Università degli Studi di Ferrara

Institute of Nuclear Physics PAN

University of Ferrara

Doctoral dissertation

Józef Borsuk

Searches for Beyond Standard Model effects in rare B decays.

Supervisors:
dr hab. Marcin Chrzaszcz, prof. IFJ PAN
prof. Massimiliano Fiorini

July 2024

Acknowledgments

I would like to express my gratitude to my supervisor dr hab. Marcin Chrzaszcz, who guided me through this challenging time and helped me to achieve my goals.

I would also like to acknowledge the National Science Center for the Preludium grant.

Finally, I would like to thank my family, especially my wife Ewa, for the constant support and care.

Summary

The aim of this dissertation was to analyze a decay of the beauty meson B^0, which is made of a quark-antiquark pair (up quark and b antiquark). It is an unstable particle, which means it decays very rapidly and this decay can proceed in very different ways. In one of the scenarios, B^0 decays to K^{*0} (down quark and strange antiquark pair, decaying to kaon K^+ and pion π^-) and a muon-antimuon pair $\mu^-\mu^+$. The angular analysis of the final state particles, which is a study of how particles propagate in a detector, can tell a lot about the B^0 decay. Along with the decay angles, the B^0 meson decay is described by a number of measurable quantities (observables) that can be compared to the SM predictions. The quantities of those observables could be modified by the presence of new, unknown particles. The analysis is based on the data collected by the LHCb experiment in 2016, 2017 and 2018 (Run2), when protons collided at the center of mass energy of 13 TeV (teraelectronovolts).

This dissertation presents an analysis of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay and is focused on the higher invariant mass range of the $K\pi$ system (1330-1530 MeV/c^2), where heavy resonant contributions could be found. The first four chapters are devoted to the introduction, the theoretical description of the Standard Model, and the description of the LHCb detector and the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay. This is followed by the analysis, whose first stage is proper data selection, because the signal candidates are also accompanied by background events, such as those with misidentified particles. The data selection is described in chapter five, which is followed by the proper reweighting of the candidates to ensure the compatibility between the data and Monte Carlo simulated samples. In order to account for a distortion in kinematic distributions caused by the data selection, the angular acceptance was corrected by weighting the candidates with the inverse of the five dimensional efficiency parametrization calculated with the simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ sample. Chapters eight and nine are devoted to the branching fraction calculations and the angular analysis of the B^0 decay. In order to extract the angular observables the method of moments was used. In chapter ten six different sources of systematic uncertainties are described. Chapters eleven and twelve show the results of the analysis where the latter compare them with previous results from Run1 analysis.

The results of the branching fraction measurements and angular moments are compatible with the previous Run1 results and have lower statistical uncertainty. Unfortunately, the level of systematic uncertainties is relatively high. The main reason for that is a low number of simulated samples of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay used in the acceptance correction calculations. New results make it possible to estimate a contribution from the resonant $K_2^*(1430)$ component, which can appear in the decay. Unfortunately, the level of the uncertainty is still high and the result is inconclusive, as more data are needed.

Streszczenie

Celem niniejszej roprawy doktorskiej jest badanie tzw. mezonu pięknego B^0, który składa się z pary kwark-antykwark (kwarka dolnego i antykwarka pięknego). Jest to cząstka nietrwała, czyli rozpadająca się bardzo szybko, a jej rozpad może przebiegać na różne sposoby. W jednym ze scenariuszy po jej rozpadzie pozostaje cząstka K^*(K-star - złożona z kwarka dolnego i antykwarka dziwnego, który z kolei rozada się na kaon K^+ i pion π^-) i para mion-antymion $\mu^-\mu^+$. Wiele o samym rozpadzie może powiedzieć analiza kątowa produktów rozpadu, czyli badanie w jaki sposób cząstki po rozpadzie mezonu B^0 propagują się w detektorze. Oprócz samych kątów rozpad jest opisywany przez pewne wielkości mierzalne (obserwable), których wartości można porównać z przewidywaniami Modelu Standardowego. Te właśnie wielkości mogą być modyfikowane poprzez obecność nowych, nieznanych do tej pory cząstek.

Analiza jest oparta o dane zebrane przez eksperyment LHCb w 2016, 2017 i 2018 roku (tzw. Run2), gdzie zderzano protony o energii 13 TeV (teraelektronowoltów) w układzie środka masy.

Niniejsza roprawa doktorska poświęcona analizie rozpadu $B^0 \to K^+\pi^-\mu^+\mu^-$ jest skoncentrowana na wyższej wartości masy niezmienniczej układu $K\pi$ (1330-1530 MeV/c^2), gdzie mogą zostać wykryte przyczynki od cięższych rezonansów. Pierwsze cztery rozdziały poświęcone są wstępowi, teoretycznemu opisowi Modelu Standardowego i opisom eksperymentu LHCb wraz z rozpadem $B^0 \to K^+\pi^-\mu^+\mu^-$. Zaraz potem przedstawiona została analiza. Pierwszy jej etap jest poświęcony selekcji danych, ponieważ cząstkom pochodzącym od sygnału towarzyszą także cząstki tła na przykład takie, gdzie cząstki zostały błędnie zidentyfikowane. Selekcja danych jest przedstawiona w rozdziale piątym, a następnie, rozdziale szóstym, została uzyskana kompatybilność między danymi eksperymentalnymi a symulowanymi. W celu wzięcia pod uwagę wpływu selekcji przypadków na rozkłady kinematyczne, została obliczona korekcja akceptancji kątowej poprzez przeważenie przypadków odwrotnością pięciowymiarowej parametryzacji wyznaczonej przy użyciu próbki symulowanych danych $B^0 \to K^+\pi^-\mu^+\mu^-$. Rozdział ósmy i dziewiąty są poświęcone obliczeniom różniczkowego współczynnika rozgałęzienia i analizie kątowej rozpadu mezonu B^0. Do wyznaczenia współczyników kątowych została użyta metoda momentów. W rozdziale dziesiątym zostało opisanych sześć różnych źródeł niepewności systematycznych. Rozdziały jedenaście i dwanaście poświęcone zostały na przedstawienie końcowych wyników, gdzie w dwunastym rozdziale uzyskane wyniki zostały porównane z poprzednimi z analizy danych Run1.

Uzyskane wyniki są kompatybilne z poprzednimi i mają mniejszą niepewność statystyczną. Niestety poziom niepewności systematycznych jest relatywnie wysoki. Powodem jest mała ilość przypadków w symulowanej próbce $B^0 \to K^+\pi^-\mu^+\mu^-$ użytej do obliczenia korekcji akceptancji kątowej. Nowe wyniki analizy kątowej pozwoliły także na dokładniejsze wyznaczenie przyczynku od rezonansu $K_2^*(1430)$, który może pojawić się w badanym rozpadzie. Jednak z powodu wysokiej niepewności pomiarowej wynik w dalszym ciągu nie jest konkluzywny, gdyż potrzebna jest większa liczba danych.

Contents

Chapter 1

Introduction

The Standard Model (SM) [1] of particle physics describes almost all phenomena in subatomic scale and is one of the most extraordinary achievements in science of the last century. This model describes three of the four fundamental interactions: strong interaction and weak with electromagnetic interactions as a unified electroweak interaction. The SM classifies all known elementary particles in to: 12 fermions and 13 gauge bosons, where the latter comprises a well known photon, taking part in electromagnetic interactions, $W^{\pm}$ and Z^0 bosons, which partake in weak interactions, 8 gluons which bind quarks within hadrons, and the so-called Higgs boson, which is related to the phenomena of giving mass to particles. As a result of the Higgs boson discovery [2], by the ATLAS and CMS Collaborations, the existence of all fundamental Standard Model particles has been proven. Despite its great success, the SM does not explain several phenomena. First of all, it does not include gravitational interactions. Secondly, the SM does not agree with several known experimental observations. Almost all of our Universe is built of matter, but according to the Big Bang theory matter and antimatter should be produced equally. CP violation, which occurs in the SM, alone cannot explain that amount of matter-antimatter asymmetry. There must exist some new mechanism that is responsible for that asymmetry but is not included in the SM. Many experiments have proven that most of our matter in the Universe interacts only gravitationally but not electromagnetically. For this reason this matter is called dark matter. Still, there is no evidence for such particles. Additionally, neutrinos were thought to be massless particles in the SM, but experimental results of neutrino oscillations have shown that they actually have a very tiny mass [3], so a new component generating neutrino masses has to be added to the framework of the SM.

Bearing these problems in mind, we need to search for a new theory extending the Standard Model, the so-called Beyond the Standard Model (BSM). In many BSM theories new particles are introduced to explain the discrepancies in measurements with the SM predictions. The reason why those new particles have not yet been observed is the fact that they can be so heavy that the present LHC experiments cannot detect them directly. However, they can be studied indirectly, as they can contribute virtually at lower energy scales via loops. The main role in searching for new phenomena is played by a high enough number of collected decays. This helps to perform more and more precise measurements.

That is why the infrastructure of each experiment at the LHC is being constantly improved in order to increase the luminosity and detection efficiency. More efficient data storage is also very important for analyses. One of the many ways to search for BSM effects is to analyze rare decays such as $B^0 \to K^{*0}\mu^+\mu^-$ [4], where $K^{*0} \to K^+\pi^{-1}$. This decay is a flavour changing neutral current (FCNC) process, where a fermion changes its flavour without changing its electric charge, which in the SM can proceed through electroweak penguin or box diagrams. In the BSM theories, angular observables and branching fraction can be modified by new, yet undiscovered particles, even if the NP scale is inaccessible for direct searches. The $B^0 \to K^+\pi^-\mu^+\mu^-$ decay provides a sufficient number of observables for an analysis of this kind and the main goal of this dissertation is to investigate the aforementioned B^0 meson decay.

The structure of this dissertation is as follows: after the introduction given above, the second chapter describes the main concepts of the particle physics theory, where mostly weak interactions are presented in terms of the Effective Field Theory, and the main ideas of BSM phenomena are outlined. The third chapter presents the angular description of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay. Chapter four contains essential information about the Large Hadron Collider and the LHCb detector.In chapter five information about the collected data samples and candidates selection, where background candidates are subtracted, is presented. Chapter six describes the methods used to obtain the compatibility between real data from the LHCb experiment and simulated Monte Carlo (MC) samples. Chapter seven deals with acceptance correction which accounts for the distortions caused by performed data selections. Chapters eight and nine provide the angular analysis and branching fraction calculations of the B^0 meson decay. Chapter ten is dedicated to describing the calculations of systematic uncertainties. Chapter eleven and twelve contain all results obtained in this analysis and is followed by chapter thirteen where the final conclusions are presented.

[1]Contribution of charge conjugate processes is implied to all decays, unless otherwise stated.

Chapter 2

Standard Model of Particle Physics

The Standard Model of Particle Physics was formulated in the 1970s. It is a quantum field theory that describes all matter that we know built from several elementary particles of spin 1/2 called fermions: 6 quarks and 6 leptons. The SM also describes fundamental interactions: electromagnetic, weak and strong, but it does not incorporate gravitational interactions. In the quantum theory an interaction is described as an exchange of gauge bosons which have integer spin. The description of fermions and bosons is provided in the next sections.

2.1 Building blocks of the Standard Model

The SM is built up from three main elements:

- a gauge group,
- fermion fields
- spontaneous symmetry breaking.

A gauge group incorporates two parts: $\mathrm{SU(3)_C}$, which represents the colour group describing strong interactions, and $\mathrm{SU(2)_W \times U(1)_Y}$, which is the electroweak group describing electromagnetic and weak interactions. The SM particle sector contains a number of fields: left-handed quark doublet ($\mathrm{Q_L}$), right-handed up-type quark ($\mathrm{U_L}$), right-handed down-type quark ($\mathrm{D_R}$), left-handed lepton doublet ($\mathrm{L_L}$), and the right-handed charged lepton ($\mathrm{E_R}$) with the following representations:

$$Q_L^i(3,2)_{1/6}, \quad U_R^i(3,1)_{2/3}, \quad D_R^i(3,1)_{-1/3}, \quad L_L^i(1,2)_{-1/2}, \quad E_R^i(1,1)_{-1}. \tag{2.1}$$

where i iterates over three generations (i=1,2,3). In addition to fermionic fields there is one additional scalar Higgs field ϕ with the representation $(1,2)_{1/2}$. The scalar Higgs field generates masses of particles through spontaneous symmetry breaking and has non-zero ground state. The excitation of the Higgs field is represented by the Higgs boson of a mass equal to $\sqrt{2\lambda\nu} \approx 125$ GeV/c^2, where $\lambda > 0$ represents the coupling of a four

particle vertex and ν is the vacuum expectation value, equal approximately to 246 GeV. This addition to the theory helps to implement spontaneous symmetry breaking of the SM group to $\mathrm{SU(3)_C \times U(1)_Q}$, where the charge Q, as a sum of hypercharge and the third component of isospin, is conserved.

2.2 Matter particles

The group of fermions is divided into two parts: quarks and leptons, and also three generations as in Table 2.1.
Quarks come in six types, the so-called flavours: up(u), down(d), charm(c), strange(s), top(t), bottom(b) and each of them has fractional charge. Those unusual names come from the process of model development. Quarks u and d are described in terms of isospin symmetry, where both have isospin 1/2 with eigenvalues $\pm 1/2$. Quark s is a part of strange particles, which were discovered in cosmic radiation. They are characterized by strange behaviour: as observations have shown, they decay very slowly through weak interactions. To explain this phenomenon, a new quantum number S(strangeness) was proposed, which is conserved in strong interactions. Later on, a c quark was discovered in produced J/ψ mesons. The charm's name match the name of the strange quark. The names of top and bottom quarks correspond to the names of up and down quarks. All observations indicate that quarks cannot be separated. Strong interactions have this special property, according to which quarks always appear in specific combinations: barions (qqq) and mesons ($\mathrm{q\bar{q}}$). There are also other possibilities such as recently discovered tetraquarks and pentaquarks [5, 6].

The second type of fermions are leptons. Three of them are negatively charged: electron, muon(μ) and lepton tau (τ), and for each charged lepton there is one neutrino partner with the same flavour (ν_e, ν_μ, ν_τ). Both muons and taus are unstable particles, but they decay differently. While muons decay to electrons and neutrinos, taus decay mostly to hadronic final state with additional neutrinos.

	I	II	III	$Q/\lvert e\rvert$
Quarks	u	c	t	+2/3
	d	s	b	-1/3
Leptons	e	μ	τ	-1
	ν_e	ν_μ	ν_τ	0

Table 2.1: List of the SM fermions, quarks and leptons divided into three generations.

2.3 Interactions

There are four fundamental types of interactions and each can be represented by a particle responsible for a given interaction (Table 2.2).

Electromagnetic force describes interactions of charged particles as an exchange of photons. Strong interactions hold quarks within hadrons as well as protons and neutrons within nucleus. The most classic example of a weak interaction is a radioactive β decay, where in atomic nucleus a neutron decays to a proton, electron and antineutrino. Weak interactions consist in the exchange of massive vector bosons $W^{\pm}$ and Z^0, which are almost hundred times as massive as a proton.

All massive particles interact gravitationally and this force is dominant in cosmic scale, but it is the weakest of all forces. It has been postulated that gravity can be described as an exchange of a spin 2 graviton, but the incorporation of gravity into the quantum description of the world is still an open subject. The last part of the SM is the Higgs boson, the only scalar particle connected to the process of giving mass to particles.

For many years scientists have debated that all elementary interactions are only a low energy manifestation of a single unified field. In the same way in high energies a specific symmetry between electromagnetic and weak interactions appear and they can be treated as one electroweak force. In even higher energies it is theorized that an electroweak interaction and strong interaction unify into singular electronuclear force, where interactions are described by one gauge symmetry with a single coupling constant. This model is known under the name of the Grand Unified Theory (GUT‘). The next step towards the unified theory is adding gravity to GUT's framework, having in result the Theory of Everything fully explaining the fundamentals of our Universe.

Force	Particle	Spin-parity
Electromagnetic	γ	1^-
Weak	$W^{\pm}$ Z^0	$1^-, 1^+$
Strong	g	1^-
gravitation	*graviton*	2^+
Scalar particle		
Higgs boson	H^0	0^+

Table 2.2: List of spin-1 gauge bosons corresponding to each fundamental type of interaction along with hipothetical spin-2 graviton and scalar particle Higgs boson.

2.4 Electromagnetic interactions

All charged particles interact with each other owing to electromagnetic force, which manifests itself as an exchange of virtual particles called photons. The photon is a massless, vector boson with spin-parity $J^P = 1^-$. The theory of quantum electrodynamics (QED) describes an interaction between charged matter particles and photons combining both quantum mechanics and the theory of special relativity.

2.4.1 Gauge theory

The QED is a gauge theory based on the symmetry group U(1) and its Lagrangian is given by:

$$\mathcal{L}_{gauge}^{QED} = \overline{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.2}$$

where ψ is the fermionic field incorporating all relevant representations from Eq. 2.1, γ^{μ} are gamma matrices and m is the mass of a particle. Lagrangian's invariance under U(1) is ensured using the covariant derivative:

$$D_{\mu} = \partial_{\mu} + ieA_{\mu}, \tag{2.3}$$

where A_{μ} is an electromagnetic four-potential. The last part of Eq. 2.2 contains electromagnetic strength tensor which prevents the gauge invariance from breaking:

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}. \tag{2.4}$$

2.5 Strong interactions

Quantum field theory, which describes interactions between quarks and gluons is called Quantum Chromodynamics (QCD). It is a non-abelian gauge theory, with a symmetry group SU(3) (special, unitary group of degree 3). Apart from charge, quarks have an additional quantum number called colour (conventionally red, green, blue) and they interact with each other thanks to the exchange of 8 massless gluons, which carry both colour and anticolour charge. Strong interaction has two very important properties:

- asymptotic freedom: in comparison to any other interaction, strong force gets weaker at short distances between particles, and when a distance increases, coupling increases as well,
- confinement: is a consequence of asymptotic freedom, where quarks are always combined in hadrons, therefore single quarks do not exist. In high energy interactions, when two quarks are being separated, it is more likely that a new pair quark-antiquark would be created. Isolation of particles with colour charge is impossible and only "colourless" hadrons exist.

2.5.1 Gauge theory

QCD Lagrangian is given by:

$$\mathcal{L}_{gauge}^{QCD} = \overline{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}\sum_{a=1}^{8} G_{\mu\nu}^{a}G_{a}^{\mu\nu}, \tag{2.5}$$

with the corresponding covariance derivative:

$$D_{\mu} = \partial_{\mu} - ig_{s}\sum_{a=1}^{8} L_{a}G_{\mu}^{a}, \tag{2.6}$$

where $G^a_{\mu\nu}$ is the gluon's field strength tensor, G^a_μ are eight gluon fields, L_a terms are the generators of the $SU(3)_C$ (the 3 x 3 Gell-Mann matrices) and g_s is a coupling constant. Index i iterates over three colour charges (i=1,2,3) and a iterates over eight gluon states (a=1...8).

2.6 Weak interactions

There are two types of weak interactions: those where involved particles change their flavour and those where the flavour is unaffected. In the first scenario a particle, for example a down quark, can transform to an up quark with the emission of a vector boson particle W^- with a mass of 80.4 GeV/c^2. This process is called flavour changing charged current and can occur also in reverse direction, so in a more general way we can write $q \to W^\pm q'$, where also an anti-W boson is presented. In the scenario where in a weak interaction the flavour of particles does not change, particles exchange the neutral vector boson Z^0 which has a mass of around 91.2 GeV/c^2. Owing to the fact that the force carriers $W^\pm$ and Z^0 have very large masses, their lifetime is very short, i.e. under 10^{-24} seconds. In contrast to a photon and gluon, $W^\pm$ and Z^0 obtain their masses through Higgs mechanism.

2.6.1 CMK matrix

The strength of flavour changing charged currents is governed by the corresponding matrix element $V_{qq'}$ of the Cabbibo, Kobayashi, Maskawa matrix (CKM matrix) [7]. It is a 3x3 complex and unitary matrix which contains four physical parameters: three real angles and one phase:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12} & s_{12}e^{u\delta} & 0 \\ -s_{12}e^{-i\delta} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23}e^{i\delta} \\ 0 & -s_{23}e^{-i\delta} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \tag{2.7}$$

where $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$ (i,j=1,2,3 and i≠j), θ_{ij} are the three rotation angles and δ is a complex phase. This complex phase is the only source of CP-violation in the SM. In CP transformation operation C stands for charge conjugation which inverts a particle's charge, and P is a parity operator changing the sign of spatial dimension in a vector. There are no theoretical predictions on the elements of the CKM matrix and they have to be measured experimentally. In this context, it is very useful to present the CKM matrix in terms of the Wolfenstein parametrisation [8], where it is expanded in powers of the small parameter λ, also known as the Cabbibo parameter $\sin\theta_{12} \approx 0.22$, where angle θ_{12} is called the Cabbibo angle:

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{2.8}$$

In this form, the CKM matrix shows its hierarchical structure. Three real parameters A, λ, ρ and one imaginary part η can be defined as follows:

$$\lambda^2 = \frac{|V_{us}|^2}{|V_{ud}|^2 + |V_{us}|^2}, A^2\lambda^4 = \frac{|V_{cb}|^2}{|V_{ud}|^2 + |V_{us}|^2}, \bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \tag{2.9}$$

where $\bar{\rho} = \rho(1 - \frac{\lambda^2}{2})$ and $\bar{\eta} = \eta(1 - \frac{\lambda^2}{2})$. Having those four parameters one can fully determine the CKM matrix in the SM. The quantities $\bar{\rho}$ and $\bar{\eta}$ determine the apex of the unitarity triangle (Figure 2.1), which is just a graphical representation of unitarity relation:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.10}$$

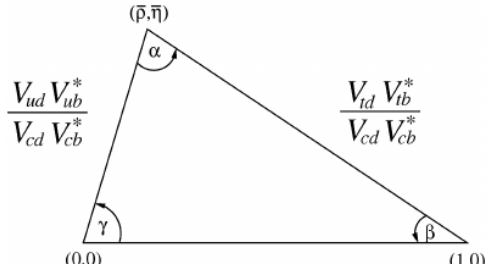


Figure 2.1: The CKM unitarity triangle.

2.6.2 Gauge theory

The weak interaction is based on the $SU(2)_W$ symmetry, to which only left-handed fermions ψ_L are sensitive. In $SU(2)_W$ the left-handed fermions are represented as doublets with weak isospin I =1/2 and right-handed fermions are represented as singlets with weak isospin I=0. The Lagrangian of weak interactions is given by:

$$\mathcal{L}_{gauge}^W = \overline{\psi} i \not{D} \psi + g \sum_{i=1}^{3} \overline{\psi} \gamma^\mu \sigma_i \psi W_\mu^i - \frac{1}{4} \sum_{i=1}^{3} W_{\mu\nu}^i W_i^{\mu\nu}, \tag{2.11}$$

where g is a coupling constant, σ_i terms are the $SU(2)_W$ generators (the Pauli matrices), W_μ^i are the three gauge fields of $SU(2)_W$ and $\not{D} = \gamma^\mu D_\mu$ is called Feynman slash notation. The last term in Eq.2.11 describes the dynamics of weak fields, where $W_{\mu\nu}^i$ is described by:

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk} W_\mu^j W_\nu^k, \tag{2.12}$$

where i iterates over three generations of fermions and ϵ^{ijk} are the $SU(2)_W$ structure constants. It should be mentioned that there is no mass term in Eq.2.11, which would not be invariant under the $SU(2)_W$ symmetry otherwise. This means that in theory fermions are massless. We know from experiments, however, that they actually have masses, and they are given to particles through Higgs mechanism.

2.6.3 Electroweak unification

Thanks to Glashow, Weinberg and Salam, new theories of electromagnetism and weak interaction have been proposed [9]. Those two theories have been unified to a single one called the electroweak theory. It is a symmetry group $SU(2)_W \times U(1)_Y$, with the weak isospin I and weak hipercharge $Y = 2(Q - I_3)$, where Q is the electric charge and I_3 is the third component of weak isospin. To have electroweak Lagrangian invariant under $SU(2)_W \times U(1)_Y$, the following covariant derivative needs to be used:

$$D_\mu = \partial_\mu + i\frac{g}{2}\sum_{i=1}^{3}\sigma_i W_\mu^i + iY\frac{g'}{2}B_\mu. \tag{2.13}$$

which is built from weak hypercharge field B and weak isospin fields W^1, W^2 and W^3. Factors g and g' are the coupling constants of the $SU(2)_W$ and $U(1)_Y$ symmetries, respectively. The covariant derivative given in Eq.2.13 is valid only for left-handed fermions, because for right-handed fermions the weak part has to be omitted. Given all above, the electroweak Lagrangian is described by:

$$\mathcal{L}_{gauge}^{EW} = \overline{\psi} i \not{D} \psi - \frac{1}{4}\sum_{i=1}^{3} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.14}$$

where $W_i^{\mu\nu}$ and $B_{\mu\nu}$ are the field strength tensors of weak isospin and weak hypercharge fields, respectively. In this description of electroweak interactions, the bosons B and W^3 combine into two neutral particles: a photon and Z^0, while the bosons W^1 and W^2 produce charged bosons $W^\pm$:

$$\begin{aligned} \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} &= \begin{pmatrix} cos\theta_W & sin\theta_W \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \\ W^\pm &= \frac{1}{\sqrt{2}}(W^1 \mp iW^2), \end{aligned} \tag{2.15}$$

where θ_W is the Weinberg mixing angle, which shows the difference between the masses of the $W^\pm$ and Z^0 bosons:

$$cos\theta_W = \frac{m_W}{m_Z}. \tag{2.16}$$

2.7 Flavour physics Lagrangian

Combining all elementary particle interactions: electromagnetic, weak and strong, apart from the gravitational one, the fermion current is given by:

$$\mathcal{L}_{gauge} = \bar{\psi} i \not{D} \psi - \frac{1}{4}\sum_{a=1}^{8} G_{\mu\nu}^a G_a^{\mu\nu} - \frac{1}{4}\sum_{i=1}^{3} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{2.17}$$

where the three components $G^{a}_{\mu\nu}$, $W^{i}_{\mu\nu}$ and $B_{\mu\nu}$ represent the field strength tensors of gluon, weak isospin and weak hypercharge fields, respectively. In the the flavour sector the covariant derivative is given by:

$$D_\mu = \partial_\mu + ig_s \sum_{a=1}^{8} G^a_\mu L^a + ig \sum_{b=1}^{3} W^b_\mu T^b + ig' B_\mu Y, \tag{2.18}$$

where G^{a}_{μ} are the eight gluon fields, W^{b}_{μ} the three bosons of the weak interaction and B_μ is a single hypercharge boson, all of which are also massless. The terms g_s , g, and g' correspond to gauge coupling constants.

In Eq. 2.2 and in Eq. 2.5 mass terms are responsible for breaking $U(1)_Y$ and $SU(3)_C$ gauge invariance, respectively. Just like in Eq. 2.11 and Eq. 2.14, mass terms should be absent, which provides an explanation why in theory fermions are initially massless.

2.8 Effective Field Theory in particle physics

As we already know weak, interactions are governed by an exchange of massive gauge bosons. As those bosons are very massive particles, their lifetime is very short. This means that the distances over which they can interact are also very small. Effective theories can be used to model processes at a given energy scale μ relevant to the particles involved in the process. The best example of this approach is the Fermi's effective theory of the β decay. In 1934 Enrico Fermi proposed to describe the $n \to p e^- \bar{\nu}$ interaction as effectively point-like, because of its short range, using the Fermi constant $G_F = g^2/m_W^2$, where g represents the weak charge and m_W is the mass of the W boson. According to this theory, weak interactions can be described by a matrix element defined as a product of two currents:

$$\langle f|\mathcal{H}_{eff}|i\rangle = G_F \sum_{a=1}^{3} J^a_{bar.} J^a_{lep.} \approx G_F(\bar{\psi}_p \gamma_a P_L \psi_n)(\bar{\psi}_e \gamma_a P_L \psi_\nu). \tag{2.19}$$

With this approach in mind, we can choose the energy scale ($\mu \approx m_W$) at which the process takes place. In other words, we integrate out all contributions from the particles of masses greater than the W boson mass.

This approach can also be applied to weak decays of heavy particles involving b quarks. In this scenario we want to focus on processes active at $\mu \approx m_b$ scale. It is also important that for example in flavour changing neutral currents QCD corrections have to be taken into account. In the case of the $b \to s\ell^+\ell^-$ decays an effective Hamiltonian is given by:

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} \frac{\alpha_s}{4\pi} \sum_{i=1}^{10,S,P} \lambda_i C_i^{(\prime)}(\mu) \mathcal{O}_i^{(\prime)}(\mu), \tag{2.20}$$

where λ_i is the relevant combination of CKM matrix elements, C_i are the so-called Wilson Coefficients (WC), which are the effective couplings and $\mathcal{O}_i$ are the effective operators.

The prime sign indicates right-handed contributions. The sum runs over all relevant operators, which are: $\mathcal{O}_{1,2}$ - tree-level, $\mathcal{O}_{3,..,6,8}$ - strong penguin, $\mathcal{O}_7$ - photon penguin, $\mathcal{O}_{9,10}$ - electroweak penguin and $\mathcal{O}_{S,P}$ - scalar, pseudoscalar contribution. Wilson coefficients are complex numbers in which high energy, short distance contribution is encoded. The $\mathcal{O}_i$, called local operators, are matrix elements describing physics acting at lower ranges. The renormalization scale μ separates the short and long distance corrections.

Processes of particle decays are described by the matrix element of a given process $\langle f|\mathcal{H}_{\text{eff}}|i\rangle$, where the Hamiltonian $\mathcal{H}_{eff}$ is defined by the theory Lagrangian. In this so-called Operator Product Expansion formalism, Wilson coefficients contain large energy scales contributions, so the effects from New Physics particles, which are expected to be heavy, would change the values of Wilson coefficients. This means that any relevant deviation from SM values of WCs would indicate some kind of NP effect.

2.9 FCNCs and GIM mechanism

As it was mentioned in Section 2.6, neutral currents do not change a particle's flavour. In fact, in the SM flavour changing neutral currents are forbidden at the tree level by the so-called GIM mechanism [10]. This concept, named after Glashow, Iliopoulos, and Maiani, tells us that, at the tree level, there are no processes between quarks with different flavour and the same charge. In neutral currents, there is no difference between up type quarks matrix and down quark type matrix. The FCNC processes are, however, allowed in higher orders in perturbation theory, when two charged currents are combined constructing a loop diagram, called penguin or box diagram. This is possible thanks to a difference between masses of up and down type quarks. Figure 2.2 shows an example of FCNC processes of $b \to s\ell^+\ell^-$, where ℓ is a lepton. Even if an NP particle is very heavy, it can enter the loop, which can be seen as a difference in the values of the corresponding WC's. The only thing one needs to do is to compare the measured WCs with their SM predictions.

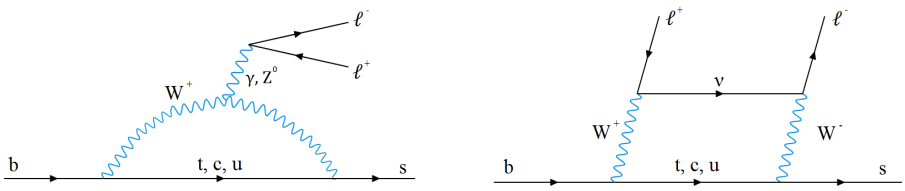


Figure 2.2: Feynman diagrams of FCNC processes of $b \to s\ell^+\ell^-$ in form of loops, electroweak penguin (left) and box diagram (right).

The effective Hamiltonian (Eq. 2.20) relevant for processes shown in Fig. 2.2 is defined with $\lambda_i = V_{tb}V_{ts}^*$ and with the following operators:

$$\mathcal{O}_7^{(\prime)} = \frac{m_b}{e}\bar{s}\sigma^{\mu\nu}P_{R(L)}bF_{\mu\nu}, \qquad \mathcal{O}_9^{(\prime)} = \bar{s}\gamma^\mu P_{L(R)}b\bar{\ell}\gamma_\mu\ell, \qquad \mathcal{O}_{10}^{(\prime)} = \bar{s}\gamma^\mu P_{L(R)}b\bar{\ell}\gamma_\mu\gamma^5\ell. \quad (2.21)$$

The term $P_{L(R)}$ corresponds to the left and right handed chiral projections, respectively, in the form of $(1 \mp \gamma^5)/2$. The right-handed versions of the operator, marked with a prime sign, are suppressed in the SM by a factor of m_s/m_b and could be enhanced by yet unknown NP effects. The operator $\mathcal{O}_7$ describes the emission of a photon, which then produces leptonic final state $\ell^+\ell^-$. The vector operator $\mathcal{O}_9$, and the axial operator $\mathcal{O}_{10}$ contribute to interactions involving the $W^\pm$ and Z^0 bosons in loops.

2.10 Beyond the Standard Model

There is no doubt that the SM is an incredibly successful theory, but it is not without flaws. One of the most clear indicators of the SM not being the final theory is that it does not contain gravity. Gravitational interaction is expected to be relevant to particle physics at Planck scale (10^{19} GeV), which is much higher than the Fermi scale (~100 GeV) accessible for current experiments.

Another problem is the fact that according to experiments most of the matter in our Universe interacts only gravitationally, and it does not interact electromagnetically. This type of matter is called dark matter and the SM does not involve a dark matter candidate.

We know that almost all of our Universe is built of matter, whereas, according to the theory of the Big Bang, matter and antimatter should be produced equally. In theory, the SM has a mechanism which could explain this asymmetry- the CP violation, but the observed CP violation contribution is not enough to explain the current matter-antimatter asymmetry.

Moreover, in the SM neutrinos are massless particles, but scientists have observed neutrino oscillations [11], which could only be possible if they had mass. There are also several recent anomalies in Lepton Universality (LU) ratios [12] and angular observables [13], which could be explained as manifestations of new heavy particles, such as leptoquarks [14] or electroweak neutral, heavy boson Z' [15]. Leptoquarks are particles carrying both lepton and barion quantum number, which allow interactions between leptons and quarks to take place. Figure 2.3 shows how $b \to s\ell^+\ell^-$ processes would look like with those NP contributions.

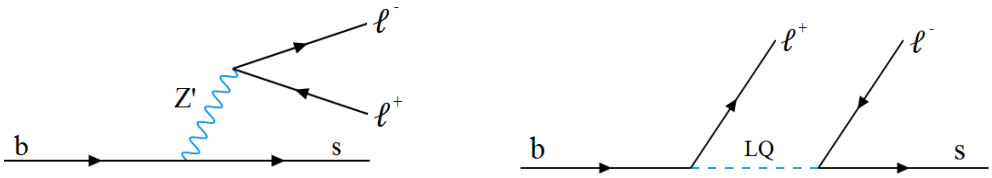


Figure 2.3: Possible NP contributions to $b \to s\ell^+\ell^-$ decays, such as electroweak heavy boson Z' (left), or leptoquark (right).

Taking into consideration these shortcomings it is clear that we need a new theory. However, based on the SM's success it makes much more sense to assume that the present form of the SM might be just a low energy effective version of some higher scale complete

theory. These BSM theories will inevitably have new degrees of freedom and the discovery of such particles will be a clearly visible signal of BSM physics or New Physics (NP). However, we have not yet observed such processes at high energy colliders, and perhaps NP particles are beyond their present energy reach. Luckily, there exists another method to infer their presence indirectly. In this method, we look at decay modes that are loop dominated and are suppressed or rare in the SM. Being loop dominated, higher degrees of freedom could contribute to them virtually and these effects could be sizable as the original amplitudes are itself suppressed in the SM.

Chapter 3

Characterization of $B^0 \to K^+\pi^-\mu^+\mu^-$ decays

The $B^0 \to K^{*0}(\to K^+\pi^-)\mu^+\mu^-$ final state (Fig.3.1) is described by the invariant mass squared of a dimuon system q^2, the three decay angles: $\cos(\theta_V)$, $\cos(\theta_\ell)$ and ϕ, and by the invariant mass of the $K^+\pi^-$ system [16]. In the B^0 rest frame hadronic and leptonic systems are back-to-back moving in opposite directions of $\hat{z}$ axis. Decay angles are defined as follows:

- θ_V - the angle between the direction of K^+ (K^-) and the direction of B^0 ($\bar{B}^0$) in K^{*0} rest frame,
- θ_ℓ - the angle between the direction of μ^+ (μ^-) and the direction opposite to that of the B^0 ($\bar{B}^0$) in the rest frame of the dimuon system,
- χ - the angle between planes constructed from the $K\pi$ system and dimuon pair in the B^0 ($\bar{B}^0$) rest frame.

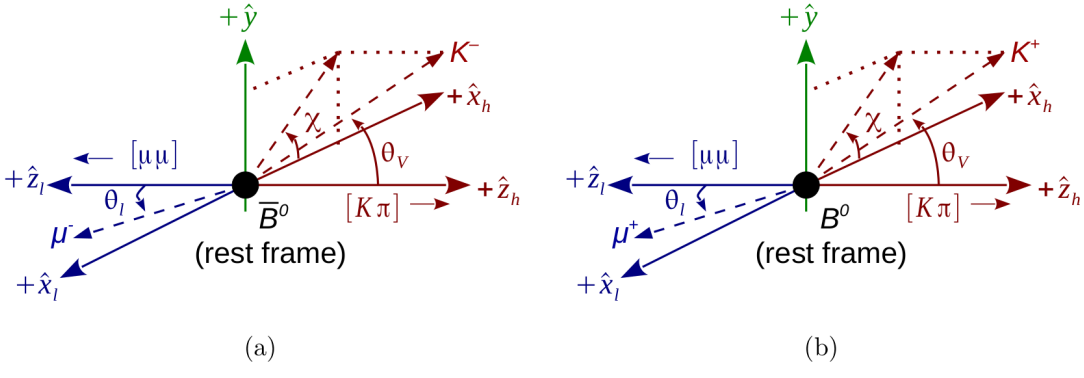


Figure 3.1: Angle conventions for the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay.

3.1 Angle definitions

The angle definitions are equivalent to the Gilman-Singleton definitions [17]. Figure 3.1 shows that there are two sets of coordinate axes: ($\hat{\mathrm{x}}_\mathrm{h}$,$\hat{\mathrm{y}}_\mathrm{h}$, $\hat{\mathrm{z}}_\mathrm{h}$) for the hadronic system [Kπ] and ($\hat{\mathrm{x}}_\ell$,$\hat{\mathrm{y}}_\ell$, $\hat{\mathrm{z}}_\ell$) for the leptonic system [$\mu\mu$]. To measure all three angles, particle momenta are first boosted to the $\mathrm{B}^0(\bar{\mathrm{B}}^0)$ rest frame.

In the case of θ_V and θ_ℓ the momenta are boosted to the [Kπ] and [$\mu\mu$] rest frames, respectively, and then the polar angles of the direction of the kaon and muon can be measured.

The dihedral angle $\chi = \phi_{[\mathrm{K}\pi]}\phi_{[\mu\mu]}$ is obtained by fixing the muon azimuthal angle $\phi_{[\mu\mu]}$ to zero in the leptonic helicity frame. Then the hadronic azimuthal angle in the hadronic helicity frame equals χ.

The explicit definitions of the $\bar{\mathrm{B}}^0$ decay angles are given below, where a superscript on any 3-vector denotes the reference frame. For $\bar{\mathrm{B}}^0 \to \mathrm{K}^-\pi^+\mu^-\mu^+$ we define the following momentum vectors:

$$\begin{aligned} \vec{P}_{\mu^+\mu^-} &= \vec{p}_{\mu^+} + \vec{p}_{\mu^-}, \\ \vec{Q}_{\mu^+\mu^-} &= \vec{p}_{\mu^+} - \vec{p}_{\mu^-}, \\ \vec{P}_{K^-\pi^+} &= \vec{p}_{K^-} + \vec{p}_{\pi^+}, \\ \vec{Q}_{K^-\pi^+} &= \vec{p}_{K^-} - \vec{p}_{\pi^+}. \end{aligned} \tag{3.1}$$

Using that, we can calculate $cos\theta_V$ and $\cos\theta_\ell$:

$$\begin{aligned} cos\theta_V &= -\frac{\vec{Q}^{K^{*0}}_{K^-\pi^+} \cdot \vec{P}^{K^{*0}}_{\mu^+\mu^-}}{|\vec{Q}^{K^{*0}}_{K^-\pi^+}||\vec{P}^{K^{*0}}_{\mu^+\mu^-}|}, \\ cos\theta_\ell &= -\frac{\vec{Q}^{\ell\ell}_{\mu^+\mu^-} \cdot \vec{P}^{\ell\ell}_{K^-\pi^+}}{|\vec{Q}^{\ell\ell}_{\mu^+\mu^-}||\vec{P}^{\ell\ell}_{K^-\pi^+|}}. \end{aligned} \tag{3.2}$$

For the dihedral angle we need to define vectors perpendicular to two decay planes ($[K\pi]$ and$[\mu\mu]$):

$$\begin{aligned} \vec{N}_{\mu^+\mu^-} &= \vec{P}^{\bar{B}^0}_{\mu^+\mu^-} \times \vec{Q}^{\bar{B}^0}_{\mu^+\mu^-}, \\ \vec{N}_{K^-\pi^+} &= \vec{P}^{\bar{B}^0}_{K^-\pi^+} \times \vec{Q}^{\bar{B}^0}_{K^-\pi^+}, \end{aligned} \tag{3.3}$$

which then allow us to define the angle χ:

$$cos\chi = -\frac{\vec{N}_{\mu^+\mu^-} \cdot \vec{N}_{K^-\pi^+}}{|\vec{N}_{\mu^+\mu^-}||\vec{N}_{K^-\pi^+}|} \quad sin\chi = (\frac{\vec{N}_{\mu^+\mu^-} \times \vec{N}_{K^-\pi^+}}{|\vec{N}_{\mu^+\mu^-}||\vec{N}_{K^-\pi^+}|})\frac{\vec{P}^{\bar{B}^0}_{\mu^+\mu^-}}{|\vec{P}^{\bar{B}^0}_{\mu^+\mu^-}|}. \tag{3.4}$$

The angles $\cos\theta_\mathrm{V}$ and $\cos\theta_\ell$ z take values from range [-1,1] and χ from range $[-\pi, \pi]$. The sign of χ is flipped during the CP conjugation.

CP conjugation does not change angular observables in terms of its corresponding amplitudes but the amplitudes are related as follows:

$$\mathcal{H}^\eta_\lambda(\delta_W, \delta_S) = \bar{\mathcal{H}}^{-\eta}_{-\lambda}(-\delta_W, \delta_S), \tag{3.5}$$

where λ is the helicity of the hadronic system, $\eta = 1(-1)$ denotes right-handed (left-handed) amplitudes, and $\delta_W(\delta_S)$ is the weak(strong) phase. To have the same observables for the B^0 and $\bar{B}^0$ decays, the sign of χ has to be flipped during CP conjugation.

3.1.1 Conventions for angular variables

Existing LHCb measurements of $B^0 \to K^{*0}\mu^+\mu^-$ provide different angle convention than that described in Sec. 3.1, namely θ_K, θ_ℓ and ϕ (e.g. [13]). Also, there are some theory papers using yet another convention [18]. Different conventions are listed in Table 3.1:

Angle	LHCb	Theory
θ_K	θ_V^{GS}	θ_V^{GS}
θ_ℓ	θ_ℓ^{GS}	$\pi - \theta_\ell^{GS}$
ϕ	$-\chi^{GS}$	χ^{GS}

Table 3.1: Translation between Gilman-Singleton [17], LHCb [13] and theory [18] angle conventions for the electroweak penguin B decay.

In order to match other LHCb analyses, throughout this dissertation the dihedral angle will be depicted as ϕ, which is equal to $-\chi$, and the hadronic system angle θ_V will be written as θ_K.

