Measurement of the decay $B_s^0 \to \phi\mu^+\mu^-$ at LHCb

Paul Nicholas Schaack

High Energy Physics
Blackett Laboratory
Imperial College London

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

May 2012

Declaration

The work presented in this thesis was carried out between October 2008 and May 2012. It is the result of my own studies, with the support of members of the Imperial College HEP group and the broader LHCb collaboration. The work of others is explicitly referenced.

Paul N. Schaack, 30 May 2012

Abstract

The measurements of two branching fraction ratios are reported in this thesis. The first one is a measurement of the ratio of the branching fractions for the decays $B_s^0 \to \psi'\phi$ and $B_s^0 \to J/\psi\,\phi$. It is based on approximately $36\,\mathrm{pb}^{-1}$ of data collected by the LHCb detector at $\sqrt{s} = 7\,\mathrm{TeV}$ during 2010. The result is $0.68 \pm 0.10(\mathrm{stat}) \pm 0.09(\mathrm{syst}) \pm 0.07(\mathcal{B})$, which is compatible with previous measurements from experiments at the Tevatron. The second measurement is the ratio of the branching fractions for the decays $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ as a function of the dimuon invariant mass. It is based on $1\,\mathrm{fb}^{-1}$ of data collected by the LHCb detector at $\sqrt{s} = 7\,\mathrm{TeV}$ during 2011. The ratio of branching fractions is measured to be $(0.558 \pm 0.070(\mathrm{stat}) \pm 0.043(\mathrm{syst}) \pm 0.006(\mathcal{B})) \times 10^{-3}$. Using the measured branching fraction of the decay $J/\psi \to \mu^+\mu^-$ gives a branching fraction of $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-) = (0.78 \pm 0.10(\mathrm{stat}) \pm 0.06(\mathrm{syst}) \pm 0.28(\mathcal{B})) \times 10^{-6}$. This is the most precise measurement of $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)$ and is in agreement with the Standard Model expectation.

Preface

I would like to thank my supervisor, Andrey Golutvin, and the people I have worked with closely during my stay at CERN, namely Vanya Belyaev, Thomas Blake and Ilya Komarov. Their help and advice has helped and taught me numerous times. I would also like to thank my family and my friends for their moral support in the past three and a half years.

Contents

Appendix **129**

List of Figures

List of Tables

Chapter 1

Introduction

This thesis describes the measurement of the decay $B_s^0 \to \phi\mu^+\mu^-$ with the LHCb detector, in particular the ratio of branching fractions for $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ decays. Measuring the ratio is advantageous as many factors and their associated uncertainties cancel. The decay $B_s^0 \to \phi\mu^+\mu^-$ is a so-called flavour changing neutral current, which is sensitive to New Physics beyond the Standard Model of particle physics. The relevant aspects of the Standard Model for this analysis are outlined in Chapter 2. The theoretical background of flavour changing neutral currents, and a description of observables, are given in Chapter 3. The LHCb detector components relevant for this thesis are described in Chapter 4. A measurement of the ratio of branching fractions for the decays $B_s^0 \to \psi'\phi$ and $B_s^0 \to J/\psi\,\phi$ carried out in 2010 is presented in Chapter 5. With the data collected by the LHCb detector during 2011 it became possible to use the same experimental techniques for a measurement of the ratio of branching fractions for the decays $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$. This analysis is presented in Chapter 6. Finally, conclusions and an outlook are given in Chapter 7.

The idea of this measurement was developed towards the end of 2010. At that point it became clear that the LHCb detector will collect a sufficient amount of data to perform this measurement within the duration of the author's PhD. The decay had been observed by the CDF collaboration already before but its branching fraction had not been measured yet. This changed in the summer of 2011 when the CDF collaboration published the differential branching fraction measurement of the $B_s^0 \to \phi\mu^+\mu^-$ decay in bins of q^2. As a result the aim was to wait for the full 2011 LHCb data set in order to reduce the statistical uncertainty and obtain a world's most precise measurement, which was achieved in time for the Moriond conference in March 2012.

Chapter 2

The Standard Model

2.1 Introduction

The Standard Model (SM) of particle physics describes the forces and interactions at the atomic and sub-atomic scale. The forces it describes are the electromagnetic, the strong and the weak force. It does not account for the gravitational force. The SM is a local gauge theory based on the gauge group:

$$SU(3)_{\mathrm{C}} \otimes SU(2)_{\mathrm{L}} \otimes U(1)_{\mathrm{Y}} \tag{2.1}$$

where $SU(3)_{\mathrm{C}}$ describes the strong interactions (C stands for *colour*) and $SU(2)_{\mathrm{L}} \otimes U(1)_{\mathrm{Y}}$ describes the electroweak interactions (L stands for *left* handedness and Y for *hypercharge*).

The forces are mediated by spin-1 gauge fields (bosons) acting on spin-$\frac{1}{2}$ matter constituents (fermions). The fermions are organised into three generations. Each generation consists of one charged lepton, one neutrino and two quarks. An overview of all discovered fermions of the SM is given in Figure 2.1.

The dynamics of the SM are described by the fundamental Lagrangian density:

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{QCD}} + \mathcal{L}_{\mathrm{EW}} + \mathcal{L}_{\mathrm{Higgs}} \tag{2.2}$$

The origin of each component of $\mathcal{L}_{\mathrm{SM}}$ will be outlined in this chapter. First a description of the gauge sector is given. This is followed by a description of the fermion sector. Finally an overview of the problems of the SM and the experimental outlook at the Large Hadron Collider (LHC) are presented.

The SM has been developed and extended continuously over the decades and it is described in various publications and books. This chapter is mainly based on material found in references [2–7]. Specific references are given within the text.

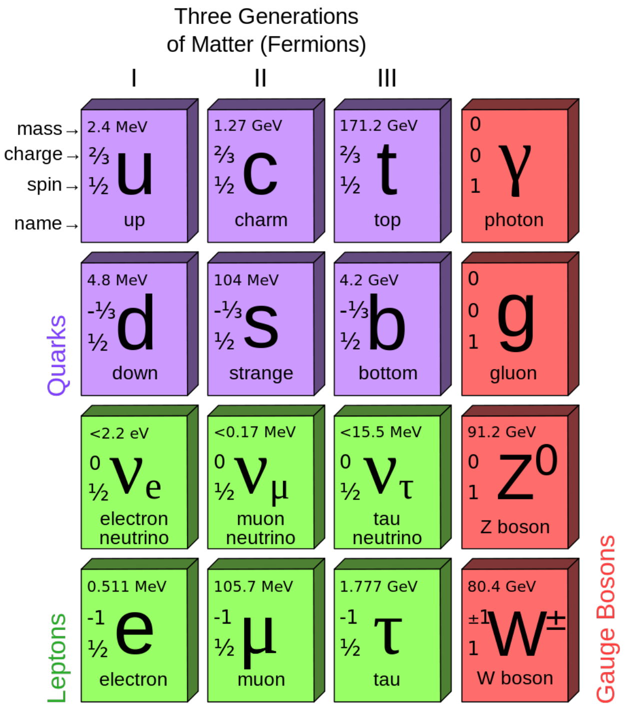


Figure 2.1: The fermions and bosons of the SM and their properties. Quarks are shaded purple, leptons green and bosons red. The fermions are grouped into the three generations. Taken from [1].

2.2 The gauge sector

In the following subsections the principles that give rise to the gauge bosons of each force are outlined.

2.2.1 The electromagnetic force

The Electromagnetic (EM) force is responsible for binding electrons and nuclei together in atoms. Its force carrier, the gauge boson of the EM force, is the photon. The photon must be massless, which also sets the range of the EM force to infinity. It is the result of *requiring* the fermion Lagrangian to be invariant under local gauge transformations, which will be outlined in the following.

The electron and positron free fermionic fields are defined as $\psi(x)$ and $\bar{\psi}(x)$. A local gauge transformation is then given by:

$$\begin{aligned} \psi(x) \longrightarrow \psi'(x) &= e^{i\alpha(x)}\psi(x) \\ \bar{\psi}(x) \longrightarrow \bar{\psi}'(x) &= e^{-i\alpha(x)}\bar{\psi}(x) \end{aligned} \tag{2.3}$$

where $\alpha(x)$ can be any function of space and/or time. The free fermion Lagrangian, given by

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi, \tag{2.4}$$

is not invariant under such transformations. In Equation 2.4 (and throughout this chapter) Greek indices denote space-time directions and imply summation and γ^{μ} are the Dirac 4×4 matrices. The local gauge transformations in Equation 2.3 introduce an extra term of the form

$$\bar{\psi}\gamma^{\mu}\psi\partial_{\mu}\alpha(x) \tag{2.5}$$

to the free fermion Lagrangian. In other words, one can say that the extra term (Equation 2.5) is the result of the locality of the field.

The Lagrangian becomes invariant by introducing a vector field A_{μ} and a covariant derivative $D_{\mu} = \partial_{\mu} - ieA_{\mu}$, where e is the electron charge and the coupling of the EM force. A_{μ} transforms as:

$$A_{\mu} \longrightarrow A'_{\mu} = A_{\mu} + \frac{1}{e}\partial_{\mu}\alpha(x) \tag{2.6}$$

This allows the free fermion Lagrangian to be rewritten as

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^{\mu}D_{\mu}\psi - m\bar{\psi}\psi \\ &= \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + e\bar{\psi}\gamma^{\mu}\psi A_{\mu}, \end{aligned} \tag{2.7}$$

which is now invariant under a local gauge transformation of ψ. The second term in Equation 2.7 can be interpreted as an interaction term between the fermion current $j = \bar{\psi}\gamma^{\mu}\psi$ and the vector field A_{μ} with strength e.

For the complete EM Lagrangian a gauge invariant kinetic term for the vector field has to be added, such that:

$$\mathcal{L}_{\mathrm{EM}} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + e\bar{\psi}\gamma^{\mu}\psi A_{\mu} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.8}$$

where $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$. $F_{\mu\nu}$ is the EM field strength tensor. The addition of a mass term for the photon field of the form $\frac{1}{2}m^2A_{\mu}A^{\mu}$ is prohibited by the principle of local gauge invariance and it is also not required as the photon is massless.

In summary the EM Lagrangian can be split into three parts: The free fermion part, the interaction part between the vector field and the fermion current with strength e and the kinetic part.

2.2.2 The strong force

The strong force exclusively acts on quarks and the theory describing the strong force is referred to as Quantum Chromodynamics (QCD). There are 6 types (or flavours) of quarks in the three families or generations of particles (2 flavours per generation: one with charge $+\frac{2}{3}e$ and the other one with $-\frac{1}{3}e$) as shown in Figure 2.1.

The equivalent of charge for the strong force is *colour*, coming in three forms: red, blue or green (the word colour here is just a label for a property that has three possible states). Therefore quarks must carry the colour charge whereas leptons do not.

The gauge bosons of the strong force are gluons, which are also massless and carry the colour charge. No free particles carry colour charge and therefore quarks and gluons are said to be confined. Because the gluons carry the charge of the theory they must also be self-interacting.

In analogy to the EM force the gluons are a result of requiring the free Lagrangian of QCD to be invariant under a local ($SU(3)$) gauge transformation. Quarks are associated with three free fermionic fields for the three colours r, b and g, arranged in a triplet:

$$\psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix} \tag{2.9}$$

$$\bar{\psi} = \begin{pmatrix} \bar{\psi}^r & \bar{\psi}^b & \bar{\psi}^g \end{pmatrix} \tag{2.10}$$

Quarks of different flavours have different masses, but quarks of the same flavour but different colour have the same mass. Thus the free QCD Lagrangian reads as:

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{2.11}$$

This is equivalent to the free EM Lagrangian in Equation 2.4, but now ψ stands for a triplet of four-component Dirac spinors.

Again, $\mathcal{L}$ is required to be invariant under a gauge transformation of the form:

$$\begin{aligned} \psi \longrightarrow \psi' &= e^{iH}\psi \\ \bar{\psi} \longrightarrow \bar{\psi}' &= \bar{\psi}e^{-iH} \end{aligned} \tag{2.12}$$

where H is a 3×3 hermitian matrix, satisfying $H^{\dagger} = H$. Thus H can be written in terms of nine real numbers ($a_1, a_2, ..., a_8$ and θ) as $H = \theta I + \lambda \cdot a$, where I is a 3×3 unit matrix. $\lambda \cdot a$ stands for the sum $\sum_{i=1}^{8} \lambda_i a_i$, where λ_i are the Gell-Mann matrices.

Therefore the transformation in Equation 2.12 can be split into a phase and a local transformation: $e^{iH} = e^{i\theta}e^{i\lambda \cdot a}$. Of interest is the latter one:

$$\psi \longrightarrow \psi' = e^{iq\lambda \cdot \phi(x)}\psi \tag{2.13}$$

where $\phi(x) = a/q$ and q is a coupling equivalent to e in the EM force. By changing the derivative, ∂_μ ($\mu = 1, 2, ..., 8$), to a covariant derivative of the form

$$D_\mu = \partial_\mu - iq\lambda \cdot A_\mu \tag{2.14}$$

the Lagrangian becomes invariant under this transformation. It takes the following form:

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \\ &= [i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi] - (q\bar{\psi}\gamma^\mu \lambda\psi) \cdot A_\mu \end{aligned} \tag{2.15}$$

where eight gauge fields, A_μ, are introduced, representing the eight gluons. Like for the EM force a gauge invariant kinetic term for the gluons is added:

$$\mathcal{L}_{\mathrm{QCD}} = [i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi] - (q\bar{\psi}\gamma^\mu \lambda\psi) \cdot A_\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.16}$$

This Lagrangian exists in 6 versions: one for each quark flavour with a distinct mass.
The strong force binds quarks into hadrons. Hadrons can either consist of two quarks (a quark-antiquark pair), called mesons, or of three quarks (either all quarks or all antiquarks), called baryons. For example the B_s^0-meson consists of a $\bar{b}s$ quark pair. Examples of baryons are protons (uud) and neutrons (udd).

2.2.3 The electroweak force

The weak force acts on all fermions. The name for the charge that 'produces' the weak force (comparable to how electric charge is responsible for the EM force and colour charge for the strong force) is the weak charge and all fermions carry it.
There are two kinds of weak interactions mediated by different gauge bosons: charged and neutral weak interactions (mediated by the W^+ and W^- and by the Z^0 respectively). Unlike the gauge bosons of the EM and the strong force the gauge bosons for the electroweak force are massive. This defines the 'short' range of the electroweak force, which is of $\mathcal{O}(10^{-18}\,\mathrm{m})$.
For weak interactions the fermionic fields for leptons and quarks can be written as left-handed $SU(2)_L$ doublets of the form:

$$\begin{aligned} l_L &= \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L \text{ or } \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \\ q_L &= \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L \text{ or } \begin{pmatrix} t \\ b \end{pmatrix}_L \end{aligned}$$

Right-handed fermions are placed in a singlet:

$$\begin{aligned} l_R &= e_R^-, \mu_R^- \text{ or } \tau_R^- \\ q_R &= u_R, d_R, c_R, s_R, t_R \text{ or } b_R \end{aligned}$$

L and R denote the left- and right-handed projections of the spinor field (chiralities), $\Psi_{L,R} = \frac{1}{2}(1 \mp \gamma_5)\Psi$.

It can be seen already that in the SM neutrinos (ν) are special. They only appear as left-handed particles and are massless. Unlike the other fermions, which can be a right-handed singlet, there are no right-handed neutrino singlets.

In a similar manner to the previous two forces the Lagrangian is required to be invariant under a local gauge transformation of the form:

$$\Psi_L \longrightarrow e^{iH}\Psi_L \tag{2.17}$$

where Ψ_L is a placeholder for either q_L or l_L. In this case H is a hermitian 2×2 matrix (satisfying $H^\dagger = H$), which can be represented in terms of four real numbers as $H = \theta I + \tau \cdot a$, where θ is a real number, I is a 2×2 unit matrix and $\tau \cdot a$ is the sum of the real numbers a_i multiplied with a set of 2×2 matrices, labelled τ.

The matrices τ for the charged weak current interactions ($\pm$) are linear combinations of the first two Pauli spin matrices:

$$\tau^\pm = \frac{1}{2}(\tau^1 \pm i\tau^2) \tag{2.18}$$

This allows one to write the charged currents in the form:

$$j_\mu^\pm = \bar{\Psi}_L \gamma_\mu \tau^\pm \Psi_L \tag{2.19}$$

For the neutral current interaction the third Pauli matrix is used, given by

$$\tau^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.20}$$

in order to obtain a neutral current of the form:

$$j_\mu^3 = \bar{\Psi}_L \gamma_\mu \tau^3 \Psi_L = \frac{1}{2}\bar{\nu}_L \gamma_\mu \nu_L - \frac{1}{2}\bar{e}_L \gamma_\mu e_L \tag{2.21}$$

This current is for left-handed fermions only. But the neutral weak interaction includes right-handed fermions. This is accounted for by a ‘weak hypercharge’ current:

$$j_\mu^Y = 2j_\mu^{\text{EM}} - 2j_\mu^3 \tag{2.22}$$

Thus the electric charge quantum number Q (in units of e) can be identified as the sum of hypercharge Y and the third component of the weak isospin T^3:

$$Q = T^3 + \frac{1}{2}Y \tag{2.23}$$

The electromagnetic current is

$$j_\mu^{\mathrm{EM}} = -\bar{e}_L\gamma_\mu e_L - \bar{e}_R\gamma_\mu e_R, \tag{2.24}$$

such that by substituting Equation 2.21 and Equation 2.24 into Equation 2.22 one obtains:

$$j_\mu^Y = -2\bar{e}_R\gamma_\mu e_R - \bar{e}_L\gamma_\mu e_L - \bar{\nu}_L\gamma_\mu\nu_L \tag{2.25}$$

The values of Q, T^3 and Y for each handedness of the particles in the first generation are given in Table 2.1.

Table 2.1: The electric charge quantum number Q, the third component of the weak isospin T^3 and the hypercharge Y for each handedness of the first generation particles.

	ν_L	e_L	e_R	u_L	d_L	u_R	d_R
Q	0	-1	-1	$+2/3$	$-1/3$	$+2/3$	$-1/3$
T^3	$+1/2$	$-1/2$	0	$+1/2$	$-1/2$	0	0
Y	-1	-1	-2	$+1/3$	$+1/3$	$+4/3$	$-2/3$

In the Glashow-Weinberg-Salam (GWS) model the weak and EM force are unified into a single gauge group (based on [8]):

$$SU(2)_L \otimes U(1)_Y \tag{2.26}$$

In $SU(2)_L$ three weak isospin currents, $j_\mu^{1,2,3}$, couple to three gauge bosons $W_\mu^{1,2,3}$ (acting on left-handed states only). In $U(1)_Y$ the weak hypercharge current, j_μ^Y, couples to a gauge boson B_μ (acting on both left- and right-handed states).
However, the observed physical vector boson, Z, of the weak neutral interaction does not only have left-handed couplings but also right-handed couplings. Therefore $SU(2)_L \otimes U(1)_Y$ must be *broken* in the GWS theory. In analogy to the electromagnetic current, j_μ^{EM}, which is a combination of the two neutral currents j_μ^Y and j_μ^3, the two neutral states W_μ^3 and B_μ combine (or "mix") to give two physical states:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \tag{2.27}$$

where A_μ can be identified as the photon, the gauge boson of the EM force, and θ_W is the weak mixing angle. The other two physical gauge bosons for the weak charged interactions are simply given by:

$$\begin{pmatrix} W_\mu^+ \\ W_\mu^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \end{pmatrix} \tag{2.28}$$

The corresponding coupling strengths are given by:

$$g_W = \frac{e}{\sin\theta_W} \tag{2.29}$$

$$g_Z = \frac{e}{(\sin\theta_W \cos\theta_W)} \tag{2.30}$$

The kinetic term of the electroweak Lagrangian can now be written as:

$$\mathcal{L}_{\text{kinetic}} = -\frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{2.31}$$

The experimental fact that the Z and $W^\pm$ gauge bosons are massive but the photon remains massless is explained by the Higgs mechanism [9–11]. In 1967 Weinberg proposed to introduce a scalar field ϕ along with a potential $V(\phi)$:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.32}$$

$$V(\phi) = \mu^2 \bar{\phi}\phi + \lambda(\bar{\phi}\phi)^2 \tag{2.33}$$

with $\mu^2 < 0$ and $\lambda > 0$. ϕ is a doublet with a weak hypercharge $Y = 1$ and a non-zero vacuum expectation value ϕ^0. The appropriate choice for the vacuum expectation value is:

$$\phi^0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.34}$$

where $v = (-\mu^2/\lambda)^{1/2}$. This choice breaks both $SU(2)_L$ and $U(1)_Y$ gauge symmetries, but the $U(1)_{\text{EM}}$ symmetry remains unbroken. The consequence of the Higgs mechanism is a massless photon and massive $W^\pm$ and Z, with masses given by:

$$m_W = \frac{g_W v}{2} \tag{2.35}$$

$$m_Z = \frac{m_W}{\cos\theta_W} \tag{2.36}$$

One also obtains a coupling to a real, massive, scalar field, known as the Higgs boson (H^0). Using the Higgs field to expand the scalar field ϕ about the vacuum expectation value v gives:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H^0 \end{pmatrix} \tag{2.37}$$

In the next section it is outlined how the Higgs mechanism also accounts for the masses of the fermions.

2.3 The fermion sector

2.3.1 Yukawa interactions

Adding the fermion masses to the SM Lagrangian ($-m\bar{\psi}\psi$) by hand would violate local gauge invariance. As outlined previously the weak gauge boson masses are generated from the coupling between the gauge fields and the Higgs. In a similar manner the fermion masses can be generated from the coupling of the fermions to the Higgs field. These couplings are known as the Yukawa couplings, based on the principles of Hideki Yukawa's work [12].

For example, the electron mass is generated by adding following gauge invariant term to the Lagrangian:

$$\mathcal{L}_{m_e} = c_e[(\bar{l}_L\phi)e_R + \bar{e}_R(\bar{\phi}l_L)] \tag{2.38}$$

where c_e is the corresponding Yukawa coupling constant, $\bar{l}_L$ is the left-handed electron doublet (containing the neutrino ν_e), e_R is the right-handed electron singlet (there is no equivalent right-handed neutrino singlet) and ϕ is the Higgs field as defined in Equation 2.32. This gives:

$$\mathcal{L}_{m_e} = c_e\Big[\begin{pmatrix} \bar{\nu}_e & \bar{e} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R \begin{pmatrix} \phi^- & \bar{\phi}^0 \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \Big] \tag{2.39}$$

Spontaneous symmetry breaking results in a Higgs field with a vacuum expectation value v of the form:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H^0 \end{pmatrix} \tag{2.40}$$

(as has been introduced in Section 2.2.3, Equation 2.37). Substituting Equation 2.40 into Equation 2.39 gives:

$$\mathcal{L}_{m_e} = c'_e v(\bar{e}_L e_R + \bar{e}_R e_L) + c'_e H^0(\bar{e}_L e_R + \bar{e}_R e_L) \tag{2.41}$$

Therefore the electron mass m_e is proportional to the vacuum expectation value of the Higgs field ($m_e = c'_e v$) and in turn the coupling to the Higgs field is proportional to the electron mass ($c'_e H^0 = \frac{m_e}{v} H^0$). In general the coupling of the Higgs field to a fermion is always proportional to the mass of the fermion.

The quark masses are generated in the same way. However, there is a slight difference because the left-handed doublet for the quarks requires two masses (for both doublet entries, e.g. the up and down quark). For the electron only one mass is required (for the electron itself) as the associated neutrino is assumed to be massless. Therefore the Higgs doublet needs to be slightly modified such that there are two couplings in total with the left-handed quark doublet.

The Yukawa interactions for all fermions are summarised and added to the SM Lagrangian as an extra component, $\mathcal{L}_{\text{Yukawa}}$.

2.3.2 The CKM matrix

Yukawa couplings between the Higgs field and fermions introduce the mass terms for the fermions. In Section 2.3.1 only one generation was considered. However, in the case of three generations the Yukawa couplings induce 3 × 3 mass matrices (with off-diagonal matrix elements) to the quark mass terms. The consequence is a mixing of quark generations or 'flavour mixing'.

For left-handed leptons this implies a transformation from the weak to the mass basis, which is performed by the unitarity matrices $U_{d,u,l}$ (for the first generation):

$$d^i_L \to U^{ij}_d d^j_L \tag{2.42}$$

$$u^i_L \to U^{ij}_u u^j_L \tag{2.43}$$

$$l^i_L \to U^{ij}_l l^j_L \tag{2.44}$$

Due to the unitarity of the matrices, the transformations have no effect on neutral interactions:

$$\bar{d}_i d_i \to \bar{d}_i U_d (U_d)^\dagger d_i = \bar{d}_i d_i \tag{2.45}$$

As a result there are no *tree-level* flavour-changing neutral currents in the SM. Such interactions, for example $b \to s$ transitions, appear only at loop level.

In the SM at tree level only the $W^\pm$ currents can change flavour, such that quark interaction terms transform like:

$$\bar{u}_i \gamma^\mu d_i \to \bar{u}_i \gamma^\mu (U_u (U_d)^\dagger) d_i \tag{2.46}$$

$$\bar{d}_i \gamma^\mu u_i \to \bar{d}_i \gamma^\mu (U_d (U_u)^\dagger) u_i \tag{2.47}$$

This leads to the 3 × 3 CKM matrix [13,14],

$$V_{\text{CKM}} = U_u (U_d)^\dagger, \tag{2.48}$$

which is named after Nicola Cabibbo, Makoto Kobayashi and Toshihide Maskawa. It is a 3 × 3 matrix in order to allow for a complex phase, which will be outlined later in Section 2.3.3.

The CKM matrix is the mixing matrix connecting the weak eigenstates of down-type quarks (d', s', b') and the corresponding mass eigenstates (d, s, b) through:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.49}$$

The unitarity of the CKM matrix assures the absence of elementary flavour-changing neutral vertices and therefore explains the suppression of such decays in the SM. The CKM matrix is a complex unitary matrix with 4 free parameters: 3 angles $(\theta_{i=1,2,3})$ and one complex phase (δ). Therefore it can be written as the product of three rotation matrices with a complex phase:

$$\begin{aligned} V_{\mathrm{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_2 & s_2 \\ 0 & -s_2 & c_2 \end{pmatrix} \times \begin{pmatrix} c_1 & s_1 & 0 \\ -s_1 & c_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i\delta} \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_3 & s_3 \\ 0 & -s_3 & c_3 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13}s_{13}e^{-i\delta} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}e^{i\delta} & s_{23}s_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -s_{23}c_{12} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{2.50}$$

where $c_{i(j)} = \cos\theta_i(\cos\theta_j)$ and $s_{i(j)} = \sin\theta_i(\sin\theta_j)$. Alternatively, one can express the CKM matrix with the Wolfenstein parametrisation as [15]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{2.51}$$

The unitarity requirement of V_{CKM} results in six conditions:

$$\begin{aligned} V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* &= 0 \\ V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* &= 0 \\ V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* &= 0 \\ V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* &= 0 \\ V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* &= 0 \\ V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* &= 0 \end{aligned}$$

These conditions can be visualised as triangles in the complex plane, which are also called the Unitarity Triangles (UTs). The UT of the first condition, sometimes referred to as the main UT, can be described by the three angles α, β and γ and the two sides R_u and R_t (the base side is usually normalised to unit length). It is shown in Figure 2.2. One aim of flavour physics is to tighten the experimental constraints on the UTs in order to test for any deviations from unitarity (for each UT there are 5 constraints: 3 angles and 2 sides). The current world averages for each of the three angles are $\alpha = (89.0^{+4.4}_{-4.2})^{\circ}$, $\beta = (21.38^{+0.79}_{-0.77})^{\circ}$ and $\gamma = (68^{+10}_{-11})^{\circ}$ [16], where the uncertainties are determined by the experimental accuracy. The sides are limited by theoretical uncertainties as discussed in [17].

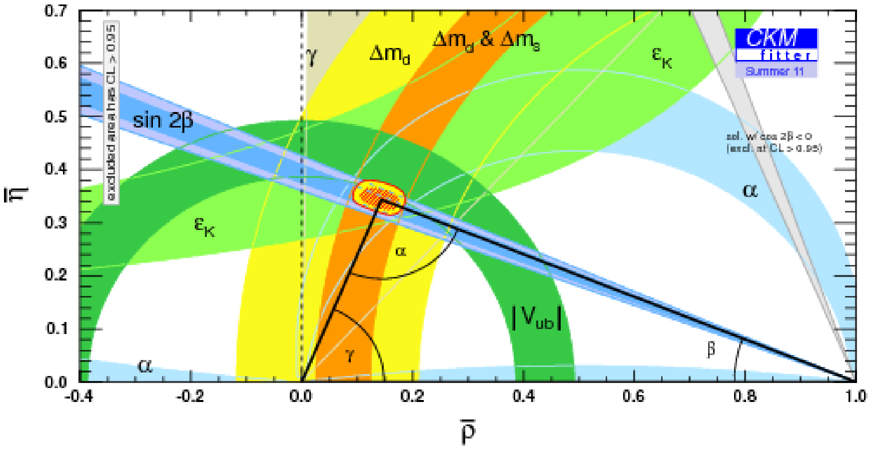


Figure 2.2: The current constraints on the Unitarity Triangle (described by the angles α, β and γ) in the complex plane. Each constraint is shown with a 95% C.L. band. Taken from [16].

2.3.3 The importance of symmetries

In physics symmetries lead to conservation laws [18]. The laws of physics should also be invariant under discrete symmetries. The symmetries of relevance in particle physics are charge conjugation (C), parity (P) and time-reversal (T). The corresponding operators are $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{T}$:

- $\mathcal{C}$ transforms any particle into its anti-particle
- $\mathcal{P}$ inverses the spatial coordinates from $(t, \bar{x})$ to $(t, -\bar{x})$
- $\mathcal{T}$ inverses the time coordinate from $(t, \bar{x})$ to $(-t, \bar{x})$

Individually these symmetries can be broken (for example P by weak interactions [19]). Combined transformations of C and P simultaneously were thought to be conserved within the relativistic quantum field theory framework. However, the CP symmetry was first observed to be broken in neutral Kaon decays [20]. CP violation in the beauty sector (involving b quarks) was observed by the B-factories [21] and recently the LHCb experiment reported on the first evidence for CP violation in the charm sector (involving c quarks) [22]. Therefore, in weak interactions both C and P can be broken whereas the EM and the strong force conserve all symmetries.

Nicola Cabibbo first came up with a 2×2 mixing matrix depending on only one real parameter (the Cabibbo angle), which therefore cannot account for CP violation [13]. In 1972 Makoto Kobayashi and Toshihide Maskawa postulated the existence of a complex phase in the framework [14]. In order to introduce a complex phase, δ, they extended the mixing matrix to a 3×3 matrix, known as the CKM matrix (V_{CKM}) today. This complex phase would allow the CP symmetry to be broken and also imply a third generation of particles (only two were known at the time). A few years later in 1975 a third generation of particles was experimentally proven to exist by the discovery of the τ lepton [23]. This was followed by the discovery of the Υ meson (a $b\bar{b}$ quark pair) in 1977 [24]. Makoto Kobayashi and Toshihide Maskawa were awarded the Nobel Prize in Physics for their work in 2008.

The Lagrangian component for the Yukawa interactions, $\mathcal{L}_{\text{Yukawa}}$, is left as the only possible source for CP violation in the SM. There are 3 types of CP violation:

- CP violation in mixing (or ‘indirect CP violation’): The wave function describing a neutral meson can be expressed as a superposition of the two flavour states $|P^0\rangle$ and $|\bar{P}^0\rangle$, such that:

  $$|\psi(t)\rangle = \psi_1(t)|P^0\rangle + \psi_2(t)|\bar{P}^0\rangle \tag{2.52}$$

  A light (L) and heavy (H) eigenstate can be constructed by adding the two flavour states in the following manner [7]:

  $$|P_L\rangle = p|P^0\rangle + q|\bar{P}^0\rangle \tag{2.53}$$

  $$|P_H\rangle = p|P^0\rangle - q|\bar{P}^0\rangle \tag{2.54}$$

  where p and q are complex mixing parameters satisfying $|p|^2 + |q|^2 = 1$. CP violation in mixing occurs when $q/p \neq 1$ and has been seen in the K^0, B^0, B_s^0 and D^0 systems [22].

- CP violation in decay (‘direct CP violation’) occurs in both charged and neutral decays when a certain process and its CP conjugated process have different magnitudes.

- CP violation induced by interference between mixing and decay.

2.4 The flavour problem

The SM has been very successful in describing the observed particle interactions so far. But it is believed only to be part of a more general theory or only to be valid up to a certain energy scale. Many theoretical models expect New Physics (NP) to enter at the TeV-scale. However, if this was the case interferences with amplitudes and observables of current measurements should have been observed *already*, especially in processes that are believed to be very sensitive.

For example Flavour Changing Neutral Current (FCNC) decays, which belong to the so-called rare decays of the SM, do not occur at tree level (the formalism of FCNC decays is given in greater detail in Chapter 3). They are also highly suppressed and have no other competing SM processes. This makes the observables of these decays very sensitive to NP. However, measurements of FCNC decays at various experiments have all been in agreement with the SM predictions. A possible explanation for the lack of signs of NP is given by the Minimal Flavour Violation (MFV) hypothesis [25]. It is sufficient to know that within the MFV framework NP is still assumed to be at the TeV-scale, but at the same time FCNC decays are ‘protected’ by symmetry principles. This implies that deviations from the SM in decay amplitudes are not greater than $\mathcal{O}(10\%)$.

2.4.1 Further problems of the SM

There are various phenomena to which the SM does not provide an answer and there are also a few open questions about the theory itself. An overview of a few selected points for both are given in the following list:

- The Higgs boson has not been discovered yet, but the mass range, in which it is predicted to be, is very narrow already due to experimental constraints.

- It has been determined experimentally that neutrinos have a mass (albeit very small). But the SM assumes the neutrinos to be massless.

- Gravity is not incorporated in the SM

- There are no dark matter candidates in the SM.

- There are 19 free parameters describing the SM, that have to be determined experimentally. These are the masses of the fermions, the CKM mixing angles, the CKM complex phase and the gauge couplings (including the Higgs parameters). Why is it those 19?

- Why are there 3 and not more particle generations?

- The matter anti-matter ratio in the universe is not even. With the current measurements of CP violation a ratio of $\mathcal{O}(10^{-20})$ is predicted - but the observed ratio is about $\mathcal{O}(10^{-9})$. Apart from other mechanisms (such as leptogenesis [26]) this could point to the fact that there are still undiscovered sources of CP violation.

2.5 Physics searches at the LHC

The SM and all its sectors are being probed by many experiments world-wide. The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland, hosts several experiments, that are designed to look for effects of NP. The experiments take different approaches to achieve this, which can be categorised into direct and indirect searches.

Direct searches are performed by the two general purpose detectors ATLAS (A Toroidal LHC ApparatuS) [27] and CMS (Compact Muon Solenoid) [28]. Many analyses at ATLAS and CMS are dedicated to the search for Supersymmetry (SUSY) particles. SUSY is considered to be a good framework to solve many of the SM problems. The SM physics searches of these two experiments include the search for the Higgs boson.

Indirect NP searches are the main focus of the Large Hadron Collider beauty Experiment (LHCb). The LHCb detector is a single-arm spectrometer. It is designed to perform precise measurements of B-meson decays in the forward direction. This allows one to search for deviations from the SM in various sectors. For example B_s^0 oscillations give access to new particles in box diagrams and rare decays can probe NP particles in loop diagrams (FCNC decays).

The LHCb detector is described in greater detail in Chapter 4.

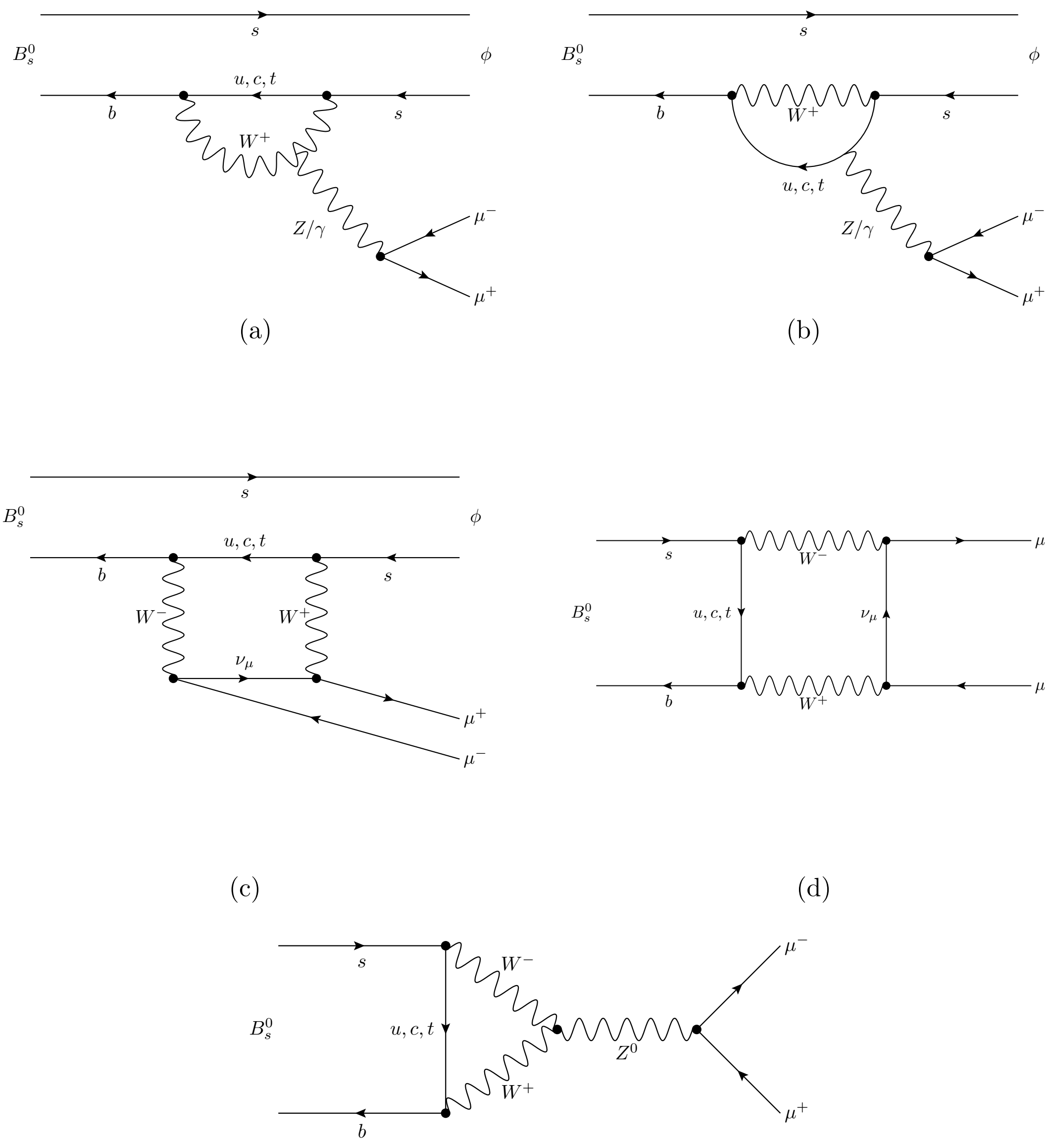
Chapter 3

Flavour Changing Neutral Currents

Flavour Changing Neutral Current (FCNC) decays are suited as the probes of NP for a variety of reasons. They are very suppressed since they are forbidden in the SM at tree level (see Section 2.3.3), with branching fractions ranging between $\mathcal{O}(10^{-5})$-$\mathcal{O}(10^{-7})$ depending on the CKM elements involved. They are mediated through either loop diagrams (sometimes referred to as penguins in the literature) or box diagrams. This makes them very sensitive to NP as couplings and masses of new particles, which can contribute virtually, can have a sizeable influence on observables.

The radiative (or electroweak) loops in the Feynman diagrams for the decays of b quarks consist of a W and an intermediate quark (u, c or t) with a radiative γ (or Z). The main contribution comes from the top quark, t, due to its high mass. The box diagrams contain two W bosons (W^+ and W^-). In theory new particles can just replace any of the intermediate quarks inside these loops and boxes and therefore influence observables indirectly. The corresponding Feynman diagrams are shown in Figure 3.1, where the decays $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ are used as examples.

One of the main aims of LHCb is to accurately measure FCNC decays in the B-meson system. Experimentally, the B-meson system has several advantages, which reduce uncertainties and increase the accuracy: The b quark is the heaviest quark, which forms stable states, that decay within a finite and therefore measurable distance. Therefore 'long distance' effects (which are non-trivial to calculate and/or not well understood; the formalism is explained later in Section 3.1) are relatively small compared to other meson systems, such as the charm and kaon sectors with a c and s quark. Therefore very precise predictions for SM processes can be made in the B-meson system. This increases the sensitivity to detect effects of NP processes. There are several B-meson FCNC decays of interest for LHCb.



