THE UNIVERSITY *of* EDINBURGH

Measurements of charmless B_s^0 meson decays at LHCb.

Adam Morris

Doctor of Philosophy
The University of Edinburgh
2017

Abstract

Using 3 fb^{-1} of proton–proton collisions, collected at centre-of-mass energies of $\sqrt{s} = 7$ and 8 TeV by the LHCb detector, several measurements of charmless B_s^0 meson decays are made. A search is also performed for a highly suppressed B^0 decay.

First, the branching fraction of the $B_s^0 \to \phi\phi$ decay is measured to be

$$\mathcal{B}(B_s^0 \to \phi\phi) = (1.84 \pm 0.05(\text{stat}) \pm 0.07(\text{syst}) \pm 0.11(f_s/f_d) \pm 0.12(\text{norm})) \times 10^{-5},$$

where the third and fourth uncertainties arise from the fragmentation fraction f_s/f_d and the branching fraction of the normalisation mode. This represents a factor of five reduction in the statistical uncertainty compared to the previous best measurement.

An upper limit on the branching fraction of the mode $B^0 \to \phi\phi$ is set at

$$\mathcal{B}(B^0 \to \phi\phi) < 2.8 \times 10^{-8} \ (90\,\%\ \text{CL}).$$

This is a factor of seven improvement over the previous best measurement.

An amplitude analysis of the $B_s^0 \to \phi K^+ K^-$ decay is performed, wherein first observations of the decay modes $B_s^0 \to \phi f_2'(1525)$ and $B_s^0 \to \phi\phi(1680)$ are made. The branching fraction of the $B_s^0 \to \phi f_2'(1525)$ decay is measured to be

$$\mathcal{B}(B_s^0 \to \phi f_2'(1525)) = \\ (1.63 \pm 0.18(\text{stat}) \pm 0.12(\text{syst}) \pm 0.29(\text{model}) \pm 0.17(\text{norm})) \times 10^{-6},$$

where the 'model' uncertainty arises from the choice of amplitude model. The longitudinal polarisation fraction of the decay $B_s^0 \to \phi f_2'(1525)$ is measured to be

$$\mathcal{F}_0 = (86.6 \pm 3.4 \pm 0.8 \pm 2.0\,(\text{model}))\,\%.$$

Lay summary

According to the current best theory of particle physics, the Standard Model, the universe should be equal parts matter and antimatter. However, it is dominated by matter. One of the key conditions in resolving this mystery is the difference in behaviour between particles and antiparticles, referred to as *CP* violation. This was discovered in the 1960s in the decays of particles called kaons, but the amount of *CP* violation in the Standard Model is far too small to account for the observed matter–antimatter asymmetry. This motivates the search for additional sources of *CP* violation that do not fit in with the Standard Model. One of the ways to do this is by studying the decays of composite particles containing a heavy and short-lived fundamental particle called the 'bottom quark', or b quark. These particles are collectively called b hadrons.

The LHCb experiment, at the Large Hadron Collider at CERN, was built specifically to detect b hadrons. The research presented in this thesis uses data collected by LHCb in 2011 and 2012 to analyse decays of a b hadron called the B_s^0 meson. Previous measurements of *CP* violation in these types of decay are already more precise than the theoretical predictions. It is therefore not clear whether the measurements agree perfectly with the Standard Model, or if there is a small effect from undiscovered physics. The methods of calculating these predictions rely, in part, on experimental input. The measurements in this thesis will help improve their precision.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own, except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in Ref. [1].

(*Adam Morris, 2017*)

Acknowledgements

I would like to thank my supervisor Matthew Needham for his excellent guidance and feedback throughout the course of my PhD. The Edinburgh PPE group has been invaluable in helping me perform my research. In particular, I am grateful to Greig Cowan, Pete Clarke, Steve Playfer and Yanyan Gao for sharing their expertise in the art of amplitude analysis.

I would like to thank CERN, the LHCb Collaboration, the University of Edinburgh, SUPA and the STFC for providing the infrastructure, training, resources and funding to make my doctoral research possible in the first place. I would also like to express my gratitude to the developers of ROOT, RooFit, RooStats, EvtGen, Pythia, Geant4, RapidFit, RapidSim and the suite of LHCb packages for providing the crucial software used to analyse and simulate the data used in my research.

For their specific contributions to the $B_s^0 \to \phi\phi$ analysis in Chapter 5, I would like to thank Haofei Luo for providing the preselected data samples and training the multivariate algorithms, Laurence Carson for calculating the particle identification efficiency and the angular acceptance systematic and for sharing the workload in getting the paper past Collaboration and Editorial Board review, Matthew Needham for calculating the tracking efficiency systematic, hadronic interaction systematic and for his major contributions to the paper, and Sean Benson for providing code to use the RooStats implementation of the CL_s method.

I am grateful to the School of Physics and Astronomy at the University of Birmingham for providing an excellent undergraduate education and, in particular, the particle physics group for equipping me with the necessary skills to perform research in experimental particle physics. Finally, I would like to thank my family, and friends from outside the world of physics, for generally being supportive and tolerating my attempts to explain what I do.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The origin of the asymmetry between the abundance of matter and antimatter in the universe is an unsolved problem in physics. In the big bang, an equal amount of matter and antimatter should have been produced, yet our universe is dominated by matter [2]. Three conditions [3] are necessary to produce this imbalance: baryon/lepton number violation, *CP* violation and a departure from thermal equilibrium. Given the experimental constraints on baryon and lepton number violation, the amount of *CP* violation in the Standard Model is many orders of magnitude too small to account for the observed matter–antimatter asymmetry [4, 5]. This motivates the search for sources of *CP* violation beyond the Standard Model. One of the main areas where this search is performed is 'B physics': the study of hadrons containing b quarks.

The primary goal of the LHCb experiment is to look for new physics in b hadron decays. This programme of research builds on the studies of B^0 and B^+ mesons conducted at the BaBar and Belle experiments [6], which were e^+e^- collider experiments. As a dedicated B-physics experiment at a hadron collider, LHCb can also study B_s^0 mesons, B_c^+ mesons and b baryons with unprecedented precision.

This thesis presents two studies of 'charmless' B_s^0 meson decays, *i.e.* fully hadronic decays which do not contain c quarks. Specifically, the decays $B_s^0 \to \phi\phi$ and $B_s^0 \to \phi K^+ K^-$ are studied, which are examples of $b \to s\bar{s}s$ transitions. In the Standard Model, these decays are forbidden at tree-level and proceed via loop diagrams. This enhances their sensitivity to new physics [7]. Since the daughter particles of charmless b hadron decays are considerably lighter than the parent particle, the energy released is typically large. This permits the use of the

perturbative QCD and QCD factorisation techniques for theoretical calculations of these decays [8]. QCD factorisation calculations rely on input from experiment for the long-range non-perturbative parts. The results of these calculations have large uncertainties and are often less precise than the measured values. In the absence of very large effects from new physics, studies of *CP* violation in charmless b hadron decays require more precise theoretical predictions with which to compare measurements. The analyses presented in Chapters 5 and 6 provide measurements which may be used to improve the precision of these calculations.

Chapter 2 of this thesis gives an overview of the theoretical background and motivation behind the work presented in Chapters 5 and 6. The topics of the Standard Model of particle physics, *CP* violation, hadronic $b \to s$ transitions and QCD calculations of hadronic charmless B decays are covered. Chapter 3 describes the LHC and the LHCb experiment, covering the detector, trigger and data processing software. Chapter 4 presents the general aspects of the event selection, common to both of the analyses presented in this thesis. Chapter 5 details a measurement of the $B^0_s \to \phi\phi$ branching fraction and a search for the decay $B^0 \to \phi\phi$. The author of this thesis performed the MVA optimisation, mass fits, calculation of the generator and selection efficiencies, the data-driven hadron trigger resimulation and the limit-setting on the $B^0 \to \phi\phi$ branching fraction. Other contributors to this analysis are acknowledged on page vii. This work is published in Ref. [1]. Chapter 6 details an amplitude analysis of the $B^0_s \to \phi K^+ K^-$ decay, in which the decays $B^0_s \to \phi f'_2(1525)$ and $B^0_s \to \phi\phi(1680)$ are observed for the first time. The analysis includes a measurement of the branching fraction and polarisation amplitudes of the $B^0_s \to \phi f'_2(1525)$ mode. The author performed the selection, peaking background estimation and mass fits, modelled the detector effects and background in the amplitude fit, developed the fitter using the RapidFit framework , performed numerous variations on the amplitude fits and calculated the systematic uncertainties. The analysis is currently under review by the collaboration. Finally, Chapter 7 summarises the results and discusses the prospects for the future of these decay modes.

The inclusion of charge-conjugate processes is implied throughout this thesis, unless the distinction is drawn between particle and antiparticle, such as in Sections 2.2 (*CP* violation), 2.4.1 ($B^0_s - \overline{B}^0_s$ mixing) and 2.4.4 (the polarisation puzzle). Where symbols for light, flavourless mesons are quoted without mass numbers, the lowest-mass state is implied: *e.g.* the symbol ϕ refers to the $\phi(1020)$. All mentions of K^{*0} refer to the $K^*(892)^0$.

Chapter 2

Theory

2.1 The Standard Model

The Standard Model of particle physics [9] describes the behaviour and properties of subatomic particles, including three of the four fundamental forces in physics: electromagnetism, the weak interaction and the strong interaction. It was developed from separate descriptions of the strong and electroweak interactions formulated during the 1960s and early 1970s. The acceptance of the Standard Model was motivated by the discoveries of the charm quark [10, 11], tau lepton [12], bottom quark [13], gluon [14–17] and the $W^{\pm}$ and Z^0 bosons [18–21], in the period from 1974 to 1983. The final three particles predicted by the Standard Model, namely the top quark, tau neutrino and Higgs boson, were discovered in 1995 [22, 23], 2000 [24] and 2012 [25, 26] respectively.

The Standard Model is a quantum field theory with twelve fermion fields (quarks and leptons), twelve gauge boson fields (eight gluons, W^+, W^-, Z^0 and γ) and a complex scalar doublet field (the Higgs boson). The particle content of the Standard Model is summarised in Figure 2.1. The theory contains the internal symmetries of the $SU(3) \times SU(2) \times U(1)$ gauge group.

Electromagnetism is a familiar interaction whose effects can be observed on a macroscopic scale. It is the interaction responsible for phenomena including electricity, magnetism, chemical interactions and the structure of matter. The charge of the interaction is called electric charge, and it is mediated by a massless vector boson called the photon (γ).

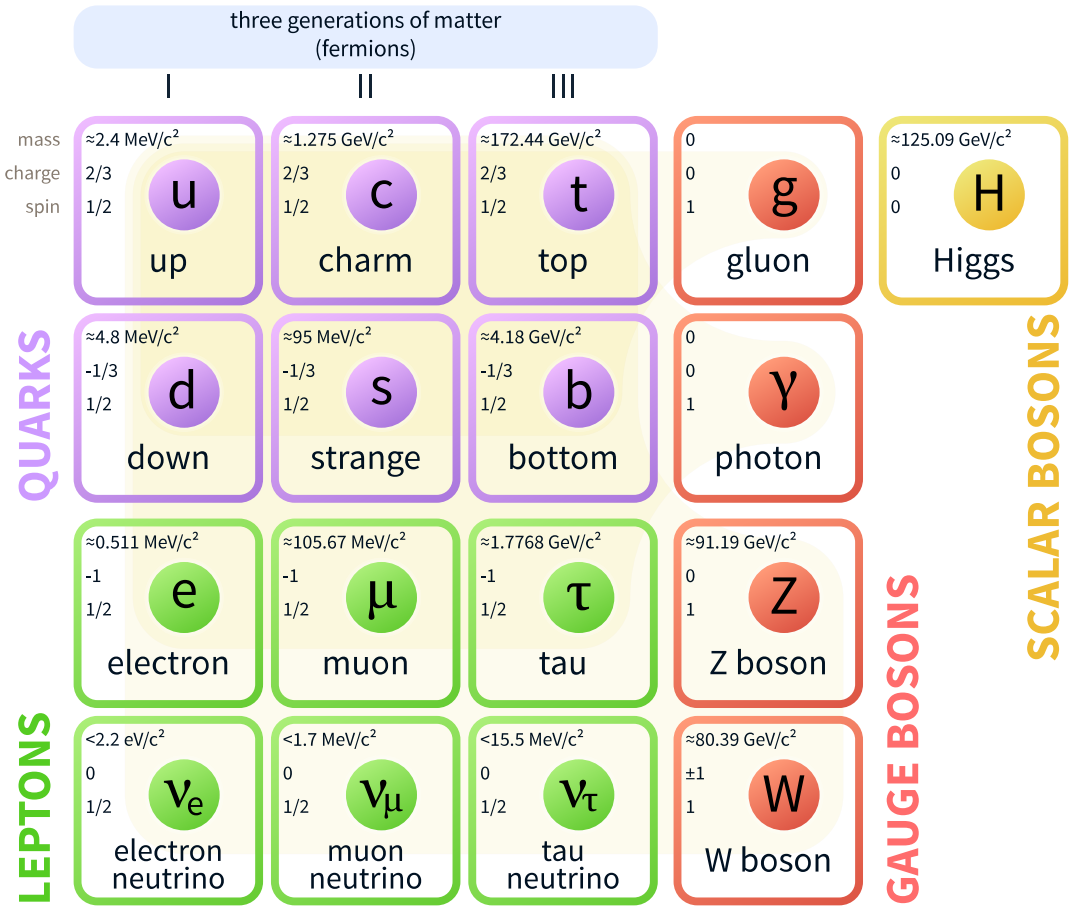


Figure 2.1 *The particle content of the Standard Model [27].*

The weak interaction is a short-range force, mediated by three massive vector bosons: W^+, W^- and Z^0. It couples to all species of fundamental fermions, but only to left-handed fermions and right-handed antifermions. Here, 'handedness' refers to the chirality of a particle. Chirality can be defined using the fifth Dirac gamma matrix $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$. A Dirac fermion field ψ can be projected onto its left-handed and right-handed components using

$$\psi_L = \frac{1}{2}\left(1-\gamma^5\right)\psi \text{ and } \psi_R = \frac{1}{2}\left(1+\gamma^5\right). \tag{2.1}$$

It is the interaction through which neutrino scattering and beta decay occur, and it plays a crucial role in the proton–proton fusion process inside stars. The charge of the weak interaction is called 'weak isospin'. Unlike the electromagnetic and strong interactions, the weak force can change quark flavour.

The electroweak interaction [28–30] is a unified description of electromagnetism

and the weak interaction. It is mediated by four massless vector-boson fields: W^1, W^2 and W^3 with symmetry $SU(2)$ and coupling g, and B with symmetry $U(1)$ and coupling g'. The observed electromagnetic and weak interactions arise due to the Higgs mechanism [31–33], which introduces a complex scalar doublet field with a potential that has a non-zero vacuum expectation value. Upon spontaneous breaking of the symmetry, the physical (massive) weak gauge bosons and the (massless) photon appear as linear combinations of the W and B fields:

$$\begin{aligned} W^+ &= \tfrac{1}{\sqrt{2}}\left(W^1 - iW^2\right), \\ W^- &= \tfrac{1}{\sqrt{2}}\left(W^1 + iW^2\right), \\ Z^0 &= B\sin\theta_W + W^3\cos\theta_W, \\ \gamma &= B\cos\theta_W - W^3\sin\theta_W, \end{aligned} \tag{2.2}$$

where θ_W is a free parameter, the Weinberg angle [29], which relates the masses of the $W^\pm$ and Z^0 bosons by $m_W = m_Z\cos\theta_W$. The quarks and charged leptons acquire mass via their coupling to the Higgs field. The electroweak interaction is discussed in further detail in Section A.1.

The strong interaction is the force which binds quarks to make hadrons and nucleons to make nuclei. The charge of the strong interaction is called 'colour', which can take the values 'red', 'green' or 'blue' (R, G, B) for quarks, or 'antired', 'antigreen' or 'antiblue' ($\overline{R}$, $\overline{G}$, $\overline{B}$) for antiquarks. Neutral combinations of colour charges may be formed from either a colour with its anticolour or the sum of all three colours or anticolours [34] *i.e.*

$$R\overline{R} = G\overline{G} = B\overline{B} = RGB = \overline{RGB} = 0. \tag{2.3}$$

The interaction is mediated by eight massless vector boson fields (gluons) with symmetry $SU(3)$ and coupling g_s. Each gluon field carries colour and anticolour in combinations corresponding to the Gell-Mann matrices, are electrically neutral and have zero weak isospin. At low energy scales, the value of g_s is large enough that perturbation theory cannot be used in calculations of strong interaction processes. The size of g_s, combined with the gluon self-interaction vertex, restricts the strong force to a short range. This is known as colour confinement [35]. The strong interaction is discussed in further detail in Section A.2.

Quarks are fundamental fermions which couple to the strong, weak and electromagnetic interactions [34]. There are six 'flavours' of quark, three of which (u, c and t) have electric charge $+2/3$ and weak isospin $+1/2$, and three (d,

s and b) have electric charge $-1/3$ and weak isospin $-1/2$. Colour confinement means that quarks are never observed in isolation, but only in colour-singlet bound states called hadrons. Hadrons are categorised as either mesons ($q\overline{q}$ states) or baryons (qqq or $\overline{qqq}$ states). The $q\overline{q}$ colour-singlet state is

$$\frac{1}{\sqrt{3}}\left(R\overline{R} + G\overline{G} + B\overline{B}\right),$$

and the colour-singlet qqq state is

$$\frac{1}{\sqrt{6}}\left(RGB - RBG + GBR - GRB + BRG - BGR\right).$$

Other colour-singlet combinations of quarks (*e.g.* $q\bar{q}q\bar{q}$, $qqqq\bar{q}$ *etc.*) are possible and are known as exotic hadrons [34].

Leptons are fundamental fermions which do not couple to the strong interaction. They are divided into three charged leptons (e, μ and τ), with electric charge -1 and weak isospin $-1/2$, and three neutrinos (ν_e, ν_μ and ν_τ), which are electrically neutral and have weak isospin $+1/2$.

The field of experimental particle physics is centred around rigorous tests of the Standard Model and searching for effects which do not fit within its framework. To date, the only observed behaviour of fundamental particles to disagree with the theory is that neutrinos oscillate between flavours [36]. This can only be explained by the introduction of non-zero neutrino masses. In addition, the Standard Model does not account for gravity, dark matter, dark energy and the matter–antimatter asymmetry of the universe.

2.2 *CP* violation

The invariance of a system under a discrete or continuous transformation is referred to as a symmetry. According to Noether's theorem [37], continuous symmetries lead to conserved quantities, *e.g.* time translation symmetry leads to the conservation of energy. Discrete symmetries are not covered by Noether's theorem and therefore do not necessarily correspond to conservation laws. Important discrete symmetries in the Standard Model are charge conjugation (C), parity (P) and time-reversal (T).

Under the charge conjugation operator C, a particle ψ is mapped to its antiparticle

$\overline{\psi}$. The parity operator P reverses the sign of Cartesian spatial coordinates: $(x, y, z) \to (-x, -y, -z)$. The time-reversal operator T reverses the sign of the time coordinate $t \to -t$. The discrete symmetries C, P and T are the conditions that an observable remains invariant under the associated transformation.

CP symmetry is the invariance under the combined operation of charge conjugation and parity:

$$CP|\psi(x, y, z)\rangle = |\overline{\psi}(-x, -y, -z)\rangle. \tag{2.4}$$

The electromagnetic and strong interactions conserve the symmetries C, P and *CP*. The weak interaction is known to violate P symmetry, as it couples to left-handed fermions but not right-handed ones [38]. Similarly, it violates C symmetry, as it couples to left-handed fermions but not left-handed antifermions. *CP* symmetry was thought to be conserved in the weak interaction until the observation of the *CP*-violating decay $K^0_{\mathrm{L}} \to \pi^+\pi^-$ by Cronin and Fitch in 1963 [39].

In the context of the matter–antimatter asymmetry of the observable universe, *CP* violation is needed to allow baryon-number-violating processes to occur more frequently in one direction than the other. This is the second Sakharov condition for baryogenesis [3].

There are three types of *CP* violation: *CP* violation in decay (direct *CP* violation), *CP* violation in mixing (indirect *CP* violation) and *CP* violation in the interference between mixing and decay [40].

Direct *CP* violation is an asymmetry in the rate of a decay and its *CP*-conjugate process. Consider some particle $|\psi\rangle$ that decays weakly to some final state $|f\rangle$. The amplitudes of this decay and its *CP*-conjugate process are

$$A_f = \langle f|H|\psi\rangle \quad \text{and} \quad \overline{A}_{\overline{f}} = \langle \overline{f}|H|\overline{\psi}\rangle, \tag{2.5}$$

where H is the weak-force Hamiltonian. Since probability is the square of decay amplitude, *CP* symmetry is violated if

$$\left|\frac{A_f}{\overline{A}_{\overline{f}}}\right|^2 \neq 1. \tag{2.6}$$

Decay amplitudes are the sum of all possible contributing intermediate processes. The individual contributing amplitudes are complex numbers which can be expressed as carrying two phases. One is the *CP*-conserving 'strong' phase, δ,

whose sign is invariant under a *CP* transformation. The other is the *CP*-violating weak phase, ϕ, which changes sign under *CP*. The origin of the weak phase in the Standard Model is described in Section 2.3. The decay amplitudes can be written as

$$A_f = \sum_k A_k = \sum_k |A_k| e^{i\delta_k} e^{i\phi_k}, \tag{2.7}$$

and

$$\overline{A}_{\overline{f}} = \sum_k \overline{A}_k = \sum_k |A_k| e^{i\delta_k} e^{-i\phi_k}. \tag{2.8}$$

If there is only one contributing process, i.e. $A_f = A_1$, then the decay probabilities of *CP*-conjugate processes must necessarily be the same:

$$|A_f|^2 = |A_1|^2 e^{i\delta_1 - i\delta_1} e^{i\phi_1 - i\phi_1} = |A_1|^2, \tag{2.9}$$

and

$$|\overline{A}_{\overline{f}}|^2 = |A_1|^2 e^{i\delta_1 - i\delta_1} e^{-i\phi_1 + i\phi_1} = |A_1|^2. \tag{2.10}$$

Therefore, direct *CP* violation can only occur in decays with multiple possible intermediate states. The simplest example is with two contributing processes:

$$\begin{aligned}
A_f &= |A_1| e^{i\delta_1} e^{i\phi_1} + |A_2| e^{i\delta_2} e^{i\phi_2}, \\
|A_f|^2 &= |A_1|^2 + |A_2|^2 + |A_1||A_2| e^{i(\Delta\delta)} e^{i(\Delta\phi)} + |A_1||A_2| e^{-i(\Delta\delta)} e^{-i(\Delta\phi)}, \\
&= |A_1|^2 + |A_2|^2 + 2|A_1||A_2| \left(\cos(\Delta\delta)\cos(\Delta\phi) - \sin(\Delta\delta)\sin(\Delta\phi)\right), \\
|\overline{A}_{\overline{f}}|^2 &= |A_1|^2 + |A_2|^2 + 2|A_1||A_2| \left(\cos(\Delta\delta)\cos(\Delta\phi) + \sin(\Delta\delta)\sin(\Delta\phi)\right),
\end{aligned} \tag{2.11}$$

where $\Delta\delta = \delta_1 - \delta_2$ and $\Delta\phi = \phi_1 - \phi_2$. The difference between the amplitudes is a function of both the strong and weak phases, but only arises as a result of the weak phase changing sign.

$$|A_f|^2 - |\overline{A}_{\overline{f}}|^2 = -4|A_1||A_2| \sin(\Delta\delta)\sin(\Delta\phi). \tag{2.12}$$

Experimentally, direct *CP* violation can be measured by the quantity $\mathcal{A}_{CP}$, defined as

$$\mathcal{A}_{CP} = \frac{|A_f|^2 - |\overline{A}_{\overline{f}}|^2}{|A_f|^2 + |\overline{A}_{\overline{f}}|^2} = \frac{\Gamma(\psi \to f) - \Gamma(\overline{\psi} \to \overline{f})}{\Gamma(\psi \to f) + \Gamma(\overline{\psi} \to \overline{f})}, \tag{2.13}$$

where $\Gamma(\psi \to f)$ is the partial decay width of the process $\psi \to f$.

The second type of *CP* violation occurs in neutral meson mixing. Species of neutral meson which have distinct flavour eigenstates, ψ^0 and $\overline{\psi^0}$, can oscillate

between particle and antiparticle without violating any conservation law. Indirect *CP* violation, or *CP* violation in mixing, is where the rate of mixing is greater in one direction than the other.

Since the *CP* operator transforms one flavour eigenstate into the other, the flavour eigenstates are not eigenstates of *CP*. The *CP* eigenstates of a neutral meson are orthogonal superpositions of the flavour eigenstates with equal amplitude:

$$|\psi_-\rangle = \frac{1}{\sqrt{2}}|\psi^0\rangle + \frac{1}{\sqrt{2}}|\overline{\psi^0}\rangle \quad \text{and} \quad |\psi_+\rangle = \frac{1}{\sqrt{2}}|\psi^0\rangle - \frac{1}{\sqrt{2}}|\overline{\psi^0}\rangle, \tag{2.14}$$

where the subscript of each of the *CP* eigenstates denotes the *CP* eigenvalue:

$$CP|\psi_+\rangle = +|\psi_+\rangle \quad \text{and} \quad CP|\psi_-\rangle = -|\psi_-\rangle. \tag{2.15}$$

As discussed in Section 2.3, the weak eigenstates of quarks are not mass eigenstates. The weak eigenstates of neutral mesons, $|\psi_1\rangle$ and $|\psi_2\rangle$, may be defined as orthogonal superpositions of $|\psi^0\rangle$ and $|\overline{\psi^0}\rangle$:

$$|\psi_1\rangle = p|\psi^0\rangle + q|\overline{\psi^0}\rangle \quad \text{and} \quad |\psi_2\rangle = p|\psi^0\rangle - q|\overline{\psi^0}\rangle, \tag{2.16}$$

where q and p are complex mixing amplitudes. *CP* symmetry is violated if the mixing amplitudes don't have the same magnitude *i.e.*

$$\left|\frac{q}{p}\right|^2 \neq 1, \tag{2.17}$$

which fulfils the condition of an asymmetry in the direction of oscillation.

The third type of *CP* violation is in the interference between mixing and decay. This occurs when a neutral particle ψ^0 and its antiparticle $\overline{\psi^0}$ can decay to the same final state f. In this case, there are two paths for ψ^0 to decay to f: directly $\psi^0 \to f$ and through mixing $\psi^0 \to \overline{\psi^0} \to f$. *CP* violation arises from the difference in the weak phases of the two paths. The first has a weak phase ϕ_D, associated with the decay. The mixing process $\psi^0 \to \overline{\psi^0}$ has a phase ϕ_M. The difference in phase between the paths is the residual phase $\phi_s = \phi_M - 2\phi_D$.

This type of *CP* violation can be quantified by defining λ as the product of the two variables introduced in Equations 2.6 and 2.17:

$$\lambda \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f} \quad \text{and} \quad \overline{\lambda} \equiv \frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_{\overline{f}}}. \tag{2.18}$$

If f is a *CP* eigenstate, then $\lambda = \overline{\lambda}$. If $\lambda \neq 1$ then *CP* symmetry is violated. This manifests itself as an asymmetry in the time-dependent decay rates:

$$\mathcal{A}_{CP}(t) = \frac{d\Gamma(\psi^0 \to f)/dt - d\Gamma(\overline{\psi^0} \to f)/dt}{d\Gamma(\psi^0 \to f)/dt + d\Gamma(\overline{\psi^0} \to f)/dt}. \tag{2.19}$$

It is important to note that *CP* may be violated through interference even if it is not individually violated in decay or mixing. In terms of λ, this would correspond to $|\lambda| = 1$ and $\arg(\lambda) \neq 0$.

2.3 The CKM matrix

The only source of *CP* violation in the Standard Model is a complex phase in the Cabibbo–Kobayashi–Maskawa (CKM) matrix [41, 42], which describes quark mixing. By convention, the weak and mass eigenstates of the up-type quarks are aligned, hence the weak eigenstate d' is that which always couples with a u quark at a flavour-changing $W^{\pm}$ vertex. The CKM matrix is defined as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.20}$$

where d', s' and b' are the weak eigenstates of the down-type quarks, the elements V_{ij} are complex numbers, and d, s and b are the mass eigenstates of the down-type quarks. The unitarity conditions of the matrix are

$$\sum_{k=d,s,b} V_{ik}^* V_{jk} = \delta_{ij} \quad \text{and} \quad \sum_{k=u,c,t} V_{ki}^* V_{kj} = \delta_{ij}. \tag{2.21}$$

The numerical values of the elements may be parametrised by four real numbers: three mixing angles (θ_{12}, θ_{23} and θ_{13}) and a *CP*-violating phase (δ). These quantities are free parameters in the Standard Model and need to be determined experimentally. The cosine and sine of each angle θ_{ij} are conventionally abbreviated to c_{ij} and s_{ij} respectively. With this parametrisation, the CKM matrix is expressed as

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \tag{2.22}$$

This can be decomposed into the product of three rotation matrices, with the phase assigned to the θ_{13} matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ s0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.23}$$

The CKM matrix can be approximated using the Wolfenstein parametrisation [43], which uses four real-valued parameters: λ, A, ρ and η. The expansion to $\mathcal{O}(\lambda^3)$ is

$$\begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}, \tag{2.24}$$

where

$$\begin{aligned} \lambda &= s_{12}, \\ A\lambda^2 &= s_{23}, \\ A\lambda^3(\rho - i\eta) &= s_{13}e^{-i\delta}. \end{aligned} \tag{2.25}$$

This parametrisation demonstrates the hierarchy of the magnitudes of the CKM matrix elements. The diagonal elements V_{ud}, V_{cs} and V_{tb} are close to unity, and elements further from the diagonal are smaller. This introduces a suppression factor for quark transitions between generations, named Cabibbo suppression.

Each unitarity condition in Equation 2.21 with $i \neq j$ can be represented by a triangle in the complex plane. By convention, these unitarity triangles are normalised such that one side has unit length. The two unitarity triangles associated with the B^0 and B_s^0 mesons are constructed from the following relations

$$V_{ud}^* V_{ub} + V_{cd}^* V_{cb} + V_{td}^* V_{tb} = 0 \quad \text{and} \quad V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} = 0. \tag{2.26}$$

2.4 Hadronic $b \to s$ transitions

This thesis presents measurements of charmless $b \to s$ transitions. These are processes that, in the Standard Model, are only allowed at loop-level. This suppression makes them sensitive to contributions from physics beyond the Standard Model [7]. A major area of interest in $b \to s\bar{s}s$ transitions is the study of *CP* violation in charmless B_s^0 decays. The decays $B_s^0 \to \phi\phi$ and $B_s^0 \to \phi K^+ K^-$

are examples of $b \to s\bar{s}s$ processes, discussed in Section 2.4.2, while $B^0 \to \phi\phi$ is an example of a $\bar{b}d \to \bar{s}s$ annihilation process, discussed in Section 2.4.6.

2.4.1 $B_s^0 - \overline{B}_s^0$ mixing

As discussed in Section 2.2, quark mixing allows neutral mesons with non-zero flavour, such as the B_s^0, to transform into their own antiparticle without violating conservation laws. This process is illustrated by the Feynman diagram in Figure 2.2 and is a source of indirect CP violation in B_s^0 decays.

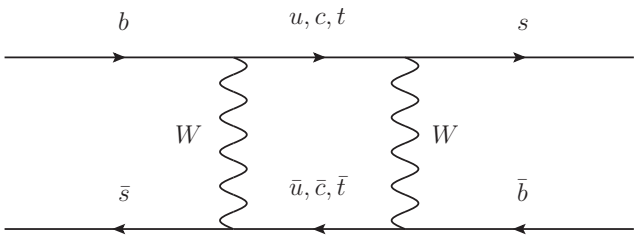


Figure 2.2 *Feynman diagram of $B_s^0 - \overline{B_s^0}$ mixing.*

The time-evolution of a B_s^0 meson is described by the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}|B(t)\rangle = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)|B(t)\rangle, \tag{2.27}$$

where $\mathbf{M}$ is the mass matrix, and $\mathbf{\Gamma}$ is the decay matrix. There are two solutions of the form

$$|B_j(t)\rangle = e^{(-im_j - \Gamma_j/2)t}|B_j(0)\rangle, \tag{2.28}$$

where j denotes the mass eigenstates, 'heavy' (H) or 'light' (L). These are related to the flavour eigenstates by

$$\begin{aligned} |B_H\rangle &= p|B_s^0\rangle + q|\overline{B}_s^0\rangle, \\ |B_L\rangle &= p|B_s^0\rangle - q|\overline{B}_s^0\rangle, \end{aligned} \tag{2.29}$$

where p and q are complex numbers. From Equations 2.28 and 2.29, the time-evolution of a B_s^0 meson produced in a particular flavour eigenstate at $t = 0$ is

$$\begin{aligned} |B_s^0(t)\rangle &= g_+(t)|B_s^0\rangle + \tfrac{q}{p}g_-(t)|\overline{B}_s^0\rangle, \\ |\overline{B}_s^0(t)\rangle &= g_+(t)|\overline{B}_s^0\rangle + \tfrac{p}{q}g_-(t)|B_s^0\rangle, \end{aligned} \tag{2.30}$$

where

$$g_{\pm}(t) = \frac{1}{2}\left(e^{(-im_H-\Gamma_H/2)t} \pm e^{(-im_L-\Gamma_L/2)t}\right). \tag{2.31}$$

The time-dependent rates of a B_s^0 decay to a final state, f, and its charge-conjugate process, are defined as

$$\begin{array}{rcccl} \Gamma(t) & = & |A(t)|^2 & \equiv & |\langle f|H|B_s^0(t)\rangle|^2 \\ \overline{\Gamma}(t) & = & \left|\overline{A}(t)\right|^2 & \equiv & \left|\langle \overline{f}|H|\overline{B}_s^0(t)\rangle\right|^2. \end{array} \tag{2.32}$$

For the decay $B_s^0 \to \phi K^+K^-$, the final state is self-charge-conjugate ($|\overline{f}\rangle = |f\rangle$). The analysis presented in Chapter 6 is time-independent and does not distinguish between B_s^0 and $\overline{B}_s^0$. Therefore, the measured decay rate is the time-integral of the sum of $\Gamma(t)$ and $\overline{\Gamma}(t)$. Under the assumptions that equal numbers of B_s^0 and $\overline{B}_s^0$ mesons are produced and that *CP* violation is small, the time-integral, in terms of the amplitudes at $t = 0$, is [44]

$$\begin{array}{rcl} \int_0^\infty \Gamma(t) + \overline{\Gamma}(t)dt & = & \frac{1}{2}\left(|A(0)|^2 + \left|\overline{A}(0)\right|^2\right)\left(\frac{1}{\Gamma_L} + \frac{1}{\Gamma_H}\right), \\ & + & \Re\left(\overline{A}(0)A(0)^*\right)\left(\frac{1}{\Gamma_L} - \frac{1}{\Gamma_H}\right). \end{array} \tag{2.33}$$

The term proportional to $\left(\frac{1}{\Gamma_L} - \frac{1}{\Gamma_H}\right)$ arises from $B_s^0 - \overline{B}_s^0$ oscillation and affects the angular distribution of the final state.

2.4.2 $b \to s\bar{s}s$ transitions

The transition $b \to s\bar{s}s$ is a flavour-changing neutral current process. In the Standard Model, there are no vertices which change quark flavour with a neutral gauge boson, so this process is forbidden at tree-level. The dominant Feynman diagram for this decay, shown in Figure 2.3, therefore contains a loop. The presence of the loop enhances the sensitivity of the decay to the existence of beyond-Standard Model particles with masses too large to be produced directly at the LHC [7].

Measurements of *CP* violation in $b \to s\bar{s}s$ transitions have been performed with decay modes such as $B^0 \to \phi K^{*0}$ [45–47], $B_s^0 \to \phi\phi$ [48, 49] and $\Lambda_b^0 \to \Lambda\phi$ [50]. These processes differ by the 'spectator' quark content, *i.e.* quarks which appear in the initial and final state and don't partake in the interaction in the leading-order diagrams. So far, these measurements agree with Standard Model predictions.

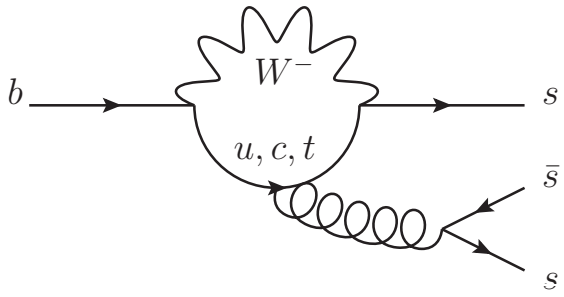


Figure 2.3 *Feynman diagram of a $b \to s\bar{s}s$ decay. This is an example of a 'penguin' diagram, characterised by a single loop formed from a W boson propagator and a quark propagator, with a leg emerging from the quark propagator.*

Since the weak phase of the diagram in Figure 2.3, ϕ_D, is half that of the diagram in Figure 2.2, ϕ_M, the residual phase, $\phi_s = \phi_M - 2\phi_D$, in longitudinally polarised $b \to s\bar{s}s$ decays involving a B_s^0 meson is expected to be close to zero. The observation of large *CP* violation in decays such as $B_s^0 \to \phi\phi$ or $B_s^0 \to \phi f_2'(1525)$ would indicate physics beyond the Standard Model [7]. The value of ϕ_s in $B_s^0 \to \phi\phi$ has been measured by LHCb using Run 1 data: $\phi_s = -0.17 \pm 0.15\,(\text{stat}) \pm 0.03\,(\text{syst})$ [49]. The uncertainty is mostly statistical and will be reduced with the addition of data collected by LHCb in Run 2 and beyond [51].

To search for physics beyond the Standard Model, precise theoretical predictions are needed. Standard Model predictions for observables in $b \to s\bar{s}s$ decays have large uncertainties due to the difficulty inherent in making QCD calculations. The technique of QCD factorisation [8], discussed further in Section 2.4.5, uses long-range form-factors which are determined empirically, *e.g.* from global fits to experimental results. The uncertainties can be reduced by making precision measurements of observables in $b \to s\bar{s}s$ transitions, such as branching fractions and polarisation amplitudes.

2.4.3 Angular distribution of $P \to VS$, VV and VT decays

When studying *CP* violation in decays where the final state is not a *CP* eigenstate, such as $B_s^0 \to \phi\phi$, angular information is used to decompose the decay into different polarisations with definite *CP* eigenvalues. The angular distribution of the final-state particles depends on the spins of intermediate states. Hence, angular information is also important to distinguish between different resonant components in an amplitude analysis. This section considers sequential two-body decays of

a pseudoscalar meson (P) to a final state containing four pseudoscalar mesons, *i.e.* $P \to X_1(\to PP)X_2(\to PP)$, where X_1 and X_2 may be scalar (S), vector (V) or tensor (T) mesons.

Helicity, λ, is the projection of angular momentum, J, onto the direction of a particle. In a two-body decay of a pseudoscalar particle, the vector sum of the angular momenta of the daughter particles must be zero. In the rest frame of the parent particle, the sum of the momentum vectors of the daughters must also be zero. The helicities of the daughters in the decay frame must therefore be equal. Since helicity is a discrete quantity, it takes integer values $-J_{\min} \leq \lambda \leq J_{\min}$, where $J_{\min}$ is the minimum angular momentum of either of the two daughter particles. When one of the intermediate daughters is a scalar meson, such as in $B_s^0 \to \phi f_0(980)$, there is only one possible polarisation: $\lambda = 0$. For the cases where both are vector mesons, as in $B_s^0 \to \phi\phi$, or a vector and a tensor meson, as in $B_s^0 \to \phi f_2'(1525)$, there are three possible values of helicity, each with a different angular distribution and an associated polarisation amplitude: $\mathcal{A}_0$, $\mathcal{A}_+$ and $\mathcal{A}_-$. The polarisation corresponding to $\lambda = 0$ is referred to as 'longitudinal', while the $|\lambda| = 1$ polarisations are 'transverse'.

The longitudinally polarised final state of the decay $B_s^0 \to \phi\phi$ is a CP eigenstate with an eigenvalue of $CP = +1$. The transverse polarisations $\mathcal{A}_+$ and $\mathcal{A}_-$ are themselves not CP eigenstates, however the linear combination $\mathcal{A}_\parallel \equiv \frac{1}{\sqrt{2}}(\mathcal{A}_+ + \mathcal{A}_-)$ has $CP = +1$, and $\mathcal{A}_\perp \equiv \frac{1}{\sqrt{2}}(\mathcal{A}_+ - \mathcal{A}_-)$ has $CP = -1$.

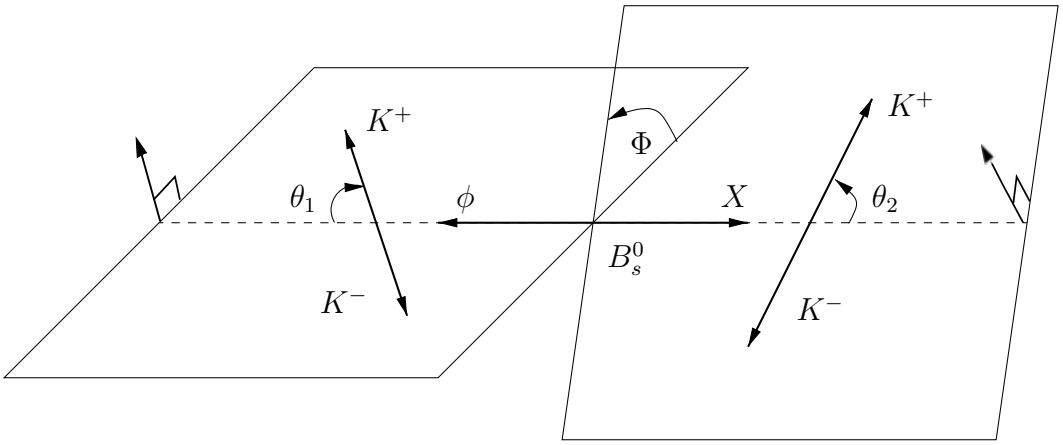


Figure 2.4 *Decay angles for the $B_s^0 \to \phi(\to K^+K^-)X(\to K^+K^-)$ decay. Each of the momentum vectors is constructed in the rest frame of its direct parent particle.*

The helicity angles, as defined for $B_s^0 \to \phi K^+K^-$, are illustrated in Figure 2.4. The angle Φ is between the decay planes of the daughters of the B_s^0 meson, as

measured in the rest frame of the B_s^0. The angle θ_1 is defined as the angle between the momentum vector of the K^+ in the ϕ rest frame and the momentum vector of the ϕ in the B_s^0 rest frame. Similarly, θ_2 is the angle between the momentum vector of the K^+ in the decay frame of the second daughter of the B_s^0 and the momentum vector of that daughter in the B_s^0 rest frame.

The helicity angle distribution of a $P \to X_1(\to PP)X_2(\to PP)$ decay can be generalised as [52]

$$F(\Phi, \theta_1, \theta_2) = \sum_{\lambda} \mathcal{A}_\lambda Y_{J_1}^{-\lambda}(\pi - \theta_1, -\Phi) Y_{J_2}^{\lambda}(\theta_2, 0), \tag{2.34}$$

where J_1 and J_2 are the angular momenta of X_1 and X_2, respectively, and Y_J^λ are the spherical harmonic functions.

2.4.4 The polarisation puzzle

In a decay of a pseudoscalar meson to a pair of vector mesons ($P \to VV$), the daughters are required by the conservation of angular momentum to have the same helicity. This requirement excludes ten of the sixteen combinations of helicities of the final state valence quarks. The remaining six combinations, of which four are longitudinally polarised, are subject to different suppression factors.

Since the weak interaction only couples to fermions with left-handed chirality, and antifermions with right-handed chirality, the outgoing quarks from a weak vertex should predominantly have negative helicity. Likewise, outgoing antiquarks from a weak vertex should predominantly have positive helicity. If a quark q produced at a weak vertex in a b decay undergoes a 'helicity flip', this introduces a suppression factor of $\sim m_q/m_b$ [53]. In $\overline{B} \to VV$ decays, one helicity flip is required for negative polarisation, and two are required for positive polarisation. Longitudinal polarisation is possible without any helicity suppression factors.

The allowed helicity configurations and the number of required helicity flips for a purely electroweak $\overline{B}_s^0 \to \phi\phi$ diagram are summarised in Table 2.1. Note that one of the longitudinally polarised configurations is favoured, and both of the transverse polarisations are suppressed. If, as in Figure 2.3, the pair $s_2\overline{s}_1$ is created by the strong interaction, parity conservation requires that these quarks have opposite helicity, hence the configuration in the penultimate column of the table requires only one helicity flip, rather than three.

Table 2.1 *The six allowed combinations of helicities of the final-state valence quarks in the decay $\overline{B}{}^0_s \to \phi\phi$. The antiquark labelled $\overline{s}_2$ is chosen to be the spectator quark. Red numbers indicate which quarks must undergo a helicity flip in a pure electroweak diagram.*

s_1 helicity	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$+\frac{1}{2}$	$+\frac{1}{2}$
$\overline{s}_1$ helicity	$-\frac{1}{2}$	$+\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$
ϕ_1 helicity	-1	0	0	0	0	$+1$
s_2 helicity	$-\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$+\frac{1}{2}$
$\overline{s}_2$ helicity	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$
ϕ_2 helicity	-1	0	0	0	0	$+1$
Helicity flips	1	0	1	2	3	2

For $\overline{B} \to VV$ decays with light particles in the final state, the expected hierarchy of the helicity amplitudes is therefore $|\mathcal{A}_0| \gg |\mathcal{A}_-| \gg |\mathcal{A}_+|$ *i.e.* the longitudinal fraction, $\mathcal{F}_0 \equiv |\mathcal{A}_0|^2 / \sum_i |\mathcal{A}_i|^2$, is large. For $B \to VV$ decays, the expected hierarchy is $|\mathcal{A}_0| \gg |\mathcal{A}_+| \gg |\mathcal{A}_-|$.

The measured values of $\mathcal{F}_0$ in various $B^0 \to VV$ and $B^0_s \to VV$ decay channels are shown in Figure 2.5. Tree-level dominated decays, such as $B^0 \to \rho^+\rho^-$, are indeed observed to be predominantly longitudinally polarised [55]. However, penguin-dominated decays, such as $B^0 \to \phi K^{*0}$, are found to have large transverse polarisations [45, 46]. This 'polarisation puzzle' has generated much discussion as to whether this discrepancy is due to effects such as penguin-annihilation contributions [53], final-state interactions [56, 57], form-factor tuning [58] or physics beyond the Standard Model (see Ref. [59] for a comprehensive list). In order to resolve this issue, further experimental study is required to reduce the uncertainties on the long-range form-factors and allow more precise theoretical calculations of the polarisation amplitudes.

2.4.5 QCD calculations

At sufficiently high energy scales, the strong coupling, g_s, becomes small enough that processes mediated by the strong interaction are calculable using perturbation theory. The technique of QCD factorisation [8] uses the assumption that a 'hard' strong-interaction process with a high momentum transfer, such as the decay of a heavy particle to light hadrons, occurs on a much shorter time-scale than the subsequent hadronisation. The calculation of the process is split in two parts: the process-dependent short-range hard scattering or decay, which can be calculated using perturbation theory, and the process-independent long-range hadronisation,

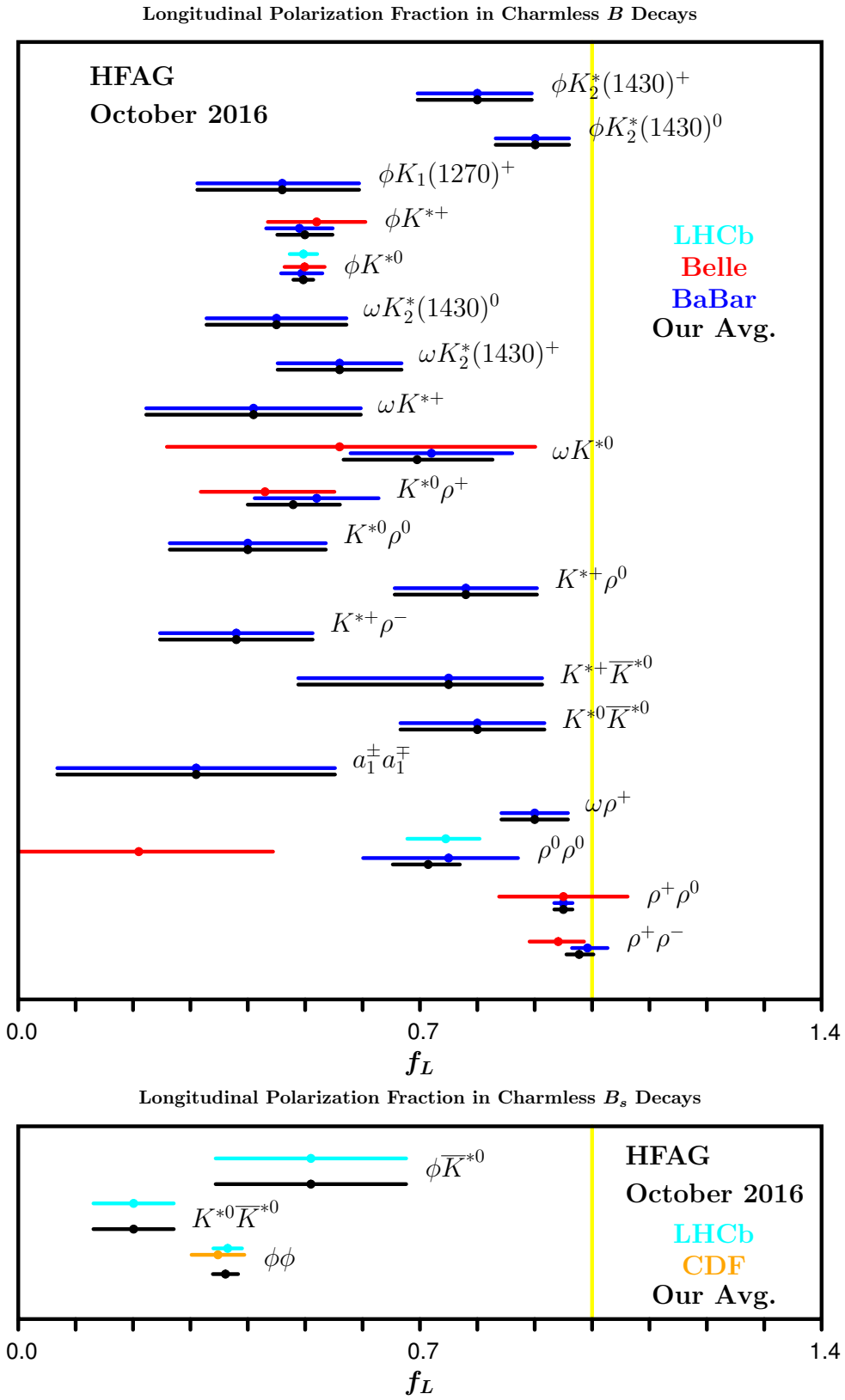


Figure 2.5 *Measurements of the longitudinal polarisation fractions of $B^0 \to VV$ (top) and $B_s^0 \to VV$ (bottom) decays [54].*

which cannot. The latter can be obtained from experiment, therefore precise measurements can serve to reduce the uncertainties on predictions.

The physical observables measured in the analyses presented in this thesis include the branching fractions and polarisation fractions of the decays $B_s^0 \to \phi\phi$ and $B_s^0 \to \phi f_2'(1525)$. These quantities have been calculated using the techniques of QCD factorisation and/or perturbative QCD.

Table 2.2 *Predictions and measurements of the $B_s^0 \to \phi\phi$ branching fraction and polarisation fractions. The uncertainties on the values from experiment are statistical and systematic. QCD factorisation and perturbative QCD are abbreviated to QCDf and pQCD respectively.*

BF ($\times 10^{-5}$)	$\mathcal{F}_0$ (%)	$\mathcal{F}_\perp$ (%)	Source
1.31	86.6	–	QCDf [60]
$1.95 \pm 1.0^{+1.3}_{-0.8}$	48^{+26}_{-27}	–	QCDf [59]
$1.67^{+0.26+1.13}_{-0.21-0.88}$	36^{+3+23}_{-4-18}	–	QCDf [61]
$1.67^{+0.89}_{-0.71}$	$34.7^{+8.9}_{-7.1}$	$31.6^{+3.4}_{-4.4}$	pQCD [62]
$2.64 \pm 0.48 \pm 0.45 \pm 0.38$	39.7 ± 16	31.2 ± 8.9	QCDf [63]
$1.91 \pm 0.26 \pm 0.16$	$34.8 \pm 4.1 \pm 2.1$	$36.5 \pm 4.4 \pm 2.7$	CDF [48]
$1.84 \pm 0.05 \pm 0.18$	$36.4 \pm 1.2 \pm 0.9$	$30.5 \pm 1.3 \pm 0.5$	LHCb [1, 49]

The $B_s^0 \to \phi\phi$ branching fraction and longitudinal polarisation fraction have been calculated using QCD factorisation [59–61, 63] and using perturbative QCD. [62]. The perpendicular polarisation fraction, $\mathcal{F}_\perp \equiv |\mathcal{A}_\perp|^2 / \sum_i |\mathcal{A}_i|^2$, has also been calculated using QCD factorisation [63] and perturbative QCD [62]. The results of these calculations are given in Table 2.2. Measurements of these observables are shown for comparison. The earliest of these calculations [60] predates the observation of $\mathcal{F}_0 \sim 0.5$ in penguin-dominated $b \to s\bar{s}s$ decays, as discussed in Section 2.4.4. Subsequent calculations are tuned to reproduce the measured $B^0 \to \phi K^{*0}$ longitudinal polarisation.

In Ref. [64], the branching fraction of the decay $B_s^0 \to \phi f_2'(1525)$ is calculated using a perturbative QCD approach to be $\left(3.1^{+1.8}_{-1.4} \pm 0.6\right) \times 10^{-6}$. The same reference also calculates the longitudinal polarisation fraction of this decay to be $\left(75.3^{+3.0+3.5}_{-3.2-1.7}\right)$ %. In light of the polarisation puzzle, it is interesting to measure this fraction in data to see whether it agrees with this prediction. This measurement is performed in the analysis described in Chapter 6. With a larger data sample, it will be possible to study *CP* violation using this channel.

2.4.6 $\overline{b}d \to s\overline{s}$ annihilations

The decay $B^0 \to \phi\phi$ is an example of the annihilation process $\overline{b}d \to s\overline{s}$, which is illustrated in Figure 2.6. This is highly suppressed in the Standard Model, due to the presence of at least one loop, Cabibbo suppression [41] and the OZI rule [65–67]. Figure 2.7 shows both the non-factorisable and factorisable contributions to the $B^0 \to \phi\phi$ branching fraction.

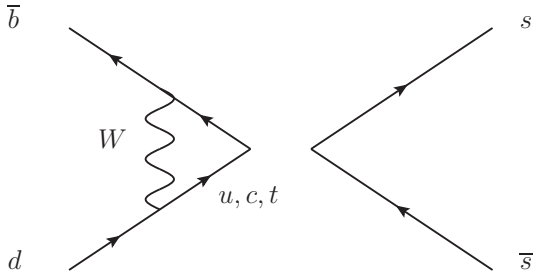


Figure 2.6 *Illustration of the $\overline{b}d \to s\overline{s}$ penguin annihilation topology.*

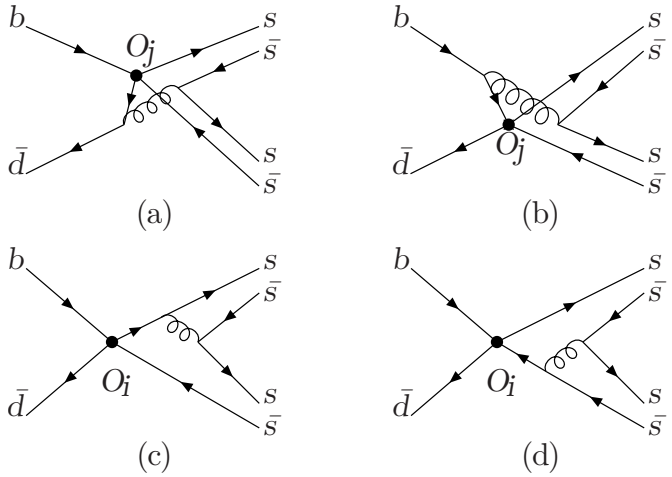


Figure 2.7 *Non-factorisable ((a) and (b)) and factorisable ((c) and (d)) diagrams for the decay $B^0 \to \phi\phi$ [68].*

The $B^0 \to \phi\phi$ branching fraction has been calculated using QCD factorisation [68] and perturbative QCD [62, 69]. These techniques are discussed further in Section 2.4.5. In Ref. [59], an upper-limit estimate is given. The results of these calculations are given in Table 2.3. It should be noted that the relative uncertainties on these values are large, and there is a variation of an order of magnitude between results obtained using different techniques.