3.2 Differential decay rate

Summing up, the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay is described by the invariant mass squared of a dimuon system q^2, invariant mass of $K^+\pi^-$ system and three decay angles $\vec{\Omega} = (cos\theta_K, cos\theta_\ell, \phi)$. The differential decay rates of the B^0 and $\bar{B}^0$ decays are given by [19]:

$$\begin{aligned} \frac{d^4\Gamma(\bar{B^0} \to K^-\pi^+\mu^+\mu^-)}{dq^2 d\overrightarrow{\Omega}} &= \frac{9}{32\pi}\sum_i I_i(q^2) f_i(\overrightarrow{\Omega}), \\ \frac{d^4\bar{\Gamma}(B^0 \to K^+\pi^-\mu^+\mu^-)}{dq^2 d\overrightarrow{\Omega}} &= \frac{9}{32\pi}\sum_i \bar{I}_i(q^2) f_i(\overrightarrow{\Omega}), \end{aligned} \tag{3.6}$$

where I_i and $\bar{I}_i$ are the angular observables and $f_i(\overrightarrow{\Omega})$ are combinations of spherical harmonics. There are twelve I_i coefficients, which can be expressed in terms of six complex amplitudes: $A_\parallel^{L,R}$, $A_\perp^{L,R}$ and $A_0^{L,R}$ corresponding to the various transversity states of the K^{*0} meson as well as left- and right-handed chiralities of the dimuon system. Muons in the final state can be considered massless when q^2 is large enough ($q^2 \gtrsim 1$ GeV^2/c^4). Two additional amplitudes can also be present: A_t for the massive case and A_S for scalar contribution. Those amplitudes are functions of the Wilson coefficients, containing information about short distance effects, and functions called form factors, governing the

$B^0 \to K^{*0}$ transition which can include long distance effects. Form factors are scalar functions of q^2, which originate from Lorentz decomposition of the hadronic transition matrix elements. Explicit formulas for differential decay rate, q^2 dependent I_i functions and complex amplitudes along with their dependence on Wilson coefficients and form factors are given in Appendix A.

It is possible to separate CP-conserving and CP-violating effects by introducing a set of observables:

$$\begin{aligned} S_i &= (I_i + \bar{I}_i)/(\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}), \\ A_i &= (I_i - \bar{I}_i)/(\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}). \end{aligned} \qquad (i = 1, .., 9) \tag{3.7}$$

In literature [13], the commonly used observables are forward-backward asymmetry A_{FB} and K^{*0} fraction of longitudinal and transverse polarization, F_L and F_T, which are related to CP-averages as follows:

$$\begin{aligned} A_{FB} &= \frac{3}{4} S_6^s, \\ F_L = S_1^c &= -S_2^c, \\ F_T &= 4 S_2^s, \end{aligned} \tag{3.8}$$

where additional superscripts c and s indicate variables' dependence on $\cos^2\theta_K$ and $\sin^2\theta_K$, respectively. CP-averages, along with the differential decay width $\frac{d\Gamma}{dq^2}$, are known as Form Factor Dependent(FFD) observables at leading order (LO), which means that they suffer from hadronic uncertainties. In order to eliminate the leading $B^0 \to K^{*0}$ form-factor uncertainties, an additional set of observables $P_{1,2,3}$ and $P'_{4,5,6,8}$ was proposed [20], which are Form Factor Independent(FFI). Those theoretically clean observables are constructed from ratios of CP-averaged angular observables S_i and thanks to that form factors are completely canceled at LO:

$$\begin{aligned} P_1 &= 2\frac{S_3}{1 - F_L}, \\ P_2 &= \frac{2}{3}\frac{A_{FB}}{1 - F_L}, \\ P_3 &= -\frac{S_9}{1 - F_L}, \\ P'_{4,5,6,8} &= \frac{S_{4,5,7,8}}{\sqrt{F_L(1 - F_L)}}. \end{aligned} \tag{3.9}$$

In many publications an angular analysis is based on the determination of the FFD angular observables S_i (or A_i) with an unbinned maximum likelihood fit of the B^0 mass and the decay angles. From the obtained CP-averages it is then possible to obtain FFI observables. In the recent past P'_5 observable became very interesting since distinct discrepancies from the SM were found in the $B^0 \to K^{*0}\mu^+\mu^-$ LHCb analysis [13].

3.3 Method of moments

It is possible to determine the angular observables without using the likelihood fit of the angular distribution. This analysis is focused on such approach, where the method of moments is applied [21]. In this method a proper angular function is used as a weight to the collected data, which in result projects out any necessary observable.

When we compare the method of moments with a likelihood fit, it appears that the former has some interesting advantages. When we study decays like $B^0 \to K^+\pi^-\mu^+\mu^-$, which is a FCNC process suppressed in the SM, we are dealing with small data sample sizes and in that case a likelihood fit can suffer from convergence issues. Those problems can be dealt with by using reparametrizations or approximations. The method of moments does not need them and gives satisfactory results for a small number of candidates. A successful fit cannot be achieved if a physical model is not fully known and some additional information, such as relations between observables, has to be used as an input. In the method of moments no additional information is needed, and what is more a correlation between angular observables is given as a result. Additionally, there is no need to formulate the likelihood and each angular observable can be measured independently. In the method of moments, for any number of candidates the estimate is unbiased, but this results in a variance increase in the sampling distribution of the estimator. Thus, the method of moments suffers from bigger uncertainties than the likelihood fit.

The method of moments starts with the definition of the probability density function (PDF), which in this case depends on q^2 and three decay angles:

$$PDF(q^2, \vec{\Omega}) = \sum_i M_i(q^2) f_i(\vec{\Omega}), \tag{3.10}$$

where $\vec{\Omega} = \{\theta_\ell, \theta_K, \phi\}$, the angular dependence is factored out as the angular function $f(\vec{\Omega})$ and coefficient M_i is referred as i-th moment of the PDF. It is implied that angular functions are orthonormal:

$$\int f_i(\vec{\Omega}) f_j(\vec{\Omega}) d\vec{\Omega} = \delta_{ij} \tag{3.11}$$

Then the observables M_i can be projected out:

$$M_i(q^2) = \int PDF(q^2, \vec{\Omega}) f_i(\vec{\Omega}) d\vec{\Omega}, \tag{3.12}$$

and the q^2 dependence integrated out:

$$\langle M_i \rangle = \int dq^2 M_i(q^2) \tag{3.13}$$

Now, the analytical integration can be replaced by the Monte Carlo estimator:

$$\langle M_i \rangle \to \widehat{\langle M_i \rangle} = \frac{1}{N} \sum_{k=1}^{N} f_i(x_k) \tag{3.14}$$

where N is the number of candidates in a dataset and according to the law of large numbers, for N $\rightarrow \infty$ data points are distributed as the corresponding PDF, so the method of moments comes down to a simple counting measurement. As it was mentioned before it is also possible to calculate the covariance matrix:

$$\widehat{Cov}[M_i, M_j] = \frac{1}{N-1}\sum_{k=1}^{N}[f_i(x_k) - \widehat{M_i}][f_j(x_k) - \widehat{M_j}], \tag{3.15}$$

from which the calculation of statistical uncertainty is straightforward.

3.4 Higher $K\pi$ mass region

Most of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decay analyses (e.g. [13, 19]) are focused on $K^*(892)^0$ resonance mass region where candidates come predominantly from the P-wave process $K^*(892)^0 \rightarrow K^+\pi^-$. However, this resonance is not the only one in the $K^+\pi^-$ mass distribution. In high energy physics the term "P-wave" stands for a particle with total angular momentum J = 1. Therefore, for spin $J = 0, 1, 2$ we can distinguish between S-, P- and D-wave, respectively. Figure [3.2] shows the candidates after preselection and BDT cut described in Chapter 5. Above the $K^*(892)^0$ resonance, in a higher mass region $1330 < m_{K\pi} < 1530$ MeV/c^2 a distinct structure is visible, where contributions from the S-, P- and D-waves could be found (Table 3.2). This analysis is a continuation of the previous one performed by the LHCb Collaboration based on Run1 data set [22] and is focused on this higher $K^+\pi^-$ invariant mass range using data from Run2 data taking period (2016, 2017, 2018).

For the limit where the muon mass can be neglected comparing to q^2, the CP-averaged differential decay rate of the $B^0 \rightarrow K^+\pi^-\mu^+\mu^-$, where the $K\pi$ system is in the S-, P- or D-wave configuration can be expanded in an orthonormal basis of 41 angular functions [16]:

$$\frac{d^4\Gamma}{dq^2 dcos\theta_K cos\theta_\ell d\phi} \equiv \sum_{i=1}^{41} M_i(q^2) f_i(cos\theta_K, cos\theta_\ell, \phi), \tag{3.16}$$

where M_i is the i-th moment. It is worth mentioning here that with the method of moments it is easily possible to derive both FFD and FFI observables. Here are several examples:

$$\begin{aligned} M_3 &= \frac{3S_1^c - 3S_1^s - S_2^c + S_2^s}{4\sqrt{10\pi}}, \\ M_{12} &= \frac{3}{10\sqrt{\pi}} S_4, \\ M_{35} &= \frac{3}{4\sqrt{5\pi}} S_5. \end{aligned} \tag{3.17}$$

Interestingly, using only the moments M_1, M_3, and M_{35}, the famous P_5' observable can be calculated when non-P-wave components are absent:

$$P_5' = \sqrt{\frac{5}{(M_1 + \sqrt{5}M_3)(M_1 - \sqrt{5}M_3/2)}} M_{35}. \tag{3.18}$$

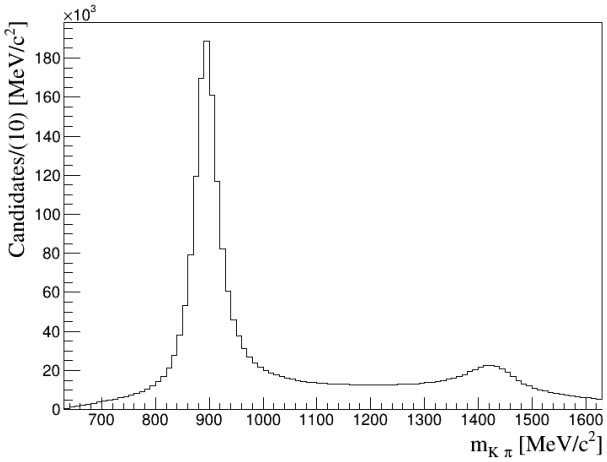


Figure 3.2: Invariant mass distribution of the $K^+\pi^-$ system for $B^0 \to (K^{*0} \to K^+\pi^-)\mu^+\mu^-$ decays after preselection and BDT cut.

Resonance	J^P	Mass (MeV/c^2)	Width (MeV/c^2)
$K^*(1410)$	1^-	1414±15	232±21
$K_0^*(1430)$	0^+	1425±50	270±80
$K_2^*(1430)$	2^+	1432.4±1.3	109±5

Table 3.2: Predicted resonant contributions in the $1330 < m_{K\pi} < 1530$ MeV/c^2 range. Each contains information about spin-parity J^P, invariant mass and decay width [23].

3.5 Methodology

There are two main goals of this study: to perform an angular measurement and differential branching fraction calculation of $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in the higher mass region of $1330 < m_{K\pi} < 1530$ MeV/c^2. As it is depicted in Table 3.3, the angular analysis (Chapter 8) is performed only in a single q^2 bin, $[1.1, 6.0]$ GeV^2/c^4 and comprises the calculation of 41 angular moments. On the other hand, the differential branching fraction measurement, described in Chapter 9, is performed in six q^2 bins.

	q^2 ranges [GeV^2/c^4]
Angular analysis	[1.1,6.0]
Differential branching fraction	[0.1,0.98], [1.1,2.5], [2.5,4.0], [4.0,6.0], [6.0,8.0], [1.1,6.0]

Table 3.3: Bins of q^2 used in both angular analysis and differential branching fraction measurements.

Chapter 4

Experimental setup/Large Hadron Collider

The Large Hadron Collider (LHC) is the biggest accelerator in the world (Fig.4.1). Its 26.7 km circular tunnel is located 100 meters beneath the France–Switzerland border near Geneva. In the tunnel proton bunches travel through two beam pipes in opposite directions, and when they are accelerated to a designated energy they collide in four main crossing points. The acceleration process comes in several stages. After the hydrogen atoms are stripped from the electrons, they are accelerated by the Linear Accelerator 2 (Linac2) to the energy of 50 MeV. Next, the protons go to the Proton Synchrotron Booster, then to the Proton Sychrotron (SP) and further to the Super Proton Sychrotron (SPS), which increases the energy up to 1.4, 25 and 450 GeV, respectively. In the final stage, the proton bunches are injected to the main LHC beam pipe and reach the energy of 13 TeV.

Currently, there are nine different experiments located at the LHC. The first four are located in the main caverns where proton bunches collide:

- LHCb (LHC beauty) - dedicated to study flavour physics,
- Alice (A Large Ion Collider Experiment) - dedicated to heavy ion collisions,
- Atlas (A Thoroidal LHC Apparatus) and CMS (Compact Muon Solenoid) - detectors of general purpose involved in the discovery of the Higgs boson,
- TOTEM - focused on the measurement of total cross section, elastic scattering, and diffraction processes,
- LHCf (LHC forward) - for studies of neutral particles generated in the forward direction of pp beam,
- MoEDAL (Monopole and Exotics Detector At the LHC)- its primary goal is to search for magnetic monopole or other highly ionizing stable massive particles (SMPs)

- FASER (ForwArd Search ExpeRiment) - searching for new light and weakly coupled elementary particles,
- SND (Scattering and Neutrino Detector) - designed to study neutrinos, and as the first attempt to detect neutrinos directly in the collider.

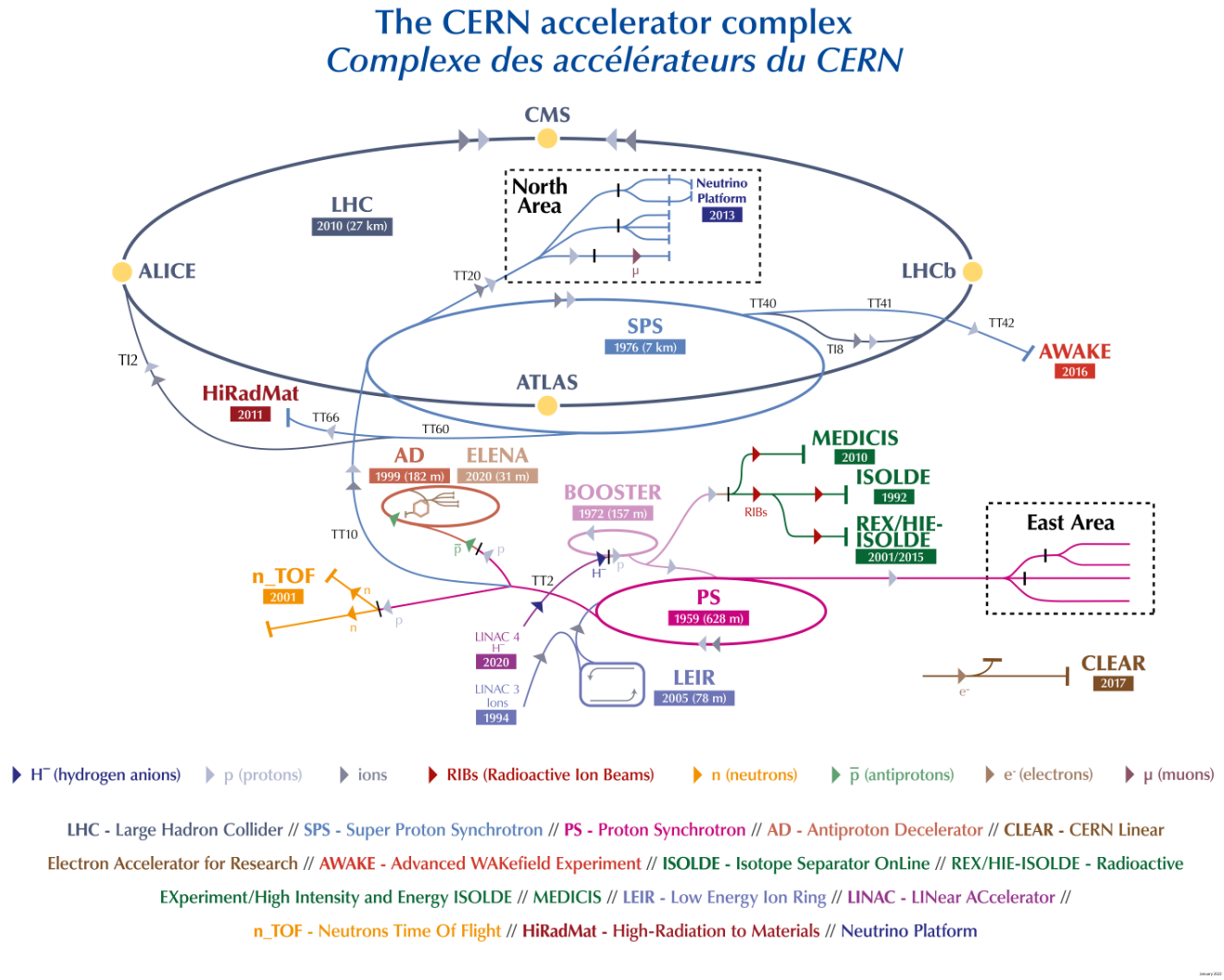


Figure 4.1: The layout of the CERN accelerator complex [24].

4.1 LHCb detector in LHC

The LHCb detector [25] is a single-arm forward spectrometer designed to study heavy flavour physics. It covers a pseudorapidity range of $2 < \eta < 5$, which corresponds to the

range of [10,300] mrad. This is dictated by the fact that most of the final state particles from $b\bar{b}$ pairs are produced in both forward and backward direction.

The LHCb detector is built up from several sub-detectors responsible for providing different information about passing particles. It contains tracking detectors of high precision (vertex locator - VELO, TT and T1-T3), ring-imaging Cherenkov detectors (RICH1, RICH2) for particle identification, hadronic and electromagnetic calorimeters (HCAL, ECAL), which measure the energies and directions of particles, and muon identification system (M1-M5). Particles inside the LHCb detector experience a bending magnetic field of around 4 Tm. All those components of the LHCb, which can be seen in Figure 4.2, are described in the next few subsections.

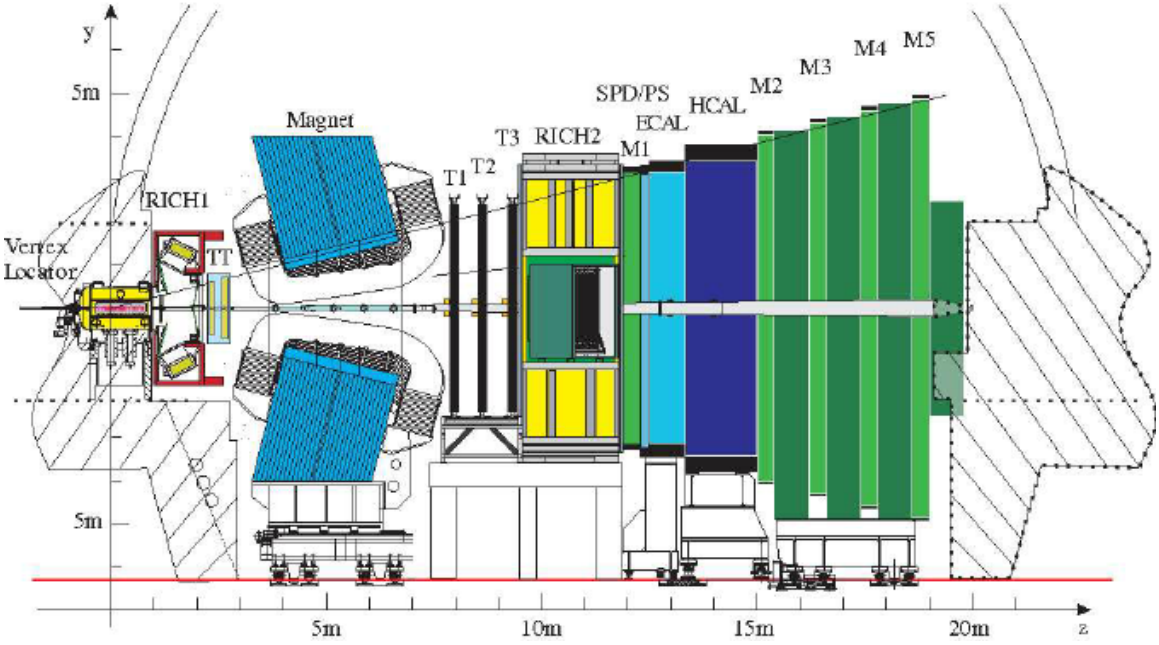


Figure 4.2: The LHCb detector [25].

4.1.1 Vertex locator

The closest to the Interaction Point (IP) is the VErtex LOcator (VELO)[26], which is responsible for measuring particle trajectories. Its main purpose is to measure the distance between IP and the point where the particles containing b quark decay. It contains 84 silicon strip sensors divided in two halves (Fig. 4.3). The strip sensors consist of two types: the R-type sensor measures the radial coordinate and the ϕ-type sensor measures the azimuthal coordinate of a particle. Both halves are separated from the primary vacuum of the LHC by a 300-μm aluminum RF-foil. In order to prevent any damage from the proton beams, VELO is able to move. When the protons are injected, two halves of VELO are in opened position, and after the injection they are moved of around 8 mm to the beam axis. VELO provides a primary vertex resolution of 10 μm transverse to the beam axis and 60 μm along the beam axis.

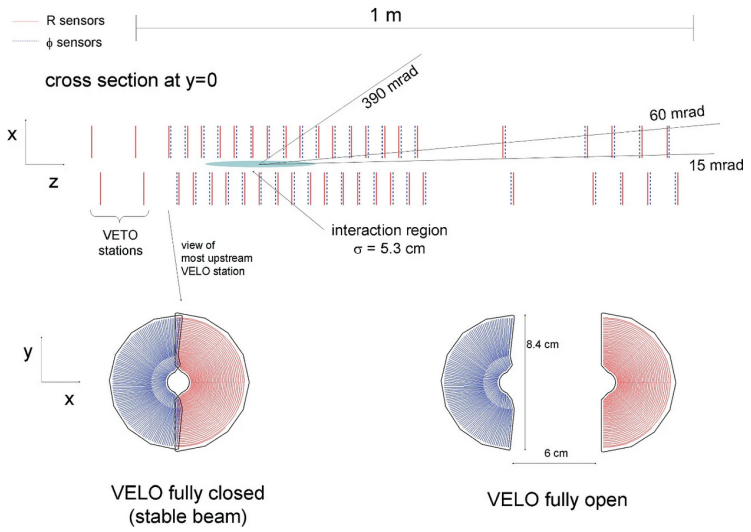


Figure 4.3: Upper side: Vertex Locator layout. Lower side: VELO in closed (left) and opened (right) position.

4.1.2 Tracking system

The tracking system [27, 28] is divided into two parts: the Tracker Turicensis (TT), which is located upstream of the dipole magnets, and the Outer Tracker (OT) composed of three tracking stations (T1-T3) located downstream of the dipole magnets. The TT is constructed out of four layers of silicon microstrip detectors with a strip pitch of approximately 200 μm and provides 50 μm resolution for a single hit. Its area of 150 cm wide and 130 cm high covers the full acceptance of the LHCb. The same silicon strips are used in the Inner Tracker (IT) located around the beam pipe in stations T1-T3, which covers a region 120 cm wide and 40 cm high. Each of four silicon trackers has four detection layers which are tilted by 5° angle.

The OT is based on straw-tubes drift chambers, with each tube of 4.9 mm in diameter containing a mixture of 70% argon and 30% CO_2. It provides a drift time faster than 50 ns, and 200 μm of drift-coordinate resolution. Each OT station contains four layers of 64 straw-tubes also tilted by 5° angle and covers the area of 5971×4850 mm^2.

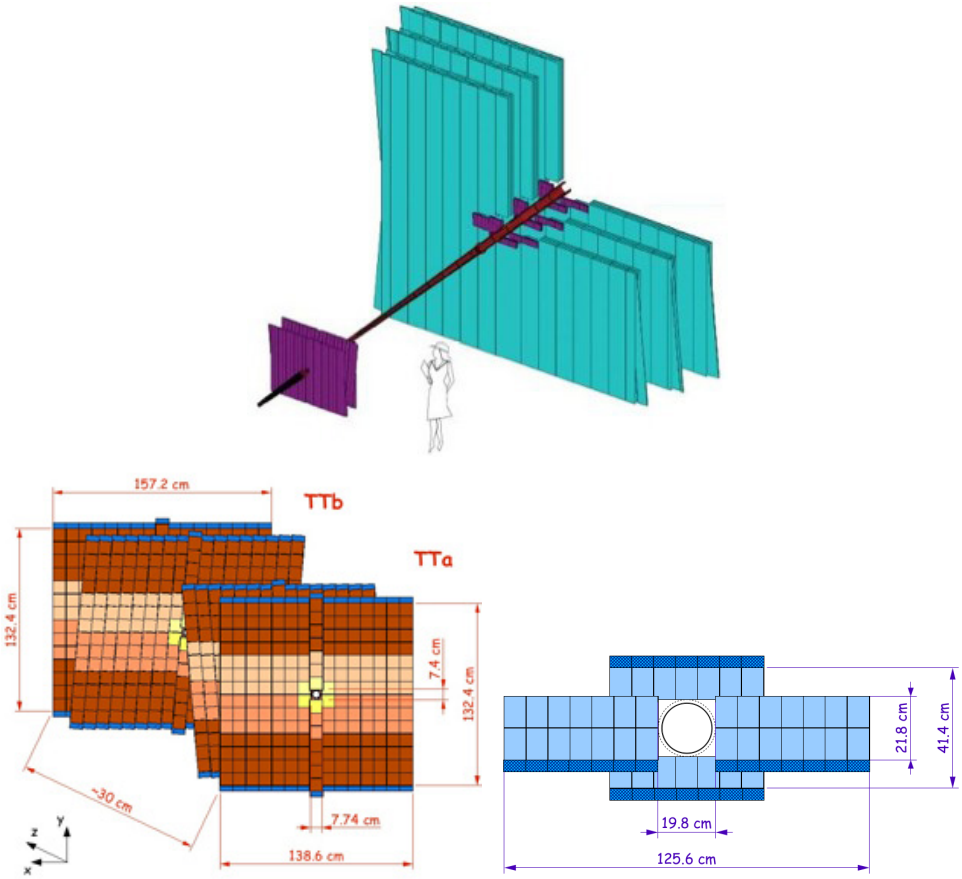


Figure 4.4: Upper side: schematic view of the OT (cyan) and TT with IT (purple). Lower side: schematic view of TT layers (left) and IT layer (right).

4.1.3 Ring Imaging Cherenkov Detectors (RICH)

Efficient Particle Identification (PID) is crucial for a selection of final state particles coming from B decays. The LHCb is equipped with two Ring Imaging Cherenkov detectors (RICH1 and RICH2) [29] to cover both low and high momentum charged particles. These detectors measure the emission of Cherenkov radiation. This phenomenon occurs when a particle travels through a dense material faster than the speed of light in this material, producing a cone of light in the direction of a particle. Both RICH detectors use a combination of spherical and flat mirrors to focus the Cherenkov light coming from passing particles which is then reflected onto the photodetectors. Hybrid Photon Detectors (HPDs) are used to detect the Cherenkov photons in the wavelength range of 195–700 nm.

The RICH1 detector is placed between VELO and the TT (Fig. 4.2). It is filled with aerogel and fluorobutane (C_4F_{10}) gas radiators and is sensitive to low momentum charged particles in the range of 1–60 GeV/c. It has a very convenient acceptance, which covers the full LHCb acceptance up to ±300 mrad horizontally and up to ±250 mrad vertically.

The RICH2 detector, which is located downstream of the dipole magnet right after the OT, is dedicated to identify particles with high momentum range from 15 GeV/c up to and beyond 100 GeV/c. Its angular acceptance covers the angular range up to ±120 mrad, horizontally and up to ±100 mrad vertically and it uses a carbon tetrafluoride CF_4 radiator.

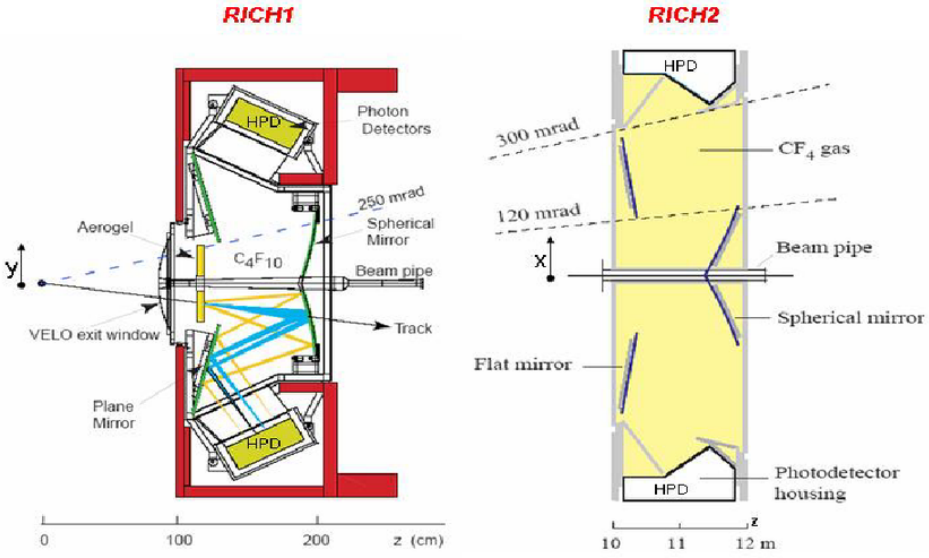


Figure 4.5: The layout of the RICH1 detector (left) and RICH2 detector (right) at LHCb.

4.1.4 Calorimeters

The calorimeter system [30] comprises four different layers: the Scintillating Pad Detector (SPD), the Pre-Shower Detector (PS), the Electromagnetic Calorimeter (ECAL), and the Hadronic Calorimeter (HCAL). The purpose of the SPD is to distinguish between charged and neutral particles while the PS gives information whether a particle is a photon or an electron. Both SPD and PS are composed of 15 mm thick scintillating pads separated by a lead converter of 2.5 X_0 (X_0 - radiation length). A signal in the form of light is transmitted using wavelength-shifting fibers (WLS) to multi-anode photomultipliers (MAPMTs). The ECAL consists of alternating layers of 4 mm thick scintillating pads and 2 mm thick lead plates, which correspond to 25 X_0. With the information from the SPD and PS, the main objective is a photon-electron separation. The HCAL detector, placed behind the ECAL, comprises the layers of 3 mm thick scintillators and 1 cm thick iron plates which are arranged parallel to the beam pipe. Its length is equivalent to 5.6 X_0.

4.1.5 Muon identification system

The LHCb detector is equipped with a crucial component for identifying the final state of the investigated B^0 decay. The muon system [31] comprises five rectangular shaped stations, which gradually increase in size and are placed at the end of the LHCb detector (Fig. 4.2). The first station M1, placed upstream of the SPD, comprises triple-GEM detectors and the rest of the stations M2-M5 are built up from multi wire proportional chambers separated by 80 cm thick iron plates. The chambers are filled with three gases combined: argon, carbon dioxide, and carbon tetrafluoride. As it is depicted in Fig. 4.6, the muon stations are divided into four regions (R1, R2, R3, R4). Each region has different granularity corresponding to particle flux in given region to keep relatively constant occupancy over the muon detector system. Information from the muon stations is used in the L0 trigger and for muon identification in the high-level trigger (HLT).

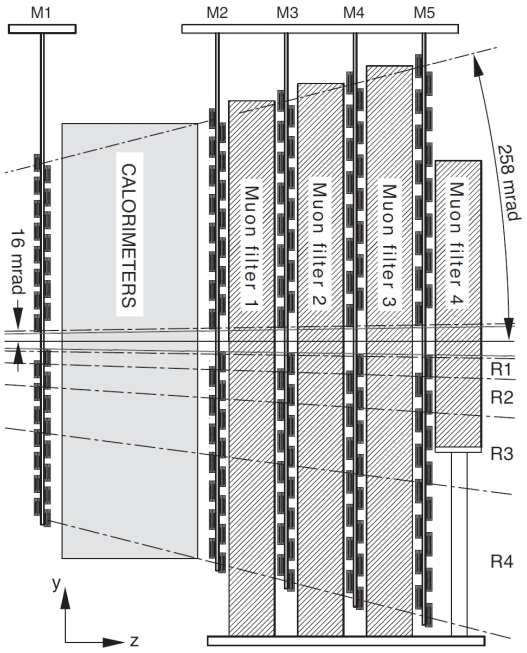


Figure 4.6: The layout of muon stations at LHCb.

4.1.6 Trigger

The frequency of proton-proton collisions at the LHC is equal to 40 MHz, but only a fraction of interesting candidates has final state particles within the acceptance of the LHCb detector. As the offline computing capacity is limited, the candidates detected by the LHCb need to be filtered before recording on tape. The trigger system [32] is dedicated to reduce the number of candidates and is divided into two stages: First Level Trigger (L0) and High Level Trigger (HLT). The L0 is a hardware trigger using information from

the calorimeters and muon system which can efficiently select high-p_T particles coming from the decays of interest. It reduces the collison rate to 1 MHz. The HLT is a software trigger divided into two stages: HLT1 and HLT2.

The HLT1 performs a partial candidate reconstruction where it confirms the high-p_T L0 candidate particles and uses information from VELO and the tracking stations to select particles with properties characteristic to a specific decay. Because of the relatively long life-time of B mesons or B baryons, the impact parameter is relatively high to the proton-proton interaction vertex.

The HLT2 performs a full candidate reconstruction comprising three steps: the track reconstruction of charged particles, the reconstruction of neutral particles, and particle identification (PID).

During Run2 the LHC provided a higher number of interesting candidates, compared with Run1, which was due to increased production cross-section. Having a higher number of candidates is challenging because of limited computing resources. That is why in 2015 the LHCb introduced the Turbo stream. In Run2 the HLT2 and offline reconstructions are the same, but selections can be different. With the Turbo stream selected candidates from HLT2 can be saved to disk and be available for analysis without additional offline reconstruction.

Chapter 5

Data Selection

The data samples for this analysis come from the LHCb experiment and were collected in 2016, 2017 and 2018, where proton-proton collisions reached the center of mass energy of 13 TeV. In those collisions pairs of b quark and b antiquark, denoted as $\bar{\mathrm{b}}$, are produced. Next, each b quark separates from the other and in consequence it hadronizes, which means that another quark-antiquark pair appears. At the LHC a $\mathrm{b\bar{b}}$ pair is dominantly produced via gluon fusion, where momenta of the incoming partons are asymmetric in a laboratory frame. Therefore, the center of mass of a $\mathrm{b\bar{b}}$ pair is boosted along the direction of a gluon with higher momentum, and consequently both produced hadrons have the same, forward or backward, direction. The forward geometry of the LHCb makes it possible to detect all products of the B^0 decay and sometimes the decay products of another b hadron can be detected. This additional detection can be used for the flavour tagging.

5.1 LHCb data

Table 5.1 shows the details about each data sample for a corresponding year. The integrated luminosity is equivalent to the collected data size. It is expressed as an inverse of the cross section ($\mathrm{fb}^{-1} \approx 7 \times 10^{13}$ pp collisions). Information about the reconstruction configuration and stripping version is also given. Stripping is a set of selections used for data filtering. Data is divided into stripping streams, which helps to speed up access to the files and to save disk space. All data samples come in two versions that correspond to different magnet polarity (MagDown, MagUp).

year	Integrated luminosity	Reconstruction	Stripping
2016	1.67/fb	reco16	Stripping28r1
2017	1.71/fb	reco17	Stripping29
2018	2.19/fb	reco18	Stripping34

Table 5.1: Data obtained for analysis. Integrated luminosity for a given year, along with the information about reconstruction and stripping. Each sample comes in two versions of magnet polarity (MagDown, MagUp).

5.2 Simulation

After obtaining the data samples, the simulated data need to be used to take into account the effects of trigger, reconstruction and selection on the angular distributions of the signal and normalization mode. Additional samples of background processes will be generated in order to estimate their contributions. In a simulation, pp collisions are generated using Pythia8 [33] with a specific LHCb configuration [34]. Decays of hadronic particles are described by EvtGen [35], in which final-state radiation is generated using Photos [36]. The interaction of the generated particles with the detector and its response are implemented using the Geant4 toolkit [37]. To efficiently model distortion effects on angular distributions caused by the trigger, reconstruction and selection, it is crucial to have the data and simulated candidates compatible with each other. All simulated samples were filtered on the `StrippingB2XMuMu` stripping line and produced with SimType Sim09. Just like real data, they come in two versions (MagDown, MagUp) for each given year. The simulated samples for this analysis are summarized in Table 5.2.

No.	Decay	Event type	Model	Number of candidates
1	$B^0 \to K^+\pi^-\mu^+\mu^-$	TGenPhaseSpace	PHSP	100 M
2	$B^0 \to K^+\pi^-\mu^+\mu^-$	11114000	PHSP	11.4 M
3	$\mathrm{B}^0 \to \mathrm{K}^{*0}\mathrm{J}/\psi$	11144001	Phys	4.0 M
4	$\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$	11114002	Phys	9.2 M
5	$\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}\pi$	11144050	PHSP	9.8 M
6	$\mathrm{B}^0 \to \psi(2\mathrm{S})\mathrm{K}\pi$	11144051	PHSP	5.2 M
7	$B^+ \to K^+\mu^+\mu^-$	12113002	Phys	0.6 M
8	$\Lambda_b \to pK\mu^+\mu^-$	15114011	PHSP	1.9 M
9	$B_s^0 \to \phi\mu^+\mu^-$	13114002	Phys	3.1 M

Table 5.2: Simulated samples obtained for analysis. Each decay channel comes in two versions (MagDown, MagUp)for each given year.

Each simulated sample has a different event type number. An event type number is a numerical representation of a specific decay channel. TGenPhaseSpace is a utility class from the ROOT package [38] that can generate all possible solutions of a given decay. The

simulation samples come in two versions: phase space and physical. Candidates of the phase space sample are generated based purely on kinematics and phase space constraints on the final state particles. The phase space sample does not incorporate any detailed physical processes. On the other hand, the physical simulation samples are more realistic and reflect representative physical effects, such as the detector responses.

5.3 Trigger requirements

The LHCb trigger system, described in Section 4.1.6, allows the most interesting candidates to be collected. There are several trigger requirements that the candidates have to pass and at each stage the candidates are required to be "triggered independent of signal" (TIS) or "triggered on signal" (TOS). A candidate is considered TIS when, if it is removed from the decay, this decay would still be accepted by the trigger selection. There is another particle in this decay that also is accepted. The candidate is considered TOS only when it is accepted by the trigger selection. Only TOS candidates are used in this analysis since they are better simulated in MC. The trigger lines are taken with the logical or, which means that in each trigger stage a particle has to pass at least one of the listed trigger lines to be considered a candidate. The trigger requirements are listed in Tab. 5.3.

Stage	Trigger lines
L0	L0Muon
HLT1	Hlt1TrackAllL0 or Hlt1TrackMuon
HLT2	Hlt2Topo[2,3,4]BodyBBDT, Hlt2TopoMu[2,3,4]BodyBBDT, Hlt2SingleMuon or Hlt2DiMuonDetached

Table 5.3: Trigger requirements for signal candidates.

5.4 Stripping

Candidates that have passed the trigger requirements are stored on a tape. Still, this produces a lot of data, so in order to make it easier to analyze stripping selection is performed, after which the data is stored on storage disks. All $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates are required to pass the stripping requirements contained in `StrippingB2XMuMu` stripping line. Each data sample from 2016, 2017 and 2018 corresponds to different stripping and reconstruction versions (Table 5.1). The stripping requirements are listed in Table 5.4.

Candidate	Stripping 28r1, 29 and 34
B^0	IP $\chi^2 < 16$ (best PV) flight distance $\chi^2 > 121$ vertex χ^2/ndf < 8 DIRA angle <14 mrad $4900 < m < 7000$ MeV/c^2
K^{*0}	m(Kπ) < 6200 MeV/c^2 vertex χ^2/ndf < 12 flight distance $\chi^2 > 16$
$\mu^+\mu^-$	m($\mu^+\mu^-$) < 7100 MeV/c^2 vertex χ^2/ndf < 9
tracks hadron muon	ghost Prob <0.5 min IP $\chi^2 > 6$ min IP $\chi^2 > 9$
muon	IsMuon DLL$_{\mu\pi} > -3$
GEC	SPD Mult.<600

Table 5.4: Selection criteria in `StrippingB2XMuMu` for stripping 28r1, 29 and 34.

In Table 5.4, the DLL$_{XY}$ refers to the Delta Log Likelihood representing the difference in the logarithm of the likelihood function between two models: first with the X particle mass hypothesis and second with the Y particle mass hypothesis.

5.5 Preselection

Further preselection is performed on the obtained stripped data sample to get rid of most of not interesting background. The preselection requirements are listed in Tab. 5.5, where θ stands for the opening angle from the beam, the θ_{pair} is the opening angle between two track pairs and the variables $< X >$, $< Y >$, $< Z >$ denote the mean position of the primary vertex.

A single wide requirement on invariant mass of the Kπ system is used, in order to cover contribution from the normalization channel $B^0 \to J/\psi K^*(892)^0$ and the higher partial waves in the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay mode:

$$630 \text{ MeV/c}^2 < m_{K\pi} < 1630 \text{ MeV/c}^2. \tag{5.1}$$

The $B^0 \to K^+\pi^-\mu^+\mu^-$ MC sample is used for the acceptance correction. This sample contains MC generator level transverse momentum requirement of $p_T > 1500$ MeV/c for the B^0 candidates. The same requirement is applied to both data and simulated samples.

Candidate	Selection
B^0	$4960 < m < 6000\ MeV/c^2$ $p_T > 1500\ MeV/c$
K^{*0}	$630 < m(K\pi) < 1630\ MeV/c^2$
track Track Pairs	$0 < \theta < 400$ mrad $\theta_{pair} > 1$ mrad
K	hasRich True $DLL_{K\pi} > -5$
π	hasRich True $DLL_{K\pi} < 25$
$\mu^+\mu^-$	IsMuon True
PV	$\lvert X- < X > \rvert < 5$ mm $\lvert Y- < Y > \rvert < 5$ mm $\lvert Z- < Z > \rvert < 200$ mm

Table 5.5: Preselection cuts applied to stripped candidates.

5.6 Peaking background vetoes

There are several contributions from decay modes that can be present at or near the signal region and can consequently distort angular and mass distributions. Those decays are described in the following subsections.

5.6.1 Strange and charm resonances

This analysis is focused on invariant mass of muon pair squared region of $0.1 < q^2 < 8.0$ $[GeV^2/c^4]$, therefore any contributions from charm resonances in the $B^0 \to (J/\psi \to \mu^+\mu^-)K^+\pi^-$ and $B^0 \to (\psi(2S) \to \mu^+\mu^-)K^+\pi^-$ decays are not present. Additionally, the chosen q^2 binning shown in Table 3.3 excludes the contribution from strange resonance in the $B^0 \to (\phi(1020) \to \mu^+\mu^-)K^+\pi^-$ decays.

5.6.2 $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decay

A peaking background from the $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decays appears when a proton is reconstructed as a pion or in another scenario a proton is identified as a kaon and simultaneously a kaon is identified as a pion. To take this into account, by using PID information and alternative mass hypotheses, new mass hypotheses have to be constructed: they are denoted as $m_{(\pi\to p)K\mu\mu}$ and $m_{(K\to p)(\pi\to K)\mu\mu}$. In the first scenario a pion is given the proton mass hypothesis, in the second scenario a kaon is given the proton mass hypothesis, and a pion is given the kaon mass hypothesis. Candidates from a given $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decay are expected to have those two new masses consistent with the known Λ_b^0 mass.

