THE UNIVERSITY *of* EDINBURGH

Measurement of the CP violating phase ϕ_s using $B_s^0 \to \psi(2S)\phi$ decays at the LHCb Experiment

Dianne Ferguson

Doctor of Philosophy
The University of Edinburgh
May 2016

Abstract

The LHCb experiment at the Large Hadron Collider (LHC) at CERN is designed to make precise measurements of processes including B and D mesons to test the Standard Model (SM) predictions for *CP* violation, and to search for new physics. From its inception one of the key aims of the LHCb collaboration has been to precisely measure the *CP* violating phase ϕ_s, the weak phase due to the interference between B_s^0-$\overline{B}_s^0$ mixing and decay.

Having collected $3\,\text{fb}^{-1}$ of data in Run 1, the combined results of LHCb measurements of ϕ_s from various decay modes are in agreement with SM predictions. The aim now is to improve the precision of the LHCb measurement to be sensitive to any small deviation from the SM prediction of ϕ_s. One strategy to achieve this, in addition to collecting more data, is to expand the number of modes used to measure ϕ_s to improve the sensitivity of the combination. This thesis presents the measurement of the *CP* violating phase ϕ_s in the yet unstudied $B_s^0 \to \psi(2S)\phi$ decay mode.

In addition to providing a measurement of ϕ_s the study of this mode presents an opportunity to confirm the lifetime difference of the B_s^0 mass eigenstates $\Delta\Gamma_s$, currently only measured in the $B_s^0 \to J/\psi\,\phi$ decay mode. The results from $3\,\text{fb}^{-1}$ of LHCb data are;

$$\phi_s = 0.23^{+0.29}_{-0.28} \pm 0.02 \text{ rad},$$
$$\Delta\Gamma_s = 0.066^{+0.041}_{-0.044} \pm 0.007 \text{ ps}^{-1}.$$

which are in agreement with the SM and the results from the LHCb measurement from $B_s^0 \to J/\psi\,\phi$ decays.

Lay Abstract

The study of *CP* violation is an important step in understanding how the Universe evolved from the time of the Big Bang. Current theories on the origin of the Universe suggest that the Big Bang created equal amounts of matter and antimatter. However, the observable Universe that is seen today is formed almost completely of matter. This suggests that there is some difference in the behaviour between matter and anti-matter, resulting in the large inequality in the amounts of the two.

One requirement for a matter/antimatter asymmetry is *CP* violation. The current model of particle physics (called the Standard Model) includes a source of *CP* violation, that has been tested extensively and has been found to be accurate. However, the amount of *CP* violation required to produce the observed Universe is much larger than the current model accounts for, so new sources of *CP* violation are required.

This thesis presents the measurement of one such *CP* violating parameter in order to test the Standard Model predictions, and search for new physics.

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.

(*Dianne Ferguson, May 2016*)

Acknowledgements

First and foremost, I would like to thank my supervisor, Peter Clarke, for his guidance and expertise throughout my entire PhD. Many thanks also to Greig Cowan for his continued advice, encouragement and endless positivity.

The Edinburgh experimental particle physics group has changed a lot over the years and I would like to thank everyone who has been part of it for making the department such a unique and fun place to work. In particular, I would like to thank Tim and Paul for sharing the PhD journey with me and keeping me sane over the years.

I would like to thank my family for being a constant source of support over the duration of my PhD. Their support has taken many forms and I am grateful for every one.

Finally, I dedicate this thesis to my sister, Karen, who is the best partner in crime I could hope to have.

Contents

Chapter 1

Introduction

One of the ongoing questions in physics comes from the knowledge that the observable Universe is composed almost entirely out of matter, a circumstance believed to have arisen in the early development of the Universe. This question arises because at a time just after the Big Bang the Universe contained equal amounts of matter and anti-matter, so there must be some process through which the Universe develops an asymmetry in the number of baryons and anti baryons, called baryogenesis. To address this issue, Andrei Sakharov proposed three conditions required at the time of the early Universe, which would result in baryogenesis. The conditions are as follows [1]:

- Violation of Baryon number (B).
- Processes that violate combined Charge - Parity (CP) symmetry.
- A departure from thermal equilibrium.

The first condition follows naturally from the observation of a matter dominated Universe as to have gone from having a baryon number of zero (assigning a value of $+1$ to baryons, and assigning a value of -1 for anti baryons), to a non-zero baryon number, there must have been baryon number violation. The second condition is required because if CP was conserved, processes that generate a baryon asymmetry would be compensated by the CP conjugate process, resulting in an overall conservation of baryon number. Finally, a departure from thermal equilibrium is required because without this the baryon number violating processes

would be balanced by the processes that violate baryon number in the opposite direction.

The main focus of the LHCb collaboration is on the second of the three Sakharov conditions. In particular the focus is on measurements of *CP* violation in B meson decays. In addition to testing the predictions for *CP* violation in this sector, outlined by the Standard Model (SM) of particle physics, this topic is particularly interesting because the amount of *CP* violation accounted for in the SM is not enough to describe the size of the matter anti matter symmetry observed in the Universe so new sources are required.

This thesis makes a contribution to the measurement of *CP* violation by presenting the measurement of the *CP* violating phase ϕ_s in $B_s^0 \to \psi(2S)\phi$ decays, where the $\psi(2S)$ decays to two muons ($\mu^+\mu^-$) and the ϕ decays to two kaons (K^+K^-) using data taken by the LHCb detector at the Large Hadron Collider (LHC).

Chapter 2

CP Violation in the Standard Model

The effect of the combination of the Charge conjugation (C) and Parity (P) operators is to transform particle to antiparticle in a mirror world where the spatial coordinates have been reversed. Once assumed to be a symmetry of nature, the experiments of Christenson et. al. showed that this symmetry is violated by the weak interaction [2]. This experiment studied the decay of neutral K mesons to π mesons. Neutral kaons are produced via the strong interaction and decay via the weak interaction to either two or three pion final states. The neutral K mesons, K^0 and its anti-particle $\overline{K}^0$, can mix via the weak force. CP eigenstates (K_1^0 and K_2^0) can be formed from linear combinations of the neutral kaons as shown in Equation 2.1.

$$\begin{aligned} K_1^0 &= \frac{1}{\sqrt{2}}(K^0 + \overline{K}^0) \\ K_2^0 &= \frac{1}{\sqrt{2}}(K^0 - \overline{K}^0) \end{aligned} \tag{2.1}$$

Both the K_1^0 and two pion final state are CP even, while the K_2^0 and three pion final state are CP odd. The choice of assigning $CP(K_1^0) = +1$ and $CP(K_2^0) = -1$ follows a well established convention [3]. This indicates that if the weak force conserved CP symmetry the K_1^0 could not decay to the three π state, and the K_2^0 could not decay to the two π state. This restriction on the final state determines the lifetimes of the neutral kaons due to the phase space available, where the K_2^0 has a long lifetime relative to the K_1^0.

All of this information combined provides the right arena in which to search for CP violation. If the weak interaction was CP invariant this would be observed experimentally in a beam of neutral kaons as seeing the decay to the two pion

state dominate close to the beam source and the three pion state exclusively far from the source.

The experimental observation found evidence for the long-lived kaon state decaying to $\pi^+\pi^-$ with a branching ratio of $\sim 2\times10^{-3}$, concluded to be an observation of *CP* violation. The inclusion of a model for *CP* violation in the SM was an important moment in particle physics and led to the proposed and subsequent discovery of the third generation of quarks [4]. This chapter discusses the description of *CP* violation in the quark sector of the SM and the *CP* violating parameter of interest to this analysis, ϕ_s.

2.1 The CKM Mechanism

The inclusion of *CP* violation in the SM is facilitated by the CKM matrix [4], shown in Equation 2.2, which describes the rotation between the weak and mass eigenstates of the down type quarks; down (d), bottom (b), and strange (s). The u, c and t indices refer to the up, charm and top quarks respectively. This matrix arises from the Yukawa couplings of the quarks to the Higgs field, included in the SM Lagrangian to generate the mass terms of the quarks. As the CKM matrix is a 3×3 unitary matrix it can be parameterised by three mixing angles and a complex phase.

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.2}$$

The CKM matrix has an inherent structure, which can be observed when expressed using the Wolfenstein parameterisation [5] (Equation 2.3), a power expansion about the sine of the Cabbibo angle (λ), the mixing angle of the first two generations of quarks [6]. The other parameters appearing in the Wolfenstein parameterisation are defined as follows; A is a constant and is approximately equal to 1, and $\bar{\rho}$ and $\bar{\eta}$ are real parameters satisfying $\bar{\rho} + i\bar{\eta} = -i(V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$ to $\mathcal{O}(\lambda^3)$. A non-zero value of $\bar{\eta}$ introduces the complex phase required for *CP* violation. In this parameterisation the elements along the diagonal are approximately equal to 1, with the elements further from the diagonal getting progressively smaller. In physical terms this means that the strongest interactions of the weak force involving quarks occur in interactions involving quarks of the same generation,

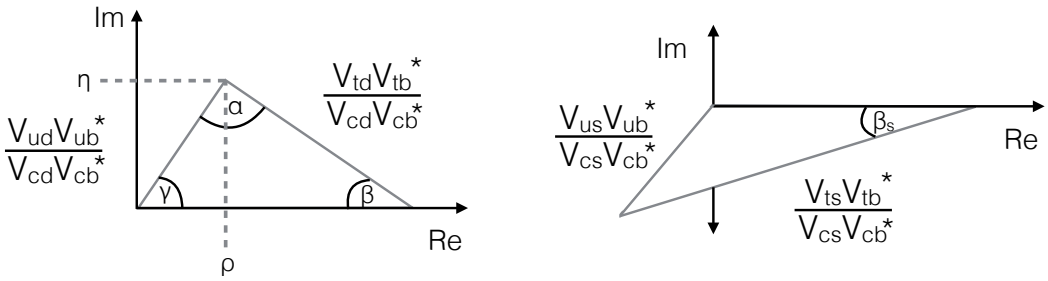


Figure 2.1 *Diagrams of two Unitarity triangles defined by constraints due to the unitarity of the CKM matrix (Equation 2.4). To the left is the B^0 triangle and the right is the B_s^0 triangle.*

e.g. the u and d quarks, or the c and s quarks.

$$V_{CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3\left(1+\frac{\lambda^2}{2}\right)(\bar{\rho}-i\bar{\eta}) \\ -\lambda - iA^2\lambda^5\bar{\eta} & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\bar{\rho}-i\bar{\eta}) & -A\lambda^2 - iA\lambda^4\bar{\eta} & 1 \end{pmatrix} + \mathcal{O}(\lambda^6) \tag{2.3}$$

The unitarity condition of the CKM matrix results in 9 relations involving the CKM matrix elements, 6 of which are vanishing. The most commonly used relations are shown in Equation 2.4. These relations can be traced as triangles in the complex plane (Figure 2.1) with equal areas given by one half of Jarlskog parameter (J= $A^2\lambda^6\eta + \mathcal{O}(\lambda^8)$) [7].

$$\begin{aligned} V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0 \\ V_{ub}^* V_{us} + V_{cb}^* V_{cs} + V_{tb}^* V_{ts} = 0 \end{aligned} \tag{2.4}$$

Measurements of the angles of these triangles, named α, β, γ and β_s, and the CKM matrix elements can be used as input to a global fit for the CKM parameters λ, A, $\bar{\rho}$ and $\bar{\eta}$, performed by the CKMfitter group [8]. The results of this fit can be used in turn to provide self-consistent predictions for measurable *CP* violating parameters, as shown in Figure 2.2. This over-constraining of the CKM matrix provides a way of testing the accuracy of the SM description of *CP* violation.

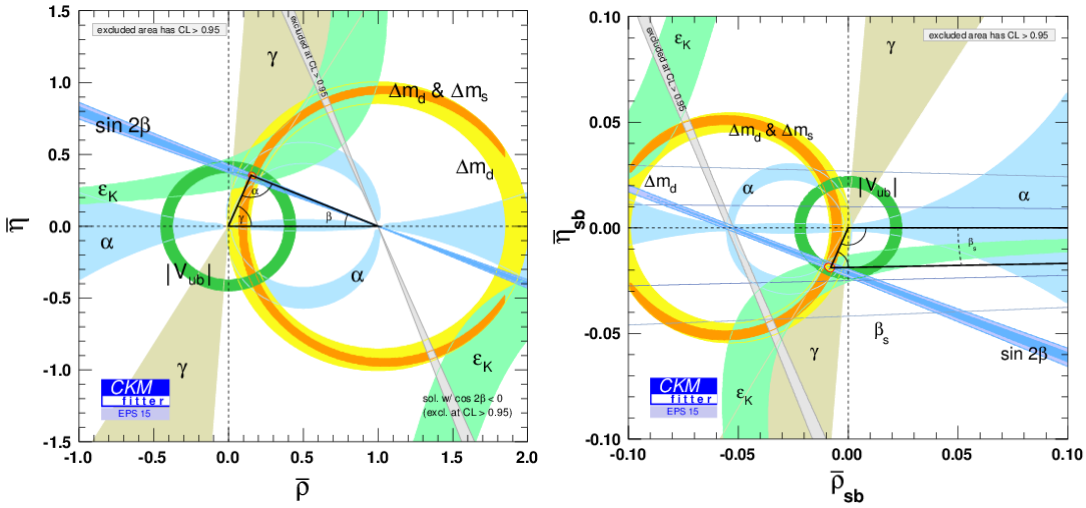


Figure 2.2 *The results of the global SM CKM fit as constraints on the CKM coordinates, $\bar{\rho}$ and $\bar{\eta}$ [8]. The two plots correspond to the unitarity triangles shown in Figure 2.1. The scale has been chosen such that $\bar{\rho}, \bar{\rho}_{sb} = 1$.*

2.2 CP violation in $B_s^0 \to \psi(2S)\phi$

In order for *CP* violation to be observed in a transition there must be more than one interfering amplitude and a difference in phase between the amplitudes. In the case of the decay mode being studied for this thesis, $B_s^0 \to \psi(2S)(\to \mu^+\mu^-)\phi(\to K^+K^-)$, *CP* violation occurs via the interference between B_s^0 mixing and decay, for which the Feynman diagrams are shown in Figure 2.3.

There are two paths available for the B_s^0 mesons to decay into the final state mesons, the $\psi(2S)$ and the ϕ. It can decay directly into the final state particles, or it can first oscillate into its antiparticle (the $\overline{B}_s^0$ meson) via the mixing diagram and then decay to the final state mesons as shown in Figure 2.4. There is an overall phase difference between the two paths, as shown in Equation 2.5, where ϕ_M and ϕ_D are the weak phases due to mixing and decay respectively. The difference in phase is called ϕ_s^{SM} and it is directly related to the CKM angle β_s. The measurement of this phase is the goal of this thesis.

$$\phi_s^{SM} = \phi_M - 2\phi_D = -2\beta_s = -2\arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \tag{2.5}$$

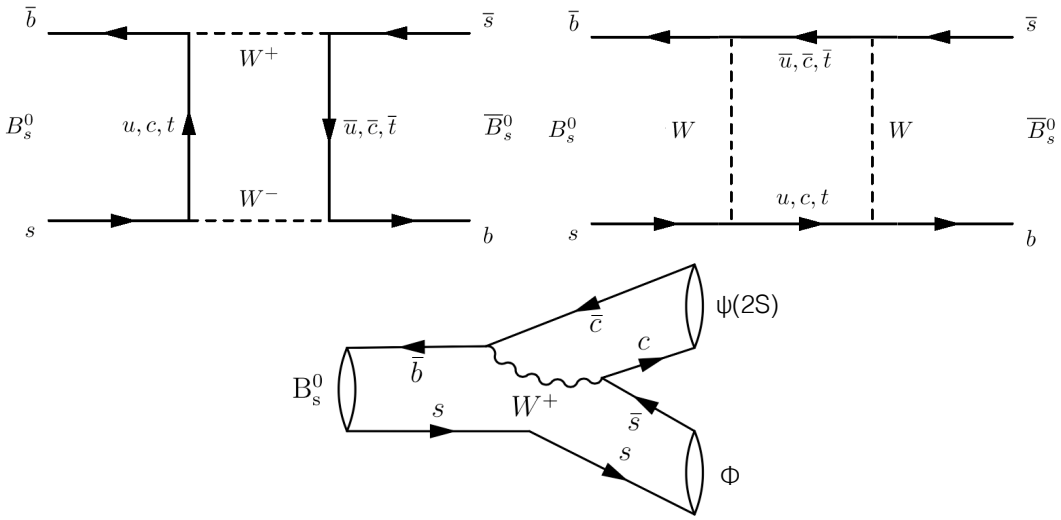


Figure 2.3 *Feynman diagrams for B_s^0-$\overline{B}_s^0$ mixing (top) and $B_s^0 \to \psi(2S)\phi$ decay (bottom).*

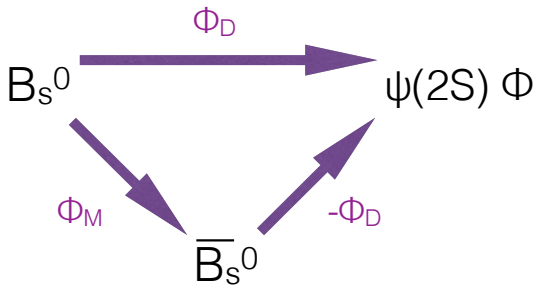


Figure 2.4 *The weak phases of the different paths available for the B_s^0 to decay to the final state, the $\psi(2S)$ and ϕ mesons.*

2.3 Phenomenology

There are a number of components required to fully describe the phenomenology of the $B_s^0 \to \psi(2S)\phi$ decay mode. This section first gives an overview of the time evolution of the B_s^0 meson system, then a description of the differential decay rate for a B_s^0 decay to a CP eigenstate with a single amplitude, and finally extends the derivation to describe $B_s^0 \to \psi(2S)\phi$ decays.

In the B_s^0 system the heavy and light mass eigenstates (B_H and B_L) can be written as linear combination of the weak eigenstates (B_s^0 and $\overline{B}_s^0$) as shown in Equation 2.6 where q and p are complex factors originating from the Hamiltonian

($\mathcal{H}$) of the mass eigenstates.

$$\begin{aligned} |B_L\rangle &= p|B_s^0\rangle + q|\overline{B}_s^0\rangle \\ |B_H\rangle &= p|B_s^0\rangle - q|\overline{B}_s^0\rangle \end{aligned} \tag{2.6}$$

The time evolution of the mass eigenstates can be written as the solution of the Schrödinger equation with eigenvalues $\lambda_{H,L}$ (Equation 2.7). The eigenvalues can be expressed in terms of the masses (m_H, m_L) and decay widths (Γ_H, Γ_L) of the heavy and light mass eigenstates.

$$\begin{aligned} |B_L(t)\rangle &= e^{\lambda_L t}|B_L(0)\rangle = e^{-i(m_L - \frac{i}{2}\Gamma_L)t}|B_L(0)\rangle \\ |B_H(t)\rangle &= e^{\lambda_H t}|B_H(0)\rangle = e^{-i(m_H - \frac{i}{2}\Gamma_H)t}|B_H(0)\rangle \end{aligned} \tag{2.7}$$

By substituting the time evolution of the mass eigenstates into Equation 2.6, and with some rearranging, an expression describing the time evolution of the weak eigenstates is shown in Equation 2.8, where the oscillation amplitudes ($g_\pm(t)$) are defined in Equation 2.9.

$$\begin{aligned} |B_s^0(t)\rangle &= g_+(t)|B_s^0(0)\rangle + \frac{q}{p}g_-(t)|\overline{B}_s^0(0)\rangle \\ |\overline{B}_s^0(t)\rangle &= \frac{p}{q}g_-(t)|B_s^0(0)\rangle + g_+(t)|\overline{B}_s^0(0)\rangle \end{aligned} \tag{2.8}$$

$$g_\pm(t) = \frac{1}{2}(e^{-i(m_L - \frac{i}{2}\Gamma_L)t} \pm e^{-i(m_H - \frac{i}{2}\Gamma_H)t}) \tag{2.9}$$

The expressions found describing the time evolution of the B_s^0 and $\overline{B}_s^0$ mesons can be used to derive expressions for the differential decay rates of B_s^0 decay modes. The following example shows a simple case, relative to $B_s^0 \to \psi(2S)\phi$ decays, where the final state is a *CP* eigenstate accessible to both the B_s^0 meson and its anti-particle with a single amplitude. The decay amplitudes of the B_s^0 and $\overline{B}_s^0$ mesons decaying to a *CP* eigenstate f are described as:

$$\begin{aligned} A_f(t) &= \langle f|\mathcal{H}|B_s^0(t)\rangle,\ A_f = \langle f|\mathcal{H}|B_s^0(0)\rangle \\ \bar{A}_f(t) &= \langle f|\mathcal{H}|\bar{B}_s^0(t)\rangle,\ \bar{A}_f = \langle f|\mathcal{H}|\bar{B}_s^0(0)\rangle \end{aligned} \tag{2.10}$$

To obtain an expression for the differential decay rate, shown here for the B_s^0 meson, the amplitude squared is calculated and shown in Equation 2.11, where the parameter λ_f is defined as $\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}$ [3]. λ_f is related to the *CP* violating parameter introduced in the previous section by $\phi_s = -\arg(\eta_f\lambda_f)$, where η_f is the

CP eigenvalue (± 1) of the final state.

$$\begin{aligned}\frac{d\Gamma_{B_s^0\to f}(t)}{dt} &= |A_f(t)|^2 = |\langle f|B_s^0(t)\rangle|^2 \\ &= |A_f|^2(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + 2\mathcal{R}e[\lambda_f g_+^*(t)g_-(t)])\end{aligned} \tag{2.11}$$

Finally, substituting the expressions for the oscillation amplitudes $g_\pm(t)$ the differential decay rate is shown in Equation 2.13 where:

$$S = -\frac{2|\lambda_f|}{1+|\lambda_f|^2}\mathcal{I}m(\lambda_f), D = -\frac{2|\lambda_f|}{1+|\lambda_f|^2}\mathcal{R}e(\lambda_f) \text{ and } C = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2} \tag{2.12}$$

The following definitions were used to simplify the expression: $\Gamma_s = \frac{1}{2}(\Gamma_H + \Gamma_L)$, $\Delta\Gamma_s = \Gamma_L - \Gamma_H$ and $\Delta m_s = m_H - m_L$.

$$\frac{d\Gamma_{B_s^0(t)\to f}}{dt} = e^{-\Gamma_s t}\left[\cosh\left(\tfrac{1}{2}\Delta\Gamma_s t\right) + D\sinh\left(\tfrac{1}{2}\Delta\Gamma_s t\right) + C\cos(\Delta m_s t) + S\sin(\Delta m_s t)\right] \tag{2.13}$$

To extend the simple case of a B_s^0 meson decaying to a CP eigenstate with a single amplitude to the $B_s^0 \to \psi(2S)\phi$ decay mode the properties of the $\psi(2S)\phi$ final state must be considered. The $B_s^0 \to \psi(2S)\phi$ decay mode is a pseudo-scalar (spin 0) to two vector (spin 1) decay mode. The total angular momentum (J), the sum of the orbital angular momentum (L) and the spin (S), of the initial B_s^0 meson is zero and conservation of momentum dictates that the total momentum of the final state must also be zero. This constrains the relative orbital angular momentum of the $\psi(2S)$ and ϕ to L $\in 0, 1, 2$. The total decay amplitude can therefore be split into three independent amplitudes which are expressed here using the polarisation amplitudes, A_0, $A_{||}$ and $A_\perp$. The amplitudes correspond to the different polarisation configurations of the $\psi(2S)$ and ϕ mesons (Figure 2.5) where the A_0 and $A_{||}$ amplitudes are CP even, and the $A_\perp$ contribution is CP odd. In addition there is one contributing S-wave amplitude (A_s) from the non-resonant K^+K^- component, to give a total of four independent amplitudes.

By examining Equation 2.13, and considering that the parameter S multiplying the oscillation term $\sin(\Delta m_s t)$ will change sign under the action of the CP operators, it can be seen that the observed amplitude of the oscillations will be reduced depending of the relative size of the CP-even and CP-odd amplitudes. To address this, the relative orientations of the four daughter particles, parameterised by three angles (θ_K, θ_μ and ϕ), can be used to separate the contributing amplitudes. For this analysis the helicity basis is used (Figure 2.6). The differential decay rate can be expressed as a sum of ten terms (due to the four squared amplitudes

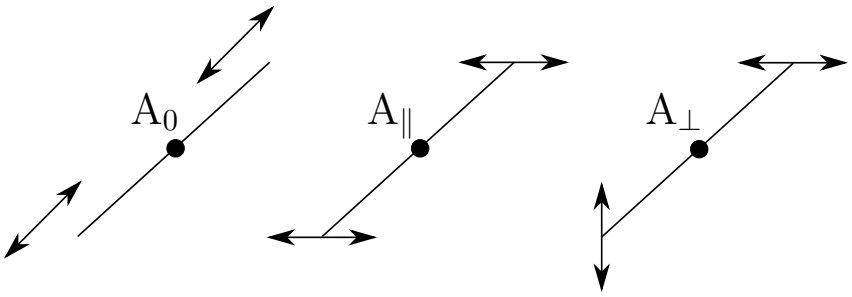


Figure 2.5 *The polarisation orientations for the three P-wave amplitudes of the $B_s^0 \to \psi(2S)\phi$ decay mode. The large arrows originating from the centre show the momentum vectors of the $\psi(2S)$ and ϕ in the B_s^0 rest frame and the small arrows show the possible orientations of the polarisation vectors.*

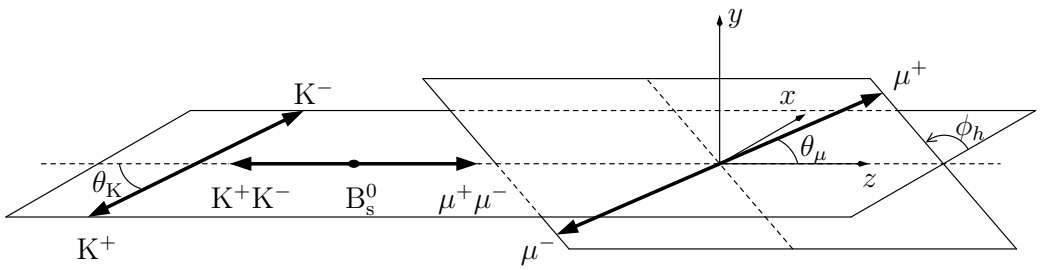


Figure 2.6 *The definition of the helicity basis where θ_μ is the angle between μ^+ and the z-axis in the $\psi(2S)$ rest frame, θ_K is the angle between the K^+ and the z-axis in the ϕ rest frame and ϕ is the relative orientation of the decay planes of the two daughter mesons. For the helicity basis the z-axis is defined as the direction of the $\psi(2S)$ meson in the B_s^0 rest frame.*

and their interference terms), dependent on the helicity angles and decay time (Equation 2.14), where $h_k(t)$ is defined in Equation 2.15 and $f_k(\Omega)$ is defined in Table 2.1.

$$\frac{\mathrm{d}^4\Gamma(B_s^0 \to \psi(2S)\phi)}{\mathrm{d}t\ \mathrm{d}\Omega} \ \propto \ \sum_{k=1}^{10} h_k(t)\ f_k(\Omega) \tag{2.14}$$

$$h_k(t) \ = \ N_k e^{-\Gamma_s t}\left[a_k \cosh\left(\tfrac{1}{2}\Delta\Gamma_s t\right) + b_k \sinh\left(\tfrac{1}{2}\Delta\Gamma_s t\right) + c_k \cos(\Delta m_s t) \ + d_k \sin(\Delta m_s t)\right], \tag{2.15}$$

The parameters $N_k,\ a_k, b_k, c_k, d_k$ are defined in Table 2.1 and depend on the polarisation amplitudes and their strong phases $(\delta_0, \delta_\perp, \delta_\parallel, \delta_\mathrm{S})$ as well as the *CP* violating parameter ϕ_s. The following conventions are adopted; $\delta_0 = 0$, $|A_\perp|^2 + |A_0|^2 + |A_\parallel|^2 = 1$ and the fraction of the S-wave amplitude $f_S = |A_s|^2/(|A_s|^2 +$

$|A_\perp|^2 + |A_0|^2 + |A_\parallel|^2) = |A_s|^2/(|A_s|^2 + 1)$. For a more complete derivation of the differential decay rate, and for the equivalent expression for $\overline{B}{}^0_s \to \psi(2S)\phi$ decays see Appendix A.

Table 2.1 *Summary of terms contributing to the differential decay rate of* $B_s^0 \to \psi(2S)\phi$.

k	$f_k(\theta_\mu, \theta_K, \phi_h)$	N_k	a_k	b_k	c_k	d_k
1	$2\cos^2\theta_K \sin^2\theta_\mu$	$\lvert A_0(0)\rvert^2$	1	D	C	$-S$
2	$\sin^2\theta_K\left(1-\sin^2\theta_\mu\cos^2\phi_h\right)$	$\lvert A_\parallel(0)\rvert^2$	1	D	C	$-S$
3	$\sin^2\theta_K\left(1-\sin^2\theta_\mu\sin^2\phi_h\right)$	$\lvert A_\perp(0)\rvert^2$	1	$-D$	C	S
4	$\sin^2\theta_K \sin^2\theta_\mu \sin 2\phi_h$	$\lvert A_\parallel(0)A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_\parallel)$	$S\cos(\delta_\perp-\delta_\parallel)$	$\sin(\delta_\perp-\delta_\parallel)$	$D\cos(\delta_\perp-\delta_\parallel)$
5	$\frac{1}{2}\sqrt{2}\sin 2\theta_K \sin 2\theta_\mu \cos\phi_h$	$\lvert A_0(0)A_\parallel(0)\rvert$	$\cos(\delta_\parallel-\delta_0)$	$D\cos(\delta_\parallel-\delta_0)$	$C\cos(\delta_\parallel-\delta_0)$	$-S\cos(\delta_\parallel-\delta_0)$
6	$-\frac{1}{2}\sqrt{2}\sin 2\theta_K \sin 2\theta_\mu \sin\phi_h$	$\lvert A_0(0)A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_0)$	$S\cos(\delta_\perp-\delta_0)$	$\sin(\delta_\perp-\delta_0)$	$D\cos(\delta_\perp-\delta_0)$
7	$\frac{2}{3}\sin^2\theta_\mu$	$\lvert A_s(0)\rvert^2$	1	$-D$	C	S
8	$\frac{1}{3}\sqrt{6}\sin\theta_K \sin 2\theta_\mu \cos\phi_h$	$\lvert A_s(0)A_\parallel(0)\rvert$	$C\cos(\delta_\parallel-\delta_\mathrm{S})$	$S\sin(\delta_\parallel-\delta_\mathrm{S})$	$\cos(\delta_\parallel-\delta_\mathrm{S})$	$D\sin(\delta_\parallel-\delta_\mathrm{S})$
9	$-\frac{1}{3}\sqrt{6}\sin\theta_K \sin 2\theta_\mu \sin\phi_h$	$\lvert A_s(0)A_\perp(0)\rvert$	$\sin(\delta_\perp-\delta_\mathrm{S})$	$-D\sin(\delta_\perp-\delta_\mathrm{S})$	$C\sin(\delta_\perp-\delta_\mathrm{S})$	$S\sin(\delta_\perp-\delta_\mathrm{S})$
10	$\frac{4}{3}\sqrt{3}\cos\theta_K \sin^2\theta_\mu$	$\lvert A_s(0)A_0(0)\rvert$	$C\cos(\delta_0-\delta_\mathrm{S})$	$S\sin(\delta_0-\delta_\mathrm{S})$	$\cos(\delta_0-\ \delta_\mathrm{S})$	$D\sin(\delta_0-\delta_\mathrm{S})$

2.4 Global Status of ϕ_s Measurements

Measurements of the *CP* violating parameter ϕ_s have been made by multiple experiments and in more than one decay channel. The channel that contributes the most sensitivity to ϕ_s is $B_s^0 \to J/\psi\,\phi$ and the CDF, D0, ATLAS, CMS and LHCb experiments have reported measurements of ϕ_s using this channel [9–13]. In addition, ϕ_s can be measured in the related modes $B_s^0 \to J/\psi\pi^+\pi^-$ and $B_s^0 \to D_s^+ D_s^-$, results for which have been reported by the LHCb experiment [14, 15]. The Heavy Flavour Averaging Group (HFAG) reports the result of ϕ_s from the combination of all these measurements as $\phi_s = -0.034 \pm 0.033$ rad [16]. The prediction, taken as the indirect determination of ϕ_s via a global fit to experimental data within the Standard Model, is $\phi_s = -0.0363^{+0.0012}_{-0.0014}$ rad [17]. The combined measurements agree with the SM prediction.

In addition, the $B_s^0 \to \psi(2S)\phi$ mode can be used to measure the difference in the lifetime of the heavy and light mass eigenstates, $\Delta\Gamma_s$. The combination as reported by HFAG of the measurements reported by ATLAS, CMS, CDF, D0 and LHCb using the $B_s^0 \to J/\psi\,\phi$ mode is $\Delta\Gamma_s = 0.082 \pm 0.006$ ps^{-1} [16]. The SM prediction for this parameter is $\Delta\Gamma_s = 0.087 \pm 0.021$ ps^{-1} [18].

Figure 2.7 shows the 2D plot of Γ_s and $\Delta\Gamma_s$ to 68% confidence level (CL) for the bands and 39% confidence level for the 2D enclosed region, with contributions from the $B_s^0 \to J/\psi\,\phi$ mode in red, and contraints from $B_s^0 \to J/\psi f_0$, $B_s^0 \to K^+K^-$ and B_s^0 to flavour specific states are shown as the green, purple and blue bands, respectively [16]. The average is shown as the grey ellipse.

Figure 2.8 shows the ϕ_s vs $\Delta\Gamma_s$ plane showing the 68% confidence levels from the ATLAS, CMS, CDF, D0 and LHCb experiments [16]. Also shown is the combination of all the measurements (grey circle) and the SM prediction as the black rectangle.

For both ϕ_s and $\Delta\Gamma_s$ the experimentally measured values agree with the SM predictions. The next goal for measurements of ϕ_s is to achieve a precision with which any small deviation from the predicted value would be observed. To do this with the data available measurements of ϕ_s can be made with new decay modes, such as $B_s^0 \to \psi(2S)\phi$, and included in the combinations discussed above.

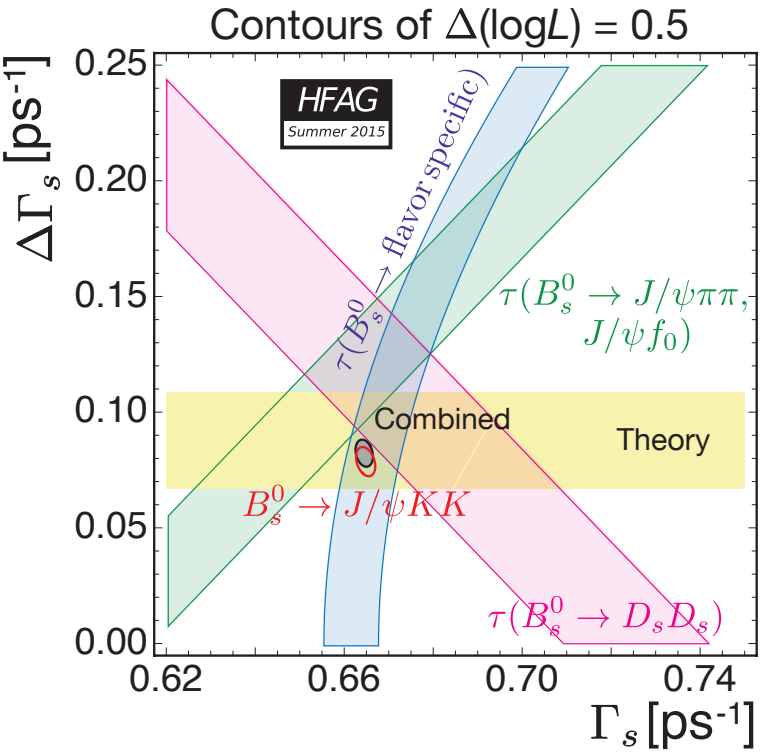


Figure 2.7 *The Γ_s vs $\Delta\Gamma_s$ plane with results included from the $B_s^0 \to J/\psi\phi$ mode in red (39% CL), and constraints from $B_s^0 \to J/\psi f_0$, $B_s^0 \to K^+K^-$ and B_s^0 to flavour specific states shown as the green, purple and blue bands, respectively (68% CL). The average is shown as the grey ellipse [16].*

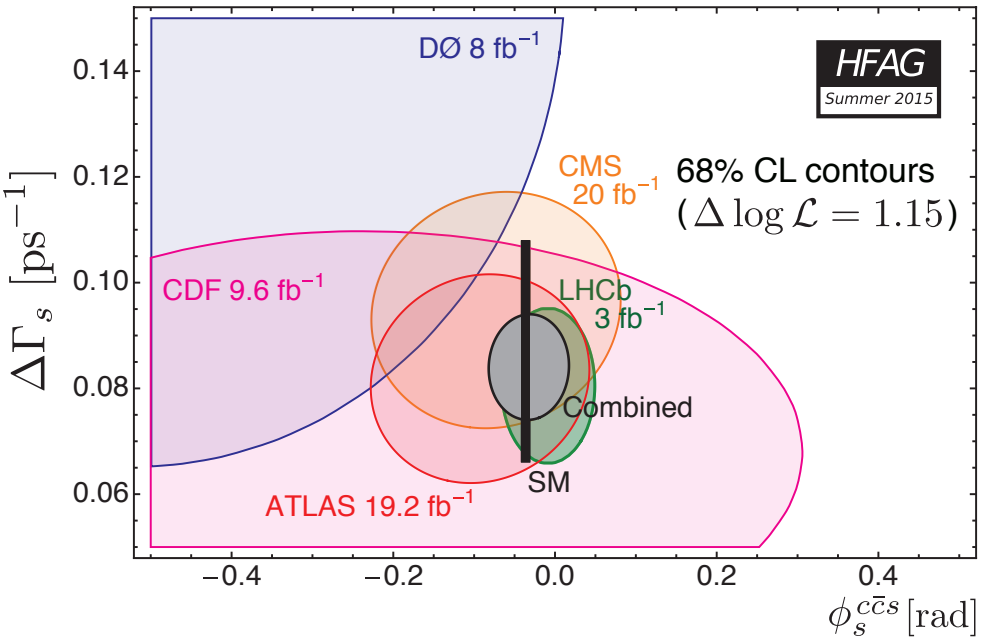


Figure 2.8 *The ϕ_s vs $\Delta\Gamma_s$ plane with the results from the ATLAS, CMS, CDF, D0 and LHCb experiments. The combination is shown in grey, and the SM prediction is shown as the black rectangle [16].*

Chapter 3

The LHCb Experiment

3.1 The Large Hadron Collider (LHC)

The LHC is a circular particle accelerator built ~ 100 m underground over the French-Swiss border (Figure 3.1). The accelerator has a circumference of 27 km and is constructed using superconducting magnets and RF cavities designed to collide bunches of protons with a centre-of-mass energy of 14 TeV. The protons, created by stripping the electron from atomic Hydrogen, pass through multiple components of the machine before they have been accelerated to design energy. The sequence of these components is as follows: LINear ACcelerator (LINAC) 2 (50 MeV), Proton Synchotron (PS) Booster (1.4 GeV), PS (25 GeV), Super PS (450 GeV) and finally the LHC ring itself (7 TeV). To create collisions in the four detectors based around the LHC ring, bunches of protons are circulated in opposite directions around the ring in separate beam pipes and collided head on in the designated interaction points of the detectors.