(a) (b) (c) (d) (e)

Figure 3.1: Feynman diagrams contributing to the decays $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$. The two diagrams (a)-(b) are the main ‘loop’ diagrams contributing to $B_s^0 \to \phi\mu^+\mu^-$. The ‘box’ diagram (c) also contributes to $B_s^0 \to \phi\mu^+\mu^-$ whereas the ‘box’ diagram (d) represents $B_s^0 \to \mu^+\mu^-$. The diagram (e) has the largest contribution to $B_s^0 \to \mu^+\mu^-$.

The two decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ have branching fractions predicted precisely by the SM as $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (0.10 \pm 0.01) \times 10^{-9}$ [29, 30]. The Feynman diagrams for the process $B_s^0 \to \mu^+\mu^-$ are given in Figure 3.1. An additional reduction to the branching fraction, compared to other CKM suppressed B-meson FCNC decays, comes from the fact that these processes are helicity suppressed.

The prominent example of helicity suppression is the weak charged pion decay $\pi^- \to l^-\bar{\nu}_l$, where $l = \mu, e$. Experimentally it has been shown that the pion favours the muon-mode over the electron-mode (by a factor of $\sim$ 8000), which is contrary to what one would naively expect (as the electron is much lighter). The explanation is that the pion has spin 0, which implies that the electron and antineutrino must emerge with opposite spins in opposite directions in the pion rest frame. Therefore they have *equal* helicities: The antineutrino is always right-handed, so the electron must be right-handed as well. Assuming the extreme case, where the electron is massless (and therefore a purely left-handed particle), this pion decay would not be allowed (as the emerging electron must be right-handed). The electron mass is much smaller and closer to zero (the massless case) than the muon mass, which is why the electron mode is more 'helicity suppressed'. NP contributions from scalar new particles, which are not helicity suppressed, to $B_s^0 \to \mu^+\mu^-$ are thought to significantly enhance the value of the branching fraction, which in turn can be measured experimentally. For example in the Constrained Minimal Supersymmetric Standard Model (CMSSM) the branching fraction is strongly correlated with $\tan^6\beta$, where $\tan\beta$ is the ratio of Higgs vacuum expectation values [31, 32]. The branching fraction is also dependent on the gaugino mass, $m_{1/2}$, and the trilinear soft supersymmetry-breaking parameter A_0 [33, 34]. The predictions of the branching fraction for various scenarios are shown in Figure 3.2.

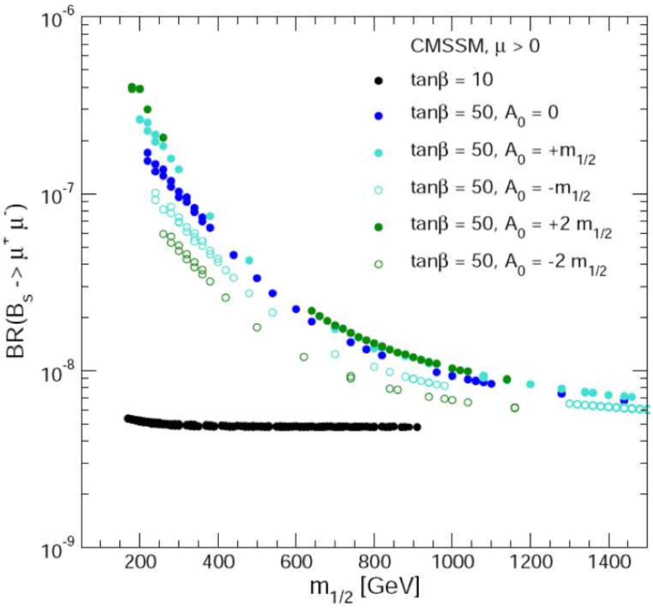


Figure 3.2: The branching fraction $\mathcal{B}(B_s^0 \to \mu\mu)$ as a function of the gaugino mass $m_{1/2}$ in the CMSSM. Indicated are the predictions for different values of $\tan\beta$ and A_0. Taken from [33].

The FCNC decay $B_s^0 \to \phi\gamma$ is not helicity suppressed. This results in larger rates compared to the previous decay and consequently in a larger data sample, enabling measurements of the helicity structure of the amplitudes. However, in practice $B_s^0 \to \phi\gamma$ is non-trivial to detect because of the primary vertex reconstruction. The ϕ decays predominantly via K^+K^-. Because there is no vertex for the photon in LHCb, the

B_s^0 decay vertex is reconstructed from only two charged tracks with a relatively small opening angle. Experimentally the reconstruction of such events introduces substantial systematic uncertainties. This is especially true for LHCb, where the occupancy is high and there are many tracks due to the nature of hadronic interactions.

The decay $B_s^0 \to \phi\mu^+\mu^-$ (where $\phi \to K^+K^-$) is also not helicity suppressed. But because it has two additional charged muon tracks it does not suffer from the same experimental problems as the decay $B_s^0 \to \phi\gamma$.
The decay $B_s^0 \to \phi\mu^+\mu^-$ has a quark transition from b to s, where the quark flavour changes but the charge does not change (hence it is a FCNC). Experimentally it is very clean to select. It has no neutral particles and the muon identification efficiency of the LHCb muon stations is high (more details on the detector performances are found in Chapter 4). The same is true for the kaons forming the ϕ. The narrow ϕ mass reduces the combinatoric background level significantly. As the decay is not helicity suppressed the rate of $B_s^0 \to \phi\mu^+\mu^-$ decays can be precisely measured and it also gives access to helicity distributions. The required formalism for this is outlined later in Section 3.1. The initial flavour of the B-meson (whether it is a B_s^0 or $\bar{B}_s^0$) cannot be determined directly. The other b quark (assuming it decays via a self-tagging mode) of the decay of a $b\bar{b}$ quark pair has to be tagged instead. A further possibility is the measurement of time-dependent CP violation in $B_s^0 \to \phi\mu^+\mu^-$ through mixing. Overall $B_s^0 \to \phi\mu^+\mu^-$ is very suited for NP searches and has several advantages over other FCNCs as outlined above.
The decay $B^0 \to K^{*0}\mu^+\mu^-$ (where $K^{*0} \to K^+\pi^-$) is very similar to $B_s^0 \to \phi\mu^+\mu^-$. The only difference is that the spectator quark is a d quark instead of a s quark. Compared to $B_s^0 \to \phi\mu^+\mu^-$ it has the advantage that it is self-tagging, because the sign of the kaon indicates the flavour of the initial B-meson (B^0 or $\bar{B}^0$).

The power of FCNC decays in general is that they can be described and therefore be studied in a model independent way. This requires the introduction of the Operator Product Expansion (OPE). The OPE facilitates the construction of observables of FCNC decays, that are sensitive to effects of NP. The following subsections give an introduction to the OPE and to the corresponding observables. The current experimental status of $B_s^0 \to \phi\mu^+\mu^-$ is given in Section 3.3.

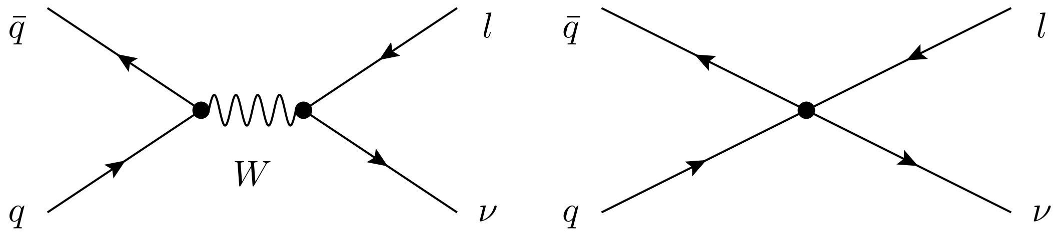
3.1 Operator Product Expansion

A model-independent framework is required, which accounts for the SM as well as for potential NP extensions. For example the SM Lagrangian $\mathcal{L}_{\mathrm{SM}}$ can be extended to a 'true' Lagrangian $\mathcal{L}$ (which can include NP) of the form:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \frac{\sum_i \mathcal{C}_i \mathcal{Q}_i}{\Lambda} \tag{3.1}$$

where $\mathcal{Q}_i$ are a set of operators, $\mathcal{C}_i$ are their respective couplings and Λ is the scale of the process. This method is referred to as Operator Product Expansion (OPE). The formalism for OPE was developed by Wilson and Zimmermann [35]. They have shown that a product of two operators can always be expanded in the form as shown in Equation 3.1.

An example is Enrico Fermi's theorem of β decays [36]. Fermi expressed β decays ($n \to pe\bar{\nu}$) as a 4-body coupling. Figure 3.3 shows how the W is absorbed into a single 4-body vertex ('effective' coupling). He therefore defined operators for the transition of $n \to p$ and for the creation of an electron and a neutrino.



(a) (b)

Figure 3.3: The full (a) and the effective picture (b) of β decay.

The resulting effective vertex has a coupling constant G_F, the Fermi coupling constant. It integrates the contribution of the W, such that [37]:

$$G_F = \frac{\sqrt{2}}{8}\frac{g^2}{m_W^2} = 1.16637 \times 10^{-5}\,\mathrm{GeV}/c^2 \tag{3.2}$$

where g is the coupling constant of the weak force and m_W the mass of the W[1].

NP is expected to interfere with the weak sector. The W gauge bosons of the weak force are very heavy compared to the final SM particles and the distance at which the weak force acts is small. Therefore the weak interaction can be modelled as an effective weak vertex instead, in the same manner as in Fermi's theorem of β decay.

The weak part of the SM Lagrangian is expressed as a model-independent sum of

---

[1] At higher energies it can be seen that the mass of the W is actually described by a Breit-Wigner function.

operators, giving a so called ‘effective’ Lagrangian:

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{QCD}\times\text{QED}} + \frac{4G_F}{\sqrt{2}} V_{\text{CKM}} \sum_i \mathcal{C}_i \mathcal{Q}_i \tag{3.3}$$

The scale Λ of this operation is equal to m_W (because the weak sector is considered). Λ is absorbed into the coefficients (in the same manner as the W contribution was integrated into G_F).

In the literature $\mathcal{L}_{\text{eff}}$ is often just given in the form of an effective Hamiltonian, $\mathcal{H}_{\text{eff}}$, instead:

$$\mathcal{H}_{\text{eff}} = \sum_i^{10} \mathcal{C}_i \mathcal{Q}_i \tag{3.4}$$

The coefficients $\mathcal{C}_i$ are the Wilson coefficients. They encode the contributions from particles heavier than the decaying particles (in analogy to how G_F encodes the contribution of the W in β decay). Lighter particles at or below the mass scale of the decaying particles contribute to the operators $\mathcal{Q}_i$. Hence the OPE is said to divide the effective vertex into two parts: the ‘short’ distance part (the Wilson coefficients) and the ‘long distance’ part (the operators).

The Wilson coefficients have been calculated for the SM [38] and for NP models, such as the Minimal Supersymmetric Standard Model (MSSM) [39]. The values of the Wilson coefficients are sensitive to the underlying model and change from the SM predictions when there are NP terms entering.

Precise measurements of the coefficients enable one to determine the ‘type’ of NP. Because each coefficient is linked to an operator with a certain gauge structure, any deviation from the SM would indicate the gauge structure of the NP.

The OPE is therefore a model-independent framework, that allows precise testing of the SM. For convenience, ‘effective ’Wilson coefficients are defined for low energies, which are the relevant quantities to describe physical observables (discussed later in Section 3.2). They are given by [38]:

$$\begin{aligned}
\mathcal{C}_7^{\text{eff}} &= \mathcal{C}_7 - \frac{1}{3}\mathcal{C}_3 - \frac{4}{9}\mathcal{C}_4 - \frac{20}{3}\mathcal{C}_5 - \frac{80}{9}\mathcal{C}_6 && (3.5)\\
\mathcal{C}_9^{\text{eff}} &= \mathcal{C}_9 + Y(q^2) && (3.6)\\
\mathcal{C}_{10}^{\text{eff}} &= \mathcal{C}_{10} && (3.7)\\
\mathcal{C}_{7,9,10}^{\prime\text{eff}} &= \mathcal{C}'_{7,9,10} && (3.8)
\end{aligned}$$

where $Y(q^2)$ is a function of the dimuon invariant mass squared (q^2) and is defined in [40]. $\mathcal{C}_i$ are the Wilson coefficients introduced above.

The operators can be classified according to their contribution to certain decays. $\mathcal{Q}_{1-2}$

are the so-called current-current operators, $\mathcal{Q}_{3-6}$ are the QCD-penguin diagram operators responsible for a quark loop with gluon emission and $\mathcal{Q}_{7-10}$ are the electroweak-penguin operators.

For example, the operators corresponding to $b \to s\bar{q}q$ transitions are given by:

$$\mathcal{Q}_7 = \frac{3}{2}(\bar{s}b)_{V-A} \sum_{q=u,d,s,c,b} e_q(\bar{q}q)_{V+A} \tag{3.9}$$

$$\mathcal{Q}_8 = \frac{3}{2}(\bar{s}_\alpha b_\beta)_{V-A} \sum_{q=u,d,s,c,b} e_q(\bar{q}_\beta q_\alpha)_{V+A} \tag{3.10}$$

$$\mathcal{Q}_9 = \frac{3}{2}(\bar{s}b)_{V-A} \sum_{q=u,d,s,c,b} e_q(\bar{q}q)_{V-A} \tag{3.11}$$

$$\mathcal{Q}_{10} = \frac{3}{2}(\bar{s}_\alpha b_\beta)_{V-A} \sum_{q=u,d,s,c,b} e_q(\bar{q}_\beta q_\alpha)_{V-A} \tag{3.12}$$

where V stands for a vector structure (γ_μ), A for an axial structure ($\gamma^\mu\gamma_5$) and α and β denote colour indices for gluon currents. The subscript $V - A$ therefore indicates charged currents with a vector minus axial structure. In the literature this is sometimes explicitly included in the terms as

$$\gamma_\mu(1-\gamma_5), \tag{3.13}$$

or simply as

$$\gamma_\mu P_L \tag{3.14}$$

for left-handed particles. The equivalent right-handed component is:

$$\gamma_\mu P_R = \gamma_\mu(1+\gamma_5) \tag{3.15}$$

For the FCNC decays of interest in this thesis ($b \to sl^+l^-$ transitions) the relevant operators are (in analogy to the above, with a similar notation):

$$\mathcal{Q}_{7\gamma} = \frac{e}{g^2}\bar{m}_b(\bar{s}\sigma_{\mu\nu}P_R b)F^{\mu\nu} \tag{3.16}$$

$$\mathcal{Q}_{8G} = \frac{1}{g}\bar{m}_b(\bar{s}\sigma_{\mu\nu}T^\alpha P_R b)G^{\mu\nu\alpha} \tag{3.17}$$

$$\mathcal{Q}_{9V} = \frac{e^2}{g^2}(\bar{s}\gamma_\mu P_L b)(\bar{l}\gamma^\mu l) = (\bar{s}b)_{V-A}(\bar{l}l)_V \tag{3.18}$$

$$\mathcal{Q}_{10A} = \frac{e^2}{g^2}(\bar{s}\gamma_\mu P_L b)(\bar{l}\gamma^\mu\gamma_5 l) = (\bar{s}b)_{V-A}(\bar{l}l)_A \tag{3.19}$$

$\mathcal{Q}_{7\gamma}$ corresponds to photon emission during a $b \to s$ transition with the photon subsequently decaying into the final state lepton pair (as shown in Figure 3.1). $\mathcal{Q}_{8G}$ is the equivalent gluonic version and therefore heavily (colour) suppressed. Its contribution can be ignored. $\mathcal{Q}_{9V}$ and $\mathcal{Q}_{10A}$ are vector and axial contributions (from loop and box diagrams). The three operators of interest also have opposite-handed partners, indicated by a prime ($\mathcal{Q}'_{7\gamma}$, $\mathcal{Q}'_{9V}$ and $\mathcal{Q}'_{10A}$).

3.2 Observables

Branching fractions and angular observables can be directly related to the Wilson coefficients. Measurements of these observables are therefore sensitive to NP (which is expected to change the values of the Wilson coefficients).

$b \to sl^+l^-$ decays, such as $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$, can be described kinematically by 4 variables: θ_l, θ_k, ϕ' and q^2.

When defining the z-axis as the B-meson flight direction, then θ_l is the angle between the l^- and the z-axis in the dilepton rest frame, θ_k is the angle between the K^- and the z-axis in the ϕ (or $\bar{K^{*0}}$) rest frame, ϕ' is the angle between the normals to the decay planes of the dilepton and the ϕ (or $\bar{K^{*0}}$) in the B-meson rest frame and q^2 is the dilepton invariant mass squared. An illustration of the angles is shown in Figure 3.4. The decay width can be written as a function of these angular observables [41]:

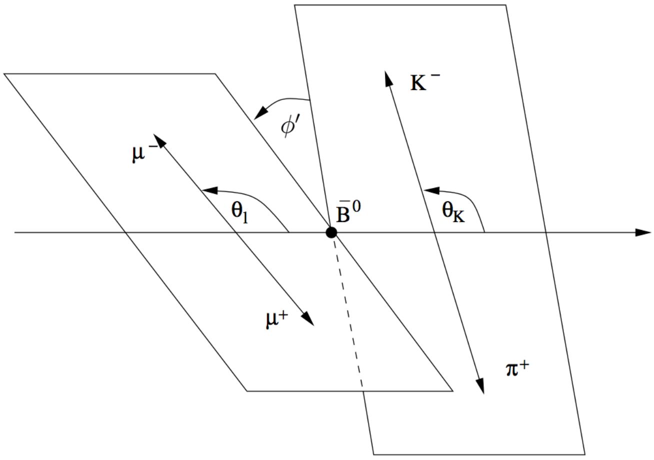


Figure 3.4: The angles θ_l, θ_k and ϕ' describing the decay $\bar{B^0} \to \bar{K^{*0}}\mu^+\mu^-$. The horizontal line indicates the B-meson flight direction, the K^- and π^+ are shown in the $\bar{K^{*0}}$ rest frame and the μ^+ and μ^- are shown in the dimuon rest frame. The two planes describing both rest frames are drawn separately for illustration purposes. Taken from [33].

$$\frac{d^4\Gamma}{dq^2 d\cos\theta_l d\cos\theta k d\phi} = \frac{9}{32\pi} I(q^2, \theta_l, \theta_k, \phi') \tag{3.20}$$

where I is the sum of a total of 9 distinct functions, I_{1-9}. For example, for the decay $B^0 \to K^{*0}\mu^+\mu^-$ the functions I_{1-9} can be written in terms of the K^* spin amplitudes $A_\perp$, $A_\parallel$, A_0 and A_t (as done in [41]). In the heavy quark and large energy limit the three spin amplitudes of interest are given at leading order by:

$$
\begin{aligned}
A_{\perp L,R} &= \sqrt{2}Nm_b(1-\hat{s})\Big[(\mathcal{C}_9^{\text{eff}} \mp \mathcal{C}_{10}^{\text{eff}}) + \frac{2\hat{m}_b}{\hat{s}}(\mathcal{C}_7^{\text{eff}} + \mathcal{C}_7^{\prime\text{eff}})\Big]\xi_\perp(E_{K^*}) && (3.21)\\
A_{\| L,R} &= -\sqrt{2}Nm_b(1-\hat{s})\Big[(\mathcal{C}_9^{\text{eff}} \mp \mathcal{C}_{10}^{\text{eff}}) + \frac{2\hat{m}_b}{\hat{s}}(\mathcal{C}_7^{\text{eff}} - \mathcal{C}_7^{\prime\text{eff}})\Big]\xi_\perp(E_{K^*}) && (3.22)\\
A_{0L,R} &= -\frac{Nm_b}{2\hat{m}_{K^*}\sqrt{\hat{s}}}(1-\hat{s})^2\Big[(\mathcal{C}_9^{\text{eff}} \mp \mathcal{C}_{10}^{\text{eff}}) + 2\hat{m}_b(\mathcal{C}_7^{\text{eff}} - \mathcal{C}_7^{\prime\text{eff}})\Big]\xi_\|(E_{K^*}) && (3.23)
\end{aligned}
$$

where $C_{7,9,10}^{(\prime)\text{eff}}$ are the Wilson coefficients introduced in Section 3.1, $\hat{s} = q^2/m_B^2$ and $\hat{m}_i = m_i/m_B$ [41].

The following subsections outline how the spin amplitudes and Wilson coefficients are linked to physical observables.

3.2.1 Branching fraction

The branching fractions of decays, in particular rare B-meson decays, were predicted to be very sensitive to effects of NP as they put constraints on the Wilson coefficients. For decays such as $B_s^0 \to \phi\mu^+\mu^-$ the shape of the branching fraction as a function of q^2, the di-muon invariant mass squared, offers access to a variety of theoretical predictions as well [42].

With a better theoretical understanding it became clear that only measuring the branching fractions will not be sufficient. Angular observables give more handles on the Wilson coefficients and thus enable one to overconstrain theoretical predictions in several ways. Most importantly, angular observables (as explained in the next Sections) are formed by taking the ratio of distributions, which cancels all hadronic form factors and their uncertainties. This is not the case for branching fraction measurements, where this inherent uncertainty is unavoidable.

However, in order to perform precise angular measurements one must understand and measure the branching fraction of a decay first, which is the focus of this thesis for the $B_s^0 \to \phi\mu^+\mu^-$ decay. The differential branching fraction of $b \to sl^+l^-$ decays, expressed in terms of Wilson coefficients, is given in [43] as:

$$
\begin{aligned}
\frac{\frac{d\Gamma(b\to s\mu^+\mu^-)}{d\hat{s}}}{\Gamma(b\to ce\bar{\nu})} &= \frac{\alpha^2}{4\pi^2}\left|\frac{V_{ts}}{V_{cb}}\right|^2\frac{(1-\hat{s})^2}{f(z)\kappa(z)}\times && (3.24)\\
&\Big[(1+2\hat{s})\big(|\mathcal{C}_{9V}^{\text{eff}}|^2+|\mathcal{C}_{10A}|^2\big)+4\big(1+\frac{2}{\hat{s}}\big)|\mathcal{C}_{7\gamma}^{\text{eff}}|^2+12|\mathcal{C}_{7\gamma}^{\text{eff}}|\mathcal{R}(\mathcal{C}_{9V}^{\text{eff}})\Big]
\end{aligned}
$$

where $\hat{s} = q^2/m_B^2$ and $l = \mu$ was used. In this expression the branching fraction is given with respect to the well-measured $b \to ce\bar{\nu}$ decay in order to cancel QCD effects. This requires a phase space factor $f(z)$ and a QCD correction factor $\kappa(z)$ (where $z = m_c/m_b$) to be included. From Equation 3.24 it can be seen that the rate of $b \to sl^+l^-$ decays is sensitive to not only the magnitude, but also to the sign of the Wilson coefficients. The branching fraction in the high q^2 $(= s)$ region is mainly sensitive to NP contributions to $\mathcal{C}_{9V}^{\text{eff}}$ and $\mathcal{C}_{10A}$, while the branching fraction in the low q^2 region also strongly depends on $\mathcal{C}_{7\gamma}^{\text{eff}}$.

The branching fraction ($\mathcal{B}$) as a function of q^2 can be measured experimentally by counting the number of signal events in bins of q^2 and correcting for detector efficiencies. Measuring the branching fraction ratio with respect to an appropriate control channel (with larger statistics) has the advantage that common systematic uncertainties cancel.

3.2.2 A_{FB}

The forward-backward asymmetry, A_{FB}, is the asymmetry between the number of decays with the positive muon going along (forward) and against (backward) the direction of the B momentum vector in the dimuon rest frame. It can be expressed as:

$$A_{FB}(q^2) = \frac{\int\limits_{0}^{1} d\cos\theta_l \frac{d^2\Gamma}{dq^2 d\cos\theta_l} - \int\limits_{-1}^{0} d\cos\theta_l \frac{d^2\Gamma}{dq^2 d\cos\theta_l}}{\int\limits_{-1}^{1} d\cos\theta_l \frac{d^2\Gamma}{dq^2 d\cos\theta_l}} \tag{3.25}$$

To measure A_{FB} it is required to flavour tag the initial B-meson (B or $\bar{B}$) because otherwise the asymmetries from both simply cancel. $B^0 \to K^{*0}(\to K^+\pi^-)\mu^+\mu^-$ is suited for this measurement as it is self-tagging, which means that the sign of the kaon indicates the flavour of the initial b quark (b or $\bar{b}$). This is not possible for $B_s^0 \to \phi\mu^+\mu^-$ as described at the beginning of this chapter. For $B_s^0 \to \phi\mu^+\mu^-$ a possibility would be to tag the other b quark of the $b\bar{b}$ quark pair decay instead, which requires a large data sample. The shapes of A_{FB} as predicted by the SM and various NP scenarios (with flipped signs for some of the Wilson coefficients) are shown in Figure 3.5. The point, where the asymmetry vanishes (the zero crossing point), is proportional to the ratio $\mathcal{C}_{9V}^{\text{eff}}/\mathcal{C}_{7\gamma}^{\text{eff}}$ and all hadronic form factor uncertainties cancel.

LHCb has measured A_{FB} for $B^0 \to K^{*0}\mu^+\mu^-$ and finds no deviation from the SM prediction, as shown in Figure 3.6 [44, 45].

3.2.3 A_{CP}

Decay rates are a handle to the time-dependent CP asymmetry, $A_{CP}(t)$, given by:

$$A_{CP}(t) = \frac{\Gamma^{B_s^0} - \Gamma^{\bar{B}_s^0}}{\Gamma^{B_s^0} + \Gamma^{\bar{B}_s^0}} \tag{3.26}$$

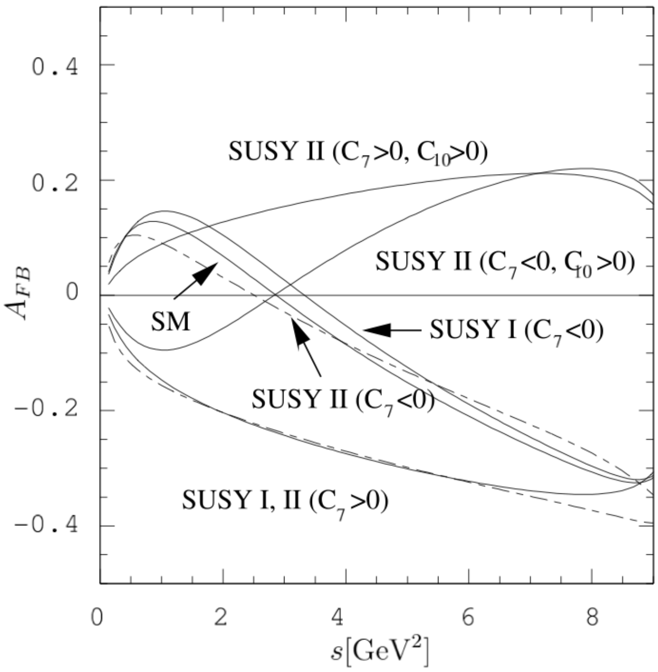


Figure 3.5: The forward-backward asymmetry (A_{FB}) for $B^0 \to K^{*0}\mu^+\mu^-$ decays as a function of the dimuon invariant mass squared (s). Indicated are the predictions from the SM and various other NP models. Taken from [46].

The functions $I_{5,6,8,9}$ (introduced at the beginning of Section 3.2, Equation 3.20) are odd under CP violation, which means that the corresponding CP asymmetry can be extracted from $d\Gamma(B) + d\Gamma(\bar{B})$ without flavour tagging the initial b quark. The formalism of this method for $B_s^0 \to \phi\mu^+\mu^-$ is outlined in [47], which also gives access to the phase of $B_s^0 - \bar{B}_s^0$ mixing. The number of signal events determines the accuracy of such a measurement.

3.2.4 Other observables

When taking the ratio of the spin amplitudes defined earlier (A_0, $A_\perp$ and $A_\parallel$), then theoretical uncertainties due to hadronic form factors cancel. Several angular and polarisation observables can be constructed in such a way.

For example the transverse amplitude, which is given by:

$$A_T^{(2)}(q^2) = \frac{|A_\perp|^2 - |A_\parallel|^2}{|A_\perp|^2 + |A_\parallel|^2} \tag{3.27}$$

$A_T^{(2)}$ is therefore directly proportional to $\mathcal{C}_{7\gamma}^{\prime\mathrm{eff}}$ (see Equations 3.21 and 3.22). Another observable is the fraction of the longitudinal polarisation of the K^* (ϕ) in $B^0 \to K^{*0}\mu^+\mu^-$ ($B_s^0 \to \phi\mu^+\mu^-$) decays, which is given by:

$$F_L(q^2) = \frac{|A_0|^2}{|A_0|^2 + |A_\perp|^2 + |A_\parallel|^2} \tag{3.28}$$

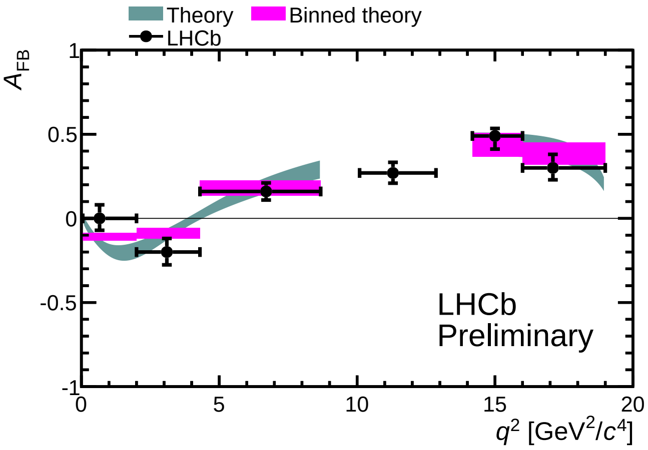


Figure 3.6: The LHCb measurement of the forward-backward asymmetry (A_{FB}) for $B^0 \to K^{*0}\mu^+\mu^-$ decays as a function of the dimuon invariant mass squared (q^2). Indicated are also the SM theory predictions. Taken from [45].

These observables can be extracted experimentally from fits to the angular distributions of events in data. If the data set is of sufficient size these measurements can also be performed in bins of q^2.

3.3 Experimental status

As outlined at the beginning of Chapter 3, the decay $B_s^0 \to \phi\mu^+\mu^-$ is a promising FCNC decay to detect effects of NP. The LHCb detector (described in greater detail in Chapter 4) is designed to measure such decays with great precision. The data set recorded by LHCb during the 2011 run corresponds to a total integrated luminosity of about $1\,\mathrm{fb}^{-1}$. In preparation for measurements of more complex observables (such as A_{CP}) the branching fraction needs to be measured first in order to understand all experimental uncertainties. The 2011 LHCb data set is suited for such an analysis for $B_s^0 \to \phi\mu^+\mu^-$ and also $B^0 \to K^{*0}\mu^+\mu^-$ decays.

The $B_s^0 \to \phi\mu^+\mu^-$ decay was first observed by the Collider Detector at Fermilab (CDF) collaboration. CDF also measured the differential branching fraction with a total 49 ± 7 events in 6 bins of q^2 [48], where the statistical error is the dominant uncertainty.

This thesis reports the most precise measurement of the differential branching fraction of $B_s^0 \to \phi\mu^+\mu^-$ with $1\,\mathrm{fb}^{-1}$ of data collected by the LHCb detector.

Chapter 4

The LHCb detector

The LHCb detector is a single-arm spectrometer, designed to perform precise measurements of b and c quark decays. It is one of the four major experiments working at the Large Hadron Collider (LHC) at CERN in Geneva, Switzerland. In this chapter an overview of CERN and the LHC is given, followed by a more detailed description of the LHCb detector. The description is focussed on the hardware components of the detector, that are of relevance for the measurement of $B_s^0 \to \phi\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$ and similar decays. Validation of the performance of these components with first data is given. This is followed by an explanation of the data processing path of LHCb, including a description of the different trigger stages (hardware- and software-based), that events have to pass in order to be stored for physics analyses. Finally, details about the operation of LHCb are given (e.g. how a good data quality is ensured) and it is explained how the data are made available to the LHCb collaboration for analyses.

4.1 CERN

CERN was founded by 12 countries in Western Europe in 1954. Today it is an international organisation with about 2400 employed and many thousand visiting scientists and engineers. Since 1954 many experiments have been carried out at CERN that have changed the field of particle physics through numerous important discoveries. Amongst many others, these include the discovery of neutral currents and the W and Z gauge bosons [49, 50].

Today CERN is the home of the LHC, the world's largest and highest-energy particle accelerator. The accelerator ring has a circumference of 27 km and is about 100 m underground beneath the French-Swiss border.

The LHC accelerates bunches of protons in both clock-wise and anti-clock-wise directions (forming two proton beams). The design energy of the LHC is 7 TeV per beam. At four points along the accelerator magnets can make both beams cross and therefore cause proton-proton collisions at high energies. The four major experiments of the LHC are located at the four interaction points. They are A Large Ion Collider Experiment (ALICE) [51], A Toroidal LHC ApparatuS (ATLAS) [27], the Compact Muon Solenoid (CMS) [28] and LHCb [52]. The different purposes of the experiments have already been briefly described in Section 2.5.

Before the protons are injected into the LHC they are pre-accelerated by a series of smaller accelerators. The source of the protons is a cylinder of hydrogen gas. The gas is ionised to obtain protons, which are then fed into the Linac2. Linac2 accelerates the protons until they reach an energy of 50 MeV before injecting them into the Proton Synchrotron Booster (PSB). When they reach 1.5 GeV the protons are fed into the Proton Synchrotron (PS), which takes the energy up to 25 GeV. Finally the Super Proton Synchrotron (SPS) increases the energy to 450 GeV before injecting the protons into the LHC. The full accelerator chain and the positions of the four main experiments are shown in Figure 4.1.

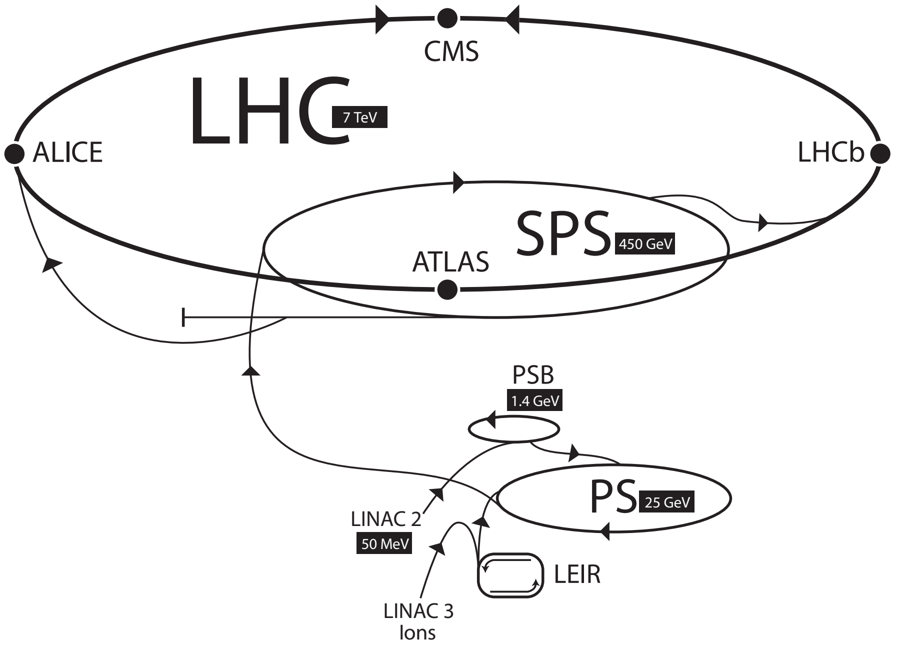


Figure 4.1: CERN's accelerator complex. When the proton beams reach an energy of 450 GeV in the SPS they are injected into the LHC. The positions of the four main experiments on the LHC ring are indicated. Taken from [53].

4.2 LHCb and its components of importance for $B_s^0 \to \phi\mu^+\mu^-$

In the following the detector components are described, which are required for an accurate reconstruction of $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays. Information from the Vertex Locator (VELO) and the tracking stations are used to reconstruct the charged tracks of the final state particles of these decays (kaons, pions and muons). The VELO also provides precise information on the position of the B-meson decay vertices and their separation from the corresponding primary vertices. The Ring Imaging Cherenkov Detectors (RICHs) provide separation power between kaon and pion tracks. The muon stations identify muon tracks.

At the LHCb interaction point the LHC optics can focus the beams, such that the instantaneous luminosity is reduced from $L \approx 10^{34}\,\mathrm{cm}^2\mathrm{s}^{-1}$ to a few $10^{32}\,\mathrm{cm}^2\mathrm{s}^{-1}$, which is closer to the design value of LHCb ($2 \times 10^{32}\,\mathrm{cm}^2\mathrm{s}^{-1}$). When the beams are focussed in this manner the number of proton-proton interactions per bunch crossing is reduced to ~ 1, as shown in Figure 4.2, such that an optimum ratio of signal over background events is reached. It also provides a cleaner environment with fewer pile-up events, which is of advantage when reconstructing rare B-meson decays.

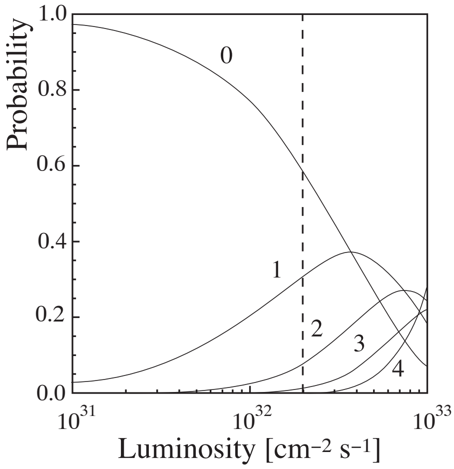


Figure 4.2: Probability of proton-proton interactions as a function of luminosity. The beams at the LHCb interaction point are focussed to deliver a luminosity of $2\times10^{32}\,\mathrm{cm}^2\mathrm{s}^{-1}$ in order to enhance single interactions. Taken from [54].

The $b\bar{b}$ production at the LHC is governed by the process of gluon fusion, $gg \to b\bar{b}$. The collision of one hard and one soft quark (or parton) and the consequent boost in the forward or backward direction means that most $b\bar{b}$ pairs are found in two narrow cones. The resulting angular correlation is shown in Figure 4.3, where θ_b ($\theta_{\bar{b}}$) is the opening angle between the flight direction of the b ($\bar{b}$) quark and the beam axis.

The angular correlation is also the motivation for the design of the experiment. The LHCb detector is a single-arm spectrometer, that covers one of the cones in the forward direction from 10 mrad to 300 mrad (250 mrad) in the horizontal (vertical) plane. Charged tracks are bent by the dipole magnet in the horizontal plane. The limit on the inner radius is due to the size of the beam pipe and the amount of radiation. The layout of the detector and its components is shown in Figure 4.4. Half of the produced $b\bar{b}$

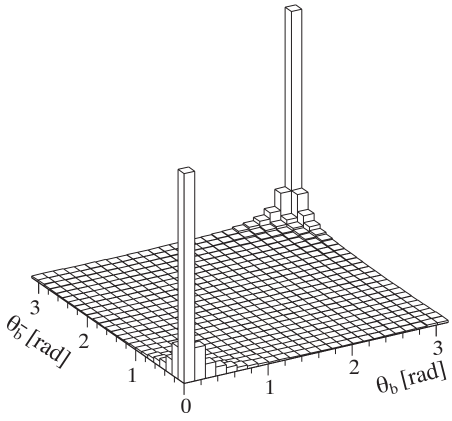


Figure 4.3: Correlation between the opening angles θ_b and $\theta_{\bar{b}}$ as predicted by Pythia 6. The opening angle is defined as the angle between the flight direction of the corresponding quark and the beam axis. The $b\bar{b}$ pairs are predominantly produced in a very narrow forward or backward cone. Taken from [54].

pairs are left undetected (they are inside the other cone), but therefore more space in the cavern is available for LHCb components, which in turn increases the accuracy of momentum measurements and the performance of the pattern recognition for tracks. A full description of the detector is given in [52].

4.2.1 Vertex Locator

The Vertex Locator (VELO) is a silicon vertex detector responsible for reconstructing primary (pp interactions) and secondary (decay) vertices. It is located at the LHC interaction point and consists of 21 silicon discs mounted in a vacuum vessel. The discs are double-sided sensors with a r-ϕ geometry. Figure 4.5 illustrates how one side is covered with radial strips and the other side with azimuthal strips. This geometry also enables fast 2-dimensional tracking, which is required for the trigger (as explained later in Section 4.3). Each disc is split vertically, such that the VELO consists of a left and a right module. To protect the modules from currents induced by the LHC beams there is a 200 μm thick aluminium foil in between the modules and the beams. The aluminium foil also preserves the primary vacuum of the beam pipe. Furthermore, both halves can be moved apart. Only when there are stable beams are both modules brought together, such that the active areas of the discs are separated from the beams by only 8 mm. During injection and acceleration of the beams the VELO modules are moved away for protection. To reduce the effects of radiation damage, which is particular high at the interaction point, the discs are kept at $-5°$C.

