Dissertation

presented to the

Combined Faculty of Natural Sciences and Mathematics
at the Ruperto Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Propound by

M.Sc. David Gerick

born in Cologne, Germany

Oral examination: 27 April 2021

Search for physics beyond the Standard Model in an angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay at LHCb

Referees:

Prof. Dr. Stephanie Hansmann-Menzemer

Prof. Dr. Klaus Reygers

Abstract

In this thesis an angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay with the subsequent decays $K^{*+} \to K^0_S\pi^+$ and $K^0_S \to \pi^+\pi^-$ is presented. The full data set collected by the LHCb experiment during the first two periods of pp collisions at the Large Hadron Collider in the years 2011 – 2018 is used. In total, 737 ± 34 signal candidates are selected, which allows to determine all angular observables and their correlations in the $B^+ \to K^{*+}\mu^+\mu^-$ decay for the first time. Furthermore, it is the first time that all angular observables are measured in a $b \to s\ell^+\ell^-$ transition other than that included in the decay of the neutral B^0 meson. A four-dimensional maximum-likelihood fit is used to extract two sets of angular observables in ten intervals of the invariant dimuon mass squared, q^2. For most observables and in most q^2 intervals the resulting values are compatible with Standard Model predictions. However, most prominently the CP-averaged (and optimised) angular observables A_{FB} (P_2) and S_5 (P'_5) show significant discrepancies from Standard Model predictions in the q^2 region below the J/ψ resonance. A global evaluation of the result in terms of the real part of the underlying vector coupling strength in the model-independent effective field theory description prefers a shift of $\Delta\mathrm{Re}(\mathcal{C}_9) = -1.9$ from the Standard Model value with a significance of 3.1σ.
The pattern of deviation is coherent with previous measurements in the isospin-partner decay of the B^0 meson. The size of the deviation largely depends on the considered q^2 regions and choice of theory nuisance parameters for the global fit.

Zusammenfassung

In dieser Thesis wird eine Winkelanalyse des $B^+ \to K^{*+}\mu^+\mu^-$ Zerfalls mit den anschließenden Zerfällen $K^{*+} \to K^0_S\pi^+$ und $K^0_S \to \pi^+\pi^-$ präsentiert. Der vollständige Datensatz, welcher am LHCb-Experiment in den ersten beiden Perioden mit pp-Kollisionen am Large Hadron Collider in den Jahren 2011 – 2018 gesammelt wurde, wird verwendet. Insgesamt werden 737 ± 34 Signalkandidaten selektiert. Mit dieser Anzahl an selektierten Kandidaten ist es erstmals möglich alle Winkelobservablen und deren Korrelationen im $B^+ \to K^{*+}\mu^+\mu^-$ Zerfall zu messen. Es ist darüberhinaus das erste Mal, dass der vollständige Satz an Winkelobservablen in einem $b \to s\ell^+\ell^-$ Übergang unabhängig dem im Zerfall des neutral B Meson gemessen wird. Ein vierdimensionaler Maximum-Likelihood-Fit wird verwendet, um zwei Sätze von Winkelobservablen in zehn Intervallen der quadratierten invarianten Masse des Muon-Paars, q^2, zu extrahieren. Zufällig gemischte Datensätze werden verwendet um die Korrelationen zwischen allen Observablen zu bestimmen.
Für die meisten Observablen und in den meisten q^2-Intervallen sind die resultierenden Werte mit den Vorhersagen des Standardmodells kompatibel. Am deutlichsten zeigen

jedoch die CP-gemittelten (und optimisierten) Winkelobservablen A_{FB} (P_2) und S_5 (P_5') signifikante Diskrepanzen zu Standardmodellvorhersagen in der q^2-Region unterhalb der J/ψ-Resonanz. Eine globale Bewertung der Ergebnisse in Bezug auf der reellen Komponente der zugrundeliegenden Vektorkopplungsstärke in der modellunabhängigen effektiven Feldtheorie bevorzugt eine Verschiebung von $\Delta\mathrm{Re}(\mathcal{C}_9) = -1{,}9$ mit einer Signifikanz von $3{,}1\sigma$ gegenüber dem Wert des Standardmodells.
Diese Spannung stimmt mit früheren Beobachtungen bei Zerfällen des Isospin-Partners, des B^0 Mesons, überein. Der Wert der Abweichung hängt jedoch weitgehend von den betrachteten q^2-Regionen und der Wahl der freien Parameter für die globale Anpassung der Theorievorhersage ab.

Preface

The measurements presented in this thesis are published as a pre-print version on arXiv and are submitted for publication in the peer-reviewed journal Physical Review Letters under the title of

"Angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay", available on arXiv.org [1]

The author of this thesis is one of the main authors of the above publication and the main person responsible for the analysis work from which they originate. This work has been carried out by the author within the LHCb collaboration, which is an international association of about 1400 scientists and engineers from 85 institutes in 19 countries. This implies the usage of common software to analyse the data collected by the LHCb experiment. Both the collection of the data and the development of common software used to analyse them result from the effort of many current and former collaboration members. The author presented the results of the above publication for the first time to a public audience at the workshop *Implications of LHCb measurements and future prospects* in October 2020 [2]. Supplemental material to the above publication is made public in Ref. [3] and a more technical and detailed description of the analysis is available internally to the LHCb collaboration in Ref. [4].

Physikalisches Institut Heidelberg, February 2021

Contents

1 Introduction **1**

2 Theoretical framework **5**

2.1 The Standard Model of Particle Physics 5

2.1.1 Particle content and forces 6

2.1.2 CKM mechanism and flavour-changing neutral currents 10

2.1.3 Physical phenomena beyond the Standard Model 12

2.2 Decays of rare B mesons 13

2.3 Kinematic description of electroweak penguin decays 17

2.3.1 Angular differential decay rate 18

2.3.2 Angular observables 21

3 Experimental apparatus **23**

3.1 The Large Hadron Collider 23

3.2 The LHCb experiment 24

3.2.1 Tracking system 26

3.2.2 Magnet 28

3.2.3 Particle identification 29

3.2.4 Calorimeters 30

3.2.5 Muon system 30

3.2.6 Trigger 31

4 Data sample **35**

4.1 Event simulation 35

4.1.1 Signal filtering of simulated candidates 36

4.1.2 Weighting of simulated candidates 36

4.2 Definition of decay variables 39

4.2.1 Momentum and energy 39

4.2.2 Particle identification 39

4.2.3 Invariant masses 40

4.2.4 Kinematical variables . . . 40
4.3 Candidate selection . . . 41
4.3.1 Trigger decisions . . . 42
4.3.2 Central event filtering . . . 42
4.3.3 One dimensional cut criteria . . . 43
4.3.4 Invariant mass fits . . . 46
4.3.5 Multi-variate classification . . . 48
4.4 Selection conclusions . . . 52

5 Determination of angular observables 57
5.1 Maximum likelihood estimation . . . 57
5.2 Fit model . . . 60
5.3 Angular acceptance correction . . . 63
5.4 Angular folding techniques . . . 66
5.5 Correlation matrices . . . 67
5.6 Fit strategy . . . 69
5.7 Fit validations . . . 72
5.7.1 Studies with pseudodata . . . 72
5.7.2 Fit to simulation . . . 77
5.7.3 Fit to reference decay . . . 77
5.8 Feldman-Cousins scans . . . 81

6 Evaluation of systematic uncertainties 85
6.1 Angular acceptance correction . . . 85
6.1.1 Angular acceptance parametrisation . . . 86
6.1.2 Statistical precision of simulation sample . . . 86
6.1.3 Simulation-data weights . . . 87
6.2 Angular resolution . . . 88
6.3 Mass model . . . 89
6.4 Angular background model . . . 91
6.5 Peaking background veto . . . 92
6.6 Simulation-data differences . . . 93
6.6.1 Kaon momentum discrepancy . . . 94
6.6.2 Trigger efficiency . . . 95
6.7 Constraint on the S-wave fraction . . . 95
6.8 Bias of the likelihood fit . . . 97
6.9 Difference between angular folds . . . 98
6.10 Summary of systematic studies . . . 98

6.11 Blinded studies on result compatibility 99

7 Results of angular observables **103**
7.1 Tabular results 103
7.2 Graphical results 103
7.3 Correlation matrices 108
7.4 EFT vector coupling strength 115

8 Conclusion **117**

A Supplemental material **121**
A.1 Simulation-data agreement 121
A.2 Generation of pseudodata 126
A.3 Pseudoexperiment results 127
A.4 Method of moments 134
A.5 Fit projections 135
A.6 Likelihood profiles of Feldman-Cousins scans 146

Bibliography **155**

Acknowledgements **166**

1 Introduction

"It is our responsibility as scientists, knowing the great progress which comes from a satisfactory philosophy of ignorance, the great progress which is the fruit of freedom of thought, to proclaim the value of this freedom; to teach how doubt is not to be feared but welcomed and discussed; and to demand this freedom as our duty to all coming generations."

– Richard Feynman, *The value of science*, 1955

With the experimental confirmation of the Higgs boson [5, 6] in 2012, the theoretical description of the Standard Model of particle physics (SM) has converged into a seemingly complete state to explain all observed interactions of elementary particles in laboratory experiments. To date, the experiments at the Large Hadron Collider (LHC), the current high-energy frontier of particle colliders, have not yet reported measurements that have deviations from the SM predictions with a degree of certainty that can claim a discovery [1] of physics beyond the SM. However, numerous results of astronomical observations and cosmological measurements cannot be explained in terms of elementary particle interactions as currently described by the SM. These shortcomings motivate the search for new physical phenomena beyond the SM.

Over the last decade, discrepancies with steadily increasing significance between measurements and theory predictions have arisen in the sector of loop-suppressed, flavour-changing neutral-current decays of B hadrons. In these decays, the b quark decays via an internal quark loop into an s quark with the additional emission of a pair of oppositely charged leptons, $\ell^+\ell^-$. The combination of a small SM transition probability and the possibility for new physical phenomena to enter the internal quark loops, make these decays sensitive to contributions from heavy mediators. A new physics process inside the loop with a mediator of heavy mass can affect the transition probability by a measurable amount, causing the measured value to deviate from the prediction of the SM. Because of their relatively small total branching fraction of below 10^{-6}, B meson decays

[1] In high-energy physics, the threshold of a discovery is set to five standard deviations, 5σ; equivalent to a statistical fluctuation probability of about 1 in 3.5 million.

with $b \to s\ell^+\ell^-$ transitions are referred to as *rare decays*.
Recent studies of decays involving $b \to s\ell^+\ell^-$ transitions by the LHCb collaboration measured consistently lower differential branching fractions [7–10] compared to predictions from the SM [11–18]. Ratios of differential branching fractions between decays with leptons of different flavour are found to deviate from the precisely predicted theoretical value [19–23]. *Angular analyses* probe the vectorial coupling strength of $b \to s\ell^+\ell^-$ transitions by measuring the size of the amplitude of angular moments in the differential decay rates. In the last years, deviations have been found in these angular analyses – especially of the $B^0 \to K^{*0}\mu^+\mu^-$ decay at LHCb [24–27], where K^{*0} denotes the $K^{*0}(892)$ meson [9, 28–34].
These deviations in decays involving $b \to s\ell^+\ell^-$ transitions are gathered under the term *flavour anomalies* [35] [2]. The flavour anomalies are theoretically interesting as they can be consistently explained by the presence of contributions from additional vector or axial-vector currents [36–56]. Yet, effects from uncertainties related to hadronic form factors or long-distance contributions cannot be ruled out [57–60].
Nevertheless, the studies performed in the $B^0 \to K^{*0}\mu^+\mu^-$ decay, with the recent angular analysis [27] by the LHCb collaboration confirming the tension with respect to the SM predictions, hint towards a consistent picture of potential new physics phenomena [43]. It is therefore expected that such new physics phenomena will also manifest themselves in other decays of B hadrons involving $b \to s\ell^+\ell^-$ transitions. Motivated by this, an angular analysis of the decay of the isospin partner B^+ is performed on the full data set recorded by the LHCb experiment. Due to the high luminosity and large production cross section of $b\bar{b}$ pairs in pp collisions, the LHCb experiment has recorded a large amount of events containing decays of B mesons in the years 2011, 2012 (Run 1) and 2015–2018 (Run 2), at centre-of-mass energies of 7, 8 and 13 TeV, respectively. The investigated data set corresponds to an integrated luminosity of $9\,\mathrm{fb}^{-1}$. The number of events selected from this data allows for the first time to determine all angular observables and their correlations in the $B^+ \to K^{*+}\mu^+\mu^-$ decay. Also, this analysis is the first to measure all angular observables in a $b \to s\ell^+\ell^-$ transition other than that included in the decay of the neutral B meson.

In the following, an angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay with $K^{*+} \to K^0_S\pi^+$ and $K^0_S \to \pi^+\pi^-$ is presented in all details; starting with the introduction to the theoretical framework and kinematic description of these rare decays in Chapter 2. The LHCb experiment and its detectors are introduced in Chapter 3.
After these two general chapters, the specific details of the angular analysis of the

[2] In addition to decays with $b \to s\ell^+\ell^-$ transitions, flavour anomalies also comprise deviations in decays containing $b \to c\ell\nu$ transitions.

$B^+ \to K^{*+}\mu^+\mu^-$ decay are explained. For this, the document follows events of pp collisions throughout their analysis process starting at the point of recording in the detectors. In Chapter 4, the selection of signal candidates is laid out. It utilises common software of the LHCb collaboration, software tools from external sources, as well as program code written by the author of this thesis. Both the signal decay sample and a reference sample using $B^+ \to J/\psi\, K^{*+}$ with $J/\psi \to \mu^+\mu^-$ decays are obtained by the same selection criteria. Effects of the selection and a non-homogeneous detector geometry on the angular distributions are corrected for by considering samples of simulated signal events. In addition, data sets of reference channel decays, of simulated events and pseudoexperiments are used to validate the functionality of the maximum-likelihood fit in Chapter 5 and to evaluate any systematic effects in Chapter 6.
At the end, the sample of selected signal candidates undergoes numerous multi-dimensional simultaneous maximum-likelihood fits to obtain two sets of angular observables, which are presented in Chapter 7. The results for the angular observables are discussed and further evaluated in terms of the underlying effective field theory coupling strengths. Finally, the document is concluded in Chapter 8, including an outlook into future prospects of the $B^+ \to K^{*+}\mu^+\mu^-$ decay at LHCb.

2 Theoretical framework around $b \to s\ell^+\ell^-$ transitions

In this chapter the theoretical framework underlying the analysis is introduced. The introduction starts with a brief overview [1] of the Standard Model of Particle Physics (SM) in Section 2.1. The list of known elementary particles and the interactions in the SM are given, together with an outlook on additional physical phenomena beyond the current Standard Model. Observations and potential theories of physics beyond the SM are discussed.

Flavour-changing neutral-current decays of B mesons with a transition of the b quark to an s quark and two oppositely charged leptons (so called $b \to s\ell^+\ell^-$ transitions) are an excellent environment to find hints of physics beyond the SM. In Sections 2.2 and 2.3, a description of these decays along with a detailed look particularly on the $B^+ \to K^{*+}\mu^+\mu^-$ decay are given.

2.1 The Standard Model of Particle Physics

Around halfway through the twentieth century the theoretical description of elementary particles and their interaction as we know it today started evolving by combining field theory and quantum mechanics into a unified theoretical framework. Previous attempts to include relativity into quantum theory such as relativistic quantum mechanics had not been successful. Following the formalism of classical electrodynamics, the new combined theory defines particles as excitation of the underlying particle fields. The theory became known as the Standard Model of Particle Physics (SM) and was mainly completed by the end of the 1970's to the theory we know today.

The reliable and general description of elementary particles and their interactions is achieved by a re-normalisable quantum field theory. All terms of the field equations are included in the *Lagrangian* of the theory, which is required to be invariant under local

[1] The reader is expected to posses a basic knowledge on symmetry groups and fundamental interactions, as a detailed and full explanation of the theory is beyond the scope of this document. For a detailed introduction to the Standard Model, usual literature like Ref. [61] are recommended.

gauge transformation of a symmetry group. From this requirement of local gauge symmetry, the particle interactions emerge and their forces are mediated by gauge bosons [2]. The symmetry group of the SM is

$$SU(3)_{\mathrm{C}} \times SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}},$$

with the indices indicating the colour charge C and the hypercharge Y and L for left-handedness. According to Noether's theorem [62], for every symmetry group a charge is conserved under the interactions of the very symmetry it is generated from. The exact combination of the three symmetry groups is motivated by the description of the experimental observations of the particle content and their interactions. Both, the particle content and the interaction forces with their corresponding charges, are discussed in the next paragraph.

2.1.1 Particle content and forces

The symmetry group of the SM defines the fundamental forces and their associated gauge bosons. An interplay between theoretical predictions and experimental observations has led to the full picture of the theory as we know it today. The last piece added to the (current version of the) puzzle was the Higgs boson discovery in 2012 by the ATLAS [5] and CMS [6] collaborations.

Particle content

All visible matter in this universe is built of atoms. The atom nuclei are composed of neutrons and protons, with *electrons* (e) in shells around the core. The neutrons and protons are made of *up* (u) and *down* (d) quarks [3]. The first generation fermions - particles with half-integer spin - are u, d and e along with the *electron neutrino* (ν_e). In total, there are three generation of fermions we know of [63]. Each generation comprises leptons (electron-like particles and neutrinos) and quarks, which will be introduced in detail in the following.
The *muon* (μ) and the *tauon* (τ) are often referred to as the big siblings of the electron, as they have the same SM quantum numbers as electrons but their masses are about 200× and 3500× larger. Both pair up with their corresponding neutrinos, the *tauon neutrino* (ν_τ) and the *muon neutrino* (ν_μ) to form the second and third generation of leptons, respectively. All six leptons are listed in the bottom left corner of the SM particle content summary table in Fig. 2.1. Electron, muon and tauon carry one negative elementary

[2] Bosons are elementary particles with an integer spin; oppose to fermions with half-integer spin ($\frac{1}{2}, \frac{3}{2}, \ldots$).
[3] In addition to consisting of up and down quarks, neutron and protons contain further particles, like gluons and randomly generated and annihilated pairs of quarks and anti-quarks.

electrical charge ($-e$), while neutrinos are electrically neutral. While the charged leptons are all massive, neutrinos are massless in the SM. Experimentally observed neutrino oscillations however show that neutrinos have to posses mass [64, 65]. Their masses are extremely small compared to all other elementary particles and have not been measured yet, but only received upper limits [66]. The minimal extension of the SM to accommodate these masses is not further discussed in this context. The observed neutrino oscillations allow lepton flavour violating transitions in the SM at extremely small and non-observable rates of below $O\left(10^{-50}\right)$ [67]. A lepton flavour quantum number is associated with the three generations of leptons and this flavour is observed to be conserved in all interactions in the SM, even though this quantum number does not origin from the local symmetry group description of the theory [68]. In addition, leptons carry a weak-isospin number of $-\frac{1}{2}$ and neutrinos of $+\frac{1}{2}$. All fermions interact via the weak interaction as explained in more detail in the next section. These six leptons have anti-particle partners, which carry the opposite electric charge ($+e$ for charged anti-leptons and still zero for anti-neutrinos) and the opposite flavour quantum number, but the identical mass [69].
Up (u) and *down* (d) quarks also have larger siblings: *charm* (c) and *truth* [4] (t) as up-type quarks; *strange* (s) and *beauty* [5] (b) for the d quark. Quarks carry an individual flavour too, which is named after the corresponding quark. However, other than for leptons, the quark flavour is not conserved in all interactions. More details on decays of quarks and their importance for this analysis is given in Section 2.1.2. Quarks do not carry integer values of the electric charge, but possess $+\frac{2}{3}e$ and $-\frac{1}{3}e$ for up- and down-type, respectively. In addition, quarks carry colour charge, which is introduced in the next paragraph along with the strong interaction force. The weak-isospin is carried by all quarks, with up type quarks ($|I, I_3\rangle = +\frac{1}{2}$) and down ($|I, I_3\rangle = -\frac{1}{2}$) quarks. For anti-quarks, which exist for all quarks, these electric, flavour and colour charges are inverted.
This sums up the twelve fermions plus their twelve anti-fermions of the Standard Model.

Fundamental interactions

Three out of the four known fundamental forces [6] are included in the SM:

- electromagnetic force
- weak nuclear force
- strong nuclear force

[4] A more common notation for the t quark is *top*.
[5] An alternative naming for the b quark is *bottom*, in analogy to *top* for the t quark.
[6] Gravity as the fourth force is not described by the SM.

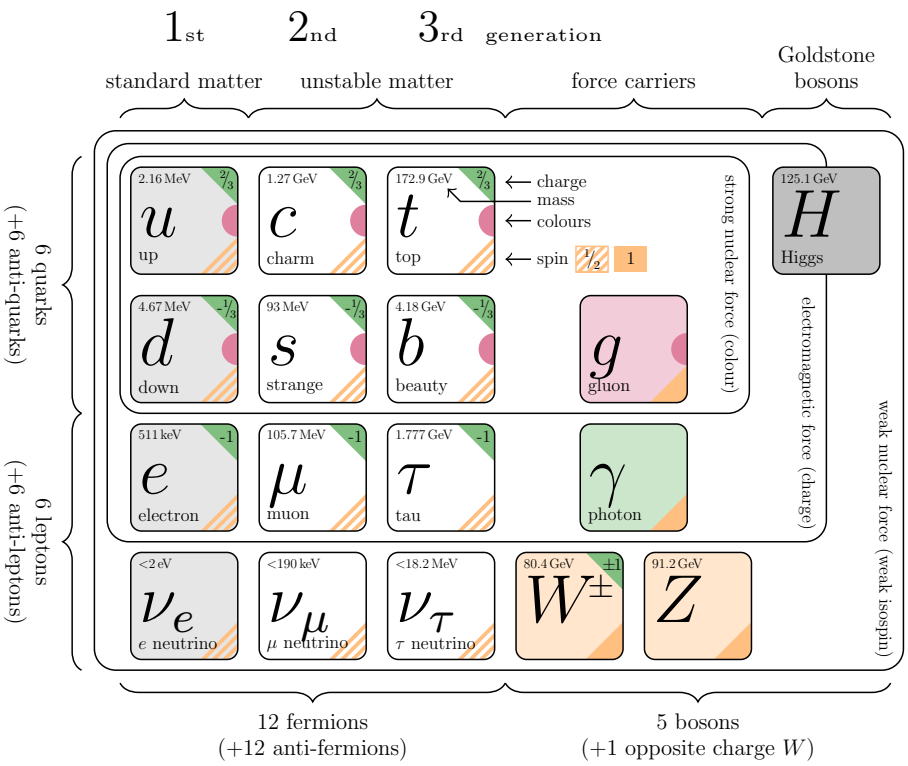

Figure 2.1: The particle content of the Standard Model. Modified from Ref. [70] and with values taken from Ref. [71].

The electromagnetic and weak nuclear force are combined in the symmetry group $SU(2)_\mathrm{L} \times U(1)_\mathrm{Y}$ of the Weinberg-Salam model [72–74], which is broken under the Higgs mechanism [7] to form the united description of the *electroweak* interactions. The four gauge bosons carrying the weak isospin-charge and hypercharge are the $SU(2)_\mathrm{L}$ bosons, $W^{0,1,2}$, and the $U(1)_\mathrm{Y}$ boson, B^0, of the unbroken Weinberg-Salam model. From these bosons, the experimentally observable mass-eigenstates of these gauge bosons are a linear combination of the charge-eigenstates. The weak gauge bosons in the mass-eigenstates are a superposition of charge-eigenstates, defined as

$$\begin{aligned} |W^\pm\rangle &= |W^1\rangle \pm i|W^2\rangle \quad \text{and} \\ |Z\rangle &= -\sin\theta_W |B^0\rangle + \cos\theta_W |W^0\rangle \,, \end{aligned} \tag{2.1}$$

with the Weinberg angle θ_W [75] [8]. The massless gauge boson $|\gamma\rangle = \cos\theta_W|B^0\rangle + \sin\theta_W|W^0\rangle$ carries the electromagnetic force. These bosons couple only to fermions which carry the corresponding charge. Left-handed fermions carry weak-isospin $\frac{1}{2}$ while right-handed

[7] The Higgs mechanism and the Higgs boson are not directly relevant for the context of this work and therefore are not further addressed in this context.

[8] The value of the Weinberg mixing angle is not predictable by theory but has been experimentally measured and is an input parameter to the SM.

fermions have weak-isospin equal to zero. Therefore, the bosons of the weak interaction, $W^{0,1,2}$, only couple to left-handed fermions, indicated by the index L of the symmetry group $SU(2)_{\mathrm{L}}$. On the other hand all fermions with hypercharge Y couple to the B^0 boson.

The $SU(3)_{\mathrm{C}}$ symmetry group generates the strong interaction and the corresponding colour charge C. Eight gauge bosons, the gluons (g), are mediators of the strong interaction and carry themselves colour charge. The mass-eigenstates of the gauge bosons of all three fundamental forces are included on the right half of the overview chart of the particle content of the SM in Fig. 2.1. Following the analogy of colour as the charge of the strong nuclear force, one refers to Quantum Chromodynamics (QCD) as the study of the field theory of strong interactions. As the potential between two particles with colour charge does not decrease with larger distance, these particles can only be found in confined states. Or in other words, quarks and gluons hadronise forming baryons and mesons, which are colour neutral objects. Like in colour theory, different colour charges can be added to form colourless objects, *i.e.* white. Mesons consist of one quark and one anti-quark while baryons contain either three quarks or three anti-quarks [9]. Another feature of QCD (compared to the electroweak interactions) is the *running* of the coupling, $\alpha_s(Q^2)$, which is inversely proportional to the transferred momentum squared in the interaction, $Q^2 = -q^2$. Only at energies significantly above the energy scale of QCD, $\Lambda_{\mathrm{QCD}} \approx 200\,\mathrm{MeV}$, the calculation of strong interactions can be done using perturbation theory, while lower energies require numerical calculations such as lattice QCD, that require large computational power and suffer from larger uncertainties. For the analysis presented in this thesis, the uncertainties on QCD calculations for the hadronic form-factors lead to the usage of an alternative definition of parameters, introduced in Section 2.3.2.

The SM shows an extraordinary performance in providing predictions for particle interactions over many orders of magnitudes and has withstood tests throughout the last decades with unprecedented precision, but still lacks the ability to fully describe of all observed phenomena, as discussed in Section 2.1.3.
The SM can be used to predict the properties of B mesons: their lifetime, their branching fractions ($\mathcal{B}$) and kinematics of their decays. Some of these predictions are tested with unprecedented precision against experimental data in the analysis of this thesis.

[9] Combinations of more than three quarks have been postulated for a long time [76] and hadrons with a quark content of four or five have been discovered in recent years, including LHCb publications [77–82]. However, these hadrons are beyond the scope of this thesis and will not be addressed.

2.1.2 CKM mechanism and flavour-changing neutral currents

In the following, the weak interaction between two quarks with different (charged current) and same electric charge (neutral current) is discussed.

Charged current

A weak interaction of two quarks in combination with a $W^\pm$ boson is called a *charged current* interaction, as the $W^\pm$ changes the electric charge of the fermionic current by $\pm e$, as shown in the Feynman graph [83] [10] in Fig. 2.2. The part of the SM Lagrangian for

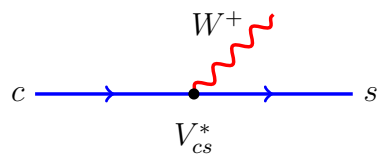

Figure 2.2: The Feynman diagram of the flavour-changing charged current of a c to an s quark with the emission of a W^+ boson.

these interactions expressed in the flavour basis is

$$\mathcal{L} \supset \frac{G_F}{2\sqrt{2}} \left[\sum_{i,j} W_\mu^+ \overline{\hat{u}'_i} \gamma^\mu \left(1-\gamma^5\right) \hat{d}'_j + \sum_{i,j} W_\mu^- \overline{\hat{d}'_i} \gamma^\mu \left(1-\gamma^5\right) \hat{u}'_j \right] \tag{2.2}$$

with the Fermi coupling constant, G_F, the Dirac matrices γ^μ and γ^5 and the field spinors $\hat{d}'$ and $\hat{u}'$ for the flavour eigenstates of down- and up-type quarks. The indices i and j run over the three quarks for each type. Experiments have shown, that the weak interaction propagator is of *vector minus axial-vector* $(V-A)$ type [84, 85] which only couples to left handed chirality

$$\frac{1}{2}\gamma^\mu \left(1-\gamma^5\right) . \tag{2.3}$$

The fields of quark flavour eigenstates $\hat{d}'$ and $\hat{u}'$ are

$$\hat{u}' = \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} \quad \text{and} \quad \hat{d}' = \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} . \tag{2.4}$$

[10] Feynman graphs are sketches of the matrix element of a decay amplitude and usually imply a time axis in positive horizontal direction. The arrows on the lines indicate (anti-)particles by running with (against) time. Interactions are occurring at vertices where lines intersect.

In experimental observations, not the flavour eigenstates but mass eigenstates of quarks are observed. The flavour eigenstates, q', are related to the mass eigenstates, q, via matrix rotations

$$\hat{u}' = \boldsymbol{U_u}\,\hat{u} \quad \text{and} \quad \hat{d}' = \boldsymbol{U_d}\,\hat{d}\,. \tag{2.5}$$

The hat denotes the vectorial description of eigenstates, $\hat{q}^{(\prime)}$ and the unitary rotation matrices, $\boldsymbol{U_{d,u}}$. In the mass basis, the Lagrangian from Eq. (2.2) reads as follows

$$\mathcal{L} \supset \frac{G_F}{2\sqrt{2}} \left[\sum_{i,j} W_\mu^+ \overline{\hat{u}_i} \gamma^\mu \left(1-\gamma^5\right) \boldsymbol{V_{ij}} \hat{d}_j + \sum_{i,j} W_\mu^- \overline{\hat{d}_i} \gamma^\mu \left(1-\gamma^5\right) \boldsymbol{V_{ij}^\dagger} \hat{u}_j \right] , \tag{2.6}$$

with the so called Cabibbo-Kobayashi-Maskawa (CKM) matrix, $\boldsymbol{V_{\mathbf{CKM}}} := \boldsymbol{V_{ij}} = \boldsymbol{U_u^\dagger U_d}$ [86, 87]. This matrix contains experimentally measured values, as theory does not provide information on the size of the elements of the matrix. The norm values of the nine complex elements of the unitary 3×3 matrix show a diagonally favoured behaviour. As these values enter the transition probability (or matrix element) of a decay, the values indicate that decays preferably occur within the same generation of fermions.

$$|\boldsymbol{V_{\mathbf{CKM}}}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.974 & 0.042 \\ 0.009 & 0.041 & 0.999 \end{pmatrix} , \tag{2.7}$$

with the values obtained by a global fit to all available measurements by the CKM fitter collaboration [88] in 2019 and their uncertainties well below the stated precision. The unitarity condition of the CKM matrix can be expressed in different ways using the elements of the matrix, as for example

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0\,. \tag{2.8}$$

Neutral current

Following the same approach as for the charged-current interaction, the Lagrangian for neutral-current interactions in the flavour basis for transitions down-to-down- and up-to-up-type quarks is

$$\mathcal{L} \supset \left[\sum_{i,j} Z_\mu \overline{\hat{d}'_i} \gamma^\mu \left(1-\gamma^5\right) \hat{d}'_j + \sum_{i,j} Z_\mu \overline{\hat{u}'_i} \gamma^\mu \left(1-\gamma^5\right) \hat{u}'_j \right] , \tag{2.9}$$

with the Z boson, the mediator of neutral-current interactions. Using the definition of the rotation matrices in Eq. (2.5) and the unitary of the rotation matrices

$$\boldsymbol{U_d^\dagger U_d} = \boldsymbol{U_u^\dagger U_u} = \boldsymbol{I}\,, \tag{2.10}$$

one can easily see in Eq. (2.11) that flavour-changing neutral currents (FCNC) are not possible at tree level [11] in the SM

$$\mathcal{L} \supset \left[\sum_{i,j} Z_\mu \overline{\hat{d}_i} \gamma^\mu \left(1-\gamma^5\right) \boldsymbol{U_d^\dagger U_d} \hat{d}_j + \sum_{i,j} Z_\mu \overline{\hat{u}_i} \gamma^\mu \left(1-\gamma^5\right) \boldsymbol{U_u^\dagger U_u} \hat{u}_j \right]. \tag{2.11}$$

Non-diagonal elements of the unitary matrices in Eq. (2.10) are zero allowing only transitions of quarks from index j to i if $i = j$, *i.e.* the quark is unchanged and we obtain a simple *scattering* process.

When considering flavour-changing neutral-currents in quark-loop transitions, the unitary condition of the CKM matrix in Eq. (2.8) would cause the contribution from all three quark generations in the quantum loops to cancel out, if the three quark masses were identical. The contributions of three quarks is indicated in the loops of both Feynman graphs in Fig. 2.3. Occurrence of flavour-changing neutral-current decays are strongly suppressed by being forbidden at tree-level and suppress at loop level, which is known as the Glashow-Iliopolus-Maiani (GIM) mechanism [89]. Typical orders of magnitude for the branching fractions are $10^{-6} - 10^{-9}$, which puts these decays into the so called regime of *rare decays*.

Studying these decays requires large samples of generated mesons.

2.1.3 Physical phenomena beyond the Standard Model

Despite the huge success of the SM, experimental observations in different fields of physics have revealed dark spots in the prediction capabilities of the SM. Most prominent indications for physical phenomena beyond the SM arise in cosmological and astrophysical observations. Baryonic matter as described by the SM is what we observe (mainly) via electromagnetic interactions. All matter, which is (as far as all conducted measurements can tell) inert – or couples only extremely weakly – to the three forces of the SM is classified as *dark matter* and there are a number of independent observations which are strongly indicating towards presence of dark matter in our universe. Two astrophysical observations are given as examples: the rotation curves of spiral galaxies and the power spectrum of background irradiation in the universe.

Observed rotation curves of stars in spiral galaxies do not reproduce the curves expected

[11] *tree level* refers to interactions with a single interaction point and without internal quantum loops.

by applying Kepler orbital kinematics on the observed distribution of visible stars. While central stars follow the predicted velocity distributions, stars further away from the galactic centre rotate significantly faster than expected. This observation can be explained by the presence of a halo made from gravitationally interacting matter [90], which is not observed via electromagnetic radiation, i.e. *dark matter*.
Evaluations of the peak spectrum in the anisotropy power spectrum of Cosmic Microwave Background (CMB) measured by the Planck collaboration [91] yield a baryonic matter density $\Omega_b h^2 = 0.0224 \pm 0.0001$ and a dark matter density $\Omega_c h^2 = 0.120 \pm 0.001$. In other words, the matter described by the SM makes up only 16% of the total matter in the universe, plus a 84% dark matter content.
With these examples suggesting the presence of matter made up from particles beyond the current SM, the big question is, what these particles are and how – if at all – they interact with the SM particles or forces.
Therefore, the search for New Physics beyond the SM is an on-going field of research at the frontiers of particle physics. Direct searches try to directly detect new particles via annihilation, generation or scattering. The benefits of these searches are, that in the case of discovery, information about the mass and coupling could directly be determined. Indirect searches on the other side profit from Heisenberg's uncertainty principle [92] that allows particles to contribute to decay amplitudes via quantum loops, despite having a (much) greater mass than the transferred energy in the decay. This allows to find hints of new particles beyond the energy scale of the particle collision.

2.2 Decays of rare B mesons and $B^+ \to K^{*+}\mu^+\mu^-$

An ideal environment to find contributions from new physical phenomena are flavour-changing neutral-current decays of B mesons. These mesons are made of a $q\overline{q}$ pair, with one (anti-)quark being b flavoured plus a second lighter flavoured (anti-)quark. In this thesis, decays of B^+ and B^- mesons are analysed,

$$|B^+\rangle = |\bar{b}u\rangle \text{ and } |B^-\rangle = |b\overline{u}\rangle. \tag{2.12}$$

With a mass of $5297.34 \pm 0.12\,\mathrm{MeV}/c^2$ [71] for the $B^\pm$ and similar masses for B^0 and B_s^0, B mesons are the heaviest flavourful mesons [12] by far. Their decays create a large phase space of how the energy can be distributed among the decay daughters. The available energy

[12] The heavier top quark does not hadronise into mesons or baryons, as its lifetime is significantly shorter than the hadronisation time, the time it takes to form hadrons. Y mesons made of $b\bar{b}$ combinations are heavier at about $10\,\mathrm{GeV}/c^2$, but are flavour-neutral.

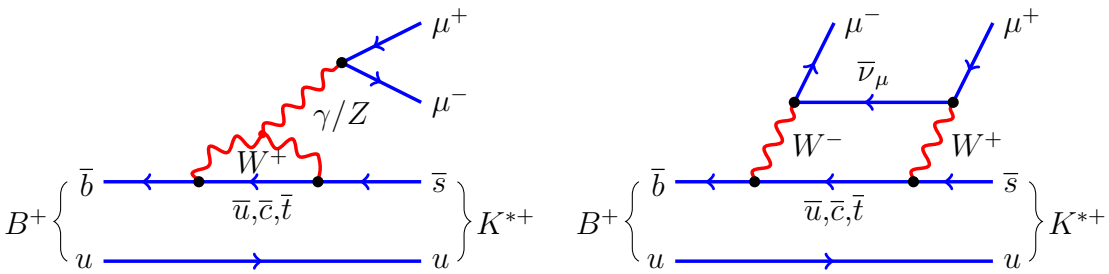

Figure 2.3: The Feynman diagrams of the two major decay amplitudes of $B^+ \to K^{*+}\mu^+\mu^-$, electroweak penguin (left) and box-diagram (right). A $b \to s\ell^+\ell^-$ transition in the form of a $\bar{b}$ to $\bar{s}$ quark is accompanied by a spectator u quark. From the internal quark loop, an oppositely charged pair of muons is emitted.

is equal to the mass differences of the decaying particle to the final state particles [13]. Flavour-changing neutral currents of a b ($\bar{b}$) to a s ($\bar{s}$) quark plus the creation of an oppositely charged pair of leptons are referred to as $b \to s\ell^+\ell^-$ decays. Fig. 2.3 shows the Feynman graphs for the $B^+ \to K^{*+}\mu^+\mu^-$ decay with the creation of the lepton pair ($\mu^+\mu^-$) in the quantum loop of the b to an s quark transition. The electroweak loop (penguin) diagram emits the pair of muons via a photon or Z boson and the box diagram via $W^\pm$ boson exchange. Also indicated are the three up-type quarks in the internal loops, which interfere destructively in the total decay amplitude and, as mentioned before in Section 2.1.2, strongly suppress the probability of the decay.

The strong suppression of the SM decay amplitude opens room for possible enhancements in the measured amplitudes by contributions from New Physics (NP) models beyond the SM. Therefore, studying decays with $b \to s\ell^+\ell^-$ transitions is a highly motivated field of research at the frontiers of high energy particle physics.

The total decay amplitude, $\mathcal{A}_{\text{tot}}$, is the sum of the SM amplitude and a (potential) NP amplitude,

$$\mathcal{A}_{\text{tot}} = \mathcal{A}_{\text{SM}} + \mathcal{A}_{\text{NP}} \,. \tag{2.13}$$

A small SM amplitude increases the sensitivity to NP amplitudes. The transition probability of a decay is the experimentally measured quantity and it is proportional to the square of the complex decay amplitudes in Eq. (2.13). Hence, the transition probability is also sensitive to possible interference between the two complex amplitudes, which can significantly enhance the effects of NP contributions. Discrepancies between the measured transition probability and the SM predictions would directly hint to a non-vanishing contribution from a NP amplitude. This counts not only for the absolute values of the

[13] In particle physics, the dimensions of mass (eV/c^2), momenta (eV/c) and energy (eV) are treated the same by setting the speed of light constant to one ($c = 1$). The unit eV is the energy equivalent to a unit charge e at an electric potential of one volt.

decay probability but is also applicable to the angular distributions of the decay daughter particles. The latter is particularly interesting for a pseudo-scalar [14] B meson (with spin 0) decaying into a vector meson with spin 1, and two leptons (each with spin $\frac{1}{2}$).
Over the past years, measurements of $b \to s\ell^+\ell^-$ transitions in the decay of $B^0 \to K^{*0}\mu^+\mu^-$ [25–27] have established a tension with the theory prediction from the SM. For the first time, the measurement of the full angular parametrisation (see Section 2.3.1) of the helicity angles is performed using $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K^0_S\pi^+$ and $K^0_S \to \pi^+\pi^-$. The decays of B^+ and B^0 differ only by the spectator quark; the u quark of the B^+ in Fig. 2.3 is exchanged by a d quark in the B^0 case, which results in very similar expectation for the decay kinematics, as the part of the decay containing the quark loops is not affected by the spectator quark. One noticeable difference of the charged B meson decay with respect to the neutral channel is weak annihilation [93], when additional contributions from $\overline{b} \to W^+\overline{u}$ transitions lead to the emission of a K^{*+} meson. The mentioned helicity angles between the four particles, K^0_S, π^+, μ^+ and μ^-, are defined in Section 2.3 following the introduction of the effective field theory used to describe $b \to s\ell^+\ell^-$ decays.
A basic introduction into the theoretical description of the angular decay rate and two examples for potential NP models, which might enter the quark loops in this decay, are given in the following paragraph.

Effective field theory description of $b \to s\ell^+\ell^-$ decays

In analogy to the Fermi theory for beta decays [94, 95], a model-independent *effective field theory* (EFT) can be used to describe $b \to s\ell^+\ell^-$ transitions, as the energy scale of the quantum loop ($\Lambda_{\mathrm{EW}} \approx m(W^\pm)$) is much larger than the exchange energy momentum ($\sqrt{q^2} < m(B^+)$). In this so called *operator product expansion* (OPE) [96] and for the limit of a short-distance expansion ($x \to 0$), the interaction point can be described by local, scale-independent operators, $\mathcal{O}_i$. However, the scale dependence is convoluted into the coupling strength coefficients, $\mathcal{C}_i(\mu)$, the so called Wilson coefficients [97] at scale μ. The energy scale of $b \to s\ell^+\ell^-$ transitions is equal to the mass of the b quark, $\mu = \lambda_b$. The effective Lagrangian is proportional to the sum of all possible operators times their coupling strengths

$$\mathcal{L}_{\mathrm{eff}} \sim \sum_i \mathcal{C}_i(\mu)\mathcal{O}_i. \tag{2.14}$$

[14] Even though, the internal structure of the meson – the two (anti-)quarks – carry non-zero spin, the total spin of the meson is zero, hence it is *pseudo-scalar*.

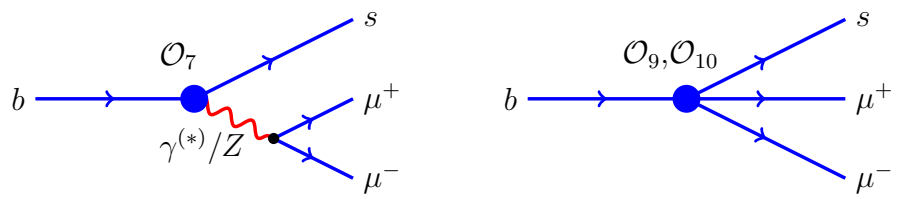

Figure 2.4: The two decay graphs for effective field theory description $b \to s\ell^+\ell^-$ transitions, along with the corresponding effective operators $\mathcal{O}_7$, $\mathcal{O}_9$ and $\mathcal{O}_{10}$. The large filled circles indicate the presence of the effective field theory with model-independent interactions inside. The asterisk in the left graph indicates an either virtual or on-shell photon.

The relevant operators [98] with a sizeable effect on $b \to s\ell^+\ell^-$ transitions are

$$\begin{aligned} \mathcal{O}_7 &= \frac{e}{g^2} m_b \left(\overline{s}\sigma_{\mu\nu} P_R b\right) F^{\mu\nu} \,, \\ \mathcal{O}_9 &= \frac{e^2}{g^2} \left(\overline{s}\gamma_\mu P_L b\right) \left(\overline{\mu}\gamma^\mu \mu\right) \,, \\ \mathcal{O}_{10} &= \frac{e^2}{g^2} \left(\overline{s}\gamma_\mu P_L b\right) \left(\overline{\mu}\gamma^\mu \gamma_5 \mu\right) \,, \end{aligned} \tag{2.15}$$

with the weak interaction parity operators $P_{L,R} = (1 \pm \gamma_5)/2$, the Dirac gamma matrices, γ_μ and γ_5, the bases of the Lie-Algebra [99], $\sigma_{\mu\nu} = \frac{1}{4}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)$, and the electromagnetic field strength tensor, $F^{\mu\nu}$. Using these effective operators, the full theory description of the $b \to s\ell^+\ell^-$ transitions as shown in both Feynman graphs of Fig. 2.3 simplifies into four particle interaction points, as illustrated in Fig. 2.4.
The SM values for the Wilson coefficients $\mathcal{C}_7$, $\mathcal{C}_9$ and $\mathcal{C}_{10}$ are calculated at the energy scale at which the EFT and the full theory meet, *i.e.* the electroweak scale, Λ_{EW}. Then they are run down by renormalisation group equations (RGE) [96] to the relevant scale of $b \to s\ell^+\ell^-$ transitions, the b-mass scale, Λ_b [100]. Wilson coefficients provide a convenient interface between model-independent measurements of low-energy probes and NP models with large energy scale. Therefore, the results of this analysis are interpreted and evaluated for their compatibility with the current SM physics via Wilson coefficients [43, 49, 101]. In analogy to Eq. (2.13), the Wilson coefficients are the sum of SM and NP contributions

$$\mathcal{C}_i(\mu) = \mathcal{C}_i^{\mathrm{SM}}(\mu) + \mathcal{C}_i^{\mathrm{NP}}(\mu) \,. \tag{2.16}$$

Potential New Physics theory models

For physics beyond the Standard Model, several theoretical models are available, which are up for validation or rejection by experimental results. Two of the possible NP contributions to $b \to s\ell^+\ell^-$ penguin decays, *leptoquarks* and Z' models are briefly outlined in the following. Fig. 2.5 shows the potential Feynman diagrams, which could directly contribute

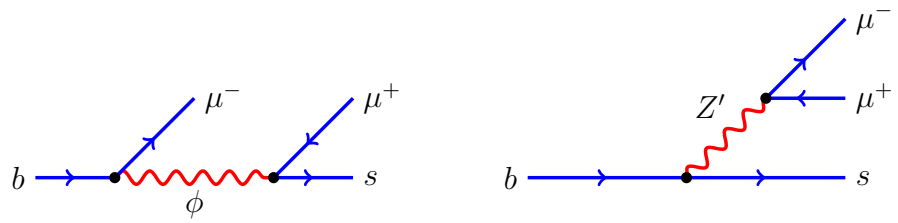

Figure 2.5: Potential tree-level FCNC transitions of $b \to s\mu^+\mu^-$ with a leptoquark ϕ (left) and a new gauge boson Z' (right).

to $b \to s\ell^+\ell^-$ decay amplitudes.

A very popular cluster of new physics models are leptoquarks, which (as the name indicates) couple a lepton directly to a quark and vice versa. In this scenario, leptoquarks may carry electrical charges equal to $\frac{1}{3}e$, $-\frac{2}{3}e$ or $\frac{4}{3}e$ [42]. Leptoquarks can be introduced as scalars via an additional U(1) symmetry group [102, 103] or as vector particles via Pati-Salam SU(4) symmetry breaking [104–106]. The former provides a currently favoured model to explain deviations from the SM in $b \to s\ell^+\ell^-$ transitions and other decays [43], due to the interpretation of constrains from lepton flavour universality tests [20, 21, 107, 108] and mixing of neutral B mesons [109–113].

In different models, the introduction of a new symmetry group U$'$(1) gives rise to new coupling terms with lepton and quark flavour violating interactions, which are mediated by the resulting gauge boson, Z'. Models can contain "family non-universal charges" [114], "little flavour model" [115] or "effective Z' " [116] with only higher dimensional operators coupling the Z' to SM particles. As before for the leptoquarks, bounds from other experimental observations already restrict the newly introduced models, requiring special adjustments in the models to prevent contradictions with existing results. For the Z', the coupling needs to be in a muonic flavour favouring regime [117].

Proving or rejecting this kind of theories is the goal of all experimental measurements in flavour physics; and the first angular analysis of $B^+ \to K^{*+}\mu^+\mu^-$ is a further contribution on finding a valid model of new physics beyond the SM.

2.3 Kinematic description of $B^+ \to K^{*+}\mu^+\mu^-$ decays

Three angles, θ_K, θ_ℓ and ϕ, and the invariant mass squared of the two leptons (q^2) describe the full kinematics of a pseudo-scalar meson ($B^\pm$) decaying into two pseudo-scalar mesons K^0_S and $\pi^\pm$ (via a vector boson $K^{*\pm}$ [15]) and two oppositely charged leptons, in this case a pair of muons, μ^+ and μ^-.

The angle θ_K (θ_ℓ) is defined as the angle between the direction of the K^0_S ($\mu^\pm$) in the K^{*+}

[15] The $K\pi$ system can decay via higher excitation states of K^* resonances, as it will be discussed in Section 2.3.1.

(dimuon) rest frame and the direction of the K^{*+} (dimuon) system in the $B^\pm$ rest frame. Additionally, ϕ describes the angle between the two decay planes given by the dimuon system on one side and by the $K^0_S\pi^+$ system on the other side, all in the rest frame of the $B^\pm$ as illustrated in a topographic sketch in Fig. 2.6.
The mathematical description of the angles are defined in Eq. (2.17) and are consistent with the definitions of previous publications in Refs. [24, 26, 27].

$$\begin{aligned}
\cos\theta_K &= \left(\hat{p}^{(K^{*\pm})}_{K^0_S}\right) \cdot \left(\hat{p}^{(B^\pm)}_{K^{*\pm}}\right) = \left(\hat{p}^{(K^{*\pm})}_{K^0_S}\right) \cdot \left(-\hat{p}^{K^{*\pm}}_{(B^\pm)}\right) \\
\cos\theta_\ell &= \left(\hat{p}^{(\mu^+\mu^-)}_{\mu^\pm}\right) \cdot \left(\hat{p}^{(B^\pm)}_{\mu^+\mu^-}\right) = \left(\hat{p}^{(\mu^+\mu^-)}_{\mu^\pm}\right) \cdot \left(-\hat{p}^{\mu^+\mu^-}_{(B^\pm)}\right) \\
\cos\phi &= \left[\left(\hat{p}^{(B^\pm)}_{\mu^\pm}\right) \times \left(\hat{p}^{(B^\pm)}_{\mu^\mp}\right)\right] \cdot \left[\left(\hat{p}^{(B^\pm)}_{K^0_S}\right) \times \left(\hat{p}^{(B^\pm)}_{\pi^\pm}\right)\right] \\
\sin\phi &= \left[\left(\hat{p}^{(B^\pm)}_{\mu^\pm}\right) \times \left(\hat{p}^{(B^\pm)}_{\mu^\mp}\right)\right] \times \left[\left(\hat{p}^{(B^\pm)}_{K^0_S}\right) \times \left(\hat{p}^{(B^\pm)}_{\pi^\pm}\right)\right] \cdot \left(\hat{p}^{(B^\pm)}_{K^{*\pm}}\right)
\end{aligned} \tag{2.17}$$

with $\hat{p}^{(Y)}_X$, the normalised vector of a particle or multi-particle system, X, in the rest frame of system Y.

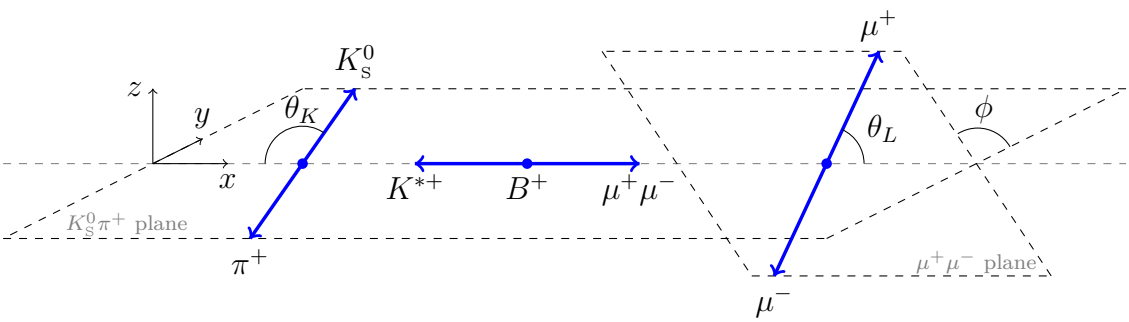

Figure 2.6: Sketch of the angular description of the $B^+ \to K^{*+}\mu^+\mu^-$ decay. The three angles θ_K, θ_ℓ and ϕ describe the decay kinematic. The indicated decay planes are spanned by the muons and hadrons. The mathematical definition of the three angles and their reference systems are given in Eq. (2.17).

Charge conjugation Throughout this thesis, charge conjugation is implied. This means, the decays of $B^+ \to K^{*+}\mu^+\mu^-$ also contain $B^- \to K^{*-}\mu^-\mu^+$ decays. As a direct consequence, all measurements are CP-averaged.

2.3.1 Angular differential decay rate

The differential angular decay rate of the $B^+ \to K^{*+}\mu^+\mu^-$ decay is defined as a function of $\cos\theta_K$, $\cos\theta_\ell$ and ϕ, which factorise to orthogonal angular moments, $f_i(\cos\theta_K, \cos\theta_\ell, \phi)$.

The angular decay rate – integrated over a given q^2 region – takes the general form

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\,\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\cos\theta_K\mathrm{d}\cos\theta_\ell\mathrm{d}\phi} \propto \sum_i S_i f_i\left(\cos\theta_K, \cos\theta_\ell, \phi\right)\ , \tag{2.18}$$

with CP-averaged angular observables, S_i, and the CP-averaged decay rate, $(\Gamma+\bar{\Gamma})$, which is normalised in each q^2 region.
More precisely, the decay rate consists of two parts: a P-wave contribution, where the $K\pi$ system originates from the $K^{*+}(892)$ resonance, and the S-wave contribution, with $K\pi$ originating from different $K\pi$ resonances, such as the $K^{*+}(1400)$. The goal of this analysis is the determination of the P-wave angular observables. The P- and S-wave components of the decay are non-distinguishable due to identical initial and final states. A different angular decay rate of the S-wave component with respect to the P-wave results in different angular distributions in $\cos\theta_\ell$, $\cos\theta_K$ and ϕ. Therefore, it has to be taken into account in the fit model used for the extraction of (P-wave) angular observables. Both fit models for the P- and S-wave are introduced in the following.

P-wave contribution

The signal component of the angular decay rate is derived in more detail in the literature (see Refs. [98, 118]) by squaring the matrix element of the decay of a pseudo-scalar to a vector boson plus two leptons. The matrix elements are summed over all possible spin states of the decay daughters plus using kinematic identities. The final angular decay rate is given by

$$\begin{aligned}
\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\,\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\cos\theta_K\mathrm{d}\cos\theta_\ell\mathrm{d}\phi}\Bigg|_{\mathrm{P}} = \\
\frac{9}{32\pi}\Big[& S_{1s}\sin^2\theta_K && + S_{1c}\cos^2\theta_K \\
& +S_{2s}\sin^2\theta_K\cos 2\theta_\ell && + S_{2c}\cos^2\theta_K\cos 2\theta_\ell \\
& +S_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi && + S_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi \\
& +S_5\sin 2\theta_K\sin\theta_\ell\cos\phi && + S_{6s}\sin^2\theta_K\cos\theta_\ell \\
& +S_7\sin 2\theta_K\sin\theta_\ell\sin\phi && + S_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi \\
& +S_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big]\ .
\end{aligned} \tag{2.19}$$

In this equation, the $\mu^+\mu^-$ system neglects contributions from scalar configurations of a pure photon pole transition in the quark loops at very low q^2 values. Also, the formula operates in the limit of considering the muons to be quasi-massless in order to fulfil the conditions $q^2 \gg 4m(\mu)^2$. The angular moments f_3, f_4, ..., f_9 are all orthogonal to each

other, theoretically making the angular observables S_3–S_9 uncorrelated. S_{1s}, S_{1c}, S_{2s} and S_{2c} are not independent and follow the normalisation conditions and correlations:

$$\frac{3}{4}\left(2S_{1s}+S_{1c}\right)-\frac{1}{4}\left(2S_{2s}+S_{2c}\right)=1\,,\ S_{2c}=-S_{1c}\ \text{and}\ S_{2s}=\frac{1}{3}S_{1s}\,. \tag{2.20}$$

With these normalisation conditions on the four non-orthogonal observables, the introduction of the fraction of the longitudinal polarisation of the K^{*+} meson

$$F_{\mathrm{L}}=1-\frac{4}{3}S_{1s}\,, \tag{2.21}$$

and the forward-backward asymmetry of the two muons

$$A_{\mathrm{FB}}=\frac{3}{4}S_{6s}\,, \tag{2.22}$$

the angular decay rate for the P-wave component reads as

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\,\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\cos\theta_K\mathrm{d}\cos\theta_\ell\mathrm{d}\phi}\bigg|_{\mathrm{P}}=&\\ \frac{9}{32\pi}\Big[\tfrac{3}{4}(1-F_{\mathrm{L}})\sin^2\theta_K &+F_{\mathrm{L}}\cos^2\theta_K\\ +\tfrac{1}{4}(1-F_{\mathrm{L}})\sin^2\theta_K\cos 2\theta_\ell &-F_{\mathrm{L}}\cos^2\theta_K\cos 2\theta_\ell\\ +S_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi &+S_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi\\ +S_5\sin 2\theta_K\sin\theta_\ell\cos\phi &+\tfrac{3}{4}A_{\mathrm{FB}}\sin^2\theta_K\cos\theta_\ell\\ +S_7\sin 2\theta_K\sin\theta_\ell\sin\phi &+S_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi\\ +S_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big]\,.&\end{aligned} \tag{2.23}$$

S-wave contribution

$K\pi$ resonances with a different angular momentum than the desired $K^{*+}(892)$ are present in the angular distributions of the decay. Events from P- and S-wave can only be separated by their angular distribution [16], hence an S-wave angular differential decay rate is added to the P-wave,

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\,\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\cos\theta_K\mathrm{d}\cos\theta_\ell\mathrm{d}\phi}\bigg|_{\mathrm{S}}=\frac{3}{16\pi}F_{\mathrm{S}}\sin^2\theta_\ell\,. \tag{2.24}$$

[16] In Section 5.2 we will learn that the invariant mass of the K^{*+} meson, $m(K^0_{\mathrm{S}}\pi^+)$, can be used to distinguish the P- from the S-wave component.

As P- and S-wave are complex amplitudes, additional interference terms arise from the combination of both, given by

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\cos\theta_K\mathrm{d}\cos\theta_\ell\mathrm{d}\phi}\bigg|_{\mathrm{PS}} = \\ \frac{3}{16\pi}\Big[S_{\mathrm{S1}}\cos\theta_K\sin^2\theta_\ell \\ +S_{\mathrm{S2}}\sin\theta_K\sin 2\theta_\ell\cos\phi \quad + S_{\mathrm{S3}}\sin\theta_K\sin\theta_\ell\cos\phi \\ +S_{\mathrm{S4}}\sin\theta_K\sin\theta_\ell\sin\phi \quad + S_{\mathrm{S5}}\sin\theta_K\sin 2\theta_\ell\sin\phi\Big]\,.\end{aligned} \tag{2.25}$$

The resulting P- and S-wave differential angular decay rate

The P-wave, S-wave and interference terms in Eqs. (2.23) to (2.25) are added to form the total differential decay rate. By definition, the P-wave part is scaled by $(1-F_{\mathrm{S}})$ as shown in Eq. (2.26) [17].

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\cos\theta_K\mathrm{d}\cos\theta_\ell\mathrm{d}\phi}\bigg|_{\mathrm{P+S}} = \\ (1-F_{\mathrm{S}})\frac{9}{32\pi}\Big[\tfrac{3}{4}(1-F_{\mathrm{L}})\sin^2\theta_K \quad + F_{\mathrm{L}}\cos^2\theta_K \\ +\tfrac{1}{4}(1-F_{\mathrm{L}})\sin^2\theta_K\cos 2\theta_\ell \quad - F_{\mathrm{L}}\cos^2\theta_K\cos 2\theta_\ell \\ +S_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi \quad + S_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi \\ +S_5\sin 2\theta_K\sin\theta_\ell\cos\phi \quad + \tfrac{3}{4}A_{\mathrm{FB}}\sin^2\theta_K\cos\theta_\ell \\ +S_7\sin 2\theta_K\sin\theta_\ell\sin\phi \quad + S_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi \\ +S_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big] \\ +\frac{3}{16\pi}\Big[F_{\mathrm{S}}\sin^2\theta_\ell \quad + S_{\mathrm{S1}}\cos\theta_K\sin^2\theta_\ell \\ +S_{\mathrm{S2}}\sin\theta_K\sin 2\theta_\ell\cos\phi \quad + S_{\mathrm{S3}}\sin\theta_K\sin\theta_\ell\cos\phi \\ +S_{\mathrm{S4}}\sin\theta_K\sin\theta_\ell\sin\phi \quad + S_{\mathrm{S5}}\sin\theta_K\sin 2\theta_\ell\sin\phi\Big].\end{aligned} \tag{2.26}$$

2.3.2 Angular observables

The measurement of the already mentioned angular observables, S_i, of the differential angular decay rate of the P-wave decays in Eq. (2.23) is the goal of this analysis. Being the scaling factors of the angular moments, $f_i(\vec{\Omega})$, with a vectorial description of the angles, $\vec{\Omega} = (\cos\theta_K, \cos\theta_\ell, \phi)$, the angular observables S_i get contributions from the Wilson coefficients (see Section 2.2) as well as hadronic form-factors. A theoretically optimised description of angular observables is the $P_i^{(\prime)}$ observable basis. By taking the ratio of the

[17] Interference terms of the form $F_{\mathrm{S}}(1-F_{\mathrm{S}})$ are convoluted into the $S_{\mathrm{S}i}$ parameters.

previously introduced S_i observables and the longitudinal polarisation fraction of the K^{*+} meson, F_{L}, hadronic form-factor uncertainties cancel out to first order. The angular observables $P_i^{(\prime)}$ are defined in Ref. [119] by

$$P_1 = 2\frac{S_3}{1-F_{\mathrm{L}}}\,,\ P_2 = \frac{1}{2}\frac{S_{6s}}{1-F_{\mathrm{L}}}\,,\ P_3 = -\frac{S_9}{1-F_{\mathrm{L}}} \text{ and } P'_{4,5,6,8} = \frac{S_{4,5,7,8}}{\sqrt{F_{\mathrm{L}}\,(1-F_{\mathrm{L}})}}\,. \quad (2.27)$$

Both sets of angular observables, S_i and $P_i^{(\prime)}$, are determined in this analysis.

