First study of the $CP$-violating phase and decay-width difference in $B_s^0 \rightarrow \psi(2S)\phi$ decays

The LHCb collaboration

Abstract

A time-dependent angular analysis of $B_s^0 \rightarrow \psi(2S)\phi$ decays is performed using data recorded by the LHCb experiment. The data set corresponds to an integrated luminosity of 3.0 fb$^{-1}$ collected during Run 1 of the LHC. The $CP$-violating phase and decay-width difference of the $B_s^0$ system are measured to be $\phi_s = 0.23^{+0.29}_{-0.28}$ ± 0.02 rad and $\Delta \Gamma_s = 0.066^{+0.041}_{-0.044} \pm 0.007$ ps$^{-1}$, respectively, where the first uncertainty is statistical and the second systematic. This is the first time that $\phi_s$ and $\Delta \Gamma_s$ have been measured in a decay containing the $\psi(2S)$ resonance.

Published in Phys. Lett. B

© CERN on behalf of the LHCb collaboration, licence [CC-BY-4.0](https://creativecommons.org/licenses/by/4.0/)

$^\dagger$Authors are listed at the end of this paper.
1 Introduction

The interference between the amplitudes of decays of $B_s^0$ mesons to $c\bar{c}X CP$ eigenstates directly or via mixing, gives rise to a $CP$-violating phase, $\phi_s$. In the Standard Model (SM), ignoring subleading penguin contributions, this phase is predicted to be $-2\beta_s$, where $\beta_s = \arg[-(V_{ts}V_{tb}^*)/(V_{cs}V_{cb}^*)]$ and $V_{ij}$ are elements of the CKM quark flavour mixing matrix $[1]$. Measurements of $\phi_s$ using $B_s^0 \to J/\psi K^+ K^-$ and $B_s^0 \to J/\psi \pi^+ \pi^-$ decays have been reported previously by the LHCb collaboration $[2]$ based upon $3.0 fb^{-1}$ of integrated luminosity collected in $pp$ collisions at a centre-of-mass energy of 7 TeV in 2011 and 8 TeV in 2012 at the LHC. Measurements of $\phi_s$ using $B_s^0 \to J/\psi \phi$ decays have also been made by the D0 $[3]$, CDF $[4]$, CMS $[5]$ and ATLAS $[6]$ collaborations. The world-average value of these direct measurements is $\phi_s = -0.033 \pm 0.033 \text{ rad}$ $[7]$. The global average from indirect measurements gives $\phi_s = -0.0376^{+0.0007}_{-0.0008} \text{ rad}$ $[8]$. Measurements of $\phi_s$ are interesting since new physics (NP) processes could modify the phase if new particles were to contribute to the box diagrams describing $B_s^0$–$\bar{B}_s^0$ mixing $[9,10]$.

In this analysis $\phi_s$ is measured using a flavour tagged, decay-time dependent angular analysis of $B_s^0 \to \psi(2S) \phi$ decays, with $\psi(2S) \to \mu^+ \mu^-$ and $\phi \to K^+ K^-$. In addition, measurements of the decay-width difference of the light (L) and heavy (H) $B_s^0$ mass eigenstates, $\Delta \Gamma_s \equiv \Gamma_L - \Gamma_H$, the average $B_s^0$ decay width, $\Gamma_s \equiv (\Gamma_L + \Gamma_H)/2$, and the polarisation amplitudes of the $B_s^0 \to \psi(2S) \phi$ decay are reported. This is the first time that a higher $c\bar{c}$ resonance is used to measure $\phi_s$.

This analysis follows very closely that of $B_s^0 \to J/\psi K^+ K^-$ decays in Refs. $[2][11]$, and only significant changes with respect to those analyses are described in this paper. Section 2 describes the phenomenology of the $B_s^0 \to \psi(2S) \phi$ decay and the physics observables. Section 3 describes the LHCb detector, data and simulated samples that are used along with the optimisation of their selection. Section 4 details the $B_s^0$ meson decay-time resolution, decay-time efficiency and angular acceptance and Section 5 describes the flavour tagging algorithms. Results and systematic uncertainties are given in Section 6 and Section 7 respectively. Conclusions are presented in Section 8.

