

Dissertation

submitted to the

Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

MSc Michael Kolpin

born in: Köln, Germany

Oral examination: 13.12.2017

Measurement of the differential
branching fraction and angular
analysis of the decay
 $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and determination
of the track reconstruction
efficiencies at the LHCb
experiment

Referees: Prof. Dr. Ulrich Uwer

Prof. Dr. André Schöning

Abstract: The measurement of the differential branching fraction and an angular analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ is performed, using a dataset collected by the LHCb experiment in proton-proton collisions in 2011 and 2012, corresponding to an integrated luminosity of 3 fb^{-1} . The results are presented as function of q^2 , the square of the dimuon invariant mass. A full set of angular observables is extracted for the first time for this decay mode, where all measured angular coefficients are consistent with Standard Model expectations. For the range of $1 < q^2 < 6 \text{ GeV}^2/c^4$ the branching fraction is found to be more than 3 standard deviations below the Standard Model prediction, which is derived using a combination of light-cone sum rules and lattice QCD calculations. This tension is compatible with observations made by other measurements of electroweak penguin decays mediated by $b \rightarrow s \ell^+ \ell^-$ transitions in this q^2 region. Possible explanations are theoretical models which predict new heavy degrees of freedom at the TeV scale, as well as underestimated contributions from $c\bar{c}$ -loop corrections in the SM calculations. An important aspect of the measurements of branching fractions is the determination of the track reconstruction efficiency. It is determined from $J/\psi \rightarrow \mu^+ \mu^-$ decays, where the data-driven tag-and-probe technique is used. During LHC Run II this measurement is performed in a “quasi-online” approach, directly after the data has been recorded. For the 2015 data-taking period of the LHCb experiment, the track reconstruction efficiency is found to be $(96.14 \pm 0.03)\%$, where the uncertainty is statistical only.

Kurzfassung: Diese Arbeit präsentiert die Messung des differentiellen Verzweigungsverhältnisses und eine Winkelanalyse des Zerfalls $B_s^0 \rightarrow \phi \mu^+ \mu^-$ mit einem Datensatz, der in den Jahren 2011 und 2012 mithilfe des LHCb Experiments aufgezeichnet wurde und einer integrierten Luminosität von 3 fb^{-1} entspricht. Die Ergebnisse werden als Funktion des Quadrates der invarianten Di-Muon-Masse q^2 angegeben. Zum ersten Mal wird für diesen Zerfallsmodus die maximale Anzahl von acht Winkelobservablen bestimmt. All diese acht Observablen stimmen gut mit Standardmodellvorhersagen überein. Im Bereich $1 < q^2 < 6 \text{ GeV}^2/c^4$ liegt der Wert des Verzweigungsverhältnisses mehr als 3 Standardabweichungen unterhalb der Standardmodellvorhersagen, die aus einer Kombination von Lichtkegel-Summationsregeln und Lattice-QCD Berechnungen stammen. Diese Abweichung stimmt mit den Messungen anderer elektroschwacher Pinguinzerfälle in diesem q^2 Bereich überein, die durch $b \rightarrow s \ell^+ \ell^-$ Übergänge vermittelt werden. Mögliche Erklärungen hierfür liefern theoretische Modelle, die neue schwere Freiheitsgrade im Bereich der TeV Skala vorhersagen, aber auch unterschätzte Beiträge von $c\bar{c}$ -Schleifenkorrekturen bei der Berechnung der Vorhersagen.

Für die Messung von Verzweigungsverhältnissen spielt die Bestimmung der Spurrekonstruktionseffizienz eine entscheidende Rolle. Diese wird mithilfe von $J/\psi \rightarrow \mu^+ \mu^-$ Zerfällen bestimmt, wobei die datengestützte Tag-and-Probe Methode verwendet wird. Während des zweiten Datennahmezeitraums des LHCb Experiments wird diese Messung “quasi online” durchgeführt, direkt nach der Aufzeichnung der Daten. Die Spurrekonstruktionseffizienz des LHCb Experiments während der Datennahme im Jahr 2015 wird auf $(96.14 \pm 0.03)\%$ bestimmt, wobei der gegebene Fehler rein statistisch ist.

Contents

Preface	v
Part I: Tracking efficiency measurement in LHCb Run II	1
1 Introduction	3
2 The LHCb Experiment	5
2.1 The Large Hadron Collider	5
2.2 The LHCb Detector	6
2.3 The Stripping framework	17
2.4 LHCb simulation	17
2.5 Datasets	18
3 Tracking in the LHCb experiment	19
3.1 Particle interaction with material	19
3.2 Properties of particle tracks	21
3.3 Kalman filters used in track fits	22
3.4 Track reconstruction algorithms in LHCb	23
3.5 Fake tracks	25
4 Determination of the track reconstruction efficiency	27
4.1 Tag-and-probe technique with $J/\psi \rightarrow \mu^+\mu^-$ decays	27
4.2 Results of the LHC Run I data-taking period	32
5 Tracking efficiency measurement in Run II	35
5.1 Trigger line implementation	35
5.2 Tracking efficiency evaluation	38
5.3 Method validation and results	39
5.4 Additional performance tests	42
5.5 Systematic uncertainties	48
6 TrackCalib: User-defined tracking efficiency evaluation	51
6.1 Features of TrackCalib	51
6.2 Data flow and tool structure	52
7 Conclusion	55
Part II: Relative branching fraction measurement and angular analysis of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$	57
8 Introduction	59
9 Theoretical Framework	61

9.1	The Standard Model of Particle Physics	61
9.2	Phenomenology of $b \rightarrow s\ell^+\ell^-$ transitions	68
9.3	New Physics models affecting FCNC transitions	76
10	Analysis Strategy	79
11	Selection of the signal candidates	83
11.1	Background sources	84
11.2	Boosted Decision Trees	85
11.3	Trigger selection	86
11.4	Stripping selections	88
11.5	Pre-selection and Multivariate Analysis	89
11.6	Fully selected data sample	93
12	Corrections of the simulated event samples	97
12.1	Simulated event samples	97
12.2	Track reconstruction efficiency corrections	98
12.3	Correction of <code>IsMuon</code> flag efficiency	100
12.4	Recreation of particle identification variables	100
12.5	Additional corrections	101
12.6	Control of improved agreement	102
13	Efficiency determination	111
13.1	Detector acceptance	111
13.2	Reconstruction and selection efficiency	112
13.3	Trigger efficiency	115
13.4	Efficiency combination	115
13.5	Angular acceptance	118
14	Relative branching fraction measurement	123
14.1	Extended Maximum Likelihood Fits	123
14.2	Signal and background models	125
14.3	Signal and normalisation fit results	126
14.4	Relative differential branching fraction calculation	128
15	Analysis of the angular distributions of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$	133
15.1	Fitting procedure and model	133
15.2	Fit validation	136
15.3	Uncertainty evaluation	137
15.4	Results	138
16	Determination of the systematic uncertainties	149
16.1	Peaking backgrounds	149
16.2	Correction of simulated events	151
16.3	Residual differences in kinematic variables	154
16.4	Model dependence of the detector acceptance	155
16.5	Branching fraction of $J/\psi \rightarrow \mu^+\mu^-$	156
16.6	Limited size of simulated event samples	156
16.7	Signal model	157
16.8	Background model and parametrisation	158

16.9	Decay time acceptance	158
16.10	Dependence of angular acceptance on q^2	160
16.11	S-wave contribution in the K^+K^- system	160
16.12	Summary	161
17	Results	163
17.1	Differential branching fraction	163
17.2	Analysis of the angular distributions	165
17.3	Implications	166
18	Conclusion	173
 Appendix		175
A 1	Angular acceptance in q^2 bins	177
A 2	Angular acceptance: two-dimensional projections	180
A 3	Parameter scans	185
A 4	Feldman-Cousins scans	190
 Bibliography		201

Preface

The measurements presented in this thesis were performed within the LHCb collaboration. This collaboration consists of more than 800 authors from many different institutes all around the globe. No part of this work would have been possible without the numerous contributions from the other current and former collaboration members and the enormous amount of help provided to the author.

The measurement of the track reconstruction efficiency of the LHCb experiment, which is presented in the first part of this thesis, is based largely on the work performed by multiple members of the LHCb collaboration during LHC Run I. For the measurement for that data-taking period the probe reconstruction algorithms have been created, and the tag-and-probe technique developed. The author was, with the supervision and support of a post-doctoral researcher at Heidelberg University, fully responsible for the implementation of the probe reconstruction and creation of dedicated trigger lines in the LHCb software trigger. This allows the determination of the tracking efficiencies in a “quasi-online” approach during LHC Run II. Furthermore, the author validated and performed the measurement of the track reconstruction efficiencies in Run II, and developed the software tool `TrackCalib` as part of this measurement. The automation of the dataset creation from the data collected by the software trigger was implemented in close collaboration with another PhD student, with the support of multiple other LHCb members.

The analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$, presented in the second part of this thesis, was completed by a working group of four people, which besides the author included two post-doctoral fellows at CERN and a Master student under the supervision of the author. The author contributed to all steps of the measurement, with main responsibilities on the correction of the simulated event samples, the efficiency determination, the fitting process, and the evaluation of the systematic uncertainties. The fitting framework used had been developed previously by one of the other proponents of this analysis for the measurement of CP violation in the decay $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ [1]. For the measurement of the differential branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$, the author adapted this fitting framework, including the implementation of the proper creation of pseudo-experiments using low-statistics datasets. The measurement described in the second part of this thesis has been published by the LHCb collaboration in JHEP 1509 (2015) 179 [2].

Part I: Tracking efficiency measurement in LHCb Run II

1 Introduction

The wish to understand how the world around us works is one of the oldest desires of humankind. From our earliest ancestors, who thought ancient gods to be responsible for the circle of night and day, through Eratosthenes calculating the earth's circumference from basic mathematical principles, to the creation of complex modern theoretical physics models, this urge to find the fundamental working principles of all things has guided us on. Today, one frontier of fundamental research is the field of Particle Physics, where the interactions of what we currently believe to be the fundamental building blocks of matter are investigated. Research in this field is typically done at particle colliders, where particles are accelerated to velocities close to the speed of light and brought to collision, recreating conditions similar to those present in the early universe. Unstable particles are created in these high-energy collisions, and their decays measured to test the established physics models and search for explanations for phenomena not described by these models.

Sub-atomic particles produced in high-energy physics experiments are typically short-lived or have relativistic velocities. Thus specific detection techniques are required in order to determine their momentum, charge, and further properties. For that reason, a detector to visualise the particle trajectories is the heart of every high energy physics experiment. However, its ability to reconstruct particle tracks is subject to inefficiencies due to multiple effects: the geometrical layout of the detector limits its acceptance for the detection of particles, and the performance of the sub-detectors to measure the interactions with the particle is imperfect. Furthermore, the flight path of the particle cannot always be successfully reconstructed from the hit information by the pattern recognition algorithms.

The precise determination of the efficiencies to detect and reconstruct particles is an essential part of the measurements of branching fractions and CP -violation parameters performed at the LHCb experiment, such as the analysis which will be presented in the second part of this thesis. Often these efficiencies are evaluated using simulated event samples. The simulation models used by the LHCb collaboration provide in general a good description of the detector, but small differences between real data and simulated event samples might persist. In order to minimise uncertainties caused by these differences, imperfections of the simulation need to be carefully investigated and accounted for.

The measurement of the track reconstruction efficiencies in the LHCb experiment is performed using the so-called tag-and-probe technique: a sample of $J/\psi \rightarrow \mu^+ \mu^-$ decay candidates is used, where one muon track, the “tag” track, is tightly selected and fully reconstructed. The second muon, the so-called “probe” signature, is reconstructed from hit information complementary to what is used by the default reconstruction. This probe signature is classified as efficient, if a track from the nominal reconstruction is found which can be associated with it by searching for shared detector hit information.

In LHC Run I this tracking efficiency determination required a set of algorithms to reconstruct the probe signature, executed after the data-taking process. The size of the available data sample used was limited due to computational limitations. For Run II the execution of these algorithms is moved into the LHCb software trigger, allowing for a “quasi online” efficiency measurement. Furthermore, this more direct approach enables the collection of a larger dataset, as computational limitations are reduced. The first part of this thesis describes the measurement of the track reconstruction efficiency during LHC Run II and is organised as follows:

The Large Hadron Collider (LHC) and the LHCb experiment based there are introduced in Chap. 2. The LHCb detector is presented, and its sub-detectors important for the scope of this thesis are described in more detail. Chapter 3 describes the general principles of particle tracks and their reconstruction, and gives an overview of the different track types and the track reconstruction algorithms as used within the LHCb experiment. Afterwards, Chap. 4 introduces the tag-and-probe technique used for the tracking efficiency measurement, and presents the results of the measurement in LHC Run I.

The implementation of the “quasi-online” measurement for LHC run II is discussed in Chap. 5, which also shows the results of the measurement of the performance of the LHCb tracking system. Additionally, Chap. 6 introduces the software tool `TrackCalib`, which allows for the customised evaluation of the track reconstruction efficiencies. Finally, Chap. 7 summarises the efforts and results of the track reconstruction efficiencies of the LHCb experiment in LHC Run II.

2 The LHCb Experiment

The usage of precise measurement instruments has been an essential aspect of experimental investigations since the early days of Physics. While for basic research, rulers, clocks and other simple devices were sufficient for this purpose, modern particle physics experiments require more intricate tools to measure the properties of sub-atomic particles. The analyses presented in this thesis were performed within the environment of the Large Hadron Collider beauty (LHCb) collaboration. This chapter will describe the LHCb experiment and detector, by which the dataset used in this thesis was collected, and the Large Hadron Collider, where the experiment is located. The general design of the detector and its elements is taken from Ref. [3]. A more detailed overview is given for detector subsystems that are especially important for the content of this thesis.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is located at the main CERN¹ facility close to Geneva, Switzerland. It is the largest particle accelerator created so far, with a circumference of about 27km and lies approximately 100m beneath the Swiss-French border. The LHC collides two circulating proton beams at centre-of-mass energies of up to 14 TeV. Four major experiments are located at dedicated interaction points: LHCb, ATLAS, CMS and ALICE. The layout of the LHC and the placement of these experiments is shown in Fig. 2.1.

A beam is divided into bunches with each bunch containing $\sim 10^{11}$ protons. The protons are accelerated through a chain of accelerators (Linac, PS, SPS, and LHC) to reach the desired energy near 7 TeV. The two beams are brought together at the locations of the experimental detectors, where some of the protons in each bunch will interact with some in the oppositely circulating bunches. In the original design, the number of bunches simultaneously being accelerated can reach up to 2808, leading to an instantaneous luminosity of $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. At the nominal time distance of 25 ns between these bunches, the rate of bunch crossings at the interaction points is about 30 MHz, due to non-uniform filling schemes. The number of bunches, their time distance and the focusing of the beams might vary between data taking periods and experiments. In addition to the acceleration of proton bunches, the LHC also can perform experiments of colliding heavy ions with each other or with proton beams.

The Large Hadron Collider was operated at centre-of-mass energies of 7 TeV and 8 TeV in the 2011 and 2012 data taking periods, respectively. In LHC run II, ongoing since summer 2015, it has been increased to 13 TeV, close to the design goal.

The individual detectors are designed for different purposes. ALICE is aiming to

¹European Organization for Nuclear Research, abbreviation derived from the name Conseil Européen pour la Recherche Nucléaire

investigate heavy ion collisions, recording proton-proton events mostly for reference compared to the ion-ion or proton-ion collisions. These heavy ion collisions aim to improve the understanding of the strong interaction as well as investigate states of extremely high energy density, such as the Quark-Gluon-Plasma. ATLAS and CMS are general purpose experiments, both trying to improve the knowledge of the Standard Model as well as searching for new physics beyond the Standard Model (BSM). As many signature decays of BSM physics models are expected to have decay products which are not directly detectable, these detectors cover nearly the full solid angle (often called 4π -detectors) in order to measure missing transversal energies. The discovery of the Higgs boson in 2012 by these two experiments [5] was the first success of the Large Hadron Collider, and completes the set of Standard Model particles.

The LHCb detector will be discussed in more detail in the following section.

2.2 The LHCb Detector

The LHCb detector is, in contrast to the other three LHC detectors, only covering a comparably small angular region along the beam pipe of 10-300 mrad in the horizontal plane and 10-250 mrad in the vertical plane. It is a single-arm forward spectrometer, with the main purpose of investigating CP violation in decays of beauty and charm hadrons and searching for contributions of physics beyond the SM in loop processes. At the centre-of-mass energies at the LHC, $b\bar{b}$ and $c\bar{c}$ pairs are created mainly by gluon-fusion and $q\bar{q}$ annihilation processes. Due to the typically very

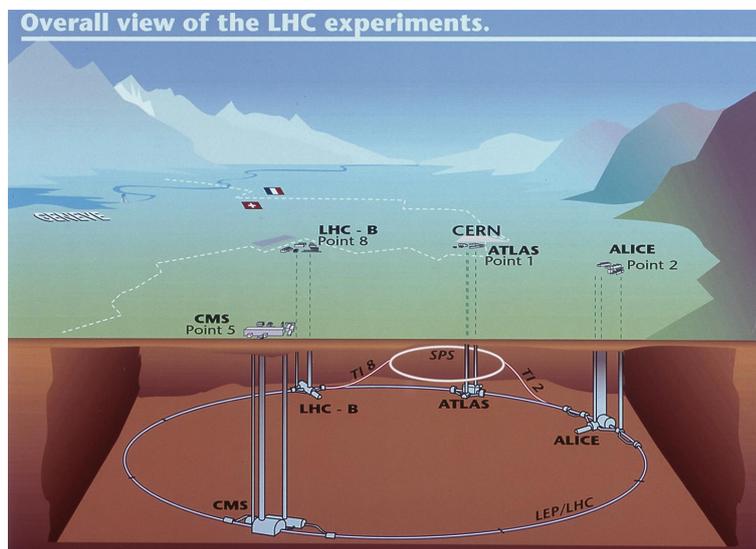


Figure 2.1: Schematic overview of the Large Hadron Collider and the location of the four major experiments performed there [4].

different momentum fractions carried by the initial partons, most quark pairs are emitted close to the beam directions. Figure 2.2 shows the angular distributions of $b\bar{b}$ pairs produced in the LHCb detector. θ_1 and θ_2 are the angles between of the flight direction of the b quark and \bar{b} quark, respectively, compared to the beam direction z . As can be seen, the distribution peaks at 0° and 180° . Due to spatial limitations of the cavern, the LHCb detector covers only the forward direction downstream of the collision. However, about 25% of $b\bar{b}$ pairs produced end up within the LHCb acceptance.

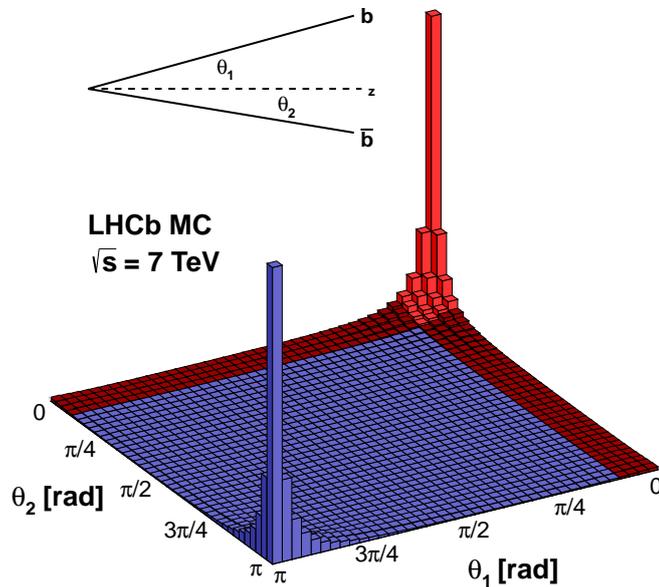


Figure 2.2: Angular distribution of $b\bar{b}$ pairs produced in proton-proton collisions in the LHCb experiment [4].

The LHCb detector is positioned in the cavern formerly used by the DELPHI detector, 100m underground at a CERN facility close to Preveessin, France. While it is oriented in the forward direction along the beam pipe, the general set and order of subdetectors is very similar to classical 4π detectors, which cover the full solid angle. In Fig. 2.3 a schematic overview of the LHCb detector and its components is shown, and the y-axis and z-axis of the LHCb coordinate system are indicated. The origin of the z-axis is placed at the interaction points and extends downstream along the beam pipe, while the y-axis points upward. Furthermore, the direction of the x-axis is pointing into the drawing plane, following the common three-dimensional axis definition. The Vertex

Locator (VELO) is a tracking detector extremely close to the beam and surrounds the interaction point in order to resolve the decay vertices of B and D hadrons. Behind the VELO, the first of two Ring-Imaging Cherenkov detectors (RICH) is placed, which delivers information used for particle identification. Farther downstream the Tracker Turicensis (TT) is located, giving additional information about tracks before they reach the magnet. The magnet bends the path of charged particles which, in combination with the tracking stations (Inner and Outer Tracker) can be used to obtain information about the track momentum. Another RICH detector adds additional particle identification data further downstream, while the calorimeters are mostly used by the trigger system. Finally, the muon system, consisting of five muon stations, is used to identify muons and provide additional tracking information for these leptons.

Key features of the LHCb detectors are a high performance in resolving primary

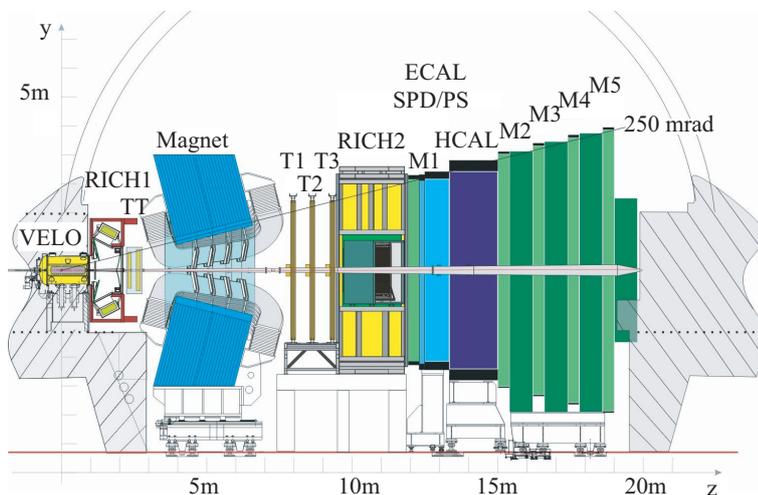


Figure 2.3: Schematic overview of the LHCb detector and its components [4].

and secondary vertices, precise momentum measurements, and good identification of the final-state particle species. The following sections will give an overview of a selection of subdetectors and how they contribute to these requirements.

2.2.1 The Vertex Locator

Due to the large boost of B and D hadrons originating from the asymmetric parton momenta, they typically traverse distances of a few cm before decaying. Many observables related to CP violation can be derived from analyses involving the decay time of neutral B mesons, which is directly connected to the flight distance of the respective particle. The Vertex Locator (VELO) is a silicon microstrip detector which provides basic tracking information close to the interaction point to allow for precise reconstruction of the primary interaction vertex as well as the decay vertices

of short-lived particles originating from the primary interaction.

In order to allow for the opening and closing mechanic, the VELO is split in two retractable halves, with a small overlap to help with the sensor alignment and ensure the full angular coverage. In the closed position it is placed 7 mm from the beam, and is moved out during periods of a wider beam profile, *e.g.* while the beam is being injected into the accelerator. Figure 2.4 schematically shows the structure of the VELO.

It is comprised of an array of silicon strip modules placed along the beam direction,

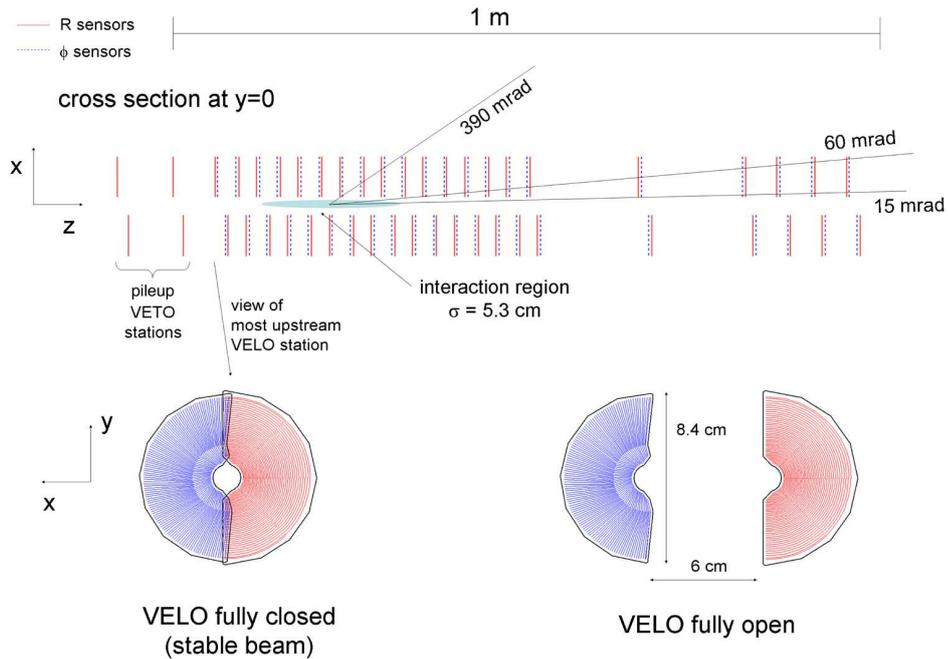


Figure 2.4: Schematic overview of the LHCb Vertex Locator [4].

where each of these modules consists of a set of two sensors, one measuring the r and one the ϕ coordinate of the hits (in a cylindrical coordinate system). Thus, each module provides precise spatial information for each hit.

During the LHC run I data taking period (2010-2012) the VELO showed excellent performance [6]. At the beginning of data taking, the sensor had a signal-to-noise ratio of about 20:1, depending on the individual strip capacitance. The single hit resolution varied between $4 \mu\text{m}$ and $20 \mu\text{m}$ for the optimum track angle of $7 - 11^\circ$. Fewer than 1% of the strips used in the VELO have been faulty during this time. Using the tag-and-probe technique with J/ψ decays, the track reconstruction efficiency of the VELO has been found to typically be above 98%, depending on the respective track kinematics [7].

Important performance parameters relying on the previously mentioned values are

the resolution of the vertex positions and the impact parameter (shortest distance) of tracks with respect to these vertices, and the decay time resolution of particles decaying within the VELO. The measured vertex resolution is heavily dependent on the number of tracks used to reconstruct the respective vertex, where for 25 tracks values of 13 μm and 71 μm are achieved in the transverse and horizontal plane, respectively. The resolution of the impact parameter is 12 μm for high momentum tracks, while, for tracks with low momentum, multiple scattering effects have a significant impact and lower it to 35 μm . The decay time resolution is derived by measuring $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decays, yielding a result of about 50 fs.

2.2.2 Ring-Imaging Cherenkov Detectors

According to the Cherenkov effect, charged particles passing through a medium at a velocity faster than the speed of light within this medium emit photons. The angle under which these photons are emitted depends on the particle velocity and the refractive index of the medium traversed. The two Ring-Imaging Cherenkov Detectors (RICH 1 and 2) used in the LHCb detector use a system of mirrors to focus the emitted photons onto a plane of photo detectors, where they appear as a circle with a radius depending on the Cherenkov angle and thus the velocity of the particle of origin. Figure 2.5 (a) shows a schematic overview of the RICH 1 detector (the RICH 2 uses a similar design). In combination with momentum information from the tracking systems, a hypothesis for the particle species can be provided. Figure 2.5 (b) shows the clear separation of protons, kaons and pions over a large range of particle momenta. The muon and pion distributions are not as clearly separated, but in combination with the muon system, the muons can be clearly identified.

The RICH system, combined with calorimeter information, provides information about the particle species in the form of difference in logarithmic likelihood (DLL) variables, comparing the likelihood of the particle being a proton, kaon or muon to being a pion. These are used by the LHCb trigger system as well as additional selection steps performed in later stages of the processing.

2.2.3 Spectrometry Stations

The LHCb tracking system consists of the VELO (see Sec. 2.2.1) and the TT upstream of a dipole magnet, and the Tracking Stations downstream of the magnet. Combining the information of track trajectories before and after them being bent by the magnetic field allows for a determination of the momentum of the charged particle tracks.

The dipole magnet has a integrated field strength of $\int B dl = 4\text{Tm}$, bending charged particle trajectories in the horizontal plane. During data taking the polarisation can be inverted, which allows for the cancellation of systematic uncertainties related to asymmetries in the detector construction.

The Tracker Turicensis is a small tracking station built from silicon strip sensors and positioned just before the magnet, providing additional information used for the

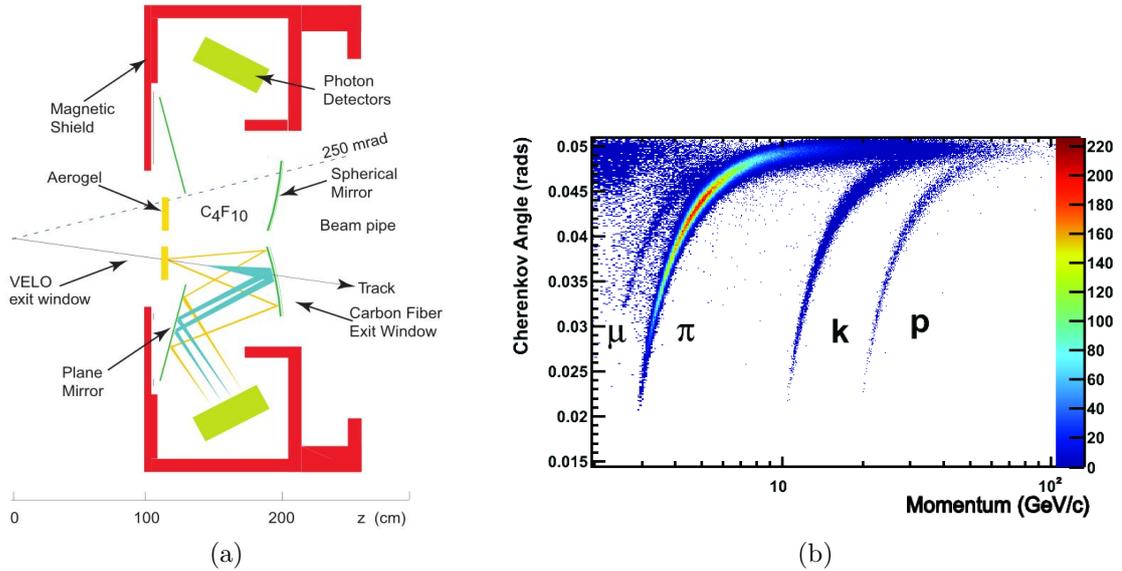


Figure 2.5: (a) Schematic overview of the RICH 1 detector; (b) Cherenkov angle - momentum relation for particle species measured by LHCb [4].

reconstruction of low momentum tracks which are bent out of the detector acceptance by the magnetic field. During the LHC Run I data taking period, it was not used in the reconstruction algorithms used for tracks that pass through both the tracking station and the VELO, so-called “Long” tracks.

The three Tracking Stations T1-T3 are placed after the magnet to provide a precise measurement of the bent particle trajectories. They are segmented into the Outer Tracker (OT) and the more granular Inner Tracker (IT) close to the beam pipe. Each OT station consists of four layers of the types X, U, and V. The strips or straws in the X-layers are oriented vertically, while the modules in the U- and V-layers are rotated by $\pm 5^\circ$ in the plane orthogonal to the beam direction with respect to the X-layers. Within each station the ordering of the layers is X-U-V-X, which results in an optimal resolution of the horizontal track direction (in which the magnet bends the tracks). The layout of the OT is shown in Fig. 2.6 (a). While the overall size of the OT layers is 6×5 m, each layer is split into two halves, called A- (access) and C-side (cryo), which are mounted on movable frames, allowing for easy access to the detector modules and electronics. These halves each consist of 4×9 straw tube modules with a length of 5 m, where each tube has a diameter of 5 mm. Figure 2.6 (b) shows the double-layer structure in which the tubes are arranged in the modules. Inside the tubes the signal is detected by gold-plated tungsten wires, acting both as a proportional counter as well as providing a drift time measurement. The position resolution achieved by these tubes is about $200 \mu\text{m}$, with a detection efficiency in the central area of more than 99.2%.

The Inner Tracker layers use the same technology as the TT, and has a dimension

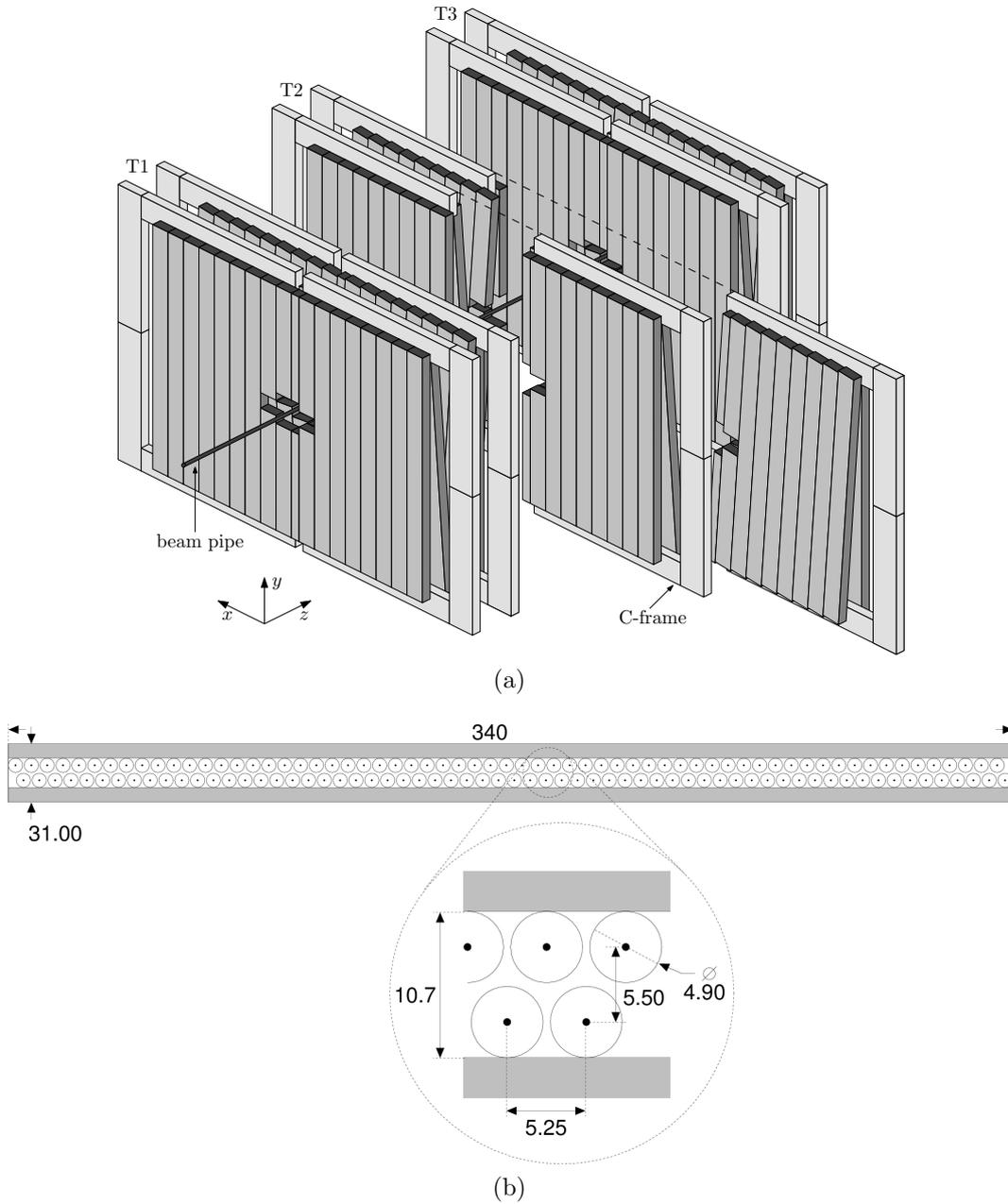


Figure 2.6: (a) Schematic overview of the LHCb Outer Tracker and (b) the positioning of the straw tubes [8].

of 126×40 cm. The layout of the IT is shown in Fig. 2.7. It is a silicon microstrip detector placed in the region closest to the beam pipe and offers both sufficient radiation hardness and a good spatial resolution. While achieving a detection

efficiency of about 99.8%, the measured single hit resolution is around $50\ \mu\text{m}$. The

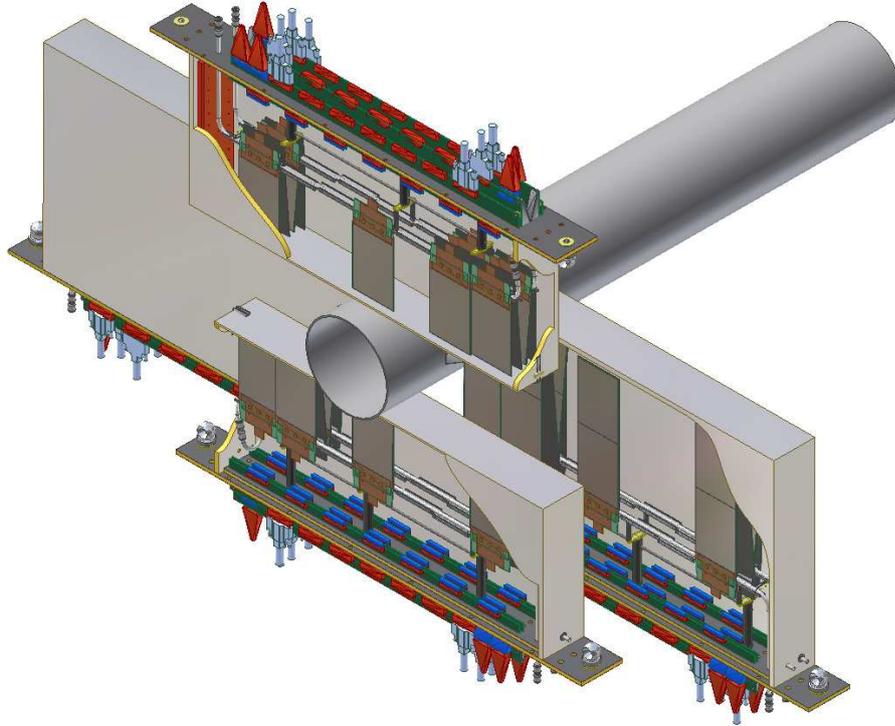


Figure 2.7: Three-dimensional sketch of the LHCb Inner Tracker [4].

IT stations are placed directly before the OT and covers the innermost region of the LHCb detector close to the beam pipe. In combination with the VELO and the TT the Tracking Stations provide an extremely efficient track reconstruction system with a minimal material budget. Efficiencies to reconstruct tracks within the LHCb acceptance have been measured to be above 96% for track momenta between 5 and 200 GeV/c [7].

2.2.4 The Muon System

Muons are typically minimal ionising particles, are due to their high mass barely affected by Bremsstrahlung, and do not interacting via the strong nuclear force. For those reasons, they pass most of the LHCb detector practically undisturbed. In order to detect muons and distinguish them from other charged particles the LHCb detector uses a system of five muon stations M1-M5. The first muon station is placed before the calorimeter, the other four downstream of it, with 80 cm thick iron absorber layers located between them. Each station is comprised of a multitude of multi-wire proportional chambers, while the innermost region of M1 uses GEMs (gas

electron multipliers), to sustain the levels of radiation in the region.

The muon system is used both for the identification of muons, by requiring hits in multiple muon stations, depending on the track kinematics, as well as in the L0 hardware trigger. It can be read out quickly and basic track trajectories are extracted, which are used to identify events with muon tracks originating from the interaction region. The first three stations M1-M3 have a high spatial resolution in the horizontal bending plane, as they are essential for the determination of the track direction, while M4 and M5 have a lower resolution, as they are mostly used to confirm the muon hypothesis for high momentum tracks penetrating all absorption layers. Muon tracks are reconstructed within the LHCb hardware trigger stage with an output event rate of 1 MHz, where the muon track segment is extrapolated to the primary interaction point by a simple p_T -dependent kink. This very simple and extremely fast momentum estimation is naturally less precise than the software-based track reconstruction. During the LHC run I data taking period a transverse momentum resolution of these muon tracks was measured to be 20%, with a muon detection efficiency of more than 99% [9].

As typically only muons can penetrate the absorption layers, the `IsMuon` flag is assigned to tracks with hits in the muon system depending on the measured track momentum: for low track momenta between 3 GeV/ c and 6 GeV/ c at least one hit in both M2 and M3 is sufficient, while for a momentum between 6 GeV/ c and 10 GeV/ c at least one additional hit in M4 or M5 is required. For high-momentum tracks with $p > 10$ GeV/ c at least one hit in each of M2, M3, M4 and M5 is needed to assign the `IsMuon` flag. It is used as an additional muon identification criterion by selections in the software trigger and offline.

2.2.5 Hardware and Software Trigger

The LHCb trigger system consists of three stages, the L0 hardware trigger and the two high-level trigger stages HLT1 and HLT2. The full trigger scheme as used in LHC run I is shown in Fig. 2.9.

Hardware and software trigger decisions are based on so-called “trigger lines”, which are a sequence of reconstruction algorithms and selection steps [10]. Events are flagged depending on which trigger line they have activated. For an event to be accepted by the trigger, it has to pass at least one trigger line in each stage.

The L0 trigger relies on information from the calorimeters and muon system to detect either energy deposits corresponding to what is expected from electrons, photons, or hadrons (L0Hadron), or high transverse momentum tracks corresponding to muons (L0(Di)Muon) originating from the interaction region. In the case of successfully passing the trigger requirements, a full detector readout is activated, reducing the readout rate to about 1 MHz from the bunch crossing rate of about 40 MHz.

The events passing the L0 trigger stage are sent to a computing cluster with about 26000 processing cores. While the first stage of the software based HLT has the full detector readout available, being operated at a rate of about 1 MHz limits the

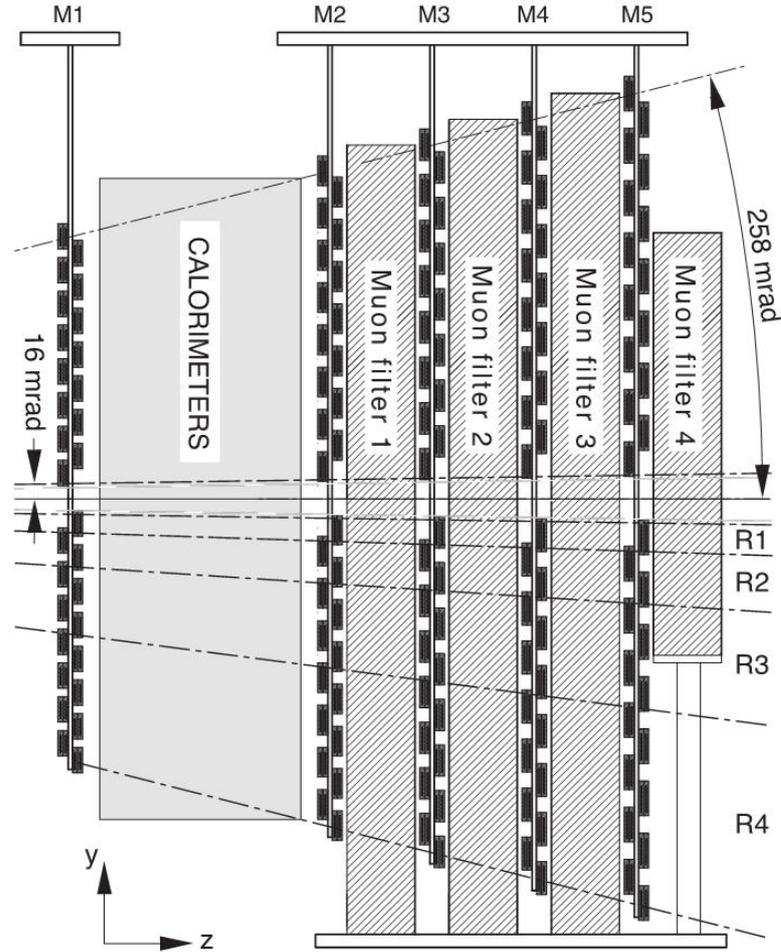


Figure 2.8: Schematic overview of the Muon Stations [4].

reconstruction detail that can be achieved. Tracks in the VELO are reconstructed and primary vertices created from at least five of these tracks. Using the pointing information from the VELO tracks with minimum momentum and transversal momentum requirements, hits are searched in the IT and OT to reconstruct Long tracks. Compared to the nominal reconstruction used offline, the material description is using a simplified model to save computing time. The HLT1 trigger lines select events where tracks with a good reconstruction quality have been found, which include dedicated lines searching for muon tracks. In addition, there are multiple lines active selecting events for calibration purposes. The output event rate is reduced from about 1 MHz to below 50 kHz during LHC Run I, and about 150 kHz during Run II by the HLT1 stage.

The lowered event rate allows for nearly offline quality reconstruction of the event

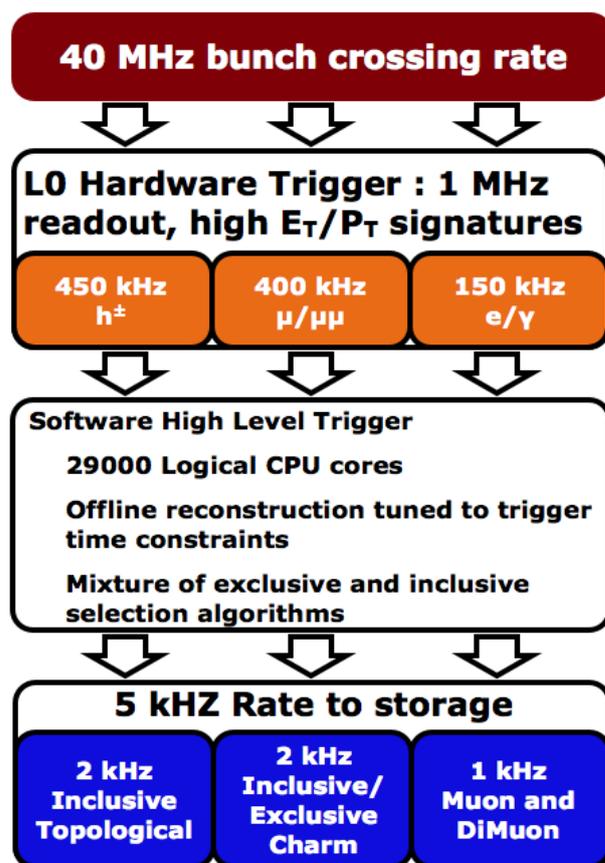


Figure 2.9: Trigger scheme of the LHCb experiment during LHC Run I [4].

in the second HLT stage, while still certain time constraints have to be taken into consideration. The reconstruction for tracks leaving hits both in the VELO and the Tracking Stations only uses one algorithm, starting with VELO tracks and searching for hits in the IT/OT, while offline an additional algorithm using standalone T station tracks and searching for matching VELO tracks is run. For that reason, the track reconstruction efficiency is 1-2% lower compared to the reconstruction run offline. Additionally, selection cuts applied to the momentum and transversal momentum are tighter than the cuts used offline. As well, the RICH particle identification algorithms are run in a simplified form, in order to provide muon, electron and photon identification.

Trigger lines used in HLT2 are mostly specifically searching for decays important for the LHCb physics programme. About 2 kHz event rate are provided by so-called topological trigger lines, searching for events with decay topologies as expected from B hadron decays to 2, 3, or 4 final-state particles, typically used for analyses with not fully reconstructible final states due to non-detectable particles. These topological trigger lines are available in two versions, with cut-based or multivariate

selection techniques. Another 2 kHz of the HLT2 output rate are occupied by trigger lines selecting exclusive decays, using the full reconstruction of specific decays. Due to the full reconstruction of the final state, these lines have typically very high selection efficiencies. The remaining 1 kHz of the HLT2 output rate is used by muon and dimuon trigger lines, which are passed by events with one or two muon tracks with high reconstruction quality. The dimuon lines search in addition for decays of well-reconstructed dimuon resonances, *e.g.* $J/\psi \rightarrow \mu^+ \mu^-$ decays, which occur in many decay modes important for the LHCb physics programme. The rates per trigger line category given refer to the trigger configuration during LHC Run I, while for Run II they are in general larger, but still subject to tuning.

The output rate of the HLT2 in LHC Run I was about 5 kHz (Run II: 12.5 kHz), and all events passing it were saved to disk for further offline processing.

2.3 The Stripping framework

All events saved to disk after passing the LHCb trigger requirements are processed by the offline reconstruction stage and subsequently by the “Stripping”. The Stripping framework is a centralised selection stage, which is used to reconstruct decay chains from the final-state particles and select candidates matching certain decay topologies. Its structure is similar to that of the trigger, where the selection algorithms are divided into an array of Stripping lines. Events passing the requirements of a Stripping line are saved to disk, including all particles of the full decay chain, reconstructed decay and primary interaction vertices, and track properties. This allows analysts to quickly access these datasets without the need of repeating those computationally intensive processing steps. The datasets saved are comparably small, as only part of the raw detector information is saved, depending on the respective Stripping line.

2.4 LHCb simulation

Simulated events are a helpful tool to estimate the efficiencies of reconstruction and selection processes, and test analysis procedures. They can be used to investigate typical observables for a large array of physics processes and decay modes and estimate the detector performance even before any real data has been collected. Furthermore, even hypothetical New Physics models can be tested and the selection optimised to improve sensitivity to these processes. For that reason, the main purpose of simulated event samples is to reproduce the real properties and conditions of specific particle decays as close as possible. This includes the simulation of the production and decay mechanisms, as well as the simulation of the response of a detector to the respective particles.

At the LHCb experiment, simulated event samples are created in several steps. The simulation of the initial proton-proton collision is performed using the Monte Carlo event generator PYTHIA [11]. The collision is generated by using parton distribution functions, and uses QCD models to describe the parton interaction. This includes

initial-state and final-state radiation and the hadronisation process of the initially created partons into hadronic particles. The decays and decay chains of the hadrons created from the PYTHIA step are simulated using the EVTGEN framework [12]. Different models can be used for these decays, including CP -violating parameters, angular dependencies, and correlations between the decay products. All parameters of initial and final-state particles generated that way are saved and can later be compared with the detector response. The response of the detector to these generator-level events is simulated using the GEANT4 package [13, 14], which includes a simulation of all hardware-based resolution effects, tuned to properly reflect the data-taking conditions. Those simulation steps result in detector-level information, which is further processed by a simulation of the hardware and software trigger, the default reconstruction algorithms, and the Stripping framework. This results in the same data structure as for real collision data, with additional information of the true generator-level properties.

2.5 Datasets

The measurements presented in this thesis use collision data collected by the LHCb experiment during LHC Run I and II.

For the determination of the track reconstruction efficiency a dataset collected during the 2015 data-taking period is used. It was recorded at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, and corresponds to an integrated luminosity of about 300 pb^{-1} .

The analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ makes use of the full dataset collected by the LHCb experiment during the 2011 and 2012 data-taking periods. During those two years, the LHC was operated at centre-of-mass energies of $\sqrt{s} = 7$ TeV and 8 TeV respectively. The datasets correspond to an integrated luminosity of about 1 fb^{-1} (2011) and 2 fb^{-1} (2012). Due to the low statistics available for the signal decay, the datasets collected in those two years are analysed simultaneously, in order to extract the full set of angular observables.

3 Tracking in the LHCb experiment

Charged particles interact with the material of the tracking detector they traverse and deposit a fraction of their energy. These interactions are saved as so-called “hits”, which represent spatial and timing information, by the detection system. Those hits are used by dedicated algorithms to reconstruct the trajectory of the particle through the detector. In presence of a magnetic field, the track trajectory is bent and this deflection can be used to measure the momentum of the respective particle. This chapter describes the general principles of the reconstruction of these trajectories, called “tracks”, and the track reconstruction algorithms used by the LHCb experiment.

3.1 Particle interaction with material

Sub-atomic particles can be measured through their interactions with other particles or detector material. All tracking detectors exploit these interactions to detect the position or even the energy loss of a particle traversing the respective detector layer. The following sections describe the most important effects that need to be considered when reconstructing the tracks of charged particles. These effects either lead to an energy loss or distort the flight path of the particle. The description given is based on the PDG review [15].

3.1.1 Ionisation

Charged particles may interact with the atoms of the materials they pass through by either exciting them into higher-energetic states or ionising them directly by removing electrons from the outer shell. The mean energy loss rate of a relativistic particle traversing a material layer of thickness x and density ρ can be described by the Bethe formula

$$-\frac{1}{\rho} \left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{m_e c^2} \cdot \frac{nz^2}{\beta^2} \cdot \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \cdot \left[\log \left(\frac{2m_e c^2 \beta^2}{I \cdot (1 - \beta^2)} \right) - \beta^2 \right], \quad (3.1)$$

where m_e and e are the mass and electric charge of the electron, z the charge of the traversing particle, n is the electron number density of the material, I is the mean excitation potential, ϵ_0 the vacuum permittivity, and $\beta = v/c$ the particle velocity relative to the speed of light. As this formula was derived using quantum mechanical perturbation theory at leading order, it only incorporates the particle charge up to the second order of z . The precision of the determination can be improved by including higher orders of z : the Barkas-Andersen-effect, proportional to z^3 , and the Bloch-correction, proportional to z^4 . In addition the non-stationary behaviour of the material electrons can be accounted for by using shell-corrections, and at very high energies density corrections, as calculated by Fermi, have to be taken into consideration.

The energy loss is minimal for particles with $\beta\gamma \approx 3.5$, where γ is the Lorentz boost factor. These particles are called minimal ionising particles (MIP), and often play an important role in the calibration of thresholds used by tracking detectors. The mean energy loss as a function of $\beta\gamma$ in different materials is shown in Fig. 3.1.

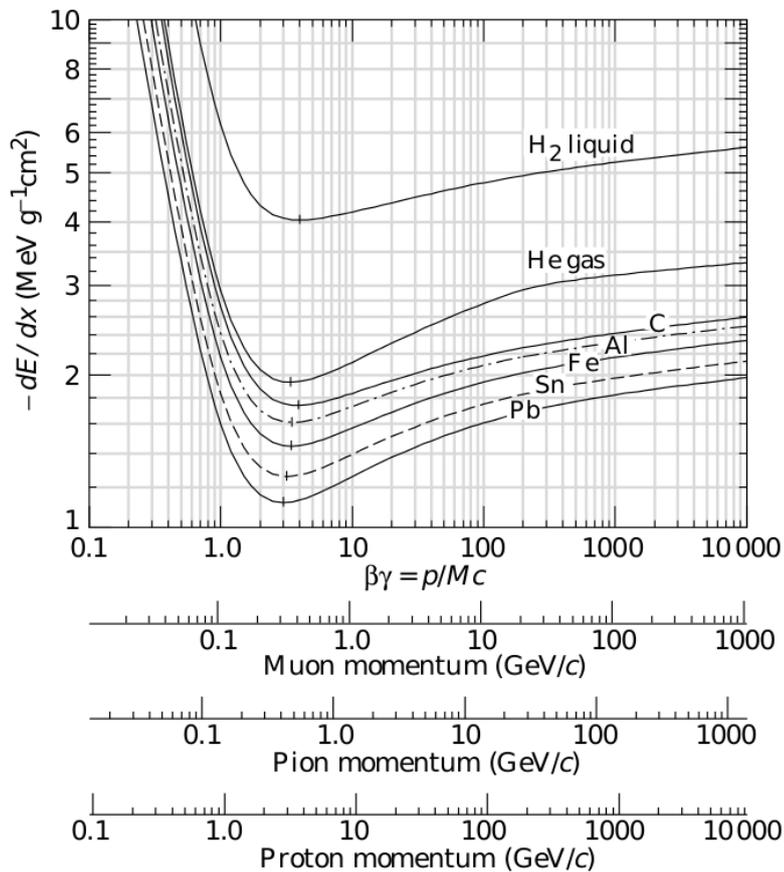


Figure 3.1: Mass stopping power $-\frac{1}{\rho} \left\langle \frac{dE}{dx} \right\rangle$ of charged particles in different materials as a function of $\beta\gamma$ [15]. Additional axes show the corresponding momenta for muons, pions, and kaons.

3.1.2 Bremsstrahlung

While low-energy electrons and positrons lose their energy mostly through ionisation, the effect of Bremsstrahlung becomes dominant at higher energies. Bremsstrahlung happens due to the electromagnetic interaction of the electron with the nuclei, where it loses energy while radiating off a photon. The energy loss can be characterised by the radiation length X_0 , which is the mean travelled distance at which the electron energy has been reduced to $1/e$ of its initial value. It is also equal to $7/9$ of the mean

free path for pair production of a high-energy photon. While Bremsstrahlung in principle occurs for all charged particles, the cross section of this process is suppressed by $(m_e/M)^2$ for particles of mass M . For that reason, Bremsstrahlung is negligible for all other species of reconstructed charged particles.

3.1.3 Coulomb scattering

Due to the interaction of a charged particle with the Coulomb field of the nuclei of the detector material, it is deflected from its original trajectory. This Coulomb (or Rutherford) scattering does not change the energy of the incoming particle significantly, but the distortion of the flight path has to be accounted for when reconstructing the trajectory, especially due to multiple-scattering effects.

3.1.4 Hadronic interactions

Hadrons passing the detector layers additionally interact with the nuclei of the detector material via the strong interaction. These hadronic interactions might lead to nuclei of the detector material or the hadron breaking up, and secondary hadrons being produced. These secondary showers are the main working principle of hadronic calorimeters, but are crucial to account for when reconstructing the tracks of hadrons, due to the stronger deflection compared to leptonic tracks. In addition, when the hadron itself is lost in the process before enough hits are left in the tracking system, no track can be reconstructed.

3.2 Properties of particle tracks

For the reconstruction of particle tracks multiple parameters are required to describe the trajectory through the detector. Besides the momentum at certain positions also the charge of the particle is needed, in order to evaluate the deflection through a magnetic field. The formalism to describe the track properties used in this section are largely taken from Ref. [16]. A track trajectory can be fully described by five parameters, with a certain freedom of choice. In the LHCb experiment those parameters at a given z-position are the x- and y-positions of the track, the slope parameters $t_x = \frac{dx}{dz}$ and $t_y = \frac{dy}{dz}$, and the particle charge relative to its momentum q/p . The choice of referring to the z-position comes naturally, as it is precisely known because the tracking stations are placed at specific z-positions within the LHCb detector. At any given position along the track trajectory a so-called “state” $\vec{x}(z)$, summarising position, momentum, and tangential direction, can be defined as

$$\vec{x}(z) = \begin{pmatrix} x \\ y \\ t_x \\ t_y \\ q/p \end{pmatrix}. \quad (3.2)$$

Along with the state itself, also the covariance matrix of all five state parameters is saved, to properly reflect the correlations of uncertainties. While the states could in principle be evaluated anywhere along the flight path of the particle, typically the states at the z-position of the tracking detectors are saved, as only there new information is added to the track. Each track state can be propagated to another z-position by using a track-propagation function f , via

$$\vec{x}(z_2) = f_{z_1 \rightarrow z_2} \cdot \vec{x}(z_1). \quad (3.3)$$

In the absence of a magnetic field, f becomes a trivial linear propagation matrix

$$f_{z_1 \rightarrow z_2} = \begin{pmatrix} 1 & 0 & z_2 - z_1 & 0 & 0 \\ 0 & 1 & 0 & z_2 - z_1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (3.4)$$

However, when a significant magnetic field is present, the propagation function can become highly non-linear, as the particle trajectory is bent. In the LHCb detector the magnetic field is strongest in the area of the dipole magnet between the TT and the T stations. But non-negligible fringe effects are also observable at the position of the TT and the first T station. Effects from multiple scattering through interactions of the particle with the detector material are taken into account by enlarging the uncertainties in the covariance matrix of the respective state. The non-trivial magnetic field parametrisation for the LHCb detector is measured, and used within the dedicated reconstruction algorithms.

3.3 Kalman filters used in track fits

The track fit is a procedure to extract the parameters of interest, *e.g.* position of origin, momentum, etc., from the various hit measurements provided by the different tracking systems. A method to perform the track fit which is very popular in high energy physics experiments is the Kalman filter [16]. It is both a very fast technique, as it does not require computationally intense matrix inversions, and can account for multiple scattering and energy-loss effects in a rather simple manner.

The Kalman filter is an iterative procedure that can be divided into three main steps:

- **Prediction:** From a given track state and with the known propagation function, the track state is propagated to the next measurement point.
- **Filtering:** Measurement information at that new point is used to update the propagated state. Filter equations are used to perform this update, and can be used to exclude extreme outliers in measurements from being added, helping in the pattern recognition.

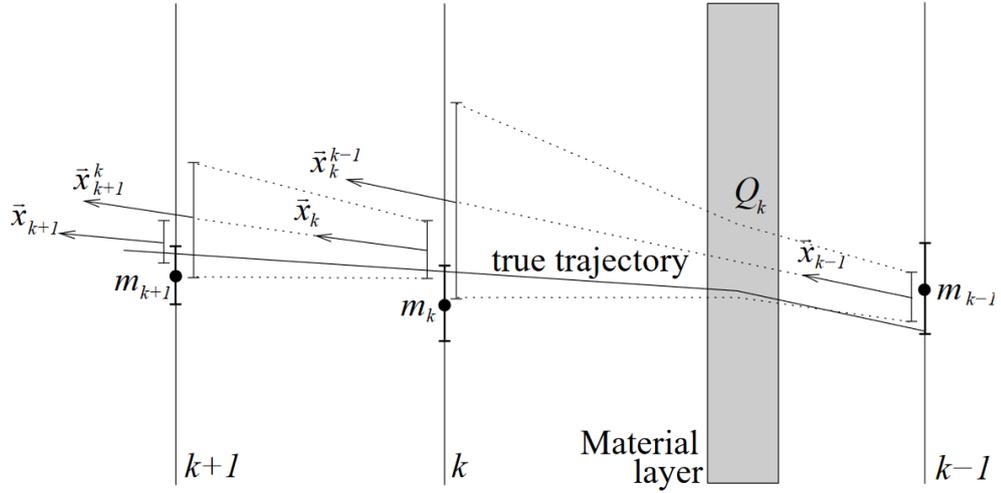


Figure 3.2: Schematic overview of the prediction and filtering steps used by a common Kalman filter [16]. Measurement points in the respective detector planes are indicated by m_i , the predicted track states by \bar{x}_j^{j+1} , and the updated track states after the filtering step by \bar{x}_j . The solid black line shows the true track trajectory which gets deflected by a material layer, indicated by the grey box. The material interaction leads to a larger search window in the following filtering step.

The prediction and filtering steps, illustrated in Fig. 3.2, are repeated until all measurements (*i.e.* hit information from each tracking detector) have been used. Afterwards, the third step is applied:

- **Smoothing:** The updated track states are used to perform the prediction and filtering steps to update the states at the previous measurement points in reverse order, starting from the last updated point. This step ensures that the states at each measurement points use all the information available from the track measurement.

Material interactions are taken into account by enlarging the search window when filtering at the new measurement point.

3.4 Track reconstruction algorithms in LHCb

The LHCb experiment uses a large number of tracking detectors, as described in Chap. 2. The optimal precision in the track reconstruction can be reached for particles that leave tracks in the VELO, close to the interaction point, and in the tracking stations behind the magnet. Additional information from the TT just before the magnet can be used to improve the accuracy of the momentum determination, as fringe-field effects can be measured. However, TT hit information is not needed

for the nominal track reconstruction algorithm, but purely optional. For muons, additional hit information from the muon station further improves the precision of the measurement of the track momentum. However, not all particles leave hits in each of the tracking detectors. Long-lived particles such as K_s^0 mesons often decay after leaving the VELO, causing the particles originating from their decay to not leave any hits in this tracking station. Low-momentum tracks are often bent out of the detector acceptance by the magnetic field and thus do not leave any hit information in the T stations. For that reason, the LHCb experiment distinguishes a number of different track types and algorithms dedicated to reconstruct them, where the most commonly used are shown schematically in Fig. 3.3.

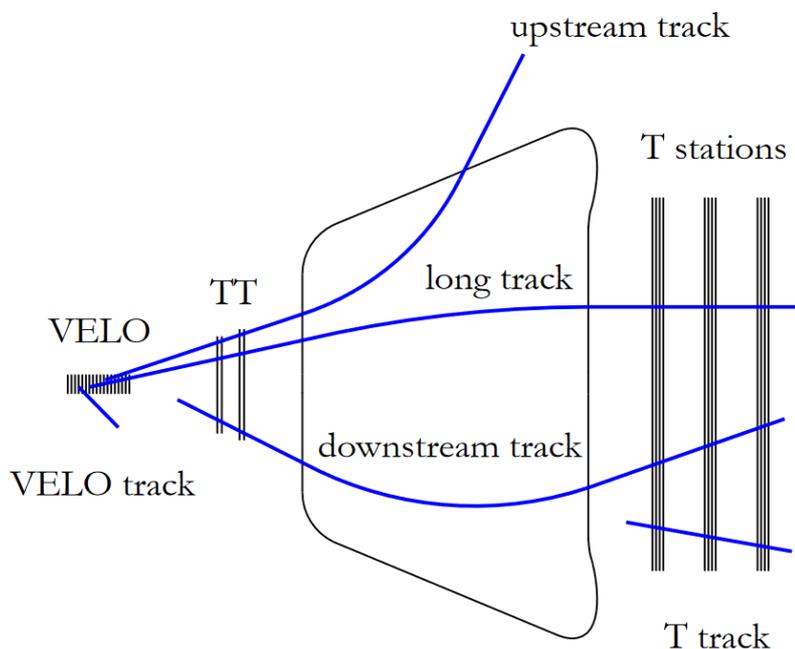


Figure 3.3: Schematic overview of the track types used by the LHCb experiment [17]. Details are given in the full text.

- **Long tracks:** The most common and important track type used by the LHCb reconstruction. They originate from a position within the VELO and traverse the detector up to at least the T stations, leaving hits in both sub-detectors. Hit information from the TT is added in some cases, but is not required for the Long track reconstruction. The Long tracks are created by two independent algorithms. In the “forward tracking” algorithm, first a VELO track is reconstructed, and the trajectory is propagated to the T stations. As the magnetic field does not extend into the VELO, the momentum information is determined

by combining the VELO track with each of the hits found in the first x-layer of the T stations, and the following layers searched for hits compatible with the trajectory of the updated track. For this propagation the bending induced by the LHCb magnet is simplified as a singular kink, similar to the typical approximation used for the optical bending induced by a lens.

The second track reconstruction algorithm used for Long track is called “Matching”. In addition to the VELO tracks also standalone tracks are created from the T station hits. All of the T-station tracks are then combined with each of the VELO tracks by extrapolating them to the magnet bending plane and comparing their compatibility. For both these algorithms hits from the TT are added only after finding a candidate from the VELO and T station hits.

- **Downstream tracks:** This track type is used for tracks originating from positions outside of the VELO. Typically these are tracks associated with particles originating from the decay of long-lived particles. Hit information from the TT and the T stations is used to reconstruct these tracks. However, as no VELO hits are available (causing the lever arm to be shorter), the momentum resolution for these tracks is in general worse compared to the Long tracks.
- **Upstream tracks:** Low momentum particles are usually bent out of the detector acceptance by the magnetic field before reaching the T stations. They are reconstructed as Upstream tracks, using the hit information from the VELO and TT only. These track types are used as part of the particle identification algorithms of the RICH 1.
- **T tracks:** These tracks are created using only hit information from the T stations. When they cannot be combined with VELO tracks via the Matching algorithm, they typically originate from very long-lived particles that have no associated hits in the TT. They are used by the particle identification algorithms of the RICH 2 and for the alignment of the T stations.
- **VELO tracks:** Similar to standalone T tracks, leftover VELO tracks could not be upgraded to Long tracks by the respective algorithms described above. They are used for the reconstruction of primary interaction vertices.

3.5 Fake tracks

Random combinations of hits, which do not originate from the same particle, can lead to the creation of fake background tracks, so called “ghosts”. Those ghost tracks can be rejected by requiring a high track fit quality of the Kalman filter. However, tight requirements on the track reconstruction quality can also reject real particle tracks, lowering the track reconstruction efficiency. In the LHCb reconstruction software, a neural network is trained to remove ghost tracks, while minimising the impact

on the tracking efficiency, as described in Ref. [18]. This neural network calculates a “ghost probability” value between 0 and 1, which is required to be below 0.4 for tracks in the nominal reconstruction. This requirement removes more than 70% of the originally created ghost tracks.

4 Determination of the track reconstruction efficiency

The track reconstruction algorithms introduced in the previous chapter are subject to inefficiencies, both due to technical limitations as well as the requirements imposed to suppress fake tracks. In order to measure the efficiency of the most common track type of the LHCb experiment, the Long tracks, a data-driven method using the tag-and-probe technique was developed during LHC Run I [7]. This chapter gives an overview over the tag-and-probe methods used and the results of the evaluation of the track reconstruction efficiencies during Run I.

4.1 Tag-and-probe technique with $J/\psi \rightarrow \mu^+ \mu^-$ decays

The tag-and-probe technique is a data-driven method that can be used to evaluate the efficiency for various processes, such as selection or reconstruction steps. For the measurement of the track reconstruction efficiency of the LHCb experiment, a sample of selected $J/\psi \rightarrow \mu^+ \mu^-$ candidates is used. The efficiency is evaluated for the most common track type used in LHCb, the Long tracks, which are reconstructed using information from the VELO and the T stations (see Sec. 3.4). For the other reconstructed track types there is no standardised procedure to determine the track reconstruction efficiencies. For analyses relying on those track types dedicated methods have to be used to determine the tracking efficiency, which will not be discussed in this thesis.

