

Measurement of branching fractions, isospin asymmetries and angular observables in exclusive electroweak penguin decays

Patrick Haworth Owen

High Energy Physics
Blackett Laboratory
Imperial College London

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy

Abstract

This thesis describes measurements of rare electroweak penguin decays performed with data collected by the Large Hadron Collider beauty experiment corresponding to 3 fb^{-1} of integrated luminosity. The purpose of these measurements is to search for physics beyond the theoretical framework known as the Standard Model (SM). Electroweak penguin decays are sensitive to virtual particles in extensions to the SM whose influence on the decay amplitude can be of similar strength to the SM contribution. The particular measurements that are described in this thesis are the differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)}\mu^+\mu^-$ decays as well as the angular observables in $B \rightarrow K\mu^+\mu^-$ decays. Although results are consistent with the SM, all the branching fractions of $B \rightarrow K^{(*)}\mu^+\mu^-$ decays tend to favour a lower value than theoretical predictions.

Declaration

The work presented in this thesis is the result of collaborations between myself and members of the LHCb collaboration which was carried out between Jan 2013 and March 2014 and was published Refs. [1–3]. Prior to this I was involved in analysis, published in Refs. [4, 5], which is not described in this thesis. All the analysis work was performed by myself.

This thesis has not been submitted for any other qualification.

Patrick Owen, May 2014

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the license terms of this work (Imperial College London PhD regulations).

Acknowledgements

Firstly, I would like to thank my supervisor, Ulrik Egede, who struck the perfect balance between freedom to get on with it, whilst being available if I needed help. I do not take for granted how lucky I was to pursue the topics he proposed during my Ph.D. Thanks also go to my colleagues whom I worked with on the LHCb experiment, in particular; Tom Blake, Greg Ciezarek, Mitesh Patel and Kostas Petridis.

I would also like to thank Darren Burton, Andrew Gilbert, Matthew Kenzie and Indrek Sepp for making the last 3 and a half years so enjoyable, it was an absolute pleasure. Finally, thanks to my family for their endless encouragement and support in getting me to this point in my life.

Contents

1. Introduction	2
2. Theoretical overview	4
2.1. The Standard Model	4
2.2. Beyond the Standard Model	8
2.3. Effective field theory in flavour physics	9
2.4. Electroweak penguin decays	10
2.5. Theoretical uncertainties in exclusive electroweak penguin decays	12
3. Experimental setup	14
3.1. LHC	14
3.2. LHCb	15
3.2.1. Tracking	16
3.2.2. Particle Identification	17
3.2.3. Trigger	20
3.2.4. Simulation	21
3.2.5. Stripping	22
4. Reconstruction efficiency of K_s^0 mesons	24
4.1. Introduction	24
4.2. Selection	26
4.3. Estimating the four-vector of the missing track	28
4.3.1. Solving using K_s^0 mass constraint	28
4.3.2. Kinematic fit	29
4.3.3. Unfolding	31
4.4. Backgrounds	32
4.4.1. Random slow pion background	33
4.4.2. Semi-combinatorial background	34

4.4.3.	$D^0 \rightarrow \phi K^{*0}$	35
4.5.	Mass fits	35
4.5.1.	Signal shape	35
4.5.2.	Partially reconstructed fits	35
4.5.3.	Fully reconstructed fits	36
4.5.4.	Vertex fit	37
4.6.	Results	37
4.6.1.	Tracking efficiency	38
4.6.2.	Vertex efficiency	41
4.7.	Cross-check using $B \rightarrow J/\psi K$ decays	42
4.8.	Summary	43
5.	Observation of a resonance in $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays at low recoil	45
5.1.	Introduction	45
5.2.	Trigger, selection and backgrounds for $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays	48
5.2.1.	Trigger response	48
5.2.2.	Peaking background	49
5.2.3.	Combinatorial background	50
5.2.4.	Mass fit to B candidates	51
5.3.	Dimuon fit	53
5.3.1.	Kinematic constraint	53
5.3.2.	Background shape	54
5.3.3.	Efficiency shape	55
5.3.4.	Signal model	56
5.4.	Results	57
5.5.	Summary	62
6.	Differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays	63
6.1.	Introduction	63
6.2.	Selection	66
6.2.1.	Pre-selection	66
6.2.2.	Multivariate selection	67
6.3.	Backgrounds	68
6.4.	Mass fits	68

6.5. Efficiency as a function of q^2	73
6.6. Systematic uncertainties	74
6.7. Branching fraction results	76
6.8. Isospin asymmetry results	78
6.9. Cross-checks	79
6.9.1. Compatibility with previous result	79
6.9.2. Stability by run period and reconstruction category	80
6.9.3. Stability with mass range	82
6.9.4. Restriction of trigger to muons	82
6.9.5. $B^0 \rightarrow \psi(2S)K_s^0/B \rightarrow J/\psi K_s^0$ branching fraction ratio	82
6.10. Summary	83
7. Angular analysis of charged and neutral $B \rightarrow K\mu^+\mu^-$ decays	84
7.1. Introduction	84
7.2. Efficiency as a function of $\cos\theta_l$	86
7.3. Angular fit	90
7.3.1. Minimisation	90
7.3.2. Coverage	91
7.4. Results	94
7.5. Systematic uncertainties	96
7.5.1. Background parametrisation	96
7.5.2. Negligible systematics	97
7.6. Summary	97
8. Conclusions	99
Bibliography	103
Appendix	109
A. Tabulated branching fractions	110
B. Tabulated isospin asymmetries	113
C. Tabulated A_{FB} and F_{H} results	115
D. Two-dimensional Feldman-Cousins intervals	117

E. License	123
List of Figures	126
List of Tables	134

Chapter 1.

Introduction

This thesis centres around measurements of a particular class of rare B meson decays known as electroweak penguin decays [6], named as such because their dominant Feynman diagram looks vaguely like a penguin and is mediated by the electroweak force. Measurements of these decays test the theoretical framework known as the Standard Model (SM), introduced in chapter 2, along with some specific issues related to predicting the behaviour of these electroweak penguin decays in the SM. These decays are sensitive to physics beyond the SM (new physics) via the influence of heavy particles which “borrow” energy from the vacuum in order to participate in the decay.

The results are based on data collected at the Large Hadron Collider (LHC) by the LHC beauty (LHCb) experiment, both of which are briefly described in chapter 3. In chapter 4, a calibration measurement of the reconstruction efficiency for long lived particles in the LHCb detector is presented. This measurement is vital for the physics analysis described later on.

The electroweak penguin decay measurements are presented in three physics analysis chapters. The first describes the measurement of a resonance found in the dimuon spectrum of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays. This resonance can be seen as a nuisance in the context of searching for new physics, and measurements of its properties are important for maximising the new physics sensitivity in the non-resonant $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay. The first chapter which describes a search for new physics is chapter 6, and details a measurement of the branching fractions of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays, which are sensitive to heavy vector and axial-vector like particles. The difference in the branching fractions of the neutral and charged versions of these decays, the so-called isospin asymmetry, is also measured. Finally, in chapter 7, an angular analysis of $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays is described, where angular observables such as the forward-backward asymmetry of the dimuon system are measured. These angular observables are sensitive to a different set of possible

new physics particles compared to the rate observables. For example, a new physics particle with zero spin can be more easily detected with an angular analysis compared to a rate analysis.

Chapter 2.

Theoretical overview

This chapter describes the theoretical motivation for studying electroweak penguin decays. The successes and short-comings to the Standard Model(SM) are described, followed by the theoretical framework used to predict the behaviour of electroweak penguin decays in the SM.

2.1. The Standard Model

The SM [7–13], describes the current understanding of electroweak and strong forces. The SM Lagrangian can be split up into two pieces as

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{Higgs}}, \quad (2.1)$$

where $\mathcal{L}_{\text{gauge}}$ obeys the local gauge symmetry

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y. \quad (2.2)$$

The group, $SU(3)_C$, is obeyed by strong interactions (QCD) and the $SU(2)_L \otimes U(1)_Y$ group is obeyed by electroweak interactions. The generators of each group, eight for QCD and four for the electroweak force, correspond to gauge bosons which mediate the SM forces.

The gauge sector of the SM has only three free parameters, the strengths of the strong, electromagnetic and weak interactions, which are all of $\mathcal{O}(1)$ in size (at least at high energies). This situation is considered to be *natural*, which is a subjective term often used to describe how

difficult it is to make a theory agree with experimental data. A theory which needs many free parameters, with wildly varying sizes, is considered to be unnatural.

This simple, predictive and highly symmetric theory can be compared to the second part of the SM Lagrangian in equation 2.1, $\mathcal{L}_{\text{Higgs}}$, responsible for the Higgs mechanism [10–13], defined as

$$\mathcal{L}_{\text{Higgs}} = -(D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi) + \mathcal{L}_Y, \quad (2.3)$$

where $-(D_\mu \phi)^\dagger (D^\mu \phi)$ is the kinetic term for a scalar particle, ϕ a doublet of complex scalar fields and

$$V(\phi^\dagger \phi) = \lambda(\phi^\dagger \phi)^2 - \mu^2 \phi^\dagger \phi + \frac{\mu^4}{4\lambda}, \quad (2.4)$$

which, for $\mu^2 > 0$, defines the Higgs potential with a non-zero minimum at

$$|\phi| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{1}{\sqrt{2}}\nu, \quad (2.5)$$

where ν is the vacuum expectation value and has been measured to be $\nu = 246 \text{ GeV}$ [14]. One can choose a *unitary gauge*, which is a particular direction of the minima,

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}, \quad (2.6)$$

which breaks electroweak symmetry and gives mass to the W and Z bosons in a renormalisable way. The Higgs boson, H , arises from fluctuations about this minimum. For flavour physics, it is important to concentrate on the Yukawa term, $\mathcal{L}_Y$ in equation 2.3, as it is responsible for all flavour violation in the SM.

The mass term for a fermion is not invariant under the SM group rotation. This problem can also be bypassed with the Higgs field using Yukawa interactions. For quarks, Yukawa terms appear in the Lagrangian as

$$\mathcal{L}_Y^q = a_{ij} \bar{q}_{Li} \phi^* u_{Rj} + b_{ij} \bar{q}_{Li} \phi d_{Rj} + h.c., \quad (2.7)$$

where, a_{ij} and b_{ij} are the Yukawa coupling strengths between the i and j quark generations, L and R denote the left and right handed chirality components, u and d denote the up and down type quarks and

$$q_{Li} = \begin{pmatrix} u_{Li} \\ d_{Li} \end{pmatrix}. \quad (2.8)$$

In the unitary gauge, the quark Yukawa terms become

$$\mathcal{L}_Y^q = \left(1 + \frac{H}{\nu}\right) (\bar{u}_{Li} m_{ij}^u u_{Rj} + \bar{d}_{Li} m_{ij}^d d_{Rj} + h.c.), \quad (2.9)$$

where $m_{ij}^{u,d}$ are the Yukawa matrices, defined as

$$m_{ij}^u = -\frac{1}{\sqrt{2}} \nu a_{ij}, \quad m_{ij}^d = -\frac{1}{\sqrt{2}} \nu b_{ij}. \quad (2.10)$$

The Higgs interaction terms are those proportional to H , and are discussed later. The Yukawa matrices are diagonalised by the unitary rotation matrices U ,

$$m_\alpha^u = (U_L^{u\dagger})_{i\alpha} m_{ij}^u (U_R^u)_{aj}, \quad m_\alpha^d = (U_L^{d\dagger})_{i\alpha} m_{ij}^d (U_R^d)_{aj}, \quad (2.11)$$

where $m_\alpha^{u,d}$ are the mass matrices for the quarks. The Lagrangian for weak interactions between quarks can be written as

$$\mathcal{L}_{NC} = ig_2[\bar{u}_{Lj}Z_\mu\gamma^u u_{Lj}] \quad \text{and} \quad (2.12)$$

$$\mathcal{L}_{CC} = \frac{ig_2}{\sqrt{2}}[W_\mu^+\bar{u}_{Lj}\gamma^u d_{Lj} + W_\mu^-\bar{d}_{Lj}\gamma^\mu u_{Lj}], \quad (2.13)$$

where Z_μ and W_μ^+ are the neutral and charged currents and g_2 is the weak coupling strength. Rotating this into the mass basis yields

$$\mathcal{L}_{NC} = ig_2[\bar{u}_{L\alpha}(U_L^u)_{\alpha j}(U_L^{u\dagger})_{j\beta}Z_\mu\gamma^u u_{L\beta}] = ig_2[\bar{u}_{L\alpha}\delta_{\alpha\beta}]Z_\mu\gamma^u u_{L\beta} \quad \text{and} \quad (2.14)$$

$$\mathcal{L}_{CC} = \frac{ig_2}{\sqrt{2}}[W_\mu^+\bar{u}_{L\alpha}(U_L^u)_{\alpha j}(U_L^{d\dagger})_{j\beta}\gamma^u d_{L\beta} + W_\mu^-\bar{d}_{Lj}(U_L^d)_{\alpha j}(U_L^{u\dagger})_{j\beta}\gamma^\mu u_{L\beta}]. \quad (2.15)$$

Due to the unitarity of the rotation matrices, there are no off diagonal neutral current interaction terms and therefore no tree-level flavour changing neutral currents (FCNC). For the charged current interactions however, there are off-diagonal terms which contribute to quark flavour mixing between generations α and β with strength

$$(U_L^u)_{\alpha j}(U_L^{u\dagger})_{j\beta} = V_{\alpha\beta} \quad (2.16)$$

where $V_{\alpha\beta}$ is the Cabbibo-Kobayoshi-Maskawa (CKM) matrix [15, 16], shown in the Wolfenstein parameterisation [17],

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 & \lambda & 0 \\ -\lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\lambda^2). \quad (2.17)$$

and describes all flavour violation in the SM. The CKM matrix has a very distinctive hierarchal structure, where the parameter λ is the sine of the Cabbibo angle has a value of about 0.2. The CKM matrix is also the only source of CP violation in the SM, which is difference in a physical system after the of the interchange of particle-antiparticle pairs (charge conjugation) and a reversal in the sign of all spatial coordinates (parity transformation).

2.2. Beyond the Standard Model

Despite the unprecedented success of the SM, it cannot explain all experimental observations:

- It only incorporates three fundamental forces, there is no explanation for gravity.
- The amount of CP violation arising from the CKM matrix is roughly ten orders of magnitude lower than what is needed to produce the matter-antimatter asymmetry of the universe after the Big Bang [18].
- Only 4.9% of the observed matter in the universe is accounted for by the SM [19].
- Neutrinos are massless in the SM, which is in contradiction to observation of neutrino oscillations [20].

These observations motivate physics beyond the SM, but at which energy scale? In addition to the experimental side, there are a few theoretical issues with the SM:

- The SM has a large number (18) of free parameters with a large range. Why is the top quark so much heavier than the up quark?
- What causes the hierarchical structure of the CKM matrix? Why are there three flavour families?
- In principle QCD should also violate the C and P symmetries, similarly to the weak force. The fact that it does not requires fine tuning at the level of 10^{-9} [21].
- Quantum corrections push the Higgs mass to a scale much larger than electroweak scale [22–25], which means there must be a very precise cancellation between the bare mass and these corrections to produce the Higgs mass that is seen in data. This issue is often used as a motivation for new physics at the TeV scale to provide a natural cancellation of these corrections.

A common theme to all these theoretical issues is that they can all be fixed with a fine-tuning of parameters. One cannot discount the possibility that the SM is effectively valid up to the Planck scale. An interesting observation is that most of the “theoretical ugliness” (e.g. 13 out of the 18 free parameters) in the SM arise in the Yukawa sector of the SM.

Decays which are forbidden at tree level are powerful probes of physics beyond the SM. An example of this is kaon mixing, which is a FCNC and so must proceed, in the SM, through a box

diagram as shown in Fig. 2.1. The kaon mixing rate is highly sensitive to physics which is over two orders of magnitude higher in energy than the kaon mass, namely the W mass and couplings. This mechanism is similar in models beyond the SM, where even heavier particles could allow additional diagrams which would alter the mixing rate. Kaon mixing is particularly sensitive as the CKM suppression is of order λ^4 , which is not necessarily the case in new physics models. If a new physics has naive $\mathcal{O}(1)$ flavour changing couplings then its scale must be over 100 TeV in order to satisfy bounds imposed by kaon mixing and CP measurements [26]. This statement is in contradiction with the argument that new physics should be at the TeV scale to avoid fine tuning of the Higgs mass. and is known as the *flavour problem*. It is bypassed by assuming that the Yukawa matrices are the only sources of flavour violation even beyond the SM, which implies that new physics also has a CKM like structure and is known the *minimal flavour violation* hypothesis. Under this hypothesis, new physics can satisfy flavour constraints whilst providing a natural cancellation of the quantum corrections to the Higgs mass. It also cancels the advantage kaon physics has over B physics due to the larger CKM suppression in kaon FCNC decays.

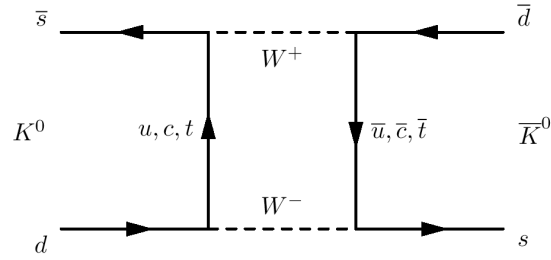


Figure 2.1: Box diagram which contributes to neutral kaon mixing in the SM.

2.3. Effective field theory in flavour physics



Figure 2.2: Diagram contributing to B mixing in the full (left) and effective (right) theories.

This section describes the theoretical framework in which flavour physics measurements are interpreted, which is known as the Operator Product Expansion (OPE) [27]. The OPE method

works by integrating out all degrees of freedom above some energy scale, Λ , which in B physics is everything at or above the electroweak scale. This approach is valid, if Λ is much larger than the relevant energy scale of the decay, which for B physics is about $5 \text{ GeV}/c^2$. In this case, the heavy degrees of freedom decouple from the rest of the decay dynamics [28]. An example of this is shown for B mixing in Fig. 2.2, where the box diagram, mediated by the heavy top quark and W boson, is collapsed into a four point interaction. This is similar to Fermi's effective theory of weak decays used before electroweak theory was devised. In practice this means one can write down the matrix element for a high energy process of an effective hamiltonian

$$\langle f | H_{\text{eff}} | i \rangle = \sum_k \frac{1}{\Lambda^k} \sum_i \mathcal{C}_{k,i} \langle f | \mathcal{O}_{k,i} | i \rangle \Big|_{\Lambda}, \quad (2.18)$$

where $\mathcal{C}_{k,i}$ are simply complex numbers which encapsulate the high energy, short distance contributions and are known as Wilson coefficients. There is a different Wilson coefficient for each operator $\mathcal{O}_{k,i}$, corresponding to a particular gauge structure. An advantage of the OPE is that by construction Wilson coefficients are independent of the underlying process. This allows one to constrain all new physics models with a particular gauge structure. It also provides a framework to combine different experimental results which have complementary sensitivity to new physics.

2.4. Electroweak penguin decays

The particular FCNC decays analysed in this thesis are $b \rightarrow s\mu^+\mu^-$ electroweak penguin transitions. At the lowest order in the SM, they are mediated by penguin and box diagrams, shown for the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$ in Fig. 2.3. The largest contribution to the SM decay rate arises from the Wilson coefficients $\mathcal{C}_{7,9,10}$ and correspond to the operators below

$$\begin{aligned} \mathcal{O}_7 &= \frac{e}{g^2} \bar{m}_b (\bar{s} \sigma_{\mu\nu} P_R b) F^{\mu\nu}, & \mathcal{O}'_7 &= \frac{e}{g^2} \bar{m}_b (\bar{s} \sigma_{\mu\nu} P_L b) F^{\mu\nu}, \\ \mathcal{O}_9 &= \frac{e^2}{g^2} (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \ell), & \mathcal{O}'_9 &= \frac{e^2}{g^2} (\bar{s} \gamma_\mu P_R b) (\bar{\ell} \gamma^\mu \ell), \\ \mathcal{O}_{10} &= \frac{e^2}{g^2} (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \gamma_5 \ell), & \mathcal{O}'_{10} &= \frac{e^2}{g^2} (\bar{s} \gamma_\mu P_R b) (\bar{\ell} \gamma^\mu \gamma_5 \ell), \end{aligned} \quad (2.19)$$

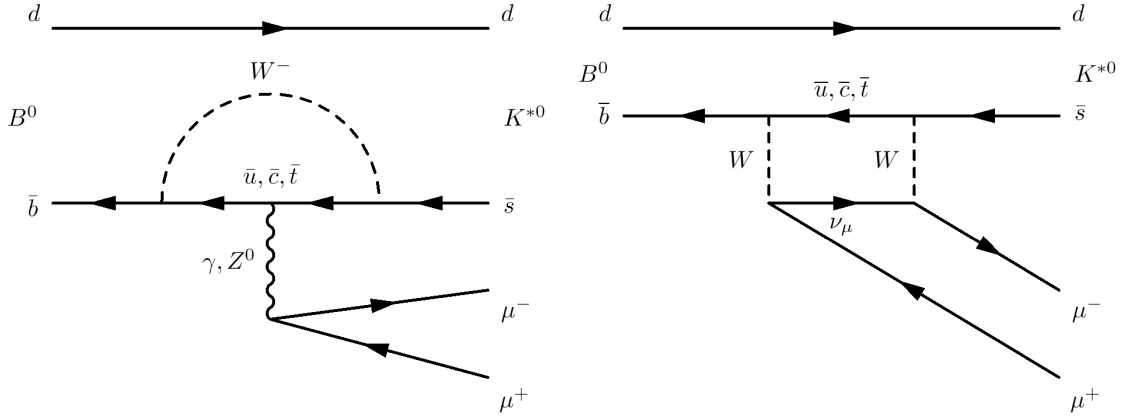


Figure 2.3: The $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay in the lowest order SM diagrams. The left is the penguin diagram, which contributes to the Wilson coefficients $C_{7,9,10}$. The right is the box diagram, which contributes to $C_{9,10}$.

where e and g are the coupling strengths of the electromagnetic and weak forces, $\sigma_{\mu\nu}$ are the Pauli spin matrices, $F^{\mu\nu}$ is the electromagnetic field strength tensor and $P_{L,R}$ are the left- and right-handed projection operators. The operator $\mathcal{O}_7$ is the electromagnetic operator, corresponding to the emission of a photon from the loop. The operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$ are the semi-leptonic vector and axial-vector operators and correspond to the Z penguin and W box diagrams. The primed operators are those with opposite chirality whose Wilson coefficients are suppressed by the factor m_s/m_b in the SM, relative to the unprimed ones.

Another set of operators can be defined as

$$\begin{aligned}
 \mathcal{O}_1^q &= (\bar{s}_i q_j)_{V-A} (\bar{q}_j b_i)_{V-A} & \mathcal{O}_2^q &= (\bar{s}_i q_i)_{V-A} (\bar{q}_j b_j)_{V-A} \\
 \mathcal{O}_3 &= (\bar{s}_i b_i)_{V-A} \sum_q (\bar{q}_j q_j)_{V-A} & \mathcal{O}_4 &= (\bar{s}_i b_j)_{V-A} \sum_q (\bar{q}_j q_i)_{V-A} \\
 \mathcal{O}_5 &= (\bar{s}_i b_i)_{V-A} \sum_q (\bar{q}_j q_j)_{V+A} & \mathcal{O}_6 &= (\bar{s}_i b_j)_{V-A} \sum_q (\bar{q}_j q_i)_{V+A} \\
 \mathcal{O}_8 &= -\frac{g_s m_b}{8\pi^2} \bar{s} \sigma \cdot G (1 + \gamma_5) b & &
 \end{aligned} \tag{2.20}$$

which provide a small contribution to the $b \rightarrow s \mu^+ \mu^-$ decay rate but are still important for this thesis. For example $\mathcal{O}_{1-6}$ are used to predict long distance contributions such as $c\bar{c}$ loops, which is the theme of chapter 5. These operators are the dominant contribution to the isospin asymmetry between $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays, discussed in chapter 6.

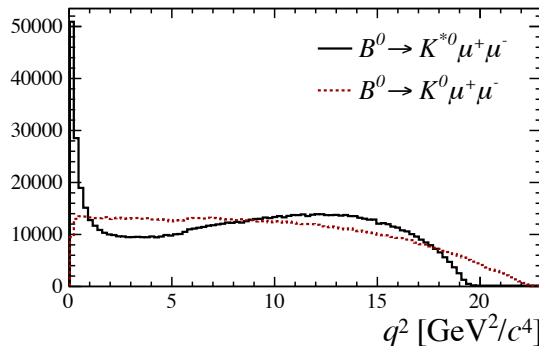


Figure 2.4: Distribution of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^0 \rightarrow K^0 \mu^+ \mu^-$ decays using the model described in Ref. [29]. There is a photon pole as $q^2 \rightarrow 0$ for the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay.

The relative contribution of each Wilson coefficient to the total decay rate varies depending on the four-momentum transferred to the $\mu^+ \mu^-$ pair, q^2 . For example, it is more likely to radiate a virtual photon at low q^2 and so the contribution from $\mathcal{C}_7$ is increased in that kinematic region. The contribution is also different depending on whether the s hadronised into a ground state (e.g. $B \rightarrow K \mu^+ \mu^-$) or excited (e.g. $B \rightarrow K^* \mu^+ \mu^-$) kaon. For the $B \rightarrow K^* \mu^+ \mu^-$ case, the $\mu^+ \mu^-$ pair can be transversely polarised and so there is a photon pole at low q^2 , shown in Fig. 2.4, where the rate of the $B \rightarrow K^* \mu^+ \mu^-$ decay rises rapidly towards that of $B \rightarrow K^* \gamma$ as $q^2 \rightarrow 0$. For $B \rightarrow K \mu^+ \mu^-$ the $\mu^+ \mu^-$ pair is fully longitudinally polarised so the contribution from a real photon is forbidden and no photon pole is seen.

2.5. Theoretical uncertainties in exclusive electroweak penguin decays

Measurements of $b \rightarrow s \mu^+ \mu^-$ decays are only useful in terms of beyond the SM physics if the long distance contributions can be controlled¹. The easiest way to achieve this is to design observables in which the long distance dependence cancels in ratios whilst retaining sensitivity to the Wilson

¹This section has been written with reference to Ref. [30].

coefficients. The isospin asymmetries and angular observables described in this thesis come under this category, in which the theoretical uncertainty is negligible compared to experimental precision and will not be discussed further.

For branching fractions, the situation is the opposite, where the theoretical uncertainty is larger than the experimental uncertainty. For the partonic part of the decay, the b quark is heavy enough to perform perturbative QCD calculations, i.e. $\alpha_s(5 \text{ GeV}) < 1$. Theoretical calculations are simplified by constructing a heavy quark effective field theory (HQET), in which the interactions of the heavy quark are soft compared to the large mass of the b quark, and the partonic process can be expanded in terms of Λ_{QCD}/m_B , where Λ_{QCD} is the QCD scale. An important assumption, known as QCD factorisation (QCDF) is that the partonic process is separate from the hadronisation of the $s\bar{q}$ pair. The latter is encoded in hadronic form-factors and is the dominant source of theoretical uncertainty for branching fraction predictions.

In the low q^2 region, HQET breaks down as the hadron has an energy which is comparable to the mass of the b quark. In this case an additional expansion can be made around Λ_{QCD}/E , which exploits the fact that interactions are soft compared to the hadron energy. This, with the addition of effects from collinear gluons forms soft collinear effective theory (SCET) [31]. SCET takes into account the hard interactions of scale m_B and soft interactions at the scale $\sqrt{m_B \Lambda_{\text{QCD}}}$ separately. Numerical values from these hard and soft interactions are calculated using light cone sum rules (LCSR) [32], which exploits the fact that the hadron is highly relativistic, as the name suggests. This is the technique used to determine signal model form factors in Fig. 2.4.

At high q^2 , another expansion is employed around $\Lambda_{\text{QCD}}/\sqrt{q^2}$, which is only valid above the open charm threshold [33]. This incorporates the contribution from $c\bar{c}$ resonances by treating it as a local interaction in the operators $\mathcal{O}_{1-6}$. In addition to extrapolating LCSR results to high q^2 , the form factors at high q^2 can be calculated using lattice QCD (LQCD) [14]. LQCD is a technique for performing non-perturbative QCD calculations by discretising space-time with a lattice spacing, a . LQCD is best used for light quarks with low energy so that partons have low momenta compared to the lattice cut-off of $1/a$. For the lattice QCD predictions used in this thesis, $1/a$ is about 2 GeV [34, 35], which restricts the lattice data points to the very high q^2 region. In this region, the form factor uncertainties calculated using LQCD are roughly three times more precise than results obtained from LCSR.

Chapter 3.

Experimental setup

In this section, a brief overview of how the data are obtained is presented. The Large Hadron Collider (LHC) accelerator and the LHC beauty (LHCb) detector are introduced, with special attention paid to aspects which are relevant to the physics described in this thesis.

3.1. LHC

The LHC [36] is the worlds most energetic particle accelerator, with a current record centre-of-mass energy of 8 TeV. It is situated inside the 27km tunnel originally used for the Large Electron Positron collider (LEP). The complex is based at the international research facility, CERN, near Geneva, Switzerland.

During a typical fill in 2011 (2012), the LHC accelerates bunches of protons to a collision energy of 7 (8) TeV from the injection energy of 450 GeV and collides them every 50 ns. The peak luminosity of the LHC was achieved in 2012 and was $7 \times 10^{33} \text{ cm}^2\text{s}^{-1}$, which is over an order magnitude above the LHCb design luminosity. To keep occupancy and pileup low at LHCb, the beams at the LHCb interaction point were offset by an amount to keep the luminosity constant during a fill. This technique, known as *luminosity levelling*, kept the LHCb luminosity at roughly $3(4) \times 10^{32} \text{ cm}^2\text{s}^{-1}$ in 2011 (2012) and maximised the data available for LHCb analyses.

The production cross section for $b\bar{b}$ pairs at the 2011 LHC centre-of mass energy of 7 TeV has been measured to be $(284 \pm 29 \pm 40) \mu\text{b}$ [37], which is over five orders of magnitude greater than in e^+e^- collisions at the $\Upsilon(4S)$ resonance, which was the production mechanism for B mesons at the BaBar [38] and Belle [39] detectors (both known as B -factories). This high cross section offsets

the difficulty of doing flavour physics at a hadron collider compared to an e^+e^- machine, where is no beam energy constraint as B mesons are only produced with a fraction of the pp energy via the strong interaction. During run 1 of the LHC during 2010-2012, LHCb collected around 3 fb^{-1} of integrated luminosity, corresponding to over 800 billion $b\bar{b}$ events.

3.2. LHCb

The LHCb detector [40], shown in Fig. 3.1, is specifically designed to study heavy flavour physics, which is reflected in excellent tracking and particle identification (PID) capabilities. Due to the kinematics of high energy proton-proton collisions, a large fraction of b quarks will decay along the beam axis, resulting in a large fraction of the physics being contained within a forward cone of the beam. The LHCb design exploits this phenomenon by only covering an angular acceptance of 15 to 300 (250) mrad in the horizontal (vertical) direction. This corresponds to roughly 2 – 5 as a pseudo-rapidity range, where pseudo-rapidity is defined as

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right), \quad (3.1)$$

where θ is the angle between the particle momentum and the beam axis. About a sixth of the signal decays described in this thesis are inside the LHCb acceptance and fully reconstructible.

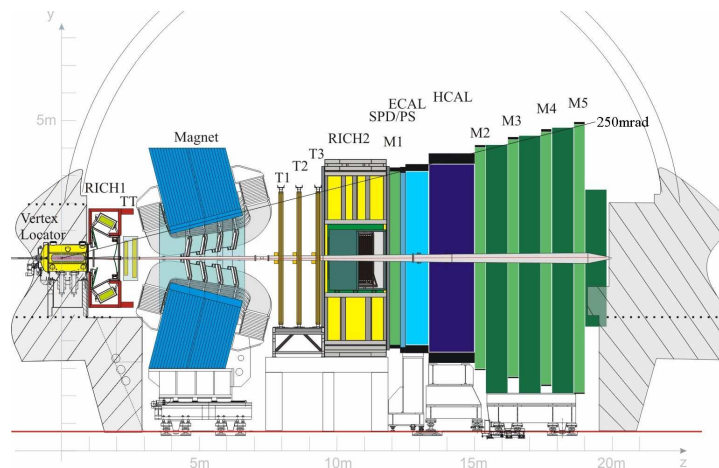


Figure 3.1: Schematic diagram of the LHCb detector.

3.2.1. Tracking

Closest to the interaction point, the VERtEX LOcator (VELO) provides precise spatial measurements in order to identify displaced vertices, a distinctive characteristic of heavy flavour decays. For example, a B^+ meson with 100 GeV/ c momentum will travel about 1 cm before decaying. The VELO consists of two halves, each with several silicon microstrip detectors. Around half of the sensors are arranged in concentric circles around the beam axis to provide r information (r sensor) and half emanate radially from the beam axis to provide ϕ information (ϕ sensor). There are around 40 r and ϕ sensors each, with a pitch of approximately 40–100 μm , depending on r . This leads to a position resolution of approximately 10-20 μm . An important feature of the VELO is its ability to retract during injection, which allows the VELO to be placed only 8 mm from the beam line and reduces the extrapolation to the primary proton-proton interaction vertex (PV). The VELO is a crucial component for the analyses in this thesis, as it allows efficient reduction of the large background coming from combinations of tracks originating from the PV.

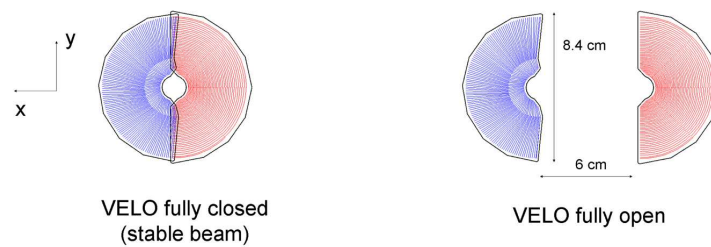


Figure 3.2: Schematic of one the VELO r - ϕ modules, showing its retractable nature.

The LHCb tracking system layout is similar to the design of a fixed target experiment, where all the tracking stations apart from the Tracking Turicensis (TT) are located downstream of a 4 Tm dipole magnet. Further downstream, the tracker is separated into the Inner Tracker (IT) and Outer Tracker (OT) due to a difference in occupancy, where the density of particles is increased closer to the beam line. The IT and TT are silicon micro-strip detectors, whereas the OT is a straw tube detector.

There are two types of tracks used in the analysis in this thesis, which are illustrated in Fig. 3.3. The most common type is a long track, where enough hits are recorded in the VELO to reconstruct a track. If a particle fails to leave enough hits in the VELO, but does leave hits in the TT, it is classed as a downstream track. For long tracks, the combined LHCb tracking system

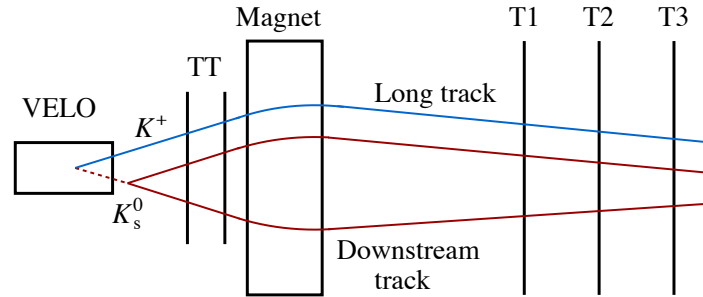


Figure 3.3: Schematic of the LHCb tracking system, showing the different track types present in the LHCb reconstruction. Analyses in this thesis use the long and downstream track types.

achieves momentum resolution, $\frac{\Delta p}{p}$, between 0.4 and 0.6% for momenta 5 GeV/ c and 100 GeV/ c , respectively [40]. This is essential to achieve a good resolution of the B meson invariant mass.

3.2.2. Particle Identification

Directly behind the VELO is a Ring Imaging CHerenkov (RICH) detector, which provides kaon/pion separation at low momentum. Another RICH detector is located upstream of the tracking system, providing separation for high momentum particles. Particles traversing the RICH interacting medium, known as the radiator, will emit a ring of Cherenkov light if the velocity of the particle is greater than the speed of light in the medium. For a given momentum, the angle of the Cherenkov light depends on the speed of the particle transversing the medium and so by combining this with the momentum measurement made in the tracker, the mass and thereby the identity can be found.

The choice of radiator is an important one as its refractive index determines the momentum range at which the PID can be determined. This momentum range is limited by the fact that lower momentum particles do not travel faster than the speed of light in the medium of a low refractive index. It is limited at the high momentum region due to the fact that there is a saturation point, where all particles regardless of mass will emit at the same Cherenkov angle and the distinguishing power is lost. These features are shown in Fig. 3.4. The RICH detectors in LHCb have three radiators; aerogel, CF_4 gas and C_4F_{10} gas which have refractive indices 1.03,

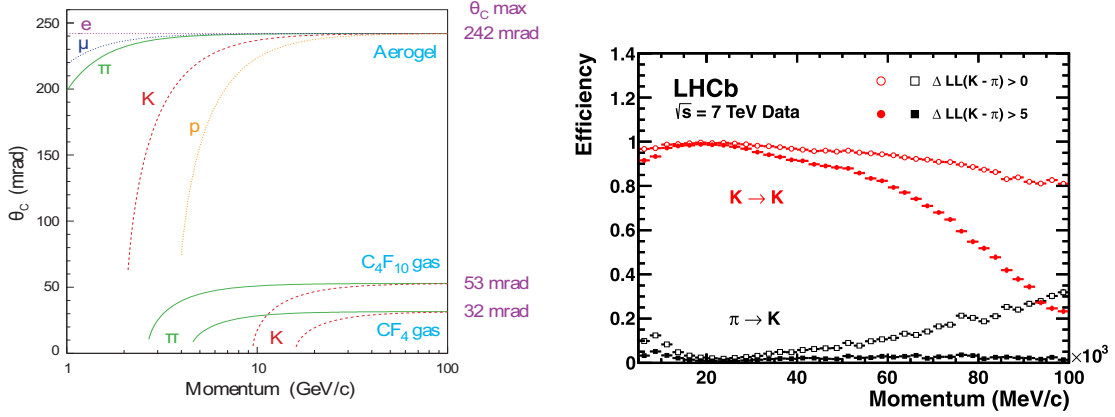


Figure 3.4: Cherenkov angle of different particle species for the three RICH radiators (left) and performance (right) as a function of momentum, taken from Ref. [41].

1.0014 and 1.0005, respectively. A combination of measurements in all three radiators allows a PID measurement to be made across a large momentum range (5-100 GeV) [41].

The RICH detectors have a system of mirrors which focus the photons emitted from a traversing particle onto a plane of Hybrid Photon Detectors (HPDs). An iterative fit then is made to the photon distribution, where a different mass hypothesis of each track is changed to maximise the likelihood the observed HPD hit distribution. There are two main ways in which this information is used in analysis at LHCb. The difference in log-likelihood between a particle and the pion hypothesis (e.g. $DLL_{K\pi}$), can be computed. This is the usual way of discriminating between particles and the behaviour of which is well understood in LHCb. Another way is to combine the RICH information in a neural network, trained on simulation, which provides good rejection against combinatorial background, and similarly to the likelihoods, is verified with calibration samples. The kaon PID performance can be measured using these calibration samples, shown in Fig. 3.4.

The electromagnetic and hadronic calorimeters (ECAL, HCAL) are placed downstream of the tracking system, and are used to measure the energy of neutral particles. The ECAL is a scintillator/lead sampling calorimeter and has an stochastic energy resolution of approximately $1\% + 10\%/\sqrt{E}$. The HCAL is a scintillator/iron sampling calorimeter and has a stochastic energy resolution of approximately $70\%/\sqrt{E}$. The calorimeters are not essential components for analysing the decays described in this thesis, apart from providing some muon PID information.