The branching fraction of this decay may be enhanced by rescattering processes [70], $\omega - \phi$ mixing or new physics such as R-parity-violating (RPV) supersymmetry [68].

Table 2.3 *Theory predictions for the $B^0 \to \phi\phi$ branching fraction.*

BF ($\times 10^{-8}$)	Approach
$0.21^{+0.16}_{-0.03}$	QCD factorisation [68]
$1.89^{+0.61}_{-0.21}$	perturbative QCD [69]
< 3	QCD factorisation [59]
$1.2^{+0.3+0.5}_{-0.2-0.4} \pm 0.1$	perturbative QCD [62]

For example, in Ref. [68], the Standard Model branching fraction of $B^0 \to \phi\phi$ is calculated to be $\left(2.1^{+1.6}_{-0.3}\right) \times 10^{-9}$, and the available bounds on RPV couplings allow for enhancement to the order of 10^{-8}–10^{-7}.

Since the d quark in Figure 2.6 may be substituted with an s quark, measurement of the $B^0 \to \phi\phi$ branching fraction provides information on the size of non-factorisable penguin-annihilation contributions to the $B^0_s \to \phi\phi$ decay.

Chapter 3

The LHCb experiment

3.1 Introduction

The Large Hadron Collider Beauty (LHCb) experiment is designed to make precision measurements of hadrons containing charm and bottom quarks with the aim of finding physics beyond the Standard Model. The LHC has a larger $b\bar{b}$ production cross-section than any previous collider, making it the largest source of all species of b hadrons [71–73]. During 2011 and 2012, the LHC produced $\mathcal{O}(10^{10})$ b hadron decays in the acceptance of the LHCb detector. Using this dataset, several key measurements have been made, such as the first evidence for the decay $B_s^0 \to \mu^+\mu^-$ [74], an angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ [75], the precise measurement of the *CP*-violating phase in $B_s^0 \to J/\psi\,\phi$ [76] and several tree-level measurements of the CKM angle γ [77]. Beyond the primary goals of the experiment, LHCb has also had success in exotic hadron spectroscopy, most notably with the discovery of pentaquarks [78] and the observation and confirmation of several exotic meson states [79, 80]. The physics programme also covers electroweak physics (*e.g.* [81]), exotics (*e.g.* [82]), soft QCD (*e.g.* [83]), heavy ion physics [84–87] and studies of beam–gas collisions [88].

3.2 The Large Hadron Collider

The Large Hadron Collider (LHC) [89] is a 27 km circular collider at CERN. Two beams, composed either of protons or lead ions, are brought to collision at four

crossing points around the ring. At each crossing point, a large detector is situated. In clockwise order, from above, these are ATLAS [90], ALICE [91], CMS [92] and LHCb [93]. The ATLAS and CMS experiments are 'general-purpose' detectors, with broad physics programmes that notably include Higgs boson, electroweak and t quark physics, together with direct searches for supersymmetric and other non-Standard-Model particles. The ALICE experiment is focused on the physics of the quark–gluon plasma produced in heavy ion collisions. The LHCb experiment is designed to study decays of hadrons containing bottom and charm quarks, with particular focus on searching for non-Standard-Model sources of *CP* violation and indirect evidence for new physics in rare decays. In addition, there are three smaller experiments. The TOTEM [94] experiment, located either side of CMS, measures the total elastic and diffuse cross section of pp collisions. The LHCf [95] experiment, located either side of ATLAS, measures neutral particle production for use in cosmic ray research. Finally, MoEDAL [96], located in the LHCb cavern, searches for magnetic monopoles.

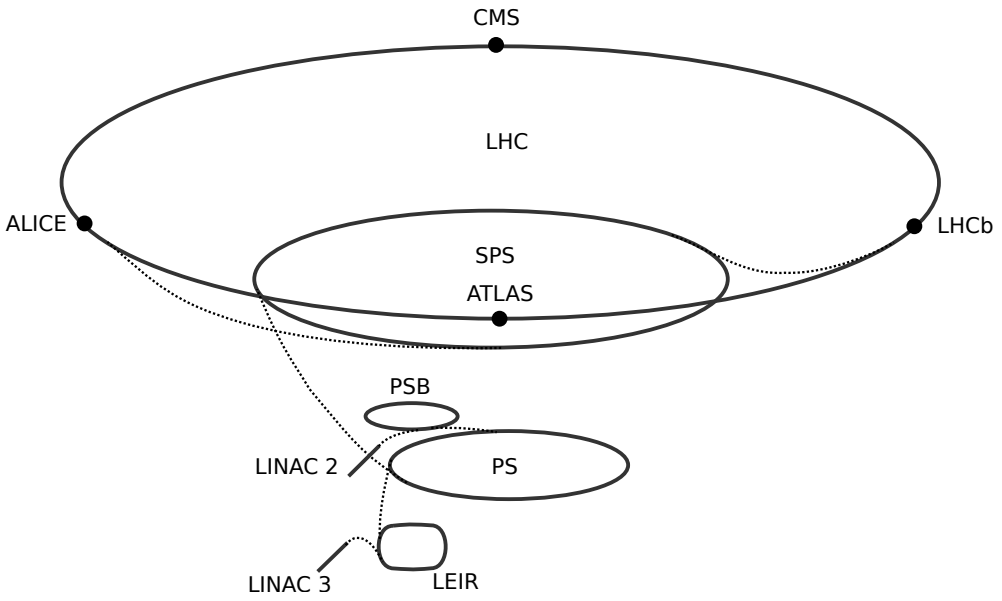


Figure 3.1 *A schematic representation of the LHC injector chain. Solid lines represent accelerators, dotted lines represent transfer lines, and solid circles represent crossing points.*

The chain of accelerators comprising the LHC complex is illustrated in Figure 3.1. The proton beam starts in a linear accelerator, called Linac2, which takes protons from a hydrogen ion source and accelerates them to 50 MeV. The beam then traverses three successively larger, circular accelerators — the 1.4 GeV Proton Synchrotron Booster (PSB), the 28 GeV Proton Synchrotron (PS) and the 450 GeV Super Proton Synchrotron (SPS). The beams are composed of bunches

of $\mathcal{O}(10^{11})$ protons, with a minimum spacing between bunches of 25 ns. An LHC beam is composed of a maximum of 2808 bunches, accounting for gaps corresponding to kicker magnet rise times in the injector chain and the beam dump system. Ion beams start at a different linear accelerator, Linac3, and traverse the Low-Energy Ion Ring (LEIR) before following the same route as proton beams through the PS and SPS.

The LHC receives 450 GeV protons from the SPS and further accelerates them using superconducting radio-frequency cavities. The design beam energy is 7 TeV, however 6.5 TeV is the highest energy achieved so far. The ring has eight bending arcs, with superconducting dipole magnets capable of generating magnetic fields up to 8.3 T. Multipole magnets are used to focus the beam. Between the bending arcs are straight sections: four have crossing points, two are used for collimation, one for acceleration and the last for the beam dump system. From a top-down perspective, beam 1 circulates clockwise and is injected just before it traverses through ALICE. Beam 2 circulates anticlockwise and is injected just before it traverses LHCb, *i.e.* 'downstream' from the detector in the coordinate system defined in Section 3.3.1.

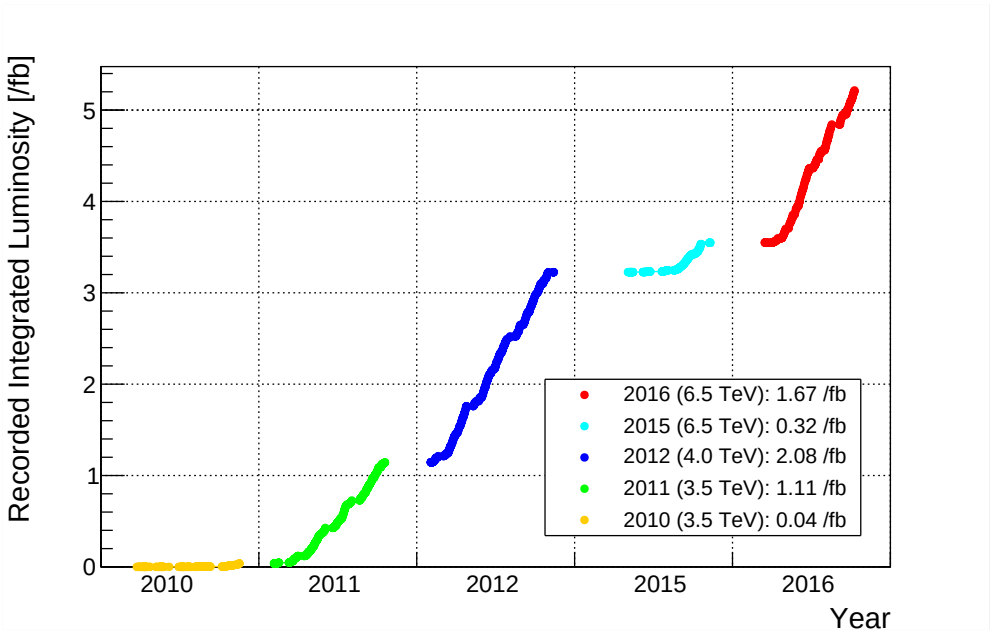


Figure 3.2 *Evolution of integrated luminosity of proton–proton collisions recorded by the LHCb detector. The legend shows the beam energy and integrated luminosity for each year.*

The first beams circulated in the LHC during 2008, with first collisions in 2009 at a beam energy of 450 GeV. The beam energy in 2010 and 2011 was 3.5 TeV, rising to 4 TeV in 2012. During this period, referred to as Run 1, a total integrated

luminosity of 3.2 fb^{-1} of proton–proton collisions was recorded by LHCb, of which 1.15 fb^{-1} was collected at a centre-of-mass energy of $\sqrt{s} = 7$ TeV and 2.1 fb^{-1} at 8 TeV. In 2015, after a two-year shutdown for maintenance and upgrades, Run 2 of the LHC started. By the end of 2016, LHCb had collected 2.0 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV. The evolution of recorded integrated luminosity is shown in Figure 3.2.

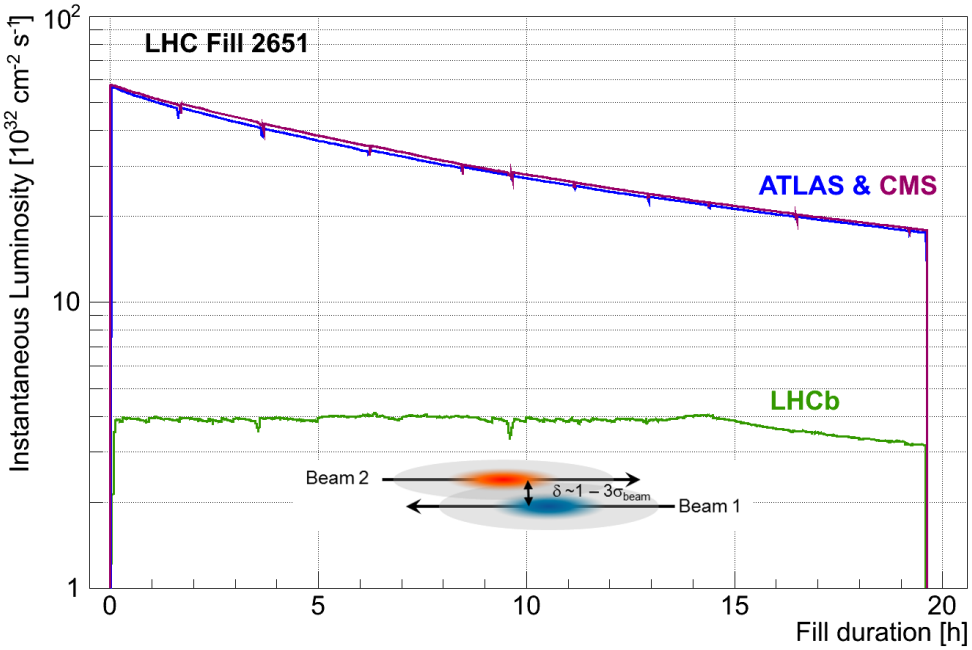


Figure 3.3 *Evolution of the instantaneous luminosity at LHCb (green), ATLAS (blue) and CMS (purple) during a typical 'fill' in Run 1. Note that the instantaneous luminosity at LHCb is constant until about 14 hours, when the beams have zero separation.*

The LHC delivers a constant instantaneous luminosity to LHCb, using a technique called 'luminosity levelling' [97]. The lateral separation of the beams at the crossing point is adjusted periodically to achieve a constant target luminosity. As the beam intensity reduces, the separation reduces accordingly. The effect of luminosity levelling is demonstrated in Figure 3.3. During 2011, the target luminosity was 3.5×10^{32} $\text{cm}^{-2}\,\text{s}^{-1}$, increasing to 4.0×10^{32} $\text{cm}^{-2}\,\text{s}^{-1}$ in 2012. Luminosity levelling is performed so that events predominantly contain a single collision, which makes them easier to reconstruct and allows for more precise measurements.

3.3 The LHCb detector

This section describes the geometry of the LHCb detector, illustrated in Figure 3.4, as well as the purpose and locations of its subdetectors. The subdetectors are described in further detail in Sections 3.4 and 3.5.

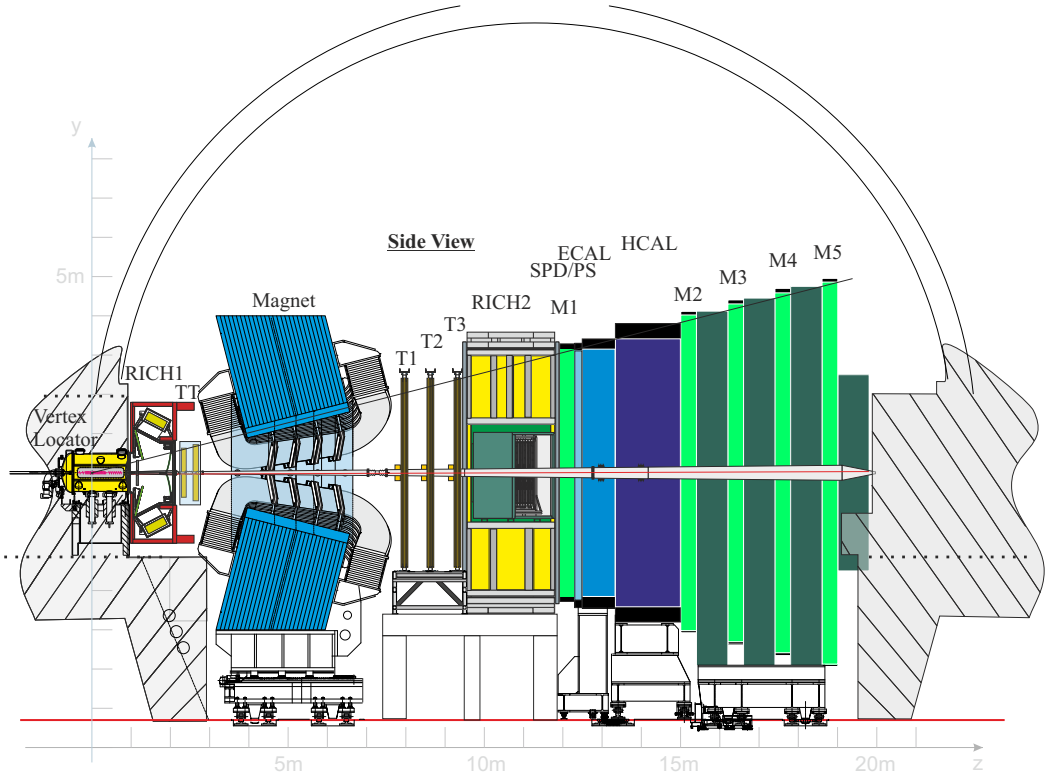


Figure 3.4 *Side view of the LHCb detector in the y–z (non-bending) plane, showing the positions of the subdetectors.*

3.3.1 Coordinate system

A right-handed Cartesian coordinate system is adopted by LHCb. The origin of the coordinate system is located at the interaction point. The z axis is aligned with the direction of beam 1 and points in the 'downstream' direction, towards the end of the detector. The y axis points vertically upwards. The x axis points horizontally towards the centre of the LHC ring.

3.3.2 Overview

In high-energy hadron collisions, the majority of heavy quark pairs are produced at a small angle with respect to the beam axis. The polar angle and pseudorapidity distributions of b quarks in simulated LHC collisions are shown in Figure 3.5. The LHCb detector [93], shown in Figure 3.4, is positioned almost entirely on one side of the crossing point in the direction of travel of beam 1, defined as the 'forward' direction. The polar angle coverage extends from 15 mrad to 300 mrad in the horizontal (bending) plane and to 250 mrad in the vertical (non-bending) plane. This translates to a solid-angle coverage of approximately 4 % and a pseudorapidity range of $2 < \eta < 5$. Roughly 25 % of the $b\bar{b}$ pairs produced in LHC collisions are contained in this region.

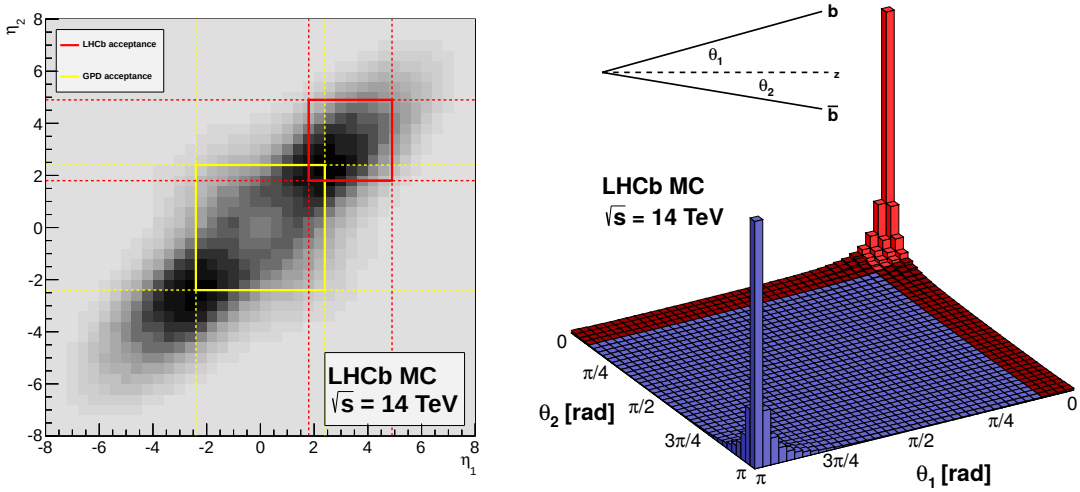


Figure 3.5 *Left: Pseudorapidity distribution of b quarks produced in simulated pp collisions. The solid lines show the acceptance of LHCb (red) compared to ATLAS and CMS (yellow). Right: Polar angle distribution of simulated b quarks. The red region is the acceptance of LHCb.*

The subdetectors of LHCb can be grouped in two categories based on their function: namely tracking and particle identification. The tracking system, detailed in Section 3.4, consists of a retractable silicon-strip vertex locator [98] around the interaction point; a silicon-strip tracking detector [99] immediately upstream from the magnet [100] and three tracking stations immediately downstream from the magnet, which consist of silicon strips in the inner part [101] and straw tubes in the outer part [102]. The particle identification system, detailed in Section 3.5, consists of the ring-imaging Cherenkov (RICH) system, the calorimeter system and the muon detection system. There are two RICH detectors [103], each designed

to discriminate between pions, kaons and protons in different momentum ranges. The first is located between the vertex locator and the magnet, while the other is located downstream from the third tracking station. The calorimetry system [104] sits downstream from the second RICH detector and consists of a scintillator pad and preshower detector, an electromagnetic calorimeter and a hadronic calorimeter. The muon system [105] consists of five stations (M1–M5) composed of multi-wire proportional counters in low occupancy regions and triple gas electron multipliers in high occupancy regions. The M1 station is located between the second RICH detector and the calorimeter. The M2 station is located immediately downstream from the calorimeter, and the rest of the stations follow, interspersed with iron absorbers.

3.4 Tracking

The tracking system is described in this section. Its purpose is to reconstruct the trajectories of charged particles (tracks) using the positions of where charged particles have interacted with the sensors of the detectors (hits). The tracking system consists of a vertex locator, a dipole bending magnet and four tracking stations: one located before the magnet and three after. The momentum of a particle is deduced from the angle of deflection through the magnetic field.

3.4.1 Magnet

In order to measure the momentum of charged particles to high precision, the tracking system needs a well-known magnetic field to bend their path. A dipole magnet is suited to this task. In order to achieve a momentum resolution of $\delta p/p = 4 \times 10^{-3}$ for 10 GeV/c particles, an integrated field of 4 Tm is required between the VELO and the tracking stations. Since the Hybrid Photon Detectors in the RICH detectors (see Section 3.5.1) are sensitive to magnetic fields, the field strength is required to be small in the region of RICH1.

A warm[1] dipole magnet, illustrated in Figure 3.6, is used to provide a magnetic field for the tracking system [100]. The magnetic field as a function of position in z is shown in Figure 3.7. The magnet is composed of two saddle-shaped coils attached to a rectangular yoke with slanted poles, with a wedge-shaped window

[1] *i.e.* not superconducting.

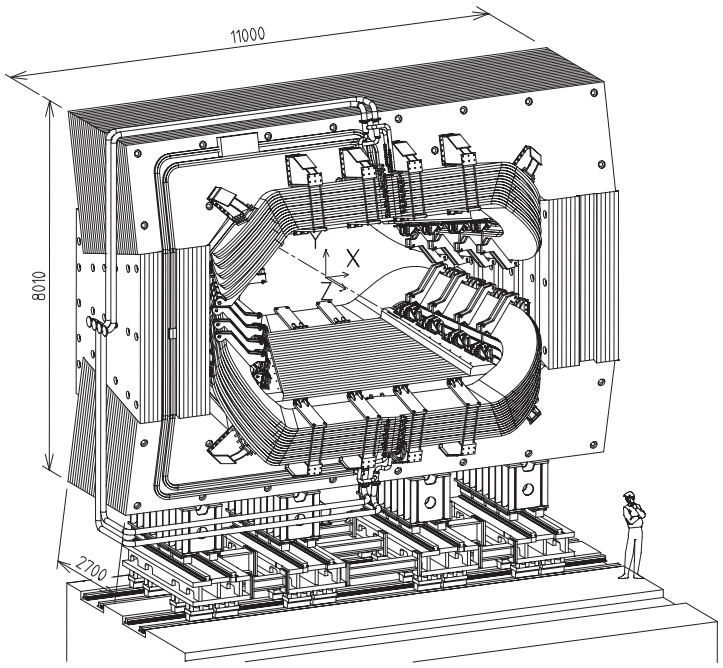


Figure 3.6 *A diagram of the LHCb dipole magnet [100]. The positive z direction is out of the page.*

through the centre that matches the acceptance of the detector. The yoke is made of twenty-seven layers of laminated low-carbon steel, each 100 mm thick and with a maximum mass of 25 tonnes. Each coil is made of fifteen layers of hollow aluminium conductor with a central channel for water-cooling. The nominal current passing through the coils is 5.85 kA.

To achieve the required momentum resolution, the magnetic field was measured to a relative precision of $\mathcal{O}(10^{-4})$ prior to data taking. In order to eliminate artificial *CP* asymmetries induced by the detector, the polarity of the magnet is reversed regularly during the operational year, so that the collected data is split evenly between the two polarities.

3.4.2 Vertex locator

The Vertex Locator (VELO) [98] is a movable silicon strip detector operated in a secondary vacuum, separated from the LHC primary vacuum by an RF-foil. It is used to make precise measurements of particle trajectories close to the interaction point. This allows for the accurate determination of the locations of proton–proton collisions (primary vertices) and heavy quark hadron decays (secondary vertices). The detector consists of 42 modules positioned in the plane perpendicular to the

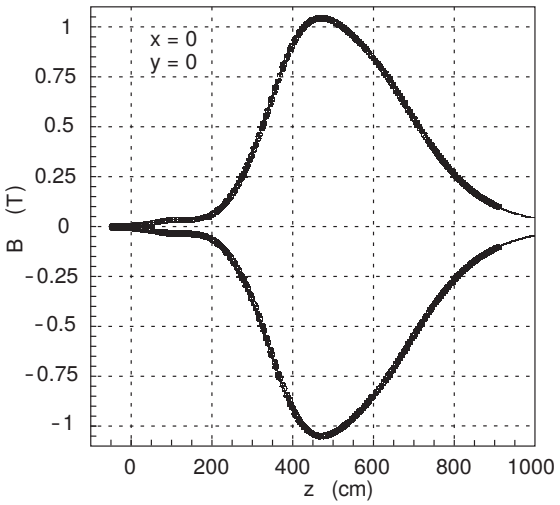


Figure 3.7 *Magnetic field as a function of z, at $x = 0$ cm and $y = 0$ cm for both polarities [100].*

beam axis, z. The VELO is split vertically into two halves, each with an equal number of modules either side of the beam.

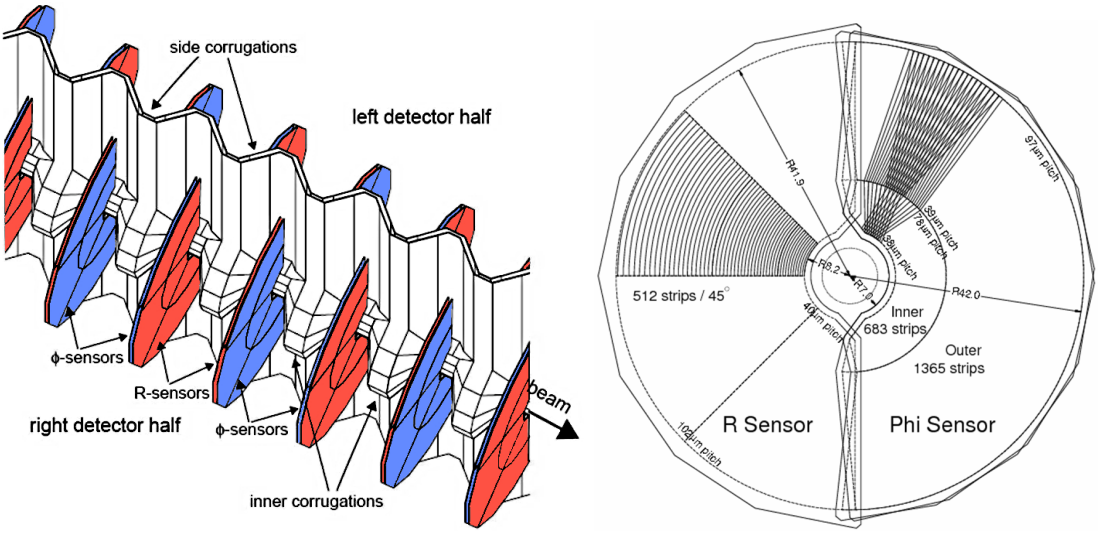


Figure 3.8 *Left: a view of the VELO sensors and RF foil in the 'closed' position. Right: the r and ϕ sensors, illustrating the size and orientation of the silicon microstrips. Note that the diagram shows two ϕ sensors overlaid. [98]*

Each half of the VELO is enclosed in a box of thin aluminium. The inner faces of the boxes (the RF-foil) are corrugated to allow for the modules to overlap when in the closed position, as shown on the left of Figure 3.8. Each module has two semi-circular silicon strip sensors, illustrated on the right of Figure 3.8, which each provide measurements of radius, r, or azimuthal angle, ϕ. The r sensor strips are concentric rings with a variable pitch that increases linearly from 38 μm at the

inner edge to 102 μm at the outer edge. The ϕ sensor strips are wedge-shaped and divided into two regions at $r = 17.25$ mm in order to reduce occupancy and avoid large strip pitches at the outer edge of the sensors. The strips in the inner region have a pitch of 38 μm at the inner edge and 78 μm at the outer edge, while the strips in the outer region have a pitch of 39 μm at the inner edge and 97 μm at the outer edge. To reduce ambiguities in the pattern recognition, the ϕ sensor strips are skewed 20° from the radial direction in the inner region and 10° in the outer region.

The modules have an inner radius of 8.2 mm, which is smaller than required by the beam during injection. To avoid damage to the sensors, the two halves of the VELO are retractable to a distance of 30 mm from the centre of the beam. The VELO is kept open during conditions when the LHC beam is diffuse. It is moved to the closed position when stable beam conditions are declared.

3.4.3 Vertexing performance

The primary vertex resolution is measured in data by comparing two independent measurements of the position of the same vertex in the same event [106]. This is done by randomly dividing the tracks into two sets and reconstructing the primary vertex using each, then taking the difference in position between the two resulting vertices. The width of the distribution of the difference between the vertex positions is $\sqrt{2}$ times the primary vertex resolution. This resolution strongly depends on the number of tracks used to reconstruct the vertex, as shown on the left of Figure 3.9.

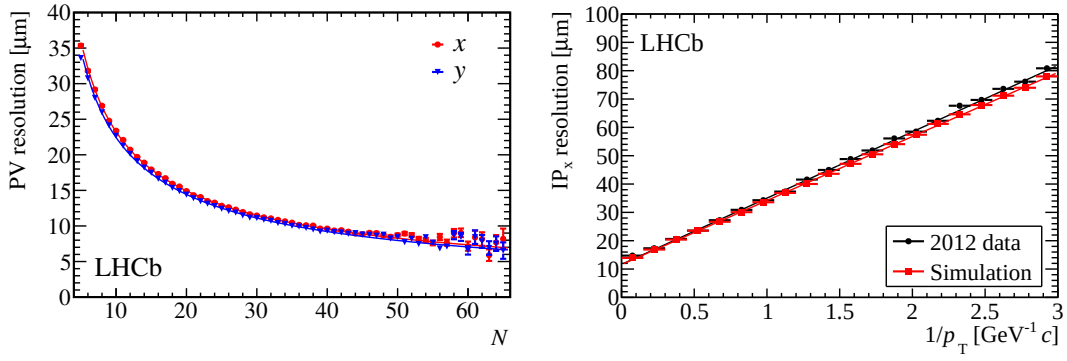


Figure 3.9 *Left: primary vertex resolution as a function of track multiplicity. Right: resolution of impact parameter in the x-direction as a function of the inverse of transverse momentum. [106]*

The impact parameter (IP) is the distance of closest approach of a particle's

trajectory to the primary vertex. The resolution on this quantity is caused by multiple scattering, spatial resolution of the VELO sensor hits, and the distance of extrapolation from the interaction point to its first hit in a sensor. It is customary to quote the resolution of the impact parameter projected onto two orthogonal axes, IP_x and IP_y, which follow Gaussian distributions, unlike the IP resolution itself. Figure 3.9 (right) shows the IP_x resolution as a function of the inverse of the transverse momentum. This dependence is due to multiple scattering in the detector material and the spatial resolution of the VELO. At high transverse momentum, typical for B decays, the IP_x resolution is about 13 μm.

3.4.4 Tracker Turicensis

The Tracker Turicensis (TT) [99] is located immediately before the dipole magnet. It is of particular importance for reconstructing tracks that originate outside of the VELO, such as those that come from K^0_S and Λ decays. The TT consists of four layers of silicon microstrip sensors with a pitch of 183 μm between the strips. The first and last layers are oriented vertically, and the second and third are rotated by $-5°$ and $+5°$ from the vertical. The sensors are housed in a light-tight, electrically insulated and thermally insulated box. The temperature inside the box is maintained below 5°C, and nitrogen gas is continuously flushed through the box to prevent condensation.

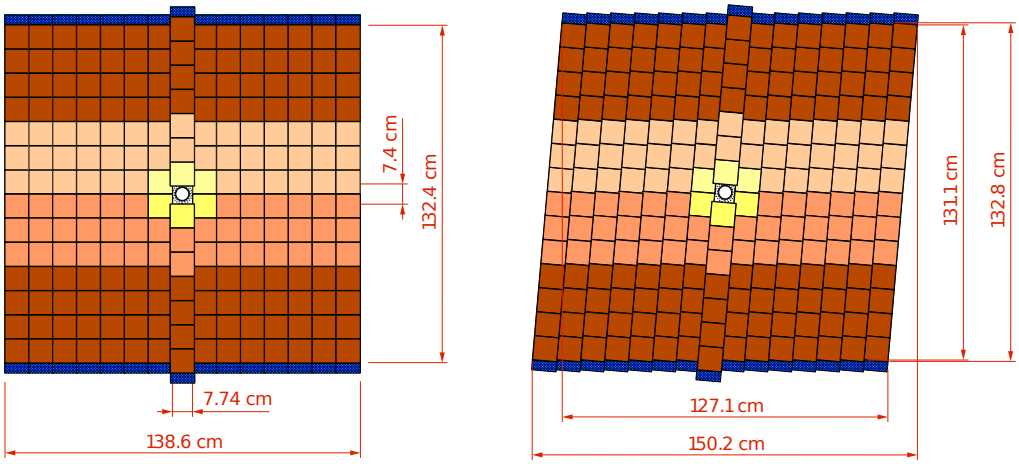


Figure 3.10 *The first two layers of the TT, illustrating the vertical and 5° orientations of the modules [99].*

Each layer of the TT is composed of modules, which are columns of 14 silicon sensors. Two half modules occupy the centre of each layer, above and below the

beam pipe. The first two layers are shown in Figure 3.10. The sensors are read out in 'sectors' of varying sizes (1 to 4) depending on their typical occupancy. Alternate adjacent modules are staggered in z and overlapped by a few millimetres to avoid gaps in the acceptance. The readout boards, structural supports and cooling system are located at the ends of the modules, outside the acceptance of the detector.

3.4.5 Inner Tracker

The Inner Tracker (IT) [101] covers the high-occupancy inner part of the three tracking (T) stations after the magnet. Each station has four boxes arranged around the beam pipe as shown in Figure 3.11. As with the TT, each of the IT stations consists of four layers of silicon microstrip sensors oriented 0°, −5°, +5° and 0° from the vertical. The pitch of the strips is 196 µm. The sensors are housed in light-tight boxes, four for each station, surrounding the beam pipe. As with the TT, the boxes are maintained below 5°C, and nitrogen gas is continuously flushed through to prevent condensation.

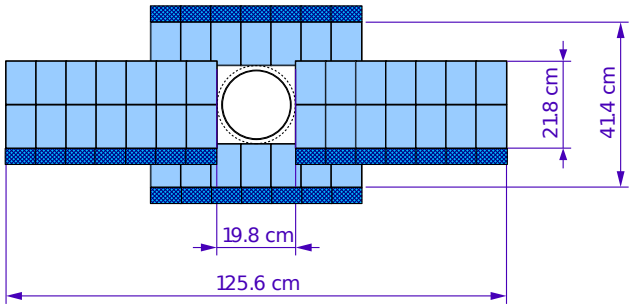


Figure 3.11 *Positions of the IT boxes around the beam pipe.*

The boxes either side of the beam pipe contain modules of two sensors each. The ones above and below the beam pipe contain single-sensor modules. The boxes are offset in z and overlap to avoid gaps in the acceptance. Unlike the TT, the readout boards, cooling system and structural supports are inside the acceptance of the detector.

3.4.6 Outer tracker

The Outer Tracker (OT) [102] is placed surrounding the IT and consists of four modules per T station. Each module consists of two staggered layers of straw tubes covering the rest of the acceptance not taken up by the IT. The modules have the same orientation as the strips in each layer of the IT and TT: the first and last are vertical, and the second and third are rotated by $-5°$ and $+5°$, respectively, from the vertical. The position and orientation of the straw tubes and modules are shown in Figure 3.12. The straw chambers are gas-tight conducting tubes with an internal diameter of 4.9 cm, filled with a mixture of 70 % argon to 30 % carbon dioxide. The tubes are formed from two layers of thin foil. The outer layer is a laminate of polyimide and aluminium, which provides gas-tightness and shielding. The inner layer is carbon-doped polyimide and acts as the cathode. A gold-plated tungsten anode wire runs through the middle of each tube.

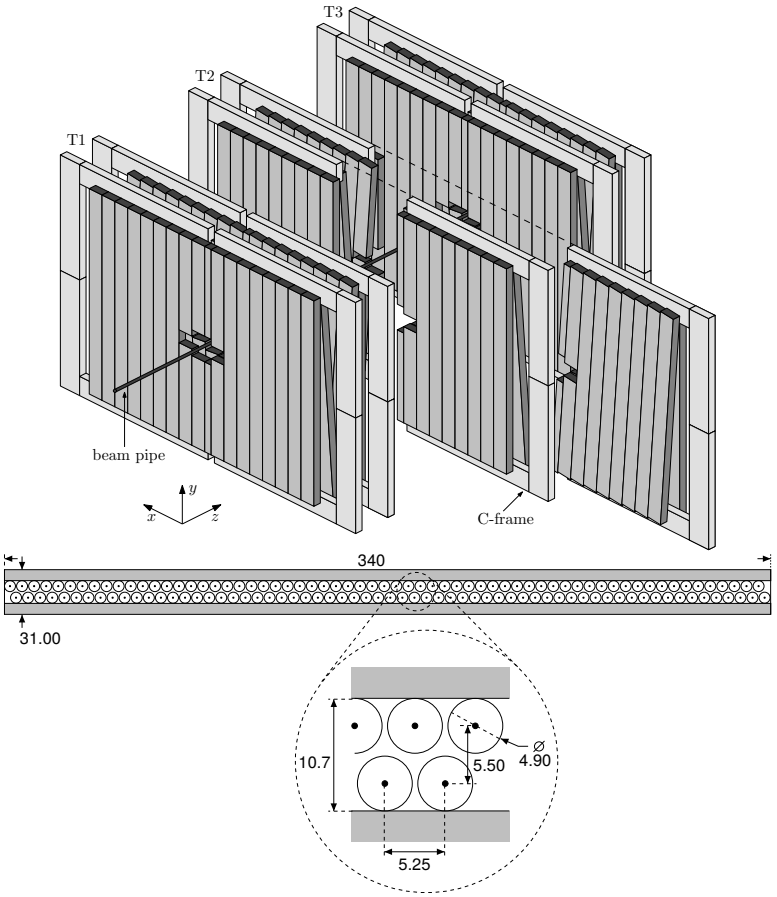


Figure 3.12 *Top: a view of the OT stations, with part of T2 retracted, showing the orientations of the straw tube modules. Note the cross-shaped gap where the IT is located. Bottom: the packing of the straw tubes within one layer of the OT. [102]*

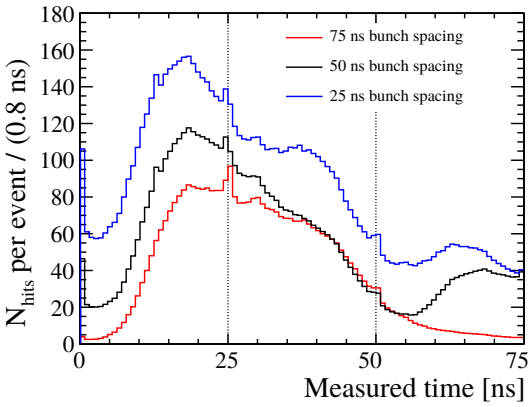


Figure 3.13 *Drift time distribution for the inner-most OT modules. [102]*

The maximum drift time within the straw tubes is around 35 ns, which is lower than the bunch crossing frequency during Run 1. However, to account for variations in time-of-flight, signal propagation and the readout electronics, a 75 ns window is read out. Figure 3.13 shows the measured drift time distribution for three different bunch spacings. The LHC operated mostly with 50 ns and 75 ns bunch spacing during Run 1, with a short period in 2012 at 25 ns.

3.4.7 Track reconstruction

Tracks are reconstructed from hits in the VELO, TT and T detectors. Tracks at LHCb are classified by which subdetectors they pass through, as illustrated in Figure 3.14. The analyses presented in this thesis use only 'long' tracks, defined as those which traverse all tracking stations. The reconstruction of long tracks starts with a search for straight line trajectories among VELO hits. These 'VELO tracks' must have hits in at least three r and three ϕ sensors. Subsequently, two complementary algorithms add information from the TT and T stations, named the 'forward tracking' and 'track matching' algorithms [106].

The forward tracking algorithm starts with a VELO track and combines it with a single hit in a T station to determine the track trajectory. Additional T station hits are searched for in a window around this trajectory. If the found hits satisfy certain quality cuts, the track is kept.

The track matching algorithm starts with a set of VELO tracks and a set of track segments reconstructed in the T stations. The algorithm attempts to match tracks from the two sets by extrapolating tracks from both sets through the magnet.

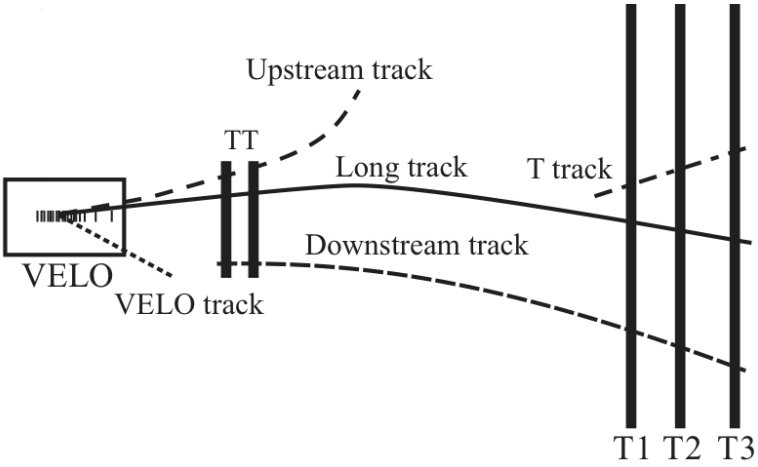


Figure 3.14 *Track types at LHCb, which are defined by the subdetectors they pass through. VELO tracks and T tracks may be combined to form long tracks.*

Quantities such as position, gradient and number of hits are used to evaluate whether or not the VELO and T tracks are compatible. The candidate tracks from both algorithms are combined, duplicates are removed, and any consistent hits in the TT are added to improve the momentum measurement.

The final step of track reconstruction is to fit the tracks with a Kalman-filter [107]. This takes into account multiple scattering and corrects for energy loss due to ionisation. Track quality is determined from the χ^2 per degree of freedom of the fit.

'Ghost' tracks are those which are reconstructed from combinations of hits that do not correspond to the trajectory of a charged particle [106]. Most of these result from incorrect matching of VELO and T tracks. The fraction of ghost tracks in minimum bias events ranges from 6.5 % to 20 % depending on the charged particle multiplicity. A neural network classifier is used to identify and remove ghost tracks. This uses the result of the track fit, track kinematics and the number of measured and expected hits in the tracking stations.

3.4.8 Tracking performance

Tracking efficiency is measured using a 'tag-and-probe' method with $J/\psi \to \mu^+\mu^-$ decays [106]. One of the daughter muons (the 'tag') is fully reconstructed, whereas the other (the 'probe') is only partially reconstructed. The $\mu^+\mu^-$ pairs must pass a J/ψ invariant mass cut. The tracking efficiency is calculated as the proportion of

probe muons which can be matched successfully to fully reconstructed long tracks. Figure 3.15 shows the tracking efficiency in 2011 and 2012 data as a function of momentum, pseudorapidity, track multiplicity and number of primary vertices. The average efficiency is above 96 % in the momentum range $5 < p < 200$ GeV/c. It is worst in high multiplicity events with $N_{\text{track}} > 200$.

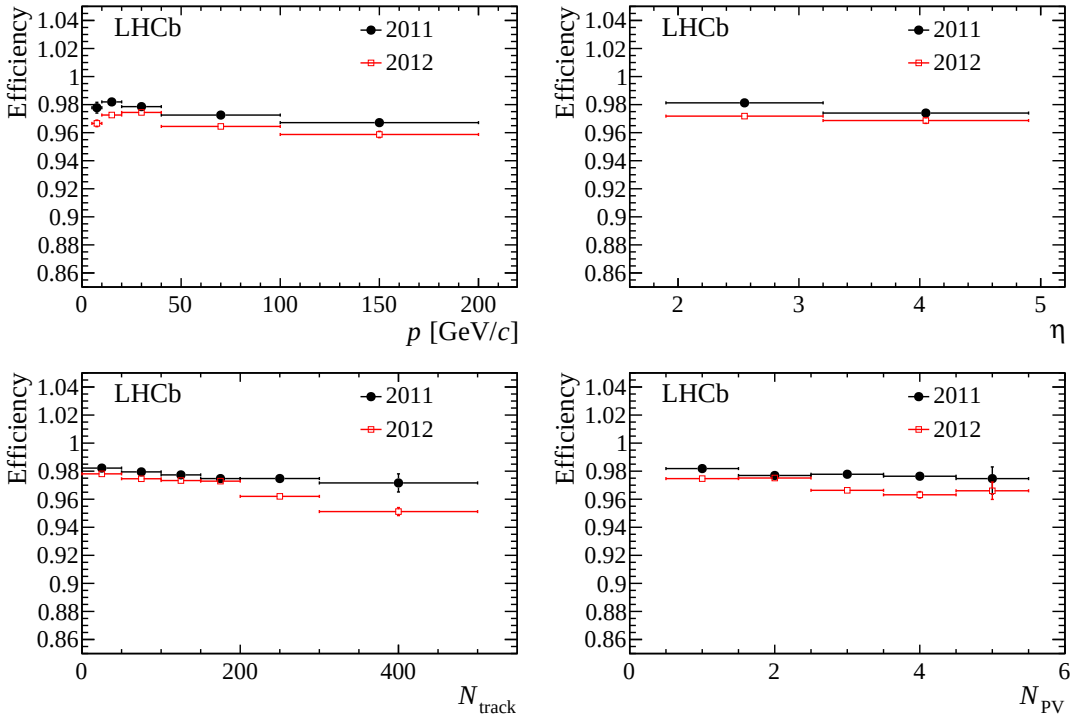


Figure 3.15 *Tracking efficiency as a function of momentum (top left), pseudorapidity (top right), multiplicity (bottom left) and number of primary vertices (bottom right) in 2011 and 2012 data [106].*

The momentum resolution for long tracks is studied using $J/\psi \to \mu^+\mu^-$ decays [106]. The mass resolution of the J/ψ meson is dominated by the momentum resolution of the muons. Ignoring the muon mass and considering the case where both muons have similar momentum, the momentum resolution ($\delta p/p$) in this case is

$$\left(\frac{\delta p}{p}\right)^2 = 2\left(\frac{\sigma_m}{m}\right)^2 - 2\left(\frac{p\sigma_\theta}{mc\theta}\right)^2,$$

where m is invariant mass, σ_m is the Gaussian width obtained from a fit to the invariant mass, θ is the opening angle between the muons, and σ_θ is the error on θ, obtained from track fits of the muons. The plot on the left of Figure 3.16 shows relative momentum resolution as a function of momentum. The resolution ranges from 0.5 % for tracks below 20 GeV/c to 1.1 % for tracks above 160 GeV/c.

Fits are performed to the invariant mass distributions of six $\mu^+\mu^-$ resonances —

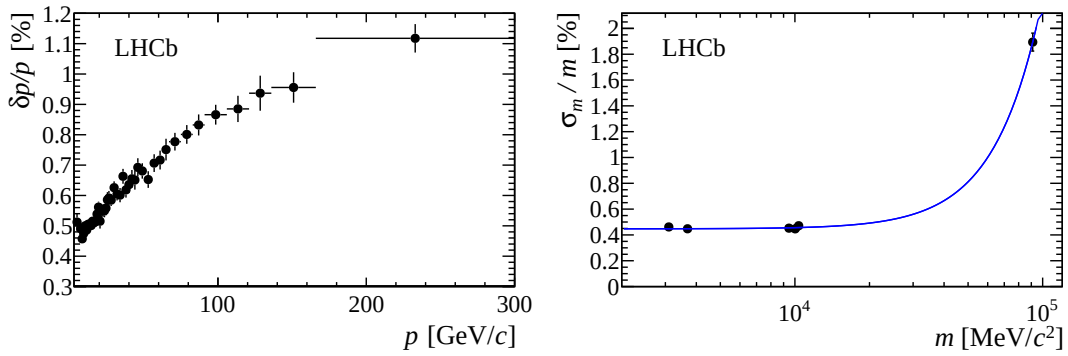


Figure 3.16 *Left: relative momentum resolution of muon tracks from $J/\psi \to \mu^+\mu^-$ decays as a function of track momentum. Right: relative mass resolution from several $\mu^+\mu^-$ resonances as a function of invariant mass. The solid curve is an empirical power-law fit through the points [106].*

the J/ψ, $\psi(2S)$, $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$ and Z^0 — in order to obtain the relative mass resolution [106]. The results are shown on the right of Figure 3.16. It can be seen that the relative mass resolution is about 0.5 % up to $\mathcal{O}$ (10 GeV/c^2).

3.5 Particle identification

In this section, the particle identification (PID) system of LHCb is described. Ring-imaging Cherenkov detectors are employed to identify long-lived charged hadrons: namely pions, kaons and protons. The calorimetry system is used to identify and measure the energy of photons, electrons and hadrons. The muon detector system is used to identify and measure the momentum of muons.

3.5.1 Ring imaging Cherenkov detectors

LHCb contains two RICH detectors [103] for the purpose of hadron identification. These detectors are optimised for distinguishing between protons, charged pions and charged kaons.

Cherenkov radiation [108] is emitted when a charged particle passes through a dielectric medium, known as a 'radiator', with a speed, β, greater than $1/n$, where n is the refractive index of the material. The light is emitted at a constant polar angle from the trajectory of the particle. This angle, known as the Cherenkov

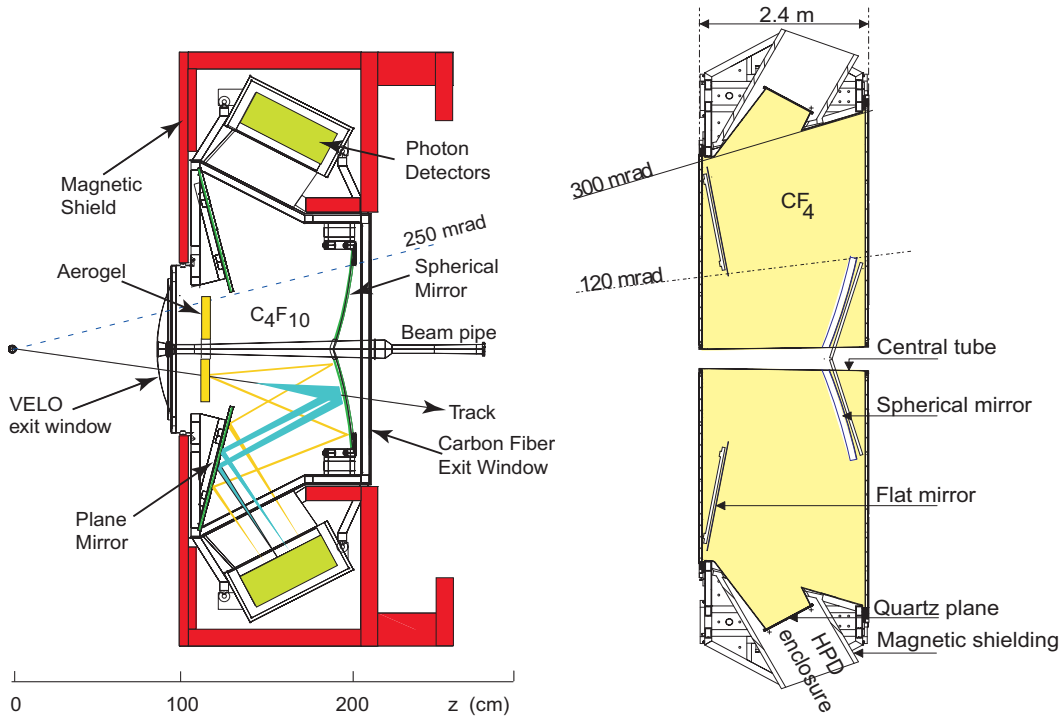


Figure 3.17 *Left: side view of RICH1, illustrating the path of Cherenkov light. Right: top-down view of RICH2. [103]*

angle, θ_c, is a function of n and β:

$$\cos\theta_c = \frac{1}{n\beta}. \tag{3.1}$$

A RICH detector employs spherical mirrors to focus the light into rings. The radius of a Cherenkov ring is the product of θ_c and the focal length of the mirror. Measurement of the radius of these rings gives information on the velocity of the associated particle. When the rings are matched to tracks, which provide momentum information, it is possible to infer the mass of the particle.

The LHCb RICH detectors, shown in Figure 3.17, have gas-tight volumes filled with fluorocarbon gas radiators. There are two detectors, separated by the TT, IT/OT tracking stations and the magnet. They have different angular coverages and cover different momentum ranges. Figure 3.18 shows the relation between Cherenkov angle and momentum for various particle species in each of the radiators used in the LHCb RICH detectors.

The RICH1 detector, shown on the left of Figure 3.17, is positioned between the VELO and the TT. During Run 1, it used both silica aerogel and C_4F_{10} gas as Cherenkov radiators. The aerogel was removed prior to Run 2. RICH1 has an

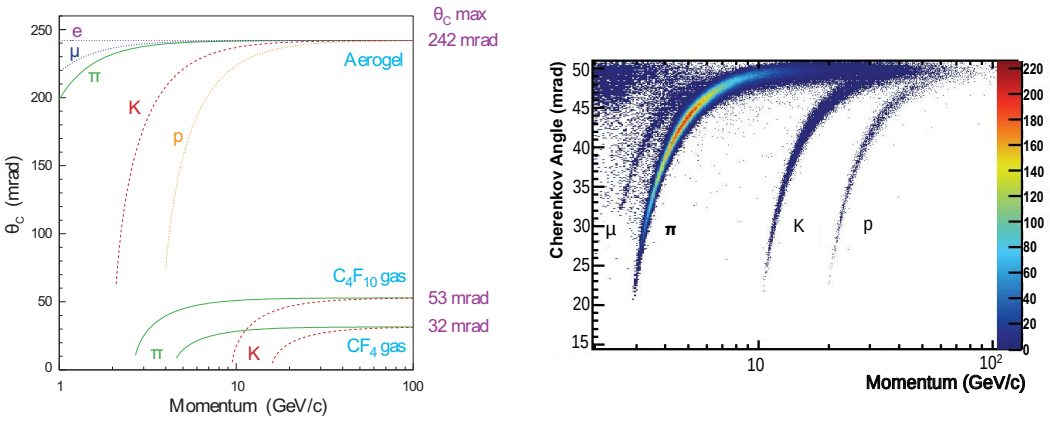


Figure 3.18 *Left: Cherenkov angle as a function of momentum for various charged particle species in each of the RICH radiators. Right: reconstructed Cherenkov angle from isolated tracks in* C_4F_{10} *as a function of track momentum, using data from RICH1 [109].*

angular coverage from 25 mrad to 250 mrad in the vertical plane and 300 mrad in the horizontal. It is designed to cover the low-momentum range from 2–60 GeV/c.

The RICH2 detector, shown on the right of Figure 3.17, is located after the third tracking station and before the first MUON plane. It uses CF_4 gas as its Cherenkov radiator. The angular coverage extends from 15 mrad to 100 mrad in the vertical plane and 120 mrad in the horizontal. This allows the detector to cover the momentum range from 15–100 GeV/c.

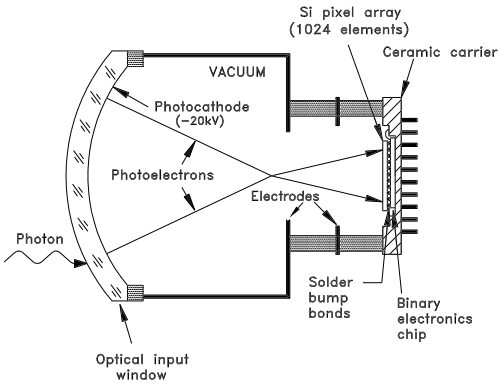


Figure 3.19 *Schematic of an HPD [103].*

Hybrid Photon Detectors (HPDs) [110] are used to detect the Cherenkov light. A diagram of an HPD can be seen in Figure 3.19. HPDs are vacuum phototubes with silicon pixel sensors placed at the anode end. The entrance window is a spherically shaped piece of quartz, with the inner surface coated with a multi-alkali photocathode. When Cherenkov photons with a wavelength in the range 200–600 nm strike the photocathode, photoelectrons are released and accelerated by

electrostatic fields to the pixel sensor, which produces an average of one electron–hole pair per 3.6 eV of deposited energy. The operating voltage of the HPDs is −20 kV, corresponding to around 5000 electron–hole pairs per photoelectron. The RICH detectors are enclosed in iron boxes to mitigate the effects of the field produced by the magnet. Furthermore, the HPDs are housed in individual cylinders of mu-metal[2] within an external box of iron. The magnetic shielding allows the HPDs to operate up to a magnetic field strength of 50 mT, well above the maximum field strength during operation (30 mT at RICH1).

