. “LHC Machine”. *J. Instrum.* 3 (2008) S08001.
DOI: [10.1088/1748-0221/3/08/s08001](https://doi.org/10.1088/1748-0221/3/08/s08001)
- [80] M. Zinser. “Search for New Heavy Charged Bosons and Measurement of High-Mass Drell-Yan Production in Proton-Proton Collisions: The Large Hadron Collider”. Springer Theses. 2018.
DOI: [10.1007/978-3-030-00650-1](https://doi.org/10.1007/978-3-030-00650-1)
- [81] G. Aad et al. (ATLAS collaboration). “The ATLAS Experiment at the CERN Large Hadron Collider”. *J. Instrum.* 3 (2008) S08003.
DOI: [10.1088/1748-0221/3/08/s08003](https://doi.org/10.1088/1748-0221/3/08/s08003)
- [82] S. Chatrchyan et al. (CMS collaboration). “The CMS experiment at the CERN LHC”. *J. Instrum.* 3 (2008) S08004.
DOI: [10.1088/1748-0221/3/08/s08004](https://doi.org/10.1088/1748-0221/3/08/s08004)
- [83] K. Aamodt et al. (ALICE collaboration). “The ALICE experiment at the CERN LHC”. *J. Instrum.* 3 (2008) S08002.
DOI: [10.1088/1748-0221/3/08/s08002](https://doi.org/10.1088/1748-0221/3/08/s08002)
- [84] A. A. Alves et al. (LHCb collaboration). “The LHCb Detector at the LHC”. *J. Instrum.* 3 (2008) S08005.
DOI: [10.1088/1748-0221/3/08/s08005](https://doi.org/10.1088/1748-0221/3/08/s08005)
- [85] C. Elsässer (LHCb collaboration). “ $\bar{b}b$ production angle plots”. 2023.
URL: https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html
- [86] T. Mombächer. “Beautiful leptons – Setting limits to New Physics with the LHCb experiment”. PhD thesis. Technische Universität Dortmund, 2020.
DOI: [10.17877/DE290R-21750](https://doi.org/10.17877/DE290R-21750)
- [87] R. Antunes-Nobrega et al. (LHCb collaboration). “LHCb reoptimized detector design and performance: Technical Design Report”. CERN-LHCC-2003-030. 2003.
URL: <https://cds.cern.ch/record/630827>
- [88] P. R. Barbosa-Marinho et al. (LHCb collaboration). “LHCb VELO (Vertex Locator): Technical Design Report”. CERN-LHCC-2001-011. 2001.
URL: <https://cds.cern.ch/record/504321>
- [89] R. Aaij et al. (LHCb VELO group). “Performance of the LHCb Vertex Locator”. *J. Instrum.* 9 (2014) P09007.
DOI: [10.1088/1748-0221/9/09/P09007](https://doi.org/10.1088/1748-0221/9/09/P09007). ARXIV: 1405.7808
- [90] S. Amato et al. (LHCb collaboration). “LHCb magnet: Technical Design Report”. CERN-LHCC-2000-007. 2000.
URL: <https://cds.cern.ch/record/424338>
- [91] P. R. Barbosa-Marinho et al. (LHCb collaboration). “LHCb inner tracker: Technical Design Report”. CERN-LHCC-2002-029. 2002.
URL: <https://cds.cern.ch/record/582793>