In order to reject contributions from both peaking backgrounds, the following vetoes are constructed using new mass hypotheses and PID information:

$$\begin{aligned}(5575 < \mathrm{m}_{(\pi\to\mathrm{p})\mathrm{K}\mu\mu} < 5665)\ \mathrm{MeV/c^2}\ \text{and}\ \pi\mathrm{DLL}_{\mathrm{p}\pi} > 0\\ (5575 < \mathrm{m}_{(\mathrm{K}\to\mathrm{p})(\pi\to\mathrm{K})\mu\mu} < 5665)\ \mathrm{MeV/c^2}\ \text{and}\ \pi\mathrm{DLL}_{\mathrm{K}\pi} > 0\end{aligned} \tag{5.2}$$

The effects of the first veto from the above equation is shown in Fig 5.1. The upper plot presents the plane of $\mathrm{m}_{(\pi\to\mathrm{p})\mathrm{K}\mu\mu}$ and $\pi\mathrm{DLL}_{\mathrm{p}\pi}$ for $\Lambda_\mathrm{b}^0 \to \mathrm{pK}^-\mu^+\mu^-$ candidates in simulation before applying the veto. The lower plot in Fig 5.1 shows the impact of the corresponding veto on $\mathrm{B}^0 \to \mathrm{K}^+\pi^-\mu^+\mu^-$ candidates in simulation.

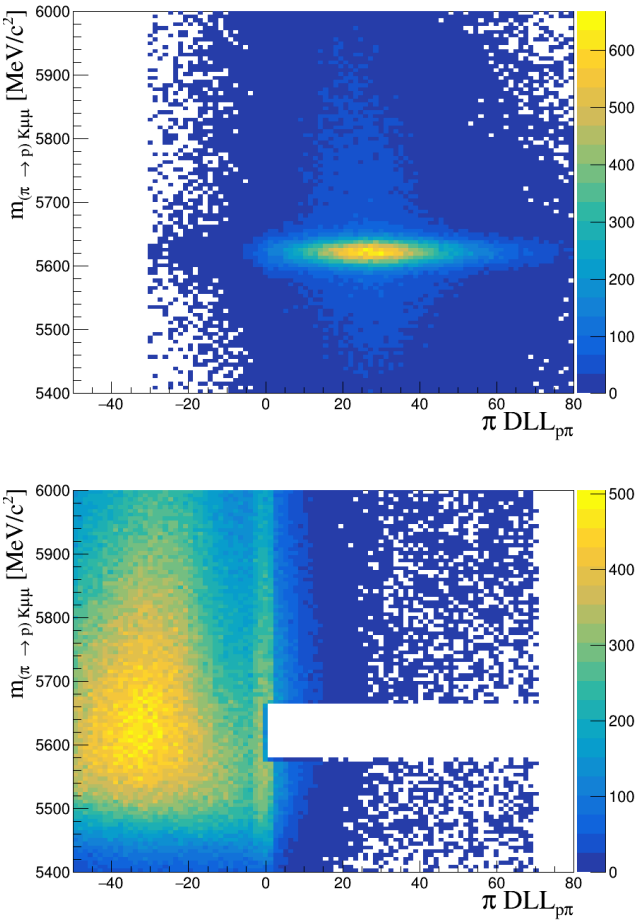


Figure 5.1: The 2D distributions of Λ_b^0 candidates mass ($\mathrm{m}_{(\pi\to\mathrm{p})\mathrm{K}\mu\mu}$) and pion candidate $\mathrm{DLL}_{\mathrm{p}\pi}$ for simulated background $\Lambda_\mathrm{b}^0 \to \mathrm{pK}^-\mu^+\mu^-$ candidates before veto (upper) and for simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates after veto (lower).

The effects of the second veto from Eq.5.2 is shown in Fig. 5.2. The upper plot presents the plane of $m_{(K\to p)(\pi\to K)\mu\mu}$ and $\pi\mathrm{DLL}_{\mathrm{K}\pi}$ for $\Lambda_\mathrm{b}^0 \to \mathrm{pK}^-\mu^+\mu^-$ candidates in simulation before applying the veto. The lower plot in Fig. 5.2 shows the impact of the corresponding veto on $\mathrm{B}^0 \to \mathrm{K}^+\pi^-\mu^+\mu^-$ candidates in simulation.

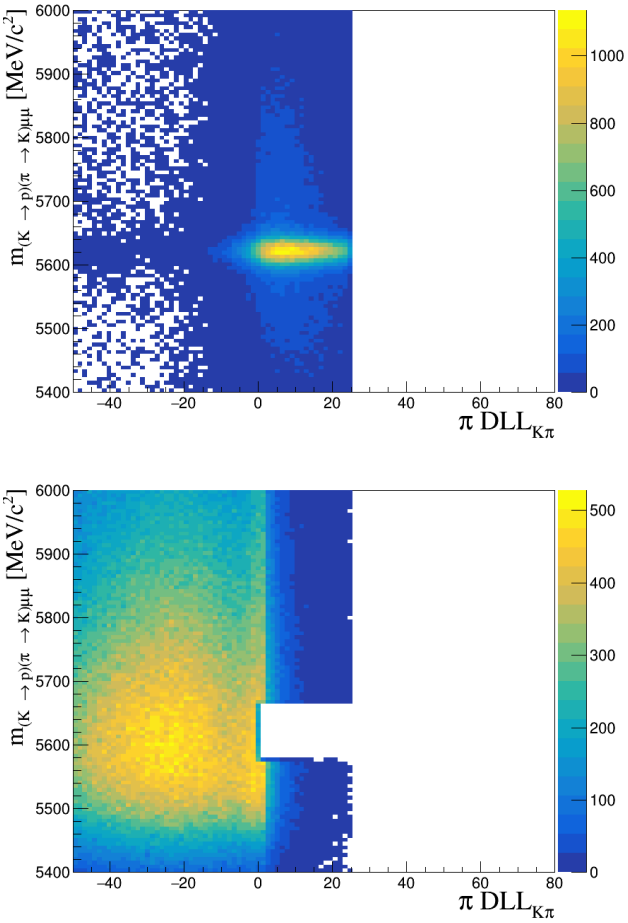


Figure 5.2: The 2D distributions of Λ_b^0 candidates mass ($\mathrm{m}_{(\mathrm{K}\to\mathrm{p})(\pi\to\mathrm{K})\mu\mu}$) and kaon candidate $\mathrm{DLL}_{\mathrm{K}\pi}$ for simulated background $\Lambda_\mathrm{b}^0 \to \mathrm{pK}^-\mu^+\mu^-$ candidates before veto (upper) and for simulated signal $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates after veto (lower).

5.6.3 $\mathrm{B}_\mathrm{s}^0 \to \phi\mu^+\mu^-$ decay

The decay $\mathrm{B}_\mathrm{s}^0 \to \phi\mu^+\mu^-$ with $\phi \to \mathrm{K}^+\mathrm{K}^-$ can contribute as peaking background if one kaon from this decay is misidentified as a pion. When the reconstructed pion is assigned the nominal mass of a kaon, the invariant mass $m_{(\pi\to K)K}$ should then be consistent with the mass of ϕ and in the same way the $\mathrm{K}\pi\mu\mu$ invariant mass $m_{(\pi\to K)K\mu\mu}$ should be

consistent with the mass of B_s^0. Knowing that the following vetoes are used:

$$\begin{aligned} 5321 < m_{(\pi\to K)K\mu\mu} < 5411\,\mathrm{MeV/c^2} \\ \text{and } 1010 < m_{(\pi\to K)K} < 1030\,\mathrm{MeV/c^2} \end{aligned} \tag{5.3}$$

with the pion satisfying $DLL_{K\pi} > -10$, or

$$\begin{aligned} 5321 < m_{(\pi\to K)K\mu\mu} < 5411\,\mathrm{MeV/c^2} \\ \text{and } 1030 < m_{(\pi\to K)K} < 1050\,\mathrm{MeV/c^2} \end{aligned} \tag{5.4}$$

with the pion satisfying $DLL_{K\pi} > 10$.
The reason for vetoing two separate regions of $m_{(\pi\to K)K}$ is to reduce the number of the removed $B^0 \to K^+\pi^-\mu^+\mu^-$ decays. The result of the applied veto is shown in Fig.5.3. The upper plot presents the plane of $m_{(\pi\to K)K}$) and $\pi DLL_{K\pi}$ for $B_s^0 \to \phi\mu^+\mu^-$ candidates in simulation before applying the veto. The lower plot in Fig 5.3 shows the impact of the corresponding veto on $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in simulation. The upper plot has a limited x axis range for better visualization.

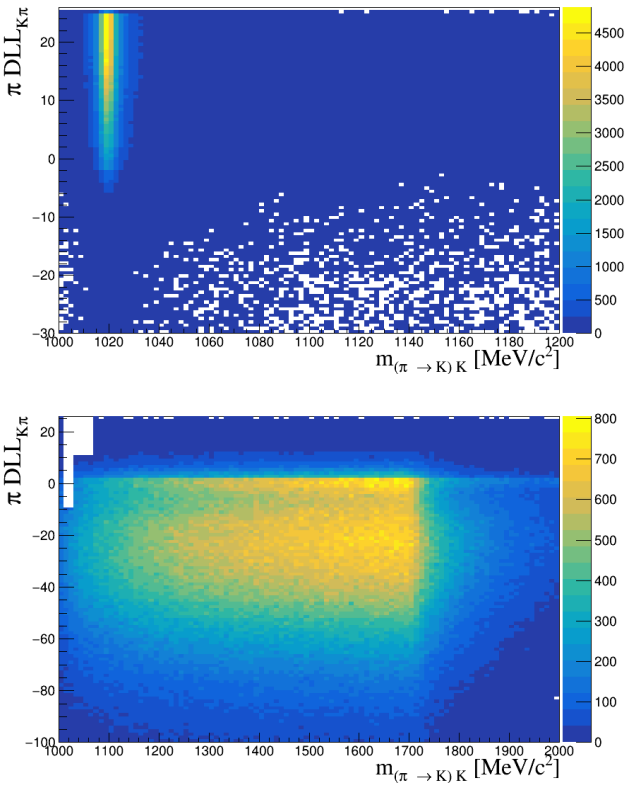


Figure 5.3: The 2D distribution of B_s^0 candidates mass $m_{(\pi\to K)K}$ and pion candidate $DLL_{K\pi}$ for simulated $B_s^0 \to \phi\mu^+\mu^-$ candidates before veto (upper) and for simulated signal $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates after veto (lower). Different x axis ranges are used for better visualization.

5.6.4 $B^+ \to K^+\mu^+\mu^-$ decay

Sometimes it is possible that a pion from elsewhere in the collision is added to a proper $B^+ \to K^+\mu^+\mu^-$ decay and forms the final state with four tracks. These candidates are expected to reside in the upper $m_{K\pi\mu\mu}$ sideband and the invariant mass $m_{K\mu\mu}$ should be consistent with the nominal B^0 mass. The following vetoes are used to reject those candidates:

$$m_{K\pi\mu\mu} > 5380\ \text{MeV/c}^2 \text{ and } 5220 < m_{K\mu\mu} < 5340\ \text{MeV/c}^2 \tag{5.5}$$

The result of the applied veto is shown in Figure 5.4. The upper plot presents the plane of $m_{K\mu\mu}$ and $m_{K\pi\mu\mu}$ for $B^+ \to K^+\mu^+\mu^-$ candidates in simulation before applying the veto. The lower plot in Fig 5.4 shows the impact of the corresponding veto on $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in simulation. The upper plot has a limited x axis range for better visualization.

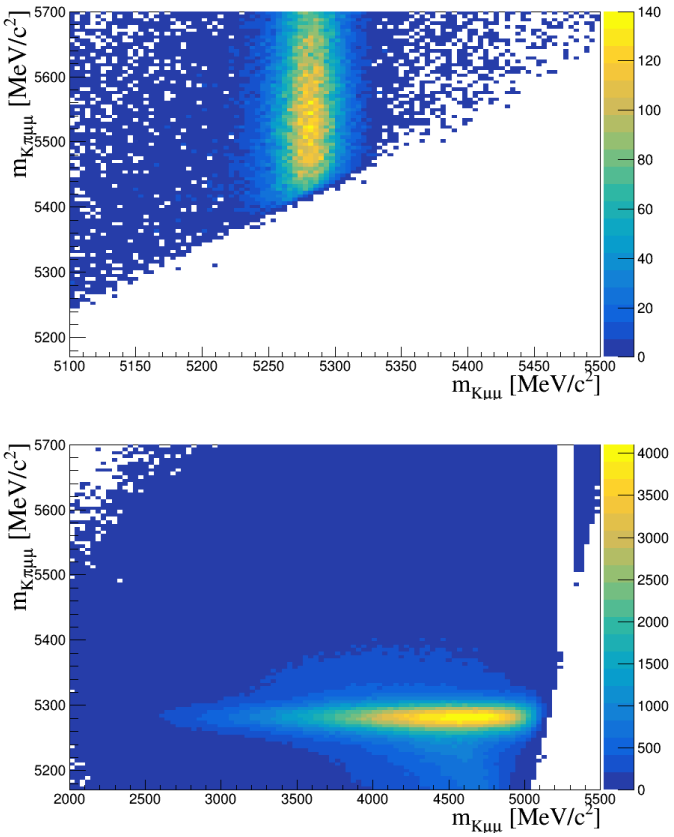


Figure 5.4: The 2D distributions of B^+ candidates mass $m_{K\mu\mu}$ and B^0 candidates mass $m_{K\pi\mu\mu}$ for simulated $B^+ \to K^+\mu^+\mu^-$ candidates before veto (upper) and simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates after veto (lower). Different x axis ranges are used for better visualization.

5.6.5 Misidentified $B^0 \to J/\psi K\pi$ and $B^0 \to \psi(2S)K\pi$

In the case when $\pi^-(K^+)$ is misidentified as $\mu^-(\mu^+)$ and $\mu^-(\mu^+)$ is misidentified as $\pi^-(K^+)$ decays $B^0 \to J/\psi K\pi$ and $B^0 \to \psi(2S)K\pi$ can contribute as a peaking background.
In the $\mu^- \leftrightarrow \pi^-$ scenario, after assigning a muon mass hypothesis to a pion, the invariant mass of π^- and the μ^+ should be consistent with the known J/ψ or $\psi(2S)$ masses. In a similar $\mu^+ \leftrightarrow K^+$ scenario, after assigning a muon mass hypothesis to a kaon , the invariant mass of K^+ and μ^- should also be consistent with the known J/ψ or $\psi(2S)$ masses. These new mass hypotheses are denoted as $m_{(\pi\to\mu)\mu}$ and $m_{(K\to\mu)\mu}$ respectively. The following vetoes to remove the misidentified $B^0 \to J/\psi K\pi$ candidates are used:

$$\begin{aligned} 2996 < m_{(\pi\to\mu)\mu} < 3196\ \mathrm{MeV/c^2} \text{ and } (\pi\mathrm{IsMuon}||\pi\mathrm{DLL}_{\mu\pi} > 0.0) \\ 2996 < m_{(K\to\mu)\mu} < 3196\ \mathrm{MeV/c^2} \text{ and } (\mathrm{KIsMuon}||\mathrm{KDLL}_{\mu\pi} > 0.0) \end{aligned} \tag{5.6}$$

The following vetoes to remove the misidentified $B^0 \to \psi(2S)K\pi$ candidates are used:

$$\begin{aligned} 3626 < m_{(\pi\to\mu)\mu} < 3746\ \mathrm{MeV/c^2} \text{ and } (\pi\mathrm{IsMuon}||\pi\mathrm{DLL}_{\mu\pi} > 5.0) \\ 3626 < m_{(K\to\mu)\mu} < 3746\ \mathrm{MeV/c^2} \text{ and } (\mathrm{KIsMuon}||\mathrm{KDLL}_{\mu\pi} > 5.0) \end{aligned} \tag{5.7}$$

5.6.6 Misidentified $B^0 \to K^+\pi^-\mu^+\mu^-$

The misidentification of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay, where two hadron hypotheses are swapped, can lead to a situation where B^0 candidates are incorrectly reconstructed as $\bar{B}^0$ candidates and vice versa. Candidates which have gone through double misidentification of the form $K \to \pi$ and $\pi \to K$ are rejected using the following cut:

$$\mathrm{KDLL}_{K\pi} - \pi\mathrm{DLL}_{K\pi} > 10, \tag{5.8}$$

whose effect on data is presented in Fig. 5.5.

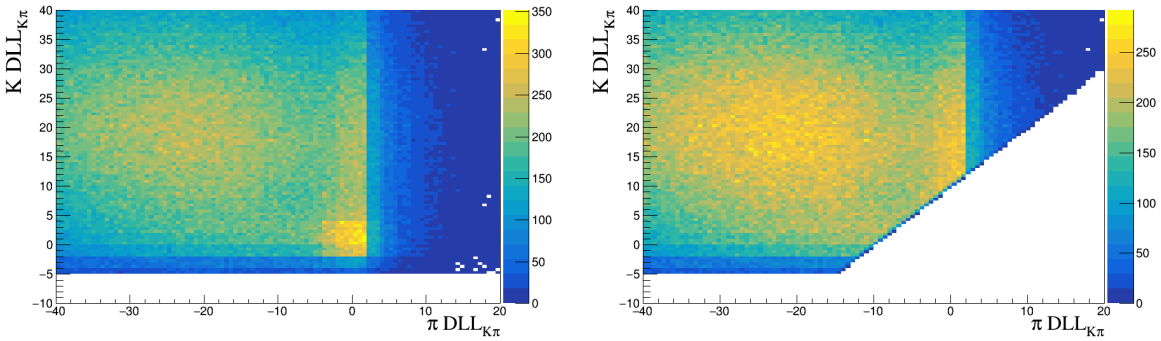


Figure 5.5: The 2D distributions of pion candidate $\mathrm{DLL}_{K\pi}$ and kaon candidate $\mathrm{DLL}_{K\pi}$ for simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates before veto (left) and after veto (right).

5.6.7 Other sources

There are two more modes which can contribute to the background. The first one is the decay $B^+ \to K^{*+}\mu^+\mu^-$ where $K^{*+} \to K^+\pi^0$. It can happen that simultaneously π^0 is not reconstructed and π^-, coming from elsewhere in the collision, is added to form the final state with four tracks. It is not possible to isolate those candidates from $B^0 \to K^{*0}\mu^+\mu^-$ with another requirement on invariant mass or DLL. No veto is used to remove this background contribution, which is expected to be at the level of 1.5% of the signal yield in the full $m_{K\pi\mu\mu}$ range and 0.5% within the signal mass window.

The second additional contribution comes from the $B^0 \to \rho^0\mu^+\mu^-$ decay, with $\rho^0(\to \pi^+\pi^-)$. It can contribute to peaking background if a pion is misidentified as a kaon. This decay is suppressed by the ratio of CKM factors:

$$|\frac{V_{cd}}{V_{cs}}|^2_{J/\psi} \approx 0.05 \ , \ |\frac{V_{td}}{V_{ts}}|^2_{\mu\mu} \approx 0.05, \tag{5.9}$$

where the additional subscripts J/ψ and $\mu\mu$ refer to the resonant and non-resonant modes, respectively. The contribution from the $B^0 \to \rho^0\mu^+\mu^-$ is expected to be less than 1% of the signal yield and for that reason it is neglected.

5.6.8 Efficiencies and residual yields

Using simulated MC samples, the efficiencies of the full set of peaking background vetoes are calculated for each relevant channel. Each efficiency, given in Tab. 5.6, is calculated for candidates fulfilling requirements given by the trigger, stripping, preselection and BDT, where the latter is described in Sec.5.9.

Even if all described vetoes were used, there would always be some remnants of peaking background candidates. To determine the residual yield of peaking background, firstly full data selection is applied to data apart from the specific veto. The yield of peaking background before applying the veto is obtained from maximum likelihood fit performed to a relevant variable. The yield is then combined with the corresponding efficiency taken from a simulation, giving in result the residual yield. In the case of $\Lambda_b^0 \to pK^-\mu^+\mu^-$ background contribution, with $p \to \pi$ swaps, the application of an inverse PID requirement of π $DLL_{p\pi} > 5$ in simulation is necessary to isolate the peaking component. To account for that, a scaling factor has to be calculated, because this additional requirement is absent in the data sample. The fraction of candidates in simulation selected by the inverse PID requirement is first determined, and its inverse is the scaling factor. The maximum likelihood fits to the peaking background candidates in data are shown in Figs. 5.6-5.11. The measured yield, efficiency of the veto, scaling factor and calculated residual yield of each peaking background component are shown in Tab. 5.6.

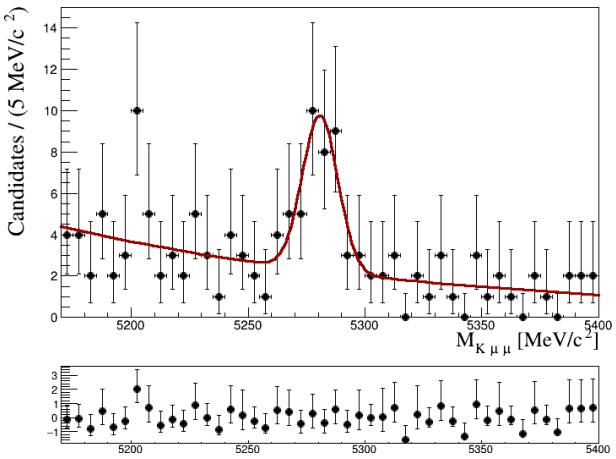


Figure 5.6: The maximum likelihood fit to extract the yield of $B^+ \to K^+\mu^+\mu^-$ not applying corresponding veto.

.

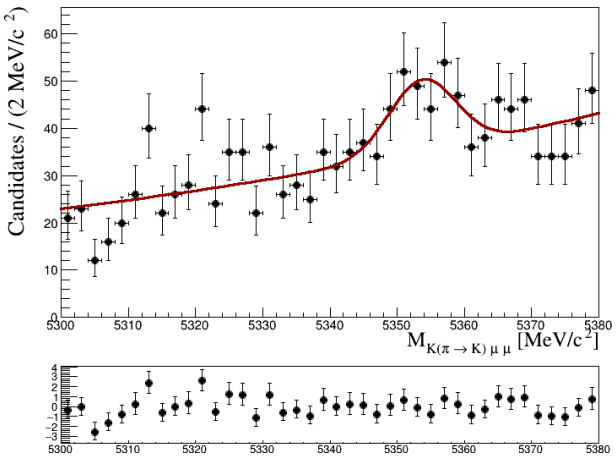


Figure 5.7: The maximum likelihood fit to extract the yield of $B_s^0 \to \phi\mu^+\mu^-$ with $(\pi \to K)$ swaps not applying corresponding veto.

.

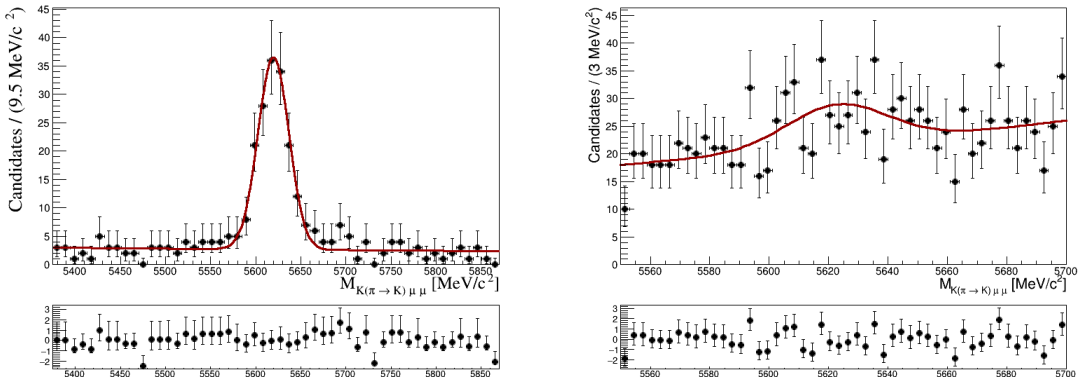


Figure 5.8: The maximum likelihood fit to extract the yield of $\Lambda_b^0 \to pK^-\mu^+\mu^-$ with $(\pi \to p)$ (left) and $(K \to p)(\pi \to K)$ (right) swaps not applying corresponding veto.

.

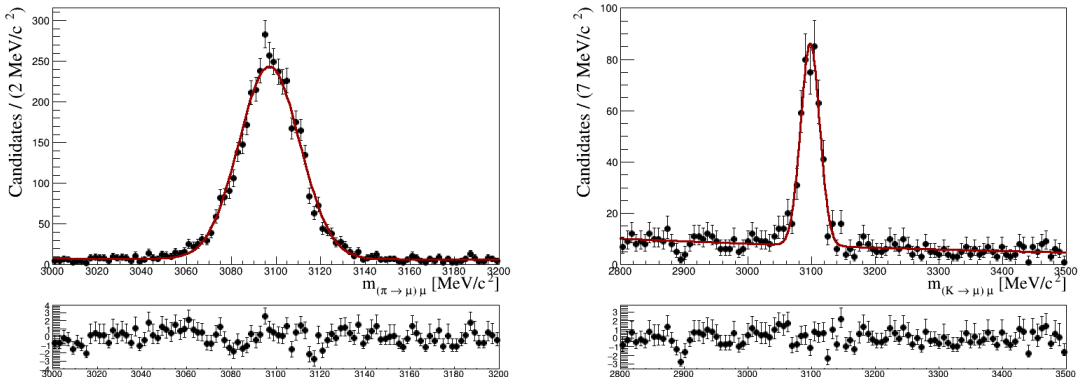


Figure 5.9: The maximum likelihood fit to extract the yield of $B^0 \to J/\psi K\pi$ with $(\pi \to \mu)$ (left) and $(K \to \mu)$ (right) swaps not applying corresponding veto.

.

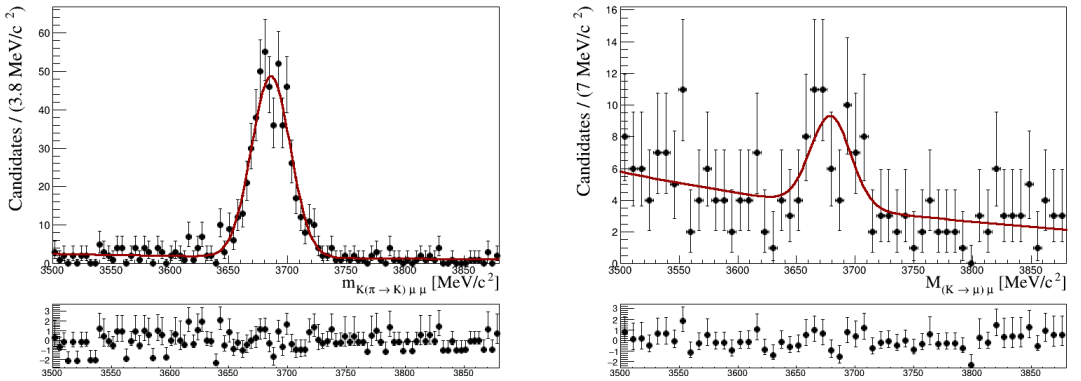


Figure 5.10: The maximum likelihood fit to extract the yield of $B^0 \to \psi(2S)K\pi$ with $(\pi \to \mu)$ (left) and $(K \to \mu)$ (right) swaps not applying corresponding veto.

.

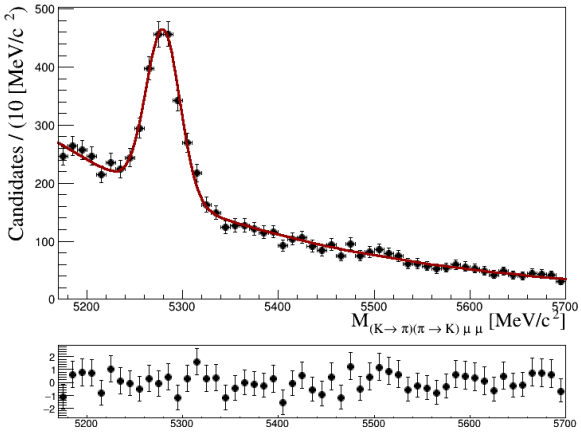


Figure 5.11: The maximum likelihood fit to extract the yield of $B^0 \to K^+\pi^-\mu^+\mu^-$ with $K \to \pi$ and $\pi \to K$ swaps not applying corresponding veto.

.

Decay	Yield	Efficiency[%]	Scaling factor	Residual yield
$\Lambda_b \to pK\mu^+\mu^-_{(p\to\pi)}$	152.8±14.0	8.17±0.07	1.05±0.01	13.1±1.2
$\Lambda_b \to pK\mu^+\mu^-_{(p\to\pi)(p\to K)}$	104.9±29.7	8.82±0.11	-	9.3±2.6
$B^0 \to K^+\pi^-\mu^+\mu^-$ swaps	1282.8±64.4	2.46±0.02	-	31.6±1.6
$B^0 \to J/\psi K\pi_{(\pi\to\mu)}$	4034.2±68.0	0.35±0.38	-	14.1±15.3
$B^0 \to J/\psi K\pi_{(K\to\mu)}$	420.2±23.4	0.34±0.16	-	1.4±0.7
$B^0 \to \psi(2S)K\pi_{(\pi\to\mu)}$	512.9±24.1	3.34±0.13	-	17.1±1.0
$B^0 \to \psi(2S)K\pi_{(K\to\mu)}$	34.8±9.3	2.57±0.09	-	0.9±0.2
$B^+ \to K^+\mu^+\mu^-$	28.8±7.3	6.26±0.15	-	1.8±0.5
$B^0_s \to \phi\mu^+\mu^-$	95.1±23.7	4.16±0.11	-	4.0±1.0

Table 5.6: The measured yield, efficiency of the veto, scaling factor and calculated residual yield of each peaking background component.

5.7 Control mode

For the purpose of removing most of the combinatorial background, the $B^0 \to J/\psi K^{*0}$ control mode with a high number of candidates is prepared.

In the data sample there are two sources of candidates: signal and background. By using likelihood fit of a known probability density function (PDF) of the so-called discriminating variables, it is possible to determine the yields of those sources. The $K\pi\mu\mu$ mass distribution model used here is a double Crystal Ball function:

$$f_{CB}(m;\mu,\sigma,\alpha,n) = N \times \begin{cases} e^{-\frac{1}{2}(\frac{m-\mu}{\sigma})^2} & \frac{m-\mu}{\sigma} > \alpha \\ (\frac{n}{|\alpha|})^n e^{-\frac{\alpha^2}{2}} \times (\frac{n}{|\alpha|} - |\alpha| - \frac{(m-\mu)}{\sigma})^{-n}) & \frac{m-\mu}{\sigma} < \alpha \end{cases}, \quad (5.10)$$

with the common mean of B^0 meson mass (μ), width (σ), tail parameters (α, n) and N as the normalization factor. Above $\alpha\sigma$ quantity Crystal Ball contains a Gaussian core and below $\alpha\sigma$ it has a power-law end tail. The combinatorial background is described by an exponential function. The total PDF is given by:

$$PDF_{tot} = n_{sig}[f_1 f_{CB}(m;\mu,\sigma_1,\alpha,n) + (1-f_1) f_{CB}(m;\mu,\sigma_2,\alpha,n)] + n_{bkg} e^{\lambda m}, \quad (5.11)$$

where n_{sig} stands for the number of fitted signal candidates, and n_{bkg} is the number of fitted background candidates, f_1 is the fraction of the first Crystal Ball function. To better model the B^0 mass distribution, an additional component is included to model the contribution from $B_s^0 \to J/\psi K^{*0}$, which has the same final state. After a fit to the mass distribution of the $B^0 \to J/\psi K^{*0}$ process (independently for 2016, 2017 and 2018 data sample Fig. 5.12) the *sPlot* [39] technique is used in order to calculate weights called *sWeights*. Those weights are derived based on the covariance matrix of the fit parameters and they correspond to the likelihood of the candidate belonging to the signal or background component. The application of the *sWeights* to the $B^0 \to J/\psi K^{*0}$ candidates in the data sample can statistically remove the majority of background without performing any cut. The data samples with $B^0 \to J/\psi K^{*0}$ candidates can be then considered as background free (Fig. 5.13).

Parameter	2016	2017	2018
n_{B^0}	370245±723	377646±722	464146 ±805
$n_{B_s^0}$	4675.18±156.04	4654.51 ± 149.95	5769.64±172.90
n_{bkg}	60136.0±482.49	58736.4±468.8	75962±526
μ [MeV/c^2]	5279.77± 0.04	5279.62±0.04	5280.49±0.04
σ_1 [MeV/c^2]	15.55±0.17	15.41±0.14	15.63±0.12
σ_2 [MeV/c^2]	26.37±0.66	26.71±0.68	27.38±0.72
α	1.38±0.01	1.35±0.01	1.34±0.01
n	100±98	120±101	120±89
f_1	0.69±0.03	0.75±0.02	0.78±0.02
λ [c^2/MeV]	-0.00350±0.00004	-0.00361±0.00004	-0.00358±0.00004

Table 5.7: Fit parameters from the fits to $B^0 \to J/\psi K^{*0}$ data candidates for each year.

Fit parameters from the $B^0 \to J/\psi K^{*0}$ fits are presented in the Tab 5.7.

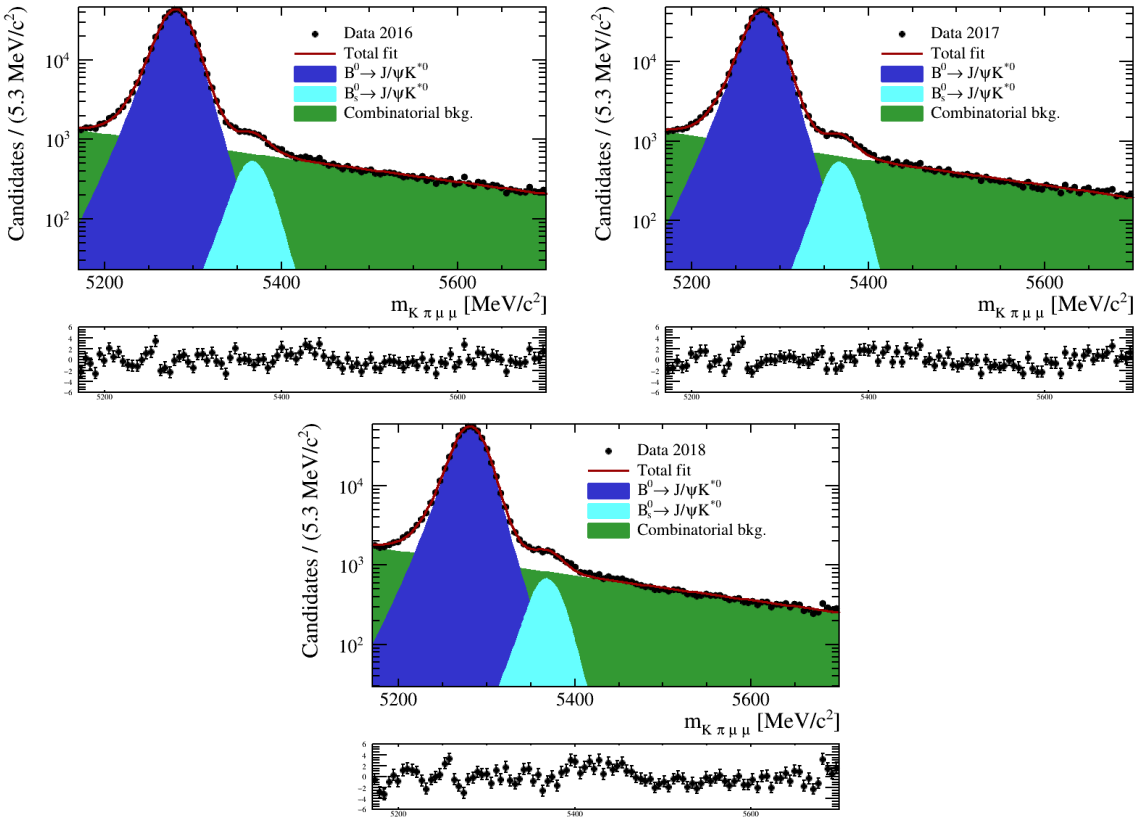


Figure 5.12: Fit to the $K\pi\mu\mu$ invariant mass of the control mode $B^0 \to J/\psi K^{*0}$ (blue area) and additional component $B_s^0 \to J/\psi K^{*0}$ (cyan area). Red line indicates total PDF. Combinatorial background is indicated by the green area. Fit performed on 2016, 2017 and 2018 data sample.

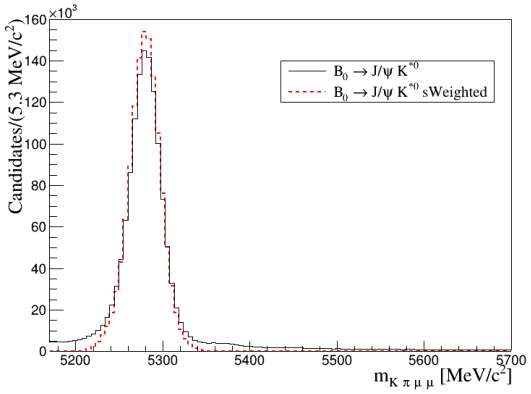


Figure 5.13: The $B^0 \to J/\psi K^{*0}$ mass distribution before (black) and after *sWeighting* (red) for the whole data sample.

5.8 Multivariate classifier

To remove the majority of the combinatorial background, a Toolkit for Multivariate Analysis (TMVA) software package is used. Classification is implemented with the boosted decision tree classifier (BDT) [40] with the Adaboost algorithm [41]. The Multivariate Analysis (MVA) presented here is an example of supervised learning where control samples need to be provided in order to distinguish between the signal and combinatorial background candidates.

To train the BDT classifier, two data samples are used. The first comprises the $B^0 \to K^{*0} J/\psi$ candidates weighted using an sPlot technique (Sec.5.7).Therefore, they are considered as free from background. The second sample comes from the upper mass sideband of the $B^0 \to K^{*0}\mu^+\mu^-$ used as combinatorial background. The discriminating variables used in the BDT are as follows:

- the B^0 candidate lifetime, momentum and transverse momentum,
- the B^0 direction angle, defined as the cosine of the angle between the direction of B^0 meson flight and the vector between the primary vertex and the B^0 decay vertex,
- the $\mathrm{K}^+\pi^-\mu^+\mu^-$ vertex χ^2, which represents the quality of the fit of the hypothesis that the final state particles originated from a common vertex,
- the $\mathrm{DLL}_{\mathrm{K}\pi}$ of the kaon and pion,
- the $\mathrm{DLL}_{\mu\pi}$ of the muons,
- the isolation variables, describing the separation of the signal tracks from tracks originating from other sources.

5.8.1 k-Folding of the data sample

For training the BDT classifier, the k-Fold technique is used. In this method all data samples are randomly divided into k=10 subsamples of equal size, both for the *sWeighted* $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$ sample, used as a signal proxy, and for the background sample. Each classifier is trained using 9 samples of signal and background and the response is evaluated on a corresponding sample which is omitted from the training (Fig. 5.14).

It is worth mentioning that in a standard case a data sample is split in half between the training and validation part. Here, instead of having 5:5 splitting, 9:1 splitting is used which helps to optimise the classifier better.

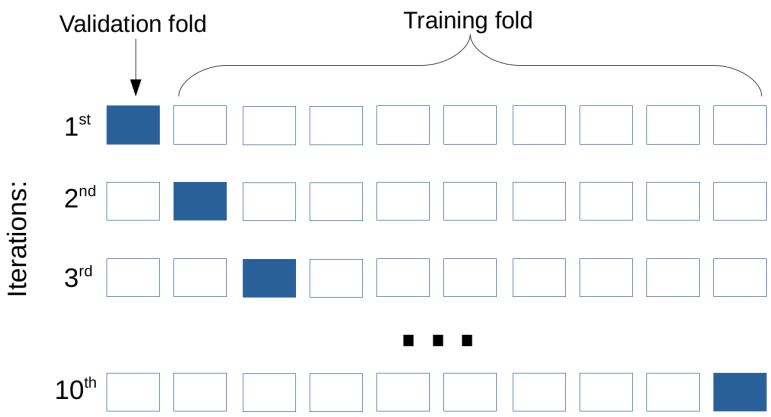


Figure 5.14: A schematic diagram of the k-Folding technique, where k=10. The whole data sample is split randomly into 10 subsets. The MVA for the first sample (blue) is trained and tested on rest of the samples (white), then this response is evaluated on this one sample. This method allows 90% of the dataset for every BDT training to be used in an unbiased way.

5.8.2 ROC curve and BDT response

The results of the training are depicted in Fig. 5.15, where an example of one of the folds is shown. The ROC curve (Receiver Operating Characteristic curve) shows the true positive rate (correctly identified signal) versus the false positive rate (background incorrectly identified as signal). The performance of the classifier is often characterized by the area under the curve (AUC), where the bigger AUC means the better classifier. The BDT response on the right side of the Fig. 5.15 shows how candidates were divided into the signal and background class. The BDT overtraining check is presented in Appendix B.

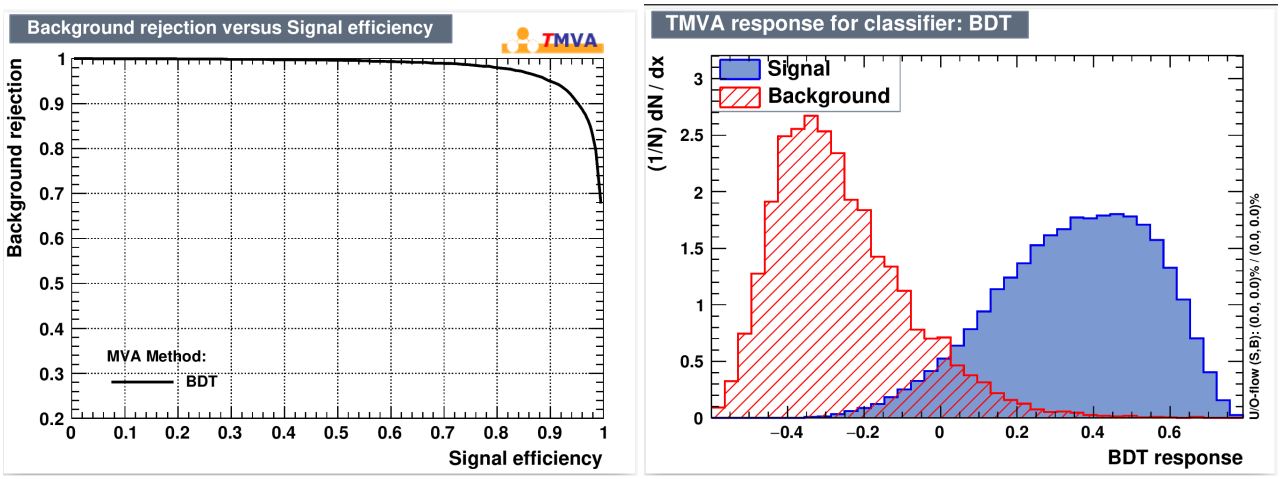


Figure 5.15: ROC curve(left) and BDT response(right) for one fold of the test set.

5.9 BDT optimization

The BDT is optimised by fitting the $B^0 \to J/\psi K^{*0}$ candidates in three regions:

- the signal region - $m_{K\pi\mu\mu} \in [5230, 5330]$ MeV/c^2,
- the lower mass sideband - $m_{K\pi\mu\mu} \in [5000, 5180]$ MeV/c^2,
- the upper mass sideband - $m_{K\pi\mu\mu} \in [5500, 7000]$ MeV/c^2,

using several BDT requirements from -0.15 to 0.5.

