Dissertation

submitted to the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of Heidelberg University, Germany

for the degree of Doctor of Natural Sciences

Put forward by

Renata Kopečná

Born: The Czech Republic

Oral examination: 15th of December 2021

Angular analysis of the decay $\mathbf{B^+ \to \left(K^{*+} \to K^+\pi^0\right)\mu^+\mu^-}$ with the LHCb experiment

Referees: Prof. Dr. Stephanie Hansmann-Menzemer
Prof. Dr. Klaus Reygers

Title: **Angular analysis of the decay $B^+ \to (K^{*+} \to K^+ \pi^0)\,\mu^+\mu^-$ with the LHCb experiment**

Author: Renata Kopečná

Supervisor: Prof. Dr. Stephanie Hansmann-Menzemer

Abstract:

In this thesis, major steps towards the first angular analysis of the rare $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ decay are presented. This study is of great interest in the context of validating the anomalies observed in recent years in the properties of b quark decays to an s quark and two muons. If confirmed, these observations could represent a sign of physics beyond the Standard Model. The analysis is based on the world-largest sample of B meson decays which was collected using the LHCb experiment at the Large Hadron Collider during the years 2011 - 2018, corresponding to an integrated luminosity of $9.0\,\mathrm{fb}^{-1}$ of pp collision data. Although less than one in a million B^+ mesons decays to $K^{*+}\mu^+\mu^-$, a total of 271±28 candidates is selected.

The main challenge of the decay mode is the presence of a neutral π^0 meson in the final state. This thesis represents the first proof of feasibility of an angular analysis involving a neutral particle in the final state using the LHCb dataset. The event selection is optimized in order to remove various sources of background pollution while maximizing the available angular acceptance. A detailed simulation of the detector is used to unfold the effects of the detector acceptance, reconstruction and selection on the angular distributions. The maximum likelihood fit to extract the angular parameters of the decay $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ has been developed and tested using thousands of pseudoexperiments. It is shown that the complicated background structure together with the low statistical power of the available data sample result in large uncertainties. However, the analysis proves to be sensitive enough to disentangle an anomaly coherent with that observed in other decays with more than two standard deviations of significance.

Keywords: LHCb, FCNC, Rare decay, Electroweak penguin, Flavor physics, Standard Model, New Physics

Titel: **Winkelanalyse des Zerfalls $\mathbf{B^+ \to (K^{*+} \to K^+\pi^0)\,\mu^+\mu^-}$ mit dem LHCb Experiment**

Autorin: Renata Kopečná

Betreuerin: Prof. Dr. Stephanie Hansmann-Menzemer

Zusammenfassung:

In dieser Arbeit werden wichtige Schritte zur ersten Winkelanalyse des seltenen $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$-Zerfalls vorgestellt. Diese Studie ist von besonderem Interesse im Zusammenhang mit der Validierung der in den letzten Jahren beobachteten Anomalien in den Eigenschaften von b-Quark-Zerfällen zu einem s-Quark und zwei Myonen. Sollten sich diese Beobachtungen bestätigen, könnten sie ein Hinweis auf eine Physik jenseits des Standardmodells sein. Die Analyse basiert auf der weltweit größten Datensatz von B Meson-Zerfällen, die mit dem LHCb-Experiment am Large Hadron Collider in den Jahren 2011 bis 2018 gesammelt wurde, was einer integrierten Luminosität von $9{,}0\,\mathrm{fb}^{-1}$ von pp-Kollisionen entspricht. Obwohl weniger als eines von einer Million B^+ Mesonen zu $K^{*+}\mu^+\mu^-$ zerfällt, wurden insgesamt 271±28 Kandidaten selektiert.

Größte Herausforderung bei diesem Zerfallsmodus ist die Anwesenheit eines neutralen π^0 Meson im Endzustand. Diese Arbeit stellt den ersten Machbarkeitsnachweis einer Winkelanalyse mit einem neutralen Teilchen im Endzustand unter Verwendung des LHCb-Datensatzes dar. Die Auswahl der Ereignisse wird optimiert, um verbleibende Hintergrundverschmutzung zu eliminieren und gleichzeitig die verfügbare Winkelakzeptanz zu maximieren. Eine detaillierte Simulation des Detektors wird verwendet, um die Auswirkungen der Detektorakzeptanz, der Rekonstruktion und der Auswahl auf die Winkelverteilungen zu entfalten. Der Maximum-Likelihood-Fit zur Extraktion der Winkelparameter des Zerfalls $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ wurde entwickelt und anhand von Tausenden von Pseudoexperimenten getestet. Es wird gezeigt, dass die komplizierte Hintergrundstruktur zusammen mit der geringen statistischen Aussagekraft der verfügbaren Daten zu großen Unsicherheiten führt. Die Sensibilität der Analyse reicht aus, um eine Anomalie, die mit der in anderen Zerfällen beobachteten übereinstimmt, mit einer Signifikanz von mehr als zwei Standardabweichungen zu beobachten.

Key words: LHCb, FCNC, Seltener Zerfall, Elektroschwacher Pinguin, Flavor-Physik, Standardmodell, Neue Physik

Acknowledgement

Preface

This work has been carried out by the LHCb collaboration, which is an international collaboration of more than 1400 scientists and engineers from 86 institutes in 19 countries. This implies the usage of common software to analyze the data collected by the LHCb experiment. Both the collection of the data and the development of common software used to analyze the data stems from the effort of many current and former collaboration members. No part of this work would have been possible without the numerous contributions of the other (ex-)members of the collaboration.

This analysis is a significant advance towards the first angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay with the K^* decaying into $K^+\pi^0$ at the LHCb. The challenging reconstruction of the π^0 resulted in unforeseen challenges leading to large statistical uncertainties. All aspects of the analysis have been carefully studied and fully implemented in pseudoexperiments. Those are then used to extract the expected sensitivity to New Physics contributing to the angular parameters. With the current dataset, no sufficient impact on the knowledge of this parameters could be reached. However, this study still provides an important insight into the $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ decay, presents solutions to many unexpected complications, and will serve as a basis for future analyses performed using the anticipated Run III dataset.

This analysis has been carried out with the supervision of post-doctoral researchers at the Heidelberg University and at the École Polytechnique Fédérale de Lausanne. The author is responsible for the implementation of the selection software and the modifications to the FCNCFITTER. The validation of this implementation has been done in close collaboration with another PhD student.

Moreover, as the collaboration task, the author worked on the development of the software tool `TrackCalib`. `TrackCalib` provides the means to correct for the track reconstruction efficiency in any LHCb analysis. The author is also responsible for the validation and modifications of the corresponding tracking efficiency trigger and stripping lines, providing the collaboration with the best possible tracking efficiency estimation. This has been done in collaboration with a post-doctoral researcher at the Heidelberg University.

Contents

List of Figures

List of Tables

List of Abbreviations

CERN .... Conseil européen pour la recherche nucléaire
CL .... Confidence level
DIRA .... Direction Angle
DOCA .... Distance of closest approach
DTF .... Decay Tree Fitter
ECAL .... Electromagnetic calorimeter
FCNC .... Flavor-changing neutral currents
FD .... Flight distance
HCAL .... Hadronic calorimeter
HLT .... High-level trigger
IP .... Impact parameter
IT .... Inner Tracker
L0 .... Level-0
LCSR .... Light Cone Sum Rule
LEP .... Large Electron Positron
LHC .... Large Hadron Collider
LS1 .... Long Shutdown 1
MAMI .... Mainz microtron
MLP .... Multilayer perceptron
MVA .... Multi-variate analyis
NP .... New Physics
OT .... Outer Tracker
PDF .... Probability density function
PHSP .... Phase-space
PID .... Particle identification
PMT .... Photomultiplier
PS .... Preshower
QCD .... Quantum chromodynamics
QED .... Quantum electrodynamics
RICH .... Ring-imagining Cherenkov
SLAC .... Stanford linear accelerator center

Introduction

The Standard Model of particles is currently the best theory describing the very basic building blocks of the universe. Except for gravity, it describes all fundamental interactions between the elementary particles. In the last decades, it has been improved, probed and many of its predictions have been confirmed. Despite the tremendous success of the Standard Model, there are several unexplained phenomena: the non-zero mass of neutrinos, the excess of matter over antimatter in the universe or the presence of dark matter in the universe.

Today, in the high-energy experiment era, the focus is not on confirming the Standard Model, but on finding inconsistencies and processes where the theory breaks down. The Large Hadron Collider at CERN, the most powerful particle accelerator up-to-date, is designed to test the Standard Model parameters and its boundaries.

There are four large experiments at the Large Hadron Collider, ALICE, ATLAS, CMS and the LHCb experiment. ALICE main design goal is to investigate the origins of the universe, ATLAS and CMS are focused on measuring the Standard Model properties, especially the ones of the Higgs and electroweak bosons, and LHCb focuses on precise measurements of the predicted Standard Model parameters, searching directly for deviations from the predictions. Possible extensions of the Standard Model, New Physics, can contribute to the quantum loops of the Standard Model. As the particles contributing to the quantum loops are not limited by the available collision energy, the energy scale probed is beyond the energy scale probed by direct searches. This approach requires Standard Model predictions or constraints with similar or better accuracy than experimental measurements.

One of the smoking guns of New Physics contribution to the Standard Model are the decays of the b quark to an s quark and a pair of leptons. These decays can be measured through the decays of a B meson into $K^*\mu^+\mu^-$. Many previous measurements of this decay show tensions with the Standard Model predictions: either the branching fraction measurements [1, 2] or the angular analyses [3, 4, 5, 6, 7, 8, 9]. This work represents a significant step towards the angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay, where the

the K^{*+} decays into $K^+\pi^0$ using the LHCb dataset. The aim is to validate the observed anomalies, adding another jigsaw puzzle piece into the physics beyond the Standard Model picture.

In this thesis, the Standard Model is introduced with an emphasis on $B \to K^*\mu^+\mu^-$ decays. Then, the LHCb experimental setup is explained. In the third section, the $B^+ \to K^{*+}\mu^+\mu^-$ decay topology and observables are described. Then, the analysis method is briefly explained with an emphasis on the difficulties of this analysis. Section five describes the methods used for the signal candidates selection. The angular acceptance resolution and correction to the signal candidates is discussed in section six. The seventh section introduces the angular modeling that is validated in section eight by fitting the simulation sample and the $B^+ \to K^{*+}J/\psi$ channel in data. The angular model is further tested by pseudoexperiments as explained in section nine. The expected statistical uncertainty of the fit to the rare $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ decay is estimated. Lastly, the tenth section is dedicated to the authors work on the tracking efficiency measurement at the LHCb experiment. Throughout the thesis, natural units are assumed.

1   The Standard Model of particle physics

The structure of matter was a subject of contemplation of philosophers and scientists since ancient history. Particle physics as we know it today, however, exists only since the beginning of the 20th century. It was driven by two breakthrough discoveries: the discovery of electrons by J. J. Thompson in 1897 [10] and the discovery of the atomic nucleus by E. Rutherford in 1911 [11]. Electrons and protons were considered the main constituents of matter until the 1932, when this picture was expanded by the discovery of the neutron by J. Chadwick [12] and by the discovery of the positron by C. Anderson [13].

Around this time, the first ideas about interactions between particles emerged: the well-known electro-magnetism was joined by the strong force holding the nuclei together and by the weak-force describing beta radioactivity, discovered in 1896 by H. Becquerel [14]. Since the weak interaction is very important for the development of the Standard Model (SM) as well as for this work, it will be in the spotlight of this chapter.

1.1   The beginning of the Standard Model

The theory of the beta decay was successfully described by E. Fermi in 1933 [15], where he predicted the existence of a neutrino[1]. In this pioneering work, he suggested a direct interaction of four spin-$1/2$ quantum fields (neutron, proton, electron and antineutrino). This reflects the fact that the weak force has essentially zero-range: unlike electromagnetism, where the photon is the interaction mediator, there is no boson mediator.

At that time this was a great approximation of the beta decay *at low energies*. Even though Fermi's description is different from quantum electrodynamic (QED), describing the electromagnetic force, Fermi used Lorentz four-vectors to describe the fermion fields appearing in bi-linear combinations (*currents*), similarly to QED. This paved the path to electroweak interaction unification.

Fermi's theory postulated only beta decays with no change of nucleonic spin ($\Delta S = 0$). However, as nuclear spectroscopy became more precise, it was clear that beta decays with a change of nucleonic spin one ($\Delta S = 1$) do occur [17]. A generalization of the Fermi theory explaining this observations was proposed by G. Gamow and E. Teller in 1936 [18]. Instead of using only vector currents as Fermi did, one can construct the four-fermion interaction using the whole set of scalars (S), vectors (V), tensors (T), axial vectors (A) and pseudoscalars (P). A- and T-couplings describe a spin-changing interaction, while S- and V- couplings are spin-zero transition operators.

---

[1]Fermi's paper suggesting the existence of neutrinos was rejected by Nature, as it was "too remote from physical reality to be of interest to the readers" [16].

At that time it was intuitively assumed parity symmetry holds in the quantum world in the same way as in our macroscopic world. However, two particles with very similar masses and the same spin were observed, τ and θ, decaying into different final states: $\tau^+ \to \pi^+ \pi^+ \pi^-$ and $\theta^+ \to \pi^+ \pi^0$. As pions have intrinsic parity of -1, τ and θ had to be different particles or the parity symmetry must be violated[2]. In 1954, R. H. Dalitz looked into the two decays and confirmed τ and ϕ are the same particle [19]. Today, we call θ and τ neutral kaon K^0.

The theoretical solution to this $\tau - \theta$ *puzzle* was found in 1956, when T. D. Lee and C. N. Yang suggested to abandon the idea of parity symmetry conservation [20]. The violation of parity symmetry was quickly observed by three independent measurements: C. S. Wu measured the rate of electrons originating from the decay of ^{60}Co atoms aligned by a uniform magnetic field. If parity is conserved, the same amount of electrons should be emitted in the direction of the nucleus spin and opposite to the spin direction of the nucleus. She observed that the electrons prefer the direction opposite to the spin of the nucleus, confirming the parity violation [21]. The other two experiments studied the pion decay to a muon and a neutrino[3]: $\pi^+ \to \mu^+ + \nu$ [22, 23]. Since pions have spin zero and spin is conserved, the muon and the neutrino have to have opposite spin sign. If parity is conserved, the polarization of the muon would be symmetric along their direction of motion. However, the muon spin direction *favors* the direction of motion.

The general lagrangian proposed by Gamov and Teller expanded by the parity violating term was rather complex. This was reduced to only the V-A component in 1958 independently by R. Feynman with M. Gell-Mann [24] and R. Marschak with E. Sudurshan [25]. They followed the idea of two-component spinor from 1920's [26], applied it to neutrinos, and postulated that any elementary fermion, regardless of its mass, can participate in weak interactions only through the left-handed chiral component of the corresponding spinor field. This lead to the universal current-current form of the weak interaction:

$$\mathcal{L}_{int}^{weak} = \frac{G_F}{\sqrt{2}} J^\rho J_\rho^\dagger, \tag{1.1}$$

where $\mathcal{L}_{int}^{weak}$ is the weak interaction lagrangian, G_F is the *universal Fermi constant* and J^ρ is the weak current (or probability flux). This implies the possibility of interaction of the weak current with itself, opening the path for intermediate vector bosons, similar to QED.

---

[2]The angular momentum J is conserved. Therefore the parity of the three-pion system is equal to $(-1)^3 (-1)^{J(\tau)}$ and the parity of two-pion system $(-1)^2 (-1)^{J(\theta^+)}$.

[3]Muons quickly decay to an electron, serving as an analyzer for the muon polarization.

Given there is no physical reason for this, in the light of today's knowledge, the idea was a rather fortunate but very successful guess. The problem of this assumption was that the observed beta decay until then preferred the spin change of either $\Delta S = 0$ or $\Delta S = 2$, which is excluded by the two-component theory. On the other hand it described very well all known particles' weak interaction and even predicted some interactions.

Another major problem of this theory is that it is *not renormalizable*. This means it does not hold anymore at high energies of about $\sqrt{\frac{2\pi\sqrt{2}}{G_F}}$ [27, 28, 29]. This problem was solved by introducing an electrically charged intermediate vector boson W. Nonetheless, this brought a new problem: production of W^-W^+ pairs in fermion-antifermion annihilation lead to a power-like unitarity violation [30].

1.2   Unification of electromagnetism with weak interaction

As the W boson has electromagnetic charge, it has to interact electromagnetically. As the weak interaction when exchanging the W boson violates parity maximally, while QED is parity conserving, one cannot just add QED and weak lagrangians together. Instead of adding terms to the lagrangians, the unification of weak interaction with QED was realized in a very different way. Taking a step back in history, in 1954 C. N. Yang and R. Mills looked into non-abelian (non-commutative) gauge invariance [31]. They showed that the transformation from global to local symmetry requires a triplet of vector fields, analogous to the photon field. However, contrary to the photon field, this field also interacts with itself.

S. Glashow showed in 1961 that the minimal representation of the electroweak unification indeed requires four gauge fields: the known photon, W^+, W^- and a new neutral vector boson [32]. The new neutral vector boson (called Z today) is required to bridge the gap between parity-conserving electromagnetism and parity-violating weak interaction. In mathematical terms, the appropriate gauge group is a not simple $SU(2) \otimes U(1)$ group[4]. This effectively means wo independent coupling constants are needed.

This idea also led to an estimation of the W boson mass to be 77.7 GeV (the currently measured value is 80.38 ± 0.012 GeV [33]). While massive W bosons canceled the main divergences mentioned in the previous subsection, new divergences appeared in the interactions between the vector-boson fields with themselves.

---

[4]It is interesting that this $SU(2) \otimes U(1)$ electroweak unification was not the only theoretical solution: by introducing new electron-type leptons, one reaches simple $SU(2)$ electroweak unification. This is however not supported by the experimental data.

1.3   Renormalization problem

The renormalization problem was solved by adding a scalar field to the electroweak theory. The initial idea was published by J. Goldstone [34], who added the so-called *mexican hat* potential $V(\varphi)$ to the lagrangian density

$$V(\varphi) = -\mu^2 \varphi \varphi^* + \lambda(\varphi\varphi*)^2\,, \tag{1.2}$$

where μ is a real parameter with dimension of mass, λ is a dimensionless coupling constant and φ is a complex scalar field. The potential is sketched in Fig. 1.1. This potential has a minimum lying on a circle in the complex plane with radius $\mu/\sqrt{\lambda}$, therefore the minimal energy is infinitely degenerate. The ratio $\mu/\sqrt{\lambda}$ is commonly denoted as v and referred to as vacuum expectation value. This means the lagrangian is no longer symmetric at its minimum. The potential effectively describes two real scalar fields with masses $\mu\sqrt{2}$ and 0. The appearance of a massles bosonic excitation (Goldstone boson) is referred to as the *Goldstone theorem* [35].

The Goldstone model was further extended by P. Higgs [36] and others [37, 38], who described the interaction with an Abelian gauge field in the frame of the Goldstone model. When gauged, the Goldstone boson disappears and the gauge field acquires a mass. This is the famous *Higgs mechanism*. It was shown later by S. Weinberg that the Higgs mechanism is actually necessary for tree-level unitarity of the electroweak theory (*i.e.* renormalizability) [39]. The application of the Higgs mechanism on the Glashow model was further expanded by A. Salam [40] and today we refer to it as the Glashow-Weinberg-Salam model. They used the Higgs mechanism to generate also lepton and fermion masses, while keeping the electromagnetic interaction parity symmetric and the weak interaction parity violating.
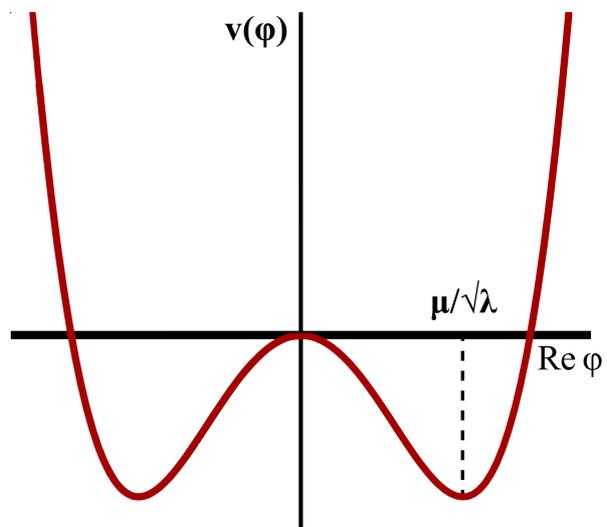
Fig. 1.1: Visualization of the Goldstone potential given by Eq. 1.2. The full potential is realized by the surface created by rotating the red curve around the y-axis.

1.4 Quark model

The picture of the Standard Model at this point in history is relying on the $SU(2) \otimes U(1)$ gauge symmetry and the Higgs mechanism realized via a complex scalar doublet.

At that time, baryons and mesons were considered to be elementary particles. That was only until 1961, when M. Gell-Mann and independently Y. Neeman proposed the *Eightfold way.* They noticed that the back-then known particles match an $SU(3)$ representation [41, 42]. Gell-Mann continued to work on this model, and in 1964 he used the word quark for the first time [43]. Independently of him, G. Zweig also proposed that "Both mesons and baryons are constructed from a set of three fundamental particles" [44, 45]. They postulated that quarks have only a partial charge of $1/3$ and $2/3$ and are fermions. They called the quarks *up*, *down* and *strange.*

In the same year, S. Glashow and J. Bjorken proposed the existence of a fourth - *charm* - quark. This was appealing at that time as the existence of τ lepton was yet to be discovered and the existence of two generations of quarks was symmetric to two generations of leptons [46].

The charm quark was later recognized by S. Glashow, J. Iliopoulos and L. Maiani (*GIM*) to be a crucial part of the Standard Model. The problem with the existence of three quarks was the interaction of quarks with the Z boson: the occurrence of strangeness-changing neutral currents was phenomenologically much smaller than expected. They added the fourth quark to the electroweak theory, allowing only for flavor-conserving neutral currents [47]. This gave the basics to the theory of flavor-changing neutral currents, where the divergences are cut-off by a heavy quark exchange in a loop. An example of such diagrams is shown in Fig. 1.2.

One of the remaining problems of the theory was CP violation. The CP violation was unexpectedly observed in 1964 in the decay of K^0 mesons [48]. Even though the community at that time was vary of accepting the quark model (the charm quark was still yet to be discovered), M. Kobayashi and T. Maskawa proposed the existence of two more quarks [49]. The model with two generations of quarks is CP conserving, while the proposed three generation model is not [49]. The matrix describing the strength of flavor-changing weak interaction is called *CKM* after N. Cabibbo[5], M. Kobayashi and T. Maskawa.

---

[5]N. Cabibbo postulated a similar matrix with two generation of quarks [50]. The motivation for such matrix was the fact that the $u \leftrightarrow d$, $e \leftrightarrow \nu_e$ and $\mu \leftrightarrow \nu_\mu$ transitions had similar measured amplitudes. On top of that, the transitions with strangeness change one ($\Delta s = 1$) have four times larger amplitude than processes with strangeness conserved. This was solved by Cabibbo by postulating weak universality and weak mixing angle θ_c.
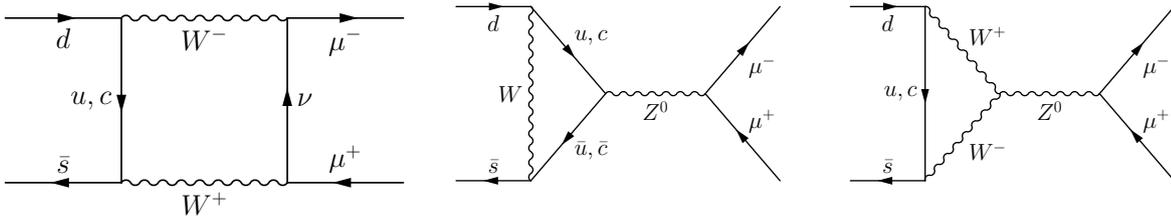
Fig. 1.2: Feynman diagrams of kaon decay to $\mu\mu$ including c-quark contribution. They were described in [51]. Note that there is also a long distnace contribution from $K^0_{\mathrm{L}} \to \gamma\gamma \to \mu^+\mu^-$.

1.5 $b \to sl^- l^+$ transitions

The exchange of heavy quarks in loops in flavor-changing neutral currents (FCNC) is a great tool to probe New Physics at high energies. The loops are sensitive to heavy particles and precision measurement of such processes could lead us to New Physics discovery, similarly as the kaon decay to muons led to the discovery of the charm quark. Higher-order transitions, such as the $b \to sl^- l+$ transition, are sensitive to New Physics, as they are even more suppressed by the GIM mechanism. The price to pay is that the interactions are rather rarely occurring. The typical decay rate for such a transition is 10^{-6}. These processes are then referred to as *rare decays*.

Experimentally reachable example of such higher-order FCNC interaction are $b \to s\ell^+\ell^-$ transitions. They occur through *box* and *penguin* diagrams, as shown in Fig. 1.3.
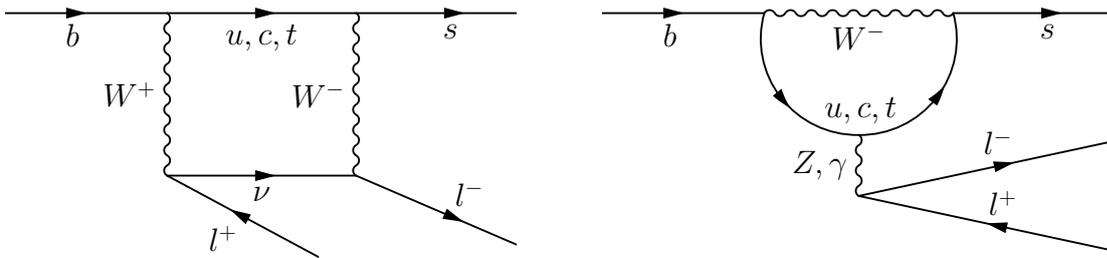
Fig. 1.3: Feynman diagrams of a $b \to s\ell^+\ell^-$ transition. The diagram on the left is referred to as *box* diagram, the right diagram is called *penguin* diagram.

The processed are mediated by γ, $W^\pm$ and Z bosons. One of the experimentally observable variables is the invariant mass squared of the lepton pair, q^2, as shown in Fig. 1.4. The $b \to s\ell^+\ell^-$ transition is dominated by several very different processes depending on the q^2 value. There are two problematic regions of q^2: around $9\,\mathrm{GeV}^2$ and $14\,\mathrm{GeV}^2$. In these regions, the process is dominated by a tree-level diagram of $b \to sJ/\psi$ and $b \to s\psi(2S)$, where J/ψ or $\psi(2S)$ decays into two leptons. As this region is dominated by a process

with different physics both theory and experiment typically omits these regions in their predictions or measurements.
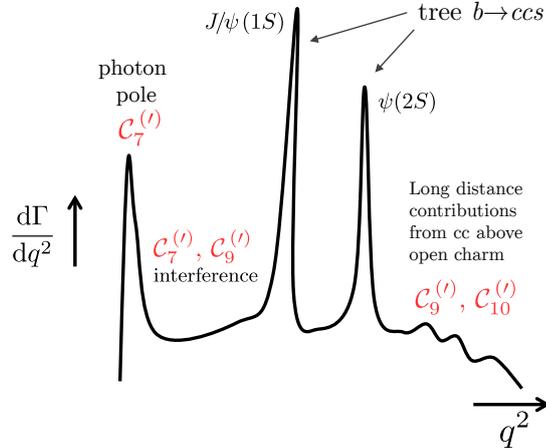
Fig. 1.4: Decay rate of $b \to s\ell^+\ell^-$ transition in depence on q^2. In the low q^2 region, the decay rate is dominated by the penguin diagram with photon exchange. With increasing q^2, contribution of other processes increases, until the decay rate is dominated by the J/ψ and $\psi(2S)$ charm resonances. At very high q^2, the decay rate is dominated by long distance contributions. For the details about the $C_i(\mu)$ variables see Eq. 1.3.
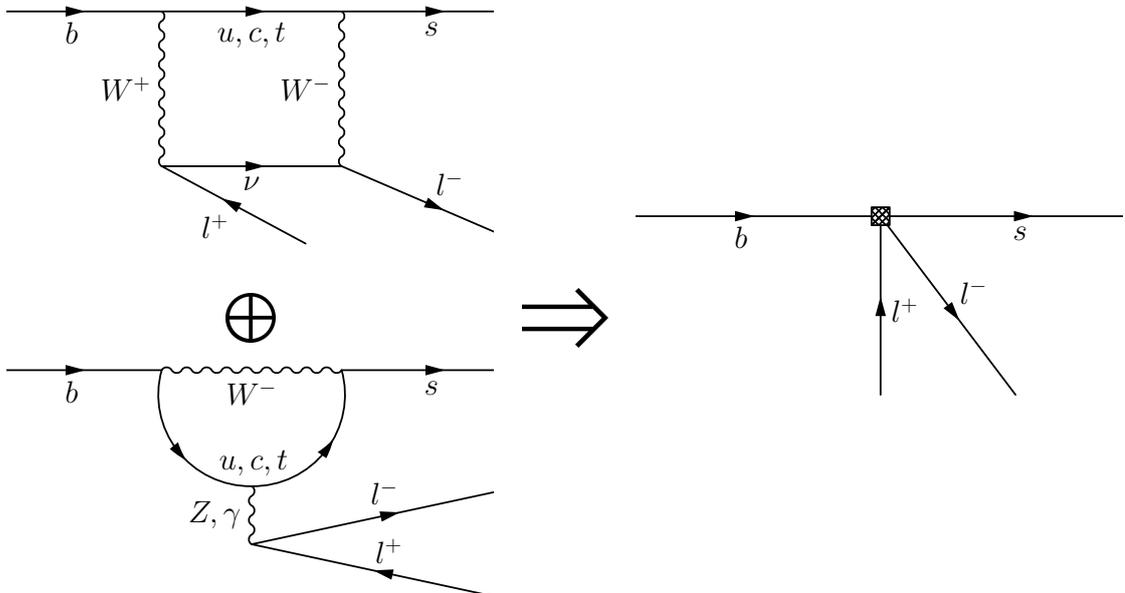
Fig. 1.5: Feynman diagrams of an effective $b \to s\ell^+\ell^-$ transition. Instead of looking at the interaction as a set of diagrams, we can describe the $b \to s\ell^+\ell^-$ transition as a point-like four-fermion interaction.

9

1.5 $b \rightarrow sl^-l^+$ transitions

Similarly as Fermi described the beta decay as one interaction of four fermions, one can apply this simplification also on these processes. The exchanged energy (smaller than the mass of the B^+ meson) is much lower than the energy scale of the quantum loop (mass of the W boson). Therefore, instead of looking at the interaction from the Standard Model point of view illustrated in Fig. 1.3, one can look at it as a point interaction, as shown in Fig. 1.5.

This description is commonly referred to as *effective theory*. The effective Hamiltonian of $b \rightarrow s\ell^+\ell^-$ transition can be expressed as:

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{\alpha_e}{4\pi} \sum_i C_i(\mu) \, \mathcal{O}_i(\mu) \, , \qquad (1.3)$$

where G_F is the weak decay constant, $V_{tb}V_{ts}^*$ are the CKM matrix elements describing the $b \rightarrow t$ and $t \rightarrow s$ transitions (the contributions of the u and c quarks to the loop is negligible), α_e is fine-structure constant, and $^1/4\pi$ comes from the loop suppression. Wilson coefficients $C_i(\mu)$ contain all information about short-distance physics in the transition above the renormalization scale μ.

The operators $\mathcal{O}_i(\mu)$ are local four-fermion operators with different Lorentz structures. These currents are all left-handed. Formally, the right-handed $\mathcal{O}'_i$ currents contribute to the Hamiltonian too, however they are very suppressed in the Standard Model due to the parity violating nature of the weak interaction described earlier.

Looking at FCNC transitions, the operator $\mathcal{O}_7$ describes the photon contribution to the decay rate and is constrained by radiative decays of $q \rightarrow q l^+ l^-$ transitions at small q^2. The operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$ are V and A currents, respectively. The operator $\mathcal{O}_8$ describes gluon contribution to the diagrams. Assuming the SM scale $\mu = M_W$, $\mathcal{O}_8$ vanishes in the Standard Model [52]. The operators are given in Eq. 1.4.

In the Eq. 1.4, e is the elementary charge, g is the strong coupling constant, and m_b is the running b quark mass. The matrices are denoted as follows: $\sigma_{\mu\nu}$ are Pauli matrices, $\gamma_{\nu,5}$ are Dirac matrices and λ^a are Gell-Mann matrices. The quark fields are denoted $\bar{s}$, b, the muon fields μ, $\bar{\mu}$, while $G^{\mu\nu\,a}$ is the gluon field tensor. The electromagnetic tensor is denoted $F^{\mu\nu}$.

$$\mathcal{O}_7 = \frac{e}{g^2} m_b (\bar{s}\sigma_{\mu\nu}\frac{1+\gamma_5}{2}b)F^{\mu\nu}\,,$$

$$\mathcal{O}_8 = \frac{1}{g} m_b (\bar{s}\sigma_{\mu\nu}\frac{\lambda^a}{2}\frac{1+\gamma_5}{2}b)G^{\mu\nu\,a}\,,$$

$$\mathcal{O}_9 = \frac{e^2}{g^2} (\bar{s}\gamma_\mu\frac{1-\gamma_5}{2}b)(\bar{\mu}\gamma^\mu\mu)\,,$$

$$\mathcal{O}_{10} = \frac{e^2}{g^2} (\bar{s}\gamma_\mu\frac{1-\gamma_5}{2}b)(\bar{\mu}\gamma^\mu\gamma_5\mu)\,.$$

(1.4)

These operators in combination with corresponding Wilson coefficients dominate in different q^2 regions, as illustrated in Fig. 1.4. As the effective theory allows for any kind of interaction, it can also describe New Physics contributions. If the measured Wilson coefficients values are different to SM expectations, it means the contribution of different SM processes is accompanied by New Physics process.

Unfortunately, this theory describes free quarks. In experiments, the quarks are bound by the strong force as depicted in Fig. 1.6, described by non-perturbative quantum chromodynamics (QCD)[6]. Despite the fact these effects are very hard to compute, there are several tools that provide these calculations. The most widely used tools are Lattice QCD [53] and Light-Cone-Sum-Rules (LCSR) [54]. Moreover, the calculations based on QCD factorisation [55] are typically performed for low q^2, for high q^2 ($q^2 \gtrsim 15\,\mathrm{GeV}^2$) the Operator Product Expansion [56] is used.
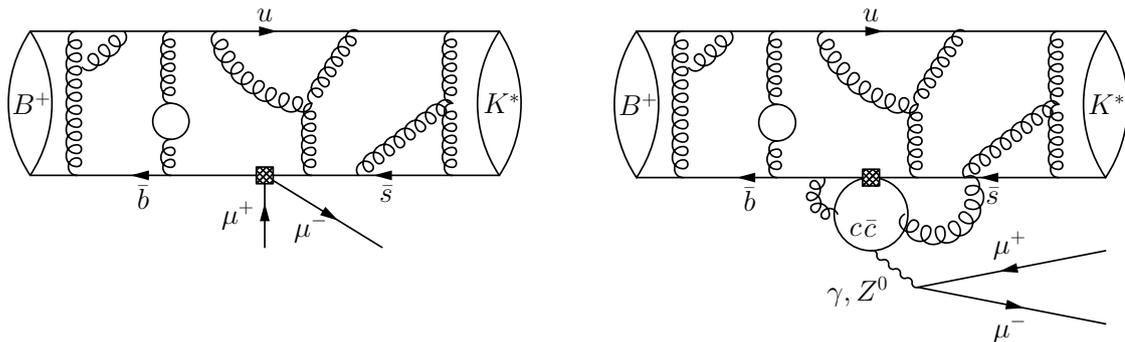
Fig. 1.6: Feynman diagrams of an effective $B^+ \to K^{*+}\mu^+\mu^-$ transition. The left diagram shows the $b \to s\ell^+\ell^-$ process in the context of the interacting quark being bound in a B^+ meson decaying into $K^{*+}\mu^+\mu^-$. There is also a non-factorizable contribution from charm loops, as shown on the right. Even though its contribution is much smaller, it needs to be correctly treated too.

---

[6]QCD is the theory of the strong interaction between quarks and gluons.

It is very hard to disentangle QCD processes from the $b \to s\ell^+\ell^-$ transition, however by choosing a convenient basis and variables, the form-factor influence can be at least removed at the first order (this will be described later in Sec. 3.2). This is a limiting factor in many theory predictions for this process, even though the calculations are constantly improved.

Besides the challenging form-factor contributions, there is another non-factorizable contribution: charm loops [57]. The process is depicted in Fig. 1.6 on the right. A charm loop is coupled to the b and s quarks and to a virtual photon decaying to the muon pair. Their contribution is much smaller than the one of form-factors. However, with increasing precision of both measurements and QCD calculations, their effect becomes significant. Additional gluons can come into play, making the theoretical calculations even harder. This process is included in the $\mathcal{C}_9$ Wilson coefficient and it is therefore important to separate this effect from possible contributions of physics beyond the Standard Model.

1.5.1 New Physics

As the effective theory allows for any kind of interaction, it can also describe New Physics (NP) contributions at large energy scale: they can be integrated out similarly to the electroweak bosons. Other operators that are negligible in the SM can contribute to the effective hamiltonian. An example of such operators $\mathcal{O}_S$, $\mathcal{O}_P$ and $\mathcal{O}_T$ that can bee added *e.g.* from Higgs penguins[7] is listed in Eq. 1.5.

$$
\begin{aligned}
\mathcal{O}_S &= \left(\bar{s}\frac{1+\gamma_5}{2}b\right)(\bar{\mu}\mu) \,, \\
\mathcal{O}_P &= \left(\bar{s}\frac{1+\gamma_5}{2}b\right)(\bar{\mu}\gamma_5\mu) \,, \\
\mathcal{O}_T &= \bar{s}\sigma_{\mu\nu}b\bar{\mu}\sigma^{\mu\nu}\mu
\end{aligned}
\tag{1.5}
$$

Moreover, the right-handed Wilson coefficients C_i' come into play. Different decays are sensitive to different coefficients, as presented in Tab. 1.1. From the table it becomes clear that the decay of $B^+ \to K^{*+}\mu^+\mu^-$ is sensitive to most of the coefficients.

Decay	$\mathcal{C}_7$	$\mathcal{C}_9$	$\mathcal{C}_{10}$	$C_{S,P}$
$B \to (X_s, K^*)\gamma$	✓	×	×	×
$B \to (X_s, K^*)\ell^-\ell^+$	✓	✓	✓	×
$B_s^0 \to \mu^+\mu^-$	×	×	✓	✓

Tab. 1.1: Sensitivity of Wilson coefficient for different decays, where X_s stands for any inclusive decay with an s quark.

---

[7]Higgs penguin is a Higgs to fermion-antifermion transition.

In most of the measurements with the potential to constrain New Physics, there is a good agreement with the SM. However, in several previous measurements, tensions of a few standard deviations appear. These measurements are discussed later in Sec. 3.6. All these tensions are in the order of 2-3 standard deviations away from the Standard Model prediction. However, they are all consistent with each other, hinting at possible New Physics contribution to the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$.

There are numerous NP approaches to explain these tensions, including the supersymmetric theory or the string theory. Among the currently most discussed theories are *portal*, *loop*, and *leptoquark* models. *Portal models* assume a particle responsible for the tensions that can be also involved in the dark matter production in the early universe [58, 59, 60]. This is typically a Z' boson. In these models, the Z' boson contributes to the operator $\mathcal{O}_9$ (and sometimes to the operator $\mathcal{O}_{10}$) with flavor violating couplings to quarks and non-universal couplings to leptons. Portal models provide corrections to $\mathcal{C}_9$, however they also mean unwanted contributions to other Wilson coefficients. An example of a Feynman graph with a Z' boson is in Fig. 1.7, left. *Loop* models postulate that the NP contribution comes from loops containing particles. These particles are in some cases potential dark matter candidates [61, 62, 63]. An example of such a hypothetical loop is in Fig. 1.7, middle. *Leptoquark* models assume the existence of two (or more) leptoquarks: a particle carrying both lepton and baryon number, allowing leptons and quarks to interact directly, as shown in Fig. 1.7, right. This model can answer the question why are neutrinos massive, but it can also explain some of the tensions in $b \to s\ell^+\ell^-$ decays [64, 65, 66, 67, 68]. There are also other models combing these approaches [69, 70, 71].
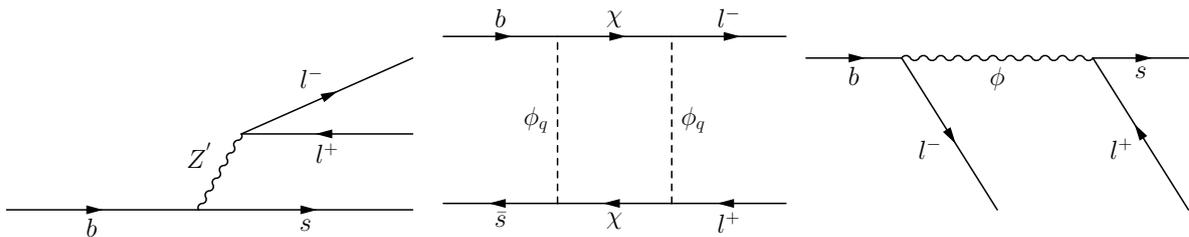
Fig. 1.7: Potential New Physics Feynman diagrams. On the left, Feynman diagram with a potential Z' gauge boson, in the middle possible loop diagram with a contribution from a dark model particles, on the right diagram of an $b \to s\ell^+\ell^-$ interaction through leptoquarks.

2   Experimental setup

In this section, the experiment used for the analysis presented here is described. The detector is placed in an accelerator facility, which is briefly introduced. The LHCb sub-detectors used in the analysis are presented, as well as the data-acquisition procedure.

2.1   The Large Hadron Collider

The story of the Large Hadron Collider (LHC) [72] reaches all the way back to 1976, when the particle physics community started to think about building the LEP (Large Electron Positron) collider. LEP was a 27 kilometer circular collider, placed in a tunnel at CERN (Conseil européen pour la recherche nucléaire) near Geneva, Switzerland. CERN is the largest physics laboratory in the world. Its main purpose it to provide *a unique range of particle accelerator facilities that enable research at the forefront of human knowledge*, *perform world-class research in fundamental physics* and *unite people from all over the world to push the frontiers of science and technology, for the benefit of all* [73].

LEP started its full operation in 1989 and was functioning until 2001, when the dismantling started, making room for the LHC [74] to be placed in the same tunnel.

The first discussions about replacing the e^--e^+ collider by a hadron collider started as early as 1984. The construction of LHC was approved 20 years later, in December 1994 [75]. The construction began in 1998 and the first collisions were delivered in 2008. However, during the initial testing, one of the superconducting magnets quenched[8]. As a result, 53 magnets were damaged, postponing the data taking to 2009.

The LHC with its circumference of 27 km is up to this day the largest and most powerful particle collider in the world. The LHC sits on the top of the CERN's accelerator complex, a succession of machines that accelerate the beam of particles to higher and higher energies, as illustrated in Fig. 2.1. The accelerated particles are in the LHC's main operation mode protons, however there are periods where other heavier ions are accelerated. The LHC itself consists of two circular storage rings where protons are injected with an energy of 450 GeV and they are accelerated to energies up to 7 TeV. The accelerated protons are collided at four main interactions points each surrounded by a large detector: ALICE, ATLAS, CMS and LHCb.

So far, the LHC has operated in two longer periods, called Run I (2010-2012) and Run II (2015-2018). In Run I, the maximal nominal energy was 3.5 TeV with increase to 4 TeV in 2012. Run I was followed by a maintenance period, long shut-down one (LS1), where the dipoles of the LHC were improved, resulting in a maximal nominal energy of

---

[8]Quenching is when a part of the superconducting coil returns to its conducting state.

6.5 TeV in Run II. At the moment, the LHC is in the second long shut-down period (LS2), where ALICE and LHCb are undergoing major upgrades. This will be followed by Run III with a maximal nominal energy of 7 TeV [76]. The duration of Run III is foreseen to be three years and the expected performance can be found in Ref. [77].
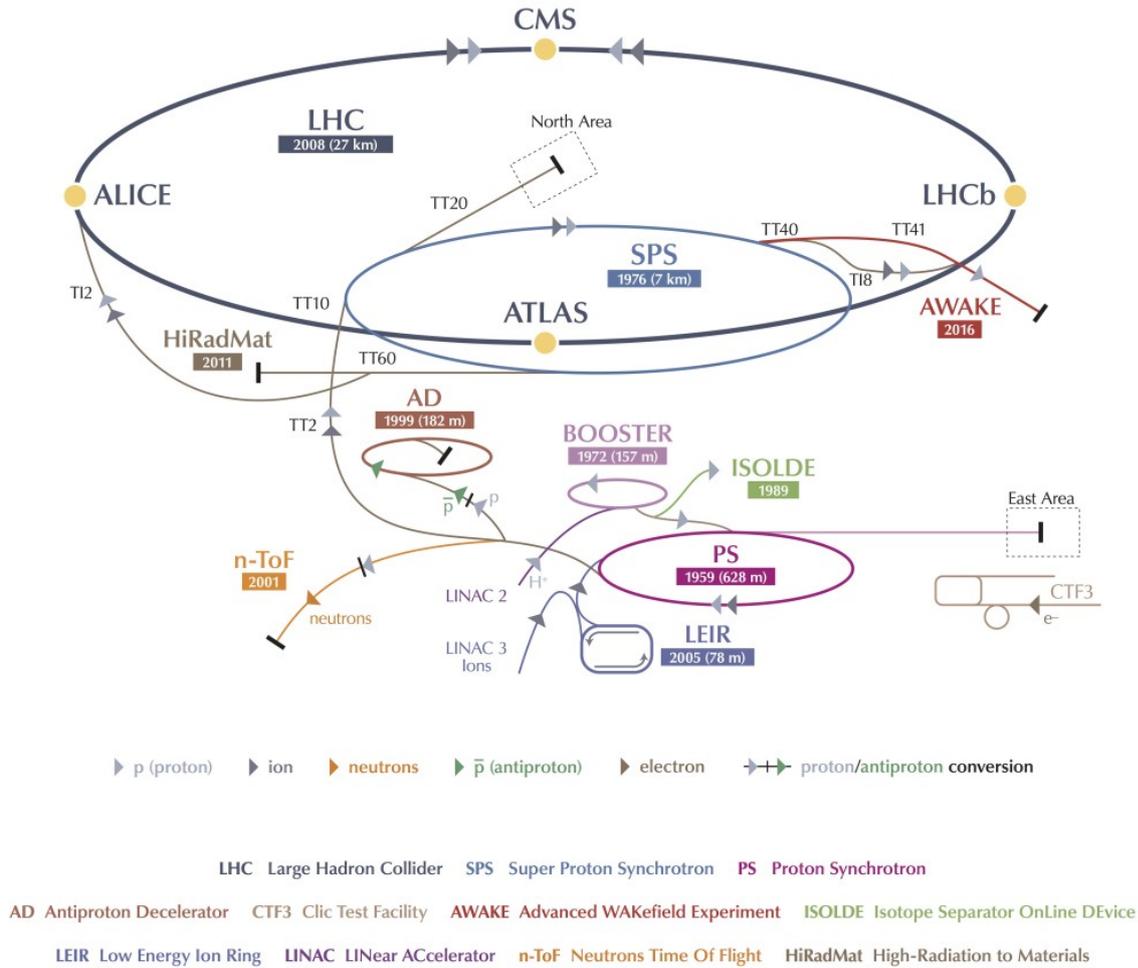
Fig. 2.1: CERN's accelerator complex. The protons are initially accelerated by **LINAC 2** and brought to **BOOSTER**. They continue to **PS**, **SPS** and from there they are finally steered into **LHC**. Heavier ions follow a similar path, they are initally accelerated in **LINAC 3**, they continue to **LEIR**. From there they are step by step brought to **PS**, **SPS** and **LHC**. Taken from Ref. [78].

2.2 The LHCb experiment

The LHCb detector [79, 80, 81] is a single arm forward spectrometer located at Point 8 (alongside the Geneva airport runway) at the LHC ring. The detector was mainly designed for precision measurements of CP violation and to study rare decays in the B and D meson systems [79].

The full LHCb design is described in Fig. 2.3. The LHCb coordinate system originates at the nominal interaction point. The z-axis is defined along the center of the beam, its positive part pointing from the interaction point into the detector and negative part pointing from the interaction point away from the detector. The y-axis is defined upwards in vertical direction from the interaction point, the x-axis similarly in horizontal direction. In order for the coordinate system to be right-handed, the positive x-axis is defined pointing to the left side, viewing in the positive direction of the z-axis. This allows for the definition of azimuthal angle ϕ, spherical angle θ and pseudorapidity[9]. For readers convenience, terms *downstream*

Fig. 2.2: The production angle of the $b\bar{b}$ pair at the center-of-mass energy of 14 TeV. Red color represents the LHCb acceptance. 27% of all produced b or $\bar{b}$ quarks are produced in the LHCb acceptance. In a standard general-purpose detector (asuming the acceptance along the beam as $-180 - 180$ mrad), 49% of b or $\bar{b}$ quarks are produced in its acceptance. Taken from [82].

(in the direction of beam into the LHCb acceptance, *i.e.* beam direction towards the ATLAS experiment) and *upstream* (beam direction towards the CMS experiment) are defined [83].

Since heavy quarks are mainly produced in the forward direction [84], the LHCb is designed to cover the very forward region as illustrated in Fig. 2.2. The LHCb geometrical acceptance is $10 - 300$ mrad in the $x - z$ plane and $10 - 250$ mrad in the $y - z$ plane.

In order to fulfill the design goals, the LHCb detector has to have very high track reconstruction efficiency, good π-K separation over a large energy range and excellent decay-time resolution. This is realized by several subsystems described in the following sections. Thanks to this universal detector design LHCb does not only excel in preci-

---

[9]Pseudorapidity η is defined as $\eta = \ln{(\tan{\theta/2})}$, where the spherical angle θ is the angle between the beam-pipe and particle's trajectory.
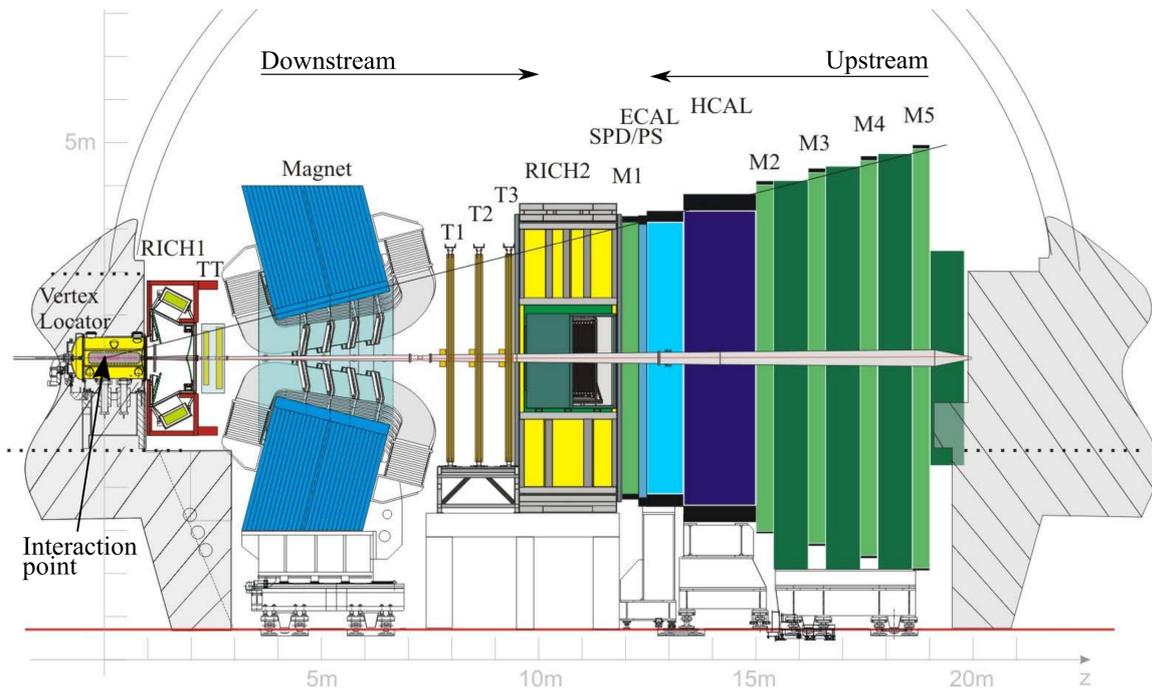
Fig. 2.3: The LHCb detector. Taken from [85].

sion measurements of B and D mesons, but also studies many new exotic states and particles [86, 87, 88, 89], and performs precision measurements of gauge boson properties [90, 91, 92]. Recent developments also allow for studies of heavy-ion collisions, for example excited $b\bar{b}$ resonance states $\Upsilon(nS)$ are observed to be suppressed in proton-lead collisions compared to proton-proton collision suppression, more so with larger n (corresponding to higher excited states) [93]. Moreover, LHCb is the only experiment at the LHC that is able to operate also in a fixed target mode. In the fixed target mode the proton beam collides with a gas target in the beam pipe. This was initially intended as a luminosity measurement [94]. Exploiting this program, *e.g.* a measurement of the antiproton production cross-section in proton-helium collisions was carried out, impacting the interpretation of results on antiproton cosmic rays from space-borne experiments [95].

2.2.1 Tracking system and vertex reconstruction

Precise vertex reconstruction is crucial for precision measurements of b hadron decays as displaced secondary vertices are typical for them. LHCb also has to have high event reconstruction efficiencies while maintaining high-speed online (trigger) selection (for more details see Sec. 2.2.5) in order to fully exploit its physics potential. Furthermore, since the main limitation for the momentum resolution is multiple scattering, the amount of material in the detector has to be minimal.

The vertex reconstruction is realized by the VELO (VErtex LOcator) detector surrounding the interaction point [96]. The VELO is consisting of two retractable halves placed along the beam direction, each consisting of 21 silicon micro-strip stations. The strips are arranged in the $r - \phi$ plane [97]. An illustration of the strips arrangement is shown in Fig. 2.4. This arrangement has the natural advantage of having the smallest segments closest to the beam. The retractable halves are open during beam setup. Once the beam in the LHC is stable, the detectors halves close around the beam, placing the closest sensors only 8 mm away from the beam itself.
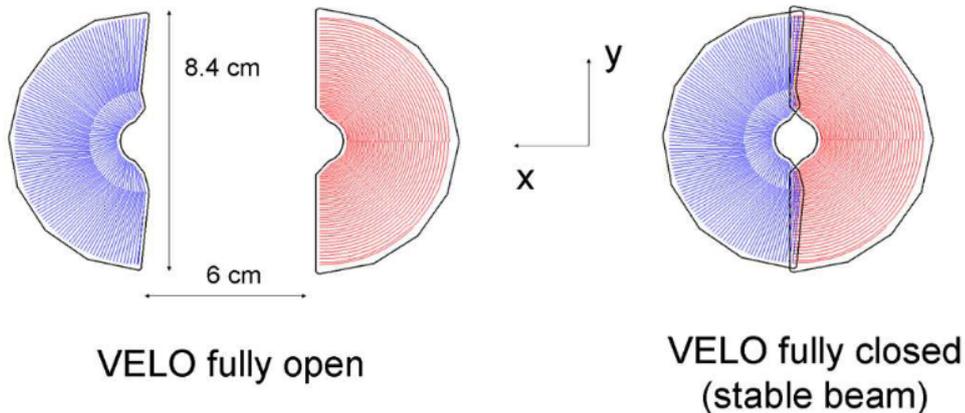
Fig. 2.4: The VELO silicon sensor sketch with R sensors in blue and ϕ sensors in red. Taken from [81].

For the physics program of LHCb it is important to reconstruct the vertex position, displaced tracks and vertices, impact parameter and decay time with high resolution and precision. The impact parameter (for the definition see Sec. 5.2) resolution in µm of the VELO in the x and y direction is $(16+24/p_T[\text{GeV}])$ µm, the vertex resolution goes down to 10 µm in the x and y direction and 50 µm in the z-direction [98]. The decay time resolution for B meson decays is around 40 fs [98]. Despite operating in an environment with very high radiation, the VELO detector's performance is stable throughout the years.

The momentum information for charged tracks is obtained by combining information from the VELO and three subdetectors downstream of the VELO: TT (trigger tracker), IT (Inner Tracker) and OT (Outer Tracker). The VELO can measure particle tracks and decay vertices, however, there is no momentum information. For this, a large magnet upstream of the VELO is used. The magnet has bending power of 4 Tm. This field is strong enough to allow the tracking system to perform momentum measurements with a good precision of tracks with momenta up to 200 GeV [99]. The magnetic field has two configurations, *down*, when the dipole field is along the positive y-axis, and *up*, when the

dipole field is along the negative y-axis. The $x - z$ plane is then referred to as *bending plane* and $y - z$ as the *non-bending plane*. The polarity of the magnetic field is periodically changed in order to control the detection asymmetries. The detection asymmetries need to be as small as possible for CP violation studies [100].

The TT detector is placed upstream of the magnet and consists of four silicon microstrip planes [101]. The TT is especially important for fast trigger selection, as explained later in Sec. 2.2.5.

The Inner Tracker [102], and the Outer Tracker [103] are located upstream of the magnet. The Inner Tracker is made of three silicon microstrip stations surrounding the beam pipe. The hit resolution of the IT is $50\,\mu$m. The OT surrounds the IT. It consists of straw tubes and has a hit resolution 170 μm. As mentioned previously, the limiting factor in the momentum resolution is multiple scattering and not the spatial resolution of the tracking detectors. The hits from these events are matched to the hits in the VELO and the TT, allowing for momentum measurement. The overall relative momentum resolution ranges from 0.4% (tracks with momentum $\sim 5\,$GeV) to 0.6% (tracks with momentum $\sim 100\,$GeV).