The Primary Vertex (PV) is reconstructed from the tracks of the prompt particles created in the proton-proton interaction. B-mesons are also produced at the PV but have a non-negligible lifetime. Due to their boost they travel a distance of ~ 7 mm inside the VELO before decaying. The Secondary Vertex (SV) is then formed from the tracks of

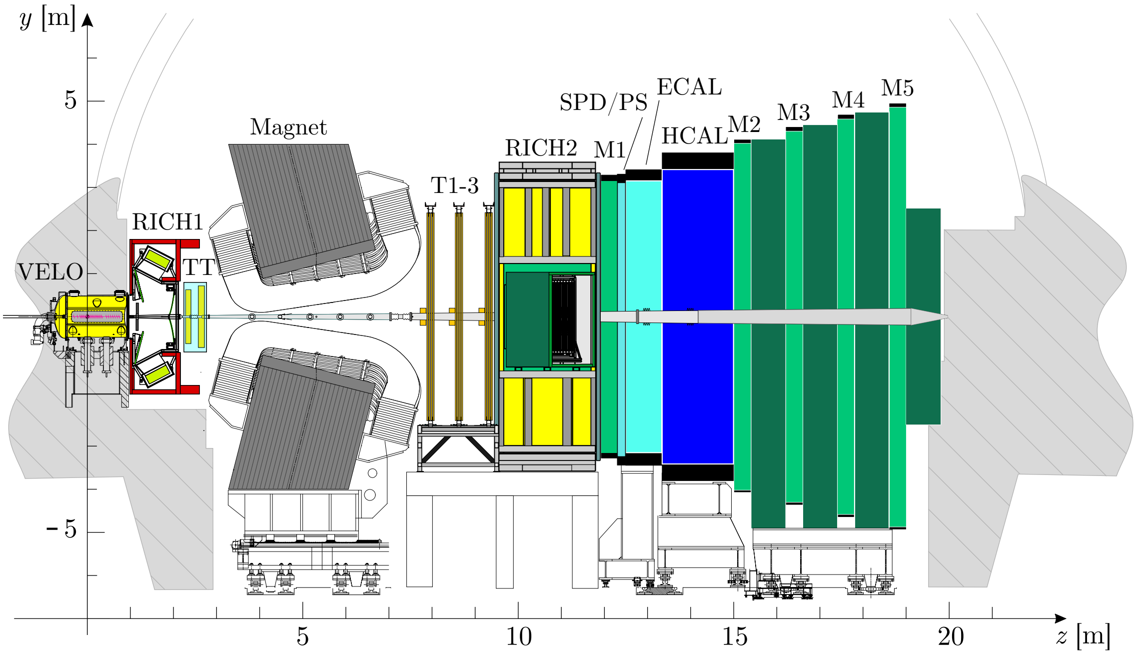


Figure 4.4: The LHCb detector and its components. The bending plane of the magnet is the x-z plane, where the z-axis points along the beam line and the x-axis comes 'out' of the page. Taken from [52].

the B-meson decay products. In the case of $B_s^0 \to \phi\mu^+\mu^-$ the SV is the origin of the K^+, K^-, μ^+ and μ^- tracks. Therefore the distance between the PV and the SV is an indirect measure of the B-meson lifetime, which is a powerful criterion to select B-meson candidates used by many analyses at LHCb.

Apart from providing the position and separation of the vertices, the VELO also reconstructs tracks. Both combined enables one to measure the Impact Parameters (IPs) of tracks. The IP of a track is the distance of closest approach between the trajectory of the track and the PV in the event. The IP is a powerful discriminator to distinguish between prompt background tracks from the PV and signal tracks from a SV. Many analyses cut on the IP of tracks and therefore a high resolution on the IP improves the accuracy and reduces systematic uncertainties.

The performance of the VELO was studied with data. Figure 4.6 shows the resolution on the PV position obtained from the 2011 data set and from Monte Carlo (MC) simulations. The PV resolution is measured by splitting the tracks of an event into two equal-sized subsets. The difference in vertex position when reconstructing a vertex for each set is taken as the resolution. Requiring that vertices are made from more than 25 tracks, average resolutions of $\sim 14\,\mu\text{m}$ and $\sim 85\,\mu\text{m}$ in the x/y- and in the z-direction are achieved, which match predictions from MC simulations [55].

The IP resolution is linked to the vertex resolution. Figure 4.7 compares the IP resolution measured for the 2011 data set against MC predictions as a function of the inverse p_{T} of a track. Only events with one reconstructed PV were considered. The PV must also

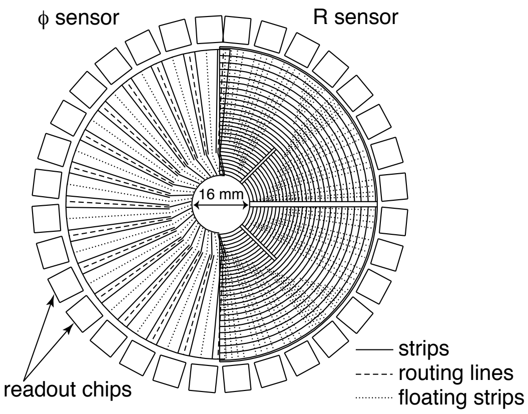
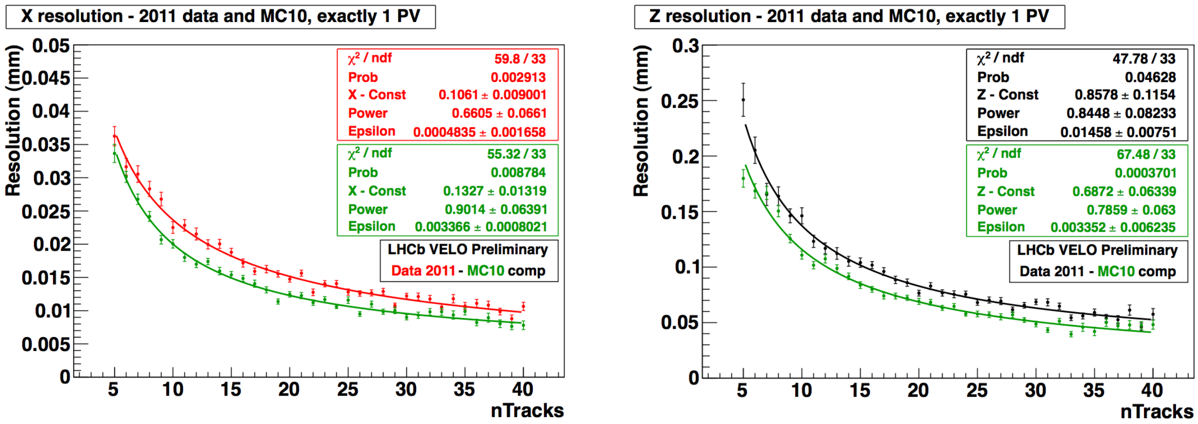


Figure 4.5: Illustration of the r-ϕ geometry of the VELO discs enabling the high spatial resolution. One side is covered with radial strips (R sensor) and the other side with azimuthal strips (ϕ sensor).



(a) (b)

Figure 4.6: PV spatial resolution in the x-direction (a) and in the z-direction (b) as a function of the number of tracks making the PV. Indicated are the distributions obtained from the 2011 data set (in red in (a) and in black in (b)) and from MC simulations (in green). Only events with one reconstructed PV were considered. Taken from [55].

be made from more than 25 tracks. The agreement between data and simulation in the relevant p_{T} range for the analysis of this thesis is good.
The discrepancies in Figures 4.6 and 4.7 between measurements and predictions are partly due to the material description (in terms of density and shape) in the simulation. The almost linear offsets between simulation and data distributions in both Figures indicate that the alignment is not tuned properly in these regions, which is the dominant factor causing the observed discrepancies.

4.2.2 Tracking stations

LHCb has four tracking stations, which are indicated in Figure 4.4: The Tracker Turicensis (TT) and three tracking stations (T1-T3). The TT is located upstream of the dipole

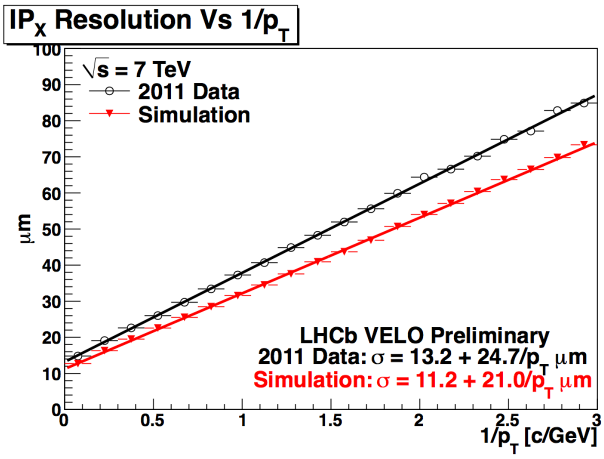


Figure 4.7: IP resolution in the x-direction as a function of the inverse p_{T} of charged tracks. Black circles represent data points from the 2011 data set and red triangles represent data points from MC simulations. The difference arises from slightly incorrect material descriptions (in terms of density and shape) and alignments in MC simulations. The agreement in the relevant p_{T} range for the analysis of this thesis ($> 2\,\mathrm{GeV}/c$) is good. The distribution for the resolution in the y-direction is similar. Taken from [55].

magnet and the three tracking stations downstream. The dipole magnet has a maximum field strength of $B_y = 1.1$T and a total magnetic field integral of about 4Tm. It bends charged particles in the x-z (horizontal) plane. Information from the VELO, TT and the T-stations are combined to form different track types. For example a ‘long’ track has hits in the VELO, TT and all three T-stations, whereas an ‘upstream track’ only has hits inside the VELO and TT before being swept out of the acceptance by the magnetic field. Figure 4.8 shows the evolution of the field strength and the different track types. The TT is a silicon strip detector, which covers the full acceptance of the detector. It has four layers and a total active area of $8.2\,\mathrm{m}^2$. Each T-station is divided into the Inner Tracker (IT) and the Outer Tracker (OT). The IT uses the same silicon micro-strip technology as used in the TT and covers an area of $120\,\mathrm{cm} \times 40\,\mathrm{cm}$ around the beam pipe. This region has the highest track occupancy. In order for the track pattern recognition to work the IT is required to have a spatial resolution of $50\,\mu\mathrm{m}$. Tracks found inside the inner region also have higher momenta than tracks further away, which is why a good spatial resolution in this region is important for a good dp/p resolution. The OT covers the area in the T-stations around the IT further away from the beam pipe. It consists of modules, where each module contains 256 straw tubes (2.4 m in length and 5 mm in diameter) filled with an argon-carbon-dioxide gas mixture [52]. With an electron drift time of about 45 ns inside the tubes a spatial resolution of about $250\,\mu\mathrm{m}$ is achieved [56].

The tracking efficiency for long tracks has been measured in data with the so-called ‘tag and probe’-method. In the decay $K_s^0 \to \pi^+\pi^-$ one pion track is reconstructed as a long track and is the *tag*. The other pion track is the *probe* and is reconstructed only partially with just VELO and calorimeter information. A high tracking efficiency is achieved when

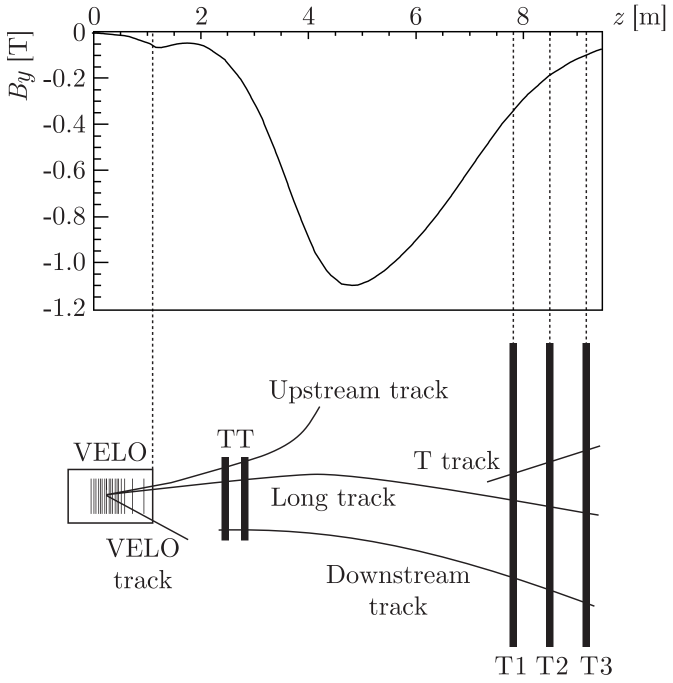


Figure 4.8: Track types at LHCb depend on the tracking components in which hits are detected. Tracks are only bent when travelling through the magnetic field. The magnetic field strength along the z-axis is indicated. Taken from [52].

there is always a matching track segment for the *probe* made out of hits in the T-stations. The tracking efficiency as a function of p_T is shown in Figure 4.9, which also shows a good agreement between MC simulations and 2010 data.

4.2.3 Ring Imaging Cherenkov Detectors

LHCb has two Ring Imaging Cherenkov Detectors (RICH1 and RICH2). RICH1 is located in between the VELO and the TT. RICH2 is in between the T3 and the first muon station (M1). The precise positions of both are shown in Figure 4.4. The purpose of both RICH1 and RICH2 is to provide Particle Identification (PID) for pions, kaons and protons over a wide momentum range. They make use of Cherenkov light[1] to measure the speed of charged particles.

The photons of the Cherenkov light are emitted in a cone described by the opening angle θ_c. θ_c is directly linked to the particle's speed βc and the medium's refractive index n through $\cos\theta_c = (n\beta)^{-1}$.

The momentum of charged tracks is obtained from the tracking (the curvature of a charged track inside the dipole magnet is inversely proportional to its momentum). By combining speed and momentum measurements of a given charged track, the mass of the particle associated with this track (and therefore the particle itself) can be identified.

---

[1] Cherenkov light is emitted by charged particles when they travel through a medium at a speed greater than the speed of light in that medium.

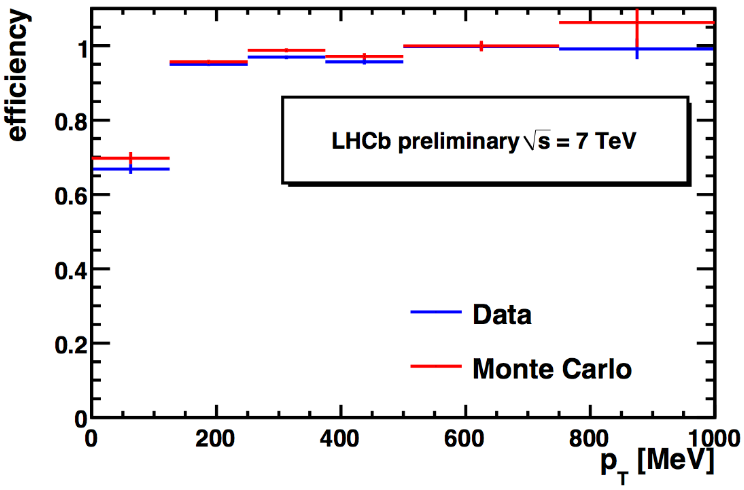


Figure 4.9: Tracking efficiency for 'long' tracks versus the p_{T} of the track in data (in blue) and MC simulations (in red). The efficiency was determined using the 'tag and probe'-method using $K_s^0 \to \pi^+\pi^-$ decays. Taken from [56].

To cover a wide momentum range different media (or radiators) with different refractive indices have to be used. RICH1 identifies particles with a momentum in the range $2 - 60\,\mathrm{GeV}/c$ by using two radiators: silica aerogel and a C_4F_{10} gas mixture. RICH2 identifies high momentum particles from $17 - 100\,\mathrm{GeV}/c$ and is therefore placed behind the magnet (which sweeps out low momentum particles). It uses CF_4 gas as radiator. For each radiator the distributions of θ_c for pions, kaons and protons as a function of momentum are shown in Figure 4.10.

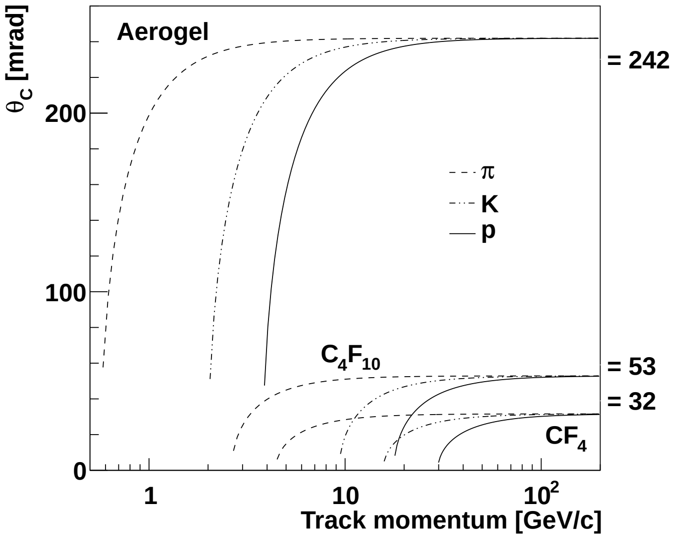


Figure 4.10: Cherenkov angle θ_c as a function of the track momentum and the radiator material. Indicated are the distributions for pions, kaons and protons. Taken from [52].

The working principles of both RICH1 and RICH2 are the same. It is illustrated in Figure 4.11 for RICH1. The particles travel through the radiator, where the Cherenkov photons are emitted. The photons are then focussed onto photon detector panels (sitting outside the acceptance) via spherical and flat mirrors. On the flat panels the photons are detected by Hybrid Photon Detectors (HPDs). The 'hits' on the HPD panels describe the shape of a ring because the photons were emitted as a cone in the radiator. These rings are identified by a pattern recognition software. From the radius of a ring the corresponding cone opening angle θ_c can be calculated. RICH1 is also shielded in an iron box to protect the sensitive HPDs from the magnetic stray field.

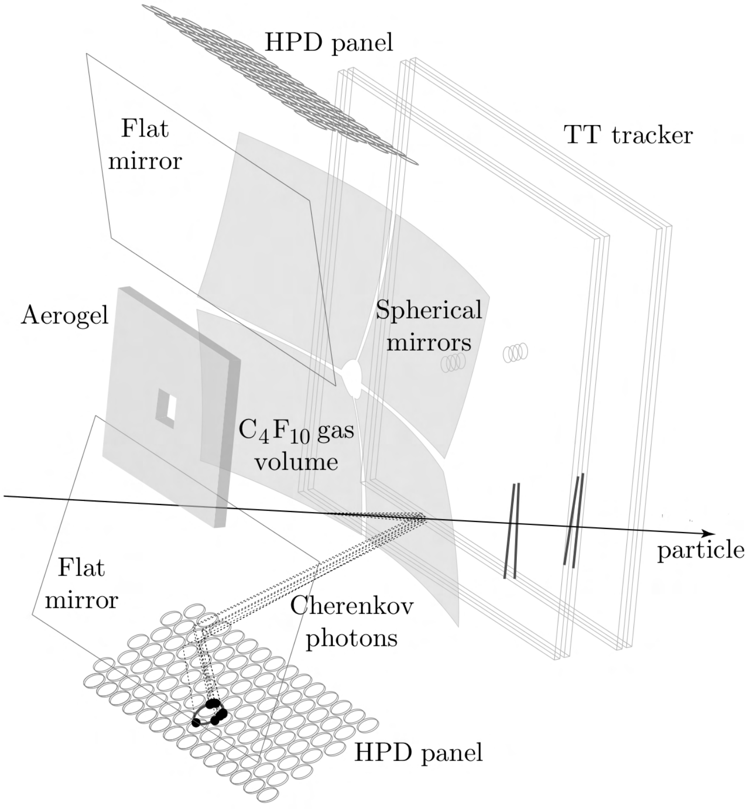
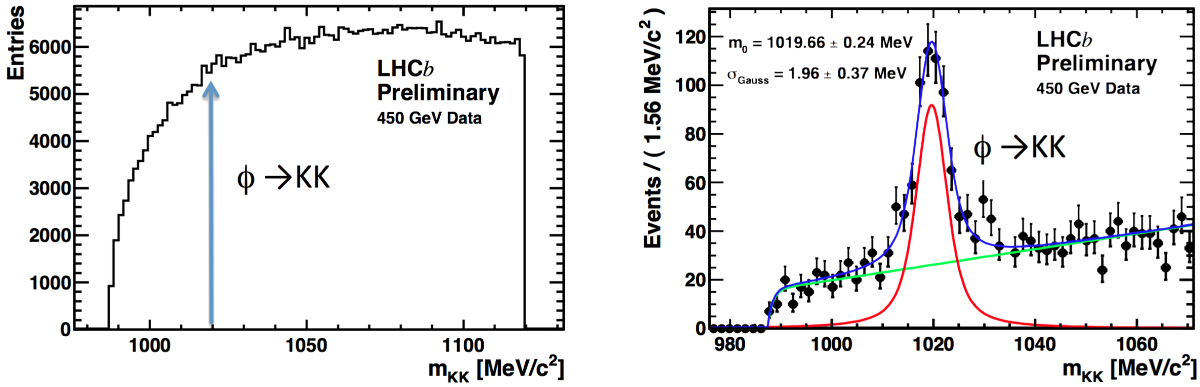


Figure 4.11: Working principle of RICH1. A charged particle emits Cherenkov photons when travelling through the C_4F_{10} gas mixture. The photons are then focussed onto HPD panels via spherical and flat mirrors. Taken from [53].

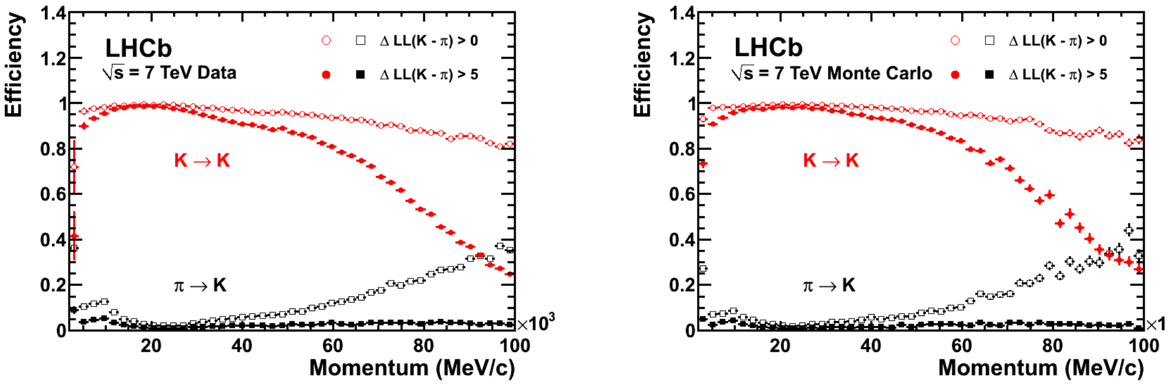
In the decay $B_s^0 \to \phi\mu^+\mu^-$ the ϕ meson is reconstructed via the mode $\phi \to K^+K^-$, which occurs in about 50% of all cases. The K^{*0} meson in $B^0 \to K^{*0}\mu^+\mu^-$ decays predominantly via the mode $K^{*0} \to K^+\pi^-$. Therefore a good separation between kaons and pions is crucial at LHCb. The PID information, that the RICH algorithms return, are the likelihoods for a track to be a certain particle. The separation power of the PID information is illustrated in Figure 4.12, which shows the invariant mass distribution of two kaon candidates forming a ϕ candidate ($\phi \to K^+K^-$) for two cases. In one case no PID information was used and the ϕ mass peak cannot be seen. In the other case it was required that the likelihood for each track to be a kaon is 10^{15} times greater than the likelihood for the track to be a pion, i.e. $\Delta \log \mathcal{L}(K - \pi) > 15$. In the latter case the ϕ mass peak becomes clearly visible.



(a) (b)

Figure 4.12: K^+K^- invariant mass distributions for when there is no RICH PID requirement (a) and for when the condition $\Delta \log \mathcal{L}(K-\pi) > 15$ has to be true for each kaon track (b). The decay $\phi \to K^+K^-$ becomes clearly in (b). Taken from [57].

The kaon efficiency (kaons identified as kaons) and the pion misidentification fraction (pions misidentified as kaons) as a function of momentum for different PID cuts are shown in Figure 4.13. As can be seen the MC predictions are in good agreement with the data.



(a) (b)

Figure 4.13: Kaon identification efficiency (red circles) and pion misidentification rate (black squares) as a function of momentum measured in data (a) and MC simulation (b). Two different $\Delta \log \mathcal{L}(K-\pi)$ requirements are shown: > 0 (hollow data points) and > 5 (solid data points). Taken from [58].

4.2.4 Muon stations

Many measurements at LHCb, such as $B_s^0 \to \mu^+\mu^-$, $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$, rely on good muon identification. For this purpose there are five muon stations (M1-M5) as shown in Figure 4.4. Muons are extremely penetrating and have flight distances greater than the diameter of the cavern. Therefore the muon stations are placed at the end of the spectrometer arm. M1 is located in between the RICH2 and the calorimeters. M2-

M5 are the sub-detectors the furthest away from the VELO and are downstream of the calorimeters.

The muon stations starting from M2 are separated by 80 cm-thick iron plates in order to absorb all remaining hadrons, that pass the calorimeters, and only let muons pass. A muon requires a momentum of $\sim 6\,\mathrm{GeV}/c$ to be able to cross all of the five muon stations. Inside the stations Multi-Wire Proportional Chambers (MWPCs) are used to observe penetrating charged particles. The MWPCs are filled with a Ar-CO_2-CF_4 gas mixture, which gives a time-resolution of about 5 ns as required by the trigger. The total active area of M1-M5 is $435\,\mathrm{m}^2$.

The flux of particles in the region close to the beam pipe is greater than on the outer edge of the acceptance. To increase the resolution in this area, smaller and more sensitive elements are used in the central region of each station. Out of the five stations M1 experiences the highest flux of particles as it is located upstream of the calorimeters. In the central region of M1 the particle rate exceeds safety limits for ageing and Gas Electron Multiplier (GEM) detectors are used instead of MWPCs.

Apart from muon PID the muon stations also provide a rough estimate of the p_{T} of a particle. By considering the relative position of 'hits' in the five stations an accuracy of 20% is achieved. This information is fast to obtain and is used in the muon trigger lines. The details are given later in Section 4.3.1.

The muon PID efficiency has been measured with the 'tag and probe'-method, which was also used to determine the tracking efficiency (described in Section 4.2.2). One muon track of the decay $J/\psi \to \mu^+\mu^-$ was used as the *tag* and the other one as the *probe*. The efficiency is very high over a wide momentum range (around 95%) in both data and MC simulations. The distributions are given in Figure 4.14. The muon misidentification rate was measured using $\Lambda \to p\pi^-$ decays with the same technique, giving a rate of 2.4% for tracks with $p > 3\,\mathrm{GeV}/c$ [59, 60]. In an analysis dedicated to the search for the $D^0 \to \mu^+\mu^-$ decay the muon misidentification rate for high momentum tracks ($p > 20\,\mathrm{GeV}/c$) was measured to be less than 1% [61].

4.3 Trigger

The number of events produced inside the LHCb detector per second is too large to be able to write the information of all of them to disc. Therefore the trigger system of LHCb has to reduce this rate significantly. At the same time the trigger needs to keep the events of interest, in this case the rare decays $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$. In the following the trigger structure and which kinematic properties it is exploiting to select these decays are explained. To conclude, details of a method are given, which can be used to estimate the trigger efficiency for decays such as $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$.

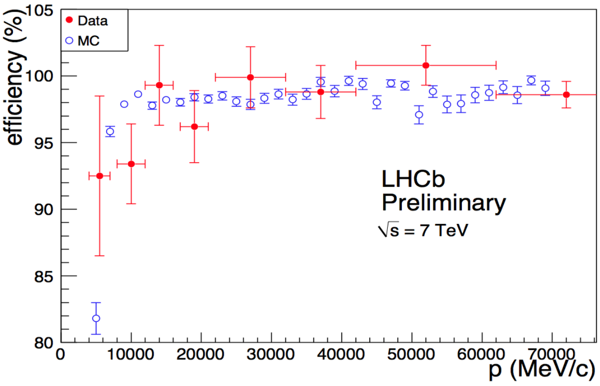


Figure 4.14: Muon identification efficiency versus momentum in data (solid data points) and in MC simulations (hollow data points). Taken from [57].

The LHC design bunch-crossing rate is 40 MHz, which corresponds to a 25 ns spacing between bunches. Due to the bunch structure the rate of visible proton-proton interactions at LHCb is closer to 10 MHz during nominal running conditions. A visible interaction is defined as an interaction which produces a minimum of two charged particles, that are reconstructed.

At a luminosity of $2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ the rate of $b\bar{b}$ pair production is about 100 kHz. Only about 15% of these events will include a b quark, that fully decays inside the LHCb acceptance, and are therefore of interest for physics analyses.

The rate at which a complete detector readout of an interaction can be written to tape is 3 kHz (the design value was 2 kHz). Therefore not all events of interest can be stored for further offline analysis. The required significant reduction in rate from 10 MHz to 3 kHz is achieved by a two-level trigger system. It is designed to keep a high fraction of events containing b quarks whilst rejecting background events.

The first level of the LHCb trigger is called the Level 0 (L0) trigger and is implemented in the hardware. The second level is referred to as the High Level Trigger (HLT), which is based on decisions made by software. The HLT is split further into two levels, called HLT1 and HLT2, for timing reasons as will be outlined later. Each trigger stage consists of certain ‘alleys’. Certain alleys are designed to decide whether an event is useful for physics analyses, control or calibration samples, based on certain kinematic features of the event. A flow-diagram of the LHCb trigger sequence and the alleys involved is shown in Figure 4.15.

4.3.1 L0

The L0 trigger is implemented in the LHCb hardware and reduces the rate of visible events down from 10 MHz to 1 MHz. Therefore it needs to be very fast in deciding whether to keep events or throw them away whilst maintaining a high efficiency for signal events. The time available is much less than it takes to read out every channel of each sub-detector

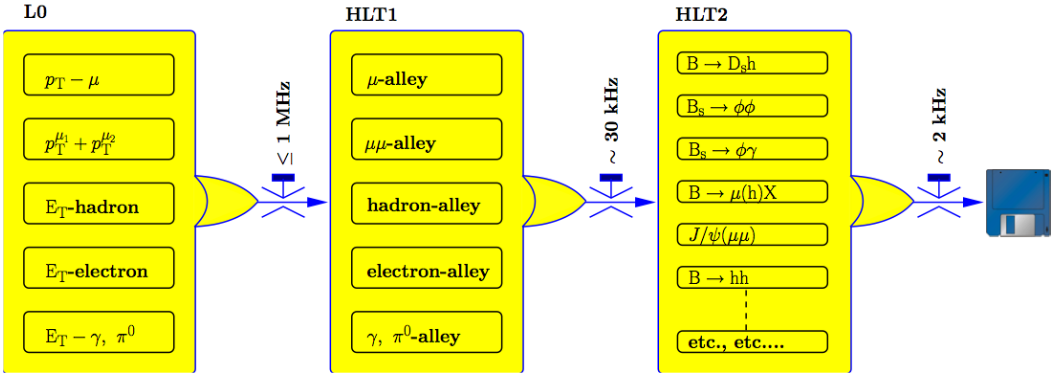


Figure 4.15: Flow-diagram of the different trigger stages. Indicated are the required rates after each stage and some of the different trigger lines per stage. Taken from [52].

and reconstruct an event fully. As a consequence the L0 trigger attempts to reconstruct only the parts of an event that are characteristic for B-meson decays. Of interest are high E_{T} clusters in the hadronic and electromagnetic calorimeters and the two highest p_{T} muons in the muon stations [52]. Additionally a pile-up veto allows one to reject events where there are several proton-proton interactions. These events would contain too many tracks to reconstruct and therefore would take a disproportionate fraction of computing power available during the later HLT stages.

The L0 is divided into three sub-systems, described below, which each provide a decision. The final L0 decision is a logical OR combination of the three sub-system decisions. A schematic of the L0 architecture is shown in Figure 4.16.

For $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays, the majority of events are selected by the lines in the muon trigger. However, due to the ϕ and K^{*0} the calorimeter trigger lines also select a non-negligible fraction of these decays.

- **The muon trigger** - The muon trigger uses the muon stations to perform a fast stand-alone muon reconstruction with a p_{T} resolution of $\sim 20\%$. Consequently, the two tracks with the highest p_{T}, which are also pointing back to the interaction region at the same time, are selected. The straight-line track finding algorithm is started as soon as any hits in the third muon station (M3) are detected (so-called seed hits). Tracks are then found by combining hit positions in all five muon stations (M1-M5). As described earlier in Section 4.2.4 each muon station is divided into several regions, R1-R4. The same regions of the five muon stations are then combined to form 192 'towers', spanning across all stations and pointing towards the interaction region. The output of each of the 192 towers is read into a separate processing element. Each processing element is capable of executing 96 tracking algorithms in parallel with the use of Field-Programmable Gate Arrays (FPGAs). The two highest p_{T} tracks in each tower are found and finally a chain of controller boards selects the two highest p_{T} tracks in the detector overall.

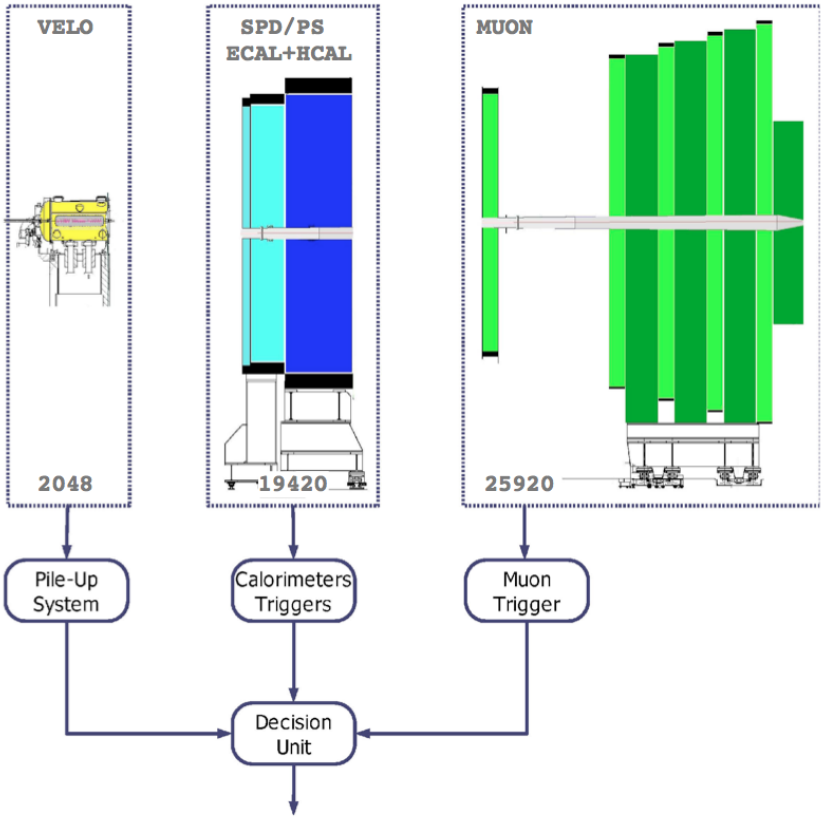


Figure 4.16: Overview of the three decision making systems of the L0 trigger. Indicated are the number of channels for each. The final trigger decision is a logical OR combination of the three decision from the pile-up system, the calorimeter trigger and the muon trigger. Taken from [52].

Currently events are triggered if there is a single muon candidate with $p_{\mathrm{T}} > 1.3\,\mathrm{GeV}/c$ or two muon candidates with a combined $p_{\mathrm{T}} > 1.5\,\mathrm{GeV}/c$ [60]. The muon trigger is the main L0 trigger for selecting $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays due to the two muons in the final state. The efficiency with respect to offline selected $B^0 \to K^{*0}\mu^+\mu^-$ events is estimated to be around 90% and is similar for $B_s^0 \to \phi\mu^+\mu^-$ events [62].

- **The calorimeter trigger** - The calorimeter trigger selects high E_{T} particles (electrons, photons or hadrons) and is therefore used by all L0 hadron alleys. It uses the information from the Scintillating Pad Detector (SPD), the Pre-Shower Detector (PSD), the Electromagnetic Calorimeter (ECAL) and the Hadron Calorimeter (HCAL). Within these sub-detectors the calorimeter trigger is looking for clusters of high E_{T} in 2×2 cells. The cluster with the largest E_{T} gets selected. A fast PID is achieved by considering in which sub-detectors clusters appear: Photons leave hits in the PSD and ECAL but not in the SPD, whereas electrons leave hits in all three of them. Hadron candidates leave hits inside the ECAL and HCAL.

  The hadronic trigger selects $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ events because of the ϕ and K^{*0} particles with an efficiency of about 27% [62].

- **The pile-up system** - As explained earlier the L0 trigger can veto events with multiple interactions in order so save CPU power during the high level trigger stage. In order to do so the pile-up system sits upstream of the VELO and consists of the first two silicon strip discs of the VELO. The influence of the magnet is negligible in this part of the detector and therefore all tracks are assumed to be a straight line. Connecting hits on the two planes and extrapolating their path and origin vertex allows one to make a fast estimation of the number of interactions during a bunch-crossing.
  Multiple decay candidates in a single event with a high track multiplicity can be problematic for various reasons. However, due to tight selection criteria for $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays used in this thesis there was usually only one candidate per event.

4.3.2 HLT1

HLT1 (as well as HLT2) is implemented in `C++` and runs on the so-called Event Filter Farm, located underground and consisting of ~ 2000 computing nodes. The advantage of a software-based trigger is that it can be modified according to changing running conditions or other unforeseen changes. As shown in Figure 4.15 HLT1 reduces the incoming rate of the L0 trigger from about 1 MHz to 30 kHz. This is achieved by categorising events into different physics alleys with different selection criteria.

The incoming rate of the HLT1 is still too high to completely reconstruct an event using the information of all sub-detectors. Therefore HLT1 is designed such that it first confirms the L0 trigger decision and then adds extra information to the event from selected sub-detectors. It adds information about the hits in the tracking stations and the VELO such that the p_{T} and IP resolution can be improved. The IP is the distance of closest approach between a track and the PV inside the VELO. As a consequence it is possible to apply a variety of additional cuts in each of the HLT1 alleys in order to reduce the event rate down to 30 kHz and to keep signal events.

In the following an overview of the HLT1 alleys is given. The two muon alleys described first are highly efficient in triggering $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays and therefore explained in greater detail.

- **Muon + track** - This alley aims to provide a high trigger efficiency for decays containing muons. It is only executed on events triggered by a single muon candidate during the L0 stage. The first step is to confirm the muon candidate. In order to do so the positions of hits in the first two muon stations (M1-M2) are used to form track ‘segments’. These track segments are consequently matched to hits in the tracking stations. In the case of a good match the track is then extrapolated back to the other muon stations (M3-M5) to confirm the track type and therefore the L0 candidate. Finally VELO hit information is used to form a 3-dimensional vertex out of the muon candidate and an other charged track in the event. In addition to a good vertex it is required that $p_{\mathrm{T}}(\mu, \mathrm{track}) > (1.0, 0.8)\,\mathrm{GeV}/c$ and $\mathrm{IP}(\mu, \mathrm{track}) > (25, 50)\,\mu\mathrm{m}$. If all criteria are met then the event is selected. This alley is highly efficient for $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays ($(87.2 \pm 0.2)\%$ with respect to selected events) [62, 63].

- **Muon** - The muon alley comprises several muon trigger lines. The first step is that L0 muons are confirmed in a similar manner to the 'muon+track' alley. One difference between the lines is that some require two muon candidates ('dimuon' events) whereas the other only one ('single muon' events). Furthermore, some lines require only one cut on the p_{T} or the IP as opposed to cutting on both. The lines cutting on only one quantity must therefore have a harder cut-value than lines with two cuts. For example single muon candidates are accepted if they have $p_{\mathrm{T}} > 1.3\,\mathrm{GeV}/c$ and $\mathrm{IP} > 0.8\,\mathrm{mm}$, but candidates with $p_{\mathrm{T}} > 6.0\,\mathrm{GeV}/c$ are accepted without any IP requirements [64, 65]. The lines in this alley are also highly efficient in triggering $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ events ($\sim 80\%$ with respect to selected events) [62].

- **Hadron** - The HLT1 hadron alley takes hadron candidates from the L0 calorimeter trigger as input. They are confirmed if the p_{T} is greater than $2.5\,\mathrm{GeV}/c$ and the IP greater than $0.1\,\mathrm{mm}$. If there is an extra track in the event forming a secondary vertex in the VELO the di-hadron line is triggered. Further details about the cut values of the HLT1 single hadron and the di-hadron lines can be found in [66].

- **Electromagnetic** - The electromagnetic alley is dedicated to select radiative decays. It consists of several lines optimised to trigger on decays with photons and electrons. L0 electron and photon candidates are confirmed and selected if the p_{T} is greater than $2.8\,\mathrm{GeV}/c$ and the IP greater than $0.15\,\mathrm{mm}$ [5, 67].