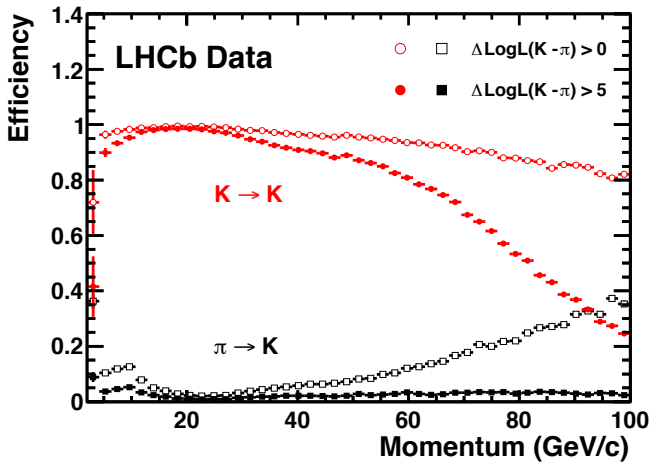


Figure 3.20 *Kaon identification efficiency (red) and pion-as-kaon misidentification rate (black) as a function of track momentum for two different cuts on* $\Delta \log \mathcal{L}(K - \pi)$ *[109].*

The performance of the RICH detectors is measured using control samples of $K^0_S \to \pi^+\pi^-$, $\Lambda \to p\pi^-$ and $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ decays, which were selected using only kinematic requirements [109]. With these, the pion, proton and kaon identification efficiencies and misidentification rates were calculated for a range of track momenta. Figure 3.20 shows the kaon identification efficiency and $\pi \to K$ misidentification rate for two different PID requirements. In the momentum range between 2 GeV/c and 100 GeV/c, for a requirement that the likelihood of the track being a kaon is larger than it being a pion, the kaon identification efficiency is 95 % with a pion misidentification rate of 10 %.

3.5.2 Calorimeters

The calorimetry system [104] is located between the first and second muon detector stations. Its purpose is to discriminate between electrons, photons and hadrons,

[2] A nickel–iron alloy with high magnetic susceptibility.

as well as provide measurements of their energy and position. It also plays an important role in providing transverse energy measurements for the hardware-level trigger described in Section 3.6.1. LHCb adopts the conventional approach of having an electromagnetic calorimeter (ECAL) followed by a hadronic calorimeter (HCAL). Two additional detectors, the Scintillator Pad Detector (SPD) and Preshower Detector (PS) are placed in front of the ECAL in order to help reduce the background for the hardware-level electron trigger caused by the large numbers of pions. In order to fully contain showers from high energy photons, the thickness of the ECAL is 25 radiation lengths, on top of the 2.5 of the SPD/PS. The HCAL thickness is 5.6 interaction lengths as a trade-off between shower containment and available space. The calorimeters have an angular coverage from 25 mrad to 250 mrad in the vertical plane and 300 mrad in the horizontal, matching that of the RICH1. Since the hit density close to the beam pipe is about two orders of magnitude larger than at the edge of the acceptance, each of the subdetectors in the calorimetry system has variable segmentation, shown in Figure 3.21, with smaller cells located in the inner region.

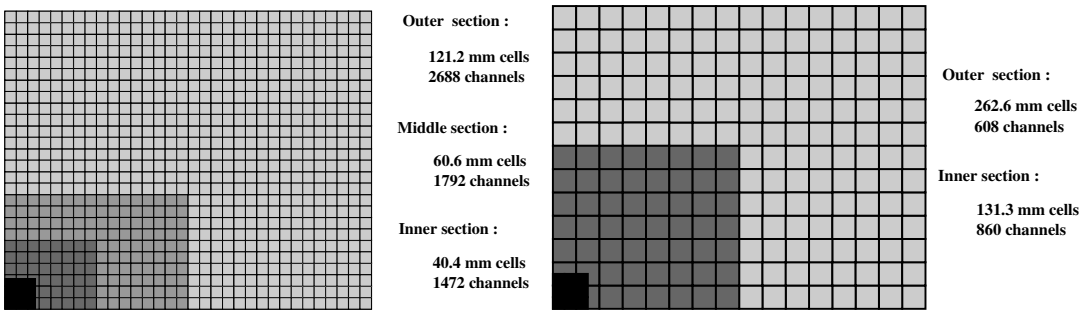


Figure 3.21 *Segmentation of (left) the SPD, PS and ECAL and (right) the HCAL [104].*

Both the ECAL and HCAL are sampling calorimeters, constructed from alternating layers of absorber and detector material. The absorber material is lead in the ECAL and iron in the HCAL. The detector material is a plastic scintillator made from doped polystyrene. Particles interact with the absorber material, generating showers of secondary particles, which in turn induce scintillation light in the detector material. This light is transmitted through wavelength-shifting (WLS) fibres to photomultipliers (PMTs). The SPD/PS cells are read out using multi-anode photomultipliers (MaPMTs) mounted outside the acceptance of the detector inside magnetic shielding. The ECAL and HCAL cells are read out by individual PMTs mounted directly on the detector modules.

The SPD/PS detector consists of two near-identical planes of rectangular

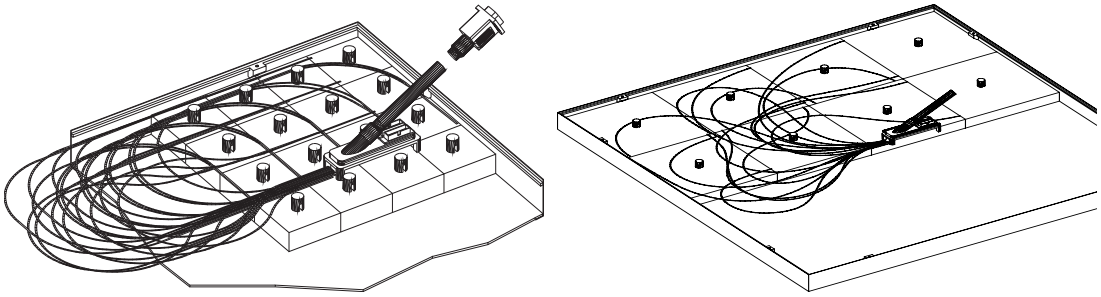


Figure 3.22 *Diagrams of (left) inner and (right) outer SPD/PS pads, showing the routing of the WLS fibres [104].*

scintillator pads, with a 15 mm layer of lead sandwiched between them. The scintillator pads are illustrated in Figure 3.22. The segmentation of the detector in θ and ϕ is chosen to exactly match that of the ECAL. The SPD detects charged particles and is used in conjunction with the ECAL to discriminate between electrons and neutral particles such as π^0. The PS, located after the layer of lead, helps distinguish between electrons and charged hadrons, such as $\pi^{\pm}$, as the latter deposit less energy than the former.

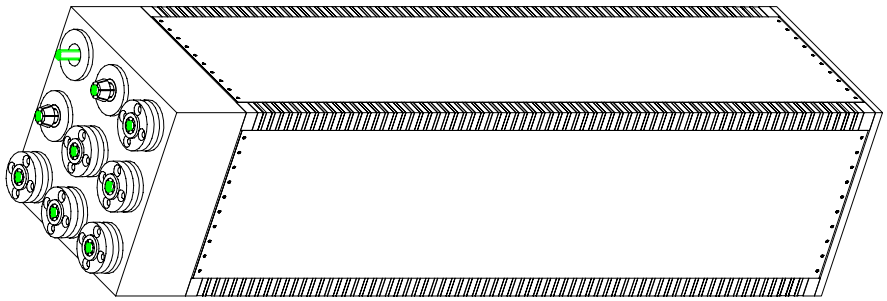


Figure 3.23 *An inner ECAL module [104]. The nine fibre bundles can be seen on the left.*

The ECAL consists of modules built from alternating layers of lead, scintillator tiles and reflecting TYVEK paper[3]. An ECAL module is shown in Figure 3.23. Modules in the inner region are divided into 9 cells; those in the middle region are divided into 4 cells, and the outer modules are a single cell. Light is collected from the scintillator tiles by WLS fibres, which traverse the entire module. The fibres from each cell are bundled together at the back end of each module and feed light into one PMT per bundle.

The HCAL is made from layers of iron absorber and scintillating tiles. Unlike most calorimeters, the tiles run parallel to the beam axis. The tiles are interspersed laterally by 10 mm of iron, whereas the length of the tiles and iron spacers

[3]TYVEK is a brand of synthetic paper manufactured by DuPont.

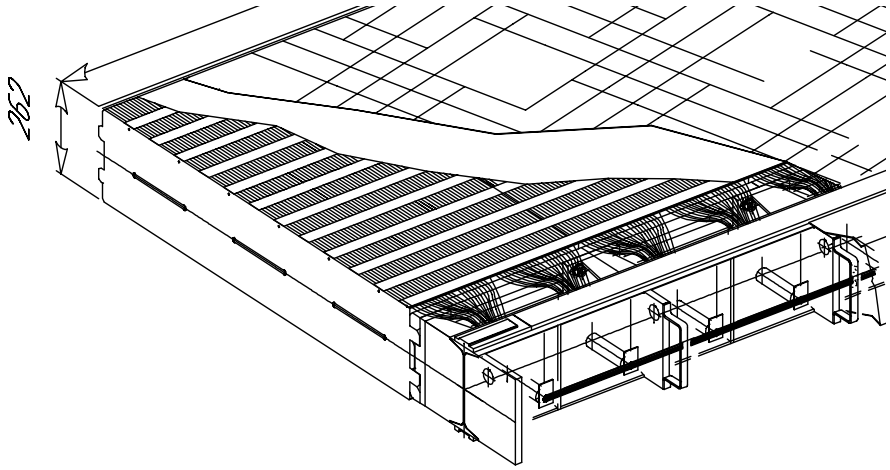


Figure 3.24 *Cut-away view of several outer HCAL submodules assembled within a full module [104].*

corresponds to the hadronic interaction length in steel. A module is composed of 216 identical periods with a thickness of 20 mm. Part of an HCAL module is shown in Figure 3.24. One period consists of an absorber structure, made from laminated steel plates, with gaps where the scintillator tiles are inserted. As with the ECAL, scintillator light is collected by WLS fibres and delivered to PMTs at the back of each module. Cells are defined as groups of tiles whose fibres are bundled together. The cells in the inner region are a quarter of the size of those in the outer region.

3.5.3 Muon detectors

The muon detection system (MUON) [105], illustrated in Figure 3.25, consists of five stations (M1–M5) of detectors. Station M1 is located between RICH2 and the calorimeters and is used to improve the measurement of transverse momentum for the trigger. The stations M2 to M5 are located immediately after the HCAL. Iron absorbers with a thickness of 80 cm are placed between each detector station after M2. The minimum momentum for a muon to traverse the entire LHCb detector to M5 is approximately 6 GeV/c. The total absorber thickness is about 20 interaction lengths.

The detectors provide spatial point measurements for charged tracks. The first three stations have good horizontal spatial resolution, translating to a transverse momentum resolution of 20 %. The transverse momentum is determined from the slope of the muon track in the horizontal plane, using knowledge of the p_{T}

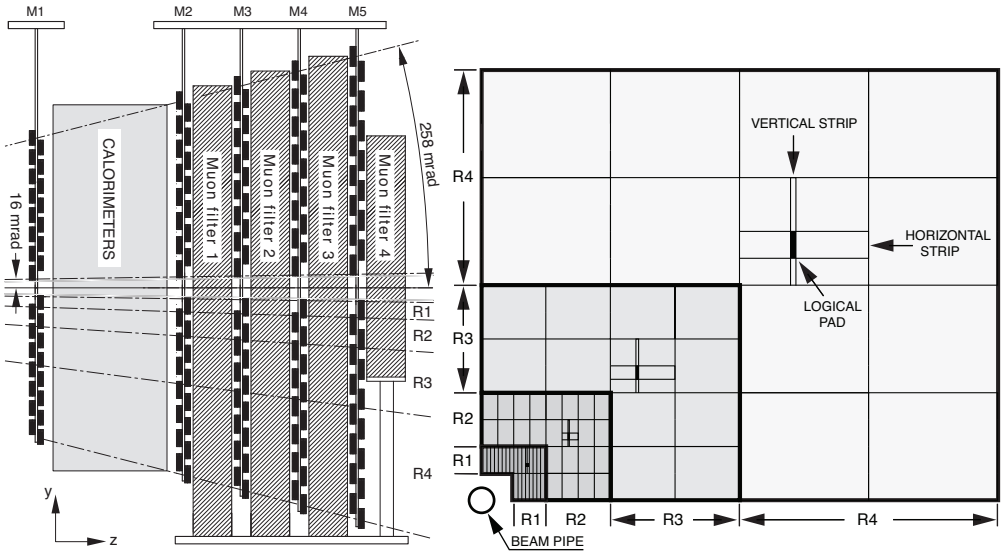


Figure 3.25 *Left: side view of the muon system. Right: segmentation of the muon stations. [105]*

kick of the magnet and assuming that the particle comes from the primary vertex. The other two stations have poor spatial resolution, and are used to identify penetrating particles.

The MUON stations are segmented into pads, each providing binary hit information to the trigger and DAQ. Since the hit density is much higher in the inner region, the stations are divided into four concentric regions from inside to out, with the inner regions being more finely segmented. The segmentation is finer in the horizontal (bending) direction than in the vertical, as illustrated on the right of Figure 3.25, since this is more important for measuring transverse momentum. The stations are scaled transversely in proportion to their distance from the beam crossing point. For all except the inner-most region of M1, the pads are composed of multi-wire proportional counters (MWPCs). The remaining pads are triple gas electron multipliers (GEMs).

3.6 Trigger

While the maximum bunch crossing frequency of the LHC is 40 MHz, the nominal frequency of interactions visible to the LHCb detector, assuming $\sqrt{s} = 14$ TeV and $\mathcal{L} = 200\ \text{nb}^{-1}\,\text{s}^{-1}$, is about 10 MHz.[4] An interaction is defined as visible if there are at least two charged particles that leave enough hits in the VELO and

[4]The target luminosity during Run 1 was about twice this. See Section 3.2.

tracking stations to be reconstructed. The rate of visible interactions is lower than the maximum bunch crossing frequency due to several factors, including 'abort gaps' in the beam[5] and the fact that many interactions are elastic or diffractive scatters.

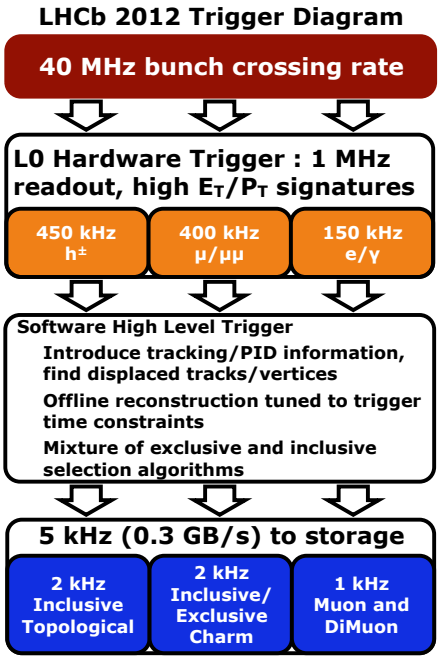


Figure 3.26 *A summary of the trigger layers, decisions and rates for Run 1.*

The role of the trigger [111] is to reduce the rate at which events are read out of the detector and written to storage to $2-5$ kHz. This is done at two levels: the hardware 'level zero' trigger (L0) and the software 'high-level' trigger (HLT). The design of the trigger is illustrated in Figure 3.26. The L0 uses custom electronics which synchronise with the 40 MHz bunch crossing frequency, whereas the HLT runs asynchronously on a processor farm. The nominal rate of $b\bar{b}$ pairs produced within the acceptance of the detector is about 100 kHz. Of these, about 15 % contain a B meson with all of its decay products contained within the acceptance. The branching fractions of decay modes of interest to the experiment are typically below 10^{-3}, hence a selective trigger is needed to select this small fraction of interesting events with high efficiency while also rejecting background events.

A set of conditions which, if satisfied, causes any level of the trigger to pass the event to the next stage, or write it to storage, is referred to as a 'trigger line'.

[5]Gaps in the beam that allow for kicker magnets in the LHC and its injectors to operate.

3.6.1 Hardware trigger

The L0 trigger reduces the rate at which the entire detector is read out to 1 MHz. It is divided into two independent systems: the calorimeter trigger and the muon trigger. These feed into the L0 decision unit [112], which makes the final decision whether to pass the full event to the HLT or not. The decision unit allows for prescaling and the overlapping of several conditions. It has 2 μs in which to make a decision, after accounting for the latency of the electronics, particle flight time and cable lengths. The calorimeter and muon triggers take advantage of the fact that the decay products of b hadrons typically have large transverse energy, E_{T}, and transverse momentum, p_{T}.

Table 3.1 *Typical L0 thresholds used in Run 1 [111].*

Decision	p_{T} or E_{T} threshold		SPD hits
	2011	2012	2011 and 2012
single muon	1.48 GeV/c	1.76 GeV/c	600
dimuon $p_{\mathrm{T}_1} \times p_{\mathrm{T}_2}$	$(1.30\ \mathrm{GeV}/c)^2$	$(1.60\ \mathrm{GeV}/c)^2$	900
hadron	3.50 GeV	3.70 GeV	600
electron	2.50 GeV	3.00 GeV	600
photon	2.50 GeV	3.00 GeV	600

The calorimeter trigger looks for high E_{T} electrons, photons, neutral pions and other hadrons. It forms clusters by adding the E_{T} of 2×2 cells in the ECAL and HCAL and selects the clusters with the highest E_{T}. Clusters are identified as electrons, photons or hadrons using information from the SPD and PS. Hadrons are identified from the HCAL cluster with the highest E_{T}, including the E_{T} from the matching ECAL cluster. Electromagnetic showers are selected from the highest E_{T} ECAL cluster with one or two hits in the PS. Electrons are discriminated from photons by requiring that there be a matching hit in the SPD. The trigger reads out the event if the E_{T} of any candidate is above a certain threshold. Typical values from Run 1 are given in Table 3.1. The total number of hits in the SPD is used to veto high-multiplicity events that would take too long to process in the HLT.

The muon trigger looks for the two highest p_{T} muons in each quadrant. The MUON system is able to perform stand-alone track reconstruction, and the first two stations are used to measure p_{T} in the trigger. Tracks are built from hits that form a straight line through all five stations and must be consistent with originating from the interaction point. An event is triggered if either the highest p_{T} of any muon candidate in the event is above a certain threshold, typically

1.5–1.8 GeV/c, or the product of the highest and second-highest p_{T} is above another threshold, typically (1.3–1.6 GeV/c)^2.

3.6.2 Software trigger

The HLT reduces the event rate from 1 MHz to 2–5 kHz using the full event data. It is split into two layers: HLT1 and HLT2 [113].

The HLT1 step refines candidates found at L0 using reconstructed candidate tracks and vertices from the VELO and tracking stations. The L0 decision unit passes information about which trigger conditions were met, and HLT1 uses this information to decide which sequence of algorithms to run. The algorithms in each sequence work to verify the decision made by the previous step using progressively more of the event information. If multiple L0 candidates were found, then the corresponding sequences will be run independently. To minimise CPU time, the HLT ensures that the same track or vertex is not reconstructed twice. Additional p_{T} and impact parameter requirements reduce the output rate of HLT1 to 80 kHz.

The 'single-track all L0' line [114], used in both analyses presented in this thesis, looks for a single detached high-momentum track. This line starts with VELO tracks and applies cuts on impact parameter, hits per track and the difference between measured and expected number of hits. The selected VELO tracks are used to seed the forward tracking algorithm, described in Section 3.4.7, with a tight search window corresponding to high-momentum tracks. The resulting long tracks are then subject to cuts on the track fit χ^2 and impact parameter χ^2.

The HLT2 step applies a set of inclusive or exclusive final-state selections based on the full event information. The input rate to HLT2 is low enough that a simplified version of the full offline event reconstruction can be run. Tracks which pass loose kinematic cuts are used to reconstruct candidates for common decays, such as $K^{*0} \to K^+\pi^-$, $\phi \to K^+K^-$, $J/\psi \to \mu^+\mu^-$, which are used in the subsequent selections.

The HLT2 stage has a large number of exclusive trigger lines, designed to search for specific processes. Since the lines require full reconstruction of decay modes, they are sensitive to the slightly poorer track resolution resulting from the simplified reconstruction that runs in the trigger. However, due to their smaller rate compared to inclusive lines, exclusive lines can use a more relaxed set of cuts.

Since the rate of visible $c\bar{c}$ production is much higher than that of $b\bar{b}$, the trigger lines designed to look for c hadrons are exclusive lines with tight cuts.

Inclusive HLT lines trigger on partially reconstructed decays. This is achieved by searching for typical distinguishing properties of b hadron decays, such as displaced decay vertices and high-p_{T} tracks, as well as final-state particles present in key channels, such as the ϕ or J/ψ. Since inclusive lines do not include cuts on the b hadron invariant mass, they are less sensitive to the quality of the track reconstruction than exclusive lines.

Inclusive 'topological' trigger lines [115] are used in both analyses presented in this thesis. These are designed to trigger on partially reconstructed b hadron decays with at least two charged particles in the final state and a displaced decay vertex. Tracks are selected using fit quality, impact parameter and particle identification information. Vertices are constructed using combinations of two, three or four of the selected tracks. Signal candidates are selected based on several kinematic and isolation variables using a multivariate algorithm, trained on simulated signal events and collision data recorded in 2010.

There are several inclusive HLT2 lines which look for particular decay products of key channels. One of these is the inclusive ϕ line, which is used in the analysis presented in Chapter 5. The trigger line constructs $\phi \to K^+K^-$ candidates from track pairs and applies selection requirements in two steps. The first step cuts on the p_{T}, impact parameter χ^2 and track fit χ^2 of the kaons, as well as the p_{T}, invariant mass, vertex fit χ^2 and distance of closest approach to the primary vertex of the ϕ candidate. Then RICH PID requirements are applied in the form of a cut on the difference in log-likelihood of the tracks being pions versus kaons.

3.6.3 TIS and TOS

Trigger decisions at LHCb are classified into three categories: 'trigger on signal' (TOS), 'trigger independently of signal' (TIS) and 'trigger on both' (TOB). TIS events are those which are triggered independently of the presence of the signal decay. TOS events are those which are triggered on the signal decay independently of the presence of the rest of the event. TOB events are those which require both the signal and the rest of the event in order to trigger. The latter are typically caused by a combination of a signal track with a ghost track forming a displaced vertex [116].

3.7 Software

The LHCb software [117] consists of several applications constructed in the Gaudi framework [118]. This section describes the applications used for event simulation (Gauss), digitisation (Boole), reconstruction (Brunel) and analysis (DaVinci). The HLT, described in Section 3.6.2, is also written in the Gaudi framework. The flow of data between the applications is illustrated in Figure 3.27.

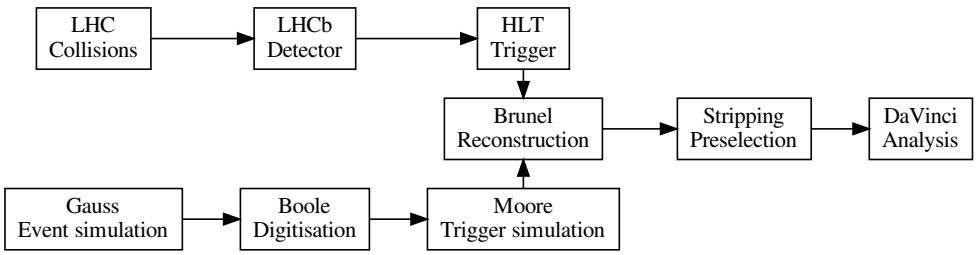


Figure 3.27 *Data flow between the LHCb data processing applications.*

3.7.1 Simulation

The LHCb event simulation is performed by two applications: Gauss [119], which generates the initial particles and simulates their passage through the LHCb detector, and Boole, which simulates the detector response and outputs data in the same format as the LHCb readout electronics.

Gauss simulates events in two independent steps: event generation and detector simulation. The event generation phase uses tools built from external libraries with a generic interface. By default, the generator simulates pp collisions using Pythia 8 [120, 121] and B meson decays using EvtGen [122]. The latter was originally developed by the BaBar collaboration for B mesons produced in the process $e^+e^- \to \Upsilon \to B\overline{B}$. It was modified by the LHCb collaboration to simulate B mesons produced in pp collisions. The output of this step is referred to in Chapters 5 and 6 as 'generator-level' events. The detector simulation phase uses Geant4 [123, 124] to simulate the interactions that take place between the detector material and the particles produced in the collision. The output of this step takes the form of 'hits' in the subdetector sensors.

Digitisation is the conversion of simulated subdetector hits to the same output

format as the readout electronics. This is performed by an application called Boole, which simulates the responses of the subdetector sensors, the readout electronics and the L0 trigger hardware. The application accounts for the effects of bunch spacing, noise, cross-talk and dead channels. After simulating the detector response, Boole then simulates the L0 trigger decision. It is possible to configure the application to output all events or just those which pass the L0. The HLT decisions are added to the event information using a package called Moore [113], which runs the same applications used during real data taking.

3.7.2 Reconstruction

The 'raw' information from the readout electronics — whether real or simulated — is reconstructed using Brunel [117]. This application has access to the same information about the detector geometry as the simulation, ensuring consistency between the two. The main reconstruction step begins with forming 'clusters' of hits in the tracking detectors, which are passed to the tracking algorithms described in Section 3.4.7. Reconstructed tracks are then passed to particle identification algorithms which use information from the calorimeters, RICH detectors and muon detectors. A second step is performed for simulated events, which attempts to associate tracks with generator-level particles. If more than a predetermined fraction of clusters used to build a given track are from the same generator-level particle, the track is associated with that particle, otherwise the track is classified as a 'ghost'. If multiple tracks are associated with the same particle, they are classified as 'clones'.

3.7.3 Analysis

DaVinci [117] is the application which converts the reconstructed events from Brunel to a format useful for physics analysis. The application reconstructs primary vertices, assigns particle identification hypotheses to tracks and calorimeter clusters and performs event selection (see Section 4.4). Several additional tools are available to analyse events within the application, including constrained vertex fitting, flavour tagging, calculation of decay angles and track momentum smearing and scaling. For the analyses presented in this thesis, DaVinci is used to construct candidate $B^0 \to \phi K^{*0}$, $B_s^0 \to \phi\phi$ and $B_s^0 \to \phi K^+ K^-$ decays.

Chapter 4

Event selection

4.1 Introduction

This chapter outlines the event selection strategy common to both analyses presented in this thesis. The selection of charmless b hadron decays at LHCb occurs in several steps. First, to reduce combinatorial background, a cut-based selection imposes requirements on track quality, vertex quality, kinematic and geometric variables. To reduce backgrounds from misidentified particles, cuts are made on particle identification variables. Specific decay modes are vetoed by removing events, or imposing stricter requirements, in invariant mass windows under the appropriate mass hypothesis. Finally, a multivariate algorithm is trained to discriminate between signal and combinatorial background.

4.1.1 Kinematic variables

Transverse momentum, p_{T}, is the component of momentum perpendicular to the beam axis. It is defined in Cartesian coordinates, with the beam axis corresponding to the z axis, as

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}.$$

Due to the large mass of b hadrons, their decay products tend to have large transverse momentum.

Pseudorapidity is a measure of the polar angle, θ, relative to the beam axis. It is

defined as

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

The invariant mass of multiple bodies (*e.g.* the decay products of a B meson) can be found from the sums of their energy and momenta. Cutting on invariant mass is useful for selecting specific parent or intermediate particles in a decay.

4.1.2 Track and vertex quality variables

The final step of the track reconstruction, described in Section 3.4.7, is a Kalman-filter based fit [107], which accounts for multiple scattering and corrects for energy losses due to ionisation. The χ^2 of this fit, known as the track fit χ^2, can be used as a measure of track quality. As described in Section 3.4.7, the probability that a given track is a ghost is determined by an artificial neural network. The output of this can be used to reduce the number background events.

A secondary vertex is reconstructed from a fit to several long tracks. The χ^2 of this fit can be used as a measure of vertex quality. Signal decay candidates, where the tracks really do originate from the secondary vertex, tend to have a better vertex χ^2 than combinatorial background events.

4.1.3 Geometric variables

Several geometric variables can be used to discriminate between signal and background. The first two variables described in this section take advantage of the displaced secondary vertices typical of B decays to identify signal candidates. The second two are measures of compatibility of a reconstructed particle with particular primary or secondary vertices.

The impact parameter (IP) of a particle is the distance of closest approach of its extrapolated trajectory to the primary vertex. This is illustrated as the length of the dotted line in Figure 4.1. The quantity χ^2_{IP} is commonly used in LHCb analyses as a measure of significant displacement from the primary vertex. It is defined as the difference in χ^2 of the primary vertex reconstructed with and without the particle in question. Particles originating from the primary vertex should have a small χ^2_{IP}.

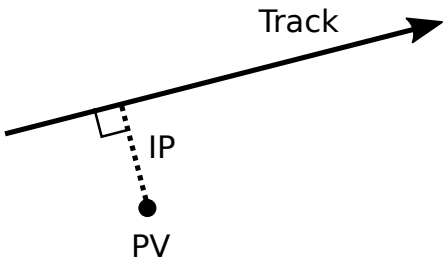


Figure 4.1 *Illustration of the impact parameter (length of the dotted line) of a track (solid line) in relation to the primary vertex (PV). The figure is drawn in the plane containing the track and the primary vertex.*

The distance between a particle's creation and decay vertices, or flight distance (FD), is a variable used to identify displaced secondary vertices. The quantity $\chi^2_{\rm FD}$ can be used as a measure of the significance of the flight distance of a particle. It is defined as the difference in χ^2 of the creation vertex reconstructed with and without the particle in question. Long-lived particles, such as B mesons, have a large $\chi^2_{\rm FD}$.

The distance of closest approach (DOCA) between two tracks can be used as a measure of how compatible they are with originating from the same decay vertex. The quantity $\chi^2_{\rm DOCA}$ is the χ^2 obtained from a vertex fit of the two tracks. It is used in the selection of $\phi \to K^+K^-$ candidates.

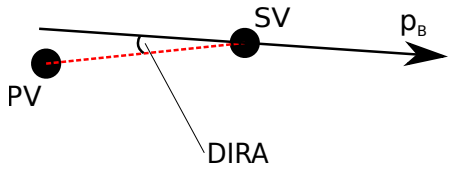


Figure 4.2 *The DIRA as defined as the angle between the B momentum vector (p_B) and the displacement vector (red dotted line) from the primary vertex (PV) to the secondary vertex (SV).*

The 'direction angle' (DIRA) is the angle between the momentum vector of a B candidate and the 3D displacement vector between the associated primary and secondary vertices. This is illustrated in Figure 4.2. It is a measure of consistency of the B candidate with the reconstructed vertices. The cosine of the DIRA is used to separate signal from combinatorial background; it should be close to one for true signal events.

4.1.4 Particle identification variables

Artificial Neural Networks are used to assign a probability that a given track is a kaon (P_K), a pion (P_π) or a proton (P_p) [106]. The input variables used to train the Neural Network contain information from the RICH detectors, combined with information from the tracking system, muon detectors and the calorimeters.

A Boolean variable called 'IsMuon' is used to remove kaons that decay in flight. This variable is built by searching for hits in the muon detector corresponding to a reconstructed track in the vertex detector. If there is at least one corresponding hit in a certain number of muon detector stations, then 'IsMuon' is set to be true [125].

4.2 Dataset

The collision data sample used in both analyses consists of 3 fb^{-1} of proton–proton collisions at $\sqrt{s} = 7$ and 8 TeV recorded by the LHCb detector during 2011 and 2012. This is hereafter referred to as the "full Run 1 dataset". Details of the integrated luminosity for each magnet polarity and value of $\sqrt{s}$ are given in Table 4.1.

Table 4.1 *The proton–proton collision data collected by LHCb during 2011 and 2012 by energy and magnet polarity.*

$\sqrt{s}$ [TeV]	Year	Integrated luminosity [pb^{-1}]	
		Magnet up	Magnet down
7	2011	434	582
8	2012	1033	1024

The Monte Carlo simulation samples used in both analyses are generated at $\sqrt{s} = 8$ TeV, as the differences between the kinematics of events generated at $\sqrt{s} = 7$ and 8 TeV are small. A list of samples and the number of generated events is given in Table 4.2.

Particle momenta in data are scaled by a correction factor, which is calibrated to give the correct pole masses for the J/ψ and B^+ resonances. The procedure ensures the momentum scale is correct to a precision of 3×10^{-4} and slightly improves the mass resolution. To ensure that the momentum resolution agrees between data and simulation, the simulated tracks are 'smeared' by multiplying

Table 4.2 *A list of Monte Carlo simulation samples used in the analyses presented in this thesis, with the number of generated events.*

Decay mode	Events generated
$B_s^0 \to \phi\phi$	2058745
$B^0 \to \phi\phi$	1016796
$B^0 \to \phi K^{*0}$	4027985
$\Lambda_b^0 \to \phi p K^-$ phase space	1052369
$B_s^0 \to \phi K^+ K^-$ phase space	2023924
$B_s^0 \to \phi \pi^+ \pi^-$ phase space	13021214

the components of momentum by random values from a Gaussian distribution with mean of one and a width that depends on the track momentum. This results in an agreement in the track momentum resolution between data and simulation on the level of 1–2 % for daughter particles from B decay modes.

4.3 Trigger lines

In order to have a well-understood efficiency, the offline event selection includes requirements that events pass particular trigger lines. Details of the LHCb trigger can be found in Section 3.6. The sets of trigger lines used in the event selections of the analyses presented in this thesis are almost identical. Both analyses require a TOS decision from the L0 hadron trigger or a TIS decision from the L0 global trigger as well as a TOS decision from the 'single-track all L0' HLT1 line [114][1]. The choice of HLT2 lines differs slightly. The $B_s^0 \to \phi\phi$ branching fraction analysis requires TOS decisions from the two-, three- and four-track 'topological' HLT2 lines [126] as well as the inclusive ϕ line. In the $B_s^0 \to \phi K^+ K^-$ amplitude analysis, the inclusive ϕ line was not included, as it was found to bias the K^+K^- invariant mass distribution.

4.4 Stripping

Selected candidates are built from the information stored in the reconstructed events. To avoid repetition of the computationally intensive process of selecting events from the full dataset, a method of preselection, called 'stripping', is used.

[1] As defined in Section 3.6, TOS means 'trigger on signal' and TIS means 'trigger independently of signal'.

This involves flagging each event in the full dataset as passing or failing various predetermined loose selection criteria ('stripping lines'). Stripping is run as a bulk job at periodic intervals. Thereafter, events can be selected for physics analyses by choosing only those which pass a particular stripping line.

The stripping lines used in the analyses presented in this thesis are designed to select quasi-two-body B meson decays with four tracks in the final state. The line used in the $B_s^0 \to \phi\phi$ branching fraction analysis in Chapter 5 is the 'B_s^0 to quasi-two-body 4π line', which selects $B_s^0 \to (\rho^0 \to \pi^+\pi^-)(\rho^0 \to \pi^+\pi^-)$, but without PID requirements on the tracks or too small an invariant mass window for the track pairs. Hence, it is suitable for selecting $B_s^0 \to \phi\phi$, $B^0 \to \phi\phi$ and $B^0 \to \phi K^{*0}$. The line used in the $B_s^0 \to \phi K^+K^-$ analysis in Chapter 6 is the '$B_s^0 \to \phi\rho^0$ line', which selects $B_s^0 \to (\phi \to K^+K^-)(\rho^0 \to \pi^+\pi^-)$ with no PID requirements on the pions and an invariant mass window for the ρ^0 candidate which covers the full phase space of a $B_s^0 \to \phi h^+h^-$ decay.

The lines use pairs of long tracks to form ϕ and ρ^0 candidates, which are in turn used to form B_s^0 candidates. The tracks are subject to cuts on ghost probability, p_{T}, χ^2_{track} and minimum χ^2_{IP}. Cuts on p_{T}, p, χ^2_{vertex}, χ^2_{IP} and invariant mass are applied to the $\rho^0 \to \pi^+\pi^-$ and $\phi \to K^+K^-$ candidates. Additional χ^2_{DOCA} and PID requirements are applied to the $\phi \to K^+K^-$ candidates. Finally, the B_s^0 candidates are subject to cuts on χ^2_{vertex}, χ^2_{IP} and invariant mass. The numerical values of these cuts can be found in Sections 5.2.3 and 6.2.3.

4.5 Multivariate selection

The data after stripping contains a large contribution from combinatorial and misidentified events. This is reduced by applying more stringent requirements on discriminating variables. Both analyses impose tighter PID requirements on the tracks, *e.g.* $P_K \times (1 - P_\pi) > 0.025$ for kaon candidates, and windows on the two- and four-body invariant masses, *e.g.* $|m(K^+K^-) - m(\phi)| < 15$ MeV/c^2 for ϕ candidates. Following this, multivariate algorithms are used to further reject combinatorial background events. Multivariate algorithms can provide much stronger background rejection and signal retention than simple 'rectangular' cuts on the same variables.

Algorithms provided by the TMVA package [127] are trained to distinguish between

signal and combinatorial background events using multiple discriminating variables. The training samples consist of simulated B decays for the signal and data in a region of invariant mass which is expected to contain purely combinatorial background. The samples are split in two, with one half used for training and the other for testing. The input variables to the MVA algorithms are chosen based on their ability to discriminate between signal and combinatorial background and how well they agree between simulation and data. Care must be taken to avoid 'overtraining', which is a form of bias where the algorithm interprets statistical fluctuations in the training sample as being significant. Overtraining can be identified by a difference in performance between the training and test samples.

The types of MVA algorithm used in this thesis are boosted decision trees (BDT) [128, 129] and multi-layer perceptrons (MLP), a type of artificial neural network (ANN) [130].

A decision tree is a graph whose nodes represent decisions, and the connections between the nodes represent the outcomes of those decisions, much like a flowchart. 'Boosting' is a technique of combining a set of 'weak' classifiers to form a strong classifier[2]. A BDT algorithm builds decision trees in an iterative process based on a weighted dataset. Misclassified events are weighted more strongly for the next tree. The result is a large number of decision trees: each individual tree is a weak classifier, but their combination should in principle be a strong classifier. An event is assigned a 'score' based on that proportion of these trees that classified it as signal.

An ANN consists of a collection of connected 'neurons', which are functions that have a number of inputs and a single output. Each neuron has an associated set of weights that are used to perform a weighted sum of the inputs. The output of the neuron is the 'activation function' evaluated on the weighted sum. Training consists of finding the optimal set of weights for each neuron. A feed-forward neural network consists of ordered 'layers' of neurons, with the outputs of the neurons from one layer feeding into the inputs of those in the next.

A MLP is a feed-forward ANN with at least 3 layers, including the input and output layers. The activation function is typically sigmoid (*i.e.* 'S'-shaped), *e.g.* $(1 + e^{-x})^{-1}$ or $\mathrm{erf}(x)$, since these are useful for generally approximating non-linear functions [131]. Training is typically performed using back-propagation [132],

[2]A classifier is an algorithm that identifies which category the input belongs to. In the context of event selection, a classifier decides whether an event is signal or background. A weak classifier is one that performs slightly better than random chance.

which is an iterative procedure that adjusts weights based on how much the output of the network differs from the desired output.

The performance of an algorithm can be determined from its curve in the space formed by background rejection and signal efficiency, called a receiver operating characteristic (ROC) curve. Better performance is indicated by the curve's proximity to the ideal performance of 100% background rejection and signal efficiency.

Optimal cuts on the MVA algorithm output are determined by maximising a figure of merit, the choice of which depends on the goal of the analysis. A common figure of merit, used in both analyses presented in this thesis, is the 'signal significance' figure of merit: the ratio of the number of signal candidates to the Poisson error on the sample size. For cases where the number of signal candidates is unknown, such as with $B^0 \to \phi\phi$, the 'Punzi' figure of merit [133] is used. This uses the signal efficiency and Poisson error on the number of background events.

Chapter 5

Measurement of the $B_s^0 \to \phi\phi$ branching fraction and search for the decay $B^0 \to \phi\phi$

5.1 Introduction

This chapter details an analysis that is published in Ref. [1].

As discussed in Chapter 2, the decay $B_s^0 \to \phi\phi$ is a flavour-changing neutral current transition. It is therefore forbidden at tree-level in the Standard Model. To leading order, this decay proceeds by a $\overline{b} \to \overline{s}s\overline{s}$ strong penguin diagram. As such, it is potentially sensitive to the effects of new particles entering the loop [7]. The central values of predictions for the $B_s^0 \to \phi\phi$ branching fraction lie in the range (1.3 to 2.6) $\times 10^{-5}$ [59–63]. The predictions are summarised in Table 2.2 and detailed in Section 2.4.5.

The first observation of the decay was made by the CDF collaboration at the Tevatron [134]. Subsequently, CDF measured the branching fraction of this mode using $B_s^0 \to J/\psi\,\phi$ for normalisation [48]. Using the current PDG value of the $B_s^0 \to J/\psi\,\phi$ branching fraction [135], the CDF result is

$$\mathcal{B}(B_s^0 \to \phi\phi) = (1.91 \pm 0.26 \ \text{(stat)} \pm 0.16 \ \text{(syst)}) \times 10^{-5}.$$

The branching fraction of $B_s^0 \to \phi\phi$ is used for normalisation in studies of other

charmless b hadron decays. Therefore, it is important to improve its precision. The Run 1 LHCb dataset contains a significantly higher number of $B_s^0 \to \phi\phi$ candidates than were collected by CDF, so an appreciable reduction in the statistical uncertainty of the branching fraction is expected. This chapter details a measurement of the $B_s^0 \to \phi\phi$ branching fraction relative to that of the $B^0 \to \phi K^{*0}$ decay. Throughout this chapter, K^{*0} refers specifically to the $K^*(892)^0$ meson.

The ratio of the $B_s^0 \to \phi\phi$ branching fraction relative to that of $B^0 \to \phi K^{*0}$ is calculated using

$$\frac{\mathcal{B}(B_s^0 \to \phi\phi)}{\mathcal{B}(B^0 \to \phi K^{*0})} = \frac{N_{B_s^0 \to \phi\phi}}{N_{B^0 \to \phi K^{*0}}} \frac{\varepsilon_{B^0 \to \phi K^{*0}}}{\varepsilon_{B_s^0 \to \phi\phi}} \frac{\mathcal{B}(K^{*0} \to K^+\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} \cdot \frac{1}{f_s/f_d}, \tag{5.1}$$

where the symbols $\mathcal{B}$ represent branching fractions, N are yields obtained from fits to data (Section 5.3), ε are efficiencies obtained through a mix of simulation and data-driven techniques (Section 5.4), and f_s/f_d is the ratio of fragmentation fractions, *i.e.* the ratio of probabilities that a b quark hadronises with a s or a d quark.

The decay $B^0 \to \phi\phi$ is unobserved and highly suppressed in the Standard Model by the OZI rule [65–67]. Branching fraction predictions lie in the range (0.2 to 3.0) $\times$ 10^{-8} [59, 68, 69]. These predictions are summarised in Table 2.3 and detailed in Section 2.4.6. However, this value may be enhanced in models of physics beyond the Standard Model [68]. The previous best upper limit for the branching fraction of the mode is

$$\mathcal{B}(B^0 \to \phi\phi) < 2.0 \times 10^{-7} \text{ at } 90\,\% \text{ confidence level,}$$

set by the BaBar collaboration [136]. Section 5.7 describes a search for this decay mode and the determination of a new upper limit on its branching fraction.

5.2 Event selection

This section describes the process of selecting data and simulation events to be used in this analysis. The goal of the selection is to remove combinatorial and misidentified background events without sacrificing signal efficiency. A more general overview of the event selection strategy is given in Chapter 4.

5.2.1 Dataset

The analysis uses the full Run 1 dataset, described in Section 4.2. Simulated $B_s^0 \to \phi\phi$, $B^0 \to \phi\phi$ and $B^0 \to \phi K^{*0}$ events are used to obtain the parameters of the invariant mass distributions of the signal and normalisation decay modes, as well as the ratio of the selection efficiencies between these modes.

A potentially important background for the signal channel arises from the decay $\Lambda_b^0 \to \phi p K^-$, where the p is misidentified as a K^+. In order to study this, a sample of simulated $\Lambda_b^0 \to \phi p K^-$ events is generated with the p and K^- produced according to phase space, *i.e.* completely non-resonant three-body $\Lambda_b^0 \to \phi p K^-$ decays.

Similarly, two important backgrounds for the normalisation channel are the decays $\Lambda_b^0 \to \phi p K^-$, where the p is misidentified as a π^+, and $\Lambda_b^0 \to \phi p \pi^-$, where the p is misidentified as a K^+. In order to study these, a mixed sample of $\Lambda_b^0 \to \phi p K^-$ and $\Lambda_b^0 \to \phi p \pi^-$ decays, hereby referred to as $\Lambda_b^0 \to \phi p h^-$, is generated according to a phase space distribution, as the resonant structures of these decays are unknown.

5.2.2 Trigger lines

The candidate events used in this analysis are required to pass a set of specific trigger lines. More information on the specifics of each line can be found in Section 4.3. Table 5.1 shows the efficiency of each of these lines on the signal and normalisation simulation samples.

5.2.3 Cut-based selection

The selection of $B_s^0 \to \phi\phi$ and $B^0 \to \phi K^{*0}$ candidates starts from a stripping line designed to select the decay mode $B_s^0 \to (\rho^0 \to \pi^+\pi^-)(\rho^0 \to \pi^+\pi^-)$. More information about the stripping selection can be found in Section 4.4. Due to the wide invariant mass windows and absence of PID requirements, this line is well suited to selecting similar quasi-two-body B meson decay modes with four charged particles in the final state.

The stripping selection proceeds as follows. Well-reconstructed tracks that traverse the entire length of the detector are chosen. Here, the quality of the track

Table 5.1 *The trigger efficiency for the signal and normalisation simulation samples. These efficiencies are calculated using truth-matched simulation events that pass the stripping selection. The total L0 efficiency is the efficiency of events to pass either of the chosen L0 trigger lines. The HLT1 efficiency is calculated after applying the L0 trigger requirements. Similarly, the HLT2 efficiencies are calculated after applying the L0 and HLT1 trigger requirements. The total efficiency is calculated requiring the candidate to pass at least one line from each layer.*

Trigger line	Fraction of events kept [%]	
	$B_s^0 \to \phi\phi$	$B^0 \to \phi K^{*0}$
L0 Hadron TOS	16.34 ± 0.08	23.21 ± 0.08
L0 Global TIS	25.55 ± 0.10	30.53 ± 0.09
Total L0	36.06 ± 0.10	44.79 ± 0.09
HLT1 Track All L0 TOS	78.52 ± 0.15	80.88 ± 0.11
HLT2 Topo 2-Body TOS	62.19 ± 0.20	69.29 ± 0.14
HLT2 Topo 3-Body TOS	84.85 ± 0.15	86.37 ± 0.11
HLT2 Topo 4-Body TOS	76.45 ± 0.17	72.81 ± 0.14
HLT2 Inclusive ϕ TOS	90.89 ± 0.12	66.99 ± 0.14
Total HLT2	98.21 ± 0.05	94.99 ± 0.07
Total	27.81 ± 0.10	34.41 ± 0.09

reconstruction is judged by the χ^2 per degree of freedom (χ^2/ndf) of the track fit, which is required to be less than 4. In addition, the ghost probability for each track, defined in Sections 3.4.7 and 4.1.2, is required to be less than 0.8. The distribution of this variable is shown in Figure 5.1 (left). Tracks originating from the primary vertex are removed by requiring each track to have an impact parameter χ^2 greater than 16. Each track is required to have a transverse momentum, p_{T}, of at least 500 MeV/c. The distribution of this variable is shown in Figure 5.1 (right). Track pairs are combined, under the $\pi^+\pi^-$ mass hypothesis, to form 'ρ^0' candidates, which are required to have $m(\pi^+\pi^-) < 1100$ MeV/c^2, $p > 1000$ MeV/c and $p_{\mathrm{T}} > 900$ MeV/c. The χ^2/ndf of the vertex fit for each ρ^0 candidate is required to be less than 16. Pairs of ρ^0 candidates are combined to form B^0 candidates, which are required to have an invariant mass between 4500 and 5700 MeV/c^2. Each B^0 candidate is required to have a vertex fit χ^2/ndf less than 12 and an impact parameter χ^2 greater than 20. The angle between the reconstructed momentum vector of the B^0 candidate and the displacement vector from the primary vertex to the decay vertex is required to be less than 0.045 radians.

The next step is a tighter cut-based selection specific to each decay mode. Kaons

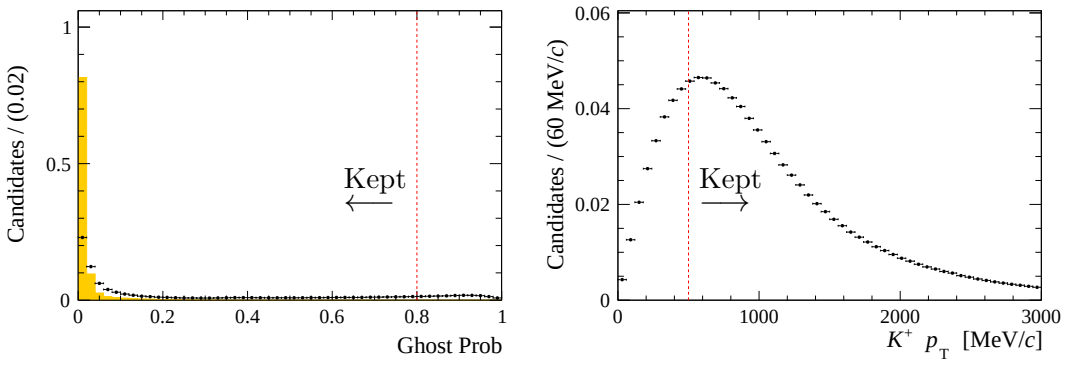


Figure 5.1 *Left: distribution of track ghost probability in selected data events (black points) and simulation (orange histogram). Right: distribution of track p_{T} of kaons from generator-level $B_s^0 \to \phi\phi$ events without cuts. The distributions are normalised to unit area. The red dashed lines indicate the value of the applied cuts.*

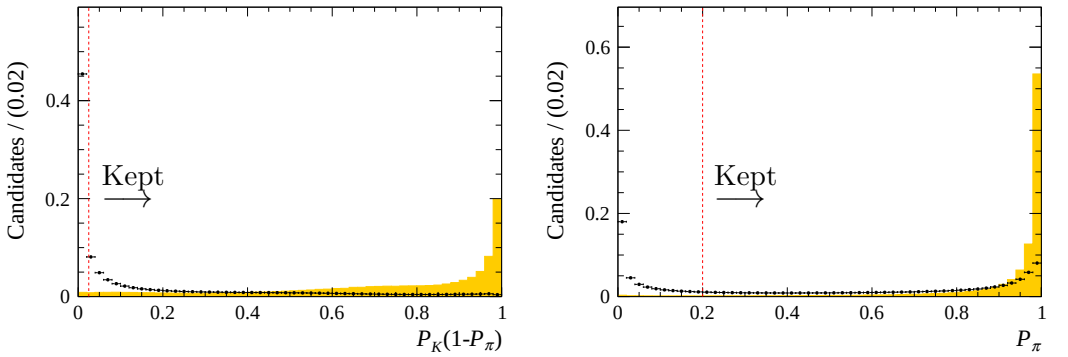


Figure 5.2 *Distributions of the PID variables $P_K(1-P_\pi)$ (left) and P_π (right) in data with loose selection (black points) and simulation (orange histogram), where the simulated particles are true kaons and pions, respectively. The red dashed lines indicate the value of the applied cuts. The distributions are normalised to unit area.*

and pions are identified using the neural-net hadron probability variables, P_h, which are described in Section 4.1.4. Kaon candidates are required to have $P_K(1-P_\pi) > 0.025$ and pion candidates to have $P_\pi > 0.2$. The pion candidate in the $B^0 \to \phi K^{*0}$ selection is chosen to be the particle with the largest value of P_π. Figure 5.2 shows the distributions of these variables for the stripped data and simulation. To ensure the two samples are mutually exclusive, if the pion candidate has $P_\pi > P_K$, it is kept in the $B^0 \to \phi K^{*0}$ sample and rejected from the $B_s^0 \to \phi\phi$ sample. Pairs of particles identified as K^+K^- and $K^+\pi^-$ are combined to form ϕ and K^{*0} candidates. The invariant mass of each ϕ candidate is required to be within 15 MeV/c^2 of the nominal value [135]. Similarly, the mass of each K^{*0} candidate must be within 150 MeV/c^2 of the nominal value [135]. These mass

windows are chosen to match those used in the LHCb analyses of $B_s^0 \to \phi\phi$ [49] and $B^0 \to \phi K^{*0}$ [47], respectively, so that previously measured S-wave fractions can be used without the need to repeat the angular fits (see Section 5.3.3). The invariant mass distributions of the signal and normalisation samples after the cut-based selection are shown in Figure 5.3. It can be seen that the $B_s^0 \to \phi\phi$ selection already results in a small combinatorial background.

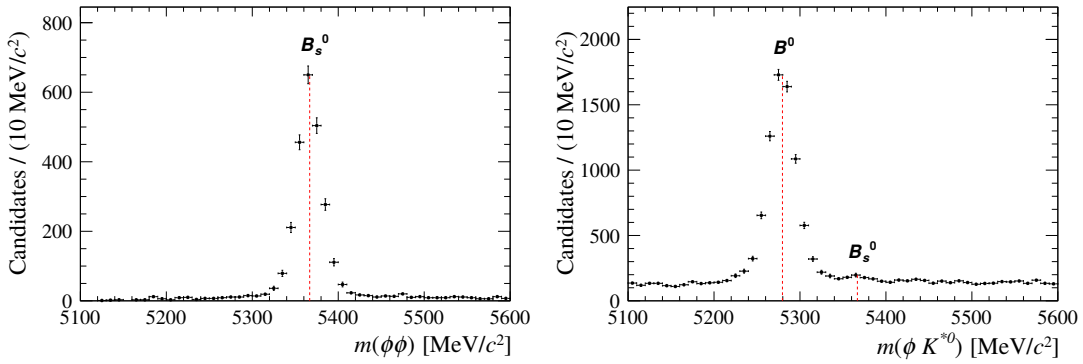


Figure 5.3 *Invariant mass distributions of the data using samples with the cut-based selection applied. Left: $B_s^0 \to \phi\phi$. Right: $B^0 \to \phi K^{*0}$. The red dotted lines indicate the B^0 and B_s^0 masses.*

5.2.4 Multivariate selection

This analysis uses a Boosted Decision Tree classifier (BDT, see Section 4.5) developed for the analysis of the $B_s^0 \to \phi\pi^+\pi^-$ decay published in Ref. [137] and detailed in Ref. [138]. This is trained to distinguish four-body hadronic b hadron decays from combinatorial background without using PID information. The training variables are the $\chi^2_{\rm IP}$, $\chi^2_{\rm Vtx}$, flight distance, lifetime, $p_{\rm T}$ and η of B_s^0 meson, the η and the minimum $p_{\rm T}$ of the K^+K^- and $\pi^+\pi^-$ pairs and the minimum $p_{\rm T}$ and minimum $\chi^2_{\rm IP}$ of the tracks. The signal sample used to train the BDT consists of simulated $B_s^0 \to \phi\pi^+\pi^-$ events, and the background sample consists of the upper sideband of data with the cut-based preselection for $B_s^0 \to \phi\pi^+\pi^-$ applied.

The BDT output in simulation is compared to data using the *sPlot* technique [139]. This is a method of obtaining the distributions of variables for separate components (*e.g.* signal and background) within a data sample. A set of weights — one per component — is assigned to each event in the sample based on a maximum likelihood fit to a 'discriminating' variable, for which the shapes of the components

are known. In this case, these 's-weights' are obtained from fits to the $m(K^+K^-K^+K^-)$ and $m(K^+K^-K^+\pi^-)$ distributions of the data samples, shown in Figure 5.3, using the models described in Section 5.3. The s-weights associated with a certain component can then be used to plot the distributions of other variables for that component only, without *a priori* knowledge of its shape. Figure 5.4 shows the BDT output in the simulation compared to the data, weighted with the s-weights associated with the signal. It can be seen that there is some disagreement between data and simulation, therefore care must be taken when quantifying the efficiency of the cut on the MVA output.