2.2.2   RICH detectors

As mentioned in the beginning of this section, for precision measurement of rare b and c decays as well as CP violation measurements, it is crucial to distinguish charged pions from charged kaons. The particle identification (PID) is achieved using two ring-imaging Cherenkov (RICH) detectors [104]. One is placed upstream the magnet, one is placed downstream the magnet, as shown in Fig. 2.3. This placement allows for PID of tracks with momentum ranging from 1 GeV to 150 GeV while covering the whole LHCb geometrical acceptance. The efficiency and fake rate of the RICH discrimination between pions and kaons is displayed in Fig. 2.5. In the figure, two configurations are shown, $\Delta LL(K-\pi) > 0$ and $\Delta LL(K - \pi) > 5^{10}$, where $\Delta LL(K - \pi)$ is the difference in logarithmic likelihood obtained by combining information from all PID detectors between the kaon and pion hypotheses: $\log \mathcal{L}_K - \log \mathcal{L}_\pi$. The reader can imagine $\Delta LL(K - \pi)$ as a measure of the probability that hypothetic kaon is not a pion.

The RICH detector does not only discriminate between pions and kaons, but also deuterons and protons. The PID of electrons, muons, and photons is obtained using the muon system and the calorimeters.

---

[10]In the analysis, $\Delta LL(K - \pi)$ is typically denoted as $\mathrm{DLL}_{K\pi}$.
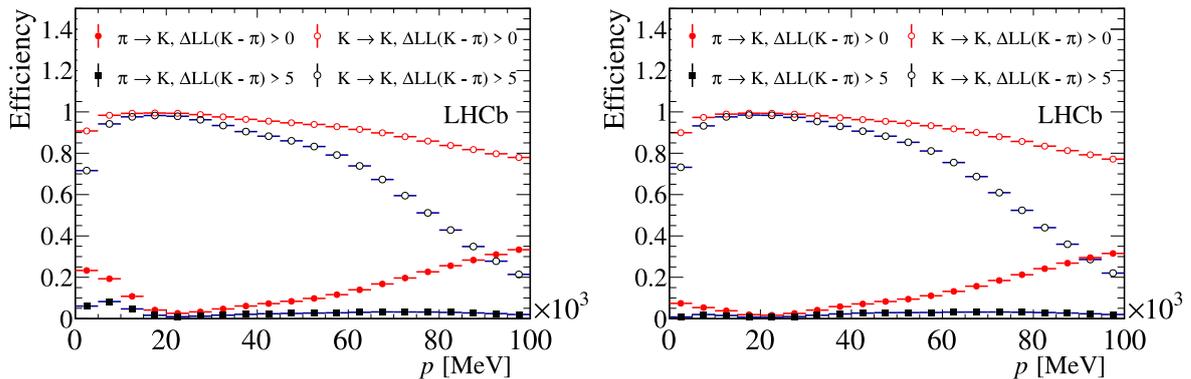
Fig. 2.5: Efficiency (open points) and fake rate (full points) of the RICH identification for the 2012 (left) and the 2016 (right) data as a function of on momentum. Two settings are shown, $\Delta LL(K - \pi) > 0$ and $\Delta LL(K - \pi) > 5$ (see the full text for the definition). There is a small improvement in Run II for particles below $15\,\text{GeV}$. Modified from [97].

2.2.3 Calorimeter system

The LHCb calorimeter system consists of four calorimeters: SPD (Scintillating Pad Detector), PS (Preshower), ECAL (electromagnetic calorimeter) and HCAL (hadronic calorimeter). The main goal of the system is fast identification and energy measurement of electrons, photons and hadrons [105]. A sketch of the calorimeter system is in Fig. 2.6.

The Scintillating pad detector is located upstream of a lead wall that creates electron and photon showers, while the PS is located downstream of the wall. They allow for clear separation between electron and photon showers, as photons do not leave a signal in the SPD [105].

The electromagnetic calorimeter is made of 66 layers of $4\,\text{mm}$ thick scintillator layers between $2\,\text{mm}$ thick lead, corresponding to 25 radiation lengths. The *shashlik* design is budget-wise, reliable and allows for fast response time ($25\,\text{ns}$ corresponding to a $40\,\text{MHz}$ readout), as the ECAL is crucial for the trigger selection. This design also has good radiation resistance [105].
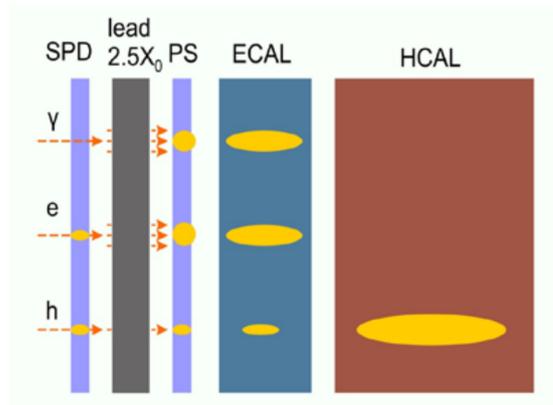
Fig. 2.6: LHCb calorimeter system. Electrons and hadrons are registered in SPD. SPD readout is limited to 0 (no hit) and 1 (hit). Electrons and photons are stopped in a lead wall (X_0 denotes the radiation lenght), creating a shower registered by PS and stopped in ECAL. Hadrons leave signal in all the detectors in the calorimeter system.

In order to separate signal from background in b decays with π^0 mesons and photons or to study lepton-flavor-universality, the photon and electron reconstruction has to be accurate: spatial and energy resolution has to be very good. The ECAL energy resolution is $\sigma/E = {}^{0.1}\!/\sqrt{E\text{[GeV]}} + 0.01$, which satisfies this requirement. The ECAL transverse granularity varies as particle flux increases towards the beam in the $x - y$ plane in order to minimize pile-up of hits in the detector, ensuring good signal-background separation [105].

The main purpose of the HCAL is to provide very fast response while having minimal detector dead-time. It measures the energy deposited by hadrons, contributing significantly to the first stage of the trigger selection of events, where the selection of high-energy events is performed.

2.2.4   Muon system

The most downstream LHCb subdetector is the muon system. It consists of five rectangular stations M1-M5: M1 is equipped with triple gas-electron-multipliers, M2-M5 are equipped with multi-wire-proportional chambers and interleaved with iron absorbers to stop very-high-energy hadrons that reach the muon station and to select penetrating muons [106].

The main purpose of this system is the trigger selection and the PID of muons. Muons detected in the muon system have minimal momentum of $3\,\text{GeV}$, as they have to pass the other LHCb subdetectors. Muons are reconstructed with an efficiency of 97%, while the pion misidentification probability varies with momentum between one and three percent [106].

2.2.5   Trigger system and data flow

At LHCb, the proton bunches collide at a rate of $40\,\text{MHz}$. With every bunch crossing, one or two inelastic proton-proton collisions occur. In 2.5‰ of the collisions a $b\bar{b}$ pair is produced. In about 15% of such events at least one B meson is produced with all its decay products in the LHCb acceptance [81]. Moreover, the typical branching ratios of B mesons used in CP violation studies are less than 10^{-3} and in the case of rare b decays the branching ratios are less than 10^{-6}. Therefore, a fast and efficient online selection of events is essential to fully exploit the data while keeping the data flow level manageable [107]. The rate of events saved for physics analysis is 2-5 kHz in Run I and 12.5 kHz in Run II [97, 98].

The LHCb online selection, commonly called *trigger*, is composed of a set of algorithms that classify (a part of) events as interesting for further analysis called *lines* [98]. The lines are applied in two stages: Hardware level-0 trigger and software high-level trigger.

The Level-0 (L0) trigger's purpose is to achieve a readout rate of 1.1 MHz with a fixed latency of 4 μs [108]. L0 trigger lines use information about the deposited energy from the calorimeters and muon stations, selecting events with high p_{T} or E_{T} signatures [109].

The High-level trigger (HLT) was significantly improved during LS1 as the computing resources doubled [110]. In Run I, the HLT was divided into two levels: HLT1 and HLT2. In HLT1, partial reconstruction of the event was performed, reducing the event rate to about 80 kHz. In HLT2 the full event reconstruction was executed. Where possible, HLT used offline-like algorithms with some simplifications due to time constraints [97]. HLT1 and HLT2 were processed independently in Run I. In Run II however, the events passing HLT1 were buffered into disk, online alignment and calibration of the detector were performed and HLT2 then performed a *full offline-like event reconstruction*. This allowed for better exploitation of *exclusive* lines (lines selecting a specific final state)[11]. As the reconstruction is performed online in a timely manner, large quantities of data can be processed fast, leading to more efficient data taking and faster publication of early measurement results, *e.g.* Ref. [111, 112]. The Run I and Run II trigger schemes are shown in Fig. 2.7.

2.2.6 Simulation

In any high energy experiment, Monte Carlo [114] simulation samples are needed to understand experimental conditions and the detector performance. Today, a simulation is a complicated project requiring vast computing power [115]. The generation of events used by the LHCb collaboration [116] is realized by the GAUSS[12] simulation framework [117]. The events are initially generated using PYTHIA[13] [118, 119]. PYTHIA simulates the proton-proton collision according to the Standard Model (although it is also possible to simulate New Physics processes) and the hadronization of the produced quarks and gluons. The decays of B mesons are generated via EVTGEN [120]. For this work, generating full dataset containing all kinds of final-state particles is not feasible. Therefore, events not containing a B^+ meson are immediately disregarded. Once a B^+ meson is found, they always decay into $K^{*+}\mu^+\mu^-$. The generated events then interact with the detector, which is simulated by GEANT4[14][121, 122].

---

[11]*Inclusive* lines select events with typologies typical for a given decay, looking for signatures such as displaced vertex or dimuons.

[12]Named after C. F. Gauss, German mathematician, making significant contributions to number theory, geometry, probability theory and other fields.

[13]Named after Pythia, Oracle of Delphi. Pythia was channeling prophecies from the Greek god Apollo.

[14]GEANT4 stands for GEometry ANd Tracking.

LHCb 2012 Trigger Diagram

40 MHz bunch crossing rate

L0 Hardware Trigger : 1 MHz readout, high E_T/P_T signatures

| 450 kHz $h^{\pm}$ | 400 kHz μ/$\mu\mu$ | 150 kHz e/γ |

Software High Level Trigger

29000 Logical CPU cores

Offline reconstruction tuned to trigger time constraints

Mixture of exclusive and inclusive selection algorithms

5 kHz (0.3 GB/s) to storage

| 2 kHz Inclusive Topological | 2 kHz Inclusive/ Exclusive Charm | 1 kHz Muon and DiMuon |

LHCb 2015 Trigger Diagram

40 MHz bunch crossing rate

L0 Hardware Trigger : 1 MHz readout, high E_T/P_T signatures

| 450 kHz $h^{\pm}$ | 400 kHz μ/$\mu\mu$ | 150 kHz e/γ |

Software High Level Trigger

Partial event reconstruction, select displaced tracks/vertices and dimuons

Buffer events to disk, perform online detector calibration and alignment

Full offline-like event selection, mixture of inclusive and exclusive triggers

12.5 kHz (0.6 GB/s) to storage

Fig. 2.7: LHCb trigger scheme in 2012 (left) and 2015 (right). The data aquisition starts at 40 MHz bunch crossing rate. This is reduced by a L0 hardware triger to 1.1 MHz by selecting events with high p_T or E_T signatures. These events are further selected by a software trigger: in 2012, this was done as a selection of inclusive and exclusive trigger lines, while in 2015, full offline-like event selection is performed thanks to full online detector calibration and alignment. In 2012, final output of 5 kHz was written to storage in three streams, in 2015 the final output was 12.5 kHz of fully-reconstructed events. Taken from Ref. [113].

The digitization of the detector response is simulated using BOOLE[15] [123]. At this step, noise is added to the detector response. The Level-0 trigger is also simulated in BOOLE, as it is purely a hardware trigger. From there on, the simulation is steered into the same flow as the real data: the high-level trigger response is emulated by MOORE[16][124], and the events are reconstructed using BRUNEL[17] [125]. The simulation then mimics the real data and its reconstruction.

---

[15]Named after G. Boole, the founder of boolean algebra.

[16]G. E. Moore is the author of Moore's law, the observation that the number of transistors on a microchip doubles every two years.

[17]I. K. Brunel was a British engineer, playing an important role in the industrial revolution.

3   Theoretical introduction to $\mathbf{B^+ \to K^{*+}\mu^+\mu^-}$

The theory of $b \to s\ell^+\ell^-$ transitions is generally introduced in Sec. 1.5. In this section, this general introduction is broadened using the specific example of $B^+ \to K^{*+}\mu^+\mu^-$ decay. Note that throughout the thesis the inclusion of opposite charged decays (*i.e.* $B^- \to K^{*-}\mu^+\mu^-$) is implied. The vector meson K^{*+} decays further in a $K^+\pi^0$ pair. The K^{*+} meson considered in this works is the $K^{*+}(892)$ resonance with natural width of $50.8 \pm 0.8\,\text{MeV}$ [33], spin 1, and parity -1.

The decay of $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ was first observed by the BaBar collaboration [126]. As the final state of the decay contains neutral particles, which are challenging to reconstruct, the angular analysis including this decay has been so far only performed by BaBar [127] and Belle [128].

The $B^+ \to K^{*+}\mu^+\mu^-$ decay is a FCNC transition. Therefore it can only occur through loop diagrams in the SM. The lowest SM diagrams that contribute to the decay are two penguin diagrams exchanging a photon or a Z boson and a box diagram exchanging two W bosons, as shown in Fig. 1.3). Looking back at Sec. 1.5, the contribution to the effective hamiltonian is polluted by QCD contributions (see Fig. 1.6). One way to validate the form-factor corrections to the decay is to change the *spectator* quark in the decay: swapping the u quark and d quark changes the decay from $B^+ \to K^{*+}\mu^+\mu^-$ to $B^0 \to K^{*0}\mu^+\mu^-$. Hence, it is important to study the $B^+ \to K^{*+}\mu^+\mu^-$ decay and compare the results to previous extensive measurements of $B^0 \to K^{*0}\mu^+\mu^-$ decay.

3.1   Decay topology

Due to the spin structure of the decay $B^+ \to K^{*+}\mu^+\mu^-$, the differential decay rate can be fully expressed using only four variables: the dimuon invariant mass squared q^2 and the three angles defined by the direction of flight of the decay products: θ_K, θ_L and ϕ. These angles are shown in Fig. 3.1. Denoting the normalized vector of a particle X in the rest frame of Y, $\hat{p}_X^{(Y)}$, the angles can be defined as in Eq. 3.1: This definition is compatible with previous LHCb measurements [3, 4, 5, 6].

$$\begin{aligned}
\cos\theta_L &= \left(\hat{p}_{\mu\pm}^{(\mu^+\mu^-)}\right) \cdot \left(\hat{p}_{\mu^+\mu^-}^{(B\pm)}\right) = \left(\hat{p}_{\mu\pm}^{(\mu^+\mu^-)}\right) \cdot \left(-\hat{p}_{(B\pm)}^{\mu^+\mu^-}\right) ,\\
\cos\theta_K &= \left(\hat{p}_{K\pm}^{(K^{*\pm})}\right) \cdot \left(\hat{p}_{K^{*\pm}}^{(B\pm)}\right) = \left(\hat{p}_{K\pm}^{(K^{*\pm})}\right) \cdot \left(-\hat{p}_{(B\pm)}^{K^{*\pm}}\right) ,\\
\cos\phi &= \left[\left(\hat{p}_{\mu\pm}^{(B\pm)}\right) \times \left(\hat{p}_{\mu\mp}^{(B\pm)}\right)\right] \cdot \left[\left(\hat{p}_{K\pm}^{(B\pm)}\right) \times \left(\hat{p}_{\pi^0}^{(B\pm)}\right)\right] ,\\
\sin\phi &= \left[\left(\hat{p}_{\mu\pm}^{(B\pm)}\right) \times \left(\hat{p}_{\mu\mp}^{(B\pm)}\right)\right] \times \left[\left(\hat{p}_{K\pm}^{(B\pm)}\right) \times \left(\hat{p}_{\pi^0}^{(B\pm)}\right)\right] \cdot \left(\hat{p}_{K^{*\pm}}^{(B\pm)}\right) .
\end{aligned} \tag{3.1}$$
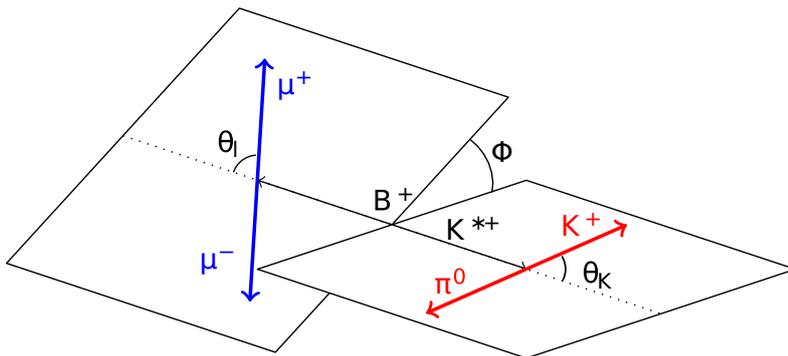
Fig. 3.1: Definition of the angles in the $B^+ \to K^{*+}\mu^+\mu^-$ decay. The angle θ_L is defined as the angle between the μ^+ flight direction in the $\mu^+\mu^-$ rest frame and the flight direction of the $\mu^+\mu^-$ pair in the B^+ meson rest frame. Similarly, θ_K is defined as the angle between the K^+ in the rest frame of K^{*+} and the flight direction of K^{*+} in the B^+ meson rest frame. Finaly, the angle ϕ is the angle between the normal vector of the $K^+\pi^0$ system and the normal vector of the $\mu^+\mu^-$ system.

3.2 Differential decay rate

As mentioned in the previous section, the differential decay rate of $B^+ \to K^{*+}\mu^+\mu^-$ can be fully expressed using only four variables: the dimuon invariant mass squared q^2 and the angles θ_K, θ_L and ϕ. The decay rate then takes the form of

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_K \mathrm{d}\cos\theta_K \mathrm{d}\phi \mathrm{d}q^2} = \frac{9}{32\pi}\sum_i J_i\left(q^2\right) f_i\left(\cos\theta_L, \cos\theta_K, \phi\right), \qquad (3.2)$$

where f_i are angular terms only depending on the three decay angles. They are given purely by the spin structure of the decay. The coefficients J_i are angular observables depending on the dimuon invariant mass squared q^2. They 'control' how much the different f_i terms contribute to the differential decay rate. For the full form, see App. A, Eq. A.1. The J_i coefficients are directly connected to the Wilson coefficients listed in Tab. 1.1. For the exact relation formulas, see Ref. [129].

Similarly, the decay rate of $B^- \to K^{*-}\mu^+\mu^-$ can be expressed in terms of $\bar{J}_i$. Given the CP asymmetry of this decay in the Standard Model is negligibly small, it is useful to define a set of *CP averaged* angular observables S_i instead of having a set of J_i and $\bar{J}_i$:

$$S_i = \frac{J_i + \bar{J}_i}{\Gamma + \bar{\Gamma}}. \qquad (3.3)$$

The dependence of the decay rate on the dimuon invariant mass squared q^2 is sketched in Fig. 1.4. At $q^2 \approx 0$ the dominating operator is $\mathcal{O}_7$. Then, between $1\,\mathrm{GeV}^2 \lesssim q^2 \lesssim 8\,\mathrm{GeV}^2$, the interference of $\mathcal{O}_7$ and $\mathcal{O}_9$ plays a role. In the region of charm resonances, the decay

is dominated by tree-level diagrams. Above the resonances, $q^2 \gtrsim 15\,\mathrm{GeV}^2$, the operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$ dominate. The observables are measured in bins of q^2. In each bin, the q^2-averaged observables are defined as

$$\langle S_i \rangle (q_{min}, q_{max}) = \frac{\int_{q_{min}}^{q_{max}} \mathrm{d}q^2 \left(J_i + \bar{J}_i \right)}{\int_{q_{min}}^{q_{max}} \mathrm{d}q^2 \frac{\mathrm{d}\left(\Gamma + \bar{\Gamma} \right)}{q^2}} \ . \tag{3.4}$$

Following Eq. 3.2, the available S_i parameters are $S_{1s,6s}$ and $S_{3,4,5,7,8,9}$. This basis is also convenient from the experimental point of view: as the $B^+ \to K^{*+}\mu^+\mu^-$ decay is a rare decay, measuring the S_i rather than the J_i and the $\bar{J}_i$ observables effectively doubles the signal yield. The S_i observables are (linearly) connected to two historical observables: the forward-backward asymmetry of the $\mu^+\mu^-$ pair, A_{FB}, and longitudinal polarization of K^*, F_L:

$$\begin{aligned} F_L &= 1 - \frac{4}{3} S_{1s} \, , \\ A_{FB} &= \frac{3}{4} S_{6s} \end{aligned} \tag{3.5}$$

As mentioned in Sec. 1.5, there is a non-negligible form-factor contribution to the decay rate. The influence of form-factor uncertainties can be transformed in a way that the theoretical uncertainties mostly cancel when studying a *single* parameter. The uncertainties are then shifted to other observables. Taking into account all correlations between the angular moments, this basis does not bring any advantage. In the scope of this work the main advantage of this basis is the possibility of a direct comparison to previous LHCb measurements and measurements of the angular observables performed by other experiments. The basis is expressed as a set of P_i' observables and F_L:

$$\begin{aligned} P_1' &= \frac{S_3}{1 - F_L} \, , & P_4' &= \frac{S_4}{\sqrt{F_L (1 - F_L)}} \, , \\ P_2' &= \frac{S_{6s}}{1 - F_L} \, , & P_5' &= \frac{S_5}{\sqrt{F_L (1 - F_L)}} \, , \\ P_3' &= \frac{S_9}{1 - F_L} \, , & P_6' &= \frac{S_7}{\sqrt{F_L (1 - F_L)}} \, , \\ & & P_8' &= \frac{S_8}{\sqrt{F_L (1 - F_L)}} \end{aligned} \tag{3.6}$$

3.3   S-wave contribution

The decay rate, as described by Eq. A.1, takes into account only the decay $B^+ \to K^{*+}\mu^+\mu^-$, where the K^* decays via $K^{*+} \to K^+\pi^0$. This is referred to as the P-wave. However, in the measurement one has to consider the possible contributions from other higher K^{*+} resonances *e.g.* $K^{*+}(1430)_0$ (S-wave). As the K^{*+} resonance is very wide in mass, it is very difficult to distinguish the S-wave component from the P-wave in the selection process. By considering only the events with the reconstructed $K^{*+}(892)$ mass being close to the $K^{*+}(892)$ rest mass ($|m_{K^*}^{reco} - m_{K^*}| < 100\,\mathrm{MeV}$), the S-wave contribution is suppressed, but not fully eliminated. The S-wave component has a different angular structure and therefore pollutes the angle θ_K, θ_L, ϕ distributions. The decay rate of the S-wave is

$$\frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{dcos}\theta_L\,\mathrm{dcos}\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{S}} = \frac{3}{16\pi}F_S\sin^2\theta_L\,. \tag{3.7}$$

However, both P-wave and S-wave amplitudes are complex. The addition of them results in interference term:

$$\frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{dcos}\theta_L\,\mathrm{dcos}\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{PS}} = \frac{3}{16\pi}\sum_1^5 S_{Si}f_{S_i}(\theta_L, \theta_K\,\phi)\,. \tag{3.8}$$

The full expression is given in Eq. A.5. For the full angular description of the $B^+ \to K^{*+}\mu^+\mu^-$ decay, both the S-wave and the interference terms have to be added to Eq. 3.2.

3.4   Angular observables

Using P-wave term from Eq. 3.2, S-wave term from Eq. 3.7, and their interference term Eq. 3.8 it is possible to construct the full angular description of $B^+ \to K^{*+}\mu^+\mu^-$. The full procedure is described in App. A. The full decay description is then given by Eq. 3.9. In total, there are eight moments related to the P-wave contribution and six moments related to the S-wave contribution and its interference with the P-wave. The observables are for readers convenience shown in red. Each of the observable is measured in bins of q^2. In order to measure all these variables, the statistical power of the measured sample is required to be rather large.

$$
\begin{aligned}
\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_L\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{S+P}} &= \frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_L\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{P}} + \frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_L\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{S}} + \frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_L\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{PS}} \\
&= (1-F_S)\frac{9}{32\pi}\bigg[\tfrac{3}{4}(1-F_L)\sin^2\theta_K \qquad + F_L\cos^2\theta_K \\
&\quad +\tfrac{1}{4}(1-F_L)\sin^2\theta_K\cos 2\theta_L \quad - F_L\cos^2\theta_K\cos 2\theta_L \\
&\quad +S_3\sin^2\theta_K\sin^2\theta_L\cos 2\phi \qquad + S_4\sin 2\theta_K\sin 2\theta_L\cos\phi \\
&\quad +S_5\sin 2\theta_K\sin\theta_L\cos\phi \qquad + \tfrac{3}{4}A_{FB}\in^2\theta_K\cos\theta_L \\
&\quad +S_7\sin 2\theta_K\sin\theta_L\sin\phi \qquad + S_8\sin 2\theta_K\sin 2\theta_L\sin\phi \\
&\quad +S_9\sin^2\theta_K\sin^2\theta_L\sin 2\phi\bigg] \\
&\quad +\frac{3}{16\pi}\bigg[F_S\sin^2\theta_L \qquad\qquad + S_{S1}\sin^2\theta_L\cos\theta_K \\
&\quad +S_{S2}\sin 2\theta_L\sin\theta_K\cos\phi \qquad + S_{S3}\sin\theta_L\sin\theta_K\cos\phi \\
&\quad +S_{S4}\sin\theta_L\sin\theta_K\sin\phi \qquad + S_{S5}\sin 2\theta_L\sin\theta_K\sin\phi\bigg].
\end{aligned}
$$

$$(3.9)$$

3.5   Folding of angles

In total, there are 14 angular observables for each q^2 bin. The full description also requires the parametrization of the mass distributions of the B^+ and K^* mesons, adding more free parameters to the decay description. Moreover, the background contribution needs to be modeled in all dimensions. The expected signal yield in this rare decay is in the order of less than a hundred events in each q^2 bin. This can result in a great instability in the data fit.

In order to improve the stability, a dedicated procedure is implemented. The *folding* of angles is an angular transformation exploiting the symmetry of the angular functions in Eq. 3.9. An example is a tranformation of $\phi \to \phi + \pi$ for signal candidates with $\phi < 0$. This results in canceling out the terms dependent on $\cos\phi$ and $\sin\phi$ and leaves the Eq. 3.9 only with observables F_L, S_3, A_{FB}, S_9 (and F_S and S_{S1}). This method has been successfully applied in previous measurements, see *e.g.* Ref. [130].

Using a total of five foldings, all observables of interest are accessible. They are listed in Eq. 3.10. This procedure greatly increases the stability of the data fit as most of the observables are canceled out. The price to pay is the loss of information about correlations between the observables.

folding 0:
$$\phi \rightarrow \phi + \pi \qquad \text{for } \phi < 0$$

folding 1:
$$\phi \rightarrow -\phi \qquad \text{for } \phi < 0$$
$$\phi \rightarrow \pi - \phi \qquad \text{for } \cos(\theta_L) < 0$$
$$\cos(\theta_L) \rightarrow -\cos(\theta_L) \qquad \text{for } \cos(\theta_L) < 0$$

folding 2:
$$\phi \rightarrow -\phi \qquad \text{for } \phi < 0$$
$$\cos(\theta_L) \rightarrow -\cos(\theta_L) \qquad \text{for } \cos(\theta_L) < 0$$

$$(3.10)$$

folding 3:
$$\cos(\theta_L) \rightarrow -\cos(\theta_L) \qquad \text{for } \cos(\theta_L) < 0$$
$$\phi \rightarrow \pi - \phi \qquad \text{for } \phi > \pi/2$$
$$\phi \rightarrow -\pi - \phi \qquad \text{for } \phi < -\pi/2$$

folding 4:
$$\cos(\theta_K) \rightarrow -\cos(\theta_K) \qquad \text{for } \cos(\theta_L) < 0$$
$$\cos(\theta_L) \rightarrow -\cos(\theta_L) \qquad \text{for } \cos(\theta_L) < 0$$
$$\phi \rightarrow \pi - \phi \qquad \text{for } \phi > \pi/2$$
$$\phi \rightarrow -\pi - \phi \qquad \text{for } \phi < -\pi/2$$

A tabular overview of the sensitivity of the angular folding to P-wave angular moments is presented in Tab. 3.1. Using all the five angular foldings gives access to all eight P-wave angular moments. The S-wave angular moments sensitivity is shown in Tab. 3.2.

observable	moment	0	1	2	3	4
S_{1s}	$\cos^2 \theta_K$	✓	✓	✓	✓	✓
S_3	$\sin^2 \theta_K \sin^2 \theta_L \cos 2\phi$	✓	✓	✓	✓	✓
S_4	$\sin 2\theta_K \sin 2\theta_L \cos \phi$	-	✓	-	-	-
S_5	$\sin 2\theta_K \sin \theta_L \cos \phi$	-	-	✓	-	-
S_{6s}	$\sin^2 \theta_K \cos \theta_L$	✓	-	-	-	-
S_7	$\sin 2\theta_K \sin \theta_L \sin \phi$	-	-	-	✓	-
S_8	$\sin 2\theta_K \sin 2\theta_L \sin \phi$	-	-	-	-	✓
S_9	$\sin^2 \theta_K \sin^2 \theta_L \sin 2\phi$	✓	-	-	-	-

Tab. 3.1: Angular folding sensitivity to P-wave angular moments.

observable	moment	0	1	2	3	4
F_S	$\sin^2\theta_L$	✓	✓	✓	✓	✓
S_{S1}	$\cos(\theta_K)\sin^2\theta_L$	✓	✓	✓	✓	-
S_{S2}	$\sin\theta_K \sin 2\theta_L \cos\phi$	-	✓	-	-	-
S_{S3}	$\sin\theta_K \sin\theta_L \cos\phi$	-	-	✓	-	-
S_{S4}	$\sin\theta_K \cos\theta_L \sin\phi$	-	-	-	✓	✓
S_{S5}	$\sin\theta_K \sin 2\theta_L \sin\phi$	-	-	-	-	-

Tab. 3.2: Angular folding sensitivity to S-wave angular moments.

3.6 Previous measurements

Experimentally, there are two main ways of studying the $b \to s\ell^+\ell^-$ transitions: measurements of branching ratios, and of angular observables[18]. First measurements of branching fractions were agreeing with the SM predictions [131, 132, 133], as the statistical power of the measurements did not allow for precision measurements. However, with more available data first discrepancies started to appear, such as in the differential branching fractions measurement of $B^+ \to K^{(*)}\mu^+\mu^-$ decays [1] or in the $B_s^0 \to \phi\mu^+\mu^-$ branching fraction measurement [2].

A discrepancy between a measurement and the SM predictions of angular observables appeared already in 2013, when LHCb analyzed $1\,\text{fb}^{-1}$ of data in the decay of $B^0 \to K^{*0}\mu^+\mu^-$ [5]. One out of 24 measurements (four P' parameters in six bins of q^2) is $3.7\,\sigma$ away from the SM prediction. The parameter is P_5'. If there is a New Physics contribution in the Wilson Coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10,}$ it would show first in the P_5': this discrepancy sparked a lot of interested.

Since then, many similar measurements have been performed [3, 6, 7, 8, 9]. These measurements are summarized in Fig. 3.2. The latest LHCb result using the $B^0 \to K^{*0}\mu^+\mu^-$ decay [3] is yet not present in the figure. In this last measurement, the P_5' discrepancy in low q^2 increased from $2.4\,\sigma$ in Ref. [6] to $2.8\,\sigma$.

Moreover, the first angular study of $B^+ \to [K_S^0\pi^+]_{K^{*+}}\mu^+\mu^-$ at LHCb [4] has been recently published. The P_5' measured in eight q^2 bins is shown in Fig. 3.3. A global evaluation of the result in terms of the real part of the Wilson coefficient $\mathcal{C}_9$ prefers a shift of $\text{Re}(\mathcal{C}_9)$=-1.9 from the Standard Model value with a significance of 3.1 standard deviations.

---

[18]$b \to s\ell^+\ell^-$ transitions are also an important tool in studying lepton flavor universality. However, this is beyond the scope of this work.
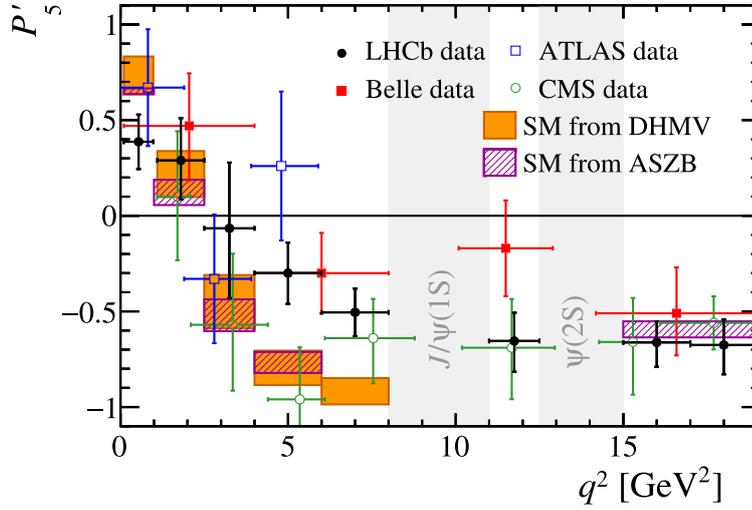
Fig. 3.2: Measurements of the optimized angular observable P_5' in bins of q^2. The shaded areas represent charmonium resonances that are dominated by tree-level diagrams. Experimental results are taken from Ref. [6, 7, 8, 9], theory predictions are taken from Ref. [134, 135, 136, 137].
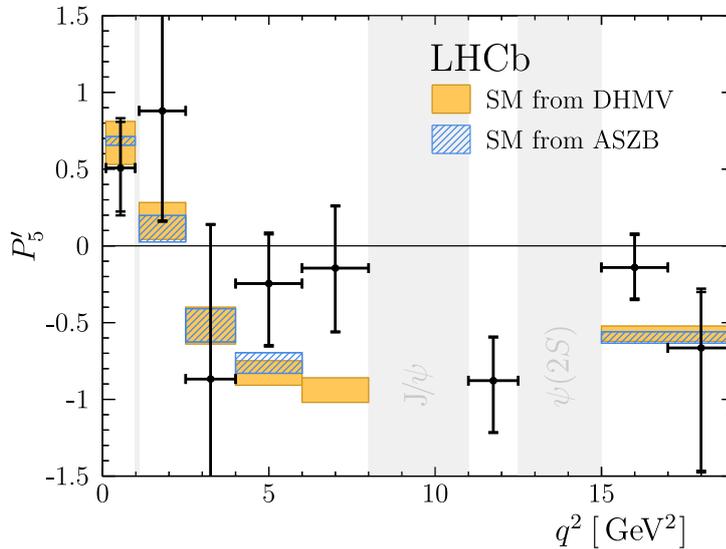
Fig. 3.3: Measurements of the optimized angular observable P_5' in bins of q^2 from LHCb in decay of $B^+ \to [K_S^0 \pi^+]_{K^{*+}} \mu^+ \mu^-$. The shaded areas represent charmonium resonances that are dominated by tree-level diagrams and ϕ pollution in the region around $1\,\text{GeV}^2$. Experimental results are taken from Ref. [4], theory predictions are obtained from Ref. [136, 138] using the FLAVIO package [139].

It is clear there are numerous independent measurements in tension with the Standard Model in the order of 2-3 σ. Similar tensions are also observed in the lepton flavor universality tests [140, 141, 142]. Moreover, there are similar tensions outside of the $b \rightarrow s\ell^+\ell^-$ scope, such as the R_D and R_{D^*} measurement [143] or the $R_{J/\psi}$ measurement [144]. They all are consistent with each other, painting a picture of possible New Physics contribution. Specifically, New Physics contribution to the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ may contribute to the anomalies [145]. An example of a global fit to all these measurements is shown in Fig. 3.4. Further measurements, such as the work presented here, and improved theory calculations will cast light on these tensions in the near future.
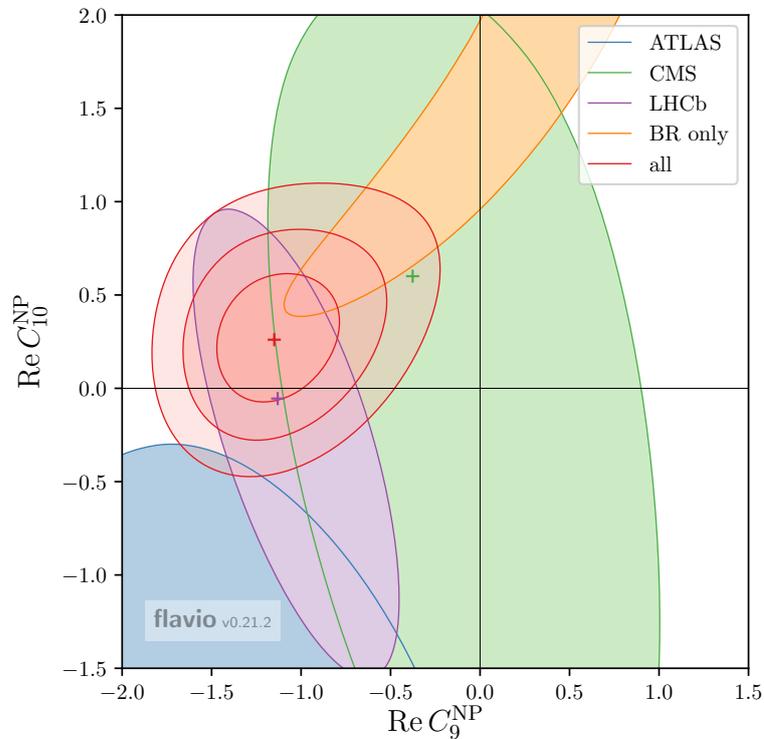
Fig. 3.4: Constraints to the New Physics contribution to Wilson coefficients C_9 and C_{10} taken from [146]. All other Wilson coefficients are assumed to have Standard Model values. The bands represent the constrains from $B \rightarrow K^*\mu^+\mu^-$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ measurements performed by listed collaborations, the countours represent one standard deviation σ. Branchingratio only measurements are shown as the yellow band. The global fit of these results is represented in red with the one, two and three σ contours. In the case of no New Physics contribution, the C_9^{NP} and C_{10}^{NP} are equal to zero. Note that the global fit is however dominated by the LHCb results and CMS measurements are compatible with the Standard Model. For the details about the global fit procedure, see Ref. [146].

4   Introduction to the analysis

The aim of this thesis is to study angular parameters S_i and P_i' in the decay of $B^+ \to K^{*+}\mu^+\mu^-$ where $K^{*+} \to K^+\pi^0$. This chapter gives an overview of the analysis procedure. The biggest obstacle in this analysis is the presence of a neutral pion in the final state. Therefore, the reconstruction of the neutral pion is discussed in detail. Then, as discussed thoroughly in Sec. 1.5, the split of data into different classes based on the dimuon invariant mass squared q^2 is explained. Lastly, the simulation samples used in this analysis are listed.

4.1   Analysis strategy

The analysis uses data from the full Run I and Run II data taking periods. During this time, the LHCb experiment collected a dataset corresponding to an integrated luminosity of $9\,\mathrm{fb}^{-1}$. The integrated luminosity $\int \mathcal{L}$, the beam energy E_{BEAM} and the center-of-mass energy $\sqrt{s}$ for each data-taking year is given in Tab. 4.1.

In the first step of this analysis the recorded data is filtered to select only events containing the signal decay. The decay of $B^+ \to K^{*+}\mu^+\mu^-$ is a rare decay with a branching fraction $\mathcal{B}_{B^+\to K^{*+}\mu^+\mu^-} = (9.6 \pm 1.0) \times 10^{-7}$ [33]. In order to get the full branching fraction of the $B^+ \to (K^{*+} \to K^+\pi^0)\,\mu^+\mu^-$ decay, the branching fraction is multiplied by the branching fraction of $\mathcal{B}_{K^{*+}\to K^+\pi^0} = 0.5$ [33]. This leads to very strict requirements on data selection: the background rejection needs to be as high as possible while keeping high signal selection efficiency.

At first, the data is selected centrally: events have to pass the trigger (online) selection and then the centrally-processed selection called *stripping*. After this, specific preselection cuts are applied. The remaining background is studied and identified with the help of

year	$\int \mathcal{L}$ [fb^{-1}]	E_{BEAM} [TeV]	$\sqrt{s}$ [TeV]
2011	1.11	3.5	7.0
2012	2.08	4.0	8.0
2015	0.33	6.5	13.0
2016	1.67	6.5	13.0
2017	1.71	6.5	13.0
2018	2.19	6.5	13.0

Tab. 4.1: Data recording conditions for LHCb in the years 2011-2018. For each year, the recorded integrated luminosity $\int \mathcal{L}$, beam energy E_{BEAM} and center-of-mass energy $\sqrt{s}$ are given.

multi-variate classifiers. The selection is validated using the $B^+ \to K^{*+}J/\psi$ decay. This *reference* decay has the same final-state particles as the *signal* decay $B^+ \to K^{*+}\mu^+\mu^-$ and very similar kinematics. Moreover, due to the branching ratio of $\mathcal{B}\,(B^+ \to K^{*+}J/\psi,$ $J/\psi \to \mu^+\mu^-)= (1.43 \times 10^{-3}) \times (5.96 \times 10^{-2}) = 8.52 \times 10^{-5}$ [33], the reference channel is ~ 200 times more abundant than the signal channel.

After choosing the optimal selection criteria, the angular acceptance correction is applied. The LHCb acceptance covers only the forward region. Moreover, all the subdetectors cover different regions in phase-space. For the accurate measurement of angles, angular acceptance corrections are crucial.

Lastly, the measured angle distributions are fitted in order to extract the S_i parameters. As one can see from Eq. 3.9, the distribution is not trivial and the fit requires careful approach, especially given the limited size of the sample. The fit is done in the reconstructed mass of the B^+ meson m_{B^+}, the reconstructed mass of the K^{*+} meson m_{K^*} and $\cos(\theta_K)$, $\cos(\theta_L)$ and ϕ dimensions. The reference channel is used to validate the fitter framework. A pseudoexperiment study is performed in order to examine the possible sensitivity of this measurement.

4.2 The neutral pion reconstruction

It is mentioned in Sec. 3 that the available measurements of the $B^+ \to K^{*+}\mu^+\mu^-$ decay are not as extensive as the measurements of the $B^0 \to K^{*0}\mu^+\mu^-$ decay. The main reason is the intrinsic property of all high-energy detectors: the reconstruction of neutral particles is non-trivial and therefore not very effective. As K^{*0} decays to $K^+\pi^-$, the final state contains four charged particles. The detection of those particles is rather simple[19]. In the case of K^{*+} meson, it either decays into $K^0\pi^+$ or $K^+\pi^0$. For LHCb, the only relevant K^0 meson is K^0_S meson, as K^0_L meson is stopped in the HCAL without leaving any signal in the tracking detectors[20]. K^0_S mesons on the other hand decay fast enough into a $\pi^+\pi^-$ pair, so the LHCb tracking system is able to register both charged pions[21].

This analysis focuses on the decay channel where $K^{*+} \to K^+\pi^0$. The K^+ meson is detected by the tracking detectors. The π^0 meson typically decays into a $\gamma\gamma$ pair (branching ratio of $\pi^0 \to \gamma\gamma$ is $\simeq 98.8\%$ [33]). Both photons are registered by the electromagnetic calorimeter. Due to the finite granularity of ECAL, the two photons can either be registered by one or two ECAL cells. A sketch of this is shown in Fig. 4.1.

---

[19]K^{*0} of course also decays to $K^0\pi^0$, which is extremely difficult to reconstruct. Therefore it is typically omitted from the measurements.

[20]Assuming a boost of $\gamma = 20$ ($v \simeq 0.999c$), a free-flying K^0_L would decay after traveling $300\,\mathrm{m}$.

[21]Once again assuming $\gamma = 20$, K^0_S decays after $53\,\mathrm{cm}$. This means depending on the K^0_S boost, it either decays inside or outside of VELO, leading to more complications in analyzing this subdecay.

The neutral pions reconstructed from photons hitting one ECAL cell are called *merged* pions, the ones reconstructed from two cells *resolved* pions.

For the reconstruction of the π^0 meson, proper settings of the electromagnetic calorimeter are essential. This is done in three steps: initial adjustment of ECAL energy scale, energy flow calibration and fine calibration of the ECAL cells [147]. The methods used for calibra-

Fig. 4.1: Illustration of resolved (left) and merged (right) π^0 mesons reconstruction in the ECAL cells.

tion are essentially the same in Run I and Run II. The main changes in Run II are the full automation of the calibration process and skipping the intermediate step.

The initial adjustment of the energy scale is done by adjusting the photo-multipliers' (PMTs) gain using the ECAL's LED monitoring system. A LED light is attached to PMTs generating a known signal. The voltage of the PMTs is adjusted to match the measured and the known signal. This adjustment leads to a precision of the cell-to-cell inter-calibration of 10%. The reason for this uncertainty is the dispersion in the photoelectron yields and the accuracy of the light yield determination. The LED-based calibration is preformed approximately once a week.

Then, the energy flow calibration is performed. This is done in order to smooth the fluctuations in the flux among neighboring cells due to initial miscalibrations. The method is rather simple: one exploits the symmetry of the energy flow of the calorimeter surface [148]. Simulations with known mis-calibration showed that the flux adjustments improves the calibration by a factor of ~ 3, assuming an initial precision of the calibration of 10%.

Aging negatively affects the ECAL's performance and has to be accounted for. This can be nicely seen from the time variation (decrease) of the reconstructed π^0 mass presented in Fig. 4.2. To account for this effect, fine calibration exploiting the π^0 mass is performed. The mean $m_{\pi^0}^{reco}$ is obtained from $\gamma\gamma$ pairs from minimum-bias events[22] with low multiplicity to remove possible pile-up events. The photons are reconstructed using 3×3 clusters with single photon signals, where the cell with the highest energy deposit is called seed. The seeds are then corrected to match the nominal π^0 mass. The effect of this correction is depicted in Fig. 4.2. This calibration is performed every LHC-runnning month.
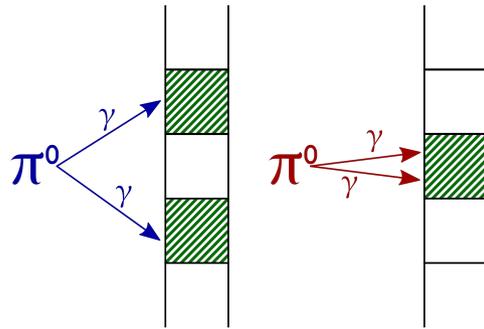
---

[22]Minimum-bias events are events with at least one charged track in the VELO detector or the downstream tracking system.
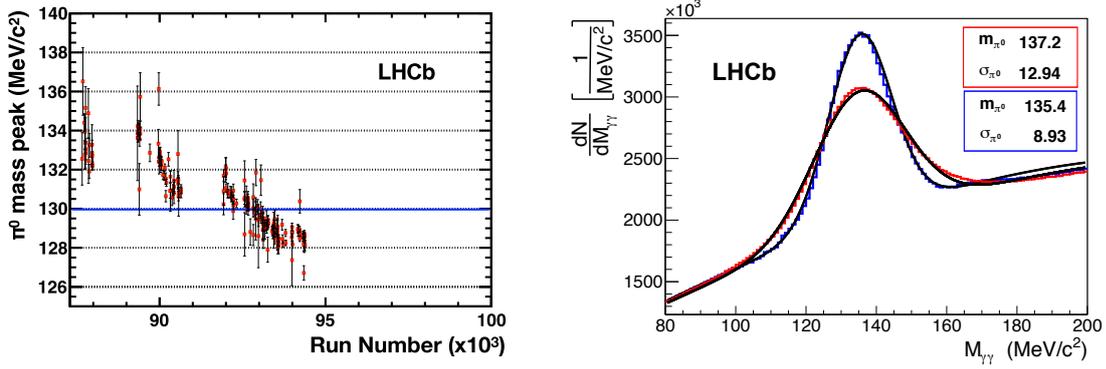
Fig. 4.2: On the left, fitted neutral pion mass m_{π^0} as a function of run number (time) using 2011 data. The π^0 mass is 135 MeV [33]. The clear decrease in the m_{π^0} value is due to the ECAL ageing. On the right, invariant mass distribution for $\pi^0 \to \gamma\gamma$ candidates used for the fine calibration. The red curve corresponds to the distribution before applying the fine calibration, the blue curve is the final m_{π^0} distribution. Values in the boxes are the mean and width of the signal peak distribution in MeV before (red box) and after (blue box) applying the m_{π^0} calibration. Taken from Ref. [149].

For this analysis, only resolved π^0 mesons are used. Merged π^0 mesons tend to have higher momenta (as the higher boost results in more collimated photons). In this work, where the π^0 mesons come from a K^{*+} meson, the statistical contribution of these events is low. The π^0 mesons originating from the K^* have typically transverse momentum of a few gigaelectronvolts. In Fig. 4.3 left the higher abundance of resolved π^0 mesons at lower momentum is shown. As merged and resolved π^0 require their own careful approach, merged π^0 mesons are not included in the analysis.
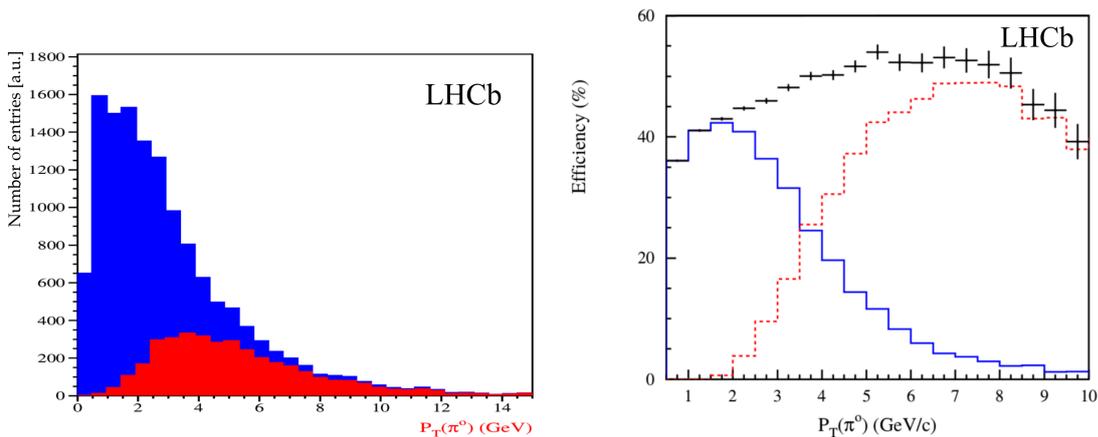
Fig. 4.3: On the left, transverse momentum distributions of merged (red) and resolved (blue) π^0 in the LHCb acceptance originating from $B^0 \to \pi^+\pi^-\pi^0$ decay. On the right, the overall merged (dashed line) and resolved (full line) π^0 efficiency (number of identified π^0/ number of π^0 in detector acceptance with $p_T^{\pi^0} > 200$ MeV). The black points represent the overall efficiency for both resolved and merged neutral pions. Taken from Ref. [149].

The mass resolution of a decay including a π^0 meson can be seen in Fig. 4.4. In this case, the $D^0 \to K^-\pi^+\pi^0$ candidate mass is reconstructed using resolved and merged pions. It is clearly visible that the mass resolution of the D^0 meson candidate is better for resolved π^0 meson.

Moreover, the ECAL resolution effects come into play. The resolution of ECAL is $\sigma/E = 0.1/\sqrt{E} \oplus 0.01$, which is a very good resolution for a sampling calorimeter. The advantage is that the resolution *decreases* with *increasing* deposited energy. However, for low-energy photons this does not bring any asset.

To summarize, we expect the neutral pion mass peak to be wide and to be affected by the combinatorial background coming from the ECAL cells. As we focus on resolved π^0, combinatorial background contributions from $\gamma\gamma$ combinations are expected. On the other hand, the usage of resolved π^0 improved the particle-identification as we have information from two cells: the probability of misidentifying a random photon as a π^0 meson is lower. The maximal efficiency to reconstruct resolved π^0 mesons is $\sim 40\%$ at low $p_T^{\pi^0}$.
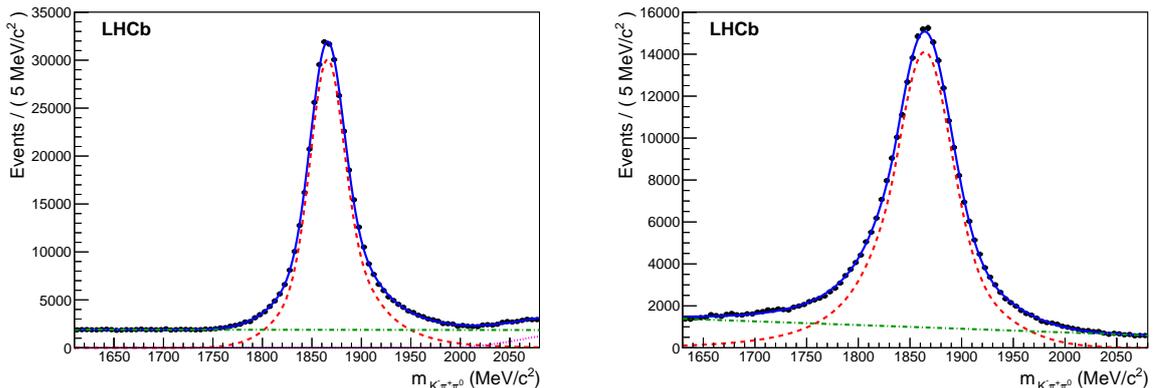
Fig. 4.4: Mass distribution of the reconstructed $D^0 \to K^-\pi^+\pi^0$ candidates with resolved π^0 (left) and merged π^0 (right) obtained from the 2011 data sample. The blue curve corresponds to a fit. The signal component of the fit function (red dashed line) and the background (green dash-dotted line) contributions are shown. One can easily see the mass resolution of the D^0 candidate is much worse for merged π^0. Taken from Ref. [149].

4.3 Binning in the dimuon invariant mass

It is explained in Sec. 1.5 that different physics processes dominate in different q^2 (dimuon invariant mass squared) regions. Therefore, a veto in the resonance regions is applied to eliminate $B^+ \to K^{*+}J/\psi$ and $B^+ \to K^{*+}\psi(2S)$ decay contributions. Moreover, an additional veto to eliminate the rare $B^+ \to K^{*+}\phi$ decays is introduced. All three resonances J/ψ, $\psi(2S)$ and ϕ are indistinguishable from the signal as they have a very short decay time and therefore they are not displaced enough from the B^+ vertex.

The range of the measured q^2 is limited by the rest mass of the two muons value and by the difference of the B^+ and K^{*+} mass squared $(m_{B+} - m_{K^{*+}})^2 = 19.25 \, \text{GeV}^2$. However, as the LHCb acceptance at very high q^2 is low, the upper limit in this measurement is set to $18 \, \text{GeV}^2$.

In this work, q^2 is segmented into four signal regions and the three vetoed regions. As a result of the challenging reconstruction there is a smaller statistical significance in the signal yield compared to the previous analyses [3, 4, 5, 6] and the q^2 segmentation is therefore coarser. These signal regions are larger than the ones chosen in previous analyses. The q^2 bins including the vetoed regions are listed in Tab. 4.2. Furthermore, a wide bin $[1.1, 6.0]$ is added on top of the four q^2 bins. The wide bin is the theoretically favored region to observe New Physics effects and it has been included also in the previous measurements [3, 4, 5, 6].

bin	$q^2 \, [\,\text{GeV}^2\,]$	veto
1	$[0.1, \ 0.98]$	
	$[0.98, \ 1.1]$	ϕ
1	$[1.1, \ 4.0]$	
2	$[4.0, \ 8.0]$	
	$[8.0, \ 11.0]$	J/ψ
3	$[11.0, \ 12.5]$	
	$[12.5, \ 15.0]$	$\psi(2S)$
4	$[15.0, \ 18.0]$	
5	$[1.1, \ 6.0]$	

Tab. 4.2: The binning scheme of the dimuon invariant mass squared q^2 in the angular analysis including the vetoed regions of resonances decaying to $\mu^+\mu^-$ pair. In the first bin, the ϕ resonance is removed.

4.4 Simulation samples

In order to study the effects of the detector response and to identify possible backgrounds, several Monte Carlo simulation samples are used in the analysis. The exhaustive list of the MC samples is presented in Tab. 4.3. The two main samples consist of the signal decay $B^+ \to K^{*+}\mu^+\mu^-$ and the reference decay $B^+ \to K^{*+}J/\psi$.

In order to study the acceptance of the detector, a PHase SPace (PHSP) Monte Carlo sample is used. This sample neglects the spin structure of the decay, reflecting only the kinematic properties of the decay. This effectively means the distributions of the angles θ_K, θ_L, ϕ are flat. Moreover, an additional requirement is imposed on the sample: the dimuon invariant mass squared q^2 distribution is generated to be flat. The sample is used

to understand the angular acceptance in four dimensions of θ_K, θ_L, ϕ and q^2.

Lastly, for the background investigation, an inclusive sample of $B^+ \to X J/\psi$ is used, where X stands for any particle that a B^+ can decay into additionally to the J/ψ meson. This is particularly useful for identifying pollutions from other decays.

MC decay, type	Year	Number of generated events per polarity
$B^+ \to K^{*+} \mu^+ \mu^-$ Signal channel	2011	50 000
	2012	50 000
	2015	100 000
	2016	100 000
	2017	115 000
	2018	120 000
$B^+ \to K^{*+} J/\psi$ Reference channel	2011	100 000
	2012	100 000
	2015	100 000
	2016	100 000
$B^+ \to K^{*+} \mu^+ \mu^-$ Phase space	2011	85 000
	2012	225 000
	2015	95 000
	2016	260 000
	2017	240 000
	2018	290 000
$B^+ \to X J/\psi$ Inclusive sample	2011	250 000
	2012	250 000
	2016	500 000

Tab. 4.3: Monte Carlo simulation samples used in this work. The bending magnet polarity is regurarly flipped during the data taking (see Sec. 2.2.1). Therefore, two samples, one for each polarity configuration, are produced.

5   Event Selection

LHC collisions occur with a frequency of 40 MHz. Storing all the data coming into LHCb would require storing data rates of 1 TB/s, which would require 3.6 petabytes per every hour of collisions[23]. However, only in about one of 400 collisions a $b\overline{b}$ quark pair is produced and the chance of a B meson decaying into $K^*\mu^+\mu^-$ is circa one in a million. Therefore, it is needed to process the data quickly and select only the required events, while maintaining very high purity and efficiency of this selection.

