The Henryk Niewodniczański
Institute of Nuclear Physics
Polish Academy of Sciences
ul. Radzikowskiego 152, 31-342 Kraków

http://www.ifj.edu.pl/publ/hab/

Kraków, August 2017

Searches for New Physics effects in *b! s`−`+* transitions

Marcin Chrząszcz

Habilitation disseration

Wydano nakładem Instytutu Fizyki Jadrowej im. Henryka Niewodniczanskiego
Polskiej Akademii Nauk
Kraków 2017

Recenzent: prof. dr hab. Mariusz Witek
ISBN 978-83-63542-84-9

Contents

Acknowledgements

Firstly, I would like to express my sincere gratitude to my friend and mentor, Prof. dr hab Tadeusz Lesiak, for the continuous support of my research, for his patience, motivation, and immense knowledge. His guidance has helped me during the whole course of my research and writing of this thesis. I could not have imagined having a better advisor, mentor and friend.

Secondly, I would like to thank my colleagues who devoted their time to reading and commenting on this monograph. My special thanks go to prof. dr Andrzej Buras, dr hab. Pawel Bruckman, prof. dr hab. Mikolaj Misiak, dr Danny van Dyk and dr Jarek Wiechczynski. Without them the monograph would not be in the shape it is now!

Last but not least, I would like to thank my parents and my fiancee for supporting me spiritually throughout writing this monograph and my life in general.

Chapter 1

Introduction

The dissertation aims at presenting the current situation in the measurements of electroweak penguin diagrams dominated decays: $b \to s\ell^-\ell^{+1}$. These decays have been a smoking gun for hunting for New Physics effects over many years, but in the last three years the research on these phenomena has intensified due to new measurements. Enormous progress has been made both on the theoretical and the experimental sides to understand the measured deviations from the current Standard Model predictions, referred to in what follows as "anomalies".

The author of this dissertation has been one of the main authors of the angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ decay in the LHCb experiment, which has been widely regarded as one of the most important results of the flavour physics sector in recent years. He has proposed a method called "the method of moments" to measure the angular terms of this decay, which he has later successfully applied in the measurement itself. Moreover, he has been the driving force behind the two other important analyses in LHCb: the measurement of the angular distribution and branching ratio of the $B^0 \to K^{*0}(1430)\mu^+\mu^-$ decay, where again the method of moments has been used to obtain the angular coefficients, and the search for the light scalar particle that can be produced in the $b \to s$ transitions and that decays to a dimuon pair. In this case no signal has been observed and the upper limits on the branching fraction have been set, later to be used for constraining the inflaton model.

The dissertation is organized as follows: the brief introduction is followed by, the second chapter devoted to a theoretical description of rare B decays, where the effective field theory formalism is introduced. Furthermore, the author discusses the current theoretical problems in calculating the Standard Model predictions for the $b \to s\ell^-\ell^+$ processes. Last but not least, the optimised angular observables that are less dependent on the form factors uncertainness are derived. The third chapter describes the experimental apparatus used in the $b \to s\ell^-\ell^+$ measurements. Special focus is put on the sub-detectors that play an important role in the studies of $b \to s\ell^-\ell^+$ transitions. Chapters 4, 5, 6 are devoted to describing the data analyses performed by the author in the LHCb experiment. In Chapter 7 the global analysis of electroweak penguin decays is presented. This kind of global analysis has become extremely popular in the past few years as it helps to constrain and pin down

[1] Charge conjugation is implied unless stated differently.

those New Physics models that are likely to be responsible for the observed anomalies. The author of this monograph is involved in one of the biggest collaborations performing New Physics fits, where he is the convenor of the Flavour Working group. Furthermore, the author presents his own study on separating the long distance effects in the $B^0 \to K^{*0}\mu^+\mu^-$ decay. This is the state of the art way of determining those contributions. The chapter ends with a description of possible New Physics models that can explain the observed discrepancies.

Chapter 2

Theoretical description of $b \to s\ell^-\ell^+$ decays

This chapter describes various theoretical aspects relevant to the $B^0 \to K^{*0}\mu^+\mu^-$ decays. Particular attention is paid to the Effective Field Theory (EFT) formalism used to describe the physics of $b \to s\ell^-\ell^+$ transitions.

2.1 Introduction to Particle Physics

The current theory describing the interactions between the fundamental particles is the so-called Standard Model (SM). This model emerged in the 1960-1970s and passed with flying colours the measurements performed by the LEP, Babar and Belle experiments. In 2012 the ATLAS and CMS experiments located at the LHC confirmed the discovery of the last missing piece of the SM: the Higgs boson [1, 2]. Currently the main task in the experimental Particle Physics is to search for phenomena that are not described by the SM. This dissertation describes such a search with the use of electroweak penguin $b \to s\ell^-\ell^+$ transitions.

2.2 Modern challenges of Particle Physics

Despite its tremendous successes, the SM cannot be the ultimate theory of Particle Physics. For instance, the SM does not include gravity, which is expected to become relevant to Particle Physics at the Planck scale (10^{19} GeV)[1]. The fact that the Planck scale is much higher than the Fermi scale (~ 100 GeV), which is a scale relevant to the SM, is known as the hierarchy problem [3]. If the particles at the Planck scale interacted with the Higgs boson, they would increase its mass by the radiative corrections by orders of magnitude, unless a fine-tuning mechanism was involved [4].

Most importantly, the SM disagrees with some experimental observations:

[1] There exist alternative theories of gravity, in particular those involving extra-dimensions where the scale of quantum gravity is significantly lower than the Planck scale.

1. Matter-antimatter asymmetry: our Universe is almost exclusively composed of matter. In the widely accepted models of inflation the standard assumption is that in the beginning the Universe was built equally from matter and anti-matter. If this assumption is correct, the SM does not offer a mechanism that can quantitatively explain this observation.

2. Dark matter and dark energy: there is striking experimental evidence that most of the matter in our Universe consists of particles not present in the SM, which do not interact electromagnetically or strongly.

3. Neutrino masses: in the SM neutrinos are massless particles; however, neutrino oscillations prove that they have a small, but non-vanishing mass[2]. It is worth pointing out that this situation can be easily accommodated with a minimal modification of the SM, by adding for example dimension five operators. After this procedure the new effective theory would no longer be the "original" SM, strictly speaking; however, the theory community still refers to this theory as the SM.

In the past few decades many theories that are trying to complete the SM, often called New Physics (NP) or Beyond the SM, have been proposed. Most of these theories predict the existence of new heavy particles that would solve the experimental and theoretical problems of the SM. Searching for microscopic evidence of NP has been the quest for the holy grail of modern Particle Physics experiments.

New particles can be produced directly if their mass is lower than the total energy in the centre-of-mass of the particle collider (this is mainly done at the LHC by the ATLAS and CMS experiments), or they can be discovered indirectly via their interaction with the SM particles (mainly done at the LHC by the LHCb experiment). In the latter case, it is not necessary to have enough energy in the centre-of-mass to produce these new particles, since they can interact as virtual particles[3], but a large number of decays have to be available to be able to perform a meaningful search for NP.

2.3 Flavour sector of the Standard Model

The Lagrangian of the SM can be divided into three pieces: a gauge sector, a fermion sector and a Higgs sector, the latter being responsible for breaking the gauge symmetry. The gauge sector is defined by a local symmetry group: $\mathcal{G}^{\rm SM}_{\rm local} = SU(3)_C \times SU(2)_L \times U(1)_Y$ and the fermion current:

$$\mathcal{L}^{\rm SM}_{\rm gauge} = \sum_{i=1\ldots3} \sum_{\psi=Q^i_L\ldots,E^i_R} \bar{\psi} i \not{D} \psi - \frac{1}{4} \sum_{a=1\ldots8} G^a_{\mu\nu} G^a_{\mu\nu} - \frac{1}{4} \sum_{a=1\ldots3} W^a_{\mu\nu} W^a_{\mu\nu} - \frac{1}{4} B_{\mu\nu} B_{\mu\nu}, \tag{2.1}$$

[2]This discovery was awarded Nobel Prize in 2015 [5].

[3]In this case, the Heisenberg uncertainty allows these particles to be produced.

where ψ is the fermionic field, D_μ is the gauge derivative, while G, W B are the gluon and electroweak gauge boson fields. The fermion content of the SM consists of five fields:

$$Q_L^i(3,2)_{+1/6}, \quad U_R^i(3,1)_{+2/3}, \quad D_R^i(3,1)_{-1/3}, \quad L_R^i(1,2)_{-1/2}, \quad E_R^i(1,1)_{-1}, \tag{2.2}$$

each appearing with the flavour replicas ($i = 1, 2, 3$). The local electroweak symmetries of $\mathcal{L}_{\rm gauge}^{\rm SM}$ are spontaneously broken by the vacuum expectation of the Higgs field, which is a $SU(2)_L$ scalar doublet:

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \qquad \langle H \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}, \tag{2.3}$$

where ν is determined by the W boson mass:

$$m_W^2 = \frac{g^2 \nu^2}{4}, \qquad \nu = \sqrt{\sqrt{2} G_f} \approx 246 \ \text{GeV}. \tag{2.4}$$

The global flavour symmetry of $\mathcal{L}_{\rm gauge}^{\rm SM}$ corresponds to unitary rotations in the flavour space of the five fermions fields in Eq. 2.2 is $U(3)^5$. This group can be decomposed:

$$\mathcal{G}_{\rm flavour} = U(3)^5 = U(1)^5 \times \mathcal{G}_q \times \mathcal{G}_\ell, \tag{2.5}$$

where

$$\mathcal{G}_q = SU(3)_{Q_L} \times SU(3)_{Q_R} \times SU(3)_{D_R}, \qquad \mathcal{G}_\ell = SU(3)_{L_L} \otimes SU(3)_{E_R}. \tag{2.6}$$

In the above two out of five $U(1)$ subgroups can be identified with the total baryon and lepton numbers which are not broken by the Yukawa sector. The third one can be associated with the hypercharge, which is broken spontaneously if $\langle H \rangle \neq 0$.

The diagonalization of each Yukawa coupling requires two independent unitary matrices:

$$V_{L,q} Y_q V_{R,q}^\dagger = \text{diag}(y_1, y_2, y_3),$$

where $q = u, d$.

The lepton sector of the SM is invariant under $\mathcal{G}_\ell$, which allows us to freely choose the two matrices that are necessary to diagonalize Y without any observable consequences. On the other hand, the quark sector does not possess this nice property. There, we can freely choose only three out of four matrices necessary to diagonalize both Y_d and Y_u. We can choose a basis in which:

$$Y_d = \lambda_d, \qquad Y_u = V^\dagger \lambda_u, \tag{2.7}$$

where

$$\lambda_d = \text{diag}(y_d, y_s, y_b), \qquad \lambda_u = \text{diag}(y_u, y_c, y_t), \qquad y_q = \frac{m_q}{\nu}. \tag{2.8}$$

In the above, V is nothing else but the CKM matrix of the SM [6].

2.4 CKM matrix

The CKM matrix can be seen as a transition matrix between the flavour eigenstates and the mass eigenstates. It can be parametrised with three rotation angles (θ_{ij}) and a complex phase (δ):

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = R(s_{12}) \times R(s_{13}; e^{i\delta}) \times R(s_{23}) \tag{2.9}$$
$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & c_{23}c_{13} \\ s_{12}s_{23} - c_{12}s_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix},$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, $R(\theta_{ij})$ is a rotation matrix for the $i-j$ family and the $R(s_{13}; e^{i\delta})$ is given by:

$$R(s_{13}; e^{i\delta}) = \begin{pmatrix} c_{13} & s_{13} & e^{i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix}. \tag{2.10}$$

The CKM matrix shows a strong hierarchical pattern. $|V_{us}|$ and $|V_{cd}|$ are close to 0.22. The elements $|V_{cb}|$ and $|V_{ts}|$ are of the order of 4×10^{-2}, whereas $|V_{ub}|$ and $|V_{td}|$ are of the order of 5×10^{-3}. The Wolfenstein parametrisation is the expression of the CKM matrix in terms of power expansion of the $\lambda = 0.22$ parameter. This parametrisation is extremely useful as it exhibits the hierarchical structure:

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{2.11}$$

where $A,\ \rho,\ \eta$ are free parameters. As can be seen, the Wolfenstein parametrisation is more transparent than the standard one. However, if one requires a sufficient level of accuracy, the terms $\mathcal{O}(\lambda^4)$ or higher would need to be taken into account. Therefore, in the numerical calculations the parametrisation of Eq. 2.10 is presently preferred.

2.5 Effective Lagrangian formalism in heavy flavour decays

Flavour physics that focuses on the decays of B, D and K mesons has built-in two energy scales within the SM. The first one is the electroweak scale that is characterized by the mass of the W boson. The second one is the strong interaction scale (the so-called Λ_{QCD} related to the hadron formation). The presence of these two scales makes calculations of the amplitudes very difficult. For instance, large logarithms of the type of $\log(m_W/\Lambda_{QCD})$ can easily appear, causing the poor convergence of the perturbation theory.

The idea to simplify the problem is to integrate out the heavy SM fields (W, Z bosons and t quark) at the electroweak scale and construct an appropriate low-energy effective field theory (EFT). In this kind of theory, only the light degrees of freedom remain.

An example of such a calculation can be shown for the above-mentioned charge current interactions. The part of the Lagrangian describing interactions is given by the formula:

$$\mathcal{L}_{cc} = \frac{g}{\sqrt{2}} J^{\mu}_{W}(x) W^{+}_{\mu}(x) + \text{h.c.}, \tag{2.12}$$

where

$$J^{\mu}_{W}(x) = V_{ij}\bar{u}^{i}_{L}(x)\gamma^{\mu}d^{j}_{L}(x) + \bar{e}^{j}_{L}(x)\gamma^{\mu}\nu^{j}_{L}(x), \tag{2.13}$$

is the weak charged current. After integrating out the W field one gets an amplitude:

$$i\mathcal{T} = i\frac{g^2}{2}\int d^4x D_{\mu\nu}(x, m_W) T\left[J^{\mu}_{W}(x), J^{\nu\dagger}_{W}(0)\right], \tag{2.14}$$

where only light fields are present. $D_{\mu\nu}(x, m_W)$ is the W propagator which can be expanded in powers of $1/m^2_W$:

$$D_{\mu\nu}(x, m_W) = \delta(x)\frac{ig_{\mu\nu}}{m^2_W} + \dots \ . \tag{2.15}$$

If we insert Eq. 2.15 into Eq. 2.12, we get at leading order:

$$\mathcal{L}^{\text{eff}}_{cc} = -\frac{4G_F}{\sqrt{2}} g_{\mu\nu} J^{\mu}_{W}(x) J^{\nu\dagger}_{W}(0). \tag{2.16}$$

The following example shows how at leading order one gets the well known Fermi theory of weak interactions in the EFT. The visualisation of such an approach is given in Fig. 2.1, where the W propagator is replaced by an effective vertex.

If one would like to describe a semileptonic decay, one would have to use a product of two currents: one quark and one leptonic:

$$\mathcal{L}^{\text{semileptonic}}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{ij}\left[\overline{u}^{i}_{L}(x)\gamma^{\mu}d^{j}_{L}(x)\right] \times \left[\overline{\nu}_{L}(x)\gamma_{\mu}e_{L}(x)\right] + \text{h.c.} \tag{2.17}$$

The above Lagrangian gives an excellent description of semileptonic decays of B or D mesons. It is worth mentioning that for the semileptonic processes, such as for example $B \to D\mu\nu$, the operator is not renormalized by the strong interactions. Thanks to this, the CKM elements like V_{cb} can be determined with high theoretical accuracy.

Unfortunately, not always the situation is as simple as in the case of semileptonic decays. A good example is the four quark interactions and flavour changing neutral currents processes, where the QCD corrections cannot be neglected. However, the general recipe for the treatment of the EFT approach remains the same:

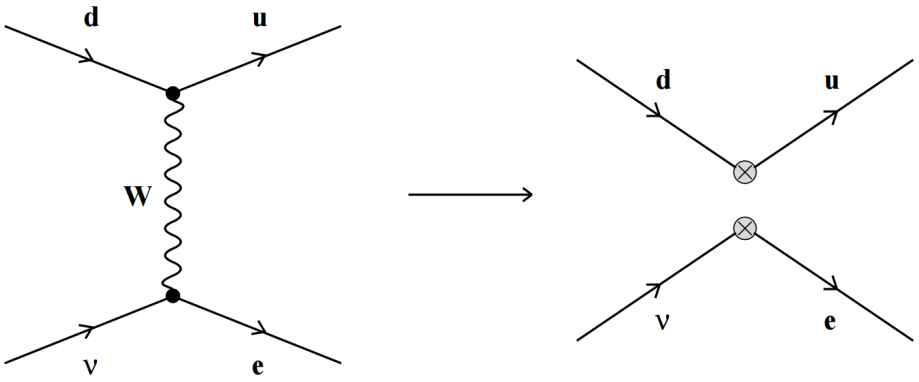


Figure 2.1: Ilustration of the EFT approach in a weak decay. The W propagator is replaced by an effective vertex [7].

- Identify the basis of local operators that are compatible with the symmetries of the system. In practice, one considers the dimension-six operators which correspond to the leading-order expansion in $1/m_W$.
- Write down the effective Lagrangian[4]:

$$\mathcal{L} = -4\frac{G_F}{\sqrt{2}} V_{CKM} \sum_i C_i(\mu) O_i(\mu), \tag{2.18}$$

  where C_i are the so-called Willson Coefficients (WC), which can be interpreted as effective couplings, O_i are the aforementioned effective operators and the sum runs over all identified operators. In Eq. 2.18, μ is the renormalization scale which acts as a seperator between the short and long distance corrections. One evaluates first C_i at the electroweak scale $\mu \sim m_W$, then using renormalization group equations [8] the WC are run down to the relevant scale (for example $\mu = m_b$).
- Finally we evaluate the matrix elements for the given process ($i \to f$):

$$\mathcal{M}(i \to f) = \langle f|\mathcal{H}_{eff}|i\rangle = 4\frac{G_F}{\sqrt{2}} V_{CKM} \sum_i C_i(\mu)\langle f|O_i|i\rangle, \tag{2.19}$$

In Eq. 2.19 the NP effect may change the values of the WC. Assuming that the NP is heavy, it can modify the initial conditions of the WC, but it will not affect the other two steps. Unfortunately, the hadronic matrix elements are affected by non-perturbative QCD effects, which are the main part of the theoretical uncertainties. We will discuss this uncertainty in Sec. 2.8.2.

[4]Some theorists prefer to work with the effective Hamiltonian instead of the Lagrangian. The Hamiltonian is defined in a very similar way: $\mathcal{H} = 4\frac{G_F}{\sqrt{2}} V_{CKM} \sum_i C_i(\mu) O_i(\mu)$.

2.6 Flavour Changing Neutral Current processes

The basic vertices of neutral currents (c.f Fig. 2.2) in the SM are flavour diagonal, which forbids the existence of Flavour Changing Neutral Currents (FCNC) at the tree level. With the help of two charge currents one can construct a loop diagram that will allow the FCNC process in the SM. The advantage of this process being mediated by a loop diagram[5] is the fact that the NP effects can enter the loop even if the NP scale is heavy. The box diagrams are represented in the EFT approach via the effective vertices that can change the flavour.

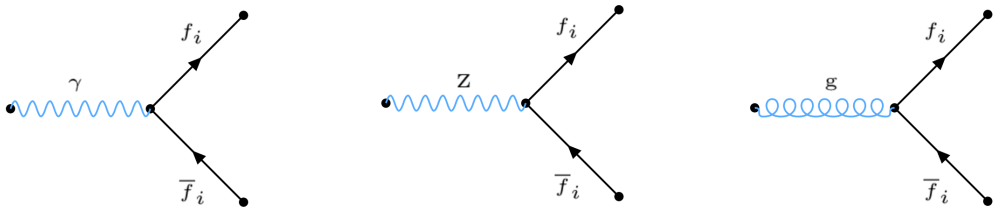


Figure 2.2: Neutral currents flavour conserving vertices in the SM.

In flavour physics there exist many FCNC processes. The prime examples are:

- $K^0_{\rm L}$ - $K^0_{\rm S}$ mixing and CP violation in the $K^0_{\rm L} \to \pi\pi$.
- Mixing of neutral $B_{s/d}$ mesons.
- Decays of type: $b \to s\ell^-\ell^+$ and $b \to s\gamma$

Each of the above-mentioned decays is sensitive to NP contributions and has its own phenomenological applications and theoretical challenges when calculating the SM prediction [9]. The discussion of all FCNC processes is beyond the scope of this monograph, so we will focus only on the processes of type: $b \to s\ell^-\ell^+$. The Lagrangian part responsible for them can be written using four effective couplings (in the SM):

$$\mathcal{L}^{\rm FCNC}_{\rm eff} = \mathcal{L}^{\rm non-lep}_{\rm eff} - 4\frac{G_F}{\sqrt{2}} V_{CKM} \left(C_7 O_7 + C_9 O_9 + C_{10} O_{10} \right),$$

where the operators[6] are:

$$O_7 = \frac{4}{16\pi^2} m_b \bar{b}^\alpha_R \sigma^{\mu\nu} F_{\mu\nu} s^\alpha_L, \qquad O_9 = \frac{1}{2} b^\alpha_L \gamma^\mu s^\alpha_L \bar{\ell} \gamma_\mu \ell, \qquad O_{10} = \frac{1}{2} b^\alpha_L \gamma^\mu s^\alpha_L \bar{\ell} \gamma_\mu \gamma_5 \ell.$$

[5] In the literature also known as box diagrams.

[6] The right-handed currents have been omitted as they are suppressed in the SM by m_s/m_b. They can appear in NP scenarios. In that case, the right-handed operators can be constructed by the $q_{L/R} \to q_{R/L}$ swap.

From the above-mentioned operators, the WC corresponding to O_{10} is the cleanest one. This is because the O_{10} operator does not mix with any four-quark operators and is not renormalized by QCD. The canonical example of an observable that is sensitive to the C_{10} WC is the $B_{s/d} \to \mu^+\mu^-$ decay, which has been measured by the LHCb and CMS collaborations [10, 11] to be consistent with the SM predictions.

2.7 Description of the $B^0 \to K^{*0}\mu^+\mu^-$ decay amplitudes

The kinematics of the decay $B^0 \to K^{*0}\mu^+\mu^-$ can be described by four independent variables: the invariant mass squared of the two muons (q^2), the angle between the directions of flight of the kaon and the B meson in the rest frame of the K^{*0} ($\cos\theta_k$), the angle between the directions of flight of the μ^- and the B meson in the dilepton rest frame ($\cos\theta_\ell$) and the azimuthal angle ϕ between the two planes defined by the $K\pi$ system and the lepton pair. The four-differential decay rate for the $K\pi$ system in the P-wave configuration, ignoring scalar[7] contributions, is given by [12]:

$$\begin{aligned}\frac{\mathrm{d}^4\Gamma[B^0 \to K^{*0}\mu^+\mu^-]}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} =& \frac{9}{32\pi}\sum_i J_i(q^2) f_i(\cos\theta_\ell, \cos\theta_K, \phi) \\ =& \frac{9}{32\pi}\Big[\, J_1^s \sin^2\theta_K + J_1^c \cos^2\theta_K + \\ & J_2^s \sin^2\theta_K \cos 2\theta_\ell + J_2^c \cos^2\theta_K \cos 2\theta_\ell + \\ & J_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi + J_4 \sin 2\theta_K \sin 2\theta_\ell \cos\phi + \\ & J_5 \sin 2\theta_K \sin\theta_\ell \cos\phi + J_6^s \sin^2\theta_K \cos\theta_\ell \; + \\ & J_6^c \cos^2\theta_K \cos\theta_\ell + J_7 \sin 2\theta_K \sin\theta_\ell \sin\phi + \\ & + J_8 \sin 2\theta_K \sin 2\theta_\ell \sin\phi + J_9 \sin^2\theta_K \sin^2\theta_\ell \sin 2\phi \,\Big]. \end{aligned} \tag{2.20}$$

As can be seen, the angular part is described by spherical harmonic functions. The twelve coefficients J_i can be calculated using six complex transversity amplitudes: $A_\perp^{L,R}$, $A_\parallel^{L,R}$, $A_0^{L,R}$, which are functions of q^2. In addition, for the massive lepton case, the A_t complex amplitude has to be included. In the non SM case the scalar amplitude A_s would have to be taken into account as well. Expressing the above-mentioned observables

[7] This refers to a scalar configuration of the dimuon system, not to be confused with an S-wave contribution to the $K\pi$ system as will be discussed later

in terms of these amplitudes yields:

$$
\begin{aligned}
J_{1s} &= \frac{(2+\beta_\ell^2)}{4}\left[|A_\perp^L|^2+|A_\parallel^L|^2+|A_\perp^R|^2+|A_\parallel^R|^2\right]+\frac{4m_\ell^2}{q^2}\mathrm{Re}\left(A_\perp^L A_\perp^{R*}+A_\parallel^L A_\parallel^{R*}\right),\\
J_{1c} &= |A_0^L|^2+|A_0^R|^2+\frac{4m_\ell^2}{q^2}\left[|A_t|^2+2\mathrm{Re}(A_0^L A_0^{R*})\right]+\beta_\ell^2\,|A_S|^2,\\
J_{2s} &= \frac{\beta_\ell^2}{4}\left[|A_\perp^L|^2+|A_\parallel^L|^2+|A_\perp^R|^2+|A_\parallel^R|^2\right],\\
J_{2c} &= -\beta_\ell^2\left[|A_0^L|^2+|A_0^R|^2\right],\\
J_3 &= \frac{1}{2}\beta_\ell^2\left[|A_\perp^L|^2-|A_\parallel^L|^2+|A_\perp^R|^2-|A_\parallel^R|^2\right],\\
J_4 &= \frac{1}{\sqrt{2}}\beta_\ell^2\left[\mathrm{Re}(A_0^L A_\parallel^{L*}+A_0^R A_\parallel^{R*})\right],\\
J_5 &= \sqrt{2}\beta_\ell\left[\mathrm{Re}(A_0^L A_\perp^{L*}-A_0^R A_\perp^{R*})-\frac{m_\ell}{\sqrt{q^2}}\mathrm{Re}(A_\parallel^L A_S^*+A_\parallel^{R*}A_S)\right],\\
J_{6s} &= 2\beta_\ell\left[\mathrm{Re}(A_\parallel^L A_\perp^{L*}-A_\parallel^R A_\perp^{R*})\right]\\
J_{6c} &= 4\beta_\ell\frac{m_\ell}{\sqrt{q^2}}\,\mathrm{Re}(A_0^L A_S^*+A_0^{R*}A_S),\\
J_7 &= \sqrt{2}\beta_\ell\left[\mathrm{Im}(A_0^L A_\parallel^{L*}-A_0^R A_\parallel^{R*})+\frac{m_\ell}{\sqrt{q^2}}\,\mathrm{Im}(A_\perp^L A_S^*-A_\perp^{R*}A_S))\right],\\
J_8 &= \frac{1}{\sqrt{2}}\beta_\ell^2\left[\mathrm{Im}(A_0^L A_\perp^{L*}+A_0^R A_\perp^{R*})\right],\\
J_9 &= \beta_\ell^2\left[\mathrm{Im}(A_\parallel^{L*}A_\perp^L+A_\parallel^{R*}A_\perp^R)\right],
\end{aligned}
\tag{2.21}
$$

where the parameter β_ℓ is given by

$$\beta_\ell = 1-\frac{4m_\ell^2}{q^2}. \tag{2.22}$$

It is worth mentioning that:

$$J_{6c}=0, \tag{2.23}$$

for the massless lepton limit or if there is no scalar amplitude (A_s). As it will be shown, LHCb is the first experiment that has measured this observable thanks to the method of moments [13].

Some theorists prefer to use the helicity amplitudes (H_i) instead of the transversity amplitudes (A_j). The relation between them is straightforward:

$$A_{\perp,\parallel} = (H_{+1}\mp H_{-1})/\sqrt{2}, \qquad A_0 = H_0.$$

Not all of the twelve observables (J_i) are independent, as has been shown in [14]. The first relation between them arises when one considers the massless limit of leptons ($\frac{m_\ell^2}{q^2}\to 0$). Under this condition the following relations are satisfied:

$$J_{1s}=3J_{2s}, \quad J_{1c}=-J_{2c}. \tag{2.24}$$

There exist in total four infinitesimal transformations of the amplitudes that do not change the physical observables. In order to introduce them, we will first write the amplitudes in the binomial forms:

$$\begin{aligned} n_{\parallel} &= \begin{pmatrix} A_{\parallel}^{L} \\ A_{\parallel}^{R*} \end{pmatrix}, \\ n_{\perp} &= \begin{pmatrix} A_{\perp}^{L} \\ -A_{\perp}^{R*} \end{pmatrix}, \\ n_{0} &= \begin{pmatrix} A_{0}^{L} \\ A_{0}^{R*} \end{pmatrix}. \end{aligned} \tag{2.25}$$

In this basis the transformation has the form of 4 rotations:

$$n_i^{'} = U n_i = \begin{bmatrix} e^{i\phi_L} & 0 \\ 0 & e^{-i\phi_R} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cosh i\tilde{\theta} & -\sinh i\tilde{\theta} \\ -\sinh i\tilde{\theta} & \cosh i\tilde{\theta} \end{bmatrix}.$$

The four angles: θ, $\tilde{\theta}$, ϕ_L and ϕ_R can be varied independently[8]. The angles θ, $\tilde{\theta}$ correspond to the transformation of a helicity $+1$ final state with a left handed current into a helicity -1 state with a right handed current. The origin of these symmetries is the fact that it is impossible to experimentally determine the helicity and handedness of the current in the same time. Until now we have written down three relations Eq. 2.23 and 2.24. The remaining fourth one is less trivial:

$$\begin{aligned} J_{2c} = &-2\,\frac{(2J_{2s}+J_3)\,(4J_4^2+\beta_\ell^2 J_7^2)+(2J_{2s}-J_3)\,(\beta_\ell^2 J_5^2+4J_8^2)}{16J_{2s}^2-(4J_3^2+\beta_\ell^2 J_{6s}^2+4J_9^2)} \\ &+4\,\frac{\beta_\ell^2 J_{6s}(J_4J_5+J_7J_8)+J_9(\beta_\ell^2 J_5J_7-4J_4J_8)}{16J_{2s}^2-(4J_3^2+\beta_\ell^2 J_{6s}^2+4J_9^2)} \end{aligned} \tag{2.26}$$

This relation holds when there is no scalar amplitude present. Further information and discussion about the inclusion of the scalar amplitudes can be found in [14]. It is worth pointing out that because of the complexity of Eq. 2.26, it is not used in the experimental measurements.

2.8 Effective Field Theory description of $B^0 \to K^{*0}\mu^+\mu^-$

Once we have defined the transversity and helicity amplitudes, we should express them in terms of the Wilson Coefficients to allow for phenomenological analysis of this decay. The transversity amplitudes read [15]:

[8]Please note that because of the fact that the left and right-handed amplitudes do not interfere, both ϕ_L and ϕ_R can be chosen arbitrary.

$$H^V = N\sqrt{q^2}\left(C_9^{\rm eff}(q^2)V^k(q^2) + \frac{2m_b}{q^2}C_7^{\rm eff}(q^2)T^k(q^2) + i\sqrt{\lambda}\frac{2m_b}{q^2}\Delta_k(q^2)\right) \tag{2.27}$$

$$H^A = N\sqrt{q^2}C_{10}V^k(q^2), \tag{2.28}$$

where $k = +, -, 0$ denotes the polarisation of the K^* meson. The helicity form factors T^k, V^k are defined as in [15] and Δ_k stands for various non-local corrections, that will be described in the sequence sections. λ is the standard phase space factor defined as follows:

$$\lambda(a,b,c) = a^2 + b^2 + c^2 - 2(ab + bc + ac),$$

and the normalization factor reads:

$$N = V_{\rm tb}V_{\rm ts}^*\left[\frac{G_F^2\alpha^2}{3\cdot 2^{10}\pi^5 m_B^3}\lambda^{1/2}\right]^{1/2}. \tag{2.29}$$

$C_i^{\rm eff}$ are the effective Wilson Coefficients. They include the factorizable[9] and perturbative QCD effects, which is the reason why they are q^2 dependent. There exists a different class of corrections called the non-factorizable corrections mostly related to the quark loops, which we will discuss in Sec. 2.8.3. The effective coefficients are given as follows [16]:

$$\begin{aligned}
C_7^{\rm eff} &= \frac{4\pi}{\alpha_s}C_7 - \frac{1}{3}C_3 - \frac{4}{9}C_4 - \frac{20}{3}C_5 - \frac{80}{9}C_6\,,\\
C_8^{\rm eff} &= \frac{4\pi}{\alpha_s}C_8 + C_3 - \frac{1}{6}C_4 + 20C_5 - \frac{10}{3}C_6\,,\\
C_9^{\rm eff} &= \frac{4\pi}{\alpha_s}C_9 + Y(q^2)\,,\\
C_{10}^{\rm eff} &= \frac{4\pi}{\alpha_s}C_{10}\,, \qquad C_{7,8,9,10}^{\prime,\rm eff} = \frac{4\pi}{\alpha_s}C'_{7,8,9,10}\,,
\end{aligned} \tag{2.30}$$

$$\begin{aligned}
\text{with}\quad Y(q^2) &= h(q^2,m_c)\left(\frac{4}{3}C_1 + C_2 + 6C_3 + 60C_5\right)\\
&\quad - \frac{1}{2}h(q^2,m_b)\left(7C_3 + \frac{4}{3}C_4 + 76C_5 + \frac{64}{3}C_6\right)\\
&\quad - \frac{1}{2}h(q^2,0)\left(C_3 + \frac{4}{3}C_4 + 16C_5 + \frac{64}{3}C_6\right)\\
&\quad + \frac{4}{3}C_3 + \frac{64}{9}C_5 + \frac{64}{27}C_6\,.
\end{aligned} \tag{2.31}$$

The one quark loop function reads:

$$h(q^2,m_q) = -\frac{4}{9}\left(\ln\frac{m_q^2}{\mu^2} - \frac{2}{3} - z\right) - \frac{4}{9}(2+z)\sqrt{|z-1|}\times\begin{cases}\arctan\dfrac{1}{\sqrt{z-1}} & z>1\\ \ln\dfrac{1+\sqrt{1-z}}{\sqrt{z}} - \dfrac{i\pi}{2} & z\leqslant 1\end{cases}. \tag{2.32}$$

[9]The factorizable corrections are the ones that can be expressed as a constant times the form factor.

z is defined as:

$$z = \frac{4m_q^2}{q^2}. \tag{2.33}$$

2.8.1 Optimized observables in $B^0 \to K^{*0}\mu^+\mu^-$

As has been pointed out by the authors of [17], at the large recoil limit ($E_{K^*} \to \infty$) and at leading order of α_s and $\frac{1}{m_b}$ the transversity amplitudes can be written as:

$$A_\perp^{L,R} = \sqrt{2}N' m_B(1-\hat{s})\left[(\mathcal{C}_9^{(\text{eff})} + \mathcal{C}_{9)}^{'(\text{eff})} \mp (\mathcal{C}_{10} + \mathcal{C}_{10)}^{'} + \frac{2\hat{m}_b}{\hat{s}}(\mathcal{C}_7^{(\text{eff})} + \mathcal{C}_{7)}^{'(\text{eff})}\right]\xi_\perp(E_{K^*}) \tag{2.34}$$

$$A_\parallel^{L,R} = -\sqrt{2}N' m_B(1-\hat{s})\left[(\mathcal{C}_9^{(\text{eff})} - \mathcal{C}_{9)}^{'(\text{eff})} \mp (\mathcal{C}_{10} - \mathcal{C}_{10)}^{'} + \frac{2\hat{m}_b}{\hat{s}}(\mathcal{C}_7^{(\text{eff})} - \mathcal{C}_{7)}^{'(\text{eff})}\right]\xi_\perp(E_{K^*}) \tag{2.35}$$

$$A_0^{L,R} = -\frac{N' m_B(1-\hat{s})^2}{2\hat{m}_{K^*}\sqrt{\hat{s}}}\left[(\mathcal{C}_9^{(\text{eff})} - \mathcal{C}_{9)}^{'(\text{eff})} \mp (\mathcal{C}_{10} - \mathcal{C}_{10)}^{'} + 2\hat{m}_b(\mathcal{C}_7^{(\text{eff})} - \mathcal{C}_{7)}^{'(\text{eff})}\right]\xi_\parallel(E_{K^*}), \tag{2.36}$$

where ξ_i are the soft form factors, $\hat{s} = \frac{q^2}{m_B^2}$ and N' is the normalization factor. At first sight, these expansions have nothing out of the ordinary, besides the reduction of the number of form factors. However, one can note that in the language of the bilinear forms introduced in Eq. 2.25 one finds that $n_{\perp,\parallel} \propto \xi_\perp$ and $n_0 \propto \xi_\parallel$. This property has been used by J.Matias et. al. [18] to construct the so-called optimized observables:

$$P_1 = \frac{|n_\perp|^2 - |n_\parallel|^2}{|n_\perp|^2 + |n_\parallel|^2} = \frac{J_3}{2J_{2s}}\ , \tag{2.37}$$

$$P_2 = \frac{\text{Re}(n_\perp^\dagger\, n_\parallel)}{|n_\parallel|^2 + |n_\perp|^2} = \beta_\ell \frac{J_{6s}}{8J_{2s}}\ , \tag{2.38}$$

$$P_3 = \frac{\text{Im}(n_\perp^\dagger\, n_\parallel)}{|n_\parallel|^2 + |n_\perp|^2} = -\frac{J_9}{4J_{2s}}\ , \tag{2.39}$$

$$P_4 = \frac{\text{Re}(n_0^\dagger\, n_\parallel)}{\sqrt{|n_\parallel|^2|n_0|^2}} = \frac{\sqrt{2}J_4}{\sqrt{-J_{2c}(2J_{2s} - J_3)}}\ , \tag{2.40}$$

$$P_5 = \frac{\text{Re}(n_0^\dagger\, n_\perp)}{\sqrt{|n_\perp|^2|n_0|^2}} = \frac{\beta_\ell J_5}{\sqrt{-2J_{2c}(2J_{2s} + J_3)}}\ , \tag{2.41}$$

$$P_6 = \frac{\text{Im}(n_0^\dagger\, n_\parallel)}{\sqrt{|n_\parallel|^2|n_0|^2}} = -\frac{\beta_\ell J_7}{\sqrt{-2J_{2c}(2J_{2s} - J_3)}}\ . \tag{2.42}$$

As can be seen, the ratios are constructed in a way that the form factors $\xi_{\perp,\parallel}$ cancel in the ratios. Of course, there exist many more of this kind of ratios, but we follow the conventions established in the literature. In addition, there exists a basis of P_i' observables [19], which are closely related to the above-defined basis. We will discuss them in more detail in Sec. 4.6.5.

2.8.2 Precision on theoretical predictions

The results that will be described in this dissertation show clear tensions w.r.t. the SM predictions. In light of this information, we will discuss the uncertainty of the theoretical calculations. The first source of theoretical uncertainty comes from neglecting the higher order perturbation corrections to the WC at the m_b scale and accuracy of the matching procedure. The uncertainty is accessed by varying the EW scale in the range of $[m_W/2, 2m_t]$, while running down the WC. The uncertainty is small and is not currently considered to be a problem. The second uncertainty comes from the knowledge of the hadronic form factors. The form factors for the decay of $B^0 \to K^{*0}\mu^+\mu^-$ come from two independent calculations. In the large recoil regime (small q^2) the form factors are known from the light-cone sum rules [20]. On the other hand, in the low recoil limit (large q^2) the form factors are calculated using the lattice QCD [21–23]. Both of these calculations produce the dominant error in the theoretical prediction. The lattice and the light-cone sum rule results have been combined using the simplified series expansion to further reduce the uncertainty [15]. The fit results are shown in Fig. 2.3. The authors of this calculation have provided the full correlation matrix for the form factor uncertainties.

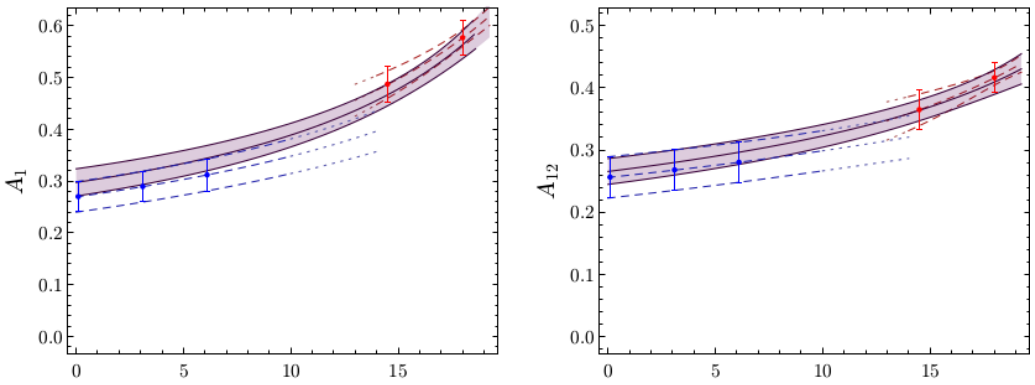


Figure 2.3: Example of the combined LCSR and lattice fit to $B \to K^*$ form factors [15]. Lattice data are marked with red lines, while the LCSR are marked in blue.

The most difficult to estimate in theoretical calculations are the effects of the non-local matrix elements of the QCD factorization framework. Following [24], the next-to-leading order corrections to the decay amplitudes ($\mathcal{T}_{\parallel,\perp}(q^2)$) read:

$$\begin{aligned}\mathcal{T}_a &= \xi_a\left(C_a^{(0)} + \frac{\alpha_s C_F}{4\pi} C_a^{(1)}\right) \\ &\quad + \frac{\pi^2}{N_c}\frac{f_B f_{K^*,a}}{M_B}\,\Xi_a \sum_{\pm}\int \frac{d\omega}{\omega}\,\Phi_{B,\pm}(\omega)\int_0^1 du\,\Phi_{K^*,a}(u)\,T_{a,\pm}(u,\omega),. \qquad (2.43)\end{aligned}$$

where $C_F = 4/3$, $N_c = 3$, $\Xi_\perp \equiv 1$, $\Xi_\parallel \equiv m_{K^*}/E$, f_{B,K^*} are the decay constants and $T_{a,\pm}(u,\omega)$ describes the hard scattering term. The functions ξ_a represent heavy-to-light

form factors [20, 25], while Φ - light-cone-distribution amplitudes. The function $T_{a,\pm}(u,\omega)$ is then expanded as:

$$T_{a,\pm}(u,\omega) = T^{(0)}_{a,\pm}(u,\omega) + \frac{\alpha_s C_F}{4\pi} T^{(1)}_{a,\pm}(u,\omega), \tag{2.44}$$

The leading order expansion occurs only through the penguin operators. The amplitudes modulus were calculated in [25] and are small compared to the leading order terms.

The next-to-leading corrections to the amplitudes $T^{(1)}_{a,\pm}(u,\omega)$ and $C^{(1)}_a$ from Eq. 2.43 are composed of two terms:

$$T^{(1)}_{a,\pm} = T^{(\mathrm{f})}_{a,\pm} + T^{(\mathrm{nf})}_{a,\pm}, \tag{2.45}$$

$$C^{(1)}_a = C^{(\mathrm{f})}_a + C^{(\mathrm{nf})}_a, \tag{2.46}$$

The first ones $\left(T^{(\mathrm{f})}_{a,\pm}/C^{(\mathrm{f})}_a\right)$, the so-called factorizable corrections, correspond to the expressions of the QCD form factors in terms of ξ_a.

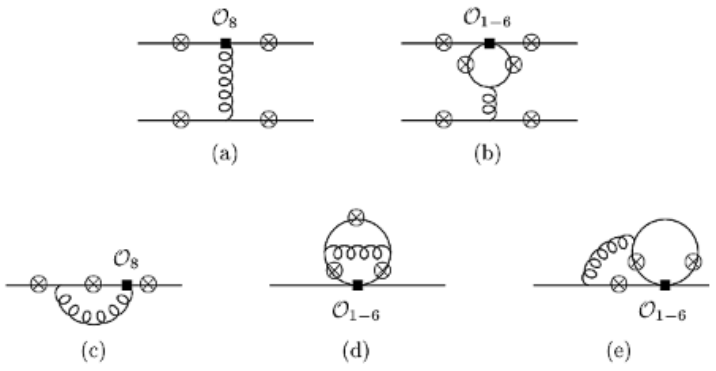


Figure 2.4: Non-factorizable contributions to the $B^0 \to K^{*0}\mu^+\mu^-$ decay [24] The circled crosses correspond to the possible insertions of the virtual photon.

The more challenging is the calculation of the non-factorizable corrections $T^{(\mathrm{nf})}_{a,\pm}/C^{(\mathrm{nf})}_a$. They are obtained by computing matrix elements of the four-quark operators and the chromomagnetic dipole operator shown in Fig. 2.4. The diagrams (a) and (b) correspond to the correction to the $T^{(\mathrm{nf})}_{a,\pm}$, while the diagrams[10] (c) to (d) - to the correction to $C^{(\mathrm{nf})}_a$. The numerical results of the non-factorizable corrections have been reported in [24].

In addition, at high q^2 the two leptons can be produced from broad charmonium resonances. The LHCb experiment has measured [26] the resonance contributions to be substantially more sizeable than anticipated [27] via naive factorisation indicating sizeable duality violations. Recently, the $e^+e^- \to$ hadrons BESII-data [28] have been used to determine them in a model independent way [27].

[10]Please note that the symmetrical diagrams of (c) and (e) are not shown.

2.8.3 Charm loop effects

Both of the above-mentioned factorisable and non-factorizable corrections at the next-to-leading power are inaccessible in the framework of the QCD factorization. A special case of those corrections called the "charm-loops" effect has been studied in [29]. In this work, the non-factorizable contributions with an additional soft gluon are discussed. The results of this study can be expressed in terms of two operators: one local and one non-local in an operator product expansion of the light-cone. The matrix elements of the local operator reproduce the results of the QCD factorization for the $C_a^{(\mathrm{nf})}$. On the other hand, the non-local operator introduces new, non form factor matrix elements, which are calculated using the light-cone sum rule at $q^2 = 1\ \mathrm{GeV}^2/c^4 \ll 4m_c^2$ and then extrapolated to $q^2 > 1\ \mathrm{GeV}^2/c^4$ via dispersion relations [30].

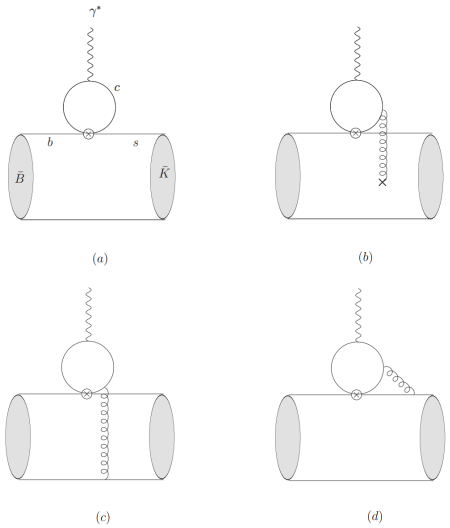


Figure 2.5: Charm-loop effect in $B^0 \to K^{*0}\mu^+\mu^-$ decays; (a)-the leading-order contribution (factorizable); (b) non-factorizale soft-gluon emission, (c),(d)-hard gluon exchange [29].

As a result of these calculations [29], the charm loops effect was found to influence $\mathcal{C}_9^{(\mathrm{eff})}$ at 20% level. The dominant contribution being the soft gluon exchange.