3.2 The LHCb Detector

The LHCb detector was built to take advantage of the large number of $b\bar{b}$ pairs created at the LHC with the aim of making measurements of *CP* violating parameters and branching fractions of rare decays of beauty and charm hadrons in search of new physics. To contribute to this field the LHCb detector must have the following properties: good vertex and momentum resolution (key to measuring

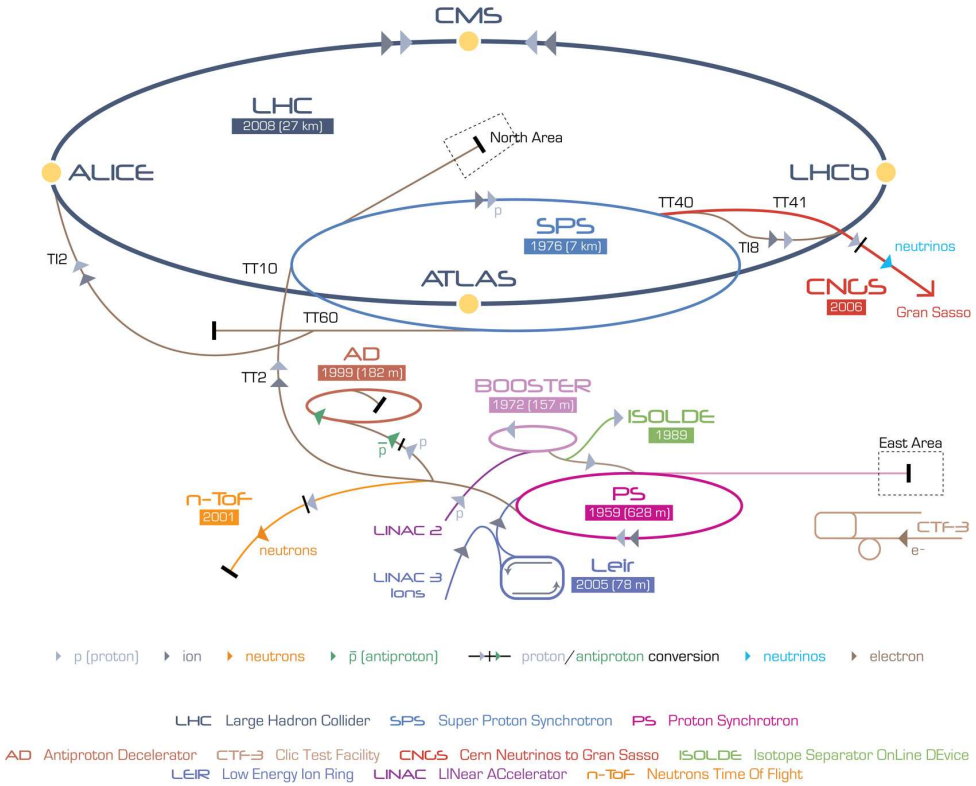


Figure 3.1 *The accelerator machines at CERN [19]*

the fast B_s^0 oscillation rate and reducing combinatorial background), good particle identification, a trigger capable of dealing with the harsh hadronic environment, and the data processing and storage facilities to keep large amounts of data for offline analysis. This chapter introduces the detector components used to fulfil these criteria.

The LHCb detector is a forward arm spectrometer (Figure 3.2) constructed at one of the interaction points of the LHC, and has a forward angular coverage of 300 (250) mrad in the bending (non-bending) plane. This coverage was chosen because at high energies the majority of the b-hadrons created travel along the direction of the beampipe. The components of the detector include; the magnet, the vertex locator (VELO), the tracking stations, Ring Imaging Cherenkov detectors (RICH), electromagnetic and hadronic calorimeters and the muon detection system. The following subsections focus on the detector components and how they are used to fulfil the required properties of the detector.

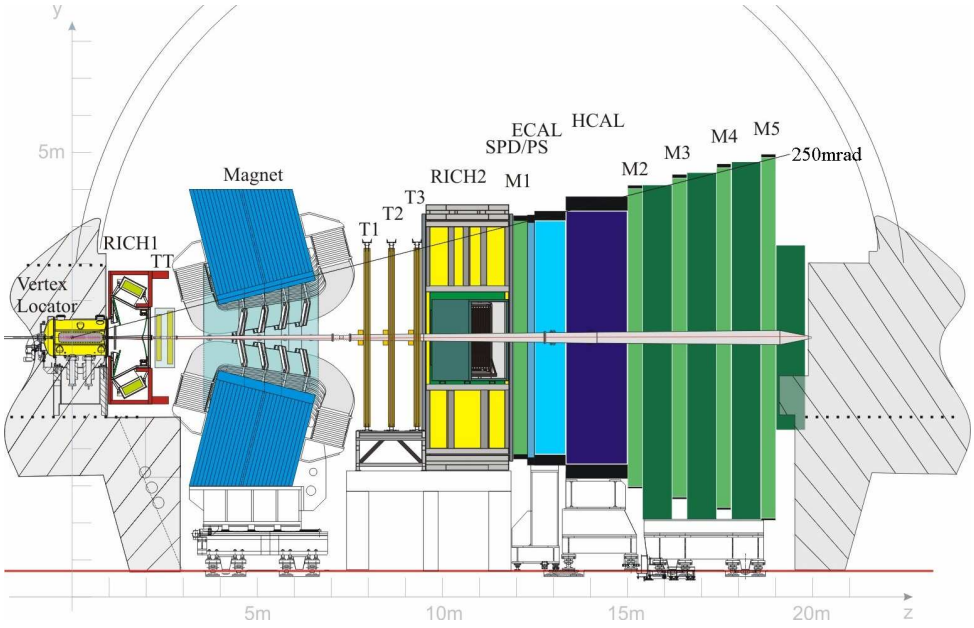


Figure 3.2 *The different components and their locations in the LHCb detector [20].*

3.3 Tracking

3.3.1 Magnet

A magnetic field is necessary to be able to make track momentum measurements and to identify oppositely charged particles. The LHCb magnet is a water cooled 4 Tm (for tracks of 10m length) dipole magnet located between the first two tracking stations (the Tracker Turicensis and T1) of the LHCb detector. A drawing of the magnet is shown in Figure 3.3 (left) while the field map is shown in Figure 3.3 (right).

3.3.2 The Vertex Locator (VELO)

The primary purpose of the VELO is to identify primary and secondary vertices by making precise track coordinate measurements. This is important for many decay processes but especially for this analysis as the B_s^0 meson has a relatively long lifetime of 1.5 ps and can travel $\sim$ 1cm in the detector before decaying into the daughter particles. This means that the tracks of the daughter particles in the VELO will converge at some point displaced from the primary interaction

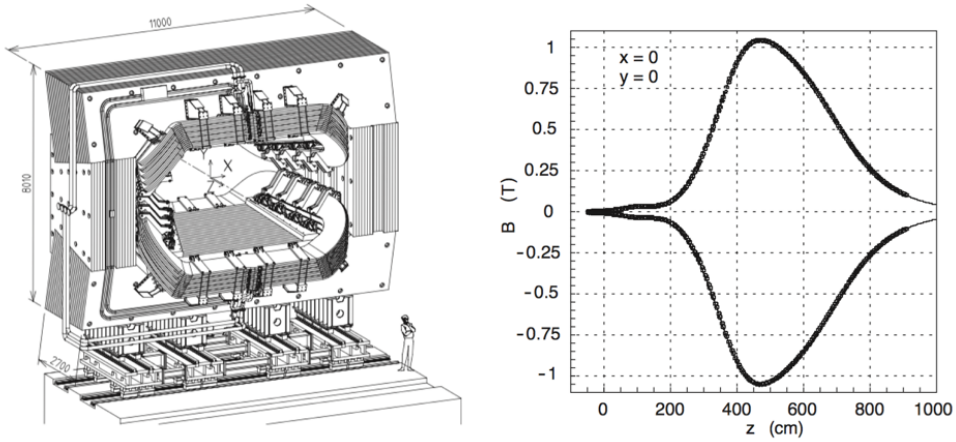


Figure 3.3 *The LHCb magnet (left) and field map along the direction of the beampipe (right) [20].*

point and the measurement of this displacement is vital in calculating the decay time of the B_s^0 meson and in suppressing prompt backgrounds.

In order to make these measurements the VELO is comprised of two sets of 21 semi-circular silicon modules that are placed on either side of the beam and can be moved mechanically to create a halo around the proton beam in the region around the interaction point (Figure 3.4). These silicon modules are placed at intervals along the direction of the beam and provide track r and ϕ measurements which can be used to reconstruct the path of charged particles through the VELO. The spatial resolution of the VELO depends on the strip pitch and the projected angle of the track, the angle between the track and the perpendicular to the strip. The best spatial resolution of the VELO is 4μm [21].

3.3.3 Tracker Turicensis (TT) and Inner Tracker (IT)

Both the TT and the IT (the inner sections of the three tracking stations T1, T2 and T3) use silicon strip detectors to record the position of charged particles as they pass through the detector. The TT (Figure 3.5 (left)) is located between RICH1 and the magnet and covers the full detector acceptance. The IT (Figure 3.5 (right)) comprises of the inner sections of the three tracking stations (T1, T2 and T3) located downstream from the magnet and in front of RICH2 (the inner purple sections of the tracking stations in Figure 3.6). Silicon sensors were chosen for these components as they have a good spatial resolution (50μm) and high signal efficiency while being able to cope with the high occupancy and irradiation

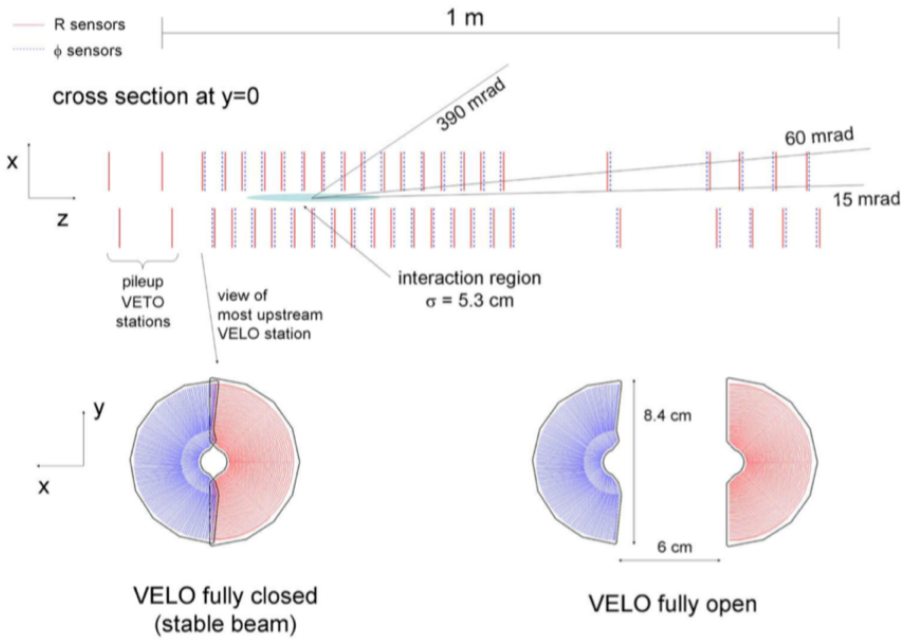


Figure 3.4 *The placement of VELO sensors along the direction of the beampipe (top), and cross-section of the VELO sensors in closed (bottom left) and open (bottom right) positions [20].*

expected close to the beampipe. The orientation and number of strips used to construct the TT and IT was chosen to minimise costly readout channels while maintaining the requirements of the trackers.

3.3.4 Outer Tracker (OT)

The OT consists of the outer sections of the three tracking stations located downstream from the magnet (Figure 3.6). These sections are constructed using panels of drift-time detectors. Each panel has two staggered layers of gas-filled straw tubes (70% Argon and 30% Carbon Dioxide) that are used to record the coordinates of charged particle tracks as they pass through the detector. These modules have a drift time of less than 50 ns and a spatial resolution of 0.2 mm.

3.4 Particle Identification (PID)

PID is a strong focus for LHCb as good pion-kaon separation is important for many key results including the topic of this thesis. The components that are responsible

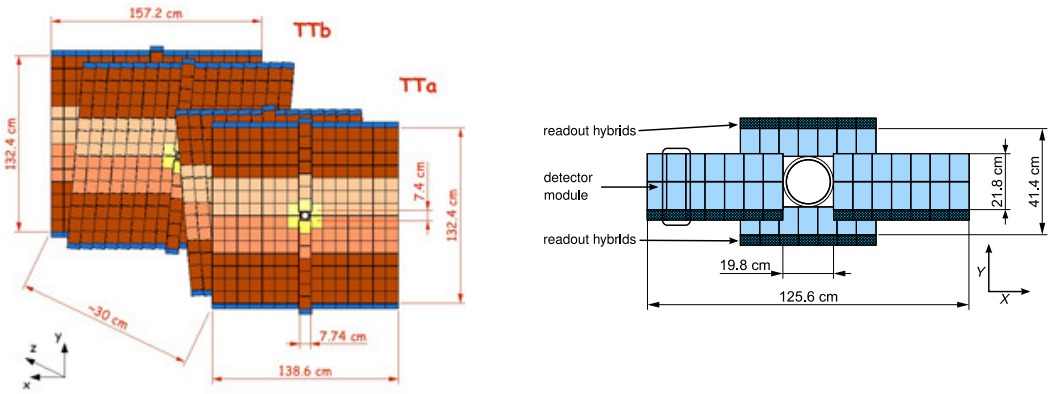


Figure 3.5 *Layout of silicon sensors of the TT (left) and IT (right) [20].*

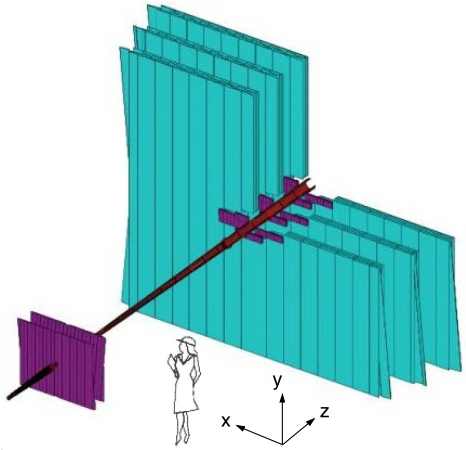


Figure 3.6 *Layout of the tracking stations; the purple sections are the TT (upstream) and IT (inner components of downstream tracking stations) and the blue sections are the OT of the downstream tracking stations [20].*

for this task are the two RICH (Ring Imaging CHerenkov) detectors (RICH1 and RICH2). These detectors use mirrors and Hybrid Photon Detectors (HPDs) to reflect and detect Cherenkov radiation, the light emitted by charged particles travelling faster than the speed of light in a medium. Cherenkov radiation is seen as cones of emitted light, which is reflected from the RICH spherical mirrors to form circles on the HPD detectors, the radii of which are related to the velocities of the charged particles. This information, used in combination with tracking information from other detector components, can be used to calculate the particle's mass and so its identity.

Due to the large momentum range of particles produced, and the large acceptance range, two RICH detectors are installed in the LHCb detector; RICH1 which is

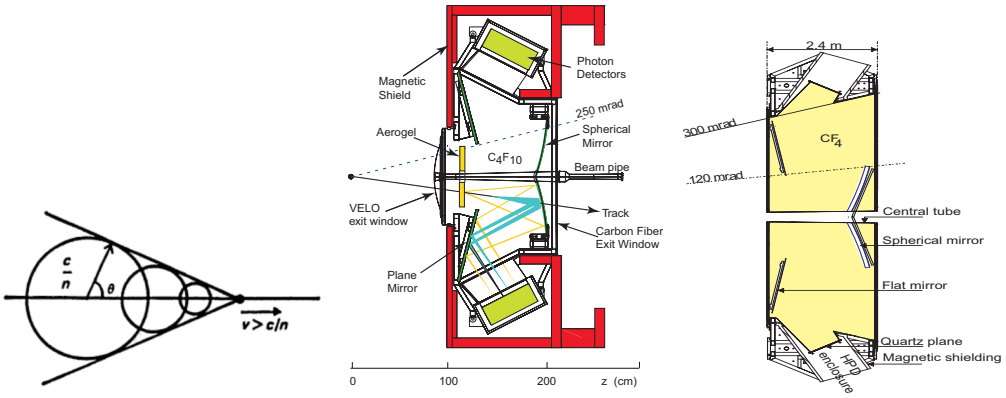


Figure 3.7 *(a) Huygens' construction to determine the direction of the propagation of the wavefront of Cherenkov radiation, where the surfaces show points of equal phase [23], (b) schematic of RICH1 [20], (c) top view of RICH2 [24].*

located upstream from the TT and RICH2 located downstream from the outer tracking stations (T1-T3). RICH1 covers the low momentum range (2 - 40 GeV/c) and total detector acceptance, while RICH2 covers the high momentum range (15 - 100 GeV/c) over the detector acceptance in the region where these particles are produced. To cover the different momentum ranges the RICH detectors have different types of radiator; C_4F_{10} gas and aerogel for RICH1, and CF_4 gas for RICH2. Finally, both RICH detectors have iron shielding to reduce the magnetic field in the HPD volume [22].

3.5 Calorimeters

There are four components to the calorimeter section of LHCb; the Scintillating Pad Detector (SPD), the Pre-Shower (PS) detector, the Electromagnetic Calorimeter (ECal) and the Hadronic Calorimeter (HCal), the purpose of which are to identify electrons, photons and hadrons and measure their energy and position. The principle behind all four detectors is the same: scintillation light is transmitted to Photo-Multiplier Tubes (PMT) by wavelength-shifting (WLS) fibres.

The first calorimeter component, located downstream from RICH2 and the first muon station, is the SPD which is used by the L0 electron trigger to reject π^0s with high transverse energy (E_{T}). A lead converter, with a depth of $2.5X_0$, separates the SPD from the downstream PS detector which is used to identify photons, electrons and charged pions. The structure of both the SPD and the PS

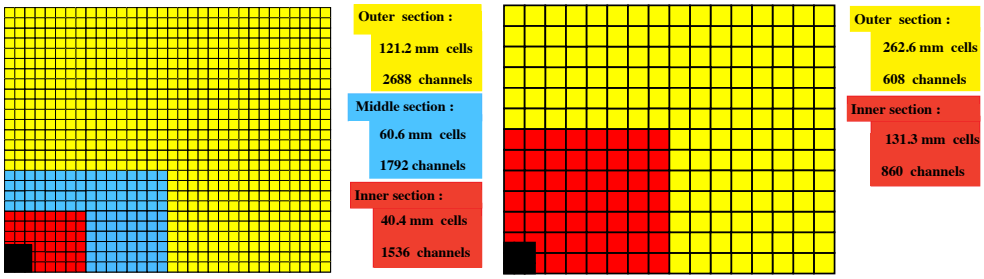


Figure 3.8 *The transverse segmentation of the SPD/PS and ECal (left) and HCal (right). The diagram shows one quarter of the total detector acceptance. The dimensions given for the left side figure are for the ECal [20].*

is similar, using scintillator pads that are read out by WLS fibers and recorded using multi-anode Photo-Multiplier Tubes (MAPMTs). To compensate for the large variation in hit density across the detector acceptance the segmentation of the SPD is variable and shown in Figure 3.8 (left).

The ECal is located downstream from the PS and is constructed using 'Shashlik' technology, which is a sampling calorimeter of scintillator tiles sandwiched between lead absorbers (Figure 3.9 (b)). The total length of the ECal is $25X_0$ to ensure no spillover to the HCal from high energy photons. The lateral segmentation of the ECal matches that of the SPD and PS (Figure 3.8 (left)).

The final calorimeter component is the HCal which is located downstream from the ECal. The HCal is comprised of scintillating tiles separated by iron plates that are read out by WLS fibres to PMTs that are located downstream from the scintillating components of the HCal, so inside the detector acceptance (Figure 3.9 (a)). The orientation of the tiles is along the direction of the beampipe, which differs from the perpendicular orientation of the tiles for the ECal. In the longitudinal direction the Iron spacers have a length equal to the hadronic interaction length (λ_I) in steel. The transverse segmentation of the HCal is shown in Figure 3.8 (right).

3.6 Muon System

The final component of the detector to discuss is the Muon system which consists of five separate tracking stations, the first of which is located downstream from RICH2, before the SPD, and the other four stations (separated by Iron absorbers $20\lambda_I$ in length) are located downstream from the HCal and are the last components

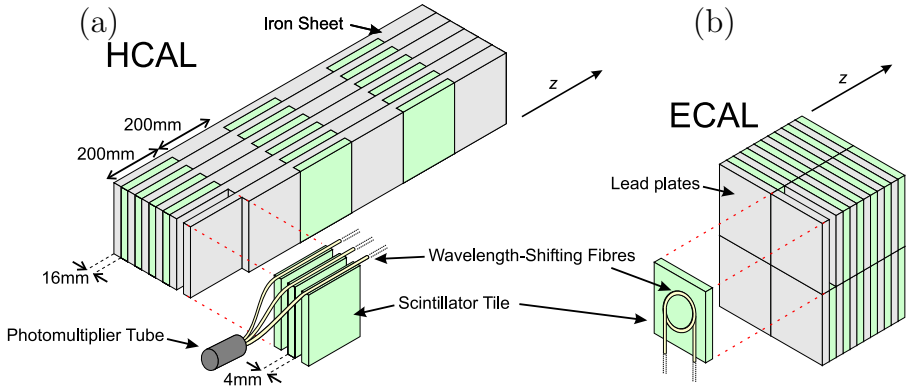


Figure 3.9 *The composition of the HCal (A) and the Ecal (b) [20].*

of the detector (Figure 3.10 (left)). The transverse segmentation of one quadrant of the Muon system is shown in Figure 3.10 (middle, right). The detection components of the Muon system are comprised of Multi-wire proportional chambers (MWPC), which rely on collecting the signal produced when charged particles pass through the chambers and ionise the gas (45% Ar, 15% CO_2, 40% CF_4) within. The inner section of M1 (M1R1) is constructed using triple-GEM detectors [25] to satisfy the ageing requirements of the LHCb detector. These detectors use thin perforated foils in a gas container as multipliers for the electrons produced from ionisation by a charged particle.

3.7 Trigger

The purpose of the LHCb trigger is to reduce the event rate to a level that can be dealt with by the offline computing facilities while attempting to preserve only the events that can be used to measure physics parameters of interest to the LHCb collaboration. This reduction in the event rate is completed in a multi-stage process (Figure 3.11) that makes use of both hardware and software filters and is briefly described in the following paragraph.

The structure of the LHCb trigger is as follows; the first level trigger (L0) is a hardware trigger that uses information from the calorimeter and muon systems to reduce the event rate from 40 MHz to 1 MHz. At this rate the information from the whole detector can be read out to a computer farm where the High Level Trigger (HLT) can be applied. The HLT is a software trigger responsible for reducing the event rate from 1 MHz to 5 kHz. The HLT is split into two

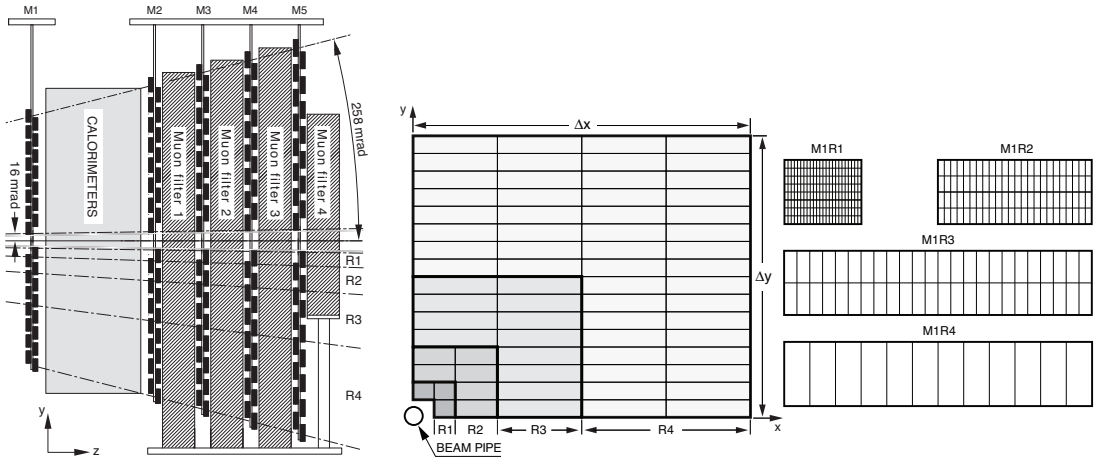


Figure 3.10 *Position of the 5 Muon tracking stations within the LHCb detector (left) and transverse segmentation of one quarter the Muon tracking stations (middle). Each rectangle from the middle image represents a single chamber. The division into logical pads for M1 is shown in the diagram to the right. For M2-M3 (M4-M5) the number of pad columns per chamber is double (half) the number for M1 [20].*

stages; HLT1 and HLT2. The purpose of HLT1 is to confirm the L0 trigger using fully reconstructed VELO tracks and information from the other tracking systems. This reduces the event rate to a level where full track reconstruction can be performed on the remaining events (30 kHz). The HLT2 trigger uses track and momentum information to form composite particles which can be used in combination with information on the impact parameter to construct candidate decay chains of particles and to reduce the event rate to 5 kHz. The events that pass these triggers are stored offline to be available for analysis.

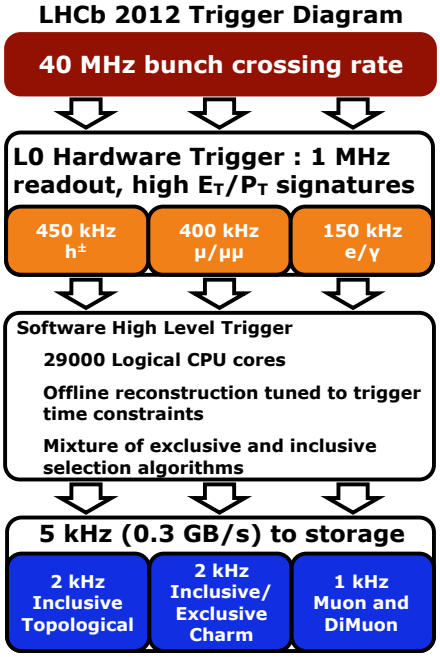


Figure 3.11 *The flow of the LHCb trigger.*

Chapter 4

Data Samples and Selection

The subject of this thesis is the measurement of the *CP* violating parameter ϕ_s in the $B_s^0 \to \psi(2S)\phi$ decay mode, where the $\psi(2S)$ decays to two muons (μ^+, μ^-) and the ϕ decays to two kaons (K^+, K^-). This and the following chapters contain the details of how the analysis was performed, including the criteria for isolating signal events, modelling the detector effects and the fit method.

Once the data has been triggered and stored offline there are further processing steps required before it is available for analysis. The first is the event reconstruction which uses information from the detector and trigger system to perform steps including forming vertices and decay chains of particles, particle identification and flavour tagging. The next stage is called the data 'stripping'. This involves identifying the categories, or 'lines', that the event might belong to based on the type of particles present or decay processes that have occurred, for example, events that contain a decay chain of a J/ψ or a $\psi(2S)$ to two muons will be identified as part of the dimuon lines. Once the stripping has been applied to the data, it is saved as a DST (Data Summary Tape) file and is available for analysis.

4.1 Simulated Data Samples

Simulated data samples were used in this analysis to model acceptance effects, possible backgrounds and detector resolution. This section introduces the samples that were used and how they were generated.

Table 4.1 *Summary of simulated samples used in the analysis.*

Collision Energy	Decay Type	Number of generated events
7 TeV	$B_s^0 \to \psi(2S)\phi$	$2,193,179$
8 TeV	$B_s^0 \to \psi(2S)\phi$	$2,040,098$
7 TeV	$B^0 \to \psi(2S)K^{*0}$	$1,000,000$
8 TeV	$B^0 \to \psi(2S)K^{*0}$	$1,000,000$
8 TeV	$\Lambda_b^0 \to \psi(2S)pK^-$	$4,000,000$

Table 4.2 *Physics parameters used to generate the simulated $B_s^0 \to \psi(2S)\phi$ samples.*

Parameter	Value
Γ_s $[\mathrm{ps}^{-1}]$	0.6654
$\Delta\Gamma_s$ $[\mathrm{ps}^{-1}]$	0.0917
ϕ_s [rad]	-0.04
$A_\perp^2$	0.16
A_0^2	0.601
$\delta_\parallel$[rad]	2.5
$\delta_\perp$[rad]	-0.17
$\Delta m_s[\mathrm{ps}^{-1}]$	17.8

Large datasets of simulated events were generated for the signal mode, $B_s^0 \to \psi(2S)\phi$, with full detector and reconstruction effects. Events were generated separately at a centre of mass energy of 7 TeV and 8 TeV, representative of data taken in 2011 and 2012, and with equal contributions from "magnet up" and "magnet down" to reflect the fact that the polarity of the LHCb magnet is periodically reversed during data taking. The same stripping and trigger configuration is performed on each simulated data sample as was implemented at the time of data taking, or the conditions for which the majority of data was taken. The number of events generated for each sample are shown in Table 4.1 and the physics parameters used to generate the signal samples are shown in Table 4.2.

In addition, fully simulated data samples were created for the $B^0 \to \psi(2S)K^{*0}$ and $\Lambda_b^0 \to \psi(2S)pK^-$ decay modes. These channels were created with the same conditions as the simulated signal sample. Both data samples were used initially to aid in checks for background events from these modes in the final data sample after selections. In addition the $B^0 \to \psi(2S)K^{*0}$ sample is used to check for differences in the decay time acceptance between the signal $B_s^0 \to \psi(2S)\phi$ mode and the channel used to model the acceptance, $B^0 \to \psi(2S)K^{*0}$.

4.1.1 PID resampling

Two types of Particle IDentification (PID) variables were used in the selection process; the DLL and ProbNN variables. The DLL variables are constructed using information from the RICH, muon and calorimeter systems to calculate the likelihood of a particle having a certain identity. Final variables are calculated by taking the difference between the log of the likelihood of the track being a particle of interest and log likelihood of it being a pion.

$$DLL_{X\pi} = \ln(\mathcal{L}(X)) - \ln(\mathcal{L}(\pi)) \tag{4.1}$$

The ProbNN variables result from combining information from the three PID systems (RICH, muon and calorimeter systems) and the tracking detectors to provide input for a neural network trained on simulated inclusive B events. This approach has the advantage that correlations between the systems are taken into account and can often result in better discrimination.

A known feature of simulated data is that the PID variables do not have the same distribution as in data, caused by the lower detector occupancy in simulation and changing operating conditions in data. This difference in the observed distributions can cause problems, particularly due to the inclusion of the ProbNN variables in the multivariate analysis (described in § 4.2.3) used to perform the final selection for this analysis. To correct for this difference a process called resampling was implemented using the PIDCalib tool.

The first step of the resampling process was to construct histograms of the PID distributions for the track types of interest (muons and kaons) in bins of the momentum of the particle, the pseudorapidity η (angle of the particle relative to the beam axis) and the number of tracks from the event in the RICH. The binning scheme used for each variable is as follows; for both the kaons and muons there are four even bins in η from $[1.5, 5]$, for both the kaons and muons the number of tracks covers the range $[0, 500]$ in four uneven bins, finally the momenta is split into 13 (18) uneven bins for the muons (kaons) over the range $[3, 100]$ GeV. The data used to create these histograms are calibration samples of $B^+ \to J/\psi\,(\mu^+\mu^-)K^+$ and $D^{*+} \to D^0(K^-\pi^+)\pi^+_s$ collected using selections with no PID requirement. For each event in the simulated dataset, a new PID value is assigned by finding the corresponding bin of the event in the PID distribution histograms and selecting a random value from that bin. Figure 4.1 shows the distributions of the PID

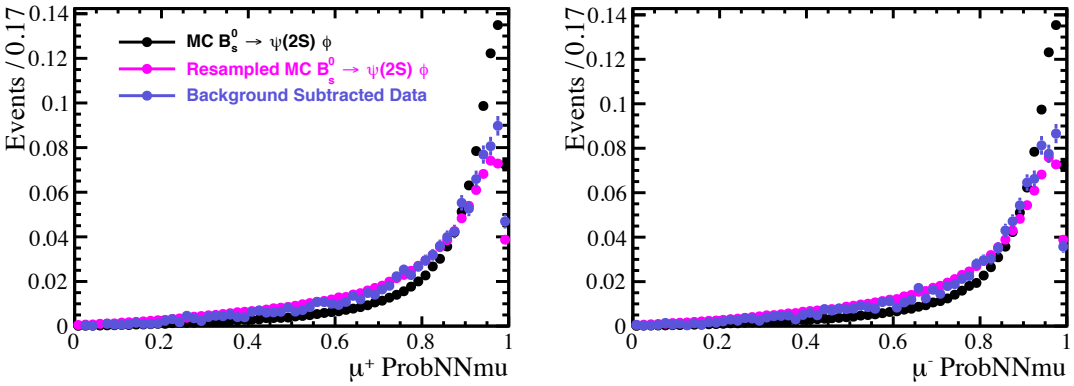


Figure 4.1 *Distribution of the PID variables in uncorrected simulated $B_s^0 \to \psi(2S)\phi$ data (black points), resampled simulated $B_s^0 \to \psi(2S)\phi$ data (pink points) and background subtracted $B_s^0 \to \psi(2S)\phi$ data (purple points). The histograms have been normalised by the number of events.*

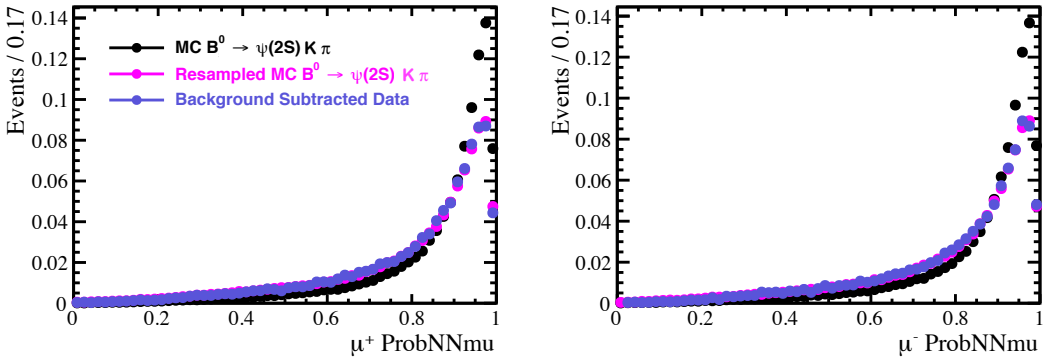


Figure 4.2 *Distribution of the PID variables in uncorrected simulated $B^0 \to \psi(2S)K^{*0}$ data (black points), resampled simulated $B^0 \to \psi(2S)K^{*0}$ data (pink points) and background subtracted $B^0 \to \psi(2S)K^{*0}$ data (purple points). The histograms have been normalised by the number of events..*

variables in $B_s^0 \to \psi(2S)\phi$ simulation before and after the resampling process has been applied along with background subtracted data (using the B_s^0 mass as the discriminating variable) and Figure 4.2 shows the same distributions for the $B^0 \to \psi(2S)K^*(892)$ mode. The method used to create the background subtracted distributions is discussed in § 4.4.

Table 4.3 *Triggers used in this analysis. All events were required to have passed the HLT1 and HLT2 triggers.*

HLT1	`Hlt1DiMuonHighMassDecision`
	or `Hlt1TrackMuonDecision`
	or `Hlt1TrackAllL0Decision`
HLT2	`Hlt2DiMuonDetachedPsi2SDecision`
	or `Hlt2DiMuonDetachedHeavyDecision`

4.2 Selection

To isolate the signal in the data for analysis a combination of cut based selections and a multi-variate analysis was used. A cut based selection is a decision to accept events that satisfy a constraint on a specific variable, or variables. The aim was to produce a final sample with a relatively small background component to optimise the measurement precision of ϕ_s.