4.1.1 Tag-and-probe methods

The candidates used for the measurement are created by combining a Long track, referred to as “tag” track, and the “probe” signature. The tag track is reconstructed with the default reconstruction algorithms used for Long tracks, and required to pass a tight selection, compatible with a muon originating from the decay of a J/ψ meson. For the identification of the probe signature, one of three custom reconstruction algorithms is used, each using hit information from a different set of sub-detectors. The selection criteria for this signature are comparably loose to minimise a potential bias of the result. In order to test the efficiency of the default Long track reconstruction algorithms, the existence of a Long track which can be associated to the probe signature is investigated. This is performed by comparing the number of common hits of the respective Long track with the probe signature in sub-detectors specific to the respective method. The so-called “overlap fraction”, the number of common hits divided by the minimum of the number of hits of the probe signature or Long track in this sub-detector, is calculated and used as the association criterion. Corresponding to the probe signature reconstruction algorithm used, three methods are distinguished for the tracking efficiency measurement, illustrated in Fig. 4.1:

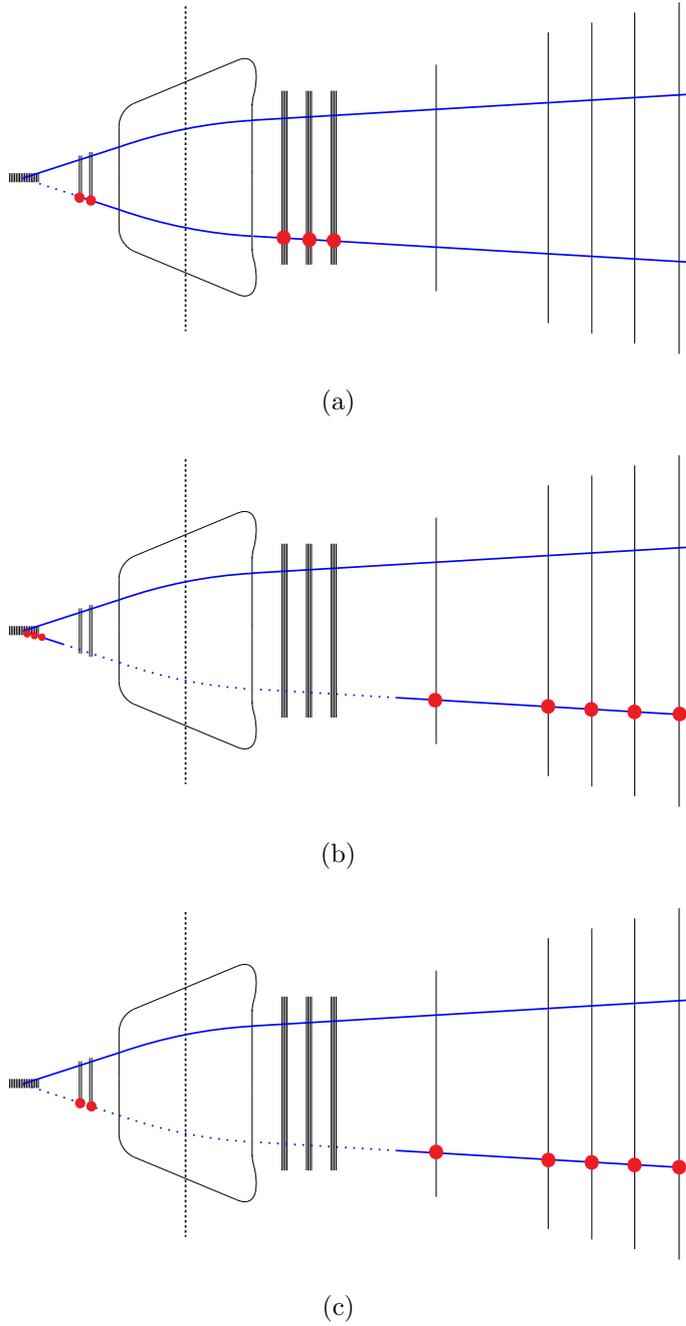


Figure 4.1: Illustration of the probe signature reconstruction algorithms of the (a) VELO method, (b) T-station method, and (c) Long method [7]. The hit information used by each algorithm to select the probe signature is indicated by red dots, while its trajectory is given by the lower blue line, dotted in the sub-detector region which is probed by the respective method. The upper blue line shows the trajectory of the tag track. The VELO is shown as the grey rectangle on the left, the TT and T stations as short and long vertical lines just before the magnet, respectively. The muon stations are indicated by the narrow vertical lines to the right. The shape of the dipole magnet is indicated between the TT and T stations, while the dashed vertical line reflects the bending plane of the magnet.

- **VELO method:** For this method the probe signatures are Downstream tracks, as described in Sec. 3.4, shown in Fig. 4.1(a). By demanding in addition a successful muon particle identification to be present for these tracks, hits in the muon station are added to them. As T-station hits are used by the Downstream track reconstruction algorithm, this method can only probe the efficiency of the VELO reconstruction of Long tracks. Candidates are considered as efficient, if a Long track is found with a minimal overlap fraction of 50% in the T stations compared to the probe signature.
- **T-station method:** This method is used complementary to the VELO method to probe exclusively the efficiency of the T-station reconstruction of Long tracks. Probe signatures for the T-station method are created by a dedicated reconstruction algorithm, for which hit information from the VELO and the muon stations is used, as sketched in Fig. 4.1(b). A candidate is considered as efficient when a Long track can be found with at least two common hits in the muon station and the same VELO segment as the probe signature.
- **Long method:** The probe signatures are reconstructed using hit information from the TT and the muon stations, which is illustrated in Fig. 4.1(c). First, the muon hits are used to create standalone muon tracks, which are subsequently matched to hits in the TT. As neither the muon stations nor hits in the TT are required to reconstruct Long tracks, this method probes directly the Long track reconstruction efficiency. A candidate is considered as efficient, when a minimum of overlapping hits is found in those two sub-detectors. An overlap fraction of 70% is required in the muon stations, while for the TT an overlapping hit fraction of 60% is sufficient. As the nominal Long track reconstruction algorithms do not require TT hits, these are added by searching for TT hits along the track trajectory after its full reconstruction.

The combination of the VELO and T-station efficiencies yields the efficiency to reconstruct Long tracks, which should result in similar values as obtained from the Long method.

During the LHC Run I data-taking period, the reconstruction of the probe signatures and selection of $J/\psi \rightarrow \mu^+ \mu^-$ candidates was performed after the data has been saved on disk and the centralised execution of the nominal LHCb reconstructions algorithms. Due to the large demand of computational power of the probe signature reconstruction algorithms for the T-station and Long methods they cannot be executed as part of the default LHCb reconstruction. Thus those custom reconstruction algorithms are only executed for events where a well-reconstructed muon track, detached from the primary vertex, was selected by the second stage of the software trigger. This single-muon trigger has due to the high abundance of high-momentum muons a very high output rate, and is for that reason configured to only record 50% of the events satisfying its requirements, randomly selected. This scaling reduces the available dataset. Furthermore, the probe signature reconstruction is performed within the centralised

selection framework of the LHCb experiment, the so-called “Stripping”. Even with the requirement of passing the trigger selection the computational requirements of the probe reconstruction is huge. Thus, it is only executed for a fraction of the events fulfilling the requirements, between 20% to 50% depending on the method.

4.1.2 Efficiency evaluation

The track reconstruction efficiency is dependent on kinematic properties of the track (*e.g.* momentum, direction) as well as the occupancy of the detector. In most analyses performed by the LHCb collaboration, the reconstruction efficiencies are evaluated using simulated event samples. While the simulation used by the LHCb experiment represents the collected data very well for most observables, there are residual differences: the distributions of the occupancies of the individual sub-detectors show discrepancies, where the mean of the distribution is shifted by up to 50% between data and simulated event samples. The tracking efficiency measurement using the tag-and-probe technique as described above is used to evaluate those efficiencies for data and simulation samples as a function of the track kinematics. The differences in occupancy are accounted for by applying weights to the simulated event sample to remove the discrepancies before evaluating the efficiencies.

The efficiency ε for each of the methods is defined by the number of $J/\psi \rightarrow \mu^+\mu^-$ decays where a Long track can be associated to the probe signature divided by the number of all $J/\psi \rightarrow \mu^+\mu^-$ decays in the sample where a probe signature was found:

$$\varepsilon = \frac{N_{\text{sig, assoc.}}}{N_{\text{sig}}}. \quad (4.1)$$

The dataset collected is affected by background from random combinations of real or fake muons, which do not originate from $J/\psi \rightarrow \mu^+\mu^-$ decays. For that reason, the signal yields are extracted from fits to the reconstructed invariant mass distributions of the J/ψ candidate, in order to distinguish real $J/\psi \rightarrow \mu^+\mu^-$ decays from this combinatorial background. The invariant mass of the J/ψ candidate is calculated from the four-momentum vectors of the tag track and probe signature via

$$m_{\text{rec}} = \sqrt{(E_{\text{tag}} + E_{\text{probe}})^2 - (\vec{p}_{\text{tag}} + \vec{p}_{\text{probe}})^2}, \quad (4.2)$$

where $E_{\text{tag/probe}}$ is the energy and $\vec{p}_{\text{tag/probe}}$ is the momentum vector of the tag track and probe signature, respectively. The reconstructed mass distributions combined from the tag track and the probe signatures selected by the three methods for the 2011 data-taking period are shown in Fig. 4.2. For each of the probe signature reconstruction algorithms, a different momentum resolution is achieved, resulting in different widths of the mass distributions. As the algorithms used for the probe signature reconstruction for the T-station and Long methods are not used in regular LHCb analyses, the worse resolution compared to Downstream or Long tracks is non-problematic. The sum of two Gaussian distributions with shared mean values is

chosen as the signal model, while the background model is given by an exponential function.

To evaluate the dependency of the efficiencies on the track kinematics and the

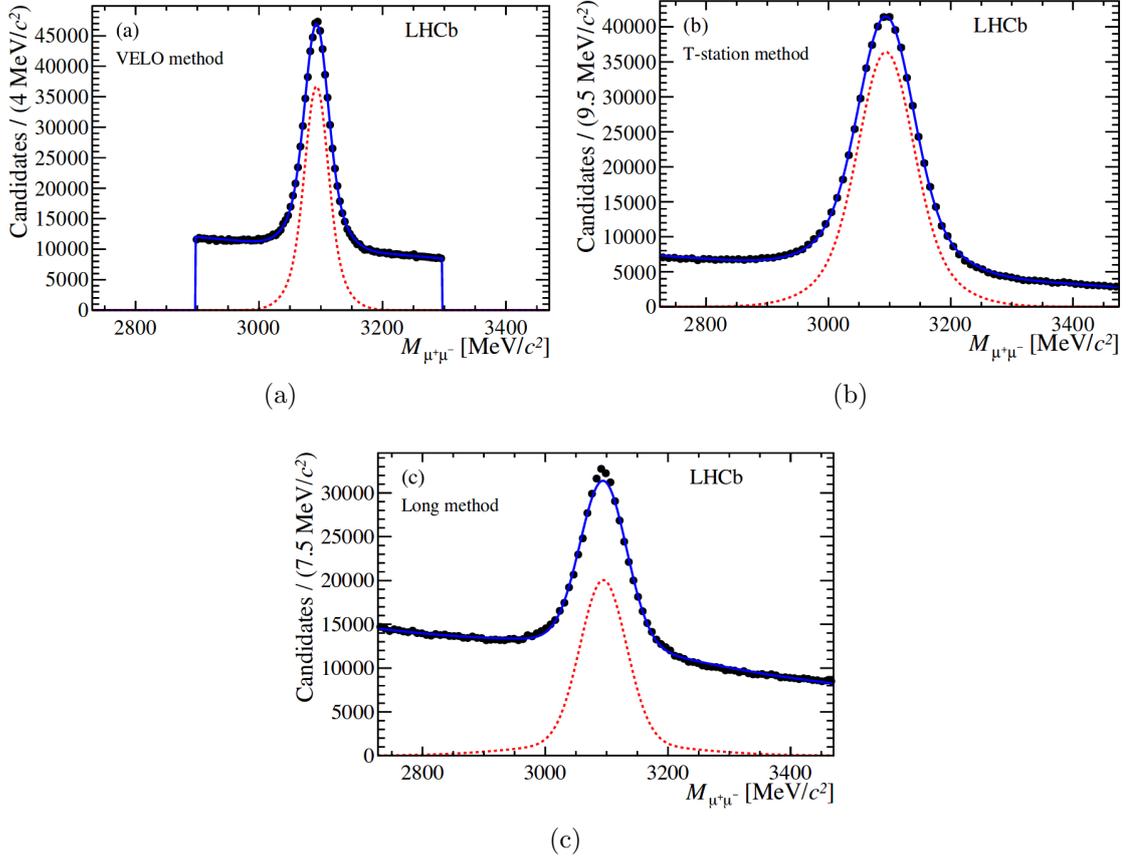


Figure 4.2: Invariant J/ψ candidate mass distributions overlaid with the fitted functions for the (a) VELO method, (b) T-station method, and (c) Long method for the dataset collected during the 2011 data-taking period [7]. Data is shown as black dots, the signal model is indicated by the red dotted line, and the full fit model by the solid blue line.

detector occupancy, these fits are performed in bins of the probe signature momentum p , pseudorapidity η , the number of tracks reconstructed for the respective event N_{track} , and the number of reconstructed primary interaction vertices N_{PV} . The efficiencies determined for the data and simulated event samples are compared in two-dimensional bins of probe signature momentum and pseudorapidity. A two-dimensional correction factor is given by the ratio R defined as

$$R(p, \eta) = \frac{\varepsilon_{\text{data}}(p, \eta)}{\varepsilon_{\text{sim}}(p, \eta)}, \quad (4.3)$$

where $\varepsilon_{\text{data}}$ and ε_{sim} are the track reconstruction efficiencies for the data and simulated event samples, respectively. The efficiencies from the VELO and T-station method are assumed to be uncorrelated except for the dependency on track kinematics and detector occupancy. Thus, the product of those efficiencies should properly reflect the Long track reconstruction efficiency. The difference to the Long method is that the latter requires the track to have hits in the TT, which is not necessary for Long tracks in general. However, those effects affect both data and simulated events and largely cancel in the ratio. The final value used to account for discrepancies between data and simulated events is the weighted average of the Long method and the combination of VELO and T-station method, defined via

$$R_{\text{final}} = \frac{1}{w_{\text{Comb}} + w_{\text{Long}}}(w_{\text{Comb}}R_{\text{Comb}} + w_{\text{Long}}R_{\text{Long}}), \quad (4.4)$$

where the weights $w_{\text{Comb}} = 1/\sigma_{R_{\text{Comb}}}^2$ and $w_{\text{Long}} = 1/\sigma_{R_{\text{Long}}}^2$ are the squares of the inverse uncertainty of the respective efficiency ratios. The statistical uncertainty of this weighted average is given by

$$\sigma_{\text{final}} = \frac{1}{\sqrt{w_{\text{Comb}} + w_{\text{Long}}}}. \quad (4.5)$$

The ratio is mapped onto a two-dimensional table in p and η and are made available to all analyses performed by the LHCb collaboration. Applying this correction factor as a per-track weight on the simulated event sample used in an analysis to determine the reconstruction efficiency for a given decay will correct for the discrepancies between data and simulation, and thus result in efficiencies properly reflecting the real dataset. The tag-and-probe method was validated on simulated event samples by comparing the derived efficiencies with efficiencies evaluated from the truth information. The track reconstruction efficiency is for simulated events typically defined as the fraction of simulated charged particles within the detector acceptance that can be associated to a reconstructed track. This association is performed by requiring a overlapping hit fraction of at least 70% in the VELO and T stations between the track and the respective simulated particle. The efficiencies obtained via the tag-and-probe methods and true efficiencies were found to agree within 1%.

4.2 Results of the LHC Run I data-taking period

The track reconstruction efficiencies for the 2010, 2011 and 2012 data-taking periods obtained from the tag-and-probe method have been published in Ref. [7]. Figure 4.3 shows the efficiencies for the combination of VELO and T-station method and of the Long method for the 2012 data-taking period. The efficiencies are given as functions of the probe signature momentum and pseudorapidity, and the number of tracks and primary interaction vertices reconstructed per event. Integrated over the full range of the probe signature kinematics, the efficiencies determined for data and simulated events agree well. However, when observing the dependency on the probe

signature momentum, significant differences are observed, showing the importance of the application of the correction for simulated events.

The only source of systematic uncertainty found to be significant for the evaluation of the track reconstruction efficiencies of the LHCb experiment during LHC Run I was the choice of the occupancy variable, in which the simulated events are weighted. This effect was evaluated by using different occupancy variables, *e.g.* the number of reconstructed primary vertices, and the resulting efficiencies compared to the nominal result, where the largest deviation of 0.4% was assigned as systematic uncertainty. Further sources of systematic uncertainties were investigated by determining the stability of the ratios when changing the fit models, and the slight difference of the efficiencies between Long method and the combination of the VELO and T-station methods. Both were found to be negligible compared to the dependency on the choice of the occupancy variable used.

As described in the previous chapter, hadrons are affected by hadronic interactions with the detector material. However, the track reconstruction efficiencies are determined by using tracks originating from muons. Studies using simulated events were used to evaluate the additional effects due to material interactions on the efficiencies for hadrons compared to muons. The additional fraction of hadronic tracks not reaching the T stations due to material interactions was found to be 11% for kaons and 14% for pions. Using the conservative estimate of 10% for the uncertainty of the material budget of the detector results in additional systematic uncertainties of 1.1% and 1.4% for kaons and pions, respectively. These uncertainties have to be taken into account when applying the corrections on simulated event samples.

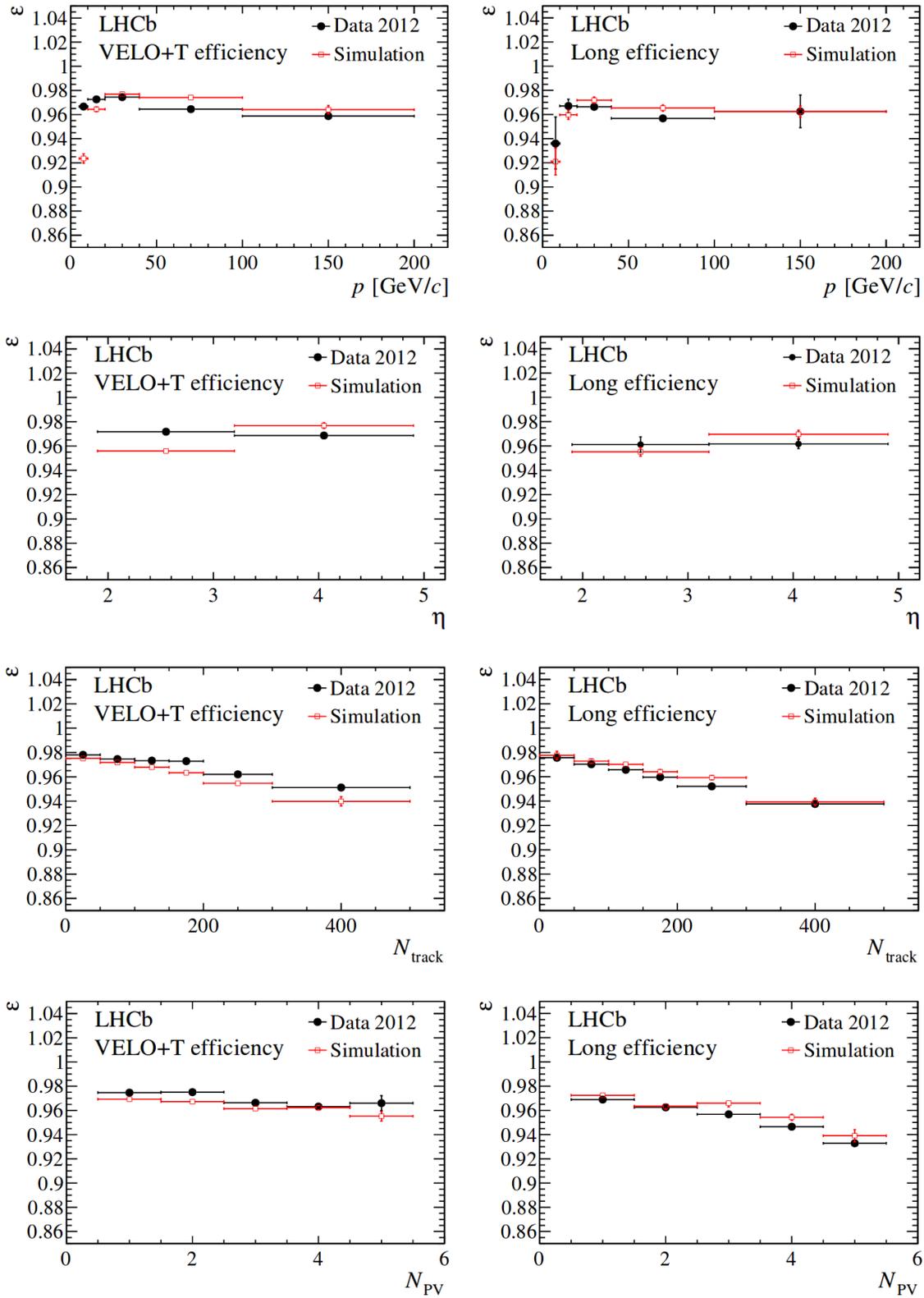


Figure 4.3: Track reconstruction efficiency for the 2012 data-taking period as function of the probe signature momentum and pseudorapidity, and the number of tracks and primary interaction vertices reconstructed per event [7]. Efficiencies measured on data are shown as solid black dots, those measured on simulated events as red squares. Uncertainties indicated are statistical only.

5 Tracking efficiency measurement in Run II

The measurement of the track reconstruction efficiency of the LHCb experiment during LHC Run I proved to be an essential component in many measurements performed during that data-taking period. However, the precision of the correction provided was limited by the size of the dataset collected. This was caused by the requirement of using events recorded with the unbiased single-muon software trigger, and the additional necessity to match the bandwidth constraints of the Stripping framework.

For the LHC Run II data-taking period many changes were implemented to the online setup of the LHCb experiment. During Run I, only a simplified track reconstruction could be performed within the software trigger due to the time constraints. Additionally, in order to achieve the best possible reconstruction quality, the detector has to be aligned and calibrated, using the recorded data.

However, for the data taking in Run II a real-time alignment and calibration of the LHCb detector was implemented. This, combined with improvements in the timing requirements of the reconstruction algorithms, enables the full track reconstruction being available in the second stage of the software trigger. For the first software trigger stage a simplified reconstruction is executed, as it has to be operated at an input event rate of 1 MHz. The full reconstruction available in HLT2 allows for analyses to be performed using the data as it is recorded without the need of the additional centralised offline reconstruction. For the purpose of exploiting the offline-quality reconstruction, the “Turbo” stream was implemented [19]. The reconstructed objects which fulfil trigger requirements and all related information can be saved and used as part of physics analyses, without the need of an additional offline reconstruction. This allows for a fast evaluation of data directly after it has been recorded. However, the measurement of the detector performance in this “quasi-online” framework has to be performed in the same manner to yield quickly available results.

In order to investigate the track reconstruction efficiencies obtained in the online reconstruction, dedicated trigger decisions have been created, which execute the probe signature reconstruction within the software trigger, and save the created tracks and J/ψ candidates for further studies. In addition, the search for efficient Long tracks corresponding to the reconstructed probe signature is done by evaluating the fraction of common hits in the respective sub-detectors, for each method introduced in the previous chapter. This allows for a “quasi-online” evaluation of the track reconstruction efficiencies of the LHCb experiment and the monitoring of this efficiency during data taking.

5.1 Trigger line implementation

In order to evaluate the efficiency of the online track reconstruction used within the software trigger, the dedicated algorithms described in the previous chapter were moved into the online framework. However, to meet the stringent time constraints

on the execution of the software trigger stage, changes are made with respect to the implementation as used in Run I, described below. For each of the three methods introduced in the previous chapter, two trigger lines are created. One line uses a positively charged probe signature and negatively charged tag track, the other line uses the tracks with the charges interchanged. This is necessary as otherwise the selection logic could not distinguish between the tag track and the probe signature, which both are identified as muons.

The first step of each trigger decision is the requirement to find a tag track, which previously passed the requirements of the `Hlt1TrackMuon` trigger line in the first stage of the software trigger. This requires a well-reconstructed track with the `IsMuon` flag (corresponding to a momentum-dependent number of hits in the muon stations). The track needs a total momentum of larger than $3 \text{ GeV}/c$, a transverse momentum of larger than $800 \text{ MeV}/c$, and has to be significantly detached from the primary interaction vertex by requiring an impact parameter significance of larger than 8. In addition to having passed this trigger decision, further selection requirements are imposed on the tag track, depending on which method is used. These additional criteria are kept close to the optimised selection found for the Run I measurement, and are listed in Table 5.1. The main difference compared to the Run I implementation is that the selection criteria are applied to the tag track before executing the probe signature reconstruction, reducing the average computation time of the latter. Only when such a tightly selected tag track is found, the probe signature reconstruction used by the respective method is executed. This reduces the average computation time of each of the six trigger decisions used for the tracking efficiency measurement to below 1 ms, well within the time constraints of the software trigger. Additional selection criteria are applied to the reconstructed probe signatures, and the J/ψ combined from the tag and probe signatures. These requirements differ for each of the three methods, and are listed in Tables 5.2 and 5.3. Compared to the Run I measurement the kinematic requirements applied to the probe signatures have been unified to provide the identical coverage for all methods. The requirements imposed upon the J/ψ are used to ensure the quality of the vertex reconstruction from the two muon tracks and reduce combinatorial background. For the Long method a requirement on the maximum detachment of the J/ψ decay vertex from the primary interaction vertex proved to be helpful in reducing background. This is ensured by requiring a maximum value of the impact parameter (IP) to the primary vertex. For the VELO method, only tracks are combined to a J/ψ candidate if their distance of closest approach (DOCA) is smaller than 5 mm, to reduce the computation time required.

The `Tesla` framework [19], which was introduced for the LHCb experiment in LHC Run II, allows to save all information of tracks and multi-particle final states of candidates which pass certain trigger line requirements. However, many analyses need additional information from other parts of the event, *e.g.* for flavour tagging algorithms. Thus this framework is only used for trigger lines where only information

Table 5.1: Selection cuts applied to the tag tracks by the software trigger.

Variable	VELO method	T-station method	Long method
$DLL_{\mu\pi}$	> -2	> -1	> -2
p	$> 5 \text{ GeV}/c$	$> 7 \text{ GeV}/c$	$> 10 \text{ GeV}/c$
p_T	$> 700 \text{ MeV}/c$	–	$> 1300 \text{ MeV}/c$
track χ^2/ndf	< 10	< 3	< 5
IP	$> 0.5 \text{ mm}$	$> 0.2 \text{ mm}$	–

Table 5.2: Selection cuts applied to the probe signatures by the software trigger.

Variable	VELO method	T-station method	Long method
p	$> 5 \text{ GeV}/c$	$> 5 \text{ GeV}/c$	$> 5 \text{ GeV}/c$
p_T	$> 500 \text{ MeV}/c$	$> 500 \text{ MeV}/c$	$> 500 \text{ MeV}/c$
track χ^2/ndf	< 10	< 5	–

of an exclusive decay is of interest, as it is the case for the tracking efficiency lines. The reconstructed tracks which do not pass the requirements of a trigger line specified to be analysed in such a way, are lost after the software trigger. In order to be able to measure the track reconstruction efficiency of the algorithms as used within the software trigger, for each method and charge another trigger decision is created. This uses the same candidates of tag track, probe signature, and composite J/ψ , and searches for Long tracks with significant overlap of $> 40\%$ in the respective sub-detectors with the reconstructed probe signature. In case such a Long track is found, all its information is saved by the Tesla framework and can be used for further analysis.

 Table 5.3: Selection cuts applied to the J/ψ reconstructed from tag and probe signatures by the software trigger.

Variable	VELO method	T-station method	Long method
$ m_{\mu^+\mu^-} - m_{J/\psi} $	$< 200 \text{ MeV}/c^2$	$< 500 \text{ MeV}/c^2$	$< 500 \text{ MeV}/c^2$
p_T	–	$> 500 \text{ MeV}/c$	$> 1000 \text{ MeV}/c$
vertex χ^2	< 5	< 2	< 2
Track DOCA	$< 5 \text{ mm}$	–	–
IP	–	–	$< 0.8 \text{ mm}$

5.2 Tracking efficiency evaluation

The dataset collected by the trigger lines described above is analysed “quasi-online” with minimal computing steps directly after the data-taking process. Due to the use of the `Tesla` framework, no additional centralised processing by the LHCb reconstruction or selection frameworks is needed, allowing for a evaluation shortly after the data is recorded. This was especially important at the start of LHC Run II, where the LHCb experiment collected a small early dataset for the measurement of multiple production processes. For the evaluation of the track reconstruction efficiency for the LHCb simulation a sample of simulated $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) X$ decays is created, where the X can be any set of additional generated particles. The hardware and software trigger steps are simulated for this sample and it is processed by the `Tesla` framework, resulting in the same data structure as it is achieved for the real dataset. As the track reconstruction efficiencies are in general dependent not only on the track kinematics but also on the detector occupancy, weights are applied to correct the difference between data and simulated events. For the Run I analysis, the number of reconstructed tracks was the observable used for this correction. However, this variable depends on the reconstruction algorithms and may change with modifications of these. For that reason, for the measurement performed using Run II data, the number of hits in the Scintillating Pad Detector (SPD) of the LHCb detector is used. As this observable is tied directly to the detector response, it is a stable measure of the detector occupancy.

The datasets of real data and simulated events are divided in one-dimensional bins of probe signature momentum p and pseudorapidity η , and the number of SPD hits N_{SPD} and reconstructed primary interaction vertices N_{PV} . Besides the investigation of the tracking efficiencies and their dependencies on the track kinematics and detector occupancy, the purpose of the measurement is to provide weights to correct simulated event samples for discrepancies compared to collision data. In order to create these two-dimensional correction tables, further two-dimensional bins in p and η of the probe signature are created. For each bin, the dataset is split depending on whether an efficient Long track was found or not, which are further referred to as “Match” and “Fail” datasets, respectively. A simultaneous extended maximum likelihood fit (as described in detail in Sec. 14.1), is performed to the invariant J/ψ candidate mass distribution of the Match and Fail datasets in each bin. For the signal model, the sum of two Crystal-Ball functions [20] is used, where the mean and the tail parameters are shared between both functions. All shape parameters are shared between the Match and Fail datasets in the simultaneous fit, while the signal yields are independent. The background model is given by an exponential function, with independent parameters used to describe both Match and Fail datasets. From the signal yields the efficiency ε is treated as a fit parameter, connecting the two datasets to which the fit is performed via

$$\varepsilon = \frac{N_{\text{sig,Match}}}{N_{\text{sig,Match}} + N_{\text{sig,Fail}}}, \quad (5.1)$$

where $N_{\text{sig,Match}}$ and $N_{\text{sig,Fail}}$ are the signal yields of the Match and Fail dataset, respectively. Due to the implementation of the efficiency as a parameter of the simultaneous fit, correlations between other parameters are properly taken into account when calculating the statistical uncertainties. The fits to the full data samples of each method collected during the nominal 2015 data-taking period are shown in Fig. 5.1. Compared to the 2011 dataset, corresponding to an integrated luminosity of about 1 fb^{-1} , the size of the collected dataset has increased by a factor of ≈ 6.5 . While the dataset collected in 2015 corresponds to only about 300 pb^{-1} , the production cross section for b hadrons has approximately doubled [21]. Combining these numbers shows the enormous improvement compared to Run I, which results in a more than ten times higher collection rate.

For the creation of the two-dimensional correction tables the efficiencies derived from the VELO and T-station methods are multiplied in each $p - \eta$ bin to give the ‘‘Combined’’ method efficiency. The final correction ratio R_{final} is created by calculating the weighted average between Combined and Long method in each bin as defined in Eq. 4.4 in the previous chapter.

5.3 Method validation and results

Due to the changes in the implementation and selection requirements applied for the Run II measurement of the tracking efficiencies, the validation described in the previous chapter is repeated with the updated procedure. The efficiencies determined from the tag-and-probe approach on simulated events are compared to the true efficiencies, using the same binning schemes as the nominal measurement. For this comparison the sample of simulated $B^+ \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) X$ decays is used. Figure 5.2 shows the comparison between true and tag-and-probe efficiencies as function of probe signature momentum and pseudorapidity for each of the three methods for the simulated event sample corresponding to the conditions of 2015 data-taking period. Excellent agreement between the two efficiency definitions is achieved, significant deviations are only observed at high pseudorapidity for the T-station and Long method. However, small discrepancies are expected to cancel between simulated events and collision data, as they are inherent to the tag-and-probe method.

The efficiencies of collision data and simulated events are shown in Figs. 5.3 to 5.6 as function of the observables mentioned above. The track reconstruction efficiencies are slightly higher for simulated events compared to data at low probe signature momentum and pseudorapidity, while at high pseudorapidity the opposite effect occurs. The ratio between the efficiencies obtained for data and simulated events in two-dimensional bins of p and η are shown in Fig. 5.7. The ratio for the Final method is made available as a correction table for physics analysis performed by the LHCb collaboration. For the phase space at low momentum and high pseudorapidity, and high momentum and low pseudorapidity only a very small amount of signal candidates can be found and no converging fit can be performed. Thus, no correction factor is given in those regions and a systematic uncertainty of 5% is assigned as a

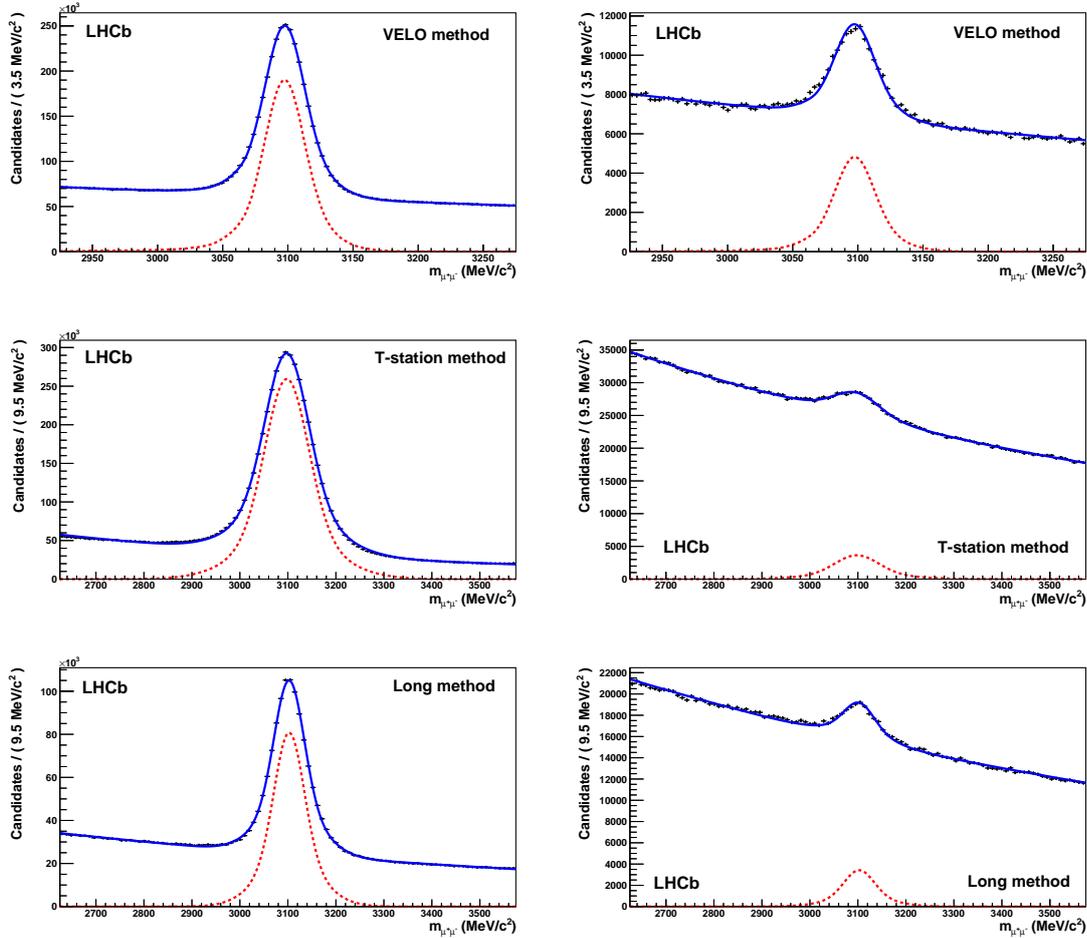


Figure 5.1: Fits to the data samples collected during the LHCb 2015 data-taking period for the three tag-and-probe methods, the Match samples are shown to left, the Fail samples to the right. Data is presented by the black points, the fitted signal model by the dashed red line, and the full fit model by the solid blue line.

conservative estimate to simulated tracks for this non-correction. Integrated over the full range of the probe signature kinematics, the Long track reconstruction efficiency for the 2015 data-taking period is measured as $(95.94 \pm 0.07)\%$ using the Long method, and $(96.18 \pm 0.03)\%$ using the combination of VELO method and T-station method. Combining these results using the weighted average defined in Eq. 4.4 yields a track reconstruction efficiency of $(96.14 \pm 0.03)\%$. The Long method and the combination of VELO and T-station methods agree well given the small differences between the efficiencies they probe, which were discussed in Sec. 4.1.2. The efficiency is about 1% lower than the efficiency measured during the 2012 data-taking period. The increased centre-of-mass energy typically causes a larger amount of particles

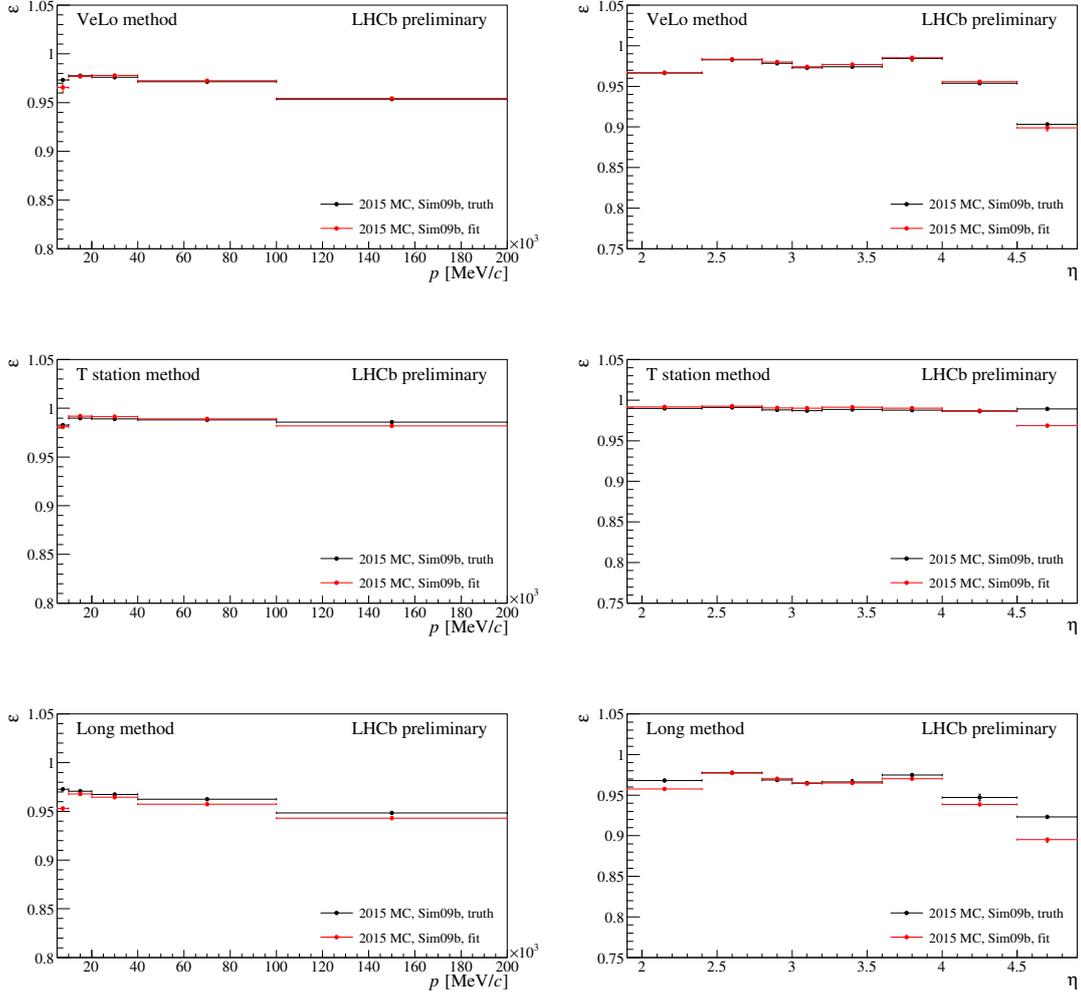


Figure 5.2: Comparison of the tracking efficiencies determined from the tag-and-probe approach (red) and truth information (black) for simulated events corresponding to the conditions of 2015 data-taking period as function of the probe signature momentum and pseudorapidity. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

originating from the primary interaction, which leads to a higher detector occupancy. This higher occupancy in turn is detrimental to the track reconstruction efficiencies, leading to slightly lower efficiencies in the Run II data taking.

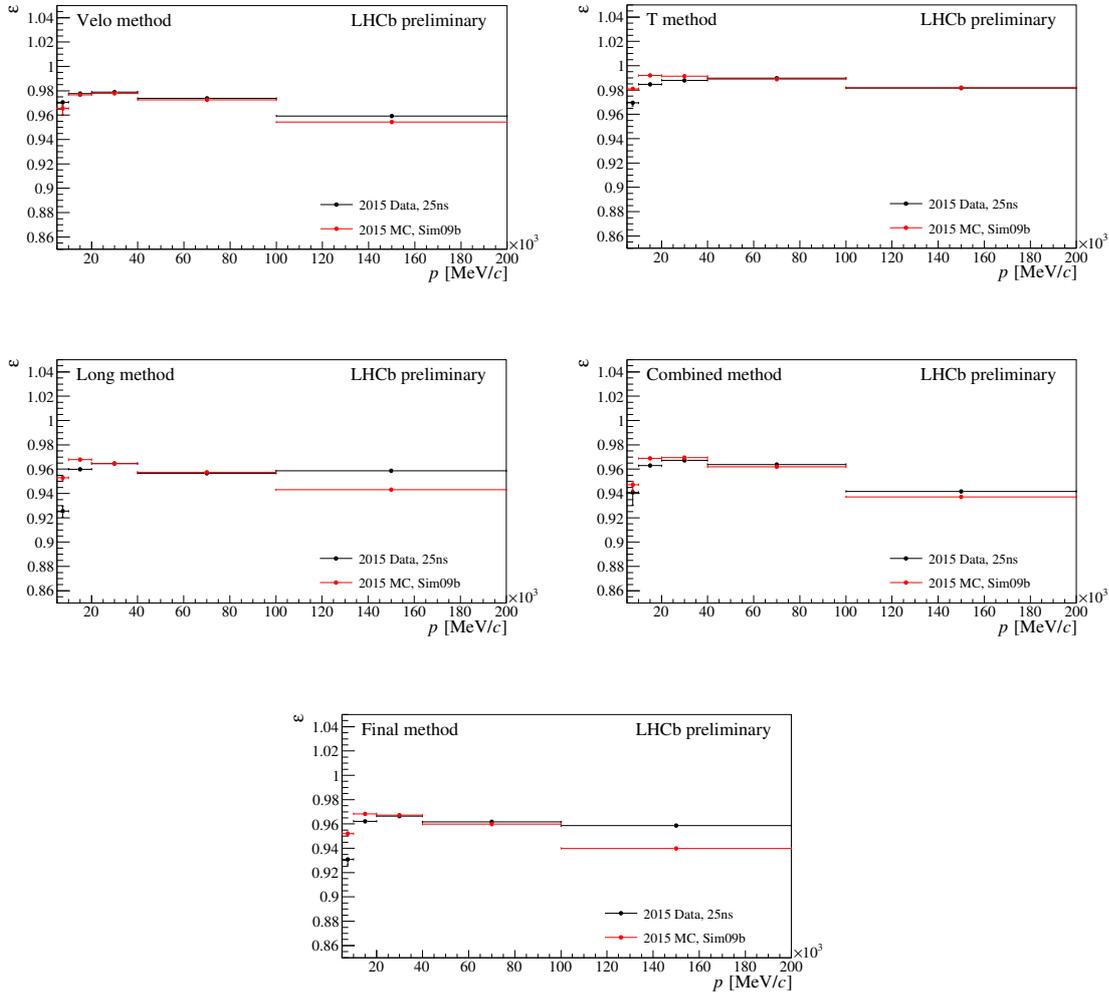


Figure 5.3: Comparison of track reconstruction efficiency for data (black) and simulated events (red) for the 2015 data-taking period as function of the probe signature momentum. The Combined method is the product of the VELO and T-station methods, the Final method the weighted average of the Long and Combined methods. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

5.4 Additional performance tests

5.4.1 Stability with different binning

The nominal ratios provided to the LHCb collaboration to correct the efficiencies determined from simulated event samples use two bins in the pseudorapidity and five bins in the momentum of the track. In this binning scheme no steep dependency

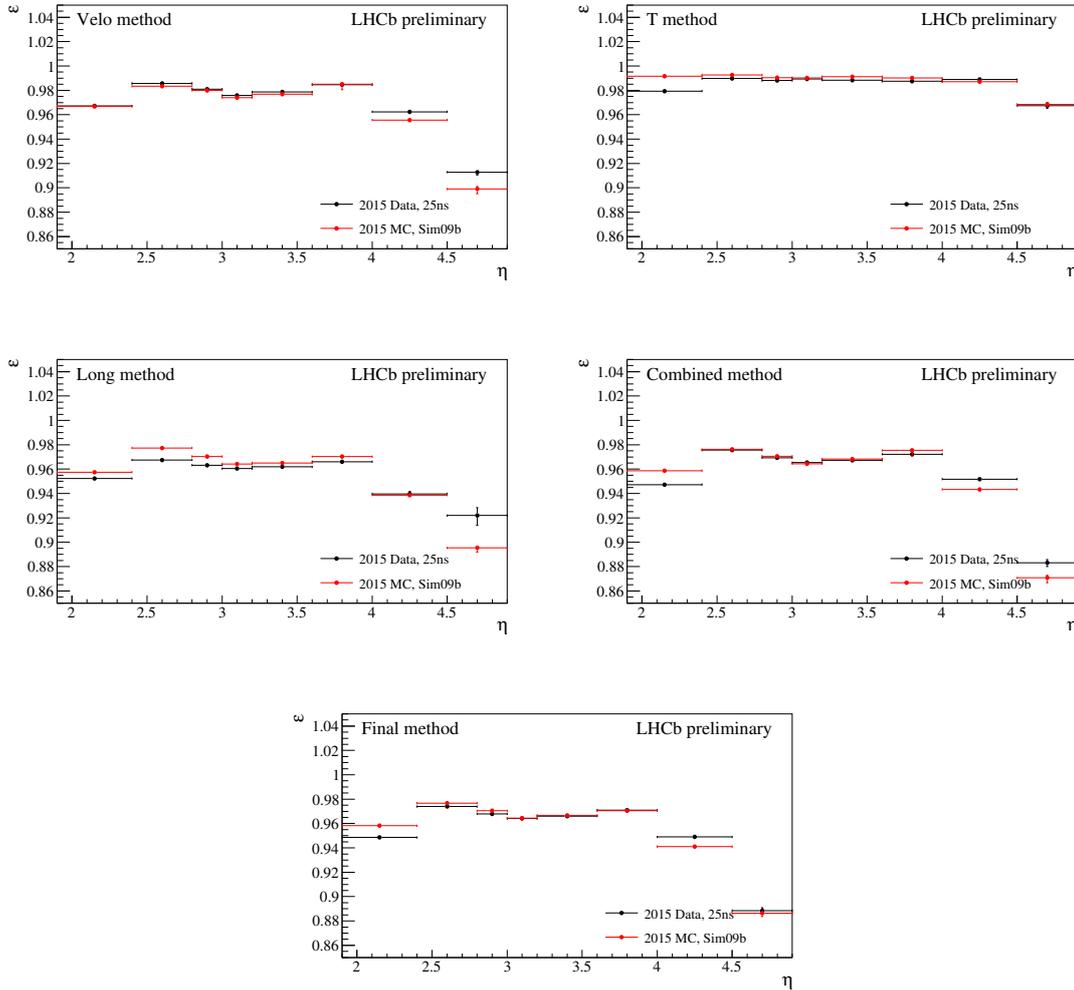


Figure 5.4: Comparison of track reconstruction efficiency for data (black) and simulated events (red) for the 2015 data-taking period as function of the probe signature pseudorapidity. The Combined method is the product of the VELO and T-station methods, the Final method the weighted average of the Long and Combined methods. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

of the ratio between the efficiencies determined for data and simulated events have been observed. However, to ensure the stability of this ratio within the respective bins, the evaluation of this ratio is repeated with a different binning scheme and compared to the nominal results. The number of bins in pseudorapidity is increased from two to eight, dividing the range of each of the nominal bins into four new bins. The resulting two-dimensional ratios are shown in Fig. 5.8. Within the range of

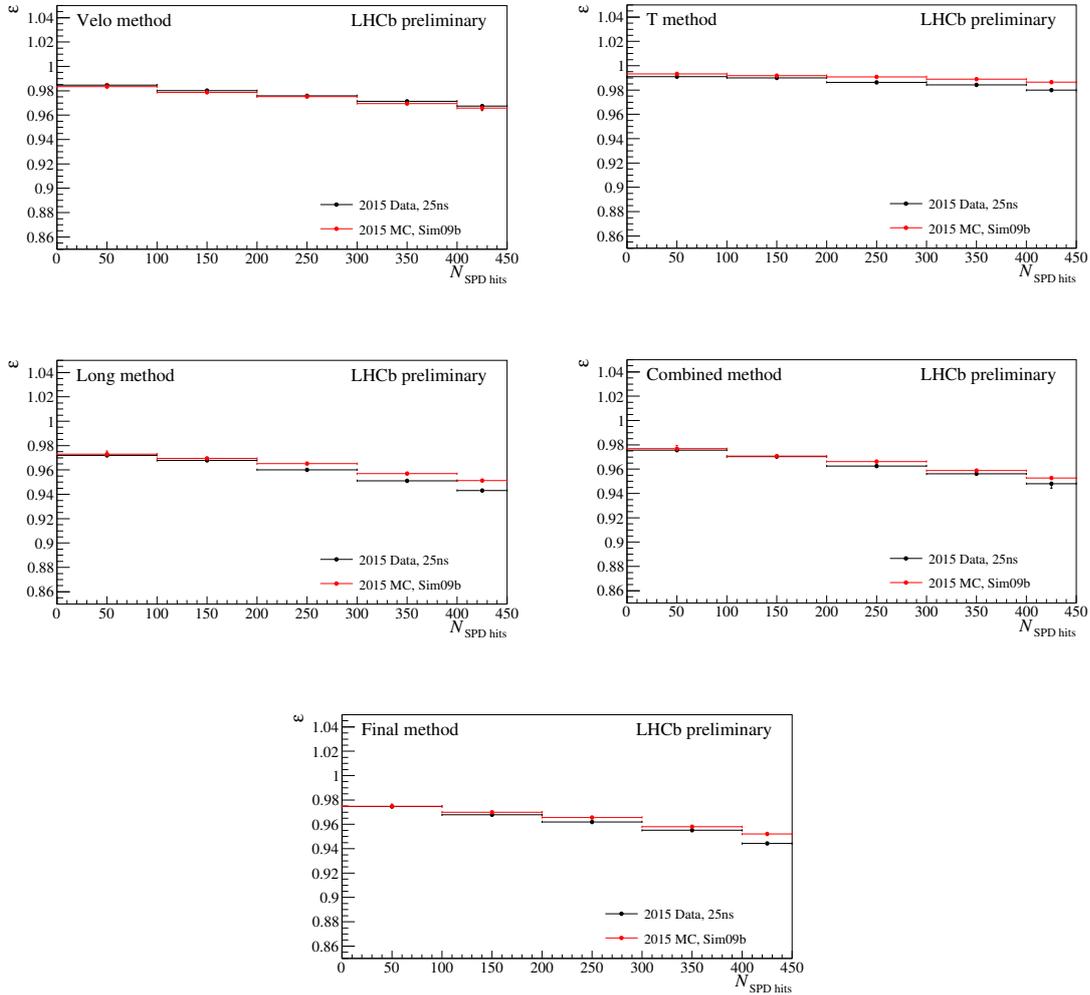


Figure 5.5: Comparison of track reconstruction efficiency for data (black) and simulated events (red) for the 2015 data-taking period as function of the number of SPD sub-detector hits. The Combined method is the product of the VELO and T-station methods, the Final method the weighted average of the Long and Combined methods. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

the original bins the efficiency ratios are mostly stable and close to unity. However, at the boundaries of the phase space coverage larger discrepancies from the value determined for the larger bins are observed. These discrepancies are accompanied by increased statistical uncertainties due to a significantly reduced number of signal candidates in these regions. Due to the small fraction of tracks in these kinematic regions, the average value of the larger bin is barely affected by the discrepancies. For

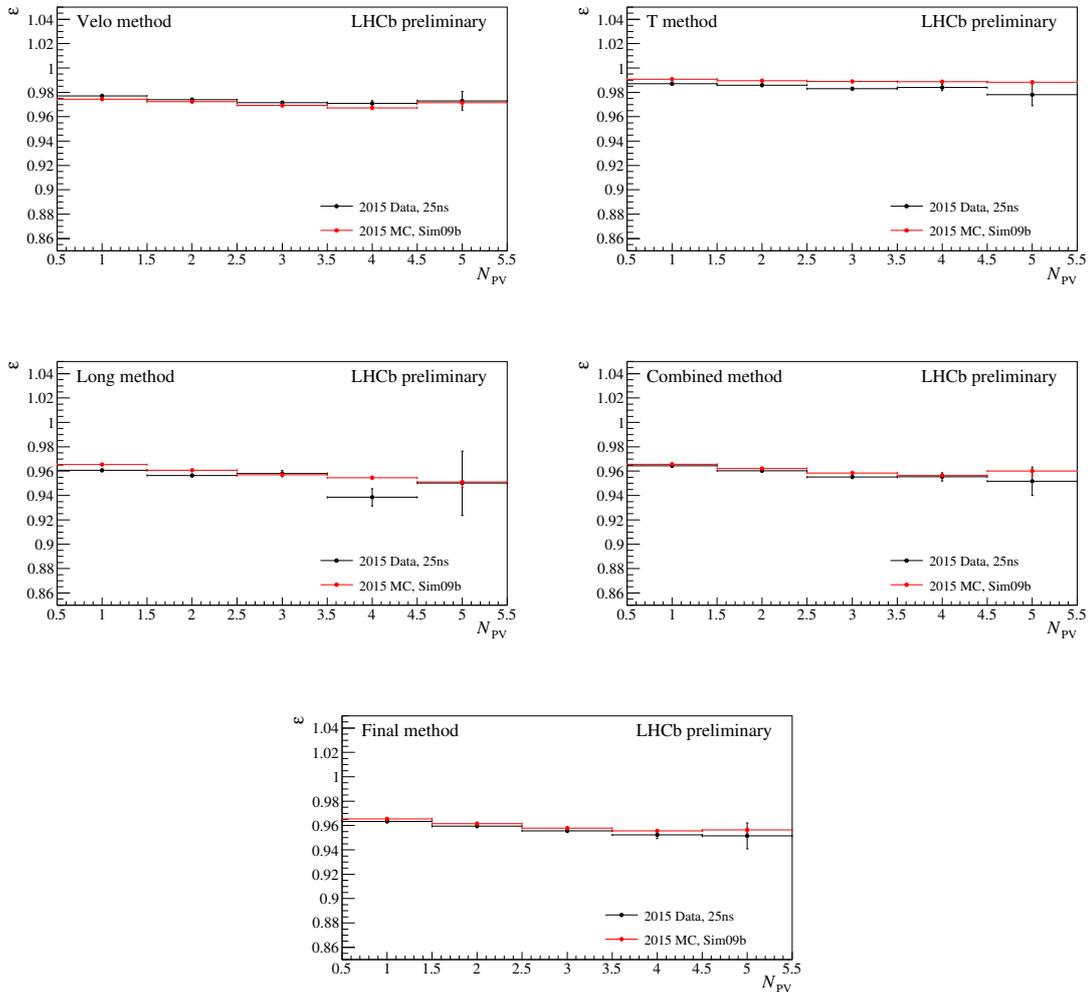


Figure 5.6: Comparison of track reconstruction efficiency for data (black) and simulated events (red) for the 2015 data-taking period as function of the number of reconstructed primary interaction vertices. The Combined method is the product of the VELO and T-station methods, the Final method the weighted average of the Long and Combined methods. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

analyses which mostly probe these kinematic regions, the usage of the finer binned corrections is advised, with the caveat of larger statistical uncertainties associated to the correction.

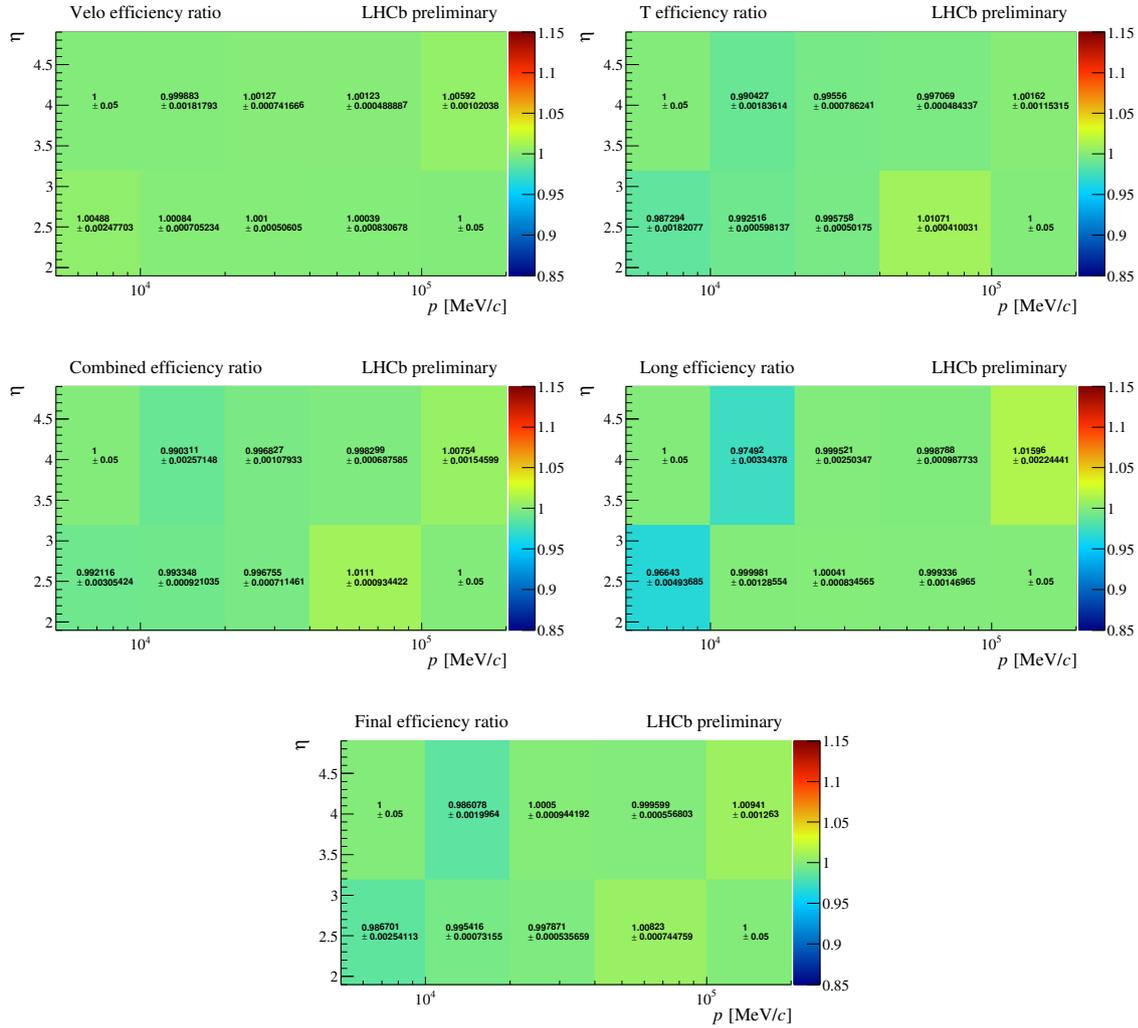


Figure 5.7: Ratio of track reconstruction efficiencies between data and simulated events for the 2015 data-taking period. The Combined method is the product of the VELO and T-station methods, the Final method the weighted average of the Long and Combined methods. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

5.4.2 Dependency on the pseudorapidity of the primary hadron

Measurements of b -hadron and c -hadron production cross sections are typically performed as a function of the kinematics of the respective hadron. The factors used to correct for differences in the reconstruction efficiencies between data and simulated events are given as a function of the kinematics of each final-state track. However, the pseudorapidity of the final-state tracks can be highly correlated to the pseudorapidity

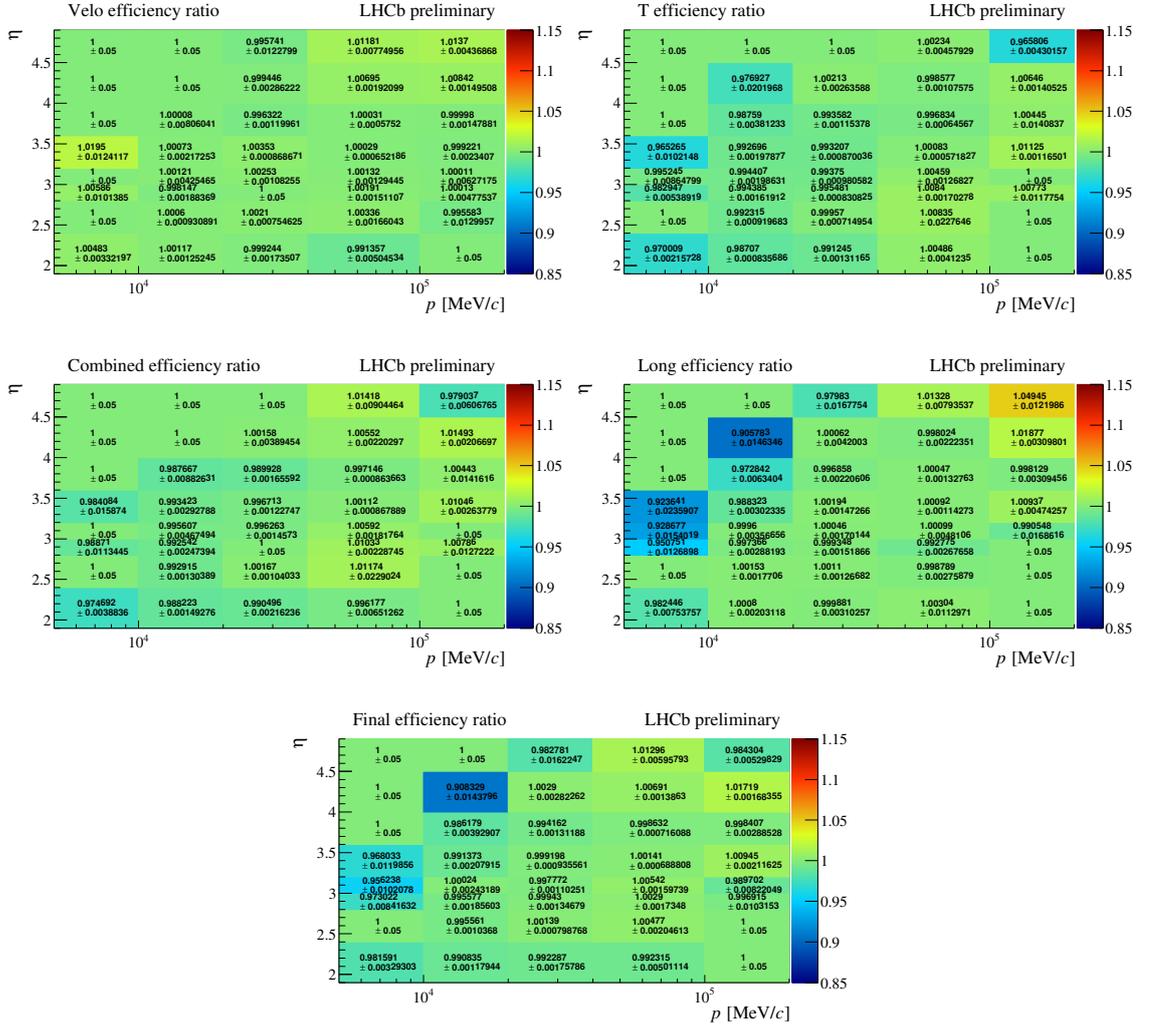


Figure 5.8: Ratio of track reconstruction efficiencies between data and simulated events for the 2015 data-taking period. The Combined method is the product of the VELO and T-station methods, the Final method the weighted average of the Long and Combined methods. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

of the decaying hadron. Due to these correlations even small per-track corrections can lead to large corrections for the combined final state. The signal candidates used for evaluation of the track reconstruction efficiencies are reconstructed as $J/\psi \rightarrow \mu^+ \mu^-$ decays, where the J/ψ -meson candidate is significantly detached from the primary interaction vertex. As the J/ψ meson is extremely short-lived, it will typically originate from the decay of a hadron produced in the primary interaction. Using the positions of the decay vertex of the J/ψ meson and the reconstructed primary

interaction vertex, the direction of the initial particle can be reproduced, and its pseudorapidity calculated. The track reconstruction efficiency for data and simulated is determined as a function of this “Mother” η , shown in Fig. 5.9. Dependencies of the efficiencies of both data and simulated events are observed, and the ratio of those two efficiencies shows significant dependencies at very low and high pseudorapidities. For analyses performed as function of hadron pseudorapidities, these corrections should be used as a function of the mother kinematics in order to account for these dependencies.

5.4.3 Choice of the overlap fractions

During the measurement of the track reconstruction efficiencies in LHC Run I, the fraction of overlap required for the association of the probe signatures with a Long track was optimised to return results agreeing with the “true” efficiency given for simulated events. For the Run II implementation in the software trigger, the fractions of common sub-detector hits with the probe signature required for a Long track to be saved is set to lower values, to allow for an investigation of the stability of the correction factors with respect to this choice. In general, requiring a higher overlapping hit fraction will result in a lower efficiency, while lowering the threshold increases the amount of mis-associated Long tracks.

For the VELO method, the fraction of common hits in the T stations and the TT has to be larger than 40%, while for the T-station method a minimum overlap of 50% of hits in the VELO and 40% in the muon stations is required. For the Long method, 40% common hits in the TT are required as well as in the muon stations. The validation with simulated events described above was performed with these efficiency criteria. For all three methods the effects when changing the required overlap fraction in the respective sub-detectors to up to 80% were found to be negligible compared to the statistical precision of the measurement. Thus, no additional tuning of the required overlap fractions is necessary.

5.5 Systematic uncertainties

Many sources of potential systematic uncertainties have been investigated during the measurement of the track reconstruction efficiencies in LHC Run I [7]. As the basic methods did not change, negligible uncertainties are not expected to have changed in a significant way for the Run II measurement. Changing the signal model used in the fits to reconstructed invariant mass to the sum of two Gaussian distributions does not change the efficiency significantly compared to the statistical uncertainty. The same negligible change is observed when using a linear background model instead of an exponential function. The small difference between the tracking efficiency derived from the Long method and the combination of VELO and T-station method is mostly covered by the statistical uncertainties of those methods, and are expected to further reduce when taking the ratio of the efficiencies from data and simulated events. Thus

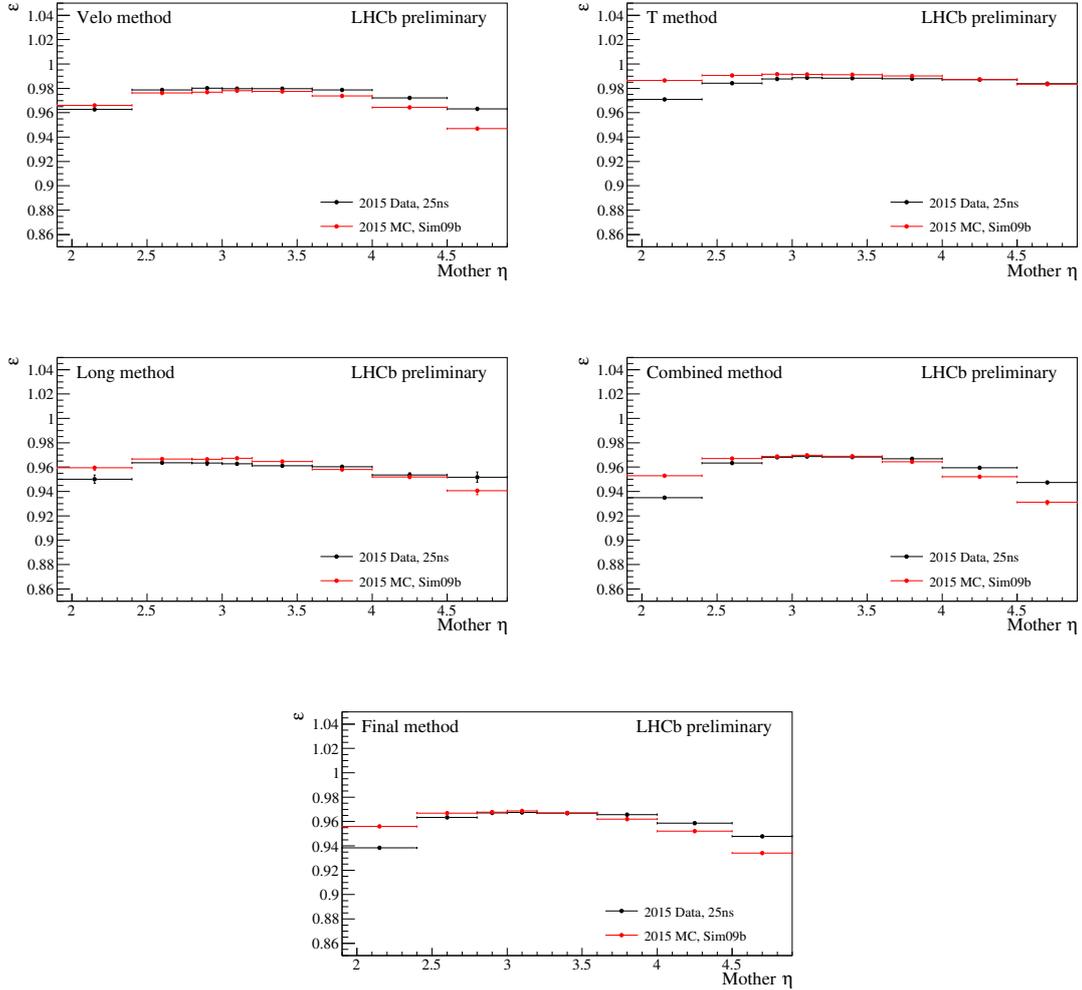


Figure 5.9: Comparison of track reconstruction efficiency for data (black) and simulated events (red) for the 2015 data-taking period as function of the pseudorapidity of the J/ψ “Mother” particle. The Combined method is the product of the VELO and T-station methods, the Final method the weighted average of the Long and Combined methods. “Sim09b” refers to the version of the simulation software used by the LHCb experiment. Uncertainties are statistical only.

no systematic uncertainty is assigned to this deviation. The uncertainties caused by the additional hadronic material interactions of kaons and pions compared to the measured muons are linked to the detector layout and the uncertainty of the material budget. As these remain unchanged for the LHCb detector in LHC Run II, the associated systematic uncertainties for the tracking efficiency measurement are 1.1% for kaons and 1.4% for pions, as described in the previous chapter. The choice of the

occupancy variable used to improve the agreement of the simulated event sample with real data was the dominant systematic uncertainty during the Run I measurement. In order to evaluate the uncertainty caused by the choice of the number of hits in the SPD, the efficiency ratios are re-evaluated when using the number of tracks, the number of Long tracks, and the number of primary interaction vertices as proxy for the detector occupancy. The distributions of the number of SPD hits and number of reconstructed tracks are shown as an example for the candidates selected for the Long method in Fig. 5.10. The largest deviation observed for the correction factors obtained from the combination of all methods in any of the two-dimensional bins is 0.6%, which is assigned as systematic uncertainty.