Chapter 3

Experimental strategies in measuring the $b \to s\ell^-\ell^+$ processes

This chapter gives a short description of experiments that are currently studying or will study the $b \to s\ell^-\ell^+$ processes in the near future. Special attention is put on the LHCb experiment, where the author has conducted his research.

3.1 Comparison of flavour physics at B-factories and hadron colliders

At present, heavy flavour physics measurements are performed in two types of colliders. The first type includes the so-called B-factories where e^+e^- are collided at the $\Upsilon(4S)$ and $\Upsilon(5S)$ thresholds[1]. The relevant B-factories are the Belle [31] and BaBar [32] experiments and they both finished data taking in 2010 and 2008, respectively. The Belle experiment is being upgraded and has started to collect data in 2017. The B-factories are considered to be a very clean environment to study B decays. In the decays of an $\Upsilon(4S)$ meson only two beauty flavoured particles are produced. Owing to this fact, the level of combinatorial background[2] is heavily suppressed. Furthermore, pile-up events do not occur because of the small cross sections, which further suppresses the background level. Last but not least, as we collide e^+ with e^-, which are elementary particles, the central-mass energy is known. This can be used as a discriminating variable in the selection or as a constraint in the analysis of the semileptonic decays.

The second type of the machines are the pp colliders. Currently, a prime example of the experiment located on such a collider is LHCb. It is worth pointing out that also the ATLAS [33] and CMS [34] experiments conduct heavy flavour measurements, although with a small number of final states accessible experimentally. In such a collider, the $b\bar{b}$ pairs are produced dominantly in a gluon-gluon fusion and $q\bar{q}$ annihilation processes [35]. The advantage of such a collider is a huge cross-section for the processes in question.

[1] $\Upsilon(4S)$ is used for $B\overline{B}$ production, while $\Upsilon(5S)$ for $B_s^0\overline{B}_s^0$ production

[2] This type of background arises when the tracks are assigned to wrong a B candidate.

Unfortunately, the price to pay for the higher cross section is the fact that besides the $b\bar{b}$ pairs a large number of particles in the interaction point are created, which increases significantly the combinatorial background. Moreover, the centre-of-mass energy of the $b\bar{b}$ system is not known. All this makes a hadron collider a much harder environment from the experimental point of view. There are decays which can reach much higher experimental precision on the hadron collider than the e^+e^- collider. This group includes mostly exclusive decays involving a $\mu^+\mu^-$ pair in the decay final state. This is because $\mu^+\mu^-$ pairs are produced not very frequently at the hadron collider so the trigger requirements for them can be much looser w.r.t. the hadronic B decays.

In this chapter a brief description of both the B-factories and pp machines with the experiments conducting heavy flavour measurements will be given.

3.2 Large Hadron Collider and LHCb detector

The Large Hadron Collider (LHC) [36] is currently the largest accelerator in the world located at Conseil Européen pour la Recherche Nucléaire (CERN) [37]. It is situated 100 m under ground in a 26.7 km circular tunnel. The LHC accelerator is composed of two beam pipes where the proton bunches travel in opposite directions in ultra-high vacuum. There are four experimental caverns in the LHC tunnel and four caverns that are used for technical maintenance and acceleration chain of the LHC. The first collisions were recorded in 2010 while the first proton beams were accelerated in 2008. The LHC uses protons that are firstly accelerated by other accelerators located at CERN. The whole accelerating chain is illustrated in Fig. 3.1.

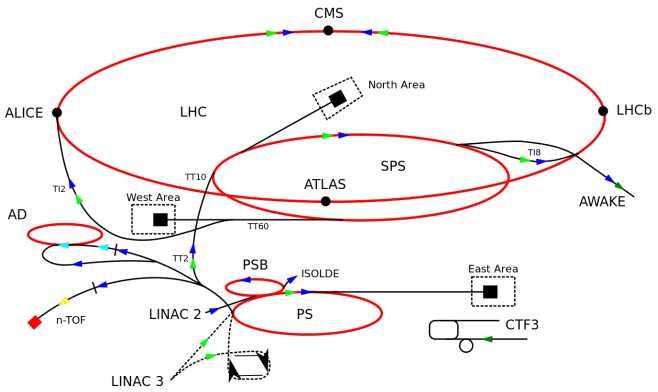


Figure 3.1: The layout of the CERN accelerator system [38].

Protons used for collisions are extracted from hydrogen atoms after stripping them out of electrons. These protons are then accelerated by Linear Accelerator 2 (LINAC2) to an energy of 50 MeV. The protons are then injected into the Proton Synchrotron Booster [39], which further increases their energy to 1.4 GeV. Next, the protons are accelerated by the Proton Synchrotron (PS) [40] and the Super Proton Synchrotron (SPS) [41], which

both increase their energy up to 25 and 450 GeV, respectively. It is worth noticing that both the PS and SPS accelerators are also used by several other experiments such as, CNGS [42], NA62 [43], or test beam facilities [44]. The LHC hosts seven experiments in total: ATLAS [33], CMS [34], LHCb [45], ALICE [46], TOTEM [47], LHCf [48] and MoEDAL [49].

The data used in the measurements described in this monograph were collected in 2011 and 2012. During this period the protons were collided with the centre-of-mass energy equal to 7 and 8 TeV, respectively. Besides the above-mentioned years, LHCb has at its disposal the data collected in 2010, 2015 and 2016. The first data set is too small for applications in rare decays. It is worth pointing out that the data sets collected in 2015 and 2016 are with a higher collision energy (13 TeV), which also entails higher $b\bar{b}$ cross sections. This datasets are currently being analysed by the author and the results are expected at the beginning of 2018.

In order to study decays of B mesons at the hadron collider, the key issue is the ability to effectively reject background. To do so, LHCb lowers its instantaneous luminosity delivered at its interaction point, compared to the other experiments (ATLAS and CMS), which is done by defocusing the proton beams. The smaller instantaneous luminosity ($\mathcal{L} = 4 \times 10^{32}$ cm^{-2} s^{-1} (Fig. 3.2 and Table 3.1) leads to a decrease in the so-called pile-up events, which in the end leads to a decrease in combinatorial background. For comparison, ATLAS and CMS have almost two orders of magnitude higher luminosity ($\mathcal{L} = 1 \times 10^{34}$ cm^{-2} s^{-1}).

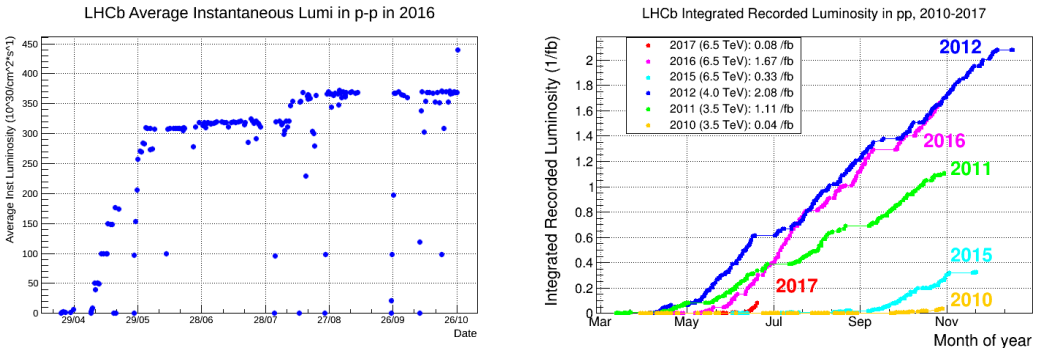


Figure 3.2: Instantaneous (a) and integrated (b) luminosity collected in the LHCb detector [50].

Table 3.1: The beam parameters at the LHCb beam crossing point in 2012 data taking.

Quantity	Achieved	Designed
No. protons in bunch [$\times 10^{11}$]	1.49	1.15
Luminosity(at LHCb) [$\times 10^{32}$ cm^{-2} s^{-1}]	4	2
Energy [TeV]	4	7
No. bunches	1380	2808

From the beginning the detector has been optimised for studying beauty flavoured mesons. It covers a pseudorapidity range: $2 < \eta < 5$, which corresponds to the geometric acceptance $(10-250)$ mrad $(10-300)$ mrad in the vertical (horizontal) plane, respectively. The unusual pseudorapidity coverage is dictated by the fact that the $b\bar{b}$ pairs are boosted in the direction of the higher energy parton, resulting in high correlations between both b's directions (cf Fig. 3.3). As a result, LHCb offers an optimum coverage for b flavoured mesons detection.

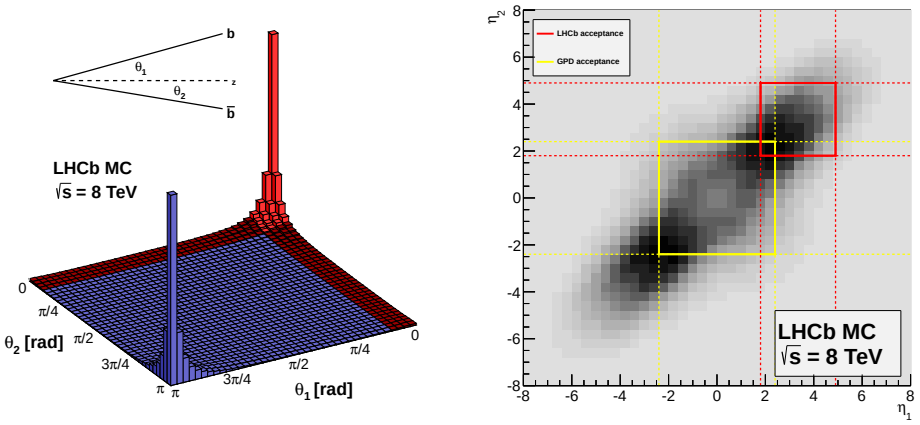


Figure 3.3: Simulated $b\bar{b}$ production angles at LHCb. Left: forward-backward production fraction as a function of b quark angles w.r.t. to the beam axis. LHCb acceptance is marked in red. Right: Distribution of pseudorapidities for $b\bar{b}$ pairs with the LHCb acceptance marked in red. [51].

The cross section of producing a $b\bar{b}$ pair for pp collisions at 7 TeV is measured to be $(75.3 \pm 5.4 \pm 13.0)\mu$ b [52]. Thanks to such a large cross-section, LHCb has already collected two orders of magnitude more b mesons than the B-factories during their full period of data taking.

The b flavoured hadrons are reconstructed in the LHCb detector as displaced secondary vertices (SV) of their decays. The b hadrons fly on average 1 cm, owing to the large boost of the $b\bar{b}$ pairs. The superb vertex resolution allows for a very efficient separation of the SV from the PV. Furthermore, the decay products of the b hadrons have a large impact parameter (IP) value w.r.t. the PV. Additionally, the SV tracks have on average higher values of transverse momenta in comparison with the PV tracks.

The LHCb detector consists of several sub-detectors (cf. Fig. 3.4), each designed for a different purpose. The reconstruction of charged tracks is performed using a Vertex Locator, a Tracker Turicensis and a magnet. Particle Identification (PID) is achieved by two Ring Imaging Cherenkov detectors and a Muon system. The energy of neutral particles is measured by a Scintillating Pad Detector, a Pre-Shower, an Electromagnetic Calorimeter and a Hadronic Calorimeter.

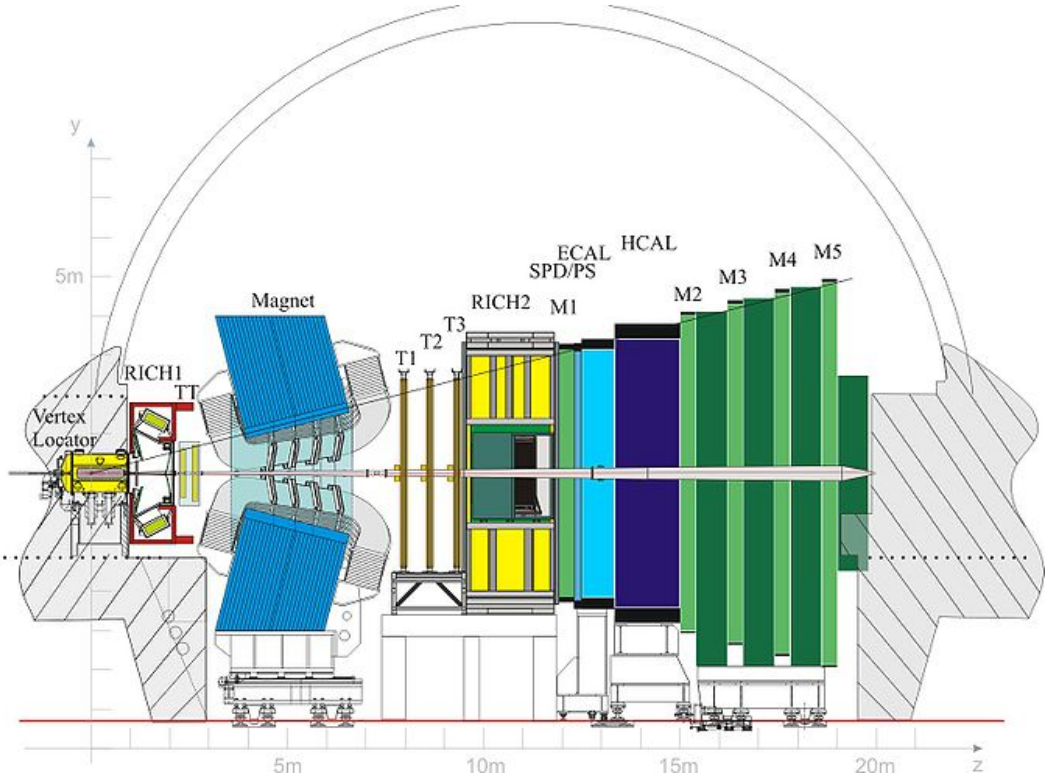


Figure 3.4: Schematic side drawing of the LHCb detector [53].

3.2.1 Vertex Locator

The Vertex Locator (VELO) is the detector located closest to the interaction point. Its purpose is to measure precisely the track position. It is built out of 42 modules of silicon microstrip sensors (Fig. 3.5), each containing the R-type and the Phi-type sensors. Both types of the sensors measure the radial and the azimuthal coordinate, respectively.

VELO is the only movable detector in all LHC experiments. During data taking the modules are located only 8 mm from the beam axis. However, during injection of the proton beams into the LHC from the SPS, the VELO modules are moved apart (the so-called open position) to ensure the safety of the detector. In the open position each module is 3 cm away from the beam axis. After the injection and ramp up of the LHC, the VELO modules are moved closer to the beam axis (the so-called closed position).

Thanks the VELO apparatus, the LHCb detector offers a superb spatial resolution: 10 µm in the x and y direction and 60 µm in z. This excellent resolution plays a crucial role in rejecting the combinatorial background. More details about the VELO detector can be found in [54].

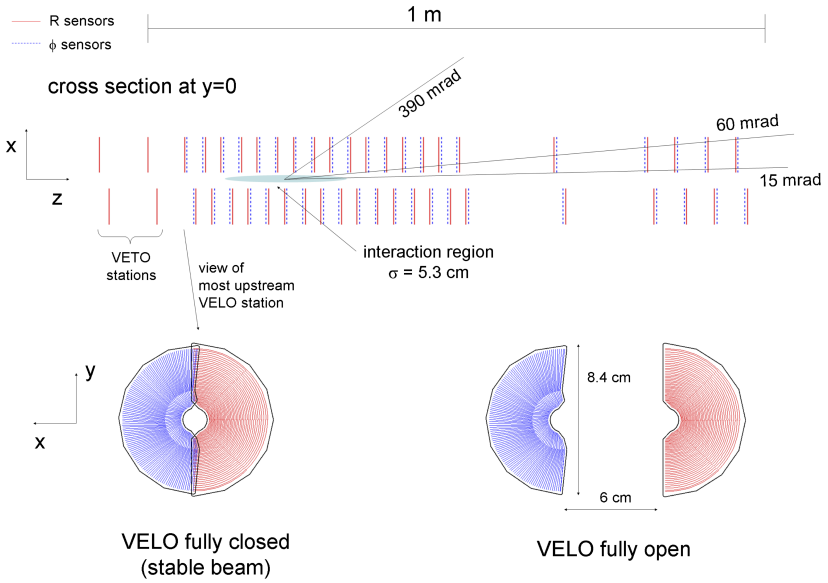


Figure 3.5: VELO module position. The lower drawings illustrate the closed (left) and open (right) configuration.

3.2.2 Tracking

The Tracker Turicensis (TT) is located between the magnet and the RICH1 detector as can be seen in Fig. 3.4. It is built out of four layers of p^+ on the n-type silicon, which are grouped in pairs (TTa and TTb), separated by 30 cm. The silicon strips that have been used for this detector have a pitch of 200 µm and thickness of $(320 - 410)$ µm. These dimensions have been optimized to achieve 50 µm hit resolution and occupancy of a few percent.

The Outer Tracker (OT) is located around 9 m from the interaction point between the magnet and RICH 2 detectors. It is built out of three stations named T1, T2 and T3. The technology used in it is drift time gas straw tubes. Each of the OT stations is composed of eight layers of drift tubes. The layers are tilted by a small angle (5 deg) to reduce the artificial hit combinatorics. The OT was designed to allow for fast track reconstruction in the trigger. Futher information about the tracking detectors can be found in [55].

A 4 Tm room temperature electromagnet bands the charge particle tracks. The curvature is used to determine the charge and the momentum of the particle. The polarity of the magnet is flipped a couple of times during the data taking period to cancel possible systematics from the detector asymmetries.

3.2.3 Ring Imaging Cherenkov detectors

The LHCb detector possesses two Ring Imaging Cherenkov (RICH) sub-detectors, marked as RICH1 and RICH2 in Fig. 3.7. They measure the particle velocity using the so-called

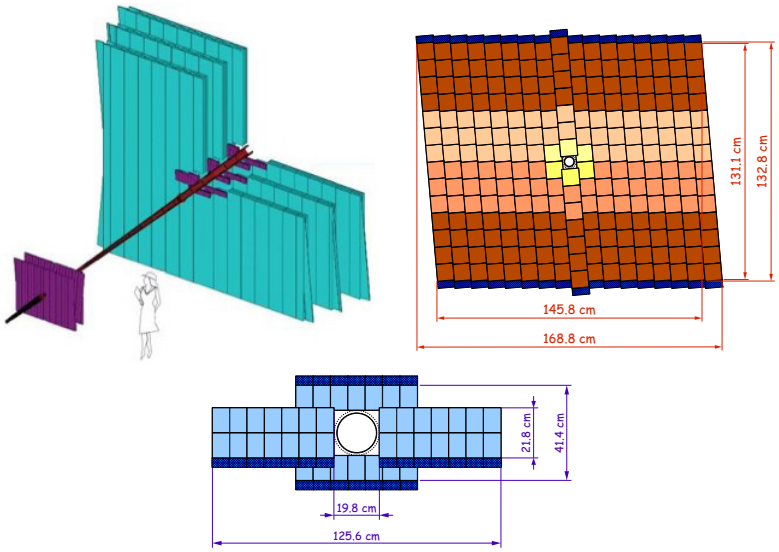


Figure 3.6: Left: schematic representation of TT (purple), IT (purple) and OT (cyan). Middle: the layout of the TT last layer. Right: the layout of the IT last layer.

Cherenkov effect. If a relativistic particle travels in a medium with a speed that exceeds the speed of light in that medium, it emits photons with an angle θ_C w.r.t. to the particle direction (the so-called Cherenkov angle). This phenomenon was discovered by P. Cherenkov in 1934 [56]. The radiation originates from medium excitation caused by the action of the field of the moving particle. The mass of the particle can be calculated using the measured velocity and the momentum of the particle (measured by the tracking detector described in Sec. 3.2.1 and 3.2.2).

The first of the RICH detectors (called RICH1) is situated after the dipole magnet. It performs particle identification for $(2-60)$ GeV/c momentum range. It uses a C_4F_{10} and aerogel radiators as Cherenkov light emitters. The light is then focused using spherical mirrors onto optical elements. Thanks to this solution the detector elements and electronics can be located outside the LHCb acceptance.

The second RICH detector (denoted as RICH2) is located futher downstream between T3 and ECAL. To identify particles it uses a different absorber: CF_4. The readout of both RICH detectors is based on Hybrid Photon Detectors (HPDs) [57], operating in the wavelength range $(200-600)$ µm. More information about the RICH system can be found in [58].

3.2.4 Calorimeters

The LHCb detector is composed of four calorimeters: Scintillating Pad Detector (SPD), Pre-Shower (PS), Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL). All of them are constructed in a "sandwich" structure, in which the absorber material layers are interchanged with the detecting material.

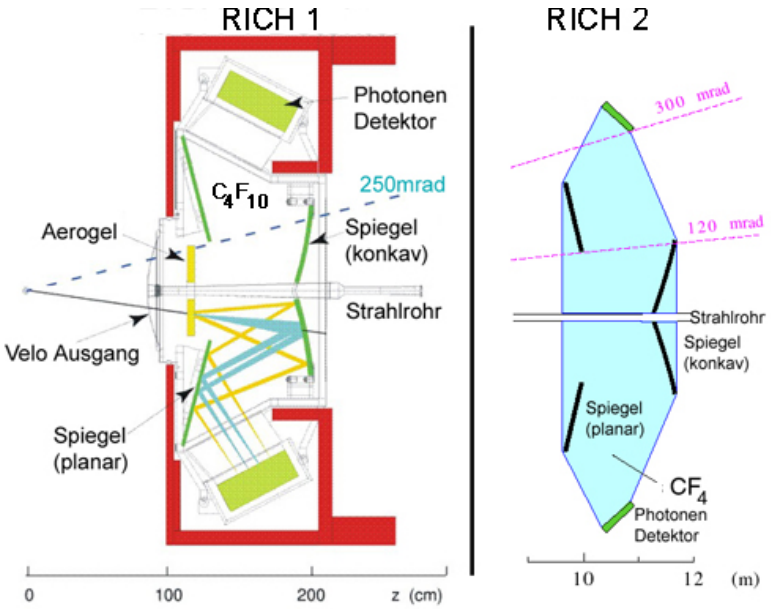


Figure 3.7: The layout of the RICH detectors

The SPD and PS detectors are the ones that are located closest to the interaction point. Their purpose is to improve the resolution of electromagnetic showers. They consist of 15 mm lead converters which are sandwiched between two identical planes of rectangular high-granularity scintillation pads. The SPD task is to distinguish which hits are coming from charge tracks and which from the neutral particles. The PS is used to reduce the backgrounds from charged pions thanks to its longitudinal segmentation.

The ECAL is built of 2 mm thick lead plates, which are alternated with 4 mm scintillating plates. The total thickness of the ECAL corresponds to 25 radiation lengths (X_0). In a similar manner the HCAL is constructed from layers of iron (16 mm) and scintillating tiles (4 mm). Further information about the calorimeter system can be found in [59].

3.2.5 Muon system

For the analysis described in this dissertation as well as for the core physics programme of LHCb the muons play the leading role. Owning to the superb performance of the muon system, the LHCb collaboration has reached the best results in rare semimuonic B decays compared to other experiments. The muons also play an important role in CP violation measurements where they are used in tagging of the B flavour [60].

As in a typical modern detector construction (such as the LHCb experiment), the muon system is located as the outermost layer w.r.t. the interaction point. In LHCb it is built of five stations: M1-M5, which are interleaved with 80 cm iron absorbers. The station itself is used in the trigger process. The modules are constructed from multi-wire

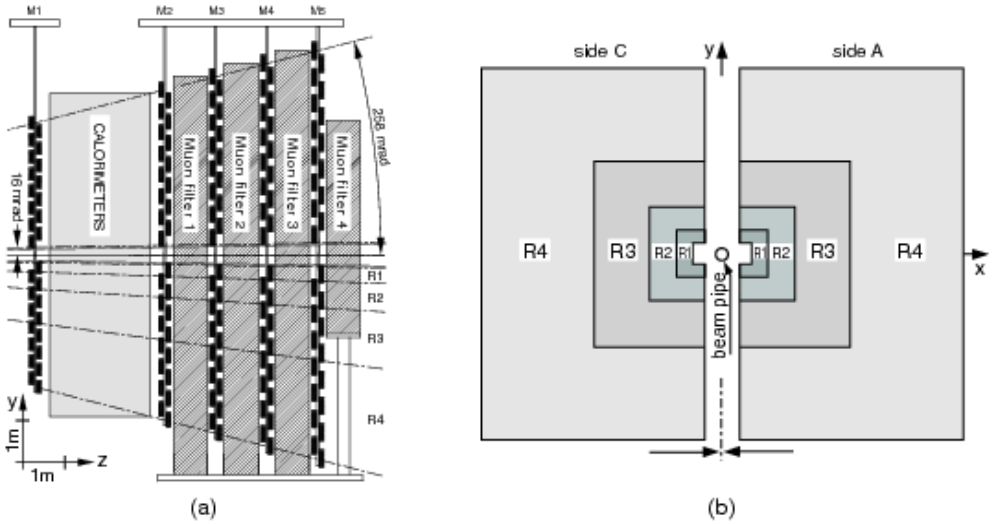


Figure 3.8: The muon system of the LHCb detector: (a) Side view, (b) station layout with four regions R1-R4 described in the text.

proportional chambers (MWPCs), with $(3-4)$ ns time resolution [61].

The LHCb muon system offers outstanding resolution in the x and y coordinates for charged tracks, which, when combined with the high detection efficiency [61] (above 95%), makes the LHCb detector an outstanding environment to study the semimuonic decays of B mesons. For more information about the muon system see [62].

3.2.6 Trigger

Given the high frequency of the collisions at the LHC, the events detected by the detectors have to be filtered before they can be stored on tape. This is the task of the LHCb trigger system. It is composed of two parts: First Level Trigger (L0) and High Level Trigger (HLT). The first one is purely a hardware trigger, which uses information from the calorimeter and muon systems, reducing the event rate from 11 MHz down to 870 kHz [63]. On the other hand the HLT is purely a software trigger and is divided into two phases: HLT1 and HLT2. HLT1 uses only partially reconstructed tracks for the event classification, while HLT2 uses the fully reconstructed events. They reduce the rate to 3 kHz. After this procedure the events are stored on tape.

3.3 CMS and ATLAS detectors

As mentioned in Sec. 3.1, also the general purpose detectors like ATLAS and CMS measure the processes like $B^0 \to K^{*0}\mu^+\mu^-$. It is worth mentioning that because of the worse resolution and lack of particle identification both experiments have not measured the full

set of angular observables of the above-mentioned decay. It is estimated that with more accumulated data this measurement will become possible in future.

3.4 ATLAS experiment

A Toroidal LHC ApparatuS (ATLAS) is the largest multi-purpose detector at the LHC accelerator. The detector has a length of 44 m and height of 25 m and weighs about 7000 tons [64]. A cut-away view of the ATLAS detector is presented in Fig. 3.9. In the most central region, the detector elements are arranged in concentric cylinders, while the endcap regions consist of discs of detectors perpendicular to the beam axis. The ATLAS detector is built of three main subsystems, namely the Inner Detector, the Calorimeter System, and the Muon Spectrometer. The detector has been designed to have a forward-backward symmetry w.r.t. the interaction point.

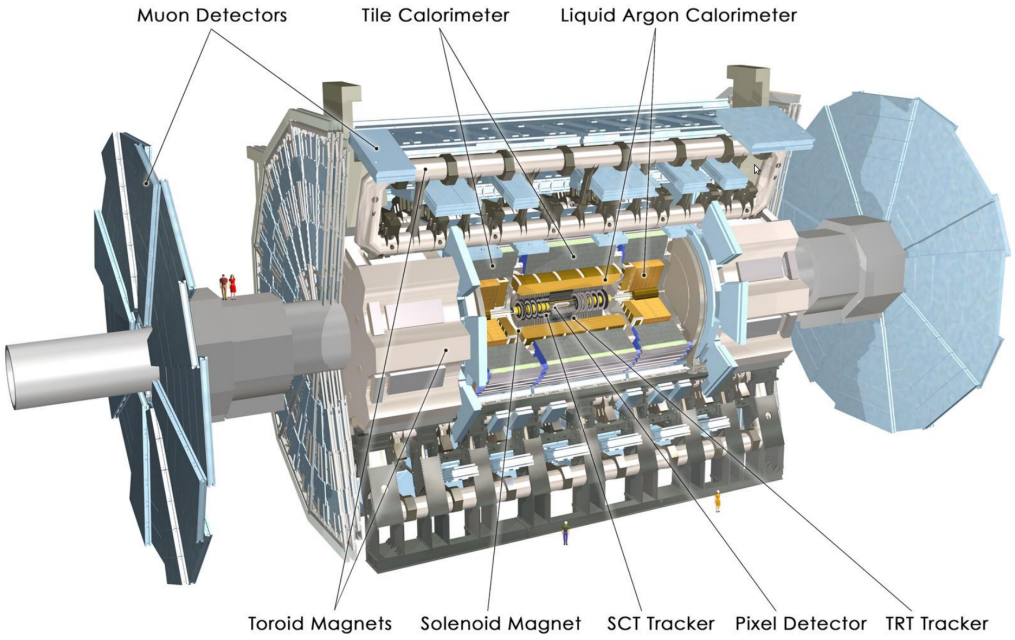


Figure 3.9: Side view of the ATLAS detector.

3.4.1 Inner detector

The Inner Detector is the detector located closest to the interaction point. It is immersed in a 2 T magnetic field oriented towards the z axis, which is generated by a superconducting cylindrical coil. It covers a central pseudorapidity region of $|\eta| < 2.5$. The Inner Detector consists of a Pixel Detector, a Semiconductor Tracker and a Transition Radiation Tracker.

Pixel Detector

The Pixel Detector is part of the detector innermost w.r.t. to the interaction point. It consists of high granularity silicon pixel sensors that are used for track and vertex reconstruction. The presence of a strong magnetic field allows also for momentum measurements. In the barrel part of the inner detector the pixels are arranged as three concentric cylinders of radii 51 mm, 89 mm and 123 mm, respectively. The pixels have an outstanding resolution of 10 μm

It is worth mentioning that for the Run2 data taking period an additional layer (called the B-layer) [65] was inserted at a distance of 3.3 cm from the beam pipe. The purpose of this upgrade was to increase the precision of the reconstruction of secondary vertices, in which ATLAS had sufficiently poorer performance compared to CMS. This layer also has a slightly wider pseudorapidity range: $|\eta| < 2.9$.

Semiconductor Tracker

Outside of the Pixel Detector there is the Semiconductor Tracker detector with a strip pitch of 80 μm. Each module consists of two layers of silicon strip detectors that are tilted by an angle of 40 mrad. In total there are four layers of these detectors in the barrel and nine in the endcaps. The total resolution of such a detector is 17 μm in the $R - \phi$ and 580 μm in the z direction.

Transition Radiation Tracker

The outermost section of the Inner Detector is the Transition Radiation Tracker. It is composed of straw drift tubes with a diameter of 4 mm filled with a $Xe/CO_2/O_2$ gas mixture. The resolution of each straw is 130 μm in the $R-\phi$ plane. Among the straw tubes there are layers of fibres and foils with different dielectric constants. Passing electrons emit light that can be absorbed by xenon in the straw tubes, resulting in a higher voltage signal when transition radiation is present. Because of this, the Transition Radiation Tracker allows for the separation of energy deposits due to electrons and other charged hadrons.

3.4.2 Calorimeter system

There are two main calorimeters in the ATLAS detector: an Electromagnetic Calorimeter and a Hadronic calorimeter. The Electromagnetic Calorimeter measures the energy of photons and electrons via absorption. It also measures partially the energy of hadrons, which are not fully stopped in the Electromagnetic Calorimeter and therefore can pass through to the Hadronic Calorimeter. A diagram showing the different parts of the calorimetry system can be seen in Fig. 3.10. As typical in modern experiments, the ATLAS calorimeters are sampling devices, where the absorbing material (the absorber) is sandwiched with the active material that measures the energy of particles produced in the absorber.

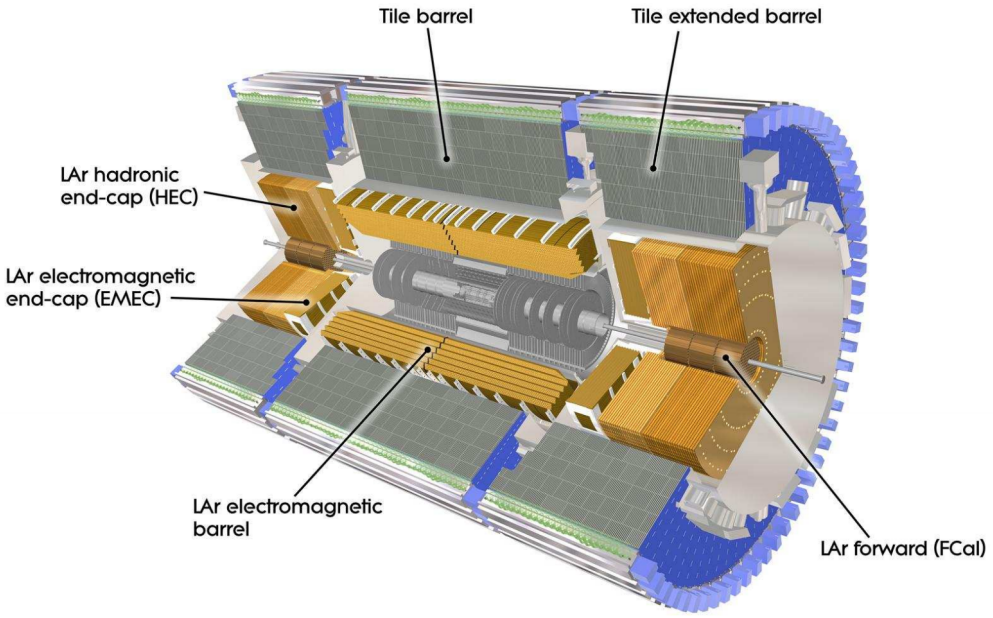


Figure 3.10: Illustration of the ATLAS calorimeter system

Electromagnetic calorimeter

The electromagnetic calorimeter uses liquid argon as active medium and consists of alternating layers of accordion-shaped kapton electrodes and lead absorber plates. The calorimeter in the barrel is greater than 22 X_0 and covers a pseudorapidty region of $|\eta| < 1.475$. The endcap part covers $1.375 < |\eta| < 3.2$ and is of the thickness of 24 X_0. The calorimeter produces an analogue signal that is sampled every 25 ns.

Hadronic calorimeter

The electromagnetic calorimeter is surrounded by the Hadronic calorimeter system. The hadronic calorimeter consists of a barrel tile calorimeter ($|\eta| < 1.7$) and two liquid argon sampling endcap calorimeters ($1.5 < |\eta| < 3.2$). The scintillation light is collected by the wavelength shifting fibres and detected using the photomultipliers. This results in a pulse with a height proportional to the energy and width of 50 ns. The hadronic calorimeters do not need such a readout time as the electromagnetic ones because they are less sensitive to pile-up events. Copper and tungsten are used as the absorption material.

3.4.3 The muon system

The muon spectrometer, shown in Fig. 3.11, is the outermost subsystem of ATLAS. It measures the energy of muons which can escape the calorimeters system. As in the case of LHCb, the muon system is used in the triggering system for the analysis described in this monograph. The superconducting air-core toroid magnets bend the muon tracks. Each

magnet consists of eight coils and produces average magnetic fields of about 0.5 T and 1 T in the barrel and the endcap regions, respectively. The barrel consists of three layers of precision chambers covering a pseudorapidity range of $|\eta| < 1.4$, while the endcaps cover a pseudorapidity region up to 2.7.

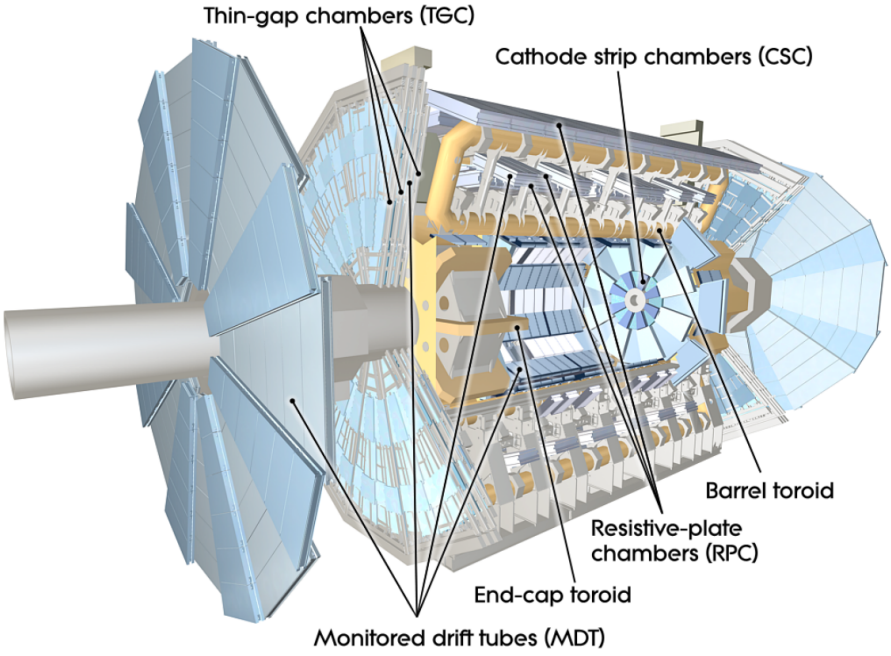


Figure 3.11: View of the ATLAS muon spectrometer.

3.5 CMS experiment

The Compact Muon Solenoid (CMS) is the second multi-purpose detector at the LHC. It is a cylindrical, 21.6 m long and 15 m high spectrometer, which weighs about 14000 tons. Similarly to ATLAS, it is composed of a barrel and two endcaps, as can be seen in Fig. 3.12. In general, the CMS detector consists of a tracker, an electromagnetic and a hadronic calorimeter and a muon system.

3.5.1 Inner Tracker

The Inter Tracker is the sub-detector that the particles produced in the collision encounter first. It is situated in a 3.8 T magnetic field generated by a superconductive solenoid. In terms of dimensions, the Inner Tracker is 5.8 m long and has a diameter of 2.5 m. This is the biggest tracking detector in all LHC experiments with a total active silicon area of about 200 m^2. Thanks to such precise apparatus, the vertexing and momentum resolution

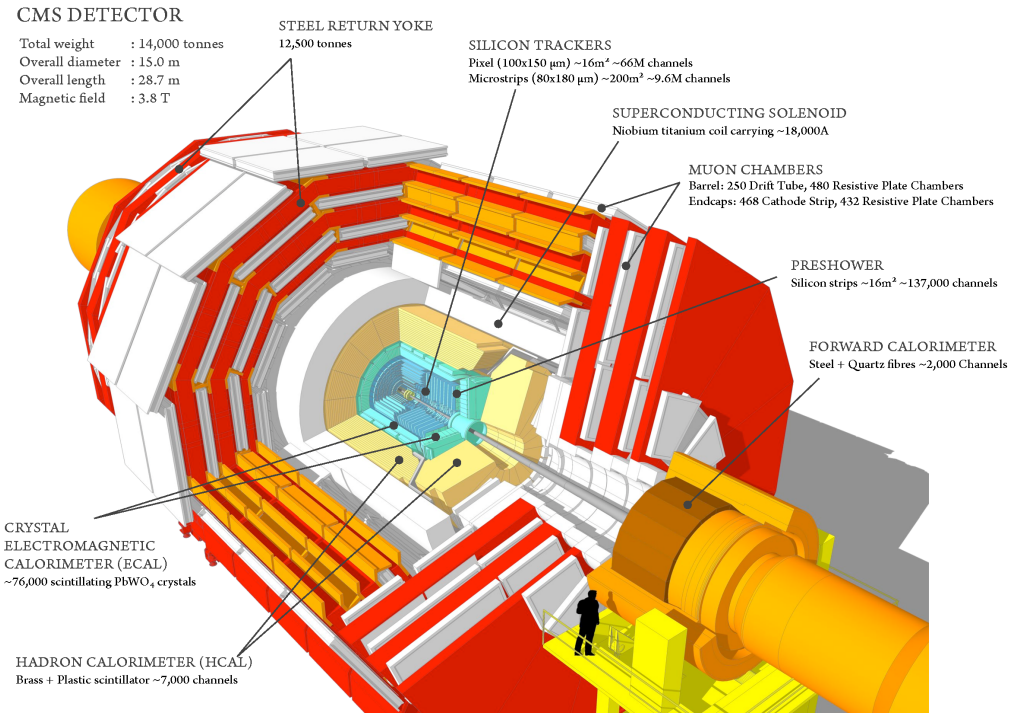


Figure 3.12: Schematic ilustration of the CMS detector.

are much superior compared to those of ATLAS. This is the main reason why heavy flavour results from CMS supersede the ones from ATLAS. The momentum resolution of the CMS inner detector is comparable to the LHCb one and is around 1 % for a transverse momentum range: $p_T < 100$ GeV. Furthermore, it covers a pseudorapidity region of $|\eta| < 2.5$. Unlike the ATLAS inner detector, the CMS detector is built entirely from silicon and consists of a Pixel Detector and a Silicon Strip Tracker, as can be seen in Fig. 3.13.

Pixel Detector

The Pixel Detector is the innermost component of the Inner Tracker. Its main purpose is to reconstruct primary and secondary vertices. The pixels used in it have the size of $100 \times 150\,\mu\text{m}^2$. This fine granularity ensures similar resolution in the $R - \phi$ plane as well as in the z direction, making the precise vertex reconstruction possible. The Pixel Detector consists of 18 million pixels with a total spatial resolution of about 20 μm.

Silicon Strip Tracker

Outside of the Pixel Detector lies the Silicon Strip Tracker. The sensors are $p - n$ silicon microstrip sensors[3]. It is composed of 15.000 silicon modules and 9.3 million readout channels. The entire system operates at the temperature below 10°C, which increases the

[3]Both single-sided and double-sided are used.

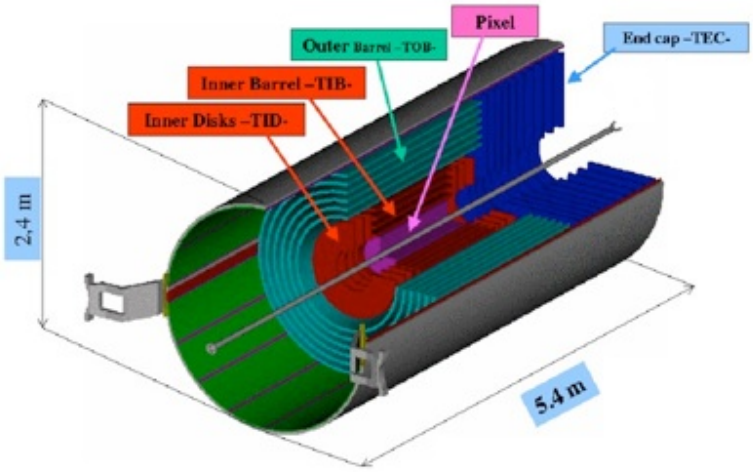


Figure 3.13: Schematic ilustration of the CMS inner detector.

radiation hardness of the detector. The sensors and electronics are mounted on a carbon or graphite fibre frame which ensures the low material budget of the overall detector. There exist over 20 different geometries of the modules used to build this detector [66]. The resolution depends on the type of the sensors and is within $(15 - 45)$ μm.

3.5.2 Electromagnetic calorimeter

The CMS Electromagnetic Calorimeter is composed of 75848 lead tungstate ($PnWO_4$) crystals exhibiting an excellent energy resolution. The barrel part of the eletromagnetic calorimeters covers the pseudorapidity region of $|\eta| < 1.5$ and the two endcaps extend the coverage up to $|\eta| < 3.0$. The barrel has the inner radius of 129 cm and is built of 36 identical modules, each composed of 1700 crystals. The endcaps are located 314 cm from the interacting point and are built of half-discs, each consisting of 3662 crystals.

3.5.3 Hadronic calorimeter

The hadronic calorimeter works together with the above-mentioned electromagnetic calorimeters to measure the energy of hadronic particles and jets as well the missing transverse energy. It uses a brass absorber interlaced with plastic scintillation layers. The unusual choice of brass was dictated by its short interaction length and non magnetic nature. Altogether the hadronic calorimeter covers the pseudorapidity region of $|\eta| < 5.2$.

3.5.4 The Muon System

Muons play an important role in the CMS experiment as its name indicates. They are the particles that are not stopped by the calorimeters system, making them very easy

to distinguish and detect. For this purpose, CMS uses its Muon System which is the outermost detector. It consists of three different types of gaseous detectors. In the barrel ($|\eta| < 1.2$) where one expects a low track multiplicity and small magnetic field, the Drift Tube Chambers have been used. On the other hand, the endcaps are equipped with the Cathode Strip Chambers, which can cope with a much higher track multiplicity. Similarly to the LHCb and ATLAS detectors, CMS also uses its Muon System as the main input for triggering muon events.

3.6 BaBar experiment

The BaBar experiment is located on the PEP-II accelerator, which operated in the years 1999-2007. The accelerator collided electrons with positrons mostly at an energy of 10.58 GeV, which corresponds to the $\Upsilon(4S)$ threshold. It is worth pointing out that this energy is only 20 MeV above the $B\overline{B}$ threshold. The PEP-II was an asymmetric e^+e^- collider, where the electrons and positrons had an energy of 9.0 GeV and 3.1 GeV, respectively. The illustration of PEP-II is given in Fig. 3.14. The asymmetric beams caused the central-mass of the system to have a boost of: $\beta\gamma = 0.56$ in the laboratory frame.

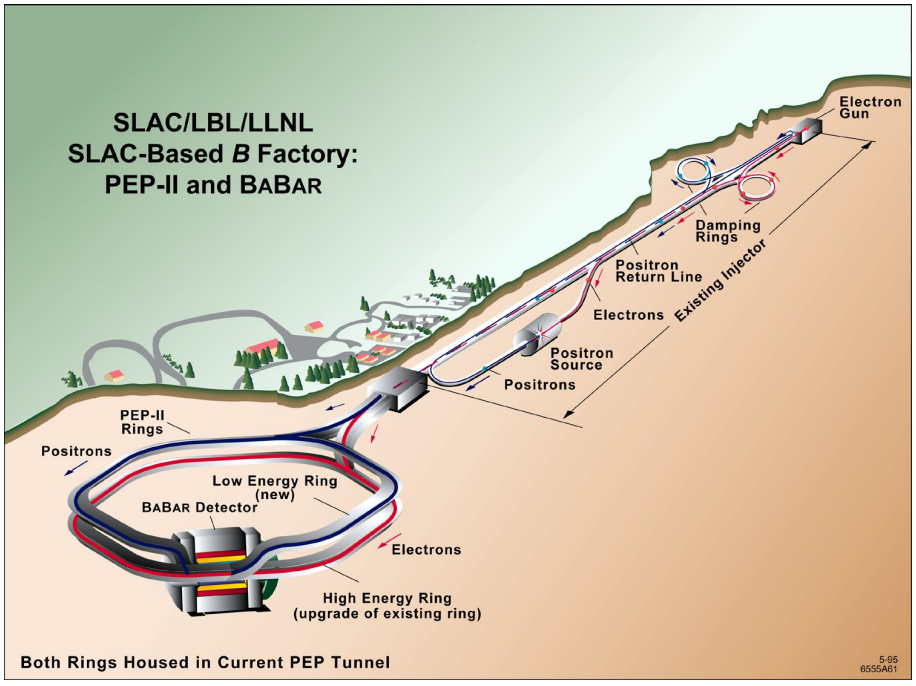


Figure 3.14: Schematic illustration of the PEP-II Linac, which accelerates electrons and positrons, and the two counter-rotating storage rings.