4.2.1 Trigger

The triggers that the data were required to have passed are identified in Table 4.3. Each event was required to have passed both the HLT1 and HLT2 trigger requirements purely on the properties of the signal candidate. All the HLT1 triggers used require either the L0Muon or L0DiMuon trigger to have been passed which puts a lower bound on the transverse momenta of the muons (trigger criteria are listed in Table 4.4). The inclusion of an upper limit to the SPD multiplicity in the L0 trigger removes busy events for which the tracking and RICH pattern recognition efficiencies are reduced while maintaining the signal efficiency [26]. There are two HLT2 triggers identified, and although it would have been preferable to use only the Hlt2DiMuonDetachedPsi2S trigger, it was not available for 2011 data, so the Hlt2DiMuonDetachedHeavy was used as an alternative. The main difference between these two HLT2 triggers is the Hlt2DiMuonDetachedHeavy has a tighter cut on the separation of the $\psi(2S)$ vertex and the primary vertex (taking into consideration the uncertainty of the vertex fit) which can be seen as a tighter cut on the decay time. The criteria for the L0 triggers required by the HLT triggers selected are shown in Table 4.4 and the cuts implemented in the HLT triggers are listed in Table 4.5.

Table 4.4 *Requirements to pass the L0Muon and L0DiMuon triggers. SPD: Scintillating Pad Detector [27].*

Trigger	Cut parameter	Value
L0 Muon	SPD Multiplicity	< 600
	p_{T}	$> 1.48\,\mathrm{GeV}/c$
L0 DiMuon	SPD Multiplicity	< 600
	$\sqrt{p_{\mathrm{T}1st\ largest} \times p_{\mathrm{T}2nd\ largest}}$	$> 1.296\,\mathrm{GeV}/c$

4.2.2 Stripping and preselection

The next step was to use the dedicated $B_s^0 \to \psi(2S)\phi$ stripping line to select a sample of prospective signal events. This stripping line applies a series of cuts to the event of interest which are detailed in Table 4.6. In addition to the cuts in the stripping line, a small set of preselection cuts were applied to the data, which are also detailed in Table 4.6. There are a number of variables used in the stripping and preselection, which include:

- $\mathrm{DLL}_{\mu\pi}$, the difference between the log likelihood of the muon hypothesis and the log likelihood of the pion hypothesis for a final state particle.
- $\mathrm{DLL}_{K\pi}$, the difference between the log likelihood of the kaon hypothesis and the log likelihood of the pion hypothesis for a final state particle.
- ProbNNk, the probability that the candidate is a kaon based on a Neural Net analysis.
- χ^2_{vtx}/nDoF, the chi-squared per degree of freedom of the vertex fit of the B_s^0, ϕ and $\psi(2S)$ mesons.
- χ^2_{track}/nDoF the chi-squared per degree of freedom of the track fit of the kaons and muons.
- p_{T}, the transverse momenta of the ϕ meson.
- DOCA χ^2, the distance of closest approach of the two muons.
- $\chi^2_{\mathrm{IP(next)}}$, the chi-squared of the Impact Parameter (IP) of the B_s^0 with the next best vertex.
- t, the decay time of the B_s^0 in ps.
- σ_t, the uncertainty on the decay time of the B_s^0 in ps.

- M, the invariant mass of certain particle combinations.

Table 4.5 *Trigger selections. For the requirements in brackets, these are the selections for 2012 data if they are different from the 2011 requirements [28].*

Trigger	Cut parameter	Value
Hlt1DiMuonHighMassDecision	L0 Muon or L0 Dimuon	pass
	χ^2_{track}/nDoF (μ^+, μ^-)	< 4
	p_{T} (μ^+, μ^-)	$> 500\,\text{MeV}/c$
	p (μ^+, μ^-)	$> 6000\,\text{MeV}/c$
	DOCA $\chi^2(\psi(2S))$	< 0.2mm
Hlt1TrackMuonDecision	L0 Muon or L0 Dimuon	pass
	VELO hits/track	> 0
	VELO missing hits/track	< 999
	hits	> 0
	χ^2_{track}/nDoF (μ)	< 2 (< 2.5)
	p (μ)	$> 8\,\text{GeV}/c$ ($> 3\,\text{GeV}/c$)
	p_{T} (μ)	$> 1\,\text{GeV}/c$
	IP (μ)	0.1 mm
	χ^2_{IP} (μ)	> 16
Hlt1TrackALLL0Decision	L0 Decision Physics	pass
	VELO hits/track	> 9
	VELO missing hits/track	< 3
	hits	> 16
	χ^2_{track}/nDoF (μ)	< 2 (< 2.5)
	p (μ)	$> 10\,\text{GeV}/c$ ($> 3\,\text{GeV}/c$)
	p_{T} (μ)	$> 1.7\,\text{GeV}/c$
	IP (μ)	– (0.1 mm)
	χ^2_{IP} (μ)	> 16
Hlt2DiMuonDetachedPsi2SDecision	χ^2_{vtx}/nDoF ($\psi(2S)$)	< 25
	Vertex separation/σ ($\psi(2S)$)	> 3
	$\lvert M(\mu^+\mu^-) - M(\psi(2S))\rvert$	$< 120\,\text{MeV}/c^2$
Hlt2DiMuonDetachedHeavyDecision	χ^2_{vtx}/nDoF ($\psi(2S)$)	< 25
	p_{T} ($\psi(2S)$)	$> 0\,\text{GeV}/c$
	Vertex separation/σ ($\psi(2S)$)	> 5
	$M(\mu^+\mu^-)$	$> 2950\,\text{MeV}/c^2$

Table 4.6 *Selection criteria for $B_s^0 \to \psi(2S)\phi$ candidates in stripping and preselection. See text for definition of parameters.*

Decay mode	Cut parameter	Stripping	Preselection
$\psi(2S) \to \mu^+\mu^-$	$\mathrm{DLL}_{\mu\pi}$ (μ^+, μ^-)	> 0	–
	$\chi^2_{\mathrm{vtx}}/\mathrm{nDoF}(\psi(2S))$	< 16	–
	$\lvert M(\mu^+\mu^-) - M(\psi(2S))\rvert$	$< 60\,\mathrm{MeV}/c^2$	$< 40\,\mathrm{MeV}/c^2$
	DOCA $\chi^2(\psi(2S))$	< 30	–
$\phi \to K^+K^-$	$\mathrm{DLL}_{K\pi}$ $(K\pm)$	> -2	–
	ProbNNk (K^+, K^-)	–	> 0.1
	$K\pm$ $\chi^2_{\mathrm{track}}/\mathrm{nDoF}$	< 5	–
	p_{T} (ϕ)	$> 1\,\mathrm{GeV}/c$	–
	$M(\phi)$	$\in [1000, 1040]\,\mathrm{MeV}/c^2$	$\in [1008, 1032]\,\mathrm{MeV}/c^2$
	$\chi^2_{\mathrm{vtx}}(\phi)$	< 16	–
$B_s^0 \to \psi(2S)\phi$	$M(B_s^0)$	$\in [5150,\ 5550]\,\mathrm{MeV}/c^2$	–
	$\chi^2_{\mathrm{IP(next)}}(B_s^0)$	–	> 50
	$\chi^2_{\mathrm{vtx}}/\mathrm{nDoF}(B_s^0)$	< 10	–
	t	$t > 0.2\,\mathrm{ps}$	$0.3 < t < 14.0\,\mathrm{ps}$
	σ_t	–	$< 0.12\,\mathrm{ps}$

4.2.3 Boosted Decision Tree (BDT)

After stripping, preselection and trigger have been applied additional selections were needed to remove the combinatorial background events remaining in the data sample (Figure 4.3). A number of multivariate (MVA) strategies were compared and it was found that the MVA with the best performance was the Boosted Decision Tree (BDT).

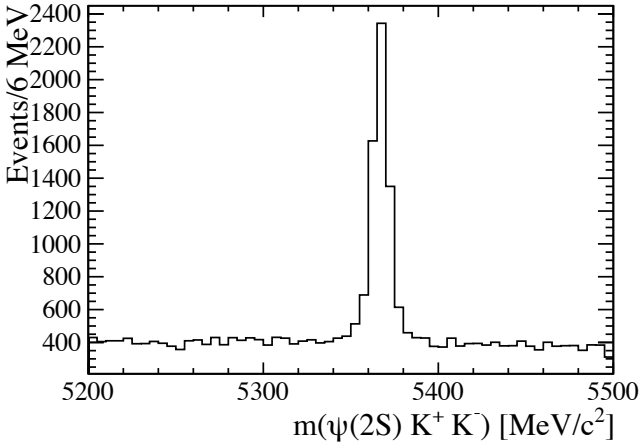


Figure 4.3 *The reconstructed $m(\psi(2S)K^+K^-)$ mass distribution after trigger, stripping and preselection has been applied.*

A decision tree is a sequence of binary nodes used to classify an event, where each node uses a variable to make a 'yes/no' decision (Figure 4.4). The result is that the observable space, constructed using the variables fed into the BDT, is split into disjoint boxes corresponding to the classification of an event as signal or background. A decision tree is normally trained on a dataset where the identity of the events is known so the classification of a subset of the observable space as either signal or background is based on the number of known signal events that are in the final 'leaf' node related to that subspace.

To enhance the performance of this method, a large number of trees can be used to produce a single decision on a subspace by combining the votes from the individual trees. This is the step that introduces the 'boosted' aspect of this MVA method. When a tree misclassifies an event, the weight of that event is increased, or 'boosted', and a new tree is formed. This process is repeated multiple times until a forest of trees has been formed, with the later trees having effectively learnt from the trees that resulted in misclassified events.

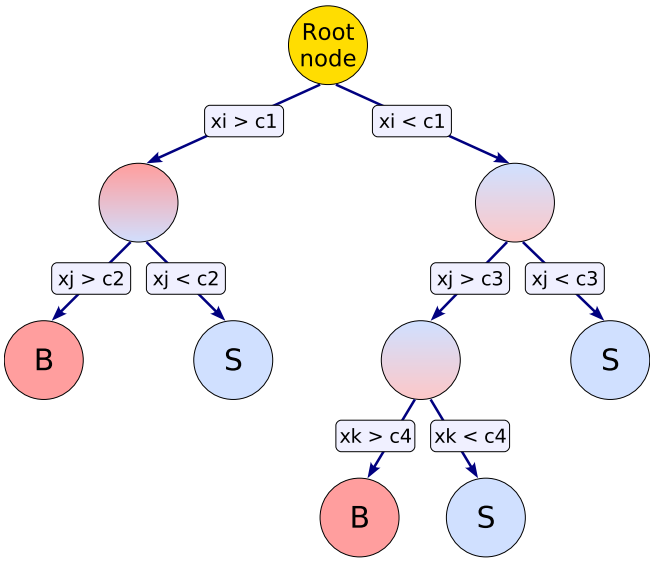


Figure 4.4 *The structure of a decision tree. For each node the discriminating variable is labeled x. The final 'leaf' nodes are labelled 'S' for signal and 'B' for background [29].*

When training the BDT fully simulated $B_s^0 \to \psi(2S)\phi$ events (details on the simulation conditions and parameters used to generate the data can be found in § 4.1) were used to model signal events and events from data, outside the reconstructed B_s^0 mass signal region ($> 5400\,\mathrm{MeV}/c^2$) were used to model the background. This region was chosen because only combinatorial events were expected to be present (Figure 4.3). For both the signal and background samples events from both 2011 and 2012 datasets were used.

The variables used to train the BDT are listed below, where (log) indicates that the logarithm of the variable has been used.

- μ^+, μ^- ProbNNmu, the probability that the particle is a muon based on a neural net analysis.
- K^+, K^-, μ^+, μ^- MINIPCHI2 (log), the χ^2 of the minimum value of the IP of the kaons and muons with respect to the primary vertex.
- B_s^0 p_T, the transverse momentum of the B_s^0 meson.
- ϕ p_T, the transverse momentum of the ϕ meson.
- B_s^0 vertex χ^2 (log), the χ^2 of the vertex fit of the B_s^0 meson.
- B_s^0 IP (log), the impact parameter of the B_s^0 meson with the primary vertex.

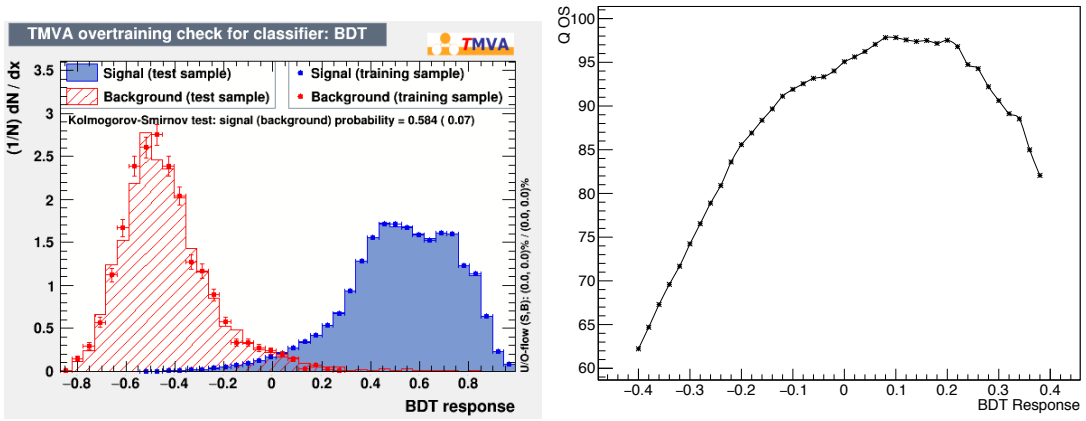


Figure 4.5 *Overtraining plot for the BDT (left) and plot showing the value of Q (Equation 4.2) vs cut on the BDT response (right). A cut of* > 0.1 *was chosen to select the final data sample.*

- B_s^0 flight distance (log), the distance the B_s^0 meson has travelled in the detector with respect to the primary vertex.
- $\psi(2S)$ endvertex χ^2, the χ^2 of the $\psi(2S)$ vertex fit.

These variables were chosen based on the difference in the distributions of these variables between signal and background (Figure B.2). To have confidence in the use of simulated data to model signal in the BDT the distributions of the variables used was compared between background subtracted data (discussed in § 4.4) and the fully simulated events (Appendix B). The majority of variables show no significant differences in the distributions between real and simulated data, except for the PID variables where the differences have been significantly reduced by the resampling process discussed in § 4.1.1.

One possible pitfall of BDT analyses is that it is very easy for the trees to become overtrained. This means that the BDT has not been trained to identify actual properties of the signal or background, but statistical variations in the training samples that were used. To check this a plot showing the distribution of the BDT response for the signal and background test and training samples was made (Figure 4.5). In this plot it is clear that the distribution of the BDT response (the combined vote of many individual decision trees) agrees between the training sample (points) and the test sample (histogram) for both signal events (blue) and background events (red), so no overtraining was observed. For further details on the BDT training, including plots showing the correlation of variables and signal/background separation see Appendix B.

The final step was to choose at which value to make a cut on the BDT response. A commonly used criterion is the signal significance, defined as $S/\sqrt{S+B}$, where S is the number of signal events in the resulting dataset, and B is the number of background events. The value at which the BDT response is cut is the value that optimises this criterion. However, the aim of this analysis was not to make an observation of this decay mode, but to make the best measurement possible of the *CP* violating parameter ϕ_s, so the best objective function to use is the inverse of the variance of this parameter [30]. After some simplifications the resulting figure of merit is shown in Equation 4.2 and is not only dependent on the signal significance, but also on how well the decay time of the events was measured (σ_t), the B_s^0-$\overline{B}_s^0$ oscillation frequency (Δm_s) and how accurately the identification of B_s^0 or $\overline{B}_s^0$ decays were made (characterised by the mistag probability, ω [1]).

$$\begin{aligned} Q &= \sum_{j=1}^{K} \frac{S_j^2}{S_j + B_j} \bar{D}_j \\ \bar{D}_j &= \frac{1}{S_j + B_j} \sum_{l=1}^{S_j+B_j} (1 - 2\omega_l)^2 e^{-(\Delta m_s \sigma_{t_l})^2} \end{aligned} \tag{4.2}$$

To calculate Q the data was split into $K = 4$ bins in decay time (chosen to ensure minimal variation of signal significance across the bin while minimising computation time), with the bin boundaries chosen so that there were an approximately equal number of signal events in each bin. For each bin in decay time a fit was made to the distribution of m($\psi(2S)K^+K^-$), using the mass model defined in § 4.4) to extract the number of signal events (S_j) and the number of background events (B_j). $\bar{D}_j$ was then calculated for each bin by iterating over all the events that fall within the bin boundaries, where Δm_s, the difference in mass between the heavy and the light B_s^0 mass eigenstates, is equal to $\Delta m_s = 17.757$ ps^{-1} [31] and ω_l is the per-event mistag probability for the OS tagging algorithm (more information on flavour tagging algorithms is included in § 5.5). The final value for Q was calculated as the sum of $\bar{D}_j$, multiplied by the square of the signal significance, over all bins in decay time. The variation of Q with the cut on the BDT response is shown in Figure 4.5 (right) and the value found to maximise the value of this function is > 0.1. The resulting distributions of m($\psi(2S)K^+K^-$), m($\mu^+\mu^-$) and m(K^+K^-) are shown in Figure 4.6, where it can be seen that there is very little background.

[1] Details of the Flavour Tagging algorithms and how they are used to identify B_s^0 or $\overline{B}_s^0$ decays are given in § 5.5

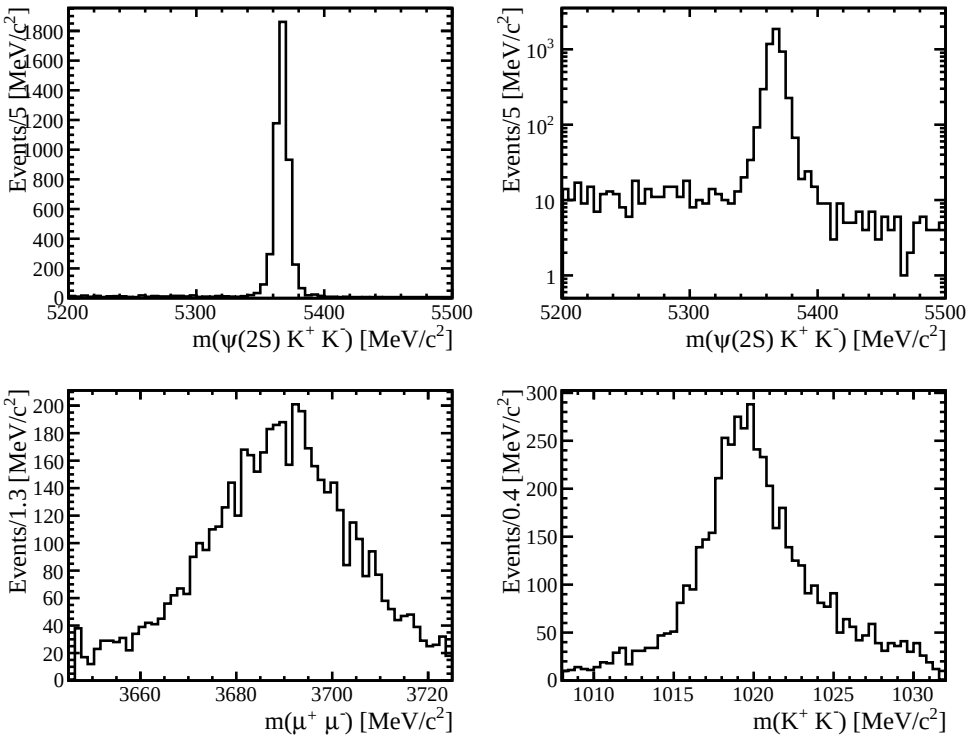


Figure 4.6 *Reconstructed mass distributions of the $m(\psi(2S)K^+K^-)$ (top plots, right is Log scale), $m(\mu^+\mu^-)$ (bottom left) and $m(K^+K^-)$ (bottom right) mesons in $3\,fb^{-1}$ of data after final selections have been applied.*

4.3 Peaking Backgrounds

One type of background that was not targeted in the selections outlined in the previous sections are backgrounds that are not flat in the $\mathrm{m}(\psi(2S)K^+K^-)$ distribution, called peaking backgrounds. In particular these can come from reflections: decays of other b-hadrons where one or more of the final state particles has been misidentified such that it has the same final state as the signal mode.

The method that was used to check for backgrounds from reflections is as follows; the events which have a reconstructed B_s^0 mass that falls outside the signal region ($\mathrm{m}(\psi(2S)K^+K^-) < 5330\,\mathrm{MeV}/c^2$ or $\mathrm{m}(\psi(2S)K^+K^-) > 5400\,\mathrm{MeV}/c^2$), called the sidebands, were reconstructed under the mass hypothesis of the reflected mode, and the b-hadron mass distribution was fit with a Gaussian for signal (with the mean fixed to the PDG value of the reflection) and a Chebyshev polynomial for background. The resulting yield only represents the number of peaking background events in the sidebands of data, so simulated events were used to scale this number to the full mass range. The two reflected modes that were searched for were the $B^0 \to \psi(2S)K^*(892)$ and $\Lambda_b^0 \to \psi(2S)pK^-$ decay modes.

4.3.1 $B^0 \to \psi(2S)K^*(892)$ Reflection

For events from this decay mode to have been reconstructed as signal, the pion from the $K^*(892) \to K\pi$ decay needs to have been misidentified as a kaon. To search for background events from this mode the upper mass sideband ($\mathrm{m}(\psi(2S)K^+K^-) > 5400\,\mathrm{MeV}/c^2$) in signal data was reconstructed under the $B^0 \to \psi(2S)K^*(892)$ hypothesis. Only the upper sideband was used because few events from this reflection were expected in the lower mass sideband, as shown by the plot of simulated $B^0 \to \psi(2S)K^*(892)$ events reconstructed and selected under the $B_s^0 \to \psi(2S)\phi$ hypothesis (Figure 4.7, left). The $\mathrm{m}(\psi(2S)K\pi)$ distribution is fit using a Gaussian for any signal, with the mean fixed to the PDG value for the B^0 mass and a polynomial for the background. The fit resulted in 14 ± 20 events, which when scaled to the full mass range equals 31 ± 45 $B^0 \to \psi(2S)K^*(892)$ events. The number of events from this reflection is small compared to the number of signal events (see later in this chapter for the signal yield) and so is not considered futher.

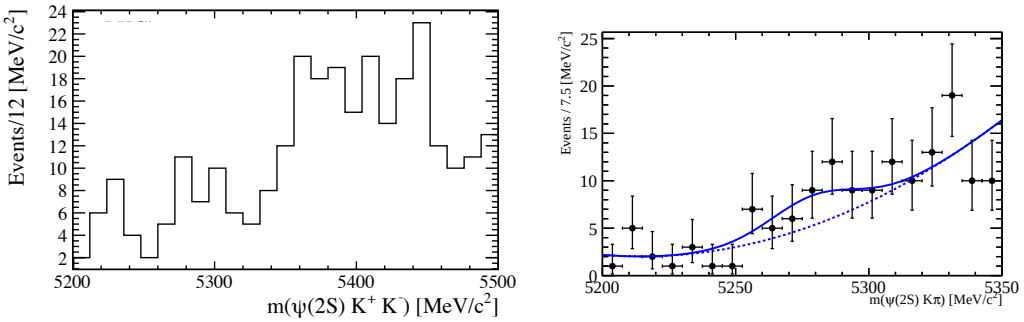


Figure 4.7 *Simulated $B^0 \to \psi(2S)K^*(892)$ decays reconstructed under the $B_s^0 \to \psi(2S)\phi$ hypothesis (left) and the fit to signal data sidebands reconstructed as $B^0 \to \psi(2S)K^*(892)$. The fit was performed using a Gaussian for the signal (with mean fixed to the expected value from PDG and the width free) and a polynomial for the background.*

4.3.2 $\Lambda_b^0 \to \psi(2S)pK^-$ Reflection

For this reflection, the p from the Λ_b^0 decay needs to have been misidentified as a K^+. To search for these events both the upper and lower mass sidebands ($\mathrm{m}(\psi(2S)K^+K^-) < 5330\,\mathrm{MeV}/c^2$ or $\mathrm{m}(\psi(2S)K^+K^-) > 5400\,\mathrm{MeV}/c^2$) of the $\psi(2S)$ constrained B_s^0 mass distribution in signal data were reconstructed under the $\Lambda_b^0 \to \psi(2S)pK^-$ hypothesis. Both the upper and lower mass sidebands in data were used because reflections from this mode are expected in both regions, as shown by the mass distribution of simulated $\Lambda_b^0 \to \psi(2S)pK^-$ reconstructed under the $B_s^0 \to \psi(2S)\phi$ hypothesis (Figure 4.8, left). A fit was performed to the signal data sidebands reconstructed under the $\Lambda_b^0 \to \psi(2S)pK^-$ hypothesis using a Gaussian for the signal with the mean fixed to the PDG value for the Λ_b^0 mass and a polynomial for the background, shown in Figure 4.8 (right). This fit resulted in 27 ± 17 $\Lambda_b^0 \to \psi(2S)pK^-$ events. Using simulated events to scale this number to the full mass range gives 37 ± 23 $\Lambda_b^0 \to \psi(2S)pK^-$ events in the signal data sample. Again, this is a small number of events compared to the number of signal events so no contribution from this decay mode was considered in the fits for ϕ_s.

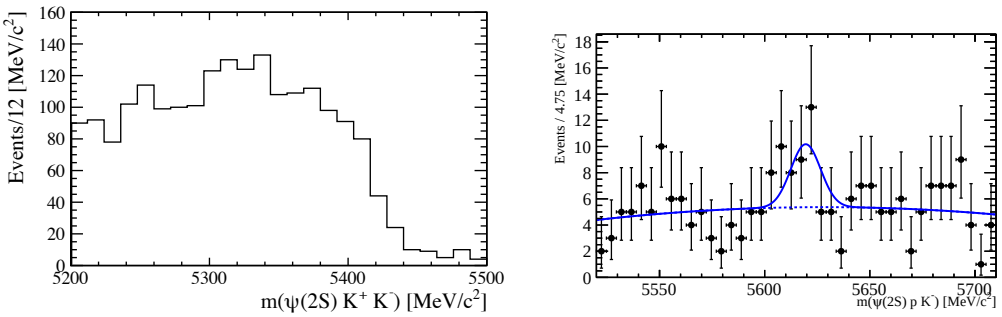


Figure 4.8 *Simulated $\Lambda_b^0 \to \psi(2S)pK^-$ decays reconstructed under the $B_s^0 \to \psi(2S)\phi$ hypothesis (left) and the fit to the signal data sidebands reconstructed as $\Lambda_b^0 \to \psi(2S)pK^-$. The fit was performed using a Gaussian for the signal (with mean fixed to the expected value from PDG and the width free) and a polynomial for the background.*

4.4 Signal Yield and sPlot method

To extract the signal and background yields in the final dataset after selections the reconstructed m($\psi(2S)K^+K^-$) distribution was fit using a sum of two Crystal Ball (Equation 4.3) functions for the signal and an exponential for the background. A Crystal Ball function has a Gaussian core with mean $\bar{x}$ and width σ and power law tail on one side, the threshold for which is controlled by the parameter α and the slope by n. The parameters A and B are defined as follows: $A = \left(\frac{n}{|\alpha|}\right) \cdot e^{\frac{|\alpha|^2}{2}}$, $B = \frac{n}{|\alpha|} - |\alpha|$, and N is the normalisation. The values for the parameters of the mass model resulting from the fit can be found in Table 4.7.

$$f(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, & \text{for } \frac{(x-\bar{x})}{\sigma} > -\alpha \\ A \cdot (B - \frac{(x-\bar{x})}{\sigma})^{-n}, & \text{for } \frac{(x-\bar{x})}{\sigma} \leq -\alpha \end{cases} \tag{4.3}$$

The fit was performed in the reconstructed m($\psi(2S)K^+K^-$) range of $5200 - 5500$ MeV/c^2 and results in 4697 ± 71 signal events. Figure 4.9 shows the plot of the fit zoomed in to the reconstructed m($\psi(2S)K^+K^-$) range of $5300 - 5400$ MeV/c^2. The data sample of simulated signal events was also fit using the signal mass model and the parameters found are shown in Table 4.7. The comparison of the signal mass model extracted from the fit to real data and simulated data is shown in Figure 4.10.

From the fit to the m($\psi(2S)K^+K^-$) mass distribution it is clear there is a small number of background events remaining in the selected data sample (Table 4.7). There are two strategies that can be used to include the knowledge of these

Table 4.7 *Results of the fit to the $m(\psi(2S)K^+K^-)$ distribution in real and simulated data samples in the reconstructed $m(\psi(2S)K^+K^-)$ range of $5200 - 5500$ MeV/c^2. The signal is modelled as a sum of two Crystal Ball functions with fraction $f_{\rm cb1}$, and the background is modelled with an exponential (not included in the fit to simulated data).*

Parameter	Data value	MC value
α_1	1.27 ± 0.20	1.967 ± 0.085
α_2	-0.94 ± 0.45	-2.024 ± 0.035
σ_1 [MeV/c^2]	4.61 ± 0.41	3.473 ± 0.069
σ_2 [MeV/c^2]	3.92 ± 0.45	5.74 ± 0.12
μ [MeV/c^2]	5366.93 ± 0.12	5367.11 ± 0.015
$f_{\rm cb1}$	0.65 ± 0.20	0.482 ± 0.032
$n_{\rm cb1}$	6.8 ± 4.0	2.056 ± 0.032
$n_{\rm cb2}$	5.9 ± 4.6	2.76 ± 0.13
$N_{\rm bkg}$	501 ± 28	–
$N_{\rm sig}$	4697 ± 71	–
slope [(MeV/c^2)$^{-1}$]	-0.0043 ± 0.00060	–

background events in the likelihood fit performed to extract values for $\Delta\Gamma_s$ and ϕ_s. First, the Probability Density Function (PDF) used in the likelihood fit to data can be constructed as a sum of the PDF used to model the signal and the PDF used to model the background. This method can be limited by the complexity of the models used in the fit, or the number of species of background that need to be considered. The second method is to statistically subtract the background in the data sample using the *sPlot* method [32]. The *sPlot* method uses the PDFs, P, of the signal and background, the yields, N, and the covariance matrix of the species yields, V, from a likelihood fit to a discriminating variable where the distributions of the signal and background components are well understood to generate a set of weights called *sWeights*, shown in Equation 4.4. The PDFs are dependent on some variables $\overrightarrow{x}$ and parameters λ. These per event weights can be used to statistically subtract the background and are included in the likelihood fit to extract the physics parameters so only the signal PDF is required. This method requires that the discriminating variable used to generate the *sWeights* is uncorrelated with the variables used in the likelihood fit to extract the physics parameters. For this analysis the *sPlot* method is used to simplify the PDF used in the fit to extract the physics parameters using *sWeights* generated by the fit to

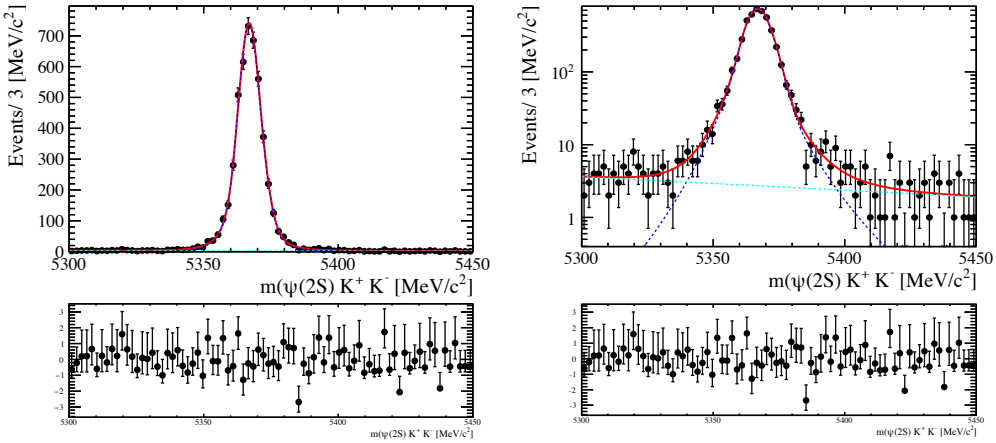


Figure 4.9 *The fit to the reconstructed $m(\psi(2S)K^+K^-)$ mass distribution where the right plot is on a log scale for the y-axis. The fit is performed using a sum of two Crystal Ball functions (blue dashed line) for the signal and an exponential for the background (cyan dashed line) and the results are shown in Table 4.7. The total PDF is shown as the dashed red line. The plots shown here are zoomed in to the mass range of* 5300-5450 MeV/c^2 *whereas the fit is performed over the full range of* 5200-5500 MeV/c^2.

the m$(\psi(2S)K^+K^-)$ distribution shown in Figure 4.9.

$$ {}_s\mathcal{P}_n(\overrightarrow{x}) = \frac{\sum_{j=1}^{N_s} V_{nj} P_j(\overrightarrow{x};\lambda)}{\sum_{k=1}^{N_s} N_k P_k(\overrightarrow{x};\lambda)} \tag{4.4} $$

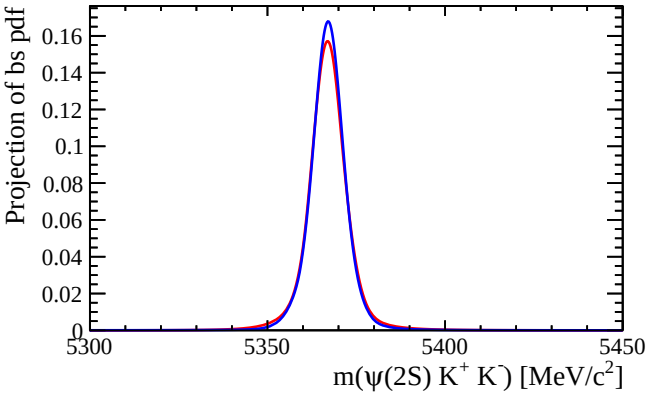


Figure 4.10 *Comparison of the B_s^0 mass fit results (PDF) for real data (red) and simulated data (blue).*

Chapter 5

Detector Acceptance and Resolution

The detector components and reconstruction algorithms introduce resolution and acceptance effects into the data. These effects must be modelled and included in the fits to data to produce reliable measurements of the underlying physics parameters. The decay time resolution is included in the fit to data by convolving the signal Probability Density Function (PDF) with the decay time resolution model and the detector acceptance is included by multiplying the signal PDF by the acceptance efficiency. The construction of the signal PDF and incorporation of resolution and acceptance effects are discussed in detail in § 6.1.

5.1 Decay Time Resolution

The LHCb detector has a finite decay time resolution of $\sim$ 50fs, which is dominated by the uncertainty of the decay length, i.e. how well the primary and secondary vertices were fitted [33]. The effect of the resolution is that the observed amplitude of an oscillation, which is proportional to $\sin(\phi_s)$ with a frequency given by Δm_s, is damped by a factor called the dilution (D) [34]. To retrieve the unbiased value of ϕ_s from a fit to data a function describing the detector resolution is convolved with the signal PDF. The model for the decay time resolution is extracted from a data sample with the same kinematics as the signal mode, selected using a stripping line and trigger that does not remove events at short decay times. The

dimuon mass distribution of this data sample was examined to look for 'prompt' J/ψ or $\psi(2S)$ events that have not come from a B_s^0 decay, but from the primary interaction. The background subtracted decay time distribution of the prompt events can then be fitted to extract the decay time resolution.

5.1.1 Prompt Data Sample

To search for a prompt $\psi(2S)$ data sample a dedicated $B_s^0 \to \psi(2S)\phi$ stripping line was used, which is similar to the stripping line used to select the signal data sample, but without the decay time biasing cut shown in Table 4.6 by the cut on the χ^2 of impact parameter to the next best vertex for the B_s^0 meson. In addition, this stripping line has a prescale applied, which means that only a fraction of the events that would pass this stripping line were flagged, in this case the prescale is 50%. A prescale is often applied to channels where a high yield is expected, where it is not necessary or feasible to record all events of that type. The triggers that were used to select events are a subset of those used for the signal data and only includes those that have no decay time biasing cuts.

Upon examination of the prompt data sample, no prompt $\psi(2S)$ peak was observed (Figure 5.1, left). Numerous different strategies were tried in an attempt to extract a clean $\psi(2S)$ peak, including applying various selections, but none were successful. As an alternative it was decided that a sample of prompt J/ψ events should be used. The prompt J/ψ data can be used to calibrate the decay time resolution model as the kinematics of the $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi(2S)\phi$ modes are similar.

The method of using prompt J/ψ data has been used in a related analysis, the measurement of ϕ_s using $B_s^0 \to J/\psi\,\phi$ [13, 35]. To select the data and calibrate the resolution model the methodology from that analysis was replicated. The stripping line used to select the data was a dedicated prescaled stripping line for $B_s^0 \to J/\psi\,\phi$ decays (without the decay time biasing selections). The triggers required were the HLT1DiMuonHighMass and J/ψ version of the HLT2DiMuon trigger, both of which do not bias the decay time distribution. The resulting dimuon mass distribution was fit using a Gaussian for the signal and an exponential for the background and is shown in Figure 5.1 (right). The background subtracted decay time distribution of this dataset is used to calibrate the decay time resolution model.