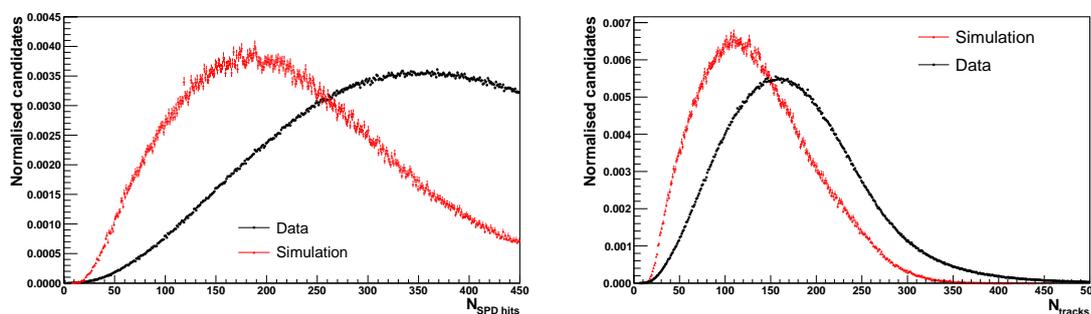


Figure 5.10: Comparison of the normalised distributions of the number of hits detected in the SPD (left) and number of reconstructed tracks (right) for real data and simulated events as selected for the Long method. Uncertainties are statistical only.

6 TrackCalib: User-defined tracking efficiency evaluation

The correction tables provided by the track reconstruction efficiency measurement are used in a wide range of analyses performed within the LHCb collaboration. However, the nominal choice of variables in which these tables are binned and the binning schemes do not fit the demands of every analysis. Furthermore, in the case that tighter reconstruction quality criteria are applied in the selection process compared to what is applied in the default reconstruction, the ratio of efficiencies of data compared to simulated events might be affected. The nominal correction tables are created regularly for new data-taking conditions and versions of the simulation software used by the LHCb collaboration. For the tracking efficiency evaluation during LHC Run I, each request of customised correction tables required a dedicated determination of those specific requirements.

In order to allow for a completely customised and accessible creation of the requested correction tables, the tool `TrackCalib` has been created for the track reconstruction efficiency determination for LHC Run II. This tool enables the determination of the track reconstruction efficiencies and creation of the corresponding correction tables from command-line options. This approach profits from the “quasi-online” evaluation of the efficiencies, which includes the automated creation of publicly available datasets shortly after being recorded by the LHCb experiment.

6.1 Features of TrackCalib

The main features included in the `TrackCalib` tool include:

- **Methods and data samples:** Users may specify to only evaluate the track reconstruction efficiency for a subset of the three tag-and-probe methods, which can be helpful in spotting problems in any of those. Furthermore, the procedure can be run on real data or simulated event samples exclusively. In that case, no correction tables can be created until both datasets have been properly processed.
- **Track quality criteria:** In the nominal LHCb reconstruction, tracks need to pass a requirement on the ghost probability. This probability is determined using a trained neural net classifier, and tracks are required to have a ghost probability value of smaller than 40% [18]. In case that tighter selection requirements are chosen, the track reconstruction efficiency decreases, which might affect real data and simulated events differently. Users may specify these tighter selection criteria to be applied to the Long tracks which are matched to the probe signature, to determine the effect these cuts have on the tracking efficiencies and the ratio of the efficiencies for data and simulation.

-
- **Occupancy variable:** While the default choice of the number of SPD hits as the proxy for the detector occupancy is independent of the reconstruction software, investigating the dependency of the correction tables on this choice can be important. Users can choose the variable which is used to weight the simulated event sample to reflect the occupancy observed in real data. Furthermore, the weighting of the simulated event sample can be disabled completely, in order to evaluate the dependency of the efficiency on detector occupancy effects.
 - **Efficiency variables:** The nominal variables in which the track reconstruction efficiencies are evaluated are the momentum and pseudorapidity of the probe signature and the number of SPD hits and reconstructed primary interaction vertices. For the correction tables two-dimensional bins of probe signature momentum and pseudorapidity are chosen. Both the choice of variables as well as the two-dimensional combinations used for the correction tables can be chosen freely by users.
 - **Binning schemes:** Per default, the binning schemes used for the tracking efficiency evaluation are the same as used in the nominal determination. Users may change those to their requirements, and, in case custom variables are used, have to define a binning scheme for those variables. With the given variables and binning schemes, the `TrackCalib` framework automatically divides the data samples and performs the simultaneous fits in those bins.
 - **Fit model:** The signal models used in the fits of the nominal track reconstruction efficiency measurement can be changed from the sum of two Crystal-Ball functions to a sum of two Gaussian distributions. This typically helps with low statistics samples (which may occur with custom binning schemes), as the tail parameters of the Crystal-Ball functions may not be properly extracted.

These features allow for a high customisability of the tracking efficiency evaluation.

6.2 Data flow and tool structure

In order to give users the option to run the track reconstruction efficiency evaluation with a customisable tool, the data samples used need to be prepared and made available in a standardised manner. Figure 6.1 shows the flow of the data after being recorded by the trigger lines described in the previous chapter, and where and how the tool appears in the processing of these data samples. The candidates selected by the trigger lines as part of the “TurboCalib” stream are processed automatically after data taking as part of the centralised processing. As all reconstructed objects and their properties as derived in the software triggered are saved, no additional reconstruction algorithms have to be executed, and this dataset is represented in a “quasi-online” state. Furthermore, the data samples are available for the determination

of the track reconstruction efficiencies shortly after the data is recorded by the LHCb detector. From the centralised processing the data samples are saved as “Tuples” in locations which can be accessed by every collaboration member. These are organised by data-taking conditions and the version of the simulation software used.

The `TrackCalib` tool accesses these central Tuples, which contain all information

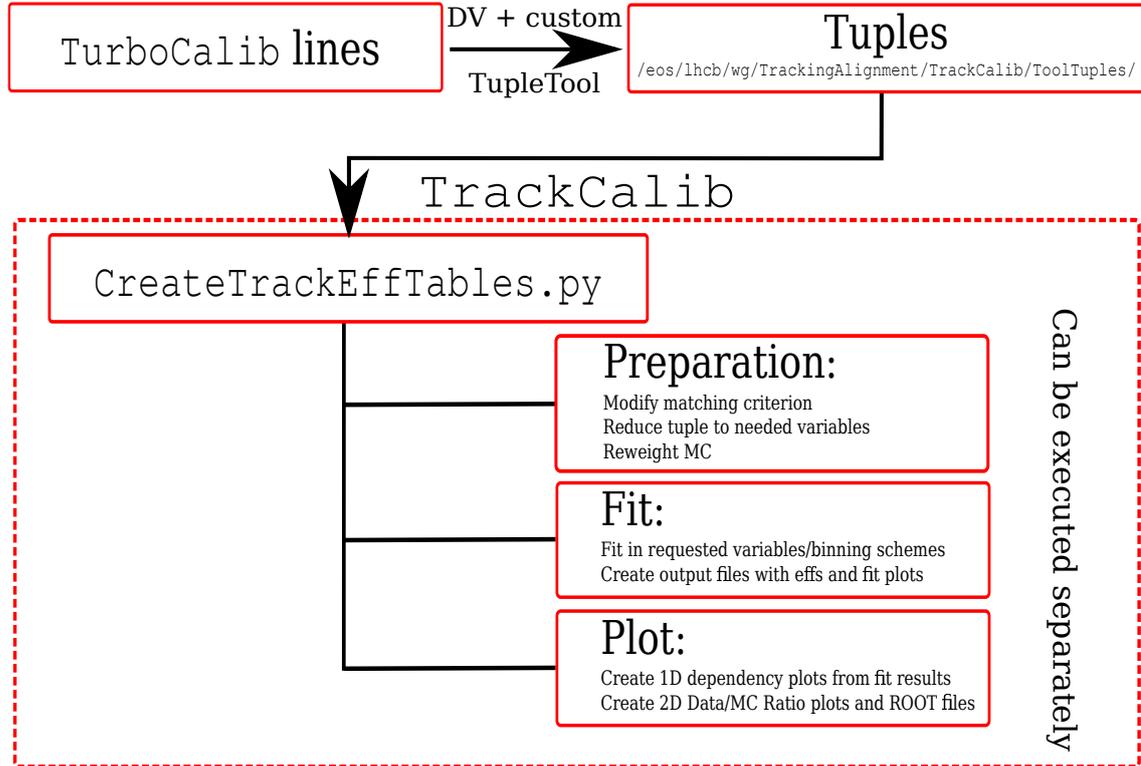


Figure 6.1: Schematic view of the processing of the data samples used for the tracking efficiency measurement and structure of the `TrackCalib` tool.

needed for its execution. The tool itself is structured into three main steps:

- **Preparation:** The central datasets, especially for real data, can be extremely large due to a large set of variables saved in them. Purpose of the preparation step is to only copy all information needed by the tool, depending on the variables requested by the user. In this step also additional track quality criteria can be applied to the Long track associated to the probe signature or the overlapping hit fraction required for association can be modified. Furthermore, for the simulated event samples weights are calculated to correct for the difference in occupancy compared to the respective data sample. The resulting datasets are stored locally by the user, allowing for easy accessibility for the following steps.

-
- **Fit execution:** The datasets created by the preparation step are divided according to the requested variables and binning schemes. In each bin, the simultaneous fit is performed with the given model, and the results saved. In the default setup, the fits are performed for all three tag-and-probe methods for real data and simulated events. For each method and variable (or variable set for two-dimensional bins) a results file is created containing the fitted distributions and efficiencies as function of the respective variable. That way, the fitted distributions can be reviewed after the fitting step to assure proper convergence.
 - **Plotting:** The efficiencies saved in the results files from the fitting step are used to create dependency plots, such as in the figures of the previous chapter. In addition, the efficiency values from the VELO and T-station methods are multiplied to create the “Combined” method efficiencies. The weighted average of the efficiencies from the Combined and Long methods are given as the “Final” method. For efficiencies as function of two-dimensional bins, the ratio between the values derived from the data and simulated event sample is calculated for all methods (including Combined and Final) and the corresponding correction table created.

While the default execution of the `TrackCalib` tool performs these three steps in the order given above, each step can also be executed individually. This is especially helpful in case that different binning schemes are investigated, or fit models tested. User can use these steps to not only create two-dimensional correction tables for any set of variable, but also investigate the dependency of the track reconstruction efficiency for variables independently.

The `TrackCalib` tool has been published for usage within the LHCb collaboration in 2017.

7 Conclusion

The first part of this thesis presents the measurement of the track reconstruction efficiency of the LHCb experiment. The reconstruction efficiency of so-called Long tracks is measured with a tag-and-probe approach. Three different methods either probe the efficiency to reconstruct Long tracks in specific sub-detectors or as a whole. The evaluation is performed for real data collected during LHC Run II and for samples of simulated events. The difference between data and simulation is mapped onto a two-dimensional correction table depending on the kinematic properties of the track, which is used by analysts to correct the track reconstruction efficiency evaluated from simulated event samples.

For the Run II data taking the LHCb collaboration implemented a real-time alignment and calibration of the LHCb detector. This enables the performance of physics analyses with the online reconstructed objects, which requires a determination of the detector performance in this online state. For that reason, the tag-and-probe methods used for the tracking efficiency measurement have been moved into the LHCb software trigger, which allows for a “quasi-online” determination of the efficiencies directly after data taking. This required changes in the selection process and the execution of the probe signature reconstruction algorithms. The updated implementation has been validated using simulated events, where the efficiencies determined with the tag-and-probe methods reflect the true track reconstruction efficiencies well.

Additionally, the processing of the dataset collected that way has been automatised, and a tool called `TrackCalib` has been implemented to provide a customised creation of the correction tables, which fit the requirements of each analysis. This is part of an ongoing effort by the LHCb experiment to increase the automation and improve the transparency of the detector calibration.

The Long track reconstruction efficiency for the 2015 data-taking period is measured to be $(96.14 \pm 0.03)\%$, which is slightly lower than what was measured during LHC Run I. This is within expectations, as the higher occupancy caused by the larger centre-of-mass energy of the proton-proton collisions is detrimental to the detector performance. Nevertheless, the performance achieved for the track reconstruction for the second phase of the LHCb experiment is extremely good and satisfies the requirements for high precision measurements.

Part II: Relative branching
fraction measurement and angular
analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$

8 Introduction

The Standard Model of Particle Physics (SM) has been one of the best established theoretical physics models created so far. Besides successfully predicting the t quark and the third generation of leptons, the discovery of the Higgs boson in 2012 by the ATLAS and CMS experiments [5] completes the observation of all particles and interactions incorporated by it. However, the Standard Model does not provide a theory of gravity or explain phenomena such as Dark Matter or Dark Energy, or the observed matter-antimatter asymmetry of the universe. Thus the search for New Physics beyond the description of the SM is one of the most important tasks of modern particle physics.

Loop-level processes provide a strong probe to test deviations from the SM, as contributions from New Physics can have a significant influence on these rarely occurring decays. A very promising field of studying loop-level transitions are electroweak penguin decays. These decays are flavour-changing neutral current (FCNC) transitions, which can in the SM only happen on loop-level and are thus heavily suppressed compared to tree-level decays. Previous measurements of the electroweak penguin decays $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ [22], $B^+ \rightarrow K^+ \mu^+ \mu^-$ [23], $B_s^0 \rightarrow \phi \mu^+ \mu^-$ [24] and $\Lambda_b^0 \rightarrow \Lambda^0 \mu^+ \mu^-$ [25] show significant deviations from what is expected from pure Standard Model processes. Furthermore, the test of lepton universality by comparing the penguin decays $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} e^+ e^-$ shows a strong tension with SM expectations [26]. The measurement of further electroweak penguin modes as well as improving precision of previously performed measurements are important to uncover the source of these discrepancies.

The second part of this thesis presents the measurement of the branching fraction and analysis of the distributions of angular observables of the electroweak penguin decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$, where the ϕ meson is reconstructed in the decay $\phi \rightarrow K^+ K^-$. Due to the loop suppression, the SM prediction for the branching fraction is comparably small, at $\mathcal{O}(10^{-7})$ [27]. As potential New Physics effects are often dependent on the squared momentum transfer q^2 of the respective process, the measurement is performed in bins of this variable, which corresponds to the dimuon mass squared. This second part is organised as follows:

Chapter 9 provides a brief introduction into the Standard Model of Particle Physics and the effective field theory used to describe electroweak penguin decays as investigated in this thesis. The phenomenology of such decays is introduced, including theoretical models used for predictions based on the Standard Model.

Chapter 10 gives an overview of the steps needed to perform the measurement of the differential branching fraction and the distributions of the decay angles. Due to the rarity of the signal decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ a careful selection is necessary in order to distinguish the signal decay mode from various background modes. This selection is described in Chap. 11 and happens in multiple stages: the hardware and software trigger stages select events with topologies similar to the signal decay, while a centralised offline selection, called Stripping, is used after the full event

reconstruction. Such selected events are further discriminated against backgrounds by using a cut-based pre-selection and a multivariate classifier.

A large sample of information about both the signal and control channel decays is derived from fully simulated events. The simulation used by the LHCb experiment models most of these decays very well, however, a few important properties are not modelled to a satisfying detail. To account for this, multiple correction steps are applied to the samples in order to better reflect the data sample. These corrections are described in Chap. 12.

Chapter 13 illustrates the determination of the efficiencies to reconstruct and select the signal and control mode decays from these corrected simulated samples. These efficiencies can be factorised into multiple sub-efficiencies: the probability for the decay products to appear within the acceptance of the LHCb detector, the efficiency for the trigger to activate on a signal event and record it, the efficiency to correctly reconstruct all final-state particles and properly identify the particle species, and the efficiency for the event to pass the additional selection steps.

With these efficiencies determined, the relative branching ratio between signal and control mode can be derived from their respective yields after all reconstruction and selection steps are applied. These yields are extracted from fits to the distribution of the reconstructed invariant mass of the B_s^0 -meson candidates, for the signal channel in bins of the dimuon mass squared q^2 . The fitting procedure as well as the computation of the relative differential branching ratio as well as the signal branching ratio integrated over the full q^2 range are described in Chap. 14.

Additionally, from analysing the distribution of the decay angles, which in combination with q^2 fully describe the decay kinematics, a multitude of CP -averages and CP -asymmetries can be derived. This evaluation is performed in bins of q^2 , and 8 observables are investigated. Detector acceptance effects affecting the shape of the distributions of these observables are determined from above mentioned corrected simulation samples, and further used in the fit to these distributions. The full procedure is illustrated in detail in Chap. 15.

All measurements of physics parameters are subject to uncertainties. This includes both the direct measurement of first-hand observables as well as the extraction of angular observables and branching fraction from these observables, and any input derived from simulated event samples. As both measurements portrayed in the previous chapters are affected by these in similar ways, the evaluation of the systematic uncertainties is summarised in a combined approach in Chap. 16.

Finally, the results of both the differential branching ratio measurement and the angular analysis are summarised in Chap. 17. The measured values are in addition compared to Standard Model predictions from Refs. [27] and [28], and the implications of these results are further discussed. Concluding remarks are given in Chap. 18.

9 Theoretical Framework

This chapter gives an overview of basics of the Standard Model and focuses further on rare flavour changing neutral current (FCNC) processes, as investigated in this thesis, and the effective theories describing them. It will also give a brief summary of models describing potential sources of “New Physics” affecting measurements of FCNC processes.

9.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) is a quantum field theory combining the electroweak theory with Quantum Chromo Dynamics (QCD). It describes the fundamental electromagnetic, weak, and strong interactions between particles. The Standard Model is a gauge theory, with a Lagrangian which is invariant under local gauge transformations of the symmetry group $SU(3)\otimes SU(2)\otimes U(1)$. Each symmetry is connected to a charge, called quantum number, which is conserved in the corresponding interaction. There are four electroweak charges and eight strong charges, and corresponding gauge bosons of the electroweak and strong interactions. The gauge bosons mediating the respective interactions are spin-1 particles, while the matter fields, the fermions, carry a spin of $1/2$. The description of the SM given in this section is based on reviews in Refs. [29, 30].

9.1.1 Fundamental interactions and particles

The electroweak theory evolved mainly through the contributions of Salam [32], Glashow [33], and Weinberg [34] in the 1960s, in order to incorporate the interactions of elementary particles and observed phenomena up to that point. It is based on the $SU(2)\otimes U(1)$ gauge group, with the corresponding gauge bosons $W_\mu^{1,2,3}$ and B_μ , and unifies the electromagnetic and the weak interactions. Through the spontaneous symmetry breaking caused by the Higgs mechanism at energies below the weak energy scale, the masses of the W^\pm and Z bosons are generated, while the photon remains massless. They are the force carriers of the electroweak interaction and are linear combinations of the gauge bosons mentioned above. $A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W$ is the photon field, $Z_\mu = -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W$ is the Z boson field, and $W_\mu^\pm = (W_\mu^1 \mp iW_\mu^2)/\sqrt{2}$ are the fields of the charged W bosons. Here, the weak mixing angle θ_W was introduced. The gauge bosons of the electroweak interaction couple to the electroweak charge of particles. This electroweak charge depends on the respective particle’s chirality, which leads to an exclusive coupling of the W^\pm bosons to left-handed particles and right-handed anti-particles. For that reason, the electroweak theory is often described as a “chiral” theory. The Higgs mechanism also predicts the presence of at least one spin-0 particle, the Higgs boson, which does not mediate an interaction. A particle has been discovered by the ATLAS and CMS experiments in 2012 at the Large Hadron Collider, compatible

Standard Model of Elementary Particles

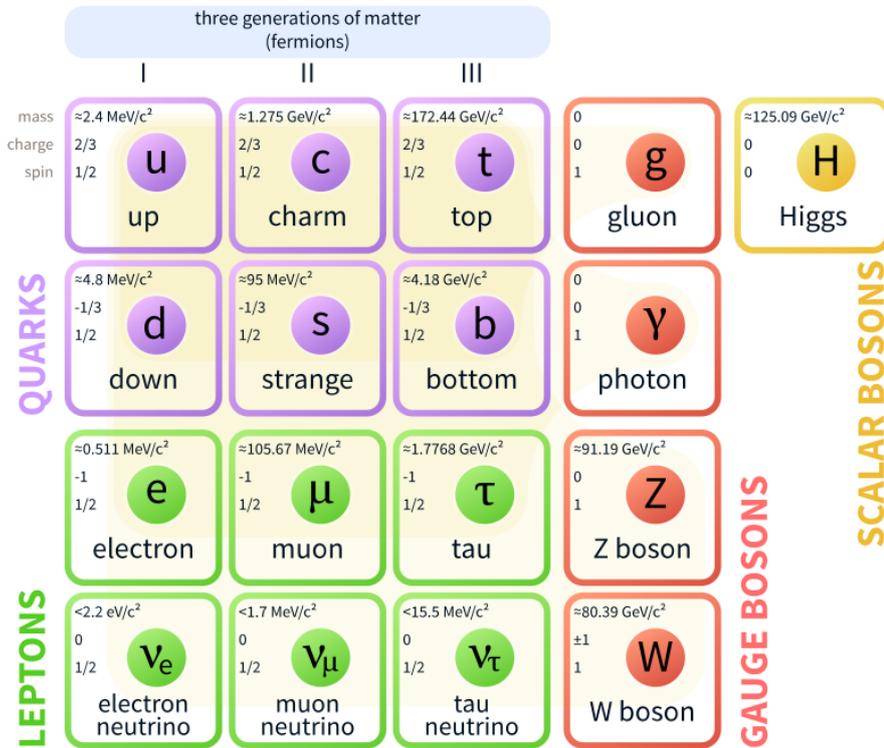


Figure 9.1: The fundamental particles of the Standard Model [31]

with such a Higgs boson [5]. However, certain properties such as cubic and quartic self-coupling still remain to be measured, to make sure it is identical to the SM Higgs boson.

The gauge theory of Quantum Chromo Dynamics, represented by the $SU(3)$ component of the Standard Model symmetry group, describes the strong interaction. Its massless gauge bosons are the eight gluons, which couple to the colour charge carried by quarks and gluons. In field theory, the effective coupling strength in each vertex depends on the interaction itself. For QCD the strong coupling $\alpha_s(Q^2)$ depends strongly on the transferred four-momentum Q^2 between the participants. It decreases for increasing Q^2 , and vanishes asymptotically. This feature of QCD is called asymptotic freedom, leading to the interaction between particles to become asymptotically weaker at increasing energy scales. This allows for perturbative QCD calculations at large momentum exchanges and short-distance interactions.

The particle content of the Standard Model is shown in Fig. 9.1, including the respective masses, charges and spins, as measured so far. The fermions, representing the mass fields of the SM, can be grouped into three generations. For each fermion

there exists a corresponding anti-particle, with mass identical to the particle, but all quantum numbers inverted (including the charge). As anti-particles behave mostly symmetric to their particle counterparts, they will not be explicitly mentioned in the following part.

There are three negatively charged leptons (e^- , μ^- , and τ^-) and corresponding neutrinos (ν_e , ν_μ , and ν_τ) with neutral charge. They leptons carry the electroweak charge but no colour, and thus do not interact via the strong force. The two leptons of each pair carry the corresponding lepton flavour number, $L_{e/\mu/\tau} = +1$. These lepton flavour numbers were for a long time assumed to be conserved in electroweak interactions, but the discovery of neutrino oscillations shows this not to be the case [35]. However, these neutrino oscillations are very long-range effects and thus play no significant role in the physics processes investigated at particle accelerator experiments. So, for simplicity, hereafter lepton flavour conservation will be assumed. There are six quarks currently known, with three up-type quarks (u , c , and t) with a charge of $+2/3$ and three down-type quarks (d , s , and b) of charge $-1/3$. They carry both the electroweak charge as well as the colour charge. In contrast to the leptons, quarks are not observed as free particles, but instead form so-called hadrons. This is thought to be caused by the asymptotic freedom of QCD, which leads an increased coupling strength between quarks at larger distances. For that reason, the increasingly large energy needed to separate two quarks leads instead to the creation of new quark pairs. This causes the so-called confinement of quarks into hadronic states. The hadrons which were first observed (due to their stability) are the baryons, which consists of three quarks or anti-quarks with each of the three colours. The most prominent baryons are the proton (two u , one d) and the neutron (one u , two d), which the usual atomic matter in the universe is made up of. The mesons consist of a quark and anti-quark, which carry opposite colour charges. Examples are pions (consisting of u and d quarks) and kaons (s and u/d quarks).

The fermions can be grouped into three generations, with the structure

$$\begin{pmatrix} u_r & u_g & u_b & \nu_e \\ d_r & d_g & d_b & e^- \end{pmatrix}, \\ \begin{pmatrix} c_r & c_g & c_b & \nu_\mu \\ s_r & s_g & s_b & \mu^- \end{pmatrix}, \\ \begin{pmatrix} t_r & t_g & t_b & \nu_\tau \\ b_r & b_g & b_b & \tau^- \end{pmatrix},$$

where each generation consists a weakly charge doublet of quarks, in three colour replicas, and a colourless weakly charged neutrino-lepton doublet. At the current state of research, there is no explanation for this structure of fermion generations.

9.1.2 The CKM mechanism

The weak interaction is the only interactions mediating flavour changing currents. These are occurring due to the exchange of the charged W^\pm bosons, allowing for

the transition of up-type to down-type quarks and vice versa. Transitions from one generation to another are caused by the physical mass or flavour eigenstates not being eigenstates of the weak interaction. The flavour changing transitions of the weak interactions are described by the Cabbibo-Kobayashi-Maskawa mechanism [36], where the change of basis from the flavour eigenstates (d, s, b) to the weak Eigenstates (d', s', b') is given by the complex and unitary CKM matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (9.1)$$

The elements of this CKM matrix enter the Standard Model Lagrangian via a charged current term:

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}} \left[\bar{u}_i \gamma^\mu \frac{1-\gamma^5}{2} V_{ij} d_j + \bar{\nu}_i \frac{1-\gamma^5}{2} e_i \right] W_\mu^+ + h.c. \quad (9.2)$$

Thus the transition probability for a process involving a i - to j -quark transition is directly proportional to the absolute value squared of the respective matrix element $|V_{ij}|^2$. The original 18 free parameters of the complex three-dimensional matrix get reduced to 9 due to the required unitarity. As relative phases are not physically observable, the five relative quark phases can be used to further reduce this to four parameters, which in the standard parametrisation of the CKM matrix are chosen as three Euler angles Θ_{12} , Θ_{13} and Θ_{23} , and one phase δ . Using these parameters V_{CKM} can be written as

$$\begin{aligned} V_{\text{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (9.3) \end{aligned}$$

where the abbreviations $c_{ij} = \cos \Theta_{ij}$ and $s_{ij} = \sin \Theta_{ij}$ are used. The phase δ is the only source of CP violation in the Standard Model.

When defining the parameters

$$\begin{aligned} \lambda &= s_{12} \\ A &= s_{23}/\lambda^2 \\ \rho + i\eta &= s_{13}e^{i\delta}/A\lambda^3, \end{aligned}$$

an expansion can be performed in orders of $\lambda \approx 0.23$. This leads to the so-called Wolfenstein parametrisation, in which the CKM matrix takes the form

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (9.4)$$

An advantage of this parametrisation is that it shows the relative magnitudes of the individual elements. All diagonal elements are close to unity, making transitions within a generation the most likely, while the magnitudes of off-diagonal elements become smaller the further from the diagonal they are. This supports both the principle of describing the quarks in generations as well as their relative ordering. Transitions between the first and second generation of quarks are suppressed by a factor of λ^2 compared to transitions within the respective generation. This suppression is often referred to as "Cabbibo suppression", as they were already described by the Cabbibo Model, which only included the first two generations of quarks. Transitions between the second and third or first and third generations is suppressed by factors of λ^4 and λ^6 , making these even less likely to occur. In a similar manner to the first to second generation transitions these are also called "CKM suppressed" when comparing them to more frequently occurring processes.

9.1.3 CP violation and neutral meson mixing

According to the criteria formulated by Andrei Sakharov [37], one of the necessities for the baryon asymmetry observed in the universe is the violation of the CP -symmetry. A C conjugation is the inversion of the quantum numbers (including the charge) of a particle or process, while P conjugation inverts the spatial coordinates. For elementary particles the CP conjugation inverts all quantum numbers and simultaneously the particle's chirality. For example, the CP -conjugated state of a left-handed neutrino is the corresponding right-handed anti-neutrino. This can also be performed for full decay chains, as for the decay mode investigated in this thesis, where the CP -conjugated decay to $B_s^0 \rightarrow \phi(\rightarrow K^+K^-)\mu^+\mu^-$ is the decay $\bar{B}_s^0 \rightarrow \phi(\rightarrow K^+K^-)\mu^+\mu^-$. As the final state is symmetric under CP conjugation, these two decays cannot be experimentally distinguished.

The Standard Model includes only one CP -violating parameter, the phase δ of the CKM-matrix described above. CP violation in the SM can be measured by comparing decays with their CP -conjugated equivalent. For neutral mesons, especially B^0 and B_s^0 mesons, the phenomenon of "mixing" can additionally be affected by CP violation. In this section neutral meson mixing will be introduced, following Ref. [38], using the B_s^0 meson system as an example.

The B_s^0 meson consists of a \bar{b} quark and a s quark, and the \bar{B}_s^0 of a b quark and a \bar{s} quark. While they are flavour eigenstates, they are not identical to the mass eigenstates, which are the eigenstates of an effective Hamiltonian \mathcal{H} , defined via

$$\mathcal{H} = M - \frac{i}{2}\Gamma. \quad (9.5)$$

Here M and Γ are hermitian 2×2 matrices. With this Hamiltonian the time evolution of any linear combination $a|B_s^0\rangle + b|\bar{B}_s^0\rangle$ of the flavour eigenstates can be described with the time-dependent Schrödinger equation

$$i\frac{d}{dt} \begin{pmatrix} a \\ b \end{pmatrix} = \mathcal{H} \begin{pmatrix} a \\ b \end{pmatrix} = \left(M - \frac{i}{2}\Gamma\right) \begin{pmatrix} a \\ b \end{pmatrix}, \quad (9.6)$$

where the basis $(|B_s^0\rangle, |\bar{B}_s^0\rangle)$ was used. The two eigenstates of \mathcal{H} are the mass eigenstates, $|B_H\rangle$ and $|B_L\rangle$, which correspond to the high and low mass, respectively. These eigenstates can themselves be described by linear combinations of the flavour eigenstates

$$|B_H\rangle = p |B_s^0\rangle - q |\bar{B}_s^0\rangle \quad (9.7)$$

and

$$|B_L\rangle = p |B_s^0\rangle + q |\bar{B}_s^0\rangle, \quad (9.8)$$

where normalisation requires $q^2 + p^2 = 1$. The eigenvalues of B_H and B_L of the matrices M and Γ are $m_{H/L}$ and $\Gamma_{H/L}$, respectively, corresponding to the mass and decay width of those states. For these states, the time evolution is then given by

$$|B_H\rangle(t) = e^{-iM_H t} e^{\frac{\Gamma_H}{2} t} |B_H\rangle \quad (9.9)$$

and

$$|B_L\rangle(t) = e^{-iM_L t} e^{\frac{\Gamma_L}{2} t} |B_L\rangle. \quad (9.10)$$

The mass and decay width differences Δm_s and $\Delta\Gamma_s$ are defined as

$$\Delta m_s = m_H - m_L \quad (9.11)$$

and

$$\Delta\Gamma_s = \Gamma_L - \Gamma_H, \quad (9.12)$$

which constrains Δm_s to be positive. A physical state, which at $t = 0$ consists of a B_s^0 or \bar{B}_s^0 meson, follows a time evolution of

$$|B_{s,phys}^0\rangle = g_+(t) |B_s^0\rangle + \frac{q}{p} g_-(t) |\bar{B}_s^0\rangle \quad (9.13)$$

and

$$|\bar{B}_{s,phys}^0\rangle = \frac{p}{q} g_-(t) |B_s^0\rangle + g_+(t) |\bar{B}_s^0\rangle, \quad (9.14)$$

where the abbreviations $g_+(t)$ and $g_-(t)$ are functions of the mass and decay width differences. In the Standard Model, these oscillations occur via box diagrams as shown in Fig. 9.2. For the measurement of the mixing parameters Δm_s and $\Delta\Gamma_s$, the initial flavour of the B_s^0 or \bar{B}_s^0 meson after its production in the primary interaction needs to be determined. Experimentally, this is done via flavour tagging algorithms, as described in Ref. [39].

While the decay mode $B_s^0 \rightarrow \phi \mu^+ \mu^-$ has a CP -symmetric final state, and cannot be distinguished from its CP -conjugate equivalent, it is still affected by the mixing parameters described above. The measured data sample is a mixture of B_s^0 and \bar{B}_s^0 decays, and due to mixing the composition from these states changes over time. For that reason, the difference in decay width $\Delta\Gamma_s$ influences the data sample in case of time-dependent selection requirements.

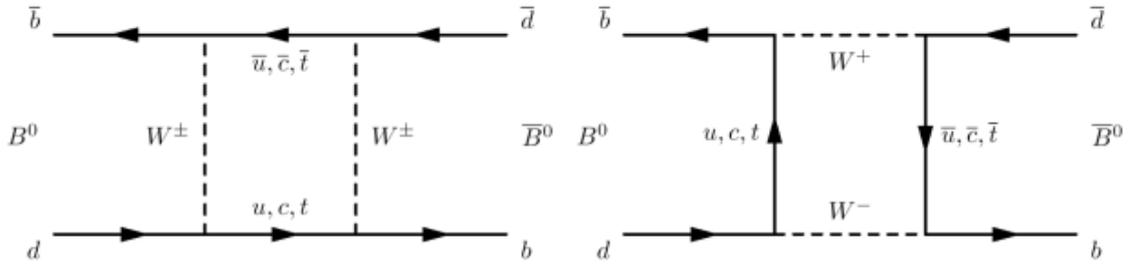
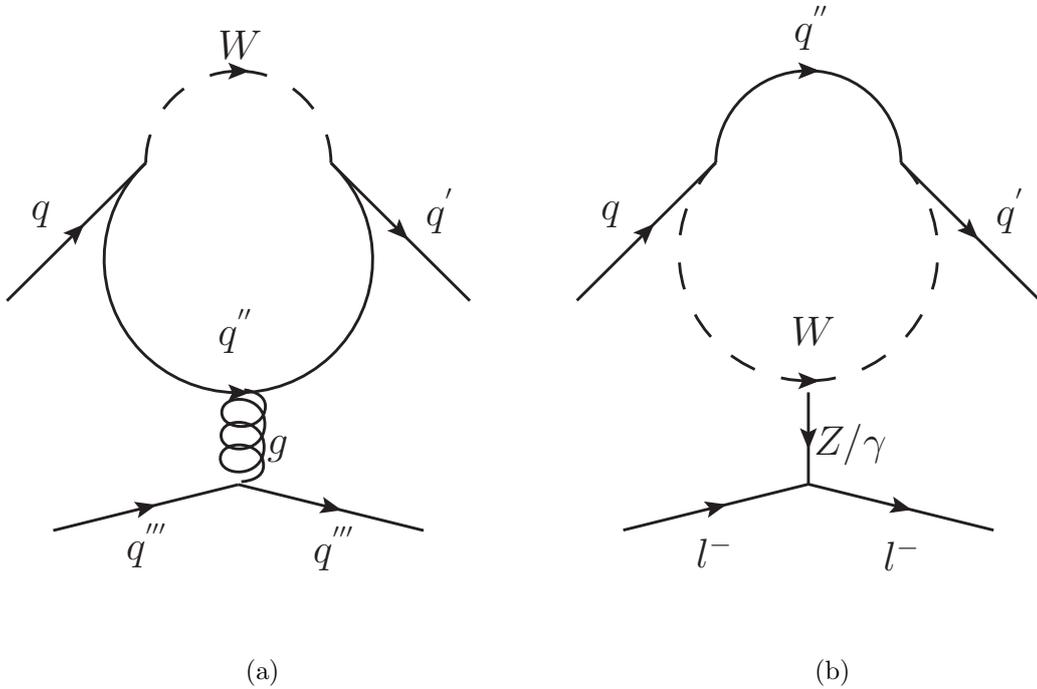

 Figure 9.2: Box diagrams contributing to $B^0 - \bar{B}^0$ mixing [4].


Figure 9.3: Feynman diagrams depicting gluonic (a) and electroweak (b) penguin decays.

9.1.4 Flavour changing neutral currents

As the Z^0 boson is its own anti-particle, its charge and all flavour quantum numbers are equal to zero, and its exchange cannot lead to a change of flavour. Thus, flavour changing neutral currents are not allowed in the Standard Model in tree-level processes. Therefore, the lowest-level processes to mediate a change of flavour without transferring charge are appearing at loop-level. The loop involves the emission of a

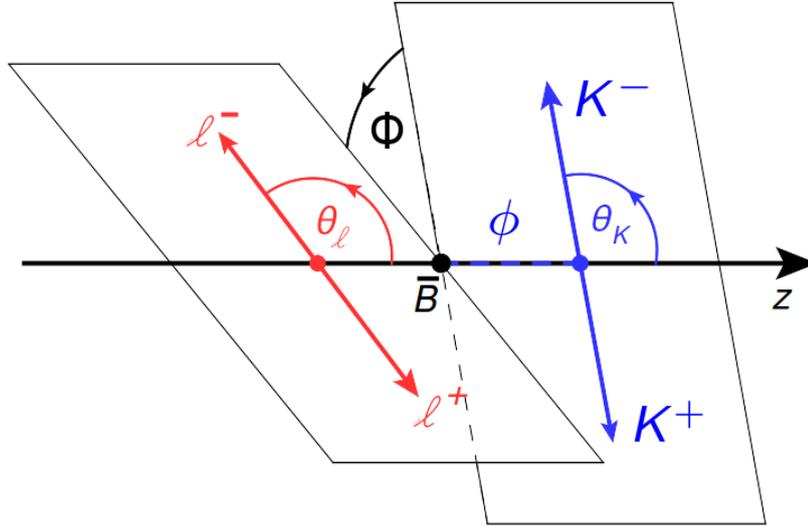


Figure 9.4: Illustration of the decay angles as defined for the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$, using a more general definition for all leptons.

W^\pm boson, which is reabsorbed after the radiation of at least one additional particle either on the boson or quark side of the transition. These processes are often referred to as "Penguin decays" [40], due to the diagram depicting the process resembling the shape of this specific animal. While only a Z^0 boson or photon can be radiated off of the vector boson side of the loop, the quark side of the loop can also emit a gluon. In this case the decay is classified as a gluonic penguin decay, otherwise an electroweak penguin. Figure 9.3 shows the corresponding Feynman diagrams for both of these possibilities.

As this thesis presents the measurement of an electroweak penguin decay, the following section will focus on the theoretical description of them, with focus on the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay investigated in this thesis.

9.2 Phenomenology of $b \rightarrow s \ell^+ \ell^-$ transitions

The phase space of the decay of a B_s^0 meson into a four-particle final state is in general described by 16 parameters. Requiring all final-state particles to be on-shell and assuming four-momentum conservation, this is reduced to eight free parameters, of which three are the Euler angles. This leads to a set of five physical parameters, further reduced to four when requiring two of the final-state particles to originate from the decay of an on-shell particle, such as the $\phi \rightarrow K^+ K^-$ decay in which the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay is reconstructed. Thus, the decay $B_s^0 \rightarrow \phi(\rightarrow K^+ K^-) \mu^+ \mu^-$ is fully described by four parameters, the invariant dimuon mass squared q^2 , and

three angles. The decay angles are defined as follows: θ_l is the angle between the negatively charged muon momentum and the B_s^0 -meson momentum in the dimuon rest frame; θ_K is the angle between the negatively charged kaon and the B_s^0 meson in the ϕ meson rest frame; Φ is the relative angle between the planes spanned by the dimuon system and the two kaons in the B_s^0 rest frame, respectively. They are illustrated in Fig. 9.4. Definitions and calculations illustrated in this section are largely taken from Ref. [41]. The angular definitions can be expressed by

$$\cos \theta_l = \frac{\vec{p}_{\mu^-}^{\mu\mu} \cdot \vec{p}_{B_s^0}^{\mu\mu}}{|\vec{p}_{\mu^-}^{\mu\mu}| \cdot |\vec{p}_{B_s^0}^{\mu\mu}|}, \quad (9.15)$$

$$\cos \theta_K = \frac{\vec{p}_{K^-}^{KK} \cdot \vec{p}_{B_s^0}^{KK}}{|\vec{p}_{K^-}^{KK}| \cdot |\vec{p}_{B_s^0}^{KK}|}, \quad (9.16)$$

$$\cos \Phi = \vec{n}_{K^-K^+}^{B_s^0} \cdot \vec{n}_{\mu^-\mu^+}^{B_s^0}, \quad (9.17)$$

and

$$\sin \Phi = (\vec{n}_{K^-K^+}^{B_s^0} \times \vec{n}_{\mu^-\mu^+}^{B_s^0}) \cdot \frac{\vec{p}_{K^-K^+}^{B_s^0}}{|\vec{p}_{K^-K^+}^{B_s^0}|}, \quad (9.18)$$

where the upper indices of the momentum and normal vectors refer to the rest frame in which they are evaluated. From the distributions of these decay angles, a large set of observables can be extracted. These observables are sensitive to physics beyond the description of the Standard Model. The following sections describe the Standard Model description of the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay from which these observables arise.

9.2.1 Effective field theory

For the description of flavour-changing neutral currents it is often advantageous to use an effective field theory in order to separate long-distance and short-distance effects. This is helpful when calculating not only the quark-level transition, but also including QCD effects of the bound hadron state. In $b \rightarrow s \ell^+ \ell^-$ transitions, the dynamics of the loop are dominated by the high masses of the virtual t quark and W boson. As they are much larger than the mass of the involved on-shell quarks and leptons, the interaction can be viewed as point-like. This is similar to the Fermi theory of the neutron decay, where the W boson mass is much larger than the mass of the quarks and leptons. For that reason, no knowledge of the W boson and its propagator is needed for a precise description of the process, and it can be calculated effectively as a point-like interaction. A theoretical model derived from such a simplification is often called an Effective Field Theory, valid at scales much smaller than the masses of the neglected virtual particles.

The effective Hamiltonian can be derived by interpreting the interaction as point-like, which corresponds to developing the short-distance expansion ($x \rightarrow 0$) of the time-ordered product of operators $T(\phi(x)\phi(0))$ into a sum of operators multiplied by

their respective effective coupling, the so-called ‘‘Wilson’’ coefficients, as described in Ref. [42]:

$$\int d^4x e^{iq \cdot x} T(\phi(x)\phi(0)) = \sum_i C_i(q^2) \mathcal{O}_i(0). \quad (9.19)$$

Each coefficient and corresponding operator represent a different Lorentz structure, where the short-distance information is encoded in the coefficient, while the dynamics of the low-energy modes is reproduced by the operator. There is a large variety of operators contributing to FCNC transitions, \mathcal{O}_1 - \mathcal{O}_{10} , \mathcal{O}_S , \mathcal{O}_P , and a corresponding \mathcal{O}'_i , for which the chirality of the intermediate current is inverted. However, $b \rightarrow s\ell^+\ell^-$ processes are only significantly affected by the C_7 , C_9 , and C_{10} coefficients and their respective operators. The operators are given by

$$\mathcal{O}_7^{(\prime)} = \frac{e}{g^2} m_b (\bar{s} \sigma_{\mu\nu} P_{R(L)} b) F^{\mu\nu}, \quad (9.20)$$

$$\mathcal{O}_9^{(\prime)} = \frac{e^2}{g^2} (\bar{s} \gamma_\mu P_{L(R)} b) (\bar{\mu} \gamma^\mu \mu), \quad (9.21)$$

and

$$\mathcal{O}_{10}^{(\prime)} = \frac{e^2}{g^2} (\bar{s} \gamma_\mu P_{L(R)} b) (\bar{\mu} \gamma^\mu \gamma_5 \mu), \quad (9.22)$$

where \bar{s} is the field of the outgoing s quark, b is the field of the incoming b quark, $F^{\mu\nu}$ the electromagnetic field strength tensor, and μ and $\bar{\mu}$ the muon fields. $P_{L/R}$ are the left- and right-handed projection operators, induced by the coupling to the weak gauge bosons. γ_i are the Dirac gamma matrices. As can be seen from these definitions, \mathcal{O}_7 represents a FCNC with the radiation of a photon, which occurs in the Standard Model via a photonic penguin decay. It affects $b \rightarrow s\ell^+\ell^-$ transitions due to the potential decay of the virtual photon to two muons. The operators \mathcal{O}_9 and \mathcal{O}_{10} correspond to electroweak penguin decays, where the current is mediated via a vector- and axial-vector current, respectively. The coefficients and corresponding operators enter the effective Hamiltonian via

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i C_i(q^2) \mathcal{O}_i(0), \quad (9.23)$$

where G_F is the Fermi coupling constant, and V_{tb} V_{ts}^* are the CKM matrix elements involved. The Wilson coefficients at the weak scale are obtained from the matching the values of the full electroweak theory to \mathcal{H}_{eff} . At scales below the W -boson mass the coefficients can be calculated using the renormalisation group evolution assuming SM dynamics [43], and at $\mu_s = m_b$ they are equal to [44]

$$\begin{aligned} C_7^{SM}(m_b) &= -0.3, \\ C_9^{SM}(m_b) &= +4.2, \\ C_{10}^{SM}(m_b) &= -4.2. \end{aligned} \quad (9.24)$$

Potential sources of Physics beyond the Standard Model would modify this Hamiltonian depending on which type of operator the New Physics current corresponds to as

$$\Delta\mathcal{H}_{NP} = \frac{c_{NP}}{\Lambda_{NP}^2} \mathcal{O}_i = C_i^{NP} \mathcal{O}_i, \quad (9.25)$$

where c_{NP} and Λ_{NP} are the coupling strength and energy scale of the New Physics model, respectively, and the index i refers to its Lorentz structure. These New Physics contributions and corresponding couplings can in general be different for each quark flavour and lepton generation, allowing for individual $C_i^{NP,q/\ell}$. In the calculation of the elements of the transition matrix, the Wilson coefficients often appear in similar combinations. For that reason, often effective Wilson coefficients are defined as linear combinations. These definitions will for simplicity not be given here, but are described in full length Ref. [41]. For the q^2 regions investigated in this thesis for the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ the relevant effective Wilson coefficients are at leading order given by $C_{7,9,10}^{eff} \approx \frac{4\pi}{\alpha_s} C_{7,9,10}$.

9.2.2 Hadronic form factors

Besides the quark-level $b \rightarrow s \ell^+ \ell^-$ transition, the QCD mechanics of gluons and virtual sea quarks affecting the inner dynamics of the B_s^0 and ϕ mesons and the $B_s^0 \rightarrow \phi$ transition are important in order to describe electroweak penguin decays. These hadronic effects can be parametrised through seven q^2 -dependent form factors: one vector-current form factor V and each three axial-vector and tensor form factors, $A_{0,1,2}$ and $T_{1,2,3}$, respectively. There are different approaches to calculate these hadronic form factors.

For large energies of the ϕ meson, *i.e.* low q^2 values, light-cone sum rules can be used, as described in Ref. [28]. Complementary to this, for energies where the ϕ meson is almost at rest in the B_s^0 -meson rest frame, corresponding to high q^2 , lattice QCD calculations can be performed [45]. To improve the precision of the form factor calculation for intermediate values of q^2 , these two methods can be combined, taking correlations into account, as done as well in Ref. [28].

The relative uncertainties of these form factor calculations are typically about 10-15%, leading to 20-50% relative uncertainties in the predicted values for the observables, and are thus the by far dominant contribution to theoretical uncertainties in Standard Model predictions.

9.2.3 Transversity amplitudes

The polarisation of the final state $\phi(\rightarrow K^+ K^-) \mu^+ \mu^-$ is described by the seven complex transversity amplitudes $A_0^{L,R}$, $A_{\parallel}^{L,R}$, $A_{\perp}^{L,R}$ and A_t [41]. They are derived by considering the decay $B_s^0 \rightarrow \phi V^*$, where the virtual Z^0 boson or photon is denoted as V^* . Here A_0 refers to longitudinal polarisations of the ϕ meson and the virtual particle, A_{\perp} to perpendicular transverse polarisations, and A_{\parallel} to parallel transverse polarisations. These polarisation modes are sketched in Fig. 9.5. As the virtual

particle is in general off-shell, it can additionally be time-like polarised, corresponding to the amplitude A_t . However, this amplitude is heavily suppressed due to the small lepton mass. The values of the transversity amplitudes are dependent on q^2 , the effective Wilson coefficients, and the hadronic form factors:

$$\begin{aligned}
A_0^{L,R} &= -\frac{N}{2m_\phi\sqrt{qsq}} \left\{ \left[(C_9^{eff} - C_9^{eff'}) \mp (C_{10}^{eff} - C_{10}^{eff'}) \right] \right. \\
&\quad \times \left[(m_{B_s^0}^2 - m_\phi^2 - q^2) \cdot (m_{B_s^0} + m_\phi) \cdot A_1(q^2) - \lambda \frac{A_2(q^2)}{m_{B_s^0} + m_\phi} \right] \\
&\quad \left. + 2m_{B_s^0} (C_7^{eff} - C_7^{eff'}) \left[(m_{B_s^0}^2 + 3m_\phi - q^2) \cdot T_2(q^2) - \frac{\lambda}{m_{B_s^0}^2 - m_\phi^2} \cdot T_3(q^2) \right] \right\}, \\
A_{\parallel}^{L,R} &= -N\sqrt{2}(m_{B_s^0}^2 - m_\phi^2) \left\{ \left[(C_9^{eff} - C_9^{eff'}) \mp (C_{10}^{eff} - C_{10}^{eff'}) \right] \frac{A_1(q^2)}{m_{B_s^0} - m_\phi} \right. \\
&\quad \left. + \frac{2m_{B_s^0}}{q^2} (C_7^{eff} - C_7^{eff'}) T_2(q^2) \right\}, \\
A_{\perp}^{L,R} &= N\sqrt{2}\lambda^{1/2} \left\{ \left[(C_9^{eff} + C_9^{eff'}) \mp (C_{10}^{eff} + C_{10}^{eff'}) \right] \frac{V(q^2)}{m_{B_s^0} + m_\phi} \right. \\
&\quad \left. + \frac{2m_{B_s^0}}{q^2} (C_7^{eff} + C_7^{eff'}) T_1(q^2) \right\}, \\
A_t &= \frac{N}{\sqrt{q^2}} \lambda^{1/2} \left\{ 2(C_{10}^{eff} - C_{10}^{eff'}) + \frac{q^2}{m_\mu} (C_P - C'_P) \right\} A_0(q^2),
\end{aligned} \tag{9.26}$$

where the L and R refer to the chirality of the virtual Z^0 boson or photon and

$$N = V_{tb}V_{ts}^* \left[\frac{G_F^2 \alpha^2}{3 \cdot 2^{10} \pi^5 m_{B_s^0}^3} q^2 \lambda^{1/2} \beta_\mu \right]^{1/2}, \tag{9.27}$$

$$\lambda = m_{B_s^0}^4 + m_\phi^4 + q^4 - 2(m_{B_s^0}^2 m_\phi^2 + m_\phi^2 q^2 + m_{B_s^0}^2 q^2), \tag{9.28}$$

and

$$\beta_\mu^2 = (1 - 4m_\mu^2/q^2) \tag{9.29}$$

are used. The effective Wilson coefficients are in general q^2 -dependent, which has been dropped in the equations above for simplification. These seven transversity amplitudes can be used to fully express the amplitude of the sequential decay $B_s^0 \rightarrow \phi V^* (\rightarrow \mu^+ \mu^-)$, including all contributions from the Wilson coefficients of interest.

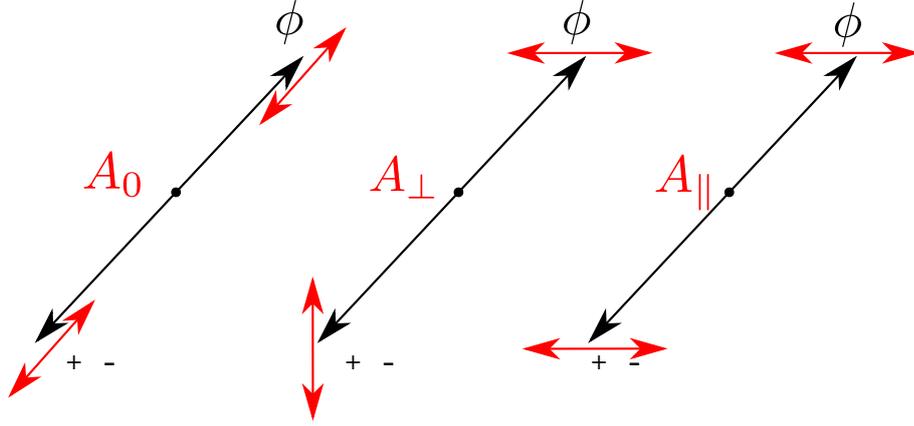


Figure 9.5: Illustration of the polarisation modes of the final state of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$.

9.2.4 Angular observables

The investigation of the distributions of the decay angles allows to access information of the Lorentz structure of the electroweak currents involved in the decay process. The full differential decay rate for the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay is given as a function of these angles by [41]:

$$\begin{aligned}
 \frac{d^4\Gamma(B_s^0 \rightarrow \phi \mu^+ \mu^-)}{dq^2 d\cos\theta_l d\cos\theta_K d\Phi} &= \sum_i J_i(q^2) f_i(\cos\theta_l, \cos\theta_K, \Phi) \\
 &= \frac{9}{32\pi} \left[J_1^s \sin^2\theta_K \cdot \left(1 + \frac{1}{3} \cos(2\theta_l)\right) + J_1^c \cos^2\theta_K \cdot \left(1 - \cos(2\theta_l)\right) \right. \\
 &\quad + J_3 \sin^2\theta_K \sin^2\theta_l \cos(2\Phi) + J_4 \sin(2\theta_K) \sin(2\theta_l) \cos\Phi \\
 &\quad + J_5 \sin(2\theta_K) \sin\theta_l \cos\Phi + J_6^s \sin^2\theta_K \cos\theta_l \\
 &\quad + J_7 \sin(2\theta_K) \sin\theta_l \sin\Phi + J_8 \sin(2\theta_K) \sin(2\theta_l) \sin\Phi \\
 &\quad \left. + J_9 \sin^2\theta_K \sin^2\theta_l \sin(2\Phi) \right].
 \end{aligned} \tag{9.30}$$

The angular functions f_i are given by spherical harmonics, while the q^2 -dependent coefficients J_i are calculated from combinations of the seven complex polarisation

amplitudes $A_0^{L,R}$, $A_{\parallel}^{L,R}$, $A_{\perp}^{L,R}$ and A_t via:

$$J_1^s = \frac{2 + \beta_\mu^2}{4} [|A_{\perp}^L|^2 + |A_{\parallel}^L|^2 + |A_{\perp}^R|^2 + |A_{\parallel}^R|^2] + \frac{4m_\mu^2}{q^2} \Re(A_{\perp}^L A_{\perp}^{R*} + A_{\parallel}^L A_{\parallel}^{R*}), \quad (9.31)$$

$$J_1^c = |A_0^L|^2 + |A_0^R|^2 + \frac{4m_\mu^2}{q^2} [|A_t|^2 + 2\Re(A_0^L A_0^{R*})], \quad (9.32)$$

$$J_3 = \frac{\beta_\mu^2}{2} [|A_{\perp}^L|^2 - |A_{\parallel}^L|^2 + |A_{\perp}^R|^2 - |A_{\parallel}^R|^2], \quad (9.33)$$

$$J_4 = \frac{\beta_\mu^2}{\sqrt{2}} [\Re(A_0^L A_{\parallel}^{L*}) + \Re(A_0^R A_{\parallel}^{R*})], \quad (9.34)$$

$$J_5 = \sqrt{2}\beta_\mu [\Re(A_0^L A_{\perp}^{L*}) - \Re(A_0^R A_{\perp}^{R*})], \quad (9.35)$$

$$J_6^s = 2\beta_\mu [\Re(A_{\parallel}^L A_{\perp}^{L*}) - \Re(A_0^R A_{\parallel}^{R*})], \quad (9.36)$$

$$J_7 = \sqrt{2}\beta_\mu [\Im(A_0^L A_{\parallel}^{L*}) - \Im(A_0^R A_{\parallel}^{R*})], \quad (9.37)$$

$$J_8 = \frac{\beta_\mu^2}{\sqrt{2}} [\Im(A_0^L A_{\perp}^{L*}) + \Im(A_0^R A_{\perp}^{R*})], \quad (9.38)$$

$$J_9 = \beta_\mu^2 [\Im(A_{\parallel}^{L*} A_{\perp}^L) + \Im(A_{\parallel}^{R*} A_{\perp}^R)], \quad (9.39)$$

For the charge-conjugated decay $\bar{B}_s^0 \rightarrow \phi\mu^+\mu^-$ the coefficients J_i are modified via $J_{1,2,3,4,7} \rightarrow \bar{J}_{1,2,3,4,7}$ and $J_{5,6,8,9} \rightarrow -\bar{J}_{5,6,8,9}$ in Eq. 9.30. Each coefficient \bar{J}_i is equal to the corresponding J_i with all weak phases conjugated. The additional minus sign in the second set of coefficients is caused by the definition of the angles θ_K and θ_l . These are based on the directions of the negatively charged final-state particles for both the B_s^0 -meson and \bar{B}_s^0 -meson decay, while a CP conjugation changes those charges. This leads to a transformation of $\theta_K \rightarrow \theta_K - \pi$ and $\theta_l \rightarrow \theta_l - \pi$ for the \bar{B}_s^0 -meson decay which in turn affects the angular functions. The q^2 -dependent J_i coefficients are tied to both the Wilson coefficients of the effective field theory used to describe the $b \rightarrow s\ell^+\ell^-$ transition and the hadronic form factors through the definitions of the transversity amplitudes.

Experimentally, the final states of B_s^0 and \bar{B}_s^0 decays cannot be distinguished, so only the averaged differential decay rate $d(\Gamma + \bar{\Gamma})/(dq^2 d\cos\theta_l d\cos\theta_K d\Phi)$ can be measured. Here Γ and $\bar{\Gamma}$ refer to the decay rates of the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and

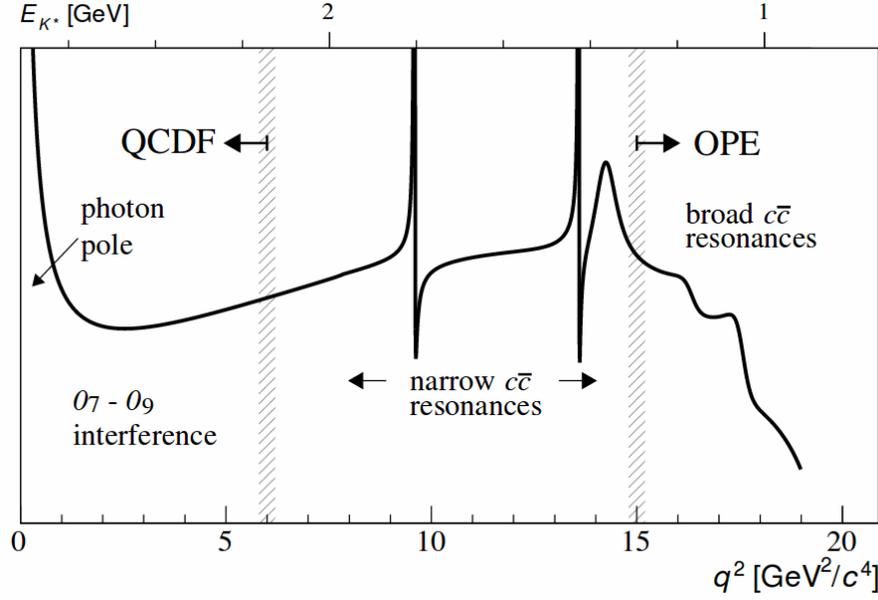


Figure 9.6: Typical q^2 dependency of the decay rate of $b \rightarrow s \ell^+ \ell^-$ processes [42].

$\bar{B}_s^0 \rightarrow \phi \mu^+ \mu^-$ decays, respectively. The averaged differential decay rate is given by

$$\begin{aligned}
 \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{dq^2 d\cos\theta_l d\cos\theta_K d\Phi} &= \sum_i (J_i(q^2) f_i(\cos\theta_l, \cos\theta_K, \Phi) + \bar{J}_i(q^2) \bar{f}_i(\cos\theta_l, \cos\theta_K, \Phi)) \\
 &= \frac{9}{32\pi} \left[\frac{3}{4} (1 - F_L) \sin^2\theta_K \cdot \left(1 + \frac{1}{3} \cos(2\theta_l)\right) \right. \\
 &\quad + F_L \cos^2\theta_K \cdot (1 - \cos(2\theta_l)) + S_3 \sin^2\theta_K \sin^2\theta_l \cos(2\Phi) \\
 &\quad + S_4 \sin(2\theta_K) \sin(2\theta_l) \cos\Phi + A_5 \sin(2\theta_K) \sin\theta_l \cos\Phi \\
 &\quad + A_6^s \sin^2\theta_K \cos\theta_l + S_7 \sin(2\theta_K) \sin\theta_l \sin\Phi \\
 &\quad \left. + A_8 \sin(2\theta_K) \sin(2\theta_l) \sin\Phi + A_9 \sin^2\theta_K \sin^2\theta_l \sin(2\Phi) \right], \tag{9.40}
 \end{aligned}$$

where the q^2 -dependent CP asymmetries $A_i = (J_i - \bar{J}_i)/(\Gamma + \bar{\Gamma})$ and CP averages $S_i = (J_i + \bar{J}_i)/(\Gamma + \bar{\Gamma})$ were introduced. Furthermore, the fraction of longitudinal polarisation of the ϕ meson $F_L := S_1^c$ was introduced and the relation $S_1^c = 3(1 - S_1^s)/4$ used. For the latter relation the muon masses are neglected, implying $\beta = 1$, which is a reasonable assumption for $q^2 > 1 \text{ GeV}^2/c^4$.

9.2.5 Explicit q^2 dependency of the decay rate

Due to the dependency on q^2 , measurements are typically performed differentially in q^2 to disentangle the influence of the different effective couplings (and thus operators)

involved. For $b \rightarrow s\ell^+\ell^-$ transitions, where q^2 is equal to the dimuon mass squared, the dependency of the decay rate on q^2 is sketched in Fig. 9.6 using the example of the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decay. In addition to the Wilson coefficients and hadronic form factors described above, corrections have to be made to include higher-order effects, such as spectator interactions. At low q^2 the energy of the ϕ meson is large in the B -meson rest frame. Thus the principle of QCD factorisation can be applied [46] as $E_\phi \gg \Lambda_{QCD}$. From the QCD factorisation framework corrections are directly calculated for the transversity amplitudes. The affected region with very low q^2 close to the threshold of $4 \cdot m_\mu^2$ is dominated by C_7 , due to the proximity to the so-called photon pole at $q^2 = 0$. In addition to this photon pole also the dimuon resonances J/ψ and $\psi(2S)$ appear in the spectrum at q^2 values of about $9.6 \text{ GeV}^2/c^4$ and $13.6 \text{ GeV}^2/c^4$, respectively. These are connected to tree-level decays involving a $b \rightarrow s\bar{c}$ transition, where the charmonium resonance decays into two muons. As these decays do not happen via $b \rightarrow s\ell^+\ell^-$ transitions, the desired Wilson coefficients cannot be extracted from measuring them, and theory predictions for the influence of the Wilson coefficients in the q^2 regions close to the charmonium resonances thus tend to have large uncertainties. For that reason, these regions are typically not taken into account in measurements.

For q^2 values above the $\psi(2S)$ mass squared the decay rate is dominantly influenced by the Wilson coefficients C_9 and C_{10} . Higher-excited charmonium resonances are not as detrimental to the measurement in this region, as their masses lie above the threshold to allow their decay into a pair of charm hadrons, which is dominant compared to the decay into two muons. Their contributions can be described by a local operator product expansion (OPE), exploiting the large mass of the b quark compared to the energy of the K^{*0} meson [47]. While the OPE does not describe the resonance contributions to the transversity amplitudes locally, their effect is captured when integrating over a sufficiently large range of q^2 , as it is typically done by binned measurements as presented in this thesis.

9.3 New Physics models affecting FCNC transitions

Deviations from Standard Model expectations have been observed in multiple electroweak penguin decays. While possible explanations include a large array of New Physics models, this section will briefly describe two potential explanations which are of particular interest: models incorporating Z' bosons and such including leptiquarks.

9.3.1 Z' boson contributions

The presence of a flavour-changing Z' boson would allow for the $b \rightarrow s\ell^+\ell^-$ transition to occur at tree-level, as sketched in Fig. 9.7(a). The influence of such a boson on the q^2 -dependent values of the Wilson-coefficients C_9 and C'_9 is reviewed in detail for multiple models in Ref. [48]. As these bosons would also affect the mixing of B_s^0 and

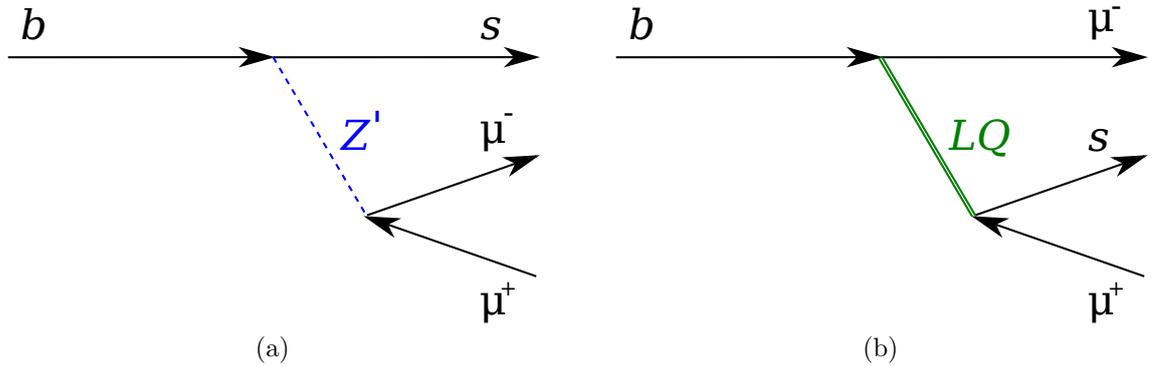


Figure 9.7: Feynman diagrams of the $b \rightarrow s\mu^+\mu^-$ transition when mediated by (a) a Z' boson and (b) a leptoquark.

\bar{B}_s^0 mesons, bounds are induced on their upper masses, suggesting them to be light, in the order of a few TeV/c^2 . However, direct searches performed by the ATLAS and CMS collaborations impose strong bounds from searches for di-lepton and di-jet resonances. For a Z' boson with an invariant mass of around $1 \text{ TeV}/c^2$ these bounds imply a coupling strength of this boson to be at least one order of magnitude below the corresponding couplings of the SM Z boson.

9.3.2 Leptoquark models

The exchange of a leptoquark can induce the $b \rightarrow s\ell^+\ell^-$ transition on tree-level by simultaneously changing quark and lepton flavours, as shown in Fig. 9.7(b). Two leptoquark models have been reviewed in Ref. [49], assuming a scalar particle which does not preserve quark flavours and couples differently to electrons and muons. For those two models the New Physics couplings are introduced in the electron and muon coupling, respectively. The invariant masses of the leptoquarks of the presented models are expected to be larger than $1 \text{ TeV}/c^2$. As these leptoquark couplings are different for electrons and muons by construction, lepton universality measurements are sensitive to them. Naturally, one would also expect their coupling to be strongest to the third generation of leptons, which makes measurements involving τ mesons especially interesting tests of leptoquark models.

10 Analysis Strategy

The goal of the measurement presented in the second part of this thesis is the determination of the differential branching ratio as well as an angular analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$. It is a rare decay with a branching fraction which is expected to be at $\mathcal{B} \approx \mathcal{O}(10^{-7})$ [27]. The tree-level decay $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ is used both as a normalisation mode for the branching fraction measurement as well as a control mode. It has the identical set of final-state particles as the signal mode, but at the same time very low background. Due to being a leading order process the decay mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ happens about 100 times more frequently than the signal mode $B_s^0 \rightarrow \phi \mu^+ \mu^-$. Figure 10.1 shows diagrams for both signal and control mode decays.

The signal decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ is investigated as a function of the dimuon mass

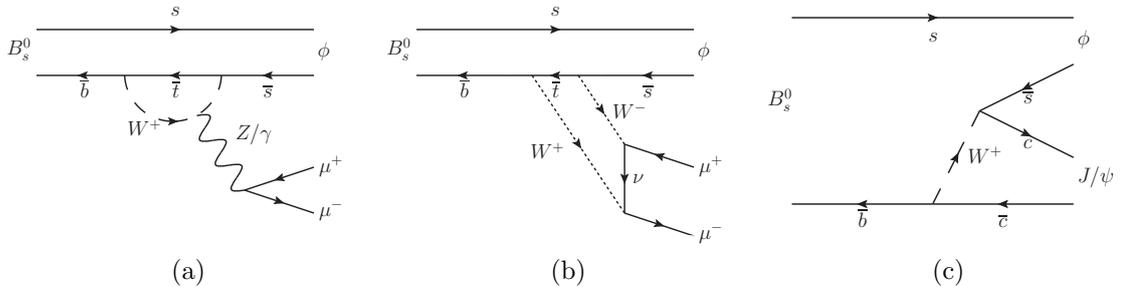


Figure 10.1: Diagrams depicting the decays of the signal mode $B_s^0 \rightarrow \phi \mu^+ \mu^-$ (a),(b), and the control and normalisation mode $B_s^0 \rightarrow J/\psi \phi$ (c).

squared q^2 , as the Wilson coefficients of the effective field theory depend on it. Furthermore, possible effects from new physics beyond the Standard Model are expected to be q^2 dependent as well. All parameters are determined in six narrow and two wide q^2 -bins, as listed in Table 10.1. The two wide bins are chosen as they span a range in which the q^2 dependence is expected to be low. In the q^2 regions of $8.0 < q^2 < 11.0 \text{ GeV}^2/c^4$ and $12.5 < q^2 < 15.0 \text{ GeV}^2/c^4$ the tree-level decays $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ and $B_s^0 \rightarrow \psi(2S) (\rightarrow \mu^+ \mu^-) \phi$ are dominant, respectively, which severely limits the theoretical understanding of the signal decay. For that reason, these bins are not used in the analysis of the signal decay. However, the decay mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$, and thus the q^2 region associated to it, is used as a normalisation mode for the determination of the relative differential branching fraction.