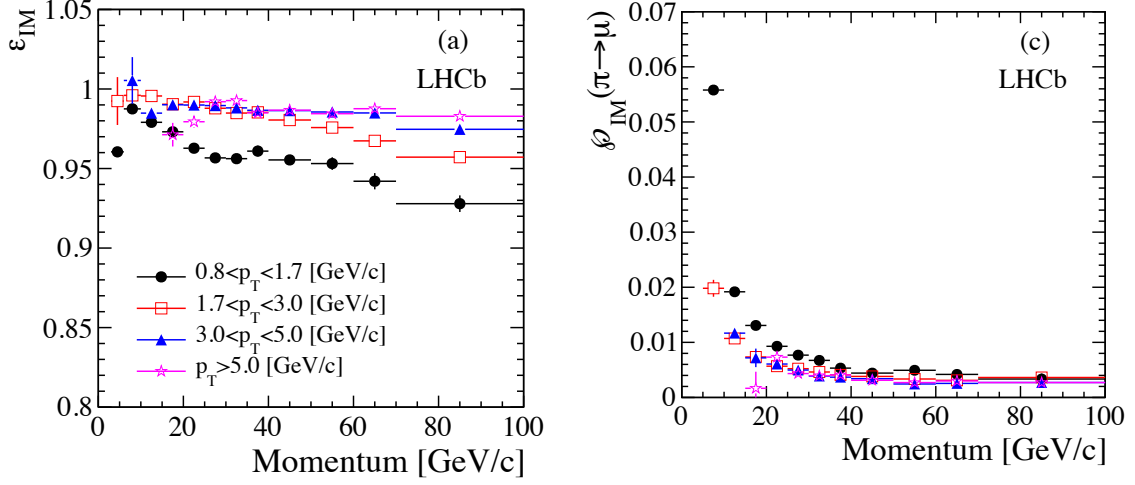


Figure 3.5: Muon identification efficiency for muons (left) and pions (right), taken from Ref. [42].

All signal decays described in this thesis are rare, and contain a pair of muons in their decay products. Thus, the very good muon identification obtained from the LHCb muon system is essential. There are four muon stations downstream of the calorimeters and one station located upstream. The stations are based on multi-wire proportional counters, with the exception of the inner region of the upstream station, which utilises triple gas electron multiplier detectors due to the high radiation level in this region [40]. A particle is identified as a muon (ISMUON), if it contains a hits in muon stations in a field of interest extrapolated from the tracker. The number of stations require to contain hits is dependent on the momentum of the particle. For example, particles with $p > 10 \text{ GeV}/c$ must have hits in four stations. Low momentum muons are likely to be absorbed by first stations and so the requirement is softer for $p < 10 \text{ GeV}/c$. The efficiency of this flag on muons and pions is shown in Fig. 3.5, which demonstrates the good muon PID performance of LHCb. The muon ID efficiency gets slightly worse at low p_T , as shown by the black curve in Fig. 3.5, due to particles falling out the acceptance. The pion mis-id is dominated by decays in flight, which is why the mis-id probability gets significantly better at higher momentum. In addition to the ISMUON flag, information from RICH and calorimeters are combined to construct a difference in log-likelihood for muons with respect to pions ($\text{DLL}_{\mu\pi}$), similar to RICH PID variables. The addition of this likelihood information typically improves the hadron to muon mis-identification by a factor three.

3.2.3. Trigger

As with any modern hadron collider experiment, the raw data rate at LHCb is too large to be archived, and must be reduced by a series of hardware and software stages known as triggers. At a typical luminosity of $4 \times 10^{32} \text{cm}^2 \text{s}^{-1}$, the production rate in the LHCb acceptance of b -hadron and c -hadrons is about 30 kHz and 500 kHz, respectively. These rates are so large that the trigger is often saturated by signal for common decay modes.

LHCb employs a two level trigger system [43]. Firstly the Level 0 (L0) reduces the data rate from 40 MHz to 1 MHz and secondly the High Level Trigger (HLT) reduces the data rate further down to the required 3-5 kHz ready to write to disk. L0 is a hardware trigger based on field programmable gate arrays and only uses information from the calorimeters and the muon system. The criteria necessary for keeping an event is a high transverse energy particle in the calorimeter or a reconstructed muon with high transverse momentum. In this thesis, transverse momentum, p_T , is defined as the momentum component perpendicular to the beam axis. L0 uses techniques such as pipelining and parallelism to increase the amount of time allowed to make a decision. For the analyses described in this thesis, at L0, one muon is required to have p_T greater than 1.48 (1.76) GeV/ c in 2011 (2012) and has a rate of around 400 kHz. The p_T requirement was tightened in 2012 due to the higher instantaneous luminosity in this period.

The HLT is fully implemented in software and runs on ~ 29000 logical CPU cores. It uses the entire event data to make a decision and is subdivided into two stages, HLT1 and HLT2. Compared to L0, HLT1 includes VELO information which allows selection on the impact parameter (IP). For the decay modes relevant to this thesis, at the HLT1 level one track is required to have p_T greater than 1.0 GeV/ c and IP greater than 100 μm . The HLT1 reduces the 1 MHz rate to about 30 kHz.

The HLT2 is a set of trigger lines, each of which read in the remaining part of the event and perform full event reconstruction. The most important trigger line for the signal decays described in this thesis is one which looks for a good quality, displaced dimuon vertex with $p_T > 1.5 \text{ GeV}/c$ and $m_{\mu^+\mu^-} > 1 \text{ GeV}/c^2$. This is most efficient HLT2 line for q^2 above $\sim 2 \text{ GeV}/c^2$. Another important line is an inclusive B trigger line, the so-called *topological trigger*, which relies on multivariate techniques to maximise signal efficiency [44]. The inputs to this multivariate classifier are properties of 2,3 or 4 track combinations, such as the corrected mass, defined as

$$m_{\text{corr}} = \sqrt{m^2 + |p_{T\text{miss}}|^2} + |p_{T\text{miss}}| \quad (3.2)$$

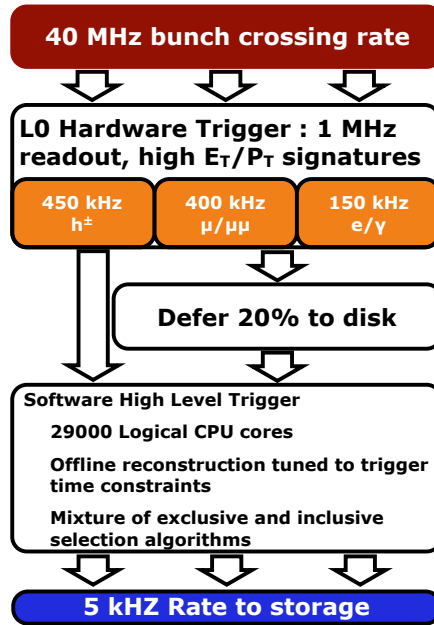


Figure 3.6: Diagram of the trigger logic for typical 2012 running conditions.

where $p_{T_{miss}}$ is the missing momentum perpendicular to the flight direction. Of all the HLT2 lines, the topological trigger tends to have the least bias to physics observables as often only a partial reconstruction is performed and the flexibility of the multivariate selection avoids tight kinematic selection criteria. Other trigger lines relevant for this thesis include one based on single high p_T muon, and an inclusive $D^{*+} \rightarrow D^0 \pi^+$ trigger, which is discussed in more detail in chapter 4.

The output rate of HLT2 was about 3.5 kHz in 2011 and 5 kHz in 2012. To achieve the 5 kHz in 2012, roughly 20% of the rate is deferred to disks situated at the farm, which are then processed between fills. The trigger logic is summarised in Fig. 3.6, shown for typical 2012 running conditions.

3.2.4. Simulation

The LHCb simulation is produced in three stages, generation, simulation and reconstruction. The generation stage is handled using the package EVTGEN [45], where decay kinematics are

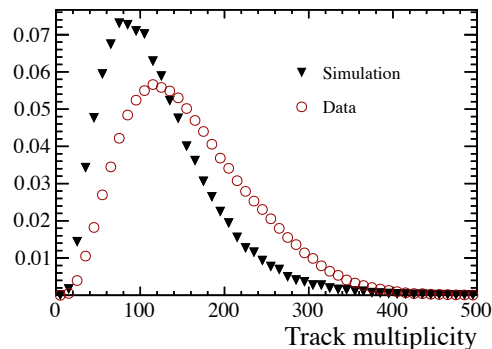


Figure 3.7: Distribution of the number of tracks in the event for simulation and data.

implemented. For example, the signal decays for this thesis use a physics model based on Ref. [29]. The rest of the pp collision is simulated using PYTHIA [46], with a specific configuration for LHCb [47]. Some discrepancies are noticeable between the data and simulation in this area. For example, the average number of tracks (track multiplicity) is lower in the simulation than in data as shown in Fig. 3.7. The impact on analyses is relatively minimal, apart from contributing to a discrepancy in the PID performance between data and simulation. The response of the LHCb detector is simulated using GEANT4 [48] as described in Ref. [49], where accurate modelling of the detector material is important to sufficiently describe the resolution performance of the detector. An example of how this can affect analyses is a discrepancy in the IP resolution of about 20%, which is corrected offline by artificially degrading it in the simulation. Recently, an updated description of the VELO was introduced which significantly reduced this discrepancy. Finally, the same reconstruction software as used for the data is applied to the simulation. None of the analyses described in this thesis are sensitive to which reconstruction version was used.

3.2.5. Stripping

The large trigger output rate of 5 kHz means that there is a huge number of events recorded. The entire dataset cannot be continuously made available to the collaboration due to finite computing resources. Instead, additional selection, known as the stripping, is applied to the full dataset twice a year which reduces the data sample by roughly an order of magnitude. This stripped dataset is the only data readily available and means that a lot of forward planning is required for any analysis on LHCb. The stripping selection is designed by analysts and for the decays in this thesis,

it requires a good quality, displaced vertex, which points back to PV. In general the stripping efficiency is very high ($> 90\%$) for signal candidates which fired the trigger.

Chapter 4.

Reconstruction efficiency of K_S^0 mesons

In this chapter, a data-driven method for estimating the reconstruction efficiency of K_S^0 mesons is presented. All of the analysis was performed by the author.

4.1. Introduction

In many LHCb analyses there is a systematic uncertainty associated with the tracking efficiency of particles. This is particularly important for analyses which rely on absolute normalisation, such as cross-section measurements. A large amount of work has been performed to measure the tracking efficiency of particles reconstructed with tracks in the VELO and T stations (long tracks) in the data. This is performed using J/ψ tag-and-probe methods which rely on partially reconstructing the probe track [50]. For example, a track reconstructed using only the TT can be used to form the J/ψ invariant mass and the long track efficiency is the ratio of signal yields with and without matching the TT track (tag) to a VELO track and T station track further downstream. The data/simulation agreement for 2011 data is shown as a function of pseudorapidity and momentum in Fig. 4.1. The good agreement is much better than the sensitivity for any of the analyses in this thesis and associated systematic uncertainties are negligible.

Unfortunately, these results can only be applied to tracks which originate close to the PV. Roughly 10% of the LHCb physics program involves long lived particles such as the K_S^0 or Λ^0 , where the daughters are reconstructed as two long tracks (LL category) or two downstream tracks (DD category). The effects of possible mis-modelling of reconstruction efficiency for long lived particles in the simulation can be reduced by using a sensible normalisation channel. However,

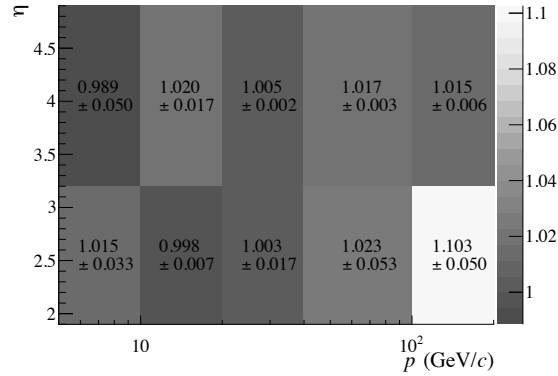


Figure 4.1: Tracking efficiency ratio between data and simulation for long tracks using the J/ψ tag-and-probe methods described in the text. The tracking efficiency ratio is consistent with unity in all bins.

it is clear that a measurement of the long lived reconstruction efficiency is important to reduce systematic uncertainties to the same level as short lived particles.

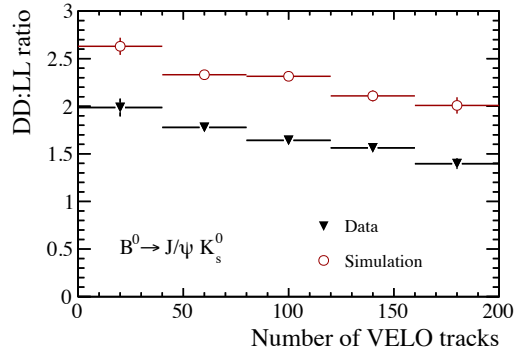


Figure 4.2: Ratio of yields between the DD and LL K_s^0 reconstruction categories for $B^0 \rightarrow J/\psi K_s^0$ decays.

The size of the disagreement for the K_s^0 reconstruction between data and simulation can be estimated by comparing the signal yield of $B^0 \rightarrow J/\psi K_s^0$ in the LL and DD reconstruction categories. The DD:LL yield ratio for data and simulation as a function of the number of VELO tracks in the event is shown in Fig. 4.2. The dependence on multiplicity is attributed to the probability to incorrectly upgrade one of the K_s^0 daughter tracks to a long track by matching it to a VELO track in the same event. This probability becomes larger in busier events. There is also a systematic

shift between the simulation and data, even within a bin of multiplicity, which suggests that the reconstruction efficiency of DD K_s^0 mesons lower in data than in simulation.

The goal of this analysis is to measure the reconstruction efficiency of K_s^0 using data and understand the systematic shift seen in Fig. 4.2 between data and simulation. The strategy is similar to that used in Ref. [51], where the yields of partially and fully reconstructed D decays are compared to measure the pion detection asymmetry. In this case, $D^{*+} \rightarrow (D^0 \rightarrow \phi K_s^0) \pi_s^+$ decays are used, where the K_s^0 has only one downstream track reconstructed. The pion originating from the D^{*+} vertex, called the *slow pion* is labelled π_s^+ . These partially reconstructed candidates are compared with the fully reconstructed decay to obtain the downstream tracking efficiency. The four vector of the missing downstream track is estimated using kinematic constraints and the difference in D^{*+} and D^0 masses is used to fit the signal. Finally, the efficiency to vertex two downstream tracks and apply the standard LHCb K_s^0 selection criteria is measured as a function of K_s^0 momentum.

4.2. Selection

Partially reconstructed decays are difficult to find, as they do not point back to the PV which is normally one of the most discriminating selection criteria. This affects the trigger, where only one set of HLT2 trigger lines, designed to inclusively select $D^{*+} \rightarrow D^0 \pi^+$ decays, is efficient on signal. Each $\phi \pi_s^+$ combination is required to be triggered by these inclusive D^{*+} lines, which leaves the trigger blind to the K_s^0 candidate. Unfortunately, this restricts the analysis to the latter 1.5 fb^{-1} of 2012 running, as the inclusive D^{*+} triggers only ran during that period. The kinematic selection criteria for the trigger are shown in Table 4.1. Note the very tight p_T cuts, which are necessary to reduce the rate of the line to the acceptable level of 500 Hz.

Candidate	Selection
D^{*+}	$p_T > 3750 \text{ MeV}/c$
π_s^+	$p_T > 300 \text{ MeV}/c$
π_s^+	$p > 3000 \text{ MeV}/c$
dihadron + π_s^+	$m_{hh\pi_s} - m_{hh} < 285 \text{ MeV}/c^2$

Table 4.1: Kinematic selection criteria for the inclusive D^{*+} trigger lines, which were used for the K_s^0 reconstruction analysis.

The stripping line used for this analysis was designed to select all $D^{*+} \rightarrow (D^0 \rightarrow hhX)\pi_s^+$ combinations, which includes over 30% of all D^0 decays. Stripping such an inclusive decay places a great strain on the retention, which is reflected in the stripping selection, shown for the $D^0 \rightarrow K^+K^-X$ case in Table 4.2. In particular, note the prescale of 0.5 and tight PID requirements on the kaons.

Candidate	Selection
kaon	$p_T > 600 \text{ MeV}/c^2$
kaon	$p > 10 \text{ GeV}/c^2$
kaon	$\chi_{\text{IP}}^2 > 16$
kaon	$\text{DLL}_{K\pi} > 3$
slow pion	$\chi_{\text{IP}}^2 < 9$
dikaon	$\text{sum } p_T > 2.5 \text{ GeV}/c^2$
dikaon	$\text{mass} < 1800 \text{ MeV}/c^2$
dikaon	$\chi_{\text{FD}}^2 > 80$
dikaon	$\cos \theta > 0.999$
dikaon	$\text{vertex } \chi^2 < 4$
dikaon + π_s^+	$m_{hh\pi_s} - m_{hh} < 250 \text{ MeV}/c^2$
event	$\text{prescale} = 0.5$

Table 4.2: Stripping selection criteria for the K_s^0 reconstruction analysis. Note the tight p_T requirements on kaons and prescale. The χ_{IP}^2 (χ_{FD}^2) variable is the increase in vertex χ^2 if the signal track(s) is combined with the PV, and θ is the angle between signal candidate direction and the line made by the decay vertex and the PV.

Offline, a downstream track is combined with the $\phi\pi_s^+$ candidate. This downstream track has the selection criteria shown in Table 4.3 applied. The p_T requirement is the most important, as it rejects a large amount of background where a true $D^0 \rightarrow \phi X$ decay is combined with a random downstream track. Given the tight p_T requirements placed on the $\phi\pi_s^+$ candidate in the trigger, downstream tracks which originate from the same decay are also expected to have high p_T , which is not the case for a random, unassociated track.

For the fully reconstructed decay, where both K_s^0 daughters are reconstructed, at least one of the tracks is required to pass the selection shown in Table 4.3. The K_s^0 candidate is then combined by simply adding up the four-vectors of the downstream tracks, rather than performing a vertex fit. This procedure separates the tracking and vertex efficiency, which is desired as the

Candidate	Selection
downstream track	$p_T > 800 \text{ MeV}/c$
downstream track	$9 < \chi_{\text{IP}}^2 < 1500$
downstream track	$\text{IP} < 100 \text{ mm}$
downstream track	$\text{ProbNNpi} > 0.2$

Table 4.3: Offline selection criteria on the downstream track. The ProbNNpi variable is a neural-net based PID criterion.

reconstruction efficiency depends on the tracking efficiency squared (as there are two tracks), while it is proportional to the vertex efficiency.

4.3. Estimating the four-vector of the missing track

Estimating the kinematics of the missing K_s^0 daughter is important as it improves the mass resolution of the signal. It also allows the tracking efficiency measurement to be binned in momentum so that the results can be applied to decay modes with different kinematics. To improve the mass resolution a simple solving method, using only pointing and K_s^0 mass requirements can be used and is described in Sect. 4.3.1. This method does not depend on the D^0 and D^{*+} masses and so can be safely used without artificially peaking the background. For the momentum estimation, a different method is used, described in Sect. 4.3.2, where all three mass constraints are combined using a kinematic fit.

4.3.1. Solving using K_s^0 mass constraint

The kinematics of the missing track can be solved exactly using the K_s^0 mass and the D^0 pointing requirements. If one follows through the algebra, a quadratic equation can be obtained with the following coefficients

$$a = 4((P_{\phi+\pi_D}^{\parallel} - P_{\phi}^{\parallel})^2 - E_{\pi_D}^2) \quad (4.1)$$

$$b = 4(P_{\phi+\pi_D}^{\parallel} - P_{\phi}^{\parallel})m_{\text{miss}} \quad (4.2)$$

$$c = m_{\text{miss}}^2 - 4(E_{\pi_D}^2 (P_{\phi+\pi_D}^{\perp 2} + m_{\pi}^2)) \quad (4.3)$$

$$m_{\text{miss}} = m_{K_S^0}^2 - m_{\pi}^2 + P_{\phi+\pi_D}^2 + P_{\phi}^2 + E_{\pi_D}^2 - 2(P_{\phi+\pi_D}^{\parallel} P_{\phi}^{\parallel} + P_{\phi+\pi_D}^{\perp 2}) , \quad (4.4)$$

where π_D is the reconstructed downstream track, ϕ is the K^+K^- pair and $P^{\parallel}$ ($P^{\perp}$) is the component of momentum which is parallel (perpendicular) to the D^0 flight direction. The solutions of which are quite obviously

$$P_{\text{solved}}^{\parallel} = \frac{-b \pm \sqrt{(b^2 - 4ac)}}{2a} . \quad (4.5)$$

Sometimes, the solutions are unphysical (i.e. $4ac < b^2$) due to resolution effects. In this case b is set to $(-b + \sqrt{(0.01b^2)})/2a$, which forces a physical solution and keeps the mass distribution centred at the nominal $D^{*+} - D^0$ mass difference.

This resolution obtained by this technique is very poor, mainly due to the bad ϕ vertex and downstream track resolution. The contribution to the resolution for each input is shown in Fig. 4.3, where different parts of the decay chain are calculated using truth information available in the simulation and the resolution reassessed. The sharp cut-off at +1 in Fig. 4.3 comes from the axis definition: both P_{reco} and P_{true} are positive and if $P_{\text{true}} \gg P_{\text{reco}}$ then $(P_{\text{reco}} - P_{\text{true}})/P_{\text{true}} \rightarrow 1$, whilst if $P_{\text{true}} \ll P_{\text{reco}}$ then $(P_{\text{reco}} - P_{\text{true}})/P_{\text{true}} \rightarrow -\infty$.

In all subsequent analysis, a loose window around m_{D^0} of $1700 - 2400 \text{ MeV}/c^2$ is used, where m_{D^0} is the solved D^0 mass. The D^0 mass resolution is not good enough to warrant fitting this variable as well as the $D^{*+} - D^0$ mass difference.

4.3.2. Kinematic fit

To estimate the momentum of the missing track, a kinematic fit is performed using the Lagrange multiplier method, where more details can be found in Ref. [52]. Each four-vector is rotated into the coordinate system parallel to the D^0 flight so that the perpendicular components, p_X and p_Y

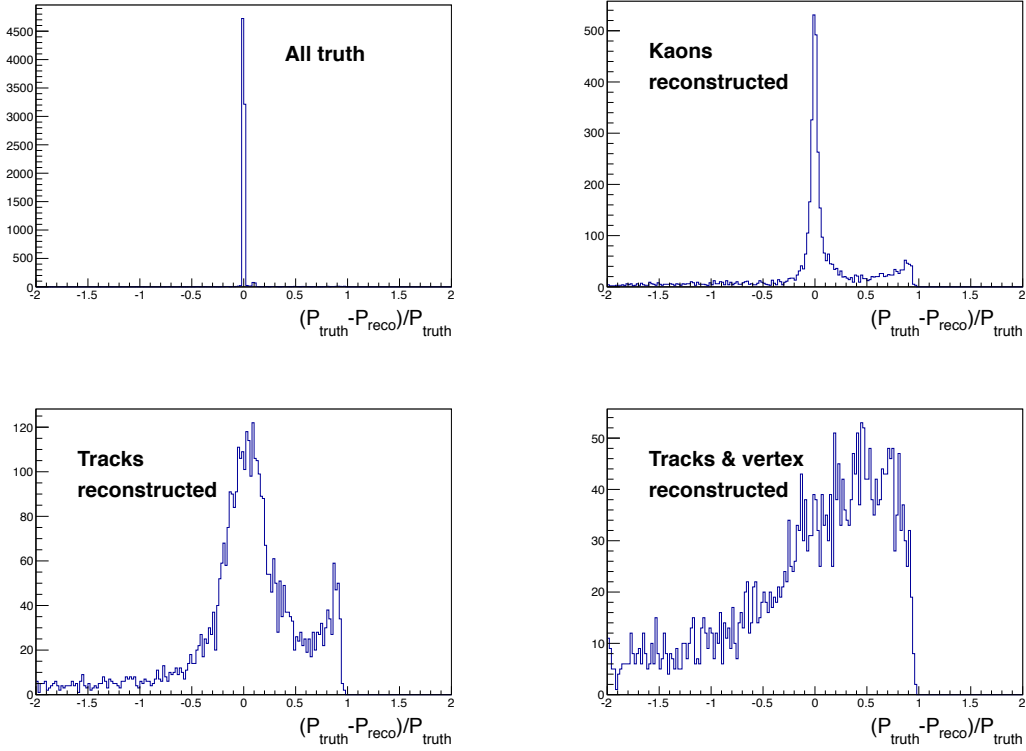


Figure 4.3: Momentum resolution of the missing downstream track for the solving method. Top left shows the resolution using the truth information (closure test). Top right uses truth for the downstream track and vertex information, bottom left uses true vertex information and bottom right uses all reconstructed quantities.

should be zero. These two constraints and the K_s^0 , D^0 and D^{*+} masses make up five constraints in the fit, with three unknowns.

The performance of the kinematic fit is significantly better than the solving method, shown in Fig. 4.4. It is still clear however that the resolution is still poor. The resolution can be improved by applying selection requirements on the fit χ^2 . However this removes a significant amount of signal and biases the mass distribution which peaks the background.

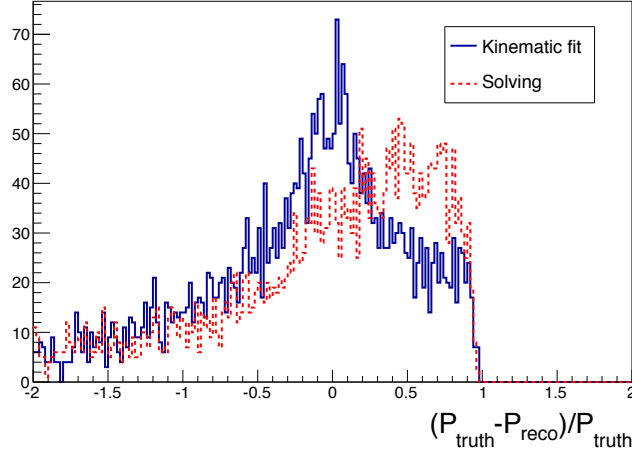


Figure 4.4: Resolution for the kinematic fitting compared to the solving method.

4.3.3. Unfolding

The fitted momentum distribution is unfolded using the fully reconstructed data. This is performed by constructing the so-called *efficiency matrix*, defined as,

$$\epsilon_{ij} = \frac{N_{\text{full}}(i, j)}{\sum_i N_{\text{full}}(i, j)} \quad (4.6)$$

$$\sum_i \epsilon_{ij} = 1$$

where N_{full} is the number of fully reconstructed data which have a reconstructed momentum in bin i but inferred momentum in bin j . The unfolded yield in each bin is then,

$$N_{\text{unfolded}}^i = \sum_j \epsilon_{ij} N_{\text{infer}}^j \quad (4.7)$$

where N_{infer}^j is the raw signal yield determined from partial reconstruction in bin j of inferred momentum. The error matrix on the quantities N_{unfolded}^i is determined by propagating through

the estimated error matrix on N_{infer}^j using,

$$U_{\text{meas}} = W V_{\text{infer}} W^T \quad (4.8)$$

where V_{infer} stands for the estimated error matrix from the maximum likelihood fit to the data (diagonal matrix). The momentum distribution of the missing track before and after the unfolding procedure is shown in Fig. 4.5. There is a large amount of bin migration, which biases the momentum distribution towards higher momenta. This is the reason why the unfolding procedure is used as it corrects for this migration.

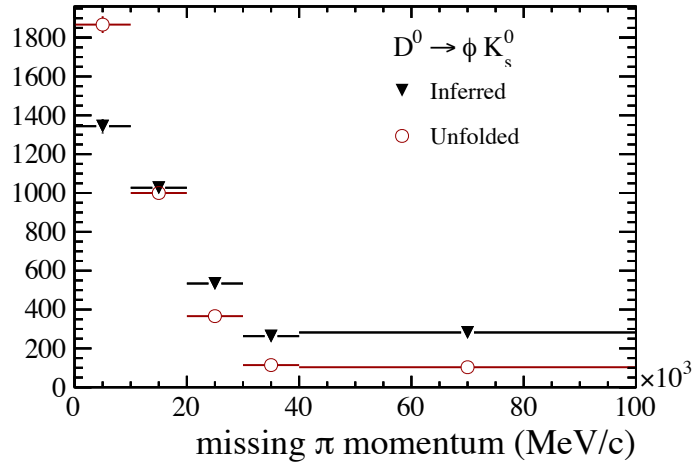


Figure 4.5: Momentum distribution of fully reconstructed data before and after unfolding. The unfolded and fully reconstructed momentum distributions agree perfectly by construction.

4.4. Backgrounds

The key to this measurement is estimating the shape of the background in the $m_{D^{*+}} - m_{D^0}$ spectrum. This is the main attraction of $D^0 \rightarrow \phi K_s^0$ decays as the data is dominated by true ϕ mesons, which reduces the list of backgrounds to consider (e.g. $K^+ \rightarrow \pi^+$ mis-id background is negligible). This also means one can reproduce the random slow pion background by mixing different events (see Sect. 4.4.1). The $D^0 \rightarrow \phi X$ branching fraction is also known, which allows background from $D^0 \rightarrow \phi X$ decays (X not reconstructed) combined with a random downstream track, to be constrained in the analysis. This background is discussed in Sect. 4.4.2 and is

called “semi-combinatorial background“ for the rest of this chapter. Finally the low Q-value for this decay ($350 \text{ MeV}/c^2$) means that there are not many backgrounds from $D^0 \rightarrow \phi K_s^0 X$, as phase-space suppression starts to take effect. Background which falls in this $D^0 \rightarrow \phi K_s^0 X$ category is $D^0 \rightarrow \phi(K^{*0} \rightarrow K_s^0 \pi^0)$ decays, discussed in Sect. 4.4.3, which turns out to be the only non-negligible source.

4.4.1. Random slow pion background

The dominant background for this analysis is where a ϕ + downstream track candidate has been randomly combined with a pion from the primary vertex and happens to sit in the narrow $m_{D^{*+}} - m_{D^0}$ window of $< 200 \text{ MeV}/c^2$. Although the amount of this background cannot be estimated, the shape can, by combining D^0 candidates and slow pions from different events. This technique works as for background the slow pion candidate is independent of the D^0 candidate as the D^0 is significantly displaced from the PV. Once these events have been mixed the trigger selection is reapplied, however this makes no difference to the shape.

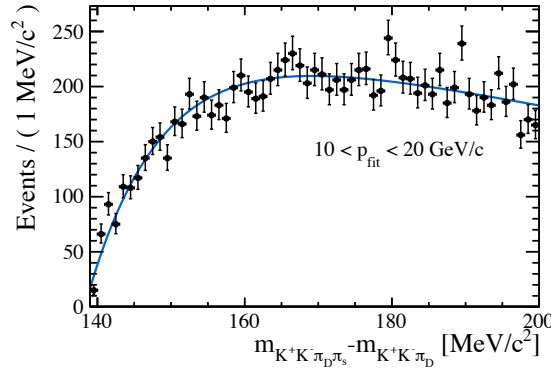


Figure 4.6: Mass fit to $m_{D^{*+}} - m_{D^0}$ in “mixed events”, where a ϕ + downstream track and π_s have been combined from different events. The solving method described in Sect. 4.3.1 is applied to these mixed candidates to form the $m_{D^{*+}} - m_{D^0}$ mass distribution.

The $m_{D^{*+}} - m_{D^0}$ distribution of these mixed candidates are then fit with an empirical threshold function, shown in Fig. 4.6. The uncertainty on the shape is then propagated to the fit to the data.

4.4.2. Semi-combinatorial background

The most dangerous background in this analysis is background from $D^0 \rightarrow \phi X$ decays, where the ϕ is combined with a random downstream track. This background peaks in $m_{D^{*+}} - m_{D^0}$ and the mass resolution is not good enough to separate it from signal in the m_{D^0} distribution.

The shape and yield of this background are estimated by fully reconstructing $D^0 \rightarrow \phi\pi^+\pi^-$ decays, at which point the signal is almost 100% clean. The $\pi^+\pi^-$ pair are then ignored and a downstream track is combined instead. In this way, there is a pure sample of the relevant background in data. The full selection chain, including estimating the missing downstream track is performed on this sample and a fit is performed to the remaining candidates. The resulting mass shape is shown in Fig. 4.7, where the shape is similar to the signal.

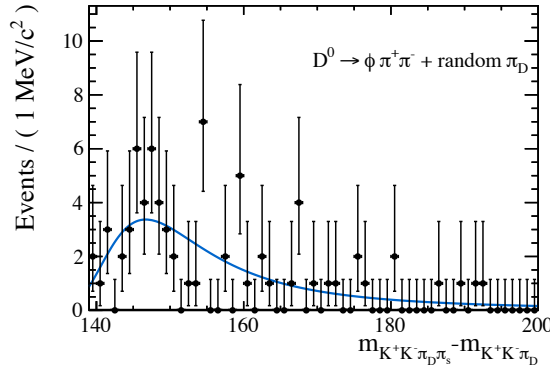


Figure 4.7: Mass fit to selected $D^0 \rightarrow \phi\pi^+\pi^-$ decays, where the ϕ has been combined with a random downstream pion (π_D) from the data. This background peaks in a similar way to signal.

The yield of this background is obtained by correcting for the reconstruction efficiency of the two pions (which are not reconstructed in the signal sample). This efficiency is obtained using simulated $D^0 \rightarrow \phi\pi^+\pi^-$ decays, and is calculated to be about 30%. There is some uncertainty associated with this number, however this is small compared to the statistical uncertainty on the background source.

Finally, to convert the $D^0 \rightarrow \phi\pi^+\pi^- + \text{random } K_s^0$ background yield into $D^0 \rightarrow \phi X + \text{random } K_s^0$, where X is not a K_s^0 , the branching fractions $\mathcal{B}(D^0 \rightarrow (\phi \rightarrow K^+K^-)X) = 0.52 \pm 0.05\%$, $\mathcal{B}(D^0 \rightarrow (\phi \rightarrow K^+K^-)\pi^+\pi^-) = (0.12 \pm 0.02)\%$ and $\mathcal{B}(D^0 \rightarrow (\phi \rightarrow K^+K^-)K_s^0) = (0.21 \pm 0.02)\%$ are used [14]. The total number of expected semi-combinatorial background present in the full signal sample is 780 ± 180 , which is about $(6 \pm 1)\%$ of the signal. This constraint is used in the

fit. For the individual momentum bins, the background sample is also split and this constraint is re-calculated.

4.4.3. $D^0 \rightarrow \phi K^{*0}$

The decay $D^0 \rightarrow \phi K^{*0}$, where the K^{*0} decays to $K_s^0 \pi^0$, has a branching fraction of $(2.6 \pm 0.5) \times 10^{-5}$, roughly 1% of the signal branching fraction. The mass shape and selection efficiency of this background is assumed to be identical to signal. As the branching fraction is low and has a big uncertainty, any effects which invalidate this assumption will have a negligible effect on the signal yield.

4.5. Mass fits

In this section, mass fits to the data, used to determine the partially and fully reconstruction yields, are described.

4.5.1. Signal shape

The signal shape is obtained by fitting fully reconstructed $D^0 \rightarrow \phi K_s^0$ data, where one of the downstream pions has been ignored and solving method has been applied. The shape is parameterised with a Crystal Ball shape [53], shown in Fig. 4.8. The uncertainty on this shape is propagated into the partially reconstructed signal fit.

4.5.2. Partially reconstructed fits

A fit to the data in bins of track momentum is performed to determine the signal yield before reconstructing the second downstream track. This fit is shown in Fig. 4.9 for fitted momenta in the range $10 - 20 \text{ GeV}/c$, where signal yield has roughly 10% precision. There are only two free parameters in this fit - the signal yield and the random slow pion background yield. Given the relatively small amount of freedom allowed in the fit, the data fits well, which gives confidence that the auxiliary measurements were good proxies and that there is no missing backgrounds to be considered.

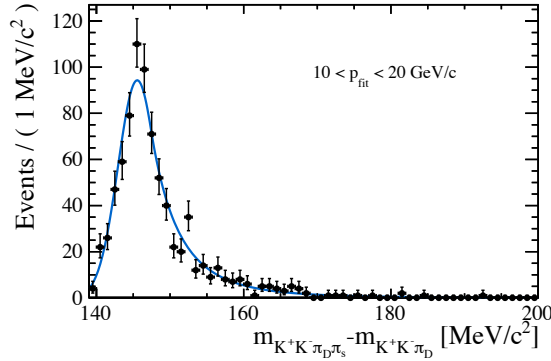


Figure 4.8: Mass fit to fully reconstructed signal, where one of the downstream pions is ignored and had its momentum estimated using the solving method as described in Sect. 4.3.1.

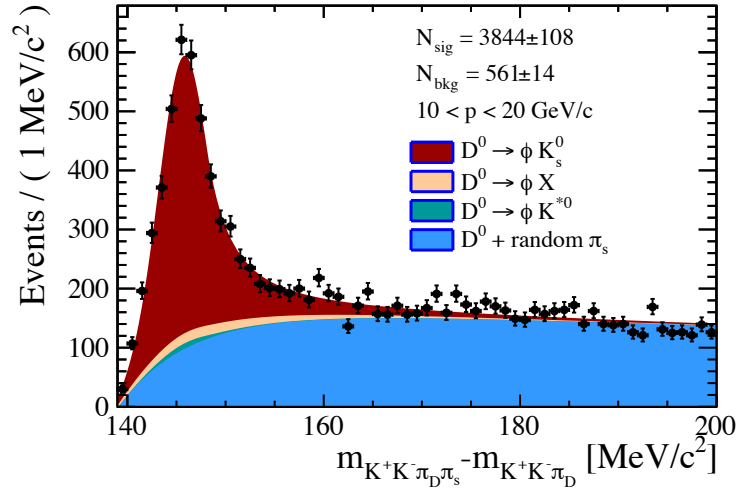


Figure 4.9: Mass fit to the partially reconstructed data for an example momentum bin.

4.5.3. Fully reconstructed fits

In this section, mass fits to fully reconstructed data, where the K_s^0 is obtained by combining the two daughter pion momenta, is performed. No vertex fit has been performed and the momenta are evaluated assuming the K_s^0 decayed in the TT. This assumption causes the K_s^0 mass resolution to be much worse than when carrying out the vertex fit, as the daughters will traverse different integrated B fields depending where the K_s^0 decayed. To calculate the correct momentum of the K_s^0 , the track momentum must be extrapolated to the decay vertex, which is most likely not in the TT.

For fully reconstructed candidates, a loose window around $m_{\phi K_s^0}$ of $1860 \pm 200 \text{ MeV}/c^2$ is applied, which substantially reduces the level of peaking background. The $D^{*+}-D^0$ mass spectrum is then fit, with the signal shape taken from the simulation and the background let free. The result is shown for an example momentum bin in Fig. 4.10.

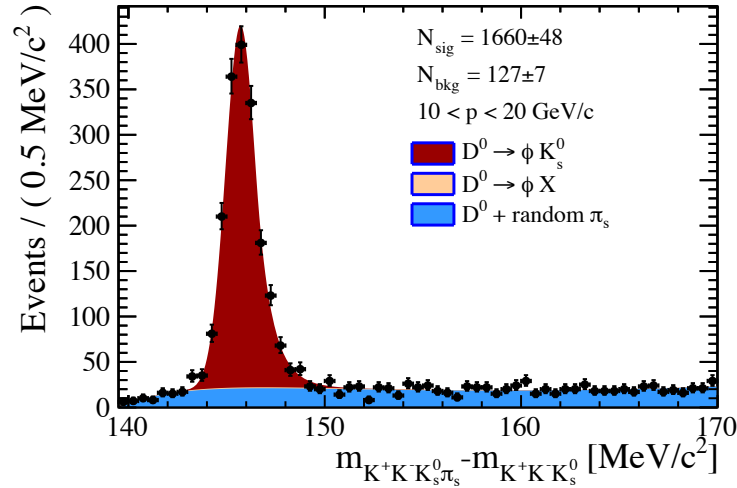


Figure 4.10: Mass fit to the fully reconstructed data in an example bin of track momentum.