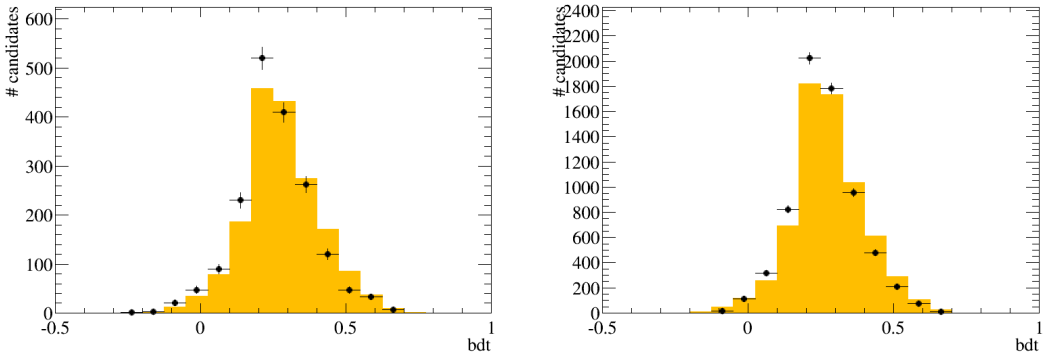


Figure 5.4 *Comparison of the BDT output between the simulation (solid orange) and s-weighted data (black points) using samples with the cut-based selection applied. Left: $B_s^0 \to \phi\phi$. Right: $B^0 \to \phi K^{*0}$.*

The choice of cut on the BDT output is optimised separately for the $B_s^0 \to \phi\phi$ branching fraction measurement and the search for the decay $B^0 \to \phi\phi$. For the branching fraction measurement, the cut on the BDT output is chosen to maximise the 'signal significance' figure of merit for the $B^0 \to \phi K^{*0}$ decay mode and then applied to both the signal and normalisation samples. The figure of merit is defined as

$$FOM = \frac{S}{\sqrt{S+B}}, \tag{5.2}$$

where S is the number of signal candidates, and B is the number of background events in a region ± 40 MeV/c^2 around the B^0 mass, corresponding to a $\sim \pm 2\sigma$ window. It is chosen to optimise for the $B^0 \to \phi K^{*0}$ signal, rather than $B_s^0 \to \phi\phi$, since the combinatorial background is larger, as can be seen in Figure 5.3. The number of $B^0 \to \phi K^{*0}$ candidates in the data before any cut is applied on the BDT output, S_0, is determined to be 6728 ± 97 by fitting the ϕK^{*0} invariant mass after the cut-based selection, using the model described in Section 5.3.2. For a given cut on the BDT output, the number of signal events is $S = S_0 \cdot \varepsilon_S$, where the efficiency to pass the BDT cut, ε_S, is determined using the $B^0 \to \phi K^{*0}$ Monte

Carlo sample. The number of background candidates, B, is obtained by counting the number of candidates in the sideband region $5600 < m(\phi K^{*0}) < 5800$ MeV/c^2 that pass the cut on the BDT output and scaling this number by the relative size of the signal and sideband windows. The signal region is 80 MeV/c^2 wide, giving a scaling factor of 0.4. This procedure assumes that the combinatorial background is flat in invariant mass, which is seen to be the case in Figure 5.5.

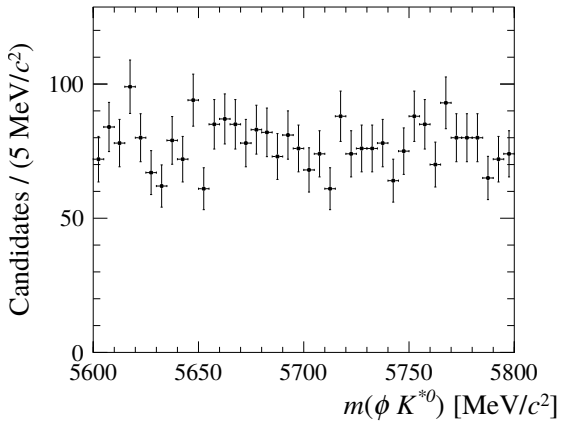


Figure 5.5 *The $m(\phi K^{*0})$ distribution of the sideband sample used in optimising the 'signal significance' figure of merit.*

The figure of merit is plotted as a function of the value of the cut on the BDT output on the left of Figure 5.6. The optimal value of the cut is BDT> -0.025. The effect of this cut on the data can be found by comparing the yields of fits to the data samples using the models described in Sections 5.3 before and after the cut. The signal yield for the $B^0_s \to \phi\phi$ ($B^0 \to \phi K^{*0}$) mode is 2341 ± 50 (6728 ± 97) before the cut and 2309 ± 49 (6680 ± 86) after. The efficiency of this cut on the $B^0_s \to \phi\phi$ ($B^0 \to \phi K^{*0}$) simulation sample is found to be $98.65 \pm 0.05\,\%$ ($98.35 \pm 0.04\,\%$). For comparison, the efficiency of this cut on the $B^0_s \to \phi\pi^+\pi^-$ training sample is $97.6\,\%$.

The efficiency of the BDT cut on the $B^0 \to \phi K^{*0}$ mode disagrees between data and simulation by a factor of $0.65\,\%$. This is assigned as a systematic uncertainty in Section 5.5.7.

In the search for $B^0 \to \phi\phi$, since there is no visible signal in Figure 5.3 and no precise prediction for the branching fraction, the Punzi figure of merit [133] is used. The aim of this figure of merit is to find a region of parameter space (in this case, simply the cut on the BDT output) that maximises the chance of observing a signal above a chosen significance level. This does not require an initial signal yield as input, making it suitable for searches for decay modes where the expected

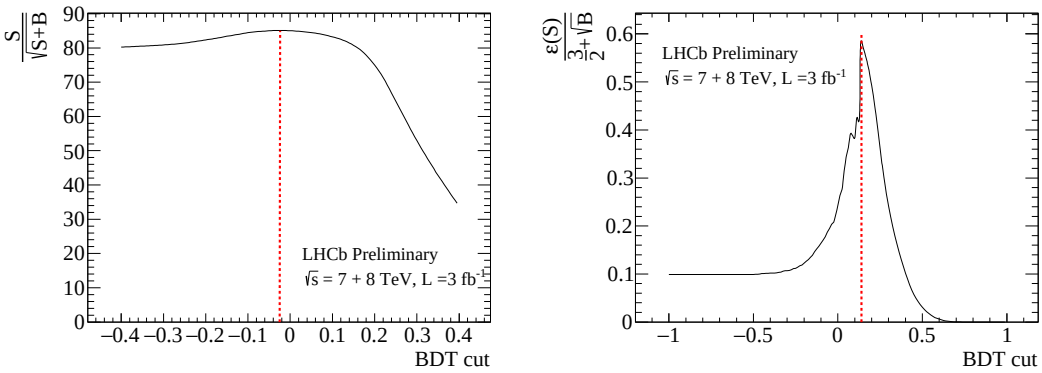


Figure 5.6 *The (left) signal significance figure of merit for $B^0 \to \phi K^{*0}$ and (right) Punzi figure of merit for $B^0 \to \phi\phi$ plotted as a function of the value of the cut on the BDT output. The red dotted lines mark the maxima of the distributions, which correspond to the optimal cut values. The distributions are normalised to the number of events in data.*

signal yield is unknown. It is defined as

$$FOM = \frac{\varepsilon_S}{\frac{n}{2} + \sqrt{B}}, \tag{5.3}$$

where ε_S is the signal efficiency, B is the number of background events, and n is the desired significance in terms of standard deviations. A value of $n = 3$ is chosen in order to optimise for a possible 3σ result, which is the conventional threshold for claiming 'evidence' of a signal. Since a fully simulated $B^0 \to \phi\phi$ sample was not available, the signal sample used to optimise this figure of merit consists of simulated $B_s^0 \to \phi\phi$ events in the window $|m(\phi\phi) - m(B_s^0)| < 40$ MeV/c^2, which relies on the assumption that the MVA cut efficiency the $B_s^0 \to \phi\phi$ and $B^0 \to \phi\phi$ modes are the same. The signal efficiency in Equation 5.3 is the proportion of events in the window that passes a given cut on the BDT output. The background sample is the upper mass sideband of the data (5800 MeV/c^2 < $m(K^+K^-K^+K^-)$ < 6000 MeV/c^2) after the cut-based selection. As with the other figure of merit, the number of background events is found by counting the number of events in this window and multiplying this by a factor of 0.4. The plot on the right of Figure 5.6 shows this FOM as a function of BDT cut. The optimal cut value is BDT> 0.14. The jagged shape of the curve is due to the low background statistics: the initial number of events in the upper mass sideband is 184, and zero background events remain at the value of the optimal cut. The efficiency of this cut is $(86.4 \pm 0.2)\,\%$ on the $B^0 \to \phi\phi$ simulation sample and $(88.3 \pm 0.1)\,\%$ on the $B_s^0 \to \phi\phi$ simulation sample.

5.3 Mass fits for the $B_s^0 \to \phi\phi$ branching fraction

The numbers of candidates in the signal and normalisation modes are found by performing unbinned maximum likelihood fits to the $\phi\phi$ and ϕK^{*0} invariant mass distributions in the fully selected data samples. In both cases, fits are first performed to the simulation samples to obtain the values of the parameters that describe the shapes of the B^0 and B_s^0 peaks. In the fits to simulation, a flat component is added to account for poorly reconstructed candidates and background candidates that have been misidentified as signal. The ratios of the widths of the components in the B^0 and B_s^0 peaks are fixed to the values found from the fits to simulation. In order to account for any difference in mass resolution between the simulation and data, the values obtained in simulation are multiplied by an overall common scaling factor, R_s, which is left free in the fit to data. This should be close to one, due to the momentum resolution smearing described in Section 4.2. The mean values of the peaks, which correspond to the pole masses of the B^0 and B_s^0 mesons, are also left free in the data fit.

5.3.1 The $\phi\phi$ invariant mass fit

The probability density function (PDF) used in the fit to the $\phi\phi$ invariant mass distribution has two components, which account for $B_s^0 \to \phi\phi$ and combinatorial background. The $B_s^0 \to \phi\phi$ signal peak is modelled by the sum of three Gaussian functions with a shared mean,

$$\begin{aligned} P(x|\mu,\sigma_1,\sigma_2,\sigma_3,f_1,f_2) = f_1 P_{\text{Gauss}}(x|\mu,\sigma_1) + f_2 P_{\text{Gauss}}(x|\mu,\sigma_2) \\ + (1-f_1-f_2) P_{\text{Gauss}}(x|\mu,\sigma_3), \end{aligned} \tag{5.4}$$

where P_{Gauss} is a Gaussian function[1], μ corresponds to the B_s^0 mass, $\sigma_{(1,2,3)}$ are the widths of the three Gaussian components in order of increasing width, and $f_{(1,2)}$ are the relative fractions of the two narrower components.

Figure 5.7 shows the fit to the simulation. The corresponding results are shown in the top half of Table 5.2. The combinatorial background is well modelled by a uniform distribution. Alternative models, namely a linear function and an exponential curve, were tried and found to give consistent results. These were

[1]See Appendix B.1 for the explicit form of this function.

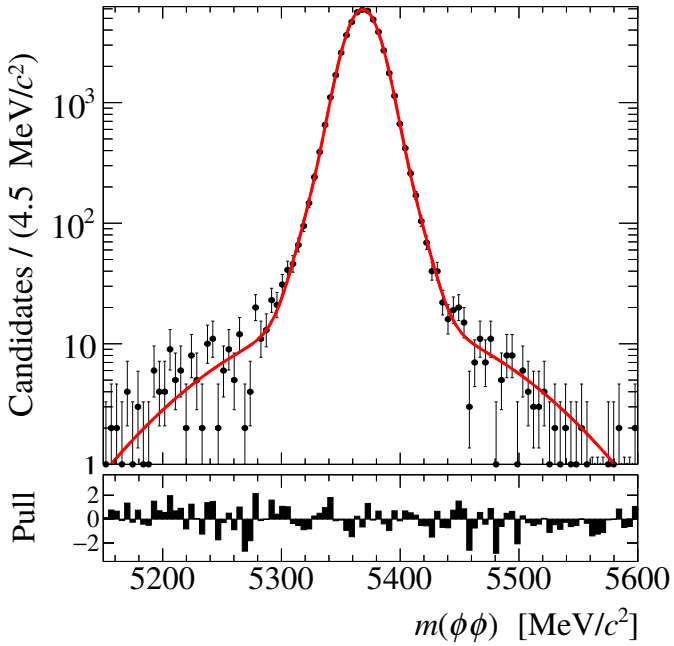


Figure 5.7 *A fit to the $\phi\phi$ invariant mass distribution in the $B_s^0 \to \phi\phi$ simulation sample. The total fitted PDF is shown as a red solid line.*

not used in order to remove an unnecessary degree of freedom from the fit. In the angular analysis of $B_s^0 \to \phi\phi$, [49], misidentified $B^0 \to \phi K^{*0}$ and $\Lambda_b^0 \to \phi p K^-$ backgrounds were also considered. When reconstructing the data, selected for $B_s^0 \to \phi\phi$, under the ϕK^{*0} mass hypothesis, no peak is found at the B^0 mass, and the contribution from this background is therefore assumed to be negligible. This is to be expected, since, as mentioned in Section 5.2.3, events which would pass the $B^0 \to \phi K^{*0}$ selection are rejected from the $B_s^0 \to \phi\phi$ sample. When reconstructing the data under the $\phi p K^-$ mass hypothesis, no peak was found at the Λ_b^0 mass. Therefore, the contribution from $\Lambda_b^0 \to \phi p K^-$ is considered to be negligible. The absence of this background is due to the tighter selection in this analysis compared to the angular analysis.

The B_s^0 mass, μ, and the yields, $N_{\{S,B\}}$, of the signal and background components are left free in the fit to data. Figure 5.8 shows the fitted PDF overlaid onto the data. The results of the fit are given in the bottom half of Table 5.2. The fitted value of μ is consistent with the world average value of $m(B_s^0) = (5366.82 \pm 0.22)$ MeV/c^2 [135]. The fitted value of R_s is consistent with unity.

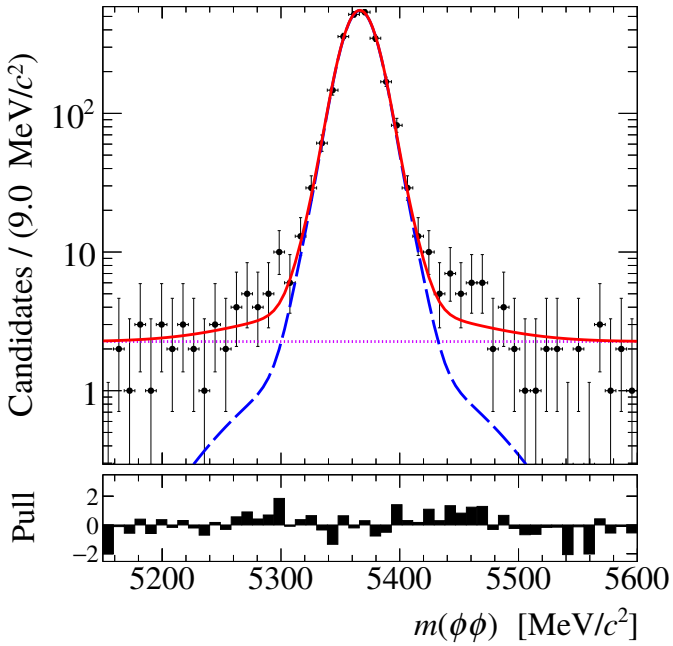


Figure 5.8 *A fit to the $\phi\phi$ invariant mass distribution in the data. The total fitted PDF as described in the text is shown by the (red) solid line, the $B_s^0 \to \phi\phi$ component by the (blue) long-dashed line, and the combinatorial background as the (purple) dotted line.*

5.3.2 The ϕK^{*0} invariant mass fit

The PDF used in the fit to the ϕK^* invariant mass distribution contains four components, which account for $B^0 \to \phi K^{*0}$, $B_s^0 \to \phi K^{*0}$, misidentified $\Lambda_b^0 \to \phi p h^-$ decays and combinatorial background. The $B^0 \to \phi K^{*0}$ peak is modelled by the sum of a Crystal Ball [140][2] and two Gaussian functions with a shared mean

$$\begin{aligned} P(x|\mu, \sigma_{\rm CB}, \sigma_1, \sigma_2, f_{\rm CB}, f_1, \alpha, n) = f_{\rm CB} P_{\rm CB}(x|\mu, \sigma_{\rm CB}, \alpha, n) + f_1 P_{\rm Gauss}(x|\mu, \sigma_1) \\ + (1 - f_{\rm CB} - f_1) P_{\rm Gauss}(x|\mu, \sigma_2), \end{aligned} \tag{5.5}$$

where the symbols for the Gaussian components have the same meaning as in Section 5.3.1. The parameters denoted with the subscript CB are associated with the Crystal Ball function.

The tail of the Crystal Ball distribution accounts for radiative decays. The widths, $\sigma_{(\rm CB,1,2)}$, and relative fractions, $f_{(\rm CB,1)}$, of these components, as well as the Crystal

[2] See Appendix B.2 for the explicit form of this function.

Table 5.2 *Results of the fits to the $B_s^0 \to \phi\phi$ simulation and data samples.*

Simulation fit parameter	Result
σ_1	12.9 ± 0.1 MeV/c^2
σ_2	21.4 ± 0.4 MeV/c^2
σ_3	73.4 ± 5.2 MeV/c^2
f_1	$(71.4 \pm 1.2)\,\%$
f_2	$(27.0 \pm 1.2)\,\%$
Data fit parameter	Result
μ	5366.4 ± 0.3 MeV/c^2
R_s	1.02 ± 0.02
N_S	2309 ± 49
N_B	113 ± 16

Table 5.3 *Results of the fits to $B^0 \to \phi K^{*0}$ simulation and data samples.*

Simulation fit parameter	Result
σ_{CB}	13.7 ± 0.1 MeV/c^2
σ_1	22.1 ± 0.4 MeV/c^2
σ_2	54.6 ± 3.7 MeV/c^2
f_{CB}	$(72.9 \pm 1.3)\,\%$
f_1	$(25.3 \pm 1.3)\,\%$
α	3.3 ± 0.1
Data fit parameter	Result
μ	5278.8 ± 0.2 MeV/c^2
R_s	1.06 ± 0.01
N_S	6680 ± 86
N_B	842 ± 72
$N_{B_s^0}$	122 ± 20
$N_{\Lambda_b^0}$	243 ± 62

Ball tail position parameter, α, are found from a fit to the fully selected $B^0 \to \phi K^{*0}$ simulation sample, shown in Figure 5.9 and the top half of Table 5.3. The Crystal Ball tail shape parameter, n, is fixed to 1.

The $B_s^0 \to \phi K^{*0}$ component shares the same shape parameters as the $B^0 \to \phi K^{*0}$ component. The difference between the masses of the two is fixed to the PDG average value of 87.04 ± 0.21 MeV/c^2 [135]. The background component from misidentified $\Lambda_b^0 \to \phi p h^-$ decays is modelled using a histogram PDF generated from a sample of generator-level phase-space simulation events. The mass of the B^0 component, μ, and the yields, $N_{\{S,B\}}$, of the signal and background components are left free in the fit to data.

Figure 5.10 shows the fitted PDF overlaid onto the data. The series of successive

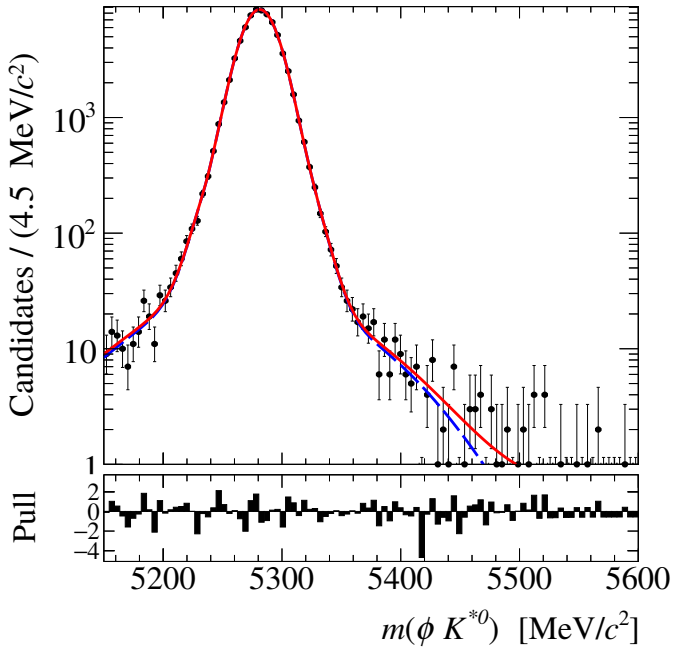


Figure 5.9 *A fit to the ϕK^{*0} invariant mass distribution in the $B^0 \to \phi K^{*0}$ simulation sample. The total fitted PDF is shown as a red solid line. The signal PDF is shown as a blue dashed line. A flat component is added to account for poorly reconstructed and misidentified candidates, but is below the scale of the plot.*

bins with negative pulls around 5500 MeV/c^2 is likely due to modelling of the $\Lambda_b^0 \to \phi p h^-$ background as non-resonant decays. The results of the fit are given in the bottom half of Table 5.3. Binning the data in 100 bins, the probability of the χ^2 of the fit is found to be 0.57. The fitted value of μ deviates from the world average value of $m(B^0) = 5279.58 \pm 0.17$ MeV/c^2 by 3.9 times the statistical uncertainty. The fitted value of R_s is not consistent with unity to within the statistical uncertainty. However, for both of these quantities, the systematic uncertainty arising from the detector calibration is similar in size to the statistical uncertainty. The ratio of yields of the B_s^0 and B^0 components is 0.018 ± 0.003, which agrees with the value found in Ref. [141] at the 2σ level.

5.3.3 S-wave subtraction

The numbers of $B_s^0 \to \phi\phi$ and $B^0 \to \phi K^{*0}$ candidates found in the fits include some fraction of K^+K^- or $K^+\pi^-$ pairs which do not come from ϕ or K^* decays. Instead, they originate from the tails of spin-0 resonances, such the $f_0(980)$ or $K_0^*(1430)^0$, or are produced in a nonresonant configuration. Collectively, the contribution to

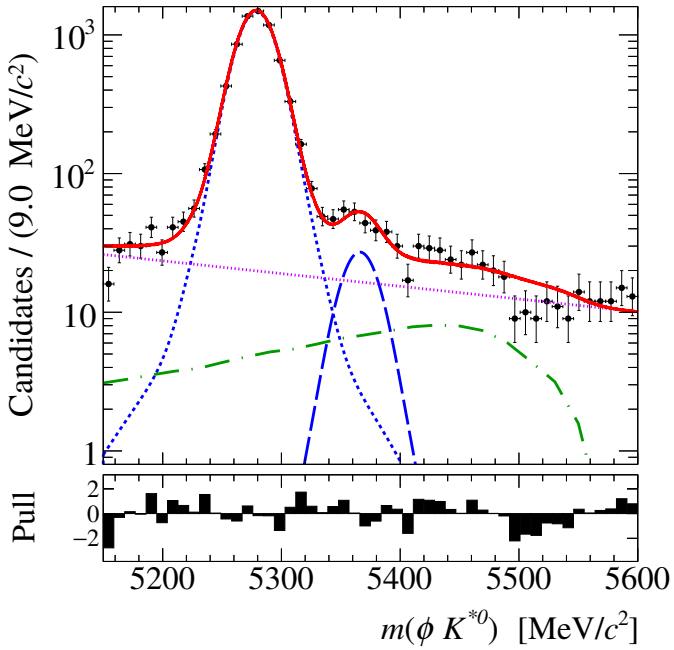


Figure 5.10 *A fit to the ϕK^{*0} invariant mass distribution in the data. The total fitted PDF is shown by the (red) solid line, the $B^0 \to \phi K^{*0}$ component by the (blue) short-dashed line, the $B_s^0 \to \phi K^{*0}$ component by the (blue) long-dashed line, the $\Lambda_b^0 \to \phi p h^-$ contribution by the (green) dashed-dotted line, and the combinatorial background by the (purple) dotted line.*

the yield from these sources is referred to as the S-wave fraction. The contribution from higher-spin resonances is negligible due to the applied K^+K^- and $K^+\pi^-$ mass windows. As mentioned in Section 5.2.3, the mass windows for the ϕ and K^* candidates are chosen so that the S-wave fraction can be taken from LHCb analyses of $B_s^0 \to \phi\phi$ [49] and $B^0 \to \phi K^{*0}$ [47]. The uncertainties on these S-wave fractions are considered as sources of systematic uncertainty on the final branching fraction result.

The $B_s^0 \to \phi\phi$ signal yield is corrected using the S-wave fraction of $(2.12 \pm 1.57)\,\%$ in a window of $|m(K^+K^-) - m(\phi)| < 15$ MeV/c^2, reported in Ref. [49]. The S-wave-subtracted yield is $2212 \pm 47\,(\text{stat}) \pm 50\,(\text{S-wave})$ events.

The $B^0 \to \phi K^{*0}$ signal yield is corrected using the S-wave fraction found in the angular analysis presented in Ref. [47]. The K^+K^- S-wave fraction in the window $|m(K^+K^-) - m(\phi)| < 15$ MeV/c^2 is $(12.2 \pm 1.3)\,\%$, and the $K^+\pi^-$ S-wave fraction in the window $|m(K^+\pi^-) - m(K^*)| < 150$ MeV/c^2 is $(14.3 \pm 1.3)\,\%$. Hence, the S-wave-subtracted yield is $5026 \pm 65\,(\text{stat}) \pm 107\,(\text{S-wave})$ events.

This approach to S-wave subtraction assumes that the angular acceptance for the S-wave and P-wave contributions is the same. The validity of this assumption is discussed in Section 5.5.6.

5.4 Detection efficiency

The branching fraction calculation, Equation 5.1, requires the ratio of the overall detection efficiencies for the signal and normalisation decay modes. In this analysis, the total detection efficiency is the product of the generator, selection and particle identification (PID) efficiencies, $\varepsilon^{\text{tot}} = \varepsilon^{\text{gen}} \cdot \varepsilon^{\text{sel}} \cdot \varepsilon^{\text{PID}}$, which are calculated independently with different methods and datasets. The generator and selection efficiencies are calculated using the simulation. Since the simulation does not accurately reproduce the PID variables, the efficiency of the cuts on these variables is calculated using a data-driven technique.

5.4.1 Generator efficiency

During the event generator phase of the simulation, cuts are applied to the events in order to reduce the computing load required to fully simulate and digitise the sample. Both the signal and normalisation simulation samples have cuts applied which ensure the final state particles are in the geometric acceptance of the detector. The $B_s^0 \to \phi\phi$ sample has an additional cut that requires the p_{T} of the final state kaons to be above 400 MeV/c, which leads to a significantly lower generator-level efficiency than for the $B^0 \to \phi K^{*0}$ sample. The generator-level efficiencies are quoted relative to the number of events that would have been generated in the full η range.

Table 5.4 *Generator-level efficiencies (in %) for the simulation samples used in the $B_s^0 \to \phi\phi$ branching fraction measurement and the $B^0 \to \phi\phi$ branching fraction limit.*

Decay Mode	Polarity	$B\ \varepsilon^{\text{gen}}$ [%]	$\overline{B}\ \varepsilon^{\text{gen}}$ [%]	Total ε^{gen}[%]
$B^0 \to \phi K^{*0}$	Up	18.69 ± 0.07	18.67 ± 0.07	18.69 ± 0.04
	Down	18.77 ± 0.07	18.63 ± 0.07	
$B_s^0 \to \phi\phi$	Up	17.10 ± 0.09	17.05 ± 0.09	17.09 ± 0.04
	Down	17.15 ± 0.09	17.06 ± 0.09	
$B^0 \to \phi\phi$	Up	19.84 ± 0.08	19.83 ± 0.08	19.81 ± 0.04
	Down	19.86 ± 0.08	19.70 ± 0.08	

Table 5.4 shows the generator-level cut efficiency of the signal and normalisation samples, calculated for each magnet polarity and B meson flavour. The total efficiency of each decay mode is taken as the average of these numbers, since they are consistent with each other. For comparison, as mentioned in Section 3.3.2, approximately 25 % of the $b\bar{b}$ pairs produced in LHC collisions are in the acceptance of the LHCb detector.

The ratio of generator efficiencies used in the $B_s^0 \to \phi\phi$ branching fraction calculation is:

$$\frac{\varepsilon^{\text{gen}}_{\phi K^{*0}}}{\varepsilon^{\text{gen}}_{\phi\phi}} = 1.094 \pm 0.004.$$

This is larger than unity due to the $p_{\text{T}} > 400$ MeV/c cut on the kaons in the $B_s^0 \to \phi\phi$ sample. The effect of this cut is compensated for in the ratio of selection efficiencies, calculated in Section 5.4.2.

5.4.2 Selection efficiency

The selection efficiency is determined by passing the simulation samples through the event selection described in Section 5.2 without the cuts on the PID variables. Fits are performed to the invariant mass distributions of the samples in order to determine the yields of selected events, which are divided by the numbers of events in the generator-level samples to give the selection efficiencies. The numbers are shown in Table 5.5.

Table 5.5 *Selection efficiencies for the MC samples used in the $B_s^0 \to \phi\phi$ branching fraction measurement.*

Decay Mode	Generated	Selected	ε^{sel} [%]
$B^0 \to \phi K^{*0}$	4027985	74758 ± 273	1.86 ± 0.01
$B_s^0 \to \phi\phi$	2058745	49211 ± 221	2.39 ± 0.02

The ratio of selection efficiencies used in the $B_s^0 \to \phi\phi$ branching fraction calculation is

$$\frac{\varepsilon^{\text{sel}}_{\phi K^{*0}}}{\varepsilon^{\text{sel}}_{\phi\phi}} = 0.776 \pm 0.006.$$

This number deviates from unity for two reasons. Firstly, as mentioned in Section 5.4.1 there is a difference in the cuts applied at the generator level: namely the p_{T} cut on the kaons in the $B_s^0 \to \phi\phi$ sample means that a larger proportion of events subsequently passes the later cut on the same variables. Secondly, the pion

from the K^* decay in the $B^0 \to \phi K^{*0}$ sample has a much softer p_{T} spectrum than the kaons in either sample, which further reduces the efficiency of the p_{T} cut.

5.4.3 Particle identification efficiency

The signal and normalisation channels have similar kinematics and have been selected using similar selection requirements, with the difference lying in the PID. Since the simulation does not accurately describe the hadron PID performance of the detector, it is necessary to use a data-driven method to calculate the efficiencies of the PID requirements. This is standard practice in LHCb analyses, and uses a tool called 'PIDCalib' [142].

A calibration sample of kaons and pions from $D^{*+} \to (D^0 \to K^{\pm}\pi^{\mp})\,\pi^+$ decays is used to calculate the efficiency as a function of p, p_{T}, and η of the track, as well as the multiplicity of tracks in the event and the polarity of the LHCb magnet. These values are translated to a per-event efficiency for each track in the $B^0_s \to \phi\phi$ and $B^0 \to \phi K^{*0}$ simulation samples after applying all the selection requirements described in Section 5.2 apart from the PID cuts.

This method of calculating the efficiencies of the PID cuts relies on three assumptions. The first is that p, p_{T}, η, multiplicity and polarity are the only dependent variables. The second is that the kinematic distributions agree between the simulation and data. The third is that the binning scheme in the dependent variables is sufficiently fine to determine the PID efficiency to the required accuracy. A systematic uncertainty, arising from the validity of these assumptions, is quantified in Section 5.5.2.

For the signal and normalisation modes, all kaons are required to have a value of $P_K \times (1 - P_\pi)$ greater than 0.025. Figure 5.11 compares the efficiencies for a range of cuts on $P_K \times (1 - P_\pi)$ between the particles in the calibration sample and simulated kaons from $B^0_s \to \phi\phi$ decays. The pion in the normalisation mode is required to have $P_\pi > 0.2$ and $P_K < 0.2$. Figures 5.12 and 5.13 compare the efficiencies for a range of cuts on P_π and P_K between calibration tracks and simulated pions from $B^0 \to \phi K^{*0}$ decays.

The simulated PID efficiency is found to be $(94.5 \pm 0.2)\,\%$, whereas the data-driven method gives a smaller efficiency of $(90.7 \pm 0.1)\,\%$. The PID cut efficiency for the $B^0 \to \phi K^{*0}$ sample, using the data-driven method, is $(84.9 \pm 0.1)\,\%$.

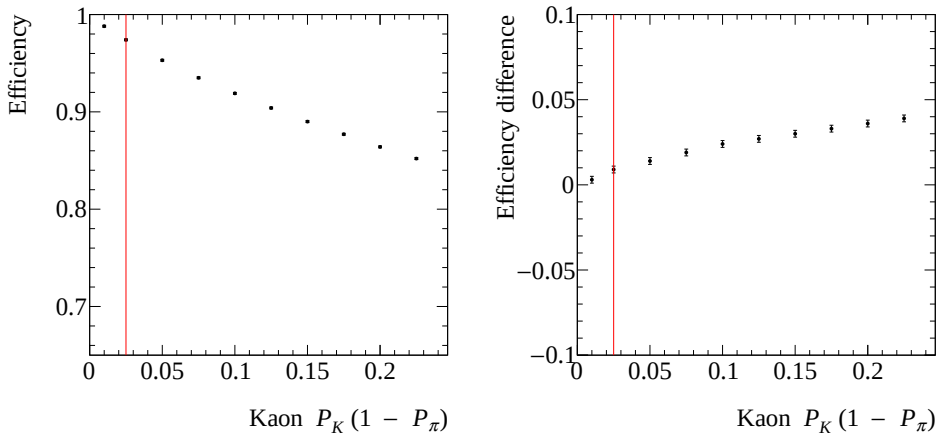


Figure 5.11 *Comparison of PID efficiencies for a range of cuts on $P_K \times (1 - P_\pi)$ between data calibration sample and simulated events for kaons from the $B_s^0 \to \phi\phi$ decay, showing (left) data-calibrated PID efficiency and (right) difference between MC and data-calibrated efficiencies. Red vertical lines are placed at the cut values used in the selection.*

The ratio of PID efficiencies is found to be

$$\frac{\varepsilon^{\text{PID}}_{B^0 \to \phi K^{*0}}}{\varepsilon^{\text{PID}}_{B_s^0 \to \phi\phi}} = 0.936 \pm 0.001,$$

using the results from the data-driven method, where the quoted uncertainty is statistical and arises due to the limited size of the calibration samples.

5.5 Systematic uncertainties

The sources of systematic uncertainties on the $B_s^0 \to \phi\phi$ branching fraction measurement are discussed in this section.

5.5.1 Choice of fit model

The uncertainty on the $B_s^0 \to \phi\phi$ branching fraction due to the choice of fit model is estimated by fitting the invariant mass distributions with different functions for the signal and background components. The largest deviation in the ratio of yields is taken as an estimate of the systematic uncertainty. This is done separately for the signal, combinatorial background and misidentified background components. The assigned uncertainty is the sum in quadrature of the three.

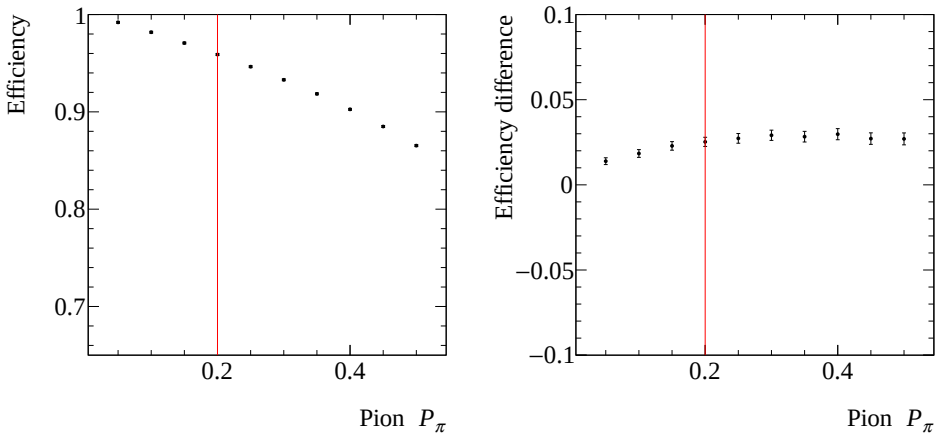


Figure 5.12 *Comparison of PID efficiencies for a range of cuts on P_π between the data calibration sample and simulated events for pions from the $B^0 \to \phi K^{*0}$ decay, showing (left) data-calibrated PID efficiency and (right) difference between MC and data-calibrated efficiencies. Red vertical lines are placed at the cut values used in the selection.*

The functions considered for the signal components are the sum of two or three Gaussian functions, a Student's T function[3], and the sum of a Crystal Ball function with one or two Gaussian functions. The largest deviation is found using the Student's T function for $B_s^0 \to \phi\phi$ and the sum of a Crystal Ball with two Gaussian functions for $B^0 \to \phi K^{*0}$, which gives a relative uncertainty of 0.59 %.

In the case of the background model, the largest deviation is found using the linear function for both the signal and normalisation modes, which gives a relative uncertainty of 0.16 %. The effect of removing the $\Lambda_b^0 \to \phi p h^-$ peaking background component from the ϕK^{*0} invariant mass model is also tested and found to give a relative uncertainty of 0.03 %. The total relative uncertainty from the choice of fit model on the ratio of yields is 0.61 %.

5.5.2 Particle identification efficiency

An uncertainty on the PID efficiencies arises from the calibration method. This is quantified by varying the binning scheme for the calibration sample. The default binning scheme has $21 \times 4 \times 4$ bins, corresponding to the three variables: p_{T}, η and track multiplicity. The number of bins is increased to $31 \times 6 \times 6$, while retaining the same range for each variable. The efficiency tables for kaons and pions are regenerated with this finer binning, and the efficiency of the PID cuts is

[3]See Appendix B.3 for the explicit form of this function.

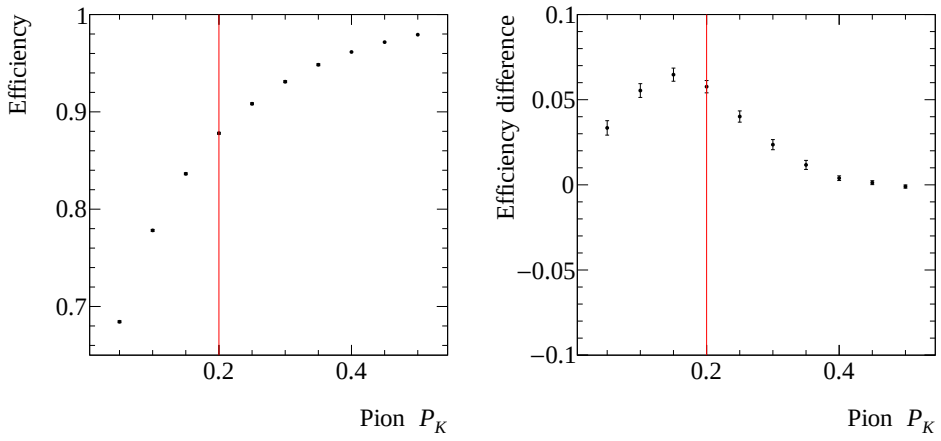


Figure 5.13 *Comparison of PID efficiencies for a range of cuts on P_K between the data calibration sample and simulated events for pions from the $B^0 \to \phi K^{*0}$ decay, showing (left) data-calibrated PID efficiency and (right) difference between MC and data-calibrated efficiencies. Red vertical lines are placed at the cut values used in the selection.*

recalculated using the method described in Section 5.4.3. This causes a change in the PID efficiency ratio of 0.3%, compared to the value calculated with the default scheme. This change is taken as the systematic uncertainty on the PID efficiency.

5.5.3 Hardware-level hadron trigger

The L0 hadron trigger is not reproduced perfectly in the simulation. A data-driven recalibration of this trigger efficiency is performed using the large sample of $D^0 \to K^{\pm}\pi^{\mp}$ decays collected during 2011 and 2012. The calibration is then applied to the $B_s^0 \to \phi\phi$ and $B^0 \to \phi K^{*0}$ simulation samples in place of the default trigger emulation. This sample consists of binned distributions of the L0 hadron trigger efficiency for charged pions and kaons as functions of charge, transverse momentum and magnet polarity.

Table 5.6 *Comparison between trigger efficiencies using the simulation and the data-driven resimulation methods for the L0 hadron trigger.*

Total Trigger Efficiency	$B^0 \to \phi K^{*0}$ [%]	$B_s^0 \to \phi\phi$ [%]	Ratio
Original simulation	34.4 ± 0.1	27.8 ± 0.1	1.237 ± 0.006
Data-driven method	29.2 ± 0.1	23.8 ± 0.1	1.227 ± 0.007

The resimulation proceeds as follows. On an event-by-event basis, four random numbers between 0 and 1 are generated and assigned to each particle and compared with the value of the trigger efficiency as found from the calibration sample.

Table 5.7 *Fraction of kaons and pions from $B^0 \to J/\psi K^*$ decays which have a hadronic interaction end-vertex before the end of the tracking system and the inferred interaction lengths.*

Particle	Fraction [%]	λ_I [m]
π^+	14.7 ± 0.6	59.2 ± 2.5
π^-	15.4 ± 0.6	56.3 ± 2.3
K^+	11.5 ± 0.5	77.0 ± 3.4
K^-	13.7 ± 0.5	63.9 ± 2.4

If the random number is less than the efficiency for any of the particles, the event is deemed to have passed the trigger. The systematic uncertainty is the relative change in the ratio of total trigger efficiencies when using the data-driven resimulation of the L0 hadron trigger, compared to the 'default' method of using the efficiencies from the original simulation. The values are given in Table 5.6, and the resulting systematic uncertainty is 1.1 %.

5.5.4 Hadronic interaction with the detector

Hadrons passing through the detector can interact strongly with the detector material. Those which are absorbed before the end of the tracking system will not pass the 'long track' requirement of the stripping line. The loss of particles as a function of distance, z, in the direction of the beam is approximated as

$$\exp(-z/\lambda_I),$$

where λ_I is the interaction length, *i.e.* the mean distance that the particle travels through a material before interacting. The pion and kaon hadronic interaction lengths differ by up to $\sim 30\,\%$ in the momentum range between 5 and 100 GeV/c [143]. The hadronic interaction lengths are not perfectly reproduced in the simulation, which leads to a systematic uncertainty on the ratio of selection efficiencies.

A sample of simulated $B^0 \to J/\psi K^*$ events is used to categorise the interaction type of kaons and pions that have end-vertices before the end of the tracking system ($z = 9.41$ m). The fractions of kaons and pions that interact hadronically with the detector in this region are shown in Table 5.7. This information is used to calculate λ_I for each hadron species.

The simulation is considered to reproduce the hadronic interaction length to an

Table 5.8 *Uncertainties on the number of kaons and pions passing the 'long track' cut due to the miss-modelling of hadronic interactions.*

Particle	Uncertainty [%]
π^+	1.44
π^-	1.51
K^+	1.10
K^-	1.33

accuracy of 10 % [144], hence the relative uncertainty on the number of remaining particles is

$$\sigma = 1 - \frac{\exp(-z/\lambda_I)}{\exp(-z/1.1\lambda_I)}.$$

Table 5.8 lists the uncertainties obtained using this formula. Since the final states differ by the exchange of a kaon with a pion, the effect on the ratio of selection efficiencies largely cancels. The residual uncertainty is calculated as the difference between the charge-averaged pion and kaon uncertainties. A relative systematic uncertainty of 0.26 % is assigned to the result.

5.5.5 Tracking efficiency

There is an approximate 2 % discrepancy in the efficiency of track reconstruction between data and simulation [145]. Since the final states of the signal and normalisation modes each have four tracks, this uncertainty largely cancels. However, the presence of the low-momentum pion in the decay $B^0 \to \phi K^{*0}$ leads to a residual uncertainty.

A look-up table of tracking efficiency ratios, and their uncertainties, binned in η and p_{T}, is provided by the LHCb tracking group. The table is illustrated as a two-dimensional histogram in Figure 5.14. A number of pseudo-experiments are performed, in which each efficiency ratio is varied randomly within its uncertainty. The events in the $B^0 \to \phi K^{*0}$ and $B_s^0 \to \phi\phi$ simulation samples are weighted by the product of the tracking efficiency ratios for each track. The ratio of selection efficiencies is then calculated from the sum of the weights of each sample. A distribution is obtained by performing 500 such pseudo-experiments. A bias of 0.5 % is found and assigned as a systematic uncertainty.

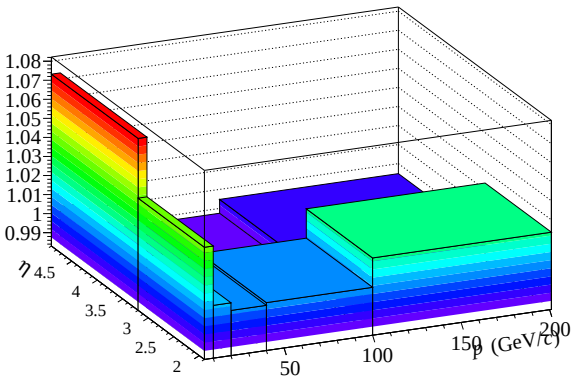


Figure 5.14 *The look-up table of tracking efficiency ratios, binned in η and p_{T}.*

5.5.6 S-wave subtraction and angular acceptance

The subtraction of the K^+K^- and $K^\pm\pi^\mp$ S-wave contributions from the $B^0 \to \phi K^{*0}$ yield is done using the results from a prior LHCb angular analysis of the decay [47]. It is assumed that the ratio of detection efficiencies for S-wave events and P-wave events is the same in this analysis. If this assumption is invalid, the S-wave fraction will differ from the values found in the angular analysis.

This effect is quantified using the acceptance calculated in the angular analysis. Three event categories are considered: those that are true $B^0 \to \phi K^{*0}$ decays; events where the K^+K^- pair is in an S-wave configuration, but the $K^\pm\pi^\mp$ pair is in a P-wave configuration and events with P-wave K^+K^- and S-wave $K^\pm\pi^\mp$. The contribution from events where both the track pairs are in an S-wave configuration is considered to be negligible. These categories differ from one another in their distributions of helicity angles and two-body invariant masses. In the $B^0 \to \phi K^{*0}$ angular analysis, the acceptance is calculated as a function of the three helicity angles and $K^\pm\pi^\mp$ invariant mass, since there was found to be no dependence on the K^+K^- invariant mass.

Toy datasets are generated for each of the event categories, using the PDF from the angular analysis with parameters fixed to the fit results. Weights are assigned to each toy event using the four-dimensional angular acceptance. The ratio of the average efficiencies for each category is taken as the ratio of the sum of these per-event weights. The weights for the two S-wave categories are each summed in the ratio given by the relevant S-wave fractions found in the angular analysis: 12.2 % for K^+K^- and 14.3 % for $K^+\pi^-$. This is then divided by the sum of the

P-wave weights, giving the efficiency of S-wave events relative to that of P-wave events.

The difference in efficiency between the S-wave and P-wave events is multiplied by the total S-wave fraction to give a relative systematic uncertainty of 1.1 % on the branching fraction result. A similar angular acceptance uncertainty is not necessary for the $B_s^0 \to \phi\phi$ mode, since the S-wave fraction is measured with an angular analysis [49] to be at the level of 1 %, and hence any effect is negligible.

5.5.7 Other uncertainties and summary

As mentioned in Section 5.2.4, the efficiency of the BDT cut on the $B^0 \to \phi K^{*0}$ mode disagrees between data and simulation. This is accounted-for by assigning a systematic uncertainty of 0.7 % on the ratio of selection efficiencies.

In the calculation of the $B_s^0 \to \phi\phi$ branching fraction, the branching fractions of $B^0 \to \phi K^{*0}$ [146, 147], $\phi \to K^+K^-$ and $K^{*0} \to K^+\pi^-$ [135], as well as the ratio of fragmentation fractions [148–150], are used as external inputs. The uncertainties on these quantities are treated as sources of systematic uncertainty on the $B_s^0 \to \phi\phi$ branching fraction result. A summary of the systematic uncertainties presented in this section is shown in Table 5.9. The total relative systematic uncertainty on the branching fraction result is taken as the sum in quadrature of the individual uncertainties. The dominant uncertainties come from the S-wave fractions, the ratio of fragmentation fractions and the normalisation channel branching fraction. The latter two uncertainties are quoted separately in the result shown in Section 5.6, in order to assist calculating the $B_s^0 \to \phi\phi$ branching fraction again, should more precise measurements become available.

5.6 The $B_s^0 \to \phi\phi$ branching fraction result

As mentioned in Section 5.1, the ratio of the $B_s^0 \to \phi\phi$ branching fraction relative to the branching fraction of the normalisation channel is calculated using Equation 5.1.

$$\frac{\mathcal{B}(B_s^0 \to \phi\phi)}{\mathcal{B}(B^0 \to \phi K^{*0})} = \frac{N_{B_s^0 \to \phi\phi}}{N_{B^0 \to \phi K^{*0}}} \frac{\varepsilon_{B^0 \to \phi K^{*0}}}{\varepsilon_{B_s^0 \to \phi\phi}} \frac{\mathcal{B}(K^{*0} \to K^+\pi^-)}{\mathcal{B}(\phi \to K^+K^-)} \cdot \frac{1}{f_s/f_d}, \tag{5.1}$$

where the symbols $\mathcal{B}$ represent branching fractions, N are yields, ε are efficiencies, and f_s/f_d is the ratio of fragmentation fractions.

Table 5.9 *Summary of systematic uncertainties for the $B_s^0 \to \phi\phi$ branching fraction measurement. The S-wave uncertainty is discussed in Section 5.3.3. The uncertainties due to the detection efficiency are discussed in Section 5.4. Uncertainties on external inputs are discussed in Section 5.6.*

Systematic	Source	Relative error [%]
Fit model	Vary fit model in data	0.6
$B^0 \to \phi K^{*0}$ BDT efficiency	Yields from data fit	0.7
Selection efficiency	Simulation statistics	0.7
Generator-level efficiency	Simulation statistics	0.3
Hadronic interaction	Tracking prescription	0.3
Tracking efficiency	Tracking prescription	0.5
L0 hadron trigger	Calibration samples	1.1
PID efficiency	Data-driven calculation	0.3
S-wave fractions	LHCb angular analysis	3.1
Angular acceptance	$B^0 \to \phi K^{*0}$ acceptance	1.1
$\mathcal{B}(\phi \to K^+K^-)$	PDG	1.0
Sum of above in quadrature		3.9
f_s/f_d	LHCb measurement	5.8
$\mathcal{B}(B^0 \to \phi K^{*0})$	PDG	6.4

The value of the $B^0 \to \phi K^{*0}$ branching fraction used in this analysis is the average of the 2008 BaBar result [146] and the 2013 Belle result [147], calculated using the Particle Data Group's prescription for averaging results [135]. In the calculation, it is assumed that the systematic uncertainties due to the fragmentation fractions and S-wave subtraction are fully correlated. The resulting value is $\mathcal{B}(B^0 \to \phi K^{*0}) = (1.00 \pm 0.04 \pm 0.05) \times 10^{-5}$. It should be noted that both analyses consistently account for the K^+K^- and $K^+\pi^-$ S-wave contributions. Additionally, the measurements were performed with the assumption that $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0) = \mathcal{B}(\Upsilon(4S) \to B^+B^-)$. However, the ratio of the two is known to deviate from unity by a few percent [151], leading to a source of systematic uncertainty which is unaccounted for [152].

The $\phi \to K^+K^-$ branching fraction is taken from the Particle Data Group's Review of Particle Physics [135], $\mathcal{B}(\phi \to K^+K^-) = 0.489 \pm 0.005$. The K^{*0} decays almost exclusively to $K\pi$ [135], so the $K^{*0} \to K^+\pi^-$ branching fraction is taken to be exactly 2/3 based on an isospin assumption. The ratio of fragmentation fractions is taken from LHCb measurements to be $f_s/f_d = 0.259 \pm 0.015$ [148–150].

The relative branching fraction is

$$\frac{\mathcal{B}(B_s^0 \to \phi\phi)}{\mathcal{B}(B^0 \to \phi K^{*0})} = 1.84 \pm 0.05 \ \text{(stat)} \pm 0.07 \ \text{(syst)} \pm 0.11\,(f_s/f_d),$$

where the uncertainty from the fragmentation fraction is quoted separately. The absolute branching fraction is

$$\mathcal{B}(B_s^0 \to \phi\phi) = (1.84 \pm 0.05 \ \text{(stat)} \pm 0.07 \ \text{(syst)} \pm 0.11\,(f_s/f_d) \pm 0.12\,(\text{norm})) \times 10^{-5},$$

where the fourth quoted uncertainty is due to the uncertainty on the branching fraction of the normalisation mode.

This result is consistent with the previous result by the CDF collaboration [48],

$$\mathcal{B}(B_s^0 \to \phi\phi) = (1.91 \pm 0.26 \ \text{(stat)} \pm 0.16 \ \text{(syst)}) \times 10^{-5},$$

and improves on the statistical uncertainty by a factor of five. The total relative uncertainty on the result is about 9 %, compared to 16 % for the CDF result. The weighted average of the two results is $(1.86 \pm 0.16) \times 10^{-5}$. The value is consistent with theory predictions, which have central values in the range (1.3 to 2.6) $\times$ 10^{-5} [59–62], although with large uncertainties.

5.7 Search for the decay $B^0 \to \phi\phi$

The suppressed decay mode $B^0 \to \phi\phi$ is searched for by fitting the $\phi\phi$ invariant mass distribution, with a selection that differs only by the choice of cut on the output of the BDT classifier: BDT> 0.14 rather than BDT> -0.025, as detailed in Section 5.2.4. From the result of this fit, shown in Figure 5.15 and Table 5.10, it is clear that there is no significant $B^0 \to \phi\phi$ signal. Hence, a limit is set on the value of the $B^0 \to \phi\phi$ branching fraction.

5.7.1 Efficiency

The trigger and reconstruction efficiencies cancel between the $B^0 \to \phi\phi$ and $B_s^0 \to \phi\phi$ modes. However, the invariant mass cuts applied by the stripping line are under the $\pi^+\pi^-\pi^+\pi^-$ mass hypothesis and are tight enough to cause a difference in the selection efficiency of the two modes. In order to obtain

the selection efficiency relative to $B_s^0 \to \phi\phi$ as a function of $\phi\phi$ invariant mass, a set of six $B_s^0 \to \phi\phi$ generator-level event samples are generated with varied $K^+K^-K^+K^-$ masses in a range 5159 MeV/c^2 to 5366 MeV/c^2. For reference, the B^0 mass is 5279.58 ± 0.17 MeV/c^2, and the B_s^0 mass is 5366.77 ± 0.24 [135]. For each sample, the ratio of the efficiency of the preselection cuts relative to that of the $m(B_s^0) = 5366$ MeV/c^2 sample is calculated. A fit is made to this distribution, and the ratio of selection efficiencies between $B^0 \to \phi\phi$ and $B_s^0 \to \phi\phi$ is taken to be the value at $m(B_s^0) = 5279$ MeV/c^2, which is 0.86.

5.7.2 Mass fit

Two fits are performed to the $\phi\phi$ invariant mass distribution using the re-optimised multivariate selection. The first is the background-only model, f_b, which is identical to the one described in Section 5.3.1. The other is the signal plus background model, f_{s+b}, which differs by the addition of a component to account for $B^0 \to \phi\phi$ decays. The same strategy is used with regards to fixing parameters to values found from fitting the simulation, as in Section 5.3. For the fit with the $B^0 \to \phi\phi$ component, the branching fraction of the decay is used as a free parameter, labelled BR in Table 5.10, to control the size of the B^0 peak. The results of the fit are shown in Figure 5.15 and Table 5.10.