2 Phenomenology

The full formalism used for this analysis can be found in Ref. $[11]$, where the $J/\psi$ is now replaced with the $\psi(2S)$ meson. The differential cross-section as a function of the signal decay time, $t$, and three helicity angles, $\Omega = (\cos \theta_\mu, \cos \theta_K, \varphi)$ (Fig. 1), is described by a sum of ten terms, corresponding to the four polarisation amplitudes (three corresponding to the $K^+ K^-$ from the $\phi$ being in a $P$-wave configuration, and one to allow for an additional non-resonant $K^+ K^-$ $S$-wave component) and their interference terms. Each term is the product of a time-dependent function and an angular function,

$$X(t, \Omega) \equiv \frac{d^4 \Gamma(B_s^0 \to \psi(2S) \phi)}{dt \ d\Omega} \propto \sum_{k=1}^{10} h_k(t) f_k(\Omega), \quad (1)$$

where the definitions of $h_k(t)$ and $f_k(\Omega)$ are given in Ref. $[11]$. The $f_k(\Omega)$ functions depend only upon the final-state decay angles. The $h_k(t)$ functions depend upon all physics parameters of interest, which are $\Gamma_s$, $\Delta \Gamma_s$, $\phi_s$, $|\lambda|$, the mass difference of the
$B_s^0$ eigenstates, $\Delta m_s$, and the polarisation amplitudes $A_i = |A_i|e^{-i\delta_i}$, where the indices $i \in \{0, \|, \perp, S\}$ refer to the different polarisation states of the $K^+K^-$ system. The sum $|A_0|^2 + |A_\||^2 + |A_\perp|^2 + |A_S|^2$ equals unity and by convention $\delta_0$ is zero. The $S$-wave fraction is defined as $F_S \equiv |A_S|^2/(|A_0|^2 + |A_\||^2 + |A_\perp|^2 + |A_S|^2)$. The parameter $\lambda$ describes CP violation in the interference between mixing and decay and is defined by $\lambda = \eta_i (q/p)(\bar{A}_i/A_i)$. The complex parameters $p = \langle B_s^0|B_{s,L} \rangle$ and $q = \langle \bar{B}_s^0|B_{s,L} \rangle$ describe the relation between flavour and mass eigenstates, where $B_{s,L}$ is the light mass eigenstate and $\eta_i$ is the CP eigenvalue of the polarisation state $i$. The CP-violating phase is defined by $\phi_s \equiv -\arg (\eta_i \lambda)$ and is assumed here to be the same for all polarisation states. In the absence of CP violation in decay it follows that $|\lambda| = 1$. In this paper CP violation in $B_s^0$-meson mixing is assumed to be negligible, following measurements in Refs. [12,13].

3 Detector, data set and selection

The LHCb detector [14, 15] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of momentum, $p$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/c. The minimum distance of a track to a primary vertex (PV), the impact parameter, is measured with a resolution of $(15 + 29/p_T) \mu$m, where $p_T$ is the component of the momentum transverse to the beam, in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

The online event selection is performed by a trigger [16], which consists of a hardware
stage, based on information from the calorimeter and the muon system, followed by a software stage. In this analysis, candidates are required to pass the hardware trigger that selects muons and muon pairs based on their transverse momentum. In the software stage, events are triggered by a $\psi(2S) \rightarrow \mu^+\mu^-$ candidate, where the $\psi(2S)$ is required to be consistent with coming from the decay of a $b$ hadron, by using either impact parameter requirements on the decay products or the detachment of the $\psi(2S)$ candidate from the PV.

In the simulation, $pp$ collisions are generated using Pythia 17 with a specific LHCb configuration 18. Decays of hadronic particles are described by EvtGen 19, in which final-state radiation is generated using Photos 20. The interaction of the generated particles with the detector, and its response, are implemented using the Geant4 toolkit 21 as described in Ref. 22.

The $B_s^0 \rightarrow \psi(2S)\phi$ candidates are first selected with loose requirements to ensure high efficiency and significant background rejection. The $\psi(2S)$ candidates are reconstructed from pairs of oppositely-charged particles identified as muons, and the $\phi$ candidates are reconstructed from pairs of oppositely-charged particles identified as kaons. The invariant mass of the muon (kaon) pair must be within 60 MeV/$c^2$ (12 MeV/$c^2$) of the known $\psi(2S)$ ($\phi$) mass 23. Reconstructed kaon tracks that do not correspond to actual trajectories of charged particles are suppressed by requiring a good track $\chi^2$ per degree of freedom. The $p_T$ of each $\phi$ candidate is required to be larger than 1 GeV/$c$.