- [92] P. R. Barbosa-Marinho et al. (LHCb collaboration). “LHCb outer tracker: Technical Design Report”. CERN-LHCC-2001-024. 2001.
URL: <https://cds.cern.ch/record/519146>
- [93] R. Arink et al. (LHCb Outer Tracker group). “Performance of the LHCb Outer Tracker”. *J. Instrum.* 9 (2014) P01002.
DOI: 10.1088/1748-0221/9/01/P01002. ARXIV: 1311.3893
- [94] The LHCb collaboration. “Tracking strategies used in LHCb”. 2023.
URL: <https://twiki.cern.ch/twiki/bin/view/LHCb/LHCbTrackingStrategies>
- [95] R. Aaij et al. (LHCb collaboration). “LHCb Detector Performance”. *Int. J. Mod. Phys. A* 30 (2015) 1530022.
DOI: 10.1142/S0217751X15300227. ARXIV: 1412.6352
- [96] S. Amato et al. (LHCb collaboration). “LHCb RICH: Technical Design Report”. CERN-LHCC-2000-037. 2000.
URL: <https://cds.cern.ch/record/494263>
- [97] M. Adinolfi et al. (LHCb RICH group). “Performance of the LHCb RICH detector at the LHC”. *Eur. Phys. J. C* 73 (2013) 2431.
DOI: 10.1140/epjc/s10052-013-2431-9. ARXIV: 1211.6759
- [98] S. Amato et al. (LHCb collaboration). “LHCb calorimeters: Technical Design Report”. CERN-LHCC-2000-036. 2000.
URL: <https://cds.cern.ch/record/494264>
- [99] C. Abellán Beteta et al. (LHCb Calorimeter group). “Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC”. CERN-LHCb-DP-2020-001. 2020.
ARXIV: 2008.11556
- [100] M. Becker. “Test of lepton flavour universality: Sensitivity study of the decay $B^+ \rightarrow K^+ \pi^+ \pi^- e^+ e^-$ with the LHCb experiment”. Master’s thesis. Technische Universität Dortmund, 2018
- [101] P. R. Barbosa-Marinho et al. (LHCb collaboration). “LHCb muon system: Technical Design Report”. CERN-LHCC-2001-010. 2001.
URL: <https://cds.cern.ch/record/504326>
- [102] A. A. Alves Jr. et al. (LHCb Muon group). “Performance of the LHCb muon system”. *J. Instrum.* 8 (2013) P02022.
DOI: 10.1088/1748-0221/8/02/P02022. ARXIV: 1211.1346
- [103] F. Archilli et al. “Performance of the Muon Identification at LHCb”. *J. Instrum.* 8 (2013) P10020.
DOI: 10.1088/1748-0221/8/10/P10020. ARXIV: 1306.0249
- [104] C. Bozzi. “LHCb Computing Resource usage in 2018”. CERN-LHCb-PUB-2019-004. 2019.
URL: <https://cds.cern.ch/record/2657833>

References

- [105] F. Alessio et al. “The LHCb data acquisition during LHC Run 1”. *J. Phys.: Conf. Ser.* 513 (2014) 012033.
DOI: 10.1088/1742-6596/513/1/012033
- [106] F. Dordei. “LHCb detector and trigger performance in Run II”. *EPJ Web Conf.* 164 (2017) 01016.
DOI: 10.1051/epjconf/201716401016
- [107] R. Aaij et al. “A comprehensive real-time analysis model at the LHCb experiment”. *J. Instrum.* 14 (2019) P04006.
DOI: 10.1088/1748-0221/14/04/P04006. ARXIV: 1903.01360
- [108] R. Aaij et al. “The LHCb trigger and its performance in 2011”. *J. Instrum.* 8 (2013) P04022.
DOI: 10.1088/1748-0221/8/04/p04022. ARXIV: 1211.3055
- [109] R. Aaij et al. (LHCb collaboration). “Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC”. *J. Instrum.* 14 (2019) P04013.
DOI: 10.1088/1748-0221/14/04/p04013. ARXIV: 1812.10790
- [110] J. Albrecht. “The LHCb Trigger System”. *Nucl. Phys. B Proceedings Supplements* 187 (2009) 237.
DOI: 10.1016/j.nuclphysbps.2009.01.033
- [111] G. Barrand et al. “Gaudi - A software architecture and framework for building HEP data processing applications”. *Comput. Phys. Commun.* 140 (2001) 45.
DOI: 10.1016/S0010-4655(01)00254-5
- [112] “The Moore project”. 2023.
URL: <https://lhcbdoc.web.cern.ch/lhcbdoc/moore/>
- [113] “The Brunel project”. 2023.
URL: <https://lhcbdoc.web.cern.ch/lhcbdoc/brunel/>
- [114] “The DaVinci project”. 2023.
URL: <https://lhcbdoc.web.cern.ch/lhcbdoc/davinci/>
- [115] R. Brun and F. Rademakers. “ROOT: An object oriented data analysis framework”. *Nucl. Instrum. Meth.* A389 (1997) 81.
DOI: 10.1016/S0168-9002(97)00048-X
- [116] “The Gauss project”. 2023.
URL: <https://lhcbdoc.web.cern.ch/lhcbdoc/gauss/>
- [117] T. Sjöstrand et al. “Pythia 6.4 physics and manual”. *J. High Energy Phys.* 05 (2006) 26.
DOI: 10.1088/1126-6708/2006/05/026
- [118] I. Belyaev et al. “Handling of the generation of primary events in Gauss, the LHCb simulation framework”. *J. Phys.: Conf. Ser.* 331 (2011) 032047.
DOI: 10.1088/1742-6596/331/3/032047