Table 5.10 *Results of the fits to $B_s^0 \to \phi\phi$ simulation and data using a model containing a $B^0 \to \phi\phi$ component.*

Simulation fit parameter	Result
σ_1	13.8 ± 0.1 MeV/c^2
σ_2	28.1 ± 0.4 MeV/c^2
σ_3	41.0 ± 1.0 MeV/c^2
f_1	$(84.7 \pm 1.2)\,\%$
f_2	$(13.6 \pm 1.2)\,\%$
Data fit parameter	Result
μ	5366.5 ± 0.3 MeV/c^2
R_s	1.02 ± 0.02
N_S	1995 ± 49
BR	$(1.1 \pm 1.7) \times 10^{-8}$
N_B	76 ± 12

The fitted value of the $B^0 \to \phi\phi$ branching fraction is found to be $\mathcal{B}(B^0 \to \phi\phi) = (1.1 \pm 1.7) \times 10^{-8}$. The significance of this peak is calculated from a scan of the log-likelihood as a function of the $B^0 \to \phi\phi$ branching fraction. Figure 5.16 shows the difference in negative log-likelihood, $-\Delta \ln(\mathcal{L})$, from the minimum value,

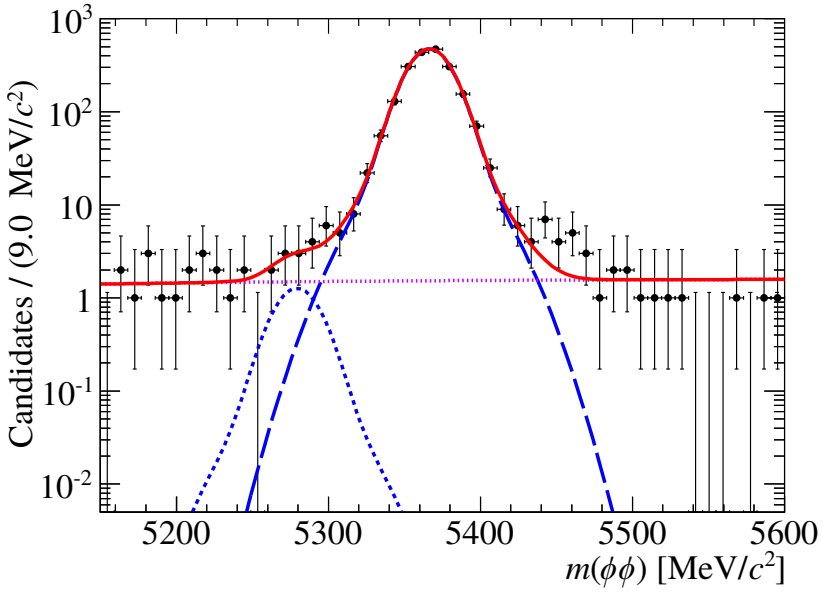


Figure 5.15 *A fit to the $\phi\phi$ invariant mass with the tight BDT selection applied. The total PDF is shown as a red solid line, $B_s^0 \to \phi\phi$ as a blue long-dashed line, $B^0 \to \phi\phi$ as a blue short-dashed line, and the combinatorial background as a purple dotted line.*

as a function of $\mathcal{B}(B^0 \to \phi\phi)$. The value of $-\Delta\ln(\mathcal{L})$ in the null hypothesis is found to be 0.87. Using Wilks' theorem [153], this translates to a significance of $\sqrt{-2\Delta\ln(\mathcal{L})} = 1.3\sigma$.

5.7.3 Limit on the $B^0 \to \phi\phi$ branching fraction

The limit on the $B^0 \to \phi\phi$ branching fraction is set using the $\mathrm{CL_s}$ method [154] as implemented in the RooStats package [155]. The $\mathrm{CL_s}$ variable is calculated as the ratio of the p-value of the signal plus background hypothesis to the p-value of the background only hypothesis. A likelihood ratio test statistic, R_L, is defined as

$$R_L = \frac{\mathcal{L}(f_{s+b}(x))}{\mathcal{L}(f_b(x))},$$

where $\mathcal{L}(f_{s+b}(x))$ is the likelihood evaluated for the signal plus background model, $\mathcal{L}(f_b(x))$ is the likelihood evaluated for the background-only model, and x is the set of parameters in the model.

At each point in a scan through the values of the $B^0 \to \phi\phi$ branching fraction, R_L is calculated from fitting the data. These fits also determines the values of

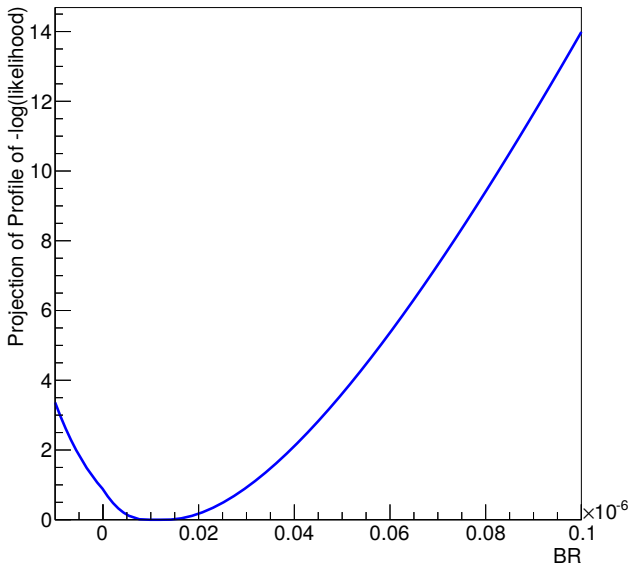


Figure 5.16 *A profile scan of minimised* $-\Delta\ln(\mathcal{L})$ *over* $\mathcal{B}(B^0 \to \phi\phi)$.

the nuisance parameters (μ, R_s, N_S and N_B) to be used when generating toys. The models f_{s+b} and f_b are each used to generate 50,000 toys. The values of $\mathrm{CL_{s+b}}$ and $\mathrm{CL_b}$ are calculated as the number of corresponding toys with a value of $-2\ln(R_L)$ above that found in data. The value of $\mathrm{CL_s}$ is then calculated as

$$\mathrm{CL_s} = \frac{\mathrm{CL_{s+b}}}{\mathrm{CL_b}}.$$

Figure 5.17 shows the results of the $\mathrm{CL_s}$ scan. The upper limit at 90 % (95 %) confidence level is taken as where the observed $\mathrm{CL_s}$ distribution falls below 0.1 (0.05). The upper limits on the $B^0 \to \phi\phi$ branching fraction are measured to be

$$\begin{aligned} \mathcal{B}(B^0 \to \phi\phi) &< 2.8 \times 10^{-8} \quad (90\,\%\ \mathrm{CL}), \\ \mathcal{B}(B^0 \to \phi\phi) &< 3.4 \times 10^{-8} \quad (95\,\%\ \mathrm{CL}). \end{aligned}$$

The limit at 90 % confidence level is a factor of seven improvement on the previous result from BaBar: $\mathcal{B}(B^0 \to \phi\phi) < 2.0 \times 10^{-7}$ [136]. It is close to the higher of the theoretical predictions, which lie in the range (0.1 to 3.0) $\times 10^{-8}$ [59, 68, 69].

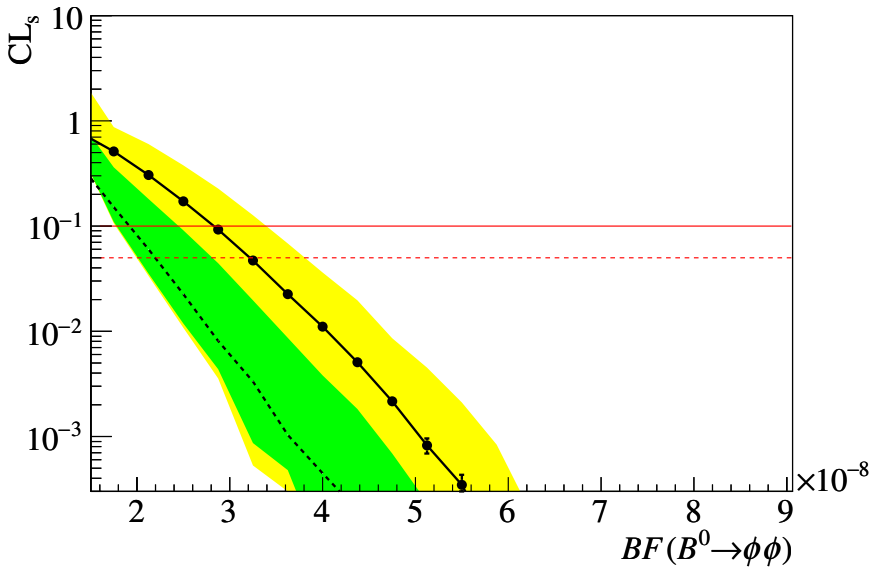


Figure 5.17 *Results of the* $\mathrm{CL_s}$ *scan with the observed* $\mathrm{CL_s}$ *values plotted as black points. The black dashed line is the expected distribution. The green and yellow bands mark the* 1σ *and* 2σ *confidence regions. The upper limit at* 90 % (95 %) *confidence level is where the observed* $\mathrm{CL_s}$ *line intercepts the red solid (dashed) horizontal line.*

Chapter 6

Amplitude analysis of the decay $B_s^0 \to \phi K^+ K^-$ in the low $m(K^+K^-)$ region

6.1 Introduction

The decay $B_s^0 \to \phi K^+ K^-$ with $m(K^+K^-) < m(D^0)$ is dominated by a $b \to s\bar{s}s$ transition. As discussed in Section 2.4.2, this process is sensitive to physics beyond the Standard Model. Measurements of this process could help resolve the 'polarisation puzzle' described in Section 2.4.4.

This chapter describes an amplitude analysis of $B_s^0 \to \phi K^+ K^-$ decays with $m(K^+K^-) < 1.8$ GeV/c^2. The largest expected resonant contributions are the decays $B_s^0 \to \phi\phi$ and $B_s^0 \to \phi f_2'(1525)$. The latter has not previously been observed. Predictions exist for its branching fraction and longitudinal polarisation fraction (see Section 2.4.5).

The event selection is described in Section 6.2. The PDF used to perform the amplitude fit contains components for signal and background. The sizes of these components are fixed to the values found from a study of the peaking backgrounds (Section 6.3) and a fit to the $K^+K^-K^+K^-$ invariant mass (Section 6.4). The PDF also accounts for the detector efficiency (Section 6.6) and resolution (Section 6.7. The amplitude formalism used in the signal model is described in Section 6.5. Justifications of the choice of K^+K^- resonances to consider in the signal model

are given in Section 6.8. Further details of the signal model, such as numerical values of parameters, are given in Section 6.9. The background model is described in Section 6.10.

The PDF is first validated using toy studies and fits to simulation, as described in Section 6.11. The best resonance model is determined from a series of fits to data, as described in Section 6.12. The systematic uncertainties are given in Section 6.13 and the key results are given in Section 6.14.

6.2 Event selection

This section describes the selection used in the analysis. The goal is to remove combinatorial background and candidates with misidentified hadrons without sacrificing signal efficiency or introducing bias in the observables used in the amplitude fit. A more general overview of the event selection strategy is given in Chapter 4.

6.2.1 Dataset

This analysis uses the full Run 1 dataset, described in Section 4.2. Simulated $B_s^0 \to \phi K^+ K^-$ events, generated using phase space, are used to train multivariate algorithms (Section 6.2.5) and model the acceptance function (Section 6.6). Simulated $B_s^0 \to \phi\phi$ events are used to verify that the amplitude fit reproduces the input values for the helicity amplitudes.

Due to the wider invariant mass window on one of the K^+K^- pairs, compared to the $B_s^0 \to \phi\phi$ analysis in Chapter 5, it is necessary to consider more potential peaking backgrounds in the data sample. For this reason, samples of simulated $B^0 \to \phi K^{*0}$ (with $K^{*0} \to K^+\pi^-$) and $\Lambda_b^0 \to \phi p K^-$ events are produced, with the latter generated according to phase space.

6.2.2 Trigger lines

At the Level 0 trigger stage, candidates are required to pass either the L0 Hadron TOS or L0 Global TIS lines (see Section 3.6). At the HLT1 stage, candidates are required to pass the 'single-track all L0' TOS line, which searches for a

single detached high-momentum track without any confirmation of a L0 trigger decision [114]. Events are subsequently required to pass either of the topological 3- or 4-body B decay TOS lines [126].

Table 6.1 *Efficiencies of the trigger lines on signal simulation samples after stripping. The column titled '$m(K^+K^-)$ cut' refers to the $B_s^0 \to \phi K^+K^-$ phase space simulation sample with a cut of $m(K^+K^-) < 1800$ MeV/c^2. The number quoted at each row is the fraction kept after applying the previous step.*

Trigger line	Fraction of events kept [%]		
	$B_s^0 \to \phi K^+K^-$	$m(K^+K^-)$ cut	$B_s^0 \to \phi\phi$
L0 Hadron TOS	24.41 ± 0.05	20.51 ± 0.08	17.09 ± 0.07
L0 Global TIS	28.60 ± 0.06	26.98 ± 0.09	28.67 ± 0.09
Total L0	44.18 ± 0.06	40.30 ± 0.10	39.19 ± 0.09
HLT1 Track All L0 TOS	80.64 ± 0.08	79.52 ± 0.13	75.10 ± 0.13
HLT2 Topo 3-Body TOS	84.69 ± 0.08	84.38 ± 0.13	80.98 ± 0.14
HLT2 Topo 4-Body TOS	70.13 ± 0.10	70.00 ± 0.16	69.66 ± 0.16
Total HLT2	86.54 ± 0.07	86.63 ± 0.12	84.53 ± 0.13
Total	30.83 ± 0.06	27.76 ± 0.09	24.88 ± 0.08

The efficiency of each trigger line, and their combinations, is calculated using simulation samples after stripping has been applied. Table 6.1 summarises the efficiency of each line for the $B_s^0 \to \phi K^+K^-$ and $B_s^0 \to \phi\phi$ samples. In order to verify that the chosen trigger lines introduce a smooth acceptance function, which can be accurately modelled by the method described in Section 6.6, the trigger efficiency is calculated in 20 bins of $m(K^+K^-)$ between 990 MeV/c^2 and 1800 MeV/c^2. These distributions are shown in Figure 6.1. The choice of trigger lines for this analysis differs from those in Chapter 5 because the inclusive ϕ and topological two-body lines preferentially select $B_s^0 \to \phi\phi$ events, causing a sharp peak in the trigger efficiency distribution at low $m(K^+K^-)$. The trigger efficiency of the $B_s^0 \to \phi\phi$ simulation sample is not guaranteed to be the same as in Table 5.1 because the efficiencies are calculated after the samples have passed two different stripping lines.

6.2.3 Cut-based selection

The $B_s^0 \to \phi\rho^0$ stripping line (Section 4.4) used in this analysis is designed to select the decay mode $B_s^0 \to \phi\pi^+\pi^-$ without particle identification or invariant mass requirements on the pions, making it suitable for studies of $B_s^0 \to \phi K^+K^-$

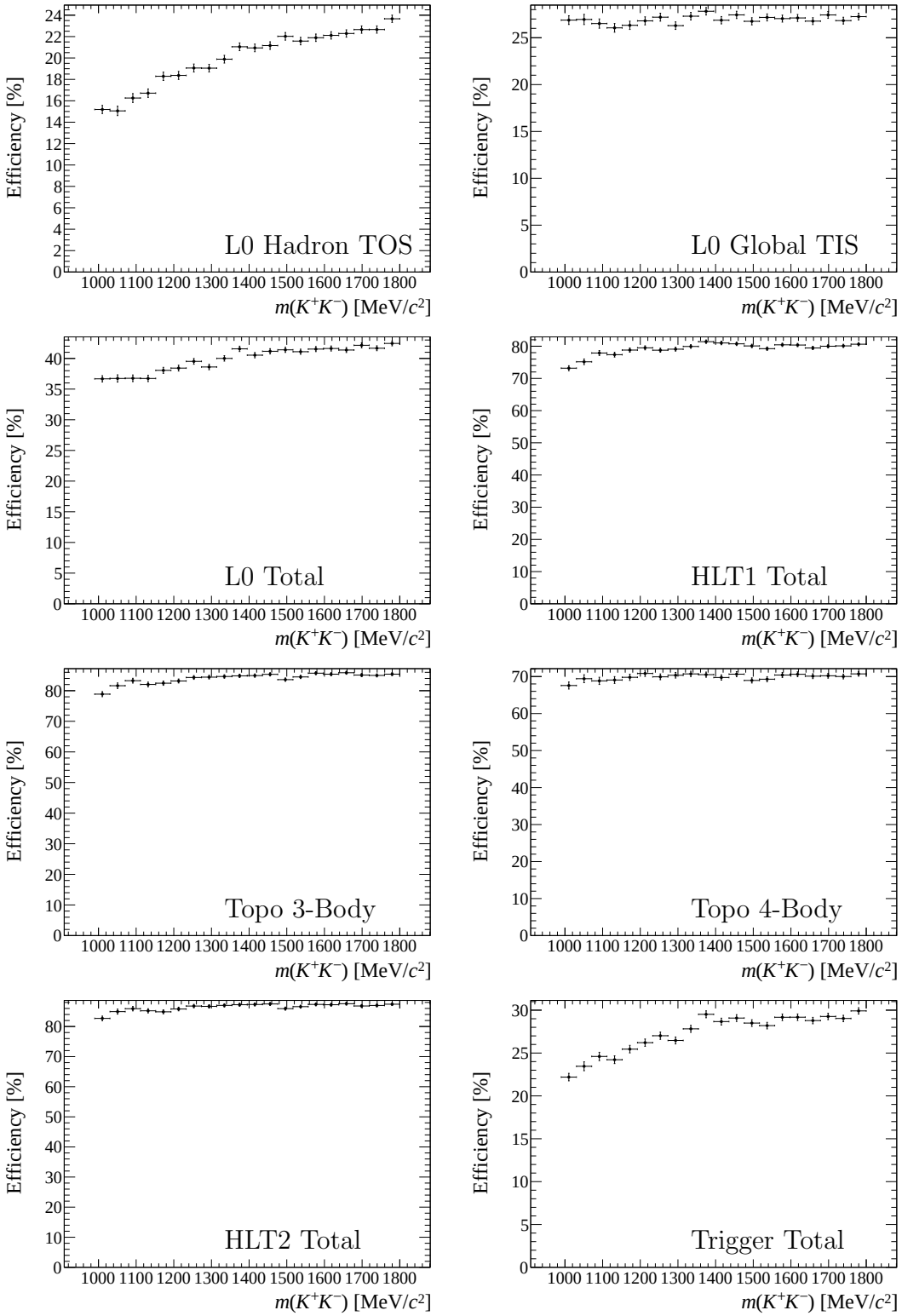


Figure 6.1 *Trigger efficiency in bins of $m(K^+K^-)$ using the $B_s^0 \to \phi K^+K^-$ simulation sample.*

decays. One track pair is combined to form a $\phi \to K^+K^-$ candidate, and the other a $\rho^0 \to \pi^+\pi^-$ candidate. These are then combined to form a B_s^0 candidate.

In the stripping, charged particles are required to have a ghost probability less than 0.5 (later tightened to 0.3) and $p_{\mathrm{T}} > 500$ MeV/c. The quality of the track reconstruction is ensured by requiring that the track fit χ^2/ndf is less than 3. The kaon tracks are required to have a log-likelihood of being a kaon greater than that of being a pion: $\Delta\ln(\mathcal{L}_{K-\pi}) > 0$, which is later imposed on all tracks. In order to exclude tracks originating from the primary vertex, the χ^2_{IP} of the kaon candidates is required to be greater than 16, which is also later imposed on the pion candidates. The ϕ and ρ^0 candidates are both required to have $p_T > 900$ MeV/c and $p > 1000$ MeV/c. The tracks that form the ϕ candidate are required to have a distance of closest approach chi-squared, χ^2_{DOCA}, less than 30. Prompt ϕ mesons are removed by requiring $\chi^2_{\mathrm{IP}} > 10$, later tightened to 16, and a loose mass window of $|m(K^+K^-) - m(\phi)| < 25$ MeV/c^2 is applied, where $m(\phi)$ comes from Ref. [135]. The vertex quality of the ϕ meson candidates is ensured by requiring $\chi^2_{\mathrm{Vtx}} < 25$. To ensure equal treatment of the daughter mesons, the same χ^2_{IP} and χ^2_{Vtx} cuts are later applied to the ρ^0 candidate. The ρ^0 candidate is also required to have $m(\pi^+\pi^-) < 4000$ MeV/c^2. Finally, the B_s^0 candidate is required to lie in the mass range $4800 < m(K^+K^-\pi^+\pi^-) < 5600$ MeV/c^2 and to have a vertex χ^2 per degree-of-freedom, χ^2_{Vtx}/ndf, less than 9 and $\chi^2_{\mathrm{IP}} < 20$ to ensure that it originates from the primary vertex.

After the stripping step, the candidates are reconstructed under the $K^+K^-K^+K^-$ mass hypothesis. In order to remove kaons that have decayed in flight, particles that have matching muon information are rejected. To suppress background from misidentified hadrons, cuts are made on the probability, P_h, for the particle to be a certain hadron, h, as described in Section 4.1. Kaon candidates are required to have $P_K \times (1 - P_\pi) > 0.025$, in order to suppress misidentified pions, and $P_K \times (1 - P_p) > 0.01$ in order to suppress misidentified protons. These two cuts retain $(95.14 \pm 0.05)\%$ and $(99.21 \pm 0.02)\,\%$ of the $B_s^0 \to \phi K^+K^-$ simulation events, respectively.

The invariant mass cut on the candidate $\phi \to K^+K^-$ decays is tightened to $|m(K^+K^-) - m(\phi)| < 15$ MeV/c^2 to agree with previous LHCb analyses of $B_s^0 \to \phi\phi$ [1, 49, 156, 157]. To ensure significant displacement from the primary vertex, the B_s^0 candidates are also required to have a flight distance χ^2, χ^2_{FD}, greater than 250. Where events have multiple candidates, they are ordered in terms of χ^2_{Vtx}. The first candidate is kept, along with any others that do not share

tracks. The above cuts are summarised in Table 6.2, along with the fraction of $B_s^0 \to \phi K^+ K^-$ and $B_s^0 \to \phi\phi$ simulation events that satisfy each condition.

Table 6.2 *Efficiencies of the offline selection cuts on signal simulation events and upper sideband data events with $m(\phi K^+ K^-) > 5500$ MeV/c^2 after applying the stripping and trigger requirements. The simulation samples have additional cuts on the background category, as described in the text.*

Cut	Fraction of events kept [%]		
	$B_s^0 \to \phi K^+ K^-$	$B_s^0 \to \phi\phi$	Sideband
$P(\text{ghost}) < 0.3$	99.084 ± 0.023	99.12 ± 0.04	87.64 ± 0.10
Remove decay-in-flight K^+	94.04 ± 0.06	94.25 ± 0.09	73.68 ± 0.13
K^+ $p_T > 500$ MeV/c	99.922 ± 0.007	99.947 ± 0.009	99.447 ± 0.022
$m(\phi) \pm 15$ MeV/c^2	95.36 ± 0.05	95.91 ± 0.08	77.66 ± 0.12
$\chi^2_{\text{FD}} > 250$	99.081 ± 0.023	98.30 ± 0.05	92.72 ± 0.08
$P_K \times (1 - P_\pi) > 0.025$	95.14 ± 0.05	95.83 ± 0.08	7.83 ± 0.08
$P_K \times (1 - P_p) > 0.01$	99.212 ± 0.021	99.20 ± 0.04	22.35 ± 0.12
Overall cut efficiency	74.61 ± 0.10	79.40 ± 0.16	1.79 ± 0.04

The simulation samples are subject to additional background category requirements [158], which use the 'truth' information from the generator. For the signal simulation samples, events are required to correspond to true signal or signal plus radiative photons, which allows for proper modelling of the radiative tail in the $m(K^+K^-K^+K^-)$ distribution. For the background simulation samples, the events are required to correspond to misidentified background, which removes any combinatorial component.

6.2.4 Removal of specific backgrounds

Despite the particle identification requirements, there remain some background events from specific decay processes. Several open charm and four-body charmless beauty decays, with a misidentified particle, are identified and vetoed.

The background contribution from the decay $B_s^0 \to D_s^- K^+$, with $D_s^- \to \phi K^-$, is removed by vetoing events with $|m(K^+K^-K^-) - m(D_s^-)| < 24$ MeV/c^2, corresponding to approximately a $\pm 3\sigma$ window around the D_s^- mass. Figure 6.2 shows the effect of this cut on the $m(K^+K^-K^-)$ distribution.

Backgrounds from several decay modes with one or more tracks misidentified as a kaon are considered. Specifically, these are $B^0 \to \phi K^{*0}$, $\Lambda_b^0 \to \phi p K^-$,

$D^-/D_s^- \to \phi\pi^-$ and $\Lambda_c^+ \to \phi p$. For each of the background sources, both of the tracks from the non-ϕ pair are considered as being misidentified pions or protons. Vetoes or tighter PID cuts are applied in invariant mass windows constructed under both charge-conjugate hypotheses.

The background contribution from $D^-/D_s^- \to \phi\pi^-$ is removed by vetoing events with $|m(K^+K^-\pi^-) - m(D_s^-)| < 24$ MeV/c^2 or $|m(K^+K^-\pi^-) - m(D^-)| < 24$ MeV/c^2. Figure 6.3 shows the effect of this cut on the $m(K^+K^-\pi^-)$ distribution. Similarly, the contribution from $\Lambda_c^+ \to \phi p$ is suppressed by requiring the non-ϕ kaon candidates to have $P_K > P_p$ for events where $|m(K^+K^-p) - m(\Lambda_c^+)| < 24$ MeV/c^2. Figure 6.4 shows the effect of this cut on the $m(K^+K^-p)$ distribution. It should be noted that the $\Lambda_c^+ \to \phi p$ veto does not remove any events with $m(K^+K^-) < 1.8$ GeV/c^2.

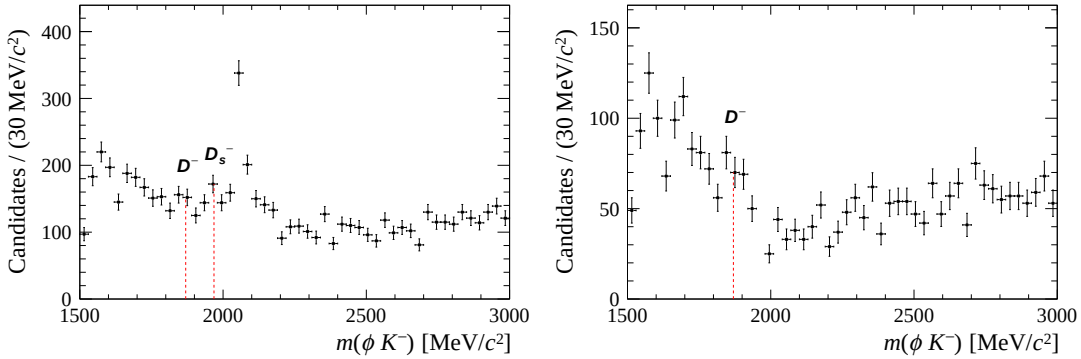


Figure 6.2 *The ϕK^- invariant mass spectrum before (left) and after (right) the mass vetoes. The prominent peak at ~ 2.05 GeV/c^2 is due to the decay $D_s^- \to \phi\pi^-$, visible in Figure 6.3.*

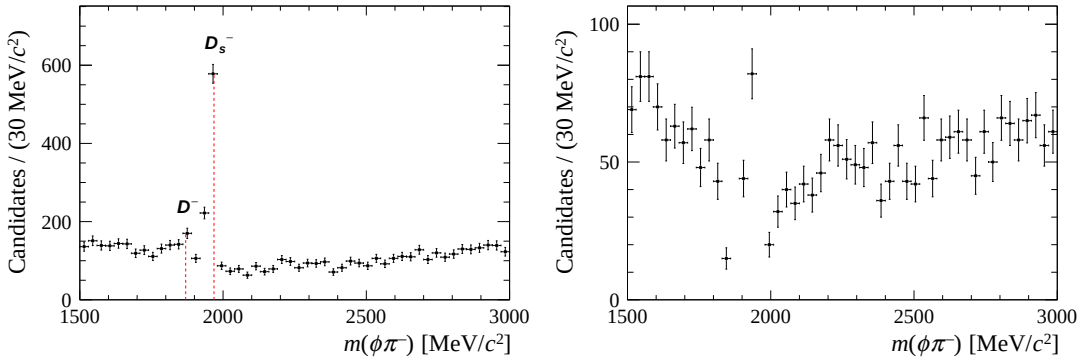


Figure 6.3 *The $\phi\pi^-$ invariant mass spectrum before (left) and after (right) the mass vetoes.*

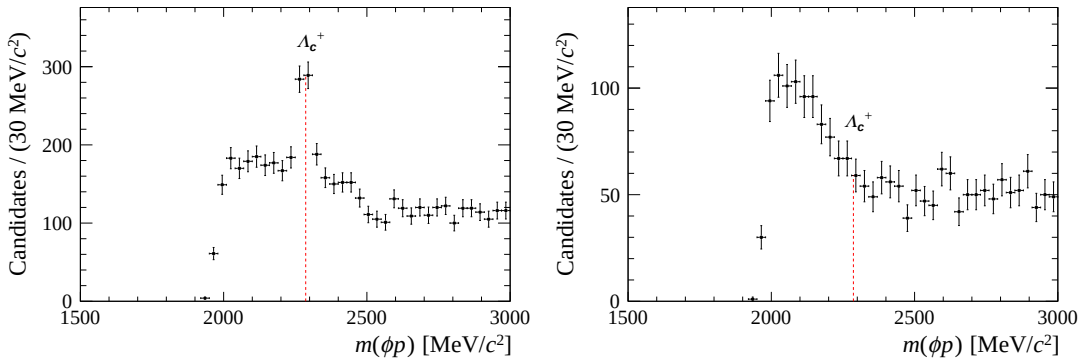


Figure 6.4 *The ϕp invariant mass spectrum before (left) and after (right) the mass vetoes.*

Two of the potential backgrounds peak under the B_s^0 peak under the four-kaon mass hypothesis. For these, tighter track particle identification requirements are used instead of mass vetoes. The values of these cuts were chosen to reject 90% of the $B^0 \to \phi K^{*0}$ and $\Lambda_b^0 \to \phi p K^-$ simulation events following the cut-based selection.

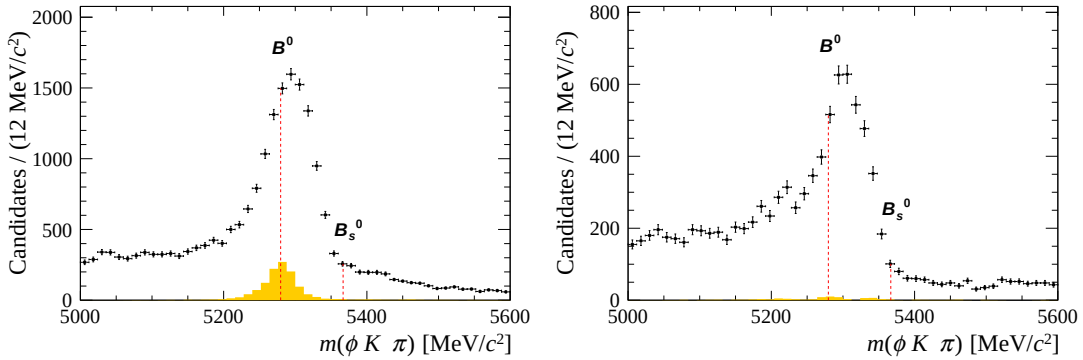


Figure 6.5 *The $\phi K^+\pi^-$ invariant mass spectrum before (left) and after (right) the tighter PID cuts. The orange histogram is the $B^0 \to \phi K^{*0}$ simulation with the same selection applied.*

The background contribution from the decay $B^0 \to \phi K^{*0}$ is rejected by requiring the non-ϕ kaon candidate to have $P_K \times (1 - P_\pi) > 0.4$ for events where $|m(K^+K^-\pi^+K^-) - m(B^0)| < 50$ MeV/c^2, corresponding to approximately a $\pm 3\sigma$ window around the B^0 mass. Figure 6.5 shows the effect of this cut on the $m(K^+K^-K^+\pi^-)$ distribution. Similarly, the contribution from $\Lambda_b^0 \to \phi p K^-$ is rejected by requiring the non-ϕ kaon candidate to have $P_K \times (1 - P_p) > 0.5$ for events where $|m(K^+K^-pK^-) - m(\Lambda_b^0)| < 50$ MeV/c^2. Figure 6.6 shows the effect

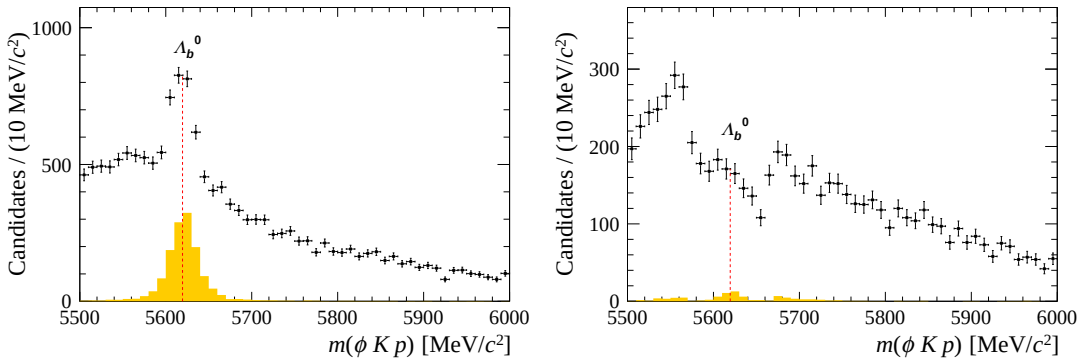


Figure 6.6 *The $\phi p K^-$ invariant mass spectrum before (left) and after (right) the tighter PID cuts. The orange histogram is the $\Lambda_b^0 \to \phi p K^-$ simulation with the same selection applied.*

of this cut on the $m(K^+K^-pK^-)$ distribution.

Table 6.3 *Efficiencies of the cuts designed to remove peaking backgrounds on signal simulation events with $m(K^+K^-) < 1.8$ GeV/c^2 after stripping, triggers and the cuts in Section 6.2.3.*

Background-specific cut	Fraction of events kept [%] $B_s^0 \to \phi K^+K^-$ simulation
$B^0 \to \phi K^{*0}$	74.94 ± 0.19
$\Lambda_b^0 \to \phi p K^-$	92.56 ± 0.12
$\Lambda_c^+ \to \phi p$	100 ± 0
$D_s^- \to \phi K^-$	99.739 ± 0.023
$D^- \to \phi \pi^-$	99.381 ± 0.035
$D_s^- \to \phi \pi^-$	99.16 ± 0.04
Combination of the above	70.47 ± 0.20

Table 6.3 summarises the efficiencies of each of the peaking background cuts on the $B_s^0 \to \phi K^+K^-$ simulation with $m(K^+K^-) < 1.8$ GeV/c^2, after the cuts described in Section 6.2.3. The overall signal rejection rate is dominated by the cut designed to remove 90% of the $B^0 \to \phi K^{*0}$ background.

6.2.5 Multivariate selection

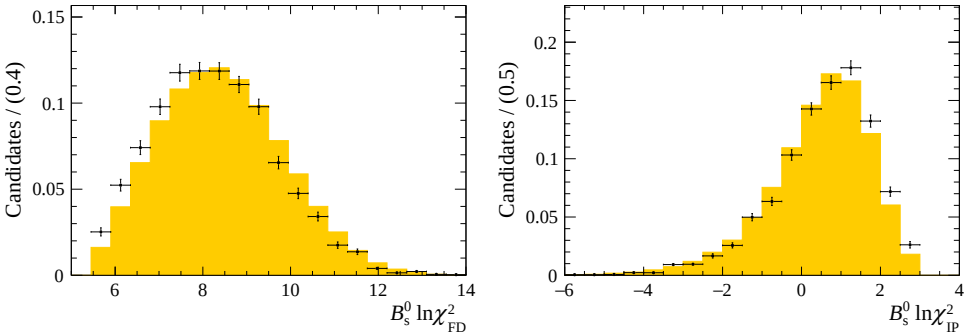
After the triggers, stripping, offline cuts and mass vetoes, multivariate analysis (MVA) algorithms are trained to further reduce the combinatorial background. Four different MVA algorithms from the TMVA package [127] were tested in order to find the best background rejection and signal efficiency. These are

three types of Boosted Decision Tree with adaptive boosting (BDT), gradient boosting (BDTG), decorrelation and adaptive boosting (BDTD) respectively, and a Multilayer Perceptron (MLP), a type of artificial neural network.

Seven discriminating variables are used to train the MVA algorithms: the minimum values of p_T and $\ln\chi^2_{\rm IP}$ out of the kaons and the p_T, η, $\ln\chi^2_{\rm IP}$, $\ln\chi^2_{\rm FD}$ and $\ln\chi^2_{\rm Vtx}$ for the B^0_s candidate. Figure 6.7 shows the distributions of these variables in the simulation and data weighted using the *sPlot* technique [139] (s-weighted data). The *sPlot* technique is described in Section 5.2.4. It can be seen that the agreement is reasonable: the discriminating variables need not agree perfectly between data and simulation in order for the classifier to distinguish signal from background.

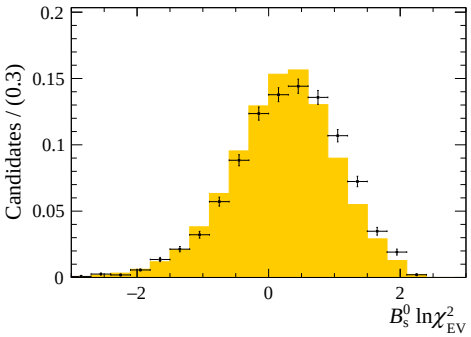
The decay $B^0_s \to \phi\phi$ is expected to be the dominant contribution to the $B^0_s \to \phi K^+K^-$ final state in the region $m(K^+K^-) < 1800$ MeV/c^2. This decay mode is already well studied [1, 48, 49, 134, 156, 157]. Hence this analysis focuses on the other resonant components, such as $B^0_s \to \phi f'_2(1525)$. The signal and background samples used in the training and optimisation have a cut of $1050 < m(K^+K^-) < 1800$ MeV/c^2. The background sample consists of the upper sideband data in the range $m(K^+K^-K^+K^-) > 5500$ MeV/c^2. Figure 6.8 shows the distributions of the discriminating variables in the signal and background training samples. It can be seen from the differences between the signal and background distributions that each of the discriminating variables has some separation power. Figure 6.9 shows the correlation matrices of the discriminating variables in the signal and background training samples. The $p_{\rm T}$ and η variables are (anti)correlated to some extent, as are the minimum kaon $\ln\chi^2_{\rm IP}$ and the B^0_s $\ln\chi^2_{\rm FD}$.

Figure 6.10 shows the background rejection versus signal efficiency (ROC curve) of the different MVA algorithm types trained. The MLP algorithm gives the best background rejection for all values of signal efficiency. Figure 6.11 shows the distribution of the MVA outputs for the MLP algorithm, with the training samples overlaid on top of the test samples. The probabilities returned by the Kolmogorov–Smirnov test [159, 160] are reasonable and suggest no overtraining. A comparison of the distributions of the MLP output in simulation and s-weighted data is shown in Figure 6.12, which shows good agreement despite the imperfect agreement of the training variables shown in Figure 6.7. Supplementary plots from the MVA training are presented in Appendix C.

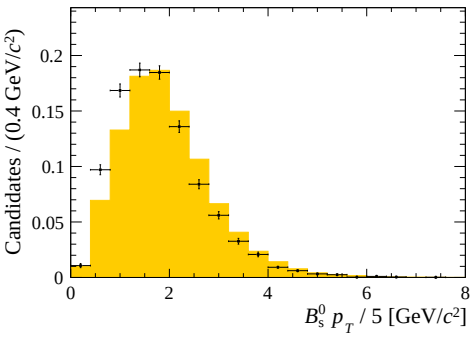


(a) B_s^0 ln $\chi^2_{\rm FD}$

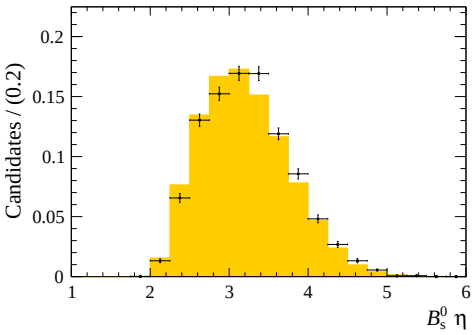
(b) B_s^0 ln $\chi^2_{\rm IP}$



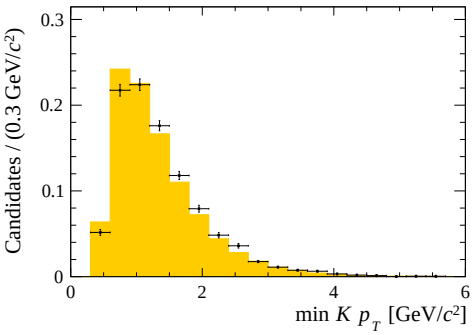
(c) B_s^0 ln $\chi^2_{\rm Vtx}$



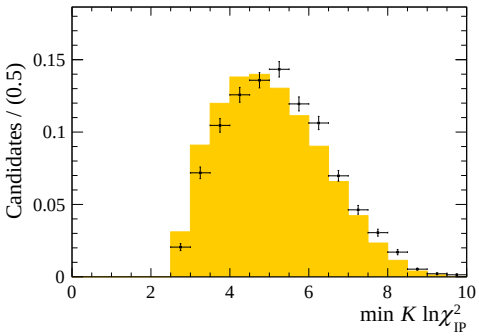
(d) B_s^0 $p_{\rm T}$



(e) B_s^0 η

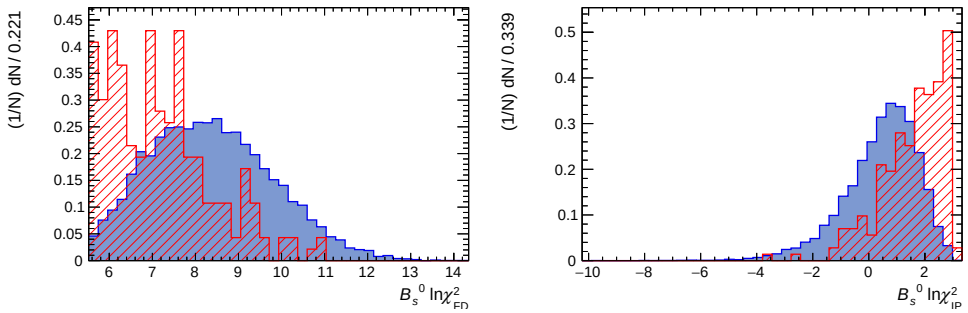


(f) *min.* $K^{\pm}$ $p_{\rm T}$



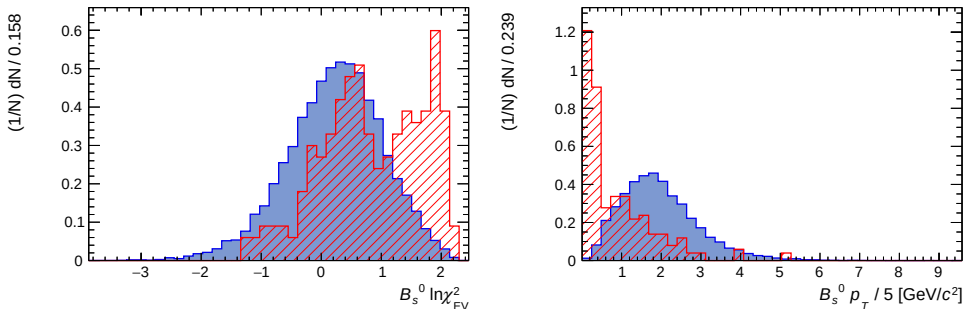
(g) *min.* $K^{\pm}$ ln $\chi^2_{\rm IP}$

Figure 6.7 *Comparison of the discriminating variables between $B_s^0 \to \phi K^+ K^-$ simulation (yellow histogram) and s-weighted data (black points). The distributions are normalised to unit area.*



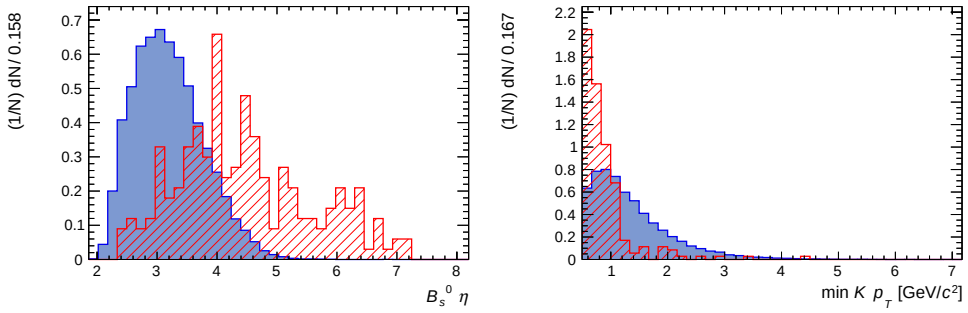
(a) B_s^0 $\ln \chi^2_{\mathrm{FD}}$

(b) B_s^0 $\ln \chi^2_{\mathrm{IP}}$



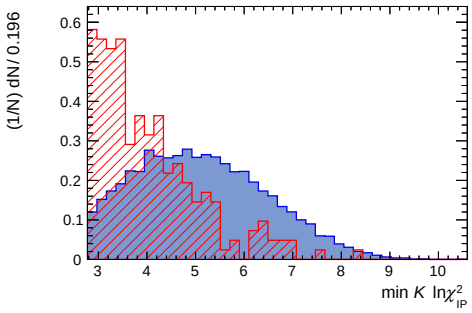
(c) B_s^0 $\ln \chi^2_{\mathrm{Vtx}}$

(d) B_s^0 p_{T}



(e) B_s^0 η

(f) *min.* $K^{\pm}$ p_{T}



(g) *min.* $K^{\pm}$ $\ln \chi^2_{\mathrm{IP}}$

Figure 6.8 *Comparison of the discriminating variables in the signal (blue) and background (red) samples used to train the MVA. The distributions are normalised to unit area.*

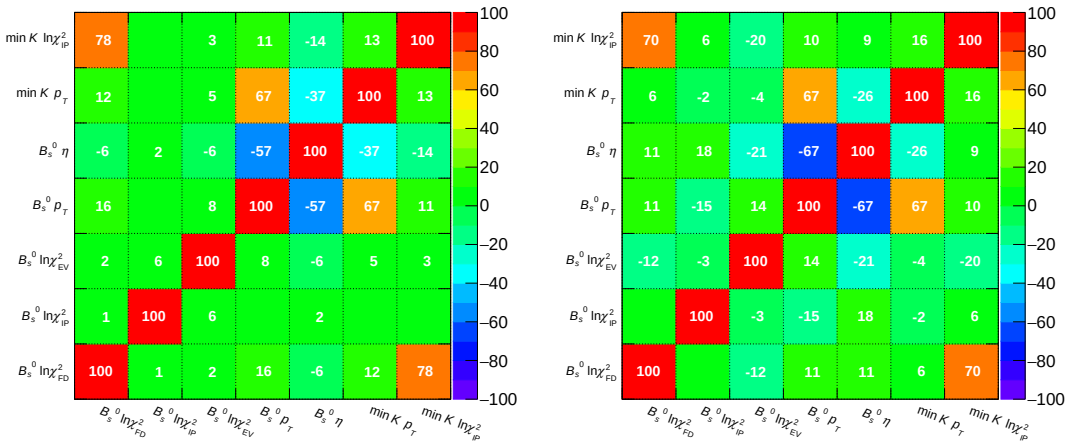


Figure 6.9 *Correlation matrices of the discriminating variables in the signal (left) and background (right) samples.*

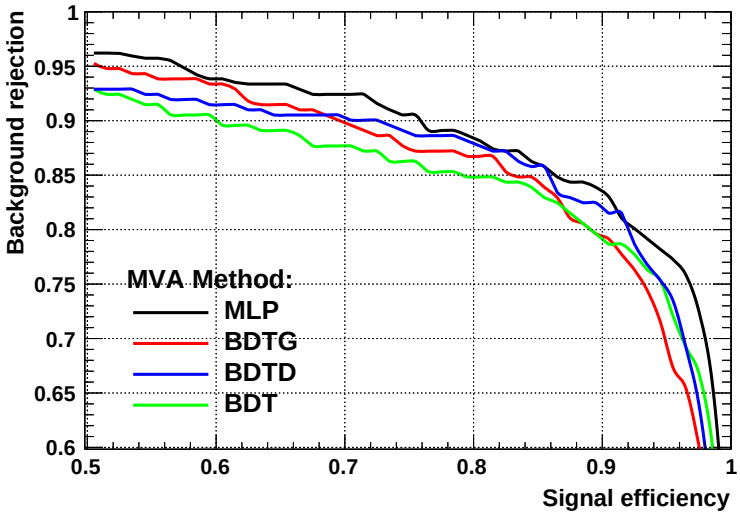


Figure 6.10 *ROC curves for the different MVA types.*

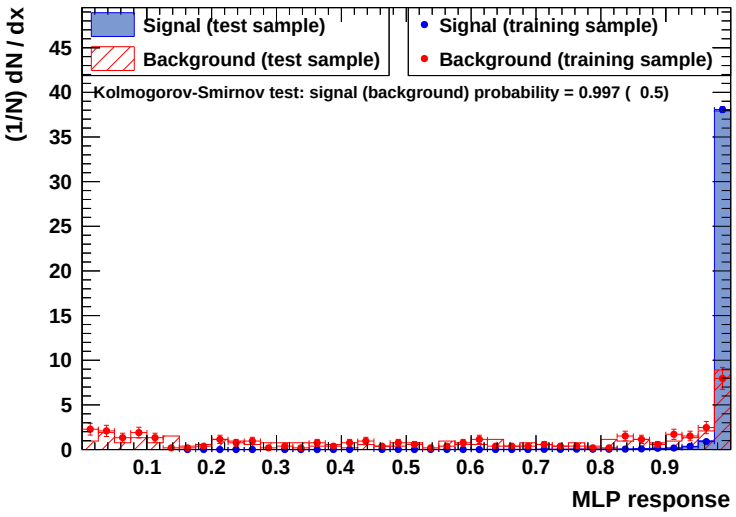


Figure 6.11 *Distribution of the MVA output for the signal and background samples. The test samples are plotted as histograms, and the training samples are plotted as points.*

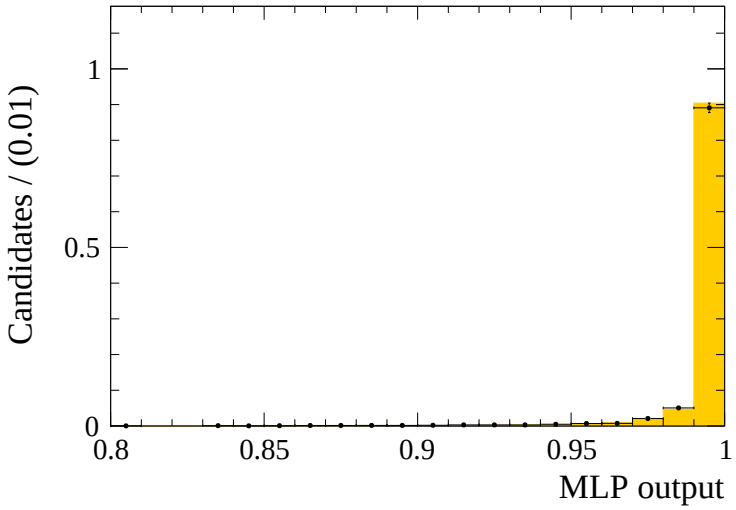


Figure 6.12 *Comparison of MLP output between $B_s^0 \to \phi K^+ K^-$ simulation (yellow histogram) and s-weighted data (black points). The distributions are normalised to unit area.*

In order to choose the optimal cut on the MVA output, the signal significance figure of merit, defined as

$$FOM = \frac{S}{\sqrt{S+B}},$$

is maximised, where $S = S_0 \cdot \varepsilon_S$ is the product of the initial signal yield, S_0, in data and the cut efficiency, ε_S, on the simulation sample; similarly, $B = B_0 \cdot \varepsilon_B$ is the product of the initial background yield, B_0 in data and the cut efficiency, ε_B on the sideband sample.

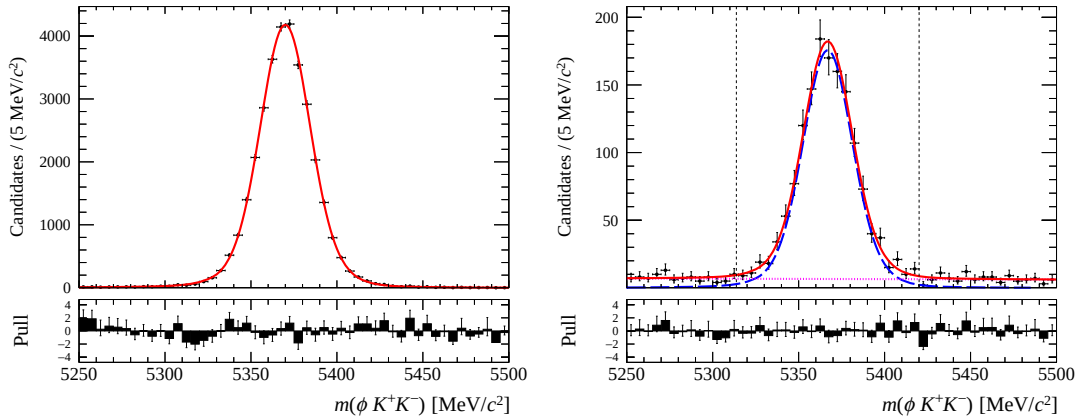


Figure 6.13 *Fits to the $m(K^+K^-K^+K^-)$ distributions of the $B_s^0 \to \phi K^+K^-$ simulation (left) and data (right) with $1.05 < m(K^+K^-) < 1.80$ GeV/c^2. The B_s^0 signal component is the blue long-dashed line, and the combinatorial background component is the purple dotted line. Vertical black dotted lines denote a window of $\pm 3\sigma$ around the mean of the B_s^0 peak.*

To determine these initial yields, fits are performed to the $m(K^+K^-K^+K^-)$ distribution of $B_s^0 \to \phi K^+K^-$ simulation and data with the cut-based selection applied and an additional cut of $1.05 < m(K^+K^-) < 1.80$ GeV/c^2. The B_s^0 signal component is modelled as a sum of a Crystal Ball function plus two Gaussian functions with a shared mean. The Crystal Ball tail parameter n is fixed to 1. The shape parameters of the signal component in the fit to data are fixed to the results of the fit to simulation. The combinatorial background component is modelled as an exponential function. These fits are shown in Figure 6.13. The values of S_0 and B_0 are found by integrating the signal and background PDFs in a window $\pm 3\sigma$ around the B_s^0 mass, which is marked by black vertical dashed lines. This gives $S_0 = 1320 \pm 40$ and $B_0 = 139 \pm 11$.

The MLP algorithm gives the highest value of significance and is chosen to be used for the rest of this analysis. The efficiency and significance curves for all four algorithms are shown in Figure C.2 in Appendix C. The optimal

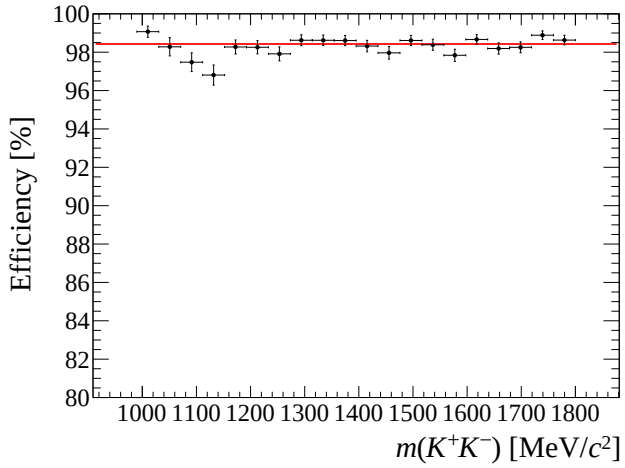


Figure 6.14 *Efficiency of the optimal cut on the MLP output as a function of $m(K^+K^-)$ using the $B_s^0 \to \phi K^+K^-$ simulation. Note that the vertical axis starts from 80 %. A fit using a constant function is superimposed.*

cut on the MLP output is at a value of 0.88. The efficiencies of this cut on the $B_s^0 \to \phi K^+K^-$ and $B_s^0 \to \phi\phi$ simulation samples are $(98.80 \pm 0.06)\,\%$ and $(98.77 \pm 0.06)\,\%$, respectively. Using the yields of fits made to the data sample with $|m(K^+K^-) - m(\phi)| < 15$ MeV/c^2, the efficiency of this cut on the signal is determined to be $(95.6 \pm 2.9)\,\%$, which is in good agreement with the simulation. The rejection rate on the sideband sample is $(65 \pm 2)\,\%$. As a further check, the efficiency of this cut as a function of $m(K^+K^-)$ for the $B_s^0 \to \phi K^+K^-$ simulation sample is shown in Figure 6.14. It can be seen that the cut does not bias the distribution.

6.3 Peaking background estimation

The decay modes $\Lambda_b^0 \to \phi p K^-$ and $B^0 \to \phi K^{*0}$, where $K^{*0} \to K^+\pi^-$, are considered as possible peaking backgrounds that enter the data sample due to one hadron misidentified as a kaon. When constructed under the ϕK^+K^- hypothesis, the invariant mass distributions of these two modes have peaks near the B_s^0 mass.