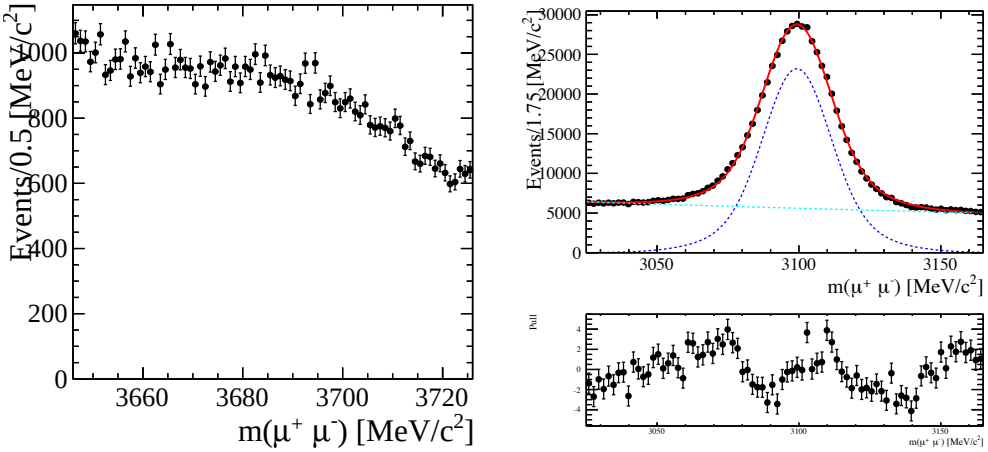


Figure 5.1 *(left) Distribution of $\mu^+\mu^-$ invariant mass for events selected using the prescaled $B_s^0 \to \psi(2S)\phi$ stripping line and triggers. No clear sign of the $\psi(2S) \to \mu^+\mu^-$ decay is visible. (right) Distribution of $\mu^+\mu^-$ invariant mass for events selected with the prescaled $B_s^0 \to J/\psi\,\phi$ stripping line and triggers. The fit is performed using a Gaussian for the signal and an exponential for the background.*

5.1.2 Calibration of the Resolution Model

The decay time resolution ($\mathcal{R}(t, \sigma_t)$) was modelled as a sum of two Gaussian functions (Equation 5.1) with a common mean, t_0, and the widths, σ_1 and σ_2, being dependent on the error of the measured decay time (σ_t), shown in Figure 5.2, where the fraction of the second Gaussian is labelled f. The error on the measured decay time is a consequence of the vertex fits performed to the primary vertex and the B_s^0 vertex, the largest contribution of which comes from the latter due to the relatively small number of tracks involved.

$$\mathcal{R}(t, \sigma_t) \propto (1-f)\frac{1}{\sigma_1(\sigma_t)} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_1(\sigma_t)}\right)^2} + f\frac{1}{\sigma_2(\sigma_t)} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_2(\sigma_t)}\right)^2} \tag{5.1}$$

In order to aid the fits to the prompt decay time distribution a parameter transformation (Equation 5.2) is performed to reduce the correlations between fit parameters as in the $B_s^0 \to J/\psi\,\phi$ 3 fb^{-1} analysis [13]. A quadratic calibration is defined for the transformed widths (Equation 5.3), and the parameters c', c'', b', b'' are found by performing a fit to the background subtracted decay time distribution of the prompt data. The parameter $\overline{\sigma_t}$ is the average decay time error which is

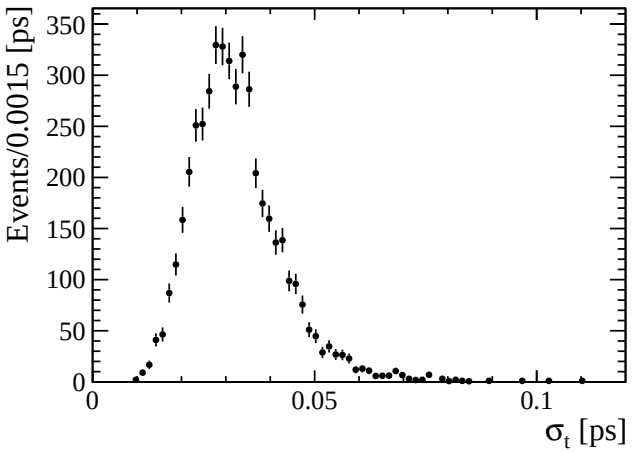


Figure 5.2 *The σ_t distribution of the prompt J/ψ data sample where the background has been subtracted statistically.*

fixed to the value in data (33 fs).

$$\begin{aligned}\sigma_1(\sigma_t) &= -\sqrt{\frac{f}{1-f}}\delta''(\sigma_t) + \delta'(\sigma_t) \\ \sigma_2(\sigma_t) &= \sqrt{\frac{1-f}{f}}\delta''(\sigma_t) + \delta'(\sigma_t)\end{aligned} \tag{5.2}$$

$$\begin{aligned}\delta'(\sigma_t) &= \sigma_t(c' + b'(\sigma_t - \overline{\sigma_t})) \\ \delta''(\sigma_t) &= \sigma_t(c'' + b''(\sigma_t - \overline{\sigma_t}))\end{aligned} \tag{5.3}$$

To perform the calibration an unbinned fit to the weighted decay time distribution of the prompt J/ψ events in the range [-1,8] ps (Figure 5.3) was performed using the function defined in Equation 5.4. This function includes components for the prompt peak, two components with lifetimes $\tau_{\rm l}$ and $\tau_{\rm s}$ to account for non-prompt events (with fractions $(1-f_{\rm s})$ and $f_{\rm s}$ respectively) all convolved with the resolution model, and a wide Gaussian ($G(t)$) to model events matched to the wrong primary vertex (PV). The number of events from each component are labelled $N_{\rm prompt}$, $N_{\rm ll}$ (where ll stands for long-lived) and $N_{\rm wpv}$ (where wpv stands for wrong primary vertex). The results of the fit are shown in Table 5.1.

$$\mathrm{PDF}(t|\sigma_t) = \mathcal{R}(t|\sigma_t) \otimes \left[N_{\rm prompt}\delta(t) + N_{\rm ll}\left(f_{\rm s}e^{-t/\tau_s} + (1-f_{\rm s})e^{-t/\tau_{\rm l}}\right)\right] + N_{\rm wpv}G(t) \tag{5.4}$$

To quantify the effect of the resolution on the result for ϕ_s the dilution is calculated

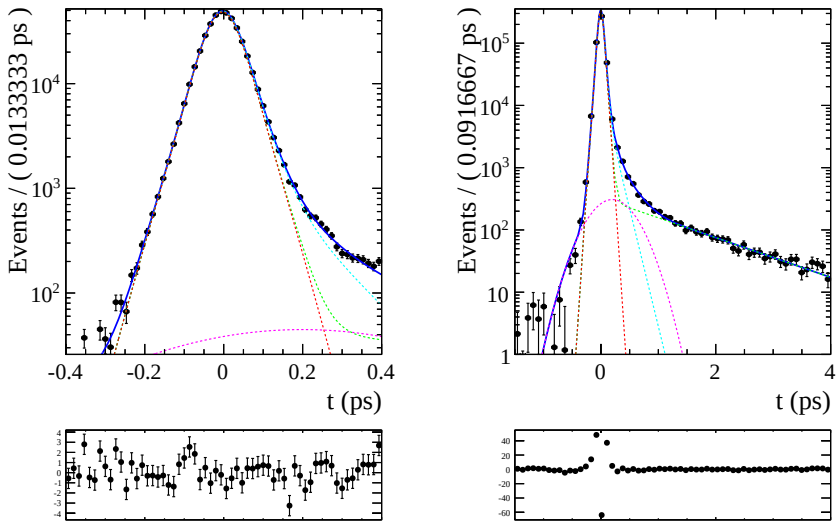


Figure 5.3 *Decay time distribution of the J/ψ prompt sample. The black points with error bars show the data and the solid blue line shows the result of an unbinned likelihood fit using the function defined in Equation 5.4. The red dashed line is the prompt component, the green and cyan dashed lines are the short and long-lived components respectively, and the pink is the wide Gaussian. Although the fit is performed in the* $[-1, 8]$ ps *range these plots show fit zoomed in to the peak,* $[-0.4, 0.4]$ ps *(left) and* $[-1, 4]$ ps *(right).*

using the equation defined in Equation 5.5.

$$D \;=\; (1-f)\; e^{\left(\frac{-\Delta m_s^2\,\sigma_1^2}{2}\right)} + f\; e^{\left(\frac{-\Delta m_s^2\,\sigma_2^2}{2}\right)} \tag{5.5}$$

The parameter Δm_s is the B_s^0-$\overline{B}_s^0$ oscillation frequency, taken as 17.757ps^{-1}. To calculate the dilution for the signal $B_s^0 \to \psi(2S)\phi$ data sample the *sWeights* generated from the fit to the $\text{m}(\psi(2S)K^+K^-)$ need to be taken into consideration. A parameter called the effective dilution is used to do this as shown in Equation 5.6 where w_i is the signal weight of each event, S_W is the sum of the weights and N is the number of events.

$$D_{effective} \;=\; \sqrt{\frac{1}{S_W}\sum_i^N w_i D_i^2} \tag{5.6}$$

The effective dilution of the $B_s^0 \to \psi(2S)\phi$ data sample from $3\,\text{fb}^{-1}$ of LHCb data is shown in Table 5.2. In addition the effective resolution was calculated, which is a measure of the width of the resolution if it was modelled using a single Gaussian function. This value matches the value reported in the $B_s^0 \to J/\psi\,\phi$ analysis [13]

Table 5.1 *The values of the fit of Equation 5.4 to the decay time distribution of J/ψ prompt data.*

Parameter	Fit Result
Resolution model calibration	
b' [ps^{-1}]	-5.122 ± 0.082
b'' [ps^{-1}]	-2.735 ± 0.090
c'	1.4678 ± 0.0018
c''	0.3871 ± 0.0028
f	0.2742 ± 0.0073
μ Gaussians [ps]	-0.003277 ± 0.000063
Long lived components	
τ_s [ps]	0.11492 ± 0.00097
τ_l [ps]	1.3639 ± 0.0052
$f_{\tau_{sl}}$	0.01092 ± 0.00017
$f_{\tau_{ll}}$	0.04396 ± 0.00080
Wide Gaussian Parameters	
σ wide	0.3615 ± 0.0071
μ wide [ps]	0.20 ± 0.30
f_{wpv}	0.00686 ± 0.00034

Table 5.2 *The effective resolution and effective dilution of the $B_s^0 \to \psi(2S)\phi$ data from $3\,fb^{-1}$ of LHCb data.*

Parameter	Value
Effective resolution	46.59 ± 0.98 fs
Effective dilution	0.7098 ± 0.0011

and is comparable to the single period for B_s^0-$\overline{B}_s^0$ oscillations ($\sim$ 56fs).

To rely on a resolution model calibrated using J/ψ prompt data it must be shown that the difference between the J/ψ prompt data and $\psi(2S)$ prompt data is negligible. To do this simulated signal events from the $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi(2S)\phi$ modes were used. A fit was performed to the difference between the true (t_{true}) and reconstructed (t_{reco}) decay time for simulated 2012 events for both modes using the double Gaussian resolution model with the same calibration as used for data. The calibration parameters resulting from the fit are shown in Table 5.3 and show no sizeable difference in the effective resolution between the two calibrations. Plots showing the fit to the difference in true and reconstructed decay time are shown in Figure 5.4. The fits to simulated datasets are used as a comparison between the two modes and the non-negligible pull distributions from these fits further motivate the implementation of a data driven calibration.

Table 5.3 *Resolution model parameters obtained from a fit to the difference between the true and reconstructed decay time using simulated 2012 $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to \psi(2S)\phi$ events.*

	$B_s^0 \to J/\psi\,\phi$ fit	$B_s^0 \to \psi(2S)\phi$ fit
b' [ps^{-1}]	1.17 ± 0.11	-0.100 ± 0.098
b'' [ps^{-1}]	1.77 ± 0.42	-6.2 ± 1.1
c'	1.3480 ± 0.0012	1.3752 ± 0.0033
c''	0.5907 ± 0.0046	0.721 ± 0.015
f_{G2}	0.00185 ± 0.00031	0.01834 ± 0.00062
μ Gaussians [ps]	-0.00040 ± 0.000036	-0.00036 ± 0.000089
Calculate the Dilution on $B_s^0 \to \psi(2S)\phi$ data		
Effective resolution	42.52 ± 0.90 fs	42.37 ± 0.89 fs
Dilution (D)	0.7517 ± 0.0015	0.7523 ± 0.0015

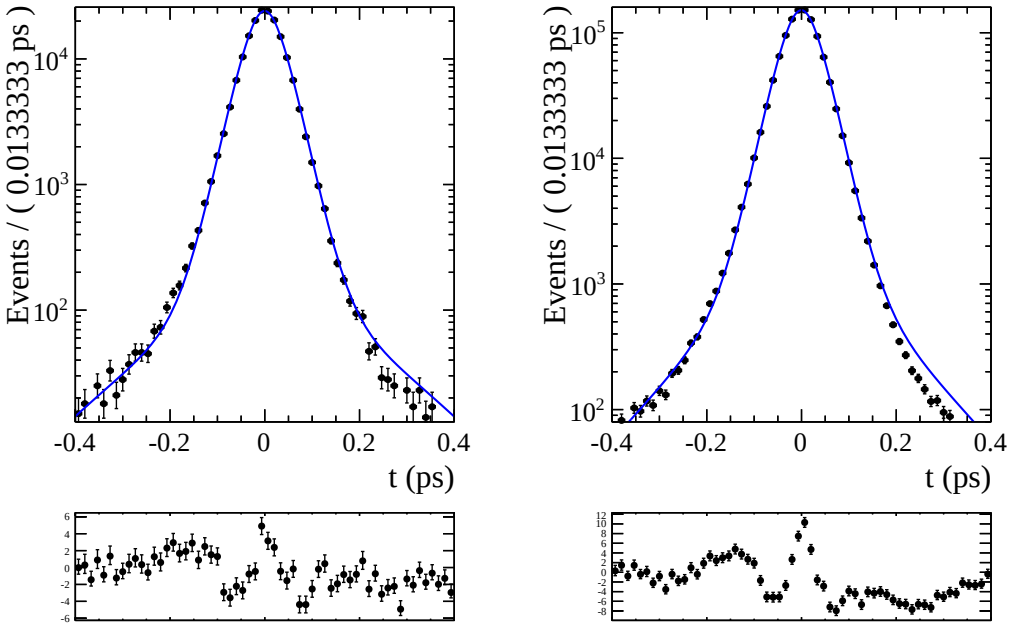


Figure 5.4 *Distribution of $t = t_{\mathrm{true}} - t_{\mathrm{reco}}$ of simulated $B_s^0 \to \psi(2S)\phi$ (left) and $B_s^0 \to J/\psi\,\phi$ (right) events. The solid blue line shows the result of a fit of the double Gaussian resolution model defined in Equation 5.1 to the data (black points with error bars).*

5.2 Decay Time Acceptance

The LHCb detector acceptance and reconstruction algorithms can bias the distribution of observables in data. Any bias of the decay time distribution in data that was not included in the fit model would affect the measured value of Γ_s and $\Delta\Gamma_s$. It was therefore important to understand the acceptance shape of the decay time. To do this a data driven method relying on the $B^0 \to \psi(2S)K^*(892)$ decay mode, for which the lifetime is well known, was used to extract a histogram modelling the decay time acceptance. This mode is appropriate to calculate the decay time acceptance for the signal $B_s^0 \to \psi(2S)\phi$ mode because the decay kinematics are similar, except that the ϕ decays to two kaons, the $K^*(892)$ decays to a kaon and a pion.

To select a sample of $B^0 \to \psi(2S)K^*(892)$ events the full $3\,\text{fb}^{-1}$ of LHCb data was used, with events selected using a dedicated stripping line, and the same trigger and selection criteria (including the BDT) as the $B_s^0 \to \psi(2S)\phi$ mode, except for making the appropriate allowance for the difference in final decay products (pion rather than a kaon). The B^0 mass distribution of the selected data was then fit to generate *sWeights* using a sum of two Crystal Ball functions for the signal and an exponential for the background, resulting in 27420 ± 230 signal events and 10940 ± 200 background events in the mass range $5200 - 5350\,\text{MeV}/c^2$. The results of the mass fit are shown in Table 5.4 and the plot is shown in Figure 5.5.

5.2.1 Peaking Backgrounds in the B^0 mass distribution

As for the $B_s^0 \to \psi(2S)\phi$ data sample, the $B^0 \to \psi(2S)K^*(892)$ data sample was examined for potential peaking backgrounds in the reconstructed B^0 mass distribution. If there were sizeable contributions from other decay modes that have a different lifetime to the B^0 mode then the effective lifetime of events in that sample would not be the true B^0 lifetime, and so the acceptance shape extracted from the data would not be correct.

Mis-identified $B_s^0 \to \psi(2S)\phi$ decays

A possible background can come from $B_s^0 \to \psi(2S)\phi$ reflections, where one of the kaons from the ϕ decay has been mis-identified as a pion. This is a potentially

Table 5.4 *Results of the fit to the $m(\psi(2S)K\pi)$ distribution in data after selections. The signal shape is the sum of two Crystal Ball functions with a common mean. In the fit to the data the background is modelled as an exponential.*

Parameter	Fit value
α_1	0.94 ± 0.15
α_2	-1.31 ± 0.11
σ_1 [MeV/c^2]	5.92 ± 0.49
σ_2 [MeV/c^2]	5.08 ± 0.11
μ [MeV/c^2]	5279.00 ± 0.06
fcb1	0.368 ± 0.091
n cb1	4.6 ± 1.5
n cb2	5.88 ± 0.57
$N_{\rm bkg}$	10940 ± 200
$N_{\rm sig}$	27420 ± 230
slope	-0.0017 ± 0.00029

dangerous background as the distribution of these events is not flat under the B^0 mass peak (as shown in Figure 5.6) and comes from a B_s^0 decay which has a different lifetime to the B^0 and so would modify the observed acceptance shape. To search for these background events the lower mass sideband, i.e. events that have a reconstructed B^0 mass value of less than 5250 MeV/c^2, of the selected $B^0 \to \psi(2S)K^*(892)$ data was reconstructed under the $B_s^0 \to \psi(2S)\phi$ mass hypothesis. The B_s^0 mass distribution was then examined for a peak at the B_s^0 mass value and no evidence of this was observed (Figure 5.7). No further steps to include this background in the B^0 mass fit were taken.

Mis-identified $B_s^0 \to \psi(2S)\pi^+\pi^-$ decays

A mis-identification of a pion as a kaon can result in $B_s^0 \to \psi(2S)\pi^+\pi^-$ background events appearing in the B^0 data sample. Figure 5.8 shows the distribution of $B_s^0 \to \psi(2S)\pi^+\pi^-$ decays under the $B^0 \to \psi(2S)K^*(892)$ mass hypothesis using fast simulation. From this plot is is clear that reflections from this background are outside the B^0 mass window (upper limit of 5350 MeV/c^2) selected for the time acceptance study so no contribution from this mode is considered.

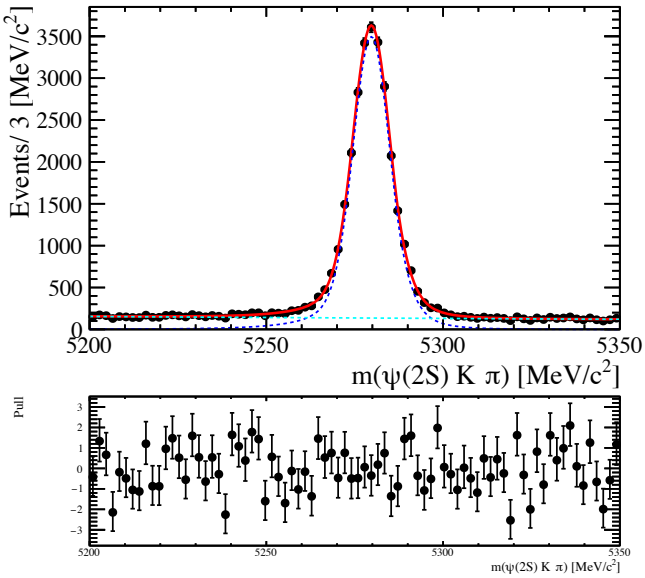


Figure 5.5 *Mass fit to the $B^0 \to \psi(2S)K^*(892)$ decay mode using $3\,fb^{-1}$ of LHCb data. The signal is modelled using a sum of two Crystal Ball functions (blue dashed line), and the background is modelled as an exponential (cyan dashed line). The total PDF is shown as the solid red line.*

Mis-identified $B^0 \to \psi(2S)\pi^+\pi^-$ decays

The same misidentification of a pion as a kaon can result in $B^0 \to \psi(2S)\pi^+\pi^-$ reflections appearing in the $B^0 \to \psi(2S)K^*(892)$ data sample. A sample of $B^0 \to \psi(2S)\pi^+\pi^-$ events, generated using fast simulation, was reconstructed under the $B^0 \to \psi(2S)K^*(892)$ mass hypothesis (Figure 5.9), and shows that a small tail appears in the mass window used for the B^0 selection. The background from this contribution is expected to be small because the ratio $\mathrm{BR}(B^0 \to \psi(2S)\pi^+\pi^-)/(B^0 \to \psi(2S)K^+\pi^-) \sim 0.1$ so a probability of 1% for the misidentification of a pion as a kaon would result in a background level from this decay of 0.1% and so is not considered further.

Mis-identified $\Lambda_b^0 \to \psi(2S)pK^-$ and $\Lambda_b^0 \to \psi(2S)p\pi^-$ decays

Another possible background can come from Λ_b^0 decays, either where the final state proton from $\Lambda_b^0 \to \psi(2S)pK^-$ decays has been misidentified as a pion, or the proton from $\Lambda_b^0 \to \psi(2S)p\pi^-$ has been misidentified as a kaon. A data sample for both modes was generated using a fast simulation, and reconstructed under the $B^0 \to \psi(2S)K^*(892)$ mass hypothesis, shown in Figure 5.10 ($\Lambda_b^0 \to \psi(2S)pK^-$)

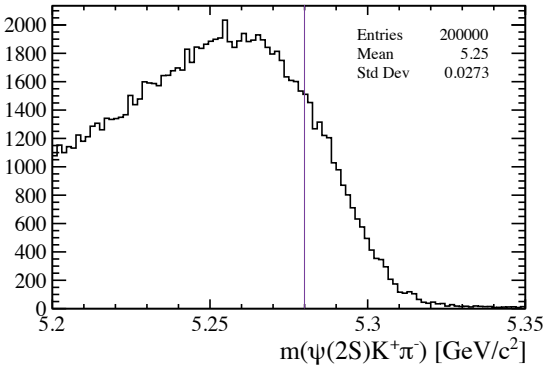


Figure 5.6 *Invariant mass distribution of $m(\psi(2S)K\pi)$ obtained from simulated $B_s^0 \to \psi(2S)\phi$ decays reconstructed under the $B^0 \to \psi(2S)K^{*0}$ hypothesis. The purple line shows the location of the known B^0 mass [31].*

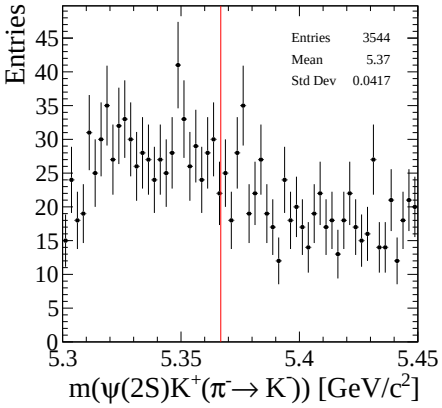


Figure 5.7 *Distribution of $m(\psi(2S)K^+K^-)$ for events in the lower sideband of the $m(\psi(2S)K^+\pi^-)$ distribution of the selected $B^0 \to \psi(2S)K^{*0}$ decays. The vertical red line shows the known position of the B_s^0 meson mass [31]. No peak is visible.*

and Figure 5.11 ($\Lambda_b^0 \to \psi(2S)p\pi^-$). These plots show that the reflections from these modes have quite a broad peak, spanning the full B^0 mass window used in the selection, and it is reasonable to conclude that any background present from these modes would have been absorbed by the combinatorial background.

A further check to look for such decays is to switch the mass hypothesis of the pion to a proton for all selected $B^0 \to \psi(2S)K^{*0}$ candidates. The invariant mass after applying this change is shown in Figure 5.12. No peak is visible at the known Λ_b^0 mass [31].

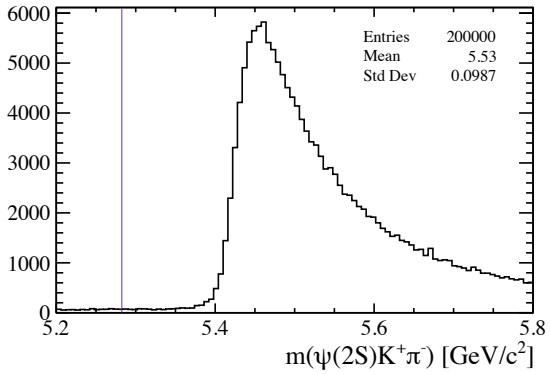


Figure 5.8 *Invariant mass distribution of $m(\psi(2S)K\pi)$ obtained from simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ decays that have been reconstructed under the $B^0 \to \psi(2S)K^{*0}$ hypothesis. The purple line shows the location of the known B^0 mass [31].*

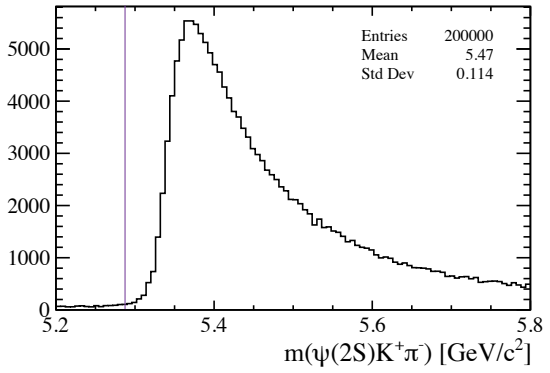


Figure 5.9 *Invariant mass distribution of $m(\psi(2S)K\pi)$ obtained from simulated $B^0 \to \psi(2S)\pi^+\pi^-$ decays that have been reconstructed under the $B^0 \to \psi(2S)K^{*0}$ hypothesis. The purple line shows the location of the known B^0 mass [31].*

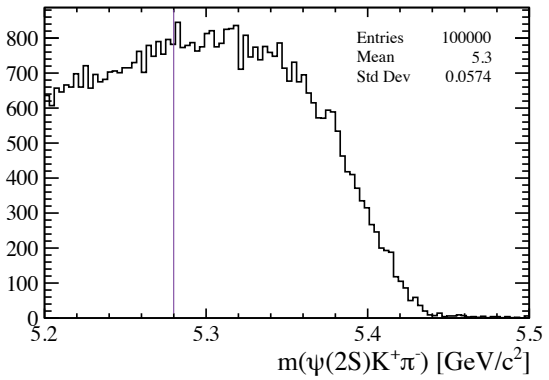


Figure 5.10 *Invariant mass distribution of $m(\psi(2S)K\pi)$ obtained from simulated $\Lambda_b^0 \to \psi(2S)pK^-$ decays that have been reconstructed under the $B^0 \to \psi(2S)K^{*0}$ hypothesis. The purple line shows the location of the known B^0 mass [31].*

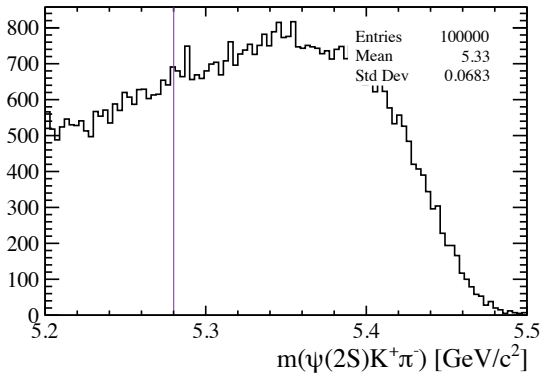


Figure 5.11 *Invariant mass distribution of $m(\psi(2S)K\pi)$ obtained from simulated $\Lambda_b^0 \to \psi(2S)p\pi^-$ decays that have been reconstructed under the $B^0 \to \psi(2S)K^{*0}$ hypothesis. The purple line shows the location of the known B^0 mass peak [31].*

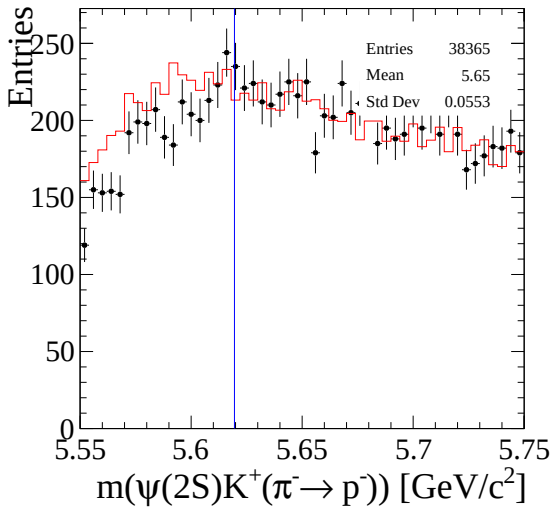


Figure 5.12 *Distribution of $m(\psi(2S)pK^-)$ for all selected $B^0 \to \psi(2S)K^{*0}$ events. The black points show the data while the red histogram shows the corresponding distribution from simulated $B^0 \to \psi(2S)K^{*0}$ decays. The vertical blue line shows the known position of the Λ_b^0 meson mass [31]. No peak is visible.*

Other backgrounds

It has been checked that backgrounds from $B^+ \to \psi(2S)K^+$ decays that are combined with an additional random pion are not present in the selected dataset. Similarly, backgrounds from cases where the true muon in $\psi(2S) \to \mu^+\mu^-$ decays is mis-identified as one of the kaons from the $K^{*0} \to K^+\pi^-$ decay are not present.

5.2.2 Resolution model

The resolution model used in the fit to the $B^0 \to \psi(2S)K^*(892)$ data decay time distribution is fixed to the results of a double Gaussian fit to prompt J/ψ data. This fit is performed in a single bin of σ_t and the widths of the Gaussians are not dependent on σ_t. The function fit to the prompt decay time is shown in Equation 5.4. The fit is shown in Figure 5.13 and the resulting parameter values are shown in Table 5.5. The parameters for the resolution model resulting from this fit are fixed in the fit to extract the decay time acceptance.

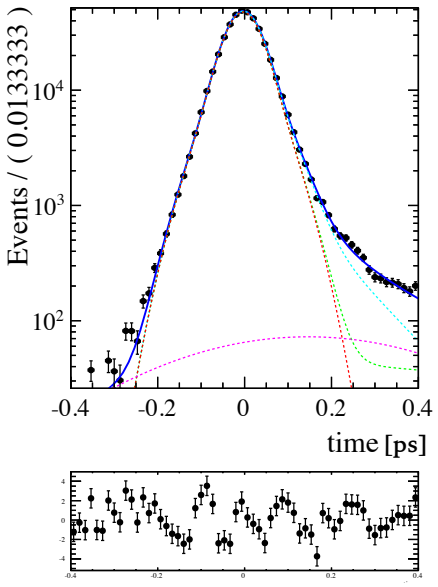


Figure 5.13 *The fit to the prompt J/ψ mass distribution in data using a double Gaussian resolution model and fit function as defined in Equation 5.4.*

Table 5.5 *The fit values of the fit of Equation 5.4 to J/ψ prompt data.*

Parameter	fit result
Resolution model calibration	
σ_1 [ps]	0.03885 ± 0.00013
σ_2 [ps]	0.07195 ± 0.00025
f_{G2}	0.349 ± 0.0042
Long lived components	
τ_s [ps]	0.1087 ± 0.0012
τ_l [ps]	1.332 ± 0.020
$f_{\tau_{sl}}$	0.04403 ± 0.00078
$f_{\tau_{ll}}$	0.01348 ± 0.00020
Wide Gaussian Parameters	
σ wide [ps]	0.3130 ± 0.0031

5.2.3 Extracting the decay time acceptance

To extract an acceptance shape for the $B^0 \to \psi(2S)K^*(892)$ decay mode the ratio of the decay time distribution of the selected $B^0 \to \psi(2S)K^*(892)$ data sample (after background subtraction by *sWeights*) and a simulated sample with no selection or detector effects included was performed. The lifetime of the simulated data was fixed to the PDG value of 1.520 ± 0.004 ps [31]. The resulting acceptance shape is shown in Figure 5.14.

Differences in the kinematics between the $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ modes, in particular the difference in mass between the ϕ and $K^*(892)$ mesons, can result in differences in the decay time acceptance curves for the two modes. To model this difference simulated data for the two modes were used. The decay time acceptance histograms were found for fully simulated $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ events, after trigger and selection, by taking the ratio of the decay time distribution with that of generator level events (simulated events without any reconstruction or detector effects included). In order to apply this process it is necessary to ensure that the number of events in the data samples for each mode (both simulated data and real data) had the same percentage composition from the different trigger categories (Table 5.6) due to the fact that the different triggers bias the data differently. It was found that the real data samples from the B_s^0 and B^0 modes had comparable contributions from each trigger category, whereas the simulated data had significant differences in their composition. To correct this, and to ensure the fraction of events from each trigger category in the simulated data samples were equal to those seen in data, weights were added to both the $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ simulated data samples to result in the same fraction of events from each trigger category as seen in $B_s^0 \to \psi(2S)\phi$ data. The decay time acceptance curves were then calculated using the weighted simulated data for the two modes (Figure 5.15, top), and the ratio of the two acceptance curves can be seen in Figure 5.15 (bottom).

The final step in calculating the acceptance was to multiply the acceptance histogram calculated using the $B^0 \to \psi(2S)K^*(892)$ data (Figure 5.14) by the histogram modelling the difference in the acceptance curves betweeen the $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ modes in simulation (Figure 5.15). The resulting acceptance curve is shown in Figure 5.16. The decay time acceptance is modelled in the final fit for ϕ_s as a histogram with 40 uneven bins in decay time from $(0.3 - 14)$ ps. The values for each bin in the histogram are shown in Table C.1.

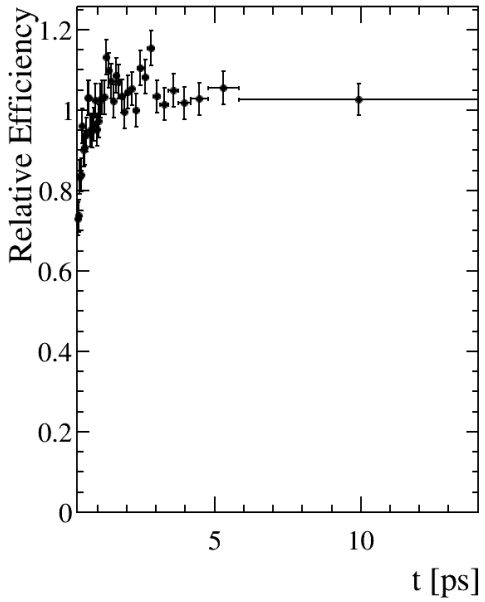


Figure 5.14 *The decay time acceptance of $B^0 \to \psi(2S)K^*(892)$ data calculated by taking the ratio of the background subtracted decay time distribution in data with the true $B^0 \to \psi(2S)K^*(892)$ decay time distribution.*

Table 5.6 *Fraction of events in each trigger category for $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ simulation and data. Here, A refers to the events coming through the* `Hlt1DiMuonHighMassDecision` *trigger while B refers to the* `Hlt1TrackMuonDecision` *and* `Hlt1TrackAllL0Decision` *triggers. The label 2011 refers to the* `Hlt2DiMuonHighMassDecision` *trigger and 2012 refers to the* `Hlt2DiMuonPsi2SDecision` *trigger.*

	$B^0 \to \psi(2S)K^*(892)$	$B_s^0 \to \psi(2S)\phi$	Difference (σ)
Data			
A 2011	0.2775 ± 0.0039	0.273 ± 0.010	0.4
B 2011	0.0456 ± 0.0014	0.0404 ± 0.0033	0.9
A 2012	0.5570 ± 0.0061	0.570 ± 0.015	1.6
B 2012	0.1199 ± 0.0024	0.1161 ± 0.0063	0.6
MC			
A 2011	0.430 ± 0.001	0.4615 ± 0.0008	-31.5
B 2011	0.0657 ± 0.0006	0.0673 ± 0.0004	-2.7
A 2012	0.408 ± 0.001	0.3838 ± 0.0008	24.2
B 2012	0.0968 ± 0.0007	0.0875 ± 0.0005	13.3

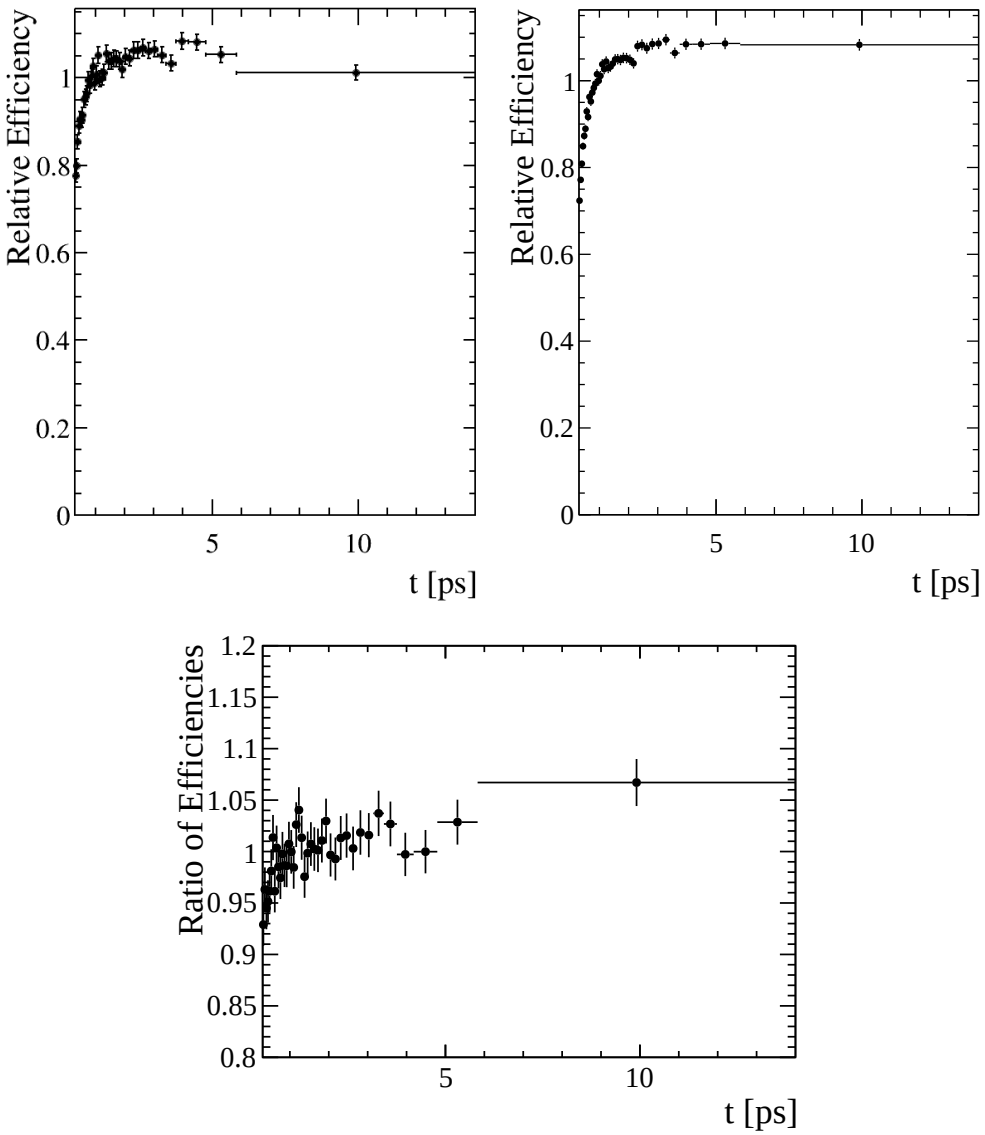


Figure 5.15 *The decay time acceptance in simulated $B^0 \to \psi(2S)K^*(892)$ events (top left) and simulated $B_s^0 \to \psi(2S)\phi$ events (top right) and the ratio of the two decay time acceptance histograms (bottom). The decay time acceptance shape for simulated $B^0 \to \psi(2S)K^*(892)$ events was found by taking the ratio of the decay time distribution in simulated with the true $B^0 \to \psi(2S)K^*(892)$ decay time distribution. To calculate the acceptance shape for simulated $B_s^0 \to \psi(2S)\phi$ the signal PDF was used to generate events using the same physics parameters as the MC dataset and the ratio of the decay time distribution in fully simulated $B_s^0 \to \psi(2S)\phi$ events was taken with this.*

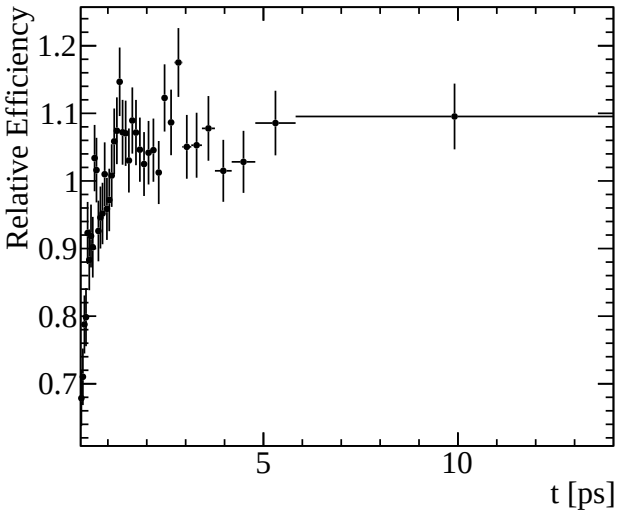


Figure 5.16 *Final decay time efficiency for the $B_s^0 \to \psi(2S)\phi$ signal events. This is found by multiplying the decay time acceptance shape found in $B^0 \to \psi(2S)K^*(892)$ data (Figure 5.14) by the ratio of the acceptance shapes found for simulated $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ events (Figure 5.15)*

.