4.5.4. Vertex fit

A vertex fit is also performed to the K_s^0 candidate and the selection criteria shown in Table 4.4 are applied. These are applied as standard in the LHCb software, and will be referred to the *standard K_s^0 selection* for the rest of this chapter. Mass fits to the fully reconstructed data are performed with and without the standard K_s^0 selection applied, where the ratio of signal yields defines the vertex efficiency shown in Sect. 4.6.2. An example of such fits are shown in Fig. 4.11, where the background has been substantially reduced by the standard K_s^0 selection.

4.6. Results

In this section, the tracking and vertex efficiency of DD K_s^0 candidates are shown as a function of momentum. For the vertex efficiency, the LL category results are also shown, which were obtained with a similar method.

Quantity	Selection
vertex fit convergence	True
mass window (pre-fit)	$\pm 80 \text{ MeV}/c^2$
mass window (post-fit)	$\pm 64 \text{ MeV}/c^2$
vertex χ^2/ndf	< 25
track DOCA χ^2	< 25

Table 4.4: Criteria to make a standard LHCb K_S^0 candidate from two downstream pions. Track DOCA is the distance of closest approach between the two daughter tracks.

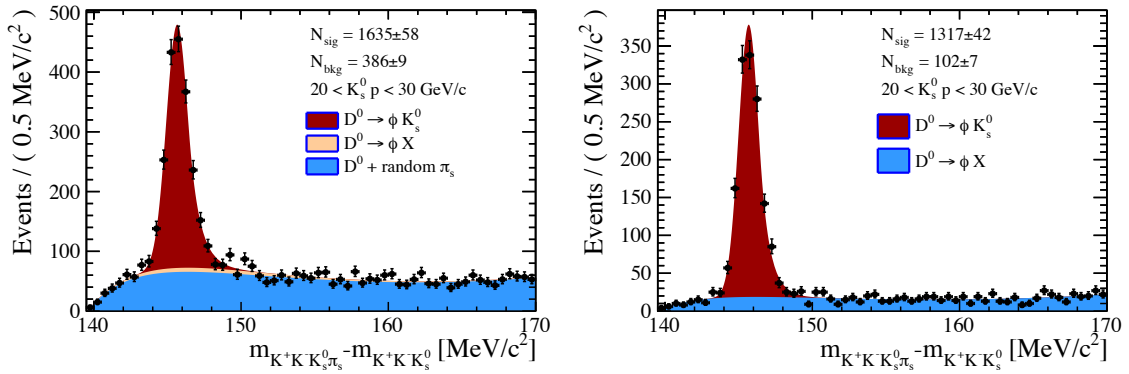


Figure 4.11: Mass fits to the fully reconstructed data before (left) and after (right) the standard K_S^0 selection criteria has been applied.

4.6.1. Tracking efficiency

The tracking efficiency is obtained by unfolding the partially reconstructed yields for each momentum bin shown in Sect. 4.5.2 and dividing them by the fully reconstructed yields shown in Sect. 4.5.3. This is then compared to the simulation, where the efficiency is trivial to measure. The tracking efficiency results for the data and simulation are shown in Fig. 4.12. The efficiency to reconstruct a second track, given the first, is about 50% and appears to slightly decrease with momentum, although this is not a significant trend.

The tracking efficiency ratio between the data and simulation is shown in Fig. 4.13. This ratio is just below unity and has no significant dependence on momentum, which means that analyses that normalise to a control sample including DD K_S^0 will not be affected by the downstream tracking efficiency difference at this level of precision. If the discrepancy is purely due to tracks

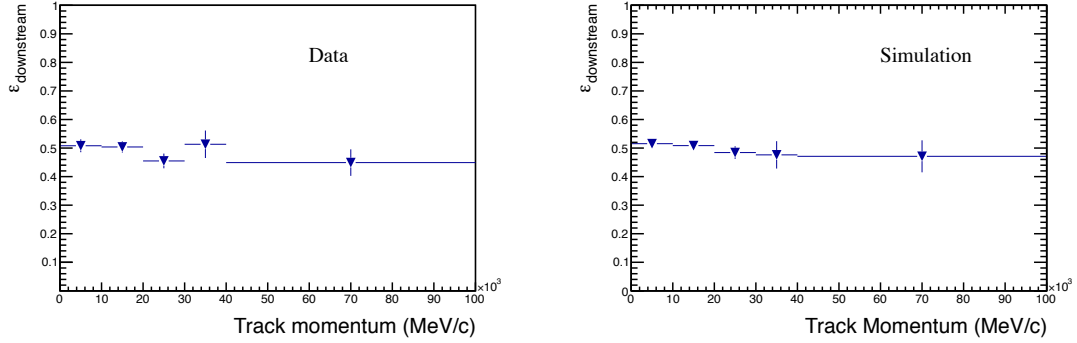


Figure 4.12: Downstream tracking efficiency in bins of track momentum. Results are shown for data (left) and simulation (right).

being incorrectly matched with a VELO track and upgraded to long, then the discrepancy should disappear when the simulation is re-weighted in occupancy. The tracking efficiency ratio after this re-weighting procedure is also shown in Fig. 4.13, where the efficiency ratio is consistent with unity.

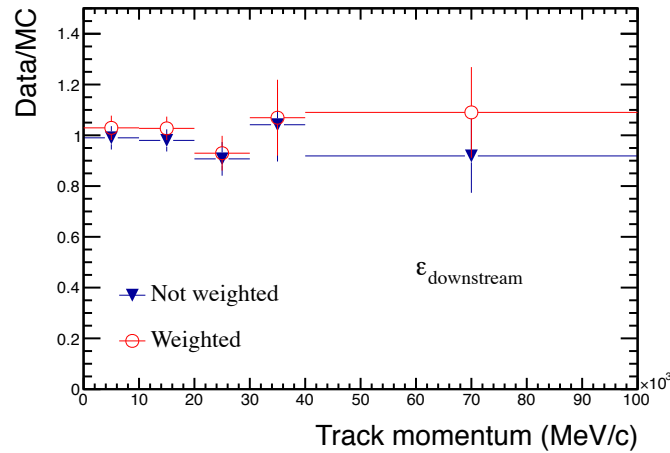


Figure 4.13: Downstream tracking efficiency ratio between data and simulation in bins of track momentum. Results are shown for raw simulation and also for simulation where the number of VELO tracks has been re-weighted to match data.

The effect of the large bin migration is estimated by repeating the measurement but using the truth level information in the simulation rather than the kinematic fit/unfolding procedure. The result can also be repeated without any unfolding to either data or simulation. These comparisons

are shown in Fig. 4.14. One possibility to take this into account as a systematic would be to re-weight the momentum spectra by the truth level curve as well as the nominal curve to assess a systematic effect. However, this is conservative as the migration effect should cancel to some extent when applied to both simulation and data. Given that the efficiency ratio is consistent with unity after re-weighting for multiplicity, it was decided that it is safer for analyses to re-weight simulation samples by VELO track multiplicity, whereby the correct momentum dependence is indirectly applied.

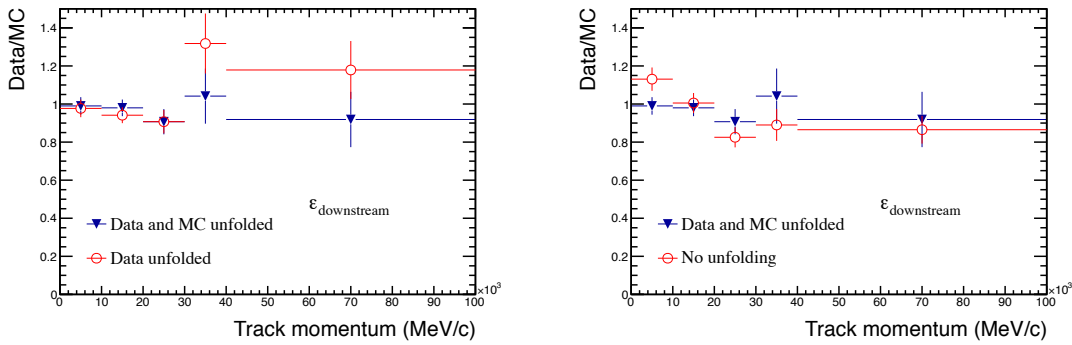


Figure 4.14: Downstream tracking efficiency in bins of track momentum. Results are shown where both data and simulation have been unfolded, where only data is unfolded, and where neither have been unfolded.

Given the tracking efficiency results suffer from large bin migration, it would be useful to improve this measurement in the future. The $\phi \rightarrow K^+ K^-$ decay is almost on threshold which is a very difficult decay vertex to measure accurately and is the main cause of the poor momentum resolution. Other D decays, such as $D^0 \rightarrow K_s^0 4\pi$ should have better resolution as there is a four track vertex with a relatively large Q-value. However it remains to be seen whether the background can be controlled to an acceptable level and the trigger efficiency will be lower as there are more particles in the final state. It should be noted that while decays such as $D^0 \rightarrow K_s^0 \pi^+ \pi^-$ have a larger branching fraction ($\sim 2\%$), they have large irreducible backgrounds from $D^0 \rightarrow K_s^0 \pi^+ \pi^- \pi^0$, which has a branching fraction of roughly 5%. As for a repetition of the $D^0 \rightarrow \phi K_s^0$ analysis, the stripping selection could be loosened to allow an inclusive ϕ trigger line, the prescale could be removed and the PID could be loosened. For the longer term, it may be beneficial to introduce a stripping line for ϕ plus slow pion combination to further improve statistics.

4.6.2. Vertex efficiency

As a reminder, the vertex efficiency is defined as the efficiency to vertex the two downstream tracks and pass the standard K_s^0 selection criteria. This measurement is carried out using fits similar to the ones shown in Figs. 4.11 in bins of K_s^0 momentum. The results for data and simulation are shown in Fig. 4.15, where the simulation has a flat vertex efficiency of roughly 96%. In data, the vertex efficiency is about 80%, with no significant dependence on momentum.

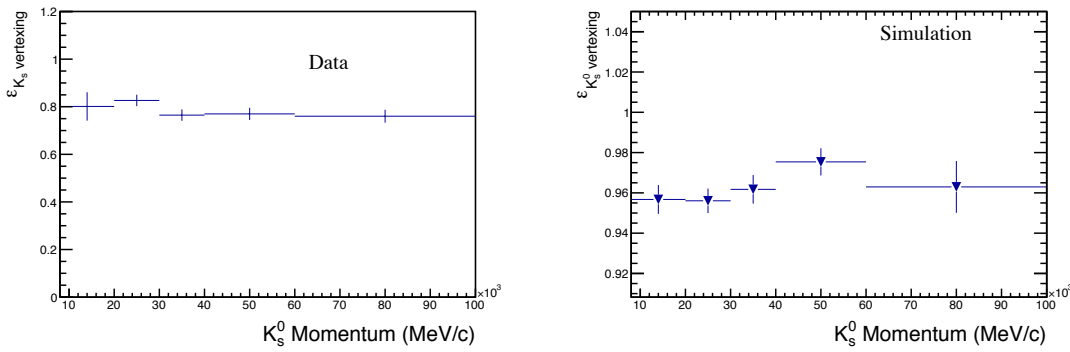


Figure 4.15: Efficiency to apply the standard K_s^0 selection criteria as a function of the K_s^0 momentum. Results are shown for both data (left) and simulation (right).

The ratio of the vertex efficiency between data and simulation is shown in Fig. 4.16. The ratio gets slightly worse with momentum which should be accounted for in the future for analyses involving K_s^0 mesons. The corresponding results for the LL category are also shown, where the ratio is much closer to one.

To investigate the cause behind the relatively large difference in the vertex efficiency, the K_s^0 vertex χ^2 is compared between data and simulation. This is shown in Fig. 4.17, where the black line indicates the cut applied in the standard K_s^0 selection. The cause of the lower selection efficiency in the data, and the discrepancy shown at the start of this chapter in Fig. 4.2, is due to badly vertexed K_s^0 candidates. It is currently unclear where these badly vertexed K_s^0 candidates originate from. One possibility is mis-alignment of the TT, which would affect the extrapolation of the daughters to the decay point and could cause mis-reconstruction of the vertex. Another possibility would be inaccuracies in the magnetic field map, which is also used to extrapolate tracks back to the decay point.

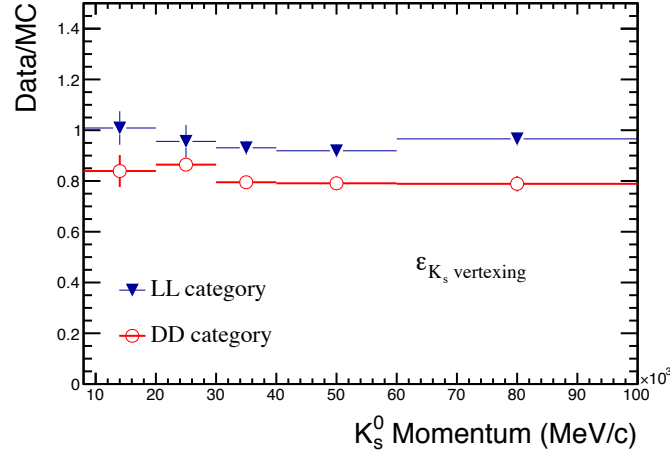


Figure 4.16: Efficiency ratio to apply the standard K_s^0 selection as a function of K_s^0 momentum between data and simulation.

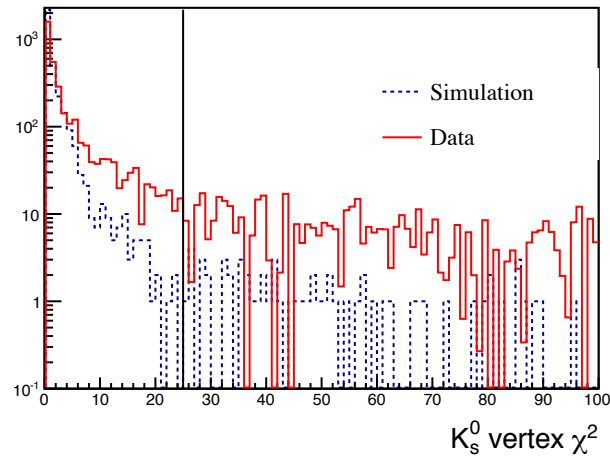


Figure 4.17: Vertex χ^2 for K_s^0 in data (s-weighted) and simulation. A black line is placed where the selection is applied for standard K_s^0 candidates.

4.7. Cross-check using $B \rightarrow J/\psi K$ decays

To cross-check the results described in Sect. 4.6, the ratio of yields between $B^0 \rightarrow J/\psi K_s^0$ and $B^+ \rightarrow J/\psi K^+$ decays is compared between data and simulation after the corrections have been applied to the K_s^0 sample. It should be noted that there are many systematics ignored for this

measurement, such as the IP resolution and long track efficiency. The PID and trigger systematics should be negligible as there is no PID applied to the K^+ , and also both decays are required to be triggered by the J/ψ candidate. The VELO track multiplicity of the $B^0 \rightarrow J/\psi K_s^0$ simulation sample is re-weighted using the difference in multiplicity obtained from $B^+ \rightarrow J/\psi K^+$ decays in data and simulation. Based on Fig. 4.13, this re-weighting procedure should account for the tracking efficiency ratio difference. The vertex efficiency ratio shown in Fig. 4.16 is also applied based on the K_s^0 momentum. The average vertex efficiency ratio is found to be $(82.4 \pm 1.5)\%$.

Once this correction has been applied, the following quantity can be measured,

$$\frac{f_d \mathcal{B}(B^0 \rightarrow J/\psi K^0)}{f_u \mathcal{B}(B^+ \rightarrow J/\psi K^+)} = 0.90 \pm 0.02 \text{ (tracking + vertexing)}. \quad (4.9)$$

Assuming the ratio of B^0 to B^+ fragmentation fractions, $\frac{f_d}{f_u}$, is unity¹, the branching fraction ratio is consistent with the ratio of B^0 and B^+ lifetimes; $\tau_0/\tau_+ = 0.93 \pm 0.01$ [14]. This is expected as the isospin asymmetry of $B \rightarrow J/\psi K$ is predicted to be small theoretically [54]. The branching fraction ratio is also consistent at the level of 1.3σ with the PDG ratio of (0.84 ± 0.04) [14]. The PDG ratio assumes equal production of $B^0 \bar{B}^0$ and $B^+ B^-$ in $\Upsilon(4S)$ decays. The consistency of the results with previous measurements and theoretical expectations gives confidence in the reconstruction efficiency results at the current level of statistical precision.

The branching fraction ratio as a function of K_s^0 momentum is shown in Fig. 4.18 for the DD and LL categories. After the correction procedure is applied, the branching fraction ratio is consistent with a flat line, as expected. The branching fraction ratio is also shown for LL K_s^0 , where the branching fraction ratio needs a much smaller vertex efficiency correction. After the corrections both categories agree nicely, which suggests that the cause of the discrepancy between the two has been found.

4.8. Summary

In summary, the downstream tracking and vertex efficiency of K_s^0 mesons have been measured using $D^{*+} \rightarrow (D^0 \rightarrow \phi K_s^0) \pi^+$ decays. At the current level of precision, there is no visible dependence

¹One would naively expect it to be slightly below unity, as there are twice as many valance u -quarks to combine with the b -quark.

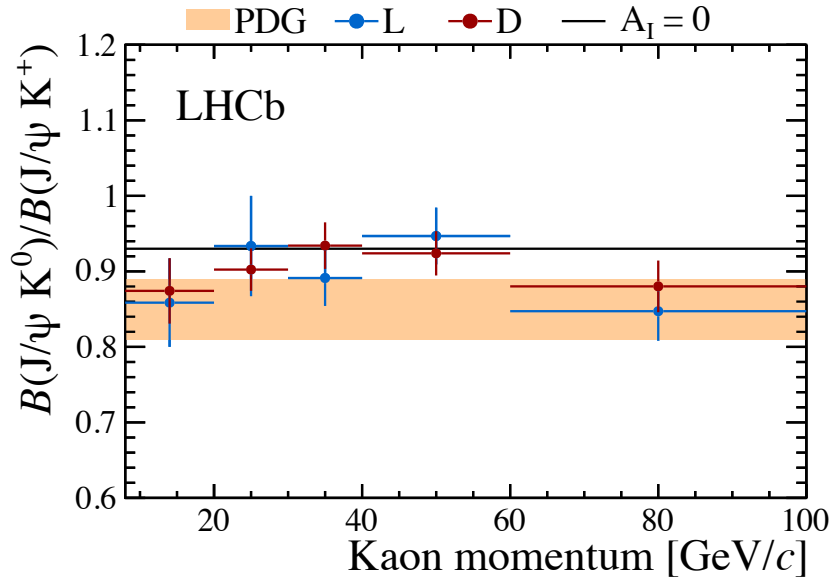


Figure 4.18: The branching fraction ratio, $\mathcal{B}(B^0 \rightarrow J/\psi K^0)/\mathcal{B}(B^+ \rightarrow J/\psi K^+)$, as a function of K_s^0 momentum. After correcting for the K_s^0 reconstruction, the DD category agrees with the world average and the theoretical expectation of zero isospin asymmetry between the modes.

of the tracking efficiency discrepancy between data and simulation. This is assuming that there are no residual effects from the unfolding, which is the weakness of this analysis. The tracking efficiency difference between the data and the simulation can be fully explained by incorrectly matching downstream tracks to a track in the VELO and upgrading it to a long track.

The vertex efficiency is about 20% worse in the data than the simulation and has a mild dependence on momentum, where there is negligible bin migration. This is mainly due to the standard vertex quality cut applied to K_s^0 candidates. The absence of a strong momentum dependence in the K_s^0 reconstruction is good news for analyses involving K_s^0 in LHCb. As long as one normalises to a control channel including DD K_s^0 , the effect should be negligible for most analyses, which is the case for the analyses described chapters 6 and 7.

Chapter 5.

Observation of a resonance in $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays at low recoil

In this chapter, the analysis of resonant dimuon structure observed in $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays, published in Ref. [1], is presented. More attention is paid to the novel parts of the analysis, rather than the components which are by now, fairly standard procedure in LHCb. All of the analysis in this chapter was performed by the author, with the exception of Fig. 5.14.

5.1. Introduction

As discussed in Chapter 2, $b \rightarrow s \mu^+ \mu^-$ transitions are excellent probes of the Standard Model. To maximise the sensitivity to new physics, it is crucial that the contamination from tree level diagrams is controlled. Most of this contamination originates from B decays into a narrow $c\bar{c}$ resonance, where the $c\bar{c}$ meson decays via a virtual photon into a pair of muons. Specifically, this refers to the J/ψ and $\psi(2S)$ mesons, which dominate the rate but are easily removed by vetoes on the dimuon mass. Prior to the work presented here, the rest of the dimuon mass region has been assumed to be free from tree-level decays. Charmonium resonances heavier than the open charm threshold, where the resonances are wide as decays to $D^{(*)}\bar{D}^{(*)}$ are allowed, are broad in nature and cannot be removed efficiently with simple vetoes. In the kinematic region where the hadron has a low recoil against a dimuon system, the *low-recoil* region, the contribution from these heavy $c\bar{c}$ states is the most important. This region is normally defined for signal with q^2 values greater than $15 \text{ GeV}^2/c^4$.

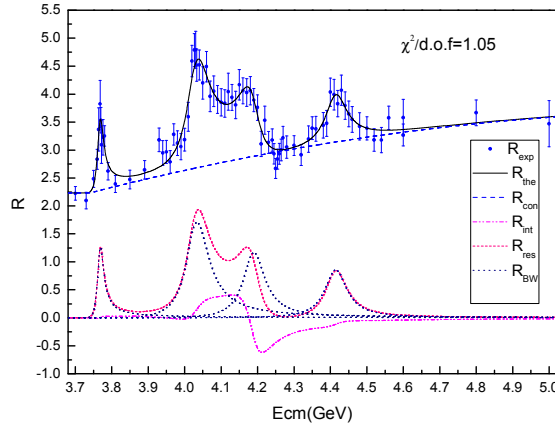


Figure 5.1: Fit to the e^+e^- hadronic cross section data. Figure and fit from Ref. [55]. Note the strong interference contribution labelled R_{int} .

Resonance	Mass [MeV/ c^2]	Width [MeV]
$\psi(3770)$	3773.2 ± 0.3	27.2 ± 1.0
$\psi(4040)$	4039.6 ± 4.3	84.5 ± 12.3
$\psi(4160)$	4191.7 ± 6.5	71.8 ± 12.3
$\psi(4415)$	4415.1 ± 7.9	71.5 ± 19.0

Table 5.1: The relevant resonance parameters of the fit to the BES data [55]. Note the rather misleading name of the $\psi(4160)$.

Nearly all available information about the heavy $c\bar{c}$ resonances, come from measurements of the cross-section ratio of $e^+e^- \rightarrow \text{hadrons}$ relative to $e^+e^- \rightarrow \mu^+\mu^-$. Among these analyses, only that of the BES collaboration in Ref. [55] takes interference and strong phase differences between the different resonances into account. The mass fit for this analysis is shown in Fig. 5.1, with the corresponding summary of the resonance parameters in Tab. 5.1. The strange name of the $\psi(4160)$ is historical, due to earlier, less precise measurements such as those from the DASP collaboration [56], which did not allow for interference between the different resonances. For this reason, the results in Tab. 5.1 are used as external information later in the analysis rather than the PDG averages, which incorrectly include results from analyses which didn't include interference.

Theoretically, the contributions from heavy $c\bar{c}$ resonances are treated as effective interactions [33]. This is valid in the low recoil region because the large value of q^2 means that charm loops are short distance compared to relatively soft energy of the hadron. The approach assumes a duality between quarks and hadrons, and results in smooth predictions even in regions where resonances are present.

The assumption is valid when predictions are integrated over a wide enough region in q^2 , where the entire low recoil is generally accepted to be wide enough. The breaking of quark-hadron duality, which converts this smooth prediction into a more realistic picture with resonance structures, is estimated to have an effect of about 2% in the low recoil region [57]. Another issue specific to the low recoil region, is the breakdown of QCDF which is the assumption that the $c\bar{c}$ and kaon systems can be treated independently, Care should therefore be taken when interpreting the results in this chapter, which implicitly assume QCDF.

The decay, $B^+ \rightarrow K^+ \mu^+ \mu^-$ was the first $b \rightarrow s \mu^+ \mu^-$ transition to be discovered [58]. Since then, it has been studied in great detail, with the most precise measurements made by the LHCb collaboration in Refs. [2,3,59]. The signal yield for this decay with the full LHCb dataset is roughly 5000, which is the largest of any $b \rightarrow s \mu^+ \mu^-$ transition in LHCb. The $\mu^+ \mu^-$ pair is also fully longitudinally polarised which simplifies the interference between loop and tree level contributions. These two aspects, make $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays the ideal place to study sub-leading QCD effects.

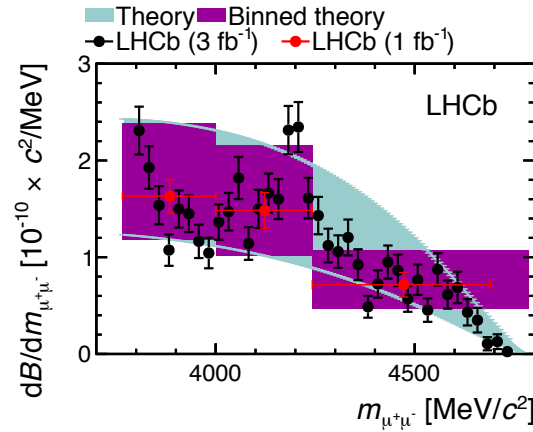


Figure 5.2: Differential branching fraction of the $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay in the low recoil region. The 3 fb^{-1} result used for this analysis is overlaid on the data from the LHCb analysis based on 1 fb^{-1} [60]. For the 3 fb^{-1} result, only statistical uncertainties are shown. The bands shown indicate the theoretical prediction for the differential branching fraction and are calculated using input from Ref. [61].

When the full 3 fb^{-1} dataset was first studied, a peak in the invariant mass distribution at around $4.2 \text{ GeV}/c^2$ of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays was noticed, shown in Fig. 5.2. The following sections detail the analysis performed with the purpose to test the compatibility of this peak to originate from a resonance.

5.2. Trigger, selection and backgrounds for $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays

The LHCb detector is well suited for $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays, which makes analysing them relatively simple. In the following sections, several favourable aspects of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays are discussed.

5.2.1. Trigger response

The trigger response for $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays, shown as a function of q^2 in Fig. 5.6, is very efficient, especially in the low recoil region. The trigger efficiency is obtained in simulation by requiring that the signal fires the set of trigger lines described in Sect. 3.2.3, which is dominated by the muons. This is why the trigger is more efficient at high q^2 , as the muons have higher p_T in this region.

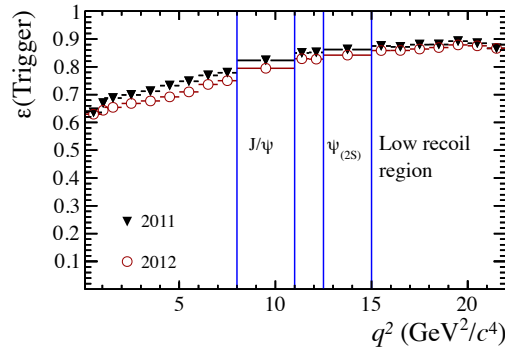


Figure 5.3: Trigger efficiency of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays as a function of q^2 . The trigger is dominated by the muons, which have higher p_T at high q^2 .

By studying the plentiful $B^+ \rightarrow J/\psi K^+$ decay, the quality of the simulation to emulate the trigger can be evaluated. This is performed using events which have been triggered independently of the signal (TIS events). Within these events, the number of events which have been triggered on the signal (TOS events) can be calculated. The trigger efficiency of signal in TIS events is known as the TISTOS efficiency. This allows the trigger efficiency to be evaluated in the data, and compared to the same technique in the simulation. The agreement, shown in Fig. 5.4, is very good for all trigger levels, and no further investigations are required.

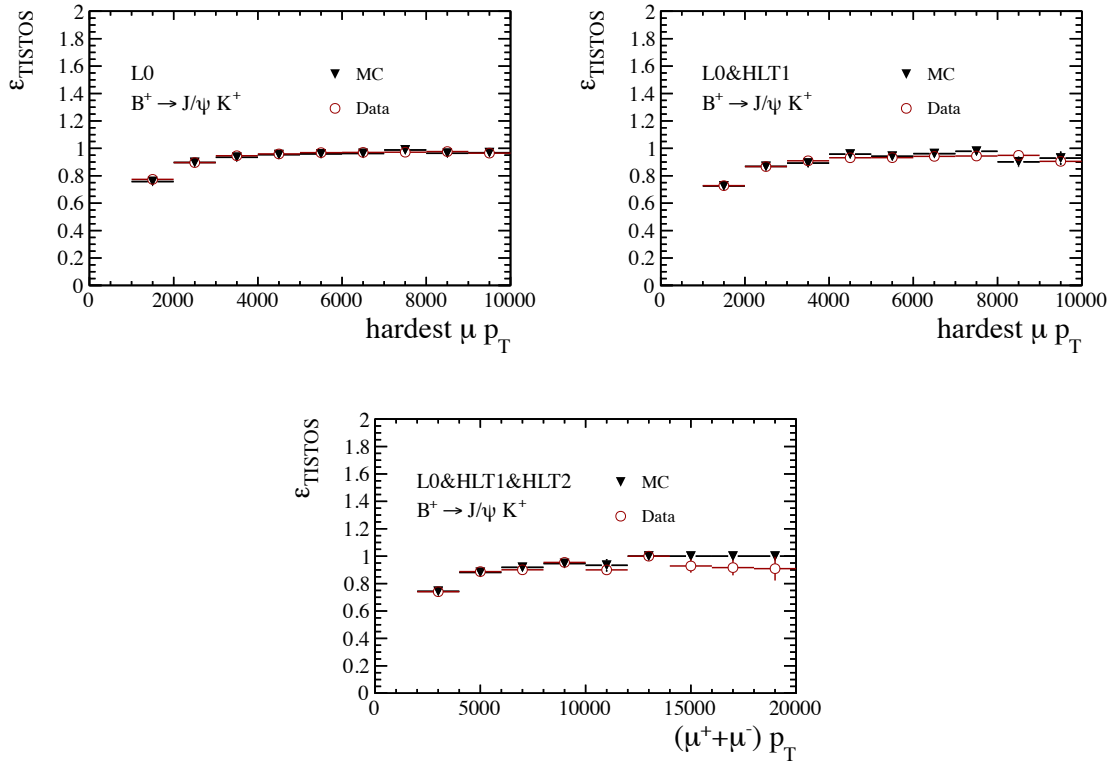


Figure 5.4: Trigger efficiency of $B^+ \rightarrow J/\psi K^+$ decays for each trigger level as a function of muon p_T . The trigger is calculated by studying events which have been triggered independently of signal.

5.2.2. Peaking background

The only relevant peaking background for $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays is $B^+ \rightarrow J/\psi K^+$ and $B^+ \rightarrow \psi(2S)K^+$, where the kaon and the same-sign muon have swapped identities. This background is investigated by looking at the $K^+ \mu^-$ invariant mass, under the $\mu^+ \mu^-$ hypothesis. If this mass is within $60 \text{ MeV}/c^2$ of the J/ψ or $\psi(2S)$ mass, then the kaon is required to fail the ISMUON requirement but be in the muon acceptance. This is over 99% efficient on signal, and reduces the background to a negligible level, as shown in Fig. 5.5.

Other backgrounds, such as $B \rightarrow (\bar{D}^0 \rightarrow K^+ \pi^-) \pi^+$ and $B \rightarrow K^+ \pi^- \pi^+$, are small as they require two $\pi \rightarrow \mu$ mis-identifications, which are highly suppressed in LHCb due to the excellent muon PID capabilities.

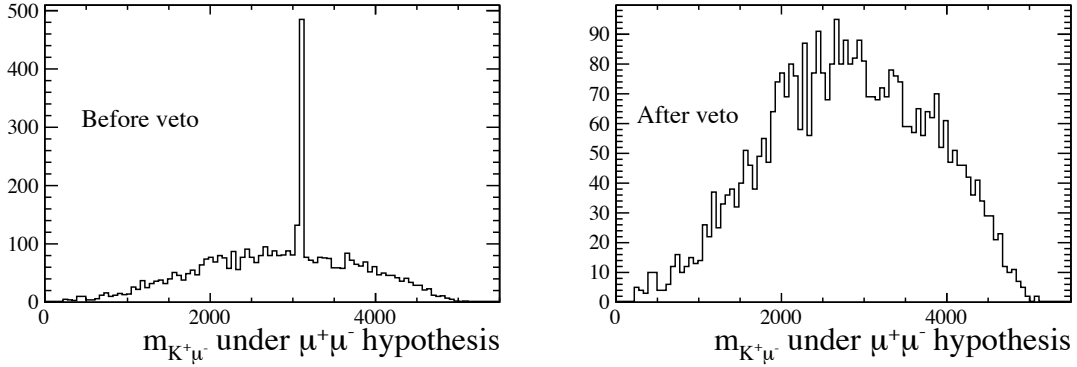


Figure 5.5: Invariant mass of the $K^+ \mu^-$ system under the $\mu^+ \mu^-$ mass hypothesis before (left) and after (right) the veto described in the text. There is no evidence for kaon-muon swaps after the veto.

5.2.3. Combinatorial background

The combinatorial background is even less of an issue than the peaking background. This is reflected in the high efficiency of the selection, which is based on a Boosted Decision Tree (BDT) [62]. The BDT uses kinematic and geometric variables, which are listed in descending order of importance in Tab. 5.2. Each of these variables are defined in chapter 4 and are well modelled in the simulation. The BDT is trained with a signal sample from simulation and a background sample consisting of 10% of the data from the sideband region, which is removed for the rest of the analysis. A cut is placed on the BDT to maximise $S/\sqrt{(S+B)}$, where S is the number of expected $B^+ \rightarrow K^+ \mu^+ \mu^-$ based on the number of $B^+ \rightarrow J/\psi K^+$ seen in data and B is the amount of background extrapolated into the signal window. The efficiency of this BDT cut on $B^+ \rightarrow K^+ \mu^+ \mu^-$ signal is 89%, whereas the efficiency for background is 6%. After this selection, the signal is very clean at high q^2 (see Sect. 5.2.4).

There is some shape to the BDT efficiency as a function of q^2 , as shown in Fig. 5.7. The BDT is less efficient at low recoil due to the inclusion of the kaon χ_{IP}^2 as one of the BDT variables. In the low recoil region, the kaon becomes aligned with the B and hence the PV, so that the kaon IP is not very significant. However, this is a small effect and the distortion is worth the extra background rejection power.

Variable
$K^+ \chi_{\text{IP}}^2$
B^+ vertex χ^2
$\mu \chi_{\text{IP}}^2$
$B^+ p_{\text{T}}$
$J/\psi \chi_{\text{IP}}^2$
B^+ IP (best PV)
$B^+ \chi_{\text{FD}}^2$
$B^+ \cos \theta$
$\mu \chi_{\text{IP}}^2$
$K^+ \text{ P}$
$B^+ \text{ P}$

Table 5.2: Variables used in the BDT ordered by importance according to the BDT. The χ_{IP}^2 (χ_{FD}^2) variable is the increase in vertex χ^2 if the signal track(s) is combined with the PV, and θ is the angle between signal candidate direction and the line made by the decay vertex and the PV.

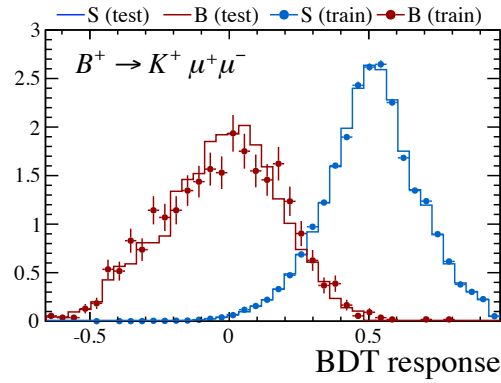


Figure 5.6: BDT response for the signal and background samples. Each sample is split into a training and testing part, the latter is used to check performance and optimise the cut placed on the BDT response.

5.2.4. Mass fit to B candidates

Finally, a mass fit to the low recoil region of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays is performed, where there is an excellent signal shape proxy in $B^+ \rightarrow \psi(2S)K^+$ decays, which have a slightly lower dimuon

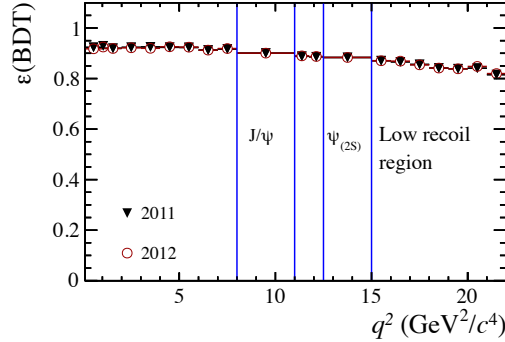


Figure 5.7: Efficiency of the BDT as a function of q^2 for $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays. In general, the BDT efficiency is very good, reflecting the favourable topology of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays.

mass than the signal. The fit to this signal proxy is shown in Fig. 5.8, which is parameterised by a sum of two Crystal Ball functions [53] with common tail parameters, but different widths.

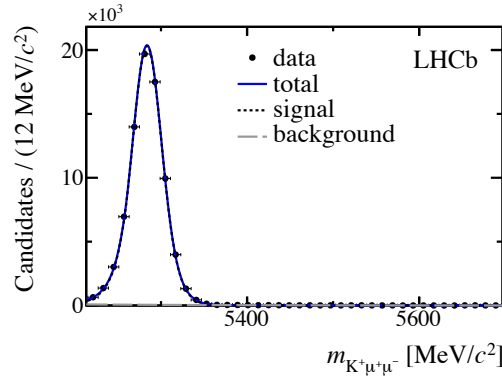


Figure 5.8: Mass fit to $B^+ \rightarrow \psi(2S)K^+$ decays, which is where the signal shape is obtained from. The lower mass sideband is very short to avoid background from partially reconstructed decays.

The shape from Fig. 5.8 is corrected by a small difference between the $B^+ \rightarrow \psi(2S)K^+$ and $B^+ \rightarrow K^+ \mu^+ \mu^-$ width and mean parameters. This correction is obtained from the simulation and applied to the low recoil data, shown in Fig. 5.9. Whether this signal shape correction is applied is irrelevant to the analysis. The small amount of background is fit with an exponential, which is used to extrapolate the amount of background in the signal region, defined as $\pm 40 \text{ MeV}/c^2$ around the nominal B^+ mass. The signal-to-background ratio in the signal region is about eight.

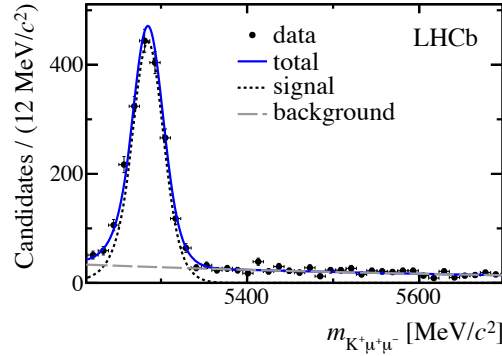


Figure 5.9: Mass fit to $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays in the low recoil region, which is used to estimate the amount of background in the $\mu^+ \mu^-$ spectrum. The lower mass sideband is very short to avoid background from partially reconstructed decays.

5.3. Dimuon fit

There is a large signal sample in the region of interest, where the small amount background has been estimated using a mass fit. Now the main part of the analysis begins, which is the construction of the fit model to the dimuon mass.