The $\psi(2S)$ and $\phi$ candidates that are consistent with originating from a common vertex are combined to create $B_s^0$ candidates. Subsequently, a kinematic fit 24 is applied to the $B_s^0$ candidates in which the $\psi(2S)$ mass is constrained to the known value 23 and the $B_s^0$ candidate is required to point back to the PV, to improve the resolution on the invariant mass $m(\psi(2S)K^+K^-)$. Combinatorial background from particles produced at the PV is reduced by requiring that the $B_s^0$ candidate decay time (computed from a vertex fit without the PV constraint) is larger than 0.3 ps. Backgrounds from the misidentification of final-state particles from other decays such as $B^0 \rightarrow \psi(2S)K^+\pi^-$ and $A_s^0 \rightarrow \psi(2S)pK^-$ are negligible.

To further improve the signal-to-background ratio, a boosted decision tree (BDT) 25, 26 is applied. The BDT is trained using simulated $B_s^0 \rightarrow \psi(2S)\phi$ events for the signal, while candidates from data with $m(\psi(2S)K^+K^-)$ larger than 5400 MeV/$c^2$ are used to model the background. Twelve variables that have good discrimination power between signal and background are used to define and train the BDT. These are: the $B_s^0$ candidate kinematic fit $\chi^2$; the $p_T$ of the $B_s^0$ and $\phi$ candidates; the $B_s^0$ candidate flight distance and impact parameter with respect to the PV; the $\psi(2S)$ candidate vertex $\chi^2$; the $\chi^2_{DP}$ of the kaon and muon candidates (defined as the change in $\chi^2$ of the PV fit when reconstructed with and without the considered particle) and the muon identification probabilities. The optimal working point for the BDT is determined using a figure of merit that optimises the statistical power of the selected data sample for the analysis of $\phi_s$ by taking account of the number of signal and background candidates, as well as the decay-time resolution and flavour-tagging power of each candidate.

Figure 2 shows the distribution of $m(\psi(2S)K^+K^-)$ for the selected $B_s^0 \rightarrow \psi(2S)\phi$ candidates. An extended maximum likelihood fit is made to the unbinned $m(\psi(2S)K^+K^-)$ distribution, where the signal component is described by the sum of two Crystal Ball functions and the small combinatorial background by an exponential function. All parameters are left free in the fit, including the yields of the signal and background
components. This fit gives a yield of $4695 \pm 71$ signal candidates and $174 \pm 10$ background candidates in the range $m(\psi(2S)K^+K^-) \in [5310, 5430] \text{MeV}/c^2$. It is used to assign per-candidate weights (sWeights) via the sPlot technique \[28\], which are used to subtract the background contribution in the maximum likelihood fit described in Section 6.

4 Detector resolution and efficiency

The resolution on the measured decay time is determined with the same method as described in Refs. \[2,11\] by using a large sample of prompt $J/\psi K^+K^-$ combinations produced directly in the $pp$ interactions. These events are selected using prompt $J/\psi \rightarrow \mu^+\mu^-$ decays via a prescaled trigger that does not impose any requirements on the separation of the $J/\psi$ from the PV. The $J/\psi$ candidates are combined with oppositely charged tracks that are identified as kaons, using a similar selection as for the signal decay. The resolution model, $R(t - t')$, is the sum of two Gaussian distributions with per-event widths. These widths are calibrated by using a maximum likelihood fit to the unbinned decay time and decay-time uncertainty distributions of the prompt $J/\psi K^+K^-$ combinations, using a model composed of the sum of a $\delta$ function for the prompt component and two exponential functions for long-lived backgrounds, all of which are convolved with the resolution function. A third Gaussian distribution is added to the total fit function to account for the small ($< 1\%$) fraction of decays that are associated to the wrong PV. The average effective resolution is $46.6 \pm 1.0 \text{fs}$. Simulated $B^0_s \rightarrow J/\psi K^+K^-$ and $B^0 \rightarrow \psi(2S)K^+K^-$ events show no significant difference in the effective decay-time resolution between the two decay modes.