Due to the suppression of the signal decay, a careful selection is required to identify it and discriminate against potential background sources. The selection happens throughout multiple steps of both the data taking process (“online”) as well as

	$q^2 [\text{GeV}^2/c^4]$
narrow	0.1 – 2.0
	2.0 – 5.0
	5.0 – 8.0
	11.0 – 12.5
	15.0 – 17.0
	17.0 – 19.0
wide	1.0 – 6.0
	15.0 – 19.0

Table 10.1: Ranges of the six narrow and two wide q^2 bins used in this analysis.

further data processing (“offline”). The selection process is optimised using the decay channel $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ as a control mode.

The determination of absolute efficiencies to detect and reconstruct a specific decay are relying heavily on simulated data samples, which can lead to large associated systematic uncertainties. Thus, the differential branching ratio of the signal mode is measured relative to the normalisation mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$. This requires the determination of relative efficiencies between signal and normalisation mode, where systematic uncertainties of the efficiency determination from simulation largely cancel.

The simulation used by the LHCb experiment and required for the efficiency determination describes most decays, their topologies and the corresponding detector response reasonably well. However, there are some observables which are not modelled perfectly. This is corrected for by the use of data-driven techniques, which evaluate particle identification and reconstruction efficiencies depending on the particle kinematics and detector occupancies.

In order to evaluate the relative differential branching ratio the relative efficiencies have to be combined with the event yields for both signal and normalisation mode. These are extracted from fits to the invariant B_s^0 mass, reconstructed from the $(K^+ K^- \mu^+ \mu^-)$ final state. Analogous to the efficiencies, the yields of the signal mode are determined individually for each q^2 bin. As partially reconstructed and misidentified background sources are expected to be very small, they are not included in the fit model, but instead treated as a source of systematic uncertainty. The differential relative branching fraction for a given q^2 bin is calculated via

$$\frac{1}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)} \frac{d\mathcal{B}(B_s^0 \rightarrow \phi \mu^+ \mu^-)}{dq^2} = \frac{\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)}{q_{max}^2 - q_{min}^2} \cdot \frac{N_{\phi \mu^+ \mu^-}}{N_{J/\psi \phi}} \cdot \frac{\varepsilon_{tot}^{J/\psi \phi}}{\varepsilon_{tot}^{\phi \mu^+ \mu^-}} \quad (10.1)$$

Here q_{max}^2 and q_{min}^2 are the upper and lower bounds of the q^2 bin, $N_{\phi \mu^+ \mu^-}$ and $N_{J/\psi \phi}$ are the yields, and $\varepsilon_{tot}^{\phi \mu^+ \mu^-}$ and $\varepsilon_{tot}^{J/\psi \phi}$ are the reconstruction and selection efficiencies of the signal and normalisation mode, respectively. Furthermore, the branching

fraction of the decay $J/\psi \rightarrow \mu^+ \mu^-$ is used as an external parameter [15].

The angular analysis investigates the distributions of the three decay angles θ_K , θ_l and Φ (as shown in Fig. ??) to derive eight q^2 dependent observables: the fraction of longitudinal polarisation F_L , three CP -averages S_3 , S_4 and S_7 , and four CP -asymmetries A_5 , A_6 , A_8 and A_9 . These variables are extracted from a four-dimensional simultaneous fit to the decay angles and the invariant B_s^0 meson candidate mass. The angular acceptance, the sensitivity to detect signal decays dependent on the three decay angles θ_l , θ_K and Φ and the dimuon mass squared q^2 , is determined using a sample of fully simulated $B_s^0 \rightarrow \phi \mu^+ \mu^-$ events. It is used to account for all detector and selection effects which distort the angular distributions.

Several sources of systematic uncertainties can affect the measurement of the differential branching fraction as well as the analysis of the angular distributions. They are investigated for each q^2 bin and observable, and compared to the statistical uncertainties.

11 Selection of the signal candidates

Physics processes are typically subject to backgrounds, either due to mis-identification of the final-state particles, or random combinations mimicking the signal decay. In order to select signal candidates and reject these background sources, a candidate selection is applied in multiple steps during data processing, which are presented in this chapter. Events are recorded that fulfil the trigger requirements, *e.g.* when decays are found fulfilling certain decay topologies or passing basic kinematic requirements to discriminate them from the most abundant background processes. All events recorded that way by the LHCb experiment are further processed by the centralised event reconstruction and the Stripping framework. Candidates which fulfil the requirements of the Stripping line used to select typical $B_s^0 \rightarrow K^+K^-\mu^+\mu^-$ decays are further required to pass an additional analysis selection. This selection consists of a slightly tightened pre-selection and a multivariate classifier, to further reduce background levels. This chapter depicts all these selection steps in detail and gives a brief overview over the final sample of selected events.

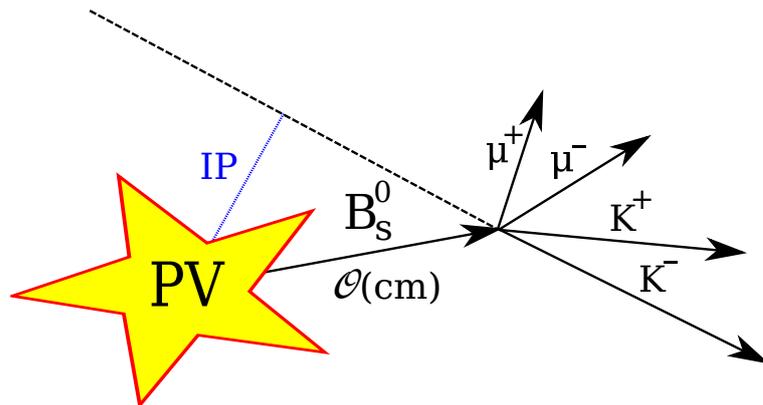


Figure 11.1: Sketch of the production and decay topologies of a typical $B_s^0 \rightarrow \phi\mu^+\mu^-$ decay.

Figure 11.1 shows a sketch of the $B_s^0 \rightarrow \phi\mu^+\mu^-$ production and decay topology. The B_s^0 meson typically travels a few centimetres in the detector before decaying to two oppositely charged muons and a ϕ meson, which decays to a K^+K^- pair. Due to the extremely short decay time of the ϕ meson of $\mathcal{O}(10^{-7} \text{ fs})$, its decay vertex cannot be separated from the B_s^0 decay vertex with the precision of the VELO, and the kaons appear to originate directly from the B_s^0 decay. Events are selected with four final-state charged tracks originating from the same displaced vertex. These tracks are required to be identified as pairs of oppositely charged muons and kaons, respectively. The decay vertex is required to be sufficiently detached from the primary interaction vertex to ensure none of the final-state particles originate from

this primary interaction instead of the signal decay. For that reason, requirements are imposed on the impact parameter (IP), the minimal distance of a particle trajectory to the PV (as shown in Fig.11.1). The four-momentum of the B_s^0 -meson candidate is calculated by summing over the measured momenta of all final-state particles, and its invariant mass is equal to the magnitude of the four-momentum vector.

11.1 Background sources

The sources of background affecting the selection of $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays can be categorised into two groups: combinatorial and peaking background.

Combinatorial background is caused by random combination of real or fake tracks, where the reconstructed invariant mass of combining a set of those tracks is close to the nominal B_s^0 -meson mass. In contrast to the signal decay, the reconstructed invariant mass distribution of those fake candidates does not have clear peaking structures. Their distributions in other observables can typically be obtained from candidates far below or above the nominal B_s^0 -meson mass, in the so-called “sidebands”. These random combinations often originate from particles created directly in the primary interaction, which is why their contributions can be reduced by requiring final-state particles to be detached from the primary interaction point and fulfil minimal (transverse) momentum requirements. Combinatorial background can be further reduced by exploiting the distributions in observables obtained from the reconstructed invariant mass sidebands, which for this analysis is done by a multivariate classifier, described below.

Background processes, where final-state particles are either mis-identified or not reconstructed can mimic the signal decay, while the reconstructed invariant mass distribution has peaking structures. Thus they are called “peaking” backgrounds. For the analysis of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$, multiple decay modes are of special importance as a source of peaking background, where the CP -conjugated equivalent is implied. The baryonic decay $\Lambda_b^0 \rightarrow \Lambda(1520)(\rightarrow pK^-)\mu^+\mu^-$ can mimic the signal decay when the final-state proton is mis-identified as a K^+ meson. The mass difference between the proton and the K^+ is close to the mass difference between the Λ_b^0 and B_s^0 meson, which is why the mis-reconstructed invariant mass distribution is very similar to the distribution of the signal decay. In a similar manner the signal decay can be faked by the electroweak penguin decay $B^0 \rightarrow K^{*0}(\rightarrow K^-\pi^+)\mu^+\mu^-$ where the final-state pion is identified as a kaon. As B^0 mesons are about four times more frequently produced than B_s^0 mesons, this decay occurs more often than the signal decay. Furthermore, fully hadronic decay modes such as $B_s^0 \rightarrow D_s^-(\rightarrow K^+K^-\pi^-)\pi^+$ can affect the measurement, when multiple final-state hadrons are mis-identified as muon tracks. While the muon identification is in general very clean, hadrons that decay in flight to a muon and muon-neutrino can lead to a false positive identification by the muon system, if the flight path of the muon is similar to the hadron track. Another possible pollution of the data sample can originate from the semileptonic cascade decays $B_s^0 \rightarrow D_s^+(\rightarrow \phi\mu^-\bar{\nu}_\mu)\mu^+\nu_\mu$ and $B^0 \rightarrow D^-(\rightarrow K^{*0}\mu^-\bar{\nu}_\mu)\mu^+\nu_\mu$. For

these decays all reconstructed final-state particles are identical to the signal decay, as the neutrinos cannot be measured by the LHCb detectors. However, due to the missing momentum of the neutrinos, the reconstructed invariant mass of the $K^+K^-\mu^+\mu^-$ system is typically lower than the nominal B_s^0 -meson mass.

11.2 Boosted Decision Trees

Boosted Decision Trees (BDT) are used in multiple steps of the selection requirements of the analysis presented. A BDT is a multivariate analysis technique, exploiting correlations between variables instead of using independent selection requirements for each variable in order to reduce background contributions.

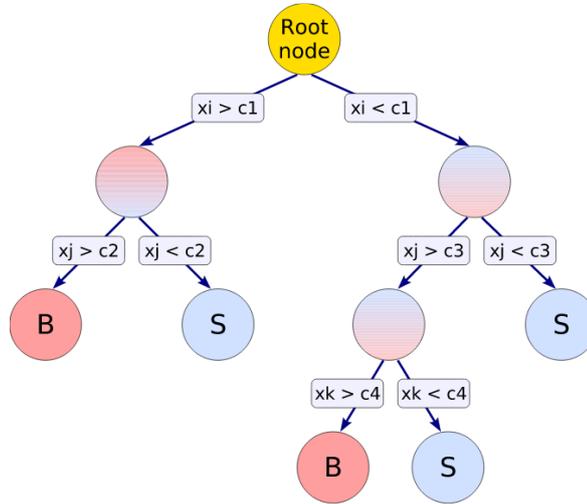


Figure 11.2: Schematic view of a decision tree [50].

Figure 11.2 shows a schematic view of a basic decision tree. It uses a sequence of binary decisions, called branches, to categorise a sample into signal-like and background-like events [51]. Each of these binary decisions splits the dataset further, where the split conditions are derived from dedicated pure signal and background samples. The process to obtain these split conditions is called “training”. The trained decision tree classifies each event according to its branches as either signal or background, corresponding to classification values of +1 and -1, respectively.

The disadvantage of using a single decision tree is the binary classification and that it can pick up statistical fluctuations of the training samples, which is called “overtraining”. This might lead to unstable results due to small differences in the samples used in the training process. The quality of the multivariate selection can be significantly improved by combining multiple decision trees. A very popular

combination technique is the so-called Boosting: The full-sized training sample is used for the first decision tree, which provides the highest separation powers with the given training samples. Afterwards falsely classified events are given an increased weight (depending on the overall number of mis-classified events) for the training of the new tree, increasing sensitivity of the training process to these events. This procedure is repeated iteratively, typically $\mathcal{O}(1000)$ times, creating a newly trained tree in each step. Each tree is assigned a score depending on the amount of correctly classified events. The final event classification of the BDT is given by the score-weighted average of all individual tree classification values for a given event, leading to a classification value between -1 (pure background) and +1 (pure signal). While the resulting distribution for the final classification value for signal and background events is not as strictly separated, the likelihood of wrong classification is strongly reduced.

11.3 Trigger selection

As described in Sec. 2.2.5, the LHCb trigger system consists of a hardware trigger stage (L0) and two software trigger stages (HLT1 and HLT2), condensing the 40 MHz rate of interactions down to a rate of about 5 kHz saved to disk. The trigger lines, sequences of selection and reconstruction algorithms, are used to select specific decay topologies. Especially in the HLT2 stage these requirements can be aimed towards very specific decay modes, which might bias variable distributions for other decay modes, *e.g.* only select decays of long-lived particles. To avoid this, the analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ uses only events that passed trigger lines corresponding to the proper decay topologies. As the control mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ consists of the same final-state particles as the signal mode, the selection of trigger lines has a comparable efficiency of selecting these decays. Table 11.1 shows a list of the trigger lines used in the respective trigger stages. Signal candidates are required to have passed at least one line in each stage, and all of the stages. Typically, trigger lines of the first and second software stage are only run on candidates passing trigger line requirements previous in the previous stages, to minimise the necessary computation time.

The two L0 trigger lines select events with one or two tracks, respectively, which originate from close to the interaction region. These tracks are identified as muons by requiring a minimal number of hits in the muon system. Additional to that, minimum transverse momentum requirements are imposed in order to suppress combinatorial background sources: in case of the L0Muon line, the muon is required to have a transverse momentum p_T of larger than 1.48 GeV/ c (in 2011, 1.76 GeV/ c in 2012), for the L0DiMuon line the product of the p_T of the two muons has to be larger than $1.68 (\text{GeV}/c)^2$ ($2.56 (\text{GeV}/c)^2$).

The trigger lines of the first software stage, HLT1, used in this analysis are searching for (muon) tracks with a high reconstruction quality. The reconstruction algorithms for these tracks uses the response of the Tracking Stations and thus is much more

Table 11.1: Trigger lines used in the different trigger stages.

Trigger stage	Line name
L0	L0Muon L0DiMuon
HLT1	Hlt1TrackAllL0 Hlt1TrackMuon Hlt1DiMuonLowMass Hlt1DiMuonHighMass Hlt1SingleMuonHighPT
HLT2	Hlt2Topo(2,3,4)BodyBBDT Hlt2TopoMu(2,3,4)BodyBBDT Hlt2SingleMuon Hlt2DiMuonDetached Hlt2DiMuonDetachedHeavy

precise than what is available in the hardware trigger. Events are required to either activate one of the track trigger lines, `Hlt1TrackAllL0` and `Hlt1TrackMuon`, or the muon trigger lines. The `Hlt1TrackAllL0` trigger line selects events containing tracks with a certain minimum momentum p and p_T that are significantly detached from the primary interaction vertex, and are thus likely to originate from a B -meson decay. The `Hlt1TrackMuon` trigger line has looser track quality and kinematic restrictions compared to `Hlt1TrackAllL0`, but requires the track to have activated one of the L0(Di)Muon triggers. The dimuon trigger lines `Hlt1DiMuonHighMass` and `Hlt1DiMuonLowMass` select events with pairs of muon tracks with good reconstruction quality, where the mass of the dimuon system is above $2700 \text{ MeV}/c^2$ and $1000 \text{ MeV}/c^2$, respectively, where the latter line also requires both muon tracks to be significantly detached from the primary vertex. Finally, the `Hlt1SingleMuonHighPT` trigger line gets activated by well-reconstructed tracks which previously passed the L0Muon requirements and have a momentum larger than $8 \text{ GeV}/c$ and a p_T of larger than $4.8 \text{ GeV}/c$. This combination of trigger lines allows for a wide coverage of range in the dimuon mass squared q^2 .

In the second stage of the software trigger, HLT2, the used trigger lines can be classified roughly into two categories, the topological triggers `Hlt2Topo(Mu)(2,3,4)BodyBBDT` and the muon triggers. The topological lines use a Boosted Decision Tree to select events where decays are found which are classified as being compatible with the decay topology of the decay of a B meson to two, three, or four final-state particles, respectively. In addition, the "TopoMu" lines specifically search for such decays with muons in the final state. All these BDT selections use the same set of kinematic and quality parameters of the final-state particles and the reconstructed meson state. The specific training results and BDT classification values required vary for each

trigger line and the time of data taken, values are given in Ref. [52]. The single-muon trigger line selects events with a well reconstructed muon track detached from the primary vertex, which is required to have activated the Hlt1TrackMuon trigger line and has p_T larger than 1300 MeV/ c . Due to the large output rate of this trigger line it is prescaled to 50%, meaning that only half of the events passing the requirements activate this line (by random choice), due to the limited computational capacities of the trigger system. The dimuon trigger lines are set up in a very similar way to the lines used in HLT1: dimuon objects are created from muons detached from the primary vertex and have a minimum p_T of 500 MeV/ c , while additional criteria on the vertex reconstruction quality are applied. The dimuon object is required to have a mass of larger than 1000 MeV/ c^2 and 2950 MeV/ c^2 for the Detached and DetachedHeavy trigger lines, respectively. The combination of these trigger lines allows coverage over a large region of the dimuon mass squared q^2 .

11.4 Stripping selections

The Stripping framework is a centralised selection stage which is executed on all events recorded by the LHCb experiment after passing the hardware and software trigger stages and being processed through the full offline event reconstruction. Similar to the trigger, the different selection algorithms in the Stripping are divided into Stripping lines. These lines select candidates for specific decays, which can be either fully or partially reconstructed. The Stripping line used for this analysis searches for $B \rightarrow X\mu^+\mu^-$ decays, where the X resonance can be any combination of two oppositely charged final-state particles. Table 11.2 lists the cuts applied in this Stripping line. The constraint on the reconstructed invariant mass of the X system is set to be smaller than 6200 MeV/ c^2 , which is a rather loose criterion. The reason for this is that this Stripping line is used to select a large array of decay modes, *e.g.* the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$, which have comparable decay topologies but include resonances of very different invariant masses. In addition the X meson candidate is required to be significantly displaced from the PV by demanding a flight distance significance, χ_{FD}^2 , of larger than 9. However, the selection of the B -meson candidate is much tighter. The reconstructed invariant mass of the B_s^0 -meson candidate m_{rec} is required to be in the range $4900 \text{ MeV}/c^2 < m_{\text{rec}} < 7000 \text{ MeV}/c^2$, and the χ^2 of the decay vertex has to be smaller than 8. The direction angle between the B flight path, determined from primary and decay vertex positions, and the momentum direction of the combined final state is required to have a cosine of larger than 0.9999. Furthermore, the B -meson candidate is required to be compatible to originating from the primary vertex by demanding the impact parameter significance squared to be smaller than 16, as well as to have travelled a significant distance in the detector before decaying, requiring χ_{FD}^2 of larger than 121.

The invariant mass requirement of the dimuon system is chosen as $m_{\mu^+\mu^-} < 7100 \text{ MeV}/c^2$ to be compatible with originating from the B decay, allowing additional resolution effects. The flight distance requirement on this system is identical

Table 11.2: Selection cuts applied in the Stripping stage for $B_s^0 \rightarrow \phi\mu^+\mu^-$ candidates.

Particle	Cut
B_s^0	$\chi_{\text{vertex}}^2 < 8$ $\chi_{\text{IP}}^2 < 16$ $\cos(DIRA) > 0.9999$ $\chi_{\text{FD}}^2 > 121$ $4900 \text{ MeV}/c^2 < m < 7000 \text{ MeV}/c^2$
X	$m < 6200 \text{ MeV}/c^2$ $\chi_{\text{FD}}^2 > 9$
$(\mu^+ \mu^-)$	$m < 7100 \text{ MeV}/c^2$ $\chi_{\text{FD}}^2 > 9$
μ^+/μ^-	$\text{DLL}_{\mu\pi} > -3$ IsMuon true $\chi_{\text{IP}}^2 > 9$ $\text{track } \chi^2/\text{ndf} < 5$

to the X resonance ($\chi_{\text{FD}}^2 < 9$). Additional cuts are implied onto the individual muons. They are required to be well reconstructed with a track fit χ^2/ndf of smaller than 5, and pass particle identification requirements by having both the `IsMuon` flag (momentum-dependent number of hits in the muon system) as well as a difference of logarithmic likelihood (DLL) between the muon and pion hypotheses of larger than -3. To ensure that the tracks do not originate from the primary interaction vertex, an IP χ^2 value of larger than 9 is required.

11.5 Pre-selection and Multivariate Analysis

The data sample achieved after the Stripping selection is, as mentioned before, still very loose in terms of requirements on the $(K^+ K^-)$ resonance. As the analysis presented in this thesis aims to investigate specifically the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$, an additional cut-based (pre-)selection step is applied before the application of a multivariate BDT. Furthermore, this pre-selection includes dedicated vetoes against mis-reconstructed background sources mimicking the investigated decay mode, as the BDT is trained to reduce only combinatorial background levels.

In the pre-selection the cut on the reconstructed invariant mass of the B_s^0 -meson candidate is tightened to $5100 \text{ MeV}/c^2 < m_{\text{rec}} < 5800 \text{ MeV}/c^2$ and the reconstructed invariant mass of the ϕ -meson candidate to $1007.455 \text{ MeV}/c^2 < m_\phi < 1031.455 \text{ MeV}/c^2$, corresponding to a $12 \text{ MeV}/c^2$ wide window around the nominal ϕ -meson mass [15]. Furthermore, the particle identification of the kaon candidates is required to fulfil

$DLL_{K\pi} > -3$. Despite the good performance of the LHCb particle identification system, there is typically a percent-level chance of mis-identifying final-state particles. When investigating rare decays such as $B_s^0 \rightarrow \phi \mu^+ \mu^-$, pollution from mis-identified candidates from much more abundant decay modes can become relevant. The mass of final-state particles is assigned depending on the particle identification by the LHCb reconstruction algorithms. For that reason, alternative mass hypotheses are assigned to them depending on the investigated background source. Subsequently, the invariant mass of composite particles is recalculated under these different particle hypotheses and the pollution from the corresponding decay mode determined. Two significant sources of mis-identification backgrounds are found: b -hadron decays involving $J/\psi \rightarrow \mu^+ \mu^-$ transitions, where one of the muons is mis-identified as a kaon, and $\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$ decays, where the proton is mis-identified as a kaon. Figure 11.3 shows the invariant mass spectra of the μK (a) and $\mu\mu KK$ (b) systems after assigning the muon and the proton mass to one of the kaon candidates, respectively. The respective mass hypothesis is applied to the kaon candidate if it fulfils the `IsMuon` flag requirements and a $DLL_{\mu\pi}$ value of larger than 5 for the $J/\psi \rightarrow \mu^+ \mu^-$ background estimation, and $(DLL_{p\pi} - DLL_{K\pi})$ of larger than 10 for the Λ_b^0 background investigation, respectively. In order to remove these background candidates from the selection, they are vetoed in case their invariant mass under the alternative hypothesis is close to the respective invariant resonance mass: a $45 \text{ MeV}/c^2$ wide window is chosen around the nominal J/ψ mass, a $50 \text{ MeV}/c^2$ wide window around the nominal Λ_b^0 mass. The width of those windows is chosen to only slightly affect the number of signal events while significantly reducing the respective background contribution.

The distribution of the $\mu^+ \mu^-$ versus the $K^+ K^- \mu^+ \mu^-$ invariant mass of the selected signal candidates is shown in Fig. 11.4. Two horizontal high-statistics bands are visible around the dimuon masses of about $3100 \text{ MeV}/c^2$ and $3700 \text{ MeV}/c^2$, corresponding to the J/ψ and $\psi(2S)$ resonances, respectively. The signal decay corresponds to events with an invariant $K^+ K^- \mu^+ \mu^-$ mass in close proximity of the nominal B_s^0 -meson mass, the signal region of $\pm 50 \text{ MeV}/c^2$ around the latter is indicated by vertical dashed lines. For the analysis of the signal decay, candidates with a dimuon mass close to the charmonium resonances are rejected from the selection by the chosen binning scheme as described in Chap. 10, indicated by vertical red lines.

After the reduction of the significant sources of peaking backgrounds, a multivariate BDT is trained to further reduce combinatorial background levels. Dedicated training samples representing the signal and background topologies are needed in the training. In order to minimise potential mis-modelling in simulated event samples that would influence the performance of the BDT classification, data is used for the training. The dominant $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decay is used as a proxy for the signal decay, as it has the same topology besides being resonant in the dimuon system. The candidates used for this training are indicated by the blue shaded area. The upper sideband in the invariant $K^+ K^- \mu^+ \mu^-$ mass of more than $200 \text{ MeV}/c^2$

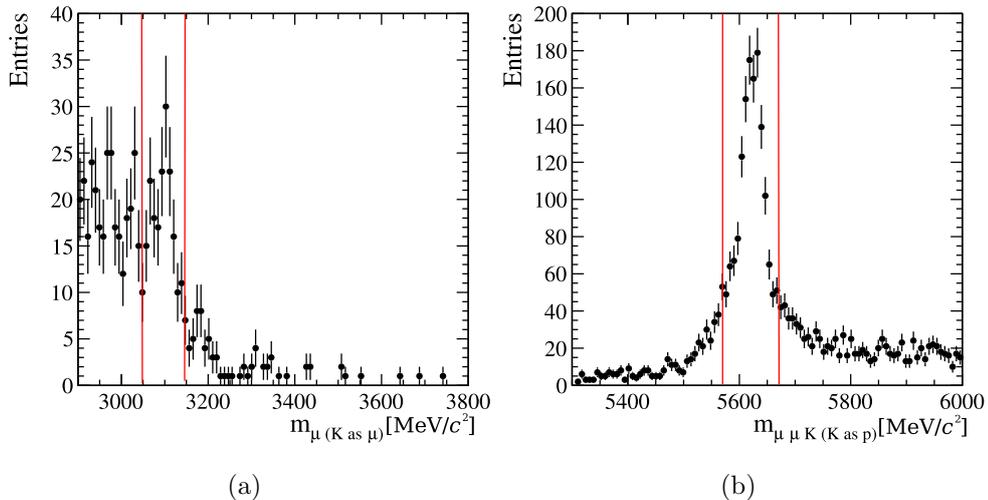


Figure 11.3: (a) Invariant μK mass spectrum after assigning the muon mass to the kaon. (b) Invariant $\mu^+\mu^-K^+K^-$ mass spectrum after assigning the proton mass to one of the kaons. Vertical red lines indicate the invariant mass range around the background resonances for which candidates are removed from the selection.

above the nominal B_s^0 -meson mass is used in the training as the background sample, indicated by the red shaded area. To avoid a biasing selection, it is not advisable to train a BDT on the sample, which is later analysed with it. This might lead to an overestimation of the separation power of the BDT, leading to an incorrect determination of the selection efficiency. For that reason, the dataset used for the training is split randomly into two halves: the half-sized signal and background samples are used individually to train two BDT classifiers. Subsequently, each of these two classifiers is applied to the other half of the sample, to investigate potential over-training effects. For the final classification of the final dataset, the two decision trees are used individually for events not used in the training of the respective BDT.

The Boosted Decision Tree is trained using the Toolkit for Multivariate Data Analysis [50]. Table 11.3 shows the variables used for the BDT training. All of these observables were used in the earlier selection steps due to a strong discriminating power against various background sources. The distributions of signal and background samples in the BDT variables are shown in Fig. 11.5. The signal distributions for the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ sample are derived from the region described above by subtracting the background using the sPlot technique [53]. Each variable shows differences between signal and background distributions, but exploiting correlations between the full set of variables allows for an even better separation. The resulting BDT response is shown in Fig. 11.6, where the two BDTs trained on one half of the

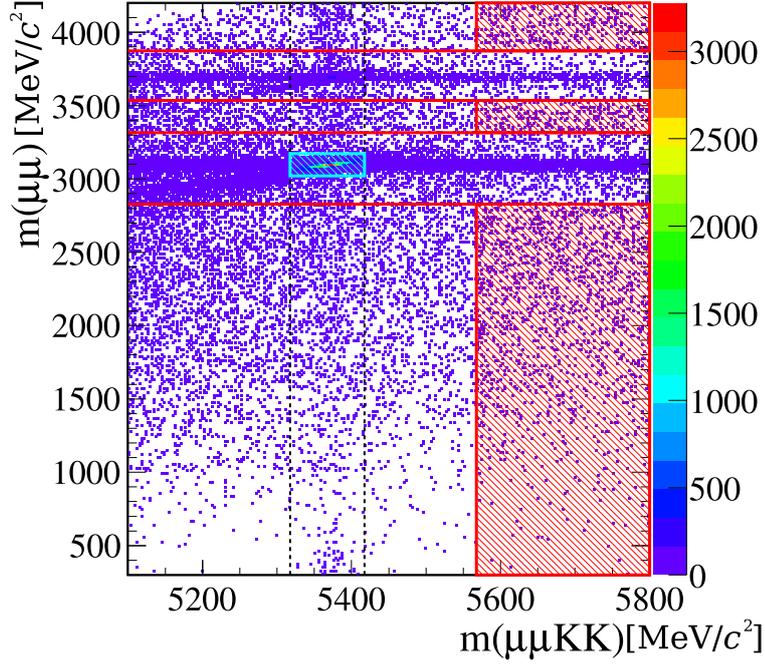


Figure 11.4: $\mu^+\mu^-$ versus the $K^+K^-\mu^+\mu^-$ invariant mass distribution after application of the pre-selection cuts. The $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+\mu^-)\phi$ candidates in the blue shaded region are used as signal in the BDT training, the red shaded areas as background.

Table 11.3: Variables used in the training of the Boosted Decision Trees.

Particle	Variable
B_s^0	χ_{vertex}^2
	χ_{IP}^2
	$\cos(DIRA)$
	χ_{FD}^2
	p_{T}
K^+/K^-	$\min(\text{DLL}_{K\pi}(K^+), \text{DLL}_{K\pi}(K^-))$
	$\max(\text{DLL}_{K\pi}(K^+), \text{DLL}_{K\pi}(K^-))$
μ^+/μ^-	χ_{IP}^2
	$\min(\text{DLL}_{\mu\pi}(\mu^+), \text{DLL}_{\mu\pi}(\mu^-))$

training sample were applied to the respective other half and the responses are compared to each other. No significant disagreement between the responses to training and test sample can be observed in any of the BDTs, indicating no overtraining effects. An important tool to evaluate the quality of a Boosted Decision Tree is the Receiver Operating Characteristic (ROC) curve, showing the background rejection over the signal efficiency. It is created by scanning through different cut values on the BDT response and evaluating both at this cut value. The ROC curve for both trained BDTs are shown in Fig. 11.7. The final working points of BDT classification values are determined by optimising the signal significance, defined as the ratio of signal events over the square root of all events, $S/\sqrt{S+B}$, where the number of signal events is scaled down from the number of $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ events using the previously measured relative branching fraction of $\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)/\mathcal{B}(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi) = 0.00113$ [54]. The final cut values chosen for the response of the two BDTs are 0.021 and 0.016, respectively. They are indicated in Fig. 11.7 by dots.

11.6 Fully selected data sample

With the full selection applied, the final data sample used in the analysis presented here contains exactly 1000 candidates for the $B_s^0 \rightarrow \phi\mu^+\mu^-$ channel, while the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ channel sample contains 70139 candidates. This includes both signal and background events. The yield of signal events for the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ channel are extracted by an extended maximum likelihood fit (see Sec. 14.1) to the invariant $K^+K^-\mu^+\mu^-$ mass distributions. Within the signal region of $\pm 50 \text{ MeV}/c^2$ around the nominal B_s^0 -meson mass, the observed yields are 432 ± 24 and 62033 ± 260 for the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ channel, respectively. Figure 11.8 shows the fits to the respective data samples. This corresponds to a signal significance for the $B_s^0 \rightarrow \phi\mu^+\mu^-$ channel of $S/\sqrt{S+B} = 18.5$ within the signal region.

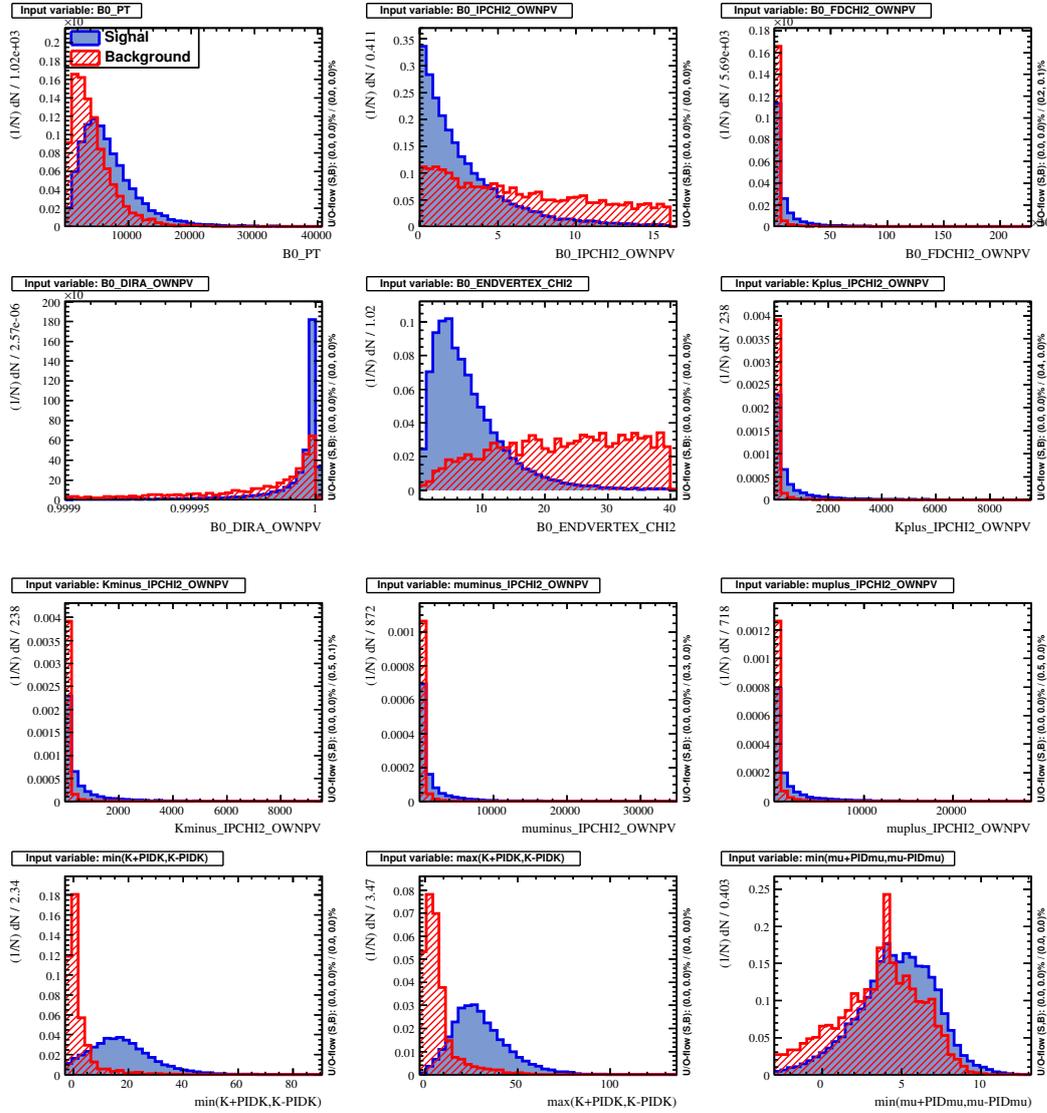


Figure 11.5: Distribution of the training samples for signal (blue) and background (red) events of the BDT variables.

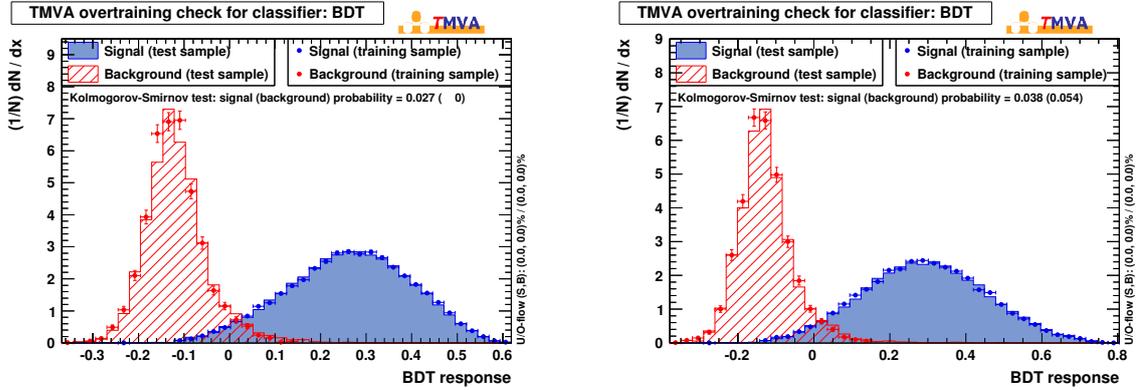


Figure 11.6: Response for signal and background events of the two trained BDTs.

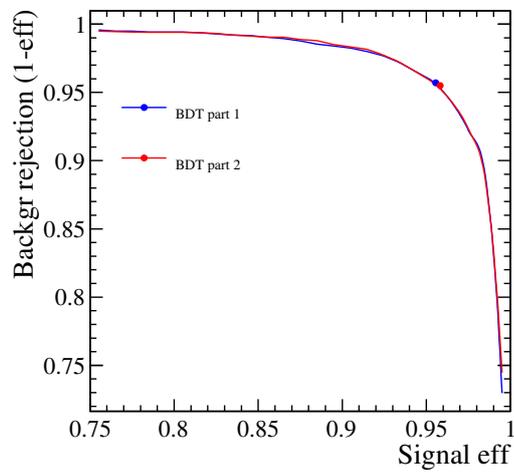


Figure 11.7: ROC curve for the two trained BDTs. The chosen cut values for optimised signal significance are indicated by dots.

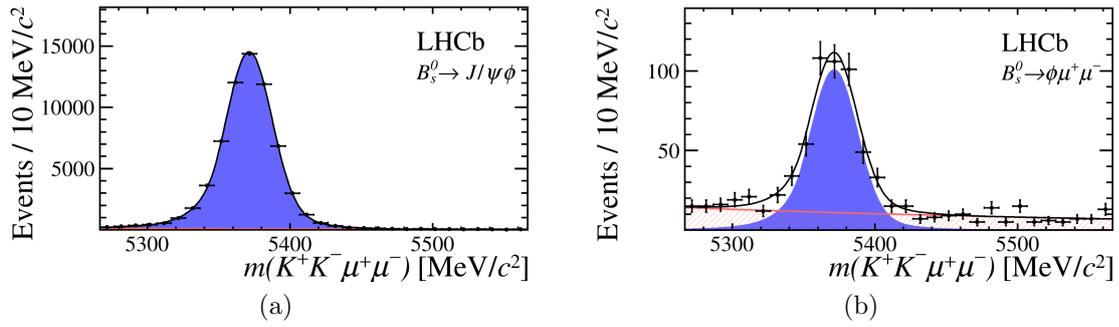


Figure 11.8: Fits to (a) the selected $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ candidates and (b) the selected $B_s^0 \rightarrow \phi\mu^+\mu^-$ candidates, the latter integrated over all six q^2 bins.

12 Corrections of the simulated event samples

In order to optimise the selection process and evaluate efficiencies in this selection process and the data collection in general, it is usually preferable to use data-driven methods in order to avoid any additional sources of uncertainties in contrast to the use of simulated event samples. However, this is not feasible in all cases, be it due to statistical limitations of potential calibration samples or the evaluation of non-observable effects. Therefore, a large part of the efficiency evaluation relies on samples of simulated events. These events are generated according to different theoretical physics models describing the dynamics of hadron production in proton-proton collisions at the LHC and the response of the LHCb detector, as described in Sec. 2.4. While the modelling of both the generation as well as the detector response agrees in general very well with what is observed, these simulations need (and allow for) a lot of tuning, and some observables remain not perfectly modelled. As the analysis presented in this thesis relies on simulated events to evaluate the efficiencies of the detection and selection process, data-driven techniques are used to further improve the agreement between data and the simulated event samples. Corrections are applied on a per-event basis, where a weighting factor is calculated from each individual correction.

12.1 Simulated event samples

In the analysis presented here a multitude of simulated event samples are used. This includes samples for the $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ decays as well as for important background sources, the latter in order to evaluate the magnitude of the respective contributions in the selected dataset. Separate samples are created for the 2011 and 2012 data-taking periods, as data-taking conditions varied between these two years, especially the configuration of the trigger and the centre-of-mass energy of the proton-proton collisions. For the $B_s^0 \rightarrow \phi\mu^+\mu^-$ channel there are two models used for the B_s^0 -meson decays: a model incorporating the full decay kinematics as described by P. Ball and R. Zwicky [55] and a model using basic phase space kinematics only. The model labelled as `EvtPVWCPLH` is included in the `EVTGEN` package [12]. It models decays of a scalar b -hadron to two vector particles, including the description of CP violation and the decay-width difference of the light and heavy mass eigenstates. It is used for the simulation of the tree-level decay $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$, as well as the similar decay $B^0 \rightarrow J/\psi K^{*0}$.

The sample sizes are chosen as a compromise between resources spent (in terms of computation time) and statistical limitations these sizes induce in terms of systematic uncertainties on the final results. Each dataset contains significantly more decays of the respective mode than what is expected to be contained in the data sample. The samples used in this analysis, the chosen decay model, and their respective simulated event numbers are listed in Table 12.1. Samples labelled with the addition “generator level” indicate that the simulation does not incorporate the detector response and

Table 12.1: Overview of the simulated event samples used.

Decay mode	Data taking conditions	Decay model	Events
$B_s^0 \rightarrow \phi \mu^+ \mu^-$	2011, $\sqrt{s} = 7$ TeV	Ball/Zwicky	4.4M
	2012, $\sqrt{s} = 8$ TeV	Ball/Zwicky	550k
	2012, $\sqrt{s} = 8$ TeV	Phase space	2.1M
	2011, $\sqrt{s} = 7$ TeV, generator level	Ball/Zwicky	5M
	2012, $\sqrt{s} = 8$ TeV, generator level	Ball/Zwicky	500k
	2012, $\sqrt{s} = 8$ TeV, generator level	Phase space	20M
$B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$	2011, $\sqrt{s} = 7$ TeV	EvtPVVCPH	37M
	2012, $\sqrt{s} = 8$ TeV	EvtPVVCPH	5.1M
	2012, $\sqrt{s} = 8$ TeV, generator level	EvtPVVCPH	5M
$B^0 \rightarrow J/\psi K^{*0}$	2012, $\sqrt{s} = 8$ TeV	EvtPVVCPH	4.5M
$B^0 \rightarrow K^{*0} \mu^+ \mu^-$	2012, $\sqrt{s} = 8$ TeV	Ball/Zwicky	520k
$B_s^0 \rightarrow D_s^+ \pi^-$	2012, $\sqrt{s} = 8$ TeV	Phase space	5M
$B_s^0 \rightarrow D_s^+ (\rightarrow \phi \mu^- \bar{\nu}_\mu) \mu^+ \nu_\mu$	2012, $\sqrt{s} = 8$ TeV, generator level	Phase space	2M
$B^0 \rightarrow D^- (\rightarrow K^{*0} \mu^- \bar{\nu}_\mu) \mu^+ \nu_\mu$	2012, $\sqrt{s} = 8$ TeV, generator level	Phase space	2M
$\Lambda_b^0 \rightarrow p K^- \mu^+ \mu^-$	2012, $\sqrt{s} = 8$ TeV	Phase space	710k
$\Lambda_b^0 \rightarrow \Lambda^0(1520) \mu^+ \mu^-$	2012, $\sqrt{s} = 8$ TeV	Phase space	511k

reconstruction steps, allowing for the investigation of observable distributions without detector effects. The simulation samples for background channels are created only for the 2012 data-taking conditions, as their expected contributions are evaluated relative to the signal decay. The variation of these relative contributions in the LHCb run I dataset¹ is expected to be negligible compared to the precision of this analysis.

12.2 Track reconstruction efficiency corrections

The track reconstruction efficiency of the LHCb detector is dependent on multiple factors. The angle and momentum with which a particle passes the tracking detectors has a direct effect in how many hits it can leave in them, and the occupancy in the detector affects the ability to separate the hits of different particles from each other and properly reconstruct the particle trajectories. Furthermore, the quality of the reconstruction is affected by the interactions of the respective particle with the detector material, which may lead to multiple scattering and energy losses. While all of these effects are crucial parts of the detector simulation, the efficiencies might differ between data and simulation, due to the complexity of the system. In order to evaluate these differences, a data-driven approach is chosen to determine the reconstruction efficiencies for both data and simulation, and map them depending

¹ p - p collisions with a centre-of-mass energy of $\sqrt{s} = 7$ TeV in 2011 and $\sqrt{s} = 8$ TeV in 2012.

on track kinematics and detector occupancy. The procedure is described in detail in Sec. 4.

Figure 12.1(a) shows the ratio of track reconstruction efficiency $R = \varepsilon_{data}/\varepsilon_{sim}$ of data compared to simulation as function of the pseudorapidity η and transverse momentum p_T , as determined for 2012 data taking conditions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. This ratio is used as a per-track correction factor for the simulated events. The correction factor per signal candidate $w_{e,track,eff.}$ is calculated by multiplying the correction factors of all final-state tracks

$$w_{e,track,eff.} = \prod_{tracks} R_{track,eff.}(\eta_{track}, p_{T,track}). \quad (12.1)$$

While the correction factors for the individual tracks are mostly close to unity, correlations between the track kinematics depending on the specific decay topology can lead to non-trivial corrections for the full signal candidate. Thus a careful evaluation of both the correction as well as the associated uncertainty has to be performed. The correction is applied to each simulated event sample, including $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decay and all background samples. The correction factors are used as event weights which appear as part of the calculation of the selection and reconstruction efficiencies or the evaluation of the background pollution, respectively.

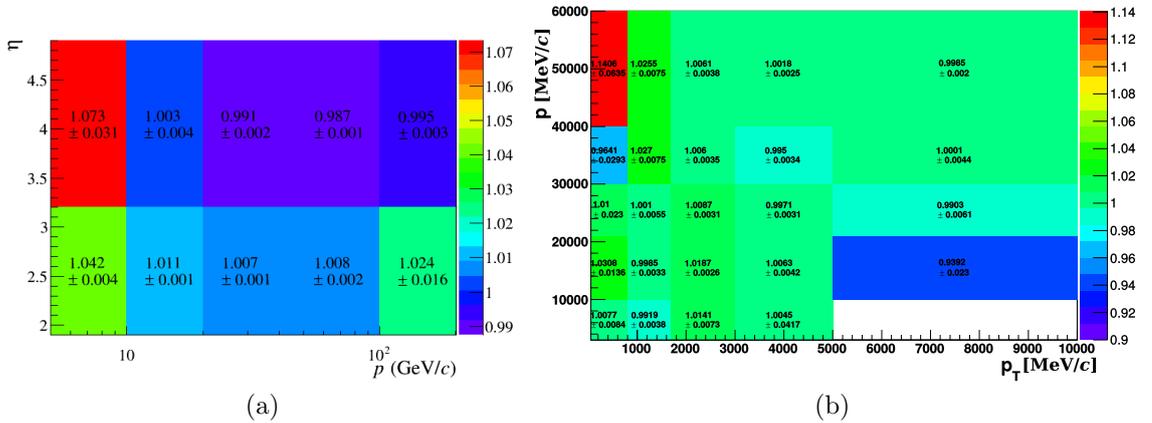


Figure 12.1: Ratio of efficiencies of data compared to simulated events for (a) track reconstruction and (b) the IsMuon criterion in the LHCb experiment for the 2012 data taking period.

12.3 Correction of IsMuon flag efficiency

The `IsMuon` flag is used as important part of the muon identification in multiple stages of the selection process. It is set if a track leaves a certain number of hits in the muon stations, depending on its momentum. Similar to the track reconstruction efficiency, the efficiency to set this flag successfully for a muon track may not fully be reproduced in simulated event samples. It is evaluated via the tag-and-probe technique, searching for $J/\psi \rightarrow \mu^+\mu^-$ candidates, where the presence of the `IsMuon` flag is checked for the final-state muon candidates. The ratio of efficiencies R_{IsMuon} between data and simulation is shown in Fig. 12.1(b) in bins of the transverse momentum p_T and the total momentum p . From this table a correction weight is calculated for each final-state muon track and all track corrections multiplied to calculate the correction factor $w_{e,\text{IsMuon}}$ per signal candidate

$$w_{e,\text{IsMuon}} = \prod_{\mu \text{ tracks}} R_{\text{IsMuon}}(p_{\text{track}}, p_{T,\text{track}}). \quad (12.2)$$

12.4 Recreation of particle identification variables

In general the simulation samples used in this analysis reflect the decay topologies and response of the LHCb detector to a high degree. However, the simulation of the particle identification is an exception to that. The particle identification algorithms use information from the RICH detectors and calorimeters to form a particle hypothesis. As the set of available observables used in these algorithms is quite large and the algorithms themselves are very complex, even small deviations might change the final particle identification variable dramatically. In addition to that, multiple experimental conditions such as temperature, gas pressure, and others, might vary from run to run and influence the detector response. To avoid the need to rely on the simulation of all these factors, the efficiencies to identify particles as the proper species and to mis-identify them as other species are measured in data. The LHCb collaboration has developed a tool to access the outcome of the data-driven evaluation of particle identification efficiencies, called PIDCalib [56]. It uses a set of decay modes with low background contributions, where the species of one or multiple of the final-state particles can be determined without the use of the particle identification variables, and compared to the detector response. The used decay modes are

- $D^{*+} \rightarrow D^0\pi^+$, where the D^0 flavour is determined by the charge of the pion and thus in its decay $D^0 \rightarrow K^-\pi^+$ the kaon and pion can be distinguished.
- $\Lambda \rightarrow p\pi^-$, where the proton particle identification is measured. The high production rate of Λ baryons yields precise results and its long decay time allows for an excellent background discrimination.

- $J/\psi \rightarrow \mu^+\mu^-$, where similar to the measurement of track reconstruction and IsMuon efficiencies one final-state muon is selected with tight criteria, while the second muon is used to probe the particle identification variables.
- $B^+ \rightarrow J/\psi(\rightarrow e^+e^-)K^+$, where in a similar manner as above the particle identification of electrons can be probed.

With these calibration samples the efficiency to identify a particle as the correct or a specific other species is determined, as a function of the track kinematics, the detector occupancy, and the used cut on the respective particle identification variable. These results are used to derive a distribution for the particle identification variables $DLL_{K\pi}$, $DLL_{\mu\pi}$, and $DLL_{p\pi}$ for the kaon candidates and $DLL_{\mu\pi}$ for the muon candidates in bins of the track kinematics and occupancy, as they occur on data. These distributions are used to resample the respective distributions for the used simulation samples. Depending on the track and detector properties, all required particle identification variables are randomly drawn according to the derived data distributions in the respective bin. As this is a statistical process, the resampled simulation sample is required to have a significant size, which is the case for all simulation samples used in the analysis presented here. The new distributions of the particle identification variables created that way are further used in the selection applied to the simulation samples, including the BDT classification, and thus yield by construction the correct selection efficiencies.

12.5 Additional corrections

With the correction and resampling steps described above, the agreement between data and simulated event samples is significantly improved. However, further investigation of the variables entering the multivariate BDT selection using the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ sample shows a remaining discrepancy in the distributions of the transverse momentum p_T and vertex χ^2 of the B_s^0 meson. Furthermore, the multiplicity of reconstructed tracks per event, which affects both track reconstruction and particle identification efficiencies, varies significantly between data and simulation. Thus the simulation samples are re-weighted to reflect the proper distributions for these three variables. “Re”-weighting refers to the normalisation of these weights over the respective simulated event sample, via:

$$\sum_{i=1}^{N_{evt}} w_{add,i}(p_T(B_s^0), \chi_{vtx}^2(B_s^0), N_{tracks}) = 1, \quad (12.3)$$

where w_i is the weight of the i -th simulated event and N_{evt} is the number of simulated events. This normalisation is applied before the selection requirements, in order to not to bias the calculated efficiencies. The weights are calculated from the difference between data and simulated events for each variable independently, as no correlations between them are observed.

12.6 Control of improved agreement

The $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ mode is used to check the quality of the applied correction steps, by comparing the cleanly selected data to the fully corrected simulation sample. All variables entering the multivariate BDT are compared, as is the response of this BDT in order to investigate variable correlations not properly reflected by the corrections. The resampled PID variables are directly used instead of the corresponding simulated values, while the corrections are applied as a per-event weight w_e , defined as:

$$w_e = w_{e,track. \text{ eff.}} \cdot w_{e, \text{ IsMuon}} \cdot w_{add,e}(p_T, \chi_{vx}^2, N_{tracks}). \quad (12.4)$$

Figures 12.2 to 12.8 show these comparisons. The four categories of corrections mentioned above are independent of each other and thus can be applied in any given order and are applied simultaneously during the analysis. However, for illustrative purpose they are applied on top of each other in the following order in the figures: the distributions for the uncorrected simulation sample are shown in green. After applying the track reconstruction and IsMuon efficiency corrections the distributions are given by the magenta line. Using in addition the resampled particle identification variables instead of the simulated values yields the blue line. Finally, after reweighting in the B_s^0 transverse momentum, vertex χ^2 and track multiplicity, the fully corrected distributions are shown by the red line. Below each figure the pull p between data and fully corrected simulation is shown. It is defined as

$$p = \frac{\mu_{\text{corrected sim}} - \mu_{\text{data}}}{\sigma_\mu}, \quad (12.5)$$

where $\mu_{\text{corrected sim}}$ and μ_{data} are the values of the respective variable for the fully corrected simulated event sample and the background subtracted dataset, and σ_μ is the quadratically added uncertainty of the these two values in each bin.

In general the correction procedure leads to a significantly improved agreement between data and simulation, allowing for a precise efficiency determination using the corrected simulated event samples. Small deviations remain in the momentum and transverse momentum and momentum of the final-state kaon candidates, which do not lead to a significant deviation for the BDT response. However, these remnant discrepancies are investigated as a source of systematic uncertainty in Chap. 16.

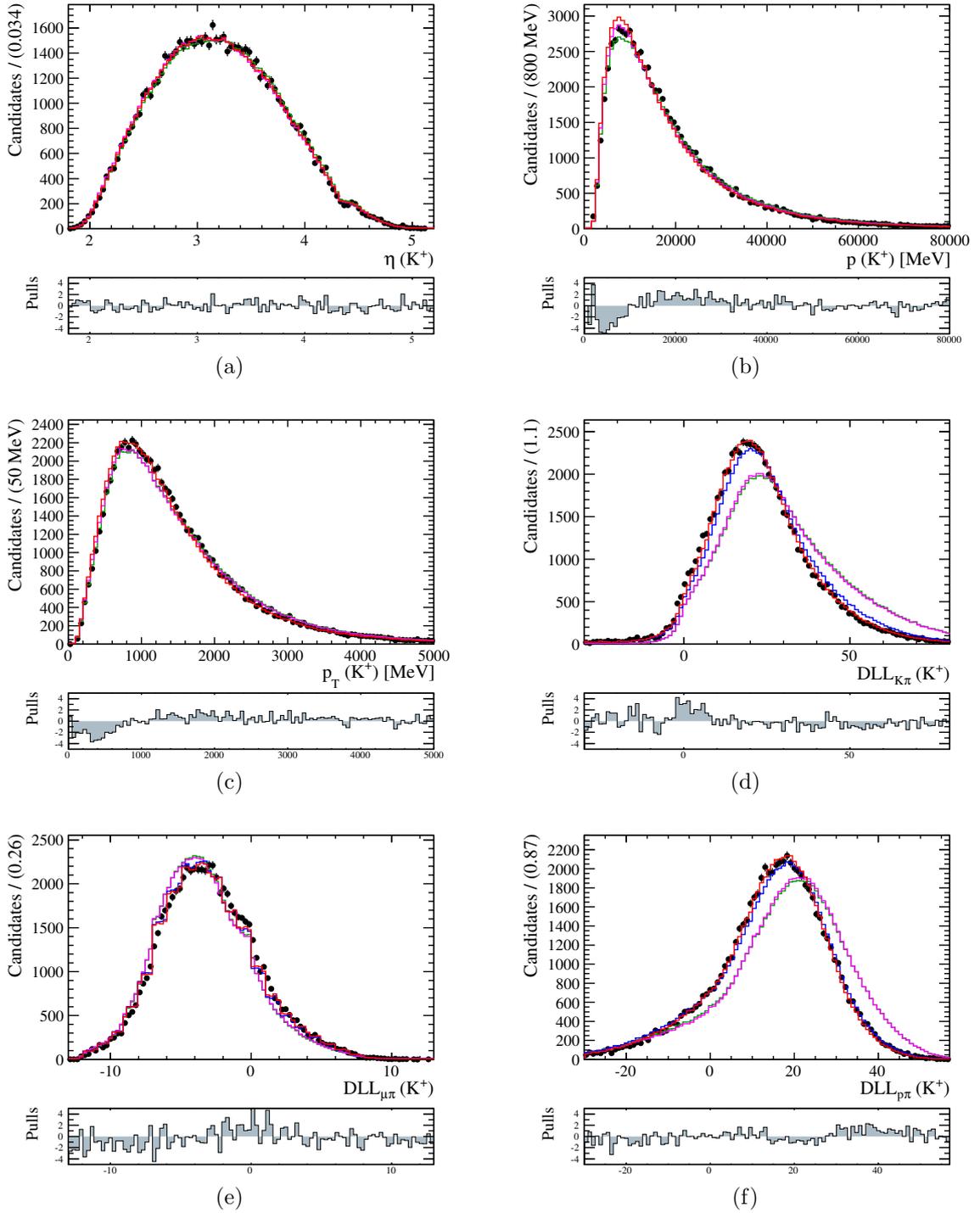


Figure 12.2: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (multiple colours, see full text for details) for the K^+ variables: η (a), momentum (b), transverse momentum (c), $DLL_{K\pi}$ (d), $DLL_{\mu\pi}$ (e), $DLL_{p\pi}$ (f).

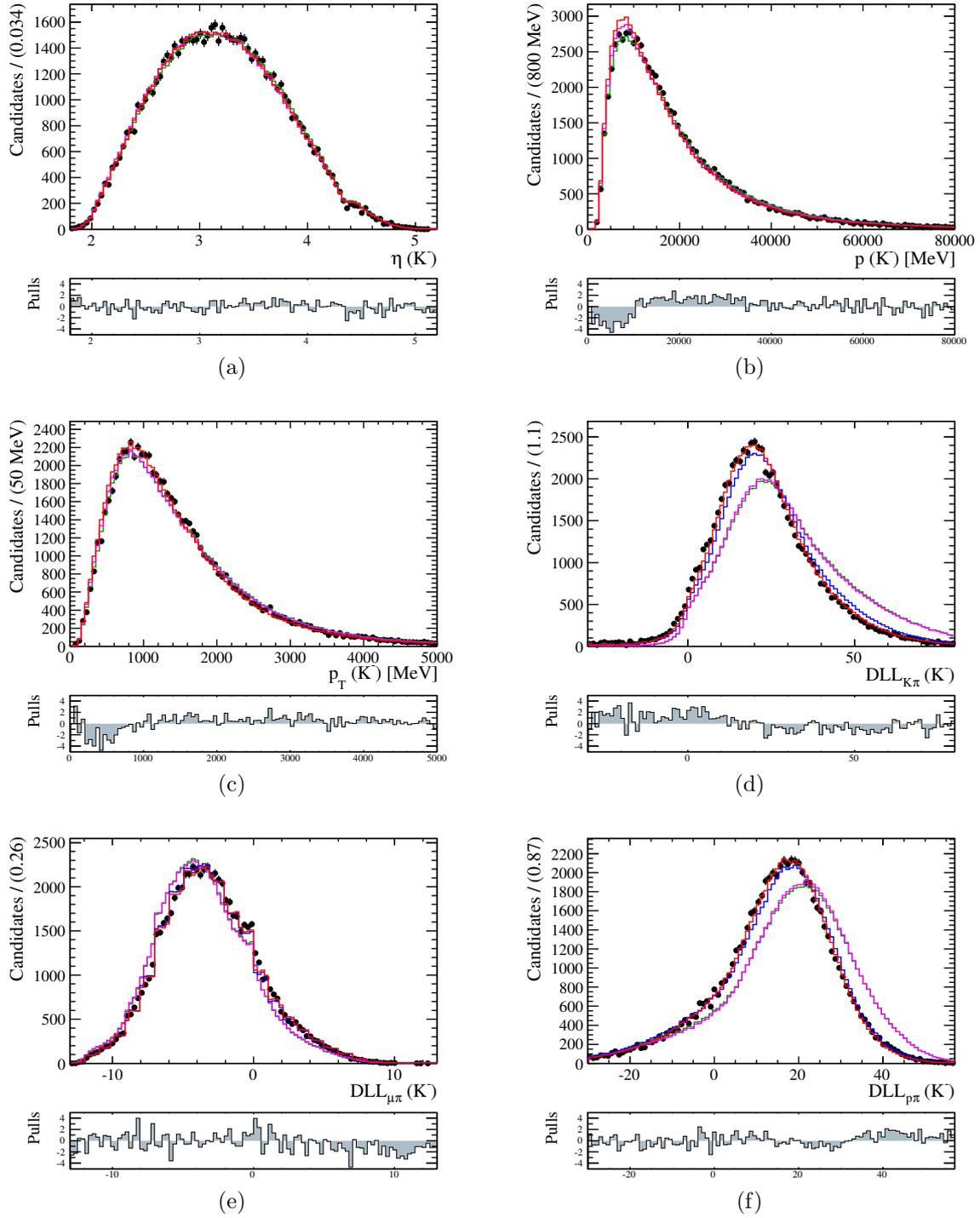


Figure 12.3: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (multiple colours, see full text for details) for the K^- variables: η (a), momentum (b), transverse momentum (c), $DLL_{K\pi}$ (d), $DLL_{\mu\pi}$ (e), $DLL_{p\pi}$ (f).

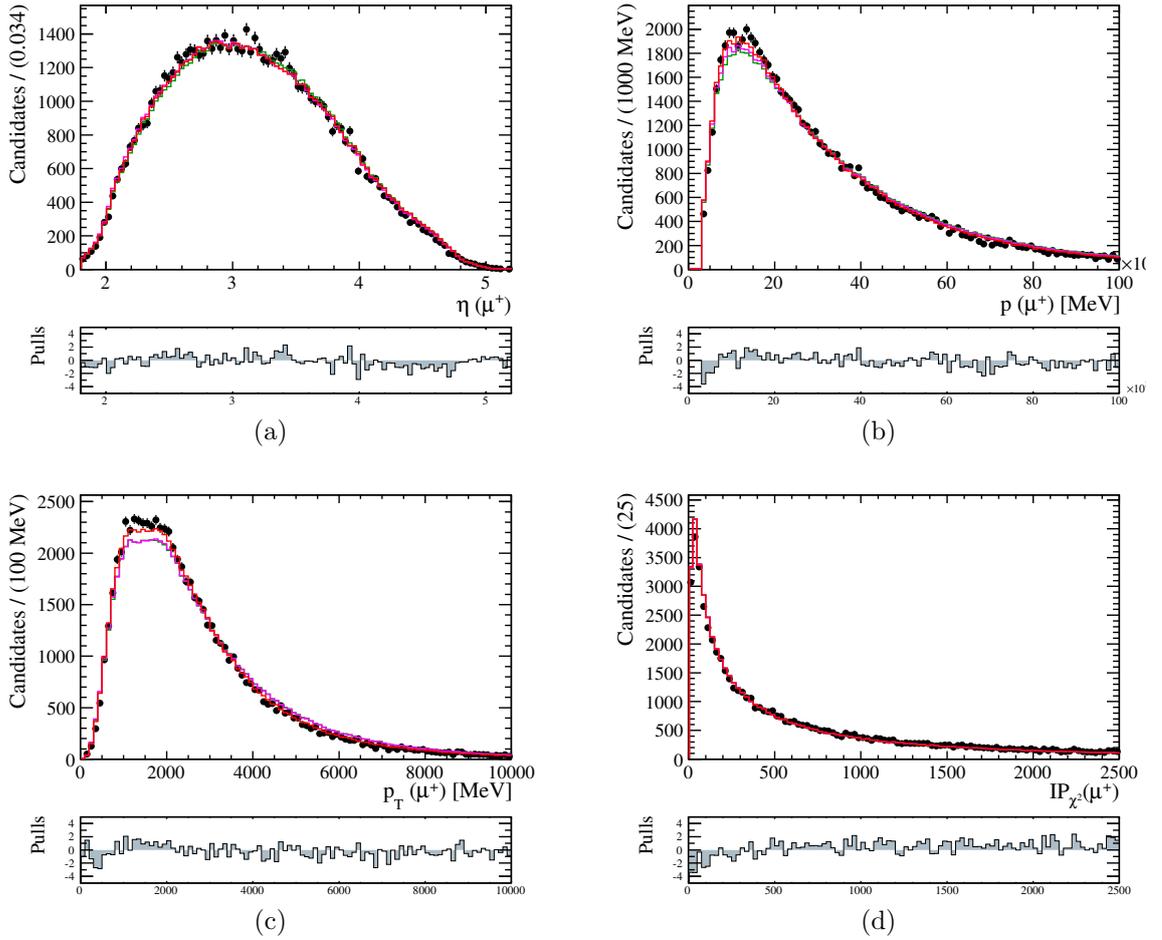


Figure 12.4: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (multiple colours, see full text for details) for the μ^+ variables: η (a), momentum (b), transverse momentum (c), χ^2_{IP} (d).

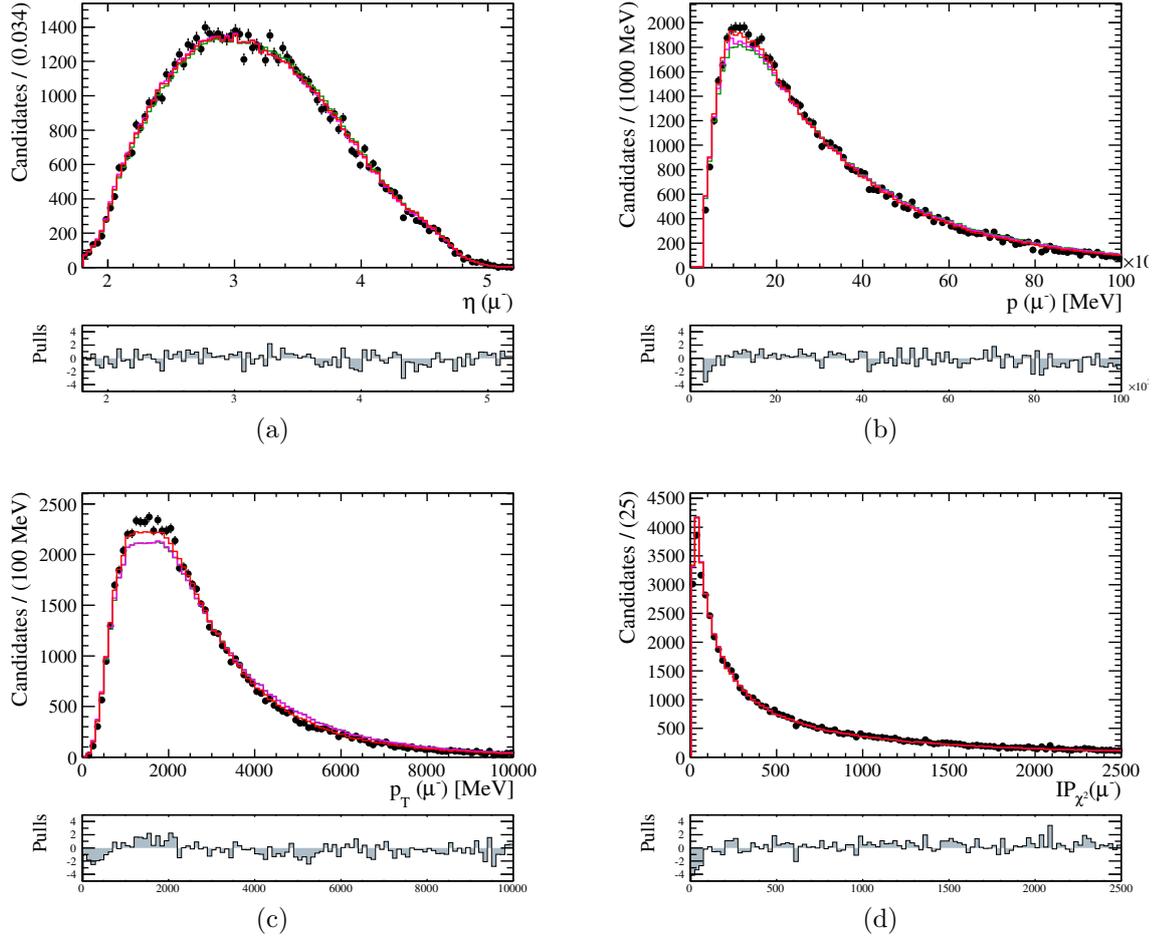


Figure 12.5: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (multiple colours, see full text for details) for the μ^- variables: η (a), momentum (b), transverse momentum (c), χ_{IP}^2 (d).