5.3 Angular Resolution

The helicity angle distributions effectively separate the contributions from different polarisation amplitudes in B_s^0 and $\overline{B}_s^0$ decays to $\psi(2S)\phi$. The precision of the measurement of the helicity angles therefore has an effect on the measured amplitudes and associated strong phases. The angular resolution was not included in the fits to data to extract the physics parameters so, to understand the effect this has on the results, the angular resolution was modelled by looking at the difference between the true and reconstructed angular distributions in simulated data and this model was used to smear the angular distributions in toy studies. Any observed bias was taken as a systematic.

Figure 5.17 shows the difference between the true and reconstructed values of the helicity angles in simulated data, labelled $\Delta\theta_K$, $\Delta\theta_\mu$ and $\Delta\phi$, fit with a triple Gaussian model for the θ_K and θ_μ resolutions (top left and top right respectively) and a sum of three Gaussian functions and a Breit-Wigner for the ϕ (bottom). The fit results for each helicity angle are shown in Table 5.7. The effective resolution of the triple Gaussian resolution model for θ_K and θ_μ is of $\mathcal{O}(10)$ mrad and $\mathcal{O}(1)$ mrad respectively.

The systematic uncertainty from not including the angular resolution in the fit is very small, and the largest impact is seen on the amplitudes $|A_\perp|^2$, $|A_0|^2$ and strong phases $\delta_\parallel$, $\delta_\perp$. The details for the evaluation of this systematic can be found in § 7.8.

Table 5.7 *Angular resolution fit parameters. A triple Gaussian function is fitted to both the θ_K and θ_μ while for the ϕ angle the fit model is the sum of three Gaussians and a Breit-Wigner function. The mean of all functions is fixed to zero.*

Parameter	θ_K (3G)	θ_μ (3G)	ϕ (3G+BW)
σ_1 [rad]	0.0625 ± 0.0016	0.003778 ± 0.000042	0.01946 ± 0.00030
σ_2 [rad]	0.02292 ± 0.00028	0.006015 ± 0.000068	0.03933 ± 0.00080
σ_3 [rad]	0.01271 ± 0.00011	0.01079 ± 0.00027	0.0969 ± 0.0028
σ_{BW} [rad]	-	-	0.307 ± 0.011
f_3	0.581 ± 0.014	0.538 ± 0.016	0.0551 ± 0.0031
f_2	0.393 ± 0.013	0.0346 ± 0.0046	0.310 ± 0.046
f_{BW}	-	-	0.1980 ± 0.0070

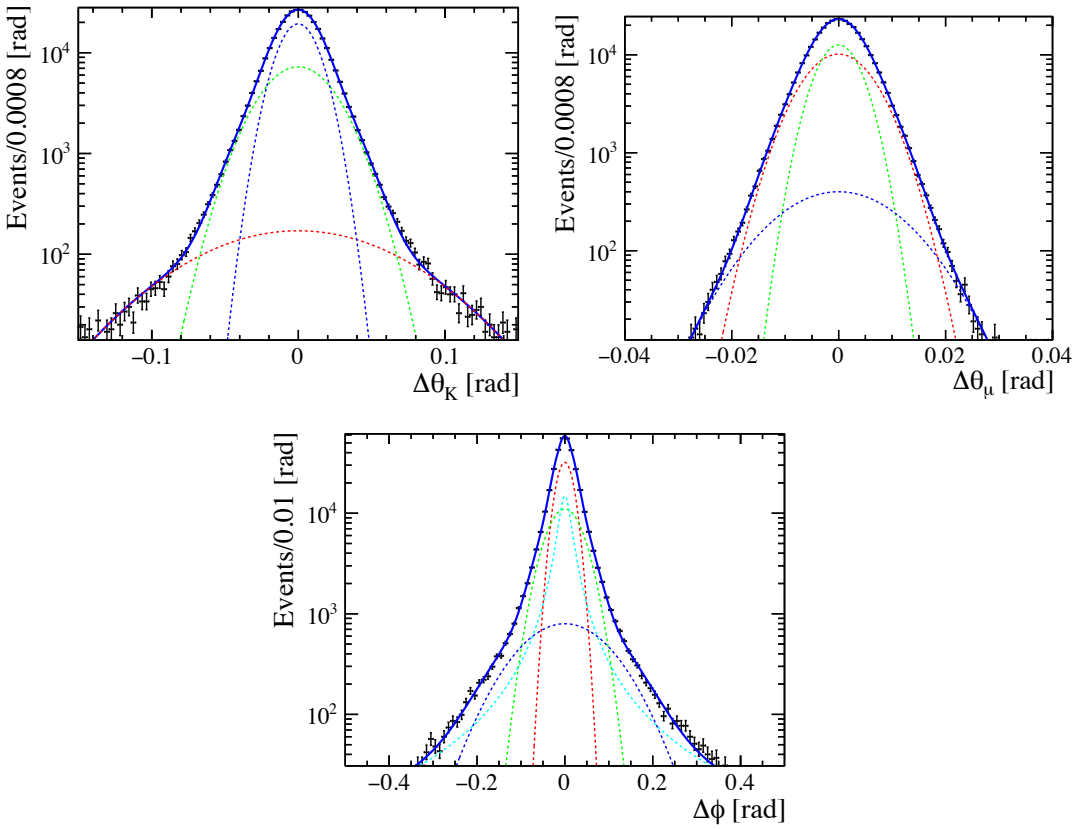


Figure 5.17 *Fit to the difference between the true and reconstructed helicity angles in simulated $B_s^0 \to \psi(2S)\phi$ events. The units are radians. A triple Gaussian function is fitted to both the θ_K and θ_μ while for the ϕ angle the fit model is the sum of three Gaussians and a Breit-Wigner function.*

5.4 Angular Acceptance

The effect of the detector acceptance and reconstruction algorithms on the helicity angle acceptance is included in the fits to data. To describe how the angular acceptance is implemented in the fit a general introduction to Negative Log-Likelihood (NLL) fits is required. This section will describe a simplified version of the full fit function to illustrate the inclusion of the angular acceptance model.

To perform an NLL fit the probability density function (PDF) must be defined. In this case, the differential decay rate of $B_s^0 \to \psi(2S)\phi$ decays is the main component. The differential decay rate was introduced in § 2.3 and has 10 angular dependent terms as defined in Equation 2.14. From this point on the differential decay rate as described in Equation 2.14, Equation 2.15 and Table 2.1 will be referred to by $\mathcal{D}(t, \Omega; \overrightarrow{P})$, where t is the observed decay time, Ω are the observed helicity angles, and $\overrightarrow{P}$ are the physics parameters. The signal PDF $S(t, \Omega; \overrightarrow{P})$ is then constructed as shown in Equation 5.7, where the numerator consists of the differential decay rate multiplied by the angular efficiency $\varepsilon_a(\Omega)$, and the denominator is the normalisation. As this section concerns the parameterisation of the angular efficiency this is the only source of efficiency that will be considered here. Resolution effects are also neglected at this point.

$$S(t, \Omega; \overrightarrow{P}) = \frac{\mathcal{D}(t, \Omega; \overrightarrow{P}) \times \varepsilon_a(\Omega)}{\iint dt\, d\Omega\, \mathcal{D}(t, \Omega; \overrightarrow{P}) \times \varepsilon_a(\Omega)} \tag{5.7}$$

The likelihood is calculated as the product over all events (N) of the values of the signal PDF. The NLL is then calculated by taking the negative log of the likelihood, turning the product into a summation and, due to the mathematical properties of the log function, the angular efficiency from the numerator can just be expressed as an additional term in the NLL (Equation 5.8). Due to the fact that the angular efficiency is not dependent upon the physics parameters this extra term acts as an offset to the NLL and does not effect where the minimum of the NLL occurs. For the fits to data this term can be ignored.

$$\text{NLL} = -\sum_i^N \log\left(\frac{\mathcal{D}(t, \Omega; \overrightarrow{P})}{\iint dt\, d\Omega\, \mathcal{D}(t, \Omega; \overrightarrow{P}) \times \varepsilon_a(\Omega)}\right) - \sum_i^N \log\Big(\varepsilon_a(\Omega)\Big) \tag{5.8}$$

The angular acceptance model in the numerator was vital however when performing PDF projections or sanity checks, and so a 3D histogram of the acceptance of the

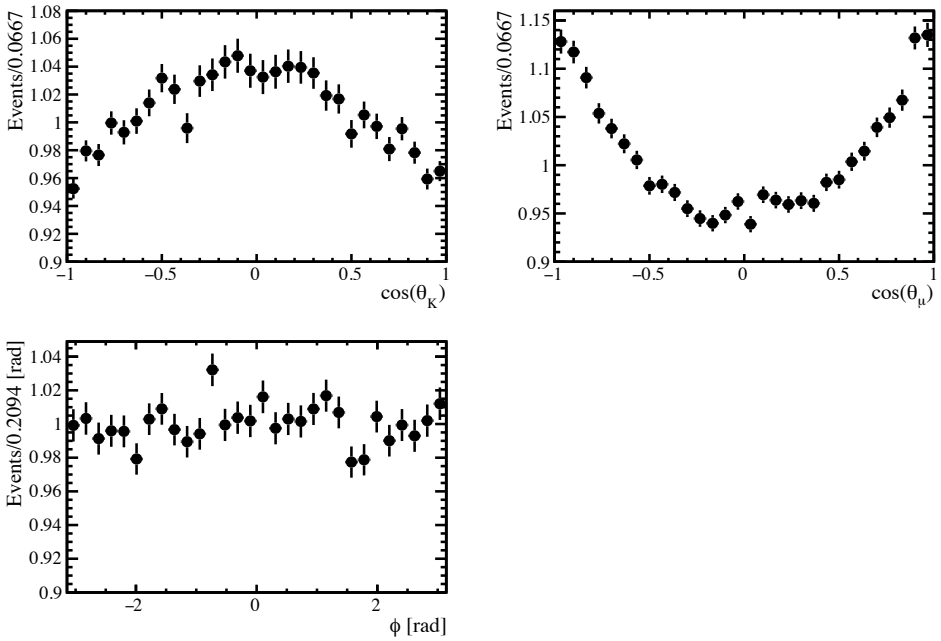


Figure 5.18 *The angular acceptance projections calculated by taking the ratio of the distributions of the helicity angles in fully simulated MC with generator level events.*

helicity angles was included for those steps. This 3D histogram was calculated by taking the ratio of the helicity angles in fully simulated signal events, and the true distribution of the helicity angles, generated using the signal PDF with the physics parameters corresponding to the simulated dataset (Table 4.2). Plots showing the projections of the acceptance for each of the helicity angles are shown in Figure 5.18.

5.4.1 Normalisation Weights

The angular acceptance was included in the PDF denominator as a set of 10 acceptance weights $\xi_j(t)$, one for each of the angular dependent terms of the signal differential decay rate $f_j(\Omega)$, see Table 2.1, as shown in Equation 5.9. It is possible to calculate the angular acceptance in this way because the decay time and angular acceptance are independent of each other and so factorise.

$$\xi_j(t) = \int d\Omega \, \varepsilon(\Omega, t) f_j(\Omega) \tag{5.9}$$

To calculate these acceptance weights from simulated $B_s^0 \to \psi(2S)\phi$ data a general derivation is followed where the efficiency $\varepsilon(\Omega, t)$ can be expressed as shown in Equation 5.10, where some extra observables $\overrightarrow{z}$ are integrated out of the

differential decay rate.

$$\varepsilon(\Omega,\ t) = \frac{\int \varepsilon\ (\Omega,\ t, \overrightarrow{z})\ \mathcal{D}(t, \Omega, \overrightarrow{z}; \overrightarrow{P})\ dz}{\mathcal{D}(t, \Omega; \overrightarrow{P})} \tag{5.10}$$

Using the fact that the following identities hold for a general function $\mathcal{G}$; $\mathcal{G}(x, y) = \int \mathcal{G}(x, y, z)dz$ and $\mathcal{G}(x, y) = \mathcal{G}(x|y)\mathcal{G}(y)$, Equation 5.10 can be substituted into Equation 5.9 to yield the expression for the angular acceptance weights shown in Equation 5.11.

$$\xi_j(t) = \int\int \frac{f_j\ (\Omega)}{\mathcal{D}(t, \Omega; \overrightarrow{P})} \varepsilon\ (\Omega,\ t, \overrightarrow{z})\ \mathcal{D}(\Omega, \overrightarrow{z}; t, \overrightarrow{P})\ dz\ d\Omega \tag{5.11}$$

To calculate the weights from a simulated data sample the term $\mathcal{D}(\Omega, \overrightarrow{z}; t, \overrightarrow{P})\ dz\ d\Omega$ in Equation 5.11 is interpreted as the probability of generating an event for a specific $(t, \overrightarrow{P})$ with observables in some small range of Ω and $\overrightarrow{z}$. Summing over the number of accepted events in the simulated data sample (N) the efficiency $\varepsilon(\Omega, \overrightarrow{z})$ can then be interpreted as having either a value of 0 or 1 depending on whether or not the individual event is accepted, resulting in Equation 5.12. To calculate the angular acceptance weights, the time is integrated out. This work has been reproduced from multiple sources and a more in depth derivation can be found in [36] and [35].

$$\xi_j = \frac{1}{N} \sum_{i=0}^{N} \frac{f_j(\Omega_i)}{\mathcal{D}(\Omega_i; t_i, \overrightarrow{P})} \tag{5.12}$$

The calculation of the angular acceptance normalisation weights was performed using simulated $B_s^0 \rightarrow \psi(2S)\phi$ data samples with the full reconstruction applied and the resulting values are shown in Table 5.8.

Table 5.8 *Angular acceptance weights calculated using fully simulated $B_s^0 \to \psi(2S)\phi$ data. The 'Value' column shows the calculated values for the weights, normalised to the first value. The final column shows the values for each term if there were no distortions to the angular distributions.*

	j	Value ξ_j/ξ_1	Perfect
1	(00)	$+1.0 \pm 0.0016$	1
2	($\parallel\parallel$)	1.0582 ± 0.0022	1
3	($\perp\perp$)	1.0585 ± 0.0024	1
4	($\parallel\perp$)	-0.0012 ± 0.0021	0
5	(0 $\parallel$)	0.0004 ± 0.0012	0
6	(0 $\perp$)	-0.0015 ± 0.0012	0
7	(SS)	1.0123 ± 0.0018	1
8	(S$\parallel$)	-0.0006 ± 0.0017	0
9	(S$\perp$)	0.0029 ± 0.0017	0
10	(S0)	-0.0049 ± 0.0033	0

5.5 Flavour Tagging

One of the key features of the LHCb detector and algorithms is the ability to infer the flavour of the mother B particle of the decay chain, either B or $\overline{B}$, called flavour tagging. This ability is important for this analysis because the terms of the differential decay rate which contain the oscillation frequency have opposite signs for B_s^0 or $\overline{B}_s^0$ decays so imperfect knowledge about the flavour of the mother results in a loss in sensitivity of ϕ_s. This section describes the methods used to enable us to make these identifications and how this information is implemented for the analysis.

There are multiple flavour tagging algorithms implemented in LHCb which use different aspects of the kinematics of a decay process to infer the flavour of a B meson. These tagging algorithms produce a tag decision (-1, 1, 0 for a $\overline{B}$, B or untagged respectively) and a value for the probability that the flavour has been incorrectly identified, the mistag probability, η. The per-event mistag probability is estimated by a neural net which is trained on simulated events and so needs to be calibrated to reflect the correct values for data, resulting in the calibrated mistag probability, ω. The performance of each tagging algorithm can be evaluated by its efficiency, ε, but due to how the flavour tagging is implemented when examining *CP* asymmetries it is better to evaluate their performance in that capacity. When

the possibility of imperfect identification of the flavour of the B is introduced the observed *CP* asymmetry, $A(t)$, dependant on the observed decay rates for the B and $\bar{B}$ mesons ($\Gamma(t)$ and $\bar{\Gamma}(t)$ respectively) is reduced compared to the true *CP* asymmetry, $A_t(t)$, by a factor called the dilution, $D = 1 - 2\omega$, as shown in Equation 5.13.

$$A(t) = \frac{\Gamma(t) - \bar{\Gamma}(t)}{\Gamma(t) + \bar{\Gamma}(t)} = (1 - 2\omega)A_t(t) = DA_t(t) \tag{5.13}$$

The uncertainty of the true *CP* asymmetry is inversely dependant on a factor called the tagging power, $\varepsilon_{eff} = \varepsilon D^2$, so to minimise the uncertainty of the parameter of interest the tagging power of the algorithm should be maximised, resulting in the optimum tagging performance.

There are two classes of flavour tagging algorithms implemented that use different pieces of information about the event to make an identification of the particle of interest (Figure 5.19). The first is the opposite side (OS) tagger [37], which uses information about the identity of the other B meson produced in association with the B particle of interest, because b quarks are produced in $b\bar{b}$ quark pairs by the pp collisions. The OS tagger can identify the flavour of the associated B meson by its daughter particles; if it has decayed semi-leptonically the charge of the final state lepton identifies the flavour of the mother meson, if there is a final state K from a $b \to c \to s$ transition the charge of the K can be used to identify the mother, and the overall charge of the secondary vertex can be used to infer the flavour of the mother. A new addition is the charm tagger [38], which uses the decay process of charm mesons to make a decision about the charge of the charm meson, and hence the flavour of the mother B meson.

The second is the same side (SS) tagger which uses information from particles that were created by the hadronisation associated with the signal B meson [39]. In the LHCb detector quarks are created in quark - antiquark pairs, so the quark that the b hadronises with to form the signal B meson has a partner that also hadronises to form another particle, for example, a pion or kaon. If this particle is charged, that charge identifies the flavour of the hadronising quark, and so the flavour of the signal B meson.

The calibration performed to extract the corrected mistag probabilities for both types of flavour tagging algorithms is done by comparing the predicted mistag probability η with the measured mistag (ω) in flavour specific decays. For the OS algorithm the calibration channel used is $B^{\pm} \to J\psi K^{\pm}$ and for the SS algorithm

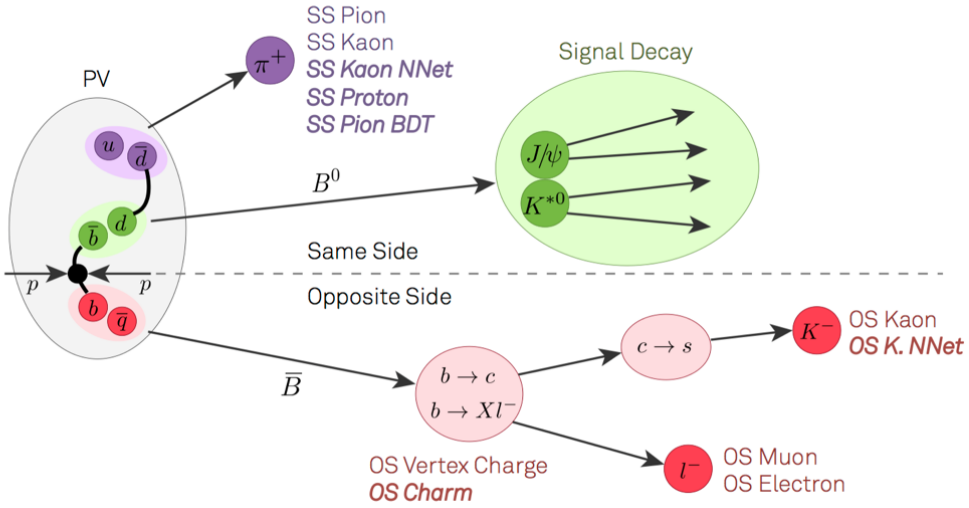


Figure 5.19 *Diagram to aid the description of flavour tagging algorigthms used at LHCb.*

the calibration channel used is $B_s^0 \to D_s^+ \pi^\mp$. The parameterisation of the mistag calibration is shown in Equation 5.14 and is performed both for the mistag probability of the B (ω) and the $\overline{B}$ ($\bar{\omega}$) for each algorithm (substitute 'OS' or 'SS' for 'alg'). The values of the calibration parameters, p_0^{alg}, p_1^{alg}, Δp_0^{alg} and Δp_1^{alg}, are shown in Table 5.9 and are the standard calibrations as provided by the Flavour Tagging group [40, 41]. The parameter $\langle \eta^{alg} \rangle$ is the average of the background subtracted mistag distribution of $B_s^0 \to \psi(2S)\phi$ data, for which the distributions are shown in Figure 5.20.

$$\begin{aligned} \omega^{alg} &= \left(p_0^{alg} + \frac{\Delta p_0^{alg}}{2} \right) + \left(p_1^{alg} + \frac{\Delta p_1^{alg}}{2} \right) (\eta^{alg} - \langle \eta^{alg} \rangle) \\ \bar{\omega}^{alg} &= \left(p_0^{alg} - \frac{\Delta p_0^{alg}}{2} \right) + \left(p_1^{alg} - \frac{\Delta p_1^{alg}}{2} \right) (\eta^{alg} - \langle \eta^{alg} \rangle) \end{aligned} \tag{5.14}$$

The systematic uncertainties of the calibration parameters are shown in Table 5.9 and are included in the single uncertainties for the mistag asymmetries ($\Delta p_0/2$, $\Delta p_1/2$). The main source of systematic uncertainty on the calibration is the difference in the kinematic distributions between the signal decay mode and the modes used for calibration. Finally, the parameters encapsulating the performance of the flavour tagging algorithms, the tagging efficiency ε, the dilution squared D^2 and the tagging power ε_{eff} are shown in Table 5.10.

Table 5.9 *Parameters used in Equation 5.14 to calibrate the per-event mistag probabilities. When two uncertainties are quoted, the first is statistical and the second is systematic.*

Parameter	OS	SS
$\langle\eta\rangle$	0.373965	0.439915
$p_0 - \langle\eta\rangle$	$0.0062 \pm 0.0019 \pm 0.0040$	$0.0070 \pm 0.0039 \pm 0.0035$
Δp_0	0.0140 ± 0.0012	$-0.0163 \pm 0.0022 \pm 0.0016$
p_1	$0.982 \pm 0.007 \pm 0.034$	$0.925 \pm 0.061 \pm 0.059$
Δp_1	0.066 ± 0.012	$-0.031 \pm 0.025 \pm 0.045$

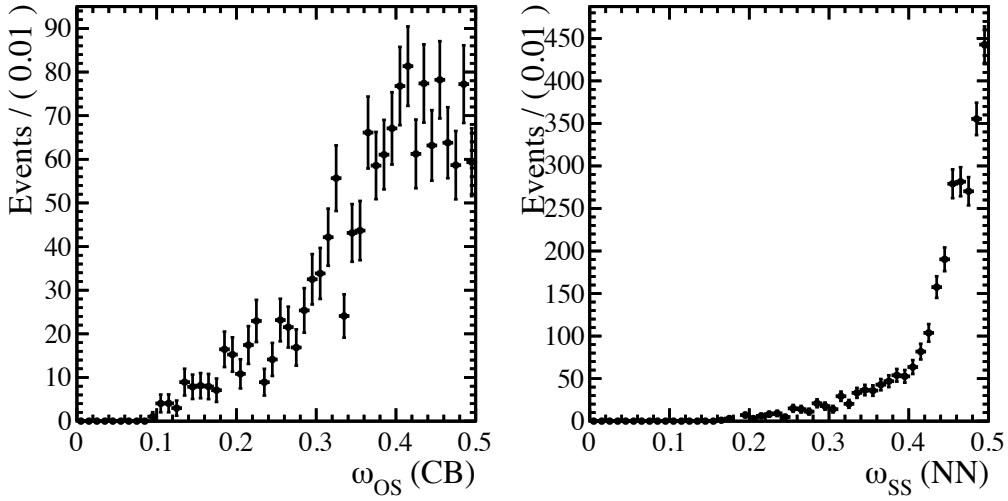


Figure 5.20 *The OS and SS mistag distributions for $B_s^0 \to \psi(2S)\phi$ signal events from $3\,fb^{-1}$ of background subtracted data.*

Table 5.10 *Summary of the tagging efficiency (ε) and tagging power (ε_{eff}) for the different type of taggers.*

Category	Fraction (%)	ε (%)	$\langle\omega\rangle$ (%)	ε_{eff} (%)
OS-only	17.56	11.52 ± 0.46	$35.20 \pm 0.15 \pm 0.37$	$1.010 \pm 0.046 \pm 0.051$
SS-only	54.68	35.89 ± 0.70	$42.00 \pm 0.41 \pm 0.32$	$0.917 \pm 0.095 \pm 0.074$
OS&SS	27.75	18.21 ± 0.56	$33.63 \pm 0.16 \pm 0.33$	$1.952 \pm 0.070 \pm 0.079$

Chapter 6

Fitting and Results

6.1 Fit Function

To extract the physics parameters a weighted negative log-likelihood (NLL) fit was performed to the data sample selected from $3\,\mathrm{fb}^{-1}$ of LHCb data, as described in chapter 4, using the *sWeights* generated from the fit to the $\mathrm{m}(\psi(2S)K^+K^-)$ mass distribution. The fit was performed using RAPIDFIT [42], a C++ code base built around the fitting framework MINUIT [43]. The signal PDF used to fit the data was constructed based on the differential decay rates for $B_s^0 \to \psi(2S)\phi$ decays, $\mathcal{D}(t, \Omega; \overrightarrow{P})$ and $\overline{B}_s^0 \to \psi(2S)\phi$, $\bar{\mathcal{D}}(t, \Omega; \overrightarrow{P})$, where t is the decay time, Ω are the helicity angles, and $\overrightarrow{P}$ are the physics parameters. The explicit function for the differential decay rate was first introduced in § 2.3 and is the sum of 10 terms dependent on the helicity angles (Table 2.1), each multiplied by a time dependent function (Equation 2.15). The full function is described in Equation 2.14. Multiple components are then introduced into the PDF to model the knowledge of the detector behaviour which has been explored in previous chapters. The following passages describe, in a step-by-step fashion, the inclusion of all of these components.

To include the flavour tagging information on whether the event was due to a B_s^0 or $\overline{B}_s^0$ decay, a tagging dependent PDF, $\mathcal{D}_{\mathrm{tag}}(t, \Omega; \overrightarrow{P})$, was created by combining the differential decay rates of the B_s^0 and $\overline{B}_s^0$ decaying to the $\psi(2S)$ and ϕ mesons, $\mathcal{D}(t, \Omega; \overrightarrow{P})$ and $\bar{\mathcal{D}}(t, \Omega; \overrightarrow{P})$ respectively, in a way that correctly combined all the information provided by the two types of flavour tagging algorithm (OS and SS). The resulting PDF is shown in Equation 6.1 where $q^{OS,SS}$ are the tagging

decisions of the OS and SS tagging algorithms, $\eta^{OS,SS}$ are the uncalibrated mistag probabilities and $\omega^{OS,SS}(\eta^{OS,SS})$ are the calibrated mistag probabilities, calculated as described in § 5.5. To compress the notation the tagging calibration parameters will be collectively referred to by $\overrightarrow{T}$ and the tagging variables will be referred to by tag.

$$\begin{aligned}\mathcal{D}_{\rm tag}(t,\Omega,tag|\overrightarrow{T};\overrightarrow{P}) =&(1+q^{OS}(1-2\omega^{OS}(\eta^{OS})))(1+q^{SS}(1-2\omega^{SS}(\eta^{SS})))\mathcal{D}(t,\Omega;\overrightarrow{P})+\\ &(1-q^{OS}(1-2\bar{\omega}^{OS}(\eta^{OS})))(1-q^{SS}(1-2\bar{\omega}^{SS}(\eta^{SS})))\bar{\mathcal{D}}(t,\Omega;\overrightarrow{P})\end{aligned} \tag{6.1}$$

The normalised signal PDF ($S(t,\Omega|\overrightarrow{r},\overrightarrow{T};\overrightarrow{P})$) was constructed by convolving the flavour tagged differential decay rate, $\mathcal{D}_{\rm tag}(t,\Omega,tag|\overrightarrow{T};\overrightarrow{P})$, with the decay time resolution, $R(t|\sigma_t;\overrightarrow{r})$, where $\overrightarrow{r}$ are the parameters of the resolution model discussed in § 5.1, and multiplying this by the acceptance efficiency, $\varepsilon(t,\Omega)$, all of which must be normalised over the phase space (Equation 6.2).

$$S(t,\Omega,tag|\overrightarrow{r},\overrightarrow{T};\overrightarrow{P}) = \frac{[\mathcal{D}_{\rm tag}(t,\Omega,tag;\overrightarrow{P})\otimes R(t|\sigma_t;\overrightarrow{r})]\times\varepsilon(t,\Omega)}{\iint dt\,d\Omega\,[\mathcal{D}_{\rm tag}(t,\Omega,tag;\overrightarrow{P})\otimes R(t|\sigma_t;\overrightarrow{r})]\times\varepsilon(t,\Omega)} \tag{6.2}$$

There were two sources of acceptance efficiency considered for this analysis, the angular acceptance and the decay time acceptance. For both cases, as the efficiencies are not dependent on the physics parameters, the efficiency in the numerator of Equation 6.2 can be ignored, except in the case where projections need to be plotted (this was shown in the discussion of the angular acceptance). The inclusion of the angular acceptance in the denominator, as a set of 10 acceptance weights multiplying each of the 10 terms of the differential decay rate, has been outlined in § 5.4.

The decay time acceptance was calculated as a histogram (H) with 40 bins ($Nbins$) spanning the range 0.3 to 14 ps as described in § 5.2. The primary reason for modelling the decay time acceptance as a histogram is that the product of the time dependent part of the differential decay rate, $h_j(t)$, with some functional form for the acceptance (dependent on t) is not analytically integrable. The decay time acceptance is therefore included in the PDF denominator by summing the integrals of the product of the differential decay rate and the decay time acceptance histogram in slices of time, as shown in Equation 6.3.

$$\sum_{n=0}^{n=Nbins-1} H_n \int_{n\,dt}^{(n+1)\,dt} \sum_{j}^{10} h_j(t)dt \tag{6.3}$$

The likelihood ($\mathcal{L}$) is defined as the product of the signal PDF values of all the events (Equation 6.4). To find the NLL, the negative log must be taken of the likelihood, as shown in the second line of Equation 6.4. To perform an *sWeighted* fit the log-likelihood of each event is multiplied by its *sWeight* (w_i), and a sum of weights correction factor, α is included to ensure the correct calculation of the variance of fit parameters.

$$\begin{aligned} \mathcal{L} &= \prod_{i}^{N} S(t_i, \Omega_i, tag_i | \overrightarrow{r}, \overrightarrow{T}; \overrightarrow{P})^{w_i} \\ -\log\Big(\mathcal{L}\Big) &= -\sum_{i}^{N} w_i \log\Big(S(t_i, \Omega_i, tag_i | \overrightarrow{r}, \overrightarrow{T}; \overrightarrow{P})\Big) \\ -\log\Big(\mathcal{L}\Big) &= -\alpha \sum_{i}^{N} w_i \log\Big(S(t_i, \Omega_i, tag_i | \overrightarrow{r}, \overrightarrow{T}; \overrightarrow{P})\Big) \\ \alpha &\approx \frac{\sum_i^N w_i}{\sum_i^N w_i^2} \end{aligned} \tag{6.4}$$

To find the physics parameters ($\overrightarrow{P}$) which minimise the NLL, the MINUIT package was used. This package uses the MIGRAD [44] minimiser which is a variable metric method (VMM) implementing an inexact line search and a stable variation of the Davidon-Fletcher-Powell metric updating scheme [45].

6.2 Error Calculation

Initial error estimation was calculated using the HESSE package (part of MINUIT), which calculates the uncertainties by calculating the second derivative matrix and inverting it, assuming parabolic errors [46]. The final uncertainties quoted for the fit parameters were calculated using MINOS which takes into account parameter correlations and non-linearities and is necessary for the parameters that do not have parabolic uncertainties [47]. The procedure for calculating the uncertainties with MINOS is to vary each parameter individually to find a value which increases the NLL by $\frac{1}{2}$, re-minimising the full NLL function in the process. This is performed for the upper and lower bound of each free parameter in the NLL function. For this thesis the majority of uncertainties quoted have been calculated using HESSE, unless explicitly indicated that they were calculated using MINOS.

6.3 External Constraints

There were several parameters included in the NLL fit that were constrained by external information. External constraints are included by adding additional terms to the NLL function of the form shown in Equation 6.5, where P is the parameter that the constraint is applied to, P_{min} is the central value, and σ_P is the uncertainty. The parameters for which external constraints are included are the flavour tagging calibration parameters and Δm_s the difference between the heavy and the light B masses. The values for these constraints are shown in Table 6.1.

$$\mathcal{L}(P) = -\left[\frac{(P - P_{min})^2}{\sigma_P^2}\right] \tag{6.5}$$

Table 6.1 *External Constraints included in the NLL fit to data.*

Parameter	Constraint
Δm_s [ps^{-1}]	17.757 ± 0.021
p_1^{OS}	0.982 ± 0.035
p_0^{OS}	0.3843 ± 0.0044
Δp_1^{OS}	0.066 ± 0.012
Δp_0^{OS}	0.014 ± 0.0012
p_1^{SS}	0.925 ± 0.085
p_0^{SS}	0.4493 ± 0.0052
Δp_1^{SS}	-0.031 ± 0.052
Δp_0^{SS}	-0.0163 ± 0.0027

6.4 Results

The results found by performing the unbinned NLL fit using the function defined in the previous section to a data sample selected from $3\,\text{fb}^{-1}$ of LHCb data (as detailed in chapter 4), including the decay time resolution, decay time and angular acceptance effects and flavour tagging calibration (with external constraints) are shown in Table 6.2. The asymmetric uncertainties that are quoted in this table have been calculated using the MINOS program. The analysis was performed blind using a blind string to disguise the true values for $\Delta\Gamma_s$ and ϕ_s in the fit, only being revealed once the systematic studies had been completed.