The reconstruction efficiency is not constant as a function of decay time due to displacement requirements made on signal tracks in the trigger and event selection. The efficiency is determined using the control channel $B^0 \rightarrow \psi(2S)K^*(892)^0$, with $K^*(892)^0 \rightarrow K^+\pi^-$, which is assumed to have a purely exponential decay-time distribution.
It is defined as

$$\varepsilon_{B^0}(t) = \varepsilon_{B^0}^{\text{data}}(t) \times \frac{\varepsilon_{B^0}(t)}{\varepsilon_{\text{sim}}(t)},$$

(2)

where $\varepsilon_{B^0}^{\text{data}}(t)$ is the efficiency of the control channel and $\varepsilon_{\text{sim}}(t)/\varepsilon_{B^0}(t)$ is the ratio of efficiencies of the simulated signal and control modes after the full trigger and selection chain has been applied. This correction accounts for the small differences in the lifetime and kinematics between the signal and control modes.

The $B^0 \to \psi(2S)K^+(892)^0$ decay is selected using a similar trigger, preselection and the same BDT training and working point as used for the signal (with appropriate changes for kaon to pion). Backgrounds from the misidentification of final-state particles from other decays such as $B_s^0 \to \psi(2S)\phi$ and $A^0_s \to \psi(2S)pK^-$ are negligible. Similarly, possible backgrounds from $B^0_s \to \psi(2S)\pi^+\pi^-$ decays where a pion is misidentified as a kaon, and $B^+ \to \psi(2S)K^+$ decays combined with an additional random pion, are negligible.

The $\psi(2S)K^+\pi^-$ invariant mass distribution is shown in Fig. 3 along with the result of a fit composed of the sum of two Crystal Ball (CB) functions for the signal and an exponential function for the background. The tail parameters and relative fraction of the two CB functions are fixed to values obtained from a fit to simulated $B^0 \to \psi(2S)K^+(892)^0$ decays. The core widths and common mean of the CB functions are free in the fit and the $B^0$ yield is found to be $28676 \pm 195$. The efficiency is defined as $\varepsilon_{B^0}^{\text{data}}(t) = N_{B^0}^{\text{data}}(t)/N_{\text{gen}}^{B^0}(t)$ where $N_{\text{data}}^{B^0}(t)$ is the number of signal $B^0 \to \psi(2S)K^+(892)^0$ decays in a given bin of decay time and $N_{\text{gen}}^{B^0}(t)$ is the number of events generated from an exponential distribution with lifetime $\tau_{B^0} = 1.520 \pm 0.004$ ps [23]. The exponential distribution is convolved with a double Gaussian resolution model, the parameters of which are determined from a fit to the decay time distribution of prompt $J/\psi K^+\pi^-$ combinations. In total $10^7$ events are generated. The sPlot [28] technique with $m(\psi(2S)K^+\pi^-)$ as discriminating variable is used to determine $N_{\text{data}}^{B^0}(t)$. The analysis is not sensitive to the absolute scale of the efficiency. The final decay-time efficiency for the $B_s^0 \to \psi(2S)\phi$ signal is shown in Fig. 4. It is relatively uniform at high values of decay time but decreases at low decay times due to selection requirements placed on the track $\chi^2_{IP}$ variables.
Figure 4: Decay-time efficiency $\varepsilon_{\text{data}}^{B_0^s}(t)$ in arbitrary units.

The efficiency as a function of the $B_0^s \to \psi(2S)\phi$ helicity angles is not uniform due to the forward geometry of the LHCb detector and the requirements imposed on the final-state particle momenta. The three-dimensional efficiency, $\varepsilon(\Omega)$, is determined with the same technique as used in Ref. [11] using simulated events that are subjected to the same trigger and selection criteria as the data. The relative efficiencies vary by up to 20%, dominated by the dependence on $\cos \theta_{\mu}$.