During its operation the BaBar detector has collected a total of 531.43 fb^{-1} of e^+e^- collisions, including runs at the $\Upsilon(2,3,4)$ thresholds.

The BaBar detector was designed as a spectrometer with almost 4π angular acceptance.

Because of the aforementioned beam energy asymmetry, the detector is also slightly asymmetric around the interaction point (shifted by 0.37 m). The track momenta of particles from B decays are usually below 1 GeV/c, so the detector needs to be built from a small amount of the material to reduce multiple scattering and increase the momentum resolution. The BaBar spectrometer consists of a Silicon Vertex Tracker, a multi-wire Drift Chamber, a Ring-Imaging Cherenkov detector, an Electromagnetic Calorimeter, a superconducting solenoid magnet and an Instrumented Flux Return. The location of all these subdetectors is shown in Fig. 3.15. We will briefly discuss the function and technology of the subdetectors.

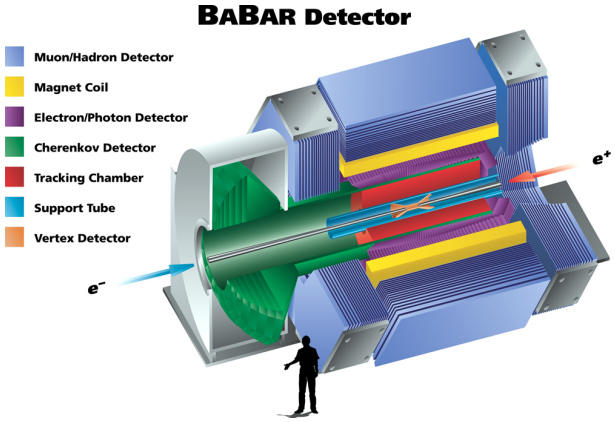


Figure 3.15: A schematic drawing of the BaBar detector.

3.6.1 Silicon Vertex Tracker

The Silicon Vertex Tracker (SVT) is built of five layers of double-sided silicon strip detectors around a beryllium beam-pipe. The strips on the inner side are oriented perpendicularly to the beam axis to provide z-coordinate measurements, while the outer side provides the ϕ coordinate. The main purpose of the SVT is to ensure precision measurements of the B decay vertex. All combined layers have a resolution of $(10-15)$ µm per track position and $(60-100)$ µm for the B vertex position, while the constituent material only contributes 4% of the radiation length.

3.6.2 Drift Chamber

The second detector responsible for tracking is the Drift Chamber (DCH). It provides the momentum measurement for particles with the transverse momentum greater than 120 MeV/c. The DCH consists of 40 layers of wires in 280 cm long cylinders with a helium-isobutane gas mixture. Amongst the aforementioned 40 layers, 24 have a small stereo angle with respect to the beam axis which enables the measurement of the longitudinal positions. When a charge particle passes through the DCH, it interacts with electrons of the atoms from the gas mixture, leaving a trail of free electrons, which drift to the wires because of

the electric field. The total collected charge provides the information of the dE/dx with an accuracy of 7%. The drift time provides the radial distance with a resolution of 140 μm.

3.6.3 Detector of Internally Reflected Cherenkov light

The Detector of Internally Reflected Cherenkov light (DIRC) provides the particle identification using Cherenkov radiation similar to the aforementioned RICH detector (cf. Sec. 3.2.3). The DIRC ensures acceptance of 94% in the azimuthal angle and 83% coverage in the polar angle in the centre-of-mass frame while adding only 17% X_0 of the material budget.

3.6.4 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (EMC) is an absorption calorimeter designed to detect electromagnetic showers within the 20 MeV − 4 GeV energy range. It is the first innermost detector capable of detecting neutral particles. The EMC is the most expensive subdetector in the BaBar experiment. It consists of 6580 cesium iodide (CsI) salt crystals. When a photon or an electron with at least a few MeV passes through the crystal, it interacts with the electric field produced by large atoms (i.e. with high atomic number Z) within the crystal material. Owing to this phenomenon, the photons create e^+e^- pairs and produce electromagnetic showers. At the same time the crystals generate scintillating light which is contained within the crystal via total internal reflection. The light is collected by the photo-diods mounted on the back of each crystal.

In total the EMC contains 5760 crystals arranged in 48 rings around the barrel and 820 crystals in the forward endcap arranged in eight azimuthal rings. It covers a total of 90% solid angle acceptance.

3.6.5 Instrumented Flux Return

The Instrumented Flux Return (IFR) is designed to identify the muons to about 600 MeV energy as well as neutral hadrons, such as K^0_{L} and neutrons. It is located outside the magnetic field and is responsible for directing the field lines for the return of the magnetic flux from the solenoid. The IFR consists of a barrel and two end-caps, each segmented into layers ranging from 2 to 10 cm in thickness for the outermost layers. Among the 19 layers of steel in the barrel (and 18 layers in the end-caps), there are either resistive plate chambers (RPCs) or limited streamer tubes (LSTs) serving to detect "streamers" from ionizing particles. An illustration of the IFR is shown in Fig. 3.16.

Thanks to the IFR the muon detection efficiency reaches the value of 83% while keeping the pion mis-identification at the level of 1.2%.

3.7 Belle experiment

The second B-factory is the Belle experiment located at the KEKB accelerator in Japan. It took data in the years 1999-2010. Similarly to the PEP-II accelerator, KEKB is an

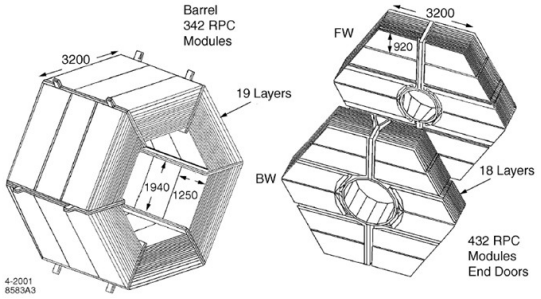


Figure 3.16: Schematic view of the IFR detector, specifically depicting the shape and instrumented layers of the (left) barrel and (right) two end caps.

asymmetric e^+e^- collider. However, contrary to the PEP-II, the energy of the beams are 8.0 GeV and 3.5 GeV for the electron and positron beams, respectively. Because of this the boost is also smaller than in PEP-II and is equal to $\beta\gamma = 0.43$. Smaller asymmetries in the energy of the beams allowed the collider to achieve larger luminosities. For this reason the Belle experiment has collected a total of 1040 fb^{-1}. In addition, 110 fb^{-1} have been collected at the $\Upsilon(5S)$ threshold.

The Belle detector is constructed in a similar manner to the BaBar one. It consists of a Silicon Vertex Detector, a Central Drift Chamber, an Aerogel Cherenkov Counter, a Time of Flight detector, an Electromagnetic calorimeter and a $K^0_{\rm L}$ and muon detector.

3.7.1 Silicon Vertex Detector

The Silicon Vertex Detector (SVD) is the detector closest to the interaction point, whose main purpose is to measure the decay vertices of the heavy flavoured particles and the momentum of their charged decay products.

The first version of the SVD was located on a 20 mm beryllium pipe, which is a part of the accelerator. It consisted of three layers located at 30.0 mm, 45.5 mm, 60.5 mm radial distance from the beam axis, respectively. They were built in double-sided silicon strip detector technology. The $n+$ strips are used to measure the z coordinate, while the $p+$ measure the coordinate on the $r - \phi$.

In 2003 the SVD detector was upgraded. The new setup now consists of four new SVD layers and a new beryllium beam pipe with 15 mm radius. The four layers are located at 20.0 mm, 43.5 mm, 70 mm and 88.0 mm radius from the beam axis, respectively. With the new configuration the solid angle coverage has increased to 92% (the old configuration had 86% solid angle coverage). A schematic drawing of the SVD detector is shown in Fig. 3.17.

Thanks to the upgrade the detector resolution in the z direction has increased dramatically from 36 µm to 27.6 µm, which has a direct impact on the time dependent measurements.

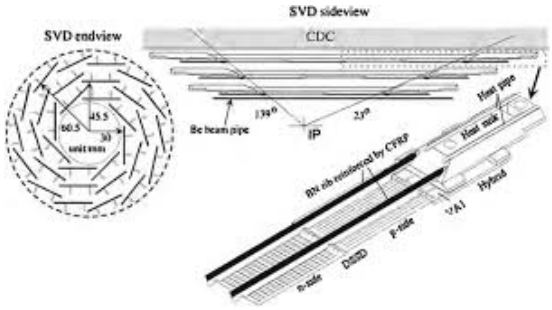


Figure 3.17: Schematic illustration of the SVD subdetector in the Belle experiment.

3.7.2 Central Drift Chamber

The main purpose of the Central Drift Chamber (CDC) is the measurement of the momentum of charged particles. The CDC has the same solid angle coverage as the SVD detector. The gas mixture used in the chamber is an equal mixture of ethane and helium. Such a choice was the result of the attempt to minimalise the multi-scattering effects in the mixture, which is crucial for low momentum track reconstruction. The typical resolution achieved by the CDC in the z direction is 130 μm, while in the $r - \phi$ direction the resolution is $(200 - 1400)$ μm depending on the particle momentum and direction.

The momentum reconstruction is done using the information from the SVD and CDC simultaneously. The overall extrapolation of the tracks from the two detectors has reached an amazing efficiency of 98%. Additionally, the CDC can help with the particle identification by measuring the ionization energy (dE/dx).

3.7.3 Aerogel Cherenkov Counter

The Aerogel Cherenkov Counter (ACC) has been built with the purpose of identifying the charge tracks with the momentum above 1 GeV/c. The detector uses the Cherenkov radiation already discussed in Sec. 3.2.3. The trick used in the Belle experiment was to choose such a refractive index of the medium so that only lighter particles such as pions could emit light. In the Belle experiment the silica aerogel has been chosen as a diffractor.

The ACC is built out of 960 modules in the barrel and 288 modules in the endcap. Each module contains five aerogel tiles stacked in aluminium boxes, as shown in Fig. 3.18 [67]. The ACC detector allows for measurement of the particle identification properties up to 4 GeV/c momentum.

3.7.4 Time of Flight detector

The Time of Flight (TOF) detector helps, to identify the particles in the low momentum range: $(1 - 2)$ GeV/c. The principle of the operation of this detector is to use scintillating counters to measure the speed of the charged particles. Combining this information with the momentum measurement from the SVD and CDC allows us to measure their mass.

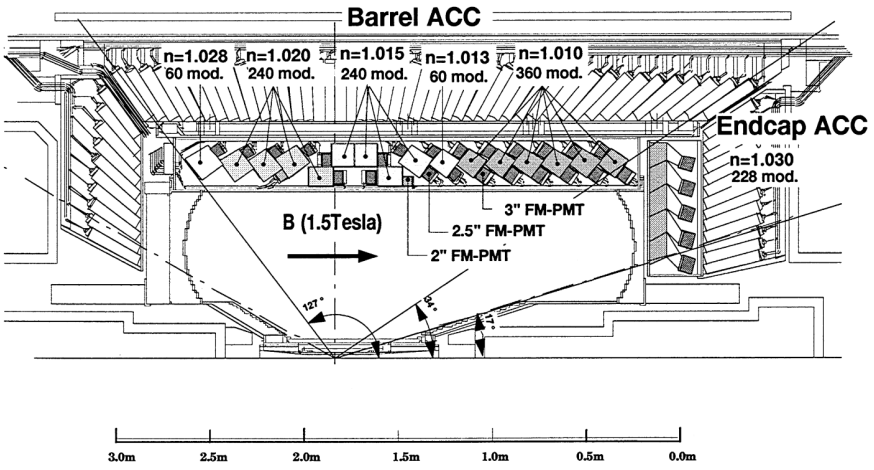


Figure 3.18: Schematic drawing of the ACC sub-detector in the Belle experiment. The regions with the same diffraction index are marked.

The counters are grouped in 64 modules located 128 cm from the interaction point. The time resolution of the TOF detector is 100 ps.

3.7.5 Electromagnetic Calorimeter

The next subdetector counting from the interaction point is the Electromagnetic Calorimeter (ECL). Its main purpose is to measure the energy of photons and electrons. NaCsI crystals are used as active material. The particles interact with the crystals producing electromagnetic cascades, which in turn generate the scintillation light inside the crystal, which is then measured with photomultipliers. The calorimeter is composed of 6624 crystalsin total, each 30 cm long. The total material budget of the ECL is 16.2 radiation lengths.

The endcap on the other hand is built of $Bi_4Ge_{30}O_{12}$ which is a more radiation hard material. The ECL endcap is also used as a beam luminosity monitor.

3.7.6 K_{L}^0 and Muon detector

The most outer subdetector of the Belle experiment is the K_{L}^0 and Muon (KLM) detector. Its main purpose is the detection of muons and long lived particles such as K_{L}^0. It consists of a barrel and two endcaps which together cover the polar angle range: $(20 \leqslant \theta \leqslant 155)$ deg. It is built of resistive plate chambers interleaved with iron absorption plates. The particles can be identified by the KLM detector if their momentum exceeds 600 MeV/c.

As mentioned above, the second purpose of the KLM detector is the identification and energy measurement of the K_L mesons. The identification of these measons is done via reconstruction of the cascades, which gives the information the energy of the particles as well as their direction.

3.8 Belle 2 experiment

In 2017 a new B-factory called the SuperKEKB (an upgrade of KEKB) will start colliding e^+e^-. In this section we will briefly discuss the Belle 2 detector. It is an upgrade of the previously described Belle detector. A comparison of the two constructions is shown in Fig. 3.19. Because of their similarities we will discuss only the major changes.

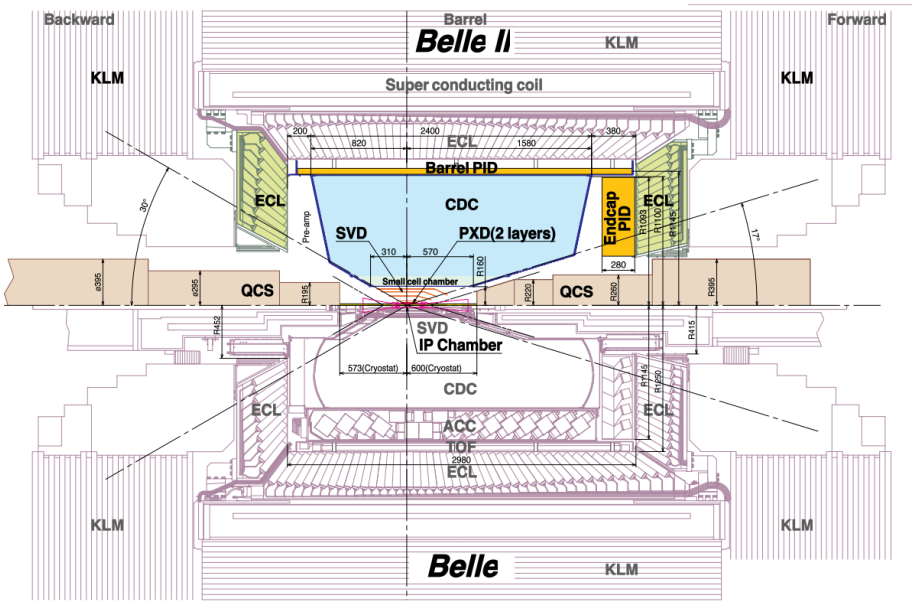


Figure 3.19: Schematic drawing of the Belle (bottom) and Belle 2 (top) detectors [68].

In order to achieve the physics goal the accelerator has gone through a major upgrade. The instantaneous luminosity will be increased from 2.1×10^{34} cm^{-2}s^{-1} to 8×10^{35} cm^{-2}s^{-1} by increasing the current and reducing the beam size. Furthermore, the crossing angle between the beams has been increased.

Because of the increased luminosity, the beams have changed to be more symmetric in terms of their energies. The B mesons have lower momenta in the laboratory rest frame: for this reasons one needs a more precise tracking detector to measure their lifetime with a precision similar to that achieved by Belle. The SVD has been upgraded with two new layers of a silicon pixel detector that are put very close to the beam pipe (14 and 22 mm respectively). Furthermore, the CDC has been extended in the radius and has more cells near the beam pipe. Additionally, faster electronics have decreased the dead time by an order of magnitude.

The main difference compared to the Belle detector is the PID system. Instead of the ACC the Belle 2 detector will be equipped with a Time of Propagation Counter detector in the barrel and an Aerogel Ring Imaging Cherenkov detector in the endcaps. This system provides less material in the detector and increases the separation power for identifying the particles. The ECL and KLM technology has stayed unchanged, except for small modifications in the readout electronics and the replacement of the degraded CsI crystals.

Chapter 4

Angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$

In this chapter the angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ will be presented. The section is structured as follows: firstly the trigger and selection requirements are presented, followed by the peaking background studies. Next, a multivariate classifier is discussed and the methods of extracting the observables are introduced. The section ends with the estimation of systematics uncertainties and the report of the results. The results of this analysis have been published in [69].

4.1 Trigger requirements

As described in Sec. 3.2.6, LHCb uses a complicated trigger system in order to save as many interesting physics events as possible. The $B^0 \to K^{*0}\mu^+\mu^-$ candidates used in this analysis are required to pass the trigger requirements listed in Table 4.1. The HLT 2 lines used here focus on finding a multi-prong decay vertex that is detached from the primary vertex. At all stages the offline-candidates are required to be triggered on signal (`TOS`), i.e. the trigger decision is due solely to the presence of the candidate in the event and not the underlying event. A brief study was performed but no significant advantage was found by allowing candidates triggered independently of the signal (`TIS`), or extra trigger lines in the updated selection. The trigger lines are taken with the logical `or`. This means that a candidate to be considered has to pass at least one of the listed HLT2 lines.

4.2 Stripping and pre-selection

Before the events are made available to be analysed, they have to pass the so-called stripping selection. The stripping selection is made centrally by the LHCb experiment around twice a year. The main reason for this is that the LHCb data is stored on tape, so analysing it would be very time-consuming. Candidates satisfying the stripping requirements are kept on storage disks and are made available to physicists. These requirements are outlined in Table 4.2. For the 2012 (2011) data, this corresponds to the stripping version `20` (`20r1`) [70] and reconstruction version `Reco14` [71]. Candidates that have passed the

Table 4.1: Triggers required for the $B^0 \to K^{*0}\mu^+\mu^-$ candidate events.

Stage	Triggers
L0	`L0Muon`
HLT 1	`Hlt1TrackAllL0` or `Hlt1TrackMuon`
HLT 2	`Hlt2Topo[2,3,4]BodyBBDT`, `Hlt2TopoMu[2,3,4]BodyBBDT`, `Hlt2SingleMuon` or `Hlt2DiMuonDetached`

Table 4.2: Stripping selection criteria in `B2XMuMu` for Stripping `20` and Stripping `20r1`.

Candidate	Selection
B meson	IP $\chi^2 < 16$ (best PV)
B meson	$4600\,\text{MeV}/c^2 < M < 7000\,\text{MeV}/c^2$
B meson	Direction Angle (DIRA) $< 14\,\text{mrad}$
B meson	flight distance $\chi^2 > 121$
B meson	vertex $\chi^2/\text{ndf} < 8$
$\mu^+\mu^-$	$m(\mu^+\mu^-) < 7100\,\text{MeV}/c^2$
$\mu^+\mu^-$	vertex $\chi^2/\text{ndf} < 9$
K^{*0}	$m(K^+\pi^-) < 6200\,\text{MeV}/c^2$
K^{*0}	vertex $\chi^2/\text{ndf} < 9$
K^{*0}	flight distance $\chi^2 > 9$
tracks	ghost Prob < 0.4
tracks	min IP $\chi^2 > 9$
muon	`IsMuon`[1]
muon	$\text{DLL}_{\mu\pi} > -3$[2]
GEC	SPD Mult. < 600

stripping line are furthermore required to pass the same additional criteria detailed explicitly in Table 4.3[3].

Finally, additional requirements are made on the invariant mass $m_{K\pi}$ of the $K\pi$ system. For this analysis, a window of $\pm 100\,\text{MeV}/c^2$ around the nominal mass of the $K^*(892)^0$ state is employed. That is:

$$795.9\ \text{MeV}/c^2 < m_{K\pi} < 995.9\ \text{MeV}/c^2. \tag{4.1}$$

This requirement was choosen to be the same as the one use in the Belle $B^0 \to J/\psi\, K^{*0}$

[3]In the Table 4.3 the `IsMuon` requirement is defined as the track that has to have hits in at least three Muon Stations. The DLL_{XY} is defined as the ratio of probabilities between the X particle hypothesis and Y particle mass hypothesis. If one writes DLL_X this means that the second particle was assumed to be a pion.

Table 4.3: Pre-selection requirements applied to stripped candidates. In this table only: θ is the opening angle from the beam; and θ_{pair} is the opening angle between two track pairs. Variables $< X >$, $< Y >$, $< Z >$ denote the mean primary vertex position.

Candidates	Selection
Track	$0 < \theta < 400$ mrad
Track Pairs	$\theta_{pair} > 1$ mrad
$\mu^+\mu^-$	`IsMuon` True
K	`hasRich`[4] True
K	$\mathrm{DLL}_{K\pi} >$ -5
π	`hasRich` True
π	$\mathrm{DLL}_{K\pi} < 25$
PV	$\lvert X - < X > \rvert < 5\,\mathrm{mm}$
PV	$\lvert Y - < Y > \rvert < 5\,\mathrm{mm}$
PV	$\lvert Z - < Z > \rvert < 200\,\mathrm{mm}$

Table 4.4: Estimated yields, and percentage relative to estimated signal yield, of peaking background events before and after the vetoes detailed in Sec. 4.3 (all for events inside the $m_{K\pi}$ window). The dominant uncertainty contributing to these numbers is in $\sigma_{b\bar{b}}$ and the estimate of $\mathrm{BR}(\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-)$. The last column is the relevant section of this dissertation.

	after preselection, before vetoes		after vetoes and selection		
Channel	Estimated events	% signal	Estimated events	% signal	§
$\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$	$(1.0 \pm 0.5) \times 10^3$	19 ± 8	51 ± 25	1.0 ± 0.4	4.3.2
$\Lambda_b^0 \to pK^-\mu^+\mu^-$	$(1.0 \pm 0.5) \times 10^2$	1.9 ± 0.8	5.7 ± 2.8	0.11 ± 0.05	4.3.2
$B^+ \to K^+\mu^+\mu^-$	28 ± 7	0.55 ± 0.06	1.6 ± 0.5	0.031 ± 0.006	4.3.5
$B_s^0 \to \phi\mu^+\mu^-$	$(3.2 \pm 1.3) \times 10^2$	6.2 ± 2.1	17 ± 7	0.33 ± 0.12	4.3.4
signal swaps	$(3.6 \pm 0.9) \times 10^2$	6.9 ± 0.6	33 ± 9	0.64 ± 0.06	4.3.3
$B^0 \to J/\psi K^{*0}$ swaps	$(1.3 \pm 0.4) \times 10^2$	2.6 ± 0.4	2.7 ± 2.8	0.05 ± 0.05	4.3.3
$B^0 \to J/\psi K^{*0}$	70 ± 22	1.35 ± 0.28	59 ± 19	1.14 ± 0.26	4.3.1

analysis.

4.3 Vetoes against specific peaking backgrounds

Several sources of specific backgrounds have been considered. The estimated yields of each background (in $3\,\mathrm{fb}^{-1}$ data sample) are summarised in Table 4.4. The selection and trigger requirements bias the angular distribution. This obviously would have an impact on the final result, hence we will correct for it at a later stage of the analysis (see Sec. 4.5).

4.3.1 Charmonia resonances

The charmonia resonances are removed in this analysis by discarding the specific q^2 regions. The bin $q^2 \in [8.0, 11.0]\,\mathrm{GeV}^2/c^4$ (containing the $B^0 \to J/\psi K^{*0}$ resonant decay mode), the

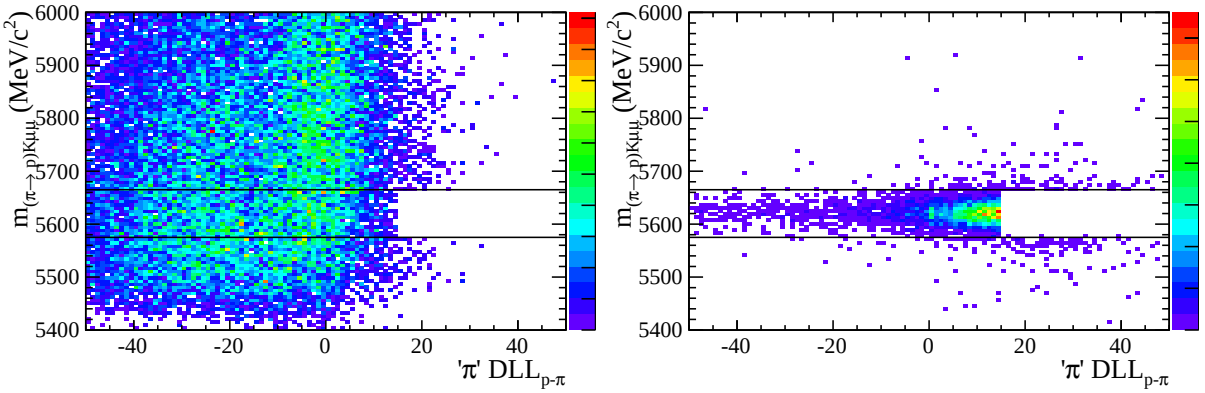


Figure 4.1: Scatter plots of Λ_b^0 candidate mass ($m_{(\pi\to p)K\mu\mu}$) and pion candidate $\mathrm{DLL}_{p\pi}$ for simulated physics signal (left) and simulated with flat phase space distribution $\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$ events (right).

bin $q^2 \in [12.5, 15.0]\,\mathrm{GeV}^2/c^4$ (containing the $B^0 \to \psi(2S)K^{*0}$ resonant decay mode) and $q^2 \in [0.98, 1.1]\,\mathrm{GeV}^2/c^4$ are removed from the data samples before training the multivariate classifier.

4.3.2 $\Lambda_b^0 \to pK^-\mu^+\mu^-$ background

The peaking background from the $\Lambda_b^0 \to pK^-\mu^+\mu^-$ (inclusive) decays arise when the p is reconstructed as either of the hadron candidates. The dominant contribution to the $pK^-\mu^+\mu^-$ final state is expected to be from the $\Lambda^*(1520)^0 \to pK^-$ resonance. The veto for those decays is done by remaking alternative mass hypotheses and removing events in the regions of mass and $\mathrm{DLL}_{p\pi}$ variable, as detailed below.

The simplest case is when p is reconstructed as the "π^+" candidate. In this case, changing the hypothesis of the π track from the pion mass to the proton mass and remaking the invariant mass, the value of $m_{(\pi\to p)K\mu\mu}$, should be consistent with the mass of Λ_b^0. The contribution from these events is removed by vetoing events with $m_{(\pi\to p)K\mu\mu}$ around the nominal Λ_b^0 mass with a proton-like $\mathrm{DLL}_{p\pi}$ discriminator value. The candidates are removed if:

$$(5575 < m_{(\pi\to p)K\mu\mu} < 5665)\,\mathrm{MeV}/c^2 \tag{4.2}$$

$$\mathrm{DLL}_{p\pi}(\pi) > 0. \tag{4.3}$$

Fig. 4.1 shows the $m_{(\pi\to p)K\mu\mu}$ and $\mathrm{DLL}_{p\pi}(\pi)$ plane for simulated signal and phase space $\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$.

The second way the $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decays can contribute is when p is reconstructed as the "K^+" candidate and K^- is reconstructed as "π^-". In this case, both hadron track mass hypotheses need to be changed. The resulting invariant mass, $m_{(K\to p)(\pi\to K)\mu\mu}$, should be consistent with that of Λ_b^0. This contribution is removed by a similar requirement on

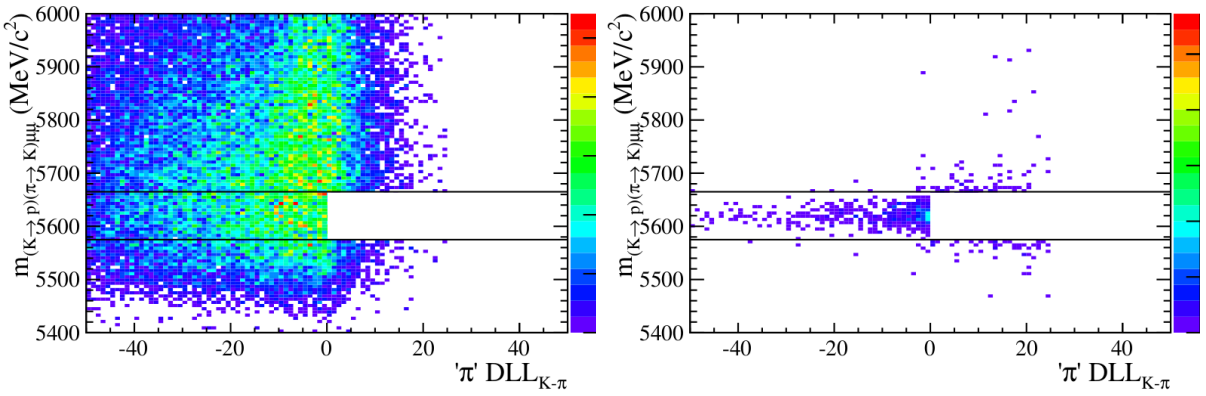


Figure 4.2: Scatter plots of Λ_b^0 candidate mass ($m_{(K\to p)(\pi\to K)\mu\mu}$) vs pion candidate DLL$_{K\pi}$ for simulated physics signal (left) and simulated phase space $\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$ events (right).

$m_{(K\to p)(\pi\to K)\mu\mu}$ and by the fact that the pion candidate has a kaon-like DLL$_{K\pi}$ value. That is:

$$(5575 < m_{(K\to p)(\pi\to K)\mu\mu} < 5665)\,\text{MeV}/c^2 \tag{4.4}$$

$$\text{DLL}_{K\pi}(\pi) > 0. \tag{4.5}$$

Fig. 4.2 shows the $m_{(K\to p)(\pi\to K)\mu\mu}$ and DLL$_{K\pi}(\pi)$ plane for simulated signal and phase space $\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$.

4.3.3 Misidentified $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi\, K^{*0}$ decay modes

Another background that has to be considered is the misidentification of $B^0 \to J/\psi\, K^{*0}$. Such misidentified candidates can be separated from the correctly identified signal candidates if the π^- and K^+ mass assignments are swapped, and the $K\pi$ invariant mass ($m_{K\leftrightarrow\pi}$) should then be consistent with that of K^{*0}. Candidates are rejected if the kaon and pion satisfy:

$$\text{DLL}_{K\pi}(K) + 10 < \text{DLL}_{K\pi}(\pi) \tag{4.6}$$

$$(795 < m_{K\leftrightarrow\pi} < 995)\,\text{MeV}/c^2. \tag{4.7}$$

In addition, a requirement is imposed on the difference in the hadrons' DLL$_{K\pi}$ variables.

$$\text{DLL}_{K\pi}(K) - \text{DLL}_{K\pi}(\pi) > 0. \tag{4.8}$$

This removes the remaining hadron PID swaps. The requirement is usually referred to as "diagonal" forming a diagonal exclusion in the DLL$_{K\pi}(K)$-DLL$_{K\pi}(\pi)$ plane.

Another way that the $B^0 \to J/\psi\, K^{*0}$ decay contributes to a peaking background, is when π^- (K^+) is misidentified as μ^- (μ^+) and μ^- (μ^+) is misidentified as π^- (K^+), which is the so-called "double-swap". The candidates where π^- and μ^- are misidentified can be

separated from signal candidates, if the π^- and μ^- mass assignments are swapped, and the $\pi\mu$ invariant mass $m_{(\pi\to\mu)\mu}$ should be consistent with that of J/ψ. The equivalent quantity for the candidates where K^+ and μ^+ are misidentified, the $K\mu$ invariant mass $m_{(K\to\mu)\mu}$, is calculated by swapping the K^+ and μ^+ mass assignments. The candidates are rejected if

$$(3036 < m_{(\pi\to\mu)\mu} < 3156)\ \text{MeV}/c^2 \tag{4.9}$$

and the pion satisfies either the `IsMuon` criteria or $\text{DLL}_{\mu\pi} > 5.0$; or if

$$(3036 < m_{(K\to\mu)\mu} < 3156)\ \text{MeV}/c^2 \tag{4.10}$$

and the kaon satisfies either the `IsMuon` criteria or $\text{DLL}_{\mu\pi} > 5.0$.

4.3.4 $B_s^0 \to \phi\mu^+\mu^-$ decay

The decay $B_s^0 \to \phi(\to K^+K^-)\mu^+\mu^-$ (hereinafter denoted as $B_s^0 \to \phi\mu^+\mu^-$) contributes to the peaking background if K^- is misidentified as π^-. If the reconstructed π^- is assigned the nominal mass of K^-, the $K\pi$ invariant mass ($m_{(\pi\to K)K}$) should then be consistent with that of ϕ. Similarly, the $K\pi\mu\mu$ invariant mass ($m_{(\pi\to K)K\mu\mu}$) should then be consistent with that of B_s^0. The candidates are rejected if

$$(5321 < m_{(\pi\to K)K\mu\mu} < 5411)\,\text{MeV}/c^2 \tag{4.11}$$

and either:

$$(1010 < m_{(\pi\to K)K} < 1030)\,\text{MeV}/c^2 \tag{4.12}$$

and the pion satisfies $\text{DLL}_{K\pi} > -10$, or

$$(1030 < m_{(\pi\to K)K} < 1075)\,\text{MeV}/c^2 \tag{4.13}$$

and the pion satisfies $\text{DLL}_{K\pi} > 10$. These two vetoes are defined in separate regions of $m_{(\pi\to K)K}$ to reduce the number of $B^0 \to K^{*0}\mu^+\mu^-$ decays that are removed.

4.3.5 $B^+ \to K^+\mu^+\mu^-$ decay with a slow π^-

A background contribution can be formed if π^- from elsewhere in the event is added to a genuine $B^+ \to K^+\mu^+\mu^-$ decay to form a four-track final state. Given that the $B^+ \to K^+\mu^+\mu^-$ decays accumulate at the nominal B^0 mass, this background can only contribute to the upper $m_{K\pi\mu\mu}$ mass sideband. These candidates are typically asymmetric in $\cos\theta_K$, which is due to a momentum imbalance between the kaon and the pion. This modifies the $\cos\theta_K$ distribution of the candidates in the upper mass sideband, potentially causing a mis-measurement of the angular observables. The candidates are removed by requiring $m_{K\pi\mu\mu} > 5380\,\text{MeV}/c^2$ and $(5220 < m_{K\mu\mu} < 5340)\,\text{MeV}/c^2$, where $m_{K\mu\mu}$ is the $K\mu\mu$ invariant mass.

4.3.6 Other sources of peaking background

The decay $B^+ \to K^{*+}\mu^+\mu^-$ contributes to the background when π^- from elsewhere in the event is added to the $K^{*+} \to K^+\pi^0$ decay (where π^0 is not reconstructed) to form a four-track final state. These candidates cannot be isolated from the $B^0 \to K^{*0}\mu^+\mu^-$ decays by imposing requirements on invariant mass distributions or DLL quantities. Therefore, no veto criteria are imposed to remove the $B^+ \to K^{*+}\mu^+\mu^-$ backgrounds. The contribution of this background is expected at the level of 1.5% of the signal yield accross the full $m_{K\pi\mu\mu}$ range and 0.5% within the signal mass window.

The $B^0 \to \rho^0(\to \pi^+\pi^-)\mu^+\mu^-$ (hereinafter denoted as $B^0 \to \rho^0\mu^+\mu^-$) decay contributes to the peaking background if π^+ is misidentified as K^+. The decay proceeds by the same process as the $B^0 \to K^{*0}\mu^+\mu^-$ decay; however, it is further suppressed by the ratio of CKM factors

$$\left|\frac{V_{cd}}{V_{cs}}\right|^2 \approx 0.05 \text{ or } \left|\frac{V_{td}}{V_{ts}}\right|^2 \approx 0.05 \tag{4.14}$$

for the resonant and non-resonant decays, respectively. Taking the rate of the $\pi \to K$ mis-identification as 10% [72], the contribution of $B^0 \to \rho^0\mu^+\mu^-$ decays is calculated to be $\ll 1\%$ of the signal yield, and is therefore neglected.

4.4 Multivariate classifier

The combinatorial background is reduced using the Toolkit for Multivariate Analysis (TMVA) software package [73]. The presented selection is based on a boosted decision tree (BDT) [74] classifier using the AdaBoost algorithm [75]. The Multivariate Analysis (MVA) needs control samples to be able to distinguish between the signal events and those from the combinatorial background. The training of the BDT classifier is done using:

- The *sPloted* [76] $B^0 \to J/\psi K^{*0}$ candidates as proxy to the signal events. A description of the selection of these events is given in Sec. 4.4.1.
- The upper mass sideband of the $B^0 \to K^{*0}\mu^+\mu^-$ signal candidate invariant mass ($m_{K\pi\mu\mu}$) is used as a sample of combinatorial background. That is the window $m_{K\pi\mu\mu} \in [5350, 7000]\,\text{MeV}/c^2$. The lower mass sideband is not used because of the presence of partially reconstructed events.

4.4.1 Use of *sPlot* technique on $B^0 \to J/\psi K^{*0}$ candidates

The *sPlot* technique [76] is employed to weight the stripped $B^0 \to J/\psi K^{*0}$ candidates. This procedure statistically removes the background contribution so that the $B^0 \to J/\psi K^{*0}$ data sample can be considered "pure". The procedure is used in two areas:

1. For the $B^0 \to J/\psi K^{*0}$ candidates that pass the preselection and peaking background vetoes before the training of the BDT.

2. For the $B^0 \to J/\psi K^{*0}$ candidates that pass the full selection to derive weights to correct the simulation (see Sec. 4.5.1).

Initially, a fit to the un-weighted $B^0 \to J/\psi K^{*0}$ candidates is performed. The model used is a double Crystal Ball for signal and a first order polynomial for the combinatorial background component.

4.4.2 Input variables for the BDT classifier

The input variables to the BDT classifiers that have been considered are the ones that exhibit good data-simulation agreement and good discriminating power. The following section will discuss these variables. The variables used can be split into kinematic/topological based, PID based and isolation based. The kinematic/topological variables are (5 in total):

- the B^0 lifetime,
- the $K\pi\mu\mu$ vertex quality (χ^2),
- the momentum of the B^0 meson,
- the transverse momentum of the B^0 meson,
- the cosine of the angle between the momentum and the direction , of flight from the primary to decay vertices (DIRA).

The PID variables are:

- the $\mathrm{DLL}_{K\pi}$ for the pion and the kaon,
- the $\mathrm{DLL}_{\mu\pi}$ for the muon tracks.

Finally, two types of isolation variables are used, one for the hadrons and one for the muons, which will be described below.

The decision to use the set of variables discussed above was reached by testing many configurations of training variables, with the five kinematic/topological variables as the baseline, and by selecting the most performant BDT as given by the Receiver Operating Characteristic (ROC) curve shown in Fig. 4.3. It must be noted here that although the artificial neutral network based PID (`ProbNN`) outperforms the standard DLLs, the decision made to use the DLLs is due to their better modelling in the simulation.

4.4.3 Hadron Isolation

Hadron isolation variables are based on the isolation used in the $B_s^0 \to \mu^+\mu^-$ analysis performed by the LHCb experiment [10]. The isolation is defined as the number of extra tracks (i.e. excluding tracks that form a $B^0 \to K^{*0}\mu^+\mu^-$ candidate), that can form a vertex with a hadron track. For each of the tracks in the event (excluding tracks forming a $B^0 \to K^{*0}\mu^+\mu^-$ candidate) the following quantities (which are depicted in Fig. 4.4) are taken into account:

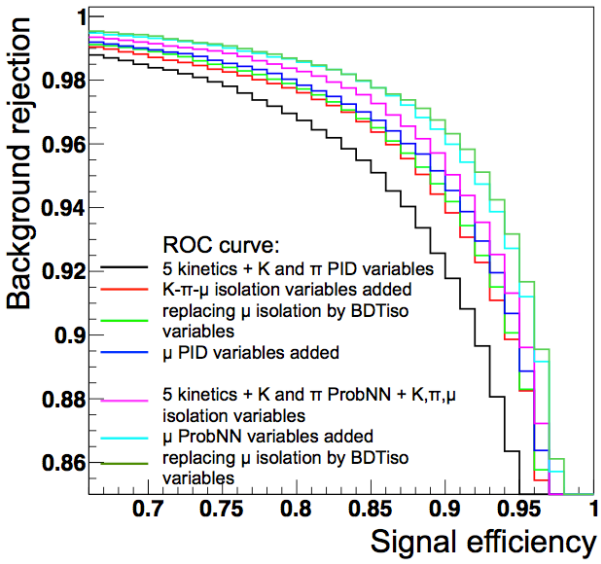


Figure 4.3: ROC curves of BDTs using differents input variables. In black the BDT was trained with the 5 kinematic variables of the B^0 meson and the PID information for the kaon and the pion. In red (light green) the isolation variables (the so-called BDTiso) of the final state particle are added to the training. In dark blue the PID information of the muons is added. In pink the training uses the 5 kinematic variables of the B^0 meson and the ProbNN information for the kaon and the pion and the isolations variables In cyan the ProbNN information of the muons are added to the BDT training. In green the muon isolation variables are replaced by the BDTiso isolation variables.

- Minimum distance between the **track** and the PV (**pvdist**),
- Minimum distance between the **track** and the $B^0 \to K^{*0}\mu^+\mu^-$ vertex (**svdist**),
- DOCA between the **hadron** and the **track** (DOCA),
- Track IP significance,
- Angle between the **hadron** and the **track** (β),
- The quantity

$$f_c = \frac{|\overrightarrow{P}_h + \overrightarrow{P}_{trk}|\alpha^{h+trk,PV}}{|\overrightarrow{P}_h + \overrightarrow{P}_{trk}|\alpha^{h+trk,PV} + P_{\mathrm{T},h} + P_{\mathrm{T},trk}}, \tag{4.15}$$

  where $\alpha^{h+trk,PV}$ is the angle between the **hadron** and the **track** candidate, $P_{\mathrm{T},h}$ and $P_{\mathrm{T},trk}$ are the transverse momentum with respect to the beam line.

The hadron isolation variables are built by counting how many extra tracks satisfy the following conditions:

- pvdist $\in [0.5, 40]$ mm,

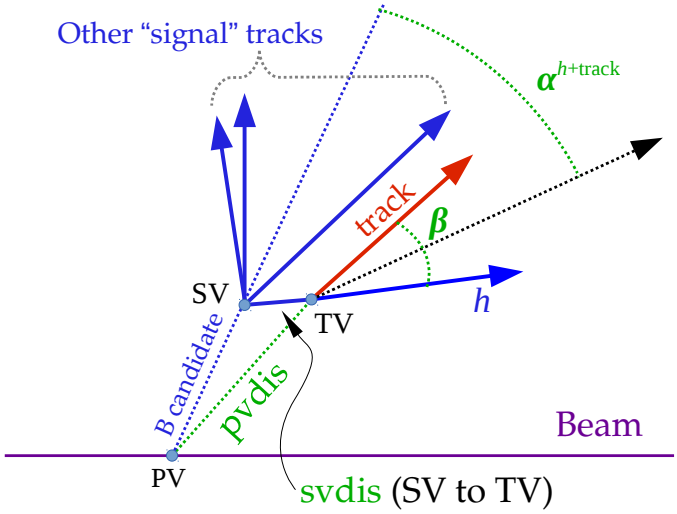


Figure 4.4: A diagram showing the variables used in the hadron isolation variable. Description is given in Sec. 4.4.3.

- svdist $\in [-0.15, 30]$ mm,
- Doca < 0.13 mm,
- Track IP significance > 3,
- $\beta < 0.27$,
- $f_c < 0.6$.

This number is stored for each $B^0 \to K^{*0}\mu^+\mu^-$ candidate and is used as a training variable for the BDT.

4.4.4 Muon Isolation

Muon isolation variables are also based on isolation used in the $B_s^0 \to \mu^+\mu^-$ analysis [10] but additionally employ a multivariate approach.

For each track in the event, the same quantities as those described in Sec. 4.4.3, replacing the hadron tracks by the muons ones, are fed into a BDT. The following variables are also added to the BDT:

- Track IP,
- Track p_{T},
- Track χ^2/ndof.

For training the MVA we use a simulated signal sample with a phase as well as an inclusive dimuon background sample corresponding to $12\,\mathrm{pb}^{-1}$ of integrated luminosity. The results of the training can be seen in Fig. 4.5. The best performance was obtained using a BDT classifier, which is henceforth referred to as BDTiso.

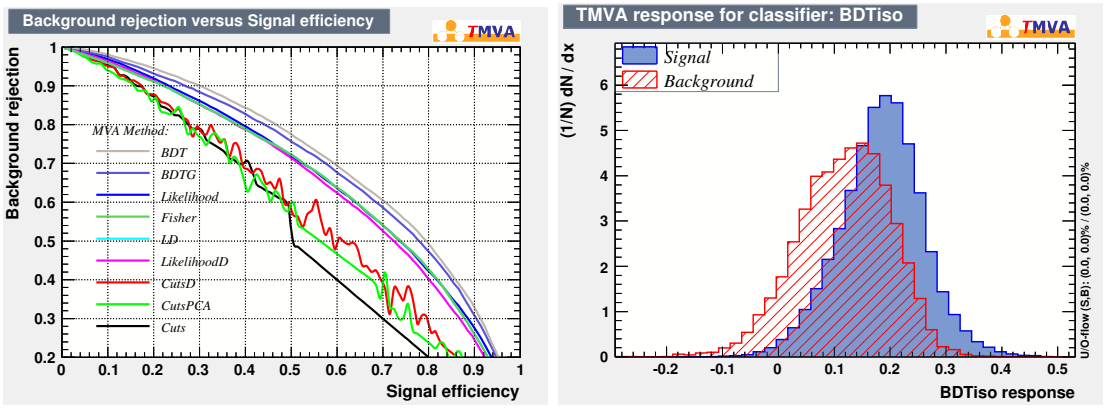


Figure 4.5: Comparison of ROC curves from TMVA training for muon isolation. BDT outperforms other classifiers.

BDTiso is defined for each track in the event (excluding the signal candidate tracks). Since the BDTiso classifier is trained on a track basis (i.e. each track in an event has a BDTiso response), the event based isolation is constructed as the number of tracks that pass a given BDTiso requirement. The BDTiso value is scanned in order to find the most discriminating requirement.

4.4.5 k-Folding of the data sample

For the training of the classifiers a k-Folding technique was used. This technique has been introduced to LHCb by the author of this dissertation and first used in this analysis. The technique employs the full set of the 2011 and 2012 data in an unbiased way and has been first adapted to be used in Particle Physics by the author of this dissertation. To cancel any influence of varying running conditions (e.g. different TCK, magnet polarities) the ordering of the data set was randomised and divided in $k = 10$ samples of equal size (both for the full background sample, $B = \bigcup_i B^i$, and for the full *sWeighted* $B^0 \to J/\psi\, K^{*0}$ sample used as signal proxy, $S = \bigcup_i S^i$).

Ten classifiers are trained[5], each using nine signal and nine background samples. More formally, the i^{th} classifier uses $\bigcup_{j,j\neq i} B^j$ and $\bigcup_{j,j\neq i} S^j$ training samples. The obtained i^{th} classifier is then applied to the i^{th} signal (S^i) and data (B^i) samples that were omitted from its training, which is illustrated in Fig. 4.6.