In order to quantify the contribution of the $B^0 \to \phi K^{*0}$ background to the data sample, the yield N is estimated using:

$$\frac{N_{B^0\to\phi K^{*0}}}{N_{B_s^0\to\phi\phi}} = \frac{N^{\rm sel}_{B^0\to\phi K^{*0}}}{N^{\rm sel}_{B_s^0\to\phi\phi}} \frac{N^{\rm gen}_{B_s^0\to\phi\phi}}{N^{\rm gen}_{B^0\to\phi K^{*0}}} \frac{\mathcal{B}(B^0\to\phi K^{*0})}{\mathcal{B}(B_s^0\to\phi\phi)} \frac{\mathcal{B}(K^{*0}\to K^+\pi^-)}{\mathcal{B}(\phi\to K^+K^-)} \frac{f_d}{f_s}, \qquad (6.1)$$

where N^{gen} is the number of generated events, corrected for the generator-level efficiency, N^{sel} is the number of simulation events that pass all the selection requirements given in Section 6.2, and f_s and f_d are the fragmentation fractions for the B_s^0 and B^0 respectively. The value of f_s/f_d is taken from an average of LHCb results as 0.259 ± 0.015 [148–150]. For the purpose of this section, the number of $B_s^0 \to \phi\phi$ events is assumed to be the signal yield from the fit to data with a cut of $m(K^+K^-) < 1.05$ GeV/c^2. The values of the $\phi \to K^+K^-$, $K^{*0} \to K^+\pi^-$ and $B^0 \to \phi K^{*0}$ branching fractions are obtained in the same way as in Section 5.6. The number of expected $B^0 \to \phi K^{*0}$ events in the full $m(K^+K^-K^+K^-)$ range of the selected data is found to be 23 ± 3. This number is used to fix the size of the $B^0 \to \phi K^{*0}$ component in the fit described in Section 6.4.

In Figure 6.6, there is a significant Λ_b^0 peak in the data reconstructed under the $\phi p K^-$ hypothesis before the vetoes are applied. The yield of this peak is determined from a fit to this distribution. The Λ_b^0 peak is modelled as the sum of a Crystal Ball [140] and a Gaussian function with a shared mean:

$$P_{\Lambda_b^0}(m) = f P_{\text{CB}}(m|\mu, \sigma_1, \alpha, n) + (1-f) P_{\text{G}}(m|\mu, \sigma_2). \tag{6.2}$$

The parameter n is fixed to 1, and the parameters f, $\sigma_{\{1,2\}}$ and α are fixed to the results of a fit to the $\Lambda_b^0 \to \phi p K^-$ simulation. The fit to the $m(K^+K^-pK^-)$ distribution in data is performed in a window of ± 60 MeV/c^2 around the Λ_b^0 mass using an exponential function to model the non-Λ_b^0 events. This results in a yield of 1240 ± 70 events, which is multiplied by the efficiency of the vetoes described in Section 6.2.4 on the $\Lambda_b^0 \to \phi p K^-$ simulation to give an expected yield of 88 ± 7 in the full $m(K^+K^-K^+K^-)$ range of the selected data. The fit is shown in Figure 6.15.

To verify this value of the $\Lambda_b^0 \to \phi p K^-$ yield to within an order of magnitude, it is estimated using

$$\frac{N_{\Lambda_b^0 \to \phi p K^-}}{N_{B_s^0 \to \phi\phi}} = \frac{N^{\text{sel}}_{\Lambda_b^0 \to \phi p K^-}}{N^{\text{sel}}_{B_s^0 \to \phi\phi}} \frac{N^{\text{gen}}_{B_s^0 \to \phi\phi}}{N^{\text{gen}}_{\Lambda_b^0 \to \phi p K^-}} \frac{\mathcal{B}(\Lambda_b^0 \to \phi p K^-)}{\mathcal{B}(B_s^0 \to \phi\phi)} \frac{1}{\mathcal{B}(\phi \to K^+K^-)} \frac{f_{\Lambda_b^0}}{f_s}, \tag{6.3}$$

where the symbols take the same meanings as in Equation 6.1, under the assumption that

$$\frac{\mathcal{B}(\Lambda_b^0 \to \phi p K^-)}{\mathcal{B}(B_s^0 \to \phi\phi)} = \frac{\mathcal{B}(\Lambda_b^0 \to J/\psi\, p K^-)}{\mathcal{B}(B_s^0 \to J/\psi\, \phi)}, \tag{6.4}$$

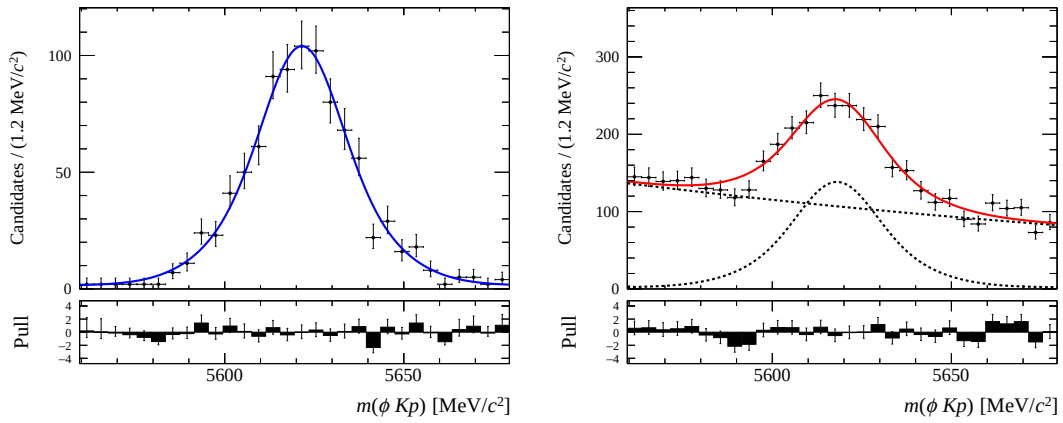


Figure 6.15 *Fits to the $m(K^+K^-pK^-)$ distributions in $\Lambda_b^0 \to \phi p K^-$ simulation (left) and data (right) with the cuts described in Section 6.2.3 applied. The 'signal' Λ_b^0 peak shape is shown as a solid blue line in the simulation fit. The two components of the data fit are the Λ_b^0 'signal' and a 'background' component (mostly misreconstructed $B_s^0 \to \phi K^+K^-$ events) modelled as an exponential curve.*

and the branching fractions for $\Lambda_b^0 \to J/\psi\, pK^-$ and $B_s^0 \to J/\psi\, \phi$ are taken from the PDG [135]. The value of $f_{\Lambda_b^0}/f_s$ is calculated as

$$\frac{f_{\Lambda_b^0}}{f_s} = \frac{f_{\Lambda_b^0}}{f_u + f_d}\frac{f_u + f_d}{f_s}, \tag{6.5}$$

where $f_s/(f_u + f_d) = 0.134 \pm 0.011$ and $f_{\Lambda_b^0}/(f_u + f_d) = (0.404 \pm 0.109)(1 - (0.031 \pm 0.005)\frac{p_{\rm T}}{1\ {\rm GeV}/c})$ [148]. The value of $p_{\rm T}$ used to calculate this fragmentation fraction is the average $p_{\rm T}$ of the Λ_b^0 candidates in the fully selected $\Lambda_b^0 \to \phi p K^-$ simulation sample: 10.4 GeV/c. This results in an expected $\Lambda_b^0 \to \phi p K^-$ yield of $\sim 41 \pm 14$, roughly half the size of the yield obtained from fitting the $m(K^+K^-pK^-)$ distribution. Considering the assumption in Equation 6.4, this estimate is in reasonable agreement with the value obtained from fitting the data.

6.4 Fit to the $K^+K^-K^+K^-$ invariant mass

A fit to the ϕK^+K^- invariant mass distribution is performed in order to determine the number of signal and background candidates in the data used for the amplitude fit. The invariant mass range $5150 < m(K^+K^-K^+K^-) < 5600$ MeV/c^2 is used so that the shape of the combinatorial background is properly modelled.

The $B_s^0 \to \phi K^+K^-$ component of the $m(K^+K^-K^+K^-)$ spectrum is modelled

using the sum of a Crystal Ball shape [140], P_{CB}, and two Gaussian distributions, P_{G}, given by:

$$P_{B_s^0}(m) = f_1 P_{\mathrm{CB}}(m|\mu, \sigma_1, \alpha, n) + (1 - f_1)\big[f_2 P_{\mathrm{G}}(m|\mu, \sigma_2) + (1 - f_2) P_{\mathrm{G}}(m|\mu, \sigma_3)\big], \tag{6.6}$$

where μ is the shared mean of the three components, $\sigma_{\{1,2,3\}}$ are the Gaussian widths of each component, and α and n describe the position and shape of the power-law tail of the Crystal Ball distribution. The parameters $f_{\{1,2\}}$ are factors that control the relative amplitude of each component. They are required to be between 0 and 1. The combinatorial background is modelled by an exponential function.

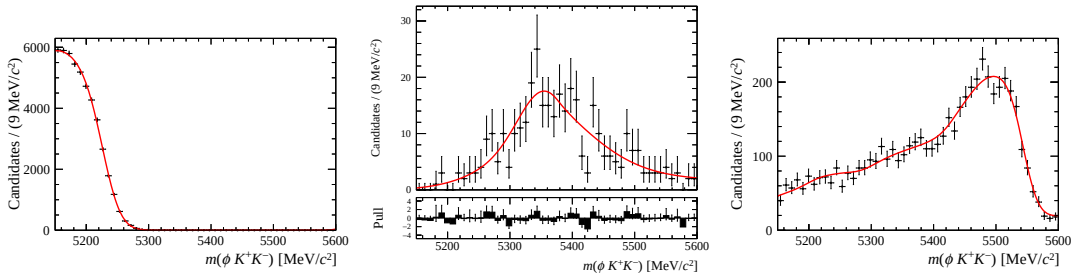


Figure 6.16 *The $m(K^+K^-K^+K^-)$ distributions used to model the $B_s^0 \to \phi f_1(1420)$ (left), $B^0 \to \phi K^{*0}$ (centre) and $\Lambda_b^0 \to \phi p K^-$ (right) peaking background components. The convolved histogram shapes are shown as solid red lines on the left and right plots, overlaid onto smeared generator-level events. The PDF used to model the $B^0 \to \phi K^{*0}$ background is shown as a red line on the centre plot.*

Three peaking background components are considered: misidentified $B^0 \to \phi K^{*0}$ and $\Lambda_b^0 \to \phi p K^-$ decays and partially reconstructed $B_s^0 \to \phi K^+ K^- \pi^0$ decays. The shape of the $B^0 \to \phi K^{*0}$ component is modelled using a Crystal Ball plus a Gaussian function with a shared mean. The parameters of this shape are fixed to the results of a fit to the fully selected $B^0 \to \phi K^{*0}$ simulation sample, shown in the centre of Figure 6.16. Due to a lack of simulation events after the full selection, the $\Lambda_b^0 \to \phi p K^-$ component is modelled using a histogram of generator-level $\Lambda_b^0 \to \phi p K^-$ events, produced using the fast generation package RapidSim [161]. The background-specific cut, applied under the $m(K^+K^-pK^-)$ hypothesis, changes the shape of the $m(K^+K^-K^+K^-)$ distribution. This effect is simulated by categorising events as falling in either, both or neither of the $m(K^+K^-pK^-)$ and $m(K^+K^-K^+\overline{p})$ windows. The cut efficiency for each of these categories is found from the full $\Lambda_b^0 \to \phi p K^-$ simulation and then used to weight the corresponding generator-level events. This results in the distribution shown on the right of

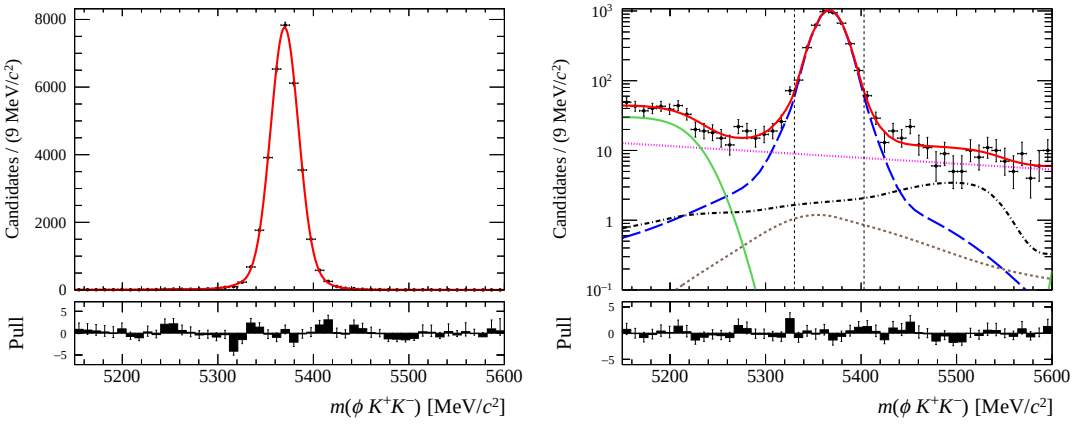


Figure 6.17 *Fits to $m(K^+K^-K^+K^-)$ using the fully selected phase-space simulation (left) and data (right). The total fit is in red, the B_s^0 component is a blue long-dashed line, the partially reconstructed background is a solid green line, the $B^0 \to \phi K^{*0}$ background is a dotted brown line, the $\Lambda_b^0 \to \phi p K^-$ background is a dot-dashed black line, and the combinatorial background component is a purple dotted line. The pull distribution is shown underneath the fit. The black vertical dashed lines mark a region of $\pm 2\sigma_{eff}$ around the fitted B_s^0 mass, where σ_{eff} is the fitted width of the peak.*

Figure 6.16. The partially reconstructed background component is modelled using a histogram of the $m(K^+K^-K^+K^-)$ distribution in $B_s^0 \to \phi f_1(1420)(\to K^+K^-\pi^0)$ generator-level events, also produced using RapidSim. This is shown on the left of Figure 6.16. The $B_s^0 \to \phi f_1(1420)$ decay mode is chosen because it is predicted to have a branching fraction of $\left(16.1^{+9.9}_{-7.6}\right) \times 10^{-6}$ [162], comparable to that of $B_s^0 \to \phi\phi$, and the decay mode $f_1(1420) \to K^+K^-\pi^0$ is dominant [135]. Both of the histogram PDFs are convolved with a Gaussian function with a width of 20 MeV/c^2 to account for the mass resolution. The yield of the $B^0 \to \phi K^{*0}$ component is fixed to the value found in Section 6.3, whereas the sizes of the $\Lambda_b^0 \to \phi p K^-$ and partially reconstructed backgrounds are left free in the fit.

The parameters $f_{\{1,2\}}$, $\sigma_{\{1,2,3\}}$ and α in the B_s^0 signal model, given in Equation 6.6, are fixed to the values found in a fit to the $B_s^0 \to \phi K^+K^-$ phase space simulation sample. The Crystal Ball tail shape parameter n is fixed to 1. In the fit to data, the value of μ, the yields of each component, and the slope of the exponential function that models the combinatorial background are left as free parameters. The results of the fits to simulation and data are summarised in Table 6.4 and Figure 6.17. From the pull distributions in Figure 6.17, the fit describes the simulation and data well. Using 50 bins in $m(K^+K^-K^+K^-)$, the χ^2/ndf of the fit to data is 0.90. The yield of the $\Lambda_b^0 \to \phi p K^-$ background is consistent with the

Table 6.4 *Results of the fit to $m(\phi K^+K^-)$ using the $B_s^0 \to \phi K^+K^-$ phase space simulation (top part) and data (bottom part).*

Parameter	Value	
α	2.79 ± 0.04	
σ_1	14.20 ± 0.24	MeV/c^2
σ_2	25 ± 2	MeV/c^2
σ_3	86 ± 9	MeV/c^2
f_1	0.83 ± 0.04	
f_2	0.927 ± 0.013	
σ_{eff}	18 ± 1	MeV/c^2
μ	5366.66 ± 0.25	MeV/c^2
N_{sig}	4160 ± 70	
N_{bkg}	380 ± 50	
N_{part}	230 ± 31	
$N_{\Lambda_b^0}$	88.0 ± 0	

value found in Section 6.3.

For the angular analysis, the signal region is defined as the window of $m(K^+K^-K^+K^-)$ between $\mu - 2\sigma_{\text{eff}}$ and $\mu + 2\sigma_{\text{eff}}$, where

$$\sigma_{\text{eff}} = f_1\sigma_1 + (1 - f_1)(f_2\sigma_2 + (1 - f_2)\sigma_3).$$

The yields of the signal, combinatorial background and peaking backgrounds are found by integrating the fitted PDFs over the signal region. This results in 3990 ± 60 signal events, 61 ± 7 combinatorial background events, 9 ± 1 $B^0 \to \phi K^{*0}$ events and 14.9 ± 0 $\Lambda_b^0 \to \phi p K^-$ events.

6.5 Amplitude formalism

The amplitude of the $B_s^0 \to \phi K^+K^-$ decay in the absence of $B_s^0 - \overline{B}_s^0$ mixing (or at $t = 0$) can be expressed as

$$A(m, \Phi, \theta_1, \theta_2) = \sum_{\text{resonances}} \left(\frac{p}{m_B}\right)^{L_B} \left(\frac{q}{m_0}\right)^{J} B'_{L_B}(p) B'_J(q) T(m) F(\Phi, \theta_1, \theta_2) \sqrt{pq} \tag{6.7}$$

where m is the invariant mass of the K^+K^- pair, m_0 and J are the mass and spin of the K^+K^- resonance, m_B is the pole mass of the B_s^0 meson, L_B is the minimum angular momentum between the ϕ and K^+K^- resonance, p is the momentum of the K^+K^- resonance in the B_s^0 decay frame, q is the momentum of a kaon in the

K^+K^- resonance decay frame, T is the resonance lineshape (see Section 6.5.1), F is the angular distribution (see Section 2.4.3), Φ, θ_1 and θ_2 are the helicity angles, and the functions B' are the Blatt–Weisskopf barrier factors [163, 164]. The first three B' functions are:

$$B'_0(p) = 1, \tag{6.8}$$

$$B'_1(p) = \sqrt{\frac{1+(p_0 r)^2}{1+(pr)^2}}, \tag{6.9}$$

$$B'_2(p) = \sqrt{\frac{9+3(p_0 r)^2+(p_0 r)^4}{9+3(pr)^2+(pr)^4}}, \tag{6.10}$$

where r is a parameter that characterises the interaction radius of the hadrons in the decay and p_0 is the momentum calculated using the pole mass of the appropriate resonance. The values of these radii in the amplitude fit to data are chosen to be $r = 1.0\,\text{GeV}^{-1}\,c$ for the B_s^0 meson and $r = 3.0\,\text{GeV}^{-1}\,c$ for the K^+K^- resonance. These values are also used in EvtGen [122] when generating the simulation samples. These are later varied to obtain a systematic uncertainty.

As discussed in Section 2.4.1, the amplitude in Equation 6.7, $A(0)$ and its *CP* conjugate, $\overline{A}(0)$, are converted to a time-integrated decay rate using [52]

$$\begin{aligned} \int_0^\infty \Gamma(t) + \overline{\Gamma}(t) dt &= \tfrac{1}{2}\left(|A(0)|^2 + \left|\overline{A}(0)\right|^2\right)\left(\tfrac{1}{\Gamma_L} + \tfrac{1}{\Gamma_H}\right) \\ &+ \Re\left(\overline{A}(0)A(0)^*\right)\left(\tfrac{1}{\Gamma_L} - \tfrac{1}{\Gamma_H}\right), \end{aligned} \tag{2.33}$$

which assumes equal numbers of B_s^0 and $\overline{B}_s^0$ mesons. The values of Γ_H and Γ_L are calculated from the PDG average values of the B_s^0 mass eigenstate lifetimes: $\tau_H = 1.661 \pm 0.032$ ps and $\tau_L = 1.045 \pm 0.025$ ps [135].

Since the helicity angles θ_1 and θ_2 are calculated relative to the K^+ mesons in each event, the *CP* conjugate amplitude $\overline{A}(0)$ is calculated as in Equation 6.7, but with the transformation $(\cos\theta_1, \cos\theta_2, \Phi) \to (-\cos\theta_1, -\cos\theta_2, -\Phi)$. The decay rate is convolved with a Gaussian function with a mass-dependent width to account for the $m(K^+K^-)$ resolution, as discussed in Section 6.7, and then multiplied by the acceptance function, discussed in Section 6.6, to give the final PDF to be fitted to the data.

6.5.1 Resonance lineshapes

The K^+K^- resonances which are significantly displaced from threshold are modelled using relativistic Breit–Wigner distributions [165], given as:

$$T(m|m_0,\Gamma_0) \propto \frac{1}{m_0^2 - m^2 - im_0\Gamma(m)},$$

where $\Gamma(m)$ is the mass-dependent width:

$$\Gamma(m) = \Gamma_0 \left(\frac{q}{q_0}\right)^{2J+1} \left(\frac{m_0}{m}\right) B'_J(q)^2,$$

where Γ_0 is the natural linewidth of the resonance, and the other symbols retain the same meanings as in Equation 6.7.

For $f_0(980) \to K^+K^-$, which is close to threshold, a Flatté lineshape [166] is used:

$$T(m|m_0, g_{\pi^+\pi^-}, g_{K^+K^-}) \propto \frac{1}{m_0^2 - m^2 - im_0\left(\Gamma_{\pi^+\pi^-}(m) + \Gamma_{K^+K^-}(m)\right)},$$

where the mass-dependent widths for the $\pi^+\pi^-$ and K^+K^- channels are

$$\Gamma_{\pi^+\pi^-}(m) = g_{\pi^+\pi^-}\sqrt{\frac{m^2}{4} - m_{\pi^+}^2}$$

and

$$\Gamma_{K^+K^-}(m) = \begin{cases} g_{K^+K^-}\sqrt{\frac{m^2}{4} - m_{K^+}^2} & m > 2m_{K^+} \\ \\ ig_{K^+K^-}\sqrt{m_{K^+}^2 - \frac{m^2}{4}} & m < 2m_{K^+} \end{cases},$$

where $g_{\pi^+\pi^-}$ and $g_{K^+K^-}$ are the couplings to the $\pi^+\pi^-$ and K^+K^- channels. The $f_0(980)$ Flatté couplings are taken from Ref. [167] to be $g_{\pi^+\pi^-} = 199 \pm 30$ MeV and $g_{K^+K^-} = (3.0 \pm 0.3) \times g_{\pi^+\pi^-}$.

6.6 Acceptance

The fit function used in this analysis accounts for the acceptance due to detecting, triggering and selecting events as a function of the decay kinematics. This is done by multiplying the total decay rate by a four-dimensional acceptance function. It is defined in terms of Legendre polynomial (P_l) and spherical harmonic (Y_l^m)

Table 6.5 *Values of the Legendre moment coefficients found using the $B_s^0 \to \phi K^+K^-$ phase-space simulation sample split by different trigger conditions.*

c^{ijkl}	TOS	not-TOS
c^{0000}	0.0642 ± 0.0003	0.0670 ± 0.0003
c^{0002}	—	-0.0035 ± 0.0006
c^{0200}	—	-0.0292 ± 0.0012
c^{0400}	-0.0135 ± 0.0015	—
c^{1000}	—	-0.0176 ± 0.0012
c^{1200}	-0.0134 ± 0.0026	-0.0164 ± 0.0025

functions of $m(K^+K^-)$ and the helicity angles as

$$\varepsilon\left(m, \theta_1, \theta_2, \Phi\right) = \sum_{i,j,k,l} c^{ijkl} P_i(\cos\theta_2) Y_j^k(\cos\theta_1, \Phi) P_l(\mu), \tag{6.11}$$

where μ is $m(K^+K^-)$ mapped onto the interval $[-1, 1]$. The coefficients c^{ijkl} are calculated from the fully selected $B_s^0 \to \phi K^+K^-$ phase-space simulation sample as the following weighted average:

$$c^{ijkl} = \frac{1}{N} \sum_{n=0}^{N} \frac{1}{p_n q_n} \frac{2i+1}{2} \frac{2l+1}{2} P_i(\cos\theta_2)_n Y_j^k(\cos\theta_1, \Phi)_n P_l(\mu)_n,$$

where n is the index of each event, and N is the number of events in the sample. In order to minimise the effect of fluctuations in the simulation sample, only coefficients with values greater than four times their statistical error are kept.

From studies of other modes (*e.g.* Ref. [47]), it is known that the shape of the acceptance is different for TIS and TOS trigger categories, as defined in Section 3.6. Therefore, the acceptance function coefficients are calculated separately for the TOS and not-TOS samples. This splitting, rather than TIS and not-TIS, is chosen as it gives a more even distribution of events between the two categories. As a cross-check, the effect of splitting by TIS and not-TIS events is also investigated in Appendix D.

A condition is imposed that the acceptance function must be zero at exactly $m(K^+K^-) = 2m(K^+)$ and rise smoothly such that the acceptance is described solely by the function given in Equation 6.11 once sufficiently far from threshold. This is achieved by multiplying the acceptance function by a function of the form

$$\varepsilon_t\left(m\right) = \mathrm{erf}[A(m - 2m(K^+))], \tag{6.12}$$

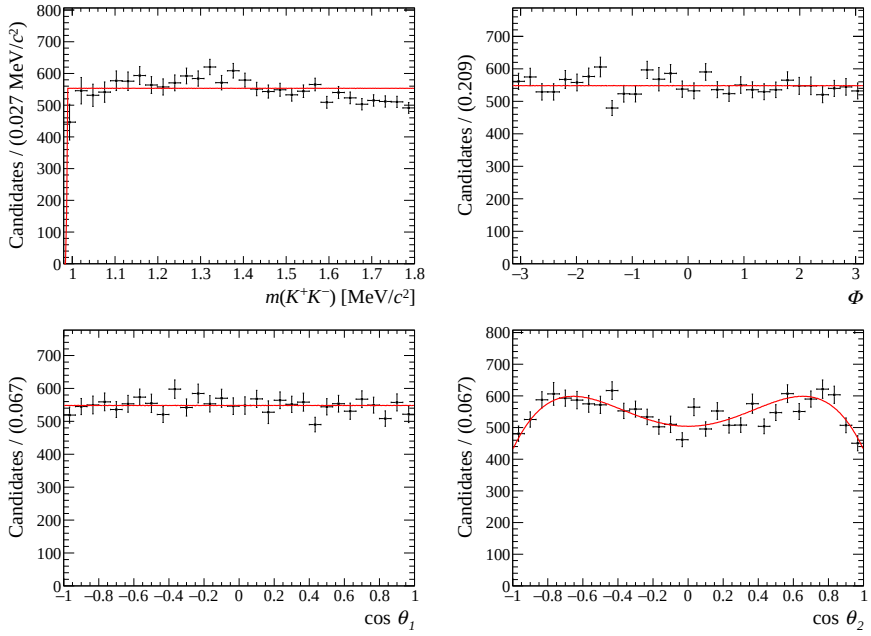


Figure 6.18 *1D projections of the acceptance function (red line) calculated for TOS events, overlaid onto the $B_s^0 \to \phi K^+ K^-$ simulation sample (black points).*

where 'erf' is the error function, and the parameter A is found from fitting the $B_s^0 \to \phi K^+ K^-$ simulation sample.

The calculated values of the coefficients are summarised in Table 6.5. The 1D projections of the acceptance function are shown in Figure 6.18 for TOS events and Figure 6.19 for not-TOS events. The 2D projections of the acceptance function can be found in Appendix E.

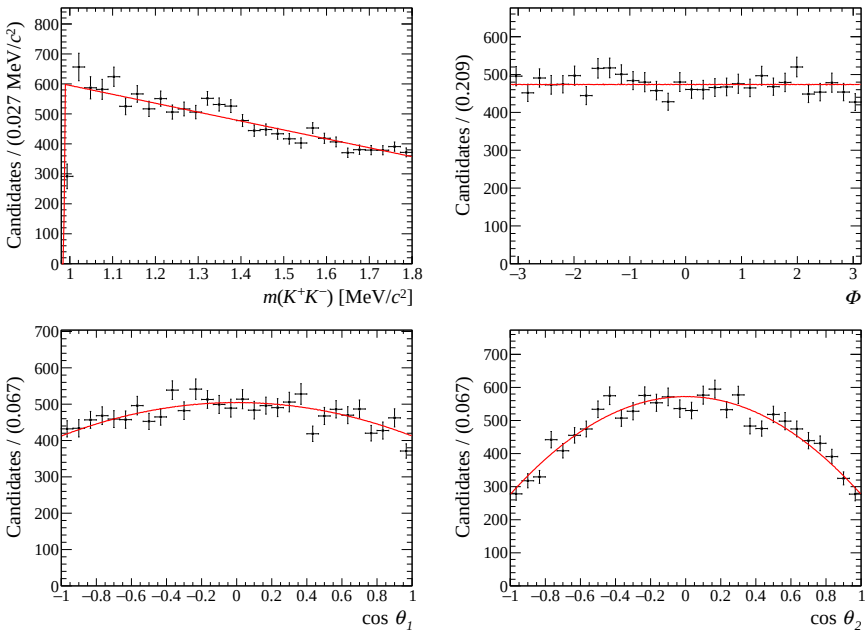


Figure 6.19 *1D projections of the acceptance function (red line) calculated for not-TOS events, overlaid onto the $B_s^0 \to \phi K^+ K^-$ simulation sample (black points).*

6.7 Resolution on $m(K^+K^-)$

When fitting the lineshapes of resonances with natural widths much larger than the mass resolution, it is sufficient to fit for the width, ignoring resolution effects. Conversely, when fitting to resonances that are much narrower than the mass resolution, it is often sufficient to ignore the natural width and fit a Gaussian function with a width equal to the resolution. In the case of resonances whose widths are of a similar order of magnitude to the mass resolution, as is the case with $\phi \to K^+K^-$ at LHCb, then both the natural width and mass resolution must be accounted for. This is done by convolving the lineshape with a Gaussian function.

The resolution is calculated using the simulation by taking the difference between the generated and simulated values of the K^+K^- invariant mass, $\Delta m(K^+K^-)$, for each event in the $B_s^0 \to \phi K^+ K^-$ simulation sample. To account for the change of the resolution with $m(K^+K^-)$, it is assumed that resolution scales with the square root of the energy release of the decay. Hence, the function

$$\sigma(m) = \sqrt{\sigma_0(m - 2m(K^+))} \tag{6.13}$$

is fitted to the distribution of $\Delta m(K^+K^-)$ plotted in nine bins of $m(K^+K^-)$. This is shown in Figure 6.20. The fitted value of σ_0 is found to be

$$\sigma_0 = (1.61 \pm 0.05) \times 10^{-5} \text{ GeV}/c^2.$$

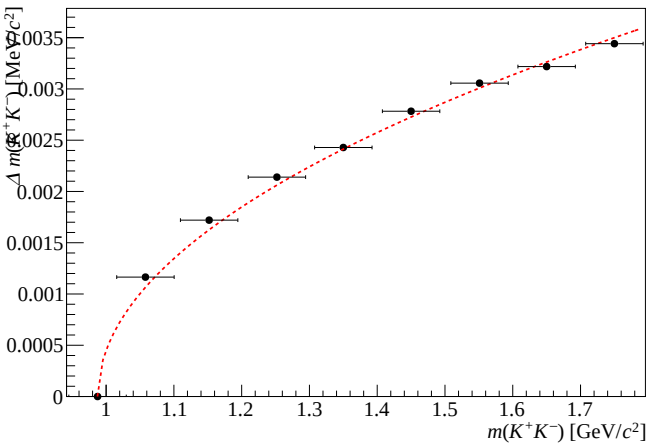


Figure 6.20 *The mass-dependence of the $m(K^+K^-)$ resolution as calculated using $B_s^0 \to \phi K^+K^-$ simulation. The red dotted line shows a fit using the function given in Equation 6.13. The resolution is forced to be zero at threshold.*

The convolution can be expressed as the following integral

$$\int_{-\infty}^{+\infty} P_{\text{G}}(m - x|0, \sigma(m))\Gamma(m, \theta_1, \theta_2, \Phi)dx,$$

where $P_{\text{G}}(m - x|0, \sigma(m))$ is a Gaussian function with a mean of zero and a mass-dependent width given in Equation 6.13, and $\Gamma(m, \theta_1, \theta_2, \Phi)$ is the time-integrated decay rate given in Equation 2.33. In the amplitude fit, this integral is done numerically with 20 samples over a range $[-3\sigma(m), +3\sigma(m)]$. The range of the integral and number of samples are chosen as a compromise between precision and computing time.

6.8 Choice of resonances

This section describes the choices of signal model with different combinations of K^+K^- resonances. Table 6.6 lists seventeen possible resonant contributions to the K^+K^- invariant mass spectrum. Since the valence quark content of ρ and a mesons has no $s\bar{s}$ component, and the decay $B_s^0 \to \phi K^+K^-$ is a $b \to s\bar{s}s$ transition,

these resonances may be safely ignored. This leaves eleven remaining $f^{(\prime)}$ and ϕ resonances, of which three spin-2 states — the $f_2(1565)$, $f_2(1640)$ and $f_2(1810)$ — are omitted from the PDG meson summary table. The $f_2(1565)$ is predominantly observed in nucleon–antinucleon annihilation and requires confirmation in other channels. There is only one analysis which reports possibly seeing the decay $f_2(1640) \to K^+K^-$ [168]. The authors state that it is only tentatively included in the fit. The existence of the $f_2(1810)$ requires confirmation, but there is some evidence of a K^+K^- channel [169]. For these reasons, these resonances are not included in the amplitude model.

Further information about which resonances to expect to contribute to the total amplitude can be taken from LHCb amplitude analyses of the $B^0_s \to J/\psi\, K^+K^-$ [170] and $B^0_s \to \phi\pi^+\pi^-$ [137] decays.

Table 6.6 *A list of resonances below 1.85* GeV/c^2 *that have been observed to decay to* K^+K^-*. Branching fractions to* K^+K^- *are quoted where available from Ref. [135].*

Resonance	Mass [MeV/c^2]	Width [MeV]	$\mathcal{B}$ ($\to K^+K^-$)
$f_0(980)$	990 ± 20	10 to 100	seen
$a_0(980)$	980 ± 20	92 ± 8	seen
ϕ (1020)	1019.461 ± 0.019	4.266 ± 0.031	$(4.89 \pm 0.05) \times 10^{-1}$
$f_2(1270)$	1275.1 ± 1.2	$185.1^{+2.9}_{-2.4}$	$(2.3 \pm 0.2) \times 10^{-2}$
$a_2(1320)$	$1318.3^{+0.5}_{-0.6}$	107 ± 5	$(2.45 \pm 0.4) \times 10^{-2}$
$f_0(1370)$	1475 ± 6	113 ± 11	seen
$a_0(1450)$	1474 ± 19	265 ± 13	seen
$f_0(1500)$	1505 ± 6	109 ± 7	$(4.3 \pm 0.5) \times 10^{-2}$
$f_2'(1525)$	1525 ± 5	73^{+6}_{-5}	$(4.44 \pm 0.11) \times 10^{-1}$
$f_2(1565)$	1562 ± 13	134 ± 8	
$f_2(1640)$	1639 ± 6	99^{+60}_{-40}	seen
ϕ (1680)	1680 ± 20	159 ± 50	seen
$\rho_3(1690)$	1696 ± 4	161 ± 10	$(7.9 \pm 1.3) \times 10^{-3}$
ρ (1700)	1720 ± 20	250 ± 100	seen
$a_2(1700)$	1732 ± 16	194 ± 40	seen
$f_0(1710)$	1723^{+6}_{-5}	139 ± 8	seen
$f_2(1810)$	1815 ± 12	197 ± 22	

The $B^0_s \to \phi\pi^+\pi^-$ analysis [137] observed and measured the branching fractions of the decays $B^0_s \to \phi f_0(980)$ and $B^0_s \to \phi f_2(1270)$ using the $B^0_s \to \phi\phi$ branching fraction result from Chapter 5 for normalisation. With knowledge of the $\pi^+\pi^-$ and K^+K^- channels of the $f_0(980)$ and $f_2(1270)$, it is possible to estimate the contribution of these intermediate resonances to the total amplitude.

In Ref. [137], a measurement is made of the ratio

$$\frac{\mathcal{B}(B_s^0 \to \phi f_0(980))\mathcal{B}(f_0(980) \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to \phi\phi)} = 0.068 \pm 0.008\,(\text{stat}) \pm 0.007\,(\text{syst}).$$

The ratio of partial widths of the $f_0(980)$ to $\pi\pi$ and $K\overline{K}$ final states [1] has been measured by the BESII collaboration to be [171]:

$$\frac{\Gamma(\pi\overline{\pi})}{\Gamma(\pi\overline{\pi}) + \Gamma(K\overline{K})} = 0.75^{+0.11}_{-0.13}.$$

From this, the expected yield of the $B_s^0 \to \phi f_0(980)$ component in the $B_s^0 \to \phi K^+K^-$ amplitude fit, relative to the $B_s^0 \to \phi\phi$ component, can be calculated using:

$$\frac{N_{B_s^0 \to \phi f_0(980)}}{N_{B_s^0 \to \phi\phi}} = \frac{\mathcal{B}(B_s^0 \to \phi f_0(980))\mathcal{B}(f_0(980) \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to \phi\phi)\mathcal{B}(\phi \to K^+K^-)} \left(\frac{\Gamma(\pi\overline{\pi}) + \Gamma(K\overline{K})}{\Gamma(\pi\overline{\pi})} - 1 \right). \tag{6.14}$$

This gives

$$\frac{N_{B_s^0 \to \phi f_0(980)}}{N_{B_s^0 \to \phi\phi}} = 0.05 \pm 0.03.$$

Assuming $\sim$ 3000 $B_s^0 \to \phi\phi$ events in the sample, this translates to $\sim$ 150 $B_s^0 \to \phi f_0(980)$ events. A $B_s^0 \to \phi f_0(980)$ component is therefore included in the fit to data.

Also in Ref. [137], a measurement of the ratio

$$\frac{\mathcal{B}(B_s^0 \to \phi f_2(1270))\mathcal{B}(f_2(1270) \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to \phi\phi)} = 0.033 \pm 0.005\,(\text{stat}) \pm 0.003\,(\text{syst})$$

is given. The branching fractions of the $f_2(1270)$ to the $\pi\pi$ and $K\overline{K}$ final states are given by the PDG to be [135]:

$$\begin{aligned} \Gamma(\pi\pi) &= \left(84.8^{+2.4}_{-1.2}\right)\%, \\ \Gamma(K\overline{K}) &= (4.6 \pm 0.4)\,\%. \end{aligned}$$

The expected yield of the $B_s^0 \to \phi f_2(1270)$ component in the $B_s^0 \to \phi K^+K^-$ amplitude fit, relative to the $B_s^0 \to \phi\phi$ component, can therefore be calculated using:

$$\frac{N_{B_s^0 \to \phi f_2(1270)}}{N_{B_s^0 \to \phi\phi}} = \frac{\mathcal{B}(B_s^0 \to \phi f_2(1270))\mathcal{B}(f_2(1270) \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to \phi\phi)\mathcal{B}(\phi \to K^+K^-)} \frac{\Gamma(K\overline{K})}{\Gamma(\pi\pi)}. \tag{6.15}$$

[1] This includes pairs of charged and neutral mesons.

The result is

$$\frac{N_{B_s^0\to\phi f_2(1270)}}{N_{B_s^0\to\phi\phi}} = 0.0037 \pm 0.007,$$

which, assuming ~ 3000 $B_s^0 \to \phi\phi$ events, converts to a yield of ~ 11. Therefore this component can be neglected.

The amplitude fit in Ref. [137] includes a high-mass component, but there is ambiguity as to whether it is the $f_0(1370)$, $f_0(1500)$ or a mixture of the two states. The $f_0(1370)$ state is not well known: the PDG quotes a range of [1.2,1.5] GeV/c^2 for the pole mass and [200,500] MeV/c^2 for the width. Measurements of the ratio of partial widths $\Gamma(K\overline{K})/\Gamma(\pi\overline{\pi})$ vary from 8 to 91 %. By contrast, the $f_0(1500)$ state is much better established. Assuming this high-mass component is entirely the $f_0(1500)$, then the fit fraction of this component is equivalent to

$$F_{B_s^0\to\phi f_0(1500)} \equiv \frac{\mathcal{B}(B_s^0 \to \phi f_0(1500))\mathcal{B}(f_0(1500) \to \pi^+\pi^-)}{\mathcal{B}(B_s^0 \to \phi\pi^+\pi^-)} = 34.7 \pm 0.34\,\%.$$

The PDG world average for the ratio of partial widths for this state is [135]

$$\frac{\Gamma(K\overline{K})}{\Gamma(\pi\overline{\pi})} = 0.241 \pm 0.028.$$

The expected yield relative to the $B_s^0 \to \phi\phi$ component is calculated using:

$$\frac{N_{B_s^0\to\phi f_0(1500)}}{N_{B_s^0\to\phi\phi}} = \frac{F_{B_s^0\to\phi f_0(1500)}\mathcal{B}(B_s^0 \to \phi\pi^+\pi^-)}{\mathcal{B}(B_s^0 \to \phi\phi)\mathcal{B}(\phi \to K^+K^-)}\frac{\Gamma(K\overline{K})}{\Gamma(\pi\overline{\pi})}. \tag{6.16}$$

This gives

$$\frac{N_{B_s^0\to\phi f_0(1500)}}{N_{B_s^0\to\phi\phi}} = 0.032 \pm 0.006.$$

This is large enough to expect ~ 100 events in the sample, hence this component needs to be considered in the fit. However, due to the ambiguity over the identity of the high-mass component in the $B_s^0 \to \phi\pi^+\pi^-$ analysis, this cannot reliably be used to predict the number of $B_s^0 \to \phi f_0(1500)$ events.

In the amplitude analysis of the decay $B_s^0 \to J/\psi\, K^+K^-$ in Ref. [170], the best fit model contains the $f_0(980)$, the $\phi(1020)$, the $f_0(1370)$, the $f_2'(1525)$, the $f_2(1640)$, the $\phi(1680)$ and the $f_2(1750)$. The $f_2(1640)$, $\phi(1680)$ and $f_2(1750)$ components are reported with fit fractions above 1 % and are important in describing the shape of the distribution above the mass of the $f_2'(1525)$. Therefore they are considered when choosing resonance models in this analysis. Although it does not appear in

the PDG, the existence of the $f_2(1750)$ is reported by Belle in Ref. [172] and is therefore considered.

This fit assumes that the 'first' K^+K^- pair — in the narrow invariant mass window — is always from a ϕ meson. Decays such as $B_s^0 \to f_0\phi$, $B_s^0 \to f_0 f_0$, nonresonant $B_s^0 \to K^+K^-\phi$ or even nonresonant $B_s^0 \to K^+K^-K^+K^-$ are neglected.

6.9 Signal model

Based on the arguments and calculations presented in Section 6.8, the K^+K^- resonances considered for the 'minimal' fit in this analysis are $f_0(980)$, $\phi(1020)$ and the $f_2'(1525)$, plus a nonresonant component to account for true three-body $B_s^0 \to \phi K^+K^-$ decays. Additional models containing components for the $f_0(1370)$, $f_0(1500)$, $f_2(1640)$, $\phi(1680)$, $f_0(1710)$ or $f_2(1750)$ states are also considered. As discussed in Section 6.5.1, the resonances are described by relativistic Breit–Wigner distributions, with the exception of the $f_0(980)$, which is described by a Flatté shape.

The shape of each Breit–Wigner resonance is controlled by the pole mass m and natural width Γ. The Flatté parameters are the coupling $g_{\pi\pi}$ and the ratio of couplings $R_g = g_{K^+K^-}/g_{\pi^+\pi^-}$. Each component in the signal model has an overall scaling factor f, which is fixed to unity for the $B_s^0 \to \phi\phi$ component. The nonresonant and spin-0 resonant amplitudes have a phase δ_0. The resonant amplitudes with spin > 0 have three complex helicity amplitudes: $\mathcal{A}_0$, $\mathcal{A}_+$ and $\mathcal{A}_-$, controlled by the fit parameters $|\mathcal{A}_0|$, $|\mathcal{A}_+|$, δ_0, δ_+ and δ_-. The magnitude $|\mathcal{A}_-|$ is calculated using the unitarity condition $\sum_\lambda |\mathcal{A}_\lambda|^2 = 1$

The natural width of the $\phi(1020)$ meson and the pole mass and natural width of the $f_2'(1525)$ meson are constrained to the PDG values, given in Table 6.6. Gaussian constraints are applied to the parameters. Asymmetric uncertainties are handled by using the average of the upper and lower uncertainty as the width of the constraint function. The pole mass of the $\phi(1020)$ meson is left free in the fit. The $f_0(980)$ Flatté couplings are taken from Ref. [167] to be $g_{\pi^+\pi^-} = 199 \pm 30$ MeV and $g_{K^+K^-} = (3.0 \pm 0.3) \times g_{\pi^+\pi^-}$. The values of the couplings are later varied to quantify a systematic uncertainty. The pole mass and natural width of the $f_0(1500)$ are fixed to the LHCb measurement in Ref. [173]. The parameters of the $f_2(1750)$ are taken from the Belle measurement in Ref. [172]. The Breit–Wigner

parameters of all other resonances considered in the fit are fixed to the central values given in Table 6.6.

The time-integrated angular distribution of a $P \to VV$ or $P \to VT$ decay, assuming small *CP* violation, is sensitive only to one difference in phase between two of the helicity amplitudes. Additionally, interference between resonances leads to sensitivity in the PDF to the difference in phase between the resonant components for a particular polarisation.[2] Since the PDF depends on the differences between phases, at least one phase in the total amplitude must be fixed. The phase of $\mathcal{A}_0$ for the $B_s^0 \to \phi\phi$ component is fixed to zero following the convention established by previous analyses of this decay, which allows for easier comparison of results. The time-integrated angular distribution of the $B_s^0 \to \phi\phi$ decay, assuming equal numbers of B_s^0 and $\overline{B}_s^0$, is sensitive to only the difference between the phases of the transverse helicity amplitudes and not their absolute values. Since the magnitude of $\mathcal{A}_-$ is known to be the smaller of the two, the phase δ_- of the $B_s^0 \to \phi\phi$ component is fixed to the value found in the LHCb angular analysis of $B_s^0 \to \phi\phi$ in Ref. [49], to simplify comparison. The phases of all other resonant components are left free, since they can be measured relative to the $B_s^0 \to \phi\phi$ component.

6.10 Background model

The background model used in the amplitude fit includes the three sources of background included in the $m(K^+K^-K^+K^-)$ fit. The relative fraction of each component is fixed to the values found in Section 6.4. The $m(K^+K^-)$ distribution of the combinatorial background is modelled a histogram of the upper sideband of the data sample, with $m(\phi K^+K^-) > 5.5$ GeV/c^2. The $m(K^+K^-)$ distribution of the $B^0 \to \phi K^{*0}$ background is modelled using a histogram from the fully selected simulation sample. Although the resonant structure of $\Lambda_b^0 \to \phi p K^-$ is unknown, this background is assumed to be dominated by $\Lambda_b^0 \to \phi\Lambda(1520)$ decays. It is therefore modelled using the $m(K^+K^-)$ distribution of misidentified $\Lambda_b^0 \to \phi\Lambda(1520)$ generator-level events with momentum-smearing applied. The $m(K^+K^-)$ distributions of the background components are shown in Figure 6.21.

The angular distribution of the entire background is expressed as a sum of spherical

[2]There is no interference between amplitudes with different polarisations.

harmonics, Y, and associated Legendre polynomials, P, given in Equation 6.17:

$$B(\theta_1, \theta_2, \Phi) = \sum_{i,j,k} c^{ijk} Y_j^k(\cos\theta_1, \Phi) P_i(\cos\theta_2), \tag{6.17}$$

where the coefficients of each term in the sum, c^{ijk}, are determined using the following sum over events in the sideband sample:

$$c^{ijk} = \frac{1}{N} \sum_{n=0}^{N} \frac{2i+1}{2} Y_j^k(\cos\theta_{1n}, \Phi_n) P_i(\cos\theta_{2n}), \tag{6.18}$$

where N is the number of events in the sample. The indices i, j and k are required to be below 3, and j must be greater than k. Only coefficients larger than 5σ are kept, where σ is calculated as:

$$\sigma = \sqrt{\frac{1}{N^2}\left((c^{ijk})^2 - \frac{(c^{ijk})^2}{N}\right)}. \tag{6.19}$$

The accepted coefficients are c^{000}, c^{002} and c^{200}; their values are summarised in Table 6.7

Table 6.7 *Values of the accepted background function coefficients.*

Coefficient	Value
c^{000}	$(7.05 \pm 0.00) \times 10^{-2}$
c^{002}	$(1.91 \pm 0.32) \times 10^{-2}$
c^{200}	$(3.71 \pm 0.68) \times 10^{-2}$

The one-dimensional projections of the background from the collision data sideband sample and the fitted background function for each of the helicity angles are shown in Figure 6.22. The two-dimensional projections of the background function are shown in Figure 6.23.

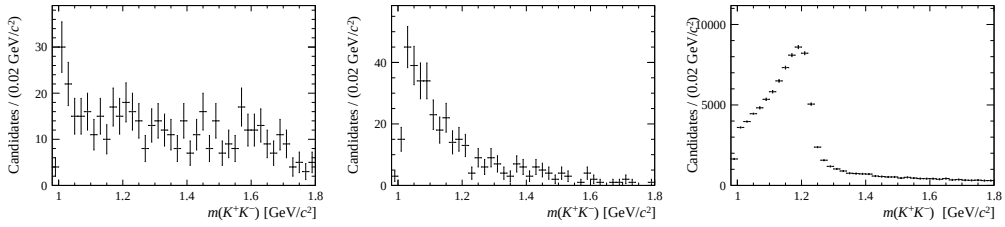


Figure 6.21 *Histograms used for the $m(K^+K^-)$-dependent part of the combinatorial (left), $B^0 \to \phi K^{*0}$ (centre) and $\Lambda_b^0 \to \phi p K^-$ (right) backgrounds.*

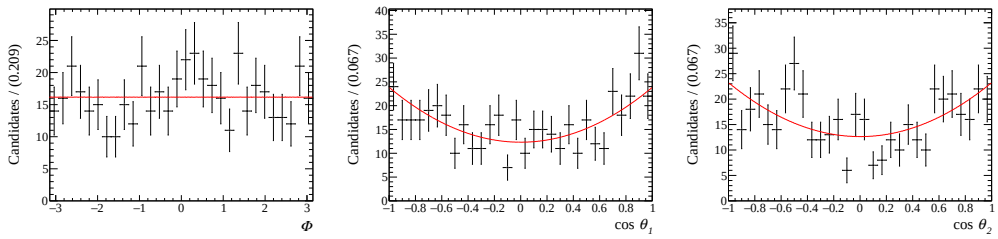


Figure 6.22 *1D projections of the angular part of the background. The black points are sideband data events, and the red line is the fitted background function.*

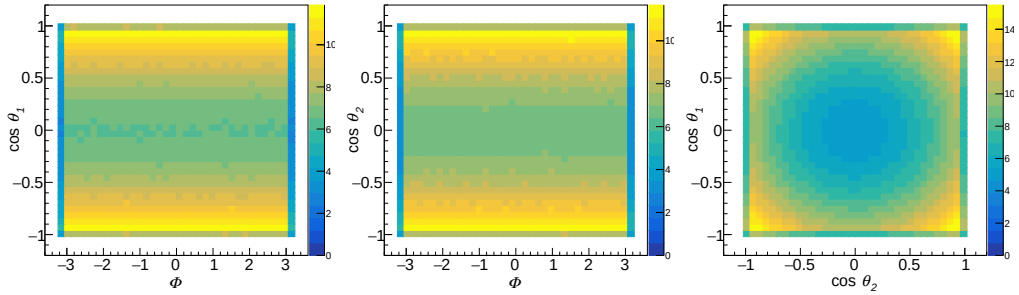


Figure 6.23 *2D projections of the angular part of the fitted background function.*

6.11 Fit model validation

The fit model is validated in three steps. First, samples of toy events are generated from the signal PDF using reasonable starting values, which are then fitted in order to check for consistency and biases. Second, fits are performed to generator-level simulation samples of $B_s^0 \to \phi\phi$ and nonresonant $B_s^0 \to \phi K^+ K^-$ decays to validate the matrix element calculation. Third, fits are performed to the fully simulated and fully selected simulation samples of the same decays, in order to validate the implementation of the mass resolution and acceptance function in the fitter.

6.11.1 Toy studies

An important step in the validation is to fit a sample of toy events generated from the PDF itself and check that the results agree with the input values. If the results are not consistent with the starting values, it indicates that there are systematic biases, lack of sensitivity to that particular fit parameter, or even mistakes in the code. The number of toys generated in each study is 4000. This is roughly the same size as the data sample, allowing for comparable statistical uncertainties. For each toy study, approximately 100 samples are generated and fitted, and a

distribution of 'pulls' is obtained for each free parameter. The pulls are calculated as

$$\text{pull} = \frac{V_{\text{obs}} - V_{\text{exp}}}{\sigma V_{\text{obs}}},$$

where V_{obs} is the observed (fitted) value, V_{exp} is the expected (generated) value, and σV_{obs} is the uncertainty on V_{obs} as reported by the fitter. The pull distribution should have a mean $\mu = 0$ and standard deviation $\sigma = 1$. Deviations from $\mu = 0$ indicate biases in the fit, and deviations from $\sigma = 1$ indicate that the incorrect uncertainty is reported by the fitter. In the case of the latter, uncertainties on fitted parameters should be calculated by running toy studies and taking the standard deviation on the distributions of V_{obs}.

Toy events are generated from the signal PDF without acceptance or resolution effects for the exclusive processes $B_s^0 \to \phi f_0(980)$, $B_s^0 \to \phi f_0(1500)$, $B_s^0 \to \phi\phi$ and $B_s^0 \to \phi f_2'(1525)$ in order to validate the Flatté and spin-0, 1 and 2 Breit–Wigner lineshapes. For the $B_s^0 \to \phi f_0(980)$ toys, the Flatté coupling parameters $g_{\pi\pi}$ and $R_g = \frac{g_{KK}}{g_{\pi\pi}}$ were floated, and the pull distributions had means and standard deviations of 0.11 ± 0.89 and 0.02 ± 1.06, respectively. The results of the toy studies using the other processes are shown in Table 6.8. The standard deviations on the pull distributions of $|\mathcal{A}_+|$ and δ_+ of the $B_s^0 \to \phi f_2'(1525)$ study indicate that the uncertainties on these parameters are over-estimated by the fitter by a factor of two.

Table 6.8 *Results of toy studies without detector effects. For each parameter, the mean, μ, and standard deviation, σ, of the pull distribution is shown.*

Parameter	$\mu \pm \sigma$ of pulls		
	$B_s^0 \to \phi f_0(1500)$	$B_s^0 \to \phi\phi$	$B_s^0 \to \phi f_2'(1525)$
m	-0.04 ± 0.95	-0.01 ± 1.02	0.17 ± 0.94
Γ	0.03 ± 0.87	0.13 ± 0.98	-0.19 ± 0.97
$\lvert\mathcal{A}_0\rvert$	–	0.05 ± 0.93	0.07 ± 0.96
$\lvert\mathcal{A}_+\rvert$	–	-0.08 ± 1.11	-0.01 ± 0.52
δ_+	–	0.03 ± 1.06	-0.09 ± 0.52

To test that the detector effects in the PDF, namely the acceptance and $m(K^+K^-)$ resolution, toy studies are performed using the $B_s^0 \to \phi\phi$ PDF with each effect included individually and together. The results of this are shown in Table 6.9.

Table 6.9 *Results of toy studies using the $B_s^0 \to \phi\phi$ PDF with detector effects. For each parameter, the mean, μ, and standard deviation, σ, of the pull distribution is shown.*

Parameter	$\mu \pm \sigma$ of pulls		
	Acceptance	Resolution	Both
m	0.06 ± 1.02	0.03 ± 1.01	0.02 ± 1.03
Γ	0.02 ± 0.93	0.02 ± 0.85	-0.14 ± 1.05
$\lvert\mathcal{A}_0\rvert$	-0.04 ± 0.84	0.02 ± 1.04	0.05 ± 0.94
$\lvert\mathcal{A}_+\rvert$	0.09 ± 0.86	-0.16 ± 0.94	-0.12 ± 0.86
δ_+	0.07 ± 0.88	-0.07 ± 0.90	-0.06 ± 0.96

6.11.2 Fits to simulation

The generator-level simulation samples are not subject to the simulated LHCb detector, hence they can be used to validate the amplitude calculation independently from the resolution and acceptance effects. The $B_s^0 \to \phi\phi$ simulation sample is generated using the `PVV_CPLH` model in EvtGen [122] with the parameters given in Table 6.10. The helicity amplitude parameters come from the CDF result [48].