5 Flavour tagging

The $B_s^0$ candidate flavour at production is determined by two independent classes of flavour tagging algorithms, the opposite-side (OS) taggers [29] and the same-side kaon (SSK) tagger [30], which exploit specific features of the production of $b\bar{b}$ quark pairs in $pp$ collisions, and their subsequent hadronisation. Each tagging algorithm gives a tag decision and a mistag probability. The tag decision, $q$, takes values $+1$, $-1$, or 0, if the signal meson is tagged as $B_0^s$, $\bar{B}_0^s$, or is untagged, respectively. The fraction of events in the sample with a nonzero tagging decision gives the efficiency of the tagger, $\varepsilon_{\text{tag}}$. The mistag probability, $\eta$, is estimated event-by-event, and represents the probability that the algorithm assigns a wrong tag decision to the event; it is calibrated using data samples of several flavour-specific $B^0$, $B^+$ and $B_{s2}^0$ decays to obtain the corrected mistag probability, $\bar{\omega}$, for an initial flavour $\bar{B}_0^s$ meson. A linear relationship between $\eta$ and $\bar{\omega}$ is used for the calibration. The effective tagging power is given by $\varepsilon_{\text{tag}}(1 - 2\omega)^2$ and for the combined taggers in the $B_s^0 \to \psi(2S)\phi$ signal sample is $(3.88 \pm 0.13 \pm 0.12)\%$, where the first uncertainty is statistical and the second systematic.

6 Maximum likelihood fit

The physics parameters are determined by a weighted maximum likelihood fit of a signal-only probability density function (PDF) to the four-dimensional distribution of $B_s^0 \to \psi(2S)\phi$ decay time and helicity angles. The negative log-likelihood function to be
minimised is given by
\[ -\ln \mathcal{L} = -\alpha \sum_{\text{events } i} W_i \ln \mathcal{P}, \]
where \( W_i \) are the sWeights computed using \( m(\psi(2S)K^+K^-) \) as the discriminating variable and the factor \( \alpha = \sum W_i / \sum W_i^2 \) is necessary to obtain the correct parameter uncertainties from the Hessian of the negative log-likelihood. The PDF, \( \mathcal{P} = \mathcal{S} / \int \mathcal{S} \, dt \, d\Omega \), is obtained from
\[ \mathcal{S}(t, \Omega, q^{\text{OS}}, q^{\text{SSK}}|\eta^{\text{OS}}, \eta^{\text{SSK}}) = \mathcal{X}(t', \Omega, q^{\text{OS}}, q^{\text{SSK}}|\eta^{\text{OS}}, \eta^{\text{SSK}}) \otimes R(t - t') \times \varepsilon_{\text{data}}^B(t), \]
where
\[ \mathcal{X}(t, \Omega, q^{\text{OS}}, q^{\text{SSK}}|\eta^{\text{OS}}, \eta^{\text{SSK}}) = (1 + q^{\text{OS}}(1 - 2\omega^{\text{OS}})) \{ 1 + q^{\text{SSK}}(1 - 2\omega^{\text{SSK}}) \} X(t, \Omega) + \}
\[ \{ 1 - q^{\text{OS}}(1 - 2\omega^{\text{OS}}) \{ 1 - q^{\text{SSK}}(1 - 2\omega^{\text{SSK}}) \} \} \overline{X}(t, \Omega), \]
which allows for the inclusion of information from both tagging algorithms in the computation of the decay rate. The function \( X(t, \Omega) \) is defined in Eq. 1 and \( \overline{X}(t, \Omega) \) is the corresponding function for \( B_s^0 \) decays. As in Ref. [11], the angular efficiency is included in the normalisation of the PDF via ten integrals, \( I_k = \int \varepsilon(\Omega) f_k(\Omega) \), which are calculated using simulated events. In contrast to Refs. [2, 11], the fit is performed in a single bin of \( m(K^+K^-) \), within 12 MeV/c^2 of the known \( \phi \) mass.

In the fit, Gaussian constraints are applied to the \( B_s^0 \) mixing frequency \( \Delta m_s = 17.757 \pm 0.021 \) ps\(^{-1}\) [7] and the tagging calibration parameters. The fitting procedure has been validated using pseudoexperiments and simulated \( B_s^0 \to \psi(2S)\phi \) decays. Due to the symmetry in the PDF there is a two-fold ambiguity in the solutions for \( \phi_s \) and \( \Delta\Gamma_s \); the solution with positive \( \Delta\Gamma_s \) is used [31]. The results of the fit to the data are shown in Tables 1 and 2 while the projections of the fit onto the data are shown in Fig. 5. The results are consistent with previous measurements of these parameters [2, 6], and the SM predictions for \( \phi_s \) and \( \Delta\Gamma_s \) [32, 34]. They show no evidence of \( CP \) violation in the interference between \( B_s^0 \) meson mixing and decay, nor for direct \( CP \) violation in \( B_s^0 \to \psi(2S)\phi \) decays as the parameter \( |\lambda| \) is consistent with unity. The likelihood profile for \( \delta_0 \) is not parabolic and the 95% confidence level range is \( [2.4, 3.9] \) rad.