For signal model, the double Crystal Ball function is used, and the exponential function models the background (Fig.5.16). The signal yield, obtained from the fit, is scaled by the ratio of the total efficiency taken from simulation and the ratio of branching fractions between $B^0 \to J/\psi K^{*0}$ and $B^0 \to K^{*0}\mu^+\mu^-$, taken from the PDG [42]:

$$\mathrm{nSig}_{\mu\mu} = \mathrm{nSig}_{J/\psi} \frac{\epsilon^{MC}_{\mu\mu}}{\epsilon^{MC}_{J/\psi}} \frac{\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{\mathcal{B}(B^0 \to J\psi K^{*0})\mathcal{B}(J\psi \to \mu^+\mu^-)}, \tag{5.12}$$

where the subscripts $\mu\mu$ and J/ψ indicate the non-resonant and resonant modes, respectively.

Background yield is obtained from the fit to both lower and upper sidebands and then extrapolated to the signal region. Fits are performed for each given BDT requirement (Appendix B.1). Using the number of signal candidates $\mathrm{nSig}_{\mu\mu}$ and the number of background candidates $\mathrm{nBkg}_{\mu\mu}$ from each fit, a figure of merit (FOM) can be calculated using the formula:

$$\mathrm{FOM} = \frac{\mathrm{nSig}_{\mu\mu}}{\sqrt{\mathrm{nSig}_{\mu\mu} + \mathrm{nBkg}_{\mu\mu}}}, \tag{5.13}$$

which determines the performance of the classification cut.

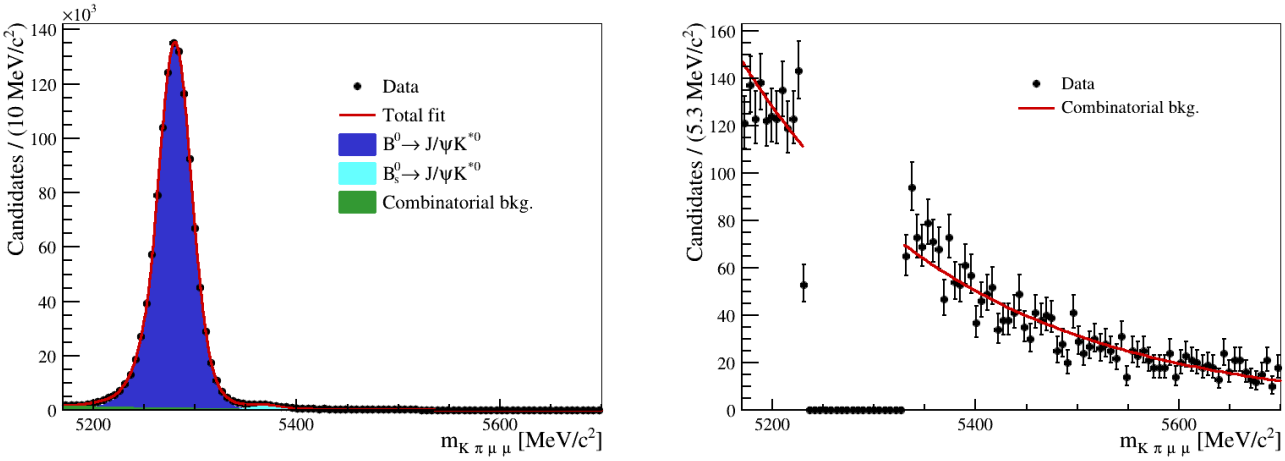


Figure 5.16: Fits to $B^0 \to J/\psi K^{*0}$ (left) and $B^0 \to K^{*0}\mu^+\mu^-$ two sidebands (right) used in the BDT optimisation. Here is an example for BDT>0.05.

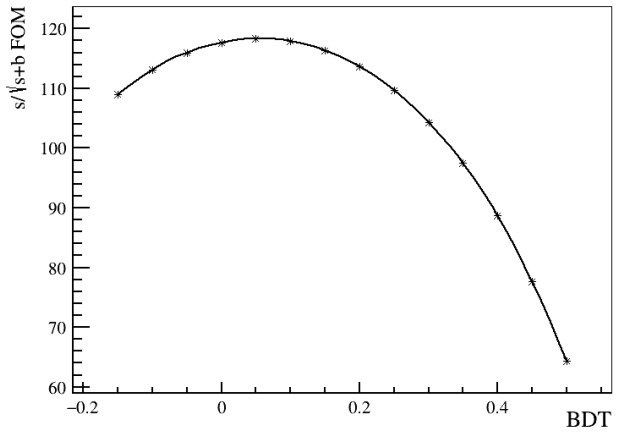


Figure 5.17: The figure of merit, for the best BDT performance.

In order to have the most efficient BDT cut with the best separation between the signal and background candidates, one has to choose the highest value of the FOM (Fig. 5.17). Since the FOM is relatively stable around the optimum value, the results of the $B^0 \to K^+\pi^-\mu^+\mu^-$ analysis are not biased by a certain BDT cut value. Small changes in the BDT cut value around the optimum region have little effect on the FOM. From all calculated significances the highest one is chosen, which corresponds to the BDT cut value equal to 0.05.

Chapter 6

Data and simulation agreement

The use of simulated data samples throughout the analysis is crucial for the best modeling of the angular distributions distortions caused by the trigger, reconstruction and selection. In order to have the compatibility between the data and simulation, the particle identification (PID) distributions of simulated candidates are corrected using the so-called PID resampling method and after that the simulated candidates are properly reweighted to match the distributions in the data.

6.1 Particle identification resampling

Each PID variable distribution of a kaon, pion and muon is resampled using the MCResampling package in Urania/PIDCalib [43]. This package provides high purity data samples containing the $D^{*+} \to D^0(\to K^-\pi^+)\pi^-$, $\Lambda \to p\pi^-$ and $J/\psi \to \mu^+\mu^-$ decays.

The resampling consists of two stages:

- histograms of PID variables are produced in bins of the number of tracks in the decay (nTracks), pseudorapidity (η) and transverse momentum (p_T),
- produced histograms are used as a PDF to draw a new value. This new value replaces the PID variable for the simulated candidate.

Next, the original PID distribution is replaced by the resampled one, which is then used throughout the whole analysis. Additionally for muons, the PID correction is done only for those particles with $p_T > 700$ GeV/c, since the corresponding PIDCalib sample holds the same requirement. For the candidates with $p_T < 700$ GeV/c the original PID is kept. The resampling procedure is applied to MC simulated samples for each year and each magnet polarity. It is important that the resampling procedure has to be done before any requirement is applied on PID variables. Therefore the resampling takes place before the preselection described in Sec. 5.5. The method is validated using sWeighted $B^0 \to J\psi K^{*0}$ candidates taken from the data and simulated $B^0 \to J\psi K^{*0}$ candidates, which is presented in Fig. 6.1 and Fig. 6.2.

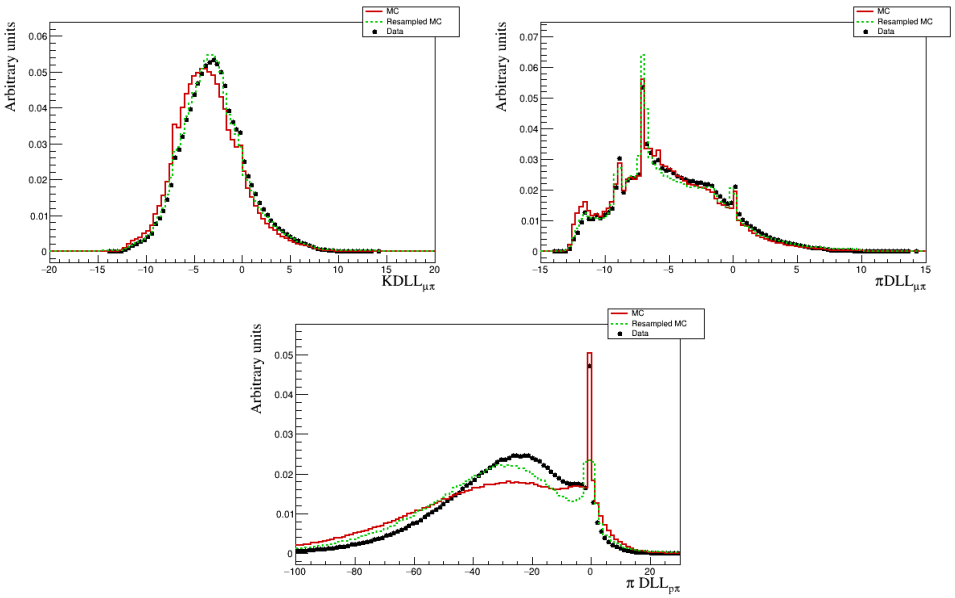


Figure 6.1: Agreement between data and simulation for the PID variables used to veto peaking background contributions. Black data points correspond to the sWeighted $B^0 \to J/\psi K^{*0}$ data. Red histogram shows the distribution in $B^0 \to J/\psi K^{*0}$ MC, and green histogram shows the same distribution after resampling.

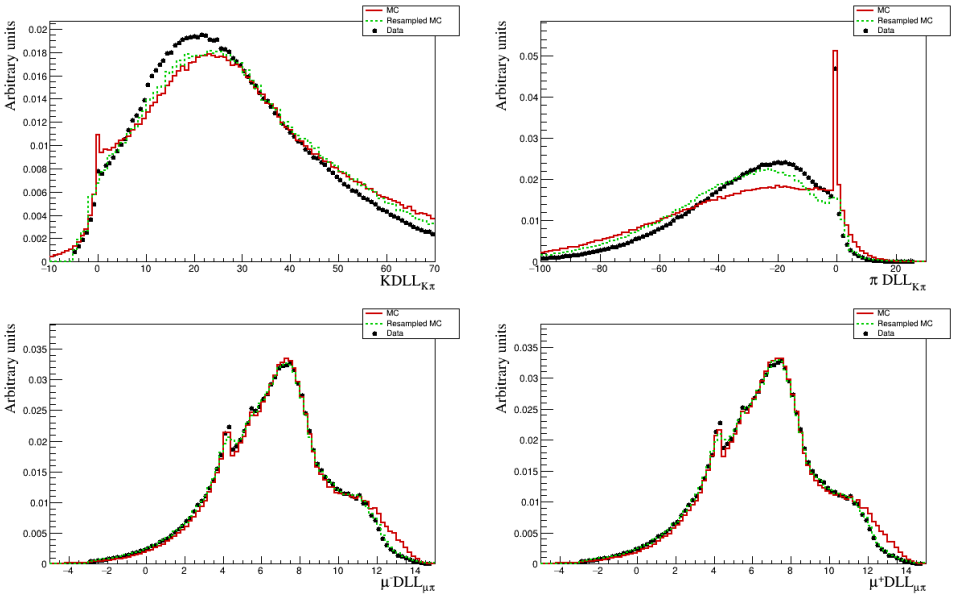


Figure 6.2: Agreement between data and simulation for the PID variables used to train the BDT classifier. Black data points correspond to the sWeighted $B^0 \to J/\psi K^{*0}$ data. Red histogram shows the distribution in $B^0 \to J/\psi K^{*0}$ MC, and green histogram shows the same distribution after resampling.

6.2 Reweighting of the candidates

To get a better agreement between the data and simulation, the candidates weights are derived by comparing the *sWeighted* decay $B^0 \to J/\psi K^{*0}$ from data with the simulated $B^0 \to J/\psi K^{*0}$ using three variables (Fig. 6.3):

- the transverse momentum p_T of the B^0 candidate,
- the number of tracks in the decay,
- quality of the $K^+\pi^-\mu^+\mu^-$ vertex.

Two requirements on the invariant mass of $K\pi$ system and q^2 have to be satisfied: $796 < m_{K\pi} < 996$ MeV/c^2 and $9.22 < q^2 < 9.96$ GeV^2/c^4. The weights are determined sequentially, where the previous weight is applied before the subsequent weight is derived. The obtained weights are applied to all simulation samples. Figure 6.4 shows the results of this method on three variables used to determine the candidate weights. Figure 6.5 shows the agreement between data and simulation of the BDT response before and after reweighting. The results for each variable used for the BDT training are presented in Appendix B.2.

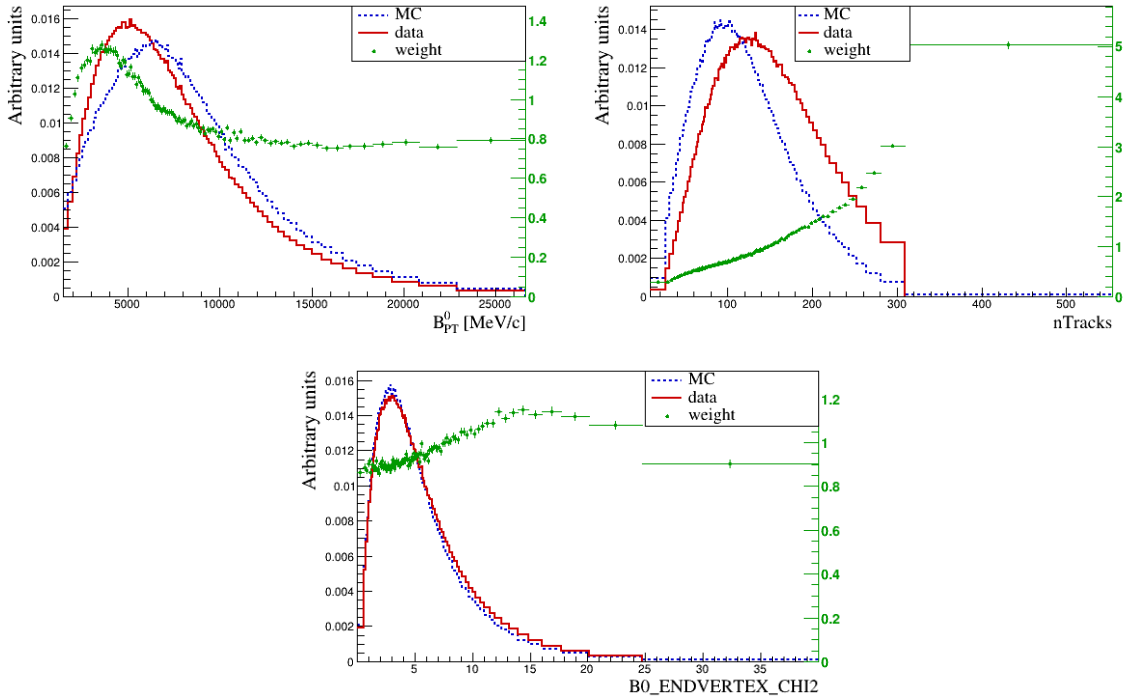


Figure 6.3: Distributions of three variables: nTracks, p_T and finals state quality vertex used to derive weights. Red histogram corresponds to *sWeighted* $B^0 \to J/\psi K^{*0}$ data, the blue histogram shows $B^0 \to J/\psi K^{*0}$ MC and the green points represent the calculated weights.

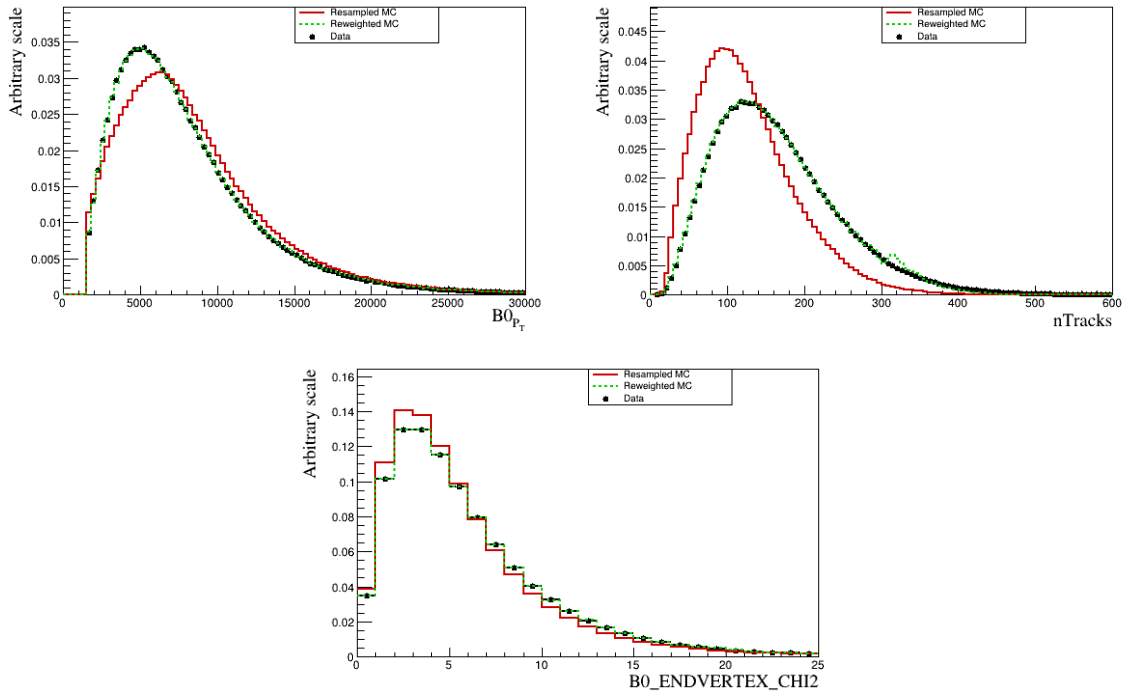


Figure 6.4: Agreement between data and simulation for three variables used to determine the candidate weights. The black points correspond to *sWeighted* $B^0 \to J/\psi K^{*0}$ data. Red histogram shows the resampled $B^0 \to J/\psi K^{*0}$ MC distribution. Green histogram shows the distribution for resampled $B^0 \to J/\psi K^{*0}$ MC with applied weights.

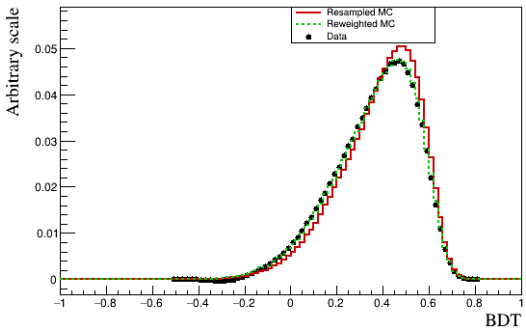


Figure 6.5: Agreement between data and simulation for the BDT response. Black points correspond to *sWeighted* $B^0 \to J/\psi K^{*0}$ data. Red histogram shows the resampled $B^0 \to J/\psi K^{*0}$ MC distribution. Green histogram shows the distribution for resampled $B^0 \to J/\psi K^{*0}$ MC with applied weights.

6.3 Cross check with the $B^0 \to J/\psi K\pi$ candidates

This analysis is focused on the ranges of $m_{K\pi}$ and q^2 different from those used for obtaining the weights. Namely, the weights are derived in the $B^0 \to J/\psi K^*(892)^0$ region and are applied to the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC candidates where the interesting region is quite different ($1330 < m_{K\pi} < 1530$ MeV/c^2 and $0.1 < q^2 < 8.0$ GeV^2/c^4). To be completely sure that the agreement between data and simulation is maintained in the higher $m_{K\pi}$ region, two sets of data are compared: *sWeighted* $B^0 \to J/\psi K\pi$ taken from data and reweighted $B^0 \to K^+\pi^-\mu^+\mu^-$ taken from simulation, where both meet the same selection requirements ($1330 < m_{K\pi} < 1530$ MeV/c^2 and $9.22 < q^2 < 9.96$ GeV^2/c^4). The BDT distribution for those two samples is shown in Fig. 6.6, where a good agreement between the two samples can be seen.

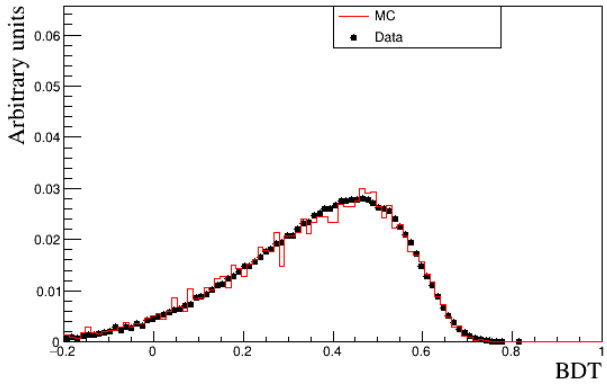


Figure 6.6: Cross check of the agreement between data and simulation of the BDT response. Black points represent the distributions for *sWeighted* $B^0 \to J/\psi K^{*0}$ data in higher $m_{K\pi}$ region and red histogram correspond to resampled $B^0 \to K^+\pi^-\mu^+\mu^-$ MC with applied weights.

Chapter 7

Angular acceptance correction

The distortions in the distributions of q^2 , $\cos\theta_K$, $\cos\theta_\ell$, ϕ and $m_{K\pi}$ caused by the triggering, reconstruction and selection are taken into account by calculating the weight as the inverse of the efficiency:

$$\omega = \frac{1}{\epsilon(q^2, cos\theta_K, cos\theta_\ell, \phi, m_{K\pi})}, \tag{7.1}$$

where ϵ is a five dimensional efficiency parameterized in terms of Legendre polynomials $P_n(x)$ of order n:

$$\epsilon(q^{2'}, cos\theta_K, cos\theta_\ell, \phi', m'_{K\pi}) = \sum_{hijkl} c_{hijkl} P_h(q^{2'}) P_i(cos\theta_K) P_j(cos\theta_\ell) P_k(\phi') P_l(m'_{K\pi}). \tag{7.2}$$

As the acceptance correction has to be used for both the differential branching fraction measurements and the angular analysis, it is calculated in a broad range of q^2 and $m_{K\pi}$ ($0.1 < q^2 < 10.0$ GeV^2/c^4, $795 < m_{K\pi} < 1530$ MeV/c^2) in order to cover the full range of the studied invariant masses. A similar weight is used in e.g. Ref.[13], but here additional $m_{K\pi}$ dimension has to be added because in the wider $m_{K\pi}$ window the efficiency distribution is no longer flat. Owing to the fact that Legendre polynomials P(x) are defined in the range $-1 < x < +1$, the observables q^2, ϕ and $m_{K\pi}$ have to be transformed to this range as well and are denoted as $q^{2'}$, ϕ' and $m'_{K\pi}$. The Legendre polynomials in $cos\theta_\ell$ and $m'_{K\pi}$ are up to third order, in $q^{2'}$ and ϕ' up to fifth order and in $\cos\theta_K$ up to seventh order.

7.1 Phase space distributions

Since the distributions of the three variables $cos\theta_K$, $cos\theta_\ell$ and ϕ are flat at the MC generator level, after the reconstruction and selection the shape of these distributions will become the shape of the efficiency. The distributions of q^2 and $m_{K\pi}$ are non-flat at the MC generator level. To get the efficiency shape for these two variables, the two dimensional distribution of q^2 and $m_{K\pi}$ (Fig. 7.1) is used from which weights are determined as the inverse of 2D distribution. Those weights help to transform the reconstructed distributions

of q^2 and $m_{K\pi}$. The MC generator level and the reconstructed distributions for those two variables before and after applying the weights are shown in Fig. 7.2 and Fig. 7.3, respectively.

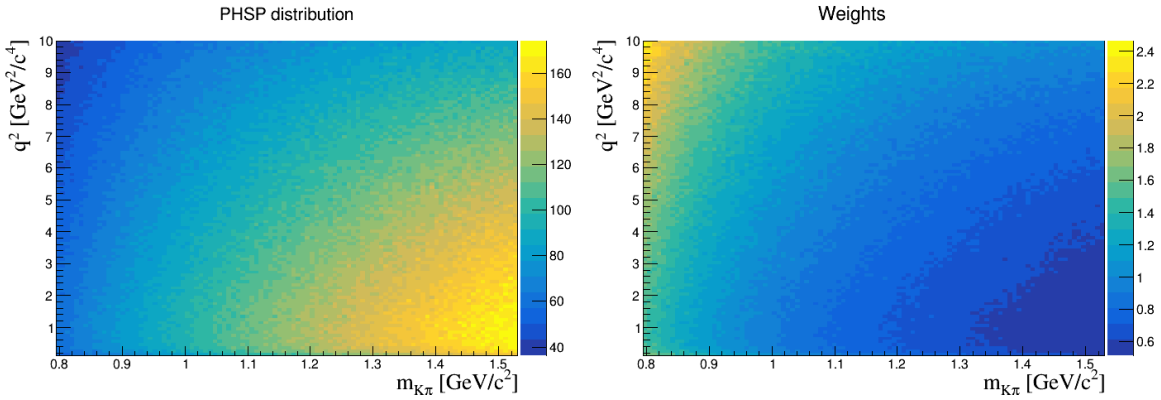


Figure 7.1: 2D distributions of $q^2 - m_{K\pi}$ (left) and derived weights (right) at MC generator level.

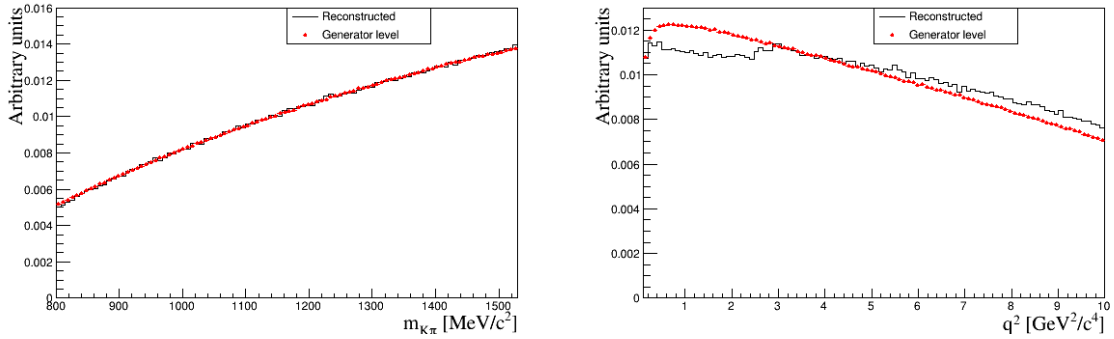


Figure 7.2: Distributions of $m_{K\pi}$ (left)and q^2 (right) before applying candidate weights to account for non flat distribution. Red points depict the MC generator level candidates and black histogram the reconstructed candidates.

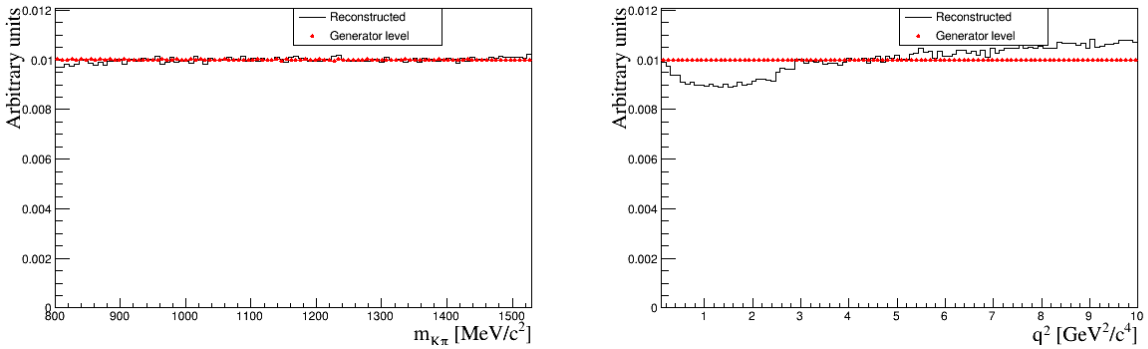


Figure 7.3: Distributions of $m_{K\pi}$ (left)and q^2 (right) after applying candidate weights to account for non flat distribution. Red points depict the MC generator level candidates and black histogram the reconstructed candidates.

7.2 Method of moments

The coefficients c_{hijkl} (Eq.7.2) are determined from the simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ decays using the method of moments:

$$c_{hijkl} = \frac{1}{\sum \omega_n} \sum_{n=0}^{N} \omega_n (\frac{2h+1}{2})(\frac{2i+1}{2})(\frac{2j+1}{2})(\frac{2k+1}{2})(\frac{2l+1}{2}) \tag{7.3}$$

$$\times P(q^{2'}, h) P(cos\theta_K, i) P(cos\theta_\ell, j) P(\phi', k) P(m'_{K\pi}, l),$$

where ω_n is the per-candidate weight which takes into account weights from the non-flat phase space distribution of q^2 and $m_{K\pi}$ (Sec.7.1), and the kinematic candidate weights (Sec.6.2). The factors (2a+1)/2 come from the orthogonality of the Legendre polynomials:

$$\int_{-1}^{1} P(x,a)P(x,a')dx = \frac{2}{2a+1}\delta_{aa'}. \tag{7.4}$$

The number of coefficients that have to be calculated depends on the orders of every Legendre polynomial. There are two relations helping to reduce the number of coefficients. Firstly, cosθ_ℓ and ϕ are assumed to be even, so only the even orders of the corresponding Legendre polynomials are kept. Secondly, both cosθ_ℓ and $m_{K\pi}$ are assumed to factorize, so only the terms with only cosθ_ℓ or only $m_{K\pi}$ are calculated. This approach reduces the number of coefficients from 4608 to just 720, thus reducing the systematic uncertainty.

7.3 Projections

The agreement between the shape of parametrisation with the reweighted simulated candidates is presented in one-dimensional projections of 5D acceptance parametrisation. The comparison of the 5D acceptance parametrisation for each of the variables is shown

in Fig. 7.4. A goood agreement is shown between the parametrisation and the reweighted simulated candidates. The one-dimensional projections in bins of $m_{K\pi}$ and q^2 are presented in Appendix C.

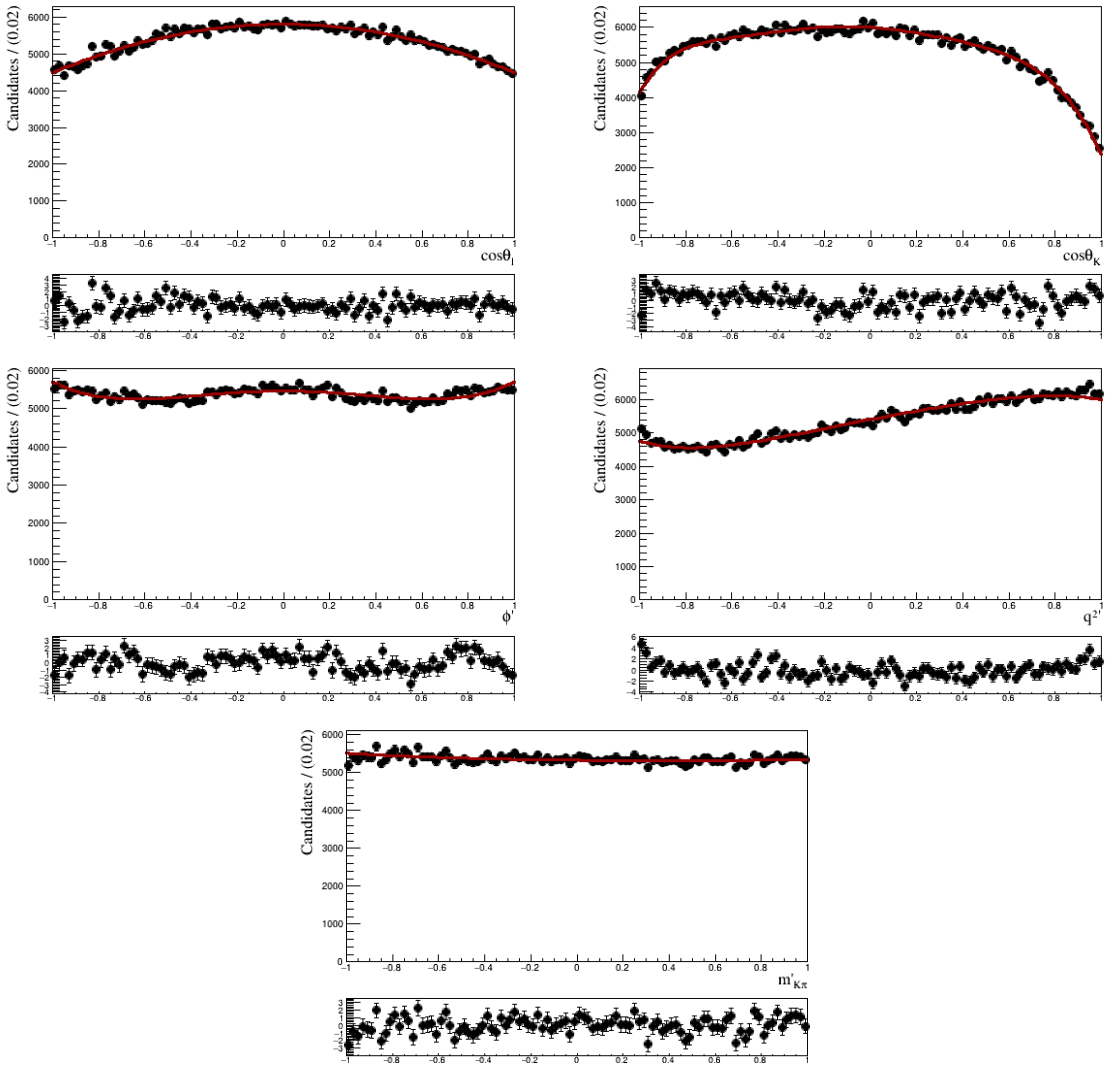


Figure 7.4: 1D projections of the 5D acceptance parametrization for q^2 , $cos\theta_K$, $cos\theta_\ell$ and ϕ, and $m_{K\pi}$. The black data points indicate the reweighted simulated candidates and the red curve shows the parametrization.

7.4 Distribution of the weights in data

It is possible that some weights can be very large or even negative in certain parts of the phase space. This is the case for $|\cos\theta_\ell| \to 1$ at low q^2 and for $\cos\theta_K \to 1$. In order to take that into account, the weights in data are required to be in the range [0,5].

As it is shown in Fig. 7.5, the weight distribution is investigated in two regions. The first region contains the $B^0 \to J/\psi K^{*0}$ candidates, and the second region contains the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates. A requirement of having weights in the range [0,5] removes 0.18% of the candidates in the first region and has no effect on the second region.

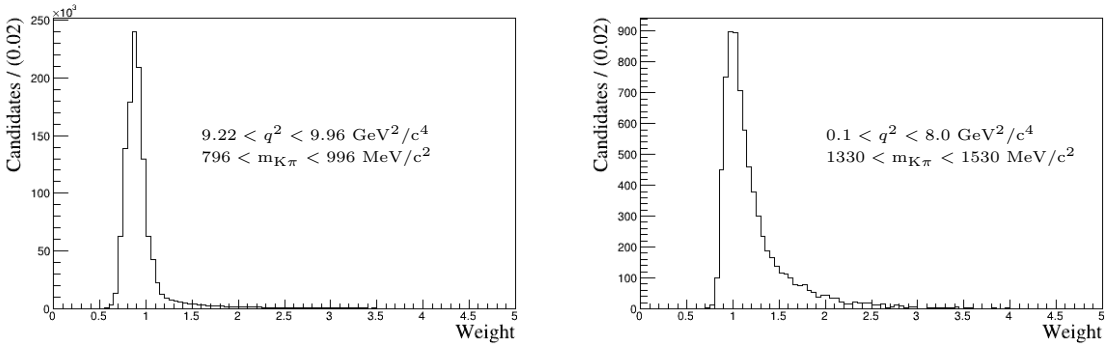


Figure 7.5: Weight distribution in data for $B^0 \to J/\psi K^{*0}$ candidates (left) and $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates (right).

7.5 Cross check with the $B^0 \to K^{*0}\mu^+\mu^-$ candidates

The acceptance correction has been cross-checked by calculating several observables: F_L, S_3, S_4, S_5, A_{FB}, S_7, S_8 and S_9 as a function of q^2 described in Sec. 3.2 and by comparing the results with the CP-averages measured in [13] (Fig. 7.6). Only statistical uncertainty is taken into account in the new results. In most q^2 regions a good agreement is shown between those two measurements. The discrepancies have their sources in the differences between the data selections and in the methods used in the analyses. In [13] the observables are obtained using likelihood fit and the current measurements are based on the method of moments described in Sec. 3.3.

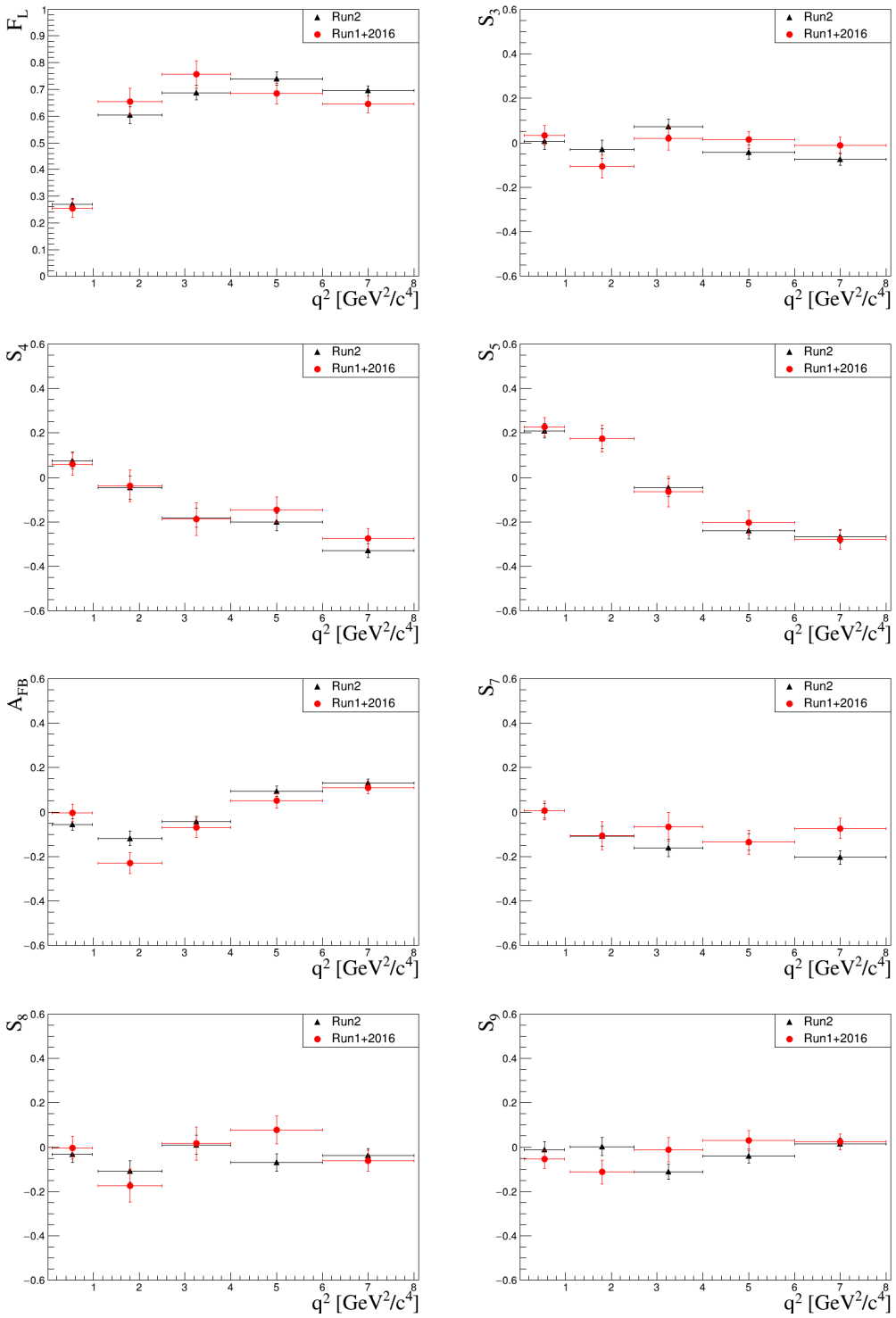


Figure 7.6: Comparison between measured CP-averaged observables from [13](red circle) and current one(black triangle).

Chapter 8

Angular analysis of B decays

The CP-averaged differential decay rate of the B^0 and $\bar{B}^0$ decays with the $K\pi$ system in the S-, P-, and D-wave configurations is expanded in an orthonormal basis of angular functions $f_i(\Omega)$ [16]:

$$\begin{aligned} \frac{d^4\Gamma}{dq^2 d\Omega} &\propto \sum_{i=1}^{41} M_i(q^2) f_i(\Omega), \\ M_i(q^2) &= M_i^L(q^2) + \eta_i^{L\to R} M_i^R(q^2) \end{aligned} \tag{8.1}$$

where $d\Omega = d\cos\theta_K d\cos\theta_\ell d\phi$, $f_i(\Omega)$ are constructed out of spherical harmonics $Y_l^m(\theta_\ell, \phi)$ and reduced spherical harmonics $P_l^m \equiv \sqrt{2\pi} Y_l^m(\theta_K, 0)$. The letters L and R stand for left- and right-handed chirality of the dimuon system. The sign of $\eta_i^{L\to R}$ depends on the f_i's sign when $\theta_\ell \to \pi + \theta_\ell$. The angular observables are measured in $1330 < m_{K\pi} < 1530$ MeV/c^2 range in a large-recoil region $1.1 < q^2 < 6.0$ GeV^2/c^4. The basis of the transversity amplitudes is defined in terms of the helicity amplitudes as:

$$\mathcal{H}_{\parallel,\perp}^J = \frac{\mathcal{H}_+^J \pm \mathcal{H}_-^J}{\sqrt{2}}, \tag{8.2}$$

and 41 moments in this basis are shown in Tab. 8.1. The transversity amplitudes correspond to the $\bar{B}^0$ decay and amplitudes for the B^0 decay are obtained by flipping the signs of the helicities and weak phases. The transversity amplitudes of the S-, P- and D-wave are denoted as $S^{\{L,R\}}$, $H_{\{0,\parallel,\perp\}}^{\{L,R\}}$ and $D_{\{0,\parallel,\perp\}}^{\{L,R\}}$, respectively.