Remark on the usage of F_{L} and A_{FB} The ambiguity in the usage of F_{L} or S_{1s} and A_{FB} or S_{6s} can be seen in some points throughout this document. For the evaluations of systematic effects in Chapter 6 and the results in Chapter 7 values for F_{L} and A_{FB} are given. However, some results of intermediate studies of the fit functionality and validation tests in Chapter 5 are given for S_{1s} and S_{6s}. Internally, the fitter always uses S_{1s} and S_{6s}. It should be noted that a simple linear transformation, as defined in Eqs. (2.21) and (2.22), can be used to convert one result to the other.
Furthermore, results for F_{L} are given for both sets of angular observables. In the S_i basis, F_{L} is measured independently, while for the $P_i^{(\prime)}$ basis, F_{L} also enters the ratios of the $P_i^{(\prime)}$ observables, as defined in Eq. (2.27). To allow a correct evaluation of the results, the F_{L} results are given for both angular bases in Chapter 7.

3 Experimental apparatus

Precision measurements of particle decay chains with a branching fraction in the order of 10^{-6} and smaller [71], as in the $B^+ \to K^{*+}\mu^+\mu^-$ decay, require a large data set of $b\bar{b}$ events with good reconstruction efficiencies. Only the most capable particle accelerators provide the high luminosity of particle collisions with a large enough interaction cross section. The decay particles are efficiently recorded and reconstructed by specialised particle detectors. The data set used for the angular analysis in this thesis is recorded by the *Large Hadron Collider beauty* (LHCb) experiment [120, 121] at the *Large Hadron Collider* (LHC) [122, 123] at the European Organisation for Nuclear Research (CERN). The collider and the experiment are introduced in Sections 3.1 and 3.2.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at the European Organisation for Nuclear Research (CERN) on – or more precisely *underneath* – the border of France and Switzerland near Geneva is the most powerful particle accelerator in the world. Inside a 27 km long circular tunnel and at a depth of about 100 m underground, two beams of protons (and heavy ions in special operation modes) circulate in opposite directions inside two vacuum pipes. Supra conducting magnets bend and focus the particle beams. The beams consist of up to about 4000 bunches of approximately 10^{11} protons each and are eventually brought to collision at four interaction points. These points are home to the four major experiments, including LHCb [120, 121]. Fig. 3.1 shows a schematic view of the LHC and these four experiments, ALICE, ATLAS, CMS and LHCb, and two of LHC's pre-accelerators, namely the Proton Synchrotron (PS) [124] and Super Proton Synchrotron (SPS) [125–127] at the end of which the protons are injected into the LHC with an energy of 450 GeV. The LHC accelerates both circulating beams to centre-of-mass energies at pp collisions of 7 TeV in 2010 and 2011, 8 TeV in 2012 and 2013 and 13 TeV from 2015–2018 [128, 129]. The first period of

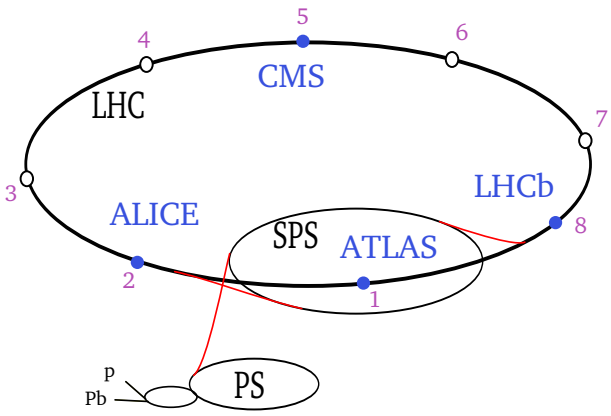

Figure 3.1: The accelerator chain for the Large Hadron Collider (LHC) at CERN with injections for both protons (p) and lead (Pb) indicated. Both the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) are increasing the beam energy prior to injection to the LHC. The four major LHC experiments are labelled accordingly to their position in the LHC ring. Modified from [130] and not to scale.

data taking in 2011–2012 is referred to as Run 1 [1] and the second period in 2015–2018 as Run 2. The resonance frequency of the LHC is 40 MHz, resulting in collisions every 25 ns [128, 129].

3.2 The LHCb experiment

The *Large Hadron Collider beauty* experiment (LHCb) [120, 121] has a unique role among the four main experiments located at the LHC. This is most prominently visible in the geometry of the detector, which is built in a single arm cone-like shape around the beam pipe towards one side of the interaction point, while all other three main experiments cover the complete sphere around the collision point to reconstruct all particles from the pp collisions. The reason for this geometry lies in the physics program of LHCb, which is dedicated to the studies of decays of hadrons containing c and b quarks. Compared to the total collision energies, the masses of these hadrons are rather small. The energy to create a $b\bar{b}$ pair from pp collisions is only a fraction of the total energy available in the collision. This increases the probability of asymmetric energies of the interacting partons [131], resulting in a strong boost along the beam axis [132]. The detectors' acceptance starts at about 10 mrad from the beam axis and reaches outwards to 250(300) mrad in the (non-)bending plane of the magnet (see Section 3.2.2). The decays of about 12% of all B mesons produced in pp collisions are fully contained [2] inside the LHCb detectors' acceptance. While the detector covers only about 2.3% of a full sphere.

[1] In the years 2010 and 2013 only insignificant data were recorded. In this analysis Run 1 refers to data from the years 2011 and 2012 only.
[2] All particles in the decay chain of the B meson are inside the active region of the LHCb detector.

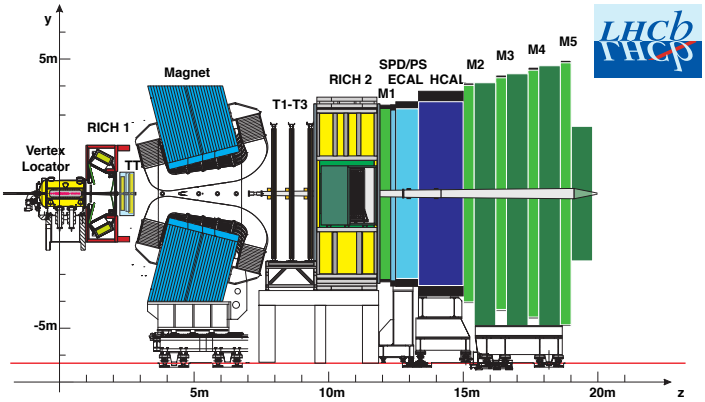

Figure 3.2: Schematic side view of the LHCb experiment along with the LHCb collaboration logo. The beam pipe runs horizontally at half of the height of the detector ($y = 0\,\mathrm{m}$) and proton beams collide inside the Vertex Locator on the very left ($z = 0\,\mathrm{m}$). Taken from Refs. [133, 134].

The LHCb experiment is an assembly of complementary particle detectors, providing particle tracking, particle identification, and momentum- and energy measurements. A side view schematic of all detector systems and the experiment's cone-like geometry are shown in Fig. 3.2. The individual detectors and their purposes are addressed in the following paragraphs.

Coordinate system

The coordinate system of the LHCb experiment originates at the nominal interaction point. Along the centre of the beam pipe into the detector, the positive z-axis is defined. Relative positions along the z-axis are expressed as up- and downstream, where upstream means towards the interaction point and downstream away from the interaction point. The positive y-axis is orientated upward in vertical direction and the horizontal x-axis is defined (as a result of the right-handedness) with the positive side towards the left when viewing downstream. Thus it points towards the centre of the LHC ring. The orientation of the y- and z-axis are visible in Fig. 3.2, with the x-axis oriented into the picture away from the reader. The spatial expansion of LHCb reaches from approximately $-5\,\mathrm{m}$ to $+5\,\mathrm{m}$ in both the x- and y-direction, and from $-1\,\mathrm{m}$ to $+20\,\mathrm{m}$ in the z-direction.
In addition to the three Cartesian coordinates, (x, y, z), the pseudorapidity for a vector $\vec{v} = (v_x, v_y, v_z)$ is defined as

$$\eta = \operatorname{arctanh}\left(\frac{v_z}{|\vec{v}|}\right) = -\log\left(\tan\frac{\theta}{2}\right) , \tag{3.1}$$

with the polar angle, θ.

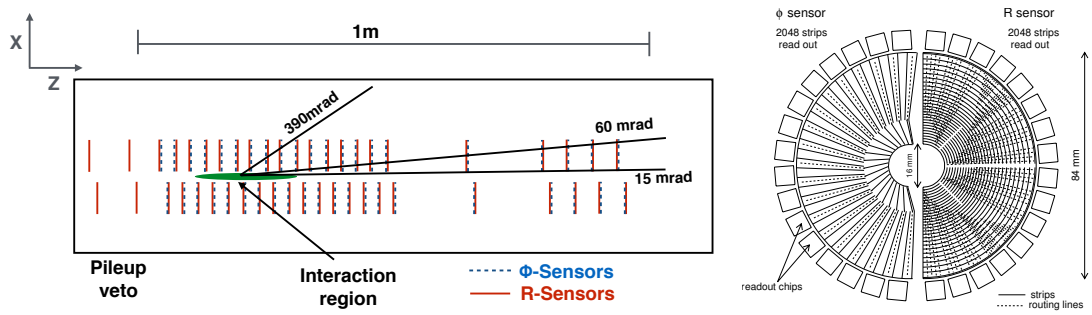

Figure 3.3: Schematic plan of VELO detector planes (left) with the *pp* collision region in green. View on the silicon strip mapping on a detector layer (right) show both R and Φ sensor routings [120].

3.2.1 Tracking system

Trajectories of charged particles from the collision are reconstructed from hits in tracking detectors. These reconstructed objects are called *tracks*. For this, LHCb uses three clusters of tracking detectors at different z-positions from the collision point, which consist of (a mix of) three detector technologies. All tracking detectors consist of planar detector layers in the x-y-plane, but use different technology of different size and granularity.

Vertex Locator (VELO)

The *Vertex Locator* [135, 136] (VELO) is located around the *pp* collision point and is separated from the LHC vacuum only by a very thin aluminium casing. Two movable halves containing each 21 semi-circular silicon strip detector planes are distanced about 7 mm in radial direction from the proton beams in closed configuration [3]. Each detector plane consists of two layers, one with radial (R-sensors) and one with azimuthal (Φ-sensors) silicon strips. In Fig. 3.3, the geometric assembly of the detector planes and a schematic of the semi-circle sensors are shown. At central (peripheral) regions, the pitch of the silicon strips is 38(102) µm.

The main tasks of the VELO detector are track finding and from this, the reconstruction of vertices. *Primary vertices* (PV) are the actual interaction point of *pp* collisions, which are determined by projecting particle tracks back to a common origin. *Secondary vertices* (SV) are decay points of long lived particles. The SV are displaced from the primary interaction point by the distance a primary particle travels during its decay time. Due to the heavy boost, SV are well separated from the PV in the order of mm to cm for *beauty* and *charm* hadrons. For longer decay cascades, tertiary vertices and even quaternary vertices are possible.

[3] To avoid damage due to the beam halo, the VELO opens up to about 58 mm distance when LHC is not in stable beam configuration.

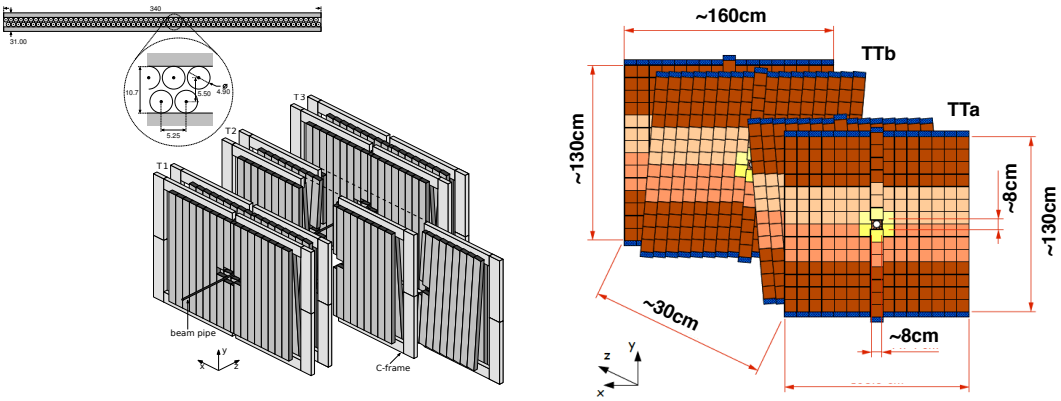

Figure 3.4: Schematics of the three stations of the Outer Tracker (left) [142], including a cross-section view of the 4.9 mm gaseous straw tubes [120]. The gaps in the OT layers around the beam-pipe is where the Inner Tracker (IT) is positioned. The Trigger Tracker (TT) is shown with all four layers – two straight and two tilted (right) [120].

Trigger Tracker (TT)

The Trigger Tracker (TT) is positioned in front of the bending magnet of LHCb at about $z = 2.5\,\mathrm{m}$ from the interaction point [137,138]. The primary use of the TT is to reconstruct long-lived neutral particles – for example K^0_S mesons – which decay outside of the VELO and low momentum particles (to about $2\,\mathrm{GeV}/c$) which are bend out of the detector acceptance by the magnet.
The $8.3\,\mathrm{m}^2$ of active area in four detector layers are covered by 198 µm silicon micro strips with a length of 33 cm. Two stereo layers are tilted by $\pm 5°$ with respect to the y-axis to achieve sensitivity on the y-position of tracks. Both, tilted and straight layers are shown in the right schematic of Fig. 3.4. Single hit resolution of the TT is about 50 µm in x-direction.

Inner Tracker (IT) and Outer Tracker (OT)

The LHCb main tracking stations T1-T3 are located right behind the magnet and in front of the calorimeters at a z-position of about 8–9 m and consist of four layers per station. Like the TT, two of the four layers are tilted by $\pm 5°$ to achieve sensitivity on the y-position of particle tracks. The layers are a composition of two different detector technologies, with the Inner Tracker (IT) [139] in the centre around the beam pipe consisting of the same silicon micro strips as the TT. The Outer Tracker (OT) [140–142] is a gaseous straw tupe detector with an outer dimension of $5\,\mathrm{m} \times 6\,\mathrm{m}$.
The IT covers the high occupancy regions around the beam pipe to prevent saturation of the detector. Similar as the TT, a resolution of about 50 µm is realised in the x-direction. Around the IT, where fewer particle tracks per area have to be recorded, the OT covers the majority of the active area of the T stations. The OT is made from 2.5 m long

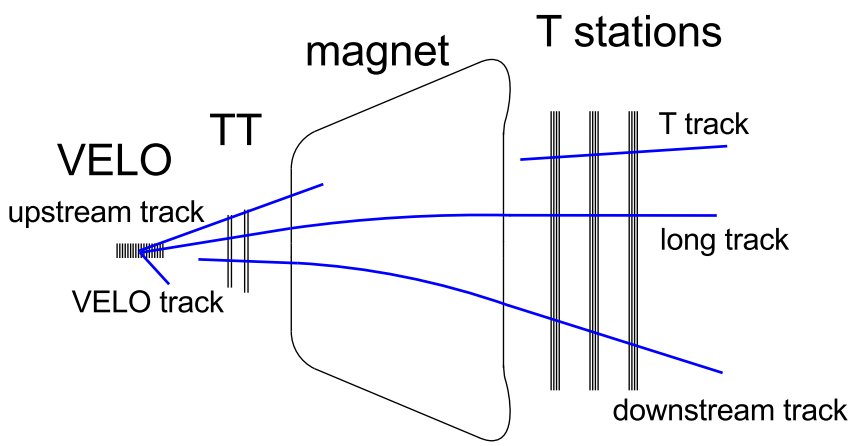

Figure 3.5: Different reconstructed track types from the LHCb track reconstruction [143] are shown in the y-z-plane. Inner- and Outer Tracker are merged into T stations. For this analysis long and downstream tracks are used to categorise the K_S^0 decay by the track type of its daughter pions. Note that long tracks are not required to be reconstructed in the TT.

gas filled straw tubes with an anode wire in the central axis. These drift tubes are proportional counters using a gas mixture of Ar, CO_2 and O_2 [142]. Ionising particle tracks are reconstructed with a spatial resolution of about 200 µm, which is less precise compared to the silicon strip detectors IT and TT, but sufficient for momentum measurements due to a long lever arm.

Track Types

In the track reconstruction of LHCb, track segments of individual detectors are combined by merge algorithms to form longer tracks that potentially pervade the full LHCb detector system. Depending on the included tracking detectors, the tracks are categorised in track types, of which the two important track types for this analysis are:

- long track: reconstructed in at least VELO and IT/OT plus optional TT hits
- down track: not present in VELO, but in both TT and IT/OT

These two and further track types are visualised in Fig. 3.5, with IT and OT displayed as a united detector, the *T stations* or T1–T3.

3.2.2 Magnet

A dipole magnet [144] is positioned at $4 < z < 7$ m, between the Trigger Tracker and the T stations, and has an integrated bending power of 4 Tm [4]. The magnet coils encircle

[4]The magnetic field of the magnet expands beyond the geometry of the magnet coils and a particle observes the full bending power of 4 Tm only when reaching to about 10 m into the detector.

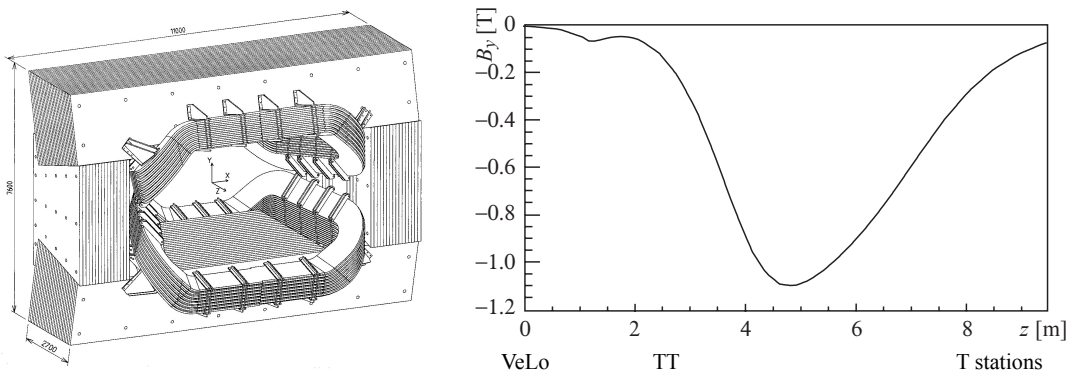

Figure 3.6: Schematic view of the magnet coils (left) [144] and the dependence of the magnet field strength as a function of the z-position (right) [137]. The labels of the three tracking detectors have been added to indicate the position with respect to the magnetic field strength.

the cone-shaped active area of the detector to not interfere with decay particles. Inside the magnet is a void only filled with air. A schematic view of the magnet along with the magnetic field strength as a function of the z-position is shown in Fig. 3.6. Charged particles are bent in x-direction by the magnetic field, which can be altered to positive (*up*) or negative (*down*) y-direction. Systematic effects of geometry-, detector- and charge-asymmetries can be evaluated due to the two different magnet polarities.
The measured curvature of tracks inside the magnet is used to determine particle momenta, achieving a relative momentum resolution of 5 ‰ at 20 GeV/c, about 8 ‰ at 100 GeV/c and 10 ‰ at 200 GeV/c [121, 145].

3.2.3 Particle identification

Two Ring Imaging Cherenkov (RICH) detectors [146–148] – one positioned before (RICH1) and one after (RICH2) the magnet – use Cherenkov light [149] from charged particles to determine the particle's velocity. In combination with momentum information from the tracking system the velocity is used to assign a the particle type hypothesis to the track of (mostly) long lived hadrons – such as pions, kaons and protons. Cherenkov light is created by particles moving faster than the speed of light in a given matter. The Cherenkov photons are emitted at an angle $\cos\theta_C = 1/\beta n$ with the particle's velocity, β, and the index of refraction of the medium, n. The photons are generated within the gas filled enclosure of the detectors, C_4F_{10} (RICH1) and CF_4 with 5% CO_2 (RICH2). Cherenkov photons are reflected out of the acceptance region by tilted spherical focusing mirrors towards hybrid photon detectors, as can be seen in the left of Fig. 3.7 for RICH1. From the ring patterns, the cone angle θ_C is calculated and assigned to a track. Using the momentum measurement along with the velocity β, the mass and therefore the particle species of this track can be determined. RICH1 provides good separation power in the

momentum range between 2 and 50 GeV/c, while RICH2 discriminates hadrons up to a momentum of 100 GeV/c.

3.2.4 Calorimeters

The energy of particles is determined in the calorimeter system [150–152] by measuring the amplitude of particle showers using scintillating light and photon detectors. In addition, the diversity of the implemented calorimeter detector types is used to distinguish light hadrons, electrons and high energetic photons. Further, the information from the calorimeters are a crucial input to the LHCb trigger system, which will be discussed in Section 3.2.6.
First in the stack of calorimeters is the Scintillating Pad Detector (SPD), which helps to identify electrons from photons, as only charged particles generate signals in form of scintillating light in this active detector layer. Behind a layer of lead follows the Pre-Shower (PS) detector, which supports the separation capabilities between electrons and light hadrons through the different shower characteristics of both.
Then the main calorimeter detectors, the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL) follow. The ECAL uses the *shashlik* technology made from 66 slices of 2 mm lead absorber plates, which are interleaved by 4 mm scintillator tiles for the active material. Wavelength-shifting fibres penetrate both lead and scintillators and guide created scintillating photons towards photon detectors, which are photo-multiplier tubes (PMT). Material and layer thickness are chosen such that showers of electrons and photons created via bremsstrahlung and e^+e^- pair production are fully contained in the ECAL.
Charged and neutral hadrons are stopped in the HCAL and generate hadronic showers. Their energy is measured by the amplitude of the shower. Here the absorber material is made from 6 mm steel plates, which are again read out by the combination of scintillator tiles, fibres and PMT.
All calorimeter detectors are shown in the central schematic of Fig. 3.7 with shower characteristics for photons γ, electrons e and hadrons h. The calorimeters are positioned downstream of RICH2.

3.2.5 Muon system

Consisting out of five stations (one in front and four downstream of the calorimeters) of multi-wire proportional chambers [153], the muon stations [154, 155] identify and track muons. Muons are the primary particles that pass through the calorimeters due to their minimal ionising characteristics. Particle tracks inside the muon stations are identified as muon tracks [156]. To further reduce pollution of the muon systems by hadrons, 80 cm thick iron shields are positioned in between the muon stations downstream of the calorimeters.

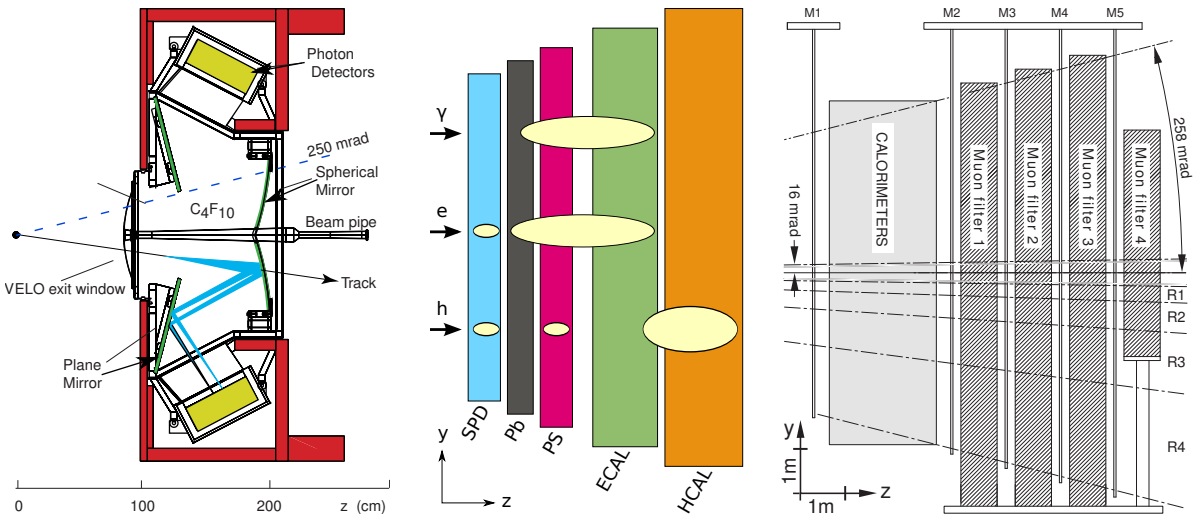

Figure 3.7: Schematic side view of RICH1 (left) [137] with the mirrors in green, hybrid photon detectors in yellow-green, the Cherenkov photon cones and the RICH enclosure in red. Schematics of the LHCb calorimeter systems (centre, *not to scale*) showing SPD, PS, ECAL and HCAL. Signatures in the calorimeters for different particle track types are indicated by the light yellow ellipses in the corresponding detectors. Side view of the muon stations (right) [155] in front and downstream of the calorimeters. Muon filters are 80 cm thick iron shields in between the muon detector planes. R1–R4 indicate the different ring segments of the muon stations.

The side view of the muon stations sandwiching the calorimeters are shown in Fig. 3.7 (right).

Due to their clean signature and importance for many analyses, as in this thesis, tracks in the muon stations are a crucial input to the LHCb trigger system, as explained in the following section.

3.2.6 Trigger

From the vast number of *pp* collisions inside the LHCb experiment, only collisions with interesting particle content are fully reconstructed and written to tape. Interesting collisions for the primary physics program of LHCb contain heavy *charm* and *beauty* hadrons or high energetic leptons. To select these collisions efficiently while making the best use of the finite computing resources, LHCb uses a three-staged trigger system [121, 157–160]. The system is staged into a level zero (L0) hardware trigger [161] and two software-based high-level triggers (HLT) [162]. The three successive steps in between the collisions and the tape storage along with their approximated throughput rates are given in Fig. 3.8 for the Run 2 configuration of the trigger system. At each trigger level, algorithms run in parallel to categorise the selected events. The algorithm configurations are called *lines*. Every step of the trigger uses more information from the detectors while reducing the throughput rate by about an order of magnitude. At the end, collisions are written to

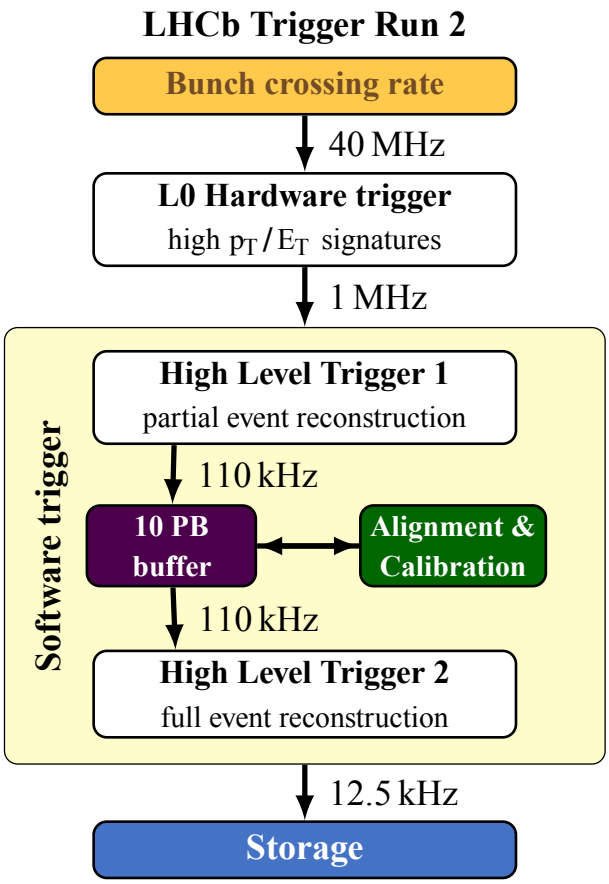

Figure 3.8: Diagram of the staged LHCb trigger system in Run 2 [160]. While the L0 trigger is implemented in hardware, the HLT1 and HLT2 are integrated in the reconstruction software. The trigger in Run 1 was similar, with the exception that alignment and calibration in the trigger was not the same as the final alignment and calibration used to reconstruct the events a second time prior to storage.

tape at a rate of 12.5 kHz.

The first level (L0) trigger is integrated into custom trigger electronics that trigger on large energy deposits in the calorimeters or large measured momenta in the muon stations at a maximum rate of 40 MHz. As shown in Fig. 3.8, the output rate of the L0 trigger is about 1 MHz (by design). The available bandwidth is mainly occupied by hadronic trigger lines and muon lines (each about 40%). Electron and photon lines making up the rest.

Selected data are passed to HLT triggers, which are software based and (partially) reconstruct particle tracks and vertices. Again different *lines* are evaluated in parallel to categorise the recorded data. HLT1 only reconstructs long tracks (see Fig. 3.5) with a momentum above 500 MeV/c along with primary vertices due to the strong time constrains at this trigger level. In HLT2, charged and neutral tracks are reconstructed. The particle identification (PID) is evaluated and the full information from all detectors are exploited. During Run 1, the reconstruction of the HLT was used solely for trigger purposes and all data were repeatedly reconstructed with a more sophisticated algorithm prior to

final storage. In Run 2, this difference between *online* and *offline* reconstruction became obsolete, due to improvements in the HLT and reconstruction software, which allowed for a full reconstruction in the HLT.
The lines of all trigger levels used in this analysis are discussed in Section 4.3.1.

4 Data sample

This analysis uses data recorded in proton-proton (pp) collisions by the LHCb experiment (see Section 3.2) in the years 2011, 2012 and 2015–2018 at centre-of-mass energies of 7 TeV, 8 TeV and 13 TeV, respectively. In this time period, LHCb recorded data equivalent to an integrated luminosity of $9\,\mathrm{fb}^{-1}$. In addition to the recorded data, simulation data with different generation conditions are used.
All data are split into four sub-categories due to different reconstruction and selection conditions. First, the data are split into a Run 1 and a Run 2 sample, as the centre-of-mass energy of the pp collisions and as a direct result the LHCb trigger performance greatly changed between the two data taking periods. Furthermore, the data are split by whether the K^0_S candidate decays inside (LL) or outside of the VELO (DD). The symbols D and L indicate the downstream and long track type (see Fig. 3.5) for each of the K^0_S daughter pions in the two cases, respectively.
Namely, the four data sub-sets are Run 1 DD, Run 1 LL, Run 2 DD and Run 2 LL. After the generation of simulated events in Section 4.1 and the commonly used decay variables in Section 4.2 are introduced, the selection of signal candidates is discussed in detail in Section 4.3.

4.1 Event simulation

Simulated events act as reference samples that are utilised in many aspects throughout all steps of the analysis. Centrally produced simulation by the LHCb collaboration is used for all simulated events [163]. The simulated – Monte Carlo (MC) – events are generated using the LHCb software framework Gauss [164] which uses Pythia [165] version 8 [166, 167] for the simulation of the pp collision and EvtGen [168, 169] for the creation of the B meson decay chain. Interactions of the decay particles with the detector are simulated using Geant4 [170, 171] and the digital detector response is generated by Boole [172]. The LHCb trigger decisions are emulated in Moore [172].
A set of three different MC samples are produced. The first two are signal decays,

$B^+ \to K^{*+}\mu^+\mu^-$, and reference decays, $B^+ \to J/\psi\, K^{*+}$, with a configuration equal to data. The simulation samples generally are of large statistics. These samples are used to obtain the shapes of mass distributions and to validate the angular fitter. They also function as input to the multi-variate classification in the selection in Section 4.3.5. In addition to the first two sample types, additional simulation of the signal decay is produced, for which the (axial-)vectorial and dipole couplings between decay particles are turned off in the generation. The angular and q^2 distributions are shaped only by phase-space kinematics. As a result, distributions in the decay angles $\cos\theta_K$, $\cos\theta_\ell$ and ϕ are flat at generation and can be used to correct effects of the detector acceptance and the candidate selection on the angular distributions. In the following, these simulation samples are referred to as *phase-space* (MC) simulation.
Due to different reconstruction efficiencies for K^0_S candidates in the trigger during the data taking in 2012, MC phase-space sample are generated separately for *early* 2012 and *late* 2012 with different trigger configurations. The differences in the angular distributions of early and late 2012 are found to be insignificant and hence the simulated data are combined to a single sample for 2012.

4.1.1 Signal filtering of simulated candidates

Simulation candidates are thoroughly checked to be correctly reconstructed and identified. For this, the BackgroundCategory [173] tool is used to require that all particles in the reconstructed decay chain are correctly matched to the generated particles. As explained in the beginning of Section 4.1, the LHCb simulation framework simulates the detector response for simulated particles. This detector simulation generates hits, which are flagged by a particle identification number [174] corresponding to the simulated particle. Tracks are reconstructed from hits and inherit their identification number [1]. The BackgroundCategory tool certifies the correctness of the particle identification number of each track in the decay chain plus the identification number of tracks upwards in the decay chain. Beside the signal decay particles, only additional soft photons may be radiated in the decay chain. If any of these conditions is not met, the simulated candidate is discarded.

4.1.2 Weighting of simulated candidates

Simulation at LHCb is continuously improving towards a better agreement with recorded data. Nevertheless, some discrepancies remain in the distributions of certain variables. For the application of MC events in critical points of the analysis, simulated and recorded data

[1] A track can also be falsely reconstructed from hits of other particles or from detector noise. This results in *ghost tracks*, which is a category that is also excluded from the simulation sample.

have to show identical behaviours for all used variables. Especially the training sample in the multi-variate classification (see Section 4.3.5) and the correction of angular acceptance effects (see Section 5.3) are sensitive to residual discrepancies.
Deviations are observed in variables sensitive to the total track multiplicity of the events as well as in variables that use the transverse component of the momentum [2]. To compensate the disagreements between MC simulation and data, all MC candidates are weighted in the two dimensions of

- number of long tracks in an event
- the transverse momentum of the B^+ candidate,

which is done separately per year and DD or LL category.
As these two variables show no significant correlation, as shown in Fig. A.1, two consecutive one-dimensional weightings are performed and the weights from both dimensions are factorised. The first correction is applied to the dimension of long track multiplicity, the second to the transverse momentum of the B^+ candidate. In both dimensions, the weights are determined from the ratio of the normalised binned distributions in background-subtracted data of $B^+ \to J/\psi K^{*+}$ decays and in simulation. These ratios, *i.e.* the distribution of weights, are shown in Fig. A.2. The effect of the weights on the distributions of p_{T} (B^+) and the number of long tracks are shown in a before-and-after comparison in Fig. 4.1. Examples of the simulation-data agreement in the 2011 samples after the application of the simulation weighting are given in Figs. A.3 and A.4 for variables used in the multi-variate classifier in Section 4.3.5 in the DD and LL category, respectively.

Background subtraction The background subtraction of data, which is utilised for the weighting of simulation events, is applied to preselected [3] data samples using the *sPlot* tool [175]. This method unfolds the signal and background contributions to a distribution in a given variable using the signal and background components of a fit to the distribution of a second variable, which has to be uncorrelated to the former. Weights are assigned to candidates in data using fits to the $m(K^0_{\mathrm{S}}\pi^+\mu^+\mu^-)$ distribution, which results in weights larger than one near the B meson peak and negative weights in the mass sidebands below and above the peak. The mass fits with the signal and background components are introduced in Section 4.3.4. The *sPlot* technique yields the pure signal distributions in the number of long tracks and in p_{T} (B^+) for the simulation weighting. The background-subtraction weights are also used for the comparison plots in Figs. A.3 and A.4.

[2] The transverse momentum in LHCb is defined as the part of the particle's momentum in transverse direction with respect to the beam axis. The total transverse momentum of all initial state particles in pp collisions is (almost) equal to zero by definition.
[3] The term *preselected* refers to data that underwent selection up to and including Section 4.3.3.

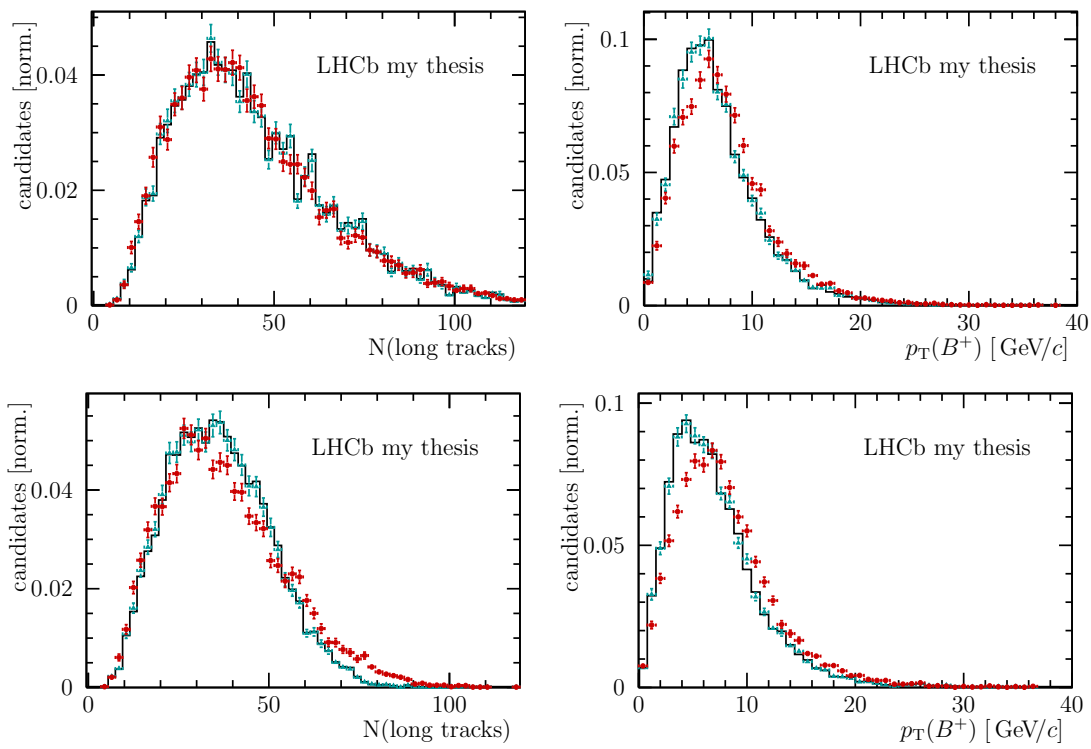

Figure 4.1: Comparison of the distributions in simulation (red) before and (cyan) after the data-simulation weights are applied. Underlaid are the (black) background-subtracted distributions from data of $B^+ \to J/\psi K^{*+}$ decays. Given are distributions of (top) 2012 and (bottom) 2017 simulation in the LL category. The variables are (left) number of long tracks and (right) p_T (B^+). It is obvious, that the weights from the number of long tracks have a bigger impact for Run 2 samples. This can be seen in the different long track distributions for (top left) 2012 and (bottom left) 2017, as well as in the weight distributions in Fig. A.2. This is a direct result of the track multiplicity limitations in the trigger for Run 1 and Run 2, as discussed in a side note in Section 4.3.2. It should be noted that the presented distributions for 2012 (2017) are representative for Run 1 (Run 2); no significant differences between DD and LL samples are observed.

Despite the weighting of the simulation to background-subtracted data, some deviations in the transverse momentum of the K_S^0 candidate in DD samples remain, which are investigated in a systematic study in Section 6.6.1.

The phase-space MC events are additionally weighted to compensate the non-flat initial distribution in the squared dimuon invariant mass, q^2, to generate a flat distribution. For this, weights are obtained from the q^2 distribution of raw signal candidates at the level of simulation generation, by taking the inverse of the generated distribution in q^2 to determine weights, which are then applied to the selected phase-space simulation samples.

4.2 Definition of decay variables

The selection of signal candidates uses *cuts*, *i.e.* requiring variables to be above or below set thresholds, on different variables – either one-dimensional in a single variable or in a multi-dimensional space constructed from many different variables. These variables will be introduced in the following. While some basic variables are directly measured by detectors, such as the energy of a particle, others are constructed from (a combination of) these basic variables. Variables can either be associated to a signal candidate or to the full event.

4.2.1 Momentum and energy

The tracking detectors in combination with the knowledge of the magnet field strength perform the measurement of the momentum for charged tracks using their curvature inside the LHCb magnet. The total momentum, p, plus the vectorial decomposition into x, y and z components are measured. From these, the transverse momentum, p_{T}, as the component in the x-y plane is determined.
The momentum for neutral hadrons and photons is determined from the energy deposition in the calorimeters. The conversion from momentum to energy requires a mass hypothesis for the particle, as described in Sections 3.2.3 and 4.2.2. Energy measurements, with the total value E as well as the transverse component, E_{T}, are given for all particles.
A more complex variable, p_{T}-ConeAsym(B), includes information from the full event and determines the transverse momentum asymmetry within a cone around the flight direction of the B meson:

$$p_{\mathrm{T}}\text{-ConeAsym}(B) = \frac{\sum_i p_{\mathrm{T}i} - \sum_j p_{\mathrm{T}j}}{\sum_i p_{\mathrm{T}i} + \sum_j p_{\mathrm{T}j}}\,, \tag{4.1}$$

with i and j the final state particles of the B meson and all particle tracks inside the cone, respectively [176]. The opening angle of the cone is 1.0(rad) in the η-ϕ-plane. In other words, the variable is a measure of the isolation of a track, quantifying how many particle tracks from outside the signal decay appear inside a cone around the B meson trajectory.

4.2.2 Particle identification

The particle identification system creates flags for when the track is matched to hits in the muon station (`isMuon`) or when a velocity measurement in at least one RICH detector is associated with the track (`hasRich`). The information from the calorimeters, muon stations and RICH detectors are combined to identify particle species as best as possible. Mostly the separation from pions is important for particle identification, as charged pions outnumber all other charged tracks in collisions at the LHC. Hence a set of variables is

defined by the difference in the likelihood of a track to be identified as either a muon, proton or kaon to the likelihood of being a pion. These variables are called $\mathrm{DLL}_{\mu\pi}$, $\mathrm{DLL}_{p\pi}$ and $\mathrm{DLL}_{K\pi}$. Even more sophisticated particle identification variables for muons, pions and kaons are a ANNPID machine learning tool [177, 178] – an artificial neural network with inputs from all detectors – to generate a single identification variable named after the particle species, ProbNN_{μ}, ProbNN_{π} or ProbNN_{K}.

4.2.3 Invariant masses

According to the results of the particle identification, mass hypotheses are assigned to long lived particles in the final state like pions, kaons and muons. Short lived particle masses are determined from the reconstructed four momenta, $P = (E, \vec{p})$, of their decay products.
In general, the mass of a B^+ meson candidate is given by the invariant mass of all five final state particles

$$m(\pi^+\pi^-\pi^+\mu^+\mu^-) = (P_{\pi^+} + P_{\pi^-} + P_{\pi^+} + P_{\mu^+} + P_{\mu^-})^2 \,. \tag{4.2}$$

However, the mass of the K^0_{S} candidate is constrained to the world averaged value [71] using DECAYTREEFITTER [179], resulting in an invariant mass of the B meson

$$m(K^0_{\mathrm{S}}\pi^+\mu^+\mu^-) = (P_{K^0_{\mathrm{S}}} + P_{\pi^+} + P_{\mu^+} + P_{\mu^-})^2 \,. \tag{4.3}$$

In addition to this mass constraint, the flight direction of the reconstructed B^+ candidate is constrained to point back at the primary vertex (PV). With these constraints, all particle momenta are recalculated using the DECAYTREEFITTER. This improves the mass resolution of the B^+ candidate slightly by about $1\,\mathrm{MeV}/c^2$, which at an absolute mass resolution of about $18\,\mathrm{MeV}/c^2$ is a relative improvement of approximately 5%.
Similar to the B^+ meson, the invariant mass of the K^{*+} meson, $m(K^0_{\mathrm{S}}\pi^+)$, is determined from the momenta of the reconstructed decay products. Furthermore, the invariant mass of the dimuon system, $m(\mu^+\mu^-)$, is used in the selection of signal candidates, to discriminate rare signal candidates from the reference sample of $B^+ \to J/\psi\, K^{*+}$ decays and in the extraction of angular observables.

4.2.4 Kinematical variables

Geometrical properties of the reconstructed particle trajectories are used to define and calculate topological variables.
Several definitions of angles between tracks or directions are utilised. Beside the already

introduced helicity angles in Section 2.3.1 – which will be used for the extraction of angular observables – a list of different and more simple angles are used in the candidate selection. The opening angle can either be defined between two tracks or between a track and the beam-pipe (*i.e.* z-axis) and is given in mrad. The cosine of the angle between a short-lived particle flight direction and direction vector between the PV and the particle's decay vertex is called the direction angle (DIRA). In a similar category falls the pseudorapidity, η, defined in Eq. (3.1), which is a measure of the *forwardness* of the trajectory.
Extrapolating the trajectories of particles back towards the PV, the distance of closest approach of a trajectory with the PV is denoted impact parameter (IP), given in mm. As a result of the limited experimental resolution, flight trajectories are assigned an uncertainty, which can relativise the meaning of the IP in absolute units. The χ^2_{IP} is defined as the difference between the χ^2 of the PV reconstructed with and without the track under consideration. More intuitively, χ^2_{IP} behaves almost like

$$\chi^2_{\mathrm{IP}} = \left(\frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}\right)^2 , \tag{4.4}$$

with the uncertainty on the IP, σ_{IP}. Hence, the χ^2_{IP} is used as a discriminator between tracks originating in the PV and those coming from other vertices. For particles that decay inside the volume of the LHCb detector both vertices – one of the particles' generation and one of the decay position – can be reconstructed. The distance between the particle's origin and its decay vertex is called *flight distance* (FD). The χ^2_{FD} for the flight distance is defined in analogy to the behaviour of χ^2_{IP} by the ratio of the absolute flight distance and its uncertainty

$$\chi^2_{\mathrm{FD}} = \left(\frac{\mathrm{FD}}{\sigma_{\mathrm{FD}}}\right)^2 . \tag{4.5}$$

Another χ^2 that is used to quantify the reconstruction quality is determined for vertices, $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$. It is the χ^2 of the vertex fit divided by the number of degrees of freedom of that fit.
The values of all kinematic variables in the selection as well as the determination of the helicity angles (see Section 2.3) are obtained by the DecayTreeFitter algorithm, described in Section 4.2.3.

4.3 Candidate selection

In the following, the selection process can be divided into a five step process:

(1) Trigger decision requirements (Section 4.3.1)

(2) Central event filtering (Section 4.3.2)

(3) One dimensional cut selection (Section 4.3.3)

(4) Invariant mass fits (Section 4.3.4)

(5) Multi-variate classification (Section 4.3.5)

All (recorded and simulated) events undergo these identical selection steps. In addition, the agreement between simulation and data is verified at different steps in the selection.

4.3.1 Trigger decisions

The LHCb trigger system as introduced in Section 3.2.6 is the first step in the candidate selection. Triggers are executed online when collisions take place. Only events passing all three trigger levels are stored for further analysis.
In this analysis, events have to pass the *L0Muon* line of the L0 trigger, which requires a single muon with a transverse momentum greater $p_{\mathrm{T}} > 1.7\,\mathrm{GeV}/c$.
In HLT1, events have to either fulfil the main HLT1 line for hadronic final states or the single muon criteria. During Run 1, the hadron line is named *Hlt1TrackAllL0* which requires well displayed tracks from any PV in the collision with a typical IP of greater 1 mm and in addition a transverse momentum $p_{\mathrm{T}} > 1.6 - 1.7\,\mathrm{GeV}/c$. For Run 2 the line was replaced by *Hlt1TrackMVA*, which uses a multi-dimensional classifier instead of one-dimensional cuts. The dedicated muonic trigger-line *Hlt1TrackMuon* requires the *L0Muon* trigger decision and a high momentum ($p_{\mathrm{T}} > 1.0\,\mathrm{GeV}/c$) track being matched with hits in the muon chambers.
At the last level, the HLT2, events are selected according to multi-variate classifiers using topological input variables. Dedicated lines for two- and three-body final-states for which zero to two [4] tracks have to be identified as muons are applied. The muons have to originate from a displaced vertex. In addition, the line *Hlt2DiMuonDetached* searching for muon tracks originating from displaced decay vertices (SV) is used. Again, all lines at the HLT2 level are used in parallel and at least one line is required to be passed.

4.3.2 Central event filtering

Similar to the parallel algorithms in the trigger system, the stored data is further processed offline by a data stripping system, also using parallel *lines* for different physics analyses. Signal candidates for the $B^+ \to K^{*+}\mu^+\mu^-$ decay are processed by the line `B2XMuMu` which consists of basic kinematic cuts to the particle trajectories. As the naming convention

[4] The topological lines with two muon tracks are only introduced for data taking in 2016 and beyond.

indicates, the line is optimised to select events containing a B meson decay into two muons plus further decay particles. Table 4.1 summarises all numerical values of the cut criteria for the central event filtering (stripping line `B2XMuMu`).
The B meson has to have a good vertex quality, a decay vertex well separated from the PV, and the reconstructed momentum has to point back towards the PV. All daughter particles have to be well separated from the best PV and must not point back to any PV in the event. The muons are required to have transverse momenta in excess of 250 MeV/c and the probability that the tracks are reconstructed from random hits in the tracking detectors must be below 50%. The likelihood to be identified as a muon has to be significantly larger than the likelihood to be a pion. The K^0_S meson has to have a proper decay time larger than 2 ps, a good vertex quality and its daughter pions must not point back at any PV in the event while having a momentum greater than 2 GeV/c.
The invariant mass of the B, K^{*+} and K^0_S mesons have to be inside a given window around the known meson masses. The dimuon invariant mass is limited to an upper value.
The total event must not have more than 600 [5] hits in the SPD detector and is required to have at least one reconstructed PV.

4.3.3 One dimensional cut criteria

The one dimensional cuts in the candidate selection are split into different categories, which are discussed in the following.

Invariant mass: The reconstructed invariant mass of the B meson is required to lie within, $5150 < m(K^0_S\pi^+\mu^+\mu^-) < 6000\,\mathrm{MeV}/c^2$, covering the region around the global average of the B^+ mass, $m(B^+) = 5279.32 \pm 0.14\,\mathrm{MeV}/c^2$ [71]. The low mass cut at $5150\,\mathrm{MeV}/c^2$ is a compromise between not cutting into the low mass tail of the signal peak while removing background from partially reconstructed events. The high upper mass limit of $6000\,\mathrm{MeV}/c^2$ leaves enough background events above to the mass peak as a background proxy for the training of the multi-variate classification.
Candidates with a $K^0_S\pi^+$ invariant mass are accepted inside a window of $\pm 100\,\mathrm{MeV}/c^2$ around the K^{*+} mass, $m(K^{*+}) = 891.76 \pm 0.25\,\mathrm{MeV}/c^2$ [71].
The invariant mass squared of the dimuon system, q^2, is used to separate the reference channel $B^+ \to J/\psi\, K^{*+}$ ($8.68 < q^2 < 10.09\,\mathrm{GeV}^2/c^4$) from the signal channel $B^+ \to K^{*+}\mu^+\mu^-$. Two additional resonances in the q^2 spectrum, $\phi(1020)$ ($0.98 < q^2 < 1.10\,\mathrm{GeV}^2/c^4$) and $\psi(2S)$ ($12.5 < q^2 < 15.0\,\mathrm{GeV}^2/c^4$), are excluded from the evaluations. The remaining non-resonant q^2 regions are binned into eight nominal plus two additional larger q^2 intervals

[5] The higher track multiplicity in collisions during Run 2 requires a stricter cut to the number of hits in the SPD detector to be lower than 450. This cut is implemented in the *L0Muon* trigger line.

Table 4.1: Central selection criteria for the `B2XMuMu` stripping line. The values from the latest version of the stripping line are given, as the values have been constantly adjusted throughout data recording in Run 1 and Run 2 due to continuous optimisation processes.

candidate	selection criteria
$B^{\pm}$	$4700 < m(B^{\pm}) < 7000\,\mathrm{MeV}/c^2$
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 8$
	$\chi^2_{\mathrm{IP}} < 16$
	$\mathrm{DIRA} > 0.9999$
	$\chi^2_{\mathrm{FD}} > 121$
$K^{*\pm}$	$592 < m(K^{*\pm}) < 1192\,\mathrm{MeV}/c^2$
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 12$
	$\mathrm{DIRA} > -0.9$
	$\chi^2_{\mathrm{FD}} > 16$
	$\min(\chi^2_{\mathrm{IP}}) > 9$
$\mu^+\mu^-$	$m(\mu^+\mu^-) < 7100\,\mathrm{MeV}/c^2$
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 12$
	$\mathrm{DIRA} > -0.9$
	$\chi^2_{\mathrm{FD}} > 9$
	$\min(\chi^2_{\mathrm{IP}}) > 9$
$\mu^{\pm}$	$p_{\mathrm{T}} > 250\,\mathrm{MeV}/c$
	track*ghost prob* < 0.5
	$\min(\chi^2_{\mathrm{IP}}) > 9$
	$\mathrm{DLL}_{\mu\pi} > -3$
	`isMuon`
K^0_{S}	$468 < m(K^0_{\mathrm{S}}) < 518\,\mathrm{MeV}/c^2$
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 25$
	$\tau > 2\,\mathrm{ps}$
$\pi^{\pm}$ from K^0_{S}	$p > 2\,\mathrm{GeV}/c^2$
	$\min(\chi^2_{\mathrm{IP}}) > 9$ (LL) or 4 (DD)
event	SPD hits < 600
	at least one PV

in which the angular observables are measured. The numerical values of the q^2 intervals are given in Table 4.2.

Particle identification All pions are required to have information from the RICH systems (the `hasRich` flag has to be true) and both muon candidates are required to be tagged as muons by the muon system (the `isMuon` flag has to be true). Both muons also are required to pass a cut on a global particle identification classifier: $\text{ProbNN}_\mu > 0.25$. All these particle identification criteria are very efficient and remove only candidates that contain clearly misidentified π or μ tracks.

Decay kinematics: To avoid modelling the reconstruction efficiency of tracks that are close to the outer edge of the detector active region, all signal candidate tracks are required to have a polar angle θ (see Section 3.2) smaller than 400.0 mrad. Furthermore, to ensure a significant spacial separation of their VELO hits, the track pairs of $\mu^+\mu^-$ and the three track pairs of the three final state pions must have a minimum opening angle of $\theta_{\text{pair}} > 1\,\text{mrad}$ with respect to each other. Latter cuts only remove less than one per mille of all candidates and do not warp the angular distributions.

Double misidentification: Double misidentification of a muon and a pion from the $B^+ \to J/\psi\, K^{*+}$ decay changes the q^2 value of the candidate, while preserving the invariant mass $m(K^0_S\pi^+\mu^+\mu^-)$ to be consistent with a B^+ decay. Therefore, the resonant decay can end up in the non-resonant q^2 region and fake a signal $B^+ \to K^{*+}\mu^+\mu^-$ decay.
This double misidentification pollution is probed by reverting the mass hypotheses of the π^+ and μ^+ candidates and evaluating the invariant masses $m(\mu^-[\pi^+]_{m_\mu})$ and $m(K^0_S[\mu^+]_{m_\pi})$. A two-dimensional cut to candidates with

$$\begin{aligned} \left|m(\mu^-[\pi^+]_{m_\mu}) - m(J/\psi)\right| &< 50\,\text{MeV}/c^2 \text{ and} \\ \left|m(K^0_S[\mu^+]_{m_\pi}) - m(K^{*+})\right| &< 30\,\text{MeV}/c^2 \end{aligned} \tag{4.6}$$

is applied to remove any source of double misidentification pollution. At the same time, the signal efficiency of this cut is determined on simulation to above 99.9%.
The higher order $c\bar{c}$ resonance, $\psi(2S)$, is tested for a potential pollution into the signal region, but due to a significantly smaller branching fraction (of about a factor 15 compared to the J/ψ resonance [71]), no contamination is observed.

Veto against peaking background structure: One source of background in the mass region above the B^+ mass peak originates from the rare decay $B^0 \to K^0_S\mu^+\mu^-$ combined with a (slow) random pion. A cut around the invariant mass of $B^0 \to K^0_S\mu^+\mu^-$ is therefore applied and events with $|m(K^0_S\mu^+\mu^-) - m(B^0)| < 50\,\text{MeV}/c^2$ are dismissed.

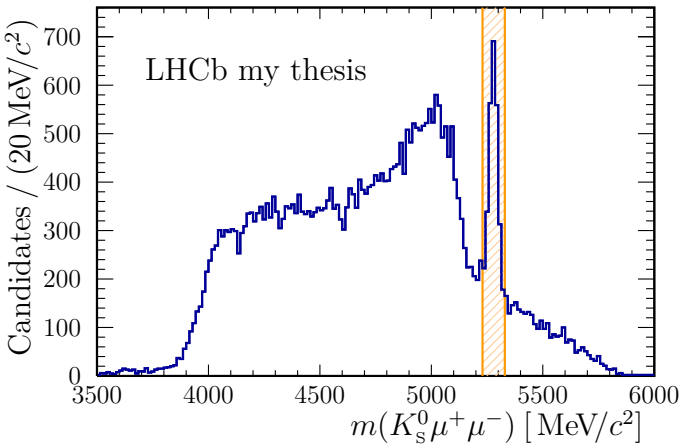

Figure 4.2: Invariant mass of the two muons and the K^0_S in the data set of 2017 with magnet polarity *up*. A clear peak at the mass of the B^0 meson is visible and overlain by a box indicating the vetoed mass region. Together with a random π^+ from the collision or fiducial particle decays, $B^0 \to K^0_S\mu^+\mu^-$ decays can mimic a signal decay.

The distribution of the invariant mass of $B^0 \to K^0_S\mu^+\mu^-$ and the discarded mass window are shown in Fig. 4.2 for 2017 events recorded with magnet polarity *up* as an example; distributions from other years and magnet polarities are nearly identical. On simulated signal events, the efficiency of this veto is found to be 99.8% and larger.
This veto completes the selection criteria cuts to one-dimensional variables. Events that pass all filters continue towards a selection using a multi-variate classification, taking a detour via fits to the invariant mass distributions, $m(K^0_S\pi^+\mu^+\mu^-)$.

4.3.4 Invariant mass fits

Fits to the invariant mass, $m(K^0_S\pi^+\mu^+\mu^-)$, of data in the q^2 region of the J/ψ resonance ($8.68 < q^2 < 10.09\,\mathrm{GeV}^2/c^4$) after the removal of peaking background contributions are used for the matching of simulated and reconstructed data (Section 4.1) and to estimate the number of signal candidates in the q^2 region of the signal channel. Furthermore, it is exploited to optimise the cut to the response of the multi-variate classifier (see Section 4.3.5).
The model used to fit the distribution comprises a two-tailed Crystal Ball (CB) [180] function for the signal and an exponential function for the background component. The CB function is an empirical description of a mass peak at x_0 with a Gaussian width, σ. However, one or both sides of the Gaussian bell are cut off at values $\alpha_{1,2}$ in units of σ. The tails are parametrised by an exponential tail of power $n_{1,2}$ to describe radiative tails due to losses from final state radiation. With these parameters, the function is defined to

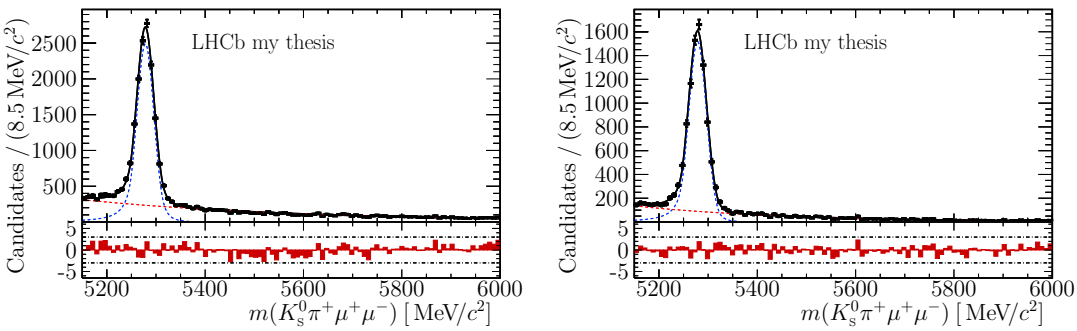

Figure 4.3: Invariant mass distribution of the B^+ candidate in $B^+ \to J/\psi K^{*+}$ decays ($8.68 < q^2 < 10.09\,\mathrm{GeV}^2/c^4$) in 2016 (left) DD and (right) LL samples after the cut-based selection. Details on the model of the added fit are given in the text.

be steady at the transition points $x_0 \pm \sigma\alpha_{1,2}$ by

$$\mathcal{P}_{CB} = N \cdot \begin{cases} A_1 \cdot \left(1 - \frac{\alpha_1}{n_1}(\alpha_1 + t)\right)^{-n_1} & \text{for} \quad t \leq -\alpha_1 \\ \exp\left(\frac{-t^2}{2}\right) & \text{for} \quad -\alpha_1 < t < \alpha_2 \\ A_2 \cdot \left(1 - \frac{\alpha_2}{n_2}(\alpha_2 + t)\right)^{-n_2} & \text{for} \quad \alpha_2 \leq t \end{cases}, \tag{4.7}$$

with the normalisation, N, the coefficient

$$A_{1,2} = \exp\left(\frac{-\alpha_{1,2}^2}{2}\right) \quad \text{and} \quad t = \frac{x - x_0}{\sigma}\,. \tag{4.8}$$

The shape of the signal peak is obtained from a fit to the corresponding simulation sample and the parameters $n_{1,2}$ and $\alpha_{1,2}$ are fixed to the values measured in simulation. The mean position and the width of the CB functions are floating in the fit.
The background model is a single exponential using only one free parameter. This model describes combinatorial background, that arises from the accidental combination of tracks originating from different decay chains. The random combination of tracks results in an exponential distribution in the invariant mass, $m(K_S^0\pi^+\mu^+\mu^-)$, with falling slope.
After after cut-based selection, including trigger- and stripping-condition, Fig. 4.3 shows the results of fits to $B^+ \to J/\psi K^{*+}$ 2016 data; the DD sample on the left and LL sample on the right. In these projections, the veto cut against B^0 plus random π^+, which cuts away some combinatorial background in the mass region 5400–5600 MeV/c^2, is visible as a signature in the pulls below the projection. The dent is indubitable caused by the veto, as it disappears when not vetoing the B^0 plus random π^+ background. This small dent is however negligible for the usage of the *sPlot* technique during MC event weighting or for

Table 4.2: Estimated signal and background yields after the cut-based selection. The signal values are determined from the yield of the reference channel $B^+ \to J/\psi K^{*+}$, as shown in Fig. 4.3. The signal yield is scaled by the ratio of branching ratios in Eq. (4.9). The background yield is taken from signal channel directly, by constraining the total number of candidates to be the sum of signal and background yield.

sample	signal estimation	background estimation
Run 1 DD	160	2040
Run 1 LL	100	450
Run 2 DD	490	4810
Run 2 LL	290	1020

the estimation of signal candidates in the $B^+ \to K^{*+}\mu^+\mu^-$ decay. For the extraction of angular observables a systematic study addresses the effect of this veto in Section 6.5.