- [119] D. J. Lange. “The EvtGen particle decay simulation package”. Nucl. Instrum. Meth. A462 (2001) 152.
DOI: 10.1016/S0168-9002(01)00089-4
- [120] P. Golonka and Z. Was. “Photos Monte Carlo: a precision tool for QED corrections in Z and W decays”. Eur. Phys. J. C45 (2006) 97.
DOI: 10.1140/epjc/s2005-02396-4
- [121] S. Agostinelli et al. “Geant4 - a simulation toolkit”. Nucl. Instrum. Methods. Phys. Res. A506 (2003) 250.
DOI: 10.1016/S0168-9002(03)01368-8
- [122] “The Boole project”. 2023.
URL: <https://lhcbdoc.web.cern.ch/lhcbdoc/boole/>
- [123] S. Tolk et al. (LHCb collaboration). “Data driven trigger efficiency determination at LHCb”. CERN-LHCb-PUB-2014-039. 2014.
URL: <https://cds.cern.ch/record/1701134>
- [124] I. Williams. “Studies of rare B-meson decays to muons at the LHCb experiment”. PhD thesis. University of Cambridge, 2022.
URL: <https://cds.cern.ch/record/2807727>
- [125] L. E. Yeomans. “A search for the $B_{(s)}^0 \rightarrow \mu^+ \mu^- (\gamma)$ decays at LHCb”. PhD thesis. University of Liverpool, 2022.
URL: <https://cds.cern.ch/record/2808355>
- [126] S. Ferreres Solé. “The beauty of the rare”. PhD thesis. Maastricht University, 2022.
URL: <https://cds.cern.ch/record/2843584>
- [127] F. Archilli et al. “Background studies for $B^0 \rightarrow \mu^+ \mu^-$ analysis optimization”. CERN-LHCb-INT-2014-047. 2014.
URL: <https://cds.cern.ch/record/1970886>
- [128] L. Anderlini et al. (LHCb collaboration). “The PIDCalib package”. CERN-LHCb-PUB-2016-021. 2016.
URL: <https://cds.cern.ch/record/2202412>
- [129] J. Therhaag. “TMVA – Toolkit for multivariate data analysis in ROOT”. Proceedings of Science 120 (2010) 510 (ICHEP 2010).
DOI: 10.22323/1.120.0510. ARXIV: physics/0703039
- [130] “The hep_ml package”. 2023.
URL: https://github.com/arogozhnikov/hep_ml
- [131] M. Pivk and F.R. Le Diberder. “sPlot: A statistical tool to unfold data distributions”. Nucl. Instrum. Meth. A555 (2005) 356.
DOI: 10.1016/j.nima.2005.08.106. ARXIV: physics/0402083

References

- [132] A. Blum, A. Kalai, and J. Langford. “Beating the hold-out: bounds for k-Fold and progressive cross-validation”. Proceedings of COLT '99 (1999) 203.
DOI: 10.1145/307400.307439
- [133] R. Aaij et al. (LHCb collaboration). “Determination of the sign of the decay width difference in the B_s^0 system”. Phys. Rev. Lett. 108 (2012) 241801.
DOI: 10.1103/PhysRevLett.108.241801. ARXIV: 1202.4717
- [134] R. Aaij et al. (LHCb collaboration). “Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \rightarrow J/\psi K^+ K^-$ and $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$ decays”. Phys. Rev. D87 (2013) 112010.
DOI: 10.1103/PhysRevD.87.112010. ARXIV: 1304.2600
- [135] T. Skwarnicki. “A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances”. PhD thesis. Institute of Nuclear Physics Krakow, 1986.
URL: <http://inspirehep.net/record/230779/>
- [136] P. A. Zyla et al. (Particle Data group). “Review of Particle Physics”. Prog. Theor. Exp. Phys. 2020 (2020) 083C01.
DOI: 10.1093/ptep/ptaa104
- [137] N. Isgur et al. “Semileptonic B and D decays in the quark model”. Phys. Rev. D39 (1989) 799.
DOI: 10.1103/PhysRevD.39.799
- [138] D. Scora and N. Isgur. “Semileptonic meson decays in the quark model: An update”. Phys. Rev. D52 (1995) 2783.
DOI: 10.1103/PhysRevD.52.2783
- [139] C. Bourrely, L. Lellouch, and I. Caprini. “Model-independent description of $B \rightarrow \pi l \nu$ decays and a determination of $|V_{ub}|$ ”. Phys. Rev. D79 (2009) 013008.
DOI: 10.1103/PhysRevD.79.013008. ARXIV: 0807.2722
- [140] Y. Amhis et al. (Heavy Flavor Averaging group). “Averages of b -hadron, c -hadron, and τ -lepton properties as of 2018”. Eur. Phys. J. C81 (2021) 226.
DOI: 10.1140/epjc/s10052-020-8156-7. ARXIV: 1909.12524
- [141] R. Aaij et al. (LHCb collaboration). “First observation of the decay $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and measurement of $|V_{ub}|/|V_{cb}|$ ”. Phys. Rev. Lett. 126 (2021) 081804.
DOI: 10.1103/PhysRevLett.126.081804. ARXIV: 2012.05143
- [142] W.-F. Wang et al. “The semileptonic decays $B/B_s \rightarrow (\pi, K)(\ell^+ \ell^-, \ell \nu, \nu \bar{\nu})$ in the perturbative QCD approach beyond the leading-order”. Phys. Rev. D86 (2012) 114025.
DOI: 10.1103/PhysRevD.86.114025. ARXIV: 1207.0265