The selection of $B^+ \to K^{*+}\mu^+\mu^-$ candidates is realized in several steps. First, the events have to pass the online (*trigger*) selection. Then, tighter selection criteria are applied in a centralized (*stripping*) selection. The criteria applied in these two steps are common for many LHCb analyses. Next step is an even tighter preselection that is specific for this work. These events are further filtered by utilizing a multi-variate analysis. In order to utilize the simulation sample in the multi-variate analysis, the simulated sample is validated and corrected to match the data. The final selection is validated using available simulation samples and exploiting the reference channel $B^+ \to K^{*+}J/\psi$.

5.1   Trigger selection

The selection of events begins at the hardware level, as described in Sec. 2.2.5. As a first step, the $B^+ \to K^{*+}\mu^+\mu^-$ candidate event has to be triggered by the L0 trigger by identifying a single muon. In HLT1, the event has to be triggered either by a single detached high p_{T} track [151] or a muon [152]. In HLT2, the events have to pass several topological criteria [153] or pass a tighter muon track cut. In Run II, the requirement on a single detached high p_{T} muon in HLT1 is replaced by a more efficient kinematic cut applied on all tracks. Moreover, topological trigger selection algorithms, or lines, using two muons as input are exploited. The full list of LHCb trigger lines used for this analysis is presented in Tab. 5.1. For detailed description of Run I trigger lines see Ref. [154].

The trigger decision can be either *triggered on signal* (TOS) or *triggered independent of signal* (TIS). That means that if the event is TOS, the signal candidate directly affected the trigger decision, while TIS means that the trigger decision is driven by a different element of the event. The simulation of TIS events is rather complicated. As the contribution of the B^+ meson TIS decisions to the signal candidates is negligible, only the B^+ meson TOS decisions are used in this analysis.

---

[23]The whole current LHCb dataset would then contain more than 30 exabytes of data. To put this into perspective, it is estimated that Google, Microsoft, Amazon and Facebook combined together store 1.2 exabytes of data [150].

L0	L0Muon
HLT1	Hlt1TrackAllL0 (Run I)
	Hlt1TrackMVA (Run II)
	Hlt1TrackMuon
HLT2	Hlt2TopoMu2BodyBBDT
	Hlt2TopoMu3BodyBBDT
	Hlt2Topo2BodyBBDT
	Hlt2Topo3BodyBBDT
	Hlt2DiMuonDetached
	Hlt2TopoMuMu3BodyBBDT (Run II)
	Hlt2TopoMuMu2BodyBBDT (Run II)

Tab. 5.1: List of applied trigger requirements. For a detailed definitions of the applied trigger selection algorithms see Ref. [154].

5.2   Central selection (stripping)

As the trigger requires events to pass only basic topological and kinematical constrains, it is necessary to filter the events selected by the trigger lines further. Due to the size of the dataset and due to computational constraints, an additional central selection is applied. This process is called *stripping* and one set of selection algorithms within stripping is called a *line*. Typically, a stripping line is used by several analyses, hence the selection is still rather loose at this step.

The cuts applied in the stripping line used in this analysis are summarized in Tab. 5.2. Most of the requirements are kinematical, however several more specific properties of the candidates are exploited:

- `IsMuon` requires a track to penetrate through the detector up to the muon stations. This reduces the probability of misidentifying a hadron as a muon to 1% while maintaining high efficiency of muon reconstruction [155]. Depending on the momentum of the track, hits in different muon stations are required. The summary of the required hits based on the track momentum is in Tab. 5.3.

track momentum	muon stations hit requirement
$3\,\mathrm{GeV} < p_\mu < 6\,\mathrm{GeV}$	M2 and M3
$6\,\mathrm{GeV} < p_\mu < 10\,\mathrm{GeV}$	M2 and M3 and (M4 or M5)
$p_\mu > 10\,\mathrm{GeV}$	M2 and M3 and M4 and M5

Tab. 5.3: Muon stations required to trigger the IsMuon decision as a function of momentum range. Taken from Ref. [155].
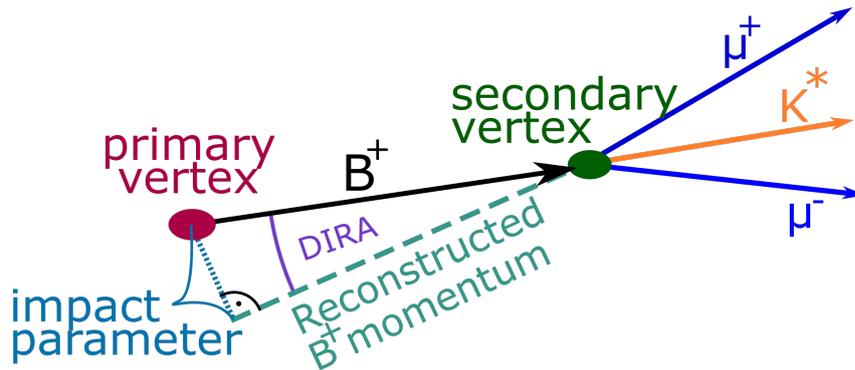
Fig. 5.1: The impact parameter definition in the specific case of $B^+ \to K^{*+} \mu^+ \mu^-$. For readers convenience, the *DIRA* angle of the B^+ meson is also shown.

- Using the established LHCb convention in notation, impact parameter (IP) is the transverse distance of closest approach between a particle trajectory and a vertex. A naive sketch of this quantity is shown in Fig. 5.1.

- *DIRA* angle (direction angle) is the angle between the reconstructed momentum of the particle and the line joining the primary vertex and the B^+ decay vertex. The B^+ meson DIRA angle is shown in Fig. 5.1.

- Despite LHCb convention in notation, χ^2_{FD} is not exactly the χ^2 of the flight distance, but the χ^2 of separation of two vertices. It is calculated as the χ^2 of the common vertex of all tracks minus the sum of χ^2 for two distinct vertices.

- At LHCb, the particles fly through the dipole magnet. Tracks are reconstructed from hits downstream and upstream of the magnet. Due to this 'gap' in the detector, the algorithm matching the tracks from subdetectors downstream and upstream of the magnet might reconstruct a track, which is not induced by a real particle flying through LHCb. These fits have typically low track fit quality. Such tracks are called *ghosts*. A dedicated variable related to the track fit quality with values between 0 and 1, *ghost probability*, is assigned to each track and represents the possibility of the track being a *ghost* track.

5.3 Preselection

The events that pass the stripping are only roughly selected and therefore further cut-based selection needs to be applied. The cuts are listed in Tab. 5.3. Visual illustration of the effect of application of the cuts is shown in Fig. 5.2.

In the case of only charged particles in the decay chain, the decay is reconstructed starting from the most downstream vertex and then built upstream (in this case it would

5.3 Preselection

candidate	selection
$B^{\pm}$	$4700\,\mathrm{MeV} < \mathrm{m}(B^{\pm}) < 7000\,\mathrm{MeV}$
	$\sum_{i \in daughters}$ daughter charge < 3
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 8$
	$\chi^2_{\mathrm{IP}} < 16$ (best PV)
	DIRA angle $< 14\,\mathrm{mrad}$
	$\chi^2_{FD} > 64$[1]
	$\min(\chi^2_{\mathrm{IP}}) > 9.0$
$\mu^{\pm}$	$p_{\mathrm{T}} > 250\,\mathrm{MeV}$
	track *ghost prob* < 0.5
	$\min(\chi^2_{\mathrm{IP}}) > 6.0$[2]
	$\mathrm{DLL}_{\mu\pi} > \text{-}3$[3]
$\mu^+\mu^-$	$m(\mu^+\mu^-) < 7100\,\mathrm{MeV}$[4]
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 12$
	DIRA angle $\in (2.69\,\mathrm{rad}, 3.59\,\mathrm{rad})$
	$\min(\chi^2_{\mathrm{IP}}) > 6.0$[5]
	flight distance $\chi^2 > 9.0$
	`isMuon`
$K^{*\pm}$	$592\,\mathrm{MeV} < \mathrm{m}(K^{*\pm}) < 1192\,\mathrm{MeV}$
K^+	track *ghost prob* < 0.5
	$\min(\chi^2_{\mathrm{IP}}) > 6.0$
	`hasRich`[6]
π^0	$105\,\mathrm{MeV} < \mathrm{m}(\pi^0) < 165\,\mathrm{MeV}$
	$p_{\mathrm{T}}(\pi^0) > 800\,\mathrm{MeV}$[7]
γ	$p_{\mathrm{T}}(\gamma) > 200\,\mathrm{MeV}$
GEC	nSPDHits$<$600
	at least one PV

[1] In S21r0p1 and S29r2 $\chi^2 > 121$.

[2] In S21r0p1 and S29r2 $\min(\chi^2_{\mathrm{IP}}) > 9.0$.

[3] For definition see Sec. 2.2.2.

[4] Only in S24r2, S28r2 and S34r0p1.

[5] In S21r0p1 and S29r2 $\min(\chi^2_{\mathrm{IP}}) > 9.0$.

[6] RICH subsystem registered a track in the candidate event.

[7] In S21r0p1, $p_{\mathrm{T}}(\pi^0) > 700\,\mathrm{MeV}$.

Tab. 5.2: Central selection (stripping) cuts for the `B2XMuMu` line. For each data-taking year, there is a dedicated version of the software. The stripping cuts slightly differ between different version of the software: S21r1p2 (2011), S21r0p1 (2012), S24r2 (2015), S28r2 (2016), S29r2 (2017) and S34r0p1 (2018).

mean starting from the K^* vertex, continuing to the B^+ meson). This means there is no propagation of information from the mother vertices to the daughter particles. In the case of neutral particles, this relation between the mother vertex and the daughter particles can contain a lot of important constraints and improve the event selection. The method exploiting these constrains, Decay Tree Fitter (DTF), was used for the first time by the BaBar collaboration to reconstruct $K^0_S \rightarrow \pi^0\pi^0$ decays [156].
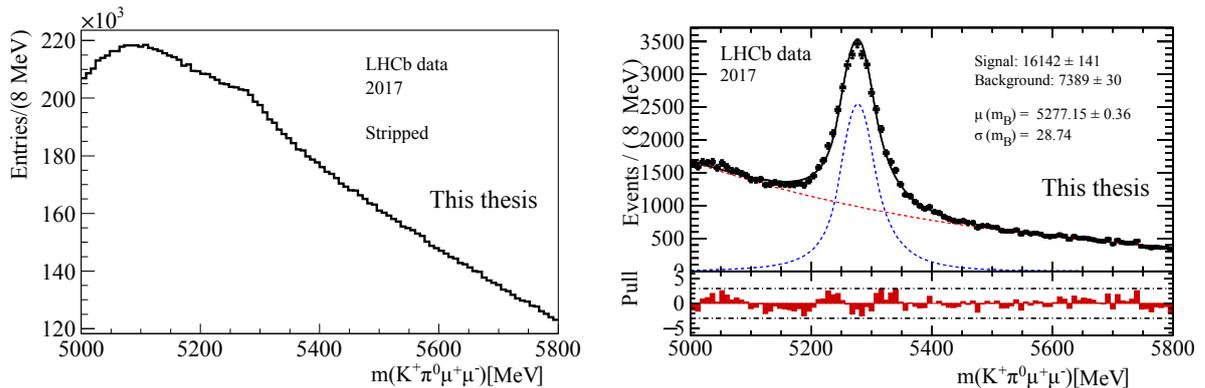
Fig. 5.2: Comparison of B^+ mass distribution before and after applying preselection cuts. On the left, B^+ mass distribution after applying *stripping* cuts is shown, on the right reconstructed B^+ mass after applying preselection cuts. The mean of the signal distribution is denoted $\mu(m_B)$, $\sigma(m_B)$ is the width of the peak. The signal shape is constrained to the signal shape of simulation samples. The signal (blue) is fitted by a two-sided Crystal Ball function (see App. B.1). The background (red) is described by an exponential function.

DTF constrains the mass of the neutral daughter (in our case π^0) and adds this information to the vertex of the mother particle. The decay chain is then parameterized in vertex position, decay length and particle momenta. A simultaneous fit of the decay is performed, taking into account all physical constrains (such as four-momentum conservation). In the case of the decay presented here, the B^+ mass resolution is significantly improved by using DTF. Therefore, momenta and mass variables used for the cut-based preselection are obtained using DTF. Moreover, to remove events where the DTF fit did not converge, only events with DTF status zero (meaning the DTF fit converged) and events with $\chi^2_{DTF} < 200$ are kept.

Moreover, a DTF-like correction to the K^{*+} mass is applied. This is done by fixing the reconstructed B^+ meson mass to its known mass 5279.34 MeV [33]. Then, the π^0 momentum is adjusted according to the fixed B^+ mass. The adjusted π^0 momentum is then used to estimate the K^{*+} mass. This has to be performed in order to remove the effects of the π^0 momentum resolution on the reconstructed K^{*+} mass. Without this adjustment, the description of the K^{*+} mass peak by the Breit-Wigner formula fails.

In order to isolate the reconstructed candidates from nearby tracks, a *cone p_T asymmetry* is defined by Eq. 5.1. The variable $p_T^{B^+}$ denotes the transverse momentum of the reconstructed B^+, while p_T^{cone} is the sum of the transverse momenta of all charged tracks *near* the reconstructed B^+. A *near* track is a track in a cone $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \leq 1.0$, where $\Delta\phi$ is the difference between the track's momentum and the B^+ meson momentum

---

[24]This function is used for partly reconstructed background in B decays, for the definition see App. B.2.
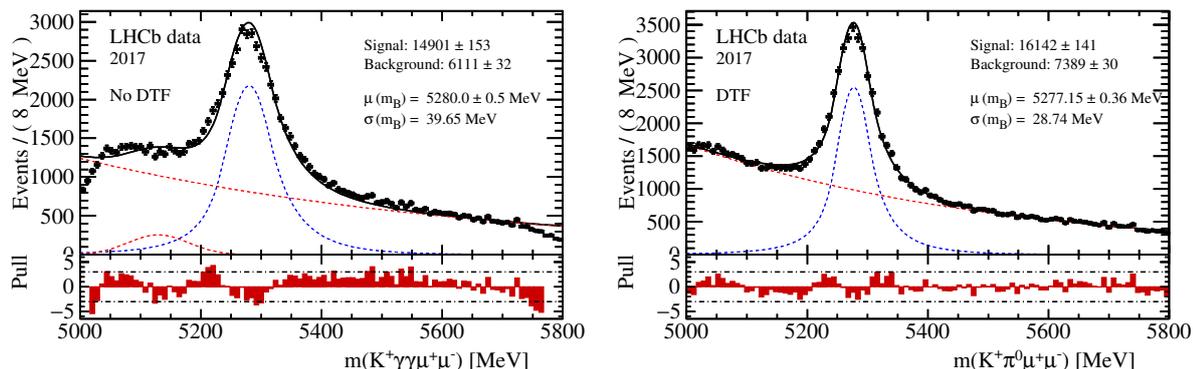
Fig. 5.3: An example of B^+ mass fit to data collected during the 2017 data-taking year. The mean of the signal distribution is denoted $\mu(m_B)$, the width of the peak is denoted $\sigma(m_B)$. The signal shape is fixed to the signal shape of simulated samples. On the left, mass calculated without the DTF is shown, signal (blue) is fitted by a two-sided Crystal Ball function, background (red) consists of an exponential combinatorial background and a function called ExpGaus[24]. The fit does not descibe the data well, the signal peak is rather wide. The sharp drop at 5000 MeV is caused by cutting on mass obtianed by the DTF. On the right, mass obtained using the DTF is shown. Signal (blue) is fitted by a two-sided Crystal Ball function, background (red) consists only of an exponential combinatorial background. The mass resolution improved significantly.

in azimuthal angle and $\Delta\eta$ is the difference in pseudorapidity. The cone p_T asymmetry is then calculated as

$$A_{p_T} = \frac{p_T^{B^+} - p_T^{cone}}{p_T^{B^+} + p_T^{cone}} \ . \tag{5.1}$$

In Sec. 2.2.2, the PID variable DLL is definied. The likelihood information from each PID subsystem (RICH, CALO, MUON) is added linearly, forming a set of combined likelihoods. Final DLL is the likelihood of a given mass hypothesis relative to the pion mass hypothesis. This does not take into account correlations between the subsystems and it does not fully exploit the non-PID information from the subdetectors. Therefore, another variable, *ProbNN* is used. ProbNN combines the PID information from the detectors and the non-PID information in a multi-variate analysis. Therefore, in the cut-based selection, ProbNN variables are used, contrary to the DLL variables in the stripping selection. The ProbNN is calculated for each type of particle, the notation is *e.g.* ProbNNmu for the muon ProbNN.

In the case of photon PID, one relies only on the information from the ECAL. The variable *confidence level* is constructed from the DLL values to indicate the confidence that the chosen assignment of particle ID is correct. It is calculated as the ratio of the likelihood of the chosen hypothesis and the sum of all hypotheses X. In the case of photon it becomes

$$CL_\gamma = \frac{DLL_{\gamma\pi}}{\sum_X DLL_{X\pi}} \ . \tag{5.2}$$

candidate	Selection criterion
B^+	$5.0\,\text{GeV} < m_{B^+} < 5.8\,\text{GeV}$
	$p_\text{T}^{B^+} > 2000\,\text{MeV}$
	Cone-p_T asymmetry $> $ -0.5
	DIRA $> 9\,\text{mrad}$
	$\chi^2_\text{IP} < 12$
	$\chi^2_{FD} > 121$
K^{*+}	$792\,\text{MeV} < m_{K^*} < 992\,\text{MeV}$
	$p_\text{T} > 1350\,\text{MeV}$
	$\chi^2_{FD} > 9$
$\mu^+\mu^-$	Angle between muons > 0.001
	ProbNNmu > 0.25
	$\chi^2_\text{IP} > 9$
	$0.1\,\text{GeV}^2 < q^2 < 21.0\,\text{GeV}^2$, q^2 binned
K^+	ProbNNk > 0.25
	Angle between K^+ and $\pi^0 > 0.001$
π^0	$p_T > 800\,\text{MeV}$
γ	$\text{CL}_\gamma > 0.15$
tracks	$\eta > 1.6$

Tab. 5.3: Preselection cuts.

5.3.1 Charmonium vetoes

The decay rate of a $b \to s\ell^+\ell^-$ transition in dependence on the dimuon mass squared q^2 shows two large excesses, as shown in Fig. 1.4. They are caused by the resonant decays of $J/\psi \to \mu^+\mu^-$ and $\psi(2S) \to \mu^+\mu^-$ coming from $B^+ \to K^{*+}J/\psi$ and $B^+ \to K^{*+}\psi(2S)$ decays. Moreover, a small contribution of $\phi \to \mu^+\mu^-$ from the rare decay $B^+ \to K^{*+}\phi$ is expected. The contributions from these resonances are depicted in Fig. 5.4. The J/ψ and $\psi(2S)$ resonances are clearly dominating the event population. As discussed in Sec. 1.5, these resonances come from tree-level processes and therefore are removed from the selection. The process $B^+ \to K^{*+}\phi$ is strongly influenced by QCD and therefore could potentially pollute the angular distribution and is removed.

5.3.2 $\text{B}^+ \to \text{K}^+\mu^+\mu^-$ veto

The decay of $B^+ \to K^+\mu^+\mu^-$ wrongly associated with an independent π^0 meson mimics the signal. The invariant $K^+\pi^0\mu^+\mu^-$ mass of these candidates is well above the the B^+ meson mass. Therefore, this background is not contributing to the signal. This is shown in Fig. 5.5.
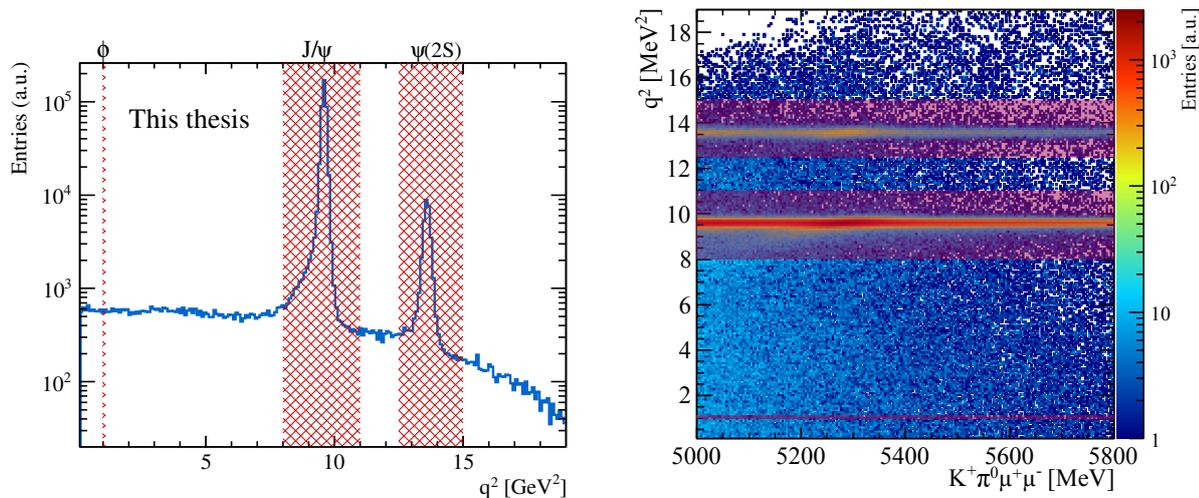
Fig. 5.4: Dimuon invariant mass squared q^2 distribution (left) and q^2 versus $K^+\pi^0\mu^+\mu^-$ invariant mass (right) from the full Run I and Run II dataset. The shaded bands represent the regions surrounding ϕ, J/ψ and $\psi(2S)$ resonances (from the bottom to the top) that are vetoed in the signal selection. The region surrounding J/ψ is further used as a control channel for validation of the fit.

However, the vetoed events account for a big part of the combinatorial background above the B^+ mass. After applying the full selection, the $B^+ \to K^+\mu^+\mu^-$ contribution even dominates this region. In order to suppress this background, a dedicated veto rejecting candidates with $K^+\mu^+\mu^-$ mass close to the B^+ mass, $|m_{B^+} - m_{K^+\mu^+\mu^-}| < 100\,\mathrm{MeV}$, is applied.

5.4   Correction to the simulation

The Monte Carlo simulation sample is used to estimate the background contribution in the data and to account for detector acceptance effects. Therefore, the distributions of variables (and the correlations between them) in data and simulation have to agree. Even though there have been many recent improvements in the Monte Carlo simulation methods, the agreement is not perfect. The main difference between the simulation and the data is multiplicity: in simulation, the underlying event is under-represented.

The correction of the simulation is done by applying weights on the simulation to match the data. In order to obtain the weights, the simulated events have to pass the same selection as the data sample. On top of this selection, only *true* signal candidates have to be selected: the reconstruction algorithms can reconstruct a track that does not correspond to any simulated particle. Such signal candidates have to be removed by the so-called *truth-matching*.
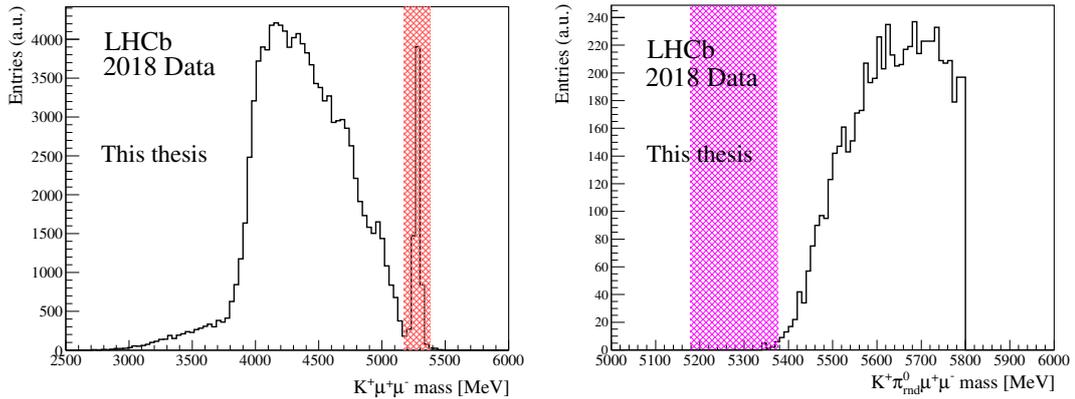
Fig. 5.5: Invatiant mass of $K^+\mu^+\mu^-$ in the 2018 data sample after cut-based selection. On the left, the $B^+ \to K^+\mu^+\mu^-$ mass is shown. There is a clear peak suggesting a contribution of $B^+ \to K^+\mu^+\mu^-$ sample to selected data. The red band represents the region $\pm 100\,\text{MeV}$ around the B^+ mass. These events are vetoed. On the right, the mass of the vetoed $K^+\mu^+\mu^-$ candidates with the associated random π^0 meson is shown. The magenta band shows the region of B^+ meson mass $\pm 100\,\text{MeV}$. The $B^+ \to K^+\mu^+\mu^-$ decay does not contribute to the signal of $B^+ \to K^{*+}\mu^+\mu^-$.

5.4.1 Matching of reconstructed signal candidates to simulated candidates

As there is a neutral pion in the decay chain, it is important to make sure our *true* candidates match the signal candidates we select in data. Events where for instance a photon is converted into an electron or one of the photons in the $\pi^0 \to \gamma\gamma$ decay is randomly assigned is still considered signal, as there is no way to asses the origin of the photon in the data sample. It is clearly visible in Fig. 5.6 there are no structures in the angular distribution of the events with a random photon included. Hence, these events can be considered as signal candidates, as they do not distort the angular distributions.

In order to select the *true* signal candidates, an ID-based selection is applied. Each particle type has it is own unique ID following the Monte Carlo Particle Numbering Scheme [33]. Each generated particle has its *true* ID and the reconstructed ID, based on the PID response of the LHCb detector. The ID-based selection is achieved by comparing the *true* ID of the particles, their mother and their grandmother ID to the reconstructed ID. This check is applied on the whole decay chain $B^+ \to (K^{*+} \to K^+\pi^0)\,\mu^+\mu^-$ except for the photons.

5.4.2 Reweighting and the *sPlot* technique

To account for the simulation imperfections listed above, a correction has to be applied. Very good agreement between the data and the simulation is achieved when the Monte Carlo simulation is weighted in $p_\text{T}^{B^+}$ and `nLongTracks`, which represents the number of

5.4 Correction to the simulation
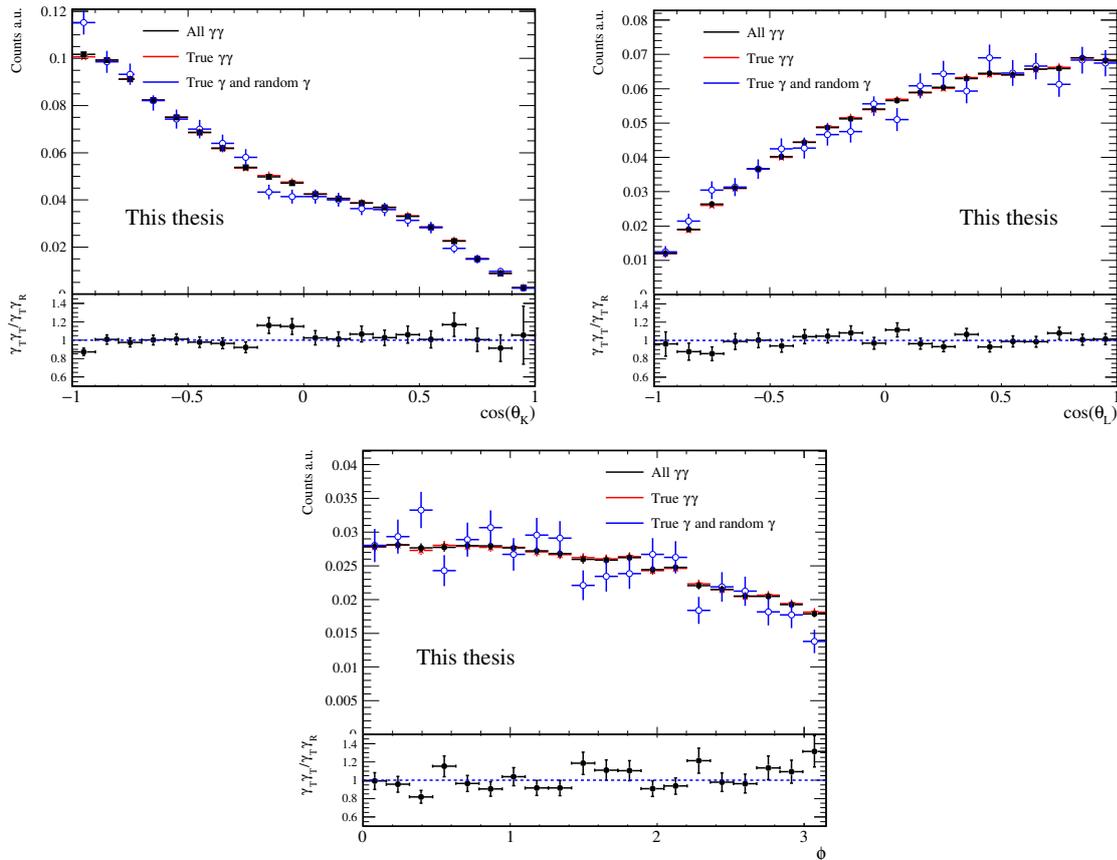
Fig. 5.6: Normalized $\cos(\theta_K)$ (left), $\cos(\theta_L)$ (middle) and ϕ (right) distributions for simulated events where both photons are either coming from $B^+ \to K^{*+}\mu^+\mu^-$, $K^{*+} \to K^+\pi^0$, $\pi^0 \to \gamma\gamma$ or one photon is a random hit in ECAL reconstructed as photon. Black squares note all events passing the *true* ID requirements, excluding photons parents' ID. Red stars are events, where both photons originate from $B^+ \to K^{*+}\mu^+\mu^-$, blue circles are events, where one photon is *true* and one is random. At the bottom of the figures, a ratio of the number of normalized events with only *true* photons over the number of normalized events with one *true* and one random photon is shown. The ratio is consistent with one.

tracks traversing VELO, TT and the T-stations. This number is strongly correlated with the overall event multiplicity. The weighting is performed as two independent weightings, as there is no correlation between $p_T^{B^+}$ and `nLongTracks`, as can be seen in Fig. 5.7.

The weights cannot be calculated directly using the data sample: the simulation sample consists only of signal candidates while in the data sample the background is also present. Therefore, the data sample has to be weighted to mimic the signal as much as possible. This is done using the *sPlot* technique [157, 158]. This technique is used to unfold the signal decay from the background by exploiting likelihood fits. The *sPlot* technique is a more general case of *sideband subtraction*: it provides a weight for every data point in a way the weighted distribution is re-sampling the background-subtracted distribution.
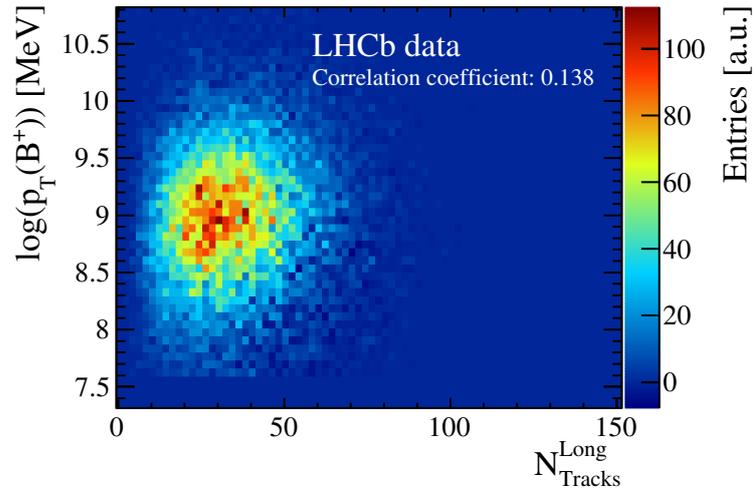
Fig. 5.7: Correlation between $p_T^{B^+}$ and the number of long tracks in the 2018 data sample. The correlation coeficcient is $\simeq 0$, proving the variables are not correlated.

Mathematically, it can be expressed using the number of signal N_s and background N_b events with probability density functions $s(d, c)$ and $b(d, c)$ respectively:

$$N_s s(d, c) + N_b b(d, c) = (N_s + N_b) f(d, c), \tag{5.3}$$

where d is the *discriminating* variable and c is the *control* variable. $f(d, c)$ is the Probability Density Function (PDF) of combined distribution of signal and background. As the *control* and *discriminating* variables are uncorrelated, one can rewrite their PDFs as

$$\begin{aligned} s(d, c) &= s(d)\, s(c)\,, \\ b(d, c) &= b(d)\, b(c)\,. \end{aligned} \tag{5.4}$$

The goal is to obtain an arbitrary weight function $w(d)$ fulfilling

$$N_s\, s(c) = (N_s + N_b) \int f(d, c) w(d)\, \mathrm{d}d = N_s s(c) \int s(d) w(d)\, \mathrm{d}d + N_n b(c) \int b(d) w(d)\, \mathrm{d}d\,. \tag{5.5}$$

Therefore, the $w(d)$ function is chosen in a way that:

$$\begin{aligned} \int s(d) w(d)\, \mathrm{d}d &= 1\,, \\ \int b(d) w(d)\, \mathrm{d}d &= 0 \end{aligned} \tag{5.6}$$

To have the smallest statistical uncertainty on the weights, the variation given by Eq. 5.7 of the weights have to be minimized

$$\int f(d, c) w(d)^2\, \mathrm{d}c\, \mathrm{d}d\,. \tag{5.7}$$

The three conditions assure an unique determination of the function $w(d)$. That allows for calculating the weights for any event with property d, resulting in signal-only distribution of the *control* variable. These weights are then called *sWeight*s.

*sWeight*ed *data* events in the resonant q^2 region are then used to obtain weights to correct the simulated sample. The data sample is dominated by the resonances (Sec. 5.3.1). Hence, the $B^+ \to K^{*+}J/\psi$ simulation sample is used to obtain the weights needed for correcting the simulation distributions. The agreement between the *sWeight*ed data and weighted simulation is crucial for the next step, the multi-variate analysis. The distributions used for the multi-variate analysis are carefully validated, see App. B.3, where the comparison of the *sWeight*ed *data* and weighted simulation for each data taking year are given. The distributions of the *sWeight*ed *data* and weighted simulation agree very well.

5.5   Multi-variate analysis selection

After the cut-based preselection, a rather large amount of combinatorial background is still present (see Fig. 5.2). To reduce the amount of background while maintaining high signal selection efficiency, a multi-variate analysis (MVA) is performed [159]. Generally, an MVA is a set of statistical methods that examine patterns in multidimensional data.
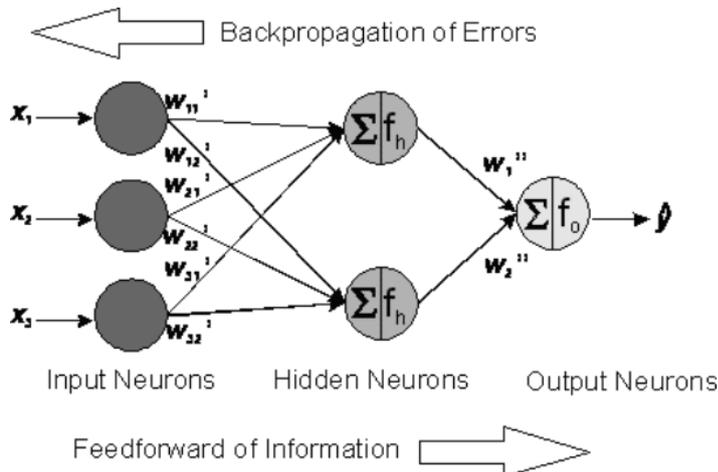
Fig. 5.8: Sketch of multilayer feedforward backpropagation neural network principles. The input layer distributes the input data by weighting them and sending them to the hidden neurons (nods). The hidden neurons sum the signal from the input neurons and project this sum on an *activation function* f_h. The activation function is typically a binary step (treshold) or rectified linear unit funcion $f(x) = max(0, x)$. The projected numbers are weighted and sent to the output layer, where they are summed again. There can be an arbitrary number of neurons and hidden layers.

There is a vast list of methods that can be considered a multi-variate analysis, the most commonly used ones are *decision trees* or *multiple regression* methods. In this analysis, the multilayer perceptron analysis is used.

A multilayer perceptron (MLP) is an artificial neural network. Neural networks were proposed as early as 1943 [160]. A simple sketch of its principle is presented in Fig. 5.8. It consists of three layers: input layer, hidden layer and output layer. Each layer consists of several (or many) nodes that are interconnected. A node receives a data item (a number) from each of its connections, multiplies it by an associated weight and returns the sum of these products.

This sum is then transformed by an *activation function*. During the training process, the associated weights are random: by examining examples with known input and/or output layer, the weights are associated in a way that the training data with same labels consistently yield similar output.

An MLP is a special kind of neural network: it is a supervised-learning network that uses backpropagation for training. It is used to distinguish data categories that are not linearly separable: in this case signal and background. Supervised learning means the neural network is trained with a set of input-output pairs (while unsupervised is trained only using the input data). Backpropagation means the gradient of the loss function with respect to the weights of the network is computed. The loss function represents the discrepancy between the desired output and the output calculated by the neural network. This *error* is then sent through the network backwards, updating the weights according to the *error*, leading to a quick reduction of the difference between the expected and calculated outputs.

The MLP tool provided by the Toolkit for Multivariate Data Analysis (TMVA) [161] is used. The samples used for training have to be clearly labeled as signal or background and be as close to the real signal and background as possible. Hence, the MLP is trained using $B^+ \to K^{*+}\mu^+\mu^-$ decay candidates in the simulation sample for signal with the requirement of the reconstructed B^+ meson mass to be close to the B^+ rest mass ($|m_{B^+}^{\mathrm{reco}} - m_{B^+}| <$ 100 MeV). The background training sample is taken from the recorded data: the B^+ meson upper-mass sideband, with the requirement of the reconstructed B^+ meson mass to be larger than 5700 MeV. The requirement of $m_{B^+}^{\mathrm{reco}} > 5700$ MeV enforces no (partially) reconstructed events in the background sample. The numbers of available signal and background events are listed in Tab. 5.4. The MLP is trained separately for Run I and Run II, as the Run conditions differed.

The list of variables that serve as an input to the MLP are presented in Tab. 5.5. These variables were identified as the variables with the largest discrimination power.

	Run I	Run II
Signal events	4531	19152
Background events	511	1748

Tab. 5.4: Number of events used for the MLP training.

The agreement between the simulated and recorded data in the listed variables becomes extremely important, as the MLP could pick up on differences between the data and simulation instead of separating background from the signal. As mentioned in Sec. 5.4, the *sWeight*ed data and weighted simulation distributions of variables listed in Tab. 5.5 are carefully checked to be in agreement. The distributions agree very well. Small discrepancies are acceptable as they are only minor and present in regions where the MLP does not differentiate between signal and background.

$$\ln p_{\mathrm{T}}^{B^+}$$
$$B^+ \ \mathrm{Cone}\text{-}p_{\mathrm{T}} \ \mathrm{asymmetry}$$
$$B^+ \ \chi^2_{\mathrm{IP}}$$
$$\ln(1 - B^+ \mathrm{DIRA})$$
$$\ln p_{\mathrm{T}}^{K^+}$$
$$|\eta(\pi^0) - \eta(K^+)|$$
$$CL_{\pi^0}$$
$$\max[\ln(p_{\mathrm{T}}^{\gamma_1}), \ln(p_{\mathrm{T}}^{\gamma_2})]$$
$$\min[\ln \mu^- \chi^2_{\mathrm{IP}}, \ln \mu^+ \chi^2_{\mathrm{IP}}]$$

Tab. 5.5: List of variables used for the MLP training. The confidence level of the neutral pion is a product of photon confidence levels, $CL_{\pi^0} = CL_{\gamma_1} CL_{\gamma_2}$. The list is identical in Run I and Run II.

In order for the MLP to select signal over background as efficiently as possible, the input variables should not be correlated among each other both in the signal and the background samples, as they lower the separation power of the MLP. The correlations between the input variables for the training signal and background samples are depicted in Fig. 5.9.

The TMVA toolkit returns MLP response value between 0 and 1, where the number represents the probability of an event being a signal event. The optimal cut value is discussed later in Sec. 5.7.

5.6 Treatment of multiple candidates

In most LHCb analyses, multiple candidates are not considered as a pollution due to the relatively precise charged track selection. In the case of π^0 reconstruction, especially
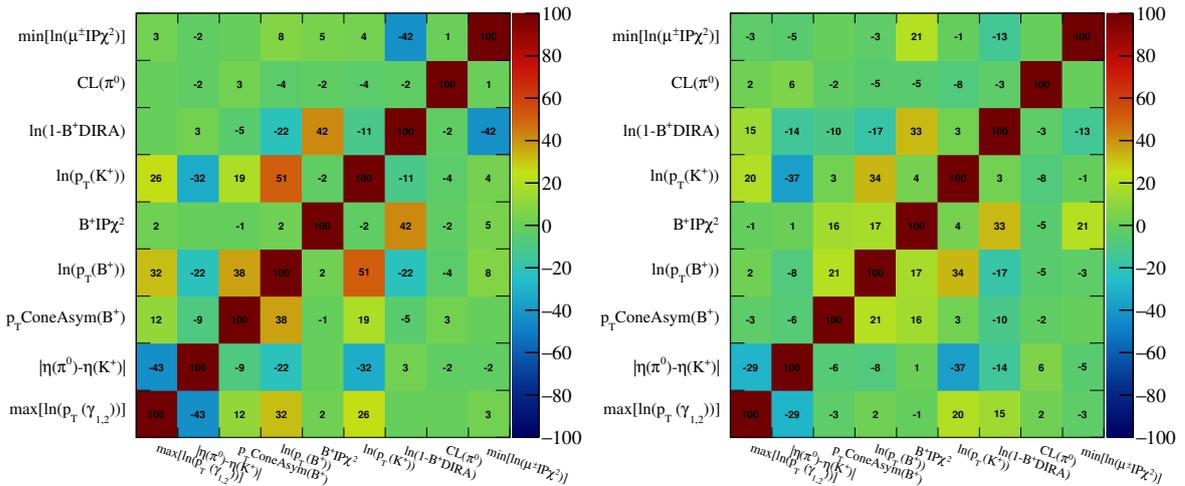
Fig. 5.9: The correlations between the input variables for the MVA training signal (left) and background samples (right). It is clear there is no significant correlation between the input variables both in the signal nad the background samples.

resolved π^0, multiple candidates are abundant. Multiple candidates refers to an event that is reconstructed with several signal candidates. This can happen *e.g.* when a π^0 meson is reconstructed using a random photon, especially in the case of very soft pions.

It is presented in Fig. 5.10 that the fraction of events containing multiple candidates decreases with increasing value of MLP response. This reflects the fact that MLP removes background events.
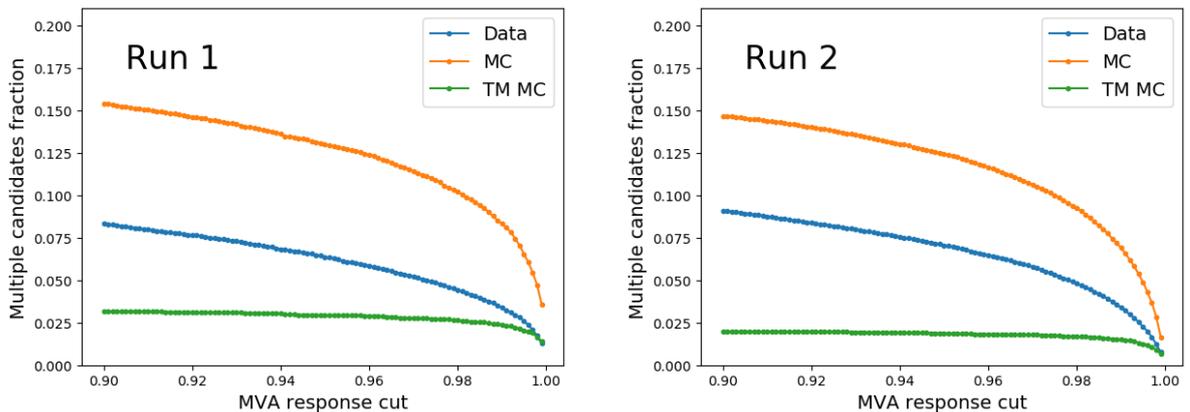
Fig. 5.10: Fraction of multiple candidates in data and simulation depending on the cut on MLP response. The fraction is defined as the number of all multiple candidates divided by the number of all events. This means that if *e.g.* there is one event with two multiple candidates in a sample of ten events, the ratio would be 0.2. This represents the number of events we actually exclude as the *fake* candidates are indistinguishable from *true* candidates. The blue points represent data, the orange points represent the simulation sample and the green points represent the truth-matched simulation sample.

Removing all multiple candidates no matter if they correspond to signal or not could negatively effect the significance $\mathcal{S}$, defined in Eq. 5.8, where S is the number of signal candidates and B is the amount of background candidates:

$$\mathcal{S} = \frac{S}{\sqrt{S + B}}.$$ (5.8)

However, as shown in Fig. 5.10, the final fraction of multiple candidates in the sample is about 10%. This means that in the worst possible case, 5% of true candidates have exactly one fake partner. In this case, the significance is worsened by a factor 0.97. The possible gain in significance if we would remove only the fake events is negligible. As a small fraction of candidates (about a 1%) have more than one fake partner, the removal of all events with *at least* one fake partner does not worsen the significance. As the disentanglement of the *true* candidate from the *fake* candidate is almost impossible and the possible loss of significance negligible, all multiple candidates are removed.

Moreover, the multiple candidates do not only affect the shape of the background. As shown in Fig. 5.11, multiple candidates typically worsen momentum resolution as they are background. As the π^0 momentum is tied to θ_K (see Fig. 3.1, θ_K is proportional to the asymmetry between K^+ and π^0 momenta), it is important to keep the π^0 resolution as good as possible. Removing multiple candidates is therefore a crucial step in this analysis, even though it is not possible to distinguish a true candidate from a fake candidate.
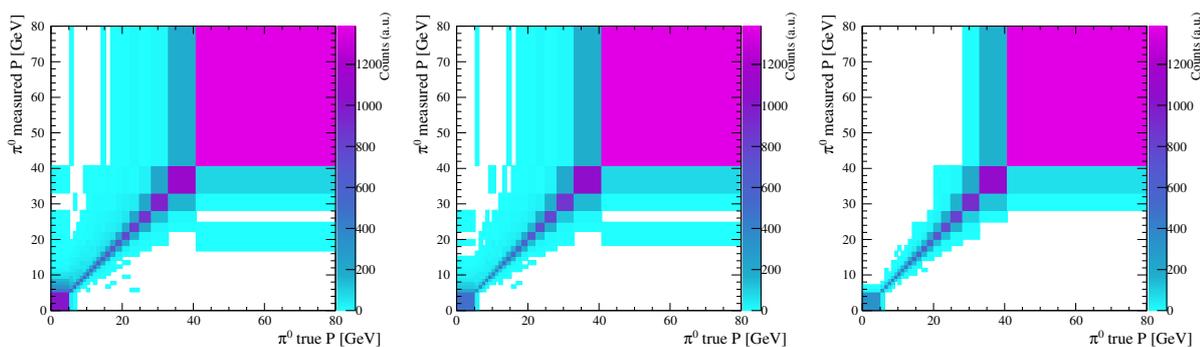
Fig. 5.11: π^0 momentum resolution in 2016 truth-matched simulation sample. The x-axis represents the *true* π^0 momentum, y-axis *measured* π^0 momentum. On the left, all events are shown. In the middle, events passing a cut on MLP response of 0.99 are shown[25]. On the right, on top of the cut on MLP response at 0.99, multiple candidates are removed. It is clear that removal of multiple candidates removes candidates with worse momentum resolution, especially for soft pions.

---

[25]This number is arbitrarily chosen, as it is clear from the MLP training the optimal MLP cut will be very close to one.

5.7   Signal estimation

In order to select the most events with the least background, the cut value on the MLP response is optimized. As the figure of merit, the significance $\mathcal{S}$ defined in Eq. 5.8 is chosen. Therefore, it is crucial to know the number of expected signal and background candidates in the data sample. Optimizing the MLP response cut using the $B^+ \to K^{*+}\mu^+\mu^-$ data sample could bias our result. Therefore, the reference decay $B^+ \to K^{*+}J/\psi$ is used to extrapolate the expected number of signal candidates.

Let the number of B^+ mesons decaying into $K^{*+}\mu^+\mu^-$ be $N_{\mu^+\mu^-}$. The number of all B^+ meson decays is denoted N_{all}. Then, branching ratio $BR_{\mu^+\mu^-}$ can be defined as

$$BR_{\mu^+\mu^-} = N_{\mu^+\mu^-}/N_{all}\,. \tag{5.9}$$

It is not possible to directly measure $N_{\mu^+\mu^-}$. Due to limited efficiency of the detector, $\varepsilon_{\mu^+\mu^-}$, the measured yield is $Y_{\mu^+\mu^-} = N_{\mu^+\mu^-} \times \varepsilon_{\mu^+\mu^-}$ events.

The branching ratio definition holds also for B^+ decaying into $K^{*+}J/\psi$. The branching ratio, $BR_{J/\psi}$, is the ratio of number of $B^+ \to K^{*+}J/\psi$ decay events, $N_{J/\psi}$, relative to all decays of the B^+ meson, N_{all}. Let the measured yield of $B^+ \to K^{*+}J/\psi$ events be $Y_{J/\psi}$ and the efficiency of detecting the decay events $\varepsilon_{J/\psi}$. Putting this together, the following formula gives the relation between the yield of $B^+ \to K^{*+}\mu^+\mu^-$ and $B^+ \to K^{*+}J/\psi$:

$$N_{all} = \frac{N_{\mu^+\mu^-}}{BR_{\mu^+\mu^-}} = \frac{Y_{\mu^+\mu^-}}{BR_{\mu^+\mu^-} \times \varepsilon_{\mu^+\mu^-}} = \frac{N_{J/\psi}}{BR_{J/\psi}} = \frac{Y_{J/\psi}}{BR_{J/\psi} \times \varepsilon_{J/\psi}}\,. \tag{5.10}$$

However, in this measurement only $B^+ \to K^{*+}J/\psi$ decays, where $J/\psi \to \mu^+\mu^-$, are considered. Therefore, the branching ratio of $J/\psi \to \mu^+\mu^-$, denoted $BR_{J/\psi\to\mu^+\mu^-}$, has to be added to Eq. 5.10. The efficiency of detecting $B^+ \to K^{*+}J/\psi$, $J/\psi \to \mu^+\mu^-$ is denoted $\varepsilon_{J/\psi,J/\psi\to\mu^+\mu^-}$. The formula used for estimation of the signal yield $Y_{\mu^+\mu^-}$ then becomes:

$$Y_{\mu^+\mu^-} = \frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi,J/\psi\to\mu^+\mu^-}} \times \frac{BR_{\mu^+\mu^-}}{BR_{J/\psi} \times BR_{J/\psi\to\mu^+\mu^-}} \times Y_{J/\psi,J/\psi\to\mu^+\mu^-}\,. \tag{5.11}$$

The used values of branching fractions, taken from Ref. [33], are presented in Tab. 5.6.

$BR_{B^+\to K^{*+}\mu^+\mu^-}$	8.668×10^{-7}
$BR_{B^+\to K^{*+}J/\psi}$	1.43×10^{-3}
$BR_{J/\psi\to\mu^+\mu^-}$	5.961×10^{-2}

Tab. 5.6: Branching ratios of $B^+ \to K^{*+}\mu^+\mu^-$, $B^+ \to K^{*+}J/\psi$ and $J/\psi \to \mu^+\mu^-$ decays used for estimating the signal yields. The values are taken from the FLAVIO package [139].

In order to estimate the yield of $B^+ \to K^{*+} \mu^+ \mu^-$, the total efficiency of both the $B^+ \to K^{*+} \mu^+ \mu^-$ and the $B^+ \to K^{*+} J/\psi$ decay selection is needed. The total efficiency depends on many factors: the detector acceptance (acc), reconstruction (rec), tracking (tr), selection (sel) and MLP efficiencies, as expressed in Eq. 5.12

$$\varepsilon_{tot} = \varepsilon_{acc} \times \varepsilon_{rec} \times \varepsilon_{tr} \times \varepsilon_{sel} \times \varepsilon_{MLP} \,. \tag{5.12}$$

However, for estimating the signal yield $Y_{\mu^+\mu^-}$, only the ratio of the signal and the reference efficiency $\varepsilon_{\mu^+\mu^-}/\varepsilon_{J/\psi}$ is needed. A lot of effects cancel out in this ratio. For practical reasons, the total efficiency is then evaluated in three steps: the acceptance efficiency ε_{acc}, the reconstruction, tracking, and selection efficiency $\varepsilon_{rec+tr+sel}$ and the efficiency of the multi-variate selection ε_{MLP}. The total efficiency then becomes

$$\varepsilon_{tot} = \varepsilon_{acc} \times \varepsilon_{rec+tr+sel} \times \varepsilon_{MLP} \,. \tag{5.13}$$

The evaluation of this efficiency is described in the following section.

5.7.1 Efficiency estimation

The efficiency for this analysis is estimated purely using the simulation sample. There are many limitations arising from this: availability of simulation samples or mismodeling of kinematic variables in the simulation. In order to obtain a signal yield estimation, the acceptance efficiency from simulation is obtained. At this level, the acceptance efficiency is approximated by the generator-level efficiency: the fraction of generated events being in the LHCb acceptance. As the resolution of the angles $\cos(\theta_K)$, $\cos(\theta_L)$ and ϕ is small, this is a good approximation.

Generator-level efficiency
Generator-level efficiencies are provided with the Monte Carlo simulation sample. Available values are summarized in Tab. 5.7. As the efficiency is studied per Run, final values are obtained by simply averaging over the magnet polarities and years. As the point of interest is the *ratio* of the efficiency of signal and reference channels, this approximation holds well enough.

Full selection efficiency
The next step is the full selection efficiency. This efficiency is the ratio of weighted truth-matched events passing the cut-based selection relative to the number of all generated events in the LHCb acceptance. The values of this efficiency for each year used to calculate the full efficiency according to Eq. 5.13 is shown in Fig. 5.12.

		2011	2012	2015	2016	2017	2018
$B^+ \to K^{*+}\mu^+\mu^-$	Down	—	—	16.15	16.10	16.09	16.05
	Up	—	—	16.08	16.11	15.95	16.09
$B^+ \to K^{*+}J/\psi$	Down	14.39	14.77	15.81	15.85	—	—
	Up	14.42	14.79	15.74	15.90	—	—

Tab. 5.7: Available generator-level efficiencies for signal and reference channels. The numbers represent the ratio of accepted signal events over generated signal events in [%]. Missing values for signal channel simulation (when these samples were produced, the generator-level efficiency was not automatically saved) samples are taken from reference channel simulation, missing values for reference simulation are taken from 2016 reference channel simulation.
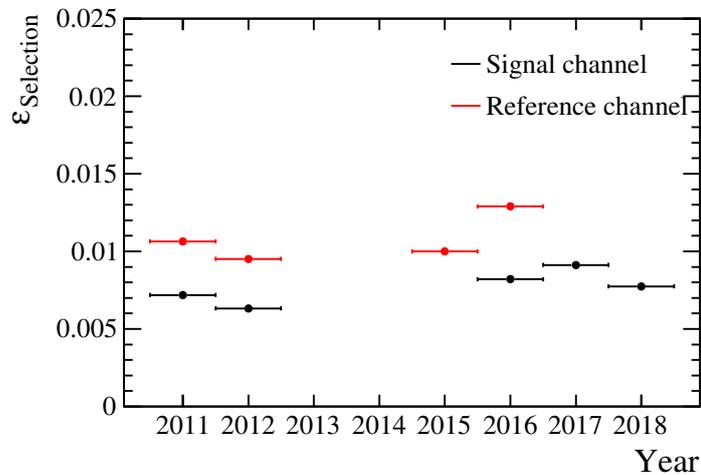
Fig. 5.12: Selection efficiency from available simulation samples per year. Red points show the efficiency for the reference $B^+ \to K^{*+}J/\psi$ channel, black points represent the signal $B^+ \to K^{*+}\mu^+\mu^-$ channel. Higher efficiency for the reference channel is caused by generally higher selection efficiency at $q^2 \sim 9\,\mathrm{GeV}^2$ (see Fig. 5.13).

A simple cross-validation of the selection process is done using a small sample of ten thousand signal events that passed only the generator-level requirements. It is shown in Fig. 5.13 that there is no significant kink in the efficiency in the resonance regions and hence no bias in the selection of $B^+ \to K^{*+}\mu^+\mu^-$ with respect to the $B^+ \to K^{*+}J/\psi$ decay is introduced. The Fig. 5.13 also explains why the reference channel $B^+ \to K^{*+}J/\psi$ efficiency is larger than the signal channel $B^+ \to K^{*+}\mu^+\mu^-$ efficiency: the q^2 selection efficiency is large around $q^2 \sim 9\,\mathrm{GeV}^2$ and therefore the selection efficiency is larger in the reference channel.