The selection efficiency is measured to be fairly constant throughout the full q^2 range, which is in agreement with previous results in Ref. [7] and hence a similar selection efficiency can be assumed for signal and reference channel decays during the optimisation of the selection efficiency. With this, the signal yield obtained from a fit to the $m(K^0_S\pi^+\mu^+\mu^-)$ distribution in the q^2 region of the J/ψ resonance at $8.68 < q^2 < 10.09\,\mathrm{GeV}^2/c^4$ is scaled by the ratio of the rare and resonant branching fractions,

$$\begin{aligned} \mathcal{B}\Big(B^+ \to K^{*+}\mu^+\mu^-\Big) &= 8.7 \cdot 10^{-7} \quad \text{and} \\ \mathcal{B}\Big(B^+ \to K^{*+}\,[\mu^+\mu^-]_{J/\psi}\Big) &= 8.5 \cdot 10^{-5}\,, \end{aligned} \tag{4.9}$$

to estimate the number of signal candidates in the $B^+ \to K^{*+}\mu^+\mu^-$ data. The branching fractions are given by theoretical predictions for the signal decay in the non-resonant q^2 region [101] and by measurements for the reference decay [71, 181, 182]. The number of estimated signal candidates after the cut-based candidate selection and prior to the multi-variate classifier (see Section 4.3.5) are given in Table 4.2.

4.3.5 Multi-variate classification

A multi-variate analysis (MVA) is performed with the goal of removing residual combinatorial background from the data by separating background and signal candidates. The MVA uses multi-dimensional decision trees constructed on a set of input variables to create a multi-dimensional classification that determines a single response variable for each candidate.

The decision trees are trained on two proxies, a pure background and a pure signal sample.

Table 4.3: After cutting in the B meson mass and q^2, the following number of events are available for the MVA training and testing. The signal sample consists of candidates from data mass sidebands, while background training samples are taken from simulation. The numbers are split equally in half for testing and training.

	Run 1 DD	Run 1 LL	Run 2 DD	Run 2 LL
Signal	32,700	16,300	61,300	33,700
Background	9000	1900	21,500	5100

By iterating the decision trees and applying weights to each branch of the trees for each iteration, the significance in separation is significantly increased throughout the iteration process. This iterative weighting is called *boosting*, therefore the used methods are also referred to as *boosted decision trees* [183]. The exact method applied in this analysis is a boosted decision tree with *gradient boost* (BDTG) [184] and uses 450 decision trees and a maximal depth of two layers. The gradient boosting is less sensitive to statistical fluctuations in the training process, as the total sample is randomly split during the iterations of the training process. After the training, the obtained classifiers are tested on a test sample, which is statistically independent of the training sample. This testing evaluates possible over-training, when the boosted decision tree exploits features of the training sample that are not representative for the total sample. A typical source of over-training are statistical fluctuations of the boosting, which are critically enhanced due to the iterative weighting. When testing and training responses are in agreement, the obtained classification is applied to all candidates to calculate a response variable for each candidate. Finally, a cut in the response variable is made. This cut is optimised to increase the statistical power of the data.
Each of the four data sub-sets, Run 1 DD to Run 2 LL, undergoes individual training and testing as well as cut optimisation. The BDTG method is implemented and applied using the build-in-ROOT [185, 186] TMVA tool-kit [187].

Training and testing samples

Simulation with weights from Section 4.1 are used as signal proxy for the training and testing of the MVA. The invariant mass, $m(K^0_S\pi^+\mu^+\mu^-)$, is required to be within a $\pm 100\,\mathrm{MeV}/c^2$ window around the known B^+ mass of $5279.34 \pm 0.12\,\mathrm{MeV}/c^2$ [71].
The background sample is taken from data. Candidates with an invariant mass greater than $5400\,\mathrm{MeV}/c^2$ are taken for the background proxy. This mass region is well separated from the signal region close to the B^+ mass. Since the background sample at this selection step consists of purely combinatorial background, it is fair to assume that the properties of

Table 4.4: List of the input variables used in the MVA training for the DD and LL samples. The DD (LL) category uses twelve (nine) variables, which are indicated by the cross-mark. The variables are identical in Run 1 and Run 2. The lists are ordered by variable categories. All decay variable are introduced in Section 4.2.

variable	DD	LL
$\ln(\chi^2_{\mathrm{IP}}(B^+))$	✓	✓
$\ln(\chi^2_{\mathrm{IP}}(K^{*+}))$	✓	
$\ln(\chi^2_{\mathrm{IP}}(\mu^+))$	✓	✓
$\ln(\chi^2_{\mathrm{IP}}(\mu^-))$	✓	✓
$\ln(\chi^2_{\mathrm{IP}}(\pi^+))$	✓	
$\mathrm{IP}(B^+)$	✓	
$\chi^2_{\mathrm{vtx}}(B^+)$		✓
$\ln(1-\mathrm{DIRA}(B^+))$	✓	
$\eta(\pi^+)$	✓	
flight distance(B^+)		✓
flight distance(K^0_{S})	✓	✓
$p_{\mathrm{T}}(K^0_{\mathrm{S}})$	✓	✓
$p_{\mathrm{T}}(B^+)$	✓	✓
p_{T}-ConeAsym(B^+)	✓	✓

other variables do not depend on the invariant mass, $m(K^0_{\mathrm{S}}\pi^+\mu^+\mu^-)$ [6], and can therefore be extrapolated from the mass sidebands to the region underneath the signal peak. The threshold of 5400 MeV/c^2 for the background sample is chosen such that the training sample of the MVA is as large as possible but to guarantee not to include signal candidates in the background proxy. In both, signal and background proxy, the q^2 regions of the J/ψ (8.68–10.09 GeV$^2/c^4$) and the $\psi(2S)$ (12.9–14.4 GeV$^2/c^4$) resonance are excluded.
In Table 4.3 the statistics of background and signal input samples are given. Each sample is split in half to be used as training and testing samples, respectively.

Input variables

Run 1 and Run 2 samples are trained individually but use the exact same set of input variables for the MVA, while some variables differ between DD and LL samples. Different variables are used due to their different discrimination power and their (dis)agreement

[6] This assumption is tested and verified for the angular dimensions after the candidate selection is finalised.

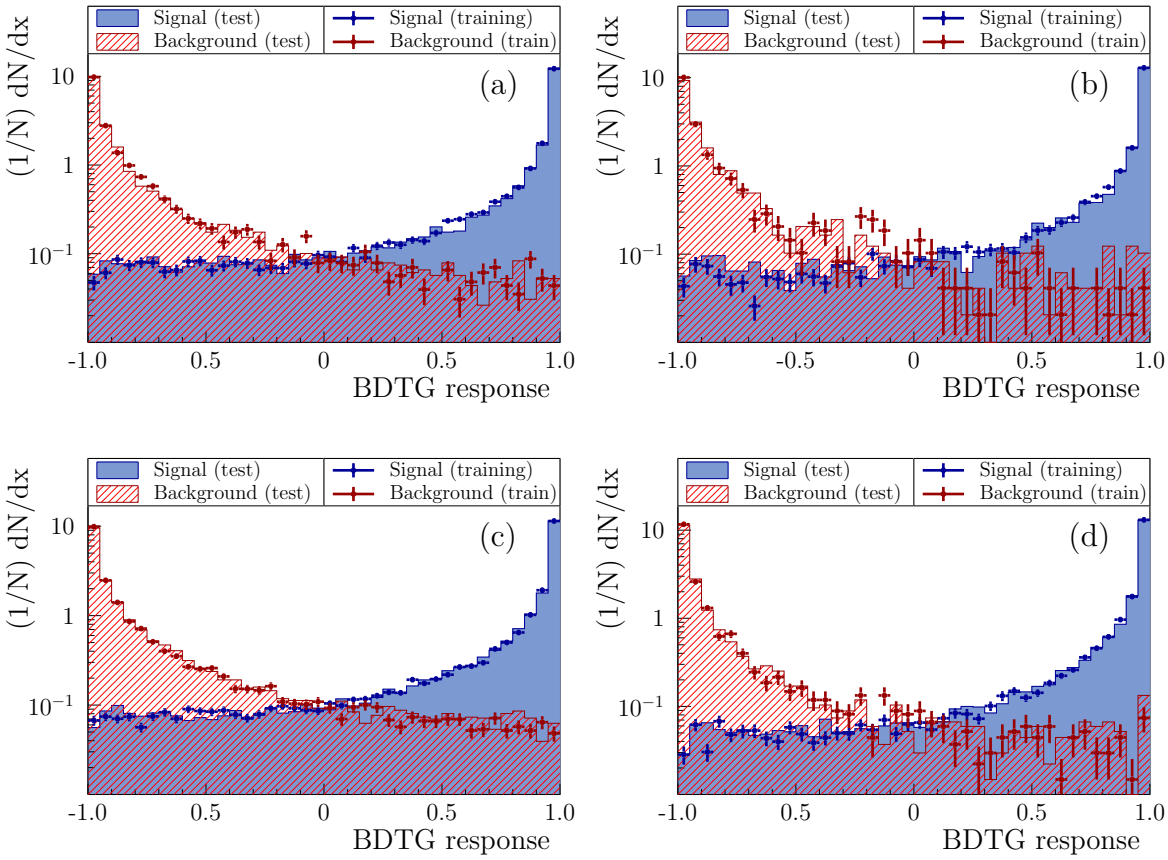

Figure 4.4: Response of the multi-variate classification (BDTG) for background and signal distributions of the Run 1 DD (a), Run 1 LL (b), Run 2 DD (c) and Run 2 LL (d) categories. Training and testing sample of the multi-variate classifier output show very good agreement in all samples, for both signal and background distributions. A significant separation power between signal and background is demonstrated by the small overlap of signal and background response distributions.

between data and simulation in the sub-sets. The variables for the DD and LL MVA training are listed in Table 4.4 and a definition of all variables is given in Section 4.2. Correlations between the input variables are tested to be sufficiently independent. In the process of selecting suitable input variables for the MVA, when finding two variables with a high correlation, the one with weaker separation power is dismissed in order to keep the training as simple as possible while preserving the separation power.

Results of the MVA selection

The overlay of training and testing results for the BDTG method are shown in Fig. 4.4. All four background and signal distributions for Run 1 and Run 2 with DD and LL categories show no signs of over-training, as training and testing distributions agree well in all cases. Fig. 4.5 shows the resulting separation power of the four classifier. All curves outperform a

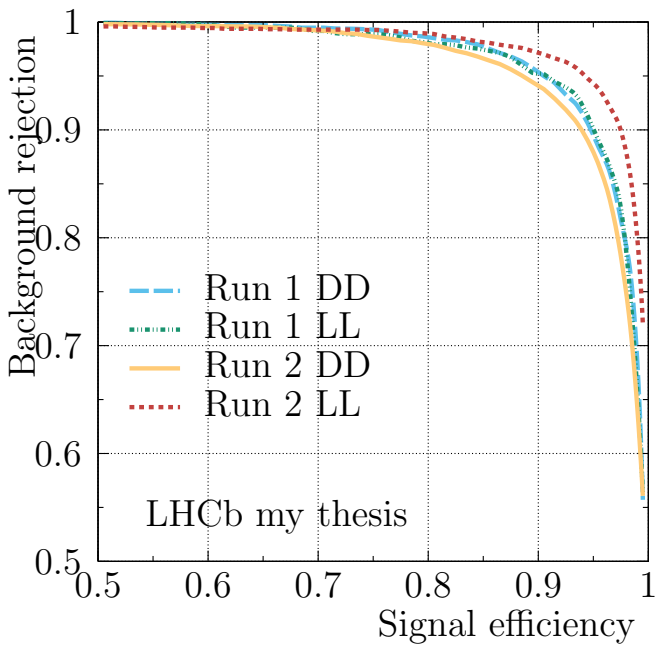

Figure 4.5: The background rejection capabilities of the BDTG as a function of its signal efficiency for the four categories. The possible cut value in the BDTG response increases from top left to bottom right through the curves. The excellent performance of the classifier to achieve both signal efficiency and background rejection well above 90% helps to separate background from signal candidates.

background rejection of 90% at 90% signal efficiency. Using the estimated number of signal and background candidates in data from Table 4.2 along with the signal and background efficiencies of the MVA, the figure of merit $S/\sqrt{S+B}$ is calculated as a function of the cut to the BDTG response variable. The figure of merit is optimised with the goal to maximise the statistical power of the sample.
The optimum cut values are determined to be 0.8 (Run 1 DD), 0.5 (Run 1 LL), 0.65 (Run 2 DD) and 0.4 (Run 2 LL).

4.4 Selection conclusions

In Fig. 4.6 all selected candidates of Run 1 and Run 2 are shown in the two dimensional plane of the invariant masses of $m(K^0_S\pi^+\mu^+\mu^-)$ and $m(\mu^+\mu^-)$. The red grid marks the region in which signal candidates of the $B^+ \to K^{*+}\mu^+\mu^-$ are expected to show. Clear structures from the $B^+ \to K^{*+}\mu^+\mu^-$ decay candidates are visible in a vertical band of clustered entries.
Beyond the candidate selection process, the data are binned in intervals of the dimuon invariant mass square, q^2. The definition of the ten q^2 intervals along with the yield of signal candidates in each interval are given in Table 4.5. The numbers for the signal

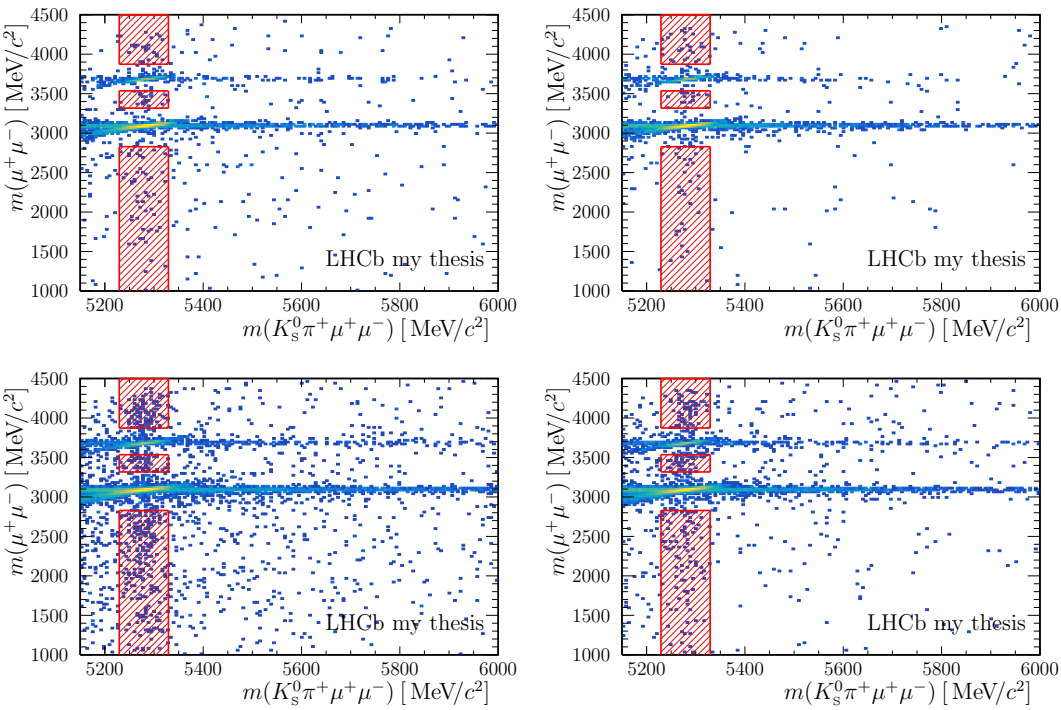

Figure 4.6: Invariant mass of the two muon system versus the invariant mass of the B meson after the final candidate selection is applied in (top) Run 1 and (bottom) Run 2, with (left) DD and (right) LL samples. The colour scheme is normalized to the total number of candidates. In all plots, the $c\bar{c}$ resonances are visible as horizontal bands of high occupation. The enhanced structure of the $B^+ \to K^{*+}\mu^+\mu^-$ decay inside the red boxes is visible.

candidates are obtained from fits to the $m(K^0_S\pi^+\mu^+\mu^-)$ distribution of data in each q^2 interval, as shown in Fig. 4.8. The fit model for signal and background components is previously introduced in Section 4.3.4. Fig. 4.7 shows the combination of all eight q^2 intervals in one projection. For this plot, the full q^2 range of the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay, $0.1 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$, excluding the $\phi(1020)$, J/ψ and $\psi(2S)$ resonances, is accumulated. The projection is the sum of all projections in Fig. 4.8. From adding these fits the total number of signal candidates is calculated to be 737 ± 34.

In summary, the selection successfully reduced the contribution of background to the sample. Peaking backgrounds from particle misidentification or pollution from other decays are removed specifically by corresponding vetoes. The topological criteria and the use of a multi-variate classifier significantly suppress the combinatorial background. The resulting number of signal candidates is sufficient to perform an angular analysis, which however requires some special techniques to obtain all angular observables and their correlations. These techniques – along with the complete maximum-likelihood fit framework – are discussed in detail in the following Chapter 5.

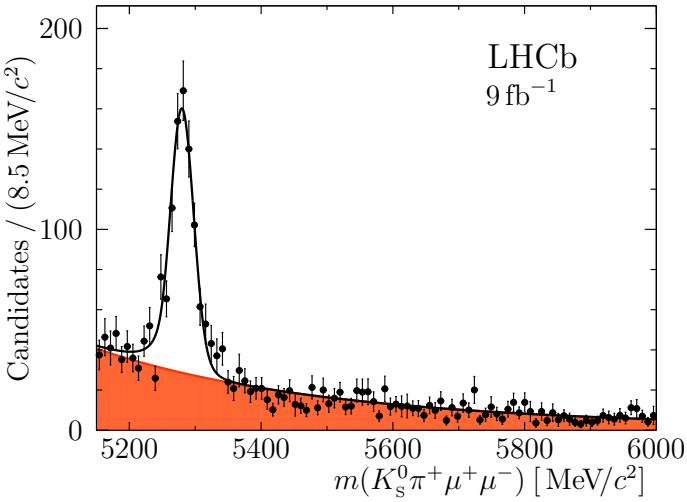

Figure 4.7: Invariant mass $m(K^0_S\pi^+\mu^+\mu^-)$ distribution of selected signal candidates from the full q^2 range without resonant regions of the combined Run 1 and Run 2 data [1]. A double-sided Crystal Ball function is used for the signal. A single exponential for the background component in orange shade is added. From this fit, the number of signal candidates, 737 ± 34, is extracted.

Table 4.5: Combined yield of signal candidates from the final fit for all eight q^2 bins and for the full data set of combined Run 1 and Run 2 and DD and LL categories. Both sets of numbers, signal and background candidates, are given over the invariant mass range of the B meson, 5150–6000 MeV/c^2. The mass fits to obtain these numbers are given in Figs. 4.7 and 4.8. The numbers for the two large q^2 intervals are summed up from the corresponding nominal intervals and are obviously not double-counted for the total numbers.

q^2 interval [GeV^2/c^4]	signal	background
$[0.1, 0.98]$	101.6 ± 12.4	186.1 ± 15.4
$[1.1, 2.5]$	49.0 ± 9.9	219.2 ± 16.5
$[2.5, 4.0]$	42.2 ± 9.7	274.9 ± 18.1
$[4.0, 6.0]$	108.8 ± 13.4	286.9 ± 19.0
$[6.0, 8.0]$	104.6 ± 13.9	297.7 ± 19.7
$[11.0, 12.5]$	110.6 ± 12.9	170.8 ± 15.0
$[15.0, 17.0]$	144.2 ± 13.4	108.7 ± 12.0
$[17.0, 19.0]$	75.7 ± 9.8	106.0 ± 11.2
$[1.1, 6.0]$	199.8 ± 19.4	781.1 ± 31.0
$[15.0, 19.0]$	220.0 ± 16.6	214.8 ± 16.5
total	737 ± 34	1650 ± 46

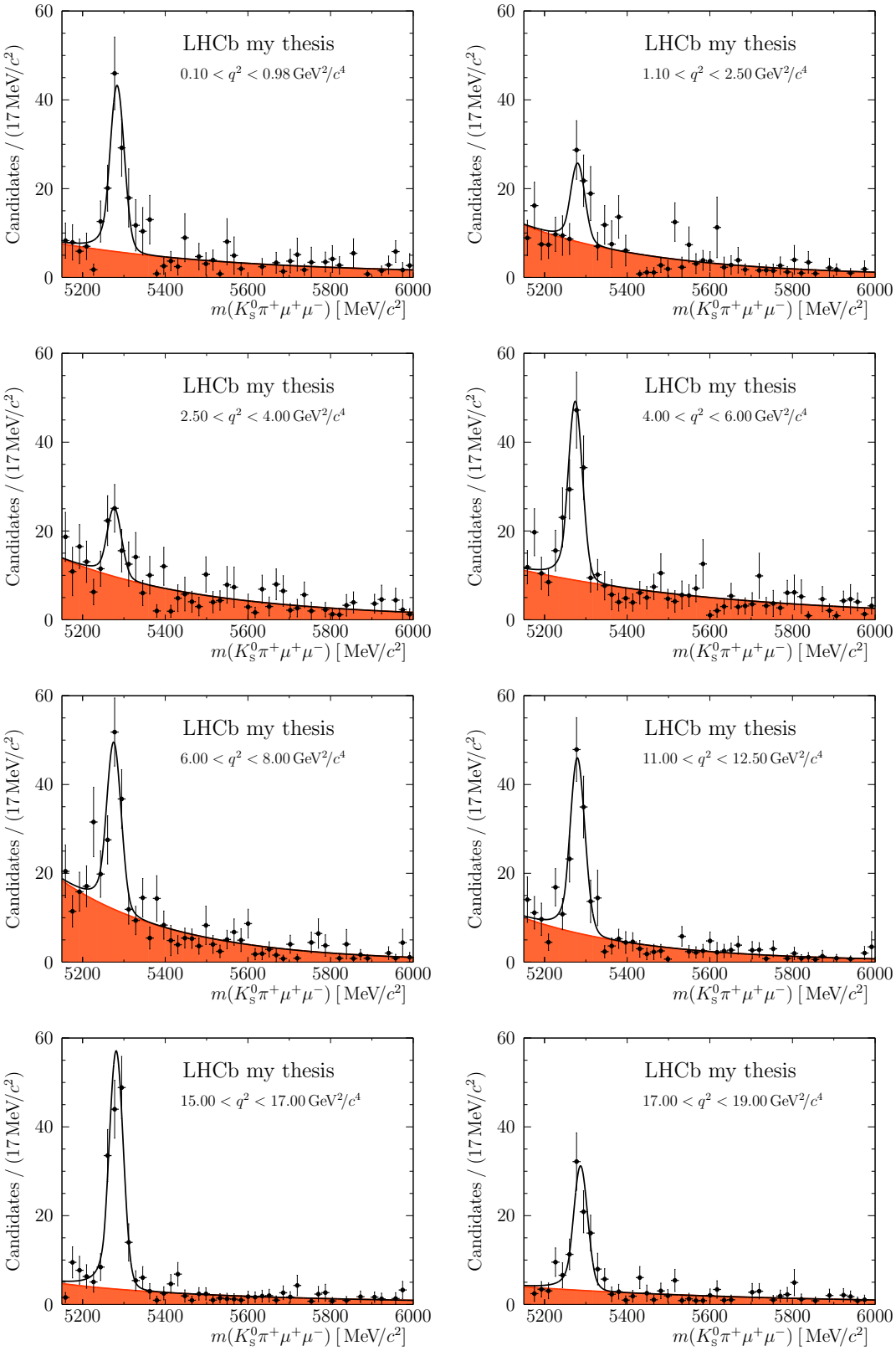

Figure 4.8: Invariant mass $m(K^0_S\pi^+\mu^+\mu^-)$ distributions of selected signal candidates from the combined Run 1 and Run 2 data. The data are split into the nominal eight q^2 intervals, with the ranges stated. A double-sided Crystal Ball function is used for the signal. A single exponential for the background component in orange shade is added. To emphasize the differences in signal yield throughout the q^2 intervals, the scale of the y-axes of all plots are identical. The numerical values of the yield of signal candidates in all q^2 intervals are given in Table 4.5.

5 Determination of angular observables

The two sets of angular observables introduced in Section 2.3.2 are extracted from selected data in intervals of q^2 by maximum-likelihood fits. The definition and functionality of fits using a maximum-likelihood estimator are introduced thoroughly in Section 5.1. Afterwards, the exact configurations of the fits, that are used for the determination of the results presented in Chapter 7, are discussed in Sections 5.2 and 5.6.
For these maximum-likelihood fits, several statistical techniques and methods are used to optimise the precision on the angular observable extraction. Acceptance effects in the candidate selection require corrections, which are explained in Section 5.3. Limited statistics in the signal decay data prevent the full set of angular observables to be determined simultaneously. To handle the low statistics, angular folding techniques in Section 5.4 are utilised to reduce the complexity of the fits. As a direct consequence of these non-simultaneous measurements of the angular observables, the determination of the correlation matrices in Section 5.5 is only possible via a bootstrapping method.
The chapter concludes with a general validation of the maximum-likelihood fit and the dedicated analysis methods, using the reference decay channel $B^+ \to J/\psi\, K^{*+}$ data, simulation data and *pseudoexperiments* [1].

5.1 Maximum likelihood estimation

Assuming a set of measurements, $\boldsymbol{X} = (X_1, X_2, ..., X_N)$, with observables $\boldsymbol{x} = (x_1, x_2, ...)$ measured for every X_i. The observables $\boldsymbol{x}$ are thought to be distributed according to a normalised probability density function, $\mathcal{P}(\boldsymbol{x}|\boldsymbol{\theta})$. When at least one of the models' parameters, $\boldsymbol{\theta} = (\theta_1, \theta_2, ...)$ is unknown, a maximum likelihood estimator can be utilised to determine the most probable values for these unknown parameters from the observed

[1] Pseudoexperiments are the studies of primitively generated simulation data – referred to as *pseudodata* – which is directly created by the fit program using the same parameters and fit model configurations as in the maximum-likelihood fit. These pseudoexperiments are discussed in detail in Section 5.7.1.

measurement results [188]. This estimation utilises the likelihood function [189] [2], which is the product of the probability density functions for all measurements evaluated at this point in the parameters space

$$\mathcal{L}(\boldsymbol{X}|\boldsymbol{\theta}) = \prod_{e=1}^{N} \mathcal{P}(\boldsymbol{x}_e|\boldsymbol{\theta}) \,. \tag{5.1}$$

Using Bayes' theorem [191] for the relations of conditional probabilities, one obtains the relation between $p(\boldsymbol{\theta}|\boldsymbol{X})$, the probability to obtain $\boldsymbol{\theta}$ given $\boldsymbol{X}$, and the likelihood function

$$p(\boldsymbol{\theta}|\boldsymbol{X}) = \frac{\mathcal{L}(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{X})} \,. \tag{5.2}$$

As the prior distribution of the parameters, $p(\boldsymbol{\theta})$ is uniform and $p(\boldsymbol{x})$ is independent from changes of the parameters, the maximum likelihood estimator coincides with the most probable Bayesian estimator. This estimator is optimised by maximising the likelihood function in Eq. (5.1).
In practise, the likelihood function is maximised by minimising the negative logarithm, the log-likelihood function

$$-2\log\mathcal{L}(\boldsymbol{X}|\boldsymbol{\theta}) = -2\sum_{e=1}^{N} \log\mathcal{P}(\boldsymbol{x}_e|\boldsymbol{\theta}) \,. \tag{5.3}$$

In the concrete case of this analysis, the input variables $\boldsymbol{x} = \left(\vec{m}, \vec{\Omega}\right)$ consist of the invariant B and K^{*+} meson masses – combined into $\vec{m}$ – and the decay angles, $\vec{\Omega} = (\cos\theta_K, \cos\theta_\ell, \phi)$. The set of fit parameters can be further divided into three groups of fit parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_\mathrm{P}, \boldsymbol{\theta}_\mathrm{S}, \boldsymbol{\theta}_\mathrm{nuis})$, with the unknown angular observables, $\boldsymbol{\theta}_\mathrm{P}$ and $\boldsymbol{\theta}_\mathrm{S}$ of the P- and S-wave, as well as additional nuisance parameters, $\boldsymbol{\theta}_\mathrm{nuis}$. With these input variables and fit parameters, and by considering the weights [3] of all candidates, w_e, the log-likelihood function from Eq. (5.3) expands to the form

$$-2\log\mathcal{L}_\mathrm{pdf} = -2\sum_{n=1}^{4}\sum_{e=1}^{N_n} w_e \log\left[\mathcal{P}_n\left(\vec{m}_e, \vec{\Omega}_e \middle| \boldsymbol{\theta}_\mathrm{P}, \boldsymbol{\theta}_\mathrm{S}, \boldsymbol{\theta}_{\mathrm{nuis},n}\right)\right] \,, \tag{5.4}$$

[2] The terms "likelihood" and "probability" have distinct meanings in statistics, which are different to what is used in common language. The probability of a sample quantifies how probable it is to obtain this particular sample given a set of values for the parameters of a probability density function. It is a property of the sample. The likelihood is a property of a point in the parameter space [190]. The numerical values of the probability and likelihood are equal.

[3] In the nominal fit configuration, the weights comprise purely the correction factor for angular acceptance effects, as detailed in Section 5.3, but may contain further contributions from systematic studies as in Sections 6.1.3, 6.6.1 and 6.6.2.

including the sum over all four sample categories of index n: Run 1 DD, Run 1 LL, Run 2 DD and Run 2 LL. The probability density functions are individual for every category. The values of the angular observables of the P-wave, $\boldsymbol{\theta}_{\mathrm{P}}$, which can be either in the S_i or $P_i^{(\prime)}$ basis, as well as the observables of the S-wave, $\boldsymbol{\theta}_{\mathrm{S}}$, including the S-wave fraction, F_{S}, are common in all categories. Other nuisance parameters – varied or fixed – can take different values in every category.
All varied parameters can be constraint to a previously measured value and uncertainty. For this, additional terms for every constraint parameter with value p_j are added to the likelihood

$$-2\log\mathcal{L}_{\mathrm{con}} = \sum_j \left(\frac{p_j - p_{j,0}}{\sigma_j}\right)^2 , \tag{5.5}$$

with the previous measurement, $p_{j,0}$, and uncertainty, σ_j.
A similar approach is used to determine the absolute numbers of signal and background yields, as they are presented in Table 4.5. For this, the signal fraction, f_{sig}, is defined as the ratio of signal and background yields

$$f_{\mathrm{sig}} = \frac{N_{\mathrm{sig}}}{N_{\mathrm{sig}} + N_{\mathrm{bkg}}} . \tag{5.6}$$

Then, a Poissonian constraint on the sum of the signal and background yields to match the total number of weighted candidates generates another contribution to the likelihood,

$$-2\log\mathcal{L}_{\mathrm{ext}} = -2\log\left[\mathcal{P}_{\mathrm{Poisson}}\left(N_{\mathrm{sig}} + N_{\mathrm{bkg}}\middle|\sum_e w_e\right)\right] . \tag{5.7}$$

The terms from the fit models in Eq. (5.4) and (if applicable) the extended and constraint terms from Eqs. (5.5) and (5.7) are summed up to the total likelihood

$$-2\log\mathcal{L}_{\mathrm{tot}} = -2\log\mathcal{L}_{\mathrm{pdf}} - 2\log\mathcal{L}_{\mathrm{con}} - 2\log\mathcal{L}_{\mathrm{ext}} . \tag{5.8}$$

The minimisation of the negative log-likelihood in Eq. (5.8) is executed by the Minuit [192] implementation within the Root [185, 186] software package. The FCNCfitter framework, which is used for this analysis and that utilises the Minuit class, was developed by colleagues in the LHCb collaboration for the analyses in Refs. [26, 27]. The program was further developed, to match the requirements of this analysis and to fully integrate the simultaneous fitting of multiple data sets, the angular folding techniques and the determination of correlation matrices with a bootstrapping method.

5.2 Fit model

In the following, the composition of probability density functions into a fit model is explained.
Generally, the fit model consists of an ensemble of signal and background probability density functions [4]. The products of both components are added together by

$$\mathcal{P} = f_{\text{sig}} \prod_d \mathcal{P}_{\text{sig},d} + (1 - f_{\text{sig}}) \prod_d \mathcal{P}_{\text{bkg},d} \,, \tag{5.9}$$

with the signal fraction, f_{sig}. The number of signal and background components depends on the included fit dimensions, indicated by the index d. The used configurations in this analysis are a two-dimensional fit of the two invariant masses, $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$ and $m(K^0_{\text{S}}\pi^+)$ as well as a four-dimensional fit to $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$ and the angular space, $\vec{\Omega}$. The models for the invariant mass of the B meson, $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$, the invariant mass of the K^{*+} meson, $m(K^0_{\text{S}}\pi^+)$, and the three-dimensional angular space, $\vec{\Omega} = (\cos\theta_K, \cos\theta_\ell, \phi)$ are discussed in the following paragraphs.

Signal components: The reconstructed invariant mass of the B^+ meson, $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$, is fitted by a two-tailed CB as introduced in Eq. (4.7). The parameters $\alpha_{1,2}$, $n_{1,2}$ and the width of the Crystal Ball function are fixed to values obtained in simulation, with an additional scaling factor of about 1.15 for the width that is obtained by comparing fits to data and simulation in the reference channel. The angular dimensions are described by the combined differential decay rate of P- and S-wave contributions, given in Eq. (2.26).

Not yet introduced is the fit model for the invariant mass of the K^{*+} meson, $m(K^0_{\text{S}}\pi^+)$. The amplitude of the P-wave component is described by a Breit-Wigner [193] model

$$\begin{aligned}\mathcal{A}_{\text{P}}\left(m(K^0_{\text{S}}\pi^+)\right) =& \sqrt{kp} \times B'_{L_B}(k, k_0, d) \left(\frac{k}{m_{B^+}}\right)^{L_B} \times B'_{L_{K^{*+}}}(p, p_0, d) \left(\frac{p}{m_{K^{*+}}}\right)^{L_{K^{*+}}} \\ &\times \frac{1}{{m_{K^{*+}}}^2 - m(K^0_{\text{S}}\pi^+)^2 - i m_{K^{*+}} \Gamma\left(m(K^0_{\text{S}}\pi^+)\right)} \,,\end{aligned} \tag{5.10}$$

with the momentum k (p) of the K^{*+} (K^0_{S}) in the rest-frame of B^+ (K^{*+}) and k_0 (p_0), the corresponding quantities at the K^{*+} resonance peak. A phase-space factor, $\sqrt{kp}$, origins from the three-body phase space integrated over the helicity angle, θ_K, and the Jacobian

[4] For fits to simulation, the signal fraction, f_{sig}, is equal to one, making the probability density functions of the background component obsolete. Similar, fits to a pure background sample do not consider signal probability functions.

of the transformation to $m(K^0_S\pi^+)$. Furthermore, the radial angular momenta L_B and $L_{K^{*+}}$ of the form $L = S, P, D, ...$ are not to be mixed up with the angular momenta of the final state particles in the decay. And finally, the relevant Blatt-Weisskopf [194] form factors B'_{L_x} are given by

$$\begin{aligned} B'_0\left(p, p_0, d\right) &= 1 \\ B'_1\left(p, p_0, d\right) &= \sqrt{\frac{1+\left(p_0 d\right)^2}{1+\left(p\ d\right)^2}}\,, \end{aligned} \tag{5.11}$$

with the size of the decaying particle fixed to $d = 1.6\,\mathrm{GeV}^{-1}$ [195], which is equivalent to 0.3 fm [5]. Considering the angular momenta of the P-wave, $L_B = 0$ and $L_{K^{*+}} = 1$, the P-wave amplitude in Eq. (5.10) simplifies to

$$\mathcal{A}_\mathrm{P}\left(m(K^0_S\pi^+)\right) = \sqrt{kp}\times\sqrt{\frac{1+\left(p_0 d\right)^2}{1+\left(p\ d\right)^2}}\times\frac{p}{m_{K^{*+}}}\times\frac{1}{{m_{K^{*+}}}^2 - m(K^0_S\pi^+)^2 - i m_{K^{*+}}\Gamma\left(m(K^0_S\pi^+)\right)}\,. \tag{5.12}$$

The S-wave component of the $m(K^0_S\pi^+)$ is modelled using the LASS parametrisation [197]

$$\begin{aligned} \mathcal{A}_\mathrm{S}\left(m(K^0_S\pi^+)\right) =& \sqrt{kp}\times B'_{L_B}(k, k_0, d)\left(\frac{k}{m_{B^+}}\right)^{L_B}\times B'_{L_{K^{*+}}}(p, p_0, d)\left(\frac{p}{m_{K^{*+}}}\right)^{L_{K^{*+}}} \\ &\times\left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B}\frac{1}{\cot\delta_R - i}\right), \end{aligned} \tag{5.13}$$

with

$$\cot\delta_B = \frac{1}{ap} + \frac{1}{2}rp \tag{5.14}$$

and

$$\cot\delta_R = \frac{{m_{K^{*+}}}^2 - m(K^0_S\pi^+)^2}{m_{K^{*+}}\Gamma\left(m(K^0_S\pi^+)\right)}\,. \tag{5.15}$$

The scattering length, $a = 1.95\,\mathrm{GeV}^{-1}$, and effective range, $r = 1.78\,\mathrm{GeV}^{-1}$, are fixed to values determined by Ref. [198]. Alternative values have been investigated by Ref. [27],

[5] The value for d (also notated as R in some literature) has been topic of discussion. In a recent LHCb analysis of the $B^+ \to J/\psi\,\rho^+$ decay [176], d was left to float in the fit, resulting in approximately $1.64\,\mathrm{GeV}^{-1}$. In a $Z(4430)$ analysis [77] the particle size parameter was studied, indicating that the fit model favours a value closer to zero. In this analysis, a fit to the reference decay channel with a free parameter d resulted in a value of 2.5 ± 0.2. As the determination of d in the signal data is not possible, the choice of $d = 1.6\,\mathrm{GeV}^{-1}$ seems to be the reasonable conclusion, which is also consistent with previous analyses of the B^0 decay [27, 196].

which showed negligible impact on the results of the angular observables.
The central value in every q^2 interval is used for the calculation of the decay width, Γ, in the Breit-Wigner model. As for the P-wave, the general form of the parametrisation in Eq. (5.13) simplifies by applying the angular momenta involved in the S-wave decay, $L_B = 1$ and $L_{K^{*+}} = 0$,

$$\mathcal{A}_{\mathrm{S}}\left(m(K^0_{\mathrm{S}}\pi^+)\right) = \sqrt{kp} \times \sqrt{\frac{1+(k_0 d)^2}{1+(k\ d)^2}} \times \frac{k}{m_{B^+}} \times \left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B}\frac{1}{\cot\delta_R - i}\right) . \tag{5.16}$$

The squared P- and S-wave amplitudes are added using the S-wave fraction, F_S, to obtain the signal probability density function in the $m(K^0_{\mathrm{S}}\pi^+)$ dimension

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^2(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}m(K^0_{\mathrm{S}}\pi^+)}\bigg|_{\mathrm{P+S}} = (1-F_S)\left|\mathcal{A}'_{\mathrm{P}}\left(m(K^0_{\mathrm{S}}\pi^+)\right)\right|^2 + F_{\mathrm{S}}\left|\mathcal{A}'_{\mathrm{S}}\left(m(K^0_{\mathrm{S}}\pi^+)\right)\right|^2 , \tag{5.17}$$

with $\mathcal{A}'_{\mathrm{S(P)}}\left(m(K^0_{\mathrm{S}}\pi^+)\right)$, the normalised decay amplitudes of S- and P-wave.

Background components: The remaining background content in the selected signal candidates is of purely combinatorial nature, which is confirmed by the excellent descriptions of the data by a single exponential curve description in the $m(K^0_{\mathrm{S}}\pi^+\mu^+\mu^-)$ dimension and a linear model in the $m(K^0_{\mathrm{S}}\pi^+)$ dimension. The angular background is individually modelled by Chebyshev polynomials in each of the three dimensions. The polynomials are included up to second order [6] in $\cos\theta_K$ and $\cos\theta_\ell$, while being flat in the angle ϕ, resulting in

$$\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_K\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\phi}\bigg|_{\mathrm{BKG}} = \left(\sum_{i=0}^{2(4)} c_i^{\theta_K} T_i(\cos\theta_K)\right) \times \left(\sum_{j=0}^{2} c_j^{\theta_\ell} T_j(\cos\theta_\ell)\right) \times \left(c_0^{\phi} T_0(\phi)\right) , \tag{5.18}$$

with the Chebyshev polynomials, $T_0 = 1$, $T_1 = x$, $T_2 = 2x^2 - 1$, $T_3 = 4x^3 - 3x$ and $T_4 = 8x^4 - 8x^2 + 1$.
Factorisability of the three angular dimensions in Eq. (5.18) is justified by statistically evaluating the correlations between the results of angular background coefficients, c_i^x, in fits to data containing pure background candidates. The correlations between coefficients

[6] The large statistics of the reference channel $B^+ \to J/\psi K^{*+}$ allowed a more detailed evaluation of the background distribution in the $\cos\theta_K$ dimension and resulted in the inclusion of up to fourth order in the fit to $B^+ \to J/\psi K^{*+}$ data.

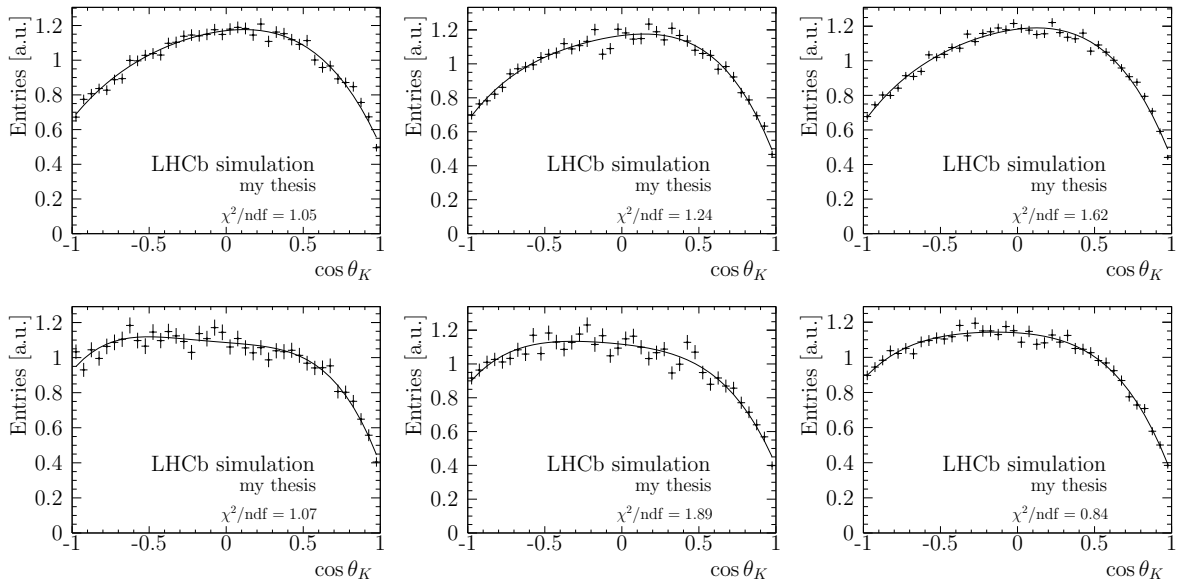

Figure 5.1: Effect of the candidate selection on the angular acceptance in the projections of $\cos\theta_K$. Selected phase-space MC candidates from the (top) DD and (bottom) LL category are further split into Run 1, 2015+2016 and 2017+2018, given from left to right. Overlaid are the parametrisation curves obtained from a moments analysis using four-dimensional Legendre polynomials. Both, the candidate distributions and the parametrisation are normalised to a flat distribution of level one and the reduced χ^2/ndf is given as a goodness of the parametrisation. More details in the text.

are found to be consistent with a correlation-free background scenario.
Normalisation conditions cancel one degree of freedom from the angular background model and constrain all three zeroth coefficients to $c_0^x = 1$.

5.3 Angular acceptance correction

The angular distributions in the selected data sets are significantly different to the raw distributions of the decay. The angular dimensions are warped by various selection and geometry constraints. The detector geometry of the LHCb experiment (see Section 3.2), the event reconstruction and the candidate selection (see Section 4.3) are the main contributors that introduce an inhomogeneity to the angular acceptance of the signal candidates. These acceptance effects have to be taken into account for the maximum-likelihood fit.
Hence, an angular acceptance correction is applied that assigns weights to every signal candidate. The weights, w, are determined as a function of q^2 and the three decay angles, $\cos\theta_K$, $\cos\theta_\ell$ and ϕ. The weight function is the inverse of the parametrised acceptance function, $\epsilon(\cos\theta_K, \cos\theta_\ell, \phi, q^2)$,

$$w(\cos\theta_K, \cos\theta_\ell, \phi, q^2) = \frac{1}{\epsilon(\cos\theta_K, \cos\theta_\ell, \phi, q^2)} \,. \tag{5.19}$$

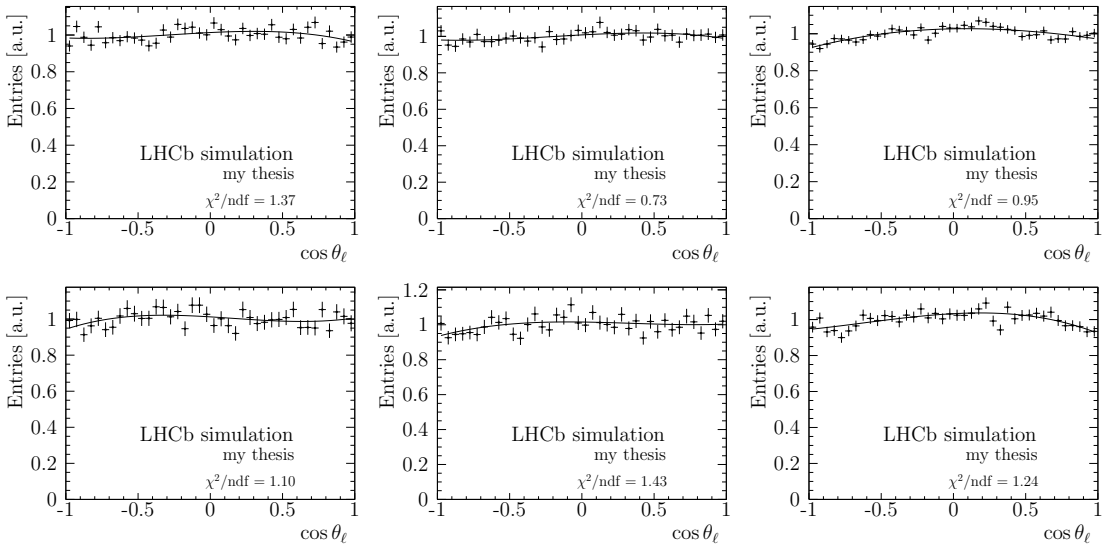

Figure 5.2: Effect of the candidate selection on the angular acceptance in the projections of $\cos\theta_\ell$. Selected phase-space MC candidates from the (top) DD and (bottom) LL category are further split into Run 1, 2015+2016 and 2017+2018, given from left to right. Overlaid are the parametrisation curves obtained from a moments analysis using four-dimensional Legendre polynomials. Both, the candidate distributions and the parametrisation are normalised to a flat distribution of level one and the reduced χ^2/ndf is given as a goodness of the parametrisation. More details in the text.

The weights from the acceptance correction are implemented into the likelihood calculation, as shown in Eq. (5.4). The acceptance function, $\epsilon(\cos\theta_K, \cos\theta_\ell, \phi, q^2)$, is determined using phase-space simulation samples, that are generated with flat distributions in these dimensions, as explained in Section 4.1. Legendre polynomials, $P_{l,m,n,o}$, parametrise the angular acceptance effects in the four-dimensional space spanned by $\cos\theta_K$, $\cos\theta_\ell$, ϕ and q^2 by

$$\epsilon(\cos\theta_K, \cos\theta_\ell, \phi, q^2) = \sum_{l=1}^{5}\sum_{m=1}^{4}\sum_{n=1}^{6}\sum_{o=1}^{7} c_{lmno} \times P_l(\cos\theta_K) \times P_m(\cos\theta_\ell) \times P_n(\phi) \times P_o(q^2)\,. \tag{5.20}$$

The maximum order of polynomials included for each of the four dimensions is investigated using a χ^2 over degrees of freedom quantifier [7] to evaluate the inclusion of additional orders to the parametrisation. The goal of this optimisation is to successfully describe the complexity of the acceptance effects, while preventing parametrisation of statistical fluctuations in the phase-space simulation samples. The optimisation converges on the following maximum orders for the Legendre polynomial parametrisation of the angular

[7] The χ^2 quantifier compares the one-dimensional projections of the parametrisation curve to the distribution of the phase-space simulation sample. The individual χ^2/ndf values are given for all projections in Figs. 5.1 to 5.4.

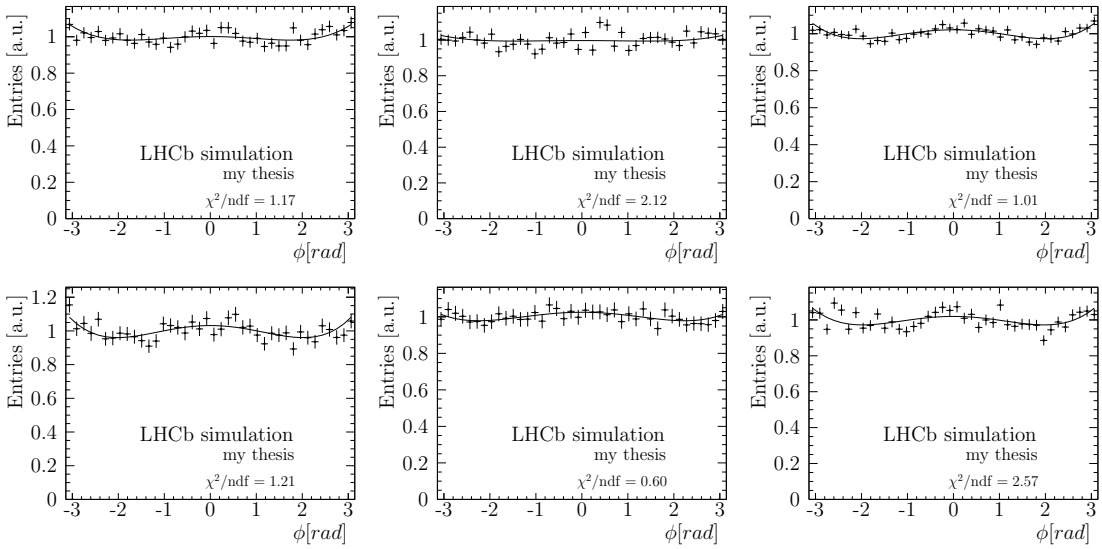

Figure 5.3: Effect of the candidate selection on the angular acceptance in the projections of ϕ. Selected phase-space MC candidates from the (top) DD and (bottom) LL category are further split into Run 1, 2015+2016 and 2017+2018, given from left to right. Overlaid are the parametrisation curves obtained from a moments analysis using four-dimensional Legendre polynomials. Both, the candidate distributions and the parametrisation are normalised to a flat distribution of level one and the reduced χ^2/ndf is given as a goodness of the parametrisation. More details in the text.

acceptance effects:

- $\cos\theta_K$: 5
- $\cos\theta_\ell$: 4
- ϕ: 6
- q^2: 7

The parametrisation is forced to be symmetric in the ϕ angle, which is motivated by the symmetric definition of ϕ as the angle between the decay planes. Reversal of the magnetic field during the data acquisition and the integration over the charge conjugated process remove any sources of potential asymmetry in acceptance of ϕ. The projections onto the ϕ dimension are shown in Fig. 5.3, with a good agreement between the simulation distribution and the parametrisation function. The symmetry is confirmed by the goodness of the parametrisation and the fit results in the reference channel (see Section 5.7.3).
The acceptance correction is performed for six sub-samples of the data set and all resulting projections of the four-dimensional parametrisation are shown in Figs. 5.1 to 5.4. The data is first split by the already discussed division between DD and LL track types, as the reconstruction efficiency of the K^0_S meson differs between the two categories and significantly warps the efficiency in the $\cos\theta_K$ dimension, as seen when comparing top

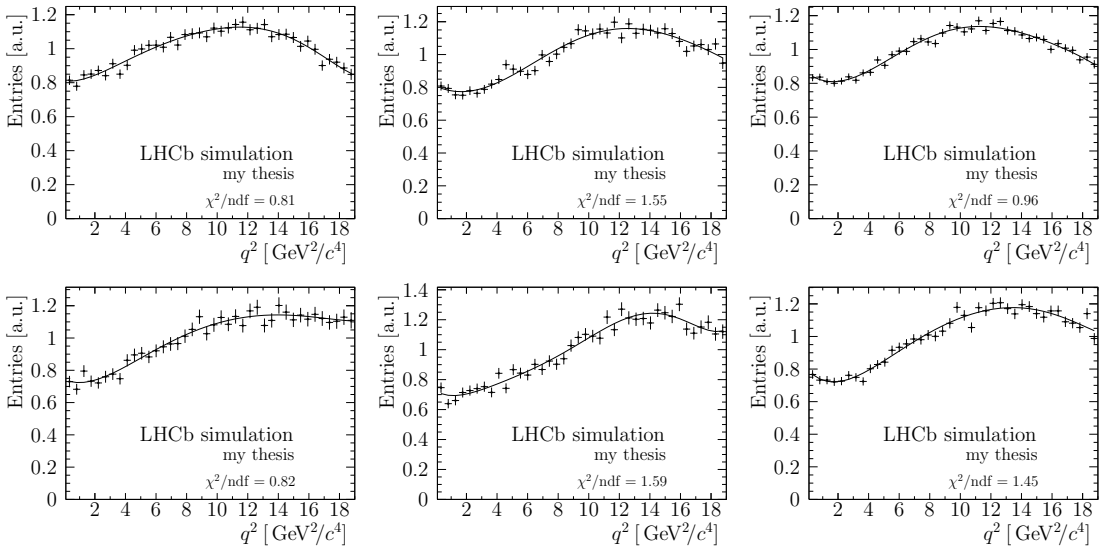

Figure 5.4: Effect of the candidate selection on the angular acceptance in the projections of q^2. Selected phase-space MC candidates from the (top) DD and (bottom) LL category are further split into Run 1, 2015+2016 and 2017+2018, given from left to right. Overlaid are the parametrisation curves obtained from a moments analysis using four-dimensional Legendre polynomials. Both, the candidate distributions and the parametrisation are normalised to a flat distribution of level one and the reduced χ^2/ndf is given as a goodness of the parametrisation. More details in the text.

and bottom rows in Fig. 5.1.

Due to different trigger configurations throughout the data taking period in Run 2, the angular acceptance effect correction is applied separately to 2015+2016 and 2017+2018. The differences are visible in the central (2015+2016) and right (2017+2018) columns of Figs. 5.1 to 5.4, especially in the projections onto q^2 in Fig. 5.4.

5.4 Angular folding techniques

Pseudoexperiments (see Section 5.7.1) with statistics equal to the signal sample sizes show significant biases in the results for most angular observables for a fit model equal to the differential decay rate in Eq. (2.26). In this *full angular* mode, *i.e.* when no angular fold is applied and all angular observables are measured simultaneously by a single fit, the number of degrees of freedom in the differential decay rate in Eq. (2.26) is too large for the limited statistics of the signal sample.

Hence, the complexity of the fit model in the three-dimensional angular space is reduced, in order to stabilise the maximum-likelihood fit and remove biases from the results of the angular observables.

For this, angular folding techniques are implemented into the FCNCfitter program. These techniques apply a set of mathematical folds on the angular variables and exploit

the periodic behaviour of trigonometric functions. As a result, some angular moments in the differential decay rate in Eq. (2.26) cancel out due to their symmetry with respect to the applied fold. Table 5.1 lists the mathematical definitions of the five folds that are used in this analysis. In Table 5.2, the sensitivity of the five folds on the angular observables is mapped out, indicating which observables cancel under a given fold and which do not. By design, all eight angular observables can be accessed by applying the five folds. For this, the maximum-likelihood fits are repeated five times – once per fold.
In practice, the angular variables are transformed accordingly to Table 5.1 prior to the input into the probability density functions, while on the other side the cancelled angular moments drop from the fit model as described by Table 5.2 [8].
These folds are a crucial part of the analysis and are needed for the successful determination of the angular observables. If not stated differently, results and uncertainty values are obtained using the angular folds during the fit procedure. The values are then obtained from the fold that is sensitive to the observable, as shown in Table 5.2, with F_{L} and S_3 (P_1) being extracted from fold 4, the fold with the least number of degrees of freedom [9]. All S-wave parameters are nuisance parameters and do not require a choice of a certain fold.
The disadvantage of using angular folding techniques and consequential non-simultaneous measurement of all angular observables in five individual folds results in a mostly blank correlation matrix, as correlations between observables measured in different folds are not directly accessible from the likelihood fit. However, this issue is addressed and solved in the following Section 5.5.

5.5 Correlation matrices

As mentioned in the previous Section 5.4 on angular folding techniques, the disadvantage of these folds is the inaccessibility of the full correlation matrix between the eight angular observables. In order to overcome this problem and to obtain the full correlation matrix, a bootstrapping [199] technique is applied.

Bootstrapping Bootstrapping is a method to statistically explore properties of a data set by randomly re-sampling the data. For this, data points are randomly selected – including double counting of points – to create a new set of random data content and of random size. The size of the random data set is generated by randomly varying the size of the genuine data set according to a Poisson distribution. The new random data set is then evaluated – in our case fitted – in the identical way

[8] The normalisation of the differential decay rates obtains a factor 2 or 4 for folds 0 and 1–4, respectively, due to the reduced angular parameter space.
[9] The choice of fold for F_{L} and S_3 (P_1) is checked to not introduce any bias using pseudoexperiments. Their values are not shared between the fits using different folds.

Table 5.1: Mathematical definition of the five angular folding techniques. Each fold exploits the symmetries in the angular differential decay rate to cancel out angular moments and to reduce the number of degrees of freedom. Note that the fold of $\cos\theta_K$ in fold number 4 depends on the sign of $\cos\theta_\ell$, not $\cos\theta_K$.

fold index	mathematical description				
0	ϕ	$\rightarrow$	$\phi+\pi$	for	$\phi<0$
1	ϕ	$\rightarrow$	$-\phi$	for	$\phi<0$
	ϕ	$\rightarrow$	$\pi-\phi$	for	$\cos\theta_\ell<0$
	$\cos\theta_\ell$	$\rightarrow$	$-\cos\theta_\ell$	for	$\cos\theta_\ell<0$
2	ϕ	$\rightarrow$	$-\phi$	for	$\phi<0$
	$\cos\theta_\ell$	$\rightarrow$	$-\cos\theta_\ell$	for	$\cos\theta_\ell<0$
3	$\cos\theta_\ell$	$\rightarrow$	$-\cos\theta_\ell$	for	$\cos\theta_\ell<0$
	ϕ	$\rightarrow$	$\pi-\phi$	for	$\phi>+\frac{\pi}{2}$
	ϕ	$\rightarrow$	$-\pi-\phi$	for	$\phi<-\frac{\pi}{2}$
4	$\cos\theta_K$	$\rightarrow$	$-\cos\theta_K$	for	$\cos\theta_\ell<0$
	$\cos\theta_\ell$	$\rightarrow$	$-\cos\theta_\ell$	for	$\cos\theta_\ell<0$
	ϕ	$\rightarrow$	$\pi-\phi$	for	$\phi>+\frac{\pi}{2}$
	ϕ	$\rightarrow$	$-\pi-\phi$	for	$\phi<-\frac{\pi}{2}$

as the genuine data set. The process of generating random data sets and fitting the very same is repeated numerous times to get a distribution of results. By definition, the distribution for any given property is then centred around the value obtained in the original data set with a spread equal to the statistical uncertainty on this very value.

To obtain all correlations – including correlations between angular observables fitted in different folds – this bootstrapping technique is applied to generate 3000 random samples. Every random data sample is fitted five times, once per fold. Two-dimensional distributions of the results from two angular observables are created to determine their correlation factor. The entire procedure is repeated in every q^2 interval and the resulting correlation matrices are given in Section 7.3 in Tables 7.3 to 7.22 for the S_i and $P_i^{(\prime)}$ angular observables. The method is validated on control pseudodata, which shows no bias in the determination of the correlation coefficients and an absolute precision on the values of 0.11.

Table 5.2: Overview on the cancellation effects of P- and S-wave angular moments of the five angular folding descriptions. The angular moments are identical to the components in Eq. (2.26). The symbol '✓' denotes angular observables which do not cancel out due to the specific fold.

observable	moment	0	1	2	3	4
F_{L}	$\cos^2\theta_K$	✓	✓	✓	✓	✓
S_3 (P_1)	$\sin^2\theta_K \sin^2\theta_\ell \cos 2\phi$	✓	✓	✓	✓	✓
S_4 (P'_4)	$\sin 2\theta_K \sin 2\theta_\ell \cos\phi$	-	✓	-	-	-
S_5 (P'_5)	$\sin 2\theta_K \sin\theta_\ell \cos\phi$	-	-	✓	-	-
A_{FB} (P_2)	$\sin^2\theta_K \cos\theta_\ell$	✓	-	-	-	-
S_7 (P'_6)	$\sin 2\theta_K \sin\theta_\ell \sin\phi$	-	-	-	✓	-
S_8 (P'_8)	$\sin 2\theta_K \sin 2\theta_\ell \sin\phi$	-	-	-	-	✓
S_9 (P_3)	$\sin^2\theta_K \sin^2\theta_\ell \sin 2\phi$	✓	-	-	-	-
F_{S}	$\sin^2\theta_\ell$	✓	✓	✓	✓	✓
S_{S1}	$\cos\theta_K \sin^2\theta_\ell$	✓	✓	✓	✓	-
S_{S2}	$\sin\theta_K \sin 2\theta_\ell \cos\phi$	-	✓	-	-	-
S_{S3}	$\sin\theta_K \sin\theta_\ell \cos\phi$	-	-	✓	-	-
S_{S4}	$\sin\theta_K \sin\theta_\ell \sin\phi$	-	-	-	✓	✓
S_{S5}	$\sin\theta_K \sin 2\theta_\ell \sin\phi$	-	-	-	-	-

5.6 Fit strategy

Up to five dimensions are used for the maximum-likelihood fits. These five dimensions and their fit model descriptions for the signal and background component are already introduced in Section 5.2. However, the data of the signal decay is not fitted in all five dimensions simultaneously. First, two-dimensional fits to the invariant masses, $m(K^0_{\mathrm{S}}\pi^+\mu^+\mu^-)$ and $m(K^0_{\mathrm{S}}\pi^+)$, are used to determine the values of the S-wave fraction, F_{S}, in three large q^2 intervals. In the consecutive fits to the four dimensions, $m(K^0_{\mathrm{S}}\pi^+\mu^+\mu^-)$ and the three decay angles $\vec{\Omega} = (\cos\theta_K, \cos\theta_\ell, \phi)$, the S-wave fractions are then constraint to the values obtained in the two-dimensional fits.
The settings of both fits are detailed in the following, with a focus on the necessity for the two-dimensional fits and the caveats that come with them.