4.3.3 HLT2

HLT2 uses the full detector readout to reconstruct events coming from the HLT1 alleys. Therefore composite particles (such as $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$) are formed at this stage and more cuts can be applied. The aim is to reduce the rate of events further down to 3 kHz, the rate at which data can be written to tape for further offline analysis.

The lines in HLT2 are either of exclusive or inclusive nature. Exclusive lines can contain custom algorithms and are designed to select specific decays for an analysis. The cut values are usually chosen to be relaxed with respect to the offline cut values of an analysis. This has the advantage that the precise cut values can be fine-tuned afterwards with the aid of multivariate methods. Both $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays are efficiently selected by inclusive HLT2 lines, such that no exclusive HLT2 lines were required. Inclusive lines in HLT2 make use of the improved mass, IP and momentum resolution to reduce the background rates of the HLT1 lines even further. For example the HLT2 topological lines have been designed to trigger inclusively on n-body b decays. The HLT2 overall trigger decision is a logical OR combination of the decisions of the exclusive and inclusive lines.

The HLT2 topological lines trigger efficiently on any b (or even c) quark decay with a minimum of two charged daughters. They are therefore responsible for selecting the majority of the incoming $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays. The exact methodology is described in greater detail in [68,69]. Inclusive muon triggers also contribute significantly. Brief descriptions of the relevant inclusive lines are given in the following:

- **Topological** - The trigger efficiency of the topological trigger lines on selected $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ events is estimated to be very high ($> 90\%$) [68]. All tracks in the VELO going in the forward direction are considered. To save CPU power and narrow the search window it is required that $p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$ and $p > 5\,\mathrm{GeV}/c$. Because the topological trigger is inclusive no tight cuts on exclusive quantities can be made. This includes the masses and IPs of candidates. The finite B-meson lifetime allows one to reject background tracks coming directly from the PV by requiring a χ^2_{IP} of greater than 4 for all tracks. χ^2_{IP} is formed by the hypothesis that the track's IP with respect to the PV is equal to zero. The quality of each track, expressed in χ^2, is required to be less than 3, except for muon candidates, where a track quality of less than 4 is required. Tracks passing these criteria are the input particles to the algorithm.

  First 2-body proto-candidates are made from two input particles, that have the same best PV. In the case of multiple PVs in an event the best PV is the one to which a particle has the smallest IP. If the sum of the transverse momenta, $\sum |p_{\mathrm{T}}|$, is greater than $3\,\mathrm{GeV}/c$ and the invariant mass is less than $7\,\mathrm{GeV}/c^2$ the proto-candidate is used as input for the 2-body topological trigger. If there is another input particle in the same event it is added to the 2-body proto-candidate (which is now treated like a single particle) to form a 3-body object. In the same manner a fourth input particle (if there is one in the event) is added to the 3-body proto-candidate to form a 4-body object. In general a n-body proto-candidate is formed from two particles (the $(n-1)$-body candidate and another particle) and not from combining n particles.

  Finally the different n-body topological lines apply a multivariate selection to the n-body filtered proto-candidates. The multivariate selection used is a Boosted Decision Tree (BDT), which performs repetitive one-dimensional splits of the data in order to find the optimal cut value (the figure of merit being the signal significance). The BDT was tuned such that its performance is superior to a simple cut-based approach, where attention was paid not to ‘overtrain’ the BDT. This stage of the topological trigger lines is where the majority of background events are rejected. The procedure is described in greater detail in [68].

- **Muon** - There are both inclusive single muon and dimuon triggers in the HLT2 serving different purposes as described in [64]. The single muon trigger requires the HLT1 muon+track trigger to be true as well as a very good track quality ($\chi^2 < 2$), p_{T} greater than $1.3\,\mathrm{GeV}/c$ and a large IP and χ^2_{IP} (IP $> 0.5\,\mathrm{mm}$ and $\chi^2_{\mathrm{IP}} > 200$). It selects semileptonic decays of B- and D-mesons. The dimuon trigger distinguishes between prompt and detached dimuons, where the first one is only run as long as HLT2 output rates allow it to. The detached dimuon line requires the dimuon candidate to have a mass above $1\,\mathrm{GeV}/c^2$, a significant χ^2_{IP} of greater than 9 for both muons and a clear separation between dimuon vertex and PV.
  The efficiency of the muon lines on selected $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0_s \to \phi\mu^+\mu^-$ events is estimated to be around 90% [62].

4.3.4 TISTOS method

The trigger efficiencies of the different stages (L0, HLT1 and HLT2) and their combinations can be calculated using the so-called TISTOS method [69–71]. In the following the term ‘signal’ refers to the combination of tracks, which are used for the offline analysis. In this case the ‘signal’ corresponds to the reconstructed $B^0_s \to \phi\mu^+\mu^-$ (or $B^0 \to K^{*0}\mu^+\mu^-$) candidates. The term ‘trigger object’ is defined as the part of an event (i.e. a collection of certain tracks) that caused one of the trigger lines to return a positive decision.
There are three categories of trigger decisions (TIS, TOS and TISTOS), which will be defined in the following. The number of entries in each category can be used to calculate the trigger efficiencies required.

- **TIS** - Triggered Independently of Signal: Events, which are triggered explicitly independent of any signal, belong in this category. Therefore the event must contain at least one trigger object, which does *not* share any tracks with the signal. To test whether there is an overlap between tracks or not the so-called LHCBIDs of the tracks are compared. LHCBIDs are track identifiers (or hits) of the different sub-detectors. If the number of shared LHCBIDs between two tracks of the trigger object and the signal is greater than 1% then both tracks are overlapping. Consequently the event would not be of the TIS category.

- **TOS** - Triggered On Signal: Events, which are explicitly triggered because of the signal, belong in this category. Therefore all tracks of the trigger object must overlap with the tracks of the signal. In this case two tracks overlap if more than 70% of the LHCBIDs are shared.

- **TISTOS**: Events belonging in this category satisfy both criteria for being TIS and TOS at the same time. Therefore the event must contain several trigger objects. In the case of two trigger objects one of them has to match the signal (TOS) and the other one has to be separated from the signal (TIS).

After defining the trigger lines for each trigger stage (L0, HLT1 and HLT2), which are responsible to select events, the TISTOS method can be applied on data. In practice attention has to be paid to changing trigger conditions. During the course of the 2011 data taking period the software-based part of the trigger has been changing ever so slightly to accommodate different collision scenarios. This has to be taken into account by a systematic uncertainty.

The trigger efficiency in general is given by:

$$\varepsilon^{\text{trigger}} = \frac{N^{\text{after}}}{N^{\text{before}}} \tag{4.1}$$

where N^{after} is the number of events, which pass the trigger. N^{before} is calculated with the TISTOS method and is given by:

$$N^{\text{before}} = N^{\text{TIS}} + N^{\text{TOS}} + N^{\text{TISTOS}} + \frac{N^{\text{TIS}} \times N^{\text{TOS}}}{N^{\text{TISTOS}}} \tag{4.2}$$

where the N^{TIS}, N^{TOS} and N^{TISTOS} are the number of events in the categories TIS, TOS and TISTOS respectively.

Furthermore, a TOS-trigger efficiency can be defined. It is the efficiency of the trigger on events, that are selected by the defined TOS trigger lines only. It is given by:

$$\varepsilon^{\text{TOS}} = \frac{N^{\text{TISTOS}}}{N^{\text{TIS}} + N^{\text{TISTOS}}} \tag{4.3}$$

An equivalent expression for ε^{TIS} can be constructed but is not of relevance.

The TISTOS method can be used to calculate the trigger efficiencies for L0, HLT1 and HLT2. The value of interest in analyses is always the efficiency of the combined trigger (L0×HLT1×HLT2). It is straightforward to show that combining the conditional efficiencies for each trigger stage leads to the total trigger efficiency as defined in Equation 4.1 [72]:

- The L0 trigger efficiency is the number of events passing L0 over the number of events before all triggers:

$$\varepsilon^{\text{L0}} = \frac{N^{\text{L0}}}{N^{\text{before}}} \tag{4.4}$$

- The conditional HLT1 trigger efficiency is the number of events passing L0 and

HLT1 over the number of events passing L0:

$$\varepsilon^{\mathrm{HLT1}} = \frac{N^{\mathrm{L0}\times\mathrm{HLT1}}}{N^{\mathrm{L0}}} \tag{4.5}$$

- The conditional HLT2 trigger efficiency is the number of events passing L0, HLT1 and HLT2 over the number of events passing L0 and HLT1:

$$\varepsilon^{\mathrm{HLT2}} = \frac{N^{\mathrm{L0}\times\mathrm{HLT1}\times\mathrm{HLT2}}}{N^{\mathrm{L0}\times\mathrm{HLT1}}} \tag{4.6}$$

- The combined L0×HLT1 trigger efficiency is the number of events passing L0 and HLT1 over the number of events before all triggers:

$$\varepsilon^{\mathrm{L0}\times\mathrm{HLT1}} = \varepsilon^{\mathrm{L0}} \times \varepsilon^{\mathrm{HLT1}} = \frac{N^{\mathrm{L0}\times\mathrm{HLT1}}}{N^{\mathrm{before}}} \tag{4.7}$$

- The total combined trigger efficiency (L0×HLT1×HLT2) is the number of events passing L0, HLT1 and HLT2 over the number of events before all triggers:

$$\varepsilon^{\mathrm{L0}\times\mathrm{HLT1}\times\mathrm{HLT2}} = \varepsilon^{\mathrm{L0}} \times \varepsilon^{\mathrm{HLT1}} \times \varepsilon^{\mathrm{HLT2}} = \frac{N^{\mathrm{L0}\times\mathrm{HLT1}\times\mathrm{HLT2}}}{N^{\mathrm{before}}} \tag{4.8}$$

As can be seen Equation 4.8 is equivalent to the general expression given in Equation 4.1. It is advantageous to be able to calculate the various trigger efficiencies (conditional or combined and total efficiency or TOS efficiency) in order to perform a variety of cross-checks.

The TISTOS method is used in Chapters 5 and 6 to determine trigger efficiencies. There more details about the choice of trigger lines for the TIS and TOS categories and the cross-checks are given.

4.4 Data Taking and Data Quality

In the following the data processing procedure at LHCb is explained. The order is chronological and describes the 'path' of the data in order to become accessible to the collaboration for physics analyses. The contributions of the author are indicated. The stripping procedure described in Section 4.4.3 is of particular importance for $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ events. The stripping is similar to a final trigger stage. Its aim is to organise the data in such a way that they are easily accessible to the collaboration on discs, where they can be read from repeatedly. Disc space and CPU power put constraints on the retention and timing of the stripping stage. The stripping is organised into several lines for different decay channels. Each line has selection requirements in order to remove background events and satisfy the retention and timing limits. The author developed two lines to specifically select $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ events, of which the details are given in Section 4.4.3.

4.4.1 Data manager

When the LHC is injecting protons and accelerating them to the desired energy of the fill the experiments need to be in a ‘ready’-state to avoid any damage. The precise fill plan for each the day is announced to the experiments in daily run-meetings in the morning. For each fill the experiments on the ring need to confirm that they are ‘ready’ for data taking with a virtual handshake.

When the LHC switches into collision mode the VELO closes and the trigger (described in Section 4.3) starts to receive input data from the sub-detectors. LHCb labels the data according to the fill (with a unique fill-number). Each fill is again divided into several runs (labelled with a run-number). For example, if a sub-detector reports problems during a fill then it can be turned off and a new run is started. Internally the new run is then flagged as one with missing sub-detector information.

At LHCb there are always at least two persons in the control room - the shift leader and the data manager. It is ensured that the two positions are always covered with a shift system. There are three 8-hour-shifts per day for each position. The sub-detectors are always covered with a piquet system, such that the data manager and shift leader can always get in contact with an expert over the phone at anytime. Most of the sub-detectors can also be controlled or maintained remotely. The author has contributed to the experiment by taking numerous data manager shifts himself during the course of 2010 and 2011.

The role of the shift leader is to be in charge of the experiment and all systems as well as being the primary contact for the LHC control room. Amongst many other things he/she has to resolve any problems with the sub-detectors and can start new runs manually. The data manager assists the shift leader with these tasks and also manages the incoming data. Managing the data means to have a first look at various distributions to ensure a good data quality. For this purpose several histograms are available to the data manager via the LHCb Online Presenter software. With this it is possible to check:

- The preliminary vertex fits of the VELO. This is especially important to determine whether the VELO closing procedure has been successful. In the case of no or too few reconstructed vertices there is most likely a problem and the run should be stopped.

- Clusters in the tracking stations. 2-dimensional histograms representing the different cells of the tracking stations show the distribution of hit clusters. For example zero entries in a cell could indicate dead tracking elements, which would require one to contact the corresponding piquet.

- Photon yields in the RICH HPDs (for both RICH1 and RICH2). A map shows each HPD and indicates the photon yields through a colour-coding scheme. Dysfunctional or dead HPDs can be identified easily.
- Hits in the calorimeters. 2-dimensional hit maps of the calorimeters make it possible to detect anomalies. For example, one always expects a higher hit-density towards the centre around the beam pipe due to small opening angles of most tracks.
- Occupancy of the muon stations. In a similar manner to the tracking stations and calorimeters an occupancy map for the five muon stations exists, on which anomalies and dead elements show up easily.
- Trigger rates. All different trigger stages are run ‘online’ and their output rate is plotted against time. If it exceeds or is below the designed rate this could indicate a severe problem.

For each distribution a reference distribution in red is shown. Any deviation from the reference distribution must be dealt with unless stated otherwise (for example in the logbook). Each sub-detector also provides an error bank, where all sorts of error messages get collected. A comprehensive documentation gives instruction to resolve errors. Furthermore, it is important that the data manager keeps a record of all problems in the logbook, such that the next data manager on shift is aware of any previous incidents.

4.4.2 Data quality

The output of the trigger system is written to storage. Before the data can be made available to the collaboration they need to be fully reconstructed. In the full reconstruction tracks are turned into ‘objects’ of the LHCb software framework, which represent particles and contain all available information about the kinematic properties. In the reconstruction the precise data taking conditions are also considered. For example missing cells or elements in one of the sub-detectors or different trigger settings are accounted for. Several LHCb specific software packages are used throughout. If there are major differences in the reconstruction algorithms and software packages used a new full reconstruction is started. New full reconstructions are assigned a new ‘reconstruction version’.

A full reconstruction of all data on tape is computational intensive and therefore time-consuming. A good data quality must be ensured beforehand. This is achieved by fully reconstructing only a small sub-set of each run. Each sub-set of each run (called the `EXPRESS`-stream) is then examined by the data quality shifter and flagged accordingly. If flagged as ‘OK’ the full reconstruction of the remaining data in the run is started. If flagged as 'BAD' the data should not be used for analyses or only in special circumstances. The data quality shifter uses the LHCb presenter software to look at kinematic distributions in the data. Divergences to the expected shapes of distributions can indicate problems. An example relevant to this thesis is the control of the J/ψ signal parameters. If the J/ψ mass peak in the dimuon mass spectrum is not clear or if it is shifted to the side then the data might have to be marked as bad. Many other distributions are available for cross-checks.

Depending on the size of the data samples a new reconstruction version including a data quality check for each run takes a couple of weeks. The author has been actively involved as a data quality shifter for different reconstruction versions in the years 2010 and 2011. The full data set of 2011 has been reconstructed with one coherent version (`Reco12`). The total integrated luminosity corresponds to $\sim 1\,\mathrm{fb}^{-1}$.

4.4.3 Stripping lines

After the reconstruction a preliminary selection is applied on the data. This selection is referred to as 'stripping' and is similar to the HLT2 framework. It consists of multiple lines, which reconstruct specific decays using the particle objects in the code, that are now fully available.

Each line applies cuts to further remove background events and satisfy retention and timing limits. Different analysis groups develop their own stripping lines to select events for their analyses. The cuts of the stripping lines are therefore usually a relaxed version of the final cuts of an analysis.

The stripping framework changes over time as new lines are added (or removed) and cut values change. With the increasing number of lines the requirements on the lines become more strict. The time it takes to execute a line (the time required to reconstruct the decay) and the rate (how many events are selected) have upper limits as described earlier. This can be challenging for multi-body decays in particular.

The stripping framework is numbered in a manner similar to the different reconstruction versions. Events rejected during the stripping remain stored on disc. They can still be selected in a new version of the stripping (for example if a new line is added). Stripping version 17 (`Stripping17`) was used for the `Reco12` 2011 data set.

The author has developed two stripping lines:

- The first line, `StrippingBs2MuMuPhi`, was designed to only select $B_s^0 \to \phi\mu^+\mu^-$ events. The selection cuts used for this stripping line are a loosened version of a set of cuts optimised to select $B_s^0 \to \phi\mu^+\mu^-$ decays. The details of the optimised cuts and the optimisation procedure are described in Appendix B.

  The optimum cuts were loosened until official LHCb timing and retention limits for stripping lines were reached. The advantage is that the freedom to adjust the optimum cut values even after the stripping line has been run on the data is maintained. Therefore these cuts are also referred to as offline cuts.

- The second line, `StrippingB2XMuMu`, was designed to select several B-meson decays, which contain two muons in the final state. Amongst others, this includes $B_s^0 \to \phi\mu^+\mu^-$, $B^0 \to \rho\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$. The advantage of selecting several decays with the same set of cuts in one common stripping line is that CPU power is saved. Furthermore, unbiased cross-checks in signal yields or other variables between these decays become possible as there is no bias introduced in the selection. The cuts and cut values for this stripping line are based on a selection for $B^0 \to K^{*0}\mu^+\mu^-$, which had already been developed previously. In order to be able to select multiple B-meson decays a few cuts, such as PID requirements, had to be removed for this purpose.

The data are available to the entire collaboration on discs after the stripping lines have been applied.

Chapter 5

Measurement of $\mathcal{B}(B_s^0 \to \psi'\phi)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

5.1 Introduction

This Chapter describes the measurement of the ratio of branching fractions for the two decays $B_s^0 \to \psi'\phi$ and $B_s^0 \to J/\psi\,\phi$. Measuring branching fractions relative to well measured decays has the advantage that common systematic uncertainties (e.g. on the luminosity) cancel. The work is based on [72], which was presented at the 13th International Conference on B-Physics at Hadron Machines (BEAUTY), Amsterdam, 4-8 April 2011. The data set for this analysis corresponds to a total integrated luminosity of $36\,\mathrm{pb}^{-1}$of pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ recorded by the LHCb detector during 2010. Many of the techniques employed in this Chapter are applicable to the measurement of the differential branching fraction of $B_s^0 \to \phi\mu^+\mu^-$, which is presented in Chapter 6.
Only decay modes where ψ', $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$ have been considered for this analysis. Thus the ratio of branching fractions is given by:

$$\frac{\mathcal{B}(B_s^0 \to \psi'\phi)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = \frac{N_{\psi'\phi}}{N_{J/\psi\,\phi}} \times \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(\psi' \to \mu^+\mu^-)} \times \frac{\varepsilon_{J/\psi\,\phi}}{\varepsilon_{\psi'\phi}} \tag{5.1}$$

where $N_{\psi'\phi}/N_{J/\psi\,\phi}$ is the ratio of the number of signal events and $\varepsilon_{J/\psi\,\phi}/\varepsilon_{\psi'\phi}$ is the ratio of the combined reconstruction, selection and trigger efficiencies. In the following the details of the measurement are given and the final result is presented.

5.2 Event selection

The motivation for the selection requirements outlined in this Section was to develop a simple, cut-based selection consisting of only a handful of variables, which rejects as much background as possible whilst maintaing a high signal efficiency.

Four charged tracks ($K^+K^-\mu^+\mu^-$) are combined to form a B_s^0 candidate, accepting the entire dimuon mass range ($2M_\mu < M_{\mu\mu} < (M_{B_s} - M_\phi)$). Each of the four charged tracks (sometimes referred to as the daughter tracks) is required to be of good quality ($\chi^2_{\text{track}}/n <$ 5 with n degrees of freedom) and to have a transverse momentum with respect to the beam axis, p_{T}, greater than 300 MeV/c. All duplicate tracks are removed by only selecting tracks with a Kullback-Liebler (KL) distance greater than 5000, where the KL distance is a measure of how much information is shared between two tracks [73]. The B_s^0 decay vertex is on average displaced by $\mathcal{O}(10\,\text{mm})$ from the Primary Vertex (PV), where the pp collision takes place in LHCb. Advantage is taken out of the B_s^0 lifetime to reject tracks coming directly from the PV. This is achieved by requiring that the χ^2_{IP} of each track is greater than 9, where the χ^2_{IP} is formed by the hypothesis that the track's Impact Parameter (IP) with respect to the PV is equal to zero. Furthermore, the K^+ and K^- candidates are required to be identified as kaons and the μ^+ and μ^- candidates as muons by the LHCb Particle Identification (PID) system. For the kaons the difference in log-likelihood ($\Delta \log \mathcal{L}$) between the kaon and pion hypotheses based on information from the RICH is required to be greater than 0. This requirement is more than 95% efficient for kaons and typically rejects about 95% of the pions (see Section 4.2.3). The muon PID requirement is also highly efficient for muons as described in Section 4.2.4. The reconstructed dikaon mass, M_{KK}, has to be within ±10 MeV of the nominal ϕ mass [37]. The requirements on the B_s^0 candidate are that its proper time (τ) is greater than 0.2 ps, its vertex (formed by the four daughter tracks) is of good quality ($\chi^2_{\text{VX}} < 40$) and its χ^2_{IP} with respect to the PV is less than 9.

A summary of the selection requirements for this analysis and the efficiencies of each cut on Monte Carlo (MC) signal events are given in Table 5.1. The efficiency shown is defined as:

$$\varepsilon^{\text{cut}} = \frac{N^{\text{selected}}}{N^{\text{selected - cut}}} \tag{5.2}$$

where N^{selected} is the number of events passing all cuts and $N^{\text{selected - cut}}$ is the number of events passing all cuts *excluding* a given cut. The simulation samples used in this study correspond to the 2010 MC (MC10) production of LHCb, in which the detector performance has been tuned to closely resemble the detector performance during the 2010 data taking. Two MC10 samples (one for $B_s^0 \to \phi\mu^+\mu^-$ and the other one for $B_s^0 \to J/\psi\,\phi$ events) have been used. At the time of the study no sample for $B_s^0 \to \psi'(\to \mu^+\mu^-)\phi$ with the 2010 configuration was available. Instead events inside the ψ' region of the MC10 $B_s^0 \to \phi\mu^+\mu^-$ sample were used. In Section 5.4.2 this method is validated by comparing efficiencies obtained for events within the J/ψ region of the $B_s^0 \to \phi\mu^+\mu^-$ sample with efficiencies obtained for events from the $B_s^0 \to J/\psi\,\phi$ sample.

The cuts shown in Table 5.1 have been optimised to reduce background events whilst maintaining a high signal efficiency. The majority of them are applied during the stripping stage (as part of the `StrippingBs2MuMuPhi` stripping line, see Section 4.4.3). The details of the cut optimisation procedure of the stripping line are given in Appendix B.

Table 5.1: Summary of the criteria to select $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ candidates. For each selection requirement the cut value and the selection efficiencies obtained from the MC10 $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \phi\mu^+\mu^-$ samples are given. For the $B_s^0 \to \phi\mu^+\mu^-$ MC sample only candidates inside the region around the nominal ψ' mass were used.

Variable	Cut value	MC10 $B_s^0 \to J/\psi\,\phi$ (%)	MC10 $B_s^0 \to \phi\mu^+\mu^-$ (%)
$K^\pm$, $\mu^\pm$ χ^2_{track}/n	< 5	100.0 ± 0.0	100.0 ± 0.0
$K^\pm$, $\mu^\pm$ p_{T}	$> 300\,\text{MeV}/c$	97.0 ± 0.3	96.3 ± 1.8
$K^\pm$, $\mu^\pm$ KL distance	> 5000	97.2 ± 0.3	96.2 ± 1.7
$K^\pm$, $\mu^\pm$ χ^2_{IP}	> 9	68.9 ± 0.6	65.2 ± 3.8
$K^\pm$ $\Delta \log \mathcal{L}$	> 0	95.2 ± 0.3	97.2 ± 1.6
M_{KK}	$1009 - 1029\,\text{MeV}/c^2$	87.8 ± 0.5	88.0 ± 3.0
B_s^0 τ	$> 0.2\,\text{ps}$	99.9 ± 0.1	100.0 ± 0.0
B_s^0 χ^2_{IP}	< 9	95.2 ± 0.3	94.8 ± 2.5
B_s^0 χ^2_{VX}	< 40	97.7 ± 0.2	96.3 ± 1.8

The mass distribution of all reconstructed B_s^0 candidates in the data set is shown in Figure 5.1. The fit is a single Gaussian for the signal on top of a linear background distribution. For the fit of the Gaussian the mean and sigma have been left floating. The fit converges at a mean (equivalent to the mass in this case) of $M_{B_s} = 5363.0 \pm 0.8\,\text{MeV}/c^2$ with a sigma of $\sigma = 18.0 \pm 0.8\,\text{MeV}/c^2$, giving 820 ± 34 signal events. From this a signal window of $M_{B_s} \in [5300 : 5430]\,\text{MeV}/c^2$, which corresponds to a width of about $\pm 3\sigma$, is defined. A lower and a slightly larger upper background window are defined to be in the range $M_{B_s} \in [5200 : 5300]\,\text{MeV}/c^2$ and $M_{B_s} \in [5430 : 5700]\,\text{MeV}/c^2$ respectively. Furthermore, the mass windows for the J/ψ and ψ' resonances are defined to be $M_{\mu\mu} \in [3040 : 3150]\,\text{MeV}/c^2$ and $M_{\mu\mu} \in [3630 : 3740]\,\text{MeV}/c^2$ respectively, corresponding to widths of about $\pm 3\sigma$. In Figure 5.2 all mass windows are indicated on a scatter plot of the reconstructed dimuon mass versus the reconstructed B_s^0 mass for all selected candidates in data.

5.3 Measurement of $N_{\psi'\phi}/N_{J/\psi\,\phi}$

By selecting candidates from data that lie inside the dimuon mass windows (as defined in the previous Section 5.2) two samples of $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ candidates are obtained. In both samples the mass of the intermediate particle (J/ψ or ψ') is constrained to be exactly its nominal mass by adjusting the 4-vector of the B_s^0 candidate. The num-

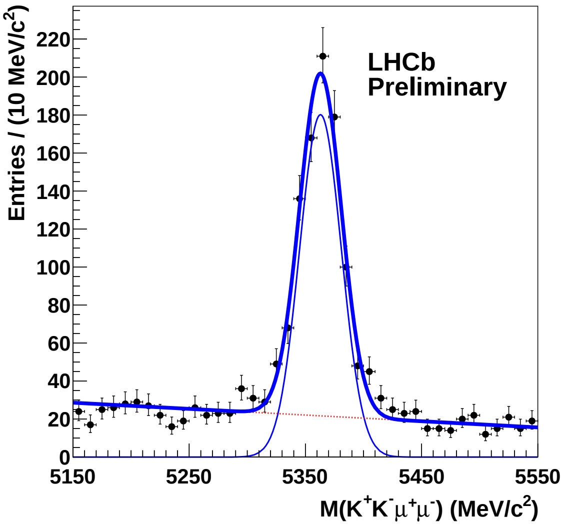


Figure 5.1: Fitted B_s^0 mass distribution for all candidates in the data set. The points represent data, the pointed red line is the linear background component, the thin solid blue line is the signal component of a Gaussian curve and the thick solid blue line is the combined fit to the data.

ber of signal candidates for each mass-constrained sample is obtained by fitting a single Gaussian for the signal on top of a linear background to the mass distribution of the B_s^0 candidates. For the fit to candidates from the $B_s^0 \to J/\psi\,\phi$ sample, the fit parameters of the Gaussian (μ and σ) have been left floating. To exclude resolution effects and biases due to the small number of signal events the fit parameters of the fit to candidates from the $B_s^0 \to \psi'\phi$ sample are fixed. The mean, μ, is fixed to the value obtained from the fit to the $B_s^0 \to J/\psi\,\phi$ sample. The width, σ, is fixed to the value from the fit to the $B_s^0 \to J/\psi\,\phi$ sample multiplied with its ratio obtained from simulations:

$$\mu_{\psi'\phi}^{\text{data}} = \mu_{J/\psi\,\phi}^{\text{data}} \tag{5.3}$$

$$\sigma_{\psi'\phi}^{\text{data}} = \sigma_{J/\psi\,\phi}^{\text{data}} \times \frac{\sigma_{\psi'\phi}^{\text{MC}}}{\sigma_{J/\psi\,\phi}^{\text{MC}}} \tag{5.4}$$

The fitted distributions are given in Figure 5.3. The converged fit parameters of the $B_s^0 \to J/\psi\,\phi$ sample are $M_{B_s}^{J/\psi\,\phi} = 5367.0 \pm 0.3\,\text{MeV}/c^2$ and $\sigma^{J/\psi\,\phi} = 7.5 \pm 0.2\,\text{MeV}/c^2$. The fixed fit parameters used for the $B_s^0 \to \psi'\phi$ sample are $M_{B_s}^{\psi'\phi} = 5367.0\,\text{MeV}/c^2$ and $\sigma^{\psi'\phi} = 5.7\,\text{MeV}/c^2$. The results for the number of signal events are $N_{J/\psi\,\phi} = 770 \pm 28$ and $N_{\psi'\phi} = 54 \pm 8$, giving a ratio of:

$$\frac{N_{\psi'\phi}}{N_{J/\psi\,\phi}} = 0.070 \pm 0.010 \tag{5.5}$$

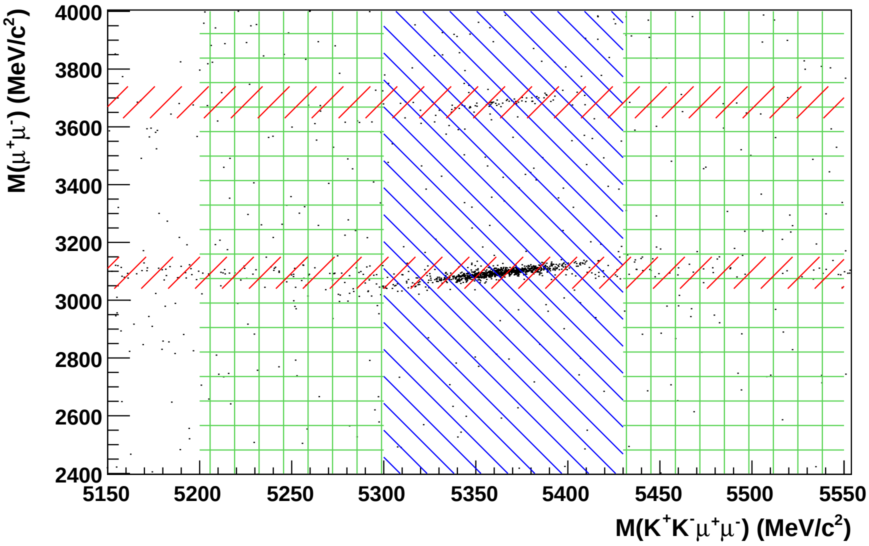
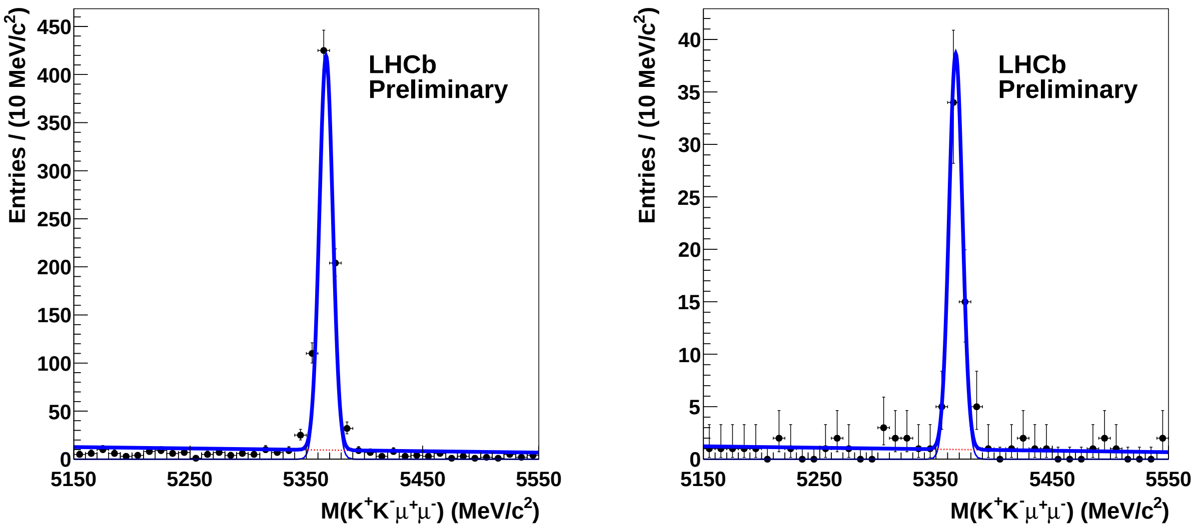


Figure 5.2: Reconstructed dimuon mass as a function of the reconstructed B_s^0 mass for all selected candidates. The signal window corresponds to the top-left to bottom-right hatched area in blue. The lower and upper background windows correspond to the square hatched areas in green. The bottom-left to top-right hatched areas in red represent the windows for the J/ψ and ψ' resonances.

where the uncertainty is statistical.

The contribution from non-resonant $B_s^0 \to \phi\mu^+\mu^-$ decays was found to be negligible. To cross-check the result of Equation 5.5 the same ratio was obtained from fits to the mass distribution of dimuon candidates, $M_{\mu\mu}$. Figure 5.4 shows the $M_{\mu\mu}$ distributions of B_s^0 candidates inside the signal and background windows for regions around the J/ψ and ψ' masses. The candidates in the signal windows correspond to $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ decays. Each fit in Figure 5.4 is a double Gaussian curve on top of a linear background. For the signal windows the fit parameters have been left floating. The values of the converged fit parameters from the signal windows are $M_{J/\psi} = 3095.1 \pm 0.6\,\mathrm{MeV}/c^2$, $\sigma^1_{J/\psi} = 11.1 \pm 0.9\,\mathrm{MeV}/c^2$, $\sigma^2_{J/\psi} = 18.2 \pm 1.1\,\mathrm{MeV}/c^2$, $M_{\psi'} = 3684.4 \pm 2.6\,\mathrm{MeV}/c^2$, $\sigma^1_{\psi'} = 10.7 \pm 2.8\,\mathrm{MeV}/c^2$ and $\sigma^2_{\psi'} = 23.7 \pm 9.7\,\mathrm{MeV}/c^2$. For the fit to candidates inside the background windows the means and widths were fixed to the values of the converged parameters from the signal windows. The results for the number of events inside the signal (background) window are $N_{J/\psi\,\phi} = 790 \pm 29$ (101 ± 13) and $N_{\psi'\phi} = 60 \pm 12$ (8 ± 4). The number of candidates inside each background window is subtracted from the number of candidates inside the corresponding signal window with a weighting of 0.35. The weighting is obtained from the relative sizes of the background and signal windows. The ratio of signal events comes out to be:

$$\frac{N_{\psi'\phi}}{N_{J/\psi\,\phi}} = 0.076 \pm 0.014(\mathrm{stat}) \tag{5.6}$$



(a) (b)

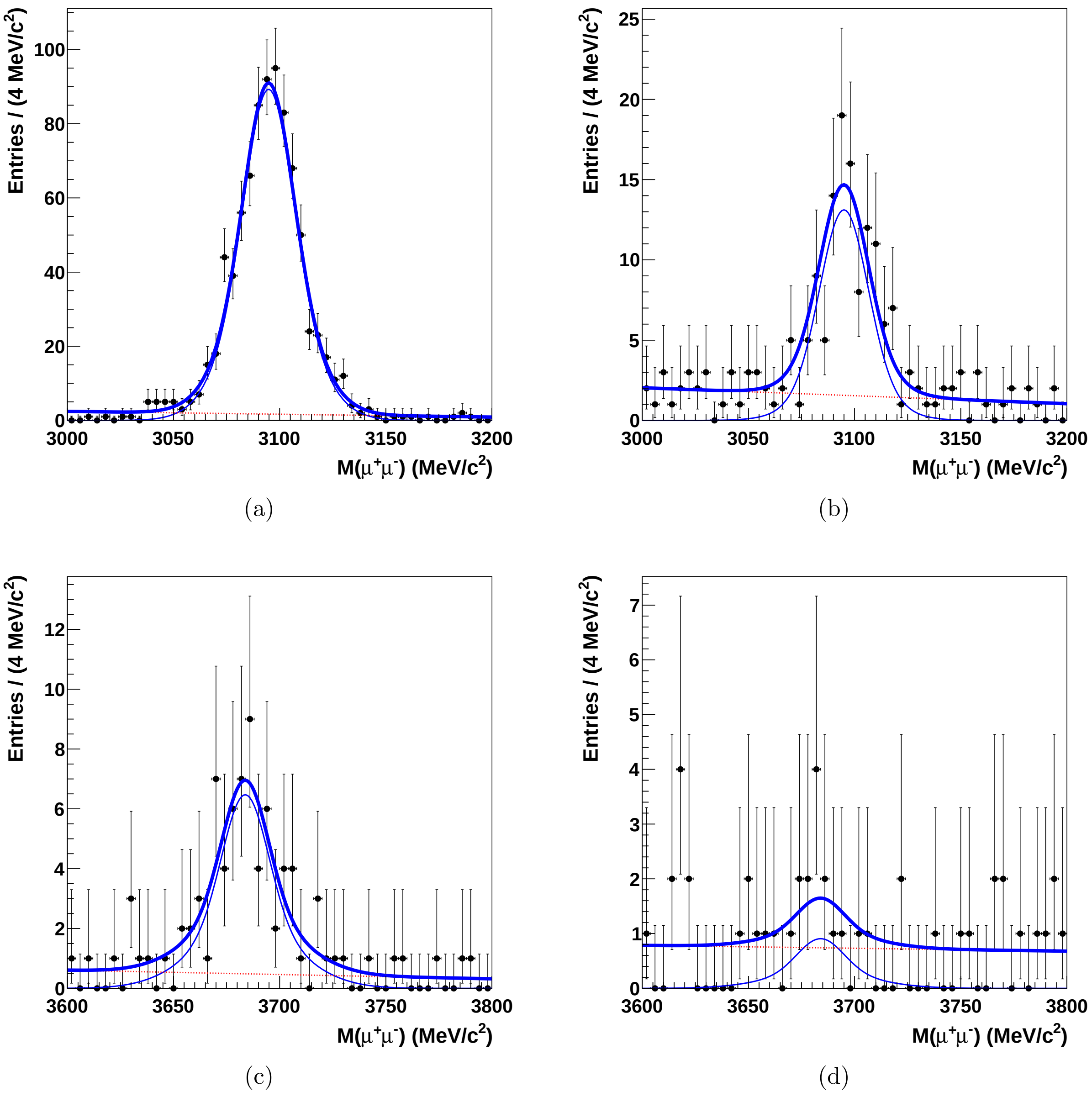
Figure 5.3: Fitted B_s^0 mass distributions for $B_s^0 \to J/\psi\,\phi$ (a) and $B_s^0 \to \psi'\phi$ (b) candidates. The mass of the intermediate J/ψ and ψ' candidates were fixed to their known values [37]. The points represent data. By fitting a Gaussian curve (thin blue line) on top of a linear background (red pointed line) the number of signal candidates for $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ were extracted from the overall fit (thick blue line).

This confirms the value obtained in Equation 5.5. In order to assign a systematic uncertainty to the ratio of number of events the difference to the non-mass constrained case was studied. Further, the systematic uncertainty associated with the signal parametrisation has been studied by fitting a Crystal Ball function instead of a single Gaussian. The Crystal Ball function is given by:

$$f(x;\alpha,n,\bar{x},\sigma) = \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}) & \text{, for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \times (B - \frac{x-\bar{x}}{\sigma})^{-n} & \text{, for } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases} \qquad (5.7)$$

where $A = (n/|\alpha|)^n \times \exp(-|\alpha|^2/2)$, $B = (n/|\alpha|) - |\alpha|$ and $\alpha, n, \bar{x}, \sigma$ are the fit parameters [74, 75]. A summary of the fit results and the corresponding number of signal events and ratios is given in Table 5.2.

The largest difference in central values between the different scenarios given in Table 5.2 is ±0.007 (10.0%), which is assigned as systematic uncertainty to the ratio of the number of signal events.



(a) (b)

(c) (d)

Figure 5.4: The dimuon mass distribution of J/ψ and ψ' candidates inside the signal window ((a) and (c)) and inside the background windows ((b) and (d)). The points represent data, the pointed red line is the linear background component, the thin solid blue line is the signal component of a double Gaussian curve and the thick solid blue line is the combined fit to the data. The fit parameters used for the fits in (b) and (d) were fixed to the converged parameter values of the fits in (a) and (c) respectively.