- [143] R. Aaij et al. (LHCb collaboration). “Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays”. *Nature Physics* 11 (2015) 743.
DOI: 10.1038/nphys3415. ARXIV: 1504.01568
- [144] R. Aaij et al. (LHCb collaboration). “Measurement of b -hadron fractions in 13 TeV pp collisions”. *Phys. Rev. D* 100 (2019) 031102(R).
DOI: 10.1103/PhysRevD.100.031102. ARXIV: 1902.06794
- [145] W. Detmold et al. “ $\Lambda_b \rightarrow p\ell\nu$ form factors from lattice QCD with static b quarks”. *Phys. Rev. D* 88 (2013) 014512.
DOI: 10.1103/PhysRevD.88.014512. ARXIV: 1306.0446
- [146] W. Detmold, Ch. Lehner, and S. Meinel. “ $\Lambda_b \rightarrow \ell\nu$ and $\Lambda_b \rightarrow \Lambda_c\ell\nu$ form factors from lattice QCD with relativistic heavy quarks”. *Phys. Rev. D* 92 (2015) 034503.
DOI: 10.1103/PhysRevD.92.034503. ARXIV: 1503.01421
- [147] A. Khodjamirian et al. “Form Factors and Strong Couplings of Heavy Baryons from QCD Light-Cone Sum Rules”. *J. High Energy Phys.* 09 (2011) 106.
DOI: 10.1007/JHEP09(2011)106. ARXIV: 1108.2971
- [148] R. Aaij et al. (LHCb collaboration). “Measurement of b hadron production fractions in 7 TeV pp collisions”. *Phys. Rev. D* 85 (2012) 032008.
DOI: 10.1103/PhysRevD.85.032008. ARXIV: 1111.2357
- [149] R. Aaij et al. (LHCb collaboration). “Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics”. *J. High Energy Phys.* 04 (2013) 001.
DOI: 10.1007/JHEP04(2013)001. ARXIV: 1301.5286
- [150] R. Aaij et al. (LHCb collaboration). “Measurement of the branching fraction of the $B^0 \rightarrow D_s^+\pi^-$ decay”. *Eur. Phys. J. C* 81 (2021) 314.
DOI: 10.1140/epjc/s10052-020-08790-2. ARXIV: 2010.11986
- [151] R. Aaij et al. (LHCb collaboration). “Measurement of f_s/f_u variation with proton-proton collision energy and B -meson kinematics”. *Phys. Rev. Lett.* 124 (2020) 122002.
DOI: 10.1103/PhysRevLett.124.122002. ARXIV: 1910.09934
- [152] R. Aaij et al. (LHCb collaboration). “Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B_s^0 decay branching fractions”. *Phys. Rev. D* 104 (2021) 032005.
DOI: 10.1103/PhysRevD.104.032005. ARXIV: 2103.06810
- [153] R. Aaij et al. (LHCb collaboration). “Measurement of the track reconstruction efficiency at LHCb”. *J. Instrum.* 10 (2015) P02007.
DOI: 10.1088/1748-0221/10/02/P02007. ARXIV: 1408.1251