Such a training has several advantages compared to standard 50:50 splitting:

[5]Hence: 10-Folding

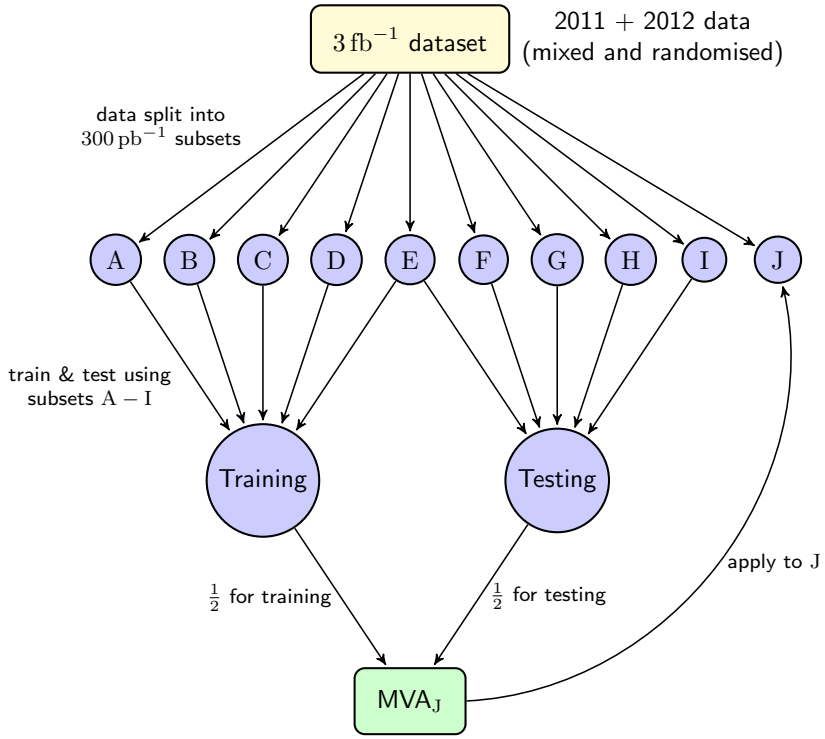


Figure 4.6: A schematic diagram of the 10-Folding technique. The arrows indicate the direction of the data "flow". The data is split randomly into 10 sub-datasets (A–J). The MVA for sample J ($\mathrm{MVA_J}$) is trained and tested on samples A–K, then this response is evaluated on sample J only. This way, 90% of the dataset is used for every BDT training in an unbiased way.

- The i^{th} and j^{th} classifiers have $\frac{8}{9}$ common events, which makes the classifiers response more similar reducing the systematic error.
- The training samples are increased in size as the 10-Folding technique allows the use of 90% of the data for training, compared to 50% in the standard case, which also leads to better optimisation of the classifier.

After training each fold has its own BDT response. All BDTs have similar responses, as expected/desired, and they are treated on data as a single classifier.

For simulation samples, which were not used for training, the following quantity is used as a BDT estimator:

$$\mathrm{BDT_{MC}} = \frac{\sum_{\mathrm{i}=1}^{10} \mathrm{BDT_{Fold\ i}}}{10}, \tag{4.16}$$

where $\mathrm{BDT_{MC}}$ is the response to a simulated event, and $\mathrm{BDT_{Fold\ i}}$ is the response to a simulated event using the $\mathrm{i^{th}}$ fold's classifier.

4.4.6 Optimization of the BDT selection

The BDT is optimised using samples orthogonal to those used in the subsequent analysis of the data in order to minimise potential statistical biases present in the analysis sample. For each BDT requirement, the signal yield in a window $m_{K\pi\mu\mu} \in [5230, 5330]$ MeV/c^2 is estimated by fitting $B^0 \to J/\psi\, K^{*0}$ events and scaling the yield obtained by the ratio of the total selection efficiency (apart from the BDT requirement) obtained from the simulation, and the ratio of branching fractions obtained from the PDG, between $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi\, K^{*0}$:

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

The background yield is estimated by fitting part of the lower and upper mass sideband regions of $B^0 \to K^{*0}\mu^+\mu^-$ events and extrapolating the yield into the signal region ($\mathrm{nBkg}^{\mathrm{BDT}}_{\mu\mu}$). The upper mass sideband is defined as $m_{K\pi\mu\mu} \in [5500, 7000]$ MeV/c^2 and the lower mass sideband is $m_{K\pi\mu\mu} \in [5000, 5180]$ MeV/c^2. Although the upper mass sideband does have a slight overlap with the final analysis sample, it helps stabilize the fit in the two regions. A double exponential is used for the fit to the background sidebands, and two Gaussian distributions with a common mean and common left side power law tail (double Crystal Ball) for the fits to $B^0 \to J/\psi\, K^{*0}$.

The figure of merit is defined as $\mathrm{nSig}^{\mathrm{BDT}}_{\mu\mu}/\sqrt{\mathrm{nSig}^{\mathrm{BDT}}_{\mu\mu} + \mathrm{nBkg}^{\mathrm{BDT}}_{\mu\mu}}$ was employed. The tightest requirement out of this cluster at 0.20 is chosen in order to keep background levels and therefore potential systematic effects under control. In addition, a tighter value is also preferable for the q^2 region $[1, 6]$ GeV/c^2, which has a lower signal over the background ratio. The optimal working point for the BDT at 0.20 gives a signal efficiency of 85% evaluated on $B^0 \to J/\psi\, K^{*0}$ candidates and a background rejection of 97% in the $B^0 \to K^{*0}\mu^+\mu^-$ upper mass sideband. Both efficiencies for the BDT classifier are given for data for which the full remaining selection was applied.

4.5 Acceptance parametrisation

To correct for the distortion caused by the reconstruction and selection of the signal decay, Monte Carlo simulated signal events are used. The acceptance can be parametrised using Legendre polynomials in four dimensions, q^2, $\cos\theta_l$, $\cos\theta_l$ and ϕ according to

$$\varepsilon(\cos\theta_\ell, \cos\theta_K, \phi, q^2) = \sum_{k,l,m,n} c_{k,l,m,n} P(\cos\theta_\ell, k) P(\cos\theta_K, l) P(\phi, m) P(q^2, n). \qquad (4.17)$$

Here $P(x, m)$ are Legendre polynomials in x of order m and $-1 \leqslant x \leqslant 1$[6]. The coefficicients $c_{k,l,m,n}$ are determined from a moment analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ flat phase space MC

[6]The variables ϕ and q^2 are rescaled to match this boundaries.

$$c_{k,l,m,n} = \frac{1}{N'} \sum_{i=1}^{N} w_i \left[\left(\frac{2k+1}{2} \right) \left(\frac{2l+1}{2} \right) \left(\frac{2m+1}{2} \right) \left(\frac{2n+1}{2} \right) \times P(\cos\theta_\ell, k) P(\cos\theta_K, l) P(\phi, m) P(q^2, n) \right]. \tag{4.18}$$

where N is the number of candidates in the MC sample, w_i is a per-candidate weight used to correct for the non-flat distribution of events in q^2 and the data-driven corrections for $p_{\mathrm{T}}(B^0)$, χ^2_{Vtx} and track multiplicity. The normalisation is given by

$$N' = \sum_{i=1}^{N} w_i. \tag{4.19}$$

The factors of $(2k+1)/2$ etc. arise from the orthonormality of the Legendre Polynomials:

$$\int_{-1}^{+1} P(x,m) P(x,m') \mathrm{d}x = \frac{2}{2m+1} \delta_{mm'} \ . \tag{4.20}$$

The acceptance is modeled using Legendre polynomials of the lowest order that give a good description of the angular acceptance effect. For q^2, Legendre polynomials of order five are used. For the decay angles polynomials of order four are used for $\cos\theta_l$, order five for $\cos\theta_K$ and order six for the angle ϕ.

This results in a total of 600 coefficients that are determined using a sample of 1.406 M reconstructed and selected Monte Carlo simulated signal events. The resulting one-dimensional projections of the efficiency on q^2 and the three decay angles are given in Fig. 4.7. A good agreement is seen for the one-dimensional projections.

4.5.1 Agreement between data and simulation

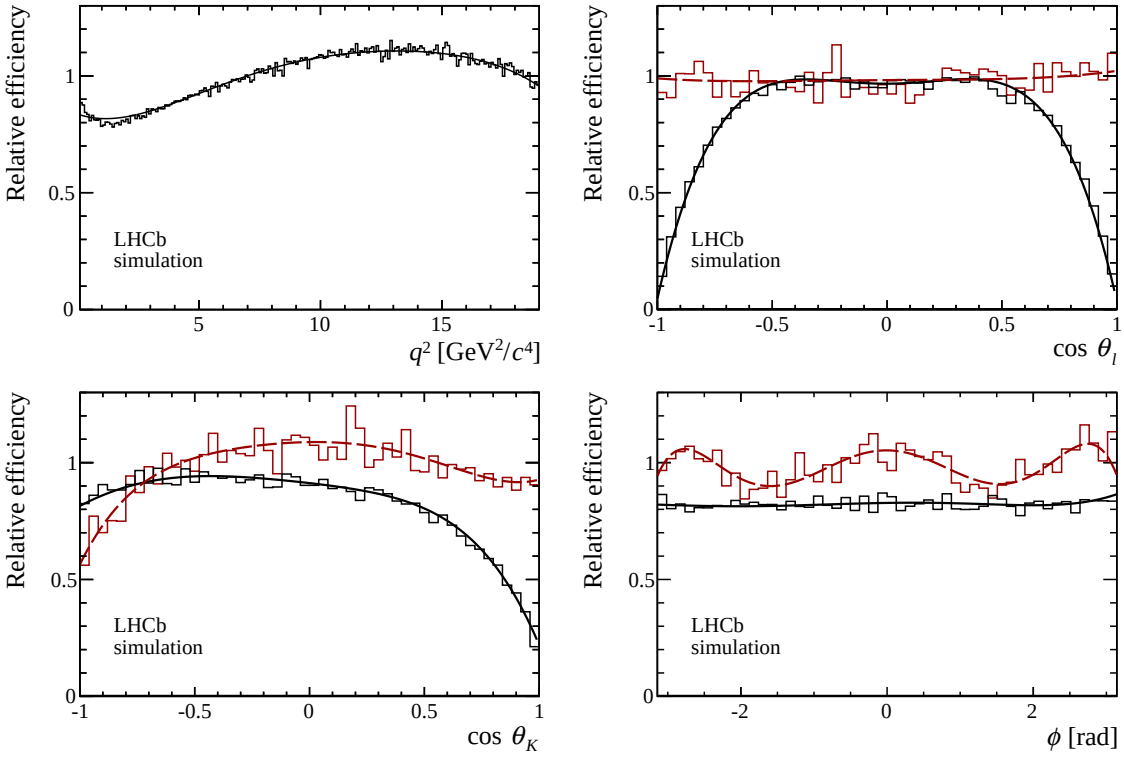
A good agreement between data and simulated events is important for all analyses of the decay $B^0 \to K^{*0}\mu^+\mu^-$ to allow the determination of acceptance effects from simulated events. Data driven corrections are employed in two steps to improve the agreement of simulation with data. The PID variables are "re-sampled", and the remaining differences are corrected by explicitly weighting distributions. These corrections are applied to all simulated samples.

4.5.2 Re-sampling the particle identification variables

The selection includes requirements on the $\mathrm{DLL}_{K\pi}$, as well as the $\mathrm{DLL}_{\mu\pi}$ and $\mathrm{DLL}_{p\pi}$ variables. In order to achieve a good agreement between data and MC, these distributions are "re-sampled".

All simulation samples are re-sampled. This procedure involves looping over all PID variables in all simulated events. For a given simulated track in the event the η, `nTracks`, and p_{T} of the track are matched to a specific calibration histogram. This histogram is then

Figure 4.7: One-dimensional projections of the four-dimensional efficiency parametrisation on q^2, $\cos\theta_l$, $\cos\theta_K$, and ϕ. The solid line for the angular distribution corresponds to the $qsq \in [0.1, 0.98]$ GeV$^2/c^4$, while the dashed red one to the $q^2 \in [18.0, 19.0]$ GeV$^2/c^4$



used as a PDF to draw a new PID value. This sampled value is stored as a replacement for the original PID variable constructed from simulation.

This procedure has the effect of improving the agreement of the PID distributions between simulation and data. However, it only preserves the correlations of the sampled variables (i.e. η, `nTracks`, and p_{T}) and breaks any others. This may have an irreducible effect on the data-simulation agreement of the BDT response. Fig. 4.8 shows the improvement before (red) and after (blue) this re-sampling procedure. The comparison is made to selected $B^0 \to J/\psi K^{*0}$ data candidates that are weighted using the *sPlot* procedure described in Sec. 4.4.1.

4.5.3 Weighting kinematic distributions of the parent B candidate

An event weighting for all MC samples is derived from the distributions of three kinematic variables that show differences between data and simulation. The event weights are calculated from comparisons of selected $B^0 \to J/\psi K^{*0}$, with offline selected MC2012. The

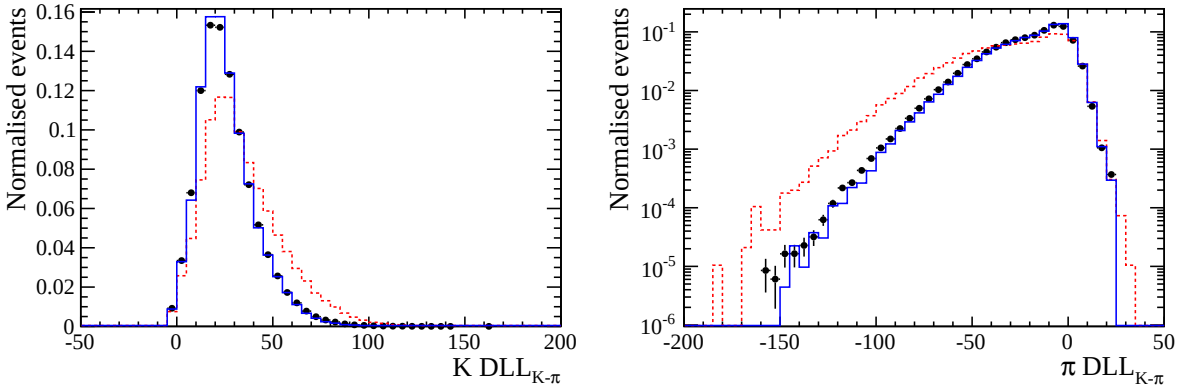


Figure 4.8: Validation plots for the re-sampling procedure. Original simulated PID (red) is re-sampled (blue) and compared to weighted data (black points) for $B^0 \to J/\psi K^{*0}$ events. Distributions shown in linear scale (left) and logarithmic scale (right). The fraction of missing events amounts to $\mathcal{O}(10\times^{-4})$ of the $B^0 \to J/\psi K^{*0}$ data events.

offline selected $B^0 \to J/\psi K^{*0}$ data candidates are *sWeighted* Agreement as described in Sec. 4.4.1.

The distributions from which the weights are derived are: the detector track occupancy (`nTracks`); the parent B candidate's p_{T} and the quality of the $K\pi\mu\mu$ vertex (χ^2/ndf). These distributions are chosen carefully so that the effect of the selection which might bias the distributions is minimal. The weights are derived sequentially because of potential small correlations between variables, with the $(N-1)$ weight being applied before deriving the N^{th} weight. That is, an event weight is derived for `nTracks` and is then applied before deriving the weight for p_{T}, the product of the event weights for `nTracks` and p_{T} is applied before deriving the weight for χ^2/ndf.

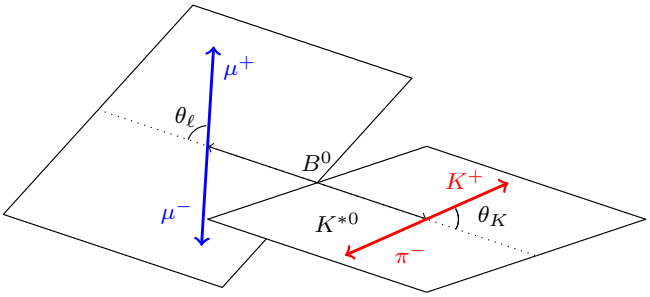
4.6 Methods of extracting angular observables

4.6.1 The angular basis

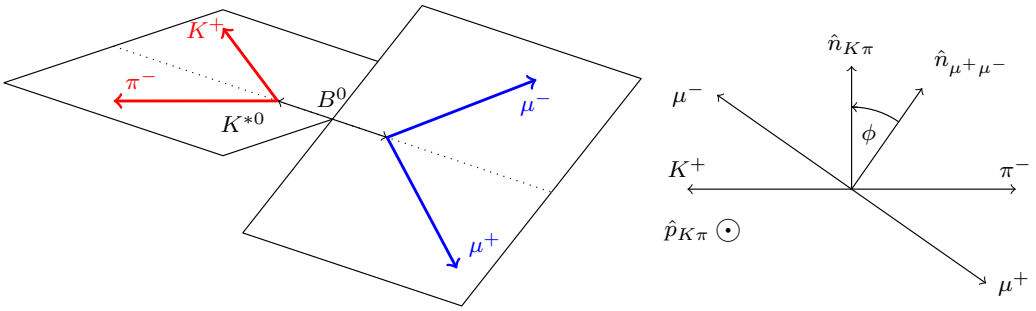
The decay angles are defined differently by experimentalists and theorists. Details on the differences between the convention chosen by the majority of theory publications [12, 77] and the experimental convention are given in Ref. [78].

The angular basis used in this paper is illustrated in Fig. 4.9. The angle θ_ℓ is defined as the angle between the direction of the μ^+ (μ^-) in the dimuon rest frame and the direction of the dimuon in the B^0 ($\overline{B}^0$) rest frame. The angle θ_K is defined as the angle between the direction of the kaon in the K^{*0} ($\overline{K}^{*0}$) rest frame and the direction of the K^{*0} ($\overline{K}^{*0}$) in the B^0 ($\overline{B}^0$) rest frame. The angle ϕ is the angle between the plane containing μ^+ and μ^- and the plane containing the kaon and the pion from K^{*0}. Explicitly, $\cos\theta_\ell$ and $\cos\theta_K$ are defined as

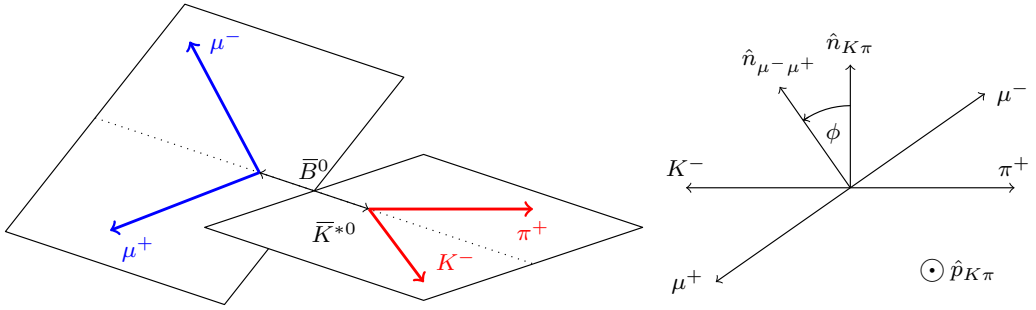
$$\cos\theta_\ell = \left(\hat{p}^{(\mu^+\mu^-)}_{\mu^+}\right) \cdot \left(\hat{p}^{(B^0)}_{\mu^+\mu^-}\right) = \left(\hat{p}^{(\mu^+\mu^-)}_{\mu^+}\right) \cdot \left(-\hat{p}^{(\mu^+\mu^-)}_{B^0}\right) \ , \tag{4.21}$$



(a) θ_K and θ_ℓ definitions for the B^0 decay



(b) ϕ definition for the B^0 decay



(c) ϕ definition for the $\overline{B}^0$ decay

Figure 4.9: Graphical representation of the angular basis used for the $B^0 \to K^{*0}\mu^+\mu^-$ and $\overline{B}^0 \to \overline{K}^{*0}\mu^+\mu^-$ decays in this paper. The notation $\hat{n}_{ab}$ is used to represent the normal to the plane containing the particles a and b in the B^0 (or $\overline{B}^0$) rest frame. An explicit description of the angular basis is given in the text.

$$\cos\theta_K = \left(\hat{p}_{K^+}^{(K^{*0})}\right) \cdot \left(\hat{p}_{K^{*0}}^{(B^0)}\right) = \left(\hat{p}_{K^+}^{(K^{*0})}\right) \cdot \left(-\hat{p}_{B^0}^{(K^{*0})}\right) \tag{4.22}$$

for the B^0 and

$$\cos\theta_\ell = \left(\hat{p}_{\mu^-}^{(\mu^+\mu^-)}\right) \cdot \left(\hat{p}_{\mu^+\mu^-}^{(\bar{B}^0)}\right) = \left(\hat{p}_{\mu^-}^{(\mu^+\mu^-)}\right) \cdot \left(-\hat{p}_{\bar{B}^0}^{(\mu^+\mu^-)}\right) \ , \tag{4.23}$$

$$\cos\theta_K = \left(\hat{p}_{K^-}^{(K^{*0})}\right) \cdot \left(\hat{p}_{K^{*0}}^{(\bar{B}^0)}\right) = \left(\hat{p}_{K^-}^{(K^{*0})}\right) \cdot \left(-\hat{p}_{\bar{B}^0}^{(K^{*0})}\right) \tag{4.24}$$

for the $\overline{B}^0$ decay. The definition of the angle ϕ is given by

$$\cos\phi = \left(\hat{p}_{\mu^+}^{(B^0)} \times \hat{p}_{\mu^-}^{(B^0)}\right) \cdot \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right) \ , \tag{4.25}$$

$$\sin\phi = \left[\left(\hat{p}_{\mu^+}^{(B^0)} \times \hat{p}_{\mu^-}^{(B^0)}\right) \times \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right)\right] \cdot \hat{p}_{K^{*0}}^{(B^0)} \tag{4.26}$$

for the B^0 and

$$\cos\phi = \left(\hat{p}_{\mu^-}^{(\bar{B}^0)} \times \hat{p}_{\mu^+}^{(\bar{B}^0)}\right) \cdot \left(\hat{p}_{K^-}^{(\bar{B}^0)} \times \hat{p}_{\pi^+}^{(\bar{B}^0)}\right) \ , \tag{4.27}$$

$$\sin\phi = -\left[\left(\hat{p}_{\mu^-}^{(\bar{B}^0)} \times \hat{p}_{\mu^+}^{(\bar{B}^0)}\right) \times \left(\hat{p}_{K^-}^{(\bar{B}^0)} \times \hat{p}_{\pi^+}^{(\bar{B}^0)}\right)\right] \cdot \hat{p}_{\overline{K}^{*0}}^{(\bar{B}^0)} \tag{4.28}$$

for the $\overline{B}^0$ decay. $\hat{p}_X^{(Y)}$ are unit vectors describing the direction of the particle X in the rest frame of the system Y. In every case the particle momenta are first boosted to the B^0 (or $\overline{B}^0$) rest frame. In this basis, the angular definition for the $\overline{B}^0$ decay is a CP transformation of that for the B^0 decay.

4.6.2 The differential decay rate

The four-differential decay rate for the $K\pi$ system in the P-wave configuration and ignoring scalar[7] contributions, is given by

[7]This refers to the scalar configuration of the dimuon system, not to be confused with the S-wave contribution to the $K\pi$ system, as will be discussed later.

$$\begin{aligned}\frac{\mathrm{d}^4\Gamma[\overline{B}^0\to\overline{K}^{*0}\mu^+\mu^-]}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} =& \frac{9}{32\pi}\sum_i J_i(q^2) f_i(\cos\theta_\ell,\cos\theta_K,\phi) \\ =& \frac{9}{32\pi}\Big[\, J_1^s\sin^2\theta_K + J_1^c\cos^2\theta_K \; + \\ & J_2^s\sin^2\theta_K\cos 2\theta_\ell + J_2^c\cos^2\theta_K\cos 2\theta_\ell \; + \\ & J_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi + J_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi \; + \\ & J_5\sin 2\theta_K\sin\theta_\ell\cos\phi + J_6^s\sin^2\theta_K\cos\theta_\ell \; + \\ & J_7\sin 2\theta_K\sin\theta_\ell\sin\phi + J_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi \; + \\ & J_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi \;\Big]. \end{aligned} \tag{4.29}$$

Here, the q^2 dependent angular observables $J_i(q^2)$ were given by Eq. 4.37.

In the limit that the decay is dominated by a vector current, the relation: $A^L_{\parallel,\perp,0} = A^R_{\parallel,\perp,0}$ holds and implies $J_{5,6,7} = 0$. The angular expression simplifies to the familiar expression for $B^0 \to J/\psi\, K^{*0}$:

$$\begin{aligned}\frac{\mathrm{d}^3\Gamma[\overline{B}^0\to J/\psi\,\overline{K}^{*0}]}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi} = \frac{9}{32\pi}\Big[\, & 2|A_0|^2\cos^2\theta_K\sin^2\theta_\ell \; + \\ & \frac{1}{2}\left(|A_\parallel|^2+A_\perp|^2\right)\sin^2\theta_K(1+\cos^2\theta_\ell) \; + \\ & \frac{1}{2}\left(|A_\perp|^2-|A_\parallel|^2\right)\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi \; + \\ & \sqrt{2}\mathrm{Re}(A_0A_\parallel^*)\sin 2\theta_K\sin 2\theta_\ell\cos\phi \; + \\ & \frac{1}{\sqrt{2}}\mathrm{Im}(A_0A_\perp^*)\sin 2\theta_K\sin 2\theta_\ell\sin\phi \; + \\ & \mathrm{Im}(A_\parallel A_\perp^*)\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big]. \end{aligned} \tag{4.30}$$

In the limit of $q^2 \gg 4m(\mu)^2$ the factor $\beta_\mu^2 \to 1$ and

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} = |A_\perp^L|^2 + |A_\parallel^L|^2 + |A_0^L|^2 + (L\to R).$$

In this limit, $J_2^c = -J_1^c$ and $J_2^s = J_1^s/3$. While the differential decay rate in Eq. 4.29 is defined for the decay of the $\overline{B}^0$ meson, the decay of the B^0 is given in complete analogy by

$$\frac{\mathrm{d}^4\Gamma[B^0\to K^{*0}\mu^+\mu^-]}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} = \frac{9}{32\pi}\sum_i \bar{J}_i(q^2) f_i(\cos\theta_\ell,\cos\theta_K,\phi). \tag{4.31}$$

The identical form of this equation compared to Eq. 4.29 is a consequence of our angular definition described in Sec. 4.6.1. Following Ref. [12], it is customary to define the CP-averaged observables S_i and CP-violating observables A_i according to

$$S_i = \frac{J_i + \bar{J}_i}{\left(\mathrm{d}\Gamma + \mathrm{d}\bar{\Gamma}\right)/\mathrm{d}q^2}, \tag{4.32}$$

$$A_i = \frac{J_i - \bar{J}_i}{\left(\mathrm{d}\Gamma + \mathrm{d}\bar{\Gamma}\right)/\mathrm{d}q^2}. \tag{4.33}$$

The normalisation condition implies $\frac{3}{4}\left(2S_1^s + S_1^c\right) - \frac{1}{4}\left(2S_2^s + S_2^c\right) = 1$. In the limit of massless leptons, the CP-averaged observables are related by $S_2^c = -S_1^c$ and $S_2^s = S_1^s/3$, as discussed above.

Often, the forward-backward asymmetry A_{FB}, and the longitudinal (transverse) polarisation fraction F_L (F_T) are referred to in the literature. These quantities are related to the CP-averaged observables S_i according to

$$\begin{aligned} A_{\mathrm{FB}} &= \frac{3}{4}S_6^s, \\ F_{\mathrm{L}} &= S_1^c = -S_2^c, \\ F_{\mathrm{T}} &= 4S_2^s. \end{aligned}$$

4.6.3 Interference with other $K^+\pi^-$ states

Eq. 4.29 is valid if the $K^+\pi^-$ system is in the P-wave configuration, as is the case for the $K^{*0}(892)$ vector meson. If the $K^+\pi^-$ system is in the S-wave configuration or in the configuration with higher angular momentum up to $J_{\max}$, the replacements

$$A(J=1)_0^{L,R} \cdot Y_1^0(\theta_K, 0) \to \sum_{i=0}^{J_{\max}} A_0^{L,R}(i) \cdot Y_i^0(\theta_K, 0) \text{ and} \tag{4.34}$$

$$A(J=1)_{\parallel,\perp}^{L,R} \cdot Y_1^0(\theta_K, 0) \to \sum_{i=1}^{J_{\max}} A_0^{L,R}(i) \cdot Y_i^{-1}(\theta_K, 0) \tag{4.35}$$

need to be made, where $Y_l^m(\theta_K)$ are spherical harmonics functions. The relevant spherical harmonics for S, P and D-wave are

$$
\begin{aligned}
Y_0^0(\theta_K) &= \frac{1}{2\sqrt{\pi}}, \\
Y_1^0(\theta_K) &= \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta_K, \\
Y_2^0(\theta_K) &= \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta_K - 1), \\
Y_1^{-1}(\theta_K) &= \frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta_K, \\
Y_2^{-1}(\theta_K) &= \frac{1}{2}\sqrt{\frac{15}{2\pi}}\sin\theta_K\cos\theta_K.
\end{aligned}
$$

4.6.4 S-wave interference

For the decay $B^0 \to J/\psi K^{*0}$ the S-wave fraction was determined to be $(6.4 \pm 0.3 \pm 1.0)\%$ in a mass window of $\pm 70\,\text{MeV}$ around the known K^{*0} mass using $1\,\text{fb}^{-1}$ of LHCb data [79]. In this analysis, the S-wave parameters are determined in the full angular analysis. Therefore, Eq. 4.29 needs to be modified according to

$$
\begin{aligned}
\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} \;\to\; & \frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} \\
& + \frac{9}{32\pi}\Big[\, J'^c_1 + J''^c_1\cos\theta_K + \\
& \qquad J'^c_2\cos 2\theta_\ell + J''^c_2\cos\theta_K\cos 2\theta_\ell + \\
& \qquad J'_4\sin 2\theta_\ell\sin\theta_K\cos\phi + \\
& \qquad J'_5\sin\theta_\ell\sin\theta_K\cos\phi + \\
& \qquad J'_7\sin\theta_\ell\sin\theta_K\sin\phi + \\
& \qquad J'_8\sin 2\theta_\ell\sin\theta_K\sin\phi\Big]\,.
\end{aligned}
\tag{4.36}
$$

with

$$
\begin{aligned}
J'^c_1 &= \frac{1}{3}|A^L_{J=0}|^2 + \frac{1}{3}|A^R_{J=0}|^2, \\
J''^c_1 &= \frac{2}{\sqrt{3}}\Big[\mathrm{Re}(A^L_{J=0}A^{L*}_0) + (L \to R)\Big], \\
J'^c_2 &= -\Big[\frac{1}{3}|A^L_{J=0}|^2 + \frac{1}{3}|A^R_{J=0}|^2\Big], \\
J''^c_2 &= -\frac{2}{\sqrt{3}}\Big[\mathrm{Re}(A^L_{J=0}A^{L*}_0) + (L \to R)\Big], \\
J'_4 &= \sqrt{\frac{2}{3}}\Big[\mathrm{Re}(A^L_{J=0}A^{L*}_\parallel) + (L \to R)\Big], \\
J'_5 &= 2\sqrt{\frac{2}{3}}\Big[\mathrm{Re}(A^L_{J=0}A^{L*}_\perp) - (L \to R)\Big], \\
J'_7 &= 2\sqrt{\frac{2}{3}}\Big[\mathrm{Im}(A^L_{J=0}A^{L*}_\parallel) - (L \to R)\Big], \\
J'_8 &= \sqrt{\frac{2}{3}}\Big[\mathrm{Im}(A^L_{J=0}A^{L*}_\perp) + (L \to R)\Big],
\end{aligned}
\tag{4.37}
$$

and

$$
\begin{aligned}
\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} &= \frac{\mathrm{d}\Gamma_P}{\mathrm{d}q^2} + \frac{\mathrm{d}\Gamma_S}{\mathrm{d}q^2} \\
&= |A^L_{J=0}|^2 + |A^L_{J=1,0}|^2 + |A^L_{J=1,\parallel}|^2 + |A^L_{J=1,\perp}|^2 + (L \to R).
\end{aligned}
\tag{4.38}
$$

The fraction of longitudinal polarisation is given by

$$
F_\mathrm{S} = |A^L_{J=0}|^2 \cdot \frac{1}{\mathrm{d}\Gamma/\mathrm{d}q^2}. \tag{4.39}
$$

4.6.5 Optimized observables

The angular observables can be reparametrised so that leading form factor uncertainties cancel to first order. The authors of Ref. [17] propose the basis consisting of F_L (or A_FB)

and the observables $P_i^{(\prime)}$ that can be calculated from the observables S_i according to

$$
\begin{aligned}
P_1 &= 2\frac{S_3}{1-F_\mathrm{L}}, \\
P_2 &= \frac{1}{2}\frac{S_6^s}{1-F_\mathrm{L}} = \frac{2}{3}\frac{A_\mathrm{FB}}{1-F_\mathrm{L}}, \\
P_3 &= -\frac{S_9}{1-F_\mathrm{L}}, \\
P_4' &= \frac{S_4}{\sqrt{F_\mathrm{L}(1-F_\mathrm{L})}}, \\
P_5' &= \frac{S_5}{\sqrt{F_\mathrm{L}(1-F_\mathrm{L})}}, \\
P_6' &= \frac{S_7}{\sqrt{F_\mathrm{L}(1-F_\mathrm{L})}}, \\
P_8' &= \frac{S_8}{\sqrt{F_\mathrm{L}(1-F_\mathrm{L})}}.
\end{aligned}
$$

The PDF used to fit the angular observables can be reparametrised to use the basis $(F_\mathrm{L}, P_{1,2,3}, P'_{4,5,6,8})$. The shapes of the allowed parameter regions are simplified compared to the basis $(F_\mathrm{L}, A_\mathrm{FB}, S_{3,4,5,7,8,9})$. However, the reparametrisation introduces additional correlations between the observables since the PDF is no longer linear in the parameters. In this analysis, the CP-averaged observables S_i and the CP-asymmetries A_i will be considered the nominal observables. However, we aim to provide also less form factor dependent observables, which can be either fit directly or constructed from the observables S_i and A_i.

We would like to note that theorists use a different convention of the P_i definitions. For completeness, we give the definition for P_i which are different:

$$
\begin{aligned}
P'_{4,theory} &= 2\frac{S_4}{\sqrt{F_\mathrm{L}(1-F_\mathrm{L})}}, \\
P'_{6,theory} &= -\frac{S_7}{\sqrt{F_\mathrm{L}(1-F_\mathrm{L})}}.
\end{aligned}
$$

4.7 Fitting for angular observables

This method of angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ determines the angular observables S_i (or A_i) in bins of q^2 using an unbinned maximum likelihood fit of the reconstructed B^0 mass and the decay angles $\vec{\Omega} = (\cos\theta_l, \cos\theta_K, \phi)$.

The analysis needs to describe the signal and background components using PDFs depending on the the angular observables and nuisance parameters. The total PDF is

given by

$$\mathcal{P}_{\text{tot}} = f_{\text{sig}}\mathcal{P}_{\text{sig}}(\vec{\Omega}, m) + (1 - f_{\text{sig}})\mathcal{P}_{\text{bkg}}(\vec{\Omega}, m). \tag{4.40}$$

Both signal and background components are assumed to factorise in the decay angles $\vec{\Omega}$ and the reconstructed B^0 mass m

$$\mathcal{P}_{\text{sig}}(\vec{\Omega}, m) = \mathcal{P}_{\text{sig}}(\vec{\Omega}) \times \mathcal{P}_{\text{sig}}(m), \tag{4.41}$$

$$\mathcal{P}_{\text{bkg}}(\vec{\Omega}, m) = \mathcal{P}_{\text{bkg}}(\vec{\Omega}) \times \mathcal{P}_{\text{bkg}}(m). \tag{4.42}$$

To determine the angular observables, the negative logarithmic likelihood

$$-\log\mathcal{L} = -\sum_{\text{events } e} \log\mathcal{P}_{\text{tot}}(\vec{\Omega}_e, m_e|\vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) \tag{4.43}$$

is minimised with respect to the physics parameters $\vec{\lambda}_{\text{phys}}$ and the nuisance parameters $\vec{\lambda}_{\text{nuisance}}$. The minimisation is performed using the MINUIT software package [80]. The uncertainties on the parameters can be either determined using the second derivative matrix (HESSE) or the $-2\Delta\log\mathcal{L} = 1$ rule (MINOS), which allows asymmetric uncertainties.

4.7.1 Angular distributions

The angular description of the signal component of the PDF is given by the differential decay rate given by Eq. 4.29. The data are binned in q^2, thereby effectively averaging the observables over the width of the q^2 bins. The resulting three-differential decay rate is given by

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_l\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\text{P}} = \tfrac{9}{32\pi}\Big[&\tfrac{3}{4}(1-F_{\text{L}})\sin^2\theta_K \\ &+ F_{\text{L}}\cos^2\theta_K + \tfrac{1}{4}(1-F_{\text{L}})\sin^2\theta_K\cos 2\theta_l \\ &- F_{\text{L}}\cos^2\theta_K\cos 2\theta_l + S_3\sin^2\theta_K\sin^2\theta_l\cos 2\phi \\ &+ S_4\sin 2\theta_K\sin 2\theta_l\cos\phi + S_5\sin 2\theta_K\sin\theta_l\cos\phi \\ &+ \tfrac{4}{3}A_{\text{FB}}\sin^2\theta_K\cos\theta_l + S_7\sin 2\theta_K\sin\theta_l\sin\phi \\ &+ S_8\sin 2\theta_K\sin 2\theta_l\sin\phi + S_9\sin^2\theta_K\sin^2\theta_l\sin 2\phi\Big].\end{aligned} \tag{4.44}$$

As discussed in Sec. 4.6.4, the inclusion of the S-wave contribution, where the $K^+\pi^-$ system is in a spin 0 configuration, leads to additional angular terms. The PDF needs to be changed to include both the S-wave and the interference between the S- and P-wave

resulting in

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{dcos}\theta_l\,\mathrm{dcos}\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{S+P}} = (1-F_S)\,\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{dcos}\theta_l\,\mathrm{dcos}\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{P}} \quad (4.45)$$
$$+\tfrac{3}{16\pi}\Big[F_S\sin^2\theta_l + S_{S1}\sin^2\theta_l\cos\theta_K$$
$$+ S_{S2}\sin 2\theta_l\sin\theta_K\cos\phi$$
$$+ S_{S3}\sin\theta_l\sin\theta_K\cos\phi$$
$$+ S_{S4}\sin\theta_l\sin\theta_K\sin\phi$$
$$+ S_{S5}\sin 2\theta_l\sin\theta_K\sin\phi\Big].$$

The background component in the maximum likelihood fit is modelled using Chebyshev polynomials. The angular parametrisation of the background is assumed to factorise in the $\cos\theta_K,\ \cos\theta_l,\ \phi$ angles. For the angular parametrisation of the background using the Chebyshev polynomials T_i of second order and lower, the PDF is given by

$$\mathcal{P}_{\mathrm{bkg}}(\cos\theta_l,\cos\theta_K,\phi) = \left[\sum_{i=0}^{2} c_i T_i(\cos\theta_l)\right]\times\left[\sum_{j=0}^{2} c_j T_j(\cos\theta_K)\right]\times\left[\sum_{k=0}^{2} c_k T_k(\phi)\right]. \quad (4.46)$$

The factorization assumption has been checked by scanning the right B side-band regions ($m\in[5380,6000]$ MeV/c^2). It has been observed that the shape of the background does not change in different m regions.

4.7.2 Mass modeling

The reconstructed B^0 mass of the signal is modelled using the sum of two Crystal Ball functions with common tail parameters for the low mass side. The parameters describing the signal mass shape are determined from a fit to the control-decay $B^0\to J/\psi\,K^{*0}$ and the q^2 dependency is accounted for by a q^2 dependent scale factor determined from Monte Carlo simulation. The mass distribution of the background is modelled using an exponential function. For the fits of the $B^0\to K^{*0}\mu^+\mu^-$ signal only τ_m, the inverse of the exponential decay constant as well as the signal fraction f_{sig} are floated. The other mass parameters are taken from the control decay and fixed in the fit of the signal decay.

4.7.3 Acceptance effect

The reconstruction and selection of the signal decay distorts the angular distributions and needs to be accounted for when determining the angular observables. This acceptance effect, depending on q^2 and the decay angles can be parameterised using multidimensional polynomials

$$\epsilon(q^2,\cos\theta_l,\cos\theta_K,\phi) = \sum_{hijk} c_{hijk}\times(q^2)^h\times(\cos\theta_l)^i\times(\cos\theta_K)^j\times(\phi)^k. \quad (4.47)$$

The determination of the polynomial coefficients and the resulting angular description was discussed in Sec. 4.5. This efficiency can be included in the fit in two ways, either by performing a weighted fit in which the events are weighted by $1/\epsilon$, or by including the effect in the signal PDF.

In the first option, the distributions are effectively unfolded, therefore the original signal PDF without acceptance can be used. It should be noted that the background component will be weighted in the same way. The per-event weight is included in the likelihood as follows

$$\begin{aligned}\mathcal{L} &= -\sum_{\text{event } e} w_e \times \log \mathcal{P}(\vec{\Omega}_e, m_e) \\ &= -\sum_{\text{event } e} \frac{1}{\epsilon(q_e^2, \vec{\Omega}_e)} \times \log \mathcal{P}(\vec{\Omega}_e, m_e).\end{aligned}$$

Special care needs to be taken for the estimation of the parameter uncertainties, since weighted fits in general are not guaranteed correct coverage. However, approximate methods exist. The corrected covariance matrix V' for the weighted fit can be calculated according to

$$V' = VC^{-1}V,$$

where V is the covariance matrix calculated with the weights w_e and C the covariance matrix calculated using the squared weights w_e^2 [81]. The unfolding using acceptance weights is the preferred approach to the large q^2 bins 1.1 $\text{GeV}^2/c^4 < q^2 < 6$ GeV^2/c^4 and 15.0 $\text{GeV}^2/c^4 < q^2 < 19.0$ GeV^2/c^4, since the method can account for a possible variation of the acceptance with q^2. Furthermore, the expected signal yield in these bins is sufficiently large to reduce a possible fluctuations from the weighting procedure.

The second option requires to include the efficiency in the signal PDF. The main difficulty with this approach is the correct determination of the norm of the signal component which will be affected by the acceptance[8]. The norm $\mathcal{N}_{\text{sig}}$ is given by

$$\begin{aligned}\mathcal{N}_{\text{sig}} &= \int \epsilon(q^2, \vec{\Omega}) \mathcal{P}_{\text{sig}}(\vec{\Omega}) \mathrm{d}\vec{\Omega} \\ &= \int \epsilon(q^2, \vec{\Omega}) \frac{9}{32\pi} \sum_i S_i f_i(\vec{\Omega}) \mathrm{d}\vec{\Omega} \\ &= \frac{9}{32\pi} \sum_i S_i \xi_i(q^2), \end{aligned} \tag{4.48}$$

with $\xi_i = \int \epsilon(q^2, \vec{\Omega}) f_i(\vec{\Omega}) \mathrm{d}\vec{\Omega}$, and where the angular terms $f_i(\vec{\Omega})$ are defined by Eqs. 4.44 and 4.45. This is the preferred approach to the $2\,\text{GeV}^2/c^4$ q^2 bins, where the acceptance does not vary significantly over the (narrow) q^2 bin.

[8]Note that the factor $\epsilon(q^2, \vec{\Omega})$ in the numerator can be omitted when determining $-\log \mathcal{L}$.

4.7.4 Physical boundaries of the observables

Eqs. 4.44 and 4.45 imply certain boundaries for the angular observables since the PDF is not allowed to become negative for any combination of angles. If the values of the angular observables lie close to these constraints, the likelihood function becomes non-Gaussian. Technically, a large penalty term is added in the fit for every event for which $\mathcal{P}_{\text{tot}}(\vec{\Omega}_e, m_e)$ becomes negative. It is instructive to explore the particular shape of the allowed parameter regions by performing parameter scans and using toys to find wether the PDF can become negative for a certain parameter set. The red points are the SM values for the seven bins of the $2\,\text{GeV}^2/c^4$ q^2 binning. Particularly striking are the constraints on the combinations of F_L with the other observables, which can be expressed by the following relations

$$|S_3| \leqslant \frac{1}{2}(1 - F_L), \tag{4.49}$$

$$|A_{\text{FB}}| \leqslant \frac{3}{4}(1 - F_L), \tag{4.50}$$

$$|S_9| \leqslant \frac{1}{2}(1 - F_L). \tag{4.51}$$

Owing to the large dependence of the allowed parameter range on F_L, the allowed regions for the parameter combinations not containing F_L are integrated over all possible F_L values (we iterate over the full available F_L range from zero to one in 40 steps). The other parameters not shown are either assumed to be zero or equal to one of the seven SM points of the $2\,\text{GeV}^2/c^4$ q^2 binning. If the studied point is allowed for one of those eight possibilities it is marked as allowed. The parameter boundaries can affect the coverage negatively. To ensure correct coverage we therefore rely on the Feldman-Cousins method described in Sec. 4.7.7.

4.7.5 CP-asymmetries $\mathbf{A_i}$ and the $P_i^{(\prime)}$ basis

There is significant theoretical interest in the CP-asymmetries A_i defined in Eq. 4.33, particularly the T-odd asymmetries $A_{7,8,9}$, where significant effects of new weak phases could be seen. To determine the angular observables, Eq. 4.44 needs to be modified, replacing $S^{(s)}_{3,\ldots,9}$ by $A_{3,\ldots,9}$ for the $\overline{B}^0$ decay and $-A_{3,\ldots,9}$ for the B^0 decay flavour.

The PDF can also be expressed in the $P_i^{(\prime)}$ basis detailed in Sec. 4.6.5. The varied signal parameters in this case are F_{L}, $P_{1,2,3}$ and $P'_{4,5,6,8}$. Since $P_i^{(\prime)}$ are nonlinear combinations of the angular coefficients S_i, the uncertainties are expected to be generally more asymmetric.

4.7.6 Fit validation using `EOS` toys

The fit is validated using simulated events generated according to an updated theory calculation based on the `EOS` package [82]. Pull studies are conducted to ensure that the fit is unbiased and estimates the uncertainties correctly. The pull of the observable p is defined as $(p_{\text{fitted}} - p_{\text{generated}})/\sigma(p)_{\text{fitted}}$. In the ideal case the pull is distributed according

to a Gaussian distribution with a mean compatible with 0 and a width compatible with 1. For low statistics and non-Gaussian PDFs this is not necessarily the case and it can lead to incorrect coverage. For the toy studies, the P-wave observables S_i, the signal fraction f_{sig}, the parameter τ_m describing the exponential shape of the combinatorial background, and six coefficients describing the angular distribution of the combinatorial background as detailed in Sec. 4.7.1 are floated. The `EOS` toys do not contain an S-wave component.

Generally the toys behave well; however, there are some observables where sizeable biases larger than 0.1 are seen. A particularly large deviation from the `EOS` value is observed for the first bin for S_1^s. This deviation is understood since the fit assumes that the lepton masses can be neglected, which is not the case close to $q^2 = 0\ \text{GeV}^2/c^4$. All other deviations seen are smaller than 0.20. To ensure correct coverage for the $2\ \text{GeV}^2/c^4$ binning, the Feldman-Cousins method will be used for the determination of the confidence intervals.

The corresponding CP-asymmetries A_i have been determined from the `EOS` toy MC as well. No CP-asymmetries show significant biases. As discussed in Sec. 4.7.5, this is due to the SM values being further away from physical parameter boundaries.