Table 5.2: Comparison of the fit results for different models and candidates. Column A contains the results of the fit of a single Gaussian to the mass-constrained candidates, B for the fit of a single Gaussian to the non mass-constrained candidates and C for the fit of a Crystal Ball function to the non mass-constrained candidates. Given are σ and the number of signal ($\mathrm{N_S}$) and background ($\mathrm{N_B}$) candidates. In each case $\sigma(\psi'\phi)$ was fixed to the value of $\sigma(J/\psi\,\phi)$ multiplied with the ratio of both σ's as predicted by simulation. For the Crystal Ball function the parameters $\alpha = 1.6$ and $n = 1.0$ were used.

	A	B	C
$\sigma(J/\psi\,\phi)$	7.5 ± 0.2	18.3 ± 0.7	17.8 ± 0.7
$\mathrm{N_S}(J/\psi\,\phi)$	770 ± 28	758 ± 29	773 ± 30
$\mathrm{N_B}$	258 ± 17	273 ± 19	258 ± 20
$\sigma(\psi'\phi)$	5.7	17.8	16.7
$\mathrm{N_S}(\psi'\phi)$	54 ± 8	58 ± 8	59 ± 8
$\mathrm{N_B}$	46 ± 7	41 ± 8	41 ± 7
$\mathrm{N}_{\psi'\phi}/\mathrm{N}_{J/\psi\,\phi}$	0.070 ± 0.010	0.077 ± 0.011	0.076 ± 0.011

5.4 Estimation of $\varepsilon_{J/\psi\,\phi}/\varepsilon_{\psi'\phi}$

The total efficiency is the product of the geometrical acceptance of the LHCb detector ($\varepsilon^{\mathrm{geo}}$), the detection, reconstruction and selection efficiencies ($\varepsilon^{\mathrm{rec\&sel}}$) and the trigger efficiency ($\varepsilon^{\mathrm{trigger}}$). The ratios of these efficiencies between $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ need to be evaluated separately in order to obtain the total efficiency ratio, given by:

$$\frac{\varepsilon_{J/\psi\,\phi}}{\varepsilon_{\psi'\phi}} = \frac{\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{geo}}_{\psi'\phi}} \times \frac{\varepsilon^{\mathrm{rec\&sel}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{rec\&sel}}_{\psi'\phi}} \times \frac{\varepsilon^{\mathrm{trigger}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{trigger}}_{\psi'\phi}} \tag{5.8}$$

The first and second terms are evaluated from simulation. The last term is assumed to be unity, which is validated with a data-driven method. In the following the evaluation of the central value and the uncertainties for each term are presented.

5.4.1 $\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{geo}}_{\psi'\phi}$

The ratio of $\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{geo}}_{\psi'\phi}$ is estimated using simulation, where $\varepsilon^{\mathrm{geo}}$ is the efficiency of all four final state particles of an event to lie within the geometrical acceptance of LHCb. The acceptance region is defined as the volume between a cone of 400 mrad and a cone of 10 mrad, which are both centred around the beam axis. The smaller cone accounts for the material of the beam-pipe. The decays $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ have the same kinematics and therefore the ratio of the acceptances for both is expected to be very close to unity. The results are $\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi} = 0.1603$ and $\varepsilon^{\mathrm{geo}}_{\psi'\phi} = 0.1609$, giving a ratio of:

$$\frac{\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{geo}}_{\psi'\phi}} = 0.996 \pm 0.004 \tag{5.9}$$

where the error is systematic. It is obtained by varying the parameters of the decay model for $B_s^0 \to \psi'\phi$ such that the two extreme polarisation states of the ψ' (CP-even and CP-odd) are considered.

5.4.2 $\varepsilon_{J/\psi\,\phi}^{\mathrm{rec\&sel}}/\varepsilon_{\psi'\phi}^{\mathrm{rec\&sel}}$

The selection and reconstruction efficiency, $\varepsilon_{J/\psi\,\phi}^{\mathrm{rec\&sel}}$, is calculated using the MC10 $B_s^0 \to J/\psi\,\phi$ sample. The result is compared to the outcome of the second MC sample, MC10 $B_s^0 \to \phi\mu^+\mu^-$, when only events inside the dimuon mass range ($M_{\mu\mu}$) around the J/ψ mass ($3040-3150\,\mathrm{MeV}/c^2$) were considered. The agreement is good ($\varepsilon_{J/\psi\,\phi}^{\mathrm{1.MCsample}} = 0.1075 \pm 0.0002$ and $\varepsilon_{J/\psi\,\phi}^{\mathrm{2.MCsample}} = 0.1079 \pm 0.0013$, where the errors are the statistical errors from the size of the MC samples). This cross-check validates using the second MC sample to obtain a value for $\varepsilon_{\psi'\phi}^{\mathrm{rec\&sel}}$ by only considering events, that lie inside the $M_{\mu\mu}$ region of the ψ' ($3630-3740\,\mathrm{MeV}/c^2$). The result is $\varepsilon_{\psi'\phi}^{\mathrm{rec\&sel}} = 0.0976 \pm 0.0010$ (from the MC10 $B_s^0 \to \phi\mu^+\mu^-$ sample). When taking $\varepsilon_{J/\psi\,\phi}^{\mathrm{rec\&sel}} = 0.1075 \pm 0.0002$ (from the MC10 $B_s^0 \to J/\psi\,\phi$ sample for larger statistics) the ratio between the two efficiencies is:

$$\frac{\varepsilon_{J/\psi\,\phi}^{\mathrm{rec\&sel}}}{\varepsilon_{\psi'\phi}^{\mathrm{rec\&sel}}} = 1.086 \pm 0.013 \tag{5.10}$$

where the error is systematic. It is determined by the number of events in the MC samples. The central value is significantly different from unity. The reason is that $\varepsilon^{\mathrm{rec\&sel}}$ varies with $M_{\mu\mu}$ due to one of the applied selection cuts. By tightening each selection cut in turn the cut on $K^\pm$ χ^2_{IP} was identified as the reason for the difference in efficiencies between the two $M_{\mu\mu}$ regions of both decays. Figure 5.5 shows the ratio between $\varepsilon^{\mathrm{rec\&sel}}$ obtained with $\chi^2_{\mathrm{IP}}(K^\pm) > 9$ and $\varepsilon^{\mathrm{rec\&sel}}$ obtained with $\chi^2_{\mathrm{IP}}(K^\pm) > 25$ as a function of $M_{\mu\mu}$. The decrease of the ratio at large $M_{\mu\mu}$ is explained by kinematical arguments. In this region the ϕ has less momentum and less p_{T} than in other regions. Therefore the produced kaons have less p_{T} available and their momentum vector is more likely to point to the primary vertex. This results in smaller values for the IP and causes the ratio in Equation 5.10 to be significantly different from unity.

The MC samples used are not describing the distributions of certain variables observed in data very well. This needs to be accounted for in the form of a systematic uncertainty for the selection and reconstruction efficiency ratio. In particular the IP resolution is worse in data than estimated by simulation. Other quantities relevant to the analysis, like the B_s^0 momentum and p_{T}, have a good agreement between data and simulation. Figure 5.6 shows how the ratio ($\varepsilon_{J/\psi\,\phi}^{\mathrm{rec\&sel}}/\varepsilon_{\psi'\phi}^{\mathrm{rec\&sel}}$) increases with larger cut values on χ^2_{IP} (applied on $K^\pm$ and $\mu^\pm$). The discrepancy between the χ^2_{IP} description in data and in MC is found to be of order 10%, which translates into a systematic uncertainty of ±0.003 (0.28%) on the ratio.

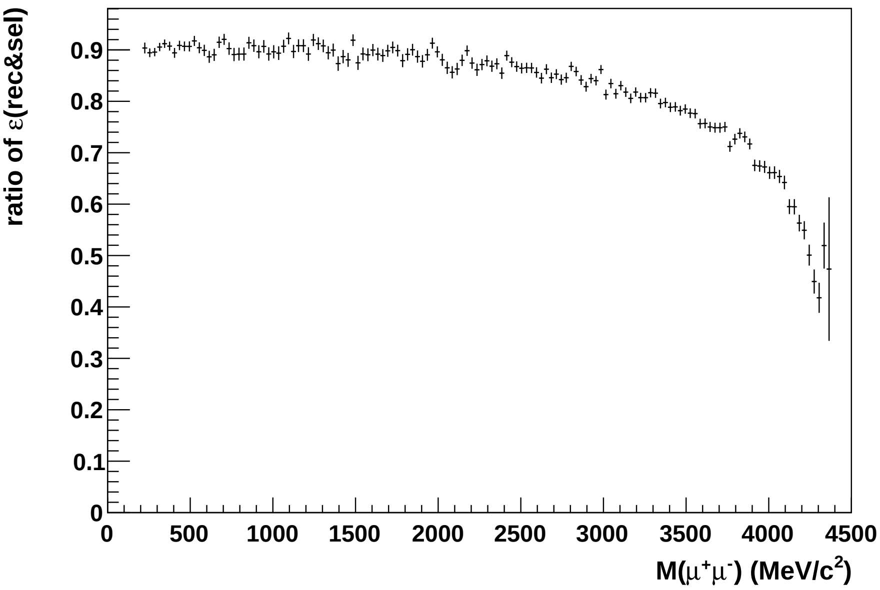


Figure 5.5: Ratio of $\varepsilon^{\text{rec\&sel}}$ obtained with $K^{\pm}$ $\chi^2_{\text{IP}} > 9$ and with $K^{\pm}$ $\chi^2_{\text{IP}} > 25$ across the dimuon mass range. The values were calculated from the $B_s^0 \to \phi\mu^+\mu^-$ MC sample.

The effect of the unknown ψ' polarisation on the $\varepsilon^{\text{rec\&sel}}$ ratio is included by calculating the ratio for the two cases where the polarisation of the ψ' is largest (CP-even and CP-odd). The maximum difference in central values is then assigned as a further systematic uncertainty. CP-even and CP-odd state enhanced regions on the $\cos\theta_K$-$\cos\theta_L$ plane are identified, where θ_K is the angle between the K^+ in the ϕ rest frame and the ϕ in the B_s^0 rest frame and θ_L is the angle between the μ^+ in the ψ' rest frame and the ψ' in the B_s^0 rest frame. In terms of the helicity amplitudes H_0, H_+ and H_- the decay rate is proportional to [76]:

$$\begin{aligned}\frac{d^3\Gamma(B_s^0 \to \psi'(\to \mu^+\mu^-)\phi(\to K^+K^-))}{d\cos\theta_L d\cos\theta_K d\varphi} \propto\ & |H_0|^2 \sin^2\theta_L \cos^2\theta_K \qquad (5.11)\\ & + \frac{1}{4}|H_+|^2 \sin^2\theta_K (1+\cos^2\theta_L)\\ & + \frac{1}{4}|H_-|^2 \sin^2\theta_K (1+\cos^2\theta_L)\\ & - \frac{1}{2}[\Re(H_+H_-^*)\cos 2\varphi - \Im(H_+H_-^*)\sin 2\varphi]\\ & \times \sin^2\theta_L \sin^2\theta_K\\ & + \frac{1}{4}[\Re(H_0(H_+^* + H_-^*))]\cos\varphi \sin 2\theta_L \sin 2\theta_K\\ & + \frac{1}{4}[\Im(H_0(H_+^* - H_-^*))]\sin\varphi \sin 2\theta_L \sin 2\theta_K\end{aligned}$$

where φ is the angle between the two decay planes of the ψ' and the ϕ. When integrating over the entire range of φ the last three terms vanish:

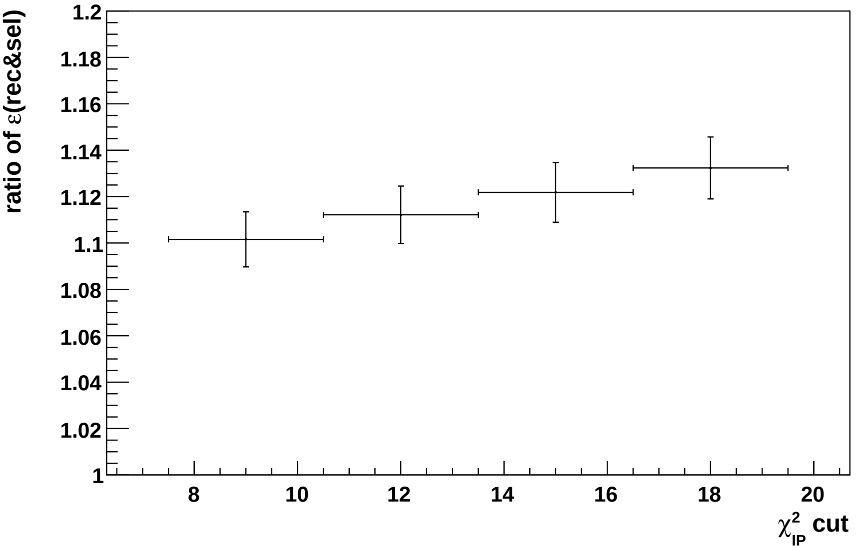


Figure 5.6: Variation of the ratio $\varepsilon^{\text{rec\&sel}}_{J/\psi\phi}/\varepsilon^{\text{rec\&sel}}_{\psi'\phi}$ with the χ^2_{IP} cut on the daughter particles $K^{\pm}$ and $\mu^{\pm}$. The values were obtained from MC simulations. For a given discrepancy between the description of χ^2_{IP} in data and in simulation an uncertainty on the ratio can be extracted.

$$\begin{aligned}\frac{d^2\Gamma(B^0_s \to \psi'(\to \mu^+\mu^-)\phi(\to K^+K^-))}{d\cos\theta_L d\cos\theta_K} \propto\ & |H_0|^2(1-\cos^2\theta_L)\cos^2\theta_K \\ & +\frac{1}{4}|H_+|^2(1-\cos^2\theta_K)(1+\cos^2\theta_L) \\ & +\frac{1}{4}|H_-|^2(1-\cos^2\theta_K)(1+\cos^2\theta_L)\end{aligned} \tag{5.12}$$

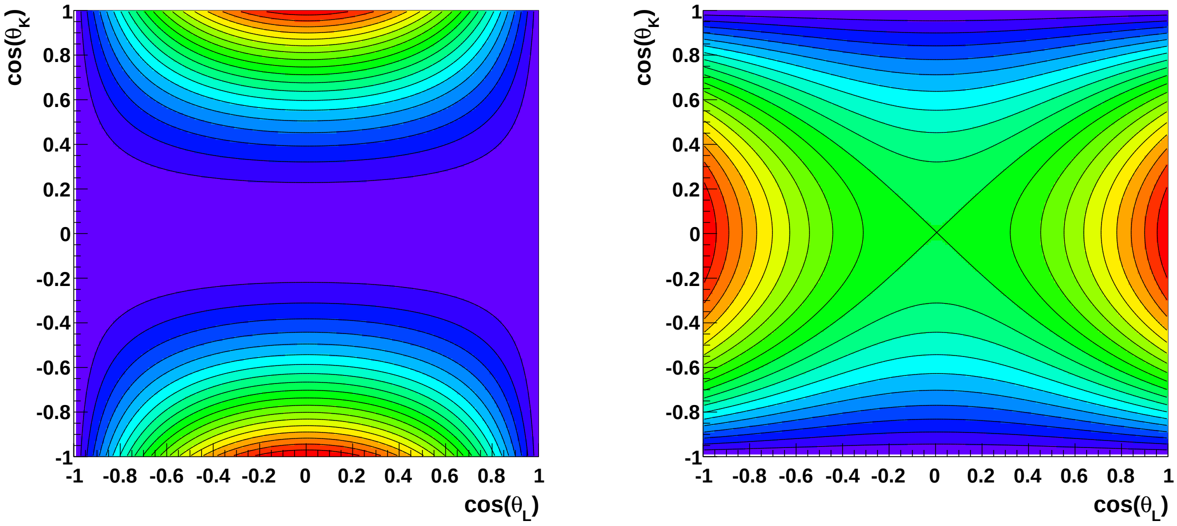
In order to get pure CP-even or pure CP-odd states the helicity amplitudes H_0, H_+ and H_- need to take the values given in Table 5.3.

Table 5.3: Helicity amplitudes for pure CP-even and pure CP-odd states.

	CP-even	CP-odd
H_0	1	0
H_+	0	$+\frac{1}{\sqrt{2}}$
H_-	0	$-\frac{1}{\sqrt{2}}$

The helicity amplitude values for CP-even and CP-odd states are substituted into Equation 5.12 and the results for both cases are represented graphically in Figure 5.7. The distributions in Figure 5.7 clearly show that the regions around $\cos\theta_L = 0$ and $\cos\theta_K = \pm 1$ are dominated by CP-even states whereas the regions around $\cos\theta_L = \pm 1$ and $\cos\theta_K = 0$

are dominated by CP-odd states.



(a) (b)

Figure 5.7: Distribution of pure CP-even (a) and pure CP-odd states (b) in the $\cos\theta_L$-$\cos\theta_K$ plane. Regions with the highest density of states are coloured red.

Calculating $\varepsilon^{\text{rec\&sel}}$ for these regions separately and comparing the resulting ratios to the central value of Equation 5.10 gives rise to an additional systematic uncertainty of ±0.04 (3.68%).

Combining all systematic uncertainties yields a total uncertainty of ±0.042 (3.87%) on the $\varepsilon^{\text{rec\&sel}}$ ratio, which is added to the overall systematic error of the final result.

5.4.3 $\varepsilon^{\text{trigger}}_{J/\psi\,\phi}/\varepsilon^{\text{trigger}}_{\psi'\phi}$

The LHCb trigger system is highly efficient in selecting B-meson decays with two muons in the final state as described in Section 4.3. The trigger efficiencies for $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ decays are expected to be very similar. Therefore the trigger efficiency ratio is assumed to be unity.

The trigger efficiency can be estimated using the TISTOS method. The details of this method are described in Section 4.3.4. To validate the assumption of the ratio being unity the trigger efficiency ratios were estimated from both data and simulation using this technique. For this purpose the data were separated into two subsets by applying mass cuts around the J/ψ and around the ψ' as done in Section 5.3.

The choice of lines for the TIS and TOS categories is based on the lines dominant in selecting candidates for this analysis. Suitable lines for the TOS category are lines that select events containing one or two muons as they are triggered more often than lines that require a certain IP for example. The complete list of the trigger lines, that were

used for the TISTOS method, is given in Appendix C. The lines contributing most are highlighted. For each category (TIS or TOS) the number of signal events is obtained by fitting a Gaussian curve to the reconstructed B_s^0 mass distribution.
The number of TIS events inside the ψ' region is not sufficient for a sensible fit. Instead all events inside a signal box, covering the M_{B_s} range $5290 - 5500\,\mathrm{MeV}/c^2$, are counted. As a cross-check and to avoid a large uncertainty from this procedure, the background events in other regions of $M_{\mu\mu}$ are used as a test. These signal-like background events can give an indication of whether the trigger efficiency is sensible or not. Several regions are defined, for which the trigger efficiencies are determined using the TISTOS method. For these regions the same signal box is used to obtain the number of candidates for each category. Figure 5.8 shows a graphical representation of these regions on the reconstructed dimuon and B_s^0 mass plane.

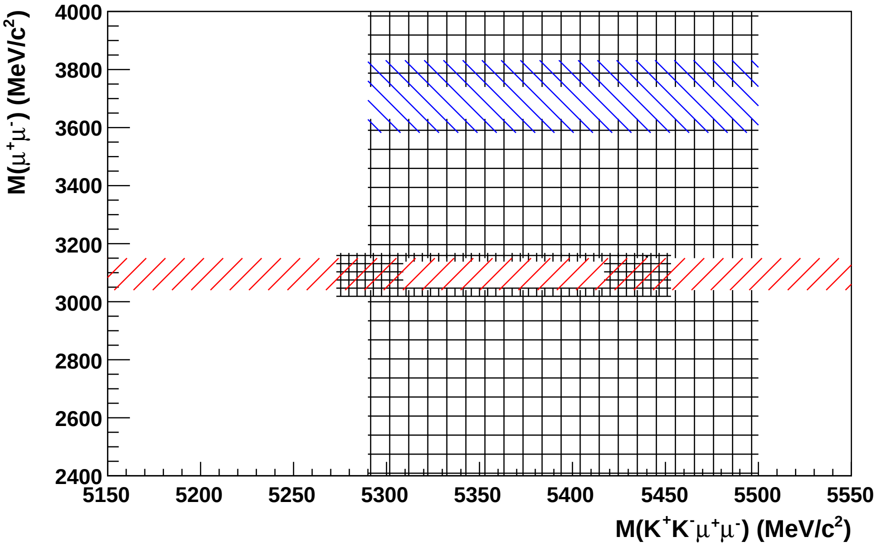


Figure 5.8: A representation of the regions used for the TISTOS method in the B_s^0 mass and dimuon mass plane. All square hatched areas (black) contain background events. The more dense hatched area represents the background region around $B_s^0 \to J/\psi\,\phi$ signal candidates. The number of $B_s^0 \to J/\psi\,\phi$ events is obtained from a fit to all candidates within the bottom-left to top-right hatched area in red. $B_s^0 \to \psi'\phi$ signal candidates are found inside the top-left to bottom-right hatched area in blue. For the blue and black areas a signal box instead of a fit is used to extract the number of signal events.

With the TISTOS method the trigger efficiency ratios for L0, HLT1 and HLT2 can be calculated. The separate efficiencies of HLT1 and HLT2 are conditional. Therefore the efficiencies of the combined triggers (L0×HLT1 and L0×HLT1×HLT2) are a direct product of the separate conditional efficiencies, as shown in Section 4.3.4.
In Figure 5.9 the L0 and HLT1 efficiencies in different dimuon mass regions for both simulation and data are given. For an estimate of the HLT2 efficiency the number of events inside the TIS and TISTOS categories in data was too small. For the entries marked by

red squares ($B_s^0 \to J/\psi\,\phi$ events in data) a fit was used to extract the number of events (for each of the TIS, TOS and TISTOS categories). For the blue ($B_s^0 \to \psi'\phi$ events in data) and black (background events in data) upward-pointing triangles the signal box was used. In addition to the MC TISTOS results (magenta downward-pointing triangles, obtained from fits) the trigger efficiencies calculated using MC truth information (turquoise circles) are also indicated. The MC10 truth information contain simulated trigger decisions of the most common LHCb trigger configuration used during the 2010 data taking period (the precise label in hexadecimal: `0x002e0002a`).

In Figure 5.9(b) the trigger efficiencies from data in regions of small $M_{\mu\mu}$ seem low for HLT1. This is because these events are combinatoric background events and not physics signal events. However, the large overlap of the red square ($B_s^0 \to J/\psi\,\phi$ events in data) with a black upward-pointing triangle (background events) confirms that within the $M_{\mu\mu}$ region of interest it is valid to use background events for the TISTOS method. Table 5.4 summarises the results of all triggers and methods for both data and simulation.

Table 5.4: Comparison of trigger efficiencies and their ratios between data and simulation for a combination of different trigger stages. The TISOS method was used on both data and simulation to obtain the values. For the simulation the efficiencies were also calculated using the MC truth information. The uncertainties are determined by the number of events in the corresponding data sample.

		L0	HLT1	HLT2	L0×HLT1	L0×HLT1×HLT2
data TISTOS	$\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$	0.974 ± 0.053	0.942 ± 0.055		0.900 ± 0.066	
	$\varepsilon_{\psi'\phi}^{\text{trigger}}$	0.968 ± 0.136	0.934 ± 0.160		1.061 ± 0.374	
	ratio	1.007 ± 0.152	1.008 ± 0.183		0.848 ± 0.306	
MC TISTOS	$\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$	0.955 ± 0.006	0.957 ± 0.006	0.995 ± 0.006	0.906 ± 0.008	0.918 ± 0.014
	$\varepsilon_{\psi'\phi}^{\text{trigger}}$	0.964 ± 0.032	0.977 ± 0.034	0.976 ± 0.035	0.903 ± 0.042	0.855 ± 0.082
	ratio	0.991 ± 0.033	0.979 ± 0.035	1.020 ± 0.038	1.004 ± 0.048	1.073 ± 0.104
MC Truth	$\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$	0.927 ± 0.003	0.914 ± 0.004	0.991 ± 0.005	0.847 ± 0.003	0.839 ± 0.003
	$\varepsilon_{\psi'\phi}^{\text{trigger}}$	0.944 ± 0.015	0.931 ± 0.021	0.989 ± 0.023	0.879 ± 0.014	0.870 ± 0.014
	ratio	0.982 ± 0.016	0.981 ± 0.023	1.002 ± 0.023	0.963 ± 0.016	0.965 ± 0.016

The ratio of the total trigger efficiencies is expected to be unity. The efficiency ratios of L0, HLT1 and L0×HLT1 obtained from data using the TISTOS method are given in Table 5.4 and they are all compatible to unity. Both MC approaches (TISTOS and using truth information) support this hypothesis. As mentioned above, the HLT2 efficiency cannot be evaluated with data because the number of events is too small for the TISTOS method to work. However, there is no reason to assume that the HLT2 efficiency is low and that the ratio is not compatible with unity. This is supported by both MC approaches calculating the HLT2 efficiency. The total efficiency ratio (of L0×HLT1×HLT2), as estimated by simulation, is also in agreement with unity.

In conclusion all methods (data-driven and using simulation) are in agreement with the ratio being unity. The slight differences to unity in the central value of the ratios obtained by the MC approaches are averaged and assigned as systematic uncertainty, such that:

$$\frac{\varepsilon^{\text{trigger}}_{J/\psi\,\phi}}{\varepsilon^{\text{trigger}}_{\psi'\phi}} = 1.00 \pm 0.04(\text{syst}) \tag{5.13}$$

5.5 Results

A summary of the results of Section 5.3 and 5.4 and the statistical and systematic uncertainties is given in Table 5.5. Using Equation 5.1 and combining these results together with the PDG values for $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.93 \pm 0.06) \times 10^{-2}$ and $\mathcal{B}(\psi' \to \mu^+\mu^-) = (7.7 \pm 0.8) \times 10^{-3}$ [37] gives:

$$\begin{aligned}\frac{\mathcal{B}(B^0_s \to \psi'\phi)}{\mathcal{B}(B^0_s \to J/\psi\,\phi)} &= \frac{N_{\psi'\phi}}{N_{J/\psi\,\phi}} \times \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(\psi' \to \mu^+\mu^-)} \times \frac{\varepsilon^{\text{geo}}_{J/\psi\,\phi}}{\varepsilon^{\text{geo}}_{\psi'\phi}} \times \frac{\varepsilon^{\text{rec\&sel}}_{J/\psi\,\phi}}{\varepsilon^{\text{rec\&sel}}_{\psi'\phi}} \times \frac{\varepsilon^{\text{trigger}}_{J/\psi\,\phi}}{\varepsilon^{\text{trigger}}_{\psi'\phi}} \\ &= 0.58 \pm 0.10(\text{stat}) \pm 0.07(\text{syst}) \pm 0.06(\mathcal{B}) \end{aligned} \tag{5.14}$$

where the first error is statistical, the second error is systematic and the third error is due to the uncertainty on the branching fractions of the J/ψ and ψ' decays. A breakdown of the total systematic error in the result is given in Table 5.6.

Table 5.5: Summary of all quantities obtained. Where applicable the statistical and systematic uncertainties are given.

Ratio	Value	(stat)	(syst)
$N_{\psi'\phi}/N_{J/\psi\,\phi}$	0.070	±0.010	±0.007
$\varepsilon^{\text{geo}}_{J/\psi\,\phi}/\varepsilon^{\text{geo}}_{\psi'\phi}$	0.996		±0.004
$\varepsilon^{\text{rec\&sel}}_{J/\psi\,\phi}/\varepsilon^{\text{rec\&sel}}_{\psi'\phi}$	1.086		±0.042
$\varepsilon^{\text{trigger}}_{J/\psi\,\phi}/\varepsilon^{\text{trigger}}_{\psi'\phi}$	1.00		±0.04

5.6 Conclusions

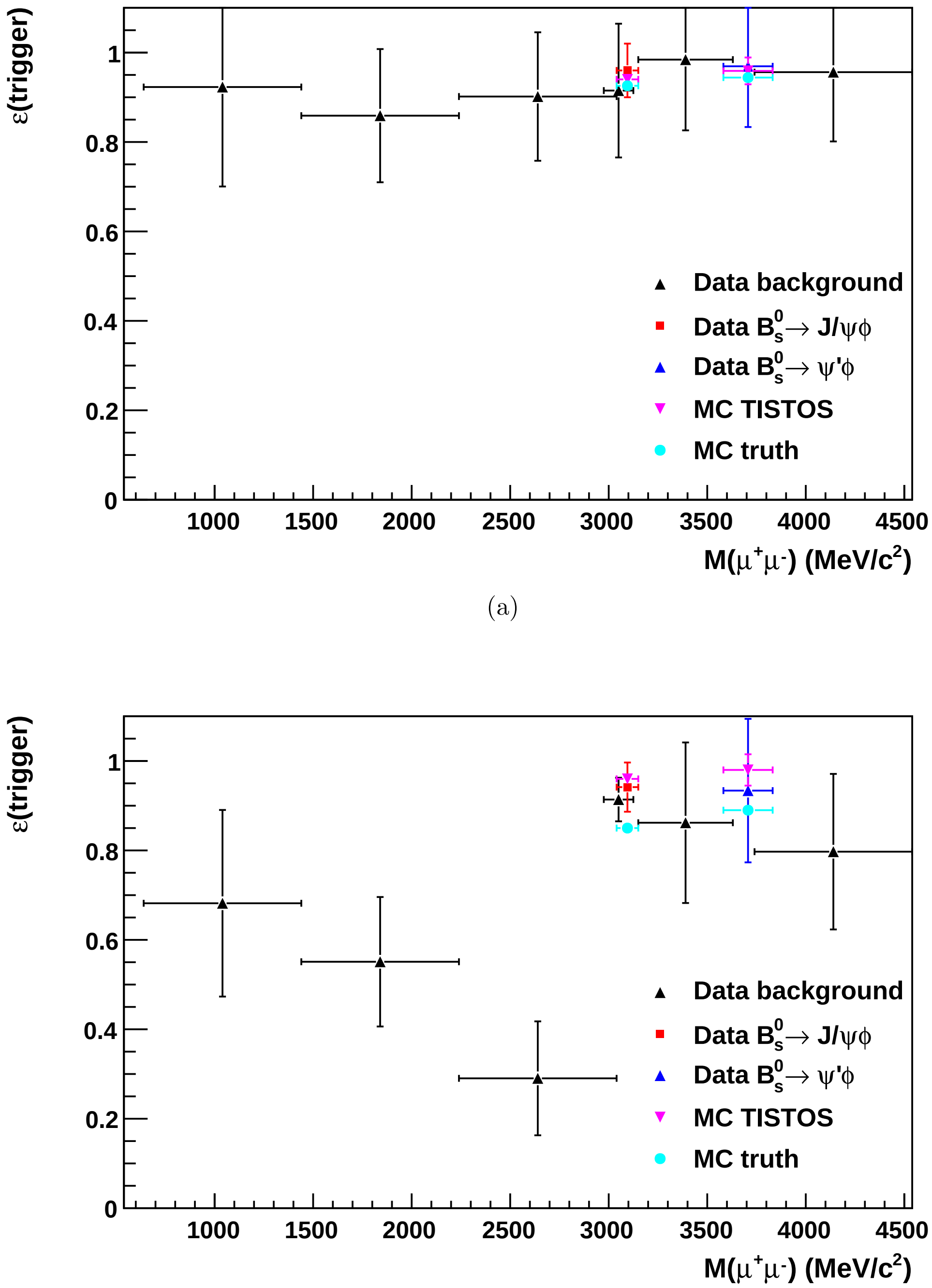
The result of the ratio of branching fractions for $B^0_s \to \psi'\phi$ and $B^0_s \to J/\psi\,\phi$ decays is compatible with measurements from the CDF and D0 experiments, which are given together with the ratios of similar decays for comparison in Table 5.7.

Table 5.6: The relative contribution of each term to the total systematic error of the final result.

Term	syst
$N_{\psi'\phi}/N_{J/\psi\,\phi}$	0.08
$\epsilon^{\text{geo}}_{J/\psi\,\phi}/\epsilon^{\text{geo}}_{\psi'\phi}$	0.01
$\epsilon^{\text{rec\&sel}}_{J/\psi\,\phi}/\epsilon^{\text{rec\&sel}}_{\psi'\phi}$	0.01
$\epsilon^{\text{trigger}}_{J/\psi\,\phi}/\epsilon^{\text{trigger}}_{\psi'\phi}$	0.03
Total	0.09

Table 5.7: Various branching fraction ratios from BABAR, CDF, D0 and the PDG as comparison.

Mode	Ratio	
$\frac{\mathcal{B}(B_s^0\to\psi'\phi)}{\mathcal{B}(B_s^0\to J/\psi\,\phi)}$ (LHCb)	$0.58 \pm 0.10(\text{stat}) \pm 0.07(\text{syst}) \pm 0.06(\mathcal{B})$	
$\frac{\mathcal{B}(B_s^0\to\psi'\phi)}{\mathcal{B}(B_s^0\to J/\psi\,\phi)}$ (CDF)	$0.52 \pm 0.13(\text{stat}) \pm 0.04(\text{syst}) \pm 0.06(\mathcal{B})$	[77]
$\frac{\mathcal{B}(B_s^0\to\psi'\phi)}{\mathcal{B}(B_s^0\to J/\psi\,\phi)}$ (D0)	$0.53 \pm 0.10(\text{stat}) \pm 0.07(\text{syst}) \pm 0.06(\mathcal{B})$	[78]
$\frac{\mathcal{B}(B^0\to\psi' K^0)}{\mathcal{B}(B^0\to J/\psi\,K^0)}$	$0.82 \pm 0.12(\text{stat}) \pm 0.13(\text{syst})$	[79]
$\frac{\mathcal{B}(B^+\to\psi' K^{*+})}{\mathcal{B}(B^+\to J/\psi\,K^{*+})}$	0.47 ± 0.10	[37]
$\frac{\mathcal{B}(B^0\to\psi' K^{*0})}{\mathcal{B}(B^0\to J/\psi\,K^{*0})}$	0.46 ± 0.04	[37]
$\frac{\mathcal{B}(B^+\to\psi' K^+)}{\mathcal{B}(B^+\to J/\psi\,K^+)}$	$0.63 \pm 0.05(\text{stat}) \pm 0.08(\text{syst})$	[78]



(a)

(b)

Figure 5.9: The trigger efficiency as a function of the dimuon mass obtained from data and simulation using the TISTOS technique. Given are the distributions for the L0 trigger (a) and the HLT1 trigger (b). The values of the points marked with an upward-pointing triangle (black for background and blue for $B_s^0 \to \psi'\phi$ candidates) were obtained from data using a signal box. The values of the points marked by a red square were obtained from data through a fit to the $B_s^0 \to J/\psi\,\phi$ candidates. The violet downward-pointing triangles represent the values obtained from simulation using the TISTOS technique. Turquoise circles represent the efficiencies obtained from simulation using the MC truth information.

Chapter 6

Measurement of $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

6.1 Introduction

This chapter describes the measurement of the ratio of branching fractions for the two decays $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$. The work is based on [80] and was presented first at the XLVIIth Recontres de Moriond session devoted to electroweak interactions and unified theories, La Thuile, 3-10 March 2012.

The theoretical prediction for $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)$ is 1.61×10^{-6} [42]. The theoretical background of $B_s^0 \to \phi\mu^+\mu^-$ and similar decays is described in Chapter 3. The branching fraction has been measured previously by the CDF experiment [48], with a limited statistical uncertainty.

Only decay modes where $\phi \to K^+K^-$ and $J/\psi \to \mu^+\mu^-$ are considered, such that the ratio of branching fractions is given by the formula:

$$\frac{\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} = \frac{N_{\phi\mu\mu}}{N_{J/\psi\,\phi}} \times \mathcal{B}(J/\psi \to \mu^+\mu^-) \times \frac{\varepsilon_{J/\psi\,\phi}}{\varepsilon_{\phi\mu\mu}} \tag{6.1}$$

where $N_{\phi\mu\mu}/N_{J/\psi\,\phi}$ is the ratio of the number of signal events and $\varepsilon_{J/\psi\,\phi}/\varepsilon_{\phi\mu\mu}$ is the ratio of the combined reconstruction, selection and trigger efficiencies. Many of the techniques employed in this chapter to calculate these terms have been developed as part of the analysis presented in Chapter 5. The branching fraction $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)$ is defined with exclusion of the J/ψ and ψ' resonances.

6.2 Data sets

The data set used in this analysis was recorded by the LHCb detector during the 2011 data taking period and corresponds to an integrated luminosity of $1\,\text{fb}^{-1}$of pp collisions at $\sqrt{s} = 7\,\text{TeV}$. A detailed description of the LHCb detector and its relevant components for this measurement is given in Chapter 4. The data have been recorded with different magnet polarities and slightly varying trigger conditions. The same set of reconstruction algorithms (version 12) and stripping selections (version 17) have been used when all data collected during 2011 were reprocessed. The simulation samples (MC11) used in this analysis have been tuned to resemble closely the detector performance during the 2011 data taking period. Samples for $B_s^0 \to \phi\mu^+\mu^-$, $B_s^0 \to J/\psi\,\phi$, $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}J/\psi$ were used. Sets for the last two decay modes were required to perform several cross-checks, which are outlined later in Section 6.4.4. In Table 6.1 a summary of the properties of the data and simulation samples used is given.

Each MC11 sample was generated with a specific underlying decay model. Several decay models, based on different theory calculations of certain parameters, exist. The default model is based on a model called `BTOSALI` and is in agreement with the SM [81]. This model uses the transition form factors for $B^0 \to K^*$, calculated from the QCD sum rule. For $B_s^0 \to \phi\mu^+\mu^-$ the B^0 mass was then subsequently changed to the B_s^0 mass and the K^* mass to the ϕ mass. This is valid as the meson masses as well as the kinematic distributions are approximately equal. The form factors for $B^0 \to K^*$ and $B_s^0 \to \phi$ are expected to be very similar. However, it is known that in the QCD sum rule the form factors are softer than in the quark model, which can cause slightly different angular distributions. For this study two non-SM decay models have been used to generate further MC11 samples. These samples were used to calculate systematic uncertainties due to the underlying decay model on the measurement. These models are different to the default SM decay file in the signs and values of the Wilson coefficients C_7 and C_{10}. They are based on a model called `BTOSLLMSEXT`, for which the theoretical background is described in [81–84]. The predicted results by these models are consistent with existing measurements.

6.3 Event selection

The motivation for the selection requirements outlined in this Section is the same as for the selection in the previous Chapter. The aim was to develop a simple, cut-based selection consisting of only a handful of variables, that have a high efficiency in selecting the few $B_s^0 \to \phi\mu^+\mu^-$ decays expected to be found in data.
All events passing the `B2XMuMu` stripping line, which was introduced in Section 4.4.3, were

Table 6.1: Overview of the MC (a) and the data (b) samples used. The reconstruction and stripping versions are given as well as the polarity of the dipole magnet and the size of each sample.

(a)

MC11a	Reco. Version	Magnet	Events
$B_s^0 \to \phi\mu^+\mu^-$	12	Up	~250k
$B_s^0 \to \phi\mu^+\mu^-$	12	Down	~250k
$B_s^0 \to J/\psi\,\phi$	12	Up	~80k
$B_s^0 \to J/\psi\,\phi$	12	Down	~80k
$B^0 \to K^{*0}\mu^+\mu^-$	12	Up	~500k
$B^0 \to K^{*0}\mu^+\mu^-$	12	Down	~500k
$B^0 \to K^{*0}J/\psi$	12	Up	~500k
$B^0 \to K^{*0}J/\psi$	12	Down	~500k

(b)

Set	Reco. Version	Stripping Version	Magnet	Luminosity
1	12	17	Down	$589.2\,\text{pb}^{-1}$
2	12	17	Up	$434.9\,\text{pb}^{-1}$

considered. The `B2XMuMu` stripping line is based on an existing set of cuts, that was previously developed to select $B^0 \to K^{*0}\mu^+\mu^-$ events. The reason for using this line over the designated `Bs2MuMuPhi` stripping line (described in Section 4.4.3) is that it also selects $B^0 \to K^{*0}\mu^+\mu^-$ events. By selecting both $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ events in an unbiased manner (with the same requirements) uncertainties due to different selections are removed. $B^0 \to K^{*0}\mu^+\mu^-$ events enable several cross-checks to validate the measurement, which are explained later in Section 6.4.4.

To further remove background events another set of cuts is applied to all candidates. These cuts are applied after the stripping and are therefore referred to as offline cuts. The offline cuts tighten requirements on variables similar to the ones cut on during the stripping stage already, and have been optimised in order to select $B_s^0 \to \phi\mu^+\mu^-$ events. The details of the optimisation are given in Appendix B.

During the stripping stage a dimuon candidate is formed combining two oppositely charged muon tracks. The kinematic allowed mass range in terms of the dimuon invariant mass squared (q^2) is $4M_\mu^2 < q^2 < 19.3\,\text{GeV}^2/c^4$. Likewise, the ϕ candidate is formed by combining two oppositely charged kaon tracks in the event. Both the dimuon and ϕ candidates are consequently combined to form a B_s^0 candidate.