Figure 6 shows values of \( F_L \equiv |A_0|^2 \), the fraction of longitudinal polarisation, for \( B_s^0 \to \phi\mu^+\mu^- \) [35], \( B_s^0 \to J/\psi \phi \) [2] and \( B_s^0 \to \psi(2S)\phi \) final states as a function of the invariant mass squared of the dimuon system, \( q^2 \). The precise measurement of \( F_L \) from \( B_s^0 \to J/\psi \phi \) at \( q^2 = 9.6 \) GeV\(^2\)/c\(^4\) is now joined by the precise measurement from this paper at \( q^2 = 13.6 \) GeV\(^2\)/c\(^4\), demonstrating a clear decrease with \( q^2 \) towards the value of 1/3, as predicted by Ref. [36].

### 7 Systematic uncertainties

Systematic uncertainties for each of the measured parameters are reported in Table 3. They are evaluated by observing the change in physics parameters after repeating the likelihood fit with a modified model assumption, or by generating pseudoexperiments in case of uncertainties originating from the limited size of a calibration sample. In general the sum in quadrature of the different sources of systematic uncertainty is less than 20% of the statistical uncertainty, except for \( \Gamma_s \) where it is close to 60%.
Table 1: Results of the maximum likelihood fit to the selected $B_s^0 \rightarrow \psi(2S)\phi$ candidates including all acceptance and resolution effects. The first uncertainty is statistical and the second is systematic, which will be discussed in Section 7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>$0.668 \pm 0.011 \pm 0.006$</td>
</tr>
<tr>
<td>$\Delta\Gamma_s$ [ps$^{-1}$]</td>
<td>$0.066^{+0.041}_{-0.044} \pm 0.007$</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp</td>
</tr>
<tr>
<td>$</td>
<td>A_0</td>
</tr>
<tr>
<td>$\delta_\parallel$ [rad]</td>
<td>$3.67^{+0.13}_{-0.18} \pm 0.03$</td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>$3.29^{+0.43}_{-0.39} \pm 0.04$</td>
</tr>
<tr>
<td>$\phi_s$ [rad]</td>
<td>$0.23^{+0.29}_{-0.28} \pm 0.02$</td>
</tr>
<tr>
<td>$</td>
<td>\lambda</td>
</tr>
<tr>
<td>$F_S$</td>
<td>$0.061^{+0.026}_{-0.025} \pm 0.007$</td>
</tr>
<tr>
<td>$\delta_S$ [rad]</td>
<td>$0.03 \pm 0.14 \pm 0.02$</td>
</tr>
</tbody>
</table>

Table 2: Correlation matrix of statistical uncertainties.

| | $\Gamma_s$ | $\Delta\Gamma_s$ | $|A_\perp|^2$ | $|A_0|^2$ | $\delta_\parallel$ | $\delta_\perp$ | $F_S$ | $\delta_S$ | $\phi_s$ | $|\lambda|$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $\Gamma_s$ | 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 |
| $|A_\perp|^2$ | 1.00 | -0.54 | -0.31 | -0.05 | 0.08 | 0.03 | -0.02 | -0.02 |
| $|A_0|^2$ | 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 |
| $\delta_S$ | 1.00 | 0.07 | -0.09 |
| $\phi_s$ | 1.00 | 0.04 |
| $|\lambda|$ | 1.00 |

Repeating the fit to $m(\psi(2S)K^+K^-)$ in bins of the decay time and helicity angles shows that the mass resolution depends upon $\cos \theta_\mu$. This breaks the assumption that $m(\psi(2S)K^+K^-)$ is uncorrelated with the observables of interest, which is implicitly made by the use of weights from the sPlot technique. The effect of this correlation is quantified by repeating the four-dimensional likelihood fit for different sets of signal weights computed from fits to $m(\psi(2S)K^+K^-)$ in bins of $\cos \theta_\mu$. The largest variation in each physics parameter is assigned a systematic uncertainty. The mass model is tested by computing a new set of sWeights, using a Student’s $t$-function to describe the signal component of the $m(\psi(2S)K^+K^-)$ distribution.