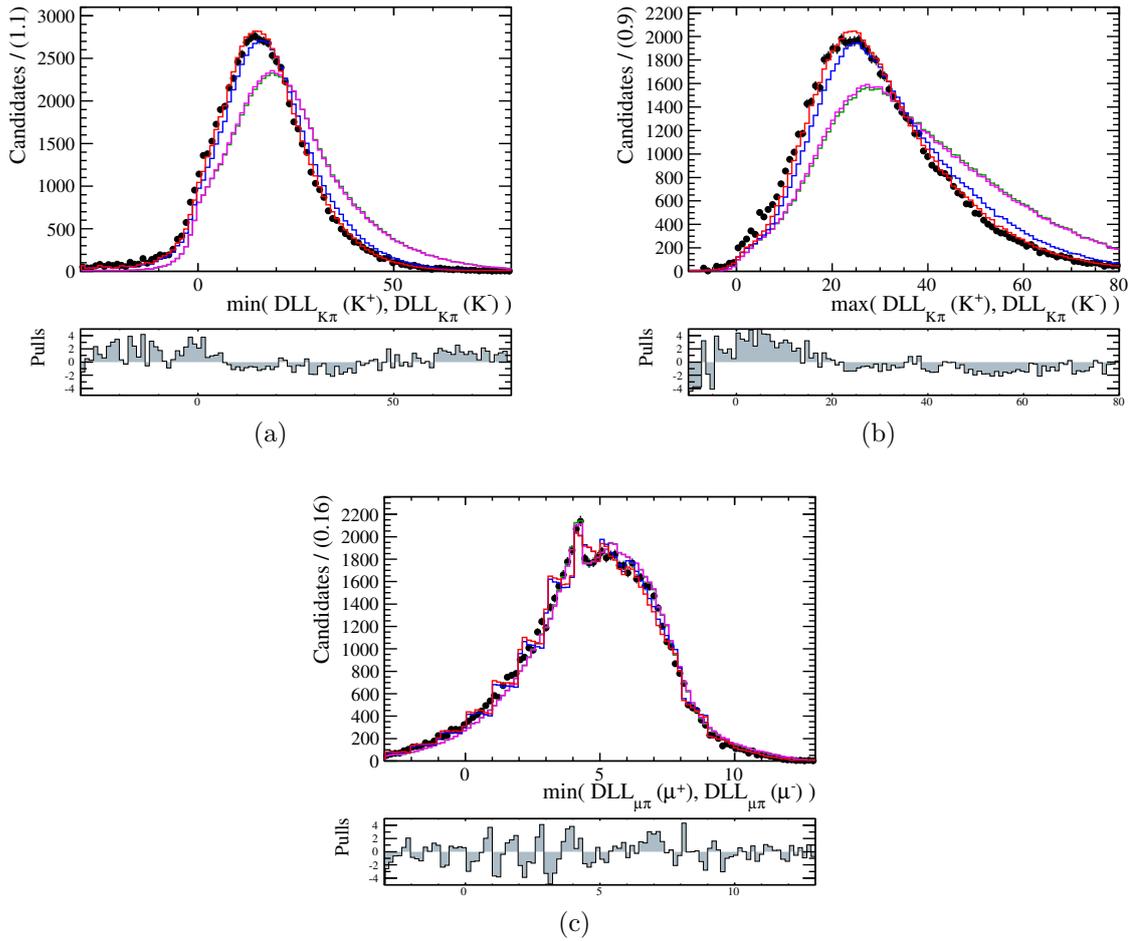


Figure 12.6: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (multiple colours, see full text for details) for the final-state particle PID variables: $\min(\text{DLL}_{K\pi}(K^+), \text{DLL}_{K\pi}(K^-))$ (a), $\max(\text{DLL}_{K\pi}(K^+), \text{DLL}_{K\pi}(K^-))$ (b), $\min(\text{DLL}_{\mu\pi}(\mu^+), \text{DLL}_{\mu\pi}(\mu^-))$ (c).

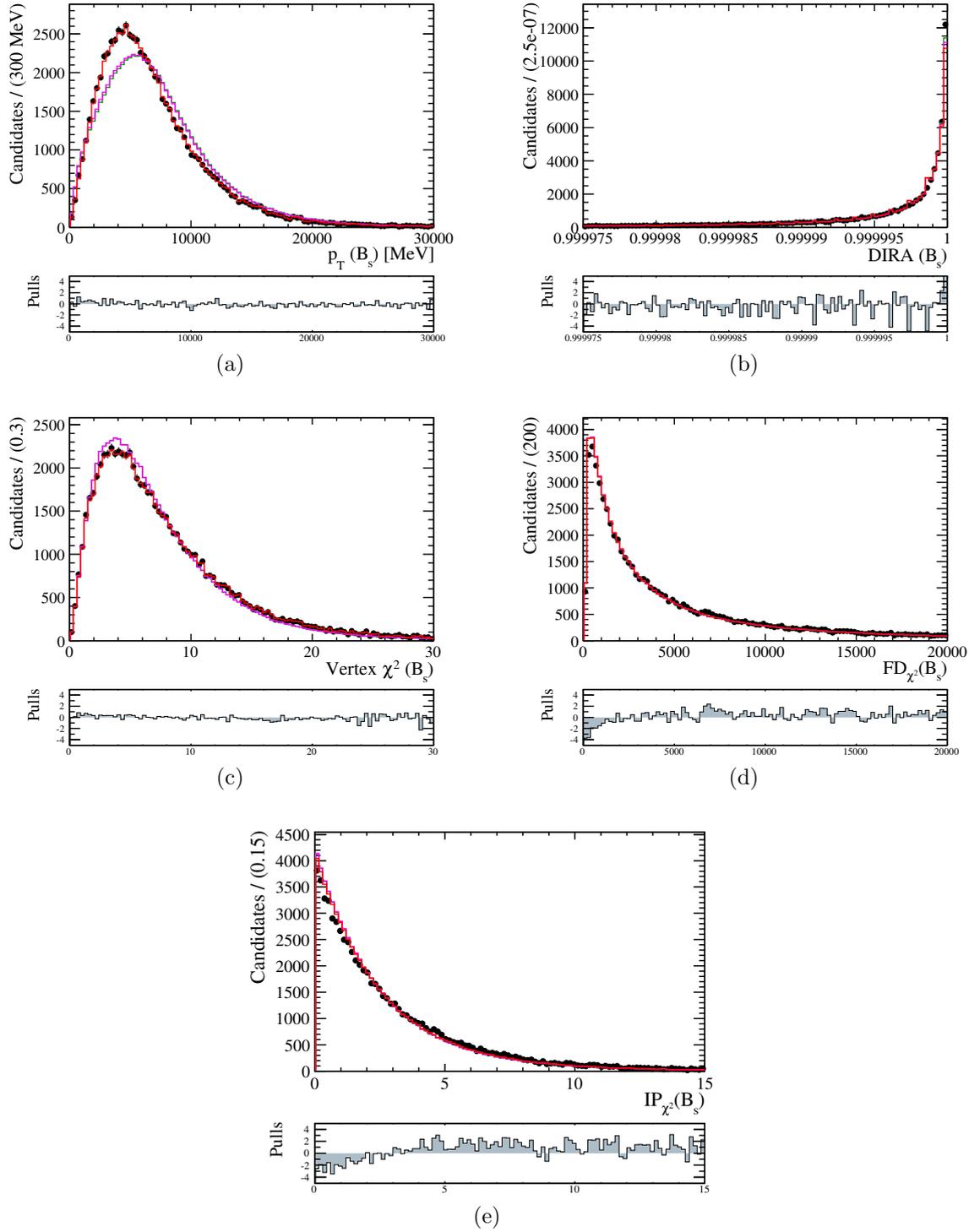


Figure 12.7: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (multiple colours, see full text for details) for the B_s^0 variables: transverse momentum (a), direction angle DIRA (b), vertex χ^2 (c), flight distance χ^2 (d), χ^2_{IP} (e).

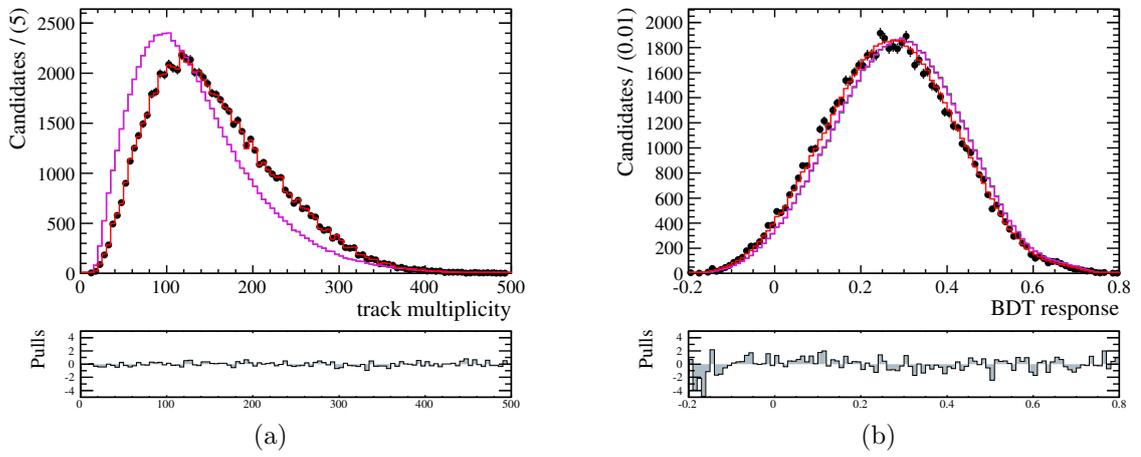


Figure 12.8: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (multiple colours, see full text for details) for the global variables: track multiplicity (a), BDT response (b).

13 Efficiency determination

Every measurement of physical properties is subject to inefficiencies, be it because of the design of the measurement device, features of the measured object, or at present day computational limitations. For the determination of the branching fraction of a decay mode, its efficiencies need to be carefully evaluated. The efficiencies to reconstruct and select the signal and control mode decays are determined from the simulation samples, where proper corrections are applied to reduce discrepancies due to imperfect detector descriptions (see Chap. 12). The total efficiency to find a desired decay can be described as the product of multiple uncorrelated efficiencies:

$$\varepsilon_{\text{tot}} = \varepsilon_{\text{det}} \cdot \varepsilon_{\text{rec|det}} \cdot \varepsilon_{\text{sel|rec}} \cdot \varepsilon_{\text{trig|sel}}, \quad (13.1)$$

where ε_{det} is the efficiency of the decay products appearing within the LHCb acceptance (limited by the detector design), $\varepsilon_{\text{rec|det}}$ is the efficiency of those decays to be reconstructed by the LHCb experiment, $\varepsilon_{\text{sel|rec}}$ the efficiency for the reconstructed decays to pass the selection requirements, and $\varepsilon_{\text{trig|sel}}$ the efficiency for those decays to have fulfilled the trigger requirements imposed. The order of those conditional efficiencies is interchangeable, as they are uncorrelated to each other. The chosen order is in line with other analyses performed by the LHCb collaboration [22–24], which allows for easy comparison of the individual efficiencies and helps to detect potential error sources.

In this chapter the determination and magnitude of these efficiencies is described, which is performed for each q^2 bin of the $B_s^0 \rightarrow \phi \mu^+ \mu^-$ mode and the $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ mode and the 2011 and 2012 data taking periods individually. In addition to the efficiency determination also the derivation of the angular acceptance used for the analysis of the angular distributions of the decay is presented.

13.1 Detector acceptance

The LHCb detector performs measurements in the forward direction close to the beam pipe. It can reconstruct particle tracks with an opening angle of 10 – 300 mrad in the x-plane and 10 – 250 mrad in the y-plane to the beam direction, originating at the primary interaction point. This limited detector acceptance affects the ability to reconstruct multi-particle final states. The preferred production of b mesons in the forward direction leads to about 25% of them to be emitted towards the detector. However, in order to fully reconstruct the multi-particle final state, all decay products are required to have a trajectory towards the active detector material. The directions in which the final-state particles are emitted depends both on the specific decay mode and the decay kinematics. Thus, the efficiency has to be determined in each q^2 bin individually, as it might differ between them.

The detector acceptance is determined purely from simulated events, as no measure-

Table 13.1: Detector acceptance efficiency for $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2011 data taking period.

$q^2 [\text{GeV}^2/c^4]$	$\varepsilon_{\text{det}} [\%]$
0.1 – 2.0	17.53 ± 0.12
2.0 – 5.0	16.74 ± 0.13
5.0 – 8.0	16.89 ± 0.12
11.0 – 12.5	16.96 ± 0.15
15.0 – 17.0	17.12 ± 0.14
17.0 – 19.0	17.18 ± 0.19
1.0 – 6.0	16.83 ± 0.10
15.0 – 19.0	17.14 ± 0.11

$q^2 [\text{GeV}^2/c^4]$	$\varepsilon_{\text{det}} [\%]$
8.0 – 11.0	16.89 ± 0.04

ment outside the detector acceptance can be performed. It is calculated via

$$\varepsilon_{\text{det}}(q^2) = \frac{N_{\text{in acceptance}}(q^2)}{N_{\text{generated}}(q^2)}, \quad (13.2)$$

where $N_{\text{in acceptance}}$ is the number of signal events where all final-state products end up in the LHCb acceptance out of $N_{\text{generated}}$ generated events. The kinematic distribution of the decay products is dependent on the hadronisation and decay models used in the simulated event sample, and the effects of using different models are accounted for as a source of systematic uncertainties in Chap. 16.

Tables 13.1 and 13.2 show the determined efficiencies for signal and control mode in the 2011 and 2012 data taking periods, respectively, where the uncertainties are of statistical nature only. The comparably larger statistical uncertainties for the signal mode efficiencies in the 2011 data taking period are caused by the lower amount of simulated events available compared to the other samples. In general, the detector acceptance for signal decays is larger for the 2012 data-taking period compared to 2011. This is likely caused by the higher centre-of-mass energy in the primary interaction which increases the average forward boost of the original B_s^0 meson and the particles originating from its decay.

13.2 Reconstruction and selection efficiency

The decays for which all final-state particles end up within the acceptance of the LHCb experiment have a finite efficiency to be properly detected by the detector

Table 13.2: Detector acceptance efficiency for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2012 data taking period.

$q^2[\text{GeV}^2/c^4]$	$\varepsilon_{\text{det}} [\%]$
0.1 – 2.0	18.00 ± 0.04
2.0 – 5.0	17.08 ± 0.04
5.0 – 8.0	16.96 ± 0.04
11.0 – 12.5	17.21 ± 0.05
15.0 – 17.0	17.40 ± 0.04
17.0 – 19.0	17.38 ± 0.06
1.0 – 6.0	17.14 ± 0.03
15.0 – 19.0	17.39 ± 0.04

$q^2[\text{GeV}^2/c^4]$	$\varepsilon_{\text{det}} [\%]$
8.0 – 11.0	17.10 ± 0.04

hardware and subsequently be correctly reconstructed by the dedicated reconstruction algorithms. In addition, the selection requirements applied in order to reduce the various background sources also remove a certain fraction of signal events. As the reconstruction algorithms also apply certain kinematic selection cuts in order to assure good quality of the reconstructed tracks, these two different steps are hard to fully disentangle and thus are treated as a single efficiency value. The reconstruction and selection efficiencies are determined from simulated event samples, where corrections are applied to account for the differences in track reconstruction and particle identification efficiencies between data and simulated events, as described in Chap. 12.

Different selection steps are applied to the data sample sequentially, which allows to decompose the reconstruction and selection efficiency via

$$\varepsilon_{\text{rec|det}} \cdot \varepsilon_{\text{sel|rec}} = \varepsilon_{\text{rec\&strip|det}} \cdot \varepsilon_{\text{preselection|rec\&strip}} \cdot \varepsilon_{\text{BDT|preselection}}, \quad (13.3)$$

where $\varepsilon_{\text{rec\&strip|det}}$ is the efficiency for a signal decay within the LHCb acceptance to be reconstructed and pass the Stripping requirements, $\varepsilon_{\text{preselection|rec\&strip}}$ for those decays to additionally pass the pre-selection, and $\varepsilon_{\text{BDT|preselection}}$ the efficiency for these signal events to be selected by the multivariate BDT. They are defined as

$$\varepsilon_{\text{rec\&strip|det}}(q^2) = \frac{N_{\text{rec.\&Strip.}}(q^2)}{N_{\text{in acceptance}}(q^2)}, \quad (13.4)$$

$$\varepsilon_{\text{preselection|rec\&strip}}(q^2) = \frac{N_{\text{preselection}}(q^2)}{N_{\text{rec.\&Strip.}}(q^2)}, \quad (13.5)$$

Table 13.3: Selection efficiencies for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2011 data taking period.

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{rec\&strip det}} [\%]$	$\epsilon_{\text{preselection rec\&strip}} [\%]$	$\epsilon_{\text{BDT preselection}} [\%]$	$\epsilon_{\text{rec det}} [\%]$
0.1 – 2.0	12.16 ± 0.05	80.79 ± 0.44	95.34 ± 0.43	9.37 ± 0.08
2.0 – 5.0	12.10 ± 0.05	80.79 ± 0.47	95.82 ± 0.46	9.37 ± 0.08
5.0 – 8.0	12.49 ± 0.05	80.47 ± 0.43	96.27 ± 0.42	9.67 ± 0.08
11.0 – 12.5	12.57 ± 0.06	80.19 ± 0.54	96.64 ± 0.52	9.74 ± 0.10
15.0 – 17.0	10.61 ± 0.05	82.29 ± 0.56	96.66 ± 0.54	8.44 ± 0.08
17.0 – 19.0	7.60 ± 0.06	82.92 ± 0.91	96.36 ± 0.89	6.07 ± 0.10
1.0 – 6.0	12.14 ± 0.04	80.72 ± 0.36	95.81 ± 0.35	9.39 ± 0.06
15.0 – 19.0	9.57 ± 0.04	82.46 ± 0.48	96.58 ± 0.46	7.62 ± 0.07

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{rec\&strip det}} [\%]$	$\epsilon_{\text{preselection rec\&strip}} [\%]$	$\epsilon_{\text{BDT preselection}} [\%]$	$\epsilon_{\text{rec det}} [\%]$
8.0 – 11.0	12.44 ± 0.01	80.16 ± 0.06	96.71 ± 0.05	9.64 ± 0.01

and

$$\epsilon_{\text{BDT|preselection}}(q^2) = \frac{N_{\text{BDT}}(q^2)}{N_{\text{preselection}}(q^2)}, \quad (13.6)$$

bin where $N_{\text{rec.\&Strip.}}$ is the number of events after reconstruction that pass the Stripping selection, $N_{\text{preselection}}$ the number of events passing the preselection step, and N_{BDT} the number of events passing the cut on the response of the multivariate classifier. Investigating these three sub-efficiencies individually for each q^2 allows to easily identify potential error sources in any of those steps.

The results of the reconstruction efficiency determination are shown in Tabs. 13.3 and 13.4, which list the individual components and total reconstruction and selection efficiency for $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ decays, separated by q^2 bins (for signal) and data taking periods. Uncertainties given are purely statistical. The efficiencies are fairly constant for low q^2 , however, the reconstruction and Stripping efficiencies turn out to be lower for high q^2 values. This effect is created by the relatively large opening angle between the two muons at high q^2 , causing at least one of them to have a steep angle to the beam line and thus leave fewer hits in the tracking system. The reconstruction and Stripping efficiency is also lower in 2012 than in 2011, which is caused by the higher detector occupancy during the 2012 data taking period, which has a detrimental effect on the track reconstruction efficiency.

Table 13.4: Selection efficiencies for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2012 data taking period.

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{rec\&strip det}} [\%]$	$\epsilon_{\text{presel rec\&strip}} [\%]$	$\epsilon_{\text{BDT presel}} [\%]$	$\epsilon_{\text{rec det}} [\%]$
0.1 – 2.0	10.66 ± 0.06	81.06 ± 0.60	94.37 ± 0.58	8.15 ± 0.09
2.0 – 5.0	10.57 ± 0.06	80.62 ± 0.64	95.17 ± 0.62	8.11 ± 0.10
5.0 – 8.0	10.90 ± 0.06	79.88 ± 0.58	95.81 ± 0.56	8.34 ± 0.09
11.0 – 12.5	11.08 ± 0.08	80.18 ± 0.73	96.20 ± 0.71	8.55 ± 0.12
15.0 – 17.0	9.51 ± 0.07	82.22 ± 0.75	96.07 ± 0.73	7.51 ± 0.11
17.0 – 19.0	6.86 ± 0.07	83.75 ± 1.23	96.13 ± 1.19	5.52 ± 0.12
1.0 – 6.0	10.55 ± 0.05	80.56 ± 0.49	95.14 ± 0.47	8.09 ± 0.07
15.0 – 19.0	8.59 ± 0.05	82.64 ± 0.64	96.09 ± 0.62	6.82 ± 0.08

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{rec\&strip det}} [\%]$	$\epsilon_{\text{presel rec\&strip}} [\%]$	$\epsilon_{\text{BDT presel}} [\%]$	$\epsilon_{\text{rec det}} [\%]$
8.0 – 11.0	11.17 ± 0.02	80.20 ± 0.16	96.21 ± 0.15	8.62 ± 0.03

13.3 Trigger efficiency

The efficiency for a signal decay to activate the desired trigger lines is determined both by the efficiency of the reconstruction algorithms used in the software and hardware trigger stages, and the cuts used in the respective trigger lines used to discriminate against background sources. The trigger efficiencies for the finally selected signal candidates is determined via

$$\epsilon_{\text{trig|sel}}(q^2) = \frac{N_{\text{trig}}(q^2)}{N_{\text{BDT}}(q^2)}, \quad (13.7)$$

where $N_{\text{trig}}(q^2)$ is the number of events passing the trigger selection. The trigger efficiencies determined from the respective simulated event samples are listed in Tab. 13.5 and 13.6.

13.4 Efficiency combination

The individual efficiencies presented in the previous sections are combined into a total efficiency defined in Eq. 13.1. Tables 13.7 and 13.8 show the resulting efficiencies for 2011 and 2012 per q^2 bin.

The rarity of the signal decay causes low statistics in most of the q^2 bins. For that reason, it is advantageous to analyse the full dataset collected in the 2011 and 2012 data-taking periods and not perform individual, less stable fits. Thus, the efficiencies

Table 13.5: Trigger efficiency for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2011 data taking period.

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{trig sel}} [\%]$
0.1 – 2.0	64.97 ± 0.43
2.0 – 5.0	71.19 ± 0.49
5.0 – 8.0	75.84 ± 0.47
11.0 – 12.5	84.50 ± 0.63
15.0 – 17.0	87.42 ± 0.65
17.0 – 19.0	88.91 ± 1.07
1.0 – 6.0	71.24 ± 0.37
15.0 – 19.0	87.83 ± 0.56

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{trig sel}} [\%]$
8.0 – 11.0	82.55 ± 0.06

Table 13.6: Trigger efficiency for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2012 data taking period.

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{trig sel}} [\%]$
0.1 – 2.0	65.44 ± 0.59
2.0 – 5.0	69.19 ± 0.66
5.0 – 8.0	74.99 ± 0.63
11.0 – 12.5	84.86 ± 0.86
15.0 – 17.0	88.98 ± 0.89
17.0 – 19.0	90.51 ± 1.45
1.0 – 6.0	69.74 ± 0.50
15.0 – 19.0	89.41 ± 0.76

$q^2 [\text{GeV}^2/c^4]$	$\epsilon_{\text{trig sel}} [\%]$
8.0 – 11.0	82.50 ± 0.18

determined independently for both data taking periods need to be combined in an overall efficiency describing this full dataset. As mentioned before, in order to reduce systematic uncertainties, the measurement of the differential branching fraction is performed relative to the control channel $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$. Thus the relative efficiencies $\frac{\epsilon_{\text{tot}}^{J/\psi\phi}}{\epsilon_{\text{tot}}^{\phi\mu^+\mu^-}(q^2)}$ between control and signal channel are needed for each q^2 bin (where the q^2 only changes for the signal channel $B_s^0 \rightarrow \phi\mu^+\mu^-$).

Table 13.7: Total efficiency for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2011 data taking period.

$q^2[\text{GeV}^2/c^4]$	$\epsilon_{\text{tot}} [\%]$
0.1 – 2.0	1.067 ± 0.013
2.0 – 5.0	1.116 ± 0.014
5.0 – 8.0	1.239 ± 0.014
11.0 – 12.5	1.396 ± 0.020
15.0 – 17.0	1.263 ± 0.017
17.0 – 19.0	0.928 ± 0.020
1.0 – 6.0	1.126 ± 0.011
15.0 – 19.0	1.148 ± 0.013

$q^2[\text{GeV}^2/c^4]$	$\epsilon_{\text{tot}} [\%]$
8.0 – 11.0	1.344 ± 0.004

Table 13.8: Total efficiency for $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays in bins of q^2 and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ mode decays for $8 < q^2 < 11 \text{ GeV}^2/c^4$ for the 2012 data taking period.

$q^2[\text{GeV}^2/c^4]$	$\epsilon_{\text{tot}} [\%]$
0.1 – 2.0	0.961 ± 0.013
2.0 – 5.0	0.958 ± 0.014
5.0 – 8.0	1.061 ± 0.013
11.0 – 12.5	1.248 ± 0.019
15.0 – 17.0	1.163 ± 0.018
17.0 – 19.0	0.869 ± 0.021
1.0 – 6.0	0.967 ± 0.010
15.0 – 19.0	1.061 ± 0.014

$q^2[\text{GeV}^2/c^4]$	$\epsilon_{\text{tot}} [\%]$
8.0 – 11.0	1.215 ± 0.005

Revisiting Eq. 10.1

$$\frac{1}{\mathcal{B}(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi)} \frac{d\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{dq^2} = \frac{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}{q_{\text{max}}^2 - q_{\text{min}}^2} \cdot \frac{N_{\phi\mu^+\mu^-}}{N_{J/\psi\phi}} \cdot \frac{\epsilon_{\text{tot}}^{J/\psi\phi}}{\epsilon_{\text{tot}}^{\phi\mu^+\mu^-}},$$

a proper combination can be derived, based only on signal candidates for $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ decays during the two data taking periods. While the (differential) branching fractions are not dependent on the centre-of-mass energy and thus identical in both 2011 and 2012, the yields and efficiencies for $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$

decays are different. Dividing the yields for both modes by years, the relative efficiency can be expressed as

$$\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}} = A \cdot \frac{N_{J/\psi\phi}(2011) + N_{J/\psi\phi}(2012)}{N_{\phi\mu^+\mu^-}(2011) + N_{\phi\mu^+\mu^-}(2012)}, \quad (13.8)$$

where $A = \frac{dB(B_s^0 \rightarrow \phi\mu^+\mu^-)}{B(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi)dq^2} \cdot \frac{q_{\text{max}}^2 - q_{\text{min}}^2}{B(J/\psi \rightarrow \mu^+\mu^-)}$ is constant for 2011 and 2012. The individual efficiencies for the two data taking periods can similarly be expressed as

$$\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}(2011) = A \cdot \frac{N_{J/\psi\phi}(2011)}{N_{\phi\mu^+\mu^-}(2011)} \quad (13.9)$$

and

$$\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}(2012) = A \cdot \frac{N_{J/\psi\phi}(2012)}{N_{\phi\mu^+\mu^-}(2012)}. \quad (13.10)$$

Combining these three equations, the full relative efficiency $\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}$ can be calculated from the individual efficiencies via

$$\begin{aligned} \frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}} = & \left[\frac{N_{J/\psi\phi}(2011)}{N_{J/\psi\phi}(2011) + N_{J/\psi\phi}(2012)} \cdot \left(\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}(2011) \right)^{-1} \right. \\ & \left. + \frac{N_{J/\psi\phi}(2012)}{N_{J/\psi\phi}(2011) + N_{J/\psi\phi}(2012)} \cdot \left(\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}(2012) \right)^{-1} \right]^{-1}. \end{aligned} \quad (13.11)$$

The relative $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ decay yields for 2011 and 2012 are determined from extended maximum likelihood fits to the reconstructed invariant mass to be

$$\frac{N_{J/\psi\phi}(2011)}{N_{J/\psi\phi}(2011) + N_{J/\psi\phi}(2012)} = 0.319 \pm 0.005$$

and

$$\frac{N_{J/\psi\phi}(2012)}{N_{J/\psi\phi}(2011) + N_{J/\psi\phi}(2012)} = 0.681 \pm 0.005,$$

respectively.

Table 13.9 shows the efficiency ratios $\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}$ for the 2011 and 2012 data taking periods and the respective combined total efficiency ratios for each q^2 bin.

13.5 Angular acceptance

While the efficiency ratios between $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays are important for the measurement of the relative branching fraction, they are not an

Table 13.9: Efficiency ratios between $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ mode decays in bins of q^2 for the 2011 and 2012 data taking periods and their combination.

q^2 [GeV ² /c ⁴]	$\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}$ (2011)	$\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}$ (2012)	$\frac{\varepsilon_{\text{tot}}^{J/\psi\phi}}{\varepsilon_{\text{tot}}^{\phi\mu^+\mu^-}}$
0.1 – 2.0	1.260 ± 0.015	1.264 ± 0.018	1.263 ± 0.015
2.0 – 5.0	1.204 ± 0.015	1.268 ± 0.019	1.247 ± 0.015
5.0 – 8.0	1.084 ± 0.013	1.145 ± 0.015	1.125 ± 0.013
11.0 – 12.5	0.963 ± 0.014	0.974 ± 0.016	0.970 ± 0.013
15.0 – 17.0	1.064 ± 0.015	0.961 ± 0.013	1.051 ± 0.014
17.0 – 19.0	1.449 ± 0.031	1.398 ± 0.035	1.414 ± 0.027
1.0 – 6.0	1.194 ± 0.012	1.257 ± 0.015	1.236 ± 0.013
15.0 – 19.0	1.171 ± 0.014	1.145 ± 0.016	1.153 ± 0.014

essential part of the analysis of the angular decay distributions. Instead, the distortion of the angular distributions by detector acceptance and trigger, reconstruction and selection effects are described by the angular acceptance, which is used in the fit to model the observed distributions. It is determined by comparing the generated angular distributions before the simulation of the detector response to the fully selected simulated event sample of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$. The latter is corrected for the known discrepancies between data and simulation, as described in Chap. 12. The angular acceptance ε_{acc} is calculated by dividing the four-dimensional distribution of the decay angles and q^2 in simulated $B_s^0 \rightarrow \phi(\rightarrow K^+K^-)\mu^+\mu^-$ events after the full detector simulation and with all selection requirements applied by that four-dimensional distribution at generator level. That way it describes the distortion induced by the detector and the signal candidate selection as a function of the decay angles and q^2 . The angular acceptance is normalised and subsequently parametrised in these four dimensions, modelled by Legendre polynomials

$$\varepsilon_{\text{acc}}(\theta_K, \theta_l, \Phi, q^2) = \sum_{i,j,k,l} c_{ijkl} L_i(\cos \theta_l) L_j(\cos \theta_K) L_k(\Phi) L_l(q^2), \quad (13.12)$$

where $L_{i,j,k,l}$ are the Legendre polynomials of the order i , j , k , and l , respectively. The coefficients c_{ijkl} are determined using the method of moments [57, pp. 467-472] via

$$c_{ijkl} = \frac{1}{N'} \sum_{e=1}^N w_e \left[\frac{2i+1}{2} \cdot \frac{2j+1}{2} \cdot \frac{2k+1}{2} \cdot \frac{2l+1}{2} \cdot L_i(\cos \theta_l) L_j(\cos \theta_K) L_k(\Phi) L_l(q^2) \right], \quad (13.13)$$

where N is the number of events in the respective q^2 bin, w_e is the correction weight for each simulated event, and $N' = \sum_{e=1}^N w_e$ is the effective number of simulated events. The method of moments is the simplest approach to parametrise distributions in terms of powers of the variables in which they are evaluated, in this case in the form

of Legendre polynomials. The main advantage compared to a maximum-likelihood fit to the distribution is that no analytical probability density function is needed. As the shape of the angular acceptance originates from a large array of detector and selection effects, deriving such a function based on proper models is close to impossible, and the method-of-moments approach preferable.

The maximum orders for each observable are chosen as $i_{max} = 4$, $j_{max} = 2$, $k_{max} = 6$, and $l_{max} = 5$, which corresponds to the lowest set of orders where adding additional orders does not improve the χ^2 when comparing parametrisation and distribution of the calculated acceptance. The exchange of the muon or kaon candidates corresponds to the transformations $\theta_l \rightarrow \theta_l - \pi$ and $\theta_K \rightarrow \theta_K - \pi$, respectively. As the detection and selection asymmetries of positively and negatively charged particles is negligible given the statistical precision of this analysis, assuming a symmetric angular acceptance under these exchanges is reasonable. For that reason, only even orders of the Legendre polynomials of the decay angles are used in the description of the acceptance description. Figure 13.1 shows the one-dimensional projections of the four-dimensional angular acceptance and the chosen parametrisation over the full q^2 range. The acceptance as function of all three angles as well as the q^2 acceptance are well described with the chosen parametrisation.

As the final fit to the angular distributions is performed in bins of q^2 , the angular acceptance for each bin is evaluated in the central q^2 value of the respective bin for computational reasons. The one-dimensional projections of the angular acceptance and its parametrisation for each of the q^2 bins are shown in App. A 1. In addition, two-dimensional projections for each combination of the decay angles can be found in App. A 2.

A potential systematic uncertainty for neglecting the q^2 dependency within the bins is evaluated and discussed together with further systematic uncertainties in Chap. 16.

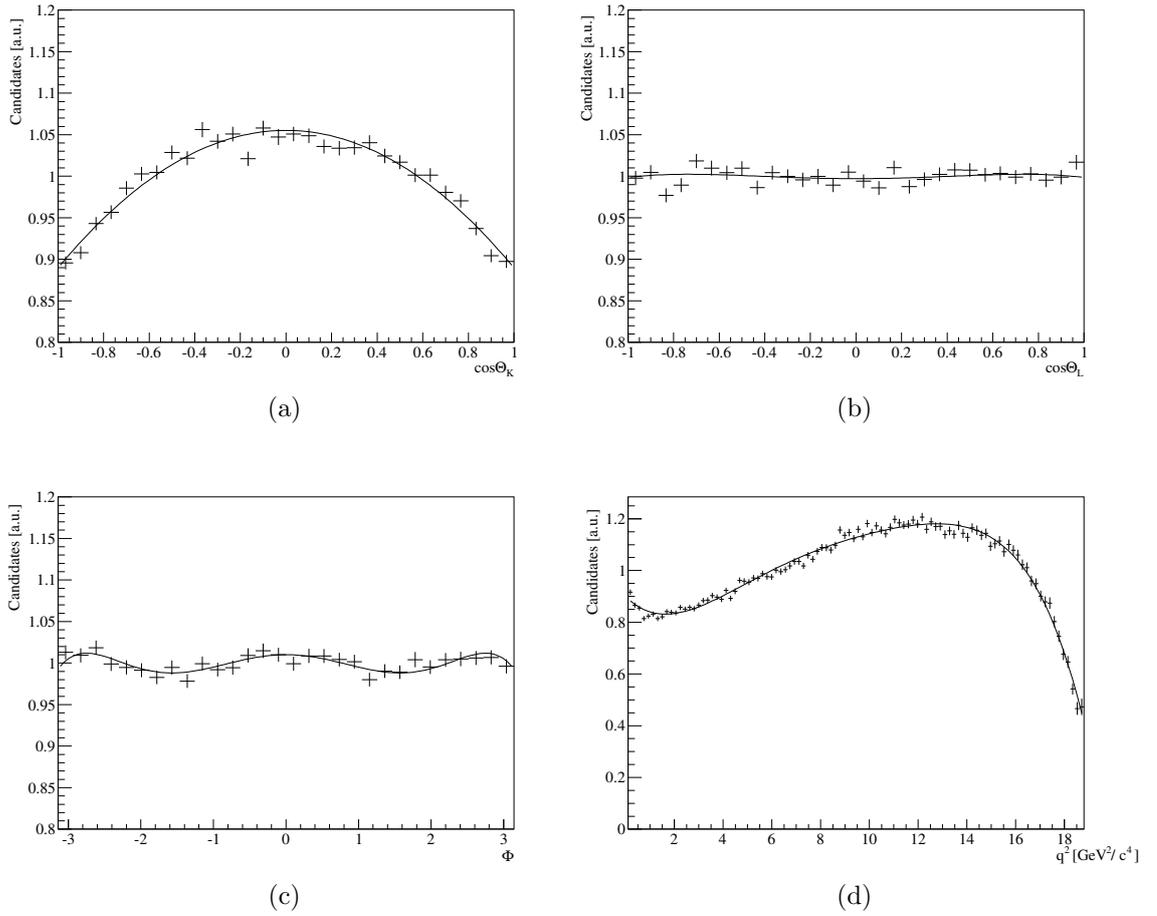


Figure 13.1: One-dimensional projections of the angular acceptance overlaid by the chosen parametrisation in the four observables: $\cos\theta_K$ (a), $\cos\theta_L$ (b), Φ (c), q^2 (d).

14 Relative branching fraction measurement

For the measurement of the relative branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ compared to the normalisation channel $B_s^0 \rightarrow J/\psi \phi$, in addition to the relative efficiencies, discussed in Chap. 13, the yields of decays of the signal and normalisation channels are required in bins of the dimuon mass squared q^2 . These yields are determined from an extended maximum-likelihood fit to the distribution of the reconstructed invariant mass of the B_s^0 -meson candidate m_{rec} , which is calculated from the sum of the four-momenta of the final-state particles via

$$m_{rec} = \sqrt{(E_{K^+} + E_{K^-} + E_{\mu^+} + E_{\mu^-})^2 - (\vec{p}_{K^+} + \vec{p}_{K^-} + \vec{p}_{\mu^+} + \vec{p}_{\mu^-})^2}. \quad (14.1)$$

Here the E_i are the energies and \vec{p}_i are the spatial momentum vectors of the final-state particles. The distribution of m_{rec} for all selected $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay candidates integrated over the six q^2 bins is shown in Fig. 14.1. In this chapter

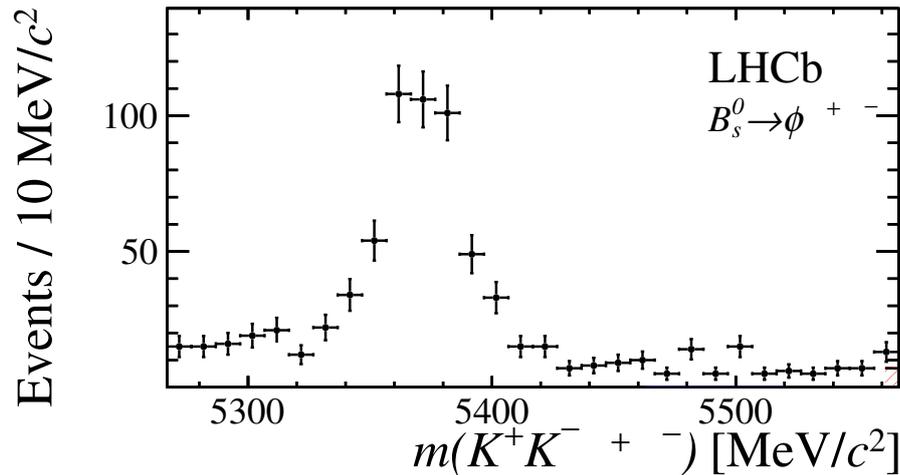


Figure 14.1: Distribution of the reconstructed invariant mass of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay candidates integrated over the full investigated q^2 range.

a short introduction is given on the topic of extended maximum likelihood fits in order to establish the fitting technique used to extract the signal yields. Signal and background models are discussed and the fit results are shown.

14.1 Extended Maximum Likelihood Fits

In order to extract the signal and normalisation mode yields, an extended maximum likelihood fit is performed to the reconstructed invariant mass in each q^2 bin. The maximum likelihood technique [58] is a fitting procedure using an i -dimensional

probability density function $p(\vec{x}|\vec{a})$, which depends on a set of n parameters $\vec{a} = (a_1, \dots, a_n)$ describing a number of i measured observables $\vec{x} = (x_1, \dots, x_i)$ for each event. The probability density function has to be normalised to fulfil

$$\int_{\Omega} p(\vec{x}|\vec{a})d\vec{x} = 1 \quad \forall \vec{a}, \quad (14.2)$$

where Ω is the allowed range for \vec{x} . For a number N of measured events, the overall likelihood function can be defined as the product of the probability density function of each event:

$$\mathcal{L}(\vec{a}) = \prod_{j=1}^N p(\vec{x}_j|\vec{a}). \quad (14.3)$$

The likelihood function is thus a function of the set of parameters \vec{a} and reflects the probability for the given set of N events to return the measured observables. Maximising this likelihood yields a set of parameters \vec{a} , which is the best estimate of the true values for those parameters, according to the maximum-likelihood-principle. Computationally, maximising the likelihood is typically achieved by minimising the negative logarithm of the likelihood

$$\mathcal{F}(\vec{a}) = -\ln \mathcal{L}(\vec{a}) = -\sum_{j=1}^N \ln [p(\vec{x}_j|\vec{a})], \quad (14.4)$$

as large sums can be evaluated much faster than products.

However, the definition of the probability density function can be defined more generally by removing the normalisation constraint from Eq. 14.2. The normalisation of this probability density function \mathcal{P} is then given by

$$\int_{\Omega} \mathcal{P}(\vec{x}|\vec{a})d\vec{x} = \mathcal{N}(\vec{a}). \quad (14.5)$$

Due to this definition, \mathcal{P} does not only describe the shape of the measured observables, but also the size of the dataset. This is advantageous when measuring a small number of events, where their observed number is best described by Poisson statistics. Due to the fluctuation of those Poisson statistics, the expected number of observed events \mathcal{N} will generally differ from the number of observed events N . The extended likelihood for a given event sample is then modified due to the Poissonian nature to

$$\mathcal{L}_e(\vec{a}) = \frac{e^{-\mathcal{N}}}{N!} \prod_{j=1}^N \mathcal{P}(\vec{x}_j|\vec{a}). \quad (14.6)$$

Calculating the negative logarithm of this extended likelihood yields

$$\mathcal{F}_e(\vec{a}) = -\ln \mathcal{L}_e(\vec{a}) = -\sum_{j=1}^N \ln [\mathcal{P}(\vec{x}_j|\vec{a})] + \mathcal{N}(\vec{a}) + \ln(N!), \quad (14.7)$$

where the difference in normalisation of \mathcal{P} compared to p is largely cancelled by the \mathcal{N} term and the constant $\ln(N!)$ term plays no role in the minimisation. Thus the extended maximum likelihood yields the same best estimate for the parameters \vec{a} , in case that the number of observed events N is equal to the expected number of observed events \mathcal{N} . In addition, allowing \mathcal{N} to float freely in the maximisation of the likelihood automatically yields the best estimate for the actual number of the observed events.

14.2 Signal and background models

Due to the rarity of the $B_s^0 \rightarrow \phi\mu^+\mu^-$ decay in the SM, the extended maximum likelihood fitting technique is best suited to extract the signal and background yields from the small event sample. The probability density function used to describe the fully selected event samples for signal and normalisation modes consists of a model for the signal component and one for the combinatorial background. Additional background sources are reduced by the selection to a level at which they can be neglected in the fit model, a remaining uncertainty tied to this assumption is investigated in Chap. 16.

The yields for signal and background in both measured decay modes is extracted from a fit to the reconstructed invariant mass distribution. The model best suited to describe the signal shape is determined using the fully corrected simulated event samples. For the signal component the sum of two Crystal-Ball functions [20] agrees well with the observed distributions. The Crystal Ball function is defined as a Gaussian distribution with a tail for values below the Gaussian peak. This models the energy losses due to final-state radiation, and thus give a better estimate than Gaussian distributions. This tail is modelled by two additional parameters α_{CB} and n_{CB} , which are shared between the two Crystal Ball functions. The full signal model is given by

$$\begin{aligned} \mathcal{P}_{\text{sig}}(m_{\text{rec}}|f_{\text{sig}}, \mu_{\text{m}}, \sigma_{m,1}, \sigma_{m,2}, \alpha_{CB}, n_{CB}) = & \mathcal{N}_{\text{sig}}[f_{\text{sig}} \cdot \mathcal{CB}(m_{\text{rec}}|\mu_{\text{m}}, \sigma_{m,1}, \alpha_{CB}, n_{CB}) \\ & + (1 - f_{\text{sig}}) \cdot \mathcal{CB}(m_{\text{rec}}|\mu_{\text{m}}, \sigma_{m,2}, \alpha_{CB}, n_{CB})], \end{aligned} \quad (14.8)$$

where the parameter f_{sig} gives the relative contribution of the first Crystal Ball function to the sum, μ_{m} the mean and $\sigma_{m,1/2}$ the widths of the Gaussian kernels of the Crystal Ball function. As both Crystal both functions are normalised, the normalisation of the full signal probability density function is equal to \mathcal{N}_{sig} . Due to the rarity of the signal decay $B_s^0 \rightarrow \phi\mu^+\mu^-$, most shape parameters are determined from the fit to the normalisation mode $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$, and fixed to these values in the signal mode fits. This method is validated using the simulated event samples of the decays $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$. However, the resolution of the reconstructed invariant mass is dependent on the momentum of the final-state tracks and thus on the dimuon mass squared q^2 . To account for this, a scale factor $s = \sigma_{m, B_s^0 \rightarrow \phi\mu^+\mu^-} / \sigma_{m, B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi}$ is determined from fits to the

Table 14.1: Scale factors of the $B_s^0 \rightarrow \phi\mu^+\mu^-$ peak widths obtained from fits to simulated event samples.

q^2 [(GeV/c ²) ²]	s
0.1 – 2.0	0.975 ± 0.006
2.0 – 5.0	0.977 ± 0.006
5.0 – 8.0	0.984 ± 0.006
11.0 – 12.5	1.017 ± 0.007
15.0 – 17.0	1.059 ± 0.007
17.0 – 19.0	1.056 ± 0.011
1.0 – 6.0	0.978 ± 0.005
15.0 – 19.0	1.058 ± 0.006

simulated event samples for each q^2 bin. Subsequently, in the fit to the $B_s^0 \rightarrow \phi\mu^+\mu^-$ mode data sample, the widths of the signal Crystal Ball functions are left floating within Gaussian constraints around the values obtained from these fits. Table 14.1 shows the values obtained for the scale factors.

The model chosen for the background component is an exponential function, where the only parameter is the exponent Λ , normalised such that the number of background events corresponds to \mathcal{N}_{bkg}

$$\mathcal{P}_{\text{bkg}}(m_{\text{rec}}|\Lambda) = \mathcal{N}_{\text{bkg}} \cdot \frac{\Lambda}{e^{-\Lambda \cdot m_{\text{min}}} - e^{-\Lambda \cdot m_{\text{max}}}} \cdot e^{-\Lambda \cdot m_{\text{rec}}}. \quad (14.9)$$

Here m_{min} and m_{max} are the lower and upper bound of the fit range chosen, respectively. It describes the reconstructed invariant mass well both below and above the signal region of $\pm 50 \text{ MeV}/c^2$ around the nominal B_s^0 mass.

The full model used in the fit to the reconstructed invariant mass is then given by

$$\mathcal{P}(m_{\text{rec}}|f_{\text{sig}}, \mu_{\text{m}}, \sigma_{m,1}, \sigma_{m,2}, \alpha_{CB}, n_{CB}, \Lambda) = \mathcal{P}_{\text{sig}} + \mathcal{P}_{\text{bkg}}, \quad (14.10)$$

which has a normalisation of $\mathcal{N} = \mathcal{N}_{\text{sig}} + \mathcal{N}_{\text{bkg}}$.

Different model choices and their respective influence on the fit results are investigated when determining the systematic uncertainties in Chap. 16.

14.3 Signal and normalisation fit results

The fits to the reconstructed invariant mass are performed individually for signal and normalisation modes, for the former both for the full sample as well as in bins of the dimuon mass squared q^2 . First the fit to the normalisation mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+\mu^-)\phi$ is performed. All model parameters are left floating, in order to fix the signal shape parameters as described above. The result of the fit is shown in Fig. 14.2, both in decimal as well as logarithmic presentation. As can be seen,

Table 14.2: Parameters of the fit to the reconstructed invariant mass of $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decay candidates.

Parameter	Best fit value	Uncertainty
\mathcal{N}_{sig}	64414.8	270.7
\mathcal{N}_{bkg}	385.7	112.9
μ_m	5371.410 MeV/ c^2	0.096 MeV/ c^2
f_{sig}	0.852	0.015
$\sigma_{m,1}$	15.85 MeV/ c^2	0.17 MeV/ c^2
$\sigma_{m,2}$	35.5 MeV/ c^2	1.7 MeV/ c^2
α_{CB}	1.480	0.069
n_{CB}	10.4	4.6
Λ	$0.00346 (\text{MeV}/c^2)^{-1}$	$0.00018 (\text{MeV}/c^2)^{-1}$

especially from the logarithmic presentation, the data points are well described by the chosen model. The extracted parameters and the respective fit uncertainties are given in Tab. 14.2. Due to the extended maximum likelihood fit, the signal and background yields are allowed to float independently and are not limited by the observed number of candidates, leading to different uncertainties for those two yields.

The values for μ_m , f_{sig} , α_{CB} , and n_{CB} are fixed from this fit for the $B_s^0 \rightarrow \phi \mu^+ \mu^-$

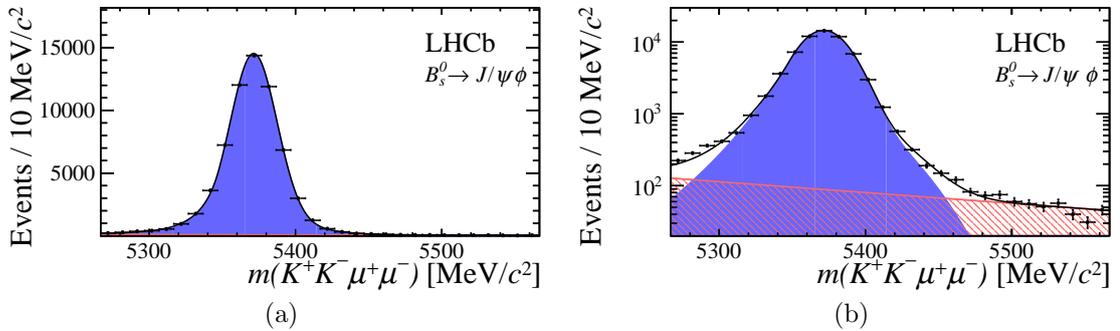


Figure 14.2: Fitted distribution of $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decay candidates in normal (a) and logarithmic (b) presentation. Data points are given in black, the projection of signal model by the blue area, the background component by the red shaded area, and the full fit model by the black line.

fits, while $\sigma_{m,1}$, $\sigma_{m,2}$ are each multiplied by the scale factors s as given in Tab. 14.1. The scale factors are left floating within Gaussian constraints, using the uncertainties obtained from the fits to simulated events. The yield parameters \mathcal{N}_{sig} and \mathcal{N}_{bkg} and the background slope parameter Λ are allowed to freely float independently for each

q^2 bin. The fitted distributions and fit projections are shown in Fig. 14.3. In general similar signal and background shapes can be observed, while both are well described by the fit model. For the q^2 bin with the range $17 < q^2 < 19 \text{ GeV}^2/c^4$ (shown in Fig. 14.3 (f)) the background is rising towards high reconstructed invariant masses, which is in clear contrast to the other bins. This behaviour is caused by the region of phase space observed in this bin. The minimum q^2 of $17 \text{ GeV}^2/c^4$ in combination with the ϕ -meson mass of about $1020 \text{ GeV}/c^2$ constrains the reconstructed invariant mass to be larger than $\approx 5.14 \text{ GeV}/c^2$. This strongly reduces the probability to reconstruct low invariant masses for the $K^+K^-\mu^+\mu^-$ combination for this q^2 bin. Table 14.3 shows the corresponding resulting parameter values and uncertainties for each q^2 bin. Due to the low statistics available in the data sample the uncertainties are in general asymmetrical, as the likelihood does not have a parabolic shape around the best fit value. Symmetrical uncertainties would correspond to the Gaussian limit of Poissonian statistics, which is only valid for sufficiently high statistics. As the background yield is not further used, the average of upper and lower uncertainty is given in the table.

14.4 Relative differential branching fraction calculation

With the yields extracted from the extended maximum likelihood fits presented in this chapter as well as the efficiencies determined in Chap. 13, only the branching fraction of the decay $J/\psi \rightarrow \mu^+\mu^-$ is left to calculate the relative differential branching fraction of $B_s^0 \rightarrow \phi\mu^+\mu^-$ compared to $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ decays, as given by Eq. 10.1. The world average measured value of $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.961 \pm 0.033)\%$ as reported by the Particle Data Group [15] is used for this calculation. Table 14.4 summarises the components for the calculation and the resulting relative differential branching fractions in bins of q^2 . All given uncertainties are purely of statistical nature. Systematic uncertainties will be discussed in Chap. 16, and the final results will be compared to theoretical expectations in Chap. 17.

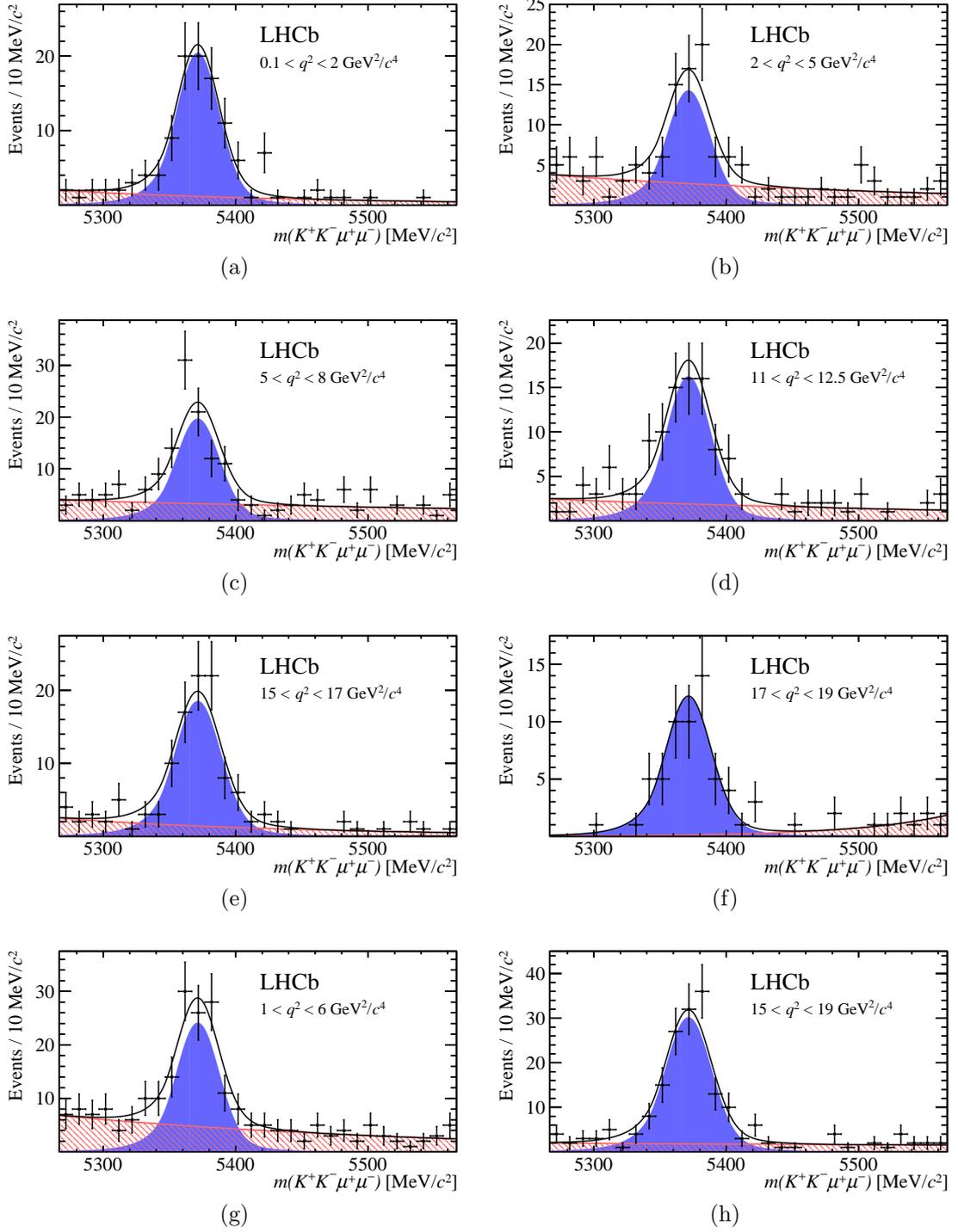


Figure 14.3: Fitted distribution of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay candidates in the eight q^2 bins. Data points are given in black, the projection of signal model by the blue area, the background component by the red shaded area, and the full fit model by the black line.

Table 14.3: Parameters of the fit to the reconstructed invariant mass of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decay candidates for each q^2 bin.

q^2 [GeV ² /c ⁴]	Parameter	Best fit value	Uncertainty
0.1 – 2.0	\mathcal{N}_{sig}	85.21	+10.57 -10.02
	\mathcal{N}_{bkg}	30.4	7.4
	Λ	0.0050 (MeV/c ²) ⁻¹	0.0023 (MeV/c ²) ⁻¹
2.0 – 5.0	\mathcal{N}_{sig}	59.62	+9.76 -9.15
	\mathcal{N}_{bkg}	71.9	10.3
	Λ	0.0032 (MeV/c ²) ⁻¹	0.0014 (MeV/c ²) ⁻¹
5.0 – 8.0	\mathcal{N}_{sig}	82.72	+11.43 -10.83
	\mathcal{N}_{bkg}	91.9	11.8
	Λ	0.0018 (MeV/c ²) ⁻¹	0.0013 (MeV/c ²) ⁻¹
11.0 – 12.5	\mathcal{N}_{sig}	70.37	+10.34 -9.75
	\mathcal{N}_{bkg}	51.5	9.3
	Λ	0.0024 (MeV/c ²) ⁻¹	0.0018 (MeV/c ²) ⁻¹
15.0 – 17.0	\mathcal{N}_{sig}	83.11	+10.41 -9.90
	\mathcal{N}_{bkg}	36.9	7.9
	Λ	0.0053 (MeV/c ²) ⁻¹	0.0021 (MeV/c ²) ⁻¹
17.0 – 19.0	\mathcal{N}_{sig}	54.22	+7.80 -7.41
	\mathcal{N}_{bkg}	13.0	4.4
	Λ	-0.0139 (MeV/c ²) ⁻¹	0.0072 (MeV/c ²) ⁻¹
1.0 – 6.0	\mathcal{N}_{sig}	101.05	+12.81 -12.20
	\mathcal{N}_{bkg}	127.0	13.8
	Λ	0.0034 (MeV/c ²) ⁻¹	0.0011 (MeV/c ²) ⁻¹
15.0 – 17.0	\mathcal{N}_{sig}	135.51	+13.17 -12.68
	\mathcal{N}_{bkg}	52.3	9.6
	Λ	0.0008 (MeV/c ²) ⁻¹	0.0018 (MeV/c ²) ⁻¹

Table 14.4: $B_s^0 \rightarrow \phi\mu^+\mu^-$ decay yields, relative total efficiencies, and resulting relative differential branching fractions in bins of q^2 . Uncertainties given are statistical only.

$q^2 [\text{GeV}^2/c^4]$	$\mathcal{N}_{sig, \phi\mu^+\mu^-}$	$\frac{\epsilon_{tot}^{J/\psi\phi}}{\epsilon_{tot}^{\phi\mu^+\mu^-}}$	$\frac{d\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{\mathcal{B}(B_s^0 \rightarrow J/\psi\phi) dq^2} [10^{-5} \text{GeV}^{-2} c^4]$
0.1 – 2.0	$85.21^{+10.57}_{-10.02}$	1.263 ± 0.015	$5.44^{+0.68}_{-0.64}$
2.0 – 5.0	$59.62^{+9.76}_{-9.15}$	1.247 ± 0.015	$2.38^{+0.39}_{-0.37}$
5.0 – 8.0	$82.72^{+11.43}_{-10.83}$	1.125 ± 0.013	$2.98^{+0.41}_{-0.39}$
11.0 – 12.5	$70.37^{+10.34}_{-9.75}$	0.970 ± 0.013	$4.37^{+0.64}_{-0.61}$
15.0 – 17.0	$83.11^{+10.41}_{-9.90}$	1.051 ± 0.014	$4.20^{+0.53}_{-0.50}$
17.0 – 19.0	$54.22^{+7.80}_{-7.41}$	1.414 ± 0.027	$3.68^{+0.53}_{-0.50}$
1.0 – 6.0	$101.05^{+12.81}_{-12.20}$	1.236 ± 0.013	$2.40^{+0.30}_{-0.29}$
15.0 – 19.0	$135.51^{+13.17}_{-12.68}$	1.153 ± 0.014	$3.75^{+0.37}_{-0.35}$

15 Analysis of the angular distributions of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$

Due to the different Lorentz structures of the processes involved in the decay of hadronic particles, analysing the angular distributions of the decay products gives insight into the particles mediating the respective decay. New Physics processes can introduce Lorentz structures which do not occur in the SM decay and thus significantly affect these angular distributions. For the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$, eight q^2 -dependent angular observables F_L , $A_{5,6,8,9}$, and $S_{3,4,7}$ can be extracted from a fit to the three-dimensional distribution in the decay angles θ_l , θ_K , and Φ , introduced in Chap. 9. As the fully selected sample of signal candidates is subject to combinatorial background, the distribution of the reconstructed invariant mass is used to distinguish the angular distributions of signal candidates from this background. In this chapter the fitting process and models used for signal and background are described, and the measured values for the angular observables presented as a function of q^2 .

15.1 Fitting procedure and model

For the determination of the eight q^2 -dependent angular observables the actual number of observed signal events plays a negligible role, as these parameters are tied to the shape of the angular distributions, and not the normalisation. Thus the standard maximum-likelihood technique (see Chap. 14.1) is used to extract their values from the data sample. Fits are performed individually for each of the eight q^2 bins. The probability density function used for those fits consists of a signal and a background component, each four-dimensional, to describe the distributions in the three angles and the reconstructed invariant mass. The latter is used to improve separation of signal and background, no yields are extracted. Furthermore, the reconstructed invariant mass distribution is considered independent of the angular distributions, which leads to a factorisation of the signal and background probability density functions

$$\mathcal{P}(m_{rec}, \cos \theta_l, \cos \theta_K, \Phi) = \mathcal{P}_m(m_{rec}) \cdot \mathcal{P}_{ang}(\cos \theta_l, \cos \theta_K, \Phi). \quad (15.1)$$

The functions used to describe the mass distributions of signal and background are using the same shape parameters as in Chap. 14, but are each normalised to 1 due to omitting the signal and background yields:

$$\begin{aligned} \mathcal{P}_{sig}(m_{rec} | f_{CB}, \mu_m, \sigma_{m,1}, \sigma_{m,2}, \alpha_{CB}, n_{CB}) &= f_{CB} \cdot \mathcal{CB}(m_{rec} | \mu_m, \sigma_{m,1}, \alpha_{CB}, n_{CB}) \\ &+ (1 - f_{CB}) \cdot \mathcal{CB}(m_{rec} | \mu_m, \sigma_{m,2}, \alpha_{CB}, n_{CB}) \end{aligned} \quad (15.2)$$

and

$$\mathcal{P}_{bkg}(m_{rec} | \Lambda) = \frac{\Lambda}{e^{-\Lambda \cdot m_{min}} - e^{-\Lambda \cdot m_{max}}} \cdot e^{-\Lambda \cdot m_{rec}}. \quad (15.3)$$

The signal model for the angular distribution is given by the differential decay rate discussed in Eq. 9.40, where the effects of the angular acceptance (see Chap. 13.5) have to be corrected for. Therefore, the probability density function, which in the (in the decay angles) unbinned maximum likelihood fit is evaluated for each event, is multiplied by the angular acceptance:

$$\mathcal{P}_{\text{sig}}(\cos \theta_l, \cos \theta_K, \Phi | \vec{a}) = N \frac{1}{\Gamma + \bar{\Gamma}} \frac{d^4(\Gamma + \bar{\Gamma})}{dq^2 d \cos \theta_l d \cos \theta_K d \Phi} \cdot \epsilon_{\text{acc}}(\theta_K, \theta_l, \Phi, q^2), \quad (15.4)$$

where $\vec{a} = (F_L, A_{5,6,8,9}, S_{3,4,7})$ corresponds to the set of angular fit parameters, and N is the proper normalisation.

In contrast, the angular background model is determined purely from data. However, due to the efficient selection requirements, the background levels are extremely low, especially when evaluated in bins of q^2 . For that reason, its shape is not determined in the nominal fit, but from candidates in the background-enriched reconstructed invariant mass region of $5466.6 < m_{\text{rec}} < 5566.6 \text{ MeV}/c^2$. With the full selection applied, there are only about 80 candidates left within this reconstructed invariant mass region. Thus the selection cut on the BDT classifier, which is used to reduce this combinatorial background is removed, increasing the statistics available to 1140 candidates. In order to validate that no correlation occurs between the BDT response and the distribution of the decay angles, the latter is compared for a BDT response of < -0.1 and > -0.1 , shown in Fig. 15.1. The angular distributions agree well within the statistical precision.

Using this background-enriched sample, the distributions of the decay angles are parametrised using a product of second order Chebyshev polynomials for each decay angle. The angular background model is then given by

$$\begin{aligned} \mathcal{P}_{\text{bkg}}(\cos \theta_l, \cos \theta_K, \Phi) &= f(\cos \theta_l) \cdot f(\cos \theta_K) \cdot f(\Phi) \\ &= [1 + c_1^{\theta_l} P_1^{Ch}(\cos \theta_l) + c_2^{\theta_l} P_2^{Ch}(\cos \theta_l)] \\ &\quad \cdot [1 + c_1^{\theta_K} P_1^{Ch}(\cos \theta_K) + c_2^{\theta_K} P_2^{Ch}(\cos \theta_K)] \\ &\quad \cdot [1 + c_1^{\Phi} P_1^{Ch}(\Phi) + c_2^{\Phi} P_2^{Ch}(\Phi)], \end{aligned} \quad (15.5)$$

where P_i^{Ch} is the Chebyshev polynomial of i -th order, and the coefficients c_i are determined from the background-enriched sample described above. They are extracted by fitting the corresponding angular distributions in bins of q^2 . The fitted distributions and corresponding parametrisations are shown in Figs. 15.2 to 15.4. The resulting coefficients are listed in Tab. 15.1 in bins of q^2 , with statistical uncertainties as returned from the fit. Due to the limited statistical precision even in the background-enriched sample, the uncertainties are significant. The coefficients determined that way are fixed in the final fit. That way, the only free parameters of the background model are the slope parameter Λ in the reconstructed invariant mass description (see Eq. 14.9) and the fraction of background events in each q^2 bin. The small size of the data set does not allow to leave the coefficients c_i floating within

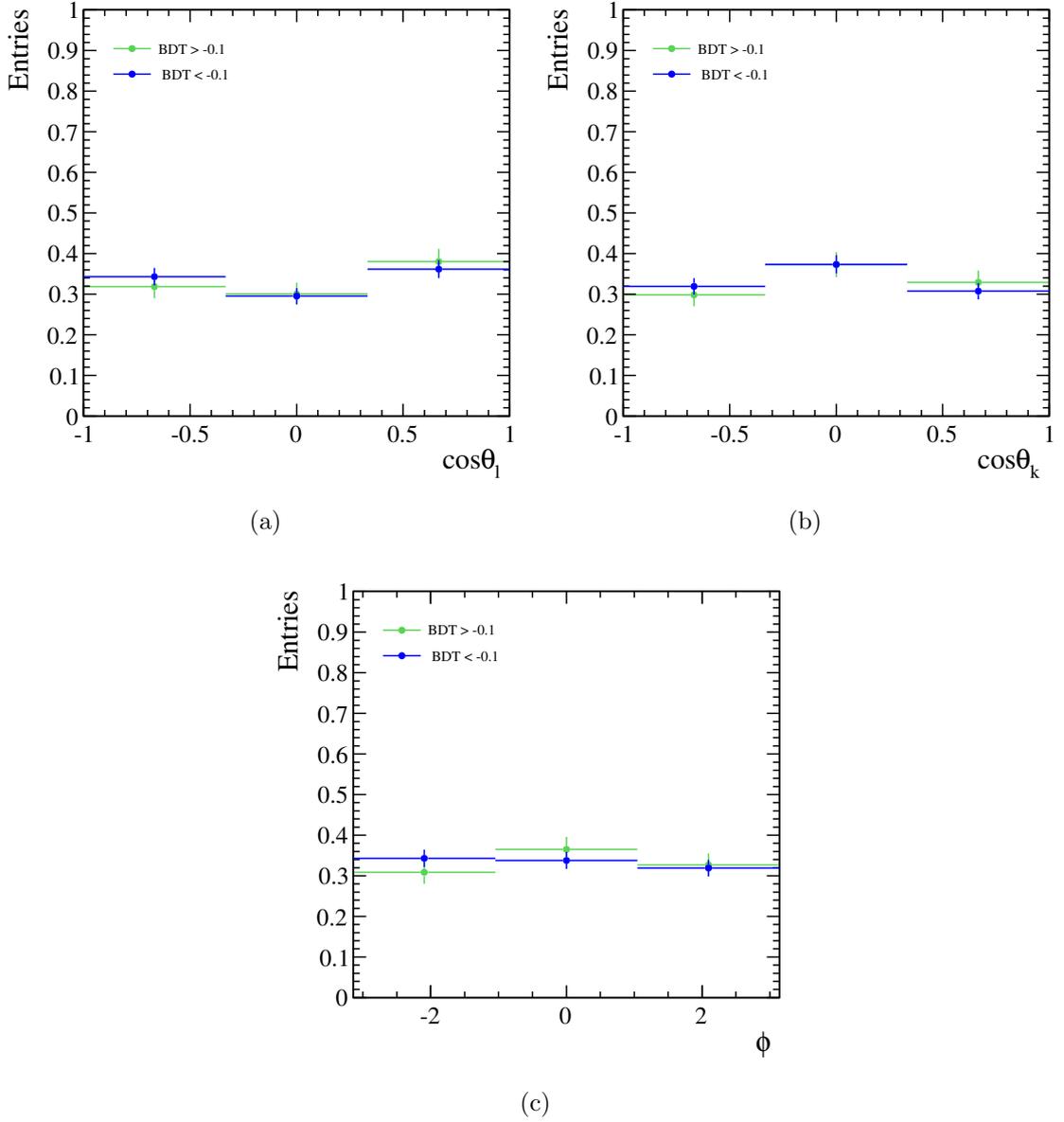


Figure 15.1: Comparison of the normalised distributions of background events with $5466.6 < m_{rec} < 5566.6 \text{ MeV}/c^2$ for low and high BDT response in the decay angles $\cos\theta_l$ (a), $\cos\theta_K$ (b), and Φ (c).

their uncertainties in the nominal fit. For that reason, the choice to use fixed values and the number of orders used in the parametrisation are investigated as a source of systematic uncertainty in Chap. 16.

Finally, these models are combined into the full, four-dimensional probability

Table 15.1: Parameters of the angular background model determined from fits to the respective distributions of a background-enriched sample in bins of q^2 .