The two muons and two kaons are also referred to as the daughter particles. A good track quality ($\chi^2_{\text{track}}/n < 4$ with n degrees of freedom) is required for all daughter tracks. The transverse momentum, p_T, with respect to the beam axis of each daughter candidate has to be greater than $250\,\text{MeV}/c$. In LHCb the B_s^0 candidate decay vertex is on average displaced by $\mathcal{O}(10\,\mathrm{mm})$ from the PV. Advantage is taken from the B_s^0 lifetime to reject tracks coming directly from the PV and therefore to reduce background events. This is achieved by requiring that the χ^2_{IP} of kaon track is greater than 9, where the χ^2_{IP} is formed by the hypothesis that the track's IP with respect to the PV is equal to zero. The χ^2_{IP} for each muon track (μ^+ and μ^-) is required to be greater than 16. Furthermore, the μ^+ and μ^- have to be identified as muons by the LHCb particle identification system. For the μ^+ and μ^- the difference in log-likelihood ($\Delta\log\mathcal{L}$) between the muon and pion hypotheses has to be greater than 0.

The requirements on the B_s^0 candidate are that its χ^2_{IP} with respect to the PV is less than 9, its decay vertex is of good quality ($\chi^2_{\mathrm{VX}} < 40$) and the angle between the B_s^0 flight direction (defined by the vector between the PV and the B_s^0 decay vertex) and the total momentum vector of the daughters, θ, is less than 8 mrad. A summary of the stripping selection requirements for this analysis can be found in Table 6.2.

The offline cuts are tighter than the stripping line cuts and put requirements on effectively the same quantities (instead of the cut on θ a requirement on the B_s^0 proper decay time is introduced). It is ensured that all duplicate tracks are removed by only selecting tracks with a Kullback-Liebler (KL) distance greater than 5000, where the KL distance is a measure of how much information is shared between two tracks [73]. The χ^2_{IP} of all daughter candidates is now required to be greater than 16. The dikaon mass, M_{KK}, is required to be within ±10 MeV of the nominal ϕ mass, which corresponds to a $\sim 5\sigma$ width [37]. The B_s^0 candidate proper decay time, τ, has to be greater than 0.4 ps, χ^2_{IP} smaller than 6 and χ^2_{VX} smaller than 15.

For the kaons the difference in log-likelihood between the kaon and pion hypotheses is now required to be greater than 5. For cut values of this order the $\Delta\log\mathcal{L}$ distributions in simulation are known not to represent the distributions in data very closely. The reason is that for example the RICH performance is very difficult to model precisely in this region. Therefore, instead of applying $\Delta\log\mathcal{L}$ cuts on the simulation samples directly, efficiency tables (obtained from data) are used for these cuts. These tables return the weight of an event based on its p_{T} and η for a given $\Delta\log\mathcal{L}$ cut, where η is the pseudorapidity. The uncertainty on the $\Delta\log\mathcal{L}$ cut due to different running conditions in data is negligible

The offline cuts and their cut values are summarised in Table 6.3. The efficiency given for each cut in Table 6.3 is defined as:

$$\varepsilon^{\text{offline cut}} = \frac{N^{\text{offline selected}}}{N^{\text{offline selected - cut}}} \tag{6.2}$$

where $N^{\text{offline selected}}$ is the number of events passing all offline cuts and $N^{\text{offline selected - cut}}$ is the number of events passing all offline cuts *excluding* the given cut. When selecting $B^0 \to K^{*0}\mu^+\mu^-$ events with these offline cuts the requirements have to be slightly adjusted. The dikaon mass, M_{KK}, is changed to be within ±100 MeV of the nominal K^{*0} mass and the

PID requirement on the K^- candidate is removed in order to be able to select π^-'s.

Table 6.2: Variables used in the stripping line to select $B^0_s \to \phi\mu^+\mu^-$ and $B^0_s \to J/\psi\,\phi$ events.

Variable	Cut value
$K^\pm$, $\mu^\pm$ χ^2_{track}/n	< 4
$K^\pm$, $\mu^\pm$ p_{T}	$> 250\,\text{MeV}/c$
$K^\pm$ χ^2_{IP}	> 9
$\mu^\pm$ χ^2_{IP}	> 16
$\mu^\pm$ $\Delta\log\mathcal{L}$	> 0
B^0_s χ^2_{IP}	< 9
B^0_s χ^2_{VX}	< 40
B^0_s θ	$< 8\,\text{mrad}$

Table 6.3: The offline cuts, which are tighter in comparison to the stripping cuts, and their efficiencies. The efficiencies are obtained with respect to the number of events after the stripping cuts.

Variable	Cut value	MC11 $B^0_s \to J/\psi\,\phi$ (%)	MC11 $B^0_s \to \phi\mu^+\mu^-$ (%)
$K^\pm$, $\mu^\pm$ KL distance	> 5000	100.0 ± 0.0	100.0 ± 0.0
$K^\pm$, $\mu^\pm$ χ^2_{IP}	> 16	94.2 ± 0.2	93.5 ± 0.2
$K^\pm$ $\Delta\log\mathcal{L}$	> 5	92.3 ± 0.3	91.2 ± 0.2
M_{KK}	$1009 - 1029\,\text{MeV}/c^2$	86.7 ± 0.3	87.6 ± 0.3
B^0_s τ	$> 0.4\,\text{ps}$	99.8 ± 0.1	99.6 ± 0.1
B^0_s χ^2_{IP}	< 6	94.4 ± 0.2	94.5 ± 0.2
B^0_s χ^2_{VX}	< 15	94.5 ± 0.3	94.3 ± 0.2
Total		59.4 ± 0.4	58.1 ± 0.3

6.4 Measurement of $N_{\phi\mu\mu}/N_{J/\psi\,\phi}$

The mass windows for the $J/\psi \to \mu^+\mu^-$ and $\psi' \to \mu^+\mu^-$ resonances are defined to be $2900 - 3176\,\text{MeV}/c^2$ and $3586 - 3766\,\text{MeV}/c^2$, which correspond to widths of about $\pm 4\sigma$. Radiative photons off the muons coming from the J/ψ and ψ' can cause the reconstructed B^0_s mass in $B^0_s \to J/\psi\,\phi$ and $B^0_s \to \psi'\phi$ decays to be shifted to the lower end of the mass spectrum. To account for this effect the lower edges of the veto regions for the resonances have to be extended to $M_{\mu\mu} > 2700\,\text{MeV}/c^2$ and $> 3400\,\text{MeV}/c^2$ when the reconstructed mass of the B^0_s candidate is less than $5250\,\text{MeV}/c^2$. The signal region for $B^0_s \to \phi\mu^+\mu^-$ events corresponds to the entire q^2 range excluding the veto regions for the resonant signals described above. The q^2 range is also divided into 6 separate bins, for which the yields are obtained. The choice of binning is consistent with previous experiments [48]. An additional q^2 bin between $1\,\mathrm{GeV}^2/c^4 < q^2 < 6\,\mathrm{GeV}^2/c^4$ is defined to compare with theoretical predictions. The signal region for $B_s^0 \to J/\psi\,\phi$ events corresponds to the J/ψ veto window.

A two-dimensional graph of the reconstructed dimuon versus the reconstructed B_s^0 mass for all candidates in data, that passed the offline selection criteria described in Section 6.3, is presented in Figure 6.1. The veto regions are also indicated. There is no contamination from radiative $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ decays as the J/ψ and the ψ' veto regions have been extended to cover these candidates as described above.

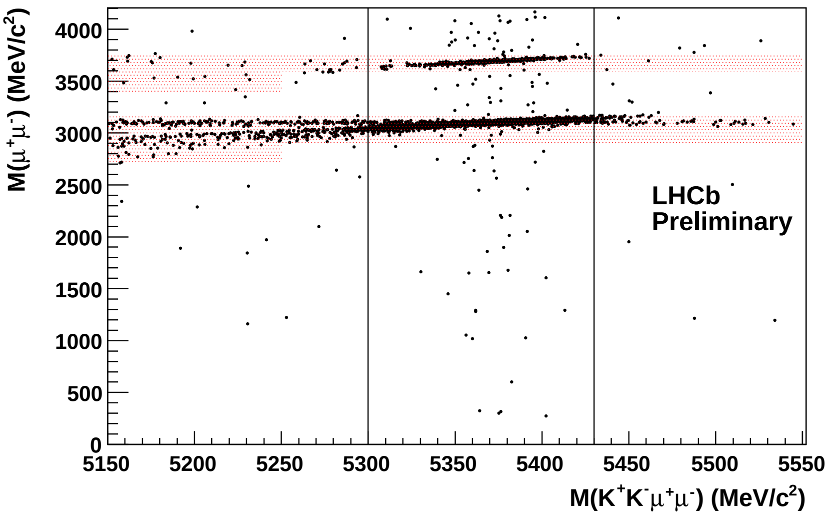


Figure 6.1: Reconstructed dimuon mass $M(\mu^+\mu^-)$ as a function of the reconstructed B_s^0 mass $M(K^+K^-\mu^+\mu^-)$ of candidates that passed all selection criteria. The two shaded regions in red are the two veto regions for the J/ψ and ψ' resonance modes. The veto regions are asymmetric and are extended in the lower B_s^0 mass band in order to account for radiative $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ events (distributed parallel to the diagonal). The B_s^0 signal band is between the two vertical lines and has a width of about $\pm 3\sigma$ around the known B_s^0 mass. The structure inside the veto region in the lower background band is where the distribution of radiative $B_s^0 \to J/\psi\,\phi$ events diverges clearly from the distribution of genuine $J/\psi \to \mu^+\mu^-$ decays, which (combined with two kaons) pass the selection.

6.4.1 Exclusive backgrounds for $B_s^0 \to \phi\mu^+\mu^-$

The level of peaking background sources other than combinatoric background needs to be estimated. In this section the following sources of peaking backgrounds are studied:

- Candidates from $B_s^0 \to J/\psi\,(\to \mu^+\mu^-)\phi(\to K^+K^-)$ decays, where a double misidentification took place. When one of the muons is misidentified as a kaon and one of the kaons as a muon, this channel can contribute to the number of observed

$B_s^0 \to \phi\mu^+\mu^-$ signal candidates. The invariant mass of all daughter candidates from this source can lie within the B_s^0 mass window.

- Candidates from $B^0 \to K^{*0}(\to K^+\pi^-)\mu^+\mu^-$ decays with the pion misidentified as a kaon.
- Candidates from $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$ decays with the proton misidentified as a kaon.
- Candidates from $B^+ \to K^+\mu^+\mu^-$ decays with an additional kaon from the event.

All of these background sources have been checked for in data. In each case the mass of one of the daughter candidates was changed and fixed to the mass of the particle it was misidentified as. The total number of reconstructed candidates inside the kinematic signal region, after the misidentification was accounted for, was obtained. If any candidates are found within this region, this source can be a peaking background and the number of candidates of this sort has to be estimated using simulation. In the following the details of each potential background listed above are given:

Peaking background from $B_s^0 \to J/\psi\,\phi$

To test the contribution from double misidentified $B_s^0 \to J/\psi\,\phi$ decays the masses of the daughters in data have been swapped ($\mu^\pm \to K^\pm$ and $K^\pm \to \mu^\pm$). Any appearances of J/ψ and ϕ resonances would indicate a contribution of this sort of background. In Figure 6.2 the distribution of the reconstructed dimuon versus the reconstructed dikaon mass is given (after the muon and kaon masses have been switched to take into account a double misidentification). The kinematic allowed region (the rectangle formed out of the J/ψ and ϕ resonance bands) is indicated. No events are found inside this region. Therefore the contribution to the number of $B_s^0 \to \phi\mu^+\mu^-$ signal candidates from this background is negligible.

Peaking background from $B^0 \to K^{*0}(\to K^+\pi^-)\mu^+\mu^-$

A similar test is performed to get an estimate of the background contribution from the decay $B^0 \to K^{*0}(\to K^+\pi^-)\mu^+\mu^-$. The full $1\,\text{fb}^{-1}$data set was used to test for its contribution. There are no events inside the kinematic allowed area, given in Figure 6.3, after changing the mass of a pion to the mass of a kaon. Therefore this source of background is also negligible.

Peaking background from $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$

In contrast to the previous two channels there are some misidentified candidates inside the kinematic allowed region coming from the decay $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$, as shown

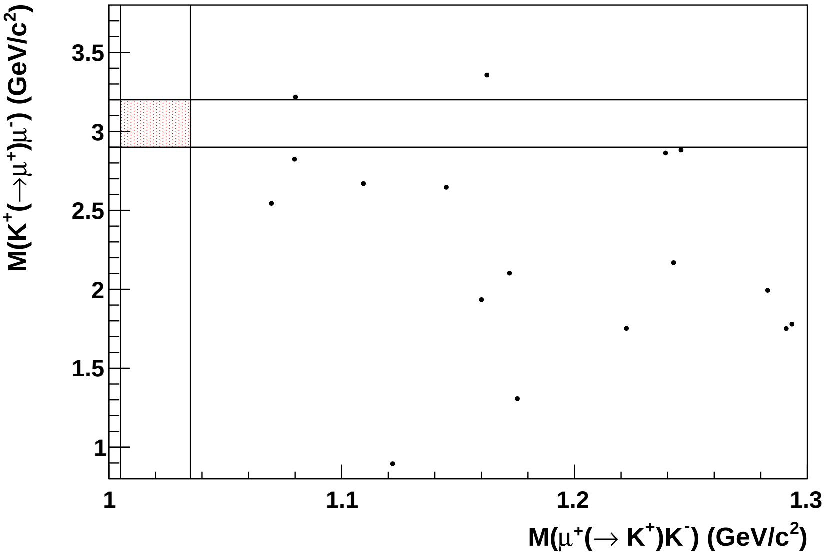


Figure 6.2: Kinematic resolved area for misidentified candidates from $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ decays. If one of the muons is misidentified as a kaon and one of the kaons is misidentified as a muon $B_s^0 \to J/\psi\,\phi$ events can form a peaking background. Switching the masses of the daughter candidates in data accordingly and applying the offline cuts gives the above entries. There are no candidates found inside the kinematic allowed region (shaded red), indicating that this kind of background is negligible.

in Figure 6.4. Therefore the number of candidates from the decay $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$, that are within the allowed region, needs to be estimated using simulation. For this purpose a $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$ sample with 6.1×10^6 MC-true events was used. After reconstruction and selection 6894 MC-true events were left (the mass of the proton was changed to the mass of a kaon before the selection cuts were applied), of which only 1510 events remained with an invariant mass close to the mass of the B_s^0-meson. Therefore the reconstruction and selection efficiency for this channel is:

$$\varepsilon^{\text{rec\&sel}}_{\Lambda_b\to\Lambda(1520)\mu^+\mu^-} = 0.0247 \pm 0.0006\% \tag{6.3}$$

The ratio of the number of $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$ and the number $B_s^0 \to \phi\mu^+\mu^-$ events is estimated by:

$$\frac{N_{\Lambda_b\to\Lambda(1520)\mu^+\mu^-}}{N_{B_s^0\to\phi\mu^+\mu^-}} = \frac{\sigma_{b\to\Lambda_b} \times \mathcal{B}(\Lambda_b \to \Lambda(1520)\mu^+\mu^-) \times \varepsilon_{\Lambda_b\to\Lambda(1520)\mu^+\mu^-} \times \varepsilon_{\text{misID}}}{\sigma_{b\to B_s^0} \times \mathcal{B}(B_s^0 \to \phi\mu^+\mu^-) \times \varepsilon_{B_s^0\to\phi\mu^+\mu^-}} \tag{6.4}$$

Assuming that the geometrical acceptance and the trigger efficiency are similar for both decays (and therefore cancel in the ratio) gives:

$$\frac{N_{\Lambda_b\to\Lambda(1520)\mu^+\mu^-}}{N_{B_s^0\to\phi\mu^+\mu^-}} = 3^{+3}_{-2} \times 10^{-4} \tag{6.5}$$

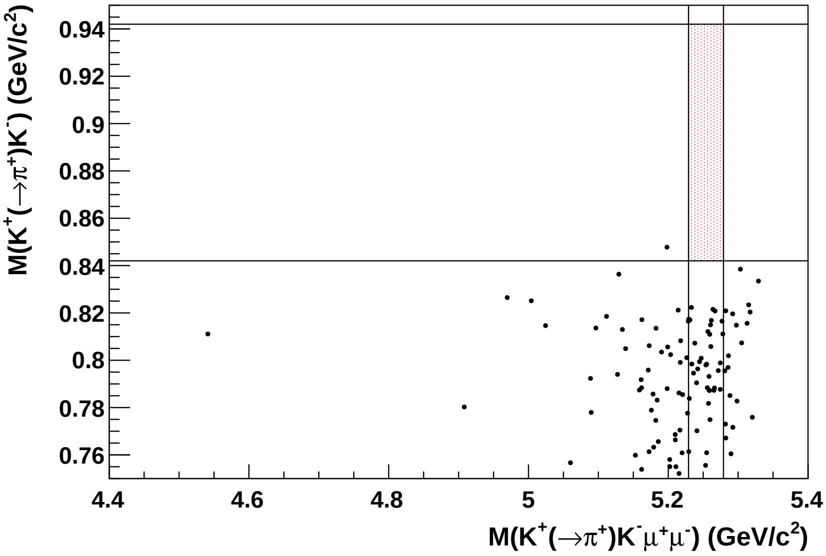


Figure 6.3: Kinematic resolved area for misidentified candidates from $B^0 \to K^{*0}(\to K^+\pi^-)\mu^+\mu^-$ decays. If the pion is misidentified as a kaon $B^0 \to K^{*0}\mu^+\mu^-$ events can be a source of background. Switching the masses of the kaon candidates in data accordingly and applying the offline cuts gives the above entries. There are no candidates found inside the kinematic allowed region (shaded red), indicating that this kind of background is negligible.

where the uncertainty is governed by the uncertainty on the branching fraction for $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$ decays. This yield of this kind of background is very small, which therefore makes it negligible.

Background from $B^+ \to K^+\mu^+\mu^-$

Decays of the type $B^+ \to K^+\mu^+\mu^-$ combined with a random kaon candidate from the event can pass the selection and form a background, which does not peak in the B_s^0 mass signal region, but can enhance the background close to the B_s^0 mass region and therefore influence the signal yield obtained through a fit. The distribution of the 3-particle invariant mass for $B^+ \to K^+\mu^+\mu^-$ candidates in data is given in Figure 6.5. There are no candidates found close to the B^+ signal region and therefore the effect of this background is negligible. This is mainly due to the narrow dikaon mass window in the selection, which rejects most dikaon combinations involving a random kaon from the event. Therefore backgrounds from this decay are also excluded.

Conclusion

Despite the large number of decays, which can contribute as a background when misidentifying one or two of the daughters, all decays have been shown to have a negligible

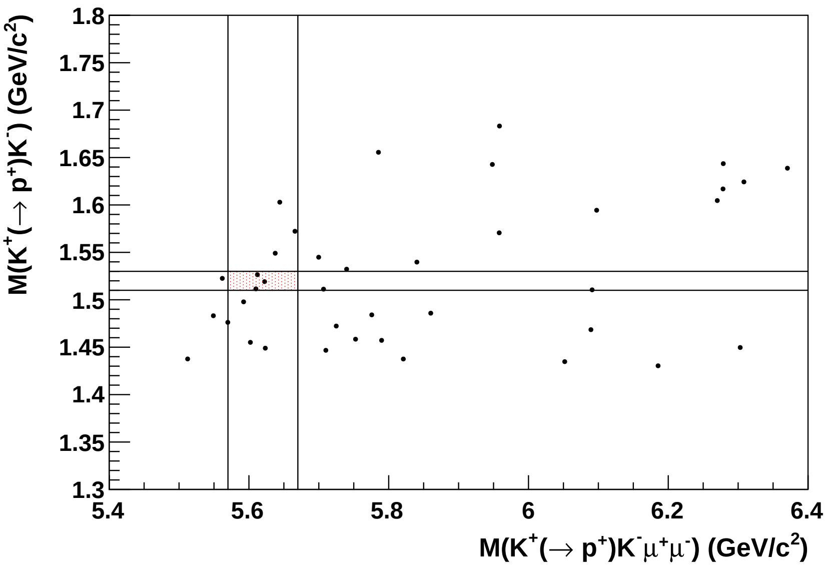


Figure 6.4: Kinematic resolved area for misidentified candidates from $\Lambda_b \to \Lambda(1520)(\to pK^-)\mu^+\mu^-$ decays. If the proton is misidentified as a kaon $\Lambda_b \to \Lambda(1520)\mu^+\mu^-$ events can be a source of background. Switching the masses of the kaon candidates in data accordingly and applying the offline cuts gives the above entries. There are three candidates found inside the kinematic allowed region (shaded red), indicating that this kind of background is not negligible.

contribution.

6.4.2 Fit model for $M_{K^+K^-\mu^+\mu^-}$

In order to obtain the number of signal candidates for both $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ decays, several unbinned maximum likelihood fits to the reconstructed B_s^0 mass distributions were performed by varying the signal and background parametrisation. Several different combinations of models are given in Figure 6.6.

Figure 6.6 shows that when the background is parametrised by an exponential shape the χ^2 of the fit is lower than when a linear shape is used. The Crystal Ball function and the double Gaussian function are also more suited to parametrise the signal than a single Gaussian function. The Crystal Ball function is described in detail elsewhere [74, 75].

However, to be aligned with other analyses of the LHCb rare decays working group a double Crystal Ball function was chosen as the fit model for the signal. The parameters of a double Crystal Ball function can be chosen such that all functions from Figure 6.6 (double Gaussian and Crystal Ball) can be 'constructed'.

The fitting strategy was the following:

1. Each Crystal Ball function is described by 4 parameters (σ, μ, α and n), giving a total of 8 parameters for a double Crystal Ball function. μ, α and n of both

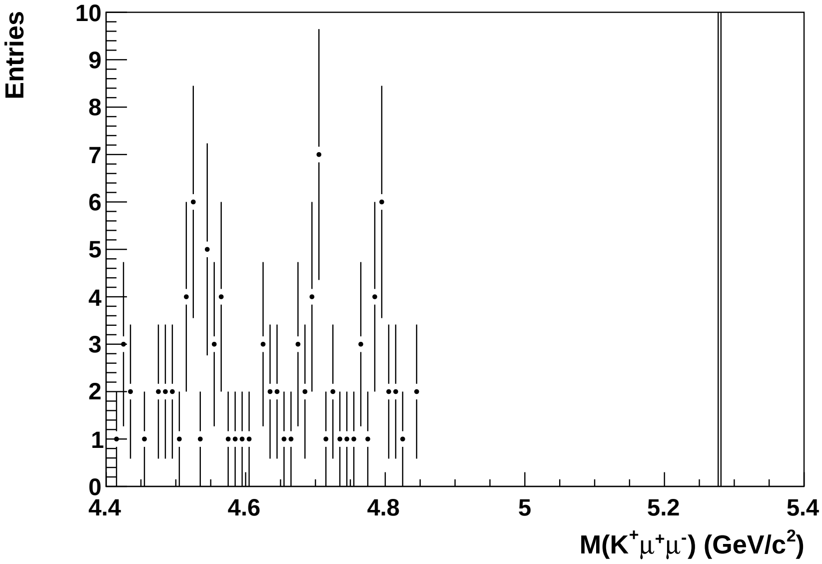


Figure 6.5: Reconstructed mass of $B^+ \to K^+\mu^+\mu^-$ candidates in data. No events are found close to $M(K^+\mu^+\mu^-) = 5279.2 \pm 0.4\,\mathrm{MeV}/c^2$, the known B^+ mass [37].

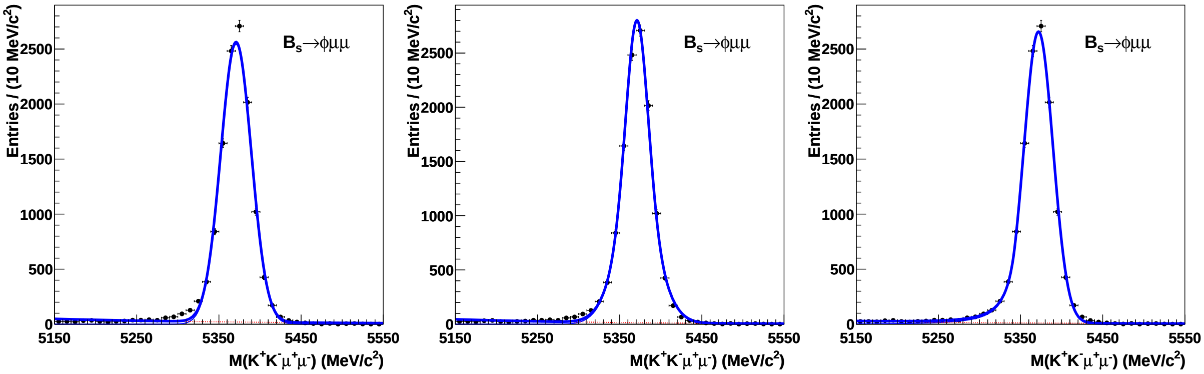
Crystal Ball functions are fixed to always be the same value. Therefore the number of parameters is effectively reduced to 5.

2. The q^2 bin containing $B_s^0 \to J/\psi\,\phi$ candidates has the largest signal statistics and is therefore used first to fit the model to the data. All parameters are left floating and converge to a value. The underlying `MINUIT`-fitter also assigns an error to each value.

3. When fitting the model to the candidates in the remaining bins of q^2, all parameters are fixed to the values obtained from the fit to $B_s^0 \to J/\psi\,\phi$ candidates in (2).

The mass resolution (σ) of a Gaussian fitted to the reconstructed B_s^0 candidate mass is not constant over the entire range of q^2. This is illustrated in Figure 6.7, which shows the variation of σ with q^2. The resolution assigned for each q^2 bin was obtained from fitting a single Gaussian curve (with floating parameters) to the reconstructed B_s^0 mass in the MC $B_s^0 \to \phi\mu^+\mu^-$ sample. Although the σ parameters of the double Crystal Ball function (used to fit the data) are not precisely equivalent to the σ parameter studied in Figure 6.7, the same q^2 dependence is assumed for data and the double Crystal Ball function. For each fit to data the resolution was consequently fixed to the relative change of σ in simulation multiplied with σ returned by the fit to the $B_s^0 \to J/\psi\,\phi$ candidates in data.

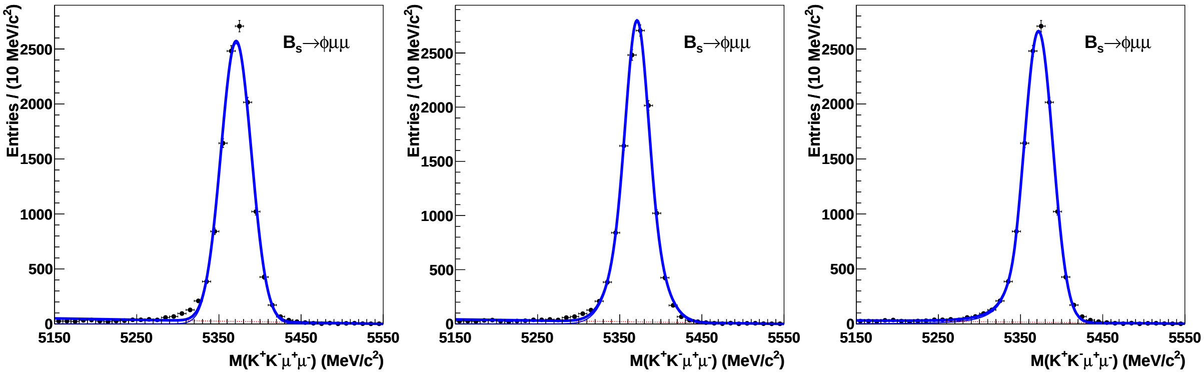
The systematic uncertainty of the fit model was obtained by varying the fixed fit parameters in the following manner (repeated for each bin):



(a)
S = Gaussian
B = exponential
χ^2 = 15.38
p = 0.15%

(b)
S = Double Gaussian
B = exponential
χ^2 = 11.61
p = 2.05%

(c)
S = Crystal Ball
B = exponential
χ^2 = 8.03
p = 15.46%



(d)
S = Gaussian
B = linear
χ^2 = 15.45
p = 0.15%

(e)
S = Double Gaussian
B = linear
χ^2 = 11.90
p = 1.81%

(f)
S = Crystal Ball
B = linear
χ^2 = 9.14
p = 10.37%

Figure 6.6: Mass distributions of $B_s^0 \to \phi\mu^+\mu^-$ candidates from simulation with the signal and background parametrised by a combination of a Gaussian, double Gaussian or Crystal Ball function and a linear or an exponential background shape (a)-(f). The χ^2 and probability p of the fit are given in each case.

1. The values of α and n are varied according to their `MINUIT`-errors (α^{error} and n^{error}). A Gaussian centred at 0 with a mean of 1 is used to generate a random value c, which is multiplied with the respective `MINUIT`-error. The new values for α and n are given by:

$$\alpha' = \alpha + \alpha^{\text{error}} \times c \tag{6.6}$$

$$n' = n - n^{\text{error}} \times c \tag{6.7}$$

   The signs are opposite because α and n are assumed to be 100% negatively correlated. This was indicated by the `MINUIT`-correlation table of the initial fit to the $B_s^0 \to J/\psi\,\phi$ candidates.

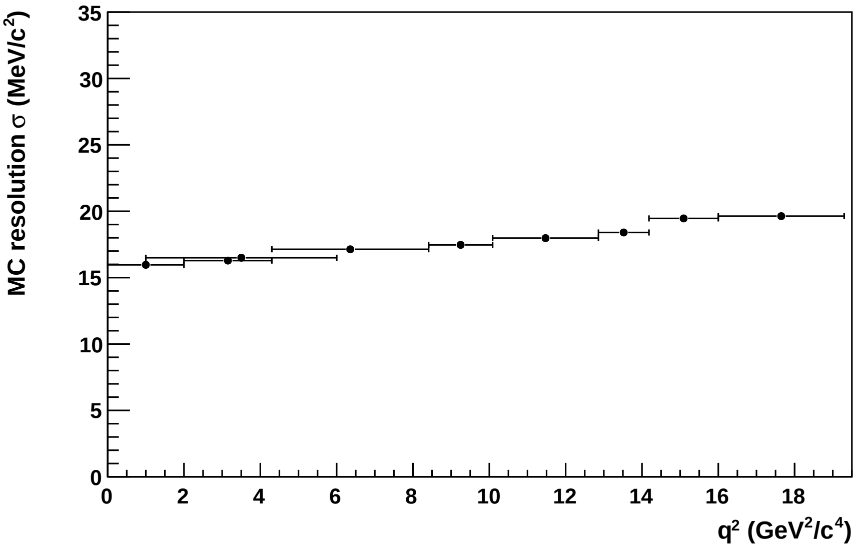


Figure 6.7: Variation of the mass resolution of $B^0_s \to \phi\mu^+\mu^-$ candidates in different bins of q^2 obtained from simulation. The mass resolution is the Gaussian parameter σ obtained from a fit to the MC sample.

2. The data are then re-fitted with the slightly changed values, α' and n'. A new number of signal events is obtained. This is repeated 250 times.

3. The distribution of signal events (with 250 entries) is fitted with a single Gaussian, which is centred around the number of signal events initially obtained. The one sigma width of this fit is assigned as the systematic uncertainty of the fit model.

During the fitting procedure special attention was paid to the q^2 bins, which cover the asymmetric veto regions (shown in Figure 6.1). In these bins the lower background band contains proportionally less events than the upper band due to the lower mass veto region for the radiative tail. This affects the shape of the fit to the background events. Events in the lower band are therefore given a larger weight during the fit than the events in the upper band. The weight is determined by the relative size of the veto regions.

When there are n candidates in a single event, each candidate is assigned a weight of $1/n$ during the fitting.

There is a kinematic upper limit on q^2 of where the combined mass of the 4 daughters exceeds the mass of the B^0_s mass (at $\sim 19.3\,\mathrm{GeV}/c^2$). Therefore no candidates are expected to be above this limit, which is found to be in the largest q^2 bin.

6.4.3 Results of the fits

Each recorded and selected candidate in data has been triggered by a certain trigger line or several lines. The LHCb trigger system is described in more detail in Section 6.5.3. For an observation of a $B^0_s \to \phi\mu^+\mu^-$ signal all candidates passing the LHCb trigger are used and a significance can be calculated from the fit. The fit functions used are described in

Section 6.4.2. The results of the fits to all $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \phi\mu^+\mu^-$ candidates in data are given in Figure 6.8. From the fitted functions the signal yields are extracted, giving $N_{J/\psi\,\phi} = 12390 \pm 119$ and $N_{\phi\mu\mu} = 80 \pm 10$ (excluding veto regions), where the uncertainties are statistical only. For the $B_s^0 \to J/\psi\,\phi$ fit the relevant parameters converge at $M_{B_s}^{J/\psi\,\phi} = 5364.0 \pm 0.3\,\mathrm{MeV}/c^2$ and $\sigma^{J/\psi\,\phi} = 19.9 \pm 0.2\,\mathrm{MeV}/c^2$. The parameters of the fit to the $B_s^0 \to \phi\mu^+\mu^-$ candidates, shown in Figure 6.8(b), are fixed to the same values. The fit has a significance of 12.31, where the significance is given by $\sqrt{2|\Delta(\min\log\mathcal{L})|}$. The term $\Delta(\min\log\mathcal{L})$ is the difference in the minimised log-likelihoods of the fit shown and the null hypothesis fit (where $N_{\phi\mu\mu}$ is set to zero).

To determine the trigger efficiency (explained in detail in Section 6.5.3) the MC samples described in Section 6.2 were used. The simulation of non-physical, beam-gas, minimum bias, rate limited and phase-space specific trigger lines is non-trivial and the effect of pile-up events during data taking cannot be accounted for in the MC samples. Therefore it is of advantage to limit the trigger lines to a small subset of lines, which are well understood and which select candidates because of the signal components (also referred to as TOS - Triggered On Signal, which is also described in more detail in Section 6.5.3). All candidates, that are used for the final branching fraction ratio measurement (which relies on the trigger efficiency being calculated from simulation), are required to be selected as TOS by following lines:

- L0:

  `Muon, DiMuon, Hadron`

- HLT1:

  `TrackMuon, TrackAllL0, DiMuonLowMass, DiMuonHighMass,`
  `SingleMuonHighPT`

- HLT2:

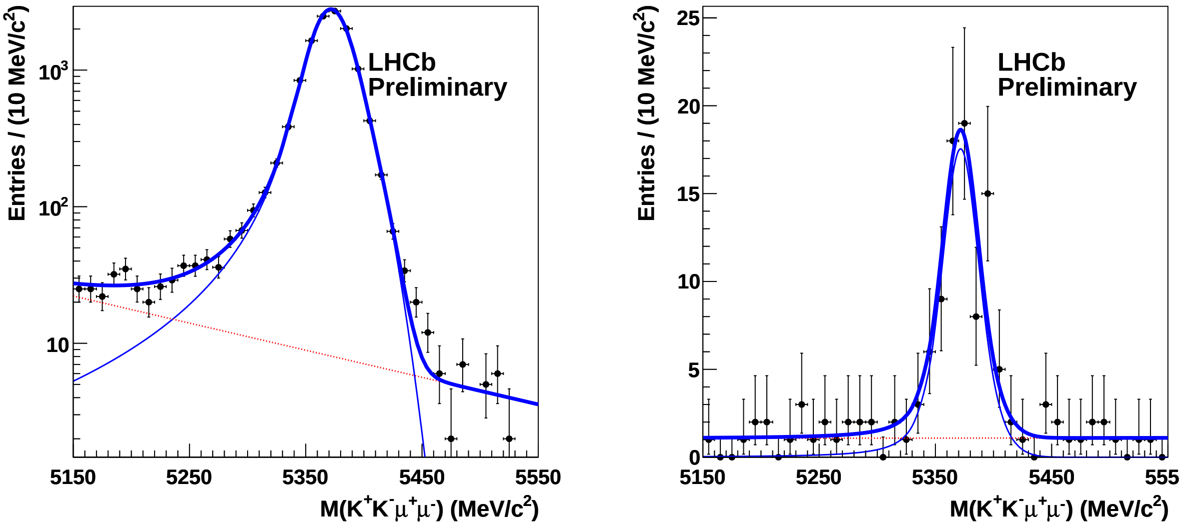
  `TopoMu[2,3,4]BodyBBDT, Topo[2,3,4]BodyBBDT, Topo[2,3,4]BodySimple,`
  `DiMuonDetached, SingleMuon, IncPhi`

where at least one line of each trigger stage has to be fired as TOS. This requirement reduces the overall number of signal candidates from $N_{\phi\mu\mu} = 80 \pm 10$ to 77 ± 10 and from $N_{J/\psi\,\phi} = 12390 \pm 119$ to 11090 ± 115. These numbers were obtained from fitting the same functions as in Figure 6.8 to the candidates, that satisfy the additional TOS requirement. The errors are statistical. The efficiency of the TOS requirement is accounted for in the trigger efficiency (described later in Section 6.5.3).

Taking the ratio of the number of TOS signal events gives:

$$\frac{N_{\phi\mu\mu}}{N_{J/\psi\,\phi}} = (6.931 \pm 0.867) \times 10^{-3} \tag{6.8}$$

where the uncertainty is statistical. An uncertainty of 0.019×10^{-3} from the fit model (introduced in Section 6.4.2) was assigned as a systematic uncertainty to the ratio. The TOS signal yields of $B_s^0 \to \phi\mu^+\mu^-$ candidates and the significances of the fits in bins of q^2 are summarised in Table 6.4. The corresponding fits are shown in Figure 6.9.



(a) (b)

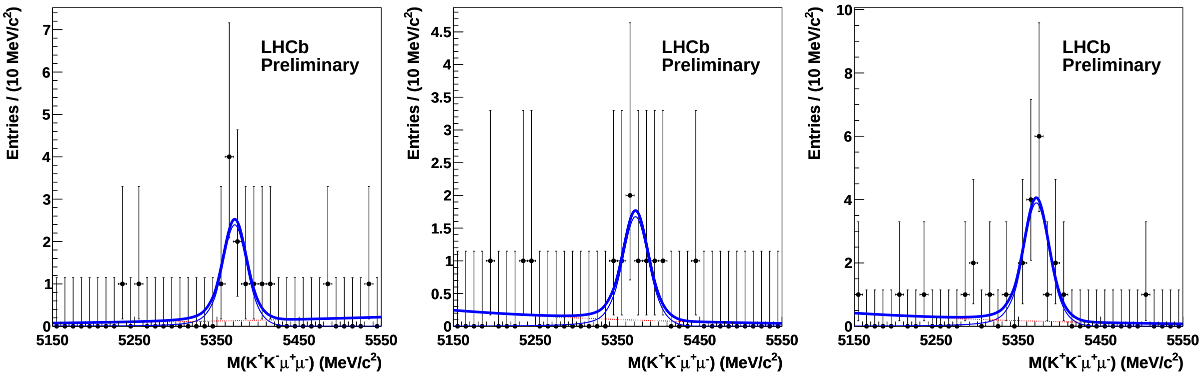
Figure 6.8: Mass distribution of B_s^0 candidates for $B_s^0 \to J/\psi\,\phi$ (a) and $B_s^0 \to \phi\mu^+\mu^-$ (b). The points represent data. A double Crystal Ball function for the signal (thin blue line) on top of an exponential background (red pointed line) was fitted to the distribution in (a). For the fit in (b) the parameters were fixed to the values obtained from the fit in (a).

Table 6.4: Summary of $B_s^0 \to \phi\mu^+\mu^-$ signal yields and significances in bins of q^2.

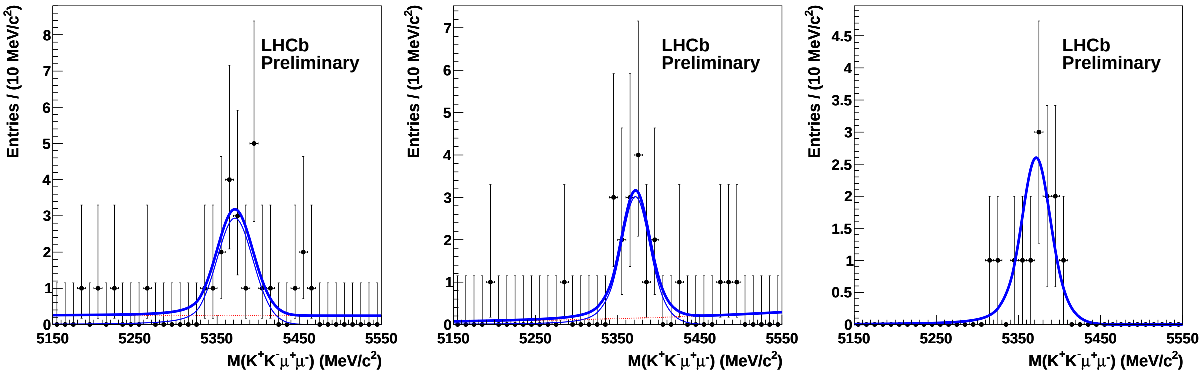
q^2 (GeV^2/c^4)	$N_{\phi\mu\mu}$	significance
0.00 – 2.00	9.8 ± 3.4	4.54
2.00 – 4.30	7.3 ± 2.9	3.82
4.30 – 8.68	17.0 ± 4.5	6.08
10.09 – 12.86	17.2 ± 4.6	5.37
14.18 – 16.00	14.9 ± 4.2	5.55
16.00 – 19.30	13.0 ± 3.6	5.90
1.00 – 6.00	15.9 ± 4.3	5.61

6.4.4 Cross-checks

The number of expected $B_s^0 \to \phi\mu^+\mu^-$ events per bin of q^2 can be estimated using the branching fractions calculated by the CDF experiment [48]. In Figure 6.10 a comparison of the LHCb yields and the expected yields (using the CDF results) is given.