8.1 Method of moments

In some previous analyses of the $B^0 \to K^{*0}\mu^+\mu^-$ decay [13, 19], the determination of angular observables is based on the simultaneous fit to the mass distribution $m(K^+\pi^-\mu^+\mu^-)$ and three decay angles $\cos\theta_K$, $\cos\theta_\ell$ and ϕ. In this analysis the method of moments is used to extract angular observables, as described in Sec. 3.3. This method takes into account the fact of orthonormality of spherical harmonics. The advantage of this approach over the maximum likelihood fit is that it can provide good results even for a small data sample. However, maximum likelihood fit gives more accurate results.

i	$f_i(\Omega)$	$M_i^{L,tr}(q^2)$	$\eta_i^{L\to R}$
1	$P_0^0 Y_0^0$	$\|H_0^L\|^2+\|H_k^L\|^2+\|H_\perp^L\|^2+\|S^L\|^2+\|D_0^L\|^2+\|D_\parallel^L\|^2+\|D_\perp^L\|^2$	+1
2	$P_1^0 Y_0^0$	$2[\frac{2}{\sqrt{5}}Re(H_0^L D_0^{L*})+Re(S^L H_0^{L*})+\sqrt{\frac{3}{5}}Re(H_\parallel^L D_\parallel^{L*}+H_\perp^L D_\perp^{L*})]$	+1
3	$P_2^0 Y_0^0$	$\frac{\sqrt{5}}{7}(\|D_\parallel^L\|^2+\|D_\perp^L\|^2)-\frac{1}{\sqrt{5}}(\|H_\parallel^L\|^2+\|H_\perp^L\|^2)+\frac{2}{\sqrt{5}}\|H_0^L\|^2+\frac{10}{7\sqrt{5}}\|D_0^L\|^2+2Re(S^L D_0^{L*})$	+1
4	$P_3^0 Y_0^0$	$\frac{6}{\sqrt{35}}[-Re(H_\parallel^L D_\parallel^{L*}+H_\perp^L D_\perp^{L*})+\sqrt{3}Re(H_0^L D_0^{L*})]$	+1
5	$P_4^0 Y_0^0$	$\frac{2}{7}[-2(\|D_\parallel^L\|^2+\|D_\perp^L\|^2)+3\|D_0^L\|^2]$	+1
6	$P_0^0 Y_2^0$	$\frac{1}{2\sqrt{5}}[(\|D_\parallel^L\|^2+\|D_\perp^L\|^2)+(\|H_\parallel^L\|^2+\|H_\perp^L\|^2)-2\|S^L\|^2-2\|D_0^L\|^2-2\|H_0^L\|^2]$	+1
7	$P_1^0 Y_2^0$	$[\frac{\sqrt{3}}{5}Re(H_\parallel^L D_\parallel^{L*}+H_\perp^L D_\perp^{L*})-\frac{2}{\sqrt{5}}Re(S^L H_0^{L*})-\frac{4}{5}Re(H_0^L D_0^{L*})]$	+1
8	$P_2^0 Y_2^0$	$[\frac{1}{14}(\|D_\parallel^L\|^2+\|D_\perp^L\|^2)-\frac{2}{7}\|D_0^L\|^2-\frac{1}{10}(\|H_\parallel^L\|^2+\|H_\perp^L\|^2)-\frac{2}{5}\|H_0^L\|^2-\frac{2}{\sqrt{5}}Re(S^L D_0^{L*})]$	+1
9	$P_3^0 Y_2^0$	$-\frac{3}{5\sqrt{7}}[Re(H_\parallel^L D_\parallel^{L*}+H_\perp^L D_\perp^{L*})+2\sqrt{3}Re(H_0^L D_0^{L*})]$	+1
10	$P_4^0 Y_2^0$	$-\frac{2}{7\sqrt{5}}[\|D_\parallel^L\|^2+\|D_\perp^L\|^2+3\|D_0^L\|^2]$	+1
11	$P_1^1\sqrt{2}Re(Y_2^1)$	$-\frac{3}{\sqrt{10}}[\sqrt{\frac{2}{3}}Re(H_\parallel^L S^{L*})-\sqrt{\frac{2}{15}}Re(H_\parallel^L D_0^{L*})+\sqrt{\frac{2}{5}}Re(D_\parallel^L H_0^{L*})]$	+1
12	$P_2^1\sqrt{2}Re(Y_2^1)$	$-\frac{3}{5}[Re(H_\parallel^L H_0^{L*})+\sqrt{\frac{5}{3}}Re(D_\parallel^L S^{L*})+\frac{5}{7\sqrt{3}}Re(D_\parallel^L D_0^{L*})]$	+1
13	$P_3^1\sqrt{2}Re(Y_2^1)$	$-\frac{6}{5\sqrt{14}}[2Re(D_\parallel^L H_0^{L*})+\sqrt{3}Re(H_\parallel^L D_0^{L*})]$	+1
14	$P_4^1\sqrt{2}Re(Y_2^1)$	$-\frac{6}{7\sqrt{2}}Re(D_\parallel^L D_0^{L*})$	+1
15	$P_1^1\sqrt{2}Im(Y_2^1)$	$3[\frac{1}{\sqrt{15}}Im(H_\perp^L S^{L*})+\frac{1}{5}Im(D_\perp^L H_0^{L*})-\frac{1}{5\sqrt{3}}Im(H_\perp^L D_0^{L*})]$	+1
16	$P_2^1\sqrt{2}Im(Y_2^1)$	$3[\frac{1}{7\sqrt{3}}Im(D_\perp^L D_0^{L*})+\frac{1}{5}Im(H_\perp^L H_0^{L*})+\frac{1}{\sqrt{15}}Im(D_\perp^L S^{L*})]$	+1
17	$P_3^1\sqrt{2}Im(Y_2^1)$	$\frac{6}{5\sqrt{14}}[2Im(D_\perp^L H_0^{L*})+\sqrt{3}Im(H_\perp^L D_0^{L*})]$	+1
18	$P_4^1\sqrt{2}Im(Y_2^1)$	$\frac{6}{7\sqrt{2}}Im(D_\perp^L D_0^{L*})$	+1
19	$P_0^0\sqrt{2}Re(Y_2^2)$	$-\frac{3}{2\sqrt{15}}[(\|H_\parallel^L\|^2-\|H_\perp^L\|^2)+(\|D_\parallel^L\|^2-\|D_\perp^L\|^2)]$	+1
20	$P_1^0\sqrt{2}Re(Y_2^2)$	$-\frac{3}{5}[Re(H_\parallel^L D_\parallel^{L*})-Re(D_\perp^L H_\perp^{L*})]$	+1
21	$P_2^0\sqrt{2}Re(Y_2^2)$	$\frac{\sqrt{3}}{2}[-\frac{1}{7}(\|D_\parallel^L\|^2-\|D_\perp^L\|^2)+\frac{1}{5}(\|H_\parallel^L\|^2-\|H_\perp^L\|^2)]$	+1
22	$P_3^0\sqrt{2}Re(Y_2^2)$	$\frac{3}{5}\sqrt{\frac{3}{7}}[Re(H_\parallel^L D_\parallel^{L*})-Re(D_\perp^L H_\perp^{L*})]$	+1
23	$P_4^0\sqrt{2}Re(Y_2^2)$	$\frac{2}{7}\sqrt{\frac{3}{5}}(\|D_\parallel^L\|^2-\|D_\perp^L\|^2)$	+1
24	$P_0^0\sqrt{2}Im(Y_2^2)$	$\sqrt{\frac{3}{5}}[Im(H_\perp^L H_\parallel^{L*})+Im(D_\perp^L D_\parallel^{L*})]$	+1
25	$P_1^0\sqrt{2}Im(Y_2^2)$	$\frac{3}{5}Im(H_\perp^L D_\parallel^{L*}+D_\perp^L H_\parallel^{L*})$	+1
26	$P_2^0\sqrt{2}Im(Y_2^2)$	$\sqrt{3}[\frac{1}{7}Im(D_\perp^L D_\parallel^{L*})-\frac{1}{5}Im(H_\perp^L H_\parallel^{L*})]$	+1
27	$P_3^0\sqrt{2}Im(Y_2^2)$	$-\frac{3}{5}\sqrt{\frac{3}{7}}Im(D_\perp^L H_\parallel^{L*}+H_\perp^L D_\parallel^{L*})$	+1
28	$P_4^0\sqrt{2}Im(Y_2^2)$	$-\frac{4}{7}\sqrt{\frac{3}{5}}Im(D_\perp^L D_\parallel^{L*})$	+1
29	$P_0^0 Y_1^0$	$-\sqrt{3}[Re(H_\perp^L H_\parallel^{L*})+Re(D_\perp^L D_\parallel^{L*})]$	-1
30	$P_1^0 Y_1^0$	$-\frac{3}{\sqrt{5}}Re(H_\perp^L D_\parallel^{L*}+H_\parallel^L D_\perp^{L*})$	-1
31	$P_2^0 Y_1^0$	$-\frac{3}{\sqrt{15}}[\frac{5}{7}Re(D_\perp^L D_\parallel^{L*})-Re(H_\perp^L H_\parallel^{L*})]$	-1
32	$P_3^0 Y_1^0$	$\frac{9}{\sqrt{105}}Re(H_\perp^L D_\parallel^{L*}+H_\parallel^L D_\perp^{L*})$	-1
33	$P_4^0 Y_1^0$	$\frac{4\sqrt{3}}{7}Re(D_\perp^L D_\parallel^{L*})$	-1
34	$P_1^1\sqrt{2}Re(Y_1^1)$	$\sqrt{\frac{3}{5}}[\sqrt{5}Re(H_\perp^L S^{L*})+\sqrt{3}Re(D_\perp^L H_0^{L*})-Re(H_\perp^L D_0^{L*})]$	-1
35	$P_2^1\sqrt{2}Re(Y_1^1)$	$3[\frac{1}{\sqrt{5}}Re(H_\perp^L H_0^{L*})+\frac{1}{\sqrt{3}}Re(D_\perp^L S^{L*})+\frac{5}{21}\sqrt{\frac{3}{5}}Re(D_\perp^L D_0^{L*})]$	-1
36	$P_3^1\sqrt{2}Re(Y_1^1)$	$\frac{6}{70}[2Re(D_\perp^L H_0^{L*})+\sqrt{3}Re(H_\perp^L D_0^{L*})]$	-1
37	$P_4^1\sqrt{2}Re(Y_1^1)$	$\frac{3\sqrt{10}}{7}Re(D_\perp^L D_0^{L*})$	-1
38	$P_1^1\sqrt{2}Im(Y_1^1)$	$-\sqrt{\frac{3}{5}}[\sqrt{5}Im(H_\parallel^L S^{L*})+\sqrt{3}Im(D_\parallel^L H_0^{L*})-Im(H_\parallel^L D_0^{L*})]$	-1
39	$P_2^1\sqrt{2}Im(Y_1^1)$	$-\sqrt{\frac{3}{5}}[\sqrt{3}Im(H_\parallel^L H_0^{L*})+\sqrt{5}Im(D_\parallel^L S^{L*})+\frac{5}{7}Im(D_\parallel^L D_0^{L*})]$	-1
40	$P_3^1\sqrt{2}Im(Y_1^1)$	$-6\sqrt{\frac{1}{70}}[2Im(D_\parallel^L H_0^{L*})+\sqrt{3}Im(H_\parallel^L D_0^{L*})]$	-1
41	$P_4^1\sqrt{2}Im(Y_1^1)$	$-\frac{3}{7}\sqrt{10}Im(D_\parallel^L D_0^{L*})$	-1

Table 8.1: The transversity-basis moments of 41 orthonormal angular functions $f_i(\Omega)$ in Eq. 8.1.

The moments in the presence of background are estimated as:

$$M_i = \sum_{k=1}^{n_{sig}} \omega_k f_i(\Omega_k) - x \sum_{k=1}^{n_{bkg}} \omega_k f_i(\Omega_k), \tag{8.3}$$

and the corresponding covariance matrix is given by:

$$C_{ij} = \sum_{k=1}^{n_{sig}} \omega_k^2 f_i(\Omega_k) f_j(\Omega_k) + x^2 \sum_{k=1}^{n_{bkg}} \omega_k^2 f_i(\Omega_k) f_j(\Omega_k), \tag{8.4}$$

where n_{sig} and n_{bkg} are the numbers of signal and background candidates, x is the ratio of the estimated number of background candidates in the signal region over the number of candidates in the background region, and the weights ω_k are the inverses of candidates efficiencies. This analysis comprises 40 normalized angular moments with 40x40 reduced covariance matrix for $i, j \in \{2, ..., 41\}$ as:

$$\bar{M}_i = \frac{M_i}{M_1}, \tag{8.5}$$

$$\bar{C}_{ij} = [C_{ij} + \frac{M_j M_j}{M_1^2} C_{11} - \frac{M_i C_{1j} + M_j C_{1i}}{M_1}] \frac{1}{M_1^2}. \tag{8.6}$$

The first moment M_1 corresponds to the total decay rate.

8.2 The consistency relations

The angular analysis formalism can be checked by taking advantage of the fact that 41 moments M_i are not independent of each other. A set of seven relations is used:

$$\begin{aligned}
&\mathcal{R}_1 \equiv M_{25} + \sqrt{\tfrac{7}{3}} M_{27} = 0, \\
&\mathcal{R}_2 \equiv M_{20} + \sqrt{\tfrac{7}{3}} M_{22} = 0, \\
&\mathcal{R}_3 \equiv \sqrt{\tfrac{3}{7}} M_{30} + M_{32} = 0, \\
&\mathcal{R}_4 \equiv M_{21} + \tfrac{1}{\sqrt{5}}(3M_{23} + M_{19}) = 0 \\
&\mathcal{R}_5 \equiv \tfrac{M_3}{\sqrt{5}} + M_8 + \tfrac{3}{5}[(\sqrt{5} M_{10} + M_5) + \tfrac{1}{3}(M_1 + \sqrt{5} M_6)] = 0, \\
&\mathcal{R}_6 \equiv \tfrac{7}{4}\sqrt{\tfrac{5}{3}} M_{28} + \tfrac{7}{12}\sqrt{\tfrac{5}{3}}(M_{24} + \sqrt{5} M_{26}) = 0, \\
&\mathcal{R}_7 \equiv \tfrac{7}{4\sqrt{3}} M_{33} + \tfrac{7}{12\sqrt{3}}(M_{29} + \sqrt{5} M_{31}) = 0.
\end{aligned} \tag{8.7}$$

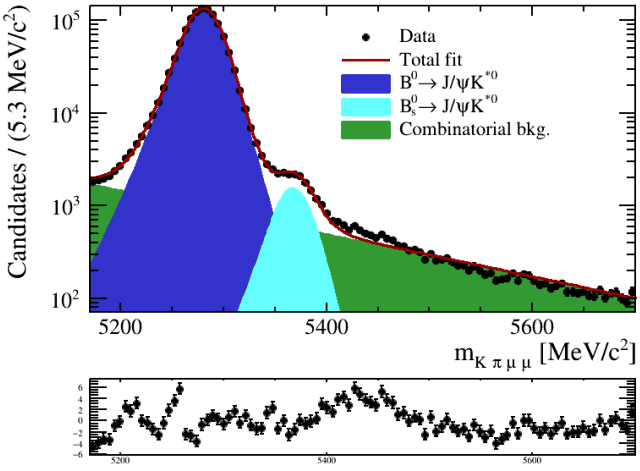
As it is depicted in Eq. 8.7, the observables $\mathcal{R}_i$ should be equal to zero and any significant deviation from this value would imply a departure from the SPD-wave model assumptions.

Chapter 9

Branching ratio of B decay

9.1 Signal yield

To discriminate between signal and background candidates, a fit to the K$\pi\mu\mu$ invariant mass distribution is performed in the region $5170 < m_{K\pi\mu\mu} < 5700$ MeV/c^2. The signal is modeled by a double Crystal Ball function with a common mean and tail parameters (Eq. 5.10). Generally, a Crystal Ball function consists of a Gaussian core and a power-law tail. The combinatorial background is described by an exponential function.



Parameter	Run2
n_{B^0}	1137900 ± 1160
$n_{B_s^0}$	13478 ± 198
n_{bkg}	58154 ± 516
μ [MeV/c^2]	5279.98 ± 0.02
σ_1 [MeV/c^2]	15.50 ± 0.07
σ_2 [MeV/c^2]	26.61 ± 0.34
α	1.36 ± 0.01
n	140 ± 27
f_1	0.75 ± 0.01
λ [c^2/MeV]	-0.00540±0.00005

Figure 9.1: Fit to the K$\pi\mu\mu$ invariant mass of the control mode $B^0 \to J\psi K^{*0}$ (blue area) with an additional component of the $B_s^0 \to J\psi K^{*0}$ (cyan area) Combinatorial background is indicated by the green area. The table on the right contains fit parameters.

9.2 High statistic control mode fit

The first thing is to perform a fit to the high statistic control mode $B^0 \to J\psi K^{*0}$ to reduce the number of degrees of freedom in the invariant mass fit to the $B^0 \to K^+\pi^-\mu^+\mu^-$ mass distribution (Fig. 9.1). As in the Sec. 5.7, an additional component of the $B_s^0 \to J\psi K^{*0}$ is added for better fitting, as it has the same final state as the control mode. The fit parameters, given in Figure 9.1, are fixed when performing a fit to the $B^0 \to K^+\pi^-\mu^+\mu^-$. The cut value on the BDT is chosen to be 0.05, based on the optimization described in Sec. 5.9.

9.3 Scaling factor

It is important to take into account possible variations in the signal mass lineshape that could be due to the difference in q^2 and $m_{K\pi}$ in the fit to the $B^0 \to J\psi K^{*0}$ ($q^2 \in [9.22, 9.96]$ GeV^2/c^4, $m_{K\pi} \in [796, 996]$ MeV/c^2) and the $B^0 \to K^+\pi^-\mu^+\mu^-$ ($q^2 \in [0.1, 8.0]$ GeV^2/c^4, $m_{K\pi} \in [1330, 1530]$ MeV/c^2). To do that a single scaling factor s_σ is applied to both σ_1 and σ_2 in the Crystal Ball function when fitting the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates. This scaling factor is determined by first fitting the $K\pi\mu\mu$ mass distribution using the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC in the $B^0 \to J\psi K^{*0}$ region. Then, all fit parameters are fixed, except for the s_σ which is allowed to float in the next few fits to the $m_{K\pi\mu\mu}$ distribution in all studied q^2 bins in the $B^0 \to K^+\pi^-\mu^+\mu^-$ signal region. All fits to the $K\pi\mu\mu$ mass distribution for obtaining the s_σ are presented in Appendix D.

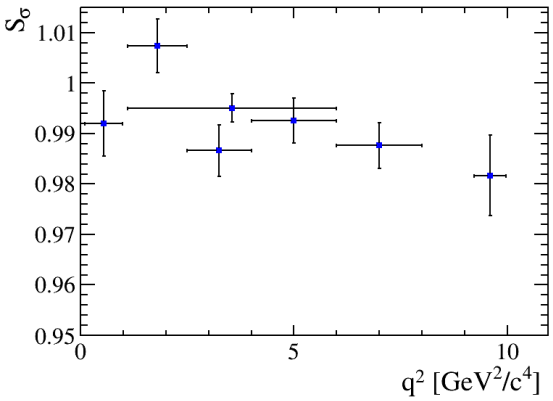


Figure 9.2: Scaling factor s_σ for 7 q^2 values for candidates in $1330 < m_{K\pi} < 1530$ MeV/c^2.

Figure 9.2 shows the distribution of the s_σ in several q^2 values within the range $m_{K\pi} \in [1330, 1530]$ MeV/c^2 and Table 9.1 shows the values of the s_σ.

q^2 [GeV^2/c^4]	s_σ
$[0.01, 0.98]$	0.992 ± 0.007
$[1.10, 2.50]$	1.007 ± 0.005
$[2.50, 4.00]$	0.987 ± 0.005
$[4.00, 6.00]$	0.993 ± 0.004
$[6.00, 8.00]$	0.988 ± 0.005
$[1.10, 6.00]$	0.995 ± 0.003
$[9.22, 9.96]$	0.982 ± 0.008

Table 9.1: Values of the scaling factors s_σ for corresponding q^2 bin for candidates in $1330 < m_{K\pi} < 1530$ MeV/c^2 .

9.4 Candidate yields

The invariant mass fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in six different q^2 bins are shown in Figure. 9.3. The determined signal and background yields are given in Table 9.2 in the regions $5170 < m_{K\pi\mu\mu} < 5700$ MeV/c^2 and $1330 < m_{K\pi} < 1530$ MeV/c^2.

q^2 [GeV^2/c^4]	Signal	Background
$[0.01, 0.98]$	268 ± 21	463 ± 25
$[1.10, 2.50]$	222 ± 23	897 ± 35
$[2.50, 4.00]$	176 ± 23	1151 ± 39
$[4.00, 6.00]$	343 ± 29	1504 ± 45
$[6.00, 8.00]$	284 ± 29	1954 ± 50
$[1.10, 6.00]$	741 ± 43	3552 ± 68

Table 9.2: Yields of signal and background of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in six q^2 bins in the regions $5170 < m_{K\pi\mu\mu} < 5700$ MeV/c^2 and $1330 < m_{K\pi} < 1530$ MeV/c^2.

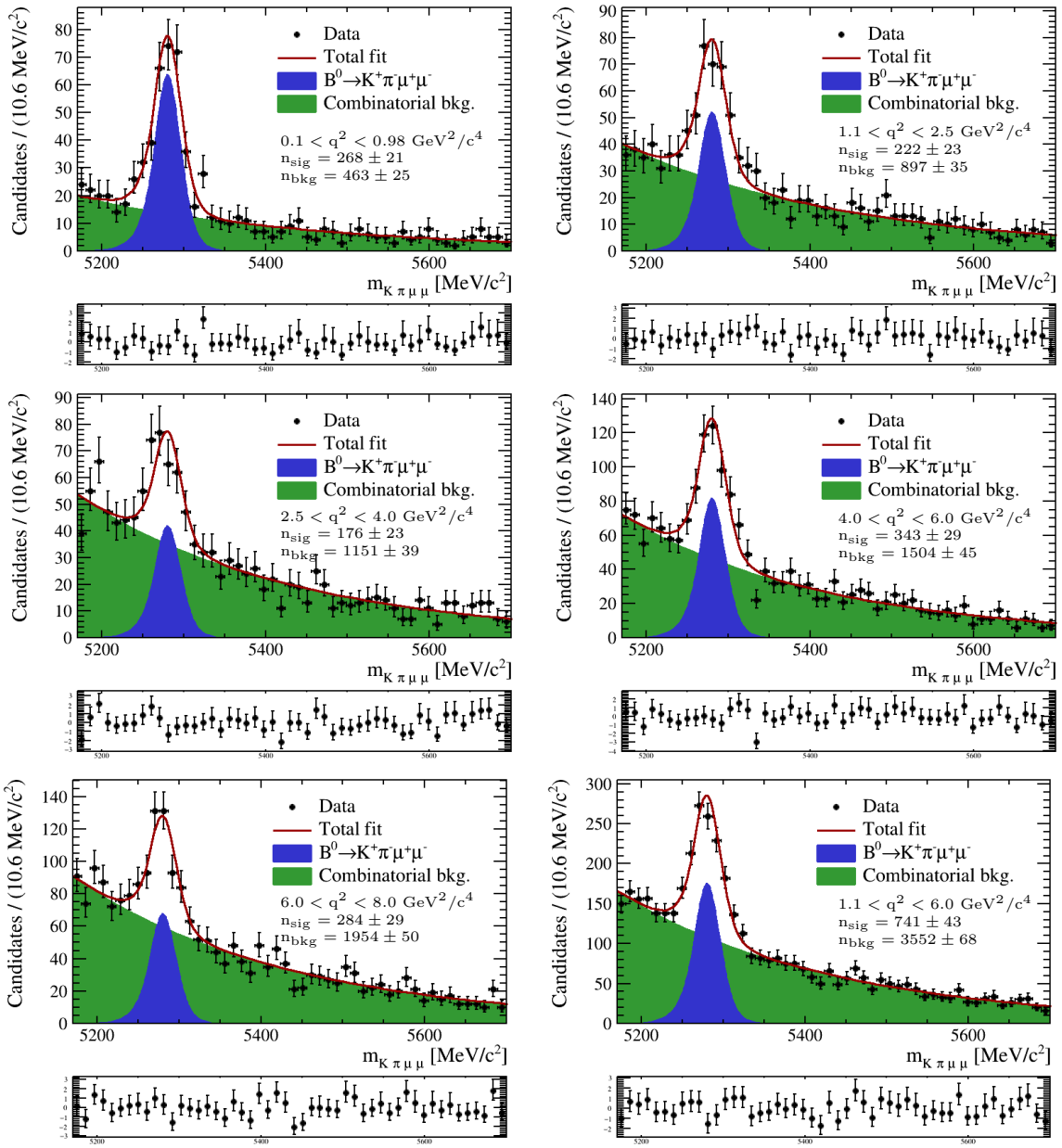


Figure 9.3: Maximum likelihood fits to the K$\pi\mu\mu$ invariant mass of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in six q^2 bins. The signal and combinatorial background are indicated by the blue and green area, respectively.

9.5 Differential branching fraction

The branching fraction $\mathcal{B}(\mathrm{B}^0 \to \mathrm{K}^+\pi^-\mu^+\mu^-)$ is a fraction of particles that decay through a given decay mode. The differential branching fraction $\mathrm{d}\mathcal{B}/\mathrm{dq}^2$ of the $\mathrm{B}^0 \to \mathrm{K}^+\pi^-\mu^+\mu^-$ decay is measured in six q^2 bins for the $\mathrm{m}_{\mathrm{K}\pi}$ region $1330 < \mathrm{m}_{\mathrm{K}\pi} < 1530$ MeV/c^2 and the decay $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$ is used as a normalization channel selected in the regions $9.22 < \mathrm{q}^2 < 9.96$ GeV^2/c^4 and $796 < \mathrm{m}_{\mathrm{K}\pi} < 996$ MeV/c^2:

$$\begin{aligned}\frac{d\mathcal{B}}{dq^2} = \frac{1}{(q^2_{max}-q^2_{min})} f_{K^*(892)^0}\mathcal{B}(B^0 \to J/\psi K^{*0})\mathcal{B}(J/\psi \to \mu^+\mu^-) \\ \times\mathcal{B}(K^*(892)^0 \to K^+\pi^-)\frac{N'_{K^+\pi^-\mu^+\mu^-}}{(1-F_S^{J/\psi K^{*0}})N'_{J/\psi K^{*0}}},\end{aligned} \tag{9.1}$$

where $N'_{K^+\pi^-\mu^+\mu^-}$ and $N'_{J/\psi K^{*0}}$ are the acceptance-corrected yields of the $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ decays, respectively. $\mathrm{F}_{\mathrm{S}}^{\mathrm{J}/\psi\mathrm{K}^{*0}}$ is an S-wave fraction of the $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$ decay. The fraction $f_{K^*(892)^0}$ is used for scaling the value of $\mathcal{B}(\mathrm{K}^*(892)^0 \to \mathrm{K}^+\pi^-)$ to the range $796 < m_{K\pi} < 996$ MeV/c^2. The values of the above fractions along with the branching fractions used in Eq. 9.1 are listed below:

$$\begin{aligned}&f_{K^*(892)^0} = 0.894 \pm 0.001,\ [22] \\ &F_S^{J/\psi K^{*0}} = 0.084 \pm 0.01,\ [22] \\ &\mathcal{B}(B^0 \to J/\psi K^{*0}) = (1.19 \pm 0.01 \pm 0.08) \times 10^{-3}\ [44],\ l \\ &\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.961 \pm 0.033) \times 10^{-2}[42], \\ &\mathcal{B}(K^*(892)^0 \to K^+\pi^-) = 2/3.\end{aligned} \tag{9.2}$$

9.6 Acceptance-corrected yield

The measured yield is corrected using the average acceptance weight, where each weight is the inverse of the candidate-by-candidate efficiency.

The average weight is given by:

$$\bar{\omega} = \frac{1}{N}\sum_i^N \omega_i, \tag{9.3}$$

where ω_{i} is the candidate-by-candidate weight and N is the number of candidates. An estimate for an error of the average weight is given by:

$$\sigma_{\bar{\omega}} = \sqrt{\frac{1}{N(N-1)}\sum_i^N(\omega_i - \bar{\omega})^2}. \tag{9.4}$$

The average weight calculated in the signal mass window is an admixture of the signal and background candidates:

$$\bar{\omega}_{mix} = \frac{N_{sig}\bar{\omega}_{sig} + N_{bkg}\bar{\omega}_{bkg}}{N_{sig} + N_{bkg}}, \tag{9.5}$$

where N_{sig} and N_{bkg} are the number of signal and background candidates in the signal mass window, respectively. In order to obtain signal corrected yield, the above formula needs to be rearranged:

$$N_{sig}\bar{\omega}_{sig} = (N_{sig} + N_{bkg})\bar{\omega}_{mix} - N_{bkg}\bar{\omega}_{bkg}. \tag{9.6}$$

The signal and background regions are defined as follows:

- signal window - $5230 < m_{K\pi\mu\mu} < 5330$ MeV/c^2,
- background window - $5350 < m_{K\pi\mu\mu} < 5700$ MeV/c^2,
- background window - $5450 < m_{K\pi\mu\mu} < 5700$ MeV/c^2 (resonant mode only).

A different background region for resonant decay mode is used in order to prevent any potential contribution from the $B^0_{(s)} \to J/\psi K^{*0}$ candidates. The values of the signal yield N_{sig} and the average signal weight $\bar{\omega}_{sig}$ along with the acceptance corrected yield for both $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ candidates are shown in Table 9.3. The corresponding yields are slightly different from those obtained in the previous part of the analysis since this time the signal and background yields are projected onto the signal mass region.

q^2[GeV^2/c^4]	N_{sig}	N_{bkg}	$\bar{\omega}_{sig}$	$N_{sig}\bar{\omega}_{sig}$
[0.10,0.98]	261±20	130±7	1.21	315±31
[1.10,2.50]	216±22	257±10	1.22	264±34
[2.50,4.00]	171±22	335±11	1.09	186±28
[4.00,6.00]	334±28	442±13	1.08	359±31
[6.00,8.00]	277±29	567±15	1.00	276±29
[1.10,6.00]	720±42	1033±20	1.12	807±54
[9.22,9.96]	1103900±1120	18620±165	0.96	1059900±1140

Table 9.3: The signal yield N_{sig}, the background yield N_{bkg}, the average signal weight $\bar{\omega}_{sig}$ and acceptance corrected yield $N_{sig}\bar{\omega}_{sig}$ for the $B^0 \to K^+\pi^-\mu^+\mu^-$ in six different q^2 bins and the $B^0 \to J/\psi K^{*0}$ candidates.

Chapter 10

Systematic uncertainties

For the differential branching fraction and angular moments measurements it is necessary to calculate the following systematic uncertainties:

- uncertainties resulting from contributions from peaking background candidates,
- statistical uncertainty on the acceptance correction,
- uncertainties resulting from the chosen order of the acceptance correction,
- uncertainties resulting from the residual mismodeling of MC simulation,
- uncertainty resulting from the resolutions in the angular variables,
- uncertainty on a chosen signal mass model.

For both the differential branching fraction and angular moment measurements, the systematic uncertainty is evaluated as a difference between the nominal sample and the sample where one or more parameters are systematically varied.

10.1 Uncertainties due to contributions from peaking background candidates

Section 5.6 is dedicated to a description of a selection applied to isolate the peaking background candidates. The effect of those peaking backgrounds is determined using pseudoexperiments, where specific peaking background components are generated along with the signal and the combinatorial background. The selections described in Section 5.6 are used to isolate the corresponding peaking background.

The observed difference in both results with and without additional peaking background contributions is considered as a systematic uncertainty.

10.2 Statistical uncertainty on the acceptance correction

As it was described in Section 7.2, the five-dimensional acceptance function is derived using the Legendre coefficients determined from a simulated sample of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates. The covariance matrix of the coefficients is also determined and is used to perform the so-called toy studies. In the toy studies, a number of new sets of coefficients are generated and each set is transformed using the Cholesky decomposition. In this method the original covariance matrix is decomposed to the product of the lower triangular matrix and its conjugate transpose. The product of the generated set of coefficients with the lower triangular matrix gives in result a new set of coefficients. These new coefficients are correlated in the same way as the original set of coefficients. The angular moments and differential branching fractions are then measured using both the original acceptance and the new one. The difference between the two measurements is distributed around 0 and the standard deviations of those distributions are considered as a systematic uncertainty. The results of the toy studies are presented in Appendices E.1, E.2.

10.3 Higher order acceptance correction

To account for uncertainties that are due to the number of the order in polynomials in Eq. 7.2, a new acceptance correction is calculated with an increased order of the polynomials by two. The angular moments and differential branching fraction are then measured using both the nominal acceptance and the new one with the higher order. The difference between the two measurements is considered as a systematic uncertainty.

10.4 Residual mismodeling of MC simulation

The candidates' weights calculated in Section 6.2 also have an influence on the acceptance correction. As it was described in Section 6.2, three variables are used to obtain a good agreement between the data and simulation: the number of tracks in the decay (nTracks), the transverse momentum p_T of the B^0 candidate and the quality of the $K\pi\mu\mu$ vertex (χ^2/ndf). To evaluate the effect of this correction on the acceptance, a new acceptance is calculated, where the reweighting is recalculated using yet another variable: the impact parameter of the B^0 meson candidate with respect to its best own parent vertex. Figure 10.1 shows the results of the candidates' weights determination. This calculated weights are incorporated into the new acceptance correction. The angular moments and the differential branching fractions are then measured using both the nominal acceptance and the new one. The difference between the two measurements is considered as a systematic uncertainty.

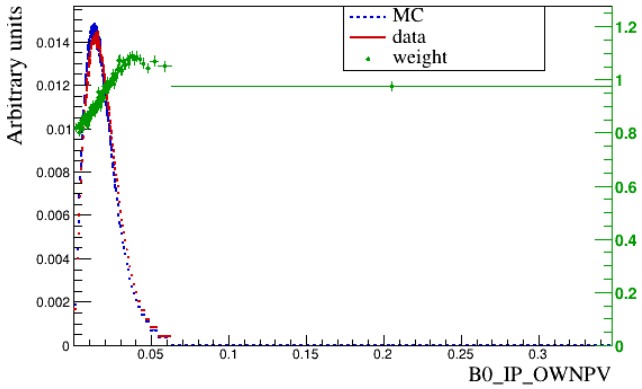


Figure 10.1: Distribution of the impact parameter of the B^0 meson candidate with respect to its best own parent vertex used to derive weights. The red histogram corresponds to the *sWeighted* $B^0 \to J/\psi K^{*0}$ data, the blue histogram shows the $B^0 \to J/\psi K^{*0}$ MC and the green points represent the calculated weights.

10.5 Uncertainty due to resolutions in the angular variables

The resolution of a particular angular variable is calculated using a simulated sample as a difference between the generated and reconstructed angle. Figure 10.2 shows the resolution histograms of three decay angles and the standard deviation of each histogram is taken as a resolution (listed in Table 10.1).

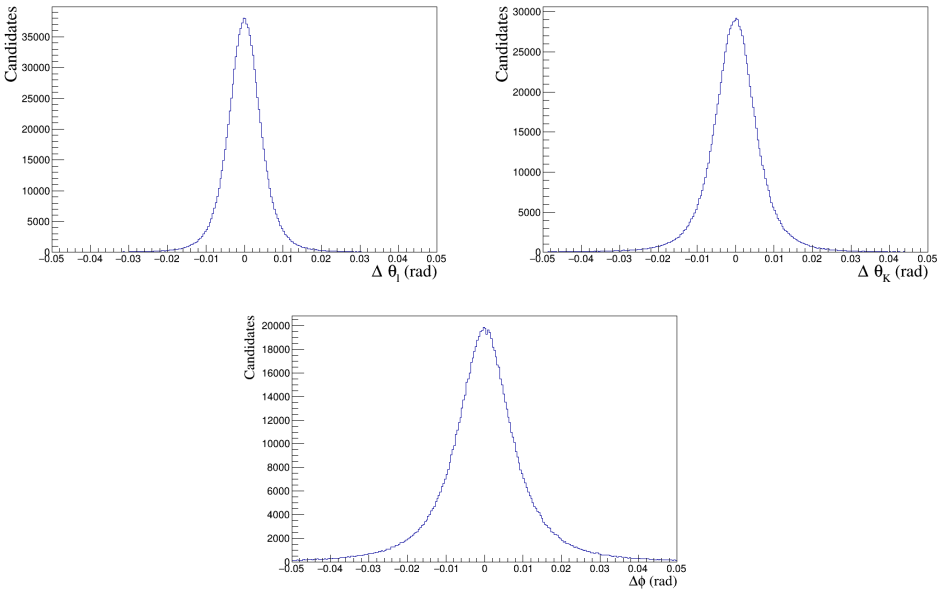


Figure 10.2: Global angular resolutions in three decay angles: θ_ℓ, θ_K and ϕ.

angle	resolution (rad.)
θ_ℓ	0.0058
$\theta_{\rm K}$	0.0078
ϕ	0.0119

Table 10.1: Global angular resolutions in three decay angles: θ_ℓ, $\theta_{\rm K}$ and ϕ. Each is taken as a standard deviation of histograms from Figure 10.5.

The determined resolutions are used to perturb each angular variable in order to obtain the deviation in Eq. 3.11:

$$\int f(\theta_\ell, \theta_K, \phi) f((\theta_\ell + \sigma(\theta_\ell), \theta_K + \sigma(\theta_K), \phi + \sigma(\phi))) d\Omega = \delta_{ij} + \Delta_{ij} \tag{10.1}$$

where $\mathrm{d}\Omega = \mathrm{dcos}\theta_{\rm K}\mathrm{dcos}\theta_\ell\mathrm{d}\phi$, $f(\theta_\ell, \theta_K, \phi)$ is a true angular distribution in data and $f((\theta_\ell + \sigma(\theta_\ell), \theta_K + \sigma(\theta_K), \phi + \sigma(\phi)))$ is a function constructed from smeared angles used to calculate the angular moments. $\Delta_{\rm ij}$ acts as a deviation from the orthonormality conditions. The angular moments are in result shifted because of a deviation from the orthonormality conditions, and the difference is interpreted as a systematic uncertainty.

10.6 Uncertainty on a signal mass model

In order to account for uncertainties that are due to the chosen signal mass model, a double Gaussian is used as an alternative model. The comparison between those two models is shown in Appendix E.3. The angular moments and the differential branching fractions are measured using both the original Crystal Ball function and the double Gaussian. The difference between the two measurements is considered as a systematic uncertainty.

10.7 Systematic uncertainties for the branching fraction measurement

The summary in Table 10.2 shows the results of the systematic uncertainties in the branching fraction measurements. The largest contribution comes from the statistical uncertainty on the acceptance correction due to the limited size of the simulated $\mathrm{B}^0 \to \mathrm{K}^+\pi^-\mu^+\mu^-$ sample. The last row in the table shows total systematic uncertainties calculated as a sum in quadrature.

q^2 [GeV^2/c^4]	[0.1,0.98]	[1.1,2.5]	[2.5,4.0]	[4.0,6.0]	[6.0,8.0]	[1.1,6.0]
Acceptance correction	0.162	0.092	0.064	0.068	0.061	0.060
Alternative mass model	0.003	0.000	0.002	0.001	0.001	0.000
Data-MC correction	0.001	0.002	0.000	0.001	0.000	0.001
Higher order	0.039	0.049	0.003	0.015	0.007	0.022
Peaking background	0.001	0.018	0.079	0.035	0.014	0.017
Sum in quadrature	0.167	0.106	0.102	0.078	0.063	0.067

Table 10.2: Systematic uncertainties for the differential branching fraction measurements in bins of q^2. All results are in units of $10^{-8}[c^4/GeV^2]$.

10.8 Systematic uncertainties for angular analysis

The summary in Table 10.3 shows the results of the systematic uncertainties in the angular moment measurements. The largest contribution comes from the statistical uncertainty on the acceptance correction due to the limited size of the simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ sample.
The last column in the table shows the total systematic uncertainties calculated as a sum in quadrature.

	Acceptance correction	Alternative mass model	Data-MC correction	Higher order	Peaking background	Angular resolution	Sum in quadrature
$\overline{M}_2$	0.050	0.000	0.001	0.003	0.003	0.011	0.051
$\overline{M}_3$	0.056	0.002	0.001	0.025	0.001	0.004	0.062
$\overline{M}_4$	0.055	0.000	0.001	0.002	0.004	0.005	0.055
$\overline{M}_5$	0.050	0.002	0.001	0.003	0.005	0.002	0.051
$\overline{M}_6$	0.053	0.004	0.000	0.018	0.007	0.001	0.057
$\overline{M}_7$	0.053	0.001	0.000	0.009	0.008	0.001	0.054
$\overline{M}_8$	0.053	0.000	0.001	0.035	0.001	0.001	0.064
$\overline{M}_9$	0.052	0.000	0.000	0.022	0.006	0.001	0.057
$\overline{M}_{10}$	0.049	0.000	0.000	0.013	0.008	0.003	0.052
$\overline{M}_{11}$	0.039	0.000	0.000	0.005	0.002	0.000	0.039
$\overline{M}_{12}$	0.043	0.001	0.000	0.011	0.001	0.002	0.044
$\overline{M}_{13}$	0.046	0.000	0.001	0.002	0.001	0.003	0.046
$\overline{M}_{14}$	0.046	0.001	0.001	0.009	0.003	0.003	0.048
$\overline{M}_{15}$	0.031	0.000	0.000	0.006	0.002	0.000	0.032
$\overline{M}_{16}$	0.035	0.000	0.000	0.000	0.005	0.002	0.035
$\overline{M}_{17}$	0.037	0.000	0.000	0.019	0.005	0.000	0.041
$\overline{M}_{18}$	0.038	0.000	0.000	0.004	0.002	0.003	0.039
$\overline{M}_{19}$	0.040	0.001	0.001	0.012	0.003	0.001	0.042
$\overline{M}_{20}$	0.042	0.000	0.000	0.010	0.004	0.002	0.043
$\overline{M}_{21}$	0.046	0.000	0.000	0.009	0.001	0.003	0.047
$\overline{M}_{22}$	0.044	0.000	0.000	0.007	0.001	0.003	0.045
$\overline{M}_{23}$	0.043	0.000	0.001	0.008	0.002	0.000	0.044
$\overline{M}_{24}$	0.025	0.000	0.000	0.002	0.001	0.002	0.025
$\overline{M}_{25}$	0.034	0.000	0.000	0.001	0.002	0.001	0.034
$\overline{M}_{26}$	0.038	0.000	0.000	0.001	0.002	0.000	0.038
$\overline{M}_{27}$	0.039	0.000	0.001	0.000	0.001	0.002	0.039
$\overline{M}_{28}$	0.038	0.001	0.000	0.001	0.001	0.000	0.038
$\overline{M}_{29}$	0.044	0.002	0.000	0.028	0.006	0.008	0.053
$\overline{M}_{30}$	0.026	0.001	0.000	0.008	0.003	0.000	0.027
$\overline{M}_{31}$	0.051	0.001	0.001	0.026	0.000	0.004	0.057
$\overline{M}_{32}$	0.050	0.001	0.001	0.028	0.006	0.002	0.057
$\overline{M}_{33}$	0.051	0.001	0.001	0.015	0.008	0.003	0.054
$\overline{M}_{34}$	0.047	0.000	0.000	0.002	0.000	0.000	0.048
$\overline{M}_{35}$	0.042	0.000	0.000	0.001	0.000	0.003	0.042
$\overline{M}_{36}$	0.046	0.001	0.000	0.001	0.002	0.002	0.046
$\overline{M}_{37}$	0.050	0.001	0.001	0.016	0.002	0.002	0.052
$\overline{M}_{38}$	0.024	0.000	0.000	0.009	0.002	0.001	0.025
$\overline{M}_{39}$	0.029	0.001	0.000	0.004	0.003	0.000	0.029
$\overline{M}_{40}$	0.038	0.000	0.000	0.006	0.005	0.004	0.039
$\overline{M}_{41}$	0.033	0.000	0.000	0.011	0.002	0.001	0.035

Table 10.3: Systematic uncertainties for 40 normalized angular moments in a single q^2 range [1.1,6.0] GeV^2/c^4.

Chapter 11

Results

The results of both the angular analysis and the differential branching fractions are presented in the subsequent sections. The measurements were performed for the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates with the invariant mass range $1330 < m(K^+\pi^-) < 1530$ MeV/c^2, where the angular analysis was focused on a single bin of dimuon invariant mass squared: $1.1 < q^2 < 6.0$ GeV^2/c^4, whereas differential branching fraction was calculated in six q^2, as it is depicted in Table 3.3. Chapter 12 covers the interpretation of the results.

11.1 Differential branching fraction

The results of the differential branching fraction are shown in Figure11.1 and Table 11.1. The first uncertainty is statistical, the second is systematic and the third is due to the uncertainty on the branching fraction of the normalization channel $B^0 \to J/\psi K^*(892)^0$.