$q^2[\text{GeV}^2/c^4]$	$c_1^{\theta_l}$	$c_2^{\theta_l}$	$c_1^{\theta_K}$	$c_2^{\theta_K}$	c_1^Φ	c_2^Φ
0.1 – 2.0	0.43 ± 0.19	0.23 ± 0.17	0.23 ± 0.19	0.30 ± 0.17	-0.26 ± 0.20	0.03 ± 0.19
2.0 – 5.0	0.05 ± 0.10	0.21 ± 0.09	0.02 ± 0.10	-0.12 ± 0.10	-0.17 ± 0.10	0.03 ± 0.09
5.0 – 8.0	0.03 ± 0.09	0.10 ± 0.09	0.07 ± 0.09	-0.05 ± 0.09	0.03 ± 0.09	0.01 ± 0.08
11.0 – 12.5	0.00 ± 0.14	-0.13 ± 0.15	-0.24 ± 0.13	-0.61 ± 0.15	0.02 ± 0.14	0.09 ± 0.12
15.0 – 17.0	0.24 ± 0.15	-0.05 ± 0.16	0.20 ± 0.16	-0.22 ± 0.17	0.27 ± 0.15	-0.08 ± 0.15
17.0 – 19.0	-0.07 ± 0.20	-0.44 ± 0.26	-0.38 ± 0.19	-0.40 ± 0.23	0.16 ± 0.20	-0.11 ± 0.19
1.0 – 6.0	0.05 ± 0.08	0.23 ± 0.07	0.05 ± 0.08	-0.07 ± 0.08	-0.13 ± 0.08	0.06 ± 0.07
15.0 – 19.0	0.12 ± 0.12	-0.18 ± 0.14	-0.01 ± 0.12	-0.27 ± 0.14	0.22 ± 0.12	-0.09 ± 0.12

density function

$$\mathcal{P}(m_{rec}, \cos \theta_l, \cos \theta_K, \Phi) = f_{\text{sig}} \cdot \mathcal{P}_{\text{sig}}(m_{rec}, \cos \theta_l, \cos \theta_K, \Phi) + (1 - f_{\text{sig}}) \cdot \mathcal{P}_{\text{bkg}}(m_{rec}, \cos \theta_l, \cos \theta_K, \Phi), \quad (15.6)$$

where f_{sig} is the fraction of signal events and $\mathcal{P}_{\text{sig/bkg}}$ are the signal and background models as described above, respectively.

15.2 Fit validation

The fitting procedure is validated by three independent checks: a fit to the fully simulated sample of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ events, a large set of pseudo-experiments, and a fit to the data sample of $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decay candidates.

15.2.1 Fit to the fully simulated event sample

The fully simulated sample of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays is fitted with the model as described in this section and the resulting fit parameters compared to the generated values. In order to extract the parameter values used in the generation, a large sample of 10 million events is generated, for which the raw angular distributions are saved without a simulation of the detector response, reconstruction or selection steps. These distributions are fitted with the angular signal model without the angular acceptance description. The results are compared to a fit to the fully simulated $B_s^0 \rightarrow \phi \mu^+ \mu^-$ events, with the full signal model including the angular acceptance as described in Chap. 13.5. The uncertainties of the parameters extracted from the fit to the generated event sample are negligible due to the high statistics. The resulting q^2 dependent values for each angular fit parameter are compared in Fig.15.5. All values agree well within statistical uncertainties.

15.2.2 Pseudo-experiments

In order to probe whether the fitting framework used for the unbinned maximum likelihood fit returns unbiased results, a set of 500 pseudo-experiments is performed.

Each pseudo-dataset consists of 50000 events, for which the signal and background events are created using the same parameters as in the default simulated event samples and according to the models described above, including all angular acceptance effects. The fraction of signal events and all fit parameters are extracted from the nominal fit to each pseudo-dataset and compared to the respective generated values. The pull p for each pseudo-experiment and parameter is defined as

$$p = \frac{\mu_{\text{fit}} - \mu_{\text{gen}}}{\sigma_{\mu}}, \quad (15.7)$$

where σ_{μ} is the statistical uncertainty of the fit result and μ_{fit} and μ_{gen} the fitted and generated values, respectively. For an unbiased fitting framework, the distribution of pulls over the full set of 500 pseudo-experiments follows a normal distribution with a mean of 0 and a width of 1, the latter due to the normalisation by the statistical uncertainty. The pull distributions for all parameters are shown in Fig. 15.6. As expected, no signs of bias is found in any of the fit parameters, validating the proper function of the fitting framework.

15.2.3 Fit of the control mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$

As a further validation of the fitting procedure and primary check of the proper description of the angular acceptance effects, the control channel $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ is investigated. For this decay mode an analysis of the distribution of the decay angles has been published before by the LHCb experiment [59], the results of which are used as reference values. As the angular acceptance is in general dependent on q^2 , which has a much wider range for $B_s^0 \rightarrow \phi \mu^+ \mu^-$ than for $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decays, the acceptance used in this check is evaluated separately from a simulated sample of $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decays. It is determined and parametrised in the same way as done for the nominal signal fit, as described in Chap. 13.5. The reference values for the fit parameters and the results of the $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ fit are listed in Tab. 15.2. All values agree within the given uncertainties. This procedure is also an independent test of the determination of the angular acceptance.

15.3 Uncertainty evaluation

Due to the many terms entering Eq. 9.40, there are certain combinations of the fit parameters for which the differential decay rate (and thus the probability density function) can take on negative values. As this is an unphysical result, this is prevented by the fitting framework by adding penalty terms to the likelihood function, which imposes boundaries upon the parameter space. Scans of the allowed range of all combinations of two parameters are shown in App. A 3. For many q^2 bins the expected values for the fit parameters from Standard Model predictions are close to the boundaries of this parameter space in terms of the expected statistical precision. Thus, due to the limited statistical precision caused by the small data sample, the

Table 15.2: Results of a fit to the control mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ candidates, compared to reference values from the formerly published LHCb measurement [59].

Parameter	Reference value	Fit result
F_L	0.506 ± 0.011	0.500 ± 0.003
S_3	0.030 ± 0.007	0.032 ± 0.004
S_4	-0.233 ± 0.008	-0.222 ± 0.004
S_7	0.0	0.002 ± 0.004
A_5	0.0	0.000 ± 0.004
A_6	0.0	0.002 ± 0.004
A_8	0.0	0.004 ± 0.005
A_9	0.0	-0.002 ± 0.004

likelihood might take on a non-parabolic shape. For that reason, the uncertainty intervals for each fit parameter are determined by profile likelihood scans using the Feldman-Cousins method [60].

For a likelihood scan each fit parameter is varied over a certain range around the nominal fit result with a specified number of steps (200 steps are used for this analysis). For each of these steps the respective parameter is fixed to the assigned value while the fit is performed with all other parameters floating, and the difference in log likelihood (DLL) is calculated with respect to the nominal fit.

A profile likelihood scan using the Feldman-Cousins method is a more precise version of a likelihood scan. The DLL at each step of the scan is not determined by one fit, but instead 500 pseudo-experiments are created. For each of these pseudo-experiments a dataset is created mimicking the data sample in all parameters but randomly fluctuated within fit uncertainties. These fluctuations take the correlations between the parameters derived from the fit into account. Subsequently, the fit is performed with the desired parameter fixed to the value assigned to the respective step and all other parameters floating, and the DLL calculated compared to a fit with all parameters being floated. The likelihood for each step of the scan is calculated as the average of the likelihood values obtained from each of the 500 pseudo-experiments. The scans result in a likelihood distribution around the nominal fit result, from which the uncertainty bands can be read of. In case of a perfectly parabolic likelihood shape, they correspond to the statistical uncertainties returned by the default fitting procedure, but otherwise the uncertainty bands might differ from those.

15.4 Results

The fit to the angular distributions and reconstructed invariant mass of the candidates is performed individually in each q^2 bin. Figures 15.7 to 15.9 show one-dimensional projections of the angular distributions overlaid with the fitted model for each q^2 bin.

Table 15.3: Fit results for the parameters F_L and $S_{3,4,7}$ in bins of q^2 . All uncertainties are statistical only.

$q^2 [\text{GeV}^2/c^4]$	F_L	S_3	S_4	S_7
0.1 – 2.0	$0.20^{+0.08}_{-0.09}$	$-0.05^{+0.13}_{-0.13}$	$0.27^{+0.28}_{-0.18}$	$0.04^{+0.12}_{-0.12}$
2.0 – 5.0	$0.68^{+0.16}_{-0.13}$	$-0.06^{+0.19}_{-0.23}$	$-0.47^{+0.30}_{-0.44}$	$-0.03^{+0.18}_{-0.23}$
5.0 – 8.0	$0.54^{+0.10}_{-0.09}$	$-0.10^{+0.20}_{-0.29}$	$-0.10^{+0.15}_{-0.18}$	$0.04^{+0.16}_{-0.20}$
11.0 – 12.5	$0.29^{+0.11}_{-0.11}$	$-0.19^{+0.20}_{-0.23}$	$-0.47^{+0.21}_{-0.29}$	$0.00^{+0.15}_{-0.10}$
15.0 – 17.0	$0.23^{+0.09}_{-0.08}$	$-0.06^{+0.16}_{-0.19}$	$-0.03^{+0.15}_{-0.15}$	$0.12^{+0.16}_{-0.13}$
17.0 – 19.0	$0.40^{+0.13}_{-0.15}$	$-0.07^{+0.23}_{-0.27}$	$-0.39^{+0.25}_{-0.34}$	$0.20^{+0.29}_{-0.22}$
1.0 – 6.0	$0.63^{+0.09}_{-0.09}$	$-0.02^{+0.12}_{-0.13}$	$-0.19^{+0.14}_{-0.13}$	$-0.03^{+0.14}_{-0.14}$
15.0 – 19.0	$0.29^{+0.07}_{-0.06}$	$-0.09^{+0.11}_{-0.12}$	$-0.14^{+0.11}_{-0.11}$	$0.13^{+0.11}_{-0.11}$

Table 15.4: Fit results for the parameters $A_{5,6,8,9}$ in bins of q^2 . All uncertainties are statistical only.

$q^2 [\text{GeV}^2/c^4]$	A_5	A_6	A_8	A_9
0.1 – 2.0	$-0.02^{+0.13}_{-0.13}$	$-0.19^{+0.15}_{-0.15}$	$0.10^{+0.14}_{-0.14}$	$0.03^{+0.14}_{-0.14}$
2.0 – 5.0	$0.09^{+0.28}_{-0.22}$	$0.09^{+0.20}_{-0.19}$	$0.19^{+0.26}_{-0.21}$	$-0.13^{+0.24}_{-0.30}$
5.0 – 8.0	$0.04^{+0.17}_{-0.17}$	$-0.01^{+0.14}_{-0.12}$	$-0.12^{+0.17}_{-0.19}$	$-0.03^{+0.17}_{-0.16}$
11.0 – 12.5	$0.08^{+0.21}_{-0.21}$	$-0.16^{+0.16}_{-0.18}$	$-0.01^{+0.15}_{-0.15}$	$-0.02^{+0.16}_{-0.15}$
15.0 – 17.0	$0.02^{+0.13}_{-0.14}$	$0.01^{+0.12}_{-0.17}$	$0.08^{+0.16}_{-0.18}$	$0.21^{+0.18}_{-0.12}$
17.0 – 19.0	$0.13^{+0.29}_{-0.27}$	$-0.04^{+0.18}_{-0.19}$	$-0.16^{+0.24}_{-0.29}$	$-0.02^{+0.19}_{-0.19}$
1.0 – 6.0	$0.20^{+0.13}_{-0.13}$	$0.08^{+0.12}_{-0.11}$	$-0.00^{+0.15}_{-0.17}$	$-0.01^{+0.13}_{-0.13}$
15.0 – 19.0	$0.11^{+0.10}_{-0.10}$	$0.00^{+0.10}_{-0.11}$	$0.03^{+0.12}_{-0.12}$	$0.12^{+0.09}_{-0.11}$

The resulting values for the fit parameters F_L , $S_{3,4,7}$, and $A_{5,6,8,9}$ are listed in Tabs. 15.3 and 15.4 for each q^2 bin. The uncertainties are evaluate from profile likelihood scans using the Feldman-Cousins method as described above. Figures of the corresponding scans are shown in App. A 4. These confidence intervals spread typically asymmetrically around the best fit value, due to the proximity of some parameters to the physical boundaries in certain q^2 bins. However, while these uncertainties are purely of statistical nature, additional sources of systematic uncertainties are evaluated separately in the following chapter.

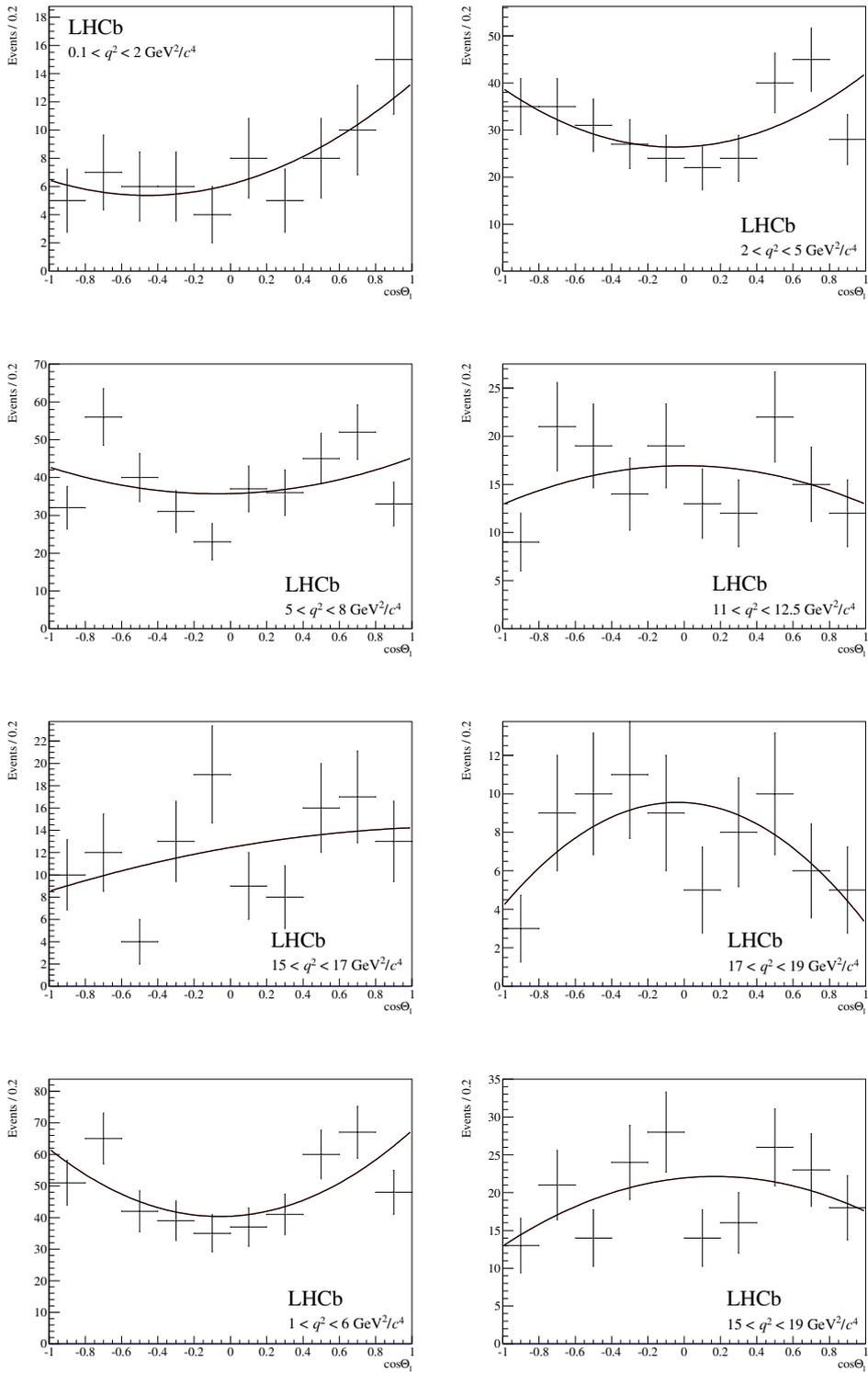


Figure 15.2: Fitted background distributions in the decay angle $\cos\theta_l$ in bins of q^2 overlaid with the respective parametrisations.

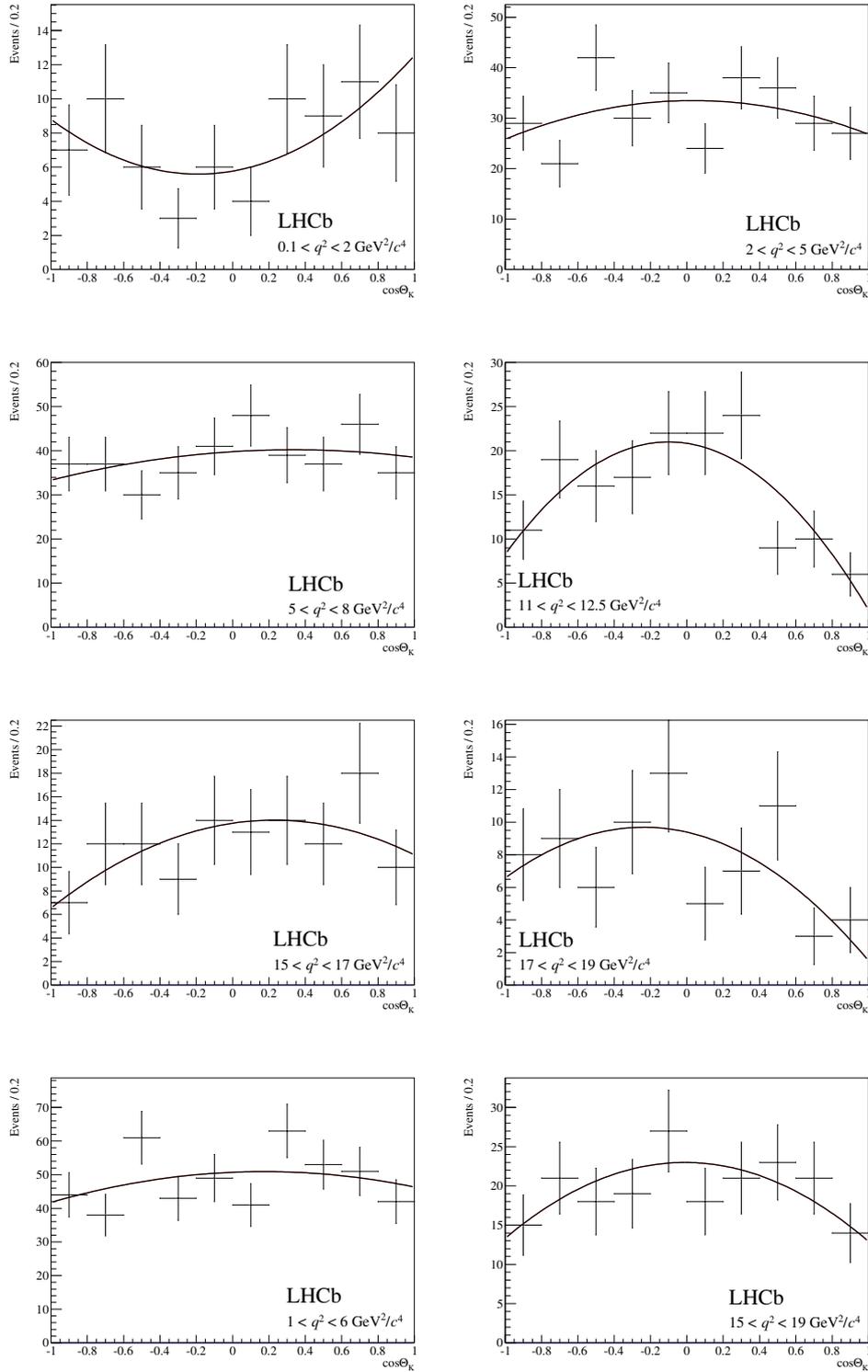


Figure 15.3: Fitted background distributions in the decay angle $\cos \theta_K$ in bins of q^2 overlaid with the respective parametrisations.

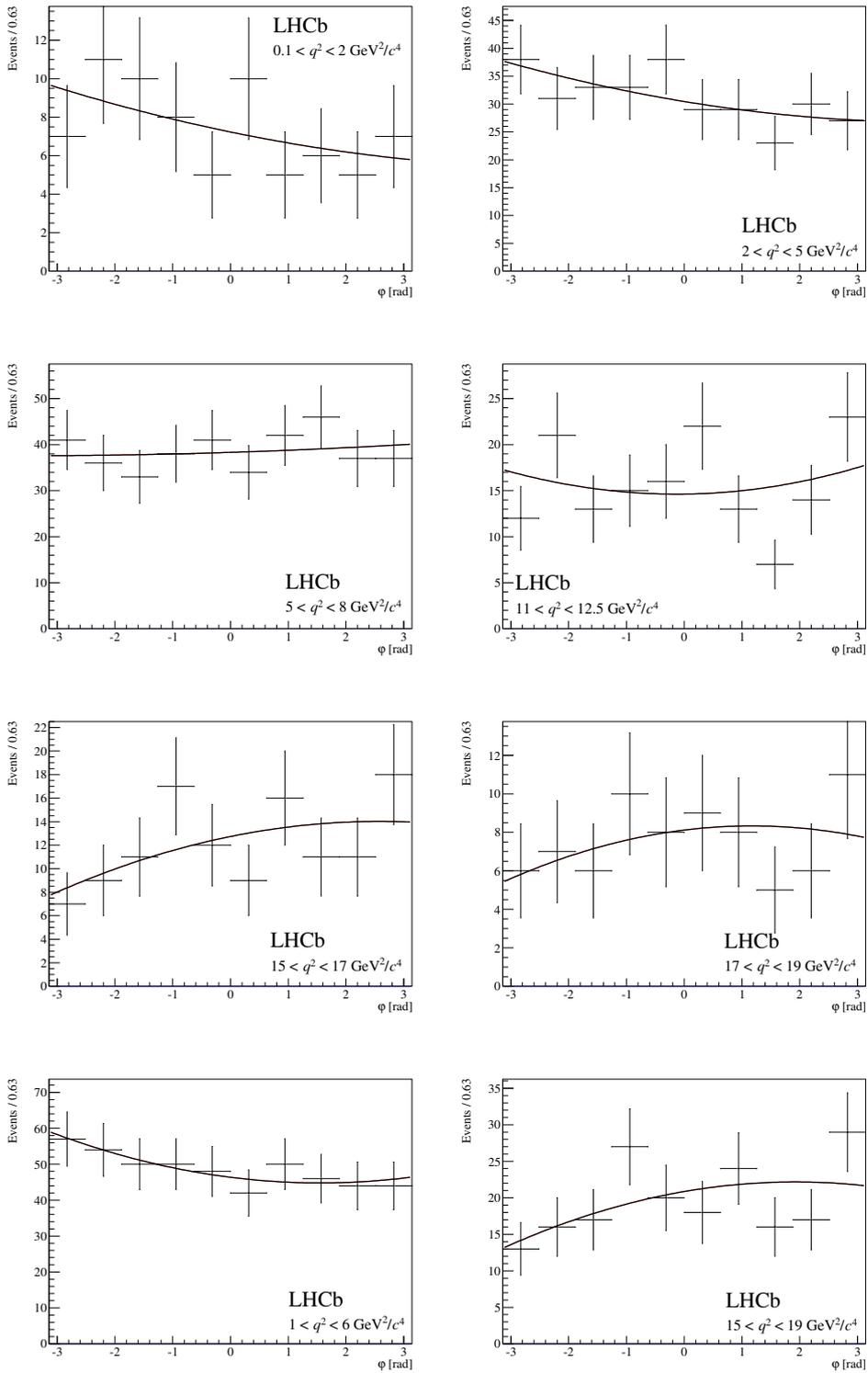


Figure 15.4: Fitted background distributions in the decay angle Φ in bins of q^2 overlaid with the respective parametrisations.

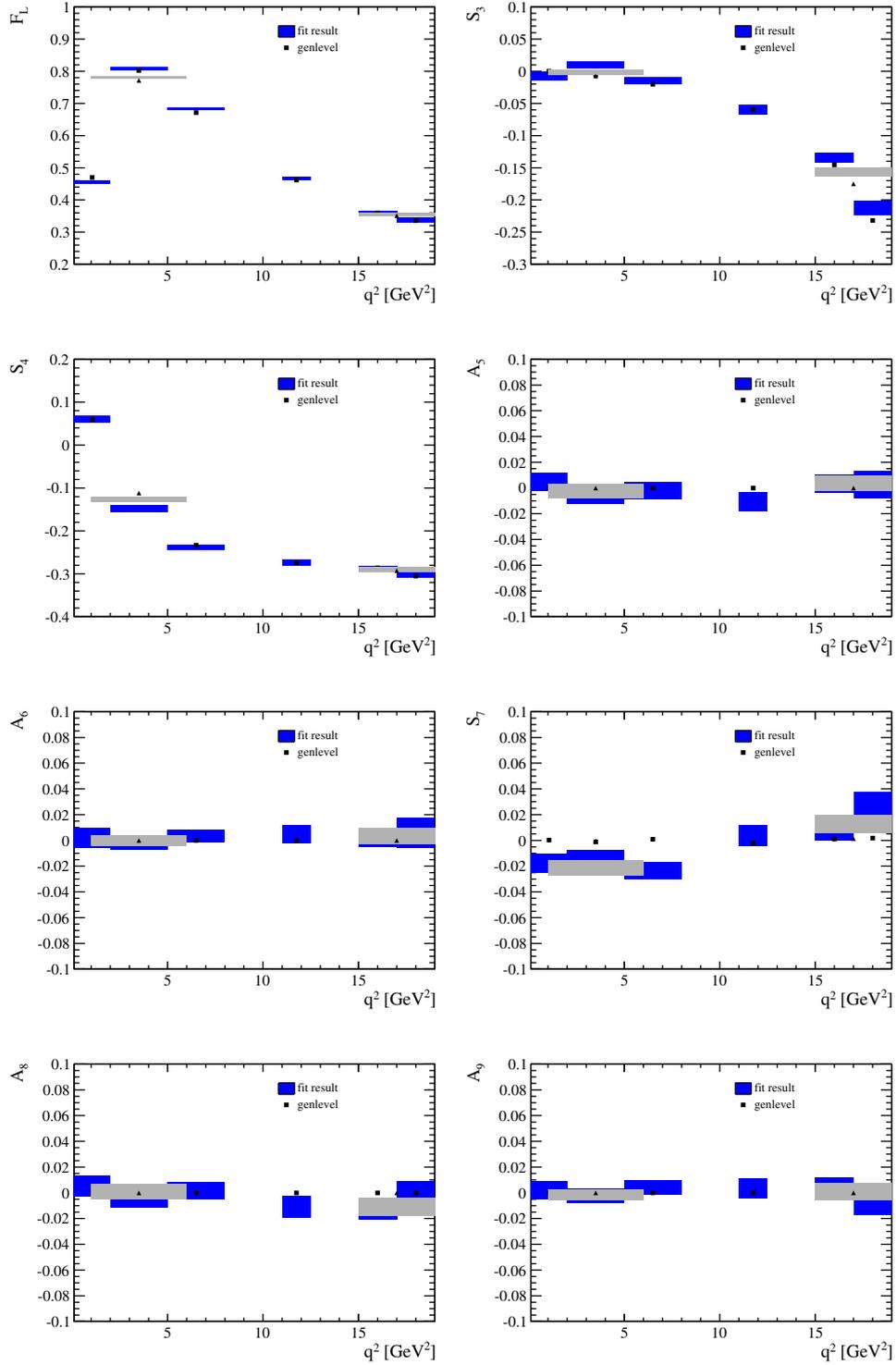


Figure 15.5: Comparison of parameters extracted from a fit to fully simulated and generated $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays as a function of q^2 . Generated values are indicated by black points and triangles, for the narrow and wide bins, respectively. Results from the fully simulated sample by blue and grey areas for the narrow and wide bins.

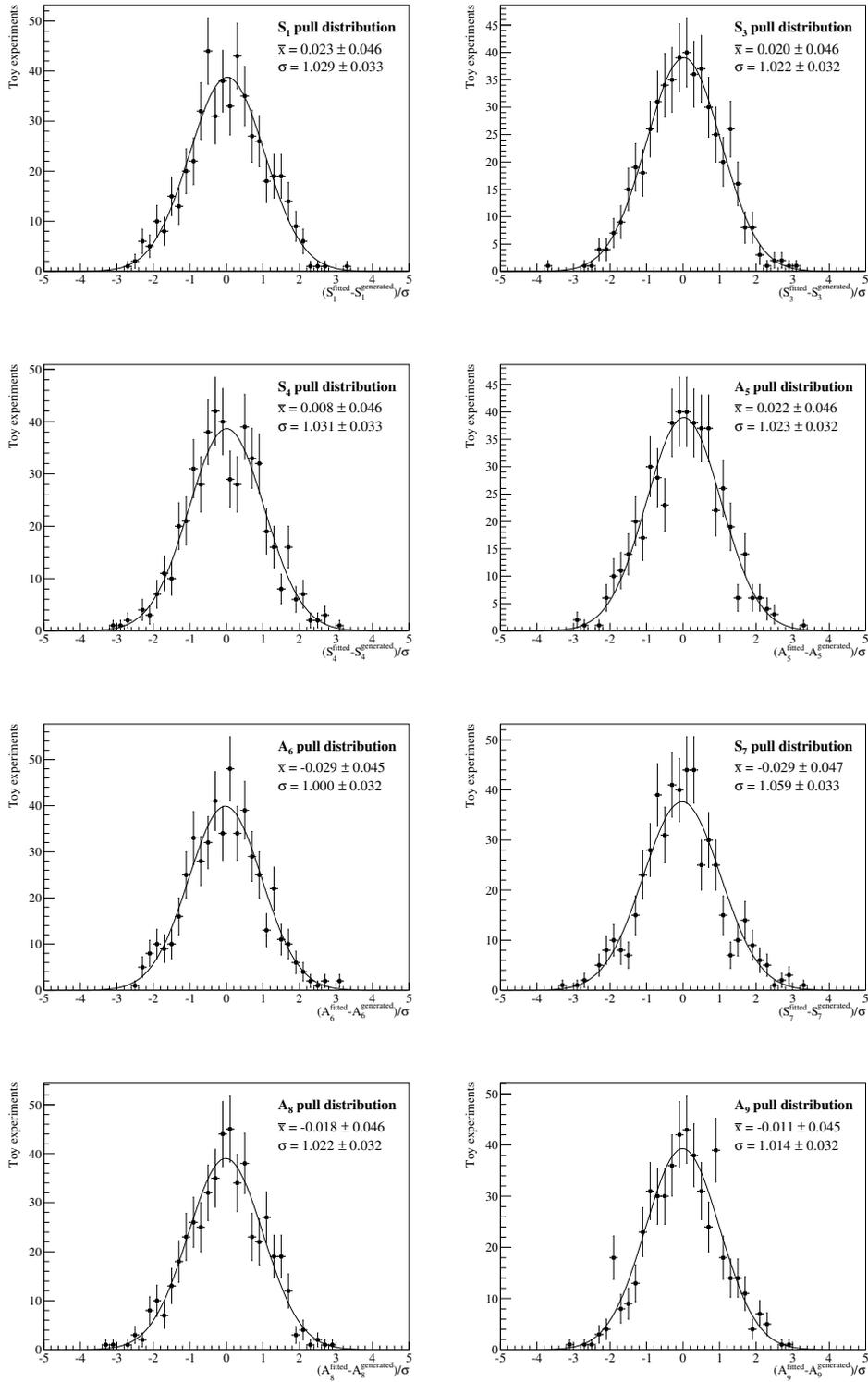


Figure 15.6: Fitted pull distributions for 500 pseudo-experiments with 50000 events each for the eight angular fit parameters. The fitted mean and widths of each distribution are shown in the individual figures.

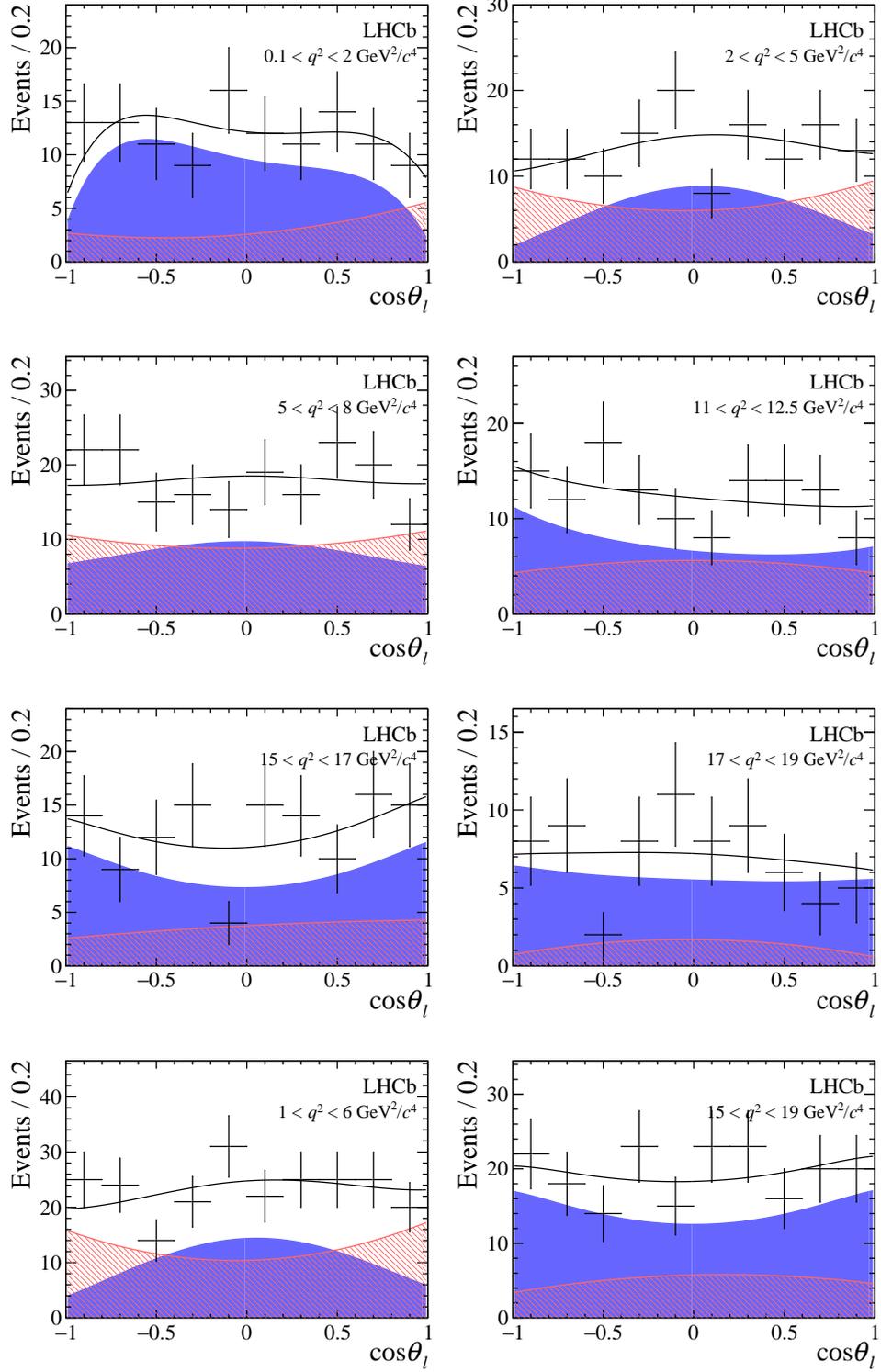


Figure 15.7: Fitted distributions of the decay angle $\cos\theta_l$ overlaid with the one-dimensional projection of the fitted model in bins of q^2 . Data points are indicated by black dots, the signal model by the blue area, the background model by the red shaded area, and the full model by the black line.

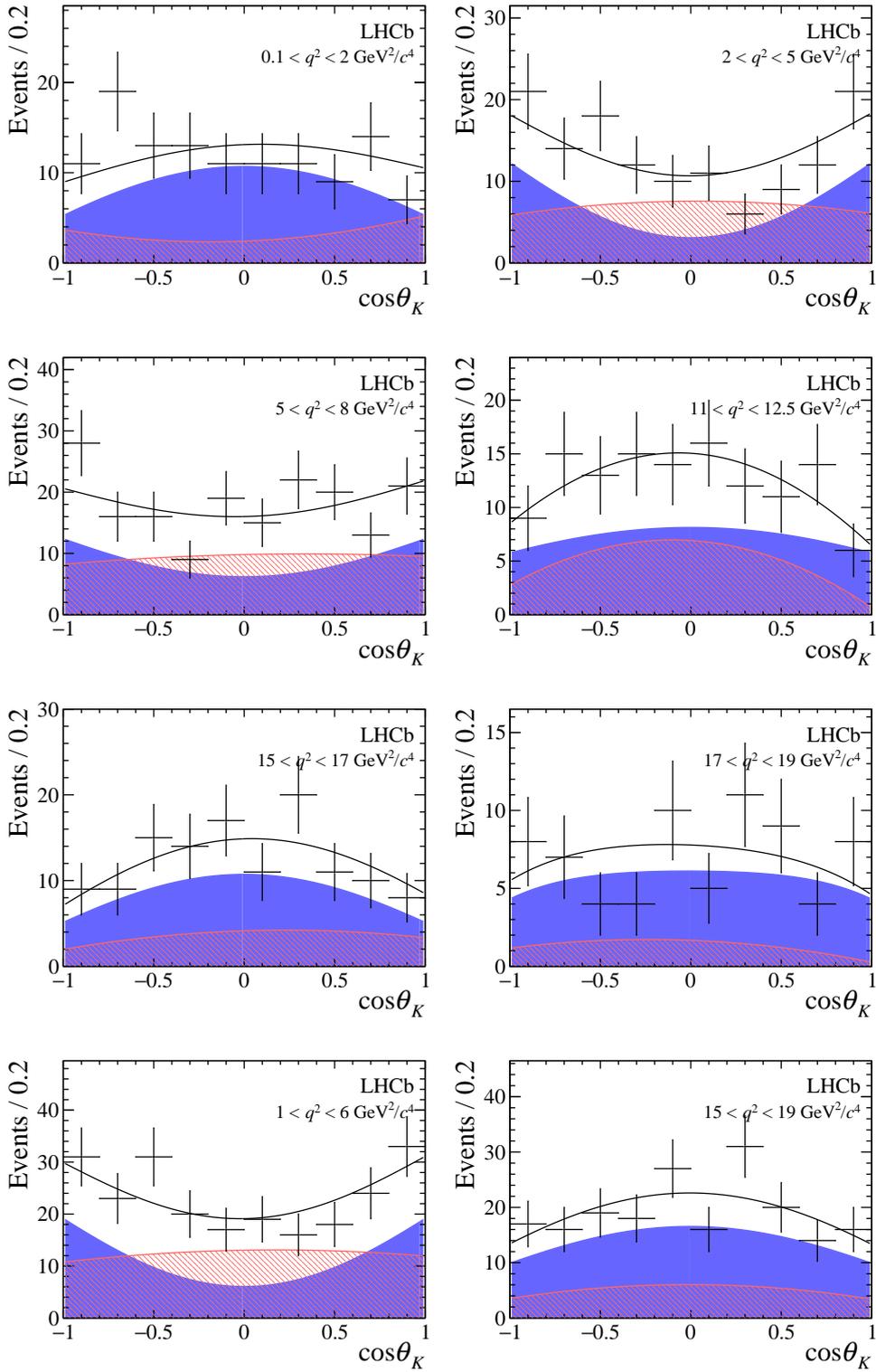


Figure 15.8: Fitted distributions of the decay angle $\cos\theta_K$ overlaid with the one-dimensional projection of the fitted model in bins of q^2 . Data points are indicated by black dots, the signal model by the blue area, the background model by the red shaded area, and the full model by the black line.

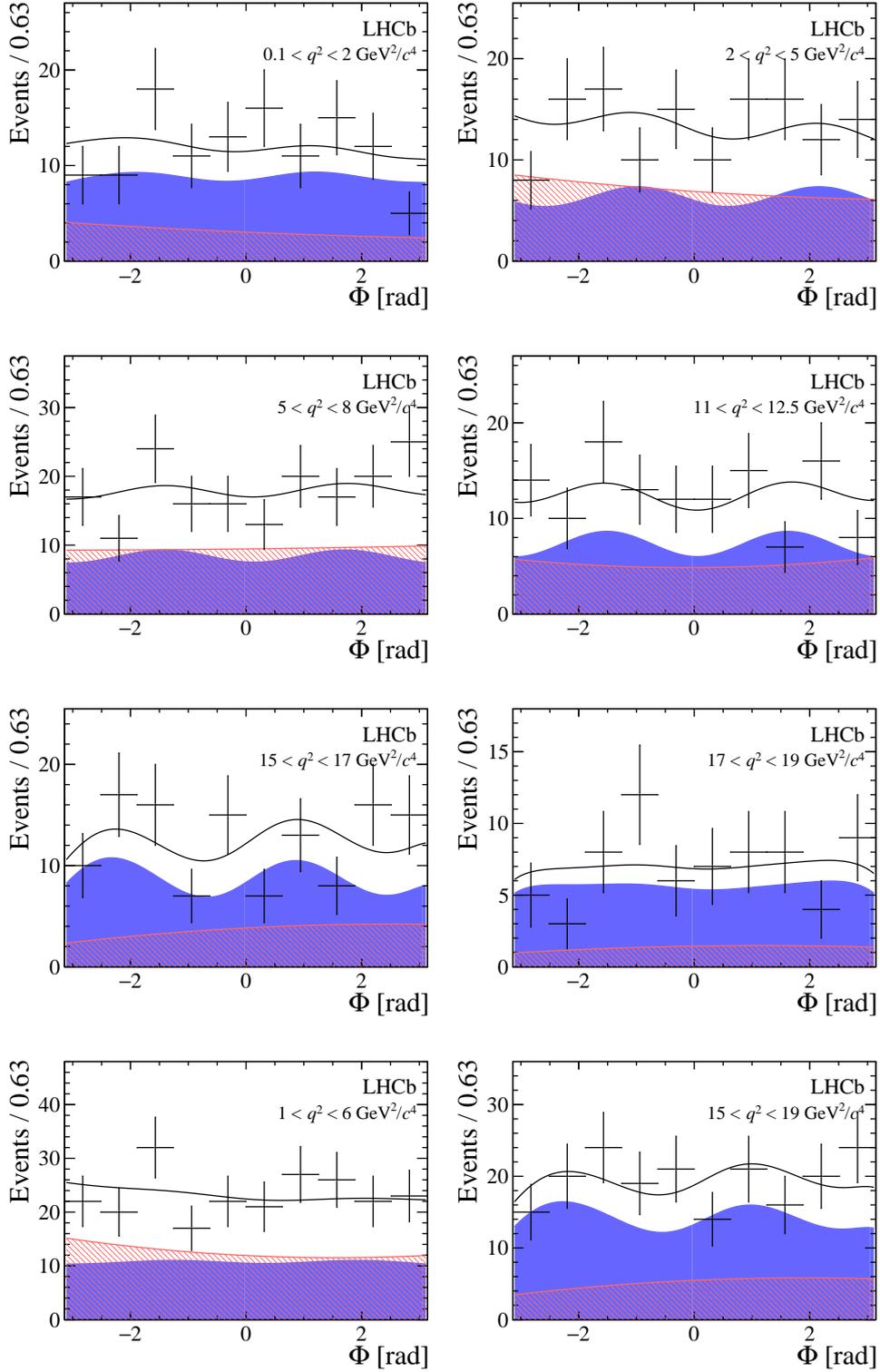


Figure 15.9: Fitted distributions of the decay angle Φ overlaid with the one-dimensional projection of the fitted model in bins of q^2 . Data points are indicated by black dots, the signal model by the blue area, the background model by the red shaded area, and the full model by the black line.

16 Determination of the systematic uncertainties

Several sources of systematic uncertainty affect the measurements of the relative differential branching fraction and the angular analysis of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$. While some of these uncertainties only play an important role for one of these analyses, most of them affect both due to the shared dataset and simulated event samples. This chapter describes the evaluation of all relevant systematic effects and summarises the final uncertainties assigned to the measurements of the relative differential branching fraction and the angular distributions. The full systematic uncertainties are compared at the end of this chapter to the statistical uncertainties derived from the fits described in the previous chapters.

For systematic studies often pseudo-experiments are used to evaluate the dependency of the fit results on different assumptions. Typically, for each of these pseudo-experiments a dataset is created that uses the parameters as extracted from the fit to the nominal dataset, varied within the statistical uncertainties, where correlations are taken into account. Effects from different models and hypothesis are tested for each pseudo-experiment and compared over the full set of them. That way statistical effects from the limited number of events in each dataset can be distinguished from the systematic uncertainty investigated.

16.1 Peaking backgrounds

As described in Chap. 11, the selection rejects most of the sources of peaking backgrounds, which originate from decays where at least one particle is misidentified and can thus mimic the signal decay. However, for two of these background decays a significant fraction may satisfy the selection requirements, which affects the measurement of the signal decay.

The rare baryonic decay $\Lambda_b^0 \rightarrow \Lambda(1520)(\rightarrow pK^-)\mu^+\mu^-$ mimics the signal decay if the proton is misidentified as a kaon. A dedicated veto is applied to reduce this background source, and the remaining pollution from this decay mode is estimated using a simulated event sample of $\Lambda_b^0 \rightarrow \Lambda(1520)(\rightarrow pK^-)\mu^+\mu^-$. As the branching fraction of this mode has not been measured yet, the measurement of the decay $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-) = (0.96 \pm 0.29) \cdot 10^{-6}$ [25] is used as an estimate, which is expected to have a similar decay rate as the investigated decay mode. The q^2 dependence is taken from a theoretical model [61] and used to distribute the expected contribution over the q^2 bins. In total, the number of background events from this decay mode passing the selection and veto is estimated to be (2.0 ± 0.8) in the signal region. The given uncertainty includes the statistical uncertainty from the simulated event sample and the uncertainty of the used branching fraction, which is reasonably large. The electroweak penguin decay $B^0 \rightarrow K^{*0}(\rightarrow K^-\pi^+)\mu^+\mu^-$ resembles the signal decay when the pion is misidentified as a kaon. The differential branching fraction as a function of q^2 is taken from the LHCb measurement [62], and the efficiencies are determined from a dedicated simulated event samples. This estimation results

Table 16.1: Changes in the relative differential branching fraction [$10^{-5} \text{ GeV}^{-2} c^4$] due to peaking backgrounds in bins of q^2 .

$q^2 [\text{GeV}^2/c^4]$	$\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$	$B^0 \rightarrow K^{*0}\mu^+\mu^-$	$B^0 \rightarrow J/\psi K^{*0}$
0.1 – 2.0	0.00	0.03	0.01
2.0 – 5.0	0.00	0.02	0.00
5.0 – 8.0	0.01	0.01	0.01
11.0 – 12.5	0.10	0.01	0.01
15.0 – 17.0	0.00	0.02	0.01
17.0 – 19.0	0.00	0.00	0.01
1.0 – 6.0	0.00	0.02	0.00
15.0 – 19.0	0.00	0.01	0.01

in an expected number of (1.7 ± 0.4) events remaining in the signal region with the full selection applied. The corresponding tree-level decay $B^0 \rightarrow J/\psi K^{*0}$ affects the yield measured for the normalisation mode in a similar way, and its contribution is estimated to (119 ± 16) events using the average of the measured branching fraction as reported by the PDG [15].

In addition, background sources from decays including $b \rightarrow c(\rightarrow s\mu^-\bar{\nu}_\mu)\mu^+\nu_\mu$ cascade transitions and purely hadronic decays such as $B_s^0 \rightarrow D_s^-(\rightarrow K^+K^-\pi^-)\pi^+$, where multiple hadrons are misidentified as muons, are investigated. They are found to be negligible in magnitude, and are not considered as a source of systematic uncertainty.

For the measurement of the relative differential branching fraction the remaining background events directly affect the measured signal and normalisation yields. As they are not incorporated in the model but peak within the signal region, they are identified as signal and normalisation mode decays, respectively. In order to evaluate their potential impact on the differential branching fraction, the q^2 -dependent background yields are inserted into Eq. 10.1 as signal and normalisation yields, respectively. For this the expected numbers of background events are distributed over the q^2 bins following the measured or expected q^2 -dependence. The resulting changes in the relative differential branching fraction are listed in Table 16.1 for each of the background sources described above. The sum of the uncertainties due to all peaking background sources are assigned as systematic uncertainty for each q^2 bin.

For the measurement of the decay-angle distributions only peaking backgrounds mimicking the signal mode are relevant as source of systematic uncertainty. In order to estimate their effects, a set of 500 pseudo-experiments is performed. In each of these pseudo-experiments two pseudo-datasets are created, one according to the nominal fit model without any peaking background events, and one dataset where background events are injected with an angular and q^2 distribution obtained from

Table 16.2: Systematic uncertainties in the angular fit parameters due to peaking backgrounds in bins of q^2 .

q^2 [GeV ² /c ⁴]	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.000	0.000	0.001	0.000	0.001	0.003	0.001	0.002
2.0 – 5.0	0.001	0.000	0.002	0.003	0.002	0.001	0.000	0.003
5.0 – 8.0	0.001	0.002	0.004	0.003	0.001	0.003	0.004	0.000
11.0 – 12.5	0.000	0.001	0.004	0.003	0.007	0.002	0.005	0.004
15.0 – 17.0	0.002	0.005	0.005	0.003	0.004	0.005	0.003	0.002
17.0 – 19.0	0.001	0.000	0.001	0.002	0.000	0.003	0.001	0.001
1.0 – 6.0	0.001	0.001	0.000	0.002	0.001	0.002	0.001	0.002
15.0 – 19.0	0.001	0.003	0.003	0.002	0.002	0.004	0.002	0.001

simulated event samples. Fits are performed to both datasets with the nominal fit model, and the difference between the extracted fit parameters from these compared. The mean difference between the fit to the nominal and background-injected datasets over the 500 pseudo-experiments is assigned as systematic uncertainty in each q^2 bin and fit parameter. The respective values are listed in Table 16.2.

16.2 Correction of simulated events

The corrections applied to the simulated event samples as described in Chap. 12 are affected by uncertainties for each correction step. These uncertainties are mostly tied to the data-driven methods used to determine the respective efficiency.

For the measurement of the relative differential branching fraction the uncertainties for the correction steps are propagated into the relative efficiencies. Their effect can be evaluated by the change in the differential branching fraction due to altered efficiencies.

The corrections to the track reconstruction efficiencies are measured using $J/\psi \rightarrow \mu^+\mu^-$ decays, measuring the efficiencies for final-state muons. The respective correction tables come with certain statistical and systematic uncertainties, depending on the track kinematics. In addition, final-state hadrons interact more strongly with the detector material than muons, which induces an additional systematic uncertainty. The overall systematic uncertainty per final-state track is 0.4% for muons, 1.17% for kaons, and 1.46% for pions. For the measurement of the relative differential branching fraction this uncertainty is propagated into the relative efficiency between signal and normalisation mode, and, as they have an identical set of final-state particles, the systematic part largely cancels. The statistical uncertainty of the tracking efficiency correction is evaluated by creating a set of ten pseudo correction tables with values randomly varied within the statistical uncertainties in each bin. This creation and application of pseudo correction tables is computationally demanding as it is performed

using the fully simulated event sample, which is why the process is only repeated ten times. However, the uncertainty of the correction is expected to affect the signal and normalisation decays in a similar way, and thus the systematic uncertainty caused by it is expected to be small compared to the statistical precision of the measurement. Each of the pseudo correction tables is used to correct the simulated event samples for signal and normalisation mode. Subsequently, the relative efficiency is re-evaluated and used to calculate the relative differential branching fraction. The variance of the results from these ten pseudo-experiments is assigned as systematic uncertainty.

The `IsMuon` flag only affects the final-state muons and thus no systematic uncertainty for hadronic interactions has to be taken into account for the `IsMuon` efficiency correction. However, the correction table is subject to statistical uncertainties. The uncertainties of these corrections are expected to be correlated for tracks with similar momentum, as the number of hits left in the muon stations required for the `IsMuon` flag is higher for tracks with a momentum of $> 10 \text{ GeV}/c$ than for lower momenta. Thus two pseudo-tables are created, maximising the effect of this threshold as a conservative estimate of the systematic uncertainty. For the first table, the entries with $p < 10 \text{ GeV}/c$ are reduced by the statistical uncertainty, while for $p > 10 \text{ GeV}/c$ they are increased. This shift is inverted for the second table. The relative efficiencies are calculated for both these pseudo-tables and the relative branching fraction ratio evaluated. The largest deviation from the nominal result is assigned as systematic uncertainty.

The particle identification variables are created using measured distributions derived from cleanly selected datasets, which are evaluated in bins of kinematic variables. Different distributions of the track kinematics within these bins of the corrected simulated event samples compared to the calibration samples might bias the correction. An infinitely fine binning scheme would remove any potential bias introduced, however, the default binning scheme is already optimised to the statistics available in the calibration samples. In order to derive a conservative estimation for the systematic uncertainties of the recreation of the PID variables, the number of bins in momentum p is cut in half and the recreation repeated. These alternatively created PID variable distributions are used in the BDT classification and selection of the simulated event sample and the selection efficiency re-evaluated. The resulting change in the relative efficiency is propagated to the branching fraction calculation, and the difference to the nominal result is assigned as systematic uncertainty.

The application of weights to the simulated event samples is used to align the distributions of the transversal momentum p_T and vertex χ^2 of the B_s^0 meson. This removes remaining differences observed in these discriminating variables of the multivariate classifier. As there is no inherent uncertainty tied to this step a conservative approach is chosen to estimate an uncertainty to this weighting step. The corrections are completely removed and the relative efficiencies and resulting relative branching fraction ratio calculated. The systematic uncertainty tied to this correction is the difference to the nominal result.

The uncertainties of all correction steps applied to the simulated event samples are

Table 16.3: Changes in the relative differential branching fraction [$10^{-5} \text{ GeV}^{-2} c^4$] due to the corrections to the simulated event samples in bins of q^2 .

$q^2 [\text{GeV}^2/c^4]$	Tracking eff.	IsMuon eff.	Weighting
0.1 – 2.0	0.01	0.01	0.00
2.0 – 5.0	0.00	0.00	0.01
5.0 – 8.0	0.00	0.00	0.00
11.0 – 12.5	0.00	0.00	0.00
15.0 – 17.0	0.02	0.00	0.04
17.0 – 19.0	0.02	0.00	0.03
1.0 – 6.0	0.00	0.00	0.00
15.0 – 19.0	0.02	0.00	0.04

 Table 16.4: Systematic uncertainties in the angular fit parameters due to the corrections to the simulated event samples in bins of q^2 .

$q^2 [\text{GeV}^2/c^4]$	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.004	0.002	0.000	0.000	0.001	0.000	0.000	0.000
2.0 – 5.0	0.003	0.002	0.001	0.001	0.002	0.001	0.001	0.001
5.0 – 8.0	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.001
11.0 – 12.5	0.004	0.002	0.001	0.001	0.001	0.001	0.001	0.001
15.0 – 17.0	0.006	0.006	0.001	0.001	0.001	0.001	0.001	0.001
17.0 – 19.0	0.011	0.004	0.004	0.003	0.002	0.003	0.003	0.003
1.0 – 6.0	0.002	0.001	0.001	0.000	0.001	0.000	0.001	0.000
15.0 – 19.0	0.007	0.005	0.001	0.001	0.001	0.001	0.001	0.001

summed quadratically to obtain an overall systematic uncertainty for the correction. The individual uncertainties are listed in Table 16.3.

All of the above correction steps do not only affect the evaluation of the relative efficiencies, but also of the angular acceptance, which is determined from the same corrected simulated event samples. The same studies as described above are used to re-evaluate the angular acceptance, and compare the fit result using this modified acceptance to the nominal fit.

In order to separate this from statistical effects, a set of 500 pseudo-experiments is performed, for which a dataset is created with the nominal acceptance and fitted with the systematically modified acceptance. The mean shift in each fit parameter and q^2 bin is assigned as systematic uncertainty due to the correction steps, listed in Table 16.4.

Table 16.5: Systematic uncertainties in the angular fit parameters due to the differences in final-state p and p_T in bins of q^2 .

q^2 [GeV ² /c ⁴]	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.007	0.001	0.000	0.000	0.001	0.000	0.000	0.000
2.0 – 5.0	0.003	0.000	0.001	0.000	0.000	0.000	0.000	0.000
5.0 – 8.0	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
11.0 – 12.5	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
15.0 – 17.0	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
17.0 – 19.0	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000
1.0 – 6.0	0.003	0.000	0.001	0.000	0.000	0.000	0.000	0.000
15.0 – 19.0	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000

16.3 Residual differences in kinematic variables

While most discrepancies between data and simulated event samples are strongly reduced by the corrections described in Chap. 12, there are small residual differences remaining in the (transversal) momenta of the final-state tracks as observed for $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ candidates. Figure 16.1 shows these normalised distributions for the μ^+ and K^- of the background-subtracted $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data and simulated event samples. The same discrepancies are observed for the μ^- and K^+ final-state tracks. These variables are not used as part of the event selection and the remaining discrepancies have no measurable effect for the measurement of the differential branching fraction. However, they enter the calculation of the decay angles and thus the deviations of the simulated distributions compared to data might bias the determination of the angular acceptance. From the distributions measured for the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ decays two-dimensional weights are determined in p and p_T to describe the remaining discrepancies as a function of these two variables. Due to the low statistics of the selected $B_s^0 \rightarrow \phi\mu^+\mu^-$ candidate sample, there is no reasonable way to determine these weights directly from the signal sample. Given the identical set of final-state particles originating from a B_s^0 -meson decay, it is a reasonable assumption that the deviations appear in a similar magnitude, at least compared to the statistical precision of the measurement. Those weights are used to correct the $B_s^0 \rightarrow \phi\mu^+\mu^-$ simulated event samples and the angular acceptance is recalculated. A set of 500 pseudo-datasets is created using the nominal acceptance and each dataset fitted once with the nominal and once with the corrected acceptance. The mean differences between the fit result for each parameter and q^2 bin are assigned as systematic uncertainty, as listed in Table 16.5.

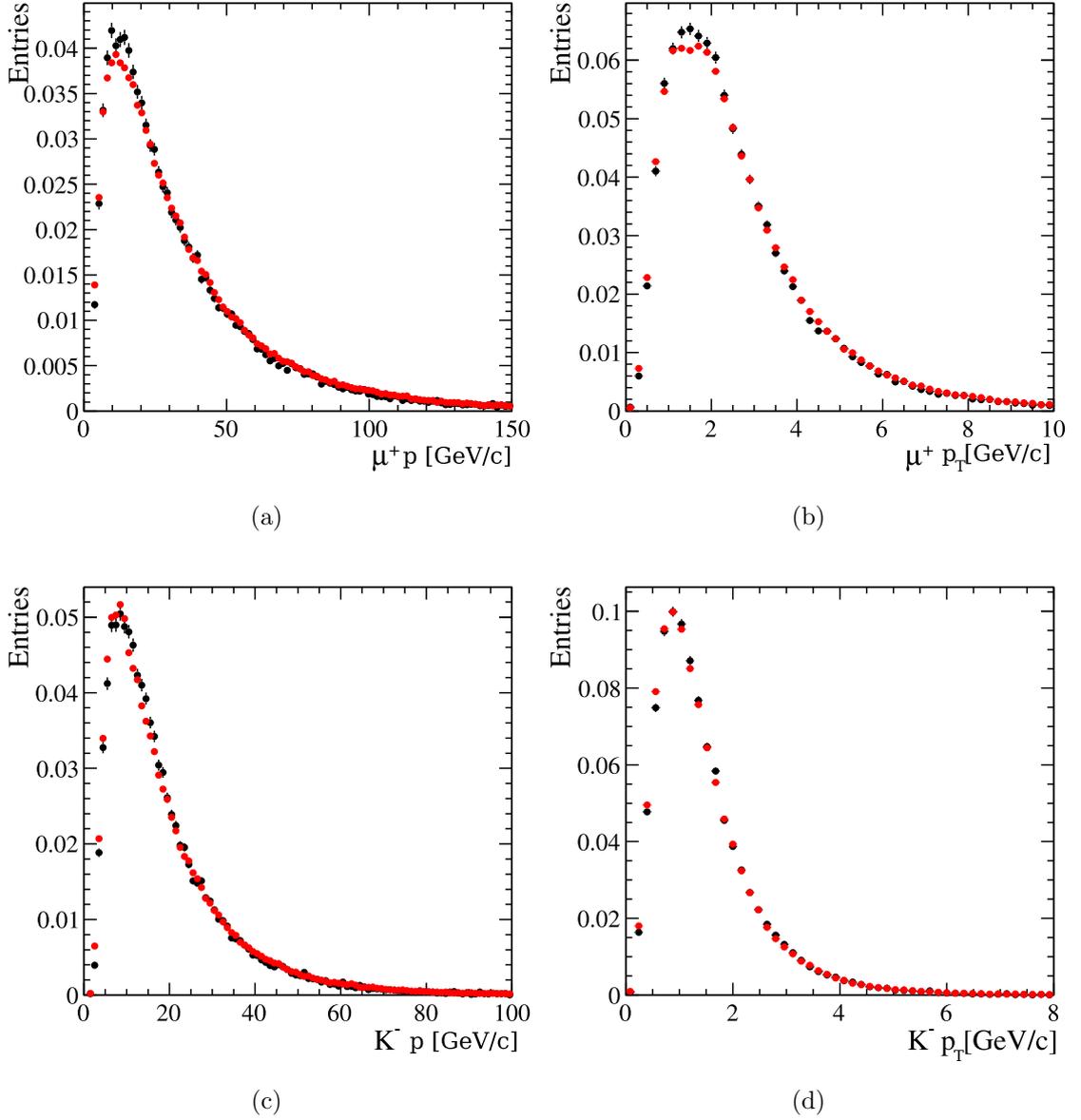


Figure 16.1: Comparison between the $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ data (black) and simulated event sample (red) for the variables: μ^+ momentum (a), μ^+ transversal momentum (b), K^- momentum (c), K^- transversal momentum (d).

16.4 Model dependence of the detector acceptance

The detector acceptance is determined by investigating the fraction of signal events where the final-state particle trajectories are pointing towards the acceptance of the LHCb detector. As this has to be done using a simulated event sample, this

acceptance is dependent on the decay model used from which the angular distributions originate. In order to evaluate the magnitude of this dependence on the decay model used, its parameters are modified and a separate simulated event sample generated. The nominal decay model uses the q^2 -dependent Wilson coefficients as predicted from the Standard Model, where the coefficients which have the largest effect on the signal decay $B_s^0 \rightarrow \phi\mu^+\mu^-$, C_7 and C_9 , are modified for this systematic study. In Ref. [63] multiple New Physics scenarios are investigated in the light of recent measurements of electroweak penguin decays. The most strongly supported scenario proposes shifts of $C_7 \rightarrow C_7 - 0.02$ and $C_9 \rightarrow C_9 - 1.5$ due to New Physics contributions. These shifts are applied to the decay model used in the generation of a simulated event sample, and the detector acceptance with this alternative model determined. The resulting shift in the detector acceptance leads to modified relative efficiencies, which are propagated to the calculation of the relative differential branching fraction. The difference to the nominal result is assigned as systematic uncertainty.

16.5 Branching fraction of $J/\psi \rightarrow \mu^+\mu^-$

The branching fraction of the decay $J/\psi \rightarrow \mu^+\mu^-$ is used in calculating the relative differential branching fraction via Eq. 10.1. Its value is taken from the measured world average value of $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-) = (5.961 \pm 0.033)\%$ [15]. The effect of the affiliated uncertainty $\sigma_{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}$ is propagated directly to the differential branching fraction via

$$\sigma_{\text{rel.diff.}\mathcal{B}} = \frac{1}{\mathcal{B}(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi)} \cdot \frac{d\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{dq^2} \cdot \frac{\sigma_{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}}{\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}, \quad (16.1)$$

leads to a relative uncertainty of 0.55% to be assigned to the branching fraction measurement, depending on the value of the differential branching fraction in the given q^2 bin.

16.6 Limited size of simulated event samples

The limited number of simulated events for $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ decays directly effect the statistical uncertainties of the determined relative efficiencies. These differ for each q^2 bin and are listed in Chap. 13. The relative efficiencies enter Eq. 10.1 in first order and thus their statistical uncertainties are propagated via

$$\sigma_{\text{rel.diff.}\mathcal{B}} = \frac{1}{\mathcal{B}(B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi)} \cdot \frac{d\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{dq^2} \cdot \frac{\sigma_{\epsilon_{\text{tot}}^{J/\psi\phi}/\epsilon_{\text{tot}}^{\phi\mu^+\mu^-}}}{\epsilon_{\text{tot}}^{J/\psi\phi}/\epsilon_{\text{tot}}^{\phi\mu^+\mu^-}}. \quad (16.2)$$

This results in relative uncertainties between 0.9% and 1.9%, depending on the q^2 bin.

The limited number of simulated events directly affects the precision in the derivation of the angular acceptance, which is determined using those events. This

Table 16.6: Systematic uncertainties in the angular fit parameters due to the statistical uncertainty of the angular acceptance parametrisation in bins of q^2 .

q^2 [GeV ² /c ⁴]	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.000	0.000	0.001	0.000	0.001	0.003	0.001	0.002
2.0 – 5.0	0.001	0.000	0.002	0.003	0.002	0.001	0.000	0.003
5.0 – 8.0	0.001	0.002	0.004	0.003	0.001	0.003	0.004	0.000
11.0 – 12.5	0.000	0.001	0.004	0.003	0.007	0.002	0.005	0.004
15.0 – 17.0	0.002	0.005	0.005	0.003	0.004	0.005	0.003	0.002
17.0 – 19.0	0.001	0.000	0.001	0.002	0.000	0.003	0.001	0.001
1.0 – 6.0	0.001	0.001	0.000	0.002	0.001	0.002	0.001	0.002
15.0 – 19.0	0.001	0.003	0.003	0.002	0.002	0.004	0.002	0.001

is reflected by the statistical uncertainty in determining the coefficients of the Legendre polynomials used for the parametrisation of the acceptance. To evaluate the systematic effect of this uncertainty a set of 500 pseudo-datasets is created using the nominal acceptance description. The fit to each of these datasets is performed twice, once with the nominal acceptance and once with the acceptance with parameters randomly fluctuated within their respective uncertainties, taking correlations into account. The width of the distribution of the difference between fits with nominal and modified acceptance parametrisation are assigned as systematic uncertainty for each extracted fit parameter. They are listed in Table 16.6.

16.7 Signal model

For the extraction of the signal yields for the measurement of the relative differential branching fraction the shape of the reconstructed invariant mass distribution for signal candidates is described by the sum of two Crystal Ball functions. A simpler model like the sum of two Gaussian distributions also describes the observed distributions reasonably well, however not as well as the nominal model. To evaluate the dependence on the chosen invariant mass model 500 pseudo-experiments are performed, where the pseudo-datasets are created according to the sum of two Gaussian distributions. The fit to each dataset is performed twice, once with the nominal and once with modified model, and the mean shift in the calculated differential branching fractions is assigned as systematic uncertainty.

As the signal model used in the analysis of the angular distributions is motivated completely by the theoretical description, all systematic effects are tied to the angular acceptance, and no explicit systematic uncertainty is caused by the model choice.

16.8 Background model and parametrisation

An exponential function is used to describe combinatorial background in the reconstructed invariant mass distribution. This is motivated by the typical shape observed in many analyses. However, as the background levels are very low due to the applied selection, its distribution can also be reasonably well described by a first-order polynomial. In order to evaluate the systematic uncertainty in the determination of the differential branching fraction the same procedure is applied as for the signal model dependence: 500 pseudo-datasets are created using the modified background model. The fit to each dataset is performed twice, once using the nominal and once using the modified background model. The mean difference in the resulting differential branching fraction is assigned as systematic uncertainty.

For the analysis of the angular distributions the background model is determined from a background-enriched data sample of events in the upper sideband of the reconstructed invariant mass distribution. This parametrisation is both affected by a statistical uncertainty in determining the coefficients of the Chebyshev polynomials as well as the choice to use second-order polynomials.

The effect of the statistical uncertainty of the nominal coefficients used for the background parametrisation is determined by performing a set of 1000 pseudo-experiments, where the dataset is created using the nominal values. As the precision of the background parametrisation is expected to be a dominant source of systematic uncertainty, a larger number of pseudo-experiments is performed in this study compared to other sources of uncertainty investigated in this chapter. Each dataset is then fitted twice, once with the nominal coefficients and once with values randomly varied within uncertainties, taking correlations between the coefficients into account. The width of the distribution of the difference between the fit parameters extracted from those two fits for all pseudo-experiments is assigned as systematic uncertainty, listed in Table 16.7.

In order to investigate the uncertainty imposed by the choice of second-order Chebyshev polynomials a similar approach is chosen. An alternative parametrisation is determined using only first-order Chebyshev polynomials, and a set of 500 pseudo-datasets created with the nominal parametrisation. Each dataset is subsequently fitted using both nominal and alternative parametrisations and the mean shift between those two fits is assigned as systematic uncertainty for each fit parameter. Table 16.8 lists the corresponding values.

16.9 Decay time acceptance

As the final state of the signal decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ is symmetrical under CP conjugation, decays originating from a \bar{B}_s^0 meson cannot be distinguished from B_s^0 -meson decays. Thus they are measured simultaneously in this analysis. However, the measurement of the differential branching fraction without taking the decay-time

Table 16.7: Systematic uncertainties in the angular fit parameters due to the statistical uncertainty of the background parametrisation in bins of q^2 .

$q^2[\text{GeV}^2/c^4]$	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.008	0.004	0.005	0.003	0.011	0.003	0.004	0.004
2.0 – 5.0	0.010	0.007	0.008	0.006	0.015	0.005	0.007	0.008
5.0 – 8.0	0.008	0.005	0.005	0.004	0.010	0.003	0.004	0.005
11.0 – 12.5	0.008	0.006	0.008	0.006	0.013	0.004	0.005	0.007
15.0 – 17.0	0.007	0.005	0.007	0.005	0.010	0.003	0.005	0.005
17.0 – 19.0	0.009	0.013	0.016	0.011	0.011	0.009	0.009	0.010
1.0 – 6.0	0.007	0.004	0.004	0.003	0.010	0.002	0.003	0.004
15.0 – 19.0	0.005	0.003	0.003	0.002	0.006	0.002	0.002	0.003

 Table 16.8: Systematic uncertainties in the angular fit parameters due to the choice of background model in bins of q^2 .

$q^2[\text{GeV}^2/c^4]$	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.013	0.000	0.001	0.000	0.002	0.000	0.000	0.001
2.0 – 5.0	0.030	0.001	0.001	0.000	0.003	0.000	0.000	0.001
5.0 – 8.0	0.013	0.000	0.000	0.000	0.001	0.000	0.000	0.000
11.0 – 12.5	0.044	0.007	0.006	0.001	0.002	0.000	0.001	0.001
15.0 – 17.0	0.013	0.001	0.002	0.000	0.000	0.000	0.000	0.000
17.0 – 19.0	0.010	0.004	0.004	0.000	0.002	0.000	0.000	0.001
1.0 – 6.0	0.024	0.003	0.002	0.000	0.002	0.000	0.000	0.000
15.0 – 19.0	0.011	0.001	0.003	0.000	0.000	0.000	0.000	0.001

difference $\Delta\Gamma_s$ between the light and heavy B_s^0 -meson mass eigenstates into account can be affected due to decay-time dependent efficiencies [64]. The LHCb trigger requirement of the final-state tracks to be significantly detached from the primary interaction vertex leads to a higher efficiency for particles with longer decay times. For that reason, the original composition of the B_s^0 -meson system into light and heavy mass eigenstates can affect the efficiencies of these selection steps on a sample consisting of an admixture of these states.