Multilayer perceptron efficiency

While the MLP is designed to separate between signal and background, it cannot be 100% *effective* and therefore a fraction of signal events is removed together with the background. The MLP efficiency is obtained from truth-matched simulation as the ratio
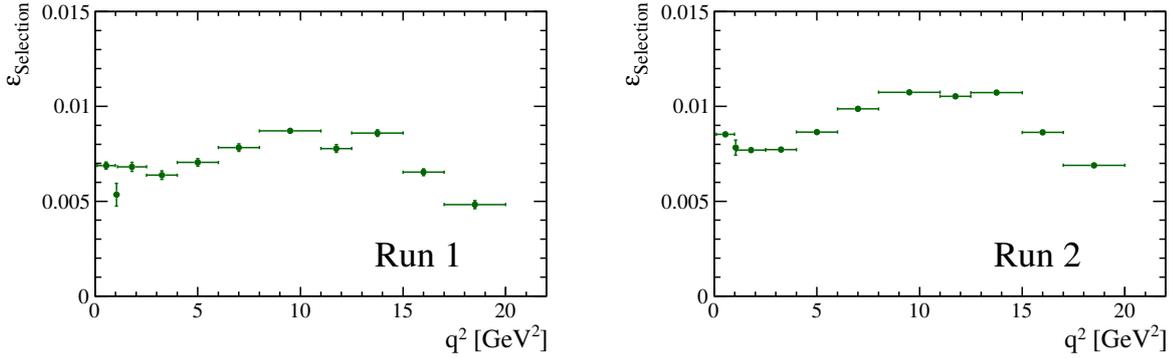
Fig. 5.13: Selection efficiency in q^2 dependence. The efficiency is estimated using ten thousand simulation events passing only generator-level requirements. The trend follows the q^2 acceptance of LHCb with no significant kink in the resonance regions ($[8.0\,\mathrm{GeV}^2, 11.0\,\mathrm{GeV}^2]$ and $[12.5\,\mathrm{GeV}^2, 15.0\,\mathrm{GeV}^2]$).

of events passing MLP response cut that are purged of multiple candidates (for details of this procedure see Sec. 5.6) over all truth-matched simulation candidates. The efficiency in dependence on MLP response is presented in Fig. 5.14.
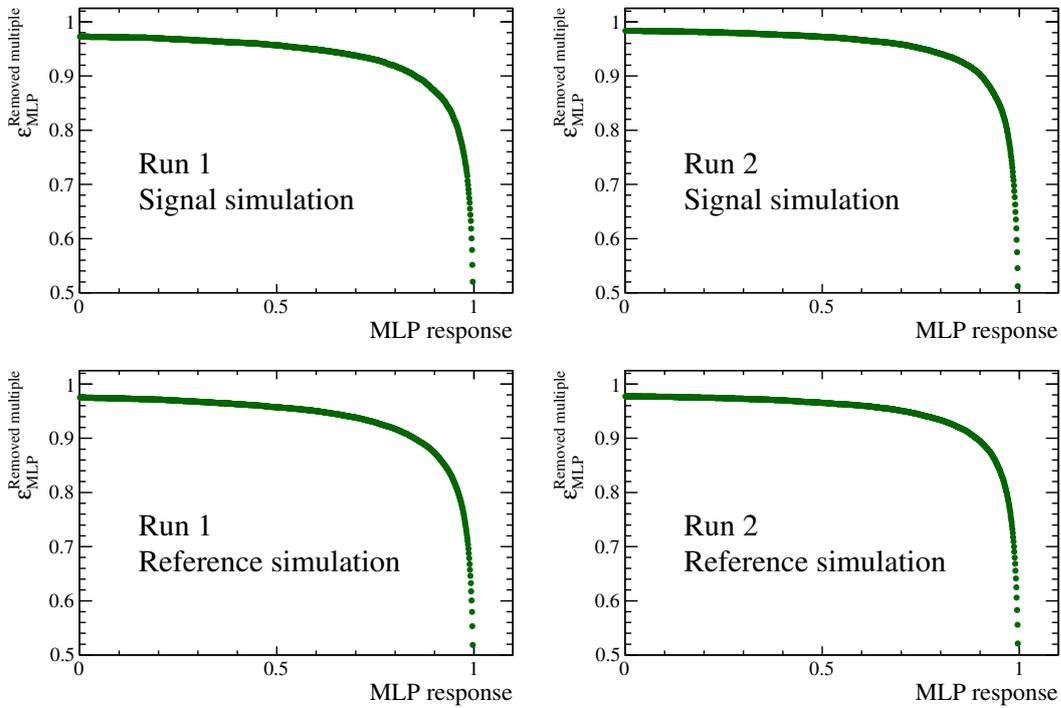
Fig. 5.14: MLP and multiple-candidates removal efficiency per Run obtained from signal and reference channels simulation sample. The offset from $\varepsilon_{MLP} = 1$ at MLP response equal to zero is caused by the removal of multiple candidates.

5.8 Checks for possible backgrounds

The main background source in this decay is combinatorial background. However, other background contributions could significantly shift the angular distributions and therefore need to be removed. It is shown already in Sec. 5.3.2 that the combinatorial background above the mass of the B^+ meson is dominated by $B^+ \to K^+\mu^+\mu^-$ decays, which are vetoed accordingly. Another check is performed to show that $B^0 \to \pi^0\mu\mu$ decays do not contribute to the signal. Contrary to the $B^+ \to K^+\mu^+\mu^-$ contribution, the reconstructed $\pi^0\mu^+\mu^-$ mass does not peak in the B^+ mass (see Fig. 5.15 and compare to Fig. 5.5).
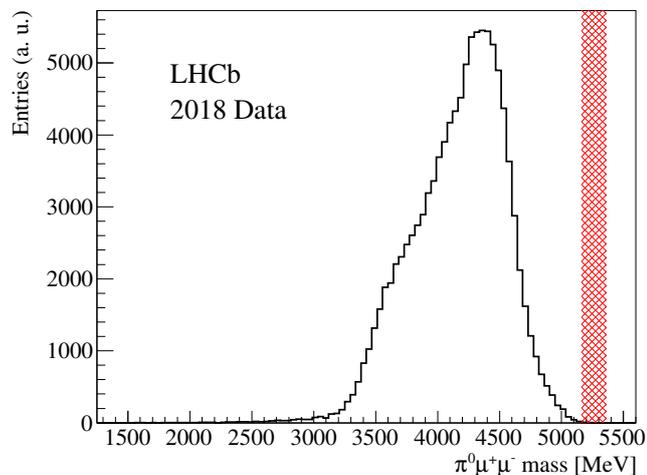
Fig. 5.15: Reconstructed $\pi^0\mu^+\mu^-$ mass in the 2018 data sample after cut-based selection. The red band represents region ± 100 MeV around the B^+ mass. These events are vetoed. There is no peak suggesting a significant contribution of the $B^0 \to \pi^0\mu\mu$ sample to the selected data. Compare with Fig. 5.5.

5.8.1 Study of the $\mathbf{B^+ \to XJ/\psi}$ sample

A validation of the selection is performed using a simulated sample of inclusive $B^+ \to XJ/\psi$ decays. This simulation is available for the years 2011, 2012 and 2016. The full selection identical to the signal selection is performed. After the selection, the true ID of the particles is checked in order to identify possible background contributions. A small contribution from B^{+*} is found. However, as this very quickly decays into B^+ meson and cannot be distinguished from the *true* signal, these events are considered as signal. In Fig. 5.16, the B^+ mass distribution from the sample of inclusive $B^+ \to XJ/\psi$ decay is shown, comparing the mass distribution from *true* $B^+ \to K^{*+}J/\psi$ events to background events. No peaking background is observed. A similar check is successfully performed in the K^* mass distribution (Fig. 5.17).

Fig. 5.16: Reconstructed B^+ meson mass in the simulation sample of inclusive $B^+ \to X J/\psi$ decay. The black line represents all candidates passing the whole selection, including the MLP response cut. The red line represents the events that are identified as background in the ID based truth-matching.
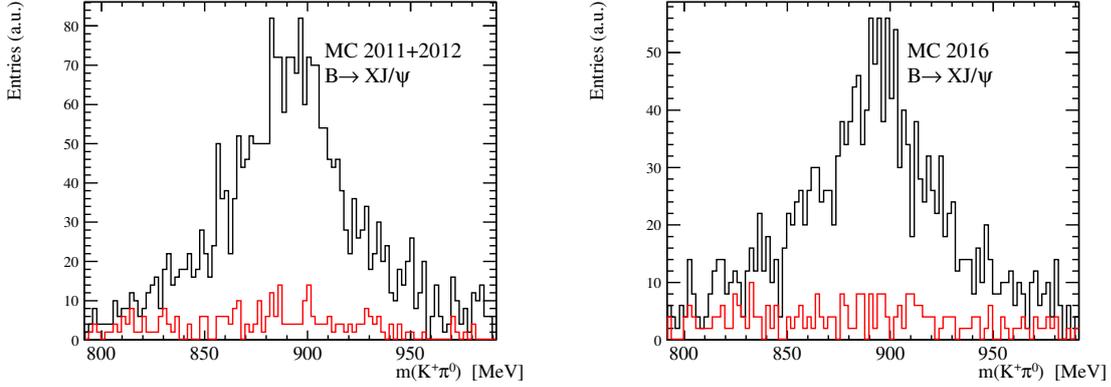
Fig. 5.17: Reconstructed K^* mass in the simulation sample of inclusive $B^+ \to X J/\psi$ decay. The black line represents all candidates passing the whole selection, including the MLP response cut. The red line represents the events that are identified as background in the ID based truth-matching.

Despite no clear contribution of peaking backgrounds, thorough checks revealed the presence of two background components: $B^+ \to K^{**} J/\psi$ and $B^+ \to K^+(Y_{c\bar{c}} \to J/\psi X)$. Both of these decays pass the selection if a π^0 meson or a photon from the decay of the excited K^{**} or charmonium $Y_{c\bar{c}}$ meson is missed. Note that both of these decays have J/ψ in their final state: they do not contribute to the signal channel, but to the reference channel only!

The shape of the $\pi^0 \mu^+ \mu^-$ mass distribution typical for these decays is shown in Fig. 5.18. While the $B^+ \to K^{**} J/\psi$ mass distribution shape is similar to the $B^+ \to K^{*+} J/\psi$ mass distribution shape, there is a clear peak at $\sim 3650\,\text{MeV}$ in the $B^+ \to K^+(Y_{c\bar{c}} \to J/\psi X)$ case. However, the overall contribution of this background is several percent, as can be seen in Fig. 5.19. In order to virtually eliminate this background and push it

to a 2% level only, a cut on the $\pi^0 \mu^+ \mu^-$ mass is proposed. This cut, corresponding to removing any event with $m_{\pi^0 \mu^+ \mu^-} < 3700\,\text{MeV}$ would also remove $\sim 12\%$ of the actual reference channel signal. Furthermore, this cut also removes virtually all events with soft π^0 mesons and therefore all events with $\cos(\theta_K) \sim 1$. This cut would then skew the angular distribution of the reference channel. This background accounts for $\sim 15\%$ of all candidates with $\cos(\theta_K) > 0.25$. Given the small number of such events and the already low selection efficiency in $\cos(\theta_K) \sim 1$, this peaking background is modeled as a part of the combinatorial background.
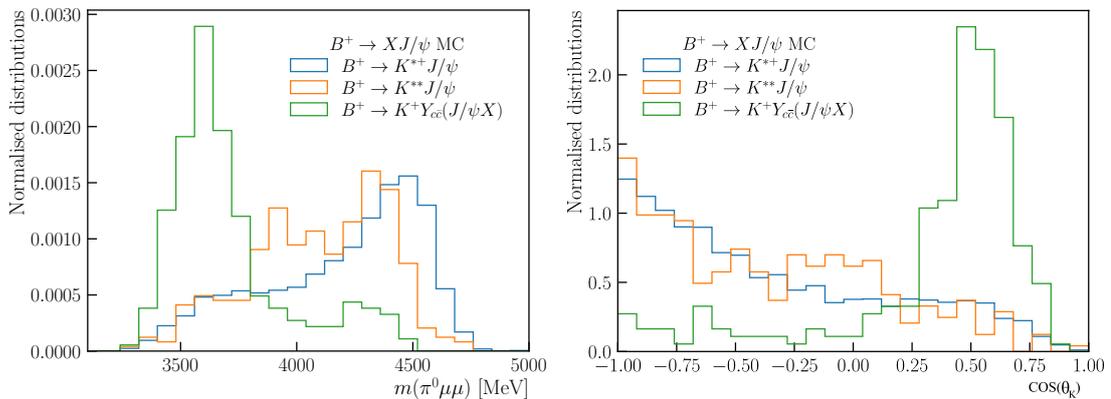
Fig. 5.18: Normalized reconstructed $\pi^0 \mu^+ \mu^-$ mass (left) and $\cos(\theta_K)$ (right) with different background contributions. The blue line represents the event candidates identified as signal, the orange line represents events coming from $B^+ \to K^{**} J/\psi$ and the green line represents the contribution from higher $c\bar{c}$ resonances $Y_{c\bar{c}}$ decaying further into J/ψ and a meson. The $B^+ \to K^{**} J/\psi$ contamination accounts for 15% of all signal candidates with $\cos(\theta_K) > 0.25$.
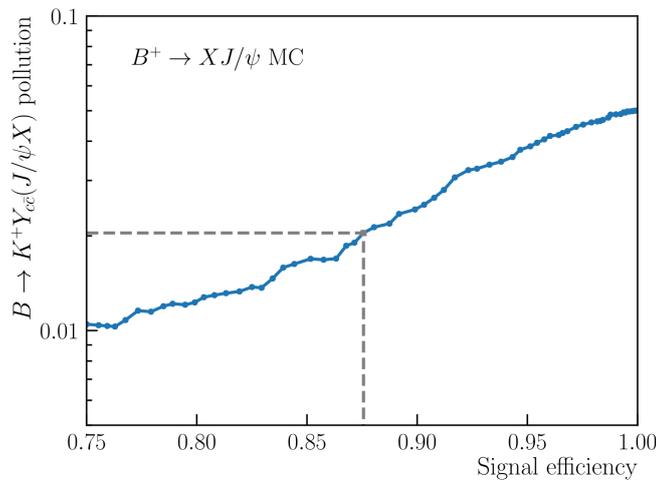
Fig. 5.19: The $B^+ \to K^+(Y_{c\bar{c}} \to J/\psi X)$ pollution against the $B^+ \to K^{*+} J/\psi$ efficiency. In order to reduce the background contribution to 2%, the signal efficiency would drop to $\sim 87\%$. This figure is produced using the inclusive $B^+ \to X J/\psi$ simulation sample.

5.9   Final selection

Using Eq. 5.8 and Eq. 5.11, the expected significance for each Run is estimated for many values of the MLP response. This is shown in Fig. 5.20. The maximum expected significance corresponds to cut on the MLP response at a value of 0.9985 for Run I and of 0.996 in Run II.
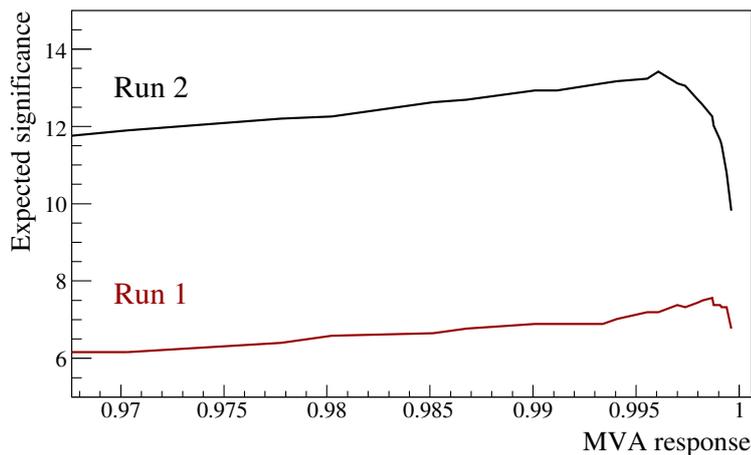
Fig. 5.20: Expected significance of the $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ decay yield. The significance is obtained using Eq. 5.8 and Eq. 5.11. The maximum significance value is 0.9985 for Run I and 0.996 for Run II.

.

The resulting mass distribution after applying the optimal MLP response cut on the full dataset obtained during Run I and Run II datat aking in the J/ψ resonance region is shown in Fig. 5.21. The expected signal and background yields obtained from the reference channel in the $B^+ \to K^{*+}\mu^+\mu^-$ channel are depicted in Tab. 5.8. The measured signal yield in the table is obtained from the fits to the signal channel presented in Fig. 5.22. It is worth noting here that the optimization was performed on the $B^+ \to K^{*+}J/\psi$ decay. As shown in Fig. 5.20, the MVA response value corresponding to the highest significance is not distinct and susceptible to fluctuations. Therefore, the expected $B^+ \to K^{*+}\mu^+\mu^-$ decay yields are higher than the actual fitted yields.

The selected $B^+ \to K^{*+}\mu^+\mu^-$ candidates are divided in *four* q^2 bins: $[0.1\text{-}4.0]\,\text{GeV}^2$ (excluding $0.98\text{-}1.1\,\text{GeV}^2$ in order to remove $\phi \to \mu^+\mu^-$ contribution), $[4.0\text{-}8.0]\,\text{GeV}^2$, $[11.0\text{-}12.5]\,\text{GeV}^2$ and $[15.0\text{-}18.0]\,\text{GeV}^2$. The measured mass distributions in these bins are presented in App. B.4. The measured signal and background yields together with their significance are shown in Fig. 5.23. The significance is also compared to a study by the CMS collaboration of Run I data exploiting $B^+ \to [K_S^0\pi^+]_{K^{*+}}\mu^+\mu^-$ [162][26] The

---

[26]This comparison is chosen as the significance in the analysis by the CMS collaboration is comparable
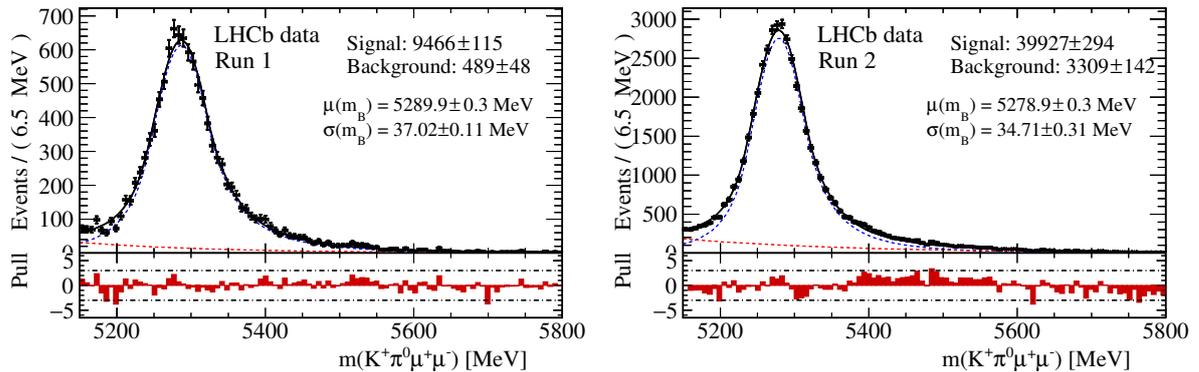
Fig. 5.21: Signal yield of the $B^+ \to K^{*+} J/\psi$ decay. The symbol $\mu(m_B)$ stands for the mean of the signal distribution, $\sigma(m_B)$ is the width of the peak. All parameters of the fit are left floating. The signal (blue) is fitted by two-sided Crystal Ball function (for the definition, see App. B.1), background (red) is described by an exponential function. The fitted signal and background yields are consired in ± 100 MeV around the B^+ meson mass.

	Expected	Fitted
Run I		
Signal	67	37±10
Background	14	49±7
S/sqrt(S+B)	7.47	4.03
Run II		
Signal	298	233±26
Background	203	262±17
S/sqrt(S+B)	13.30	10.49

Tab. 5.8: Expected and measured signal yields in the $B^+ \to K^{*+}\mu^+\mu^-$ decay. The expected $B^+ \to K^{*+}\mu^+\mu^-$ decay yields are higher than the actual fitted yields due to the optimization of the MVA response cut using only the reference $B^+ \to K^{*+} J/\psi$ decay.

$K^{*+} \to K^0_S \pi^+$, $K^0_S \to \pi^+\pi^-$ channel is detected more efficiently with better resolution, as the final state consists of charged particles only. The measured significance is higher than the one in the study done by the CMS collaboration, proving the potential of this analysis to measure all angular observables mentioned in Eq. 3.3. It is worth noting here that the CMS collaboration successfully measured only the F_L and A_{FB} angular parameters.

The numbers of signal candidates per data-taking year are given in Tab. 5.9 for the reference $B^+ \to K^{*+} J/\psi$ and in Tab. 5.10 for the signal $B^+ \to K^{*+}\mu^+\mu^-$ channel. To put the final number of candidates in perspective, the number of candidates after each selection step is included as well.

---

to the significance presented here. The other previous measurements discussed in Sec. 3.6 were performed either in experimentally cleaner environment or with only charged particles in the final state, reaching higher significance values.
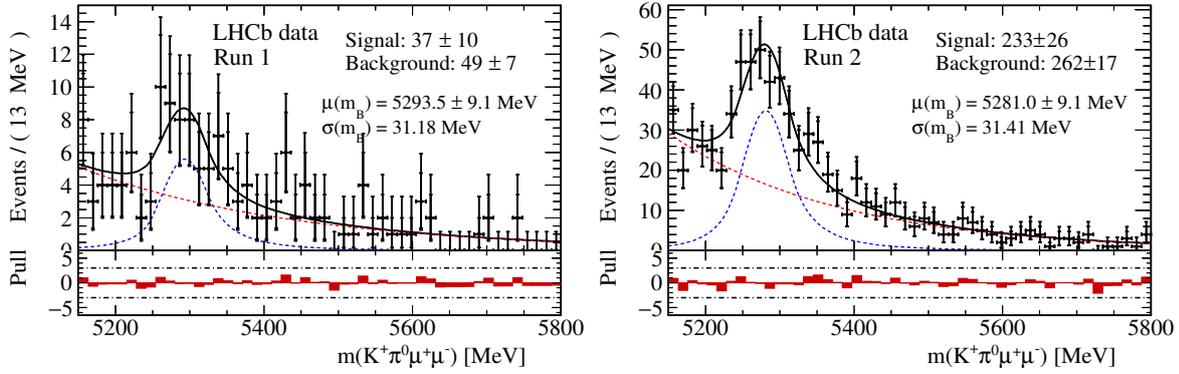
Fig. 5.22: Signal yield of the $B^+ \to [K^+\pi^0]_{K^{*+}}\mu^+\mu^-$ decay. $\mu(m_B)$ stands for the mean of the signal distribution, $\sigma(m_B)$ is the width of the peak. The signal shape is constrained to the shape of the signal yield in the resonance region shown in Fig. 5.21. The signal (blue) is fitted by two-sided Crystal Ball (for the definition, see App. B.1), background (red) is described by exponential function.
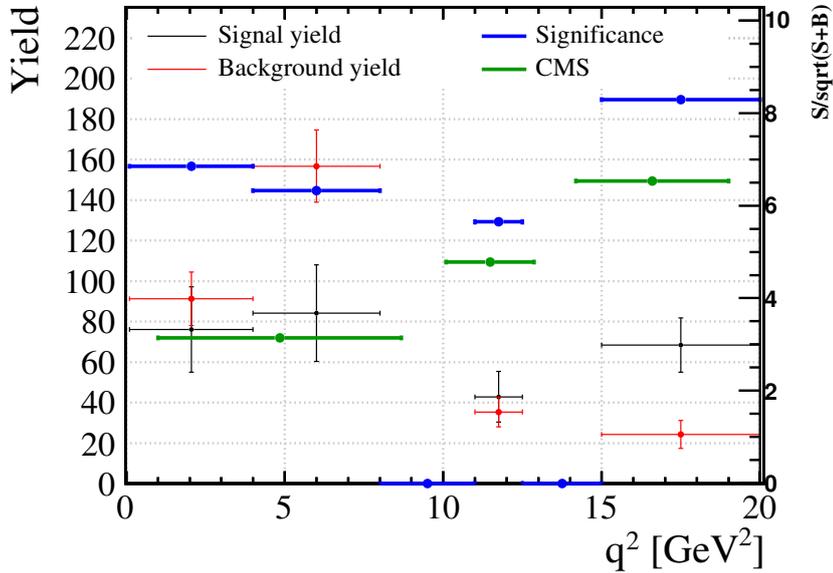
Fig. 5.23: Signal (red) and background (black) yields and measured **significance** (blue) in the combined Run I and Run II dataset. In the first bin $q^2 \in [0,4]$, the $\phi \to \mu\mu$ contribution ($q^2 \in [0.98, 1.1]$) is removed. The **green** data are taken from a study done by CMS. The study uses the $B^+ \to [K^0_S\pi^+]_{K^{*+}}\mu^+\mu^-$ decay data from Run I [162].

selection \ year	2011	2012	2015	2016	2017	2018
Trigger and online	23 718 772	58 047 021	9 822 137	57 955 614	32 702 706	54 868 587
Preselection	31197	67191	13769	89310	90460	90660
MLP selection	4409	8637	2483	16475	17885	17659

Tab. 5.9: Number of the **reference channel** event candidates after the trigger and the stripping selection, preselection and MLP selection. Event candidate is any event passing the selection step, therefore this includes background candidates. In the last step, the background is mostly combinatorial. Note the discrepancy between the events passing the trigger and online selection in 2017 data-taking year compared to 2016 and 2018: this is a result of stricter stripping selection. For details see Tab. 5.2.

selection \ year	2011	2012	2015	2016	2017	2018
Trigger and online	10 972 833	28 455 565	5 322 454	31 999 312	16 969 963	30 897 345
Preselection	3134	6881	1288	10017	10016	10090
MLP selection	42	101	39	216	242	241

Tab. 5.10: Number of the **signal channel** event candidates after the trigger and the stripping selection, preselection and MLP selection. Event candidate is any event passing the selection step, therefore this includes background candidates. In the last step, the background is mostly combinatorial. Note the discrepancy between the events passing the trigger and online selection in 2017 data-taking year compared to 2016 and 2018: this is a result of stricter stripping selection. For details see Tab. 5.2.

6 Acceptance effects

An accurate measurement of the angular distribution of the $B^+ \to K^{*+}\mu^+\mu^-$ decay in different q^2 regions requires a correction of the acceptance effects. The non-flat angular acceptance is given by the geometry of the LHCb detector and by the event selection. A dedicated simulation sample is used to ascertain this effect. In this simulation sample, the events are generated with a pure phase-space (PHSP) coupling in the decay chain. Therefore, this set of simulated events is per construction generated with flat distributions in the 4D space of θ_K, θ_L, ϕ and q^2.

The PHSP simulation sample is used to validate that the event selection does not strongly bias the θ_K, θ_L, ϕ distributions. Then, the PHSP simulation is corrected to match the data in the same way as described in Sec. 5.4. Last step is the 4D-parametrization of θ_K, θ_L, ϕ and q^2 distributions, resulting in weights applied to the data.

6.1 Simulation with homogeneous phase-space distribution

In order to precisely describe the acceptance, a PHSP simulation sample is generated for each data-taking year. The number of events passing the event selection are summarized in Tab. 6.1.

The trigger, central and cut-based selections are heavily influenced by the detector geometry. A dedicated cross-check is done to verify the multi-variate selection does not depend on the decay angles. This is validated by establishing the efficiency of the MLP ε_{MLP} in a same way as described in Sec. 5.7.1.

The MLP efficiency ε_{MLP} in dependence on q^2, $\cos(\theta_K)$, $\cos(\theta_L)$ and ϕ is shown in Fig. 6.1. In the top row, $\varepsilon_{MLP} = \varepsilon_{MLP}(q^2)$ is shown. Careful reader will notice a small dip at $\sim 3\,\mathrm{GeV}^2$. This is purely caused by the detectors acceptance, similarly to the *roof*-like trend in $\varepsilon_{MLP} = \varepsilon_{MLP}(\cos(\theta_L))$. A large effect is visible in the MLP efficiency in the very-high $\cos(\theta_K)$ region. As $\cos(\theta_K)$ is proportional to the asymmetry between the momentum of K^{*+} meson decay products $\frac{p_{K^+}-p_{\pi^0}}{p_{K^+}+p_{\pi^0}} \simeq \cos(\theta_K)$, events with very low π^0 momentum are more affected by background contributions and hence the efficiency drops at $\cos(\theta_K) \sim 1$.

year	2011	2012	2015	2016	2017	2018
events	6965	15836	8431	28631	31589	36307

Tab. 6.1: Number of PHSP signal candidates passing the full selection. For the number of generated PHSP events, see Tab. 4.3.

Fig. 6.1: MLP efficiency per Run obtained from PHSP simulation sample as a function of $\cos(\theta_K)$, $\cos(\theta_L)$, ϕ and q^2. The dip of ε_{MLP} at $\cos(\theta_K) \sim 1$ is caused by events with very low neutral pion momentum p_{π^0}.

6.2 Parametrization of the angular acceptance

In order to obtain the correction weights, the PHSP sample, flat in θ_K, θ_L, ϕ and q^2, are parametrized using a Legendre polynomial [163]. As the four observables do not factorize, the polynomial takes the form of

$$\epsilon(\cos(\theta_L), \cos(\theta_K), \phi, q^2) = \sum_{l,m,n,o} c_{lmno} \times P_l(q^2) \times P_m(\cos(\theta_L)) \times P_n(\cos(\theta_K)) \times P_o(\phi) , \quad (6.1)$$

where $P_{l,m,n,o}$ are Legendre polynomials of orders l, m, n and o. The maximal order of the polynomial is chosen in a way that the polynomial describes the acceptance well while preventing picking-up statistical fluctuations in the PHSP simulation sample. Moreover, the parametrization in ϕ is forced to be symmetric. The possible asymmetry in the ϕ distribution is smeared out by integrating over B^+ and B^- meson decays as well as the reversal of polarity of the bending magnet.

The parametrization is obtained individually for each Run. The maximal order of the polynomial is optimized using a χ^2-goodness of the parametrization (see Fig. 6.2) and visual inspection of the projections (see Fig. 6.3, Fig. 6.4 and App. C). It is clear from Fig. 6.2 that there is no clear best maximal order of the polynomial. This should be taken into account as a systematic uncertainty and it is discussed later in Sec. 9.4. The order of the Legendre polynomial describing the PHSP simulation sample well is found to be six in $\cos(\theta_K)$, three in $\cos(\theta_L)$, flat in ϕ and seven in q^2. The higher order of the $\cos(\theta_K)$ polynomial is caused by the very low acceptance in the high $\cos(\theta_K)$ region arising from the high background contribution in the low π^0 momentum region. The acceptance at very high $\cos(\theta_K)$ is essentially zero. This leads to huge weights destabilizing the angular fit later on. For this reason, the $\cos(\theta_K)$ range is limited to $[-1.0, 0.8]$. The $\cos(\theta_K)$ range is further limited in the case of applying folding 4 defined in Sec. 3.5: in order to be able to fold in the $\cos(\theta_K)$ dimension, only candidates with $\cos(\theta_K) \in [-0.8, 0.8]$ are considered.

The final form of the Legendre polynomial takes the form of

$$\epsilon(\cos(\theta_K), \cos(\theta_L), \phi, q^2) = \sum_{l=1}^{6} \sum_{m=1}^{3} \sum_{o=1}^{7} c_{lmno} \times P_l(\cos(\theta_K)) \times P_m(\cos(\theta_L)) \times P_1(\phi) \times P_o(q^2) .$$
$$(6.2)$$

Finally, in order to correct the data for the angular acceptance, each event is weighted with the weight w

$$w(\cos(\theta_L), \cos(\theta_K), \phi, q^2) = \frac{1}{\epsilon(\cos(\theta_L), \cos(\theta_K), \phi, q^2)} . \quad (6.3)$$

6.2 Parametrization of the angular acceptance

χ^2 scan of max. order of polynoms (Run1)

$q^2-\phi$	5-2	5-3	5-4	6-2	6-3	6-4	7-2	7-3	7-4
8-3	1.075	1.030	1.042	1.076	1.036	1.047	1.077	1.038	1.047
8-2	1.069	1.025	1.029	1.072	1.026	1.038	1.078	1.027	1.042
8-1	1.069	1.025	1.029	1.072	1.026	1.038	1.078	1.027	1.042
7-3	1.090	1.029	1.041	1.083	1.034	1.044	1.082	1.036	1.044
7-2	1.069	1.025	1.032	1.069	1.026	1.034	1.070	1.028	1.032
7-1	1.069	1.025	1.032	1.069	1.026	1.034	1.070	1.028	1.032
6-3	1.076	1.030	1.173	1.075	1.033	1.052	1.074	1.035	1.045
6-2	1.068	1.026	1.036	1.068	1.027	1.035	1.068	1.028	1.032
6-1	1.068	1.026	1.036	1.068	1.027	1.035	1.068	1.028	1.032

($\theta_K^{-}\theta_L$)

χ^2 scan of max. order of polynoms (Run2)

$q^2-\phi$	5-2	5-3	5-4	6-2	6-3	6-4	7-2	7-3	7-4
8-3	1.159	1.024	1.022	1.160	1.026	1.026	1.162	1.030	1.031
8-2	1.161	1.027	1.027	1.161	1.027	1.028	1.162	1.029	1.030
8-1	1.161	1.027	1.027	1.161	1.027	1.028	1.162	1.029	1.030
7-3	1.158	1.023	1.021	1.158	1.025	1.024	1.160	1.028	1.028
7-2	1.160	1.027	1.027	1.160	1.027	1.027	1.161	1.029	1.029
7-1	1.160	1.027	1.027	1.160	1.027	1.027	1.161	1.029	1.029
6-3	1.156	1.022	1.021	1.157	1.023	1.024	1.159	1.027	1.028
6-2	1.159	1.026	1.026	1.159	1.027	1.027	1.160	1.028	1.029
6-1	1.159	1.026	1.026	1.159	1.027	1.027	1.160	1.028	1.029

($\theta_K^{-}\theta_L$)

Fig. 6.2: Angular acceptance parametrization χ^2-goodness scan for Run I (left) and Run II (right). The numbers on the axis correspond to the applied order of the Legendre polynomial for the given variables. Note that the χ^2 values for each of the parametrization are very close to each other: there is no preference of the order of the polynomial in the orders considered here.

LHCb simulation
this thesis

LHCb simulation
this thesis

LHCb simulation
this thesis

LHCb simulation
this thesis
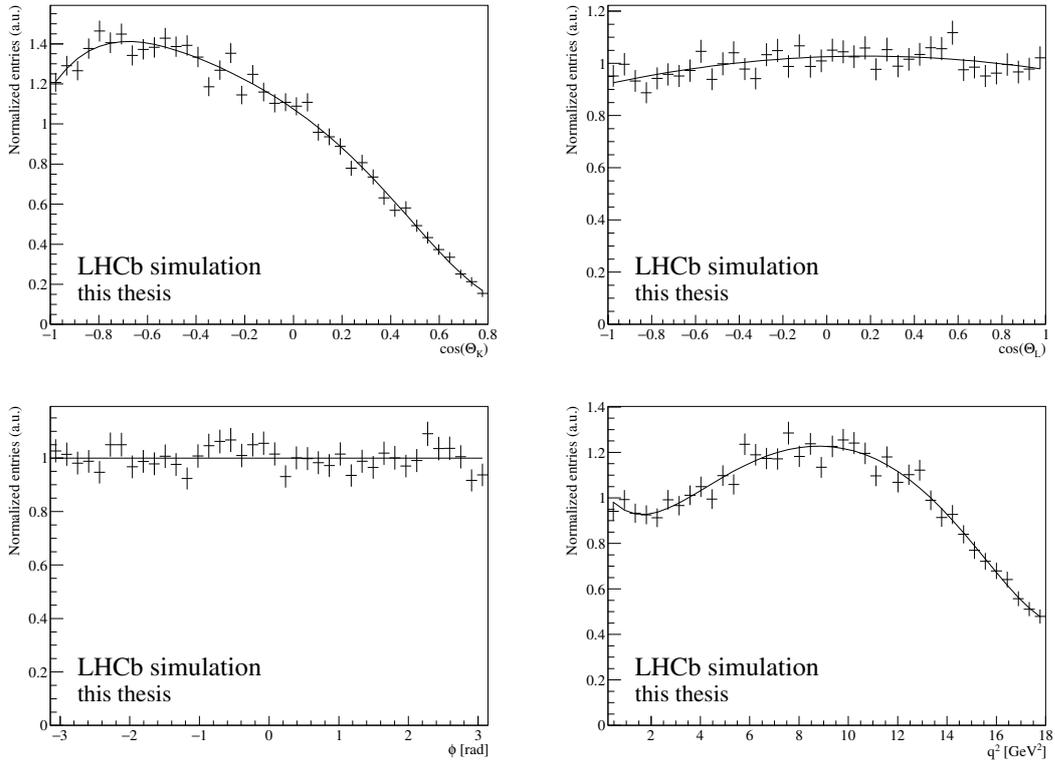
Fig. 6.3: One-dimensional projections of the angular acceptance. The data points are Run I PHSP simulation, the solid curve is the four dimensional Legendre-polynomial parametrization described by Eq. 6.2.

74
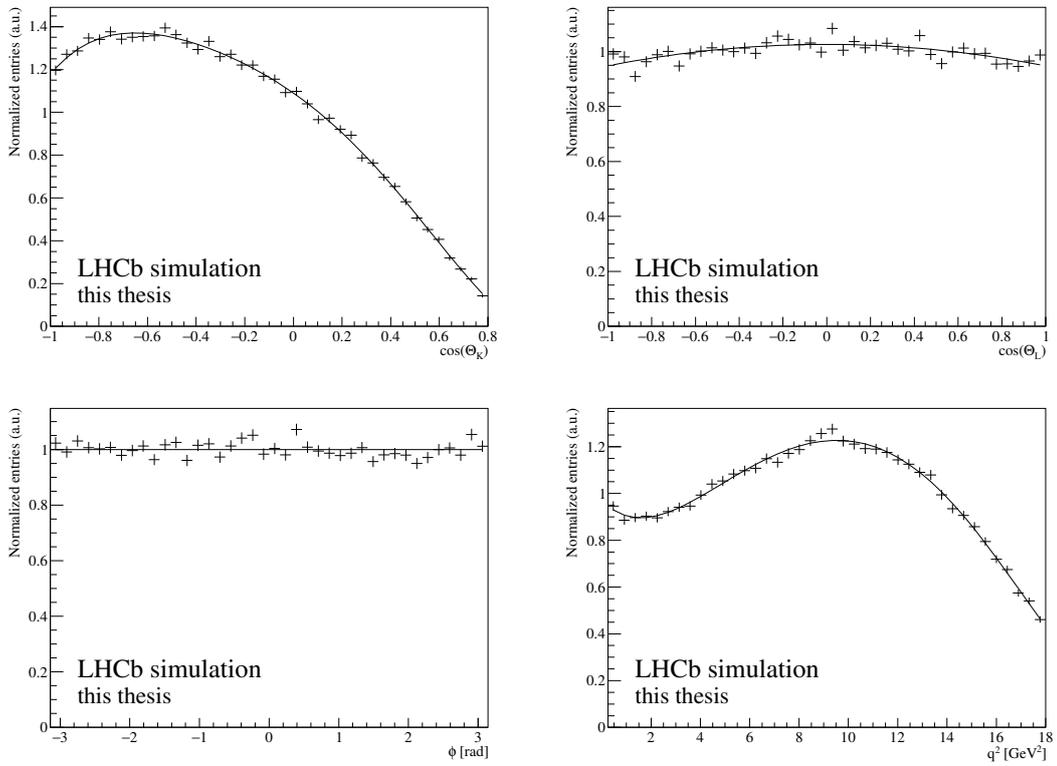
Fig. 6.4: One-dimensional projections of the angular acceptance. The data points are Run II PHSP simulation, the solid curve is the four dimensional Legendre-polynomial parametrization described by Eq. 6.2.

7 Parameter measurement

The angular moments F_L, S_3, S_4, S_5, A_{FB}, S_7, S_8 and S_9 described in Eq. 3.9 are extracted from the selected events using a C++ based fitter framework FCNCFITTER. It performs maximum-likelihood fits using the TMINUIT [164] minimization class of the ROOT framework [165][27]. This framework has been developed in the LHCb collaboration and successfully used in previous analyses described in Ref. [3, 4, 6]. The framework is further developed to accommodate for the needs of this particular analysis and to improve the user experience.

7.1 Maximum likelihood

Maximum-likelihood is a method of estimating the parameters of a probability distribution. As the name suggests, the method maximizes the likelihood function in a way that the assumed statistical model of the observed data is most probable.

In order to fully understand the concept of maximum-likelihood in the multi-dimensional fit, it is useful to start with a simple example of a one parameter fit and expand it to the multi-dimensional space.

Let h represent a hypothesis, $\{D\} = x_1, x_2, ..., x_n$ measured data. Then, using bayes theroem [166], the probability density function PDF representing that the hypothesis is valid given the data $\{D\}$ can be rewritten as

$$\mathrm{PDF}(h \,|\, \{D\}) = \frac{\mathrm{PDF}(\{D\} \,|\, h)\,\mathrm{PDF}(h)}{\mathrm{PDF}(\{D\})}\,. \tag{7.1}$$

The first part of the numerator, $\mathrm{PDF}(\{D\} \,|\, h)$, is refereed to as *likelihood*, the second part, $\mathrm{PDF}(h)$, as *prior*. The prior is assumed to be a uniform distribution. The numerator represents the probability of the data averaged over all parameters. As $\mathrm{PDF}(\{D\})$ is a constant that is not relevant for maximizing the probability that the hypothesis is true.

As the goal is to maximize the probability $\mathrm{PDF}(h \,|\, \{D\})$, the first order derivative at observable μ_0 has to be zero and the second order derivative positive:

$$\left.\frac{\partial \mathrm{PDF}}{\partial h}\right|_{\mu_0} = 0\,, \qquad \left.\frac{\partial^2 \mathrm{PDF}}{\partial h^2}\right|_{\mu_0} > 0\,. \tag{7.2}$$

In this work, an unbinned maximum likelihood fit is performed. Extending the previous to a multidimensional space, let N be the number of events, each assigned a weight

---

[27]The discrepancy between maximum likelihood and its minization will be explained in the following section.

w_n. Let $\vec{\mu}$ be the parameter vector and $\{\vec{D}\} = \vec{x_1}, \vec{x_2}, ..., \vec{x_n}$ measured data. The likelihood takes the form of

$$\mathcal{L}(\{\vec{D}\}|\vec{\mu}) = \text{PDF}(\{\vec{D}\}|\vec{\mu}) = \prod_{n=1}^{N} \text{PDF}(\vec{D_n}|\vec{\mu})^{w_n} , \qquad (7.3)$$

where $\text{PDF}(\vec{D_n}|\vec{\mu})$ is the normalized probability density function according to which the data is distributed.

The maximization problem is often reduced to a much simpler problem. Instead of maximizing the likelihood (for simplicity denoted $\mathcal{L}$) itself, it is possible to *minimize* a negative logarithm of the likelihood $-\ln\mathcal{L}$. Looking at Eq. 7.3, the minimization problem becomes:

$$-\ln(\mathcal{L}(\{D\}|\vec{\mu})) \propto - \sum_{n=1}^{N} w_n \ln\left(\text{PDF}(\vec{D_n}|\vec{\mu})\right) . \qquad (7.4)$$

In the unbinned fit, the weights for each event w_n are often not normalized. A factor of $N/\sum_n w_n$ then has to be added to the logarithm of the likelihood in order to extract an unbiased uncertainty.

Some of the parameters can be constrained to previously known value v with some uncertainty u. For every such parameter p_i, an additional term is added to the likelihood

$$\ln\mathcal{L}_{\text{constr}} = \sum_{j} \left(\frac{p_i - v_i}{u_i}\right) . \qquad (7.5)$$

The biggest advantage of minimizing the negative logarithm of likelihood instead of maximizing the likelihood directly is that the logarithm can be expanded using the Taylor expansion [167] in the maximum likelihood estimator $\vec{\mu_0}$:

$$\ln(\mathcal{L}(\vec{\mu})) = \mathcal{L}(\vec{\mu_0}) + \left.\frac{\partial \mathcal{L}}{\partial \vec{\mu}}\right|_{\vec{\mu_0}} (\vec{\mu} - \vec{\mu_0}) + \left.\frac{\partial^2 \mathcal{L}}{\partial \vec{\mu}^2}\right|_{\vec{\mu_0}} \frac{(\vec{\mu} - \vec{\mu_0})^2}{2} + \omega_3 ,$$

where the ω_3 denotes the higher order contributions, which are *typically* negligible.

The first element in the expansion is a constant, therefore not interesting for the minimization. The second element is equal to zero from Eq. 7.2. Therefore, the PDF that has to be minimized, denoted for simplicity $\mathcal{P}$, becomes

$$\mathcal{P} = - \left.\frac{\partial^2 \mathcal{L}}{\partial \vec{\mu}^2}\right|_{\vec{\mu_0}} \frac{(\vec{\mu} - \vec{\mu_0})^2}{2} . \qquad (7.6)$$

7.2 Fit model

The FCNCFITTER framework offers a wide variety of fit models. Despite its versatility, further improvements are made in order to adapt for this analysis, especially due to the limited $\cos(\theta_K)$ availability and the presence of a complicated background component. The fit model used to extract the parameters F_L, S_3, S_4, S_5, A_{FB}, S_7, S_8 and S_9 consists of two main components: signal and background probability density functions P_{sig} and P_{bkg}. The PDF then can be generally described using signal and background probability density functions:

$$\text{PDF} = f_{sig} \prod_{i=1}^{D} \text{P}_{sig}^i + (1 - f_{sig}) \prod_{i=1}^{D} \text{P}_{bkg}^i \,, \tag{7.7}$$

where the D represents the dimension of the fit and f_{sig} is the fraction of signal candidates in the dataset to all candidates.

The fit is performed in four dimensions of B^+ meson mass, $\cos(\theta_L)$, $\cos(\theta_K)$ and ϕ. In addition in the case of the $B^+ \rightarrow K^{*+} J/\psi$ decay, in order to extract the F_S parameter (see Eq. 3.9) another fit is performed in two dimension of B^+ meson mass and K^* mass.

As both the collision and the detector conditions differ between Run I and Run II, the datasets are treated separately. However, all angular observables noted in Eq. 3.9 are independent of those conditions. Hence these parameters are shared between the two datasets in the fit. Moreover, to further stabilize the fit, the angular background parameters are also shared between the two datasets.

7.2.1 Signal component

The reconstructed B^+ meson mass is described by a double-sided Crystal Ball function defined in App. B.1. The parameters of the Crystal Ball function $\alpha_{1,2}$ and $n_{1,2}$ are fixed to the B^+ meson mass shape in the simulation of the reference channel decay $B^+ \rightarrow K^{*+} J/\psi$. This is due to the shape of the tails of the Crystal Ball function: even when fitting the simulated B^+ meson mass sample of the rare channel decay $B^+ \rightarrow K^{*+} \mu^+ \mu^-$, the parameters $\alpha_{1,2}$ and $n_{1,2}$ show large uncertainties and the fit becomes unstable. On top of this, the mean of the B^+ meson mass peak in the rare $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ decay is fixed to the one obtained by fitting the reference decay channel data.

Due to low statistics of the signal sample, the width of the Crystal Ball function $\sigma_{\text{rare}}^{Data}$ is fixed to the width in the reference decay channel data fit $\sigma_{\text{ref}}^{Data}$ multiplied by a scaling factor obtained by fitting the simulated B^+ meson mass in the rare and reference decay channels, $\sigma_{\text{rare}}^{MC}$ and σ_{ref}^{MC}

$$\sigma_{\text{rare}}^{Data} = \sigma_{\text{ref}}^{Data} \frac{\sigma_{\text{rare}}^{MC}}{\sigma_{\text{ref}}^{MC}} \,. \tag{7.8}$$

7.2 Fit model

For the fit of the K^* mass, the P-wave amplitude is described by Breit-Wigner model [168]:

$$\mathcal{A}_{\mathrm{P}}\left(m_{K^+\pi^0}\right) = \sqrt{kp} \times B'_{L_B}(k, k_0, d) \left(\frac{k}{m_{B^+}}\right)^{L_B} \times B'_{L_{K^{*+}}}(p, p_0, d) \left(\frac{p}{m_{K^{*+}}}\right)^{L_{K^{*+}}}$$
$$\times \frac{1}{m_{K^{*+}}{}^2 - \left(m_{K^+\pi^0}\right)^2 - im_{K^{*+}}\Gamma\left(m_{K^+\pi^0}\right)} . \tag{7.9}$$

The momentum of K^{*+} in the rest-frame of B^+ meson meson is denoted k with a mean peak value of k_0, the momentum of K^+ in the rest frame of K^{*+} is denoted p with a mean peak value of p_0. L denotes the angular momentum of the corresponding meson. The factors B_L' are so-called Blatt-Weisskopf form-factors [169]

$$B'_0\left(p, p_0, d\right) = 1 \, ,$$
$$B'_1\left(p, p_0, d\right) = \sqrt{\frac{1 + (p_0 d)^2}{1 + (p\,d)^2}} \, , \tag{7.10}$$

where d represents the size of the decaying particle. This parameter is reported in Ref. [170] to be $d = 1.6\,\mathrm{GeV}^{-1}$ (or $0.3\,\mathrm{fm}$). This is also consistent with a $B^+ \to J/\psi\rho^+$ branching fraction measurement, where the fit favored $d = 1.64\,\mathrm{GeV}^{-1}$ [171]. However, recent LHCb study of $Z(4430)$ favored $d \sim 0$ [172]. As the determination of the d parameter is not possible in this analysis, the value is fixed to $d = 1.6\,\mathrm{GeV}^{-1}$. This is also consistent with previous B^0 decay analyses [3, 6] and the $B^+ \to [K_S^0\pi^+]_{K^{*+}}\mu^+\mu^-$ decay analysis [4].

As the angular momentum of the P-wave is $L_B = 0$ and $L_{K^*} = 1$, the Eq. 7.9 becomes

$$\mathcal{A}_{\mathrm{P}}(m_{K^+\pi^0}) = \sqrt{kp} \times \sqrt{\frac{1 + (p_0 d)^2}{1 + (pd)^2}} \times \frac{p}{m_{K^{*+}}} \times \frac{1}{m_{K^{*+}}{}^2 - m_{K^+\pi^0}{}^2 - im_{K^{*+}}\Gamma(m_{K^+\pi^0})} . \tag{7.11}$$

For the description of the S-wave in the $m_{K^+\pi^0}$, the LASS parametrization [173] is used

$$\mathcal{A}_{\mathrm{S}}(m_{K^+\pi^0}) = \sqrt{kp} \times B'_{L_B}(k, k_0, d) \left(\frac{k}{m_{B^+}}\right)^{L_B} \times B'_{L_{K^{*+}}}(p, p_0, d) \left(\frac{p}{m_{K^{*+}}}\right)^{L_{K^{*+}}}$$
$$\times \left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B}\frac{1}{\cot\delta_R - i}\right) ,$$
$$\cot\delta_B = \frac{1}{ap} + \frac{1}{2}rp \, , \tag{7.12}$$
$$\cot\delta_R = \frac{m_{K^{*+}}{}^2 - m_{K^+\pi^0}{}^2}{m_{K^{*+}}\Gamma(m_{K^+\pi^0})} .$$

The parameter a represents the scattering length and r is the effective range parameter.

Their values $a = 1.95$ and $r = 1.78$ are taken from Ref. [174]. A study about the influence of these parameters is done in Ref. [3]. The impact of varying these two parameters on the angular observables is negligible.

For the S-wave, angular momenta is $L_B = 0$ and $L_{K^{*+}} = 0$. This simplifies Eq. 7.12 to

$$\mathcal{A}_{\mathrm{S}}(m_{K^+\pi^0}) = \sqrt{kp} \times \sqrt{\frac{1 + (k_0 d)^2}{1 + (k\ d)^2}} \times \frac{k}{m_{B^+}} \times \left(\frac{1}{\cot \delta_B - i} + e^{2i\delta_B} \frac{1}{\cot \delta_R - i} \right). \qquad (7.13)$$

The final amplitude in the $m_{K^+\pi^0}$ dimension is then a combination of the squared normalized P- and S-wave amplitudes, using the S-wave fraction F_S

$$\left. \frac{\mathrm{d}\Gamma}{\mathrm{d}m_{K^+\pi^0}} \right|_{\mathrm{S+P}} = (1 - F_S) \left| \mathcal{A}'_{\mathrm{P}}(m_{K^+\pi^0}) \right|^2 + F_S \left| \mathcal{A}'_{\mathrm{S}}(m_{K^+\pi^0}) \right|^2. \qquad (7.14)$$

7.2.2 Background component

As the background contribution is high especially at large $\cos(\theta_K)$, as discussed in Sec. 6.2, a dedicated study on a predominantly background data sample is done. This sample consists of all events passing the selection described in Sec. 5 in the resonant J/ψ dimuon invariant mass squared interval with the reconstructed B^+ meson mass higher than $5629\,\mathrm{MeV}$, corresponding to the mass of a B^+ meson $+350\,\mathrm{MeV}$. This rather strict cut is applied in order to make sure the signal tail does not significantly contribute to the background sample.

In the B^+ meson mass dimension, the background mostly consists of random accidental track combinations and is described by an exponential with one free parameter. Similarly, in the case of $m_{K^+\pi^0}$, a linear model describes this combinatorial background well. The fit projections of the B^+ and K^* reconstructed mass distributions are depicted in Fig. 7.1.

In the dimension of the decay angles $\cos(\theta_K)$, $\cos(\theta_L)$, ϕ, the background is parameterized using Chebyshev polynomials T_i [175]. The background in each dimension is described by a dedicated Chebyshev polynomial. This factorization is possible, as a study of the background sample shows no correlation between the angles, as shown in Fig. 7.2.

The $\cos(\theta_L)$ angular background is modeled with a polynomial of order two, $\cos(\theta_K)$ is modeled with a polynomial of order five in the reference channel and order of two in the signal channel (this is explained in the next paragraph), and ϕ angular background is flat. The angular background is then described by Eq. 7.15:

$$\left. \frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{d}\cos(\theta_K)\,\mathrm{d}\cos(\theta_L)\,\mathrm{d}\phi} \right|_{\mathrm{BKG}} = \left(\sum_{i=0}^{5(2)} c_i^{\theta_K} T_i(\cos(\theta_K)) \right) \times \left(\sum_{j=0}^{2} c_j^{\theta_L} T_j(\cos(\theta_L)) \right) \times \left(c_0^{\phi} T_0(\phi) \right).$$
$$(7.15)$$
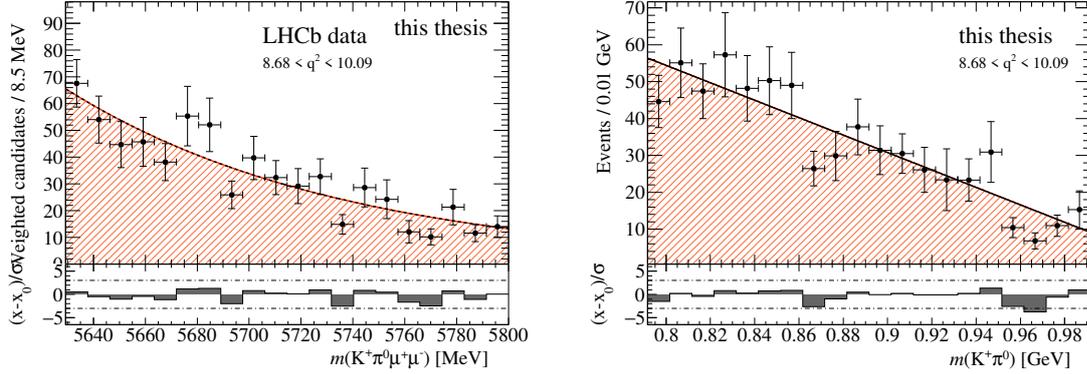
Fig. 7.1: Fit to the background sample in the B^+ meson and K^* mass dimensions. The black markers represent the data, the red area represents the background fit model. The reconstructed B^+ meson mass distribution (left) is fitted with an exponential function, the K^* meson mass distribution (right) with a linear function.
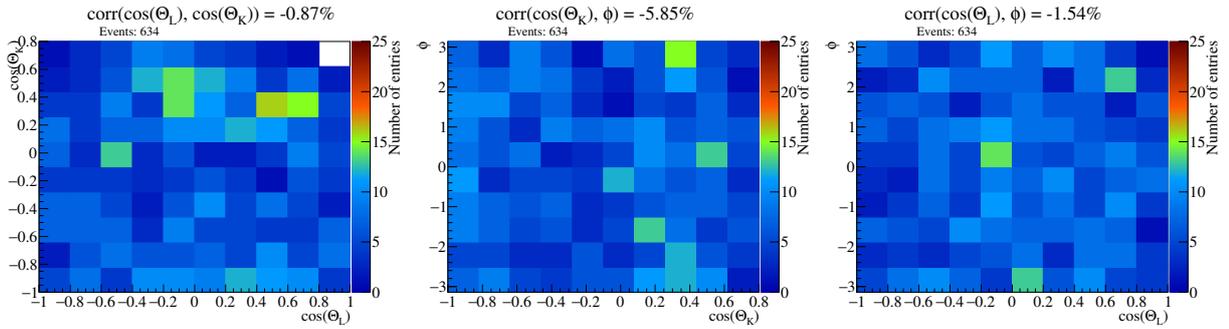
Fig. 7.2: Correlation between the decay angles in predominantly background sample. The correlation between the angular coefficients is in the order of couple of percent, proving there is no correlation. Hence three independent Chebyshev polynomials can be used for the description of the angular background.

The fit projections to the decay angles are shown in Fig. 7.3. The crucial part of the description is the peaking structure at high $\cos(\theta_K)$ values. Correct modeling of this peak is crucial for extracting the angular moments. The lowest possible order of the Chebyshev polynomial describing the $\cos(\theta_K)$ background well is the order of five. However, even when considering the large statistical sample of the $B^+ \to K^{*+}J/\psi$ decay, the five free parameters tend to overfit the data. The χ^2 of the fit to the background sample in the $\cos(\theta_K)$ is equal to only 0.346. This can be avoided either by cutting even harder on the high $\cos(\theta_K)$ or by describing the background by a lower order Chebyshev polynomial. Cutting away more events with high $\cos(\theta_K)$ leads to lower sensitivity to the angular parameters, especially the parameter F_L. Lower order polynomial does not describe the shape of the background well, especially in the regions at $\cos(\theta_K) \approx -1$ and $\cos(\theta_K) \approx 0.8$. This problem disappears when considering the low statistical power

of the rare channel: Chebyshev polynomial of order of two is sufficient to describe the background contribution. The overfitting is present only in the reference $B^+ \to K^{*+} J/\psi$ decay and manifested by the third order polynomial parameter running into the boundary of this parameter[28]. This parameter controls the shape of the plateau at $\cos(\theta_K) \approx -0.4$. As this is a nuisance parameter and a wide range of values describes the background well, this parameter is left floating in the fit to the reference $B^+ \to K^{*+} J/\psi$ decay.
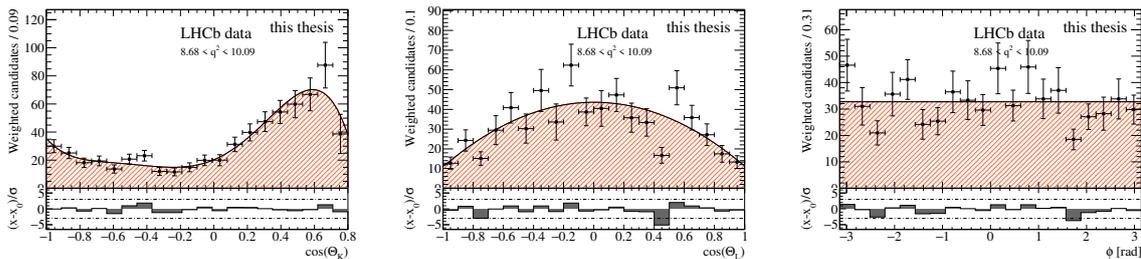
Fig. 7.3: Angular fit of the background sample. The sample is obtained from the $B^+ \to K^{*+} J/\psi$ data taken during both Run I and Run II. The black markers represent the data, the red area represents the background model described by Eq. 7.15. On the left, the $\cos(\theta_K)$ distribution is presented, in the middle the $\cos(\theta_L)$ distribution, and on the right ϕ distribution is shown.