(a) $0.00 < q^2 < 2.00$ (b) $2.00 < q^2 < 4.30$ (c) $4.30 < q^2 < 8.68$



(d) $10.09 < q^2 < 12.86$ (e) $14.18 < q^2 < 16.00$ (f) $16.00 < q^2 < 19.30$

Figure 6.9: Mass distributions of $B_s^0 \to \phi\mu^+\mu^-$ candidates in bins of q^2 (a)-(f). The points represent data. By fitting a single Crystal Ball function for the signal (thin blue line) on top of an exponential background (red pointed line) the number of signal candidates is obtained. The resolution of the fit in each bin is fixed to the resolution of the fit to the $B_s^0 \to J/\psi\,\phi$ candidates multiplied by a factor, which takes into account the variation of the resolution with q^2 observed in simulation.

Selecting $B^0 \to K^{*0}\mu^+\mu^-$ candidates (as well as the J/ψ and ψ' resonance modes) with the $B_s^0 \to \phi\mu^+\mu^-$ offline cuts is a further interesting comparison to make. For this cross-check the selection had to be slightly modified (removing PID requirements on the K^- candidate and adjusting mass windows). The yields of $B^0 \to K^{*0}\mu^+\mu^-$, $B^0 \to K^{*0}J/\psi$, $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ candidates are given in Table 6.5. The fitted functions, from which the $B^0 \to K^{*0}\mu^+\mu^-$ candidate yields were obtained, are given in Figure 6.12. The distributions of the number of selected $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ signal candidates in bins of q^2 are given in Figure 6.11 as a comparison. Jumping ahead and using the efficiencies calculated in the next sections, the efficiency corrected yields have been calculated, which are also given in Table 6.5. From the various yields given in Table 6.5 several ratios can be calculated and compared, as done in Table 6.6. The ratios between the B_s^0 and B^0 system are in agreement within errors, which gives confidence in the validity of the measurement.

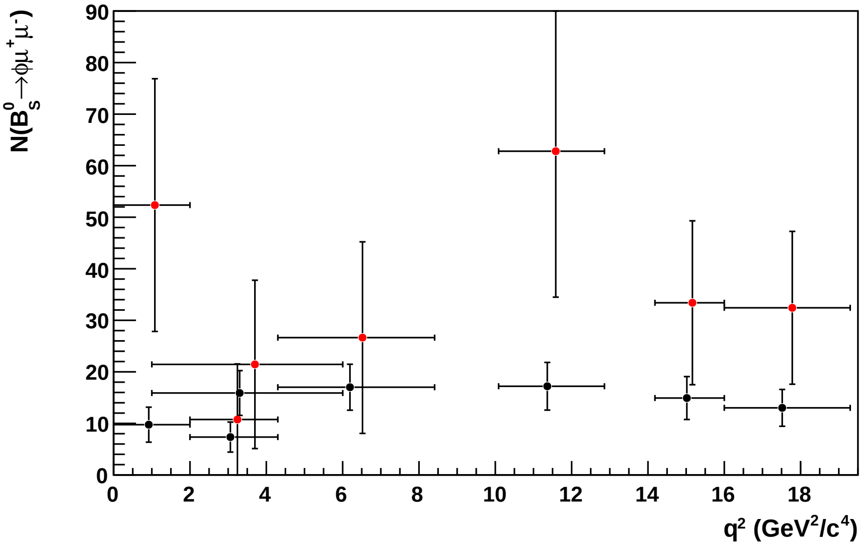
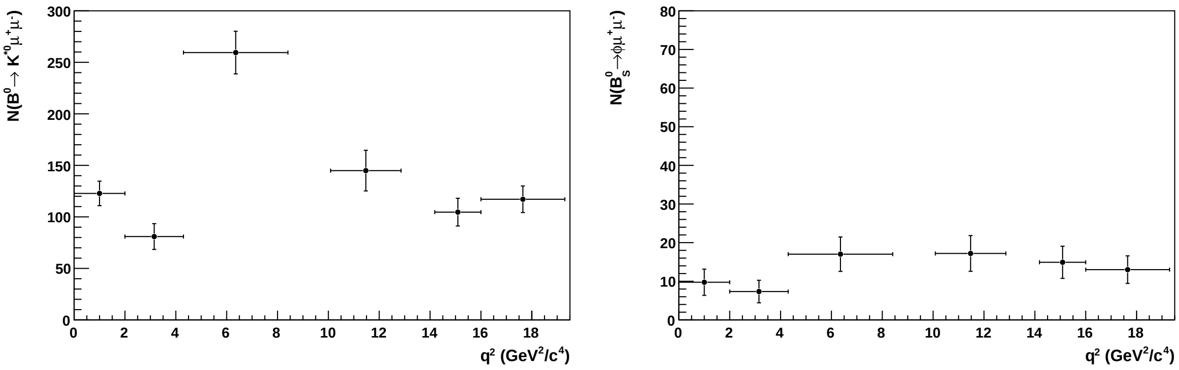


Figure 6.10: The observed number of $B_s^0 \to \phi\mu^+\mu^-$ candidates in black and the expected number of candidates in red in bins of q^2. For the expected number of candidates the branching fractions measured by the CDF experiment were used [48].

The large number of $B_s^0 \to J/\psi\,\phi$ candidates in data enables accurate comparisons with simulation. A table comparing the cut efficiencies of the offline cuts described in Section 6.3 is given in Appendix D. Several figures, showing the ratio of the distributions of kinematic variables in data over simulation normalised to unit area, are also given in Appendix D. This includes the opening angle $\theta_{K^+K^-}$ between the two kaons. This angle is very small in low q^2 regions as the kaons have a large momentum, which could affect the reconstruction. However, all cross-checks indicate that the MC samples describe the variables of relevance as found in data with a good accuracy.



(a) (b)

Figure 6.11: The number of measured candidates for $B^0 \to K^{*0}\mu^+\mu^-$ (a) and $B_s^0 \to \phi\mu^+\mu^-$ (b) in bins of q^2.

Table 6.5: Comparison of the $B^0 \to K^*X$ and $B_s^0 \to \phi X$ candidate yields (a) in $1\,\mathrm{fb}^{-1}$, where $X = \mu\mu$ or $X = J/\psi$. Apart from adjusted PID requirements and mass windows the candidates have been selected with the same offline cuts. (b) contains the efficiency corrected yields.

(a)

	$X = \mu\mu$	$X = J/\psi$
$B^0 \to K^*X$	806.5 ± 40.2	$96,056 \pm 2,650$
$B_s^0 \to \phi X$	79.5 ± 10.0	$12,309 \pm 119$

(b)

	$X = \mu\mu$	$X = J/\psi$
$B^0 \to K^*X$	$124,928 \pm 14,581$	$10,628,398 \pm 532,615$
$B_s^0 \to \phi X$	$14,001 \pm 1904$	$1,487,713 \pm 81,995$

6.5 Estimation of $\varepsilon_{J/\psi\,\phi}/\varepsilon_{\phi\mu\mu}$

The overall efficiency is the product of the geometrical acceptance of LHCb ($\varepsilon^{\mathrm{geo}}$), the detection, reconstruction and selection efficiencies ($\varepsilon^{\mathrm{rec\&sel}}$) and the trigger efficiency ($\varepsilon^{\mathrm{trigger}}$). For each efficiency the ratio between $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \phi\mu^+\mu^-$ decays has to be calculated. The overall efficiency ratio is given by:

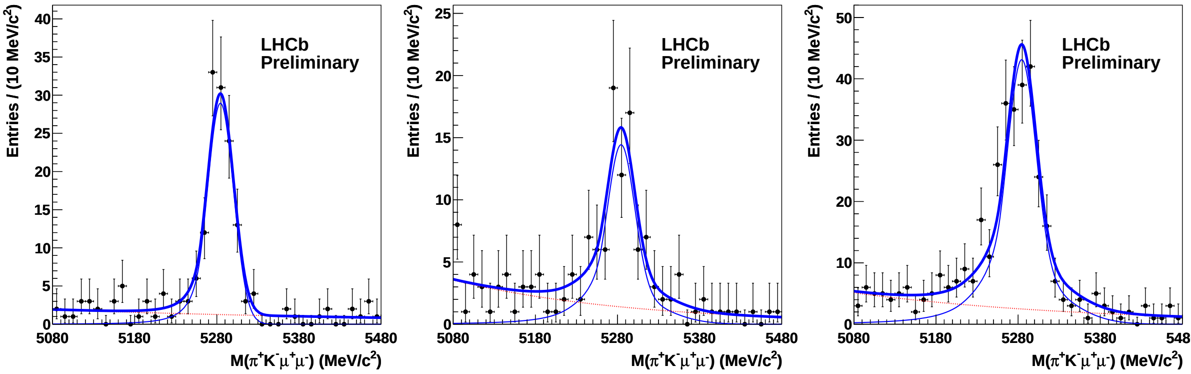
$$\frac{\varepsilon_{J/\psi\,\phi}}{\varepsilon_{\phi\mu\mu}} = \frac{\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{geo}}_{\phi\mu\mu}} \times \frac{\varepsilon^{\mathrm{rec\&sel}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{rec\&sel}}_{\phi\mu\mu}} \times \frac{\varepsilon^{\mathrm{trigger}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{trigger}}_{\phi\mu\mu}} \tag{6.9}$$

where all terms are evaluated from simulation. For the last term the trigger efficiency of $B_s^0 \to J/\psi\,\phi$ decays has been cross-checked with a data-driven method. In the following the errors on each efficiency ratio are treated as systematic uncertainties, which are combined and added to the final result.

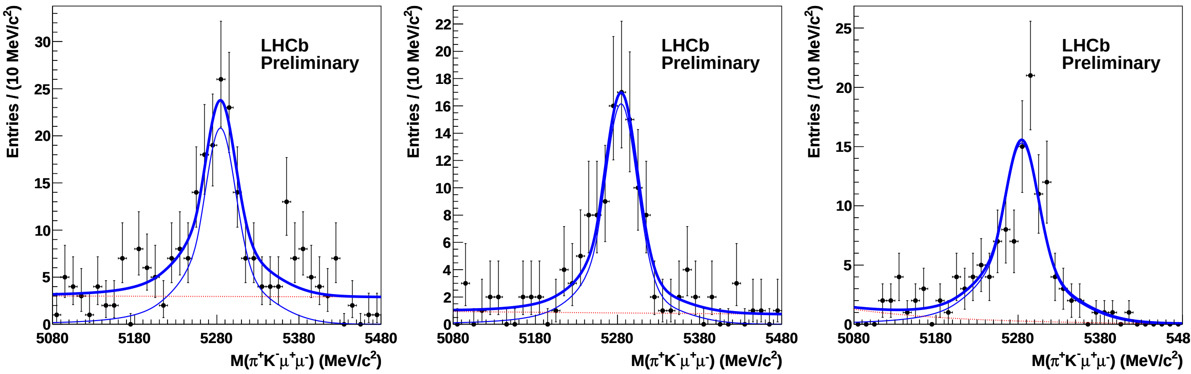
6.5.1 $\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{geo}}_{\phi\mu\mu}$

The ratio of $\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{geo}}_{\phi\mu\mu}$ is estimated using simulation, where $\varepsilon^{\mathrm{geo}}$ is the efficiency for all four final state particles in an event to lie within the geometrical acceptance of LHCb. The acceptance region is defined in Section 5.4.1. $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \phi\mu^+\mu^-$ decays are equivalent topologically and therefore the ratio of both acceptance efficiencies is expected to be close to unity. The results for the entire q^2 range, $\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi} = 0.160$ and $\varepsilon^{\mathrm{geo}}_{\phi\mu\mu} = 0.162$, give a ratio of:

$$\frac{\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{geo}}_{\phi\mu\mu}} = 0.987 \pm 0.005 \tag{6.10}$$



(a) $0.00 < q^2 < 2.00$ (b) $2.00 < q^2 < 4.30$ (c) $4.30 < q^2 < 8.68$



(d) $10.09 < q^2 < 12.86$ (e) $14.18 < q^2 < 16.00$ (f) $16.00 < q^2 < 19.30$

Figure 6.12: Mass distributions of $B^0 \to K^{*0}\mu^+\mu^-$ candidates in bins of q^2 (a)-(f). The points represent data. By fitting a single Crystal Ball function for the signal (thin blue line) on top of an exponential background (red pointed line) the number of signal candidates is obtained. The resolution of the fit in each bin is fixed to the resolution of the fit to the $B^0 \to K^{*0}J/\psi$ candidates multiplied with a factor, which takes into account the variation of the resolution with q^2 observed in simulation.

where the error is systematic. It is obtained by changing the couplings of the default decay model to non-SM couplings, such as anomalous Wilson coefficients C_7 and C_{10} [83]. The details of the non-SM decay models are described in Section 6.2. The largest difference between the default and the other models is assigned as systematic uncertainty. The effect of changing the decay model parameters for $B_s^0 \to J/\psi\,\phi$ decays is also accounted for in the error. To illustrate the difference between the decay models the q^2 distributions after the Gauss generation stage (a specific LHCb software using Pythia [85, 86]) are given in Figure 6.13. The variation of $\varepsilon^{\text{geo}}_{\phi\mu\mu}$ in bins of q^2 is given in Figure 6.14. The ratios of $\varepsilon^{\text{geo}}_{J/\psi\,\phi}/\varepsilon^{\text{geo}}_{\phi\mu\mu}$ in bins of q^2 are summarised in Table 6.7.

Table 6.6: Ratios between the yields of $B^0 \to K^* X$ and $B_s^0 \to \phi X$ candidates (a) taken from Table 6.5, where $X = \mu\mu$ or $X = J/\psi$. The ratios between the efficiency corrected yields are given in (b).

(a)

Mode	Ratio
$B_s^0 \to \phi\mu^+\mu^- / B_s^0 \to J/\psi\,\phi$	0.006 ± 0.001
$B^0 \to K^{*0}\mu^+\mu^- / B^0 \to K^{*0} J/\psi$	0.008 ± 0.003
$B_s^0 \to \phi\mu^+\mu^- / B^0 \to K^{*0}\mu^+\mu^-$	0.096 ± 0.014
$B_s^0 \to J/\psi\,\phi / B^0 \to K^{*0} J/\psi$	0.129 ± 0.052

(b)

Mode	Ratio
$B_s^0 \to \phi\mu^+\mu^- / B_s^0 \to J/\psi\,\phi$	0.009 ± 0.001
$B^0 \to K^{*0}\mu^+\mu^- / B^0 \to K^{*0} J/\psi$	0.012 ± 0.001
$B_s^0 \to \phi\mu^+\mu^- / B^0 \to K^{*0}\mu^+\mu^-$	0.111 ± 0.020
$B_s^0 \to J/\psi\,\phi / B^0 \to K^{*0} J/\psi$	0.140 ± 0.010

Table 6.7: Summary of the geometrical acceptance efficiency ratios in bins of q^2.

q^2 (GeV^2/c^4)	$\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{geo}}_{\phi\mu\mu}$
0.00 – 2.00	0.943 ± 0.010
2.00 – 4.30	1.002 ± 0.007
4.30 – 8.68	1.018 ± 0.010
10.09 – 12.86	0.991 ± 0.007
14.18 – 16.00	0.979 ± 0.003
16.00 – 19.30	0.990 ± 0.007
1.00 – 6.00	1.003 ± 0.007

6.5.2 $\varepsilon^{\mathrm{rec\&sel}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{rec\&sel}}_{\phi\mu\mu}$

The selection and reconstruction efficiency, $\varepsilon^{\mathrm{rec\&sel}}$, is calculated from simulation. The MC $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \phi\mu^+\mu^-$ samples contain both the simulated data and truth information of the data. The simulated data are the outcome of a simulation of the LHCb detector and its reconstruction and stripping algorithms. The truth information contains the particle properties at generator level. The difference between both is the reconstruction efficiency. The selection efficiency is calculated by applying all offline selection cuts described in Section 6.3 to the MC samples. The results are $\varepsilon^{\mathrm{rec\&sel}}_{\phi\mu\mu} = 0.0438 \pm 0.0021$ and $\varepsilon^{\mathrm{rec\&sel}}_{J/\psi\,\phi} = 0.0581 \pm 0.0028$, giving a ratio of:

$$\frac{\varepsilon^{\mathrm{rec\&sel}}_{J/\psi\,\phi}}{\varepsilon^{\mathrm{rec\&sel}}_{\phi\mu\mu}} = 1.326 \pm 0.090 \tag{6.11}$$

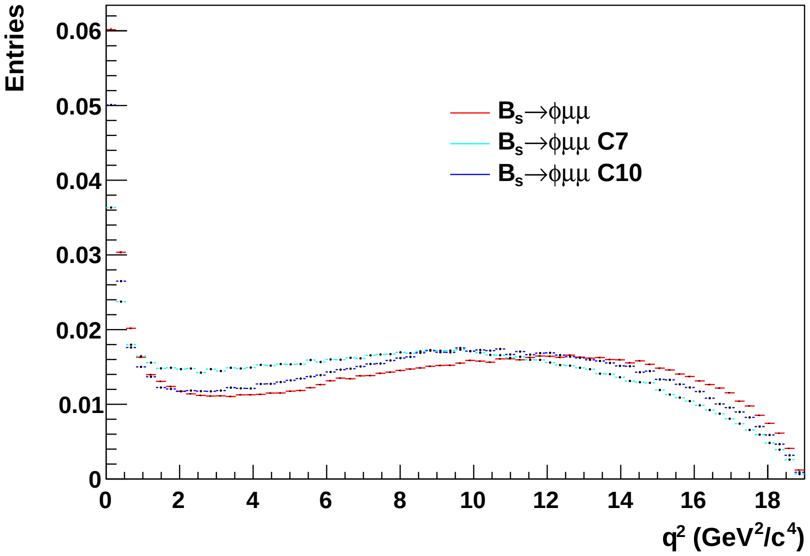


Figure 6.13: The true q^2 distributions for different models. The default model used in simulation is shown in red. The distributions for two models with non Standard Model couplings, where the Wilson coefficients C_7 and C_{10} have been changed [83], are shown in turquoise and blue respectively. For each model 1×10^6 events were generated. All three distributions are normalised to unit area.

where the error is systematic and accounts for differences in the descriptions of variables between data and simulation. In particular the effect of the description of the IP resolution on the efficiency was studied (in a similar manner as described in Section 5.10). The variation of $\varepsilon_{\phi\mu\mu}^{\text{rec\&sel}}$ as a function of different χ^2_{IP} selection cut values is shown in Figure 6.15. Indicated are the variations for the lowest and highest q^2 bin in order to show the variation of the two extreme cases. Using the gradient of the straight line fit a systematic uncertainty can be obtained: Assuming a (conservative) ±10% difference between the description of χ^2_{IP} in simulation and in data the gradient can be used to estimate an uncertainty on $\varepsilon_{\phi\mu\mu}^{\text{rec\&sel}}$. A ±10% shift around $\chi^2_{\text{IP}} = 16$ corresponds to a change of ±1.6 along the x-axis, which in turn corresponds to a ±4.8%(±1%) shift in $\varepsilon_{\phi\mu\mu}^{\text{rec\&sel}}$ within the highest (lowest) q^2 bin. To be on the conservative side a systematic uncertainty of ±4.8% is assigned to the efficiency in every bin of q^2. Other quantities relevant to the analysis (for example the B_s^0 momentum and p_{T}) are in good agreement between data and simulation and have a negligible systematic uncertainty.

The effect of the uncertainty of the PID performance was also taken into account. As explained in Section 6.3, the cuts on the PID of the daughters were applied by using weights, representing the data performance, instead of the simulated PID values. The weights are binned in p_{T} and η and are summarised in look-up tables for different PID cut values. Therefore a unique weight for a given PID cut can be obtained for each

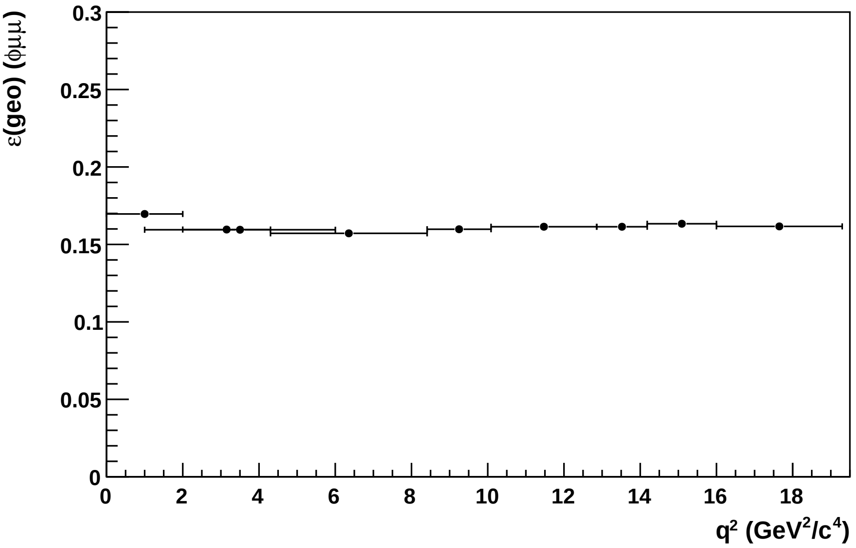


Figure 6.14: Distribution of $\varepsilon^{\rm geo}_{\phi\mu\mu}$ in bins of q^2.

event. A systematic uncertainty arises from different tables, where each table is tuned to represent slightly different data configurations. The effect on the reconstruction and selection efficiency ratio was found to be negligible.

The efficiency ratio is significantly different from unity. The reason is that in the MC $B_s^0 \to \phi\mu^+\mu^-$ sample $\varepsilon^{\rm rec\&sel}$ is a function of q^2 due to one of the applied selection cuts. Figure 5.5 in the previous chapter shows the ratio between $\varepsilon^{\rm rec\&sel}$ obtained with $\chi^2_{\rm IP}(K^\pm)$ > 9 and $\varepsilon^{\rm rec\&sel}$ obtained with $\chi^2_{\rm IP}(K^\pm) > 25$ as a function of q^2. The decrease of the ratio at large q^2 is explained by kinematical arguments. In this region the ϕ has less momentum and less $p_{\rm T}$ than in other regions. Therefore the kaons produced have less $p_{\rm T}$ available and their momentum vector is more likely to point to the primary vertex. This results in smaller values for the IP and causes the ratio in Equation 6.11 to be significantly different from unity. As in the previous measurement (Chapter 5) this bias is unavoidable. Figure 6.16 shows the variation of $\varepsilon^{\rm rec\&sel}_{\phi\mu\mu}$ with bins of q^2. The ratios of $\varepsilon^{\rm rec\&sel}_{J/\psi\,\phi}/\varepsilon^{\rm rec\&sel}_{\phi\mu\mu}$ in bins of q^2 are summarised in Table 6.8.

6.5.3 $\varepsilon^{\rm trigger}_{J/\psi\,\phi}/\varepsilon^{\rm trigger}_{\phi\mu\mu}$

The LHCb trigger is highly efficient in selecting $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ decays with two muons in the final state. In Section 4.3 the details of the trigger stages (L0, HLT1 and HLT2) and trigger lines are given. Due to the topological similarity the trigger efficiencies are expected to be very close for both $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ decays. The trigger efficiency can be estimated using the TISTOS technique. The details of this method are given in Section 4.3.4 and in [87]. The method can also be used to calculate the trigger efficiency for TOS events only. The TOS efficiency of all trigger stages

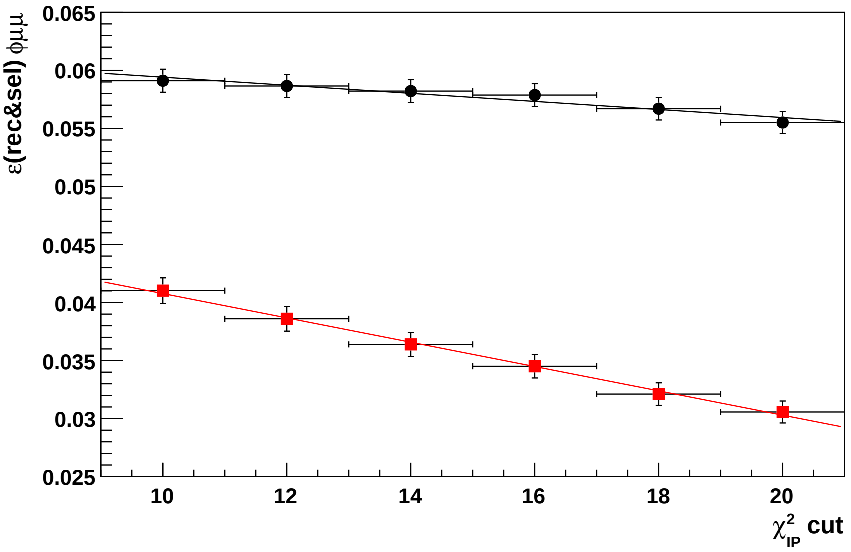


Figure 6.15: Variation of $\varepsilon^{\rm rec\&sel}_{\phi\mu\mu}$ with the $\chi^2_{\rm IP}$ cut on the daughter particles $K^{\pm}$ and $\mu^{\pm}$. The black circles show the values for the lowest q^2 bin ($0-2\ \mathrm{GeV}^2/c^4$) and the red squares for the highest q^2 bin ($16.0-19.3\ \mathrm{GeV}^2/c^4$). The gradients of the straight line fits come out to be -5.6×10^{-4} and -16.8×10^{-4} respectively.

Table 6.8: Summary of the reconstruction and selection efficiency ratios in bins of q^2.

q^2 (GeV^2/c^4)	$\varepsilon^{\rm rec\&sel}_{J/\psi\,\phi}/\varepsilon^{\rm rec\&sel}_{\phi\mu\mu}$
0.00 – 2.00	1.003 ± 0.068
2.00 – 4.30	1.012 ± 0.069
4.30 – 8.68	0.989 ± 0.067
10.09 – 12.86	1.038 ± 0.070
14.18 – 16.00	1.281 ± 0.087
16.00 – 19.30	1.682 ± 0.114
1.00 – 6.00	1.011 ± 0.069

combined (L0×HLT1×HLT2) is the relevant quantity for this analysis (as explained in Section 6.4.3 only signal events selected as TOS by a subset of well-understood trigger lines are considered). The disadvantage of this method is that it requires large statistics in each bin of q^2, especially in order to determine the HLT2 efficiencies correctly. The current statistics in data are only sufficient for calculating the trigger efficiency of $B^0_s\to J/\psi\,\phi$ candidates. Therefore the trigger efficiency has to be obtained from simulation. The trigger decision for each candidate in the MC samples was simulated (using the LHCb software MOORE [52]) for the main trigger conditions (TCKs) used during the 2011 data taking period. This allows the TISTOS method to be used on simulation in the same manner as on data. The TCKs, that were used on most of the selected data, were simulated in the MC samples (in hexadecimal: `0x006d0032` and `0x00770037`). Each trigger efficiency

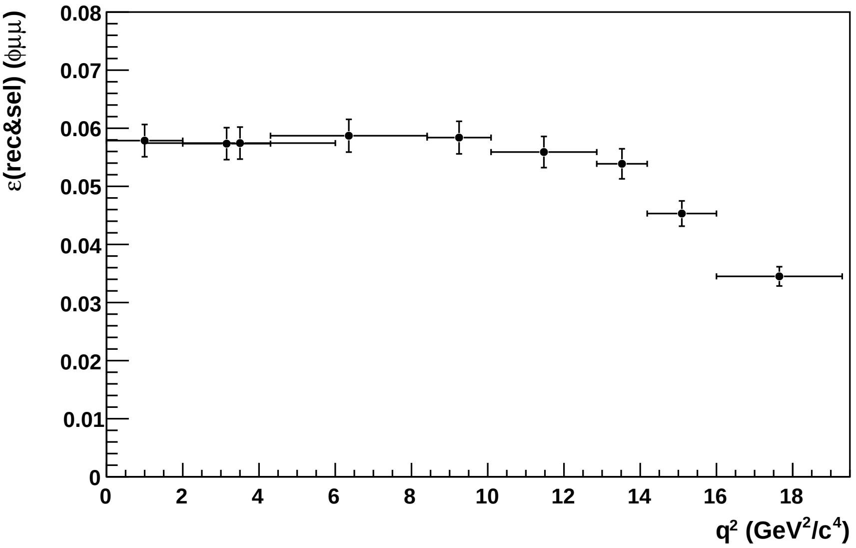


Figure 6.16: Distribution of $\varepsilon_{\phi\mu\mu}^{\text{rec\&sel}}$ in bins of q^2.

obtained from simulation (L0, HLT1, L0× HLT1, etc.) can be compared to the equivalent one from data (where statistics allows). The MC samples also enable one to calculate the true MC trigger efficiency directly (using the MC truth trigger decisions), which is used as the central value for this analysis. The efficiency calculated from the TISTOS method on data (in a bin with sufficient statistics, e.g. the one containing $B_s^0 \to J/\psi\,\phi$ candidates) can be used to validate the MC truth efficiency. This allows to quantify an uncertainty coming from the various methods to obtain trigger efficiencies in simulation and data. The trigger lines, that are used to select TOS events in Section 6.4.3, are the same ones as used for the TISTOS method.

The MC truth global TOS efficiencies are $\varepsilon_{J/\psi\,\phi}^{\text{trigger}} = 0.803 \pm 0.020$ and $\varepsilon_{\phi\mu\mu}^{\text{trigger}} = 0.773 \pm 0.019$, giving a ratio of:

$$\frac{\varepsilon_{J/\psi\,\phi}^{\text{trigger}}}{\varepsilon_{\phi\mu\mu}^{\text{trigger}}} = 1.038 \pm 0.037 \tag{6.12}$$

where the error is systematic. It is the combination of the uncertainty due to the method used (more details in the next paragraph), due to the limited number of generated MC events and due to the different TCKs, where the contribution of the last is the smallest. Figure 6.17 shows the variation of $\varepsilon_{\phi\mu\mu}^{\text{trigger}}$ for various bins of q^2. The ratios of $\varepsilon_{J/\psi\,\phi}^{\text{trigger}}/\varepsilon_{\phi\mu\mu}^{\text{trigger}}$ in bins of q^2 are summarised in Table 6.9.

The number of $B_s^0 \to J/\psi\,\phi$ candidates in data is sufficient in size to use the TISTOS method to calculate the TOS trigger efficiency ($\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$). The result can be compared to the one when $\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$ is calculated from MC samples instead (using the same method) and to the true $\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$ accessible in the MC samples. An overview of these numbers is given in Table 6.10. These results give confidence to rely on simulation to obtain the TOS trigger efficiency. The difference in efficiency when obtained from data and MC truth, 2.3% (see Table 6.10), is assigned as a systematic uncertainty to each bin.

Table 6.9: Summary of the trigger efficiency ratios in bins of q^2.

q^2 (GeV^2/c^4)	$\varepsilon^{\text{trigger}}_{J/\psi\,\phi}/\varepsilon^{\text{trigger}}_{\phi\mu\mu}$
0.00 – 2.00	1.182 ± 0.042
2.00 – 4.30	1.119 ± 0.040
4.30 – 8.68	1.052 ± 0.037
10.09 – 12.86	0.970 ± 0.034
14.18 – 16.00	0.934 ± 0.033
16.00 – 19.30	0.904 ± 0.032
1.00 – 6.00	1.104 ± 0.039

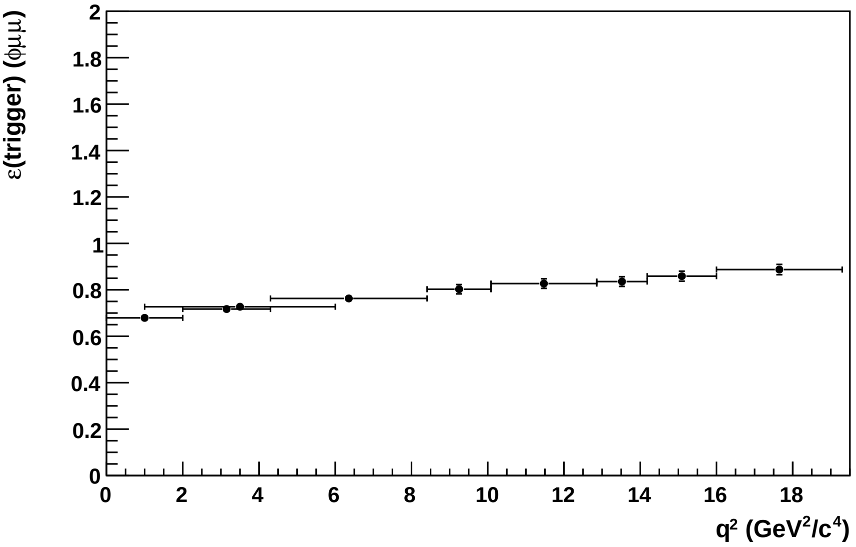


Figure 6.17: Distribution of $\varepsilon^{\text{trig}}_{\phi\mu\mu}$ in bins of q^2.

6.6 Results

Using Equation 6.1 and combining the results of Section 6.4 and 6.5 with the PDG value of $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.93 \pm 0.06) \times 10^{-2}$ [37] gives:

Table 6.10: Overview of the total and the TOS trigger efficiencies for $B_s^0 \to J/\psi\,\phi$ candidates obtained with different methods. The TISTOS method can only be applied on the q^2 bin containing the $B_s^0 \to J/\psi\,\phi$ candidates as this is the only bin with sufficient candidates for the method to work. The good agreement between the values obtained by the TISTOS method (in data and simulation) and from MC truth gives confidence in relying on the TOS efficiencies obtained from MC truth for the other q^2 bins.

	total $\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$	TOS $\varepsilon_{J/\psi\,\phi}^{\text{trigger}}$
MC truth	$90.1 \pm 2.3\%$	$80.1 \pm 2.1\%$
TISTOS method on MC	$96.8 \pm 4.7\%$	$85.0 \pm 5.1\%$
TISTOS method on data	$95.4 \pm 1.2\%$	$82.4 \pm 2.9\%$

$$\begin{aligned}\frac{\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to J/\psi\,\phi)} &= \frac{N_{\phi\mu\mu}}{N_{J/\psi\,\phi}} \times \mathcal{B}(J/\psi \to \mu^+\mu^-) \times \frac{\varepsilon_{J/\psi\,\phi}^{\text{geo}}}{\varepsilon_{\phi\mu\mu}^{\text{geo}}} \times \frac{\varepsilon_{J/\psi\,\phi}^{\text{rec\&sel}}}{\varepsilon_{\phi\mu\mu}^{\text{rec\&sel}}} \times \frac{\varepsilon_{J/\psi\,\phi}^{\text{trigger}}}{\varepsilon_{\phi\mu\mu}^{\text{trigger}}} \\ &= (0.558 \pm 0.070(\text{stat}) \pm 0.043(\text{syst}) \pm 0.006(\mathcal{B})) \times 10^{-3}\end{aligned} \tag{6.13}$$

where the first error is statistical, the second error is systematic and the third error accounts for the uncertainty on $\mathcal{B}(J/\psi \to \mu^+\mu^-)$. The results in bins of q^2 are listed in Table 6.11. The results for the differential branching fractions for $B_s^0 \to \phi\mu^+\mu^-$ with respect to q^2 are shown in Figure 6.18. A summary of the results of the previous sections, including the statistical and systematic uncertainties discussed, is given in Table 6.12. A breakdown of the total systematic error of the result is given in Table 6.13.

6.7 Conclusions

The absolute branching fraction of $B_s^0 \to \phi\mu^+\mu^-$ is obtained by multiplying by the world average branching fraction $\mathcal{B}(B_s^0 \to J/\psi\,\phi) = (1.4 \pm 0.5) \times 10^{-3}$ as provided by the PDG [37]. Thus, LHCb measures the branching fraction for $B_s^0 \to \phi\mu^+\mu^-$ decays to be $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-) = (0.78 \pm 0.10(\text{stat}) \pm 0.06(\text{syst}) \pm 0.28(\mathcal{B})) \times 10^{-6}$. The difference between the relative branching fraction measured by LHCb and CDF is less than 3σ [48].

Rather than using the PDG value for $\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ to obtain the absolute branching fraction for $B_s^0 \to \phi\mu^+\mu^-$ decays the natural choice would be to use a value obtained by LHCb. For example common systematic uncertainties would cancel and increase the overall accuracy of the result. However, a measurement of $\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ by LHCb

Table 6.11: Summary of the branching fraction ratios in bins of q^2. The bin ranging from $1-6\,\mathrm{GeV}^2/c^4$ is of interest from a theoretical point of view (it is 'clean' as it is far away from $0\,\mathrm{GeV}^2/c^4$ and the J/ψ resonance) and is therefore shown in addition to the other bins.

q^2 (GeV^2/c^4)	$\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)/\mathcal{B}(B_s^0 \to J/\psi\,\phi)$ (10^{-3})
$0.00-2.00$	0.058 ± 0.018
$2.00-4.30$	0.045 ± 0.018
$4.30-8.68$	0.096 ± 0.026
$10.09-12.86$	0.092 ± 0.026
$14.18-16.00$	0.094 ± 0.027
$16.00-19.30$	0.105 ± 0.030
$1.00-6.00$	0.095 ± 0.027
full range	0.558 ± 0.082

Table 6.12: Summary of all quantities obtained. Where applicable the statistical and systematic uncertainties are shown.

Term	Value	(stat)	(syst)
$N_{\phi\mu\mu}/N_{J/\psi\,\phi}$	6.931×10^{-3}	$\pm0.867\times10^{-3}$	$\pm0.019\times10^{-3}$
$\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{geo}}_{\phi\mu\mu}$	0.987		±0.005
$\varepsilon^{\mathrm{rec\&sel}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{rec\&sel}}_{\phi\mu\mu}$	1.326		±0.090
$\varepsilon^{\mathrm{trigger}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{trigger}}_{\phi\mu\mu}$	1.038		±0.037

was not publicly available at the time when the analysis was approved and the thesis submitted.

As outlined in Section 3.2.1 this measurement is necessary as it forms the foundation for angular analyses of the $B_s^0 \to \phi\mu^+\mu^-$ decay to be carried out at LHCb in the future. Angular observables offer greater sensitivity to the Wilson coefficients than branching fractions as form factors cancel, but also require larger data sets. The measurement of $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)$ in bins of q^2 made the optimum use of the amount of data available to provide a result, that is of interest from a theoretical point of view with the given uncertainties. The result is in agreement with the SM.

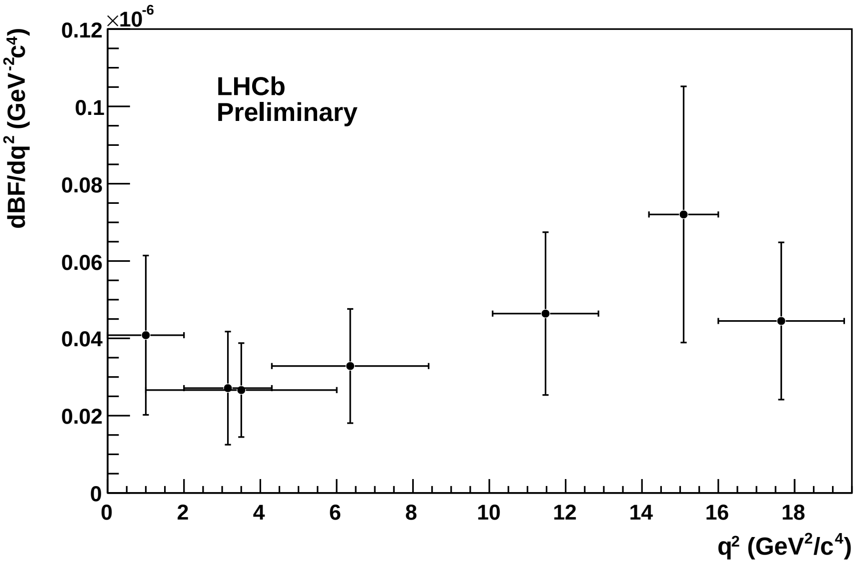


Figure 6.18: Distribution of $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)$ as a function of q^2. The vertical bars on each point are the combined statistical and systematic uncertainties. The horizontal bars indicate the width of the corresponding q^2 region. The region of the entry ranging from $1-6\,\mathrm{GeV}^2/c^4$ is of interest from a theoretical point of view (it is 'clean' as it is far away from $0\,\mathrm{GeV}^2/c^4$ and the J/ψ resonance) and is therefore shown in addition to the other entries.