Two-dimensional fit to $m(K^0_{\mathrm{S}}\pi^+\mu^+\mu^-)$ and $m(K^0_{\mathrm{S}}\pi^+)$ As already mentioned in Section 2.3.1, pollution from S-wave components of the $K^0_{\mathrm{S}}\pi^+$ system cannot be separated

from the P-wave in the candidate selection, but can only be distinguished statistically in either the invariant mass distribution, $m(K^0_S\pi^+)$, or in the three-dimensional decay angle space. Both, P- and S-wave components, are therefore included in the fit model of all fits to the angular dimensions and to $m(K^0_S\pi^+)$, as shown in the total differential angular decay rate in Eq. (2.26) and in the $m(K^0_S\pi^+)$ fit model in Eq. (5.17). The P-wave component in the total differential decay rate is scaled by a factor $(1 - F_S)$, which diminishes the sensitivity on the angular observables. The uncertainty on F_S and the additional degrees of freedom introduced to the fit by the S-wave component further decreases the sensitivity on the angular observables (of the P-wave component).

Hence, the precision on the angular observables relies on the precision on F_S. A study using pseudodata performed in Ref. [200] shows the improved precision on F_S by including the invariant mass of the K^{*+} meson, $m(K^0_S\pi^+)$, into the multi-dimensional fit. In the studies, identical precision on F_S is achieved in two-dimensional fits to the invariant masses $m(K^0_S\pi^+\mu^+\mu^-)$ and $m(K^0_S\pi^+)$, as compared to including $m(K^0_S\pi^+)$ as a fifth dimension to fits of $m(K^0_S\pi^+\mu^+\mu^-)$ and the three decay angles. In other words, the sensitivity on F_S mainly arises from the fits to the $m(K^0_S\pi^+)$ dimension. The simplicity of the fit in two-dimensions with identical precision on F_S is the first reason to utilise two-dimensional fits.

But more importantly, a two-dimensional mass fit easily allows to increase the size of q^2 intervals to improve the statistical precision, as the P- and S-wave probability density functions are independent of q^2 in $m(K^0_S\pi^+)$, oppose to the q^2 dependence in the angular dimensions. For the two-dimensional fits, the q^2 spectrum is split into three regions: always one below, in between and above the narrow $c\bar{c}$ resonances, while the very low q^2 region below the $\phi(1020)$ resonance is excluded. The resulting, large q^2 intervals are $1.1 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$, $11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$ and $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$. The results of F_S in the lowest and highest q^2 interval are then scaled to the nominal q^2 intervals, which are defined in Table 4.5. The scaling is done using the theoretical prediction of the fraction of longitudinally polarised K^{*+} mesons, F_L. The ratio of F_L between the larger and the nominal q^2 intervals is taken as the scaling factor

$$F_S^{\mathrm{nom}} = F_S^{\mathrm{large}} \times \frac{F_L^{\mathrm{nom}}}{F_L^{\mathrm{large}}} = F_S^{\mathrm{large}} \times F_L^{\mathrm{ratio}} \,. \tag{5.21}$$

The obtained values for F_S in the large q^2 intervals, along with the scaling factors obtained from the F_L ratios are given in Table 5.3. The larger q^2 intervals allow a measurement of F_S with an absolute precision better than 0.1, which is about a factor two better compared to the precision observed in sensitivity studies using the nominal q^2 intervals. The scaling using F_L assumes a similar dependence of F_L and F_S on the squared invariant dimuon mass, q^2. This statement is broadly compatible with findings in Ref. [10]. Also

Table 5.3: Scaling of the S-wave fraction as measured in the two-dimensional fit to the invariant masses in the larger q^2 intervals, $F_{\mathrm{S}}^{\mathrm{large}}$, to values of F_{S} in the nominal q^2 intervals. The ratio of F_{L} between the nominal and large intervals, as defined in Eq. (5.21) is used for this scaling. The resulting values of $F_{\mathrm{S}}^{\mathrm{nom}}$ are used to put constraints on the S-wave fraction in the four-dimensional fit, with more details given in the text.

q^2 intervals [GeV^2/c^4]		$F_{\mathrm{S}}^{\mathrm{large}}$	$F_{\mathrm{L}}^{\mathrm{ratio}}$	$F_{\mathrm{S}}^{\mathrm{nom}}$
	[0.98–1.1]	scale from [1.1–8.0]	0.43	0.11 ± 0.04
[1.1–8.0]	[1.1–2.5]	0.25 ± 0.10	1.09	0.27 ± 0.11
	[2.5–4.0]		1.13	0.29 ± 0.12
	[4.0–6.0]		1.01	0.26 ± 0.10
	[6.0–8.0]		0.86	0.22 ± 0.09
[11.0–12.5]	[11.0–12.5]	0.01 ± 0.15	1.0	0.01 ± 0.15
[15.0–19.0]	[15.0–17.0]	0.02 ± 0.09	1.02	0.02 ± 0.09
	[17.0–19.0]		0.96	0.02 ± 0.09

note, that the scaling of F_{S} does not change the value from the larger q^2 intervals by more than 10%, with the exception of the lowest q^2 bin, as can be seen by the values in Table 5.3. The absolute effect of the scaling is smaller than the statistical uncertainty on the measurements. The values for F_{L} in the large and nominal q^2 intervals are calculated using the Flavio software package [101].

In summary, the two-dimensional fits to $m(K_{\mathrm{S}}^0\pi^+\mu^+\mu^-)$ and $m(K_{\mathrm{S}}^0\pi^+)$ are used to put a constraint on F_{S} in the subsequent four-dimensional fits. To achieve the best possible sensitivity, F_{S} is measured in larger q^2 intervals and then scaled via the ratio of the values for the fraction of longitudinally polarised K^{*+} mesons, F_{L}.

This method is clearly a compromise between different strategies. Possible considerations range from ignoring the S-wave pollution completely (as done in Ref. [25]) over constraining it to predictions from B^0 results to fixing it to purely theoretical calculations of F_{S} values. The benefit of the introduced scaling is that the S-wave fraction is measured on signal decay data and uses only minimal input from theory predictions for the scaling factors of F_{L}. A systematic study in Section 6.7 is dedicated to the potential impact of this scaling on the results of the angular observables.

For the two extra wide q^2 intervals, $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ and $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$, no scaling of the determined S-wave fraction is required and the values in the four-dimensional fit are constraint directly to the values obtained from the two-dimensional mass fits in the large intervals.

Four-dimensional fit to $m(K^0_S\pi^+\mu^+\mu^-)$, $\cos\theta_K$, $\cos\theta_\ell$ and ϕ The fits to the four dimensions are the heart of the angular analysis. One hundred fits to data determine the central values of the angular observables in ten q^2 intervals (see Table 4.5) with five angular folding techniques – which are explained in Section 5.4 – and for two observable bases, S_i and $P_i^{(\prime)}$. Again, the reference to Section 5.2 is made, in which the fit models for all four dimensions are explained.
The data is split in the four categories Run 1 DD, Run 1 LL, Run 2 DD and Run 2 LL, and fitted simultaneously in all of them. As explained in Section 5.1, the angular observable parameters are shared among the different categories. For the fits every candidate is assigned a weight to correct for angular acceptance effects, as explained in Section 5.3 and Eq. (5.4).
Before the results are presented in Chapter 7, the fitter configuration with all statistical techniques is tested on pseudodata, simulation data and the reference channel, $B^+ \to J/\psi K^{*+}$.

At last, the projections of the four-dimensional fit model in the three angular dimensions are given in Appendix A.5, for angular folds labelled 0 to 4 in Figs. A.5 to A.14. As the mass dimension is not directly impacted by the angular folds, exemplary projections are given in Fig. 4.8 and not specifically for every fold configuration.

5.7 Fit validations

All parts of the fit program are tested for potential biases in the results and for a correct determination of the uncertainty on the measurements. This counts especially for the correct implementation and functionality of the previously introduced and fitter-integrated methods like the bootstrapping techniques, angular foldings and corrections of angular acceptance effects.
In the following three paragraphs, the fitter is tested on pseudodata in Section 5.7.1, on official LHCb simulation samples in Section 5.7.2 and on the reference decay channel, $B^+ \to J/\psi K^{*+}$, in Section 5.7.3.
In general, the fitter mostly works as expected and the utilisation to extract the angular observables in the ten q^2 intervals is possible. However, some systematic effects observed in the tests require the assignment of additional uncertainties to the measurements.

5.7.1 Studies with pseudodata

In a first step, the correct functionality of the maximum-likelihood fitting software is investigated on *pseudodata*. These are very simple events, which can be generated with

specifically chosen parameter values. Usually, only directly required variables are generated, which in this analysis are $m(K^0_S\pi^+\mu^+\mu^-)$, $m(K^0_S\pi^+)$, q^2, $\cos\theta_K$, $\cos\theta_\ell$ and ϕ. The variables are generated using a random number generator and the probability density functions of the maximum-likelihood fit. The generation process is introduced in detail in Appendix A.2. Due to the simplicity and incomplete generation (with respect to full simulation data, as introduced in Section 4.1), these candidates are called *pseudodata* and their studies are referred to as *pseudoexperiments*.

The main objectives in these studies are the correct functionality of the minimisation to determine bias-free results of the free parameters [10]. Furthermore, the coverage of the statistical uncertainty is tested. Both, potential bias of the central value and the coverage of the statistical uncertainty are studied by evaluating *pull* distributions of pseudoexperiments.

Pull distributions The pull of a measurement with the measured value, x, and the uncertainty on the measurement, σ, is determined by

$$\frac{x - x_0}{\sigma}\,, \tag{5.22}$$

with the generated value, x_0. By definition and due to the central limit theorem, the shape of pull distributions for statistically independent random variables is distributed like a standard Gaussian shape with a mean position equal to zero and a width equal to one [201].

Pull distributions are created for all pseudoexperiments and are used to evaluate the functionality of the fitter at a given configuration of parameters, sample sizes, fit models and other settings like acceptance correction weights and angular folding techniques. Fig. 5.5 shows three excellent examples of pull distributions for angular observables, which are well described by the symmetric Gaussian model. The mean position and the width of the overlaid fits to the distributions are compatible with the nominal values. Fig. 5.6 however shows asymmetric distributions. The normal Gaussian model does not describe the distribution sufficiently and both, mean position and width, are off from their nominal values. Reasons for these distortions are diverse and can originate from various sources, such as boundary limits in the parameter space, a general bias in the maximum-likelihood fit, limited statistical precision or a technical problem with the minimisation. In the following paragraph, the configurations and findings of the conducted pseudoexperiments are described.

[10] In the investigations, not only the results of the angular observables are investigated, but also the bias-free determination of all nuisance parameters in the S-wave and background components. No mentionable biases or issues are found for these parameters and hence will not be the subject of further investigations.

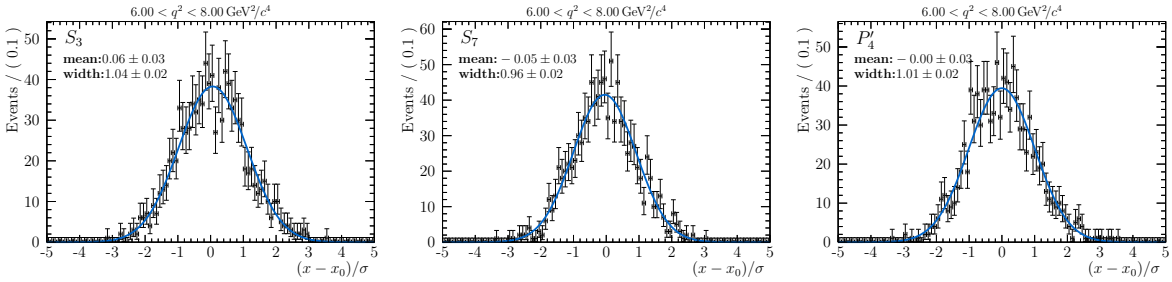

Figure 5.5: Positive examples of pull distributions from pseudoexperiment results using generated values equal to SM predictions for (left) S_3, (centre) S_7 and (right) P_4' in the interval $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$. The normal Gaussian model describes the data points well, with nominal values for mean and width. The distributions are symmetric and show no signs of boundary effects. When all pull distributions looked like this, no further systematic uncertainties or Feldman-Cousins scans for the statistical uncertainty are needed.

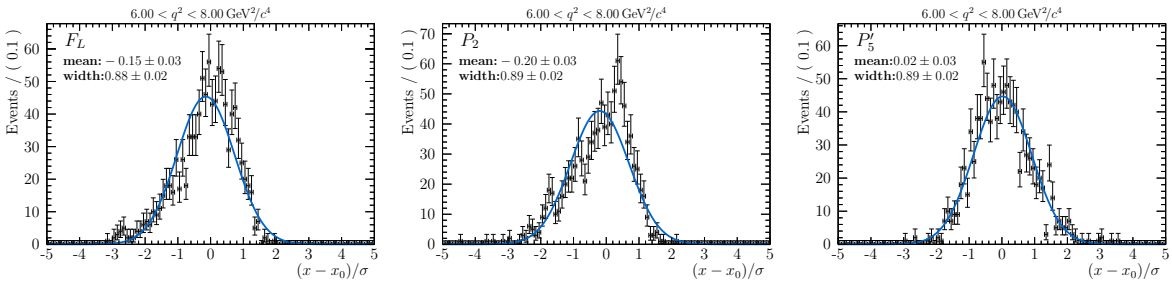

Figure 5.6: Negative examples of pull distributions from pseudoexperiment results using generated values equal to SM predictions for (left) F_L, (centre) P_2 and (right) P_5' in the interval $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$. The normal Gaussian model does not sufficiently describe the clearly asymmetric distributions. Values for mean and width obtained by the fit do not represent the true values of the distributions and are not compatible with the nominal values. As a result of these distributions, systematic studies are performed in Sections 6.7 and 6.8 and the statistical uncertainty of the measurements is determined from Feldman-Cousins scans in Section 5.8.

Pseudoexperiment agenda In a first series of pseudoexperiments, it is shown that the implementation of the special analysis techniques – for the angular acceptance correction, the angular folding techniques and the determination of the correlation matrices – does not introduce any bias to the measurements. Also more general, the minimisation of the negative logarithm of the likelihood is confirmed to operate bias-free at large candidate statistics. For these first studies, pseudodata samples with some thousand candidates per q^2 interval are used, to be insensitive on effects due to low statistical precision.

The main pseudoexperiments are configured identically to the fits on data. This means, a two-dimensional fit to the $m(K_\mathrm{S}^0\pi^+\mu^+\mu^-)$ and $m(K_\mathrm{S}^0\pi^+)$ invariant mass is followed by a four-dimensional fit to $m(K_\mathrm{S}^0\pi^+\mu^+\mu^-)$, $\cos\theta_K$, $\cos\theta_\ell$ and ϕ, as explained in Section 5.6.

The candidate statistics are equal to the statistics in data.
In total, pseudodata are generated for three scenarios, with different values of the angular observables at generation:

- predictions of a pure Standard Model scenario
- contributions from physics beyond the SM to the vector coupling strength of $\mathrm{Re}(\mathcal{C}_9^{\mathrm{NP}}) = -1$
- the best-fit results from the fits to data

The values for the angular observables in the SM and $\mathrm{Re}(\mathcal{C}_9^{\mathrm{NP}}) = -1$ scenario are determined using the Flavio software package.
In total, pseudoexperiments are performed in all ten q^2 intervals and for both sets of angular observables, S_i and $P_i^{(\prime)}$. The resulting values for the widths and mean positions of the pull distributions of all sixteen angular observables in the three different scenarios are given in Appendix A.3.
Fig. 5.7 shows the comparison of the pulls of the angular observables P_1–P_4' in the interval $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$ between the three scenarios. The three columns reveal different shapes of the pull distributions.
In general, some shifts in the mean position away from 0.0 are visible for some angular observables in all three scenarios. Unfortunately, modelling the bias as a function of the generated value for each observable is not possible. Hence, the bias cannot be corrected in the result, and the conservative approach of assigning a systematic uncertainty to this bias is used, which is given in Table 6.11 in Section 6.8. In addition, the Gaussian widths differ from the nominal value of 1.0 in several cases. A correct coverage by the Hesse [11] uncertainty determination is not achieved. As a result, Feldman-Cousins profile scans are utilised to determine the correct confidence interval of all angular observables. These scans are explained in Section 5.8.

Conclusions from pseudoexperiments In summary, the studies using pseudodata prove the functionality of the fitter, with no biases observed in pseudoexperiments using statistics in the order of 10,000 candidates. This statement holds for the implementation of the angular folding techniques, as well as for the correction of angular acceptance effects using per-event weights, as introduced in Sections 5.3 and 5.4, respectively. The studies with statistics equal to data, however, revealed some major issues that require further measures. Firstly, the usage of previous measurements of F_{S}, which in addition are scaled

[11] The standard statistical uncertainties from the maximum-likelihood fit are obtained using the *Hesse* determination [202], which derives a symmetric statistical uncertainty from the second derivative of the likelihood function at the best-fit value.

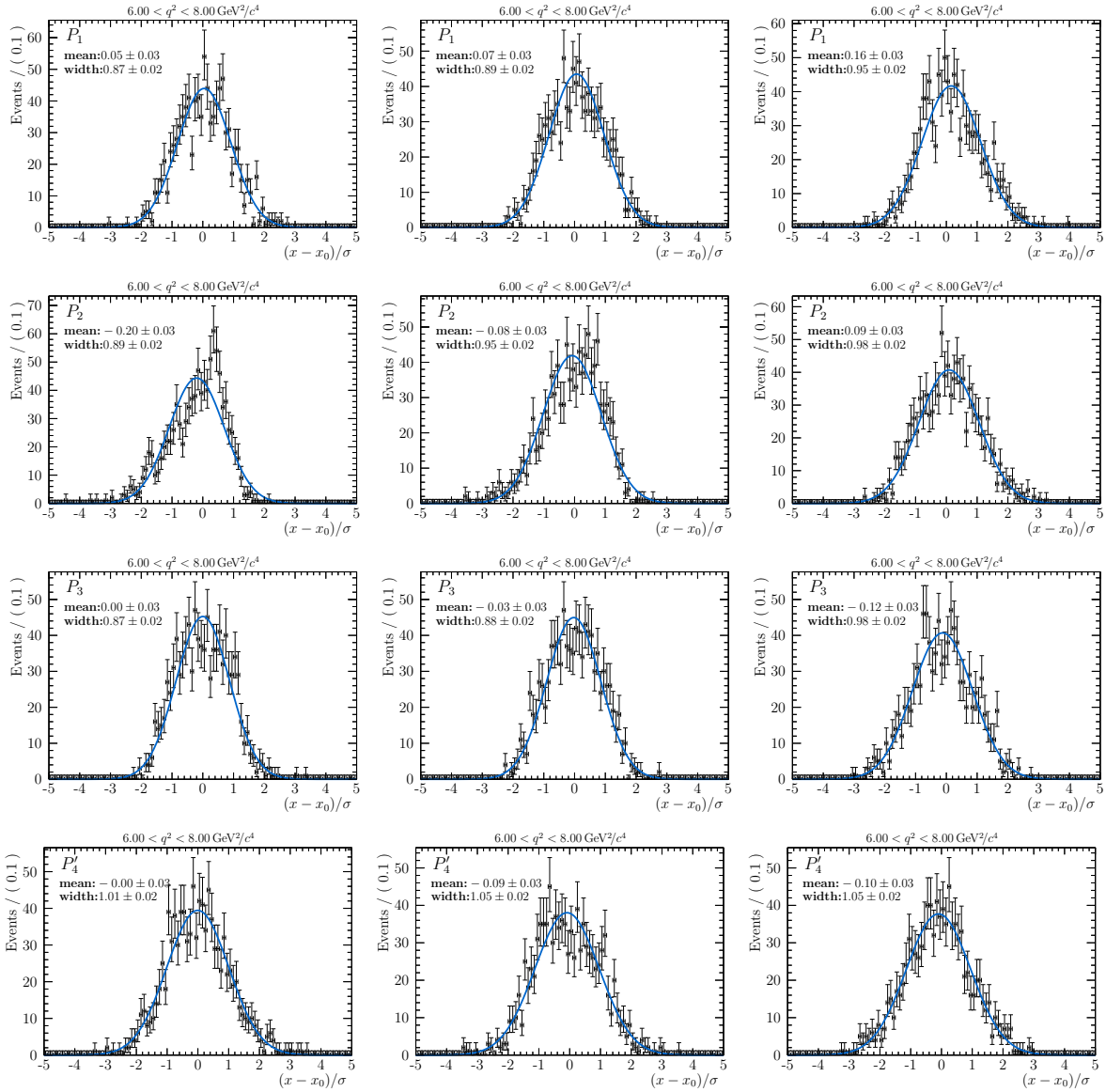

Figure 5.7: Comparison of pull distributions from different generation scenarios. On the left, SM values are used to create pseudodata, in the centre a new physics scenario with a shift in the Wilson coefficient $\Delta\mathrm{Re}(\mathcal{C}_9) = -1$ and the best-fit results as generation values on the right. The angular observables P_1–P_4' are shown in the interval $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$. The distributions show individual features for the three scenarios.

from a different q^2 binning scheme (as explained in Section 5.6), creates biases, which require the assignment of a systematic uncertainty in Section 6.7.

Also, pseudoexperiments with generated values equivalent to the best-fit results or to a new physics scenario with a shift in the real part of the vector coupling strength, $\Delta\mathrm{Re}(\mathcal{C}_9) = -1$, show non-vanishing biases, which are also assigned as a systematic uncertainty in Section 6.8.

Finally, the studies show significant problems in the uncertainty coverage, as the widths of

pull distributions are either too wide or too narrow. As a result and to be consistent through all angular observables, Feldman-Cousins scans are used to determine the statistical uncertainty of the measurement results in Section 5.8.

5.7.2 Fit to simulation

The validation of the fitter continues with fits to the official LHCb simulation, with a focus on verifying the angular acceptance correction and the angular folding. The methods are tested over the full q^2 range of the $B^+ \to K^{*+}\mu^+\mu^-$ decay.
As described in Section 4.1, recorded data and simulation samples are selected in an identical way. Using the selected candidate samples, six maximum-likelihood fits per nominal q^2 interval are performed to extract the angular observables in the S_i basis. Five fits correspond the five folds being applied to the angular distributions, while the values of the decay angles in the sixth fit are pristine. Because simulation comprises of only signal candidates without background contributions and no S-wave pollution, the fit model consists of a pure angular signal model in three dimensions. The fit to the invariant mass is not required, as signal and background does not require to be distinguished from one another [12]. This greatly reduces the complexity of the fit and isolates the focus on the determination of the angular observables in an idealised scenario.
In Fig. 5.8, the results for the eight S_i angular observables are shown as a trend of q^2. The different marker and line colour indicates the angular folds. Almost perfectly hiding are the underlying results of the full angular fit and the fit to generator level simulation in grey. The latter fit to simulation samples prior to reconstruction or selection extracts the original values that the simulation samples are generated with [13]. This generator-level sample is free from any effects due to acceptance filter, no reconstruction and no candidate selection.

5.7.3 Fit to reference decay

The validation of the angular fitter peaks in the fits to the reference decay, $B^+ \to J/\psi K^{*+}$. For this, data is selected from the q^2 interval 8.68–10.09 GeV^2/c^4. It is then split into the four nominal samples, Run 1 DD, Run 1 LL, Run 2 DD and Run 2 LL and simultaneously fitted by a maximum-likelihood in five dimensions, the invariant masses of the B and K^{*+} meson and the angular dimensions, $\vec{\Omega}$. The large statistics of the reference channel allow

[12] A fit to the invariant mass dimension $m(K^0_S\pi^+\mu^+\mu^-)$ is performed outside the scope of these fitter validations in order to extract the parameter values of the signal peak of the B^+ meson, which are then used in the fit to data, as explained in Section 5.6.

[13] The method of extracting the generated values through fits to generator-level simulation is the most practical approach, as the values for the underlying vector couplings and hadronic form factors that are used for the simulation, are deeply buried inside the LHCb simulation framework.

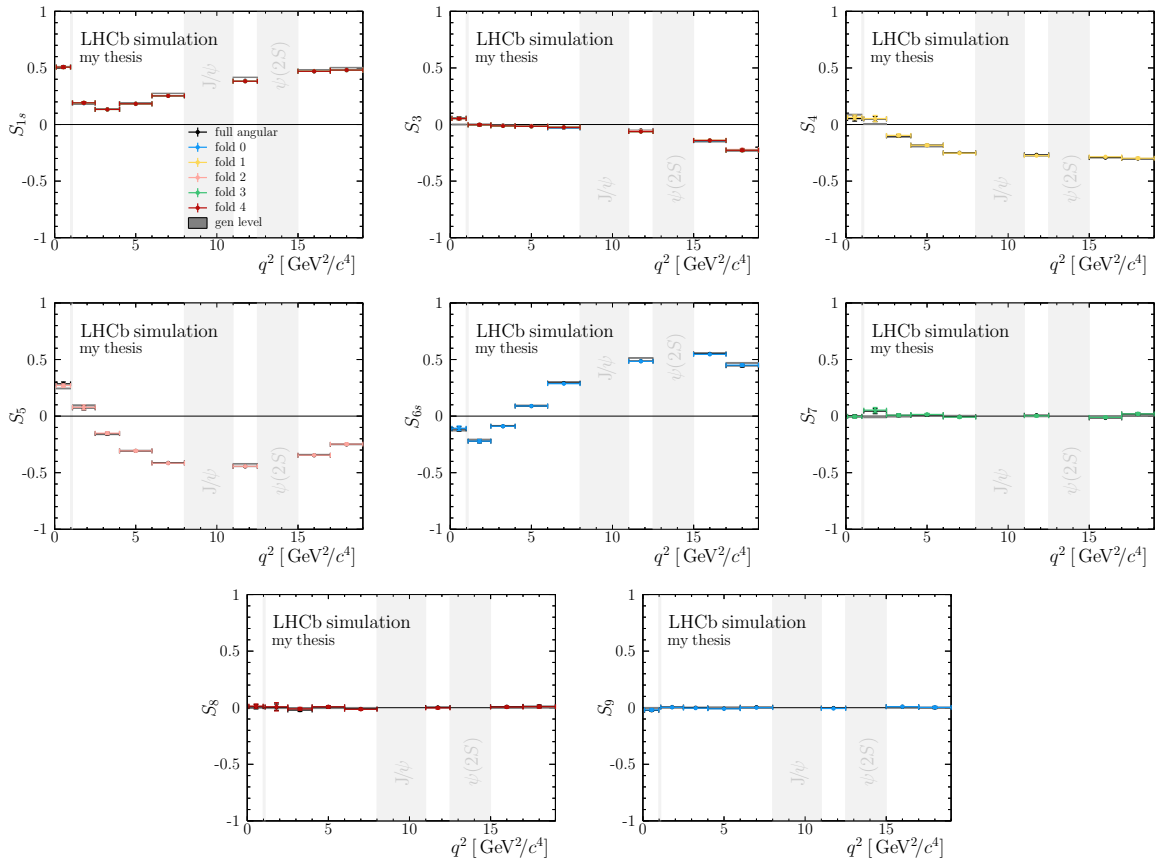

Figure 5.8: Comparison of results from multiple fits to the simulation samples in the nominal q^2 interval scheme. Fits using the angular folds (coloured) are compared to the fit without angular folding applied (black) and results from a fit to the true information of simulation at generator level (grey box). The large statistics of simulation samples result in the statistical uncertainty mostly disappearing under the marker. All results are in very good agreement and demonstrate the functionality of the angular acceptance correction throughout the full range of q^2.

to directly fit the S-wave fraction F_{S}. All fit models are identical to the signal channel and previously explained in detail in Section 5.2.

The fit projections in all five dimensions along with the data distributions are shown in Fig. 5.9, where the projections and the data are combined from four sub-samples. A very good description of the data distribution by the fit model is achieved. The asymmetric distribution of $\cos\theta_K$ clearly shows the necessity of including the S-wave component into the fit.

The results from the fit to the reference decay channel $B^+ \to J/\psi K^{*+}$ are compared to the results from Refs. [203–205] in Table 5.4. Results of all angular observables are in good agreement with the dedicated analyses of the $B \to J/\psi K^*$ decays, and the achieved precision is clearly sufficient for an analysis of the rare decay $B^+ \to K^{*+}\mu^+\mu^-$. Especially, the agreement of S_5, S_{6s} and S_7, which are by definition equal to zero, with the results

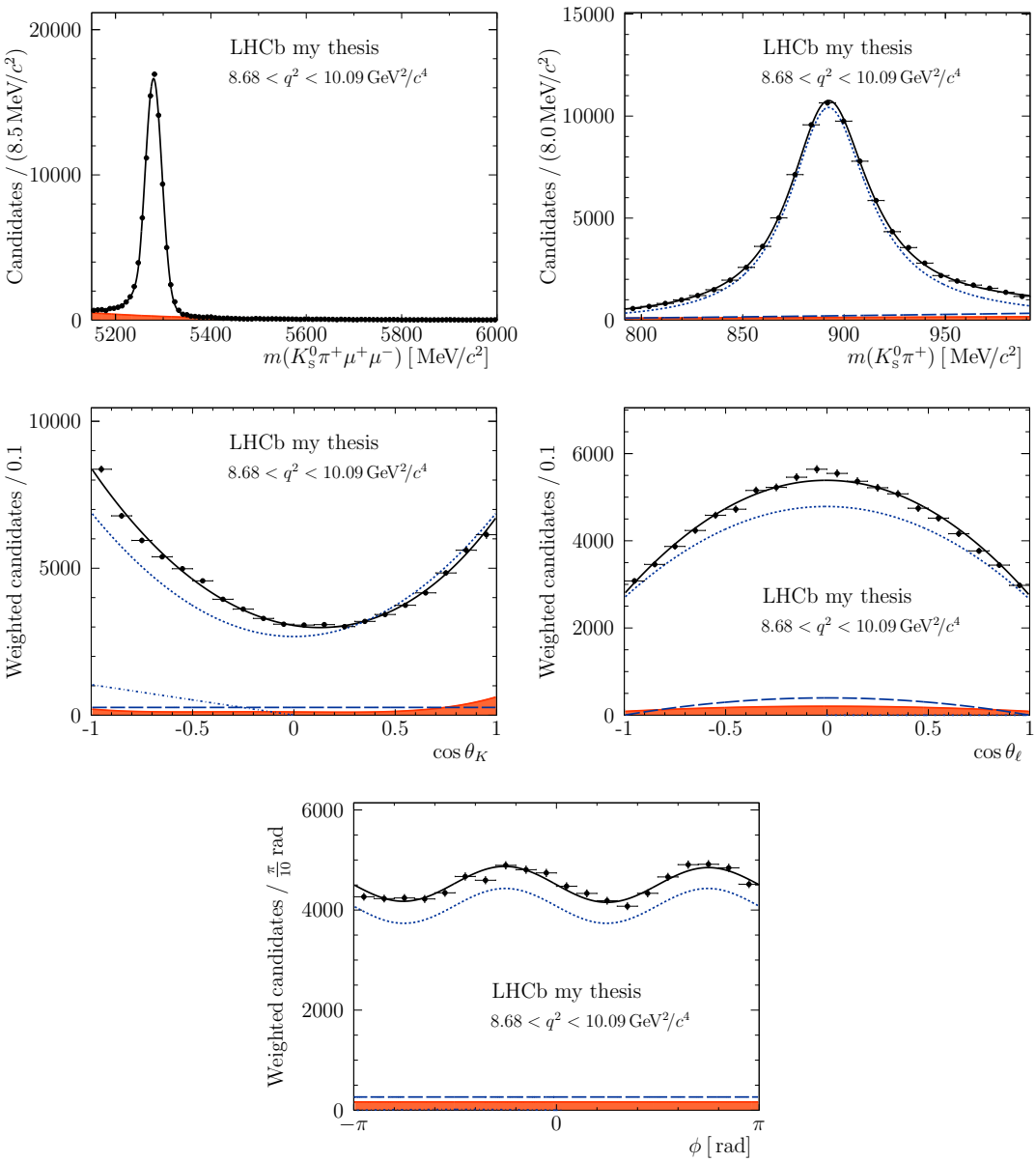

Figure 5.9: Five-dimensional fit to the reference decay channel, $B^+ \to J/\psi K^{*+}$, in the (top left) $m(K^0_S\pi^+\mu^+\mu^-)$, (top right) $m(K^0_S\pi^+)$, (bottom left) $\cos\theta_K$, (bottom centre) $\cos\theta_\ell$ and (bottom right) ϕ dimension. Data is separated from the signal decay by the given cut in q^2 and comprises the full selection for Run 1 and Run 2 for both DD and LL. The total fit model is given in solid black, with the (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference components given individually in the angular dimensions. The measured values of the angular observables are compared to previous measurements by LHCb, BaBar and Belle in Table 5.4.

from Belle, BaBar and LHCb are a strong confirmation of the functionality of the fitter and the acceptance correction.

The agreement between angular folds and the non-folded full angular fit are again investigated in the reference channel. In addition to the very consistent results of the folded

Table 5.4: Comparison of the results of our full angular fit to $B^+ \to J/\psi K^{*+}$ data to results of existing publications from BaBar [203], Belle [204] and LHCb [205]. These publications provide measurements in different combinations of B^0 and B^+ decays. The results are originally stated as polarisation amplitudes, which have been converted into the angular observable basis, S_i, to be comparable to the results in this thesis. The uncertainties from the polarisation amplitudes are propagated to the angular observables using 100,000 randomly generated toys. All results are in agreement.

	this analysis	Belle (B^+)	BaBar ($B^+ + B^0$)	LHCb (B^0)
S_{1s}	0.321 ± 0.004	0.297 ± 0.011	0.333 ± 0.007	0.321 ± 0.006
S_3	-0.002 ± 0.007	-0.018 ± 0.017	0.011 ± 0.011	-0.013 ± 0.010
S_4	-0.246 ± 0.008	-0.255 ± 0.010	-0.237 ± 0.007	-0.250 ± 0.006
S_5	-0.003 ± 0.008	0	0	0
S_{6s}	-0.003 ± 0.006	0	0	0
S_7	-0.001 ± 0.008	0	0	0
S_8	-0.063 ± 0.008	-0.037 ± 0.018	-0.058 ± 0.015	-0.048 ± 0.007
S_9	-0.084 ± 0.007	-0.041 ± 0.016	-0.095 ± 0.014	-0.084 ± 0.006

Table 5.5: Results of the five-dimensional fits to the $B^+ \to J/\psi K^{*+}$ data in a full angular configuration and all five folds. The results show good agreement within their uncertainties for all angular observables of the P- and S-wave between folds and full angular configuration.

	full angular	fold 0	fold 1	fold 2	fold 3	fold 4
S_{1s}	0.321 ± 0.004	0.321 ± 0.005	0.321 ± 0.004	0.321 ± 0.005	0.321 ± 0.005	0.321 ± 0.002
S_3	-0.002 ± 0.007	-0.003 ± 0.007	-0.002 ± 0.007	-0.003 ± 0.007	-0.003 ± 0.007	-0.004 ± 0.003
S_4	-0.246 ± 0.008	-	-0.246 ± 0.008	-	-	-
S_5	-0.003 ± 0.008	-	-	-0.002 ± 0.008	-	-
S_{6s}	-0.003 ± 0.006	-0.004 ± 0.006	-	-	-	-
S_7	-0.001 ± 0.008	-	-	-	-0.002 ± 0.008	-
S_8	-0.063 ± 0.008	-	-	-	-	-0.061 ± 0.004
S_9	-0.084 ± 0.007	-0.084 ± 0.007	-	-	-	-
F_S	0.055 ± 0.013	0.055 ± 0.013	0.055 ± 0.013	0.055 ± 0.013	0.055 ± 0.013	0.055 ± 0.005
S_{S1}	-0.193 ± 0.014	-0.195 ± 0.014	-0.193 ± 0.014	-0.195 ± 0.014	-0.195 ± 0.014	-
S_{S2}	0.024 ± 0.009	-	0.024 ± 0.008	-	-	-
S_{S3}	-0.001 ± 0.008	-	-	0.001 ± 0.008	-	-
S_{S4}	-0.004 ± 0.008	-	-	-	-0.003 ± 0.008	-0.002 ± 0.004
S_{S5}	-0.074 ± 0.009	-	-	-	-	-

fits to the simulation in Fig. 5.8, the results from the reference decay for all observables of the P- and S-wave component in Table 5.5 are in very good agreement to other folds and to the fit without any angular fold. In a last fit validation test, random sets of the $B^+ \to J/\psi K^{*+}$ data are sampled. The random sets are the size equal to the statistics of the $B^+ \to K^{*+}\mu^+\mu^-$ decay. It should be noted, that the amount of signal candidates in these random samples is larger than in the $B^+ \to K^{*+}\mu^+\mu^-$ samples due to a purer signal sample. The signal fraction, f_{sig}, is approximately 0.9 in $B^+ \to J/\psi K^{*+}$ data and can be

Table 5.6: Average results obtained by fits to random samples of $B^+ \to J/\psi K^{*+}$ data of size equal to the $B^+ \to K^{*+}\mu^+\mu^-$ sample size (rare size) in the right column are compare to results from the nominal five-dimensional fit to the full $B^+ \to J/\psi K^{*+}$ data set in the left column. The low statistics results are obtained using angular folds and their uncertainties are determined by the spread of the results from the one thousand fits. Again, the results are consistent for all eight angular observables. More details are given in the text.

	full data set	rare size (average)
S_{1s}	0.321 ± 0.002	0.316 ± 0.047
S_3	-0.002 ± 0.007	-0.001 ± 0.067
S_4	-0.246 ± 0.008	-0.256 ± 0.095
S_5	-0.003 ± 0.008	-0.008 ± 0.078
S_{6s}	-0.003 ± 0.006	-0.006 ± 0.062
S_7	-0.001 ± 0.008	-0.005 ± 0.085
S_8	-0.063 ± 0.008	-0.060 ± 0.086
S_9	-0.084 ± 0.007	-0.085 ± 0.070

as low as 0.5 in the $B^+ \to K^{*+}\mu^+\mu^-$ case. As this effect is hard to correct [14] and the method is meant as an additional cross-check, no further optimisations are applied. In total, one thousand random samples are fitted using the fit configuration as for the signal samples – including angular folds, and the averaged results are given in Table 5.6, where they are compared to the nominal fit results from the $B^+ \to J/\psi K^{*+}$ decay, as already shown in Tables 5.4 and 5.5. All results are in very good agreement.

5.8 Feldman-Cousins scans

The pseudoexperiment studies in Section 5.7.1 revealed issues with the correct coverage of the statistical uncertainty obtained by the maximum-likelihood fit for some of the angular observables and in some q^2 intervals. These individual pull distributions, as exampled in Fig. 5.6, show signs of *over*- or *under*-coverage, that is where the widths of the pull distribution are smaller or larger than one, respectively. In other words, the statistical uncertainty is *over*- or *under*-estimated by the fit. These problems appear in cases when the likelihood profile of the minimisation is warped from an ideal parabolic shape, which can be the case in fits to samples of finite statistics. The likelihood profile around the minimum can contain shoulder-like features, be slightly stretched or compressed, or simply be asymmetric. Also edges from the boundaries of the parameter space can show up in the shape of the likelihood profile.

[14] Simply scaling the sample size by the ratio of the signal fractions does not make up for the differences, as the number of background candidates would drop in the same manner.

As a result of these issues, the statistical uncertainties are determined using one-dimensional Feldman-Cousins (FC) [15] scans [206].

Feldman-Cousins technique The statistical method, that is referred to as Feldman-Cousins technique [206], uses Neyman constructions [207] to ensure correct coverage of a measured parameter, θ [189, 208]. Neyman constructions are a frequentist method, as opposed to the bayesian method of the maximum-likelihood estimator introduced in Section 5.1. They are used to construct an confidence interval of a given confidence level $\mathcal{C}$ around a measured value θ_0. The method scans the coverage probability in steps of θ_j of the parameter θ. While for the traditional Neyman construction the coverage probability at every step θ_j is determined as the frequency of pseudoexperiment for which the confidence interval contains the actual value of θ_0, in the Feldman-Cousins technique the coverage probability is determined by the frequency of pseudoexperiments with a smaller delta-log-likelihood as a reference fit to the actual data. This determination is explained in the following.

In practice, a Feldman-Cousins likelihood profile scan can be performed individually for every free parameter [16] of a converged likelihood fit. The parameter-in-question is varied by a defined step size over a suitable range around the minimum value obtained by the fitter. The step size is a trade-off between granularity in the scan and available computing resources. The range has to be wide enough to cover the confidence interval, to reach the confidence level $\mathcal{C}$. At every step, the data sample is fitted again, with the parameter-in-question fixed to the value at this scan step. All other parameters are fixed, constrained or floated according to the genuine fit configuration. The total likelihood (see Eq. (5.8)) is obtained from the fit at step j. From this, the likelihood difference, the delta-log-likelihood, with respect to the best-fit scenario is determined as

$$-2\Delta \log \mathcal{L}(j) = -2 \log \mathcal{L}_{\text{fixed}}(j) + 2 \log \mathcal{L}_{\text{best-fit}} \,. \tag{5.23}$$

The delta-log-likelihood from the fits to data is shown in blue in the scan profiles of Fig. 5.10, after being converted to a confidence level, $\mathcal{C}$,

$$\mathcal{C}(j) = \text{erf}\left(\sqrt{-\Delta \log \mathcal{L}(j)}\right) \,, \tag{5.24}$$

[15] The abbreviation FC is not be confused with the glorious football club 1. FC Köln.

[16] Multi-dimensional FC scans are possible. Several parameters can be scanned at the same time, but the size of the parameter space grows with the power of the dimension. The precision gain on the uncertainties by including one or two dimensions is doubtful, while varying all parameters exceeds the computational capacities of any available computing grid. Hence, in this analysis, only one-dimensional FC scans are used.

with the Gaussian error function, erf. The delta-log-likelihood from data is needed for the next part of Feldman-Cousins scans. The algorithms and methods so far correspond to a standard delta-log-likelihood profile scan.
For the Feldman-Cousins method, pseudoexperiments (see details on the generation of pseudodata in Appendix A.2) are run at every scan step. The pseudodata is generated with all parameter values equivalent to the values obtained by the best-fit results on the genuine fit to data. Only exception to this is the parameter-in-question, θ, which is (logically) set to the value θ_j at this scan step.
The pseudodata samples are then fitted twice, first with the nominal fit configuration, *i.e.* all parameters are either fixed, constrained or floated as in the original fit to data. And in a second fit, the only difference is that the parameter-in-question is fixed to the generated value at this scan step. From two fits to the pth pseudodata sample and at the scan step j, the delta-log-likelihood is determined by

$$-2\Delta \log \mathcal{L}(j,p) = -2 \log \mathcal{L}_{\text{fixed}}(j,p) + 2 \log \mathcal{L}_{\text{floated}}(j,p)\,. \tag{5.25}$$

At every step j, the delta-log-likelihoods obtained from all pseudoexperiments at this step are compared to the delta-log-likelihood obtained from data as in Eq. (5.23). The fraction of pseudodata delta-log-likelihoods smaller than the delta-log-likelihood on data defines the confidence level at this step. Combining all steps, the Feldman-Cousins profile scan is generated. The resulting distributions are exemplarily shown in black in Fig. 5.10, with all profiles being attached in Appendix A.6. The values for the confidence interval, which by definition is 68.3% (in alignment with a Gaussian distribution), are directly obtained from the distributions, as indicated by the vertical red lines in Fig. 5.10. The reach of this limit also defines the most minimal size of the scan range, as mentioned during the explanation of the FC method.

In the application of this method to the sixteen angular observables, a step granularity of 0.01 is chosen over a default range of ± 1.0 around the best-fit value obtained by the maximum-likelihood fit. The range of many observables is extended a posteriori to cover the full 68.3% confidence interval. For each step, 500 samples of pseudodata are generated for the FC method. Again, the number is a trade-off between statistical precision and the computational resources available.

As already mentioned, six profiles of FC scans are exemplarily shown in Fig. 5.10, to highlight different features of the profiles. The figure also contains profiles in the bottom row, which are the result of a parabolically shaped delta-log-likelihood profile. For the latter profiles, the statistical uncertainty from the fit perfectly matches the values obtained

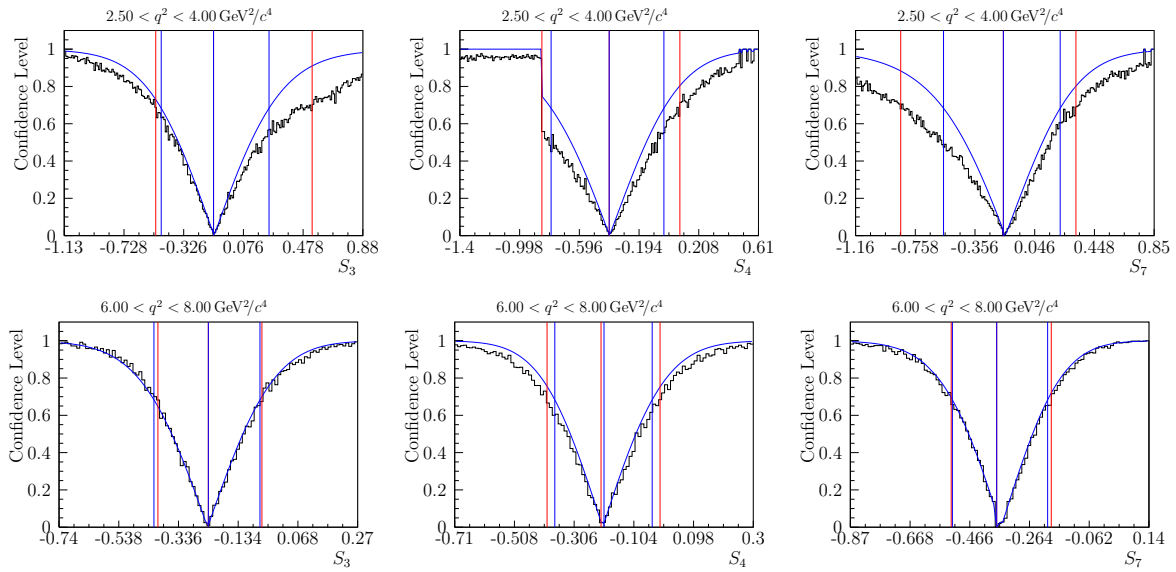

Figure 5.10: Exemplary profiles of Feldman-Cousins scans of the angular observables (left) S_3, (centre) S_4 and (right) S_7 in the q^2 intervals (top) $2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$ and (bottom) $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$. The black histogram shows the results of Feldman-Cousins scans around the central value of the observable as obtained by the maximum-likelihood fit, which is situated at the minimum of the profiles and indicated by a vertical line. Overlaid are the $\Delta \log \mathcal{L}$ profiles in blue. The additional two vertical lines indicate the 68.3% confidence intervals – corresponding to a 1σ interval of a Gaussian shaped likelihood. Clear deviations between the black and blue curve are visible in the top row, which are signs of non-Gaussian confidence intervals. The edge in the top centre profile show the effect of a boundary in the allowed parameter space. The scan profiles in the bottom row show good agreement for both profiles; $\Delta \log \mathcal{L}$ and FC scans. Also the profiles are equivalents of Gaussian-like likelihood profiles possessing near-perfect symmetry. More details in the text.

by the FC scans, making the FC profile scans optional for these parameters. For these well behaved parameters, the FC scans are only performed to create consistency throughout all angular observables.

The FC profile scans for all sixteen angular observables are given in Figs. A.15 to A.30 in Appendix A.6.

All final results of the two sets of angular observables in Chapter 7 possess asymmetric statistical uncertainties determined by Feldman-Cousins profile scans.

6 Evaluation of systematic uncertainties

In addition to the statistical uncertainty due to the finite size of the data sample, systematic effects add uncertainties to the measured values of the angular observables. Some systematic effects cannot be corrected, hence systematic uncertainties are estimated by evaluating pseudoexperiments with significantly larger statistics than the data sample. In this way, one systematic effect can be isolated from other effects of both statistical and systematic nature.

In the following, potential sources of systematic uncertainties are introduced, explained and their effect is quantified. The numerical values of all studies are given for all observables in all q^2 intervals. In Section 6.10, the systematics are combined and the total values of the systematic uncertainties for all angular observables are given.

In general, the systematic studies fit into one of three categories, which differ in how the uncertainty is estimated. In the first category, the effect in-question is used to create different sets of randomly varied pseudodata samples, which are all fitted with the nominal fit configuration. The uncertainty is given by the spread of values of angular observables. In the second category, the configuration of the fit changes by a given effect. Sets of pseudoexperiments are fitted once with the nominal configuration and once with the alternative configuration. The averaged difference between results from the two fits is assigned as the systematic uncertainty. In the last category, pseudoexperiments mimic the data sample in their size and in the values of generated parameters. These pseudoexperiments are introduced and discussed in Section 5.7.1. Average shifts between generated and measured values are assigned as systematic uncertainties.

6.1 Angular acceptance correction

The correction of angular acceptance effects, as introduced in Section 5.3, is crucial for the measurements of the angular observables. However, the strong warping of the angular phase space by the application of angular acceptance weights gives rise to potential

systematic effects. Therefore, the maximum order of polynomials in the parametrisation of the acceptance correction (see Section 5.3), the limited size of the simulation sample used for the parametrisation and the effect of data-simulation weighting (see Section 4.1.2) on the simulation sample are addressed in the following Sections 6.1.1 to 6.1.3.

6.1.1 Angular acceptance parametrisation

The parametrisation of the angular acceptance effects in the four-dimensional space of $\cos\theta_K$, $\cos\theta_\ell$, ϕ and q^2 by Legendre polynomials are defined in Eq. (5.20) in Section 5.3. The finite order of the polynomials is chosen to minimise the number of coefficients and the complexity of the parametrisation, as well as preventing the parametrisation to describe statistical fluctuations. At the same time, the orders are required to reproduce the shape of the acceptance effect in four dimensions. In order to estimate the effect of a finite description on the angular observables, the maximum order in each dimension is increased in a test scenario by two with respect to the nominal orders. There is no certainty whether one of the two configurations is more accurate than the other. Therefore, the differences between both are assigned as systematic uncertainties.
For this, the angular observables from about 200 samples of pseudodata with high candidate statistics are extracted by two maximum-likelihood fits, one with weights obtained by the nominal parametrisation of the acceptance, one with weights from the acceptance parametrisation using the incremented maximum order of polynomials. The absolute value of the average shift in the difference between the two methods is assigned as the systematic uncertainty and the numerical values are given in Table 6.1.

6.1.2 Statistical precision of simulation sample

Unfortunately, official LHCb simulation are not available in infinite sample sizes. The phase-space MC simulation, which is used to parametrise the angular acceptance effects, is no exception.
Even though several million of events are generated, the candidate selection yields only some ten thousands of candidates in each simulation sample. Spreading the given statistics onto the four-dimensional angular phase space in order to parametrise the acceptance effect results in low occupation of some regions. This consequently creates statistical fluctuations in the parametrisation, which are studied using the bootstrapping method. The method is introduced and applied to the determination of the correlation matrices of all angular observables in Section 5.5. Every randomly shuffled sample of the phase-space MC simulation is parametrised to extract a series of sets of angular acceptance correction coefficients. This process is applied individually to all six sub-samples of the angular acceptance correction, which are defined in Section 5.3.

Table 6.1: Systematic uncertainty assigned to the limited polynomial order in the parametrisation of the angular acceptance correction. See text for more details.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.033	0.009	0.012	0.017	0.002	0.004	0.017	0.010
$[1.1, 2.5]$	0.025	0.022	0.022	0.042	0.032	0.019	0.030	0.003
$[2.5, 4.0]$	0.049	0.006	0.001	0.009	0.012	0.022	0.013	0.007
$[4.0, 6.0]$	0.017	0.001	0.014	0.019	0.001	0.010	0.020	0.011
$[6.0, 8.0]$	0.005	0.014	0.001	0.011	0.004	0.003	0.016	0.003
$[11.0, 12.5]$	0.004	0.011	0.018	0.019	0.002	0.010	0.008	0.004
$[15.0, 17.0]$	0.009	0.014	0.006	0.021	0.001	0.013	0.006	0.007
$[17.0, 19.0]$	0.016	0.028	0.052	0.032	0.011	0.005	0.014	0.011
$[1.1, 6.0]$	0.005	0.002	0.001	0.004	0.007	0.003	0.007	0.004
$[15.0, 19.0]$	0.024	0.023	0.006	0.010	0.013	0.004	0.006	0.001

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.034	0.088	0.001	0.003	0.023	0.092	0.006	0.011
$[1.1, 2.5]$	0.008	0.061	0.001	0.008	0.016	0.040	0.002	0.019
$[2.5, 4.0]$	0.013	0.069	0.014	0.015	0.021	0.023	0.016	0.003
$[4.0, 6.0]$	0.019	0.032	0.001	0.004	0.012	0.026	0.002	0.011
$[6.0, 8.0]$	0.011	0.046	0.002	0.007	0.012	0.013	0.002	0.009
$[11.0, 12.5]$	0.020	0.028	0.001	0.002	0.005	0.016	0.001	0.001
$[15.0, 17.0]$	0.012	0.039	0.001	0.003	0.021	0.013	0.005	0.010
$[17.0, 19.0]$	0.030	0.074	0.001	0.004	0.007	0.035	0.014	0.001
$[1.1, 6.0]$	0.002	0.025	0.001	0.007	0.007	0.009	0.023	0.025
$[15.0, 19.0]$	0.017	0.040	0.004	0.001	0.039	0.013	0.008	0.006

One set of pseudodata with large statistics is fitted repeatedly for all obtained sets of coefficients for the angular acceptance correction. The results for each angular observable are collected in histograms. The absolute value of the spread in the result distribution for each angular observable is assigned as the systematic uncertainty. The value of the spread is obtained from a unbinned maximum-likelihood fit using a Gaussian bell shape. In Table 6.2, the uncertainty values of this systematic effect are listed.

6.1.3 Simulation-data weights

Simulation are weighted to remove residual discrepancies between simulation and data as much as possible, as described in Section 4.1.2. The effect of the weights being applied to phase-space MC is investigated. For this, 200 samples of pseudodata are fitted twice, once with and once without the simulation-data weights applied. The systematic uncertainty is the averaged difference in the results of angular observables between both fits and is given in Table 6.3. The complete removal of simulation-data weights for the study is a slight exaggeration of the systematic. But the values of the systematic due to simulation-data weights of the phase-space simulation are not larger than other sources of systematic uncertainties and, hence, justifies this conservative estimation.

Table 6.2: Systematic uncertainty associated to the limited statistics of the phase-space MC samples. The uncertainty is obtained using a bootstrapping method.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.022	0.019	0.033	0.028	0.026	0.012	0.018	0.038
$[1.1, 2.5]$	0.022	0.006	0.018	0.016	0.011	0.007	0.013	0.003
$[2.5, 4.0]$	0.009	0.006	0.014	0.012	0.005	0.005	0.007	0.001
$[4.0, 6.0]$	0.011	0.005	0.014	0.013	0.006	0.005	0.009	0.001
$[6.0, 8.0]$	0.009	0.005	0.008	0.008	0.005	0.003	0.003	0.001
$[11.0, 12.5]$	0.011	0.005	0.009	0.008	0.008	0.004	0.008	0.001
$[15.0, 17.0]$	0.009	0.008	0.011	0.009	0.006	0.003	0.006	0.001
$[17.0, 19.0]$	0.017	0.015	0.016	0.018	0.011	0.010	0.017	0.003
$[1.1, 6.0]$	0.024	0.007	0.025	0.016	0.022	0.011	0.021	0.006
$[15.0, 19.0]$	0.012	0.008	0.021	0.015	0.010	0.005	0.008	0.003

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.019	0.053	0.024	0.038	0.077	0.062	0.026	0.040
$[1.1, 2.5]$	0.021	0.054	0.024	0.006	0.040	0.039	0.017	0.032
$[2.5, 4.0]$	0.010	0.059	0.022	0.007	0.038	0.034	0.013	0.017
$[4.0, 6.0]$	0.011	0.036	0.012	0.004	0.031	0.030	0.011	0.019
$[6.0, 8.0]$	0.009	0.026	0.005	0.002	0.017	0.018	0.006	0.008
$[11.0, 12.5]$	0.011	0.020	0.002	0.001	0.019	0.018	0.009	0.015
$[15.0, 17.0]$	0.010	0.025	0.003	0.002	0.022	0.020	0.007	0.013
$[17.0, 19.0]$	0.017	0.046	0.009	0.005	0.032	0.033	0.021	0.035
$[1.1, 6.0]$	0.024	0.046	0.042	0.033	0.055	0.035	0.024	0.045
$[15.0, 19.0]$	0.012	0.025	0.010	0.005	0.046	0.031	0.011	0.017

6.2 Angular resolution

The determination of the decay angles, $\cos\theta_K$, $\cos\theta_\ell$ and ϕ is limited by the precision of the reconstruction of the particle momenta. Using the *true* information [1] from simulated decays, the angular resolution is determined as the spread of the distribution $x - x_\mathrm{true}$, with x being $\cos\theta_K$, $\cos\theta_\ell$ or ϕ. The resolutions of all three angles are shown in Fig. 6.1 for 2016 DD and LL simulation along with the fit of a double Gaussian bell shape model. The fit is used to extract the numerical value of the resolution as the widths of the distributions. Values for the 2016 sample are given in the plots and are representative for all years. Resolution in LL samples is slightly more precise with respect to DD samples, as expected, due to the better quality of long tracks in the reconstruction.

For the study of a potential systematic effect of this resolution on the angular observables, generated angles in pseudoexperiments are smeared randomly by their measured resolutions. This method is repeated numerously and the systematic uncertainty is taken from the spread of the results. This uncertainty is the lowest of all investigated systematic effects

[1] Analogous to the particle identification number in Section 4.1, the information on the generated values of energy and momentum of generated particles is assigned to the reconstructed tracks and can be utilises for resolution measurements.

Table 6.3: Values of the systematic uncertainty from the weighting of phase-space simulation. To remove residual discrepancies between simulation and data, the weights are applied. The effect of this weighting is quantified by these systematic uncertainties.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.018	0.002	0.006	0.019	0.007	0.002	0.038	0.038
$[1.1, 2.5]$	0.003	0.001	0.003	0.003	0.002	0.001	0.001	0.001
$[2.5, 4.0]$	0.001	0.001	0.002	0.002	0.001	0.001	0.001	0.001
$[4.0, 6.0]$	0.003	0.001	0.006	0.003	0.001	0.001	0.002	0.001
$[6.0, 8.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[11.0, 12.5]$	0.001	0.001	0.002	0.002	0.001	0.001	0.001	0.001
$[15.0, 17.0]$	0.001	0.002	0.002	0.002	0.001	0.001	0.001	0.001
$[17.0, 19.0]$	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001
$[1.1, 6.0]$	0.003	0.001	0.003	0.002	0.002	0.001	0.001	0.001
$[15.0, 19.0]$	0.001	0.001	0.002	0.002	0.001	0.001	0.001	0.001

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.037	0.010	0.018	0.010	0.082	0.031	0.054	0.078
$[1.1, 2.5]$	0.009	0.102	0.062	0.144	0.094	0.050	0.069	0.006
$[2.5, 4.0]$	0.026	0.127	0.031	0.165	0.035	0.050	0.046	0.057
$[4.0, 6.0]$	0.016	0.051	0.056	0.124	0.038	0.061	0.042	0.041
$[6.0, 8.0]$	0.004	0.043	0.049	0.035	0.015	0.030	0.023	0.003
$[11.0, 12.5]$	0.016	0.072	0.039	0.026	0.055	0.033	0.019	0.033
$[15.0, 17.0]$	0.007	0.040	0.022	0.007	0.056	0.037	0.006	0.039
$[17.0, 19.0]$	0.010	0.022	0.022	0.007	0.090	0.054	0.019	0.057
$[1.1, 6.0]$	0.003	0.009	0.004	0.010	0.018	0.004	0.002	0.003
$[15.0, 19.0]$	0.001	0.004	0.005	0.001	0.004	0.004	0.001	0.003

and their values are given in Table 6.4.

6.3 Mass model

The effective description of the signal mass peak by a two-tailed CB function, defined in Eq. (4.7), shows excellent agreement with the observed distribution of the invariant mass of the signal decay, as proved in Figs. 4.3, 4.7, 4.8 and 5.9. However, a systematic uncertainty arising from the choice of a signal mass model is probed by generating pseudodata with a Gaussian bell shape as mass model for the signal. The 200 generated samples are fitted with the nominal fit model, i.e. a two-tailed CB function, and the resulting shift in the mean position of the angular observables is assigned as a systematic uncertainty. The numerical values of this study are given in Table 6.5.