The motivation for the study of this mode was to make a measurement of the *CP* violating parameter ϕ_s and the difference in the decay widths of the heavy and light B_s^0 eigenstates $\Delta\Gamma_s$. The value for ϕ_s extracted by the study performed in this thesis is $\phi_s = 0.23^{+0.29}_{-0.28}$ rad, which is in agreement with both the Standard Model prediction as provided by the CKMfitter group, $\phi_s = -0.0363^{+0.0012}_{-0.0014}$ rad [17], and the combined measurements of ϕ_s from other decay modes as provided by the Heavy Flavour Averaging Group, $\phi_s = -0.034 \pm 0.033$ rad [16]. Although the statistical uncertainty of the result is large compared with most LHCb measurements from other modes, in particular when compared to the measurement using $B_s^0 \to J/\psi\,\phi$ decays for which the statistical uncertainty is 0.049 rad [13], this measurement provides a useful cross check as well as aggregating with other channels with a relatively low number of events to provide a useful addition to the precision.

The value for $\Delta\Gamma_s$ found from the NLL fit to $B_s^0 \to \psi(2S)\phi$ events as discussed in this thesis is $\Delta\Gamma_s = 0.066^{+0.041}_{-0.044}\,\text{ps}^{-1}$ which is in agreement with the LHCb measurement using $B_s^0 \to J/\psi\,\phi$ decays of $\Delta\Gamma_s = 0.0805 \pm 0.0091 \pm 0.0033\,\text{ps}^{-1}$ [13]. In addition the NLL fit extracts values for average of the decay widths of the heavy and light B_s^0 eigenstates, Γ_s, the square of the amplitudes contributing to $B_s^0 \to \psi(2S)\phi$ decays and the associated strong phases, $A_\perp^2$, A_0^2 and $\delta_\parallel$, $\delta_\perp$, the fraction of the S-wave contribution and associated strong phase, F_S and δ_S and the parameter $|\lambda|$.

Table 6.3 shows the correlations between the physics parameters resulting from the NLL fit where the largest correlations are observed between $\Delta\Gamma_s$ and the amplitudes, $A_\perp^2$ and A_0^2. The fit projections of the decay time and helicity angle distributions are shown in Figure 6.1. The one-dimensional likelihood profiles of the fit parameters can be found in Figure 6.2. These profiles are useful for

ensuring that the fit has not found a local minima, but a global one. The 2D likelihood scan of ϕ_s and $\Delta\Gamma_s$ can be found in Figure 6.3.

Table 6.2 *Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to \psi(2S)\phi$ candidates including all acceptance and resolution effects. The tagging calibration parameters and Δm_s are Gaussian constrained in the fit. The uncertainty is statistical only.*

Parameter	Value
Γ_s $[\,\mathrm{ps}^{-1}]$	0.670 ± 0.011
$\Delta\Gamma_s$ $[\,\mathrm{ps}^{-1}]$	$0.066^{+0.041}_{-0.044}$
$\lvert A_\perp \rvert^2$	$0.264^{+0.024}_{-0.023}$
$\lvert A_0 \rvert^2$	0.422 ± 0.014
$\delta_\parallel$ [rad]	$3.67^{+0.13}_{-0.18}$
$\delta_\perp$ [rad]	$3.29^{+0.43}_{-0.39}$
ϕ_s [rad]	$0.23^{+0.29}_{-0.28}$
$\lvert\lambda\rvert$	$1.045^{+0.069}_{-0.050}$
F_S	$0.061^{+0.026}_{-0.025}$
δ_S [rad]	0.03 ± 0.14

Table 6.3 *Correlation between fit parameters of the full NLL fit. The correlations larger than 0.5 are highlighted in bold.*

	Γ_s	$\Delta\Gamma_s$	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel$	$\delta_\perp$	F_S	δ_S	ϕ_s	$\lvert\lambda\rvert$
Γ_s	1.00	−0.40	0.35	−0.27	−0.08	−0.02	0.15	0.02	0.02	−0.04
$\Delta\Gamma_s$		1.00	**−0.66**	**0.60**	0.02	−0.04	−0.10	−0.02	0.19	0.03
$\lvert A_\perp\rvert^2$			1.00	**−0.54**	−0.31	−0.05	0.08	0.03	−0.02	−0.02
$\lvert A_0\rvert^2$				1.00	0.05	−0.02	−0.15	−0.02	0.07	0.03
$\delta_\parallel$					1.00	0.26	−0.26	−0.01	0.00	0.08
$\delta_\perp$						1.00	−0.21	−0.25	−0.06	0.59
F_S							1.00	0.02	0.05	−0.25
δ_S								1.00	0.07	−0.09
ϕ_s									1.00	0.04
$\lvert\lambda\rvert$										1.00

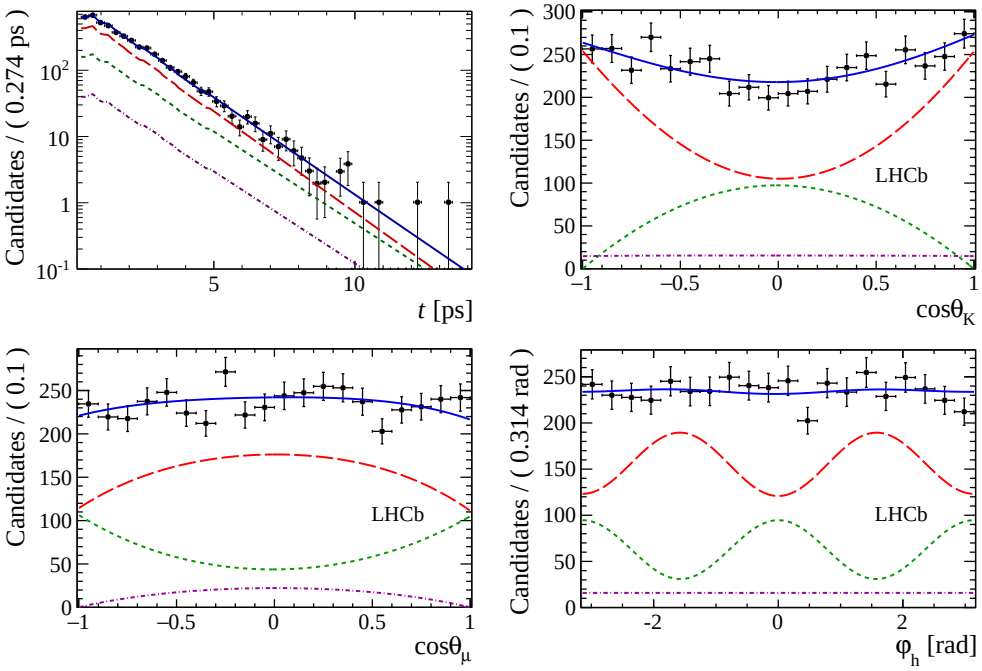


Figure 6.1 *Decay time and helicity-angle distributions for selected $B_s^0 \to \psi(2S)\phi$ events (data points) with the one-dimensional projections of the PDF at the maximal likelihood point. The solid blue line shows the total signal contribution, which is composed of CP-even (long-dashed red), CP-odd (short-dashed green) and S-wave (dash-dotted purple) contributions.*

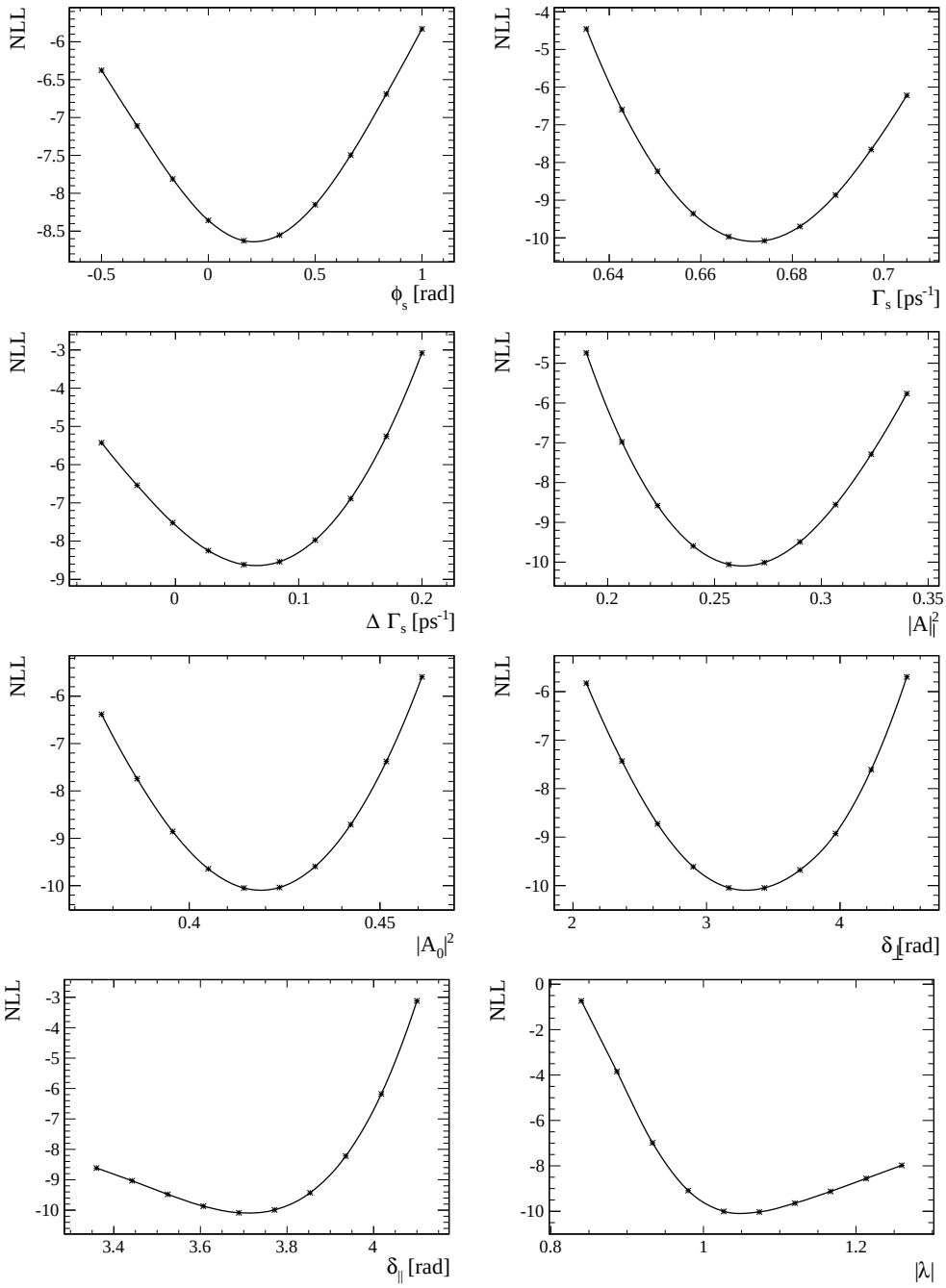


Figure 6.2 *Negative Log likelihood scans of fit parameters.*

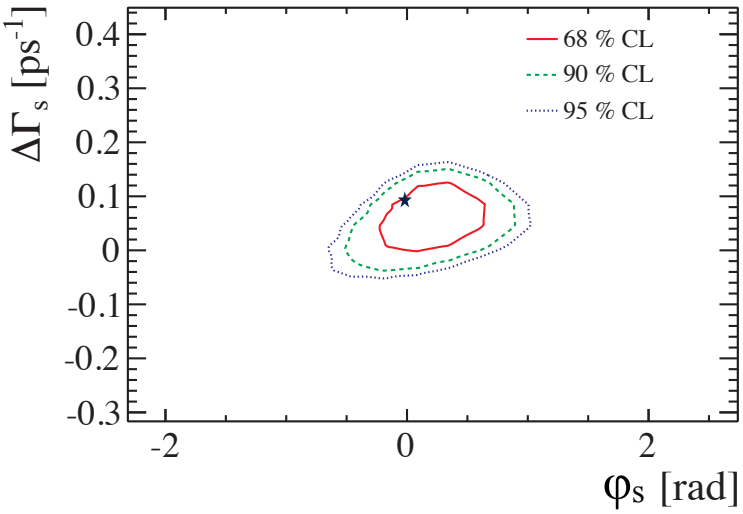


Figure 6.3 *2D Negative Log likelihood scan of ϕ_s and $\Delta\Gamma_s$. The star indicates the position of the central values for the SM prediction.*

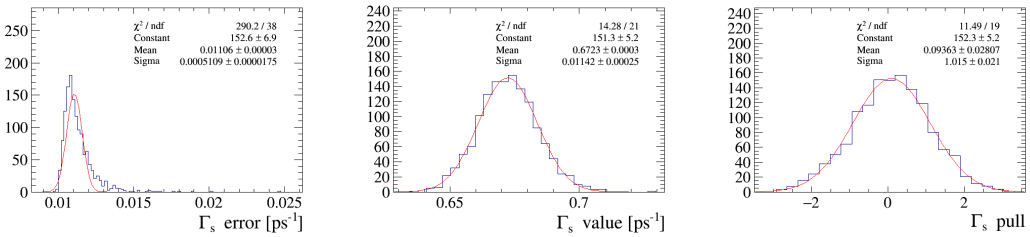
6.5 Sensitivity Studies

The expected sensitivity to the physics parameters in the NLL fit was found by performing toy studies. These studies were performed by generating 1400 sets of data using the signal PDF, with the number of events in each dataset equal to the number of signal $B_s^0 \to \psi(2S)\phi$ events found in the available LHCb data, and fitting with the same PDF. The parameters used to generate the toy datasets were those found in the fits to data (Table 6.2). The same decay time acceptance, decay time resolution and angular acceptance were used in the generation and fit stages. The mistag parameters were fixed to their average values in data.

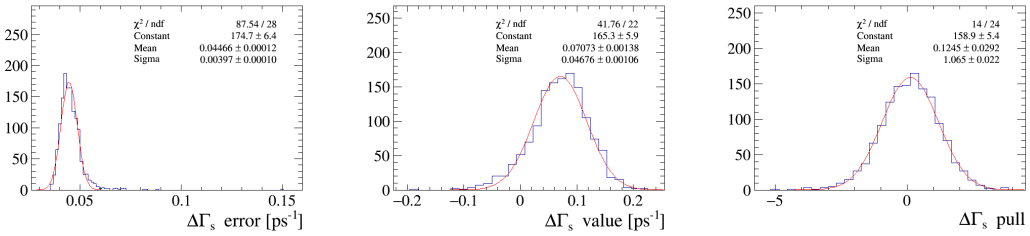
The results from the toy studies are shown in Table 6.4 where the sensitivity was found by the RMS of the fitted physics parameter values from the toy studies, and the pull was calculated as the difference between the generated values, and the mean fit value divided by the uncertainty. Plots showing the distribution of fit results, uncertainties and pulls are shown in Figure 6.4 and Figure 6.3. The expected sensitivity of most of the physics parameters matches that seen in data, with some small deviations on the second significant figure. The only exception is for $|\lambda|$ where the spread of the uncertainty shows that the difference in the error is acceptable. The evaluation of the fit bias systematic is discussed in § 7.9.

Table 6.4 *Parameter sensitivities and mean and widths of the pull distributions values from 1400 toy studies, with the number of events expected in data, fixed mistag and event resolution, and decay time and angular acceptance.*

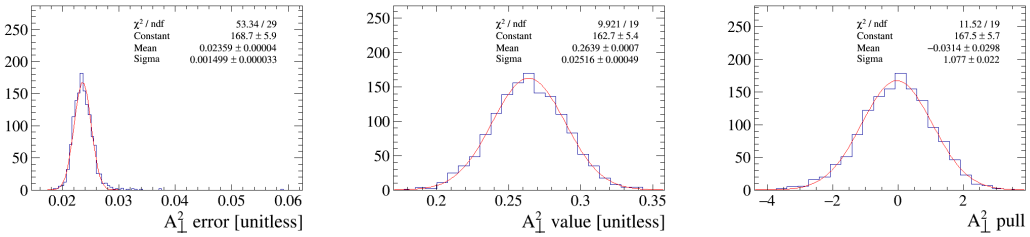
Parameter	Sensitivity	Pull mean	Pull width
Γ [ps^{-1}]	0.011	0.094 ± 0.028	1.01 ± 0.02
$\Delta\Gamma$ [ps^{-1}]	0.045	0.125 ± 0.029	1.06 ± 0.02
$\lvert A_\perp \rvert^2$	0.024	−0.031 ± 0.030	1.08 ± 0.02
$\lvert A_0 \rvert^2$	0.014	0.101 ± 0.028	1.03 ± 0.02
$\delta_\parallel$ [rad]	0.15	0.076 ± 0.028	0.89 ± 0.02
$\delta_\perp$ [rad]	0.32	0.001 ± 0.029	1.07 ± 0.02
F_S	0.025	0.104 ± 0.027	0.98 ± 0.02
δ_S [rad]	0.15	0.046 ± 0.025	0.91 ± 0.02
ϕ_s [rad]	0.25	−0.011 ± 0.026	0.95 ± 0.02
$\lvert\lambda\rvert$	0.046	0.009 ± 0.031	1.00 ± 0.02



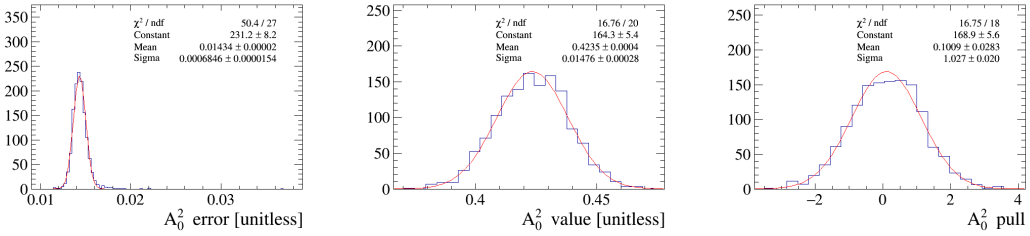
(a) Γ_s *uncertainty, value and pull*



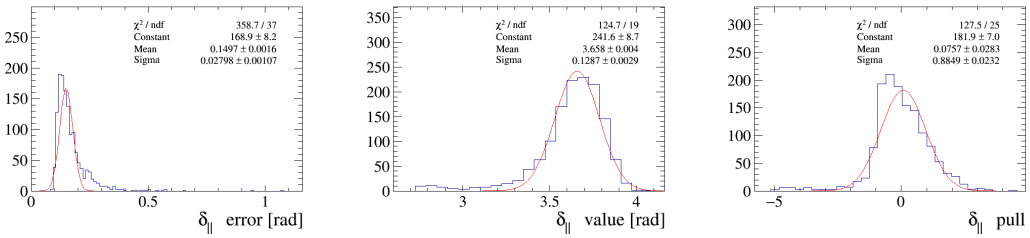
(b) $\Delta\Gamma_s$ *uncertainty, value and pull*



(c) $A_\perp^2$ *uncertainty, value and pull*

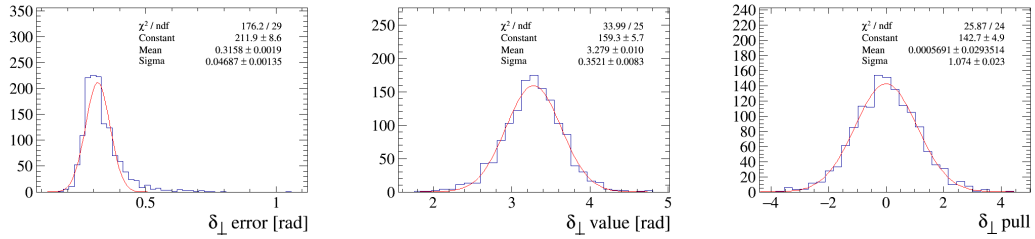


(d) A_0^2 *uncertainty, value and pull*

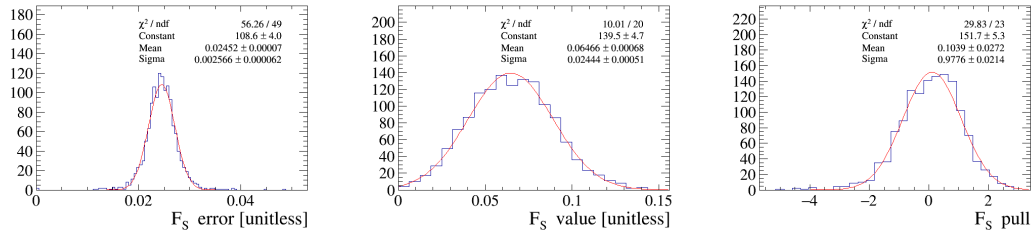


(e) $\delta_\parallel$ *uncertainty, value and pull*

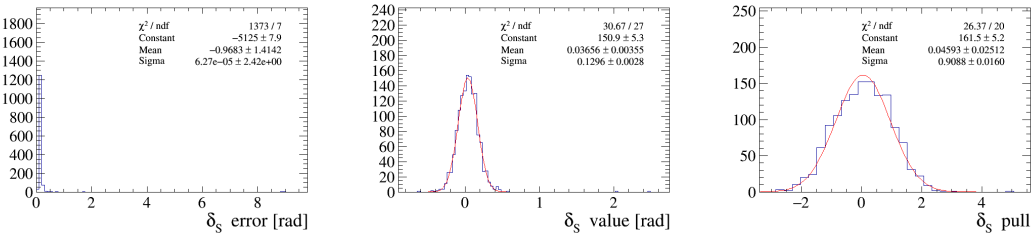
Figure 6.4 *Distribution of values, uncertainty and pulls from 1400 toy studies for* Γ_s, $\Delta\Gamma_s$, $A_\perp^2$, A_0^2 *and* $\delta_\parallel$.



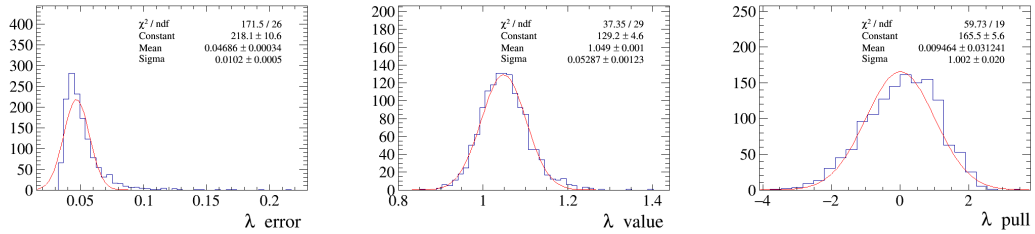
(f) *$\delta_\perp$ uncertainty, value and pull*



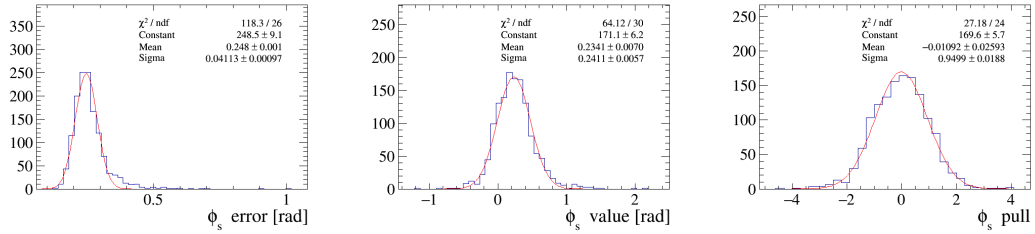
(g) *F_S uncertainty, value and pull*



(h) *δ_S uncertainty, value and pull*



(i) *$|\lambda|$ uncertainty, value and pull*



(j) *ϕ_s uncertainty, value and pull*

Figure 6.3 *Distribution of values, uncertainties and pulls from 1400 toy studies for $\delta_\perp$, F_S, δ_S, $|\lambda|$ and ϕ_s.*

Chapter 7

Systematics

Systematic uncertainties are intended to express how the choice of models or methods used in the fitting process can effect the results. This section describes the sources of systematic uncertainty that were considered for this measurement and how they were evaluated. A contribution to the systematic uncertainties from the flavour tagging calibration was not considered as the calibration parameters are allowed to float in the fit (Gaussian constrained within their uncertainty) and thus the systematic uncertainty was propagated into the statistical uncertainties reported for the results.

7.1 Factorisation of $m(\psi(2S)K^+K^-)$ with the decay time and angles

When using the $m(\psi(2S)K^+K^-)$ distribution as the variable to determine the *sWeights* for the fit there is an inherent assumption that the reconstructed B_s^0 mass distribution is uncorrelated with the decay time and helicity angle distributions. To test this assumption the data was split into 3 bins in each of the helicity angle and decay time distributions and a fit to the B_s^0 mass was performed in each bin. For this study a simplified mass model was implemented, using a student's T-function for the signal and an exponential for the background. Examining the fit parameters of the mass model a trend was found between the width (σ) of the student's T-function and the bin of $\cos(\theta_\mu)$, as can be seen in Table 7.1. This was verified using simulated data as shown in Table 7.2. To assign a systematic for

Table 7.1 *Results of the fits in bins of* $\cos(\theta_\mu)$ *to the reconstructed* $m(\psi(2S)K^+K^-)$ *distribution in data, using a Student's T-function for the signal and an exponential for the background.*

$\lvert\cos\theta_\mu\rvert$	$\in (0, 0.25)$	$\in (0.25, 0.70)$	$\in (0.70, 1.00)$
σ [MeV/c^2]	3.923 ± 0.151	4.06 ± 0.12	5.29 ± 0.19
μ [MeV/c^2]	5366.85 ± 0.14	5366.98 ± 0.11	5367.13 ± 0.17
n	3.82 ± 0.67	3.94 ± 0.53	3.79 ± 0.67

Table 7.2 *Mass model parameters from fits to the simulated* $m(\psi(2S)K^+K^-)$ *distribution using a Student's T-function. The strongest variation is observed in the mass resolution (*σ*) as a function of* $\cos\theta_\mu$*.*

$\lvert\cos\theta_\mu\rvert$	$\in (0, 0.22)$	$\in (0.22, 0.44)$	$\in (0.44, 0.66)$	$\in (0.66, 0.88)$	$\in (0.88, 1.00)$
σ [MeV/c^2]	3.287 ± 0.013	3.523 ± 0.014	3.937 ± 0.017	4.444 ± 0.020	4.954 ± 0.032
μ [MeV/c^2]	5367.06 ± 0.014	5367.00 ± 0.015	5367.08 ± 0.017	5367.05 ± 0.020	5367.06 ± 0.032
n	2.487 ± 0.022	2.476 ± 0.022	2.592 ± 0.025	2.743 ± 0.029	2.716 ± 0.042

this effect for each of the 3 bins in $\cos(\theta_\mu)$ the mass fit parameters were used to generate *sWeights* for the full dataset and the full fit was performed. The results of the fits are shown in Table 7.3. The largest difference in fit results from the three fits and the nominal results was taken as a systematic uncertainty. This source contributes a systematic to most of the measured physics parameters and has the largest effect on Γ_s, for which the associated systematic is 27% of the statistical uncertainty. To reduce this in future the *sWeights* for the analysis could be generated in bins of $\cos(\theta_\mu)$.

7.2 $m(\psi(2S)K^+K^-)$ mass model

A systematic uncertainty for the choice of mass model was evaluated by using a different model for the signal. In this case a student's T-function was chosen and is shown in Equation 7.1 where μ is the mean, σ is the width and n is the number of degrees of freedom. A fit was performed to the $m(\psi(2S)K^+K^-)$ distribution in data to generate a new set of *sWeights* (Figure 7.1).

$$\begin{aligned} f(t) &= \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \\ t &= \frac{x-\mu}{\sigma} \end{aligned} \tag{7.1}$$

Table 7.3 *Fit results when sWeights are generated using a mass shape from a fit to $m(\psi(2S)K^+K^-)$ in a single bin of $\cos\theta_\mu$. The end column shows the maximal difference with respect to the nominal fit result.*

		$\lvert\cos\theta_\mu\rvert$			
Parameter	Nominal	$\in (0, 0.25)$	$\in (0.25, 0.70)$	$\in (0.70, 1.00)$	Max diff
Γ_s [ps^{-1}]	0.672 ± 0.011	0.67 ± 0.011	0.671 ± 0.011	0.675 ± 0.011	0.003
$\Delta\Gamma_s$ [ps^{-1}]	0.071 ± 0.043	0.071 ± 0.043	0.071 ± 0.043	0.073 ± 0.043	0.002
$A_\perp^2$	0.263 ± 0.024	0.263 ± 0.024	0.263 ± 0.024	0.262 ± 0.024	0.001
A_0^2	0.419 ± 0.014	0.42 ± 0.014	0.419 ± 0.014	0.419 ± 0.015	0.001
$\delta_\parallel$ [rad]	3.71 ± 0.14	3.7 ± 0.14	3.71 ± 0.14	3.73 ± 0.13	0.02
$\delta_\perp$ [rad]	3.30 ± 0.42	3.3 ± 0.42	3.3 ± 0.42	3.30 ± 0.40	0.00
F_S	0.063 ± 0.025	0.064 ± 0.026	0.063 ± 0.026	0.060 ± 0.024	0.003
δ_S [rad]	0.04 ± 0.13	0.04 ± 0.13	0.04 ± 0.13	0.03 ± 0.12	0.01
ϕ_s [rad]	0.35 ± 0.29	0.34 ± 0.29	0.34 ± 0.29	0.34 ± 0.29	0.01
$\lvert\lambda\rvert$	1.050 ± 0.058	1.050 ± 0.059	1.051 ± 0.059	1.049 ± 0.057	0.001

The full NLL fit was performed on the dataset with new *sWeights*, and the difference between the fit results and the nominal was taken as a systematic uncertainty. This is expected to be small and contributes a systematic of less than 10% of the statistical uncertainty to the parameters Γ_s, $\Delta\Gamma_s$, and $|\lambda|$.

In addition, a systematic due to the statistical uncertainties of the mass model was evaluated. The standard method of doing this is to shift each parameter of the signal model by a randomly generated Gaussian shift with the mean equal to the parameter value and width equal to the parameter uncertainty. To take correlations between the signal model parameters into consideration a transformation was first performed into the diagonal basis of the correlation matrix. The shifts were then calculated in the transformed basis and then the values were transformed back into the original basis. It was found that when performing this method with the nominal signal model (a sum of two Crystal Ball functions) the high correlations between parameters were causing large shifts in the model parameter values. As an alternative, the method was performed using the student's T-function as the signal model. The new values of the mass model were used to generate a set of *sWeights*, and the full NLL fit was performed on the dataset. This was repeated 300 times, and the distribution of the results of each parameter was fit with a Gaussian (Figure 7.2). The width of the Gaussian is taken as a systematic and for all physics parameters this is less than 5% of the statistical uncertainties and so is not included for the final systematic uncertainties.

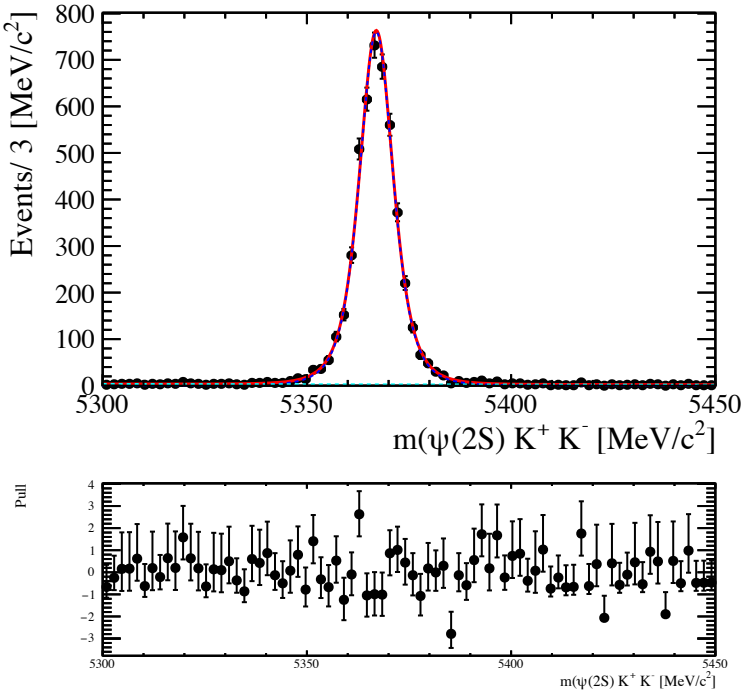


Figure 7.1 *Fit to the reconstructed B_s^0 mass distribution in 2011+2012 data after selections using a student's T-function for the signal and an exponential for the background.*

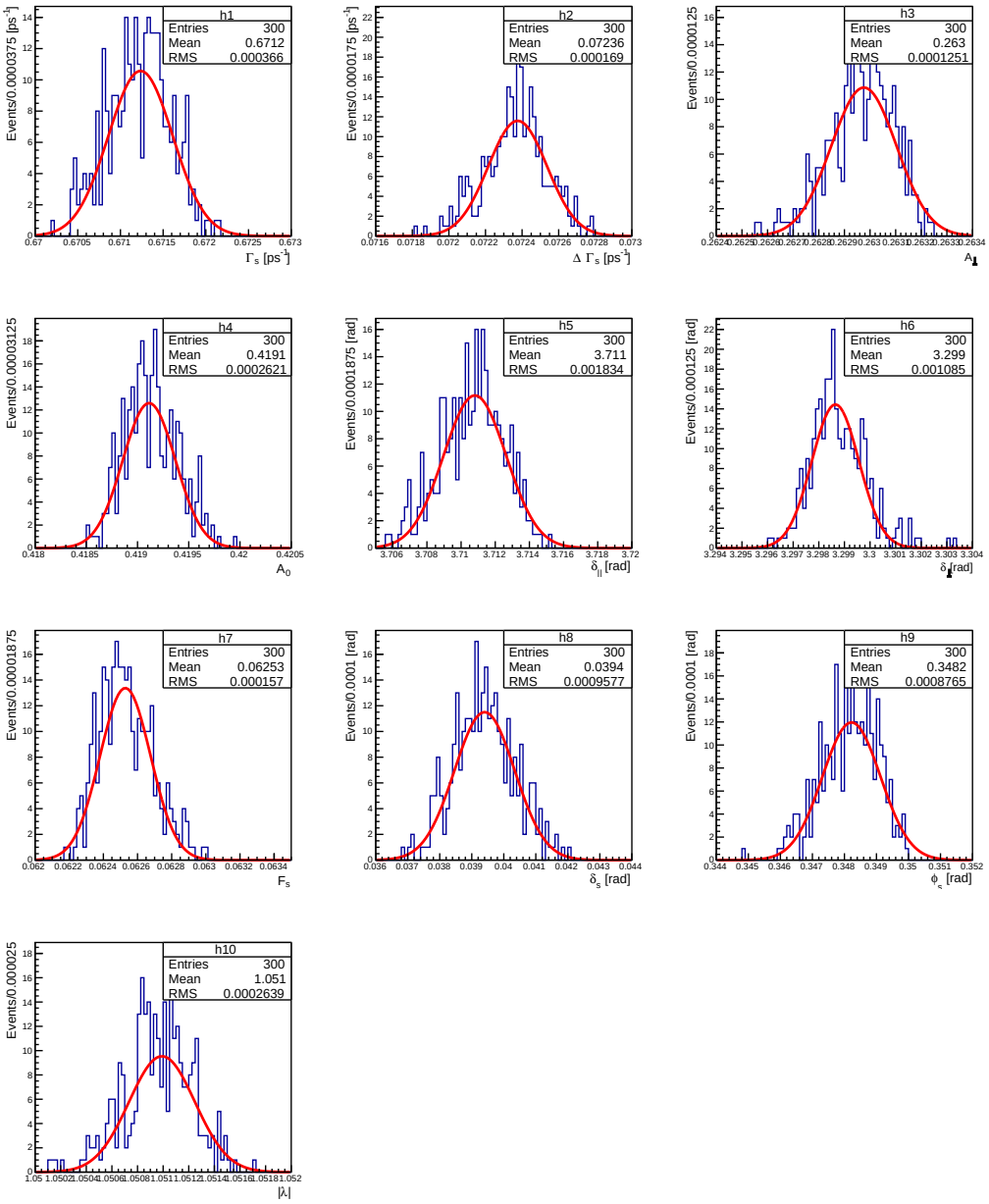


Figure 7.2 *The results of the physics parameters from 300 fits to data, varying the values of the student's T-function mass model within their uncertainty (taking into account correlations between parameters). The RMS of these distributions is taken as a systematic uncertainty.*

7.3 Angular Acceptance

To estimate a systematic due to the statistical uncertainties of the normalisation weights the same procedure as implemented for the systematic due to the statistical uncertainties of the mass model was performed. As the angular acceptance weights were correlated a transformation was performed to the diagonal basis of the correlation matrix. The transformed values of the acceptance weights were shifted by a Gaussian distributed random number with the mean equal to the transformed weight value and width equal to the transformed weight uncertainty. The shifted weights were then transformed back into their original basis and the full NLL fit was performed using the newly generated angular acceptance weights. This was repeated 300 times and the distribution of physics results were fit with a Gaussian function with the width taken as a systematic, all of which are less than 20% of the statistical uncertainties. The plots showing the distribution of results from the fits with the fluctuated values for the acceptance weights can be found in Figure 7.3.