Table 6.10 *Physics parameters used in generating the $B_s^0 \to \phi\phi$ simulation sample. The symbol τ denotes lifetime, L and H denote the light and heavy B_s^0 mass eigenstates, and $r_{B_s^0}$ and $r_{K^+K^-}$ are Blatt–Weisskopf barrier factor radii.*

Parameter	Value
$\lvert\mathcal{A}_\parallel\rvert$	0.536
$\lvert\mathcal{A}_0\rvert$	0.590
$\lvert\mathcal{A}_\perp\rvert$	0.604
$\delta_\parallel$	2.71
δ_0	0.0
$\delta_\perp$	2.39
m_ϕ	1.019461 GeV/c^2
τ_ϕ	1.5451×10^{-22} s
τ_L	1.405 ps
τ_H	1.661 ps
$r_{B_s^0}$	1.0 GeV$^{-1}\,c$
$r_{K^+K^-}$	3.0 GeV$^{-1}\,c$

The signal PDF is, in principle, sensitive to two of the magnitudes of the helicity amplitudes, one of the phases, the mass and width of the ϕ, the width splitting of the B_s^0 mass eigenstates, $\Delta\Gamma_s = \Gamma_L - \Gamma_H$, and the barrier factor radii. A fit is made to a sample of generator-level $B_s^0 \to \phi\phi$ events with these parameters free in

the fit, with the exception of $\Delta\Gamma_s$ and the B_s^0 barrier factor radius. The results are shown in Table 6.11.

Table 6.11 *Results of the fit to the $B_s^0 \to \phi\phi$ generator-level simulation*

Parameter	Fit result	σ from input
m	1019.448 ± 0.004 MeV/c^2	-2.9
Γ	4.269 ± 0.009 MeV	0.33
$\lvert\mathcal{A}_0\rvert$	0.7924 ± 0.0007	-2.2
$\lvert\mathcal{A}_+\rvert$	0.5899 ± 0.0007	-0.13
δ_+	2.532 ± 0.004	-1.8
$r_{K^+K^-}$	$3.08 \pm 0.08\,\text{GeV}^{-1}\,c$	1.1

The difference between the generator-level and fully simulated and fully selected simulation samples is due to acceptance and resolution. To check how well the PDF models these effects, a fit is performed to the fully selected $B_s^0 \to \phi\phi$ simulation sample with the $\phi(1020)$ mass, width and helicity amplitude parameters left free. The results are shown in Table 6.12.

Table 6.12 *Results of the fit to the full $B_s^0 \to \phi\phi$ simulation*

Parameter	Fit result	σ from input
m	1019.42 ± 0.02 MeV/c^2	-2.2
Γ	4.3 ± 0.3 MeV	1.4
$\lvert\mathcal{A}_0\rvert$	0.7917 ± 0.0027	-1.6
$\lvert\mathcal{A}_+\rvert$	0.5904 ± 0.0027	0.13
δ_+	2.508 ± 0.016	-2.1

The deviations on the parameters m, $|\mathcal{A}_0|$ and δ_+ suggest that there are biases in the fit model which were not evident from Section 6.11.1. These could be due to *e.g.* a difference in formalism between EvtGen and the PDF. This is an issue that must be resolved before the analysis can be published.

6.12 Fits to data

Several amplitude fits are performed to the data in the invariant mass window $|m(K^+K^-K^+K^-) - m(B_s^0)| < 2\sigma_{\text{eff}}$, where σ_{eff} is defined in Section 6.4. The background model is described in Section 6.10. The signal is modelled using the amplitude formalism described in Section 6.5 with the choice of resonances built using an iterative method. If a fit with an additional spin-1 or 2 resonance fails to converge, results in amplitudes close to zero or has a non-positive-definite

covariance matrix, then it is repeated with a reduced number of free parameters by fixing the magnitudes and/or phases of the helicity amplitudes.

The figure of merit used to evaluate fit performance is the Bayesian information criterion (BIC) [174], which penalises the maximised likelihood, $\mathcal{L}$, using the number of free parameters, p, and the sample size, n: $\text{BIC} = p\ln(n) - 2\ln(\mathcal{L})$. A smaller value of the BIC indicates a more suitable model. For comparison, a similar quantity, called the Akaike information criterion ($\text{AIC} = 2p - 2\ln(\mathcal{L})$), as well as the negative log-likelihood ($-\ln(\mathcal{L})$) were calculated for each fit and found to give similar results.

The 'minimal' model from Section 6.8, containing components for the $f_0(980)$, $\phi(1020)$, $f_2'(1525)$ and nonresonant K^+K^- pairs, is first fitted to the data. This fit is then compared with models that include an extra $f_0(1370)$, $f_0(1500)$, $f_2(1640)$, $\phi(1680)$, $f_0(1710)$ or $f_2(1750)$ resonance to determine which, if any, best describes the data. The model including a $\phi(1680)$ component is the only one to decrease the BIC, by 48 units. The decrease in negative log-likelihood compared to the minimal model is 46 units, indicating that this component is required.

Next, fits are performed using the model from the previous step with the smallest BIC (minimal + $\phi(1680)$) with an additional $f_0(1370)$, $f_0(1500)$, $f_2(1640)$, $f_0(1710)$ or $f_2(1750)$ component. The best model was the one including the $f_0(1500)$, which decreased the BIC by 7.5 and the negative log-likelihood by 12 units. The second best model from this step was the inclusion of the $f_0(1710)$, which increased the BIC by 0.1 and decreased the negative log-likelihood by 8.2 units.

Finally, the addition of a $f_2(1640)$, $f_0(1710)$ or $f_2(1750)$ were tried. None of these were found to decrease the value of the BIC. Therefore, the best fit model is chosen to be the one with components for the $f_0(980)$, $\phi(1020)$, $f_0(1500)$, $f_2'(1525)$, $\phi(1680)$ and nonresonant $B_s^0 \to \phi K^+K^-$ decays.

As systematic variations, the three models with the best values of BIC are considered. Hereafter, these are referred to as Models I, II and III. Model I contains the fewest components. Model II is formed by adding a $f_0(1500)$ component to Model I. Likewise, Model III is formed by adding a $f_0(1710)$ component to Model I. The fit results of the three best models are shown in Table 6.13, and their performance in terms of $-\ln(\mathcal{L})$, AIC and BIC are shown in Table 6.14. The projections of the best-fit model (Model II) are shown in Figure 6.24. The performance of all of the models tried in the first two steps, as well as the projections of Models I and III, are given in Appendix F.

Table 6.13 *Fit results using the three best fit models.*

Fit parameter	Result with statistical uncertainty		
	Model II	Model III	Model I
Breit–Wigner parameters			
$\phi(1020)$ m [MeV/c^2]	1019.52 ± 0.06	1019.52 ± 0.04	1019.52 ± 0.06
$\phi(1020)$ Γ [MeV]	4.27 ± 0.03	4.267 ± 0.021	4.267 ± 0.030
$f_2'(1525)$ m [MeV/c^2]	1522 ± 4	1517.7 ± 2.4	1518 ± 4
$f_2'(1525)$ Γ [MeV]	75 ± 5	75.8 ± 2.7	77 ± 5
Fit fractions			
$f_0(980)$	$(3.0 \pm 0.5)\%$	$(6.3 \pm 0.7)\%$	$(5.0 \pm 0.7)\%$
$\phi(1020)$	$(74.0 \pm 0.8)\%$	$(73.7 \pm 0.8)\%$	$(73.7 \pm 0.9)\%$
$f_0(1500)$	$(1.0 \pm 0.4)\%$	—	—
$f_2'(1525)$	$(6.0 \pm 0.6)\%$	$(7.0 \pm 0.7)\%$	$(6.9 \pm 0.6)\%$
$\phi(1680)$	$(4.1 \pm 0.9)\%$	$(3.9 \pm 0.7)\%$	$(3.5 \pm 0.9)\%$
$f_0(1710)$	—	$(2.6 \pm 0.7)\%$	—
nonresonant	$(8.9 \pm 1.1)\%$	$(7.4 \pm 1.2)\%$	$(9.1 \pm 1.0)\%$
interference	$(3.2 \pm 0.9)\%$	$(-0.6 \pm 0.9)\%$	$(1.9 \pm 1.0)\%$
Magnitudes			
$\phi(1020)$ $\lvert A_+ \rvert$	0.798 ± 0.009	0.798 ± 0.009	0.798 ± 0.009
$\phi(1020)$ $\lvert A_0 \rvert$	0.603 ± 0.009	0.603 ± 0.009	0.603 ± 0.009
$f_2'(1525)$ $\lvert A_+ \rvert$	0.30 ± 0.08	0.33 ± 0.03	0.28 ± 0.07
$f_2'(1525)$ $\lvert A_0 \rvert$	0.93 ± 0.03	0.915 ± 0.012	0.930 ± 0.026
$\phi(1680)$ $\lvert A_+ \rvert$	0.81 ± 0.06	0.87 ± 0.03	0.81 ± 0.06
$\phi(1680)$ $\lvert A_0 \rvert$	0.09 ± 0.13	0.16 ± 0.08	0.13 ± 0.15
Phases [rad]			
nonresonant δ_0	1.40 ± 0.10	1.40 ± 0.06	1.40 ± 0.10
$f_0(980)$ δ_0	-1.62 ± 0.23	-1.97 ± 0.07	-1.75 ± 0.19
$\phi(1020)$ δ_+	2.71 ± 0.07	2.71 ± 0.04	2.72 ± 0.07
$f_0(1500)$ δ_0	-0.2 ± 0.3	—	—
$f_0(1710)$ δ_0	—	2.52 ± 0.08	—
$f_2'(1525)$ δ_+	2.5 ± 0.4	2.15 ± 0.19	2.0 ± 0.3
$f_2'(1525)$ δ_0	-0.06 ± 0.26	-0.17 ± 0.09	-0.36 ± 0.22
$f_2'(1525)$ δ_-	1.9 ± 0.4	1.56 ± 0.20	1.5 ± 0.4
$\phi(1680)$ δ_+	-0.40 ± 0.24	-0.75 ± 0.09	-0.63 ± 0.23
$\phi(1680)$ δ_0	-2.0 ± 1.1	-2.5 ± 0.5	-2.4 ± 0.9
$\phi(1680)$ δ_-	2.97 ± 0.32	2.57 ± 0.17	2.6 ± 0.3

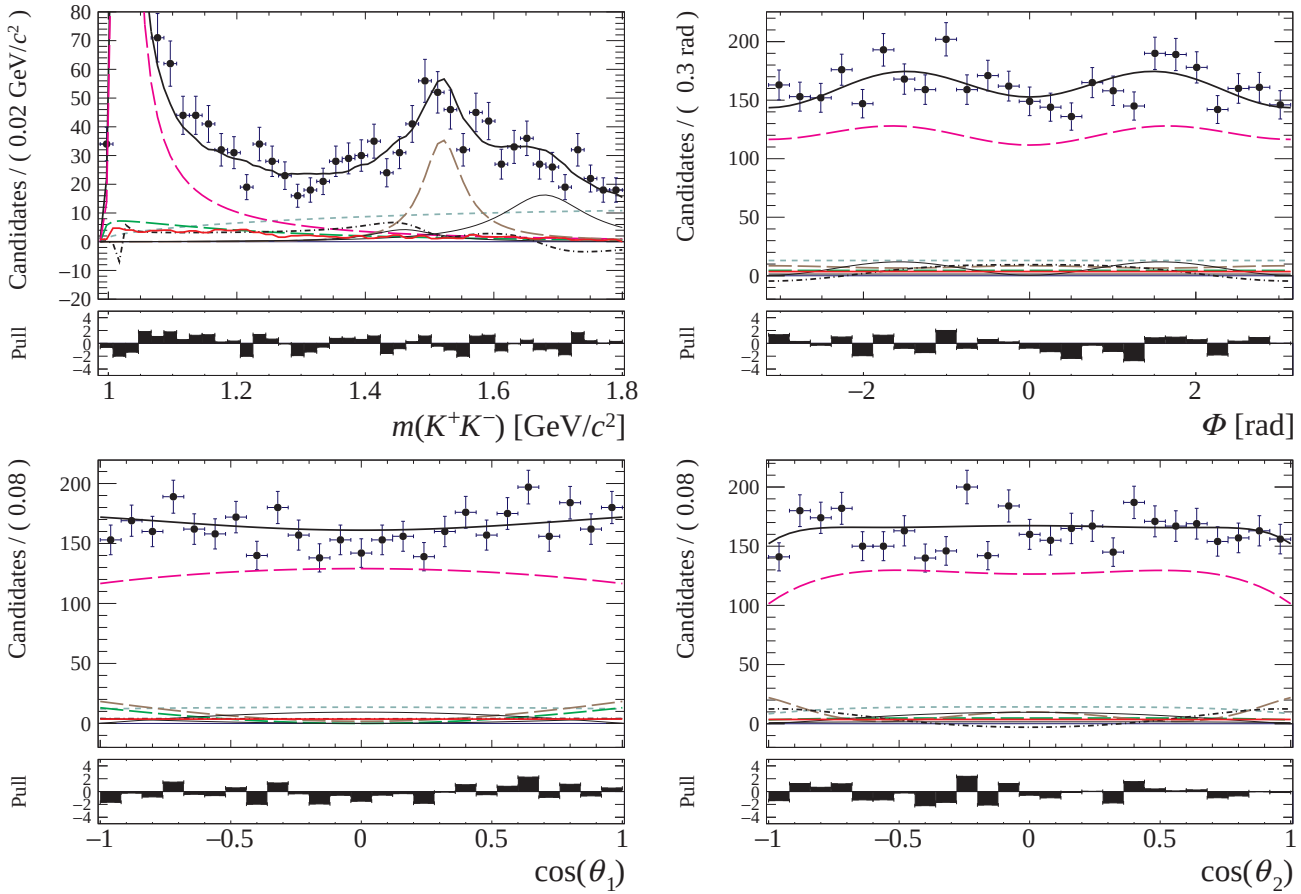


Figure 6.24 *Fit projections of Model II in $m(K^+K^-)$ and the helicity angles, overlaid onto the data (black points). The total PDF is a solid black line, the $\phi(1020)$ is a purple long-dashed line, the $f_0(980)$ is a green long-dashed line, the $f_2'(1525)$ is a brown long-dashed line, the nonresonant component is a light blue short-dashed line, the interference is a black dot-dashed line, the $f_0(1500)$ and $\phi(1680)$ are thin solid black lines, and the background is a red histogram.*

Table 6.14 *Comparison of the $-\ln(\mathcal{L})$, AIC and BIC between the three best models.*

Model	Resonance content	$-\ln(\mathcal{L})$	AIC	BIC
I	$f_0(980) + \phi(1020) + f_2'(1525) + \phi(1680)$	3836.7	7719.4	7864.6
II	Model I + $f_0(1500)$	3824.7	7699.4	7857.1
III	Model I + $f_0(1710)$	3828.5	7707.0	7864.7

Fit fractions are calculated from the signal PDF, without background components or detector effects, as the integral of one component divided by the integral of the total PDF. The sum of fit fractions can differ from unity due to interference. This difference is quoted in Table 6.13 as the 'interference' fit fraction. The statistical uncertainties on the fit fractions are obtained by generating and fitting ~ 100 random toy samples, each the same size as the dataset and with starting values set to the results of the fit to data, and taking the standard deviation of the results.

Table 6.15 *Change in negative log-likelihood caused by removing each component from Model II.*

Component removed	$-\ln(\mathcal{L})$	$\Delta\ln(\mathcal{L})$	n_σ
None	3824.7	—	—
Nonresonant K^+K^-	3827.9	3.2	2.1
$f_0(980)$	3855.9	31.2	7.6
$f_0(1500)$	3836.7	12.0	4.5
$f_2'(1525)$	3923.1	98.4	12.9
$\phi(1680)$	3879.1	54.4	9.4

The statistical significance of each component in Model II, with the exception of the $\phi(1020)$, is estimated from the change in $-\ln(\mathcal{L})$ caused by removing that component and repeating the fit to data. The results are shown in Table 6.15. Wilks' theorem [153] allows for an estimate of the statistical significance, expressed as an equivalent number of Gaussian standard deviations, n_σ, as

$$n_\sigma = \sqrt{2}\mathrm{erfc}^{-1}\left(\mathrm{prob}\left(-2\Delta\ln(\mathcal{L}), \Delta\mathrm{ndf}\right)\right), \tag{6.20}$$

where the 'erfc' is the cumulative error function, 'prob' calculates the probability of χ^2 for a given number of degrees of freedom (ndf), and $\Delta\mathrm{ndf} = 2$ for the spin-0 components, 6 for the $\phi(1680)$ (3 complex amplitudes) and 8 for the $f_0(1525)$ (3 amplitudes, plus mass and width). This is valid in the limit of an infinitely large sample size, and the calculation of $-\Delta\ln(\mathcal{L})$ neglects systematic uncertainties. Although this approximation may lead to an over-estimate of the statistical significance, the decays $B_s^0 \to \phi f_2'(1525)$ and $B_s^0 \to \phi\phi(1680)$ are

significant enough to claim observation.

Table 6.16 *S-wave fractions calculated from Model II in different windows around the $\phi(1020)$ mass. The uncertainties are statistical.*

Window	S-wave fraction
± 10 MeV/c^2	$(0.29 \pm 0.11)\%$
± 15 MeV/c^2	$(0.45 \pm 0.13)\%$
± 20 MeV/c^2	$(0.60 \pm 0.15)\%$
± 25 MeV/c^2	$(0.76 \pm 0.18)\%$

The polarisation parameters of the $B_s^0 \to \phi\phi$ decay are measured to be $\mathcal{F}_0 = (35.3 \pm 1.0)\%$, $\mathcal{F}_\perp = (31.9 \pm 0.6)\%$, and $\delta_\parallel = 2.71 \pm 0.05$, where the uncertainties are statistical. In the LHCb Run 1 angular analysis of $B_s^0 \to \phi\phi$ [49], the polarisation parameters were measured to be $\mathcal{F}_0 = (36.4 \pm 1.2)\%$, $\mathcal{F}_\perp = (30.5 \pm 1.3)\%$, and $\delta_\parallel = 2.54 \pm 0.07$. The agreement between these results is good, considering the differences in the selection and that the angular analysis did not fit the K^+K^- invariant mass. The parameter $\delta_\parallel$ disagrees the most of the three, which could be due to interference from resonant states that were not included in the angular analysis, or it could be due to a bias in the fit model as suggested by the fit to simulation in Section 6.11.2.

The S-wave fraction in ± 10, ± 15, ± 20 and ± 25 MeV/c^2 windows around the $\phi(1020)$ pole mass are obtained by calculating the $B_s^0 \to \phi\phi$ fit fraction within the mass window and assuming that what remains is S-wave. The results are shown in Table 6.16.

6.12.1 Angular moments

The distributions of angular moments depend on the spins of the resonant components and the interference between them. The angular moments $\langle P_l \rangle$ are defined as the $m(K^+K^-)$ distribution weighted by the associated Legendre polynomial function $P_l(\cos\theta_2)$,

$$\langle P_l \rangle = \int_{-1}^{1} d\Gamma(m, \theta_2) P_l(\cos\theta_2) d\cos\theta_2. \tag{6.21}$$

If there is no contribution from resonances with spin > 2, the moments with $l > 4$ should be zero. Each of the distributions can be interpreted as the following: $\langle P_0 \rangle$ is the distribution of events, $\langle P_1 \rangle$ is the sum of the interference between S- and P-wave and P- and D-wave, $\langle P_2 \rangle$ is the sum of P-wave, D-wave and the

interference between S- and D-wave, $\langle P_3 \rangle$ is the interference between P- and D-wave, $\langle P_4 \rangle$ is the D-wave, and $\langle P_5 \rangle$ is the F-wave. The averaging over B_s^0 and $\overline{B}_s^0$ causes a cancellation in the interference terms involving longitudinal P-wave and transverse D-wave, which causes $\langle P_1 \rangle$ and $\langle P_3 \rangle$ to be zero.

The angular moments of the data and Model II are shown for $m(K^+K^-) < 1.05$ GeV/c^2 in Figure 6.25 and for $1.05 < m(K^+K^-) < 1.8$ GeV/c^2 in Figure 6.26. It can be seen from these plots that the model describes the first six angular moments of the data well. The dip in the $\langle P_2 \rangle$ distribution around 1.63 GeV/c^2 was found to be best accommodated by models that include the $\phi(1680)$.

6.12.2 Toy study

Since the values of BIC between the three best fit models do not differ significantly, an alternative approach is taken to compare the goodness of fit. Toy samples are generated from each model with all parameters fixed to the result of the fit to data. Each sample is the same size as the fully selected data set used in the amplitude fit. The samples are fitted using each of the three models, giving a total of nine fits. This exercise is repeated to build up the distributions of $-\ln(\mathcal{L})$ shown in Figure 6.27.

The distributions are fitted with Gaussian functions, and the distance from the mean of the distribution to the value of $-\ln(\mathcal{L})$ observed in data is quantified in terms of the width (σ) of each Gaussian. If the observed $-\ln(\mathcal{L})$ values lie within 1 or 2σ of the mean for each of the fits to samples generated from a given model, then this would suggest that the model used to generate the toys gives a good description of the data. The observed $-\ln(\mathcal{L})$ values are consistently $> 3\sigma$ away from the mean in the same direction. This supports the conclusion from comparing BIC values that none of the three performs significantly better than the others. It also indicates that none of the models describe the data well. One possible reason might be a poor fit to the $B_s^0 \to \phi\phi$ component, as suggested by the results of the fits to simulation in Section 6.11.2.

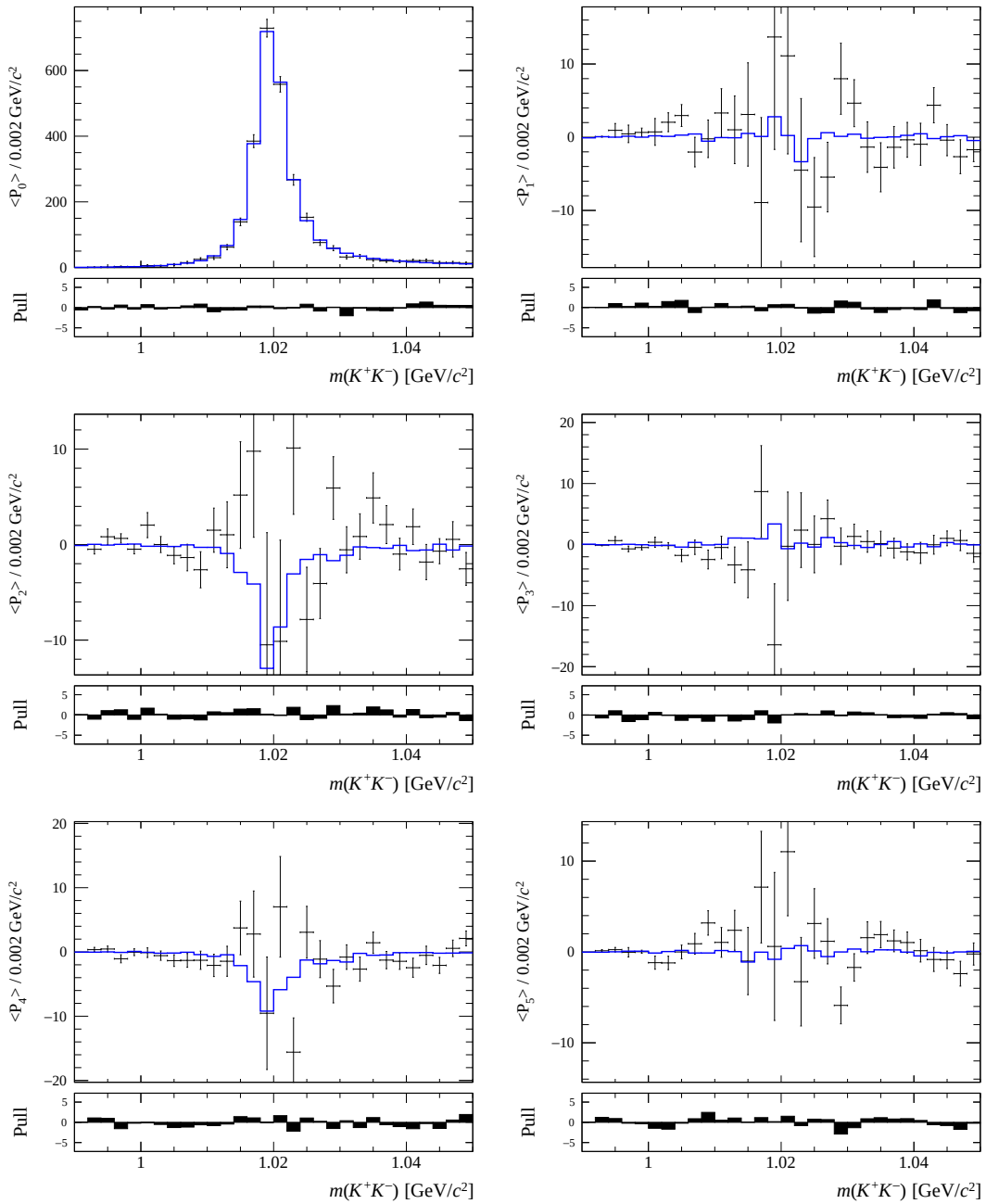


Figure 6.25 *Legendre moment distributions with $l < 6$ in the invariant mass range $m(K^+K^-) < 1.05$ GeV/c^2. The black points are data, and the blue histograms are toy events generated from the PDF.*

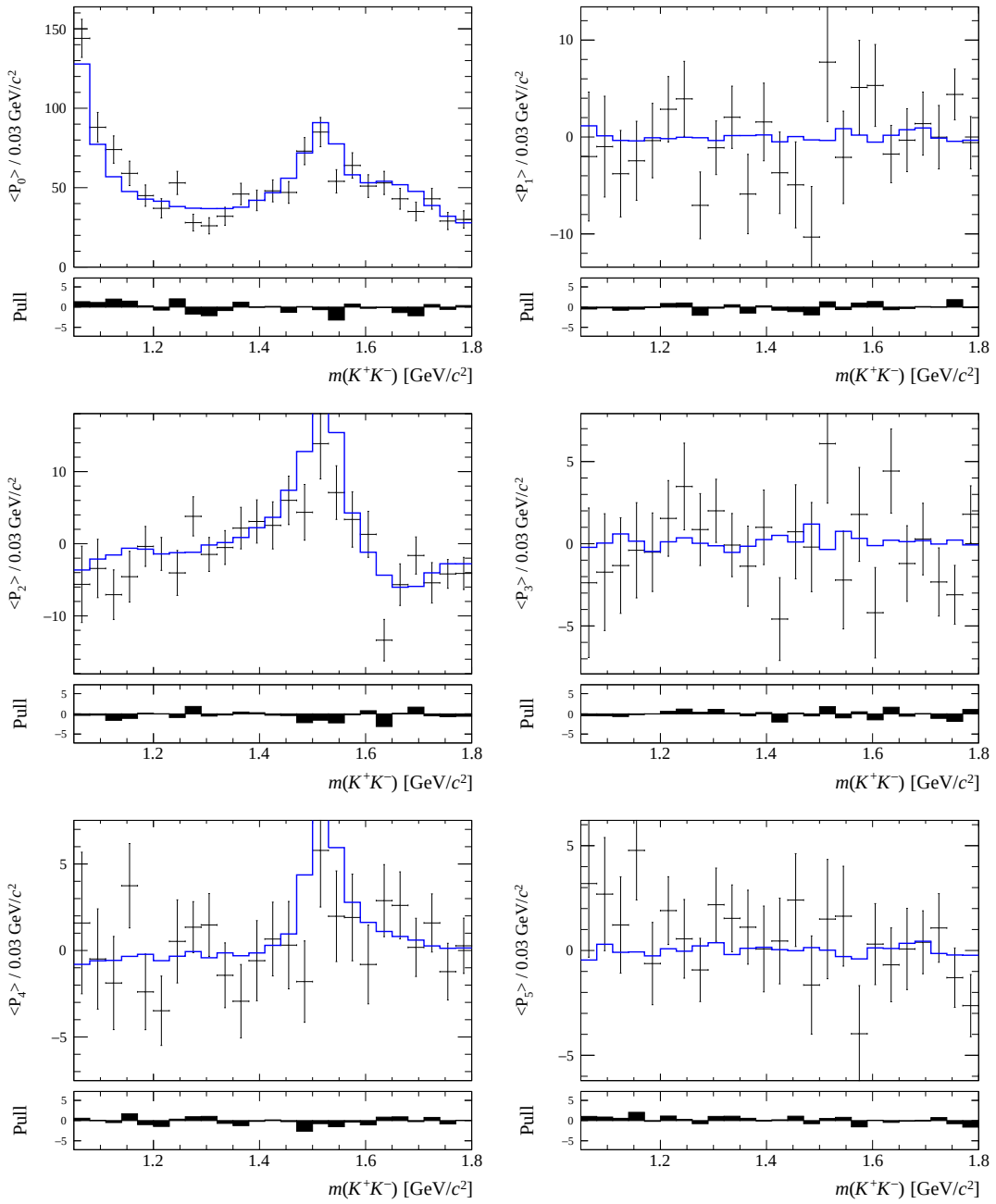
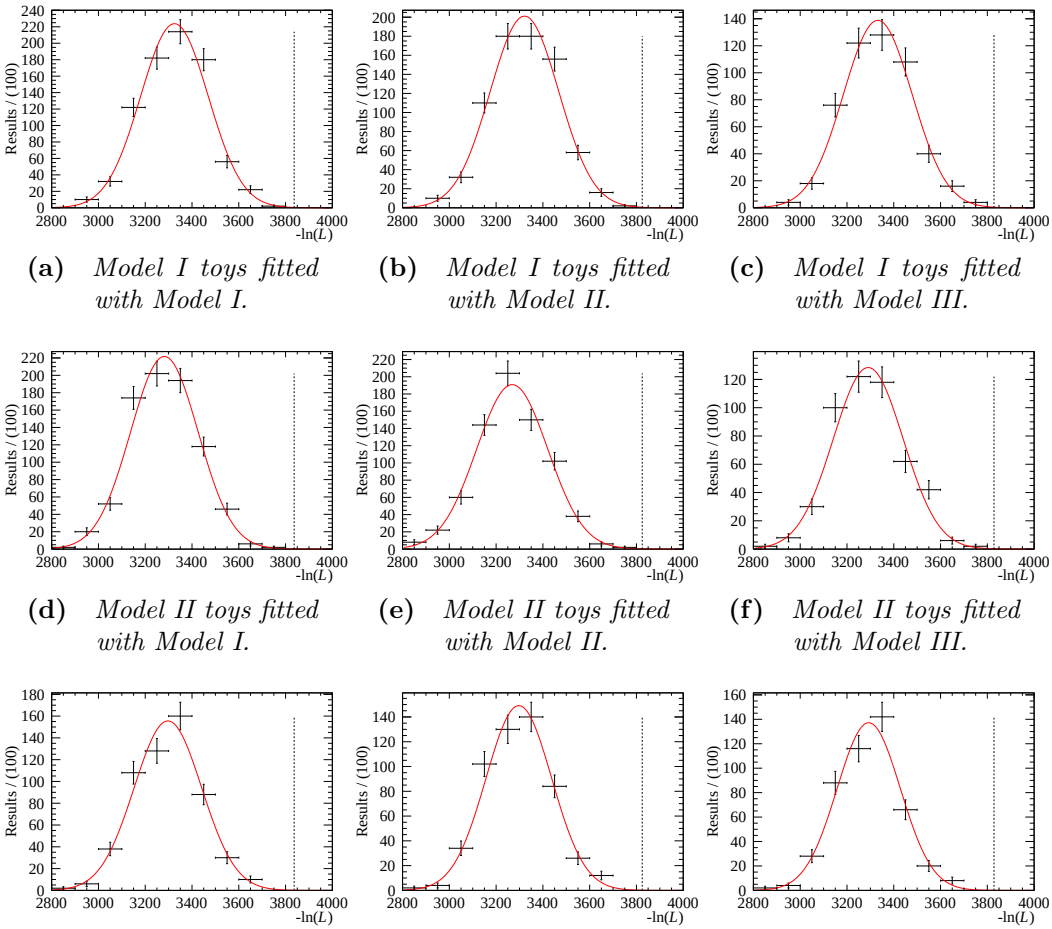


Figure 6.26 *Legendre moment distributions with $l < 6$ in the invariant mass range $1.05 < m(K^+K^-) < 1.8$ GeV/c^2. The black points are data, and the blue histograms are toy events generated from the PDF.*



(a) *Model I toys fitted with Model I.* **(b)** *Model I toys fitted with Model II.* **(c)** *Model I toys fitted with Model III.*

(d) *Model II toys fitted with Model I.* **(e)** *Model II toys fitted with Model II.* **(f)** *Model II toys fitted with Model III.*

(g) *Model III toys fitted with Model I.* **(h)** *Model III toys fitted with Model II.* **(i)** *Model III toys fitted with Model III.*

Figure 6.27 *The distributions of* $-\ln(\mathcal{L})$ *(black points) found by generating toy samples from each of the three best models and then fitting each sample with each model. A fit using a Gaussian function (red line) is overlaid. The value of* $-\ln(\mathcal{L})$ *found from the fits to data is marked as a vertical dashed line.*

6.13 Systematic uncertainties

The systematic uncertainties on each fit parameter are summarised in Table 6.17. The uncertainties are grouped together as arising from the choice of resonances in the signal model, choice of resonance parameters, modelling the detector effects and modelling the background.

The systematic uncertainty arising from the choice of resonances is found by taking the largest deviation on each fit parameter when using either of the two next-best fit models. These are summarised in the 'model' column of Table 6.17. This is the dominant source of systematic uncertainty for the fit fraction and $|\mathcal{A}_0|$ of the

$B_s^0 \to \phi f_2'(1525)$ component. Therefore, it is quoted as a separate uncertainty on the results presented in Section 6.14.

The uncertainty arising from the acceptance function is found by generating a toy sample of 100,000 events from Model II and fitting it 100 times with each coefficient, given in Table 6.5, varied randomly within its statistical uncertainty. This results in distributions for each of the fit parameters. The standard deviations of these distributions are taken as the uncertainty due to the acceptance coefficients. The uncertainty arising from the resolution function is found by performing two fits to the data with the parameter σ_0, defined in Equation 6.13, fixed to the upper or lower bound of its uncertainty. The largest deviation on each parameter from either of these fits is taken as a systematic uncertainty. The acceptance and resolution systematic uncertainties are added in quadrature to give the 'detector' uncertainty in Table 6.17.

The fit is repeated with varied Blatt–Weisskopf barrier factor radii for the B_s^0 meson and the K^+K^- resonance. Two sets of values were chosen: $r_{B_s^0} = 1.0\,\text{GeV}^{-1}\,c$ and $r_{K^+K^-} = 1.8\,\text{GeV}^{-1}\,c$, commonly used by BaBar, and $r_{B_s^0} = 1.5\,\text{GeV}^{-1}\,c$ and $r_{K^+K^-} = 5.0\,\text{GeV}^{-1}\,c$, used in the LHCb analysis of $B_s^0 \to J/\psi\,K^+K^-$ in Ref. [170]. The largest deviation on each parameter is taken as the systematic uncertainty. The fit is repeated with two sets of Flatté parameters for the $f_0(980)$ resonance: $m = 953 \pm 20$ MeV, $g_{\pi\pi} = 329 \pm 96$ MeV and $R_g = 2$, measured by CLEO [175], and $m = 965 \pm 10$ MeV, $g_{\pi\pi} = 165 \pm 18$ MeV and $R_g = 4.21 \pm 0.33$ measured by BES [176]. The largest deviation on each fit parameter is taken as a systematic uncertainty. The mass and width of the $\phi(1680)$, which are fixed to the PDG averages in the nominal fit, are allowed to float with Gaussian constraints, and the deviations from the nominal result are taken as systematic uncertainties. Finally, the mass and width of the $f_0(1500)$, by default fixed to an LHCb measurement [173], are fixed to the PDG averages, and the deviations from the nominal result are taken as systematic uncertainties. The uncertainties from the choice of Blatt–Weisskopf barrier factor radii and other resonance parameters are summed in quadrature to give the 'params' uncertainty in Table 6.17.

The systematic uncertainty due to the background model is found by repeating the fit to data with the fraction of signal candidates and the size of each peaking background constrained to the values found in the fit to the $K^+K^-K^+K^-$ invariant mass, given in Section 6.4. Another fit is performed assuming the background is entirely combinatorial. The deviations on each parameter are summed in quadrature and given as the 'background' uncertainty in Table 6.17.

Table 6.17 *Summary of the systematic uncertainties on each of the fit parameters for the best fit model. The column 'model' lists the uncertainties due to the choice of resonances, 'params' refers to the choice of resonance parameters, 'detector' refers to the uncertainties due to acceptance and resolution, and 'background' refers to the uncertainties due to the size of the combinatorial and peaking backgrounds.*

Fit parameter	Model	Params	Detector	Background
Breit–Wigner parameters				
$\phi(1020)$ m [MeV/c^2]	0.0010	0.0020	0.0010	0.0010
$\phi(1020)$ Γ [MeV]	0.003	0.005	0.0021	0.0008
$f_2'(1525)$ m [MeV/c^2]	4	1.3	0.08	0.07
$f_2'(1525)$ Γ [MeV]	2	0.5	0.4	0.06
Fit fractions				
nonresonant	0.015	0.021	0.0008	0.0011
$f_0(980)$	0.033	0.007	0.0003	0.0015
$\phi(1020)$	0.004	0.021	0.005	0.0009
$f_0(1500)$	—	0.004	0.0003	0.0005
$f_2'(1525)$	0.011	0.004	0.0009	0.0005
$\phi(1680)$	0.006	0.005	0.0007	0.0006
Magnitudes				
$\phi(1020)$ $\lvert A_+\rvert$	1.0×10^{-6}	1.4×10^{-6}	0.0026	1.0×10^{-6}
$\phi(1020)$ $\lvert A_0\rvert$	2.0×10^{-6}	2.5×10^{-6}	0.003	1.0×10^{-6}
$f_2'(1525)$ $\lvert A_+\rvert$	0.03	0.016	0.004	0.004
$f_2'(1525)$ $\lvert A_0\rvert$	0.016	0.006	0.0018	0.0011
$\phi(1680)$ $\lvert A_+\rvert$	0.06	0.015	0.0015	0.0031
$\phi(1680)$ $\lvert A_0\rvert$	0.07	0.03	0.0016	0.004
Phases				
nonresonant δ_0	0.0022	0.05	0.0016	0.0030
$f_0(980)$ δ_0	0.4	0.22	0.009	0.04
$\phi(1020)$ δ_+	0.005	0.012	0.0008	0.0011
$f_0(1500)$ δ_0	—	0.7	0.010	0.018
$f_2'(1525)$ δ_+	0.4	0.16	0.006	0.014
$f_2'(1525)$ δ_0	0.31	0.11	0.009	0.027
$f_2'(1525)$ δ_-	0.2	0.14	0.012	0.023
$\phi(1680)$ δ_+	0.4	0.08	0.006	0.016
$\phi(1680)$ δ_0	0.5	0.4	0.004	0.04
$\phi(1680)$ δ_-	0.4	0.22	0.007	0.030

6.14 Results

The fit fractions of the components in the signal model are calculated using the signal-only PDF without modelling the acceptance and resolution effects. This removes the need to correct for these effects when calculating ratios of $B_s^0 \to \phi X(\to K^+K^-)$ branching fractions. The ratio of fit fractions of the $B_s^0 \to \phi\phi$ and $B_s^0 \to \phi f_2'(1525)$ components in the amplitude fit is therefore equivalent to the quantity

$$\frac{N_{B_s^0 \to \phi f_2'(1525)}}{N_{B_s^0 \to \phi\phi}} \equiv \frac{\mathcal{B}(B_s^0 \to \phi f_2'(1525))\mathcal{B}(f_2'(1525) \to K^+K^-)}{\mathcal{B}(B_s^0 \to \phi\phi)\mathcal{B}(\phi \to K^+K^-)}.$$

Using the PDG averages of the $f_2'(1525) \to K^+K^-$ and $\phi \to K^+K^-$ branching fractions [135]:

$$\mathcal{B}(f_2'(1525) \to K^+K^-) = (44.4 \pm 1.1)\,\%,$$

$$\mathcal{B}(\phi \to K^+K^-) = (48.9 \pm 0.5)\,\%,$$

the branching fraction of $B_s^0 \to \phi f_2'(1525)$ relative to $B_s^0 \to \phi\phi$ is

$$\frac{\mathcal{B}(B_s^0 \to \phi f_2'(1525))}{\mathcal{B}(B_s^0 \to \phi\phi)} = (8.9 \pm 1.0 \pm 0.7 \pm 1.6\ (\text{model}))\ \%,$$

where the first two uncertainties are statistical and systematic, and the uncertainty labelled 'model' is from the choice of fit model. The uncertainties due to the $f_2'(1525) \to K^+K^-$ and $\phi \to K^+K^-$ branching fractions are included in the systematic uncertainty. This is significantly lower than the ratio $\mathcal{B}(B_s^0 \to J/\psi\, f_2'(1525))/\mathcal{B}(B_s^0 \to J/\psi\,\phi) = (25 \pm 5)\,\%$, measured in Ref. [170].

Using the result for the $B_s^0 \to \phi\phi$ branching fraction from Chapter 5, the absolute branching fraction of the decay $B_s^0 \to \phi f_2'(1525)$ is

$$\mathcal{B}(B_s^0 \to \phi f_2'(1525)) = (1.63 \pm 0.18 \pm 0.12 \pm 0.29\ (\text{model}) \pm 0.17\ (\text{norm})) \times 10^{-6},$$

where the uncertainty labelled 'norm' arises from the uncertainty on the $B_s^0 \to \phi\phi$ branching fraction. To prevent the uncertainty due to the $\phi \to K^+K^-$ branching fraction being accounted for twice, it is subtracted from the uncertainty on the normalisation branching fraction. This result is consistent with the theoretical prediction from pQCD: $\mathcal{B}(B_s^0 \to \phi f_2'(1525)) = \left(3.1^{+1.8}_{-1.4} \pm 0.6\right) \times 10^{-6}$ [64].

In the absence of a measurement of $\mathcal{B}(\phi(1680) \to K^+K^-)$, the following quantity

is measured for the decay $B_s^0 \to \phi\phi(1680)$:

$$\frac{\mathcal{B}(B_s^0 \to \phi\phi(1680))\mathcal{B}(\phi(1680) \to K^+K^-)}{\mathcal{B}(B_s^0 \to \phi\phi)} = (2.7 \pm 0.6 \pm 0.4 \pm 0.4 \text{ (model)}) \ \%.$$

The longitudinal polarisation fraction of the $B_s^0 \to \phi f_2'(1525)$ decay is simply the square of the magnitude $|\mathcal{A}_0|$ of the $B_s^0 \to \phi f_2'(1525)$ component in the fit:

$$\mathcal{F}_0 = (86.6 \pm 3.4 \pm 0.8 \pm 2.0 \text{ (model)}) \ \%.$$

This is consistent with the theoretical prediction of $\left(75.3^{+3.0+3.5}_{-3.2-1.7}\right)$ % [64].

The ratio of the $B_s^0 \to \phi f_0(980)$ and $B_s^0 \to \phi\phi$ fit fractions is measured to be

$$\frac{N_{B_s^0 \to \phi f_0(980)}}{N_{B_s^0 \to \phi\phi}} = (4.00 \pm 0.24 \pm 0.35) \ \%,$$

where the second (systematic) uncertainty includes the uncertainty due to the choice of model. This agrees with the value of (5 ± 3) % predicted in Section 6.8.

The ratio of the $B_s^0 \to \phi f_0(1500)$ and $B_s^0 \to \phi\phi$ fit fractions is measured to be

$$\frac{N_{B_s^0 \to \phi f_0(1500)}}{N_{B_s^0 \to \phi\phi}} = (1.3 \pm 0.6 \pm 0.5) \ \%,$$

where no uncertainty is calculated from the choice of model, since Models I and III do not include the $f_0(1500)$. This is significantly smaller than the prediction from Section 6.8 of (3.2 ± 0.6) %, which was calculated under the assumption that the high mass peak seen in Ref. [137] was entirely the $f_0(1500)$, rather than the $f_0(1370)$ or a linear combination of the two.

Chapter 7

Conclusion

As of summer 2017, the sample of b hadron decays collected by LHCb in Run 2 already exceeds the sample used for the analyses presented in this thesis. By the end of Run 2 in late 2018, the LHCb experiment is expected to have a sample of between ~ 5 and ~ 6 fb^{-1} of proton–proton collisions collected at $\sqrt{s} = 13$ TeV [177]. Repetition of the analyses presented in this thesis using data from Run 2, and beyond, will lead to increased precision on the measurements. Additionally, more precise measurements of external inputs to the branching fraction calculations can also reduce the uncertainties.

The $B_s^0 \to \phi\phi$ branching fraction is measured in Chapter 5 to be

$$\mathcal{B}(B_s^0 \to \phi\phi) = (1.84 \pm 0.05(\text{stat}) \pm 0.07(\text{syst}) \pm 0.11(f_s/f_d) \pm 0.12(\text{norm})) \times 10^{-5}.$$

The statistical uncertainty on this measurement is already smaller than the uncertainties from other sources. The third-largest source of uncertainty, which dominates the quoted systematic uncertainty, arises from the S-wave fractions found in angular analyses of the signal [49] and normalisation [47] channels. The precision of these S-wave fractions can be improved by repeating these analyses using Run 2 data. To significantly improve the precision of this branching fraction, more precise measurements of the fragmentation fraction f_s/f_d and the branching fraction of the normalisation channel $B^0 \to \phi K^{*0}$ are required.

The upper limit on the $B^0 \to \phi\phi$ branching fraction is calculated in Chapter 5 to be

$$\mathcal{B}(B^0 \to \phi\phi) < 2.8 \times 10^{-8} (90\,\% \text{ CL}).$$

Since this is already close to the larger of the theoretical predictions, it is likely that LHCb or Belle II will make the first observation of this mode in the coming years.

The $B_s^0 \to \phi f_2'(1525)$ branching fraction is measured in Chapter 6 to be

$$\mathcal{B}(B_s^0 \to \phi f_2'(1525)) = \\ (1.63 \pm 0.18(\text{stat}) \pm 0.12(\text{syst}) \pm 0.29(\text{model}) \pm 0.17(\text{norm})) \times 10^{-6},$$

and the longitudinal polarisation fraction of $B_s^0 \to \phi f_2'(1525)$ is measured to be

$$\mathcal{F}_0 = (86.6 \pm 3.4 \ \ (\text{stat}) \pm 0.8 \ \ (\text{syst}) \pm 2.0 \ (\text{model}))\,\%.$$

Repetition of this analysis using Run 2 data will reduce the statistical uncertainties on both of these values. The increased sample size may allow for the inclusion of more resonant components, such as the $f_2(1270)$, and better discrimination between models, leading to a reduction in the uncertainties arising from the choice of model. The region in $m(K^+K^-)$ above the $f_2'(1525)$ mass can be studied in further detail to give a more precise measurement of the size of the contribution from $B_s^0 \to \phi\phi(1680)$ decays. With a large enough sample, a time-dependent study of *CP* violation in the decay $B_s^0 \to \phi f_2'(1525)$ will be possible. Additionally, an amplitude analysis of the decay $B_s^0 \to K^+K^-K^+K^-$, without imposing invariant mass windows on the kaon pairs, can be performed to search for decays such as $B_s^0 \to f_2'(1525) f_2'(1525)$.

As discussed in Section 2.4.5, the results presented above can be used to reduce the uncertainties on the experimental input used to describe the long-range non-perturbative part of QCD factorisation calculations. In turn, this allows for more precise calculations of *CP* violation in charmless b hadron decays, against which experimental measurements can be compared in the search for non-Standard Model sources of *CP* violation.

Appendix A

The Standard Model

A.1 The electroweak interaction

Before breaking electroweak symmetry, the interaction is mediated by four massless vector-boson fields: W^1, W^2 and W^3 with symmetry $SU(2)$ and coupling g, and B with symmetry $U(1)$ and coupling g'. The kinetic energy term of the electroweak Lagrangian is

$$\mathcal{L}_G = -\frac{1}{4}\sum_{i=1}^{3} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{A.1}$$

where $B_{\mu\nu}$ and $W^i_{\mu\nu}$ are the field strength tensors of the B and W^i fields, given by

$$\begin{aligned} W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\varepsilon_{ijk} W^j_\mu W^k_\nu, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned} \tag{A.2}$$

The interaction term of the Lagrangian is

$$\mathcal{L}_I = i\sum_{j=1}^{3} \left(\overline{\ell}^j_L \not{D}^L \ell^j_L + \overline{e}^j_R \not{D}^R e^j_R + \overline{q}^j_L \not{D}^L q^j_L + \overline{u}^j_R \not{D}^R u^j_R + \overline{d}^j_R \not{D}^R d^j_R, \right) \tag{A.3}$$

where the index j denotes the generations of fermions, ℓ^j_L and q^j_L are left-handed lepton and quark doublets, and e^j_R, u^j_R and d^j_R are the right-handed singlets of charged leptons, up-type quarks and down-type quarks, respectively.[1] The two

[1] There are no right-handed neutrinos in the Standard Model, so there is no $\overline{\nu}^j_R \not{D}^R \nu^j_R$ term.

covariant derivatives acting on left-handed and right-handed fermions are

$$D_\mu^L = \partial_\mu + i\frac{g}{2}\sum_{i=1}^{3}\sigma^i W_\mu^i + i\frac{g'}{2}B_\mu, \tag{A.4}$$

$$D_\mu^R = \partial_\mu + i\frac{g'}{2}B_\mu, \tag{A.5}$$

where σ^i are the Pauli matrices.

So far, this description of the electroweak interaction deals with massless fields, which do not correspond to the particles observed in nature. Electromagnetism is mediated by a massless gauge boson, γ, while the weak interaction is mediated by three massive gauge bosons: W^+, W^- and Z^0. The charged fermions are also massive and carry both left- and right-handed chirality.[2]

The Higgs mechanism provides a way to mix the massless W^i and B fields to produce the observed electroweak bosons and also grant masses to the charged fermions. It does this by introducing a doublet of complex scalar fields

$$\phi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}. \tag{A.6}$$

In terms of real-valued scalar fields, $\varphi^+ = \varphi_1 + i\varphi_2$ and $\varphi^0 = \varphi_3 + i\varphi_4$. The conjugate of this doublet, which will appear later when granting mass to the up-type quarks, is defined as

$$\phi^c = \begin{pmatrix} \overline{\varphi}^0 \\ -\varphi^- \end{pmatrix}. \tag{A.7}$$

The dynamics of this doublet is contained in the Lagrangian term

$$\mathcal{L}_H = \left|D_\mu^H \phi\right|^2 - \mu^2 |\phi|^2 + \lambda |\phi|^4, \tag{A.8}$$

where μ and λ are constants which describe the shape of the Higgs potential. The covariant derivative is

$$D_\mu^H = \partial_\mu + ig\sum_{i=1}^{3} W_\mu^i + i\frac{g'}{2}B_\mu. \tag{A.9}$$

[2]Neutrino oscillation implies they are also massive, but the problem of adding neutrino masses to the Standard Model will not be covered here.

The charged fermions couple to the scalar doublet through the Yukawa interaction terms of the electroweak Lagrangian

$$\mathcal{L}_Y = -\sum_{j=1}^{3} \left(y_e^j \overline{\ell}_L^j \phi e_R^j + y_u^j \overline{q}_L^j \phi^c u_R^j + y_d^j \overline{q}_L^j \phi d_R^j + h.c. \right), \tag{A.10}$$

where y_e^j, y_u^j and y_d^j are the Yukawa couplings for the charged leptons, up-type quarks and down-type quarks for each generation j, and the term $h.c.$ represents the Hermitian conjugates of the three terms shown.

With a choice of $\lambda < 0$ and $\mu^2 < 0$, the Higgs potential has minima at a value of ϕ given by

$$|\phi|^2 = \frac{\mu^2}{2\lambda}. \tag{A.11}$$

A vacuum state is chosen from the set of possible minima such that

$$\langle 0|\varphi_1|0\rangle = \langle 0|\varphi_3|0\rangle = \langle 0|\varphi_4|0\rangle = 0, \tag{A.12}$$

and the vacuum expectation value of the field φ_3 becomes

$$|\langle 0|\varphi_3|0\rangle|^2 = \frac{\mu^2}{\lambda} \equiv v^2. \tag{A.13}$$

The scalar doublet becomes

$$\phi = \begin{pmatrix} 0 \\ \varphi_3 \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{A.14}$$

where H is the field of the Higgs boson.

The full electroweak Lagrangian is the sum of the terms defined in Equations A.1, A.3, A.8 and A.10. By inserting the value of ϕ from Equation A.14 into the full Lagrangian, the physical electroweak gauge bosons now appear as

$$\begin{aligned} W_\mu^+ &= \tfrac{1}{\sqrt{2}} \left(W_\mu^1 - i W_\mu^2 \right), \\ W_\mu^- &= \tfrac{1}{\sqrt{2}} \left(W_\mu^1 + i W_\mu^2 \right), \\ Z_\mu^0 &= B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W, \\ \gamma_\mu &= B_\mu \cos\theta_W - W_\mu^3 \sin\theta_W, \end{aligned} \tag{A.15}$$

where θ_W is a free parameter called the Weinberg angle, which relates the masses of the $W^\pm$ and Z^0 bosons by $m_W = m_Z \cos\theta_W$. The Yukawa term in the Lagrangian

now provides Dirac mass terms for the fermions, *e.g.* for the electron

$$\mathcal{L}_{Ye} = -\frac{y_e^1}{\sqrt{2}}(v+H)(\overline{e}_R e_L + \overline{e}_L e_R) = -\frac{y_e^1}{\sqrt{2}}(v+H)\,\overline{e}e, \tag{A.16}$$

where the electron mass is given by

$$m_e = \frac{y_e^1 v}{\sqrt{2}}. \tag{A.17}$$

A.2 The strong interaction

In the Standard Model, the strong interaction is described by a gauge theory called Quantum Chromodynamics (QCD), which is mediated by eight massless vector boson fields, G, with symmetry $SU(3)$ and coupling g_s. The QCD Lagrangian is given by

$$\mathcal{L}_S = -\frac{1}{4}\sum_{i=1}^{8} G^i_{\mu\nu}G^{i\mu\nu} + i\sum_f \overline{Q}^f\left(\not{D}^S - m_f\right)Q^f, \tag{A.18}$$

where Q^f are quark triplets of a particular flavour $f \in \{u, c, t, d, s, b\}$, $G_{\mu\nu}$ is the gluon field strength tensor

$$G^i_{\mu\nu} = \partial_\mu G^i_\nu - \partial_\nu G^i_\mu - g_s f_{ijk} G^j_\mu G^k_\nu, \tag{A.19}$$

where f_{ijk} are the $SU(3)$ structure constants, and the covariant derivative is

$$D^S_\mu = \partial_\mu - i\frac{g_s}{2}\sum_{i=1}^{8}\lambda^i G^i_\mu, \tag{A.20}$$

where λ^i are the Gell-Mann matrices.

Appendix B

Fit functions

B.1 The Gaussian distribution

The Gaussian distribution is given by

$$f(x|\mu,\sigma) = \frac{1}{N} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{B.1}$$

where μ is the mean and σ is the width of the distribution. The normalisation parameter N is

$$N = \sqrt{2\sigma^2\pi}.$$

B.2 The Crystal Ball distribution

The Crystal Ball function [140] consists of a Gaussian core that transitions to a power-law tail at a certain threshold. It is given by

$$f(x|\mu,\sigma,\alpha,n) = \frac{1}{N} \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) & \frac{x-\mu}{\sigma} > -\alpha, \\ A(B - \frac{x-\mu}{\sigma})^{-n} & \frac{x-\mu}{\sigma} \leq -\alpha, \end{cases} \tag{B.2}$$

with

$$A = \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}},$$

and

$$B = \frac{n}{|\alpha|} - |\alpha|\,,$$

where μ is the mean, σ is the width of the Gaussian core, α controls the position of the tail and n controls the shape of the tail. The normalisation parameter N is given by

$$N = \frac{n\sigma}{|\alpha|\,(n-1)} e^{-\frac{|\alpha|^2}{2}} + \sqrt{\frac{\pi\sigma^2}{2}} \left[1 + \mathrm{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right], \tag{B.3}$$

where 'erf' is the error function.

B.3 The Student-T distribution

The Student-T distribution is a weighted sum of an infinite number of Gaussian distributions with a common mean, μ. It can be written as

$$f(x|\mu, s, n) = \frac{1}{N}\left(1 + \frac{(x-\mu)^2}{ns^2}\right)^{-\frac{n+1}{2}}, \tag{B.4}$$

where n is the number of degrees of freedom and s is a parameter that controls the width of the distribution. The normalisation parameter N is

$$N = \sqrt{\pi n s^2}\frac{\Gamma(n/2)}{\Gamma((n+1)/2)}, \tag{B.5}$$

where Γ is the gamma function. If $n > 2$, the distribution has a variance σ^2 of

$$\sigma^2 = \frac{ns^2}{n-2}. \tag{B.6}$$

In the limit $n \to \infty$, the distribution tends to a Gaussian distribution.