---

[28]Of course this can be avoided by enlarging the range of this free parameter. However, this lead to other parameters running into the boundary.

7.3   Extraction of the $\mathbf{F_S}$ parameter

As mentioned in Sec. 3.3, it is impossible to distinguish the contribution of the P-wave and the S-wave in the $K^+\pi^0$ system at the selection level. However, using the reconstructed mass of K^* meson, statistical selection is possible. This can be done by performing a two-dimensional fit in the K^* meson mass and the B^+ meson mass. First, the 2D fit is performed in the reference $B^+ \to K^{*+}J/\psi$ channel, as it is much more abundant than the rare $B^+ \to K^{*+}\mu^+\mu^-$ channel. The projections in the m_{B^+} and $m_{K^+\pi^0}$ dimensions are shown in Fig. 7.4.
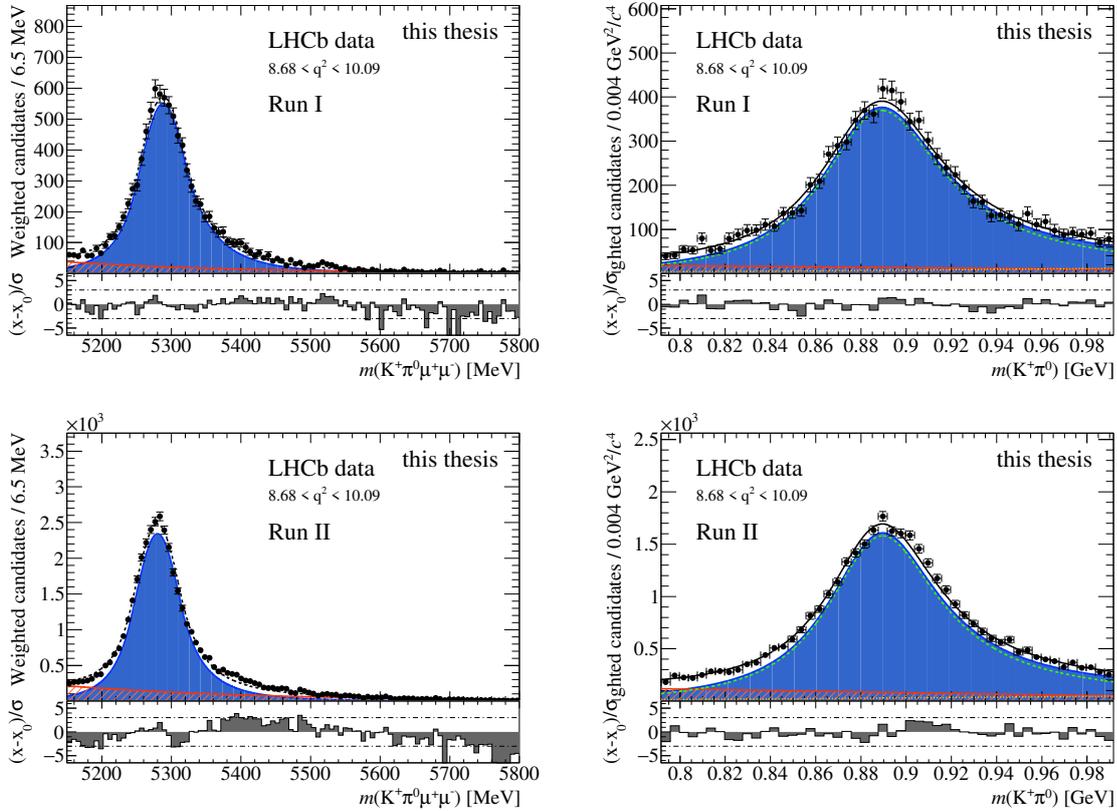
Fig. 7.4: Fit projections to the reference decay channel in reconstructed B^+ meson (left) and K^* meson (right) masses. The top row represents Run I sample, the bottom row Run II sample. The black markers represent the data, the blue space represents the signal. Red surface represents the background contribution. The green dashed line represents the P-wave, the green dotted line, present under the background area, represents the S-wave contribution.

The study of $B^+ \to K^{*+}\mu^+\mu^-$ is performed in multiple q^2 bins. Looking at Eq. 7.11 and Eq. 7.13, the distributions are q^2 independent. Despite the effort to fit the F_S in the $B^+ \to K^{*+}\mu^+\mu^-$ channel using the full q^2 range (excluding the resonance regions), the statistical power of the sample is not large enough to find a contribution of the S-wave in the $K^+\pi^0$ system and the fit prefers $F_S = 0$. The fit projections in the m_{B^+} and $m_{K^+\pi^0}$ dimensions are given in Fig. 7.5.

A dedicated study of pseudoexperiments is done in order to establish the sensitivity of the fit on the F_S parameter in the rare $B^+ \to K^{*+}\mu^+\mu^-$ channel. It is shown in Sec. 9.4 that fixing F_S and the interference terms S_{Si} to zero does not introduce bias on the measured P-wave angular parameters.
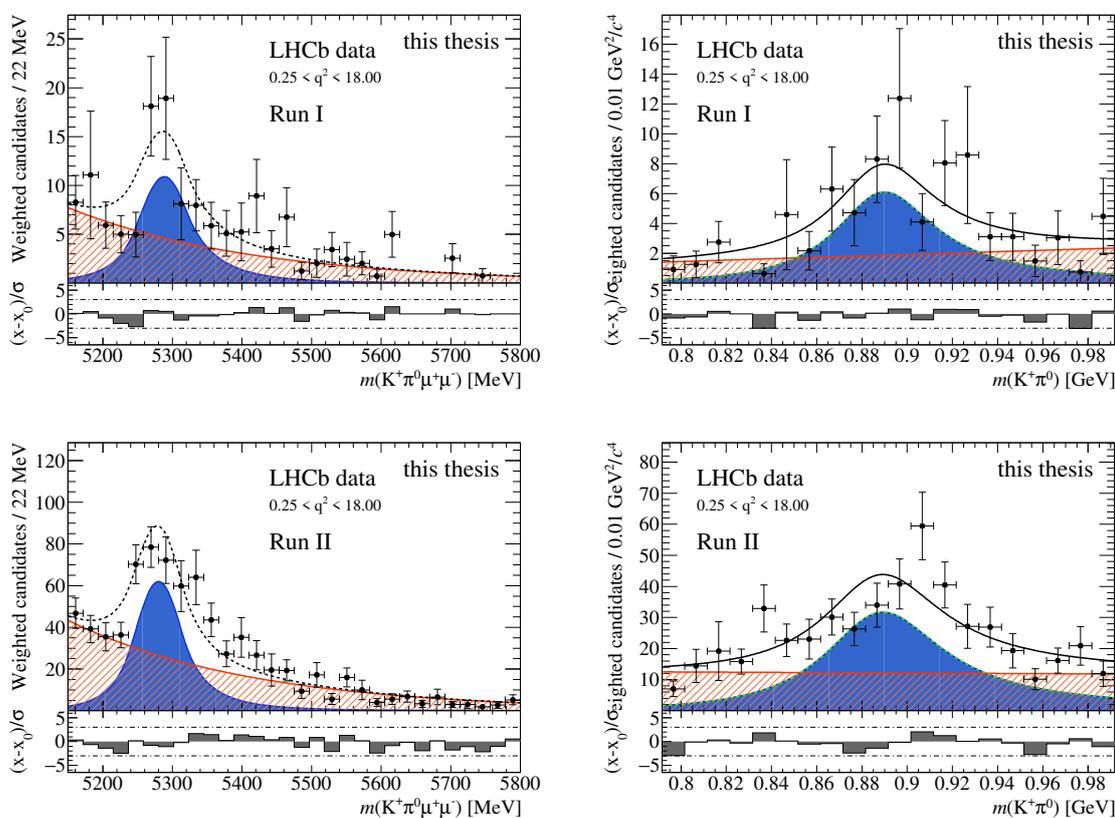
Fig. 7.5: Fit projections to the signal channel in reconstructed B^+ meson (left) and K^* meosn (right) masses. Top row represents Run I sample, bottom row Run II sample. The black markers represent the data, the blue space represents the signal. Red surface represents the background contribution. As the fit prefers $F_S = 0$, the S-wave contribution is not plotted. The green dashed line represents the P-wave contribution to the K^* mass.

8   Tests with large statistics

In order to validate the FCNCFITTER framework and the functionality of the angular acceptance correction, tests with large statistical samples are performed. First, the simulation samples introduced in Sec. 4.4 are fitted. As the Monte Carlo simulation uses a form-factor model BTOSLLBALL [176], the extraction of the initial angular moments F_L, S_3, S_4, S_5, A_{FB}, S_7, S_8 and S_9 very complicated. Therefore, instead of extracting the angular moments from the BTOSLLBALL model itself, a *generator-level* simulation is studied in order to obtain the values of the angular parameters used at generation. Generator level sample is free of any acceptance, reconstruction or selection effects. The agreement of the angular moments of the full Monte Carlo simulation and the generator level simulation is a crucial validation of the angular acceptance correction procedure.

The full simulation sample is also used to validate the folding procedure described in Sec. 3.5. The large statistics allows for a full angular fit. The angular observables obtained by the full angular fit are compared to the results of the folded fit.

The last step of the validation is the fit to the reference $B^+ \rightarrow K^{*+} J/\psi$ channel. The statistical power of the reference channel allows to test the functionality of the FCNCFITTER framework also on the data with present background contribution. The fit is validated by comparing the measured angular moments to previous measurements. Moreover, the $B^+ \rightarrow K^{*+} J/\psi$ decay data sample is also used to further validate the folding method.

8.1   Fit to the simulation sample

The simulation sample used for the validation is treated the same way as the real data: the events pass the LHCb acceptance and the full selection. The angular acceptance correction is applied. The fit is performed using the same q^2 binning as in the data sample. There is no S-wave pollution present in the simulation sample: only the P-wave is fitted. The fit projections are shown in Fig. 8.1. The fit converges and describes the data very well.

8.1 Fit to the simulation sample
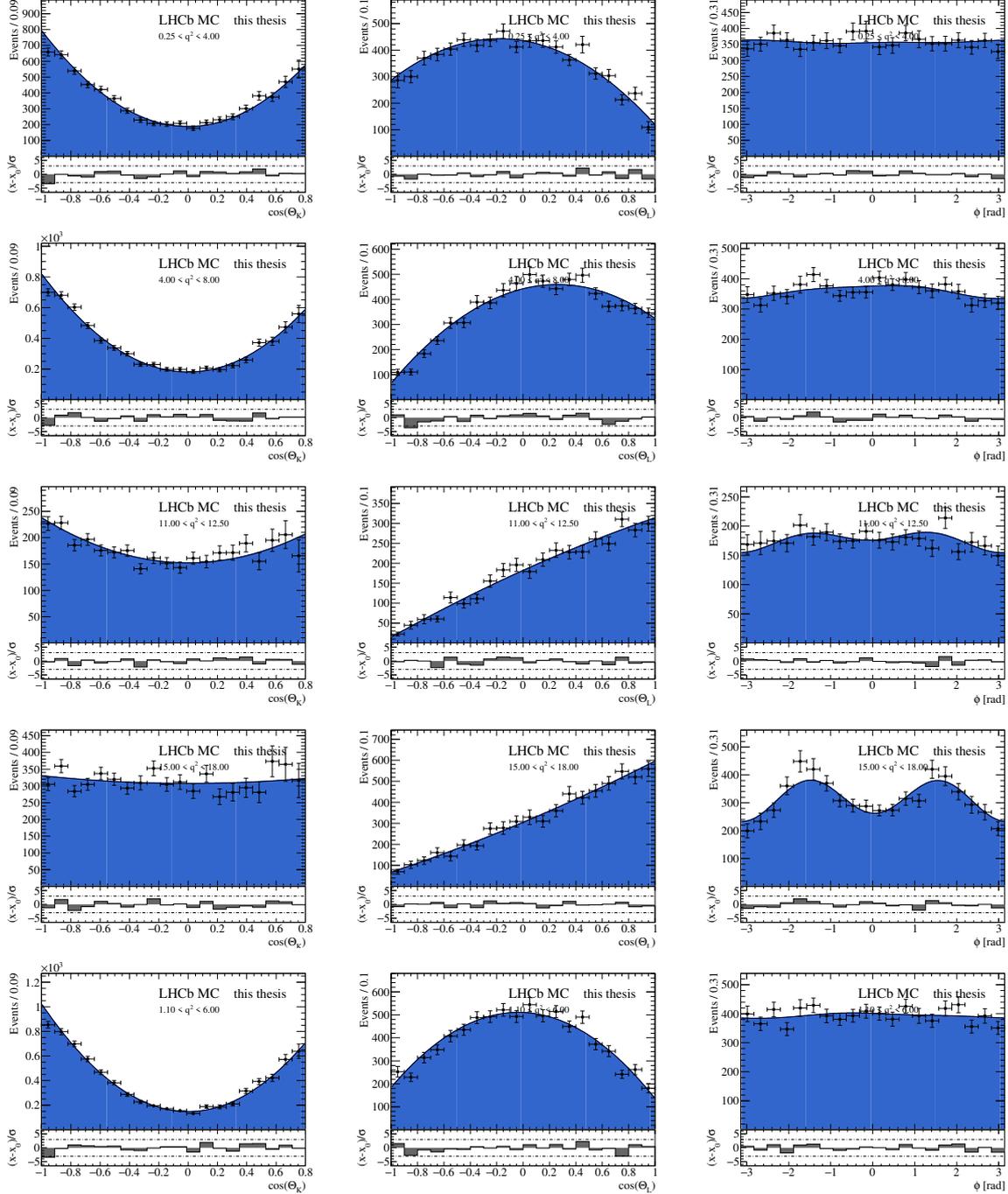
Fig. 8.1: Projections of the fit to the simulated $B^+ \to K^{*+}\mu^+\mu^-$ decay sample. All events are weighted according to the acceptance correction function. The black markers represent the data, blue area represents the fit. Each figure represents one q^2 interval (q^2 range of the interval is denoted in the figures in GeV^2). On the left, $\cos(\theta_K)$ projecitons are shown, in the middle $\cos(\theta_L)$ projections and on the right ϕ projecitons are shown.

8.2   Generator level simulation fit

In order to extract the angular parameters used for the LHCb Monte Carlo simulation, an independent generator level sample of $200\,000$ B^+ mesons decaying at rest to $K^{*+}{}_{[K^+\pi^0]}\mu^+\mu^-$ have been generated. In Fig. 8.2, the measured values of P-wave angular moments are shown. The agreement between the LHCb simulation and the generator level event simulation is very good. The difference between them is below three standard deviations in all bins and all variables, showing the functionality of the angular acceptance corrections in all q^2 regions.

8.3   Validation of the folding method

In order to validate the FCNCFITTER framework's folding method classes, a check using the LHCb simulation sample of $B^+ \to K^{*+}\mu^+\mu^-$ is performed. The sample is fitted using the full angular description as well as using all five folding techniques described in Sec. 3.5. The results of the full fit are compared to the fit results using the five folding methods. As shown in Fig. 8.3, there is a perfect agreement between all five folding methods and the full angular fit in all q^2 bins and all angular observables.

Fig. 8.2: Fit to the generator level simulation compared to the fit to the LHCb simulation results. The brown stripes represent the resonant q^2 regions. These regions are excluded from the fit. The red boxes represent the difference between the processed LHCb simulation and the generator level simulation in terms of standard deviations, σ. The fitted values are in agreement, proving the functionality of the angular acceptance correction in all q^2 regions.

90
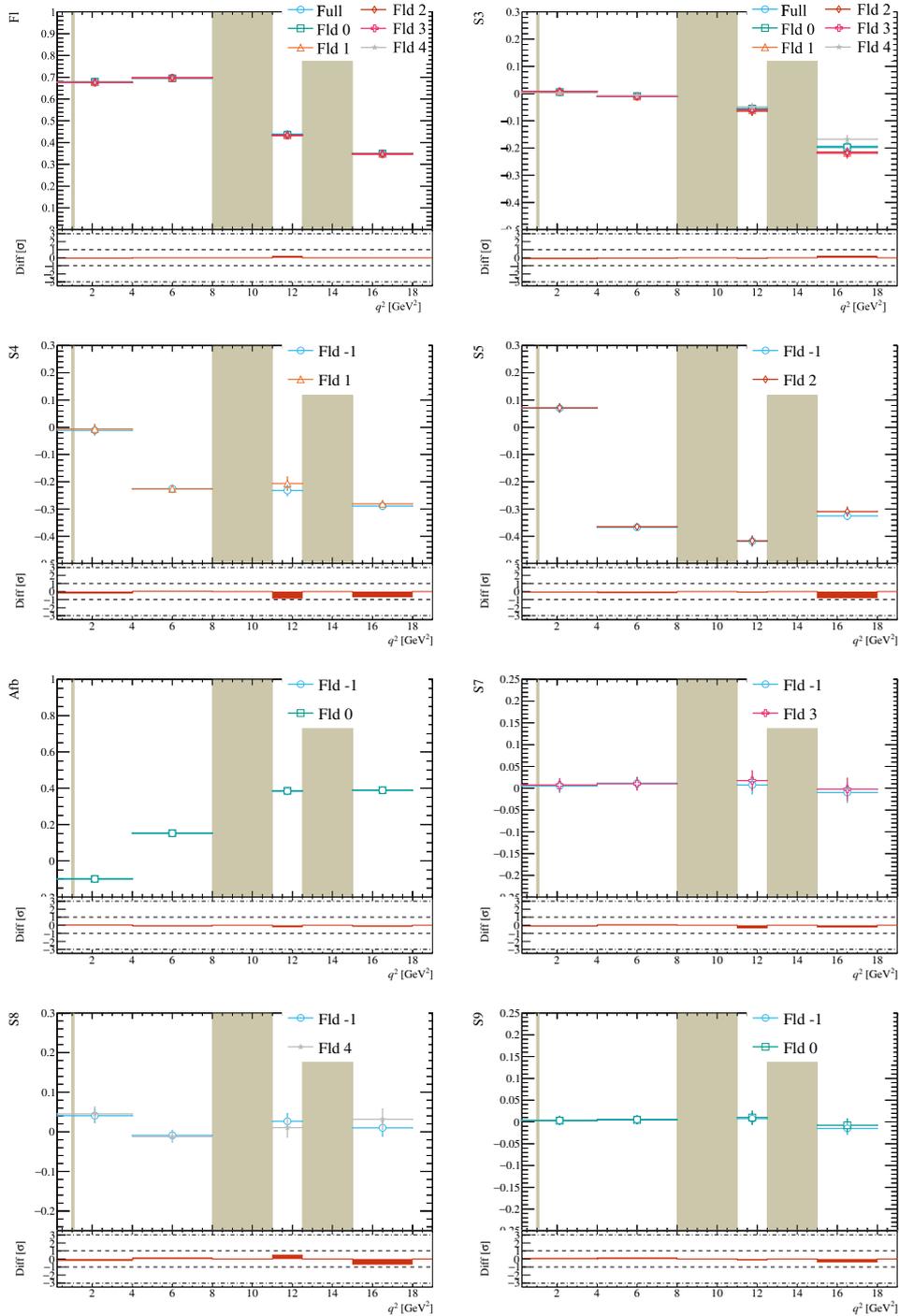
Fig. 8.3: Full angular fit compared to fits using angular folding method as indicated in each figure by 'Fld'. The brown stripes represent the resonance q^2 regions. These regions are excluded from the fit. The red boxes represent the difference between the processed LHCb simulation and the generator level simulation in terms of standard deviations σ. The results from the full angular fit and the fits using the folding methods are in perfect agreement in all q^2 regions. This proves the functionality of the angular acceptance correction in the folded fits.

91

8.4   Fit to the $\mathbf{B^+ \to K^{*+} J/\psi}$

As the statistical power of the rare $B^+ \to K^{*+} \mu^+ \mu^-$ channel is highly limited, tests are performed using the $B^+ \to K^{*+} J/\psi$ data sample. Candidates with a dimuon invariant mass squared between 8.68 and 10.09 GeV^2 are considered. The data sample is split between the Run I and Run II samples. The angular parameters are shared between both samples. The two samples are fitted simultaneously in four dimensions of the B^+ meson reconstructed mass m_{B^+} and the angles θ_K, θ_L, ϕ exploiting the maximum-likelihood method. The parameter F_S is extracted from a two-dimensional fit to the B^+ meson mass and the K^{*+} mass. The statistical power of this sample is large enough to test the functionality of the FCNCFITTER framework as well as all corrections applied to the data.

The projection of the full fit is presented in Fig. 8.4. The results of this fit are compared to previous measurements by BaBar studying both decays of B^+ and B^0 mesons to $K^* \mu^+ \mu^-$ [177], Belle focusing on the $B^+ \to K^{*+} J/\psi$ decay [178], and LHCb measurements of the $B^0 \to J/\psi K^{*0}$ decay [179] and of the $B^+ \to K^{*+} \mu^+ \mu^-$ decay with $K^{*+} \to K^0_S \pi^+$ [4], where the fit to $B^+ \to K^{*+} J/\psi$ is also performed as an important check of the fitter framework. The results of the full angular fit are shown in Tab. 8.1. The statistical uncertainty is approximated by the HESSE calculation (see Sec. 9.3) and scaled according to the pull widths listed in Tab. 9.1. For readers convenience, the difference between this measurement and the measurement listed in Tab. 8.1 are shown in terms of standard deviations in Tab. 8.2.

The measured values from the full angular fit agree very well with the other measurements. The only exception is the parameter S_4 that is consistently below the values measured by other experiments. This can be caused by the complicated modeling of the angular acceptance: the parameter S_4 is very sensitive to the symmetry of the $\cos(\theta_L)$ distribution. However, the angular acceptance description does not always prefer a symmetry in $\cos(\theta_L)$, see App. C. A dedicated test by varying the order of the polynomial used to correct the angular acceptance in $\cos(\theta_L)$ can be done.

Moreover, the reference channel $B^+ \to K^{*+} J/\psi$ is used to further test the angular folding method. The values obtained using the full angular fit are compared to the values obtained via folding methods 0-4 listed in Tab. 8.3. The agreement between the results is almost perfect. The projections of these fits are shown in Fig. 8.5.
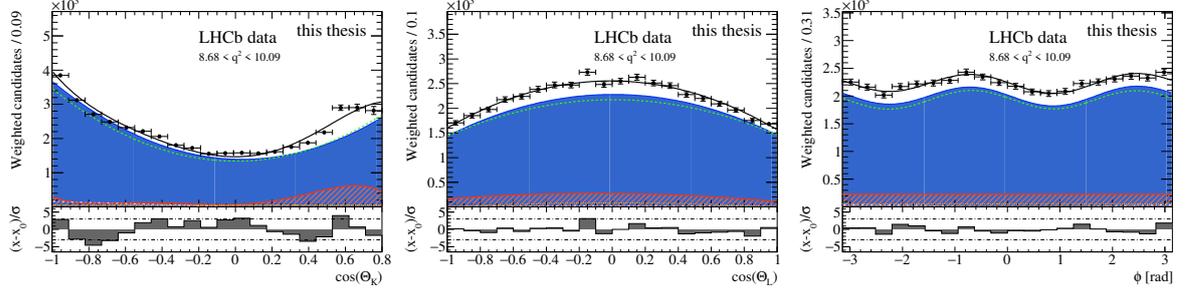
Fig. 8.4: Full angular fit to the reference channel $B^+ \to K^{*+} J/\psi$. The black markers represent the data. The black line shows the full fit model. The blue space represents the signal contribution. From left to right, $\cos(\theta_K)$, $\cos(\theta_L)$ and ϕ projections are shown. Red surface represents the background contribution. The green dashed line shows only the P-wave component, the orange dotted line represents the S-wave components and the pink dot-and-dash line depicts the interference between the P-wave and the S-wave.

	this fit	LHCb B^+	Belle (B^+)	BaBar $(B^+ + B^0)$	LHCb B^0
F_L	0.563±0.014	0.572±0.005	0.604±0.015	0.556±0.009	0.572±0.008
S_3	0.014±0.011	-0.002±0.007	-0.018±0.017	0.011±0.011	-0.013±0.010
S_4	-0.211±0.014	-0.246±0.008	-0.255±0.010	-0.237±0.007	-0.250±0.006
S_5	-0.013±0.015	-0.003±0.008	0.000±0.000	0.000±0.000	0.000±0.000
A_{FB}	0.002±0.007	-0.002±0.005	0.000±0.000	0.000±0.000	0.000±0.000
S_7	0.002±0.014	-0.001±0.008	0.000±0.000	0.000±0.000	0.000±0.000
S_8	-0.062±0.015	-0.063±0.008	-0.037±0.018	-0.058±0.015	-0.048±0.007
S_9	-0.074±0.011	-0.084±0.007	-0.041±0.016	-0.095±0.014	-0.084±0.006

Tab. 8.1: Comparison of the full angular fit to reference channel $B^+ \to K^{*+} J/\psi$ to previously done measurements by BaBar, Belle and two LHCb measurements [177, 178, 179, 4], focusing on different combinations of B^+ and B^0 meson decays, as indicated. The measurements are published in the form of polarization amplitudes. The amplitudes are converted into the S_i angular observables and the uncertanities are propagated to the basis using 100 000 randomly generated samples. The full angular fit results are in agreement with the previously published measurements. The statistical uncertainty of this result is scaled according to Tab. 9.1, as the pseudoexperiment studies showed an overestimation of the statistical uncertainities (for the details see Sec. 9). Dark green represents an agreement better than one standard deviation, lime represents an agreement better than two standard deviations and yellow represents an agreement better than three standard deviations.

93

	LHCbB^+	Belle (B^+)	BaBar ($B^+ + B^0$)	LHCbB^0
F_L	-0.61	-2.00	0.42	-0.56
S_3	1.23	1.58	0.19	1.82
S_4	2.17	2.56	1.66	2.56
S_5	-0.59	-0.87	-0.87	-0.87
A_{FB}	0.46	0.29	0.29	0.29
S_7	0.19	0.14	0.14	0.14
S_8	0.06	-1.07	-0.19	-0.85
S_9	0.77	-1.70	1.18	0.80

Tab. 8.2: The difference between the full angular fit to reference channel $B^+ \to K^{*+} J/\psi$ and the previously done measurements by BaBar, Belle and two LHCb measurements [4, 177, 178, 179] in terms of the standard deviations. The measurements are published in the form of polrization amplitudes. The amplitudes are converted into the S_i angular observables and the uncertanities are propagated to the basis using 100 000 randomly generated samples. The full angular fit results are in agreement with the previously published measurements.

	Full angular	Folding 0	Folding 1	Folding 2	Folding 3	Folding 4
F_L	0.563±0.014	0.563±0.012	0.565±0.011	0.564±0.011	0.564±0.012	0.564±0.019
S_3	0.014±0.011	0.015±0.005	0.015±0.005	0.016±0.005	0.016±0.005	0.016±0.005
S_4	-0.211±0.014	—	-0.218±0.008	—	—	—
S_5	-0.013±0.015	—	—	-0.012±0.007	—	—
A_{FB}	0.002±0.007	0.001±0.004	—	—	—	—
S_7	0.002±0.014	—	—	—	0.002±0.007	—
S_8	-0.062±0.015	—	—	—	—	-0.069±0.010
S_9	-0.074±0.011	-0.074±0.005	—	—	—	—

Tab. 8.3: Comparison of the full angular fit results to reference channel $B^+ \to K^{*+} J/\psi$ to the fits using angular folding method. The results are in perfect agreement. This proves the functionality of the folding methods.
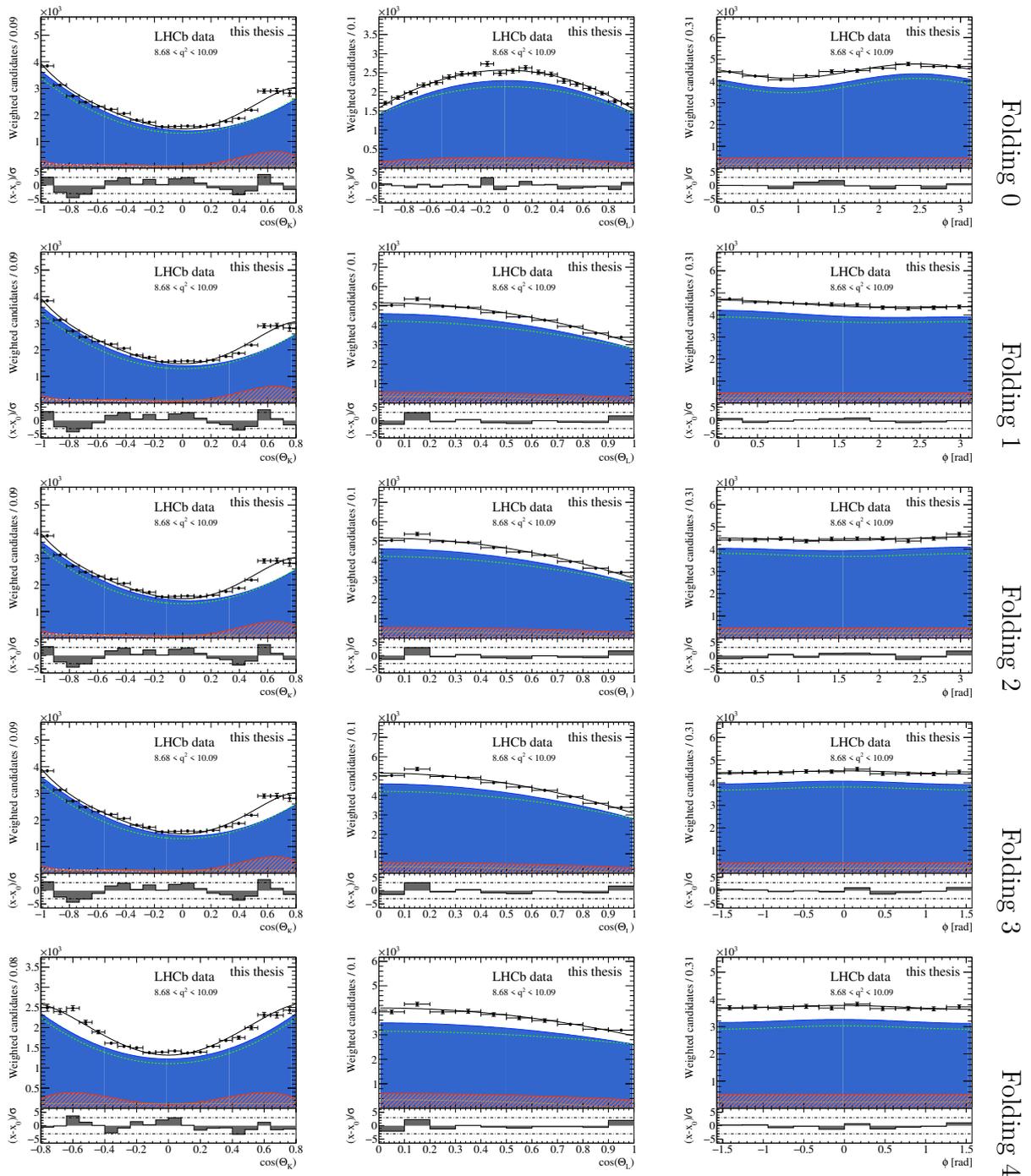
Fig. 8.5: Full angular fit to the reference channel $B^+ \to K^{*+} J/\psi$ for the five folding methods. The black markers represent the data, the blue space represents the signal contribution. On the left, $\cos(\theta_K)$ projeciton is shown, in the middle $\cos(\theta_L)$ projection and on the right ϕ projection is shown. Red surface represents the background contribution. The green dashed line shows only the P-wave component, the orange dotted line represents the S-wave components and the pink dot-and-dash line depicts the interference between the P-wave and the S-wave.

9   Pseudoexperiments

In order to validate the FCNCFITTER framework and its settings, dedicated tests on pseudoexperiments are performed. These tests are designed to verify the estimation of the fitted values as well as their associated statistical uncertainties. The tests are necessary as the FCNCFITTER framework is a rather complex tool. Moreover, the limited statistical power of the sample calls for many constraints on the fitted parameters. These constrains also have to be thoroughly validated.

Therefore, dedicated sets of *pseudoexperiments* are created. *Pseudoexperiment* is a randomly-generated set of *pseudoevents*. The pseudoevents are events generated according to a simplified model. Such simplified model, or a *toy* model, allows to study specific cases of the framework configuration as well as the influence of specific parameter values on the final fit result. Moreover, these studies can be done on arbitrarily large samples, minimizing the influence of limited statistical power.

The goal of the pseudoexperiment studies is to validate and correct the functionality of the FCNCFITTER framework and therefore to obtain a bias-free result with good coverage of statistical uncertainty. The focus of these studies is on the angular P-wave parameters, however, the coverage of statistical uncertainty and potential biases are also studied for all other free parameters. The events are generated following the distributions of $\cos(\theta_K)$, $\cos(\theta_L)$, ϕ and q^2 without the angular acceptance weights applied. In order to study the influence of the acceptance weights, the acceptance weights are applied during the fit.

The pseudoexperiments are validated by studying the *pull distributions* of the parameters that are free in the fit. Such distributions represent the difference between the measured value, x, and the generated value, x_0, divided by the uncertainty of the measurement σ,

$$p(x, \sigma) = \frac{x - x_0}{\sigma} \, . \tag{9.1}$$

Using the central limit theorem, the shape of such distribution for any statistically independent random variable follows a Gaussian distribution. The width of such distribution is equal to one, the mean is equal to zero. Any shift from the mean value of zero indicates a bias of the measured value in terms of standard deviations. If the width is larger than one, it suggest the undercoverage of the uncertainty, width smaller than one signals overcoverage of the uncertainty. As an example, a width of 0.25 means the actual statistical uncertainty should be four times as large as the measured one. There are also many reasons why the pull distribution can not be described by a Gaussian distribution, such as that the free parameter is at its limit or there is a technical problem with the minimization.

9.1 Generation of pseudoexperiments

For this analysis, the fitted variables m_{B^+}, $m_{K^{*+}}$, $\cos(\theta_K)$, $\cos(\theta_L)$, ϕ are directly generated using a random number generator. For the generation, the ROOT class `TRandom3` is used. This class is used to generate equi-distributed pseudorandom floating-point numbers in the interval $(0, 1\rangle$ via the Mersenne Twister algorithm [180]. This algorithm has been devoloped in the late 90's and it is widely used as it is relatively fast while passing most statistical tests designed to measure the quality of a random number generator[29].

The `TRandom3` class provides generation of uniform distributions. However, the desired distributions are non-uniform. There are several methods how to convert the uniformly distributed random numbers to produce non-uniform distributions: simplest of them being a simple transformation of the distribution. However, this is possible only for distributions with an analytical integral. Unfortunately, this is rarely the case. A simple method used to generate a non-uniform distribution, *rejection method* [183], is used in this analysis. The method is very similar to numerical computation of integrals.

9.1.1 Rejection method

The goal is to generate a sample of random numbers with a non-uniform distribution function $p(x)$, where x is defined and non-zero in a certain range $(A, B\rangle$. First step is to select a *comparison function* $f(x)$. The comparison function has to be larger $p(x)$ for all $x \in (A, B\rangle$:

$$f(x) > p(x) \quad \text{for } \forall x \in (A, B\rangle. \tag{9.2}$$

Fig. 9.1: Rejection method illustration. The desired generated distribution is denoted $p(x)$, the *comparison function* $f(x)$. The comparison function is always more than the function $p(x)$. The second random deviate of any x_0 is then used to decide whether to accept or reject the point at x_0. If rejected, new second random deviate of f is found. Taken from Ref. [183].

Then, the area below the comparison function $f(x)$ is populated uniformly with random points denoted $[u_x, u_y]$. For each point, the value $f(u_x)$ and $p(u_x)$ is calculated. If $u_y > p(u_x)/f(u_x)$ the point $[u_x, u_y]$ is *rejected*. New point is generated until the condition is satisfied. The ratio of rejected/accepted points is then equal to the ratio of the area between $f(x)$ and $p(x)$ to the area under $p(x)$. Hence, the accepted points follow the distribution $p(x)$. An illustration of this procedure is shown in Fig. 9.1.

---

[29]Among standardized tests are *e.g.* DieHard [181] or the U01 tests [182]. The Mersenne Twister algorithm passes all of DieHard tests and vast majority of the U01 tests.

The main advantage of this method is its variability and simplicity. The distribution function $p(x)$ always has to have a maximum, as the integration over its domain has to be equal to one, is positive and is continuous. Therefore, it is always possible to construct the comparison function as a 'rectangle' above the desired distribution $f(x) = \max_{x \in (A,B)}\{p(x)\}$. On the other hand, this leads to the main disadvantage of the method: when the area below the comparison function $f(x)$ is much larger than the area below $p(x)$, the count of rejected points will be very large, leading to a long computing time. Good comparison function is crucial for an effective generation of non-uniform distributions.

9.2   Validation of the generation of the pseudoexperiments

The FCNCFITTER framework also provides the tools to generate the pseudoexperiments, exploiting the rejection method. In order to validate this functionality of the framework, pseudoexperiments with $B^+ \to K^{*+} J/\psi$ decays are generated.

The first test is to generate only events with the signal component. The events are generated following the distributions explained in Sec. 7.2.1. The mass parameters and the parameter F_S are taken from the fit to the reference decay $B^+ \to K^{*+} J/\psi$, the angular components are generated with values based on the study of the $B^+ \to K^{*+} J/\psi$ decay done for Ref. [4]. According to the mass fit to the reference channel, there are about 52 000
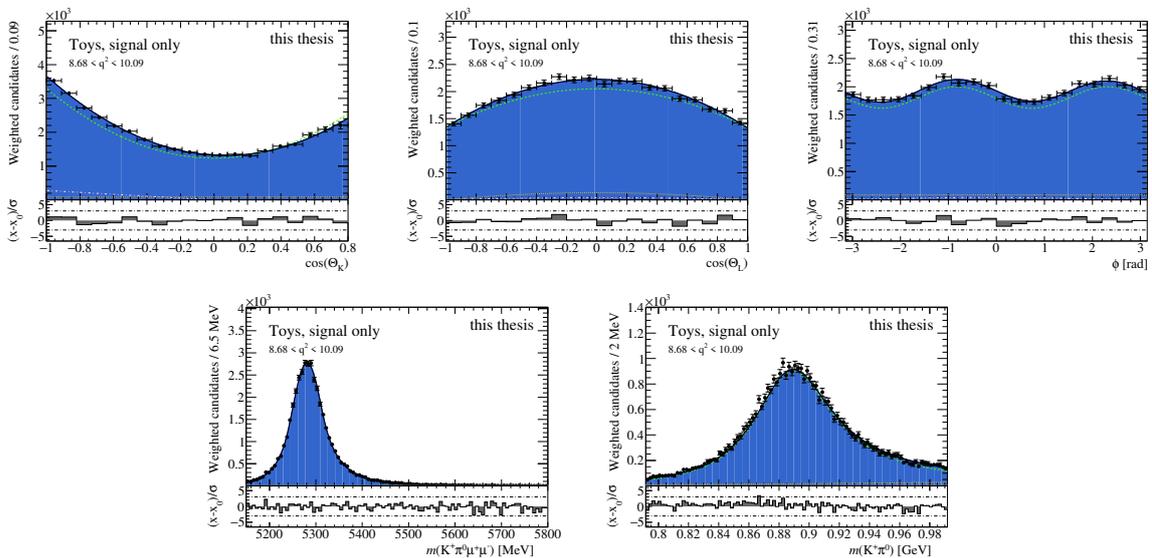
Fig. 9.2: Fit to 52 000 pseudoevents with only the signal component. The projections in $\cos(\theta_K)$, $\cos(\theta_L)$, ϕ, B^+ meson mass, and K^{*+} meson mass are shown. The black points represent the generated pseudoevents, the black line is the fitted model. The blue space represents the signal component. The green dashed line shows only the P-wave component, the orange dotted line represents the S-wave components and the pink dot-and-dash line depicts the interference between the P-wave and the S-wave.

signal candidates in the data sample. Therefore, for this validation 52 000 pseudoevents are generated. The fit to this pseudoexperiment is given in Fig. 9.2. The FCNCFITTER generates the desired distributions and also successfully fits them.

Next step is the validation of the folding technique. As the FCNCFITTER was initially designed without the cut on $\cos(\theta_K)$ in mind, this is a crucial check, especially for folding four (see Eq. 3.10). The fit to the pseudoevents with the folding applied is shownin Fig. C.7 in the appendix C.1. The generated pseudoevents and the fit to the pseudoevents fulfill the expectations and agree with each other.

Similarly, a pseudoexperiment with only the background component is performed. The pseudoevents are generated according to the distributions given in Sec. 7.2.2 and are shown with their corresponding fit in Fig. 9.3. Following the mass fit of the $B^+ \rightarrow K^{*+} J/\psi$ decay, the expected background yield is 13 000 thousand events. Hence, 13 000 events are generated in each pseudoexperiment. The parameters used in the generation are taken from the fit of the background data sample described in Sec. 7.2.2. In the case of the background component, the validation of the folding technique is even more important than in the case of signal only due to the complicated shape of the $\cos(\theta_K)$ background not previously implemented in the FCNCFITTER. As shown in Fig. C.8 in the appendix, the background is successfully generated, folded and fitted. Once again, 500 pseudoexperiments are created and the pull distributions of the free parameters are investigated. The pull distributions are normalized and centralized at zero.
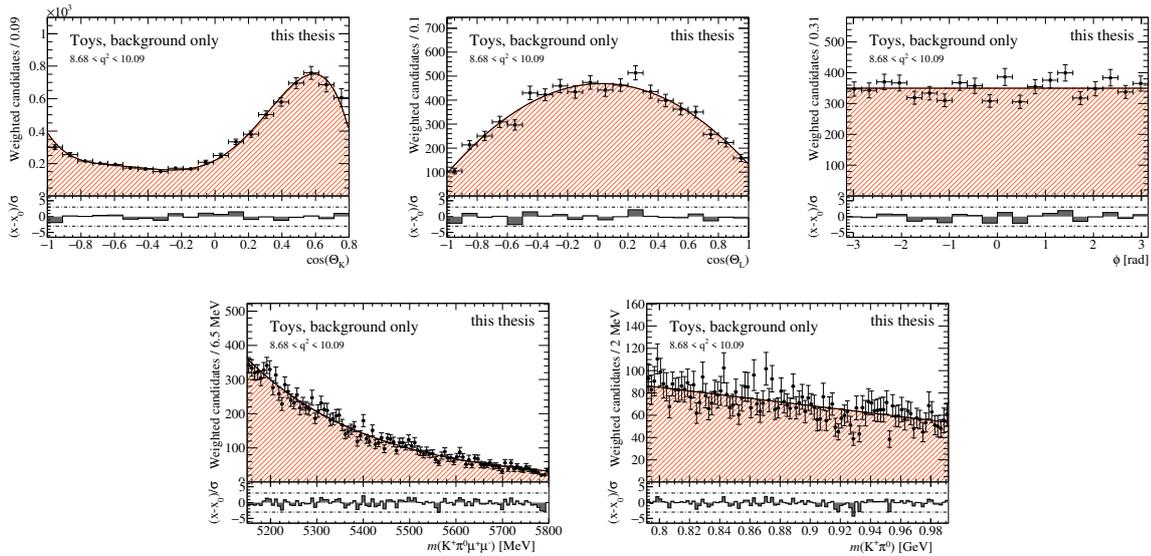
Fig. 9.3: Fit to 13 000 pseudoevents with only the background component. The black points represent the generated pseudoevents, the black line represents the fitted model. The red area represents the background component.

9.3 Large scale pseudoexperiments

Once the pseudoexperiment generation, folding and fitting is validated, the next step is to validate the fit technique using both the signal and the background components. To avoid potential biases of the minimization introduced by low statistical power of the sample, a pseudoexperiment with $65\,000$ thousand events is generated. The fraction of signal events, the mass parameters and the F_S parameters are generated with values obtained by the mass fit to the reference channel $B^+ \to K^{*+}J/\psi$ described in Sec. 7.3, the background parameters are generated with values obtained by the fit to the background sample, described in Sec. 7.2.2. The angular parameters are initialized based on the study of the $B^+ \to K^{*+}J/\psi$ decay done for Ref. [4]. In order to study the pull distributions of these pseudoexperiments, five hundred pseudoexperiments are generated for each folding.

The pull distributions are shown in Fig. 9.4 and Fig. 9.5 and presented in Tab. 9.1. There are no large biases present, the largest one is 6% of the statistical uncertainty in the case of the F_S parameter. However, it is clear that the errors are significantly overestimated: the width of the pull distribution is in the order of 20% for all parameters. This is caused by only an approximate estimation of the uncertainty in the fitter. The standard statistical uncertainties from the maximum-likelihood fit are obtained using the HESSE determination, which derives a symmetric statistical uncertainty by inverting the second derivative of the likelihood function at the best-fit value [184]. For any real physical problem, the covariance matrix has to be positive-definite. However, in the presence of correlated free parameter(s), this matrix can fail the positive-definite requirement. In this case, HESSE forms a positive-definite approximation [164]. Due to modeling of the complicated structure of the $\cos(\theta_K)$ background by the Chebyshev polynomial of order five, discussed Sec. 7.2.2, the parameters describing the $\cos(\theta_K)$ background are highly correlated and therefore only an approximation of the statistical uncertainty is available.

Moreover, the presence of larger acceptance correction weights prevents the simplification of the full covariance matrix expression to the inverse Hessian. Hence HESSE uncertainty determination can no longer guarantee providing correct coverage.

This can be improved by using MINOS [164], which takes into account the parameter correlations and non-linearities. MINOS varies each parameter, minimizing the fit function with respect to the other parameters. This procedure however requires a good previous error estimation and is computationally very intense: running several hundred pseudoexperiments would require significant amount of CPU time. As the shape of the pull distribution using the HESSE approximation is Gaussian, the widths of the pull distributions can be used to correct the statistical uncertainty in the fit of the data.

parameter	mean	width
F_L	-0.008 ± 0.009	0.116 ± 0.006
S_3	0.013 ± 0.018	0.234 ± 0.013
S_4	-0.004 ± 0.014	0.185 ± 0.010
S_5	0.003 ± 0.017	0.229 ± 0.012
A_{FB}	0.008 ± 0.014	0.178 ± 0.010
S_7	0.010 ± 0.017	0.221 ± 0.012
S_8	-0.009 ± 0.016	0.214 ± 0.011
S_9	0.038 ± 0.017	0.218 ± 0.012
F_S	-0.060 ± 0.006	0.074 ± 0.004
S_{S1}	0.053 ± 0.009	0.121 ± 0.007
S_{S2}	-0.038 ± 0.015	0.204 ± 0.011
S_{S3}	0.022 ± 0.016	0.211 ± 0.011
S_{S4}	0.021 ± 0.017	0.220 ± 0.012
S_{S5}	-0.071 ± 0.015	0.199 ± 0.011

Tab. 9.1: The widths and the means of the pull distributions of the angular moments in reference-like pseudoexperiments. 500 pseudoexperiments have been generated, mimicking the refrence $B^+ \to K^{*+} J/\psi$ decay. Each pseudoexperiment consists of $65\,000$ pseudoevents.

Another five sets of pseudoexpriments are created in order to validate the pull distribution of fits exploiting the folding technique. The corresponding pull distributions are presented in App. D.1. The effect of the complicated background structure in $\cos(\theta_K)$ can be nicely seen in Fig. D.5: the odd orders of the polynomials cancel out, leaving only orders of two and four. The correlations between the coefficient is therefore weaker and the width of the pull distribution is close to one.

9.4   Realistic scale pseudoexperiments

Lastly, the statistical properties of the fit are investigated by creating a set of 500 pseudoexperiments designed to resemble the $B^+ \to K^{*+} \mu^+ \mu^-$ decay. The fraction of signal pseudoevents and the mass parameters are generated with values obtained by the mass fit to the $B^+ \to K^{*+} \mu^+ \mu^-$ channel described in Sec. 7.3. The fraction of signal pseudoevents for each q^2 bin and each Run are listed in Tab. 9.2. The background parameters are generated with values obtained by the fit to the background sample, described in Sec. 7.2.2. The background is generated according to Eq. 7.15: the order of the Chebyshev polynomial in the $\cos(\theta_K)$ dimension is five. The P-wave angular parameters are initialized to

f_{sig}	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
Run I	0.33	0.35	0.51	0.53	0.25
Run II	0.36	0.28	0.43	0.61	0.32

Tab. 9.2: Fraction of signal pseudoevents in the pseudoexperiments for each of the q^2 bins and for each Run. The fraction is based on the mass fit to the data sample of the $B^+ \to K^{*+} \mu^+ \mu^-$ decay candidates.
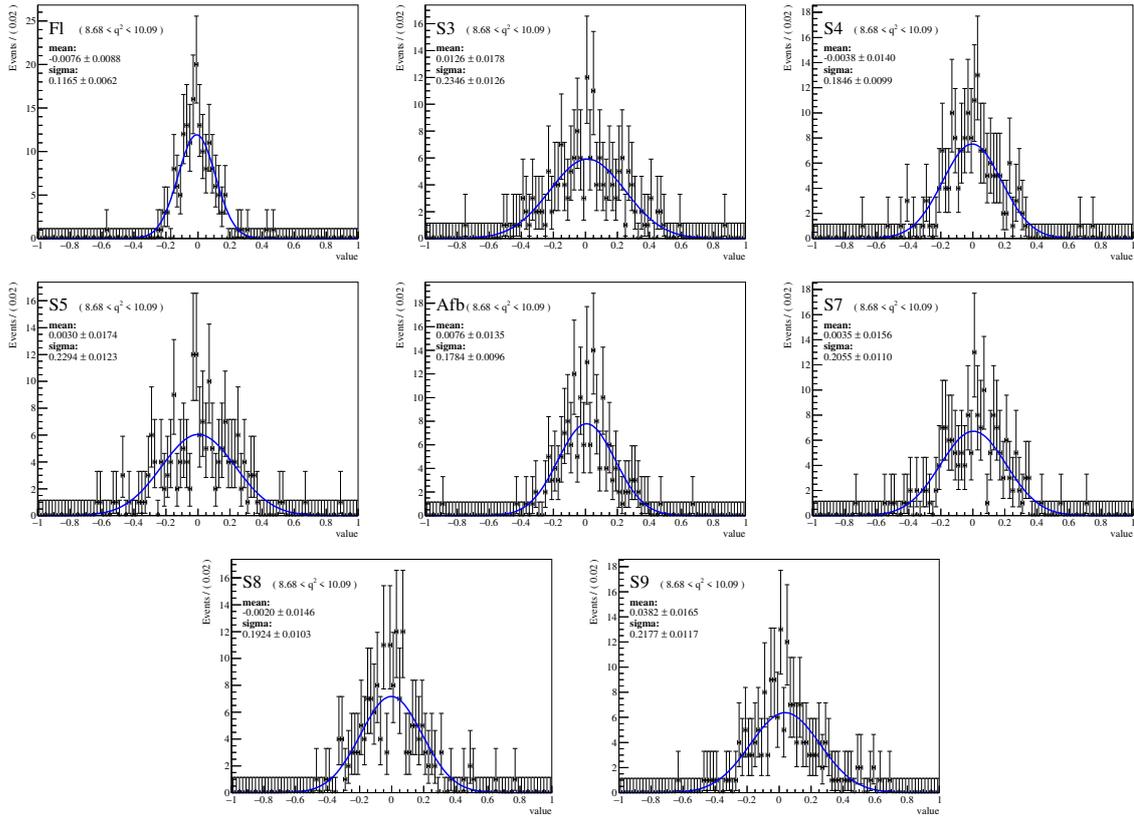
Fig. 9.4: Pull distributions of the P-wave angular moments. 500 pseudoexperiments have been generated, mimicking the refrence $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 65 000 pseudoevents.
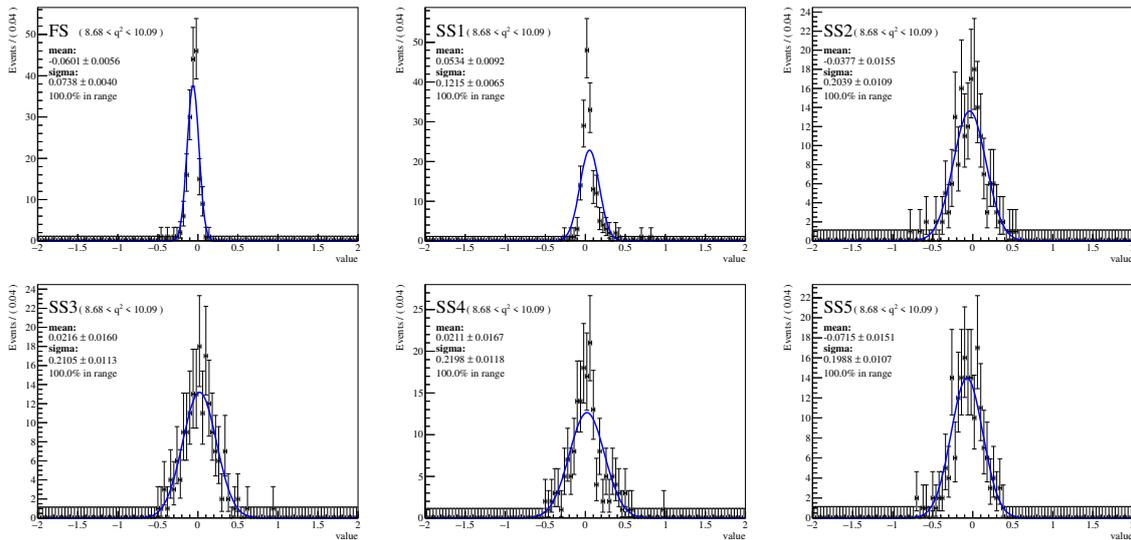
Fig. 9.5: Pull distributions of the S-wave angular moments. 500 pseudoexperiments have been generated, mimicking the refrence $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 65 000 pseudoevents.

the Standard Model values obtained using the FLAVIO package [139]. The F_S is initialized to be 0.25, based on the *maximal* value of F_S observed in Ref. [4]. The interference angular parameters are initialized according to the values measured in Ref. [4]. For each pseudoexperiment, 871 pseudoevents are generated. This is motivated by the number of selected signal candidates in Tab. 5.10. In the fit, the free parameters are the P-wave angular parameters, the angular background parameters, the fraction of signal pseudoevents to all pseudoevents, and the exponential parameter describing the B^+ mass background.

Due to the low number of available candidates in the sample, it is not possible to use the complicated model of the $\cos(\theta_K)$ background shown in Sec. 7.2.2. Hence, a dedicated test has been done to show that a Chebyshev polynomial of order of two is sufficient to describe the $\cos(\theta_K)$ background shape. In this test, despite the generated distribution following the Chebyshev polynomial order of five, only the parameters up to the order of two are left floating, the rest is set to zero. The pull distributions created from 500 pseudoevents are summarized in Tab. 9.3. The simplified background description does not bring any significant bias to the floating angular parameters[30]. Furthermore, by lowering the order of the Chebyshev polynomial, the overestimation of the statistical uncertainty is reduced to $\sim 90\%$.

The largest discrepancy of the mean of the pull distribution from zero is present for the parameter F_L, the bias reaches up to 24% of the statistical uncertainty in the bin in $q^2 \in [11.0, 12.5]\,\text{GeV}^2$. The folding technique is also tested and the results are shown in Tab. 9.4. The bias in F_L can be reduced by using the folding technique, but it does not disappear. The parameter F_L is very sensitive to $\cos(\theta_K)$ and therefore small bias caused by not properly describing the background shape is expected. Another more prominent bias in the full angular fit is present in the case of A_{FB} in $q^2 \in [4.0, 8.0]\,\text{GeV}^2$. However, this bias disappears by using the folding 0, see Tab. 9.4.

Also in the case of folded fit, the width of the pulls is ≈ 0.9. The only exception is S_3 and S_8 with folding four applied: in folding four, the $\cos(\theta_K)$ is also folded. In this case, parabolic fit was not describing the folded background well enough, hence the fourth Chebyshev polynomial parameter is added. This leads to a correlation present in the background description, resulting in the smaller width of the pulls. Contrary to this, the width of the F_L pulls are up to 1.340. However, using this folding, the measurement of parameter F_L is also biased by up to 45% of the statistical uncertainty. This indicates the large sensitivity of F_L on proper background description.