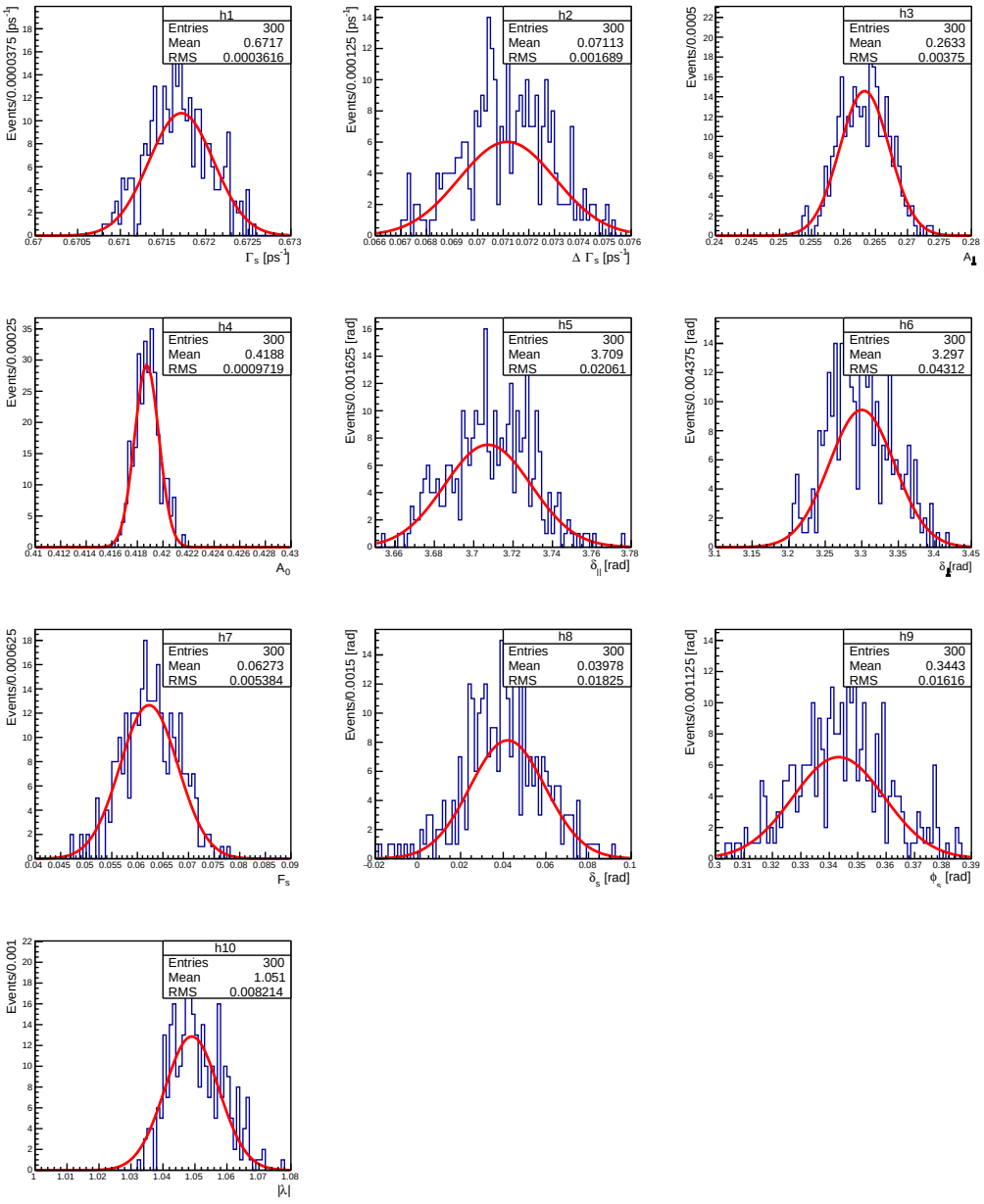


Figure 7.3 *The results of the physics parameters from 300 fits to data, varying the values of the angular acceptance weights within their uncertainty (taking into account correlations between weights). The RMS of these distributions is taken as a systematic uncertainty.*

7.4 Decay time resolution

Two sources of systematic uncertainty were considered for the decay time resolution; first due to the statistical uncertainties of the decay time resolution model and secondly due to the resolution model chosen.

For the systematic due to the statistical uncertainties of the decay time resolution model the same method as for the mass model and the angular acceptance weights was implemented. By varying the parameters of the resolution model by a randomly generated Gaussian value with the mean fixed to the parameter value and the width fixed to the uncertainty in data, taking into account correlations between the model parameters in the same way as discussed previously, and fitting the nominal dataset with the new resolution (repeated 300 times), the width of the Gaussian fit to the results gives the associated systematic. The plots showing the distribution of results from the fits with the fluctuated resolution model parameters can be found in Figure 7.4. The largest systematics from this source are seen on the strong phase $\delta_\perp$ and $|\lambda|$, and both are less than 5% of the statistical uncertainties. A systematic uncertainty due to the choice of resolution model was evaluated by fitting the prompt J/ψ data sample by the function shown in Equation 5.4 but with a single Gaussian resolution model. The width of the single Gaussian resolution model was constructed to be linearly dependent on the decay time uncertainty, as shown in Equation 7.2. The parameters c and b were fixed from the fit to the prompt J/ψ data (Figure 7.5). A set of toy experiments were then performed, generating the dataset with the nominal (double Gaussian) decay time resolution model and fitting with the single Gaussian model and the mean shift in the fit values was taken as a systematic. The decay time resolution model is the largest source of systematic uncertainty for ϕ_s (7% of the statistical uncertainty).

$$\sigma_1(\sigma_t) = c + b\sigma_t \tag{7.2}$$

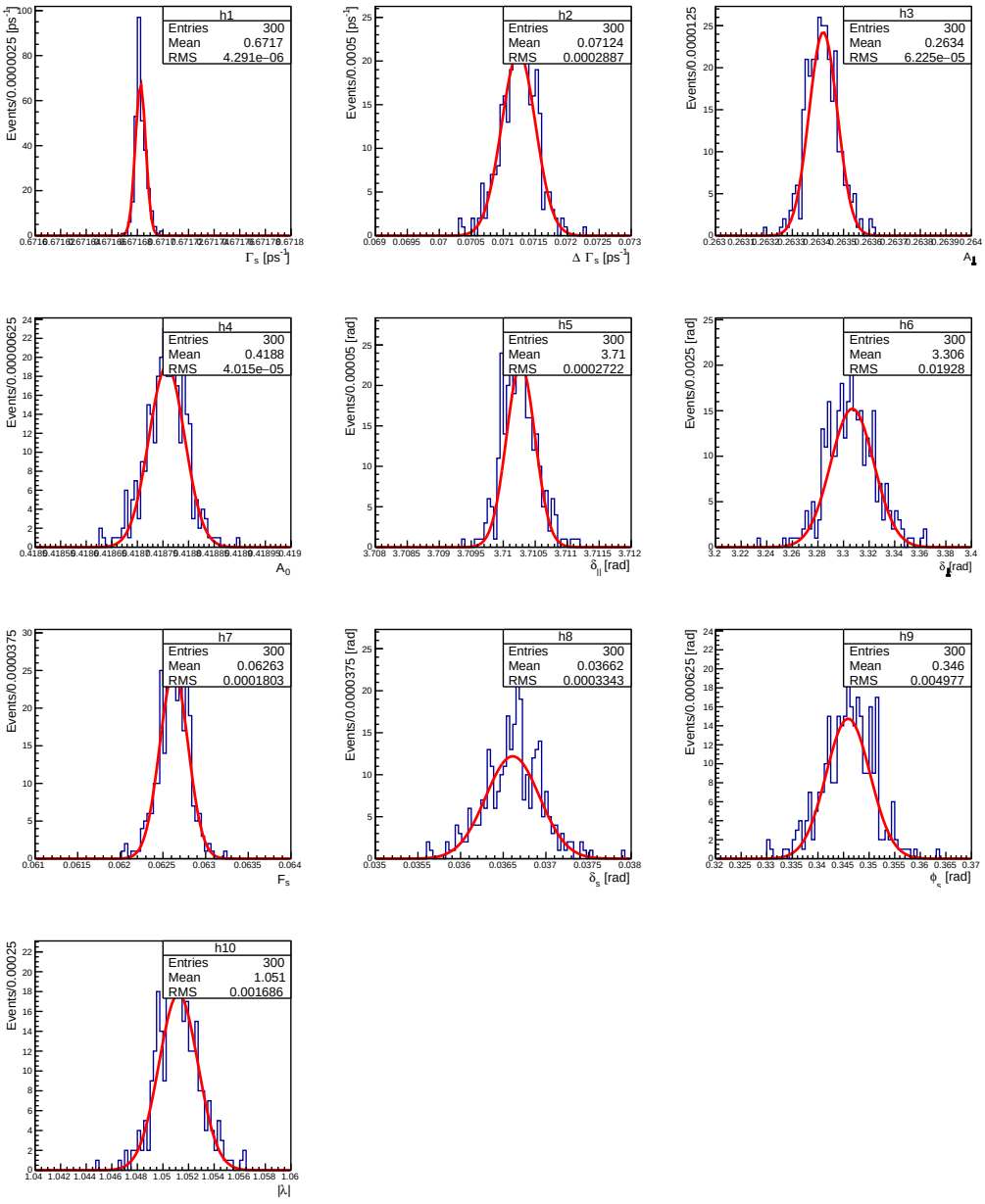


Figure 7.4 *The results of the physics parameters from 300 fits to data, varying the parameters of the time resolution within their uncertainties, and taking into account correlations between the parameters. The RMS of these distributions is taken as a systematic.*

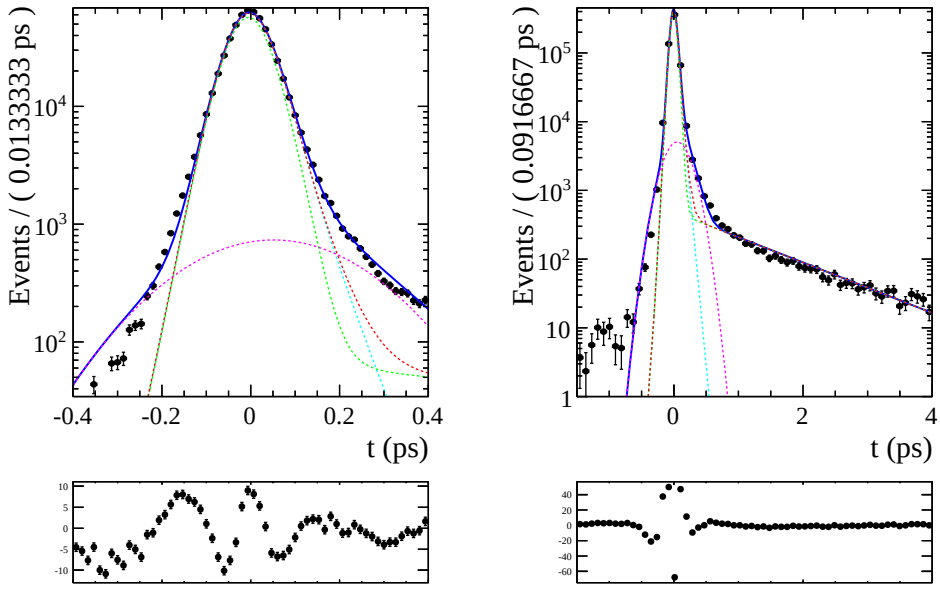


Figure 7.5 *Plot showing the fit to prompt J/ψ data using a single gaussian (linearly dependent on σ_t) resolution model. The cyan is the prompt contribution, the green and the red are the short and long lived components respectively and the magenta is the wide gaussian included to account for events with the wrong primary vertex.*

7.5 Decay Time Acceptance

A systematic due to the statistical uncertainty of the calculated decay time acceptance was estimated by repeating the NLL fit 300 times with a new decay time acceptance histogram each time. The new decay time acceptance histograms were created by generating a shifted value for each bin using a randomly generated number from a Gaussian PDF, with the mean fixed to the nominal bin value, and the width fixed to the nominal bin uncertainty. As with the other systematics from statistical sources, the distribution of the results of the physics parameters were each fit with a Gaussian and the width was taken as a systematic. The plots showing the distribution of results from the fits with the fluctuated decay time acceptance bins can be found in Figure 7.6. As the decay time acceptance was constructed using multiple sources of information, there is more than one contribution to the systematic uncertainty. A source of systematic uncertainty that was considered was due to the model used to fit the B^0 mass distribution in data. To assign a systematic uncertainty from this source, the model used for the B^0 signal was changed to a student's T-function (Figure 7.7), and the *sWeights* were recalculated and used in the process for calculating the decay time acceptance for data (outlined in § 5.2). The fit to $3\,\text{fb}^{-1}$ of LHCb data was performed using the new decay time acceptance, and the difference in fit results from the nominal was taken as a systematic.

In addition, a systematic is calculated to reflect the uncertainty of the knowledge of the B^0 lifetime. To find the decay time acceptance from B^0 data the ratio of the decay time distribution of *sWeighted* data is taken with a histogram of the decay time distribution in simulated data, generated without reconstruction or acceptance effects with the B^0 lifetime as reported in the PDG as 1.520 ± 0.004ps [48]. To evaluate this systematic the generator level events are generated using a value of the lifetime shifted $\pm 1\sigma$. The largest variation between the physics parameters obtained with this modified decay time acceptance histogram and the nominal values were assigned as a systematic uncertainty.

The sensitivity to the BDT selection was also studied. Multiple strategies were used to study how the choice of BDT cut impacts the shape of the decay time acceptance in both the $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ modes and hence the fit results. The first method was to adjust the BDT cut around the nominal point for both the signal and control samples. The decay time acceptance was then recalculated and the full tagged time-angle fit repeated. Figure 7.8 shows

the variation of the fit results for Γ_s and $\Delta\Gamma_s$ caused by changing the cut on the BDT score. No trend is observed and there is no variation outside of statistical expectation.

The second strategy implemented was to keep the BDT cut on the signal mode at the nominal point but adjust the cut on the control mode such that the ratio of the decay time acceptance shapes in the signal and control mode using simulated data is flat. The decay time acceptance histogram was then recalculated and the full fit repeated. This resulted in a variation on Γ_s of $0.0004\,\text{ps}^{-1}$ and a variation on $\Delta\Gamma_s$ of $0.00004\,\text{ps}^{-1}$, both of which are negligible compared to the statistcal uncertainties and so no systematic uncertainty was assigned related to the BDT selection.

The final test made was to check the maximum scale of the dependence of the results upon simulation. This was done by creating the decay time acceptance histogram purely from $B^0 \to \psi(2S)K^*(892)$ data, and not considering the correction due to the difference in acceptance shapes between the $B_s^0 \to \psi(2S)\phi$ and $B^0 \to \psi(2S)K^*(892)$ modes in simulation. No significant observation was observed apart from a shift to Γ_s of $0.004\,\text{ps}^{-1}$, which is 36% of its statistical uncertainty. This test was performed to give an idea of the maximum shift caused by the dependence on simulation and is an extreme check to include as a systematic and so is not considered.

7.6 Length and momentum scales

The LHCb collaboration estimates the uncertainty on the LHCb detector length scale to be at most 0.020% [49]. This uncertainty translates directly to an uncertainty on Γ_s, $\Delta\Gamma_s$ and Δm_s of 0.020%. In addition there is an uncertainty on the momentum scale of at most 0.022% [49]. The resulting effect on Γ_s and $\Delta\Gamma_s$ is negligible because it affects both the reconstructed momentum and mass of the B_s^0 meson and so cancels.

7.7 Contribution from B_c^+ decays

In addition to being created directly from p-p collisions, the signal B_s^0 meson can be created from $B_c^+ \to B_s^0\pi^+$ decays. As a consequence the reconstructed decay

time of the B_s^0 meson is overestimated because the B_c^+ also has an associated decay time. The $B_c^+ \to B_s^0\pi^+$ decay mode has been observed at LHCb and has been reported as $\sigma(B_c^+)/\sigma(B_s^0) \times B(B_c^+ \to B_s^0\pi^+) = [2.37 \pm 0.31 \pm 0.11^{+0.17}_{-0.13}] \times 10^{-3}$ [50]. Using this information, a measurement of ϕ_s performed using the $B_s^0 \to J/\psi\phi$ decay mode reported the contribution of signal events from $B_c^+ \to B_s^0\pi^+$ decays as 0.8% in Ref. [13]. This analysis also performed a study to examine the effect of this contribution on the fit results and found that only Γ_s was affected, and reported the bias found on the central value ($20 \pm 6\%$ of the statistical uncertainty) as a systematic. To use this information to assign a systematic for this analysis the assumption was made that the ratio of efficiencies for selecting $B_s^0 \to \psi(2S)\phi$ decays either promptly or via the decay of a B_c^+ meson is the same as that for $B_s^0 \to J/\psi\phi$ decays. For the final result, the central value of Γ_s is shifted to compensate for the known bias from this source ($0.002 \pm 0.001\,\text{ps}^{-1}$) and the error on this bias is assigned as a systematic.

7.8 Angular resolution

The angular resolution was not included in the NLL fit to data, so toy studies were performed to examine the effect of ignoring this. To do this toy datasets were generated using the signal PDF with the same number of events as in data. The angular distribution in the generated dataset was then smeared by some random shift generated using the resolution model described in § 5.3, resulting in a new dataset. The datasets were then fit using the nominal PDF, and the distribution of the difference between fit results, on a toy-by-toy basis, was examined for evidence of bias. Any bias resulting from the smearing of the helicity angles is taken as a systematic. The largest systematics due to ignoring the angular resolution are seen on $|A_\perp|^2$, $\delta_\perp$ and $\delta_\parallel$, however these are all only a few percent of the statistical uncertainty, 4%, 6% and 2% respectively and are not the largest contributions to the systematic uncertainty. The plots showing the distribution of results from the fits with the smeared helicity angles can be found in Figure 7.9.

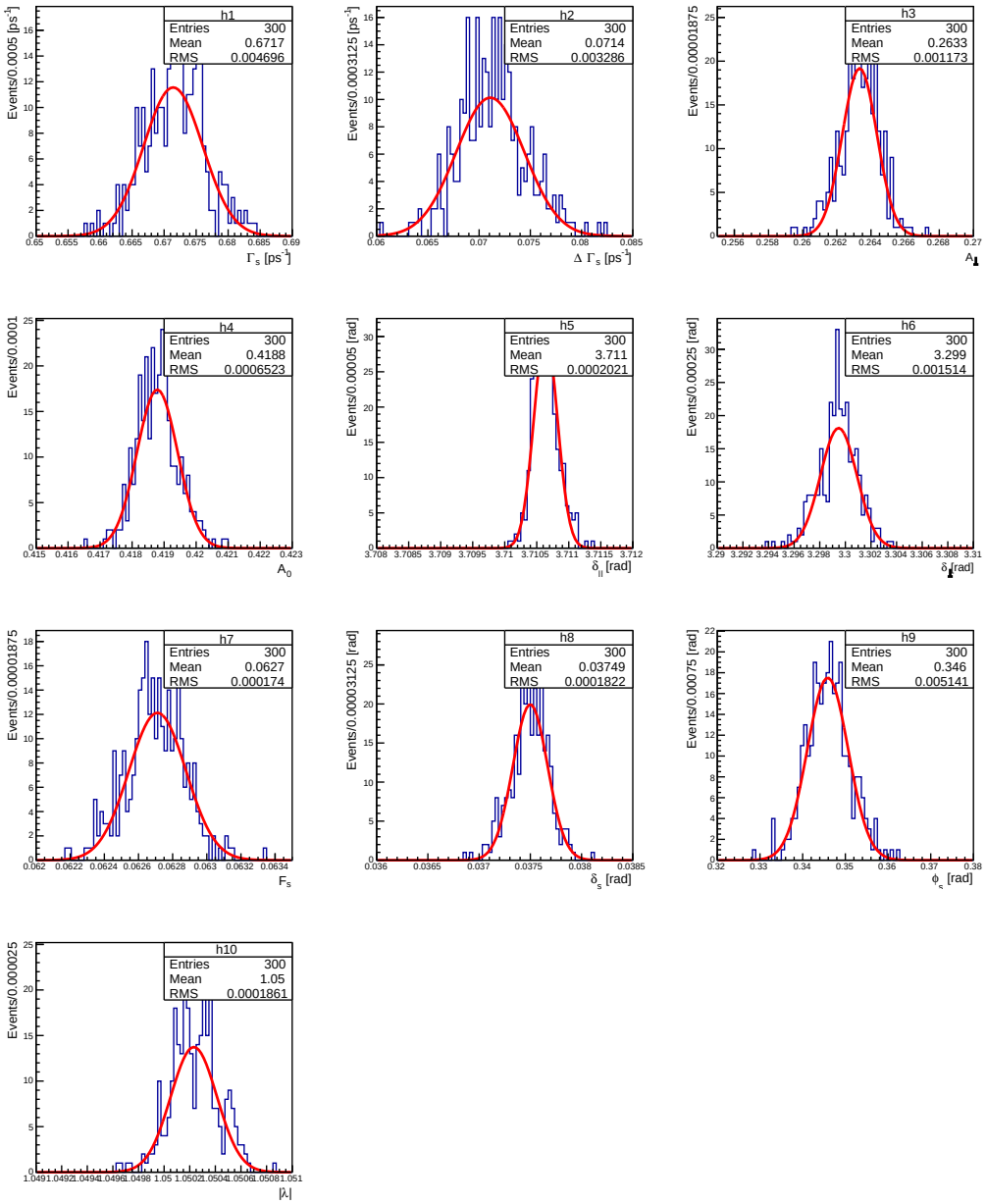


Figure 7.6 *The results of the physics parameters from 300 fits to data, varying the bins of the time acceptance histogram within their uncertainties. The RMS of these distributions is taken as a systematic uncertainty.*

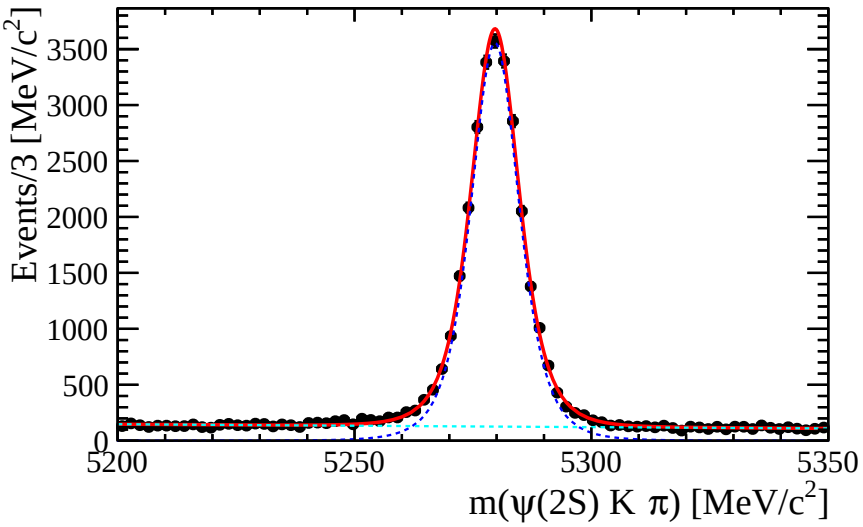


Figure 7.7 *Fit to the $m(\psi(2S)K\pi)$ distribution using a student's T-function.*

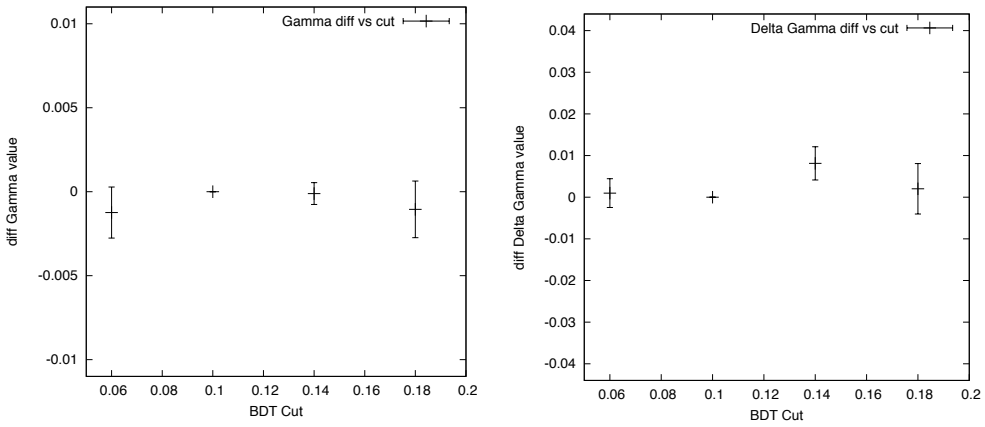


Figure 7.8 *Variation in the fit result for Γ_s (left) and $\Delta\Gamma_s$ (right) between the nominal fit and the fit performed with a different cut on the BDT for both the signal and control channel. The nominal cut is > 0.1. The y-axis range was of each plot corresponds to the nominal statistical uncertainty for Γ_s or $\Delta\Gamma_s$. The error bar for each point is calculated as the quadrature difference of the uncertainty between the two fit results, in order to remove statistical correlations.*

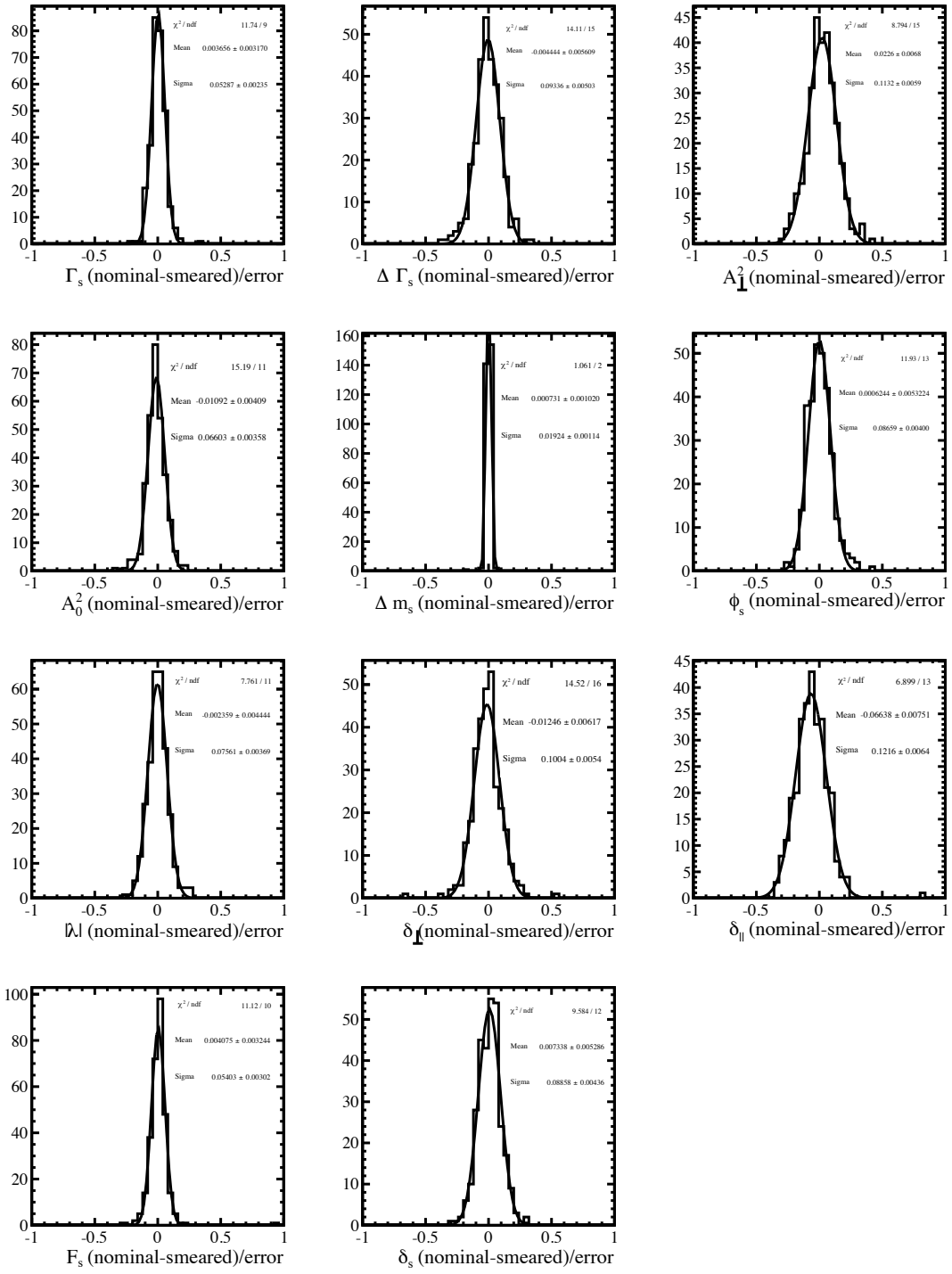


Figure 7.9 *Distributions of the difference of the fitted parameters divided by their statistical uncertainty obtained from a fit to a dataset first without (nominal) and then with (smeared) angular resolution applied to the dataset. 300 toy experiments were used, each containing 4750 events with time efficiency included in the fit and equal numbers of tagged and untagged events and average mistags as observed in the data.*

7.9 Fit bias

The possibility of introducing a bias on the fit results by the fitting procedure itself was examined using toy studies. These toy studies are detailed in § 6.5. Any parameter for which the mean of the pull is larger that 2σ is included as a systematic in Table 7.8.

7.10 Nuisance CP asymmetries

It is possible that the measurement of CP violation in this mode, caused by interference between B_s^0–$\overline{B}_s^0$ mixing and decay could be altered by other CP asymmetries, for example, CP violation in the mixing, direct CP violation in the decay, production asymmetry or tagging asymmetry. To reduce the risk of these external factors contaminating this measurement a number of precautions were implemented including; the fit parameter $|\lambda|$ being used, using separate tagging calibrations for B_s^0 and $\overline{B}_s^0$ decisions and separate normalisations of the PDF for each tagging decision. The contamination from this source was shown to be negligible in a related analysis [35].

7.11 Further checks

Additional checks on the data consistency were performed by repeating the nominal fit to data in bins of year of data taking (Table 7.4), magnet polarity (Table 7.5), number of primary vertices (Table 7.6) in each event and B_s^0 p_{T} (Table 7.7). All sub-sets were found to be consistent with the nominal fit to the full data set. For some parameters in Table 7.4, where the fit is performed in bins of run period, the nominal result does not lie in between the values found by the binned fits. This is due to the difference in the number of events in the two datasets (the 2012 dataset has roughly double the number of events) and also because the 2011 dataset has only ~ 1600 events which affects the stability of the fit.

Table 7.4 *Fit results in the run periods of 2011 and 2012.*

Parameter	Nominal	2011	2012
Γ_s [ps^{-1}]	0.672 ± 0.011	0.669 ± 0.023	0.676 ± 0.014
$\Delta\Gamma_s$ [ps^{-1}]	0.071 ± 0.043	0.05 ± 0.10	0.067 ± 0.055
$A_\perp^2$	0.263 ± 0.024	0.257 ± 0.044	0.271 ± 0.029
A_0^2	0.419 ± 0.014	0.440 ± 0.032	0.407 ± 0.017
$\delta_\parallel$ [rad]	3.71 ± 0.14	3.76 ± 0.24	3.75 ± 0.16
$\delta_\perp$ [rad]	3.30 ± 0.42	3.71 ± 0.80	3.44 ± 0.49
F_S	0.063 ± 0.025	0.059 ± 0.066	0.052 ± 0.031
δ_S [rad]	0.04 ± 0.13	-0.20 ± 0.27	0.12 ± 0.18
ϕ_s [rad]	0.35 ± 0.29	-0.02 ± 0.84	0.41 ± 0.34
$\lvert\lambda\rvert$	1.050 ± 0.058	1.13 ± 0.20	1.050 ± 0.073

Table 7.5 *Fit results in bins of magnet polarity.*

Parameter	Nominal	Magnet Up	Magnet Down
Γ_s [ps^{-1}]	0.672 ± 0.011	0.691 ± 0.019	0.672 ± 0.015
$\Delta\Gamma_s$ [ps^{-1}]	0.071 ± 0.043	-0.032 ± 0.095	0.079 ± 0.057
$A_\perp^2$	0.263 ± 0.024	0.290 ± 0.036	0.287 ± 0.034
A_0^2	0.419 ± 0.014	0.397 ± 0.020	0.419 ± 0.020
$\delta_\parallel$ [rad]	3.71 ± 0.14	3.81 ± 0.18	2.85 ± 0.47
$\delta_\perp$ [rad]	3.30 ± 0.42	3.38 ± 0.53	3.15 ± 0.60
F_S	0.063 ± 0.025	0.061 ± 0.042	0.092 ± 0.040
δ_S [rad]	0.04 ± 0.13	0.06 ± 0.21	-0.03 ± 0.15
ϕ_s [rad]	0.35 ± 0.29	-0.7 ± 1.1	0.83 ± 0.49
$\lvert\lambda\rvert$	1.050 ± 0.058	1.15 ± 0.12	0.980 ± 0.059

Table 7.6 *Fit results in bins of number of PVs.*

Parameter	Nominal	1	2	3 or more
Γ_s [ps^{-1}]	0.672 ± 0.011	0.668 ± 0.021	0.727 ± 0.056	0.689 ± 0.022
$\Delta\Gamma_s$ [ps^{-1}]	0.0713 ± 0.043	0.059 ± 0.071	0.29 ± 0.13	-0.001 ± 0.089
$A_\perp^2$	0.263 ± 0.024	0.227 ± 0.040	0.309 ± 0.043	0.272 ± 0.043
A_0^2	0.419 ± 0.014	0.419 ± 0.024	0.418 ± 0.027	0.402 ± 0.026
$\delta_\parallel$ [rad]	3.71 ± 0.14	3.95 ± 0.17	3.73 ± 0.20	3.39 ± 0.45
$\delta_\perp$ [rad]	3.30 ± 0.42	4.40 ± 0.65	2.36 ± 0.54	2.93 ± 0.51
F_S	0.063 ± 0.025	0.069 ± 0.048	0.016 ± 0.022	0.093 ± 0.047
δ_S [rad]	0.04 ± 0.13	-0.01 ± 0.25	0.57 ± 0.44	-0.19 ± 0.22
ϕ_s [rad]	0.35 ± 0.29	0.35 ± 0.47	1.44 ± 0.34	-0.30 ± 0.70
$\lvert\lambda\rvert$	1.050 ± 0.058	1.06 ± 0.12	1.06 ± 0.16	1.06 ± 0.074

Table 7.7 *Fit results in bins of B_s^0 $p_{\rm T}$ [MeV/c]; 1) <3600, 2) 3600<$p_{\rm T}$ <6500; 3) >6500.*

Parameter	Nominal	1	2	3
Γ [ps^{-1}]	0.672 ± 0.011	0.641 ± 0.023	0.677 ± 0.022	0.710 ± 0.015
$\Delta\Gamma$ [ps^{-1}]	0.0713 ± 0.043	0.056 ± 0.078	0.155 ± 0.085	0.039 ± 0.074
$A_\perp^2$	0.263 ± 0.024	0.264 ± 0.048	0.187 ± 0.038	0.326 ± 0.038
A_0^2	0.419 ± 0.014	0.399 ± 0.028	0.428 ± 0.024	0.413 ± 0.024
$\delta_\parallel$ [rad]	3.71 ± 0.14	3.57 ± 0.35	4.00 ± 0.14	2.85 ± 0.38
$\delta_\perp$ [rad]	3.30 ± 0.42	4.58 ± 0.76	3.13 ± 0.58	2.64 ± 0.40
F_S	0.063 ± 0.025	0.039 ± 0.041	0.042 ± 0.031	0.117 ± 0.038
δ_S [rad]	0.04 ± 0.13	-0.27 ± 0.45	-0.07 ± 0.30	0.08 ± 0.13
ϕ_s [rad]	0.35 ± 0.29	-0.20 ± 0.58	0.00 ± 0.33	0.44 ± 0.37
$\lvert\lambda\rvert$	1.050 ± 0.058	1.21 ± 0.13	1.10 ± 0.083	0.978 ± 0.050

7.12 Summary

Table 7.8 shows the different sources of systematic uncertainty and the size of these systematic uncertainties on the measured physics parameters. For all the parameters measured the uncertainty is dominated by the statistical uncertainty due to the relatively small number of $B_s^0 \to \psi(2S)\phi$ events in $3\,\text{fb}^{-1}$ of LHCb data. The parameter that has the largest systematic, in terms of percentage of the statistical uncertainty, is Γ_s where the total systematic uncertainty is 55% of the statistical uncertainty. This systematic is driven largely by the decay time acceptance systematic from the number of $B^0 \to \psi(2S)K^*(892)$ events available in the LHCb data sample and also by the systematic contribution from the mass factorisation check.

For the parameter of interest to this thesis, ϕ_s, the largest systematic comes from the uncertainty assigned due to the choice of decay time resolution model. Even though this is the largest systematic source for ϕ_s, it is small compared to the statistical uncertainty (7%) which dominates the total uncertainty for this parameter.

Another parameter of interest is $\Delta\Gamma_s$, as the measurement of this parameter using the $B_s^0 \to \psi(2S)\phi$ decay mode is the first to verify the measurement using $B_s^0 \to J/\psi\,\phi$ events. The uncertainty of this measurement is dominated by the statistical uncertainty, and the total systematic uncertainty is 16% of the statistical.