Unfortunately, the simulated event samples used in this analysis for the decay modes $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ are created with $\Delta\Gamma_s = 0\text{ ps}^{-1}$ and $\Delta\Gamma_s = 0.092\text{ ps}^{-1}$, respectively. The large difference of the former value from the currently measured value of $\Delta\Gamma_s = (0.086 \pm 0.006)\text{ ps}^{-1}$ [15] could bias the relative efficiencies used in the measurement of the differential branching fraction. To investigate this effect, multiple samples of simulated $B_s^0 \rightarrow \phi\mu^+\mu^-$ events are created, where the proper decay-time difference and following from this decay-time dependent

Table 16.9: Systematic uncertainties in the angular fit parameters due to the q^2 dependence of the angular acceptance in bins of q^2 .

q^2 [GeV ² /c ⁴]	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.0 – 5.0	0.003	0.000	0.001	0.000	0.000	0.000	0.000	0.000
5.0 – 8.0	0.002	0.002	0.001	0.000	0.000	0.000	0.000	0.000
11.0 – 12.5	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000
15.0 – 17.0	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000
17.0 – 19.0	0.003	0.002	0.000	0.000	0.000	0.000	0.000	0.000
1.0 – 6.0	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
15.0 – 19.0	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000

decay rates are implemented, as described in Refs. [65] and [41]. In these samples the decay time acceptance determined from the fully simulated $B_s^0 \rightarrow \phi\mu^+\mu^-$ decay sample is used. The fraction of CP -even and CP -odd eigenstates (and thus light and heavy mass eigenstates) is modified by changing the Wilson coefficients by $\pm 50\%$ (as a conservative approach), which are directly related to the composition. For each variation the relative difference in events passing the selection requirements compared to the nominal simulated event sample is determined, in the range of 1.0% to 1.6%. The largest deviation of 1.6% is assigned as a relative systematic uncertainty for differential branching fraction measurement.

16.10 Dependence of angular acceptance on q^2

As described in Sec. 13.5, the angular acceptance used for the analysis of the angular distributions is dependent on q^2 and evaluated for each q^2 bin at the central value of the respective bin. This ignores the potential variations of the angular acceptance within the q^2 bins. As an alternative approach, a three-dimensional acceptance in the three decay angle is evaluated for each bin, integrated over the respective q^2 range. A set of 500 pseudo-experiments is performed, where each dataset is created using the nominal acceptance model, and subsequently fitted with the nominal and the alternative model. The mean shift between the results from these fits over the full set of pseudo-experiments is assigned as systematic uncertainty. The respective values are listed in Table 16.9.

16.11 S-wave contribution in the K^+K^- system

The ϕ meson of the $B_s^0 \rightarrow \phi\mu^+\mu^-$ decay is reconstructed in the decay $\phi \rightarrow K^+K^-$, corresponding to the resonant P-wave (spin-1) state of the K^+K^- system. However, the K^+K^- system can also be produced in the non-resonant S-wave state, a spin-0 system. These states can bias the measured signal yields extracted from the fit to the

Table 16.10: Systematic uncertainties in the angular fit parameters due to the S-wave contribution in bins of q^2 .

q^2 [GeV ² /c ⁴]	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.002	0.000	0.004	0.001	0.001	0.001	0.000	0.000
2.0 – 5.0	0.010	0.002	0.003	0.001	0.000	0.002	0.001	0.000
5.0 – 8.0	0.005	0.000	0.003	0.001	0.000	0.002	0.001	0.000
11.0 – 12.5	0.001	0.002	0.004	0.001	0.000	0.000	0.002	0.000
15.0 – 17.0	0.004	0.007	0.005	0.002	0.001	0.001	0.001	0.000
17.0 – 19.0	0.002	0.008	0.005	0.001	0.001	0.000	0.000	0.000
1.0 – 6.0	0.008	0.000	0.002	0.002	0.001	0.001	0.001	0.001
15.0 – 19.0	0.002	0.006	0.006	0.001	0.001	0.000	0.001	0.001

reconstructed invariant mass distribution. For the decay mode $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ the fraction of S-wave states in a ± 12 MeV/c² window around the nominal ϕ meson mass was found to be $(2.2 \pm 1.2 \pm 0.7)\%$ [66]. Within the statistical precision of the measurement it is reasonable to assume that these S-wave states affect $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decays in the same manner. For that reason, the S-wave contribution should largely cancel in the measurement of the relative differential branching fraction. For the analysis of the angular distributions the result might be slightly biased, as the angular distributions for these non-resonant decays are different compared to the signal decay. To investigate the potential size of this effect, a set of 500 pseudo-datasets are created where a 2% fraction of S-wave states is added, where the angular distributions of those states are taken from the LHCb measurement of $B_s^0 \rightarrow J/\psi (\rightarrow \mu^+ \mu^-) \phi$ decays. Each dataset is then fitted with the nominal model and the mean shift of the fit parameters is assigned as systematic uncertainty, listed in Table 16.10.

16.12 Summary

The systematic uncertainties as described above for the measurement of the relative differential branching fraction are summarised in Table 16.11. The overall systematic uncertainty is the quadratic sum of the individual sources. While the individual sources of systematic uncertainty have varying impact depending on the q^2 bin, the most dominant source is the decay time dependence of the selection efficiency. This shows the importance of the proper implementation of the decay-time dependence of the decay rate in the simulation for future analyses. However, due to the small size of the data sample, the statistical uncertainties shown in Chap. 14 are clearly dominant in this measurement. For analyses using the larger datasets collected in LHC run II and beyond the careful evaluation of these effects will become increasingly important.

For the analysis of the angular distributions the full systematic uncertainty, given by the quadratic sum of the uncertainties from all sources, is listed in Table 16.12.

Table 16.11: Summary of the systematic and statistical uncertainties on the relative differential branching fraction [$10^{-5} \text{ GeV}^{-2} c^4$] in bins of q^2 [GeV^2/c^4].

Source	[0.1, 2]	[2, 5]	[5, 8]	[11, 12.5]	[15, 17]	[17, 19]	[1, 6]	[15, 19]
Peak. bkg.	0.04	0.02	0.03	0.12	0.03	0.01	0.02	0.02
Corrections	0.02	0.02	0.01	0.00	0.05	0.04	0.00	0.05
Model dep.	0.04	0.00	0.01	0.00	0.01	0.06	0.00	0.01
$\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$	0.03	0.01	0.02	0.02	0.02	0.02	0.01	0.02
Sim. stat.	0.06	0.03	0.03	0.06	0.06	0.07	0.02	0.04
Sig. mod.	0.02	0.01	0.03	0.03	0.03	0.00	0.05	0.05
Bkg. mod.	0.02	0.02	0.02	0.02	0.03	0.05	0.01	0.06
Time acc.	0.09	0.04	0.05	0.07	0.07	0.06	0.04	0.06
Overall	0.13	0.06	0.07	0.15	0.11	0.13	0.07	0.12
Statistical	+0.68 -0.64	+0.39 -0.37	+0.41 -0.39	+0.64 -0.61	+0.53 -0.50	+0.53 -0.50	+0.30 -0.29	+0.37 -0.35

Table 16.12: Full systematic uncertainties of the angular fit parameters in bins of q^2 .

q^2 [GeV^2/c^4]	F_L	S_3	S_4	A_5	A_6	S_7	A_8	A_9
0.1 – 2.0	0.018	0.005	0.007	0.003	0.011	0.004	0.004	0.005
2.0 – 5.0	0.034	0.008	0.009	0.007	0.016	0.006	0.007	0.009
5.0 – 8.0	0.017	0.006	0.007	0.005	0.010	0.005	0.006	0.005
11.0 – 12.5	0.045	0.011	0.012	0.007	0.015	0.005	0.008	0.008
15.0 – 17.0	0.017	0.013	0.010	0.006	0.011	0.006	0.006	0.006
17.0 – 19.0	0.020	0.024	0.019	0.014	0.013	0.011	0.011	0.012
1.0 – 6.0	0.027	0.006	0.005	0.004	0.010	0.003	0.004	0.005
15.0 – 19.0	0.017	0.010	0.008	0.003	0.007	0.005	0.003	0.004

The largest contribution to this total uncertainty comes from the background model chosen and the precision of the parametrisation of the background distributions. As those are determined in a data-driven way, these uncertainties can only be reduced by increasing the dataset, and thus will automatically decrease along with the statistical uncertainties. Similar to the measurement of the differential branching fraction, statistical uncertainties are by far dominant compared to the total systematic uncertainties for all parameters and q^2 bins.

17 Results

The results and systematic uncertainties for the measurements of the relative differential branching fraction (Chap. 14) and the analysis of the angular distributions (Chap. 15) are combined with systematic uncertainties derived in Chap. 16. This chapter presents the final results of both measurements. For the differential branching fraction measurement, the relative value is also propagated to an absolute branching fraction, and extrapolated to the full q^2 range. The results are compared to predictions based on Standard Model calculations, and their agreement is discussed.

17.1 Differential branching fraction

Table 17.1 lists the results for the relative differential branching fraction in bins of the invariant dimuon mass squared q^2 , including statistical and systematic uncertainties. In order to extract the differential branching fraction of the decay mode $B_s^0 \rightarrow \phi\mu^+\mu^-$ from the relative value calculated from Eq. 10.1, the branching fraction of the decay mode $B_s^0 \rightarrow J/\psi\phi$ is used. As this branching fraction enters the calculation in first order, its relative uncertainty directly propagates to the final result of the branching fraction.

17.1.1 Branching fraction of $B_s^0 \rightarrow J/\psi\phi$

At the time this analysis was performed, the published world average was $\mathcal{B}(B_s^0 \rightarrow J/\psi\phi) = (11.0 \pm 0.9) \cdot 10^{-4}$ [15], where the uncertainty is dominated by the most recent LHCb result of $\mathcal{B}(B_s^0 \rightarrow J/\psi\phi) = (10.50 \pm 0.13 \pm 0.64 \pm 0.82) \cdot 10^{-4}$ [66]. The first uncertainty is statistical, the second systematic, and the third due to the uncertainty

Table 17.1: Relative differential branching fraction of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$ compared to $B_s^0 \rightarrow J/\psi\phi$ in bins of q^2 . The first uncertainty given is statistical, the second systematic.

$q^2 [\text{GeV}^2/c^4]$	$\frac{1}{\mathcal{B}(B_s^0 \rightarrow J/\psi\phi)} \frac{d\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{dq^2} [10^{-5} \text{GeV}^{-2}c^4]$
0.1 – 2.0	$5.44_{-0.64}^{+0.68} \pm 0.13$
2.0 – 5.0	$2.38_{-0.37}^{+0.39} \pm 0.06$
5.0 – 8.0	$2.98_{-0.39}^{+0.41} \pm 0.07$
11.0 – 12.5	$4.37_{-0.61}^{+0.64} \pm 0.15$
15.0 – 17.0	$4.20_{-0.50}^{+0.53} \pm 0.11$
17.0 – 19.0	$3.68_{-0.50}^{+0.53} \pm 0.13$
1.0 – 6.0	$2.40_{-0.29}^{+0.30} \pm 0.07$
15.0 – 19.0	$3.75_{-0.35}^{+0.37} \pm 0.12$

of the measurement of hadronisation fractions at the LHC of $f_s/f_d = 0.256 \pm 0.020$. However, after the publication of this measurement, the value for these hadronisation fractions has been updated to $f_s/f_d = 0.259 \pm 0.015$ [67], reducing the relative uncertainty significantly. With this more recent value, the LHCb measurement can be updated to $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) = (10.38 \pm 0.13 \pm 0.63 \pm 0.60) \cdot 10^{-4}$ and the world average value recalculated to $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) = (10.76 \pm 0.81) \cdot 10^{-4}$, using the weighted averaging technique as described by the PDG review [15]. This corresponds to a relative uncertainty of 7.5%, which is directly propagated to a relative uncertainty of the (differential) branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$.

17.1.2 Differential branching fraction

With this updated value for the branching fraction of $B_s^0 \rightarrow J/\psi \phi$, the differential branching fraction of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ is calculated. The results are listed in Tab. 17.2, where the last uncertainty is due to the uncertainty of the normalisation mode. The measured results are in good agreement with the former LHCb measurement [24], which used the dataset recorded by LHCb in 2011. Figure 17.1 shows the q^2 dependence of the differential branching fraction and compares it with predictions based on Standard Model calculations [27, 28, 45]. The uncertainties for the Standard Model predictions are dominated by form factor uncertainties, which are calculated using lattice QCD for high q^2 and light cone sum rules for the low q^2 regions, which is where the respective methods are most precise. For the wide high q^2 bin covering the range of $15 < q^2 < 19 \text{ GeV}^2/c^4$ an additional prediction is available using only form factors from lattice QCD calculations. No predictions are available for the regions close to and in between the charmonium resonances, where no precise calculations can be performed due to those. Both in the narrow bin of $2.0 < q^2 < 5.0 \text{ GeV}^2/c^4$ and the wide bin of $1.0 < q^2 < 6.0 \text{ GeV}^2/c^4$ the measured branching fraction results are significantly lower than the theoretical prediction of $(4.81 \pm 0.56) \times 10^{-8} \text{ GeV}^{-2} c^4$, the significance for the wider bin is calculated to be 3.3 standard deviations. At high q^2 the prediction using exclusively lattice QCD calculations are slightly higher than the combined prediction and the measured value. However, the significance of this deviation is low.

The discrepancy between the measurement and the prediction made from Standard Model calculations fit well into what was observed in measurements of other electroweak penguin decays by the LHCb experiment [68–70], which will be further discussed later in this chapter.

17.1.3 Extrapolation to the full q^2 range

The total relative and absolute branching fractions of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ can be calculated by extrapolating the differential branching fraction to the full q^2 range. This extrapolation is performed by integrating the branching fraction over the six narrow q^2 bins, and multiplying it by a extrapolation factor f_{extra} to account for the fraction of signal events in vetoed q^2 regions. This extrapolation factor is determined

q^2 [GeV ² /c ⁴]	$\frac{d\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{dq^2}$ [10^{-8} GeV ⁻² c ⁴]
0.1 – 2.0	$5.85_{-0.69}^{+0.73} \pm 0.14 \pm 0.44$
2.0 – 5.0	$2.56_{-0.39}^{+0.42} \pm 0.06 \pm 0.19$
5.0 – 8.0	$3.21_{-0.42}^{+0.44} \pm 0.08 \pm 0.24$
11.0 – 12.5	$4.71_{-0.65}^{+0.69} \pm 0.16 \pm 0.36$
15.0 – 17.0	$4.52_{-0.54}^{+0.57} \pm 0.12 \pm 0.34$
17.0 – 19.0	$3.96_{-0.54}^{+0.57} \pm 0.14 \pm 0.30$
1.0 – 6.0	$2.58_{-0.31}^{+0.33} \pm 0.08 \pm 0.19$
15.0 – 19.0	$4.04_{-0.38}^{+0.39} \pm 0.13 \pm 0.30$

Table 17.2: Differential branching fraction of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$ in bins of q^2 . The first uncertainty given is statistical, the second systematic, and the third due to the uncertainty of $\mathcal{B}(B_s^0 \rightarrow J/\psi\phi)$.

using calculations from Ref. [71], including updated form factors from Ref. [72], to generate the q^2 distributions of simulated $B_s^0 \rightarrow \phi\mu^+\mu^-$ events. To evaluate the systematic uncertainty of this extrapolation, the generation is repeated with the calculated form factor values and Wilson coefficients varied within the theoretical uncertainties, and the largest deviation assigned as uncertainty. The extrapolation factor is determined to be $f_{\text{extra}} = 1.520 \pm 0.003 \pm 0.043$, where the first uncertainty given is statistical due to the number of events generated, and the second systematic. This results in the relative and absolute branching fractions of

$$\frac{\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-)}{\mathcal{B}(B_s^0 \rightarrow J/\psi\phi)} = (7.41_{-0.40}^{+0.42} \pm 0.20 \pm 0.21) \cdot 10^{-4}$$

and

$$\mathcal{B}(B_s^0 \rightarrow \phi\mu^+\mu^-) = (7.97_{-0.43}^{+0.45} \pm 0.22 \pm 0.23 \pm 0.60) \cdot 10^{-7},$$

where the first uncertainty is statistical, the second systematic, and the third due to the extrapolation to the full q^2 range. For the branching fraction a further uncertainty is assigned due to the uncertainty of $\mathcal{B}(B_s^0 \rightarrow J/\psi\phi)$.

17.2 Analysis of the angular distributions

The values for the angular fit parameters, including all evaluated uncertainties are listed in Tabs. 17.3 and 17.4 in bins of q^2 . Figure 17.2 shows these results as function of q^2 and in comparison to predictions from a combination of light cone sum rules and lattice QCD calculations [27, 28]. For the CP -averaged parameters $A_{5,6,8,9}$ and the parameter S_7 no predictions are available, as their deviation from 0 in the Standard Model is multiple orders of magnitude smaller than the experimental precision. All

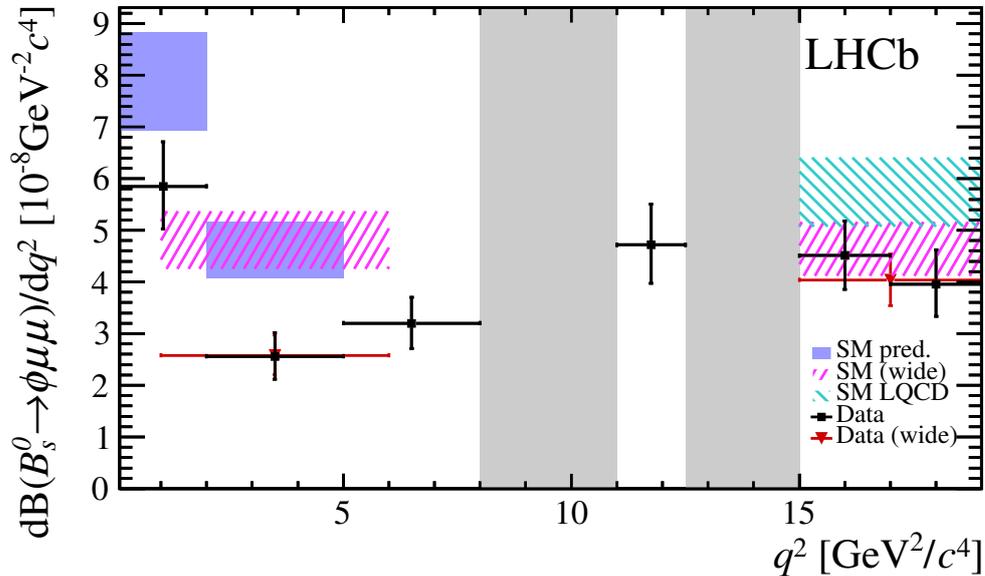


Figure 17.1: Differential branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$ as a function of the dimuon mass squared q^2 . Measurements are shown by black points for narrow and red triangles for wide bins. Standard Model predictions using form factors from a combination of lattice QCD and light cone sum rules calculations [27, 28] are shown by blue and hatched magenta areas. Additionally, for the wide high q^2 bin the prediction exclusively from lattice QCD calculations [45] is indicated by the hatched turquoise area. The vetoed regions around the charmonium resonances are indicated by the grey areas.

measured parameters agree well with the theoretical predictions. This is the first measurement of the parameters S_4 and S_7 for the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$, as the previous LHCb measurement [24] used a folding technique in the angle Φ to account for the lower amount of data available. The parameters A_5 and A_6 are measured for the first time for any rare $b \rightarrow s \ell^+ \ell^-$ transition, and provide new constraints for global fits in the search for New Physics in FCNC processes.

17.3 Implications

The results presented in this chapter provide new constraints to the theoretical understanding of electroweak $b \rightarrow s \ell^+ \ell^-$ transitions. Of special interest is the differential branching fraction measurement at low q^2 , which is more than 3σ below the expectation from Standard Model calculations, consistent with measurements of other $b \rightarrow s \ell^+ \ell^-$ decays. Nevertheless, even the parameters measured by analysing the distributions of the decay angles are used in so-called “global” fits, as presented in Refs. [73–76]. These global fits use the measurements of differential decay rates and

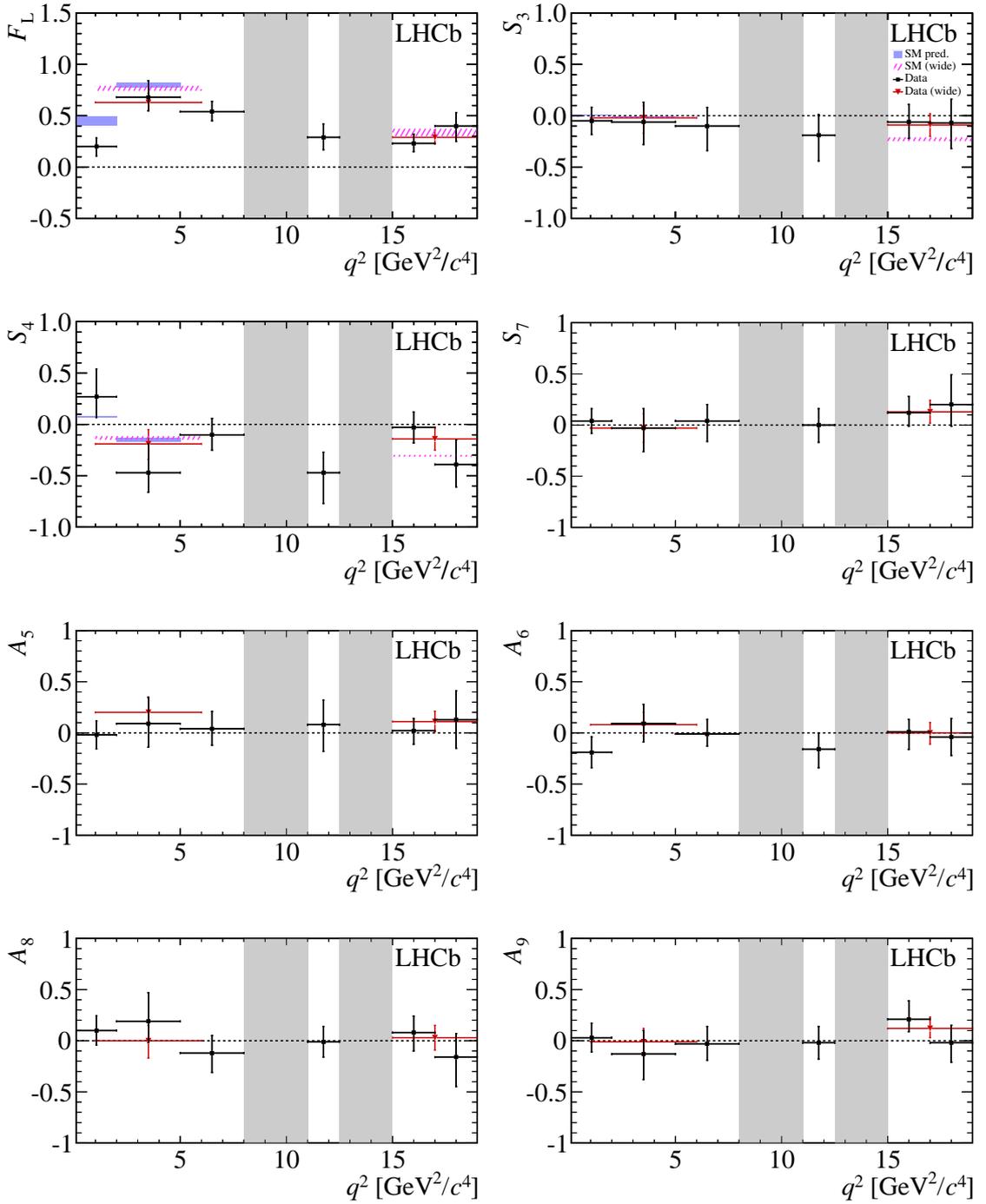


Figure 17.2: Angular parameters F_L , $S_{3,4,7}$ and $A_{5,6,8,9}$ as function of the dimuon mass squared q^2 . Measurements are shown by black points for narrow and red triangles for wide bins. Standard Model predictions using form factors from a combination of lattice QCD and light cone sum rules calculations [27, 28] are shown by blue and hatched magenta areas. The vetoed regions around the charmonium resonances are indicated by the grey areas.

Table 17.3: Results for the parameters F_L and $S_{3,4,7}$ in bins of q^2 . The first uncertainty is statistical, the second systematic.

q^2 [GeV ² /c ⁴]	F_L	S_3	S_4	S_7
0.1 – 2.0	$0.20^{+0.08}_{-0.09} \pm 0.02$	$-0.05^{+0.13}_{-0.13} \pm 0.01$	$0.27^{+0.28}_{-0.18} \pm 0.01$	$0.04^{+0.12}_{-0.12} \pm 0.00$
2.0 – 5.0	$0.68^{+0.16}_{-0.13} \pm 0.03$	$-0.06^{+0.19}_{-0.23} \pm 0.01$	$-0.47^{+0.30}_{-0.44} \pm 0.01$	$-0.03^{+0.18}_{-0.23} \pm 0.01$
5.0 – 8.0	$0.54^{+0.10}_{-0.09} \pm 0.02$	$-0.10^{+0.20}_{-0.29} \pm 0.01$	$-0.10^{+0.15}_{-0.18} \pm 0.01$	$0.04^{+0.16}_{-0.20} \pm 0.01$
11.0 – 12.5	$0.29^{+0.11}_{-0.11} \pm 0.04$	$-0.19^{+0.20}_{-0.23} \pm 0.01$	$-0.47^{+0.21}_{-0.29} \pm 0.01$	$0.00^{+0.15}_{-0.10} \pm 0.01$
15.0 – 17.0	$0.23^{+0.09}_{-0.08} \pm 0.02$	$-0.06^{+0.16}_{-0.19} \pm 0.01$	$-0.03^{+0.15}_{-0.15} \pm 0.01$	$0.12^{+0.16}_{-0.13} \pm 0.01$
17.0 – 19.0	$0.40^{+0.13}_{-0.15} \pm 0.02$	$-0.07^{+0.23}_{-0.27} \pm 0.02$	$-0.39^{+0.25}_{-0.34} \pm 0.02$	$0.20^{+0.29}_{-0.22} \pm 0.01$
1.0 – 6.0	$0.63^{+0.09}_{-0.09} \pm 0.03$	$-0.02^{+0.12}_{-0.13} \pm 0.01$	$-0.19^{+0.14}_{-0.13} \pm 0.01$	$-0.03^{+0.14}_{-0.14} \pm 0.00$
15.0 – 19.0	$0.29^{+0.07}_{-0.06} \pm 0.02$	$-0.09^{+0.11}_{-0.12} \pm 0.01$	$-0.14^{+0.11}_{-0.11} \pm 0.01$	$0.13^{+0.11}_{-0.11} \pm 0.01$

Table 17.4: Results for the parameters $A_{5,6,8,9}$ in bins of q^2 . The first uncertainty is statistical, the second systematic.

q^2 [GeV ² /c ⁴]	A_5	A_6	A_8	A_9
0.1 – 2.0	$-0.02^{+0.13}_{-0.13} \pm 0.00$	$-0.19^{+0.15}_{-0.15} \pm 0.01$	$0.10^{+0.14}_{-0.14} \pm 0.00$	$0.03^{+0.14}_{-0.14} \pm 0.01$
2.0 – 5.0	$0.09^{+0.28}_{-0.22} \pm 0.01$	$0.09^{+0.20}_{-0.19} \pm 0.02$	$0.19^{+0.26}_{-0.21} \pm 0.01$	$-0.13^{+0.24}_{-0.30} \pm 0.01$
5.0 – 8.0	$0.04^{+0.17}_{-0.17} \pm 0.01$	$-0.01^{+0.14}_{-0.12} \pm 0.01$	$-0.12^{+0.17}_{-0.19} \pm 0.01$	$-0.03^{+0.17}_{-0.16} \pm 0.01$
11.0 – 12.5	$0.08^{+0.21}_{-0.21} \pm 0.01$	$-0.16^{+0.16}_{-0.18} \pm 0.01$	$-0.01^{+0.15}_{-0.15} \pm 0.01$	$-0.02^{+0.16}_{-0.15} \pm 0.01$
15.0 – 17.0	$0.02^{+0.13}_{-0.14} \pm 0.01$	$0.01^{+0.12}_{-0.17} \pm 0.01$	$0.08^{+0.16}_{-0.18} \pm 0.01$	$0.21^{+0.18}_{-0.12} \pm 0.01$
17.0 – 19.0	$0.13^{+0.29}_{-0.27} \pm 0.01$	$-0.04^{+0.18}_{-0.19} \pm 0.01$	$-0.16^{+0.24}_{-0.29} \pm 0.01$	$-0.02^{+0.19}_{-0.19} \pm 0.01$
1.0 – 6.0	$0.20^{+0.13}_{-0.13} \pm 0.00$	$0.08^{+0.12}_{-0.11} \pm 0.01$	$-0.00^{+0.15}_{-0.17} \pm 0.00$	$-0.01^{+0.13}_{-0.13} \pm 0.01$
15.0 – 19.0	$0.11^{+0.10}_{-0.10} \pm 0.00$	$0.00^{+0.10}_{-0.11} \pm 0.01$	$0.03^{+0.12}_{-0.12} \pm 0.00$	$0.12^{+0.09}_{-0.11} \pm 0.00$

angular parameters from a large set of $b \rightarrow s\ell^+\ell^-$ decays to determine the agreement with the Standard Model predictions of the Wilson coefficients. Additionally, scenarios of physics beyond what is described by the Standard Model are tested by adding a floating New Physics coefficient C_i^{NP} to each SM Wilson coefficient, which is a free parameter of the global fit. A common outcome of those global fits is that the current measurements seem to favour New Physics contributions to the Wilson coefficient C_9 , which corresponds to a vector-vector coupling in the interaction.

Figure 17.3 shows the two-dimensional constraints derived from one of the most recent publication [76] regarding contributions from BSM physics to Wilson coefficients C_9 , and C_{10} or C'_9 , respectively, introduced via a modification of the SM coefficients $C_9 \rightarrow C_9 + C_9^{NP}$ and $C_{10} \rightarrow C_{10} + C_{10}^{NP}$, respectively. Additionally the presence of

right-handed currents is tested by allowing the corresponding Wilson coefficients C'_i to float in fit, which are identical to 0 in the SM. For the fits presented in Fig. 17.3 the New Physics contributions C_9^{NP} and C_{10}^{NP} (or C'_9) are free parameters, while all other Wilson coefficients are assumed to be SM-like. The Standard Model predictions without any New Physics effects would correspond to the C_i^{NP} values to be exactly 0. For the measurements performed by each of the ATLAS, CMS, and LHCb collaborations the contours corresponding to 1σ confidence intervals are shown. Additionally, the combination of only the differential branching fraction measurements of all three collaborations is shown with a 1σ contour. The combination of all measurements is presented with contours corresponding to 1, 2, and 3σ confidence intervals, which shows a clear tension with the SM prediction of $C_9^{NP} = C_{10}^{NP} = 0$. However, as no significant deviations are observed for the New Physics contributions C_{10}^{NP} or C'_9 , the significance of this tension is also evaluated by assuming them to be SM-like. This would correspond to a New Physics contribution C_9^{NP} with a significance of about $4 - 5\sigma$, which would completely resolve the tensions observed with SM predictions.

New Physics models, which could lead to such a shift in the Wilson coefficient

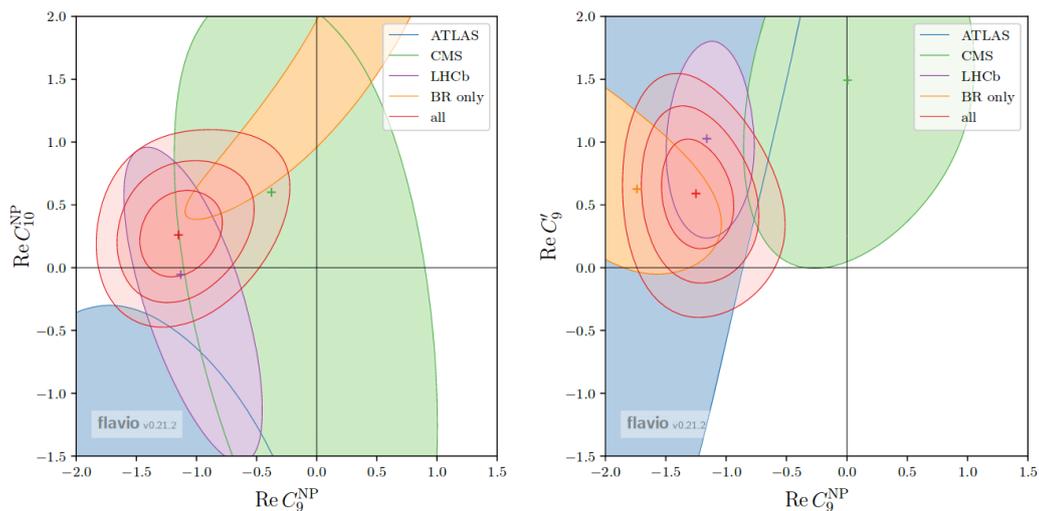


Figure 17.3: Two-dimensional confidence intervals in the real part of New Physics contributions C_9^{NP} vs. C_{10}^{NP} in the Wilson coefficients C_9 vs. C_{10} (left) and C_9^{NP} vs. C'_9 (right) [76], assuming all other Wilson coefficients are at Standard Model values. The 1σ contours are given for the measurements of individual experiments and using only branching fraction results. The combination of all results shows 1, 2, and 3σ contours.

C_9 , are models introducing a Z' boson [48] or leptoquarks [49] mediating flavour changing neutral currents at tree level, as shown in Fig. 17.4(a).

However, the origin of the tensions between observations and the predictions from

Standard Model calculations do not necessarily need to be physics beyond the Standard Model. Instead, underestimated contributions from hadronic $c\bar{c}$ loops might explain the observed discrepancy. Figure 17.4(b) shows sketches of these potential effects. An important probe to distinguish New Physics effects from the potential influence from charmonium loops is the measurement of sensitive variables as a function of q^2 . While the New Physics effects are in general independent of q^2 , the $c\bar{c}$ contributions are larger for q^2 close to the mass squared of the charmonium resonances. Figure 17.5 shows the value of C_9 extracted from a global fit [76] as a function of q^2 . With the current experimental precision, there are no hints of a q^2 dependence.

Another way to understand the origin of the anomalies observed in the electroweak

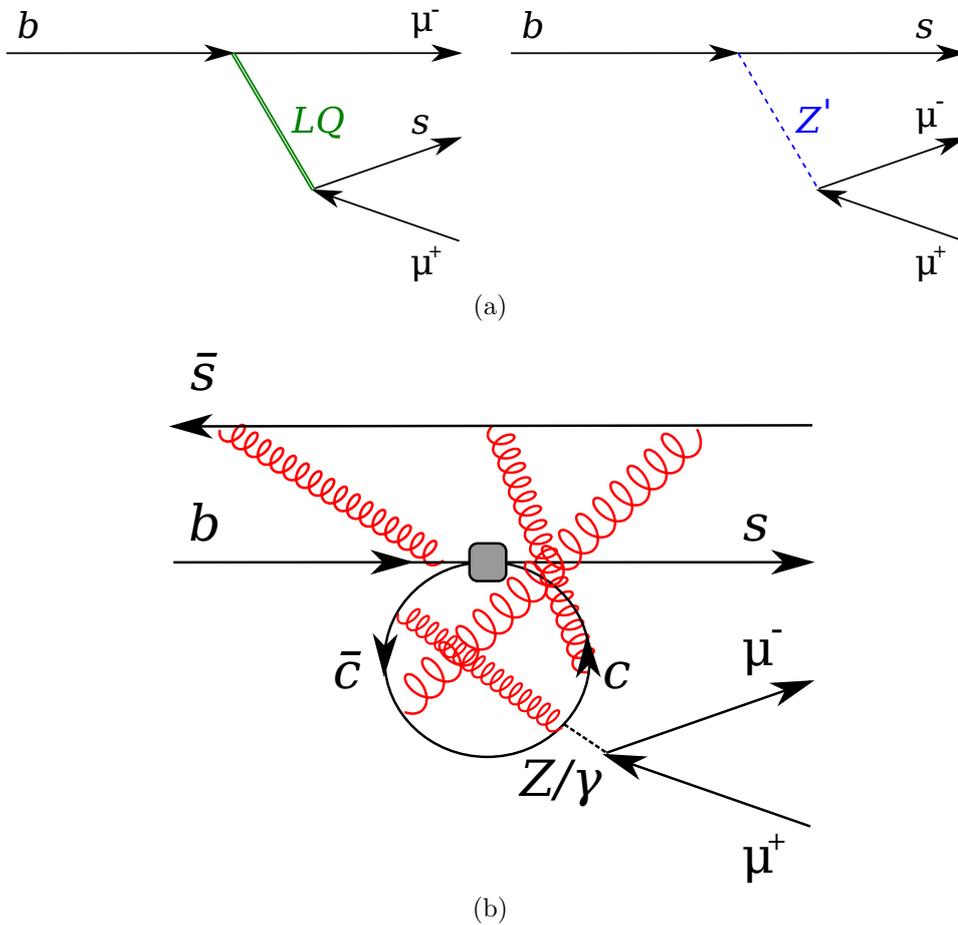


Figure 17.4: (a) Feynman diagrams of tree-level $b \rightarrow s \ell^+ \ell^-$ transitions mediated by Z' and lepto-quarks. (b) $c\bar{c}$ contributions to the Standard Model penguin transition. The grey box symbolises the four-point interaction described by the effective field theory.

penguin decays are measurements of lepton universality, as the charmonium loops

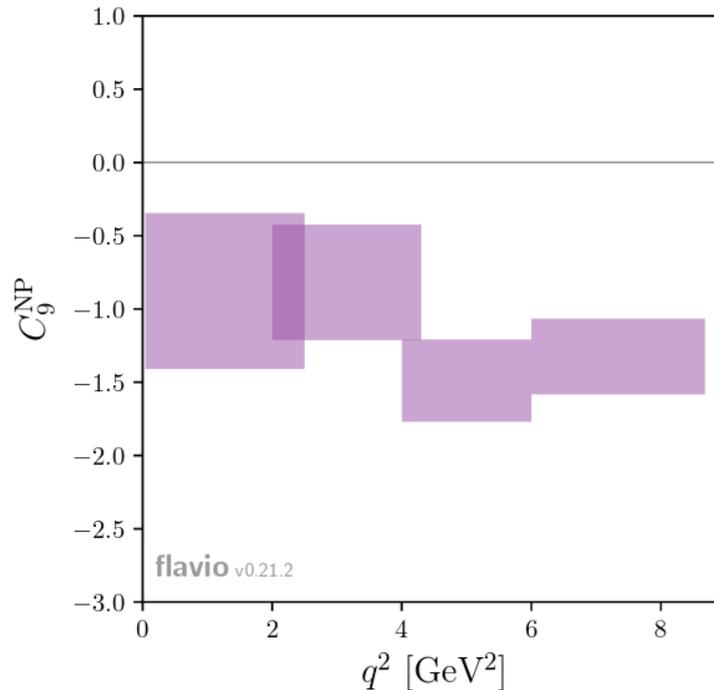


Figure 17.5: Extracted 1σ confident intervals for a shift relative to the SM prediction of the Wilson coefficient C_9 as a function of q^2 [76].

affect all lepton generations in the same manner. Measurements of lepton universality in electroweak penguin decays are performed by comparing decays with muons and electrons in the final state, *e.g.* $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} e^+ e^-$. Measuring the q^2 dependent ratio of the branching fraction of those decays allows the comparison to very precise theoretical predictions, as many uncertainties cancel. The main challenge to perform these lepton universality measurements at the LHCb experiment is the reconstruction of the electrons, as they lose a large fraction of their energy due to Bremsstrahlung. For that reason the reconstructed invariant mass distribution of the primary b hadron has long tails towards lower masses, complicating the background treatment. Also the reconstruction efficiency for electrons is significantly lower than for muons, reducing the available data sample size for the already rare decay further. Two lepton universality measurements in electroweak penguin decays have been published by the LHCb experiment [26, 77] and further decay modes are under investigation.

Global fits are performed including those lepton-universality measurements, where the New Physics contributions are left floating only for the muon sector, by introducing shifts to the SM Wilson coefficients $C_9 \rightarrow C_9 + C_9^\mu$ and $C_{10} \rightarrow C_{10} + C_{10}^\mu$, where the coupling to the electrons is assumed to be SM-like. Figure 17.6 shows the two-dimensional constraints on these New Physics contributions C_9^μ and C_{10}^μ . The

significance of the deviation from Standard Model predictions increases to $5-6\sigma$. The lepton flavour universality measurement investigated on its own shows a significance of about 4σ , due to the clean theoretical predictions based on SM calculations. Future measurements will contribute to better understanding the picture of lepton universality, and larger amounts of available data allow to additionally measure the ratio of angular observables. For those ratios typically uncertainties from form factor calculations largely cancel, which allows for very precise theoretical predictions.

The larger datasets that will be collected by the LHCb experiment in LHC run II

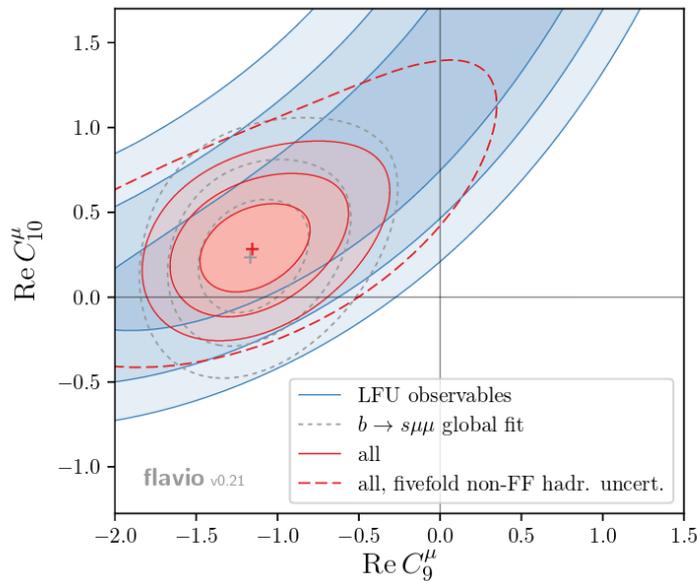


Figure 17.6: Two-dimensional confidence intervals in the real part of New Physics contributions C_9^μ vs. C_{10}^μ [75], assuming all other Wilson coefficients are at Standard Model values. 1, 2, and 3σ contours are shown for the lepton flavour universality constraints alone and the combination with branching fraction and angular analyses.

and beyond will help immensely in narrowing down on the origin of the anomalies observed, and might point towards specific models of physics beyond the description of the Standard Model.

18 Conclusion

The second part of this thesis presents the measurement of the differential branching fraction and an analysis of the distributions of the decay angles of the electroweak penguin decay $B_s^0 \rightarrow \phi\mu^+\mu^-$. The full dataset of 3fb^{-1} collected by the LHCb during LHC Run I is used for this measurement, which provides the first full angular analysis of the investigated decay mode. An extended maximum likelihood fit is performed to the reconstructed invariant mass distribution of the selected candidates. The yields of the signal mode and the normalisation channel $B_s^0 \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\phi$ are extracted from this fit, and used to calculate the relative differential branching fraction between those decays as a function of the dimuon mass squared q^2 . Furthermore, a four-dimensional maximum likelihood fit is performed to distributions of the reconstructed invariant mass and the three decay angles θ_l , θ_K , and Φ . From this fit the eight angular parameters F_L , $S_{3,4,7}$, and $A_{5,6,8,9}$ are extracted as a function of q^2 , which are tied to the Wilson coefficients of the effective field theory describing $b \rightarrow s\ell^+\ell^-$ transitions. Possible sources of systematic uncertainties are evaluated. However, due to the rare occurrence of the signal decay, statistical uncertainties are dominating the results.

The results of this measurements agree with predictions made from Standard Model deviations for all angular observables, but for the differential branching fraction a significant deviation of 3.3σ is observed for the dimuon mass squared region of $1 < q^2 < 6\text{GeV}^{-2}c^4$. This observations is in agreement with similar deviations seen in the low- q^2 region of other electroweak penguin decays, and the anomaly observed in the angular observable P'_5 of the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$.

While these deviations are compatible with New Physics models, where the flavour changing neutral current could be mediated on tree level by a Z' boson or leptoquarks, the underestimation of contributions from charmonium loops in the theoretical predictions cannot be excluded yet. Measurements of lepton universality in electroweak penguin decays also disagree with the Standard Model.

In order to fully resolve the anomalies in $b \rightarrow s\ell^+\ell^-$ more precise measurements will be needed. These will be possible by using the larger dataset collected by the LHCb experiment in LHC run II and beyond, as well as future b -physics experiments such as BELLE2. During Run II the LHCb experiment is expected to collect an additional integrated luminosity of about 8fb^{-1} , while the b -hadron production cross section has roughly doubled compared to Run I due to the higher centre-of-mass energy of $\sqrt{s} = 13\text{TeV}$. Naive scaling leads to an expected increase in the $B_s^0 \rightarrow \phi\mu^+\mu^-$ candidates by about 530%. Assuming that the systematic uncertainties which are tied to the dataset size scale similarly would reduce the experimental uncertainties by about 60%. With these uncertainties, theoretical uncertainties as for the current predictions would be dominant. However, the anomaly in the low q^2 regions could be measured with close to 4σ sensitivity. Including more decay modes, such as $B_s^0 \rightarrow \phi e^+e^-$, will allow to more precisely study the lepton universality of those anomalies, given additional constraints to New Physics models.

Appendix

A 1 Angular acceptance in q^2 bins

This section contains one-dimensional and two-dimensional projections of the angular acceptance and the chosen parametrisation in q^2 bins. They are shown in Figs. A 1.1 to A 1.8.

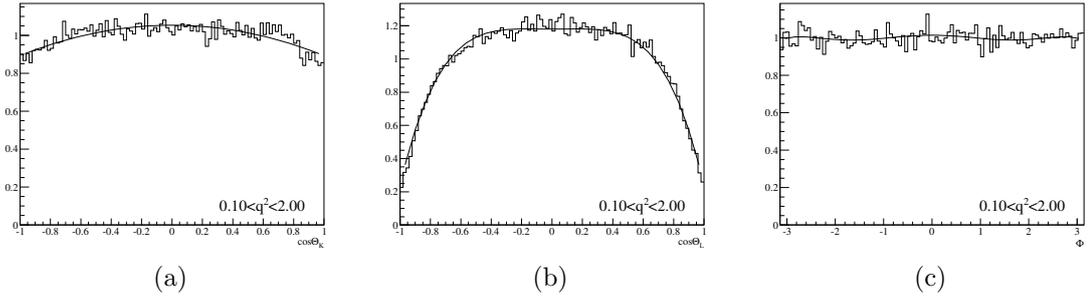


Figure A 1.1: One-dimensional projections of the angular acceptance for $0.1 < q^2 < 2.0 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos \theta_K$ (a), $\cos \theta_K$ (b), Φ (c).

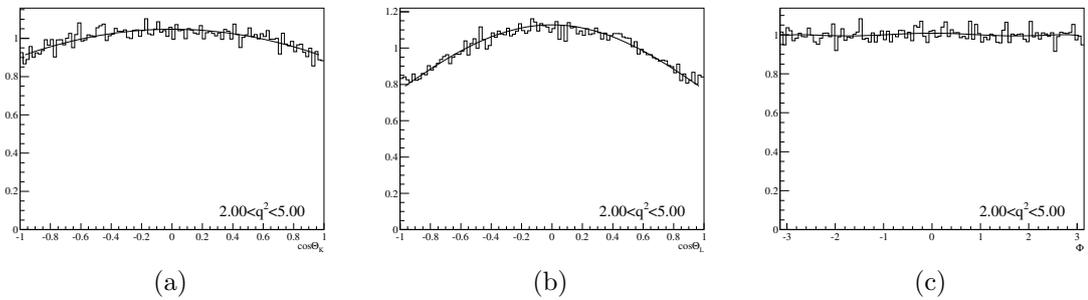


Figure A 1.2: One-dimensional projections of the angular acceptance for $2.0 < q^2 < 5.0 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos \theta_K$ (a), $\cos \theta_K$ (b), Φ (c).

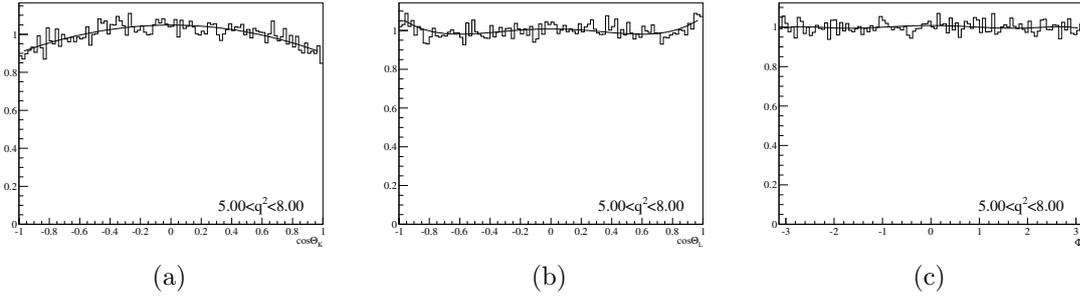


Figure A 1.3: One-dimensional projections of the angular acceptance for $5.0 < q^2 < 8.0 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos\theta_K$ (a), $\cos\theta_K$ (b), Φ (c).

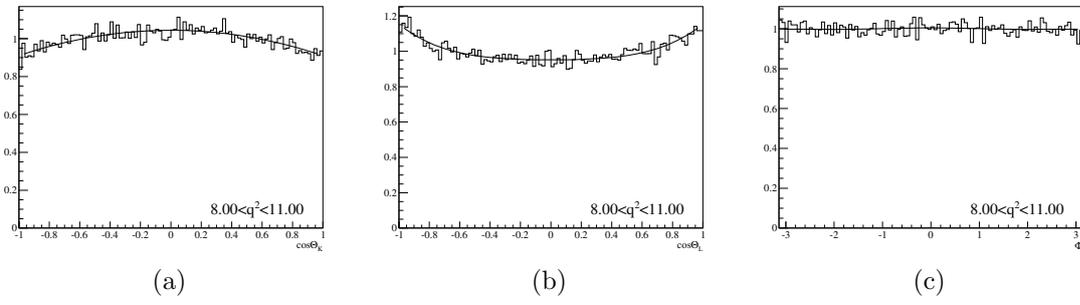


Figure A 1.4: One-dimensional projections of the angular acceptance for $11.0 < q^2 < 12.5 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos\theta_K$ (a), $\cos\theta_K$ (b), Φ (c).

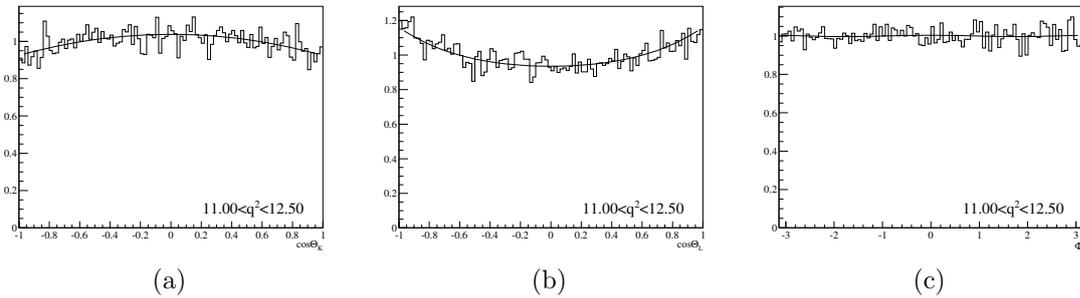


Figure A 1.5: One-dimensional projections of the angular acceptance for $15.0 < q^2 < 17.0 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos\theta_K$ (a), $\cos\theta_K$ (b), Φ (c).

A 1. ANGULAR ACCEPTANCE IN Q^2 BINS

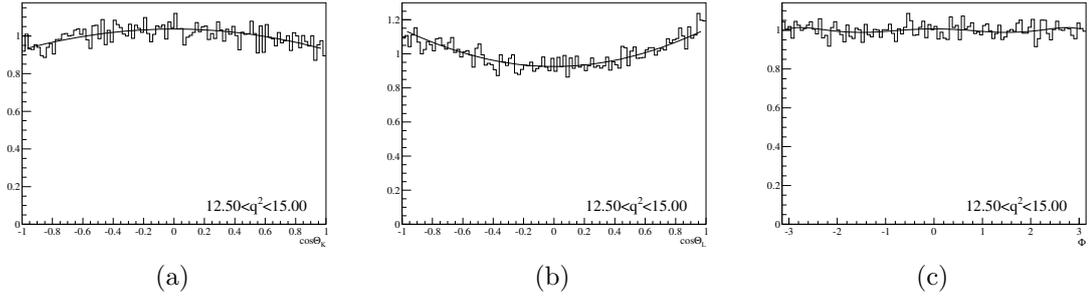


Figure A 1.6: One-dimensional projections of the angular acceptance for $17.0 < q^2 < 19.0 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos \theta_K$ (a), $\cos \theta_K$ (b), Φ (c).

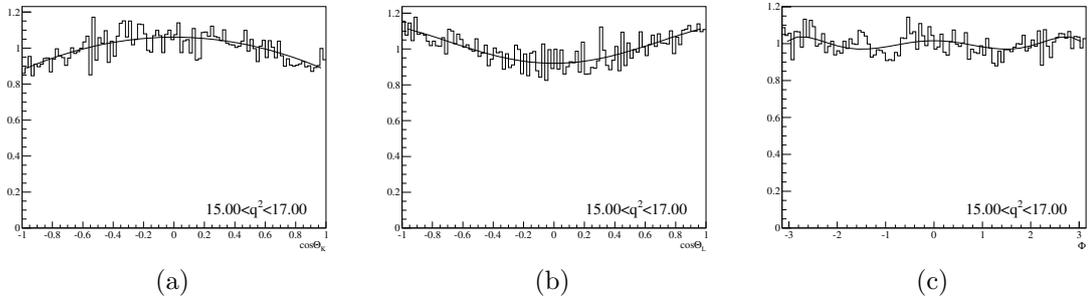


Figure A 1.7: One-dimensional projections of the angular acceptance for $1.0 < q^2 < 6.0 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos \theta_K$ (a), $\cos \theta_K$ (b), Φ (c).

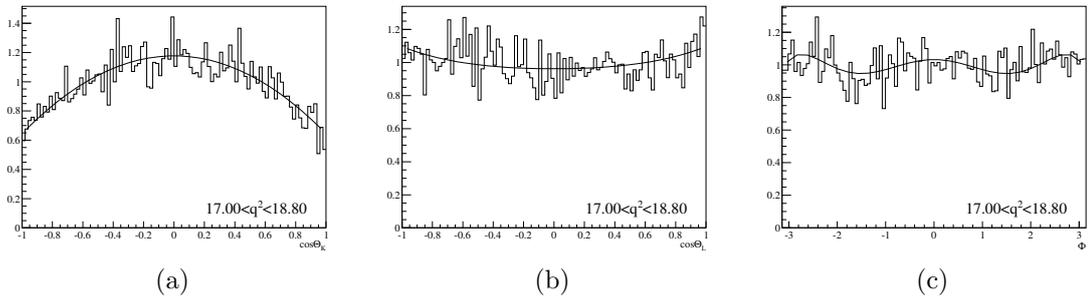


Figure A 1.8: One-dimensional projections of the angular acceptance for $15.0 < q^2 < 19.0 \text{ GeV}^2/c^4$ overlaid by the chosen parametrisation in the four observables: $\cos \theta_K$ (a), $\cos \theta_K$ (b), Φ (c).

A 2 Angular acceptance: two-dimensional projections

Figure A 2.1 shows the two-dimensional angular acceptance projections of all combinations of the four parameters in which the angular acceptance is evaluated. In addition, Figs. A 2.2 to A 2.9 show the two-dimensional projections for combinations of the decay angles in the chosen q^2 bins.

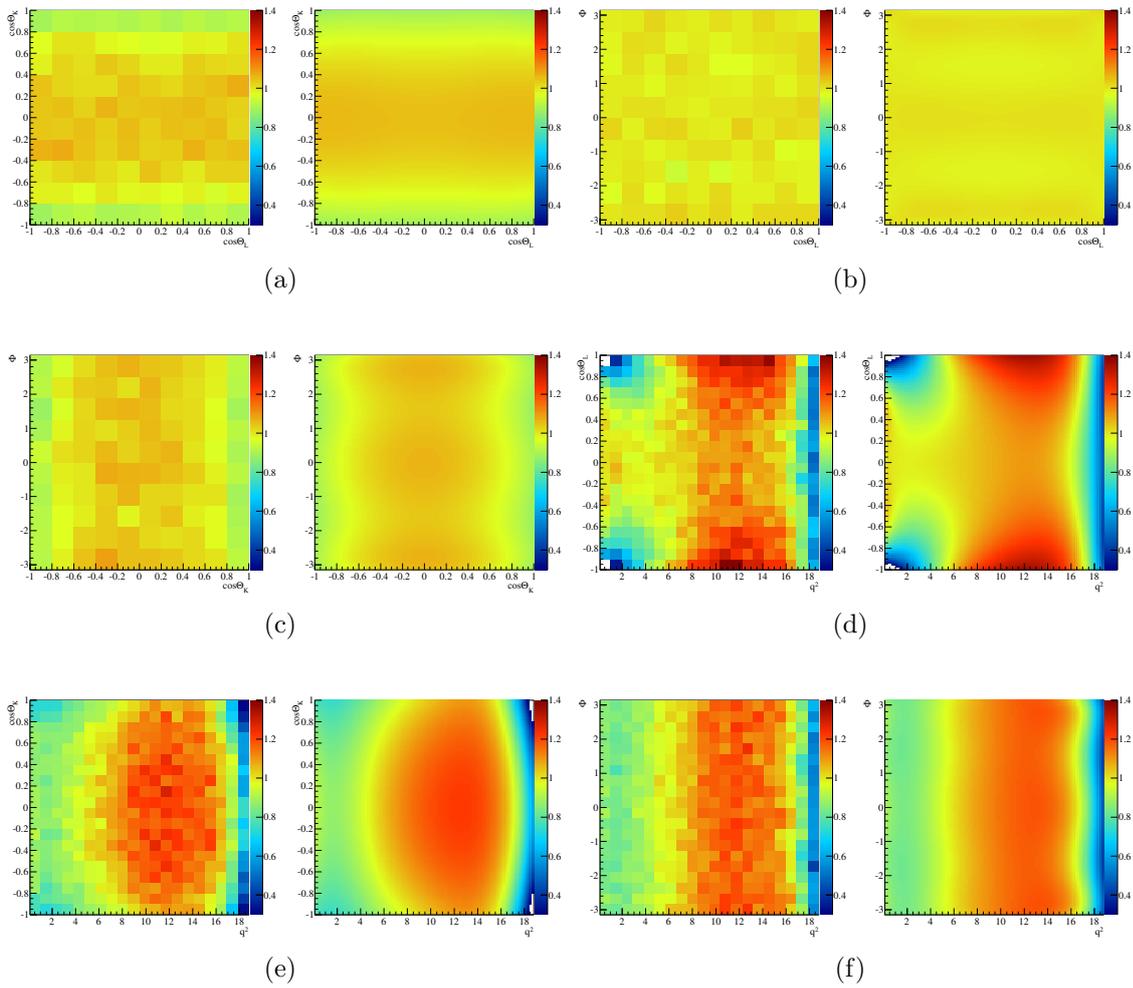


Figure A 2.1: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation in the six observable combinations: $\cos\theta_l$ vs. $\cos\theta_K$ (a), $\cos\theta_l$ vs. Φ (b), $\cos\theta_K$ vs. Φ (c), q^2 vs. $\cos\theta_l$ (d), q^2 vs. $\cos\theta_K$ (e), q^2 vs. Φ .

A 2. ANGULAR ACCEPTANCE: TWO-DIMENSIONAL PROJECTIONS

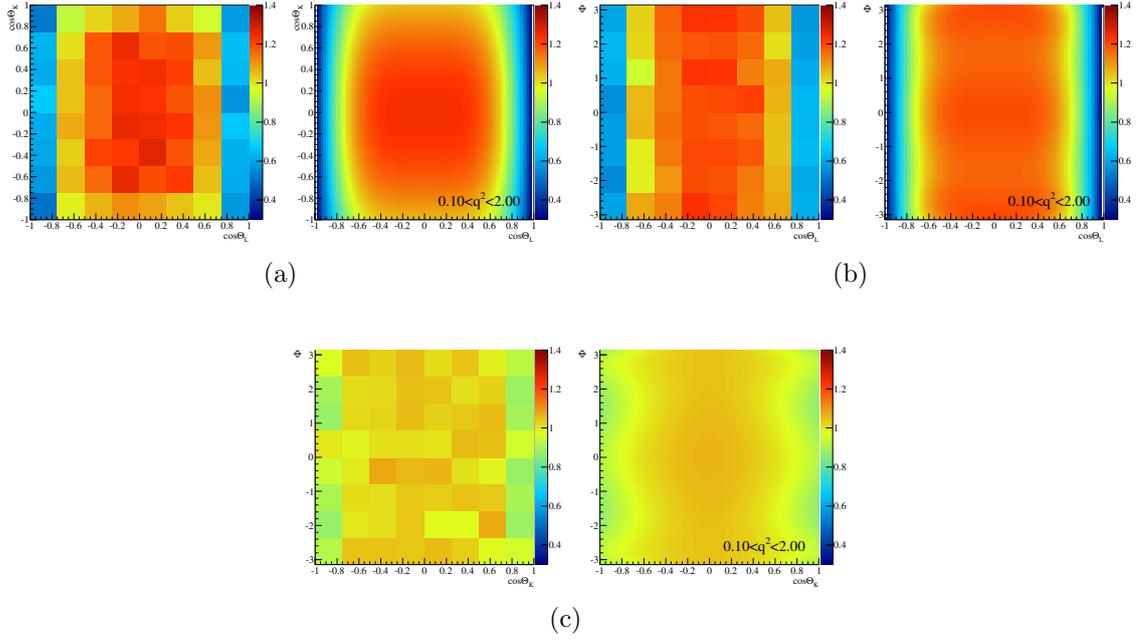


Figure A 2.2: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $0.1 < q^2 < 2.0 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos\theta_l$ vs. $\cos\theta_K$ (a), $\cos\theta_l$ vs. Φ (b), $\cos\theta_K$ vs. Φ (c).

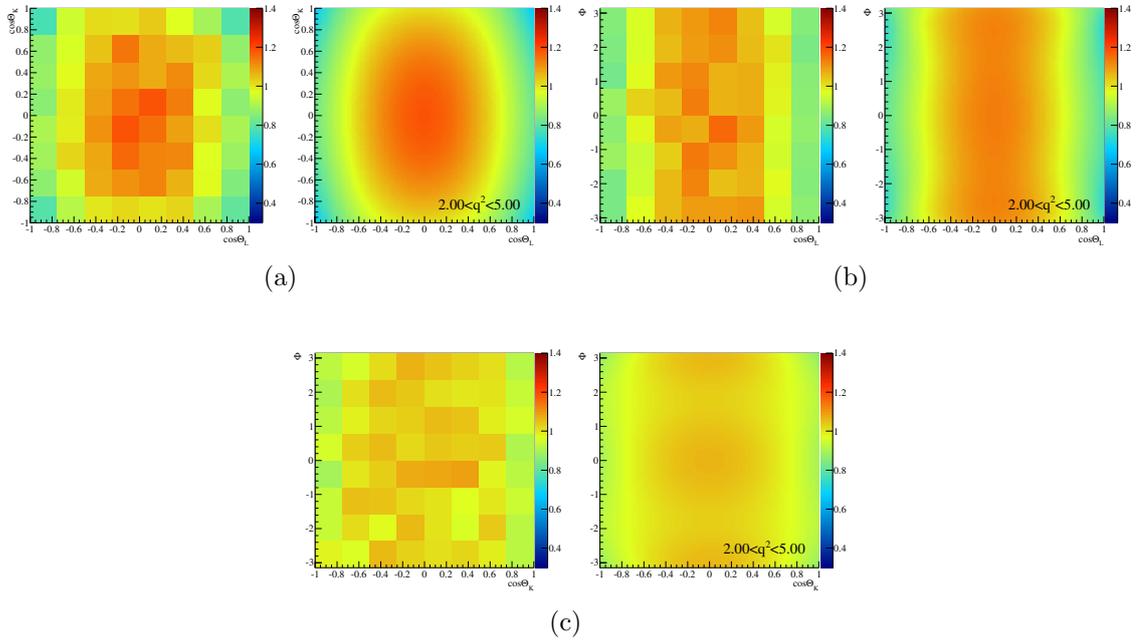


Figure A 2.3: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $2.0 < q^2 < 5.0 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos\theta_l$ vs. $\cos\theta_K$ (a), $\cos\theta_l$ vs. Φ (b), $\cos\theta_K$ vs. Φ (c).

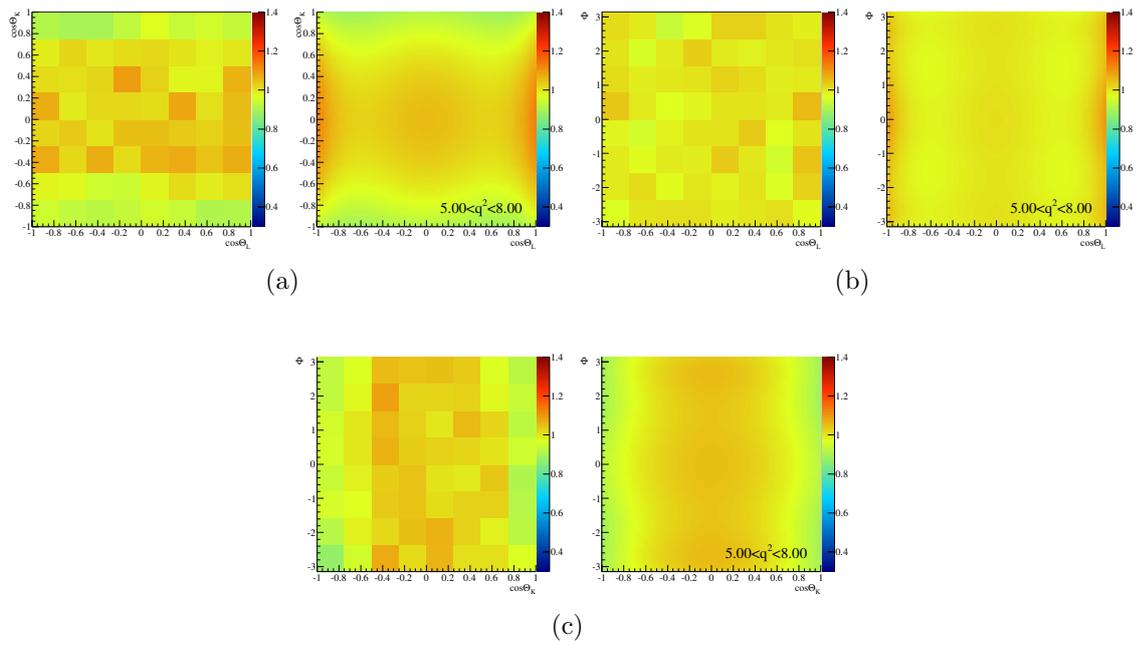


Figure A 2.4: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $5.0 < q^2 < 8.0 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos\theta_l$ vs. $\cos\theta_K$ (a), $\cos\theta_l$ vs. Φ (b), $\cos\theta_K$ vs. Φ (c).

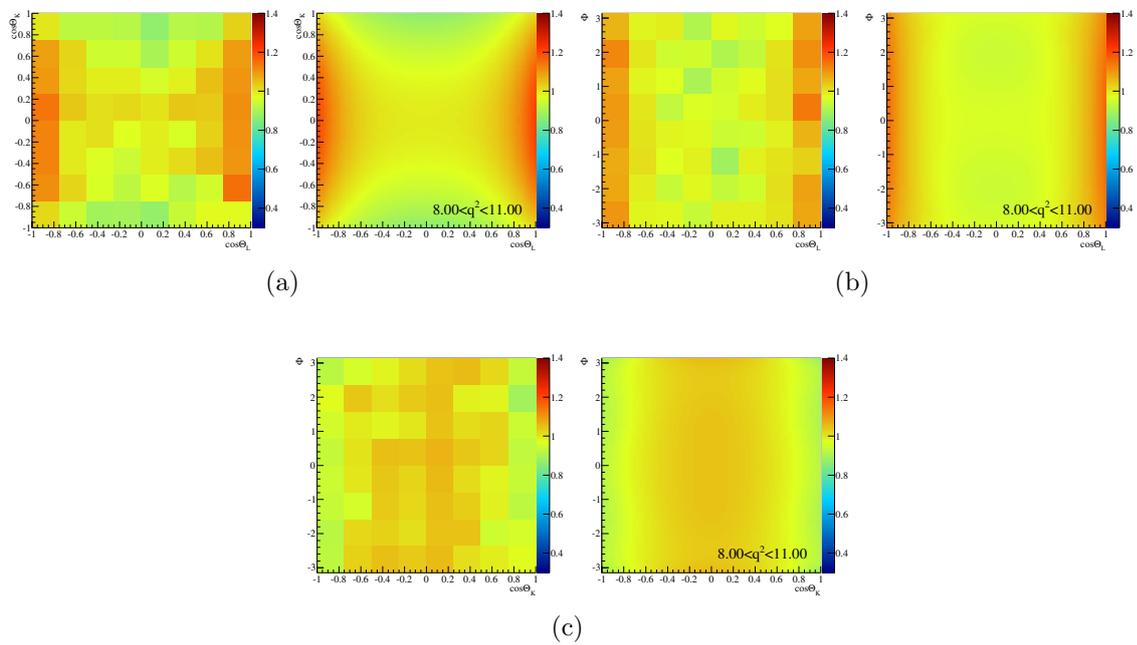


Figure A 2.5: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $11.0 < q^2 < 12.5 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos\theta_l$ vs. $\cos\theta_K$ (a), $\cos\theta_l$ vs. Φ (b), $\cos\theta_K$ vs. Φ (c).