References

- [154] K. Cranmer. “Kernel estimation in high-energy physics”. *Comput. Phys. Commun.* 136 (2001) 198.
DOI: 10.1016/S0010-4655(00)00243-5. ARXIV: hep-ex/0011057
- [155] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. *Ann. Math. Stat.* 9 (1938) 60.
DOI: 10.1214/aoms/1177732360
- [156] A. L. Read. “Presentation of search results: the CLs technique”. *J. Phys. G: Nucl. Part. Phys.* 28:10 (2002) 2693.
DOI: 10.1088/0954-3899/28/10/313
- [157] “The GammaCombo project”. 2023.
URL: <https://gammacombo.github.io/>
- [158] R. Aaij et al. (LHCb collaboration). “Measurement of the CKM angle γ from a combination of LHCb results”. *J. High Energy Phys.* 12 (2016) 087.
DOI: 10.1007/JHEP12(2016)087. ARXIV: 1611.03076
- [159] The LHCb collaboration. “The LHCb upgrade I”. CERN-LHCb-DP-2022-002. 2023.
ARXIV: 2305.10515
- [160] The LHCb collaboration. “Future physics potential of LHCb”. CERN-LHCb-PUB-2022-012. 2022.
URL: <https://cds.cern.ch/record/2806113>
- [161] D. Martínez Santos and F. Dupertuis. “Mass distributions marginalized over per-event errors”. *Nucl. Instrum. Meth.* A764 (2014) 150.
DOI: 10.1016/j.nima.2014.06.081. ARXIV: 1312.5000
- [162] J. Albrecht et al. “Performance of the LHCb High Level Trigger in 2012”. *J. Phys.: Conf. Ser.* 513 (2014) 012001.
DOI: 10.1088/1742-6596/513/1/012001. ARXIV: 1310.8544
- [163] T. Chen and C. Guestrin. “XGBoost: A Scalable Tree Boosting System”. *Proceedings of KDD '16* (2016) 785.
DOI: 10.1145/2939672.2939785
- [164] G. Punzi. “Sensitivity of searches for new signals and its optimization”. 2003.
ARXIV: physics/0308063
- [165] A. Wald. “Tests of statistical hypotheses concerning several parameters when the number of observations is large”. *Trans. Amer. Math. Soc.* 54 (1943) 426.
DOI: 10.2307/1990256

Acknowledgements

An dieser Stelle möchte ich mich bei Menschen bedanken, ohne deren Zuspruch und Unterstützung diese Arbeit nicht zustande gekommen wäre.

Mein erster Dank gilt Prof. Johannes Albrecht, bei dem ich vor über acht Jahren eine meiner ersten Vorlesungen zur Teilchenphysik gehört habe. Fasziniert von diesem Teilgebiet der Physik schrieb ich infolgedessen sowohl meine Bachelor- als auch meine Masterarbeit innerhalb seiner Arbeitsgruppe. Daran anschließend bot sich mir die Möglichkeit, meine Studien im Rahmen einer Promotion fortzuführen. Danke Johannes, dass du mir diesen Weg ermöglicht hast! Auch danke ich Prof. Kevin Kröniger, der sich freundlicherweise bereiterklärt hat, als weiterer Gutachter diese Arbeit zu bewerten.

In der Experimentalphysik wird selten alleine geforscht und so wären die Ergebnisse dieser Dissertation nicht zustande gekommen, hätte ich nicht die Möglichkeit gehabt, mit ganz wunderbaren Menschen zusammenzuarbeiten. Ich bedanke mich beim ganzen Team der $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ Analyse. Danke Flavio, Harry, Francesco, Fabio, Jacco, Silvia, Titus, Mick, Matteo, Marco, Ifan und Lauren für all die langen und produktiven gemeinsamen Meetings, die gerade in Zeiten des Corona-Lockdowns so wichtig waren. Insbesondere möchte ich Titus danken, der mich in Dortmund an die Analyse herangeführt hat und keiner naiven Frage von mir ausgewichen ist. Auch an die $B_{(s)}^0 \rightarrow p\bar{p}\mu^+\mu^-$ Analyse habe ich mich nicht alleine herangewagt. Danke Vitalii und Lukas für die guten Diskussionen und die hilfreichen Ratschläge.

Neben der wissenschaftlichen Arbeit war mir auch immer ein gutes Miteinander innerhalb der Dortmunder Arbeitsgruppe wichtig. Danke Quentin, Julian, Alex und insbesondere Gerwin und Vukan, dass ihr mich als Bürokollegen ausgehalten habt. Danke auch an Nicole, Michelle, Henning, Jan, Jan, Louis und all die anderen Mitglieder der Arbeitsgruppe für die gemeinsame Zeit.

Ein besonderer Dank gilt auch Simone und Kevin für die freundschaftliche Zeit im Studium und danach.

Zum Schluss danke ich natürlich meiner Familie, ohne deren grenzenlose Unterstützung ich jetzt nicht an diesem Punkt angekommen wäre. Danke, dass es euch gibt!