5.3.1. Kinematic constraint

Before the fit is performed, when the dimuon mass is calculated, a kinematic fit is performed [63]. In the fit, the $K^+ \mu^+ \mu^-$ mass is constrained to the B^+ PDG value and the B candidate is required to point back to the PV. These constraints dramatically improve the dimuon mass resolution, as shown in Fig 5.10. The main source of improvement is due to the B mass constraint rather than the PV constraint. For the unconstrained dimuon mass, the resolution gets worse as a function of $m_{\mu^+ \mu^-}$, as tracks have higher momentum at higher mass. For the constrained dimuon mass, the resolution gets better with $m_{\mu^+ \mu^-}$, as there is less phase space in the decay and so the B mass constraint is more powerful. Following this constraint, the dimuon mass resolution is good enough to expand the low recoil region (nominally at $15 \text{ GeV}^2/c^4 = 3873 \text{ MeV}/c^2$) down to $3770 \text{ MeV}/c^2$ and include part of the $\psi(3770)$ resonance. The region cannot be further expanded without contaminating the data with $\psi(2S)$ decays, the tail of which is difficult to model and would introduce a systematic uncertainty. It is also safe to neglect the resolution in the mass fit, as it is small compared to the widths of the resonances that are subsequently analysed.

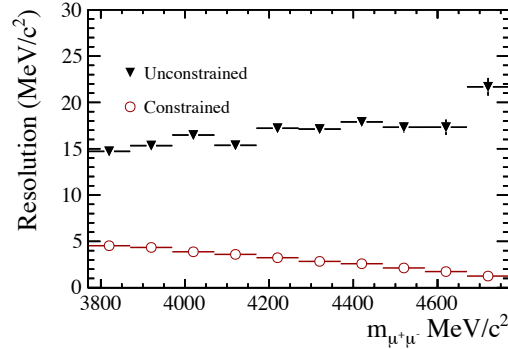


Figure 5.10: Dimuon mass resolution of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays as a function of mass, obtained with the simulation. Results are shown before and after the kinematic constraint described in the text.

5.3.2. Background shape

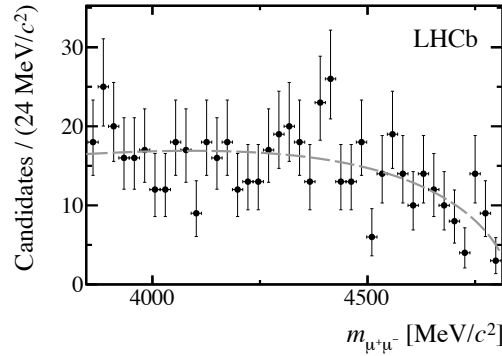


Figure 5.11: Fit to the unconstrained dimuon mass in the sideband region of the $K^+ \mu^+ \mu^-$ mass. The shape is parameterised by an ARGUS function [64].

The shape of the background in dimuon mass is estimated by fitting the unconstrained dimuon mass in the sideband. The data are fit with an ARGUS function [64], where all parameters are allowed to vary. In the signal region and under the kinematic constraint, the threshold of the ARGUS function is fixed to the $B^+ - K^+$ mass difference. The reason why the unconstrained dimuon mass is used is because the kinematic constraint warps the background shape when the $K^+ \mu^+ \mu^-$ is far from the PDG B^+ mass. Near the B^+ mass, it is assumed that the background shape is not significantly affected by the kinematic constraint. This assumption is checked by constructing a mass window as wide as the signal region in the sideband, and constraining the

dimuon mass in that window to the centre of it. The background shape is consistent before and after the constraint which suggests that the shape obtained from the unconstrained mass can be trusted. The statistical uncertainty on the shape, and the correlation of shape parameters is propagated to the dimuon mass fits in the signal region.

5.3.3. Efficiency shape

The relative efficiency between $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ decays as a function of mass is shown in Fig. 5.12. It drops to zero at high mass due an χ^2_{IP} cut on the K^+ in the stripping; close to the kinematic end point, the kaon is stationary in the B rest frame and so becomes collinear with the B flight direction and hence the PV. It is same reason why the BDT efficiency is worse at high q^2 .

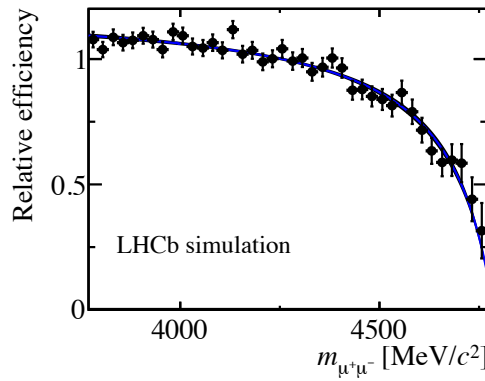


Figure 5.12: Relative efficiency between $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ decays as a function of dimuon mass. The efficiency drops at high mass due to IP selection requirements.

The reliability of the efficiency shape depends on how reproducible the IP resolution is in the simulation. Unfortunately, the IP resolution is not perfectly well modelled, at least for the simulation samples which are used for this analysis. For this reason the IP resolution is artificially degraded by approximately 20%, after which the $K^+ \chi^2_{\text{IP}}$ matches well between data and simulation, as shown in Fig. 5.13.

The signal PDF is multiplied by efficiency as a function of dimuon mass. The efficiency shape has a small effect on the result as the efficiency only drops in the kinematic region where there is a low signal yield due to phase-space suppression. This means that systematic effects associated with the efficiency are rendered negligible.

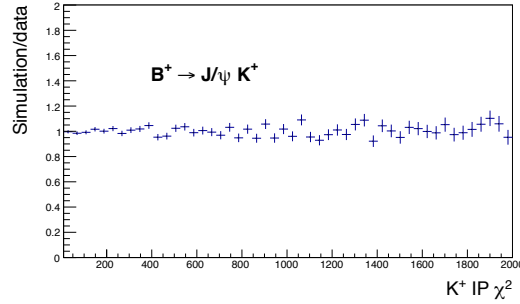


Figure 5.13: Ratio of the $K^+ \chi_{\text{IP}}^2$ distributions in the data and simulation for $B^+ \rightarrow J/\psi K^+$ decays.

5.3.4. Signal model

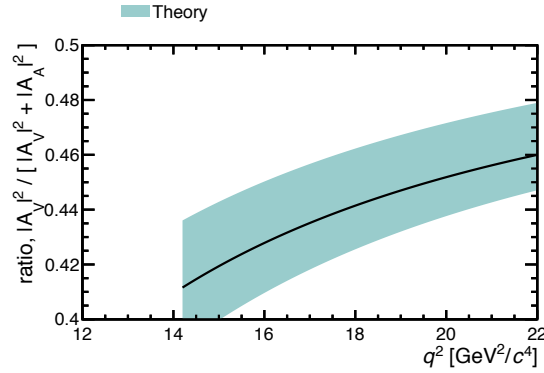


Figure 5.14: Vector fraction of the non-resonant $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay as a function of q^2 . Compared to the experimental sensitivity, the fraction is essentially constant with dimuon mass.

The tree-level decay, $B^+ \rightarrow K^+(\psi \rightarrow \mu^+ \mu^-)$, shares the same initial and final state as the non-resonant $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay. This means that the resonant and non-resonant components will interfere and this must be taken into account in the mass fit. The non-resonant $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay is composed of vector and axial-vector components, corresponding to the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$, respectively. Assuming the resonances are vector-like, only the vector part of the non-resonant amplitude will interfere. For this analysis, the vector fraction of the non-resonant part is assumed to be the SM value. This vector fraction is calculated using the EOS flavour tool described in Ref. [61], and is shown as a function of q^2 in Fig. 5.14. Given the small variation with dimuon mass, the non-resonant vector fraction can be safely assumed to be constant across the low recoil region, which leads to the following expression for the signal PDF,

$$\mathcal{P}_{\text{sig}} \propto P(m_{\mu^+ \mu^-}) |\mathcal{A}|^2 f^2(m_{\mu^+ \mu^-}^2), \quad (5.1)$$

$$|\mathcal{A}|^2 = |A_{\text{nr}}^{\text{V}} + \sum_k e^{i\delta_k} A_{\text{r}}^k|^2 + |A_{\text{nr}}^{\text{AV}}|^2, \quad (5.2)$$

where A_{nr}^{V} and $A_{\text{nr}}^{\text{AV}}$ are the vector and axial vector amplitudes of the non-resonant decay. The shape of the non-resonant signal in $m_{\mu^+ \mu^-}$ is driven by phase space, $P(m_{\mu^+ \mu^-})$ which is given by,

$$P(m_{\mu^+ \mu^-}) \propto \left(\frac{p}{m_{B^+}} \right) \left(\frac{q}{m_{\mu^+ \mu^-}} \right) m_{\mu^+ \mu^-} \quad (5.3)$$

where p is the momentum of the K^+ in the B^+ rest frame and q is the μ^- momentum in the $\mu^+ \mu^-$ rest frame. The dimuon mass dependence of the form factor, $f(m_{\mu^+ \mu^-}^2)$ is described using the parametrisation given in Ref. [65]. This form factor parametrisation is consistent with recent lattice calculations [35, 66]. The uncertainty on the form-factor parameterisation is taken from Ref. [61]. The total vector amplitude is formed by summing the vector amplitude of the non-resonant signal with a number of Breit-Wigner amplitudes, A_{r}^k . Each Breit-Wigner amplitude is rotated by a phase, δ_k , which represents the strong phase difference between the non-resonant vector component and the resonance with index k . Such phase differences are expected [65].

It is obviously not ideal to assume the SM when studying a decay sensitive to new physics. Unfortunately, it is not possible to allow the size of the non-resonant vector fraction and branching fraction of the resonance to vary simultaneously, as they are almost 100% correlated with each other. This is because the interference which has a very similar shape to the resonance itself, so that if the vector fraction is increased, it can be almost perfectly compensated by a corresponding reduction in the size of the resonance.

5.4. Results

For each mass fit, a component to describe the $\psi(3770)$ resonance is introduced, where the mass and width is constrained to the world average values [14]. The phase and size of the $\psi(3770)$ resonance are allowed to vary.

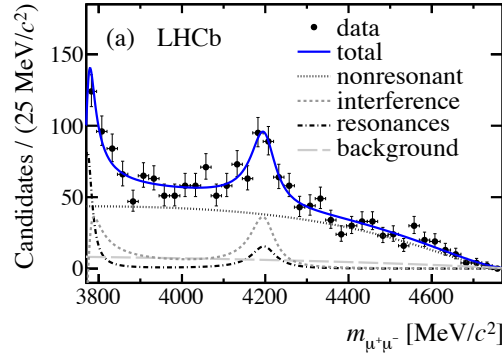


Figure 5.15: Dimuon mass distribution of data with fit results overlaid for the fit that includes contributions from the non-resonant vector and axial vector components, the $\psi(3770)$ meson, and one additional resonance with free width and mean parameters. Interference terms are included and the relative strong phases are left free in the fit.

Initially a fit with a single resonance in addition to the $\psi(3770)$ and non-resonant terms is performed, shown in Fig. 5.15. This additional resonance has its phase, mean and width left free. The parameters of the resonance returned by the fit are a mass of 4191_{-8}^{+9} MeV/ c^2 and a width of 65_{-16}^{+22} MeV/ c^2 . Branching fractions are determined by integrating the square of the Breit-Wigner amplitude returned by the fit, normalising to the $B^+ \rightarrow J/\psi K^+$ yield, and multiplying with the product of branching fractions, $\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$ [14]. The product $\mathcal{B}(B^+ \rightarrow X K^+) \times \mathcal{B}(X \rightarrow \mu^+ \mu^-)$ for the additional resonance, X, is determined to be $(3.9_{-0.6}^{+0.7}) \times 10^{-9}$. The uncertainty on this product is calculated using the profile likelihood. A compatible value of the branching fraction is found if the vector fraction of the non-resonant component is allowed to vary, but the error is inflated by more than a factor of four reflecting the greater freedom allowed in the interference between the resonance and the non-resonant component.

The significance of the resonance is obtained by simulating pseudo-experiments that include the non-resonant, $\psi(3770)$ and background components. The distribution of log likelihood ratios between fits that include and exclude a resonant component for 6×10^5 such samples is shown in Fig. 5.16. None of the samples have a higher ratio than observed in data and the tail of the distribution is fit with an exponential to extrapolate the distribution beyond the observed value gives a significance of the signal above six standard deviations.

To determine which resonance has been observed a 2D likelihood scan of the mass and width is performed, shown in Fig. 5.17. The data is consistent with the heavy $c\bar{c}$ resonance, $\psi(4160)$ as

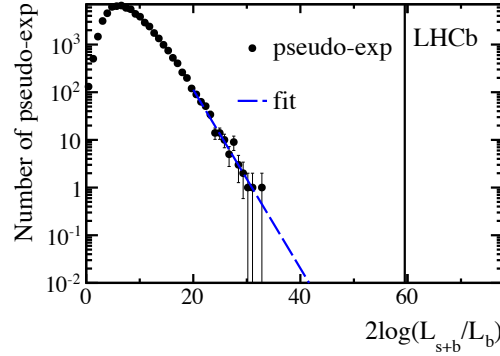


Figure 5.16: Distribution of the log likelihood ratio from fits that include and exclude a single additional resonance component for pseudo-experiments that only include the non-resonant, $\psi(3770)$ and background components. The dashed blue line corresponds to the fit of an exponential function which is used to extrapolate to larger values of $2\log(L_{s+b}/L_b)$. The vertical black line denotes the observed likelihood ratio. A signal significance exceeding six standard deviations is observed.

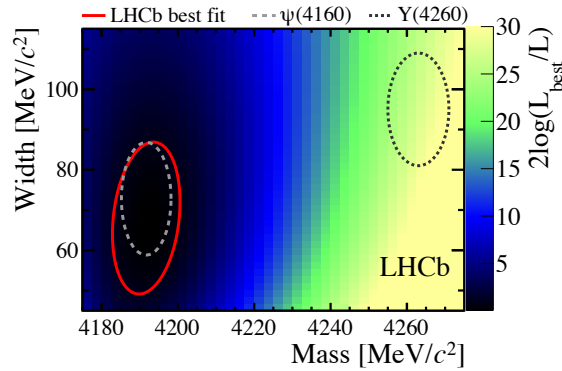


Figure 5.17: Profile likelihood as a function of mass and width of a fit with a single extra resonance. At each point all other fit parameters are re-optimised. The three ellipses are (red-solid) the best fit and previous measurements of (grey-dashed) the $\psi(4160)$ [55] and (black-dotted) the $Y(4260)$ [14] states.

measured in Ref. [55], but rejects the more exotic explanation of the $Y(4260)$ by over four standard deviations.

Table 5.3: Parameters of the dominant resonance for fits where the mass and width are unconstrained and constrained to those of the $\psi(4160)$ meson [55], respectively. The branching fractions are for the B^+ decay followed by the decay of the resonance to muons.

	Unconstrained	$\psi(4160)$
$\mathcal{B}[\times 10^{-9}]$	$3.9^{+0.7}_{-0.6}$	$3.5^{+0.9}_{-0.8}$
Mass [MeV/ c^2]	4191^{+9}_{-8}	4190 ± 5
Width [MeV/ c^2]	65^{+22}_{-16}	66 ± 12
Phase [rad]	-1.7 ± 0.3	-1.8 ± 0.3

To test the hypothesis that ψ resonances well above the open charm threshold are observed, another fit including the $\psi(4040)$ and $\psi(4160)$ resonances is performed. The mass and width of the two are constrained to the measurements from Ref. [55]. The data have no sensitivity to a $\psi(4415)$ contribution. The fit describes the data well and the parameters of the $\psi(4160)$ meson are almost unchanged with respect to the unconstrained fit. The fit overlaid on the data is shown in Fig. 5.18 and Table 5.3 reports the fit parameters.

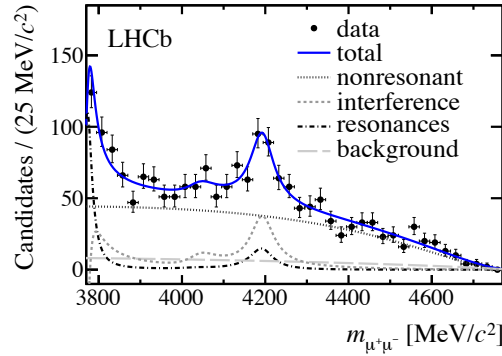


Figure 5.18: Dimuon mass distribution of data with fit results overlaid for the fit that includes contributions from the non-resonant vector and axial vector components, and the $\psi(3770)$, $\psi(4040)$ and $\psi(4160)$ resonances. Interference terms are included and the relative strong phases are left free in the fit.

The resulting profile likelihood ratio compared to the best fit as a function of branching fraction can be seen in Fig. 5.19. In the fit with the three ψ resonances, the $\psi(4160)$ meson is visible with $\mathcal{B}(B^+ \rightarrow \psi(4160)K^+) \times \mathcal{B}(\psi(4160) \rightarrow \mu^+\mu^-) = (3.5^{+0.9}_{-0.8}) \times 10^{-9}$ but for the $\psi(4040)$ meson, no significant signal is seen, and an upper limit is set. The limit $\mathcal{B}(B^+ \rightarrow \psi(4040)K^+) \times \mathcal{B}(\psi(4040) \rightarrow$

$\mu^+ \mu^-) < 1.3(1.5) \times 10^{-9}$ at 90(95)% confidence level is obtained by integrating the likelihood ratio compared to the best fit and assuming a flat prior for any positive branching fraction.

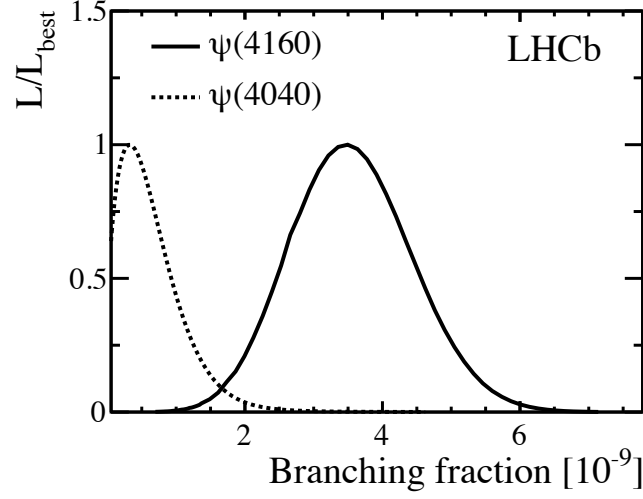


Figure 5.19: Profile likelihood ratios for the product of branching fractions $\mathcal{B}(B^+ \rightarrow \psi K^+) \times \mathcal{B}(\psi \rightarrow \mu^+ \mu^-)$ of the $\psi(4040)$ and the $\psi(4160)$ mesons. At each point all other fit parameters are reoptimised.

5.5. Summary

A resonance, compatible with the $\psi(4160)$ meson, has been observed in the low recoil region of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays. This is the first experimental evidence of tree-level contamination in electroweak penguins outside the J/ψ and $\psi(2S)$ regions and is also the first observation of the decays $B^+ \rightarrow \psi(4160)K^+$ and $\psi(4160) \rightarrow \mu^+ \mu^-$. The resonance contribution makes up a large portion (20%) of the signal yield in the low recoil region, which has highlighted the importance of controlling these effects for branching fraction measurements and angular analyses in this region.

The observed resonance spectrum is not the same as seen the $e^+ e^-$ cross-section ratio, shown in Fig. 5.1, where the $\psi(4040)$ has a bigger contribution than the $\psi(4160)$ resonance. This is evidence that the $\psi(4040)$ may be more exotic than first thought, although studies in different decay modes such as $B \rightarrow \psi(4040) \rightarrow D^{(*)} \bar{D}^{(*)} X$ and improved measurements of the leptonic width would be needed to confirm.

Another issue is the one already discussed in Sect. 5.1, which is that of QCD factorisation. All the results in this chapter assume QCDF, most clearly when the product of branching fractions are written, but also when the form factor is assumed to be the same for the non-resonant and resonant components in equation 5.2. Factorisation is known to break down to some extent, however given that the signal PDF fits the data well, the effect is appears to be small compared to the statistical sensitivity.

Looking ahead, if the branching fraction of the resonance was known, the vector fraction of the non-resonant component could be measured, which would be a novel search for new physics. It may be possible in the future to use the interference of the J/ψ , $\psi(2S)$ and perhaps even low mass resonances to determine the non-resonant vector fraction, which depends on the size of the Wilson Coefficient $\mathcal{C}_9$ relative to $\mathcal{C}_{10}$. This however requires a very good understanding of the dimuon resolution, as the J/ψ and $\psi(2S)$ are narrow. Such an analysis may be feasible, but not without significant effort.

Chapter 6.

Differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays

This chapter describes the measurement of the differential branching fraction for four $b \rightarrow s$ $\mu^+ \mu^-$ transitions; $B^+ \rightarrow K^+ \mu^+ \mu^-$, $B^0 \rightarrow K^0 \mu^+ \mu^-$, $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} \mu^+ \mu^-$. This analysis was performed entirely by the author and was published in Ref. [2].

6.1. Introduction

The rare decay of a B meson into a ground state or excited kaon, accompanied by a $\mu^+ \mu^-$ pair is a FCNC transition and as described in Chapter 2 is sensitive to physics beyond the SM via the influence of virtual particles. In particular, the branching fractions of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays are highly sensitive to the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$. The theoretical uncertainties in the decay rate suffer from relatively large uncertainties due to form factor estimates, although recent progress by lattice calculations [34, 35, 66, 67] has significantly improved the precision at high dimuon invariant mass squared (q^2). Nevertheless, to maximise sensitivity, observables can be constructed out of ratios or asymmetries where the leading form factor uncertainties cancel. The CP averaged isospin asymmetry (A_I) is such an observable and is defined as

$$\begin{aligned}
 A_I &= \frac{\Gamma(B^0 \rightarrow K^{(*)0}\mu^+\mu^-) - \Gamma(B^+ \rightarrow K^{(*)+}\mu^+\mu^-)}{\Gamma(B^0 \rightarrow K^{(*)0}\mu^+\mu^-) + \Gamma(B^+ \rightarrow K^{(*)+}\mu^+\mu^-)} \\
 &= \frac{\mathcal{B}(B^0 \rightarrow K^{(*)0}\mu^+\mu^-) - \frac{\tau_0}{\tau_+}\mathcal{B}(B^+ \rightarrow K^{(*)+}\mu^+\mu^-)}{\mathcal{B}(B^0 \rightarrow K^{(*)0}\mu^+\mu^-) + \frac{\tau_0}{\tau_+}\mathcal{B}(B^+ \rightarrow K^{(*)+}\mu^+\mu^-)},
 \end{aligned} \tag{6.1}$$

where $\Gamma(f)$ and $\mathcal{B}(f)$ are the partial width and branching fraction of the $B \rightarrow f$ decay and τ_0/τ_+ is the ratio of the lifetimes of the B^0 and B^+ mesons. In the SM, isospin asymmetries arise from diagrams where the spectator quark¹ radiates a photon, which carry a different amplitude due to the difference in electric charge between the up and down quarks. An example diagram for this is shown in Fig. 6.1, in which the short distance part corresponds to the Wilson Coefficient $\mathcal{C}_8$. Another contribution originates from a difference in the coupling between up and down quarks to the short distance operators. This occurs in annihilation diagrams, shown in Fig. 6.2, where a B^+ meson can annihilate at tree level through a W boson, whereas a B^0 meson cannot. These contributions are small, which is reflected in the SM prediction for A_I which is $\mathcal{O}(\%)$ in the q^2 region below the J/ψ resonance [68–70]. There is no precise prediction for A_I for the q^2 region above the J/ψ , but it is expected to be even smaller than the low q^2 case as the contribution from the photon, where isospin contributions propagate, is reduced.

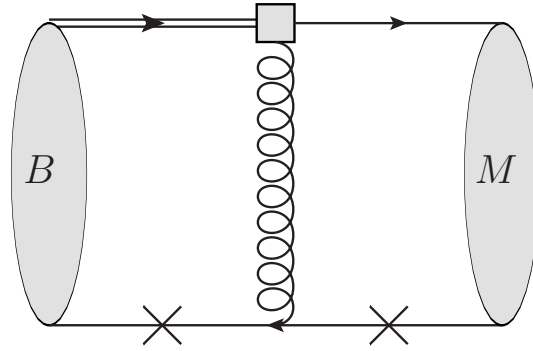


Figure 6.1: Diagram for $b \rightarrow s\mu^+\mu^-$ involving the Wilson Coefficient $\mathcal{C}_8$, taken from Ref. [68]. The difference in electric charge between the up quark and down quarks contributes to the isospin asymmetry. The crosses denote where a photon can be radiated to produce the dimuon pair.

Extensions to the SM can predict isospin asymmetries via the exchange of a particle which can change the flavour of the b quark at tree level. An example of such a transition is shown in

¹The term “spectator” refers to the up or down quark, even if it participates in the decay.

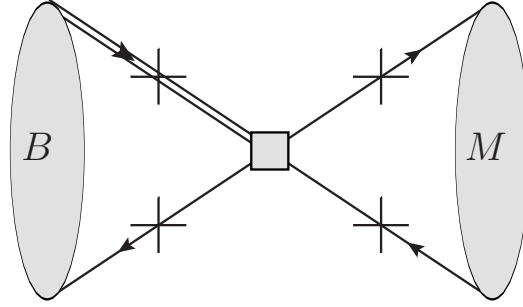


Figure 6.2: Diagram for $b \rightarrow s\mu^+\mu^-$ involving the Wilson Coefficients $\mathcal{C}_{1,2}$, taken from Ref. [68]. The crosses denote where a photon can be radiated to produce the dimuon pair.

Fig. 6.3, which is proposed in Ref. [71]. Like the SM contribution, sizeable isospin asymmetry is most likely at low q^2 , where the photon coupling is higher.

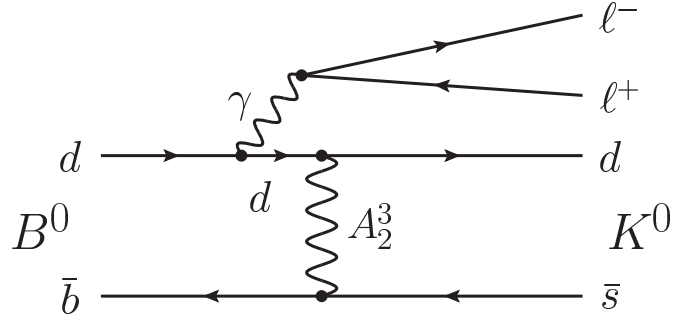


Figure 6.3: Diagram contributing to the isospin asymmetry which involves the new physics model proposed in Ref. [71]. The particle labelled as the A_2^3 is a gauge boson which can change flavours between the second and third quark generations.

Measurements of A_I have been previously made by the BaBar [72], Belle [73] and LHCb [5] collaborations, with results that tend to be negative. In particular, the $B \rightarrow K\mu^+\mu^-$ isospin asymmetry measured by the LHCb experiment with 1 fb^{-1} deviated by over 4σ from zero. In addition to this, the combined $B \rightarrow K\mu^+\mu^-$ and $B \rightarrow K^*\mu^+\mu^-$ isospin asymmetries measured by the BaBar collaboration deviated by 3.9σ [74], although this measurement has been since superseded by Ref. [72]. For $B \rightarrow K^*\mu^+\mu^-$, A_I is consistent with zero, and is expected to agree with the $B \rightarrow K^{*0}\gamma$ measurement of $5 \pm 3\%$ [75] as q^2 approaches zero. A summary of the most recent published results before this analysis is shown in Fig. 6.4.

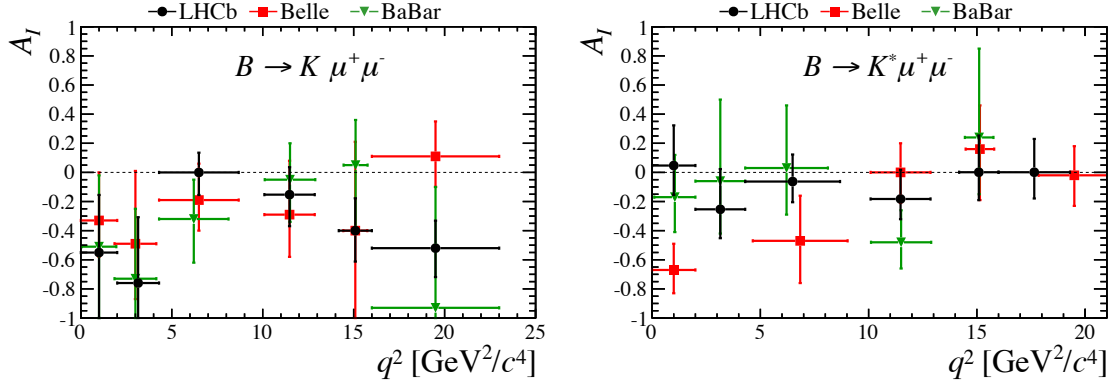


Figure 6.4: Most precise measurements of the isospin asymmetry prior to this analysis. The $B \rightarrow K\mu^+\mu^-$ results tend to be negative whereas the $B \rightarrow K^*\mu^+\mu^-$ are consistent with the SM prediction of no isospin asymmetry.

This chapter describes a measurement of the isospin asymmetry in $B \rightarrow K\mu^+\mu^-$ and $B \rightarrow K^*\mu^+\mu^-$ decays based on data corresponding to an integrated luminosity of 3 fb^{-1} taken with the LHCb detector at a centre-of-mass energy of 7 TeV in 2011 and 8 TeV in 2012. The isospin asymmetries are determined by measuring the differential branching fractions of $B^+ \rightarrow K^+\mu^+\mu^-$, $B^0 \rightarrow K^0\mu^+\mu^-$, $B^0 \rightarrow (K^{*0} \rightarrow K^+\pi^-)\mu^+\mu^-$ and $B^+ \rightarrow (K^{*+} \rightarrow K_s^0\pi^+)\mu^+\mu^-$; the decays involving a K_L^0 or π^0 are not considered².

6.2. Selection

6.2.1. Pre-selection

The stripping selection for the signal candidates is the same as described in chapter 5: the final state products originating directly from the B are required to have an impact parameter χ^2 , $\chi_{\text{IP}}^2 > 9$, where χ_{IP}^2 is defined as the minimum change in χ^2 of the vertex fit to any of the PVs in the event when the particle is added to that PV; the dimuon pair, as well as the $K^+\pi^-$ and $K_s^0\pi^-$ pairs which form the K^* candidates, have a vertex fit $\chi^2 < 9$. Additionally, an associated PV is chosen based on the lowest χ_{IP}^2 of the B candidate. For this PV it is required that: the B candidate has $\chi_{\text{IP}}^2 < 16$; the vertex fit χ^2 must increase by more than 121 when including the B candidate daughters; and the angle between the B candidate momentum and the direction from the PV to

²The $K^{*0} \rightarrow K^+\pi^-$ and $K^{*+} \rightarrow K_s^0\pi^+$ decays have a branching fraction of approximately two-thirds.

the decay vertex is below 14 mrad. Finally the B candidate is required to have a vertex fit $\chi^2 < 24$. For candidates containing a K_s^0 in the final state, the K_s^0 is required to have a lifetime larger than 2 ps. Candidate K_s^0 mesons are reconstructed in the $\pi^+\pi^-$ decay mode and are required to have a di-pion mass within 30 MeV/ c^2 of the nominal K_s^0 mass. Candidate K^* mesons are required to have a mass within 100 MeV/ c^2 of the nominal K^* mass.

6.2.2. Multivariate selection

Each signal mode is selected using the same strategy as described for the $B^+ \rightarrow K^+\mu^+\mu^-$ decay in Chapter. 5, where BDTs which contain kinematic and geometric information are trained on simulation as a signal sample and the extreme upper mass sideband as background. Separate BDTs are employed for each signal decay. Additionally, for decays involving a K_s^0 , two independent BDTs are trained for the LL and DD categories. The most difficult decay to select is the $B^0 \rightarrow K_s^0\mu^+\mu^-$ decay, as it is difficult to determine whether a K_s^0 meson originated from the secondary vertex (SV) rather than the PV due to its long lifetime. For example the BDT selection efficiency for $B^0 \rightarrow K_s^0\mu^+\mu^-$ is only 66 (48)% for the LL (DD) K_s^0 reconstruction category, whereas for $B^+ \rightarrow K^+\mu^+\mu^-$ it is 90%. There is some distinguishing power between the PV and SV for K_s^0 mesons in the LL category, which can be visually seen in Fig. 6.5, which shows the BDT response for the signal and background samples. The overlap between the signal and background responses in the LL category is smaller than the DD category.

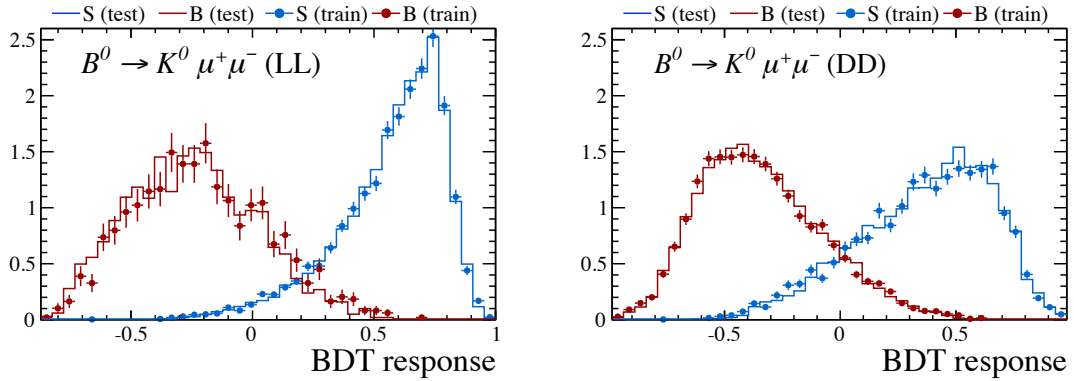


Figure 6.5: BDT response for the signal and background samples for the $B^0 \rightarrow K_s^0\mu^+\mu^-$ decay in the LL category (left) and DD category (right). The samples are split into training and testing samples, the latter of which is used to optimise the BDT selection cut.

The $B \rightarrow K^*\mu^+\mu^-$ channels do not suffer from as much background as the $B \rightarrow K\mu^+\mu^-$ modes and the subsequent BDT selections are all roughly 90% efficient.

6.3. Backgrounds

Due to the long lifetime of the K_s^0 meson, there are very few sources of exclusive background that can be mistakenly identified as $B^0 \rightarrow K_s^0\mu^+\mu^-$ or $B^+ \rightarrow K^{*+}\mu^+\mu^-$ decays. The largest fully reconstructed background is from $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays, where the proton from the $\Lambda \rightarrow p\pi^-$ decay is incorrectly identified as a π^+ . This background is removed by rejecting K_s^0 meson candidates if their $m(\pi^+\pi^-)$ mass is consistent, within $10 \text{ MeV}/c^2$ ($15 \text{ MeV}/c^2$) for LL (DD) candidates, with that of a Λ baryon. This veto is $\sim 95\%$ efficient on genuine K_s^0 meson decays and removes more than 99% of Λ baryons. Similarly to the $B^+ \rightarrow K^+\mu^+\mu^-$ case described in section 5.2.2, for $B^+ \rightarrow K^{*+}\mu^+\mu^-$ there is background from $B^+ \rightarrow J/\psi K^{*+}$ decays, where the pion and the same-sign muon have swapped identities. This background is rejected in the same way as $B^+ \rightarrow K^+\mu^+\mu^-$ and is reduced to a negligible level.

Of the four signal decays, $B^0 \rightarrow K^{*0}\mu^+\mu^-$ suffers from the largest array of peaking backgrounds. Most of the vetoes were first employed for the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ angular analysis and are described in detail in Ref. [76]. The only update to this is to reject peaking contribution from the decay $\Lambda_b^0 \rightarrow pK^+\mu^+\mu^-$, using requirements on pK invariant mass and applying selection to the PID of the π^+ candidate. All vetoes are about 99% efficient on signal and reduce residual peaking background to a negligible level.

The two-dimensional distribution of dimuon mass against $K^{(*)}\mu^+\mu^-$ mass is shown for the four signal decays in Fig. 6.6 after the full selection. The signal is visible as a vertical band centred at the B mass. Dimuon pairs originating from the J/ψ and $\psi(2S)$ resonances are also visible as horizontal bands, which are subsequently removed using the q^2 binning scheme.

6.4. Mass fits

The mass fit strategy is similar to the one described in Chapter 5, where the relevant $B \rightarrow J/\psi K^{(*)}$ decay is used as a signal shape proxy, parameterised by a sum of two Crystal Ball functions [53] with common tail parameters, but different widths. This shape is fit to the $B^0 \rightarrow J/\psi K_s^0$ data with all shape parameters allowed to vary and is shown in Fig. 6.7 for the LL and DD categories.

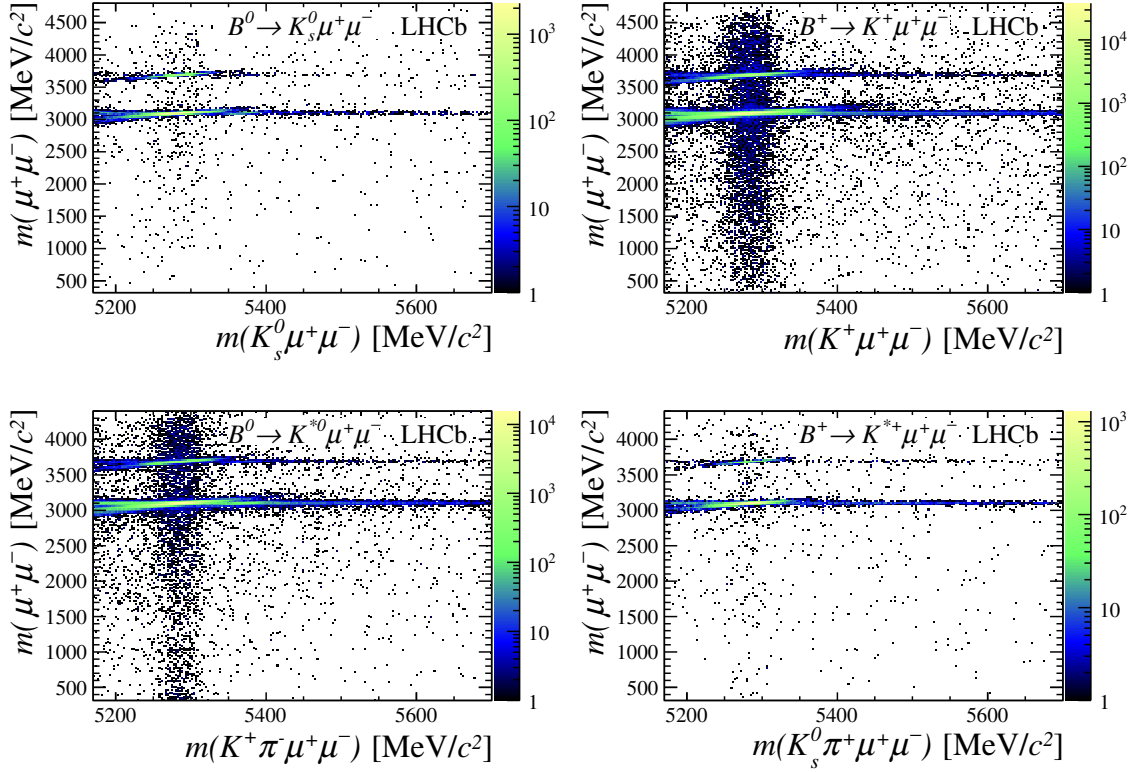


Figure 6.6: Dimuon mass versus $K^{(*)}\mu^+\mu^-$ mass distribution of the four signal decays.

The fit quality is good, which means the signal shape parameterisation is adequate to describe the signal. For the channels involving a K_s^0 , simultaneous fits are made to the LL and DD categories. The combinatorial background is parameterised by an exponential function, which is allowed to vary for each fit independently.

There is a small dependence on the mean and width parameters on q^2 , which is shown for $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ in Fig. 6.8 and is obtained by fitting simulation in bins of q^2 . The width increases with q^2 as the muons have higher momentum and hence worse momentum resolution. The increase in the mean of the distribution is currently not understood, however the effect is below the sensitivity for the analysis. The difference between in the mean and widths between signal in a bin of q^2 and the J/ψ bin is used to convert the J/ψ shape into the signal shape. As only the difference in mean and widths is taken from the simulation, the dependence on the absolute momentum scale is reduced. Ultimately, these corrections make very little difference to the signal yields, as shown in Fig. 6.9 for $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ as an example.