The statistical uncertainty on the angular efficiency is propagated by repeating the fit using new sets of the ten integrals, $I_k$, systematically varied according to their covariance.
Figure 5: Decay-time and helicity-angle distributions for $B^0_s \to \psi(2S)\phi$ decays (data points) with the one-dimensional projections of the fitted PDF. 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.

The effect of assuming perfect angular resolution in the likelihood fit is studied using pseudoexperiments. There is a small effect on the polarisation amplitudes and strong phases while all other parameters are unaffected.

The decay-time resolution is studied by generating pseudoexperiments using the nominal double Gaussian model and subsequently fitting them using a single Gaussian model, the parameters of which have been calibrated on the prompt $J/\psi K^+ K^-$ sample. In addition, the nominal model parameters are varied within their statistical uncertainties and the fit repeated.

The decay-time efficiency introduces a systematic uncertainty from three different sources. First, the contribution due to the statistical error on the determination of the decay-time efficiency from the control channel is determined by repeating the fit multiple times after randomly varying the parameters of the time efficiency within their statistical uncertainties. The statistical uncertainty is dominated by the size of the $B^0 \to \psi(2S)K^*(892)^0$ control sample. Second, a Student’s $t$-function is used as an alternative mass model for the $m(\psi(2S)K^+\pi^-)$ distribution and a new decay-time efficiency function is produced. Finally, the efficiency function is recomputed with the lifetime of the $B^0$ modified by $\pm 1\sigma$. In all cases the difference in fit results arising from the use of the new efficiency function is taken as a systematic uncertainty. The sensitivity to the BDT selection is studied by adjusting the working point around the optimal position.
\[ |A_0|^2 \] as a function of the invariant mass squared of the dimuon system, \( q^2 \). Data points are taken from Ref. [35] (\( B_s^0 \rightarrow \phi \mu^+ \mu^- \), circles), Ref. [2] (\( B_s^0 \rightarrow J/\psi \phi \), diamond) and this paper (square).

Table 3: Summary of statistical and systematic uncertainties. Fields containing a dash (–) correspond to systematic uncertainties that are negligible.

| Source                              | \( \Gamma_s \) [ps\(^{-1}\)] | \( \Delta \Gamma_s \) [ps\(^{-1}\)] | \( |A_\perp|^2 \) | \( |A_0|^2 \) | \( \delta_{||} \) [rad] | \( \delta_{\perp} \) [rad] | \( \phi_s \) [rad] | \( |\lambda| \) | \( F_S \) | \( \delta_S \) [rad] |
|-------------------------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Stat. uncertainty                   | 0.011            | +0.041           | +0.024           | 0.014            | +0.13            | +0.33            | +0.29            | +0.069           | +0.026           | 0.14            |
| Mass factorisation                 | 0.003            | 0.002            | 0.001            | 0.002            | 0.02             | 0.01             | 0.001            | 0.003            | 0.003            | 0.003            |
| Mass model                         | 0.001            | 0.001            | –                | –                | –                | –                | –                | –                | –                | –                |
| Angular eff. (stat.)               | –                | 0.001            | 0.001            | 0.002            | 0.02             | 0.03             | 0.01             | 0.006            | 0.005            | 0.01            |
| Angular resolution                 | –                | –                | 0.001            | –                | 0.01             | –                | –                | –                | –                | –                |
| Time resolution                    | –                | 0.001            | –                | –                | –                | 0.02             | 0.02             | 0.002            | 0.002            | 0.002            |
| Time resolution (stat.)            | –                | –                | –                | –                | –                | 0.02             | 0.02             | –                | –                | –                |
| Time eff. (stat.)                  | 0.005            | 0.003            | 0.001            | 0.001            | –                | –                | –                | 0.002            | –                | –                |
| Time eff. (mass model)             | 0.001            | 0.001            | –                | –                | –                | –                | –                | –                | –                | –                |
| Time eff. (\( \tau_{B^0} \))      | 0.002            | –                | –                | –                | –                | –                | –                | –                | –                | –                |
| \( B_s^+ \) feed-down              | 0.001            | –                | –                | –                | –                | –                | –                | –                | –                | –                |
| Fit bias                           | 0.001            | 0.006            | 0.001            | 0.01             | –                | –                | –                | –                | –                | –                |
| 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            | +0.045           | +0.023           | 0.014            | –0.18            | –0.39            | –0.29            | –0.050           | –0.026           | 0.14            |

equally for both signal and control channel, and also differently for each channel in order to make the ratio \( \varepsilon_{B_s^0}^{\psi(2S)}(t)/\varepsilon_{B_s^0}^{J/\psi}(t) \) uniform. The efficiency is recomputed in each case and the fit repeated. No significant change in the physics parameters is observed.