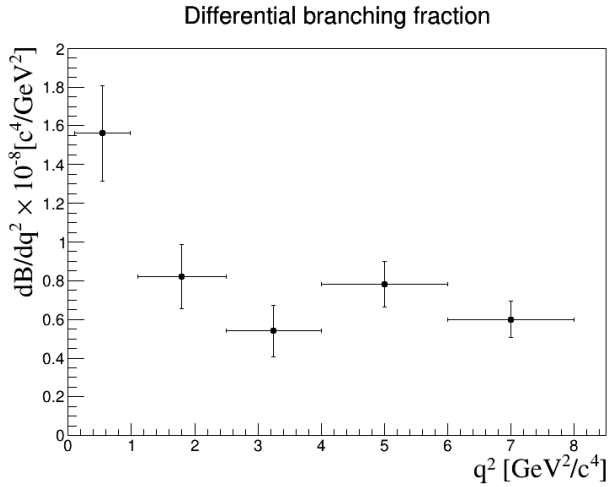


Figure 11.1: The differential branching fraction of the $B^0 \to K^+\pi^-\mu^+\mu^-$ in bins of q^2 in $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 range. The error bars correspond to the sums in quadrature of the statistical and systematic uncertainties.

$q^2[GeV^2/c^4]$	$d\mathcal{B}/dq^2 10^{-8}$ $[c^4/GeV^2]$
$[0.01, 0.98]$	$1.56 \pm 0.15 \pm 0.17 \pm 0.10$
$[1.10, 2.50]$	$0.82 \pm 0.11 \pm 0.11 \pm 0.06$
$[2.50, 4.00]$	$0.54 \pm 0.08 \pm 0.10 \pm 0.04$
$[4.00, 6.00]$	$0.78 \pm 0.07 \pm 0.08 \pm 0.05$
$[6.00, 8.00]$	$0.60 \pm 0.06 \pm 0.06 \pm 0.04$
$[1.10, 6.00]$	$0.72 \pm 0.05 \pm 0.07 \pm 0.05$

Table 11.1: The differential branching fraction of the $B^0 \to K^+\pi^-\mu^+\mu^-$ in bins of q^2 in $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 range. The first uncertainty is statistical, the second is systematic and the third is due to the uncertainty on the branching fraction of $B^0 \to J/\psi K^*(892)^0$ channel.

11.2 Angular analysis

The results of the angular analysis where 40 normalized moments are calculated are shown in Figure.11.2 and Table 11.2. The first uncertainty is statistical and the second is systematic. The statistical uncertainty is calculated based on the reduced covariance matrix from Equation.8.6 (Fig. 11.3). Figure 11.4 shows the results of calculating correlation matrix, where no strong correlation between the angular moments is observed.

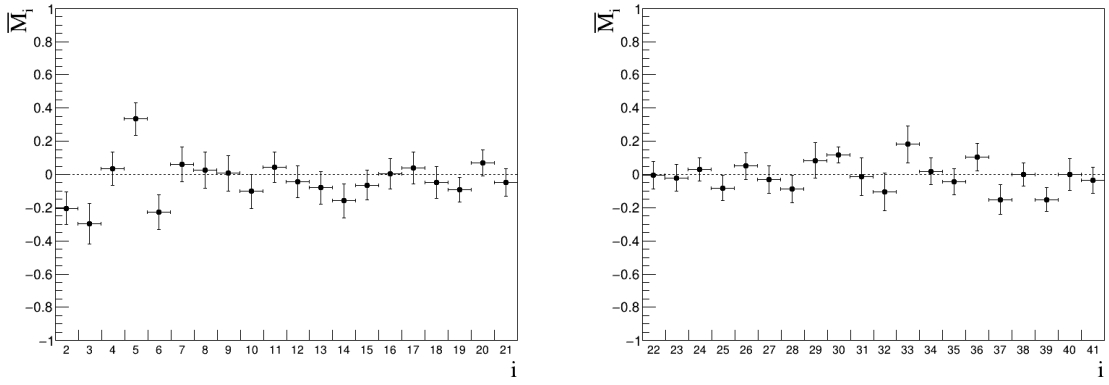


Figure 11.2: Measurement of 40 normalized moments, $\overline{M}_i$, of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay in the $1.1 < q^2 < 6.0$ GeV^2/c^4 and $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 ranges. The error bars correspond to the sums in quadrature of the statistical and systematic uncertainties.

Obs.	Value	Obs.	Value
$\overline{M}_2$	-0.20±0.08±0.05	$\overline{M}_{22}$	-0.01±0.07±0.05
$\overline{M}_3$	-0.30±0.11±0.06	$\overline{M}_{23}$	-0.02±0.07±0.04
$\overline{M}_4$	0.04±0.08±0.06	$\overline{M}_{24}$	0.03±0.06±0.03
$\overline{M}_5$	0.33±0.08±0.05	$\overline{M}_{25}$	-0.08±0.07±0.03
$\overline{M}_6$	-0.23±0.09±0.06	$\overline{M}_{26}$	0.05±0.07±0.04
$\overline{M}_7$	0.06±0.09±0.05	$\overline{M}_{27}$	-0.03±0.07±0.04
$\overline{M}_8$	0.03±0.09±0.06	$\overline{M}_{28}$	-0.09±0.07±0.04
$\overline{M}_9$	0.01±0.09±0.06	$\overline{M}_{29}$	0.08±0.09±0.05
$\overline{M}_{10}$	-0.10±0.09±0.05	$\overline{M}_{30}$	0.12±0.04±0.03
$\overline{M}_{11}$	0.05±0.08±0.04	$\overline{M}_{31}$	-0.01±0.10±0.06
$\overline{M}_{12}$	-0.05±0.08±0.04	$\overline{M}_{32}$	-0.10±0.10±0.06
$\overline{M}_{13}$	-0.08±0.09±0.05	$\overline{M}_{33}$	0.18±0.10±0.05
$\overline{M}_{14}$	-0.16±0.09±0.05	$\overline{M}_{34}$	0.02±0.07±0.05
$\overline{M}_{15}$	-0.06±0.08±0.03	$\overline{M}_{35}$	-0.04±0.07±0.04
$\overline{M}_{16}$	0.00±0.08±0.04	$\overline{M}_{36}$	0.11±0.07±0.05
$\overline{M}_{17}$	0.04±0.09±0.04	$\overline{M}_{37}$	-0.15±0.07±0.05
$\overline{M}_{18}$	-0.05±0.09±0.04	$\overline{M}_{38}$	0.00±0.07±0.03
$\overline{M}_{19}$	-0.09±0.06±0.04	$\overline{M}_{39}$	-0.15±0.07±0.03
$\overline{M}_{20}$	0.07±0.07±0.04	$\overline{M}_{40}$	0.00±0.09±0.04
$\overline{M}_{21}$	-0.05±0.07±0.05	$\overline{M}_{41}$	-0.04±0.07±0.04

Table 11.2: Measurement of 40 normalized moments, $\overline{M}_i$, of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay in the $1.1 < q^2 < 6.0$ GeV^2/c^4 and $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 ranges. The first uncertainty is statistical and the second is systematic.

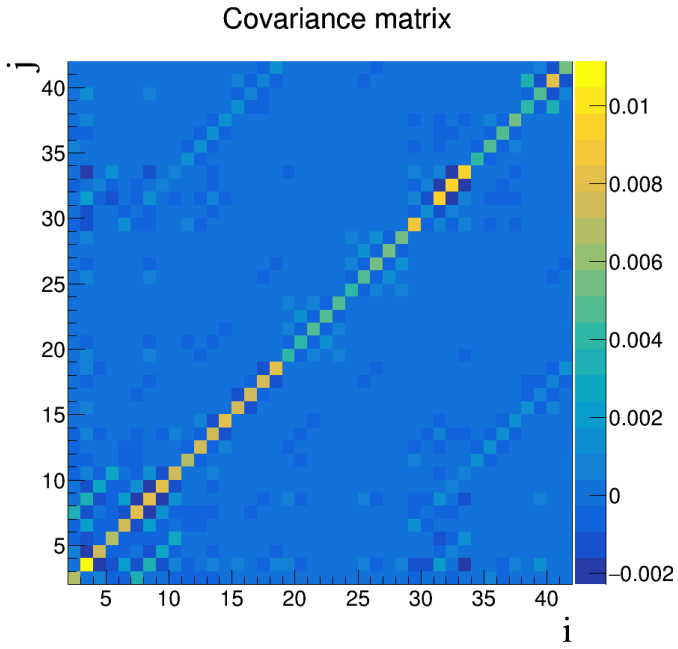


Figure 11.3: The reduced covariance matrix C_{ij} (i,j$\in$ $\{2,..,41\}$) used to calculate the statistical uncertainty of the angular moments, based on Equation 8.6.

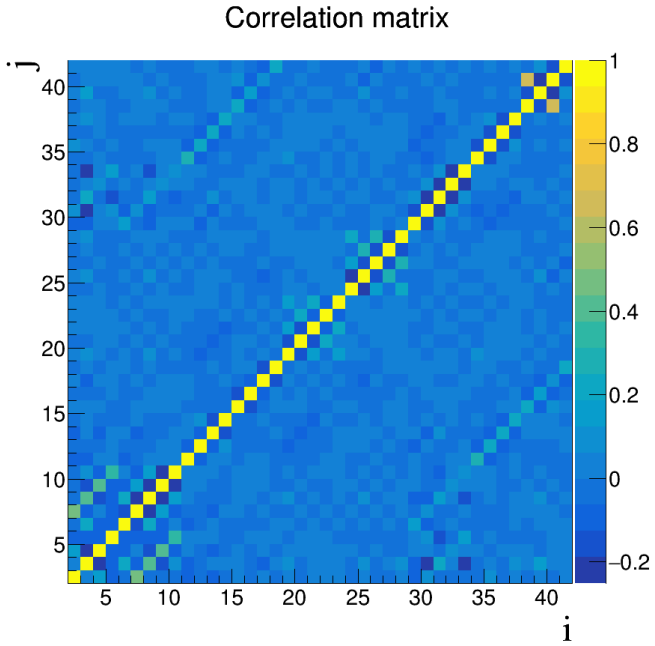


Figure 11.4: The reduced correalation matrix $Corr_{ij}$ (i,j$\in$ $\{2,..,41\}$) derived from the covariance matrix C_{ij}.

Chapter 12

Run1 and Run2 comparison

12.1 Differential branching fraction

Based on the Run1 data set, the differential branching fraction distribution shows a regular shape. A slightly different result is obtained with the Run2 data set, where there is an excess in two q^2 regions, [4,6] and [6,8] GeV^2/c^4, when compared to the lower q^2 ranges. However the compatibility with the Run1 results is maintained. The smaller error bars in these new results are due to the lower statistical uncertainty.

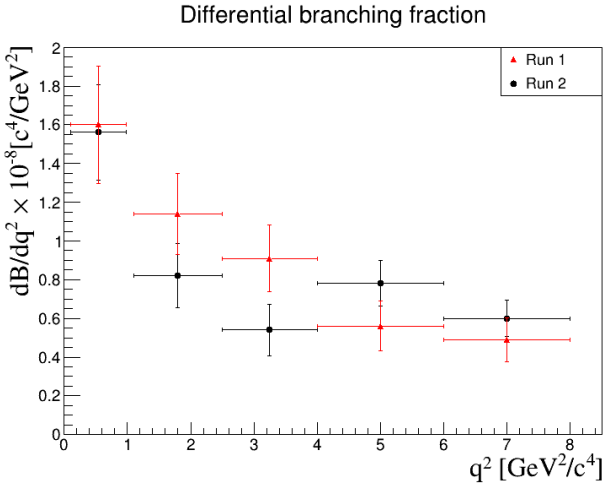


Figure 12.1: Comparison between Run1(red triangles) and Run2(black circles) differential branching fraction measurements of $B^0 \to K^+\pi^-\mu^+\mu^-$ in bins of q^2 in $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 range. The error bars correspond to the sums in quadrature of the statistical and systematic uncertainties.

12.2 Angular analysis

A comparison between the results from the analyses based on the Run1 data set [22] and new Run2 data set shows the consistency between them. It is clearly visible that the Run2 data analysis is more precise because of a higher number of candidates.

Upon analysis of the results, a few features can be observed directly, even without any information about the physical model. Still, the values of normalized moments $\overline{M}_2$ and $\overline{M}_3$ may indicate the presence of interference effects between the contributions from the S-, P- and D-waves. Some of the moments are proportional to $\sin\phi$. These moments correspond to the complex right-handed admixture in the weak hadronic current, absent in the SM. In line with Ref. [16], any statistical significance found in the moments corresponding to $\mathrm{Im}(Y_l^{m\neq 0})$ would lead to tensions with the SM. After checking those moments, namely: $\overline{M}_{15} - \overline{M}_{18}, \overline{M}_{24} - \overline{M}_{28}$ and $\overline{M}_{38} - \overline{M}_{41}$, no clear contribution was found.

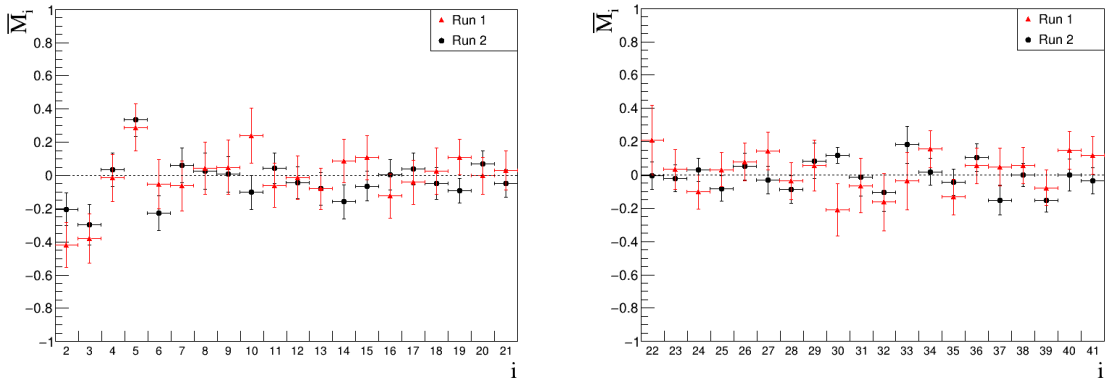


Figure 12.2: A comparison between Run1(red triangles) and Run2(black circles) measurements of 40 normalized moments, $\overline{M}_i$, of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay in the $1.1 < q^2 < 6.0$ GeV^2/c^4 and $1330 < m(K^+\pi^-) < 1530$ MeV/c^2 ranges. The error bars correspond to the sums in quadrature of the statistical and systematic uncertainties.

12.2.1 The consistency relations

As it is described in Sec. 8.2, normalized angular moments are not independent variables but they are related through seven formulas (Eq. 8.7). Each $\mathcal{R}_i$ should be equal to zero and any significant deviation from that value would suggest a departure from assumptions of the SPD-wave model in which the considered $m(K^+\pi^-)$ region contains contribution from the S-wave $K_0(1430)$, the P-wave $K_1^*(1410)$ and the D-wave $K_2^*(1430)$. Figure 12.3 presents the values and significances of the seven $\mathcal{R}_i$ variables and they are all consistent with zero.

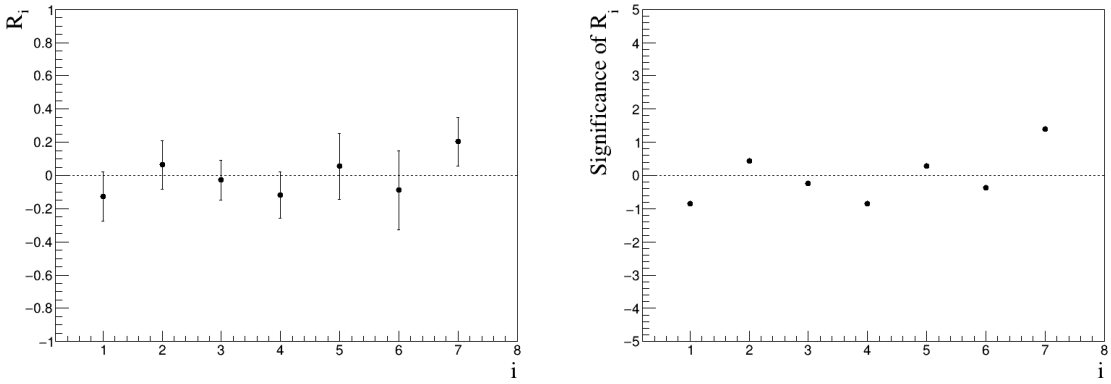


Figure 12.3: Values (left) and significances (right) for the consistency relations $\mathcal{R}_i$ defined in Eq. 8.7, integrated over the ranges $q^2 \in [1.1, 6.0]$ GeV^2/c^4 and $m_{K\pi} \in [1330, 1530]$ MeV/c^2.

12.2.2 D-wave fraction

As it is depicted in Table 8.1, the amplitudes can be derived using angular moments. It is not possible to extract the S- or the P-wave component alone, since those two amplitudes are always together. That is not the case of the D-wave component, and information from Table 8.1 helps to calculate the following amplitudes:

$$|D|_0^2 = \frac{7}{18}(M_5 - 2\sqrt{5}M_{10}), \quad |D|_{\perp,\parallel}^2 = -\frac{7}{12}(M_5 + \sqrt{5}M_{10} \pm \sqrt{\frac{5}{3}}M_{23}). \tag{12.1}$$

The formulas above tell us that the fraction of the D-wave component F_D can be calculated using only the moments $\overline{M}_5$ and $\overline{M}_{10}$:

$$F_D = -\frac{7}{18}(2\overline{M}_5 + 5\sqrt{5}\overline{M}_{10}). \tag{12.2}$$

The analysis based on the Run1 data set [22] gave the F_D fraction equal to -1.26±0.78. Thus, only an upper limit was provided on the D-wave fraction of $F_D < 0.29$ at 95% confidence level because of a limited number of candidates. This analysis gives better results and for the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay the D-wave fraction is equal to 0.18 ± 0.45. However, taking into account the level of uncertainty, this result is unfortunately still inconclusive.

Chapter 13

Conclusions

This dissertation is devoted to measuring 40 normalized angular moments and the differential branching fraction of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay in the $K^+\pi^-$ invariant mass range $1330 < m_{K\pi} < 1530$ MeV/c^2.

The new results based on the Run2 data set are consistent with the previous Run1 analysis. No clear signal from NP effects was observed despite the fact of using a bigger data sample than in the previous analysis. The D-wave fraction F_D was calculated using normalized moments $\overline{M}_5$ and $\overline{M}_{10}$ and is equal to 0.18 ± 0.45.

The FCNC processes, such as the presented one, are rare decays and their analysis can pose a challenge because of a low number of candidates detected by the experiment. An increase in this number may benefit future results. That is why a natural continuation would be to perform a combined analysis of Run1 and Run2, and further in the future, even the Run3 data set. A bigger data set will result in a higher number of signal candidates and will allow us to explore the higher mass of the $K^+\pi^-$ system with greater precision. For example, instead of exploring the presented wide range of the mass spectrum, it will be possible to do it in several smaller ranges, separating P-wave contribution from S- and D-wave contributions.

Although this work has brought only raw results without any comparison with the SM predictions, it is enough to explore several characteristics of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay. However, additional information from the theory would definitely increase our understanding of rare $b \to s\mu^+\mu^-$ processes and, perhaps, it would lead us to yet unexplored fields of New Physics beyond the Standard Model.

Appendix A

Decay description

A.1 Angular distribution

The differential decay rate of the $\overline{\mathrm{B}}^0 \to \overline{\mathrm{K}}^* \ell^+ \ell^-$ can be written as:

$$\begin{aligned}\frac{d^4\Gamma}{dq^2 d\cos\theta_K \cos\theta_\ell d\phi} = \frac{9}{32\pi}[&I_{1s}sin2\theta_K + I_{1c}cos2\theta_K + (I_{2s}sin2\theta_K + I_{2c}cos2\theta_K)cos2\theta_\ell \\ &+ I_3 sin2\theta_K sin2\theta_\ell cos2\phi + I_4 sin2\theta_K sin2\theta_\ell cos\phi + I_5 sin2\theta_K sin\theta_\ell cos\phi \\ &+ (I_{6s}sin2\theta_K + I_{6c}cos2\theta_K)cos\theta_\ell + I_7 sin2\theta_K sin\theta_\ell sin\phi \\ &+ I_8 sin2\theta_K sin2\theta_\ell sin\phi + I_9 sin2\theta_K sin2\theta_\ell sin2\phi].\end{aligned} \tag{A.1}$$

Twelve q^2 dependent angular coefficients $I_i(q^2)$ from Eq. A.1 in terms of transversity amplitudes (A_i) are given by:

$$I_{1s} = \frac{2+\beta_\ell^2}{4}[|A_\perp^L|^2 + |A_\parallel^L|^2 + |A_\perp^R|^2 + |A_\parallel^R|] + \frac{4m_\ell^2}{q^2} Re(A_\perp^L A_\perp^{R^*} + A_\parallel^L A_\parallel^{R^*}), \tag{A.2}$$

$$I_{1c} = |A_0^L|^2 + |A_0^R|^2 + \frac{4m_\ell^2}{q^2}[|A_t|^2 + 2Re(A_0^L A_0^{R^*})] + \beta_\ell^2 |A_S|^2, \tag{A.3}$$

$$I_{2s} = \frac{\beta_\ell^2}{4}[|A_\perp^L|^2 + |A_\parallel^L|^2 + |A_\perp^R|^2 + |A_\parallel^R|^2], \tag{A.4}$$

$$I_{2c} = -\beta_\ell^2[|A_0^L|^2 + |A_0^R|^2], \tag{A.5}$$

$$I_3 = \frac{1}{2}\beta_\ell^2[|A_\perp^L|^2 - |A_\parallel^L|^2 + |A_\perp^R|^2 - |A_\parallel^R|^2], \tag{A.6}$$

$$I_4 = \frac{1}{\sqrt{2}}\beta_\ell^2[Re(A_0^L A_\parallel^{L^*}) + Re(A_0^R A_\parallel^{R^*})], \tag{A.7}$$

$$I_5 = \sqrt{2}\beta_\ell[Re(A_0^L A_\perp^{L^*}) - Re(A_0^R A_\perp^{R^*}) - \frac{m_\ell}{\sqrt{q^2}} Re(A_\parallel^L A_S^* + A_\parallel^R A_S^*)], \tag{A.8}$$

$$I_{6s} = 2\beta_\ell[Re(A_\parallel^L A_\perp^{L^*}) - Re(A_\parallel^R A_\perp^{R^*})], \tag{A.9}$$

$$I_{6c} = 4\beta_\ell \frac{m_\ell}{\sqrt{q^2}}[Re(A_0^L A_S^*) + Re(A_0^R A_S^*)], \tag{A.10}$$

$$I_7 = \sqrt{2}\beta_\ell[Im(A_0^L A_\parallel^{L^*}) - Im(A_0^R A_\parallel^{R^*}) + \frac{m_\ell}{\sqrt{q^2}} Im(A_\perp^L A_S^* + A_\perp^R A_S^*)], \tag{A.11}$$

$$I_8 = \frac{1}{\sqrt{2}}\beta_\ell^2[Im(A_0^L A_\perp^{L^*}) + Im(A_0^R A_\perp^{R^*})], \tag{A.12}$$

$$I_9 = \beta_\ell^2[Im(A_\parallel^{L^*} A_\perp^L) + Im(A_\parallel^{R^*} A_\perp^R)], \tag{A.13}$$

and

$$\beta_\ell = \sqrt{1 - \frac{4m_\ell^2}{q^2}}. \tag{A.14}$$

A.2 Decay amplitudes

The decay amplitudes shown in A.1 can be expressed in terms of the most relevant Wilson coefficients $\mathcal{C}_{7,9,10}^{\text{eff}}$ and q^2 dependent form factors $V(q^2)$, $A_{0,1,2}(q^2)$ $T_{1,2,3}(q^2)$. The effective Wilson coefficients in their effective form contain contributions from four quark operators:

$$\begin{aligned} A_\perp^{L,R} = N\sqrt{2}\lambda^{1/2}\Big[&\{(\mathcal{C}_9^{eff} + \mathcal{C}_9^{'eff}) \mp (\mathcal{C}_{10}^{eff} + \mathcal{C}_{10}^{'eff})\}\tfrac{V(q^2)}{m_B + m_{K^*}} \\ &+ \tfrac{2m_b}{q^2}(\mathcal{C}_7^{eff} + \mathcal{C}_7^{'eff})T_1(q^2)\Big] \end{aligned} \tag{A.15}$$

$$\begin{aligned} A_\parallel^{L,R} = -N\sqrt{2}(m_B^2 - m_{K^*}^2)\Big[&\{(\mathcal{C}_9^{eff} - \mathcal{C}_9^{'eff}) \mp (\mathcal{C}_{10}^{eff} - \mathcal{C}_{10}^{'eff})\}\tfrac{A_1(q^2)}{m_B - m_{K^*}} \\ &+ \tfrac{2m_b}{q^2}(\mathcal{C}_7^{eff} + \mathcal{C}_7^{'eff})T_2(q^2)\Big] \end{aligned} \tag{A.16}$$

$$\begin{aligned} A_0^{L,R} = &-\tfrac{N}{2m_{K^*}\sqrt{q^2}}\Big[\{(\mathcal{C}_9^{eff} - \mathcal{C}_9^{'eff}) \mp (\mathcal{C}_{10}^{eff} - \mathcal{C}_{10}^{'eff})\}\times \\ &\times\Big\{(m_B^2 - m_{K^*}^2 - q^2)(m_B + m_{K^*})A_1(q^2) - \tfrac{\lambda A_2(q^2)}{m_B + m_{K^*}}\Big\} + \\ &+2m_b(\mathcal{C}_7^{eff} - \mathcal{C}_7^{'eff})\Big\{(m_B^2 + 3m_{K^*}^2 - q^2)T_2(q^2) - \tfrac{\lambda}{m_B^2 - m_{K^*}^2}T_3(q^2)\Big\}\Big], \end{aligned} \tag{A.17}$$

$$A_t = N\frac{\lambda^{1/2}}{\sqrt{q^2}}\Big\{2(\mathcal{C}_{10}^{eff} - \mathcal{C}_{10}^{'eff})\Big\}A_0(q^2), \tag{A.18}$$

where

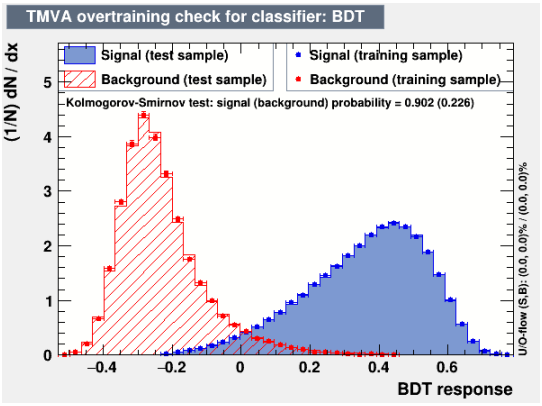
$$N = \sqrt{\frac{G_F^2\alpha^2}{3 * 2^{10}\pi^5 m_B^3}|V_{tb}V_{ts}^*|^2 q^2 \lambda^{1/2}\sqrt{1 - \frac{4m_\ell^2}{q^2}}} \tag{A.19}$$

$$\lambda = m_B^4 + m_{K^*}^4 + q^4 - 2(m_B^2 m_{K^*}^2 + m_{K^*}^2 q^2 + m_B^2 q^2) \tag{A.20}$$

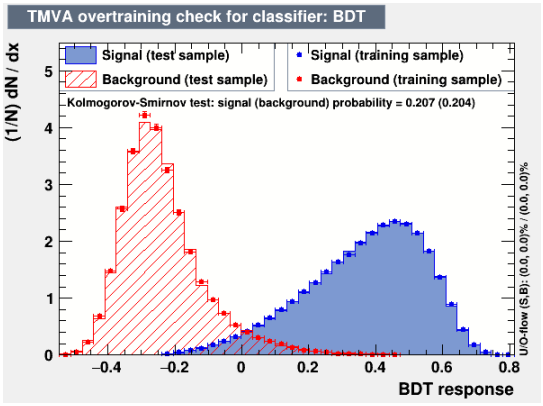
Appendix B

BDT overtraining check

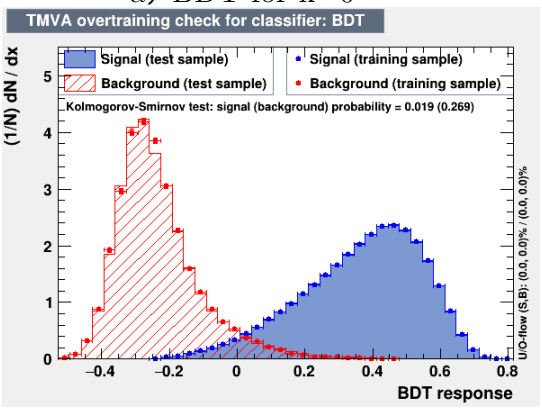
The dataset is divided into k=10 subsets(folds) of an approximately equal size, with which the BDT classifier is trained 10 times, each time using 9 folds for training and 1 fold for testing. This is repeated for each fold that is used for testing exactly once. The overtraining check is basically a comparison of the classifier's performance of the training and testing data sets. Overfitting manifests itself in a significantly better performance on the training dataset compared to the testing set.



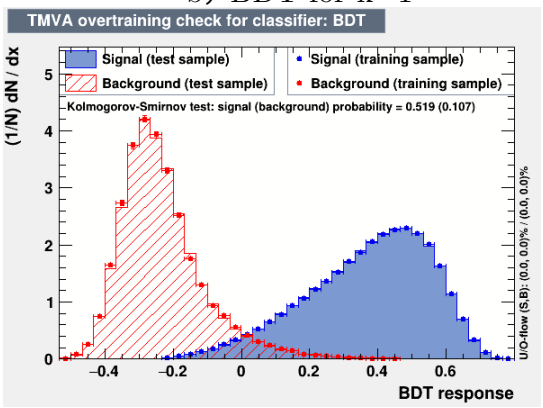
a) BDT for k=0



b) BDT for k=1

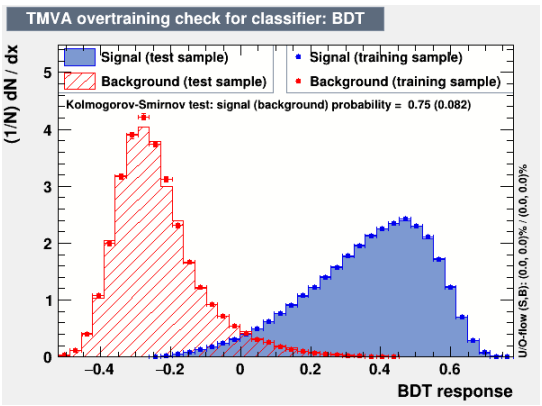


c) BDT for k=3

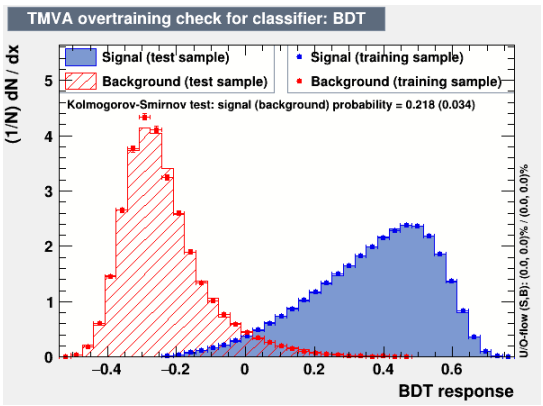


d) BDT for k=3

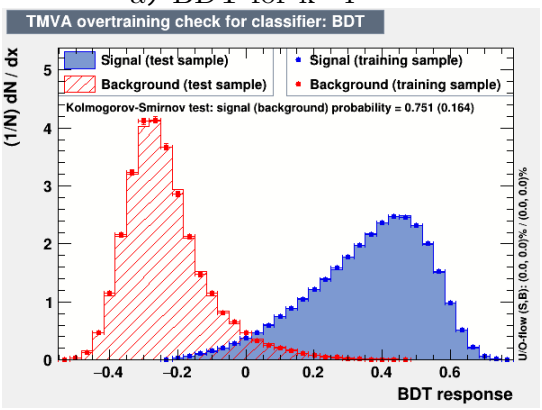
Figure B.1: The TMVA overtraining check for k= 0-5. Red colour (hashed) indicates the background and blue colour (shadowed) indicates the signal candidates.



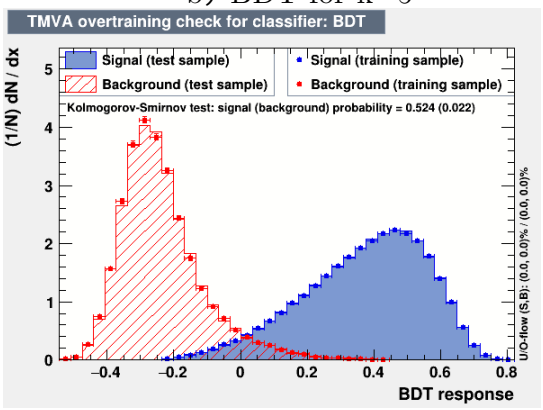
a) BDT for k=4



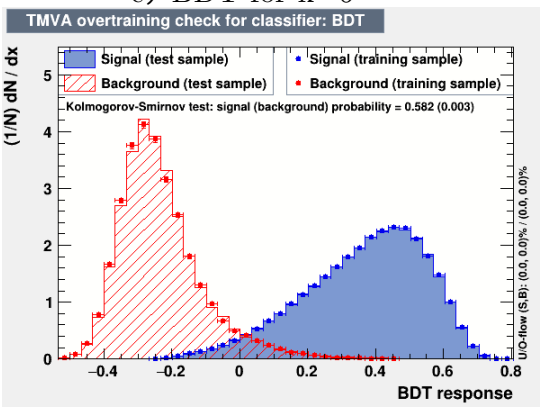
b) BDT for k=5



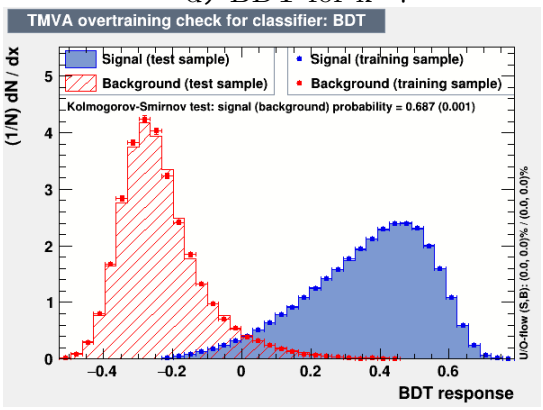
c) BDT for k=6



d) BDT for k=7



e) BDT for k=8



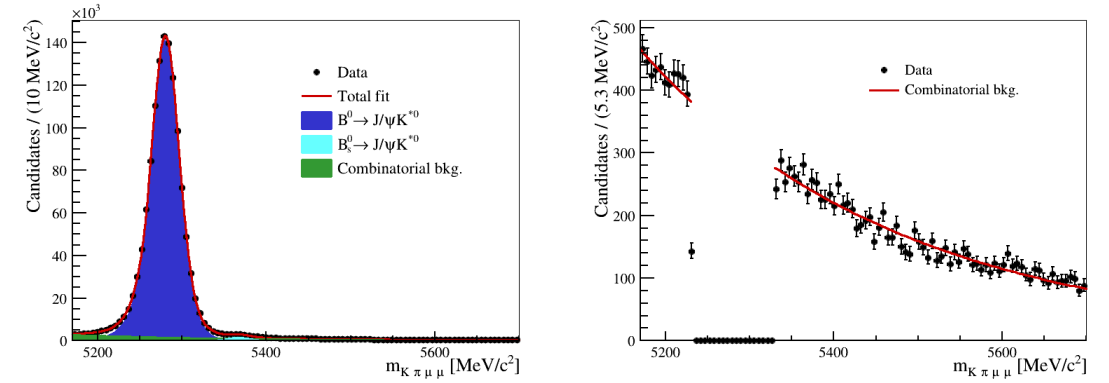
f) BDT for k=9

Figure B.2: The TMVA overtraining check for k= 6-9. Red colour (hashed) indicates the background and blue colour (shadowed) indicates the signal candidates.

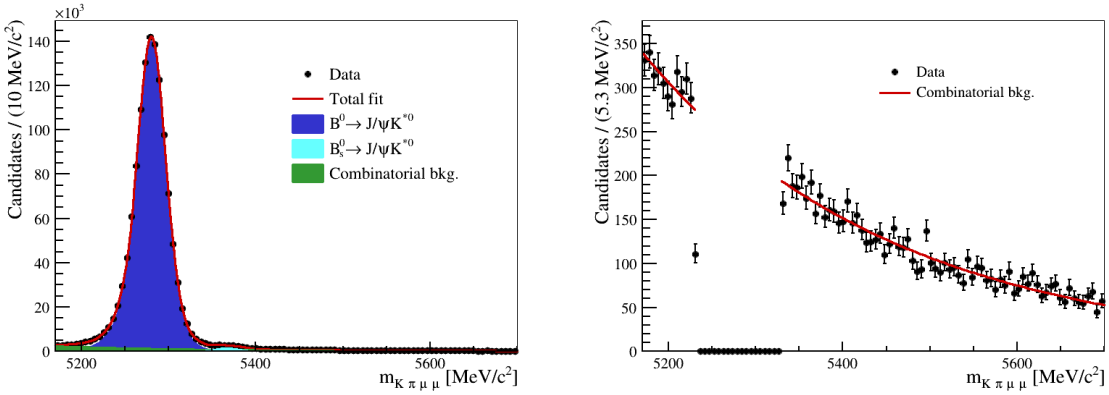
B.1 BDT optimization

As it is described in Section.5.9, the BDT cut value is optimized to check which cut is the best for extracting signal candidates. The invariant mass distribution of the B0 →

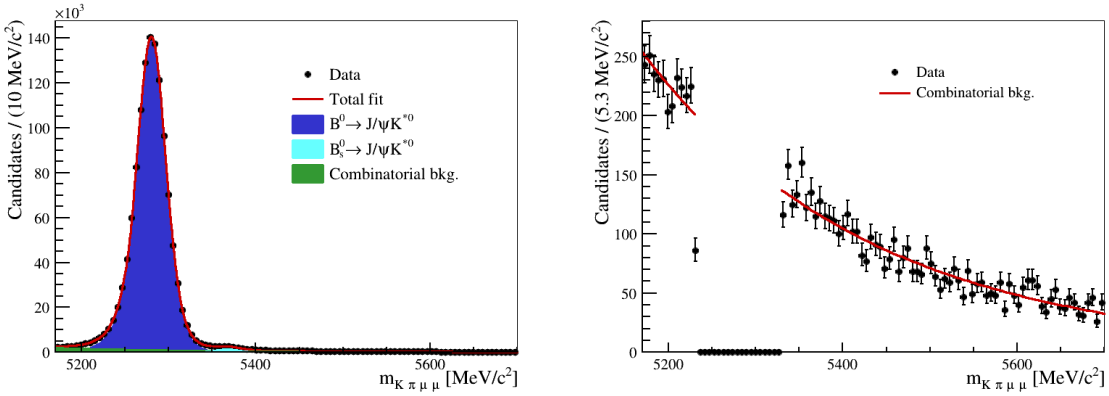
$K^+\pi^-\mu^+\mu^-$ candidates in the ranges $796 < m_{K\pi} < 996$ MeV/c^2 and $9.22 < q^2 < 9.96$ GeV^2/c^4 is fitted with the BDT cut value increasing from -0.15 to 0.5. This is presented in Figs. B.3-B.7. The number of signal and the number of background candidates obtained from the fit is used to calculate the significance $S\sqrt{S+B}$ (Eq. 5.13), which is presented as a function of BDT cut values in Fig. 5.9.



a) BDT >-0.15

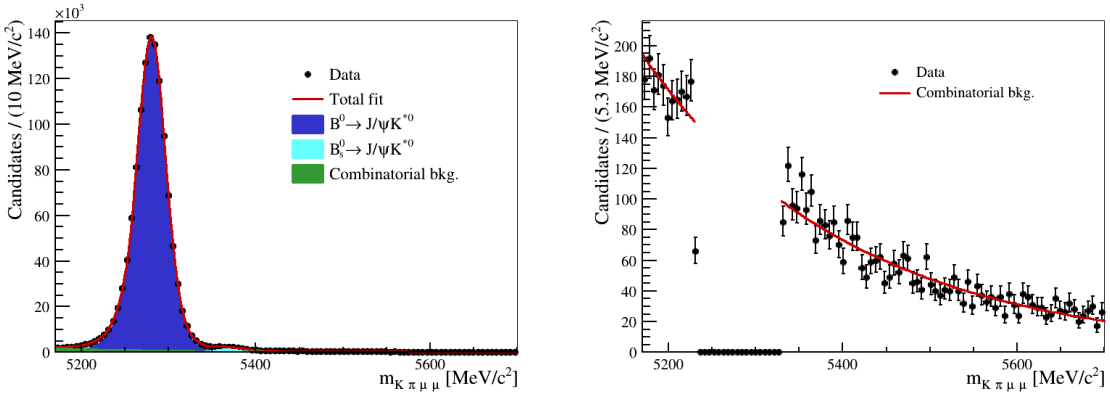


b) BDT >-0.10

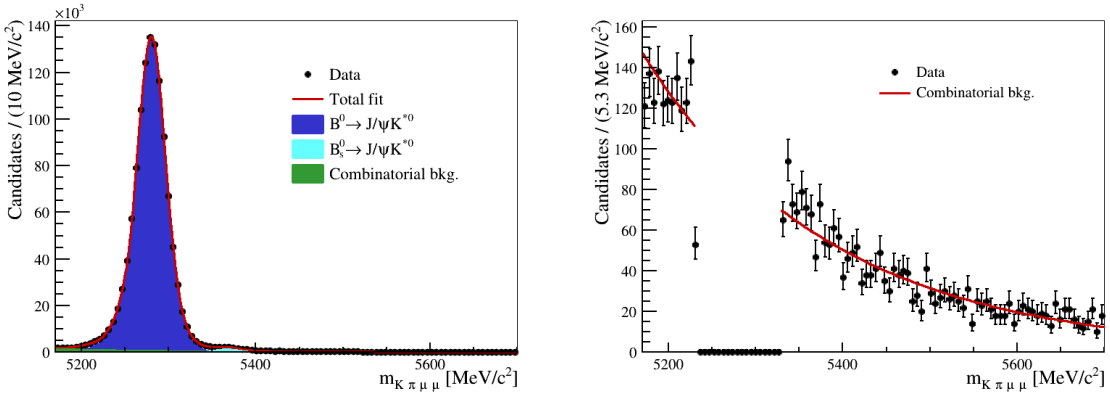


c) BDT >-0.05

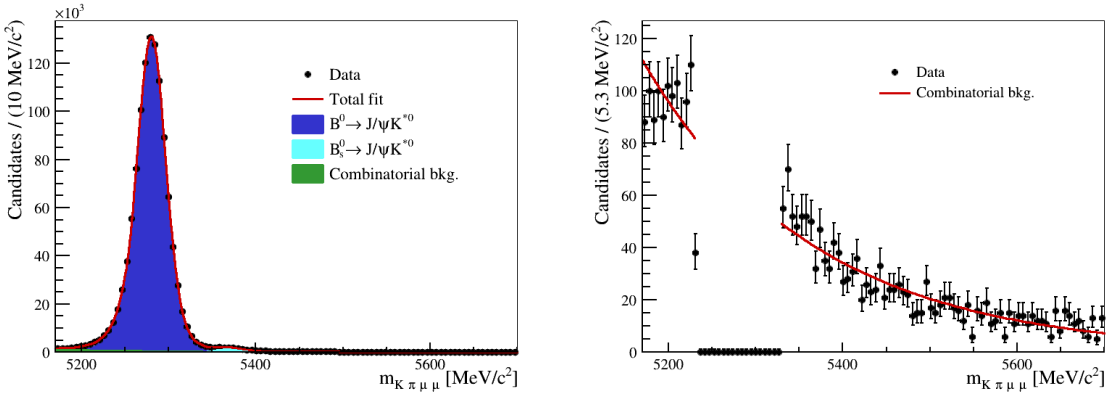
Figure B.3: Fits to the $B^0 \to J/\psi K^{*0}$ (left) and the $B^0 \to K^{*0}\mu^+\mu^-$ sidebands (right) used in the BDT optimization for several different BDT cuts. The BDT optimization is described in Sec.5.9.