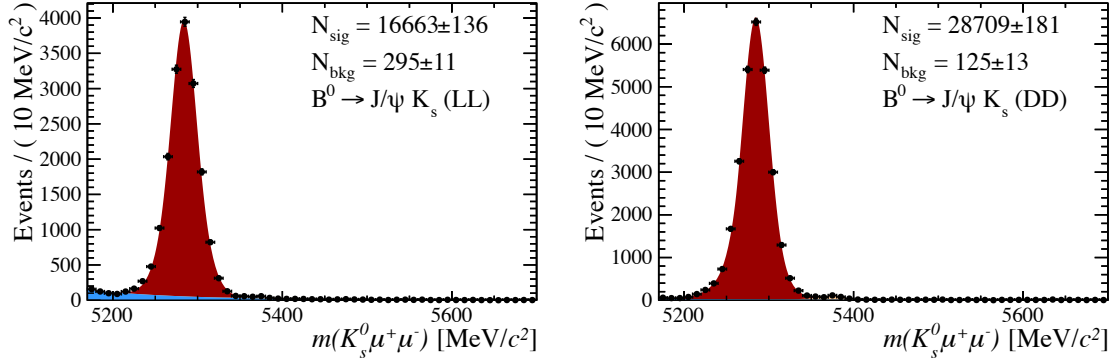


Figure 6.7: Mass fits to $B^0 \rightarrow J/\psi K_s^0$ decays in the LL category (left) and DD category (right). The fit model is described in the text.

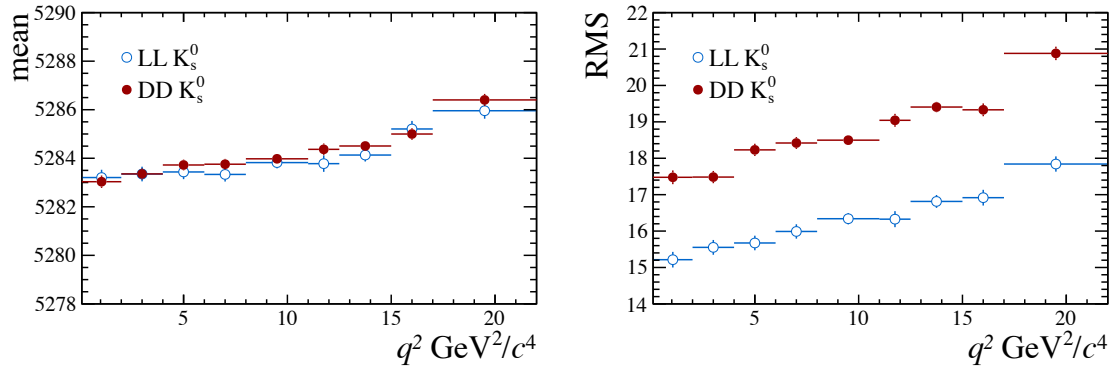


Figure 6.8: Mean and width of the $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ mass distribution in the simulation as a function of q^2 .

The mass fits for the four signal channels are shown in Fig. 6.10, where the LL and DD K_s^0 categories have been combined and the fits are performed in separate q^2 bins and subsequently merged. The corresponding number of signal candidates for each channel is given in Table 6.1. The decays involving a K_s^0 have typically 20-30 times less signal than the K^+ channels. This is mainly due to a lower visible branching fraction and a lower reconstruction efficiency for long lived particles.

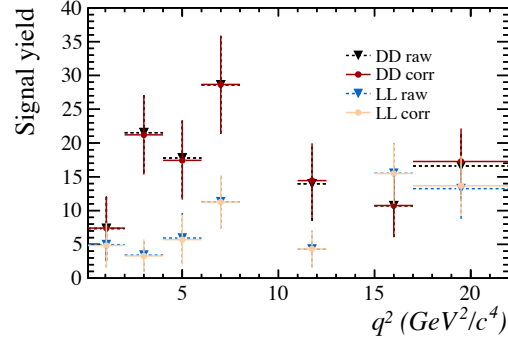


Figure 6.9: Signal yields for $B^0 \rightarrow K_S^0 \mu^+ \mu^-$ in the LL and DD categories before and after the correction to the changes in mean and width as a function of q^2 . The corrections to the mass shape make a very small difference to the signal yield.

Table 6.1: Observed yields of the four signal channels. See text for details.

Decay mode	Signal yield
$B^+ \rightarrow K^+ \mu^+ \mu^-$	4746 ± 81
$B^0 \rightarrow K_S^0 \mu^+ \mu^-$	176 ± 17
$B^+ \rightarrow K^{*+} (\rightarrow K_S^0 \pi^+) \mu^+ \mu^-$	162 ± 16
$B^0 \rightarrow K^{*0} (\rightarrow K^+ \pi^-) \mu^+ \mu^-$	2361 ± 56

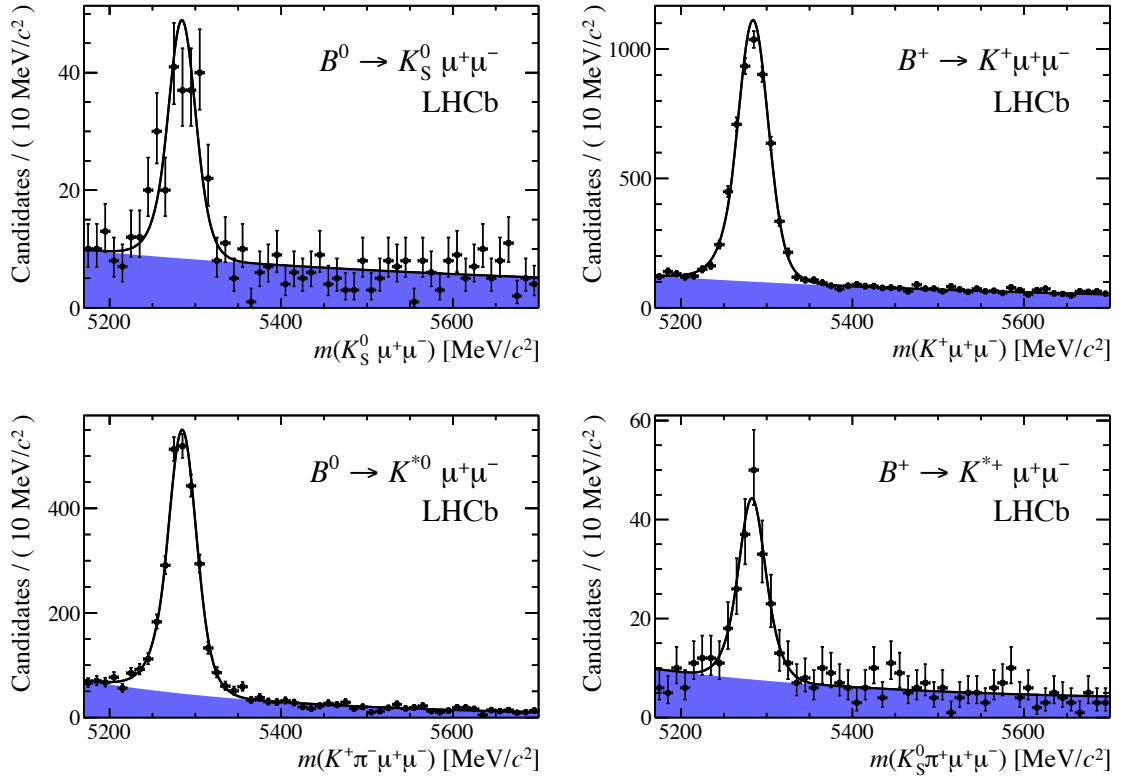


Figure 6.10: Reconstructed B candidate mass for the four signal modes. The data are overlaid with the result of the fit described in the text. The long and downstream K_S^0 categories have been combined. The fits are performed in separate q^2 bins and subsequently merged.

6.5. Efficiency as a function of q^2

Each signal mode is normalised to the $B \rightarrow J/\psi K^{(*)}$ channel, where the J/ψ decays into two muons. These decays have branching fractions which are approximately two orders of magnitude higher than those of the signal decays. Each normalisation channel has similar kinematics and the same final state particles as the signal modes, which simplifies the calculation of systematic uncertainties. There are separate normalisations for the long and downstream K_s^0 reconstruction categories to further cancel potential sources of systematic uncertainty.

Similarly to chapter 5, corrections to the IP resolution, PID and B kinematics are applied to the simulation, after which, the kinematic distributions of the normalisation channels agree with the data. The simulation samples are subsequently used to calculate relative efficiency as a function of q^2 , the shape of which is dominated by effects associated with the trigger, shown earlier in Fig. 5.6, where the muons have increased (decreased) p_T at high (low) q^2 and consequently have a higher (lower) trigger efficiency. For $B^0 \rightarrow K_s^0 \mu^+ \mu^-$, the shape is slightly steeper, as the trigger has a low efficiency on K_s^0 mesons and therefore relies more heavily than $B^+ \rightarrow K^+ \mu^+ \mu^-$ on the trigger efficiency of the muons. Similarly to the situation described in Sect. 5.3.3, at high q^2 , where the hadron will be almost at rest in the B frame and, like the B , will point back to the PV in the lab frame. This is expected to be well modelled in the simulation as discussed in Sect. 5.3.3.

The K_s^0 channels have an additional effect due to the acceptance of the two reconstruction categories. For example, K_s^0 mesons are more likely to be reconstructed in the long category if they have low momentum, which favours the high q^2 region. This effect is expected to be well reproduced as the K_s^0 momentum distributions are in good agreement between data and simulation, shown in Fig. 6.11 for $B^0 \rightarrow J/\psi K_s^0$ decays as an example.

The efficiency as a function of q^2 for all four signal channels is shown in Fig. 6.12, where the lower trigger efficiency for $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ can be seen at low q^2 , and K_s^0 acceptance between the LL and DD categories seen at high q^2 . The $B \rightarrow K^* \mu^+ \mu^-$ shape is flatter than the $B \rightarrow K \mu^+ \mu^-$ shape as the heavier K^* mass reduces the correlation between the hadron kinematics and q^2 . For example, the K^+ in $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decays is never at rest in the B frame.

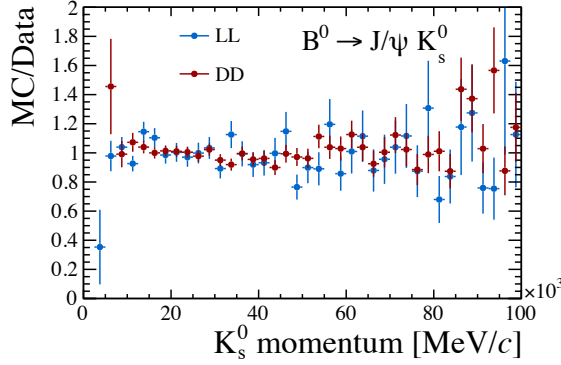


Figure 6.11: Ratio of the K_s^0 momentum distributions in $B^0 \rightarrow J/\psi K_s^0$ decays between data and simulation.

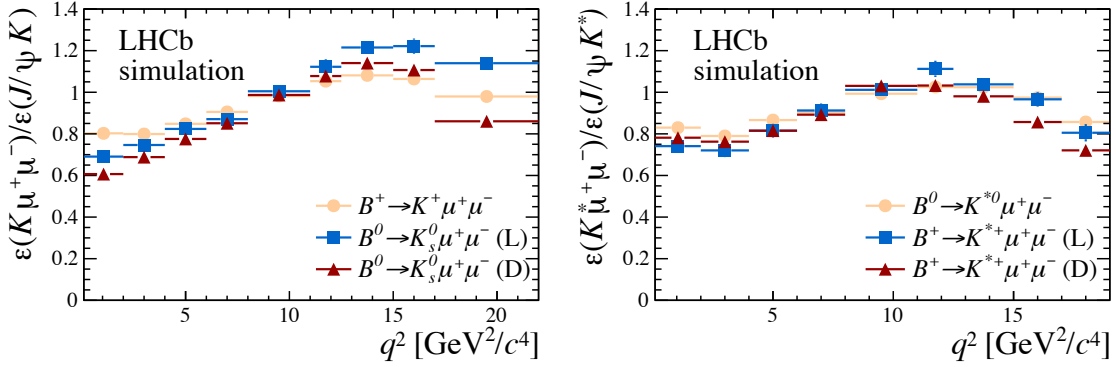


Figure 6.12: Relative efficiency between the signal and normalisation channels as a function of q^2 , obtained from the simulation.

6.6. Systematic uncertainties

The branching fraction measurements of the normalisation modes from the B -factory experiments assume that the B^+ and B^0 mesons are produced with equal proportions at the $\Upsilon(4S)$ resonance [77–79]. In contrast, in this paper isospin asymmetry is assumed between the $B \rightarrow J/\psi K^{(*)}$, implying that the $B^+ \rightarrow J/\psi K^+$ ($B^+ \rightarrow J/\psi K^{*+}$) and $B^0 \rightarrow J/\psi K^0$ ($B^0 \rightarrow J/\psi K^{*0}$) decays have the same partial width. The branching fractions used in the normalisation are obtained by: taking the most precise branching fraction results from Ref. [77] and translating them into partial widths; averaging the partial widths of the K^+ , K^0 and the K^{*+} , K^{*0} modes, respectively; and finally translating the widths back to branching fractions. The calculation only requires knowledge of the ratio of B^0 and B^+ lifetimes for which we use 0.93 ± 0.01 [14]. Statistical uncertainties are treated

as uncorrelated while systematical uncertainties are conservatively treated as fully correlated. The resulting branching fractions of the normalisation channels are

$$\begin{aligned}\mathcal{B}(B^+ \rightarrow J/\psi K^+) &= (0.998 \pm 0.014 \pm 0.040) \times 10^{-3}, \\ \mathcal{B}(B^0 \rightarrow J/\psi K^0) &= (0.928 \pm 0.013 \pm 0.037) \times 10^{-3}, \\ \mathcal{B}(B^+ \rightarrow J/\psi K^{*+}) &= (1.431 \pm 0.027 \pm 0.090) \times 10^{-3}, \\ \mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) &= (1.331 \pm 0.025 \pm 0.084) \times 10^{-3},\end{aligned}$$

where the first uncertainty is statistical and the second systematic. The uncertainty on the branching fractions of the normalisation modes constitutes the dominant source of systematic uncertainty on the branching fraction measurements presented for this analysis while it cancels in the isospin measurements.

The remaining systematic uncertainties are very small compared to the statistical uncertainties. A systematic uncertainty is assigned to account for the imperfect knowledge of the q^2 spectrum in the simulation. For example the recent observation of a resonance in the high q^2 region of $B^+ \rightarrow K^+\mu^+\mu^-$ decays [1], could alter the selection efficiencies in that region. This uncertainty is found to be 1 – 2% depending on channel and q^2 bin.

As described in chapter 4, data-driven corrections of the long and downstream tracking efficiencies in the simulation are determined using tag-and-probe techniques in $J/\psi \rightarrow \mu^+\mu^-$ and $D^0 \rightarrow \phi K_s^0$ decays respectively. Due to the robust nature of the normalisation strategy, the corresponding correction factors are negligible and no systematic uncertainty is considered. The systematic uncertainty that arises from all remaining corrections to the simulation varies between 1% and 3% depending on channel and q^2 bin.

The presence of an S-wave contribution to the $K^+\pi^-$ and $K_s^0\pi^+$ systems of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^+ \rightarrow K^{*+}\mu^+\mu^-$ decays respectively, complicates the analysis of these channels. This effect is of the order of a few percent and can be safely neglected in $B^+ \rightarrow K^{*+}\mu^+\mu^-$ decays due to the current statistical precision. The larger signal yield of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ however merits a systematic uncertainty of 7% designed to cover this effect, which corresponds to the S-wave fraction measured in $B^0 \rightarrow J/\psi K^{*0}$ decays. The S-wave fraction can in principle vary across q^2 [80], however to assess this effect a dedicated analysis of the S-wave contribution is required, which will be the focus of a future LHCb publication.

6.7. Branching fraction results

The averaged differential branching fraction over a q^2 bin of width $q_{\min}^2 - q_{\max}^2$, can be written as

$$\frac{d\mathcal{B}}{dq^2} = \frac{N(B \rightarrow K^{(*)}\mu^+\mu^-)}{N(B \rightarrow J/\psi K^{(*)})} \cdot \frac{\varepsilon(B \rightarrow J/\psi K^{(*)})}{\varepsilon(B \rightarrow K^{(*)}\mu^+\mu^-)} \cdot \frac{\mathcal{B}(B \rightarrow J/\psi K^{(*)})\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)}{(q_{\max}^2 - q_{\min}^2)}, \quad (6.2)$$

where $N(B \rightarrow K^{(*)}\mu^+\mu^-)$ is the number of signal candidates in the bin, $N(B \rightarrow J/\psi K^{(*)})$ is the number of normalisation candidates, the product of $\mathcal{B}(B \rightarrow J/\psi K^{(*)})$ and $\mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)$ is the visible branching fraction of the normalisation channel [14], $\varepsilon(B \rightarrow K^{(*)}\mu^+\mu^-)/\varepsilon(B \rightarrow J/\psi K^{(*)})$ is the relative efficiency between the signal in the bin and the normalisation channel.

The differential branching fraction for the K_s^0 channels is determined by simultaneously fitting the long and downstream categories of the signal channels. The branching fraction in bins of q^2 for $B^+ \rightarrow K^+\mu^+\mu^-$ and $B^0 \rightarrow K^0\mu^+\mu^-$ decays are shown in Fig. 6.13 and given in Tables A.2 and A.1 in the appendix. Theoretical predictions [61, 81] are superimposed in green on Figs. 6.13 and 6.14. In the low q^2 region, these predictions rely on the QCD factorisation approaches from Refs. [82, 83] for $B \rightarrow K^*\mu^+\mu^-$ and Ref. [84] for $B \rightarrow K\mu^+\mu^-$ which lose accuracy when approaching the J/ψ resonance. In the high q^2 region, an operator product expansion in the inverse b -quark mass, $1/m_b$, and in $1/\sqrt{q^2}$ is used based on Ref. [33]. This expansion is only valid above the open charm threshold. A dimensional estimate is made of the uncertainty from expansion corrections [85]. For the predictions in green, the form factor calculations for $B \rightarrow K^*\mu^+\mu^-$ and $B \rightarrow K\mu^+\mu^-$ are taken from Refs. [86] and [65] respectively. Predictions based on form factor estimates from lattice calculations are also overlaid, shown in yellow [87]. These form factors lead to a high correlation in the uncertainty of the predictions across q^2 .

The sample size for $B^+ \rightarrow K^+\mu^+\mu^-$ is large enough to see structure in the data. The structure at high q^2 is due to the $\psi(4160)$ resonance, which was the subject of chapter 5. The origin of the structure at low q^2 is currently unknown, although the ρ and ω resonances can contribute there. The branching fraction of $B^+ \rightarrow K^{*+}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays is shown in Fig. 6.14 and given in Table A.3 in the appendix.

The branching fractions integrated over q^2 are obtained by extrapolating to the full q^2 range assuming the distribution based on Ref. [29]. The correction factors to the branching fractions due to this extrapolation are 1.39 and 1.50 for $B \rightarrow K\mu^+\mu^-$ and $B \rightarrow K^*\mu^+\mu^-$, respectively.

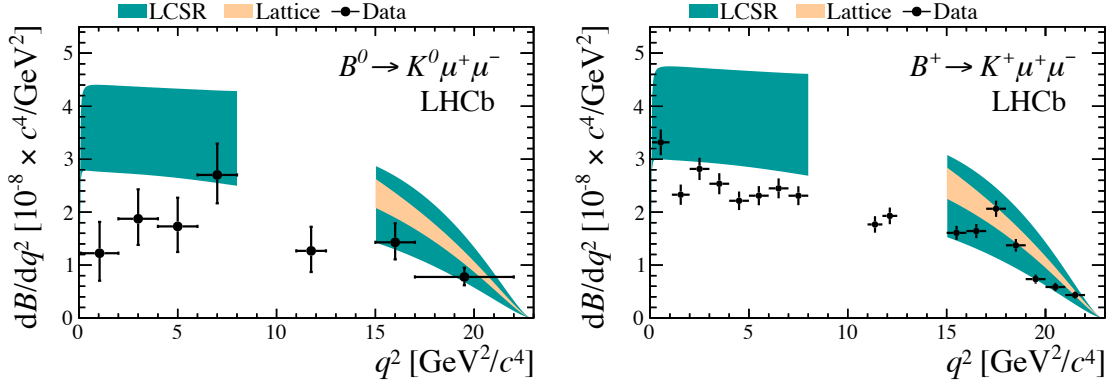


Figure 6.13: Differential branching fraction for $B^0 \rightarrow K^0 \mu^+ \mu^-$ (left) and $B^+ \rightarrow K^+ \mu^+ \mu^-$ (right) decays. The shaded region illustrates the theoretical prediction with its uncertainty (see text for details).

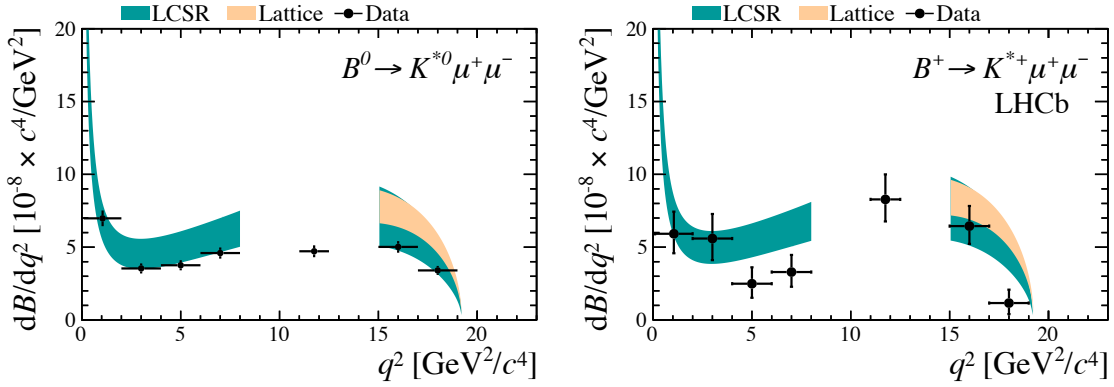


Figure 6.14: Differential branching fraction for $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ (left) and $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ (right) decays. The shaded regions illustrate the theoretical prediction with its uncertainty (see text for details).

No uncertainty is assigned to these corrections. Summing the q^2 bins and correcting for the extrapolation, the integrated branching fractions become

$$\begin{aligned}\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-) &= (4.29 \pm 0.07 \text{ (stat)} \pm 0.21 \text{ (syst)}) \times 10^{-7}, \\ \mathcal{B}(B^0 \rightarrow K^0 \mu^+ \mu^-) &= (3.27 \pm 0.34 \text{ (stat)} \pm 0.17 \text{ (syst)}) \times 10^{-7}, \\ \mathcal{B}(B^+ \rightarrow K^{*+} \mu^+ \mu^-) &= (9.24 \pm 0.93 \text{ (stat)} \pm 0.67 \text{ (syst)}) \times 10^{-7}, \\ \mathcal{B}(B^0 \rightarrow K^{*0} \mu^+ \mu^-) &= (9.14 \pm 0.20 \text{ (stat)} {}^{+0.55}_{-0.83} \text{ (syst)}) \times 10^{-7}.\end{aligned}$$

These measurements are more precise than the current world averages [14].

The $B \rightarrow K^{(*)}\mu^+\mu^-$ branching fractions integrated over the low recoil region are summarised in Tab. 6.2. Each measurement is compared to predictions obtained from LQCD. All branching fraction measurements are below the SM prediction.

Table 6.2: Integrated branching fractions in the low recoil region. For $B \rightarrow K^*\mu^+\mu^-$ this is defined as 15-19 GeV²/c⁴ whereas for $B \rightarrow K\mu^+\mu^-$ it is 15-22 GeV²/c⁴. Predictions are obtained using the form factors calculated in LQCD as described in Refs. [34, 35, 66, 67].

Branching fraction ($\times 10^{-8}$)			
Decay mode	Measurement	Prediction	Significance from SM
$B^+ \rightarrow K^+\mu^+\mu^-$	$8.5 \pm 0.3 \text{ (stat)} \pm 0.4 \text{ (syst)}$	10.7 ± 1.2	1.7
$B^0 \rightarrow K^0\mu^+\mu^-$	$6.7 \pm 1.1 \text{ (stat)} \pm 0.4 \text{ (syst)}$	9.8 ± 1.0	2.0
$B^+ \rightarrow K^{*+}\mu^+\mu^-$	$15.8^{+3.2}_{-2.9} \text{ (stat)} \pm 1.1 \text{ (syst)}$	26.8 ± 3.6	2.2
$B^0 \rightarrow K^{*0}\mu^+\mu^-$	$16.5 \pm 0.8 \text{ (stat)}^{+1.0}_{-1.5} \text{ (syst)}$	24.7 ± 3.2	2.4

6.8. Isospin asymmetry results

The assumption of no isospin asymmetry in the $B \rightarrow J/\psi K^{(*)}$ modes is equivalent to measuring the difference in isospin asymmetry between the rare $B \rightarrow K^{(*)}\mu^+\mu^-$ and $B \rightarrow J/\psi K^{(*)}$ decays. Compared to using the PDG values for the branching fractions of the $B \rightarrow J/\psi K^{(*)}$ modes this approach shifts A_I in each bin by approximately 4%. The isospin asymmetries are shown in Fig. 6.15 for $B \rightarrow K\mu^+\mu^-$ and $B \rightarrow K^*\mu^+\mu^-$ and given in Tables B.1 and B.2 in the appendix.

Since there is no knowledge on the shape of A_I outside the SM, the $A_I = 0$ hypothesis is tested against the simplest alternative, which is a constant value different from zero. The difference in χ^2 between the two hypotheses is used as a test statistic and is compared to an ensemble of pseudo-experiments which are generated with zero isospin asymmetry. As discussed in Sect. 6.1, in the SM A_I is predicted to be $\mathcal{O}(1\%)$ [68–70], which means the hypothesis of $A_I = 0$ is a good approximation to the SM given the current statistical precision. The p -value for the $B \rightarrow K\mu^+\mu^-$ isospin asymmetry with respect to zero is 11%, corresponding to a significance of 1.5σ . The test statistic used here differs from that of Ref. [5] which simply used the product of the p -values of individual q^2 bins with respect to zero to determine the total significance of the result.

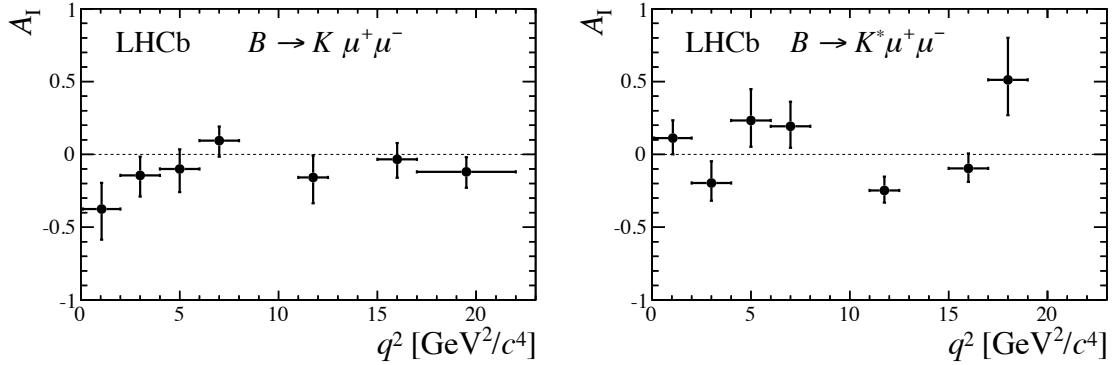


Figure 6.15: Isospin asymmetries for $B \rightarrow K \mu^+ \mu^-$ (left) and $B \rightarrow K^* \mu^+ \mu^-$ (right) decays.

Although the isospin asymmetry for $B \rightarrow K \mu^+ \mu^-$ decays is negative in all but one q^2 bin, results are more consistent with the SM compared to the previous measurement in Ref. [5] which quoted a 4.4σ significance to differ from zero. This reduction from 4.4σ to 1.5σ is due to four effects: the change of the test statistic in the calculation of the significance itself, which reduces the previous result to 3.5σ ; the assumption that the isospin asymmetry of $B \rightarrow J/\psi K^{(*)}$ is zero which reduces the significance further to 3.2σ ; a re-analysis of the 2011 data with the updated reconstruction and event selection that reduces the significance to 2.5σ ; and finally the inclusion of the 2012 data set reduces the significance further to 1.5σ . The measurements of A_I in the individual q^2 bins obtained from the re-analysis of the 2011 data set are compatible with those obtained in the previous analysis, the p -value of a χ^2 test being 93%.

6.9. Cross-checks

This section describes several cross-checks which are performed to assess the robustness of the isospin asymmetry results. Particular attention is paid towards the $B \rightarrow K \mu^+ \mu^-$ results.

6.9.1. Compatibility with previous result

The isospin asymmetry from Ref. [5], which is based on an analysis of LHCb data from 2011 alone, can be compared to the 2011 part of the new result to test the compatibility. To do this, the overlap of events in the signal region is calculated for each q^2 bin and is assumed to be the

correlation between each pair of measurements. Although the q^2 binning has changed, the nearest bin is used to calculate the compatibility. The results are overlaid on Fig. 6.16, where the χ^2 probability, taking into account the overlap of events, is 93%. For consistency, both analyses use the normalisation method as was used for Ref. [5], which is take the $B \rightarrow J/\psi K^{(*)}$ branching fractions from Ref. [14].

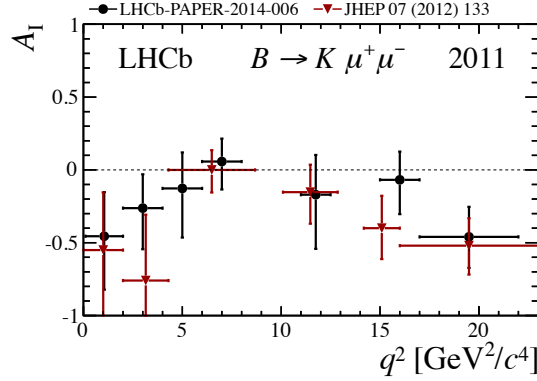


Figure 6.16: Isospin asymmetry of the old 2011 analysis and the 2011 part of the new analysis. The compatibility between the two is very good (93%).

6.9.2. Stability by run period and reconstruction category

The isospin asymmetry results are shown for the DD and LL categories separately. The χ^2 agreement between the two categories is good, with a p-value of 15%.

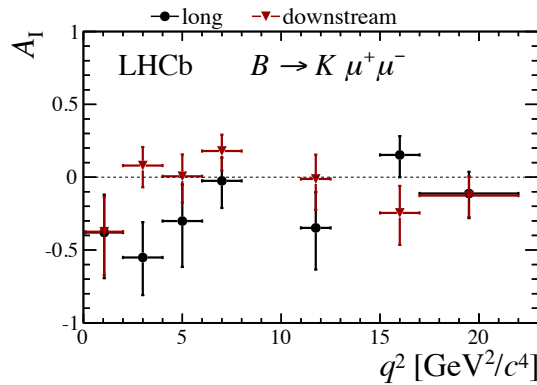


Figure 6.17: Isospin asymmetry results as measured LL and DD categories.

The isospin asymmetry results for the 2011 and 2012 datasets are shown in Fig. 6.18. For each bin, the 2012 dataset yields more positive central values than the 2011 dataset.

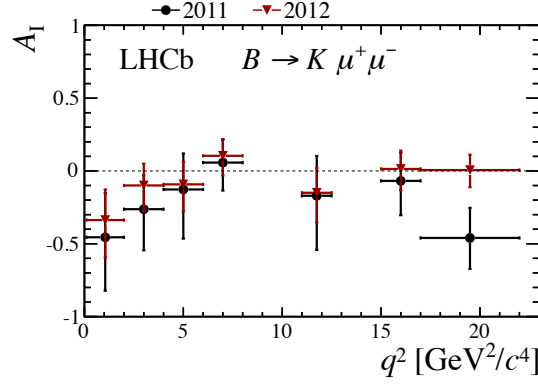


Figure 6.18: Isospin asymmetry for the 2011 and 2012 datasets. The 2012 results are consistently above the 2011 ones.

The data is also split into the K_s^0 reconstruction categories and then compared between 2011 and 2012 for the categories separately. This is shown in Fig. 6.19, where there is no evidence for a systematic shift. For the LL category there are five bins in a row where the 2012 result is higher but that is not significant compared to the seven in the combination. The DD category looks perfectly compatible between the 2011 and 2012 datasets. Given that neither category shows a clear trend, this is evidence that the systematic shift of all seven bins for the combination is indeed a statistical fluctuation as the two categories are essentially independent measurements.

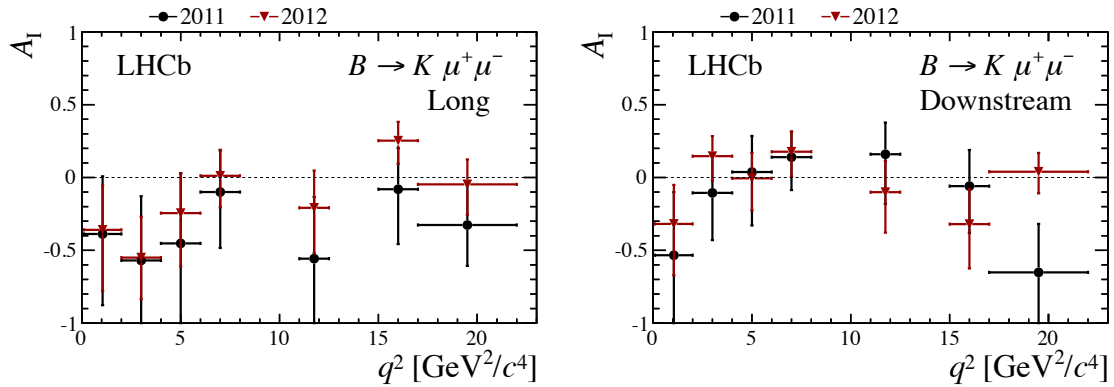


Figure 6.19: Isospin asymmetry for the 2011 and 2012 datasets for the LL (left) and DD (right) K_s^0 reconstruction categories.

6.9.3. Stability with mass range

The stability of the isospin result with different mass ranges is shown in Fig. 6.20. There is no significant change which implies that the background model is a good approximation to the data.

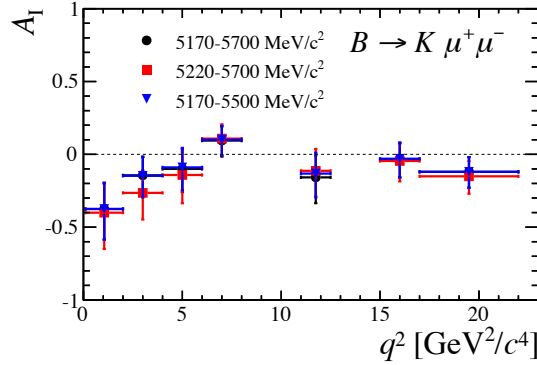


Figure 6.20: Isospin asymmetry under different mass ranges (5170-5700 MeV/ c^2 is nominal).

6.9.4. Restriction of trigger to muons

The isospin asymmetry results are compared when restricting the trigger decision to the dimuon candidate, leaving it independent of the kaon. This ensures that the trigger efficiency between $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ cancel as they both trigger on the muons. The comparison is shown in Fig. 6.21 and only a visible effect is seen at low q^2 as expected as that is where the K_s^0 and K^+ candidates participate the most. It is however a small effect and no systematic trend is observed.

6.9.5. $B^0 \rightarrow \psi(2S) K_s^0 / B \rightarrow J/\psi K_s^0$ branching fraction ratio

The ratio of branching fractions between $B^0 \rightarrow \psi(2S) K_s^0$ and $B^0 \rightarrow J/\psi K_s^0$ decays can be measured in the different run periods and K_s^0 reconstruction categories. The relative efficiency between the two is calculated using the measurements shown in Fig. 6.12 in Sect. 6.5. The results are given in Table. 6.3, where all four branching fractions agree with each other, the 2011 LL case is slightly high, but if this was a systematic effect, it would be in the opposite direction to the non-resonant case, where the LL 2011 category has a negative isospin asymmetry.

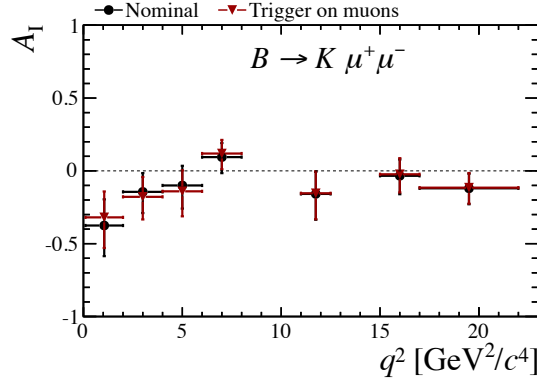


Figure 6.21: Isospin asymmetry results when applying the trigger requirements on the dimuon candidate.

Table 6.3: Ratio of branching fractions of $B^0 \rightarrow \psi(2S) K_s^0$ to $B^0 \rightarrow J/\psi K_s^0$ decays, where the $J/\psi \rightarrow \mu^+ \mu^-$ and $\psi(2S) \rightarrow \mu^+ \mu^-$ branching fractions are not taken into account.

Category	Branching fraction ratio
2011 LL	0.083 ± 0.004
2012 LL	0.075 ± 0.003
2011 DD	0.077 ± 0.003
2012 DD	0.076 ± 0.002

6.10. Summary

The most precise measurements of the differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays have been performed using a dataset corresponding to 3 fb^{-1} of integrated luminosity collected by the LHCb detector. The isospin asymmetries of the $B \rightarrow K \mu^+ \mu^-$ and $B \rightarrow K^* \mu^+ \mu^-$ decays are both consistent with SM expectations. However, the branching fractions of the four decay modes all lie below SM predictions.

Chapter 7.

Angular analysis of charged and neutral $B \rightarrow K \mu^+ \mu^-$ decays

In this chapter, an angular analysis of $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays is described. The analysis was performed entirely by the author and was published in Ref. [3].

7.1. Introduction

In addition to the branching fraction, diagrams contributing to the $B \rightarrow K \mu^+ \mu^-$ decay amplitude can influence the angular distribution of the $\mu^+ \mu^-$ pair. Similarly to the isospin asymmetry, angular observables are relatively free from hadronic uncertainties and are sensitive to different combinations of Wilson Coefficients to rate observables. The differential decay rate of the B^+ (B^-) decay as a function of $\cos \theta_l$, the angle between the μ^- (μ^+) and the K^+ (K^-) in the rest frame of the dimuon system can be written [29, 84] as

$$\frac{1}{\Gamma} \frac{d\Gamma}{d \cos \theta_l} = \frac{3}{4}(1 - F_H)(1 - \cos^2 \theta_l) + \frac{1}{2}F_H + A_{FB} \cos \theta_l \quad , \quad (7.1)$$

where θ_l is the angle between the direction of the μ^- (μ^+) lepton and the K^+ (K^-) meson. The differential decay rate depends on two parameters, the forward-backward asymmetry of the dilepton system, A_{FB} , and a “flat” parameter F_H . The width, Γ , A_{FB} and F_H depend on the dimuon invariant mass squared (q^2).