Table 6.13: Systematic uncertainty contributions to the integrated result for the ratio of branching fractions.

Contribution	systematic uncertainty (%)
$N_{\phi\mu\mu}/N_{J/\psi\,\phi}$	0.14
$\varepsilon^{\mathrm{geo}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{geo}}_{\phi\mu\mu}$	0.35
$\varepsilon^{\mathrm{rec\&sel}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{rec\&sel}}_{\phi\mu\mu}$	4.74
$\varepsilon^{\mathrm{trig}}_{J/\psi\,\phi}/\varepsilon^{\mathrm{trig}}_{\phi\mu\mu}$	2.49
$\mathcal{B}(J/\psi \to \mu^+\mu^-)$	0.01
Total	7.73

Chapter 7

Conclusions

The result of the ratio of branching fractions for the decays $B_s^0 \to \psi'\phi$ and $B_s^0 \to J/\psi\,\phi$ is $0.68 \pm 0.10(\text{stat}) \pm 0.09(\text{syst}) \pm 0.07(\mathcal{B})$. The result of the ratio of branching fractions for the decays $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ is $(0.558 \pm 0.070(\text{stat}) \pm 0.043(\text{syst}) \pm 0.006(\mathcal{B})) \times 10^{-3}$. Both results are compatible with previous measurements by other experiments. The branching fraction ratio of the latter is measured to an accuracy of $\sim 15\%$, making it the world's most precise measurement. Using the measured branching fraction of the decay $J/\psi \to \mu^+\mu^-$ gives a branching fraction of $\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-) = (0.78 \pm 0.10(\text{stat}) \pm 0.06(\text{syst}) \pm 0.28(\mathcal{B})) \times 10^{-6}$.

In order to achieve this level of precision the sources of systematic uncertainties have been understood and the trigger has been studied extensively. The total uncertainty is dominated by the statistical uncertainty. With more collected data this component will decrease and the systematic uncertainties will have to be reduced accordingly in order to maintain this precision.

The work presented in this thesis forms the foundation for a series of measurements to be carried out in the future. This includes the angular analysis of $B_s^0 \to \phi\mu^+\mu^-$ decays, as described in Chapter 3. An upgraded LHCb detector could collect a total integrated luminosity of up to $50\,\text{fb}^{-1}$ in the next decade [88]. This would yield ~ 4000 selected $B_s^0 \to \phi\mu^+\mu^-$ candidates. Assuming a 2% efficiency in determining the flavour of the initial b quark, a total of ~ 80 events could be obtained, that could be used for angular analyses. The decay $B_s^0 \to \phi\mu^+\mu^-$ could then be used to observe effects of New Physics beyond the Standard Model.

Bibliography

[1] MissMJ, 2006. `http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`, accessed 24 November, 2011.

[2] D. Griffiths, “Introduction to Elementary Particles”, John Wiley & Sons, New York, USA, 1987.

[3] F. Halzen and A. Martin, “Quarks & Leptons: An introductory course in modern particle physics”, John Wiley & Sons, New York, USA, 1984.

[4] D. H. Perkins, “Introduction to High Energy Physics; 4th Edition”, Cambridge University Press, Cambridge, 2000.

[5] F. Soomro, B. Cameron, and A. Golutvin, “Radiative decays of B mesons at LHCb”, PhD thesis, Imperial College London, London, 2011. Presented 3 June, 2011.

[6] W. R. Reece and U. Egede, “Exploiting angular correlations in the rare decay $B^0 \to K^{*0}\mu^+\mu^-$ at LHCb”, PhD thesis, Imperial College London, London, 2010. Presented on 14 May, 2010.

[7] A. Powell, N. Harnew, and G. Wilkinson, “Measuring the CKM Angle γ with $B^\pm \to DK^\pm$ Decays at LHCb and a Determination of the $D \to K\pi\pi\pi$ Coherence Factor at CLEO-c.”, PhD thesis, Oxford University, Oxford, 2009. Presented on 22 May 2009.

[8] S. Weinberg, “A model of leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266.

[9] G. S. Guralnik, C. R. Hagen, and T. W. Kibble, “Global Conservation Laws and Massless Particles”, *Physical Review Letters* **13** (November, 1964) 585–587.

[10] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Physical Review Letters* **13** (October, 1964) 508–509.

[11] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Physical Review Letters* **13** (August, 1964) 321–323.

[12] H. Yukawa, “On the Interaction of Elementary Particles. I *”, *Progress of Theoretical Physics Supplement* **1** (1955) 1–10.

[13] N. Cabibbo, “Unitary symmetry and leptonic decays”, *Phys. Rev. Lett.* **10** (1963), no. 12, 531–533.

[14] M. Kobayashi and T. Maskawa, “CP-Violation in the Renormalizable Theory of Weak Interaction”, *Progress of Theoretical Physics* **49** (1973), no. 2, 652–657.

[15] L. Wolfenstein, “Parametrization of the Kobayashi-Maskawa Matrix”, *Phys. Rev. Lett.* **51** (Nov, 1983) 1945–1947.

[16] CKMfitter Group, J. Charles *et al.*, “CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories”, *Eur. Phys. J.* **C41** (2005) 1–131. Updated results and plots available at: `http://ckmfitter.in2p3.fr`.

[17] Heavy Flavor Averaging Group, D. Asner *et al.*, “Averages of b-hadron, c-hadron, and τ-lepton Properties”, 2010. [TT arXiv:1010.1589].

[18] A. E. Noether, “Invariante Variationsprobleme”, *Gott. Nachr.* **1918** (1918) 235–257.

[19] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, “Experimental test of parity conservation in beta decay”, *Phys. Rev.* **105** (Feb, 1957) 1413–1415.

[20] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, “Evidence for the 2π Decay of the K^0 Meson”, *Phys. Rev. Lett.* **13** (1964) 138–140.

[21] BABAR collaboration, B. Aubert *et al.*, “Observation of CP violation in the B^0-meson system”, *Phys.Rev.Lett.* **87** (2001) 091801. [TT hep-ex/0107013].

[22] LHCb collaboration, R. Aaij *et al.*, “Evidence for *CP* violation in time-integrated $D^0 \to h^- h^+$ decay rates”, 2011. [TT arXiv:1112.0938].

[23] M. L. Perl *et al.*, “Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation”, *Phys. Rev. Lett.* **35** (Dec, 1975) 1489–1492.

[24] S. W. Herb, D. C. Hom, L. M. Lederman, J. C. Sens, H. D. Snyder, J. K. Yoh, J. A. Appel, B. C. Brown, C. N. Brown, W. R. Innes, K. Ueno, T. Yamanouchi, A. S. Ito, H. Jöstlein, D. M. Kaplan, and R. D. Kephart, “Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions”, *Physical Review Letters* **39** (Aug., 1977) 252–255.

[25] A. J. Buras, “Minimal flavor violation”, *Acta Phys.Polon.* **B34** (2003) 5615–5668. [TT hep-ph/0310208].

[26] M. Fukugita, "Barygenesis without grand unification", *Physics Letters B* **174** (1986), no. 1, 45–47.

[27] ATLAS collaboration, "ATLAS: Technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN", LHC Tech. Proposal, CERN, Geneva, 1994.

[28] CMS collaboration, "Technical proposal", LHC Tech. Proposal, CERN, Geneva, 1994.

[29] A. J. Buras, "Minimal flavour violation and beyond: Towards a flavour code for short distance dynamics", *Acta Phys.Polon.* **B41** (2010) 2487–2561. [TT arXiv:1012.1447].

[30] A. J. Buras, M. V. Carlucci, S. Gori, and G. Isidori, "Higgs-mediated FCNCs: Natural Flavour Conservation vs. Minimal Flavour Violation", *JHEP* **10** (2010) 009. [TT arXiv:1005.5310].

[31] M. Drees, R. Godbole, and P. Roy, "Theory and phenomenology of sparticles: An account of four-dimensional n=1 supersymmetry in high energy physics", World Scientific, Singapore, 2004.

[32] J. R. Ellis, T. Falk, G. Ganis, K. A. Olive, and M. Srednicki, "The CMSSM parameter space at large $\tan\beta$", *Phys.Lett.* **B510** (2001) 236–246. [TT hep-ph/0102098].

[33] LHCb collaboration, "Letter of Intent for the LHCb Upgrade", CERN, March, 2011. Tech. Rep., CERN-LHCC-2011-001.

[34] J. R. Ellis, K. A. Olive, and V. C. Spanos, "On the interpretation of $B_s^0 \to \mu^+\mu^-$ in the CMSSM", *Phys.Lett.* **B624** (2005) 47–59. [TT hep-ph/0504196].

[35] K. G. Wilson and W. Zimmermann, "Operator product expansions and composite field operators in the general framework of quantum field theory", *Commun. Math. Phys.* **24** (1972) 87–106.

[36] E. Fermi, "An attempt of a theory of beta radiation. 1", *Z. Phys.* **88** (1934) 161–177.

[37] Particle Data Group, K. Nakamura *et al.*, "Review of particle physics", *J.Phys.G* **G37** (2010) 075021.

[38] C. Bobeth, M. Misiak, and J. Urban, "Photonic penguins at two loops and m(t) dependence of $\mathcal{B}(B \to X(s)l^+l^-)$", *Nucl.Phys.* **B574** (2000) 291–330. [TT hep-ph/9910220].

[39] C. Bobeth, A. J. Buras, and T. Ewerth, "$\bar{B} \to X(s)l^+l^-$ in the MSSM at NNLO", *Nucl.Phys.* **B713** (2005) 522–554. [TT hep-ph/0409293].

[40] W. Altmannshofer, P. Paradisi, and D. M. Straub, "Model-Independent Constraints on New Physics in $b \to s$ Transitions", 2011. [TT arXiv:1111.1257].

[41] F. Kruger and J. Matias, "Probing new physics via the transverse amplitudes of $B^0 \to K^{*0}(\to K^-\pi^+)l^+l^-$ at large recoil", *Phys.Rev.* **D71** (2005) 094009. [TT hep-ph/0502060].

[42] C. Q. Geng and C. C. Liu, "Study of $B_s^0 \to (\eta, \eta', \phi)l\bar{l}$ decays", *J. Phys.* **G29** (2003) 1103–1118. [TT hep-ph/0303246].

[43] A. J. Buras and R. Fleischer, "Quark mixing, CP violation and rare decays after the top quark discovery", *Adv.Ser.Direct.High Energy Phys.* **15** (1998) 65–238. [TT hep-ph/9704376]. To appear in Heavy Flavours II, World Scientific (1997), eds. A.J. Buras and M. Linder.

[44] LHCb collaboration, R. Aaij, C. Abellan Beteta, B. Adeva, M. Adinolfi, C. Adrover, *et al.*, "Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$", 2011. [TT arXiv:1112.3515].

[45] LHCb collaboration, "Differential branching fraction and angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay", CERN, 2012. LHCb-CONF-2012-008.

[46] P. Ball, R. Fleischer, G. Tartarelli, P. Vikas, G. Wilkinson, *et al.*, "B Decays at the LHC", 2000. [TT hep-ph/0003238].

[47] C. Bobeth, G. Hiller, and G. Piranishvili, "CP Asymmetries in $\bar{B} \to \bar{K}^*(\to \bar{K}\pi)\bar{l}l$ and Untagged $\bar{B}_s^0, B_s^0 \to \phi(\to K^+K^-)\bar{l}l$ Decays at NLO", *JHEP* **07** (2008) 106. [TT arXiv:0805.2525].

[48] CDF collaboration, T. Aaltonen *et al.*, "Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$", *Phys.Rev.Lett.* **107** (2011) 201802. [TT arXiv:1107.3753].

[49] Gargamelle Neutrino collaboration, F. Hasert *et al.*, "Observation of Neutrino Like Interactions without Muon or Electron in the Gargamelle Neutrino Experiment", *Nucl.Phys.* **B73** (1974) 1–22.

[50] G. Arnison *et al.*, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $s = 540\,\mathrm{GeV}$", *Physics Letters B* **122** (1983), no. 1, 103 – 116.

[51] ALICE collaboration, “ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC”, LHC Tech. Proposal, CERN, Geneva, 1995.

[52] LHCb collaboration, A. Alves *et al.*, “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005.

[53] C. G. Blanks and U. Egede, “V^0 production ratios at LHCb and the alignment of its RICH detectors”, PhD thesis, Imperial College London, London, 2011. Presented 14 October, 2011.

[54] LHCb collaboration, S. Amato *et al.*, “LHCb technical proposal”, CERN, 1998. CERN-LHCC-98-04.

[55] P. M. Bjørnstad, “Performance of the LHCb vertex locator”, *Journal of Instrumentation* **6** (2011), no. 12, C12024.

[56] LHCb collaboration, S. Borghi, “Performance of the Tracking System at the LHCb Experiment”, LHCb-TALK-2010-094, Presented at the *35th International Conference on High Energy Physics*, Paris, France, July, 2010.

[57] LHCb collaboration, A. Powell, “Particle ID at LHCb”, CERN, 2011. LHCb-PROC-2011-008.

[58] A. Powell, 2011. `https://lbtwiki.cern.ch/bin/view/RICH/RichPIDPerformancePlots`, accessed 31 December, 2011.

[59] LHCb collaboration, G. Lanfranchi *et al.*, “The Muon Identification Procedure of the LHCb Experiment for the First Data”, CERN, 2009. LHCb-PUB-2009-013.

[60] LHCb collaboration, A. Sarti, S. Furcas, G. Lanfranchi, and M. Palutan, “Calibration Strategy and Efficiency measurement of the Muon Identification procedure at LHCb”, CERN, 2010. LHCb-PUB-2010-002.

[61] LHCb collaboration, W. Bonivento and F. Dettori, “Search for the $D^0 \to \mu^+\mu^-$ decay with $922\,\mathrm{pb}^{-1}$ at LHCb”, CERN, 2011. LHCb-ANA-2011-017.

[62] The LHCb collaboration, B. Adeva *et al.*, “Roadmap for selected key measurements of LHCb”, CERN, 2009. LHCB-PUB-2009-029.

[63] A. Perez-Calero and H. Ruiz, “The μ+track alley of the LHCb HLT”, CERN, October, 2008. Tech. Rep., CERN-LHCb-2008-075.

[64] LHCb collaboration, R. Aaij and J. Albrecht, “Muon triggers in the HLT of LHCb”, CERN, 2011. LHCb-PUB-2011-017.

[65] S. Amato *et al.*, “HLT1 Muon Alley Description”, CERN, November, 2008. Tech. Rep., CERN-LHCb-2008-058.

[66] J. A. Hernando Morata *et al.*, “The Hadron Alley of the LHCb HLT”, CERN, April, 2010. Tech. Rep., CERN-LHCb-2009-034.

[67] LHCb collaboration, K. Senderowska *et al.*, “HLT1 Electromagnetic Alley”, CERN, 2009. LHCb-PUB-2009-001.

[68] LHCb collaboration, M. Williams *et al.*, “The HLT2 topological lines”, CERN, 2011. LHCb-PUB-2011-002.

[69] LHCb collaboration, V. Gligorov, C. Thomas, and M. Williams, “The HLT inclusive B triggers”, CERN, 2011. LHCb-PUB-2011-016.

[70] LHCb collaboration, B. Souza de Paula, “Studies on Systematic Effects of the Trigger on Flavour Tagging at the Generator Level”, CERN, 2009. LHCb-PUB-2009-014.

[71] H. Dijkstra, N. Tuning, and N. Brook, “Some Remarks on Systematic Effects of the Trigger and Event Generator Studies”, CERN, December, 2003. Tech. Rep., LHCb-2003-157.

[72] LHCb collaboration, “Measurement of $\mathcal{B}(B_s^0 \to \psi(2S)\phi)/\mathcal{B}(B_s^0 \to J/\psi\phi)$”, CERN, 2011. LHCb-CONF-2011-014.

[73] M. Needham, “Clone Track Identification using the Kullback-Liebler Distance”, CERN, January, 2008. Tech. Rep., CERN-LHCb-2008-002.

[74] 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.

[75] J. Gaiser, “Charmonium spectroscopy from radiative decays of the J/ψ and ψ'”, PhD thesis, SLAC, Stanford, 1982. SLAC-R-0255.

[76] T. du Pree, M. H. M. Merk, and H. G. Raven, “Search for a Strange Phase in Beautiful Oscillations”, PhD thesis, Vrije Universiteit, Amsterdam, 2010. Presented on 22 Oct 2010.

[77] CDF collaboration, A. Abulencia *et al.*, “Observation of $B_s^0 \to \psi(2S)\phi$ and measurement of ratio of branching fractions $\mathrm{B}(B_s^0 \to \psi(2S)\phi)$ / $\mathrm{B}(B_s^0 \to J/\psi\phi)$”, *Phys. Rev. Lett.* **96** (2006) 231801. [TT hep-ex/0602005].

[78] D0 collaboration, V. M. Abazov *et al.*, "Relative rates of B meson decays into $\psi(2S)$ and J/ψ mesons", *Phys. Rev. D* **79** (June, 2009) 111102.

[79] BABAR collaboration, B. Aubert *et al.*, "BABAR collaboration, Measurement of branching fractions for exclusive B decays to charmonium final states", *Phys. Rev.* **D65** (2002) 032001.

[80] LHCb collaboration, "Measurement of the ratio of branching fractions for $B_s^0 \to \phi\mu\mu$ and $B_s^0 \to J/\psi\phi$", CERN, 2012. LHCb-CONF-2012-003.

[81] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, "Comparative study of the decays $B^0 \to (K, K^{*0})l^+l^-$ in the standard model and supersymmetric theories", *Phys. Rev. D* **61** (Mar, 2000) 074024.

[82] D. Melikhov, N. Nikitin, and S. Simula, "Probing the standard model and its extensions in rare exclusive $b \to sl^+l^-$ decays", *Nucl. Phys. Proc. Suppl.* **75B** (1999) 97–99.

[83] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, "The Decay $\bar{B} \to \bar{K}l^+l^-$ at Low Hadronic Recoil and Model-Independent $\Delta B = 1$ Constraints", *JHEP* **1201** (2012) 107. [TT arXiv:1111.2558].

[84] D. Becirevic and E. Schneider, "On transverse asymmetries in $B^0 \to K^{*0}l^+l^-$", *Nucl. Phys.* **B854** (2012) 321–339. [TT arXiv:1106.3283].

[85] I. Belyaev *et al.*, "Handling of the generation of primary events in Gauss, the LHCb simulation framework", in "Nuclear Science Symposium Conference Record (NSS/MIC)," p. 1155–1161. IEEE, 2010.

[86] M. Clemencic *et al.*, "The LHCb Simulation Application, Gauss: Design, Evolution and Experience", *Journal of Phys: Conf. Ser.* **331** (2011) 032023.

[87] LHCb collaboration, "Observation of double J/ψ production in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV", CERN, 2011. LHCb-CONF-2011-009.

[88] LHCb collaboration, "Letter of Intent for the LHCb Upgrade", CERN, 2011. CERN-LHCC-2011-001.

Appendix A

List of Acronyms

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC ApparatuS
BDT	Boosted Decision Tree
CDF	Collider Detector at Fermilab
CERN	European Organization for Nuclear Research
CMS	Compact Muon Solenoid
CMSSM	Constrained Minimal Supersymmetric Standard Model
ECAL	Electromagnetic Calorimeter
EM	Electromagnetic
FCNC	Flavour Changing Neutral Current
FPGA	Field-Programmable Gate Array
GEM	Gas Electron Multiplier
GWS	Glashow-Weinberg-Salam
HCAL	Hadron Calorimeter
HEP	High Energy Physics
HPD	Hybrid Photon Detector
IP	Impact Parameter
IT	Inner Tracker
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty Experiment

MC	Monte Carlo
MFV	Minimal Flavour Violation
MWPC	Multi-Wire Proportional Chamber
NP	New Physics
OPE	Operator Product Expansion
OT	Outer Tracker
PID	Particle Identification
PS	Proton Synchrotron
PSD	Pre-Shower Detector
PSB	Proton Synchrotron Booster
PV	Primary Vertex
QCD	Quantum Chromodynamics
RICH	Ring Imaging Cherenkov Detector
SPD	Scintillating Pad Detector
SPS	Super Proton Synchrotron
SM	Standard Model
SUSY	Supersymmetry
SV	Secondary Vertex
TT	Tracker Turicensis
UT	Unitarity Triangle
VELO	Vertex Locator

Appendix B

Optimisation of cuts for $B_s^0 \to \phi\mu^+\mu^-$

In order to select $B_s^0 \to \phi\mu^+\mu^-$ signal events and reject background events a cut-based selection was developed. It consists of several cuts on kinematic variables. The aim was to keep the number of cuts to a minimum in order to reduce uncertainties and not to favour certain regions of the q^2 spectrum. The latter also ensures that the decays $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi'\phi$ are selected with the same set of cuts. The procedure of finding the optimum cut value for the cuts and the results are presented in this section.

To find the optimal cut value for a cut on a certain variable the distributions of signal (S) and background (B) events of that variable were compared. MC simulated events for $B_s^0 \to \phi\mu^+\mu^-$ were used to obtain distributions for S. Events lying inside background regions in data were used to obtain distributions for B (e.g. events lying inside non-peaking regions). As the figure of merit $S/\sqrt{B}$ was used. For variables with a large discrimination power between S and B the figure of merit will peak sharply. When the distribution of the figure of merit shows a 'plateau'-like feature the optimum cut value can be chosen to be one of the values along the range of the 'plateau'.

Several kinematic variables describing the daughter particles ($K^\pm$ and $\mu^\pm$), variables describing the B_s^0 decay vertex and variables related to the finite B_s^0 flight distance were considered. The optimisation procedure was as follows:

1. A broad mass window of $\pm 1\,\mathrm{GeV}/c^2$ around the B_s^0 mass was applied on both the S and the B sample
2. For each variable the distributions of S and B were normalised to unit area and the figure of merit ($S/\sqrt{B}$) was calculated
3. From the position of the peak of the figure of merit an initial cut value was determined for each variable

To confirm the cut value of a certain variable all initial cuts were applied on the S and B samples except for the given variable. The figure of merit was recalculated using the cut versions of S and B. If the position of the peak (and therefore the optimal cut value) was significantly different to the previous outcome the initial cut value was adjusted. This was repeated for each variable.

The distributions of S, B and $S/\sqrt{B}$ for the main variables are shown in Figure B.1. The final choice of cuts and the optimised cut values are summarised in Table B.1(a). In most cases the chosen cut value does not correspond to the exact position of the maxima of the figure of merit as indicated in Figure B.1. This is either due to a flat distribution of the figure of merit or due to manual adjustments (as outlined earlier). In future the selection can be improved by using more advanced techniques, such as boosted decision tree.

In Figure B.1 the distributions for S were obtained from a MC sample, which had one specific trigger configuration from 2010 simulated. The events for B were obtained from 2010 data, where several different trigger configurations were used. This is the reason for small discrepancies observed in the various distributions - for example the different starting values of the two distributions for B_s^0 τ in Figure B.1(f). However, the effect of this on the chosen optimal cut values is negligible.

For the stripping line `StrippingBs2MuMuPhi` the optimum cut values derived above were loosened for reasons explained in Section 4.4.3. The loosened values used for this stripping line are summarised Table B.1(b).

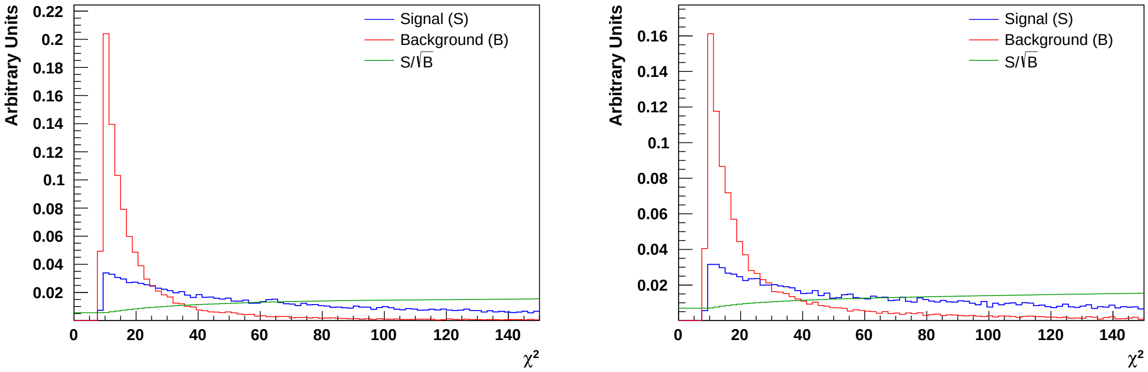
Table B.1: The cuts chosen to select $B_s^0 \to \phi\mu^+\mu^-$ events (a) and the loosened versions of the same cuts (b), that are used in the stripping line `StrippingBs2MuMuPhi`.

(a)

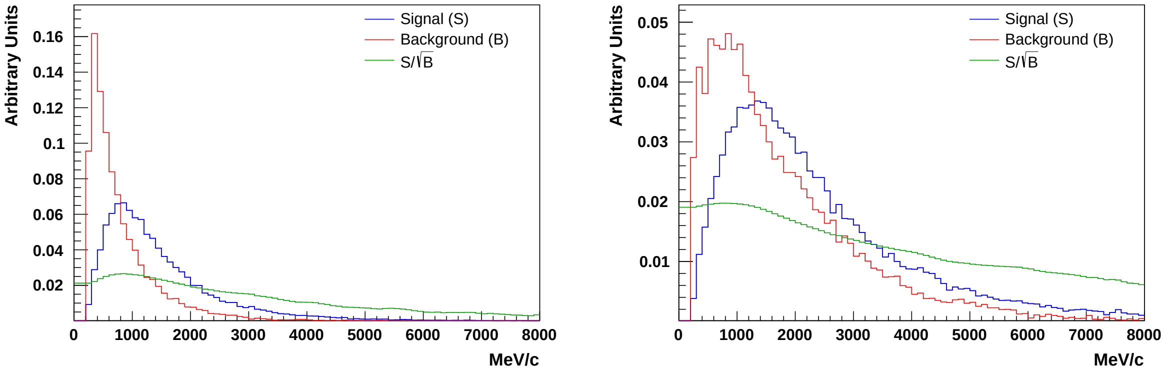
Variable	Cut value
$K^\pm$, $\mu^\pm$ χ^2_{IP}	> 16
$K^\pm$ $\Delta\log\mathcal{L}$	> 5
$K^\pm$, $\mu^\pm$ p_{T}	$> 300\,\mathrm{MeV}/c$
B_s^0 τ	$> 0.4\,\mathrm{ps}$
B_s^0 χ^2_{VX}	< 15
B_s^0 χ^2_{IP}	< 6

(b)

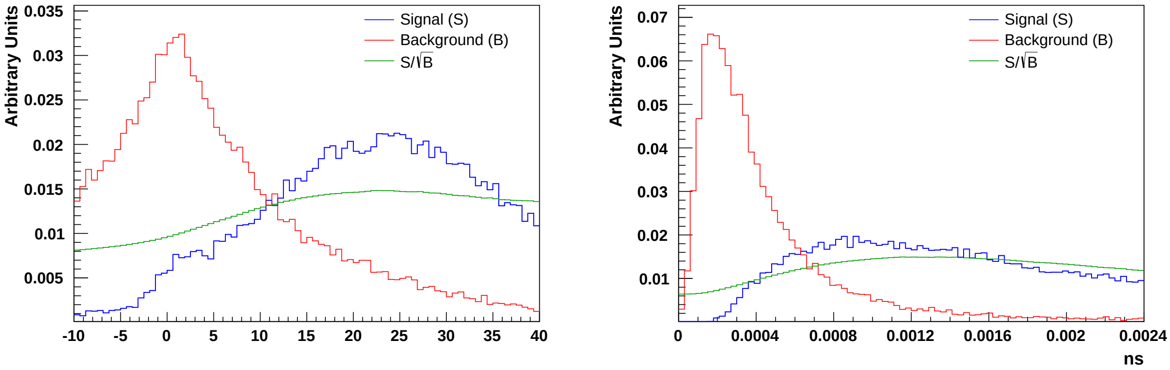
Variable	Cut value
$K^\pm$, $\mu^\pm$ χ^2_{IP}	> 9
$K^\pm$ $\Delta\log\mathcal{L}$	> 0
B_s^0 τ	$> 0.2\,\mathrm{ps}$
B_s^0 χ^2_{VX}	< 40
B_s^0 χ^2_{IP}	< 9



(a) K^+ χ^2_{IP} (b) μ^+ χ^2_{IP}

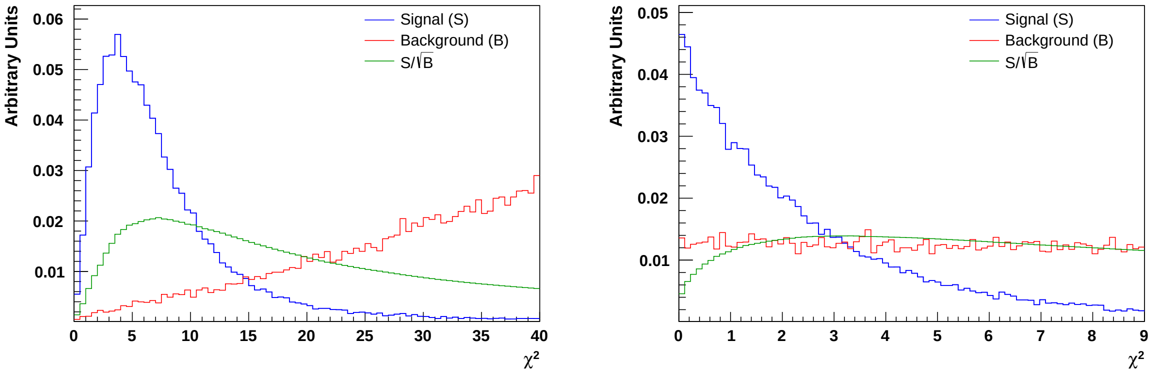


(c) K^+ p_{T} (d) μ^+ p_{T}



(e) K^+ $\Delta \log \mathcal{L}$ (f) B_s^0 τ

Figure B.1: (continued on the next page)



(g) B_s^0 χ^2_{VX} (h) B_s^0 χ^2_{IP}

Figure B.1: The distributions of signal (S, in blue) and background (B, in red) events for different variables (a)-(h). For S events from $B_s^0 \to \phi\mu^+\mu^-$ simulations were used. For B events from data were used, that lie in the background regions. Indicated on each plot is also the figure of merit $S/\sqrt{B}$ (in green), from which the final offline cut value was determined. All distributions are normalised to unit area as only the relative shapes of the distributions are of importance.

Appendix C

The TIS and TOS lines of each trigger level used for the TISTOS method. The order of the lines corresponds to the number of candidates selected by each line. The superscript $^{1}(^{2})$ denotes the lines, that select the majority of candidates ($> 90\%$) for the TOS (TIS) category:

L0 lines (TIS and TOS):
`Muon`$^{1\,2}$
`DiMuon`$^{1\,2}$
`MuonHigh`$^{1\,2}$
`Hadron`

Additional L0 TIS lines:
`Photon`^{2}
`Electron` ^{2}

HLT1 lines (TIS and TOS):
`TrackMuon`$^{1\,2}$
`SingleMuonNoIPL0`$^{1\,2}$
`TrackAllL0`$^{1\,2}$
`DiMuonNoIP2L0`^{1}
`DiMuonNoIPL0Di`^{1}
`SingleMuon4BsMuMu`
`SingleMuonNoIPL0HighPT`
`SingleMuonIPCL0`
`MuTrack`
`DiMuonNoIPL0Seg`
`DiMuonIPCL0Seg`
`DiMuonIPC2L0`
`DiMuonIPCL0Di`

Additional HLT1 TIS lines:
`DiHadron`^{2}
`SingleHadron`^{2}
`TrackPhoton`^{2}
`SingleElectronWithIP`
`ElectronTrackNoIP`
`PhotonDiTrackFromEle`

HLT2 lines (TIS and TOS):
`DiMuonUnbiasedJPsi`^{1}
`DiMuonBiasedJPsi`^{1}
`BiasedDiMuonMass`^{1}
`BiasedDiMuonIP`^{1}
`IncPhi`$^{1\,2}$
`MuTrack`$^{1\,2}$
`PromptJPsi`^{1}
`PromptJPsiHighPT`
`TopoOSTF3Body`^{2}
`TopoOSTF2Body`^{2}
`UnbiasedDiMuon`
`TopoOSTF4Body`^{2}
`SingleMuon`
`MuTrackNoIP`
`IncPhiSidebands`

Additional HLT2 TIS lines:
`CharmOSTF2Body`
`CharmOSTF3Body`^{2}
`IncDiProton`
`Bs2PhiGamma`
`DiMuonUnbiasedBmm`
`DiMuonSameSign`
`CharmD02PiPiForD02MuMu`
`SingleTFElectron`
`UnbiasedTFDiElectronLowMass`
`UnbiasedTFPsi2ee`
`UnbiasedTFDiElectron`
`TFElectronPlusTrack`
`TFBs2JpsieePhiSignal`
`IncPhiTrackFit`
`IncPhiRobust`
`DiMuonUnbiasedPsi2S`
`CharmOSTF3BodyWideMass`
`BiasedTFDiElectron`

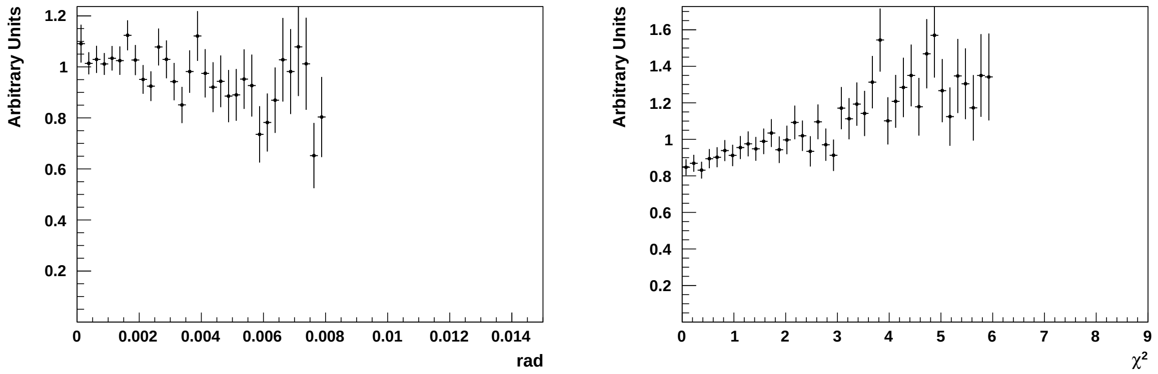
Appendix D

Comparison of data and simulation using $B_s^0 \to J/\psi\,\phi$

The large number of $B_s^0 \to J/\psi\,\phi$ candidates in data enable accurate comparisons between data and simulation. The offline cut efficiencies evaluated using data and simulation are given in Table D.1. The ratios of distributions between data and simulation for various kinematic variables are given in Figure D.1. Variables for which the description slightly varies between data and the MC samples used are B_s^0 χ^2_{IP}, B_s^0 χ^2_{VX} and $K^\pm$ $\Delta\log\mathcal{L}$. This difference is taken into account for in the analysis by assigning conservative systematic uncertainties for this effect and using look-up tables for PID values, that are based on PID performaces measured in data.

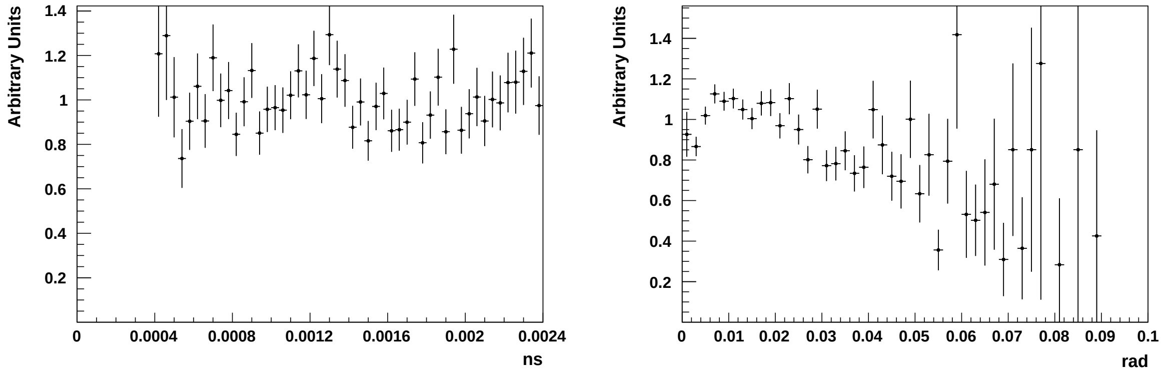
Table D.1: The offline cut selection efficiencies determined for $B_s^0 \to J/\psi\,\phi$ candidates in data and simulation (using the MC11a $B_s^0 \to J/\psi\,\phi$ sample). The efficiency was calculated according to Equation 6.2.

Offline cut		Data	Simulation
$K^\pm$ χ^2_{IP}	> 16	0.943 ± 0.002	0.941 ± 0.002
$\mu^\pm$ χ^2_{IP}	> 16	0.954 ± 0.002	0.948 ± 0.002
$K^\pm$ $\Delta\log\mathcal{L}$	> 5	0.835 ± 0.003	0.921 ± 0.003
B_s^0 τ	$> 0.4\,\text{ps}$	0.996 ± 0.001	0.997 ± 0.001
B_s^0 χ^2_{IP}	< 6	0.913 ± 0.003	0.945 ± 0.003
B_s^0 χ^2_{VX}	< 15	0.895 ± 0.003	0.944 ± 0.002



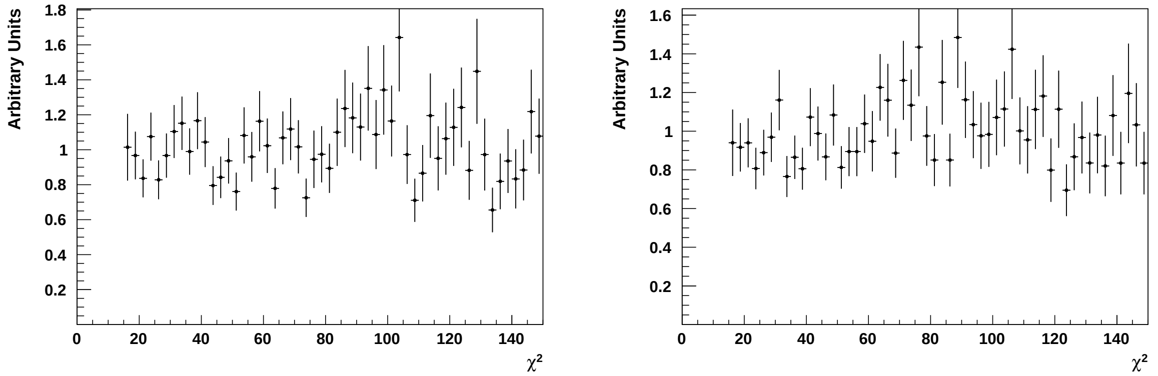
(a) B_s^0 θ

(b) B_s^0 χ^2_{IP}



(c) B_s^0 τ

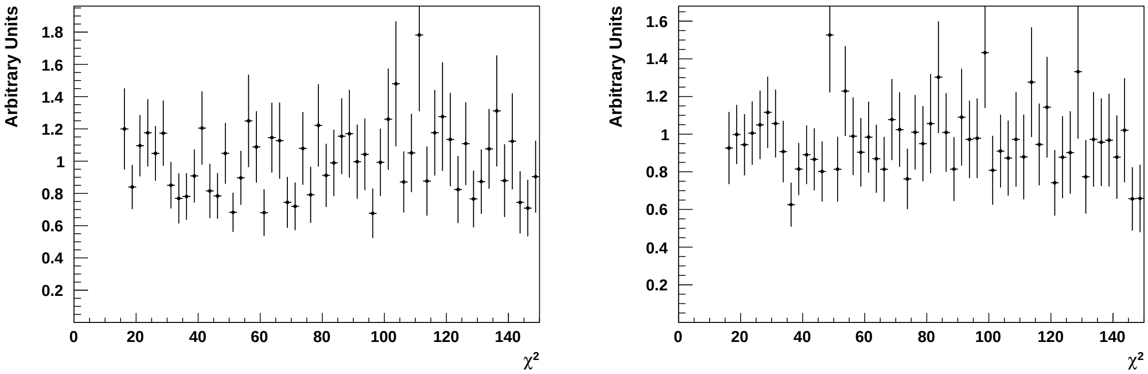
(d) $\theta_{K^+K^-}$



(e) K^- χ^2_{IP}

(f) K^+ χ^2_{IP}

Figure D.1: (continued on the next page)



(g) μ^- χ^2_{IP} (h) μ^+ χ^2_{IP}

Figure D.1: The ratios of various kinematic variables for $B_s^0 \to J/\psi\,\phi$ candidates between data and simulation (a)-(h).