4.7.7 Coverage correction

To guarantee correct coverage even for non-Gaussian PDFs, the Feldman-Cousins method [83] is employed. This method is a specific Neyman construction using likelihood ratios as an ordering principle. The nuisance parameters are included using the plugin method [84].

Technically, the parameter of interest is scanned at a number of equidistant points. For every point the likelihood ratio on data, $\Delta \log \mathcal{L}_{\text{data}} = \log \mathcal{L}^{\text{data}}_{\text{fixed}} - \log \mathcal{L}^{\text{data}}_{\text{floated}}$, is determined, where the parameter of interest is fixed at the point for $\log \mathcal{L}^{\text{data}}_{\text{fixed}}$, but allowed to float for $\log \mathcal{L}^{\text{data}}_{\text{floated}}$. Then N_{toys} toys are thrown for the point, determining N_{toys} toy likelihood ratios $\Delta \log \mathcal{L}_{\text{toy }i} = \log \mathcal{L}^{\text{toy}}_{\text{fixed }i} - \log \mathcal{L}^{\text{toy}}_{\text{floated }i}$. The confidence level of the point under study is then given by the fraction of toys for which $\Delta \log \mathcal{L}_{\text{toy }i} > \Delta \log \mathcal{L}_{\text{data}}$.

Fig. 4.10 shows the results for a single toy generated with `EOS` with signal and background yields corresponding to the $3\,\text{fb}^{-1}$ data sample. The observable S_5 is determined in seven bins of q^2. For every bin, 500 toys are generated for 100 equidistant points of the observable of interest. The resulting coverage-corrected 68.3% confidence interval is given in red. The blue line denotes the coverage from the likelihood method, the 68.3% confidence interval from the likelihood is given by the blue vertical line. As is apparent for bins three and seven, for certain parameter configurations the likelihood method undercovers.

4.7.8 Fit validation on data using $B^0 \to J/\psi\, K^{*0}$

The angular distributions of the tree-level decay $B^0 \to J/\psi\, K^{*0}$ were studied previously by the BaBar [85], Belle [86] and CDF [87] experiments. Most recently, LHCb has analysed the decay using $1\,\text{fb}^{-1}$ of the data recorded in 2011 [79]. The decay $B^0 \to J/\psi\, K^{*0}$ is selected using the full selection for the $B^0 \to K^{*0}\mu^+\mu^-$ signal decay and requiring that the invariant mass of the dimuon system is $\pm 60\,\text{MeV}/c^2$ around the known J/ψ mass. The

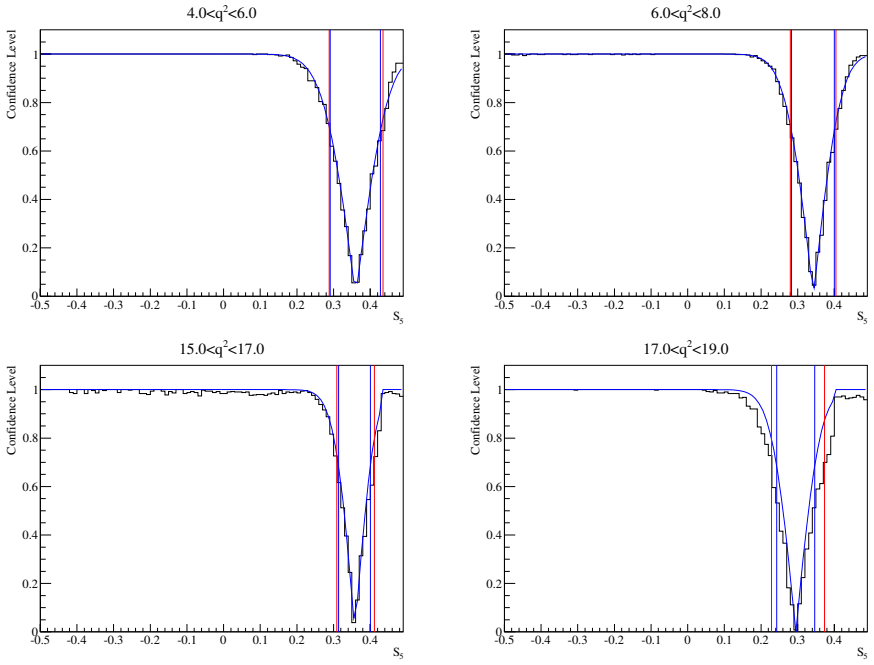


Figure 4.10: Feldman-Cousins results for the observable S_5 using an `EOS` toy in seven bins of q^2. The Feldman-cousins confidence level is given by the black histogram. The red vertical lines denote the 68.3% confidence interval from the Feldman-Cousins method. As comparison the blue line gives the confidence level using the likelihood method. The blue vertical lines give the 68.3% from the likelihood method.

parameter n_{CB} is fixed to 4.23, determined from a fit in the q^2 region $[8.0, 11.0]\,\mathrm{GeV}^2/c^4$ where the fit is more stable. In contrast to the fit of the signal decay $B^0 \to K^{*0}\mu^+\mu^-$, the contribution from the B_s^0 decay $B_s^0 \to J/\psi\, K^{*0}$, which is suppressed by $f_s/f_d|V_{\mathrm{cd}}/V_{\mathrm{cs}}|^2$, is modelled in the fit as well. Its angular and mass distribution are assumed to be identical to $B^0 \to J/\psi\, K^{*0}$, with the mass distribution of the B_s^0 shifted by Δm.

The parametrisation described in Sec. 4.5 is used as angular acceptance, included in the fit as discussed in Sec. 4.7.3. Table 4.5 gives the result of a full angular fit in different bins of $m_{K\pi}$. For comparison, Table 4.6 gives the angular terms that were previously measured by LHCb [79]. Since Ref. [79] gives the magnitudes of the amplitudes $|A_{0,\parallel,\perp,S}|$ and the strong phases $\delta_{\parallel,\perp,S}$ as results, they are converted to the angular observables according to Eqs. 4.37. The observables are found to be consistent with the previous measurement of $B^0 \to J/\psi\, K^{*0}$.

4.7.9 Constraining the S-wave using the $m_{K\pi}$ distribution

According to Eq. 4.45, all P-wave parameters are scaled by the factor $(1 - F_S)$ which is not known a priori. Neglecting the S-wave in the fit and correcting the P-wave parameters using F_S from the dedicated S-wave analysis, since it partially uses the same data distributions

Table 4.5: Results of the angular fit of the decay $B^0 \to J/\psi K^{*0}$ in different bins of $m_{K\pi}$, using the full available data set corresponding to $3\,\mathrm{fb}^{-1}$. The angular terms that have been previously determined in Ref. [79] are given in Tab. 4.6.

	$m_{K\pi}$ range in MeV/c^2					
parameter	$[795.9, 995.9]$	$[825.9, 965.9]$	$[826.0, 861.0]$	$[861.0, 896.0]$	$[896.0, 931.0]$	$[931.0, 966.0]$
S_{1s}	0.331 ± 0.001	0.329 ± 0.001	0.326 ± 0.004	0.324 ± 0.002	0.333 ± 0.002	0.334 ± 0.003
S_3	-0.000 ± 0.002	0.000 ± 0.002	-0.009 ± 0.006	0.000 ± 0.003	0.000 ± 0.003	0.004 ± 0.006
S_4	-0.255 ± 0.002	-0.255 ± 0.002	-0.258 ± 0.007	-0.258 ± 0.003	-0.254 ± 0.003	-0.251 ± 0.006
S_5	-0.001 ± 0.002	-0.002 ± 0.002	-0.002 ± 0.007	0.000 ± 0.003	-0.007 ± 0.003	0.002 ± 0.006
S_6^s	0.000 ± 0.002	0.000 ± 0.002	-0.005 ± 0.006	-0.000 ± 0.003	0.001 ± 0.003	0.004 ± 0.005
S_7	0.001 ± 0.002	0.001 ± 0.002	-0.000 ± 0.007	0.001 ± 0.003	0.002 ± 0.003	0.002 ± 0.006
S_8	-0.053 ± 0.002	-0.052 ± 0.002	-0.064 ± 0.007	-0.055 ± 0.003	-0.051 ± 0.003	-0.045 ± 0.006
S_9	-0.089 ± 0.002	-0.089 ± 0.002	-0.088 ± 0.007	-0.084 ± 0.003	-0.094 ± 0.003	-0.090 ± 0.006
F_S	0.087 ± 0.003	0.072 ± 0.003	0.12 ± 0.01	0.051 ± 0.005	0.061 ± 0.005	0.119 ± 0.009
S_{S1}	-0.234 ± 0.003	-0.233 ± 0.004	-0.75 ± 0.01	-0.363 ± 0.006	-0.091 ± 0.006	0.15 ± 0.01
S_{S2}	0.023 ± 0.002	0.027 ± 0.002	0.159 ± 0.007	0.065 ± 0.004	-0.006 ± 0.004	-0.091 ± 0.007
S_{S3}	0.003 ± 0.002	0.003 ± 0.002	-0.004 ± 0.007	0.003 ± 0.003	0.004 ± 0.004	0.009 ± 0.006
S_{S4}	0.001 ± 0.002	0.001 ± 0.002	0.015 ± 0.007	-0.003 ± 0.003	0.000 ± 0.004	0.007 ± 0.006
S_{S5}	-0.068 ± 0.002	-0.064 ± 0.002	0.037 ± 0.008	-0.031 ± 0.004	-0.091 ± 0.004	-0.166 ± 0.007

Table 4.6: Results of the full angular fit of the decay $B^0 \to J/\psi K^{*0}$ in Ref. [79], translated to the angular observables.

	$m_{K\pi}$ range in MeV/c^2				
parameter	$[825.9, 965.9]$	$[826.0, 861.0]$	$[861.0, 896.0]$	$[896.0, 931.0]$	$[931.0, 966.0]$
S_1^s	0.321 ± 0.006	0.321 ± 0.006	0.321 ± 0.006	0.321 ± 0.006	0.321 ± 0.006
S_3	-0.013 ± 0.010	-0.013 ± 0.010	-0.013 ± 0.010	-0.013 ± 0.010	-0.013 ± 0.010
S_4	-0.250 ± 0.006	-0.250 ± 0.006	-0.250 ± 0.005	-0.250 ± 0.006	-0.250 ± 0.006
S_5	0	0	0	0	0
S_6^s	0	0	0	0	0
S_7	0	0	0	0	0
S_8	-0.048 ± 0.007	-0.048 ± 0.007	-0.048 ± 0.007	-0.048 ± 0.007	-0.048 ± 0.007
S_9	-0.084 ± 0.006	-0.084 ± 0.006	-0.084 ± 0.006	-0.084 ± 0.006	-0.084 ± 0.006
F_S	0.064 ± 0.010	0.115 ± 0.021	0.049 ± 0.008	0.052 ± 0.011	0.105 ± 0.016
S_{S1}	-	-0.887 ± 0.082	-0.514 ± 0.030	-0.216 ± 0.044	0.035 ± 0.096
S_{S2}	-	0.192 ± 0.018	0.100 ± 0.007	0.022 ± 0.012	-0.045 ± 0.021
S_{S3}	-	0	0	0	0
S_{S4}	-	0	0	0	0
S_{S5}	-	0.028 ± 0.023	-0.034 ± 0.012	-0.105 ± 0.015	-0.176 ± 0.013

($m_{K\pi\mu\mu}$ and $\cos\theta_K$). A possibility to circumvent these difficulties is to include the $m_{K\pi}$ projection in a simultaneous fit. Since the P-wave is peaking in $m_{K\pi}$ while the S-wave contribution is relatively flat, this gives an additional constraint on F_S and therefore also allows a better determination of the P-wave observables. Ref. [88] gives details on the dependence of the decay amplitudes on $m_{K\pi}$. To parameterize the $m_{K\pi}$ dependence of the

P-wave a Breit-Wigner distribution is used

$$\mathcal{A}_P(m_{K\pi}) = \sqrt{pq} \times B'_{L_B}(p, p_0, d) \left(\frac{p}{m_B}\right)^{L_B} \times B'_{L_{K^*}}(q, q_0, d) \left(\frac{q}{m_{K\pi}}\right)^{L_{K^*}} \times \frac{1}{m^2_{K^*} - m^2_{K\pi} - i m_{K^*}\Gamma(m_{K\pi})}, \tag{4.52}$$

where p (q) denotes the K^{*0} (K^+) momentum in the B^0 (K^{*0}) rest frame, p_0 (q_0) are the corresponding quantities at the resonance peak. L_B (L_{K^*}) are the orbital angular momenta and B'_{L_B} ($B'_{L_{K^*}}$) the Blatt-Weisskopf functions. For the S-wave component the LASS parameterisation [89] is used

$$\mathcal{A}_S(m_{K\pi}) = \sqrt{pq} \times B'_{L_B}(p, p_0, d) \left(\frac{p}{m_B}\right)^{L_B} \times B'_{L_{K^*_0}}(q, q_0, d) \left(\frac{q}{m_{K\pi}}\right)^{L_{K^*_0}} \times \left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B} \frac{1}{\cot\delta_R - i}\right), \tag{4.53}$$

where $\cot\delta_B = \frac{1}{aq} + \frac{1}{2}rq$ and $\cot\delta_R = (m^2_{K^*_0} - m^2_{K\pi})/(m_{K^*_0}\Gamma_0(m_{K\pi}))$. Accounting for the $m_{K\pi}$ dependence, Eq. 4.45, integrated over the three decay angles $\cos\theta_l$, $\cos\theta_K$ and ϕ, becomes

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2} \frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}m_{K\pi}}\bigg|_{\mathrm{S+P}} = (1-F_S) \sum_{i=1}^{9} \tfrac{9}{32\pi} \xi_i S_i^{(s,c)} |\mathcal{A}'_P(m_{K\pi})|^2 \\ + \tfrac{3}{16\pi}\Big[F_S \xi_{F_S} |\mathcal{A}'_S(m_{K\pi})|^2 \\ + (S_{S1}\xi_{S1} + S_{S2}\xi_{S2} + S_{S3}\xi_{S3})\, \Re\left(\mathcal{A}'_S(m_{K\pi})\mathcal{A}^{*\prime}_P(m_{K\pi})\right) \\ + (S_{S4}\xi_{S4} + S_{S5}\xi_{S5})\, \Im\left(\mathcal{A}'_S(m_{K\pi})\mathcal{A}^{*\prime}_P(m_{K\pi})\right)\Big],\end{aligned} \tag{4.54}$$

where $\xi_{(S)i}$ denote the angular integrals $\xi_{(S)i} = \int \epsilon(\cos\theta_l, \cos\theta_K, \phi) f_{(S)i}(\cos\theta_l, \cos\theta_K, \phi) \mathrm{d}\vec{\Omega}$ and the amplitudes are appropriately normalized according to

$$\mathcal{A}'_P(m_{K\pi}) = \frac{\mathcal{A}_P(m_{K\pi})}{\sqrt{\int_{795.9\,\mathrm{MeV}/c^2}^{995.9\,\mathrm{MeV}/c^2} |\mathcal{A}_P(m_{K\pi})|^2 \mathrm{d}m_{K\pi}}},$$

$$\mathcal{A}'_S(m_{K\pi}) = \frac{\mathcal{A}_S(m_{K\pi})}{\sqrt{\int_{795.9\,\mathrm{MeV}/c^2}^{995.9\,\mathrm{MeV}/c^2} |\mathcal{A}_S(m_{K\pi})|^2 \mathrm{d}m_{K\pi}}}.$$

The simultaneous fit of the angles and the $m_{K\pi}$ projection are tested using the control decay $B^0 \to J/\psi\, K^{*0}$. Fig. 4.11 shows the corresponding projections on the decay angles, $m_{K\pi\mu\mu}$ and $m_{K\pi}$. The result is in good agreement with the results in Tab. 4.5 where only the decay angles are used.

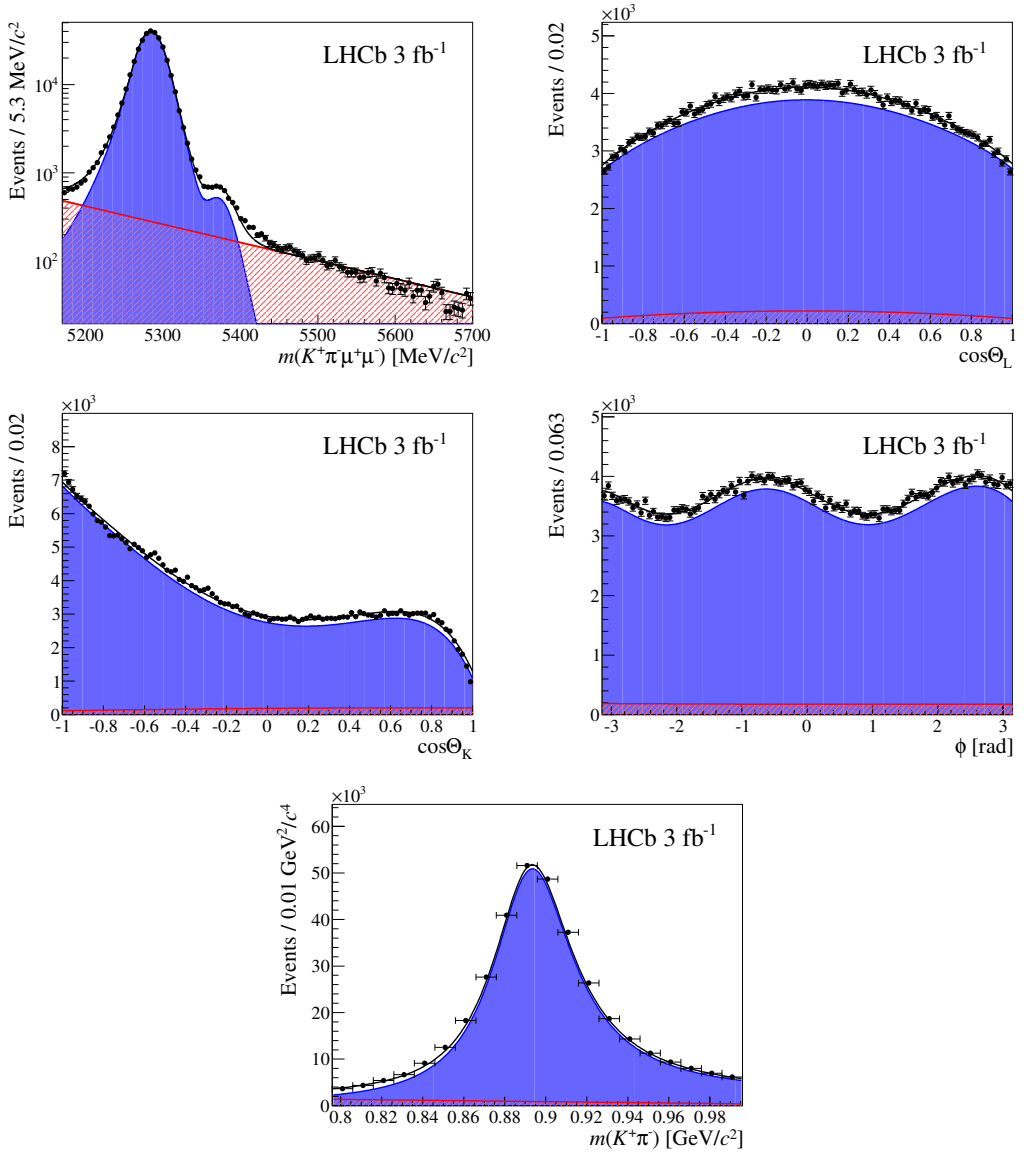


Figure 4.11: Angular, $m_{K\pi\mu\mu}$ and $m_{K\pi}$ projections after the fit of the full $B^0 \to J/\psi K^{*0}$ data sample. The fit is performed as described in Sec. 4.7.9, simultaneously in the decay angles, $m_{K\pi\mu\mu}$, and $m_{K\pi}$, in the $m_{K\pi}$ mass range $[795.9, 995.9]$ MeV/c^2. The slight mismodeling of the reconstructed B^0 mass is due to the narrow $\pm 60\,\text{MeV}/c^2$ requirement around the known J/ψ mass which removes away the radiative tails. This, however, does not affect the angular observables significantly.

4.8 Extracting angular observables using the method of moments

Thanks to the orthogonality of the spherical harmonics (and consequently the angular terms), it is possible to extract the angular observables from a moment analysis [13, 90, 91]. The angular distribution has the form

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\mathrm{d}\vec{\Omega}} = \frac{9}{32\pi}\sum_i S_i(q^2) f_i(\vec{\Omega}) \ , \tag{4.55}$$

which can be averaged over q^2 bin:

$$\frac{\mathrm{d}^3\Gamma}{\mathrm{d}\vec{\Omega}} = \frac{9}{32\pi}\sum_i S_i f_i(\vec{\Omega}) \ , \tag{4.56}$$

and is normalised so that

$$\int \frac{\mathrm{d}^3\Gamma}{\mathrm{d}\vec{\Omega}}\mathrm{d}\vec{\Omega} = \int_{-1}^{+1}\int_{-1}^{+1}\int_{-\pi}^{+\pi} \frac{\mathrm{d}^3\Gamma}{\mathrm{d}\cos\theta_K\,\mathrm{d}\cos\theta_l\,\mathrm{d}\phi}\mathrm{d}\cos\theta_K\,\mathrm{d}\cos\theta_l\,\mathrm{d}\phi = 1, \tag{4.57}$$

where $\vec{\Omega} = (\cos\theta_K, \cos\theta_l, \phi)$ and

$$\begin{aligned}
f_{1s}(\cos\theta_K, \cos\theta_l, \phi) &= \sin^2\theta_K \\
f_3(\cos\theta_K, \cos\theta_l, \phi) &= \sin^2\theta_K \sin^2\theta_l \cos 2\phi \\
f_4(\cos\theta_K, \cos\theta_l, \phi) &= \sin 2\theta_K \sin 2\theta_l \cos\phi \\
f_5(\cos\theta_K, \cos\theta_l, \phi) &= \sin 2\theta_K \sin\theta_l \cos\phi \\
f_{6s}(\cos\theta_K, \cos\theta_l, \phi) &= \sin^2\theta_K \cos\theta_l \\
f_7(\cos\theta_K, \cos\theta_l, \phi) &= \sin 2\theta_K \sin\theta_l \sin\phi \\
f_8(\cos\theta_K, \cos\theta_l, \phi) &= \sin 2\theta_K \sin 2\theta_l \sin\phi \\
f_9(\cos\theta_K, \cos\theta_l, \phi) &= \sin^2\theta_K \sin^2\theta_l \sin 2\phi
\end{aligned} \tag{4.58}$$

Since the angular functions are orthogonal we have:

$$\int f_i(\vec{\Omega}) f_j(\vec{\Omega})\mathrm{d}\vec{\Omega} = \alpha\delta_{ij}, \tag{4.59}$$

for $i = 3 \ldots 9$ and where α is a normalisation constant. The mean (or expectation value) of the f_i can be used to determine the S_i, i.e.

$$M_i = \int \frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}} f_i(\vec{\Omega})\mathrm{d}\vec{\Omega} = \begin{cases} \frac{8}{25} S_{i=3,4,8,9} \\ \frac{2}{5} S_{i=5,6,7} \ . \end{cases} \tag{4.60}$$

Note that the f_i for f_{1s}, f_{2s}, f_{1c} and f_{2c} are not orthogonal. The corresponding moments are the linear combinations of S_{1s}, S_{2s}, S_{1c} and S_{2c}, with

$$M_{1s} = \frac{2}{5}(2 - F_{\mathrm{L}}) \quad , \tag{4.61}$$

under the assumption that $q^2 \gg 4m(\mu)^2$. Re-arranging gives

$$F_{\mathrm{L}} = 2 - \frac{5}{2}M_{1s}, \tag{4.62}$$

$$S_i = \frac{5}{2}M_{5,6,7}, \tag{4.63}$$

$$S_i = \frac{25}{8}M_{3,4,8,9}. \tag{4.64}$$

$$\tag{4.65}$$

In the absence of background, the parameters M_i can be estimated as

$$\langle M_i \rangle = \frac{1}{N} \sum_{\text{event } e} f_i(\vec{\Omega}_e), \tag{4.66}$$

where N is the number of events in the data sample. An estimate for the error can be evaluated as a normal variance

$$\delta\langle M_i \rangle = \sqrt{\frac{1}{N(N-1)} \sum_{\text{event } e} \Big(\langle M_i \rangle - f_i(\vec{\Omega}_e)\Big)^2}. \tag{4.67}$$

4.8.1 Measurement of S_{6c} observable

In the SM, the term S_{6c} is vanishingly small. It is relevant only in the presence of large scalar operators and is suppressed by $m_\mu/\sqrt{q^2}$. The method of moments can be used to determine the S_{6c} coefficient. The corresponding angular term is

$$f_{6c}(\cos\theta_K, \cos\theta_l, \phi) = \cos^2\theta_K \cos\theta_l \ , \tag{4.68}$$

which appears mixed with S_{6s} when evaluating the raw moments:

$$M_{6c} = \frac{1}{20}(3S_{6c} + 2S_{6s}), \tag{4.69}$$

$$M_{6s} = \frac{1}{10}(S_{6c} + 4S_{6s}). \tag{4.70}$$

$$\tag{4.71}$$

The solution to this linear system is

$$S_{6c} = 2(4S_{6c} - S_{6s}), \tag{4.72}$$

$$S_{6s} = 2S_{6c} + 3S_{6s}. \tag{4.73}$$

$$\tag{4.74}$$

This allows to determine both S_{6c} and S_{6s}.

4.8.2 Method of moments in the presence of background

In the presence of background, the moments M_i in the signal mass window will be an admixture of the moments for pure signal ($M_{i,\text{sig}}$) and pure background ($M_{i,\text{bkg}}$). The mixed moment reads:

$$M_{i,\text{mix}} = \frac{N_{\text{sig}} M_{i,\text{sig}} + N_{\text{bkg}} M_{i,\text{bkg}}}{N_{\text{sig}} + N_{\text{bkg}}} \ , \tag{4.75}$$

where N_{sig} and N_{B} are the number of signal and background events in the signal mass window, respectively. The yields N_{sig} and N_{bkg} can be estimated from an extended unbinned maximum likelihood fit to the $K^+\pi^-\mu^+\mu^-$ invariant mass and $M_{i,\text{bkg}}$ can be estimated from the upper mass sideband. The upper mass sideband is chosen to be $m(K^+\pi^-\mu^+\mu^-) > 5350$ MeV/c^2. It would also be possible to determine $M_{i,\text{sig}}$ by sWeighting the events.

4.8.3 Acceptance corrections of the method of moments

When including the angular acceptance, the measured moments (raw moments) are no longer proportional to the observables S_i. To correct for the acceptance, each event is weighted according to a weight

$$w_e = \frac{1}{\epsilon(\vec{\Omega}_e, q_e^2)} \ , \tag{4.76}$$

where $\epsilon(\vec{\Omega}_e, q_e^2)$ is the efficiency function derived in Sec. 4.5. The corresponding formula to obtain the raw moments is then

$$\widehat{M}_i = \frac{1}{\sum_e w_e} \sum_{\text{event } e} w_e f_i(\vec{\Omega}_e) \ . \tag{4.77}$$

The angular acceptance does not need to be treated as constant over the q^2 bin and the full q^2 dependence can be accounted for. The absolute normalisation of the weights does not matter, since it appears in the numerator and denominator of Eq. 4.77, i.e. the weights can be re-scaled for an arbitrary constant.

An estimate for the uncertainty on the moments can be derived from the weighted variance

$$V_{ij} = \frac{\sum_e w_e}{(\sum_e w_e)((\sum_e w_e)^2 - \sum_e w_e^2)} \sum_e w_e \left(\widehat{M}_i - f_i(\vec{\Omega}_e)\right)\left(\widehat{M}_j - f_j(\vec{\Omega}_e)\right) . \tag{4.78}$$

4.8.4 Toy studies for the method of moments

Toy studies for the method of moments were performed in the same way as for the fits for the angular observables. Signal events were generated using `EOS` predictions for different q^2 bins. Toy studies were performed in three different configurations:

- pure signal without detector acceptance,

- mixture of signal and background without detector acceptance,
- mixture of signal and background with detector acceptance.

In all cases the method of moments did not shown any sign of bias, apart for the evident large bias in the value of $F_{\rm L}$. This bias, at the level of 0.3 standard deviations in the $(0.1 < q^2 < 0.98)$ GeV$^2/c^4$ bin, comes from neglecting lepton mass terms in the angular distribution and not the method itself.

4.8.5 Method of moments applied to $B^0 \to J/\psi\, K^{*0}$

In order to check our method with data, we use the decay $B^0 \to J/\psi\, K^{*0}$ as a control channel. The measured angular observables are presented in Tab. 4.7. A good agreement is observed in the P-wave angular observables between these measurements and the angular parameters determined in Ref. [79].

Table 4.7: Results of the angular fit of the decay $B^0 \to J/\psi\, K^{*0}$ in different bins of $m(K^+\pi^-)$, using the full available data set corresponding to $3\,\text{fb}^{-1}$. The angular terms that have been previously determined in Ref. [79] are given in Tab. 4.6.

	$m_{K\pi}$ range in MeV/c^2					
parameter	$[795.9, 995.9]$	$[825.9, 965.9]$	$[826.0, 861.0]$	$[861.0, 896.0]$	$[896.0, 931.0]$	$[931.0, 966.0]$
$F_{\rm L}$	0.558 ± 0.003	0.558 ± 0.003	0.566 ± 0.006	0.561 ± 0.004	0.549 ± 0.004	0.562 ± 0.005
S_3	0.000 ± 0.002	0.001 ± 0.002	-0.006 ± 0.006	0.000 ± 0.004	0.001 ± 0.003	0.004 ± 0.006
S_4	-0.280 ± 0.003	-0.282 ± 0.004	-0.278 ± 0.007	-0.288 ± 0.005	-0.279 ± 0.004	-0.275 ± 0.006
S_5	-0.002 ± 0.003	-0.002 ± 0.003	-0.004 ± 0.007	0.000 ± 0.005	-0.006 ± 0.003	0.003 ± 0.006
S_6^s	0.001 ± 0.003	0.002 ± 0.003	-0.004 ± 0.008	0.001 ± 0.003	0.003 ± 0.004	0.003 ± 0.005
S_7	0.001 ± 0.003	0.001 ± 0.003	-0.003 ± 0.007	0.001 ± 0.004	0.001 ± 0.004	0.007 ± 0.006
S_8	-0.053 ± 0.003	-0.054 ± 0.003	-0.072 ± 0.008	-0.058 ± 0.004	-0.051 ± 0.004	-0.047 ± 0.006
S_9	-0.089 ± 0.003	-0.088 ± 0.004	-0.089 ± 0.008	-0.086 ± 0.004	-0.091 ± 0.004	-0.086 ± 0.006
F_S	0.080 ± 0.004	0.068 ± 0.003	0.100 ± 0.012	0.053 ± 0.006	0.061 ± 0.005	0.108 ± 0.009
S_{S1}	-0.240 ± 0.004	-0.245 ± 0.004	-0.700 ± 0.01	-0.387 ± 0.007	-0.109 ± 0.006	0.160 ± 0.010
S_{S2}	0.003 ± 0.003	0.007 ± 0.003	0.140 ± 0.008	0.045 ± 0.004	-0.028 ± 0.004	-0.108 ± 0.006
S_{S3}	0.004 ± 0.003	0.004 ± 0.003	-0.005 ± 0.007	0.003 ± 0.003	0.004 ± 0.003	0.012 ± 0.006
S_{S4}	0.001 ± 0.003	0.001 ± 0.003	0.014 ± 0.008	-0.003 ± 0.003	0.000 ± 0.004	0.005 ± 0.006
S_{S5}	-0.065 ± 0.003	-0.061 ± 0.003	0.040 ± 0.008	-0.027 ± 0.004	-0.091 ± 0.004	-0.157 ± 0.007

4.8.6 Measuring asymmetries with the method of moments

The method described here is also used to measure the *CP*-asymmetries (the A_i). These observables are defined as the asymmetries of the corresponding J_i for B^0 and $\overline{B}^0$, normalised with respect to the total width $\Gamma_{\rm tot}$ as defined in Eq. 4.33. In order to measure these observables, the B^0 candidates (only) are multiplied by a factor (-1) for the angular terms $f_{i=4\ldots9}$ when determining the moments.

4.8.7 Expected difference between the likelihood fit and the method of moments

The method of moments estimator is strongly, but not completely, correlated to the maximum likelihood estimator. This correlation between the two estimators was studied using the `EOS` SM toy MC. Example scatter plots for S_5 and S_7 in one q^2 bin are given in Fig. 4.12. Whilst the distributions are strongly correlated, the spread of the data points is larger than one might naively expect (approximately 50% of the statistical uncertainty). This effect is mainly statistical and is largely independent from the level of background and from the acceptance effects.

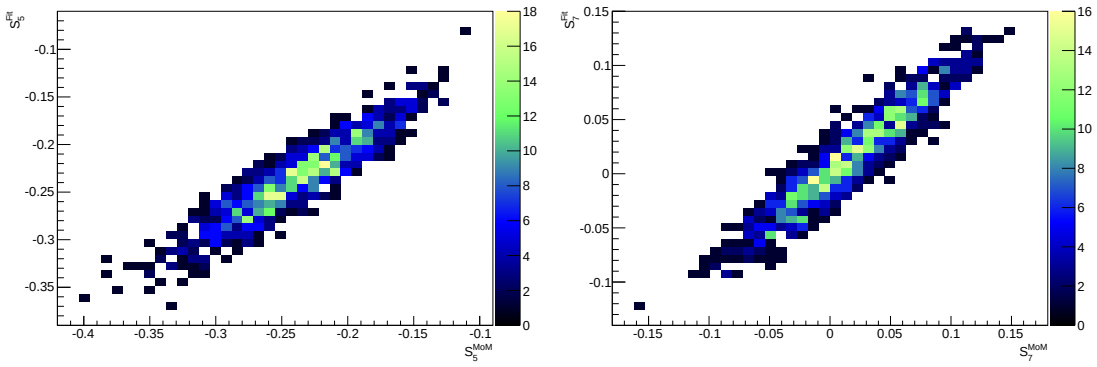


Figure 4.12: An example illustration of the difference between the maximum likelihood and method of moments, when applied to SM MC.

Small differences between the two estimators are also seen in the data. For a global comparison, we calculate the difference between the two estimates of each observable in every q^2 bin, then divide it by the expected difference from toy MC. The resulting distribution is consistent with having a mean of zero and a width of one, i.e. the moments/likelihood fit are consistent with each other when accounting for the expected differences between the two methods.

4.9 Systematic uncertainties

The systematic uncertainties are composed of effects that can distort the mass or angular distributions of either the signal or background candidates. They are estimates performing pseudoexperiments.

The main systematic uncertainties are related to:

- Four dimensional acceptance correction.
- Differences between data and MC.
- Peaking backgrounds.

- Signal mass modelling.
- B^0 production asymmetry.
- Detector asymmetry.

The acceptance corrections systematics uncertainty is composed of two parts. First one is related to the limited MC sample size that is used to calculate the efficiency function. The systematic is accessed by varying the parameters of the function accordingly to the calculated covariance matrix of the c_{ijkl} parameters (see Sec. 4.5). The second systematic is related to the choice of the maximum order of the Legendre polynomials. The systematic due to this choice is assigned by calculating a second acceptance function increasing the polynomial order by two. Then a high statistics pseudoexperiment is performed using two functions and the differences in the angular observable are assigned as systematic.

The second systematic is related to the difference between the data and MC simulation. As described in Sec. 4.5.1, the main differences are corrected in the analysis. Small differences remain in the p_T distributions of the four tracks. The systematic is assigned via calculating the new acceptance correlation with the additional reweighting of the p_T spectrum and performing a high statistics pseudoexperiment. The differences of the angular observables between the nominal acceptance and the one with the p_T correction are assigned as systematic.

The signal mass parametrisation has been explained in Sec. 4.7.2. The systematic is accessed by generating pseudoexperiments with nominal PDF (double Crystal Ball), which are then fitted with double Gaussian function. The averaged difference between the two fits is taken as s systematic uncertainty.

The last systematic is related to the B^0 production and detector asymmetries. The effects have been estimated by the LHCb collaboration in [92]. In the same study the detector asymmetry has been estimated. Both estimations are used as systematic uncertainties.

The systematic uncertainty related to the peaking backgrounds is similarly studied with the pseudoexperiments. The experiments are constructed by injecting the expected background events to the signal sample. The averaged differences in the angular observables are taken as a systematic uncertainty.

As a result it is found that the systematic uncertainties are significantly smaller than the statistical uncertainties (the largest total systematic being around 0.02, while the statistical errors are $0.08 - 0.10$) The dominant systematic uncertainty comes from the residual difference between data and MC. The systematics uncertainties will be added in quadrature to the statistic uncertainties for the final result.

4.10 Results

This section contains the results of the angular analysis on data. Because the analysis has produced over 80 tables with numerical results, we will not give them here in order to save space. They can be found in the publication [69] and have also been submitted by the

author to the HepData portal [93]. We will give the results in a graphical representation only.

A comparison of the values of the *CP*-averaged and *CP*-asymmetric angular observables obtained using the likelihood fit and the method of moments is presented in Figs. 4.13 and 4.14. As mentioned in Sec. 4.8.7, pseudoexperiments have been performed and they have demostrated that both the method of moments and likelihood fit exhibit an excellent agreement. The SM predictions for the *CP*-averaged are taken from Ref. [94][9]. There are currently no predictions for the *CP*-asymmetric observables but they are expected to be very small in the SM.

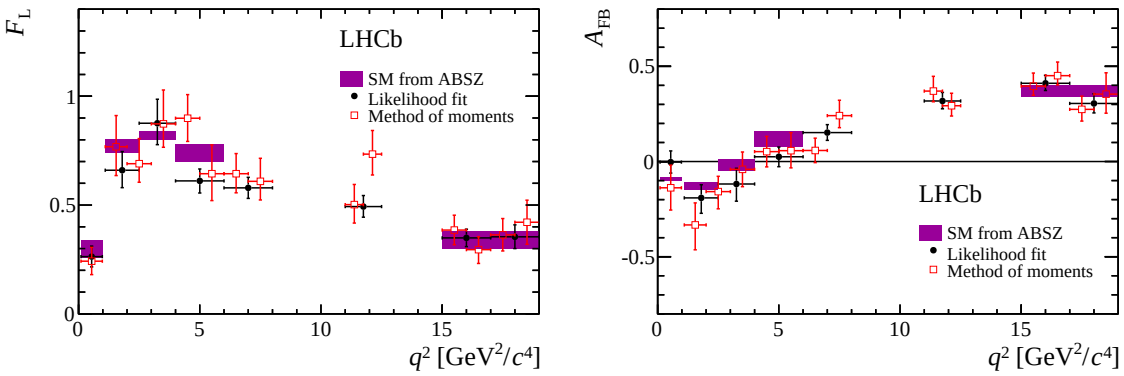


Figure 4.13: The *CP*-averaged observables in bins of q^2 determined in a maximum likelihood fit to the data and by the method of moments. The shaded boxes show the SM predictions based on the prescription of Ref. [94].

4.11 Comparison with other studies

The angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ has been also performed by the Belle Collaboration [95]. Recently, ATLAS [96] and CMS [97] have presented their preliminary measurements. Unfortunately, the results form the other three collaborations have a much larger statistical uncertainty. In spite of this, they seem to point to the same deviations from the SM as the results from LHCb, as can be seen in Fig. 4.15.

The other experiments could not perform the full angular analysis as LHCb did, because they had fewer $B^0 \to K^{*0}\mu^+\mu^-$ decays at their disposal. In this case the collaborations decided to use the folding technique [98], which integrates out some of the angles when performing the fit. In this way only a group of observables can be extracted at the same time and the correlation is lost.

[9]The authors do not give the predictions for $S_{7,8,9}$; however, those observables are expected to be extremely small in the SM.

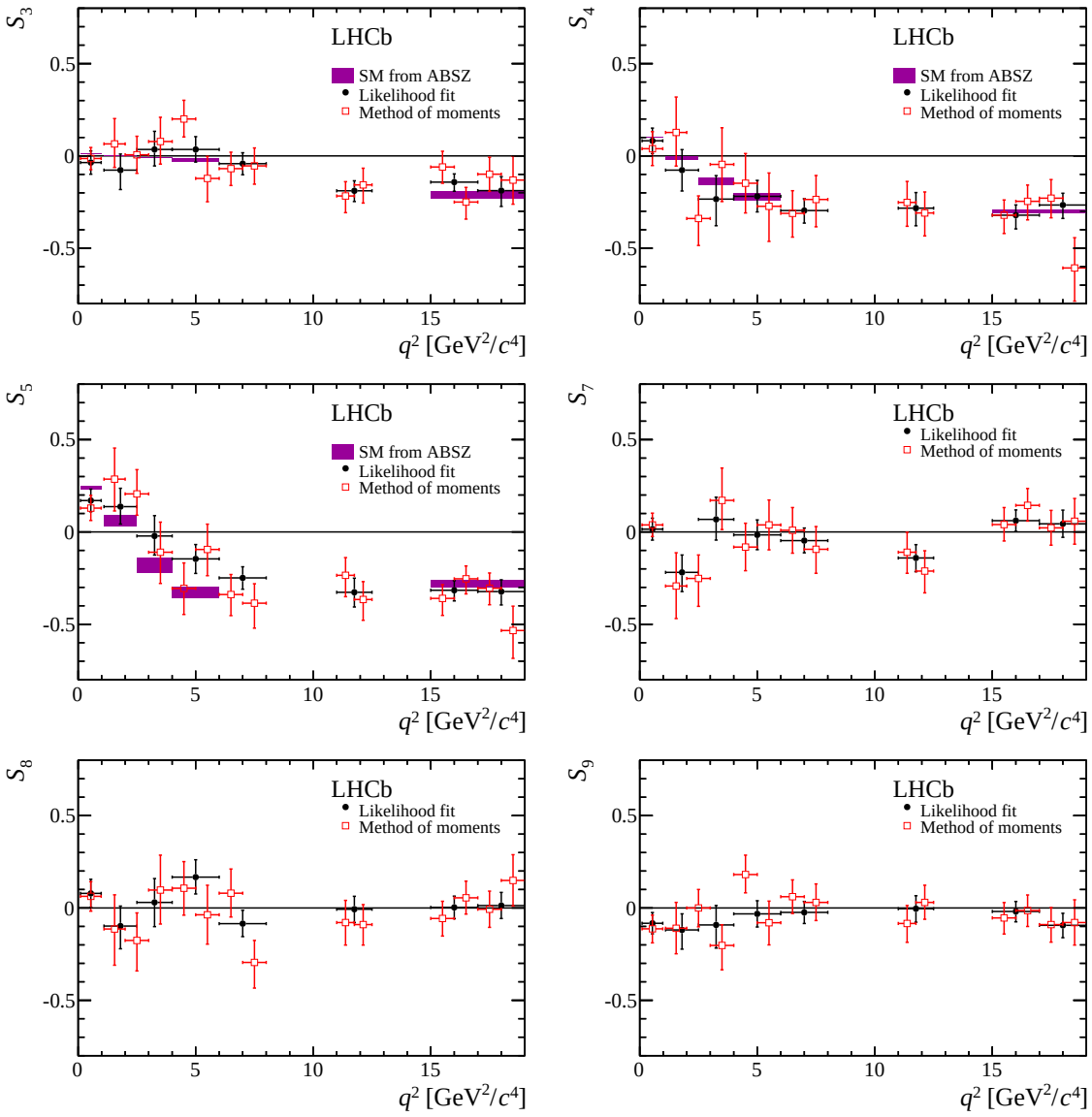


Figure 4.14: The *CP*-averaged observables in bins of q^2 determined in a maximum likelihood fit to the data and by the method of moments. The shaded boxes show the SM predictions based on the prescription of Ref. [94].

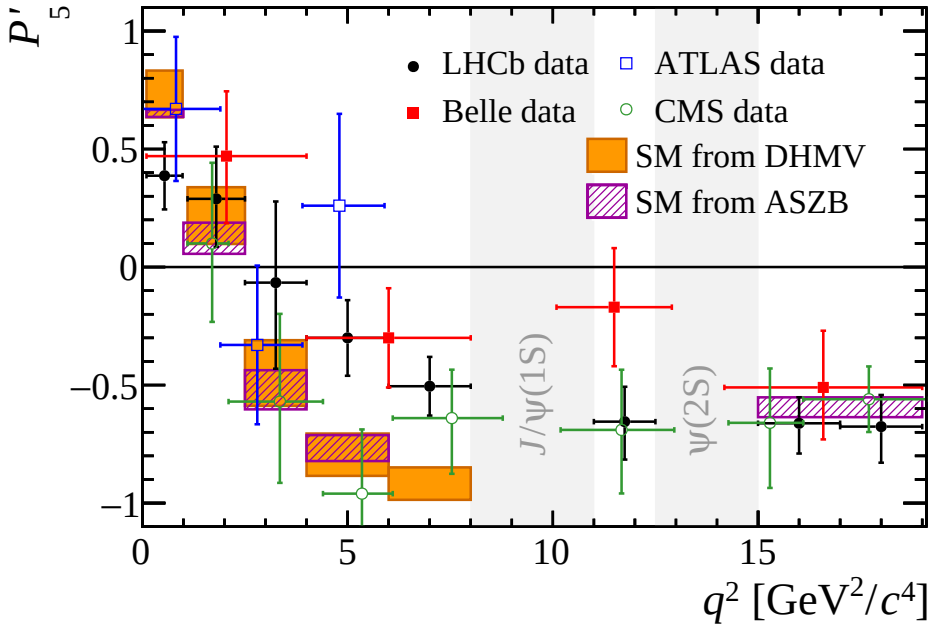


Figure 4.15: Comparison of P'_5 measurement by different collaborations..

Chapter 5

Branching ratio and angular analysis of $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region

In this section the branching ratio and angular analysis of $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ resonance region is presented. The analysis is similar to the one described in Sec. 4. Owing to this fact only differences in the selections will be presented. The results of this study have been published in [99].

Trigger requirements, selection, multivariate classifier requirements and systematic determinations are exactly the same as described in Sec. 4. The only difference is the $m_{K\pi}$ region that is selected: in this analysis we have chosen: 1330 MeV/$c^2 < m_{K\pi} <$ 1530 MeV/c^2. In this $m_{K\pi}$ range there are a number of K^{*0} states that can be present in this mass range that are listed in Table 5.1. As one can see, this region contains the S-wave, P-wave and D-wave states. It has been checked that the same BDT that has been applied to the $B^0 \to K^{*0}\mu^+\mu^-$ is equally highly discriminant to K^{*0} as to $K^*_{0,2}(1430)^0$.

Table 5.1: Known K^* states that can contribute to the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay over the $m_{K\pi}$ range of interest in this analysis. Adapted from Ref. [100].

	J^P	Mass (MeV/c^2)	Full width (MeV/c^2)	$\Gamma_{K\pi}/\Gamma$ (%)
$K^*(1410)$	1^-	1414 ± 15	232 ± 21	6.6 ± 1.3
$K^*_0(1430)$	0^+	1425 ± 50	270 ± 80	93 ± 10
$K^*_2(1430)$	2^+	1432.4 ± 1.3	109 ± 5	49.9 ± 1.2
$K^*(1680)$	1^-	1717 ± 27	322 ± 110	38.7 ± 2.5

The differential branching ratio measurement will be performed as a function of q^2 in the five q^2 bins. The angular analysis will be performed in a single q^2 bin, $[1.1, 6.0]$ GeV$^2/c^4$. The analysis will consist of measuring 40 normalized angular moments, $M_2 - M_{41}$ [101], corresponding to the $[K\pi]$ system including partial waves up to spin-2 (D-wave).