Since the B^0 and $\bar{B}^0$ meson can decay to the same $K_s^0\mu^+\mu^-$ final-state, it is not possible to determine the flavour of the B meson from the decay products. Without tagging the flavour of the neutral B meson at production¹, it is therefore not possible to unambiguously choose the correct muon to determine θ_l . In this situation any visible A_{FB} would indicate that there is either a large difference in the number of B^0 and $\bar{B}^0$ mesons produced, large CP violation in the decay or that the A_{FB} of the B^0 and $\bar{B}^0$ decay differ. Any residual asymmetry can be cancelled by performing the analysis in terms of $|\cos\theta_l|$,

$$\frac{1}{\Gamma} \frac{d\Gamma}{d|\cos\theta_l|} = \frac{3}{2}(1 - F_{\text{H}})(1 - |\cos\theta_l|^2) + F_{\text{H}}. \quad (7.2)$$

In contrast to the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$, A_{FB} is zero up-to tiny corrections in the SM. A sizable value of A_{FB} is possible in models that introduce large (pseudo)scalar- or tensor-like couplings [84,88]. The so called “flat” parameter, F_{H} , is also close to zero in the SM, but similarly can be enhanced in new physics models with (pseudo-) scalar or tensor like couplings.

The angular distribution of $B^+ \rightarrow K^+\mu^+\mu^-$ decays has previously been studied by BaBar [89], Belle [90], CDF [91] and LHCb [60], where all measurements are consistent with the SM. The LHCb results are shown in Fig. 7.1, which already dominated the world precision for A_{FB} and F_{H} despite only using a third of the dataset used in this analysis.

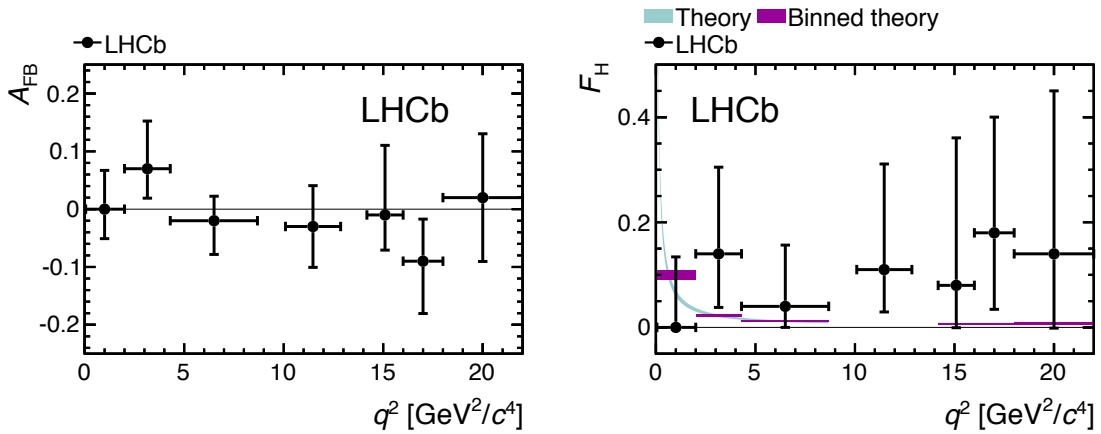


Figure 7.1: Most precise measurements of the angular observables of $B^+ \rightarrow K^+\mu^+\mu^-$ decays prior to the analysis described in this chapter, taken from Ref. [60].

¹The effective tagging power at LHCb is about 3% which prevents a time dependent tagged analysis.

This chapter describes an update to the angular analysis of $B^+ \rightarrow K^+\mu^+\mu^-$ decays, and the first angular analysis of $B^0 \rightarrow K^0\mu^+\mu^-$ decays, using a combined dataset corresponding to 3 fb^{-1} of integrated luminosity. The angular distributions of $B^+ \rightarrow K^+\mu^+\mu^-$ and $B^0 \rightarrow K^0\mu^+\mu^-$ are expected to be identical within the experimental precision. Despite the small impact $B^0 \rightarrow K^0\mu^+\mu^-$ would have on a combination of the two, its inclusion is interesting since a non-zero isospin asymmetry as discussed in the previous chapter, could lead to a different angular distribution for the charged and neutral decay. The selection and peaking backgrounds for this analysis are the same as described as in Chapters. 5 and 6. Correcting for the detector acceptance as a function of $\cos\theta_l$ is crucial to avoid a bias to A_{FB} and F_{H} and is described in Sect. 7.2. The main challenge for this analysis is dealing with the mathematical boundary that Eq. 7.1 imposes, where the PDF becomes negative if $2|A_{\text{FB}}| > F_{\text{H}}$. This mathematical boundary is shown in Fig. 7.2, where the likelihood surface for an example q^2 bin is shown. This leads to complications in the angular fit, described in Sect. 7.3, where care must be taken when determining central values and statistical uncertainties of the observables. Systematic uncertainties are described in Sect. 7.5, however the most crucial part of this analysis is estimating the statistical uncertainties correctly, as the size of the systematic uncertainties are small.

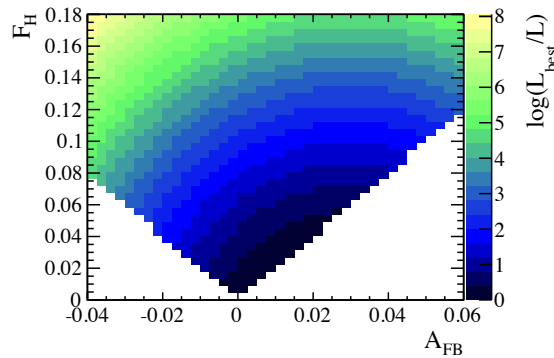


Figure 7.2: Typical likelihood surface for the angular observables A_{FB} and F_{H} for the $B^+ \rightarrow K^+\mu^+\mu^-$ mode. The PDF becomes negative in the white region.

7.2. Efficiency as a function of $\cos\theta_l$

Equations 7.1 and 7.2 describe only the physics of the signal decays. In reality the situation is complicated by the acceptance of the detector, where decays in a certain region of $\cos\theta_l$ are more likely to be reconstructed, triggered and selected than another region. This effect is included in

the analysis by multiplying equations 7.1 and 7.2 by the acceptance as a function of $\cos\theta_l$, which is calculated using the simulation. The efficiency is assumed to be symmetric around $\cos\theta_l = 0$, which implies that it is independent of the relative charge between the muon and kaon used to measure θ_l . This is a good assumption as charge asymmetries are at the % level in LHCb, and are not correlated between particles.

The simulation is corrected using the same methods as described in chapter 6. These corrections make little difference to the efficiency shape as a function of $\cos\theta_l$, as most badly modelled variables are uncorrelated to $\cos\theta_l$. An example is the PID efficiency, shown in Fig. 7.3 for $B^+ \rightarrow K^+\mu^+\mu^-$ decays. Whether the PID efficiency is accounted for in the fit is irrelevant to the analysis, as it is essentially flat in $\cos\theta_l$.

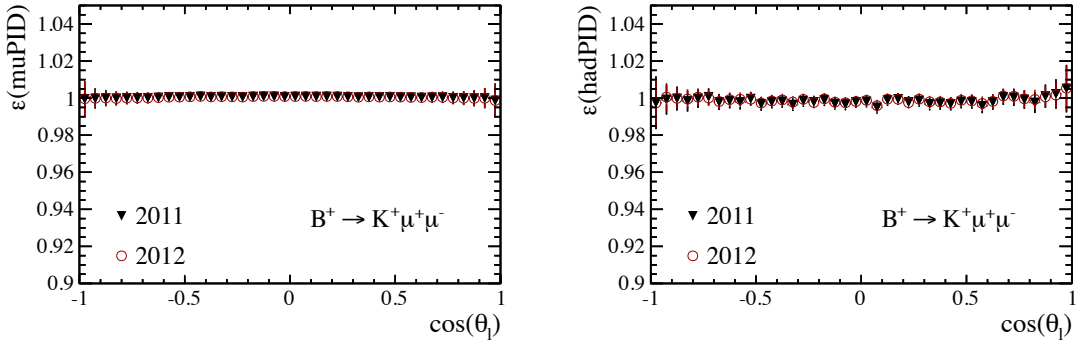


Figure 7.3: Efficiency of the muon (left) and hadron (right) PID selection criteria as a function of $\cos\theta_l$. The shape can be considered a flat line when considering the statistical precision of the data.

As described in Sect. 5.2.2, there is a peaking background veto which removes candidates with a $\mu^- K^+$ mass consistent with originating from a $\bar{D}^0 \rightarrow K^+ \pi^-$ decay. As $B^+ \rightarrow K^+\mu^+\mu^-$ is a three body decay, at a given q^2 mass, the $\mu^- K^+$ mass has a one-to-one correspondence with $\cos\theta_l$ as there is only one degree of freedom left in the system. This can be clearly seen for $B^+ \rightarrow J/\psi K^+$ data in Fig. 7.4, which means the veto must be taken into account in the fit. The width and position of this dip depends on q^2 , which is estimated for each q^2 bin by comparing $B^+ \rightarrow K^+\mu^+\mu^-$ simulation with and without the $\bar{D}^0$ veto applied. The ratio of the $\cos\theta_l$ distributions of the two is then fit using a step function.

The efficiency as a function of $\cos\theta_l$ for the two wide q^2 bins is shown in Fig. 7.5. The function is parameterised by a forth order polynomial, where the odd terms are set to zero to enforce

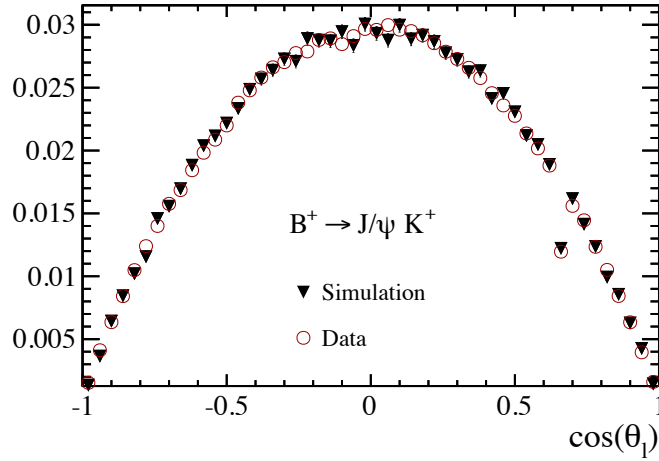


Figure 7.4: Distribution of $\cos \theta_l$ for $B^+ \rightarrow J/\psi K^+$ in data and simulation. There is dip corresponding to the $B^+ \rightarrow \bar{D}^0 \pi^+$ veto.

symmetry around $\cos \theta_l = 0$. The efficiency shape is similar between the K_s^0 and K^+ decay modes as $\cos \theta_l$ is not highly correlated to the kaon kinematics.

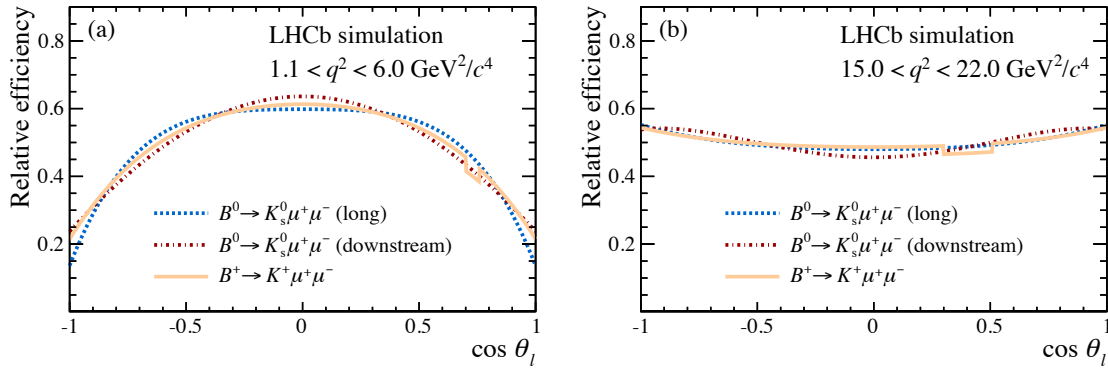


Figure 7.5: Efficiency as a function of $\cos \theta_l$ for $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays in the wide q^2 bins. The step function is required in the $B^+ \rightarrow K^+ \mu^+ \mu^-$ case to reproduce the effect of the $B^+ \rightarrow \bar{D}^0 \pi^+$ veto.

At extreme $\cos \theta_l$ values in the low q^2 region, one muon will be pointing opposite to the B direction and will have very low momentum in the lab frame. For example, below dimuon masses of $1 \text{ GeV}/c^2$ with $\cos \theta_l < -0.8$, the muon with the same sign as the kaon has an average momentum of $7 \text{ GeV}/c$. This muon will be more easily bent out of the muon acceptance by the magnet and

is not easily reconstructed. The trigger efficiency is also lower for these decays, as there is only one muon with large p_T to trigger on. Both these effects are reproduced in the simulation with an accuracy beyond the level of precision required, which is reflected in the good agreement between data and simulation for the muon kinematics in $B^+ \rightarrow J/\psi K^+$ decays, shown in Fig. 7.6. At high q^2 , the pointing of the muon is not so dependent on $\cos\theta_l$ which results in a relatively flat efficiency shape.

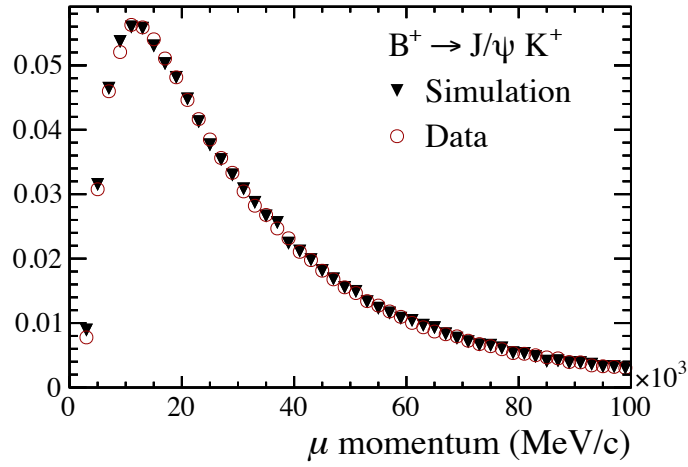


Figure 7.6: Distribution of the muon momentum for $B^+ \rightarrow J/\psi K^+$ decays in data and simulation. There is a good agreement across the full kinematic range.

7.3. Angular fit

The angular observables are determined using a 2D fit to $\cos\theta_l$ and $K\mu^+\mu^-$ mass. The $m(K^+\mu^+\mu^-)$ and $m(K_s^0\mu^+\mu^-)$ invariant mass distributions of candidates that pass the full selection procedure are shown in Fig. 7.7, for two q^2 intervals. The long and downstream categories are combined for the decay $B^0 \rightarrow K_s^0\mu^+\mu^-$. The angular distribution of the candidates is shown in Fig. 7.8. The background is parameterised by a Chebychev polynomial, the parameters of which are allowed to float. For the $B^+ \rightarrow K^+\mu^+\mu^-$ channel, the background is parameterised with three parameters. For the $B^0 \rightarrow K_s^0\mu^+\mu^-$ channel, two and one (zero) parameters are used for the DD and LL categories, respectively. The LL category uses one free parameter $\cos\theta_l$ for the first two q^2 bins, where the background level is relatively high, and zero free parameters for the highest three q^2 bins, where the background level is low. A background parameterisation with zero free parameters assumes that the background distribution is flat in $\cos\theta_l$. The number of parameters used is driven by the amount of background left at the end of the selection. In general the maximum number of parameters which does not spoil fit stability is chosen. The bias introduced by assuming a background parameterisation is the only important systematic for this analysis and is discussed in Sect. 7.5.

7.3.1. Minimisation

The physical boundary means that quite often a simple minimiser, such as Minuit, will fail to find the right minimum. For $B^+ \rightarrow K^+\mu^+\mu^-$, as there are only two dimensions, a brute force method is used to minimise. The likelihood is evaluated at a point in the physical region and eight adjacent points. The point with the best likelihood is then used for the next iteration, where another eight adjacent points are tested, but with a smaller distance between points. This is repeated, reducing the distance between points each time until the difference in log-likelihood (DLL) less than 5×10^{-3} . This technique converges relatively quickly, and avoids the problem of convergence when close to the boundary, as the angular observables are fixed for the entire procedure. Given that the likelihood surfaces are well behaved (i.e. no second minima), the technique is expected to work well and indeed the central values agree with simply manually scanning the entire 2D likelihood surface and picking the lowest value. For $B^0 \rightarrow K_s^0\mu^+\mu^-$, a simple profile likelihood is performed and the best DLL is chosen as there is only one angular observable in this case.

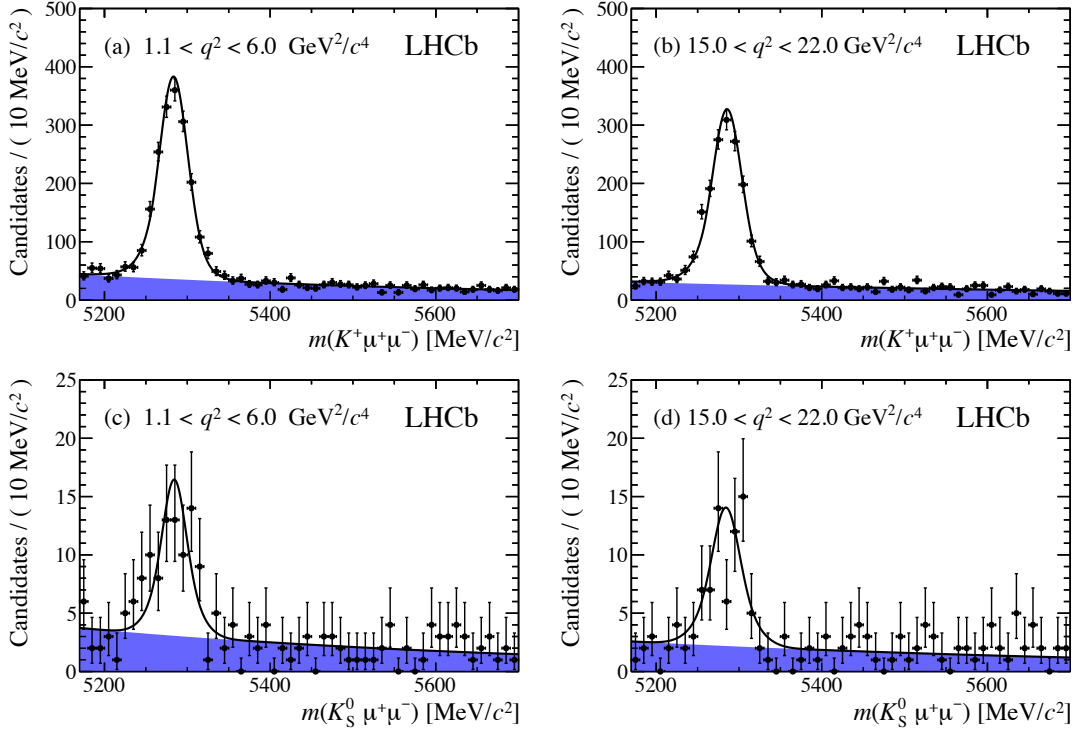


Figure 7.7: Top, reconstructed mass of $B^+ \rightarrow K^+\mu^+\mu^-$ candidates in the ranges (a) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (b) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. Bottom, reconstructed mass of $B^0 \rightarrow K_S^0\mu^+\mu^-$ candidates in the ranges (c) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (d) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The data are overlaid with the result of the fit described in the text. The shaded region indicates the background contribution in the fit.

7.3.2. Coverage

The Feldman-Cousins procedure [92] is performed to estimate 68% uncertainties. For each parameter of interest (POI), a profile likelihood scan is performed. At each scan point, pseudo-experiments are generated and fits are performed to this data with the POI fixed and also allowed to vary. This is done many times at each point and the DLL is recorded. The number of DLLs which are greater than the DLL obtained from the profile likelihood in data defines the p-value at a given point. The one σ intervals are then defined where the p-value reaches 32%. In this procedure A_{FB} and F_{H} are treated independently. For example, when determining the confidence interval for F_{H} , A_{FB} is treated as if it were a nuisance parameter. Nuisance parameters are dealt with using the so-called “plug-in” method [93]. At each value of F_{H} or A_{FB} that is considered, the maximum likelihood estimate of the nuisance parameters is used when generating the pseudo-experiments.

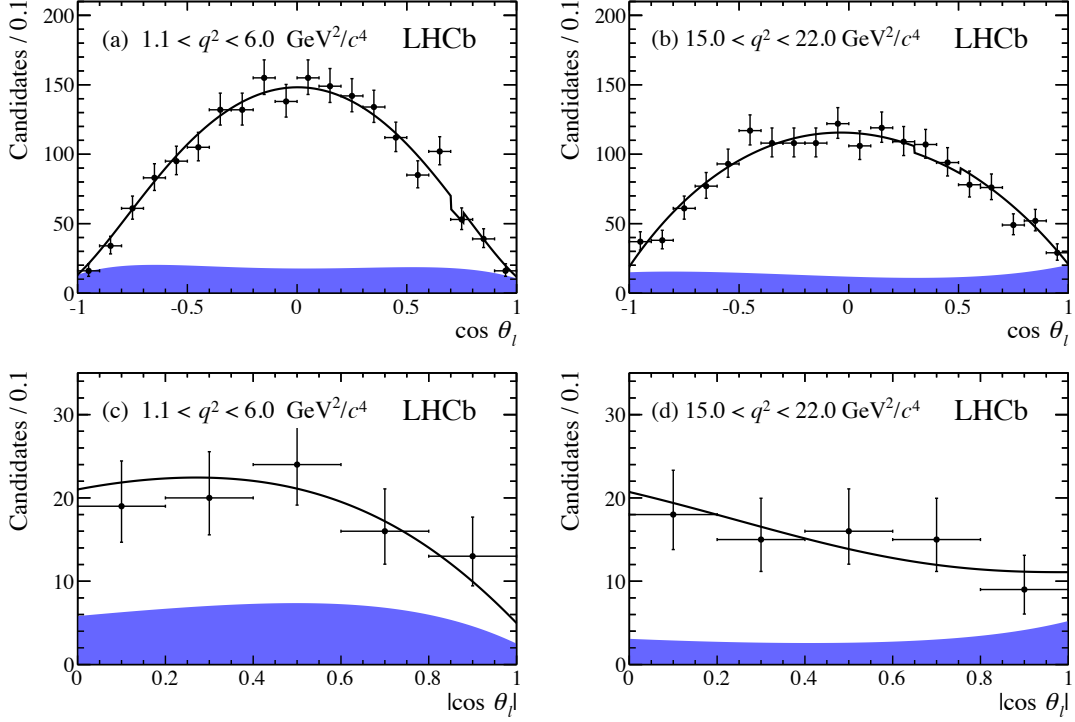


Figure 7.8: Top, angular distribution of $B^+ \rightarrow K^+\mu^+\mu^-$ candidates with (a) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (b) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. Bottom, angular distribution of $B^0 \rightarrow K_s^0\mu^+\mu^-$ candidates with (c) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (d) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. Only candidates with a reconstructed mass within $\pm 50 \text{ MeV}/c^2$ of the known B^+ or B^0 mass are shown. The data are overlaid with the result of the fit described in the text. The shaded region indicates the background contribution in the fit.

For $B^+ \rightarrow K^+\mu^+\mu^-$, the unavoidably large correlation between A_{FB} and F_{H} due to the physical boundary means that coverage is not guaranteed when A_{FB} or F_{H} is “plugged in” as a nuisance parameter. This issue is treated by providing 2D Feldman-Cousins confidence intervals in addition to the 1D ones.

The p -value for A_{FB} and F_{H} in the wide $1\text{--}6 \text{ GeV}^2/c^4$ q^2 bin for $B^+ \rightarrow K^+\mu^+\mu^-$ is shown in Fig. 7.9. The A_{FB} coverage is symmetric, whereas for F_{H} the upper uncertainty tends to be larger. In general, choosing the uncertainty where the profile likelihood is 0.5 tends to over-cover near the boundary, shown in Fig. 7.10, where the Feldman-Cousins intervals are smaller than the ones obtained from the profile likelihood.

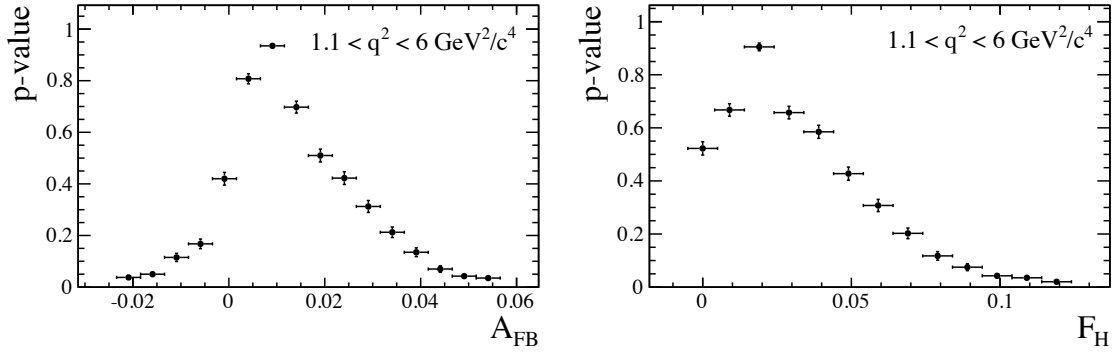


Figure 7.9: Fraction of toy datasets which have a DLL bigger than the one in the data for $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the first q^2 bin.

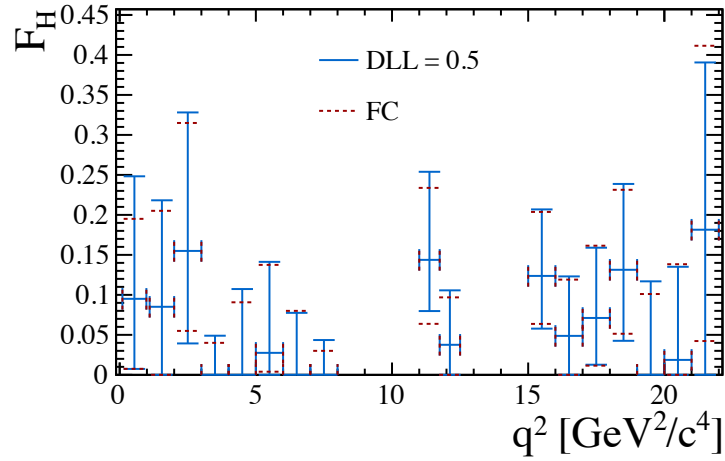


Figure 7.10: Statistical uncertainties for F_H in bins of q^2 obtained from the profile likelihood and obtained using the Feldman-Cousins technique.

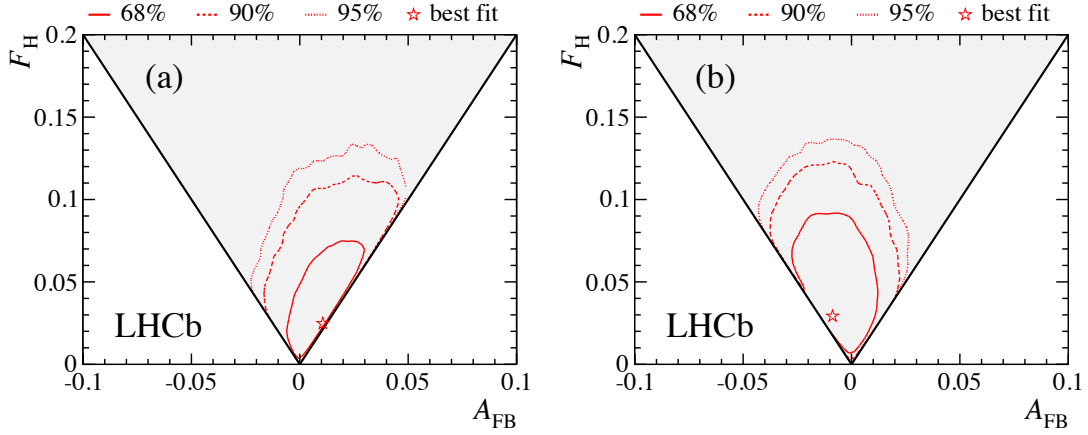


Figure 7.11: Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+\mu^+\mu^-$ in the q^2 ranges (a) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (b) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.

7.4. Results

Two-dimensional confidence regions, for the q^2 ranges $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$, are shown in Fig. 7.11, and for the other q^2 bins are provided in Appendix D. The result of the fits for the decays $B^+ \rightarrow K^+\mu^+\mu^-$ and $B^0 \rightarrow K_s^0\mu^+\mu^-$ are shown in Figs. 7.12 and 7.13, respectively. Tabulated results can be found in Tables C.2 and C.1 in Appendix C. Results are presented in 17 (5) bins of q^2 for the $B^+ \rightarrow K^+\mu^+\mu^-$ ($B^0 \rightarrow K_s^0\mu^+\mu^-$). They are also presented in two wide bins of q^2 : one at low hadronic-recoil above the open charm threshold and one at large recoil, below the J/ψ meson mass.

The fitting procedure for $B^+ \rightarrow K^+\mu^+\mu^-$ ($B^0 \rightarrow K_s^0\mu^+\mu^-$) decays is also tested using samples of $B^+ \rightarrow J/\psi K^+$ ($B^0 \rightarrow J/\psi K_s^0$) decays where $A_{\text{FB}} = F_{\text{H}} = 0$, due to the vector nature of the J/ψ . These samples are more than one hundred times larger than the signal samples. Tests are also performed splitting these samples into sub-samples of comparable size to the data sets in the individual q^2 bins. No indication of any bias is seen in the fitting procedure in either set of tests. The fits to the $B^0 \rightarrow K_s^0\mu^+\mu^-$ decay have also been repeated allowing for a non-zero A_{FB} . The value of A_{FB} returned by these fits is consistent with zero, as expected, and the best fit value of F_{H} compatible with that of the baseline fit.

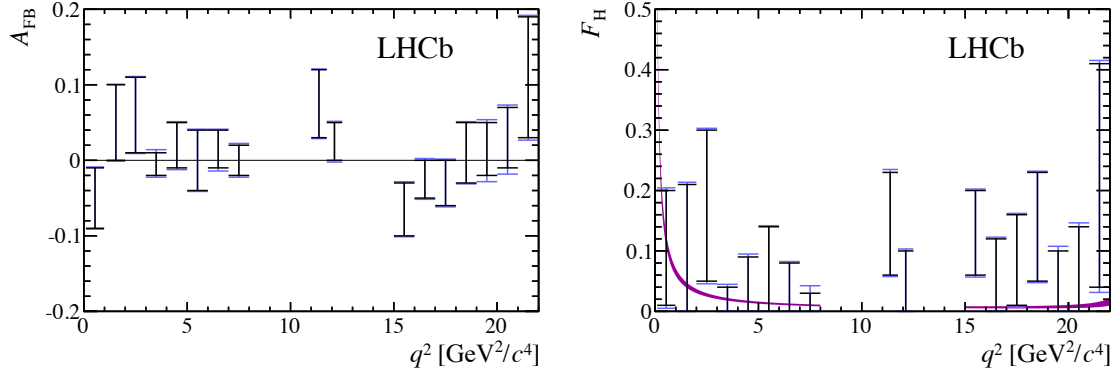


Figure 7.12: Dimuon forward-backward asymmetry, A_{FB} , and the parameter F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ as a function of the dimuon invariant mass squared, q^2 . The inner horizontal bars indicate the one-dimensional 68% statistical confidence intervals. The outer horizontal bars include contributions from systematic uncertainties (described in the text). The confidence intervals for F_{H} are overlaid with the SM theory prediction (narrow band). Data are not presented for the regions around the J/ψ and $\psi(2S)$ resonances.

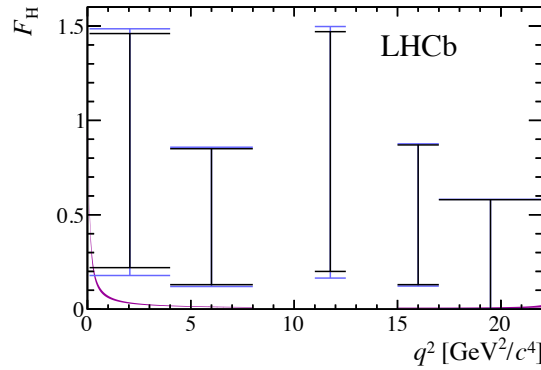


Figure 7.13: Results for the parameter F_{H} for the decay $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ as a function of the dimuon invariant mass squared, q^2 . The inner horizontal bars indicate the one-dimensional 68% statistical confidence intervals. The outer horizontal bars include contributions from systematic uncertainties (described in the text). The confidence intervals are overlaid with the SM theory prediction (narrow band). Data are not presented for the regions around the J/ψ and $\psi(2S)$ resonances.

7.5. Systematic uncertainties

7.5.1. Background parametrisation

The effect of assuming a background shape with a finite number of parameters is estimated by relaxing the BDT cut, and fitting the $\cos\theta_l$ distribution with a larger number of parameters. An example of this can be shown in Fig. 7.14, along with the corresponding nominal fit (with the nominal BDT cut). Toy datasets are generated with the background parameterisation with the relaxed BDT cut and this dataset is fit with the generated PDF and the nominal (simpler) background parameterisation. These toy datasets are generated from the central values of the angular observables and have the same mass fit parameters as the fit to the actual data. For each pseudo-experiment, the differences between the best fit values for A_{FB} and F_H for the two background parameterisations are recorded. These differences for an ensemble of toys are shown in Fig. 7.15 for an example q^2 bin. The RMS or bias of the distribution is chosen as the associated systematic uncertainty, depending which is largest.

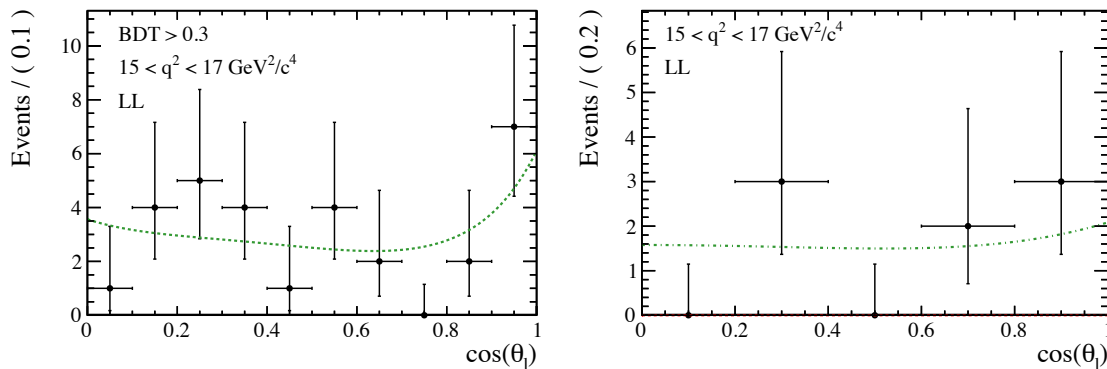


Figure 7.14: Fits to the $\cos\theta_l$ distribution in the background region where the BDT is loosened (left) and the nominal fit (right). The loosened BDT PDF has four parameters whereas the nominal is assumed to be flat (reminder: the PDF is multiplied by the efficiency and so does not look perfectly flat here). For the associated systematic uncertainty, datasets are generated from the left PDF and fit with both PDFs, where the difference in F_H is used to quantify the uncertainty.

This procedure assumes that the background shape does not depend heavily on the BDT cut, which is checked in the data by calculating the correlation between the BDT and $\cos\theta_l$, which is found to be at the level of 1×10^{-3} (averaged over all BDT response and $\cos\theta_l$ values). The parameterisation of the loose BDT data also fits the data with the nominal BDT cut well.

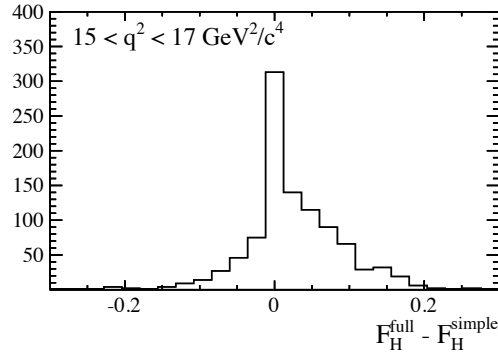


Figure 7.15: Difference between the F_H best fit results in $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ toy datasets for two different background parameterisations. The toy datasets are generated from the “full” parameterisation. The “simple” parameterisation is used for the final result, where a few parameters are let free. The “full” parameterisation is fixed from a fit to the background at a looser BDT cut, where four parameters are used.

7.5.2. Negligible systematics

Systematic uncertainties related to the efficiency as a function of $\cos \theta_l$ are negligible as the main effect originates from geometry which is expected to be well modelled in the simulation. This is shown to be the case earlier in Fig. 7.6, where the agreement of the muon kinematics is much better than needed for this analysis. The only other effect which has a significant effect on the efficiency shape is the trigger, which is also shown to be very well modelled, as shown earlier in Fig. 5.4.

As discussed in Sects. 5.2.2 and 6.3, peaking backgrounds are negligible for these decays after the specific vetoes have been applied. The effect of the D^0 veto is small and is also expected to be well reproduced by the simulation. Finally, the mass model is expected to describe the signal to a much better precision that is needed (for example see Fig. 5.8). Furthermore, corrections to the mass shape as a function of q^2 have a negligible effect on the signal yield, shown in Fig. 6.9.

7.6. Summary

In summary, the angular distribution of charged and neutral $B \rightarrow K\mu^+\mu^-$ decays is studied using a data set, corresponding to an integrated luminosity 3 fb^{-1} , collected by the LHCb experiment. The

angular distribution of the decays is parameterised in terms of the forward-backward asymmetry of the decay, A_{FB} , and a parameter F_{H} .

The measurements of A_{FB} and F_{H} presented for the decays $B^+ \rightarrow K^+\mu^+\mu^-$ and $B^0 \rightarrow K_s^0\mu^+\mu^-$ are the most precise to date. They are consistent with SM predictions ($A_{\text{FB}} \approx 0$ and $F_{\text{H}} \approx 0$) in every bin of q^2 . The results are also compatible between the decays $B^+ \rightarrow K^+\mu^+\mu^-$ and $B^0 \rightarrow K_s^0\mu^+\mu^-$. The largest difference with respect to the SM prediction is seen in the range $11.00 < q^2 < 11.75 \text{ GeV}^2/c^4$ for the decay $B^+ \rightarrow K^+\mu^+\mu^-$. Even in this bin, the SM point is included at 95% confidence level when taking into account the systematic uncertainties on the angular observables.

Chapter 8.

Conclusions

This thesis describes several measurements of the rare electroweak penguin decays $B \rightarrow K^{(*)}\mu^+\mu^-$, with the main purpose to search for physics beyond the Standard Model (SM). These measurements include the world's most precise and include branching fractions, isospin asymmetries and angular observables. The sensitivity of the branching fraction measurements are saturated by theoretical uncertainties in most decay modes whereas for the isospin and angular observables the theoretical uncertainty is negligible.

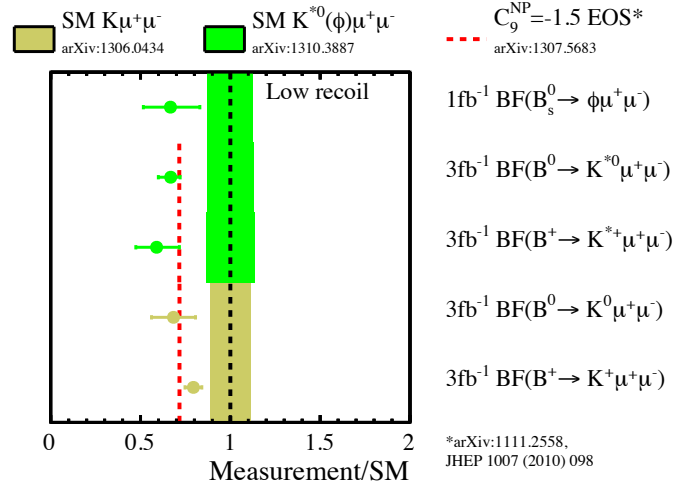


Figure 8.1: Ratio of the measured branching fractions to the SM prediction for $B \rightarrow K^{(*)}\mu^+\mu^-$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays. The red line is the result a global fit based on the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ angular distribution.