a) BDT >0

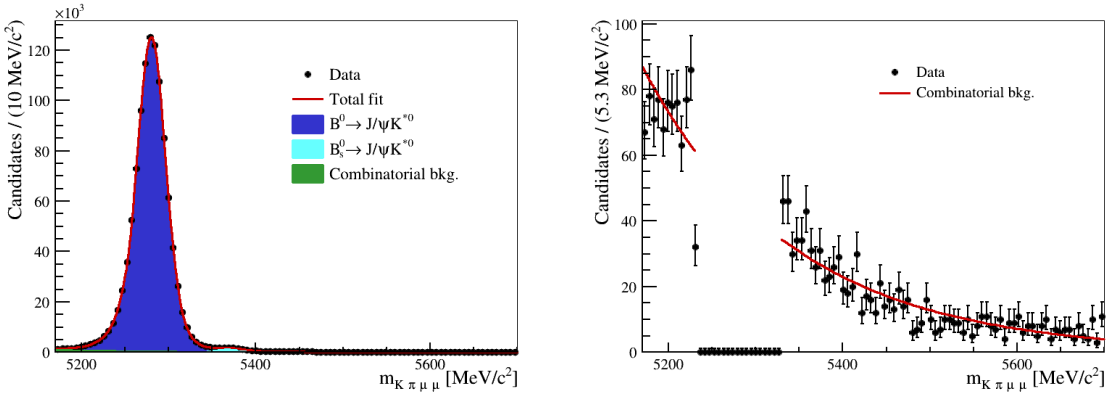


b) BDT >0.05

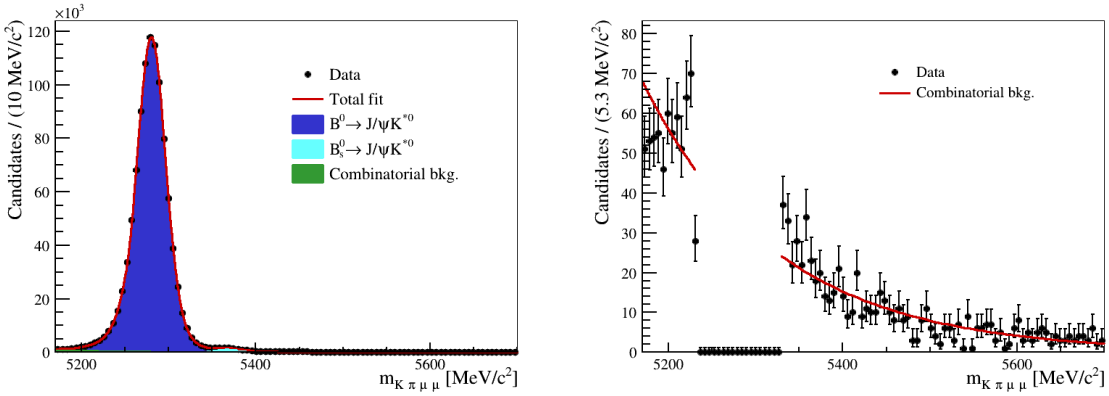


c) BDT >0.10

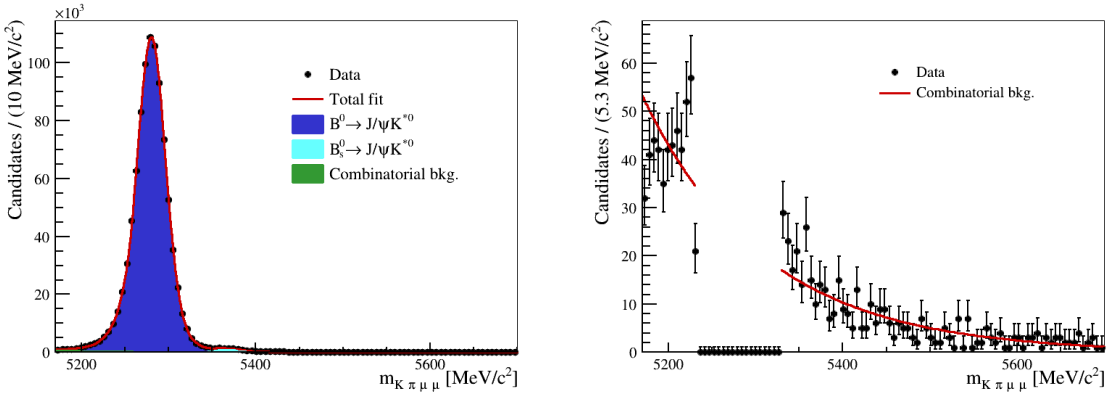
Figure B.4: Fits to the $B^0 \to J/\psi K^{*0}$ (left) and the $B^0 \to K^{*0}\mu^+\mu^-$ sidebands (right) used in the BDT optimization for several different BDT cuts. The BDT optimization is described in Sec.5.9.



a) BDT >0.15

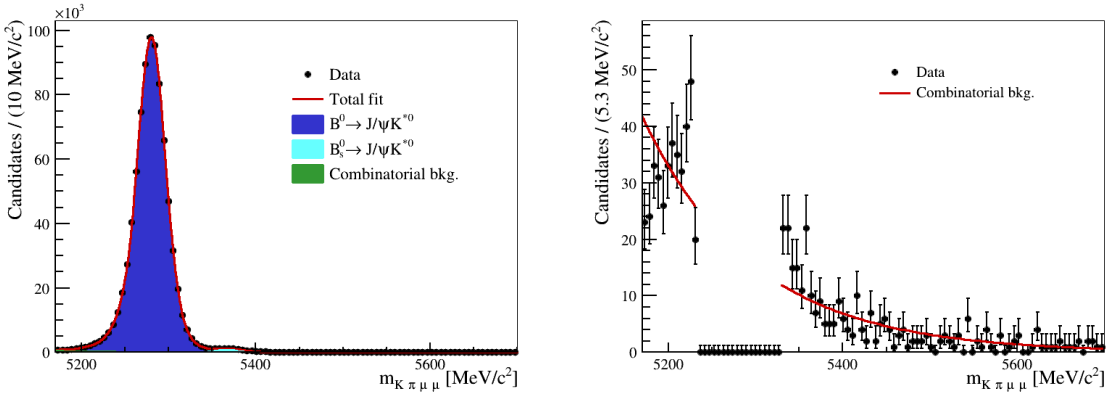


b) BDT >0.2

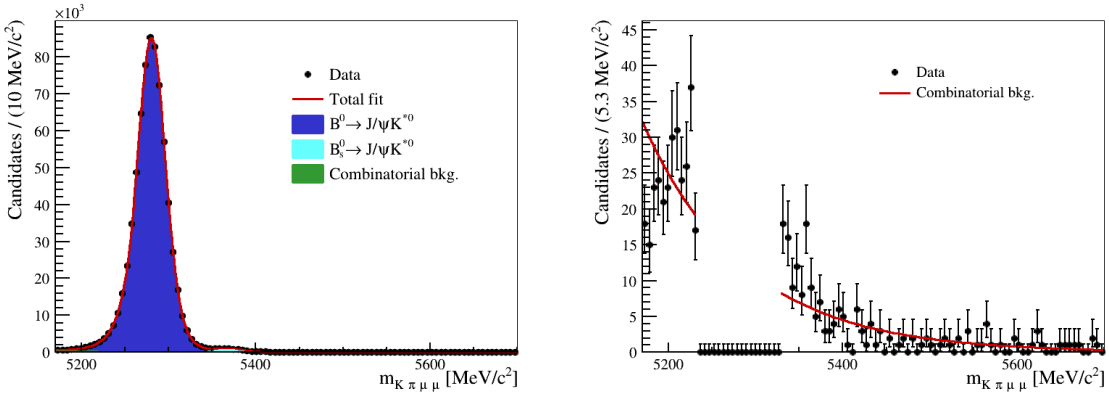


c) BDT >0.25

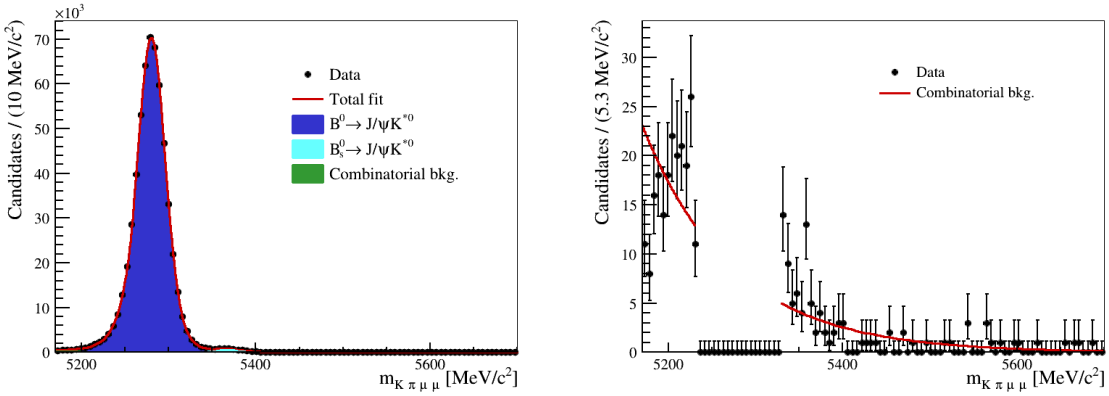
Figure B.5: Fits to the $B^0 \to J/\psi K^{*0}$ (left) and the $B^0 \to K^{*0}\mu^+\mu^-$ sidebands (right) used in the BDT optimization for several different BDT cuts. The BDT optimization is described in Sec.5.9.



a) BDT >0.3

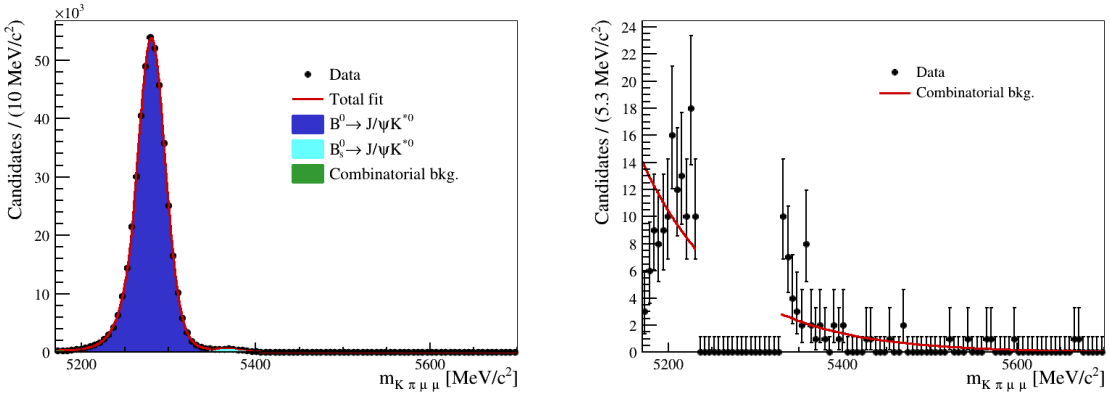


b) BDT >0.35

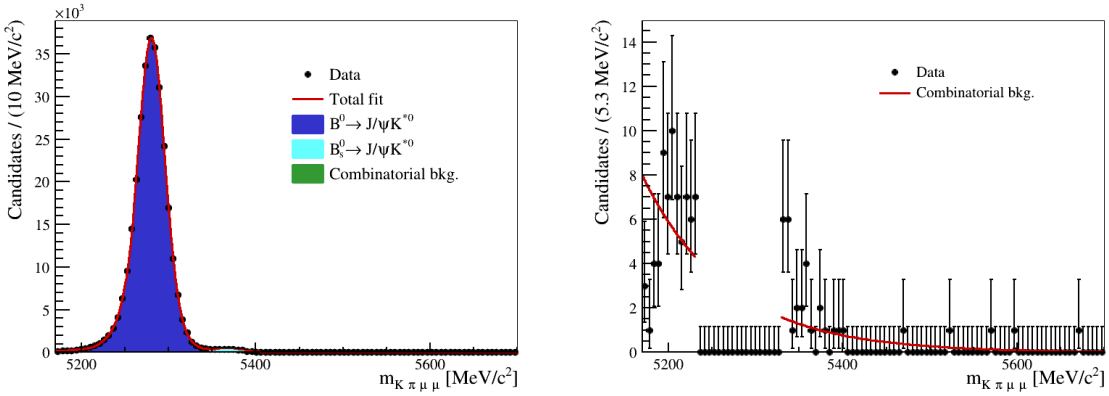


c) BDT >0.4

Figure B.6: Fits to the $B^0 \to J/\psi K^{*0}$ (left) and the $B^0 \to K^{*0}\mu^+\mu^-$ sidebands (right) used in the BDT optimization for several different BDT cuts. The BDT optimization is described in Sec.5.9.



a) BDT >0.45



b) BDT >0.5

Figure B.7: Fits to the $B^0 \to J/\psi K^{*0}$ (left) and the $B^0 \to K^{*0}\mu^+\mu^-$ sidebands (right) used in the BDT optimization for several different BDT cuts. The BDT optimization is described in Sec.5.9.

B.2 Agreement between data and simulation for BDT input variables

Figure B.2 presents the results of the reweighting of the candidates described in Section 6.2 for each variable used as an input to the BDT training.

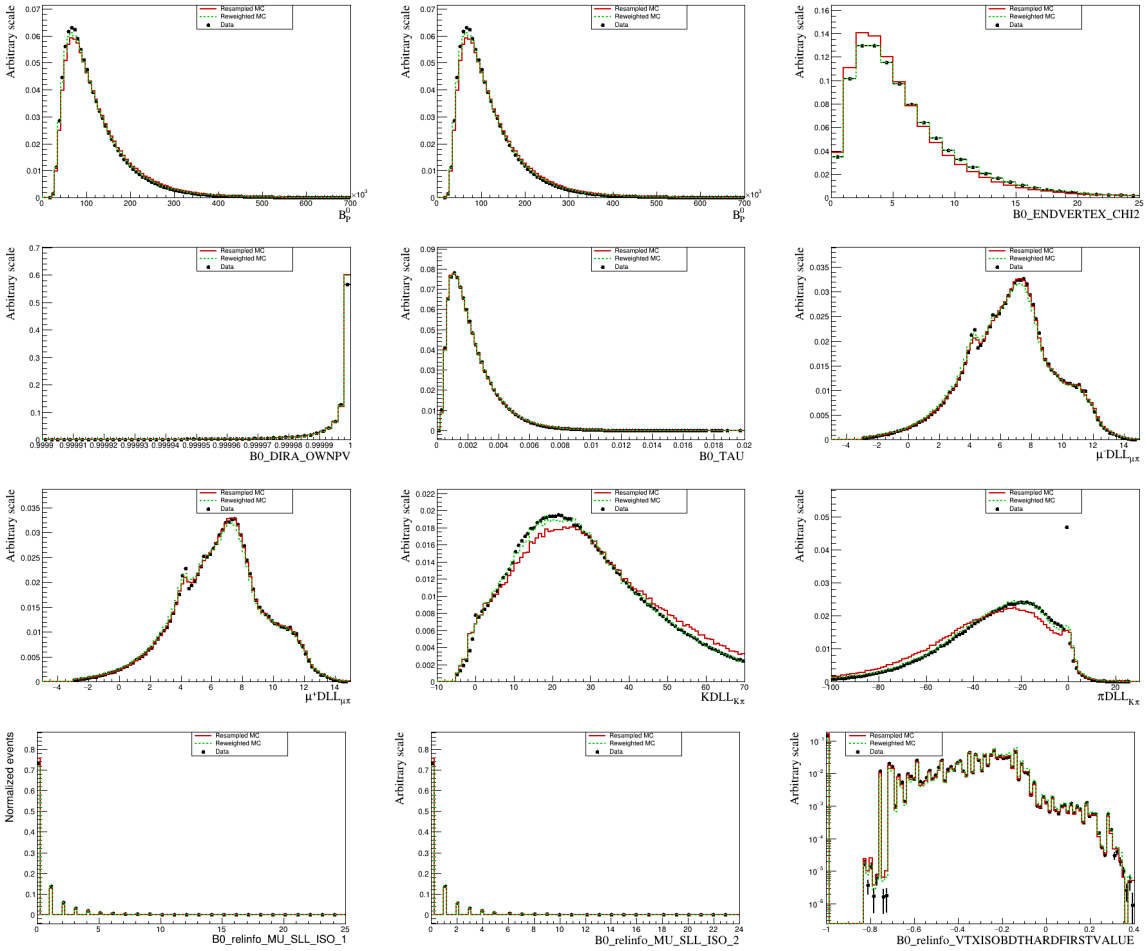


Figure B.8: Agreement between data and simulation for all variables used as an input to the BDT training. Black points correspond to *sWeighted* $B^0 \to J/\psi K^{*0}$ data. Red histogram shows the resampled $B^0 \to J/\psi K^{*0}$ MC distribution. Green histogram shows the distribution for resampled $B^0 \to J/\psi K^{*0}$ MC with applied candidates weights.

Appendix C

1D projections

Figures C.1-]C.8 presents 1-dimensional projections of the acceptance parametrization, described in Chapter 7, for variables $q^{2\prime}$, $\cos\theta_K$, $\cos\theta_\ell$, ϕ', and $m_{K\pi}$ in bins of $m_{K\pi}$ and q^2.

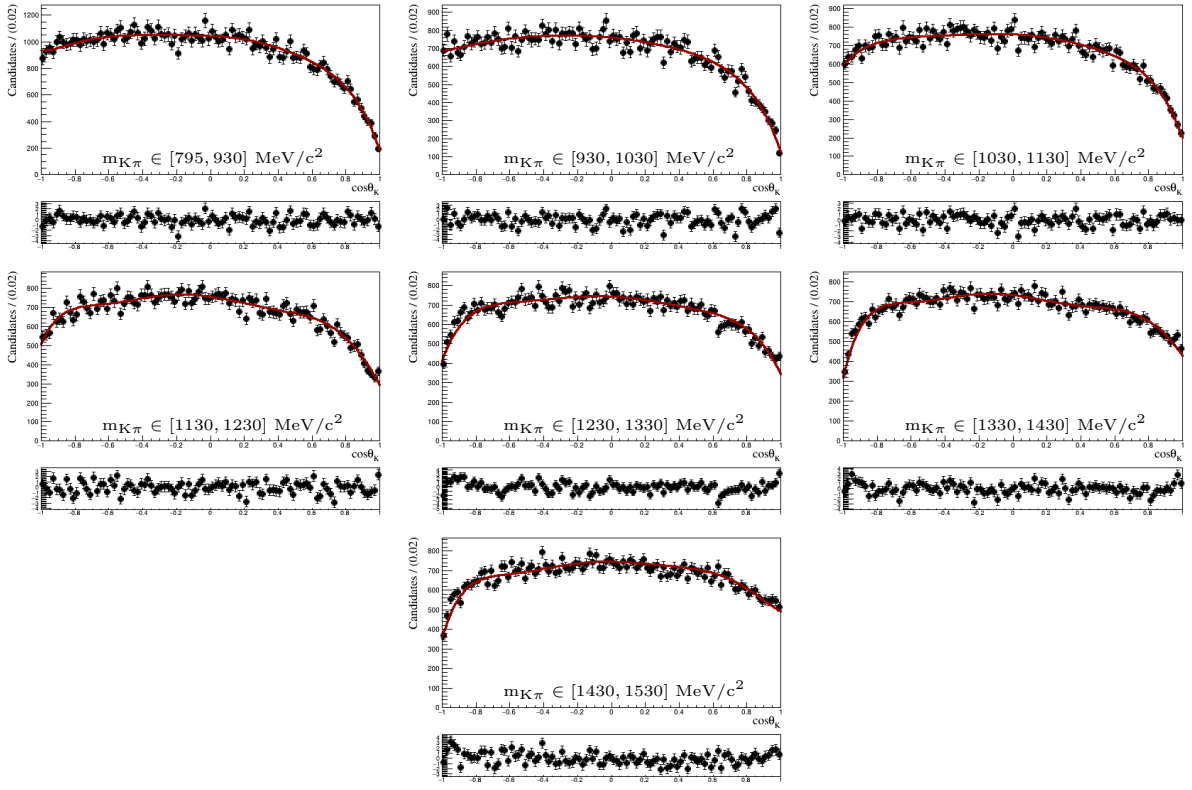


Figure C.1: The 1D projections of the 5D acceptance parametrization for $\cos\theta_K$ in bins of $m_{K\pi}$. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

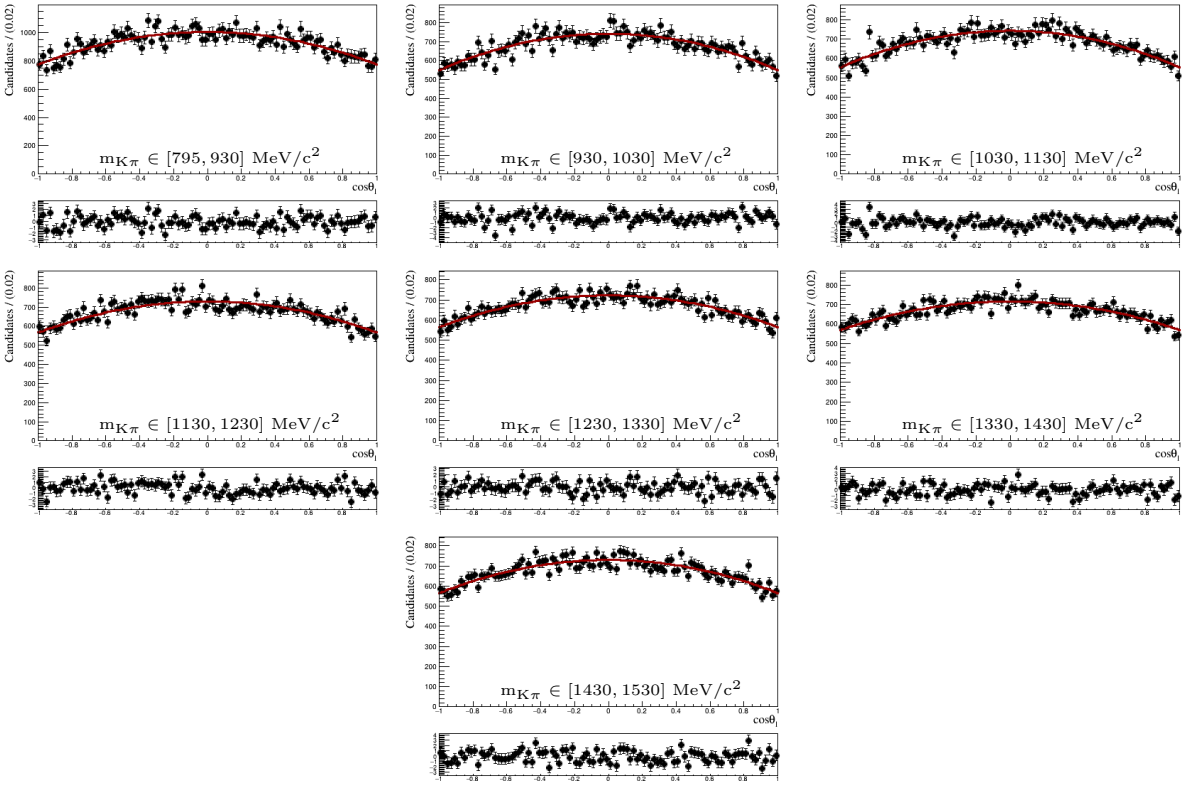


Figure C.2: The 1D projections of the 5D acceptance parametrization for $\cos\theta_\ell$ in bins of $m_{K\pi}$. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

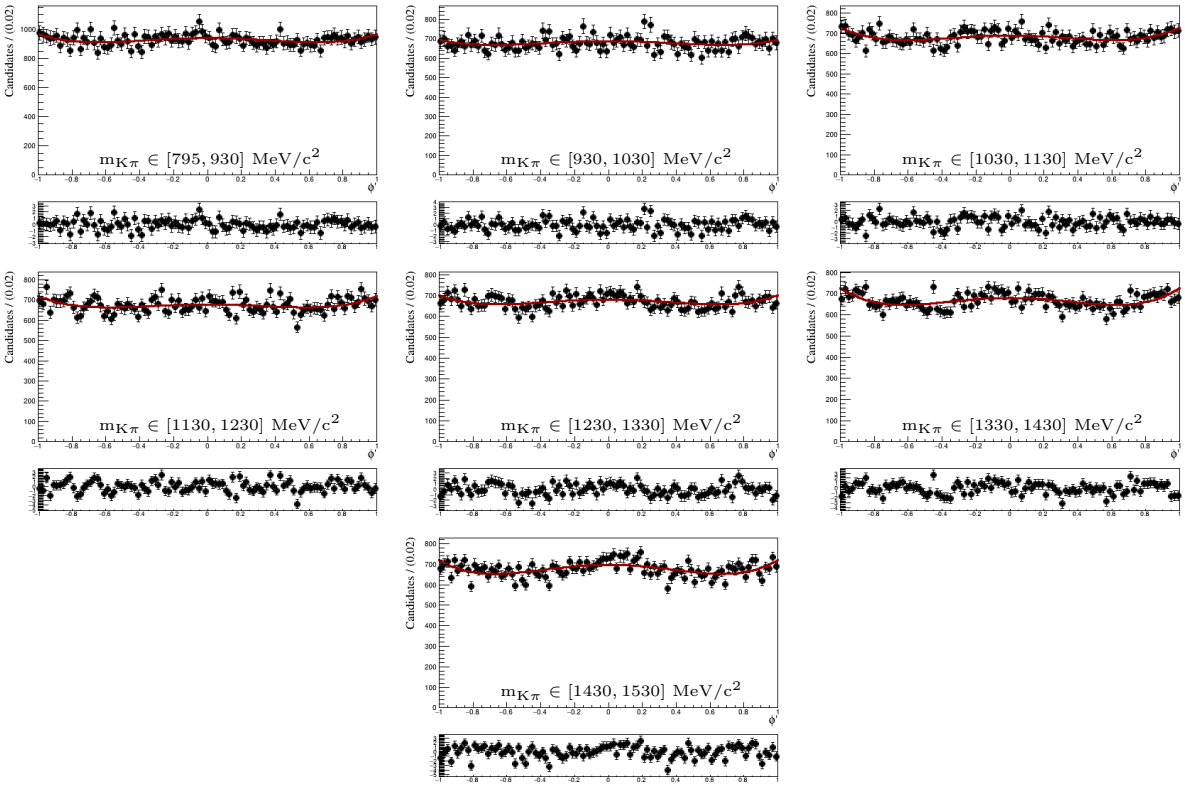


Figure C.3: The 1D projections of the 5D acceptance parametrization for ϕ' in bins of $m_{K\pi}$. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

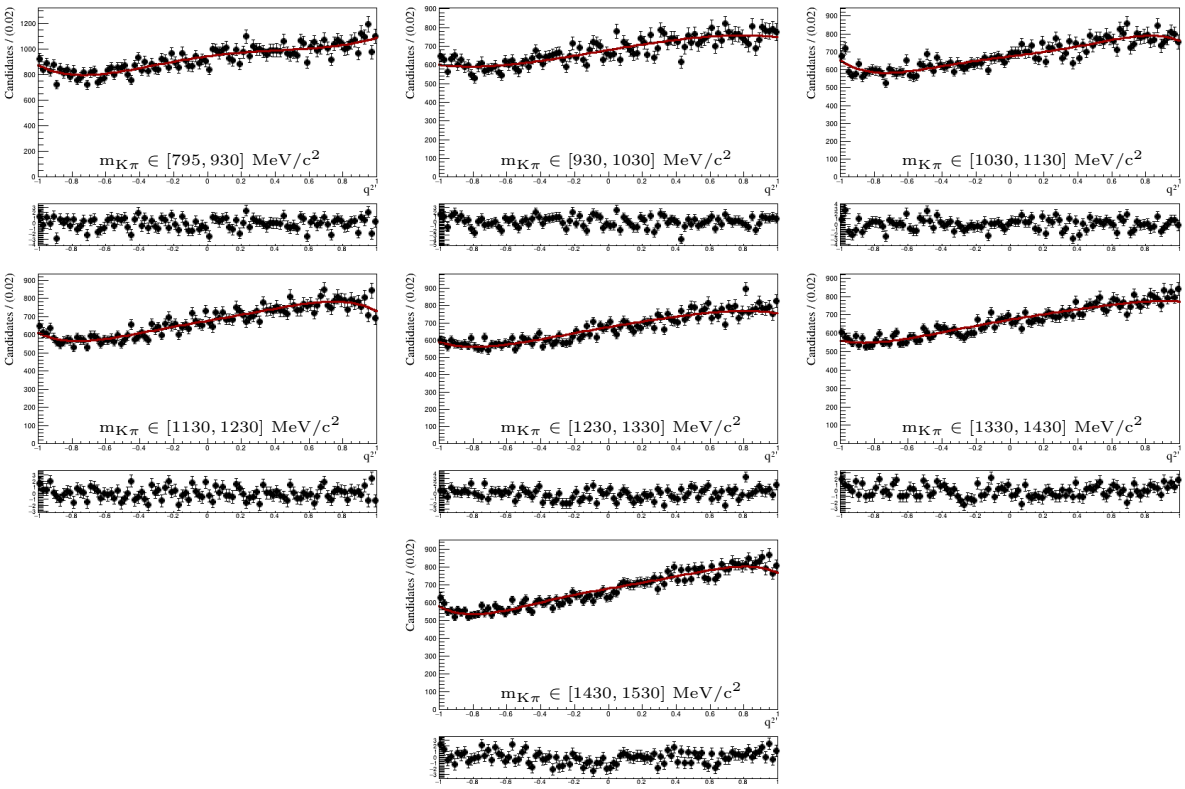


Figure C.4: The 1D projections of the 5D acceptance parametrization for $q^{2\prime}$ in bins of $m_{K\pi}$. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

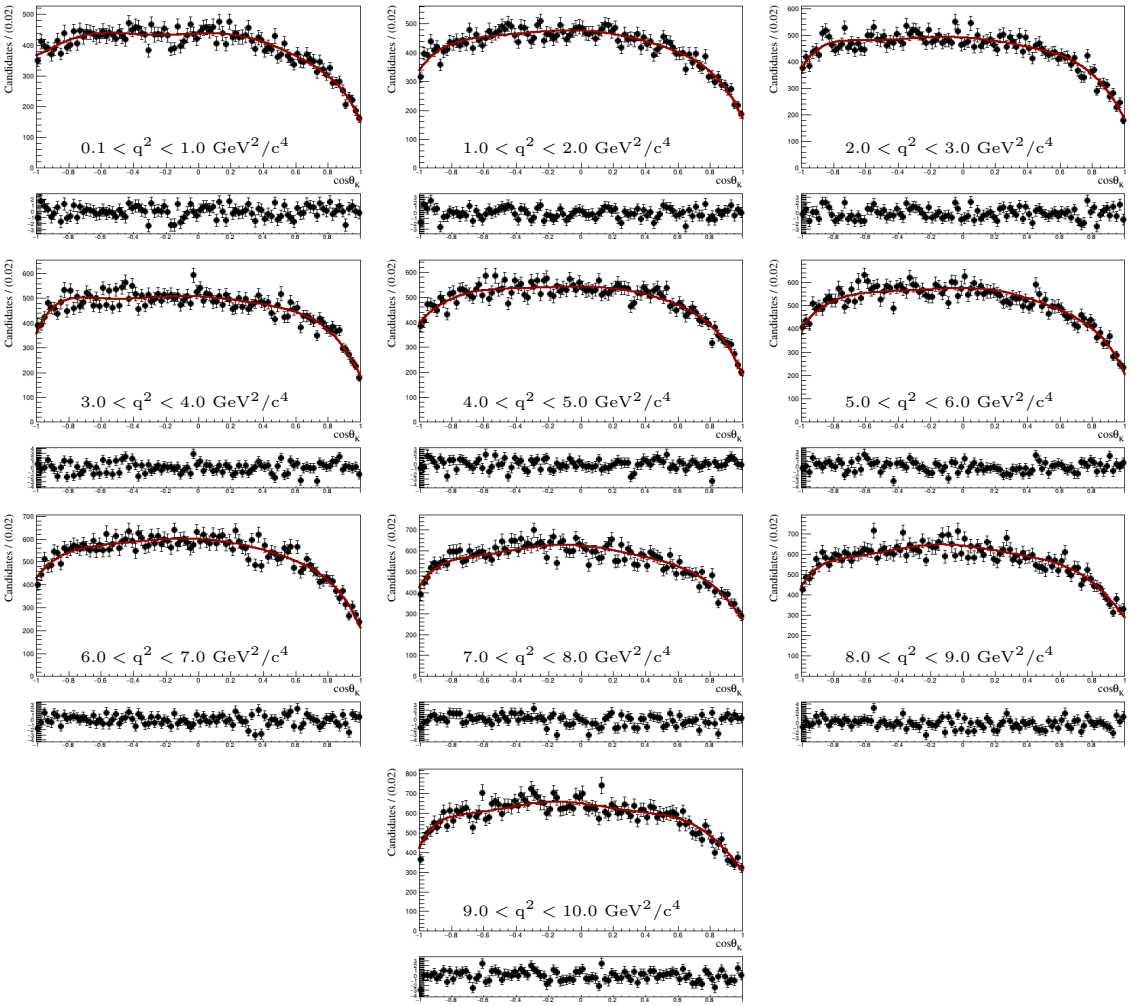


Figure C.5: The 1D projections of the 5D acceptance parametrization for $\cos\theta_K$ in bins of q^2. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

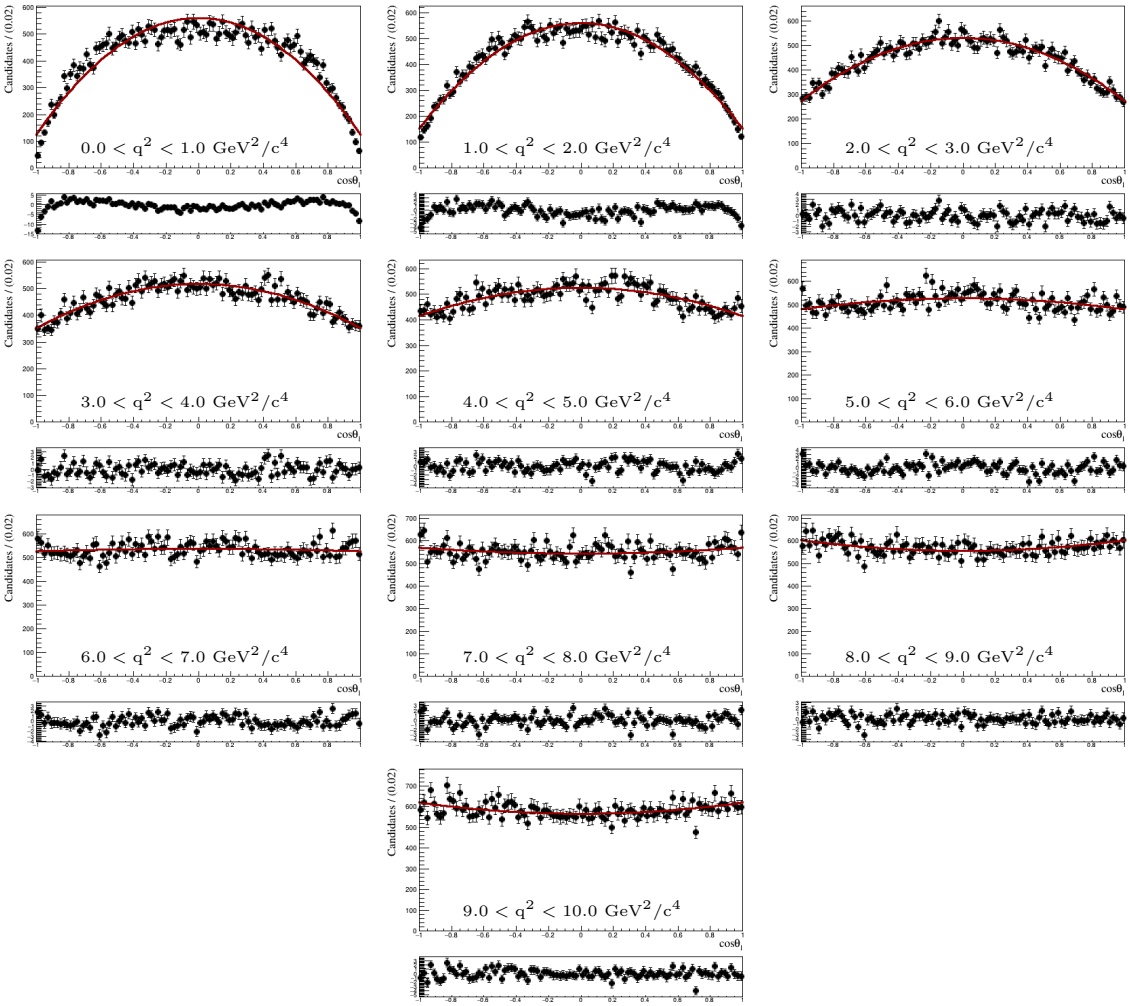


Figure C.6: The 1D projections of the 5D acceptance parametrization for $\cos\theta_\ell$ in bins of q^2. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

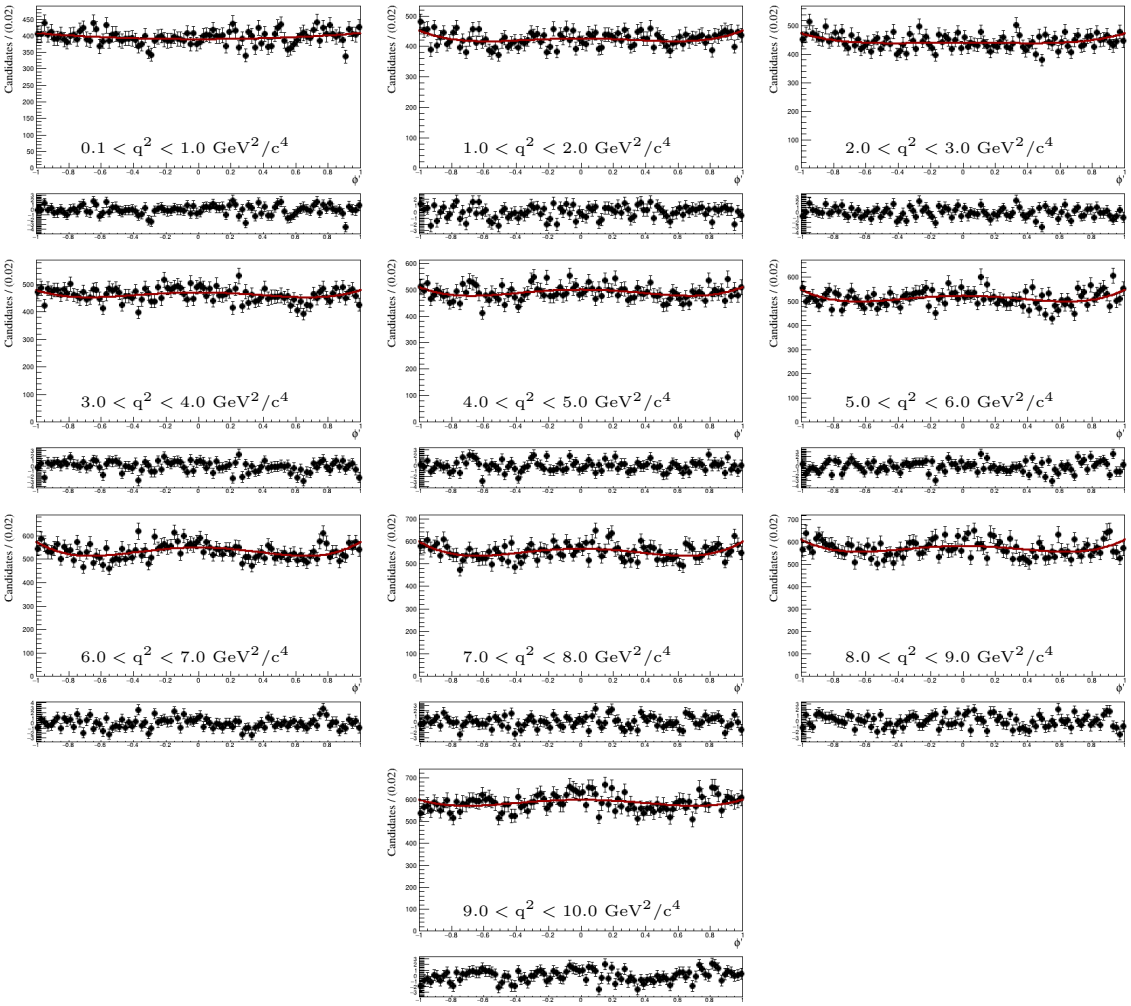


Figure C.7: The 1D projections of the 5D acceptance parametrization for ϕ' in bins of q^2. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

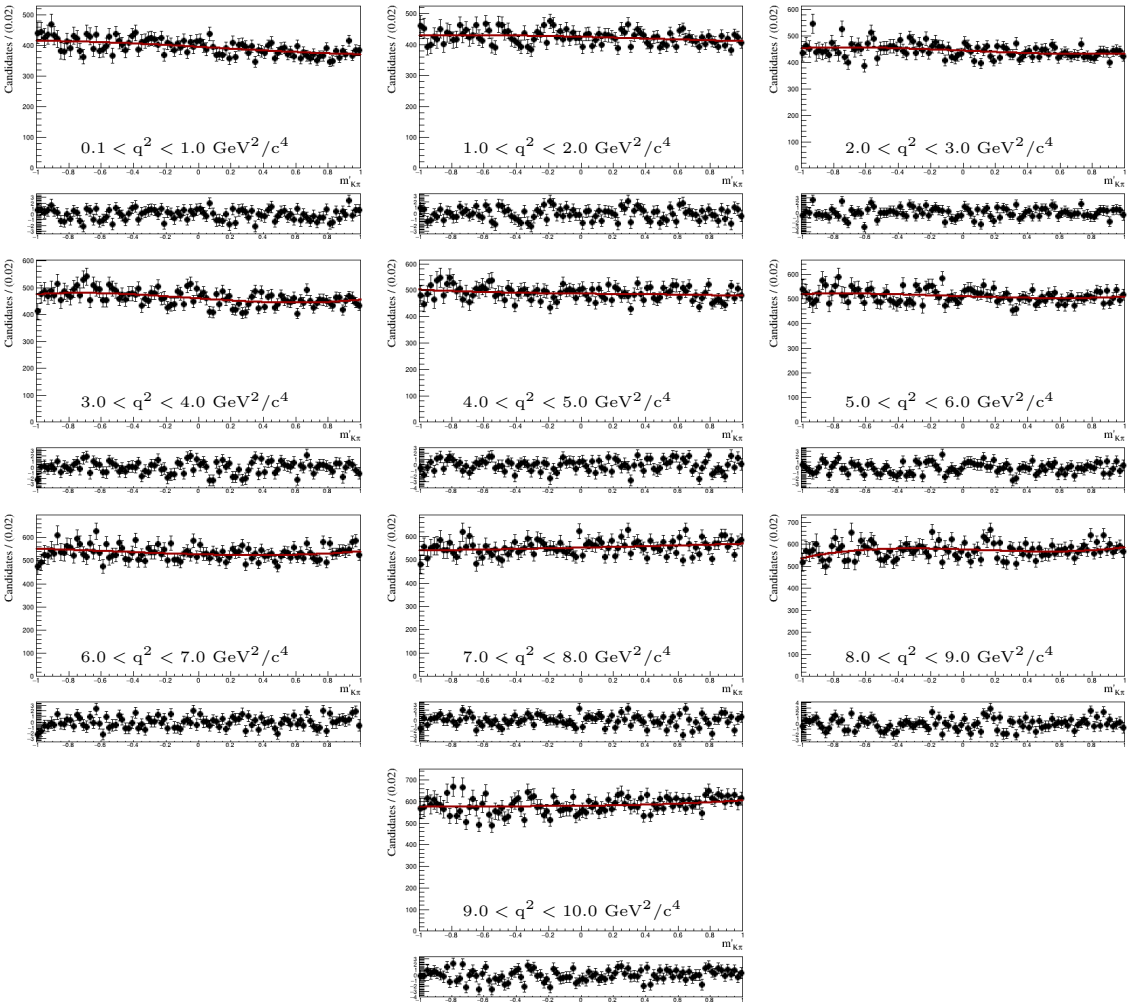


Figure C.8: The 1D projections of the 5D acceptance parametrization for $m'_{K\pi}$ in bins of q^2. The black data points show the reweighted simulated $B^0 \to K^+\pi^-\mu^+\mu^-$ MC and the red curve shows the parametrization.