---

[30]This does not hold for samples with more statistical power: a dedicated test is done using pseudoexperiments with 10 times more pseudoevents than currently present in the data. The simplified background description using the Chebyshev polynomial in $\cos(\theta_K)$ of order of two does not describe the background well enough anymore and the pulls of the angular parameters evince biases of up to 55% of the standard statistical uncertainty.

means					
parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L	0.164 ± 0.045	0.002 ± 0.047	0.237 ± 0.042	0.163 ± 0.043	0.184 ± 0.043
S_3	-0.025 ± 0.039	-0.017 ± 0.040	-0.022 ± 0.039	-0.051 ± 0.038	0.017 ± 0.042
S_4	0.072 ± 0.044	-0.113 ± 0.038	0.043 ± 0.038	-0.035 ± 0.041	-0.066 ± 0.038
S_5	0.030 ± 0.042	-0.046 ± 0.036	0.061 ± 0.039	0.029 ± 0.040	0.008 ± 0.041
A_{FB}	-0.137 ± 0.044	0.210 ± 0.039	-0.042 ± 0.039	0.060 ± 0.039	-0.084 ± 0.040
S_7	-0.038 ± 0.041	-0.041 ± 0.037	0.012 ± 0.038	0.011 ± 0.041	0.078 ± 0.040
S_8	-0.081 ± 0.043	-0.053 ± 0.041	0.086 ± 0.043	0.050 ± 0.040	0.010 ± 0.042
S_9	-0.006 ± 0.039	-0.003 ± 0.040	0.031 ± 0.038	-0.003 ± 0.042	-0.098 ± 0.041

widths					
parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L	1.012 ± 0.032	1.050 ± 0.033	0.926 ± 0.030	0.929 ± 0.030	0.952 ± 0.030
S_3	0.869 ± 0.027	0.893 ± 0.028	0.865 ± 0.028	0.837 ± 0.027	0.934 ± 0.030
S_4	0.980 ± 0.031	0.854 ± 0.027	0.801 ± 0.027	0.880 ± 0.029	0.858 ± 0.027
S_5	0.922 ± 0.029	0.805 ± 0.026	0.821 ± 0.028	0.866 ± 0.028	0.924 ± 0.029
A_{FB}	0.977 ± 0.031	0.864 ± 0.028	0.737 ± 0.028	0.800 ± 0.028	0.889 ± 0.028
S_7	0.904 ± 0.029	0.833 ± 0.026	0.833 ± 0.027	0.912 ± 0.029	0.898 ± 0.028
S_8	0.962 ± 0.031	0.913 ± 0.029	0.946 ± 0.031	0.891 ± 0.029	0.941 ± 0.030
S_9	0.880 ± 0.028	0.897 ± 0.028	0.833 ± 0.027	0.931 ± 0.030	0.914 ± 0.029

Tab. 9.3: The means and widths of the pull distributions for the angular moments in rare-like pseudoexperiments. 500 pseudoexperiments have been generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. The color scheme ranges from red (overestimation by 50% of the uncertainty) through green (ideal value) to blue (underestimation by 50% of the uncertainty), changing in steps of 10% for readers convenience.

It was discussed already in Sec. 7.3 that obtaining the F_S parameter using the fit to the reconstructed K^{*+} mass is not possible due to the low statistical power of the sample. Hence, a dedicated test is done: the pseudoevents are generated with the $F_S = 0.25$ and with non-zero interference terms, but in the modeling of the pseudoevents both F_S and the interference terms are fixed to zero. The background is treated the same way as in the previous case: in the pseudoevents the generated shape $\cos(\theta_K)$ follows a Chebyshev polynomial of order of five, but in the fit only order of two is allowed. The mean and width of the pulls obtained by studying 500 pseudoexperiments is listed in Tab. 9.5. The parameters S_3, S_4, S_7, S_8 and S_9 show none or little bias and their statistical uncertainty is estimated well. On the other hand, parameters F_L, S_5 and A_{FB} show rather large biases, especially in bins between and above the charmonium resonances. Looking at Tab. 9.2, those two bins also have the highest background fraction. In these bins, the parameters show a tendency to run into the given parameter limit despite their value at generation being very far from it.

The folding technique is applied also in this pseudoexperimental setup. The results are given in Tab. 9.6. The folding technique significantly reduces the bias in the A_{FB} parameter, the parameter F_L also can be reduced using the folding technique. Unfortu-

means

parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L (0)	0.183 ± 0.044	0.062 ± 0.046	0.147 ± 0.041	0.139 ± 0.044	0.252 ± 0.047
F_L (1)	0.201 ± 0.047	0.009 ± 0.048	-0.066 ± 0.044	-0.039 ± 0.043	0.216 ± 0.047
F_L (2)	0.302 ± 0.044	-0.011 ± 0.050	-0.029 ± 0.046	-0.161 ± 0.041	0.199 ± 0.047
F_L (3)	0.242 ± 0.046	0.091 ± 0.049	-0.087 ± 0.044	-0.023 ± 0.040	0.240 ± 0.046
F_L (4)	0.204 ± 0.049	0.364 ± 0.053	0.371 ± 0.057	0.452 ± 0.068	0.252 ± 0.047
S_3 (0)	0.043 ± 0.041	-0.056 ± 0.042	-0.029 ± 0.043	-0.105 ± 0.046	0.001 ± 0.041
S_3 (1)	-0.026 ± 0.045	0.006 ± 0.041	0.041 ± 0.035	-0.091 ± 0.041	0.048 ± 0.042
S_3 (2)	0.012 ± 0.043	-0.035 ± 0.039	-0.095 ± 0.041	-0.044 ± 0.038	-0.033 ± 0.043
S_3 (3)	-0.082 ± 0.044	-0.023 ± 0.040	-0.041 ± 0.040	0.013 ± 0.041	-0.031 ± 0.043
S_3 (4)	-0.060 ± 0.033	-0.008 ± 0.033	0.036 ± 0.033	0.048 ± 0.038	0.005 ± 0.033
S_4	-0.081 ± 0.045	-0.127 ± 0.041	0.049 ± 0.040	-0.040 ± 0.042	-0.080 ± 0.043
S_5	-0.017 ± 0.042	-0.082 ± 0.039	0.002 ± 0.038	-0.010 ± 0.039	-0.059 ± 0.040
A_{FB}	-0.098 ± 0.043	0.105 ± 0.042	0.037 ± 0.041	-0.047 ± 0.043	0.028 ± 0.046
S_7	0.016 ± 0.043	-0.057 ± 0.040	-0.049 ± 0.038	0.011 ± 0.041	0.073 ± 0.042
S_8	-0.019 ± 0.036	-0.018 ± 0.035	-0.046 ± 0.039	0.025 ± 0.041	-0.055 ± 0.036
S_9	0.025 ± 0.041	0.028 ± 0.039	-0.013 ± 0.042	0.016 ± 0.043	0.044 ± 0.041

widths

parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L (0)	0.985 ± 0.031	1.033 ± 0.033	0.902 ± 0.029	0.950 ± 0.031	1.040 ± 0.033
F_L (1)	1.047 ± 0.033	1.080 ± 0.034	0.968 ± 0.031	0.924 ± 0.030	1.056 ± 0.033
F_L (2)	0.989 ± 0.031	1.127 ± 0.036	0.996 ± 0.032	0.887 ± 0.029	1.040 ± 0.033
F_L (3)	1.019 ± 0.032	1.099 ± 0.035	0.957 ± 0.031	0.861 ± 0.028	1.027 ± 0.033
F_L (4)	1.072 ± 0.034	1.166 ± 0.037	1.160 ± 0.040	1.340 ± 0.049	1.057 ± 0.034
S_3 (0)	0.917 ± 0.029	0.931 ± 0.029	0.950 ± 0.030	1.018 ± 0.033	0.920 ± 0.029
S_3 (1)	0.996 ± 0.031	0.915 ± 0.029	0.781 ± 0.025	0.916 ± 0.029	0.933 ± 0.029
S_3 (2)	0.961 ± 0.030	0.864 ± 0.027	0.901 ± 0.029	0.840 ± 0.027	0.961 ± 0.030
S_3 (3)	0.992 ± 0.031	0.898 ± 0.028	0.877 ± 0.028	0.904 ± 0.029	0.962 ± 0.030
S_3 (4)	0.746 ± 0.024	0.739 ± 0.023	0.739 ± 0.024	0.848 ± 0.027	0.746 ± 0.024
S_4	1.013 ± 0.032	0.910 ± 0.029	0.870 ± 0.028	0.933 ± 0.030	0.956 ± 0.030
S_5	0.928 ± 0.029	0.872 ± 0.028	0.817 ± 0.027	0.871 ± 0.028	0.885 ± 0.028
A_{FB}	0.960 ± 0.031	0.941 ± 0.030	0.823 ± 0.029	0.921 ± 0.031	1.023 ± 0.032
S_7	0.954 ± 0.030	0.899 ± 0.028	0.846 ± 0.027	0.912 ± 0.029	0.931 ± 0.029
S_8	0.810 ± 0.026	0.773 ± 0.024	0.860 ± 0.027	0.905 ± 0.029	0.809 ± 0.026
S_9	0.916 ± 0.029	0.881 ± 0.028	0.930 ± 0.030	0.941 ± 0.030	0.922 ± 0.029

Tab. 9.4: The means and widths of the pull distributions of the angular moments in rare-like pseudoexperiments. 500 pseudoexperiments have been generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. In the fit to the pseudodata, folding is applied. The number at the parameters F_L and S_3 indicate the applied folding, as the two parameters can be measured using all five folding techniques. For the rest of the parameters, folding sensitive to the parameter is used. The color scheme ranges from red (overestimation by 50% of the uncertainty) through green (ideal value) to blue (underestimation by 50% of the uncertainty), changing in steps of 10% for readers convenience.

nately, the folding technique decreases the bias in the parameter S_5 only minimally and S_5 remains the only problematic parameter in the q^2 region below the J/ψ resonance. As the potential worst emerging bias from setting the S-wave contribution to zero is up to 35% of the statistical uncertainty, the S-wave parameters can be omitted in the fit to

means

parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L	0.061 ± 0.046	0.023 ± 0.045	0.421 ± 0.048	0.378 ± 0.046	-0.045 ± 0.045
S_3	0.046 ± 0.043	0.050 ± 0.041	-0.003 ± 0.045	0.094 ± 0.046	-0.012 ± 0.042
S_4	-0.016 ± 0.044	0.187 ± 0.042	0.169 ± 0.045	0.046 ± 0.046	-0.023 ± 0.045
S_5	-0.006 ± 0.044	0.404 ± 0.042	0.363 ± 0.046	0.325 ± 0.046	0.115 ± 0.039
A_{FB}	0.088 ± 0.043	-0.187 ± 0.043	-0.252 ± 0.048	-0.355 ± 0.044	-0.018 ± 0.045
S_7	0.062 ± 0.043	-0.033 ± 0.040	-0.067 ± 0.041	-0.030 ± 0.044	-0.046 ± 0.043
S_8	0.023 ± 0.045	0.065 ± 0.040	0.069 ± 0.045	0.052 ± 0.043	0.031 ± 0.043
S_9	-0.040 ± 0.042	-0.013 ± 0.039	-0.015 ± 0.042	-0.102 ± 0.043	0.014 ± 0.041

widths

parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L	1.027 ± 0.032	1.002 ± 0.032	1.064 ± 0.034	1.011 ± 0.032	1.009 ± 0.032
S_3	0.966 ± 0.031	0.916 ± 0.029	0.988 ± 0.032	1.020 ± 0.033	0.935 ± 0.030
S_4	0.970 ± 0.031	0.931 ± 0.030	0.965 ± 0.032	0.995 ± 0.033	1.010 ± 0.032
S_5	0.993 ± 0.031	0.927 ± 0.029	0.979 ± 0.033	1.001 ± 0.032	0.865 ± 0.027
A_{FB}	0.960 ± 0.031	0.958 ± 0.030	0.959 ± 0.034	0.925 ± 0.031	0.993 ± 0.032
S_7	0.960 ± 0.030	0.885 ± 0.028	0.907 ± 0.029	0.968 ± 0.031	0.951 ± 0.030
S_8	1.008 ± 0.032	0.894 ± 0.028	0.978 ± 0.032	0.950 ± 0.030	0.950 ± 0.030
S_9	0.928 ± 0.029	0.865 ± 0.027	0.933 ± 0.030	0.956 ± 0.031	0.913 ± 0.029

Tab. 9.5: The means and widths of the pull distributions in rare-like pseudoexperiments. 500 pseudoexperiments have been generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. The pseudoexperiments are generated with the $F_S = 0.25$ and with non-zero interference moments S_{S_i}. In the fit, the F_S and S_{S_i} moments are set to zero. The color scheme ranges from red (overestimation by 50% of the uncertainty) through green (ideal value) to blue (underestimation by 50% of the uncertainty), changing in steps of 10% for readers convenience.

data. It is worth noting that this assumes $F_S = 0.25$ for each pseudoexperiment, which is the maximal observed value in any bin in Ref. [4]: the actual value of F_S can be very well below this value.

Using the fits to the pseudoexperiments, an estimation of the statistical uncertainty in the fit to the data in the rare $B^+ \to K^{*+}\mu^+\mu^-$ channel is done. The statistical uncertainty is estimated by fitting the statistical uncertainty distribution for the given parameter using 500 pseudoexperiments with a Gaussian function. The expected statistical uncertainty is the mean of this Gaussian distribution. In the case of F_L and A_{FB}, the error distribution deviates from a pure Gaussian distribution, as the parameters run into the limit, and a right tail is present. As the deviation is not large, the Gaussian distribution is used to describe also the statistical uncertainty for these two parameters. Together with the very narrow pull distributions due to the complicated shape of the background in $\cos(\theta_K)$ in the sample with large statistical power (see Tab. 9.1), using the Feldman-Cousins technique [185] to ensure the correct coverage of the angular parameters will be necessary in the future measurements of this channel.

means

parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L (0)	-0.029 ± 0.047	-0.025 ± 0.047	0.218 ± 0.043	0.327 ± 0.044	-0.089 ± 0.047
F_L (1)	0.136 ± 0.049	0.074 ± 0.047	0.092 ± 0.047	0.214 ± 0.045	0.049 ± 0.043
F_L (2)	0.130 ± 0.049	-0.016 ± 0.051	0.179 ± 0.043	0.234 ± 0.047	0.014 ± 0.047
F_L (3)	0.064 ± 0.045	-0.082 ± 0.052	0.200 ± 0.042	0.157 ± 0.042	-0.040 ± 0.043
F_L (4)	0.080 ± 0.047	0.129 ± 0.055	0.367 ± 0.066	0.478 ± 0.067	-0.024 ± 0.046
S_3 (0)	0.012 ± 0.041	0.047 ± 0.042	0.094 ± 0.045	0.304 ± 0.046	0.006 ± 0.043
S_3 (1)	0.082 ± 0.043	0.043 ± 0.040	0.070 ± 0.038	0.211 ± 0.042	-0.037 ± 0.043
S_3 (2)	0.009 ± 0.044	0.011 ± 0.041	0.076 ± 0.039	0.244 ± 0.046	-0.039 ± 0.046
S_3 (3)	-0.034 ± 0.044	0.066 ± 0.039	0.054 ± 0.039	0.185 ± 0.042	0.024 ± 0.040
S_3 (4)	-0.048 ± 0.038	0.084 ± 0.035	0.089 ± 0.040	0.273 ± 0.044	0.021 ± 0.040
S_4	0.046 ± 0.044	0.190 ± 0.044	0.172 ± 0.041	0.043 ± 0.044	0.105 ± 0.043
S_5	0.022 ± 0.045	0.391 ± 0.046	0.344 ± 0.045	0.263 ± 0.050	0.181 ± 0.046
A_{FB}	0.025 ± 0.044	-0.093 ± 0.042	-0.335 ± 0.043	-0.396 ± 0.044	-0.038 ± 0.045
S_7	0.066 ± 0.042	0.043 ± 0.040	-0.032 ± 0.041	0.065 ± 0.044	0.054 ± 0.042
S_8	-0.022 ± 0.041	0.034 ± 0.038	-0.003 ± 0.045	0.045 ± 0.051	-0.030 ± 0.041
S_9	-0.049 ± 0.041	-0.035 ± 0.043	0.007 ± 0.041	-0.004 ± 0.047	-0.043 ± 0.040

widths

parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L (0)	1.036 ± 0.033	1.045 ± 0.033	0.943 ± 0.030	0.974 ± 0.031	1.036 ± 0.033
F_L (1)	1.092 ± 0.035	1.045 ± 0.033	1.042 ± 0.033	1.000 ± 0.032	0.962 ± 0.030
F_L (2)	1.027 ± 0.035	1.064 ± 0.036	0.895 ± 0.031	0.971 ± 0.033	0.988 ± 0.034
F_L (3)	0.996 ± 0.032	1.168 ± 0.037	0.938 ± 0.030	0.927 ± 0.029	0.971 ± 0.031
F_L (4)	1.024 ± 0.033	1.190 ± 0.039	1.340 ± 0.047	1.376 ± 0.048	0.997 ± 0.032
S_3 (0)	0.905 ± 0.029	0.928 ± 0.030	0.991 ± 0.032	1.003 ± 0.032	0.955 ± 0.031
S_3 (1)	0.971 ± 0.031	0.895 ± 0.028	0.851 ± 0.027	0.926 ± 0.029	0.961 ± 0.030
S_3 (2)	0.918 ± 0.031	0.858 ± 0.029	0.804 ± 0.027	0.945 ± 0.032	0.949 ± 0.032
S_3 (3)	0.976 ± 0.031	0.873 ± 0.028	0.867 ± 0.028	0.936 ± 0.030	0.904 ± 0.029
S_3 (4)	0.829 ± 0.027	0.762 ± 0.025	0.872 ± 0.028	0.964 ± 0.031	0.882 ± 0.028
S_4	0.992 ± 0.031	0.984 ± 0.031	0.913 ± 0.029	0.987 ± 0.031	0.960 ± 0.030
S_5	0.939 ± 0.032	0.950 ± 0.032	0.921 ± 0.032	1.041 ± 0.036	0.968 ± 0.033
A_{FB}	0.955 ± 0.031	0.923 ± 0.030	0.890 ± 0.031	0.945 ± 0.031	0.994 ± 0.032
S_7	0.932 ± 0.030	0.886 ± 0.028	0.904 ± 0.029	0.978 ± 0.031	0.938 ± 0.030
S_8	0.904 ± 0.029	0.828 ± 0.027	0.970 ± 0.032	1.101 ± 0.036	0.903 ± 0.029
S_9	0.896 ± 0.029	0.959 ± 0.031	0.902 ± 0.029	1.027 ± 0.033	0.879 ± 0.028

Tab. 9.6: The means and widths of the pull distributions in rare-like pseudoexperiments. 500 pseudoexperiments have been generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. In the fit to the pseudodata, folding is applied. The number at the parameters F_L and S_3 indicate the applied folding, as the two parameters can be measured using all folding techniques. The color scheme ranges from red (overestimation by 50% of the uncertainty) through green (ideal value) to blue (underestimation by 50% of the uncertainty), changing in steps of 10% for readers convenience.

The expected statistical uncertainties of the angular parameters F_L, S_3, S_4, S_5, A_{FB}, S_7, S_8 and S_9 are given in Tab. 9.7. In order to put these absolute uncertainties into perspective, the Standard Model value for each parameter is given together with the expected uncertainty. The Standard Model values are obtained using the FLAVIO package. The standard statistical uncertainty is obtained from fits to pseudoexperiments using the

folding technique. For the parameter F_L folding 1 is used, for the parameter S_3 folding 3 is used. This choice of the folding is made based on the results in Tab. 9.6: these foldings have the smallest bias. For the rest of the parameters, the folding sensitive to the parameter is used. The pull distributions are shown in App. D.1. Comparing the uncertainty to the $B^+ \to [K_S^0 \pi^+]_{K^{*+}} \mu^+ \mu^-$ measurement of the angular moments [4], the only two comparable q^2 intervals are $q^2 \in [1.1, 6.0]$ and $q^2 \in [15.0, 18.0]$. In the interval below the J/ψ resonance, the expected statistical uncertainties in this work are up to two times larger, in the high q^2 interval the uncertainty is up to three times larger compared to the one in the $B^+ \to [K_S^0 \pi^+]_{K^{*+}} \mu^+ \mu^-$ measurement. Note that the statistical uncertainty in Ref. [4] is estimated with the Feldman-Cousins technique, which returned larger statistical uncertainty values than HESSE.

Due to the large uncertainty of the parameter F_L and the potential bias of this parameter, the fit to the P'_i basis, defined in Eq. 3.6, has not been performed.

parameter	[0.25–4.00]	[4.00–8.00]	[11.00–12.50]	[15.00–18.00]	[1.10–6.00]
F_L	0.67 ± 0.15	0.66 ± 0.12	0.43 ± 0.16	0.34 ± 0.14	0.75 ± 0.13
S_3	0.00 ± 0.15	-0.03 ± 0.16	-0.09 ± 0.23	-0.19 ± 0.19	-0.01 ± 0.15
S_4	-0.03 ± 0.21	-0.24 ± 0.19	-0.28 ± 0.25	-0.30 ± 0.21	-0.15 ± 0.19
S_5	0.04 ± 0.19	-0.37 ± 0.19	-0.41 ± 0.26	-0.30 ± 0.19	-0.19 ± 0.19
A_{FB}	-0.09 ± 0.14	0.19 ± 0.12	0.39 ± 0.18	0.39 ± 0.16	0.01 ± 0.12
S_7	-0.02 ± 0.19	-0.01 ± 0.19	-0.00 ± 0.26	-0.00 ± 0.22	-0.02 ± 0.19
S_8	-0.01 ± 0.23	-0.00 ± 0.22	0.00 ± 0.29	0.00 ± 0.24	-0.01 ± 0.21
S_9	-0.00 ± 0.15	-0.00 ± 0.16	0.00 ± 0.23	0.00 ± 0.19	-0.00 ± 0.15

Tab. 9.7: The Standard Model values of the angular parameters with their expected standard statistical uncertainty from the fit to the data. The Standard Model values are obtained using the FLAVIO package [139]. The standard statistical uncertainty is obtained from fits to pseudoexperiments using the folding technique. For the parameter F_L folding 1 is used, for the parameter S_3 folding 3 is used. For the rest of the parameters, the folding sensitive to the parameter is used.

The statistical uncertainty estimation can be used to obtain the expected sensitivity of the measurement of the real value of the vector coupling $\mathrm{Re}(\mathcal{C}_9)$. A likelihood scan as a function of $\mathrm{Re}(\mathcal{C}_9)$ is performed. This is done using the FLAVIO package and shown in Fig. 9.6. A pseudomeasurement is generated using the predictions for a New Physics model with $\mathrm{Re}(\mathcal{C}_9)$=-2 as central value, which is approximately the value preferred by the fit in Ref. [4]. The q^2 bins used are in the range $[1.1, 6.0]\,\mathrm{GeV}^2$ and $[15.0, 18.0]\,\mathrm{GeV}^2$, as the predictions close to J/ψ are affected by the $c\bar{c}$ loops (see Sec. 1.5). One unit on the y-axis $[-2\Delta \log \mathrm{L}]$ corresponds to one standard deviations squared. Assuming the value of $\mathrm{Re}(\mathcal{C}_9)$=-2, the expected deviation from the Standard Model value is ≈ 2.4 standard

deviations. It is important to stress this estimation is done only using the expected *statistical* uncertainty on the parameters. Performing systematic studies especially related to the background shape can increase this uncertainty. Moreover, performing a Feldman-Cousin scan is necessary as the pull distributions show that the uncertainty estimation is volatile. The uncertainties obtained from the Feldman-Cousins scan can be larger than the ones presented here.
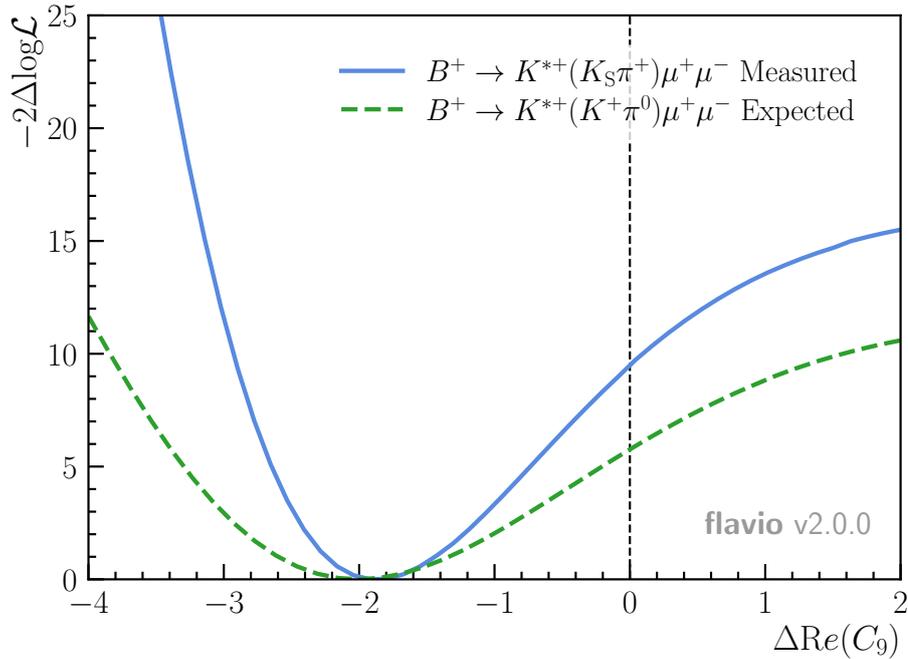
Fig. 9.6: Likelihood scan of the shift of $\mathrm{Re}(\mathcal{C}_9)$ from its SM value. The expected likelihood scan is represented by the dashed green line. The prediction is compared to the measured likelihood scan in the decay $B^+ \rightarrow \left(K^{*+} \rightarrow K_S^0 \pi^+\right) \mu^+ \mu^-$, given by the solid blue line. The preditions are taken from the FLAVIO package [139].

10   Tracking efficiency measurement

The measurement of track reconstruction efficiency of the LHCb detector and its recon-
struction algorithms efficiency is essential for most analyses carried out by the LHCb
collaboration. Contrary to the other large LHC experiments, the physics program of
LHCb focuses on relatively low momentum particles. Hence, it is optimized to contain as
little material as possible in order to reduce their scattering and hence be able to perform
very precise measurements. Therefore, any redundancy in the LHCb tracking system is
removed. The downside of this approach is sensitivity to the performance of tracking
detectors and the tracking algorithms. The track reconstruction efficiency is vital for the
physics program of LHCb.

10.1   Track reconstruction algorithms in LHCb

As shown in Sec. 2.2.1, the LHCb detector is designed with large gaps between the tracking
detectors. The particles can escape or be created at any point in the active detector
area, leaving hits only in some of the detectors. Hence there are many possible track
types to be reconstructed. A sketch of the most used types at the LHCb experiment is
given in Fig. 10.1. The tracks are reconstructed using algorithms that correspond to the
different track types. These algorithms are independent and therefore a particle crossing
the detector is typically reconstructed by several of these algorithms: for an example, a
particle bent out of the LHCb acceptance by the magnet can be reconstructed by the
upstream tracking algorithm as well as the VELO tracking algorithm.

(a)
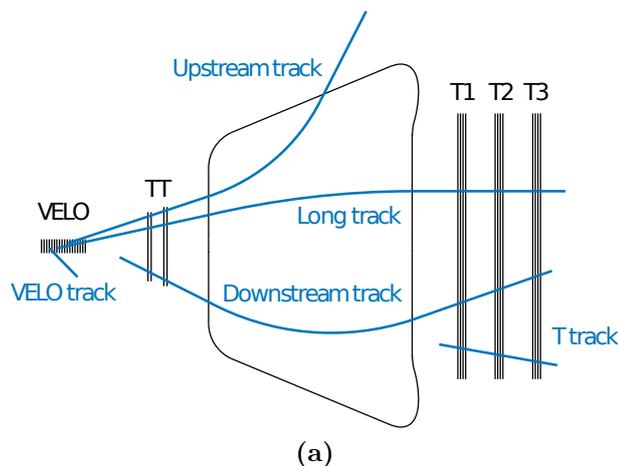
Fig. 10.1: The most common track types used at the LHCb experiment. The most valuable
track type to the LHCb physics analyses are *long tracks* as they have the best momentum
resolution. Long tracks have hits in all LHCb tracking detectors.

- **Long track:** A track with origin in the VELO that also transverses the T stations. Hit information from the TT *can* be added, but it is not required. This type of track is the most common track type in LHCb studies as it has the best momentum resolution. There are two independent algorithms used to obtain the long track: *forward tracking* and *matching*. The *forward tracking* algorithm propagates VELO track's trajectory to the T stations, taking into account the bending of the trajectory by the magnet. In the *matching* algorithm standalone T tracks are created and combined with VELO tracks, also taking the bending into account. The two algorithm results can be compared and a combined set of best long tracks is obtained. TT hits are added only after finding a track candidate from the VELO and T station hits.

- **Velo track:** A track that only consists of hits in the VELO detector: they are independent of the forward tracking. These tracks are used for primary vertices reconstruction.

- **Upstream track:** A track reconstructed using VELO and TT hits. As these tracks have only poor momentum information, they are rarely used in analyses. If no other algorithm reconstructs this track, it corresponds to a particle with low enough momentum so it is bent out of the LHCb acceptance by the magnet.

- **Downstream track:** A track reconstructed using TT and T station hits. As there is no VELO information, the momentum resolution is worse compared to long tracks. Long-lived particle decay products leave downstream tracks in the detector.

- **T track:** A track reconstructed only using hits in T stations. Similarly to VELO tracks, there is no momentum information. If no other algorithm reconstructs this track, it typically represents a very long-lived particle decay product.

The algorithms searching for hits from different detectors and combining them can sometimes combine hits that do not originate from the same particle. Such tracks, called *ghost* tracks, contribute to the background. Most of the ghost tracks can be rejected by requiring a high track fit quality χ^2. However, this can also lead to rejection of real particle tracks, modifying the track reconstruction efficiency. To resolve this issue, a dedicated neural networks is trained in the LHCb reconstruction software. This neural network is designed to remove most ghost tracks while minimizing the impact on real tracks. It returns a value between 0 and 1, a "ghost probability", which is typically required to be below 0.4, corresponding to removal of more than 70% of reconstructed ghost tracks with hardly any loss in efficiency.

10.2   Determination of the track reconstruction efficiency

In most analyses carried out by the LHCb collaboration, the track reconstruction efficiency is estimated using a Monte Carlo simulation. While the simulation is a very good representation of the real data, it is not perfect. The main discrepancy between the real data and simulation is the detector occupancy distributions (see Sec. 5.4), but there can be other imperfections in the kinematic variables as well. The track reconstruction efficiency depends mainly on the kinematic properties of the track (momentum, direction, and position in the detector) as well as on the occupancy of the detector. While the discrepancies between the data and the simulation in the kinematic and occupancy quantities can be corrected for, the track reconstruction efficiency also depends on the placement and the amount of dead channels, inactive materials and others. These effects are very hard to simulate or correct for in the Monte Carlo simulation. Hence, the track reconstruction efficiency obtained using purely a simulation sample is a good approximation, but it does not meet the required precision. Measurement using a data-driven technique is necessary.

The track reconstruction efficiency measurement is executed using a data-driven tag-and-probe method exploiting the decay of $J/\psi \to \mu^+\mu^-$. Tag-and-probe technique is widely used in high energy physics [186, 187, 188] to measure the efficiency of various processes, typically reconstruction or selection. The method exploits two-product decays of a well-know resonance. One of the decay products, the *tag*, is a well identified track, while the other, the *probe*, is an unbiased track. The *probe* track then either passes or fails the reconstruction or selection criteria for which the efficiency is to be measured. The ratio of track passing this criteria to all reconstructed unbiased tracks is the reconstruction or selection efficiency ε:

$$\varepsilon = \frac{N_{\text{passing criteria}}}{N_{\text{all unbiased}}} . \tag{10.1}$$

The method used at the LHCb experiment was developed during Run I and further advanced during Run II [189]. It exploits the decay of J/ψ mesons to a muon pair. Recently, a new technique exploiting the decay of $J/\psi \to e^- e^+$ has been developed [190]. However, the focus in this work is on the $J/\psi \to \mu^+\mu^-$ decay only. In the following section, the tag-and-probe method used to measure the track reconstruction efficiency in LHCb is explained in detail.

10.2.1   Tag-and-probe technique using $J/\psi \to \mu^+\mu^-$ decays

The muon decay of J/ψ meson is used for track reconstruction efficiency determination muons transverse the whole LHCb detector region and leave hits also in the muon stations. Moreover, they do not interact hadronically.

The *tag* track is a muon track reconstructed using standard long track reconstruction algorithm and passing a tight selection to make sure it is a decay product of a *J/ψ* resonance, such as a momentum requirement. The *probe* track is reconstructed using one of three dedicated algorithms designed in such a way that each probes one (or two) of the three tracking detectors of the LHCb: the VELO, the T stations and the TT. These algorithms are very loose in order to minimize the potential bias imposed by any selection on the final result.

The criteria to determine whether the *probe* track is efficiently reconstructed or not is the existence of a long track that can be associated with the *probe* track. The matching is performed by checking the amount of common hits between the *probe* and the long track in the tracking detectors. The *overlap fraction* is used as the association criterion. The overlap fraction is the the number of common hits N_{common} divided by the minimum number of hits in the subdetector required by the long track reconstruction algorithm N_{required}:

$$\text{overlap fraction} = \frac{N_{\text{common}}}{N_{\text{required}}} \,, \tag{10.2}$$

Using the overlap fraction as the association criterion, the Eq. 10.1 then becomes

$$\varepsilon_{\text{tr}} = \frac{N_{\text{assoc}}(J/\psi \rightarrow \mu^+\mu^-)}{N_{\text{all}}(J/\psi \rightarrow \mu^+\mu^-)} \,, \tag{10.3}$$

where N_{assoc} denotes the number of *probe* tracks passing the association criteria and N_{all} denotes all unbiased *probe* tracks.

Depending on the algorithm used to reconstruct the *probe* track, there are three methods to obtain the track reconstruction efficiency. The methods, illustrated in Fig. 10.2, are:

- **Long method** The *probe* track is reconstructed at first using muon station hits to create a standalone muon track. This track is then matched to the hits in the TT. Note that the long track reconstruction algorithm described in Sec. 10.1 does require hits neither in the muon stations nor the TT. Hence, this method directly probes the track reconstruction efficiency of long tracks. The *probe* track is considered efficient, when the overlap fraction is at least 70% for the muon stations and 60% for the TT.

- **VELO method** The *probe* track is reconstructed as a downstream track with added muon station hits in order to identify the particle as a muon. This method probes the VELO track reconstruction efficiency. The *probe* track is considered efficient when the overlap fraction in the T stations is at least 50%.

- **T station method** The *probe* track is reconstructed by a dedicated algorithm from hits in the VELO and the muon stations. This method probes the T station track reconstruction efficiency. The *probe* track is considered efficient when there are at least two common hits in the muon stations and the same VELO segment as the *probe* signature.

Different methods probe different track reconstruction algorithms. When combining the VELO and T station efficiency, a careful reader may notice this is equivalent to probing the long track reconstruction efficiency. This is further exploited by two methods:

- **Combined method** This method represents the combination of VELO and T station efficiencies. The efficiencies from these two methods are uncorrelated, with the exception of track kinematics and detector occupancy. The combined method efficiency is then simply

$$\varepsilon_{\mathrm{Comb}} = \varepsilon_{\mathrm{VELO}} \times \varepsilon_{\mathrm{Tstation}} \ . \tag{10.4}$$

- **Final method** The Final method is the weighted average of Long and Combined methods. The weights are inverse squares of the uncertainty for each method, $w_{\mathrm{Comb}} = 1/\sigma_{\mathrm{Comb}}^2$, $w_{\mathrm{Long}} = 1/\sigma_{\mathrm{Long}}^2$. The weighted average is then

$$\varepsilon_{\mathrm{Final}} = \frac{w_{\mathrm{Comb}}\varepsilon_{\mathrm{Comb}} + w_{\mathrm{Long}}\varepsilon_{\mathrm{Long}}}{w_{\mathrm{Comb}} + w_{\mathrm{Long}}} \ . \tag{10.5}$$

The statistical uncertainty of the average is

$$\sigma_{\mathrm{Final}} = \frac{1}{\sqrt{w_{\mathrm{Comb}} + w_{\mathrm{Long}}}} \ . \tag{10.6}$$

This method represents the most precise track reconstruction efficiency at LHCb, as it exploits the information from all three available methods.

It is worth noting here that using muon tracks as the *tag* and *probe* does not take into account the material absorption effects on the track reconstruction efficiency. The track reconstruction efficiency presented here represents the probability of a particle crossing the full active detector area to be reconstructed. However, the hadronic interactions can be taken into account in the form of systematic uncertainty. The uncertainty is equal the fraction of hadrons that cannot be reconstructed due to hadronnic interactions multiplied by the material budget uncertainty. As the cross-section of the hadronic interaction depends on the given particle, this has to be evaluated separately for each hadron. Moreover, it also depends on the momentum of the hadron: the fraction of hadrons that cannot be

(a)

(b)

(c)
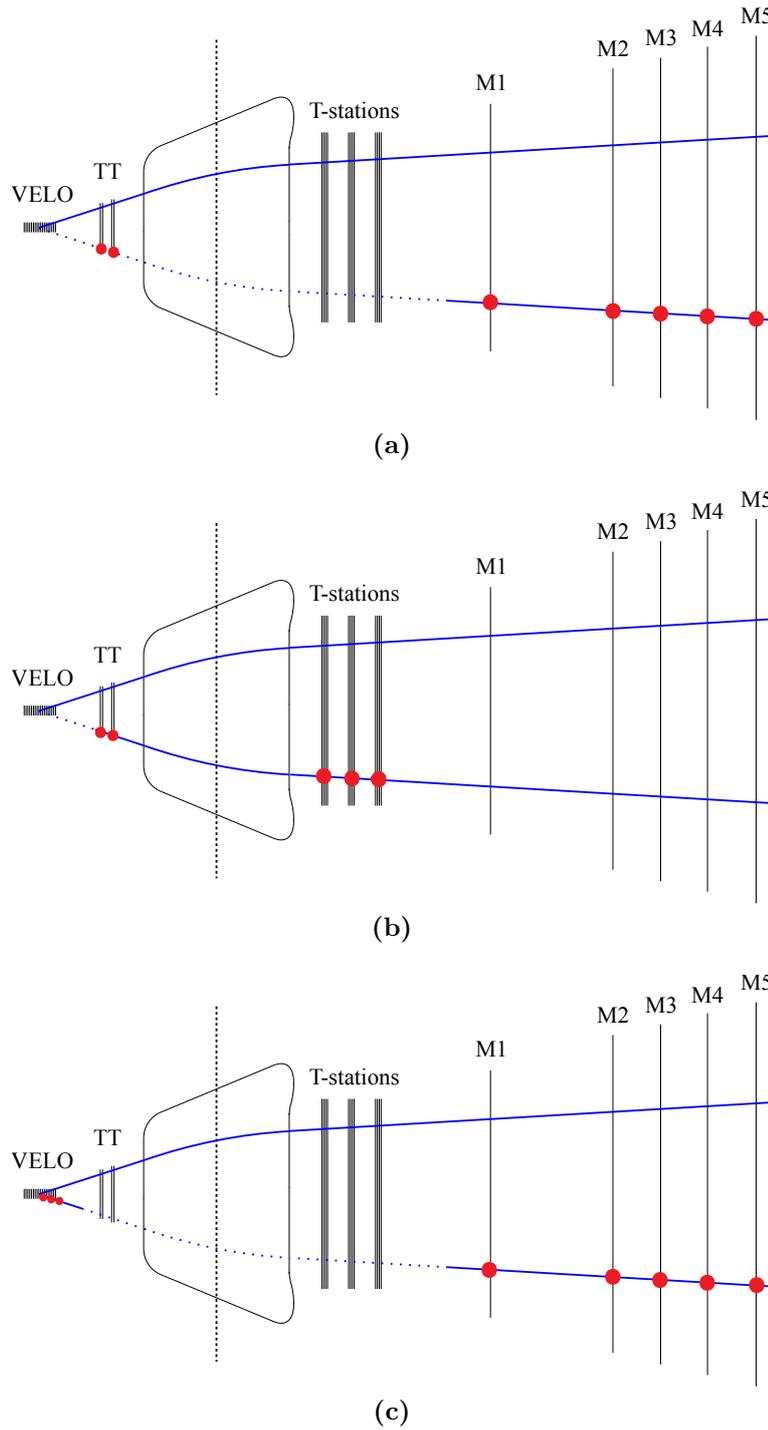
Fig. 10.2: Illustration of the probe track reconstruction algorithms: (a) Long method, (b) VELO method, and (c) T station method [189]. Red dots indicate the hit information used by each algorithm to select the *probe* track. Solid blue line represent the trajectory of the *tag* (upper line) and *probe* (lower line) tracks. The blue dotted line represents the sub-detector region which is probed by the respective method. The dashed vertical line shows the bending plane of the magnet.

reconstructed but to the hadronic interactions can be estimated for each process using the simulation.

10.2.2   Efficiency evaluation

When selecting the J/ψ candidates, there is a contribution from random combinations of real or even fake muon tracks. The number of signal $J/\psi \to \mu^+\mu^-$ events N_{all}, as well as the number of associated events N_{assoc} in Eq. 10.3 needs to be extracted from a mass fit to the J/ψ candidate. The J/ψ candidate's mass is calculated as

$$m_{\mathrm{rec}} = \sqrt{(E_{\mathrm{tag}} + E_{\mathrm{probe}})^2 - (\vec{p}_{\mathrm{tag}} + \vec{p}_{\mathrm{probe}})^2}, \qquad (10.7)$$

where E is energy of the *tag* or *probe* track and $\vec{p}$ is its momentum.

The reconstructed J/ψ candidates are split into two sets. *Matched* candidates fulfill the association criteria, *failed* candidates do not fulfill the association criteria. These two sets are simultaneously fitted with a sum of two Crystal Ball functions with shared mean value as the signal component and an exponential function as the background component. The two sets share all signal shape parameters. The yields and the background shape in the *matched* and *failed* sets are independent. Following Eq. 10.3, the track reconstruction efficiency can be expressed as a function of the signal yields of the *matched*, $N_{\mathrm{sig}}^{\mathrm{Match}}$, and *failed*, $N_{\mathrm{sig}}^{\mathrm{Fail}}$, samples

$$\varepsilon_{\mathrm{tr}} = \frac{N_{\mathrm{sig}}^{\mathrm{Match}}}{N_{\mathrm{sig}}^{\mathrm{Match}} + N_{\mathrm{sig}}^{\mathrm{Fail}}}. \qquad (10.8)$$

It is important to consider only the *signal* yields, as combinatorial background is present in the measurement. An example of the mass distributions for the matched and failed candidates is given in Fig. 10.3.

In this approach, the efficiency is treated as a fit parameter[31]. This scheme guarantees that the correlations between the parameters are properly treated in the calculation of the statistical uncertainties.

As discussed in Sec. 10.2, the efficiencies obtained from the simulation are a good approximation of the actual efficiencies, however they are not perfect. On the other hand, the tag-and-probe method of obtaining the efficiencies is rather lengthy and performing this study for every analysis is not feasible. Therefore, a correction factor R is calculated:

$$R = \frac{\varepsilon_{\mathrm{data}}}{\varepsilon_{\mathrm{sim}}}, \qquad (10.9)$$

---

[31]In the fit to the J/ψ mass, the floated parameters are then the track reconstruction efficiency and the total yield of $N_{\mathrm{sig}} < all = N_{\mathrm{sig}}^{\mathrm{Match}} + N_{\mathrm{sig}}^{\mathrm{Fail}}$.

Fig. 10.3: Invariant reconstructed J/ψ candidate mass distribution for the Long method. On the left, the invariant mass of the matched candidates is shown, on the right are failed candidates. The black points represent a subset of the data obtained during the data taking year 2018. The black line represents the full fit model. The red dashed line represents the signal component.

where the $\varepsilon_{\text{data}}$ represents the tag-and-probe efficiency in data and ε_{sim} represents the tag-and-probe efficiency in simulation. This ratio R can be used independently by many analyses to 'correct' the efficiency obtained directly from simulation. Moreover, first order uncertainties are canceled out. Therefore, the ratio R is the ultimate goal of the tracking efficiency measurement.

In order to accomodate the requirements of most LHCb analyses, the track reconstruction efficiency is measured in bins of pseudorapidity, η, momentum, p, number of hits in the SPD detector, $N_{SPD\text{hits}}$, and the number of primary vertices present in the event N_{PV}. The ratio R is measured in two dimensions in bins of pseudorapidity and momentum. This is referred to later as *correction table*.

10.3   Tracking efficiency measurement in Run II

The track reconstruction has been already measured (and published [189]) in Run I. While the efficiency measurement is a crucial part of any Run I analysis, the precision was limited by the available data size: any selected event was required to pass an unbiased single muon software trigger. This trigger could perform only a simple track reconstruction due to time constrains. On top of this, the alignment and calibration was performed using the recorded data and the trigger bandwidth had to match the stripping bandwidth.

This procedure was fundamentally changed for Run II. The real-time alignment and calibration together with the full reconstruction available in the high-level trigger allowed

Variable	VELO method	T station method	Long method
$\mathrm{DLL}_{\mu\pi}$	> -2	> -1	> -2
p	$> 5\,\mathrm{GeV}$	$> 7\,\mathrm{GeV}$	$> 10\,\mathrm{GeV}$
p_{T}	$> 700\,\mathrm{MeV}$	$-$	$> 1300\,\mathrm{MeV}$
track χ^2/ndf	< 10	< 3	< 5
IP	$> 0.5\,\mathrm{mm}$	$> 0.2\,\mathrm{mm}$	$-$

Tab. 10.1: Selection cuts applied to the ***tag*** tracks by the software trigger.

for the full track reconstruction in real time[32]. This allows for fast evaluation of data directly after it has been recorded. Hence, dedicated trigger software has been created in order to perform the tag-and-probe efficiency evaluation at the HLT level.

10.3.1   Trigger lines implementation

As explained in Sec. 10.2.1, a *J/ψ* meson reconstructed from two muon tracks has to be found. Therefore, the first stage of the trigger selection is searching for good-quality muon tracks. Such track fulfills the requirements of the `Hlt1TrackMuon` trigger line: the track has to have $p_{\mathrm{T}} > 800\,\mathrm{MeV}$, fulfill the `IsMuon` requirement (see Sec. 5.2) and has to have the impact parameter significance larger than eight. The last requirement ensures the track is detached from the primary vertex to reduce the background pollution.

In order to speed up the computation process, first a *tag* track has to be find. Such track has to fulfill also additional criteria listed for each method in Tab. 10.1. The tag track is reconstructed using the standard LHCb tracking. Only when such a track is found, the dedicated track reconstruction is performed to search for *probe* tracks, accelerating the computation process significantly. Also the *probe* tracks have to fulfill additional loose requirements, see Tab. 10.2. Average decision time for each of the trigger decisions is below 1 ms, satisfying the software trigger requirements. All the requirements are optimized to have the largest possible kinematic coverage while keeping the coverage identical for the three methods. There are six trigger lines implemented: two lines per method. As the trigger has to distinguish between the *tag* and *probe* tracks, the charge of the muons is exploited. One line reconstructs the *J/ψ* candidates using positively charged *tag* muon track and negatively charged *probe* track, the other one uses the tracks with opposite charges.

Lastly, criteria listed in Tab. 10.3 are applied on the *J/ψ* meson. These conditions are optimized in order to reduce the combinatorial background and to make sure the two-

---

[32] *"Real time is defined as the interval between the collision in the detector and the moment the data are sent to permanent storage."* [191]

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

Tab. 10.2: Selection cuts applied to the ***probe*** tracks by the software trigger.

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

Tab. 10.3: Selection cuts applied to the **J/ψ** reconstructed from tag and probe tracks by the software trigger.

muon vertex has a good quality. Moreover, in the case of the VELO method, the distance of closest approach (DOCA) condition is added in order to speed up the computation process.

In order to measure the track reconstruction efficiency, the overlap fraction also has to be saved. Therefore, another trigger line is added: the selection is identical to the selection described above plus a requirement of finding a long track associated to the probe track with an overlap fraction fulfilling the criteria for each method is added. The existence of two lines allows for online tracking efficiency calculation.

10.3.2   Trigger lines online monitoring

The full tracking efficiency estimation in real time also allows for real time monitoring of the track reconstruction efficiency. A dedicated online monitoring tool has been implemented in 2017. For the monitoring purposes, the output of the trigger lines described in the previous section is saved in a form of three one-dimensional histograms in mass, momentum and pseudorapidity distributions. This is saved for both *probe* and *tag* muon tracks as well as the *J/ψ* candidates.

In order to estimate the online efficiency, a fit to the *J/ψ* reconstructed mass is performed. The fit consists of Gaussian distribution for signal and exponential function for background. This fit is performed for every *run* that lasted at least 45 minutes in order to have sufficiently large data sample available. A run is a set of data taken during the same detector settings that lasted maximum of one hour. The yields of these fits can be used to estimate the tracking efficiency for each method.

Variable	VELO method	T station method	Long method
Tag selection criteria			
$\mathrm{DLL}_{\mu\pi}$	> -2	> -1	> -2
p	$> 5\,\mathrm{GeV}$	$> 7\,\mathrm{GeV}$	$> 10\,\mathrm{GeV}$
p_T	$> 0.7\,\mathrm{GeV}$	$> 0.0\,\mathrm{GeV}$	$> 1.3\,\mathrm{GeV}$
track χ^2/ndf	< 10	< 3	< 5
IP	$> 0.5\,\mathrm{mm}$	$> 0.2\,\mathrm{mm}$	$--$
Probe selection criteria			
$p > 5\,\mathrm{GeV}$	$> 5\,\mathrm{GeV}$	$> 5\,\mathrm{GeV}$	
p_T	$> 0.5\,\mathrm{GeV}$	$> 0.5\,\mathrm{GeV}$	$> 0.5\,\mathrm{GeV}$
track χ^2/ndf	< 10	< 5	$--$
J/ψ candidates selection criteria			
$|m_{\mu^+\mu^-} - m_{J/\psi}|^{\mathrm{precomb}}$	$< 2000\,\mathrm{MeV}$	$< 1000\,\mathrm{MeV}$	$< 1000\,\mathrm{MeV}$
$|m_{\mu^+\mu^-} - m_{J/\psi}|^{\mathrm{postcomb}}$	$< 200\,\mathrm{MeV}$	$< 500\,\mathrm{MeV}$	$< 500\,\mathrm{MeV}$
p_T	$--$	$> 0.5\,\mathrm{GeV}$	$> 1\,\mathrm{GeV}$
vertex χ^2	< 5	< 2	< 2
Track DOCA	$< 5\,\mathrm{mm}$	$--$	$--$
IP	$-$	$--$	$< 0.8\,\mathrm{mm}$

Tab. 10.4: Selection cuts applied to the *tag* track, *probe* track and the reconstructed J/ψ candidate by the stripping selection. `Precomb` and `postcomb` denote cuts applied before and after the vertex fit respectively.

10.3.3   Stripping lines implementation

While the full reconstruction allows for faster and more efficient determination of the track reconstruction efficiencies, when a trigger line fails, there is no data available for measuring the track reconstruction. Therefore, dedicated stripping lines for each method have been implemented also for Run II. This has been proven to be useful in 2017, when a part of the trigger line for the T station method was overwritten by an output from a different trigger line. With available stripping lines, the T station method has been successfully recovered. The price to pay is the same as in Run I: smaller amount of J/ψ candidates available and longer processing times. However, the datasample taken in 2017 was large enough to fully recover the T station method. The measured track reconstruction efficiency measurements via the recovered T station method is given in Fig. 10.4.

The stripping lines perform very similar calculations as the trigger lines, however, the workflow does not allow to search for the *tag* track first, resulting in longer computation times. This is slightly improved by imposing a mass requirement on the combination of the muon candidates before the vertex fit. The cuts used in the stripping lines are identical to the cuts applied in the trigger selection and they are listed in Tab. 10.4.

Moreover, the selection criteria applied in the trigger selection might not match re-
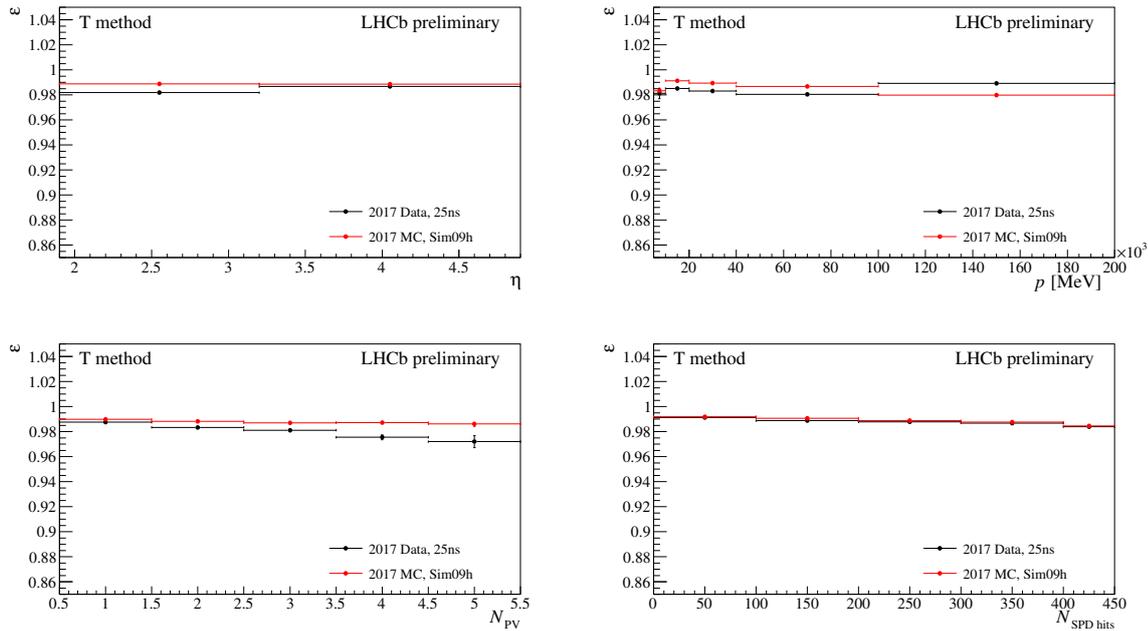
Fig. 10.4: Track reconstruction efficiency using the T method in pseudorapidity η (top left), in momentum p (top right), in the number of primary vertices N_{PV} (bottom left) and in number of hits in the SPD $N_{SPDhits}$ (bottom right) for the 2017 data-taking period (25ns bunch spacing). This sample is obtained using stripping in order to recover faulty trigger selection of the T method in 2017. Sim09h denotes the used version of the simulation software used by the LHCb experiment. The uncertainties are statistical only.

quirements of analyses exploring the edges of the available phase-space. An example is a dedicated set of stripping lines allowing to study the track reconstruction efficiency in events with minimal detector occupancy. This has been added in 2017. This analysis is beyond the scope of this work, for the details, see Ref. [192].

10.4 TrackCalib package

The ultimate goal of the track reconstruction efficiency measurement is to measure the ratio R defined in Eq. 10.9. The ratio obtained directly from the trigger selection can be used by a wide range of LHCb analyses to correct the track reconstruction efficiencies obtained from the simulation. However, many analyses require a dedicated approach: tighter selection than the one presented here, different detector occupancy measure to weight their simulation sample, special binning in momentum p and pseudorapidity η or estimating the track reconstruction efficiency or the ratio R in some other variable. In order to simplify the dedicated measurement of the track reconstruction efficiency tailored to the needs of any analysis, a dedicated tool `TrackCalib` has been created and made available to the collaboration in 2017. This PYTHON tool allows to evaluate the track

reconstruction efficiency using command line options. In Run I, such customization has not been possible. However, exploiting the stripping lines described in Sec. 10.3.3, these options have been recently extended also to the Run I sample. The tool documentation is available online at [193].

The `TrackCalib` package works in three main steps: data preparation, the fit of the data and the plotting of the efficiencies. It is possible to run each step or run all three steps together. The user can decide

- what method is used

- which simulation version is used

- whether data, simulation or both are used

- whether only one magnet polarity or both are used

for the track reconstruction efficiencies or R evaluation. Moreover, the user can decide in dependence on what variables should the efficiency or the ratio R be evaluated. In each of the three steps, additional options can be set.

10.4.1   Data preparation

In the first step of `TrackCalib` tool, the dataset used for the tracking efficiency calculation is selected. The full datasample obtained from the trigger lines is rather large (especially in real data) not just due to the large amount of saved events, but also because of many variables being saved. Therefore, a smaller datafile consisting only of relevant variables is created. Additional selection criteria, typically ghost track probability cut, can be set by the user. This criteria can be applied only on the *probe* track, *tag* track or both. The default requirements used by the `TrackCalib` are listed in Tab. 10.5.

Moreover, the required overlap fraction needed to associate the tracks can be modified. The variable used for the weighting of the simulation sample can be chosen. Lastly, maximum number of event candidates per method and charge settings[33] can be used. The selected dataset is locally stored in order to be readily available for the next two steps.

---

[33]Simirarly to what is done in the trigger selection, the `TrackCalib` tool either uses a μ^+ track as a *probe* and μ^- as a *tag* or the opposite charge configuration.

Variable	VELO method	T station method	Long method
Tag selection criteria			
$\text{DLL}_{\mu\pi}$	$--$	$--$	$--$
p	$> 5\,\text{GeV}/c$	$> 7\,\text{GeV}/c$	$> 10\,\text{GeV}/c$
p_{T}	$> 0.7\,\text{GeV}/c$	$> 0.5\,\text{GeV}/c$	$> 1.3\,\text{GeV}/c$
track χ^2/ndf	< 5	< 5	< 2
IP	$--$	$> 0.2\,\text{mm}$	$--$
Probe selection criteria			
p	$> 5\,\text{GeV}/c$	$> 5\,\text{GeV}/c$	$> 5\,\text{GeV}/c$
p_{T}	$> 0.7\,\text{GeV}/c$	$> 0.1\,\text{GeV}/c$	$> 0.1\,\text{GeV}/c$
track χ^2/ndf	$--$	$--$	$-$
J/ψ candidates selection criteria			
$|m_{\mu^+\mu^-} - m_{J/\psi}|$	$< 200\,\text{MeV}/c^2$	$< 500\,\text{MeV}/c^2$	$< 500\,\text{MeV}/c^2$
p_{T}	$--$	$--$	$> 0\,\text{GeV}/c$
vertex χ^2	< 5	< 5	< 5
Track DOCA	$--$	$--$	$--$
IP	$-$	$--$	$< 0.8\,\text{mm}$

Tab. 10.5: Selection cuts applied applied to the *tag* track, *probe* track and the reconstructed J/ψ candidate by the default `TrackCalib` selection.