In addition to the systematic uncertainty from modelling the signal component, the upper limit of the $m(K^0_\mathrm{S}\pi^+\mu^+\mu^-)$ acceptance window at $6000\,\mathrm{MeV}/c^2$ is investigated. For this, pairs of pseudodata samples – always one with the nominal mass cut at $6000\,\mathrm{MeV}/c^2$ and one at $7000\,\mathrm{MeV}/c^2$ – are generated. The increased upper mass sideband is thought

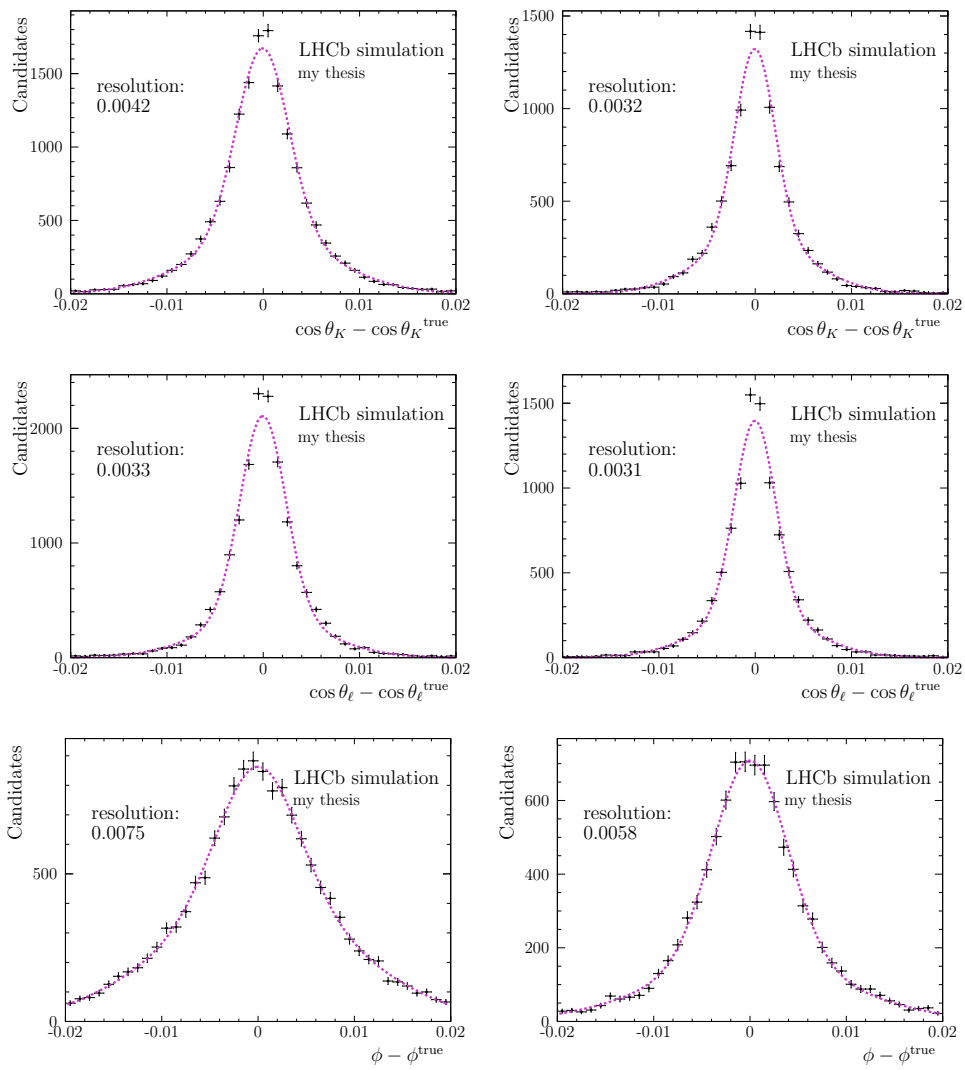

Figure 6.1: Angular resolution of the three angles, (top) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (bottom) ϕ in 2016 simulation as obtained from comparing the reconstructed with the true angles. The columns show the resolution for the (left) DD and (right) LL sample. The precision is slightly better in $\cos\theta_K$ for LL data and more significantly in the angle ϕ. The effect of this resolution is studied and found to be very small compared to other systematic effects. The results in Table 6.4 are nevertheless added to the total systematic uncertainty.

to stabilise the determination of the shape of the combinatorial background component, therefore the signal fraction, f_{sig}, and consequently improve the statistical precision on the angular observables. The tests however showed no improvement in the precision on the angular observables. As a result of this, the upper limit at 6000 MeV/c^2 for $m(K^0_S\pi^+\mu^+\mu^-)$ is established.

Table 6.4: Systematic uncertainties assigned to the non-zero resolution of the decay angles. The systematic effect is the smallest of all systematic uncertainties. Most values in this table would be zero if not rounded up to the lowest significant digit given in the tables.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.001	0.001	0.001	0.001	0.002	0.001	0.003	0.001
$[1.1, 2.5]$	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001
$[2.5, 4.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[4.0, 6.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[6.0, 8.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[11.0, 12.5]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[15.0, 17.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[17.0, 19.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[1.1, 6.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[15.0, 19.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.001	0.003	0.002	0.001	0.004	0.003	0.003	0.007
$[1.1, 2.5]$	0.001	0.003	0.002	0.002	0.003	0.002	0.002	0.004
$[2.5, 4.0]$	0.001	0.004	0.001	0.002	0.002	0.002	0.002	0.002
$[4.0, 6.0]$	0.001	0.003	0.001	0.002	0.002	0.001	0.001	0.002
$[6.0, 8.0]$	0.001	0.002	0.001	0.001	0.002	0.001	0.001	0.001
$[11.0, 12.5]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[15.0, 17.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[17.0, 19.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$[1.1, 6.0]$	0.001	0.003	0.001	0.002	0.002	0.001	0.001	0.002
$[15.0, 19.0]$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

6.4 Angular background model

The factorisable Chebyshev polynomial description of the background component in the angular dimensions is limited to second order in $\cos\theta_K$ and $\cos\theta_\ell$ and is uniform in the angle ϕ. This description is chosen to be minimal, while still satisfyingly describing the background component. The utilisation is confirmed by evaluating the fit results from pure background samples in the lower, upper and combined mass sidebands. For this evaluation, the background coefficients from different background samples are found to be statistically compatible. And correlations between coefficients from different angles are compatible with a correlation-free hypothesis.

In addition to these confirmations, the effect of higher order polynomial components in the background model on the angular observables is tested by generating pseudodata with the maximum order of the background component increased by one in each dimension, *i.e.* third order polynomials in $\cos\theta_K$ and $\cos\theta_\ell$, and a linear component in ϕ. The extra polynomials are scaled by an absolute coefficient of 0.1, which is a conservative estimation of possible contributions from higher order polynomials. These samples are fitted with

Table 6.5: Systematic uncertainties obtained from the differences in the results of angular observables when generating pseudodata with a Gaussian bell model for the mass peak of the B meson and fitting the $m(K^0_S\pi^+\mu^+\mu^-)$ distribution with a two-tailed CB function.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
$[0.10, 0.98]$	0.001	0.001	0.003	0.001	0.005	0.001	0.001	0.002
$[1.1, 2.5]$	0.001	0.002	0.002	0.009	0.002	0.002	0.001	0.001
$[2.5, 4.0]$	0.001	0.004	0.001	0.001	0.003	0.002	0.001	0.001
$[4.0, 6.0]$	0.002	0.001	0.003	0.006	0.001	0.003	0.001	0.004
$[6.0, 8.0]$	0.003	0.001	0.009	0.003	0.003	0.009	0.004	0.001
$[11.0, 12.5]$	0.003	0.003	0.007	0.008	0.001	0.001	0.001	0.001
$[15.0, 17.0]$	0.001	0.004	0.008	0.007	0.004	0.001	0.001	0.001
$[17.0, 19.0]$	0.002	0.005	0.004	0.005	0.005	0.004	0.001	0.002
$[1.1, 6.0]$	0.001	0.002	0.002	0.006	0.004	0.003	0.001	0.002
$[15.0, 19.0]$	0.001	0.003	0.006	0.002	0.003	0.002	0.005	0.003

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.003	0.005	0.001	0.005	0.003	0.011	0.001	0.007
$[1.1, 2.5]$	0.002	0.033	0.005	0.008	0.005	0.005	0.001	0.002
$[2.5, 4.0]$	0.001	0.052	0.010	0.002	0.003	0.008	0.006	0.012
$[4.0, 6.0]$	0.003	0.038	0.001	0.009	0.014	0.006	0.001	0.012
$[6.0, 8.0]$	0.002	0.021	0.005	0.009	0.003	0.001	0.002	0.005
$[11.0, 12.5]$	0.002	0.011	0.003	0.003	0.006	0.012	0.007	0.001
$[15.0, 17.0]$	0.003	0.007	0.006	0.003	0.009	0.011	0.004	0.008
$[17.0, 19.0]$	0.001	0.013	0.002	0.001	0.013	0.009	0.002	0.004
$[1.1, 6.0]$	0.001	0.041	0.007	0.001	0.011	0.006	0.006	0.012
$[15.0, 19.0]$	0.004	0.005	0.001	0.004	0.012	0.016	0.005	0.009

the nominal configuration and the shift in the mean of the resulting values of the angular observables is assigned as a systematic uncertainty, with the numbers given in Table 6.6.

6.5 Peaking background veto

The veto against peaking background pollution from $B^0 \to K^0_S\mu^+\mu^-$ decays with an additional random pion is described in the particle identification paragraph of Section 4.3.3 and the cut to the invariant mass, $m(K^0_S\mu^+\mu^-)$, is illustrated in Fig. 4.2. As seen in latter figure, the veto successfully removes contribution from the peaking background, but in addition cuts away combinatorial background underneath the $B^0 \to K^0_S\mu^+\mu^-$ peak. This results in a dent in the upper mass sideband of the invariant mass of the B^+ meson, $m(K^0_S\pi^+\mu^+\mu^-)$, in the region between 5400–5600 MeV/c^2, as shown in the fit to the reference channel in Figs. 4.3 and 5.9. Gladly, no edges or non-continuous distortions in the angular distributions of the background are introduced by the veto cut, and the parametrisation of the angular background is studied in Section 6.4.

In the following, solely the effect of a dent in the background region above the signal peak

Table 6.6: Systematic uncertainty estimated for an under-estimation of the complexity of angular background shapes. The bias in the observables obtained from fits with a nominal background model in pseudoexperiments that use an additional polynomial order in the angular background shapes at generation.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.002	0.004	0.002	0.003	0.002	0.003	0.003	0.009
$[1.1, 2.5]$	0.001	0.006	0.006	0.006	0.001	0.004	0.002	0.014
$[2.5, 4.0]$	0.001	0.005	0.003	0.002	0.003	0.003	0.003	0.018
$[4.0, 6.0]$	0.001	0.004	0.002	0.001	0.001	0.004	0.001	0.017
$[6.0, 8.0]$	0.003	0.004	0.004	0.005	0.003	0.004	0.006	0.015
$[11.0, 12.5]$	0.001	0.001	0.001	0.002	0.002	0.001	0.006	0.013
$[15.0, 17.0]$	0.002	0.001	0.002	0.005	0.002	0.003	0.005	0.011
$[17.0, 19.0]$	0.003	0.006	0.003	0.001	0.002	0.001	0.002	0.010
$[1.1, 6.0]$	0.001	0.002	0.001	0.003	0.003	0.003	0.001	0.011
$[15.0, 19.0]$	0.003	0.003	0.004	0.001	0.001	0.004	0.001	0.020

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.004	0.003	0.003	0.011	0.001	0.009	0.003	0.004
$[1.1, 2.5]$	0.002	0.022	0.016	0.093	0.006	0.015	0.001	0.009
$[2.5, 4.0]$	0.002	0.056	0.004	0.113	0.004	0.014	0.003	0.004
$[4.0, 6.0]$	0.001	0.030	0.006	0.060	0.006	0.001	0.005	0.007
$[6.0, 8.0]$	0.001	0.018	0.012	0.033	0.008	0.017	0.001	0.001
$[11.0, 12.5]$	0.003	0.001	0.004	0.016	0.003	0.011	0.004	0.001
$[15.0, 17.0]$	0.004	0.002	0.004	0.011	0.011	0.006	0.006	0.004
$[17.0, 19.0]$	0.003	0.014	0.002	0.010	0.007	0.006	0.003	0.006
$[1.1, 6.0]$	0.002	0.012	0.005	0.025	0.005	0.012	0.007	0.010
$[15.0, 19.0]$	0.004	0.018	0.001	0.031	0.001	0.017	0.013	0.014

in the invariant mass distribution of the B meson is investigated. For this, an artificial dent in the upper mass sideband between 5400–5600 MeV/c^2 is implemented via subtracting a Gaussian bell shape from the exponential curve in the generation of pseudodata. The size of the artificial dent is approximated to match the amount of removed combinatorial background candidates by the veto in data. The averaged bias of the nominal fit to hundreds of pseudodata sets is assigned as the systematic uncertainty to this veto, with the numerical values given in Table 6.7.

6.6 Simulation-data differences

In Section 4.1.2 the efforts to remove discrepancies between simulation and data are explained. However, in addition to the already discussed effect of this weighting on the angular acceptance correction in Section 6.1.3, two more systematic studies are performed in the context of simulation-data discrepancies.

Table 6.7: Systematic uncertainty assigned to the effect of the veto against peaking background applied in the paragraph on particle identification in Section 4.3.3. The veto cuts away combinatorial background, leaving a small dent in the upper mass sideband, as seen in Figs. 4.3 and 5.9. The effect of an artificially created dent in the mass sideband is investigated and the numerical values of the corresponding systematic uncertainty are given.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.001	0.001	0.001	0.001	0.003	0.001	0.006	0.008
$[1.1, 2.5]$	0.003	0.006	0.006	0.005	0.006	0.005	0.007	0.001
$[2.5, 4.0]$	0.005	0.004	0.001	0.001	0.001	0.008	0.004	0.004
$[4.0, 6.0]$	0.001	0.001	0.002	0.003	0.001	0.005	0.001	0.005
$[6.0, 8.0]$	0.001	0.004	0.002	0.001	0.003	0.007	0.001	0.002
$[11.0, 12.5]$	0.002	0.001	0.002	0.002	0.001	0.001	0.005	0.002
$[15.0, 17.0]$	0.003	0.001	0.001	0.002	0.003	0.006	0.008	0.003
$[17.0, 19.0]$	0.001	0.004	0.006	0.001	0.001	0.002	0.001	0.005
$[1.1, 6.0]$	0.001	0.002	0.001	0.003	0.002	0.001	0.005	0.002
$[15.0, 19.0]$	0.001	0.003	0.008	0.002	0.001	0.001	0.002	0.002

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.002	0.006	0.002	0.003	0.005	0.004	0.001	0.003
$[1.1, 2.5]$	0.001	0.037	0.010	0.032	0.004	0.004	0.001	0.001
$[2.5, 4.0]$	0.002	0.052	0.016	0.003	0.020	0.001	0.007	0.011
$[4.0, 6.0]$	0.002	0.061	0.003	0.005	0.040	0.017	0.008	0.001
$[6.0, 8.0]$	0.002	0.059	0.001	0.001	0.026	0.035	0.001	0.002
$[11.0, 12.5]$	0.005	0.021	0.001	0.010	0.019	0.027	0.003	0.005
$[15.0, 17.0]$	0.001	0.001	0.001	0.002	0.023	0.020	0.003	0.006
$[17.0, 19.0]$	0.003	0.001	0.001	0.001	0.016	0.012	0.005	0.007
$[1.1, 6.0]$	0.002	0.008	0.011	0.003	0.009	0.002	0.001	0.001
$[15.0, 19.0]$	0.001	0.004	0.001	0.005	0.009	0.018	0.014	0.001

6.6.1 Kaon momentum discrepancy

Even after the weighting of simulation, as explained in Section 4.1.2, some residual differences in the transverse momentum distribution of the K^0_S candidate are observed between simulation and data for DD tracks. The discrepancy at low p_T of the K^0_S candidate is thought to result from the difficulty of modelling low energetic pions from K^0_S decays, which are only reconstructed downstream of the VELO. Fig. 6.2 shows the simulation and data distribution of $p_\mathrm{T}(K^0_\mathrm{S})$ in the DD category for the years 2012 and 2016. To test the impact of this discrepancy, weights are generated from the ratio of the transverse momenta, $p_\mathrm{T}(K^0_\mathrm{S})$, in simulation and background-subtracted data in Fig. 6.2. Always two fits to each of some hundred pseudodata samples are performed, once with the additional $p_\mathrm{T}(K^0_\mathrm{S})$ weights and once with the nominal fit configuration without these weights. The averaged difference in the results is assigned as a systematic uncertainty. The values of this systematic uncertainty for all angular observables are given in Table 6.8.

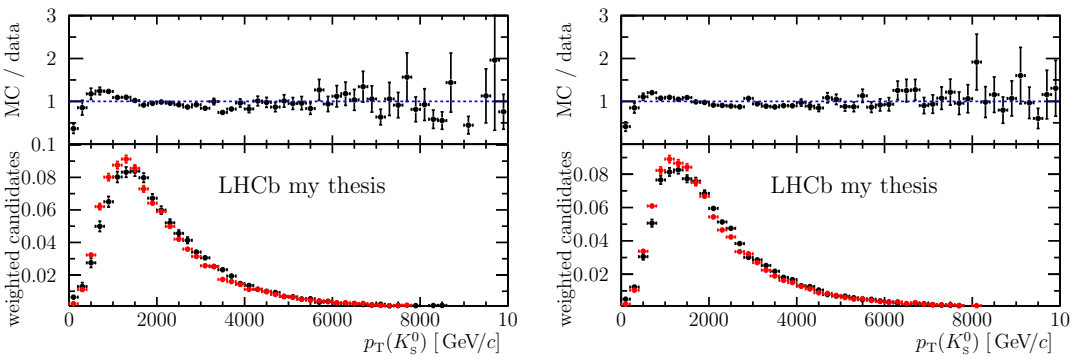

Figure 6.2: Ratio and overlaid distributions of the transverse momentum (p_T) of the K^0_S candidate in (red) weighted simulation of the signal decay and (black) background-subtracted data including the $c\bar{c}$ resonant regions. Both samples are DD category in (left) 2012 and (right) 2018. The discrepancy at low p_T is clearly visible and representative for samples from all years. The ratios are used to determine additional weights for simulation, which are used in a systematic study. The corresponding uncertainties from this study are given in Table 6.8.

6.6.2 Trigger efficiency

Slight discrepancies are observed in the trigger line `L0Muon` of the first trigger stage between the emulation of the trigger in the generation of simulation and the actual trigger used in data recording. The efficiency of this trigger line as a function of the maximum transverse momenta of the two muons is slightly higher for simulation as it is in data, especially towards lower maximal p_T of μ^+ and μ^-. As before in Section 6.6.1, weights are used to counteract the difference in the efficiencies and their effect on the results of the angular observables is assigned as a systematic uncertainty. The values are determined from the averaged difference of two fits to 200 pseudodata samples, one fit with the weights to compensate the trigger efficiency difference and one fit with the nominal configuration. The results are given in Table 6.9.

6.7 Constraint on the S-wave fraction

The S-wave fraction, F_S, is constraint in the four-dimensional maximum-likelihood fit to a value obtained in a two-dimensional fit to $m(K^0_S\pi^+\mu^+\mu^-)$ and $m(K^0_S\pi^+)$, as explained in Section 5.6. In between the two fits, the value of F_S is scaled from the measurements in three larger q^2 intervals to the nominal intervals. A priori, this strategy of fitting the data twice is considered sensitive to introducing bias effects along with over- or under-coverage of the statistical uncertainty. In Section 5.6, however, the need for this compromise in the fit strategy is explained. Therefore, pseudoexperiments of identical statistics as estimated in data are performed to investigate the effect of this, as described in Section 5.7. For the

Table 6.8: Systematic uncertainty due to the residual discrepancy in the transverse momentum of the K^0_S candidate between simulation and data. Weights are used to compensate this discrepancy and their effect on the angular observables is given by these values.

q^2 [GeV$^2/c^4$]	F_L	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
$[0.10, 0.98]$	0.001	0.001	0.005	0.001	0.001	0.001	0.004	0.001
$[1.1, 2.5]$	0.003	0.008	0.003	0.001	0.001	0.003	0.003	0.002
$[2.5, 4.0]$	0.002	0.005	0.005	0.001	0.001	0.004	0.001	0.004
$[4.0, 6.0]$	0.001	0.005	0.002	0.001	0.002	0.001	0.002	0.004
$[6.0, 8.0]$	0.005	0.003	0.001	0.001	0.001	0.002	0.001	0.003
$[11.0, 12.5]$	0.003	0.004	0.003	0.002	0.003	0.001	0.001	0.002
$[15.0, 17.0]$	0.006	0.003	0.001	0.002	0.001	0.003	0.001	0.001
$[17.0, 19.0]$	0.004	0.001	0.007	0.004	0.003	0.001	0.002	0.001
$[1.1, 6.0]$	0.002	0.001	0.003	0.001	0.002	0.004	0.005	0.001
$[15.0, 19.0]$	0.001	0.001	0.001	0.003	0.001	0.001	0.002	0.003

q^2 [GeV$^2/c^4$]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.002	0.005	0.002	0.001	0.001	0.009	0.001	0.003
$[1.1, 2.5]$	0.002	0.053	0.001	0.006	0.006	0.001	0.008	0.025
$[2.5, 4.0]$	0.001	0.001	0.002	0.003	0.003	0.012	0.004	0.001
$[4.0, 6.0]$	0.001	0.002	0.003	0.001	0.020	0.013	0.010	0.012
$[6.0, 8.0]$	0.004	0.022	0.002	0.002	0.001	0.012	0.002	0.005
$[11.0, 12.5]$	0.002	0.003	0.001	0.006	0.002	0.007	0.009	0.007
$[15.0, 17.0]$	0.006	0.001	0.003	0.003	0.004	0.004	0.003	0.005
$[17.0, 19.0]$	0.001	0.001	0.004	0.001	0.011	0.010	0.003	0.006
$[1.1, 6.0]$	0.006	0.022	0.002	0.005	0.001	0.005	0.011	0.030
$[15.0, 19.0]$	0.004	0.025	0.001	0.005	0.007	0.001	0.001	0.005

investigation of potential bias in the results, samples of pseudodata are generated with values of the angular observables equal to SM predictions from Ref. [101] and with the identical statistics as in data. Every sample is fitted with the nominal fit configuration, a two-dimensional fit followed by a four-dimensional fit. As explained in Section 5.7, pull distributions are generated from the results of pseudoexperiment studies. In these distributions, some biases in their mean positions are observed. Consequently, a systematic uncertainty equivalent to the shift in residual distributions ($x_{meas} - x_{gen}$) between measured and generated values is assigned to the method of constraining the S-wave fraction, F_S. The observed bias in these pseudoexperiment studies originates indubitable from the constraint on F_S, because when artificially constraining F_S in the four-dimensional fits with unrealistically high precision, the bias is fully removed from the pseudoexperiment studies.

All values of the systematic uncertainties on the angular observables in the eight narrow q^2 intervals are given in Table 6.10, as the two extra large bins do not require the scaling of F_S.

Table 6.9: Systematic uncertainty due to simulation-data differences in the selection efficiency of the `L0Muon` line in the first trigger level. The effect is investigated by applying weights to counteract the efficiency discrepancy and determining the resulting shift of the angular observable values.

q^2 [GeV$^2/c^4$]	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
$[0.10, 0.98]$	0.003	0.002	0.002	0.003	0.004	0.001	0.004	0.001
$[1.1, 2.5]$	0.003	0.001	0.005	0.001	0.001	0.002	0.006	0.001
$[2.5, 4.0]$	0.002	0.002	0.002	0.001	0.002	0.001	0.003	0.001
$[4.0, 6.0]$	0.001	0.001	0.002	0.004	0.001	0.002	0.006	0.002
$[6.0, 8.0]$	0.001	0.001	0.003	0.002	0.001	0.008	0.003	0.001
$[11.0, 12.5]$	0.002	0.004	0.004	0.001	0.001	0.003	0.001	0.001
$[15.0, 17.0]$	0.001	0.001	0.001	0.001	0.001	0.005	0.002	0.001
$[17.0, 19.0]$	0.001	0.001	0.004	0.001	0.001	0.004	0.002	0.003
$[1.1, 6.0]$	0.004	0.004	0.001	0.002	0.001	0.003	0.001	0.001
$[15.0, 19.0]$	0.002	0.004	0.001	0.004	0.002	0.002	0.003	0.001

q^2 [GeV$^2/c^4$]	F_{L}	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.001	0.004	0.001	0.001	0.007	0.002	0.002	0.004
$[1.1, 2.5]$	0.001	0.001	0.007	0.011	0.002	0.001	0.004	0.013
$[2.5, 4.0]$	0.004	0.022	0.015	0.010	0.022	0.004	0.003	0.001
$[4.0, 6.0]$	0.004	0.019	0.008	0.003	0.007	0.022	0.005	0.007
$[6.0, 8.0]$	0.004	0.004	0.009	0.018	0.003	0.005	0.004	0.012
$[11.0, 12.5]$	0.001	0.006	0.001	0.004	0.006	0.001	0.001	0.002
$[15.0, 17.0]$	0.002	0.001	0.003	0.004	0.002	0.002	0.003	0.001
$[17.0, 19.0]$	0.001	0.006	0.003	0.004	0.004	0.002	0.011	0.001
$[1.1, 6.0]$	0.006	0.011	0.007	0.007	0.003	0.010	0.003	0.001
$[15.0, 19.0]$	0.001	0.004	0.001	0.003	0.003	0.003	0.004	0.001

6.8 Bias of the likelihood fit

In analogy to the studies with pseudoexperiments in the previous paragraph on the constraints of the S-wave fraction, F_{S}, the generated values of the parameters are set to best-fit values from the ultimate fits to data. In total, 1000 samples of pseudodata are evaluated for each angular fold and the resulting values of the angular observables are compared to the generated value and the mean and spread of the results is investigated. Shifts in these residual distributions, $x_{meas} - x_{gen}$, are observed throughout most angular observables and in most q^2 intervals. Unfortunately, the biases cannot be parametrised as a function of the generated value of the observables, hence a systematic uncertainty is assigned to the largest observed bias in pseudoexperiments generated with the best-fit values or parameters equivalent to a new physics scenario with a shift in the real part of the effective vector coupling strength, $\Delta\mathrm{Re}(\mathcal{C}_9)$. The values of the systematic uncertainty due to the bias of the maximum-likelihood fit are given in Table 6.11.

Table 6.10: Using the constraints on the S-wave fraction, F_{S}, obtained in the two-dimensional mass fit in larger q^2 intervals and scaled by the ratio of the fraction of polarised K^{*+} mesons, F_{L}, a systematic uncertainty of this method is estimated. Details on the strategy of the F_{S} constraints are given in Section 5.6. The scaling is only applied to the fits in the narrower q^2 intervals and consequently a systematic uncertainty is applied only there.

q^2 [GeV^2/c^4]	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
$[0.10, 0.98]$	0.031	0.001	0.005	0.002	0.002	0.003	0.004	0.003
$[1.1, 2.5]$	0.042	0.001	0.012	0.005	0.016	0.001	0.007	0.001
$[2.5, 4.0]$	0.039	0.001	0.021	0.010	0.001	0.018	0.015	0.004
$[4.0, 6.0]$	0.001	0.008	0.014	0.014	0.004	0.003	0.010	0.001
$[6.0, 8.0]$	0.004	0.008	0.004	0.010	0.001	0.007	0.005	0.007
$[11.0, 12.5]$	0.007	0.007	0.004	0.007	0.001	0.007	0.004	0.011
$[15.0, 17.0]$	0.001	0.003	0.001	0.006	0.003	0.008	0.008	0.003
$[17.0, 19.0]$	0.001	0.014	0.013	0.019	0.005	0.006	0.001	0.008

q^2 [GeV^2/c^4]	F_{L}	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.019	0.023	0.007	0.010	0.001	0.026	0.008	0.004
$[1.1, 2.5]$	0.007	0.005	0.134	0.042	0.037	0.008	0.014	0.001
$[2.5, 4.0]$	0.019	0.129	0.103	0.002	0.010	0.034	0.008	0.033
$[4.0, 6.0]$	0.004	0.030	0.011	0.004	0.032	0.003	0.001	0.005
$[6.0, 8.0]$	0.006	0.007	0.015	0.006	0.020	0.031	0.008	0.003
$[11.0, 12.5]$	0.003	0.016	0.004	0.017	0.015	0.013	0.008	0.009
$[15.0, 17.0]$	0.007	0.001	0.006	0.004	0.027	0.031	0.013	0.007
$[17.0, 19.0]$	0.006	0.010	0.036	0.002	0.082	0.051	0.001	0.005

6.9 Difference between angular folds

As shown in Table 5.2, the angular observables F_{L} and S_3 (P_1) are accessible in all five folds, but only one value is stated as the final result in Chapter 7. With the help of pseudodata samples, it is shown, that the values of F_{L}, S_3 and P_1 do not depend on the choice of the fold. No bias between the folds is observed. The choice, which fold the stated values is taken from, is made by the fact, that the fold with index 4 cancels most fit parameters, as some angular background coefficients cancel in the $\cos\theta_K$ dimension only for this fold. Due to this lower number of degrees of freedom, the fit shows the most stable behaviour in fold with index 4. No systematic uncertainty is assigned to this decision.

6.10 Summary of systematic studies

As the uncorrelated systematic uncertainties presented in the previous paragraphs, in Sections 6.1.1 to 6.8, are orthogonal to one another, they are added up quadratically to a total systematic uncertainty for each measurement of an angular observable in every q^2

Table 6.11: Using the best-fit values obtained in data as values at generation of pseudoexperiments, the pull distributions of some angular observables show significant bias. As the bias cannot be parametrised in terms of the initial values of the generated values, a systematic uncertainty as large as the largest bias in the studies with pseudodata is assigned.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.026	0.006	0.006	0.004	0.002	0.005	0.002	0.008
$[1.1, 2.5]$	0.035	0.007	0.008	0.012	0.007	0.012	0.021	0.002
$[2.5, 4.0]$	0.003	0.008	0.027	0.010	0.003	0.003	0.014	0.004
$[4.0, 6.0]$	0.002	0.003	0.042	0.005	0.001	0.001	0.002	0.008
$[6.0, 8.0]$	0.002	0.009	0.008	0.001	0.004	0.012	0.008	0.005
$[11.0, 12.5]$	0.005	0.003	0.004	0.001	0.050	0.003	0.001	0.018
$[15.0, 17.0]$	0.011	0.019	0.003	0.013	0.006	0.004	0.002	0.009
$[17.0, 19.0]$	0.009	0.004	0.010	0.005	0.001	0.004	0.007	0.003
$[1.1, 6.0]$	0.002	0.002	0.007	0.007	0.001	0.001	0.003	0.001
$[15.0, 19.0]$	0.002	0.013	0.052	0.002	0.035	0.003	0.003	0.005

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.028	0.008	0.010	0.021	0.027	0.023	0.011	0.023
$[1.1, 2.5]$	0.023	0.278	0.005	0.003	0.023	0.069	0.037	0.013
$[2.5, 4.0]$	0.025	0.003	0.001	0.033	0.126	0.052	0.013	0.012
$[4.0, 6.0]$	0.013	0.269	0.020	0.057	0.049	0.032	0.006	0.005
$[6.0, 8.0]$	0.004	0.032	0.006	0.019	0.046	0.012	0.004	0.007
$[11.0, 12.5]$	0.007	0.011	0.017	0.026	0.038	0.012	0.002	0.005
$[15.0, 17.0]$	0.003	0.035	0.009	0.006	0.012	0.016	0.011	0.006
$[17.0, 19.0]$	0.006	0.010	0.002	0.008	0.029	0.089	0.016	0.025
$[1.1, 6.0]$	0.003	0.020	0.001	0.005	0.017	0.003	0.002	0.005
$[15.0, 19.0]$	0.001	0.030	0.033	0.004	0.083	0.009	0.015	0.005

interval.

$$\sigma_{\mathrm{t}ot} = \sqrt{\sum_{i \in \mathrm{sys}} \sigma_i^2} \tag{6.1}$$

Table 6.12 shows the resulting numerical values of the total systematic uncertainties. No individual systematic effect sets the total size of the uncertainties, but different uncertainties are of similar size and impact angular observables individually.
The total uncertainty on the results are clearly dominated by the statistical precision as it is shown in tables and figures of Chapter 7.

6.11 Blinded studies on result compatibility

Before the results are presented in the next Chapter 7, sub-categories of the complete sample are evaluated. The results of complementary sub-samples are compared as a *blind* cross-check. The term *blind* denotes that differences and ratios in the results are considered

Table 6.12: Total systematic uncertainties for all angular observables in all q^2 intervals. The values combined by adding the individual uncertainties from Tables 6.1 to 6.11 in quadrature. No individual systematic uncertainty is dominating. These values are added orthogonal to the statistical uncertainty of the results, given in Chapter 7. It has to be noted that the precision of the results is clearly limited by the statistical precision.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.060	0.023	0.037	0.038	0.028	0.015	0.046	0.056
$[1.1, 2.5]$	0.065	0.027	0.034	0.048	0.038	0.024	0.041	0.015
$[2.5, 4.0]$	0.064	0.015	0.038	0.022	0.015	0.030	0.026	0.021
$[4.0, 6.0]$	0.021	0.013	0.049	0.029	0.008	0.014	0.025	0.023
$[6.0, 8.0]$	0.013	0.021	0.016	0.018	0.010	0.021	0.020	0.018
$[11.0, 12.5]$	0.016	0.017	0.023	0.024	0.051	0.014	0.014	0.026
$[15.0, 17.0]$	0.018	0.026	0.015	0.029	0.011	0.018	0.015	0.016
$[17.0, 19.0]$	0.026	0.038	0.058	0.042	0.017	0.015	0.023	0.019
$[1.1, 6.0]$	0.025	0.009	0.027	0.021	0.024	0.014	0.024	0.014
$[15.0, 19.0]$	0.028	0.029	0.058	0.025	0.039	0.011	0.012	0.021

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.064	0.106	0.033	0.047	0.119	0.122	0.062	0.092
$[1.1, 2.5]$	0.034	0.317	0.151	0.180	0.113	0.104	0.082	0.050
$[2.5, 4.0]$	0.044	0.223	0.113	0.204	0.141	0.092	0.054	0.071
$[4.0, 6.0]$	0.031	0.291	0.063	0.150	0.092	0.086	0.047	0.051
$[6.0, 8.0]$	0.018	0.103	0.055	0.056	0.063	0.065	0.026	0.020
$[11.0, 12.5]$	0.030	0.086	0.044	0.045	0.074	0.055	0.025	0.039
$[15.0, 17.0]$	0.020	0.071	0.026	0.017	0.075	0.061	0.022	0.045
$[17.0, 19.0]$	0.037	0.093	0.044	0.017	0.132	0.127	0.038	0.073
$[1.1, 6.0]$	0.026	0.076	0.045	0.045	0.067	0.043	0.037	0.062
$[15.0, 19.0]$	0.022	0.066	0.035	0.033	0.105	0.048	0.031	0.027

without looking at the absolute value of individual results. The agreement between two sub-samples is tested by comparing the total χ^2 to a test statistic from pseudoexperiments and determining the p-value. The sub-categories are

- data taking periods: Run 1 *vs.* Run 2
- track categories: DD *vs.* LL
- magnet polarity: up *vs.* down
- evaluation methods: maximum-likelihood fit *vs.* method of moments [209].

In Appendix A.4, an introduction to the method of moments is given. The comparison of maximum-likelihood fit and the method of moments is the only cross-check performed on the complete data set. As a result, it also is the only cross-check performed in the nominal q^2 binning scheme, as defined in Table 4.5. The other three cross-checks are only performed in the two larger q^2 intervals, 1.1–6.0 and 15.0–19.0 GeV^2/c^4. The limited signal

statistics do not allow a further splitting of the already low numbers of candidates in the nominal q^2 intervals.
In practise, the total χ^2_{tot}, is defined as the sum over all χ^2 of all angular observables, X_i, in all q^2 intervals, b, via

$$\chi^2_{tot} = \sum_{i,b} \frac{(X_{i,b,A} - X_{i,b,B})^2}{\sigma^2_{i,b,A} + \sigma^2_{i,b,B}} \tag{6.2}$$

for two sub-samples or methods, A and B.
The value of χ^2_{tot} obtained in data is compared to the test statistics of χ^2_{tot} obtained from 1000 pseudodata sets, which are generated with the same configuration as the data samples. From this the p-value is calculated as the fraction of χ^2_{tot} from simulation with a larger χ^2_{tot} than data. A p-value greater than 0.05 is considered a proof of the good agreement between the two methods or in the two sub-samples.
In all tests for the S_i and $P_i^{(\prime)}$ angular observables and in the four comparison categories, all p-values are significantly above the previously set threshold of 0.05. With this final validation, the results are presented in the following Chapter 7.

7 Results of angular observables

After the selection of signal candidates in Section 4.3, the two sets of CP-averaged angular observables, as introduced in Section 2.3.2, are extracted from data by a four-dimensional maximum-likelihood fit as defined in Chapter 5. The statistical uncertainties of the measurements are determined using Feldman-Cousins scans in Section 5.8 and different sources of systematic uncertainties are evaluated in Chapter 6. In the following, all results are presented in tabular form in Section 7.1 and in graphical form in Section 7.2. The graphical presentation also includes comparisons of the measurement results with two theoretical predictions for the SM values. The correlation matrices between angular observables in all ten q^2 intervals – determined with the methods described in Section 5.5 – are listed in tables in Section 7.3.
Finally, the measurement results are evaluated in terms of the vector coupling strength of the underlying effective field theory in Section 7.4.

7.1 Tabular results

Tables 7.1 and 7.2 present the measurement results for all S_i and $P_i^{(\prime)}$ angular observables, respectively. The mean values, obtained from the four-dimensional maximum-likelihood fit, as introduced in Section 5.6, are accompanied by their asymmetric statistical uncertainties, determined by Feldman-Cousins scan (Section 5.8), and their systematic uncertainties from Chapter 6.

7.2 Graphical results

The results in the eight nominal q^2 intervals from Tables 7.1 and 7.2 are presented in graphical form in Figs. 7.1 and 7.2 as a trend of q^2. Theoretical predictions for the SM values of S_i and $P_i^{(\prime)}$ angular observables for most [1] q^2 intervals are underlaid.

[1] Not all q^2 regions are theoretically precisely predictable. The caveats, especially in the q^2 regions around the $c\bar{c}$ resonances are discussed in detail in Section 7.4.

Table 7.1: Results for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 in the eight nominal and two wide q^2 intervals. The first uncertainties are statistical and the second systematic.

q^2 [GeV^2/c^4]	F_{L}	S_3	S_4	S_5
$[0.10, 0.98]$	$0.34\,^{+0.10}_{-0.10} \pm 0.06$	$0.14\,^{+0.15}_{-0.14} \pm 0.02$	$-0.04\,^{+0.17}_{-0.16} \pm 0.04$	$0.24\,^{+0.12}_{-0.15} \pm 0.04$
$[1.1, 2.5]$	$0.54\,^{+0.21}_{-0.18} \pm 0.06$	$0.37\,^{+0.97}_{-0.41} \pm 0.03$	$0.29\,^{+0.34}_{-0.27} \pm 0.03$	$0.44\,^{+0.38}_{-0.32} \pm 0.05$
$[2.5, 4.0]$	$0.17\,^{+0.23}_{-0.32} \pm 0.06$	$-0.12\,^{+0.66}_{-0.39} \pm 0.02$	$-0.39\,^{+0.48}_{-0.45} \pm 0.04$	$-0.35\,^{+0.41}_{-0.31} \pm 0.02$
$[4.0, 6.0]$	$0.67\,^{+0.12}_{-0.14} \pm 0.02$	$-0.20\,^{+0.16}_{-0.19} \pm 0.01$	$-0.37\,^{+0.20}_{-0.13} \pm 0.05$	$-0.12\,^{+0.14}_{-0.19} \pm 0.03$
$[6.0, 8.0]$	$0.39\,^{+0.20}_{-0.21} \pm 0.01$	$-0.24\,^{+0.18}_{-0.17} \pm 0.02$	$-0.21\,^{+0.20}_{-0.18} \pm 0.02$	$-0.07\,^{+0.16}_{-0.20} \pm 0.02$
$[11.0, 12.5]$	$0.39\,^{+0.24}_{-0.17} \pm 0.02$	$-0.10\,^{+0.13}_{-0.13} \pm 0.02$	$-0.31\,^{+0.14}_{-0.17} \pm 0.02$	$-0.43\,^{+0.14}_{-0.16} \pm 0.02$
$[15.0, 17.0]$	$0.41\,^{+0.21}_{-0.14} \pm 0.02$	$-0.26\,^{+0.12}_{-0.11} \pm 0.03$	$-0.16\,^{+0.10}_{-0.11} \pm 0.02$	$-0.07\,^{+0.10}_{-0.10} \pm 0.03$
$[17.0, 19.0]$	$0.34\,^{+0.12}_{-0.11} \pm 0.03$	$-0.13\,^{+0.20}_{-0.17} \pm 0.04$	$-0.27\,^{+0.14}_{-0.15} \pm 0.06$	$-0.32\,^{+0.16}_{-0.34} \pm 0.04$
$[1.1, 6.0]$	$0.59\,^{+0.09}_{-0.09} \pm 0.03$	$-0.10\,^{+0.11}_{-0.11} \pm 0.01$	$-0.20\,^{+0.13}_{-0.14} \pm 0.03$	$-0.04\,^{+0.12}_{-0.12} \pm 0.02$
$[15.0, 19.0]$	$0.40\,^{+0.13}_{-0.11} \pm 0.03$	$-0.21\,^{+0.09}_{-0.09} \pm 0.03$	$-0.19\,^{+0.10}_{-0.13} \pm 0.06$	$-0.12\,^{+0.07}_{-0.07} \pm 0.02$

q^2 [GeV^2/c^4]	A_{FB}	S_7	S_8	S_9
$[0.10, 0.98]$	$-0.05\,^{+0.12}_{-0.12} \pm 0.03$	$-0.01\,^{+0.19}_{-0.17} \pm 0.01$	$0.21\,^{+0.22}_{-0.20} \pm 0.05$	$0.28\,^{+0.15}_{-0.12} \pm 0.06$
$[1.1, 2.5]$	$-0.21\,^{+0.19}_{-0.23} \pm 0.04$	$0.15\,^{+0.32}_{-0.72} \pm 0.02$	$0.06\,^{+0.40}_{-0.37} \pm 0.04$	$0.05\,^{+0.37}_{-0.30} \pm 0.02$
$[2.5, 4.0]$	$0.03\,^{+0.28}_{-0.26} \pm 0.01$	$-0.15\,^{+0.49}_{-0.69} \pm 0.03$	$0.04\,^{+0.75}_{-0.58} \pm 0.03$	$0.31\,^{+0.39}_{-0.36} \pm 0.02$
$[4.0, 6.0]$	$-0.08\,^{+0.09}_{-0.10} \pm 0.01$	$-0.04\,^{+0.18}_{-0.20} \pm 0.01$	$-0.07\,^{+0.21}_{-0.22} \pm 0.03$	$-0.18\,^{+0.22}_{-0.33} \pm 0.02$
$[6.0, 8.0]$	$-0.05\,^{+0.11}_{-0.12} \pm 0.01$	$-0.36\,^{+0.18}_{-0.15} \pm 0.02$	$-0.19\,^{+0.18}_{-0.16} \pm 0.02$	$-0.11\,^{+0.21}_{-0.20} \pm 0.02$
$[11.0, 12.5]$	$0.54\,^{+0.21}_{-0.18} \pm 0.05$	$-0.05\,^{+0.14}_{-0.14} \pm 0.01$	$0.06\,^{+0.14}_{-0.14} \pm 0.01$	$0.19\,^{+0.24}_{-0.19} \pm 0.03$
$[15.0, 17.0]$	$0.40\,^{+0.04}_{-0.09} \pm 0.01$	$-0.24\,^{+0.11}_{-0.11} \pm 0.02$	$-0.17\,^{+0.12}_{-0.11} \pm 0.02$	$0.14\,^{+0.12}_{-0.09} \pm 0.02$
$[17.0, 19.0]$	$0.14\,^{+0.12}_{-0.07} \pm 0.02$	$0.06\,^{+0.16}_{-0.16} \pm 0.01$	$0.17\,^{+0.18}_{-0.16} \pm 0.02$	$-0.08\,^{+0.15}_{-0.15} \pm 0.02$
$[1.1, 6.0]$	$-0.08\,^{+0.07}_{-0.08} \pm 0.02$	$-0.10\,^{+0.11}_{-0.13} \pm 0.01$	$0.02\,^{+0.13}_{-0.14} \pm 0.02$	$-0.05\,^{+0.11}_{-0.12} \pm 0.01$
$[15.0, 19.0]$	$0.31\,^{+0.06}_{-0.06} \pm 0.04$	$-0.14\,^{+0.08}_{-0.09} \pm 0.01$	$-0.06\,^{+0.09}_{-0.09} \pm 0.01$	$0.04\,^{+0.08}_{-0.06} \pm 0.02$

The predictions in blue in Figs. 7.1 and 7.2 are generated using the Flavio [101] software package (version 2.0.0) and are available for all angular observables. The software is based on the work of Ref. [11] and uses calculations of the hadronic form factors from Refs. [12, 13, 210]. Furthermore, predictions for the optimised angular observables, $P_i^{(\prime)}$, are provided by Refs. [211, 212] with form-factor calculations from Ref. [15]. The latter are added in orange shaded boxes to Fig. 7.2.

Most measured values of angular observables are in agreement with the theoretical SM predictions. For the q^2 regions around $3.0\,\mathrm{GeV}^2/c^4$ the number of signal candidates in the data set is lowest, as can be seen in the mass projections in Fig. 4.8. In this region the poor statistical precision of the measurements completely eliminates any sensitivity on the

Table 7.2: Results for the optimised observables F_{L} and P_1–P_8' in the eight nominal and two wide q^2 intervals. The first uncertainties are statistical and the second systematic.

q^2 [GeV^2/c^4]	F_{L}	P_1	P_2	P_3
$[0.10, 0.98]$	$0.34\,^{+0.10}_{-0.10} \pm 0.06$	$0.44\,^{+0.38}_{-0.40} \pm 0.11$	$-0.05\,^{+0.12}_{-0.12} \pm 0.03$	$-0.42\,^{+0.20}_{-0.21} \pm 0.05$
$[1.1, 2.5]$	$0.54\,^{+0.18}_{-0.19} \pm 0.03$	$1.60\,^{+4.92}_{-1.75} \pm 0.32$	$-0.28\,^{+0.24}_{-0.42} \pm 0.15$	$-0.09\,^{+0.70}_{-0.99} \pm 0.18$
$[2.5, 4.0]$	$0.17\,^{+0.24}_{-0.14} \pm 0.04$	$-0.29\,^{+1.43}_{-1.04} \pm 0.22$	$0.03\,^{+0.26}_{-0.25} \pm 0.11$	$-0.45\,^{+0.50}_{-0.62} \pm 0.20$
$[4.0, 6.0]$	$0.67\,^{+0.11}_{-0.14} \pm 0.03$	$-1.24\,^{+0.99}_{-1.17} \pm 0.29$	$-0.15\,^{+0.19}_{-0.20} \pm 0.06$	$0.52\,^{+0.82}_{-0.62} \pm 0.15$
$[6.0, 8.0]$	$0.39\,^{+0.20}_{-0.21} \pm 0.02$	$-0.78\,^{+0.61}_{-0.69} \pm 0.10$	$-0.06\,^{+0.12}_{-0.13} \pm 0.05$	$0.17\,^{+0.33}_{-0.31} \pm 0.06$
$[11.0, 12.5]$	$0.39\,^{+0.23}_{-0.16} \pm 0.03$	$-0.32\,^{+0.44}_{-0.52} \pm 0.09$	$0.62\,^{+0.55}_{-0.14} \pm 0.04$	$-0.32\,^{+0.29}_{-0.65} \pm 0.05$
$[15.0, 17.0]$	$0.41\,^{+0.18}_{-0.14} \pm 0.02$	$-0.88\,^{+0.41}_{-0.67} \pm 0.07$	$0.45\,^{+0.03}_{-0.07} \pm 0.03$	$-0.23\,^{+0.16}_{-0.20} \pm 0.02$
$[17.0, 19.0]$	$0.34\,^{+0.11}_{-0.12} \pm 0.04$	$-0.40\,^{+0.58}_{-0.57} \pm 0.09$	$0.14\,^{+0.10}_{-0.10} \pm 0.04$	$0.12\,^{+0.21}_{-0.21} \pm 0.02$
$[1.1, 6.0]$	$0.59\,^{+0.10}_{-0.10} \pm 0.03$	$-0.51\,^{+0.56}_{-0.54} \pm 0.08$	$-0.13\,^{+0.13}_{-0.13} \pm 0.05$	$0.12\,^{+0.27}_{-0.28} \pm 0.04$
$[15.0, 19.0]$	$0.40\,^{+0.13}_{-0.11} \pm 0.02$	$-0.70\,^{+0.35}_{-0.43} \pm 0.07$	$0.34\,^{+0.09}_{-0.07} \pm 0.04$	$-0.07\,^{+0.12}_{-0.13} \pm 0.03$

q^2 [GeV^2/c^4]	P_4'	P_5'	P_6'	P_8'
$[0.10, 0.98]$	$-0.09\,^{+0.36}_{-0.35} \pm 0.12$	$0.51\,^{+0.30}_{-0.28} \pm 0.12$	$-0.02\,^{+0.40}_{-0.34} \pm 0.06$	$0.45\,^{+0.50}_{-0.39} \pm 0.09$
$[1.1, 2.5]$	$0.58\,^{+0.62}_{-0.56} \pm 0.11$	$0.88\,^{+0.70}_{-0.71} \pm 0.10$	$0.25\,^{+1.22}_{-1.32} \pm 0.08$	$0.12\,^{+0.75}_{-0.76} \pm 0.05$
$[2.5, 4.0]$	$-0.81\,^{+1.09}_{-0.84} \pm 0.14$	$-0.87\,^{+1.00}_{-1.68} \pm 0.09$	$-0.37\,^{+1.59}_{-3.91} \pm 0.05$	$0.12\,^{+7.89}_{-4.95} \pm 0.07$
$[4.0, 6.0]$	$-0.79\,^{+0.47}_{-0.28} \pm 0.09$	$-0.25\,^{+0.32}_{-0.40} \pm 0.09$	$-0.09\,^{+0.40}_{-0.41} \pm 0.05$	$-0.15\,^{+0.44}_{-0.48} \pm 0.05$
$[6.0, 8.0]$	$-0.43\,^{+0.41}_{-0.45} \pm 0.06$	$-0.15\,^{+0.40}_{-0.41} \pm 0.06$	$-0.74\,^{+0.29}_{-0.40} \pm 0.03$	$-0.39\,^{+0.30}_{-0.39} \pm 0.02$
$[11.0, 12.5]$	$-0.63\,^{+0.30}_{-0.34} \pm 0.07$	$-0.88\,^{+0.28}_{-0.34} \pm 0.05$	$-0.11\,^{+0.28}_{-0.29} \pm 0.03$	$0.13\,^{+0.29}_{-0.30} \pm 0.04$
$[15.0, 17.0]$	$-0.32\,^{+0.23}_{-0.22} \pm 0.08$	$-0.14\,^{+0.21}_{-0.20} \pm 0.06$	$-0.48\,^{+0.21}_{-0.21} \pm 0.02$	$-0.34\,^{+0.23}_{-0.22} \pm 0.04$
$[17.0, 19.0]$	$-0.57\,^{+0.29}_{-0.36} \pm 0.13$	$-0.66\,^{+0.36}_{-0.80} \pm 0.13$	$0.12\,^{+0.33}_{-0.33} \pm 0.04$	$0.36\,^{+0.37}_{-0.33} \pm 0.07$
$[1.1, 6.0]$	$-0.41\,^{+0.28}_{-0.28} \pm 0.07$	$-0.07\,^{+0.25}_{-0.25} \pm 0.04$	$-0.21\,^{+0.23}_{-0.23} \pm 0.04$	$0.03\,^{+0.26}_{-0.28} \pm 0.06$
$[15.0, 19.0]$	$-0.39\,^{+0.18}_{-0.21} \pm 0.10$	$-0.24\,^{+0.16}_{-0.16} \pm 0.05$	$-0.28\,^{+0.19}_{-0.14} \pm 0.03$	$-0.11\,^{+0.19}_{-0.18} \pm 0.03$

angular observables. However, also in regions with significantly better statistical precision, theory predictions and measurement results mostly agree well.

The exception of these agreements are two (pairs of) angular observables, A_{FB} ($P_{2)}$ and S_5 (P_5), which show significant deviations between SM predictions in the q^2 region between 4.0–8.0 GeV^2/c^4, just below the J/ψ resonance, as shown in Figs. 7.1 and 7.2. A local deviation of P_2 in the interval $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$ is quantified to 3.1 standard deviations. The pattern of deviations in these observables is consistent with previous results from angular analyses of the B^0 meson decay at LHCb [26, 27].

In Section 7.4, the results are evaluated in a fit to the vector coupling strength of an underlying effective field theory.

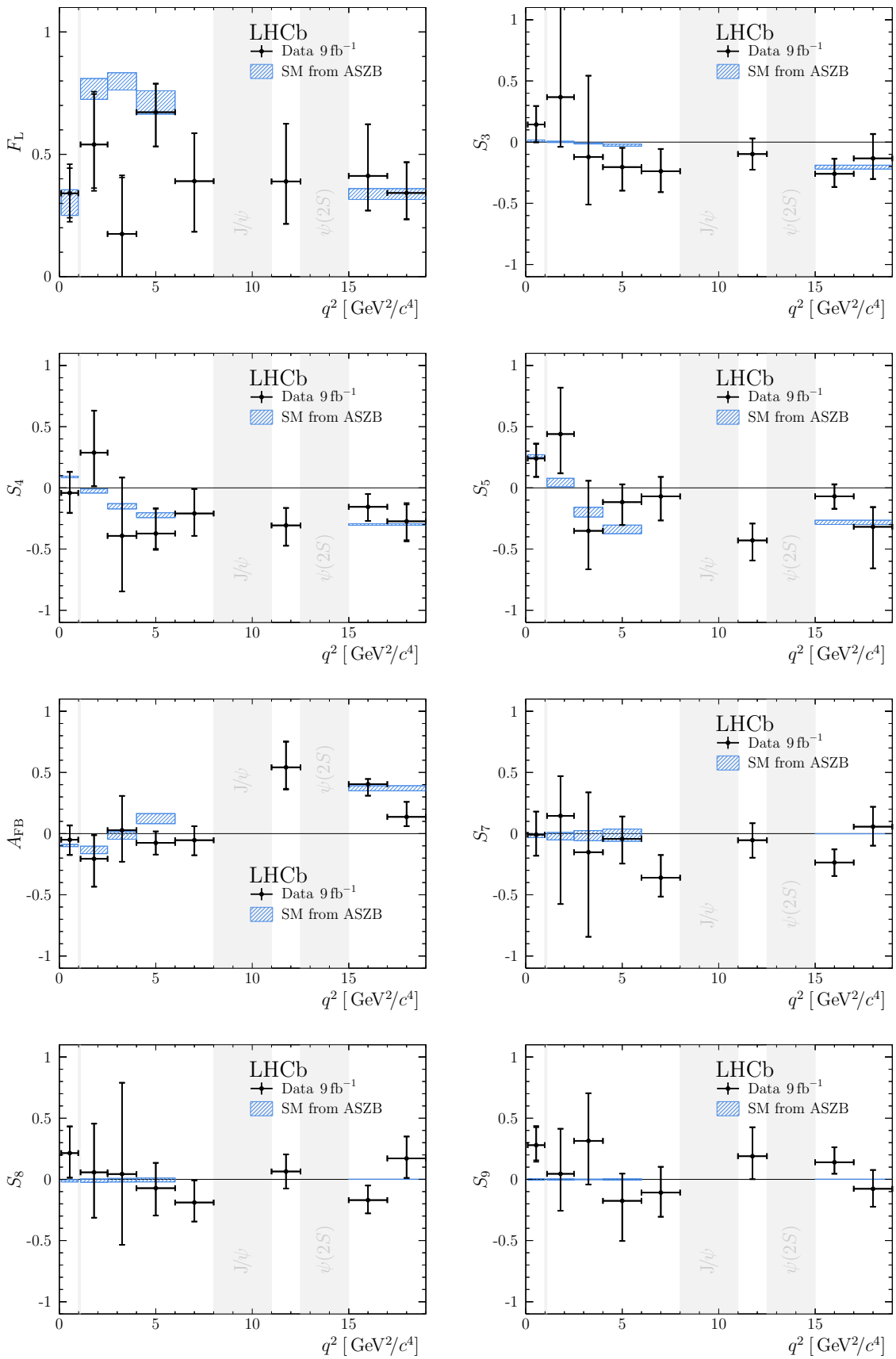

Figure 7.1: The *CP*-averaged angular observables F_L, A_{FB} and S_3–S_9 in the nominal q^2 intervals [1]. Theory predictions for a pure SM scenario are added from Refs. [11,12] with hadronic form-factors by Refs. [13,210]. The first uncertainty is purely statistical, the second bar represents the total uncertainty. Grey regions indicate the excluded resonant regions.

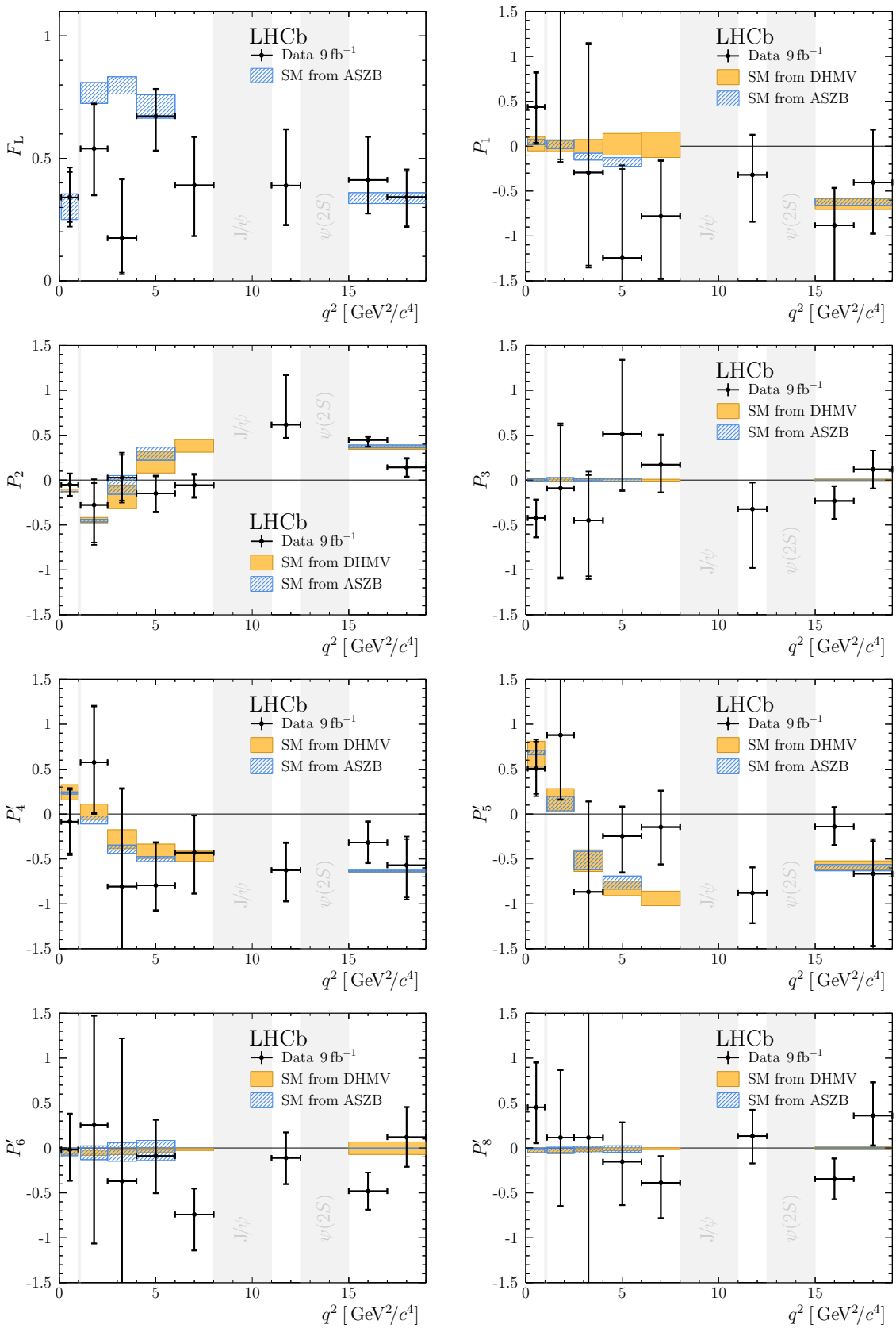

Figure 7.2: The optimised angular observables F_L and P_1–P'_8 in the nominal q^2 intervals [1]. Two theory predictions for a pure SM scenario from Refs. [11, 12] with hadronic form-factors by Refs. [13, 210] in blue and from Refs. [211, 212] with form-factors from Ref. [15] in orange are added. The first uncertainty of the data points is purely statistical, the second bar represents the total uncertainty. Grey regions indicate the excluded resonant regions.

7.3 Correlation matrices

The following pages contain the correlations between all angular observables, for the S_i basis in Tables 7.3 and 7.12 and for the $P_i^{(\prime)}$ basis in Tables 7.13 and 7.22. The values are determined by using a bootstrapping method, as described in Section 5.5.
The angular observables are mostly uncorrelated throughout all q^2 intervals, as the absolute values of most correlations is below 0.1. Only some exceptions – like the correlation between F_{L} and A_{FB} in the regions around the $c\bar{c}$ resonances – are observed. These findings are broadly consistent with measurements in the B^0 channel [26, 213] and are expected from the orthogonal definition of the angular moments in Section 2.3. The information from the correlation matrix are a crucial point in global fit to the measurement results, as presented in Section 7.4. Omitting the correlations between angular observables in a global evaluation significantly decreases the sensitivity on fit model parameters, like the Wilson coefficient $\mathcal{C}_9$.