5.1 Agreement between data and simulation

Similarly to the aforementioned analysis of $B^0 \to K^{*0}\mu^+\mu^-$, a good agreement between data and simulation is necessary in order to be able to accurately model the distortion of the angular distributions caused by the trigger, reconstruction and selection. The acceptance correction, described in detail in Sec. 5.3, is determined from simulated flat phase space $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates. Data driven techniques are used to improve the agreement between data and simulation (the same way as in Sec. 4.5.2). The PID distributions in simulation are corrected using a method known as "resampling". The procedure has already been described in the previous section and will not be repeated here. In this case, besides checking the agreement on $B^0 \to J/\psi K^{*0}$, an additional check is made on the $B^0 \to J/\psi K\pi$. As can be seen in Fig. 5.1, also in the second case a good agreement is observed.

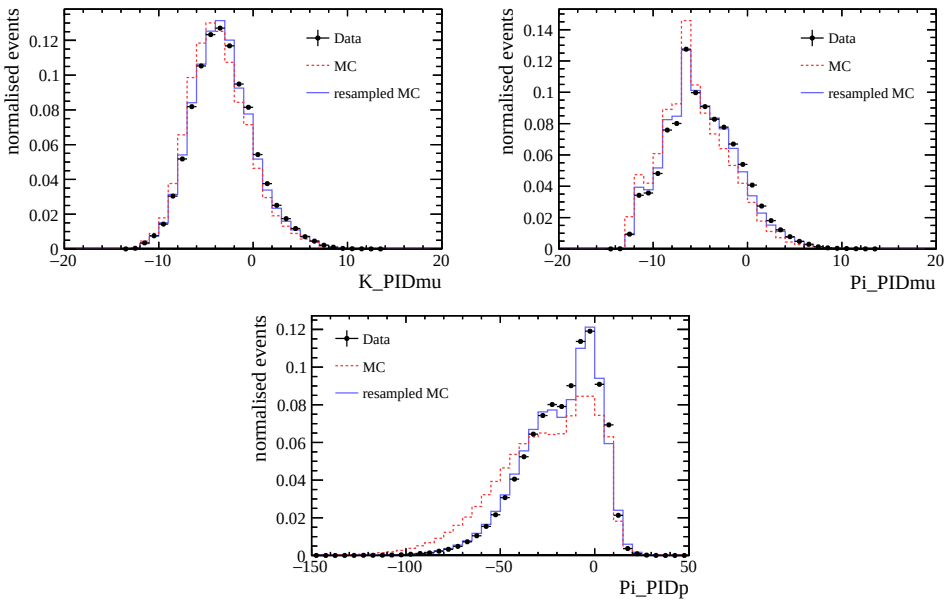


Figure 5.1: Data-simulation agreement for the PID variables used to veto peaking backgrounds. The black data points show the distributions for *sWeighted* $B^0 \to J/\psi K^{*0}$ data. The red dashed histograms show the nominal distribution in $B^0 \to J/\psi K^{*0}$ MC. The blue histograms show the distribution for the resampled $B^0 \to J/\psi K^{*0}$ MC.

5.2 Mass fits

In this analysis, the $K\pi\mu\mu$ invariant mass distribution is described by a signal model comprising a double Crystal Ball function with common mean (μ) and tail parameters (α, n) and a background model comprising of an exponential function, in line with [69].

In order to help to constrain the invariant mass fit to the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates, the number of degrees of freedom is reduced by first fitting the high statistics control mode $B^0 \to J/\psi\, K^{*0}$. For the control mode, an additional component is included to model the contribution from $B_s^0 \to J/\psi\, K^{*0}$. The signal fit parameters determined from the fit to $B^0 \to J/\psi\, K^{*0}$ are fixed when performing the fit to $B^0 \to K^+\pi^-\mu^+\mu^-$.

To take into account possible variations in the signal mass lineshape due to the different region of $m_{K\pi}$ and q^2 in the fits to $B^0 \to J/\psi\, K^{*0}$ ($q^2 \in [9.22, 9.96]$ GeV$^2/c^4$, $m_{K\pi} \in [796, 996]$ MeV$/c^2$) and $B^0 \to K^+\pi^-\mu^+\mu^-$ ($q^2 \in [0.1, 8.0]$ GeV$^2/c^4$, $m_{K\pi} \in [1330, 1530]$ MeV$/c^2$), a single scaling factor, s_σ, is applied to both σ_1 and σ_2 during the fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates. This factor is determined by first fitting the $m_{K\pi\mu\mu}$ distribution in the $B^0 \to K^+\pi^-\mu^+\mu^-$ flat phase space MC in the $B^0 \to J/\psi\, K^{*0}$ region. All fit parameters are then fixed except for s_σ, which is allowed to float in the subsequent fits to the $m_{K\pi\mu\mu}$ distribution in each of the q^2 bins in the $B^0 \to K^+\pi^-\mu^+\mu^-$ signal region. The numerical values are given in Tab. 5.2.

q^2 [GeV$^2/c^4$]	Scaling factor
$[0.10, 0.98]$	0.972 ± 0.004
$[1.10, 2.50]$	0.981 ± 0.003
$[2.50, 4.00]$	0.980 ± 0.003
$[4.00, 6.00]$	0.987 ± 0.002
$[6.00, 8.00]$	0.981 ± 0.002
$[1.10, 6.00]$	0.983 ± 0.002
$[9.22, 9.96]$	0.989 ± 0.004

Table 5.2: Scaling factor s_σ in bins of q^2 for candidates in $m_{K\pi} \in [1330, 1530]$ MeV$/c^2$.

In order to cross check the method, the scaling factor is determined both from $B^0 \to K^+\pi^-\mu^+\mu^-$ flat phase space MC and from data in the region $q^2 \in [9.22, 9.96]$ GeV$^2/c^4$, $m_{K\pi} \in [1330, 1530]$ MeV$/c^2$. The scaling factor determined from MC is 0.989 ± 0.004, which is in good agreement with the scaling factor determined from data 0.989 ± 0.005.

The invariant mass fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in each of the q^2 bins are shown in Fig. 5.2. The signal and background yields in each of the q^2 bins are given in Tab 5.3 for the full range of $m_{K\pi\mu\mu} \in [5170, 5700]$.

5.3 Acceptance correction

Similarly to the previous analysis, the triggering, reconstruction and selection of signal candidates all cause distortions to the distributions of q^2, $\cos\theta_\ell$, $\cos\theta_K$, ϕ and $m_{K\pi}$, and this needs to be corrected for. This effect will be accounted for by reweighting each event by the inverse of the efficiency:

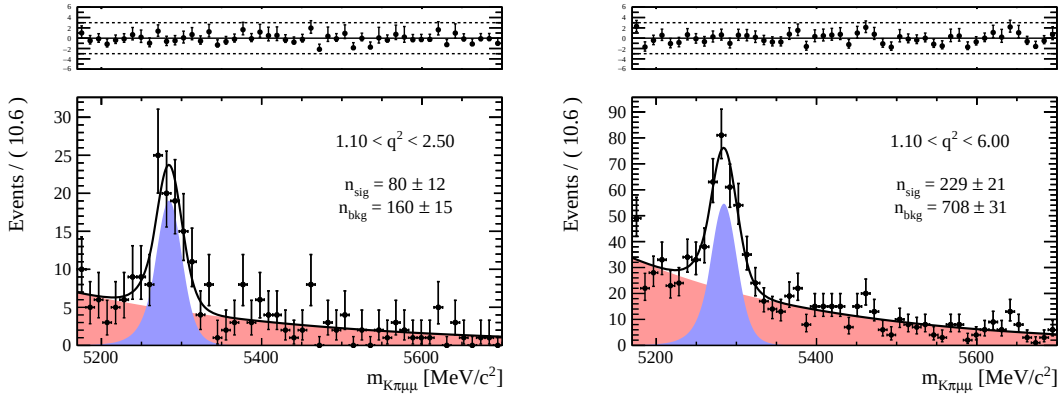


Figure 5.2: Example fits to the $K\pi\mu\mu$ invariant mass of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in q^2 bins.

q^2 [GeV$^2/c^4$]	Signal yield	Background yield
$[0.10, 0.98]$	67 ± 10	93 ± 11
$[1.10, 2.50]$	80 ± 12	160 ± 15
$[2.50, 4.00]$	75 ± 12	213 ± 17
$[4.00, 6.00]$	75 ± 13	334 ± 21
$[6.00, 8.00]$	60 ± 14	476 ± 25
$[1.10, 6.00]$	229 ± 21	708 ± 31

Table 5.3: Signal and background yields of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in q^2 bins

$$w = \frac{1}{\varepsilon(q^2, \cos\theta_\ell, \cos\theta_K, \phi, m_{K\pi})}, \tag{5.1}$$

where $\varepsilon(q^2, \cos\theta_\ell, \cos\theta_K, \phi, m_{K\pi})$ is a five dimensional efficiency parametrisation determined from the simulated flat phase space events. The main change compared to the previous measurement of $B^0 \to K^{*0}\mu^+\mu^-$ is the 5^{th} dimension, which is the $m_{K\pi}$. The change has to be made as in the wider $m_{K\pi}$ window of this analysis the efficiency is no longer flat.

As the acceptance correction will be used for both the differential branching fraction measurement and the angular analysis, it is calculated from $q^2 \in [0.1, 10.0]$ GeV$^2/c^4$ GeV$^2/c^4$ and $m_{K\pi} \in [795, 1530]$ MeV$/c^2$ in order to encompass the full range of invariant masses studied.

5.3.1 Legendre polynomial parameterisation

As described in [69], the efficiency is parametrised as the product of Legendre polynomials:

$$\varepsilon(q^{2'}, \cos\theta_\ell, \cos\theta_K, \phi', m'_{K\pi}) = \sum_{h,i,j,k,l} c_{h,i,j,k,l}\ P(q^{2'}, h)P(\cos\theta_\ell, i)P(\cos\theta_K, j)P(\phi', k)P(m'_{K\pi}, l), \tag{5.2}$$

where $P(x, a)$ are the Legendre polynomials in x of order a and $-1 < x < 1$. Owing to the range of the Legendre polynomials the variables $\phi \in [-\pi, \pi]$, $q^2 \in [0.1, 10.0]$ GeV$^2/c^4$ and $m_{K\pi} \in [0.795, 1.53]$ GeV$/c^2$ are transformed into the correct ranges $[-1, 1]$. The coefficients, $c_{h,i,j,k,l}$, are determined from the $B^0 \to K^+\pi^-\mu^+\mu^-$ flat phase space MC using the moments method:

$$c_{h,i,j,k,l} = \frac{1}{\sum w_n}\sum_{n=0}^{N} w_n \left(\frac{2h+1}{2}\right)\left(\frac{2i+1}{2}\right)\left(\frac{2j+1}{2}\right)\left(\frac{2k+1}{2}\right)\left(\frac{2l+1}{2}\right) \times P(q^{2'}, h)P(\cos\theta_\ell, i)P(\cos\theta_K, j)P(\phi', k)P(m'_{K\pi}, l), \tag{5.3}$$

where w_n is the per-event weight taking into account both the non-flat phase space distribution of q^2, $m_{K\pi}$ and the kinematic event weights. The factors of $(2a + 1)/2$ arise from the orthonormality of the Legendre polynomials,

$$\int_{-1}^{+1} P(x, a)P(x, a')\mathrm{d}x = \frac{2}{2a+1}\delta_{aa'}. \tag{5.4}$$

The order of polynomial used in each case is determined as the lowest order possible that gives a good agreement between the parametrisation and the simulated flat phase space events. These are: $n_{q^2} < 7$, $n_{\cos\theta_\ell} < 5$, $n_{\cos\theta_K} < 9$, $n_\phi < 7$, $n_{m_{K\pi}} < 5$. Several relations are used to reduce the number of coefficients necessary to be determined. Both $\cos\theta_\ell$ and ϕ are assumed to be even and as such only even orders are used. Also, $\cos\theta_\ell$ and $m_{K\pi}$ are assumed to factorise. Therefore, any term containing both $\cos\theta_\ell$ and $m_{K\pi}$ is ignored. This reduces the terms along with the number of coefficients to be determined from 11025 to 1764.

5.3.2 Covariance matrix

The covariance matrix of the legendre coefficients is also determined. The variance of the coefficients is defined as

$$s^2 = \frac{\sum_i w_i(x_i - \bar{x})(y_i - \bar{y})}{\sum_i w_i} \tag{5.5}$$

and the variance on the mean is defined as

$$u^2 = \frac{s^2}{n_{\text{eff}}}, \tag{5.6}$$

where n_{eff} is the effective number of events

$$n_{\text{eff}} = \frac{(\sum_i w_i)^2}{\sum_i w_i^2}. \tag{5.7}$$

The covariance matrix will be used to determine the systematic uncertainty, which is due to the limited statistics MC sample used to calculate the legendre coefficients.

5.4 Differential branching fraction measurement

The differential branching fraction for $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in the $m_{K\pi}$ range $m_{K\pi} \in [1330, 1530]$ MeV/c^2 will be measured in five q^2 bins. The decay $B^0 \to J/\psi\, K^{*0}$ is used as the normalisation channel with candidates selected in the ranges $q^2 \in [9.22, 9.96]$ GeV$^2/c^4$ and $m_{K\pi} \in [796, 996]$ MeV/c^2.

The partial branching fraction, BR($B^0 \to K^+\pi^-\mu^+\mu^-$), in a specific q^2 bin can be estimated by comparing the yield of $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in that q^2 bin to the number of $B^0 \to J/\psi\, K^{*0}$ candidates. The partial branching fraction is then given by

$$\begin{aligned}\text{BR}(B^0 \to K^+\pi^-\mu^+\mu^-) = f_{K^*(892)^0} \times \text{BR}(B^0 \to J/\psi\, K^*(892)^0) \times \text{BR}(J/\psi \to \mu^+\mu^-) \\ \times \text{BR}(K^*(892)^0 \to K^+\pi^-) \frac{N'_{K\pi\mu\mu}}{(1 - F_S^{J/\psi\, K^{*0}}) N'_{J/\psi\, K^{*0}}},\end{aligned}$$

where $N'_{K\pi\mu\mu;k}$ is the number of acceptance corrected $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates in bin k, $N'_{J/\psi\, K^{*0}}$ is the number of acceptance corrected $B^0 \to J/\psi\, K^{*0}$ candidates in the full data sample. The number of $B^0 \to J/\psi\, K^{*0}$ candidates is corrected for the S-wave fraction, $F_S^{J/\psi\, K^{*0}}$. The fraction $f_{K^*(892)^0}$ is used to scale the value of BR($B^0 \to J/\psi\, K^*(892)^0$) to the $m_{K\pi}$ range $796 < m_{K\pi} < 996$ MeV/c^2.

5.4.1 Acceptance corrected yields

To avoid making any assumptions about the unknown distributions of the $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates, the event-by-event efficiencies described in Sec. 5.3 are used to correct the measured yields by calculating the average acceptance weight, where each weight is the reciprocal of the event-by-event efficiency.

In the case where there are only signal candidates present, the average weight would simply be calculated as

$$\overline{w} = \frac{1}{N} \sum_i^N w_i, \tag{5.8}$$

where w_i is the event-by-event acceptance and N is the number of candidates. An estimate for the error on the average weight is given by

$$\delta_{\overline{w}} = \sqrt{\frac{1}{N(N-1)}\sum_{i}^{N}(w_i - \overline{w})^2}. \tag{5.9}$$

Because of the presence of background, the average weight calculated in the signal mass window will be an admixture of the average weight for both signal candidates ($\overline{w}_{sig}$) and background candidates ($\overline{w}_{bkg}$),

$$\overline{w}_{mix} = \frac{N_{sig}\overline{w}_{sig} + N_{bkg}\overline{w}_{bkg}}{N_{sig} + N_{bkg}}, \tag{5.10}$$

where N_{sig} and N_{bkg} are the numbers of signal and background events in the signal mass window, respectively. This can be rearranged to give the average weight for the signal candidates,

$$\overline{w}_{sig} = \frac{(N_{sig} + N_{bkg})\overline{w}_{mix} - N_{bkg}\overline{w}_{bkg}}{N_{sig}}. \tag{5.11}$$

However, what is needed for both $B^0 \to K^+\pi^-\mu^+\mu^-$ and $B^0 \to J/\psi\, K^{*0}$ is the acceptance corrected yield $\overline{w}_{sig}N_{sig}$. This is given by

$$\overline{w}_{sig}N_{sig} = (N_{sig} + N_{bkg})\overline{w}_{mix} - N_{bkg}\overline{w}_{bkg}, \tag{5.12}$$

where the errors are propagated as

$$\sigma^2_{\overline{w}_{sig}N_{sig}} = (N_{sig}+N_{bkg})^2\sigma^2_{\overline{w}_{mix}} + (-N_{bkg})^2\sigma^2_{\overline{w}_{bkg}} + (\overline{w}_{mix})^2\sigma^2_{N_{sig}} + (\overline{w}_{mix}-\overline{w}_{bkg})^2\sigma^2_{N_{bkg}}. \tag{5.13}$$

The signal region is defined as $m_{K\pi\mu\mu} \in [5230, 5330]$ MeV/c^2 and the background region as $m_{K\pi\mu\mu} \in [5350, 5700]$ MeV/c^2. For the resonant mode, the background region is altered to $m_{K\pi\mu\mu} \in [5450, 5700]$ MeV/c^2 in order to prevent any potential pollution from $B^0 \to J/\psi\, K^{*0}$ or $B^0_s \to J/\psi\, K^{*0}$ candidates.

5.4.2 Pseudoexperiments studies

Toy studies are performed for the extraction of $\overline{w}_{sig}N_{sig}$ with different numbers of signal and background candidates. In each toy N_{sig}, N_{bkg} are Poisson fluctuated. The nominal mass models, described in Sec. 5.2, are used to generate signal and background candidates. The weights for both signal and background are sampled from two gaussian functions with different means. The pulls for the extraction of $\overline{w}_{sig}N_{sig}$ are shown in Fig. 5.3. No bias is observed.

The extracted values of the signal yield N_{sig}, the average signal weight $\overline{w}_{sig}$ and the acceptance corrected yield $\overline{w}_{sig}N_{sig}$ for both the $B^0 \to J/\psi\, K^{*0}$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates are shown in Tab. 5.4.

The resulting differential branching fraction in the q^2 bin is then given by:

$$\frac{d\mathrm{BR}(B^0 \to K^+\pi^-\mu^+\mu^-)}{dq^2} = \frac{1}{q^2_{\mathrm{max.};k} - q^2_{\mathrm{min.};k}}\mathrm{BR}(B^0 \to K^+\pi^-\mu^+\mu^-) \quad .$$

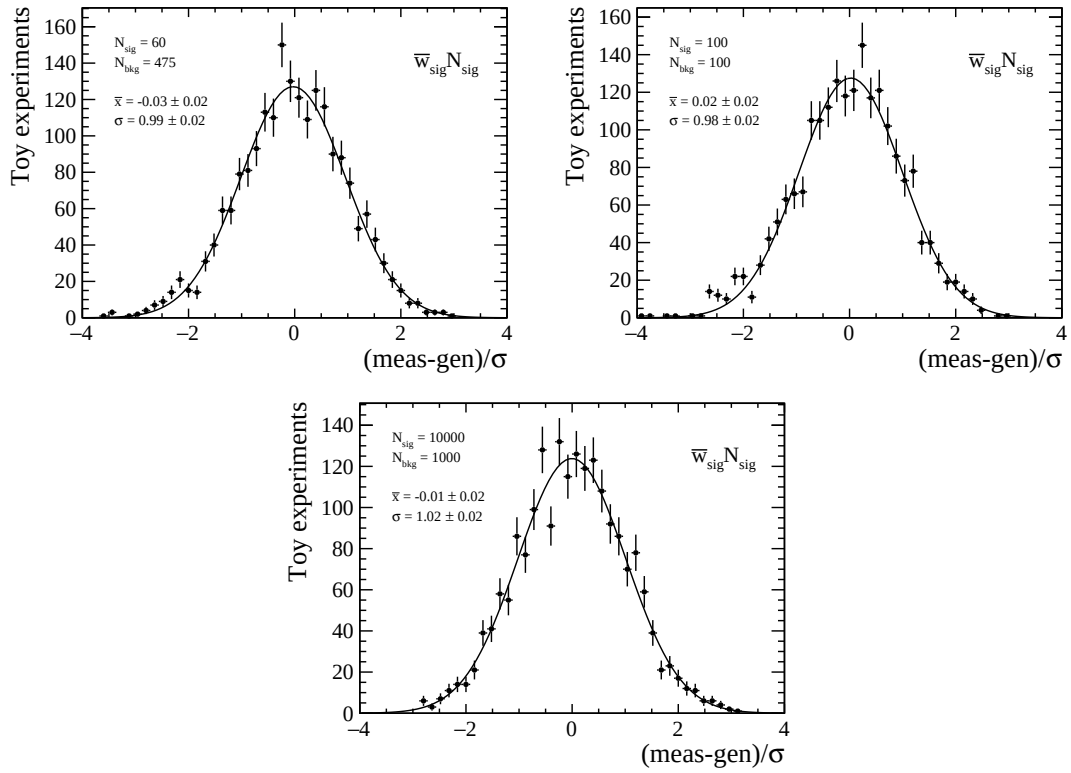


Figure 5.3: Pull plots for the extraction of $N_{sig}/\overline{\varepsilon}_{sig}$ with different numbers of signal and background candidates.

bin	N_{sig}	$\overline{w}_{sig}$	$\overline{w}_{sig}N_{sig}$
$B^0 \to J/\psi\, K^{*0}$	304491.75	0.88	266650.04 ± 560.23
$[0.10, 0.98]$	65.18	1.25	81.44 ± 14.04
$[1.10, 2.50]$	77.34	1.19	92.25 ± 15.61
$[2.50, 4.00]$	72.49	1.09	79.09 ± 13.43
$[4.00, 6.00]$	72.66	0.90	65.16 ± 13.45
$[6.00, 8.00]$	57.91	0.97	56.08 ± 13.22
$[1.10, 6.00]$	222.24	1.05	232.68 ± 24.59

Table 5.4: The signal yield (N_{sig}), the average signal weight $\overline{w}_{sig}$ and the acceptance corrected yield $\overline{w}_{sig}N_{sig}$ for both the $B^0 \to J/\psi\, K^{*0}$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ candidates.

5.4.3 Branching fraction results

Some external inputs are needed in order to determine the branching fraction, which will be discussed in this section. The first input is the branching fraction for the normalisation channel $B^0 \to J/\psi\, K^{*0}$, which is taken from [102],

Table 5.5: Differential branching fraction of $B^0 \to K^+\pi^-\mu^+\mu^-$ in bins of q^2 for the range $1330 < m_{K\pi} < 1530\,\mathrm{MeV}/c^2$. The first uncertainty is statistical, the second is systematic, and the third is due to the uncertainty on the $B^0 \to J/\psi\, K^*(892)^0$ and $J/\psi \to \mu^+\mu^-$ branching fractions.

q^2 $[\mathrm{GeV}^2/c^4]$	$\mathrm{dBR}/\mathrm{d}q^2 \times 10^{-8}$ $[c^4/\mathrm{GeV}^2]$
$[0.10, 0.98]$	$1.60 \pm 0.28 \pm 0.04 \pm 0.11$
$[1.10, 2.50]$	$1.14 \pm 0.19 \pm 0.03 \pm 0.08$
$[2.50, 4.00]$	$0.91 \pm 0.16 \pm 0.03 \pm 0.06$
$[4.00, 6.00]$	$0.56 \pm 0.12 \pm 0.02 \pm 0.04$
$[6.00, 8.00]$	$0.49 \pm 0.11 \pm 0.01 \pm 0.03$
$[1.10, 6.00]$	$0.82 \pm 0.09 \pm 0.02 \pm 0.06$

$$\mathrm{BR}(B^0 \to J/\psi\, K^{*0}) = (1.19 \pm 0.01 \pm 0.08) \times 10^{-3}. \tag{5.14}$$

The remaining branching fraction is taken from the PDG average,

$$\mathrm{BR}(J/\psi \to \mu^+\mu^-) = (5.961 \pm 0.033)\%. \tag{5.15}$$

The value of $F_S^{J/\psi\, K^{*0}}$ is obtained from Ref. [79] and is adjusted to the $m_{K\pi}$ range $m_{K\pi} \in [796, 996]$ MeV/c^2, which leads to:

$$F_S^{J/\psi\, K^{*0}} = 0.084 \pm 0.01. \tag{5.16}$$

The value of $f_{K^*(892)^0}$ is calculated by integrating the $K^*(892)^0$ lineshape given in Ref. [102] in the $m_{K\pi}$ range $796 < m_{K\pi} < 996$ MeV/c^2,

$$f_{K^*(892)^0} = 0.894 \pm 0.001. \tag{5.17}$$

Taking the above numerical values and the results of our analysis, one can calculate the differential branching fraction, which is shown in Table 5.5 and Fig. 5.4. This is the first observation of this decay.

5.5 Extracting the angular coefficients with the method of moments for $B^0 \to K^+\pi^-\mu^+\mu^-$

The angular formalism used in this analysis is described in detail in [101]. The most relevant information is summarised in the following section. The four-differential decay rate can be written as:

$$\frac{1}{\mathrm{d}(\Gamma + \bar{\Gamma})/\mathrm{d}q^2} \frac{\mathrm{d}(\Gamma + \bar{\Gamma})}{\mathrm{d}\cos\theta_l\, \mathrm{d}\cos\theta_K\, \mathrm{d}\phi} = \sum_i b_i f_i(\Omega), \tag{5.18}$$

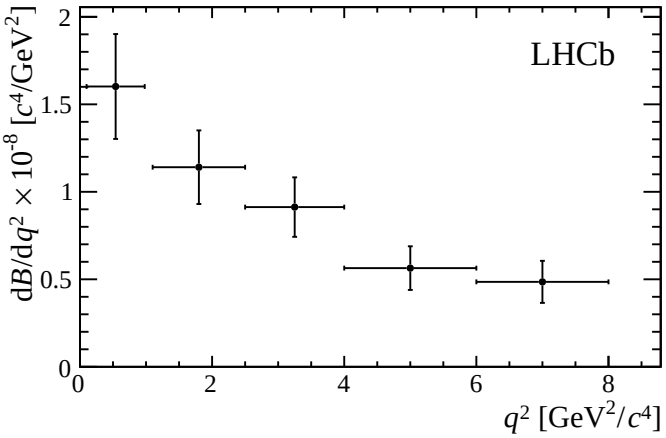


Figure 5.4: Differential branching fraction of $B^0 \to K^+\pi^-\mu^+\mu^-$ in bins of q^2 for the range $1330 < m_{K\pi} < 1530\,\mathrm{MeV}/c^2$. The error bars indicate the sums in quadrature of the statistical and systematic uncertainties.

where f_i are spherical harmonics (see [101] for details). The moments b_i can be measured from the data as

$$b_i = \tilde{f}_i = \sum_{k=1}^{N_{\mathrm{data}}} f_i(\Omega_k), \tag{5.19}$$

while the corresponding covariance can be measured from the data as

$$\tilde{C}_{ij} = \sum_{k=1}^{N_{\mathrm{data}}} f_i(\Omega_k) f_j(\Omega_k). \tag{5.20}$$

In the presence of background, $\tilde{f}_i$ and $\tilde{C}_{ij}$ become

$$\begin{aligned} \tilde{f}_i &= \left(\sum_{k=1}^{N_{\mathrm{sig}}} f_i(\Omega_k)\right) - x\left(\sum_{k=1}^{N_{\mathrm{bkgd}}} f_i(\Omega_k)\right), \\ \tilde{C}_{ij} &= \left(\sum_{k=1}^{N_{\mathrm{sig}}} f_i(\Omega_k) f_j(\Omega_k)\right) + x^2\left(\sum_{k=1}^{N_{\mathrm{bkgd}}} f_i(\Omega_k) f_j(\Omega_k)\right), \end{aligned} \tag{5.21}$$

where x is the ratio of the number of background events in the signal region over the number of background events in the background region, $N_{\mathrm{bkgd}}^{\mathrm{sig}}/N_{\mathrm{bkgd}}^{\mathrm{bkgd}}$, and is used for the background subtraction. Furthermore, in the presence of detector inefficiency, $\tilde{f}_i$ and $\tilde{C}_{ij}$ become

$$\begin{aligned}\tilde{f}_i &= \left(\sum_{k=1}^{N_{\rm sig}} w_k f_i(\Omega_k)\right) - x\left(\sum_{k=1}^{N_{\rm bkgd}} w_k f_i(\Omega_k)\right), \\ \tilde{C}_{ij} &= \left(\sum_{k=1}^{N_{\rm sig}} w_k^2 f_i(\Omega_k) f_j(\Omega_k)\right) + x^2\left(\sum_{k=1}^{N_{\rm bkgd}} w_k^2 f_i(\Omega_k) f_j(\Omega_k)\right),\end{aligned} \tag{5.22}$$

where w_k is the event-by-event weight that accounts for acceptance effects.

In this analysis the moments, $\tilde{f}_i$, and their corresponding covariance matrix, C_{ij}, will be determined using Eq. 5.22. The ratio x is determined in each q^2 bin from a maximum likelihood fit to the $m_{K\pi\mu\mu}$ distribution. The signal region is defined as $m_{K\pi\mu\mu} \in [5230, 5330]$ MeV/c^2 and the background region as $m_{K\pi\mu\mu} \in [5350, 5700]$ MeV/c^2. The event-by-event weights are obtained from the acceptance parametrisation described in Sec. 5.3.

5.6 Normalised moments and the reduced covariance matrix

In this analysis, $\tilde{f}_1$ corresponds to the sum of the weights, so we define the normalized moments for $i \in \{2, ..., 41\}$ as

$$M_i = \frac{\tilde{f}_i}{\tilde{f}_1}. \tag{5.23}$$

The covariance matrix on the set of normalised moments can be estimated as

$$\bar{C}_{ij} = \left[\tilde{C}_{ij} + \frac{\tilde{f}_i\tilde{f}_j}{\tilde{f}_1^2}\tilde{C}_{11} - \frac{\tilde{f}_i\tilde{C}_{1j} + \tilde{f}_j\tilde{C}_{1i}}{\tilde{f}_1}\right]\frac{1}{\tilde{f}_1^2}, \quad i, j \in [2, 41]. \tag{5.24}$$

The set of 40 normalized moments and the full set of 40×40 covariance matrix elements comprise the main results we aim to publish for the angular analysis.

5.7 Angular analysis results

The results for the angular analysis are shown in Tab. 5.6 and Fig. 5.5. The covariance and correlation matrices for the measured moments are shown in Fig. 5.6. The systematic uncertainness has been evaluated in the same way as in the previous analysis (see Sec. 6.5). They have been found to be much below the statistical uncertainty as well.

Unfortunately at this moment no SM predictions exist for the measured observables. This is mostly due to the lack of knowledge of form factors of the higher K^{*0} states. With development on the theory side, this decay should have sensitivity to the same Wilson Coefficients as $B^0 \to K^{*0}\mu^+\mu^-$.

$\overline{\Gamma}_i$	Value	$\overline{\Gamma}_i$	Value
$\overline{\Gamma}_2$	$-0.42 \pm 0.13 \pm 0.03$	$\overline{\Gamma}_{22}$	$0.21 \pm 0.12 \pm 0.01$
$\overline{\Gamma}_3$	$-0.38 \pm 0.15 \pm 0.01$	$\overline{\Gamma}_{23}$	$0.03 \pm 0.12 \pm 0.01$
$\overline{\Gamma}_4$	$-0.02 \pm 0.14 \pm 0.01$	$\overline{\Gamma}_{24}$	$-0.10 \pm 0.10 \pm 0.01$
$\overline{\Gamma}_5$	$0.29 \pm 0.14 \pm 0.02$	$\overline{\Gamma}_{25}$	$0.03 \pm 0.10 \pm 0.01$
$\overline{\Gamma}_6$	$-0.05 \pm 0.14 \pm 0.04$	$\overline{\Gamma}_{26}$	$0.08 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_7$	$-0.06 \pm 0.15 \pm 0.03$	$\overline{\Gamma}_{27}$	$0.14 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_8$	$0.04 \pm 0.16 \pm 0.01$	$\overline{\Gamma}_{28}$	$-0.04 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_9$	$0.05 \pm 0.16 \pm 0.02$	$\overline{\Gamma}_{29}$	$0.06 \pm 0.15 \pm 0.04$
$\overline{\Gamma}_{10}$	$0.24 \pm 0.17 \pm 0.02$	$\overline{\Gamma}_{30}$	$-0.21 \pm 0.15 \pm 0.04$
$\overline{\Gamma}_{11}$	$0.06 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{31}$	$-0.07 \pm 0.16 \pm 0.01$
$\overline{\Gamma}_{12}$	$-0.01 \pm 0.13 \pm 0.02$	$\overline{\Gamma}_{32}$	$-0.16 \pm 0.17 \pm 0.02$
$\overline{\Gamma}_{13}$	$-0.08 \pm 0.12 \pm 0.01$	$\overline{\Gamma}_{33}$	$-0.04 \pm 0.17 \pm 0.02$
$\overline{\Gamma}_{14}$	$0.09 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{34}$	$0.15 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{15}$	$0.11 \pm 0.13 \pm 0.00$	$\overline{\Gamma}_{35}$	$-0.13 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{16}$	$-0.12 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{36}$	$0.05 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{17}$	$-0.04 \pm 0.13 \pm 0.01$	$\overline{\Gamma}_{37}$	$0.05 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{18}$	$0.03 \pm 0.14 \pm 0.01$	$\overline{\Gamma}_{38}$	$0.06 \pm 0.11 \pm 0.00$
$\overline{\Gamma}_{19}$	$0.11 \pm 0.11 \pm 0.01$	$\overline{\Gamma}_{39}$	$-0.08 \pm 0.11 \pm 0.00$
$\overline{\Gamma}_{20}$	$-0.00 \pm 0.11 \pm 0.01$	$\overline{\Gamma}_{40}$	$0.15 \pm 0.11 \pm 0.01$
$\overline{\Gamma}_{21}$	$0.03 \pm 0.12 \pm 0.01$	$\overline{\Gamma}_{41}$	$0.12 \pm 0.11 \pm 0.01$

Table 5.6: Results of the angular analysis. The first uncertainty is statistical and the second one is systematic.

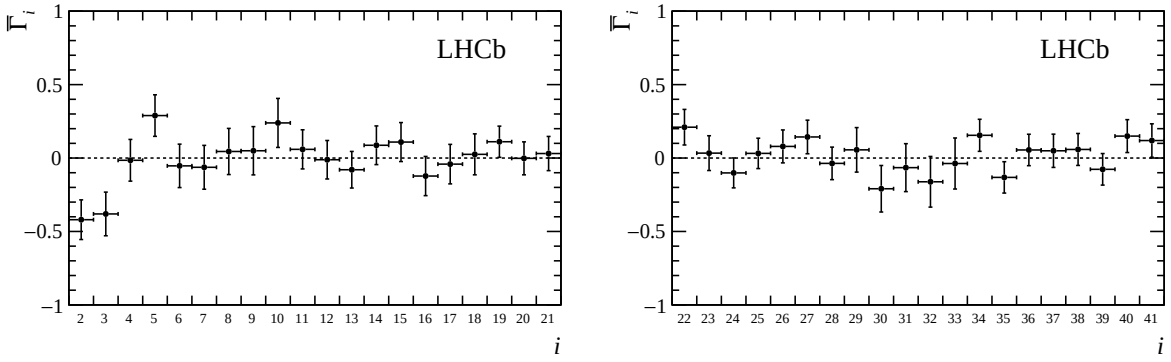


Figure 5.5: Results of the angular analysis.

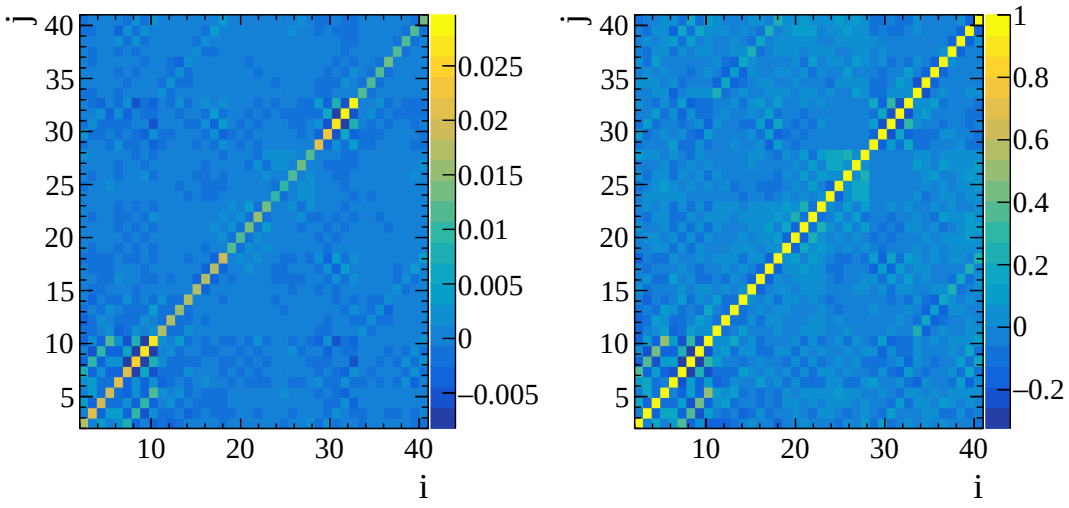


Figure 5.6: Covariance and correlation matrices for the measured moments.

Chapter 6

Search for long-lived particles decaying into two muons

In this chapter a search for a light long-lived particle χ produced in the $b \to s$ transitions will be presented. The result of this study have been published in [103]. The channel that the search will be performed in is: $B^+ \to K^+\chi$, where χ decays to two opposite charged muons. Because of its different topology, when compared to the previously discussed analysis, the selection had to be adjusted and will be discussed in detail in this chapter. This kind of long-lived particles makes them perfect candidates of dark matter. A possible extension of the SM is to add dark matter particles that are gauge singlets. Those particles would only be able to communicate with known particles via weakly interacting messenger particles through one of four portals: the vector, axion, Higgs, and neutrino portals.

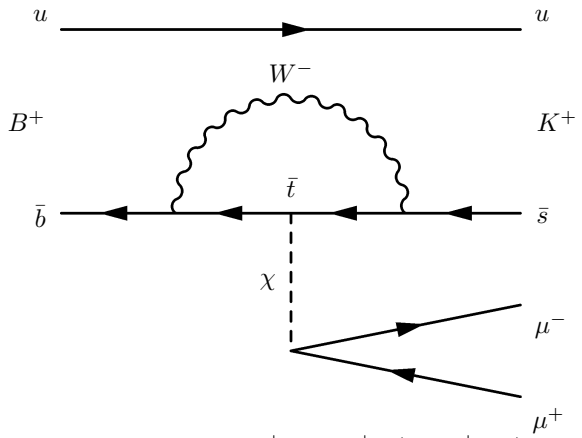


Figure 6.1: Feynman diagram of the decay $B^+ \to K^+\chi(\to \mu^+\mu^-)$, where χ interacts by mixing with the Higgs and then decays into a couple of muons.

The Higgs portal is constructed in such a way that the new scalar particle can mix with the recently discovered Higgs boson. Prime examples of such models are the so-called inflaton models where a new field was responsible for cooling down the Universe after the Big Bang [104, 105]. These models also help to solve the hierarchy problem and explain

the baryon asymmetry in the Universe [106, 107]. In these models, the inflaton mass and the lifetime are weakly constrained and the inflaton mass can be below the B meson mass, thus the decay of $B^+ \to K^+\chi$ is a prime candidate to look for such phenomena at LHCb. In such a scenario the inflaton couples to the t quark in the loop, as shown in Fig. 6.1, and can also be produced in the $b \to s$ transitions.

Besides the Dark Matter, the SM suffers from other unexplained phenomena. One of them is the so-called strong-CP problem. If one writes explicitly the QCD part of the SM Lagrangian, the following Lagrangian term, containing the free parameter θ, is allowed:

$$\mathcal{L} = \theta \frac{1}{16\pi^2} F^a_{\mu\nu} \epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}. \tag{6.1}$$

In this equation the θ parameter is responsible for breaking the CP symmetry. Given the experimental observation one sees that there is fine-tuning of this parameter to be extremely close to zero. This is known as the strong-CP problem. The solution to this problem has been proposed by R.Perci and H.Quinn [108] and is based on the introduction of a new axion pseudoscalar field with an additional $U(1)$ symmetry. Unlike the inflaton (which mixes with the Higgs), the new field can directly couple to the quarks.

6.1 $B^+ \to K^+\chi$ event selection

In this section we will discuss the selection requirements for the $B^+ \to K^+\chi$ search. Because we do not know a priori the mass and lifetime of the new particle χ, the selection has to be designed in a way that is efficient for the whole possible parameters space.

6.1.1 Trigger requirements

The trigger lines used are listed in Table 6.1. Candidates are required to be TOS in all stages of the trigger and they should be triggered on the B candidate. The dominant Hlt2 lines are `DiMuonDetached`, `TopoMu2Body` and `Hlt2SingleMuon` lines. Hlt2 `Topo3` lines contribute with a negligible signal efficiency but they increase the combinatorial background by 4 % and are therefore excluded from the requirement. It is observed that the trigger efficiency depends on m_χ, however muon trigger efficiency is very well simulated in the MC.

6.1.2 Stripping and preselection

The stripping line used to perform the analysis is the `B2KX2MuMuDarkBosonLine`. A summary of the requirements included in the stripping line can be found in Table 6.2.

After the trigger and stripping stages, a loose preselection is applied. These requirements are listed as well in Table 6.2 and are found to be between 97 and 98 % efficient on the signal MC after the stripping. In the following efficiency computation these requirements will be considered together with the stripping efficiency.

Table 6.1: Trigger requirements. For each level, candidates are required to be TOS in at least one line.

Level	Requirement
L0	`L0Muon` or `L0DiMuon`
Hlt1	`Hlt1TrackMuon` or `Hlt1DiMuonLowMass` or `Hlt1TrackAllL0`
Hlt2	`Hlt2SingleMuon` or `Hlt2DiMuonDetached` or `Hlt2TopoMu2BodyBBDT`

6.1.3 B-mass signal region

The analysis is performed with the candidates that have the full reconstructed B-meson invariant mass in a window of 100 MeV/c^2 around the B^+ mass: $[5230, 5330]$ MeV/c^2. This range corresponds to the selection of almost $\pm 3\sigma$ around the B-mass and is blinded for the full execution of the analysis. The B-mass resolution is shown in Fig. 6.2 for a MC signal of $m_\chi = 2500$ MeV/c^2 and $\tau_\chi = 1$ ps, the distribution is fitted with a double Crystal-Ball function and the total resolution is found to be (18.1 ± 0.3) MeV/c^2.

The upper sideband range is defined as $m_B \in [5450, 5800]$ MeV/c^2 and is used as a pure combinatorial background sample to train the multivariate selection described below.

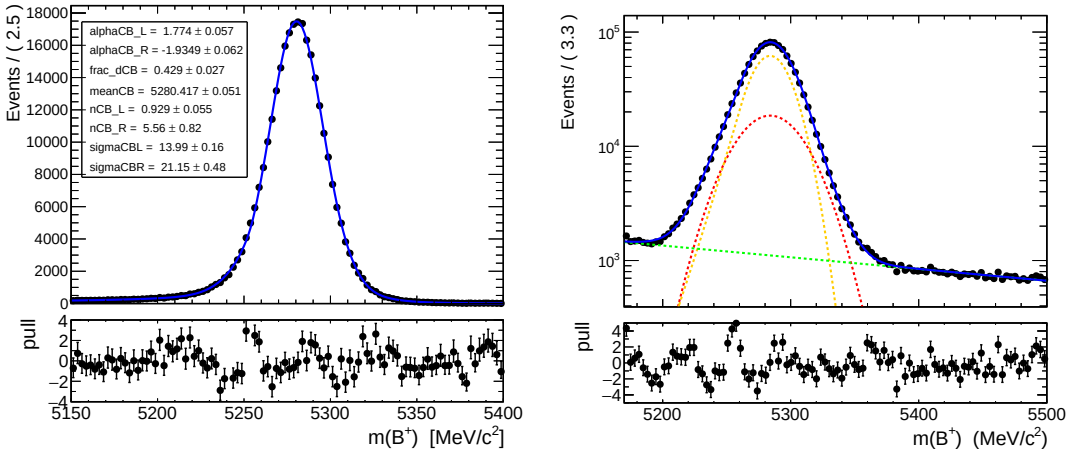


Figure 6.2: Left: Invariant mass distribution of the B^+ candidate for a MC signal of $m_\chi = 2500$ MeV/c^2 and $\tau_\chi = 1$ ps. The distribution is fitted with a double Crystal Ball. Right: Normalization channel $B^+ \to J/\psi K^+$; invariant mass fit for candidates with $m(\mu^+\mu^-) \in [3047, 3147]$ MeV/c^2. A double-sided double Crystal Ball is used to fit the peak while an exponential function is used for the background.

Table 6.2: List of requirements made on the $B^+ \to K^+\chi$ candidates at the stripping level.

Stripping		
Candidate	Variable	Requirement
B^+	M	[4800, 5800] MeV/c^2
	p_{T}	> 1000 MeV/c
	τ	> 0.2 ps
	$\cos\theta_{\mathrm{dir}}$	> 0
	vertex χ^2/ndf	< 25
	IP χ^2	< 50
tracks	$\mathcal{P}_{\mathrm{ghost}}$	< 0.3
	IP χ^2	> 9
	χ^2_{track}/ndf	< 3
K^+	p	> 2000 MeV/c
	p_{T}	> 250 MeV/c
	ProbNN$_K$	> 0.1
χ	$\cos\theta_{\mathrm{dir}}$	> 0
	χ^2_{FD} (w.r.t. B)	> 25
	χ^2_{vtx}/ndf	< 10
di-μ	M	< 5000 MeV/c^2
	p_{T}	> 250 MeV/c
	`doca`	< 0.2 mm
	χ^2_{doca}	< 25
μ	`isMuon`	`True`
	p_{T}	> 100 MeV/c
	DLL$_\mu$	> -5

Preselection		
Variable		Requirement
χ^2_{vtx} (B^+)	$<$	15
χ^2_{IP} (B^+)	$<$	25
$\cos\theta_{\mathrm{dir}}$ (B^+)	$>$	0.999

6.1.4 Multivariate selection

A multivariate selection is applied to further discriminate the signal from the background events. A k-folding technique is applied, with $k = 11$, in order to exploit at most the limited available background data set. The classifier is trained using a background sample consisting of around 200000 events taken from the right B-mass sideband introduced in the previous section. The contributions from the vetoed region of the analysis (see Sec. 6.3.1) are excluded also from the training background sample. Finally, each training background samples for the 11 folds consists of $\mathcal{O}(10^5)$ events. The classifier that gives the best performance among the traditional MVA algorithms is found to be the Gradient Boosted Decision Tree (BDTG). All the classifiers of the 11 folds are found to be statistically equivalent.

The variables used for the multivariate selection are summarised in Table 6.3. These variables are designed to pick up the tracks that are close to our signal tracks and that

can form a vertex with them. For the combinatorial background one expects more such tracks than for the signal tracks.

Table 6.3: Input variables for the multivariate selection.