Table 7.8 *Statistical and systematic uncertainties. Systematic uncertainties larger than 20% of the statistical uncertainty are highlighted in bold.*

Source	Γ_s [ps^{-1}]	$\Delta\Gamma_s$ [ps^{-1}]	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel$ [rad]	$\delta_\perp$ [rad]	ϕ_s [rad]	$\lvert\lambda\rvert$	F_S	δ_S [rad]
Stat. uncertainty	0.011	$^{+0.041}_{-0.044}$	$^{+0.024}_{-0.023}$	0.014	$^{+0.13}_{-0.18}$	$^{+0.43}_{-0.39}$	$^{+0.29}_{-0.28}$	$^{+0.069}_{-0.050}$	$^{+0.026}_{-0.025}$	0.14
Mass factorisation ($\cos\theta_\mu$)	**0.003**	0.002	0.001	0.001	0.02	–	0.01	0.001	0.003	0.01
Mass model	0.001	0.001	–	–	–	–	–	0.001	–	–
Mass sWeights (stat.)	–	–	–	–	–	–	–	–	–	–
Ang. acc. (stat.)	–	0.001	0.001	0.002	0.02	0.03	0.01	0.006	0.005	0.02
Time resolution model	–	0.001	–	–	–	0.02	0.02	0.002	0.002	–
Time resolution (stat.)	–	–	–	–	–	0.02	–	0.002	–	–
Time acc. (stat.)	**0.005**	0.003	0.001	0.001	–	–	–	–	0.002	–
Time acc. (B^0 mass model)	0.001	0.001	–	–	–	–	–	–	–	–
Time acc. ($\tau_{B^0} \pm \sigma$)	0.002	–	–	–	–	–	–	–	–	–
Length, mom. scales	–	–	–	–	–	–	–	–	–	–
B_c^+ background	0.001	–	–	–	–	–	–	–	–	–
Angular resolution	–	–	0.001	–	0.01	0.01	–	–	–	–
Fit bias	0.001	0.006	–	0.001	0.01	–	–	–	0.003	–
Quad. sum of syst.	0.006	0.007	0.002	0.003	0.03	0.04	0.02	0.007	0.007	0.02
Total uncertainties	0.013	$^{+0.042}_{-0.045}$	$^{+0.024}_{-0.023}$	0.014	$^{+0.13}_{-0.18}$	$^{+0.43}_{-0.39}$	$^{+0.29}_{-0.28}$	$^{+0.069}_{-0.050}$	$^{+0.027}_{-0.026}$	0.14

Chapter 8

Conclusion

This thesis presents the first measurement of the physics parameters ϕ_s and $\Delta\Gamma_s$ in $B_s^0 \to \psi(2S)\phi$ decays using $3\,\text{fb}^{-1}$ of LHCb data. These results provide an important cross-check of the multiple preceeding measurements of ϕ_s in different decay channels and provides the only corroboration of the measured value of $\Delta\Gamma_s$ in $B_s^0 \to J/\psi\,\phi$ decays. The final results, where the first uncertainty is statistical and the second is systematic, are shown in Table 8.1. These results were found by performing an unbinned maximum likelihood fit to selected $B_s^0 \to \psi(2S)\phi$ candidates including detector acceptance and resolution effects. The flavour tagging parameters and the B_s^0-$\overline{B}_s^0$ oscillation frequency Δm_s were Gaussian constrained in the fit. The statistical uncertainty dominates the total uncertainty of the measured results, a situation which will improve with the next run of data taking at the LHC.

The final results are in good agreement with the SM prediction and the combination of previous measurements provided by the Heavy Flavour Averaging Group (HFAG), covering the values of ϕ_s and $\Delta\Gamma_s$ within 1σ of the statistical uncertainty. This indicates that there is no observable evidence of new physics in the measurement of these physics parameters in $B_s^0 \to \psi(2S)\phi$ decays. The comparison of this result and the combined results from previous measurements provided by HFAG can be seen in Figure 8.1.

With the inclusion of the next run of data taken at the LHC, which will increase the size of the LHCb dataset to $50\,\text{fb}^{-1}$, the predicted sensitivity to ϕ_s in $B_s^0 \to \psi(2S)\phi$ decays will improve to 0.04 rad providing a valuable contribution to the precision of the combination of measurements. Due to the much larger branching fraction

of the $B_s^0 \to J/\psi\,\phi$ decay mode with respect to others, the sensitivity of the combination will be driven by the reduction in the statistical uncertainty from this channel. This additional data is expected to improve the sensitivity to ϕ_s in $B_s^0 \to J/\psi\,\phi$ decays to 0.008 rad, which is comparable to the uncertainty of the value provided by the CKMfitter group of $\pm^{0.0012}_{0.0014}$ rad. With this improved precision the LHCb collaboration will be in the position to measure any small deviation from the SM prediction caused by new physics.

Table 8.1 *Results of the unbinned maximum likelihood fit to the selected $B_s^0 \to \psi(2S)\phi$ candidates including all acceptance and resolution effects. The tagging calibration parameters and Δm_s are Gaussian constrained in the fit. The first uncertainty is statistical and the second is systematic. The shift in the central value of Γ_s has been included to account for the bias from the B_c^+ contribution.*

Parameter	Value
Γ_s [ps^{-1}]	$0.668 \pm 0.011 \pm 0.006$
$\Delta\Gamma_s$ [ps^{-1}]	$0.066^{+0.041}_{-0.044} \pm 0.007$
$\lvert A_\perp\rvert^2$	$0.264^{+0.024}_{-0.023} \pm 0.002$
$\lvert A_0\rvert^2$	$0.422 \pm 0.014 \pm 0.003$
$\delta_\parallel$ [rad]	$3.67^{+0.13}_{-0.18} \pm 0.04$
$\delta_\perp$ [rad]	$3.29^{+0.43}_{-0.39} \pm 0.04$
ϕ_s [rad]	$0.23^{+0.29}_{-0.28} \pm 0.02$
$\lvert\lambda\rvert$	$1.045^{+0.069}_{-0.050} \pm 0.007$
$F_{\rm S}$	$0.061^{+0.026}_{-0.025} \pm 0.007$
$\delta_{\rm S}$ [rad]	$0.03 \pm 0.14 \pm 0.02$

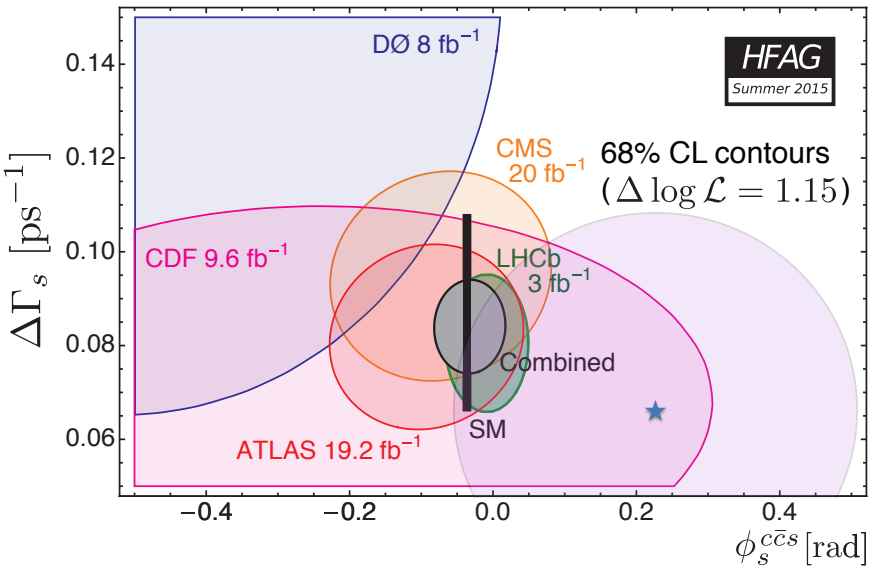


Figure 8.1 *Comparison of the ϕ_s vs. $\Delta\Gamma_s$ plane of the combined results of previous measurements as performed by the HFAG and the measurement from this analysis. The blue star indicates the $B_s^0 \to \psi(2S)\phi$ result and the purple circle indicates the size of the total uncertainty.*

Appendix A

Time evolution of the B_s^0 system

This appendix gives a more detailed derivation of the differential decay rate of $B_s^0 \to \psi(2S)\phi$ and $\overline{B}_s^0 \to \psi(2S)\phi$, beginning with a description of B_s^0-$\overline{B}_s^0$ mixing.

A.1 Neutral Meson Mixing

Neutral mesons can oscillate between particle and antiparticle states via the weak interaction. To describe this oscillation the state can be descibed as an admixture of B_s^0 and $\overline{B}_s^0$ with some factors (a(t) and b(t)) which are dependent on time (Equation A.1).

$$|\Psi(t)\rangle = a(t)|B_s^0\rangle + b(t)|\overline{B}_s^0\rangle \tag{A.1}$$

The time evolution of this state can then be described by the Schrodinger equation as shown in Equation A.2, where the Hamiltonian (H) can be expressed in terms of a mass matrix (M) and a lifetime matrix (Γ), both of which are hermitian.

$$i\begin{pmatrix} d \\ dt \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = H\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \left(M - \frac{i}{2}\Gamma\right)\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \tag{A.2}$$

The matrices H, M and Γ are 2×2 matrices, where the constraint that $H_{11} = H_{22} = H_0$ comes from CPT invariance, and $H_{12} = H_{21}^*$ because the matrix is

hermitian.

$$H = \begin{pmatrix} H_0 & H_{12} \\ H_{21} & H_0 \end{pmatrix}, \; M = \begin{pmatrix} M & M_{12} \\ M_{12}^* & M \end{pmatrix}, \; \Gamma = \begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix}$$

Solving the Schrodinger equation gives the eigenvalues ($\lambda_\pm$) as shown in Equation A.3.

$$\lambda_\pm = H_0 \pm \sqrt{H_{12}H_{21}} \tag{A.3}$$

The Heavy and Light mass eigenstates (B_L and B_H) can be expressed as a linear combination of the flavour eigenstate (B_s^0 and $\overline{B}_s^0$) as shown in Equation A.4 where the factors p and q are constrained by the following relation; $|p|^2 + |q|^2 = 1$.

$$\begin{aligned} |B_L\rangle &= p|B_s^0\rangle + q|\overline{B}_s^0\rangle \\ |B_H\rangle &= p|B_s^0\rangle - q|\overline{B}_s^0\rangle \end{aligned} \tag{A.4}$$

The eigenvalues of the mass eigenstates are shown in Equation A.5 taking the convention that the 'Heavy' (H) state has the higher mass.

$$\begin{aligned} \lambda_H &= H_0 + \sqrt{H_{12}H_{21}} \\ \lambda_L &= H_0 - \sqrt{H_{12}H_{21}} \end{aligned} \tag{A.5}$$

Taking the real and imaginary parts of the eigenvalues gives the mass and decay widths respectively: $m_{H,L} = \mathcal{R}e(\lambda_{H,L})$, $\Gamma_{H,L} = 2\mathcal{I}m(\lambda_{H,L})$. Using these definitions the mass difference (Δm) of the Heavy and Light states, and the difference in the decay width ($\Delta\Gamma$) are defined as follows:

$$\begin{aligned} \Delta m &= m_H - m_L = 2\mathcal{R}e(\sqrt{H_{12}H_{21}}) \\ \Delta\Gamma &= \Gamma_L - \Gamma_H = 4\mathcal{I}m(\sqrt{H_{12}H_{21}}) \end{aligned} \tag{A.6}$$

The parameters p and q can be related to the matrix elements of the hamiltonian by the expressions described in Equation A.7 and Equation A.8.

$$\begin{pmatrix} H_0 & H_{12} \\ H_{21} & H_0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \lambda_{\pm} \begin{pmatrix} p \\ q \end{pmatrix} \tag{A.7}$$

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \tag{A.8}$$

A.2 Time evolution of states

The time evolution of the Heavy and the Light mass eigenstates is described by the equations shown in Equation A.9.

$$\begin{aligned} |B_L(t)\rangle &= e^{\lambda_L t}|B_L(0)\rangle = e^{-i(m_L - \frac{i}{2}\Gamma_L)t}|B_L(0)\rangle \\ |B_H(t)\rangle &= e^{\lambda_H t}|B_H(0)\rangle = e^{-i(m_H - \frac{i}{2}\Gamma_H)t}|B_H(0)\rangle \end{aligned} \tag{A.9}$$

The expressions for the time dependent mass eigenstates can be substituted into Equation A.4, and with some rearranging yields expressions for the time evolution of the flavour states, the B_s^0 and $\overline{B}_s^0$ (Equation A.10).

$$\begin{aligned} |B_s^0(t)\rangle &= \frac{1}{2p}\left[e^{-i(m_H - \frac{i}{2}\Gamma_H)t}|B_H(0)\rangle + e^{-i(m_L - \frac{i}{2}\Gamma_L)t}|B_L(0)\rangle\right] \\ &= \frac{1}{2p}\left[e^{-i(m_H - \frac{i}{2}\Gamma_H)t}(p|B_s^0(0)\rangle - q|\overline{B}_s^0(0)\rangle) + e^{-i(m_L - \frac{i}{2}\Gamma_L)t}(p|B_s^0(0)\rangle + q|\overline{B}_s^0(0)\rangle)\right] \\ &= \Big(\frac{e^{-i(m_L - \frac{i}{2}\Gamma_L)t} + e^{-i(m_H - \frac{i}{2}\Gamma_H)t}}{2}\Big)|B_s^0(0)\rangle + \frac{q}{p}\Big(\frac{e^{-i(m_L - \frac{i}{2}\Gamma_L)t} - e^{-i(m_H - \frac{i}{2}\Gamma_H)t}}{2}\Big)|\overline{B}_s^0(0)\rangle \\ |\overline{B}_s^0(t)\rangle &= \frac{p}{q}\Big(\frac{e^{-i(m_L - \frac{i}{2}\Gamma_L)t} - e^{-i(m_H - \frac{i}{2}\Gamma_H)t}}{2}\Big)|B_s^0(0)\rangle + \Big(\frac{e^{-i(m_L - \frac{i}{2}\Gamma_L)t} + e^{-i(m_H - \frac{i}{2}\Gamma_H)t}}{2}\Big)|\overline{B}_s^0(0)\rangle \end{aligned} \tag{A.10}$$

The following expressions (Equation A.11) can be made for the oscillation amplitudes $g_{\pm}(t)$, resulting in the simplified equations for the time evolution of the flavour states shown in Equation A.12.

$$g_{\pm}(t) = \frac{1}{2}(e^{-i(m_L - \frac{i}{2}\Gamma_L)t} \pm e^{-i(m_H - \frac{i}{2}\Gamma_H)t}) \tag{A.11}$$

$$
\begin{aligned}
|B_s^0(t)\rangle &= g_+(t)|B_s^0(0)\rangle + \frac{q}{p}g_-(t)|\overline{B}_s^0(0)\rangle \\
|\overline{B}_s^0(t)\rangle &= \frac{p}{q}g_-(t)|B_s^0(0)\rangle + g_+(t)|\overline{B}_s^0(0)\rangle
\end{aligned}
\tag{A.12}
$$

To calculate the decay rate for B_s^0 or $\overline{B}_s^0$ decaying to a final state f with a single amplitude, the decay amplitude ($A_f(t)$) must be calculated, as defined in Equation A.13. The terms A_f and $\bar{A}_f$ are also defined here.

$$
\begin{aligned}
A_f(t) &= \langle f|\mathcal{H}|B_s^0(t)\rangle,\ A_f = \langle f|\mathcal{H}|B_s^0(0)\rangle \\
\bar{A}_f(t) &= \langle f|\mathcal{H}|\bar{B}_s^0(t)\rangle,\ \bar{A}_f = \langle f|\mathcal{H}|\bar{B}_s^0(0)\rangle
\end{aligned}
\tag{A.13}
$$

$$
\begin{aligned}
\frac{d\Gamma_{B_s^0\to f}(t)}{dt} &= |A_f(t)|^2 = |\langle f|B_s^0(t)\rangle|^2 = \langle f|B_s^0(t)\rangle\langle f|B_s^0(t)\rangle^\dagger \\
&= |g_+|^2|A_f|^2 + g_+(t)g_-^*(t)\Big(\frac{q}{p}\Big)^* A_f\bar{A}_f^* + \Big(\frac{q}{p}\Big)g_-(t)g_+^*(t)\bar{A}_f A_f^* + \Big|\frac{q}{p}\Big|^2|g_-(t)|^2|\bar{A}_f|^2 \\
&= |A_f|^2(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + 2\mathcal{R}e[\lambda_f g_+^*(t)g_-(t)])
\end{aligned}
\tag{A.14}
$$

$$
\frac{d\Gamma_{\bar{B}_s^0\to f}(t)}{dt} = \Big|\frac{p}{q}\Big|^2|g_+|^2|A_f|^2 + g_+(t)g_-^*(t)\Big(\frac{p}{q}\Big)A_f\bar{A}_f^* + \Big(\frac{p}{q}\Big)^* g_-(t)g_+^*(t)\bar{A}_f A_f^* + |g_-(t)|^2|\bar{A}_f|^2
\tag{A.15}
$$

By making the substitution that $\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}$, and using the definitions in Equation A.16 the decay rate can be expressed as shown in Equation A.17.

$$
\begin{aligned}
|g_+(t)|^2 &= \frac{e^{-\Gamma_s t}}{2}\Big(\cos(\Delta m_s\, t) + \cosh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big)\Big) \\
|g_-(t)|^2 &= \frac{e^{-\Gamma_s t}}{2}\Big(-\cos(\Delta m_s\, t) + \cosh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big)\Big) \\
g_+^*(t)g_-(t) &= \frac{e^{-\Gamma_s t}}{2}\Big(i\sin(\Delta m_s\, t) - \sinh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big)\Big) \\
g_+(t)g_-^*(t) &= \frac{e^{-\Gamma_s t}}{2}\Big(-i\sin(\Delta m_s\, t) - \sinh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big)\Big)
\end{aligned}
\tag{A.16}
$$

$$\begin{aligned}\frac{d\Gamma_{B_s^0\to f}(t)}{dt} =&|A_f|^2(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + 2\mathcal{R}e[\lambda_f g_+^*(t)g_-(t)]) \\ =&\frac{e^{-\bar{\Gamma}_s}}{2}|A_f|^2\Big((\cos(\Delta m_s\, t) + \cosh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big) \\ &+ |\lambda_f|^2(-\cos(\Delta m_s\, t) + \cosh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big) \\ &+ 2\mathcal{R}e\Big(\lambda_f\Big(-i\sin(\Delta m_s\, t) - \sinh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big)\Big)\Big)\Big)\end{aligned} \tag{A.17}$$

By collecting the common terms, the decay rate is found to the following:

$$\begin{aligned}\frac{d\Gamma_{B_s^0\to f}(t)}{dt} =&\frac{e^{-\bar{\Gamma}_s t}}{2}|A_f|^2\Big(\Big((1 + |\lambda_f|^2)\cosh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big)\Big) \\ &+ (1 - |\lambda_f|^2)\cos(\Delta m_s\, t) \\ &- 2\mathcal{R}e(\lambda_f)\sinh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big) \\ &- 2\mathcal{I}m(\lambda_f)\sin(\Delta m_s\, t)\Big)\end{aligned} \tag{A.18}$$

Repeating the procedure for the decay of the $\overline{B}_s^0$ meson to the final state f gives the decay rate given in Equation A.19.

$$\begin{aligned}\frac{d\Gamma_{\bar{B}_s^0\to f}(t)}{dt} =&\Big|\frac{p}{q}\Big|^2|A_f|^2\frac{e^{-\bar{\Gamma}_s t}}{2}\Big(\Big((1 + |\lambda_f|^2)\cosh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big)\Big) \\ &- (1 - |\lambda_f|^2)\cos(\Delta m_s\, t) \\ &- 2\mathcal{R}e(\lambda_f)\sinh\Big(\frac{\Delta\Gamma_s}{2}\, t\Big) \\ &+ 2\mathcal{I}m(\lambda_f)\sin(\Delta m_s\, t)\Big)\end{aligned} \tag{A.19}$$

For both derived decay rates the parameter λ_f encapsulates all the *CP* information of the decay mode. Taking $CP(\lambda_f) = \eta_f\lambda_f$ where η_f, the *CP* eigenvalue is ± 1.

Appendix B

BDT Details

This section presents further details of the BDT training process. Plots showing the correlations between the BDT variables for the signal and background samples are shown in Figure B.1 top and bottom respectively. Figure B.2 shows the distribution of the signal and background sample for the variables used to train the BDT. The signal is shown as the blue histogram and the background as the red. To train the BDT simulated data was used to model the signal. A condition of using simulated data to train the BDT is that the variables used must be well represented compared to the signal distributions from data. Plots showing the comparison between simulated data and real data for the variables used in the BDT are shown in Figure B.3.

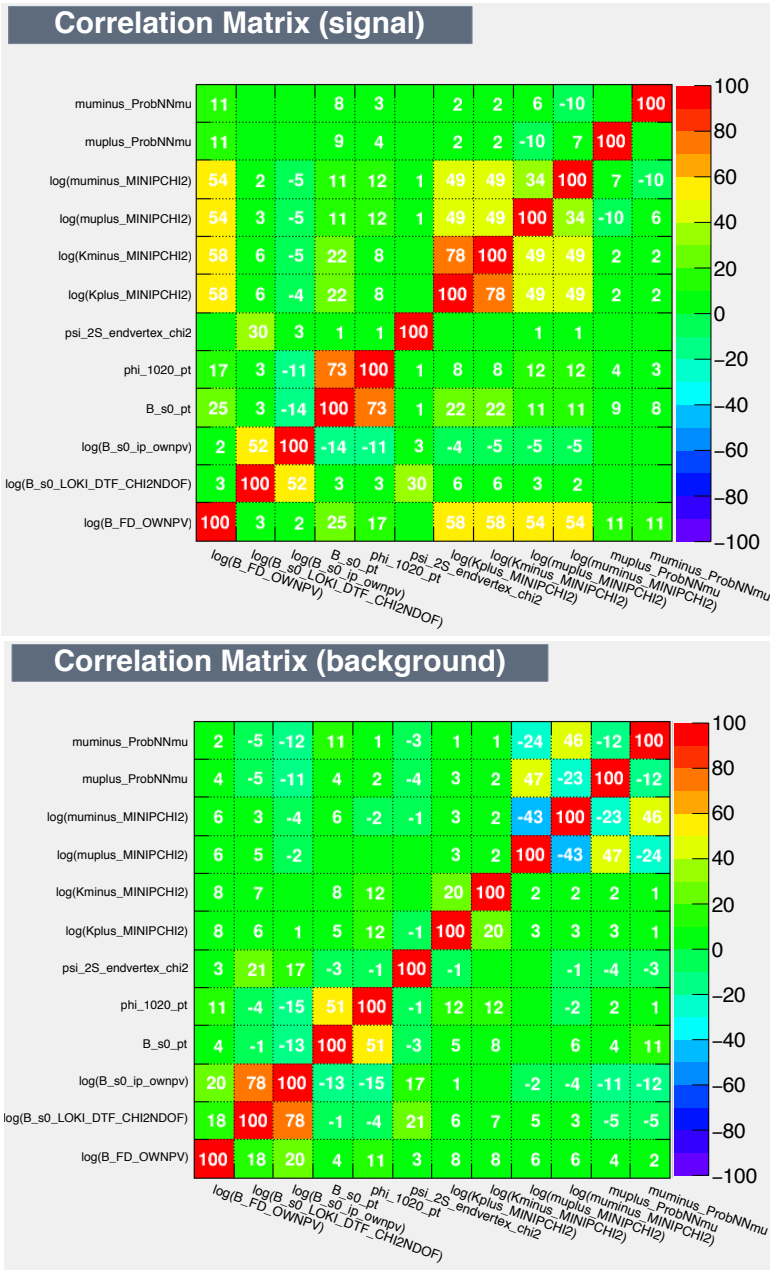


Figure B.1 *Correlation matrices for the signal (top) and background (bottom) samples.*

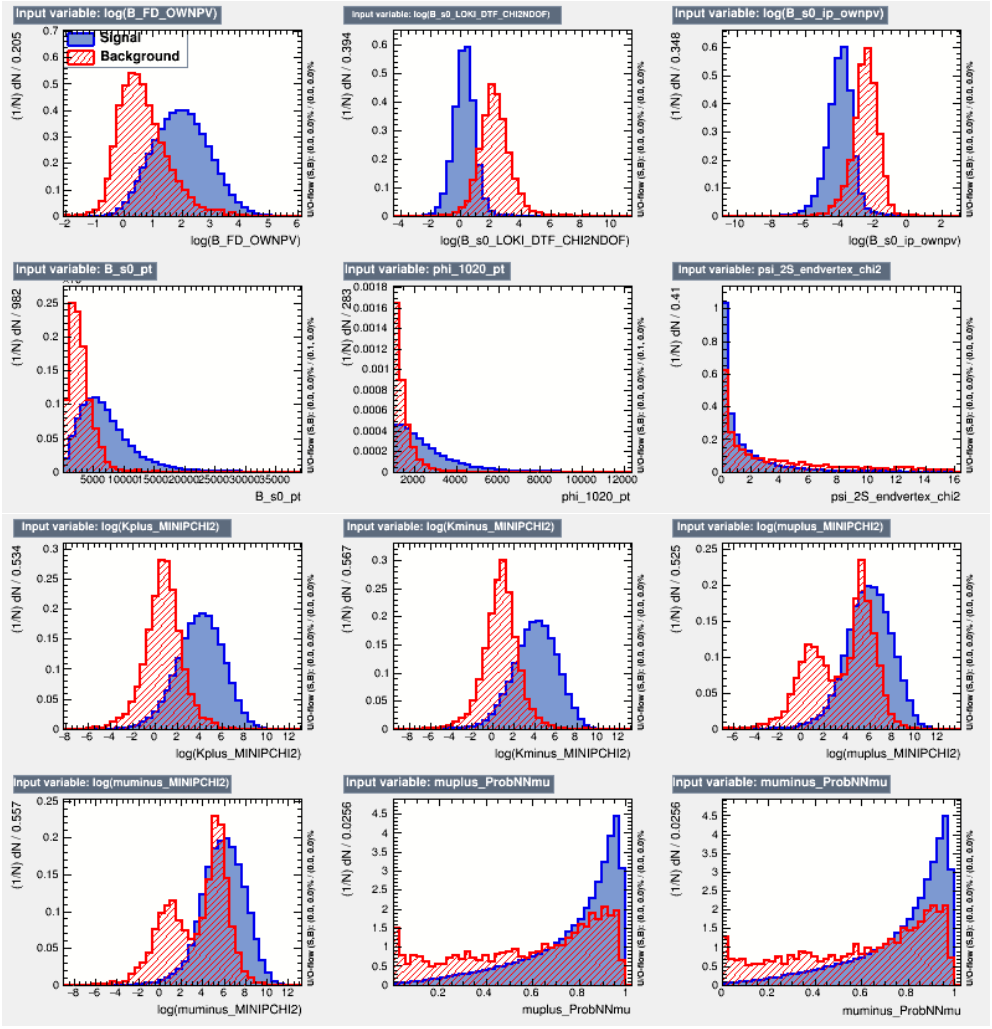


Figure B.2 *Input variable distributions for signal (blue) and background (red) for the BDT.*

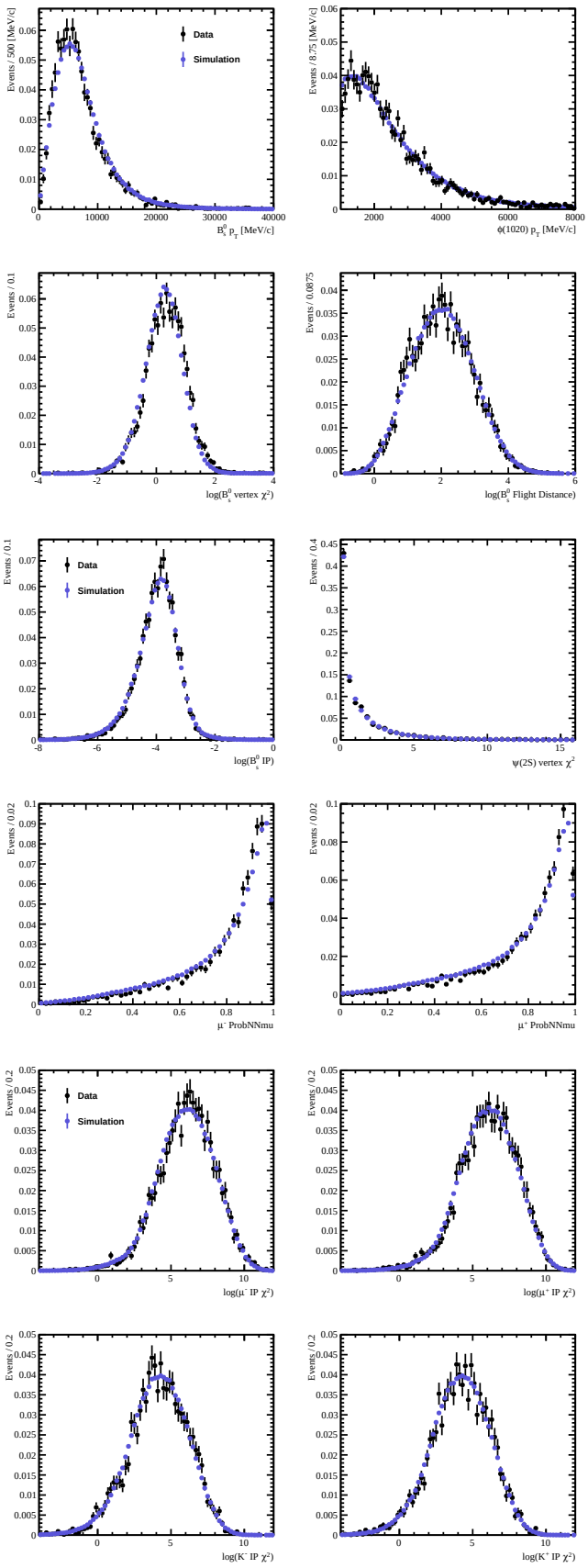


Figure B.3 *Comparison of between simulated events and background subtracted-data of the distribution of variables used for the BDT training. The real data are shown as the black points, and the simulated events as the blue.*

Appendix C

Decay Time Acceptance

Table C.1 *The values for the decay time acceptance histogram used in the log-likelihood fits, shown in Figure 5.16.*

Bin	Lower Edge	Value	Bin	Lower Edge	Value
1	0.3	0.679	21	1.340	1.072
2	0.338	0.710	22	1.417	1.071
3	0.377	0.788	23	1.498	1.030
4	0.420	0.798	24	1.583	1.089
5	0.458	0.923	25	1.674	1.072
6	0.500	0.883	26	1.771	1.046
7	0.544	0.919	27	1.874	1.025
8	0.589	0.902	28	1.986	1.042
9	0.635	1.034	29	2.106	1.046
10	0.682	1.016	30	2.236	1.012
11	0.731	0.926	31	2.379	1.123
12	0.782	0.946	32	2.537	1.087
13	0.835	0.952	33	2.714	1.175
14	0.889	1.010	34	2.914	1.050
15	0.946	0.959	35	3.145	1.053
16	1.005	0.972	36	3.418	1.078
17	1.066	1.008	37	3.752	1.015
18	1.130	1.059	38	4.183	1.028
19	1.197	1.074	39	4.791	1.086
20	1.266	1.147	40	5.827	1.095

List of Figures

List of Tables

Bibliography

[1] Andrei Dmitrievich Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Physics-Uspekhi*, 34(5):392–393, 1991.

[2] 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.*, 13:138–140, Jul 1964.

[3] G.C. Branco, L. Lavoura, and J.P. Silva. *CP Violation*. International series of monographs on physics. Clarendon Press, 1999.

[4] M. Kobayashi and T. Maskawa. *CP* violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652, 1973.

[5] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.

[6] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531, 1963.

[7] C. Jarlskog. Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation. *Phys. Rev. Lett.*, 55:1039–1042, Sep 1985.

[8] J. Charles et al. Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics. *Phys. Rev.*, D91(7):073007, 2015. with updated results and plots available at `http://ckmfitter.in2p3.fr`.

[9] T. Aaltonen et al. Measurement of lifetime and decay-width difference in $B_s^0 \to J/\psi\phi$ decays. *Phys. Rev. Lett.*, 100:121803, 2008.

[10] V. M. Abazov et al. Lifetime difference and CP-violating phase in the B_s^0 system. *Phys. Rev. Lett.*, 98:121801, 2007.

[11] Georges Aad et al. Time-dependent angular analysis of the decay $B_s^0 \to J/\psi\phi$ and extraction of $\Delta\Gamma_s$ and the CP-violating weak phase ϕ_s by ATLAS. *JHEP*, 12:072, 2012.

[12] Vardan Khachatryan et al. Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_\mathrm{s}$ using the $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV. 2015.

[13] R. Aaij et. al. Precision measurement of CP violation in $B_s^0 \to J/\psi K^+K^-$ decays. *Phys. Rev. Lett.*, 114(arXiv:1411.3104. LHCB-PAPER-2014-059. CERN-PH-EP-2014-271):041801. 19 p, Nov 2014.

[14] Roel Aaij et al. Measurement of the CP-violating phase ϕ_s in $\overline{B}_s^0 \to J/\psi\pi^+\pi^-$ decays. *Phys. Lett.*, B736:186–195, 2014.

[15] Roel Aaij et al. Measurement of the CP-violating phase ϕ_s in $\bar{B}_s^0 \to D_s^+D_s^-$ decays. *Phys. Rev. Lett.*, 113(21):211801, 2014.

[16] Y. Amhis et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014. `http://www.slac.stanford.edu/xorg/hfag/osc/summer_2015/#BETAS`.

[17] J. Charles et al. Predictions of selected flavour observables within the Standard Model. *Phys. Rev.*, D84:033005, 2011.

[18] Alexander Lenz and Ulrich Nierste. Numerical Updates of Lifetimes and Mixing Parameters of B Mesons. In *CKM unitarity triangle. Proceedings, 6th International Workshop, CKM 2010, Warwick, UK, September 6-10, 2010*, 2011.

[19] Christiane Lefvre. The CERN accelerator complex. Complexe des accelerateurs du CERN. Dec 2008.

[20] The LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, Aug 2008.

[21] R Aaij, A Affolder, K Akiba, M Alexander, S Ali, RB Appleby, M Artuso, A Bates, A Bay, O Behrendt, et al. Performance of the LHCb vertex locator. *Journal of Instrumentation*, 9(09):P09007, 2014.

[22] The LHCb RICH Collaboration. Performance of the LHCb RICH detector at the LHC, journal=.

[23] M.S. Longair. *High Energy Astrophysics: Volume 1, Particles, Photons and Their Detection*. High Energy Astrophysics. Cambridge University Press, 1992.

[24] The LHCb Collaboration. *LHCb RICH: Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2000.

[25] *LHCb muon system: second addendum to the Technical Design Report*. Technical Design Report LHCb. CERN, Geneva, 2005. Submitted on 9 Apr 2005.

[26] O Callot, M Ferro-Luzzi, and P Perret. Using the SPD multiplicity in the Level-0 trigger. Technical Report LHCb-2003-022, CERN, Geneva, Apr 2003.

[27] Roel Aaij and Johannes Albrecht. Muon triggers in the High Level Trigger of LHCb. Technical Report LHCb-PUB-2011-017. CERN-LHCb-PUB-2011-017, CERN, Geneva, Sep 2011.

[28] Albert Puig. The LHCb trigger in 2011 and 2012. Technical Report LHCb-PUB-2014-046. CERN-LHCb-PUB-2014-046, CERN, Geneva, Nov 2014.

[29] A Hoecker, P Speckmayer, J Stelzer, J Therhaag, E von Toerne, and H Voss. Tmva 4; toolkit for multivariate data analysis with root. *arXiv preprint physics/0703039*, 2007.

[30] Y Xie. Principles to optimize event selections for measurements of CP asymmetries. Technical Report LHCb-INT-2009-012. CERN-LHCb-INT-2009-012, CERN, Geneva, Jul 2009.

[31] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014.

[32] Muriel Pivk and Francois R. Le Diberder. SPlot: A Statistical tool to unfold data distributions. *Nucl.Instrum.Meth.*, A555:356–369, 2005.

[33] LHCb Collaboration et al. LHCb detector performance. *International Journal of Modern Physics A*, 2015.

[34] H. G. Moser and A. Roussarie. Mathematical methods for B0 anti-B0 oscillation analyses. *Nucl. Instrum. Meth.*, A384:491–505, 1997.

[35] R. Aaij et. al. Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ decays. *Phys. Rev. D*, 87(arXiv:1304.2600. CERN-PH-EP-2013-055. LHCB-PAPER-2013-002):112010. 39 p, Apr 2013. Comments: 33 pages, 16 figures.

[36] Tristan du Pree, M H M Merk, and H G Raven. *Search for a Strange Phase in Beautiful Oscillations*. PhD thesis, Vrije U., Amsterdam, Amsterdam, 2010. Presented on 22 Oct 2010.

[37] R. Aaij et al. Opposite-side flavour tagging of B mesons at the LHCb experiment. *Eur. Phys. J.*, C72:2022, 2012.

[38] Roel Aaij et al. B flavour tagging using charm decays at the LHCb experiment. *JINST*, 10(10):P10005, 2015.

[39] Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data, Nov 2012.

[40] Dorigo, M. and others. FT workshop, intro and status. `https://indico.cern.ch/event/257864/session/0/contribution/37/material/slides/0.pdf`.

[41] Dorigo, M. and others. FT twiki page. `https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/FlavourTagging4Reco14`.

[42] Robert A Currie. *CP-Violation in Beautiful-Strange Oscillations at LHCb*. PhD thesis, Edinburgh U., May 2014. Presented 24 Oct 2013.

[43] Fred James, Matthias Winkler, et al. Minuit users guide. 2004. `http://seal.web.cern.ch/seal/documents/minuit/mnusersguide.pdf`.

[44] F James and MINUIT Roos. Minuit-a system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications*, 10(6):343–367, 1975.

[45] Roger Fletcher. A new approach to variable metric algorithms. *The computer journal*, 13(3):317–322, 1970.

[46] Fred James. The interpretation of errors. 2004. `http://lmu.web.psi.ch/docu/manuals/software_manuals/minuit2/mnerror.pdf`.

[47] Frederick James. Statistical methods in experimental physics, 2006. `http://cds.cern.ch/record/1019859/files/9789812567956_TOC.pdf`.

[48] K. A. Olive et al. Review of particle physics. *Chin. Phys.*, C38:090001, 2014. and 2015 update.

[49] R. Aaij et al. Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^+\pi^-$, April 2013.

[50] R. Aaij et al. Observation of the decay $B_c^+ \to B_s^0\pi^+$. *Phys. Rev. Lett.*, 111:181801, 2013.