10.4.2   Fit execution

In this part of the `TrackCalib` tool, the previously created dataset is divided based on the selected variables. The user can decide what binning in the desired variables is used[34]. In the case of low statistics sample, instead of performing a simultaneous fit to the matched and failed J/ψ candidates, as explained in Sec. 10.2.1, a fit to the matched and *all* J/ψ candidates is performed. Due to very high track reconstruction efficiency, the failed sample has very little signal component (see Fig. 10.3): by avoiding the fit to the failed sample, the fit stability improves. To further improve the fit stability, the Crystal Ball function used to describe the signal component can be replaced by a sum of two Gaussian distributions. Lastly, the user can also execute an ubinned fit to the J/ψ mass.

For each method and each variable bin, a dedicated file containing the calculated efficiency as well as the fitted distribution is saved. Moreover, an output file is created, where the fit status and the fitted parameter values are saved. For user's convenience, another warning file is created, where only failed fit statuses and variables with large or zero uncertainty are saved. This allows for quick recongnition of failed fits.

---

[34]This can be done both by requiring a certain number of bins with the same width or by specifying the bin edges.

year	Simulation versions
2015 (EM)	Sim08h, Sim09b
2015	Sim09a, Sim09b
2016	Sim09a, Sim09b, Sim09d
2016 (strip)	Sim09h
2017	Sim09h
2017 (strip)	Sim09h
2018	Sim09h
2018 (strip)	Sim09h

Tab. 10.6: Available simulation samples for the track reconstruction efficiency measurements. In the first months of data taking in 2015, the settings of the machine were different to the rest of the year. This sample is denoted *early measurements* (EM). Samples produced using the stripping selection instead of the trigger selection are denoted as (strip).

10.4.3   Plotting

Last part of the `TrackCalib` tool is the plotting of the track reconstruction efficiency dependency plots and the creation of correction tables: the ROOT files (the correction tables), where the ratio R is saved in bins of the desired variables (the default is pseudorapidity and momentum). In this step, the three methods are also combined into the Combined and Final methods.

10.4.4   Simulation samples

In order to apply the tag-and-probe method on the simulated sample, several sets of the decay $B^+ \to J/\psi(\to \mu^+\mu^-)X$, where X is any particle, are created. These simulated samples are then treated the same way as the record data. In order to correct the difference between the simulation and the data in the occupancy, weights based on the number of hits in the SPD are applied.

The LHCb software is constantly evolving and improving. Therefore, it is important to evaluate the track reconstruction efficiency correctly for each version of the software. The available simulation samples used in `TrackCalib` for each data taking year are listed in Tab. 10.6.

10.5   Systematical uncertainties

The sources of potential systematic uncertainties have been investigated for the Run I track reconstruction efficiency measurement [189]. The method of measuring the track reconstruction efficiencies remained unchanged, hence the uncertainties are not expected to significantly change in the Run II measurement.

Changing the J/ψ mass signal model from the sum of two Crysta-ball distributions to the sum of two Gaussian distributions does not change the efficiency significantly compared to the statistical uncertainty. Similarly, changing the background model from an exponential distribution to a linear one leads only to a negligible change in the track reconstruction efficiency.

Another source of the systematical uncertainty could be the difference between the long method and the combined method. However, the difference is observed to be small relative to the statistical uncertainty and is further reduced in the ratio of the track reconstruction efficiency in data to the efficiency in the simulation.

The dominant systematical uncertainty in the Run I measurement originates from the choice of the occupancy variable used to improve the agreement of the simulated event sample with the real data. The uncertainty is evaluated by using the number of hits in the SPD, the number of long tracks in the event and the number of primary interaction vertices as the occupancy variables. The largest deviation observed in Run II for the correction factors obtained from the combination of all methods in any of the two-dimensional correction tables is 0.8%.

10.6 Results

The tracking efficiencies for the 2018 data-taking period obtained form data and from the simulation version Sim09h depending on pseudorapidity η, momentum p, number of primary vertices N_{PV} and the number of hits in the SPD detector $N_{SPDhits}$ are shown in Figs. 10.5, 10.6, 10.7 and 10.8. All the shown track reconstruction efficiencies are 94% or higher. The agreement between the track reconstruction efficiency obtained form the data sample and the simulation is very good. The two-dimensional correction table in momentum p and pseudorapidity η is given in Fig. 10.9. The ratio is very close to one in all bins. The two-dimensional ratio for the Final method is made available as a correction table for physics analyses performed by the LHCb collaboration.

The results for all the data-taking years 2011-2018 can be found at Ref. [193]. As all the results are consistent throughout the years, they are not shown in this work. Alternatively, more results are published in Ref. [194].
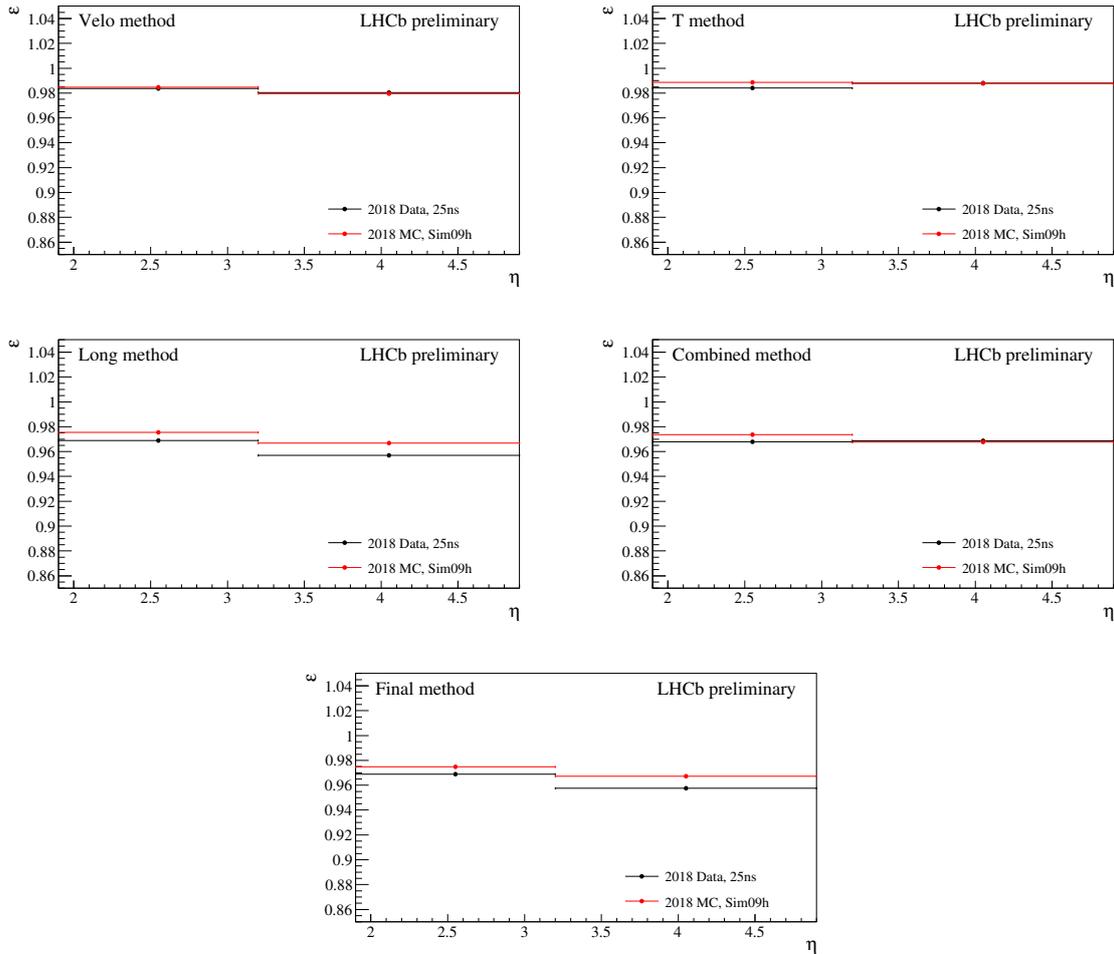
Fig. 10.5: Track reconstruction efficiency versus pseudorapidity η for the 2018 data-taking period (25ns bunch spacing). Sim09h denotes the used version of the simulation software used by the LHCb experiment. The track reconstruction efficiency is obtained using the VELO method (top left), the T method (top right), the Long method (middle left), the Combined method (middle right), product of the VELO and T method and the Final method (bottom), weighted average of the Long and Combined methods. The uncertainties are statistical only.
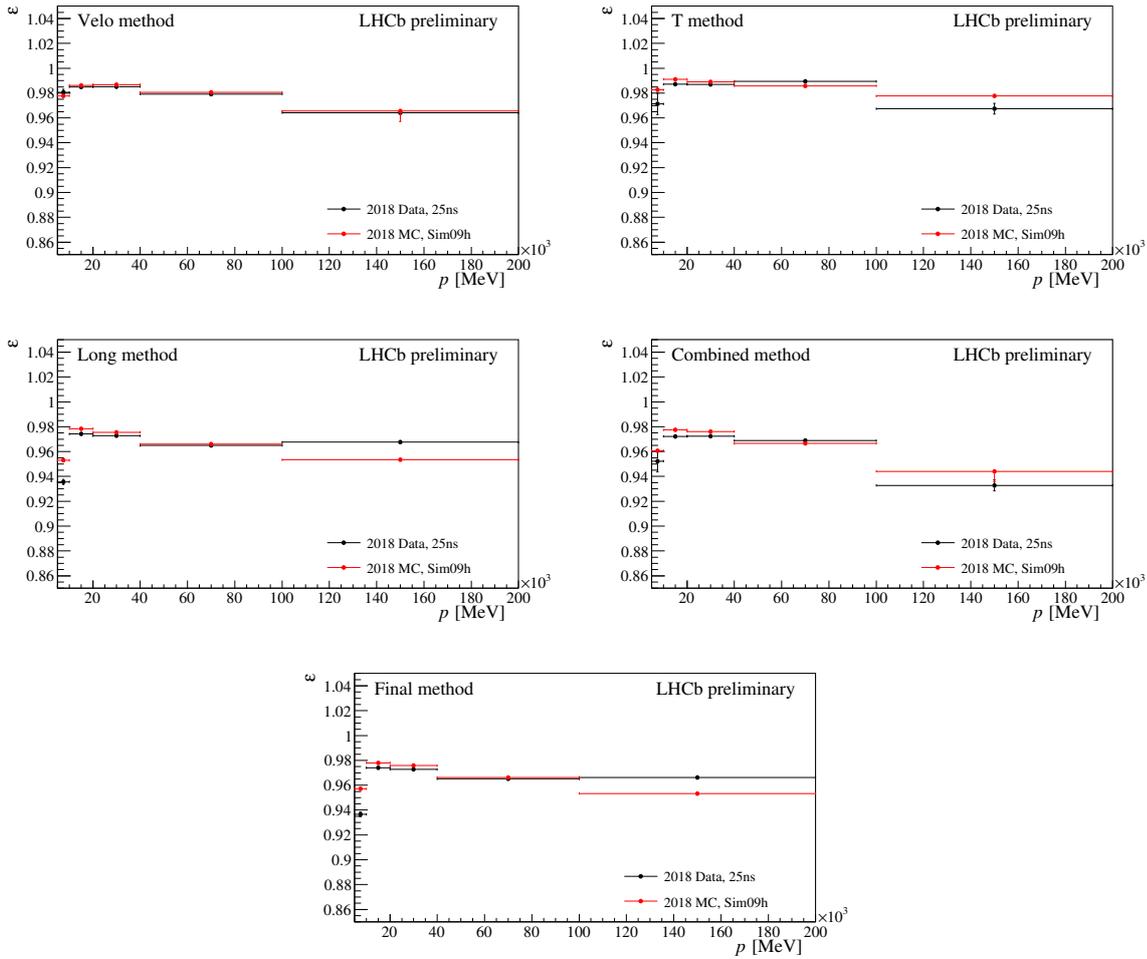
Fig. 10.6: Track reconstruction efficiency versus momentum p for the 2018 data-taking period (25ns bunch spacing). Sim09h denotes the used version of the simulation software used by the LHCb experiment. The track reconstruction efficiency is obtained using the VELO method (top left), the T method (top right), the Long method (middle left), the Combined method (middle right), product of the VELO and T method and the Final method (bottom), weighted average of the Long and Combined methods. The uncertainties are statistical only.
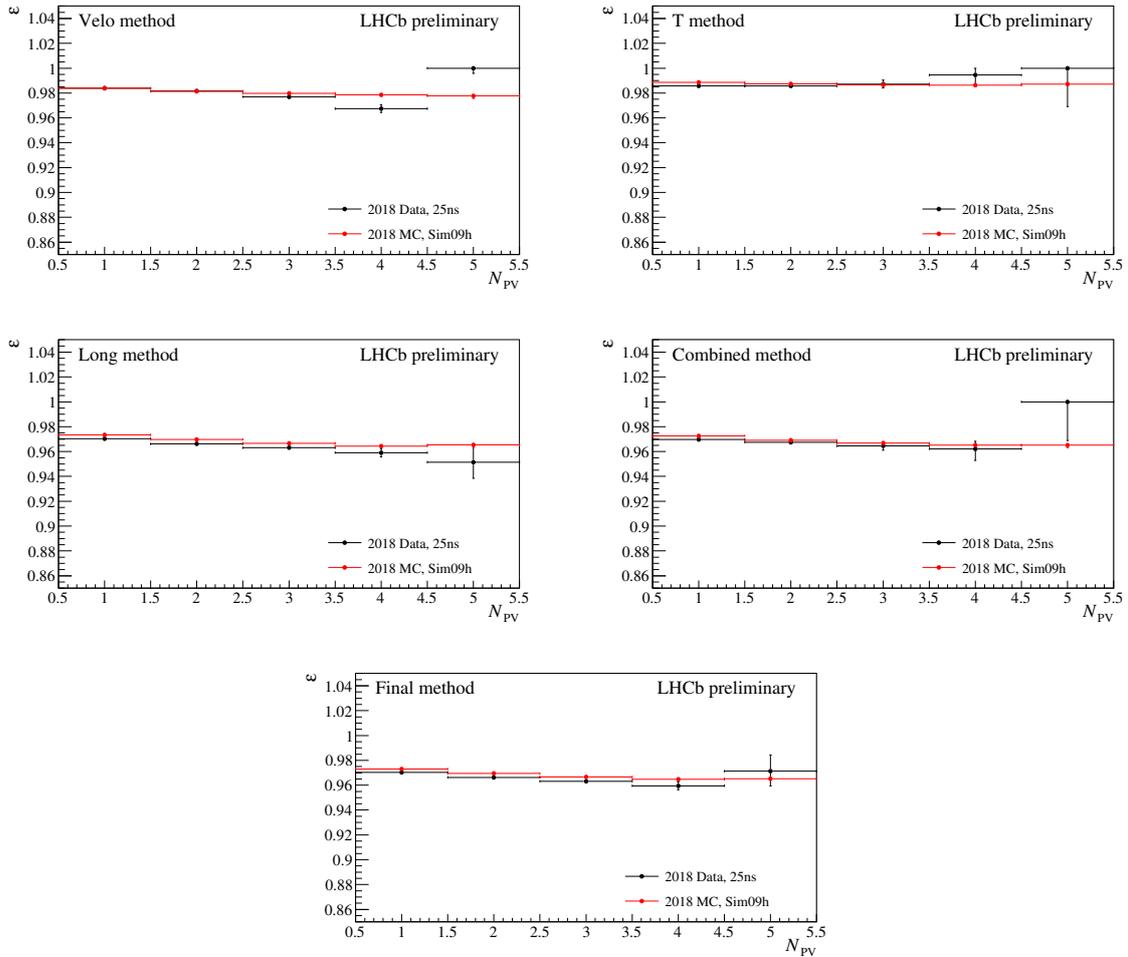
Fig. 10.7: Track reconstruction efficiency versus the number of primary vertices N_{PV} for the 2018 data-taking period (25ns bunch spacing). Sim09h denotes the used version of the simulation software used by the LHCb experiment. The track reconstruction efficiency is obtained using the VELO method (top left), the T method (top right), the Long method (middle left), the Combined method (middle right), product of the VELO and T method and the Final method (bottom), weighted average of the Long and Combined methods. The uncertainties are statistical only.
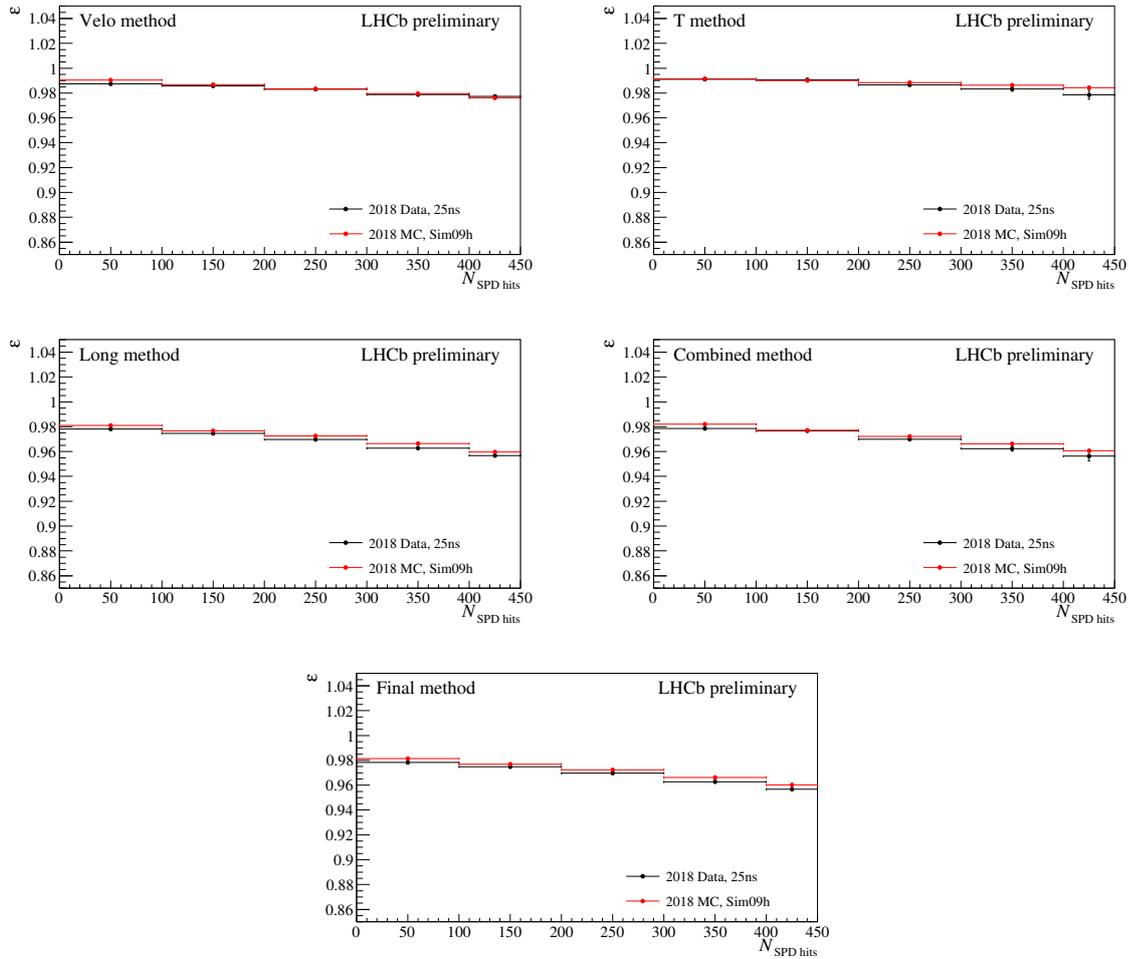
Fig. 10.8: Track reconstruction efficiency versus the number of hits in the SPD detector $N_{SPDhits}$ for the 2018 data-taking period (25ns bunch spacing). Sim09h denotes the used version of the simulation software used by the LHCb experiment. The track reconstruction efficiency is obtained using the VELO method (top left), the T method (top right), the Long method (middle left), the Combined method (middle right), product of the VELO and T method and the Final method (bottom), weighted average of the Long and Combined methods. The uncertainties are statistical only.
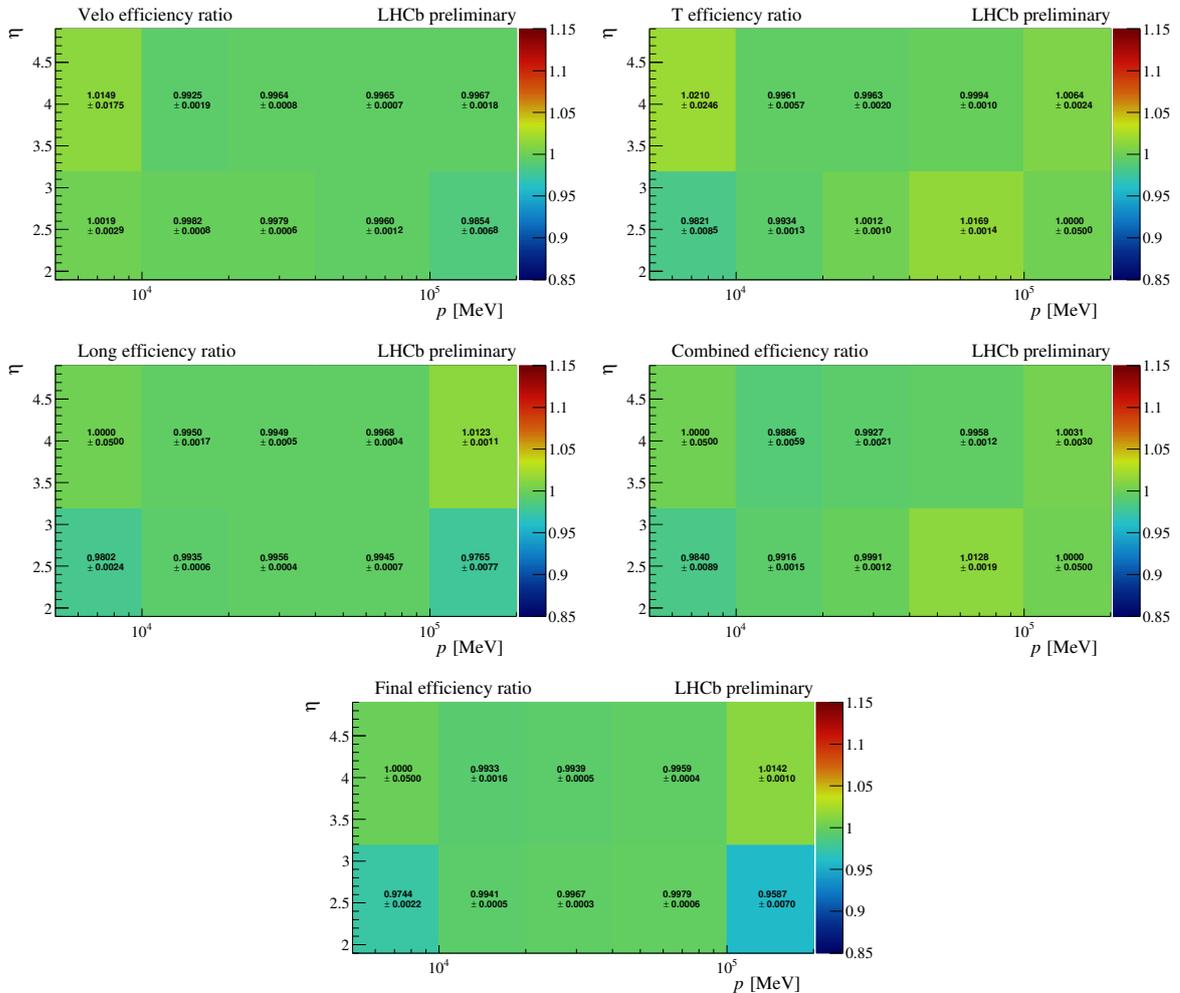
Fig. 10.9: Track reconstruction efficiency ratio of data to simulation in momentum p and pseudorapidity η for the 2018 data-taking period (25ns bunch spacing). Sim09h denotes the used version of the simulation software used by the LHCb experiment. The track reconstruction efficiency is obtained using the VELO method (top left), the T method (top right), the Long method (middle left), the Combined method (middle right), product of the VELO and T method and the Final method (bottom), weighted average of the Long and Combined methods. The ratio for the Final method is made available as a correction table for physics analyses performed by the LHCb collaboration. The uncertainties are statistical only.

Conclusions

In this thesis, the first study of the rare $B^+ \to (K^{*+} \to K^+ \pi^0) \, \mu^+ \mu^-$ is presented. This work is the first attempt to perform the angular analysis with a neutral particle in the final state with the LHCb dataset. The full dataset of $9 \, \mathrm{fb}^{-1}$ collected by the LHCb experiment is utilized in this measurement.

Due to the challenging reconstruction of this decay channel, the selection criteria is carefully chosen and tested in order to maximally suppress the background contributions while preserving high signal efficiency and an even angular acceptance. On top of applying linear cuts in the selection, a multi-variate analysis is used to suppress the background pollution even more. On top of applying a set of simple requirements in the selection, a multi-variate analysis is used to suppress the background pollution even more. In total, 271±28 signal candidates are selected.

The fit model and the shape of the angular background present in the selected data sample is thoroughly investigated using simulation and data samples with large number of signal candidates. The angular shape of the background is modeled to maximize the fit stability and to avoid introducing biases in the angular parameters. The fit model is validated using a fit to the reference resonant $B^+ \to K^{*+} J/\psi$ decay. The results of the fit to the $B^+ \to K^{*+} J/\psi$ decay agree with previous measurements at other experiments.

A detailed study of the maximum likelihood fit to the rare signal channel is done using pseudoexperiments. Multiple angular foldings are employed to maximize the stability of the fit. However, it is shown that the complicated background structure together with the low statistical power of the current data sample results in large uncertainties. The precision of the angular observables to be measured is estimated in five intervals of the dimuon invariant mass squared. Using the FLAVIO package, a likelihood scan as a function of the real part of the Wilson coefficient $\mathrm{Re}(\mathcal{C}_9)$ is performed, assuming a New Physics scenario with $\mathrm{Re}(\mathcal{C}_9)$=-2, as observed in the previous studies. The best possible sensitivity to the deviation of the Wilson coefficient from the Standard Model value is estimated to be ≈ 2.4 standard deviations.

References

[1] 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`.

[2] LHCb collaboration, R. Aaij *et al.*, *"Differential branching fraction and angular analysis of the decay $B_s^0 \to \phi\mu^+\mu^-$ "*, JHEP **07** (2013) 084, `arXiv:1305.2168`.

[3] 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`.

[4] LHCb collaboration, R. Aaij *et al.*, *"Angular Analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ Decay"*, Phys. Rev. Lett. **126** (2021) 161802, `arXiv:2012.13241`.

[5] 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`.

[6] 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`.

[7] 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`.

[8] CMS collaboration, A. M. Sirunyan *et al.*, *"Measurement of angular parameters from the decay $\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s}$ = 8 TeV"*, Phys. Lett. B **781** (2018) 517, `arXiv:1710.02846`.

[9] 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 **10** (2018) 047, `arXiv:1805.04000`.

[10] J. J. Thomson, *"Cathode rays"*, The Electrician **39** (1897) 104.

[11] E. Rutherford, *"The scattering of alpha and beta particles by matter and the structure of the atom"*, Phil. Mag. Ser. 6 **21** (1911) 669.

[12] J. Chadwick, *"Possible Existence of a Neutron"*, Nature **129** (1932) 312.

[13] C. D. Anderson, *"The positive electron"*, Phys. Rev. **43** (1933) 491.

[14] A. H. Bécquerel, *"Sur les radiations invisibles emises par les corps phosphorescents"*, C. R. Acad. Sci. Paris **122** (1896) 501.

REFERENCES

[15] E. Fermi, *"Tentativo di una teoria dell'emissione dei raggi beta"*, Ric. Sci. **4** (1933) 491.

[16] C. Bernardini and L. Bonolis, *Enrico Fermi. His work and legacy.*, Springer Berlin Heidelberg, 2004.

[17] F. N. D. Kurie, J. R. Richardson, and H. C. Paxton, *"The radiations emitted from artificially produced radioactive substances. i. the upper limits and shapes of theβ-ray spectra from several elements"*, Physical Review **49** (1936) 368.

[18] G. Gamow and E. Teller, *"Selection rules for the β-disintegration"*, Phys. Rev. **49** (1936) 895.

[19] R. H. Dalitz, *"Isotopic spin changes in tau and theta decay"*, Proc. Phys. Soc. A **69** (1956) 527.

[20] T. D. Lee and C. N. Yang, *"Question of Parity Conservation in Weak Interactions"*, Phys. Rev. **104** (1956) 254.

[21] C. S. Wu *et al.*, *"Experimental Test of Parity Conservation in β Decay"*, Phys. Rev. **105** (1957) 1413.

[22] J. I. Friedman and V. L. Telegdi, *"Nuclear emulsion evidence for parity nonconservation in the decay chain $\pi^+ \rightarrow \mu^+ \rightarrow e^+$"*, Phys. Rev. **105** (1957) 1681.

[23] R. L. Garwin, L. M. Lederman, and M. Weinrich, *"Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon"*, Phys. Rev. **105** (1957) 1415.

[24] R. P. Feynman and M. Gell-Mann, *"Theory of Fermi interaction"*, Phys. Rev. **109** (1958) 193.

[25] E. C. G. Sudarshan and R. E. Marshak, *"Chirality invariance and the universal fermi interaction"*, Phys. Rev. **109** (1958) 1860.

[26] H. Weyl, *"Gruppentheorie und quantenmechanik"*, Monatshefte für Mathematik und Physik **36** (1929) A48.

[27] T. D. Lee and C. N. Yang, *"Some considerations on global symmetry"*, Phys. Rev. **122** (1961) 1954.

[28] B. L. Ioffe, L. B. Okun, and A. P. Rudik, *Weak interactions in colliding electron beam*, AEC-TR-6592, 1964.

[29] T. Appelquist and J. D. Bjorken, *"On weak interactions at high energies"*, Phys. Rev. D **4** (1971) 3726.

[30] M. Gell-Mann, M. L. Goldberger, N. M. Kroll, and F. E. Low, *"Amelioration of divergence difficulties in the theory of weak interactions"*, Phys. Rev. **179** (1969) 1518.

[31] C. N. Yang and R. L. Mills, *"Conservation of Isotopic Spin and Isotopic Gauge Invariance"*, Phys. Rev. **96** (1954) 191.

[32] S. L. Glashow, *"Partial-symmetries of weak interactions"*, Nuclear Physics **22** (1961) 579.

[33] Particle Data Group, P. A. Zyla *et al.*, *"Review of Particle Physics"*, Progress of Theoretical and Experimental Physics **2020** (2020) , 083C01.

[34] J. Goldstone, *"Field Theories with Superconductor Solutions"*, Nuovo Cim. **19** (1961) 154.

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

[36] P. W. Higgs, *"Broken symmetries and the masses of gauge bosons"*, Phys. Rev. Lett. **13** (1964) 508.

[37] F. Englert and R. Brout, *"Broken symmetry and the mass of gauge vector mesons"*, Phys. Rev. Lett. **13** (1964) 321.

[38] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *"Global Conservation Laws and Massless Particles"*, Phys. Rev. Lett. **13** (1964) 585.

[39] S. Weinberg, *"A model of leptons"*, Phys. Rev. Lett. **19** (1967) 1264.

[40] A. Salam, *"Weak and Electromagnetic Interactions"*, Conf. Proc. C **680519** (1968) 367.

[41] M. Gell-Mann, *"The Eightfold Way: A Theory of strong interaction symmetry"*, CTSL-20, TID-12608, 1961.

[42] Y. Ne'eman, *"Derivation of strong interactions from a gauge invariance"*, Nucl. Phys. **26** (1961) 222.

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

REFERENCES

[44] G. Zweig, *"An SU(3) model for strong interaction symmetry and its breaking. Version 1"*, CERN-TH-401 (1964).

[45] G. Zweig, in *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, D. B. Lichtenberg and S. P. Rosen, eds., pp. 22–101, 1964.

[46] J. D. Bjorken and S. L. Glashow, *"Elementary Particles and SU(4)"*, Phys. Lett. **11** (1964) 255.

[47] S. L. Glashow, J. Iliopoulos, and L. Maiani, *"Weak interactions with lepton-hadron symmetry"*, Phys. Rev. D **2** (1970) 1285.

[48] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *"Evidence for the 2π Decay of the K_2^0 Meson"*, Phys. Rev. Lett. **13** (1964) 138.

[49] M. Kobayashi and T. Maskawa, *"CP-Violation in the Renormalizable Theory of Weak Interaction"*, Progress of Theoretical Physics **49** (1973) 652.

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

[51] M. K. Gaillard and B. W. Lee, *"Δ I = 1/2 Rule for Nonleptonic Decays in Asymptotically Free Field Theories"*, Phys. Rev. Lett. **33** (1974) 108.

[52] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *"Weak decays beyond leading logarithms"*, Rev. Mod. Phys. **68** (1996) 1125, `arXiv:hep-ph/9512380`.

[53] K. G. Wilson, *"Confinement of quarks"*, Physical Review D **10** (1974) 2445.

[54] V. M. Braun and I. E. Filyanov, *"QCD Sum Rules in Exclusive Kinematics and Pion Wave Function"*, Sov. J. Nucl. Phys. **50** (1989) 511.

[55] M. Beneke, T. Feldmann, and D. Seidel, *"Systematic approach to exclusive $B \to Vl^+l^-$, $V\gamma$ decays"*, Nucl. Phys. B **612** (2001) 25, `arXiv:hep-ph/0106067`.

[56] B. Grinstein and D. Pirjol, *"Exclusive rare $b \to K^*\ell^+\ell^-$ decays at low recoil: Controlling the long-distance effects"*, Phys. Rev. D **70** (2004) 114005.

[57] 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`.

[58] D. Aristizabal Sierra, F. Staub, and A. Vicente, *"Shedding light on the $b \to s$ anomalies with a dark sector"*, Phys. Rev. D **92** (2015) 015001, `arXiv:1503.06077`.

[59] A. Falkowski, S. F. King, E. Perdomo, and M. Pierre, *"Flavourful Z′ portal for vector-like neutrino Dark Matter and $R_{K^{(*)}}$"*, JHEP **08** (2018) 061, `arXiv:1803.04430`.

[60] G. Arcadi, T. Hugle, and F. S. Queiroz, *"The Dark $L_\mu - L_\tau$ Rises via Kinetic Mixing"*, Phys. Lett. B **784** (2018) 151, `arXiv:1803.05723`.

[61] J. Kawamura, S. Okawa, and Y. Omura, *"Interplay between the b→ sℓℓ anomalies and dark matter physics"*, Phys. Rev. D **96** (2017) 075041, `arXiv:1706.04344`.

[62] J. M. Cline and J. M. Cornell, *"$R(K^{(*)})$ from dark matter exchange"*, Phys. Lett. B **782** (2018) 232, `arXiv:1711.10770`.

[63] L. Dhargyal, *"A simple model to explain observed muon sector anomalies and small neutrino masses"*, J. Phys. G **46** (2019) 125002, `arXiv:1711.09772`.

[64] D. Aristizabal Sierra, M. Hirsch, and S. G. Kovalenko, *"Leptoquarks: Neutrino masses and accelerator phenomenology"*, Phys. Rev. D **77** (2008) 055011, `arXiv:0710.5699`.

[65] Y. Cai *et al.*, *"From the trees to the forest: a review of radiative neutrino mass models"*, Front. in Phys. **5** (2017) 63, `arXiv:1706.08524`.

[66] F. J. Botella, G. C. Branco, and M. Nebot, *"Singlet Heavy Fermions as the Origin of B Anomalies in Flavour Changing Neutral Currents"*, IFIC-17-60, CFTP-17-010 (2017) `arXiv:1712.04470`.

[67] J. Heeck and D. Teresi, *"Pati-Salam explanations of the B-meson anomalies"*, JHEP **12** (2018) 103, `arXiv:1808.07492`.

[68] C. Hati, G. Kumar, J. Orloff, and A. M. Teixeira, *"Reconciling B-meson decay anomalies with neutrino masses, dark matter and constraints from flavour violation"*, JHEP **11** (2018) 011, `arXiv:1806.10146`.

[69] P. Agrawal, S. Blanchet, Z. Chacko, and C. Kilic, *"Flavored dark matter, and its implications for direct detection and colliders"*, Phys. Rev. D **86** (2012) 055002, `arXiv:1109.3516`.

[70] R. Allahverdi, P. S. B. Dev, and B. Dutta, *"A simple testable model of baryon number violation: Baryogenesis, dark matter, neutron–antineutron oscillation and collider signals"*, Phys. Lett. B **779** (2018) 262, `arXiv:1712.02713`.

REFERENCES

[71] B. Fornal and B. Grinstein, *"Dark Matter Interpretation of the Neutron Decay Anomaly"*, Phys. Rev. Lett. **120** (2018) 191801, `arXiv:1801.01124`, [Erratum: Phys. Rev. Lett. **124** (2020) 219901].

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

[73] *"What is CERN's mission?"*, `https://home.cern/about/who-we-are/our-mission/`, CERN. Accessed: 02-11-2020.

[74] CERN press release, *"The LEP story"*, `home.cern/news/press-release/cern/lep-story`, CERN, 2000. Accessed: 22-07-2020.

[75] C. L. Smith, *"Genesis of the Large Hadron Collider"*, Philos. Trans. R. Soc. A **373** (2015) 20140037.

[76] LHC commissioning group, *"Longer term lhc schedule"*, `https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm`, 2020. Accessed: 14-11-2020.

[77] LHC commissioning group, *"Run III performance estimate"*, `https://lhc-commissioning.web.cern.ch/performance/Run-3-performance.htm`, CERN, 2020. Accessed: 14-11-2020.

[78] J. Haffner, *"The CERN accelerator complex."*, OPEN-PHO-ACCEL-2013-056, 2013, General Photo.

[79] LHCb collaboration, *LHCb : Technical Proposal*, Tech. Proposal, CERN, Geneva, 1998.

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

[81] LHCb collaboration, *"The LHCb detector at the LHC"*, JINST **3** (2008) S08005.

[82] LHCb collaboration, C. Elsässer, *"$b\bar{b}$ production angle plots"*, `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html`, CERN, 2012. Accessed: 24-07-2020.

[83] C. e. a. Ilgner, *"The Beam Conditions Monitor of the LHCb Experiment"*, arXiv:1001.2487, 2010, `http://cds.cern.ch/record/1233669`.

[84] G. A. Schuler, *"Theory of heavy quark production in p-anti p collisions"*, Nucl. Phys. B - Proceedings Supplements **16** (1990) 343.

[85] R. Lindner, *"LHCb layout"*, LHCb-PHO-GENE-2008-002, 2008.

[86] LHCb collaboration, R. Aaij *et al.*, *"Observation of five new narrow Ω_c^0 states decaying to $\Xi_c^+ K^-$"*, Phys. Rev. Lett. **118** (2017) 182001, `arXiv:1703.04639`.

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

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

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

[90] LHCb collaboration, R. Aaij *et al.*, *"Measurement of forward $W \to e\nu$ production in pp collisions at $\sqrt{s} = 8\,TeV$"*, High Energ. Phys. **2016** (2016) 30, `arXiv:1608.01484`.

[91] LHCb collaboration, R. Aaij *et al.*, *"Measurement of forward W and Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 8$ TeV"*, High Energ. Phys. **2016** (2016) 131, `arXiv:1605.00951`.

[92] LHCb collaboration, R. Aaij *et al.*, *"First observation of forward $Z \to b\bar{b}$ production in pp collisions at $\sqrt{s} = 8\,TeV$"*, Physics Letters B **776** (2018) 430, `arXiv:1709.03458`.

[93] LHCb collaboration, R. Aaij *et al.*, *"Study of Υ production in pPb collisions at $\sqrt{s_{NN}} = 8.16$ TeV"*, JHEP **11** (2018) 194, `arXiv:1810.07655`, [Erratum: JHEP **02**, (2020) 093].

[94] LHCb collaboration, R. Aaij *et al.*, *"Precision luminosity measurements at LHCb"*, JINST **9** (2014) P12005, `arXiv:1410.0149`.

[95] LHCb collaboration, R. Aaij *et al.*, *"Measurement of Antiproton Production in pHe Collisions at $\sqrt{s_{NN}} = 110$ GeV"*, Phys. Rev. Lett. **121** (2018) 222001, `arXiv:1808.06127`.

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

REFERENCES

[97] LHCb collaboration, *"LHCb detector performance"*, International Journal of Modern Physics A **30** (2015) 1530022.

[98] R. Aaij *et al.*, *"Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC"*, JINST **14** (2019) P04013.

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

[100] M. Vesterinen, *"Considerations on the LHCb dipole magnet polarity reversal"*, CERN-LHCb-PUB-2014-006, 2014, On behalf of the LHCb collaboration.

[101] J. Gassner, M. Needham, and O. Steinkamp, *"Layout and Expected Performance of the LHCb TT Station"*, CERN-LHCb-2003-140, 2004.

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

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

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

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

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

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

[108] LHCb collaboration, *"LHCb online system, data acquisition and experiment control: Technical Design Report"*, CERN-LHCC-2001-040, 2001.

[109] LHCb collaboration, *"LHCb Trigger and Online Technical Design Report"*, CERN-LHCC-2014-016, 2014.

[110] M. Cattaneo, P. Charpentier, P. Clarke, and S. Roiser, *"Recent and planned changes to the LHCb computing model"*, J. Phys. Conf. Ser. **513** (2014) 032017.

[111] LHCb collaboration, R. Aaij *et al.*, *"Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV"*, JHEP **10** (2015) 172, `arXiv:1509.00771`, [Erratum: JHEP **05** (2017) 063].

[112] LHCb collaboration, R. Aaij *et al.*, *"Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV"*, JHEP **03** (2016) 159, `arXiv:1510.01707`, [Erratum: JHEP **09** (2016) 013 , Erratum: JHEP **05** (2017) 074].

[113] LHCb collaboration, *"Trigger schemes"*, `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html`, CERN, 2015. Accessed: 28-07-2020.

[114] N. Metropolis and S. Ulam, *"The Monte Carlo Method"*, Journal of the American Statistical Association **44** (1949) 335.

[115] LHCb collaboration, *"LHCb Upgrade Software and Computing"*, CERN-LHCC-2018-007, 2018.

[116] LHCb collaboration, *"LHCb computing: Technical Design Report"*, CERN-LHCC-2005-019, 2005.

[117] I. B. et. al, *"Handling of the generation of primary events in Gauss, the LHCb simulation framework"*, Journal of Physics: Conference Series **331** (2011) 032047.

[118] Torbjörn Sjöstrand and Stephen Mrenna and Peter Skands, *"PYTHIA 6.4 physics and manual"*, Journal of High Energy Physics **2006** (2006) 026, `arXiv:hep-ph/0603175`.

[119] Torbjörn Sjöstrand and Stephen Mrenna and Peter Skands, *"A brief introduction to PYTHIA 8.1"*, Computer Physics Communications **178** (2008) 852, `arXiv:0710.3820`.

[120] D. J. Lange, *"The EvtGen particle decay simulation package"*, Nucl. Instrum. Meth. A **462** (2001) 152.

[121] GEANT4 collaboration, S. Agostinelli *et al.*, *"GEANT4–a simulation toolkit"*, Nucl. Instrum. Meth. A **506** (2003) 250.

[122] J. Allison *et al.*, *"Geant4 developments and applications"*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[123] LHCb collaboration, *"LHCb digitization application"*, `http://lhcbdoc.web.cern.ch/lhcbdoc/boole/`, CERN. Accessed: 29-07-2020.

REFERENCES

[124] LHCb collaboration, *"LHCb High Level Trigger application"*, `http://lhcbdoc.web.cern.ch/lhcbdoc/moore/`, CERN. Accessed: 29-07-2020.

[125] LHCb collaboration, *"LHCb event reconstruction application"*, `http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/`, CERN. Accessed: 29-07-2020.

[126] BaBar collaboration, B. Aubert *et al.*, *"Evidence for the rare decay $b \to K^* \ell^+ \ell^-$ and measurement of the $b \to K \ell^+ \ell^-$ branching fraction"*, Phys. Rev. Lett. **91** (2003) 221802, `arXiv:hep-ex/0308042`.

[127] BaBar collaboration, B. Aubert *et al.*, *"Angular Distributions in the Decays $B \to K^* l^+ l^-$"*, Phys. Rev. D **79** (2009) 031102, `arXiv:0804.4412`.

[128] 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`.

[129] T. Hurth, C. Langenbruch, and F. Mahmoudi, *"Direct determination of Wilson coefficients using $B^0 \to K^{*0} \mu^+ \mu^-$ decays"*, JHEP **11** (2017) 176, `arXiv:1708.04474`.

[130] T. Blake *et al.*, *"Angular analysis of $B^0 \to K^{*0} \mu^+ \mu^-$ at LHCb with 1 fb^{-1}"*, LHCb-ANA-2012-051, 2013, Linked to LHCB-PAPER-2013-019.

[131] Belle collaboration, J.-T. Wei *et al.*, *"Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \to K^{(*)} \ell^+ \ell^-$"*, Phys. Rev. Lett. **103** (2009) 171801, `arXiv:0904.0770`.

[132] 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`.

[133] LHCb collaboration, R. Aaij *et al.*, *"Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0} \mu^+ \mu^-$"*, Phys. Rev. Lett. **108** (2012) 181806, `arXiv:1112.3515`.

[134] 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`.

[135] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *"On the impact of power corrections in the prediction of $B \to K^* \mu^+ \mu^-$ observables"*, JHEP **12** (2014) 125, `arXiv:1407.8526`.

[136] W. Altmannshofer and D. M. Straub, *"New physics in $b \to s$ transitions after LHC Run 1"*, Eur. Phys. J. C **75** (2015) 382, `arXiv:1411.3161`.

[137] A. Bharucha, D. M. Straub, and R. Zwicky, *"B → Vℓ⁺ℓ⁻ in the Standard Model from light-cone sum rules"*, JHEP **08** (2016) 098, `arXiv:1503.05534`.

[138] B. Capdevila *et al.*, *"Patterns of New Physics in b → sℓ⁺ℓ⁻ transitions in the light of recent data"*, JHEP **01** (2018) 093, `arXiv:1704.05340`.

[139] D. M. Straub, *"flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond"*, `arXiv:1810.08132`.

[140] LHCb collaboration, R. Aaij *et al.*, *"Test of lepton universality in beauty-quark decays"*, `arXiv:2103.11769`, LHCb-PAPER-2021-004, CERN-EP-2021-042.

[141] 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`.

[142] 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`.

[143] LHCb collaboration, R. Aaij *et al.*, *"Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$"*, Phys. Rev. Lett. **115** (2015) 111803, `arXiv:1506.08614`, [Erratum: Phys.Rev.Lett. 115, 159901 (2015)].

[144] LHCb collaboration, R. Aaij *et al.*, *"Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$"*, Phys. Rev. Lett. **120** (2018) 121801, `arXiv:1711.05623`.

[145] A. K. Alok, A. Dighe, S. Gangal, and D. Kumar, *"Continuing search for new physics in b → sμμ decays: two operators at a time"*, JHEP **06** (2019) 089, `arXiv:1903.09617`.

[146] W. Altmannshofer, C. Niehoff, P. Stangl, and D. M. Straub, *"Status of the $B \to K^*\mu^+\mu^-$ anomaly after Moriond 2017"*, Eur. Phys. J. C **77** (2017) 377, `arXiv:1703.09189`.

[147] C. A. Beteta *et al.*, *"Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC"*, CERN-LHCb-DP-2020-001, 2020.

[148] A. Martens, *"Towards a measurement of the angle γ of the Unitarity Triangle with the LHCb detector at the LHC (CERN): calibration of the calorimeters using an energy flow technique and first observation of the $B_s^0 \to \bar{D}^0\bar{K}^{*0}$ decay"*, CERN-THESIS-2011-128, 2011, Presented 09 Sep 2011.

REFERENCES

[149] I. M. Belyaev, E. M. Govorkova, V. Y. Egorychev, and D. V. Savrina, *"Study of π^0/γ reconstruction and selection efficiency in the LHCb experiment"*, Moscow Univ. Phys. Bull. **70** (2015) 497.

[150] G. Mitchell, *"How much data is on the internet?"*, `https://www.sciencefocus.com/future-technology/how-much-data-is-on-the-internet/`, 2019. Accessed: 20-09-2021.

[151] V. V. Gligorov, *"A single track HLT1 trigger"*, CERN-LHCb-PUB-2011-003. LHCb-INT-2010-053, 2011.

[152] R. Aaij and J. Albrecht, *"Muon triggers in the High Level Trigger of LHCb"*, CERN-LHCb-PUB-2011-017, 2011.

[153] M. Williams *et al.*, *"The HLT2 Topological Lines"*, CERN-LHCb-PUB-2011-002, 2011.

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

[155] F. Archilli *et al.*, *"Performance of the Muon Identification at LHCb"*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[156] W. D. Hulsbergen, *"Decay chain fitting with a Kalman filter"*, Nucl. Instrum. Meth. A **552** (2005) 566, `arXiv:physics/0503191`.

[157] M. Pivk and F. R. Le Diberder, *"sPlot: A Statistical tool to unfold data distributions"*, Nucl. Instrum. Meth. A **555** (2005) 356, `arXiv:physics/0402083`.

[158] M. Pivk, *"sPlot: A Quick introduction"*, in *Statistical Problems in Particle Physics, Astrophysics and Cosmology*, 173–177, 2006, `arXiv:physics/0602023`.

[159] A. P. Dempster, *"An overview of multivariate data analysis"*, Journal of Multivariate Analysis **1** (1971) 316.

[160] W. S. McCulloch and W. Pitts, *"A logical calculus of the ideas immanent in nervous activity"*, The Bulletin of Mathematical Biophysics **5** (1943) 115.

[161] A. Hocker *et al.*, *"TMVA - Toolkit for Multivariate Data Analysis"*, CERN-OPEN-2007-007, `arXiv:physics/0703039`.

[162] CMS collaboration, A. M. Sirunyan *et al.*, *"Angular analysis of the decay $B^+ \rightarrow K^*(892)^+\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV"*, JHEP **04** (2021) 124, `arXiv:2010.13968`.

[163] A. M. Legendre, *"Sur l'attraction des sphéroides."*, Mém. Math. et Phys. présentés à l'Ac. r. des. sc. par divers savants **22** (1785).

[164] F. James and M. Roos, *"Minuit - a system for function minimization and analysis of the parameter errors and correlations"*, Computer Physics Communications **10** (1975) 343.

[165] I. Antcheva *et al.*, *"ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization"*, Comput. Phys. Commun. **180** (2009) 2499, `arXiv:1508.07749`.

[166] T. Bayes, Rev. *"An essay toward solving a problem in the doctrine of chances"*, Phil. Trans. Roy. Soc. Lond. **53** (1764) 370.

[167] B. Taylor, *Methodus incrementorum directa et inversa*, 1715.

[168] G. Breit and E. Wigner, *"Capture of Slow Neutrons"*, Phys. Rev. **49** (1936) 519.

[169] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics*, Springer New York, 1979.

[170] Belle collaboration, K. Chilikin *et al.*, *"Experimental constraints on the spin and parity of the $Z(4430)^+$"*, Phys. Rev. D **88** (2013) 074026, `arXiv:1306.4894`.

[171] LHCb collaboration, R. Aaij *et al.*, *"Measurement of the branching fraction and CP asymmetry in $B^+ \rightarrow J/\psi\rho^+$ decays"*, Eur. Phys. J. C **79** (2019) 537, `arXiv:1812.07041`.

[172] 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`.

[173] D. Aston *et al.*, *"A study of $K^-\pi^+$ scattering in the reaction $K^-p \rightarrow K^-\pi^+n$ at 11 GeV/c"*, Nuclear Physics B **296** (1988) 493.

[174] B. Dunwoodie, *"Fits to $K\pi$ I=1/2 S-wave Amplitude and Phase Data"*, `https://www.slac.stanford.edu/~wmd/kpi_swave/kpi_swave_fit.note`, SLAC, 2013. Accessed: 18-08-2021.

[175] P. L. Chebyshev, *Théorie des mécanismes connus sous le nom de parallélogrammes*, Imprimerie de l'Académie impériale des sciences, 1853.

REFERENCES

[176] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, *"A Comparative study of the decays $B \to (K, K^*) \ell^+ \ell^-$ in standard model and supersymmetric theories"*, Phys. Rev. D **61** (2000) 074024, `arXiv:hep-ph/9910221`.

[177] 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. D **76** (2007) 031102, `arXiv:0704.0522`.

[178] 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`.

[179] LHCb collaboration, R. Aaij *et al.*, *"Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays"*, Phys. Rev. D **88** (2013) 052002, `arXiv:1307.2782`.

[180] M. Matsumoto and T. Nishimura, *"Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator"*, ACM Trans. Model. Comput. Simul. **8** (1998) 3.

[181] G. Marsaglia and W. W. Tsang, *"Some difficult-to-pass tests of randomness"*, Journal of Statistical Software **7** (2002) 1.

[182] P. L'Ecuyer and R. Simard, *"TestU01: A C Library for Empirical Testing of Random Number Generators"*, ACM Trans. Math. Softw. **33** (2007) 1.

[183] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge University Press, 2007.

[184] W. C. Thacker, *"The role of the hessian matrix in fitting models to measurements"*, Journal of Geophysical Research **94** (1989) 6177.

[185] G. J. Feldman and R. D. Cousins, *"A Unified approach to the classical statistical analysis of small signals"*, Phys. Rev. D **57** (1998) 3873, `arXiv:physics/9711021`.

[186] CMS collaboration, S. Chatrchyan *et al.*, *"Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV"*, JINST **7** (2012) P10002, `arXiv:1206.4071`.

[187] CMS collaboration, A. M. Sirunyan *et al.*, *"Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV"*, JINST **13** (2018) P06015, `arXiv:1804.04528`.

[188] ATLAS collaboration, G. Aad *et al.*, *"Muon reconstruction performance of the AT-LAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV"*, Eur. Phys. J. C **76** (2016) 292, `arXiv:1603.05598`.

[189] LHCb Collaboration, R. Aaij *et al.*, *"Measurement of the track reconstruction efficiency at LHCb"*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[190] LHCb collaboration, R. Aaij *et al.*, *"Measurement of the electron reconstruction efficiency at LHCb"*, JINST **14** (2019) P11023, `arXiv:1909.02957`.

[191] R. Aaij *et al.*, *"Tesla: an application for real-time data analysis in High Energy Physics"*, Comput. Phys. Commun. **208** (2016) 35, `arXiv:1604.05596`.

[192] B. Audurier *et al.*, *"Measurement of the coherent J/ψ production in ultra-peripheral Pb-Pb collisions at $\sqrt{s_{NN}} = 5$ TeV with the LHCb experiment"*, CERN-LHCb-ANA-2017-044, 2017.

[193] F. Archilli, G. Frau, and R. Kopecna, *"TrackCalib Package Twiki"*, `https://twiki.cern.ch/twiki/bin/view/LHCb/TrackCalib`, 2017. Accessed: 03-09-2021.

[194] F. Archilli and R. Kopecna, *"Muon track reconstruction efficiency at LHCb"*, LHCB-FIGURE-2020-010, 2020.

[195] T. Skwarnicki, *"A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances"*, DESY-F31-86-02, 1986.

[196] E. D. Bloom and C. Peck, *"Physics with the Crystal Ball Detector"*, Ann. Rev. Nucl. Part. Sci. **33** (1983) 143.

[197] D. Watts, *"The Crystal Ball programme at MAMI"*, EPJ Web of Conferences **37** (2012) 01027.

A   Theoretical introduction to the $\mathbf{B^+ \to K^{*+} \mu^+ \mu^-}$ decay

A.1   Decay rate

The full form of Eq. 3.2 with explicitly stated f_i is a rather lengthy Eq. A.1:

$$
\begin{aligned}
\frac{\mathrm{d}^4 \Gamma}{\mathrm{d}\cos\theta_K \mathrm{d}\cos\theta_K \mathrm{d}\phi \mathrm{d}q^2} &= \frac{9}{32\pi} \sum_i J_i\left(q^2\right) f_i \left(\cos\theta_L, \cos\theta_K, \phi\right) = \\
&= \frac{9}{32\pi} \{ \, J_{1s} \sin^2\theta_K \\
&\qquad + J_{1c} \cos^2\theta_K \\
&\qquad + J_{2s} \sin^2\theta_K \cos 2\theta_L \\
&\qquad + J_{2c} \cos^2\theta_K \cos 2\theta_L \\
&\qquad + J_3 \sin^2\theta_K \sin^2\theta_L \cos 2\phi \\
&\qquad + J_4 \sin 2\theta_K \sin 2\theta_L \cos\phi \\
&\qquad + J_5 \sin 2\theta_K \sin\theta_L \cos\phi \\
&\qquad + J_{6s} \sin^2\theta_K \cos\theta_L \\
&\qquad + J_7 \sin 2\theta_K \sin\theta_L \sin\phi \\
&\qquad + J_8 \sin 2\theta_K \sin 2\theta_L \sin\phi \\
&\qquad + J_9 \sin^2\theta_K \sin^2\theta_L \sin 2\phi \, \} \, .
\end{aligned}
\tag{A.1}
$$

This lengthy formula can be simplified by neglecting the muon mass. This is a good assumptions for $q^2 \gtrsim 1\,\mathrm{GeV}^2$. Under this assumption, the following relations can be obtained:

$$
\begin{aligned}
J_1^c &= 1 - \frac{4}{3} J_1^s \, , \\
J_2^s &= \frac{1}{3} J_1^s \, , \\
J_2^c &= \frac{4}{3} J_1^s - 1 \, .
\end{aligned}
\tag{A.2}
$$

In some cases it is also convenient to define *CP-asymmetric* angular observables A_i besides the usual S_i variables (see Eq. 3.3)

$$
A_i = \frac{J_i - \bar{J}_i}{\Gamma + \bar{\Gamma}} \, .
\tag{A.3}
$$

Rewriting the Eq. A.1 using the *CP-asymmetric* S_i variables results in the following formula:

$$
\left. \frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{d}\cos\theta_L \, \mathrm{d}\cos\theta_K \, \mathrm{d}\phi} \right|_{\mathrm{P}} = \frac{9}{32\pi} \left[\frac{3}{4}(1 - F_{\mathrm{L}}) \sin^2\theta_K \right.
$$
$$
+ F_{\mathrm{L}} \cos^2\theta_K + \frac{1}{4}(1 - F_{\mathrm{L}}) \sin^2\theta_K \cos 2\theta_L
$$
$$
- F_{\mathrm{L}} \cos^2\theta_K \cos 2\theta_L + S_3 \sin^2\theta_K \sin^2\theta_L \cos 2\phi
$$
$$
+ S_4 \sin 2\theta_K \sin 2\theta_L \cos\phi + S_5 \sin 2\theta_K \sin\theta_L \cos\phi
$$
$$
+ \frac{4}{3} A_{\mathrm{FB}} \sin^2\theta_K \cos\theta_L + S_7 \sin 2\theta_K \sin\theta_L \sin\phi
$$
$$
\left. + S_8 \sin 2\theta_K \sin 2\theta_L \sin\phi + S_9 \sin^2\theta_K \sin^2\theta_L \sin 2\phi \right]
$$

A.2 S-wave decay rate

The decay rate of the S-wave is

$$
\left. \frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{d}\cos\theta_L \, \mathrm{d}\cos\theta_K \, \mathrm{d}\phi} \right|_{\mathrm{S}} = \frac{3}{16\pi} F_S \sin^2\theta_L \, . \tag{A.4}
$$

The P-wave and S-wave interference term can be parameterized as follows:

$$
\left. \frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{d}\cos\theta_L \, \mathrm{d}\cos\theta_K \, \mathrm{d}\phi} \right|_{\mathrm{PS}} = \frac{3}{16\pi} \left[S_{S1} \sin^2\theta_L \cos\theta_K \right.
$$
$$
+ S_{S2} \sin 2\theta_L \sin\theta_K \cos\phi
$$
$$
+ S_{S3} \sin\theta_L \sin\theta_K \cos\phi
$$
$$
+ S_{S4} \sin\theta_L \sin\theta_K \sin\phi
$$
$$
\left. + S_{S5} \sin 2\theta_L \sin\theta_K \sin\phi \right] \, . \tag{A.5}
$$

These two terms need to be added to the full measured decay rate. The full angular description of the decay is then in Eq. 3.9.