Table 7.3: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $0.10 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.04	−0.01	0.03	0.04	0.12	−0.00	−0.11
S_3		1	−0.02	0.12	−0.02	0.02	0.06	0.02
S_4			1	−0.27	−0.09	−0.25	0.24	−0.06
S_5				1	0.10	0.22	−0.18	0.06
A_{FB}					1	0.19	−0.27	−0.06
S_7						1	−0.35	0.22
S_8							1	−0.08
S_9								1

Table 7.4: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.16	−0.05	0.11	0.11	0.04	−0.10	−0.03
S_3		1	0.06	0.09	−0.02	0.13	−0.01	−0.12
S_4			1	−0.02	0.17	0.05	0.33	0.09
S_5				1	0.20	0.22	−0.06	0.04
A_{FB}					1	0.20	0.11	0.12
S_7						1	0.06	0.16
S_8							1	0.22
S_9								1

Table 7.5: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.02	−0.01	0.06	−0.08	−0.02	−0.07	0.04
S_3		1	0.02	−0.06	−0.01	−0.03	0.07	0.02
S_4			1	0.00	−0.06	0.10	−0.05	−0.00
S_5				1	0.01	−0.07	0.00	−0.11
A_{FB}					1	0.05	0.06	−0.16
S_7						1	0.26	−0.14
S_8							1	−0.09
S_9								1

Table 7.6: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.20	−0.09	−0.09	0.07	0.01	0.16	−0.03
S_3		1	−0.08	−0.10	0.03	0.11	0.17	0.03
S_4			1	−0.08	−0.15	0.07	−0.04	0.05
S_5				1	−0.17	−0.02	0.09	−0.02
A_{FB}					1	−0.04	−0.03	−0.01
S_7						1	0.09	0.09
S_8							1	−0.08
S_9								1

Table 7.7: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.26	−0.01	0.07	0.01	−0.04	0.06	0.05
S_3		1	0.01	−0.03	−0.05	0.08	−0.04	−0.00
S_4			1	0.35	0.02	−0.05	−0.03	−0.10
S_5				1	0.02	−0.11	−0.07	−0.17
A_{FB}					1	−0.05	−0.19	−0.13
S_7						1	−0.10	−0.06
S_8							1	0.04
S_9								1

Table 7.8: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.09	0.03	0.09	−0.44	−0.09	−0.13	−0.08
S_3		1	−0.08	−0.13	−0.08	−0.04	−0.04	−0.19
S_4			1	0.08	0.06	−0.05	−0.09	0.12
S_5				1	−0.30	0.05	−0.04	−0.10
A_{FB}					1	0.10	0.11	0.15
S_7						1	0.05	−0.07
S_8							1	−0.07
S_9								1

Table 7.9: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.19	0.04	0.07	−0.28	−0.06	−0.13	−0.07
S_3		1	−0.09	−0.06	0.04	0.01	−0.06	0.01
S_4			1	0.27	0.07	0.10	0.06	0.14
S_5				1	−0.15	0.09	−0.06	−0.13
A_{FB}					1	0.07	−0.02	0.16
S_7						1	0.23	0.02
S_8							1	0.00
S_9								1

Table 7.10: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	−0.10	−0.02	−0.10	−0.10	0.07	−0.10	−0.01
S_3		1	−0.08	0.06	0.09	0.01	0.08	−0.02
S_4			1	−0.06	−0.07	0.00	0.06	0.04
S_5				1	−0.19	−0.03	0.09	0.02
A_{FB}					1	0.17	0.01	−0.07
S_7						1	−0.17	0.10
S_8							1	−0.19
S_9								1

Table 7.11: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.17	−0.00	−0.02	0.01	0.04	0.08	0.06
S_3		1	−0.01	−0.02	−0.02	0.04	−0.03	−0.05
S_4			1	−0.03	0.06	−0.02	0.19	−0.01
S_5				1	0.01	0.14	0.04	0.04
A_{FB}					1	−0.05	0.04	0.05
S_7						1	0.17	−0.02
S_8							1	−0.01
S_9								1

Table 7.12: Correlation matrix for the *CP*-averaged observables F_{L}, A_{FB} and S_3–S_9 from the maximum-likelihood fit in the interval $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
F_{L}	1	0.13	−0.05	−0.02	−0.17	−0.02	0.03	−0.05
S_3		1	−0.07	−0.00	−0.02	0.12	0.10	−0.05
S_4			1	0.05	−0.14	0.06	0.05	−0.02
S_5				1	0.05	−0.07	0.07	−0.05
A_{FB}					1	−0.10	−0.03	0.10
S_7						1	0.15	−0.01
S_8							1	−0.09
S_9								1

Table 7.13: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $0.10 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	−0.14	0.02	−0.18	−0.03	0.00	0.12	−0.01
P_1		1	−0.00	−0.01	−0.03	0.17	0.01	0.03
P_2			1	0.02	−0.08	0.09	0.19	−0.24
P_3				1	0.06	−0.03	−0.21	0.04
P_4'					1	−0.22	−0.23	0.15
P_5'						1	0.18	−0.18
P_6'							1	−0.25
P_8'								1

Table 7.14: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	0.03	0.02	−0.01	−0.06	−0.01	0.05	−0.08
P_1		1	−0.05	−0.01	0.06	−0.09	−0.03	0.04
P_2			1	−0.05	0.15	0.12	0.10	0.13
P_3				1	−0.08	−0.07	−0.02	−0.13
P_4'					1	0.03	−0.01	0.22
P_5'						1	0.09	−0.08
P_6'							1	−0.01
P_8'								1

Table 7.15: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	0.00	0.00	0.02	−0.02	−0.03	−0.06	−0.03
P_1		1	0.00	0.04	0.04	0.00	−0.04	−0.06
P_2			1	0.07	−0.01	0.04	0.04	−0.03
P_3				1	−0.03	0.02	0.06	−0.01
P_4'					1	0.07	0.06	0.08
P_5'						1	−0.02	−0.09
P_6'							1	0.21
P_8'								1

Table 7.16: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	0.16	−0.10	0.02	−0.02	−0.08	0.02	0.08
P_1		1	−0.03	0.02	−0.08	0.03	0.03	0.08
P_2			1	0.04	−0.12	−0.14	−0.03	−0.05
P_3				1	−0.02	−0.02	−0.05	0.09
P_4'					1	−0.11	−0.01	−0.10
P_5'						1	−0.04	0.07
P_6'							1	0.05
P_8'								1

Table 7.17: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	0.11	−0.10	0.01	−0.03	0.05	−0.05	−0.00
P_1		1	−0.04	0.01	0.01	−0.02	0.02	−0.06
P_2			1	0.12	−0.01	0.02	−0.05	−0.17
P_3				1	0.04	0.12	0.00	−0.04
P_4'					1	0.25	−0.03	−0.01
P_5'						1	−0.08	−0.06
P_6'							1	−0.05
P_8'								1

Table 7.18: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	−0.05	0.35	−0.09	0.00	0.04	−0.06	−0.14
P_1		1	−0.05	0.17	−0.09	−0.14	−0.03	−0.02
P_2			1	−0.15	0.12	−0.14	−0.00	0.07
P_3				1	−0.09	0.06	0.07	0.09
P_4'					1	0.04	−0.03	−0.10
P_5'						1	0.05	−0.01
P_6'							1	0.06
P_8'								1

Table 7.19: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	0.07	0.15	−0.09	0.08	0.09	0.00	−0.09
P_1		1	0.01	−0.05	0.00	−0.01	−0.00	−0.06
P_2			1	−0.23	0.10	−0.06	0.07	−0.03
P_3				1	−0.15	0.10	−0.03	0.02
P_4'					1	0.27	0.09	0.05
P_5'						1	0.09	−0.07
P_6'							1	0.21
P_8'								1

Table 7.20: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	−0.10	0.09	0.07	0.02	−0.10	0.06	−0.08
P_1		1	0.06	0.04	−0.10	0.02	−0.01	0.06
P_2			1	0.07	−0.07	−0.16	0.13	−0.00
P_3				1	−0.08	0.03	−0.08	0.17
P_4'					1	−0.08	−0.03	0.05
P_5'						1	0.00	0.08
P_6'							1	−0.12
P_8'								1

Table 7.21: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	0.11	−0.19	0.01	−0.01	−0.02	0.02	0.08
P_1		1	−0.05	0.07	0.01	−0.01	0.00	−0.04
P_2			1	−0.06	0.04	0.00	−0.05	0.01
P_3				1	0.01	−0.04	0.01	0.01
P_4'					1	−0.03	−0.02	0.18
P_5'						1	0.14	0.04
P_6'							1	0.17
P_8'								1

Table 7.22: Correlation matrix for the *CP*-averaged and optimised observables F_{L} and P_1–P_8' from the maximum-likelihood fit in the interval $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
F_{L}	1	−0.00	0.03	−0.01	0.00	0.01	0.03	0.05
P_1		1	0.01	0.04	−0.06	0.04	0.03	0.08
P_2			1	−0.07	−0.13	−0.00	−0.12	−0.03
P_3				1	0.03	0.04	0.02	0.08
P_4'					1	−0.00	0.12	0.04
P_5'						1	−0.09	0.07
P_6'							1	0.17
P_8'								1

7.4 EFT vector coupling strength

In addition to one-dimensional comparisons of the angular observable results to SM predictions in Figs. 7.1 and 7.2, the numerical results of the angular observables given in Tables 7.1 and 7.2 are evaluated in a global fit to the underlying couplings of the EFT using the software package Flavio [101] (version 2.0.0).

The theoretical models face caveats in the prediction of values for the angular observables, which already caused the absence of predictions in some q^2 regions in Figs. 7.1 and 7.2. These caveats are discussed in the following and in the context of a global evaluation.

Broad charmonium resonances and long-distant effects challenge the precision of theoretical predictions at low hadronic recoil, *i.e.* $q^2 \approx m_b^2$ [214–216], which corresponds to the q^2 regions around and above the J/ψ and $\psi(2S)$ resonance. As a result, the q^2 interval in between the $c\bar{c}$ resonances is not included in the global fit and the result for the wide $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$ interval is used instead of the two narrower $2\,\mathrm{GeV}^2/c^4$-intervals. Virtual charm-quark loops introduce larger theoretical uncertainties to the q^2 region at and below the J/ψ resonance [15]. The effects reach down into the q^2 region below $8.0\,\mathrm{GeV}^2/c^4$, hence the q^2 interval 6.0–$8.0\,\mathrm{GeV}^2/c^4$ is also omitted from the evaluations.

In summary, these limitations in precision on theoretical predictions result in only the four q^2 intervals up to $6.0\,\mathrm{GeV}^2/c^4$ and the broad q^2 interval between 15.0–$19.0\,\mathrm{GeV}^2/c^4$ to be included in the global fit.

In the fit, the real part of the vector coupling strength of an effective field theory, $\mathrm{Re}(\mathcal{C}_9)$, as introduced in Section 2.2, is varied. This is motivated by the discrepancies between measurements and SM predictions for this parameter in Refs. [26,27] and global evaluations performed by Refs. [11,43,50,54,55]. The default Flavio hadronic form-factor parameters are used; as well as subleading corrections to account for long-distance QCD interference effects with the charmonium decay modes [11, 12]. The resulting likelihood profile for $\Delta\mathrm{Re}(\mathcal{C}_9) = \mathrm{Re}(\mathcal{C}_9) - \mathrm{Re}(\mathcal{C}_9^{\mathrm{SM}})$ is given in Fig. 7.3. The best-fit value, $\Delta\mathrm{Re}(\mathcal{C}_9) = -1.9$ is favoured by 3.1σ with respect to a pure SM scenario.

The significance of this deviation from SM predictions shows a clear tension between the measurement results given in Section 7.1 and SM predictions. The result also underlines the observed deviations of some angular observables as shown in Figs. 7.1 and 7.2. The value of the significance of this deviation from SM predictions is however dependent on the exact configuration of the global fit, on the choice of nuisance parameters, which coupling coefficients are varied and which q^2 ranges are included.

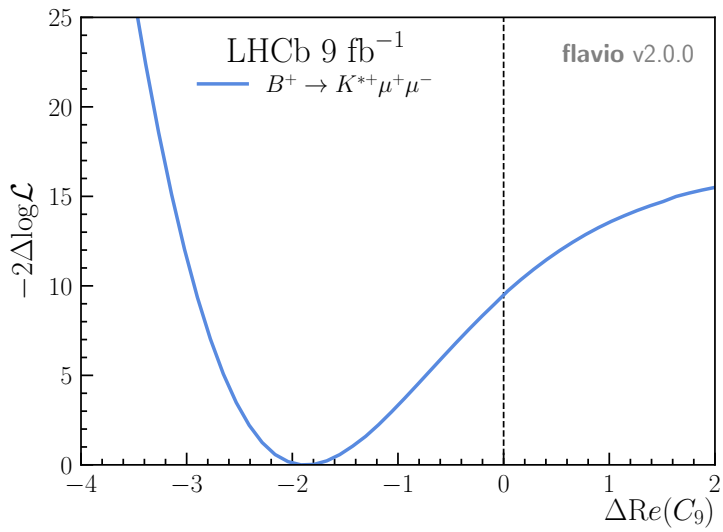

Figure 7.3: Likelihood profile scan of the real part of the vector coupling strength of an effective field theory description of $b \to s\ell^+\ell^-$ transitions, the Wilson coefficient $\mathcal{C}_9$. The profile is obtained from a global fit to a selection of results from Table 7.1 using the Flavio [101] software package. The exact input data and configuration of the fit are discussed in the text. A pure SM scenario is indicated by the vertical dashed line. The scenario with beyond SM contributions to the vector coupling is favoured by 3.1σ [3] over a pure SM scenario.

8 Conclusion

The angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay with the subsequent decay $K^{*+} \to K^0_S\pi^+$ shows the great capabilities of the LHCb experiment and its unique precision in studying rare flavour-changing neutral-current decays of B mesons involving $b \to s\ell^+\ell^-$ transitions. In total, 737 ± 34 signal candidates of the $B^+ \to K^{*+}\mu^+\mu^-$ decay have been selected from the full data set recorded during LHC Run 1 and Run 2 in the years 2011–2018, equivalent to an integrated luminosity of $9\,\text{fb}^{-1}$. The candidate selection uses the LHCb trigger system and central event filtering. Selection criteria of different variables are applied to remove specific sources of background, before a multi-variate classifier finalises the candidate selection by significantly suppressing the residual combinatorial background.

Two sets of eight angular observables in the S_i and $P_i^{(\prime)}$ basis are determined from the data using a four-dimensional maximum-likelihood fit in ten intervals of the dimuon mass squared, q^2. In order to utilise the identical q^2 interval scheme as used in previous analyses of the B^0 decay channel on a data set with significantly less signal candidates, the symmetries of the angular differential decay rate are exploited by dedicated angular folding techniques. These angular folds reduce the number of free parameters and the complexity in the fit, thereby stabilising the maximum-likelihood fits. Five angular folds are used in ten q^2 intervals for two sets of angular observables, resulting in a total of one hundred fits.
A further analysis technique is deployed to correct acceptance effects of the detector geometry and the candidate selection on the angular distributions. Despite the angular folding techniques, the full correlation matrices between all angular observables are determined using a bootstrapping method. Pseudoexperiments reveal that the statistical uncertainties as determined directly by the maximum-likelihood fits suffer from over- and under-coverage for some observables. Consequently, Feldman-Cousins profile scans are used to determine the correct confidence interval of the measured values. An extensive list of sources of systematic effects is studied and the corresponding systematic uncertainties are quantified

using pseudoexperiments. It is observed that no individual source of systematic effects dominates the values of the combined systematic uncertainty. The precision of the angular observable results are clearly dominated by the precision of the statistical uncertainty.

The angular observables results are compared with (up to) two independent predictions for pure SM scenarios. For most angular observables and in most q^2 intervals, the measured values are compatible with predictions from theoretical calculations. Partially, this is due to the large (statistical) uncertainties, but also due to good agreement of the central values.
However, deviations are observed, especially the angular observables S_5 (P_5) and A_{FB} (P_2) in the q^2 region below the J/ψ resonance. The observed pattern of deviations in the q^2 trend of the angular observables is in line with the observations previously made in the angular analyses of the decays of the isospin-partner B^0 [26, 27]. And it is those deviations that drive the global evaluation of the results in terms of the real part of the coupling strength of the underlying effective field theory vector operator, the Wilson coefficient $\mathcal{C}_9$, away from the SM predicted value. A fit using the Flavio [101] software package, which varies only the Wilson coefficient $\mathcal{C}_9$, favours a deviation from a pure SM scenario by about 3.1σ at a difference in the coupling strength, $\Delta\mathrm{Re}(\mathcal{C}_9) = -1.9$. The significance of this discrepancy greatly depends on included q^2 regions, the choice of nuisance parameters and which coupling strengths are varied in the fit.
A definite conclusion on what causes these discrepancies is beyond the scope of this analysis and relies on interpretation of the results by elementary particle theoreticians. Not only new physical phenomena in the form of leptoquarks [36–43], a Z' boson [44–50] or Higgs-doublet models [51] can explain the deviations [52–56], but also effects from uncertainties on the hadronic form factors calculations and long-distance contributions in the internal quark loops [57–60].

With a publication in the peer-reviewed journal *Physical Review Letters* pending, the analysis is already presented to the public in a conference presentation [2]. The measurement results are well received by the *flavour physics* community, which is expressed in rapidly published theory interpretations of the results [217, 218].

The potential of the data collected by LHCb during Run 1 and Run 2 is not yet fully exploited by the presented analysis, as works on an angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay with subsequent decays $K^{*+} \to K^+\pi^0$ and $\pi^0 \to \gamma\gamma$ are ongoing. The challenging reconstruction of the π^0 meson from a $\gamma\gamma$ pair in the LHCb experiment results in a lower signal yield, poorer mass and angular resolutions and more complex background components. The analysis will nevertheless add information to the evaluations

of the $B^+ \to K^{*+}\mu^+\mu^-$ decay in terms of the underlying vector coupling strength.
For the near future, the $B^+ \to K^{*+}\mu^+\mu^-$ decay channel will greatly profit from the significant increase in luminosity in the LHCb phase 1 upgrade [219]. As Tables 7.1 and 7.2 show, the uncertainties on the presented results are clearly dominated by their statistical precision. With a projection of collecting data equivalent to an integrated luminosity of $50\,\mathrm{fb}^{-1}$ at the LHCb experiment, these statistical uncertainties are expected to be more than halved by the year 2030. This can help to narrow down the source of deviations observed in decays with $b \to s\ell^+\ell^-$ transitions, along with measurements in the B^0 decay and other b hadrons, lepton flavour universality test and searches for lepton flavour violating decays.
The main competitor to LHCb in the field of *flavour physics*, the recently upgraded Belle II experiment, is expected to contribute similar exciting results in $B \to K^*\mu^+\mu^-$ decays and update their analyses of the neutral B^0 meson decay [30] and extend the program to the charge B^+ meson decay. Also, the CMS collaboration has published results on $B \to K^*\mu^+\mu^-$ decays [31,34] and is expected to improve the statistical precision on their results in the next years.

Whether it turns out to be the internal structure of the quark-loops or new physical phenomena beyond the SM, the $B^+ \to K^{*+}\mu^+\mu^-$ decay – along with other decays containing $b \to s\ell^+\ell^-$ transitions – is and will be one of the most exciting and promising fields of research at the precision and high energy frontiers of particle physics; now and in the years to come.

A Supplemental material

The following sections contain additional material in form of tables and plots to the analysis of $B^+ \to K^{*+}\mu^+\mu^-$, which is not strictly needed in the main part of the document. Due to the split into individual years of data taking and the DD and LL category during the selection and repetitive applications of the angular fit, figures are usually created in six-, twelve- or even thirtyfold. The similarity between different data taking years makes this repetition meaningless and therefore mostly only one or two examples of any distribution are given in this document.
In Appendix A.1, additional plots from the weighting of simulation samples are given. The correlations between the two weighting dimensions and the distributions of weights in these two variables are given. For the validation of the agreement, variable distributions are compared between simulation and data after the MC weighting is applied. The generation of pseudodata for pseudoexperiments is explained in Appendix A.2 and the results from pull evaluations of these studies is given in Appendix A.3. An alternative method to determine the angular observables, which is used as a validation of the maximum-likelihood fit is the method of moments, which is explained in Appendix A.4. Also, fit projections of the four-dimensional fit are presented in the angular dimensions for the eight nominal q^2 intervals and for the five folds in Appendix A.5. At the end of this appendix, the likelihood profiles of the Feldman-Cousins [206] scans as discussed in Section 5.8 are presented in Appendix A.6.

A.1 Simulation-data agreement

In Section 4.1.2, the weighting of simulation samples to match the distributions in data is explained. As the two variables, $p_{\mathrm{T}}(B^+)$ and number of long tracks, are uncorrelated, as shown in Fig. A.1 for 2012 simulation, the weighting is done subsequently in these two dimensions.
The ratios of background-subtracted data and simulation for the years 2012 and 2017 in the LL category are shown in Fig. A.2. These weights are equivalent to the values of the

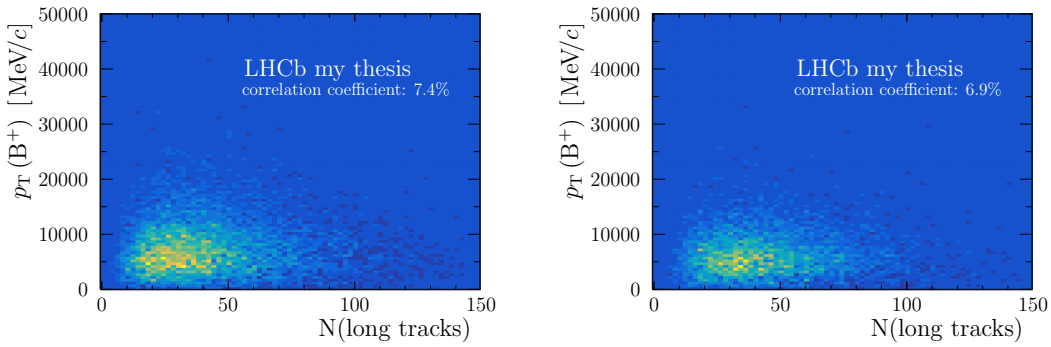

Figure A.1: The transverse momentum of the B meson candidate versus the number of long tracks in the event in the 2012 samples of the (left) DD and (right) LL category. The correlation is shown to be well below 10%, which is true for the samples of every year. Distributions in other years look very similar. No significant correlation allows for a more simple consecutive weighting of simulation from two one-dimensional distributions. More details on the weighting are given in the text.

ratios.

After the weighting and prior to the training of the multi-variate classifier in Section 4.3.5, all input variables to the training are compared between weighted simulation and background-subtracted data. Figs. A.3 and A.4 show the distribution for the 2011 sample in the DD and LL category, respectively. While only 2011 data is shown, the distributions are representative for all years 2011–2018. Variations between the years are minor and do not impact the the conclusion drawn from the comparisons in Figs. A.3 and A.4. The distributions are in good agreement and are suitable for the multi-variate classification. The selection with a classifier is a pure optimisation step in this analysis, as the absolute efficiency of the selection is of no concern to the latter four-dimensional maximum-likelihood fit. The discrepancy in the transverse momentum of the K_S^0 candidate in DD category, as seen in Fig. 6.2 is evaluated in a systematic study in Section 6.6.1.

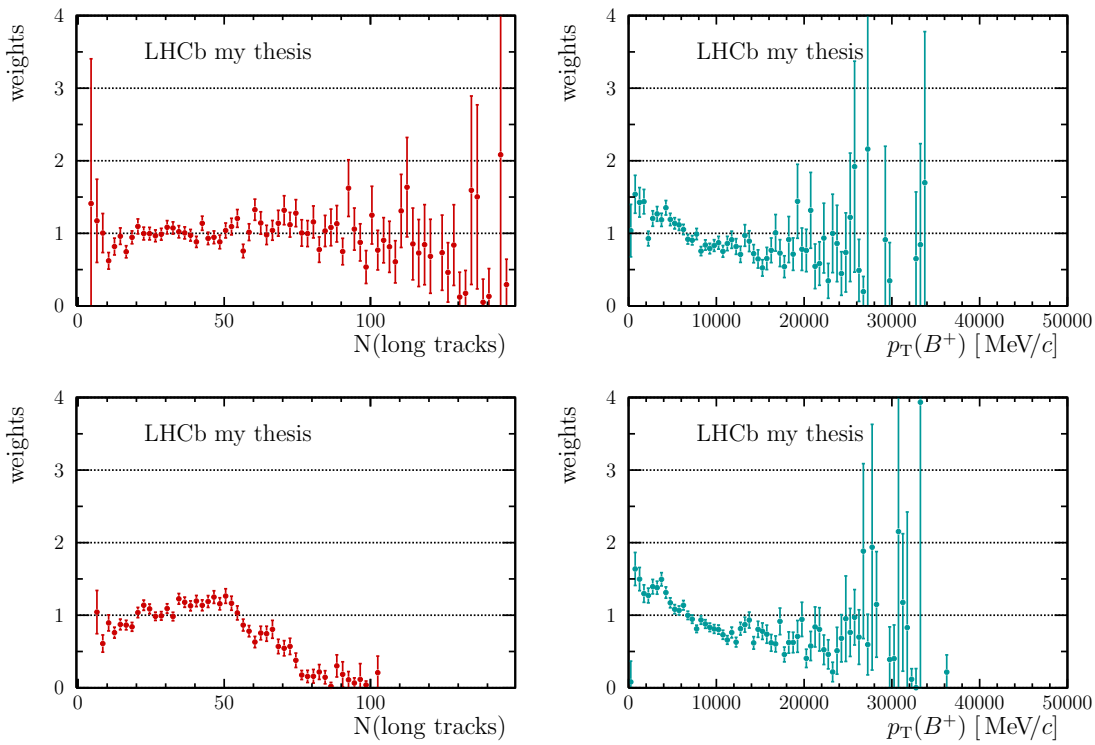

Figure A.2: The simulation weights to improve simulation-data agreement for the (top) 2012 and (bottom) 2017 samples in the LL category with (left) number of long tracks and (right) p_T (B^+) ratios. The weights are obtained from the ratio of the distributions in background-subtracted data of $B^+ \to J/\psi K^{*+}$ decays and in simulation. The large uncertainties on some ratios emerge due to low occupancy in the boundary regions of the distributions. It is obvious to see the different long track weight distributions for (top left) 2012 and (bottom left) 2017. This is a direct result of the track multiplicity cuts in the trigger for Run 1 and Run 2, as discussed in a side note in Section 4.3.2. The presented weight distributions for 2012 (2017) are representative for Run 1 (Run 2); no significant differences between DD and LL samples are observed.

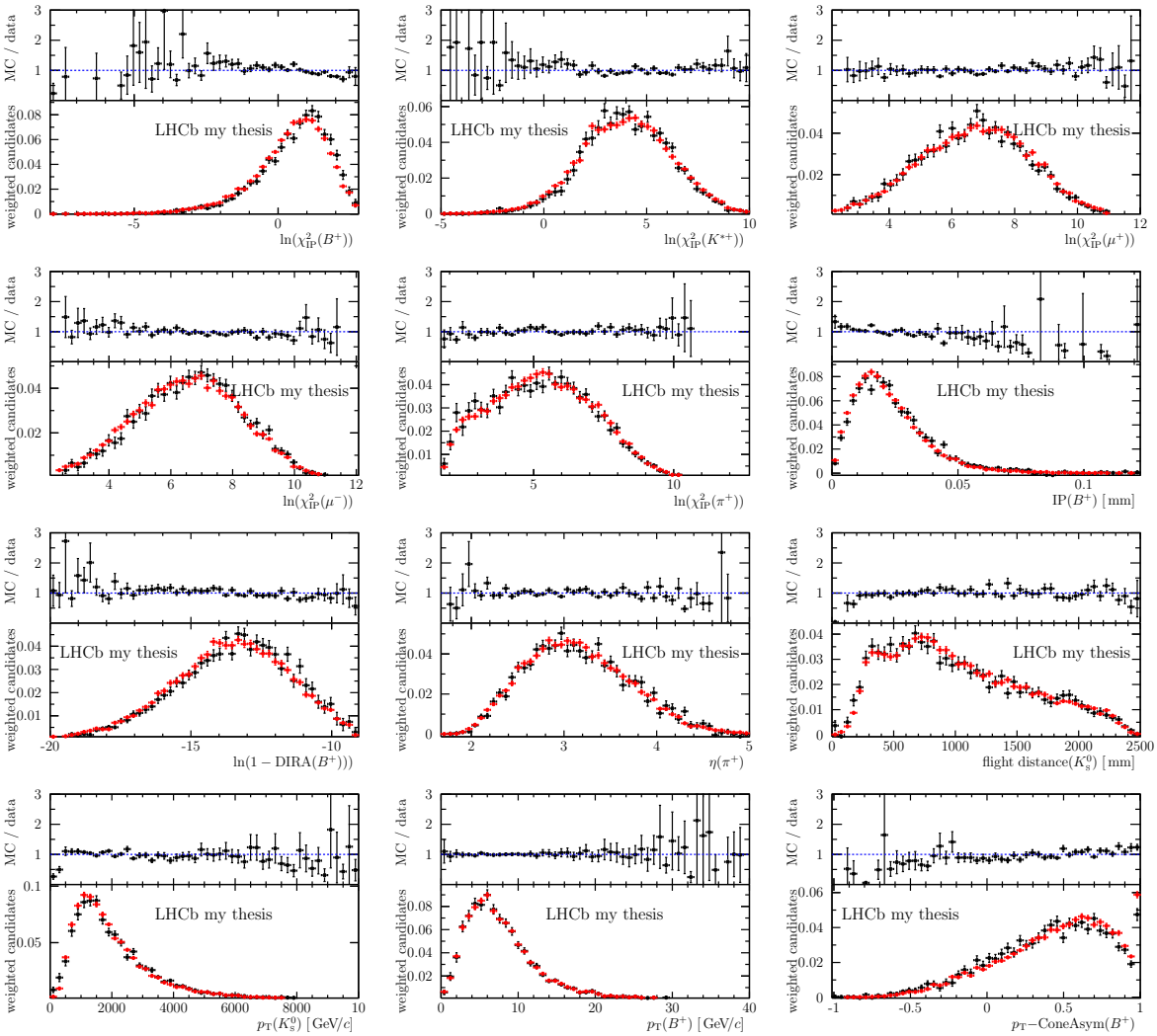

Figure A.3: Comparisons of the distribution of all variables used for the training of the multi-variate classifier in Section 4.3.5 of the DD category given for data including the $c\bar{c}$ resonant regions and simulation of the signal decay from the year 2011. The distribution of weighted simulation is given in red with the background-subtracted data in black. The ratio of both distribution is also given. All distributions are in fair agreement and allow the usage of simulation as the signal proxy sample for the training of the multi-variate classifier. This distributions of 2011 stand exemplarily for all years.

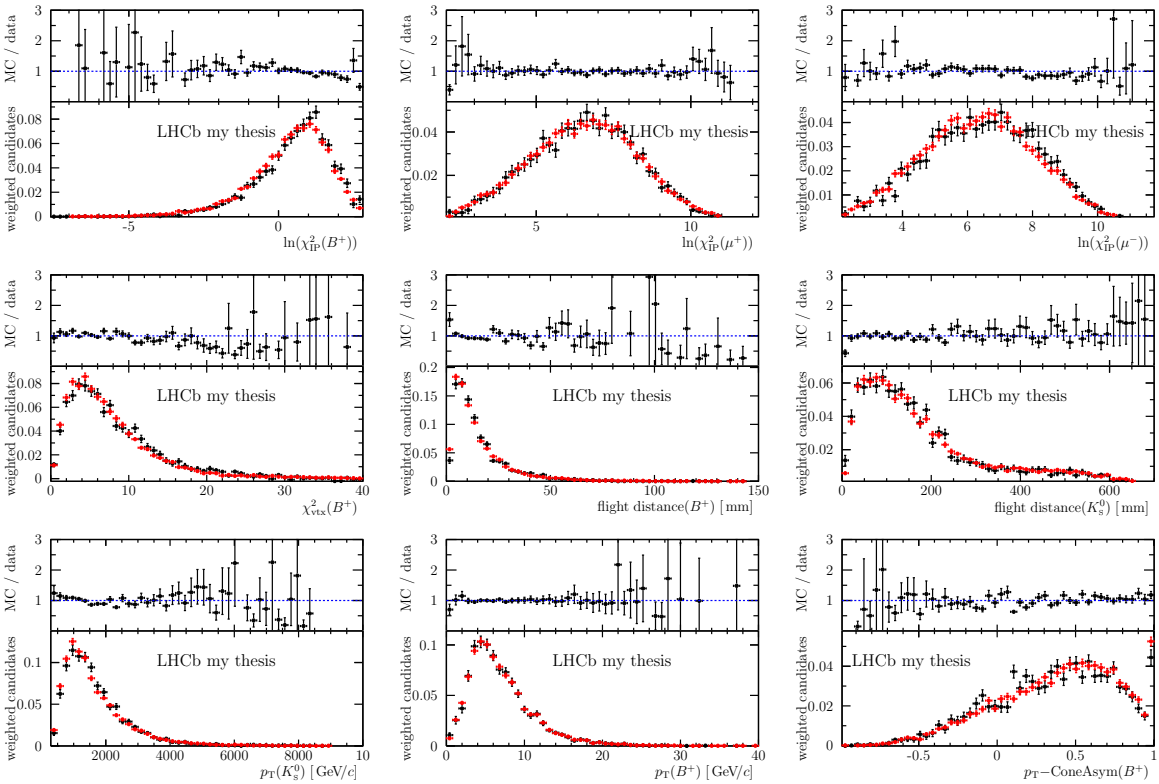

Figure A.4: Comparisons of the distribution of all variables used for the training of the multi-variate classifier in Section 4.3.5 of the LL category given for data including the $c\bar{c}$ resonant regions and simulation of the signal decay from the year 2011. The distribution of weighted simulation is given in red with the background-subtracted data in black. The ratio of both distribution is also given. All distributions are in fair agreement and allow the usage of simulation as the signal proxy sample for the training of the multi-variate classifier. This distributions of 2011 stand exemplarily for all years.

A.2 Generation of pseudodata

Pseudodata comprises of simulated candidates, which only contain a minimal number of variables and which are generated using the fit program itself. The pseudocandidates are generated from random numbers and the probability density functions of the fit models to create values for all needed variables. This process is done in loops and over several sequential steps.

At first, a random number within the range from zero to one defines if a candidate is signal or background by whether the number is smaller or larger than the signal fraction, f_{sig}, respectively. This will decide whether the probability density functions of the signal and background model are considered for the generation of the variables.

The invariant mass of the B meson, $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$, is generated by randomly drawing a number within the defined range, 5150–6000 MeV/c^2. Subsequently, a second random number is drawn and compared to the value of the probability density function (of either background or signal model) at the formerly chosen value of $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$. For this, the second random number is normalised to the maximum value of the probability density function inside the (invariant mass) range. If the normalised random number is smaller than the value of the probability density function, the value for $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$ is accepted and assigned to the pseudocandidate. Else, the process of picking a random value for $m(K^0_{\text{S}}\pi^+\mu^+\mu^-)$ and comparing a second random value to the p.d.f. is repeated until a normalised random number is smaller.

The probability density functions for $m(K^0_{\text{S}}\pi^+)$ and the three angular dimensions are connected via the composition of P- and S-wave contributions. Also, the angular dimensions are connected to q^2, via the four-dimensional angular acceptance correction, which has to be considered for the generation of pseudodata. Conclusively, these five values are generated in combination.

Five random numbers select values for each of the five dimensions from their corresponding range and the process as described for the B meson mass dimension is repeated. The values of the one- and three-dimensional p.d.f in $m(K^0_{\text{S}}\pi^+)$, and $\cos\theta_K$, $\cos\theta_\ell$ and ϕ are determined and multiplied with one another. Then the efficiency of the angular acceptance corrections for the values of q^2, $\cos\theta_K$, $\cos\theta_\ell$ and ϕ is determined according to the parametrisation in Eq. (5.20) and multiplied with the product of the two probabilities. Again, an additional random number, which is normalised to the maximum probability is compared to the product of the determined efficiency and probabilities. If accepted, *i.e.* the random number is smaller, the generation of a pseudocandidate is completed, else the generation of values for $m(K^0_{\text{S}}\pi^+)$, q^2, $\cos\theta_K$, $\cos\theta_\ell$ and ϕ is repeated.

The generation loop is repeated until the pseudodata sample reaches the desired size.

A.3 Pseudoexperiment results

The results from studies of the pull distributions of pseudoexperiments with three scenarios of parameter values are given on the following pages. In Tables A.1, A.3 and A.5, the mean positions of the pseudoexperiment pull distributions are given. Tables A.2, A.4 and A.6 lists the widths of the pull distributions. All values from these six tables are obtained using a maximum-likelihood fit of a Gaussian bell shape to the pull distributions. The consequences and interpretation of these results is discussed in Section 5.7.1.

Table A.1: Results of the mean positions of the pull distributions from pseudoexperiments with statistics equal to the statistics in data and with generated angular observable values according to SM predictions obtained by the FLAVIO software package. Deviations from the nominal value, zero, are observed for some observables. A systematic uncertainty is assigned for this bias, as discussed in Section 6.7. More discussion on these results is given in Section 5.7.1.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.25 ± 0.03	-0.01 ± 0.03	-0.03 ± 0.03	-0.03 ± 0.03	-0.00 ± 0.03	-0.03 ± 0.03	-0.01 ± 0.04	-0.02 ± 0.03
$[1.1, 2.5]$	0.05 ± 0.03	-0.02 ± 0.03	-0.06 ± 0.03	-0.03 ± 0.03	0.02 ± 0.03	-0.06 ± 0.03	0.03 ± 0.03	0.00 ± 0.03
$[2.5, 4.0]$	-0.07 ± 0.03	-0.01 ± 0.03	-0.05 ± 0.03	-0.01 ± 0.03	0.04 ± 0.03	0.06 ± 0.03	-0.04 ± 0.03	0.01 ± 0.03
$[4.0, 6.0]$	-0.08 ± 0.03	0.06 ± 0.03	-0.10 ± 0.03	0.14 ± 0.03	-0.05 ± 0.03	0.01 ± 0.03	0.04 ± 0.03	0.01 ± 0.03
$[6.0, 8.0]$	-0.07 ± 0.03	0.06 ± 0.03	-0.01 ± 0.03	0.12 ± 0.03	-0.04 ± 0.03	-0.05 ± 0.03	-0.04 ± 0.03	-0.04 ± 0.03
$[11.0, 12.5]$	-0.13 ± 0.03	0.08 ± 0.03	0.05 ± 0.03	0.17 ± 0.03	-0.11 ± 0.03	-0.03 ± 0.03	-0.02 ± 0.03	-0.08 ± 0.03
$[15.0, 17.0]$	-0.00 ± 0.03	0.06 ± 0.03	0.03 ± 0.03	0.10 ± 0.03	-0.10 ± 0.03	-0.06 ± 0.03	0.06 ± 0.03	0.02 ± 0.03
$[17.0, 19.0]$	-0.01 ± 0.03	0.10 ± 0.03	-0.02 ± 0.03	-0.05 ± 0.04	-0.08 ± 0.03	-0.03 ± 0.03	0.01 ± 0.03	0.03 ± 0.03
$[1.1, 6.0]$	-0.09 ± 0.03	-0.01 ± 0.03	-0.08 ± 0.03	0.11 ± 0.03	0.04 ± 0.03	-0.03 ± 0.03	0.06 ± 0.03	0.02 ± 0.03
$[15.0, 19.0]$	-0.01 ± 0.03	0.09 ± 0.03	-0.08 ± 0.03	0.24 ± 0.03	0.10 ± 0.03	-0.08 ± 0.03	-0.04 ± 0.03	-0.02 ± 0.03

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.16 ± 0.03	-0.06 ± 0.03	-0.05 ± 0.03	0.05 ± 0.03	-0.02 ± 0.03	0.04 ± 0.03	-0.03 ± 0.04	-0.02 ± 0.03
$[1.1, 2.5]$	-0.07 ± 0.03	-0.02 ± 0.03	0.50 ± 0.03	0.08 ± 0.03	-0.05 ± 0.03	-0.06 ± 0.03	0.03 ± 0.03	-0.01 ± 0.03
$[2.5, 4.0]$	0.10 ± 0.03	0.11 ± 0.03	0.33 ± 0.03	0.01 ± 0.03	0.07 ± 0.03	0.14 ± 0.03	0.01 ± 0.03	0.03 ± 0.03
$[4.0, 6.0]$	-0.18 ± 0.03	0.04 ± 0.03	-0.14 ± 0.03	0.02 ± 0.03	0.04 ± 0.03	0.16 ± 0.03	-0.00 ± 0.03	0.02 ± 0.03
$[6.0, 8.0]$	-0.15 ± 0.03	0.05 ± 0.03	-0.20 ± 0.03	0.00 ± 0.03	-0.00 ± 0.03	0.02 ± 0.03	0.02 ± 0.03	0.02 ± 0.03
$[11.0, 12.5]$	-0.16 ± 0.03	0.08 ± 0.03	-0.22 ± 0.03	0.07 ± 0.03	0.01 ± 0.03	0.09 ± 0.03	0.00 ± 0.03	0.02 ± 0.03
$[15.0, 17.0]$	-0.16 ± 0.03	0.05 ± 0.03	-0.08 ± 0.03	0.01 ± 0.03	-0.02 ± 0.03	-0.01 ± 0.03	0.05 ± 0.03	-0.02 ± 0.03
$[17.0, 19.0]$	-0.11 ± 0.03	0.05 ± 0.03	0.03 ± 0.03	0.03 ± 0.04	-0.06 ± 0.03	-0.03 ± 0.03	-0.00 ± 0.03	0.02 ± 0.03
$[1.1, 6.0]$	-0.14 ± 0.03	0.06 ± 0.03	-0.04 ± 0.03	0.01 ± 0.03	-0.16 ± 0.03	0.03 ± 0.03	0.02 ± 0.03	0.01 ± 0.03
$[15.0, 19.0]$	-0.07 ± 0.03	0.11 ± 0.03	0.09 ± 0.03	0.00 ± 0.03	-0.08 ± 0.03	0.15 ± 0.03	-0.08 ± 0.03	-0.03 ± 0.03

Table A.2: Results of the Gaussian width of the pull distributions from pseudoexperiments with statistics equal to the statistics in data and with generated angular observable values according to SM predictions obtained by the FLAVIO software package. Deviations from the nominal value, zero, are observed for some observables. A systematic uncertainty is assigned for this bias, as discussed in Section 6.7. More discussion on these results is given in Section 5.7.1.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	1.08 ± 0.02	1.03 ± 0.02	1.01 ± 0.02	1.06 ± 0.02	0.98 ± 0.02	1.01 ± 0.02	1.11 ± 0.02	0.95 ± 0.02
$[1.1, 2.5]$	1.05 ± 0.02	1.03 ± 0.02	1.05 ± 0.02	0.98 ± 0.02	1.01 ± 0.02	1.00 ± 0.02	1.01 ± 0.02	1.00 ± 0.02
$[2.5, 4.0]$	0.99 ± 0.02	0.99 ± 0.02	1.02 ± 0.02	0.96 ± 0.02	0.98 ± 0.02	0.99 ± 0.02	1.01 ± 0.02	1.00 ± 0.02
$[4.0, 6.0]$	0.94 ± 0.02	1.01 ± 0.02	1.05 ± 0.02	1.00 ± 0.02	1.02 ± 0.02	1.02 ± 0.02	1.06 ± 0.02	1.04 ± 0.02
$[6.0, 8.0]$	0.95 ± 0.02	1.04 ± 0.02	1.04 ± 0.02	1.00 ± 0.02	1.07 ± 0.02	0.96 ± 0.02	1.08 ± 0.02	1.00 ± 0.02
$[11.0, 12.5]$	1.03 ± 0.02	1.00 ± 0.02	1.02 ± 0.02	0.97 ± 0.02	1.06 ± 0.02	1.02 ± 0.02	1.01 ± 0.02	1.04 ± 0.02
$[15.0, 17.0]$	1.10 ± 0.02	1.04 ± 0.02	1.07 ± 0.02	1.00 ± 0.02	1.07 ± 0.02	1.04 ± 0.02	1.00 ± 0.02	1.10 ± 0.02
$[17.0, 19.0]$	1.08 ± 0.02	1.02 ± 0.02	1.06 ± 0.02	1.14 ± 0.03	1.07 ± 0.02	1.09 ± 0.02	1.09 ± 0.02	1.10 ± 0.02
$[1.1, 6.0]$	0.91 ± 0.02	1.01 ± 0.02	1.00 ± 0.02	0.99 ± 0.02	0.98 ± 0.02	1.00 ± 0.02	1.03 ± 0.02	0.97 ± 0.02
$[15.0, 19.0]$	1.04 ± 0.02	1.01 ± 0.02	1.03 ± 0.02	1.00 ± 0.02	1.02 ± 0.02	1.02 ± 0.02	0.99 ± 0.02	0.97 ± 0.02

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.96 ± 0.02	0.92 ± 0.02	1.01 ± 0.02	0.96 ± 0.02	1.06 ± 0.02	0.98 ± 0.02	1.03 ± 0.02	1.01 ± 0.02
$[1.1, 2.5]$	1.38 ± 0.02	0.81 ± 0.02	0.87 ± 0.02	0.75 ± 0.02	0.93 ± 0.02	0.93 ± 0.02	0.86 ± 0.02	0.94 ± 0.02
$[2.5, 4.0]$	1.52 ± 0.02	0.73 ± 0.02	0.77 ± 0.02	0.75 ± 0.02	0.89 ± 0.02	0.89 ± 0.02	0.85 ± 0.02	0.91 ± 0.02
$[4.0, 6.0]$	0.87 ± 0.02	0.80 ± 0.02	0.83 ± 0.02	0.74 ± 0.02	0.99 ± 0.02	0.94 ± 0.02	0.89 ± 0.02	0.92 ± 0.02
$[6.0, 8.0]$	0.88 ± 0.02	0.87 ± 0.02	0.89 ± 0.02	0.87 ± 0.02	1.01 ± 0.02	0.89 ± 0.02	0.99 ± 0.02	1.00 ± 0.02
$[11.0, 12.5]$	1.11 ± 0.02	0.93 ± 0.02	0.90 ± 0.02	0.94 ± 0.02	0.95 ± 0.02	0.98 ± 0.02	0.98 ± 0.02	0.98 ± 0.02
$[15.0, 17.0]$	1.16 ± 0.02	0.96 ± 0.02	0.98 ± 0.02	1.02 ± 0.02	1.04 ± 0.02	1.02 ± 0.02	1.00 ± 0.02	1.05 ± 0.02
$[17.0, 19.0]$	1.17 ± 0.02	0.99 ± 0.02	1.03 ± 0.02	1.06 ± 0.03	1.03 ± 0.02	1.08 ± 0.02	0.99 ± 0.02	1.04 ± 0.02
$[1.1, 6.0]$	0.87 ± 0.02	0.90 ± 0.02	0.91 ± 0.02	0.83 ± 0.02	0.97 ± 0.02	0.95 ± 0.02	0.94 ± 0.02	0.97 ± 0.02
$[15.0, 19.0]$	1.12 ± 0.02	1.00 ± 0.02	0.97 ± 0.02	0.99 ± 0.02	1.07 ± 0.02	1.02 ± 0.02	0.98 ± 0.02	1.05 ± 0.02

Table A.3: Results of the mean position of the pull distributions from pseudoexperiments with statistics equal to the statistics in data and with generated angular observable values according to predictions of a new physics contribution of $\mathrm{Re}(\mathcal{C}_9^{\mathrm{NP}}) = -1$, as determined by the Flavio software package. Deviations from the nominal value, zero, are observed for some observables. A systematic uncertainty is assigned for this bias, as discussed in Section 6.8. More discussion on these results is given in Section 5.7.1.

q^2 [GeV^2/c^4]	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
$[0.10, 0.98]$	0.20 ± 0.03	-0.07 ± 0.03	-0.07 ± 0.03	0.01 ± 0.03	-0.00 ± 0.03	-0.02 ± 0.03	0.05 ± 0.03	0.00 ± 0.03
$[1.1, 2.5]$	0.09 ± 0.03	0.01 ± 0.03	-0.03 ± 0.03	-0.01 ± 0.03	0.09 ± 0.03	-0.03 ± 0.03	0.07 ± 0.03	0.00 ± 0.03
$[2.5, 4.0]$	-0.11 ± 0.03	0.00 ± 0.03	0.00 ± 0.03	-0.01 ± 0.03	0.01 ± 0.03	-0.01 ± 0.03	0.02 ± 0.03	0.02 ± 0.03
$[4.0, 6.0]$	-0.09 ± 0.03	-0.01 ± 0.03	-0.04 ± 0.03	0.04 ± 0.03	0.01 ± 0.03	0.07 ± 0.03	-0.00 ± 0.03	-0.00 ± 0.03
$[6.0, 8.0]$	-0.10 ± 0.03	-0.01 ± 0.03	-0.03 ± 0.03	0.05 ± 0.03	0.00 ± 0.03	0.05 ± 0.03	-0.02 ± 0.03	0.01 ± 0.03
$[11.0, 12.5]$	-0.09 ± 0.03	0.07 ± 0.03	0.03 ± 0.03	0.14 ± 0.03	-0.11 ± 0.03	0.00 ± 0.03	-0.01 ± 0.03	-0.08 ± 0.03
$[15.0, 17.0]$	-0.01 ± 0.03	0.14 ± 0.03	0.01 ± 0.03	0.09 ± 0.03	-0.14 ± 0.03	0.00 ± 0.03	0.04 ± 0.03	0.01 ± 0.03
$[17.0, 19.0]$	-0.04 ± 0.04	0.07 ± 0.03	-0.03 ± 0.04	0.02 ± 0.03	-0.03 ± 0.03	0.02 ± 0.03	0.02 ± 0.03	0.02 ± 0.03
$[1.1, 6.0]$	-0.11 ± 0.03	0.01 ± 0.03	-0.09 ± 0.03	0.08 ± 0.03	0.01 ± 0.03	-0.00 ± 0.03	-0.01 ± 0.03	-0.03 ± 0.03
$[15.0, 19.0]$	0.05 ± 0.03	0.10 ± 0.03	-0.06 ± 0.04	0.17 ± 0.03	0.15 ± 0.03	-0.05 ± 0.03	-0.00 ± 0.03	0.05 ± 0.03

q^2 [GeV^2/c^4]	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
$[0.10, 0.98]$	0.23 ± 0.03	-0.03 ± 0.03	-0.04 ± 0.03	-0.02 ± 0.03	-0.02 ± 0.03	-0.07 ± 0.03	0.03 ± 0.03	-0.02 ± 0.03
$[1.1, 2.5]$	0.01 ± 0.03	-0.08 ± 0.03	0.42 ± 0.03	0.01 ± 0.03	0.04 ± 0.03	-0.09 ± 0.03	-0.01 ± 0.03	0.00 ± 0.03
$[2.5, 4.0]$	-0.12 ± 0.03	-0.04 ± 0.03	0.29 ± 0.03	0.01 ± 0.03	-0.02 ± 0.03	0.08 ± 0.03	0.01 ± 0.03	0.03 ± 0.03
$[4.0, 6.0]$	-0.22 ± 0.03	0.06 ± 0.03	-0.04 ± 0.03	-0.01 ± 0.03	-0.01 ± 0.03	0.06 ± 0.03	0.02 ± 0.03	-0.03 ± 0.03
$[6.0, 8.0]$	-0.16 ± 0.03	0.07 ± 0.03	-0.08 ± 0.03	-0.03 ± 0.03	-0.09 ± 0.03	0.05 ± 0.03	-0.01 ± 0.03	0.02 ± 0.03
$[11.0, 12.5]$	-0.15 ± 0.03	0.07 ± 0.03	-0.13 ± 0.03	0.00 ± 0.03	-0.01 ± 0.03	0.09 ± 0.03	0.03 ± 0.03	-0.02 ± 0.03
$[15.0, 17.0]$	-0.05 ± 0.03	0.08 ± 0.03	-0.04 ± 0.03	0.03 ± 0.03	0.03 ± 0.03	0.03 ± 0.03	-0.00 ± 0.03	-0.01 ± 0.03
$[17.0, 19.0]$	-0.08 ± 0.04	0.11 ± 0.03	-0.03 ± 0.04	-0.01 ± 0.03	-0.04 ± 0.03	-0.03 ± 0.03	0.05 ± 0.03	0.03 ± 0.03
$[1.1, 6.0]$	-0.17 ± 0.03	-0.04 ± 0.03	-0.09 ± 0.03	-0.01 ± 0.03	-0.18 ± 0.03	0.13 ± 0.03	-0.02 ± 0.03	0.02 ± 0.03
$[15.0, 19.0]$	-0.05 ± 0.03	0.14 ± 0.03	0.04 ± 0.04	0.03 ± 0.03	-0.10 ± 0.03	0.18 ± 0.03	-0.04 ± 0.03	0.03 ± 0.03

Table A.4: Results of the Gaussian width of the pull distributions from pseudoexperiments with statistics equal to the statistics in data and with generated angular observable values according to predictions of a new physics contribution of $\mathrm{Re}(\mathcal{C}_9^{\mathrm{NP}}) = -1$, as determined by the FLAVIO software package. Deviations from the nominal value, zero, are observed for some observables. A systematic uncertainty is assigned for this bias, as discussed in Section 6.8. More discussion on these results is given in Section 5.7.1.

q^2 [GeV^2/c^4]	F_{L}	S_3	S_4	S_5	A_{FB}	S_7	S_8	S_9
$[0.10, 0.98]$	1.00 ± 0.02	0.98 ± 0.02	1.11 ± 0.02	1.03 ± 0.02	1.00 ± 0.02	1.03 ± 0.02	1.07 ± 0.02	0.97 ± 0.02
$[1.1, 2.5]$	1.04 ± 0.02	0.98 ± 0.02	1.02 ± 0.02	0.96 ± 0.02	1.02 ± 0.02	0.97 ± 0.02	1.06 ± 0.02	1.01 ± 0.02
$[2.5, 4.0]$	0.98 ± 0.02	0.97 ± 0.02	1.04 ± 0.02	0.97 ± 0.02	1.04 ± 0.02	0.98 ± 0.02	1.03 ± 0.02	1.00 ± 0.02
$[4.0, 6.0]$	0.97 ± 0.02	1.03 ± 0.02	1.02 ± 0.02	0.97 ± 0.02	1.04 ± 0.02	0.97 ± 0.02	1.03 ± 0.02	1.00 ± 0.02
$[6.0, 8.0]$	0.96 ± 0.02	1.01 ± 0.02	1.10 ± 0.02	1.01 ± 0.02	1.02 ± 0.02	1.02 ± 0.02	1.02 ± 0.02	1.02 ± 0.02
$[11.0, 12.5]$	1.09 ± 0.02	0.98 ± 0.02	1.03 ± 0.02	1.00 ± 0.02	1.10 ± 0.02	1.02 ± 0.02	0.99 ± 0.02	0.97 ± 0.02
$[15.0, 17.0]$	1.05 ± 0.02	0.98 ± 0.02	1.08 ± 0.02	1.08 ± 0.02	1.06 ± 0.02	1.01 ± 0.02	1.05 ± 0.02	1.04 ± 0.02
$[17.0, 19.0]$	1.12 ± 0.03	1.01 ± 0.02	1.10 ± 0.02	1.10 ± 0.02	1.06 ± 0.02	1.05 ± 0.02	1.08 ± 0.02	1.07 ± 0.02
$[1.1, 6.0]$	0.92 ± 0.02	0.98 ± 0.02	0.97 ± 0.02	1.02 ± 0.02	1.02 ± 0.02	1.03 ± 0.02	1.00 ± 0.02	1.00 ± 0.02
$[15.0, 19.0]$	1.03 ± 0.02	1.00 ± 0.02	1.11 ± 0.02	1.01 ± 0.02	1.00 ± 0.02	1.00 ± 0.02	1.03 ± 0.02	1.02 ± 0.02

q^2 [GeV^2/c^4]	F_{L}	P_1	P_2	P_3	P_4'	P_5'	P_6'	P_8'
$[0.10, 0.98]$	0.93 ± 0.02	0.97 ± 0.02	1.06 ± 0.02	0.96 ± 0.02	1.06 ± 0.02	0.96 ± 0.02	1.00 ± 0.02	1.02 ± 0.02
$[1.1, 2.5]$	1.17 ± 0.02	0.85 ± 0.02	0.81 ± 0.02	0.78 ± 0.02	0.92 ± 0.02	0.91 ± 0.02	0.87 ± 0.02	0.95 ± 0.02
$[2.5, 4.0]$	1.20 ± 0.02	0.80 ± 0.02	0.81 ± 0.02	0.71 ± 0.02	0.94 ± 0.02	0.83 ± 0.02	0.94 ± 0.02	0.92 ± 0.02
$[4.0, 6.0]$	0.92 ± 0.02	0.79 ± 0.02	0.81 ± 0.02	0.82 ± 0.02	0.98 ± 0.02	0.91 ± 0.02	0.89 ± 0.02	0.96 ± 0.02
$[6.0, 8.0]$	0.93 ± 0.02	0.89 ± 0.02	0.95 ± 0.02	0.88 ± 0.02	1.05 ± 0.02	0.99 ± 0.02	1.01 ± 0.02	1.03 ± 0.02
$[11.0, 12.5]$	1.17 ± 0.02	0.91 ± 0.02	0.87 ± 0.02	0.93 ± 0.02	1.00 ± 0.02	0.98 ± 0.02	0.99 ± 0.02	1.03 ± 0.02
$[15.0, 17.0]$	1.23 ± 0.02	0.97 ± 0.02	0.96 ± 0.02	1.03 ± 0.02	1.03 ± 0.02	1.04 ± 0.02	1.00 ± 0.02	0.98 ± 0.02
$[17.0, 19.0]$	1.16 ± 0.03	1.00 ± 0.02	1.03 ± 0.02	1.03 ± 0.02	1.10 ± 0.02	1.03 ± 0.02	1.04 ± 0.02	1.02 ± 0.02
$[1.1, 6.0]$	0.85 ± 0.02	0.90 ± 0.02	0.93 ± 0.02	0.90 ± 0.02	0.96 ± 0.02	0.98 ± 0.02	0.94 ± 0.02	0.96 ± 0.02
$[15.0, 19.0]$	1.09 ± 0.02	0.99 ± 0.02	0.99 ± 0.02	1.00 ± 0.02	1.06 ± 0.02	1.03 ± 0.02	0.97 ± 0.02	1.02 ± 0.02

Table A.5: Results of the mean position of the pull distributions from pseudoexperiments with statistics equal to the statistics in data and with generated angular observable values according to result values obtained by the maximum-likelihood fit, equal to the central values in the results in Tables 7.1 and 7.2. Deviations from the nominal value, zero, are observed for some observables. A systematic uncertainty is assigned for this bias, as discussed in Section 6.8. More discussion on these results is given in Section 5.7.1.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	0.21 ± 0.03	-0.05 ± 0.03	-0.03 ± 0.03	-0.04 ± 0.03	-0.00 ± 0.03	0.05 ± 0.03	0.01 ± 0.03	0.01 ± 0.03
$[1.1, 2.5]$	0.07 ± 0.03	-0.07 ± 0.03	-0.06 ± 0.03	-0.10 ± 0.03	0.09 ± 0.03	0.02 ± 0.03	0.07 ± 0.03	-0.01 ± 0.03
$[2.5, 4.0]$	-0.01 ± 0.04	-0.02 ± 0.03	0.17 ± 0.03	0.12 ± 0.03	-0.01 ± 0.03	0.04 ± 0.03	0.06 ± 0.03	-0.08 ± 0.03
$[4.0, 6.0]$	-0.09 ± 0.03	0.05 ± 0.03	0.32 ± 0.03	0.04 ± 0.03	0.03 ± 0.03	0.01 ± 0.03	0.02 ± 0.03	0.10 ± 0.03
$[6.0, 8.0]$	-0.05 ± 0.03	0.08 ± 0.03	-0.03 ± 0.04	0.01 ± 0.03	0.06 ± 0.03	0.11 ± 0.03	0.06 ± 0.03	-0.01 ± 0.03
$[11.0, 12.5]$	-0.11 ± 0.03	0.02 ± 0.03	0.08 ± 0.03	0.14 ± 0.03	-0.74 ± 0.04	0.02 ± 0.03	-0.01 ± 0.03	-0.23 ± 0.03
$[15.0, 17.0]$	-0.14 ± 0.03	0.20 ± 0.03	-0.02 ± 0.03	-0.09 ± 0.03	-0.16 ± 0.04	0.06 ± 0.03	0.01 ± 0.03	-0.12 ± 0.03
$[17.0, 19.0]$	-0.09 ± 0.04	-0.02 ± 0.03	-0.01 ± 0.03	0.03 ± 0.03	-0.01 ± 0.03	0.01 ± 0.03	0.04 ± 0.03	-0.00 ± 0.03
$[1.1, 6.0]$	-0.02 ± 0.03	-0.00 ± 0.03	-0.07 ± 0.03	0.10 ± 0.03	0.02 ± 0.03	0.04 ± 0.03	0.03 ± 0.03	0.01 ± 0.03
$[15.0, 19.0]$	-0.05 ± 0.03	0.19 ± 0.03	0.62 ± 0.03	-0.01 ± 0.03	-0.65 ± 0.03	-0.03 ± 0.03	0.03 ± 0.03	-0.07 ± 0.03

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.24 ± 0.03	-0.01 ± 0.03	-0.05 ± 0.03	-0.00 ± 0.03	-0.07 ± 0.03	0.03 ± 0.03	-0.05 ± 0.03	0.03 ± 0.03
$[1.1, 2.5]$	-0.02 ± 0.03	-0.48 ± 0.03	0.20 ± 0.03	0.05 ± 0.03	-0.13 ± 0.03	-0.02 ± 0.03	0.02 ± 0.03	0.00 ± 0.03
$[2.5, 4.0]$	0.25 ± 0.04	0.05 ± 0.03	-0.01 ± 0.03	0.14 ± 0.03	-0.04 ± 0.03	0.24 ± 0.03	0.09 ± 0.03	-0.03 ± 0.03
$[4.0, 6.0]$	-0.22 ± 0.03	0.41 ± 0.03	0.19 ± 0.03	-0.29 ± 0.03	0.06 ± 0.03	-0.00 ± 0.03	0.01 ± 0.03	0.01 ± 0.03
$[6.0, 8.0]$	-0.15 ± 0.03	0.16 ± 0.03	0.09 ± 0.04	-0.12 ± 0.03	-0.10 ± 0.03	-0.02 ± 0.03	0.07 ± 0.03	0.00 ± 0.03
$[11.0, 12.5]$	-0.21 ± 0.03	0.08 ± 0.03	-0.74 ± 0.03	0.30 ± 0.03	-0.03 ± 0.04	0.12 ± 0.03	0.02 ± 0.03	-0.04 ± 0.03
$[15.0, 17.0]$	-0.09 ± 0.03	0.18 ± 0.03	-0.14 ± 0.03	0.09 ± 0.03	-0.03 ± 0.04	-0.05 ± 0.03	0.10 ± 0.03	0.04 ± 0.03
$[17.0, 19.0]$	0.01 ± 0.04	0.07 ± 0.03	-0.06 ± 0.03	-0.07 ± 0.03	0.02 ± 0.03	-0.10 ± 0.03	0.03 ± 0.03	0.00 ± 0.03
$[1.1, 6.0]$	-0.10 ± 0.03	0.05 ± 0.03	0.05 ± 0.03	-0.00 ± 0.03	-0.08 ± 0.03	0.03 ± 0.03	0.02 ± 0.03	-0.03 ± 0.03
$[15.0, 19.0]$	-0.03 ± 0.03	0.16 ± 0.03	-0.56 ± 0.03	0.03 ± 0.03	0.49 ± 0.03	-0.04 ± 0.03	-0.08 ± 0.03	0.03 ± 0.03

Table A.6: Results of the Gaussian width of the pull distributions from pseudoexperiments with statistics equal to the statistics in data and with generated angular observable values according to result values obtained by the maximum-likelihood fit, equal to the central values in the results in Tables 7.1 and 7.2. Deviations from the nominal value, zero, are observed for some observables. A systematic uncertainty is assigned for this bias, as discussed in Section 6.8. More discussion on these results is given in Section 5.7.1.

q^2 [GeV^2/c^4]	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$[0.10, 0.98]$	1.05 ± 0.02	0.96 ± 0.02	1.07 ± 0.02	1.04 ± 0.02	1.00 ± 0.02	1.02 ± 0.02	1.09 ± 0.02	1.04 ± 0.02
$[1.1, 2.5]$	1.08 ± 0.02	1.02 ± 0.02	0.98 ± 0.02	0.91 ± 0.02	1.00 ± 0.02	1.00 ± 0.02	1.08 ± 0.02	1.01 ± 0.02
$[2.5, 4.0]$	1.10 ± 0.02	0.96 ± 0.02	0.99 ± 0.02	0.97 ± 0.02	1.00 ± 0.02	0.97 ± 0.02	1.01 ± 0.02	0.98 ± 0.02
$[4.0, 6.0]$	0.98 ± 0.02	0.99 ± 0.02	1.04 ± 0.02	0.99 ± 0.02	1.04 ± 0.02	0.98 ± 0.02	1.03 ± 0.02	1.03 ± 0.02
$[6.0, 8.0]$	1.10 ± 0.02	1.01 ± 0.02	1.11 ± 0.02	1.00 ± 0.02	1.02 ± 0.02	1.02 ± 0.02	1.01 ± 0.02	1.03 ± 0.02
$[11.0, 12.5]$	1.08 ± 0.02	0.98 ± 0.02	1.02 ± 0.02	0.94 ± 0.02	1.25 ± 0.03	0.94 ± 0.02	1.00 ± 0.02	1.11 ± 0.02
$[15.0, 17.0]$	1.07 ± 0.02	1.04 ± 0.02	1.05 ± 0.02	1.05 ± 0.02	1.11 ± 0.02	1.01 ± 0.02	0.99 ± 0.02	1.08 ± 0.02
$[17.0, 19.0]$	1.15 ± 0.03	1.02 ± 0.02	1.10 ± 0.02	1.09 ± 0.02	1.08 ± 0.02	1.06 ± 0.02	1.07 ± 0.02	1.10 ± 0.02
$[1.1, 6.0]$	1.00 ± 0.02	1.03 ± 0.02	1.05 ± 0.02	0.98 ± 0.02	1.00 ± 0.02	0.96 ± 0.02	1.03 ± 0.02	0.99 ± 0.02
$[15.0, 19.0]$	1.06 ± 0.02	0.96 ± 0.02	0.99 ± 0.02	1.00 ± 0.02	1.08 ± 0.02	0.96 ± 0.02	1.01 ± 0.02	1.02 ± 0.02

q^2 [GeV^2/c^4]	F_L	P_1	P_2	P_3	P'_4	P'_5	P'_6	P'_8
$[0.10, 0.98]$	0.97 ± 0.02	0.93 ± 0.02	1.03 ± 0.02	0.94 ± 0.02	1.07 ± 0.02	1.00 ± 0.02	1.01 ± 0.02	1.08 ± 0.02
$[1.1, 2.5]$	1.01 ± 0.02	1.00 ± 0.02	0.93 ± 0.02	0.85 ± 0.02	1.04 ± 0.02	0.96 ± 0.02	0.97 ± 0.02	1.02 ± 0.02
$[2.5, 4.0]$	1.05 ± 0.02	0.93 ± 0.02	0.95 ± 0.02	0.90 ± 0.02	1.05 ± 0.02	0.98 ± 0.02	1.00 ± 0.02	0.92 ± 0.02
$[4.0, 6.0]$	0.93 ± 0.02	0.86 ± 0.02	0.84 ± 0.02	0.87 ± 0.02	1.05 ± 0.02	0.94 ± 0.02	0.92 ± 0.02	0.97 ± 0.02
$[6.0, 8.0]$	1.24 ± 0.02	0.95 ± 0.02	0.98 ± 0.02	0.98 ± 0.02	1.05 ± 0.02	1.01 ± 0.02	1.01 ± 0.02	1.01 ± 0.02
$[11.0, 12.5]$	1.29 ± 0.02	0.95 ± 0.02	1.18 ± 0.02	1.02 ± 0.02	1.03 ± 0.03	0.99 ± 0.02	1.01 ± 0.02	0.99 ± 0.02
$[15.0, 17.0]$	1.09 ± 0.02	0.94 ± 0.02	1.03 ± 0.02	1.05 ± 0.02	1.04 ± 0.02	1.04 ± 0.02	1.03 ± 0.02	1.05 ± 0.02
$[17.0, 19.0]$	1.10 ± 0.03	1.02 ± 0.02	1.00 ± 0.02	1.01 ± 0.02	1.05 ± 0.02	1.04 ± 0.02	1.01 ± 0.02	1.04 ± 0.02
$[1.1, 6.0]$	0.90 ± 0.02	0.96 ± 0.02	0.94 ± 0.02	0.96 ± 0.02	1.00 ± 0.02	0.97 ± 0.02	0.99 ± 0.02	1.03 ± 0.02
$[15.0, 19.0]$	1.11 ± 0.02	0.97 ± 0.02	0.99 ± 0.02	1.01 ± 0.02	1.03 ± 0.02	1.04 ± 0.02	0.99 ± 0.02	1.05 ± 0.02

A.4 Method of moments

The method of moments [209] is a statistical tool to extract angular observables from data without the use of a likelihood fit. In this analysis, it is used as a validation cross-check to confirm the results obtained by the maximum-likelihood fit, as described in Section 6.11. For this, the angular moments of the differential angular decay rate, as defined in Eq. (2.18) are multiplied with the full differential decay rate, as in Eq. (2.26), to obtain the product

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\cos\theta_K\mathrm{d}\cos\theta_\ell\mathrm{d}\phi}\bigg|_{\mathrm{P+S}} \times f_i(\cos\theta_K,\cos\theta_\ell,\phi)\,. \tag{A.1}$$

By integrating this product over the full angular space, $\vec{\Omega}=(\cos\theta_K,\cos\theta_\ell,\phi)$, one obtains the expectation value M_i of each angular moment f_i

$$\begin{aligned} M_i &= \int\limits_{\vec{\Omega}} \frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}^3\vec{\Omega}} \times f_i(\vec{\Omega})\,\mathrm{d}^3\vec{\Omega} \\ &= \frac{9}{32\pi}\int\limits_{\vec{\Omega}} \sum_j S_j f_j(\vec{\Omega}) \times f_i(\vec{\Omega})\,\mathrm{d}^3\vec{\Omega}\,. \end{aligned} \tag{A.2}$$

This relation provides a link between the expectation values M_i and the angular observables S_i. Due to the orthogonality of angular moments, most terms cancel out under integration [1]. Solving the integral for all angular observables S_i, one obtains the direct relations

$$\begin{aligned} S_{1s} &= \tfrac{15}{8}\left(M_{1s}-\tfrac{2}{3}\right) & S_{6s} &= 3M_{6c}-2M_{6s} \\ S_3 &= \tfrac{25}{8}M_3 & S_7 &= \tfrac{5}{2}M_7 \\ S_4 &= \tfrac{25}{8}M_4 & S_8 &= \tfrac{25}{8}M_8 \\ S_5 &= \tfrac{5}{2}M_5 & S_9 &= \tfrac{25}{8}M_9 \end{aligned} \tag{A.3}$$

From these, the angular observables are determined by calculating the expectation values M_i from background-subtracted data. Summing over all candidates, e, the expectation value is derived by

$$M_i = \frac{1}{\sum_{e=1}^{N} w_e}\sum_{e=1}^{N} w_e\,,\, f_i(\vec{\Omega}_e) \tag{A.4}$$

[1] For S_{1s} and S_{6s} a linear equation system has to be solved to obtain the relations in Eq. (A.3).

with the candidate weight, w_e, which includes the weights according to the background subtraction algorithm. The uncertainties on M_i are estimated by the weighted variance

$$\delta M_i = \frac{1}{\sum_{e=1}^{N} w_e} \sqrt{\sum_{e=1}^{N} w_e^2 \left(M_i - f_i(\vec{\Omega}) \right)^2} . \tag{A.5}$$

A.5 Fit projections

The fit projections to the invariant mass of the B meson, $m(K_S^0 \pi^+ \mu^+ \mu^-)$, are already presented at the end of the candidate selection in Fig. 4.8 in Chapter 4. In addition to the mass dimension, which is not impacted by the angular folding method from Section 5.4, the projections from all five folds, in all ten q^2 intervals and in the angular dimension of $\cos\theta_K$, $\cos\theta_\ell$ and ϕ are given in the following Figs. A.5 to A.14.