The branching fraction results for all four $B \rightarrow K^{(*)}\mu^+\mu^-$ modes are below the SM prediction, a trend also seen in $B_s^0 \rightarrow \phi\mu^+\mu^-$ decays [94], as shown in Fig. 8.1. This is true for the entire kinematic region but is particularly interesting at high q^2 , where the theoretical uncertainty is reduced using LQCD calculations.

This low branching fraction is what is expected if the value of the Wilson coefficient C_9 is smaller than predicted in the SM, which is one of the interpretations [95–97] of an anomaly seen in the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ angular distribution [98]. The theoretical uncertainties associated with the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ angular observables vary considerably between the different papers depending on the treatment of the sub-leading Λ/m_b corrections and thus have different levels of tension with the SM. It is encouraging to see a similar trend at high q^2 where the theoretical treatment is different.

However, the prominent nature of the resonance described in chapter 5 in the low recoil region complicate the interpretation of inclusive branching fraction results in the low recoil region for $B^+ \rightarrow K^+\mu^+\mu^-$ decays. Depending on exactly where the low recoil region is defined, branching fraction measurements can include contribution from the $\psi(3770)$, which is about 6% of the total rate in that region. Given the smooth nature of the theoretical predictions, this is currently not taken into account and will need to be addressed in the future. The branching fraction of the $B \rightarrow K\mu^+\mu^-$ decay mode is closer to the SM prediction than the $B \rightarrow K^*\mu^+\mu^-$. This could be a statistical fluctuation, but could also indicate a problem with the $c\bar{c}$ contribution predicted by the OPE, where the 20% contribution measured in Ref. [1] is claimed to be larger than predicted in theory. Since the publication of Ref. [1], there has been some, still unresolved, discussion about the accuracy of this statement, however, if the resonance contribution is larger than theoretical predictions then it will artificially increase the $B^+ \rightarrow K^+\mu^+\mu^-$ branching fraction towards the SM value. For $B^0 \rightarrow K^{*0}\mu^+\mu^-$, there is no visible structure and the $c\bar{c}$ contribution appears to be less of an issue.

There is also an issue with the $B \rightarrow K\mu^+\mu^-$ predictions from the lattice described in [66], which overestimate the branching fraction by roughly 10% due to a missed two-loop virtual correction to the effective value of C_9 . If this correction is applied then the tension is reduced significantly for the $B \rightarrow K\mu^+\mu^-$ case.

Another possibility, proposed in Ref. [96], is that C'_9 is non-zero, i.e. that new physics affecting C_9 has a right-handed component. For $B^+ \rightarrow K^+\mu^+\mu^-$, there is no freedom in the polarisation and so an C'_9 component fully interferes with C_9 and can cancel the reduction of branching fraction. For $B^0 \rightarrow K^{*0}\mu^+\mu^-$ the situation is more complicated as a C'_9 component will have an opposite interference with the C_9 part depending on the polarisation state of the $\mu^+\mu^-$ pair.

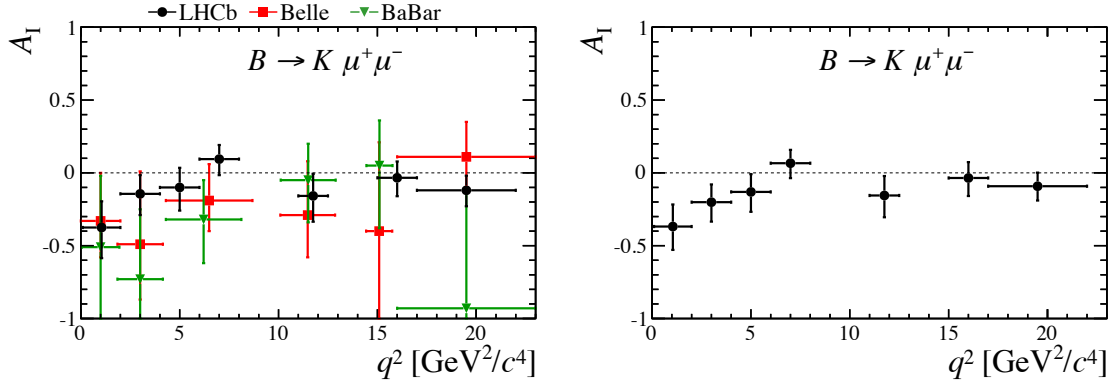


Figure 8.2: Isospin asymmetry in $B \rightarrow K \mu^+ \mu^-$ decays as a function of q^2 as obtained from the LHCb [2], Belle [73] and BaBar [72] collaborations. The right figure shows a naive combination of these three measurements, assuming uncorrelated systematic uncertainties.

Although the isospin asymmetry for $B \rightarrow K \mu^+ \mu^-$ decays is negative in all but one q^2 bin, results are more consistent with the SM compared to the previous measurement in Ref. [5]. Despite the reduction in significance compared to the previous result, measurements are still negative at low q^2 , which agree with the results from the B -factories, as shown in Fig. 8.2, which includes a naive combination of the results from the three experiments. The low q^2 region is the one which is the most interesting theoretically, as it where isospin asymmetries are more likely to propagate through the emission of a photon. There is also the possibility that this could be related to the peaking structure shown in Fig. 6.13 for $B^+ \rightarrow K^+ \mu^+ \mu^-$ at very low q^2 . The enhancement seen in the first bin appears to be too large to be from resonances. Such structure is possible in new physics models which violate isospin symmetry at low q^2 , such as the model proposed in Ref. [71], which predicts a neutral heavy gauge boson, and appears to fit with the anomaly observed in the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ angular distribution [98]. In summary, the combination of the isospin asymmetry measurements and the structure of the $B^+ \rightarrow K^+ \mu^+ \mu^-$ branching fraction at low q^2 is intriguing. More precise experimental measurements will be of great interest when LHCb collects more data.

In chapter 7, the angular distribution of charged and neutral $B \rightarrow K \mu^+ \mu^-$ decays were studied with a dataset corresponding to 3 fb^{-1} of luminosity. The angular observables A_{FB} and F_{H} in $B^+ \rightarrow K^+ \mu^+ \mu^-$ were updated with respect to Ref. [60], and an angular analysis has been performed for $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ for the first time.

The angular analysis is the first of its kind to provide correlation information on the angular observables, in the form of the two-dimensional Feldman-Cousins intervals. This is not only the correct way to present such observables, but provides crucial information to the theoretical

community, as it allows consistency checks with the SM without having to assume that A_{FB} and F_{H} are independent. This could be even more critical in the more complicated $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ angular analyses, where the theoretical uncertainties are comparable with the experimental precision. In this case it would allow the possibility for a more sophisticated treatment of theoretical nuisance parameters, where the correlations between measurements can provide input.

The angular analysis results place constraints on (pseudo)scalar and tensor amplitudes, which are vanishingly small in the SM but can be enhanced in many extensions of the SM. In Ref. [88], it is stated that only contributions from new tensor like particles could be visible at the current level of experimental precision, as pseudoscalar and scalar amplitudes were already highly constrained by measurements of the branching fraction of the decay $B_s^0 \rightarrow \mu^+ \mu^-$ [99, 100]. The results presented here, however, also rule out the possibility of large accidental cancellations between the left- and right-handed couplings of the (pseudo)scalar amplitudes to the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction. The absence of other operators mean that these measurements provide complimentary information to other modes, such as the well known $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay. It is currently unclear whether tensor contributions could also provide the deviation in the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ angular distribution discussed earlier, where these angular analysis measurements will be invaluable in dis-entangling the issue.

In summary, there is a consistent picture emerging from the measurements of these rare electroweak penguin decays, which suggests that the value of $\mathcal{C}_9$ is low compared to the SM prediction. A possible new physics model which fits into this is a Z' particle with a mass of $\sim 7 \text{ TeV}/c^2$ [101]. More data will be needed to confirm this anomaly, and future updates to the branching fractions of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ (with the S-wave removed), $B_s^0 \rightarrow \phi \mu^+ \mu^-$, $\Lambda_b \rightarrow \Lambda \mu^+ \mu^-$ and $\Lambda_b \rightarrow p K \mu^+ \mu^-$ will be very important, particularly if the form factor calculations between each hadron are somewhat uncorrelated. An improved theoretical estimate of the $c\bar{c}$ contribution at high q^2 is also desirable, perhaps with help of measurements in data similar to the one described in chapter 5. Finally, a measurement of the inclusive $b \rightarrow s \mu^+ \mu^-$ decay rate would also be very useful, as theoretical uncertainties should also be somewhat uncorrelated to the exclusive ones.

Bibliography

- [1] LHCb collaboration, R. Aaij *et al.*, *Observation of a resonance in $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays at low recoil*, Phys. Rev. Lett. **111** (2013) 112003, [arXiv:1307.7595](#).
- [2] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays*, [arXiv:1403.8044](#).
- [3] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of charged and neutral $B \rightarrow K \mu^+ \mu^-$ decays*, JHEP **05** (2014) 082, [arXiv:1403.8045](#).
- [4] LHCb collaboration, R. Aaij *et al.*, *Search for lepton number violating decays $B^+ \rightarrow \pi^- \mu^+ \mu^+$ and $B^+ \rightarrow K^- \mu^+ \mu^+$* , Phys. Rev. Lett. **108** (2012) 101601, [arXiv:1110.0730](#).
- [5] LHCb collaboration, R. Aaij *et al.*, *Measurement of the isospin asymmetry in $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays*, JHEP **07** (2012) 133, [arXiv:1205.3422](#).
- [6] M. A. Shifman, *Foreword to ITEP lectures in particle physics*, [arXiv:hep-ph/9510397](#).
- [7] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. (1961), no. 22 579.
- [8] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [9] A. Salam, *Weak and electromagnetic interactions*, in *Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lurem, Sweden, Stockholm*, pp. 367–377, 1968.
- [10] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.
- [11] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156.
- [12] C. R. H. G. S. Guranlnik and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.
- [13] T. W. B. Kibble, *Symmetry breaking in non-abelian gauge theories*, Phys. Rev. **155** (1967) 1554.
- [14] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.
- [15] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

- [16] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [17] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Physical Review Letters **51** (1983) 1945.
- [18] A. D. Sakharov, *Violation of CP in variance, C asymmetry, and baryon asymmetry of the universe*, Sov. Phys. Usp. **34** (1991), no. 5 392.
- [19] Planck Collaboration, P. A. R. Ade *et al.*, *Planck 2013 results. XVI. Cosmological parameters*, arXiv:1303.5076.
- [20] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.
- [21] R. Peccei, *The Strong CP problem and axions*, Lect. Notes Phys. **741** (2008) 3, arXiv:hep-ph/0607268.
- [22] S. Weinberg, *Implications of dynamical symmetry breaking*, Phys. Rev. D **13** (1976) 974.
- [23] S. Weinberg, *Implications of dynamical symmetry breaking: An addendum*, Phys. Rev. D **19** (1979) 1277.
- [24] E. Gildener, *Gauge-symmetry hierarchies*, Phys. Rev. D **14** (1976) 1667.
- [25] L. Susskind, *Dynamics of spontaneous symmetry breaking in the weinberg-salam theory*, Phys. Rev. D **20** (1979) 2619.
- [26] G. Isidori, Y. Nir, and G. Perez, *Flavor Physics Constraints for Physics Beyond the Standard Model*, Ann. Rev. Nucl. Part. Sci. **60** (2010) 355, arXiv:1002.0900.
- [27] K. G. Wilson, *Non-lagrangian models of current algebra*, Phys. Rev. **179** (1969) 1499.
- [28] T. Appelquist and J. Carazzone, *Infrared singularities and massive fields*, Phys. Rev. D **11** (1975) 2856.
- [29] A. Ali, P. Ball, L. Handoko, and G. Hiller, *A Comparative study of the decays $B \rightarrow (K, K^*)\ell^+\ell^-$ in standard model and supersymmetric theories*, Phys. Rev. **D61** (2000) 074024, arXiv:hep-ph/9910221.
- [30] T. Mannel, *Effective Field Theories in Flavor Physics*, Springer Tracts Mod. Phys. **203** (2004) 1.
- [31] C. W. Bauer *et al.*, *An effective field theory for collinear and soft gluons: heavy to light decays*, Phys. Rev. **D63** (2001) 114020, arXiv:hep-ph/0011336.
- [32] V. M. Braun, *Light cone sum rules*, arXiv:hep-ph/9801222.
- [33] B. Grinstein and D. Pirjol, *Exclusive rare $B \rightarrow K^*\ell^+\ell^-$ decays at low recoil: controlling the long-distance effects*, Phys. Rev. **D70** (2004) 114005, arXiv:hep-ph/0404250.

- [34] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Lattice QCD calculation of form factors describing the rare decays $B \rightarrow K^* \ell^+ \ell^-$ and $B_s \rightarrow \phi \ell^+ \ell^-$* , [arXiv:1310.3722](#).
- [35] C. Bouchard *et al.*, *Rare decay $B \rightarrow K \ell \ell$ form factors from lattice QCD*, Phys. Rev. **D88** (2013) 054509, [arXiv:1306.2384](#).
- [36] L. Evans and P. Bryant, *LHC machine*, Journal of Instrumentation **3** (2008), no. 08 S08001.
- [37] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \rightarrow b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. **B694** (2010) 209, [arXiv:1009.2731](#).
- [38] BaBar Collaboration, B. Aubert *et al.*, *The BaBar detector*, Nucl. Instrum. Meth. **A479** (2002) 1, [arXiv:hep-ex/0105044](#).
- [39] A. Abashian *et al.*, *The Belle Detector*, Nucl. Instrum. Meth. **A479** (2002) 117.
- [40] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [41] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759](#).
- [42] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, [arXiv:1211.1346](#).
- [43] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, [arXiv:1211.3055](#).
- [44] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, [arXiv:1210.6861](#).
- [45] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [46] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [47] I. Belyaev *et al.*, *Handling of the generation of primary events in GAUSS, the LHCb simulation framework*, Nuclear Science Symposium Conference Record (NSS/MIC) **IEEE** (2010) 1155.
- [48] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [49] M. Clemencic *et al.*, *The LHCb simulation application, GAUSS: design, evolution and experience*, J. Phys.: Conf. Ser. **331** (2011) 032023.
- [50] R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, LHCb-DP-2013-002, in preparation.

- [51] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $D_s^+D_s^-$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B713** (2012) 186, [arXiv:1205.0897](#).
- [52] P. Avery, *Applied fitting theory 1*, <http://www.phys.ufl.edu/avery/fitting/>.
- [53] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.
- [54] R. Fleischer and T. Mannel, *General analysis of new physics in $B \rightarrow J/\psi K$* , Phys. Lett. **B506** (2001) 311, [arXiv:hep-ph/0101276](#).
- [55] BES collaboration, M. Ablikim *et al.*, *Determination of the $\psi(3770)$, $\psi(4040)$, $\psi(4160)$ and $\psi(4415)$ resonance parameters*, Phys. Lett **B660** (2008) 315, [arXiv:0705.4500](#).
- [56] DASP collaboration, R. Brandelik *et al.*, *Total cross-section for hadron production by e^+e^- annihilation at center-of-mass energies between 3.6 GeV and 5.2 GeV*, Phys. Lett. **B76** (1978) 361.
- [57] M. Beylich, G. Buchalla, and T. Feldmann, *Theory of $B \rightarrow K^{(*)}\ell^+\ell^-$ decays at high q^2 : OPE and quark-hadron duality*, Eur. Phys. J. **C71** (2011) 1635, [arXiv:1101.5118](#).
- [58] Belle collaboration, K. Abe *et al.*, *Observation of the decay $B \rightarrow K\ell^+\ell^-$* , Phys. Rev. Lett. **88** (2002) 021801, [arXiv:hep-ex/0109026](#).
- [59] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in $B^+ \rightarrow K^+\mu^+\mu^-$ decays*, Phys. Rev. Lett. **111** (2013) 151801, [arXiv:1308.1340](#).
- [60] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the $B^+ \rightarrow K^+\mu^+\mu^-$ decay*, JHEP **02** (2013) 105, [arXiv:1209.4284](#).
- [61] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, *The decay $B \rightarrow K\ell^+\ell^-$ at low hadronic recoil and model-independent $\Delta B = 1$ constraints*, JHEP **01** (2012) 107, [arXiv:1111.2558](#).
- [62] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, [arXiv:physics/0408124](#).
- [63] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, [arXiv:physics/0503191](#).
- [64] ARGUS collaboration, H. Albrecht *et al.*, *Search for Hadronic $b \rightarrow u$ Decays*, Phys. Lett. **B241** (1990) 278.
- [65] A. Khodjamirian, Th. Mannel, A. A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in $B \rightarrow K^{(*)}\ell^+\ell^-$ and $B \rightarrow K^*\gamma$* , JHEP **09** (2010) 089, [arXiv:1006.4945](#).
- [66] C. Bouchard *et al.*, *Standard Model predictions for $B \rightarrow K \ell\ell$ with form factors from lattice QCD*, Phys. Rev. Lett. **111** (2013) 162002, [arXiv:1306.0434](#).
- [67] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Calculation of $B \rightarrow K^*\mu^+\mu^-$ and*

- $B_s \rightarrow \phi \mu^+ \mu^-$ using form factors from lattice QCD, [arXiv:1310.3887](#).
- [68] J. Lyon and R. Zwicky, *Isospin asymmetries in $B \rightarrow (K^*, \rho) \gamma / l^+ l^-$ and $B \rightarrow K l^+ l^-$ in and beyond the Standard Model*, Phys. Rev. **D88** (2013) 094004, [arXiv:1305.4797](#).
- [69] A. Khodjamirian, T. Mannel, and Y. Wang, *$B \rightarrow K \ell^+ \ell^-$ decay at large hadronic recoil*, JHEP **1302** (2013) 010, [arXiv:1211.0234](#).
- [70] T. Feldmann and J. Matias, *Forward-backward and isospin asymmetry for $B \rightarrow K^{(*)} l^+ l^-$ decay in the Standard Model and in supersymmetry*, JHEP **01** (2003) 074, [arXiv:hep-ph/0212158](#).
- [71] H. Ishida and Y. Koide, *Effective Valence Quark Model and a Possible Dip in $d\text{Br}(B \rightarrow K \ell \bar{\ell})/dq^2$* , [arXiv:1305.7342](#).
- [72] BaBar collaboration, *Measurement of branching fractions and rate asymmetries in the rare decays $B \rightarrow K^{(*)} \ell^+ \ell^-$* , submitted to Phys. Rev. D, [arXiv:1204.3933](#).
- [73] Belle collaboration, J.-T. Wei *et al.*, *Measurement of the differential branching fraction and forward-backward asymmetry for $B \rightarrow K^{(*)} l^+ l^-$* , Phys. Rev. Lett. **103** (2009) 171801, [arXiv:0904.0770](#).
- [74] BaBar collaboration, B. Aubert *et al.*, *Direct CP, lepton flavor and isospin asymmetries in the decays $B \rightarrow K^{(*)} l^+ l^-$* , Phys. Rev. Lett. **102** (2008) 091803, [arXiv:0807.4119](#).
- [75] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of early 2012*, [arXiv:1207.1158](#), updated results and plots available at: <http://www.slac.stanford.edu/xorg/hfag/>.
- [76] C. Parkinson, *The angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay at LHCb*, PhD thesis, Imperial College London, 2013.
- [77] BaBar Collaboration, B. Aubert *et al.*, *Measurement of branching fractions and charge asymmetries for exclusive B decays to charmonium*, Phys. Rev. Lett. **94** (2005) 141801, [arXiv:hep-ex/0412062](#).
- [78] Belle collaboration, K. Abe *et al.*, *Measurement of branching fractions and charge asymmetries for two-body B meson decays with charmonium*, Phys. Rev. **D67** (2003) 032003, [arXiv:hep-ex/0211047](#).
- [79] Belle collaboration, K. Abe *et al.*, *Measurements of branching fractions and decay amplitudes in $B \rightarrow J/\psi K^*$ decays*, Phys. Lett. **B538** (2002) 11, [arXiv:hep-ex/0205021](#).
- [80] D. Becirevic and A. Tayduganov, *Impact of $B \rightarrow K_0^* \ell^+ \ell^-$ on the New Physics search in $B \rightarrow K^* \ell^+ \ell^-$ decay*, Nucl. Phys. **B868** (2013) 368, [arXiv:1207.4004](#).
- [81] C. Bobeth, G. Hiller, and D. van Dyk, *More benefits of semileptonic rare B decays at low recoil: CP Violation*, JHEP **07** (2011) 067, [arXiv:1105.0376](#).

- [82] M. Beneke, T. Feldmann, and D. Seidel, *Systematic approach to exclusive $B \rightarrow V\ell^+\ell^-$, $V\gamma$ decays*, Nucl. Phys. **B612** (2001) 25, [arXiv:hep-ph/0106067](#).
- [83] M. Beneke, T. Feldmann, and D. Seidel, *Exclusive radiative and electroweak $b \rightarrow d$ and $b \rightarrow s$ penguin decays at NLO*, Eur. Phys. J. **C41** (2005) 173, [arXiv:hep-ph/0412400](#), 29 pages, LaTeX Report-no: PITHA 04/19.
- [84] C. Bobeth, G. Hiller, and G. Piranishvili, *Angular distributions of $B \rightarrow K^*\ell^+\ell^-$ decays*, JHEP **12** (2007) 040, [arXiv:0709.4174](#).
- [85] U. Egede *et al.*, *New observables in the decay mode $\bar{B}_d \rightarrow \bar{K}^{*0}\ell^+\ell^-$* , JHEP **0811** (2008) 032, [arXiv:0807.2589](#).
- [86] P. Ball and R. Zwicky, *$B_{d,s} \rightarrow \rho, \omega, K^*, \phi$ decay form factors from light-cone sum rules reexamined*, Phys. Rev. **D71** (2005) 014029, [arXiv:hep-ph/0412079](#).
- [87] C. Bouchard *et al.*, *The rare decay $B \rightarrow K\ell^+\ell^-$ form factors from lattice QCD*, [arXiv:1306.2384](#).
- [88] A. K. Alok, A. Dighe, and S. U. Sankar, *Large forward-backward asymmetry in $B \rightarrow K\mu^+\mu^-$ from new physics tensor operators*, Phys. Rev. **D78** (2008) 114025, [arXiv:0810.3779](#).
- [89] BaBar collaboration, B. Aubert *et al.*, *Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays $B \rightarrow K\ell^+\ell^-$ and $B \rightarrow K^*\ell^+\ell^-$* , Phys. Rev. **D73** (2006) 092001, [arXiv:hep-ex/0604007](#).
- [90] Belle, J.-T. Wei *et al.*, *Measurement of the differential branching fraction and forward-backward asymmetry for $B \rightarrow K^{(*)}\ell^+\ell^-$* , Phys. Rev. Lett. **103** (2009) 171801, [arXiv:0904.0770](#).
- [91] CDF Collaboration, T. Aaltonen *et al.*, *Measurements of the angular distributions in the decays $B \rightarrow K^{(*)}\mu^+\mu^-$ at CDF*, Phys. Rev. Lett. **108** (2012) 081807, [arXiv:1108.0695](#).
- [92] G. J. Feldman and R. D. Cousins, *A Unified Approach to the Classical Statistical Analysis of Small Signals*, Phys. Rev. D (1997) [arXiv:physics/9711021](#).
- [93] B. Sen, M. Walker, and M. Woodroffe, *On the unified method with nuisance parameters*, Statistica Sinica **19** (2009) 301.
- [94] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B_s^0 \rightarrow \phi\mu^+\mu^-$* , JHEP **07** (2013) 84, [arXiv:1305.2168](#).
- [95] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $B \rightarrow K^*\mu^+\mu^-$ Anomaly*, Phys. Rev. **D88** (2013) 074002, [arXiv:1307.5683](#).
- [96] W. Altmannshofer and D. M. Straub, *New physics in $B \rightarrow K^*\mu\mu$?*, Eur. Phys. J. **C73** (2013) 2646, [arXiv:1308.1501](#).
- [97] F. Beaujean, C. Bobeth, and D. van Dyk, *Comprehensive Bayesian Analysis of Rare*

- (Semi)leptonic and Radiative B Decays, [arXiv:1310.2478](#).
- [98] LHCb collaboration, R. Aaij *et al.*, *Measurement of form factor independent observables in the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$* , Phys. Rev. Lett. **111** (2013) 191801, [arXiv:1308.1707](#).
- [99] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction and search for $B^0 \rightarrow \mu^+ \mu^-$ decays at the LHCb experiment*, Phys. Rev. Lett. **111** (2013) 101805, [arXiv:1307.5024](#).
- [100] CMS Collaboration, S. Chatrchyan *et al.*, *Measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction and search for $B^0 \rightarrow \mu^+ \mu^-$ with the CMS experiment*, Phys. Rev. Lett. **111** (2013) 101804, [arXiv:1307.5025](#).
- [101] R. Gauld, F. Goertz, and U. Haisch, *On minimal Z' explanations of the $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ anomaly*, Phys. Rev. **D89** (2014) 015005, [arXiv:1308.1959](#).

Appendix A.

Tabulated branching fractions

Table A.1: Differential branching fraction results (10^{-9}) for $B^0 \rightarrow K^0 \mu^+ \mu^-$ including statistical and systematic uncertainties.

q^2 range (GeV^2/c^4)	central value	stat	syst
$0.1 < q^2 < 2.0$	12.2	$+5.9$ -5.2	0.6
$2.0 < q^2 < 4.0$	18.7	$+5.5$ -4.9	0.9
$4.0 < q^2 < 6.0$	17.3	$+5.3$ -4.8	0.9
$6.0 < q^2 < 8.0$	27.0	$+5.8$ -5.3	1.4
$11.0 < q^2 < 12.5$	12.7	$+4.5$ -4.0	0.6
$15.0 < q^2 < 17.0$	14.3	$+3.5$ -3.2	0.7
$17.0 < q^2 < 22.0$	7.8	$+1.7$ -1.5	0.4
$1.1 < q^2 < 6.0$	18.7	$+3.5$ -3.2	0.9
$15.0 < q^2 < 22.0$	9.5	$+1.6$ -1.5	0.5

Table A.2: Differential branching fraction results (10^{-9}) for $B^+ \rightarrow K^+ \mu^+ \mu^-$ including statistical and systematic uncertainties.

q^2 range (GeV^2/c^4)	central value	stat	syst
$0.1 < q^2 < 0.98$	33.2	1.8	1.7
$1.1 < q^2 < 2.0$	23.3	1.5	1.2
$2.0 < q^2 < 3.0$	28.2	1.6	1.4
$3.0 < q^2 < 4.0$	25.4	1.5	1.3
$4.0 < q^2 < 5.0$	22.1	1.4	1.1
$5.0 < q^2 < 6.0$	23.1	1.4	1.2
$6.0 < q^2 < 7.0$	24.5	1.4	1.2
$7.0 < q^2 < 8.0$	23.1	1.4	1.2
$11.0 < q^2 < 11.8$	17.7	1.3	0.9
$11.8 < q^2 < 12.5$	19.3	1.2	1.0
$15.0 < q^2 < 16.0$	16.1	1.0	0.8
$16.0 < q^2 < 17.0$	16.4	1.0	0.8
$17.0 < q^2 < 18.0$	20.6	1.1	1.0
$18.0 < q^2 < 19.0$	13.7	1.0	0.7
$19.0 < q^2 < 20.0$	7.4	0.8	0.4
$20.0 < q^2 < 21.0$	5.9	0.7	0.3
$21.0 < q^2 < 22.0$	4.3	0.7	0.2
$1.1 < q^2 < 6.0$	24.2	0.7	1.2
$15.0 < q^2 < 22.0$	12.1	0.4	0.6

Table A.3: Differential branching fraction results (10^{-9}) for $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ including statistical and systematic uncertainties.

q^2 range (GeV^2/c^4)	central value	stat	syst
$0.1 < q^2 < 2.0$	59.2	$^{+14.4}_{-13.0}$	4.0
$2.0 < q^2 < 4.0$	55.9	$^{+15.9}_{-14.4}$	3.8
$4.0 < q^2 < 6.0$	24.9	$^{+11.0}_{-9.6}$	1.7
$6.0 < q^2 < 8.0$	33.0	$^{+11.3}_{-10.0}$	2.3
$11.0 < q^2 < 12.5$	82.8	$^{+15.8}_{-14.1}$	5.6
$15.0 < q^2 < 17.0$	64.4	$^{+12.9}_{-11.5}$	4.4
$17.0 < q^2 < 19.0$	11.6	$^{+9.1}_{-7.6}$	0.8
$1.1 < q^2 < 6.0$	36.6	$^{+8.3}_{-7.6}$	2.6
$15 < q^2 < 19.0$	39.5	$^{+8.0}_{-7.3}$	2.8

Appendix B.

Tabulated isospin asymmetries

Table B.1: Isospin asymmetry results for $B \rightarrow K\mu^+\mu^-$ decays.

q^2 range (GeV^2/c^4)	central value	stat	syst
$0.1 < q^2 < 2.0$	-0.37	$^{+0.18}_{-0.21}$	0.02
$2.0 < q^2 < 4.0$	-0.15	$^{+0.13}_{-0.15}$	0.02
$4.0 < q^2 < 6.0$	-0.10	$^{+0.13}_{-0.16}$	0.02
$6.0 < q^2 < 8.0$	0.09	$^{+0.10}_{-0.11}$	0.02
$11.0 < q^2 < 12.5$	-0.16	$^{+0.15}_{-0.18}$	0.03
$15.0 < q^2 < 17.0$	-0.04	$^{+0.11}_{-0.13}$	0.02
$17.0 < q^2 < 22.0$	-0.12	$^{+0.10}_{-0.11}$	0.02
$1.1 < q^2 < 6.0$	-0.10	$^{+0.08}_{-0.09}$	0.02
$15.0 < q^2 < 22.0$	-0.09	$^{+0.08}_{-0.08}$	0.02

Table B.2: Isospin asymmetry results for $B \rightarrow K^* \mu^+ \mu^-$ decays.

q^2 range (GeV^2/c^4)	central value	stat	syst
$0.1 < q^2 < 2.0$	0.11	$^{+0.12}_{-0.11}$	0.02
$2.0 < q^2 < 4.0$	-0.20	$^{+0.15}_{-0.12}$	0.03
$4.0 < q^2 < 6.0$	0.23	$^{+0.21}_{-0.18}$	0.02
$6.0 < q^2 < 8.0$	0.19	$^{+0.17}_{-0.15}$	0.02
$11.0 < q^2 < 12.5$	-0.25	$^{+0.09}_{-0.08}$	0.03
$15.0 < q^2 < 17.0$	-0.10	$^{+0.10}_{-0.09}$	0.03
$17.0 < q^2 < 19.0$	0.51	$^{+0.29}_{-0.24}$	0.02
$1.1 < q^2 < 6.0$	0.00	$^{+0.12}_{-0.10}$	0.02
$15.0 < q^2 < 19.0$	0.06	$^{+0.10}_{-0.09}$	0.02

Appendix C.

Tabulated A_{FB} and F_{H} results

Table C.1: The parameter F_{H} for the decay $B^0 \rightarrow K_{\text{s}}^0 \mu^+ \mu^-$ in q^2 bins. In addition to the narrow binning used in the analysis, results are also given in the theoretically favoured $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ bin and a single wide bin at low recoil, $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The first uncertainty is statistical and the second systematic in nature.

$q^2 (\text{GeV}^2/c^4)$	F_{H} (stat)	F_{H} (syst)
0.1 – 4.0	[+0.22, +1.46]	± 0.28
4.0 – 8.0	[+0.13, +0.85]	± 0.08
11.0 – 12.5	[+0.20, +1.47]	± 0.20
15.0 – 17.0	[+0.12, +0.77]	± 0.07
17.0 – 22.0	[+0.00, +0.58]	± 0.04
1.1 – 6.0	[+0.32, +1.24]	± 0.09
15.0 – 22.0	[+0.09, +0.59]	± 0.03

Table C.2: Forward-backward asymmetry, A_{FB} , and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 bins used in this analysis. These parameters are also given in the theoretically favoured $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ bin and a single wide bin at low recoil, $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The first uncertainty is statistical and the second systematic in nature.

$q^2 (\text{GeV}^2/c^4)$	F_{H} (stat)	F_{H} (syst)	A_{FB} (stat)	A_{FB} (syst)
0.10 – 0.98	[+0.01, +0.20]	± 0.03	[−0.09, −0.01]	± 0.01
1.10 – 2.00	[+0.00, +0.21]	± 0.03	[+0.00, +0.10]	± 0.01
2.00 – 3.00	[+0.05, +0.30]	± 0.03	[+0.01, +0.11]	± 0.01
3.00 – 4.00	[+0.00, +0.04]	± 0.02	[−0.02, +0.01]	± 0.01
4.00 – 5.00	[+0.00, +0.09]	± 0.03	[−0.01, +0.05]	± 0.01
5.00 – 6.00	[+0.00, +0.14]	± 0.02	[−0.04, +0.04]	± 0.01
6.00 – 7.00	[+0.00, +0.08]	± 0.02	[−0.01, +0.04]	± 0.01
7.00 – 8.00	[+0.00, +0.03]	± 0.03	[−0.02, +0.02]	± 0.01
11.00 – 11.75	[+0.06, +0.23]	± 0.03	[+0.03, +0.12]	± 0.01
11.75 – 12.50	[+0.00, +0.10]	± 0.02	[+0.00, +0.05]	± 0.01
15.00 – 16.00	[+0.06, +0.20]	± 0.02	[−0.10, −0.03]	± 0.01
16.00 – 17.00	[+0.00, +0.12]	± 0.02	[−0.05, +0.00]	± 0.01
17.00 – 18.00	[+0.01, +0.16]	± 0.02	[−0.06, +0.00]	± 0.01
18.00 – 19.00	[+0.05, +0.23]	± 0.02	[−0.03, +0.05]	± 0.01
19.00 – 20.00	[+0.00, +0.10]	± 0.04	[−0.02, +0.05]	± 0.02
20.00 – 21.00	[+0.00, +0.14]	± 0.04	[−0.01, +0.07]	± 0.02
21.00 – 22.00	[+0.04, +0.41]	± 0.05	[+0.03, +0.19]	± 0.02
1.10 – 6.00	[+0.00, +0.06]	± 0.02	[−0.01, +0.02]	± 0.01
15.00 – 22.00	[+0.00, +0.07]	± 0.02	[−0.03, +0.00]	± 0.01

Appendix D.

Two-dimensional Feldman-Cousins intervals

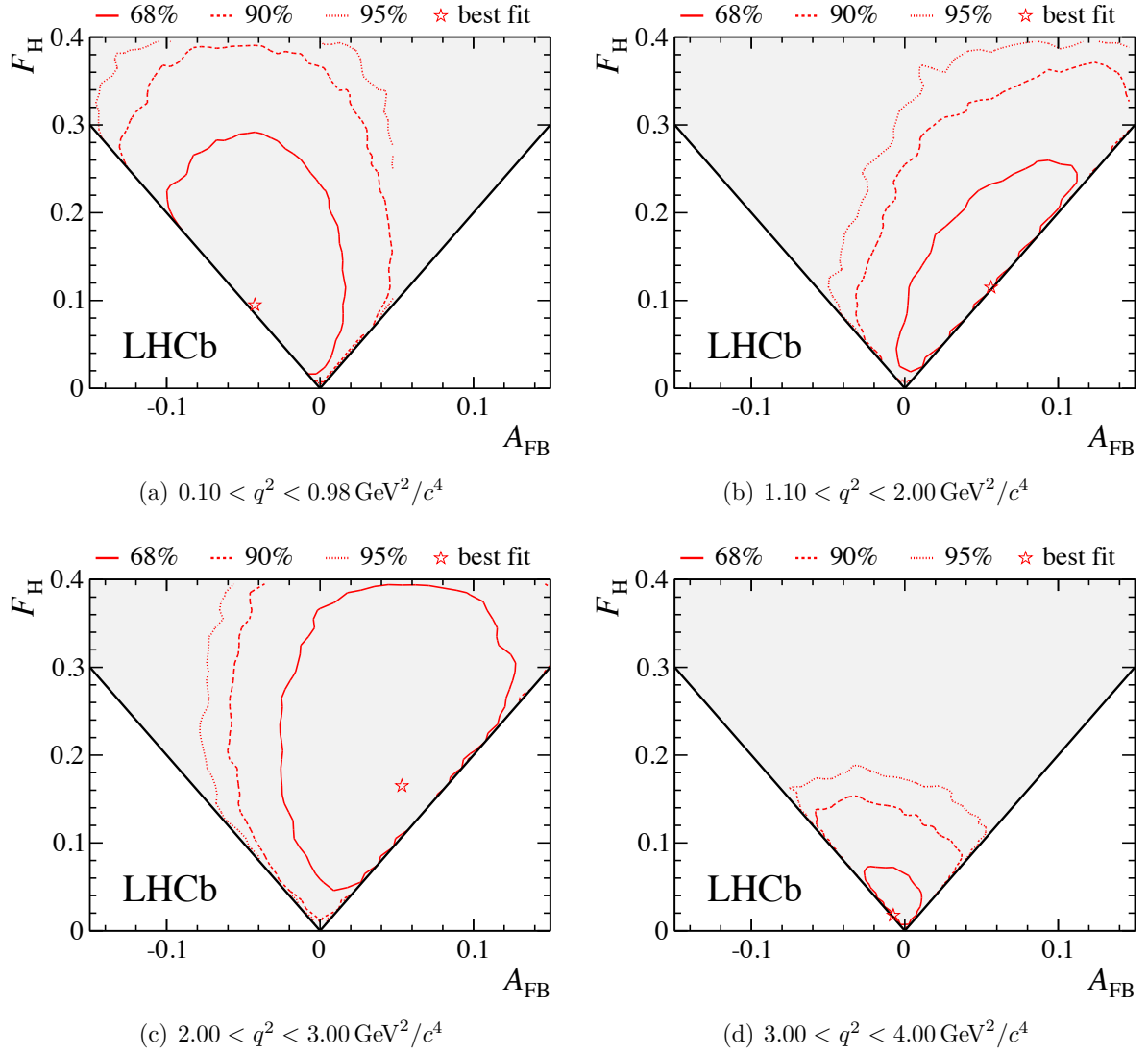


Figure D.1: Two dimensional confidence regions for A_{FB} and F_H for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $0.10 < q^2 < 0.98 \text{ GeV}^2/c^4$, (b) $1.10 < q^2 < 2.00 \text{ GeV}^2/c^4$, (c) $2.00 < q^2 < 3.00 \text{ GeV}^2/c^4$ and (d) $3.00 < q^2 < 4.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_H over which the signal angular distribution remains positive in all regions of phase-space.