Input variables	
decay-time (B^+)	$\chi^2_{\rm IP}$ (K^+)
$p_{\rm T}$ (B^+)	$\chi^2_{\rm FD}$ (χ)
$\cos\theta_{\rm dir}$ (B^+)	$\chi^2_{\rm IP}$ (μ)
$\chi^2_{\rm IP}$ (B^+)	$\mu^+\mu^-$ `doca`
$\chi^2_{\rm vtx}$ (B^+)	`track-iso65` (μ)
`MVA-iso` (χ)	

Three different options are considered as the signal sample for the training of the classifier: $\tau_\chi = 1, 10, 100$ ps. The goal is to have a performance as uniform as possible in the considered range of the (m_χ, τ_χ) plane. Fig. 6.3 shows the signal efficiency as a function of the signal lifetime for the three considered training configurations, all corresponding to $m_\chi = 2500$ MeV/c^2. All efficiencies are shown for two fixed background rejection rates, 0.9 and 0.99 respectively.

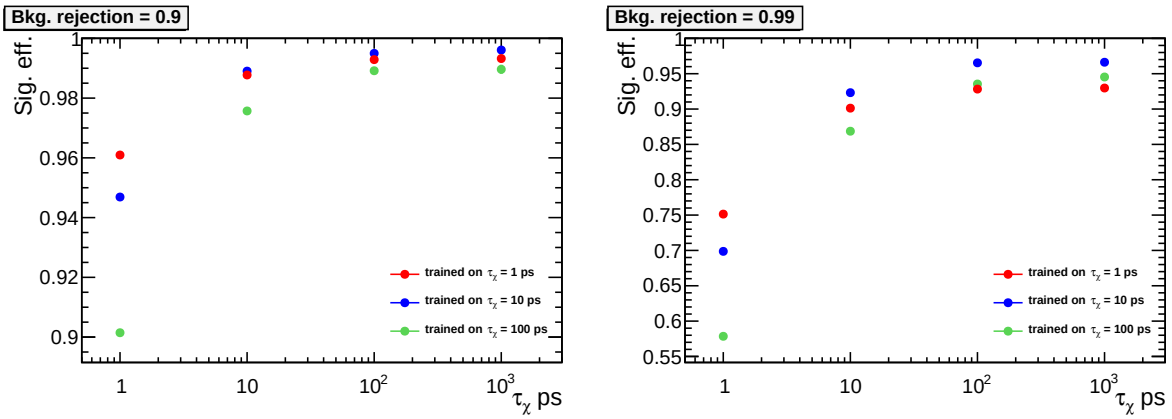


Figure 6.3: Signal efficiency for different lifetime of signal samples with $m_\chi = 2500$ MeV/c^2 (top), three training configuration are shown corresponding to different signal lifetimes. The efficiencies correspond to a background rejection of 0.9 (left top) and 0.99 (right top).

Configurations that were trained with a long lifetime signal showed a considerable loss of efficiency for short lifetime samples, while the configuration trained with the signal sample with $\tau_\chi = 1$ ps gave the best overall performance and has been chosen for the multivariate selection.

6.2 Strategy of the search

This analysis consists in a search for a new particle of unknown mass and lifetime through the decay $B^+ \to K^+\chi$. This search is performed by scanning the dimuon invariant mass spectrum looking for an excess of events. Since the lifetime of the new particle can be long compared to the detector resolution, the dimuon vertex is allowed; however, it is not required, to be displaced from the B-meson decay vertex. In order to increase the sensitivity to different lifetimes, three bins of the time of decay of the χ candidate are considered separately. Afterwards, the information from the three different bins is combined in a statistical way. The result of this analysis is presented as a 95% confidence level (CL) upper limit, as a function of the mass and lifetime of the new particle χ. The considered mass and lifetime range goes between 250 and 4700 MeV/c^2 and between 0.1 and 1000 ps, respectively.

The division in the three bins for the time of decay of the candidate reflects the following criteria:

- The first bin corresponds to a small time of decay and contains events that can be identified as prompt decays. It is meant to contain all the irreducible SM $B^+ \to K^+\mu^+\mu^-$ electroweak penguin decays that are the main background of this analysis (see Sec. 6.3.6). It is characterized by a high signal efficiency but also by a huge background contamination.

- The second bin contains displaced dimuon decays and is supposed to be most sensitive to long lifetimes; it has lower reconstruction efficiency compared to the first bin, but it is affected by a low background contamination; only combinatorial background can populate this region (see Sec. 6.3.7).

- The third bin contains very displaced dimuon candidates and is chosen so that this region corresponds to a "zero background" search. This third bin is mainly added in order to increase the sensitivity to a very long lifetime signal in the event that the observation is made.

The bins boundaries are optimized using pseudoxperiments to maximize the CLs value. The obtained values for each bin are:

- 1^{st} bin: $t < 1$ ps, BDTG> 0.6,
- 2^{nd} bin: $t \in [1 < t < 10]$ ps, BDTG $> f(m)$, where $f(m)$ is plotted in Fig. 6.4,
- 3^{rd} bin: $t > 10$ ps, BDTG > 0.8.

6.2.1 Searching in the mass dimension

The analysis aims at searching for a new particle χ decaying into a pair of opposite charged muons, that is produced in the decay of: $B^\pm \to K^\pm\chi$. The new particle is of unknown

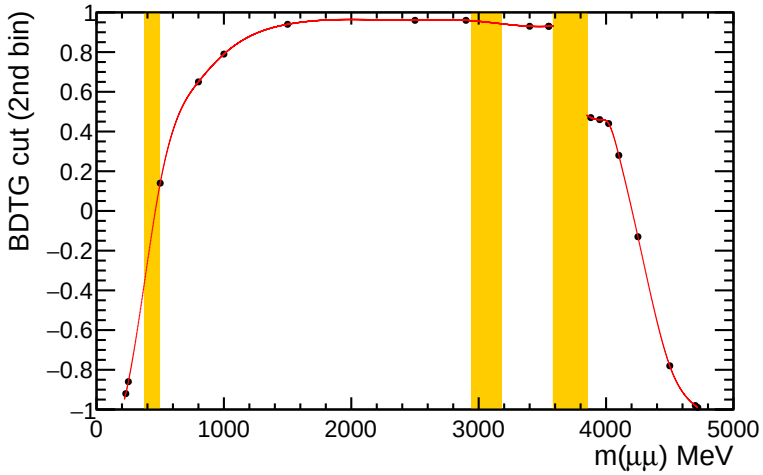


Figure 6.4: Value of the tightest allowed selection in the second bin as a function of mass. The points corresponds to masses for which the requirement described in the text was tested with toys. The $BDTG_{2nd\,bin}^{max\text{-}cut}$ is then interpolated with a cubic spline. Yellow bands correspond to vetoed regions in the search.

mass and lifetime, so the search is performed in different regions of those parameters. The search for the new particle is the dimuon invariant mass in step of $\frac{1}{2}\sigma(m)$, where $\sigma(m)$ is the local dimuon mass resolution (see Sec. 6.4.2). For each tested mass m_{test}, a signal region is defined as

$$|m - m_{\text{test}}| < 2\sigma(m) \tag{6.2}$$

and a background region is the sidebands of the signal region corresponding to

$$3\sigma(m) < |m - m_{\text{test}}| < (2x + 3)\sigma(m). \tag{6.3}$$

Therefore, the parameter x is defined as the ratio between the size of the background and the signal regions.

For each m_{test} a test statistic is performed: the background plus signal and the background only hypothesis are compared using the CLs method, where the information from the three bins is combined into a single likelihood. For each bin the expected number of signal events is taken from the MC and normalized to the yield of normalization channel: $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$, while the expected number of background events is obtained directly from the data with a linear fit to the background region (dimuon sidebands).

The choice of the parameter x is driven by the principle of keeping x as large as possible, but in such a way that the approximation of local linearity is still valid.

The regions around the ϕ, J/ψ, $\psi(2S)$ and $\psi(3770)$ dimuon resonances are vetoed from the search (see Sec. 6.3.1). Moreover, if the vetoed region is inside the sideband used to extrapolate the background yield in the signal region, the opposite sideband to the one containing the vetod region is enlarged, so that the total sideband size is kept to be x. The linear approximation of background yield extraction in the above scenario is tested with

pseudo-experiment studies. It has been found that $x = 10$ provides the best compromise between the statistical uncertainty and bias created by the linear approximation. The largest biases that have been found in the pseudo-experiments were 8 % of the statistical error.

6.3 Background

This section describes the peaking background considerations. Fake rates of particle misidentification are measured using data driven methods from the PIDCalib package [109]. The expected background events are reported after the full selection was applied and computed with the formula

$$N_{channel} = N^{obs}_{B^+\to J/\psi K^+} \times \frac{\varepsilon_{channel}}{\varepsilon_{B^+\to J/\psi K^+}} \times \frac{\mathrm{BR}(channel)}{\mathrm{BR}(B^+ \to J/\psi K^+)}, \tag{6.4}$$

where the decay $B \to J/\psi(\to \mu^+\mu^-)K^+$ is a normalization channel.

6.3.1 SM charmonium resonances

The decay $B^+ \to \phi K^+$ and the charmonium decays $B^+ \to J/\psi K^+$, $B^+ \to \psi(2S)K^+$, $B^+ \to \psi(3770)K^+$ and $B^+ \to \psi(4160)K^+$, where ϕ, J/ψ, $\psi(2S)$, $\psi(3770)$ and $\psi(4160)$ decay into two opposite sign muons, have the same final state of the signal. These candidates are removed by applying a veto on the $\mu^+\mu^-$ mass in the ranges $985 < m_{\mu^+\mu^-} < 1055$, $2946 < m_{\mu^+\mu^-} < 3176$, $3586 < m_{\mu^+\mu^-} < 3850$ and $4103 < m_{\mu^+\mu^-} < 4270$ MeV/c^2. The upper end of the $\psi(2S)$ veto falls around the $\psi(3770)$ mass, so they are included into a single vetoed region. Since the expected contribution in the displaced region of the analysis of the $B^+ \to \phi K^+$ and $B^+ \to \psi(4160)K^+$ decays is negligible and no peaks corresponding to the ϕ and $\psi(4160)$ resonances are observed in the combinatorial background in the right B-mass sideband after the selection, the corresponding vetoes are applied only in the first proper life time bin of the search. Despite the fact that the $\psi(4160)$ resonance is much larger than the signal resolution (see Figure 6.9), a significant bias in the background prediction is observed and therefore a veto is applied.

6.3.2 $B^+ \to J/\psi K^+$ with a $K^+ \leftrightarrow \mu^+$ swap

The decay $B^+ \to J/\psi K^+$ can fall out of the J/ψ veto if the kaon swaps mass hypothesis with the same-sign muon. This background is rejected by computing the $\mu^- K^+$ mass under the $\mu^-\mu^+$ hypothesis, $m_{(K\mu\leftrightarrow\mu\mu)}$, and by requiring the kaon to fail the isMuon flag for the candidates with a mass $m_{(K\mu\leftrightarrow\mu\mu)}$ within the range $3000 < m_{(K\mu\leftrightarrow\mu\mu)} < 3200$ MeV/c^2 (cf. Fig. 6.5).

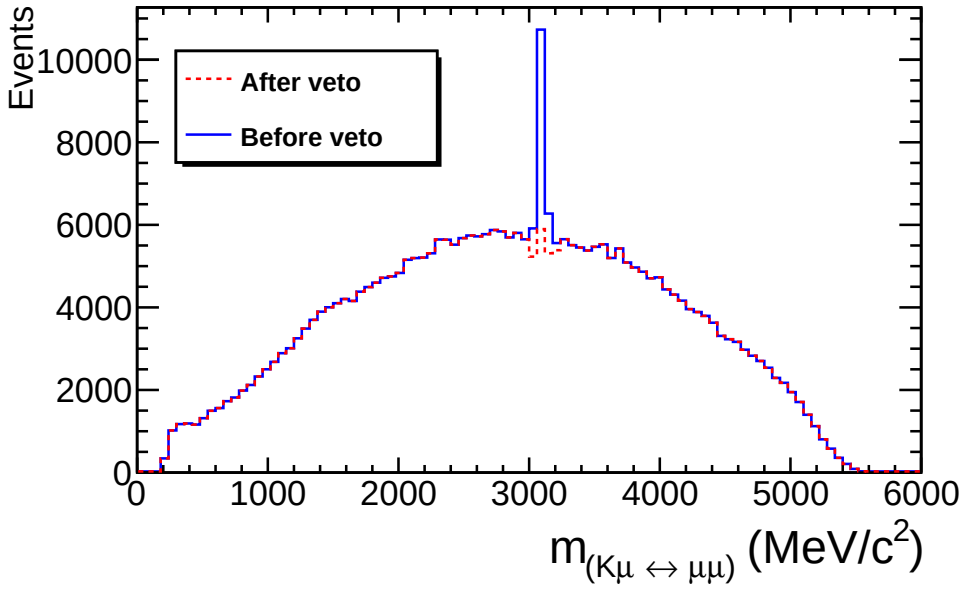


Figure 6.5: The peaking background from $B^+ \to J/\psi K^+$ with a $K^+ \leftrightarrow \mu^+$ swap is shown before and after the veto described in Sec. 6.3.2 is applied. The distribution is shown for data in the right B-mass sideband.

6.3.3 Hadronic B^+ decay

Four purely hadronic decays have been investigated as potential sources of peaking background: $B^+ \to K^+K^+K^-$, $B^+ \to K^+K^-\pi^+$, $B^+ \to K^+\pi^+\pi^-$ and $B^+ \to \pi^+\pi^+\pi^-$. These decays can form a source of peaking backgroud when the pions or kaons are mis-identified as muons. The mis-identification rate and the final number of events expected to pass the full selection are listed in Table 6.4. The mis-identification rate is computed with the PIDCalib package, the fraction of events within the B-mass signal region comes from the MC, and the final expected number of the events follows the formula in Equation 6.4.

The remaining peaking background, such as the 26 ± 4 $B^+ \to K^+\pi^+\pi^-$ expected events, will not cause problems, because this decay is dominated by wide resonances that are not dangerous to this study as the linear background extrapolation will be valid. The decay mode like $B^+ \to K^+\phi$, $B^+ \to K^+\omega$, $B^+ \to K^+\rho^0$ or $B^+ \to \pi^+\rho^0$ are even more suppressed and are found to be negligible.

6.3.4 $B^+ \to D$ decays

The decays $B^+ \to D^0\pi$ and $B^+ \to D^0K$, with $D^0 \to K^+\pi^-$, $D^0 \to K^+K^-$ or $D^0 \to \pi^+\pi^-$ can be particularly dangerous,which is due to the presence of the displaced secondary vertex of the D^0 meson decay. $B^+ \to D^0(\to KK)X$ and $B^+ \to D^0(\to \pi\pi)X$ are Cabibbo suppressed and are found to be negligible, while a small contribution of $B^+ \to D^0(\to K\pi)\pi$ and $B^+ \to D^0(\to K\pi)K$ is expected in data (see Table 6.5). These backgrounds are

Table 6.4: Fully hadronic B^+ background decays.

Decay	Branching ratio	Misidentification rate	Fraction within signal region	Expected number of events
$B^+ \to K^+K^+K^-$	$(3.40 \pm 0.14) \times 10^{-5}$	$(2.00 \pm 0.05) \times 10^{-4}$	0.0076 ± 0.0013	0.58 ± 0.14
$B^+ \to K^+K^-\pi^+$	$(5.0 \pm 0.7) \times 10^{-6}$	$(1.40 \pm 0.05) \times 10^{-4}$	0.17 ± 0.01	1.6 ± 0.3
$B^+ \to K^+\pi^+\pi^-$	$(5.1 \pm 0.3) \times 10^{-5}$	$(9.40 \pm 0.08) \times 10^{-5}$	0.758 ± 0.011	26 ± 4
$B^+ \to \pi^+\pi^+\pi^-$	$(1.52 \pm 0.14) \times 10^{-5}$	$(1.5 \pm 0.1) \times 10^{-5}$	0.57 ± 0.02	2.3 ± 0.4

rejected by the introduction of a more stringent muon PID requirement for candidates with invariant mass, after the correct mass hypothesis assignment $m_{(\mu\mu\leftrightarrow K\pi)}$, in a window around the D^0 mass. The requirement at ProbNNmu(μ) > 0.4 is sufficient to remove these sources of background and is applied to events with $m_{(\mu\mu\leftrightarrow K\pi)} \in [1840, 1890]$ MeV/c^2.

Table 6.5: Potential B decay backgrounds. The last column reports the expected yields before the introduction of the dedicated requirement described in the text.

Decay	Branching ratio	Misidentification rate	Fraction within signal region	Expected number of events
$B^+ \to D^0(K^+\pi^-)K^+$	$(1.43 \pm 0.07) \times 10^{-5}$	$(1.80 \pm 0.04) \times 10^{-4}$	0.148 ± 0.012	3.0 ± 0.5
$B^+ \to D^0(K^+K^-)K^+$	$(1.46 \pm 0.07) \times 10^{-6}$	$(2.60 \pm 0.05) \times 10^{-4}$	0.013 ± 0.004	0.03 ± 0.01
$B^+ \to D^0(\pi^+\pi^-)K^+$	$(5.2 \pm 0.3) \times 10^{-7}$	$(1.20 \pm 0.04) \times 10^{-4}$	0.76 ± 0.02	0.51 ± 0.08
$B^+ \to D^0(K^+\pi^-)\pi^+$	$(1.87 \pm 0.06) \times 10^{-4}$	$(3.50 \pm 0.06) \times 10^{-5}$	0.26 ± 0.02	15.3 ± 3.0
$B^+ \to D^0(K^+K^-)\pi^+$	$(1.90 \pm 0.07) \times 10^{-5}$	$(4.00 \pm 0.06) \times 10^{-5}$	0.062 ± 0.008	0.43 ± 0.08
$B^+ \to D^0(\pi^+\pi^-)\pi^+$	$(6.7 \pm 0.2) \times 10^{-6}$	$(2.40 \pm 0.05) \times 10^{-5}$	0.62 ± 0.03	1.00 ± 0.18

6.3.5 K^0_S, Λ^0 resonances

Peaks corresponding to the $K^0_S \to \pi^+\pi^-$ and $\Lambda^0 \to p\pi$ are observed in the data sidebands and removed by a veto around the K^0_S and Λ^0 mass after the correct mass hypothesis assignment. Because of a small difference in the mass between a muon and a pion, the veto around the K^0_S mass is reflected as a complete loss of efficiency for the signal with m_χ close to the K^0_S mass. For this reason the window $[380, 495]$ MeV/c^2 is excluded from the search.

A big difference between the proton and the muon mass never causes any signal in the studied mass range to fall into the Λ^0 vetoed region.

6.3.6 $b \to s\ell^-\ell^+$ background

The $B^+ \to K^+\mu^+\mu^-$ decay has the same final state of the signal and represents an irreducible source of background. The dimuon pairs are expected to come from a prompt decay

of the B-meson and, as already mentioned, these events represent the main background in the first bin of the search.

Besides the dominant $B^+ \to K^+\mu^+\mu^-$, other decays that contain two opposite sign muons are considered:

- $B^+ \to \pi^+\mu^+\mu^-$: in this case the pion should be mis-identified as a kaon, but the two real muons enhance the mis-identification rate. However, the low branching fraction and stripping efficiency return it negligible.
- $B^0 \to K^*\mu^+\mu^-$ with one missing track from $K^* \to K\pi$. Here the missing track reduces the probability for the total invariant mass to be reconstructed within the B-mass signal region.

The decay modes with the relative rates are listed in Table 6.6. Other possible contributions from the B_s and Λ_b decays have been considered but they are always found to be negligible.

Table 6.6: B decays with two opposite sign muons in the final state, thus comprising a possible source of background. For the $B^0 \to K^*(\to K\pi)\mu^+\mu^-$ decay with one missing track the notation $K^*(\to \pi)$ corresponds to a missing kaon and oppositely $K^*(\to K)$ to a missing pion.

Decay	Branching ratio	Misidentification rate	Fraction within signal region	Expected number of events
$B^+ \to \pi^+\mu^+\mu^-$	$(2.4 \pm 0.6) \times 10^{-8}$	0.128 ± 0.0016	0.448 ± 0.003	1.1 ± 0.3
$B^0 \to K^*(\to K)\mu^+\mu^-$	$(1.05 \pm 0.1) \times 10^{-6}$	0.8096 ± 0.0013	0.00016 ± 0.00008	1.1 ± 0.5
$B^0 \to K^*(\to \pi)\mu^+\mu^-$	$(1.05 \pm 0.1) \times 10^{-6}$	0.0767 ± 0.0004	0.038 ± 0.008	6.5 ± 1.6

6.3.7 Combinatorial background

The combinatorial background is the only source of background that can populate the *displaced* region of the analysis. Fig. 6.6 shows the distribution of the decay time of the dimuon system for the combinatorial background where the events are taken from the right B-mass sideband. The tail of the distribution extends up to tens of ps. We reject the events with life time $t < -1$ ps. The requirement value was chosen in such a way that it takes into account the time resolution of the detector. This requirement rejects more than 30 % of the combinatorial background in the first bin of the search.

The combinatorial background yield is extrapolated from data with a fit to the right B-mass sideband, and the expected yields after the selection are listed in Table 6.7 for each bin of the analysis.

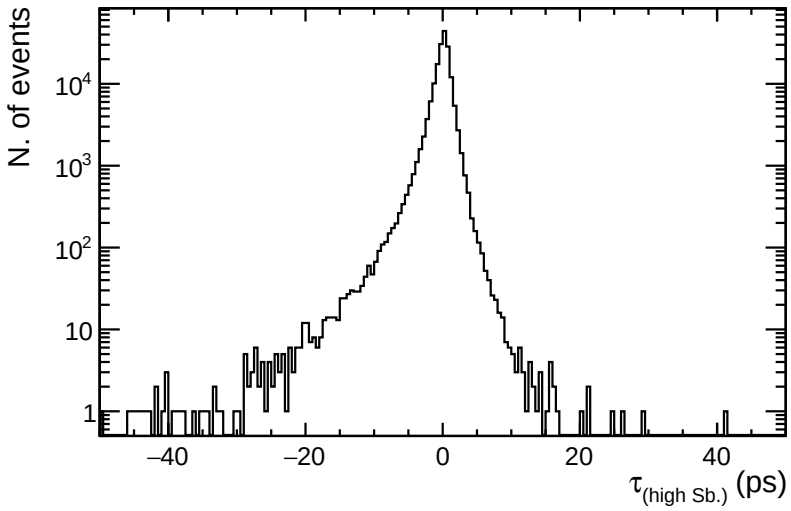


Figure 6.6: Time of decay distribution for combinatorial background. The events are taken from the right B-mass sideband.

Table 6.7: Expected number of combinatorial background events after the selection.

	Combinatorial background
1^{st} bin	981 ± 50
2^{nd} bin	171 ± 23
3^{rd} bin	2.6 ± 2.4

6.4 Efficiency determination

This section presents the evaluation of the total signal efficiency, used to normalize the signal to the observed normalization yield with the formula

$$N_{sig} = N^{obs}_{B^+ \to J/\psi K^+} \times \frac{\varepsilon_{sig}}{\varepsilon_{B^+ \to J/\psi K^+}} \times \frac{\mathrm{BR}(sig)}{\mathrm{BR}(B^+ \to J/\psi K^+)}. \tag{6.5}$$

6.4.1 MC corrections

Two kinds of corrections are applied to the MC: the correction for the number of tracks in the event and for particle identification variables, which both are not reproduced well in the simulation. These MC corrections are applied to all the MC samples after the stripping, pre-selection and trigger requirements. Fig. 6.7 shows the distribution of the number of tracks before and after the reweighting of the MC. The normalization channel $B^+ \to J/\psi K^+$ is used as a reference for data distribution. Fig. 6.8 proves that

the distribution of the number of tracks is about the same for all the MC signals and the normalization channel.

The PID identification corrections follow the same strategy as in the previously described analysis. For details see Sec. 4.5.2.

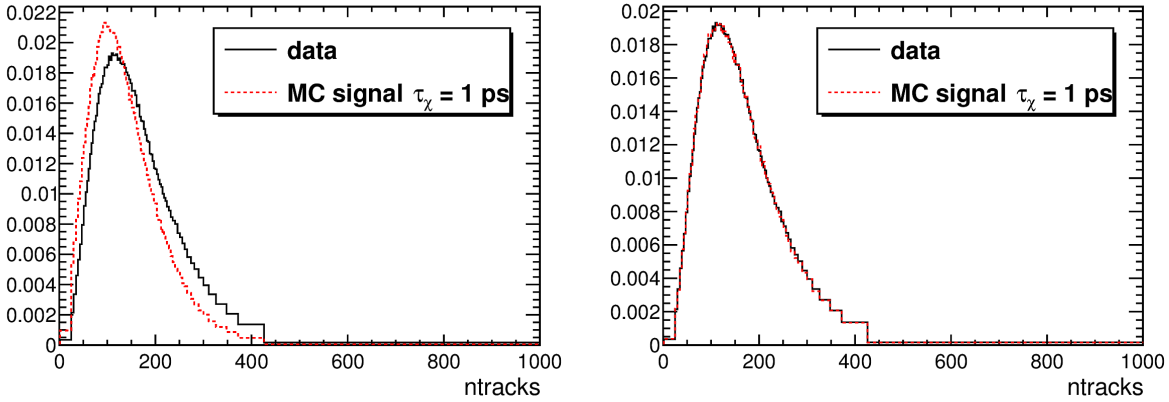


Figure 6.7: Distributions of the number of tracks for the normalization sample in data and for one MC signal sample of reference. The distributions are shown before (left) and after (right) the correction of the MC.

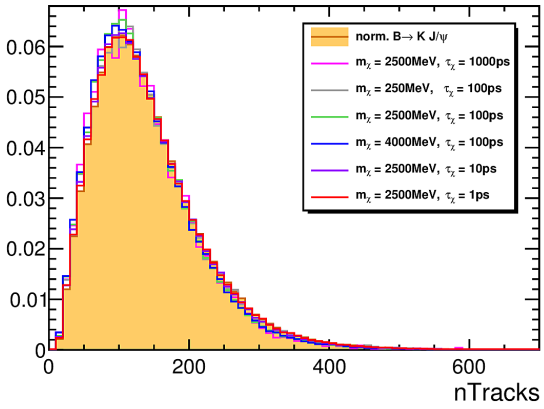


Figure 6.8: Distributions of the number of tracks for the MC normalization sample and for some MC signal samples.

6.4.2 χ mass resolution

All the reconstructed invariant masses in this analysis are computed using `DecayTreeFitter` [110]. In order to improve the dimuon mass resolution, the candidates in the signal region, $5230 < m_{B^+} < 5330$ MeV/c^2, have the mass of the B^+ meson

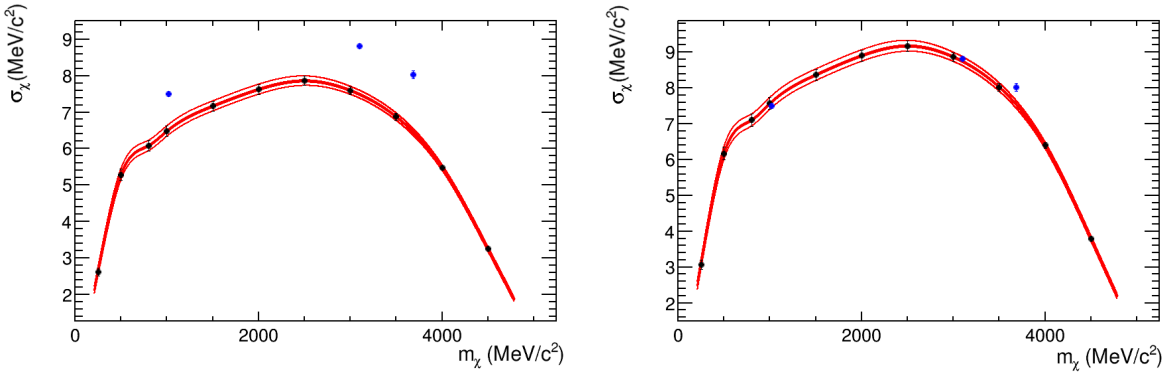


Figure 6.9: Mass resolution as a function of m_χ before (left) and after (right) scaling to the resolution on data. MC signal (black) points are interpolated with a cubic spline. The resolution from data for the J/ψ and $\psi(2s)$ resonances is also shown as blue points. The mass of the B^+ meson is constrained to its nominal value [111].

constrained to its world average value in the vertex fit [111]. The dimuon mass resolutions vary as a function of m_χ and the resolution for each simulated sample is found by fitting the mass spectrum to a double Gaussian function. All the dimuon mass resolutions are computed after the stripping, pre-selection, trigger, PID, and B-mass signal region requirement. The normalization channel is used to cross-check the resolution obtained from the MC. It has been found that the MC slightly underestimates the resolution, therefore the resolution from all the MC m_χ is scaled by the factor:

$$\frac{\sigma^{DATA}_{J/\psi\to\mu^+\mu^-}}{\sigma^{MC}_{J/\psi\to\mu^+\mu^-}} = 1.167. \tag{6.6}$$

This kind of procedure has been tested and applied to various analyses in LHCb. The validity of the obtained scale factor is tested with the resolution of the ϕ and $\psi(2S)$ resonances measured on data, as shown in Fig. 6.9.

After the rescaling, the ϕ resolution is compatible with the interpolation of the m_χ resolution at the corresponding mass, while the $\psi(2S)$ resolution is found to be within two sigma deviation. This deviation corresponds to 6% of the scale factor. It is assigned as the uncertainty on the scale factor and added as a source of systematics.

6.4.3 Signal efficiency

The overall signal efficiency, called ε_{sig}, is factorized in several components: ε_{acc}, acceptance efficiency for tracks to be within the LHCb acceptance, $\varepsilon_{reco\&strip}$, reconstruction efficiency for the signal to pass the stripping and pre-selection requirements, $\varepsilon_{trigger}$, trigger efficiency, ε_{PID}, particle identification efficiency, $\varepsilon^{B\text{-}meson}_{sig.Reg.}$, efficiency to lie into the signal region defined around the B-meson mass, $\varepsilon^{dimuon}_{sig.Reg.}$, efficiency to lie into the signal region around each tested dimuon mass, ε_{BDTG}, efficiency of the multivariate selection,

$$\varepsilon_{sig} = \varepsilon_{acc} \times \varepsilon_{reco\&strip} \times \varepsilon_{trigger} \times \varepsilon_{PID} \times \varepsilon^{B\text{-}meson}_{sig.Reg.} \times \varepsilon^{dimuon}_{sig.Reg.} \times \varepsilon_{BDTG}. \tag{6.7}$$

The total efficiency is always defined by combining the three bins of the search, while the partial components are calculated with respect to the number of events passing all the previous requirements. The efficiency ε_{PID} includes the PID requirements contained in the stripping: `isMuon` and $\mathrm{DLL}_\mu > -5$ for the muons and $\mathrm{ProbNN}_K > 0.1$ for the kaon. The muon PID efficiency is obtained directly from the MC, while the PID efficiency of the $\mathrm{ProbNN}_K > 0.1$ is computed with PidCalib package. The uncertainty on ε_{PID} is dominated by the muon component limited by the MC size. All the other efficiency components and corresponding uncertainties are evaluated with the MC. Fig. 6.10 shows the total signal efficiency for all the MC generated masses.

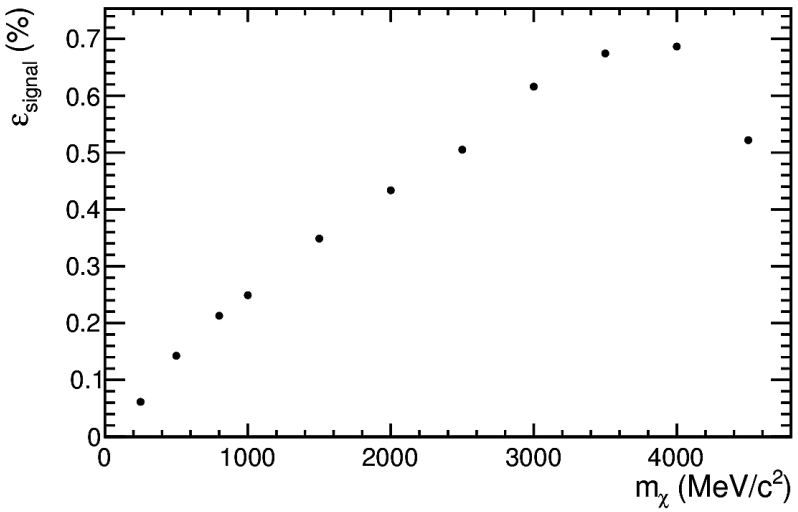


Figure 6.10: Total signal efficiency as a function of mass for MC samples generated with $\tau_\chi = 100$ ps.

6.5 Systematic uncertainties

Estimation of systematic uncertainties is required for setting upper limits which are included as nuisance parameters in the search. The following sources of systematics are considered:

- Uncertainties on the signal efficiency computed on the MC.
- Mismodelling of the signal mass resolution.
- Uncertainty on the normalization branching ratio.
- Mismodelling of the background mass shape.

The relative contributions are listed in Table 6.8 at the end of this section. The impact of these uncertainties on the excluded limit is found to be minimal, the excluded limit is increased by only 2 % on average in the considered (m_χ, τ_χ) plane.

Table 6.8: List of systematic uncertainties.

Source	Uncertainty
Signal resolution	$(1.5 \div 2)\%$
MC size	$(2 \div 6)\%$
MC lifetime reweighting	$(0 \div {}^{+0}_{-20})\%$
Background mass shape mismodelling	$0.08 \times$ stat. err.
Normalization branching ratio	3 %

6.5.1 Signal efficiency modelling

The signal efficiency is computed based on the available MC samples as described in Sec. 6.4 and a systematic uncertainty must be introduced because of the finite size of the MC sample. The limited size of the MC especially affects the two extreme lifetime hypotheses considered in the analysis, $\tau_\chi = 0.1$ ps and 1000 ps, when large weights are assigned to a very small number of events. For these two extreme configurations the systematic uncertainty associated to the signal efficiency is found to be around 6 %. In all the intermediate hypotheses this systematic uncertainty is 2 %. Moreover, the reweighting procedure overestimates by around 20 % the signal efficiency for very long lifetimes, $\mathcal{O}(10^3)$ ps. To take this effect into account, an asymmetric systematic uncertainty is added to the signal samples with $\tau_\chi > 100$ ps.

6.5.2 Signal resolution modelling

The MC signal resolution is scaled to data as explained in Sec. 6.4.2. The systematic associated to the uncertainty of the mass resolution is given by two contributions: on the one hand the uncertainty on the resolution as it comes from the fit to the MC samples, which is obtained by varying the resolution within its uncertainty, and by considering the difference on the resulting efficiency. On the other hand, the relative error of 6 % assigned to the scale factor (see Sec. 6.4.2) is found to be the dominant one and results in the uncertainty on the signal efficiency between 1.5 and 2 %, depending on the signal mass .

6.6 Results

The upper limits are set using the CLs method as described is Sec. 6.2. After unblinding the analysis, the B-mass signal region is explored. Fig. 6.11 shows the dimuon distribution for events in the B-signal region in the first and second bin of the analysis. Since no significant excess from the SM background is found, a 95% CL upper limit is set on the $\mathrm{BR}(B^+ \to K^+\chi(\mu^+\mu^-))$.

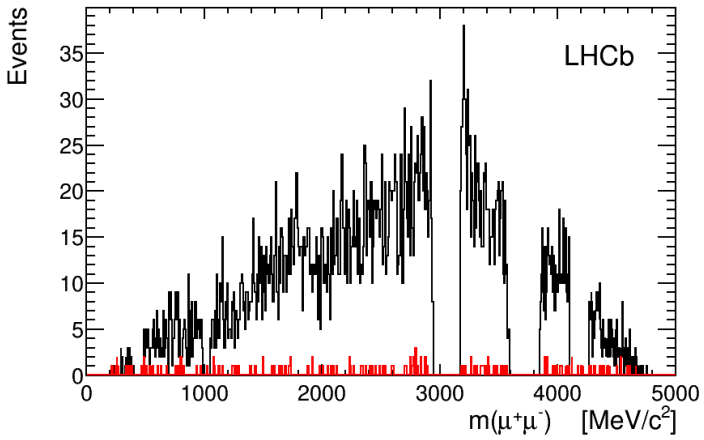


Figure 6.11: Dimuon distribution for events in the B-signal region in the first (black) and second (red) bin of the search. The binning scheme corresponds to one sigma of the mass-dependent mass-resolution model. No events are observed in the third bin of the analysis.

The upper limits are then set on the branching fraction BR($B^+ \to K^+\chi(\mu^+\mu^-)$) at 95% CL. The result is shown in Figure 6.12 as a function of m_χ and τ_χ.

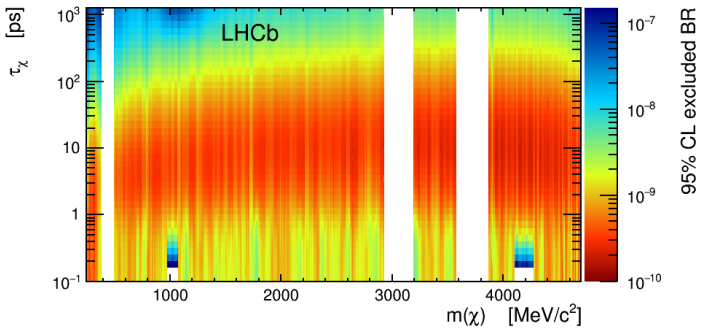


Figure 6.12: 95 % CL excluded limit in the studied (m_χ, τ_χ) plane.

6.7 Interpretation of the results in the inflaton model

In this section we will give a detailed description of the inflaton model. In the literature there exists a widely known inclusive formula for the branching fraction [104]:

$$\mathrm{Br}(B \to X_s\chi) \sim 10^{-6}\left(1-\frac{m_\chi^2}{m_b^2}\right)\left(\frac{\beta}{\beta_0}\right)\left(\frac{300\,\mathrm{MeV/c^2}}{m_\chi}\right). \tag{6.8}$$

We propose an interpretation of our results based on [112]. There explicit formulas for

the K^* and K case are given:

$$\mathrm{BR}(B^+ \to K^+\chi) = 4 \times 10^{-7} \left(\frac{\theta}{10^{-3}}\right)^2 F_K(m_\chi)\sqrt{\lambda}, \tag{6.9}$$

$$\mathrm{BR}(B^0 \to K^{*0}\chi) = 5 \times 10^{-7} \left(\frac{\theta}{10^{-3}}\right)^2 F_{K^*}(m_\chi)\sqrt{\lambda}^3, \tag{6.10}$$

where F_i are the form factors, and λ is the phase space suppression. The form factors used in the above equations have been updated to the latest available ones. One finds that the branching fraction to K and K^* is of the order of 10 % of the inclusive productions (X_s). Taking this into account in this calculation, we can draw the exclusion of the parameters space in the inflaton model (see Fig. 6.13). As can be seen, in some regions of the parameters space, our results improves on the previous $B^0 \to K^*\chi$ [113] by almost an order of magnitude.

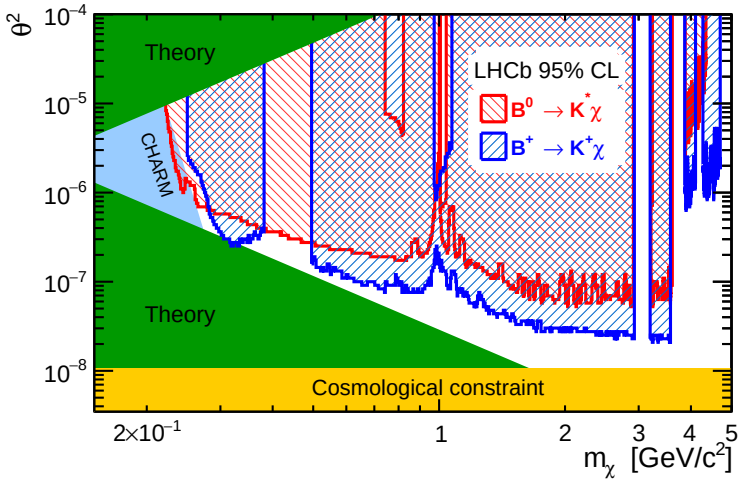


Figure 6.13: Exclusion plane (m_ξ, θ^2). The blue curve corresponds to the exclusion by $B^+ \to K^+\chi$, the red one to $B^0 \to K^{*0}\chi$. The theoretical constraints of the model and the region excluded by the CHARM [114] experiment are shown as well.

Chapter 7

Interpretation of $b \to s\ell^-\ell^+$ anomalies

In this chapter we will present the interpretation of the observed anomalies in the $b \to s\ell^-\ell^+$ transitions. We will start by presenting the results of global fits to the WC and long distance effects. The chapter ends with list of possible NP scenarios that can explain the cause of the observed deviations from the SM prediction.

7.1 Wilson Coefficient fit

Presently, there are two groups that perform the state of the art global WC fits, including a complete set of $b \to s\ell^-\ell^+$ decays. The first group involves B. Capdevila, A. Crivellin, S. Descotes-Genon, J. Matias and J. Virto [115, 116] and will be denoted as CCDMV from the first letters of the authors' names. The second group includes W. Altmannshofer, C. Niehoff, P. Stangl, D. Straub and, in analogy, will be called ANSS [117]. We will shortly present the results of both groups.

7.1.1 CCDMV fit

In the CCDMV fit [115] the following experimental results are taken into account:

- Branching ratio measurement of $B^0 \to K^{*0}\mu^+\mu^-$ done by LHCb with full Run1 data set [118].
- Angular observables of $B^0 \to K^{*0}\mu^+\mu^-$ in the P_i basis Sec. 2.8.1 measured by the LHCb [69], Belle [95], ATLAS [96] and CMS [97] experiments.
- Branching ratio of the $B^+ \to K^+\mu^+\mu^-$ decay from LHCb [119].
- Branching ratio and angular observables of $B_s^0 \to \phi\mu^+\mu^-$ [120].
- Inclusive measurements of $B \to X_s\ell\ell$ and $B \to X_s\gamma$ [121, 122].
- Branching ratio measurement of $B^0 \to K^{*0}\gamma$ [123].

- Branching ratio of $B_s^0 \to \mu^+\mu^-$ [11].
- The LFU measurements of R_K and R_{K^*} [124, 125].

The global fit is performed minimalizing χ^2:

$$\chi^2(\mathcal{C}_k) = \sum_{i,j=1}^{N_{\rm obs}} \left[O_i^{\rm exp} - O_i^{\rm th}(C_k)\right] (C_{\rm exp} + C_{\rm th})_{ij}^{-1} \left[O_j^{\rm exp} - O_j^{\rm th}(C_k)\right] , \tag{7.1}$$

where $N_{\rm obs}$ is the number of observables, $O_i^{\rm exp/th}$ are the values of the i-th observable measured by the experiment and calculated by the theory respectively, and $C_{\rm exp/th}$ are experimental ant theoretical covariance matrices.

The CCDMV authors performed two types of WC fits. In the first fits only one WC is allowed to float, while the others are fixed to the SM values. In the second fits two WC are allowed to float. The results of the fits are shown in Table 7.1. The authors also considered a subset of observables, which are R_K and R_{K^*}.

	All					LFUV				
1D Hyp.	Best fit	1 σ	2 σ	Pull_{SM}	p-value	Best fit	1 σ	2 σ	Pull_{SM}	p-value
$C9\mu^{\rm NP}$	-1.10	$[-1.27, -0.92]$	$[-1.43, -0.74]$	5.7	72	-1.76	$[-2.36, -1.23]$	$[-3.04, -0.76]$	3.9	69
$C9\mu^{\rm NP} = -C10\mu^{\rm NP}$	-0.61	$[-0.73, -0.48]$	$[-0.87, -0.36]$	5.2	61	-0.66	$[-0.84, -0.48]$	$[-1.04, -0.32]$	4.1	78
$C9\mu^{\rm NP} = -C9\mu'$	-1.01	$[-1.18, -0.84]$	$[-1.33, -0.65]$	5.4	66	-1.64	$[-2.12, -1.05]$	$[-2.52, -0.49]$	3.2	31
$C9\mu^{\rm NP} = -3C9e^{\rm NP}$	-1.06	[-1.23,-0.89]	[-1.39,-0.71]	5.8	74	-1.35	$[-1.82, -0.95]$	$[-2.38, -0.59]$	4.0	71

	All			LFUV		
2D Hyp.	Best fit	Pull_{SM}	p-value	Best fit	Pull_{SM}	p-value
$(C_{9\mu}^{\rm NP}, C_{10\mu}^{\rm NP})$	(-1.17,0.15)	5.5	74	(-1.13,0.40)	3.7	75
$(C_{9\mu}^{\rm NP}, C_7')$	(-1.05,0.02)	5.5	73	(-1.75,-0.04)	3.6	66
$(C_{9\mu}^{\rm NP}, C_{9'\mu})$	(-1.09,0.45)	5.6	75	(-2.11,0.83)	3.7	73
$(C_{9\mu}^{\rm NP}, C_{10'\mu})$	(-1.10,-0.19)	5.6	76	(-2.43,-0.54)	3.9	85
$(C_{9\mu}^{\rm NP}, C_{9e}^{\rm NP})$	(-0.97,0.50)	5.4	72	(-1.09,0.66)	3.5	65

Table 7.1: Results of the CCDMV 1-dim(top) and 2-dim fits. [115]. The "All" column refers to all the above-mentioned observables used in the fit, while the "LFUV" refers to the fit results that have been done only using the measurements of R_K and R_{K^*}.

As can be seen, when including all of the observables the authors have observed above 5 σ deviation from the SM prediction of the C_9 WC. It is worth mentioning at this point that some of the observables in the fit suffer from charm-loops effects (see Sec. 2.8.3 for details), so it is hard to quantify how reliable the obtained significances are. For this reason the authors have decided to add the fit to the subset of the observables which do not suffer from these effects and are the cleanest observables from the theoretical point of view[1]. In this case the significance is reduced to above 3 σ. Interestingly, the central values of the fit to the clean observables and all observables converge to the compatible central value (cf. Fig. 7.1). This might be an indication that indeed the observed anomalies are related to NP rather than to charm-loop effects.

[1]The so-called lepton universality measurements.

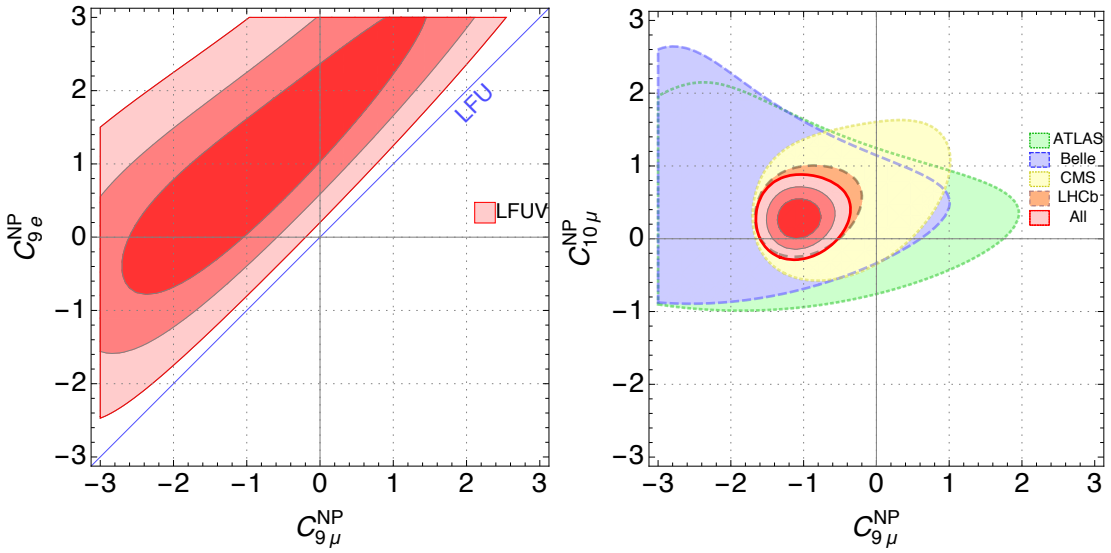


Figure 7.1: Results of 2-dim fit of CCDMV for full set of observables (right) and "LFUV" (left).