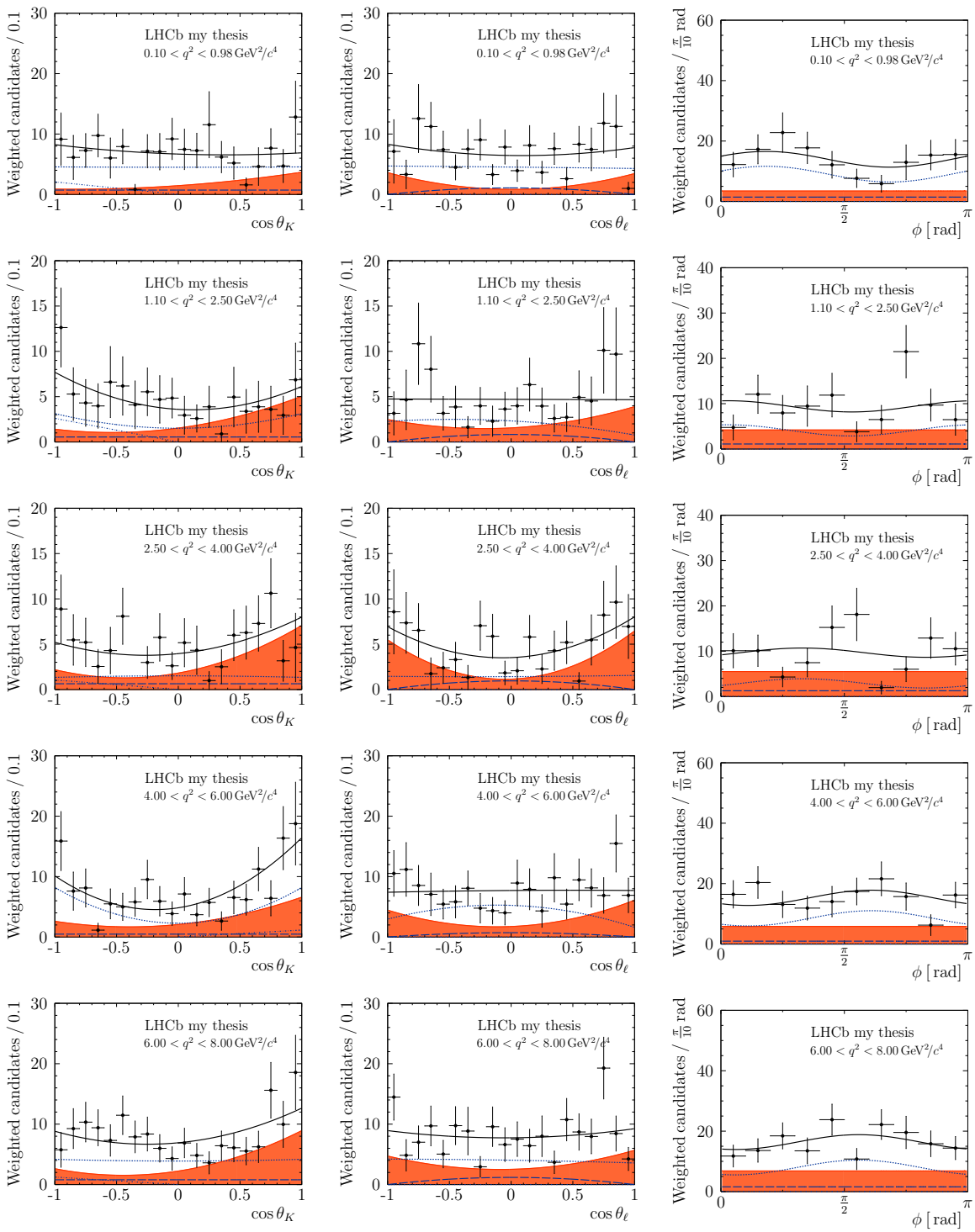

Figure A.5: Angular folding number 0: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in q^2 intervals below $8.0\,\mathrm{GeV}^2/c^4$. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave.

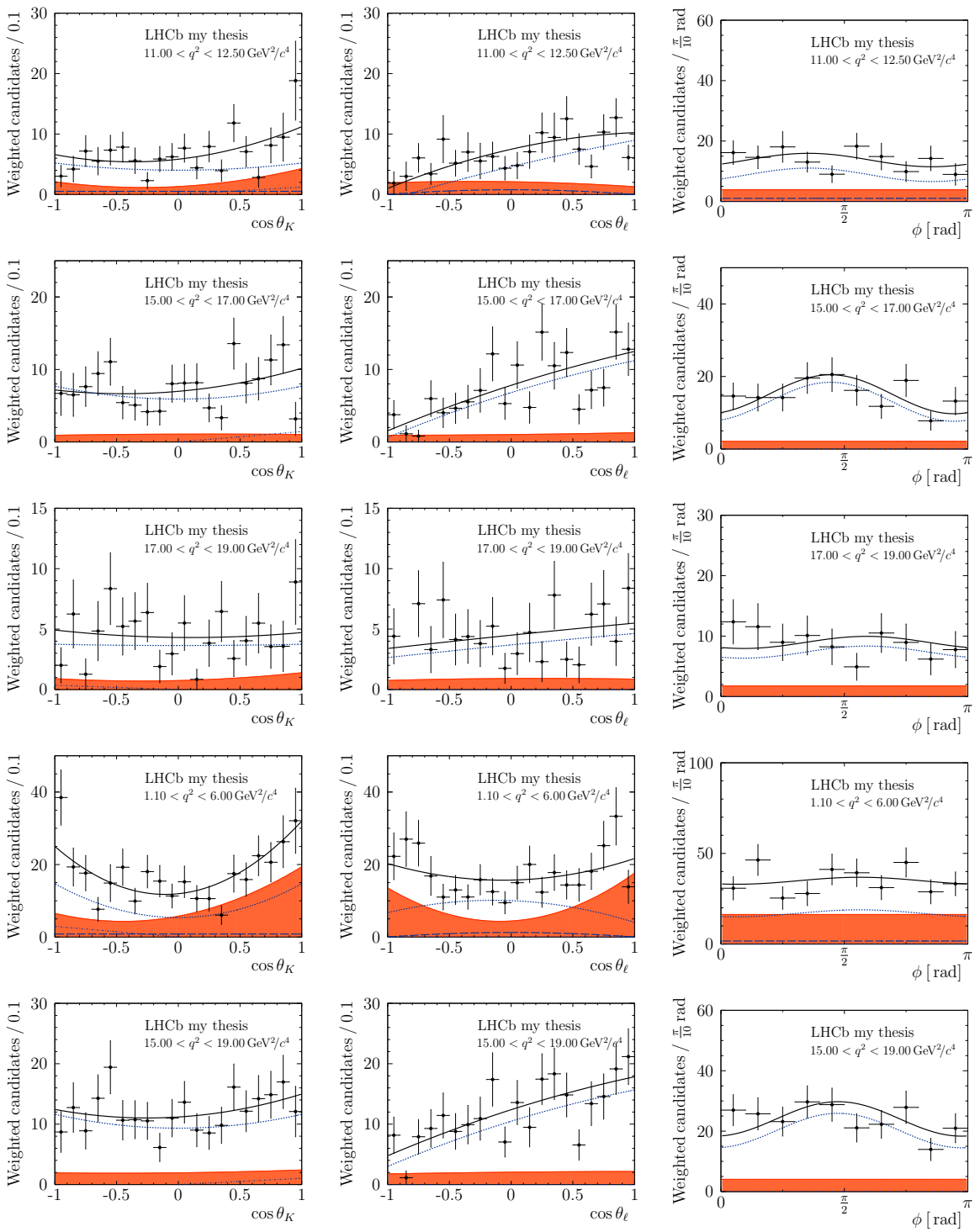

Figure A.6: Angular folding number 0: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in the nominal q^2 intervals above $8.0\,\mathrm{GeV}^2/c^4$ and in the two larger q^2 intervals. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave.

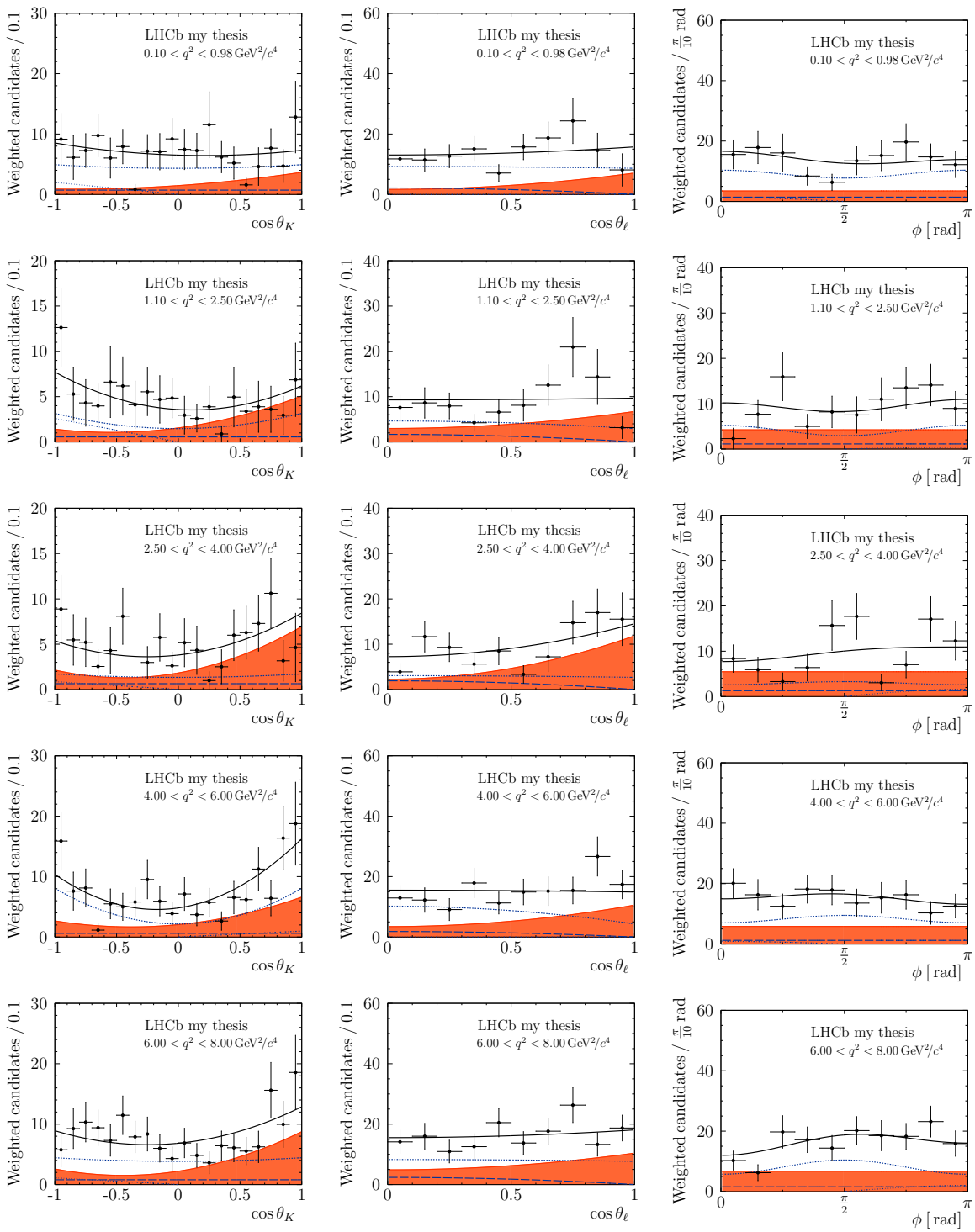

Figure A.7: Angular folding number 1: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in q^2 intervals below $8.0\,\mathrm{GeV}^2/c^4$. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

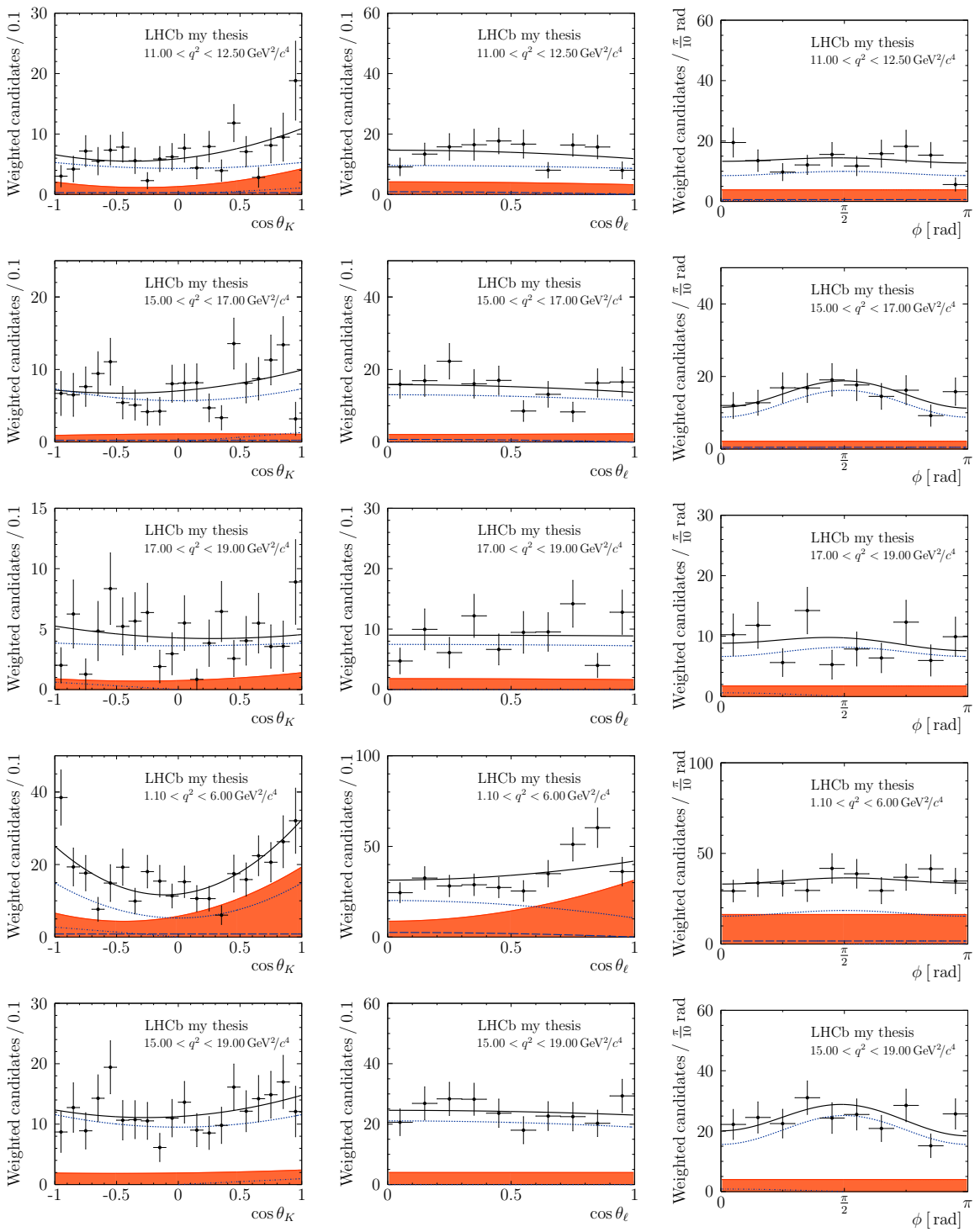

Figure A.8: Angular folding number 1: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in the nominal q^2 intervals above $8.0\,\mathrm{GeV}^2/c^4$ and in the two larger q^2 intervals. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

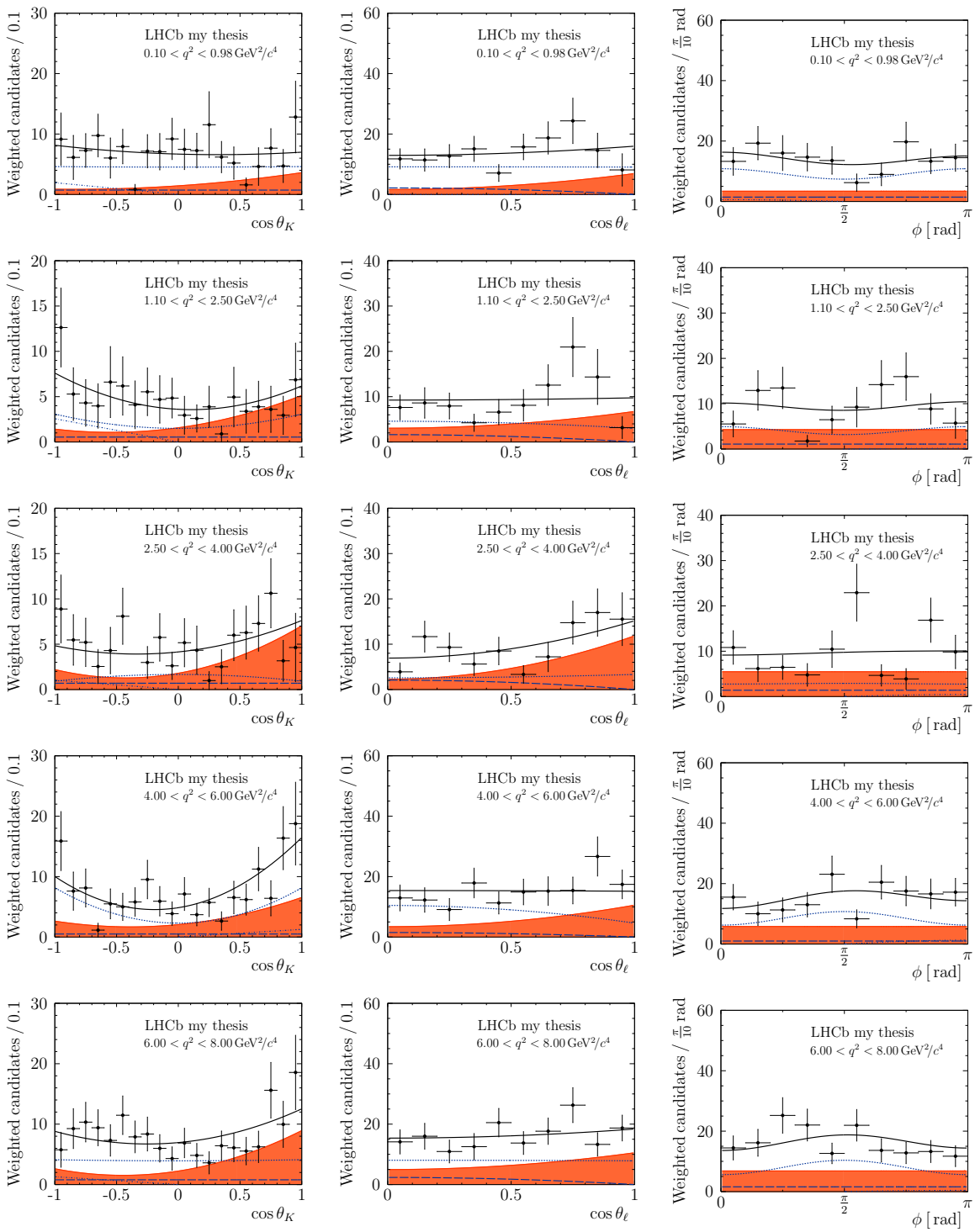

Figure A.9: Angular folding number 2: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in q^2 intervals below $8.0\,\mathrm{GeV}^2/c^4$. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

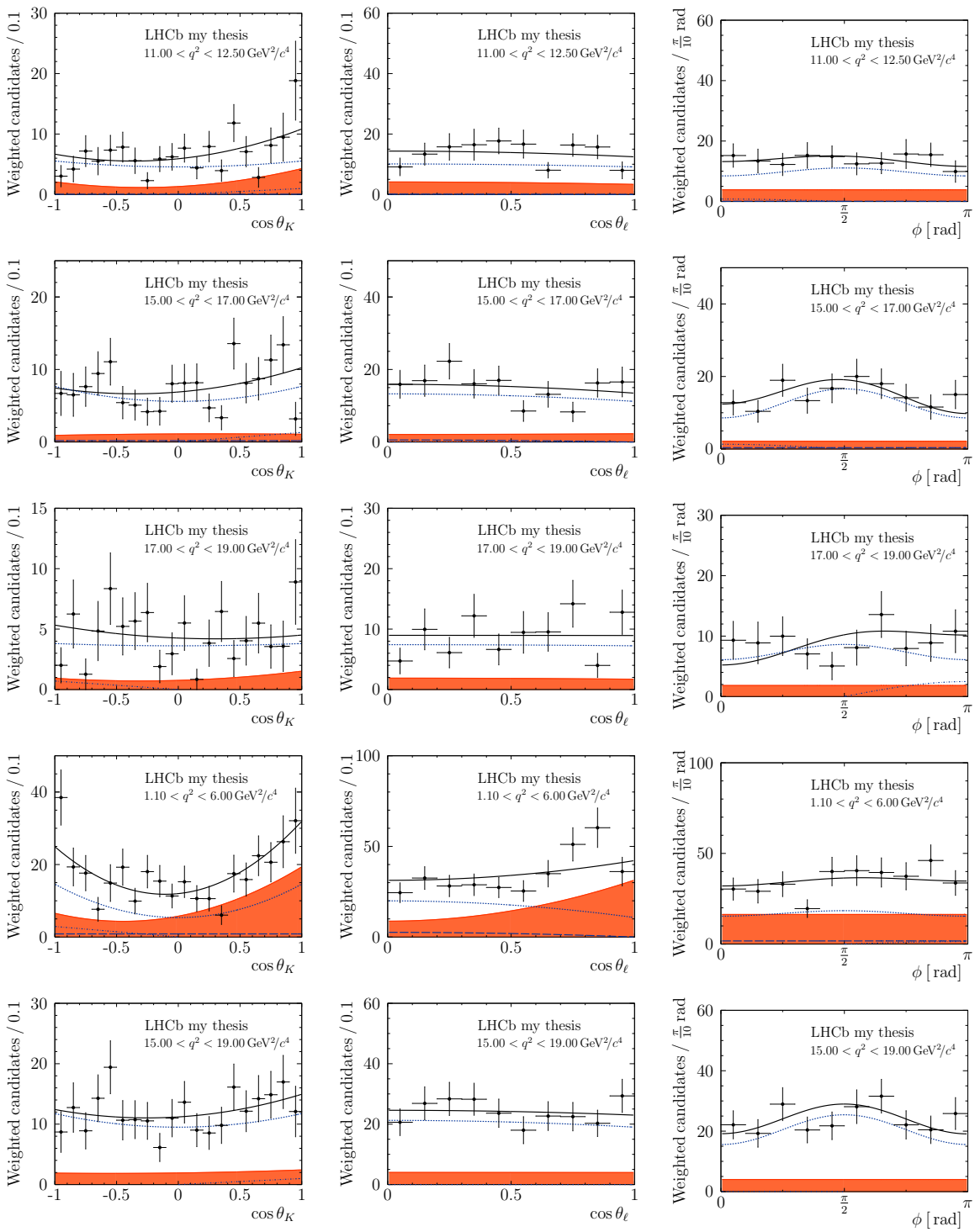

Figure A.10: Angular folding number 2: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in the nominal q^2 intervals above $8.0\,\mathrm{GeV}^2/c^4$ and in the two larger q^2 intervals. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

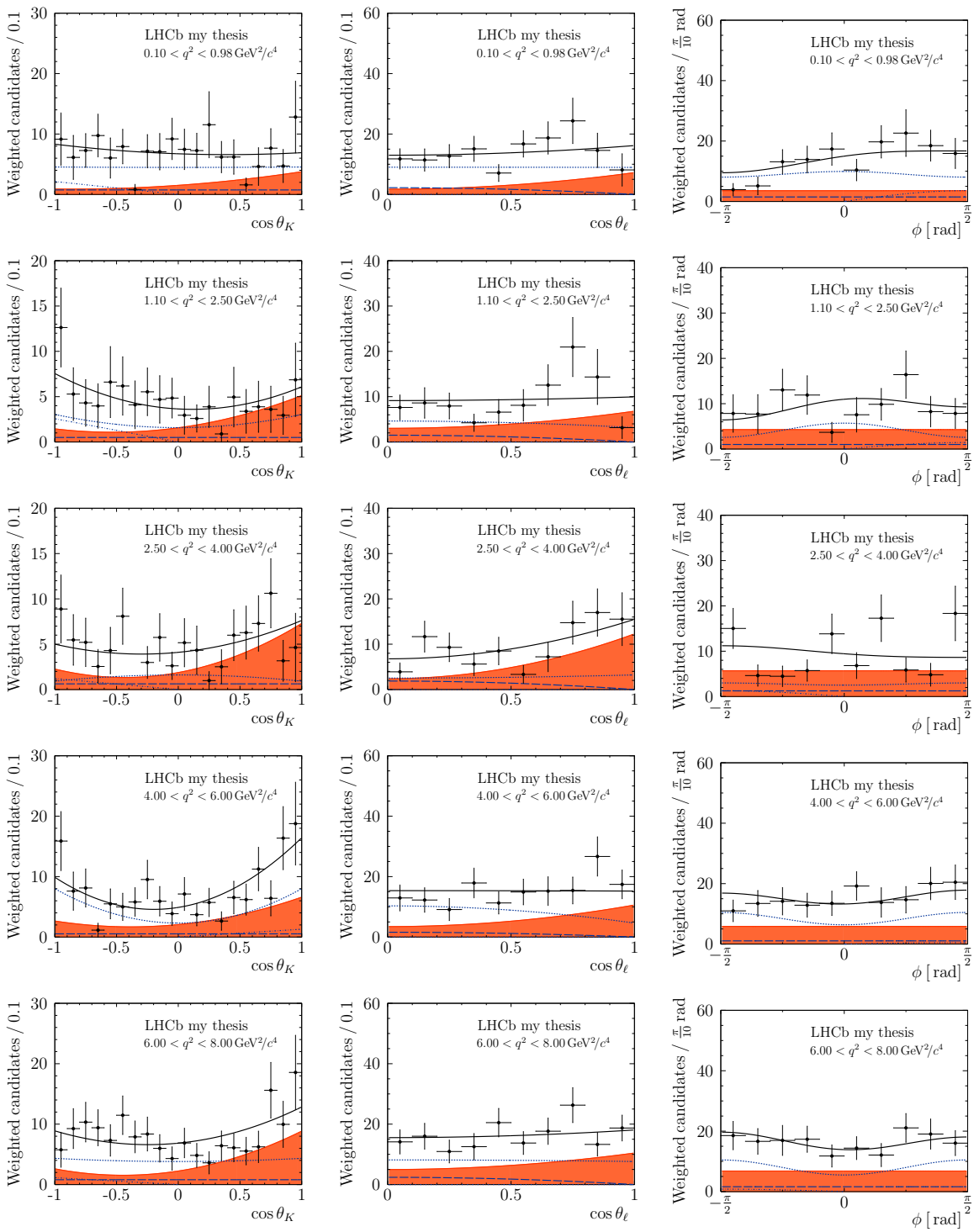

Figure A.11: Angular folding number 3: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in q^2 intervals below $8.0\,\mathrm{GeV}^2/c^4$. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

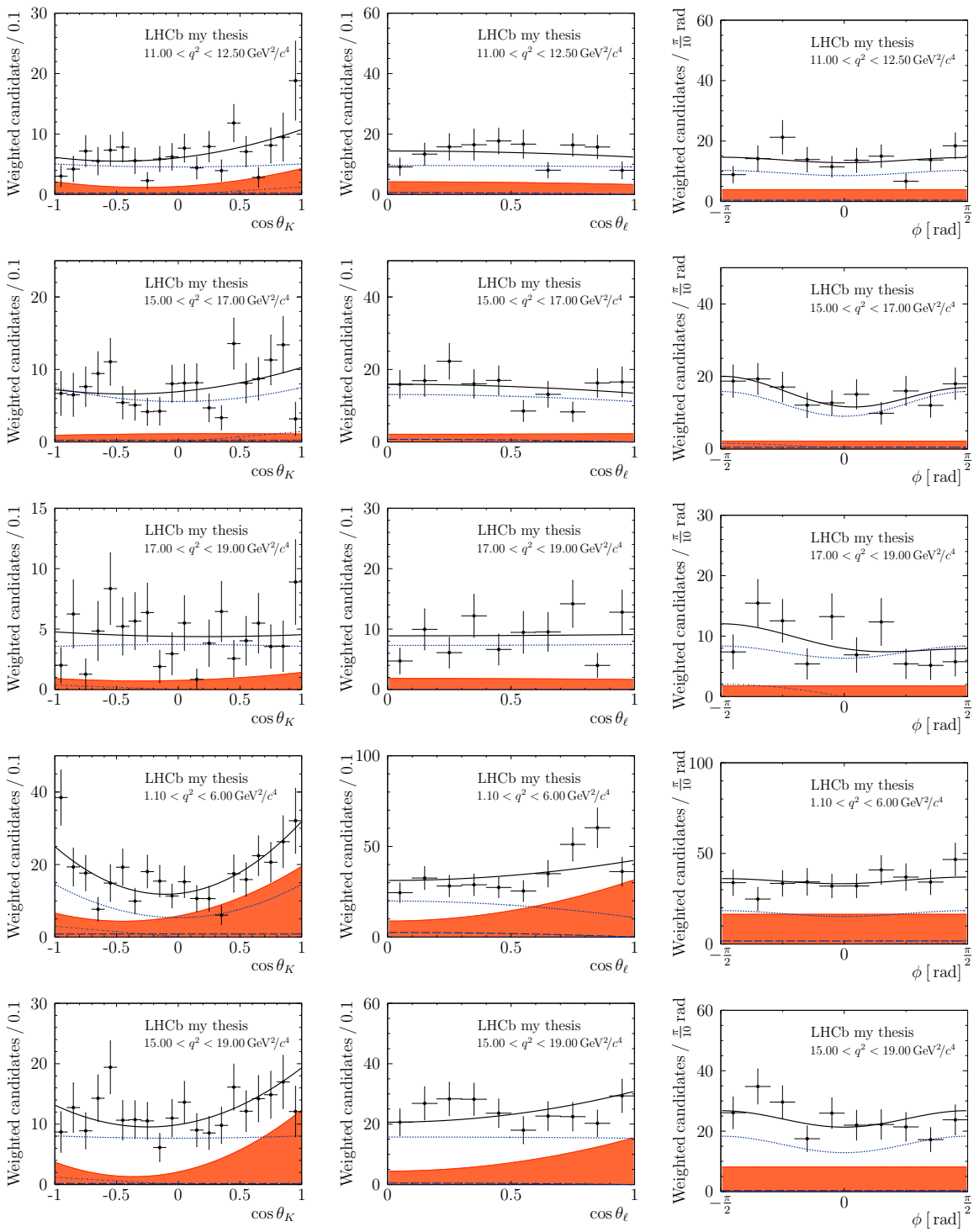

Figure A.12: Angular folding number 3: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in the nominal q^2 intervals above $8.0\,\mathrm{GeV}^2/c^4$ and in the two larger q^2 intervals. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

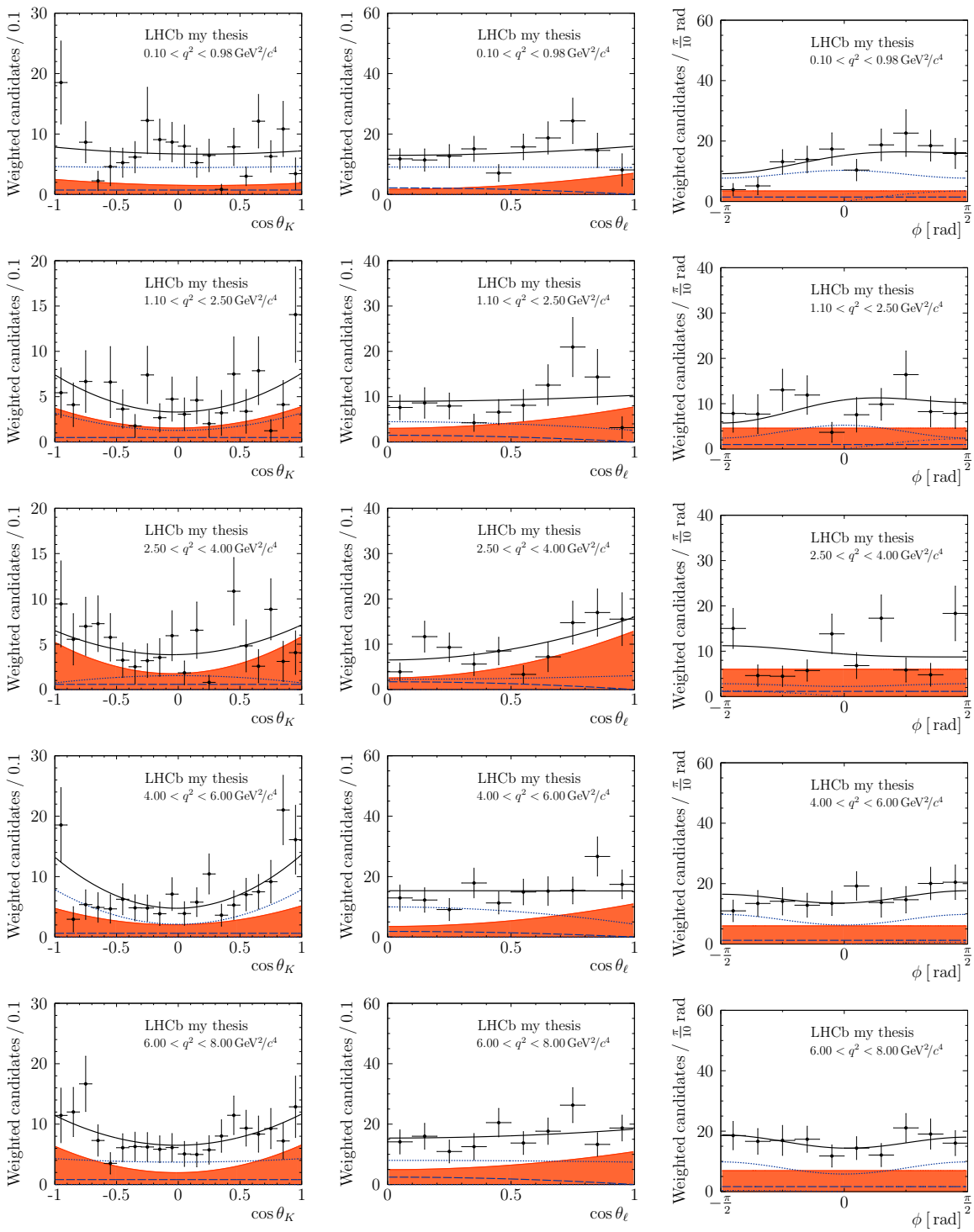

Figure A.13: Angular folding number 4: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in q^2 intervals below $8.0\,\mathrm{GeV}^2/c^4$. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

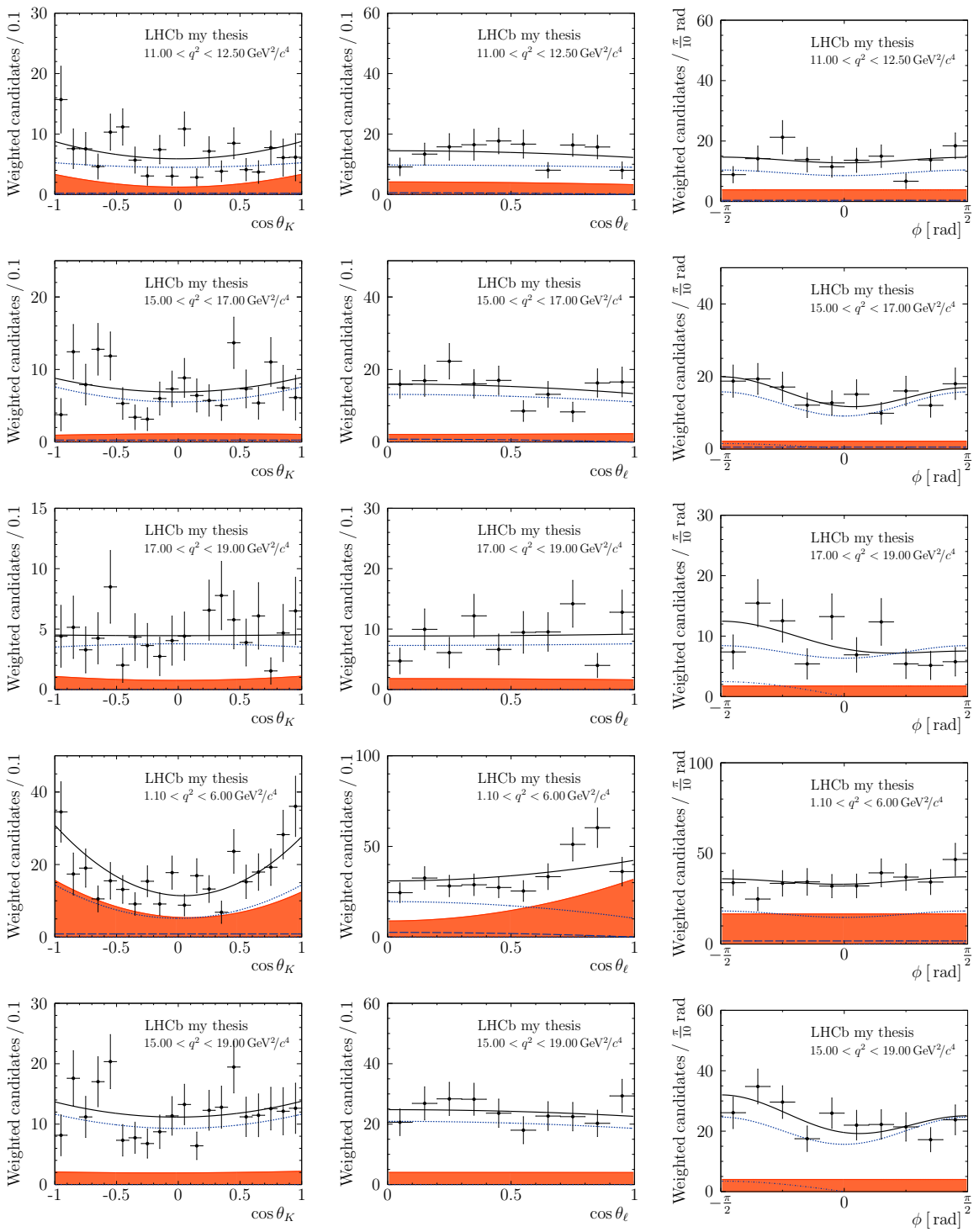

Figure A.14: Angular folding number 4: Projections from the four-dimensional maximum-likelihood fit in the (left) $\cos\theta_K$, (centre) $\cos\theta_\ell$ and (right) ϕ dimensions in the nominal q^2 intervals above $8.0\,\mathrm{GeV}^2/c^4$ and in the two larger q^2 intervals. The data set is fit simultaneously in the four sub categories and recombined for these projections. Data in the black markers is described by the total fit model in solid black line individual fit components for (shaded orange) background, (fine dashed blue) P-wave, (long dashed blue) S-wave and (dotted dashed blue) interference between P- and S-wave. The data is restricted to a signal window of $\pm 50\,\mathrm{MeV}/c$ around the $m(B^+)$ peak.

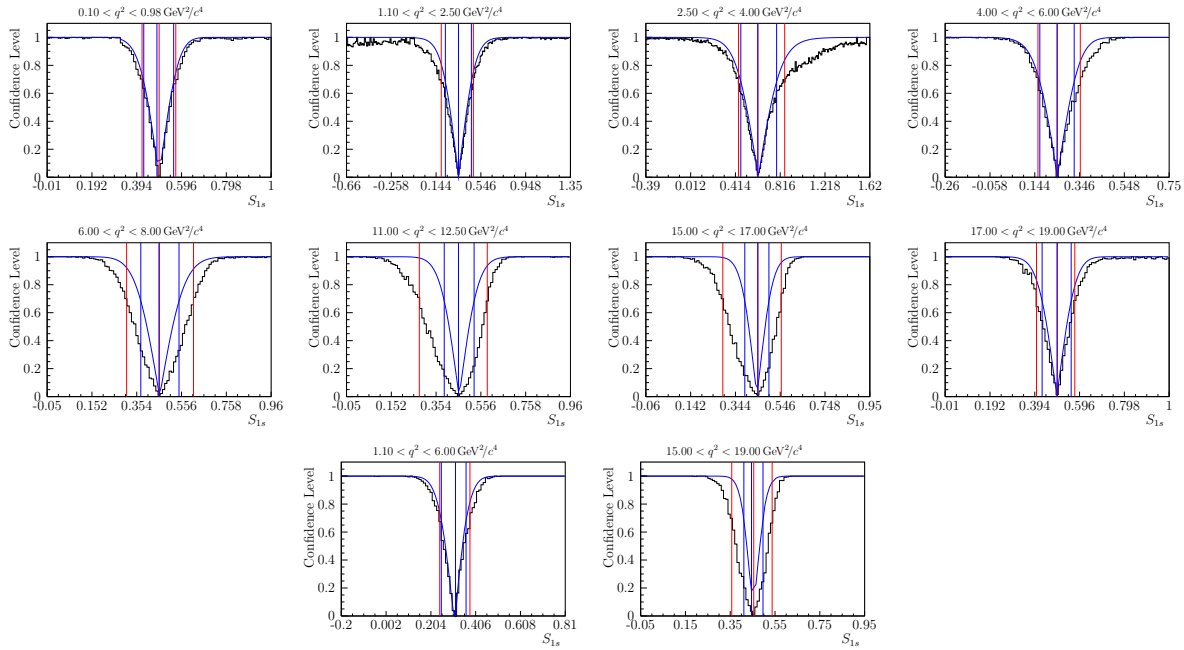

Figure A.15: Likelihood profiles Feldman-Cousins scans for the S_{1s} observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$). Note that the result and uncertainties of S_{1s} are converted to F_{L} in Chapter 7.

A.6 Likelihood profiles of Feldman-Cousins scans

The following Figs. A.15 to A.22 and Figs. A.23 to A.30 present all profiles of the Feldman-Cousins scans in the ten q^2 intervals for the S_i and $P_i^{(\prime)}$ angular observables, respectively. The scans usually are performed in a range of ± 1.0 around the central value obtained from the maximum-likelihood fit, but may be extended to reach the 68.3% level for individual observables. A record-breaking range from about -5.0 to $+8.0$ is scanned for the P_8 angular observable in the q^2 interval $2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$, as seen in Fig. A.30.
The details on the scan technique are explained in Section 5.8.

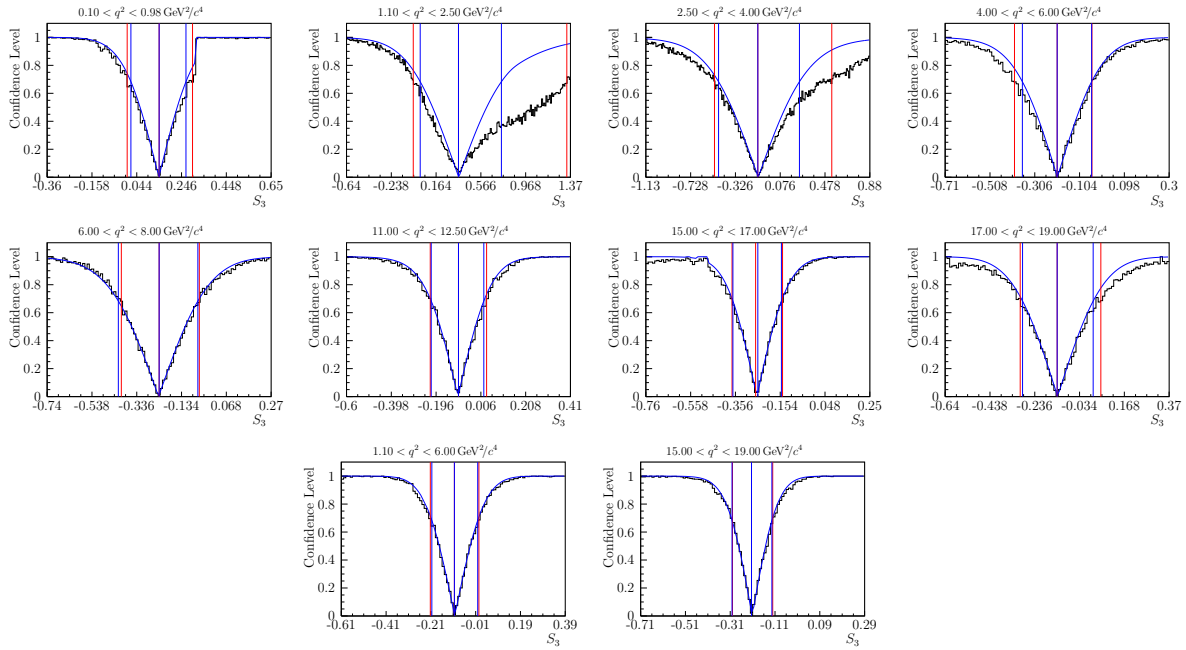

Figure A.16: Likelihood profiles Feldman-Cousins scans for the S_3 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

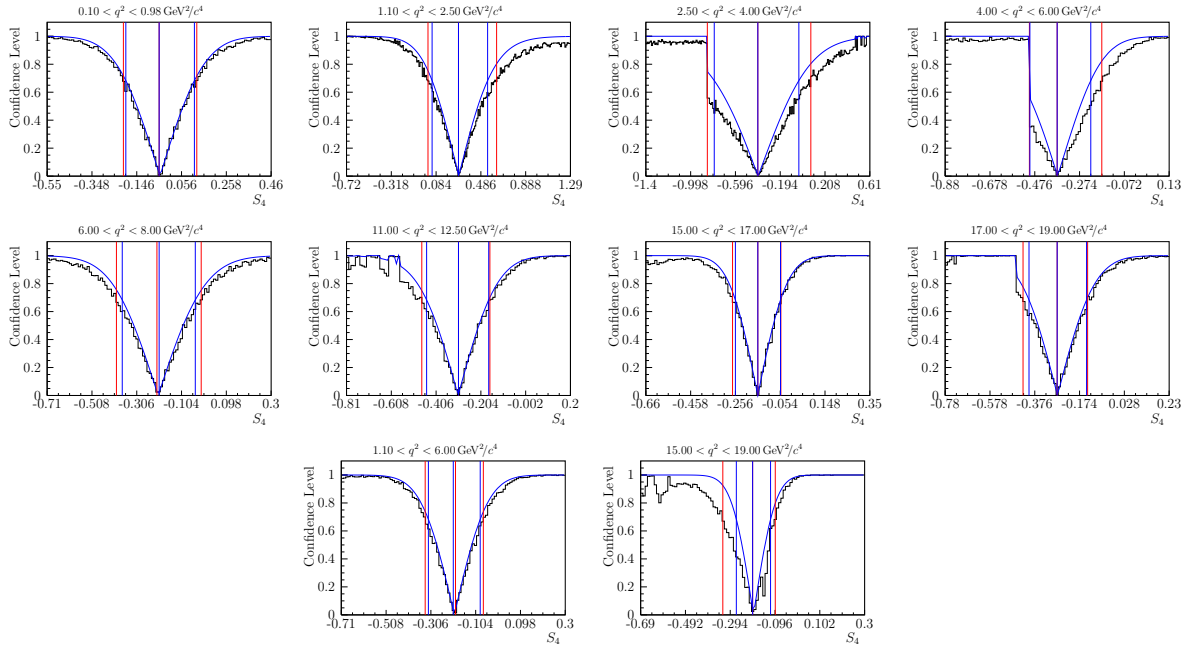

Figure A.17: Likelihood profiles Feldman-Cousins scans for the S_4 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

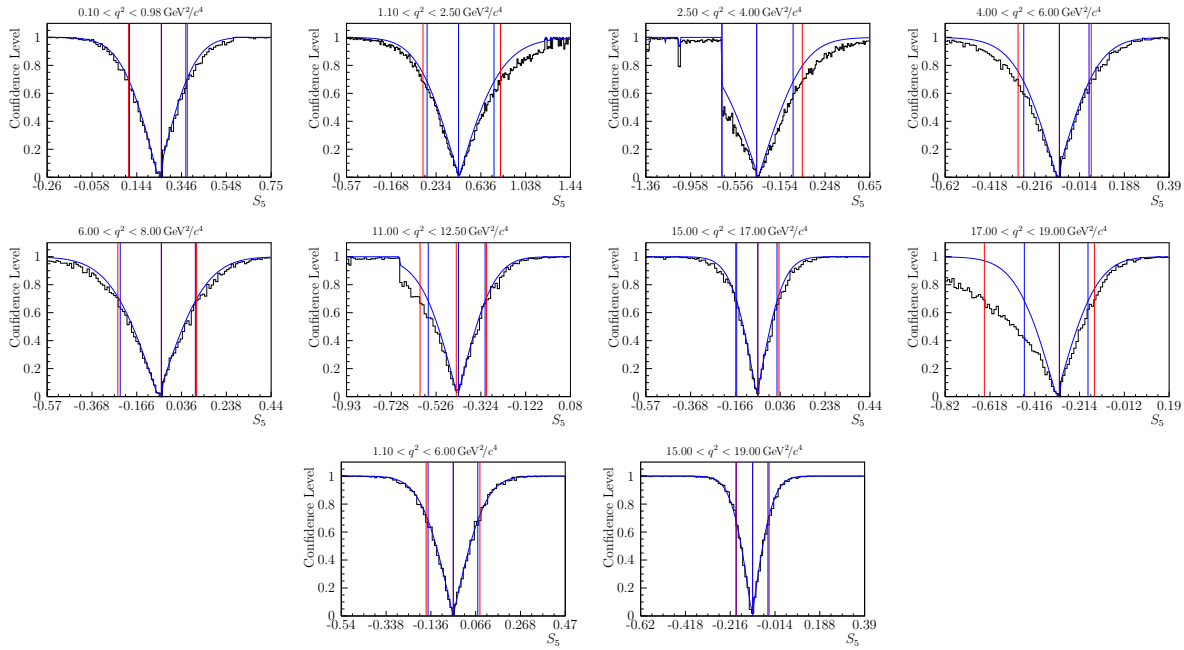

Figure A.18: Likelihood profiles Feldman-Cousins scans for the S_5 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the Δlog-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

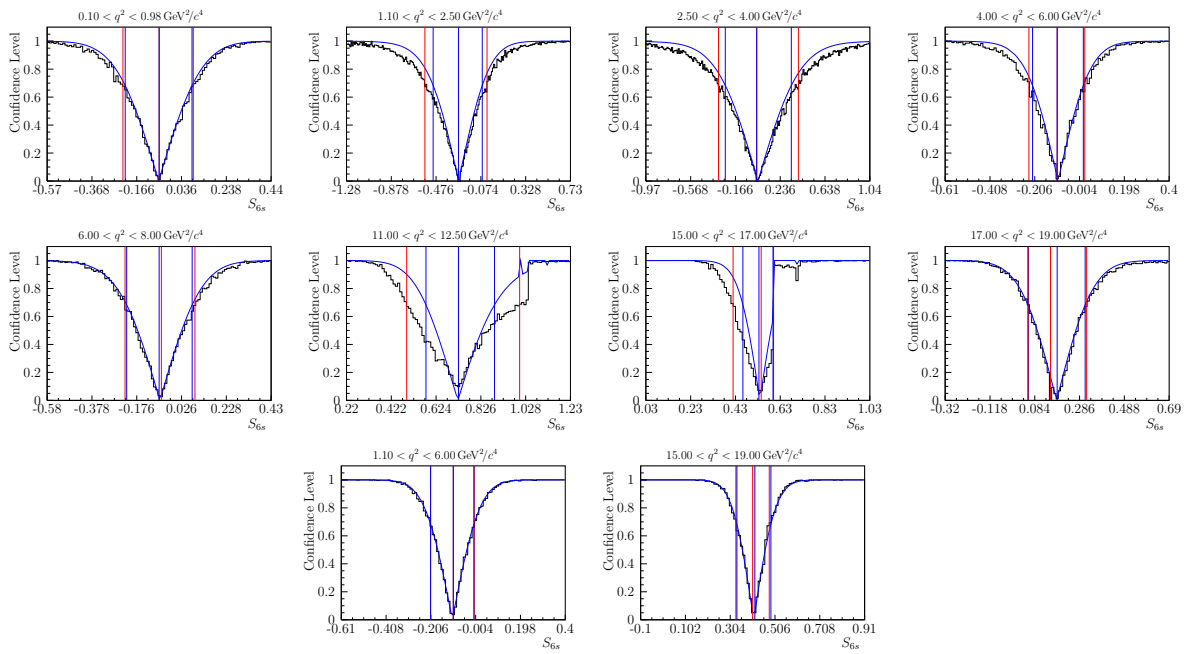

Figure A.19: Likelihood profiles Feldman-Cousins scans for the S_{6s} observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the Δlog-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$). Note that the result and uncertainties of S_{6s} are converted to A_{FB} in Chapter 7.