Appendix D

Scaling factor calculations

Figure D.1 depicts the fits to the $m_{K\pi\mu\mu}$ distribution in the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC for scaling factor calculations described in Sec. 9.3.

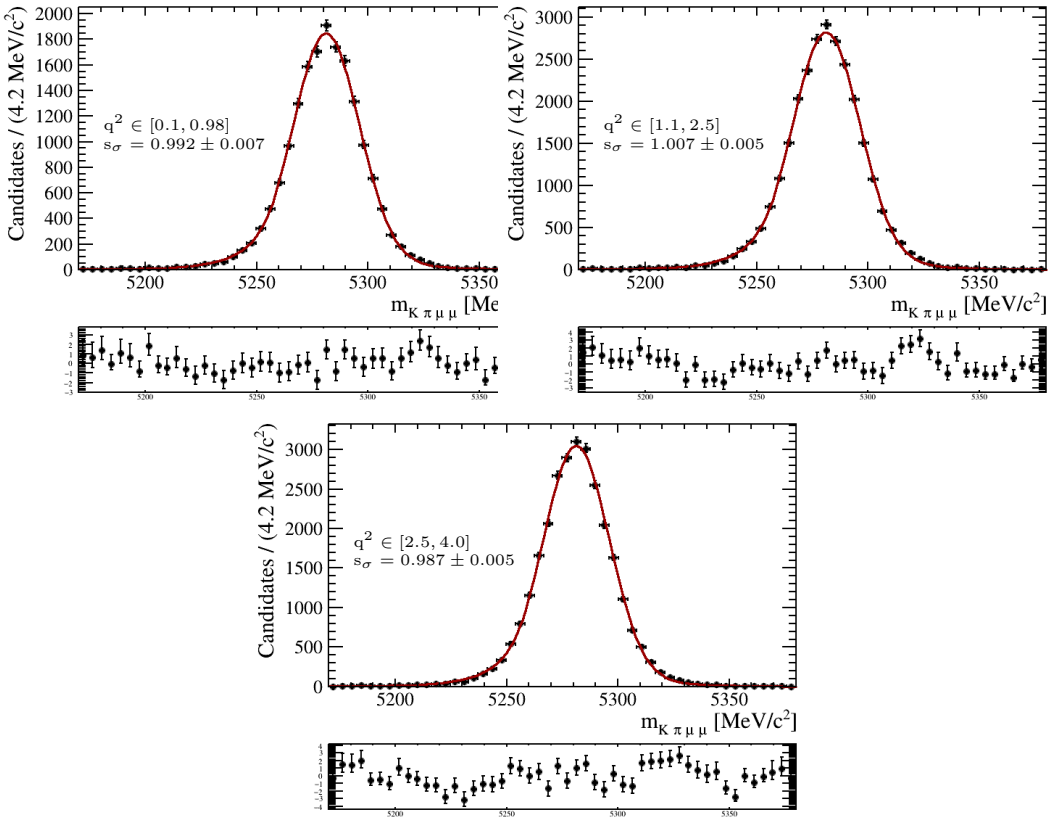


Figure D.1: Fits to the $K\pi\mu\mu$ invariant mass of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in three q^2 bins using MC simulation.

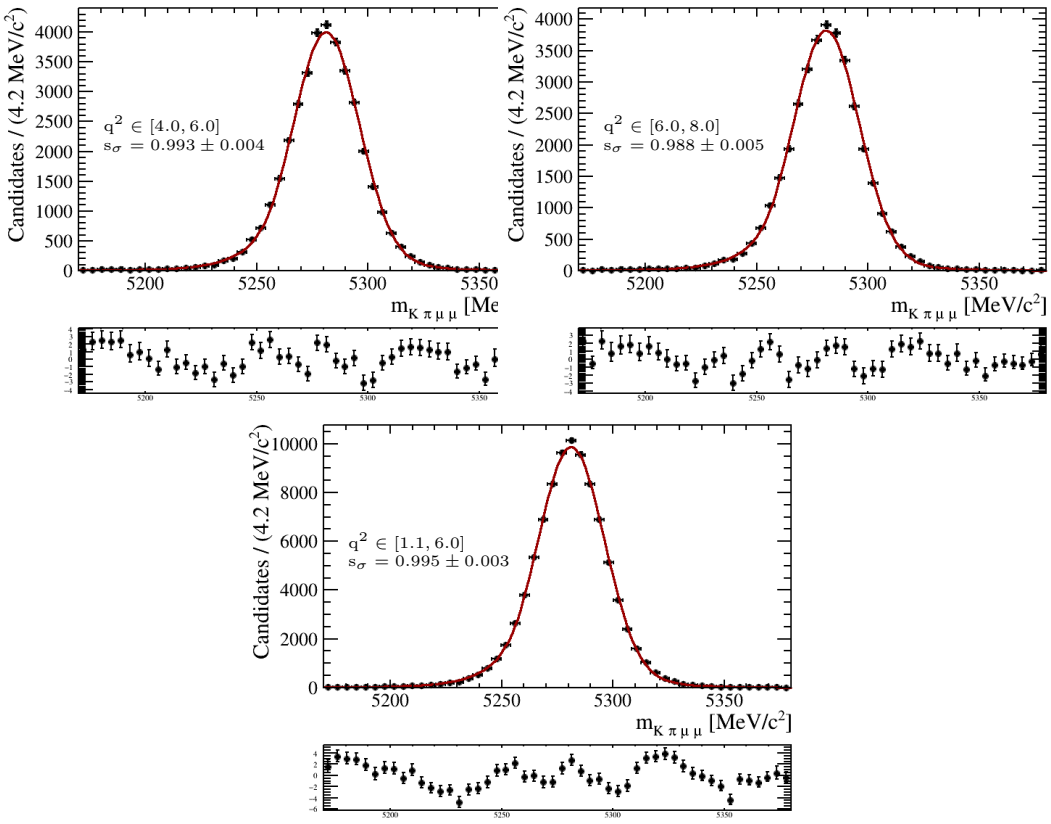


Figure D.2: Fits to the K$\pi\mu\mu$ invariant mass of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in three q^2 bins using MC simulation.

Appendix E

Systematic plots

The figures in E.1 and E.2 show the results of the toy studies for the branching fraction calculations and the angular analysis, as it is described in Section 10.2.

E.1 Toy studies for the branching fractions

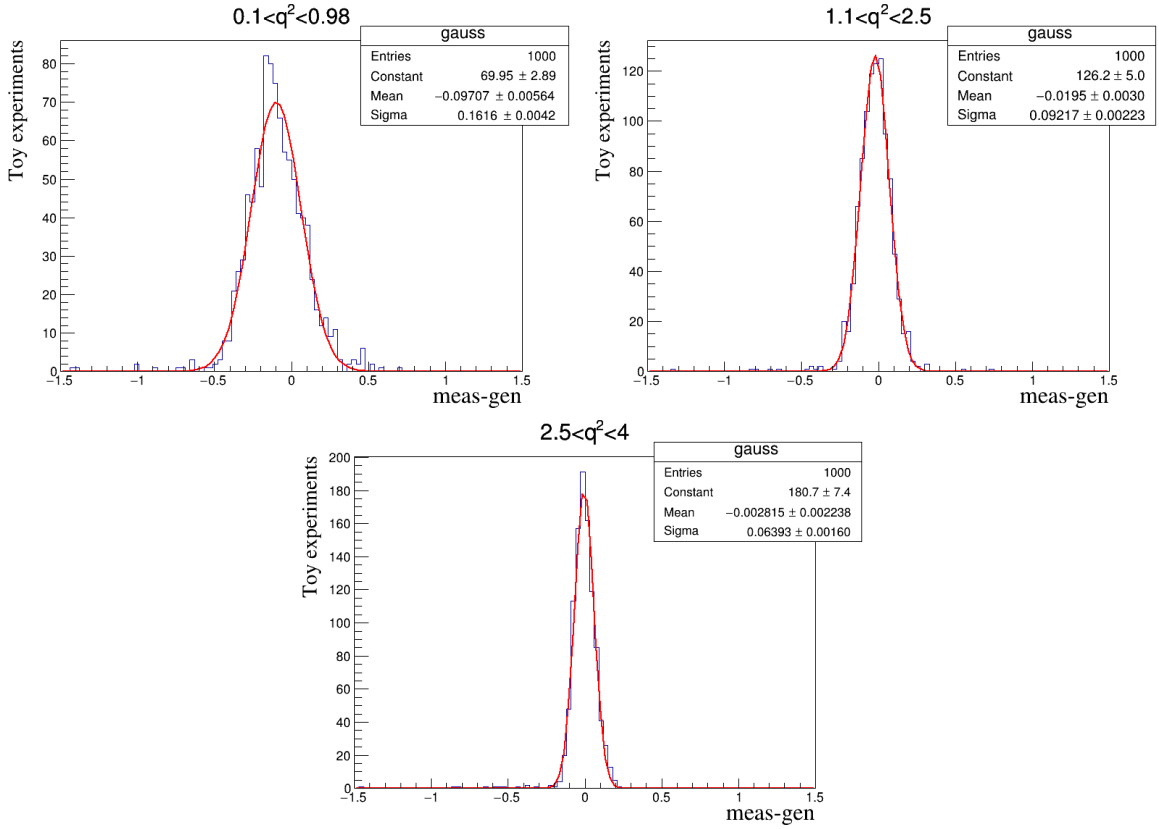


Figure E.1: Distributions of a difference between the measured differential branching fraction with original acceptance and the results based on generated toy datasets with new acceptance.

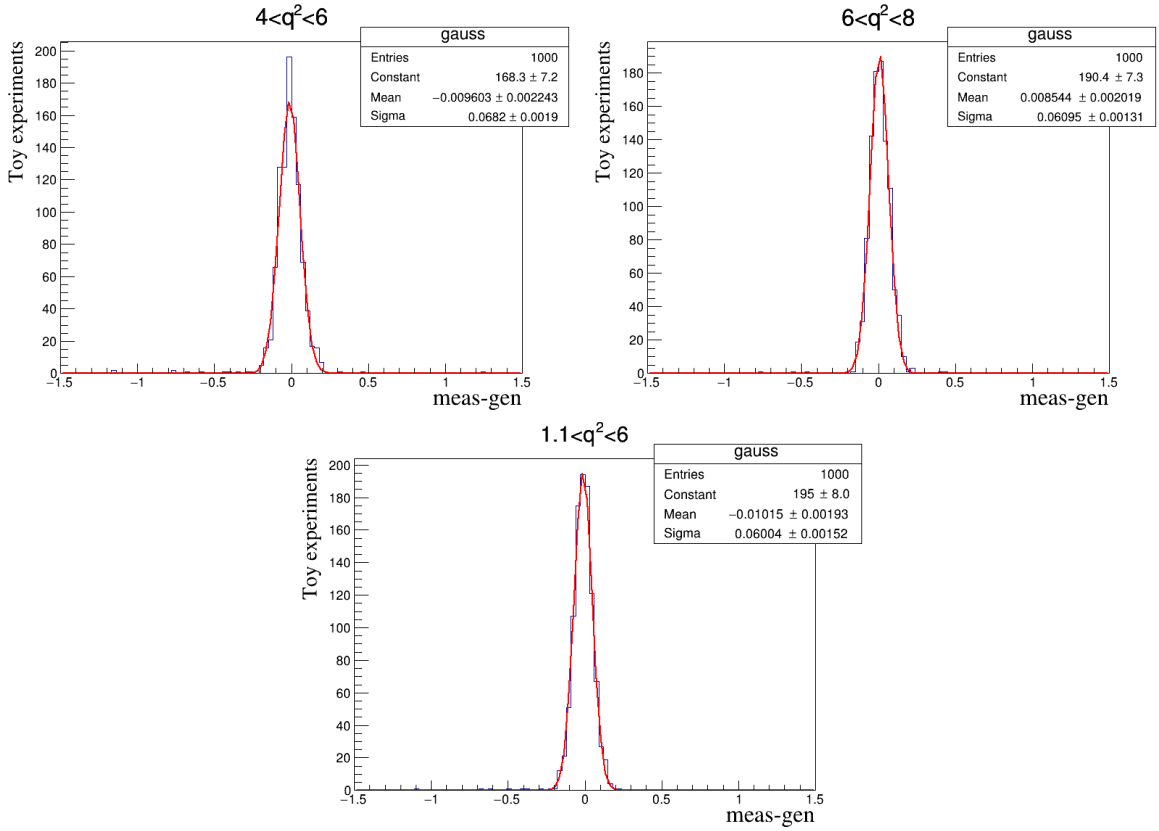


Figure E.2: Distributions of a difference between the measured differential branching fraction with original acceptance and the results based on generated toy datasets with new acceptance.

E.2 Toy studies for the angular analysis

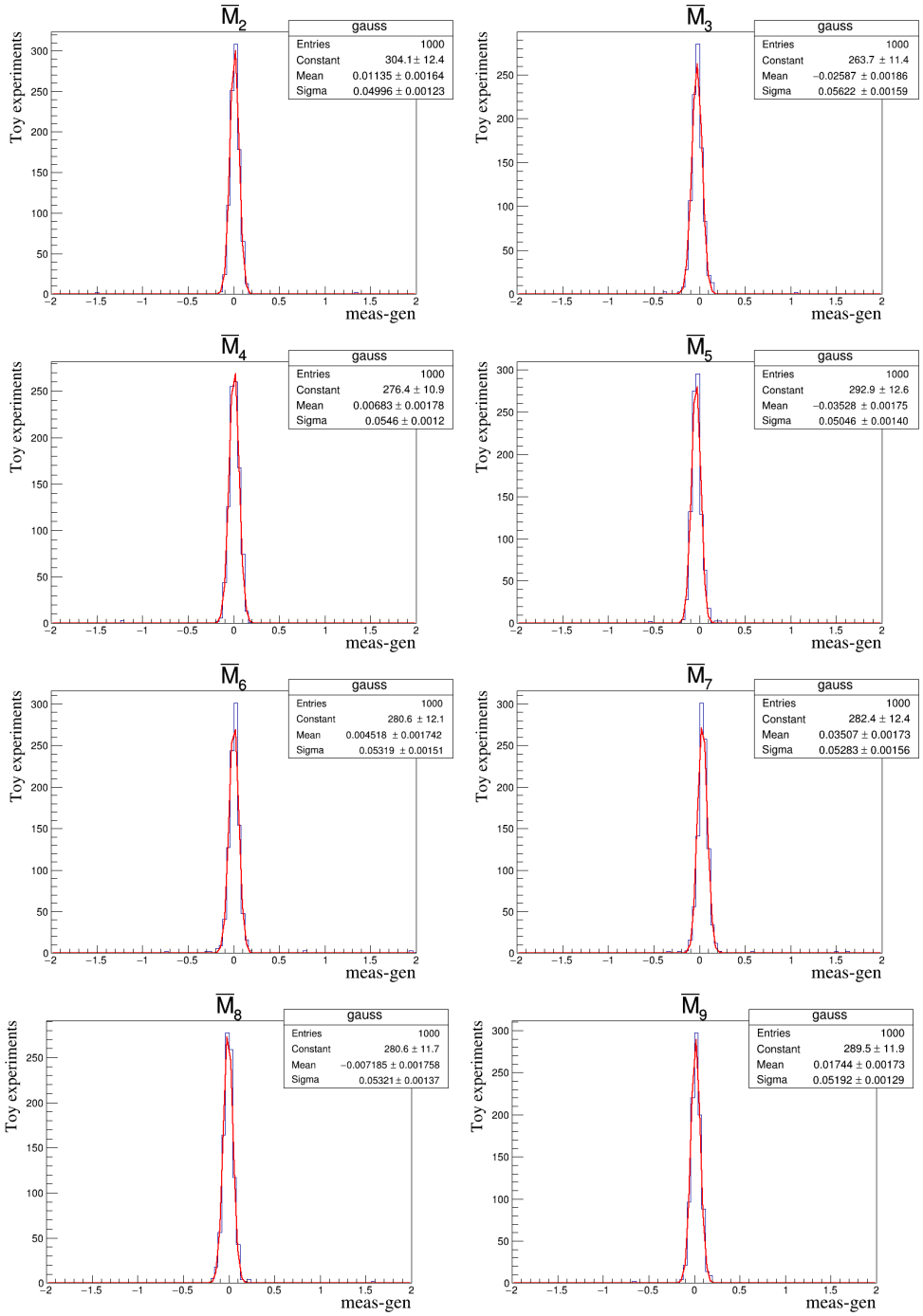


Figure E.3: Distributions of a difference between the measured angular moments with original acceptance and the results based on generated toy datasets with new acceptance.

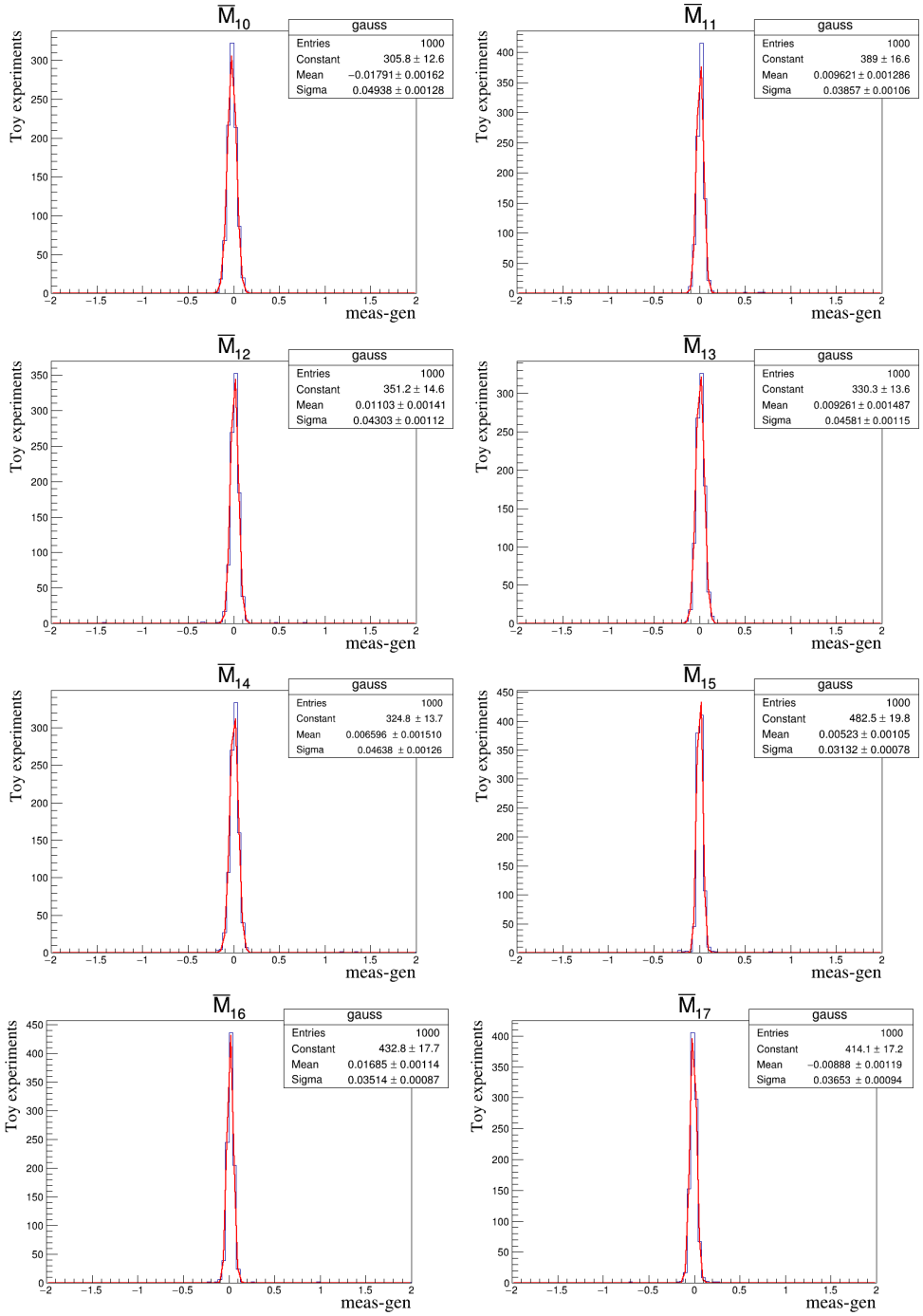


Figure E.4: Distributions of a difference between the measured angular moments with original acceptance and the results based on generated toy datasets with new acceptance.

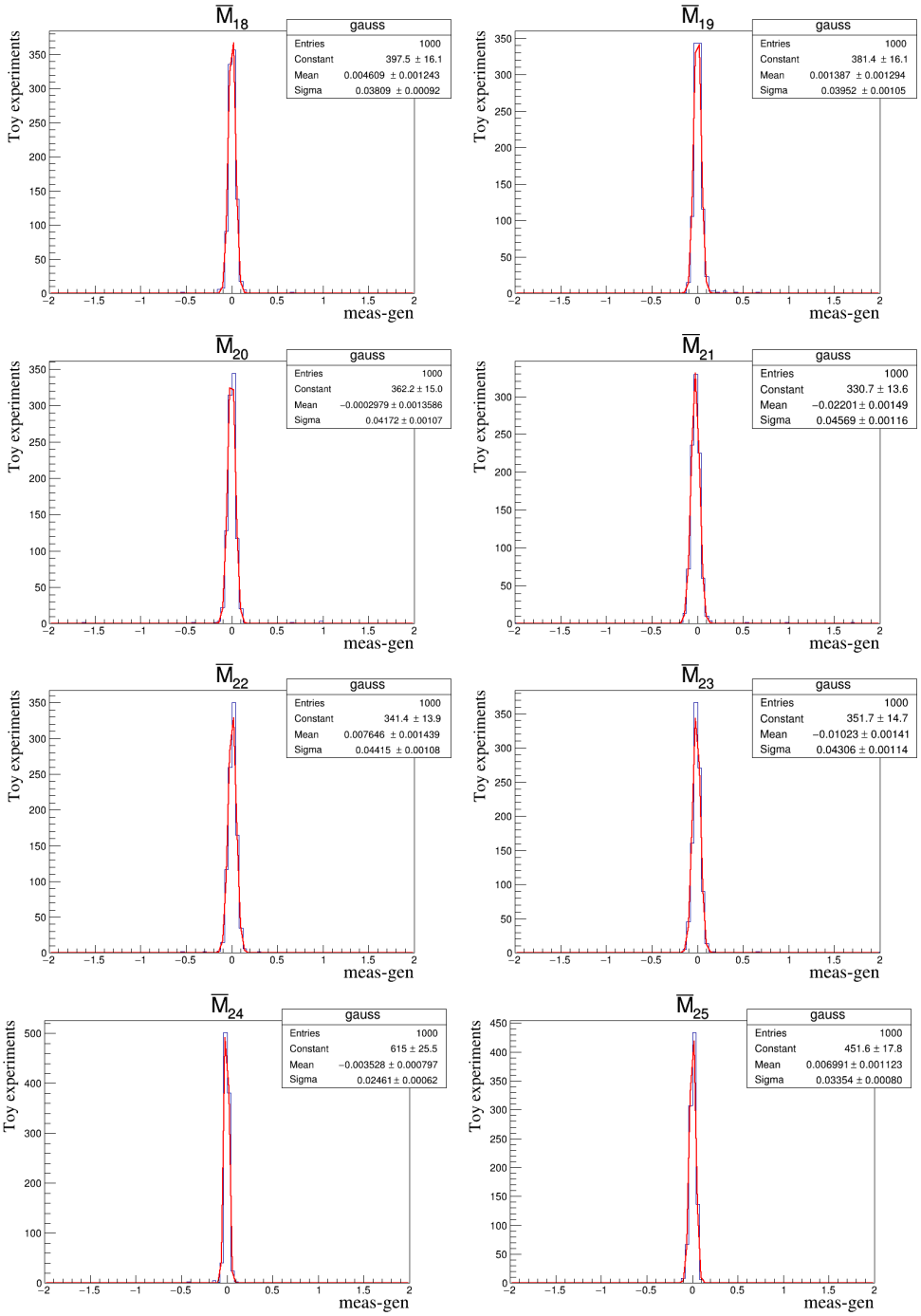


Figure E.5: Distributions of a difference between the measured angular moments with original acceptance and the results based on generated toy datasets with new acceptance.

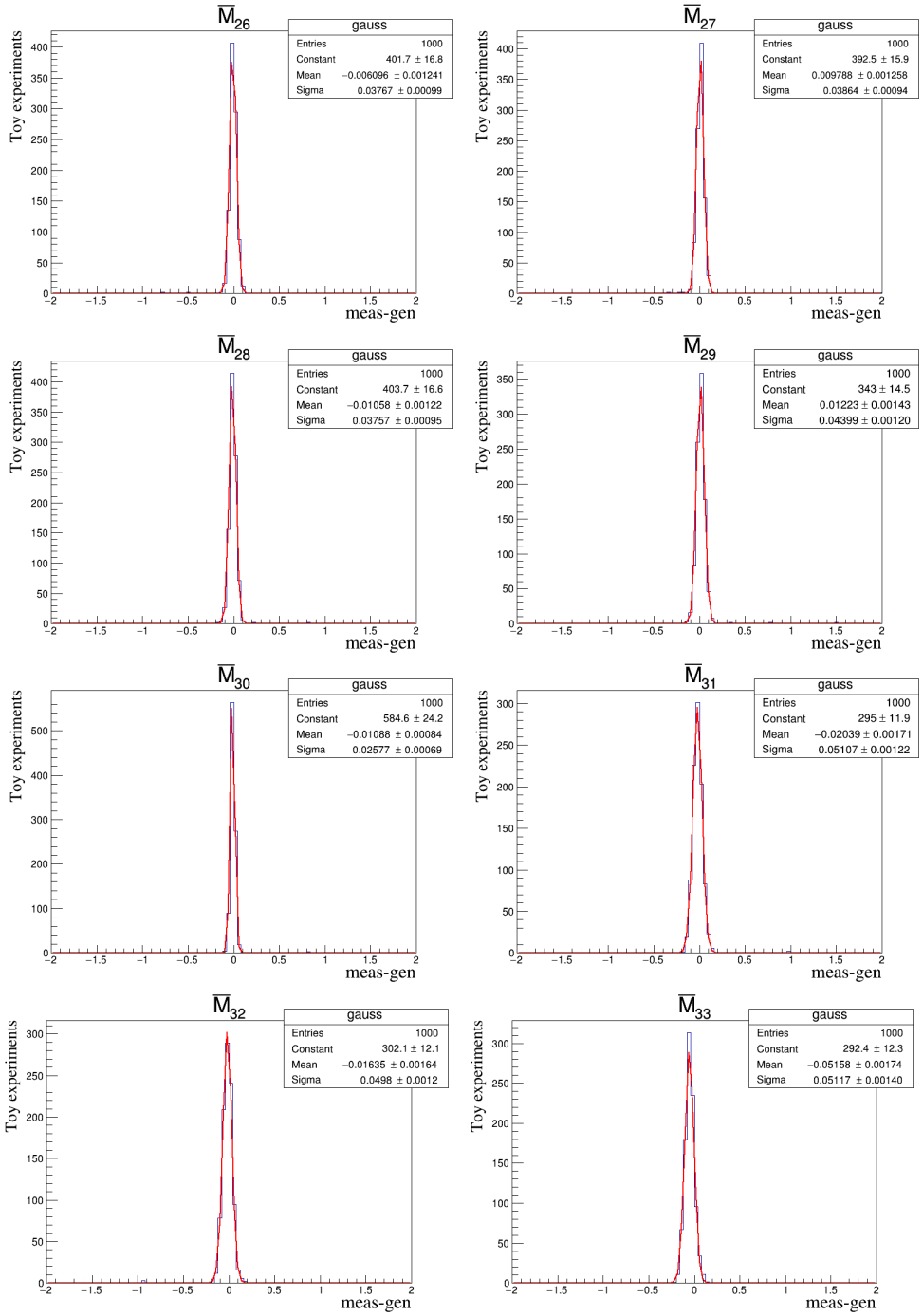


Figure E.6: Distributions of a difference between the measured angular moments with original acceptance and the results based on generated toy datasets with new acceptance.

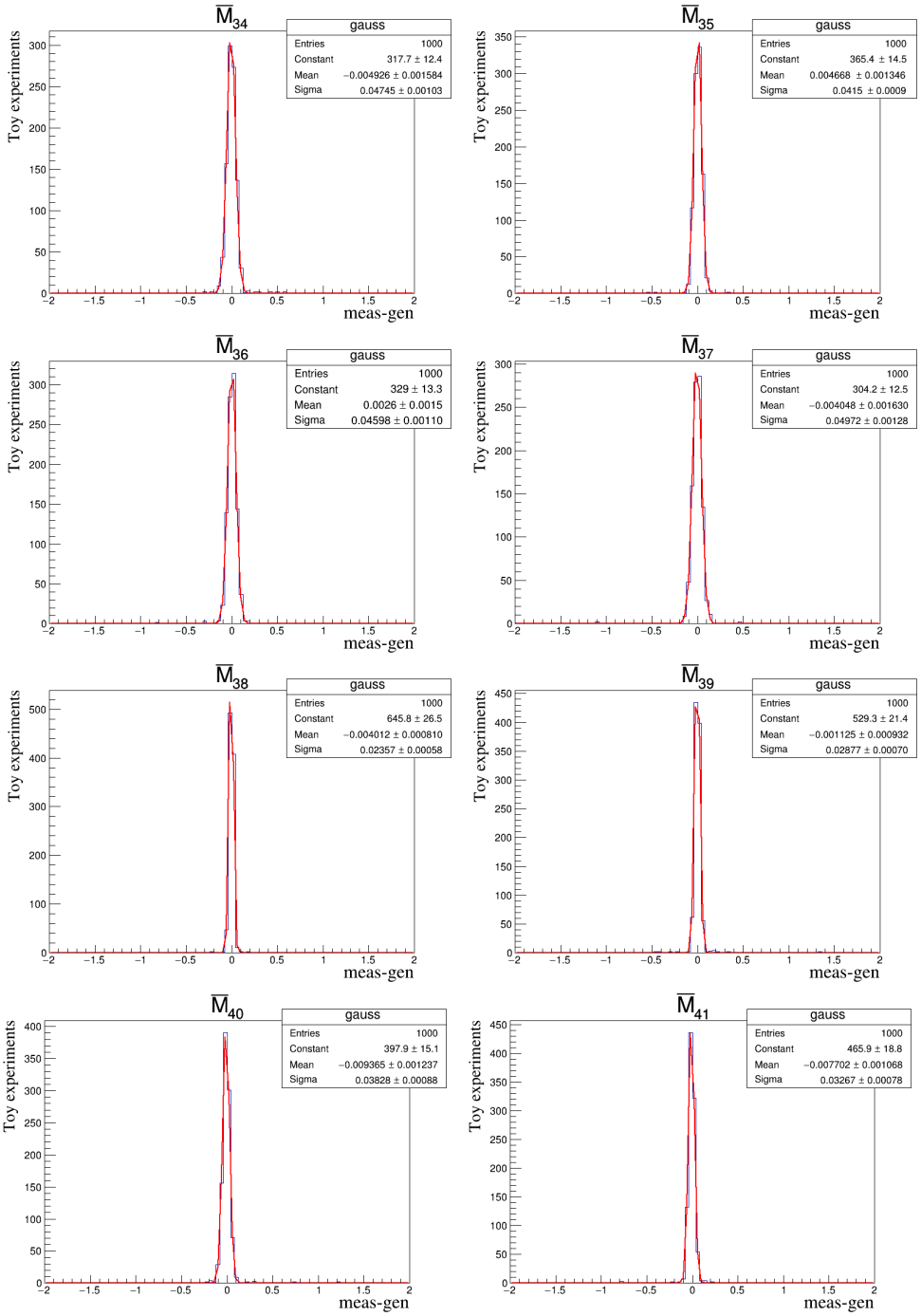


Figure E.7: Distributions of a difference between the measured angular moments with original acceptance and the results based on generated toy datasets with new acceptance.

E.3 Alternative mass model

Figure E.3 shows the results of the comparison between two mass models: Crystal Ball function and Gaussian function, as it is described in Section 10.6.

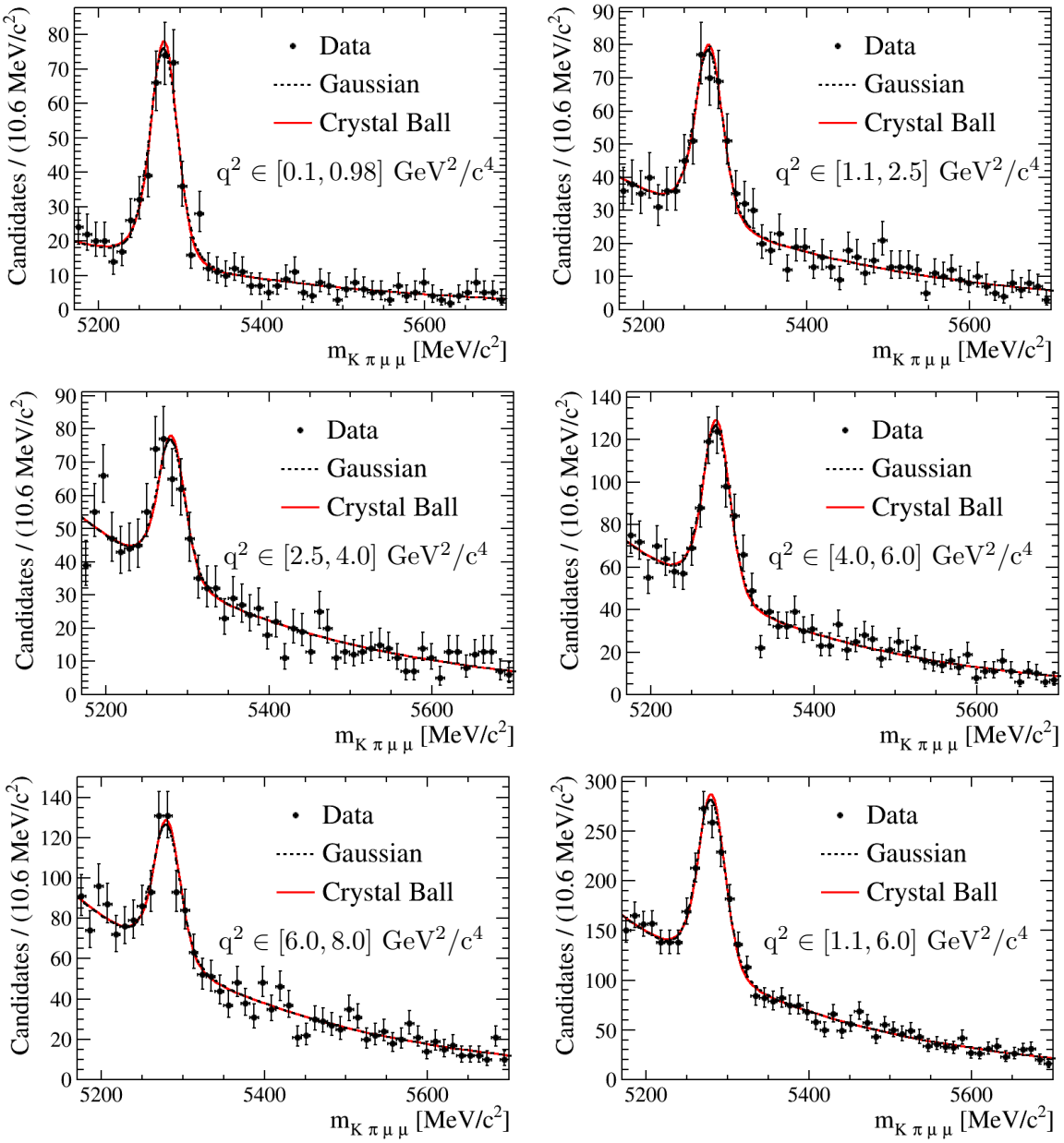


Figure E.8: Comparison between two mass models in the fit to the K$\pi\mu\mu$ invariant mass of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates. The black, dotted line corresponds to the double Gaussian function and the red solid line depicts the double Crystal Ball function.

Bibliography

[1] Robert Mann. An Introduction to Particle Physics and the Standard Model, CRC Press, ISBN 9781420082982 , 2009.

[2] Y. Sirois R. Nicolaidou. The Higgs boson discovery and measurements. *Comptes Rendus Physique*, 16:379–393, 2015.

[3] Y. Ashie et al. Evidence for an oscillatory signature in atmospheric neutrino oscillations. *Phys. Rev. Lett.*, 93:101801, Sep 2004.

[4] J. Matias, F. Mescia, M. Ramon et al. Complete anatomy of $\bar{B}_d \to K^{*0}(\to K\pi)\ell^+\ell^-$ and its angular distribution. *JHEP*, 04:104, 2012.

[5] LHCb Collaboration. Observation of structure in the J/ψ-pair mass spectrum. *Science Bulletin*, 65(23):1983–1993, 2020.

[6] LHCb Collaboration. Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$. *Phys. Rev. Lett.*, 122(22):222001, 2019.

[7] Toshihide Kobayashi, Makoto Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[8] Lincoln Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.

[9] Steven Weinberg. Conceptual Foundations of the Unified Theory of Weak and Electromagnetic Interactions. *Rev. Mod. Phys.*, 52:515–523, 1980.

[10] J. Maiani L. Glashow, S. L. Iliopoulos. Weak interactions with lepton-hadron symmetry. *Phys. Rev. D*, 2:1285–1292, Oct 1970.

[11] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Physical Review Letters*, 81, 08 1998.

[12] Y. S. Amhis et al. (HFLAV). Preliminary average of R(D) and R(D*) for Moriond 2024.

[13] LHCb Collaboration. Measurement of CP-Averaged Observables in the $B^0 \to K^{*0}\mu^+\mu^-$ Decay. *Phys. Rev. Lett.*, 125:011802, Jul 2020.

[14] K. Schoenwald G. Hiller, D. Loose. Leptoquark flavor patterns and B decay anomalies. *Journal of High Energy Physics*, 2016(12), Dec 2016.

[15] L. Matias J. Nierste U. Pokorski S. Rosiek J. Crivellin, A. Hofer. Lepton-flavor violating b decays in generic $Z^{'}$ models. *Phys. Rev. D*, 92:054013, Sep 2015.

[16] Biplab Dey. Angular analyses of exclusive $\overline{B} \to X\ell_1\ell_2$ with complex helicity amplitudes. *Phys. Rev. D*, 92:033013, Aug 2015.

[17] Frederick J. Gilman and Robert L. Singleton. Analysis of semileptonic decays of mesons containing heavy quarks. *Phys. Rev. D*, 41:142–150, Jan 1990.

[18] U. Egede and J. Ramon M. Reece W. Hurth, T. Matias. New physics reach of the decay mode $\bar{B} \to \bar{K}^{*0}\ell^+\ell^-$. *JHEP*, 10:056, 2010.

[19] Roel Aaij et al. Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity. *JHEP*, 02:104, 2016.

[20] Matias J. Ramon M. et al. Descotes-Genon, S. Implications from clean observables for the binned analysis of $B \to K^{*0}\mu^+\mu^-$ at large recoil. *JHEP*, 01:48, 2013.

[21] Serra Nicola Beaujean Frederik, Chrzaszcz Marcin and van Dyk Danny. Extracting angular observables without a likelihood and applications to rare decays. *Phys. Rev. D*, 91:114012, Jun 2015.

[22] LHCb Collaboration. Differential branching fraction and angular moments analysis of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region. *Journal of High Energy Physics*, 2016(12), Dec 2016.

[23] Cai-Dian Lu and Wei Wang. Analysis of $B \to K^*_J(\to K\pi)\mu^+\mu^-$ in the higher kaon resonance region. *Phys. Rev. D*, 85:034014, 2012.

[24] https://cds.cern.ch/record/2800984/.

[25] LHCb Collaboration. LHCb detector performance. *International Journal of Modern Physics A*, 30(07):1530022, 2015.

[26] LHCb Collaboration. LHCb VELO TDR: Vertex locator. Technical design report, CERN-LHCC-2001-011. 5 2001.

[27] LHCb Collaboration. LHCb: Inner tracker technical design report, CERN-LHCC-2002-029. 11 2002.

[28] LHCb Collaboration. LHCb: Outer tracker technical design report, CERN-LHCC-2001-024. 9 2001.

[29] LHCb Collaboration. LHCb: RICH technical design report, CERN-LHCC-2000-037. 9 2000.

[30] LHCb Collaboration. LHCb calorimeters: Technical design report, CERN-LHCC-2000-036. 9 2000.

[31] LHCb Collaboration. LHCb muon system technical design report, CERN-LHCC-2001-010. 5 2001.

[32] LHCb Collaboration. Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC. *Journal of Instrumentation*, 14(04):P04013, apr 2019.

[33] P. Skands T. Sjöstrand, S. Mrenna. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, 2008.

[34] I. Belyaev et. al. Handling of the generation of primary events in Gauss, the LHCb simulation framework. *Journal of Physics: Conference Series*, 331(3):032047, Dec 2011.

[35] David J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1):152–155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[36] P. Golonka and Z. Was. Photos monte carlo: a precision tool for qed correctionsin z and w decays. *European Physical Journal C*, 45:97–107, 01 2006.

[37] J. Allison et al. Geant4 collaboration. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, 2006.

[38] https://root.cern/manual/.

[39] Muriel Pivk and Francois R. Le Diberder. SPlot: A Statistical tool to unfold data distributions, 2005.

[40] L. Breiman et al. Classification and regression trees, Wadsworth international group, https://doi.org/10.1201/9781315139470. 1984.

[41] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting, 1997.

[42] P. A. Zyla et al. (Particle Data Group). Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2020(8):083C01, 08 2020.

[43] L. Anderlini et al. The PIDCalib package, LHCb-PUB-2016-021; CERN-LHCb-PUB-2016-021.

[44] K. Chilikin et al. Belle collaboration. Observation of a new charged charmoniumlike state in $\overline{B}^0 \to J/\psi K^- \pi^+$ decays. *Phys. Rev. D*, 90:112009, Dec 2014.