7.1.2 ANSS fit

Similarly to the CCDMV fit, the ANSS group uses a large number of observables to elucidate the possible NP contribution precisely:

- Branching ratio measurement of $B^0 \to K^{*0}\mu^+\mu^-$ done by LHCb [118], CMS [126] and CDF [127].
- Angular observables of $B^0 \to K^{*0}\mu^+\mu^-$ in the S_i basis 2.7 [69], ATLAS [96] and CMS [97].
- Branching ratio of the $B^+ \to K^+\mu^+\mu^-$ decay from LHCb [119].
- Branching ratio and angular observables of $B_s^0 \to \phi\mu^+\mu^-$ [120].
- Inclusive measurements of $B \to X_s\ell\ell$ [121].
- The LFU measurements of R_K and R_{K^*} [124, 125].

On the theoretical side, there is also a major difference between the two groups. The first one exploits the optimized variables P_i, which are calculated with the soft-form factor approximation, while the ANSS group uses the full form factors. It is worth mentioning that the ANSS fit is based on a public code `flavio` [128].

As can be seen in Table 7.2 and Fig. 7.2, the results of the ANSS group fit show over 4σ deviation from the SM, which is in agreement with the previous CCDMV fit. In addition, the fits done only to the theoretically clean observables R_K and R_{K^*} are in agreement with the other observables, as was the case of the previous fit.

Let us stress that despite the fact that in the fits one sees the $(4-5)\ \sigma$ deviations from the SM, the flavour community did not claim the discovery of NP. This is because the

Coeff.	best fit	1σ	2σ	pull
C_9^μ	-1.59	$[-2.15, -1.13]$	$[-2.90, -0.73]$	4.2σ
C_{10}^μ	$+1.23$	$[+0.90, +1.60]$	$[+0.60, +2.04]$	4.3σ
C_9^e	$+1.58$	$[+1.17, +2.03]$	$[+0.79, +2.53]$	4.4σ
C_{10}^e	-1.30	$[-1.68, -0.95]$	$[-2.12, -0.64]$	4.4σ
$C_9^\mu = -C_{10}^\mu$	-0.64	$[-0.81, -0.48]$	$[-1.00, -0.32]$	4.2σ
$C_9^e = -C_{10}^e$	$+0.78$	$[+0.56, +1.02]$	$[+0.37, +1.31]$	4.3σ
$C_9^{\prime\mu}$	-0.00	$[-0.26, +0.25]$	$[-0.52, +0.51]$	0.0σ
$C_{10}^{\prime\mu}$	$+0.02$	$[-0.22, +0.26]$	$[-0.45, +0.49]$	0.1σ
$C_9^{\prime e}$	$+0.01$	$[-0.27, +0.31]$	$[-0.55, +0.62]$	0.0σ
$C_{10}^{\prime e}$	-0.03	$[-0.28, +0.22]$	$[-0.55, +0.46]$	0.1σ

Table 7.2: Puls of the ANSS fit [117].

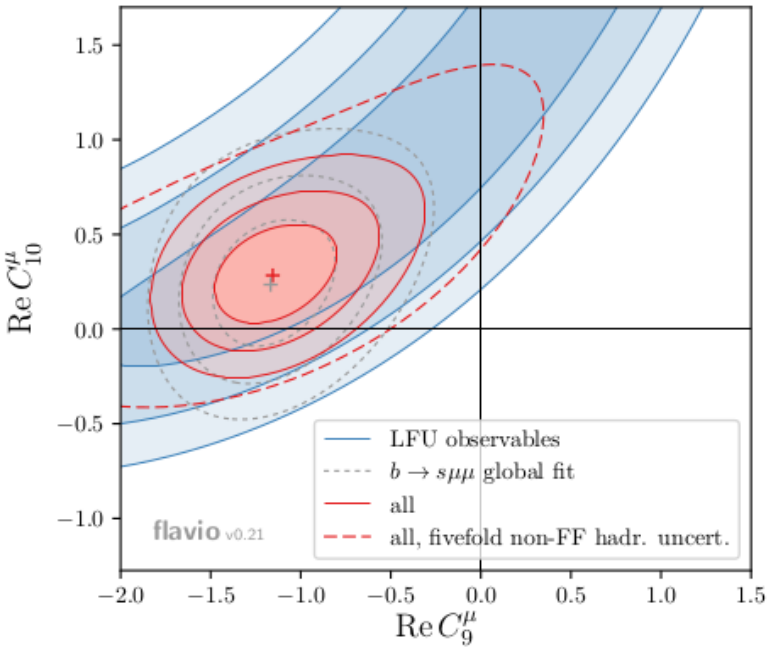


Figure 7.2: Results of ANSS fit [117].

observed discrepancies can be due to the QCD effects rather than NP. With more collected data by the LHCb collaboration this matter might be addressed in a more definite way.

7.1.3 GAMBIT WC fit

The author of this dissertation is a member of the GAMBIT (The Global and Modular Beyond-the-Standard-Model Inference Tool) collaboration [129], where he is the convenor of the Flavour Working group [130]. The GAMBIT collaboration is mainly focused on performing fits parametrising the searches for supersymmetry [131–133]. However, the FlavBit module in GAMBIT that is responsible for calculating the likelihood based on the Flavour observables, has the possibility of performing the fits of WC in a manner similar to the previously presented fits [115, 117]. In the results of the GAMBIT fit the following measurements have been used:

- Angular observables of $B^0 \to K^{*0}\mu^+\mu^-$ in the S_i basis 2.7 [69].
- Branching ratios of $B_{s/d} \to \mu\mu$ [10, 11].
- Inclusive branching ratio of $b \to s\gamma$ [134].

We have found that our fit, similarly to the previous ones, also prefers the negative modification of C_9 WC, as can be seen in Fig. 7.3.

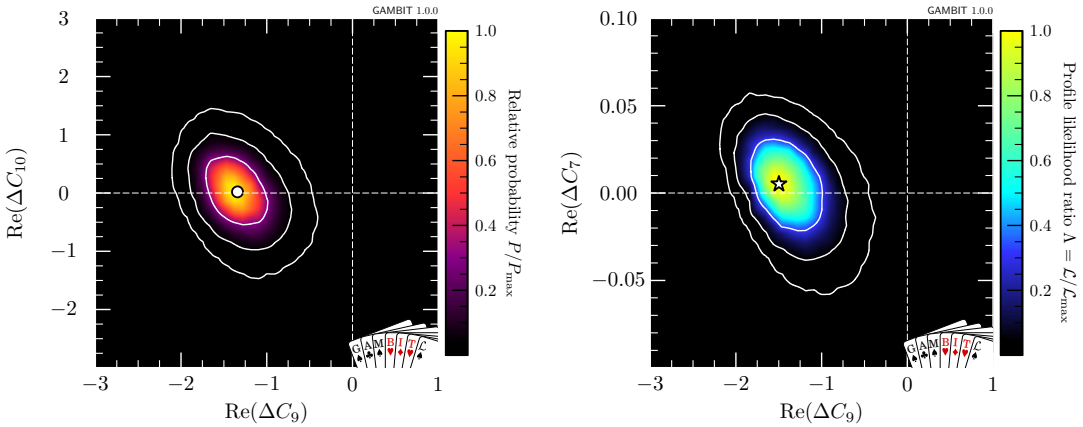


Figure 7.3: Results of WC fit using FlavBit [130].

In future, FlavBit will be extended to other observables to reach a similar complexity as the aforementioned fits.

7.1.4 Fitting the nonfactoriable corrections

As was presented in Sec. 7.1.1 and 7.1.2, the measured deviations from the SM have generated a lot of discussion on the robustness of the theory predictions. In a natural way, this leads to the idea of extracting the non factorizable corrections from the LHCb measurements directly. The first attempt has been done by the group led by L. Silvestrini [135] (called CFFMPSV after the authors' surnames). The idea behind this parametrisation is to describe the charm-loop amplitude in a similar way as [136]:

$$
\begin{aligned}
h_\lambda(q^2) &= \frac{\epsilon^*_\mu(\lambda)}{m_B^2}\int d^4x e^{iqx}\langle \bar{K}^*|T\{j^\mu_{\rm em}(x)\mathcal{H}^{\rm had}_{\rm eff}(0)\}|\bar{B}\rangle \\
&= h^{(0)}_\lambda + \frac{q^2}{1\,{\rm GeV}^2}h^{(1)}_\lambda + \frac{q^4}{1\,{\rm GeV}^4}h^{(2)}_\lambda\,, \qquad (7.2)
\end{aligned}
$$

which are related to the amplitudes calculated by A. Khodjamirian et. al. [29] via:

$$
\begin{aligned}
\tilde{g}^{\mathcal{M}_1} &= -\frac{1}{2C_1}\frac{16m_B^3(m_B+m_{K^*})\pi^2}{\sqrt{\lambda(q^2)}V(q^2)q^2}\left(h_-(q^2)-h_+(q^2)\right)\,, \\
\tilde{g}^{\mathcal{M}_2} &= -\frac{1}{2C_1}\frac{16m_B^3\pi^2}{(m_B+m_{K^*})A_1(q^2)q^2}\left(h_-(q^2)+h_+(q^2)\right)\,, \qquad (7.3)\\
\tilde{g}^{\mathcal{M}_3} &= \frac{1}{2C_1}\left[\frac{64\pi^2 m_B^3 m_{K^*}\sqrt{q^2}(m_B+m_{K^*})}{\lambda(q^2)A_2(q^2)q^2}h_0(q^2)\right. \\
&\quad \left.-\frac{16m_B^3\pi^2(m_B+m_{K^*})(m_B^2-q^2-m_{K^*}^2)}{\lambda(q^2)A_2(q^2)q^2}\left(h_-(q^2)+h_+(q^2)\right)\right]\,.
\end{aligned}
$$

The new feature is that the next coefficient to the Taylor expansion is added to account for possible effects when approaching the charm threshold. Next, the authors use the measurements presented in Sec. 4 to extract the h^i_λ coefficients. The results of the fit are shown in Fig. 7.4.

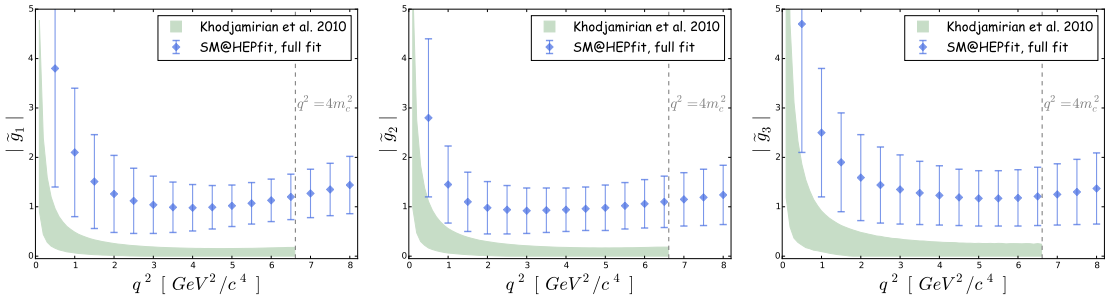


Figure 7.4: Results of the fit to the to $\tilde{g}^{\mathcal{M}_i}$ functions [135].

As can be seen, in order to fit the data correctly one has to have a significantly larger charm-loop contribution than the one calculated by theoreticians [29]. It is worth pointing out that it is not surprising that the 18 free parameters added to the fit do improve the fit quality to experimental data. Furthermore, one has to note that the fit prefers a rather constant contribution to the $\tilde{g}^{\mathcal{M}_i}$, which would have the same effect as if the NP modifying the WC, which is q^2 independent. This kind of approach has not provided the definite statement about the nature of the observed effects and more development on the theory and the experimental side is needed.

7.1.5 Fitting the analytic structure of long distance effects

In the previous Sec. 7.1.4 we have shown the first attempt to extract the long distance effect by parametrising the local correlator with third order polynomial. Here, we will present a more robust and model independent parametrisation, conducted by the author of this monograph with C. Bobeth, D. van Dyk and J. Virto in [137]. We will start from a generic hadronic transition matrix element [29]:

$$\mathcal{H}^\mu(p,q) \equiv i \int \mathrm{d}^4x\, e^{iq\cdot x} \langle \bar{K}^*(p-q,\eta)|T\{J^\mu_{\text{e.m.}}(x), C_1O_1(0) + C_2O_2(0)\}|\bar{B}(p)\rangle\,, \tag{7.4}$$

which we decompose into the vector valued correlators $\mathcal{H}_i$:

$$\mathcal{H}^\mu(p,k) \equiv M_B^2\, \eta^*_\alpha \left[S_\perp^{\alpha\mu}\mathcal{H}_\perp(q^2) - S_\parallel^{\alpha\mu}\mathcal{H}_\parallel(q^2) - S_0^{\alpha\mu}\mathcal{H}_0(q^2)\right]\,, \tag{7.5}$$

where $S_\lambda^{\alpha\mu}$ are the Lorentz structures:

$$\begin{aligned}
S_\perp^{\alpha\mu} &= \frac{\sqrt{2}M_B}{\sqrt{\lambda}}\varepsilon^{\alpha\mu kq}\,,\\
S_\parallel^{\alpha\mu} &= \frac{iM_B}{\sqrt{2}\lambda}\Big[\lambda g^{\alpha\mu} + 4M_{K^*}^2 q^\alpha q^\mu - 4(q\cdot k)\, q^\alpha k^\mu\Big]\,,\\
S_0^{\alpha\mu} &= -\frac{i\,4M_{K^*}(M_B+M_{K^*})}{\lambda\sqrt{q^2}}\Big[(q\cdot k)\, q^\alpha q^\mu - q^2\, q^\alpha k^\mu\Big]\,.
\end{aligned}$$

In the previous approach (Sec. 7.5), the h_λ defined in Eq. 7.2 enter the $\mathcal{H}_i$ correlators in the following way:

$$H_V^\lambda = -iN\left\{C_9^{\text{eff}}\tilde{V}_{L\lambda} + \frac{m_B^2}{q^2}\left[\frac{2m_b}{m_B}C_7^{\text{eff}}\tilde{T}_{L\lambda} - 16\pi^2 h_\lambda(q^2)\right]\right\}\,, \tag{7.6}$$

The correlators $\mathcal{H}_i$ are functions of complex plain with branch cuts and poles, while the CFFMPSV parametrisation does not have any of those properties. We have proposed a much more adequate parametrisation. In order to explore the properties of the poles and branch cuts we, transform q^2:

$$z(q^2) \equiv \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}\,, \tag{7.7}$$

where $t_+ = 4M_D^2$ and $t_0 = t_+ - \sqrt{t_+(t_+ - M^2_{\psi(2S)})}$. This transformation maps the $c\bar{c}$ branch cut in the q^2 plane to the unit circumference $|z| = 1$, and the entire first Riemann sheet in the q^2 plane to the interior of the unit circle $|z| < 1$. This has been expired by the

z-expansion for the form factors [138]. Now the $\mathcal{H}^{\mu}(p,k)$ can be Taylor expanded around $z=0$. We propose the following parametrisation:

$$\mathcal{H}_{\lambda}(z) = \frac{1 - z\, z^{*}_{J/\psi}}{z - z_{J/\psi}} \frac{1 - z\, z^{*}_{\psi(2S)}}{z - z_{\psi(2S)}} \hat{H}_{\lambda}(z), \tag{7.8}$$

where

$$\mathcal{H}_{\lambda}(z) = \hat{H}_{\lambda}(z) = \Big[\sum_{k=0}^{K} \alpha_{k}^{(\lambda)} z^{k} \Big] \mathcal{F}_{\lambda}(z), \tag{7.9}$$

and $\mathcal{F}$ are the form factors and $\alpha_{k}^{(\lambda)}$ are complex coefficients. We end the expansion after $k = 2$[2]. The fit to this model is performed using both theoretical and experimental results.

On the theoretical side we use the calculation results from [29] for $\mathcal{H}_{\lambda}(z)$ for points $q^2 = [-7, -5, -3, -1]$ GeV$^2/c^4$. These calculations are used as observables to constrain the $H_{\lambda}(q^2)/\mathcal{F}_{\lambda}(q^2)$. The ratios are later used to constrain the Eq. 7.8.

From the experimental constraints we use the amplitudes of decays: $B \to K^{*}\psi_{n}$, where $\psi_1 = J/\psi$ and $\psi_2 = \psi(2S)$. According to the LSZ formula, the decay amplitudes are defined by the residual of the functions $H_{\lambda}(q^2)$ pole:

$$\mathcal{H}_{\lambda}(q^2 \to M^2_{\psi_n}) \sim frac M_{\psi_n} f^{*}_{\psi_n} \mathcal{A}^{\psi_n}_{\lambda} M_B^2 (q^2 - M^2_{\psi_n}) + \cdots , \tag{7.10}$$

where $\mathcal{A}^{\psi_n}$ are transversity amplitudes that have been measured by the Babar [85, 139], Belle [86, 102, 140] and LHCb [141] experiments. There are three transversity amplitudes, so we can produce five observables[3] in total:

$$|r^{\psi_n}_{\perp}|,\ |r^{\psi_n}_{\parallel}|,\ |r^{\psi_n}_{0}|,\ \arg\{r^{\psi_n}_{\perp} r^{\psi_n *}_{0}\},\ \arg\{r^{\psi_n}_{\parallel} r^{\psi_n *}_{0}\}, \tag{7.11}$$

where

$$r^{\psi_n}_{\lambda} \equiv \operatorname*{Res}_{q^2 \to M^2_{\psi_n}} \frac{\mathcal{H}_{\lambda}(q^2)}{\mathcal{F}_{\lambda}(q^2)} \sim \frac{M_{\psi_n} f^{*}_{\psi_n} \mathcal{A}^{\psi_n}_{\lambda}}{M_B^2\, \mathcal{F}_{\lambda}(M^2_{\psi_n})} \ . \tag{7.12}$$

The numerical values of these observables are calculated using pseudo-experiments.

A fit is then performed to the above-mentioned experimental and theoretical constraints using Eq. 7.8. As a result , one obtains the set of $\alpha_{k}^{(\lambda)}$ parameters and the correlation matrix of them. This allows us to calculate the SM prediction for full set of observables in the $B^0 \to K^{*0}\mu^{+}\mu^{-}$ decay. As can be seen in Fig. 7.5, the results of the calculation show a clear tension w.r.t. to the measured values described in Chapter 4 of this monograph. It is worth stressing out that this is the most precise determination of the SM prediction of the $B^0 \to K^{*0}\mu^{+}\mu^{-}$ decay.

We perform the second fit including the observables from Sec. 4 to determine the value of C_9 WC, using as prior the previously obtained correlator function. Both likelihood (LLH) and method of moments (MOM) results are being taken into account. Furthermore,

[2]The value was chosen as the lowest value that provides good fit to the data.

[3]One phase can be chosen arbitrarily.

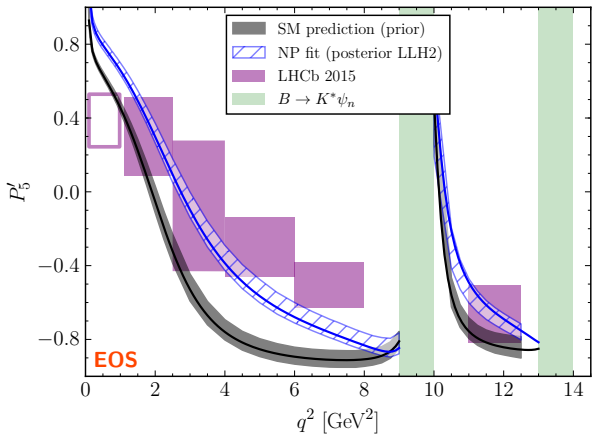


Figure 7.5: Prior and posteriori predictions for P_5' with the SM prediction and NP fit. The data points correspond to the ones measured in Chapter 4 of this monograph.

we study two scenarios: the first one does not include the iterresonance bin[4], while the second one does. The results of this fit read:

$$(\text{LLH}) : C_9 = 2.51 \pm 0.29\,, \tag{7.13}$$
$$(\text{LLH2}) : C_9 = 3.01 \pm 0.25\,, \tag{7.14}$$
$$(\text{MOM}) : C_9 = 2.81 \pm 0.37\,, \tag{7.15}$$
$$(\text{MOM2}) : C_9 = 3.20 \pm 0.31\,, \tag{7.16}$$

which show a clear tension w.r.t. the SM prediction for C_9. This is illustrated in Fig. 7.6.

This fits are in perfect agreement with the other global fits in Sec. 7.1.1 - 7.1.3. It is worth stressing out that this is the first fit that considers the fundamental structure of the long distance effects, as was done in the above described fit. Last but not least, this approach can be extended to other decays of the type $b \to s\ell^-\ell^+$.

7.2 Possible New Physics contributions

In this section we will give a short overview of possible NP scenarios which could explain the observed anomalies. There are many models that fit individually or some groups of anomalies quite well. However, if NP is to be claimed in the flavour physics, it has to form a consistent pattern and provide a global solution to the anomalies. Because of rich literature and space limitations of this dissertation, we will just list the most popular models and give references that the reader can follow if interested.

One of the most popular groups of models is the one postulating the existence of the leptoquarks [142–151]. This type of models introduces a new kind of particles (called

[4]This is the bin between the J/ψ and $\psi(2S)$ resonances.

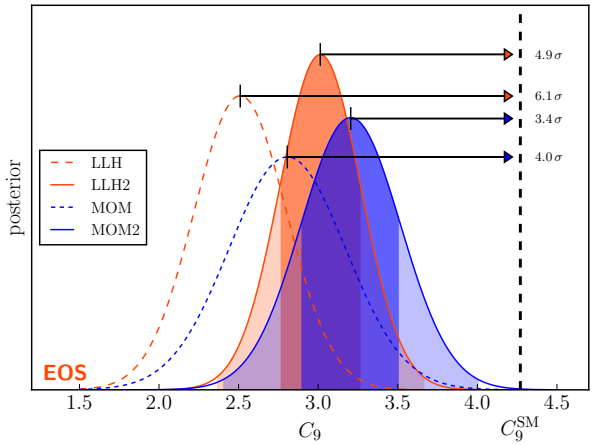


Figure 7.6: Posteriori distributions for the C_9 WC for the fit.The dark and light regions correspond to 68 % and 99 % probability band.

leptoquarks) that allow the leptons and quarks to interact with each other. Thus the leptoquarks carry both the lepton and baryon number. An example of a Feynman diagram with leptoquarks is shown in Fig. 7.7, left side.

The second most popular theory includes the non-universal and flavour changing Z' models [152–156]. Those models have a small disadvantage compared to leptoquarks: they cannot explain the observed anomaly of $R(D^*)$ [122]. A possible Feyman diagram explaining $b \to s\ell^-\ell^+$ in this framework is shown in Fig. 7.7, right side.

The review of other models can be found in [157].

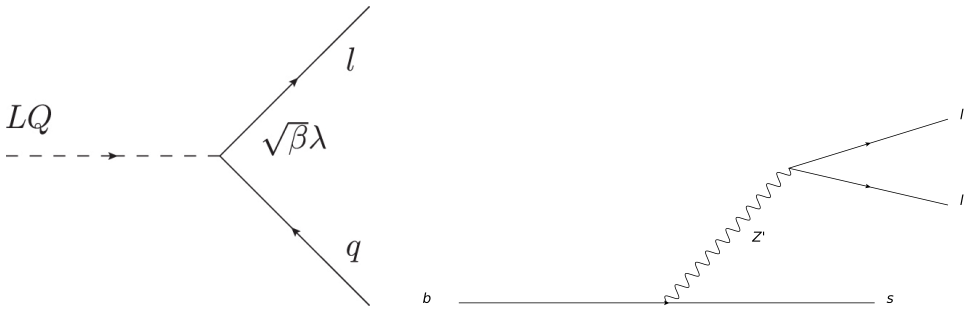


Figure 7.7: Example of a leptoquark interaction(left) and Z' (right).

Chapter 8

Conclusion

In this monograph three experimental studies of the $b \to s\ell^-\ell^+$ transitions performed by the author have been presented. The angular analysis $B^0 \to K^{*0}\mu^+\mu^-$ showed a tension w.r.t. the SM predictions, which is especially profound in the so-called P_5' observable, that to a large extend is free from form factor uncertainties. The analysis is one of the golden channels in LHCb and has generated a huge interest in the flavour physics community. The second study was dedicated to the measurement of higher moments and branching fraction of the $B^0 \to K^+\pi^-\mu^+\mu^-$ decay in the $K^*_{0,2}(1430)^0$ resonance region. This has been the first observation of this decay and with advances on the theory side it has the potential to be a complementary measurement in terms of NP sensitivity to the $B^0 \to K^{*0}\mu^+\mu^-$. The third analysis was a search for a hypothetical new particle called inflaton. In this case we have managed to exclude most of the inflaton parameters space.

When performing the above-mentioned studies, the author has introduced several novel experimental techniques such as k-Folding, which has now become a standard in the LHCb collaboration. Moreover, he has authored a paper proposing the application of the method of moments to rare decays, later used by him in the measurement. In this way the author for the first time has measured the angular observable S_{6c}.

The angular distribution of $B^0 \to K^{*0}\mu^+\mu^-$ together with the other $b \to s\ell^-\ell^+$ measurements forms a consistent pattern of deviations from the SM. The global analysis has led to the significance over 5σ. Unfortunately, the SM predictions suffer from QCD corrections that could not have been taken into account properly while computing the SM predictions and uncertainties. The future measurements together with improvements on the theory side will hopefully be able to shed light on whether we are on the brick of discovering NP or whether nature has played a cruel trick on us with the existence of large charm-loops effects.

Bibliography

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

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

[3] L. Susskind, *Dynamics of spontaneous symmetry breaking in the weinberg-salam theory*, Phys. Rev. D **20** (1979) 2619.

[4] R. Barbieri and G. F. Giudice, *Upper bounds on supersymmetric particle masses*, Nuclear Physics B **306** (1988), no. 1 63 .

[5] *Nobel Prize 2015*, http://www.nobelprize.org/nobel_prizes/physics/laureates/2015/.

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

[7] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, in *Probing the standard model of particle interactions. Proceedings, Summer School in Theoretical Physics, NATO Advanced Study Institute, 68th session, Les Houches, France, July 28-September 5, 1997. Pt. 1, 2*, pp. 281–539, 1998. arXiv:hep-ph/9806471.

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

[9] A. J. Buras, *Flavour Changing Neutral Current Processes*, .

[10] LHCb, CMS, V. Khachatryan *et al.*, *Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature **522** (2015) 68, arXiv:1411.4413.

[11] LHCb, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, Phys. Rev. Lett. **118** (2017), no. 19 191801, arXiv:1703.05747.

[12] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \to K^* \mu^+ \mu^-$ Decays in the Standard Model and Beyond*, JHEP **0901** (2009) 019, arXiv:0811.1214.

[13] F. Beaujean, M. Chrzaszcz, N. Serra, and D. van Dyk, *Extracting Angular Observables without a Likelihood and Applications to Rare Decays*, accepted by Phys. Rev. D. (2015) arXiv:1503.04100.

[14] U. Egede *et al.*, *New physics reach of the decay mode $\bar{B} \to \bar{K}^{*0} \ell^+ \ell^-$*, JHEP **10** (2010) 056, arXiv:1005.0571.

[15] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V \ell^+ \ell^-$ in the Standard Model from Light-Cone Sum Rules*, arXiv:1503.05534.

[16] A. J. Buras, M. Misiak, M. Munz, and S. Pokorski, *Theoretical uncertainties and phenomenological aspects of B —¿ X(s) gamma decay*, Nucl. Phys. **B424** (1994) 374, arXiv:hep-ph/9311345.

[17] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, *Optimizing the basis of $B \to K^* \ell^+ \ell^-$ observables in the full kinematic range*, JHEP **1305** (2013) 137, arXiv:1303.5794.

[18] J. Matias, F. Mescia, M. Ramon, and J. Virto, *Complete Anatomy of $\bar{B}_d \to \bar{K}^* (\to K\pi) \ell^+ \ell^-$ and its angular distribution*, JHEP **1204** (2012) 104, arXiv:1202.4266.

[19] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean observables for the binned analysis of $B- > K * \mu^+ \mu^-$ at large recoil*, JHEP **01** (2013) 048, arXiv:1207.2753.

[20] J. Charles *et al.*, *Heavy to light form-factors in the heavy mass to large energy limit of QCD*, Phys. Rev. **D60** (1999) 014001, arXiv:hep-ph/9812358.

[21] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Rare B decays using lattice QCD form factors*, PoS **LATTICE2014** (2015) 372, arXiv:1501.00367.

[22] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Lattice QCD calculation of form factors describing the rare decays $B \to K^* \ell^+ \ell^-$ and $B_s \to \phi \ell^+ \ell^-$*, Phys. Rev. **D89** (2014), no. 9 094501, arXiv:1310.3722.

[23] J. Flynn *et al.*, *Hadronic form factors for rare semileptonic B decays*, PoS **LATTICE2015** (2016) 345, arXiv:1511.06622.

[24] M. Beneke, T. Feldmann, and D. Seidel, *Systematic approach to exclusive $B \to V l^+ l^-$, $V\gamma$ decays*, Nucl. Phys. **B612** (2001) 25, arXiv:hep-ph/0106067.

[25] M. Beneke and T. Feldmann, *Symmetry breaking corrections to heavy to light B meson form-factors at large recoil*, Nucl. Phys. **B592** (2001) 3, arXiv:hep-ph/0008255.

[26] LHCb, R. Aaij *et al.*, *Observation of a resonance in $B^+ \to K^+\mu^+\mu^-$ decays at low recoil*, Phys. Rev. Lett. **111** (2013), no. 11 112003, arXiv:1307.7595.

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

[28] BES, M. Ablikim *et al.*, *Determination of the psi(3770), psi(4040), psi(4160) and psi(4415) resonance parameters*, eConf **C070805** (2007) 02, arXiv:0705.4500, [Phys. Lett.B660,315(2008)].

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

[30] F. Kruger and L. M. Sehgal, *Lepton polarization in the decays b —¿ X(s) mu+ mu- and B —¿ X(s) tau+ tau-*, Phys. Lett. **B380** (1996) 199, arXiv:hep-ph/9603237.

[31] https://belle.kek.jp/.

[32] http://www.slac.stanford.edu/BFROOT/.

[33] ATLAS Collaboration. home.web.cern.ch/about/experiments/atlas.

[34] CMS Collaboration. home.web.cern.ch/about/experiments/cms.

[35] ATLAS, CMS, V. P. Andreev, *B production at the LHC / QCD aspects*, in *Proceedings, 42nd Rencontres de Moriond on QCD and High Energy Hadronic Interactions: La Thuile, Italy, March 17-24, 2007*, 2007. arXiv:0706.1789.

[36] CERN. /home.cern/topics/large-hadron-collider.

[37] CERN. home.web.cern.ch/.

[38] CERN accelerator Complex. http://public.web.cern.ch/public/en/research/AccelComplex-en.html.

[39] Booster. home.web.cern.ch/about/accelerators/proton-synchrotron-booster.

[40] PS. home.web.cern.ch/about/accelerators/proton-synchrotron.

[41] SPS. home.web.cern.ch/about/accelerators/super-proton-synchrotron.

[42] CNGS. https://home.cern/about/accelerators/cern-neutrinos-gran-sasso.

[43] NA62 Collaboration. https://na62.web.cern.ch/na62/.

[44] Test Beam Areas. http://sba.web.cern.ch/sba/.

[45] LHCb Collaboration. home.web.cern.ch/about/experiments/lhcb.

[46] ALICE Collaboration. home.web.cern.ch/about/experiments/alice.

[47] TOTEM Collaboration. home.web.cern.ch/about/experiments/totem.

[48] LHCf Collaboration. home.web.cern.ch/about/experiments/lhcf.

[49] MOEDAL Collaboration. home.web.cern.ch/about/experiments/moedal.

[50] LHCb Operation Group. http://lhcb-operationsplots.web.cern.ch/lhcb-operationsplots/index.htm.

[51] LHCb Speakers Bureau. https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[52] LHCb Collaboration, R. Aaij *et al.*, *Measurement of* $\sigma(pp \to b\bar{b}X)$ *at* $\sqrt{s} = 7$ TeV *in the forward region*, Phys. Lett. **B694** (2010) 209, arXiv:1009.2731.

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

[54] LHCb Collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb VELO (VErtex LOcator): Technical Design Report*, .

[55] LHCb Collaboration, P. R. Barbosa-Marinho *et al.*, *LHCb inner tracker: Technical Design Report*, , revised version number 1 submitted on 2002-11-13 14:14:34.

[56] P. A. Cherenkov, *Visible emission of clean liquids by action of radiation*, Doklady Akademii Nauk SSSR **2** (1934) 451+.

[57] Hybrid Photon Detectors. tilde-gys.web.cern.ch/~gys/LHCb/PixelHPDs.htm.

[58] LHCb Collaboration, S. Amato *et al.*, *LHCb RICH: Technical Design Report*, .

[59] LHCb Collaboration, S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*, .

[60] LHCb, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.

[61] J. Alves *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) 2022P, 1211.1346.

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

[63] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, 1211.3055.

[64] ATLAS, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[65] M. Capeans *et al.*, *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, Sep, 2010.

[66] E. Focardi *et al.*, *The cms silicon strip tracker*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **435** (1999), no. 1 102 .

[67] T. Sumiyoshi *et al.*, *Silica aerogel Cherenkov counter for the KEK B factory experiment*, Nucl. Instrum. Meth. **A433** (1999) 385.

[68] Belle-II, T. Abe *et al.*, *Belle II Technical Design Report*, arXiv:1011.0352.

[69] LHCb, R. Aaij *et al.*, *Angular analysis of the* $B^0 \to K^{*0}\mu^+\mu^-$ *decay*, arXiv:1512.04442.

[70] DAVINCI. lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci.

[71] BRUNEL. lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel.

[72] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, European Physical Journal C **73** (2013) 2431, arXiv:1211.6759.

[73] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, physics/07.

[74] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[75] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119.

[76] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[77] C. Bobeth, G. Hiller, and G. Piranishvili, *CP Asymmetries in bar* $B \to \bar{K}^*(\to \bar{K}\pi)\bar{\ell}\ell$ *and Untagged* $\bar{B}_s$, $B_s \to \phi(\to K^+K^-)\bar{\ell}\ell$ *Decays at NLO*, JHEP **07** (2008) 106, arXiv:0805.2525.

[78] T. Blake and C. Langenbruch, *Angular conventions for the decays* $B^0 \to K^{*0}\mu^+\mu^-$ *and* $B_s^0 \to \phi\mu^+\mu^-$, LHCb-INT-2012-021, 2012.

[79] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes in* $B^0 \to J/\psi K^*(892)^0$ *decays*, Phys. Rev. **D88** (2013) 052002, arXiv:1307.2782.

[80] MINUIT. https://seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/.

[81] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, ArXiv Physics e-prints (2003) arXiv:physics/0306116.

[82] D. van Dyk *et al.*, *EOS — A HEP program for Flavor Observables*, 2016. https://eos.github.io.

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

[84] B. Sen, M. Walker, and M. Woodroofe, *On the unified method with nuisance parameters*, Statistica Sinica **19** (2009) 301.

[85] BaBar collaboration, B. Aubert *et al.*, *Measurement of decay amplitudes of $B \to J/\psi K^*, \psi(2S)K^*$, and $\chi_{c1}K^*$ with an angular analysis*, Phys. Rev. **D76** (2007) 031102, arXiv:0704.0522.

[86] Belle collaboration, R. Itoh *et al.*, *Studies of CP violation in $B^0 \to J/\psi K^*$ decays*, Phys. Rev. Lett. **95** (2005) 091601, arXiv:hep-ex/0504030.

[87] CDF collaboration, *Angular analysis of $B_s^0 \to J/\psi\phi$ and $B^0 \to J/\psi K^*$ decays and measurement of $\Delta\Gamma_s$ and ϕ_s*, CDF public note 8950, www-cdf.fnal.gov/physics/new/bottom/bottom.html.

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

[89] D. Aston *et al.*, *A Study of K- pi+ Scattering in the Reaction K- p —¿ K- pi+ n at 11-GeV/c*, Nucl. Phys. **B296** (1988) 493.

[90] F. James, *Statistical methods in experimental physics*, .

[91] A. S. Dighe, I. Dunietz, and R. Fleischer, *Extracting CKM phases and $B_s - \bar{B}_s$ mixing parameters from angular distributions of nonleptonic B decays*, Eur. Phys. J. **C6** (1999) 647, arXiv:hep-ph/9804253.

[92] LHCb, R. Aaij *et al.*, *Measurement of the $\bar{B}^0 - B^0$ and $\bar{B}_s^0 - B_s^0$ production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218, arXiv:1408.0275.

[93] HepData. https://hepdata.net/record/ins1486676.

[94] W. Altmannshofer and D. M. Straub, *New physics in $b \to s$ transitions after LHC run 1*, Eur. Phys. J. **C75** (2015), no. 8 382, arXiv:1411.3161.

[95] Belle, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^* \ell^+ \ell^-$*, Phys. Rev. Lett. **118** (2017), no. 11 111801, arXiv:1612.05014.

[96] ATLAS Collaboration, *Angular analysis of $B_d^0 \to K^* \mu^+ \mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2017-023, CERN, Geneva, Apr, 2017.

[97] CMS Collaboration, *Measurement of the P_1 and P'_5 angular parameters of the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Tech. Rep. CMS-PAS-BPH-15-008, CERN, Geneva, 2017.

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

[99] LHCb, R. Aaij *et al.*, *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*, JHEP **12** (2016) 065, arXiv:1609.04736.

[100] C.-D. Lu and W. Wang, *Analysis of $B \to K^*_J(\to K\pi)\mu^+\mu^-$ in the higher kaon resonance region*, Phys. Rev. **D85** (2012) 034014, arXiv:1111.1513.

[101] B. Dey, *Angular analyses of exclusive $\overline{B} \to x\ell_1\ell_2$ with complex helicity amplitudes*, Phys. Rev. D **92** (2015) 033013.

[102] Belle, K. Chilikin *et al.*, *Observation of a new charged charmoniumlike state in $\bar{B}^0 \to J/\psi K^-\pi^+$ decays*, Phys. Rev. **D90** (2014), no. 11 112009, arXiv:1408.6457.

[103] LHCb, R. Aaij *et al.*, *Search for long-lived scalar particles in $B^+ \to K^+\chi(\mu^+\mu^-)$ decays*, Phys. Rev. **D95** (2017), no. 7 071101, arXiv:1612.07818.

[104] F. Bezrukov and D. Gorbunov, *Light inflaton Hunter's Guide*, JHEP **1005** (2010) 010, arXiv:0912.0390.

[105] F. Bezrukov and D. Gorbunov, *Relic Gravity Waves and 7 keV Dark Matter from a GeV scale inflaton*, Phys. Lett. **B736** (2014) 494, arXiv:1403.4638.

[106] M. P. Hertzberg and J. Karouby, *Generating the Observed Baryon Asymmetry from the Inflaton Field*, Phys. Rev. **D89** (2014), no. 6 063523, arXiv:1309.0010.

[107] M. P. Hertzberg and J. Karouby, *Baryogenesis from the Inflaton Field*, Phys. Lett. **B737** (2014) 34, arXiv:1309.0007.

[108] R. D. Peccei and H. R. Quinn, *CP conservation in the presence of pseudoparticles*, Physical Review Letters **38** (1977) 1440.

[109] PIDCalib. twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage.

[110] DecayTreeFitter. https://twiki.cern.ch/twiki/bin/view/LHCb/DecayTreeFitter.

[111] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

[112] B. Batell, M. Pospelov, and A. Ritz, *Multi-lepton Signatures of a Hidden Sector in Rare B Decays*, Phys. Rev. **D83** (2011) 054005, arXiv:0911.4938.

[113] LHCb Collaboration, R. Aaij, *et al.*, *Search for hidden-sector bosons in $B^0 \to K^{*0}\mu^+\mu^-$ decays*, Phys. Rev. Lett. **115** (2015) 161802.

[114] F. Bergsma *et al.*, *Search for axion-like particle production in 400 gev proton-copper interactions*, Physics Letters B **157** (1985), no. 5 458 .

[115] B. Capdevila *et al.*, *Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data*, arXiv:1704.05340.

[116] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *Global analysis of $b \to s\ell\ell$ anomalies*, arXiv:1510.04239.

[117] W. Altmannshofer, C. Niehoff, P. Stangl, and D. M. Straub, *Status of the $B \to K^*\mu^+\mu^-$ anomaly after Moriond 2017*, Eur. Phys. J. **C77** (2017), no. 6 377, arXiv:1703.09189.

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

[119] LHCb, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the $B^+ \to K^+\mu^+\mu^-$ decay*, JHEP **02** (2013) 105, arXiv:1209.4284.

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

[121] BaBar, J. P. Lees *et al.*, *Measurement of the $B \to X_s l^+ l^-$ branching fraction and search for direct CP violation from a sum of exclusive final states*, Phys. Rev. Lett. **112** (2014) 211802, arXiv:1312.5364.

[122] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, arXiv:1612.07233.

[123] BaBar, B. Aubert *et al.*, *Measurement of Branching Fractions and CP and Isospin Asymmetries in $B \to K^*(892)\gamma$ Decays*, Phys. Rev. Lett. **103** (2009) 211802, arXiv:0906.2177.

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

[125] LHCb, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, arXiv:1705.05802.

[126] CMS, V. Khachatryan *et al.*, *Angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B753** (2016) 424, arXiv:1507.08126.

[127] CDF, T. Aaltonen *et al.*, *Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b \to \Lambda\mu^+\mu^-$*, Phys. Rev. Lett. **107** (2011) 201802, arXiv:1107.3753.

[128] D. Straub, *flav-io/flavio v0.20.3*, Mar., 2017. https://doi.org/10.5281/zenodo.438351, doi: 10.5281/zenodo.438351.

[129] GAMBIT, P. Athron *et al.*, *GAMBIT: The Global and Modular Beyond-the-Standard-Model Inference Tool*, arXiv:1705.07908.

[130] GAMBIT, F. U. Bernlochner *et al.*, *FlavBit: A GAMBIT module for computing flavour observables and likelihoods*, arXiv:1705.07933.

[131] GAMBIT, P. Athron *et al.*, *Global fits of GUT-scale SUSY models with GAMBIT*, arXiv:1705.07935.

[132] GAMBIT, P. Athron *et al.*, *Status of the scalar singlet dark matter model*, arXiv:1705.07931.

[133] GAMBIT, P. Athron *et al.*, *A global fit of the MSSM with GAMBIT*, arXiv:1705.07917.

[134] M. Misiak and M. Steinhauser, *Weak Radiative Decays of the B Meson and Bounds on $M_{H^\pm}$ in the Two-Higgs-Doublet Model*, Eur. Phys. J. **C77** (2017), no. 3 201, arXiv:1702.04571.

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

[136] S. Jger and J. Martin Camalich, *On $B \to V \ell\ell$ at small dilepton invariant mass, power corrections, and new physics*, JHEP **05** (2013) 043, arXiv:1212.2263.

[137] C. Bobeth, M. Chrzaszcz, D. van Dyk, and J. Virto, *Long-distance effects in $B \to K^* \ell\ell$ from Analyticity*, arXiv:1707.07305.

[138] C. G. Boyd, B. Grinstein, and R. F. Lebed, *Model independent extraction of —V(cb)— using dispersion relations*, Phys. Lett. **B353** (1995) 306, arXiv:hep-ph/9504235.

[139] BaBar, B. Aubert *et al.*, *Measurement of branching fractions and charge asymmetries for exclusive B decays to charmonium*, Phys. Rev. Lett. **94** (2005) 141801, arXiv:hep-ex/0412062.

[140] Belle, K. Chilikin *et al.*, *Experimental constraints on the spin and parity of the Z(4430)$^+$*, Phys. Rev. **D88** (2013), no. 7 074026, arXiv:1306.4894.

[141] LHCb, R. Aaij *et al.*, *Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays*, Phys. Rev. **D88** (2013) 052002, arXiv:1307.2782.

[142] B. Chauhan, B. K. and, and A. Narang, *A Leptoquark explanation for $(g-2)_\mu$, R_K, $R_{K^\star}$ and, IceCube PeV events*, arXiv:1706.04598.

[143] O. Sumensari, *Leptoquark models for the B-physics anomalies*, 2017. arXiv:1705.07591.

[144] Y. Cai, J. Gargalionis, M. A. Schmidt, and R. R. Volkas, *Reconsidering the One Leptoquark solution: flavor anomalies and neutrino mass*, arXiv:1704.05849.

[145] D. Beirevi and O. Sumensari, *A leptoquark model to accommodate* $R_K^{\rm exp} < R_K^{\rm SM}$ *and* $R_{K^*}^{\rm exp} < R_{K^*}^{\rm SM}$, arXiv:1704.05835.

[146] A. Crivellin, D. Mller, and T. Ota, *Simultaneous Explanation of* $R(D^{(*)})$ *and* $b \to s\mu^+\mu^-$*: The Last Scalar Leptoquarks Standing*, arXiv:1703.09226.

[147] C.-H. Chen, T. Nomura, and H. Okada, *Excesses of muon* $g-2$*,* $R_{D^{(*)}}$*, and* R_K *in a leptoquark model*, arXiv:1703.03251.

[148] P. Cox, A. Kusenko, O. Sumensari, and T. T. Yanagida, *SU(5) Unification with TeV-scale Leptoquarks*, JHEP **03** (2017) 035, arXiv:1612.03923.

[149] R. Barbieri, C. W. Murphy, and F. Senia, *B-decay Anomalies in a Composite Leptoquark Model*, Eur. Phys. J. **C77** (2017), no. 1 8, arXiv:1611.04930.

[150] O. Popov and G. A. White, *One Leptoquark to unify them? Neutrino masses and unification in the light of* $(g-2)_\mu$*,* $R_{D^{(\star)}}$ *and* R_K *anomalies*, arXiv:1611.04566.

[151] M. Bauer and M. Neubert, *Minimal Leptoquark Explanation for the* $R_{D^{(*)}}$ *,* R_K *, and* $(g-2)_g$ *Anomalies*, Phys. Rev. Lett. **116** (2016), no. 14 141802, arXiv:1511.01900.

[152] W. Altmannshofer, C.-Y. Chen, P. S. Bhupal Dev, and A. Soni, *Lepton flavor violating Z explanation of the muon anomalous magnetic moment*, Phys. Lett. **B762** (2016) 389, arXiv:1607.06832.

[153] M. Jung, *Non-universal Z' models with protected flavour-changing interactions*, in *Proceedings, 51st Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Italy, March 12-19, 2016*, pp. 335–340, 2016. arXiv:1606.09191.

[154] C.-W. Chiang, X.-G. He, J. Tandean, and X.-B. Yuan, $R_{K^{(*)}}$ *and related* $b \to s\ell\bar{\ell}$ *anomalies in minimal flavor violation framework with* Z' *boson*, arXiv:1706.02696.

[155] D. Banerjee and S. Sahoo, *Analysis of* $\Lambda_b \to \Lambda l^+ l^-$ *rare decays in a non-universal* Z' *model*, arXiv:1705.10458.

[156] S. Di Chiara *et al.*, *Minimal flavor-changing* Z' *models and muon* $g-2$ *after the* R_{K^*} *measurement*, arXiv:1704.06200.

[157] D. Buttazzo, A. Greljo, G. Isidori, and D. Marzocca, *B-physics anomalies: a guide to combined explanations*, arXiv:1706.07808.