Appendix C

Supplementary MVA plots

This section shows a comparison of the four MVA algorithms investigated in the $B_s^0 \to \phi K^+ K^-$ selection, detailed in Section 6.2.5. Figure C.1 shows the comparison of the MVA outputs for the signal and background training samples, where half of each sample is used for training (plotted as points) and the other half is used for testing (plotted as histograms). There is good agreement between the signal training and testing samples for all algorithms. The agreement between training and testing samples for the background sample is good for the BDTG and MLP algorithms, but not for the BDT and BDTD algorithms. Figure C.2 shows the efficiency curves for each of the algorithms, as well as the figure of merit $S/\sqrt{S+B}$, as a function of the MVA output. The maximum value of the figure of merit is very similar for all algorithms, varying between 34.1 and 34.4.

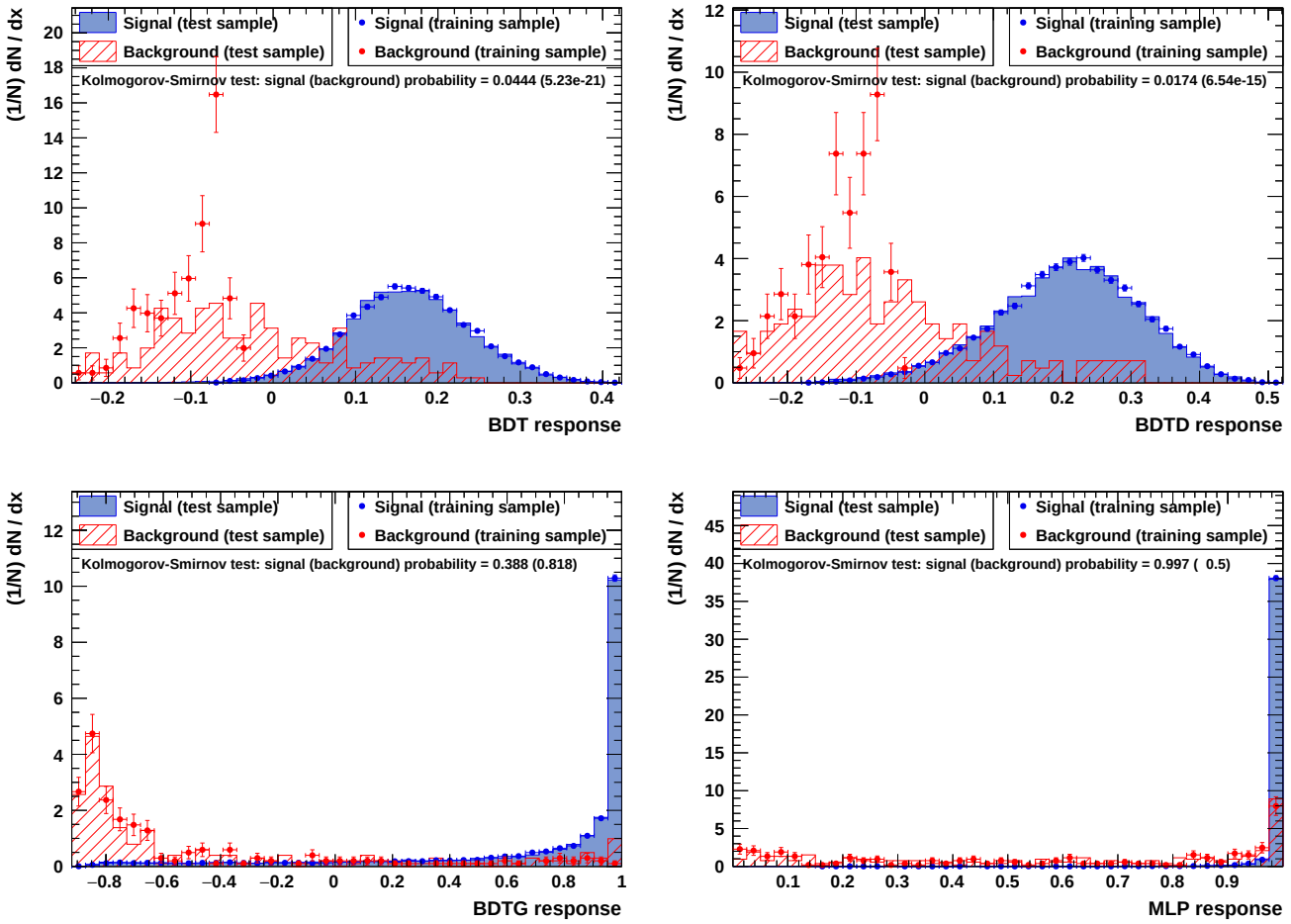


Figure C.1 *Distributions of MVA outputs for the signal and background samples. The test samples are plotted as histograms, and the training samples are plotted as points.*

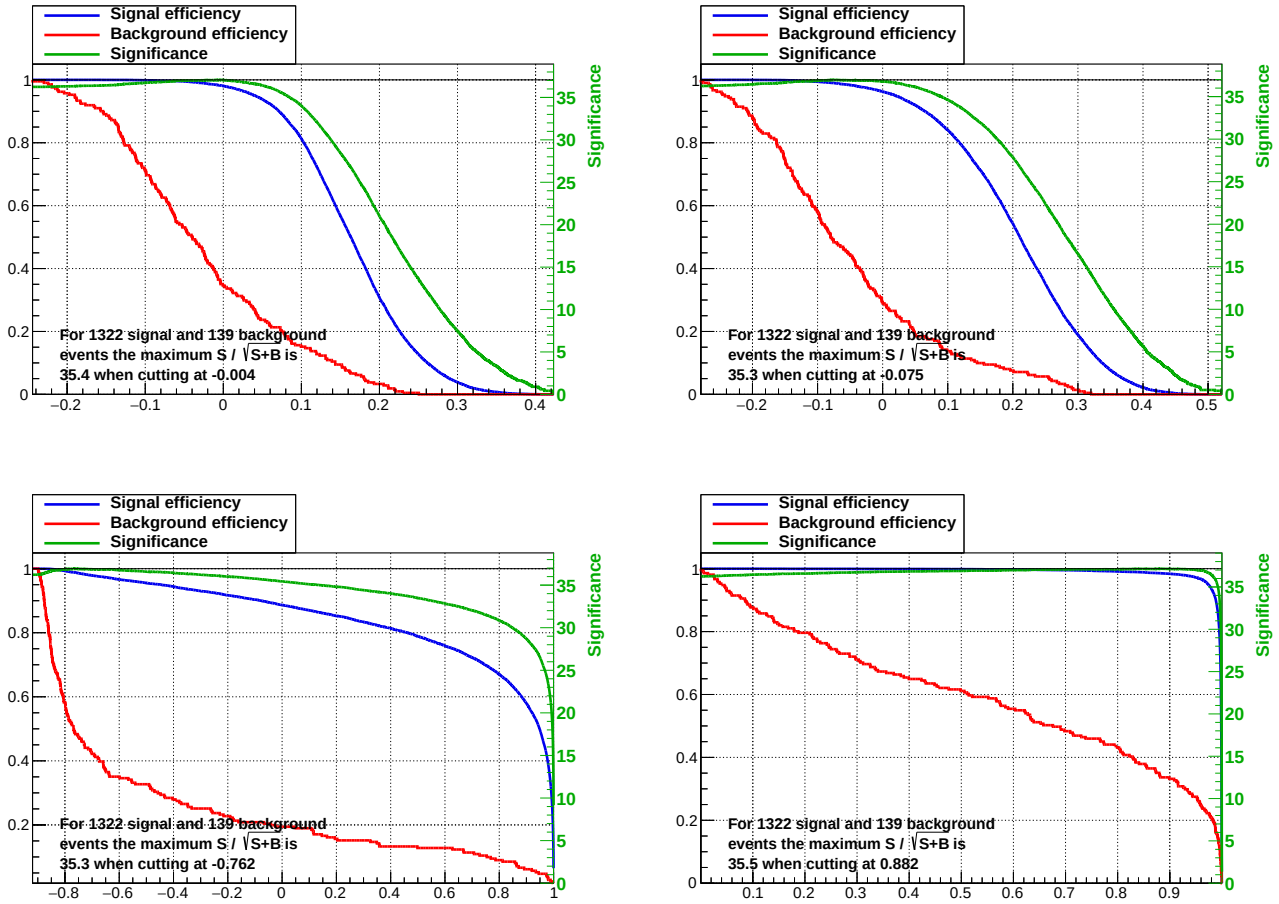


Figure C.2 *Efficiency curves as a function of cut on the MVA output. The green line on each plot is the signal significance figure of merit.*

Appendix D

Alternative angular acceptance

As mentioned in Section 6.6, the angular acceptance is different between TIS and TOS trigger categories. Since these categories are not mutually exclusive, there are two ways to split the events: TOS and not-TOS or TIS and not-TIS. The former was chosen because it splits the data more evenly. The effect of splitting by TIS and not-TIS is explored in this section.

The coefficients of the acceptance function, defined in Equation 6.11, are given in Table D.1. The 1D projections of the acceptance function are shown in Figure D.1 for TIS events and Figure D.2 for not-TIS events.

A fit to data is performed with the best fit model (Model II) using this alternative trigger category splitting and these angular acceptance coefficients. The fit projections are shown in Figure D.3, and the deviation of each parameter from the default fit, in terms of the statistical error on that parameter (σ), is given in Table D.2. Since none of the parameters deviate by as much as $\pm 1\sigma$, the results are very consistent with the default fit.

Table D.1 *Values of the Legendre moment coefficients found using the $B_s^0 \to \phi K^+ K^-$ phase-space simulation sample split by different trigger conditions.*

c^{ijkl}	TIS	not-TIS
c^{0000}	0.0664 ± 0.0002	0.0638 ± 0.0003
c^{0002}	-0.0025 ± 0.0005	—
c^{0200}	-0.0220 ± 0.0010	—
c^{0400}	-0.0070 ± 0.0014	-0.0148 ± 0.0018
c^{1000}	-0.0144 ± 0.0010	—
c^{1200}	-0.0140 ± 0.0021	-0.0133 ± 0.0026

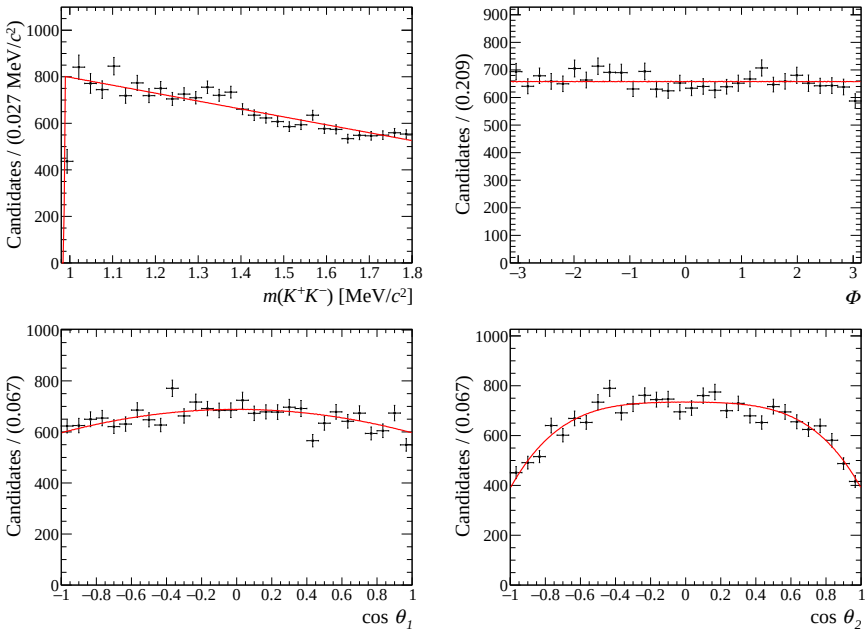


Figure D.1 *1D projections of the acceptance function calculated for TIS events, overlaid onto the $B_s^0 \to \phi K^+ K^-$ simulation sample.*

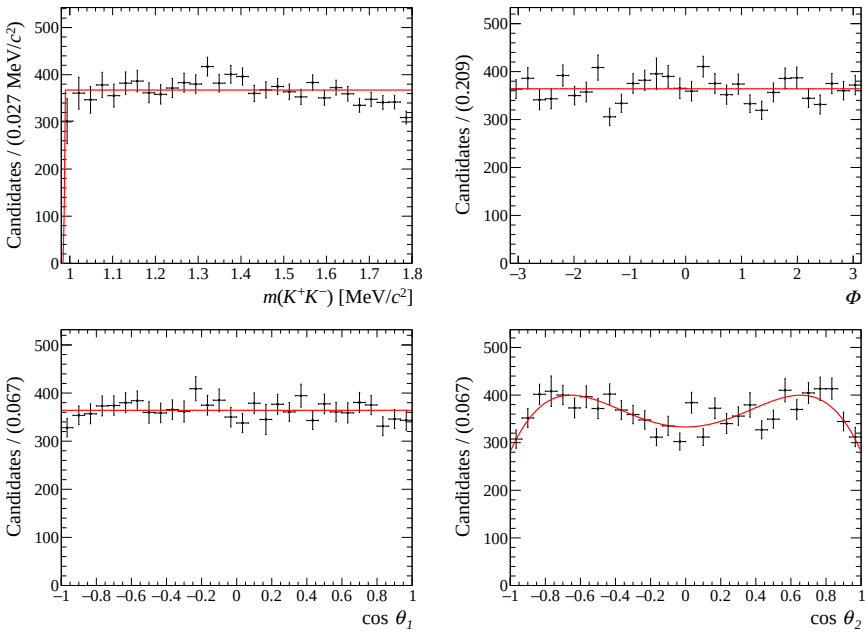


Figure D.2 *1D projections of the acceptance function calculated for not-TIS events, overlaid onto the $B_s^0 \to \phi K^+ K^-$ simulation sample.*

Table D.2 *Deviations of the fit results (in multiples of their statistical error) from using the alternative trigger category splitting.*

Fit parameter	σ from default
Breit-Wigner parameters	
$\phi(1020)$ m [MeV/c^2]	0.000
$\phi(1020)$ Γ [MeV]	0.006
$f_2'(1525)$ m [MeV/c^2]	0.030
$f_2'(1525)$ Γ [MeV]	0.006
Fit fractions	
$f_0(980)$	−0.10
$\phi(1020)$	−0.48
$f_0(1500)$	0.14
$f_2'(1525)$	0.54
$\phi(1680)$	0.03
nonresonant	0.09
interference	−0.07
Magnitudes	
$\phi(1020)$ $\lvert A_+\rvert$	−0.25
$\phi(1020)$ $\lvert A_0\rvert$	0.19
$f_2'(1525)$ $\lvert A_+\rvert$	−0.13
$f_2'(1525)$ $\lvert A_0\rvert$	0.20
$\phi(1680)$ $\lvert A_+\rvert$	−0.005
$\phi(1680)$ $\lvert A_0\rvert$	0.029
Phases [rad]	
nonresonant δ_0	0.03
$f_0(980)$ δ_0	−0.16
$\phi(1020)$ δ_+	0.03
$f_0(1500)$ δ_0	−0.18
$f_2'(1525)$ δ_+	0.009
$f_2'(1525)$ δ_0	−0.05
$f_2'(1525)$ δ_-	0.004
$\phi(1680)$ δ_+	−0.03
$\phi(1680)$ δ_0	0.002
$\phi(1680)$ δ_-	−0.007

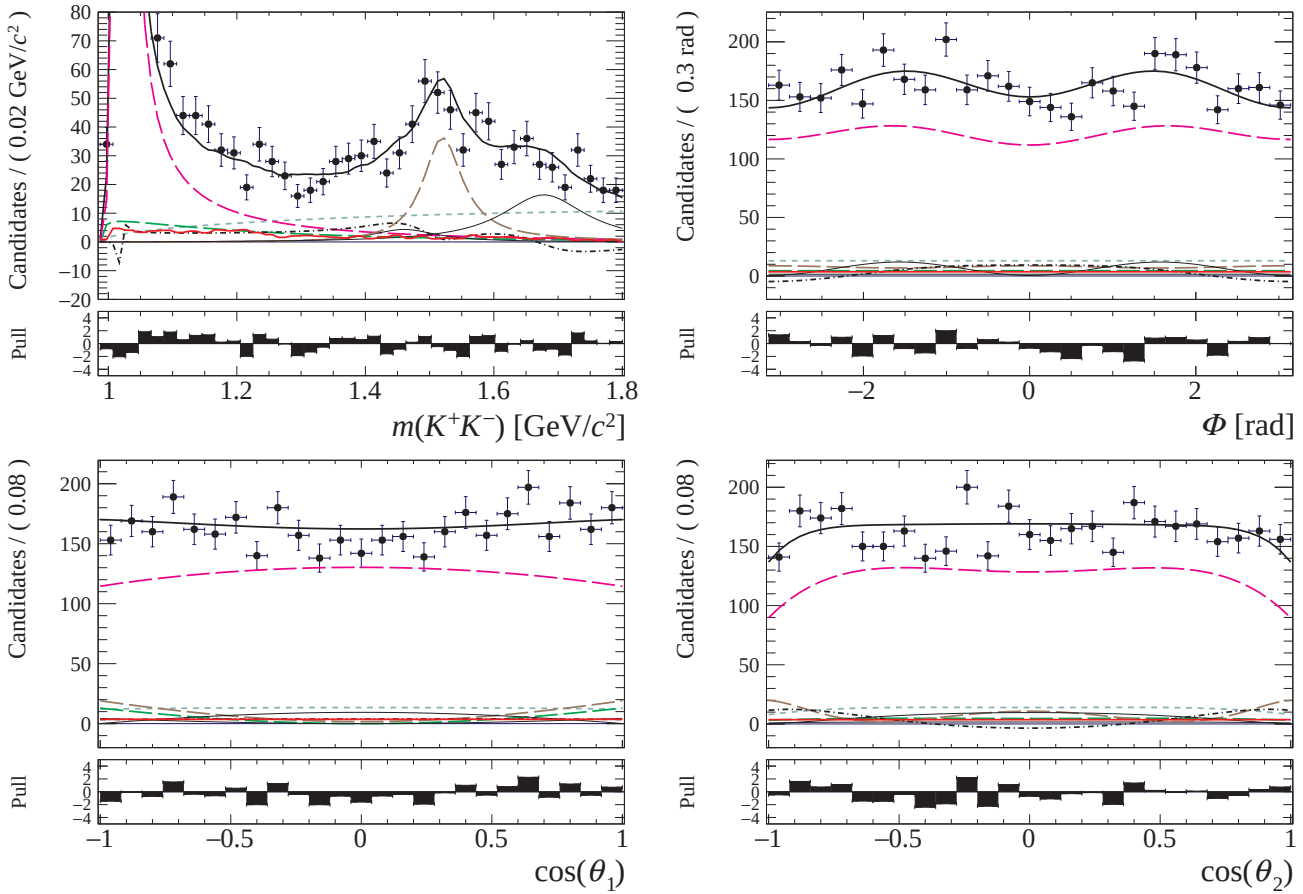


Figure D.3 *Fit projections of Model II using the alternative trigger category splitting in $m(K^+K^-)$ and the helicity angles, overlaid onto the data (black points). The total PDF is a solid black line, the $\phi(1020)$ is a purple long-dashed line, the $f_0(980)$ is a green long-dashed line, the $f_2'(1525)$ is a brown long-dashed line, the nonresonant component is a light blue short-dashed line, the interference is a black dot-dashed line, the $f_0(1500)$ and $\phi(1680)$ are thin solid black lines, and the background is a red histogram.*

Appendix E

2D acceptance plots

This section shows the two-dimensional projections of the four-dimensional acceptance function used in the $B_s^0 \to \phi K^+ K^-$ amplitude analysis, detailed in Section 6.6. Figure E.1 shows the 2D projections of the acceptance function calculated using TOS events, and Figure E.2 shows the same but calculated using not-TOS events.

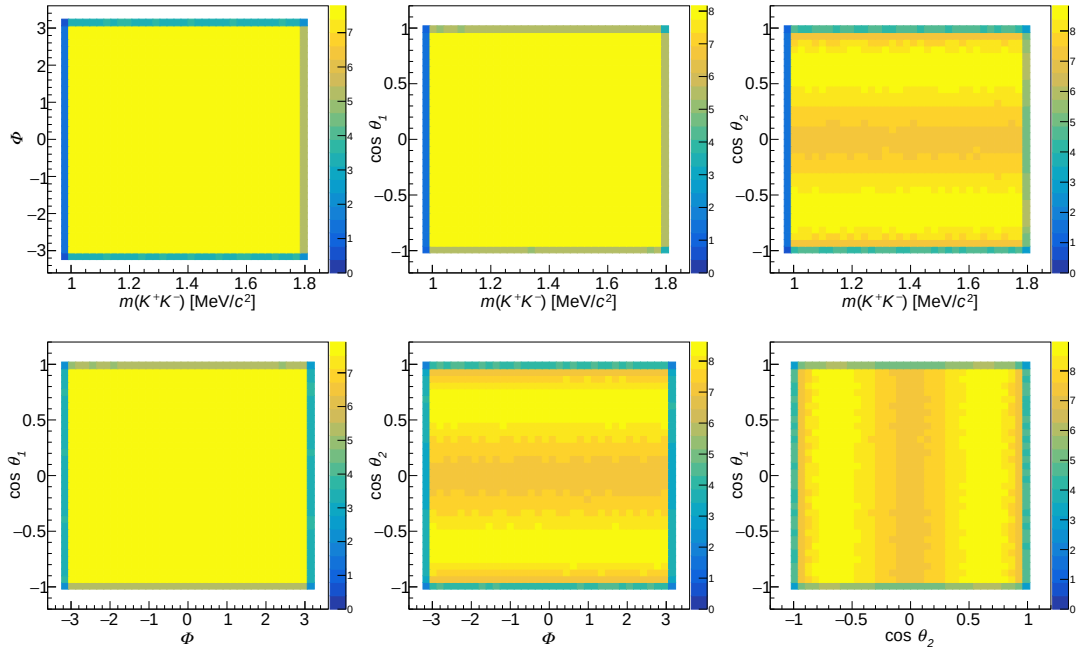


Figure E.1 *2D projections of the acceptance function for TOS events.*

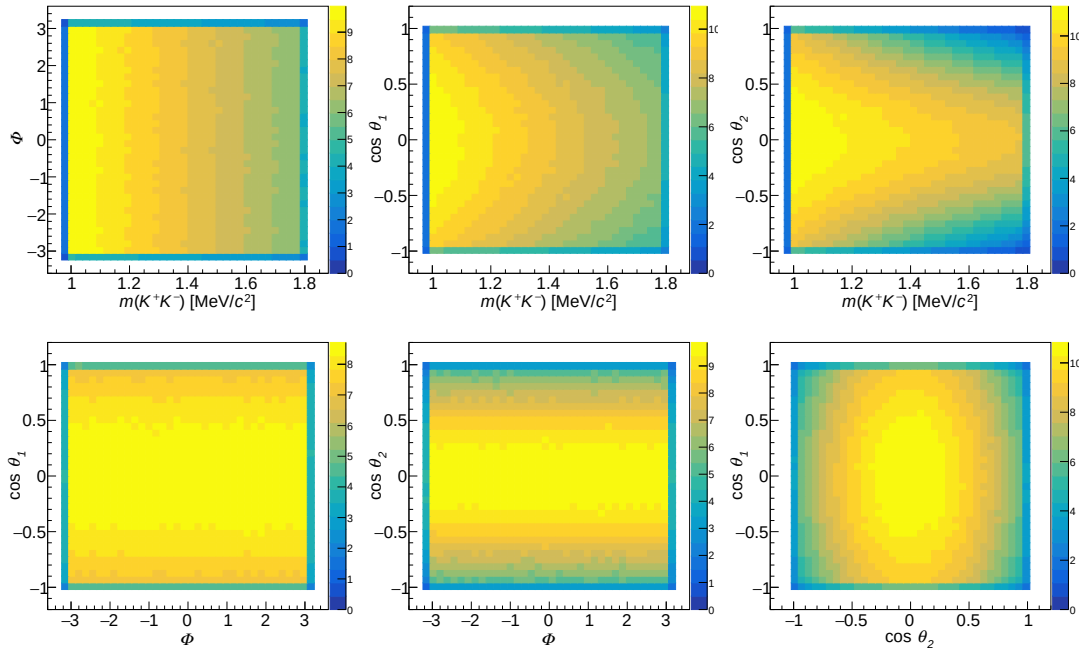


Figure E.2 *2D projections of the acceptance function for not-TOS events.*

Appendix F

Alternative choices of resonances

Table F.1 shows the fit model performance, in terms of $-\ln(\mathcal{L})$, AIC and BIC (defined in Section 6.12) for the first two steps of the iterative procedure used to find the best fit model. Figures F.1 and F.2 show the projections of the fits to data using the two next-best fit models in terms of BIC (Models I and III, respectively).

Table F.1 *Comparison between the different choices of resonances in terms of* $-\ln(\mathcal{L})$*, AIC and BIC.*

Model	$-\ln(\mathcal{L})$	AIC	BIC
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$	3884.7	7803.5	7910.8
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $f_2(1640)$	3883.5	7813.1	7958.2
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $f_2(1750)$	3883.2	7812.4	7957.5
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $f_0(1500)$	3879.1	7796.2	7916.1
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $f_0(1710)$	3881.3	7800.6	7920.5
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $\phi(1680)$	3836.7	7719.4	7864.6
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $\phi(1680)$ $f_2(1640)$	3834.5	7727.0	7910.1
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $\phi(1680)$ $f_2(1750)$	3830.7	7719.3	7902.4
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $\phi(1680)$ $f_0(1500)$	3824.7	7699.4	7857.1
$f_0(980)$ $\phi(1020)$ $f_2'(1525)$ $\phi(1680)$ $f_0(1710)$	3828.5	7707.0	7864.7

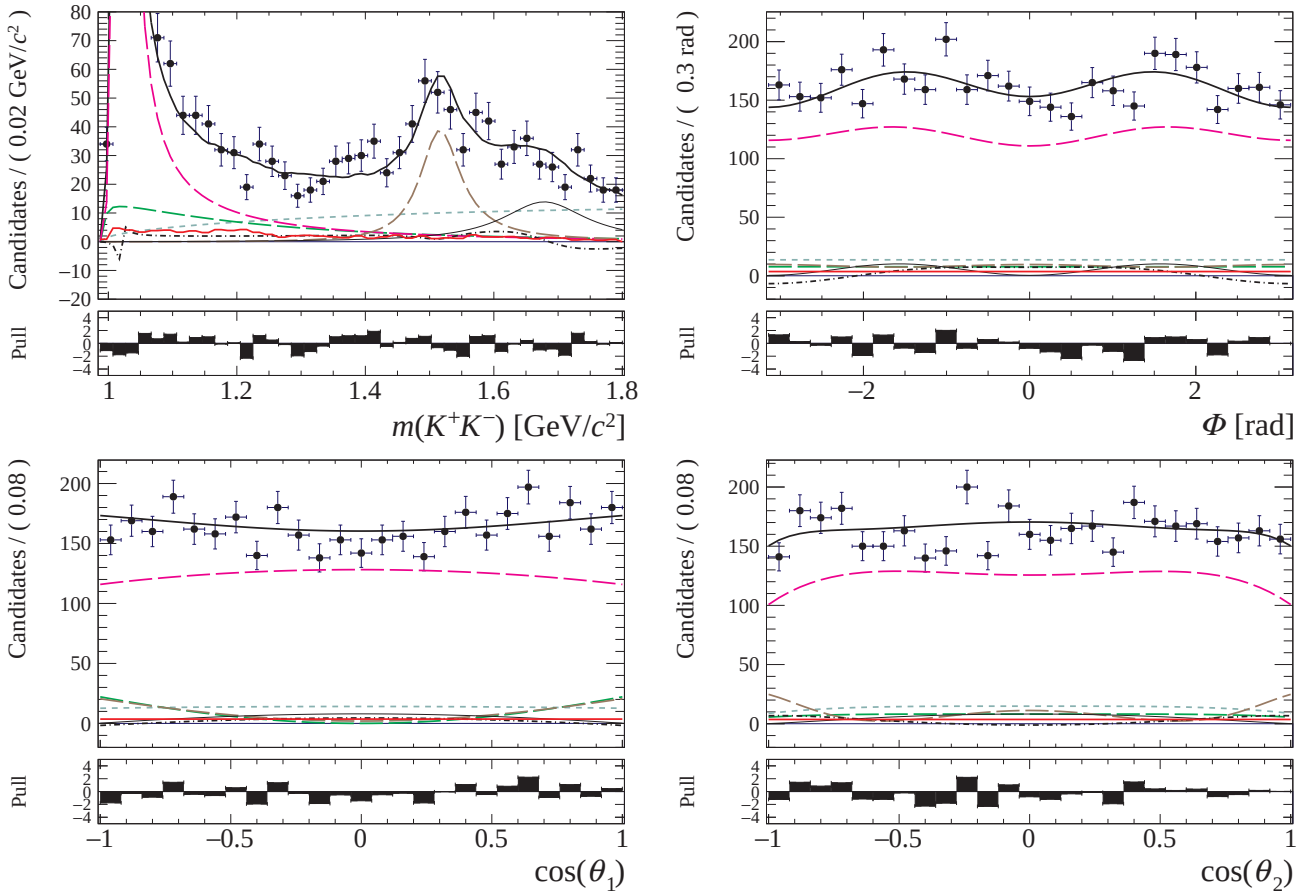


Figure F.1 *Fit projections of Model I in $m(K^+K^-)$ and the helicity angles, overlaid onto the data (black points). The total PDF is a solid black line, the $\phi(1020)$ is a purple long-dashed line, the $f_0(980)$ is a green long-dashed line, the $f_2'(1525)$ is a brown long-dashed line, the nonresonant component is a light blue short-dashed line, the interference is a black dot-dashed line, the $\phi(1680)$ is a thin solid black line, and the background is a red histogram.*

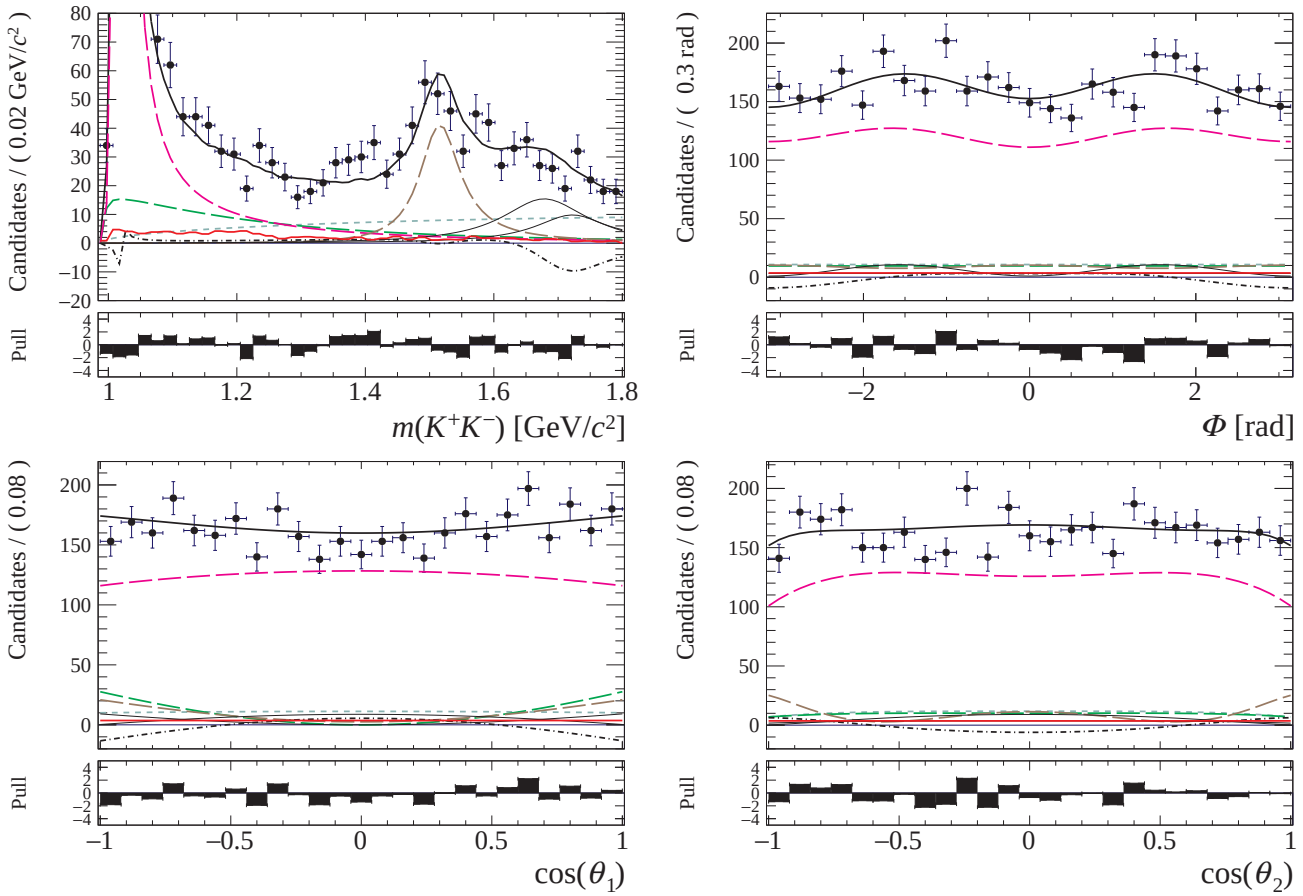


Figure F.2 *Fit projections of Model III in $m(K^+K^-)$ and the helicity angles, overlaid onto the data (black points). The total PDF is a solid black line, the $\phi(1020)$ is a purple long-dashed line, the $f_0(980)$ is a green long-dashed line, the $f_2'(1525)$ is a brown long-dashed line, the nonresonant component is a light blue short-dashed line, the interference is a black dot-dashed line, the $f_0(1710)$ and $\phi(1680)$ are thin solid black lines, and the background is a red histogram.*

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \phi\phi$ branching fraction and search for the decay $B^0 \to \phi\phi$*, JHEP **10** (2015) 053, arXiv:1508.00788.

[2] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and Antimatter in the Universe*, New J. Phys. **14** (2012) 095012, arXiv:1204.4186.

[3] A. D. Sakharov, *Violation of CP Invariance, C-Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[4] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, *Progress in electroweak baryogenesis*, Ann. Rev. Nucl. Part. Sci. **43** (1993) 27, arXiv:hep-ph/9302210.

[5] A. Riotto, *Theories of baryogenesis*, in *Proceedings, Summer School in High-energy physics and cosmology: Trieste, Italy, June 29-July 17, 1998*, pp. 326–436, 1998. arXiv:hep-ph/9807454.

[6] Belle, BaBar, A. J. Bevan *et al.*, *The Physics of the B Factories*, Eur. Phys. J. **C74** (2014) 3026, arXiv:1406.6311.

[7] M. Raidal, *CP asymmetry in $B \to \phi K_S$ decays in left-right models and its implications on B_s decays*, Phys. Rev. Lett. **89** (2002) 231803, arXiv:hep-ph/0208091.

[8] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1, arXiv:hep-ph/0409313.

[9] E. S. Abers and B. W. Lee, *Gauge Theories*, Phys. Rept. **9** (1973) 1.

[10] SLAC-SP-017 collaboration, J. E. Augustin *et al.*, *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406, [Adv. Exp. Phys.5,141(1976)].

[11] E598 collaboration, J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[12] M. L. Perl *et al.*, *Evidence for Anomalous Lepton Production in e^+e^- Annihilation*, Phys. Rev. Lett. **35** (1975) 1489.

[13] S. W. Herb *et al.*, *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.

[14] PLUTO collaboration, C. Berger *et al.*, *Jet Analysis of the* $\Upsilon(9.46)$ *Decay Into Charged Hadrons*, Phys. Lett. **B82** (1979) 449.

[15] PLUTO collaboration, C. Berger *et al.*, *Evidence for Gluon Bremsstrahlung in* e^+e^- *Annihilations at High-Energies*, Phys. Lett. **B86** (1979) 418.

[16] TASSO collaboration, R. Brandelik *et al.*, *Evidence for Planar Events in* e^+e^- *Annihilation at High-Energies*, Phys. Lett. **B86** (1979) 243.

[17] JADE collaboration, W. Bartel *et al.*, *Observation of Planar Three Jet Events in* e^+e^- *Annihilation and Evidence for Gluon Bremsstrahlung*, Phys. Lett. **B91** (1980) 142.

[18] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} =$ 540 *GeV*, Phys. Lett. **B122** (1983) 103.

[19] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/*c^2 *at the CERN SPS Collider*, Phys. Lett. **B126** (1983) 398.

[20] UA2 collaboration, P. Bagnaia *et al.*, *Evidence for* $Z^0 \to e^+e^-$ *at the CERN* $\overline{p}p$ *Collider*, Phys. Lett. **B129** (1983) 130.

[21] UA2 collaboration, M. Banner *et al.*, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN* $\overline{p}p$ *Collider*, Phys. Lett. **B122** (1983) 476.

[22] D0 collaboration, S. Abachi *et al.*, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632, `arXiv:hep-ex/9503003`.

[23] CDF collaboration, F. Abe *et al.*, *Observation of top quark production in* $\bar{p}p$ *collisions*, Phys. Rev. Lett. **74** (1995) 2626, `arXiv:hep-ex/9503002`.

[24] DONUT collaboration, K. Kodama *et al.*, *Observation of tau neutrino interactions*, Phys. Lett. **B504** (2001) 218, `arXiv:hep-ex/0012035`.

[25] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, `arXiv:1207.7214`.

[26] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, `arXiv:1207.7235`.

[27] Various contributors, *Standard Model of Elementary Particles*, Wikimedia Commons, 2006-2017. Taken from `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg` under Creative Commons licence. The source cites PBS, FermiLab and the PDG.

[28] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[29] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[30] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.

[31] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[32] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

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

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

[35] H. Fritzsch and P. Minkowski, *Colored Gluons, Bjorken Scaling and Quark Confinement*, Phys. Lett. **B49** (1974) 462.

[36] SNO, Q. R. Ahmad *et al.*, *Measurement of the rate of* $\nu_e + d \to p + p + e^-$ *interactions produced by* ^8B *solar neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301, `arXiv:nucl-ex/0106015`.

[37] E. Noether, *Invariant variation problems*, Transport Theory and Statistical Physics **1** (1971), no. 3 186, `arXiv:physics/0503066`, M. A. Tavel's English translation of Noether's Theorems (1918), reproduced by Frank Y. Wang.

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

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

[40] G. C. Branco, L. Lavoura, and J. P. Silva, *CP Violation*, Int. Ser. Monogr. Phys. **103** (1999) 1.

[41] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[42] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[43] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[44] L. Zhang and S. Stone, *Time-dependent Dalitz-plot formalism for* $B_q \to J/\psi\ h^+h^-$, Phys. Lett. **B719** (2013) 383, `arXiv:1212.6434`.

[45] Belle collaboration, K. F. Chen *et al.*, *Measurement of branching fractions and polarization in $B \to \phi K^{(*)}$ decays*, Phys. Rev. Lett. **91** (2003) 201801, arXiv:hep-ex/0307014.

[46] BaBar collaboration, B. Aubert *et al.*, *Rates, polarizations, and asymmetries in charmless vector-vector B meson decays*, Phys. Rev. Lett. **91** (2003) 171802, arXiv:hep-ex/0307026.

[47] LHCb collaboration, R. Aaij *et al.*, *Measurement of polarization amplitudes and CP asymmetries in $B^0 \to \phi K^*(892)^0$*, JHEP **05** (2014) 069, arXiv:1403.2888.

[48] CDF collaboration, T. Aaltonen *et al.*, *Measurement of Polarization and Search for CP Violation in $B_s^0 \to \phi\phi$ Decays*, Phys. Rev. Lett. **107** (2011) 261802, arXiv:1107.4999.

[49] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B_s^0 \to \phi\phi$ decays*, Phys. Rev. **D90** (2014) 052011, arXiv:1407.2222.

[50] LHCb collaboration, R. Aaij *et al.*, *Observation of the $\Lambda_b \to \Lambda\phi$ decay*, Phys. Lett. **B759** (2016) 282, arXiv:1603.02870.

[51] LHCb collaboration, R. Aaij *et al.* , and A. Bharucha *et al.*, *Implications of LHCb measurements and future prospects*, Eur. Phys. J. **C73** (2013) 2373, arXiv:1208.3355.

[52] A. Datta *et al.*, *Study of Polarization in $B \to VT$ Decays*, Phys. Rev. **D77** (2008) 114025, arXiv:0711.2107.

[53] A. L. Kagan, *Polarization in $B \to VV$ decays*, Phys. Lett. **B601** (2004) 151, arXiv:hep-ph/0405134.

[54] Heavy Flavour Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, arXiv:1612.07233.

[55] BaBar collaboration, B. Aubert *et al.*, *A Study of $B^0 \to \rho^+\rho^-$ Decays and Constraints on the CKM Angle α*, Phys. Rev. **D76** (2007) 052007, arXiv:0705.2157.

[56] H.-Y. Cheng, C.-K. Chua, and A. Soni, *Final state interactions in hadronic B decays*, Phys. Rev. **D71** (2005) 014030, arXiv:hep-ph/0409317.

[57] P. Colangelo, F. De Fazio, and T. N. Pham, *The Riddle of polarization in $B \to VV$ transitions*, Phys. Lett. **B597** (2004) 291, arXiv:hep-ph/0406162.

[58] H. Li, *Resolution to the $B \to \phi K^*$ polarization puzzle*, Phys. Lett. **B622** (2005) 63, arXiv:hep-ph/0411305.

[59] M. Beneke, J. Rohrer, and D. Yang, *Branching fractions, polarisation and asymmetries of $B \to VV$ decays*, Nucl. Phys. **B774** (2007) 64, arXiv:hep-ph/0612290.

[60] X. Q. Li, G.-r. Lu, and Y. D. Yang, *Charmless $\overline{B_s} \to VV$ decays in QCD factorization*, Phys. Rev. **D68** (2003) 114015, arXiv:hep-ph/0309136, [Erratum: Phys. Rev.D71,019902(2005)].

[61] H.-Y. Cheng and C.-K. Chua, *QCD Factorization for Charmless Hadronic B_s Decays Revisited*, Phys. Rev. **D80** (2009) 114026, arXiv:0910.5237.

[62] Z.-T. Zou *et al.*, *Improved Estimates of The $B_{(s)} \to VV$ Decays in Perturbative QCD Approach*, Phys. Rev. **D91** (2015) 054033, arXiv:1501.00784.

[63] C. Wang, Q.-A. Zhang, Y. Li, and C.-D. Lu, *Charmless $B_{(s)} \to VV$ Decays in Factorization-Assisted Topological-Amplitude Approach*, arXiv:1701.01300.

[64] F. Simanjuntak, C. S. Kim, R. H. Li, and Z. T. Zou, *Charmless $B_{u,d,s} \to VT$ decays in perturbative QCD approach*, Phys. Rev. **D88** (2013) 014031, arXiv:1303.6036.

[65] S. Okubo, *ϕ meson and unitary symmetry model*, Phys. Lett. **5** (1963) 165.

[66] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, in *Developments in the quark theory of hadrons. Vol. 1. 1964 - 1978* (D. B. Lichtenberg and S. P. Rosen, eds.), pp. 22–101. 1964.

[67] J. Iizuka, *Systematics and phenomenology of meson family*, Prog. Theor. Phys. Suppl. **37** (1966) 21.

[68] S. Bar-Shalom, G. Eilam, and Y.-D. Yang, *$B \to \phi\pi$ and $B^0 \to \phi\phi$ in the standard model and new bounds on R parity violation*, Phys. Rev. **D67** (2003) 014007, arXiv:hep-ph/0201244.

[69] C.-D. Lu, Y.-l. Shen, and J. Zhu, *$B^0 \to \phi\phi$ decay in perturbative QCD approach*, Eur. Phys. J. **C41** (2005) 311, arXiv:hep-ph/0501269.

[70] M. Gronau, D. London, and J. L. Rosner, *Rescattering Contributions to rare B-Meson Decays*, Phys. Rev. **D87** (2013), no. 3 036008, arXiv:1211.5785.

[71] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. **B694** (2010) 209, arXiv:1009.2731.

[72] LHCb collaboration, R. Aaij *et al.*, *Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s} = 7$ TeV*, JHEP **08** (2013) 117, arXiv:1306.3663.

[73] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, arXiv:1612.05140, Erratum in preparation.

[74] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment*, Phys. Rev. Lett. **111** (2013) 101805, arXiv:1307.5024.

[75] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay* $B^0 \to K^{*0}\mu^+\mu^-$, JHEP **08** (2013) 131, arXiv:1304.6325.

[76] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase* ϕ_s *in the decay* $B_s^0 \to J/\psi\,\phi$, Phys. Rev. Lett. **108** (2012) 101803, arXiv:1112.3183.

[77] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle* γ *from a combination of LHCb results*, JHEP **12** (2016) 087, arXiv:1611.03076.

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

[79] LHCb collaboration, R. Aaij *et al.*, *A model-independent confirmation of the* $Z(4430)^-$ *state*, Phys. Rev. **D92** (2015) 112009, arXiv:1510.01951.

[80] LHCb collaboration, R. Aaij *et al.*, *Observation of exotic* $J/\psi\,\phi$ *structures from amplitude analysis of* $B^+ \to J/\psi\,\phi K^+$ *decays*, Phys. Rev. Lett. **118** (2016) 022003, arXiv:1606.07895.

[81] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward* W *and* Z *boson production in pp collisions at* $\sqrt{s} = 8\ TeV$, JHEP **01** (2016) 155, arXiv:1511.08039.

[82] LHCb collaboration, R. Aaij *et al.*, *Search for long-lived particles decaying to jet pairs*, Eur. Phys. J. **C75** (2015) 152, arXiv:1412.3021.

[83] LHCb collaboration, R. Aaij *et al.*, *Measurement of charged particle multiplicities and densities in pp collisions at* $\sqrt{s} = 7\ TeV$ *in the forward region*, Eur. Phys. J. **C74** (2014) 2888, arXiv:1402.4430.

[84] LHCb collaboration, R. Aaij *et al.*, *Study of* $\psi(2S)$ *production cross-sections and cold nuclear matter effects in pPb collisions at* $\sqrt{s_{NN}} = 5\ TeV$, JHEP **03** (2016) 133, arXiv:1601.07878.

[85] LHCb collaboration, R. Aaij *et al.*, *Study of* Υ *production and cold nuclear matter effects in pPb collisions at* $\sqrt{s_{NN}} = 5\ TeV$, JHEP **07** (2014) 094, arXiv:1405.5152.

[86] LHCb collaboration, R. Aaij *et al.*, *Study of* J/ψ *production and cold nuclear matter effects in pPb collisions at* $\sqrt{s_{NN}} = 5\ TeV$, JHEP **02** (2014) 072, arXiv:1308.6729.

[87] LHCb collaboration, R. Aaij *et al.*, *Observation of* Z *production in proton-lead collisions at LHCb*, JHEP **09** (2014) 030, arXiv:1406.2885.

[88] LHCb collaboration, L. Massacrier, *First LHCb results from pA and Pb-Pb collisions*, in *4th Large Hadron Collider Physics Conference (LHCP 2016) Lund, Sweden, June 13-18, 2016*, 2016. arXiv:1609.06477.

[89] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[90] ATLAS collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[91] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[92] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[93] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[94] TOTEM collaboration, G. Anelli *et al.*, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.

[95] LHCf collaboration, O. Adriani *et al.*, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.

[96] MoEDAL collaboration, *Technical Design Report of the MoEDAL Experiment*, Tech. Rep. CERN-LHCC-2009-006, CERN, Geneva, 2009. MoEDAL-TDR-001.

[97] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, , contribution to the ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders, CERN, Geneva, Switzerland, 18-22 Mar 2013.

[98] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011. LHCb-TDR-005.

[99] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. LHCb-TDR-009.

[100] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007. LHCb-TDR-001.

[101] LHCb collaboration, *LHCb Inner Tracker: Technical Design Report*, CERN-LHCC-2002-029. LHCb-TDR-008.

[102] LHCb collaboration, *LHCb Outer Tracker: Technical Design Report*, CERN-LHCC-2001-024. LHCb-TDR-006.

[103] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037. LHCb-TDR-003.

[104] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036. LHCb-TDR-002.

[105] LHCb collaboration, *LHCb Muon system: Technical Design Report*, CERN-LHCC-2001-010. LHCb-TDR-004.

[106] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[107] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[108] P. A. Cherenkov, *Visible emission of clean liquids by action of γ radiation*, Dokl. Akad. Nauk SSSR **2** (1934) 451.

[109] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[110] S. Eisenhardt, *Production and tests of Hybrid Photon Detectors for the LHCb RICH detectors*, Nucl. Instrum. Meth. **A595** (2008) 142.

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

[112] R. Cornat, J. Lecoq, and P. Perret, *Level-0 decision unit for LHCb*, Tech. Rep. LHCb-2003-065, CERN, Geneva, Sep, 2003.

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

[114] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

[115] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.

[116] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, LHCb-PUB-2011-016. CERN-LHCb-PUB-2011-016.

[117] LHCb collaboration, *LHCb computing: Technical Design Report*, CERN-LHCC-2005-019. LHCb-TDR-011.

[118] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.

[119] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[120] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[121] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[122] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[123] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

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

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

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

[127] A. Höcker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis with ROOT: Users guide*, Tech. Rep. CERN-OPEN-2007-007, CERN, Geneva, Mar, 2007.

[128] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[129] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, arXiv:physics/0408124.

[130] C. M. Bishop, *Pattern recognition and machine learning*, springer, 2006.

[131] K. Hornik, M. Stinchcombe, and H. White, *Multilayer feedforward networks are universal approximators*, Neural Networks **2** (1989), no. 5 359.

[132] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Neurocomputing: Foundations of research*, ch. Learning Representations by Back-propagating Errors, pp. 696–699. MIT Press, Cambridge, MA, USA, 1988.

[133] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.

[134] CDF collaboration, D. Acosta *et al.*, *First evidence for $B_s^0 \to \phi\phi$ decay and measurements of branching ratio and A_{CP} for $B^+ \to \phi K^+$*, Phys. Rev. Lett. **95** (2005) 031801, arXiv:hep-ex/0502044.

[135] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.

[136] BaBar collaboration, B. Aubert *et al.*, *Searches for B meson decays to $\phi\phi$, $\phi\rho$, $\phi f_0(980)$, and $f_0(980)f_0(980)$ final states*, Phys. Rev. Lett. **101** (2008) 201801, arXiv:0807.3935.

[137] LHCb collaboration, R. Aaij *et al.*, *Observation of the decays $B_s^0 \to \phi\pi^+\pi^-$ and $B \to \phi\pi^+\pi^-$*, Phys. Rev. **D95** (2017) 012006, arXiv:1610.05187.

[138] H. Luo, *Searching for $B^0_{d,s} \to \phi\pi^+\pi^-$ decays*, PhD thesis, University of Edinburgh, 2016.

[139] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[140] 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.

[141] LHCb collaboration, R. Aaij *et al.*, *First observation of the decay $B^0_s \to \phi\overline{K}^{*0}$*, JHEP **11** (2013) 092, arXiv:1306.2239.

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

[143] S. Miglioranzi and G. Corti, *Material Interaction Cross Section Studies*, Tech. Rep. LHCb-INT-2011-002. CERN-LHCb-INT-2011-002, CERN, Geneva, Jan, 2011.

[144] LHCb collaboration, R. Aaij *et al.*, *Prompt K^0_S production in pp collisions at $\sqrt{s} = 0.9$ TeV*, Phys. Lett. **B693** (2010) 69, arXiv:1008.3105.

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

[146] BaBar collaboration, B. Aubert *et al.*, *Time-Dependent and Time-Integrated Angular Analysis of $B \to \phi K_S\pi^0$ and $B \to \phi K^+\pi^-$*, Phys. Rev. **D78** (2008) 092008, arXiv:0808.3586.

[147] Belle collaboration, M. Prim *et al.*, *Angular analysis of $B^0 \to \phi K^*$ decays and search for CP violation at Belle*, Phys. Rev. **D88** (2013), no. 7 072004, arXiv:1308.1830.

[148] LHCb collaboration, R. Aaij *et al.*, *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. **D85** (2012) 032008, arXiv:1111.2357.

[149] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.

[150] LHCb collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, LHCb-CONF-2013-011.

[151] Heavy Flavor Averaging Group (HFAG), Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, arXiv:1412.7515.

[152] M. Jung, *Branching ratio measurements and isospin violation in B-meson decays*, Phys. Lett. **B753** (2016) 187, arXiv:1510.03423.

[153] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

[154] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. **G28** (2002) 2693.

[155] L. Moneta *et al.*, *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv:1009.1003.

[156] LHCb collaboration, R. Aaij *et al.*, *First measurement of the CP-violating phase in $B_s^0 \to \phi\phi$ decays*, Phys. Rev. Lett. **110** (2013) 241802, arXiv:1303.7125.

[157] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes and triple product asymmetries in the $B_s^0 \to \phi\phi$ decay*, Phys. Lett. **B713** (2012) 369, arXiv:1204.2813.

[158] V. V. Gligorov, *Reconstruction of the channel $B_d^0 \to D^+\pi^-$ and background classification at LHCb (revised)*, LHCb-2007-044.

[159] N. Smirnov, *Table for estimating the goodness of fit of empirical distributions*, Ann. Math. Statist. **19** (1948) 279.

[160] A. Mahdavi, *Robust goodness of fit test based on the forward search*, American Journal of Applied Mathematics and Statistics **1** (2013), no. 1 6.

[161] G. A. Cowan, D. C. Craik, and M. D. Needham, *RapidSim: an application for the fast simulation of heavy-quark hadron decays*, Comput. Phys. Commun. **C214** (2017) 239, arXiv:1612.07489.

[162] X. Liu, Z.-J. Xiao, and Z.-T. Zou, *Nonleptonic decays of $B \to (f_1(1285), f_1(1420))V$ in the perturbative QCD approach*, Phys. Rev. **D94** (2016), no. 11 113005, arXiv:1609.01024.

[163] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics*, Springer-Verlag, 1979.

[164] F. Von Hippel and C. Quigg, *Centrifugal-barrier effects in resonance partial decay widths, shapes, and production amplitudes*, Phys. Rev. **D5** (1972) 624.

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

[166] S. M. Flatté, *Coupled - Channel Analysis of the $\pi\eta$ and $K\overline{K}$ Systems Near $K\overline{K}$ Threshold*, Phys. Lett. **B63** (1976) 224.

[167] LHCb collaboration, R. Aaij *et al.*, *Analysis of the resonant components in $\overline{B}_s^0 \to J/\psi\pi^+\pi^-$*, Phys. Rev. **D86** (2012) 052006, arXiv:1204.5643.

[168] Crystal Barrel, C. Amsler *et al.*, *Study of $K\overline{K}$ resonances in $\overline{p}p \to K^+K^-\pi^0$ at 900* MeV/c *and 1640* MeV/c, Phys. Lett. **B639** (2006) 165.

[169] G. Costa *et al.*, *An amplitude analysis of the K^+K^- system produced in the reaction $\pi^-p \to K^+K^-n$ at 10* GeV/c, Nucl. Phys. **B175** (1980) 402.

[170] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis and branching fraction measurement of $\overline{B}^0_s \to J/\psi K^+K^-$*, Phys. Rev. **D87** (2013) 072004, arXiv:1302.1213.

[171] BES collaboration, M. Ablikim *et al.*, *Partial wave analysis of $\chi_{(c0)} \to \pi^+\pi^-K^+K^-$*, Phys. Rev. **D72** (2005) 092002, arXiv:hep-ex/0508050.

[172] Belle, K. Abe *et al.*, *Measurement of K^+K^- production in two photon collisions in the resonant mass region*, Eur. Phys. J. **C32** (2003) 323, arXiv:hep-ex/0309077.

[173] LHCb collaboration, R. Aaij *et al.*, *Measurement of resonant and CP components in $\overline{B}^0_s \to J/\psi\pi^+\pi^-$ decays*, Phys. Rev. **D89** (2014) 092006, arXiv:1402.6248.

[174] G. Schwarz, *Estimating the dimension of a model*, Ann. Statist. **6** (1978) 461.

[175] CLEO, G. Bonvicini *et al.*, *Dalitz plot analysis of the $D^+ \to \pi^-\pi^+\pi^+$ decay*, Phys. Rev. **D76** (2007) 012001, arXiv:0704.3954.

[176] BES, M. Ablikim *et al.*, *Resonances in $J/\psi \to \phi\pi^+\pi^+$ and ϕK^+K^-*, Phys. Lett. **B607** (2005) 243, arXiv:hep-ex/0411001.

[177] A. Hicheur, *Review of recent LHCb results and prospects for Run II*, in *International School on High Energy Physics: Session C: Workshop in HEP (LISHEP 2015) Manaus, Amazonas, Brazil, August 2-9, 2015*, 2015. arXiv:1509.07708.