A 2. ANGULAR ACCEPTANCE: TWO-DIMENSIONAL PROJECTIONS

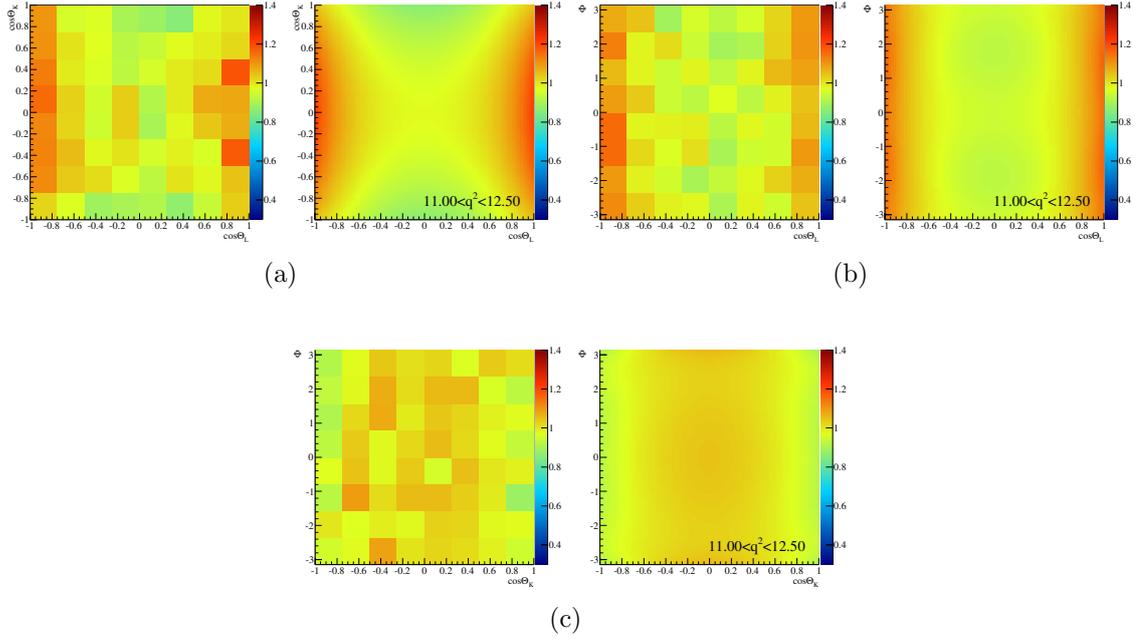


Figure A 2.6: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $15.0 < q^2 < 17.0 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos\theta_l$ vs. $\cos\theta_K$ (a), $\cos\theta_l$ vs. Φ (b), $\cos\theta_K$ vs. Φ (c).

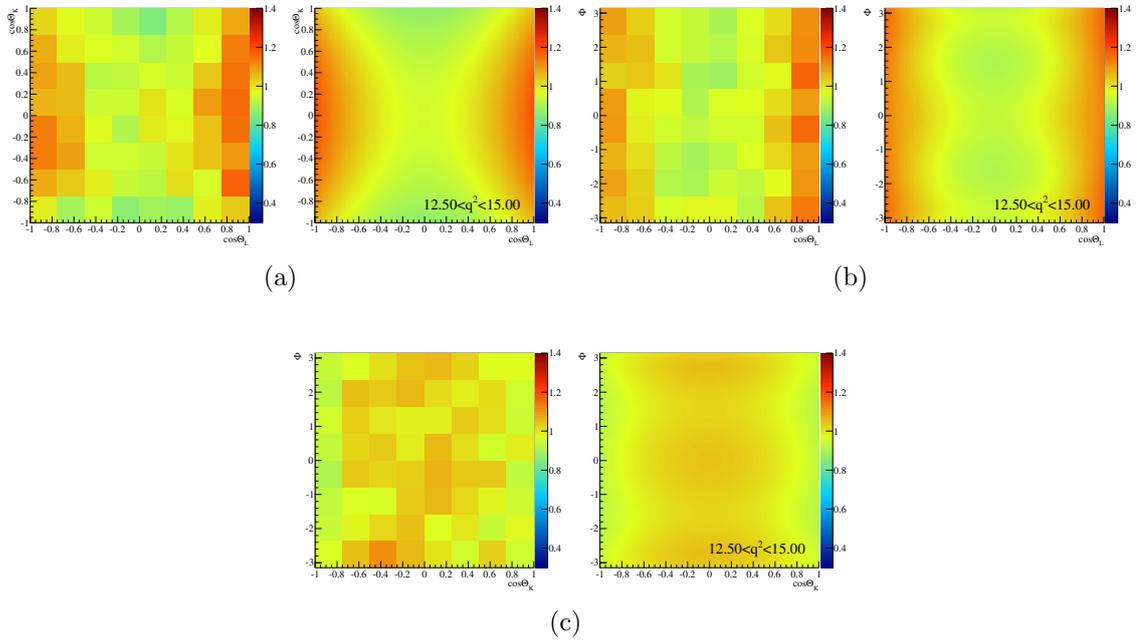


Figure A 2.7: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $17.0 < q^2 < 19.0 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos\theta_l$ vs. $\cos\theta_K$ (a), $\cos\theta_l$ vs. Φ (b), $\cos\theta_K$ vs. Φ (c).

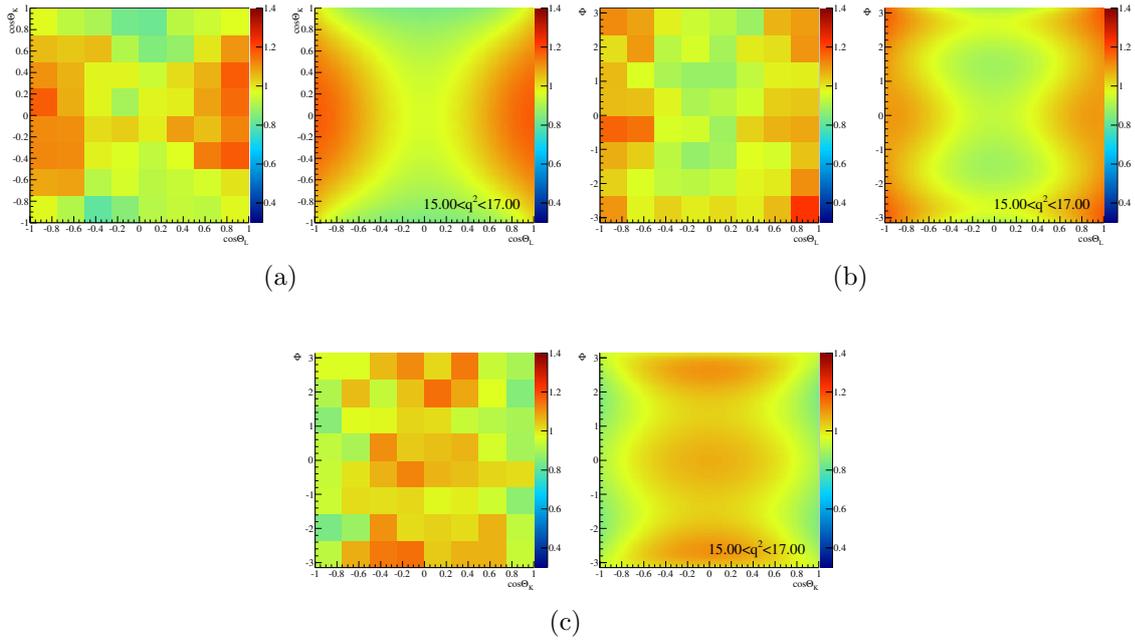


Figure A 2.8: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $1.0 < q^2 < 6.0 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos \theta_l$ vs. $\cos \theta_K$ (a), $\cos \theta_l$ vs. Φ (b), $\cos \theta_K$ vs. Φ (c).

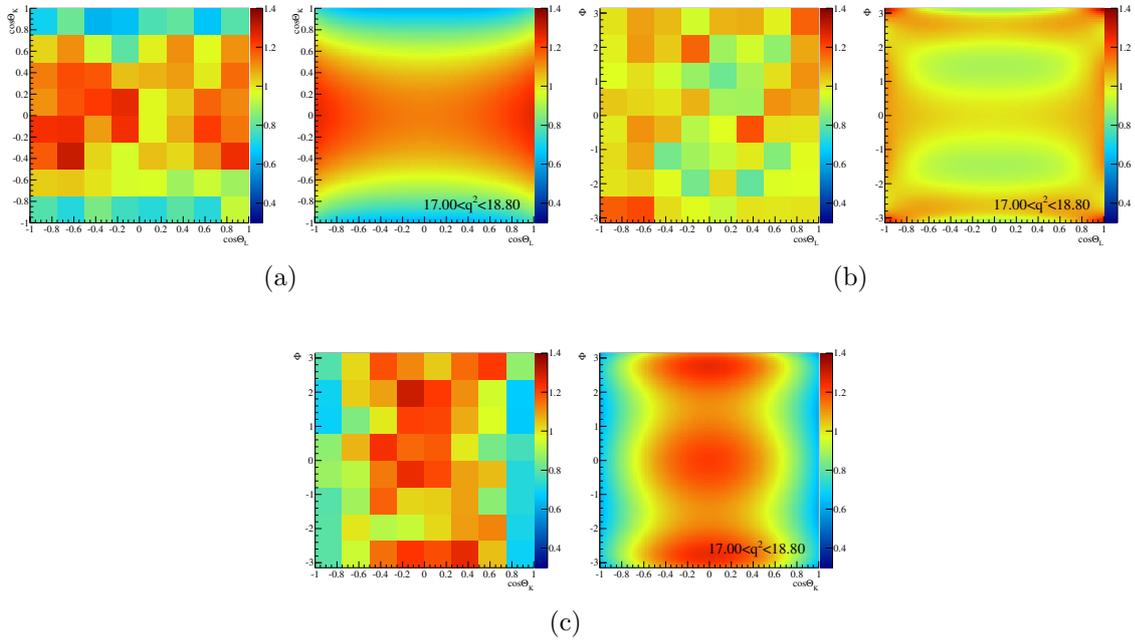


Figure A 2.9: Two-dimensional projections of the angular acceptance overlaid by the chosen parametrisation for $15.0 < q^2 < 19.0 \text{ GeV}^2/c^4$ in the three observable combinations: $\cos \theta_l$ vs. $\cos \theta_K$ (a), $\cos \theta_l$ vs. Φ (b), $\cos \theta_K$ vs. Φ (c).

A 3 Parameter scans

Figures A 3.1 to A 3.5 show the regions for each angular fit parameter combination for which the differential decay rate takes on positive values. They are obtained by varying one of the two parameters and evaluating the allowed region for the respective other parameter shown, while keeping all other parameters on the nominal values as expected from Standard Model predictions. The red squares indicate the SM prediction for the six narrow q^2 bins, numbered from the lowest to the highest q^2 values.

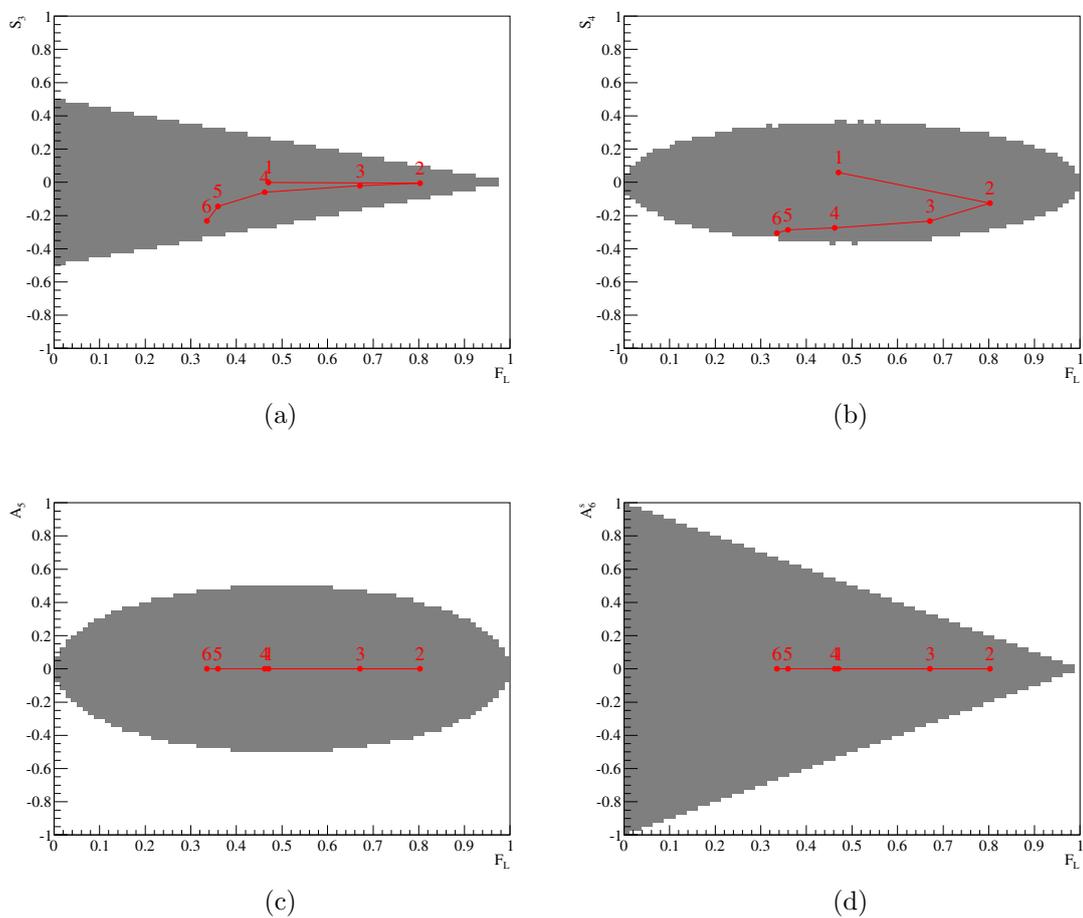


Figure A 3.1: Physically allowed regions for different two-dimensional combinations of angular fit parameters. Red dots indicate the nominal values used in the generation of the simulated event sample for each of the six narrow q^2 bins.

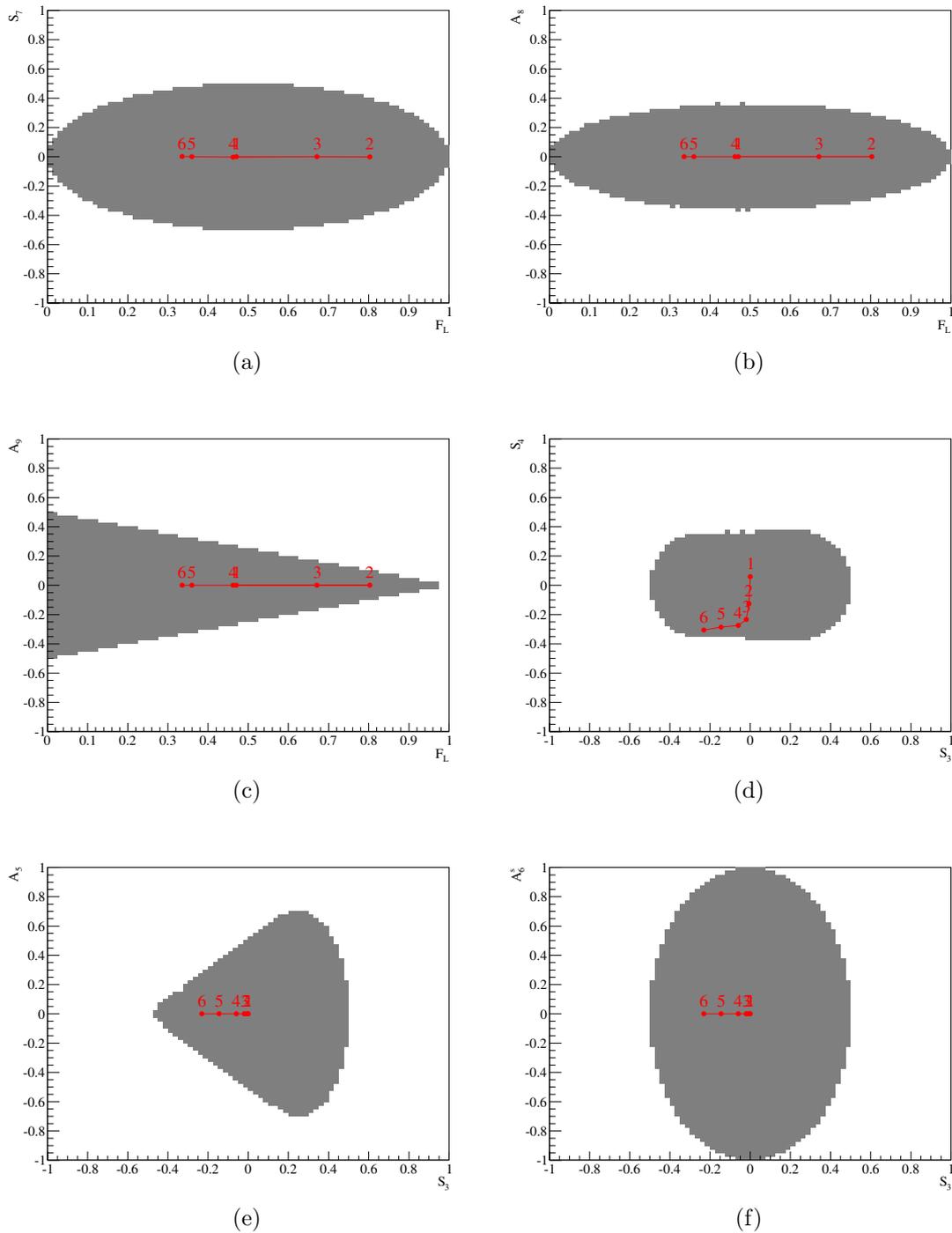


Figure A 3.2: Physically allowed regions for different two-dimensional combinations of angular fit parameters. Red dots indicate the nominal values used in the generation of the simulated event sample for each of the six narrow q^2 bins.

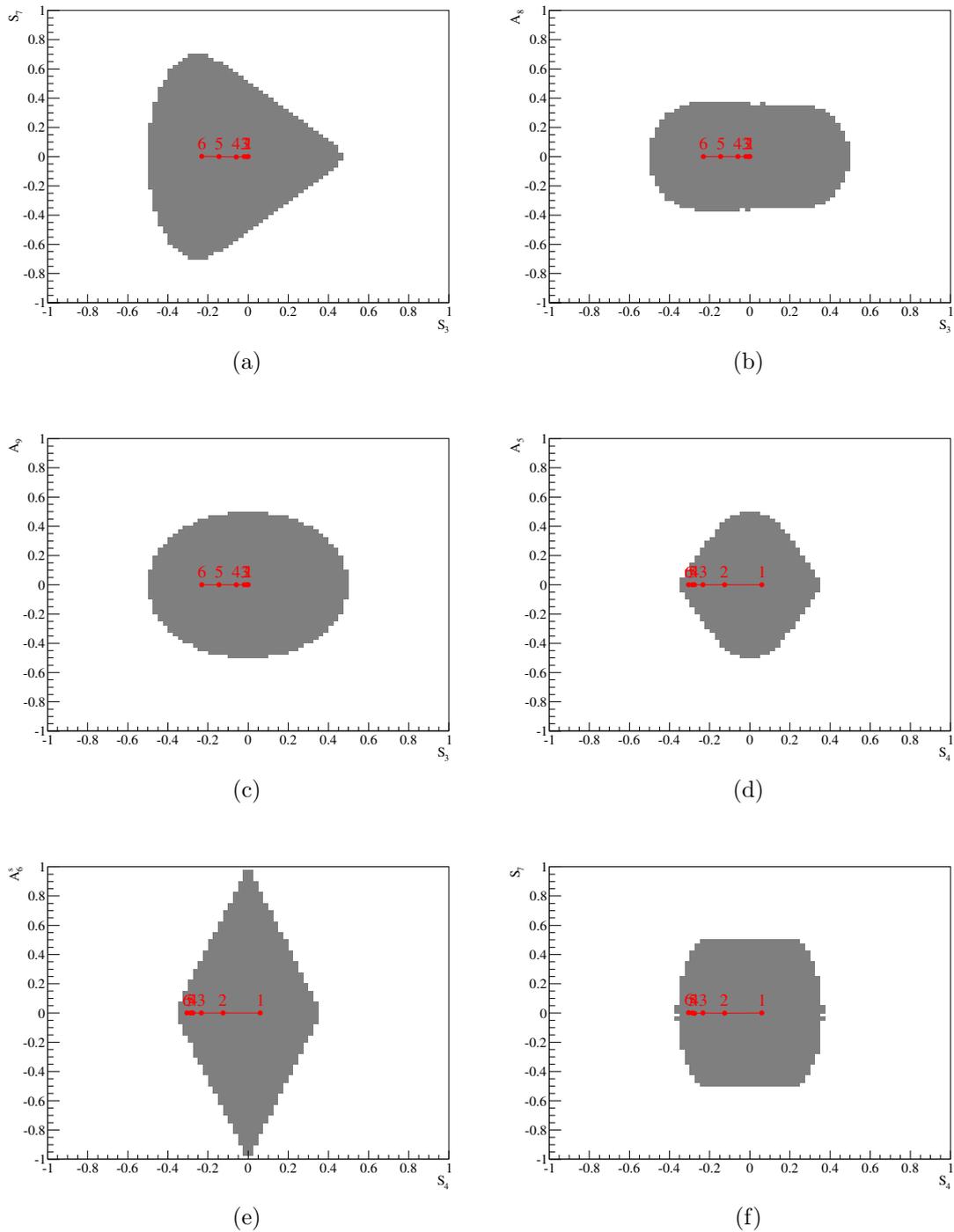


Figure A 3.3: Physically allowed regions for different two-dimensional combinations of angular fit parameters. Red dots indicate the nominal values used in the generation of the simulated event sample for each of the six narrow q^2 bins.

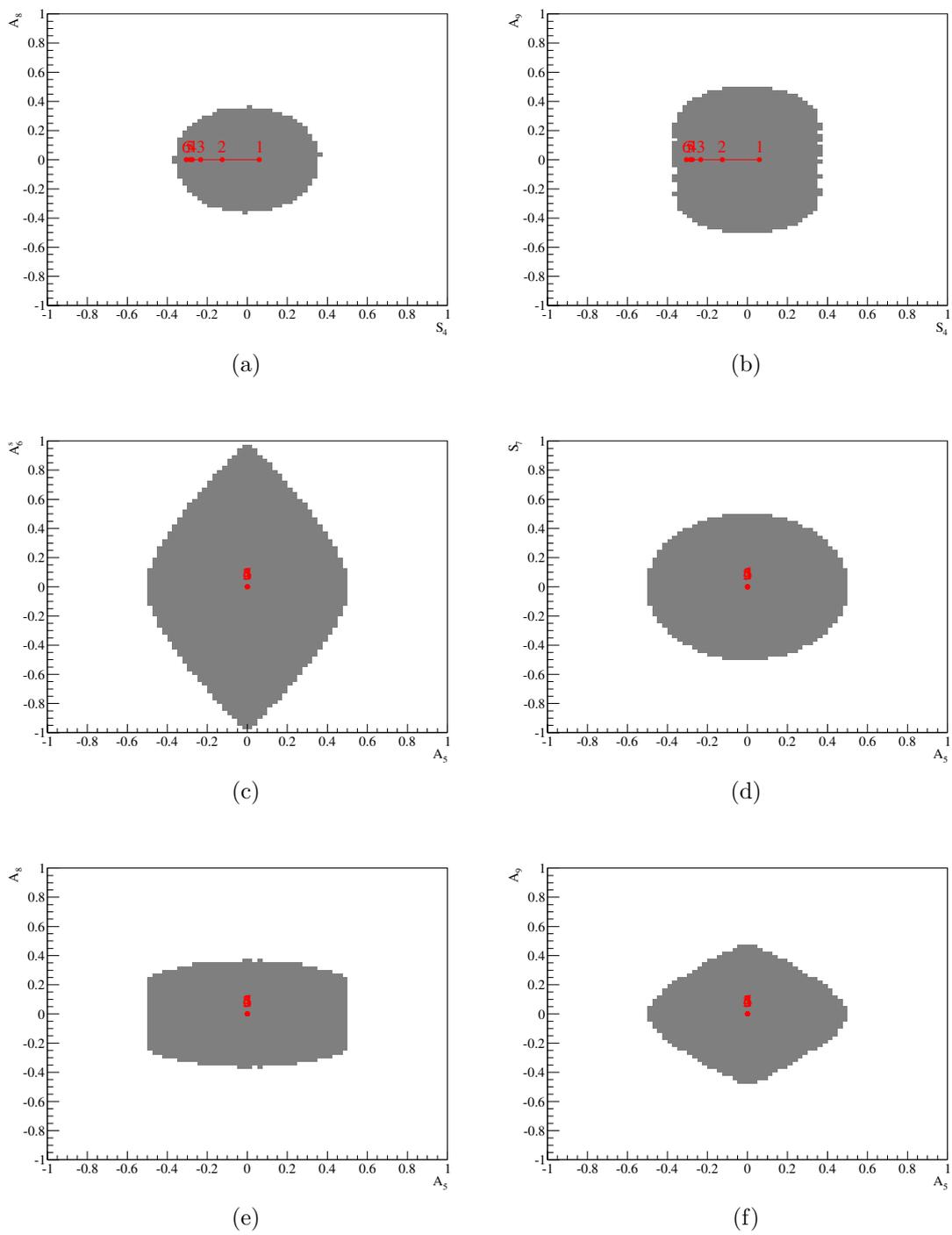


Figure A 3.4: Physically allowed regions for different two-dimensional combinations of angular fit parameters. Red dots indicate the nominal values used in the generation of the simulated event sample for each of the six narrow q^2 bins.

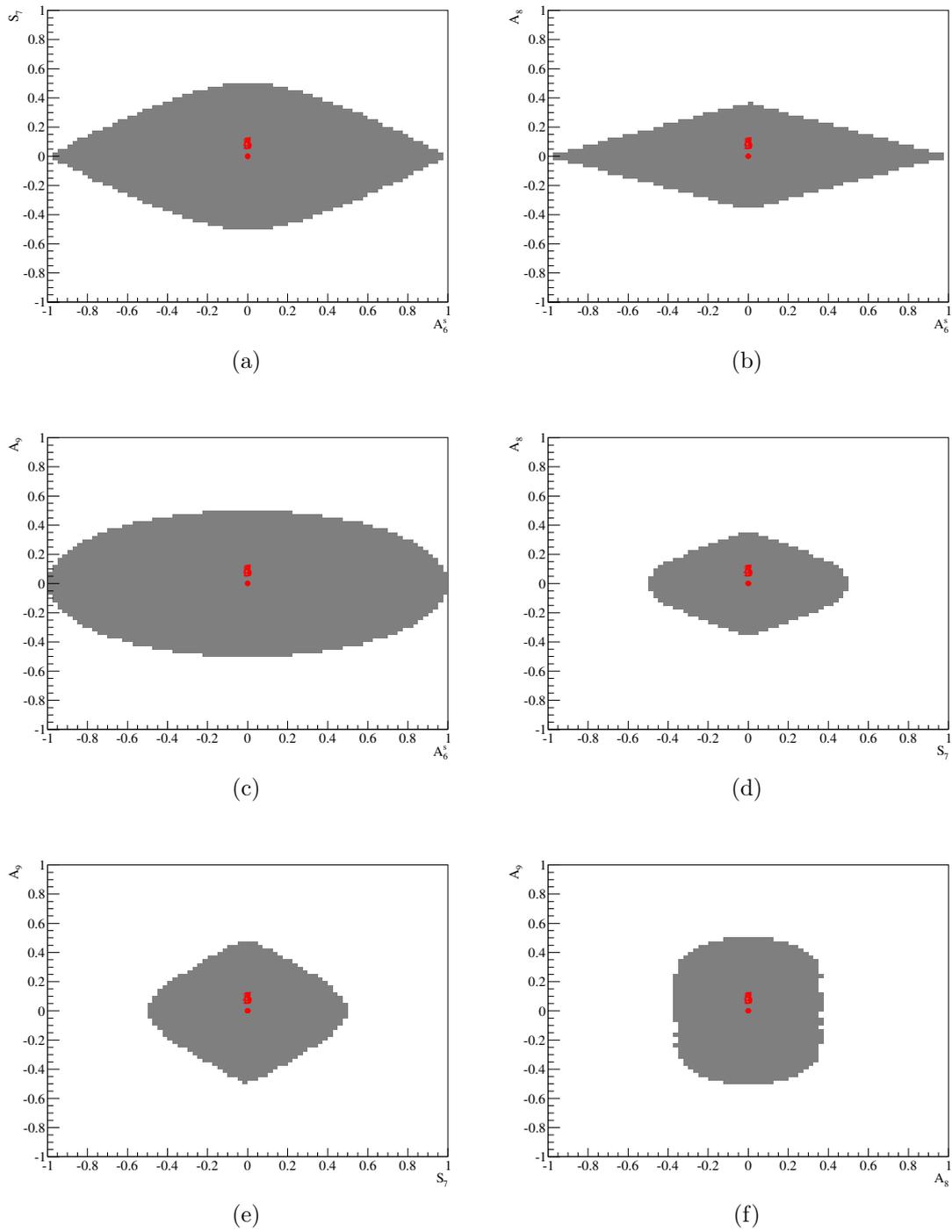


Figure A 3.5: Physically allowed regions for different two-dimensional combinations of angular fit parameters. Red dots indicate the nominal values used in the generation of the simulated event sample for each of the six narrow q^2 bins.

A 4 Feldman-Cousins scans

Figures A 4.1 to A 4.8 show the profile likelihood scans using the Feldman-Cousins method for each angular fit parameter in bins of q^2 . Intervals indicating a confidence level of 69% are marked by red horizontal lines. The blue horizontal lines indicate the 1σ confidence intervals as returned by the default fit.

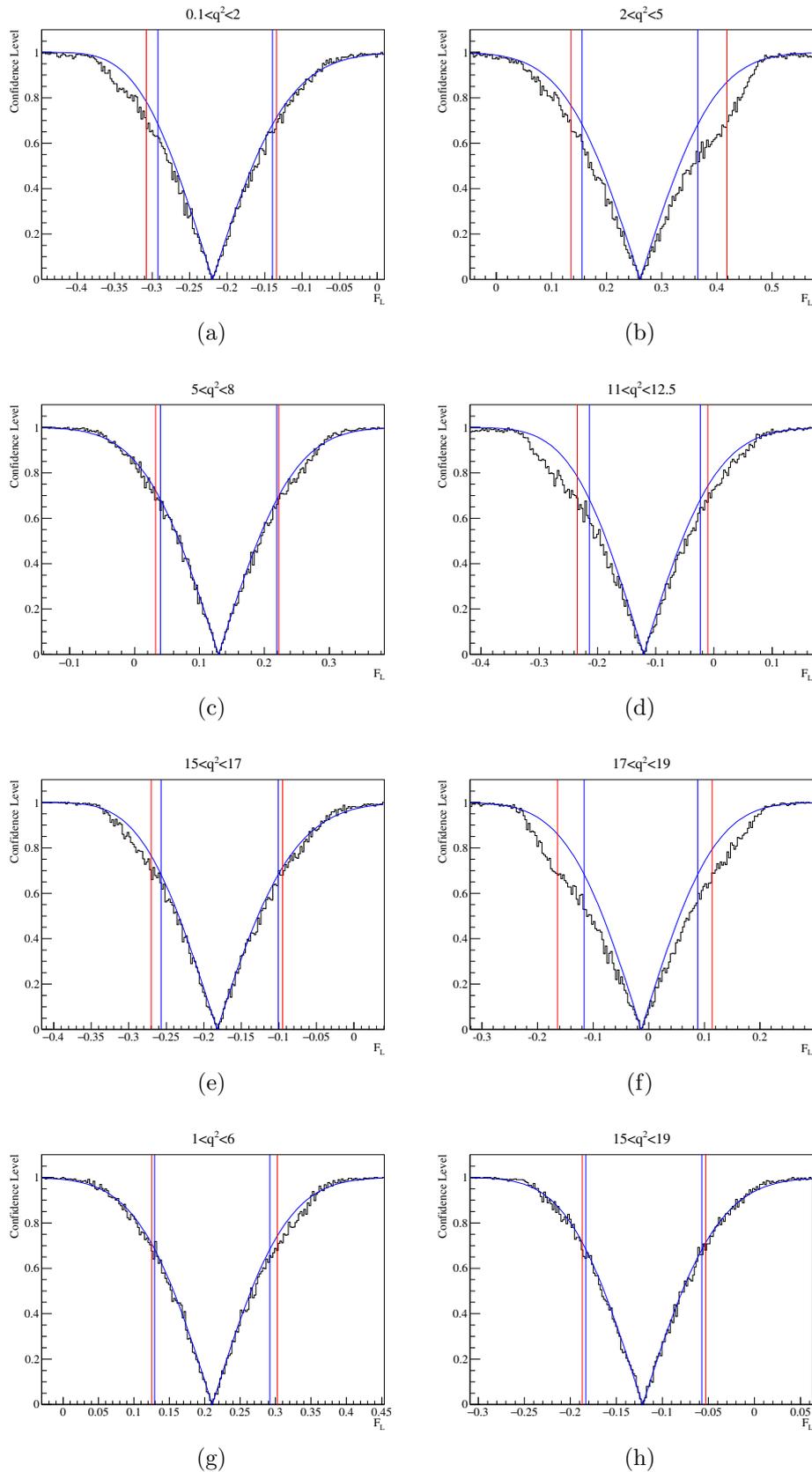


Figure A 4.1: Feldman-Cousins profile likelihood scans for the angular fit parameter F_L . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

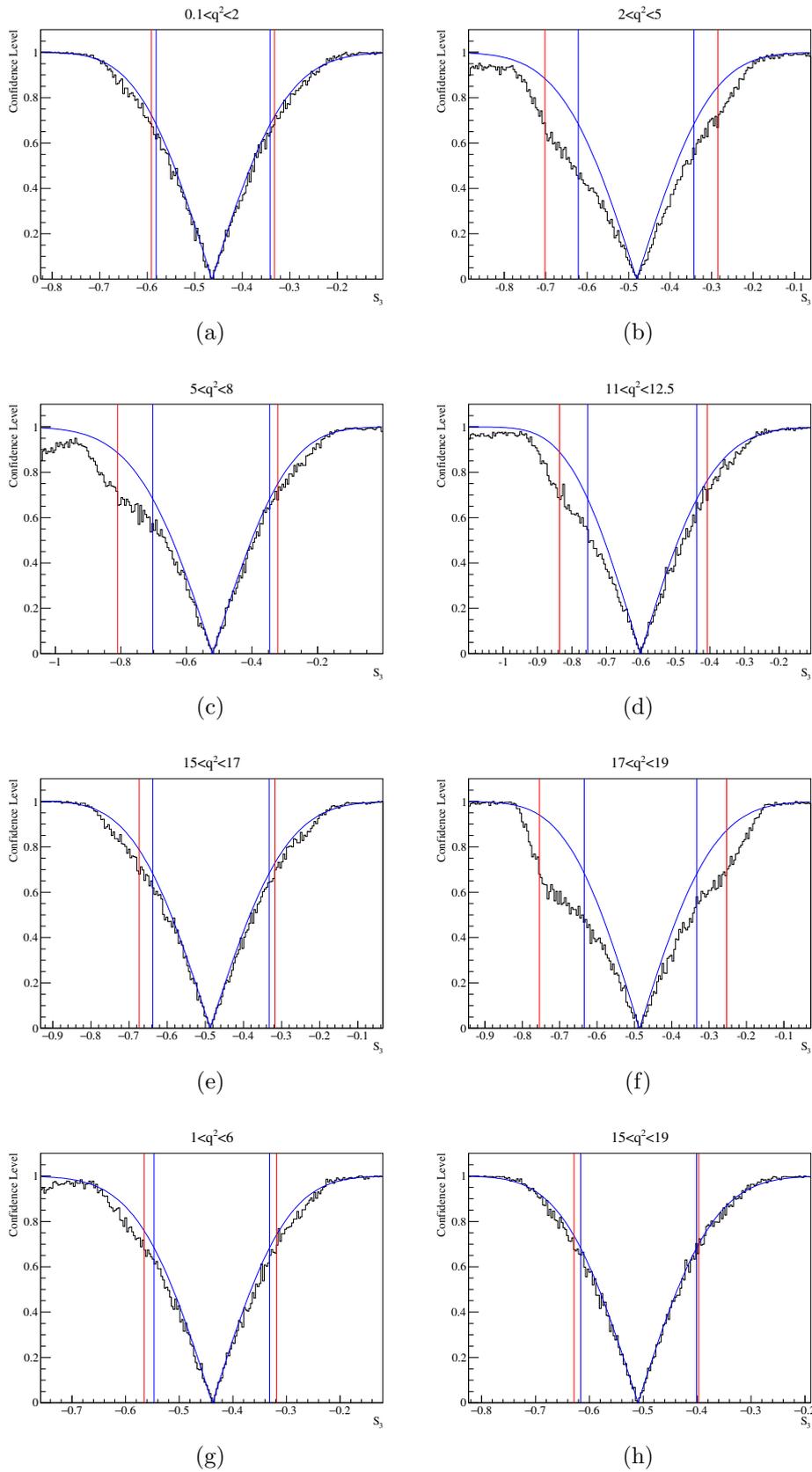


Figure A 4.2: Feldman-Cousins profile likelihood scans for the angular fit parameter S_3 . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

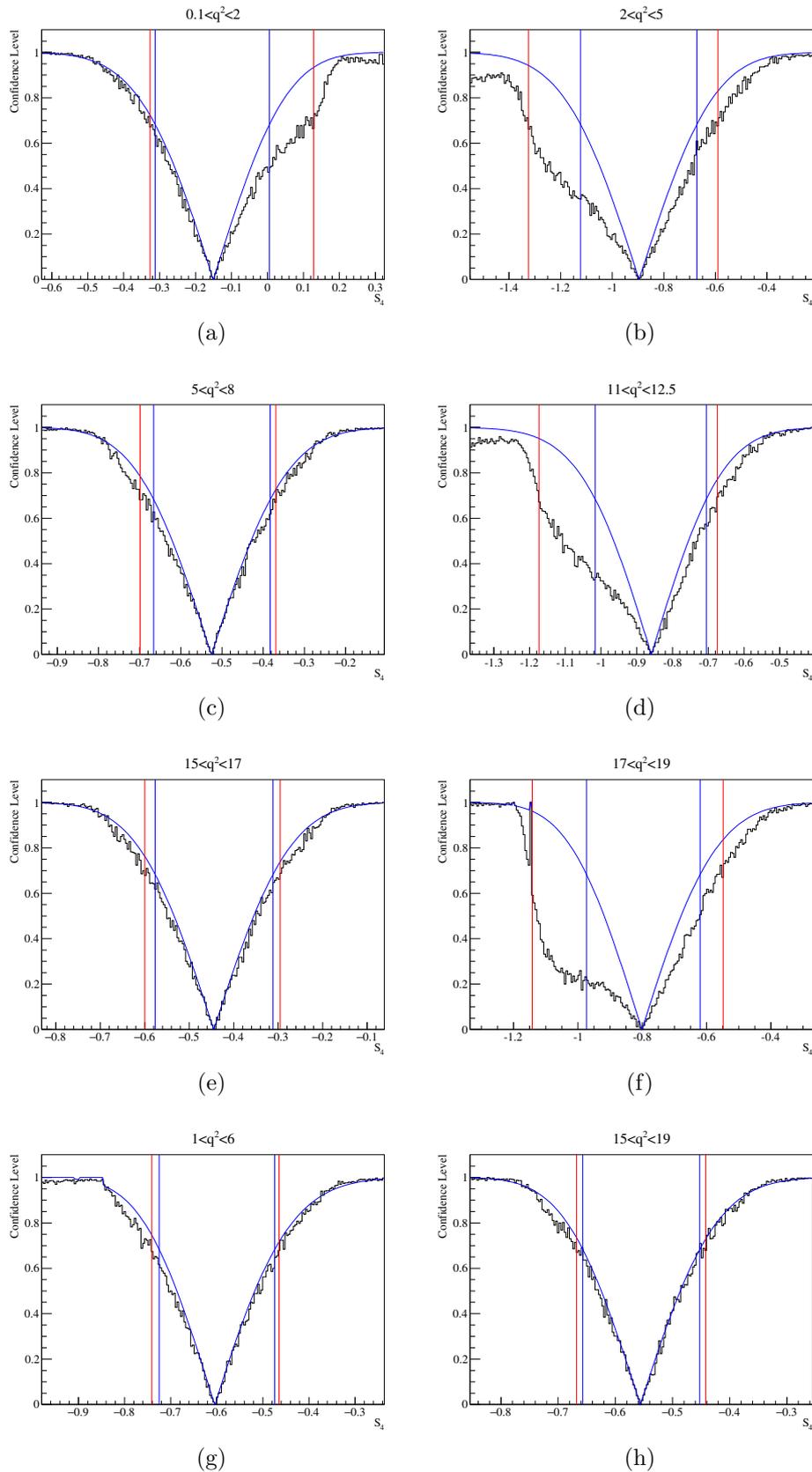


Figure A 4.3: Feldman-Cousins profile likelihood scans for the angular fit parameter S_4 . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

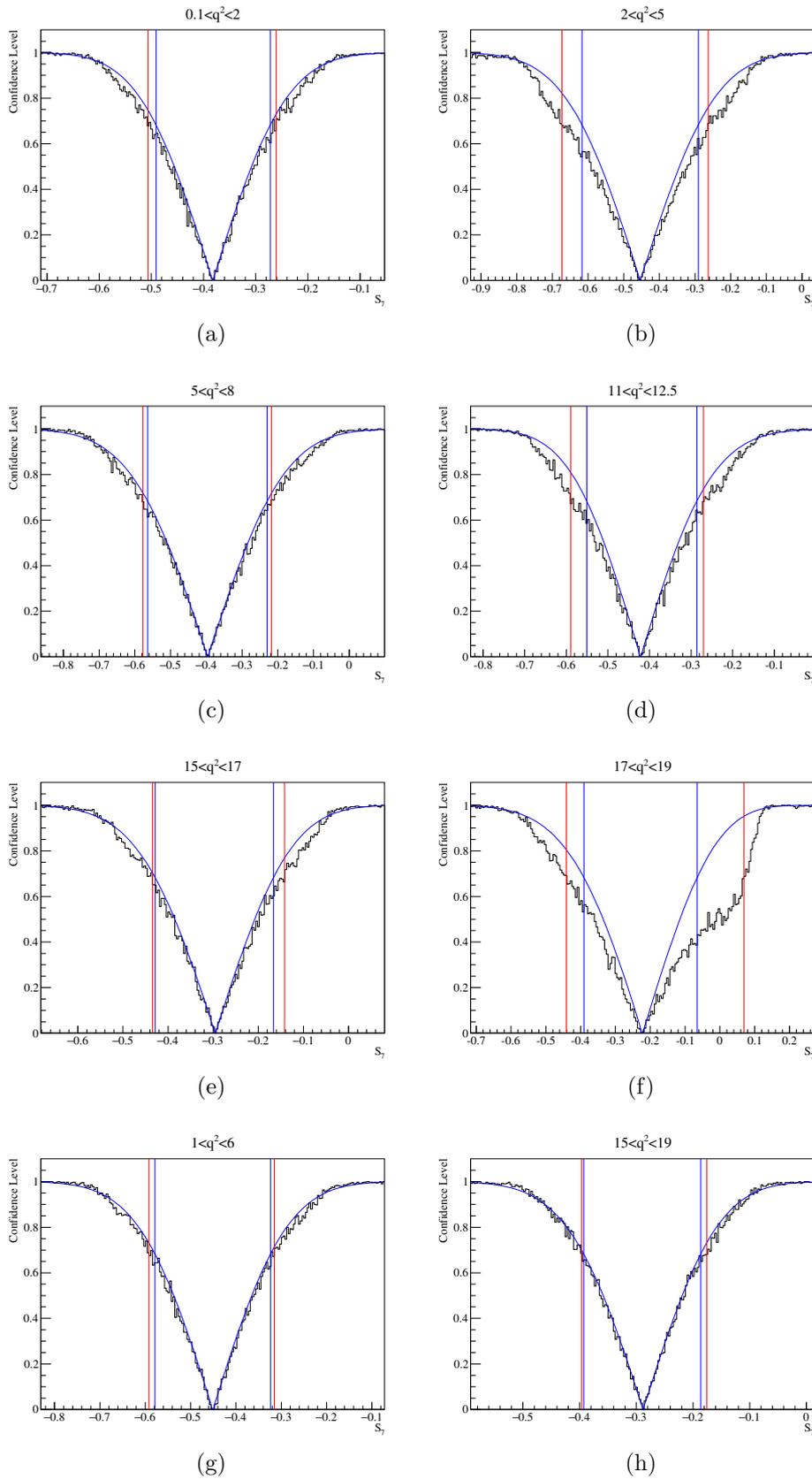


Figure A.4.4: Feldman-Cousins profile likelihood scans for the angular fit parameter S_7 . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

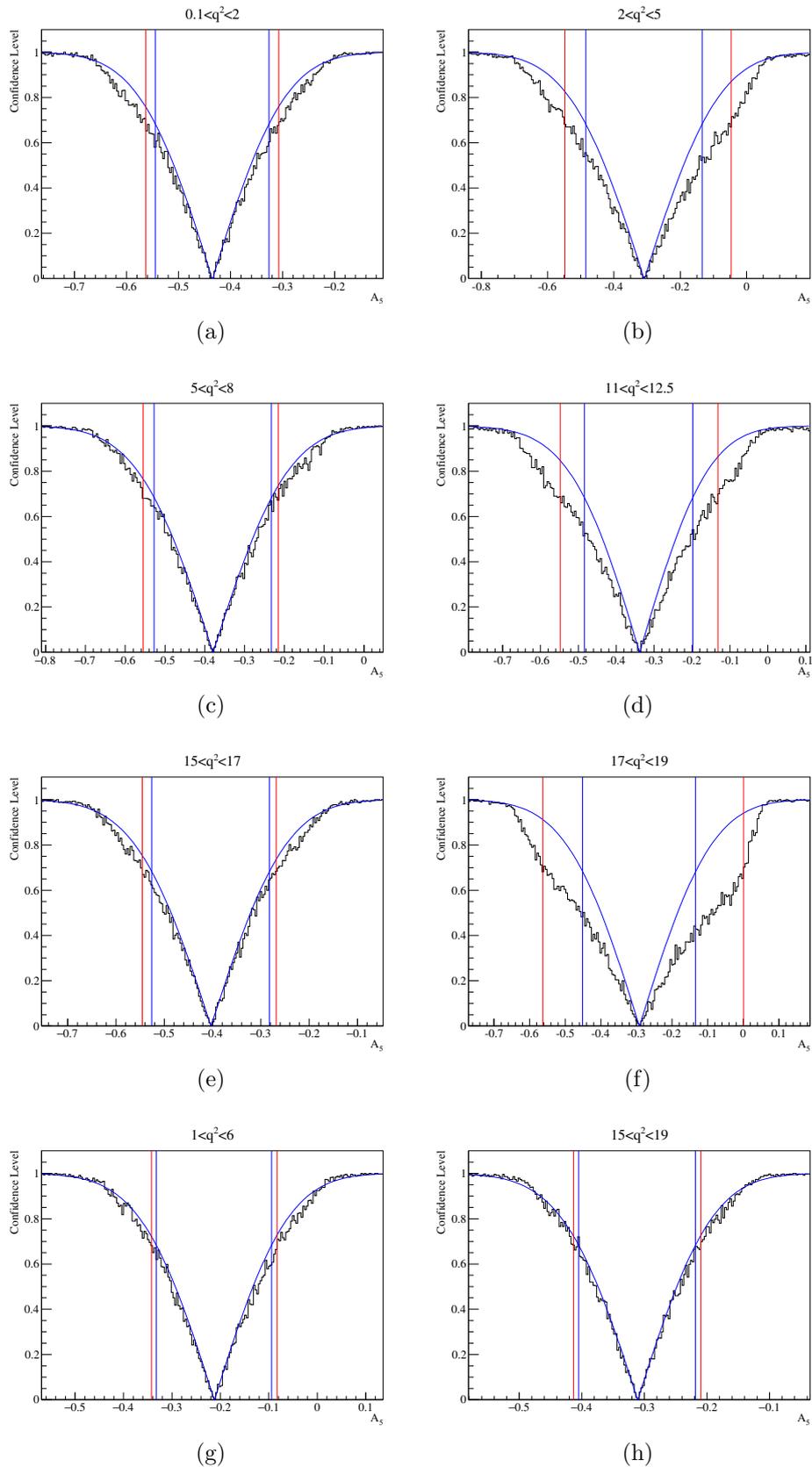


Figure A 4.5: Feldman-Cousins profile likelihood scans for the angular fit parameter A_5 . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

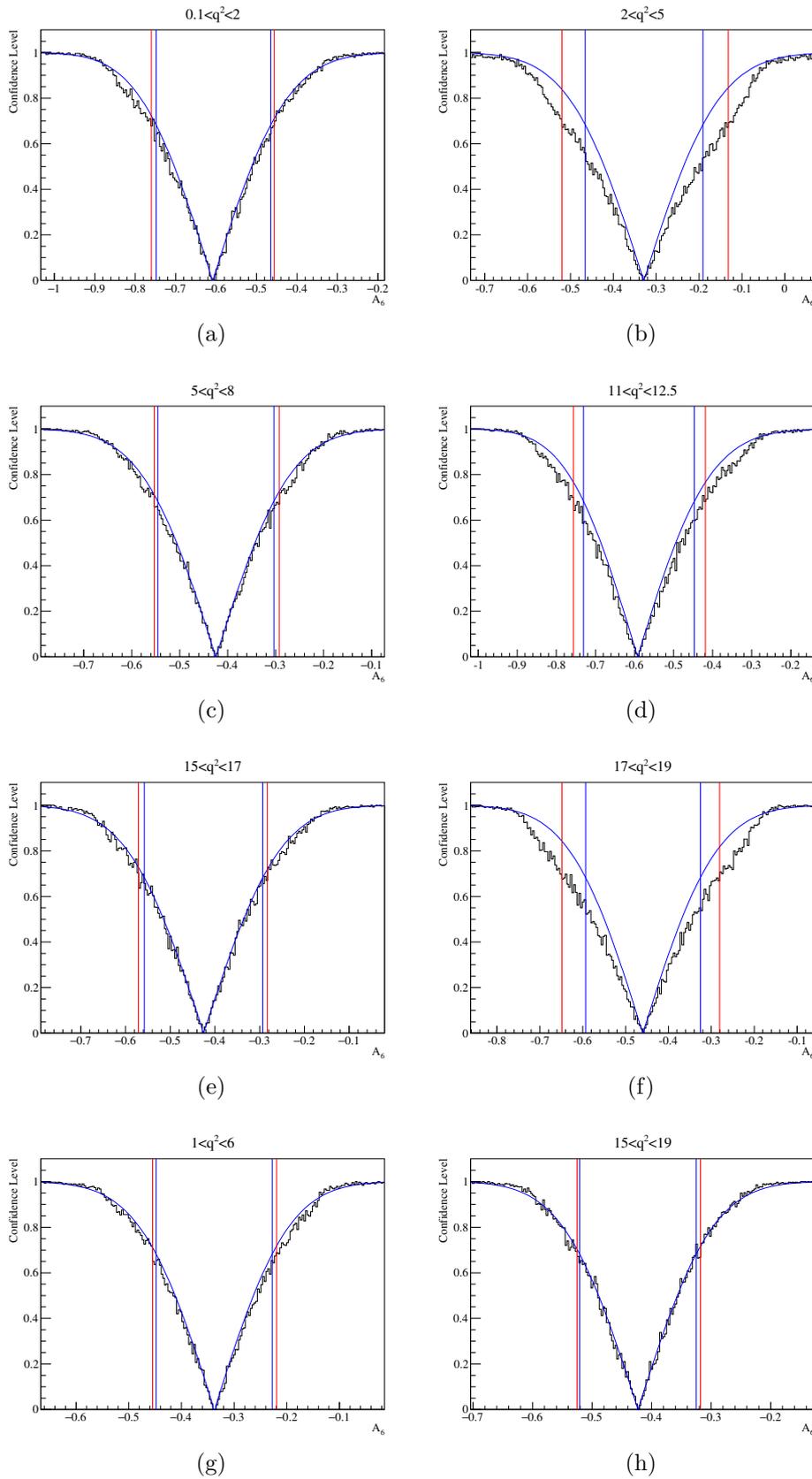


Figure A 4.6: Feldman-Cousins profile likelihood scans for the angular fit parameter A_6 . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

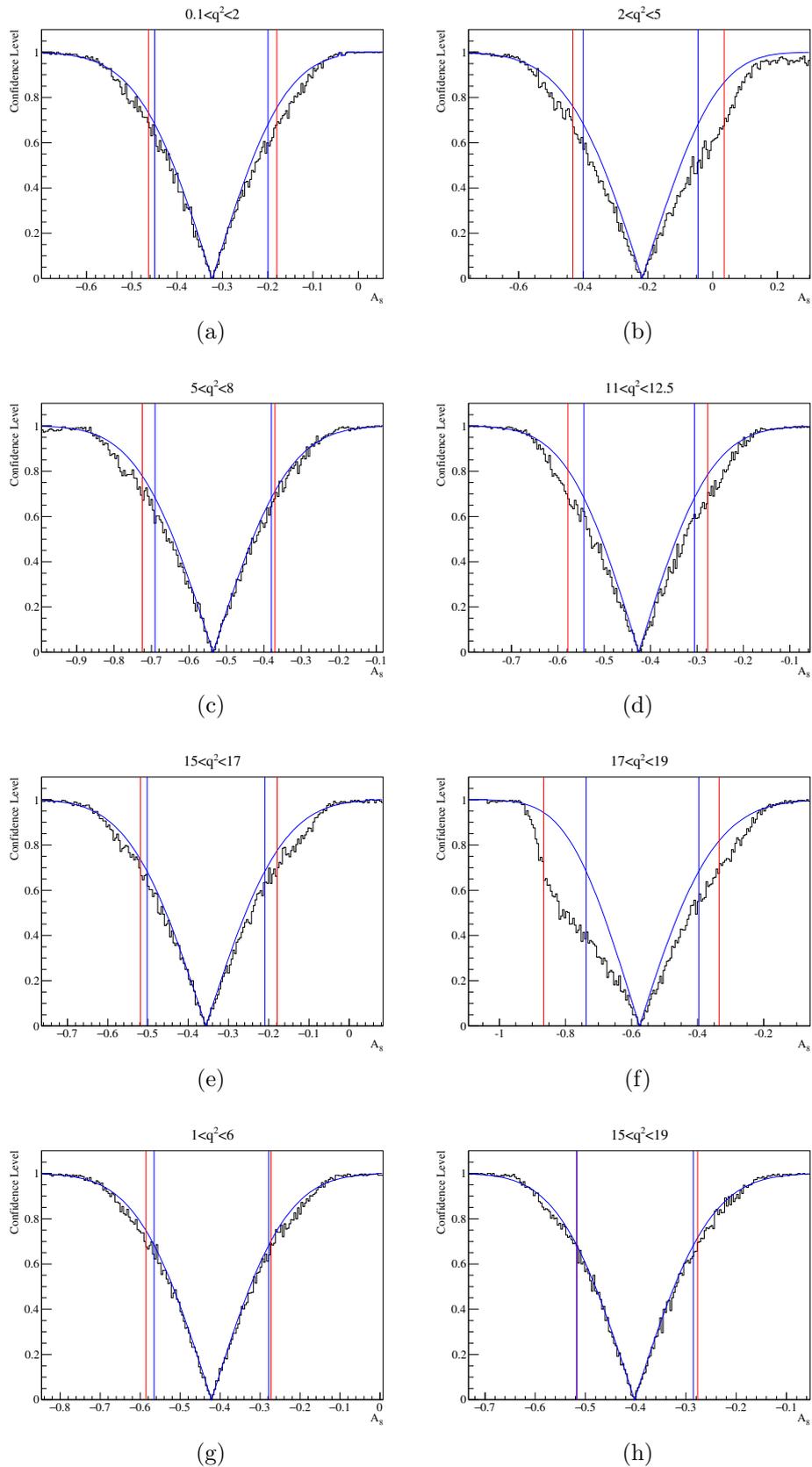


Figure A 4.7: Feldman-Cousins profile likelihood scans for the angular fit parameter A_8 . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

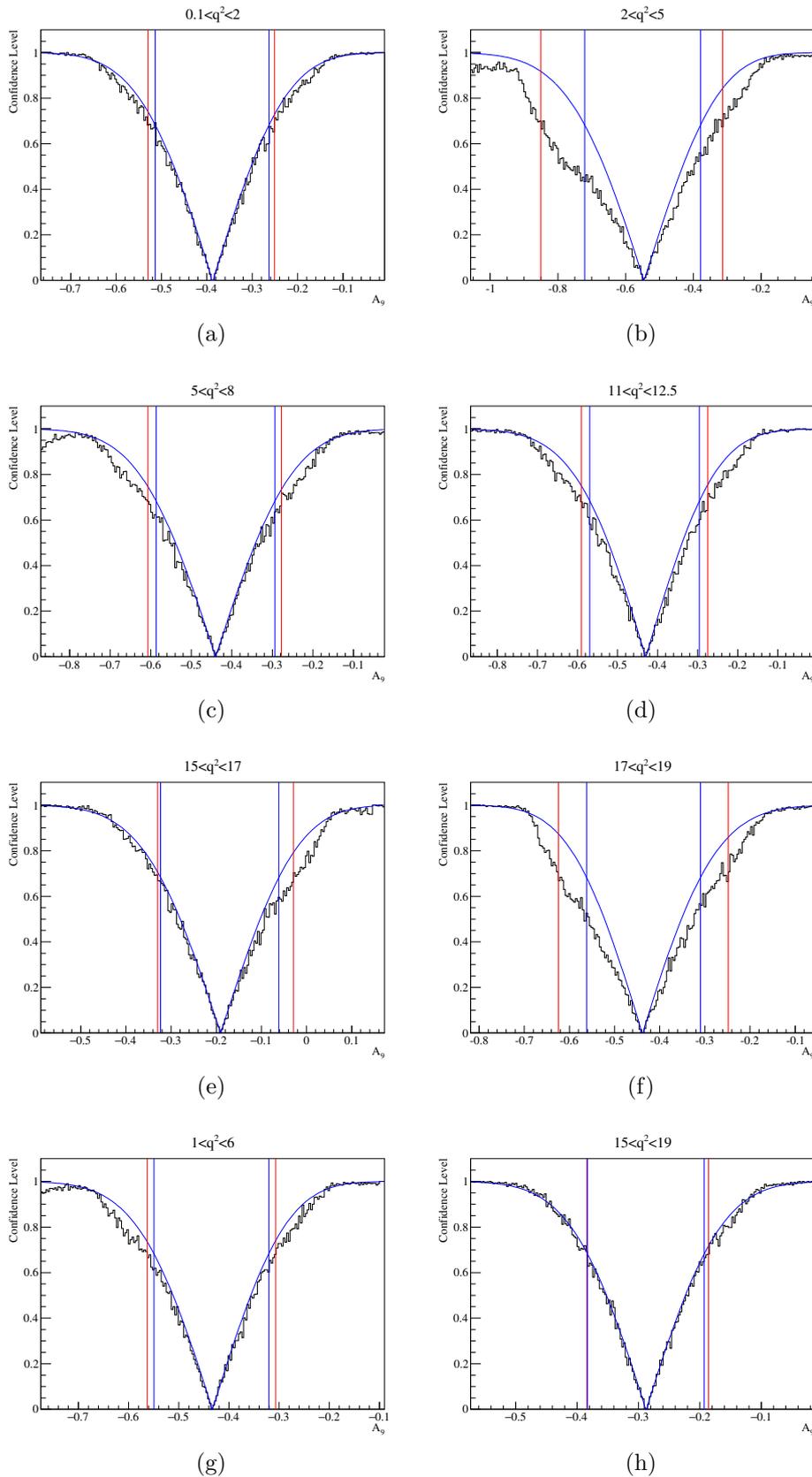


Figure A 4.8: Feldman-Cousins profile likelihood scans for the angular fit parameter A_9 . Horizontal blue lines indicate the error bands returned by the default fit, horizontal red lines the 69% confidence levels of the likelihood scan.

Bibliography

-
- [1] C. Langenbruch and U. Uwer, *Measurement of the B_s^0 mixing phase in the decay $B_s^0 \rightarrow J/\psi\phi$ with the LHCb experiment*, PhD thesis, Heidelberg U., Heidelberg, 2011, Presented 09 Nov 2011.
- [2] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$* , JHEP **09** (2015) 179 LHCb-PAPER-2015-023, CERN-PH-EP-2015-145, arXiv:1506.0877.
- [3] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.
- [4] http://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html.
- [5] T. A. 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) 1-29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- [6] L. V. Group *et al.*, *Performance of the lhcb vertex locator*, 1405.7808.
- [7] T. L. collaboration, *Measurement of the track reconstruction efficiency at lhcb*, Journal of Instrumentation **10** (2015), no. 02 P02007.
- [8] LHCb Outer Tracker group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, J. Instrum. **9** (2013) P01002. 30 p, Comments: 30 pages, 20 figures.
- [9] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.
- [10] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.
- [11] T. Sjstrand, S. Mrenna, and P. Skands, *A brief introduction to pythia 8.1*, 0710.3820.
- [12] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [13] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [14] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.
- [15] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

-
- [16] J. Van Tilburg and M. Merk, *Track simulation and reconstruction in LHCb*, 2005. Presented on 01 Sep 2005.
- [17] M. De Cian, U. Straumann, O. Steinkamp, and N. Serra, *Track Reconstruction Efficiency and Analysis of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ at the LHCb Experiment*, Sep, 2013. Presented 14 Mar 2013.
- [18] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection*, Tech. Rep. LHCb-PUB-2017-011. CERN-LHCb-PUB-2017-011, CERN, Geneva, Mar, 2017.
- [19] R. Aaij *et al.*, *Tesla : an application for real-time data analysis in high energy physics*, 1604.5596.
- [20] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
- [21] L. collaboration *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 tev pp collisions*, 1612.5140.
- [22] LHCb collaboration, R. Aaij *et al.*, *Measurement of form-factor-independent observables in the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$* , Phys. Rev. Lett. **111** (2013) 191801 LHCb-PAPER-2013-037, CERN-PH-EP-2013-146, arXiv:1308.1707.
- [23] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the $B^+ \rightarrow K^+\mu^+\mu^-$ decay*, JHEP **02** (2013) 105 LHCb-PAPER-2012-024, CERN-PH-EP-2012-263, arXiv:1209.4284.
- [24] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$* , JHEP **07** (2013) 084 CERN-PH-EP-2013-078, LHCb-PAPER-2013-017, arXiv:1305.2168.
- [25] LHCb collaboration, R. Aaij *et al.*, *Measurement of the differential branching fraction of the decay $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$* , Phys. Lett. **B725** (2013) 25 LHCb-PAPER-2013-025, CERN-PH-EP-2013-094, arXiv:1306.2577.
- [26] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \rightarrow K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601 LHCb-PAPER-2014-024, CERN-PH-EP-2014-140, arXiv:1406.6482.
- [27] W. Altmannshofer and D. M. Straub, *New physics in $b \rightarrow s$ transitions after lhc run 1*, 1411.3161.
- [28] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \rightarrow V\ell^+\ell^-$ in the standard model from light-cone sum rules*, 1503.5534.

- [29] G. Altarelli, *The standard electroweak theory and beyond*, 2000.
- [30] G. Altarelli, *The standard model of particle physics*, 2005.
- [31] http://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg.
- [32] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.
- [33] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.
- [34] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [35] R. Z. Funchal, B. Schmauch, and G. Giesen, *The physics of neutrinos*, 2013.
- [36] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652.
- [37] A. D. Sakharov, *SPECIAL ISSUE: Violation of CP in variance, C asymmetry, and baryon asymmetry of the universe*, Soviet Physics Uspekhi **34** (1991) 392.
- [38] H. R. Quinn and P. F. Harrison, *The BaBar physics book: physics at an asymmetric B factory*, SLAC, Stanford, CA, 1998.
- [39] M. Grabalosa and M. Musy, *Flavour Tagging developments within the LHCb experiment*, PhD thesis, Barcelona U., Mar, 2012, Presented 15 May 2012.
- [40] J. Ellis, M. Gaillard, D. Nanopoulos, and S. Rudaz, *The phenomenology of the next left-handed quarks*, Nuclear Physics B **131** (1977), no. 2 285 .
- [41] W. Altmannshofer *et al.*, *Symmetries and asymmetries of $b \rightarrow k^* \mu^+ \mu^-$ decays in the standard model and beyond*, 0811.1214.
- [42] T. Blake, T. Gershon, and G. Hiller, *Rare b hadron decays at the lhc*, 1501.3309.
- [43] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, hep-ph/95.
- [44] C. Bobeth, M. Misiak, and J. Urban, *Photonic penguins at two loops and m_t -dependence of $\mathcal{B}[B^- \rightarrow X_s l^+ l^-]$* , hep-ph/99.
- [45] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Calculation of $b^0 \rightarrow k^{*0+-}$ and $b_s^0 \rightarrow ^{+-}$ observables using form factors from lattice qcd*, 1310.3887.
- [46] M. Beneke, T. Feldmann, and D. Seidel, *Systematic approach to exclusive $B \rightarrow V l^+ l^-$, $V \gamma$ decays*, hep-ph/01.

-
- [47] M. Beylich, G. Buchalla, and T. Feldmann, *Theory of $B \rightarrow K^* l^+ l^-$ decays at high q^2 : OPE and quark-hadron duality*, 1101.5118.
- [48] W. Altmannshofer and D. M. Straub, *New physics in $b \rightarrow k^* ?$* , 1308.1501.
- [49] G. Hiller and M. Schmaltz, *r_k and future $b \rightarrow s l l$ bsm opportunities*, 1408.1627.
- [50] A. Hoecker *et al.*, *Tmva - toolkit for multivariate data analysis*, physics/0.
- [51] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **543** (2005), no. 23 577 .
- [52] A. Puig, *The LHCb trigger in 2011 and 2012*, Tech. Rep. LHCb-PUB-2014-046. CERN-LHCb-PUB-2014-046, CERN, Geneva, Nov, 2014.
- [53] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- [54] CDF Collaboration, T. Aaltonen *et al.*, *Observation of the baryonic flavor-changing neutral current decay $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$* , Phys. Rev. Lett. **107** (2011) 201802.
- [55] P. Ball and R. Zwicky, *$B_{d,s} \rightarrow \rho, \omega, K^*, \phi$* , Phys. Rev. D **71** (2005) 014029.
- [56] L. Anderlini *et al.*, *The PIDCalib package*, Tech. Rep. LHCb-PUB-2016-021. CERN-LHCb-PUB-2016-021, CERN, Geneva, Jul, 2016.
- [57] S. Kotz and N. L. Johnson, *Encyclopedia of Statistical Sciences Volume 5*, Wiley, New York, NY, USA, 1985.
- [58] R. Barlow, *Extended maximum likelihood*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **297** (1990), no. 3 496 .
- [59] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays*, Phys. Rev. Lett. **114** (2015) 041801 LHCb-PAPER-2014-059, CERN-PH-EP-2014-271, arXiv:1411.3104.
- [60] G. J. Feldman and R. D. Cousins, *A unified approach to the classical statistical analysis of small signals*, physics/9.
- [61] L. Mott and W. Roberts, *Rare dileptonic decays of Λ_b in a quark model*, 1108.6129.
- [62] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , JHEP **08** (2013) 131 CERN-PH-EP-2013-074, LHCb-PAPER-2013-019, arXiv:1304.6325.

- [63] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $b \rightarrow K^* \mu^+ \mu^-$ anomaly*, Phys. Rev. D **88** (2013) 074002.
- [64] K. D. Bruyn *et al.*, *Branching ratio measurements of B_s decays*, 1204.1735.
- [65] C. Bobeth, G. Hiller, and G. Piranishvili, *C_p asymmetries in $bar{b} \rightarrow bar{k} \pi^* l+l-$ and untagged $bar{B}_s, B_s \rightarrow \phi(- \rightarrow K+K-)l+l-$ decays at nlo*, 0805.2525.
- [66] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis and branching fraction measurement of $\bar{B}_s^0 \rightarrow J/\psi K^+ K^-$* , Phys. Rev. **D87** (2013) 072004 LHCb-PAPER-2012-040, CERN-PH-EP-2013-008, arXiv:1302.1213.
- [67] LHCb collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, LHCb-CONF-2013-011.
- [68] LHCb collaboration, R. Aaij *et al.*, *Measurement 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*, JHEP **11** (2016) 047 LHCb-PAPER-2016-012, CERN-EP-2016-141, arXiv:1606.0473.
- [69] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \rightarrow \Lambda \mu^+ \mu^-$ decays*, JHEP **06** (2015) 115 LHCb-PAPER-2015-009, CERN-PH-EP-2015-078, arXiv:1503.0713.
- [70] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \rightarrow K^* \mu^+ \mu^-$ decays*, JHEP **06** (2014) 133 LHCb-PAPER-2014-006, CERN-PH-EP-2014-055, arXiv:1403.8044.
- [71] A. Ali, E. Lunghi, C. Greub, and G. Hiller, *Improved model-independent analysis of semileptonic and radiative rare b decays*, arXiv:hep-ph/0112300.
- [72] P. Ball and R. Zwicky, *$B_{d,s} \rightarrow \rho, \omega, K^*, \phi$ decay form factors from light-cone sum rules revisited*, arXiv:hep-ph/0412079.
- [73] T. Hurth, F. Mahmoudi, and S. Neshatpour, *On the anomalies in the latest lhcb data*, 1603.0865.
- [74] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *Global analysis of $b \rightarrow s \ell \ell$ anomalies*, 1510.4239.
- [75] B. Capdevila *et al.*, *Patterns of new physics in $b \rightarrow s \ell^+ \ell^-$ transitions in the light of recent data*, 2017.
- [76] W. Altmannshofer, C. Niehoff, P. Stangl, and D. M. Straub, *Status of the $B \rightarrow K^{*+} \mu^+ \mu^-$ anomaly after moriond 2017*, 1703.9189.

-
- [77] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \rightarrow K^{*0}\ell^+\ell^-$ decays*, [arXiv:1705.0580](#) LHCb-PAPER-2017-013, CERN-EP-2017-100, [arXiv:1705.0580](#), submitted to JHEP.

Acknowledgements

In conclusion I want to thank the many, many people who supported me in life and during the nearly four years of creating this thesis!

While I will most likely forget multiple (which I already feel extremely sorry for) I would like to highlight some of those people:

- Prof. Dr. Ulrich Uwer, the main supervisor of my thesis work. He supported me since my Masters studies, had to endure many complicated talks, and was at all time a great help with his massive experience in Particle Physics.
- Prof. Dr. André Schöning, the second referee of my thesis, who already corrected my Master thesis and encouraged me to continue as a PhD student.
- Prof. Dr. Stephanie Hansmann-Menzemer, who was as the second supervisor of my thesis work always very supportive in pointing out problems in the analysis work, while providing ideas how to improve.
- Dr. Klaus Reygers, who was the third supervisor of my thesis work.
- Dr. Christoph Langenbruch, Dr. Christian Linn, and Dr. Michel De Cian, who I worked with intensely. Without their collaboration and support I would never have been able to get this work done.
- The LHCb Heidelberg group. Besides being great people to work with, I could not have enjoyed more spending five years of my life with such a large variety of great and friendly people.
- The whole LHCb collaboration.
- Last but definitely not least: my friends and family. Without your support I would not be here to write this. Be it by guiding me through all stages of life, being at my side at tough times, or just having a friendly chat and distracting me. You inspire me every day, some are enjoying Physics as much as I do and continue to work in this field, others give me puzzled looks whenever I start to drift off, yet still support me in my dreams.