B   Event selection

B.1   Crystal Ball function

The Crystal Ball function is a probability density function widely used to model processes with losses [195]. It consists of a gaussian core and one power-law low end tail, that describes the loss, typically from the final state radiation. The function got its name from the Crystal Ball collaboration [196]. The experiment was placed at the SPEAR accelerator at SLAC National Laboratory and designed as a spark chamber surrounded by an almost complete sphere (covering 98% of the solid angle) made of scintillating crystals. Therefore, the detector got its prophetic name. The detector is operating until today. It is located in Mainz, placed at the MAMI microtron [197].

The Crystal Ball function is then defined as

$$\mathcal{P}(x; \alpha, n, \bar{x}, \sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n} & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases} , \tag{B.1}$$

where A and α and n describe the tail, μ and σ are the mean and the width of the peak. N is a normalization factor, A and B are constants defined as:

$$\begin{aligned} A &= \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) , \\ B &= \frac{n}{|\alpha|} - |\alpha| . \end{aligned} \tag{B.2}$$

B.1.1   Double sided Crystal Ball function

The Crystal Ball function can be extended to contain a gaussian core and two power-law low end tails. The double sided Crystal Ball function is then defined as

$$\mathcal{P}_{CB}(x; x_{peak}, \sigma, n_1, n_2, \alpha_1, \alpha_2) = N \cdot \begin{cases} A_1 \cdot (B_1 - \frac{x-x_{peak}}{\sigma})^{-n_1} & \text{for } \frac{x-x_{peak}}{\sigma} \leq -\alpha_1 \\ \exp(\frac{-(x-x_{peak})^2}{2\sigma^2}) & \text{for } -\alpha_1 \leq \frac{x-x_{peak}}{\sigma} \leq \alpha_2 \\ A_2 \cdot (B_2 - \frac{x-x_{peak}}{\sigma})^{-n_2} & \text{for } \alpha_2 \leq \frac{x-x_{peak}}{\sigma} \end{cases} , \tag{B.3}$$

N is a normalization factor, $A_{1,2}$ and $B_{1,2}$ are constants defined as:

$$\begin{aligned} A_{1,2} &= (\frac{n_{1,2}}{\|n_{1,2}\|})^{n_{1,2}} \cdot \exp(\pm \frac{\alpha_{1,2}^2}{2}), \\ B_{1,2} &= \frac{n_{1,2}}{\|\alpha_{1,2}\|} - \|\alpha_{1,2}\| . \end{aligned} \tag{B.4}$$

B.2 ExpGaus function

ExpGaus function is a function used to describe partially reconstructed backgrounds in B meson decays. The definition is in Eq. B.5. The μ denotes the mean of the distribution, σ is the variance of the function, D is a constant representing the decay of the B meson.

$$f_{EG}(x) = \begin{cases} \exp\left(-\frac{\mu - D}{\sigma^2}\left(x - D\right)\right)\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) & \text{if } x \leq D \\ \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) & \text{otherwise} \end{cases}. \quad \text{(B.5)}$$

B.3 Reweighted distributions of parameters used for the MLP training
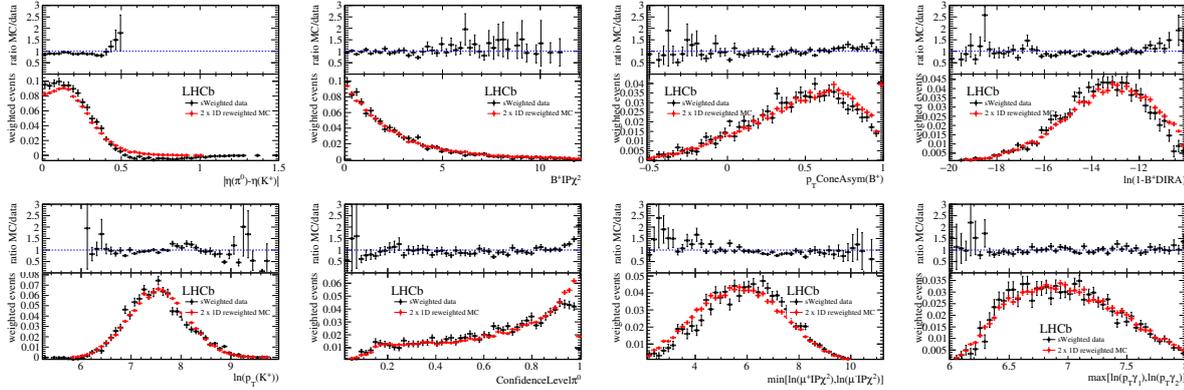
Fig. B.1: Reweighted distributions of variables used for training the MLP for the year 2011. The black points represent *sWeight*ed data in the q^2 region of J/ψ, the red points are $2 \times 1D$ reweighted MC events of $B^+ \rightarrow K^{*+}J/\psi$.
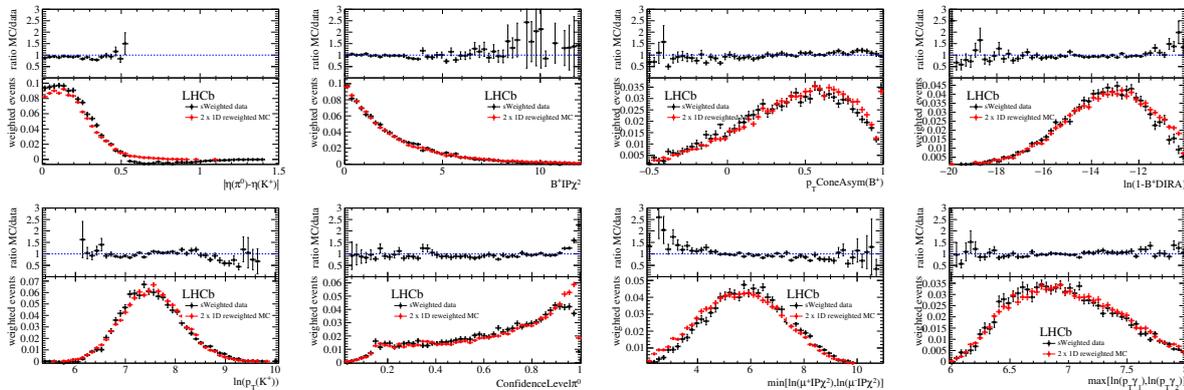
Fig. B.2: Reweighted distributions of variables used for training the MLP for the year 2012. The black points represent *sWeight*ed data in the q^2 region of J/ψ, the red points are $2 \times 1D$ reweighted MC events of $B^+ \rightarrow K^{*+}J/\psi$.
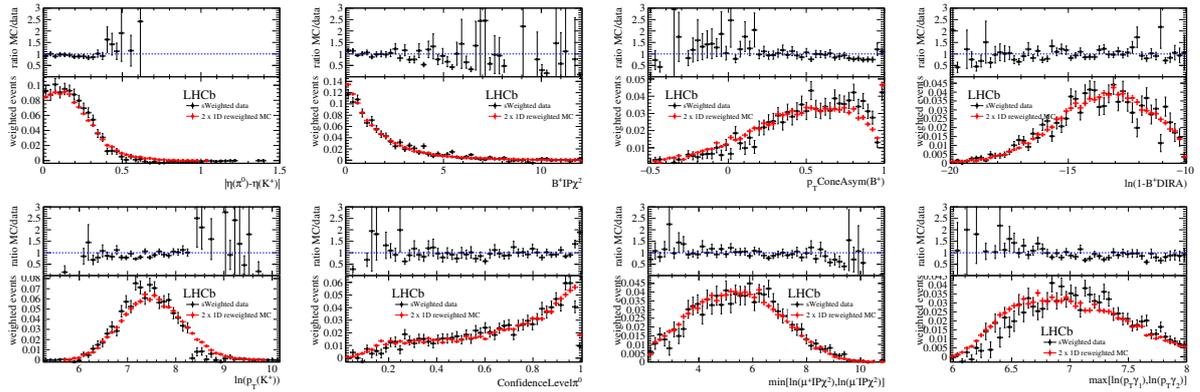
Fig. B.3: Reweighted distributions of variables used for training the MLP for the year 2015. The black points represent *sWeight*ed data in the q^2 region of J/ψ, the red points are $2 \times 1D$ reweighted MC events of $B^+ \to K^{*+} J/\psi$.
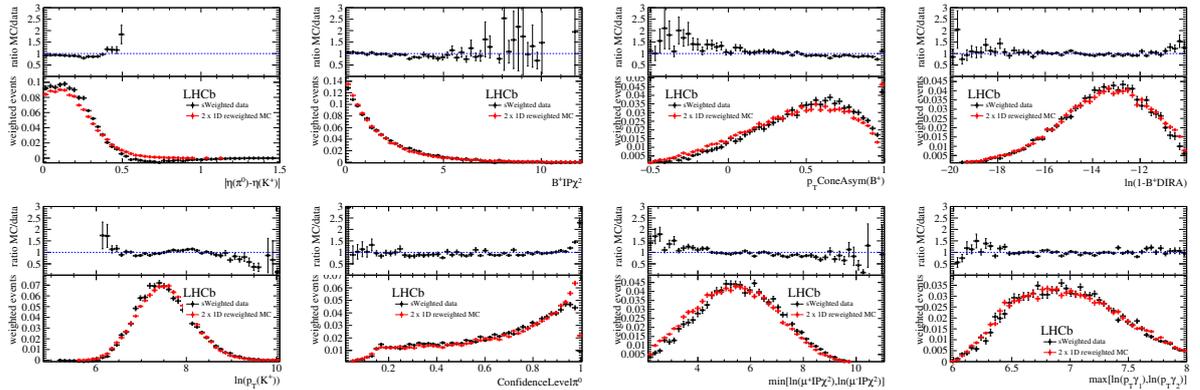
Fig. B.4: Reweighted distributions of variables used for training the MLP for the year 2016. The black points represent *sWeight*ed data in the q^2 region of J/ψ, the red points are $2 \times 1D$ reweighted MC events of $B^+ \to K^{*+} J/\psi$.

Fig. B.5: Reweighted distributions of variables used for training the MLP for the year 2017. The black points represent *sWeight*ed data in the q^2 region of J/ψ, the red points are $2 \times 1D$ reweighted MC events of $B^+ \to K^{*+}J/\psi$. As the data-taking conditions were very similar in 2016-2018 and $B^+ \to K^{*+}J/\psi$ MC sample is not available for 2017, a sample of 2016 MC is used.
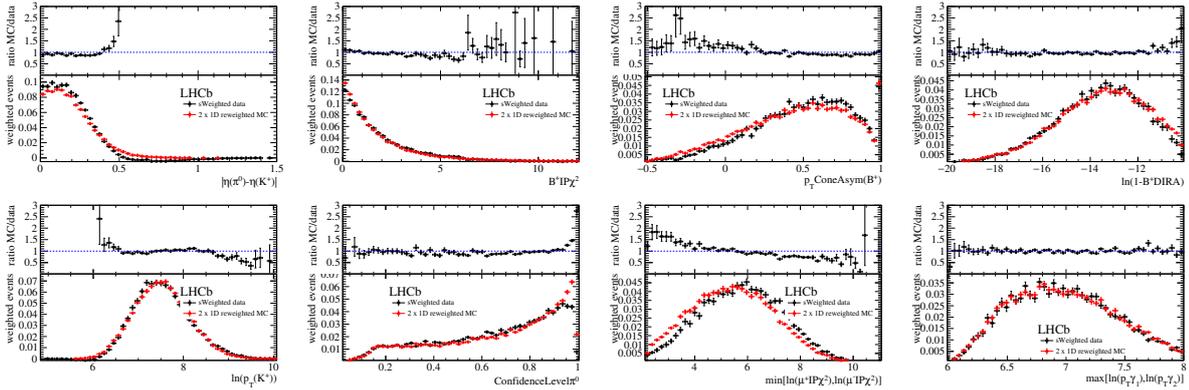
Fig. B.6: Reweighted distributions of variables used for training the MLP for the year 2018. The black points represent *sWeight*ed data in the q^2 region of J/ψ, the red points are $2 \times 1D$ reweighted MC events of $B^+ \to K^{*+}J/\psi$. As the data-taking conditions were very similar in 2016-2018 and $B^+ \to K^{*+}J/\psi$ MC sample is not available for 2018, a sample of 2016 MC is used.

B.4   Signal yield in bins of the dimuon invariant mass squared
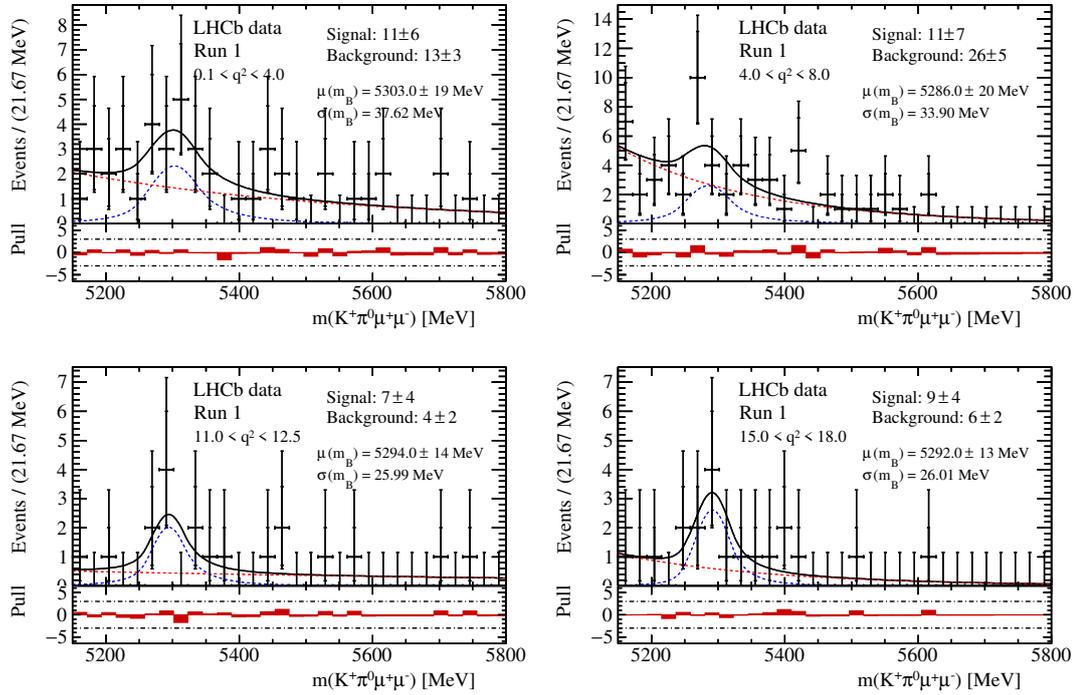
Fig. B.7: Signal channel yield in q^2 bins in Run I.
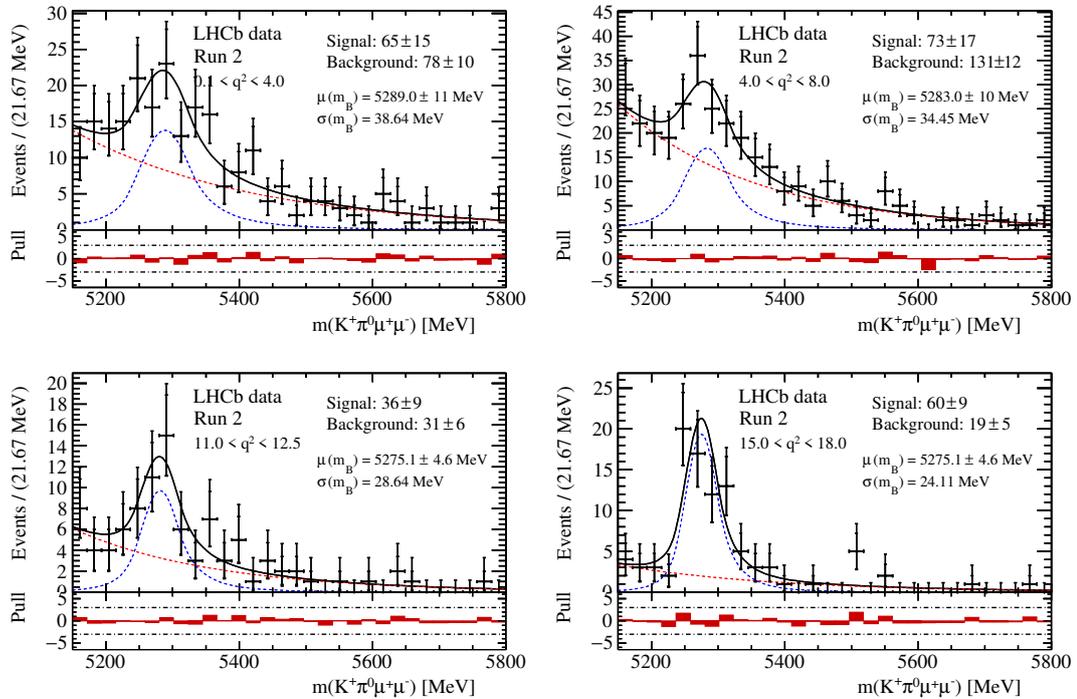
Fig. B.8: Signal channel yield in q^2 bins in Run II.

C   Angular parametrization
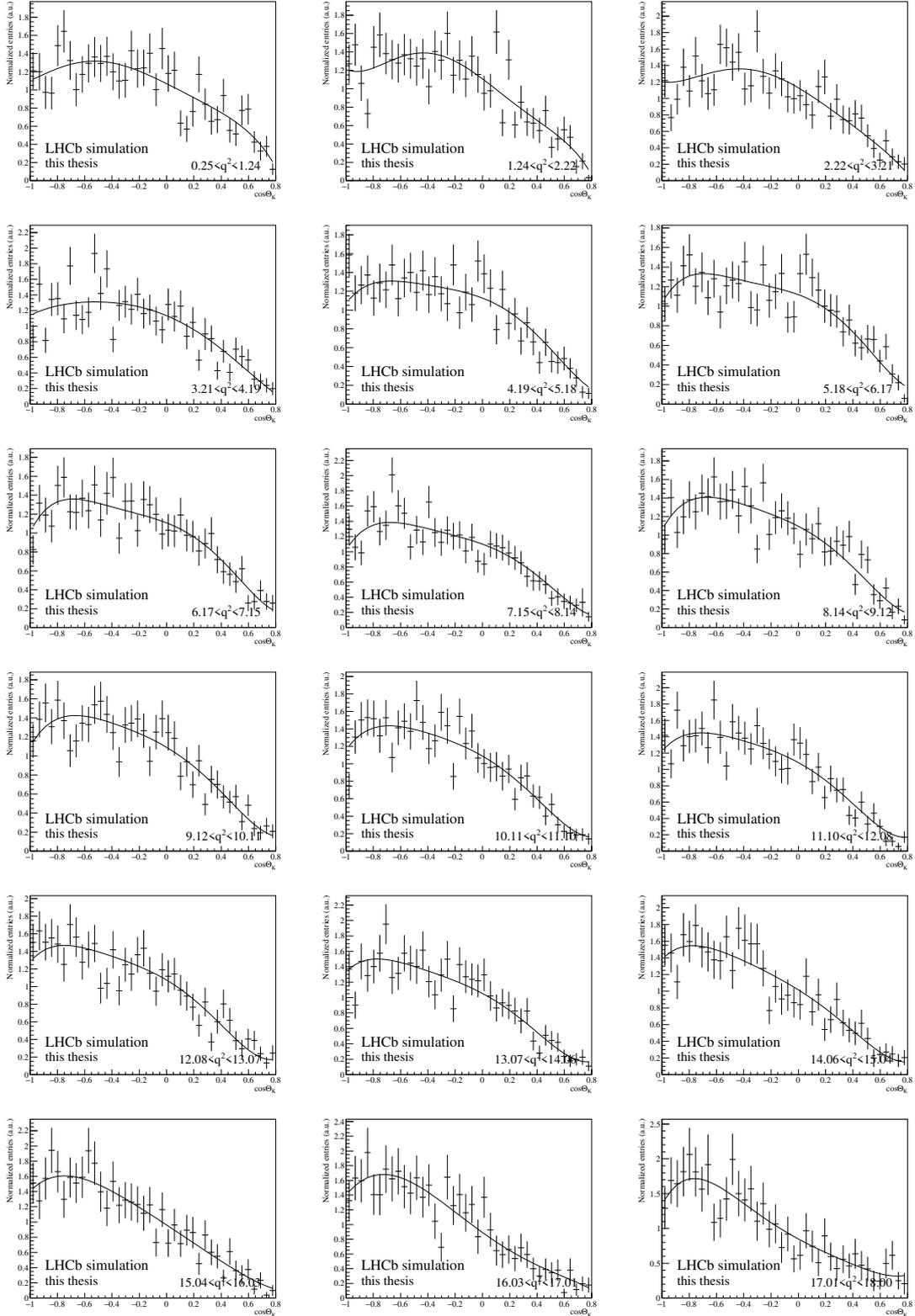
Fig. C.1: One-dimensional projections of the angular acceptance in the dimension of $\cos(\theta_K)$ in 18 bins of q^2. The data points are Run I PHSP MC sample, the solid curve is the four dimensional Legendre-polynomial parametrization.
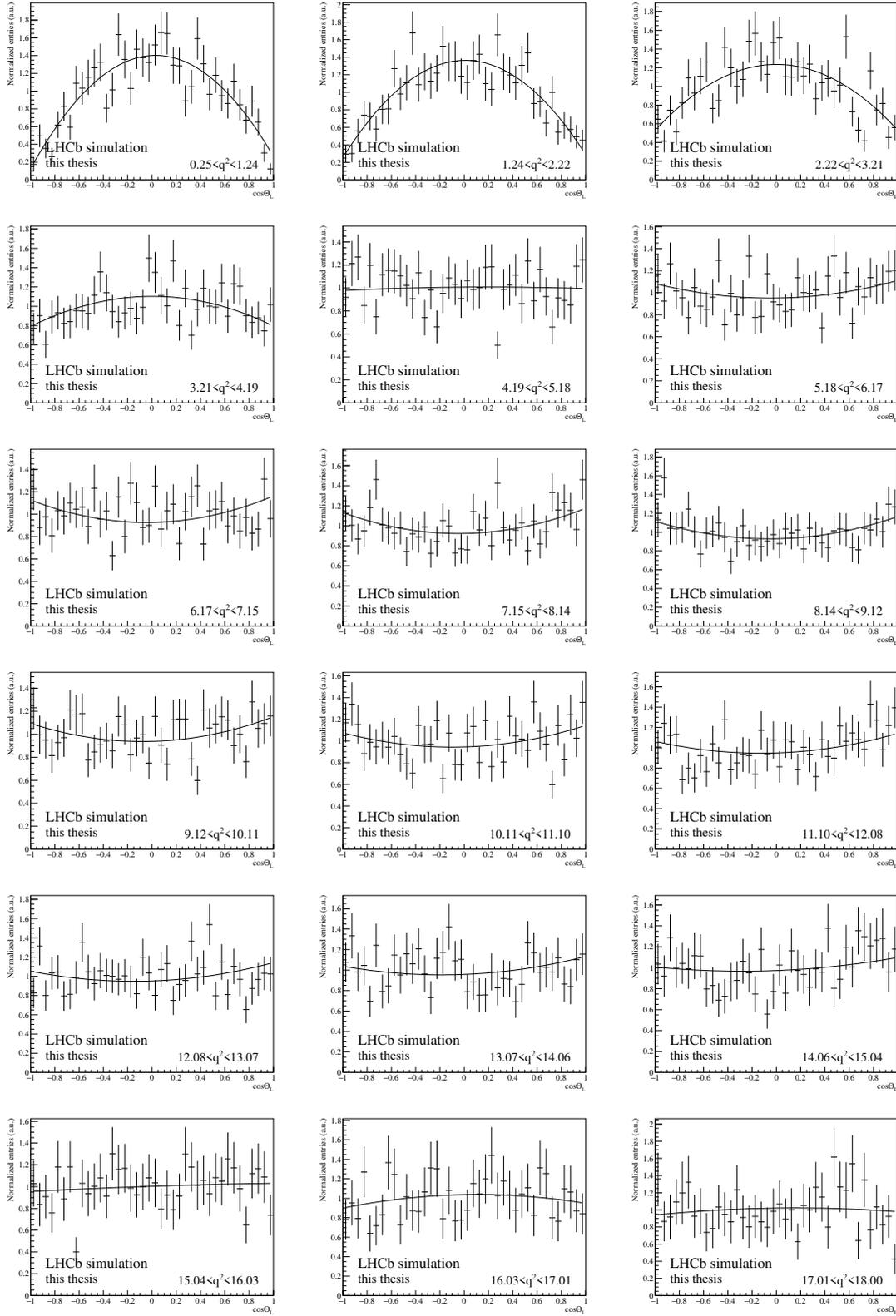
Fig. C.2: One-dimensional projections of the angular acceptance in the dimension of $\cos(\theta_L)$ in 18 bins of q^2. The data points are Run I PHSP MC, the solid curve is the four dimensional Legendre-polynomial parametrization.

160

Fig. C.3: One-dimensional projections of the angular acceptance in the dimension of ϕ in 18 bins of q^2. The data points are Run I PHSP MC, the solid curve is the four dimensional Legendre-polynomial parametrization.
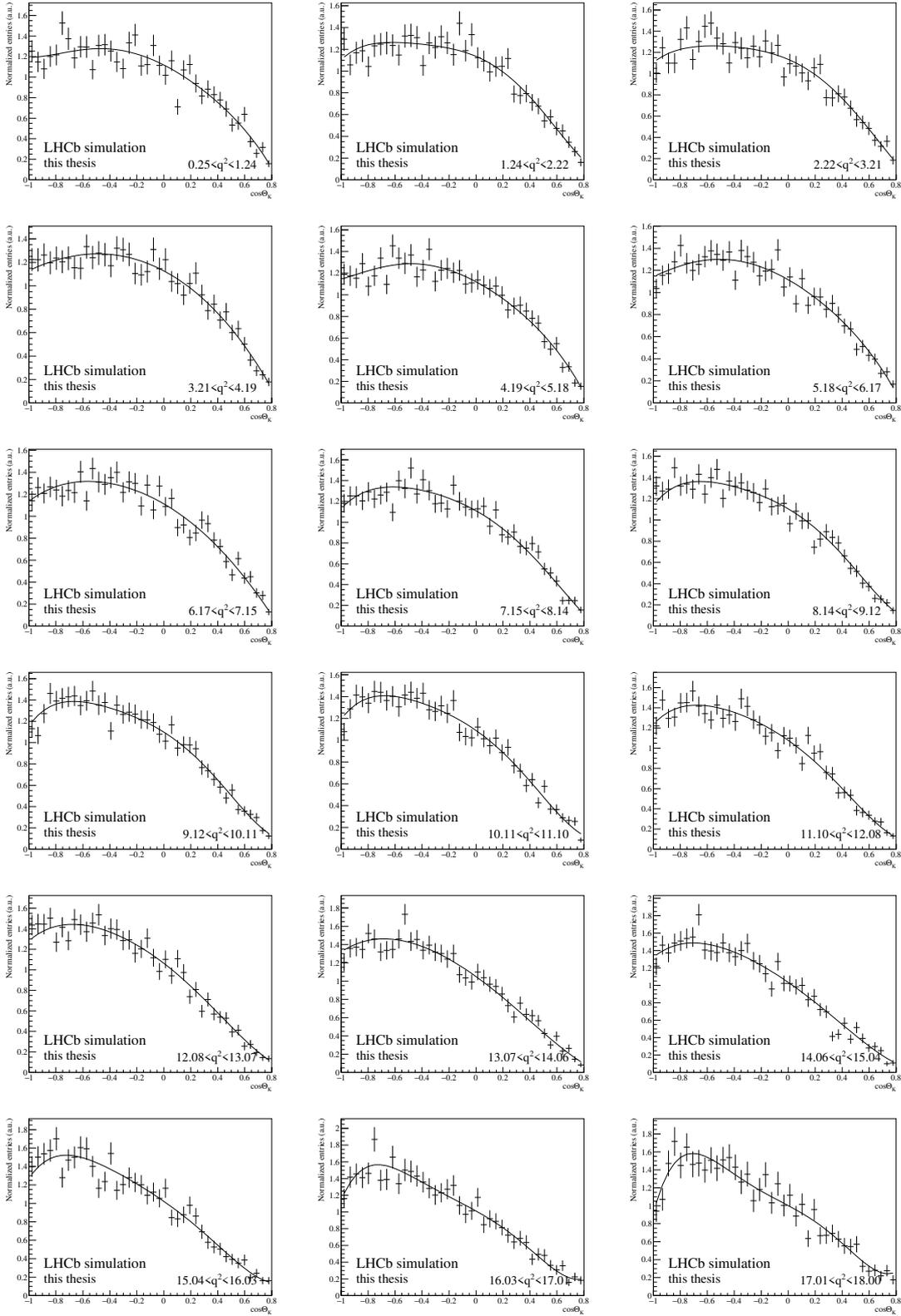
Fig. C.4: One-dimensional projections of the angular acceptance in the dimension of $\cos(\theta_K)$ in 18 bins of q^2. The data points are Run II PHSP MC, the solid curve is the four dimensional Legendre-polynomial parametrization.

162

Fig. C.5: One-dimensional projections of the angular acceptance in the dimension of $\cos(\theta_L)$ in 18 bins of q^2. The data points are Run II PHSP MC, the solid curve is the four dimensional Legendre-polynomial parametrization.
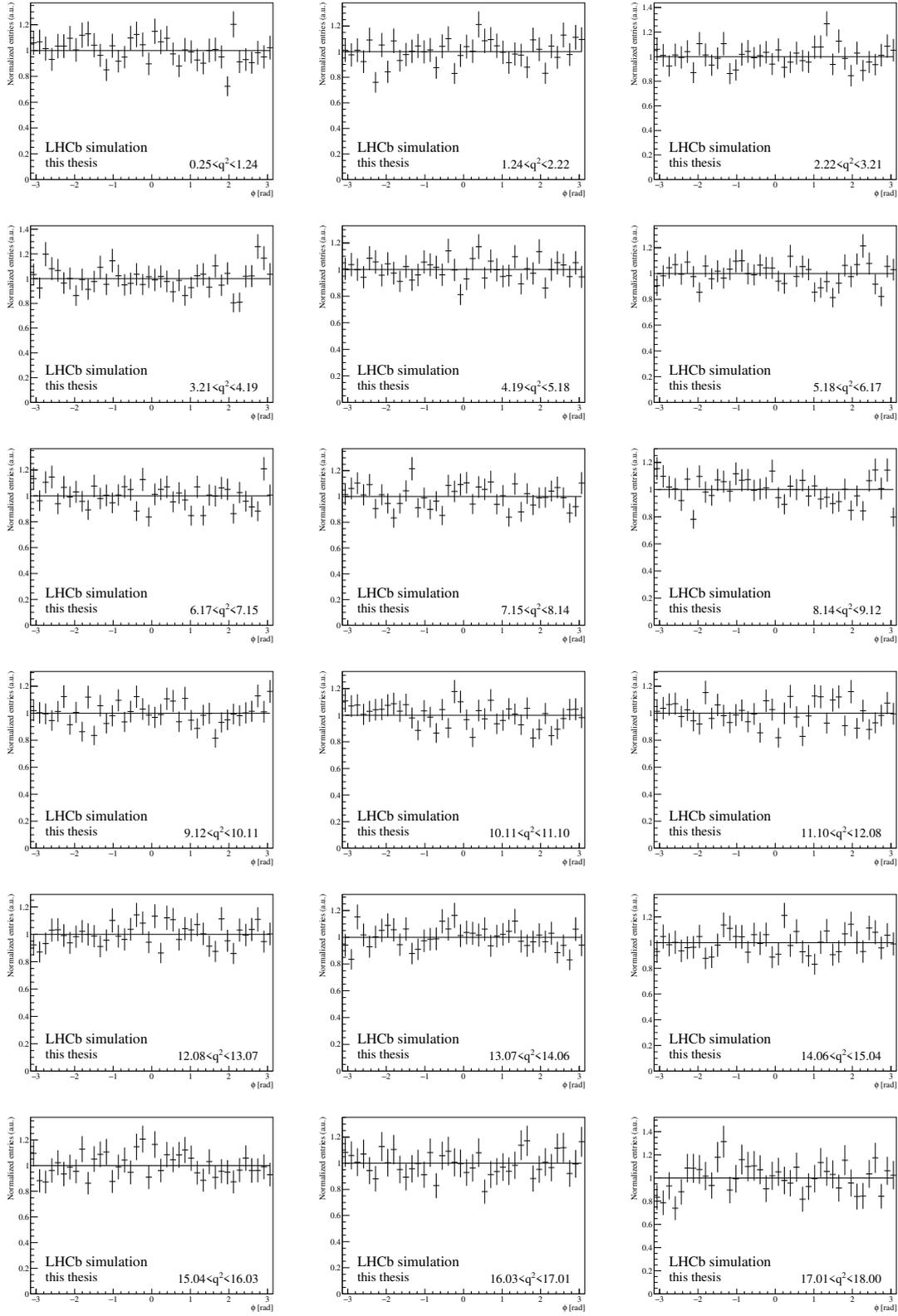
Fig. C.6: One-dimensional projections of the angular acceptance in the dimension of ϕ in 18 bins of q^2. The data points are Run II PHSP, the solid curve is the four dimensional Legendre-polynomial parametrization.

C.1 Validation of the generation of the pseudoexperiments
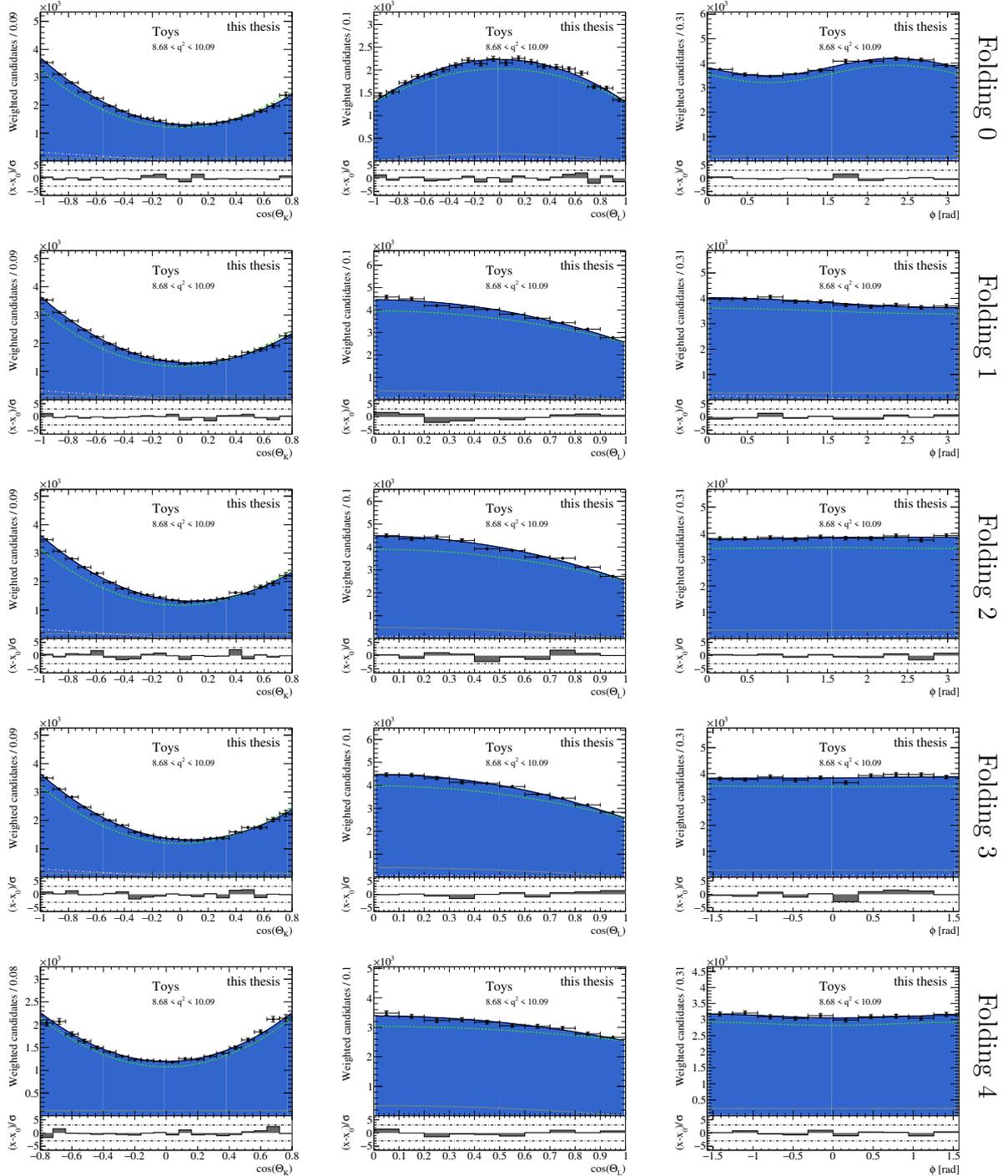
Fig. C.7: Fit with applied folding to 52 000 pseudoexperiments with only the signal component. The black points represent the generated pseudoexperiments, the black line represents the fitted model. The blue space represents the signal component. The green dashed line shows only the P-wave component, the orange dotted line represents the S-wave components and the dot-and-dash line depicts the interference between the P-wave and the S-wave.

165
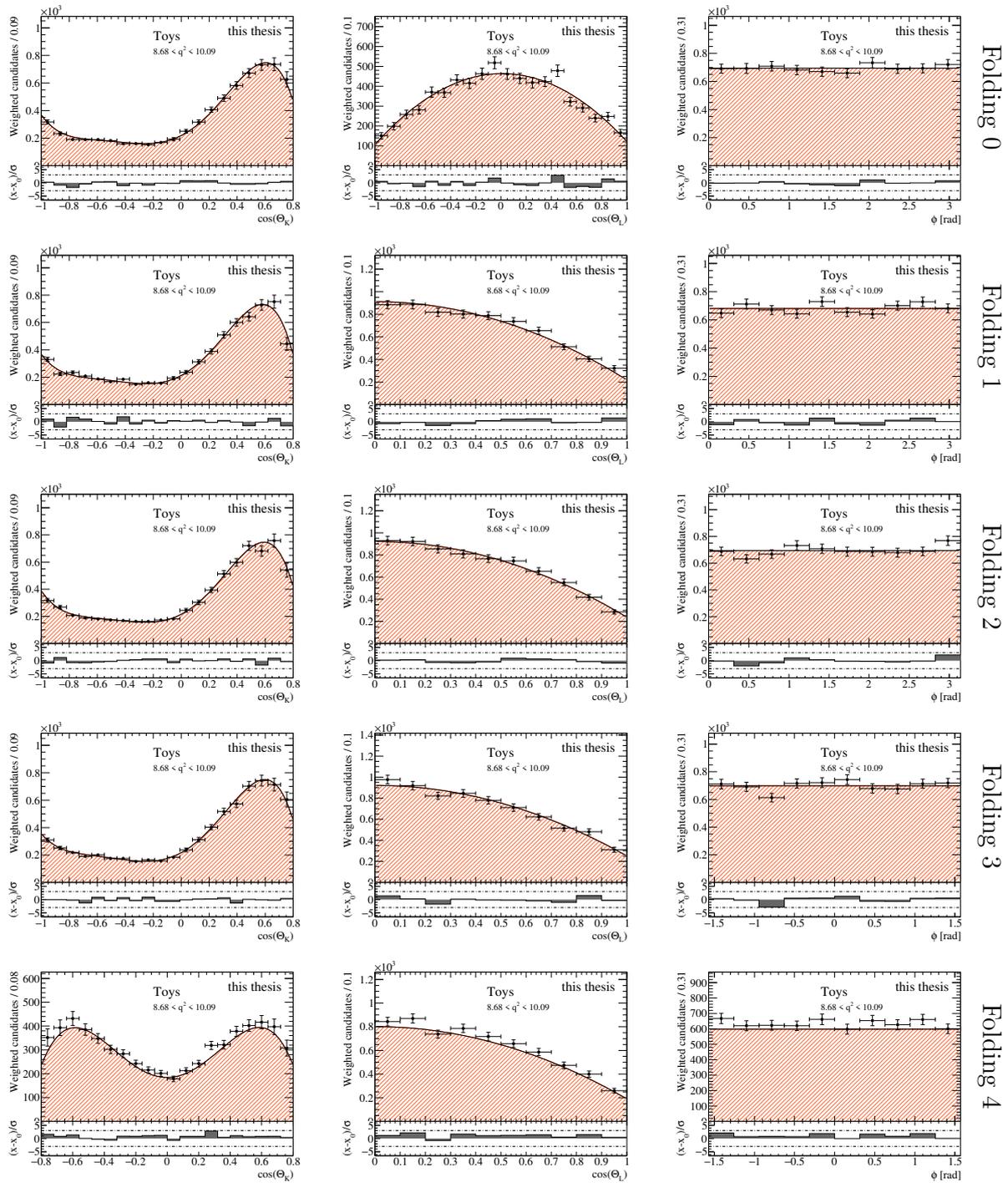
Fig. C.8: Fit with applied folding to 13 000 pseudoexperiments with only the background component. The black points represent the generated pseudoexperiments, the black line represents the fitted model. The red space represents the background component.

D   Pseudoexperiments

D.1   Large scale pseudoexperiments
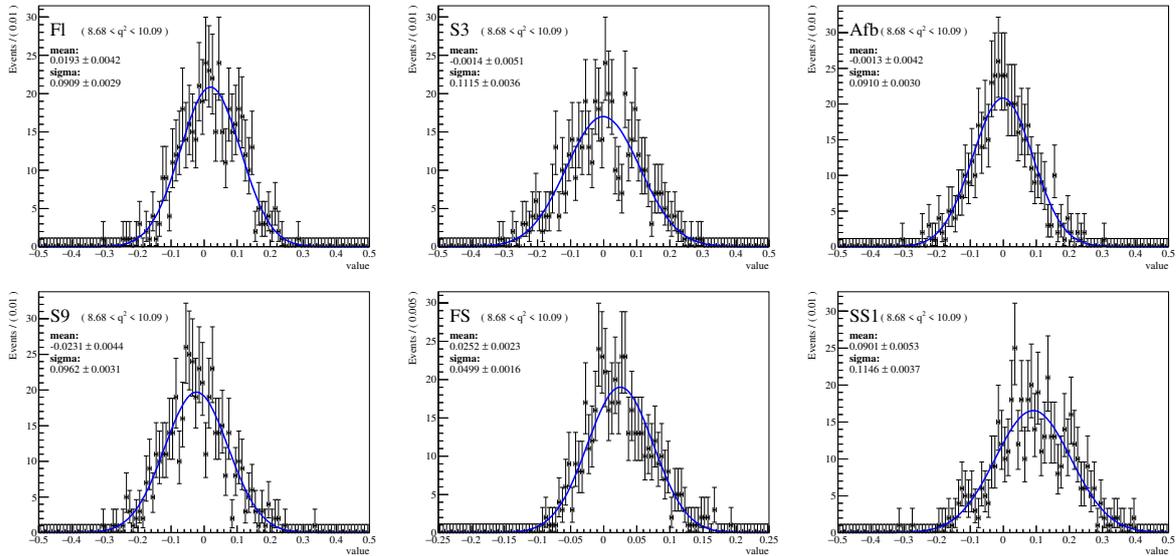
Fig. D.1: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the refrence $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 65 000 pseudoevents. In the fit to the pseudoexperiments, folding 0 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.


Fig. D.2: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the refrence $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 65 000 pseudoevents. In the fit to the pseudoexperiments, folding 1 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.

167
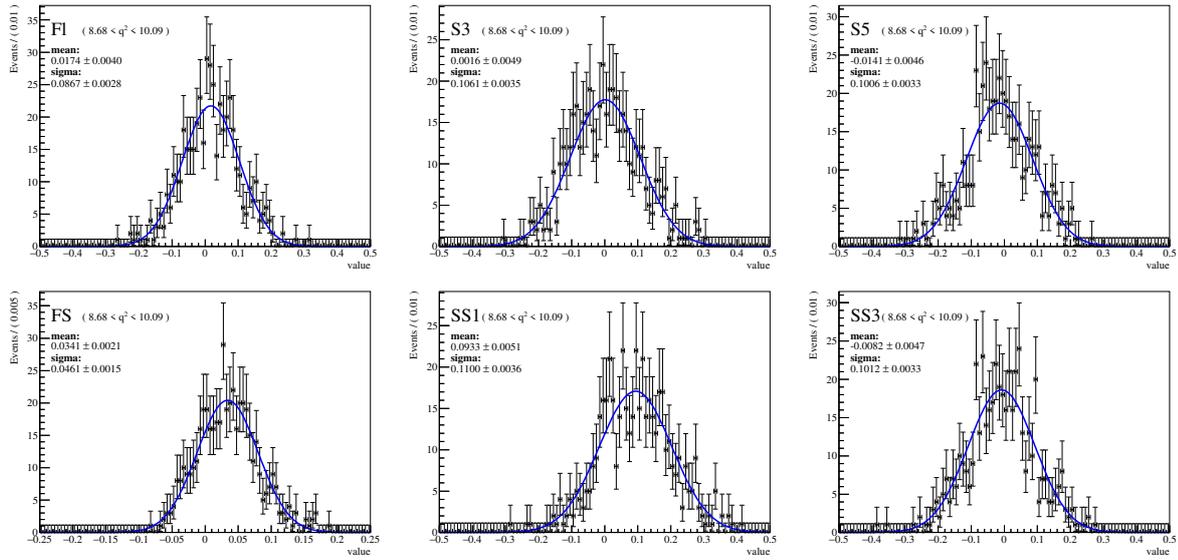
Fig. D.3: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the refrence $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 65 000 pseudoevents. In the fit to the pseudoexperiments, folding 2 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.


Fig. D.4: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the refrence $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 65 000 pseudoevents. In the fit to the pseudoexperiments, folding 3 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.
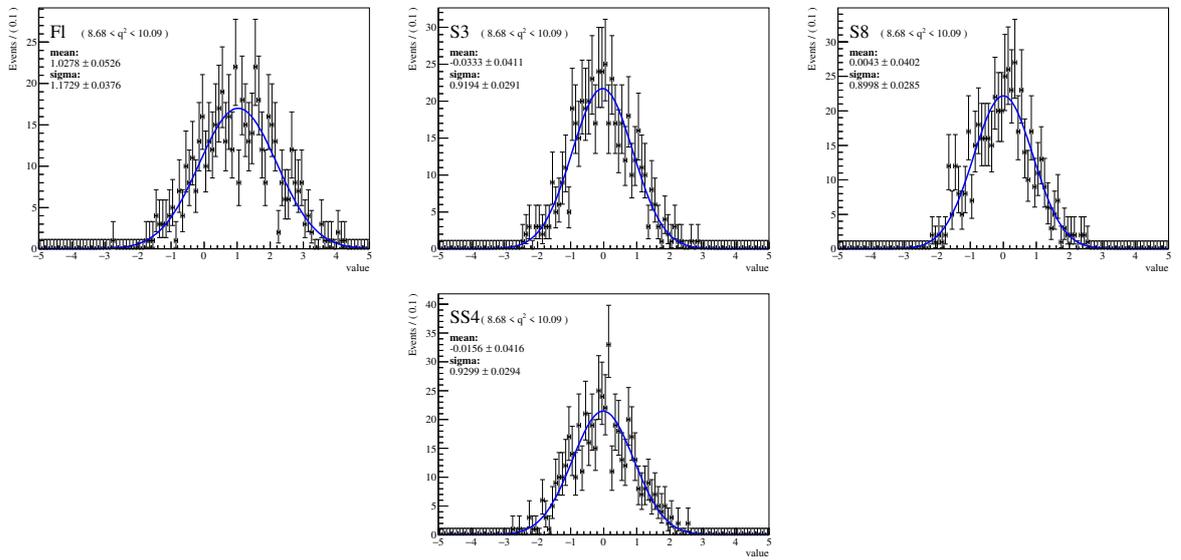
Fig. D.5: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the refrence $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 65 000 pseudoevents. In the fit to the pseudoexperiments, folding 4 is applied.

D.2   Realistic scale pseudoexperiments
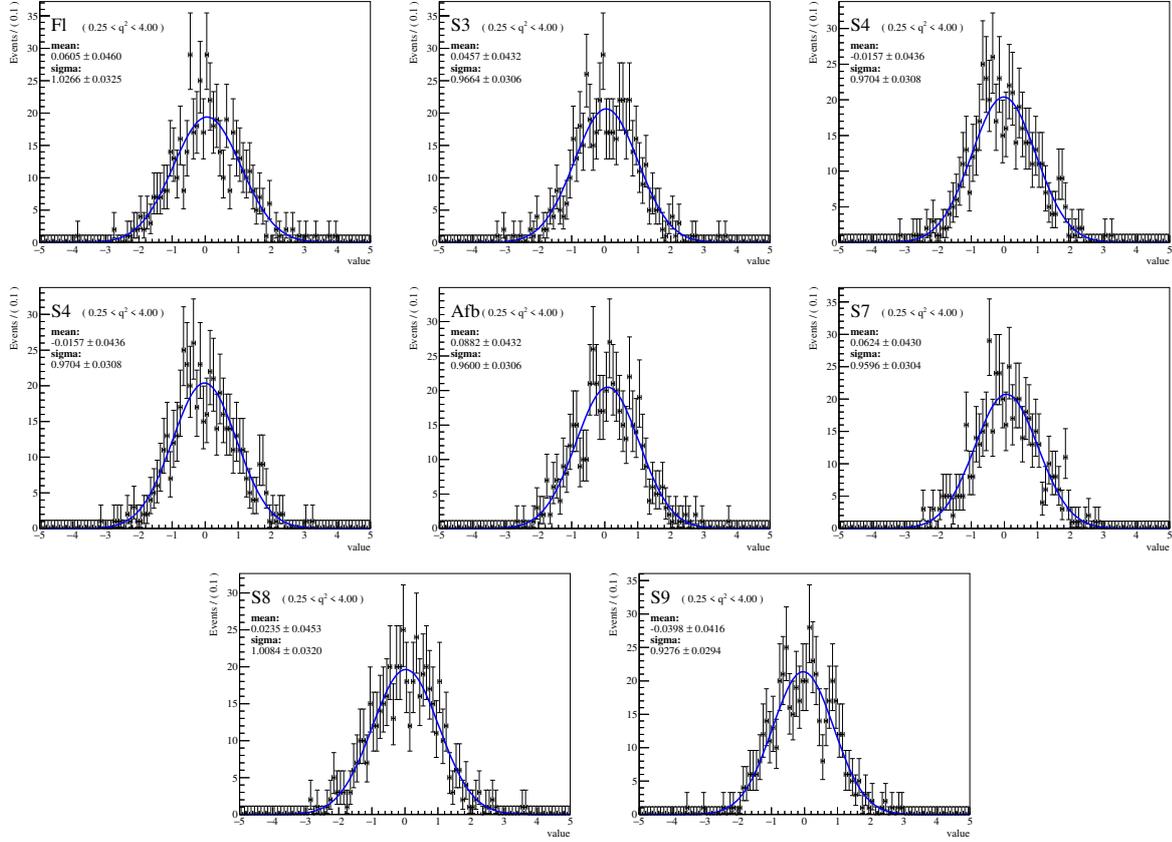
Fig. D.6: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. Each pseudoexperiment consists of 871 pseudoevents distributed among q^2 bins. Full angular fit is performed. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.
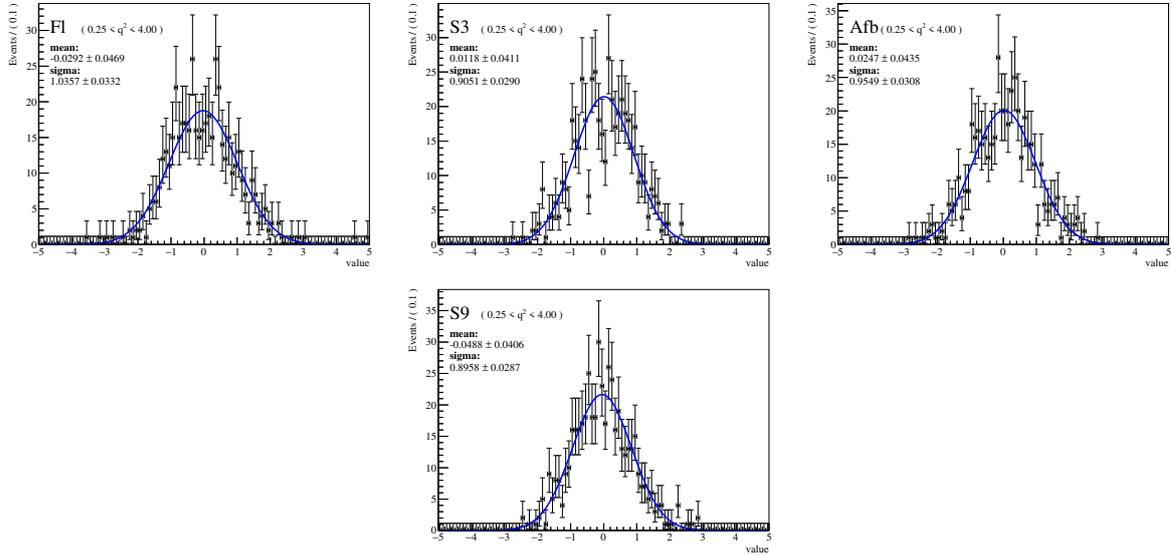
Fig. D.7: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. Each pseudoexperiment consists of 871 pseudoevents distributed among q^2 bins. In the fit to the pseudoexperiments, folding 0 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.
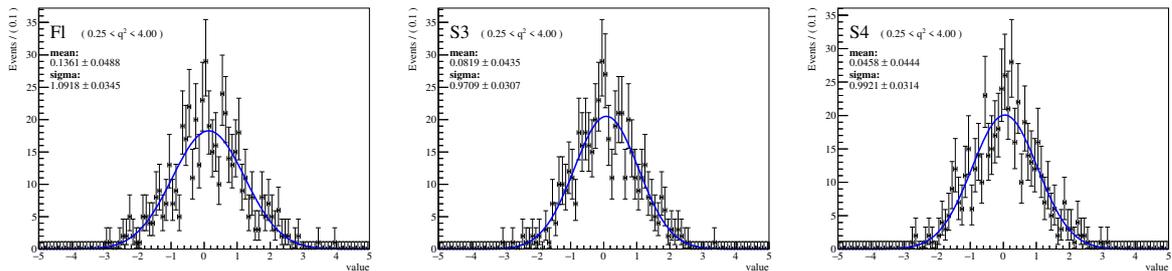

Fig. D.8: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the rare $B^+ \to K^{*+}J/\psi$ decay. Each pseudoexperiment consists of 871 pseudoevents distributed among q^2 bins. In the fit to the pseudoexperiments, folding 1 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.
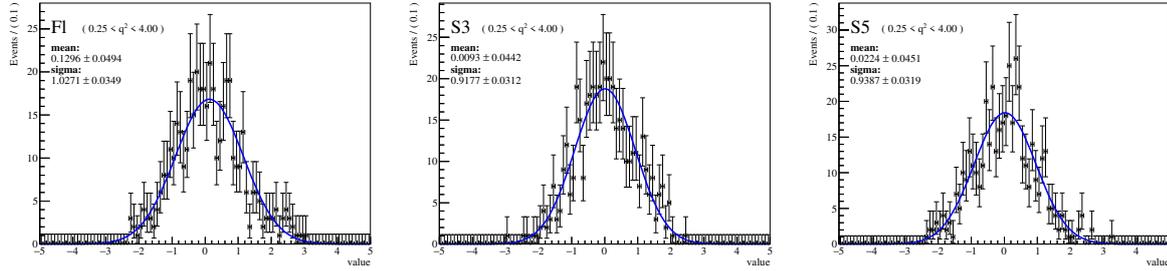
Fig. D.9: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. Each pseudoexperiment consists of 871 pseudoevents distributed among q^2 bins. In the fit to the pseudoexperiments, folding 2 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.


Fig. D.10: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. Each pseudoexperiment consists of 871 pseudoevents distributed among q^2 bins. In the fit to the pseudoexperiments, folding 3 is applied. Note the small range of the x-axis: seemingly large shift of the mean value is in the order of units of percent.


Fig. D.11: Pull distributions of the P-wave angular parameters. 500 pseudoexperimentsare generated, mimicking the rare $B^+ \to K^{*+}\mu^+\mu^-$ decay. Each pseudoexperiment consists of 871 pseudoevents distributed among q^2 bins. In the fit to the pseudoexperiments, folding 4 is applied.