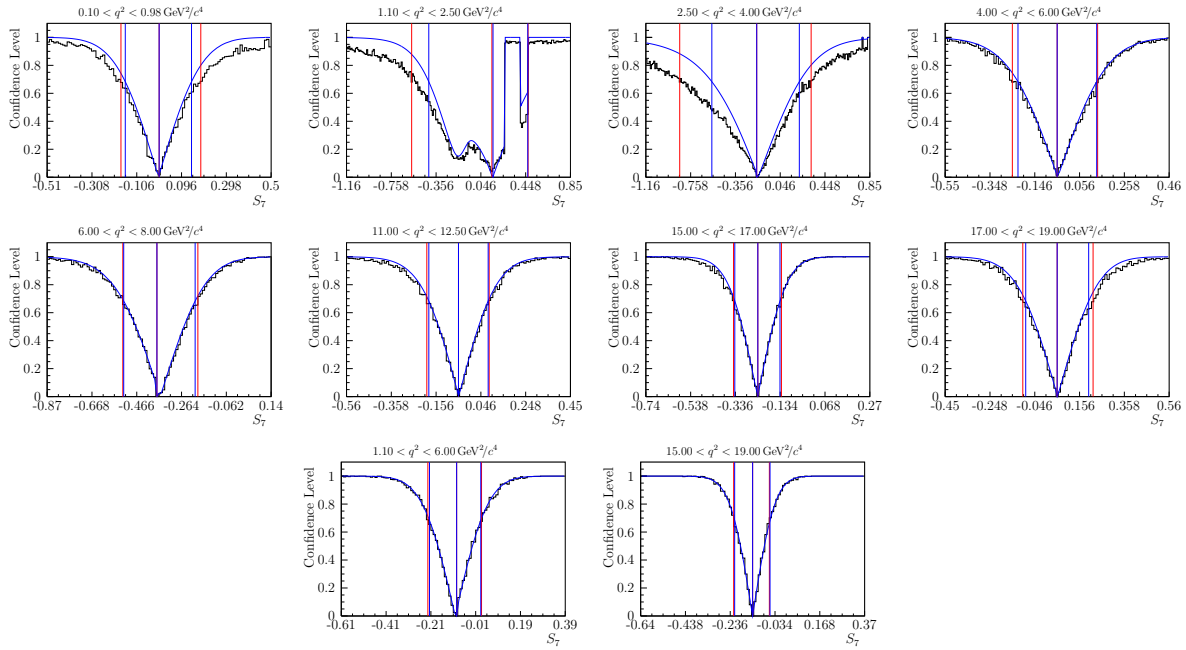

Figure A.20: Likelihood profiles Feldman-Cousins scans for the S_7 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

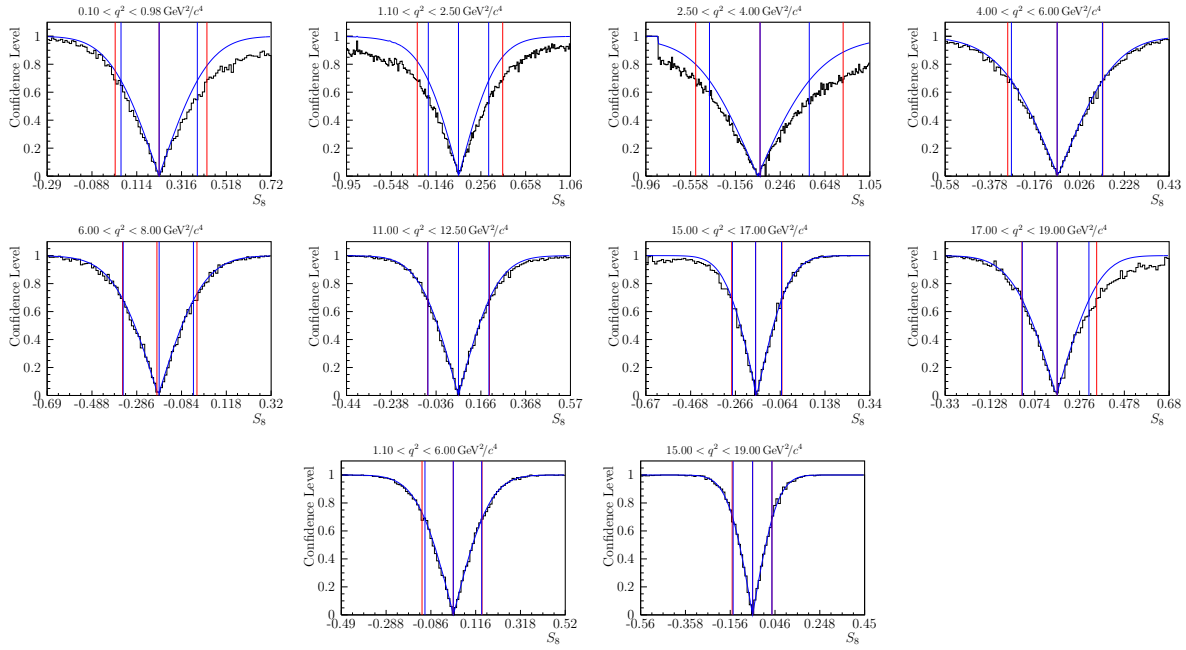

Figure A.21: Likelihood profiles Feldman-Cousins scans for the S_8 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

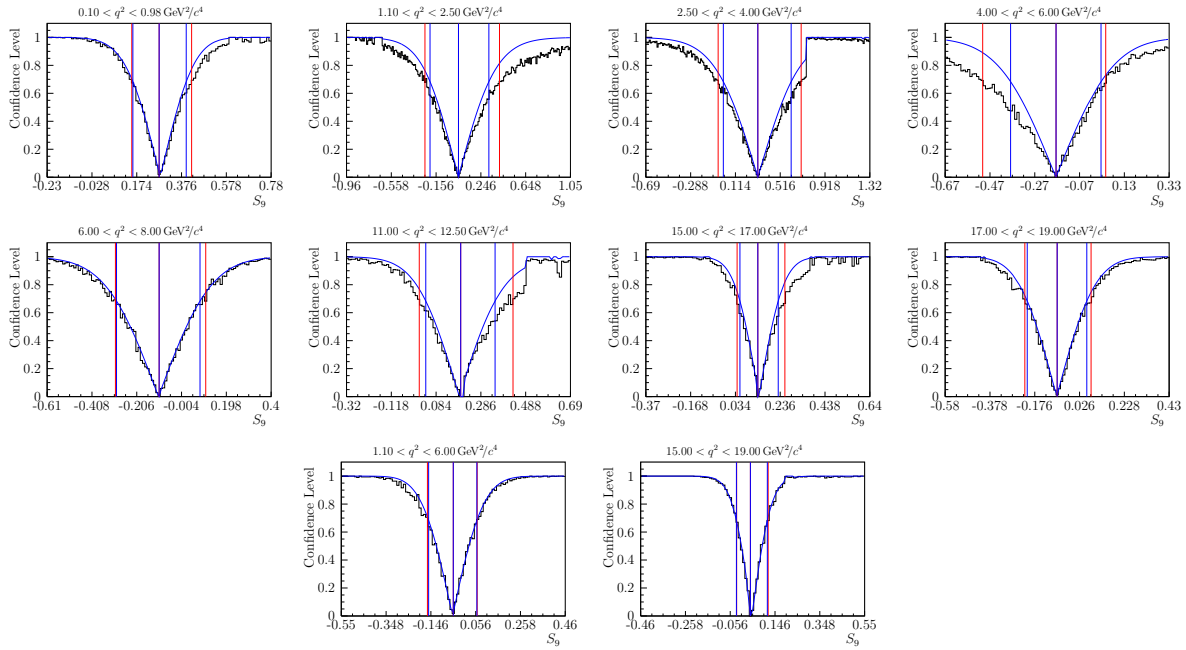

Figure A.22: Likelihood profiles Feldman-Cousins scans for the S_9 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the Δlog-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

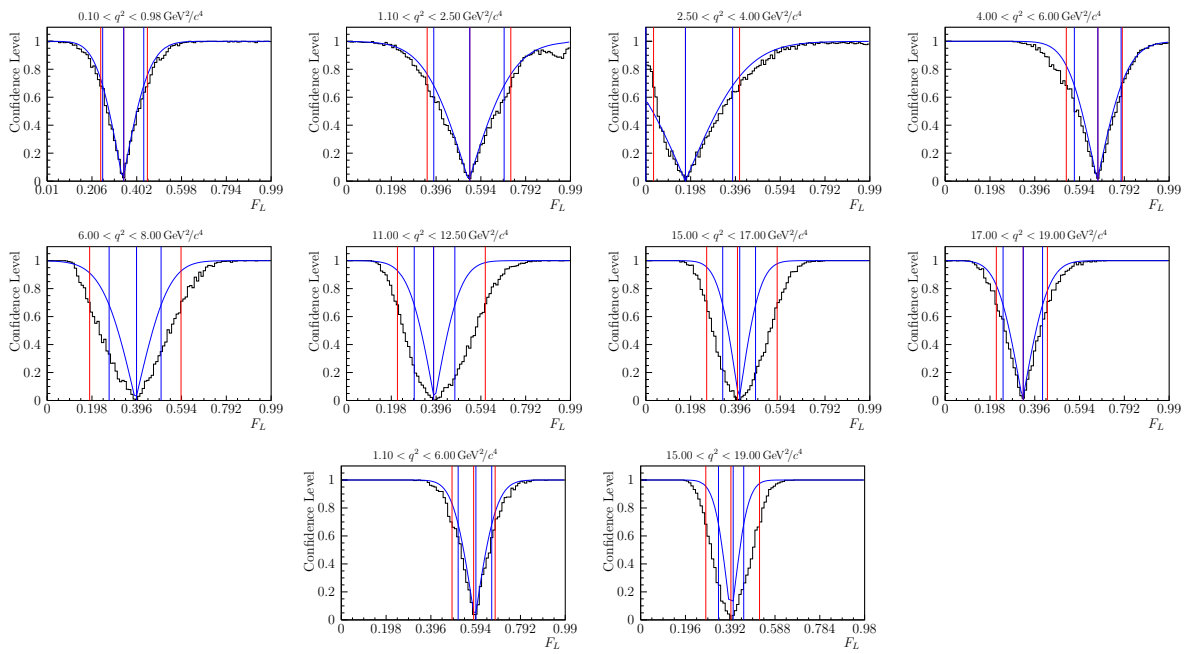

Figure A.23: Likelihood profiles Feldman-Cousins scans for the F_L observable in fits of the $P_i^{(\prime)}$ basis in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the Δlog-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

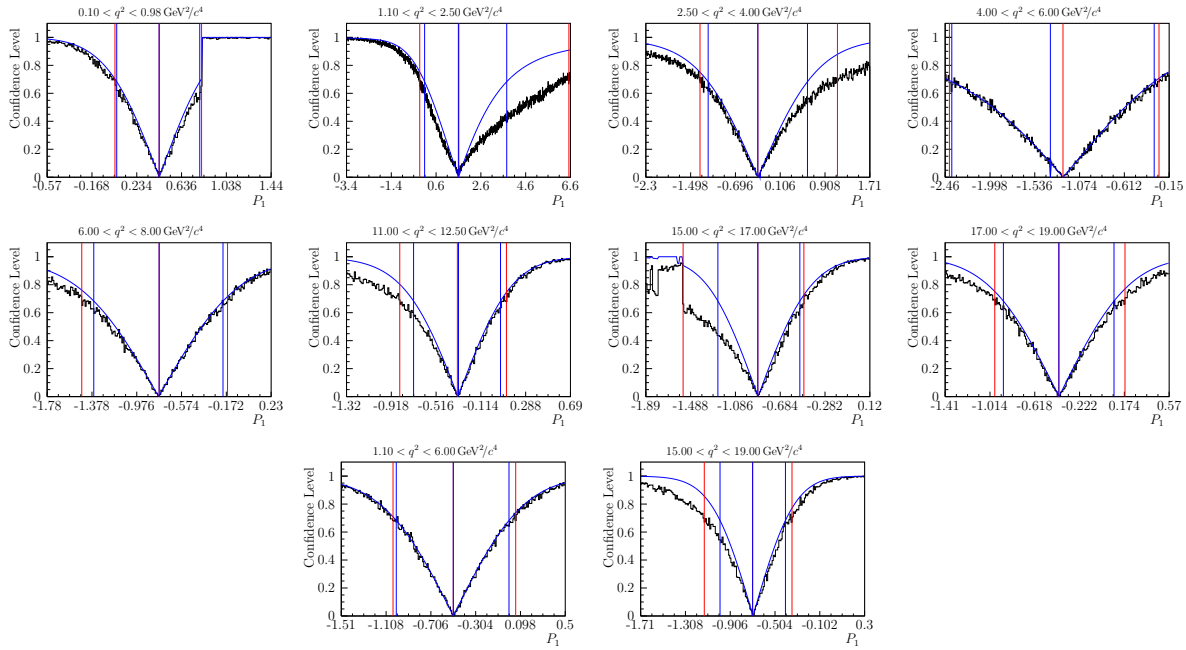

Figure A.24: Likelihood profiles Feldman-Cousins scans for the P_1 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

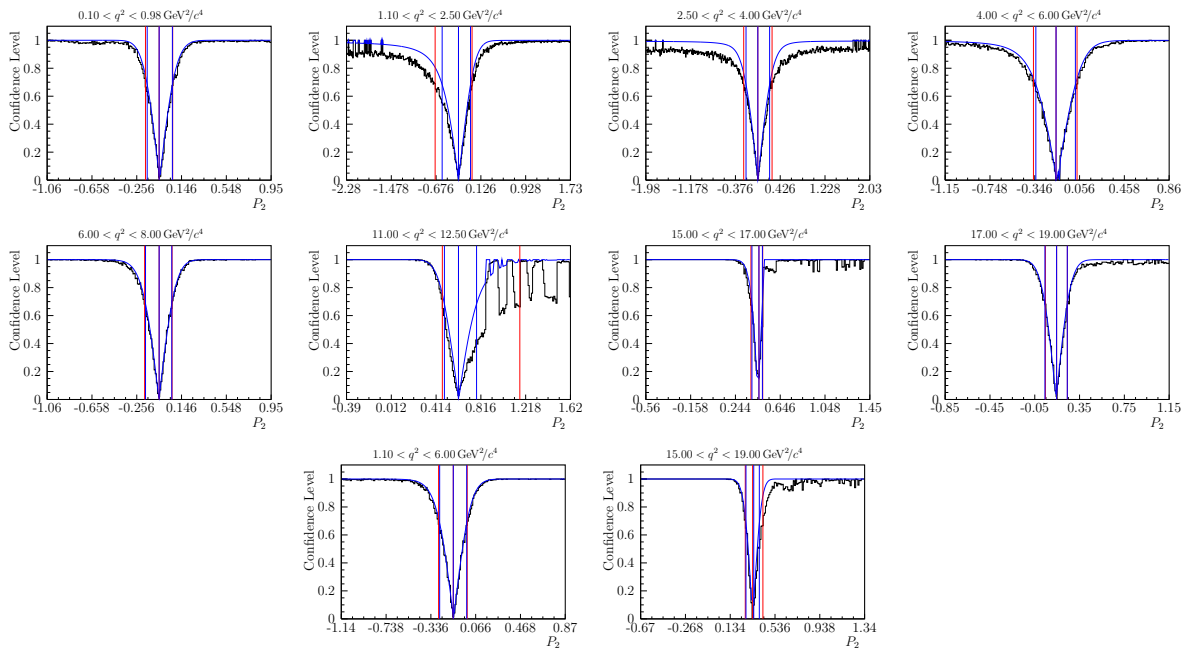

Figure A.25: Likelihood profiles Feldman-Cousins scans for the P_2 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

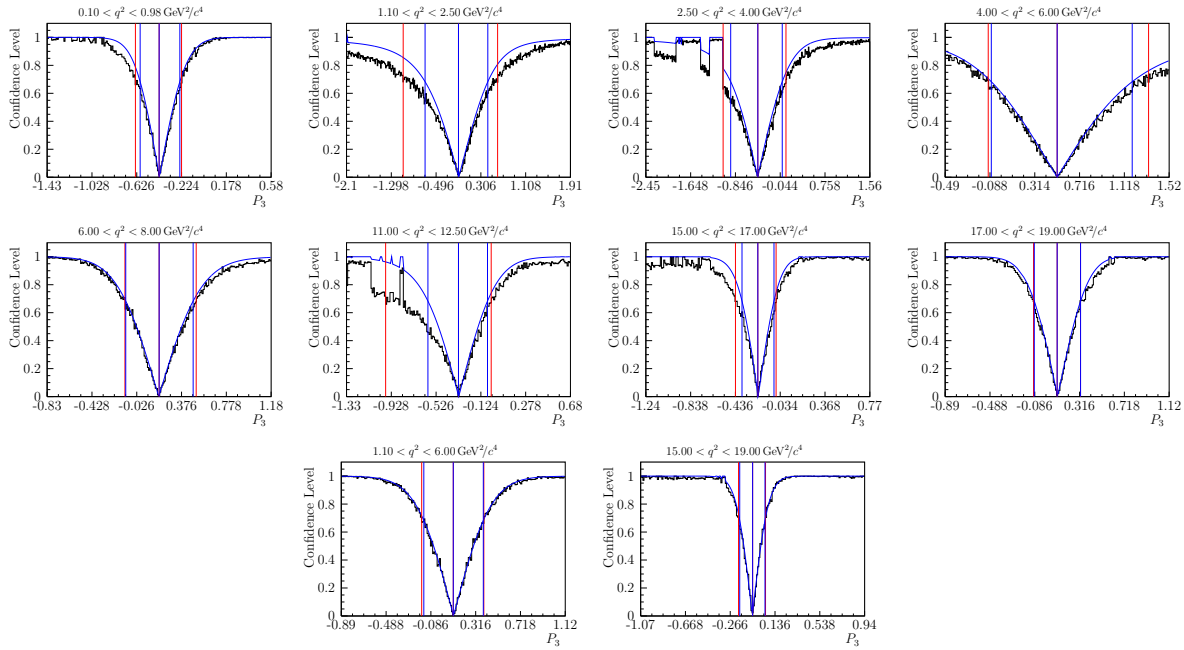

Figure A.26: Likelihood profiles Feldman-Cousins scans for the P_3 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

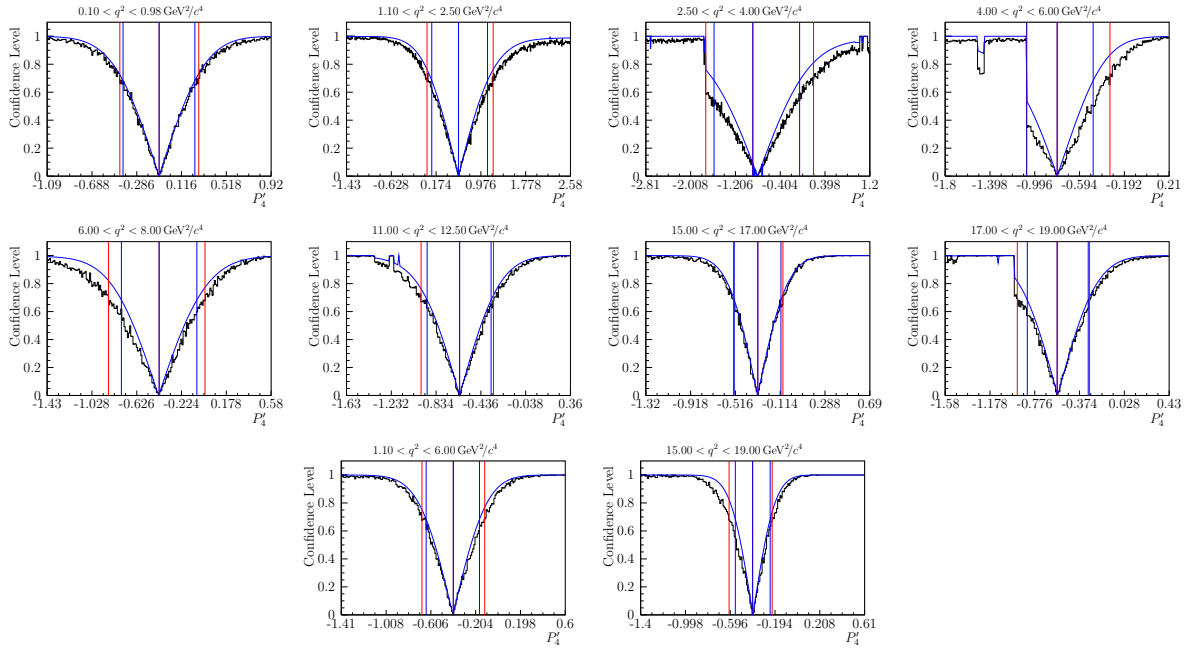

Figure A.27: Likelihood profiles Feldman-Cousins scans for the P_4 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

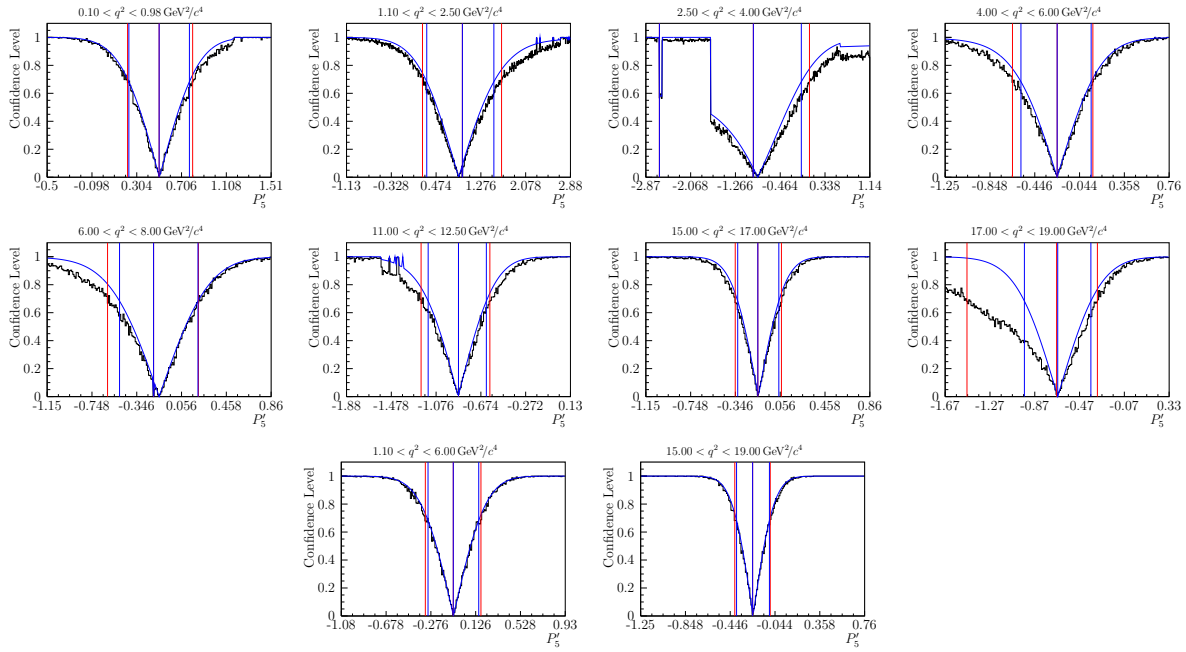

Figure A.28: Likelihood profiles Feldman-Cousins scans for the P_5 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

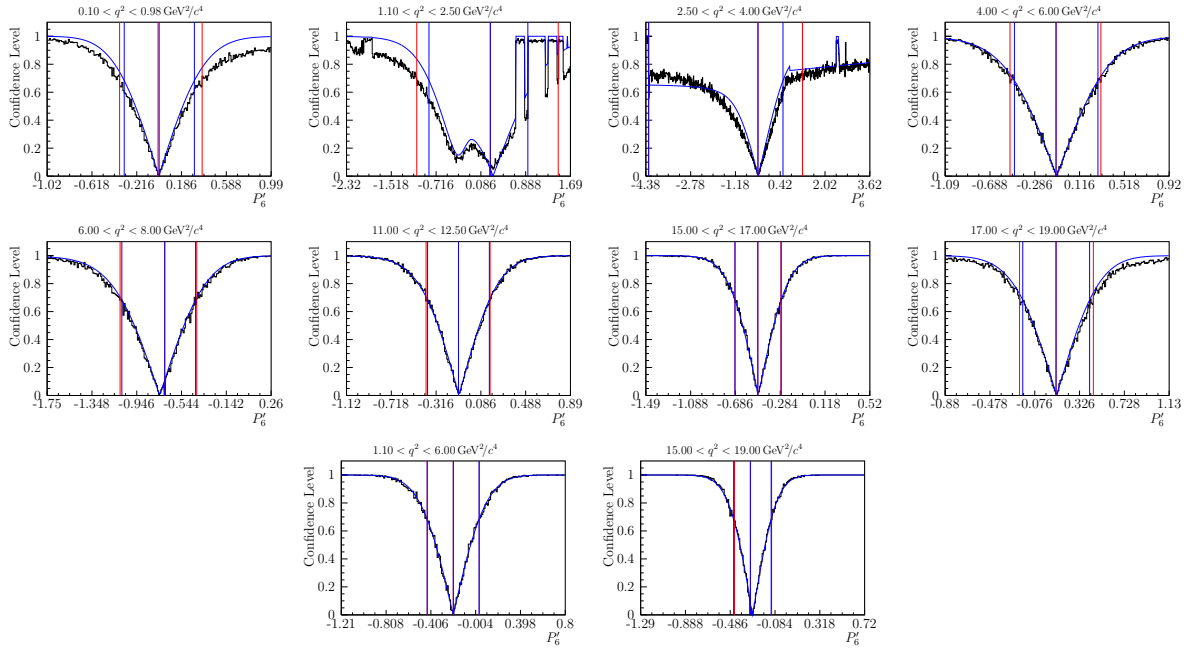

Figure A.29: Likelihood profiles Feldman-Cousins scans for the P_6 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the $\Delta\log$-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

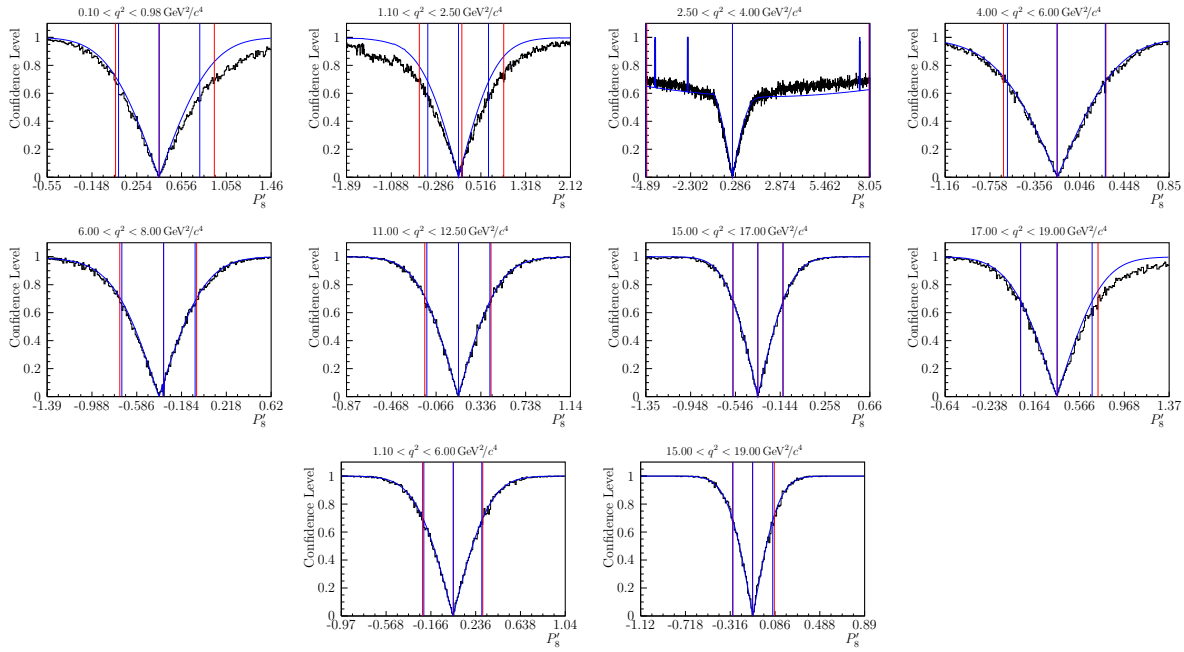

Figure A.30: Likelihood profiles Feldman-Cousins scans for the P_8 observable in ten q^2 intervals. The black histogram represents the results obtained by the Feldman-Cousins (FC) method and pseudoexperiments, as discussed in Section 5.8. In addition, the Δlog-likelihood profile is given in blue. The confidence intervals equivalent to 1σ at 68.3% are given in red (blue) for the FC ($\Delta\log\mathcal{L}$).

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay*, `arXiv:2012.13241`, submitted to Phys. Rev. Lett.

[2] D. Gerick, *Rare Decays at LHCb: recent results*, LHCb-TALK-2020-186, Implications of LHCb measurements and future prospects, CERN, Geneva (*virtual*), 2020.

[3] LHCb collaboration, R. Aaij *et al.*, *Supplementary Material to LHCb-PAPER-2020-041*, `https://cds.cern.ch/record/2748330`, 2020.

[4] M. De Cian, D. Gerick, R. Kopecna, and M. Borsato, *Angular analysis of the decay $B^+ \to K^{*+}\mu^+\mu^-$ with $K^{*+} \to K_s^0\pi^+$ using 9.1 fb^{-1}*, LHCb-ANA-2018-022, CERN, 2018.

[5] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, `arXiv:1207.7214`.

[6] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of* 125 GeV *with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, `arXiv:1207.7235`.

[7] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, `arXiv:1403.8044`.

[8] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, JHEP **06** (2015) 115, Erratum ibid. **09** (2018) 145, `arXiv:1503.07138`.

[9] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **09** (2015) 179, `arXiv:1506.08777`.

[10] LHCb collaboration, R. Aaij *et al.*, *Measurements of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction*, JHEP **11** (2016) 047, Erratum ibid. **04** (2017) 142, `arXiv:1606.04731`.

[11] W. Altmannshofer and D. M. Straub, *New physics in $b \to s$ transitions after LHC run 1*, Eur. Phys. J. **C75** (2015) 382, `arXiv:1411.3161`.

[12] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098, `arXiv:1503.05534`.

[13] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Lattice QCD calculation of form factors describing the rare decays $B \to K^*\ell^+\ell^-$ and $B_s \to \phi\ell^+\ell^-$*, Phys. Rev. **D89** (2014) 094501, `arXiv:1310.3722`.

[14] P. Ball and R. Zwicky, *$B_{(d,s)} \to \rho, \omega, K^*, \phi$ decay form-factors from light-cone sum rules revisited*, Phys. Rev. **D71** (2005) 014029, `arXiv:hep-ph/0412079`.

[15] A. Khodjamirian, T. Mannel, A. A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in $B \to K^{(*)}\ell^+\ell^-$ and $B \to K^*\gamma$*, JHEP **09** (2010) 089, `arXiv:1006.4945`.

[16] HPQCD collaboration, C. Bouchard *et al.*, *Rare decay $B \to K\ell^+\ell^-$ form factors from lattice QCD*, Phys. Rev. **D88** (2013) 054509, `arXiv:1306.2384`.

[17] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Calculation of $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ observables using form factors from lattice QCD*, Phys. Rev. Lett. **112** (2014) 212003, `arXiv:1310.3887`.

[18] HPQCD collaboration, C. Bouchard *et al.*, *Standard Model Predictions for $B \to K\ell^+\ell^-$ with Form Factors from Lattice QCD*, Phys. Rev. Lett. **111** (2013) 162002, `arXiv:1306.0434`.

[19] BaBar collaboration, J. P. Lees *et al.*, *Measurement of branching fractions and rate asymmetries in the rare decays $B \to K^{(*)}\ell^+\ell^-$*, Phys. Rev. **D86** (2012) 032012, `arXiv:1204.3933`.

[20] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, `arXiv:1705.05802`.

[21] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **122** (2019) 191801, `arXiv:1903.09252`.

[22] Belle collaboration, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ decays*, `arXiv:1908.01848`, submitted to JHEP.

[23] Belle collaboration, S. Wehle *et al.*, *Test of lepton-flavor universality in $B \to K^*\ell^+\ell^-$ decays at Belle*, `arXiv:1904.02440`, submitted to Phys. Rev. Lett.

[24] LHCb collaboration, R. Aaij *et al.*, *Measurement of form-factor-independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013) 191801, `arXiv:1308.1707`.

[25] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$*, JHEP **08** (2013) 131, `arXiv:1304.6325`.

[26] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,fb^{-1}$ of integrated luminosity*, JHEP **02** (2016) 104, `arXiv:1512.04442`.

[27] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-averaged observables in the $B^0 \to K^{*0}\mu^+\mu^-$ decay*, Phys. Rev. Lett. **125** (2020) 011802, `arXiv:2003.04831`.

[28] BaBar collaboration, B. Aubert *et al.*, *Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$*, Phys. Rev. **D73** (2006) 092001, `arXiv:hep-ex/0604007`.

[29] CDF collaboration, T. Aaltonen *et al.*, *Measurements of the angular distributions in the decays $B \to K^{(*)}\mu^+\mu^-$ at CDF*, Phys. Rev. Lett. **108** (2012) 081807, `arXiv:1108.0695`.

[30] Belle collaboration, S. Wehle *et al.*, *Lepton-flavor-dependent angular analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017) 111801, `arXiv:1612.05014`.

[31] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B781** (2018) 517, `arXiv:1710.02846`.

[32] ATLAS collaboration, M. Aaboud *et al.*, *Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **2018** (2018) 47, `arXiv:1805.04000`.

[33] LHCb collaboration, R. Aaij *et al.*, *Angular moments of the decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ at low hadronic recoil*, JHEP **09** (2018) 146, `arXiv:1808.00264`.

[34] CMS collaboration, A. M. Sirunyan *et al.*, *Angular analysis of the decay $B^+ \to K^*(892)^+\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, arXiv:2010.13968, Submitted to JHEP.

[35] E. Graverini, *Flavour anomalies: a review*, J. Phys. Conf. Ser. **1137** (2019) 012025, arXiv:1807.11373.

[36] G. Hiller and M. Schmaltz, *R_K and future $b \to s\ell\ell$ physics beyond the standard model opportunities*, Phys. Rev. **D90** (2014) 054014, arXiv:1408.1627.

[37] B. Gripaios, M. Nardecchia, and S. A. Renner, *Composite leptoquarks and anomalies in B-meson decays*, JHEP **05** (2015) 006, arXiv:1412.1791.

[38] I. de Medeiros Varzielas and G. Hiller, *Clues for flavor from rare lepton and quark decays*, JHEP **06** (2015) 072, arXiv:1503.01084.

[39] R. Barbieri, C. W. Murphy, and F. Senia, *B-decay anomalies in a composite leptoquark model*, Eur. Phys. J. **C77** (2017) 8, arXiv:1611.04930.

[40] A. Crivellin, D. Müller, and T. Ota, *Simultaneous explanation of $R(D^{(*)})$ and $b \to s\mu^+\mu^-$: the last scalar leptoquarks standing*, JHEP **09** (2017) 040, arXiv:1703.09226.

[41] J.-H. Sheng, R.-M. Wang, and Y.-D. Yang, *Scalar leptoquark effects in the lepton flavor violating exclusive $b \to s\ell_i^-\ell_j^+$ Decays*, Int. J. Theor. Phys. **58** (2019) 480, arXiv:1805.05059.

[42] G. Hiller, D. Loose, and I. Nišandžić, *Flavorful leptoquarks at hadron colliders*, Phys. Rev. **D97** (2018) 075004, arXiv:1801.09399.

[43] J. Aebischer *et al.*, *B-decay discrepancies after Moriond 2019*, Eur. Phys. J. **C80** (2020) 252, arXiv:1903.10434.

[44] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, Phys. Rev. **D89** (2014) 095033, arXiv:1403.1269.

[45] A. Celis, J. Fuentes-Martín, M. Jung, and H. Serôdio, *Family nonuniversal Z' models with protected flavor-changing interactions*, Phys. Rev. **D92** (2015) 015007, arXiv:1505.03079.

[46] A. Falkowski, M. Nardecchia, and R. Ziegler, *Lepton flavor non-universality in B-meson decays from a $U(2)$ flavor model*, JHEP **11** (2015) 173, arXiv:1509.01249.

[47] F. Sala and D. M. Straub, *A new light particle in B decays?*, Phys. Lett. **B774** (2017) 205, arXiv:1704.06188.

[48] P. Ko, Y. Omura, Y. Shigekami, and C. Yu, *LHCb anomaly and B physics in flavored Z' models with flavored Higgs doublets*, Phys. Rev. **D95** (2017) 115040, arXiv:1702.08666.

[49] M. Algueró *et al.*, *Emerging patterns of new physics with and without Lepton flavour universal contributions*, Eur. Phys. J. **C79** (2019) 714, arXiv:1903.09578.

[50] K. Kowalska, D. Kumar, and E. M. Sessolo, *Implications for new physics in $b \to s\mu\mu$ transitions after recent measurements by Belle and LHCb*, Eur. Phys. J. **C79** (2019) 840, arXiv:1903.10932.

[51] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Explaining $h \to \mu^\pm\tau^\mp$, $B \to K^*\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to Ke^+e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$*, Phys. Rev. Lett. **114** (2015) 151801, arXiv:1501.00993.

[52] T. Hurth, F. Mahmoudi, D. Martínez Santos, and S. Neshatpour, *Lepton nonuniversality in exclusive $b \to s\ell\ell$ decays*, Phys. Rev. **D96** (2017) 095034, arXiv:1705.06274.

[53] A. Arbey, T. Hurth, F. Mahmoudi, and S. Neshatpour, *Hadronic and new physics contributions to $b \to s$ transitions*, Phys. Rev. **D98** (2018) 095027, arXiv:1806.02791.

[54] A. Arbey *et al.*, *Update on the $b \to s$ anomalies*, Phys. Rev. **D100** (2019) 015045, arXiv:1904.08399.

[55] M. Ciuchini *et al.*, *New physics in $b \to s\ell^+\ell^-$ confronts new data on lepton universality*, Eur. Phys. J. **C79** (2019) 719, arXiv:1903.09632.

[56] A. K. Alok, A. Dighe, S. Gangal, and D. Kumar, *Continuing search for new physics in $b \to s\mu\mu$ decays: Two operators at a time*, JHEP **06** (2019) 089, arXiv:1903.09617.

[57] S. Jäger and J. Martin Camalich, *Reassessing the discovery potential of the $B \to K^*\ell^+\ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities*, Phys. Rev. **D93** (2016) 014028, arXiv:1412.3183.

[58] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$?*, arXiv:1406.0566.

[59] M. Ciuchini *et al.*, *$B \to K^*\ell^+\ell^-$ decays at large recoil in the Standard Model: a theoretical reappraisal*, JHEP **06** (2016) 116, arXiv:1512.07157.

[60] C. Bobeth, M. Chrzaszcz, D. van Dyk, and J. Virto, *Long-distance effects in $B \to K^*\ell\ell$ from analyticity*, Eur. Phys. J. **C78** (2018) 451, arXiv:1707.07305.

[61] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.

[62] E. Noether, *Invariant Variation Problems*, Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen (1918) 235, arXiv:physics/0503066, translated to English by M.A. Tavel in Transp. Theory Statist. Phys. **1** (1971) 186.

[63] ALEPH, DELPHI, L3, OPAL and SLD collaborations, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, Physics Reports **427** (2006) 257, arXiv:hep-ex/0509008.

[64] Super-Kamiokande collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.

[65] SNO collaboration, Q. R. Ahmad *et al.*, *Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv:nucl-ex/0204008.

[66] KATRIN collaboration, M. Aker *et al.*, *Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN*, Phys. Rev. Lett. **123** (2019) 221802, arXiv:1909.06048.

[67] J. I. Illana and T. Riemann, *Charged lepton flavor violation from massive neutrinos in Z decays*, Phys. Rev. **D63** (2001) 053004, arXiv:hep-ph/0010193.

[68] A. Pich and M. Ramsey-Musolf, *Tests of Conservation Laws*, http://pdg.lbl.gov/2019/reviews/rpp2019-rev-conservation-laws.pdf, 2019.

[69] P. A. M. Dirac, *The quantum theory of the electron*, Proc. R. Soc. Lond. A **117** (1928) 610.

[70] C. Burgard and D. Galbraith, *A standard diagram of the current standard model of physics*, http://www.texample.net/tikz/examples/model-physics/, 2016.

[71] Particle Data Group, M. Tanabashi *et al.*, *Review of Particle Physics*, Phys. Rev. **D98** (2018) 030001.

[72] J. Goldstone, A. Salam, and S. Weinberg, *Broken symmetries*, Phys. Rev. **127** (1962) 965.

[73] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[74] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.

[75] S. Weinberg, *Implications of dynamical symmetry breaking*, Phys. Rev. **D13** (1976) 974.

[76] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.

[77] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014) 222002, `arXiv:1404.1903`.

[78] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.03414`.

[79] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$*, Phys. Rev. Lett. **122** (2019) 222001, `arXiv:1904.03947`.

[80] LHCb collaboration, R. Aaij *et al.*, *Observation of structure in the J/ψ-pair mass spectrum*, Science Bulletin **65** (2020) 1983, `arXiv:2006.16957`.

[81] LHCb collaboration, R. Aaij *et al.*, *Model-independent study of structure in $B^+ \to D^+D^-K^+$ decays*, Phys. Rev. Lett. **125** (2020) 242001, `arXiv:2009.00025`.

[82] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of the $B^+ \to D^+D^-K^+$ decay*, Phys. Rev. **D102** (2020) 112003, `arXiv:2009.00026`.

[83] R. P. Feynman, *Space-time approach to quantum electrodynamics*, Phys. Rev. **76** (1949) 769.

[84] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[85] M. Goldhaber, L. Grodzins, and A. W. Sunyar, *Helicity of neutrinos*, Phys. Rev. **109** (1958) 1015.

[86] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[87] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[88] CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, `arXiv:hep-ph/0406184`.

[89] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. **D2** (1970) 1285.

[90] K. G. Begeman, A. H. Broeils, and R. H. Sanders, *Extended rotation curves of spiral galaxies: dark haloes and modified dynamics*, Monthly Notices of the Royal Astronomical Society **249** (1991) 523.

[91] Planck collaboration, N. Aghanim *et al.*, *Planck 2018 results - VI. Cosmological parameters*, A&A **641** (2020) A6, `arXiv:1807.06209`.

[92] W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, Zeitschrift für Physik **43** (1927) 172.

[93] P. Ball, G. W. Jones, and R. Zwicky, *$B \to V\gamma$ beyond QCD factorization*, Phys. Rev. **D75** (2007) 054004, `arXiv:hep-ph/0612081v3`.

[94] E. Fermi, *Tentativo di una Teoria dell'emissione dei raggi β*, La Ricerca Scientifica **4** (1933) 491.

[95] E. Fermi, *Versuch einer Theorie der β-Strahlen. I*, Zeitschrift für Physik **88** (1934) 161.

[96] S. Weinberg, *The Quantum Theory of Fields*, vol. 2, Cambridge University Press, 1996.

[97] K. G. Wilson, *Quantum field - theory models in less than 4 dimensions*, Phys. Rev. **D7** (1973) 2911.

[98] W. Altmannshofer *et al.*, *Symmetries and asymmetries of $B \to K^* \mu^+ \mu^-$ decays in the Standard Model and beyond*, JHEP **01** (2009) 019, `arXiv:0811.1214`.

[99] M. S. Lie, *Die Grundlagen für die Theorie der unendlichen kontinuierlichen Transformationsgruppen. I. Treatise*, Leipz. Ber., 1891.

[100] J. Aebischer, J. Kumar, and D. M. Straub, *Wilson: a Python package for the running and matching of Wilson coefficients above and below the electroweak scale*, Eur. Phys. J **C78** (2018) 1026, `arXiv:1804.05033`.

[101] D. M. Straub, *Flavio: A python package for flavour and precision phenomenology in the Standard Model and beyond*, `arXiv:1810.08132`.

[102] J. M. Arnold, B. Fornal, and M. B. Wise, *Simplified models with baryon number violation but no proton decay*, Phys. Rev. **D87** (2013) , `arXiv:1212.4556`.

[103] D. M. Straub, *Flavour anomalies: status and implications*, 2018. Seminar talk at the Max-Planck-Institut für Kernphysik, Heidelberg, `https://www.mpi-hd.mpg.de/lin/seminar_theory/talks/Talk_Straub_050218.pdf`.

[104] B. Fornal, S. A. Gadam, and B. Grinstein, *Left-right SU(4) vector leptoquark model for flavor anomalies*, Phys. Rev. **D99** (2019) 055025, `arXiv:1812.01603`.

[105] L. Di Luzio, A. Greljo, and M. Nardecchia, *Gauge leptoquark as the origin of B-physics anomalies*, Phys. Rev. **D96** (2017) 115011, `arXiv:1708.08450`.

[106] N. Assad, B. Fornal, and B. Grinstein, *Baryon number and lepton universality violation in leptoquark and diquark models*, Phys. Lett. **B777** (2018) 324, `arXiv:1708.06350`.

[107] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, `arXiv:1406.6482`.

[108] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \to pK^- \ell^+ \ell^-$ decays*, JHEP **05** (2020) 040, `arXiv:1912.08139`.

[109] LHCb collaboration, R. Aaij *et al.*, *Measurement of the B_s^0–$\overline{B}_s^0$ oscillation frequency Δm_s in $B_s^0 \to D_s^-(3)\pi$ decays*, Phys. Lett. **B709** (2012) 177, `arXiv:1112.4311`.

[110] LHCb collaboration, R. Aaij *et al.*, *Measurement of the B^0–$\overline{B}^0$ oscillation frequency Δm_d with the decays $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi K^{*0}$*, Phys. Lett. **B719** (2013) 318, `arXiv:1210.6750`.

[111] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP violating phase and decay-width difference in $B_s^0 \to \psi(2S)\phi$ decays*, Phys. Lett. **B762** (2016) 253, `arXiv:1608.04855`.

[112] LHCb collaboration, R. Aaij *et al.*, *Resonances and CP-violation in B_s^0 and $\overline{B}_s^0 \to J/\psi K^+ K^-$ decays in the mass region above the $\phi(1020)$*, JHEP **08** (2017) 037, `arXiv:1704.08217`.

[113] HFLAV, Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, `arXiv:1909.12524`, updated results and plots available at `https://hflav.web.cern.ch/`.

[114] P. Langacker and M. Plümacher, *Flavor changing effects in theories with a heavy Z' boson with family nonuniversal couplings*, Phys. Rev. **D62** (2000) 013006.

[115] S. Sun, D. B. Kaplan, and A. E. Nelson, *Little flavor: A model of weak-scale flavor physics*, Phys. Rev. **D87** (2013) 125036.

[116] P. J. Fox, J. Liu, D. Tucker-Smith, and N. Weiner, *An effective Z'*, Phys. Rev. **D84** (2011) 115006.

[117] W. Altmannshofer and D. M. Straub, *New physics in $B \to K^*\mu\mu$?*, Eur. Phys. J. **C73** (2013) 2646, `arXiv:1308.1501`.

[118] D. Becirevic and A. Tayduganov, *Impact of $B \to K_0^*\ell^+\ell^-$ on the New Physics search in $B \to K^*\ell^+\ell^-$ decay*, Nucl. Phys. **B868** (2013) 368, `arXiv:1207.4004`.

[119] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, *Optimizing the basis of $B \to K^*\ell^+\ell^-$ observables in the full kinematic range*, JHEP **05** (2013) 137, `arXiv:1303.5794`.

[120] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[121] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[122] O. S. Brüning *et al.*, *LHC Design Report*, CERN-2004-003, CERN, Geneva, 2004. doi: 10.5170/CERN-2004-003-V-1.

[123] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.

[124] J. B. Adams, *The CERN Proton Synchrotron*, Nature **185** (1960) 568.

[125] CERN, *The* 300 GeV *programme*, CERN-1050-E, CERN, Geneva, 1972.

[126] J. B. Adams, *A design of the European* 300 GeV *research facilities*, MC-70-60; CERN-SPC-299, CERN, Geneva, 1970.

[127] J. B. Adams, *The CERN* 400 GeV *Proton Synchrotron (CERN SPS)*, HEACC 77, pp. 17-29, CERN, Protvino, 1977.

[128] R. Alemany-Fernandez *et al.*, *Operation and Configuration of the LHC in Run 1*, CERN-ACC-NOTE-2013-0041, CERN, ATS Department, 2013.

[129] J. Wenninger, *Operation and Configuration of the LHC in Run 2*, CERN-ACC-NOTE-2019-0007, CERN, ATS Department, 2019.

[130] A. Horvath, *The LHC experiments (Large Hadron Collider) and the preaccelerators*, LHC_02.svg, Wikimedia commons, 2012.

[131] J. D. Bjorken, *Asymptotic sum rules at infinite momentum*, Phys. Rev. **179** (1969) 1547.

[132] H. Dijkstra, H. J. Hilke, T. Nakada, and T. Ypsilantis, *LHCb Letter of Intent, LHCb collaboration*, LHCb-95-001, CERN, Geneva, 1995.

[133] R. Lindner, *LHCb layout and schema*, LHCb-PHO-GENE-2008-002, CERN, 2008. LHCb Collection.

[134] D. Mitzel, *Search for new physics in rare four-body charm decays at LHCb*, PhD thesis, Universität Heidelberg, Heidelberg, 2019, CERN-THESIS-2018-376, doi: 10.11588/heidok.00025790.

[135] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011, 2001.

[136] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[137] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030, 2003.

[138] C. Abellan Beteta *et al.*, *Monitoring radiation damage in the LHCb Tracker Turicensis*, `arXiv:1809.05063`.

[139] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, 2002.

[140] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, 2001.

[141] The LHCb Outer Tracker group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, `arXiv:1311.3893`.

[142] The LHCb Outer Tracker group, P. d'Argent *et al.*, *Improved performance of the LHCb outer tracker in LHC run 2*, JINST **12** (2017) P11016, `arXiv:1708.00819`.

[143] B. Storaci, *Optimization of the LHCb track reconstruction*, J. Phys. Conf. Ser. **664** (2015) 072047. 6 p LHCb-PROC-2015-012.

[144] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007, 2000.

[145] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2018) 1, `arXiv:1803.00824`.

[146] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037, 2000.

[147] A. Papanestis and C. D'Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, Nucl. Instrum. Meth. **A876** (2017) 221.

[148] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[149] P. A. Cherenkov, *Visible Emission of Clean Liquids by Action of γ Radiation*, 1934.

[150] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, 2000.

[151] R. Aaij *et al.*, *Performance of the LHCb calorimeters*, LHCb-DP-2013-004, in preparation.

[152] P. Ghez and S. T'Jampens, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, LHCb-DP-2020-001.

[153] G. Charpak *et al.*, *The Use of Multiwire Proportional Counters to Select and Localize Charged Particles*, Nucl. Instrum. Meth. **62** (1968) 262.

[154] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010, 2001.

[155] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[156] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[157] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031, 2003.

[158] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[159] A. Puig, *The LHCb trigger in 2011 and 2012*, LHCb-PUB-2014-046, CERN, Geneva, 2014.

[160] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, `arXiv:1812.10790`.

[161] A. Martin Sanchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 Calorimeter Trigger*, LHCb-PUB-2011-026, CERN, Geneva, 2012.

[162] J. Albrecht, L. de Paula, A. Pérez-Calero Yzquierdo, and F. Teubert, *Commissioning and Performance of the LHCb HLT1 muon trigger*, LHCb-PUB-2011-006, CERN, Geneva, 2011.

[163] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[164] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[165] T. Sjöstrand, S. Mrenna, and P. Skands, *Pythia 6.4 physics and manual*, JHEP **2006** (2006) 026–026.

[166] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

[167] T. Sjöstrand *et al.*, *An introduction to Pythia 8.2*, Comput. Phys. Commun. **191** (2015) 159–177.

[168] A. Ryd *et al.*, *EvtGen: A Monte Carlo Generator for B-Physics*, EVTGEN-V00-11-07.

[169] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[170] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[171] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4—a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[172] P. Koppenburg, *Reconstruction and analysis software environment of LHCb*, Nucl. Phys. **B156** (2006) 213, Proceedings of the 10th International Conference on B-physics at Hadron Machines.

[173] V. V. Gligorov, *Reconstruction of the Channel $B_d^0 \to D^+\pi^-$ and Background Classification at LHCb (revised)*, LHCb-2007-044, CERN, Geneva, 2007. revised version submitted on 2008-01-24 12:46:44.

[174] F. Krauss, S. Navas, P. Richardson, and T. Sjöstrand, *43. Monte Carlo particle numbering scheme*, Monte Carlo Numbering.

[175] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356–369, `arXiv:physics/0402083v3`.

[176] LHCb collaboration, R. Aaij *et al.*, *Measurement of the branching fraction and CP asymmetry in $B^+ \to J/\psi\,\rho^+$ decays*, Eur. Phys. J. **C79** (2019) 537, `arXiv:1812.07041`.

[177] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, EPJ Tech. Instrum. **6** (2019) 1, `arXiv:1803.00824`.

[178] L. Anderlini *et al.*, *Computing strategy for PID calibration samples for LHCb Run 2*, LHCb-PUB-2016-020, CERN, Geneva, 2016.

[179] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, `arXiv:physics/0503191`.

[180] 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.

[181] BaBar collaboration, B. Aubert *et al.*, *Measurement of Branching Fractions and Charge Asymmetries for Exclusive B Decays to Charmonium*, Phys. Rev. Lett. **94** (2005) 141801, `arXiv:hep-ex/0412062`.

[182] BES collaboration, J. Z. Bai *et al.*, *A measurement of J/ψ decay widths*, Phys. Lett. **B355** (1995) 374.

[183] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, `arXiv:physics/0408124`.

[184] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[185] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[186] F. Rademakers *et al.*, *root-project/root: v6.20/06*, 2020. doi: 10.5281/zenodo.3895852.

[187] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[188] S. Stemmle, *CP violation and lifetime measurements in the decay $B_s^0 \to J/\psi\,\phi$ with the LHCb experiment*, PhD thesis, Universität Heidelberg, Heidelberg, 2019, doi: 10.11588/heidok.00026861.

[189] G. Cowan, *Statistical data analysis*, Oxford science publications, Clarendon Press, Oxford, 1998.

[190] S. Valavanis, *Econometrics: an introduction to maximum likelihood methods*, Economics handbook series, McGraw-Hill, New York, 1959.

[191] T. Bayes, *An essay towards solving a problem in the doctrine of chances*, Phil. Trans. R. Soc. **53** (1763) 370, letter to John Canton.

[192] F. James and M. Roos, *Minuit – a system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun. **10** (1975) 343.

[193] G. Breit and E. Wigner, *Capture of slow neutrons*, Phys. Rev. **49** (1936) 519.

[194] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics*, Springer, New York, 1952.

[195] K. Chilikin *et al.*, *Experimental constraints on the spin and parity of the $Z(4430)^+$*, Phys. Rev. **D88** (2013) .

[196] LHCb collaboration, R. Aaij *et al.*, *Quantum numbers of the $X(3872)$ state and orbital angular momentum in its $\rho^0 J/\psi$ decays*, Phys. Rev. **D92** (2015) 011102(R), `arXiv:1504.06339`.

[197] D. Aston *et al.*, *A Study of $K^-\pi^+$ scattering in the reaction $K^-\pi^+ \to K^-\pi^+ n$ at* 11 GeV/c, Nucl. Phys. **B296** (1988) 493.

[198] B. Dunwoodie, *Fits to $K\pi$ I=1/2 S-wave Amplitude and Phase Data*, kpi_swave_fit.note, SLAC, 2013.

[199] B. Efron, *Bootstrap Methods: Another Look at the Jackknife*, Ann. Statist. **7** (1979) 1.

[200] T. Nikodem, *Angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ decays using $3\,fb^{-1}$ of integrated luminosity at LHCb*, PhD thesis, Universität Heidelberg, Heidelberg, 2016, doi: 10.11588/heidok.00020274.

[201] L. Demortier and L. Lyons, *Everything you always wanted to know about pulls*, CDF/ANAL/PUBLIC/5776, CERN, 2008. version 3.00.

[202] W. Eadie *et al.*, *Statistical methods in experimental physics*, JASA **68** (1973) 494.

[203] BaBar collaboration, B. Aubert *et al.*, *Measurement of decay amplitudes of $B \to J/\psi K^*, \psi(2S)K^*$, and $\chi_{c1}K^*$ with an angular analysis*, Phys. Rev. **D76** (2007) 031102, `arXiv:0704.0522`.

[204] Belle collaboration, R. Itoh *et al.*, *Studies of CP violation in $B \to J/\psi K^*$ decays*, Phys. Rev. Lett. **95** (2005) 091601, `arXiv:hep-ex/0504030`.

[205] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes in $B^0 \to J/\psi\, K^*(892)^0$ decays*, Phys. Rev. **D88** (2013) 052002, `arXiv:1307.2782`.

[206] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873, `arXiv:physics/9711021`.

[207] J. Neyman, *Outline of a theory of statistical estimation based on the classical theory of probability*, Phil. Trans. R. Soc. **A236** (1937) 333.

[208] E. Gross, *LHC Statistics for Pedestrians*, in *PHYS-STAT Workshop on Statistical Issues for LHC Physics*, No. CERN-2008-001 in PHYSTAT LHC, 205, CERN, 2008.

[209] F. Beaujean, M. Chrząszcz, N. Serra, and D. van Dyk, *Extracting angular observables without a likelihood and applications to rare decays*, Phys. Rev. **D91** (2015) 114012, `arXiv:1503.04100`.

[210] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Rare B decays using lattice QCD form factors*, PoS (2015) 372, `arXiv:1501.00367`.

[211] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *Global analysis of $b \to s\ell\ell$ anomalies*, JHEP **06** (2016) 92, `arXiv:1510.04239`.

[212] B. Capdevila *et al.*, *Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data*, JHEP **01** (2018) 093, `arXiv:1704.05340`.

[213] LHCb collaboration, R. Aaij *et al.*, *Strong constraints on the $b \to s\gamma$ photon polarisation from $B^0 \to K^{*0}e^+e^-$ decays*, JHEP **12** (2020) 081, `arXiv:2010.06011`, to appear in JHEP.

[214] B. Grinstein and D. Pirjol, *Exclusive rare $B \to K^*\ell^+\ell^-$ decays at low recoil: Controlling the long-distance effects*, Phys. Rev. **D70** (2004) 114005, `arXiv:hep-ph/0404250`.

[215] M. Beylich, G. Buchalla, and T. Feldmann, *Theory of $B \to K^{(*)}\ell^+\ell^-$ decays at high q^2: OPE and quark-hadron duality*, Eur. Phys. J. **C71** (2011) 1635, `arXiv:1101.5118`.

[216] S. Braß, G. Hiller, and I. Nisandzic, *Zooming in on $B \to K^*\ell\ell$ decays at low recoil*, Eur. Phys. J. **C77** (2017) 16, `arXiv:1606.00775`.

[217] M. Ciuchini *et al.*, *Lessons from the $B^{0,+} \to K^{*0,+}\mu^+\mu^-$ angular analyses*, Phys. Rev. **D103** (2021) 015030, `arXiv:2011.01212`.

[218] T. Hurth, F. Mahmoudi, and S. Neshatpour, *Model independent analysis of the angular observables in $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^{*+}\mu^+\mu^-$*, `arXiv:2012.12207`.

[219] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007, 2012.

Acknowledgements

The work presented in this thesis would have not been possible without the help and support from many people over the last six years. I am very thankful to everyone who participated or otherwise contributed to my PhD project and cannot thank you enough.
First of all I want to thank Prof. Dr. Stephanie Hansmann-Menzemer for the possibility to do my PhD studies in the LHCb group of the Physikalisches Institut in Heidelberg. Thank you for your support and the continuous advise over the years – including my Master's studies. I would also like to express my gratitude for providing the possibilities of conducting experiments and attending meetings at CERN; and for participating in summer schools, conferences and workshops all around the world. I want to thank Prof. Dr. Klaus Reygers as the second referee to this thesis, along Prof. Dr. Susanne Westhoff and Prof. Dr. Ulrich Platt for their examination of my thesis work.
Thanks to all members of the Heidelberg LHCb group, in particular Prof. Dr. Ulrich Uwer, for the great working environment and positive atmosphere. A big impact on my analysis had of course Dr. Michel De Cian and Dr. Martino Borsato. Thank you two for the long term advise, the patience with bug finding and seemingly never ending needs for fit improvements, and for making it possible to run this great analysis on the limited statistics of data that we have in our decay channel. Without your supervision, I would have never succeeded in getting these fantastic results out. Amazing to have you both available at almost any time, especially when considering that Michel spent the last years already in his next job in Lausanne. Thanks to Renata Kopečná for taking over the even more challenging decay chain of the $B^+ \to K^{*+}\mu^+\mu^-$ channel. Amazing what kind of improvements you managed to implement into your analysis. I am grateful for all the tiding up after me and the shiny formatting of plots that you added to our shared code repository.
And then we have all the other important people that did not directly contributed to the analysis, but helped me through the studying time in the last decade. To my fellow PhD graduates, Dr. Dominik Mitzel, Dr. Matthieu Kecke, Dr. Philippe d'Argent und Dr. Sebastian Ohmer, thanks for the long companion through the undergraduate and PhD studies. I have learned a lot from you and actually found the way to particle physics only because of you. I am glad to call you my friends, that we spent and spend holidays

together and I am sure to always meet you again.
Thanks to Dr. Xiaoxue Han and Dr. Blake Leverington for all the help during my studies of silicon photomultipliers during my Master thesis and my early PhD studies. The test-beams with you, and also with Dominik, Matthieu, Max, Axel and Olivier, have been a lot of fun - even the night shifts inside the dusty and noisy containers at the SPS test-beam facilities.
I also want to thank all ex- and non-physics friends from Heidelberg that I got to know over the years. Especially Andi, Lars, Felix and Timo for all the parties, wine- and beer-festivals and hikings.
Cheers to all friends from Cologne and surroundings, who I did not see nearly as much as I wanted over the lasts years, but who are still looking forward to my return to the Rhineland someday.
I am very thankful for my supportive parents, Christoph and Sabine, and siblings, Felix and Pia, and grandmother, Margret. Thank you for always giving me the freedom to do what I like and always believing in me. Also, my younger brother, Dr. Felix Gerick, who started his PhD studies over two and a half years after me, just so he could finish before me.
Today, I am in the fantastic situation of having a small family for my own. Thank you Jane (Dr. Christiane Gerick, to be consistent) for being my wonderful wife, for always being there for me and for always making me smile. I love you so much. And last, Marlon, thank you for being the cutest kid in the world.
Sorry to everybody that should be on this list but who I forgot in the rush of writing a thesis.