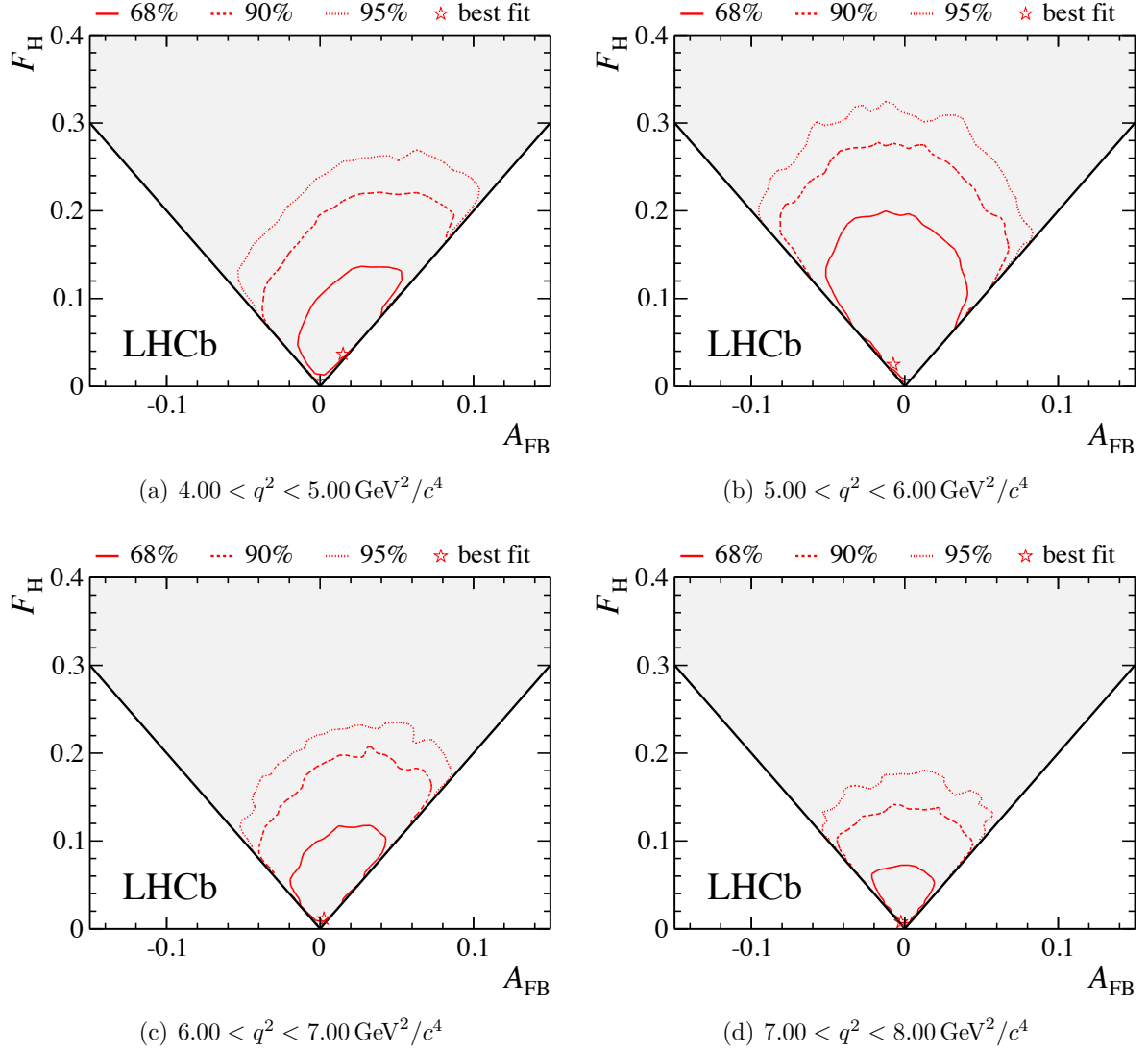


Figure D.2: Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $4.00 < q^2 < 5.00 \text{ GeV}^2/c^4$, (b) $5.00 < q^2 < 6.00 \text{ GeV}^2/c^4$, (c) $6.00 < q^2 < 7.00 \text{ GeV}^2/c^4$ and (d) $7.00 < q^2 < 8.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.

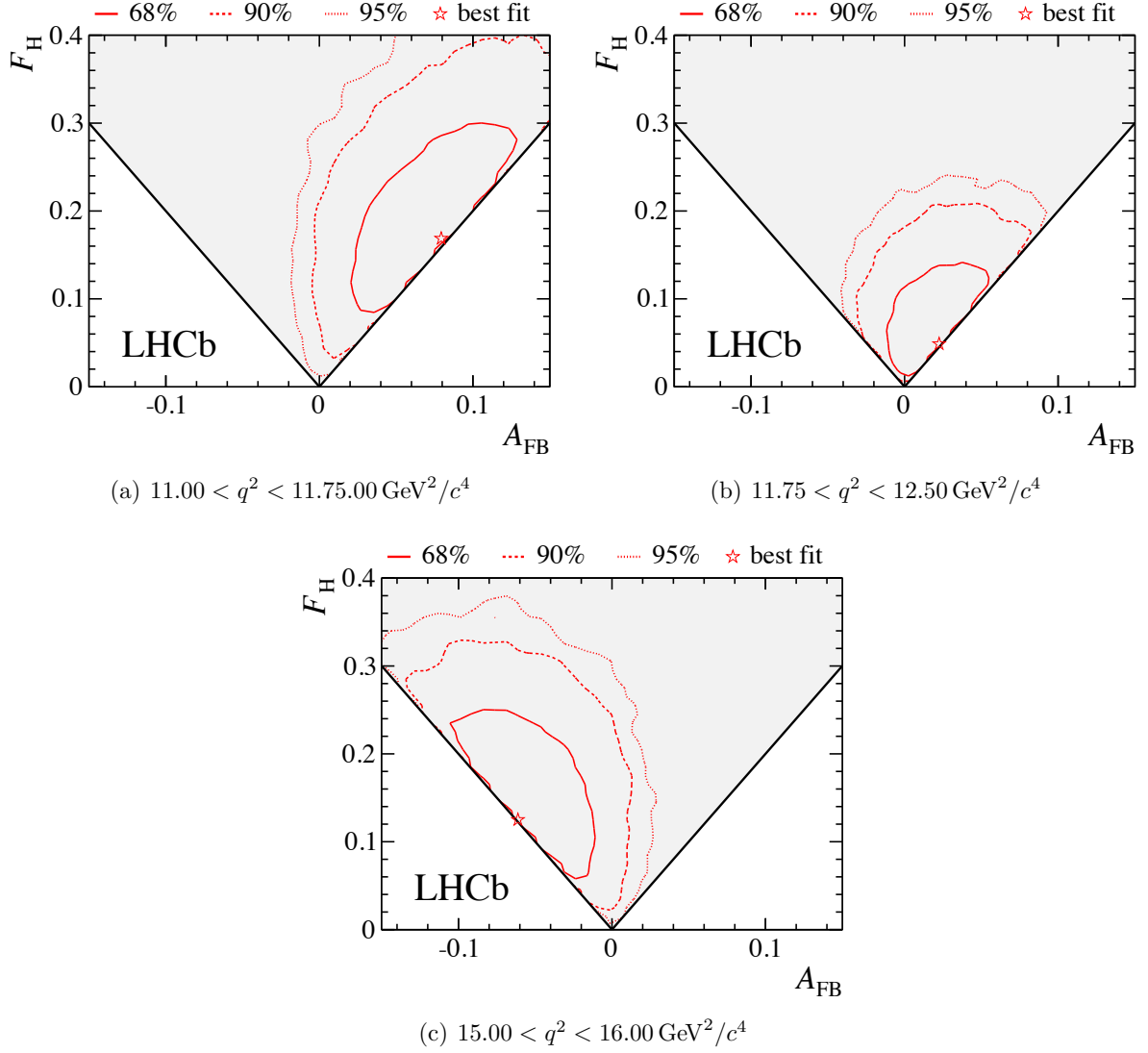


Figure D.3: Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $11.00 < q^2 < 11.75 \text{ GeV}^2/c^4$, (b) $11.75 < q^2 < 12.50 \text{ GeV}^2/c^4$ and (c) $15.00 < q^2 < 16.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.

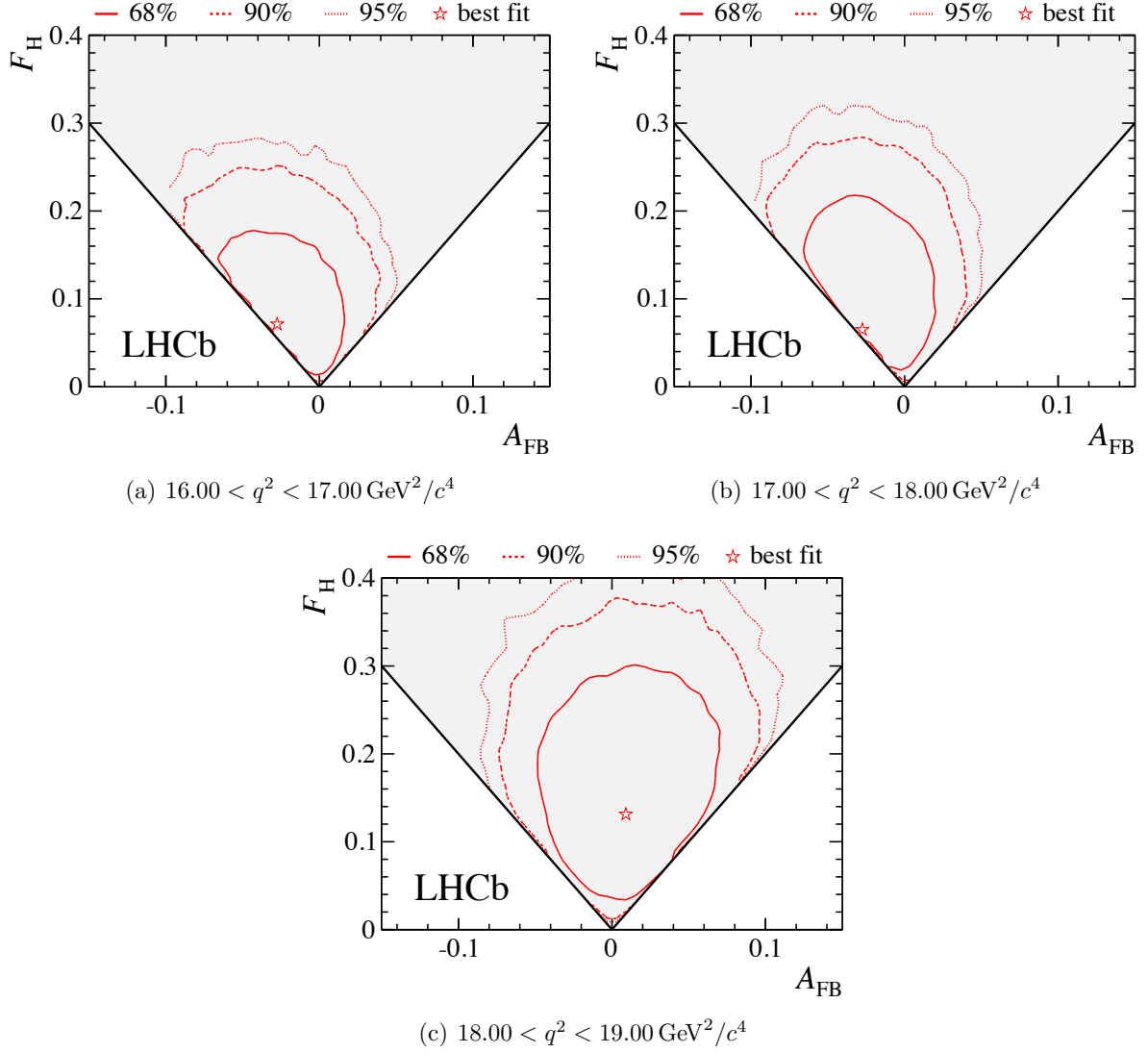


Figure D.4: Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $16.00 < q^2 < 17.00 \text{ GeV}^2/c^4$, (b) $17.00 < q^2 < 18.00 \text{ GeV}^2/c^4$ and (c) $18.00 < q^2 < 19.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.

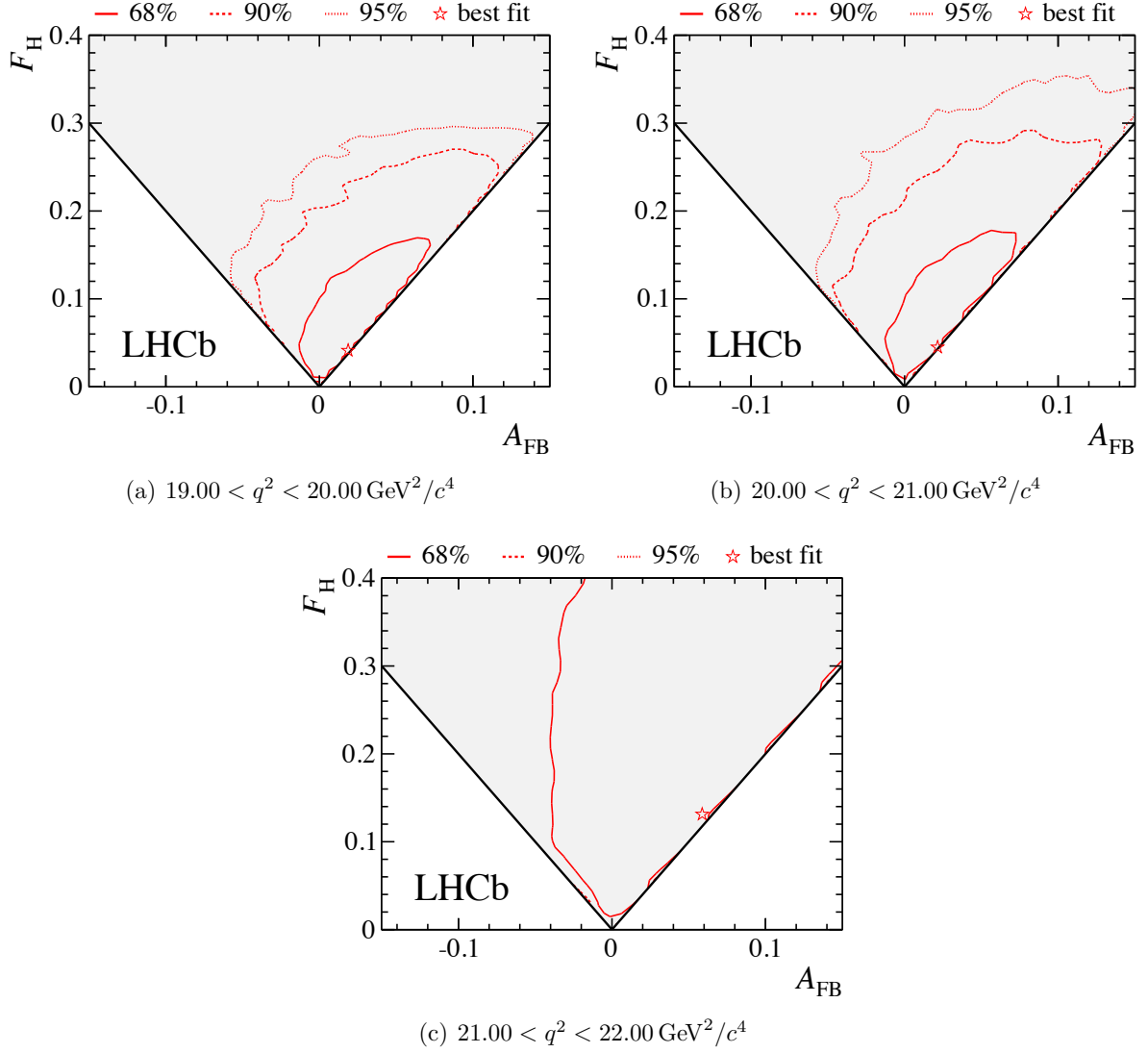


Figure D.5: Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $19.00 < q^2 < 20.00 \text{ GeV}^2/c^4$, (b) $20.00 < q^2 < 21.00 \text{ GeV}^2/c^4$ and (c) $21.00 < q^2 < 22.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.

Appendix E.

License

3/14/2014

Rightslink Printable License

ELSEVIER LICENSE TERMS AND CONDITIONS

Mar 14, 2014

This is a License Agreement between Patrick Owen ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	Patrick Owen
Customer address	8 Rahere Road Oxford, OX4 3QG
License number	3347621337112
License date	Mar 14, 2014
Licensed content publisher	Elsevier
Licensed content publication	Physics Letters B
Licensed content title	Determination of the $\psi(3770)$, $\psi(4040)$, $\psi(4160)$ and $\psi(4415)$ resonance parameters
Licensed content author	M. Ablikim, J.Z. Bai, Y. Ban, X. Cai, H.F. Chen, H.S. Chen, H.X. Chen, J.C. Chen, Jin Chen, Y.B. Chen, Y.P. Chu, Y.S. Dai, L.Y. Diao, Z.Y. Deng, Q.F. Dong, S.X. Du, J. Fang, S.S. Fang, C.D. Fu, C.S. Gao, Y.N. Gao, S.D. Gu, Y.T. Gu, Y.N. Guo, Z.J. Guo, F.A. Harris, K.L. He, et al.
Licensed content date	28 February 2008
Licensed content volume number	660
Licensed content issue number	4
Number of pages	5
Start Page	315
End Page	319
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No

Figure E.1: License to re-use Fig. 5.1 in chapter 5.



AMERICAN PHYSICAL SOCIETY

One Physics Ellipse, College Park, MD 20740 · <http://www.aps.org>

March 27, 2014

Patrick Owen
patrickowen22@gmail.com**Ref # 26287**

Thank you for your permission request dated Mar. 18, 2014. We are pleased to grant you a non-exclusive, non-transferable permission, English rights, limited to **print and electronic format**, provided you meet the criteria outlined below. Permission is for a one-time use and does not include permission for future editions, updates, databases, translations, or any other matters. Permission must be sought for each additional use. This permission does not include the right to modify APS material.

Please print the required copyright credit line on the first page that the material appears: "Reprinted (abstract/excerpt/figure) with permission from [FULL REFERENCE CITATION] as follows: authors names, journal title, volume number, page number and year of publication. Copyright (YEAR) by the American Physical Society.

The following language must appear somewhere on the website: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or part, without prior written permission from the American Physical Society."

Provide a hyperlink from the reprinted APS material (the hyperlink may be embedded in the copyright credit line). APS's link manager technology makes it convenient and easy to provide links to individual articles in APS journals. For information, see: <http://link.aps.org/>.

You must also obtain permission from at least one of the authors for each separate work, if you haven't done so already. The author's name and address can be found on the first page of the published Article.

Use of the APS material must not imply any endorsement by the American Physical Society.

Permission is granted for use of the following APS material only:

- Fig. 1, Phys. Rev. D 88, 094004 – Published 11 November 2013

Permission is limited to the single title specified of the publication as follows:

A thesis entitled "Measurement of branching fractions, isospin asymmetries and angular observables in exclusive electroweak penguin decays" to be published by Imperial College London.

If you have any questions, please refer to the Copyright FAQ at: <http://publish.aps.org/copyrightFAQ.html> or send an email to assocpub@aps.org.

Sincerely,

A handwritten signature in black ink that reads "Jamie L. Casey".

Jamie L. Casey
Circulation and Fulfillment Assistant

Figure E.2: License to re-use Figs. 6.1 and 6.2 in chapter 6.

List of Figures

2.1.	Box diagram which contributes to neutral kaon mixing in the SM.	9
2.2.	Diagram contributing to B mixing in the full (left) and effective (right) theories. .	9
2.3.	The $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay in the lowest order SM diagrams. The left is the penguin diagram, which contributes to the Wilson coefficients $C_{7,9,10}$. The right is the box diagram, which contributes to $C_{9,10}$	11
2.4.	Distribution of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^0 \rightarrow K^0 \mu^+ \mu^-$ decays using the model described in Ref. [29]. There is a photon pole as $q^2 \rightarrow 0$ for the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay. . . .	12
3.1.	Schematic diagram of the LHCb detector.	15
3.2.	Schematic of one the VELO $r\text{-}\phi$ modules, showing its retractible nature.	16
3.3.	Schematic of the LHCb tracking system, showing the different track types present in the LHCb reconstruction. Analyses in this thesis use the long and downstream track types.	17
3.4.	Cherenkov angle of different particle species for the three RICH radiators (left) and performance (right) as a function of momentum, taken from Ref. [41].	18
3.5.	Muon identification efficiency for muons (left) and pions (right), taken from Ref. [42].	19
3.6.	Diagram of the trigger logic for typical 2012 running conditions.	21
3.7.	Distribution of the number of tracks in the event for simulation and data.	22
4.1.	Tracking efficiency ratio between data and simulation for long tracks using the J/ψ tag-and-probe methods described in the text. The tracking efficiency ratio is consistent with unity in all bins.	25
4.2.	Ratio of yields between the DD and LL K_S^0 reconstruction categories for $B^0 \rightarrow J/\psi K_S^0$ decays.	25

4.3.	Momentum resolution of the missing downstream track for the solving method. Top left shows the resolution using the truth information (closure test). Top right uses truth for the downstream track and vertex information, bottom left uses true vertex information and bottom right uses all reconstructed quantities.	30
4.4.	Resolution for the kinematic fitting compared to the solving method.	31
4.5.	Momentum distribution of fully reconstructed data before and after unfolding. The unfolded and fully reconstructed momentum distributions agree perfectly by construction.	32
4.6.	Mass fit to $m_{D^{*+}} - m_{D^0}$ in “mixed events”, where a ϕ + downstream track and π_s have been combined from different events. The solving method described in Sect. 4.3.1 is applied to these mixed candidates to form the $m_{D^{*+}} - m_{D^0}$ mass distribution.	33
4.7.	Mass fit to selected $D^0 \rightarrow \phi\pi^+\pi^-$ decays, where the ϕ has been combined with a random downstream pion (π_D) from the data. This background peaks in a similar way to signal.	34
4.8.	Mass fit to fully reconstructed signal, where one of the downstream pions is ignored and had its momentum estimated using the solving method as described in Sect. 4.3.1.	36
4.9.	Mass fit to the partially reconstructed data for an example momentum bin.	36
4.10.	Mass fit to the fully reconstructed data in an example bin of track momentum.	37
4.11.	Mass fits to the fully reconstructed data before (left) and after (right) the standard K_s^0 selection criteria has been applied.	38
4.12.	Downstream tracking efficiency in bins of track momentum. Results are shown for data (left) and simulation (right).	39
4.13.	Downstream tracking efficiency ratio between data and simulation in bins of track momentum. Results are shown for raw simulation and also for simulation where the number of VELO tracks has been re-weighted to match data.	39
4.14.	Downstream tracking efficiency in bins of track momentum. Results are shown where both data and simulation have been unfolded, where only data is unfolded, and where neither have been unfolded.	40
4.15.	Efficiency to apply the standard K_s^0 selection criteria as a function of the K_s^0 momentum. Results are shown for both data (left) and simulation (right).	41
4.16.	Efficiency ratio to apply the standard K_s^0 selection as a function of K_s^0 momentum between data and simulation.	42
4.17.	Vertex χ^2 for K_s^0 in data (s-weighted) and simulation. A black line is placed where the selection is applied for standard K_s^0 candidates.	42

4.18. The branching fraction ratio, $\mathcal{B}(B^0 \rightarrow J/\psi K^0)/\mathcal{B}(B^+ \rightarrow J/\psi K^+)$, as a function of K_s^0 momentum. After correcting for the K_s^0 reconstruction, the DD category agrees with the world average and the theoretical expectation of zero isospin asymmetry between the modes.	44
5.1. Fit to the e^+e^- hadronic cross section data. Figure and fit from Ref. [55]. Note the strong interference contribution labelled R_{int}	46
5.2. Differential branching fraction of the $B^+ \rightarrow K^+\mu^+\mu^-$ decay in the low recoil region. The 3fb^{-1} result used for this analysis is overlaid on the data from the LHCb analysis based on 1fb^{-1} [60]. For the 3fb^{-1} result, only statistical uncertainties are shown. The bands shown indicate the theoretical prediction for the differential branching fraction and are calculated using input from Ref. [61].	47
5.3. Trigger efficiency of $B^+ \rightarrow K^+\mu^+\mu^-$ decays as a function of q^2 . The trigger is dominated by the muons, which have higher p_T at high q^2	48
5.4. Trigger efficiency of $B^+ \rightarrow J/\psi K^+$ decays for each trigger level as a function of muon p_T . The trigger is calculated by studying events which have been triggered independently of signal.	49
5.5. Invariant mass of the $K^+\mu^-$ system under the $\mu^+\mu^-$ mass hypothesis before (left) and after (right) the veto described in the text. There is no evidence for kaon-muon swaps after the veto.	50
5.6. BDT response for the signal and background samples. Each sample is split into a training and testing part, the latter is used to check performance and optimise the cut placed on the BDT response.	51
5.7. Efficiency of the BDT as a function of q^2 for $B^+ \rightarrow K^+\mu^+\mu^-$ decays. In general, the BDT efficiency is very good, reflecting the favourable topology of $B^+ \rightarrow K^+\mu^+\mu^-$ decays.	52
5.8. Mass fit to $B^+ \rightarrow \psi(2S)K^+$ decays, which is where the signal shape is obtained from. The lower mass sideband is very short to avoid background from partially reconstructed decays.	52
5.9. Mass fit to $B^+ \rightarrow K^+\mu^+\mu^-$ decays in the low recoil region, which is used to estimate the amount of background in the $\mu^+\mu^-$ spectrum. The lower mass sideband is very short to avoid background from partially reconstructed decays.	53
5.10. Dimuon mass resolution of $B^+ \rightarrow K^+\mu^+\mu^-$ decays as a function of mass, obtained with the simulation. Results are shown before and after the kinematic constraint described in the text.	54
5.11. Fit to the unconstrained dimuon mass in the sideband region of the $K^+\mu^+\mu^-$ mass. The shape is parameterised by an ARGUS function [64].	54

- 5.12. Relative efficiency between $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ decays as a function of dimuon mass. The efficiency drops at high mass due to IP selection requirements. 55
- 5.13. Ratio of the $K^+ \chi_{\text{IP}}^2$ distributions in the data and simulation for $B^+ \rightarrow J/\psi K^+$ decays. 56
- 5.14. Vector fraction of the non-resonant $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay as a function of q^2 . Compared to the experimental sensitivity, the fraction is essentially constant with dimuon mass. 56
- 5.15. Dimuon mass distribution of data with fit results overlaid for the fit that includes contributions from the non-resonant vector and axial vector components, the $\psi(3770)$ meson, and one additional resonance with free width and mean parameters. Interference terms are included and the relative strong phases are left free in the fit. 58
- 5.16. Distribution of the log likelihood ratio from fits that include and exclude a single additional resonance component for pseudo-experiments that only include the non-resonant, $\psi(3770)$ and background components. The dashed blue line corresponds to the fit of an exponential function which is used to extrapolate to larger values of $2\log(L_{\text{s+b}}/L_{\text{b}})$. The vertical black line denotes the observed likelihood ratio. A signal significance exceeding six standard deviations is observed. 59
- 5.17. Profile likelihood as a function of mass and width of a fit with a single extra resonance. At each point all other fit parameters are re-optimised. The three ellipses are (red-solid) the best fit and previous measurements of (grey-dashed) the $\psi(4160)$ [55] and (black-dotted) the $Y(4260)$ [14] states. 59
- 5.18. Dimuon mass distribution of data with fit results overlaid for the fit that includes contributions from the non-resonant vector and axial vector components, and the $\psi(3770)$, $\psi(4040)$ and $\psi(4160)$ resonances. Interference terms are included and the relative strong phases are left free in the fit. 60
- 5.19. Profile likelihood ratios for the product of branching fractions $\mathcal{B}(B^+ \rightarrow \psi K^+) \times \mathcal{B}(\psi \rightarrow \mu^+ \mu^-)$ of the $\psi(4040)$ and the $\psi(4160)$ mesons. At each point all other fit parameters are reoptimised. 61
- 6.1. Diagram for $b \rightarrow s \mu^+ \mu^-$ involving the Wilson Coefficient $\mathcal{C}_8$, taken from Ref. [68]. The difference in electric charge between the up quark and down quarks contributes to the isospin asymmetry. The crosses denote where a photon can be radiated to produce the dimuon pair. 64
- 6.2. Diagram for $b \rightarrow s \mu^+ \mu^-$ involving the Wilson Coefficients $\mathcal{C}_{1,2}$, taken from Ref. [68]. The crosses denote where a photon can be radiated to produce the dimuon pair. 65

6.3. Diagram contributing to the isospin asymmetry which involves the new physics model proposed in Ref. [71]. The particle labelled as the A_2^3 is a gauge boson which can change flavours between the second and third quark generations.	65
6.4. Most precise measurements of the isospin asymmetry prior to this analysis. The $B \rightarrow K\mu^+\mu^-$ results tend to be negative whereas the $B \rightarrow K^*\mu^+\mu^-$ are consistent with the SM prediction of no isospin asymmetry.	66
6.5. BDT response for the signal and background samples for the $B^0 \rightarrow K_s^0\mu^+\mu^-$ decay in the LL category (left) and DD category (right). The samples are split into training and testing samples, the latter of which is used to optimise the BDT selection cut.	67
6.6. Dimuon mass versus $K^{(*)}\mu^+\mu^-$ mass distribution of the four signal decays.	69
6.7. Mass fits to $B^0 \rightarrow J/\psi K_s^0$ decays in the LL category (left) and DD category (right). The fit model is described in the text.	70
6.8. Mean and width of the $B^0 \rightarrow K_s^0\mu^+\mu^-$ mass distribution in the simulation as a function of q^2	70
6.9. Signal yields for $B^0 \rightarrow K_s^0\mu^+\mu^-$ in the LL and DD categories before and after the correction to the changes in mean and width as a function of q^2 . The corrections to the mass shape make a very small difference to the signal yield.	71
6.10. Reconstructed B candidate mass for the four signal modes. The data are overlaid with the result of the fit described in the text. The long and downstream K_s^0 categories have been combined. The fits are performed in separate q^2 bins and subsequently merged.	72
6.11. Ratio of the K_s^0 momentum distributions in $B^0 \rightarrow J/\psi K_s^0$ decays between data and simulation.	74
6.12. Relative efficiency between the signal and normalisation channels as a function of q^2 , obtained from the simulation.	74
6.13. Differential branching fraction for $B^0 \rightarrow K^0\mu^+\mu^-$ (left) and $B^+ \rightarrow K^+\mu^+\mu^-$ (right) decays. The shaded region illustrates the theoretical prediction with its uncertainty (see text for details).	77
6.14. Differential branching fraction for $B^0 \rightarrow K^{*0}\mu^+\mu^-$ (left) and $B^+ \rightarrow K^{*+}\mu^+\mu^-$ (right) decays. The shaded regions illustrate the theoretical prediction with its uncertainty (see text for details).	77
6.15. Isospin asymmetries for $B \rightarrow K\mu^+\mu^-$ (left) and $B \rightarrow K^*\mu^+\mu^-$ (right) decays.	79
6.16. Isospin asymmetry of the old 2011 analysis and the 2011 part of the new analysis. The compatibility between the two is very good (93%).	80
6.17. Isospin asymmetry results as measured LL and DD categories.	80

6.18. Isospin asymmetry for the 2011 and 2012 datasets. The 2012 results are consistently above the 2011 ones.	81
6.19. Isospin asymmetry for the 2011 and 2012 datasets for the LL (left) and DD (right) K_s^0 reconstruction categories.	81
6.20. Isospin asymmetry under different mass ranges (5170-5700 MeV/ c^2 is nominal). . .	82
6.21. Isospin asymmetry results when applying the trigger requirements on the dimuon candidate.	83
7.1. Most precise measurements of the angular observables of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays prior to the analysis described in this chapter, taken from Ref. [60].	85
7.2. Typical likelihood surface for the angular observables A_{FB} and F_H for the $B^+ \rightarrow K^+ \mu^+ \mu^-$ mode. The PDF becomes negative in the white region.	86
7.3. Efficiency of the muon (left) and hadron (right) PID selection criteria as a function of $\cos \theta_l$. The shape can be considered a flat line when considering the statistical precision of the data.	87
7.4. Distribution of $\cos \theta_l$ for $B^+ \rightarrow J/\psi K^+$ in data and simulation. There is dip corresponding to the $B^+ \rightarrow \bar{D}^0 \pi^+$ veto.	88
7.5. Efficiency as a function of $\cos \theta_l$ for $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ decays in the wide q^2 bins. The step function is required in the $B^+ \rightarrow K^+ \mu^+ \mu^-$ case to reproduce the effect of the $B^+ \rightarrow \bar{D}^0 \pi^+$ veto.	88
7.6. Distribution of the muon momentum for $B^+ \rightarrow J/\psi K^+$ decays in data and simulation. There is a good agreement across the full kinematic range.	89
7.7. Top, reconstructed mass of $B^+ \rightarrow K^+ \mu^+ \mu^-$ candidates in the ranges (a) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (b) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. Bottom, reconstructed mass of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ candidates in the ranges (c) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (d) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The data are overlaid with the result of the fit described in the text. The shaded region indicates the background contribution in the fit. . .	91
7.8. Top, angular distribution of $B^+ \rightarrow K^+ \mu^+ \mu^-$ candidates with (a) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (b) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. Bottom, angular distribution of $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ candidates with (c) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (d) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. Only candidates with a reconstructed mass within $\pm 50 \text{ MeV}/c^2$ of the known B^+ or B^0 mass are shown. The data are overlaid with the result of the fit described in the text. The shaded region indicates the background contribution in the fit.	92
7.9. Fraction of toy datasets which have a DLL bigger than the one in the data for $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the first q^2 bin.	93

7.10. Statistical uncertainties for F_H in bins of q^2 obtained from the profile likelihood and obtained using the Feldman-Cousins technique.	93
7.11. Two dimensional confidence regions for A_{FB} and F_H for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ and (b) $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_H over which the signal angular distribution remains positive in all regions of phase-space.	94
7.12. Dimuon forward-backward asymmetry, A_{FB} , and the parameter F_H for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ as a function of the dimuon invariant mass squared, q^2 . The inner horizontal bars indicate the one-dimensional 68% statistical confidence intervals. The outer horizontal bars include contributions from systematic uncertainties (described in the text). The confidence intervals for F_H are overlaid with the SM theory prediction (narrow band). Data are not presented for the regions around the J/ψ and $\psi(2S)$ resonances.	95
7.13. Results for the parameter F_H for the decay $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ as a function of the dimuon invariant mass squared, q^2 . The inner horizontal bars indicate the one-dimensional 68% statistical confidence intervals. The outer horizontal bars include contributions from systematic uncertainties (described in the text). The confidence intervals are overlaid with the SM theory prediction (narrow band). Data are not presented for the regions around the J/ψ and $\psi(2S)$ resonances.	95
7.14. Fits to the $\cos \theta_l$ distribution in the background region where the BDT is loosened (left) and the nominal fit (right). The loosened BDT PDF has four parameters whereas the nominal is assumed to be flat (reminder: the PDF is multiplied by the efficiency and so does not look perfectly flat here). For the associated systematic uncertainty, datasets are generated from the left PDF and fit with both PDFs, where the difference in F_H is used to quantify the uncertainty.	96
7.15. Difference between the F_H best fit results in $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ toy datasets for two different background parameterisations. The toy datasets are generated from the “full” parameterisation. The “simple” parameterisation is used for the final result, where a few parameters are let free. The “full” parameterisation is fixed from a fit to the background at a looser BDT cut, where four parameters are used.	97
8.1. Ratio of the measured branching fractions to the SM prediction for $B \rightarrow K^{(*)} \mu^+ \mu^-$ and $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays. The red line is the result a global fit based on the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ angular distribution.	99
8.2. Isospin asymmetry in $B \rightarrow K \mu^+ \mu^-$ decays as a function of q^2 as obtained from the LHCb [2], Belle [73] and BaBar [72] collaborations. The right figure shows a naive combination of these three measurements, assuming uncorrelated systematic uncertainties.	101

D.1. Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $0.10 < q^2 < 0.98 \text{ GeV}^2/c^4$, (b) $1.10 < q^2 < 2.00 \text{ GeV}^2/c^4$, (c) $2.00 < q^2 < 3.00 \text{ GeV}^2/c^4$ and (d) $3.00 < q^2 < 4.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.	118
D.2. Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $4.00 < q^2 < 5.00 \text{ GeV}^2/c^4$, (b) $5.00 < q^2 < 6.00 \text{ GeV}^2/c^4$, (c) $6.00 < q^2 < 7.00 \text{ GeV}^2/c^4$ and (d) $7.00 < q^2 < 8.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.	119
D.3. Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $11.00 < q^2 < 11.75 \text{ GeV}^2/c^4$, (b) $11.75 < q^2 < 12.50 \text{ GeV}^2/c^4$ and (c) $15.00 < q^2 < 16.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.	120
D.4. Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $16.00 < q^2 < 17.00 \text{ GeV}^2/c^4$, (b) $17.00 < q^2 < 18.00 \text{ GeV}^2/c^4$ and (c) $18.00 < q^2 < 19.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.	121
D.5. Two dimensional confidence regions for A_{FB} and F_{H} for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 ranges (a) $19.00 < q^2 < 20.00 \text{ GeV}^2/c^4$, (b) $20.00 < q^2 < 21.00 \text{ GeV}^2/c^4$ and (c) $21.00 < q^2 < 22.00 \text{ GeV}^2/c^4$. The purely statistical confidence intervals are determined using the Feldman-Cousins technique. The shaded (triangular) region illustrates the range of A_{FB} and F_{H} over which the signal angular distribution remains positive in all regions of phase-space.	122
E.1. License to re-use Fig. 5.1 in chapter 5.	124
E.2. License to re-use Figs. 6.1 and 6.2 in chapter 6.	125

List of Tables

4.1.	Kinematic selection criteria for the inclusive D^{*+} trigger lines, which were used for the K_s^0 reconstruction analysis.	26
4.2.	Stripping selection criteria for the K_s^0 reconstruction analysis. Note the tight p_T requirements on kaons and prescale. The χ_{IP}^2 (χ_{FD}^2) variable is the increase in vertex χ^2 if the signal track(s) is combined with the PV, and θ is the angle between signal candidate direction and the line made by the decay vertex and the PV.	27
4.3.	Offline selection criteria on the downstream track. The ProbNNpi variable is a neural-net based PID criterion.	28
4.4.	Criteria to make a standard LHCb K_s^0 candidate from two downstream pions. Track DOCA is the distance of closest approach between the two daughter tracks.	38
5.1.	The relevant resonance parameters of the fit to the BES data [55]. Note the rather misleading name of the $\psi(4160)$	46
5.2.	Variables used in the BDT ordered by importance according to the BDT. The χ_{IP}^2 (χ_{FD}^2) variable is the increase in vertex χ^2 if the signal track(s) is combined with the PV, and θ is the angle between signal candidate direction and the line made by the decay vertex and the PV.	51
5.3.	Parameters of the dominant resonance for fits where the mass and width are unconstrained and constrained to those of the $\psi(4160)$ meson [55], respectively. The branching fractions are for the B^+ decay followed by the decay of the resonance to muons.	60
6.1.	Observed yields of the four signal channels. See text for details.	71
6.2.	Integrated branching fractions in the low recoil region. For $B \rightarrow K^* \mu^+ \mu^-$ this is defined as $15-19 \text{ GeV}^2/c^4$ whereas for $B \rightarrow K \mu^+ \mu^-$ it is $15-22 \text{ GeV}^2/c^4$. Predictions are obtained using the form factors calculated in LQCD as described in Refs. [34,35,66,67].	78
6.3.	Ratio of branching fractions of $B^0 \rightarrow \psi(2S) K_s^0$ to $B^0 \rightarrow J/\psi K_s^0$ decays, where the $J/\psi \rightarrow \mu^+ \mu^-$ and $\psi(2S) \rightarrow \mu^+ \mu^-$ branching fractions are not taken into account.	83

A.1. Differential branching fraction results (10^{-9}) for $B^0 \rightarrow K^0 \mu^+ \mu^-$ including statistical and systematic uncertainties.	110
A.2. Differential branching fraction results (10^{-9}) for $B^+ \rightarrow K^+ \mu^+ \mu^-$ including statistical and systematic uncertainties.	111
A.3. Differential branching fraction results (10^{-9}) for $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ including statistical and systematic uncertainties.	112
B.1. Isospin asymmetry results for $B \rightarrow K \mu^+ \mu^-$ decays.	113
B.2. Isospin asymmetry results for $B \rightarrow K^* \mu^+ \mu^-$ decays.	114
C.1. The parameter F_H for the decay $B^0 \rightarrow K_s^0 \mu^+ \mu^-$ in q^2 bins. In addition to the narrow binning used in the analysis, results are also given in the theoretically favoured $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ bin and a single wide bin at low recoil, $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The first uncertainty is statistical and the second systematic in nature.	115
C.2. Forward-backward asymmetry, A_{FB} , and F_H for the decay $B^+ \rightarrow K^+ \mu^+ \mu^-$ in the q^2 bins used in this analysis. These parameters are also given in the theoretically favoured $1.1 < q^2 < 6.0 \text{ GeV}^2/c^4$ bin and a single wide bin at low recoil, $15.0 < q^2 < 22.0 \text{ GeV}^2/c^4$. The first uncertainty is statistical and the second systematic in nature.	116