A small fraction of \( B_s^0 \rightarrow \psi(2S)\phi \) signal candidates comes from the decay of \( B_c^+ \) mesons, causing an average positive shift in the reconstructed decay time of the \( B_s^0 \) meson. This fraction was estimated as 0.8\% in Ref. [2] and pseudoexperiments were used to assess the impact of ignoring such a contribution. Only \( \Gamma_s \) was affected, with a bias on its central value of \((+20 \pm 6)\%\) of its statistical uncertainty. The assumption is made
that the ratio of efficiencies for selecting $B_s^0 \rightarrow \psi(2S)\phi$ decays either promptly or via the decay of $B_c^+$ mesons is the same as that for $B_s^0 \rightarrow J/\psi \phi$ decays. This leads to a bias of $+0.002 \pm 0.001 \text{ps}^{-1}$ in $\Gamma_s$. The central value of $\Gamma_s$ is therefore reduced by 0.002 ps$^{-1}$ and a systematic uncertainty of 0.001 ps$^{-1}$ is assigned.

A test for a possible bias in the fit procedure is performed by generating and fitting many simulated pseudoexperiments of equivalent size to the data sample. The resulting biases are small and those that are not compatible with zero within two standard deviations are quoted as systematic uncertainties.

The uncertainty from knowledge of the LHCb detector’s length and momentum scale is negligible as is the statistical uncertainty from the sWeights. The tagging parameters are allowed to float in the fit using Gaussian constraints according to their uncertainties, and thus their systematic uncertainties are propagated into the statistical uncertainties reported on the physics parameters themselves. The systematic uncertainties for $\phi_s$, $\Delta\Gamma_s$ and $\Gamma_s$ can be treated as uncorrelated between this result and those in Ref. [2].

8 Conclusions

Using a dataset corresponding to an integrated luminosity of 3.0 fb$^{-1}$ collected by the LHCb experiment in $pp$ collisions during LHC Run 1, a flavour tagged, decay-time dependent angular analysis of approximately 4700 $B_s^0 \rightarrow \psi(2S)\phi$ decays is performed. The analysis gives access to a number of physics parameters including the $CP$-violating phase, average decay-width and decay-width difference of the $B_s^0$ system as well as the polarisation amplitudes and strong phases of the decay. The effective decay-time resolution and effective tagging power are approximately 47 fs and 3.9%, respectively. This is the first measurement of the $CP$ content of the $B_s^0 \rightarrow \psi(2S)\phi$ decay and first time that $\phi_s$ and $\Delta\Gamma_s$ have been measured in a final state containing the $\psi(2S)$ resonance. The results are consistent with previous measurements [2, 6], the SM predictions [32–34], and show no evidence of $CP$ violation in the interference between $B_s^0$ meson mixing and decay. The parameter $|\lambda|$ is consistent with unity, implying no evidence for direct $CP$ violation in $B_s^0 \rightarrow \psi(2S)\phi$ decays. The fraction of longitudinal polarisation in the $B_s^0 \rightarrow \psi(2S)\phi$ decay is measured to be lower than that in the $B_s^0 \rightarrow J/\psi \phi$ decay, consistent with the predictions of Ref. [36].

Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FANO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (USA). We are indebted to
the communities behind the multiple open source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).

References


Università di Genova, Genova, Italy
Università di Milano Bicocca, Milano, Italy
Università di Roma Tor Vergata, Roma, Italy
Università di Roma La Sapienza, Roma, Italy
AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
Hanoi University of Science, Hanoi, Viet Nam
Università di Padova, Padova, Italy
Università di Pisa, Pisa, Italy
Università degli Studi di Milano, Milano, Italy
Università di Urbino, Urbino, Italy
Università della Basilicata, Potenza, Italy
Scuola Normale Superiore, Pisa, Italy
Università di Modena e Reggio Emilia, Modena, Italy
Iligan Institute of Technology (IIT), Iligan, Philippines